الإحصاء التطبيقي

منشورات جامعة دمشق
كلية العلوم السياسية

الإحصاء التطبيقي

الدكتور
عدنان غانم
أستاذ مساعد في قسم الإحصاء التطبيقي
كلية الاقتصاد

الدكتور
نبيل علي
أستاذ مساعد في قسم الاقتصاد والإدارة العامة
كلية العلوم السياسية

١٤٢٧ - ١٤٢٨ هـ

جامعة دمشق

٢٠٠٦ - ٢٠٠٧ م

الفهرس

كما أنه يحتل مكانة في **الاقتصاد** وبخاصة في تطوير الاقتصاد ونموه لأن توفر البيانات الإحصائية عن اقتصاد دولة ما يشكل الأساس في وضع السياسات الاقتصادية الملائمة للنظريات الاقتصادية.

يُعد الإحصاء في **العلوم الاجتماعية** وسيلة مهمة لقياس مدى رفاهية الناس وتقدم المجتمع ورقي مستوى الأفراد ثقافياً واجتماعياً، حيث يمكن جمع البيانات الإحصائية المتعلقة بالحالة الثقافية " مثقفين، أميين" عددهم ونسبتهم في المجتمع وتوزيعهم حسب النوع أو المناطق.

أما أهمية الإحصاء في **علم السياسة** فتنبع أهم من خلال استخدام أدواته في العديد من الموضوعات التي يمكن أن تخضع للقياس والتعبير عنها كمياً كالسلوك التصويتي والانتماء الحزبي، وأثر الحملة الانتخابية في السلوك الانتخابي، وقياس ميول الرأي العام إزاء برنامج إصلاحي معين والعلاقة بين مستوى الدخل ومستوى المشاركة السياسية.

في الآونة الأخيرة بدأ المهتمين بالسياسة بتطوير النظريات والقوانين المتعلقة بسلوكيات هذا المجال التخصصي سواءً من حيث الكم أم الكيف والاستفادة من تطور مناهج البحث العلمي في التخصصات الأخرى ذات الصلة بها لعرض بعض نتائجهم في صورة مؤشرات تلخيصية وصفية أو استدلالية وتفسيرها واتخاذ القرارات سليمة بناءً عليها .

نأمل في فصول هذا الكتاب، أن نكون قد قدمنا مفهوماً واضحاً لعلم الإحصاء الرياضي والاستدلال الإحصائي ومدى فائدتهما لتخصص السياسة الذي يعاني من نقص واضح في الإنتاج الفكري في مجال المناهج الإحصائية وتطبيقاتها في السياسة، كما يفقد هذا التخصص الإطار النظري الرياضي المتطور الذي يساعد على تفسير نشاطاته .

مقدمة

لكلمة **الإحصاء** معاني كثيرة ومتباينة فعند سماع هذه الكلمة يتبادر إلى أذهان البعض من الناس مفهوم الجداول والأرقام الخاصة بعدد مفردات معينة كعدد السكان أو عدد المرشحين أو الناخبين وعدد المراكز الانتخابية، عدد صناديق الاقتراع، والبعض الآخر من الناس يُبني مفهوم الإحصاء لديه بناءً على ما تنشره الصحف من بيانات مختلفة أو بيانات استطلاعات الرأي أو تلك البيانات التي توضح المواد المصدرة والمستوردة في بلد ما .. إلى غير ذلك من الأمثلة المتداولة يومياً. بدأ استخدام كلمة إحصاء لأول مرة في مجالات متعلقة بإعداد جداول ذات طابع تنظيمي في التنظيمات الحكومية مثل جداول جمع الضرائب وغيرها.

أما في الوقت الحاضر فقد تعددت أساليب استخدام الإحصاء في معظم المجالات الحياتية وأصبحت طرائق البحث العلمي تتطلب حضور الإحصاء بشكل كثير لفائدته في تصميم التجارب العلمية وإمكانية الحصول على النتائج التي يسعى الباحثون إلى تحقيقها عن طريق هذا العلم، بل أصبح الإلمام بأساليب استخدام الإحصاء البسيطة وكيفية تطبيقها ضرورة ملحة من مقتضيات الحضارة المعاصرة .

يشكل الإحصاء في **التخطيط** أساساً يعتمد عليه، في رسم أية خطة مستقبلية تحددها السياسة العليا للدولة لتحقيق أهدافاً معينة في مختلف المجالات والأنشطة، ومن المعروف أنه لا يمكن أن توضع أية خطة دون توفر البيانات الإحصائية، وكلما كانت البيانات دقيقة ومفصلة كلما كانت الخطة أسلم وأقرب إلى الواقع .

لقد توخينا البساطة والسهولة في العرض الرياضي، وابتعدنا عن البراهين المعقدة، ليكون هذا الكتاب في متناول جميع المهتمين، لاسيما طلبة كلية العلوم السياسية .

يتكون هذا الكتاب من ثمانية فصول توزعت على النحو التالي :

1 ــ مبادئ الاحتمالات .

2 ــ المتغيرات العشوائية والتوزيعات الاحتمالية

3 ــ توزيعات المعاينة

4 ــ التقدير الإحصائي للعينات الكبيرة الحجم

5 ــ اختبار الفروض الإحصائية للعينات الكبيرة الحجم

6 ــ التقدير الإحصائي واختيار الفروض للعينات الصغيرة

7 ــ الإحصاء اللامعلمي

8 ــ تحليل التباين

وقد ضمنّا كل فصل من الفصول السابقة عدداً من التمارين المحلولة وأخرى غير محلولة لإتاحة الفرصة للطالب كي يختبر مدى تفهمه وإدراكه للمعلومات المعطاة وألحقت في نهاية الكتاب الجداول الإحصائية اللازمة، بالإضافة إلى قائمة بالمراجع العربية والأجنبية وأخرى بالمصطلحات الإحصائية .

إن إعداد هذا الكتاب، هو إعداد جماعي ومسؤولية مشتركة، لكل من **الدكتور عدنان غانم والدكتور نبيل علي .**

أننا نضع هذا الكتاب بين أيديكم، نأمل منكم جميعاً طلاباً وباحثين، أن يتفضلوا بتقديم أي نقد بنّاء يساهم في إغنائه وإخراجه مستقبلاً بشكل نجعل منه

أكثر نفعاً وفائدة، وبما يسمح لنا في متابعة مسيرة البحث العلمي، وأخيراً نقدم جزيل الشكر والتقدير لكل من ساهم في إخراج هذا الكتاب.

والله ولي التوفيق .

دمشق في 2006/3/14 . المؤلفان

الفصل الأول

مبادئ الاحتمالات

- مقدمة
- التجربة الاحتمالية والحوادث
- تعاريف هامة
- خواص الاحتمالات
- الاحتمال الشرطي
- نظرية بايز
- تمارين غير محلولة

1-1- مقدمة :

إن كلمة " احتمال " تستعمل كثيراً في حياتنا اليومية، كأن نقول مثلاً: " احتمال فوز مرشح ما في الانتخابات " أو " احتمال وصول الطائرة قبل ساعة من الموعد المحدد " أو "احتمال ارتفاع الحرارة عن معدلها " أو " احتمال أن يكون منتج ما جيد أو رديء الصنع " .. الخ. وكلمة احتمال تستعمل للتعبير عن أمر معين، لا يمكن التنبؤ بنتيجة حدوثه سلفاً، لذلك فإن العديد من الناس يستخدمون كلمة " احتمال" في تعبيرات تتفاوت درجة الاعتقاد في حدوث حادثة ما من شخص لآخر لكون هذه التعبيرات مبنية على أحكام ذاتية لكل شخص بعينه وعلى مقدار المعلومات المتعلقة بالحدث المتوفرة لديه .

بدأ الاحتمال تاريخياً في القرن السابع عشر، بما يعرف بالألعاب القائمة على الدور الذي تلعبه الصدفة (الفرصة) كالنرد وقذف قطعة نقود متوازنة أو سحب ورقة من مجموعة أوراق اللعب ... الخ. في مثل هذه الألعاب تكون نتيجة محاولة معينة غير مؤكدة الحدوث ولا يمكن التنبؤ بنتيجتها مسبقاً .

ونتيجةً للتباين في استخدام كلمة " احتمال" من شخص ولآخر تطلب وضع مقاييس رقمية للاحتمال للتغلب على مشكلة التنبؤات المختلفة حول حدوث أو عدم حدوث حادثة ما، من خلال دراسة هذه المقاييس الاحتمالية والعلاقات التي تربط فيما بينها نشأ ما يعرف بعلم الاحتمال.

يُعد علم الاحتمال من الوسائل الهامة جداً في علم الإحصاء. فالإحصاء الرياضي بني أساساً على نظرية الاحتمال، حيث يهتم بالتنبؤ بأحداث مستقبلية غير مؤكدة الحدوث، أي يتضمن عنصر المخاطرة والاحتمال. وتبرز أهمية علم الاحتمال وحيويته عند اتخاذ قرارات حول قضايا في كثير من المجالات العلمية والاقتصادية والإدارية. مثل هذه القضايا تتطلب تعيين احتمالات

موضوعية لأحداثها المختلفة ودراستها يساعد على اتخاذ قرارات ذكية ومناسبة عندما يوجد هناك مخاطرة أو عدم تأكد.

لقد برزت أهمية نظرية الاحتمالات بشكل كبير بعد ظهور الثورة الصناعية والتقدم التقاني في العمليات الإنتاجية، حيث تلعب نظرية الاحتمال دوراً مهماً في دراسة الحوادث والمتغيرات والظواهر التي لا يمكن التنبؤ بحدوثها سلفاً، إضافة إلى ذلك يوجد العديد من العلوم التطبيقية والنظريات المبنية أساساً على نظرية الاحتمال مثل نظرية الإحصاء (Theory of Statistics) ونظرية القرارات (Theory of Decisions) ، ونظرية الطوابير (Queuning Theory) ونظرية المعلومات (Information Theory) وكثير من العلوم التقنية والهندسية والعسكرية والطبية... إلخ. ومن المفيد هنا الإشارة إلى أن هناك عدة تفسيرات للاحتمال فقد فسر الاحتمال بأنه التكرار النسبي أو الاحتمال التجريبي (Relative Frequency) ، حيث تستخدم التكرارات النسبية التي حدثت في الماضي كاحتمالات. وتكمن الصعوبة في هذا التفسير أننا نحصل على احتمالات مختلفة (تكرارات نسبية) عندما يختلف عدد المحاولات أو التجارب. وهناك أيضاً تفسير آخر للاحتمال يطلق عليه التفسير الكلاسيكي.

والفرق بين الاحتمال الكلاسيكي والاحتمال التجريبي هو أن التكرار النسبي عبارة عن نسبة بين عدد المرات التي يقع فيها حدث ما فعلاً وبين العدد الكلي للنواتج الفعلية أو المشاهدات، وكلما زاد عدد المحاولات أو التجارب (مثلاً عدد مرات رمي العملة) كلما اقترب التكرار النسبي من الاحتمال الكلاسيكي، إضافة إلى ذلك هناك تفسير ثالث للاحتمال يسمى الاحتمال الشخصي (أو الموضوعي) وهذا يشير إلى درجة الاعتقاد لدى شخص ما بأن حدث ما سوف يحدث، وهذا الاحتمال يمثل الحكم الشخصي أو

الذاتي على حدوث شيء ما أو عدم حدوثه. أي يعتمد على اعتقاد الشخص وعلى ما يتوفر له من معلومات. ومن عيوب هذا التفسير أن كل شخص يعطي احتمالاً مخالفاً لغيره عند مواجهتهم لنفس الموقف.

تقوم نظرية الاحتمال عموماً على مفهومين هما:

- التجربة الاحتمالية أو العشوائية (Random Experiment) : بالتعريف هي كل عملية تؤدي إلى ملاحظة أو قياس ظاهرة ما.
- الحادثة (Event): هي الناتج من التجربة العشوائية والتي لا يمكن معرفة نتائجها قبل إجراؤها.

1-2- التجربة الاحتمالية والحادثة :

كما هو معروف، إن احتمال وقوع A هو P حيث أن A حادثة ما مثل احتمال أن يفوز المرشح بالانتخابات .... الخ وحيث أن P هو عدد أو صفة تصف إمكانية الحدوث أو عدم إمكانية الحدوث للحادثة A ويعتمد على سلسلة من الظروف المحيطة، هذه السلسلة تسمى تجربة أو محاولة. ونتيجة التجربة تسمى الحادثة وتتميز التجربة العشوائية بأننا :

- لا نستطيع التنبؤ بصورة أكيدة عن نتيجة التجربة مسبقاً.
- نستطيع فقط عمل قائمة بكافة النتائج الممكنة .

تعريف :

مجموعة كل النتائج الممكنة لتجربة عشوائية ما تسمى فراغ العينة ويرمز لها بالرمز S.

الحادث هو مجموعة جزئية من فراغ العينة. ولتوضيح هذا التعريف لدينا الأمثلة التالية :

مثال (1) :

تجربة قذف قطعة نقد متوازنة مرة واحدة. هذه التجربة تؤدي إلى إحدى نتيجتين إما صورة النقد (Head) ويرمز لها عادة بالرمز H أو إلى الكتابة على النقد (Tail) ويرمز لها عادة بالرمز T. عندئذ فإن فراغ العينة لهذه التجربة هو . $S= \{H, T\}$ وقد نكون مهتمين بأن تكون نتيجة التجربة إحدى الحالات التالية :

(a) هي H

(b) إما H أو T

(c) كلاهما T, H (وهذا غير محتمل حدوثه) .

(d) ليست H.

مثال (2) :

طلب مدرس من 6 تلاميذ في مدرسة معينة أن يختار كل واحد منهم بصورة عشوائية رقماً صحيحاً من الأرقام الموجبة الـــ 52 الأولى من مجموعة الأعداد الصحيحة الموجبة: في هذه الحالة نهتم فيما إذا كان 2 على الأقل من هذه الأرقام الستة المختارة متساوية (M) أو جميعها مختلفة (D). عندئذ نجد أن فراغ العينة هو :

$S = \{ M , D \}$

مثال (3) :

علبة بسكويت تحتوي على واحدة من أربع جوائز مختلفة لتشجيع الزبائن على الشراء. شراء علبة بسكويت واحدة يؤدي إلى واحدة من هذه الجوائز كنتيجة

للاختيار العشوائي. أي أن فراغ العينة لهذه التجربة هو { كل الجوائز الأربع المختلفة } = S .

مثال (4) :

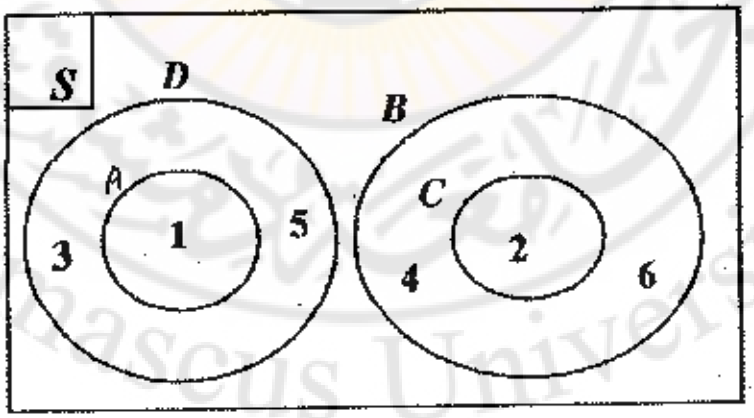
في تجربة قذف حجر النرد. نعلم أن هناك 6 نتائج ممكنة. فراغ العينة لهذه التجربة هو : $S = \{1, 2, 3, 4, 5, 6\}$.

يمكن أن نعرف الحوادث الآتية على فراغ هذه التجربة :

A هي حادثة الحصول على العدد 1 و B هي حادثة الحصول على عدد زوجي و C هي حادثة الحصول على عدد زوجي لا تتجاوز 3 وكذلك D هي حادثة عدد فردي، ويمكن تمثيل تجربة قذف حجر النرد في شكل فن المبين في الشكل المرفق وعلى ذلك يكون لدينا:

(1) $A = \{ 1 \}$ (2) $B = \{2, 4, 6 \}$

(3) $C = \{ 2 \}$ (4) $D = \{ 1, 3, 5\}$



ملاحظة :

يرمز للحوادث (المجموعة الجزئية) بالحروف الكبيرة: A,B,C, . . .

ويرمز لنقاط العينة أو العناصر بالحروف الصغيرة a,b,c. . . .

تعريف :

كل نقطة من فضاء العينة تسمى حادثة بسيطة أو حادثة ابتدائية، ونعني بالحادثة البسيطة تلك الحادثة التي تحتوي على عنصر واحد من عناصر فضاء العينة، أما الحادثة المركبة فهي التي تحتوي على عنصرين أو أكثر من عناصر فضاء العينة .

ولمعرفة ماذا يقصد باحتمال الحادثة A والذي يرمز له بالرمز P(A)، لدينا المثال التالي:

بفرض رغبت إحدى المدارس اختيار 3 أرقام من بين 1000 رقم متسلسل للطلاب، يتألف كل منها من ثلاث خانات، فإن اختيار الأرقام يتم بصورة عشوائية، وعندئذ نتوقع أن كل رقم من الـ 1000 له نفس الفرصة في الاختيار أي $\frac{1}{1000}$ ، وحصلنا مثلاً أثناء الاختيار العشوائي على الأرقام $A = \{211, 121, 112\}$ فهذا يعني أن احتمال الحادث A هو $P(A) = \frac{3}{1000}$ وهكذا .

وهناك أيضاً حوادث أخرى، مثل الحادث الأكيد وهو الذي يظهر عند إجراء التجربة مثل: (ظهور رقم 5 عند رمي حجر النرد) والحادث المستحيل فهو الذي لا يظهر أبداً عند إجراء التجربة ويرمز له (ϕ)، بينما الحادث الممكن هو الحادث الذي لا يمكن أن نجزم بحدوثه مثل: (ظهور رقم 5 لا يعني أننا عرفنا مسبقاً بحدوثه).

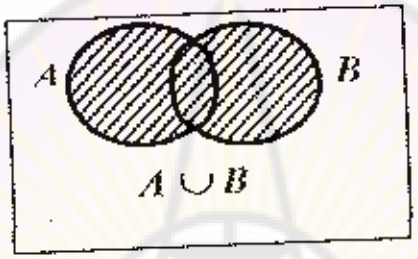
1-3- تعاريف هامة :

– الاجتماع union ورمزه U :

إن اجتماع مجموعتين B, A هو مجموعة العناصر التي تنتمي إلى A أو إلى B أو إليهما معاً. ويرمز لذلك بالرمز $A \cup B$، حيث إن $A \cup B$ هو مجموعة العناصر التي تنتمي إلى واحدة منها على الأقل .

ففي مثالنا السابق (4) نجد أن :

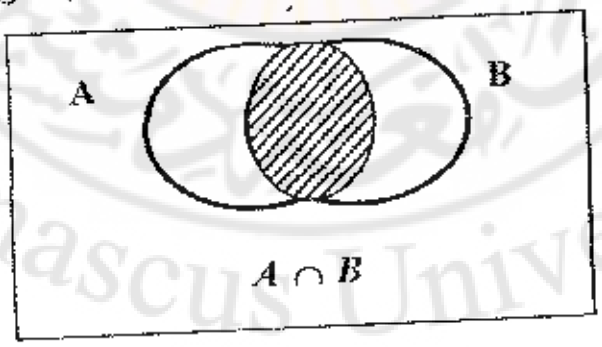
$A \cup B = \{1,2,4,6\}$ و $B \cup C = \{2,4,6\}$



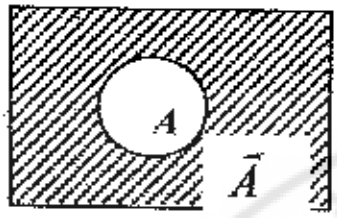
– التقاطع Intersection ورمزه ∩ :

إن تقاطع مجموعتين A و B هو مجموعة العناصر التي تنتمي إلى A وإلى B بالوقت ذاته، ويرمز له بالرمز $A \cap B$، ولو عدنا للمثال (4) السابق لوجدنا أن :

$A \cap B = \phi$ و $B \cap C = \{2\}$ (أحياناً نكتب $A \cap B$ بالصورة AB).



ـ متممة مجموعة (أو تتمة مجموعة) :



إن متممة مجموعة A ويرمز لها بالرمز A^c (أو $\bar{A}$) .

هي مجموعة العناصر الموجودة في المجموعة الشاملة ولا تنتمي إلى A أي أن $\bar{A}$ هي نفي A (أو ليس A).

من الواضح أنه إذا كانت S هي المجموعة الشاملة فإن $A \cup \bar{A} = S$. ففي مثالنا (4) السابق نلاحظ أن $\bar{A} = \{2,3,4,5,6\}$.

ـ المجموعة الخالية :

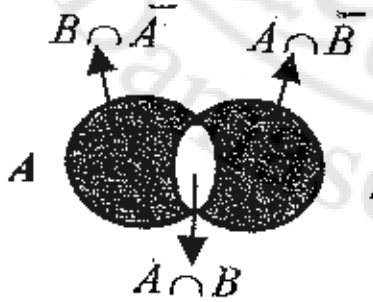
هي تلك المجموعة التي لا تتضمن أي عناصر فيها ويرمز لها بالرمز ϕ أو بالرمز { } ونجد في مثال (4) أن $D \cap C = \phi$.

ـ المجموعات الجزئية :

إذا كان كل عنصر ينتمي إلى A ينتمي أيضاً إلى B فيقال أن A محتواه في B وتكتب بالصورة $A \subset B$ (وعند التساوي نكتب) $A \subseteq B$، $A \subset B$ وعليه كالتالي : A محتواة في B أو A مجموعة جزئية في B.

ـ الفرق بين مجموعتين :

إن الفرق بين المجموعتين B و A هو مجموعة العناصر التي تنتمي إلى A ولا تنتمي إلى B. ويرمز لذلك بالرمز $A \cap \bar{B}$ وهذا موضح في شكل فن المقابل أي :



$A \cup B = A\bar{B} \cup AB \cup B\bar{A}$ حيث

إن :

$\bar{A}B = \bar{A} \cap B$

وبالمثل $A\bar{B} = A \cap B$ وكذلك $AB = A \cap B$.

العمليات على المجموعات :

إذا كانت لدينا المجموعات الجزئية A, B, C من S فإن عمليات المجموعات تحقق عدة خواص منها الآتي :

ـ قوانين التبديل :

$$A \cup B = B \cup A$$

$$A \cap B = B \cap A$$

ـ قوانين الدمج أو التجميع :

$$(A \cup B) \cup C = A \cup (B \cup C)$$

$$(A \cap B) \cap C = A \cap (B \cap C)$$

ـ قوانين التوزيع :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

ـ قوانين دي مورغان :

$$(\bar{A} \cup \bar{B}) = 1 - [(A \cap B)]$$

$$(\bar{A} \cap \bar{B}) = 1 - [(A \cup B)]$$

1-4- خواص الاحتمال:

ليكن (P) القياس الاحتمالي المعرف على (S, ξ) والمعطى :

$$P : \xi \to [0,1]$$

وهو يحقق الخواص التالية :

(1) $0 \le P(A) \le 1$ (الاحتمال موجب ودائماً أكبر من الصفر)

(2) $P(\phi) = 0$ ، $P(S)=1$

(3) إذا كانت $A_1, A_2, \ldots,$ تجمعاً من الحوادث المنفصلة في ξ بحيث إن :

$A_i \cap A_j = \phi \quad \forall\, i \neq j$ فإن : $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

وما يمكن ملاحظته من هذا التعريف هو أن الاحتمال P عدد ليس له وحدات ولا يمكن أن يكون سالب القيمة ولا يمكن أيضاً أن يتجاوز الواحد الصحيح .

ملاحظة :

إذا وضعنا في الخاصية (3) الحوادث المنفصلة A_1, A_2، وباقي الحوادث $A_i = \phi$ لجميع $i = 3, 4, 5, \ldots$ فإن الخاصية (3) تعني :

$$P(A_1 \cup A_2) = (P(A_1) + P(A_2) + P(\phi) + \ldots + P(\phi)$$
$$= P(A_1) + P(A_2)$$

أي أنه :

إذا كان A_1, A_2 منفصلة في ξ فإن $P(A_1 \cup A_2) = P(A_1) + P(A_2)$ ،

ونلاحظ أن A_1, A_2 حوادث منفصلة تعني $A_1 \cap A_2 = \phi$،

أو أنهما حادثتان متنافيتان أي أن حدوث إحداهما ينفي حدوث الأخرى .

مثال (5) :

عند قذف قطعة من النقد (ممكن أن تكون غير متوازنة أي متحيزة)، فإن فراغ العينة لهذه التجربة هو :

$S = \{ H, T \}$

والصف من الحوادث هو :

$\xi = \{ \phi , H , T , S \}$

والقياس الاحتمالي الممكن هو $P : \xi \to [0,1]$. عندئذ يكون لدينا الآتي :

$P(S) = 1 \quad , \quad P(\phi) = 0$

$P(H) = P \quad , \quad P(T) = 1 - P$

حيث إن P عدد حقيقي $(0 \le P \le 1)$ موجود في المجال [0,1]، وكانت $P = \frac{1}{2}$ ، فإن قطعة النقد هي قطعة متوازنة أو غير متحيزة عدا ذلك فهي متحيزة أو (غير متوازنة).

نظرية I :

إن احتمال الحادث المتمم للحادث A يعطى :

$$P(\overline{A}) = 1 - P(A)$$

ولإثبات ذلك، نعلم بأن:

$$\phi = \overline{A} A \cap \quad و \quad \overline{A} S = A \cup$$

وبالتالي يكون:

$$P(S) = P(A \cup \overline{A})$$

$$1 = P(A) + P(\overline{A})$$

$$= P(A) + P(\overline{A})$$

$$\Rightarrow P(A) = 1 - P(\overline{A})$$

مثال (6) :

سحبت ورقة من مجموعة أوراق لعب مكونة من 52 ورقة، فما هو احتمال أن تكون الورقة المسحوبة ولداً (J) (يترك للطالب).

مثال (7) :

بفرض أننا حصلنا في 100 رمية لعملة نقود متوازنة على H(53) و T(47). فيكون التكرار النسبي لـ H هو $\left(\frac{53}{100}\right)$ وهو ما يطلق والاحتمال التجريبي والذي يجب تمييزه عن الاحتمال الكلاسيكي (الفرص المتساوية) :

أي $P(H)=\frac{1}{2}$

ومع تزايد عدد الرميات إلى ما لانهاية فإن التكرار النسبي أو الاحتمال التجريبي يقترب من الاحتمال الكلاسيكي أي يقترب من $\frac{1}{2}$ ، فعلى سبيل المثال قد يكون التكرار النسبي أو الاحتمال التجريبي 0.517 في 1000 رمية و 0.508 في 10000 رمية وهكذا ...

نظرية II : إن احتمال الحادث المستحيل يساوي الصفر، أي: $P(\phi)=0$:

البرهان :

لتكون $A=\phi$ و $A^c=S$ وبالتالي فإن :

$$P(A)=1-P(\overline{A})$$
$$=1-P(S)\Rightarrow$$

$$P(\phi)=1-1=0$$

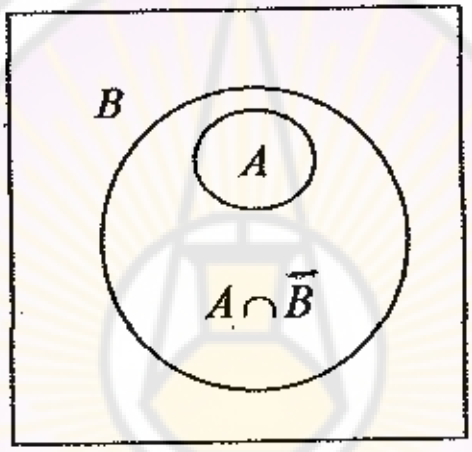
نظرية III : إذا كانت لدينا الحوادث A, B وكان $A \subset B$ فإن : $P(A) \le P(B)$:

البرهان : حيث أن :

$$B = A \cup (B \cap \overline{A}), A \cap (B \cup \overline{A}) = A$$

فنجد أن (انظر شكل فن):

$$P(B) = P(A) + P(B \cap \overline{A}) \ge P(A)$$



نظرية V :

لتكن A حادثة ، فإن احتمالها $P(A) \le 1$

البرهان :

بما أن $A \subset S$ فإن $P(A) \le P(S)$ وبالتالي يكون :

$$B = (A \cap B) \cup (B \cap \overline{A}), \quad 0 \le P \le 1$$

نظرية IV : إن احتمال اجتماع حادثين A, B يعطى :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

البرهان :

من شكل فن المرافق نجد أن : $A, B \cap A$ منفصلتان ، وأن

منفصلتان وأن :$A \cup \overline{B} = A \cup (B \cap A)$ وكذلك $B \cap A , A \cap \overline{B}$

$B = (A \cap B) \cup (B \cap A)$ وبالتالي فإن :

$$P(A \cup B) = P(A) + P(B \cap A) \quad (1)$$

$$P(B) = P(A \cap B) + P(B \cap A) \quad (2)$$

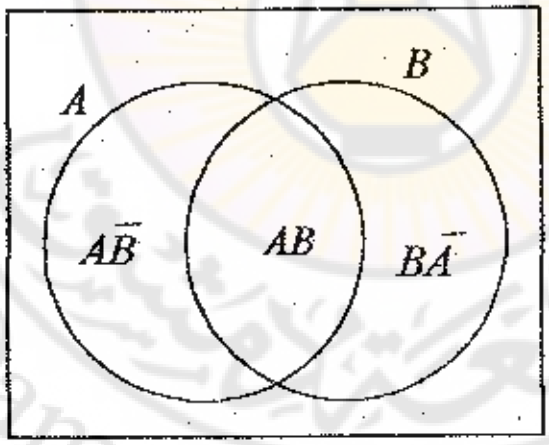
من المعادلة (2) نجد أن:

$$P(B \cap A) = P(B) - P(B \cap A) \quad (3)$$

بالتعويض من (3) في (1) نحصل على المطلوب أي :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

وهو ما يسمى بقانون الجمع، انظر الشكل التالي :



مثال (8) :

لدينا قاعدتان لإطلاق صواريخ أرض/جو، وفجأة ظهرت طائرة معادية في مجال القاعدتين حيث أطلقت كل قاعدة صاروخاً واحداً على الطائرة، إذا كان:

$P(A) = 0.93$, $P(B) = 0.89$ وكذلك $P(A \cap B) = 0.87$

المطلوب : أوجد احتمال إصابة الطائرة المعادية لصاروخ واحد على الأقل ؟

الحل :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= 0.93 + 0.89 - 0.87 = 0.95$$

مثال (9) :

صندوق به 5 كرات اثنتان بيضاوان وكرة واحدة سوداء وكرتان حمراوان، سحبت كرة من هذا الصندوق بصورة عشوائية، فما احتمال أن تكون الكرة المسحوبة بيضاء أو سوداء ؟ .

الحل :

لتكن A حادثة الحصول على كرة بيضاء و B حادثة الحصول على كرة سوداء .

بما أن هذه الحوادث متنافية الحدوث أي بيضاء أو سوداء، فإننا نطبق قانون الجمع فنحصل على :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{2}{5} + \frac{1}{5} - 0$$

$$= \frac{3}{5}$$

بفرض أن الصندوق السابق به 200 كرة منها 150 بيضاء، 30 سوداء و20 حمراء، فما هو احتمال سحب كرة بيضاء أو حمراء، عند سحب كرة واحدة عشوائياً من الصندوق ؟ .

حيث أن الحوادث السابقة متنافية تبادلياً فإن :

$$P(A \cup B) = \frac{150}{200} + \frac{30}{200}$$
$$= \frac{9}{10}$$

نظرية VI :

إذا كانت C و B و A ثلاثة حوادث فإن :

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - p(A \cap B)$$

$$P(A \cap C) - P(B \cap C) - P(A \cap B \cap C)$$

البرهان : نكتب :

$$A \cup B \cup C = A \cup (B \cup C)$$

تفاصيل البرهان متروكة كتمرين .

1-5- الاحتمالي الشرطي Condition Probability :

يقوم هذا الاحتمال على إيجاد احتمالات معينة إذا علم تحقق حوادث أخرى ويُعبر عنه كالتالي: P(A/B) ويقرأ احتمال حدوث (A) علماً بأن الحادث (B) قد وقع، ويعرّف بالعلاقة التالية:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

بحيث : $P(B) > 0$

وبالمثل :

إذا كان $P(A) > 0$ فإن $P(B/A) = \frac{P(A \cap B)}{P(A)}$ أحياناً $P(A/B)$ يقرأ " احتمال حدوث A شريطة أن B قد حدث .

مثال 10 :

لدينا 20 كرة متماثلة من حيث الشكل، وضعت في صندوقين، وكانت (13) منها بلون أحمر (R) و (7) بلون أصفر (Y) تبعاً للجدول الآتي:

اللون	الصندوق الأول (A)	الصندوق الثاني (B)	المجموع
R	5	8	13
Y	3	4	7
المجموع	8	12	20

اختير منها كرة عشوائياً :

(1) أوجد احتمال أنها حمراء .

(2) إذا كانت الكرة المختارة من اللون الأحمر فما هو احتمال أنها من الصندوق الأول.

الحل :

(1) $P(R)=(1/2)(5/8)+(1/2)(8/12)$
$=(5/16)+(8/24)=0.64$

(2)

$$P(A/R)=\frac{P(A\cap R)}{P(R)}=\frac{P(A)\,P(R/A)}{P(R)}$$

$$=\frac{(1/2)(5/8)}{0.64}=0.48$$

مثال (11) :

إذا كان لدينا : $P(A\cap B)=0.3$ ، $P(B)=0.5$ ، $P(A)=0.4$

فإن : $P(A/B)=\frac{P(A\cap B)}{P(B)}=\frac{0.3}{0.5}=0.6$

وبالمثل نجد أن : $P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{0.3}{0.4}=0.75$

مثال (12) :

إذا كان لدينا : $P(A\cap B)=0.2$ ، $P(B)=0.3$ ، $P(A)=0.7$

والمطلوب : أوجد احتمال كل من $P(B/A)$ و $P(A/B)$

$$P(A/B)=\frac{P(A\cap B)}{P(B)}=\frac{0.2}{0.3}=2/3$$

$$P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{0.2}{0.7}=2/7$$

خواص الاحتمال الشرطي :

$P(A/B)$ ، $P(B)\geq 0$ (1)

$P(A/A)=1$ (2)

(3) إذا كان $A_1, A_2, \ldots$ حوادث متنافية:

$A_i\cap A_j=\phi \quad \forall\, i\neq j$ فإن :

$$P(A_1\cup A_2\cup\ldots/B)=P(A_1/B)+P(A_2/B)+\ldots$$

مثال (13) :

بفرض عائلة لديها طفلين، ما هو احتمال أن يكون كلاهما ذكور، علماً أن واحداً منهما على الأقل هو ذكر؟

الحل : إذا رمزنا للولد بالرمز B وللبنت بالرمز G، عندها يصبح فراغ العينة لهذه التجربة هو :

$S=\{GG, GB, BG, BB\}$

باستخدام مبدأ الإمكانيات المتساوية الفرص نجد أن :

$$P(GG) = P(GB) = P(BG) = P(BB) = \frac{1}{4}$$

بتطبيق تعريف الاحتمال الشرطي نحصل على :

P (BB / GB ∪ BG ∪ BB) (ولد واحد على الأقل) = P (BB /

= P (BB ∩ (GB ∪ BG ∪ BB) / P (BB) / P (GB) + P (G) + P(BB)

$$= P(BB) / P(GB) + P(BG) + P(BB) = \frac{1/4}{3/4} = \frac{1}{3}$$

ملاحظات :

(1) من خلال تعريف الاحتمال الشرطي يمكن بسهولة إدراك أن :

P (A ∩ B) = P (A) P (B / A)

أو :

P (A ∩ B) = P (B) P (A / B)

(2) لأي حادثين B, A فإن :

P (A) = P (A / B) P (B) + P (A / B^c) P (B^c)

(3) إذا كان $B_1, B_2, \ldots, B_n$ تجزيئة لفراغ عينة S فإن :

$$P(A) = \sum_{i=1}^{n} P(A/B_i)P(B_i)$$

مثال (14) :

صندوقان I و II بهما كرات ملونة، الصندوق I يحتوي كرتين بيضاء (W) و3 كرات حمراء (R)، والصندوق II يحتوي 3 كرات بيضاء (W) و4 كرات حمراء (R)، وقد سحبت كرة واحدة عشوائياً من الصندوق I ووضعت

في الصندوق II ثم سحبت كرة واحدة عشوائياً من الصندوق II لمعرفة لونها، فما هو احتمال أنها حمراء (R) ؟

الحل : لتكن A حادثة كون الكرة المسحوبة في النهاية حمراء، ولتكن B حادثة كون الكرة المسحوبة أولاً حمراء (R).

عند سحب كرة من الصندوق I ووضعها في الصندوق II يصبح في الصندوق II :

5R, 3w أو 4R, 4w حسب كون الكرة المسحوبة حمراء أو بيضاء .

إذاً الاحتمالات الشرطية في هذه الحالة هي :

$$P(A/B) = P(A/5R,3W \quad \text{يحتوي } II) = \frac{5}{8}$$

$$P(A/B^c) = P(A/4R,4w \text{ يحتوي } \quad II) = \frac{4}{8} = \frac{1}{2}$$

وكذلك :

$$P(B) = \frac{3}{5}$$

$$P(B^c) = 1 - P(B) = 1 - \frac{3}{5} = \frac{2}{5}$$

وباستخدام العلاقة :

$$P(A) = P(A \cap B) + P(A \cap \overline{B})$$

أي :

$$P(A) = P(B)P(A/B) + P(\overline{B})P(A/\overline{B})$$

نحصل على :

$$P(A) = \left(\frac{3}{5}\right)\left(\frac{5}{8}\right) + \left(\frac{2}{5}\right)\left(\frac{1}{2}\right) = \frac{23}{40}$$

تعريف :

لتكن A و B حادثتان مستقلتان $\Longleftrightarrow P(A \cap B) = P(A)\,P(B)$ فيما عدا ذلك فإن A و B حادثتان غير مستقلتين .

يطلق أحياناً على الحوادث المستقلة بأنها مستقلة إحصائياً أو مستقلة عشوائياً وأحياناً أخرى مستقلة في منطق الاحتمال (أو احتمالياً) ولكن في معظم الأحوال نستخدم أحداثاً مستقلة بدون استخدام التعبيرات السابقة عندما لا يوجد سوء فهم .

ملاحظة :

من الواضح أن التعريف دائماً يتحقق إذا كان $P(A) = 0$ أو $P(B) = 0$ وذلك بسبب أن :

$P(A \cap B) = 0$ لأن $(A \cap B) \subset A$ و $A \cap B \subset B$. وعليه فإن الطرف الأيسر والأيمن للعلاقة $P(A \cap B) = P(A)\,P(B)$ كلاهما يساوي الصفر وبالتالي كل منهما يساوي الآخر.

ونلاحظ من خلال ذلك أنه إذا كان A, B منفصلتين؛ فإنهما يكونان مستقلتين.

وفي حالة $P(A) = 0$ أو $P(B)=0$ لأنه في حالة الانفصال $P(A \cap B) = 0$ وبسبب الاستقلال $P(A \cap B) = P(A).P(B)$

فنحصل على النتيجة المطلوبة .

مثال (15) :

وعاء يحتوي أربع كرات مرقمة 1,2,3,4. اختيرت كرة واحدة عشوائياً من هذا الوعاء. وكانت الحوادث A و B و C معرفة كما يلي :

$A = \{1,2\}$ ، $B = \{1,3\}$ ، $C = \{1,4\}$ عندئذ نجد أن :

$$P(A) = P(B) = P(C) = \frac{2}{4} = \frac{1}{2}$$

ونجد أن :

$$P(A \cap B) = \frac{1}{4}$$

$$P(A \cap C) = \frac{1}{4}$$

$$P(B \cap C) = \frac{1}{4}$$

ومن تعريف استقلال حادثتين نجد أن C, B, A هي حوادث مستقلة مثنى مثنى حيث :

$$P(A \cap B) = P(A).P(B)$$
$$= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

من ناحية أخرى $A \cap B \cap C = \{ 1 \}$ ولذلك فإن :

$$P(A \cap B \cap C) = P(A).P(B).P(C) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

هذا يعني أن هذه النتيجة تفتقر إلى الاستقلال الكلي للحوادث A و B و C.
هذا المثال يوضح السبب في وضع الشرط الثاني في التعريف التالي :

تعريف :

إذا كانت الحوادث A, B, C حوادث مستقلة تبادلياً، فإن هذه الحوادث تكون مستقلة مثنى مثنى : أي :

$$\left.\begin{aligned} P(A \cap B) &= P(A)P(B), \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \end{aligned}\right\}$$

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

التعريف السابق يمكن تعميمه إلى أربعة أو أكثر من حوادث مستقلة .

مثال (16) فإذا فشل الجهاز A، وكان 0.15، علماً أن ه الأجهزة عبارة عن حوادث مستقلة، وفما هو احتمال ؟ $\overline{C}\,\overline{B}\,\overline{A}$

مثال (17) :

ثلاثة صناديق : I, II, III يحتوي كل منها على كرات حمراء (R) وبيضاء (W) كالتالي:

I : 2R , 3W
II: 2R, 1W
III: 1R, 3W

عند المحاولة i، سحبنا كرة واحدة من الصندوق i حيث أن j= I, II, III.
وبفرض أن المحاولات مستقلة، أوجد :

$$P(\{R, R, R\}) \quad P(\{W, R, W\}) \quad , \quad P(\{R, W, W\}) \quad ,$$

حيث الحروف وضعت بالترتيب ، أي أن {R, W, W} تعني حمراء في المحاولة الأولى وبيضاء في الثانية وبيضاء في الثالثة .

الحل : باستعمال تعريف الاستقلال :

$$P(\{R,R,R\}) = \left(\frac{2}{5}\right)\left(\frac{2}{3}\right)\left(\frac{1}{4}\right) = \frac{4}{60} = \frac{1}{15}$$

$$P(\{W,R,W\}) = \left(\frac{3}{5}\right)\left(\frac{2}{3}\right)\left(\frac{3}{4}\right) = \frac{18}{60} = \frac{9}{30} = \frac{3}{10}$$

$$P(\{R,W,W\}) = \left(\frac{2}{5}\right)\left(\frac{1}{3}\right)\left(\frac{3}{4}\right) = \frac{6}{60} = \frac{1}{10}$$

1-6- : قاعدة بايز Bayes' Theorem :

هي إحدى الطرائق المناسبة لحساب احتمال حادث (E) إذا كان هذا الحادث يأتي من قبل عدة حوادث متعددة، أي علينا تحديد كل مرحلة من مراحل التجربة الإحصائية ولتقريب فكرة هذه القاعدة إلى الأذهان نأخذ المثال الآتي :

مثال (18) :

ثلاثة صناديق تحتوي على دوائر الكترونية صغيرة جداً، الصندوق B_1 يحتوي على دائرتين بلون أحمر (R) و 4 بيضاء (W)، والصندوق B_2 يحتوي على دائرة واحدة حمراء (R) ودائرتين بلون أبيض (W)، كما أن الصندوق B_3 يحتوي على 5 دوائر حمراء (R) و4 بيضاء (W)، وبفرض أن احتمالات اختيار الصناديق غير متساوية، وكانت كمايلي:

$P(B_1) = \frac{1}{3}$ $\quad P(B_2) = \frac{1}{6}$ $\quad P(B_3) = \frac{1}{2}$ حيث أن B_1, B_2, B_3 هي حوادث اختيار الصناديق على الترتيب . هذه التجربة تتكون من اختيار صندوق معين بالاحتمالات السابقة وبعد ذلك اختيار دائرة معينة عشوائياً من ذلك الصندوق. دعنا الآن نحسب احتمال الحادثة R (سحب دائرة حمراء) وليكن P(R). لاحظ أن P (R) يعتمد على اختيار أي من هذه الصناديق، وبعد

ذلك على احتمال سحب دائرة حمراء من الصندوق المختار، هذا يعني أن الحادثة R هي عبارة عن اجتماع الحوادث المتنافية الآتية:

$B_1 \cap R, B_2 \cap R, B_3 \cap R$

وبالتالي فإن :

$$P(R) = P(B_1 \cap R) + P(B_2 \cap R) + P(B_3 \cap R)$$

ومن تعريف الاحتمال الشرطي نحصل على :

$$P(R) = P(B_1)\,P(R/B_1) + P(B_2)\,P(R/B_2) + P(B3)\,P(R/B_3)$$

$$= \left(\frac{1}{3}\right)\left(\frac{2}{6}\right) + \left(\frac{1}{6}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{5}{9}\right) = \frac{4}{9}$$

وعلى افتراض أن نتيجة التجربة كانت فعلاً دائرة الكترونية حمراء فمن أي صندوق تم اختيارها؟ . وللإجابة على هذا التساؤل فإننا نحسب الاحتمالات الشرطية التالية :

$P(B_1/R), P(B_2/R), P(B_3/R)$ وهي :

$$P(B_1/R) = \frac{P(B_1 \cap R)}{P(R)}$$

$$= \frac{P(B_1)P(R/B_1)}{P(B_1)P(R/B_1) + P(B_2)P(R/B_2) + P(B_3)P(R/B_3)}$$

$$= \frac{(1/3)(2/6)}{(1/3)(1/6) + (1/2)(4/9) + (1/2) + (5/6)}$$

$$= \frac{2}{8} = \frac{1}{4}$$

$$P(B_2/R)=\frac{P(B_2)P(R/B_2)}{P(B_1)P(R/B_1)+P(B_2)P(R/B_2)+P(B_3)P(R/B_3)}$$
$$=1/8$$

الاحتمال الأخير :

$$P(B_3/R)=\frac{P(B_3)P(R/B_3)}{P(B_1)P(R/B_1)+P(B_2)P(R/B_2)+P(B_3)P(R/B_3)}$$
$$=\frac{(1/2)(5/9)}{\frac{4}{9}}=\frac{5}{8}$$

تمارين غير محلولة

تمرين (1) :

في أحد الصفوف (30) طالباً وطالبة، منهم (10) طلاب ذكور والباقي إناث، وإن نصف أعين الطلاب والطالبات من اللون الأزرق، اختير أحد أعضاء الصف بطريقة عشوائية، فما هو احتمال أن يكون العضو المختار طالباً وعيناه زرقاوين.

تمرين (2) :

لدى تاجر ثلاث مجموعات من المصابيح .

أ ــ تحتوي المجموعة الأولى على (500) مصباح منها (2%) معطوبة.

ب ــ تحتوي المجموعة الثانية على (300) مصباح منها (5%) معطوبة.

ج ــ تحتوي المجموعة الثالثة على (200) مصباح منها (1%) معطوبة .

وقد تم خلط تلك المصابيح جيداً ثم اختير مصباحاً بصورة عشوائية، فتبين أنه معطوباً.

والمطلوب :

ما احتمال انتماء ذلك المصباح إلى كل مجموعة من المجموعات الثلاث .

تمرين (3) :

يحتوي صندوق على (4) كرات حمراء مرقمة من (1) إلى (4)، و (3) كرات بيضاء مرقمة من (1) إلى (3)، وجميع هذه الكرات من الحجم نفسه .

والمطلوب :

أ ــ سحبت بصورة عشوائية كرة من الصندوق، فما احتمال أن تكون هذه الكرة حمراء أو تحمل رقم زوجي ؟.

ب ــ أعيدت الكرة المسحوبة سابقاً، ثم سحبت الكرات بالتتالي، ما احتمال تعاقب الكرات الحمراء والبيضاء في عملية السحب.

تمرين (4) :

ثلاث صيادين أطلقوا النار على هدف، الأول رمى طلقة واحدة والثاني رمى طلقتين والثالث أربع طلقات .

والمطلوب :

ــ ما احتمال إصابة الهدف .

ــ ما احتمال عدم إصابة الهدف .

تمرين (5) :

أطلق رامي ثلاث طلقات متلاحقة على هدف، إذا علمت أن احتمال إصابة الهدف (0.6) والمطلوب:

ــ ما هو احتمال إصابة الهدف بطلقة واحدة فقط .

ــ احسب عدد الطلقات الواجب رميها على الهدف ليكون احتمال إصابة الهدف (0.9) على الأقل.

تمرين (6) :

ثلاثة رجال يطلقون النار على هدف واحد وكل منهم يطلق طلقة واحدة فإذا كان احتمال إصابة الأول للهدف (0.4) واحتمال إصابة الثاني لهدف (0.5) واحتمال إصابة الثالث لهدف (0.1) والمطلوب :

ــ ما احتمال إصابة الهدف بطلقة واحدة .

ــ ما احتمال إصابة الهدف بطلقتين فقط .

ــ ما احتمال إصابة الهدف بثلاث طلقات .

ـــ ما احتمال إصابة الهدف بطلقة واحدة على الأقل .

تمرين (7) :

يوجد في أحد الصفوف الدراسة (30) طالب وطالبة، (12) طالب من بينهم (4) طلاب متفوقين و (18) طالبة من بينهن (8) طالبات متفوقات، اختير أحد الطلبة بصورة عشوائية ليكون ممثلاً عن هذا الصف، **والمطلوب** : ما احتمال أن يكون الممثل :

ـــ طالب متفوق .

ـــ طالبة متفوقة .

ـــ متفوق (طالب أو طالبة).

تمرين (8) :

في أحد المختبرات (10) حاسبات، (6) منها تعمل بلغة البيسك فقط و(3) تعمل بلغة الفورتران فقط و (1) تعمل بكلتا اللغتين، فإذا اختيرت إحدى الحاسبات بصورة عشوائية، ما احتمال:

ـــ أن تكون الحاسبة تعمل بلغة البيسك ولاتعمل بلغة الفورتران.

ـــ أن تكون الحاسبة تعمل بلغة البيسك والفورتران معاً.

ـــ أن تكون الحاسبة تعمل بلغة البيسك .

ـــ أن تكون الحاسبة تعمل بلغة الفورتران .

ـــ أن تكون الحاسبة تعمل بلغة البيسك أو الفورتران .

تمرين (9) :

ثلاثة متبارين a,b,c يصوبون معاً إلى هدف واحد واحتمال إصابتهم للهدف على الترتيب 1/4, 1/2, 2/5 .

والمطلوب :

ـــ ما احتمال إصابة اثنين منهم للهدف عندما يصوب الثلاثة معاً بنفس الوقت.

تمرين (10) :

لنفترض أنه لدينا مصنع للغزل والنسيج، حيث أن نسبة القطع المعطوبة في مرحلة الغزل هي (20%) وفي مرحلة النسيج (15%) وفي مرحلة الصباغة (10%)، سحب ثوب من الإنتاج الكلي عشوائياً فكان معطوباً.

والمطلوب :

ما احتمال أن يكون هذا العطب قد حصل في مرحلة الغزل .

تمرين (11) :

ليكن لدينا مصنع مكون من (3) مجموعات من الآلات M_1, M_2, M_3 ولنفترض أن (90%) من إنتاج المجموعة الأولى سلع جيدة و (80%) من إنتاج المجموعة الثانية هو أيضاً سلع جيدة و (25%) من إنتاج المجموع الثالثة هو سلع معطوبة .

والمطلوب :

أ ــ أثناء تجميع الإنتاج الكلي للسلع، سحبت سلعة من هذا الإنتاج عشوائياً، ما احتمال أن تكون السلعة المسحوبة معطوبة .

ب ــ أثناء تجميع الإنتاج الكلي للسلع سحبت سلعة عشوائياً فكانت معطوبة، ما احتمال أن تكون هذه السلعة قد أنتجت من المجموعة الأولى .

تمرين (12) :

يتم تصويب بندقيتين (I ، II) على نفس الهدف، أطلقت (9) طلقات من البندقية (I) في نفس الوقت الذي أطلقت فيه (10) طلقات من البندقية (II)، فإذا علمت أنه بالمتوسط من كل (10) طلقات من البندقية (I) (8) فقط تصيب الهدف ومن البندقية (II) (7) فقط تصيب الهدف، وأثناء التصويب أصيب الهدف بطلقة ولم يكن معلوماً من أي بندقية جاءت الطلقة.

والمطلوب :

١ ــ ما هو احتمال إصابة الهدف من البندقية II.

٢ ــ ما هو احتمال إصابة الهدف من البندقية I.

٣ ــ ما هو احتمال إصابة الهدف بطلقة .

تمرين (13) :

لون الوجهان الأول والثاني من حجر النرد باللون الأحمر (R) والوجهان الثالث والرابع باللون الأصفر (J) والوجهان الخامس والسادس باللون الأزرق (B) .

والمطلوب :

1 ــ حدد الفضاء الاحتمالي لتجربة رمي حجر النرد مرتين متتاليتين وحسب احتمال الحصول على كل من نتائج هذه التجربة .

تمرين (14) :

صندوقان B, A يحوي الصندوق A على (9) أوراق مرقمة من (1 إلى 9) ويحوي الصندوق B على (17) ورقة مرقمة من (1 إلى 17)، اختير صندوقاً من الصندوقين بطريقة عشوائية واختيرت منه ورقة، والمطلوب :

1 ــ ما احتمال أن تكون الورقة المختارة تحمل رقم زوجي .

2 ــ إذا كان رقم الورقة المختارة زوجياً، فما احتمال أن تكون الورقة المختارة من الصندوق A .

تمرين (15):

شركة أحذية لها ثلاثة مصانع إنتاجية، المصنع الأول ينتج (40%) من إنتاج الشركة وفيه (2%) من البضاعة معيبة، والمصنع الثاني ينتج (35%) من إنتاج الشركة وفيه (10%) من البضاعة معيبة، أما المصنع الثالث فينتج (25%) من إنتاج الشركة وفيه (40%) من البضاعة تكون معيبة، فإذا اشتريت حذاء من إنتاج هذه الشركة، فما احتمال أن يكون :

1 ــ معيباً 2 ــ غير معيب 3 ــ إذا اشتريت حذاء من إنتاج هذه الشركة ووجدته معيباً، فما هو احتمال أن يكون من المصنع الأول أو الثاني .

تمرين (16) :

إذا علمت أن الصندوق (A) فيه (3) كرات بيضاء و (5) كرات سوداء وفي الصندوق (B) (5) كرات سوداء و (3) كرات بيضاء، سحبنا عشوائياً مع الإعادة كرتين من الصندوق (A)، وسحبنا بدون إعادة كرتين من الصندوق (B).

والمطلوب :

1 ــ احسب احتمال الحصول على (4) كرات بيضاء .

2 ــ احسب احتمال الحصول على كرة سوداء واحدة على الأقل .

تمرين (17) :

يحتفظ مشفى بسيارتي إسعاف للطوارئ واحتمال أن تكون السيارة جاهزة للتحرك عند الحاجة إليها هو (0.9)، وإذا علمت أن توفر أحد السيارتين مستقل عن الأخرى .

والمطلوب :

1 ــ أوجد احتمال أن لا تتوفر أيّ منهما .

2 ــ ما هو احتمال تلبية الطلب لسيارة إسعاف في حالة الطوارئ .

تمرين (18) :

إذا كان لدينا احتمال هطول مطر في يوم معين (0.1)، واحتمال وجود رياح نشطة في ذلك اليوم (0.05)، واحتمال وجود مطر ورياح نشطة (0.03).

والمطلوب :

1 ــ أوجد احتمال وجود مطر أو رياح نشطة في ذلك اليوم.

2 ــ أوجد احتمال هطول مطر في ذلك اليوم علماً أن الرياح نشطة .

تمرين (19) :

كيس يحتوي على (25) كرة متماثلة منها (10) كرات بيضاء و (5) كرات سوداء والباقي حمراء، سحبت منه ثلاث كرات الواحدة تلو الأخرى عشوائياً، أوجد احتمال أن تكون الكرات الثلاث من نفس اللون .

إذا كان : 1 ــ السحب مع الإعادة ؟ 2 ــ السحب بدون إعادة .

تمرين (20) :

يتنافس أحمد، محمد، حسين على المرتبة الأولى في امتحان مقرر الإحصاء، فإذا كان احتمال فوز أحمد = (2/3) من احتمال فوز محمد، واحتمال فوز حسين ضعف احتمال فوز محمد، فما هو احتمال فوز كل منهم ؟.

تمرين (21) :

قذفت قطعة نقود عدة مرات في الهواء فكان احتمال ظهور الصورة ضعف احتمال ظهور الكتابة في كل مرة، والمطلوب : أوجد احتمال ظهور الصورة والكتابة.

تمرين (22) :

إذا كانت (B, A) حادثتين ، حيث كان $P(A\cup B)=3/4$، $P(B)2/5$، $P(A)=x$. أوجد قيمة x في كل من الحالات التالية :

1 ــ إذا كان A و B متنافيتين .

2 ــ إذا كان B, A مستقلتين .

تمرين (23):

يتضمن صندوق (6) قطع حمراء مرقمة من (1) إلى (6) و(6) قطع بيضاء مرقمة من (1) إلى (6) سحبنا قطعة بصورة عشوائية. **والمطلوب :**

١ ــ ما احتمال أن تكون حمراء .

٢ ــ ما احتمال أن تكون عليها رقم زوجي.

٣ ــ ما احتمال أن تكون حمراء وعليها رقم زوجي .

تمرين (24) :

اختير شخصان على التوالي من مجموعة بها (7) رجال و (5) نساء، فإذا كان الشخص الثاني امرأة، فما احتمال أن يكون الشخص الأول امرأة أيضاً .

تمرين (25) :

ثلاثة صفوف في كل منها (25) طفلاً، أطفال الصف الأول كلهم أولاد والثاني كلهم بنات والثالث مناصفة، اختير صفاً عشوائياً ومنه اختير طفلاً، فما هو احتمال :

1 ــ أن يكون ولد .

2 ــ إذا علمنا أنه ولد، فما هو احتمال أن يكون من الصف الأول .

تمرين (26) :

صندوقان في كل منهما (3) كرات بيضاء و (2) كرة حمراء نقلت كرة من الصندوق الأول إلى الصندوق الثاني وسحبت كرة من الصندوق الثاني، فما هو احتمال أن تكون الكرة المسحوبة بيضاء.

تمرين (27) :

تعرّف الحوادث التالية على فضاء العينة (S) رمي حجر النرد لمرة واحدة .

الحادثة (A) : حصولنا على عدد زوجي .

الحادثة (B) : حصولنا على عدد يقبل القسمة على (3) بدون باقي .

والمطلوب : أوجد الاحتمالات الآتية :

1- $P(A \cup B)$

2- $P(A \cap B)$

تمرين (28) :

بين وجه الخطأ في كل من العبارات التالية :

1 ــ الاحتمال التجريبي = $\frac{\text{عدد الحالات الملائمة}}{\text{عدد الحالات الممكنة}}$

2 ــ الاحتمال $0 \le P(A) \le 1$

3 ــ الاحتمال لا يكون سالباً .

4 ــ الاحتمال النظري مبني على المنطق الرياضي وليس نتيجة التجربة.

5 ــ احتمال النجاح + احتمال الإخفاق ≠ ١ .

6 ــ الاحتمال يستخدم للتعبير عن حدث مؤكد الحدوث

7ــ الإحصاء الرياضي مبني على نظرية الاحتمالات ويهتم بالمستقبل غير المؤكد

8 ــ الحادثة كل عملية تؤول إلى ملاحظة أو قياس ظاهرة ما .

9 ــ في الحوادث المتنافية بالتبادل :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

10 ــ عندما يكون عدد مرات التجربة كبيراً فإن الاحتمال التجريبي يقترب من الاحتمال النظري.

11 ــ التكراري النسبي = $\frac{\text{عدد الحالات الملائمة}}{\text{عدد الحالات الممكنة}}$

12 ــ احتمال وقوع الحادث الأكيد يساوي الصفر .

13 ــ ظهور رقم (7) عند رمي حجر النرد هو حادث أكيد .

14 ــ في الحوادث المستقلة $P(A \cap B) = P(A)\,P(B)$

15 ــ المتغير العشوائي هو نتائج التجربة الإحصائية غير المعروفة مسبقاً.

16 ــ المتغير العشوائي المنفصل يأخذ قيماً صحيحة فقط .

17 ــ من شروط تابع التوزيع الاحتمالي

18 ــ الأمل (التوقع) الرياضي هو قيمة المتغير مضروباً باحتمال الحصول عليها .

19 ــ إن فضاء العينة لرمي حجر النرد لمرة واحدة .

$S : \{ 1,2,3,4,5,6, \}$

20 ــ عن فضاء العينة لرمي قطعة نقود مرتين متتاليتين هو :

$S : \{ HT , TH , TT \}$

21 ــ إن فضاء العينة لرمي قطعة نقود سوية ثلاث مرات متتالية هو على النحو التالي:

S : {(HHH), (HHT), (HTH), (HTT), (TTH), (THH), (THT) }

22 ــ من عيوب التعريف الإحصائي للاحتمال (التكرار النسبي)، أن يكون فضاء العينة منتهياً، ولكل عنصر من عناصر فضاء العينة نفس احتمال الحدوث .

23 ــ إذا عرفنا على فضاء العينة لرمي حجر النرد مرة واحدة، الحادث A على النحو التالي : A = { 7 }، فإن هو حادثاً أكيداً.

24 ــ احتمال أن ينجح خالد في امتحان الفيزياء هو (-0.95) .

25 ــ احتمال أن يفوز الفريق الوطني لكرة القدم في مباراته القادمة مع منتخب الجامعة هو (0,75) واحتمال أن يتعادل (0,09) واحتمال أن يتعادل أو يفوز (0,95) .

26 ــ احتمال أن ينجح عارف في الإحصاء هو (0,9) واحتمال أن ينجح في مقرري الإحصاء والرياضيات (0,95) .

27 ــ احتمال أن تستقبل عيادة طبيب أقل من (5) مرضى في فترة ما قبل الظهر هو (0,62) واحتمال أن تستقبل (5) مرضى أو أكثر هو (0,25) .

28ــ $P(A \cap B) = 1 - P(A \cup B)$

29ــ $$P(A/\overline{B}) = \frac{P(A) - P(A \cap B)}{1 - (\overline{B})}$$

30ــ $$P(B/A) = \frac{(P(A \cap B)}{P(A)}$$

تمرين (29) :

أجب باختصار عن الأسئلة التالية :

1 ــ ما الفرق بين الحادثة والتجربة ؟

2 ــ ما الفرق بين الاحتمال النظري والاحتمال التجريبي .

3 ــ ما الفرق بين المتغيرات العشوائية المنفصلة والمتغيرات العشوائية المتصلة ؟

4 ــ ميز بين المفاهيم التالية :

الحادث الأكيد، الحادث المستحيل، الحادث الممكن .

5 ــ متى يكون التابع ، تابع كثافة احتمالي ؟

الفصل الثاني

المتغيرات العشوائية والتوزيعات الاحتمالية

Random Varibles and Probablity Distribution's

- مقدمة
- المتغيرات العشوائية : المتغيرات المتقطعة والمستمرة
- تابع التوزيع التجميعي (التراكمي)
- التوزيعات الاحتمالية المتقطعة (الخاصة)

1 ـ محاولة بيرنولي .

2 ـ توزيع ثنائي الحدين

3 ـ توزيع بواسون .

- التوزيعات الاحتمالية (المستمرة) .
- التوزيع الطبيعي وخواصه :
  - التوزيع الطبيعي المعياري
  - تحديد قيم Z باحتمال معلوم $(1-\alpha)$
- تمارين غير محلولة.

2 ــ 1 ــ مقدمة :

عند إجراء التجارب الإحصائية وتعيين فضاء العينة لكل منها، ليس من الضروري دراسة نقاط فضاء العينة بالتفصيل ، وبمعنى آخر لا يكون الاهتمام منصب على النتائج في حد ذاتها، وإنما على القيم العددية المرتبطة بهذه النتائج الممكنة، لذلك فإن هذه القيم العددية هي ما نعبر عنه بقيم المتغير العشوائي، الذي هو عبارة عن نتائج التجربة الإحصائية غير المعروفة النتائج مسبقاً، مثل :

ــ عدد المرضى الذين يترددون إلى إحدى العيادات الطبية يومياً، ما هو إلا متغير عشوائي يمكن أن يأخذ رقماً ما يعتمد بصورة عامة على المصادفة؛

ــ مبيعات مؤسسة تجارية ما يومياً يعتبر متغيراً عشوائياً ؛

ــ تجربة رمي حجري النرد، يمكن أن نقرن كل نتيجة بمجموع الرقمين اللذين يظهران على الوجهين العلويين لحجري النرد وهذا المجموع يأخذ واحدة من القيم التالية :

$$2 , 3 , 4 , 5 , 6, 7 , 8 , 9 , 10 , 11 , 12$$

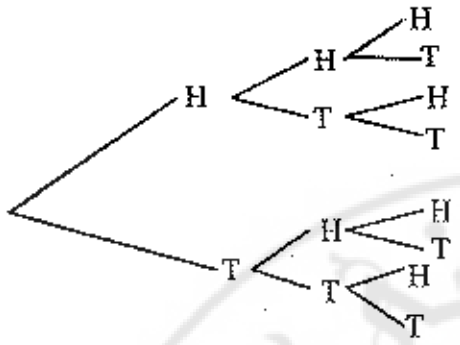
ويعرف المتحول العشوائي رياضياً : بأنه تابع أو تطبيق منطلقه فضاء العينة ومستقره مجموعة الأعداد الحقيقية R .

$$X : S \to R$$

ونرمز إلى المتغير العشوائي بأحد الحروف x, z, y .

تمرين (1) :

ألقيت قطعة نقود ثلاث مرات وكان المتغير العشوائي x يمثل عدد مرات ظهور صورة. والمطلوب : وضح القيم الممكنة للمتغير العشوائي **بطريقة الشجرة البيانية .**



$HHH \rightarrow X = 3$
$HHT \rightarrow X = 2$
$HTH \rightarrow X = 2$
$HTT \rightarrow X = 1$
$THH \rightarrow X = 2$
$THT \rightarrow X = 1$
$TTH \rightarrow X = 1$
$TTT \rightarrow X = 0$

من خلال ذلك، نلاحظ أنه عند إلقاء قطعة النقود مرة واحدة، فإن فضاء العينة:

$$S = \{ x : xi \in \{ H , T \}]$$

واحتمال الحصول على كل من نتائج هذه التجربة :

X_i	0	1	2	3
$f(x_i)$	1/8	3/8	3/8	1/8

وبذلك نقول عن فضاء عينة (S) أنه فضاء عينة متقطع إذا كانت S مجموعة منتهية أو مجموعة غير منتهية ولكنها قابلة للعد .

المتغيرات العشوائية :

تقسم المتغيرات العشوائية إلى نوعين رئيسيين :

ـ المتغيرات العشوائية المتقطعة .

ـ المتغيرات العشوائية المستمرة .

ونقول عن متحول عشوائي (x) بأنه متحول عشوائي متقطع، إذا كان منطلقه (S) متقطعاً، أي إذا كانت (S) مجموعة منتهية أو قابلة للعدد، لذلك

يأخذ المتغير العشوائي المتقطع قيماً صحيحة فقط. مثال { حجر النرد، أوراق اللعب، عدد أفراد الأسرة، عدد سكان مدينة ما }.

وفي خلاف ذلك نقول عن متحول عشوائي (x) بأنه متحول عشوائي مستمر إذا كان فضاء العينة مستمراً مثل أوزان الطلاب أو أطوالهم، عدد الكريات في الدم .. الخ .

2 ـ 2 ـ التوزيع الاحتمالي للمتحول العشوائي المتقطع :

إذا كان x متحولاً عشوائياً متقطعاً معرفاً على فضاء العينة S حيث صورته مجموعة منتهية:

$$X(s) = \{ x_1, x_2, x_3, \ldots . x_n \}$$

فإن تابع الاحتمال للمتحول العشوائي (X) يكتب على الشكل التالي :

$$f(x_i) = P(X = x_i)$$

حيث x_i تقابل P_i :

أو على شكل جدول :

X	X_1	X_2	X_3	X_n
$f(x_i)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	$f(x_n)$

إن الجدول يمكن تسميته بجدول الكثافة الاحتمالية للمتحول العشوائي (x) ونسمي التابع $f(x_i) = P(x = x_i)$ بتابع الكثافة الإحتمالية أو تابع التوزيع الاحتمالي، إذا تحقق الشرطين التاليين :

1 - $0 \le f(x_i) \le 1$

2- $\sum f(x_i) = 1$

لنعد إلى مثالنا السابق: عند إلقاء قطعة النقود ثلاث مرات، كان المتغير العشوائي (x) يمثل عدد مرات ظهور الصورة، **والمطلوب :**

1 ــ تحديد فضاء العينة ؟

2 ــ كتابة تابع الاحتمال

الحل :

ــ فضاء العينة $S = \{ HHH, HHT, HTH, HTT, THH, THT, TTT\}$

حيث أن $S = 8$ عناصر .

تابع الإحتمال :

X_i	0	1	2	3
$f(x_i)$	1/8	3/8	3/8	1/8

وأنه يحقق الشرطين السابقين .

$1- \sum f(x_i) = 1/8 + 3/8 + 3/8 + 1/8 = 1$

$2- 0 \le f(x_i) \le 1$

تمرين (2) :

صندوقان يحوي الصندوق الأول أربع بطاقات مرقمة كالتالي (1,2,3,4) ويحوي الصندوق الثاني ثلاث بطاقات مرقمة (1,2,3)، نسحب بطاقة من كل صندوق ونرمز بـــ x للمتغير العشوائي الدال على مجموع الرقمين، **والمطلوب :**

1 ــ تحديد فضاء العينة .

2 ــ كتابة تابع الاحتمال .

الحل :

فضاء العينة على النحو التالي :

	1	2	3	4
1	1,1	1,2	1,3	1,4
2	2,1	2,2	2,3	2,4
3	3,1	3,2	3,3	3,4

وتابع الكثافة الاحتمالي

X_i	2	3	4	5	6	7
$f(x_i)$	1/12	2/12	3/12	3/12	2/12	1/12

وهو يحقق الشرطين لتابع الكثافة الاحتمالية :

$1- \sum f(x_i) = 1/12 + 2/12 + 3/12 + 3/12 + 2/12 + 1/12 = 12/12 = 1$

$2- 0 \le f(x_i) \le 1$

2 ــ 3 ــ الأمل الرياضي Mathematical Expecation :

يعد الأمل الرياضي للمتغير العشوائي من أهم المواضيع في الاحتمالات، ولتقريب فكرة الأمل الرياضي إلى الأذهان . نفترض أن (3) قطع نقدية رميت (10) مرات.

ــ فإذا كان عدم ظهور (H) قد حدث مرتين .

ــ وظهور (1H) حدث ثلاث مرات .

ــ وظهور (2H) حدث ثلاث مرات .

ــ وظهور (3H) حدث مرتين .

فلو أخذنا متوسط ظهور H في الرقعة الواحدة لقطع النقود الثلاث لوجدنا أن:

$$\frac{(0\times 2)+(1\times 3)+(2\times 3)+(3\times 2)}{10}=1.5H$$

وباختصار فإن الأمل الرياضي هو عبارة عن قيمة المتحول مضروباً باحتمال الحصول عليها، ويعطى التوقع الرياضي لمحتول عشوائي متقطع x يأخذ القيم x_i المقابلة للإحتمالات P_i بالعلاقة :

$$E(x)=\sum_{i=1}^{n} x_i\ f(x_i) \qquad i = 1, 2, 3, 4, \ldots n$$

ويطلق أحياناً اسم المتوسط للمتحول العشوائي x على المقدار E (x).

ففي تمريننا الأول يكون الأمل الرياضي أو التوقع لعدد مرات ظهور صورة من خلال ثلاث رميات هو :

E (x) = [(0). (1/8)] + (1).(3/8) + (2).(3/8) + (3). (1/8) = (12/8) = 3/2

والأمل الرياضي في تمريننا (2) هو :

E (x) = (2. 1/12) + (3. 2/12) + (4. 3/12) + (5. 3/12) + (6. 2/12) + (7. 1/12)

= 54 / 12 = 4.5

التباين : Variance

إن تباين المتغير العشوائي المتقطع X الذي يأخذ القيم x_i والمقابلة للاحتمالات ذات $f(x_i)$ هو المقدار:

$$\mathrm{var}(x)=\sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)$$

حيث إن : $E(x)=\mu$

ولحساب التباين في تمريننا (1) السابق:

x_i	$f(x_i)$	$x_i - \mu$	$(x_i - \mu)^2$	$(x_i-\mu)^2 f(x_i)$
0	1/8	0-3/2 = -3/2	2.25	0.28125
1	3/8	1 – 3/2 = -0.5	0.25	0.09735
2	3/8	2-3/2 = 0.5	0.25	0.09735
3	1/8	3-3/2 = 1.5	2.25	0.28125

$$\sum_{i=1}^{n} f(x_i) = 1 -$$

$$- \text{var}(x) = \sigma^2 = 0.28125 + (0.09735) \,.\, 2 + 0.28125 = 0.76$$

ويمكن تبسيط العلاقة السابقة :

$$\text{var}(x) = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i)$$

$$= \sum_{i=1}^{n} (x_i^2 - 2x_i\mu + \mu^2) f(x_i)$$

$$= \sum_{i=1}^{n} \left[x_i^2 f(x_i) - 2\mu \sum x_i f(x_i) + \mu^2 f(x_i)\right]$$

$$= \sum x_i^2 f(x_i) - 2\mu \sum x_i f(x_i) + \mu^2 \sum f(x_i)$$

حيث أن :

$$\sum x_i^2 f(x_i) = E(x^2)$$
$$\sum x_i f(x_i) = \mu$$
$$\sum f(x_i) = 1$$

لذلك تصبح العلاقة السابقة :

$$\text{var}(x) = E(x^2) - 2\mu^2 + \mu^2$$

$= E(x^2) - \mu^2$

وبتطبيق هذه العلاقة على تمريننا (1) السابق في حساب التباين نجد :

$$E(x^2) = \sum_{i=i}^{n} x_i^2 f(x_i)$$

$$= (0.\ 1/8) + (\ 1.3/8) + (4.\ 3/8) + (9.1/8) = 3$$

$$var(x) = E(x^2) - \mu^2$$

$$= 3 - (1.5)^2 = 3 - 2.25 = 0.75$$

وعليه يعرّف الانحراف المعياري لمحتول عشوائي متقطع (X)، بأنه الجذر التربيعي الموجب لتشتت (التباين) هذا المحتول ويرمز له بـــ δ_x ويمكن كتابة

$$\delta_x = \sqrt{var(x)} = \sqrt{0.75} = 0.866$$

2 ــ 4 ــ التوزيع المتجمع (التراكمي) :

إن تابع التوزيع المتجمع الاحتمالي F(x) لمتحول عشوائي متقطع تابع كثافته الاحتمالي $f(x_i)$ يعطى بالعلاقة :

$$F(x) = P(x \le x_i) = \sum_{i=1}^{n} f(x_i)$$

تمرين (3) :

لنفرض أن متغيراً عشوائياً منقطعاً له التوزيع التالي :

x_i	1	2	3	4
$f(x_i)$	0.1	0.2	0.3	0.4

اكتب تابع التوزيع المتجمع ؟.

الحل :

يعطي تابع التوزيع المتجمع بالجدول التالي :

x_i	1	2	3	4
$f(x_i)$	0.1	0.2	0.3	0.4
$F(x)$	0.1	0.3	0.6	1

أي أن :

$F(1) = f(x_1) = 0.1$
$F(2) = f(x_1) + f(x_2) = 0.1 + 0.2 = 0.3$
$F(3) = f(x_1) + f(x_2) + f(x_3) = 0.1 + 0.2 + 0.3 = 0.6$
$F(4) = f(x_1) + f(x_2) + f(x_3) + f(x_4)$. $1+0.2 + 0.3 +0.4 = 1$

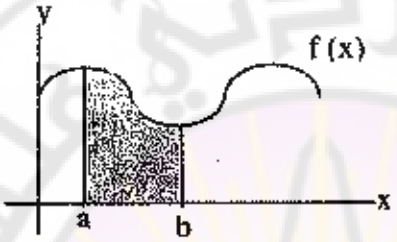
2 ـ 5 : المتغيرات العشوائية المستمرة :

فيما سبق أشرنا إلى أن الأوزان، والأطوال، وكميات الأمطار الهاطلة، هي أمثلة على المتغيرات العشوائية المستمرة، ومثل هذه المتحولات تشكل بدون انقطاع نقاطاً على محور موجه أو مجالات بين قيمة وأخرى. ولايمكننا تخصيص احتمال $f(x_i) = P_i$ لكل قيمة من قيم المتحول العشوائي x_i كما في حالة المتحولات العشوائية المتقطعة، وإن تابع الكثافة الاحتمالية للمتغير العشوائي المتقطع $f(x_i)= P(X = x_i)$ يصبح عديم المعنى بالنسبة للمتغيرات العشوائية المستمرة، أي أن المتغير المستمر يأخذ قيمة في مدى محدد، لذلك توجد لأي متغير عشوائي مستمر دالة $f(x)$ بحيث أن المساحات تحت منحني هذه الدالة (التابع) تعطي الاحتمالات المناظرة للفترات التابعة لها على المحور الأفقي .

تابع الكثافة الاحتمالية للمتحول العشوائي المستمر (X) :

يعبر عن احتمال الحادث $P (a \le x \le b)$ بالنسبة للمتحول العشوائي المستمر بالعلاقة:

$P(a \le x \le b) = \int_a^b f(x).dx$ ، وإنه لإيجاد الاحتمال أو المساحة $P(a \le x \le b)$ والمعرفة بالدالة f(x) لابد من حساب المساحة تحت منحني هذه الدالة شكل (1 ــ 3) .



الشكل (1 ــ 3)

وإن احتمال الحادثة :

$$P(x = a) = \int_a^b f(x).dx$$

وإن مجموعة الحوادث :

$$a < x < b\,,\ a < x \le b\,,\ a \le x < b\,,\ a \le x \le b$$

لها جميعها نفس قيمة الاحتمال والمساو $\int_a^b f(x).dx$

ونقول عن تابع ما أنه تابع كثافة احتمالية لمتغير عشوائي مستمر، إذا كانت قيمة التابع f(x) تحقق الشرطين التاليين :

1 – $\int_{-\infty}^{+\infty} f(x).dx = 1$

2 – $f(x) \ge 0$

تمرين (4) :

ليكن لدينا تابعاً معيناً معطى بالعلاقة :

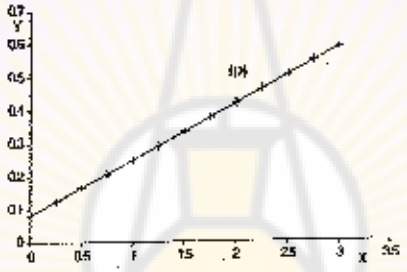
$$f(x)=\begin{cases}1/6x+1/12 & \forall x:0\le x\le 3\\ 0 & \text{فيما عدا ذلك}\end{cases}$$

والمطلوب :

1 ــ ارسم هذا التابع .

2 ــ أثبت أن هذا التابع هو تابع كثافة احتمالي لمتحول عشوائي مستمر .

الحل :



ــ لكي يكون هذا التابع تابع كثافة احتمالي لمتحول عشوائي مستمر، لابد أن يحقق الشرطين التاليين :

1 ــ $f(x)\ge 0$ وذلك من أجل جميع قيم x_i

2 ــ $\int_0^3 f(x).dx=1$

ــ كما نلاحظ أن الشرط الأول محقق وذلك من أجل جميع قيم x .

$$P(0 \le x \le 3) = \int_0^3 f(x).dx$$

$$= \int_0^3 (\tfrac{1}{6}x + \tfrac{1}{12}).dx = \tfrac{1}{6}\int_0^3 (x + \tfrac{1}{2}).dx$$

$$= \tfrac{1}{6}\left[\tfrac{x^2}{2} + \tfrac{x}{2}\right]_0^3 = \tfrac{1}{12}\left[x^2 + x\right]_0^3$$

$$= \tfrac{1}{12}\left[(9+3)-(0)\right] = \tfrac{1}{12}[12] = \tfrac{12}{12} = 1$$

تمرين (5) :

إذا كان لدينا التابع المعرّف بالعلاقة التالية :

$$f(x)\begin{cases} 2/9\ X & \forall : 0 \le x \le 3 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

أثبت أن تابع كثافة احتمالي لمتحول عشوائي مستمر .

الحل :

الشرط الأول محقق :

$$f(x) \ge 0$$

$$\int_0^3 \tfrac{2}{9}x.dx = \tfrac{2}{9}\int_0^3 x.dx$$

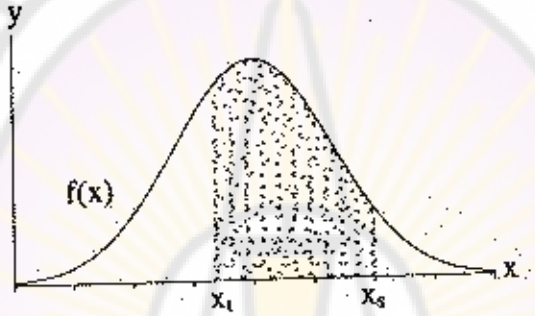
$$= \tfrac{2}{9}\left[\tfrac{x^2}{2}\right]_0^3 = \tfrac{2}{18}[(9)-(0)] = \tfrac{18}{18} = 1$$

2 ــ 6 ــ تابع التوزيع التجميعي (التراكمي) :

إذا كان x متحولاً مستمراً ومحدداً بقانون التوزيع f(x) فيمكننا أن نحدد تابع التوزيع التجميعي F(x) بالعلاقة :

$$F(x_s) = P(x \le x_s) \int_{-\infty}^{X_s} f(x).d$$

وهذا التابع يحدد الاحتمال أو المساحة الواقعة على يسار النقطة X_s ويستخدم التابع F(x) في حساب المساحة أو الاحتمال لأية قيمة أقل من الحد الأعلى وأكبر من الحد الأدنى، ويمكننا حساب هذا الاحتمال بطريقتين:



الشكل (2 ــ 3)

1 – $$F(x) = \int_{X_1}^{X_s} f(x).dx$$

2 – $$F(x) = \int_{X_1}^{+\infty} f(x).dx - \int_{X_s}^{X_2} f(x).dx$$

ومن أجل حساب الاحتمال أو المساحة :

$$P(x \geq x_s) = \int_{X_s}^{+\infty} f(x).dx$$

$$= \int_{-\infty}^{+\infty} f(x).dx - \int_{X_s}^{+\infty} f(x).dx$$

$$= 1 - F(x_s)$$

خصائص تابع التوزيع التجميعي :

1 ــ إن تابع التوزيع التجميعي هو تابع متزايد دوماً أي أن $F(x) \geq 0$ ؛

2 ــ $F(+\infty) = 1$ ، $x \to +\infty$ ؛ $F(-\infty) = 0$ ، $x \to -\infty$

3 ــ إذا كان (Xs) قابلاً للاشتقاق في مجال تعريفه، فإن مشتقه بالنسبة لـ x يساوي تابع الكثافة عند هذه النقطة، أي :

$$F'(x_s) = \frac{\Delta F(x_s)}{\Delta x} = f(x_s)$$

تمرين (6) :

احسب الاحتمالات التالية :

$P(x \leq 1)$; $P(2 \leq x \leq 3)$; $P(1 \leq x \leq 2)$

وذلك بالنسبة إلى التمرين (4) السابق ؟ .

الحل :

$$P(1 \le x \le 2) = \int_1^2 (\tfrac{1}{6}x + \tfrac{1}{12}).dx$$

$$\tfrac{1}{6}[x + \tfrac{1}{2}]_1^2 = \tfrac{1}{6}[\tfrac{x^2}{2} + \tfrac{x}{2}]_1^2$$

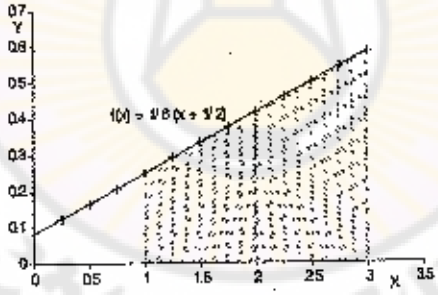
$$= \tfrac{1}{12}[x^2 + x]_1^2 = \tfrac{1}{12}[(4+2) - (1+1)]$$

$$= \tfrac{1}{12}[6 - 2] = \tfrac{4}{12} = \tfrac{1}{3}$$

$$P(2 \le x \le 3) = \tfrac{1}{12}[x^2 + x]_2^3 = \tfrac{1}{12}\{(9+3) - (4+2)]$$

$$= \tfrac{1}{12}[12 - 6] = \tfrac{6}{12} = \tfrac{1}{2}$$

$$P(x \le 1) = \tfrac{1}{12}[1+1] = \tfrac{2}{12}$$



الشكل (2 ــ 4)

تمرين (7) :

إذا كان لدينا التابع التالي :

$$F(x)\begin{cases} 2/9\ X & \forall : 0 \le x \le 3 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

احسب : $p(x \le 2.5)$ ، $P(x \ge 2.5)$

الحل :

$$P(x \le 2.5) = \int_{a} f(x).dx = \frac{2}{9}\int_{0}^{2.5} x.dx$$

$$= \frac{2}{9}\left[\frac{x^2}{2}\right]_0^{2.5}$$

$$= \frac{2}{9}\left[\frac{6.25}{2}\right] = \frac{6.25 \times 2}{18} = \frac{6.25}{9}$$

$$P(x \ge 2.5)$$

$$P(x \ge 2.5) = \int_{2.5}^{3} \frac{2}{9} x.dx = \left[\frac{x^2}{2}\right]_{2.5}^{3}$$

$$= \frac{2}{9}\left[\frac{9}{2} - \frac{6.25}{2}\right] = \frac{2}{9}\left[\frac{2.75}{2}\right] = \frac{5.5}{18}$$

الأمل الرياضي :

نعرّف الأمل الرياضي بالنسبة للمتحول العشوائي المستمر x وكثافته f(x)، بالعلاقة:

$$\mu = E(x) = \int_{D} x f(x).dx$$

حيث D تشير إلى حدود التكامل المعرّف

$$\mathrm{var}(x) = \delta_x^2 = E(x^2) - \mu^2$$

$$E(x^2) = \int_{D} x f(x^2).dx$$

$$\delta_x = \sqrt{\delta_x^2}$$

تمرين (8) :

أوجد قيمة التوقع E(x) والانحراف المعياري بالنسبة لتابع الكثافة الاحتمالي الآتي:

$$f(x) = \begin{cases} 2/9\, x & \forall : 0 \le x \le 3 \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

الحل :

ـ الأمل الرياضي :

$$E(x) = \int_D x f(x).dx = \tfrac{2}{9} \int_0^3 x^2 .dx$$

$$= \tfrac{2}{9}\left[x^3/3\right]_0^3 = \tfrac{2}{27}[27] = \tfrac{54}{27} = 2$$

ـ التباين :

$$\mathrm{var}(x) = \int_D x^2 f(x).dx - \mu^2$$

$$= \tfrac{2}{9}\int_0^3 x^3 .dx - \mu = \tfrac{2}{9}\left[x^4/4\right]_0^3 - \mu^2$$

$$= \tfrac{2}{9}\left[81/4\right] - (2)^2 = 167/36 - 4 = 0.5$$

ـ الانحراف المعياري :

$$\delta = \sqrt{\mathrm{var}(x)} = \sqrt{0.5} = 0.7$$

تمرين :

لدينا التابع المعرف كما يلي :

$$F(x) = \begin{cases} X & \forall : 0 \le x \le 1 \\ 2 - x & \forall : 1 \le x \le 2 \\ 0 & \text{خلاف ذلك} \end{cases}$$

والمطلوب :

1 ــ برهن أن هذا التابع هو تابع كثافة احتمالي .

2 ــ احسب كلاً من التوقع الرياضي، والانحراف المعياري .

3 ــ احسب $P(x \le 1)$ و $P(1/2 \le x \le 3/2)$

الحل :

ــ نلاحظ أن المتراجحة $f(x) \ge 0$ محققة من أجل جميع قيم x.

ــ الشرط الثاني :

$$f(x) = \int_0^1 x.dx + \int_0^1 (2-x).dx$$

$$= \left[x^2/2\right]_0^1 + \left[(2x - 2/2) - (2x - 2/2)\right]_1^2$$

$$= \left[1/2\right] + \left[(4 - 4/2) - (2 - 1/2)\right] = 1$$

ومنه الشرط الثاني لتابع الكثافة محقق، لذلك نستنتج أن f(x) يمثل تابع كثافة احتمالي.

ـ التوقع الرياضي :

$$E(x) = \int_D x.f(x).dx = \int_0^1 x^2.dx + \int_1^2 (2x - x^2).dx$$

$$= \left[\frac{x}{3}\right]_0^1 + \left[\frac{2x^2}{2} - \frac{x^3}{3}\right]_1^2 = \left[\frac{1}{3}\right] + \left[(4 - \frac{8}{3}) - (1 - \frac{1}{3})\right]$$

$$= \frac{1}{3} + \left[\frac{4}{3} - \frac{2}{3}\right] = 1$$

ـ التباين :

$$E(x^2) = \int_D x^2 f(x).dx = \int_0^1 x^3.dx + \int_1^2 (2x^2 - x^3).dx$$

$$= \frac{1}{4}\left[x^4\right]_0^1 + \left[\frac{2}{3}x - \frac{x^4}{4}\right]_1^2 = \frac{19}{6}$$

$$\text{var}(x) = E(x^2) - \mu^2 = \frac{19}{6} - 1 = \frac{13}{6}$$

الانحراف المعياري :

$$\delta = \sqrt{\text{var}(x)} = \sqrt{\frac{13}{6}}$$

$$P(x \le 1) = \int_0^1 x.dx = \frac{1}{2}$$

$$P(\frac{1}{2} \le x \le \frac{3}{2}) = \int_{\frac{1}{2}}^1 x.dx + \int_1^{\frac{3}{2}} (2 - x).dx$$

$$\frac{1}{2}[x]_{\frac{1}{2}}^1 + [2x - \frac{x^3}{3}]_1^{\frac{3}{2}} = \frac{3}{4}$$

2 ـ 7 ـ التوزيعات الاحتمالية المتقطعة الخاصة :

سبق لنا دراسة المتغيرات العشوائية المتقطعة، وتابع كثافته الإحتمالية، وتابع التوزيع التراكمي، والأمل الرياضي، والتباين، لذا فإننا سندرس هنا بعض التوزيعات المتقطعة الخاصة لما لها من أهمية في الحياة العملية، بتفسير كثير من الظواهر .

2-7-1- التوزيع المنتظم : يعرّف التوزيع المنتظم، بأنه التوزيع الذي يكون فيه لجميع قيم المتغير العشوائي نفس الاحتمال .

لنفرض أن المتغير العشوائي X يأخذ القيم $X = \{ x_1, x_2, \ldots . x_n \}$ باحتمال متساوٍ ، فإن تابع الكثافة الاحتمالية يعطى بالعلاقة :

$$f(x) = \begin{cases} 1/n & \forall : x = x_1, x_2 \ldots\ldots x_n \\ 0 & \text{فيما عدا ذلك} \end{cases}$$

أي أن : $P(X = x_i) = P_i = 1/n$

حيث تمثل n عدد القيم الممكنة للمتحول العشوائي X.

خواص التوزيع المنتظم :

1 ـ $P_i > 0$

2 ـ $\sum P_i = \sum 1/n = n.1/n = 1$

3 ـ $F(x) = P(x \le x_i) = \sum_{i=1}^{s} P_i = s/n$

4 ـ الأمل الرياضي $E(x) = \sum P_i x_i = 1/n \sum x_i = \bar{x}$

5 ـ التباين $\mathrm{var}(x) = \sum P_i (x-\mu)^2 = 1/n \sum (x-\bar{x})^2$

6 ـ الانحراف المعياري $\delta_x = \sqrt{\mathrm{var}(x)}$

2-7-2- محاولة بيرنولي :

ثمة تجارب كثيرة يكون لها نتيجتان فقط، تحقق الحالة المرغوبة أو عدم تحققها. فإما أن يكون الإنتاج سليماً أو معطوباً، فنتيجة رمي قطعة النقد في الهواء نحصل على صورة أو كتابة، فلو رمزنا لتحقق الحالة المرغوبة بـــ P فيكون عدم تحققها مساوياً إلى $q = 1-p$.

فكل تجربة لها نتيجتان فقط، إما q أو P وتسمى مثل هذه التجارب بمحاولة بيرنولي، وإن فراغ العينة لتجربة بيرنولي لها نتيجتان فقط ، ونرمز لحالة تحقق الحادثة المرغوبة بـــ $P(x) = 1$ ولحالة عدم تحقق الرغبة بـــ $P(x) = 0$، أي أن $X = \{ 0, 1 \}$.

وعليه يجب أن تحقق محاولة بيرنولي الشروط التالية :

1 ــ كل تجربة لها نتيجتان فقط .

2 ــ كل التجارب مستقلة، أي أن نتيجة أي تجربة ليس لها أي تأثير على نتائج التجارب الأخرى .

3 ــ احتمال النجاح والفشل يظل ثابتاً في جميع المحاولات .

إن تابع الكثافة الاحتمالية يعطى بالعلاقة :

$$f(x)=\begin{cases} P^{x}(1-P)^{1-x} & x=\{0,1\} \\ 0 & \text{خلاف ذلك} \end{cases}$$

أو يعطي تابع الكثافة على شكل جدول :

Xi	0	1
Pi	(1-P)	P

تمرين (10) :

لنفرض أننا رمينا حجر النرد وكانت الحادثة المرغوبة هي ظهور الوجه {x =6}، أوجد تابع الكثافة لمحاولة بيرنولي :

$$f(x) = (1/6)^1 . (1- 1/6)^{1-1} = 1/6$$

أو :

$$f(x) = 0\ (1- 1/6) + 1.\ (1/6) = 1/6$$

خواص محاولة بيرنولي :

1 ـ $P + q = 1$

2 ـ $0 \le P \le 1$

$$E\ (x) = \sum x_i P_i = 0\ (\ 1\text{-}P) + 1.P = P$$

3 ـ الأمل الرياضي $E\ (x) = P$

4 ـ التباين : $\mathrm{var}(x) = E(x^2) - \mu^2 = \sum P_i x_i^2 - \mu^2$

$$= 0.\ (1\text{-}P) + 1.p - P^2 = P - P^2 = P\ (\ 1 - P\ ) = P\ .q$$

5 ـ الانحراف المعياري :

$$\delta = \sqrt{P(1-P)} = \sqrt{P.q}$$

2-7-3- توزيع ثنائي الحدين The Binomial distribution :

نفرض أن تجربتنا العشوائية تتكون من تكرار محاولة بيرنولي عدداً من المرات، (nمرة) وتعرف التجارب من هذا النوع بأنها تجارب ثنائية. وحينما ندرس تجربة ثنائية ذات (n) مرة، فإن كل خصائص هذه التجربة ترتبط بوسيطي هذه التجربة الثنائية (P,n) وهما احتمال النجاح، وعدد مرات تكرار التجربة، ويجب أن تحقق التجربة الإحصائية الشروط التالية:

1 ـ تتألف التجربة من عدد (n) من التكرارات المتماثلة .

2 ـ نتيجة كل محاولة للتجربة أحد ناتجين: " نجاح " أو " فشل " .

3 ــ إن احتمال النجاح (P) أو الفشل (q) يبقى ثابتاً في جميع المحاولات المكررة للتجربة.

4 ــ التكرارات مستقلة عن بعضها البعض، أي ليس هناك من تأثير لنتيجة أي محاولة على نتائج المحاولات الأخرى .

5 ــ نهتم فقط بعدد النجاحات (K) أو الحالات المرغوبة ، خلال جميع المحاولات (n).

تمرين (11) :

مجتمع يحتوي (20) شخصاً، (12) منهم مدخنين و (8) غير مدخنين، أخذنا بشكل عشوائي (5) أشخاص، فما احتمال أن يكون الثلاثة الأوائل مدخنين؟ .

الحل :

A = { غير مدخن، غير مدخن، مدخن، مدخن، مدخن }

بما أن الحوادث مستقلة عن بعضها البعض فإن احتمال تحقق الحادثة :

$$P(A) = P \cdot P \cdot P \cdot q \cdot q = P^3 q^2 .$$

$$= (12/20)^3 \cdot (8/20)^2 .$$

وبصورة عامة نجد، أنه عندما يكون لدينا ظاهرة ما تأخذ إحدى نتيجتين فقط {نجاح، فشل} وقمنا بـــ (n) محاولة فإن حصولنا على :

n وفقط k مرة نجاح يقتضي بالضرورة حصولنا على (n-k) حالة فشل، وإن ذلك يتم بعدد من الترتيبات المساوية إلى C_n^k ترتيباً ، وإن احتمال كل ترتيب مساوٍ $P^k q^{n-k}$ وإن تابع التوزيع الثنائي يعطي بالعلاقة :

$$f(x)=\begin{cases} C_k^n P^k q^{n-k} & k=0,1,2,.....n \\ 0 \end{cases}$$

وإن $0 \le k \le n$

وهو ما يعرف بالتوزيع الثنائي أو بتوزيع ذي الحدين ويمكن عرض تابع الكثافة على الصيغة الجدولية التالية :

k	0	1	2	k	n
P	q^n	$C_1^n P^1 q^{n-1}$	$C_2^n P^2 q^{n-2}$	$C_k^n P^k q^{n-k}$	P^n

ـ التوقع (الأمل الرياضي) $E(x) = nP$

ـ التباين

$Var(x) = np$

الانحراف المعياري :

$$\delta = \sqrt{var(x)} = \sqrt{nPq}$$

تابع التوزيع التراكمي :

$$F(x) = P(x = k \le m) = \sum_{j=0}^{m} C_k^n P^k q^{n-k}$$

تمرين (12) :

إذا كان احتمال إصابة الهدف لشخص ما $P = 1/5$، وأتيحت له فرصة الرماية $n=(5)$ مرات.

والمطلوب :

1 ـ أثبت أن هذا التوزيع الإحتمالي هو تابع ثنائي الحدين ؟

الحل :

لكي يكون هذا التوزيع ثنائي الحدين يجب أن يحقق العلاقة :

$$\sum_{k=0}^{n} C_k^n \; P^k \; q^{n-k} = 1$$

$$\sum_{k=0}^{5} C_k^n \left(\tfrac{1}{5}\right)^k \left(\tfrac{4}{5}\right)^{n-k}$$

$$C_0^5\left(\tfrac{1}{5}\right)^0\left(\tfrac{4}{5}\right)^5 + C_1^5\left(\tfrac{1}{5}\right)^1\left(\tfrac{4}{5}\right)^4 + C_2^5\left(\tfrac{1}{5}\right)^2\left(\tfrac{4}{5}\right)^3 + C_3^5\left(\tfrac{1}{5}\right)^3\left(\tfrac{4}{5}\right)^2 + C_4^5\left(\tfrac{1}{5}\right)^4\left(\tfrac{4}{5}\right)^1 + C_5^5\left(\tfrac{1}{5}\right)^5\left(\tfrac{4}{5}\right)^0 =$$

$$= 0.32768 + 0.4096 + 0.2048 + 0.0512 + 0.064 + 0.00032 = 1$$

والشرط الأول محقق حيث :

$$\sum_{k=0}^{n} C_k^n P^k q^{n-1} = 1$$

والشرط الأول محقق حيث : $0 \le C_k^n P^k q^{n-k} \le 1$

2 ــ احسب احتمال اصابة الهدف مرتين : $P(x = 2)$

$$P(x=2) = C_2^5\left(\tfrac{1}{5}\right)^2\left(\tfrac{4}{5}\right)^3 = 0.2048$$

3 ــ احسب احتمال الإصابة لأكثر أو يساوي مرتين $F(x) = P(x \ge 2)$

$$P(x \ge 2) = P(x=2) + P(x=3) + P(x=4) + P(x=5)$$

$$= 1 - [P(x=0) + P(x=1)]$$

حيث أن :

$$P(x=0) = C_0^5\left(\tfrac{1}{5}\right)^0\left(\tfrac{4}{5}\right)^5 = 0.32768$$

$$P(x=1) = C_1^5\left(\tfrac{1}{5}\right)^1\left(\tfrac{4}{5}\right)^4 = 0.4096$$

وعليه يكون الاحتمال المطلوب :

$$= 1 - [0.32768 + 0.4096] = 0.26272$$

4 ــ احسب التوقع الرياضي والتباين والانحراف المعياري للمتحول العشوائي x

$n = 5 \quad , \; P = 1/5 \quad , \quad q = 4/5$

$E(x) = nP = 5\left(1/5\right) = 1$

$\mathrm{var}(x) = nPq = 5\left(1/5\right)\left(4/5\right) = 4/5$

$\delta = \sqrt{4/5} = 0.894$

تمرين (13):

لوحظ لفترة طويلة من الزمن أن صياداً يصيب الهدف باحتمال قدره (0.8) فإذا أطلق أربع طلقات على الهدف، فما هو احتمال :

1 ــ إصابة الهدف مرتين تماماً $P(x = 2)$ ؟

الحل :

$P(x = 2) = C_k^n P^k q^{n-k}$

$= C_2^4 (0.8)^2 (0.2)^2 = 0.1536$

2 ــ إصابة الهدف مرتين على الأقل $P(x \geq 2)$ ؟

$P(x \geq 2) = P(x = 2) + P(x = 3) + P(x = 4)$

$= 1 - [\, P(x = 0) + P(x = 1) \,]$

$= 1 - [\, 0.0016 + 0.0256 \,] = 0.9728$

3 ــ إصابة الهدف أربع مرات تماماً $P(x = 4)$ ؟

تمرين (14) :

ماهو احتمال وجود (x) من الذكور في أسرة مؤلفة من (5) أطفال؟ ثم أحسب الاحتمالات الموافقة لقيم (x) التالية :

$$X = \{ 0 , 1, 2, 3 , 4, 5 \}$$

الحل :

$$P(x = x) = C_x^5 \left(\tfrac{1}{2}\right)^x \left(\tfrac{1}{2}\right)^{5-x}$$

$$P(x = 0) = C_x^5 \left(\tfrac{1}{2}\right)^0 \left(\tfrac{1}{2}\right)^5 = \tfrac{1}{32}$$

ويترك للطالب البقية .

تمرين (15) :

إذا كان احتمال نجاح الطالب في كلية العلوم السياسية في مادة الإحصاء **(٠,٣)** فماهو :

1 ــ احتمال نجاح طالب واحد من مجموع أربعة طلاب ؟

$$P(x=1) = C_x^5 q^3 P^1 .$$

$$= C_1^4 (0.7)^3 . (0.3)^1 .$$

2 ــ احتمال نجاح طالب واحد على الأقل من مجموعة أربعة طلاب ؟

$$P(x \geq 1) = C_1^4 .(0.3)^1 (0.7)^3 + C_2^4 .(0.3)^2 (0.7)^2 + C_3^4 (0.3)^3 (0.7)^1 + C_4^4 (0.3)^4 (0.7)^0$$

3 ــ احتمال نجاح ثلاث طلاب على الأكثر ؟

$$P(x \geq 3) = C_0^4 .(0.3)^0 (0.7)^4 + C_1^4 .(0.3)^1 (0.7)^3 + C_2^4 (0.3)^2 (0.7)^2 + C_3^4 (0.3)^3 (0.7)^1$$

تمرين (16) :

إذا كان لدينا أربعة مرشحين للانتخابات، وكان احتمال فوز كل واحد منهم يساوي (1/2)، فما هو احتمال أن يكون :

1 ــ جميعهم قد فازوا بالانتخابات؟

$$P(x=4)=C_4^4\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0$$

2 ــ واحد منهم على الأقل فاز بالانتخابات

$$P(x\geq 1)=C_1^4.(1/2)^1(1/2)^3+C_2^4.(1/2)^2(1/2)^2+C_3^4(1/3)^3(1/2)^1+C_4^4(1/2)^4(1/2)^0$$

3 ــ عدد الفائزين يماثل عدد غير الفائزين بالانتخابات .

(يترك الحل للطالب) .

2-7-4- توزيع بواسون Possion is distribution :

يصف توزيع بواسون المتغيرات العشوائية المتقطعة النادرة الحدوث ويعطينا عدد النجاحات في فترات زمنية معينة أو في منطقة محددة، والفترة الزمنية يمكن أن تكون ثانية أو دقيقة أو يوماً والمنطقة المحددة يمكن أن تكون وحدة قياس للمساحة أو الحجم، فعدد المرضى الذين يراجعون غرفة الإسعاف خلال ساعة، وعدد حوادث السيارات خلال أسبوع في منطقة محددة، وعدد المكالمات الهاتفية التي تصل إلى مكتب الاتصال كل عشرة دقائق، وعدد الكريات الحمراء في عيِّنة من الدم، كلها أمثلة على تجارب بواسون .

ويجب أن تحقق التجارب في توزيع بواسون الشروط التالية :

1 ــ عدد مرات النجاح (x) قليل خلال عدد كبير من محاولات بيرنولي (n).

2 ــ احتمال النجاح (P) صغير جداً من خلال عدد كبير من محاولات بيرنولي .

3 ــ في عدد فترات زمنية منفصلة عن بعضها البعض فإن حدوث النجاحات أي فترة مستقل عن حدوث النجاحات في أي فترة أخرى .

ويعطى التوزيع الإحتمالي لبواسون، بالعلاقة التالية :

$$P(X=K)=\frac{e^{-\lambda}.\lambda^{k}}{K!}$$

حيث λ : هي معدل عدد النجاحات في الفترة الزمنية المعينة، أو المنطقة المحددة .. ،

$e = 2.718$

$x = k = 0, 1, 2, \ldots$

خصائص توزيع بواسون :

$1 - \lambda > 0$

$2- \sum_{k=0}^{\infty} P_k = 1$

$3- E(x) = var(x)\ \lambda = nP$

تمرين (17) :

إذا كان متوسط عدد المكالمات الهاتفية في أحد المراكز في الدقيقة الواحدة (2) مكالمة.

المطلوب:

1 ــ ما احتمال أن يتلقى المركز (4) مكالمات هاتفية خلال دقيقتين ؟

الحل :

$\lambda = 4$ متوسط المكالمات خلال دقيقتين .

$K=4$ عدد المرات أو المكالمات .

$$P(K=4)=e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$=e^{-4}.\frac{4^4}{4!}$$

$$=\frac{1}{e^4}.\frac{4^4}{4\times3\times2}=0.194$$

2 ــ ما احتمال أن يتلقى المركز (4) أو (5) مكالمات خلال (3) دقائق ؟

$\lambda = 6$

- $K = 4$

$$P(K=4)=e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$=e^{-6}.\frac{6^4}{4!}=0.049$$

- K=5

$$P(K=5)=e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$=e^{-6}.\frac{\lambda^4}{K!}$$

$$=e^{-6}.\frac{6^5}{5!}=0.58$$

- $P(K=4)\cup P(K=5)=0.049+0.058=0.107$

تمرين (18):

لنفرض أنه في أحد المستشفيات الصغيرة يوجد غرفة للإسعاف، وأن الطاقة الاستيعابية لهذه الغرفة (5) مرضى في الساعة الواحدة **والمطلوب :**

1 ــ ما إحتمال عدم قدوم أي مريض إلى هذه الغرفة ؟.

$K = 0 \qquad , \lambda = 5$

الحل :

$$P(K=0) = e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$= e^{-5}.\frac{5^0}{0!}$$

2 ــ ما احتمال قدوم مريضين إلى هذه الغرفة ؟

$K = 2 \qquad , = 5\ \lambda$

$$P(K=0) = e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$= e^{-5}.\frac{5^2}{2!} = 0.00672.\frac{25}{2}$$

$$= 0.984$$

3 ــ ما احتمال قدوم مريضين أو ثلاثة مرضى إلى هذه الغرفة ؟.

$K = 3$

$$P(K=3) = e^{-\lambda}.\frac{\lambda^k}{k!}$$

$$= e^{-5}.\frac{5^3}{3!} = 0.00672.\frac{125}{6}$$

$$= 0.13999$$

$$P(K=3) \cup P(K=2) = 0.084 + 0.14042 = 0.22399$$

4 ــ ما احتمال قدوم مريضين على الأقل إلى هذه الغرفة ؟

$$P(K \geq 2) = 1 - \left[e^{-5}.\frac{5^0}{0!} + e^{-5}.\frac{5^1}{1!} \right] = 0.959$$

تمرين (19):

محطة للبنزين تقوم بخدمة الزبائن بمعدل سيارة واحدة كل (5) دقائق، أوجد احتمال أن تقوم هذه المحطة بخدمة (5) سيارات خلال ربع ساعة ؟

الحل :

λ : معدل عدد النجاحات في (5) دقائق = سيارة واحدة .

λ : معدل عدد النجاحات في ربع ساعة = ثلاث سيارات

$K = 5$

$$P(K=5) = e^{-3} \cdot \frac{3^5}{5!} =$$

$$P(K=5) = (2.718)^{-3} \frac{3^5}{5!}$$

$$P(K=5) = \frac{1}{(2.718)^3} \cdot \frac{243}{5 \times 4 \times 3 \times 2} = 0.09$$

2 ــ 8 ــ التوزيع الطبيعي : Normal Disbritoution :

يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية المستمرة وهو عبارة عن مجموعة المنحنيات البيانية أو التوزيعات التكرارية، التي تعد في غاية الأهمية في مجال البحوث طالما أن العديد من الظواهر تتبع التوزيع الطبيعي في مجتمع الدراسة، ويمكن من خلال هذا التوزيع الذي يشبه الجرس المقلوب إيجاد النسبة المئوية للحالات المدروسة التي تقع بين نقطتين فعندما يكون المتحول العشوائي (x) مستمراً، فإن عدد القيم الممكنة يصبح لانهائياً، لذلك فإن توزيع الاحتمالات على قيم المتحولات يعطي بدلالة تابع الكثافة الاحتمالية لمتغير عشوائي مستمر وهي من الشكل :

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2}$$

ونلاحظ أن لهذا التوزيع وسيطين هما الوسط الحسابي (μ) والانحراف المعياري (δ) وإن مجال تحوله $-\infty < x < +\infty$

تمرين (20):

أنشئ نموذجاً رياضياً لتوزيع أطوال مجتمع الطلاب الجامعيين إذا علمت أن

$\mu = 168$ c.m $\qquad \sigma = 12$ cm

$$f(x)=\frac{1}{12\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-168}{12}\right)^2}$$

إننا نرمز عادة لهذا التوزيع $N(\mu , \delta^2)$، حيث تمثل (μ) الوسط الحسابي و(δ) الانحراف المعياري، وهو يشكل منحنياً متناظراً حول (μ) .

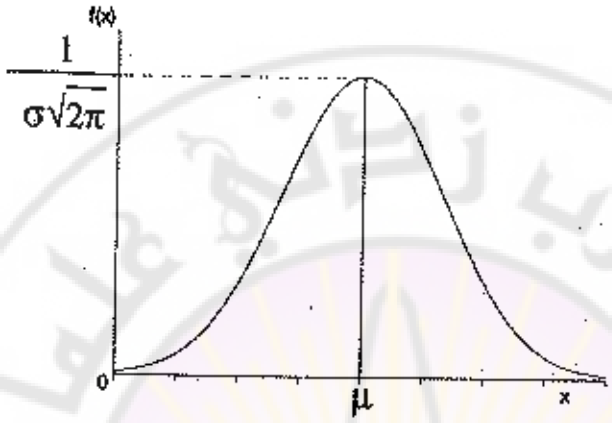
خواص التوزيع الطبيعي :

1 ــ إن المساحة التي يحددها تابع الكثافة f(x) على المحور (0x) تساوي الواحد الصحيح، أي:

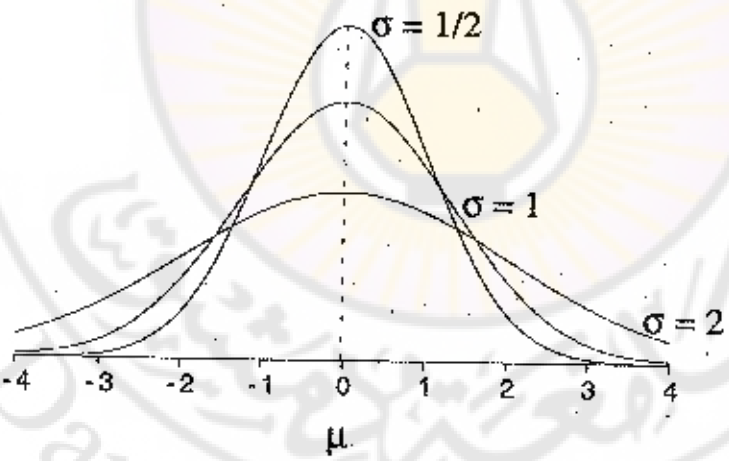
$$\int_{-\infty}^{+\infty} f(x).dx = 1$$

2 ــ إن منحني هذا التابع متناظر حول النقطة $x=\mu$ وهي النقطة التي يتمركز عندها التوزيع أو يبلغ نهايته العظمى، ثم ينتشر على جانبيها بصورة متناظرة وله نقطتي إنعطاف في الموضعين $x = \mu - \delta$ والنقطة $x = \mu + \delta$ ونلاحظ أنه من أجل قيم صغيرة لــ δ يكون انتشار المنحني على جانبي μ ضئيلاً، فيما يمتد إلى مسافات أبعد ويأخذ شكلاً أكثر انبساطاً كلما ازدات δ .

الشكل رقم (2 ــ 6) وإن المساحة التي يحددها المنحني مساوية الواحد في جميع الأحوال .



الشكل رقم (2-5).



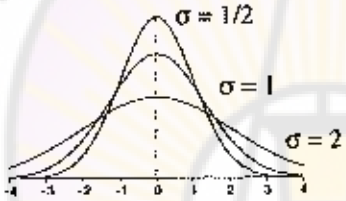
الشكل رقم (2-6).

3 ــ الأمل الرياضي :

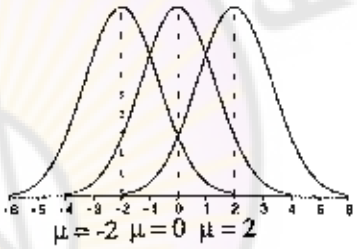
$$E(x) = \int_{-\infty}^{+\infty} f(x).dx = 1$$

4 ــ إن تابع الكثافة لمنحني التوزيع الطبيعي يحوي كما رأينا وسيطين (δ ، μ) ومن أجل قيم مختلفة لهما، نحصل على منحنيات متعددة لها جميعاً شكل الجرس المقلوب ويمثل الشكلان رقم (2 ــ 7) و (2 ــ 8) مجموعتين من المنحنيات الطبيعية .

تختلف منحنيات المجموعة الأولى بعضها عن بعض بمتوسطها كما في الشكل رقم (2 ــ 7) فقط بينما تختلف المجموعة الثانية بانحرافاتها المعيارية الشكل رقم (2 ــ 8).

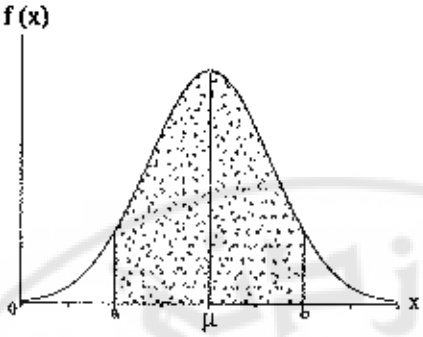


الشكل (2 ــ 7)



الشكل (2 ــ 8)

وإن احتمال أن يقع x في المجال [a , b] هو عبارة عن المساحة المحدودة بهذا المنحني والمحور (0x) والمستقيمين x=a و x = b حيث a < b الشكل رقم (2 ــ 9) .



الشكل رقم (2 ــ 9).

ولحساب الاحتمال :

$$P(a \le x \le b) = \int_a^b \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2} .dx$$

إذا كانت الظاهرة تخضع لمنحني توزيع طبيعي في المجتمع. فإن المساحات تتوزع وفق النسب التالية :

$P(\mu . \delta \le x \le \mu + 1. \delta) = 68.27\%$

$P(\mu - 1.96\,\delta \le x \le \mu + 1.69\,\delta) = 95\%$

$P(\mu = 3\delta \le x \le \mu + 3\delta) = 99.73\%$

تمرين (21) :

إذا كانت علامات الطلاب الناجحين بمقرر مبادئ الإحصاء في إحدى الكليات في الجامعة تتوزع طبيعياً حول متوسط $\mu = 70$ درجة وانحراف معياري $\delta = 10$ درجات،

والمطلوب :

احسب الاحتمالات التالية :

$P(x \ge 90)$, $P(80 \le x \le 90)$, $P(60 \le x \le 80)$, $P(x \le 55)$

الحل :

ولحساب هذه الاحتمالات نقوم بتمثيلها على المنحني الطبيعي المعياري كما في الشكل

(10 – 2)

الشكل رقم (2-10) .

$$P(x \geq 90) = \int_{90}^{+\infty} \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)^2} .dx$$

$$P(80 \leq x \leq 90) = \int_{80}^{90} \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2} .dx$$

$$P(60 \leq x \leq 80) = \int_{60}^{80} \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2} .dx$$

$$P(x \leq 55) = \int_{-\infty}^{55} e \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-70}{10}\right)^2} .dx$$

وفي كل مرة نحتاج فيها إلى حساب المساحة أو الاحتمال لابد من إجراء هذه التكاملات الشاقة، والسبب في ذلك يعود إلى اختلاف الأوساط الحسابية للمنحنيات وانحرافاتها المعيارية، من هنا جاءت فكرة البحث عن منحني طبيعي له شكل واحد مهما اختلفت قيم المتحول العشوائي x، وهو ما يعرف بالتوزيع الطبيعي المعياري .

وقبل البحث في مفهوم هذا التوزيع ، لابد من التعرض لمفهوم الدرجات المعيارية، التي هي عبارة عن وحدة لقياس الانحراف المعياري عن الوسط الحسابي، وتستخدم لمقارنة قيمتين مأخوذتين من مجموعتين إحصائيتين مختلفتين وتعطى الدرجة المعيارية بالعلاقة الآتية:

$$Z = \frac{|x - \mu|}{\delta} \quad \text{أو} \quad Z = \frac{|x - \bar{x}|}{S}$$

تمرين (22):

كانت درجتا طالبين في أحد المقررات في شعبتين مختلفتين على النحو التالي:

المقاييس	الطالب الأول (A)	الطالب الثاني (B)
درجة الطالب	86	64
متوسط درجات الطلاب	75	58
الانحراف المعياري	10	4

والمطلوب : أيهما كان تحصيله أفضل بالنسبة لمستوى الطلاب؟.

الحل :

$$Z = \frac{|x - \bar{x}|}{S}$$

$$Z_{(A)} = \frac{|x - \bar{x}|}{S} = \frac{|86 - 75|}{10} = 1.1$$

$$Z_{(B)} = \frac{|x - \bar{x}|}{S} = \frac{|64 - 58|}{4} = 1.5$$

من خلال المقارنة نجد أن : $Z_{(B)} > Z_{(A)}$ أي أن تحصيل الطالب (B) أفضل من تحصيل الطالب (A) على الرغم أنه قبل التحويل إلى درجات معيارية كانت درجات الطالب (A) أفضل من درجات الطالب (B) .

2 ـ 9 ـ التوزيع الطبيعي المعياري :

Standard Normal Distribution

يعرف التوزيع الطبيعي المعياري بأنه التوزيع الطبيعي الذي وسطه الحسابي يساوي الصفر وتباينه يساوي الواحد الصحيح أي أن المتغير العشوائي (Z) يخضع للتوزيع الطبيعي ذا الوسط الحسابي $\mu=0$ و $\delta^2=1$ ونعبر عنه $Z: N(0,1)$. إذا كان (x) يخضع للتوزيع الطبيعي وسطاؤه (μ) و(δ^2)، فإن المتغير (Z) الذي نحصل عليه من التحويل $Z = \left|\frac{x-\mu}{\delta}\right|$ يكون توزيعه توزيعاً طبيعياً معيارياً وكل قيمة من قيم (x) يقابلها قيمة من قيم (Z).

ـ خصائص التوزيع الطبيعي المعياري :

1 ـ المساحة التي يحددها التوزيع تحت المنحني تساوي الواحد الصحيح .

2 ـ منحني متناظر وله وسطه الحسابي ($\mu=0$) و ($\delta^2=1$)

3 ـ البيانات تتوزع بشكل منتظم حول الوسط الحسابي .

ولإيجاد تابع الكثافة لمنحني التوزيع الطبيعي المعياري $N \sim (0,1)$ نجري عملية التحويل لتابع الكثافة لمنحني التوزيع الطبيعي $N(\mu, \delta^2)$

$$Z^2\left(\frac{x-\mu}{\delta}\right)^2$$

والإنحراف المعياري $\delta = 1$ إذا يصبح تابع الكثافة لمنحني التوزيع الطبيعي من الشكل :

$$\varphi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

وهو ما يعرف بالتوزيع الطبيعي المعياري .

حيث أن :

$$Z = \frac{x-\mu}{\delta}$$

$$-\infty < Z < +\infty$$

إن هذا التابع متناظر حول وسطه الحسابي حيث ($\mu = 0$) كما أن انحرافه المعياري $\delta = 1$.

نسمي المتغير العشوائي الجديد (Z)، المتغير المعياري أو الدرجة المعيارية، ولحساب قيمة الاحتمال أو المساحة تطبق العلاقة :

$$P(a \le x \le b = \int_a^b \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\delta}\right)} dx$$

نحول المتغير العشوائي (x) إلى المتغير المعياري (Z) في تابع الكثافة الاحتمالية

$N \sim (\mu, \delta^2)$ ليصبح $N \sim (0,1)$.

ولإيجاد التكامل السابق، لابد أيضاً من تحويل حدود التكامل إلى الدرجات المعيارية أو المتغير المعياري .

عندما x=a يصبح الحد الأدنى للتكامل $Z = \frac{a-\mu}{\delta}$

عندما x=b يصبح الحد الأعلى للتكامل $Z = \frac{b-\mu}{\delta}$

وإن : $dz = \frac{dx}{\delta} \Rightarrow dx = dz.\delta$

وبالتعويض في علاقة منحنى التوزيع الطبيعي $N(\mu, \delta^2)$

$$P(a \le x \le b) = \int_{\frac{a-\mu}{\delta}}^{\frac{b-\mu}{\delta}} \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \delta.dz$$

وبما أن $\delta = 1$ تصبح العلاقة :

$$P(a \le x \le b) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\delta}}^{\frac{b-\mu}{\delta}} e^{-\frac{1}{2}z^2} \delta.dz$$

تمرين (23) :

احسب المساحة المحصورة $P(0 \le x \le 2)$

الحل :

$$P(0 \le x \le 2) = \frac{1}{2\pi} \int_0^2 e^{-\frac{1}{2}z^2}.dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^2 \left(1 - \frac{Z^2}{2} + \frac{Z^4}{4\times 2} + \frac{Z^6}{8\times 2\times 3} + \frac{Z^8}{16\times 2\times 3\times 4} + \frac{Z^{10}}{32\times 2\times 3\times 4\times 5}\right).dZ$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^2 \left(1 - \frac{Z^2}{2} + \frac{Z^4}{8} - \frac{Z^6}{48} + \frac{Z^8}{384} + \frac{Z^{10}}{3840}\right).dZ$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^2 \left(\frac{3840 - 1920Z^2 + 480Z^4 + 80Z^6 + 10Z^8 + Z^{10}}{3840}\right).dZ$$

$$= \frac{1}{3840\sqrt{2\pi}} \left[3840Z - 1920\frac{Z^3}{3} + 480\frac{Z^5}{5} + 80\frac{Z^7}{7} + 10\frac{Z^9}{9} + \frac{Z^{11}}{11}\right]_0^2$$

$$= \frac{1}{3840 \times 2.5}\left[3840 \times 2 - 1920\frac{8}{3} + 480\frac{32}{5} + 80\frac{128}{7} + 10\frac{512}{9} + \frac{2048}{11}\right]$$

$$= \frac{4551.85}{9600} = 0.474151$$

وكما لاحظنا من خلال التمرين (23) السابق: إن إجراء هذا التكامل شاق ويتطلب جهداً كبيراً، لذلك تم وضع جدول خاص لحساب هذه التكاملات ويدعى بجدول المساحات تحت منحني التوزيع الطبيعي المعياري $N\sim(0,1)$ وقد حسبت المساحات في هذا الجدول لنصف المنحني على اعتبار أن منحني التوزيع القياسي متناظر حول النقطة $\mu = 0$ وهي عبارة عن الوسط الحسابي وتقسم المنحني إلى قسمين متساويين ومتناظرين تماماً، وقد حسبت المساحات بدءاً من النقطة $Z = 0$ وبتزايد مقداره 0.01 ويقسم جدول المساحات، إلى سطر وعمود وتبدأ قيم (Z) في السطر بالقيمة $Z = 0$ وتزايد مقداره 0,01 حتى القيمة 0,09 وتبدأ قيم العمود من القيمة $Z = 0$ وتزايد مقداره 0.1 حتى القيمة $Z = 3$ حيث المساحة المحصورة $P(Z \le 0) = P(Z \ge 0) = 0.5$ وأن الهدف من الاستعراض السابق هو معرفة نسبة أو احتمال المفردات الواقعة في مجال محدد أو خارجه تحت المنحني الطبيعي من خلال التوزيع الطبيعي المعياري بطريقة رياضية تعتمد على تحويل القيم الطبيعية (x) قيم طبيعية معيارية (Z) وذلك لمعرفة مدى ابتعادها عن الوسط الحسابي μ أو $\bar{x}$ بالانحرافات المعياري وذلك من خلال جدول المساحات (Z) والموجود في نهاية الكتاب.

تمرين (24) :

احسب المساحة المحصورة بين القيمتين $(Z = 0 , Z = 1.96)$ أي $P(0 \le Z \le 1.96)$ باستخدام جدول المساحات .

الحل :

لإيجاد المساحة المطلوبة نبحث في عمود (Z) عن القيمة 1.9 وفي سطر (Z) عن القيمة 0.06 ونقطة تقاطع السطر والعمود هي عبارة عن المساحة المطلوبة .

$P(0 \le Z \le 1.96) = 0.475$

تمرين (25):

أحسب احتمال وقوع (Z) بين (0, 0.98) أي حساب الإحتمال $P(0 \le Z \le 0.98)$. لهذه الغاية نبحث في عمود (Z) عن القيمة (0,9) وفي السطر عن القيمة (0,08) ونقطة التقاطع و المساحة المطلوبة :

$P(0 \le Z \le 0.98) = 0.3365$

تمرين (26):

ماهي قيمة Z1 التي تحقق العلاقة :

$P(0 \le Z \le Z1) = 0.3849$

إذا نظرنا في الجدول نجد أن المساحة 0,3849 تقابل (1,2) درجة معيارية .

$|Z| = 1.2$.

تمرين (27) :

إن طول الشخص هو متحول عشوائي مستمر (x) يخضع للتوزيع الطبيعي بمتوسط x=170 وانحراف معياري $\delta = 5$

1 ــ ماهو احتمال أن يقع x في المجال $P(160 \le x \le 185)$

2 ــ ماهو احتمال أن يقع x في المجال $P(x \ge 170)$

3 ــ ماهو احتمال أن يقع x في المجال $P(x \le 160)$

الحل :

$$P(160 \le x \le 185) = P\left(\frac{160-170}{5} \le Z \le \frac{185-170}{5}\right)$$

$= P(-2 \le Z \le 3)$

$= P(0 \le Z \le 2) + P(0 \le Z \le 3)$

$= 0.4772 + 0.49865 = 0.97585$

2- $P(Z \ge 0) = P(0 \le Z \le 3) = 0.5$

3- $P(Z \le 2) = [P(Z \ge 0) - P(0 \le Z \le 2)$

$= 0.5 - 0.4772 = 0.0228$

تمرين (28) :

إذا كان (Z) متغيراً عشوائياً مستمراً، يتبع التوزيع الطبيعي المعياري N (0,1) والمطلوب : احسب الاحتمالات التالية :

1 ـ $P(0 \le Z \le 1.54)$

2 ـ $P(1 \le Z \le 2)$ 5 ـ $P(Z \ge 1.25)$

3 ـ $P(Z \le -1.2)$ 6 ـ $P(-1.8 \le Z \le 3)$

4 ـ $P(-1 \le Z \le 1.28)$ 7 ـ $P(-2 \le Z \le -1)$

الحل :

$1 - P(0 \le Z \le 1.54) =$

$P(Z \le 1.54) = 0.4382$

$$2 - P(1 \le Z \le 2) = P(Z \le 2) - P(Z \le 1)$$

$$= 0.4712 - 0.3413 = 0.1359$$

$$3 - P(Z \le -1.2)$$

نحن نعلم أن جدول المساحات لايعطي المساحة للقيم السالبة ونظراً لتماثل المنحني في كل من طرفيه الموجب والسالب فإن :

$$P(Z \le -1.2) = P(Z \ge 1.2)$$

$$= P(Z \ge 0) - P(Z \le 1.2)$$

$$= 0.5 - 0.3849 = 0.1151$$

ويمكن أن نقوم بحساب بقية الإحتمالات بنفس الطريقة، وتترك واجباً للطالب.

تحديد قيم Z باحتمال معلوم $(1-\alpha)$:

نرغب أحياناً في تحديد مجالات تحوي (Z) باحتمال معلوم، فإذا كان لدينا توزيعاً $N(0,1)$، وكنا نرغب في تعيين مجال متناظر حول $\mu = 0$ كالمجال $(-Z_1, Z_1)$ والذي يحتوي على 98% من قيم التوزيع مثلاً :

$$P(-Z_1 \le Z \le Z_1) = 0.98$$
$$P(0 \le Z \le Z_1) = 0.49$$

وبالرجوع إلى جدول المساحات نجد أن قيمة $Z_1 = 2.33$ أي أن المجال [-2.33, 2.33] يحتوي على (98%) من قيم التوزيع، وبالطريقة ذاتها نجد أن المجال [-1.96, 1.96] يحتوي على 95% من قيم التوزيع .

وبصورة عامة إذا رغبنا في تعيين مجال يحوي ($1 - \alpha$) من قيم التوزيع نطبق العلاقة:

$$P(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}) = 1-\alpha$$

حيث تمثل $Z_{\alpha/2}$ فاصلة النقطة (OZ) التي تقع على يمينها $\alpha / 2$ من قيم التوزيع ويمكن تعميم ذلك من أجل مجتمع طبيعي عادي $N(\mu, \delta^2)$.

فإذا رغبنا في تعيين مجالات متناظرة حول μ تحوي $(1-\alpha)$ من قيم التوزيع نبدل (Z) بقيمتها بدلالة (x) من العلاقة السابقة فتصبح :

$$P\left(-Z_{\alpha/2} \le \frac{x-\mu}{\delta} \le Z_{\alpha/2}\right) = 1-\alpha$$

أو :

$$P(\mu - \delta_{z\alpha/2} \le x \le \mu + \delta_{z\alpha/2}) = 1-\alpha$$

ويمكننا تعيين القيم المختلفة لـ $Z_{\alpha/2}$ حسب مستويات الثقة $(1-\alpha)$ المطلوبة والجدول التالي يعطينا بعض هذه القيم :

مستوى الدلالة α	31.75%	10%	5%	2%	1%	0.27%
الاحتمال $1-\alpha$	%68.27	90%	95%	98%	99%	99.73%
$Z_{\frac{\alpha}{2}}$	± 1	± 1.64	± 1.96	2,33	± 2.58	± 3

تمرين (29):

يتوزع المتغير العشوائي (x) توزعاً طبيعياً N (120,400)، والمطلوب تشكيل مجال متناظر حول (μ) بحيث يكون احتمال وقوع (x) فيه 94% ؟.

الحل :

إذا رمزنا بـــ (Z) للمتغير المعياري المرافق لـــ (x)، فإن المجال المطلوب حول (μ) يصبح مجالاً متناظراً حول (0) بالنسبة للمتغير (Z) .

$$P(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}) = 0.94$$

نبحث عن $Z_{\alpha/2}$ الموافقة للاحتمال 0,94 فنجد :

$$Z_{\alpha/2} = 1.88$$

لكن :

$$Z = \frac{x-120}{20}$$

$$P\left(-1.88 \le \frac{x-120}{20} \le 1.88\right) = 0.94$$

$$1.88) = 0.94 \times (120 - 20 * 1.88 \le x \le 120 + 20$$

$$(82.4 \le x \le 157.6) = 0.94$$

والمجال المطلوب هو : (157,6 ، 82,4)

تمارين غير محلولة

تمرين (1) :

وجد أحد الباحثين الاجتماعيين أن مشاهدي التلفزيون يفضلون القناة A بنسبة (75%)، بينما (25%) يفضلون القناة B ، سئل أربعة مشاهدين .

والمطلوب :

1 ـــ احسب احتمال أن نجد بينهم اثنين يفضلون مشاهدة القناة (A) ؟

2 ـــ احسب احتمال أن نجد بينهم واحد على الأقل يفضل مشاهدة القناة (A) ؟

3 ـــ أحسب التوقع والتباين والانحراف المعياري للمشاهدين الذين يفضلون مشاهدة القناة (A)؟

تمرين (2) :

إذا كان من المعلوم ، أن الأطفال يفضلون برنامجي الأطفال B, A بالتساوي، سئل ثمانية أطفال، فأجاب ستة منهم أنهم يفضلون البرنامج A ،

والمطلوب :

1 ـــ ماهو احتمال هذه النتيجة .

2 ـــ ماهو التوقع والانحراف المعياري .

تمرين (3) :

أوجد الأمل أو التوقع الرياضي للتوزيع الاحتمالي التالي :

X	-2	-1	0	1	3
F(x)	0.1	0.2	0.1	0.35	0.25

تمرين (4) :

إذا كانت نسبة المؤمّنين على الحريق في منطقة (0,03)، أخذت عينة عشوائية مكونة من (5) أشخاص. والمطلوب : ما احتمال أن يكون (3) منهم فقط مؤمنيّن ضد الحريق ؟

تمرين (5) :

من خلال دراسة أجريت على سجلات الطلاب في كلية العلوم السياسية تبين أن نسبة النجاح (50%)، قام أحد الطلاب بامتحان (12) مقرر والمطلوب :

1 ــ ماهو احتمال نجاح الطالب بـ (10) مقررات على الأكثر؟

2 ــ ماهو احتمال نجاح الطالب بـ (5) مقررات على الأقل ؟

3 ــ ماهو احتمال نجاح الطالب بجميع المقررات .

تمرين (6) :

إذا كان أطوال طلاب الكلية يتبع توزيعاً طبيعياً $N\sim(168 , 36)$، وسحبنا عينة عشوائية من الطلاب بحجم $n = 500$ طالباً .

والمطلوب :

1 ــ احسب عد د الطلاب الذين تزيد أطوالهم عن 184 سم.

2 ــ احسب عدد الطلاب الذين يقل أطوالهم عن 156 سم.

3 ــ احسب عدد الطلاب الذين تزيد أطوالهم عن 165 وتقل عن 174 سم.

تمرين (7) :

إذا علمت أن فترة الحمل عند المرأة تخضع لتوزيع طبيعي $\mu = 266$ يوماً وانحراف معياري (12) يوماً .

والمطلوب :

احسب الاحتمالات التالية :

1 ــ أن لاتزيد فترة الحمل عن 270 يوماً ولاتقل عن 260 يوماً.

2 ــ أن تزيد الفترة عن 270 يوماً .

3 ــ أن تقل الفترة عن 260 يوماً .

تمرين (8) :

كان الإعتقاد السائد لدى المسؤولين عن الصناعات البترولية أن متوسط الأجر الاسبوعي للعامل 3100 وحدة نقدية بانحراف معياري 250 وحدة، وإذا علمت أن الأجور كانت تتبع توزيعاً طبيعياً في هذه الصناعة، وسحبنا بصورة عشوائية عاملاً من عمال هذه الصناعة، فما هو احتمال أن يكون أجره :

1 ــ $P(3600 \leq x \leq 3350)$

2 ــ $P(3350 \leq x \leq 3850)$

3 ــ $P\, x \leq 2455)$

على اعتبار أن الأجر هو متغير عشوائي مستمر ونرمز له بـــ (x) .

تمرين (9) :

تبين من الدراسات الاحصائية أن احتمال إصابة الأطفال بالشلل هو (0,016) فإذا افترضنا أن عدد الأطفال حديثي الولادة (250) خلال فترة زمنية محددة، المطلوب :

1 ــ ما احتمال أن يوجد (5) أطفال من بين هذا العدد مصابين بهذا المرض؟

2 ــ ما احتمال أن يوجد أكثر من ثلاثة أطفال مصابين بهذا المرض من بين (5) أطفال.

تمرين (10) :

ليكن لدينا x متغيراً عشوائياً متقطعاً له جدول التوزيع الاحتمالي التالي :

X	0	1	2	3
P(X=x)	0.1	0.35	0.4	0.15

والمطلوب :

ـــ احسب التوقع الرياضي للمتغير العشوائي x ، ثم احسب التباين والانحراف المعياري.

الفصل الثالث

توزيعات المعاينة

Sampling Distribution's

- مقدمة
- أخطاء المعاينة : أخطاء التحيز – الأخطاء الاحتمالية.
- المجتمعات المحدودة والمجتمعات اللامحدودة .
- توزيع المعاينة للأوساط الحسابية .
- توزيع المعاينة للنسب المئوية .
- توزيع المعاينة للفرق بين متوسطي مجتمعين .
- تمارين غير محلولة.

3 ـ 1 ـ مقدمة :

لنفرض أنه لدينا مجتمعاً N من المفردات، يتبع توزيعاً احتمالياً معيناً (سواء كان هذا المجتمع كبيراً أو محدوداً)، وإننا بصدد سحب عينة حجمها n من هذا المجتمع، لحساب بعض المقاييس الإحصائية، يسمى كل من هذه المقاييس إذا كان محسوباً من العينة بـــ (التابع إحصائي) وسواءً أكان وسطاً حسابياً أو وسيطاً أو انحرافاً معيارياً، ويسمى نفس المقياس بمعلمة أو الثابت الإحصائي فيما إذا كان محسوباً بدلالة المجتمع الإحصائي .

والآن نفرض أننا سحبنا عينة حجمها n_1 من هذا المجتمع، وحسبنا من هذه العينة مقياساً معيناً، وليكن وسطاً حسابياً أو نسبة مئوية أو انحرافاً معيارياً، ثم سحبنا عينة ثانية من نفس الحجم ومن نفس المجتمع n_2 وحسبنا منها نفس المقياس وعينة ثالثة ورابعة وهكذا بالنسبة لجميع العينات التي يمكن سحبها من هذا المجتمع، سنجد أمامنا عدداً من القيم لنفس المقياس، فإذا كان المقياس المحسوب هو الوسط الحسابي سنجد أنه لدينا عدداً من الأوساط الحسابية بنفس عدد العينات المسحوبة، وإذا كان المقياس هو الانحراف المعياري، فإنه سنجد لدينا عدداً من الانحرافات المعيارية بعدد العينات المحسوبة من المجتمع، وتتغير قيمة هذا التابع الاحصائي بتغير العينة، ولن تكون جميع القيم لهذا التابع الإحصائي التي حصلنا عليها من العينات متساوية، وإنما ستكون مختلفة عن بعضها البعض وتكون توزيعاً احتمالياً يدعى توزيع المعاينة. ويعرف بأنه التوزيع التكراري لأحد التوابع الإحصائية المحسوب في العينات العشوائية ذات الحجم الواحد والتي يمكن سحبها من مجتمع إحصائي واحد .

إن هذه الاختلافات في قيم التابع الإحصائي بين العينات المسحوبة من نفس المجتمع ناتجة بسبب نوعين من الأخطاء تدعى بأخطاء المعاينة، وهي

أخطاء التحيّز وأخطاء الحظ والصدف **(الأخطاء الاحتمالية) والتي تنتج عن الفروق بين قيم التوابع للعينات والقيم الحقيقية في المجتمع (الثوابت) مع العلم** أنه في الحياة العملية، لا نسحب عينات كثيرة من نفس المجتمع، وإنما نسحب عينة واحدة فقط لاستخدامها في التحليل الإحصائي.

1 ـ أخطاء التحيز (Bais):

هي مجموعة الأسباب التي تميل لإحداث أخطاء في نفس الاتجاه سواءً أكان هذا الاتجاه سالباً أم موجباً وتبقى أخطاء التحيّز ثابتة، بالرغم من زيادة حجم العينة وتنشأ أخطاء التحيّز بسبب عدم مراعاة الأساليب العلمية في سحب العينات، أو الاختيار العمدي للعينة أو عدم تنفيذ العينة بصورة كاملة أو الإستعاضة، وتعتبر أخطاء التحيز مبطلة لصلاحية العينة، لأن الكشف عنها وتحديد طبيعة التحيز وحجمه من الأمور الصعبة جداً بل والمستحيلة أحياناً.

2 ـ الأخطاء الاحتمالية :

إن الأخطاء الاحتمالية مفهوم عام، يشمل أسباباً متعددة، وإن تأثير هذه الأخطاء في الإتجاه السلبي أو الإيجابي متكافئ الاحتمال، وإن هذه الأخطاء تميل في المدى البعيد، لإلغاء بعضها بعضاً، فنحن نعلم مسبقاً أن احتمال ظهور قطعة النقد على أحد الوجهين مساوٍ لـ (1/2) بحسب تعريف الاحتمال النظري، فإذا رمينا قطعة من النقد 100 مرة وظهرت على أحد الوجوه 45 مرة وعلى الوجه الأخر 55 مرة، فإن هذا الفرق بين الاحتمال النظري والاحتمال التجريبي ناتج عن صغر حجم العينة وهو ما ندعوه **بالخطأ الاحتمالي والذي يتوزع وفق توزيع احتمالي محدد.**

3 ـ 2 ـ المجتمعات اللامحدودة والمجتمعات المحدودة :

في دراستنا لتوزيعات المعاينة لابد لنا من أن نفرق بين العينات المسحوبة من مجتمعات كبيرة لانهائية (غير محدودة) وبين المجتمعات المحدودة، وذلك بسبب اختلاف خصائص توزيعات المعاينة، للعينات المسحوبة من المجتمعات الكبيرة عن تلك المسحوبة من المجتمعات المحدودة أو الصغيرة.

فعند سحب عينة من المصابيح من إنتاج مصنع معين، فإننا بلا شك نسحب من مجتمع كبير وهو إنتاج المصنع ولكن عند سحب عينة من طلبة كلية العلوم السياسية، فإننا بلا شك نسحب من مجتمع محدود، وكذلك عند سحب عينة مكونة من 10 مفردات من بين مجتمع مكون من مائة مفردة، في مثل هذه الحالة علينا أن نميز عما إذا كان السحب مع الإعادة، أي بالإمكان سحب المفردة أكثر من مرة، وعندها يمكن اعتبار المجتمع غير محدود، أو لانهائي، أما إذا كان السحب بدون إعادة فيمكن اعتبار سحبها من مجتمع منته هو عبارة عن توافيق C_n^N، ومثال ذلك أن عدد العينات التي يمكن سحبها من مجتمع محدود أو منته N=12 وعينه حجمها n=2 يساوي $C_2^{12} = 66$ عينة أما إذا كان السحب مع الإعادة، فإن عدد العينات مساوٍ $N^n = 12^{12} =$ 144 عينة، وإن احتمال تمثيل العينة في حالة المجتمع المحدود $\frac{1}{C_n^N}$ وفي الحالة الثانية $\frac{1}{N^n}$

3 ـ 3 ـ توزيع المعاينة للأوساط الحسابية :

لدينا مجتمع إحصائي متقطع حجمه N=4 مفردات، مكون من القيم $\{0,2,4,6\}$.

فإن :

$$\mu = \frac{0+2+4+6}{4} = 3$$

$$\delta^2 = \frac{(0-3)^2+(2-3)^2+(3-3)^2+(4-3)^2}{4} = 5$$

نفرض أننا سحبنا جميع العينات الممكنة مع الإعادة ذات الحجم n=2 ثم حسبنا $\bar{x}$ لكل عينة، فإن عدد العينات الممكن سحبها مع الإعادة يعطى بالعلاقة :

$$N^n = 4^2 = 16$$

وإن قيم $\bar{x}$ للعينات العشوائية المحسوبة تتأرجح بين (0، 6) أنظر الجدول رقم (1) المرفق :

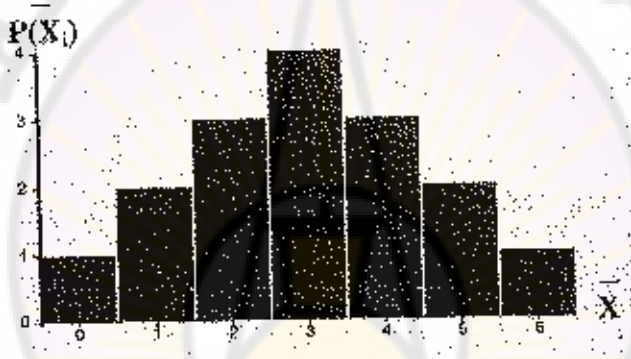
الجدول رقم (1)

رقم العينة	العينة		$\bar{X}$	رقم العينة	العينة		X
1	0	0	0	9	4	0	2
2	0	2	1	10	4	2	3
3	0	4	2	11	4	4	4
4	0	6	3	12	4	6	5
5	2	0	1	13	6	0	3
6	2	2	2	14	6	2	4

7	2	4	3	15	6	4	5
8	2	6	4	16	6	6	6

وإن الجدول الاحتمالي لتوزيع معاينة الأوساط الحسابية $\overline{X}_i$:

$\overline{X}_i$	0	1	2	3	4	5	6
$P(\overline{X}_i)$	1/16	2/16	3/16	4/16	3/16	2/16	1/16



شكل (7-1)

من خلال ذلك، نلاحظ أن توزيع المعاينة للأوساط الحسابية $(\overline{X}_i)$ للعينات يمكن أن يقترب وبشكل جيد من منحني التوزيع الطبيعي والذي يرمز لمتوسطه بـ μ_x وتشتته بـ δ_x^2 وفيما يلي حساب هاتين الإحصائيتين .

$$\mu_{\overline{x}} = \sum \overline{X}_i P(\overline{X}_i) = 0.\left(\tfrac{1}{16}\right) + 1.\left(\tfrac{2}{16}\right) + 2.\left(\tfrac{3}{16}\right) + 3.\left(\tfrac{4}{16}\right)$$

$$+ 4.\left(\tfrac{3}{16}\right) + 5.\left(\tfrac{2}{16}\right) + 6.\left(\tfrac{1}{16}\right) = 3$$

وهي نفس قيمة μ .

إذاً : $\mu_{\overline{X}} = \mu$

لذلك نقول أن ($\mu_{\overline{X}}$) هو تقدير غير متحيز لـــ (μ) أي أن الوسط الحسابي لتوزيع معاينة الأوساط الحسابية هو تقدير غير متحيز للوسط الحسابي للمجتمع وذلك مهما كان عدد العينات المسحوبة منه عشوائياً .

ويحسب التباين من العلاقة :

$$\delta_{\overline{x}} = \sum (\overline{X_i} - \mu_{\overline{X}})^2 P(\overline{X_i}) = (0-3)^2 \left(\tfrac{1}{16}\right) + (1-3)^2 \left(\tfrac{2}{16}\right) + (2-3)^2 \left(\tfrac{3}{16}\right)$$

$$+ (3-3)^2 \left(\tfrac{4}{16}\right) + (4-3)^2 \left(\tfrac{3}{16}\right) + (5-3)^2 \left(\tfrac{2}{16}\right) + (6-3)^2 \left(\tfrac{1}{16}\right) = \tfrac{5}{2}$$

ويمكن أن نلاحظ بسهولة أن توزيع معاينة الأوساط الحسابية لـــ 64 عينة من الحجم n=3 ومجتمع N=4 مختارة مع الإعادة، تكون أقرب إلى التوزيع الطبيعي بمتوسط $\mu_{\overline{X}} = 3$ وتشتت $\delta^2 \overline{x} = \tfrac{5}{3}$، وكما نلاحظ فإن الوسط الحسابي $\mu \overline{x}$ يساوي دوماً متوسط المجتمع الذي سحبت منه العينات، أما بالنسبة لتشتت المتحولات العشوائية $(\overline{X_i})$ فهو يساوي دائماً تشتت المجتمع الأصلي مقسوماً على n أي :

$$\delta^2 \overline{x} = \frac{\delta^2}{n} = \frac{5}{3}$$

ـــ أما في حالة السحب بدون إعادة يعطى عدد العينات من علاقة التوافيق التالية: C_n^N

$$C_2^4 = \frac{4!}{2!(4-2)!} = 6$$

رقم العينة	1	2	3	4	5	6
العينات	(2 ، 0)	(0 ،4)	(6 ، 0)	(4 ، 2)	(6 ، 2)	(4،6)
أوساط العينات	1	2	3	3	4	5

وإن الجدول الإحتمالي لتوزيع معاينة الأوساط الحسابية ($\overline{X}_i$) هو :

$\overline{X}_i$	1	2	3	4	5
$P(\overline{X}_i)$	1/6	1/6	2/6	1/6	1/6

وإن :

$$\mu_{\overline{X}} = \sum \overline{X}_i P(\overline{X}_i)$$

$$= \tfrac{1}{6}(1) + \tfrac{1}{6}(2) + \tfrac{1}{6}(3) + \tfrac{1}{6}(4) + \tfrac{1}{6}(5) = 3$$

التباين :

$$\delta^2_{\overline{X}} = \sum (\overline{X}_i - \mu_{\overline{X}}) P(\overline{X}_i)$$

$$= (1-3)^2 . \tfrac{1}{6} + (2-3)^2 . \tfrac{1}{6} + (3-3)^2 . \tfrac{2}{6} + (4-3)^2 . \tfrac{1}{6}$$

$$+ (5-3)^2 . \tfrac{1}{6} = \tfrac{10}{6}$$

وهي مساوية للمقدار :

$$\delta^2_{\overline{X}} = \frac{\delta^2}{n} . \frac{N-n}{N-1}$$

$$= \tfrac{5}{2} . \frac{4-2}{4-1} = \frac{5}{3}$$

وهذا يعني أن الانحراف المعياري لتوزيع معاينة الأوساط الحسابية (الخطأ المعياري لتوزيع معاينة الأوساط الحسابية) يختلف عكساً مع حجم العينة وبزيادة حجم العينة سوف يقل مقدار الخطأ المرتكب.

إن المقدار $\frac{N-n}{N-1}$ يدعى بمعامل التصحيح، وتقترب قيمة معامل التصحيح من الواحد الصحيح كلما كبرت N مقارنة بـــ n أي عندما يكون المجتمع منتهياً أو إذا كانت $\frac{n}{N} < 5\%$

فإن معامل التصحيح $\frac{N-n}{N-1} = 1$

4-3- مبرهنة النهاية المركزية :

إذا كان لدينا مجتمعاً احصائياً توقعه (μ) وانحرافه المعياري (δ) وسحبنا منه جميع العينات العشوائية الممكنة ذات الحجم n، فإن متوسطات هذه العينات تتوزع على وجه التقريب توزيعاً طبيعياً بمتوسط $\mu_{\bar{X}} = \mu$ وانحراف معياري $\delta_{\bar{X}} = \frac{\delta_x}{\sqrt{n}}$ وإذا كانت الظاهرة تخضع لتوزيع ذي الحدين، فإن هذه النسب تتوزع طبيعياً بمتوسط $\mu_{\bar{X}} = P$ وانحراف معياري $\delta_{\bar{P}} = \sqrt{\frac{P-q}{n}}$ وإن قيمة (Z) تعطى بالعلاقة :

$$Z = \frac{\overline{X}_i - \mu_{\bar{X}}}{\delta_{\bar{X}}} = \frac{\overline{X} - \mu}{\delta / \sqrt{n}}$$

وبالنسبة لتوزيع (النسبة المئوية):

$$Z = \frac{\overline{P_i} - \mu_{\overline{X}}}{\delta_{\overline{P}}} = \frac{\overline{P} - P}{\sqrt{\frac{P.q}{n}}}$$

وهذه المقادير تنتهي للتوزيع الطبيعي عندما $n \to \infty$، لذا لايختلف شكل المجتمع كثيراً عن التوزيع الطبيعي عندما $n>30$ أما إذا كان شكل المجتمع الأصلي طبيعياً، فإن توزيع العينة يتبع التوزيع الطبيعي مهما كان حجمها .

تمرين (1):

إن أعمار الناخبين في إحدى المدن يتبع التوزيع الطبيعي بوسط حسابي $\bar{x} = 30$ وانحراف معياري قدره $(S=9)$، أوجد توزيع المعاينة لعينة عشوائية من (36) ناخب .

الحل :

نحن لانعلم توزيع أعمار الناخبين قيد الدراسة، لكن بالإعتماد على نظرية الحد المركزي نعلم أنه عندما يكون حجم العينة $n>30$ سيكون توزيع الأعمار تقريباً توزيع طبيعي، لذلك نقول أن $\overline{X}$ يتوزع طبيعياً، بوسط = 30 وخطأ معياري $\delta_{\bar{x}} = \frac{9}{\sqrt{36}}$

3 ـ 5 ـ توزيع معاينة النسب المئوية :

للتعرف على هذا النوع من توزيع المعاينة نورد المثال التالي :

تمرين (2) :

ينتج مصنع أنابيب معدنية ، وتعتبر الأنابيب صالحة، إذا وقعت أقطارها ضمن حدود معينة، غير صالحة، إذا وقعت أقطارها خارج تلك الحدود، سحبنا عينة عشوائية من إنتاج المصنع حجمها (n) وجدنا منها (x) أنبوباً صالحاً للاستعمال، إن النسبة $\overline{P} = \frac{x}{n}$، هي عبارة عن وسط حسابي للمتغير العشوائي ذي الحدين ، فإذا كررنا التجربة عدة مرات، أي سحبنا العديد من هذه العينات العشوائية سوف نحصل على قيم لـــ $(\overline{P})$ بعدد العينات المسحوبة وبالتالي فإن:

$$E\left(\frac{x}{n}\right) = \frac{1}{n}E(x) = \frac{1}{n}.nP = P$$

وبالتالي نقول أن لقيمة $(\overline{P})$ هي تقدير غير متحيز لـــ (P) ، أي لنسبتها في المجتمع، أما

تباين توزيع معاينة النسب يعطى:

$$\delta_{\overline{P}}^2 = \mathrm{var}\left(\frac{x}{n}\right) = \frac{1}{n^2}\mathrm{var}(x) = \frac{1}{n^2}.nPq = \frac{P.q}{n}$$

والانحراف المعياري :

$$\delta_{\overline{P}} = \sqrt{\frac{P.q}{n}}$$

وبمعنى آخر أنه عندما تكون $n > 100$ و $0 \le P \le 1$ فإن $(\overline{P})$ تقترب من التوزيع الطبيعي بوسط (P) وتباين $\frac{p.q}{n}$.

3 ــ 6 ــ توزيع معاينة الفرق بين متوسط مجتمعين ($\bar{x}_1 , \bar{x}_2$):

إذا كان لدينا ، مجتمعان إحصائيان متوسطهما (μ_1, μ_2) وانحرافهما δ_1,δ_2 على الترتيب. وليكن ($\bar{x}_1$) متوسط العينة ذات الحجم (n_1) والمسحوبة من المجتمع الأول و($\bar{x}_2$) متوسط العينة ذات الحجم (n_2) والمسحوبة من المجتمع الثاني. فمن أجل جميع القيم الممكنة لـــ $\bar{X}_1 , \bar{X}_2$ سيكون توزيع الفرق لـــ $\bar{x}_1 - \bar{x}_2$ هو :

$$\delta_{\bar{X}_1 \ \bar{X}_2} = \sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}$$

إن اختبار معنوية الفروق للأوساط الحسابية المسحوبة من عينات عشوائية n > 30 تتبع توزيعاً طبيعياً Z(0,1) يعطى بالعلاقة :

$$Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^3}{n_2}}}$$

ولإنشاء مجال ثقة لتوزيع معنوية الفروق للأوساط الحسابية باحتمال $(1-\alpha)$.

$$P\left(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\right) = 1 - \alpha$$

$$P = \left(-Z_{\alpha/2} \le \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}} \le Z_{\alpha/2} \right)$$

$$P=\left((\mu_1-\mu_2)-Z\%_2\sqrt{\frac{\delta_1^2}{n_1}+\frac{\delta_2^2}{n_2}}\le(\overline{X_1}-\overline{X_2})\le(\mu_1-\mu_2)+Z\%_2\sqrt{\frac{\delta_1^2}{n_1}+\frac{\delta_2^2}{n_2}}\right)=1-\alpha$$

تمرين (3):

سحبت عينة حجمها n=50 من مجتمع إحصائي توقعه μ=40 وانحرافه المعياري δ=10 والمطلوب :

1 ــ ما هو احتمال أن يقع متوسط هذه العينة من القيمتين 42, 41

2 ــ ماهو احتمال أن يزيد متوسط هذه العينة على 39 .

الحل :

(1

$$P(41\le\overline{X}\le42)=P\left(\frac{\overline{X}-\mu}{\delta/\sqrt{n}}\le Z\le\frac{\overline{X}-\mu}{\delta/\sqrt{n}}\right)$$

$$=P\left(\frac{41-40}{10/\sqrt{50}}\le Z\le\frac{42-40}{10/\sqrt{50}}\right)=P(0.7\le Z\le1.41)$$

$$=P(Z\le1.41)-P(Z\le0.7)$$
$$=0.4207-0.258=0.1627$$

(2

$$P(\overline{X}\ge39)$$
$$=P(Z\ge0)+P(Z\le0.7)$$
$$=0.5+0.258=0.758$$

تمرين (4) :

إذا كان أطوال الطلاب في إحدى الجامعات السورية يتبع توزيعاً طبيعياً بمتوسط $\mu=170$ وانحراف معياري $\delta=8$ سحبت عينة عشوائية مكونة من 64 طالباً **والمطلوب :** ما احتمال أن يكون متوسط أطوالهم :

1 ــ أكبر من 172 سم . 2 ــ أقل من 169 سم 3 ــ أقل من 173 سم

4 ــ بين القيمتين (172 و 169).

الحل :

(1

$$P(\overline{X} \geq 172) = P\left(Z \geq \frac{172-170}{8/\sqrt{64}} \right) = P(Z \geq 2)$$

$$= P(Z \geq 0) - P(Z \leq 2) = 0.228$$

(2

$$P(\overline{X} \leq 169) = P\left(Z \geq \frac{169-170}{8/\sqrt{64}} \right) = P(Z \leq -1)$$

(3

$$P(Z \geq 1) = P(Z \geq 0) - P(Z \leq 1) = 0.15865$$

$$P(\overline{X} \leq 173) = P(Z \leq 0) + P(Z \leq 3) =$$

$$0.5 + 0.49865 = 0.99865$$

(4

$$P(169 \leq \overline{X} \leq 172) = P(-1 \leq Z \leq 2)$$

$$P(Z \leq 1) + P(Z \leq 2) = 0.818$$

تمرين (5) :

يدعي المكتب المركزي للإحصاء في سورية أن 60% من أسر مدينة دمشق تتألف من أربعة أفراد فأكثر، سحبت عينة مؤلفة من (600) أسرة، فإذا علمت أن هذه النسبة تخضع لمنحني توزيع طبيعي.

والمطلوب :

1 ــ ما احتمال أن يكون هناك 396 أسرة يزيد متوسط عددها أفرادها عن أربعة ؟

2 ــ ما احتمال أن يكون هناك 396 أسرة يقل متوسط عدد أفرادها عن أربعة.

الحل :

$$\overline{P} = \frac{x}{n} = \frac{396}{600} \times 100 = 66\%$$

$$(\overline{P} \geq 66) = P\left(Z \geq \frac{\overline{P} - P}{\sqrt{\frac{P.q}{n}}} \right)$$

(1

$$Z = \frac{\overline{P} - P}{\sqrt{\frac{P.q}{n}}} = \frac{66 - 60}{\sqrt{\frac{60 \times 40}{600}}} = 3$$

$$P(Z \geq 3) = P(Z \geq 0) - P(Z \leq 3)$$

$$= 0.50 - \frac{0.9973}{2} = 0.00135$$

(2

$$= 1 - P(Z \geq 3)$$

$$= 1 - 0.00135 = 0.99865$$

أو :

$$P(\overline{P} \leq 66) = P(Z \leq 3)$$

$$= P(Z \leq 0) + P(Z \leq 3)$$

$$= 0.5 + \frac{0.9973}{2} = 0.9986$$

تمرين (6) :

إذا كانت نسبة حملة الشهادة الجامعية في إحدى الشركات 3%، وقام أحد الباحثين بسحب عينة عشوائية من 400 عامل من عمال هذه الشركة، فما هو احتمال :

1 ـ أن يجد الباحث 20 عاملاً على الأقل من حملة الشهادة الجامعية ؟

الحل :

$$\overline{P} = \left(x/n\right) = 20/400 \times 100 = 5\,\%$$

$$P(\overline{P} \geq 5) = P\left(Z \geq \frac{\overline{P} - P}{\sqrt{\frac{P.q}{n}}}\right)$$

$$= P\left(Z \geq \frac{5-3}{\sqrt{\frac{3\times 97}{400}}}\right) = P(Z \geq 2.35)$$

$$= P(Z \geq 0) - P(Z \leq 2.35) = 0.5 - 0.4906 = 0.094$$

ملاحظات :

1ـ إن التوزيع الطبيعي من وجهة النظر الواقعية مستحيلاً، غير أن نظرية الاحتمالات سمحت لنا باستنتاجات: أن مختلف توزيعات المعاينة للعينات الكبيرة يكون لها شكل التوزيع الطبيعي اعتماداً على قانون الأعداد الكبيرة الذي يعتبر الأساس في نظرية الاحتمالات وينص على أنه: " كلما ازداد حجم العينة فإن الاحتمال يقترب من اليقين والفرق المشاهد بين الاحصاء والمعلمة يصبح أصغر مايمكن".

2 ـ عندما يكون التوزيع الطبيعي توزيعاً للمعاينة، يمكننا القول أن هناك (68.26%) من العينات التي يكون الفرق بين الاحصاء والمعلمة أقل من خطأ معياري واحد وهناك (95.5%) من العينات، الفرق بين الاحصاء والمعلمة لها أقل من خطئين معياريين وهكذا ... وإن هناك (27%) من العينات العشوائية غير اعتيادية واقعة على نهايتي المنحني.

3 ـ عند تعميم النتائج من خلال العينة يجب أن نأخذ بالحسبان الأخطاء الاحتمالية وليس هناك ضرورة للقول بأن معالجة الأخطاء الاحتمالية يتم بتكبير حجم العينة لأن هذا الأمر مرفوض لذلك نحسب أثر الأخطاء الاحتمالية على النتائج التي توصلنا إليها ثم نأخذ تلك الأخطاء على المجتمع، إلا أنه لكي نحدد هذه الأخطاء لابد لنا من مقياس نعتبره مثالي تقاس على أساسه العينات للعشوائية لمعرفة مقدار الخطأ الذي أصيبت به العينة وهذا المقياس هو توزيع المعاينة .

تمارين غير محلولة

تمرين (1) :

إذا كان العمر الوسطي لمصباح كهربائي ($\mu = 5$) سنوات وانحراف معياري $\delta=1$ سنة وكانت أعمار المصابيح تتبع التوزيع الطبيعي .

احسب احتمال أن يقع متوسط عينة حجمها ($n=9$) من هذه المصابيح بين القيمتين (5.2 و 4.4).

تمرين (2) :

إذا كان متوسط الدخل الأسبوعي لمجموعة كبيرة من العمال المهرة $\mu=2728$ وحدة نقدية بانحراف معياري مقداره 100 وحدة نقدية، فما هو إحتمال أن يكون متوسط دخل عينة حجمها ($n=100$) من هؤلاء العمال أكثر من 2756 وحدة نقدية .

تمرين (3) :

سحبت عينة عشوائية حجمها ($n=64$) من مجتمع إحصائي متوسطه $\mu=10$ وانحراف معياري ($\delta=3$)

والمطلوب: ما احتمال :

1 ــ أن يقع متوسط هذه العينة بين القيمتين (11 و 9)

2 ــ ماهو احتمال أن يزيد متوسط هذه العينة على القيمة (11).

تمرين (4) :

سحبت عينة ($n=10$) من مجتمع إحصائي متوسطه = 80 وانحراف معياري ($\delta = 8$) فماهو احتمال ألا يختلف متوسط هذه العينة عن μ بأكثر من (± 2) .

تمرين (5) :

إذا كان (x) متغيراً عشوائياً يمثل أطوال الطلاب في كلية العلوم السياسية بجامعة دمشق وكان (x) خاضعاً لتوزيع طبيعي وسطه (160) سم وانحرافه المعياري (20) سم. أخذت عينة عشوائية حجمها (16) طالباً من هذه الكلية.

المطلوب:

احسب احتمال أن يقل الوسط الحسابي لأطوال أفراد هذه العينة عن (163) سم .

تمرين (6):

إذا كانت الأجور اليومية لعمال إحدى الشركات في دمشق تخضع لتوزيع طبيعي N(5 ، 0.5) وكانت الأجور اليومية لعمال شركة مماثلة في حلب تخضع لتوزيع طبيعي N(4 ،0.25) أخذت عينة عشوائية من الشركة الأولى حجمها (20) وعينة عشوائية من الشركة الثانية حجمها (10)، فإذا رمزنا لوسط أجور العينة الأولى بالرمز $(\overline{X}_1)$ والثانية $(\overline{X}_2)$ **المطلوب:**

أوجد : $P(\bar{x}_1 - \bar{x}_2 \leq 0.75)$

الفصل الرابع

نظرية التقدير للعينات الكبيرة الحجم

Theory Of Estimation for Big Samples

- مقدمة (التقدير المجالي ، التقدير النقطي)
- خصائص التقدير الأمثل (عدم التحيّز، التناسق، الكفاءة، الكفاية) .
- تقدير المتوسط لمجتمع إحصائي عندما يكون تباين المجتمع σ^2 معلوماً
- تقدير النسبة لمجتمع إحصائي.
- تقدير الفرق بين متوسطي مجتمعين إحصائيين.
- تقدير الفرق بين نسبتي مجتمعين إحصائيين.

4 ـ 1 ـ مقدمة :

يعالج الاستدلال الإحصائي مسألتين هامتين هما اختبار الفروض والتقدير الإحصائي الذي يعرّف بأنه استنتاج المعالم (البارامترات) والخصائص الأساسية للظواهر التي تخص المجتمع بدلالة المقاييس الإحصائية المحسوبة من العينة (الإحصاءات) مع الأخذ بعين الاعتبار للأخطاء الاحتمالية.

وتتلخص نظرية التقدير بأنه إذا كان لدينا متغيراً عشوائياً يتبع توزيعاً احتمالياً $f(x,\theta)$ حيث θ معلمة (بارا متر) التوزيع مجهولة القيمة ويراد الحصول على تقدير لها بناءً على اختيار عينة عشوائية $X_1, X_2, \ldots X_n$ من هذا التوزيع، وذلك باختيار تابع لهذه المتغيرات يمكن استخدامها كمقدر للمعلمة θ ، ويسمى هذا التابع المستخدم في تقدير θ بالمقدر (Estimator) والقيمة التي نحصل عليها من هذا التابع هي قيمة المقدر وتسمى (Estimate). وحيث أن المقدر تابع في قيم العينة العشوائية عليه سيكون المقدر نفسه متغيراً عشوائياً يتغير بتغير العينة، كما أن القيمة التي سنحصل عليها كتقدير لقيمة المعلمة غالباً ما تختلف عن القيمة الحقيقية للمعلمة، لأن تقديرنا معتمد على عينة عشوائية، ولايمكن الحصول على القيمة الحقيقية للمعلمة إلا إذا تم اختيار المجتمع كله. ولكن اختيار المجتمع ككل ليس ممكناً من جميع النواحي مثل التكاليف والجهد والوقت والدقة، وهناك طريقتان للحصول على تقدير لبارامتر (معلمة) المجتمع (التوزيع) هما :

ـ التقدير النقطي (أو تقدير النقطة) : Point estimation

ـ والتقدير المجالي (أو التقدير بالفترة) : Interval estimation

ونلاحظ في حياتنا اليومية جمل أو عبارات مستخدمة تدل على هاتين الطريقتين، فمثلاً، نقول إن متوسط دخل العامل في مصنع ماهو 15000 ليرة

أو نقول إن متوسط دخل العامل في مصنع ما يتراوح مابين 10000 إلى 18000 ليرة (وقد تكتب كفترة مثل : (18000 ــ 10000) . حيث استعملنا في التعبير الأول التقدير النقطي، أي أننا قدرنا متوسط الدخل بقيمة واحدة أي نقطة واحدة في فراغ كل النقاط الممكنة، أما في التعبير الثاني فقد قدرنا متوسط الدخل بفترة ، نعتقد أن القيمة الحقيقية تقع داخل هذه الفترة، وبهذه الطريقة ونفس المفهوم، نقدر المعلمة المجهولة إما تقديرها بنقطة أو تقديرها بفترة.

4 ــ 2 ــ خصائص التقدير الامثل :

أ ــ عدم التحيّز Unbiased estimators :

تعرف خاصية عدم التحيّز على أنها إذا كانت القيمة المتوقعة للمقدر مساوية للقيمة الحقيقية، فإن المقدر يسمى مقدراً غير متحيِّزاً، عدا ذلك يقال عن المقدر بأنه مقدراً متحيّزاً، أي إذا كان $X_1, X_2, \ldots X_n$ عينة عشوائية مأخوذة من توزيع احتمالي بتابع كثافة احتمالي $f(x;\theta)$ حيث θ هي معلمة (بارامتر) التوزيع، فإنه يقال إن المقدر $T = t(x_1, x_2, \ldots x_n)$ متجمع ويتركز حول قيمة θ الحقيقية .

مثال (1):

إذا كان متوسط وتباين عينة عشوائية مأخوذة من التوزيع الطبيعي $N(\mu, \sigma^2)$ هما على الترتيب $\bar{X}$ و S^2 حيث أن $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ وأن $S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$

بيّن أن $\bar{X}$ مقدر غير متحيّز لمتوسط المجتمع μ وأن S^2 مقدر غير متحيز لتباين المجتمع σ^2، أي أن :

$$E(\bar{X}) = \mu \qquad E(S^2) \neq \sigma^2$$

الحل : حسب نظرية النهائية المركزية، فإن :

$$E(\overline{X}) = E\left(\sum_{i=1}^{n} x_i / n\right) = \frac{1}{n}\sum E(x_i)$$

$$= \frac{1}{n}[E(x_1) + E(x_2) + ...E(x_n)]$$

$$= \frac{1}{n}[\mu + \mu + ... + \mu] = \frac{1}{n} \times n\mu = \mu$$

إذاً $\overline{X}$ مقدر غير متحيز لمتوسط المجتمع μ ، وباستخدام نظرية النهاية المركزية أيضاً نجد أن:

$$E(S^2) = E\left[\frac{\sum(x_i - \overline{x})^2}{n}\right] = \frac{1}{n}E\left[\sum(x_i - \mu + \mu - \overline{x})^2\right]$$

$$= \frac{1}{n}E\left[\sum[(x - \mu) - (\overline{x} - \mu)]^2\right]$$

$$= \frac{1}{n}E\left[\sum(x_i - \mu)^2 - 2(\overline{x} - \mu)\sum(x_i - \mu) + n(\overline{x} - \mu)^2\right]$$

$$= \frac{1}{n}E\left[\sum(x_i - \mu) - 2n(\overline{x} - \mu)^2 + n(\overline{x} - \mu)^2)\right]$$

أي أن :

$$E(S^2)=\frac{1}{n}\left[\sum E(x_i-\mu)^2-nE(\bar{x}-\mu)^2\right]=\frac{1}{n}E\left[\sum\sigma^2-n\frac{\sigma^2}{n}\right]$$

$$=\frac{1}{n}\left[n\sigma^2-\sigma^2\right]=\frac{1}{n}[n-1]\sigma^2=\frac{(n-1)}{n}\sigma^2$$

وعليه فإن S^2 مقدر متحيز لتباين المجتمع σ^2، من الواضح أنه لو كان حجم العينة كبيراً يكون S^2 مقدراً غير متحيزاً حتى بالقسمة على n (حجم العينة الكبير).

وإذا كان حجم العينة صغيراً يكون S^2 مقدراً غير متحيزاً إذا تمت القسمة على (n-1) بدلاً من n، وقد سُمي في هذه الحالة المقدر غير متحيزاً .

أي عندما $n\to\infty$ فإن $E(S^2)\to\sigma^2$.

ب ــ التناسق : Consistency

إذا كان لدينا قطعة نقود وكان احتمال ظهور H هو P، فإذا ألقيت قطعة النقود n مرة مستقلة فإن عدد ظهور H هو x مرة. في هذه الحالة يمكن تقدير P بالمقدر :

$$T=\frac{x}{n}$$

من الواضح أنه كلما زادت قيمة n كلما اقتربت القيمة $\frac{x}{n}$ من القيمة الحقيقية للمعلمة p وحيث أن T متغير عشوائي، فهذا يعني باستخدام لغة الاحتمالات أن احتمال الفرق بين قيمة المقدر والقيمة الحقيقية للمعلمة P يكون صغيراً جداً، حيث يقترب من الواحد كلما زادت قيمة n، ويساوي الواحد عندما $n\to\infty$ ، أي أن : $P(|T-P|\le c\le 1)$.

هذه الخاصية تسمى " بالتناسق " . وعلى ذلك، إذا كـــان المقـــدر Tn مقـــدراً متناســـــقاً للمعلمـــــــة θ فيكـــــون : $\lim_{n\to\infty} P(|T_n-\theta|\le\varepsilon)=1$ أو :

$$\lim_{n\to\infty} P(|T_n-\theta|\le\varepsilon)=0$$

ج ــ الكفاءة (الفعالية) : Efficiency

لقد اتضح لنا من خاصية عدم التحيّز أنه يمكن الحصول على أكثر من مقدر واحد غير متحيز للمعلمة، وبالتالي كيف نستطيع أن نحدد أو نفضل أو نختار أحدهما على الآخرين.

ونحن سوف نختار المقدر الأكثر تجمعاً حول المعلمة. أي المقدر الأقل اختلافاً أو الأقل تشتتاً هو المقدر الأفضل. وحيث أن التباين يعتبر أحسن مقاييس التشتت، وبالتالي يكون المقدر الأفضل هو الأقل تبايناً. فإذا كان لدينا T_1, T_2 مقدرين غير متحيّزين للمعلمة θ وكان T_1 أكثر تجمعاً أي أقل اختلافاً وأقل تشتتاً حول θ من المقدر T_2 يكون T_1 أكثر كفاءة من T_2 .

من الواضح أن المقدر T_1 أقل تشتتاً أي أكثر تجمعاً حول المعلمة θ من المقدر T_2، فهو بالتالي أكثر كفاءة. وتعرف الكفاءة النسبية E للمقدر T_1 كما يلي :

$$E=\frac{\mathrm{var}(T_1)}{\mathrm{var}(T_2)}$$

فإذا كان $\mathrm{Var}(T_1) < \mathrm{Var}(T_2)$ تكون E أقل من الواحد ويكون T_1 أكثر كفاءة من T_2،

وإذا كانت E أكبر من الواحد يكون T_2 أكثر كفاءة من T_1 .

مثال (2):

المقدر $T = \frac{x}{n}$ كمقدر للمعلمة P للتوزيع الثنائي تزيد كفاءته كلما زاد حجم العينة .

لنبيّن أن $T = \frac{x}{n}$ مقدر غير متحيّز للمعلمة (الثابت) P، فإن :

$$E(T) = E\left(\frac{x}{n}\right) = \frac{1}{n}E(x) = \frac{1}{n}np = p$$

وكذلك نجد أن :

$$\text{var}(T) = \text{var}\left(\frac{x}{n^2}\right) = \frac{1}{n^2}\text{var}(x) = \frac{1}{n^2}.npq = \frac{pq}{n}$$

وعليه كلما زاد حجم العينة كلما كان للمقدر تباين أقل أي أكثر كفاءة .

د ــ الكفاية : Sufficiency

تعني أن المقدر يحتوي على جميع المعلومات التي توفرت من العينة ولايمكن لأي مقدر آخر أن يحتوي على معلومات ليست موجودة بالمقدر الأصلي، فيقال أن هذا المقدر كاف لتقدير المعلمة حيث لايوجد مقدر آخر سوف يأتي من العينة المسحوبة بمعلومات أكثر.

4 ــ 3 ــ تقدير المتوسط لمجتمع إحصائي (μ) عندما يكون تباين المجتمع σ^2 معلوماً:

إذا أخذت عينة عشوائية $X_1, X_2, \ldots, X_2$ حجمها n من مجتمع طبيعي متوسطه غير معلوم μ وتباينه σ^2 بحيث أن σ^2 معلوم، فيمكن إيجاد عدد، وليكن $Z_{\alpha/2}$ من جدول لتوزيع الطبيعي المعياري بحيث أن :

$$P\left(-z_{\alpha/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}\right) = 1-\alpha$$

حيث أن $\sigma > 0$ فإن المتراجحة السابقة يمكن كتابتها بصورة مكافئة كما يلي :

$$-z_{\alpha/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < Z_{\alpha/2}$$

$$-z_{\alpha/2} < \left(\frac{\sigma}{\sqrt{n}}\right) \le \bar{X}-\mu \le Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

إذا كان الاحتمال للأول هو $(1-\alpha)$، فإن الاحتمال للآخر يجب أن يكون أيضاً مساوٍ لنفس الاحتمال أي $(1-\alpha)$ هذا يعني أن :

$$P\left[-z_{\alpha/2} < \left(\frac{\sigma}{\sqrt{n}}\right) \le \bar{X}-\mu \le Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)\right] = 1-\alpha$$

ومنه فإن الاحتمال (عند مستوى ثقة $(1-\alpha)\%$ هو :

$$\left[\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\alpha/2}\frac{\sigma}{n}\right] = \bar{X} \pm Z_{\alpha/2}\frac{\sigma}{n}$$

سوف يتضمن المتوسط غير المعلوم (μ)حسب نظرية النهاية المركزية عندما $(n\ge30)$.

مثال (3) :

إذا كان X هو العمر الزمني لمصباح كهربائي بفرض أن X متغير عشوائي طبيعي $X\sim N(\mu, 1296)$. تم اختيار عينة عشوائية حجمها $n=37$

مصباحاً استخدمت حتى تلاشت من جراء الاستخدام معطية متوسط عينة $\bar{x} = 1478$ ساعة والمطلوب :

قدر باحتمال ٩٥% متوسط عمر المصابيح في المجتمع (مجتمع المصابيح) .

$$P\left[\bar{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le x \le \bar{x} + \left(\frac{\sigma}{\sqrt{n}}\right)\right] = 1 - \alpha$$

$$P\left[1478 - 1.96\left(\frac{36}{\sqrt{37}}\right) \le x \le 1478 + 1.96\left(\frac{36}{\sqrt{37}}\right)\right] = 0.95$$

$$P[1478 - 11.60 \le x \le 1478 + 11.60] = 0.95$$

$$P[1478 - 11.60 \le x \le 1478 + 11.60] = 0.95 \Rightarrow$$

$$[1466.6, 1489.6]$$

وبالتالي يمكن القول أن متوسط عمر المصابيح لن يقل عن (1466.6) ولن يزيد عن (1489.6) ساعة .

ويمكن تفسير ذلك أنه لو أخذنا (37) عينة عشوائية وحسبنا فترة الثقة من كل عينة لوجدنا أن 95 فترة تتضمن قيم الوسط الحسابي للمجتمع μ، و5 فترات لاتتضمن تلك القيم، مع العلم أنه يمكن إنشاء فترات ثقة أخرى عند مستويات مختلفة بإتباع نفس الأسلوب واستخدام جدول التوزيع الطبيعي المعياري للحصول على قيم Z، إلا أن أكثر فترات الثقة شيوعاً واستخداماً هي:

ـ عند 90% ومستوى ثقة يكون :

$$1-\alpha=0.90 \Rightarrow \alpha=0.10 \Rightarrow Z_{0.10/2}=Z_{0.05} \Rightarrow \bar{x} \pm 1.645\frac{\sigma}{\sqrt{n}}$$

ـ عند 95% مستوى ثقة يكون :

$$1-\alpha=0.95 \Rightarrow \alpha=0.05 \Rightarrow Z_{0.05/2}=Z_{0.025} \Rightarrow \bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}}$$

ـ عند 99% مستوى ثقة يكون :

$$1-\alpha=0.99 \Rightarrow \alpha=0.01 \Rightarrow Z_{0.01/2}=Z_{0.005} \Rightarrow \bar{x} \pm 2.58\frac{\sigma}{\sqrt{n}}$$

من خلال ذلك نجد أن زيادة مستوى الثقة يصاحبه زيادة في قيم Z، الأمر الذي يعني أن مجال الثقة يصبح أكثر وسعاً كلما زاد مستوى الثقة، وهذه الخاصية لفترة الثقة توضح لنا أنه كلما ازدادت الثقة فإن تقدير مجال الثقة التي تحوي قيمة μ الحقيقية، سيكون أكبر .

4 ـ 4 ـ التقدير المجالي لنسبة المجتمع الإحصائي (P) :

عند إجراء تجربة (n) مرة من محاولات بيرنولي فإن الاحتمال (P) من النجاحات عند كل محاولة، هو: $P=\frac{Y}{n}$ حيث أن Y عدد النجاحات و n: كبيرة كبراً كافياً

فإن قيمة : $Z=\frac{Y-np}{\sqrt{np(1-p)}}=\frac{(Y/n)-p}{\sqrt{P(1-p)/n}} \approx N(0,1)$

ومن أجل الاحتمال (1-α) يمكنا إيجاد $Z_{\alpha/2}$ بحيث أن :

$$P\left[-Z\%_2 \le \frac{(Y/n)-P}{\sqrt{P(1-P)/n}} \le Z\%_2\right] = 1-\alpha$$

$$P\left[\frac{Y}{n} - Z_{\alpha/2}\sqrt{\frac{P(1-P)}{n}} \le P \le \frac{Y}{n} + Z_{\alpha/2}\sqrt{\frac{P(1-P)}{n}}\right] = 1-\alpha$$

حيث (P) غير معلوم. بفرض أن التقدير النقطي لـــ P هو $\hat{P} = y/n$ فتكون فترة الثقة للمعلمة P هي :

$$\hat{P} \pm Z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

وبإدخال %(1-α) فترة الثقة يمكن كتابة العلاقة السابقة على النحو التالي:

$$P\left[\frac{Y}{n} - Z_{\alpha/2}\sqrt{\frac{(Y/n)(1-Y/n)}{n}} \le P \le \frac{Y}{n} - Z_{\alpha/2}\sqrt{\frac{(Y/n)(1-Y/n)}{n}}\right] = 1-\alpha$$

أو ما يكافئها :

$$p\left[\frac{Y}{n} \pm Z_{\alpha/2}\sqrt{\frac{(Y/n)(1-Y/n)}{n}}\right] = 1-\alpha$$

مثال (4):

في استطلاع للرأي ، تم سؤال عينة مكونة من (200) شخص عن أهم مصدر للمعلومات لديهم فأجاب (36) منهم بأنه التلفزيون، والمطلوب :

قدّر باحتمال (95%) نسبة الأفراد في المجتمع الذين يعتبرون التلفزيون أهم مصدر للمعلومات لديهم؟

$$\hat{P} = \frac{Y}{n} = \frac{36}{200} = 0.18$$

نحن نعلم أن :

$$1-\alpha = 0.95 \Rightarrow Z_{\alpha/2} = Z_{\alpha/2} = Z_{0.025} = 1.96$$

$$0.18 \pm 1.96\sqrt{\frac{(0.18)(0.82)}{200}} = [0.127, 0.233]$$

وبالتالي يمكن القول أن نسبة الأفراد في المجتمع الذين يعتبرون التلفزيون أهم مصدر للمعلومات لديهم لن تقل عن (0.12) ولن تزيد عن (0.23).

4 ــ 5 ــ تقدير الفرق(μ_1-μ_2) بين متوسطي مجتمعين إحصائيين معلومي التباين:

لنفرض أن عينتين عشوائيتين بحجم n_1, n_2 قد سحبتا من مجتمعين مستقلين طبيعيين غير معلومة متوسطاتها μ_1, μ_2، فإذا كان $\overline{X}_1$ هو متوسط العينة الأولى و $\overline{X}_2$ متوسط العينة الثانية. فإن تقدير الفرق $\mu_1 - \mu_2$ هو $\overline{X}_1 - \overline{X}_2$. وللحصول على تقدير فترة للفرق μ_1-μ_2، فإننا سنتعامل مع الحالة التي يكون فيها تباين المجتمعين σ_1^2, σ_2^2 معلومين. وأن العينتين عشوائيتين مأخوذتين من مجتمع يتوزع طبيعياً وإن $\overline{X}_1 - \overline{X}_2$ تتوزع طبيعياً أيضاً وكذلك الفرق $\overline{X}_1 - \overline{X}_2$ له توزيع طبيعي بوسط (μ_1 - μ_2) وإن تباين $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ وبناءً على ذلك فإن الإحصاءه تعطى بالعلاقة التالية :

$$Z = \left[(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)\right] \Big/ \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

تخضع للتوزيع الطبيعي N(0,1) . وعليه من أجل الاحتمال (1-α) يمكن كتابة العلاقة السابقة على النحو التالي :

$$P\left[-Z_{\alpha/2} \le \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \le Z_{\alpha/2}\right] = 1 - \alpha$$

$$P[(\overline{X}_1 - \overline{X}_2) - Z_{\alpha/2}\sigma_w \le (\mu_1 - \mu_2) \le (\overline{X}_1 - \overline{X}_2) + Z_{\alpha/2}\sigma_w] = 1 - \alpha$$

حيث أن : $\sigma_w = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$ وتكون فترة الثقة للفرق $(\mu_1-\mu_2)$ وبمعلومية

(σ_1, σ_2) على الصورة التالية :

$$P[(\overline{X}_1 - \overline{X}_2) - Z_{\alpha/2}\sigma_w \le (\mu_1 - \mu_2) \le (\overline{X}_1 - \overline{X}_2) + Z_{\alpha/2}\sigma_w] = 1 - \alpha$$

أو بصورة مكافئة $\overline{X}_1 - \overline{X}_2 \pm Z_{\alpha/2}\sigma_w$ فترة ثقة للفرق بين المتوسطين $(\mu_1 - \mu_2)$ والتباينين σ_1^2, σ_2^2 معلومين ووفقاً لنظرية النهاية المركزية فإن النتيجة يمكن استخدامها لعينات عشوائية مستقلة من مجتمعات غير طبيعية معروفة تبايناتها σ_1^2, σ_2^2 شريطة أن يكون حجم العينات كبيراً كبراً كافياً $(n_1, n_2 \ge 30)$.

وعندما تكون σ_1^2, σ_2^2 غير معلومة ولكن أحجام العينات كبيرة $(n_1, n_2 \ge 30)$ فإنه يتم استبدال σ_1^2, σ_2^2 بتقديراتهم غير المتحيّزة S_1^2, S_2^2 ، على الترتيب ويصبح لدينا في هذه الحالة $S_w = \sqrt{S_1^2/n_1 + S_2^2/n_2}$ وعليه فإن فترة الثقة للفرق $\mu_1 - \mu_2$ تكون :

$$[(\overline{X_1}-\overline{X_2})-Z_{\alpha/2}S_w,(\overline{X_1}-\overline{X_2})-Z_{\alpha/2}S_w]=(\overline{X_1}-\overline{X_2})\pm Z_{\alpha/2}S_w$$

مثال (5) :

في دراسة حول درجات الطلاب في مقرر الإحصاء، أخذت عينتان مستقلتان، أظهرت النتائج الآتي :

$$n_1=150\ ,\ n_2=80\ ,\ \overline{x_1}=75{,}3\ ,\ \overline{x_2}=70,\quad S_1^2=60,\quad S_2^2=40$$

والمطلوب: قدر باحتمال (90%) الفرق الحقيقي بين وسط درجات هؤلاء الطلاب.

الحل :

$$1.645\sigma_w=1.645\sqrt{\frac{60}{150}+\frac{40}{80}}$$

$$=1.560$$

$$[+5+1.560\ ,+5-1.560]=[\ 6.5\ ,+3.5]$$

وبالتالي يمكن القول أنه باحتمال (90%) فإن الفرق بين متوسطي المجتمعين لن يقل (3.5) ولن يزيد عن (6.5) درجة .

4 ــ 6 ــ تقدير الفرق بين نسبتي مجتمعين إحصائيين بحيث ينصب الاهتمام :

في هذه الحالة يكون لدينا مجتمعان بحيث ينصب الاهتمام في استقراء وجود فروق بين النسب للمجتمعين بالاعتماد على عينتين عشوائيتين بسيطتين مستقلتين. بفرض أن : نسبة المجتمع الأول هي P_1 ونسبة المجتمع الثاني هي P_2 وأن $\hat{P}_1=\frac{Y_1}{n_1}$ هو تقدير لنسبة المجتمع الأول وأن $\hat{P}_2=\frac{Y_2}{n_2}$ هو أيضاً تقدير لنسبة المجتمع الثاني. حيث أن (n_1,n_2) عينتين عشوائيتين قد سحبتا من مجتمعين مستقلين طبيعين وأن تقدير الفرق P_1-P_2 هو $(\hat{P}_1-\hat{P}_2)$ وبناءً على ذلك نوجد فترة الثقة للفرق P_1-P_2 من خلال :

$$Z = \frac{(Y_1/n_1)-(Y_2/n_2)-(P_1-P_2)}{\sqrt{P_1(1-P_1)/n_1 + P_2(1-P_2)/n_2}}$$

تخضع للتوزيع الطبيعي N(0,1). وباحتمال %(1-α) يمكن كتابة :

$$P\left[\left(\frac{y_1}{n_1}-\frac{y_2}{n_2}\right)\pm Z_{\alpha/2}\sqrt{\frac{(y_1/n_1)(1-y_1/n_1)}{n_1}+\frac{(y_2/n_2)(1-y_2/n_2)}{n_2}}\right]=1-\alpha$$

مثال (6):

أجريت دراسة في إحدى الجامعات السورية حول الموقف من تسلم المرأة مناصب عليا في الدولة، أخذت عينتان عشوائيتان (n_1=100) طالب (n_2=120) طالبة، فوجد أن (40) طالباً أيدوا تسلم المرأة مناصب عليا في الدولة، فيما أيدت (60) طالبة هذا الموقف.

والمطلوب : أوجد حدود النقطة (95%) (قدر باحتمال 95%) الفرق الحقيقي لمواقف الطلاب والطالبات في هذه الجامعة .

الحل :

الطلاب P_1=40/100=40% q_1=60%

الطالبات P_2=60/120=50% q_2=50%

$$P\left[(0.5-0.4)\pm 1.96\sqrt{\frac{(0.4)(0.6)}{100}+\frac{(0.5)(0.5)}{120}}\right]=0.95$$

حيث أن : (P_1-P_2) هي :

P [0.04 , 0.24] = 0.95

وبالتالي يمكن القول أنه باحتمال (95%) فإن الفرق بين نسبتي المجتمعين لن يقل عن (0.04) ولن يزيد عن (0.24).

الفصل الخامس

اختبار الفروض الإحصائية

للعينات الكبيرة الحجم

Testing Hypotheses for Big Samples

- مقدمة
- أنواع الفروض الإحصائية
- اختبار الفرضيات لمتوسط مجتمع.
- منطقة الرفض، منطقة القبول، المنطقة الحرجة.
- اختبار الفرضيات لنسبة مجتمع .
- اختبار الفرضيات للفرق بين نسبتين مئويتين (P_1-P_2).
- اختبار الفرضيات للفرق بين متوسطي مجتمعين$(\mu_1-\mu_2)$.

1-5 ـ مقدمة :

إن اختبار الفروض عبارة عن تحديد عما إذا كانت بيانات العينة تؤيد اعتقاداً معيناً في المجتمع أم لا، أي مدى تطابق الإحصاءات مع الفروق هي أخطاء احتمالية. والفرضية هي إدعاء أو اعتقاد حول خاصية أو أكثر من خصائص مجتمع ما، فعلى سبيل المثال: يدعى أحد دكاترة الإحصاء في إحدى الكليات بأنه استخدم طريقتين تدريسيتين A,B مع طلابه، وأن الطريقة A بيّنت أنها أفضل بكثير من الطريقة B حيث وجد أن متوسط النجاح من هذه الطريقة هو 85%. في هذا المثال، من الواضح أننا لا نستطيع التسليم بإدعائه ما لم نقوم باختبار هذا الإدعاء، الذي يعطينا حكماً على قبول أو رفض هذا الإدعاء. ويتم ذلك من خلال وضع الفرضية التالية: $\mu=85$ درجة (حيث μ هو المتوسط الحقيقي للنجاح في هذه المادة) . وللتأكد من ذلك، يقوم الباحث باستخدام بيانات ناتجة عن عينة من الكليات بتجربة تلك الطريقة التدريسية وذلك لبناء قرار من خلال فرضيتين : الأولى أن $\mu=85$ درجة أي أن إدعاء الدكتور كان صائباً والفرضية الأخرى أن $\mu < 85$ درجة أي أن معدل (متوسط) النجاح بعد استخدام الطريقة المذكورة أقل من النسبة التي تنتج عادة عند التدريس بالطرق المتبعة، وبالتالي فإن إدعاء هذا الدكتور غير صحيحاً، أي أن اختبار الفروض يهدف للتأكد من نتائج العينة ومدى مطابقتها أو مخالفتها للتخمينات التي وضعها الباحث عن المجتمع.

2-5 ـ أنواع الفروض الإحصائية:

يقوم اختبار الفرضيات على فرضيتين متناقضتين لبيان أياً منهما تكون صحيحة. ونحن نفترض سلفاً أن إحدى الفرضيتين هي الصحيحة وبإجراء

الاختبار عليها نستطيع إما رفضها لصالح الفرضية الأخرى (البديلة) أو التأكد من صحتها. ونحن لدينا فرضيتان:

- الفرضية الصفرية أو الإبتدائية (Null Hypothesis) ويرمز لها عادة بالرمز H_0، هي الفرضية الأساسية التي نرغب في اختبارها ويفترض صحتها.
- الفرضية البديلة (Alternative Hypothesis) ويرمز لها بالرمز H_1.وهي الفرضية التي تمثل البديل في حالة رفضت فرضية العدم (H_0) مع العلم أن الفرض الإحصائي قد يكون بسيطاً أي يحدد توزيع المتغير بشكل كامل، وقد يكون مركباً أي لايحدد توزيع المتغير بشكل كامل.

ولإجراء اختبار على H_0 مقابل H_1، فإن الفرضية H_0 سترفض لصالح H_1 فقط، إذا توفر دليل قاطع من خلال العينة على أن H_0 خاطئة، وإذا لم يتوفر ذلك الدليل فلا يمكن رفضها. ويمكن القول أن النتيجتين المحتملتين هما إما رفض H_0 أو قبولها.

ولبيان صحة أو خطأ فرضية ما، يمكن اختيار عينة عشوائية من المجتمع المراد دراسته وإجراء الاختبار عليها وبالتالي سيكون لدينا إحدى الحالات التالية :

(1) H_0 صحيحة ويتم قبولها.

(2) H_0 صحيحة ويتم رفضها

(3) H_0 خاطئة ويتم قبولها

(4) H_0 خاطئة ويتم رفضها

في اختبار الفرضيات ونتيجة لاستخدام معلومات عينة بدلاً من المجتمع كله، فإن الإستقراءات تكون عرضة للوقوع في نوعين من الخطأ هما:

- يسمى الخطأ الناتج عن رفض H_0 عندما تكون صحيحة، خطأ من النوع الأول (Type I Error) .
- يسمى الخطأ الناتج عن قبول H_0 عندما تكون خاطئة، خطأ من النوع الثاني
(Type II Error)

علماً أنه لاتوجد أي طريقة معلومة تضمن عدم الوقوع في أي من الخطأين.
ويمكن إعادة صياغة ما سبق ذكره في شكل جدول كما هو مبين :

نوع القرار / طبيعة الفرضية	قبول H_0	رفض H_0
H_0 صحيحة	قرار صائب	خطأ من النوع الأول Type I error
H_0 خاطئة	خطأ من النوع الثاني Type II error	قرار صائب

ويسمى احتمال الوقوع في خطأ من النوع الأول (يرمز له عادة بالرمز α) بمستوى دلالة الاختبار أو التجربة (Significance Level).
ويرمز لاحتمال وقوع خطأ من النوع الثاني بالرمز β.

3-5- اختبار الفرضيات حول متوسط المجتمع ووسيط التوزيع الثنائي (النسبة المئوية):

تباين المجتمع σ^2 معلوم :

في هذه الحالة فإن قيمة الاختبار Z تخضع للتوزيع الطبيعي المعياري وتعطى بالعلاقة التالية:

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

حيث أن $\bar{x}$ هو متوسط العينة العشوائية المأخوذة من المجتمع قيد الدراسة وأن μ_0 متوسط افتراضي (أو ابتدائي) للمجتمع المراد اختبار متوسطه الحقيقي μ وأن σ^2 تباين هذا المجتمع كما أن $\sigma/\sqrt{n}$ هي الانحراف المعياري لمتوسط العينة $\bar{x}$ وتسمى الخطأ المعياري للمتوسط ويرمز له بالرمز $\sigma_{\bar{x}}$، وأن الفرضيات المراد اختبارها من أجل متوسط المجتمع μ يمكن أن تأتي في ثلاث إمكانات، إما أن μ أكبر أو أصغر أو لايساوي القيمة الابتدائية μ_0. ويمكن تلخيص اختبار الفرضيات حول متوسط مجتمع في الجدول التالي :

المنطقة التي ترفض فيها H_0 منطقة الرفض أو المنطقة الحرجة	الفرضية البديلة H_1	الفرضية الصفرية H_0
$Z \geq Z_\alpha$ أو $\bar{x} \geq \mu_0 + Z_\alpha \frac{\sigma}{\sqrt{n}}$	$\mu > \mu_0$	$\mu = \mu_0$
$Z \leq -Z_\alpha$ أو $\bar{x} \leq \mu_0 - Z_\alpha \frac{\sigma}{\sqrt{n}}$	$\mu < \mu_0$	$\mu = \mu_0$
$\lvert Z \rvert \leq Z_{\alpha/2}$ أو $\lvert \bar{x} \pm \mu_0 \rvert \geq Z_\alpha \frac{\sigma}{\sqrt{n}}$	$\mu \neq \mu_0$	$\mu = \mu_0$

حيث أن $Z \geq Z_\alpha$ و $Z \leq -Z_\alpha$ و $|Z| \leq Z_{\alpha/2}$ هي على الترتيب اختبار ذي طرف أيمن وأيسر وذي جهتين .

مثال (1) :

بينت الدراسات التي أجرتها وزارة الصحة في إحدى الدول أن الكمية الموصى بها طبياً ويومياً من الزنك في غذاء الأفراد الذين أعمارهم أكثر من (50) سنة هي $\mu=15$ ملغ/يوم وبانحراف معياري $\sigma=6.43$ وللتأكد من ذلك سحبت عينة من الأفراد حجمها (115) فرداً تراوحت أعمارهم بين (65-74) سنة، فوجد أن ($\bar{x}=11.3$) وبانحراف معياري (8) ملغ/يوم **والمطلوب:**

هل تبين هذه البيانات أن المتوسط اليومي μ من الزنك لهؤلاء المسنين هو أقل مما هو موصى به.

الحل :

في هذا المثال كون النقصان في كمية الزنك هو موضع الشك بالنسبة للكمية الموصى بها فأن الفرضية الواجب اختبارها في مثل هذه الحالة هي :

$$H_0 : \mu = 15$$
$$H_1 : \mu < 15$$

وسنرفض الفرضية الصفرية إذا توفر لنا دليل قوي أن $\mu < 15$.

نلاحظ هنا أولاً أن $\bar{x} = 11.3$ هي بالفعل قيمة أصغر من 15، ولكن هذا الفرق قد يحدث بسبب تغاير المعاينة، لكن السؤال، هل الفرق بين 11.3 و 15 كبيراً إلى الدرجة التي يتوجب معرفة السبب الحقيقي وراء ذلك؟

للإجابة على هذا السؤال نقوم بإجراء الاختبار :

ـ صياغة الفرضية :

$$H_0 : \mu = 15$$
$$H_1 : \mu < 15$$

ـ الاختبار الإحصائي :

$$Z = \frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}} \sim N(0,1)$$
$$= \frac{|11.3 - 15|}{6.43/\sqrt{115}}$$
$$= \frac{|-3.7|}{0.6} = |-6.16|$$

ـ القرار:

نلاحظ أن القيمة تقع في أسفل ذيل منحنى Z وتعني أن $\bar{x}$ تبعد بأكثر من 6 انحرافات معيارية عن الوسط الحقيقي μ . هذه النتيجة تعني أن العينة قد أظهرت دليلاً مقنعاً جداً ضد H_0 بالتالي نرفضها لصالح الفرضية البديلة. أي أن الزنك المستهلك في غذاء المسنين في الفترة العمرية 74-65 لا ترقى إلى المعدل المطلوب وهو 15 ملغ / اليوم. ولنفرض أن الدراسة قد أخذت بعين الإعتبار قياس معدل فيتامين E في غذاء نفس المجموعة من الأفراد المسنين، وأن الكمية الموصى بها من هذا الفيتامين هي 10ملغ/ اليوم وأن σ=8.58 أخذت عينة حجمها (115) فرد . أظهرت النتائج أن : $\bar{x} = 9.4$ وانحراف معياري S=6 ملغ/يوم

المطلوب:

هل تبين هذه البيانات أن المتوسط اليومي (μ) من الفيتامين (E) لهؤلاء المسنين هو أقل من 10 ملغ/اليوم.

الحل :

ـ صياغة الفرضية :

$$H_0 : \mu = 10$$
$$H_1 : \mu < 10$$

ـ الاختبار الإحصائي:

$$Z = \frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}} = \frac{|9.4 - 10|}{8.58/\sqrt{115}} = 0.75$$

ـــ القرار :

وهذا يعني أن $\bar{x}$ لاتبتعد سوى 0.75 انحرافاً معيارياً عن الوسط الحقيقي μ الذي يجعل H_0 صحيحة. وبالتالي يمكن القول أنه لايوجد دليل قوي على أن معدل الاستهلاك الفيتامين E بين مجتمع المسنين في هذه الفئة العمرية أقل مما هو موصى به .

كما لاحظنا في المثال السابق كانت القيمتان $Z = 0.75$, $Z = 6.17$ مختلفتين إلى الحد الذي كان فيه قرار رفض أو قبول H_0 واضح المعالم تماماً. لكن في حالات أخرى قد لا يكون القرار واضحاً. لاسيما إذا أدخلنا مفهوم الاحتمال (درجة الثقة أو مفهوم مستوى المعنوية (α) لمعرفة وقوع القيمة (Z) ضمن منطقة الرفض أو منطقة القبول.

5-4- منطقة الرفض، منطقة القبول، المنطقة الحرجة :

تعريف : تسمى مجموعة قيم الاختبار التي تقود إلى رفض الفرضية الصفرية بمنطقة الرفض، وتسمى مجموعة قيم الاختبار التي تقود إلى قبول الفرضية الصفرية بمنطقة القبول. وتسمى قيم الاختبار التي تفصل منطقة الرفض عن منطقة القبول بالقيم الحرجة. وتحدد القيم الحرجة عن طريق اختبار مستوى الثقة (أو الدلالة α) وتحديد القيم الجدولية من جدول التوزيع الطبيعي المعياري التي تجعل المساحة في منطقة الرفض. انظر الشكل (1، 5).

ونحن نميز ثلاثة خيارات في عملية اختبار الفرضية البديلة هي :

- **اختبار ذو طرفين : (Two-Tailed Test)**

فإذا كان الاهتمام ينحصر على تقرير، ما إذا كان متوسط مجتمع μ يختلف عن قيمة محددة ولتكن μ_0 فإن الفرضية البديلة يجب أن تكون:

$$H_1 = \mu \neq \mu_0$$

اختبار فرضية لفرضية بديلة من هذا النوع يسمى باختبار ذي طرفين.

- **اختبار طرف أيسر (left - Tailed)**

إذا كان الاهتمام ينحصر على تقرير ما إذا كان متوسط مجتمع μ أقل من قيمة محددة ولتكن μ_0 فإن الفرضية البديلة يجب أن تكون :

$$H_1 : \mu < \mu_0$$

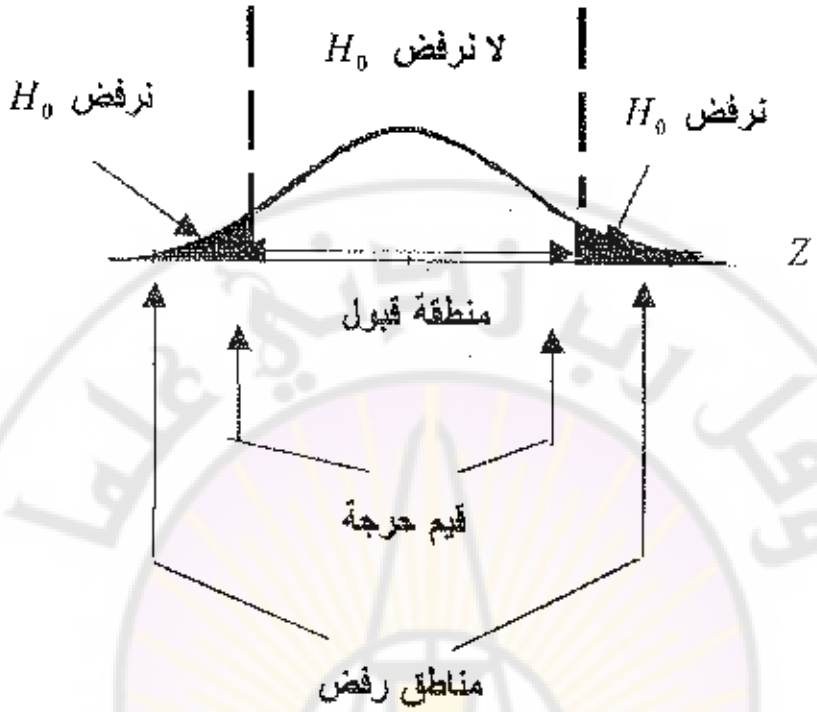
وفي هذه الحالة يسمى باختبار ذي طرف أيسر .

- **اختبار ذو طرف أيمن (Right - tailed test) :**

إذا كان الاهتمام ينحصر على تقرير، ما إذا كان متوسط مجتمع μ أكبر من قيمة ولتكن μ_0 فإن الفرضية البديلة يجب أن تكون : $H_1 : \mu > \mu_0$

وسواء كان الاختبار ذو طرف أيمن أو أيسر فإن الاختبار يسمى ذو طرف واحد (One Tailed Test). **وبناءً على ذلك فإن :**

- الفرضية الصفرية في اختبار ذي الطرفين سترفض كلما كانت قيمة الاختبار صغيرة جداً أو كبيرة جداً، أي أن منطقة الرفض في مثل هذا الاختبار ستحوي جزأين الأول إلى اليسار والأخر إلى اليمين .
- الفرضية الصفرية في اختبار طرف أيسر سترفض فقط إذا كانت قيمة الاختبار صغيرة جداً، أي أن منطقة الرفض في مثل هذا الاختبار ستحوي جزءاً واحداً ويجب أن يكون إلى اليسار.
- الفرضية الصفرية في اختبار طرف أيمن سترفض فقط إذا كانت قيمة إحصاءه الاختبار كبيرة جداً، أي أن منطقة الرفض في مثل هذا الاختبار ستحوي جزأين الأول إلى اليسار والأخر إلى اليمين .



شكل (1، 5)

ويمكن أن نلخص خطوات إجراء إختبار الفرضيات :

1 ــ تحديد الفرضية (H_0) :

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

أو

$$H_1 : \mu > \mu_0$$

2 ــ **تحديد مستوى المعنوية (الدلالة)** (α) أو درجة الثقة (إحتمال الدقة) ($1-\alpha$).

3 ــ **تحديد قيمة** (Z_α أو $Z_{\alpha/2}$) النظرية المقابلة لمستوى المعنوية المتخذ من جداول (Z).

ـ ($Z_{\alpha/2}$) تعني أن مستوى المعنوية موزع على طرفي التوزيع الطبيعي.

ـ (Z_{α}) تعني أن مستوى المعنوية يقع بالكامل على طرف واحد من التوزيع.

4 ـ **تحديد قيمة الاختبار (Z) :**

ـ (Z) عبارة عن الفرق بين متوسط العينة ومتوسط المجتمع معبراً عنه بوحدات من الخطأ المعياري :

$$Z = \frac{|\overline{x} - \mu|}{\sigma / \sqrt{n}}$$

5 ـ **بالمقارنة واتخاذ القرار** : من خلال المقارنة نجد أن :

ـ إذا كانت قيمة (الجدولية) Z_{α}، $Z < Z_{\alpha/2}$ (المحسوبة) فإننا نقبل فرضية العدم (H_0)

ـ إذا كانت قيمة (الجدولية) Z_{α}، $Z > Z_{\alpha/2}$ (المحسوبة) فإننا نرفض فرضية العدم (H_1)

مثال (2):

رغب أحد المستثمرين إنشاء مجمع سكني على هيئة شقق مفروشة فاخرة للإيجار، وأبلغته إحدى الشركات المتخصصة أن متوسط الإيجار اليومي لمثل هذه الشقق يبلغ 600 وحدة نقدية وقبل أن يأخذ قراره، أراد التأكد من صحة معلومات تلك الشركة، فقام بسحب عينة عشوائية حجمها (36) شقة مماثلة لتلك الشقق فكان متوسط الإيجار اليومي (620) وحدة وبانحراف معياري (85) وحدة نقدية.

المطلوب: هل معلومات الشركة المتخصصة صحيحة، عند مستوى معنوية (α=0.05)؟

ـ **صياغة الفرضية** : (معلومات الشركة صحيحة) $H_0 : \mu = 587$

(معلومات الشركة غير صحيحة) $H_1 : \mu \neq 587$

اختبار ذو طرفين وبالتالي فالقيم الحرجة هي:

$$\pm Z_{\alpha/2} = \pm 1.96$$

ـ الاختبار الاحصائي :

$$Z = \frac{|600 - 620|}{85/\sqrt{36}} =$$

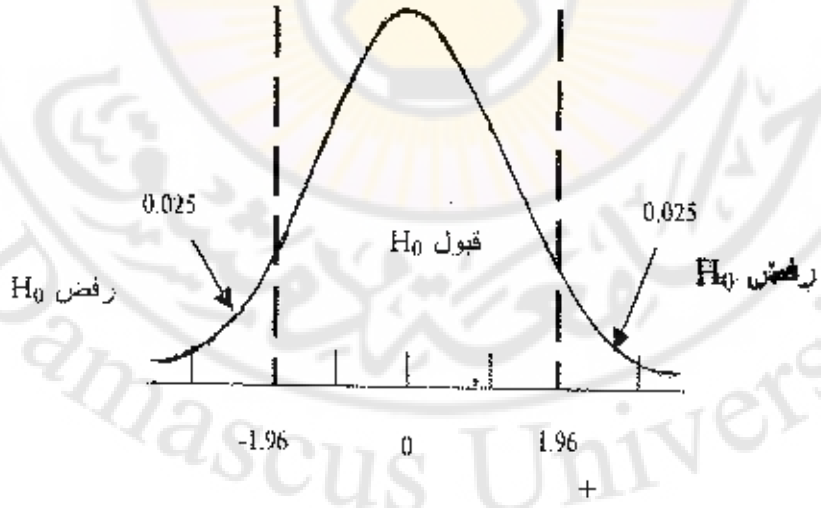
$$\frac{|20|}{85/6} = |-1.4|$$

ـ اتخاذ القرار : من خلال المقارنة نجد أن :

$$Z < Z_{\alpha/2}$$

$$1.41 < 1.96$$

وبالتالي فإن H_0 مقبولة، أي أن بيانات إيجارات الشقق التي حصل عليها المستثمر أكدت صحة ما ادعته الشركة المتخصصة.



5.5- اختبار الفرضيات لنسبة مجتمع :

اختبار الفرضيات للنسبة في حالة العينات الكبيرة يتبع نفس الخطوات التي تمت مع متوسط المجتمع تماماً مع الأخذ بعين الاعتبار فروق المسميات والرموز التي ناقشناها سابقاً. وفيما يلي ملخصاً لخطوات إجراء مثل هذا الاختبار .

ـ صياغة الفرضية :

$$H_0 : P = P_0$$
$$H_1 : P \neq P_0$$

أو

$$H_1 : P > P_0$$
$$H_1 : P < P_0$$

الاختبار الإحصائي :

$$Z = \frac{|\hat{P} - P_0|}{\sqrt{P_0(1-P_0)/n}}$$

القرار : نرفض H_0 إذا وقعت قيمة إحصاءه الاختبار في منطقة الرفض ولانرفض فيما عدا ذلك.

مثال (3) :

تنافس مرشحان الأول (x) والثاني (y) على مقعد مجلس الشعب في إحدى الدوائر الإنتخابية في مدينة دمشق قبل إجراء الإنتخابات بيوم واحد وقد أراد فريق عمل المرشح (x) معرفة حجم شعبية مرشحهم وفيما إذا كان سيفوز أم لا. أخذت عينة عشوائية من المسجلين في الدائرة الذين يحق لهم التصويت حجمها 3100 فرداً أجاب 1650 منهم بأنهم يعتزمون التصويت لصالح هذا المرشح (x). بمستوى معنوية 5%، **المطلوب :** هل هذه المعلومات كافية لحصول المرشح على غالبية الأصوات؟.

الحل :

قبل أن نتم الحسابات لابد لنا من التأكد أن شروط إجراء الاختبار متوفرة. فنجد:

$$nP_0 = 0.5 * 3100 = 1550$$
$$n(1-P_0) = 3100\ (1-0.5) = 1550$$

وكلاهما أكبر من 0.5 وتصبح بناءاً على ذلك خطوات الاختبار قابلة للتطبيق:

الحل :

ـ صياغة الفرضية:

$H_0 : P = 0.5$ (لن يحصل المرشح (x) على غالبية الأصوات)

$H_0 : P > 0.5$ (سيحصل المرشح (x) على غالبية الأصوات)

$$Z_\alpha = Z_{0.05} = 1.645 \Leftarrow \alpha = 0.05$$

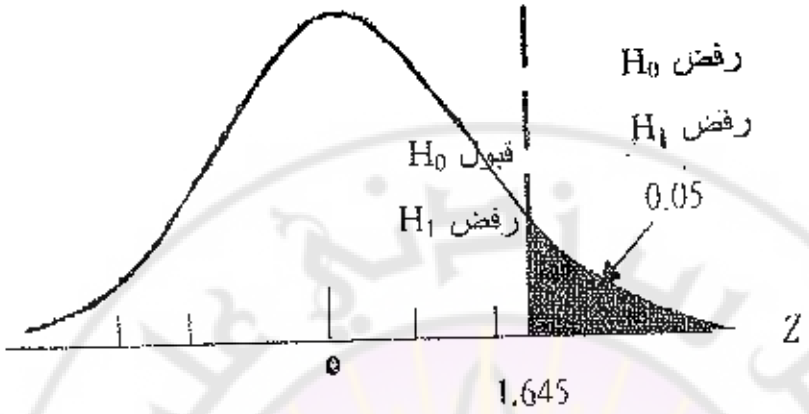
$$\hat{P} = \frac{x}{n} = \frac{1650}{3100} = 0.53$$

ـ الاختبار الإحصائي :

$$Z = \frac{|\hat{P} - P_0|}{\sqrt{P_0(1-P_0)/n}}$$
$$= \frac{|0.53 - 0.5|}{\sqrt{\frac{(0.5)(0.5)}{3100}}} = 3.75$$

ـ اتخاذ القرار : من خلال المقارنة نجد أن: $Z > Z_\alpha$ $(3.75 > 1.645)$

وبالتالي H_0 مرفوضة مما يعطي أن الاستبيان الذي أجراه فريق عمل المرشح (x) دليل على أن مرشحهم سيحصل على غالبية الأصوات.



6-5- اختبار الفرضيات للفرق بين نسبتين مئويتين (P_1 ، P_2):

كما هو الحال في اختبار الفرضيات فإننا نستخدم المتغير العشوائي الطبيعي المعياري لمقارنة قيمة إحصاء العينة مع القيمة الافتراضية لمعلمة المجتمع، ولاختبار الفرضيات المتعلقة بنسب مجتمعين، يأخذ Z الصورة التالية :

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - d_0}{\sigma_{\hat{P}_1 - \hat{P}_2}}$$

حيث أن $d_0 = P_1 - P_2$ وأن $\sigma_{\hat{P}_1 - \hat{P}_2}$ هو الانحراف المعياري للفروق النسب المئوية $(\hat{P}_1 - \hat{P}_2)$، بناءً عليه نستخدم العلاقة :

$$S_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{P}_1)}{n_1} - \frac{\hat{p}_2(1-\hat{P}_2)}{n_2}}$$

لتقدير $\sigma_{\hat{P}_1 - \hat{P}_2}$، ولكن عندما تكون الفرضية الصفرية لاختبار فرضية حول الفرق بين نسب مجتمعين هي :

$$\sigma_{\hat{p}_1-\hat{p}_2}=\sqrt{\frac{P(1-P)}{n_1}+\frac{P(1-P)}{n_2}}$$

$$=\sqrt{P(1-P)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

أي أننا نحتاج فقط إلى تقدير واحد للقيمة P لتقدير $\sigma_{\hat{p}_1-\hat{p}_2}$، ويمكن حسابه مباشرة من العلاقة:

النسبة أو الاحتمال التجميعي :

$$\hat{P}=\frac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}$$

ولتقدير الانحراف المعياري للفرق $\hat{p}_1-\hat{p}_2$ فإننا نستخدم بدلاً من $\sigma_{\hat{p}_1-\hat{p}_2}$ العلاقة التالية:

$$S_{\hat{p}_1-\hat{p}_2}=\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

مثال (4): في دراسة لنسبة الناخبين المؤيدين للمرشح (x) في إحدى المدن، سحبت عينتين عشوائيتين، عينة عشوائية في المنطقة (A) أظهرت أن (60) من (400) ناخب يؤيدون المرشح (x) وعينة عشوائية في المنطقة (B) أظهرت أن (40) من (300) ناخب يؤيدون هذا المرشح **والمطلوب** : اختبر فرضية ليس هناك اختلاف بين نسبة الذين يؤيدون هذا المرشح في المنطقتين عند مستوى معنوية ($\alpha = 0.05$) .

الحل :

P_1 = نسبة الناخبين المؤيدين للمرشح (x) في المنطقة (A)

P_2 = نسبة الناخبين المؤيدين للمرشح (x) في المنطقة (B).

ـ صياغة الفرضية :

$$H_0 : P_1 - P_2 = 0 \text{ أو } P_1 = P_2$$

$$H_1 : P_1 - P_2 \neq 0 \text{ أو } P_1 \neq P_2$$

عند $\alpha = 0.05$ فإن $Z_{\alpha/2} = 1.96$ والنسبة والاحتمال للمنطقتين (العينتين):

ـ للمنطقة A : $\hat{P}_1 = \frac{60}{400} = 0.15$

ـ للمنطقة B $\hat{P}_2 = \frac{40}{300} = 0.13$

فيكون لدينا، النسبة أو الاحتمال التجميعي هو :

$$\hat{P} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} = \frac{400(0.15) + 300(0.13)}{400 + 300} = 0.14$$

الاختبار الاحصائي :

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}} = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\hat{P}(1 - \hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\frac{(0.15 - 0.13) - (0)}{\sqrt{(0.14)(1 - 0.14)\left(\frac{1}{400} + \frac{1}{300}\right)}} = 0.76$$

القرار : وبالتالي نقبل H_0 لأن $0.76 < 1.96$ بالتالي ليس هناك فرقاً معنوياً بين نسب المؤيدين للمرشح (x) في المنطقتين (A) و (B) .

7-5- اختبار الفرضيات للفرق بين متوسطي مجتمعين(μ_1، μ_2) :

يعتمد اختبار الفرضيات في هذه الحالة على قيم توابع العينة ($\bar{x}_1 - \bar{x}_2$) بدلاً من $\bar{x}$ في حالة عينة المجتمع الواحد، وبناءً على ذلك فإننا نستخدم معاينة الفرق $\bar{x}_1 - \bar{x}_2$ لإجراء اختبار الفرضية. وقد علمنا أن توزيع الفرق يمكن تقريبه بتوزيع احتمالي طبيعي، إذا كان حجم العينة كبيراً، وبذلك تكون القيمة Z متغيراً عشوائياً يتبع التوزيع الطبيعي المعياري وهي الاختبار المطلوب لإجراء اختبار الفرضيات في مثل هذه الحالة .

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (d_0)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_1}}}$$

حيث أن: $d_0 = \mu_1 - \mu_2$ وأن: $\bar{x}_1 - \bar{x}_2$ القيمة الملاحظة لـ $(\mu_1 - \mu_2)$ للفروق بين الوسطين الحقيقيين، هي :

$$\sigma_{\bar{x}1-\bar{x}2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$\sigma_{\bar{x}1-\bar{x}2}$ هو الخطأ المعياري للفرق $\bar{x}_1 - \bar{x}_2$ وإذا كانت σ_1^2, σ_2^2 غير معلومتين، فإنه بالإمكان الاستعاضة عنهما بتباينات العينات S_1^2، S_2^2 مادام أن الأحجام لهذه العينات كبيرة، أي $n_1, n_2 \geq 30$.

مثال (5) :

في دراسة حول متوسط عمر المرأة عند الزواج الأول في إحدى الدول، سحبت من مدينتين عينتين عشوائيتين فكانت النتائج في الجدول المرفق .

المؤشر / العينة	العينة الأولى	العينة الثانية
حجم العينة	$n_1=248$	$n_2 = 225$
متوسط عمر المرأة عند الزواج الأول	$\bar{x}_1 = 24$	$\bar{x}_2 = 26$
الانحراف المعياري حول عمر المرأة عند الزواج (بالسنوات)	$S_1 = 3$	$S_2 = 4$

المطلوب :

باستخدام هذه المعلومات هل هناك فرقاً بين متوسط أعمار البناء المتزوجات في هاتين المدينتين؟.

صياغة الفرضية :

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 \neq \mu_2 \neq 0$$

باستخدام مستوى دلالة $\alpha = 0.05$ واختبار ذي طرفين، فإن قيم Z الحرجة كما نعلم هي $Z_{\alpha/2} = \pm 1.96$

ـ الاختبار الإحصائي :

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (d_0)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_1}}} = \frac{(24-26)-0}{\sqrt{\frac{(3)2}{248} + \frac{(4)2}{225}}} = |6.45|$$

القرار : وبالتالي نرفض H_0 لأن $6.45 > 1.96$ ونستنتج أن هناك فرقاً معنوياً بين متوسطي أعمار النساء المتزوجات في هاتين المدينتين.

تمهيد :

عند اختيار أسلوب العينات لإجراء الدراسة الإحصائية فإن السؤال الذي يتبادر إلى الذهن، ماهو حجم العينة اللازم لإجراء الدراسة ؟

وبالرغم من أن هذا السؤال لاتوجد له إجابة قاطعة، إلا أنه وكما نعلم أن حجم العينة يتناسب عكساً مع خطأ المعاينة، وكلما كبرنا حجم العينة، كلما قللنا من مقدار الخطأ الاحتمالي أو خطأ المعاينة، لكننا بنفس الوقت لا نريد أن نزيد من حجم العينة إلى الحد الذي نفقد معه مزايا استخدام أسلوب العينات في الدراسة. وبعبارة أخرى أنه إذا كبرنا حجم العينة إلى الحد الذي نصل بها إلى حجم المجتمع، نكون قد ضحينا بأسلوب العينات، وإذا صغرنا من حجم العينة لن تكون صادقة في تمثيلها للمجتمع المسحوبة منه.

5-9- ـــ تحديد حجم العينة :

بشكل عام عند تحديد حجم العينة لابد من مراعاة مايلي:

1 ـــ الاعتبارات المادية.

2 ـــ الاعتبارات العلمية.

بالنسبة للاعتبارات المادية : فإن حجم العينة يتحدد في ضوء الميزانية المتاحة للدراسة من كوادر بشرية، ومالية، ووقت، وفي ضوء هذا الاعتبار يمكن التحكم في تحديد حجم العينة، الأمر الذي يقودنا إلى تصغير أخطاء المعاينة.

أما بالنسبة للاعتبارات العلمية : إن العلمية في تحديد حجم العينة تتوقف على مدى الدقة المطلوب توافرها في إحصاءات العينة أولاً، وتشتت القيم في المجتمع ثانياً، فمثلاً بالنسبة للدقة، حين نريد أن نقدّر متوسط دخل الفرد من عينة، وكان المطلوب تحديد هذا المتوسط تحديداً دقيقاً كان لزاماً علينا اختيار

عينة كبيرة، وإذا كان المطلوب تحديد هذا المتوسط بالتقريب، فإن عينة صغيرة تكون كافية وتفي بالغرض في هذه الحالة. أما بالنسبة للتشتت، فإذا كان المجتمع الأصلي الذي سحبت منه العينة متبايناً تبايناً شديداً فإن حجم العينة لابد أن يكون كبيراً، وإذا كان الانحراف المعياري للمجتمع صغيراً، فإن عينة صغيرة تكون كافية وتفي بالغرض.

إذن نلاحظ أن الدقة معياراً أساسياً في تحديد حجم العينة، ويجب أن نحدد أولاً مدى الدقة المطلوبة، أو بعبارة أخرى ماهو حجم الخطأ الذي نستطيع أن نقبله، في تقدير معالم المجتمع باستخدام إحصاءات العينة، وبالإضافة إلى ذلك لابد من معرفة الانحراف المعياري للمجتمع، وهو أمر يتعذر علينا معرفته في أغلب الأحيان، لذلك نعتبر الانحراف المعياري للعينة تقديراً غير متحيزاً للانحراف المعياري للمجتمع.

من خلال الاستعراض السابق نستنتج أنه لتحديد حجم العينة نتبع الخطوات التالية:

1 ـ نحدد الحد الأقصى للخطأ المسموح به، ولنفرض أنه مساوٍ (E)، أي أننا نأمل في الحصول على عينة يكون الفرق بين وسطها الحسابي والوسط الحسابي للمجتمع، محل الدراسة، في حدود مقدار معين من الخطأ هو (E).

فإذا أردنا تحديد نسبة البطالة في بلد ما، وكنا نرغب في أن لاتزيد قيمة الحد الأقصى للخطأ المسموح عن (3%)، وكانت النسبة الحقيقية للبطالة في المجتمع مساوية (20%)، فإن النسبة المئوية للبطالة في العينة يجب أن لاتقل عن (17%) وأن لاتزيد عن (23%).

إن الخطأ الأعظمي المرتكب هو :

$$E = Z_{\alpha/2}\sqrt{\frac{pq}{n}}$$

وذلك باحتمال قدره :

$$(1-\alpha)\%$$

ومن العلاقة السابقة نجد : $E^2 = Z^2_{\alpha/2}\frac{\overline{p}\overline{q}}{n}$

وبالإصلاح فإن :

$$n = \frac{Z^2_{\alpha/2}\overline{p}\overline{q}}{E^2}$$

وتفسير هذه العلاقة أننا واثقون باحتمال $(1-\alpha)\%$ أن الخطأ المرتكب لن يزيد على القيمة (E) إذا كان حجم العينة :

$$n = \frac{Z^2_{\alpha/2}\overline{p}\overline{q}}{E^2}$$

وبما أن $\overline{P}(1-\overline{P})$ غير معلومة فإننا نعتبر أن $\overline{q} = 1/2$ ، $\overline{p} = 1/2$

وإن $\overline{P}(1-\overline{P})$

يأخذ قيمته العظمى عندما $\overline{p} = 1/2$.

إذاً يمكن القول أن الخطأ المرتكب لن يتجاوز (E) إذا كان حجم العينة :

$$n = \frac{Z^2_{\alpha/2}}{4(E^2)}$$

تمرين (6) :

في استفتاء سريع حول أحد المرشحين وجد أن (960) من أصل (1500) تم اختيارهم عشوائياً، كانوا يؤيدون ذلك المرشح .

والمطلوب :

1 ــ أوجد مجال ثقة 98% للنسبة الحقيقية لمؤيدي هذا المرشح .

2 ــ أوجد الخطأ الأعظمي المرتكب في قولنا أن نسبة المؤيدين الحقيقية تساوي نسبة المؤيدين في هذه العينة بثقة 98% .

3 ــ ماهو حجم العينة الواجب سحبها حتى لا يتجاوز الخطأ 0.02 زيادة أو نقصاناً وذلك بثقة 98% .

الحل :

$$\bar{p} = x/n = \frac{960}{1500} = 0.64$$

من جدول (Z) نجد أن:

$$Z_{\alpha/2} = Z_{0.49} = 2.33$$

1 ــ

$$\bar{p} - Z_{\alpha/2}\sqrt{\frac{\bar{p}.\bar{q}}{n}} \le p \le \bar{p} + Z_{\alpha/2}\sqrt{\frac{\bar{p}.\bar{q}}{n}}$$

$$0.64 - 2.33\sqrt{\frac{0.64(0.36)}{1500}} \le P \le 0.64 + 2.33\sqrt{\frac{0.64(0.36)}{1500}}$$

$$(0.61, 0.67)$$

يمكن القول أنه باحتمال 98% لن تقل نسبة المؤيدين لهذا المرشح في المجتمع الذي سحبت منه العينة عن 61% ولن تزيد عن 67% .

2 ــ نحن نعلم أن $\overline{p} = 0.64$:

$$E = Z_{\alpha/2}\sqrt{\frac{\overline{P}(1-\overline{P})}{n}}$$

$$E = 2.33\sqrt{\frac{0.64(0.36)}{1500}} = 0.03$$

3 ــ حتى لايتجاوز الخطأ 0.02 زيادة أو نقصاناً بثقة 98% فإن حجم العينة المطلوب هو:

$$n = \frac{Z^2_{\alpha/2}}{4E^2} = \frac{(2.33)^2}{4(0.02)^2} = 3393$$

وفي حال تقدير متوسط المجتمع (μ) من متوسط العينة $\overline{X}$ فإن الخطأ الأعظمي المرتكب، يعطى بالعلاقة :

$$E = Z_{\alpha/2}\frac{\delta}{\sqrt{n}}$$

وذلك باحتمال : %(1-α) .

وبالتالي فإن حجم العينة الواجب سحبها حتى لا يتجاوز الخطأ في التقدير مقداراً (E) وذلك باحتمال %(1-α) هو :

$$n = \frac{Z^2_{\alpha/2}\delta^2}{E^2}$$

تمرين (7) :

في عينة من 36 طالباً كان متوسط النقاط التي حازها كل منهم وانحرافها المعياري S = 0.3 **والمطلوب :**

1 ــ أوجد 99% حد الثقة لمتوسط المجتمع .

2 ــ كم يجب أن يكون حجم العينة الواجب سحبها عشوائياً بحيث لايتجــاوز الخطأ في تقدير الوسط الحسابي للمجتمع 0.05 وذلك باحتمال 99% من الثقة.

الحل :

بما أن $n > 30$ يمكن اعتبار : $S = \delta = 0.3$

ومن جدول (Z) نجد أن :

$$Z_{\alpha/2} = Z_{0.945} = 2.58$$

ويكون مجال الثقة هو :

$$\bar{x} - Z_{0.945}\frac{S}{\sqrt{n}} \le \mu \le \bar{x} + Z_{0.945}\frac{S}{\sqrt{n}}$$

$$2.6 - 2.58\frac{0.3}{\sqrt{36}} \le \mu \le 2.6 + 2.58\frac{0.3}{\sqrt{36}}$$

$$2.47 \le \mu \le 2.73$$

والتفسير أنه باحتمال 99% فإن متوسط النقاط في مجتمع الطلبة (μ) الذي سحبت منه العينة لن يقل عن 2.47 ولن يزيد عن 2.73 .

$$n = \left[\frac{(2.58)(0.3)}{0.05}\right]^2 = 238.7$$ 2 ــ طالباً

والتفسير :

إننا واثقون باحتمال 99% بأنه لن يختلف μ بأكثر من 0.05 إذا كان حجــم العينة مساو لــ 239 طالباً .

ملاحظة :

في الحالات العملية في حال عدم إعطاء احتمال الدقة (درجة الثقة) فإننا نأخذ أكبر احتمال ممكن وهو (99.73%) ويقابل في جدول (Z) القيمة (3).

ملاحظة :

عند تحديد حجم العينة للنوعيات (النسبة المئوية) وكانت (P) غير معلومة كنا نفترض أن $(\bar{p} = 1/2)$ و $(\bar{q} = 1/2)$ ، لكن ليس هذا هو الحال في أغلب الحالات العملية فقد تتوافر لدينا معلومات عن قيمة تحقق الظاهرة المدروسة في المجتمع وأنها تتراوح بين قيمتين دنيا وعليا مثلاً: $(10 \le P \le 20)$، فعند حساب حجم العينة نأخذ القيمة الأقرب إلى الـ (50%) ومن ثم نقوم بتطبيق العلاقة التالية لحساب حجم العينة:

$$n = \frac{Z^2 . P.q}{E^2}$$

تمرين (8) :

لقد كان من المرغوب به تحديد نسبة المرشحين للانتخابات من حملة الإجازة الجامعية وقد قدرت هذه النسبة بين %(12-22) المطلوب: ماهو حجم العينة الواجب سحبها عشوائياً بغية تقدير النسبة الحقيقية لحملة الإجازة الجامعية لهؤلاء المرشحين على أن لا يزيد الخطأ المرتكب في التقدير عن (0.05) كما أن احتمال الدقة لا يقل عن (99.5%)

الحل :

P = 22%

$$n = \frac{Z^2 . \bar{P}.\bar{q}}{E^2} = \frac{(2)^2 (0.22)(0.78)}{(0.05)} = 275$$ مرشح.

تمارين غير محلولة

ـ تمرين (1) :

رغب أحد الباحثين في تقدير نسبة الأسر التي لايزيد عدد أفرادها عن (3) أشخاص، وكان الاعتقاد بأن نسبة هذه الأسر تتراوح بين (60%) و(70%) **والمطلوب:**

ماهو حجم العينة المراد سحبها عشوائياً لتقدير النسبة الحقيقية للأسر التي لايزيد عدد أفرادها عن (3) أشخاص، وألا يزيد الخطأ عن (4%) وإحتمال الدقة لايقل عن (95.5%)

تمرين (2) :

يحوي صندوقاً عدداً غير معروفاً من الكرات الحمراء والبيضاء، سحبت منه عينه من (60) كرة فوجد أن 70% منها حمراء .

والمطلوب :

1 ـ أوجد النسبة الحقيقية لعدد الكرات الحمراء الموجودة في الصندوق وذلك باحتمال ثقة 90% .

2 ـ أوجد الخطأ الأعظمي المرتكب في قولنا أن النسبة الحقيقية لعدد الكرات الحمراء تساوي نسبة العينة وذلك باحتمال 90% .

3 ـ احسب حجم العينة الواجب سحبها عشوائياً حتى لاتختلف النسبة الحقيقية للكرات الحمراء عن نسبتها في العينة بأكثر من 0.05 وذلك باحتمال 90% .

تمرين (3) :

تدعي مجلة سياسية أن (25%) من قرائها هم طلبة جامعيون. أخذت عينة عشوائية من (200) قارئ لها، ووجد أن (42) فقط هم طلبة جامعيون.

والمطلوب:

1 ــ اختبر الفرض القائل بأن نسبة القراء من الطلبة الجامعيين ومازالت (25%) عند مستوى معنوية $\alpha = 0.05$

2 ــ باستخدام بيانات العينة، أنشئ فترة (95%) ثقة لنسبة القراء من المجتمع الذين هم طلبة جامعيون .

تمرين (4):

ترغب إحدى القنوات التلفزيونية في معرفة، ما إذا كان درجة شعبية برنامجين تلفزيونيين متساوية، سحبت عينة عشوائية من (400) من المشاهدين المتوقعين للبرنامج الأول (A) فوجد أن (220) منهم يشاهدون البرنامج بانتظام كما سحبت عينة أخرى مستقلة من حجم مماثل للمشاهدين المتوقعين للبرنامج الثاني (B) فوجد أن (260) منهم يشاهدون البرنامج بانتظام.

المطلوب: اختبر الفرض القائل بتساوي النسبتين عند مستوى معنوية $(\alpha=0.05)$.

تمرين (5):

يعتقد أن أكثر من (50%) من مشاهدي التلفزيون يشاهدون نشرة الأخبار المسائية في القناة الثانية. سحبت عينة عشوائية من (1600) مشاهد وجد أن (840) منهم يشاهدون نشرة الأخبار المسائية في القناة الثانية.

المطلوب: اختبر عند مستوى معنوية $(\alpha=0.01)$ صحة هذا الاعتقاد.

تمرين (6) :

قبل إجراء انتخابات رئاسة اتحاد الطلاب سحبت عينة عشوائية مكونة من (1600) عضو من أعضاء اتحاد الطلاب الذين لهم حق الانتخاب فوجد أن (50%) يؤيدون أحد المرشحين. والمطلوب :

1 ــ أوجد فترة ثقة (0.95) للنسبة الحقيقية (P) لمؤيدي هذا المرشح من بين من لهم حق الانتخاب .

تمرين (7):

ترغب وزارة العمل في إحدى الدول تقدير متوسط عمر العاطلين عن العمل باستخدام عينة عشوائية بحيث لاتزيد الفرق بين متوسط العينة والمتوسط الحقيقي عن (2) وذلك بدرجة ثقة (0.99).

المطلوب: حدد حجم العينة المناسب إذا علمت أن الانحراف المعياري لأعمار العاطلين عن العمل هو (5) سنوات.

تمرين (8):

يرغب مدير أحد الفنادق الكبرى في تقدير نسبة النزلاء الذين يقضون أكثر من ثلاثة أيام بالفندق بحيث لايزيد الفرق بين نسبة العينة والنسبة الحقيقية عن (0.02) بدرجة ثقة (0.90).

المطلوب: حدد أكبر حجم عينة يحقق رغبة المدير.

تمرين (9):

قررت إدارة شركة تجارية لها فرعين أن تقدر الفرق في متوسط حجم المبيعات للفرعين أخذت عينتين مستقلتين من مبيعات يوم ما في الفرعين ووجدت النتائج الآتية

الفرع الأول	*الفرع الثاني*
$n_1=32$	$n_2=36$
$\overline{X}_1=5000$	$\overline{X}_2=3750$
$s_1=1500$	$s_2=1300$

والمطلوب: أنشئ تقدير للفرق بين متوسط المبيعات لدى الفرعين.

تمرين (10) :

كان متوسط درجات امتحان القبول في إحدى الجامعات (900) درجة وبانحراف معياري يساوي (80). في كل عام يتم أخذ عينة من نتائج امتحان القبول لدراسة ما إذا كان المستوى للمتقدمين للالتحاق قد تغير أم لا؟ في هذا العام تم أخذ عينة نتائج (60) متقدماً وتبين أن متوسط درجاتهم بلغ (924).

المطلوب : هل يمكن القول أن المستوى التعليمي للمتقدمين للامتحان في هذا العام أفضل من الأعوام السابقة، استخدام مستوى معنوية (α=0.05)

تمرين (11) ــ أثناء أزمة مياه في إحدى المدن رأت الشركة المختصة إجراء معاينة عشوائية على عدادات مياه منطقة سكنية بغية مراقبة الاستهلاك اليومي في المياه، في يوم معين أظهرت عينة من (50) عداداً أن متوسط الاستهلاك $\bar{x} = 25$ بانحراف معياري $S=5m^3$ أنشئ تقدير فترة (95%) ثقة لمتوسط استهلاك مجتمع تلك المدينة من المياه.

تمرين (12) :

سحبت عينة عشوائية من (1000) من المثقفين وعينة أخرى من (1000) من غير المثقفين وعقدت مقابلات شخصية لهم لمعرفة رأيهم تجاه موضوع تعدد الزوجات، فوجد أن (450) مثقفاً و (500) من غير المثقفين يؤيدون الفكرة.

المطلوب: هل تدل هذه البيانات على أن نسبة مؤيدي تعدد الزوجات بين المثقفين وغير المثقفين متساوية عند مستوى معنوية (α=0.05).

تمرين (13) :

في عينة عشوائية مكونة من (900) شخص من مشاهدي التلفزيون تبين أن (180) من بينهم يشاهدون برنامجاً تلفزيونياً معيناً. أوجد فترة الثقة (0.99) للنسبة الحقيقية لمشاهدي التلفزيون الذين يشاهدون هذا البرنامج .

تمرين (14) :

إذا علمت أن الانحراف المعياري لدرجات اختبار الذكاء يساوي (16) درجة فما هو حجم العينة الذي يؤدي إلى الحصول على خطأ للتقدير لا يزيد عن (5) درجات بدرجة ثقة (0.95) .

تمرين (15) :

بفرض أن الانحراف المعياري لأوزان الأطفال الحديثي الولادة يساوي (2) كغ. يرغب مدير أحد المشافي تقدير متوسط وزن الأطفال الحديثي الولادة في إحدى المدن. فقام باختيار عينة عشوائية بحيث لايختلف متوسط العينة عن المتوسط الحقيقي بأكثر من (2) كغ بدرجة ثقة (0.95) .

المطلوب : حدد حجم العينة المناسب .

تمرين (16):

إذا كانت نسبة البيوت المؤجرة في إحدى المناطق هي (40%). ماهو حجم العينة الواجب سحبها عشوائياً، إذا علمت أن حدود الثقة (95%) والخطأ للتقدير لايزيد عن (10%).

تمرين (17) :

تدعي عمادة كلية العلوم السياسية في جامعة دمشق أن نسبة الطلاب في الكلية هي (98%) والطالبات (52%)، وللتأكد من ذلك، سحبت عينة عشوائية من (100) طالب وطالبة أظهرت النتائج أن نسبة الطلاب هي (40%).

والمطلوب: هل كان الادعاء صحيحاً، عند مستوى معنوية (α=0.05) .

تمرين (18) :

مصنع للنسيج أدخل أحد المهندسين تحسينات عليه مدعياً تخفيف نسبة التلف إلى (4%)، للتأكد من ذلك أخذت عينة عشوائية من (200) قطعة ظهر (12) قطعة تالفة.

المطلوب: هل كان هذا الإدعاء صحيحاً عند مستوى معنوية (α=0.05).

تمرين (19):

في دراسة لمجموعتين من الطلاب تتكون المجموعة الأولى من (100) طالب والمجموعة الثانية من (80) طالباً، فكان الوسط الحسابي لعلامات المجموعة الأولى (70) درجة بانحراف معياري (10) درجات والوسط الحسابي للمجموعة الثانية (66) درجة بانحراف معياري (8) درجات.

المطلوب: هل هناك إختلافاً معنوياً في أداء الطلاب لهاتين المجموعتين عند مستوى معنوية (α=0.01).

الفصل السادس

نظرية التقدير

واختبار الفروض للعينات الصغيرة

Estimation and Testing of Hypotheses
For Small Samples

- مقدمة
- توزيع t ستودنت
- التقدير المجالي لمعلمات المجتمع في العينات الصغيرة عندما يكون التباين σ^2 غير معلوم.
- تقدير الفرق بين متوسطي مجتمعين غير معلومي التباين.
- اختبارات الفرضيات للعينات الصغيرة (اختبار الفروض للمتوسط ـ حالة مجتمع واحد، اختبار معنوية الفروق للأوساط ـ حالة مجتمعين) .
- توزيع كاي مربع (χ^2)
- التقدير المجالي لمعلمة المجتمع
- اختبار الفروض للتباين σ^2.
- اختبار (F) فيشر.
- اختبار المقارنة بين تباينى مجتمعين .
- تمارين غير محلولة .

6-1- مقدمة :

سبق أن درسنا توزيعات المعاينة للمتوسط وأيضاً للنسب المئوية وللفرق بين متوسطي مجتمعين ... الخ، وبينّا أنه عند تطبيق نظرية النهاية المركزية نحتاج لمتوسط وتباين المجتمع μ, σ^2 . وقد لاحظنا أن توزيع المعاينة لـ (μ) يتبع التوزيع الطبيعي، إذا كان المجتمع المدروس يخضع للتوزيع الطبيعي والانحراف المعياري للمجتمع معلوماً، لكن في كثير من الحالات يمكن حساب تباين المجتمع σ^2، لذا نقبل من أجل عينات $n \geq 30$ تباين S^2 كتقدير جيد للتباين σ^2 وبالتالي نظرية النهاية المركزية محققة هنا وتصبح الدرجة المعيارية (Z) بالشكل:

$$Z = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

أما إذا كانت $n<30$ فإن (Z) في هذه الحالة تتبع توزيعاً جديداً يدعى توزيع (t) ستورنت.

6-2- توزيع (t) ستودنت :

لقد بينّا أنه عندما يكون حجم العينة $n \leq 30$ والإنحراف المعياري للمجتمع المدروس (δ) غير معلوم، في هذه الحالة، فإن البيانات تتبع توزيعاً آخر يدعى توزيع (t ستودنت)، ويمكن القول أن هذا التوزيع يشبه التوزيع الطبيعي من حيث التماثل، ولكنه أكثر انبساطاً منه، وتتقارب نهايتيه من الصفر عند ($+\infty$) ويقترب من التوزيع الطبيعي المعياري، عندما تزداد (n) حجم العينة، بمعنى آخر أن توزيع ستودنت هو عبارة عن مجموعة من المنحنيات المتماثلة أو المتناظرة حول المعلمة، ويتغير شكل هذه المنحنيات تبعاً لتغير عدد درجات الحرية، التي تعرف بأنها عدد المشاهدات المستقلة التي تتكون منها العينة أو حجم العينة ناقصاً عدد المعلمات الاحصائية، التي يجب تقديرها

منهـــــا، ويرمـــــز لهـــــا بـــــالرمز (v)، حيـــــث V = n-1، وعندما تكون درجات الحرية قليلة، فإن المنحنـي يصـبح أكثـر تفلطحاً وتشتتاً بالمقارنة مع التوزيع الطبيعي المعيـاري، ويتطـابق توزيـع ستودنت مع التوزيع الطبيعي المعياري N(0,1) عندما يصبح درجات الحرية $n \geq 30$.

يمكن اعتبار الانحراف المعياري المحسوب من عينة عشـوائية (S) كتقـدير غير متحيز للانحراف المعياري بعد ضرب الانحراف المعياري للعينة بالعامل المصحح $\sqrt{\frac{n}{n-1}}$.

إن أهم تطبيقات توزيع ستودنت هي إمكانية اسـتخدامه كتوزيـع للمعاينـة، للعينات الصغيرة الحجم. وتعطى (t) بالعلاقة التالية:

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

من خلال هذه العلاقة يمكن ملاحظة أن توزيع (t) يشابه التوزيـع الطبيعـي المعياري من حيث كون متوسطه $\mu=0$ إلا أن تباينه لايساوي الصفر .

ـ خصائص التوزيع t:

1 ـ القيمة المتوقعة للدرجة المعيارية (t) هي : $E(t)=0$

2 ـ التباين لتوزيع t عندما $n \geq 30$: $Var(t) = \frac{v}{v-2}$

3 ـ يستخدم توزيع (t) بصورة خاصة عند عدم معلوميـة تبـاين المجتمـع الإحصـائي المدروس والذي سحبت منه العينة العشوائية.

4 ـ تتقارب قيم (t) من بعضها البعض عندما تكون أحجام العينات كبيرة.

5 ـ يعتمد توزيع (t) على شرط التوزيـع الطبيعـي المجتمـع الإحصـائي المدروس، وبخلاف ذلك قد لا تكون نتائجه دقيقة.

3-6- التقدير المجالي لمعلمات المجتمع في العينات صغيرة الحجم (تقدير متوسط المجتمع μ) عندما يكون التباين σ^2 غير معلوم:

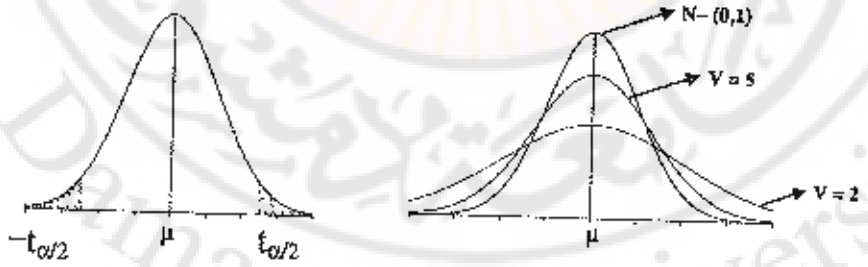
إذا رغبنا في تقدير متوسط مجتمع إحصائي (μ) وكان $n \leq 30$ يمكن كتابة المجال الاحتمالي، لـ (μ) وفق الصيغة التالية :

$$P\left(-t_{\alpha/2} < t < t_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-t_{\alpha/2} < \frac{\bar{x}-\mu}{S/\sqrt{n}} < t_{\alpha/2}\right) = 1-\alpha$$

$$P\left(\bar{x} - t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)\right) = 1-\alpha$$

وبالتالي يمكن القول، بأن (μ) تقع ضمن المجال السابق باحتمال قدره (1-α) وتعطى قيم $t_{\alpha/2}$ من جداول (t) في نهاية الكتاب، حيث تعطى قيم t لكل درجة من درجات الحرية وذلك من أجل مستويات دلالة محدودة .



شكل (1-7)

تمرين (1) :

في إنتخابات نقابة المعلمين في إحدى المدن السورية تم اختيار (16) مرشحاً عشوائياً، فوجد أن متوسط عمر المرشح $\bar{x} = 30$ سنة وبانحراف معياري (5) سنوات.

والمطلوب :

أوجد مجال الثقة لمتوسط عمر المرشحين باحتمال 99% أو قدر باحتمال 99% متوسط عمر المرشحين لنقابة المعلمين في هذه المدينة.

الحل :

$$P\left(\bar{x} - t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)\right) = 1 - \alpha$$

وبالتعويض :

$$P\left(30 - 1.753\left(\frac{5}{\sqrt{16}}\right) < \mu < 30 + 1.753\left(\frac{5}{\sqrt{16}}\right)\right) = 0.99$$

إن قيمة (T) الجدولية من أجل :

$$\alpha = 5\%$$

$$V = 16 - 1 = 15$$

$$T_{\alpha/2} = 1.735$$

وعليه يكون :

$$[\,27.80 < \mu < 32.19]$$

وبالتالي يمكن القول أنه باحتمال 95% إن متوسط عمر المرشح في مجتمع المعلمين لن يقل عن 27.80 سنة ولن يزيد عن 32.19 سنة.

تمرين (2) :

لقياس السرعة القصوى لسيارة جديدة، أجرينا (16) تجربة فحصلنا على متوسط للسرعة القصوى يساوي $x = 200$ كم / ساعة وانحراف معياري (10) كم / ساعة .

والمطلوب :

إيجاد مجال الثقة لمتوسط السرعة القصوى لهذا النوع الجديد من السيارات باحتمال 99%.

الحل :

$$P\left(\bar{x} - t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)\right) = 1 - \alpha$$

وبالتعويض :

$$P\left(200 - 2.602\left(\frac{10}{\sqrt{16}}\right) < \mu < 200 + 2.602\left(\frac{10}{\sqrt{16}}\right)\right) = 0.99 \Rightarrow$$

$$[193.495 \quad , \quad 206.505]$$

وبالتالي يمكن القول أنه باحتمال 99% فإن متوسط السرعة لن يقل عن 193.495 كم /ساعة، ولن يزيد عن 206.5 لهذا النوع الجديد من السيارات.

4-6- تقدير الفرق بين متوسطي مجتمعين إحصائيين (μ_1-μ_2) غير معلومي التباين وحجم العينات صغيراً :

إذا كان $(n_1, n_1) \leq 30$ و (δ_1^2, δ_2^2) غير معلومين، في هذه الحالة نقوم بإنشاء لـ (μ_1- μ_2) وفقاً لتوزيع (t) ستودنت، بشرط $\delta_1^2 = \delta_2^2$ والتباين المشترك الذي يعطى بالعلاقة :

$$S_p^2 = \frac{S_1^2.(n-1) + S_2^2.(n-1)}{n_1 + n_2 - 2}$$

والخطأ المعياري لـ (μ_1-μ_2) للعينات الصغيرة يعطى بالعلاقة :

$$\delta_{\bar{x}1-\bar{x}2} = SP\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

ونحن نعلم أنه من أجل احتمال (1-α) فإن (t) تتوزع وفق العلاقة :

$$P\,[\,-t_{\alpha/2} < t < t_{\alpha/2}] = 1 - \alpha$$

وإن الاختبار (t) يأخذ الشكل :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}}$$

وبالتعويض نجد أن :

$$p\left[-t_{\alpha/2} < \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}} < t_{\alpha/2}\right] = 1 - \alpha$$

وبالتالي تصبح العلاقة :

$$P\left[(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu - \mu < (\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right] = 1 - \alpha$$

تمرين (3):

بفرض أنه لدينا عينتين عشوائيتين مستقلتين من الكتب في مجال الاقتصاد والعلوم السياسية، وقمنا بحساب البيانات : $n_1=10$ $\bar{x}_1=3.5$ دولار $S_1^2=25$ التالية (العلوم السياسية):

$$S_2^2=22 \ , \bar{x}2=3 \ n_2=10$$

والمطلوب : أوجد باحتمال قدرة 95% التقدير المجالي للفرق، بين متوسطي أسعار الكتب في هذين المجتمعين، بافتراض تساوي التباينات في المجتمع وأسعار الكتب تتبع التوزيع الطبيعي.

الحل :

التباين المشترك $S_p=\sqrt{\dfrac{S_1^2.(n-1)+S_2^2.(n-1)}{n_1+n_2-2}}$

$$=\sqrt{\frac{25(10-1)+22(10-1)}{10+10-2}}=4.847$$

$$P\left[(3.5-3)-2.101(4.847)\sqrt{\frac{1}{10}+\frac{1}{10}}<\mu_1-\mu_2<(3.5-3)+2.101(4.847)\sqrt{\frac{1}{10}+\frac{1}{10}}\right]=0.95$$

إن قيمة (t) الجدولية من أجل :

V = 18

$\alpha = 0.05$

$T_{\alpha/2} = 2.101$

$$[-4.054 < \mu_1 - \mu_2 < 5.054]$$

وبالتالي يمكن القول أنه باحتمال 95% فإن المجال [-4.054,5.054] يحتوي القيمة الحقيقية للفرق بين المتوسطين.

6-5- إختبار الفرضيات في حالة العينات الصغيرة :

اختبار الفروض المتوسط (μ) (حالة مجتمع واحد) :

عندما يكون لدينا عينة صغيرة الحجم $n \leq 30$ وأن (δ^2) غير معلوم، في هذه الحالة نبني قرارنا اعتماداً على توزيع (t) ستودنت من أجل عدد من درجات الحرية $v=n-1$ وتكون صياغة الفرضية (H_0 و H_1) مشابهاً تماماً للعينات الكبيرة. $H_0: \mu=\mu_0$ $H_1: \mu \neq \mu_0$

ويتم الاختبار الإحصائي يتم من خلال العلاقة :

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

تمرين (4):

تدعي إحدى المكتبات أن متوسط سعر الكتب المرجعية هذه السنة هو (20) دولاراً وبغية التأكد من ذلك سحبت عينة من (16) كتاباً فوجد أن $\bar{x} = 24$ دولاراً وبانحراف معياري قدره (3) دولارات . بفرض أن أسعار الكتب تتبع التوزيع الطبيعي. **المطلوب** :هل تدل هذه البيانات على أن متوسط سعر الكتاب المرجعي $\bar{x} = 20$ عند مستوى معنوية $\alpha = 0.05$

الحل :

1 ــ صياغة الفرض $H_0 : \mu = 20$

$$H1: \mu \neq 20$$

2 ـ إجراء الاختبار الإحصائي :

$$t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

$$= \frac{|24 - 20|}{3/\sqrt{16}} = 5.33$$

3 ـ إتخاذ القرار :

من خلال المقارنة بين $t = 5.33$ وقيمة t الجدولية من أجل ($v=15$) و ($\alpha =$ 5%) فإن $t=1.753$ وبالتالي ($t > t_{\alpha/2}$)، نرفض الفرضية (H_0) ونقـول أن متوسط سعر الكتاب لم يعد يساوي ($\bar{x} = 20$) دولاراً.

6-6- اختبار معنوية الفـروق للأوسـاط الحسـابية (μ_1-μ_2) (حالـة مجتمعين):

في هذه الحالة إن قيم (δ_1^2, δ_2^2) غير معلومتين وإن $(n_1 , n_2) \leq 30$ والمجتمعين يتوزعان توزيعاً طبيعياً، وإن رفض أو قبول الفرض (H_0) يعتمد على قيمة (t) توزيع ستودنت من أجل عدد درجات حريـة ($V=n_1+n_2-2$) ضمن شرط تساوي التباينات في المجتمعين أي أن :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{S_p\sqrt{1/n_1 + 1/n_2}}$$

حيث : $d_0 = \mu_1 - \mu_2$

$$SP = \sqrt{\frac{S_1^2(n-1) + S_2^2(n-1)}{n_1 + n_2 - 2}}$$

تمرين (5) :

رغب أحد دكاترة مقرر الإحصاء في إحدى الكليات معرفة إذا كان لطريقتين مختلفتين في التدريس (الطريقة العادية، الطريقة المتلفزة) نفس التأثير، سحبت عينتان مستقلتان من طلاب تلك الكلية عشوائياً وتم تدريس كل منهما بطريقة من الطريقتين السابقتين وبعد إنتهاء مدة التدريس أجرى لهما امتحاناً موحداً لقياس مستوى الطلاب، فحصلنا على البيانات الآتية :

$$n_1 = 18, \quad \overline{x_1} = 80, \quad S_1^2 = 36$$

$$n_2 = 12, \quad \overline{x_2} = 85, \quad S_2^2 = 34$$

المطلوب : اختبر الفرض القائل بأن هاتين الطريقتين بالتدريس متساويتان عند مستوى معنوية $\alpha=0.05$ وتحت افتراض أن التوزيع التكراري للمجتمعين طبيعياً والتباينات متساوية.

الحل :

1 ـ صياغة الفرضية :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

2 ـ الاختبار الإحصائي :

$$t = \frac{(\overline{x_1} - \overline{x_2}) - d_0}{S_p \sqrt{1/n_1 + 1/n_2}}$$

$$S_p = \sqrt{\frac{S_1^2.(n-1)+S_2^2.(n-1)}{n_1+n_2-2}}$$

$$S_p = \sqrt{\frac{36(18-1)+34(12-1)}{18+12-2}} = 5.93$$

$$t = \frac{(85-80)}{5.93\sqrt{1/18+1/12}} = 2.27$$

3 ـ إتخاذ القرار :

من خلال المقارنة بين $t = 2.27$ وقيمة t الجدولية من أجل $V = 28$ و $\alpha =$ 5% والاختبار من إتجاهين فإن $t = 2.048$ وبالتالي : $|t| > t$ فإننا نـــرفض الفرضية H_0 ونقبل الفرضية البديلة H_1 ونقول بأن هناك فرقاً حقيقيـــاً بـــين الطريقتين.

تمرين (6) :

يرغب مدير إحدى رياض الأطفال قياس قدرة الطفل على تعلم كلمات جديدة، وللتأكد من ذلك، سحبت عينتان مستقلتان من الأطفال في إحدى رياض الأطفال وتم قياس قدرة كل منهم على تعلم الكلمات الجديدة وبعـــد الإنتهـــاء، أجري لهما امتحاناً موحداً لقياس قدرة أطفال العينتين، فحصلنا على البيانـــات التالية :

العينة الأولى	العينة الثانية
$n_1 = 10$	$n_2 = 8$
$\bar{x}_1 = 95$	$\bar{x}_2 = 97$
$S_1^2 = 40$	$S_2^2 = 36$

وبافتراض أن درجات اختبار جميع الأطفال تتبع التوزيع الطبيعي والتباينات متساوية، **المطلوب:** اختبر الفرض القائل بتساوي قدرة أطفال العينتين على تعلم الكلمات الجديدة، عند مستوى دلالة $\alpha = 5\%$.

الحل :

أ ــ صياغة الفرض :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

2 ــ إجراء الاختبار الإحصائي :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{S_p \sqrt{1/n_1 + 1/n_2}}$$

$$S_p = \sqrt{\frac{S_1^2.(n-1) + S_2^2.(n-1)}{n_1 + n_2 - 2}}$$

$$S_p = \sqrt{\frac{40(10-1) + 36(8-1)}{10+8-2}} = 6.36$$

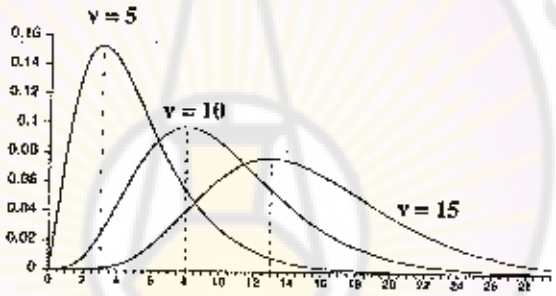
$$t = \frac{(85-80) - 0}{6.36\sqrt{1/10 + 1/8}} = 0.66$$

3 ــ اتخاذ القرار :

من خلال المقارنة بين t=6.36 وقيمة t الجدولية من أجل v=16 و $\alpha = 5\%$ والاختبار من إتجاهين فإن t=2.119 وبالتالي : t<t ونقبل الفرضية H_0 ونقول ليس هناك فرقاً بين قدرة أطفال العينتين على تعلم الكلمات الجيدة.

7-6- توزيع كاي مربع (x^2) :

يتألف توزيع كاي مربع (χ^2) مثل توزيع (t) ستودنت، من مجموعة من المنحنيات، التي يتغير شكلها بتغير عدد درجات الحرية (v)، فعندما يكون ($v \leq 2$)، فإن منحني التوزيع يأخذ شكل حرف الراء المقلوبة، وبزيادة عدد درجات الحرية يصبح شكل المنحني ملتوياً نحو اليمين، ويقترب شكل المنحني من منحني التوزيع الطبيعي عندما يصبح عدد درجات الحرية ($v > 30$)



وتتوزع الكمية $\sqrt{2\chi^2}$ توزيعاً يقترب من الطبيعي بمتوسط $\sqrt{2n-1}$ وانحراف معياري = 1، وعلاقة المتحول المعياري تأخذ الشكل :

$$(t) = (Z) = \frac{\sqrt{2\chi^2} - \sqrt{2n-1}}{1}$$

يتضح مما سبق، أنه باستخدام (χ^2) نحتاج إلى عدد من الجداول تختلف باختلاف عدد درجات الحرية، وللتغلب على هذه الصعوبة، حسبت بعض القيم المهمة لـ (χ^2)، عند درجات حرية من (1) إلى (30) . ومستويات دلالة محددة.

نحن نعلم أن الأوساط الحسابية للعينات الصغيرة تتوزع وفق التوزيع (t)، أما بالنسبة للانحرافات المعيارية والمحسوبة من عينات صغيرة، فإن شكل توزيعها الاحتمالي ملتوياً نحو اليمين (موجب الإلتواء)، وتتوزع وفق (χ^2) الذي يعتمد على معلمتين أساسيتين هما، تباين المجتمع الاحصائي (δ^2) وحجم العينة المدروسة (n) وتعطى قيمة الاختبار بالعلاقة :

$$\chi^2 = \frac{(n-1)S^2}{\delta^2}$$

ـ خصائص توزيع كاي مربع :

1 ـ $\chi^2 \geq 0$

2 ـ توزيع كاي مربع يعتمد على متغير واحد فقط هو درجة الحرية (V=n-1)

3 ـ $E(\chi^2)=v$

4 ـ $var(\chi^2)=2v$

6-8- التقدير المجالي لمعلمة المجتمع (δ^2) :

لتقدير قيمة تباين المجتمع الاحصائي δ^2 والذي يتميز بتوزيعه الطبيعي، يمكن الاستفادة من تعريف المتغير العشوائي χ^2_α الذي يفترض أن تكون العينة العشوائية مأخوذة من مجتمع إحصائي ذي توزيع طبيعي.

$$\chi^2 \leq \frac{(n-1)S^2}{\delta^2}$$

وبالتالي يمكن تقدير قيمة المعلمة (δ^2) عندما تكون مجهولة، بدلالة قيمة (S^2) والمحسوب من عينة عشوائية صغيرة الحجم، بدرجة احتمال (1- α) وعدد من درجات الحرية (V=n-1). على النحو التالي :

$$P\left(a \le \frac{(n-1)S^2}{\delta^2} \le b\right) = 1-\alpha$$

وبحل المتراجحات السابقة نحصل على :

$$P\left(\frac{(n-1)S^2}{b} \le \delta^2 \le \frac{(n-1)S^2}{a}\right) = 1-\alpha$$

تمرين (7):

عينة عشوائية حجمها n=19 مسحوبة من $N(\mu, \sigma^2)$ متوسطها $\bar{x} = 89.7$ وتباينها $S^2=11.72$ **والمطلوب:** أوجد 95% فترة الثقة لـ δ^2 . أو (قدر باحتمال 95% تباين المجتمع δ^2).

الحل :

$$(n-1)S^2 = 18 * 11.72 = 210.96$$

من جداول (χ^2) نجد أن :

$$a = \chi^2_{1-\alpha/2}(n-1)$$

$$q = \chi^2\ 0.975 * (18) = 8.231$$

$$b = \chi^2_{\alpha/2}(n-1) = \chi^2_{0.025}(18) = 31.53$$

وبالتعويض والاختصار نجد أن :

$$[6.69 \le \sigma^2 \le 25.63]$$

وبالتالي يمكن القول أن σ^2 لن تقل عن 6.69 ولن تزيد عن 25.63 باحتمال 95% .

6-9- إختبار الفروض للتباين σ^2 :

لإجراء الاختبار بين (S^2) و (δ^2) لمعنوية الفرق عند مستوى دلالة (α) وعدد من درجات الحرية $(V=n-1)$ يأخذ المتغير χ^2 الصيغة التالية:

$$\chi^2 = \frac{(n-1)S^2}{\delta^2}$$

وبمقارنة χ^2 المحسوبة مع χ^2_α الجدولية عند مستوى دلالة محدد (α) وعدد درجات حرية $(V=n-1)$ يتم اتخاذ القرار.

تمرين (8) :

بفرض أن أعمار الطلاب بالجامعات السورية تتبع التوزيع الطبيعي بانحراف معياري قدره (6) سنوات، سحبت عينة عشوائية من (25) من طلاب إحدى الجامعات السورية فوجد أن الانحراف المعياري للأعمار في العينة يساوي (5). **المطلوب** : هل تدل هذه البيانات على أن تباين أعمار الطلاب بهذه الجامعة يختلف عن تباين أعمار الطلاب في الجامعات السورية عند مستوى دلالة $(\alpha = 0.05)$.

الحل :

1 ـ صياغة الفرضية :

$$H_0 : S^2 = \delta^2$$

$$H_1 : S^2 \neq \delta^2$$

2 ــ الاختبار الاحصائي :

$$\chi^2 = \frac{(n-1)S^2}{\delta^2} = \frac{24(5)^2}{(6)^2} = 16.66$$

3 ــ اتخاذ القرار :

نبحث عن قيمة $\chi^2_{0.05}$ من أجل V = 24 في جداول (χ^2) فنجد أن :

$$\chi^2_{0.05} = 36.415$$

من خلال المقارنة نجد أن : $\chi^2=(16.66) < \chi^2=(36.415)$، وبالتالي نقبل الفرضية H_0 ونقول ليس هناك من فارق معنوي بين تباين أعمار الطلاب بالجامعة وتباين أعمار الطلاب في الجامعات .

تمرين (9):

تدعي إحدى شركات صناعة بطاريات السيارات بأن الانحراف المعياري لمنتجاتها يساوي (0.9) سنة، وللتأكد من ذلك أخذت عينة عشوائية من (10) بطاريات إمكان انحرافها المعياري (1.2). والمطلوب : هل تدل هذه البيانات على أن الانحراف المعياري لمنتجات الشركة أكبر من (0.9)، عند مستوى معنوية (دلالة) $\alpha = 0.05$

الحل :

1 ــ صياغة الفرضية :

$$H_0 : \sigma 2 = 0.81$$
$$H_1 : \sigma 2 > 0.81$$

2 ــ الاختبار الإحصائي :

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{9(1.44)}{0.81} = 16.00$$

3 ـ المقارنة وإتخاذ القرار:

من جداول (χ^2) نجد أن :

$$\chi^2_\alpha = 16.919$$

$$V = 9$$

$$\alpha = 0.05$$

من خلال المقارنة نجد أن : $\chi^2 = (16) < \chi^2 = (16.919)$

وبالتالي نقبل فرضية العدم (H_0) ونقول بأن إدعاء الشركة كان صحيحاً وأن الانحراف المعياري لمنتجات الشركة لم يكن أكبر من (0.9).

6-10 توزيع فيشر (F) :

عند مناقشتنا للاختبارات الخاصة بـ ($\mu_1 - \mu_2$) في توزيع (t) ستودنت، افترضنا دوماً تجانس التباينات في المجتمعات المسحوبة منها العينات، أي أن التباين في العينة الأولى مساوٍ للتباين في العينة الثانية، لذلك وقبل إجراء اختبار معنوية الفروق للأوساط الحسابية للعينات الصغيرة الحجم لابد من إجراء اختبار تجانس التباين في المجتمعات الإحصائية المسحوبة منها العينات، من خلال النسبة $\frac{S_1^2}{S_2^2}$ التي تمثل توزيعاً تكرارياً غير متماثلاً وهو ملتوياً نحو اليمين، فكلما اقتربت النسبة من الواحد الصحيح اقترب تبايني المجتمعين الإحصائين المدروسين من بعضهما البعض، وبالعكس كلما ابتعدت عن الواحد ابتعد تبايني المجتمعين عن بعضهما البعض.

إن توزيع فيشر (F) يستخدم لاختبار تساوي تبايني مجتمعين وقيمة الاختبار تعطى بالعلاقة التالية :

$F = \frac{S_1^2}{S_2^2}$ وهو عبارة عن نسبة التباين الأكبر إلى التباين الأصغر

وإن قيمة الاختبار (F) تأخذ قيماً موجبة دوماً.

وعند المقارنة واتخاذ القرار نقارن قيمة (F) المحسوبة من العلاقة السابقة مع قيمة (F) الجدولية، عند مستوى دلالة محدد وعدد من درجات الحرية للتباين في المقام ($V=n_1-1$) ونحصل عليه من سطر جداول (F) وعدد من درجات الحرية للتباين البسط ($V=n_2-1$) ونحصل عليه من عمود جداول (F)، بمعنى أننا نحصل على قيمة (F) الجدولية من نقطة تقاطع السطر مع العمود، مع أنه في بعض الحالات تؤخذ أقرب قيمة لدرجات الحرية في حال عدم توافرها في الجداول.

ويمكن القول، أنه كلما اقترب حاصل القسمة في علاقة (F) من الواحد الصحيح، كلما ازداد الاحتمال بأن يكون الفرق غير معنوياً بين التباينات، وبالعكس كلما ازدادت قيمة (F) كلما ازداد الاحتمال بأن يكون الفرق معنوياً. إن توزيع (F) مشابه لتوزيع (χ^2) من حيث توزع المنحنيات التي يتغير شكلها بتغير عدد درجات الحرية وإن التوزيعان (t) و (χ^2) يعتمدان اعتماداً كاملاً على معلمة واحدة فقط هي درجات الحرية بينما التوزيع (F) على معلمتين (V_1) و (V_2).

6-11- اختبار المقارنة بين تباين مجتمعين :

لنفرض أننا سحبنا عينتين عشوائيتين من مجتمعين إحصائيين مستقلين وحصلنا على البيانات التالية :

العينة الأولى	العينة الثانية
$n_1 = 12$	$n_2 = 10$

$\bar{x}_1 = 85$ $\bar{x}_2 = 81$

$S_1 = 4$ $S_2 = 5$

والمطلوب :

اختبر الفرض القائل بأن تبايني المجتمعين متساويين $\delta_1^2 = \delta_2^2$ عند مستوى دلالة $\alpha=0.05$

الحل :

1ـ صياغة الفرضية :

$$H_0 : \delta_1^2 = \delta_2^2$$
$$H_1 : \delta_1^2 \neq \delta_2^2$$

2 ـ الاختبار الإحصائي :

$$F = \frac{S_2^2}{S_1^2} = \frac{(5)^2}{(4)^2} = \frac{25}{16} = 1.5625$$

3 ـ المقارنة والقرار :

من جداول F نجد أن : F (3.11)

V= (11,9)

α=0.05

من خلال المقارنة نجد أن : F < F وبالتالي نقبل فرضية العدم (H_0) والمجتمعين متجانسين أي أن $\delta_1^2 = \delta_2^2$

تمرين (10) :

أرادت إحدى المؤسسات الاجتماعية أن توضح فيما إذا كان أسلوب النقاش الجماعي يؤدي إلى تقارب الآراء حول موضوع معين. ثم اختيار: المجموعة الأولى $n_1=6$ أشخاص وطرحت عليهم الموضوع، وطلبت منهم إعلان رأيهم النهائي بعد النقاش الجماعي، وبعد إعطاء الدرجات التقييمية

للآراء المطروحة فوجد أن $S_1^2 = 194.2$ ، والمجموعة الثانية $n_2=10$ أشخاص وطلب منهم إبداء الرأي على الموضوع الأول دون إجراء النقاش الجماعي فكان $S_2^2 = 77.9$

والمطلوب :

اختبر الفرضية القائلة بأن النقاش يؤدي إلى تقليل التباين في الآراء أي $H_0 : \delta_1^2 = \delta_2^2$ مقابل $H_1 : \delta^2 > \delta^2$ ، حيث أن : σ_1^2 تباين المجتمع الأول الذي يستعمل النقاش و σ_2^2 تباين المجتمع الثاني الذي لايستعمل النقاش عند مستوى معنوية $(\alpha = 0.05)$

الحل :

1 ــ الفرضية :

$$H_0 : \delta_1^2 = \delta_2^2$$

$$H_1 : \delta_1^2 > \delta_2^2$$

2 ــ الإختبار : $F = \frac{194.2}{77.9} = 2.49$

3 ــ المقارنة والقرار :

نبحث عن قيمة (F) الجدولية من أجل $(\alpha = 0.05)$ وعدد من درجات الحرية:

(في السطر) $V = n_1 - 1 = 5$

(في العمود) $V = n_2 - 1 = 9$

ومن نقطة تقاطع السطر مع العمود في جداول (F) نجد :

$$F_{0.05} = 3.48$$

وبالمقارنة نجد أن :

$$F < F_{0.05}$$

$$2.49 < 3.48$$

وبالتالي نقبل الفرضية (H_0)، ونقول بأن النقاش الجماعي حول هذا الموضوع لم يكن كافياً لتقليل التباين فيما لو لم يتم النقاش .

تمارين غير محلولة

تمرين (1) :

يعتقد أن متوسط عمر الناخبين في إحدى المدن لايزيد عن (30) سنة. ويدعي أحد مراقبي الانتخابات بأن المتوسط يزيد عن (30) سنة. وللتأكد من ذلك أخذت عينة عشوائية مكونة من (25) ناخب من الناخبين في هذه المدينة، فوجد أن متوسط عمر الناخب هو $\bar{x} = 32$ وبانحراف معياري $S=2$ سنة.

المطلوب : هل إدعاء هذا المراقب كان صحيحاً، بفرض أن أعمار الناخبين يتبع التوزيع الطبيعي عند مستوى معنوية $\alpha = 0.05$.

تمرين (2) :

من المعلوم أن متوسط الانفاق الشهري على الثقافة في إحدى المدن خلال العام الماضي هو 300 وحدة نقدية في الشهر ويتوقع البعض زيادة الانفاق على الثقافة هذا العام بنسبة (10%) شهرياً، أي أن يكون (330) وحدة نقدية في الشهر. أخذت عينة عشوائية من (16) أسرة في هذه المدينة، فوجد أن متوسط الانفاق الشهري على الثقافة بهذه العينة يساوي (340) وحدة نقدية وبانحراف معياري قدره $S=25$ ، **المطلوب** :

هل تؤدي هذه البيانات إلى رفض الفرض القائل بأن متوسط الانفاق على الثقافة في هذه المدينة هذا العام لن يزيد عن (10%) عند مستوى دلالة $\alpha = 0.05$ وعلماً أن الانفاق على الثقافة يتبع التوزيع الطبيعي .

تمرين (3) :

ترغب إحدى دور النشر الكبرى في معرفة ما إذا كان متوسط ربحها في المطبوعات الدورية أقل من 3500 وحدة نقدية شهرياً سحبت عينة عشوائية

من (16) فاتورة من الفواتير السابقة، فوجد أن متوسط الربح في العينة هو (3200) وحدة نقدية وبانحراف معياري S=125 وحدة نقدية، **المطلوب :**

1 ــ هل تؤيد هذه البيانات الفرض القائل بأن متوسط الربح يقل عن (3500) وحدة نقدية عند $\alpha = 0.05$ وذلك بافتراض أن أرباح المطبوعات تتبع التوزيع الطبيعي.

2 ــ استخدم بيانات العينة لإيجاد فترة ثقة 0.95 للمتوسط الحقيقي للأرباح في المجتمع.

تمرين (4):

ترغب إحدى الكليات في معرفة متوسط المسافة التي يقطعها الطلبة من منازلهم إلى الكلية، فأخذت عينة عشوائية من (16) طالباً فوجد أن $\bar{x}=12$ كيلو متراً وتباينها $S^2=36$

والمطلوب:

1 ــ اختبر فرض العدم القائل بأن الطالب يجب أن يقطع مسافة لاتقل عن (14) كيلو متراً عند $\alpha = 0.05$. بفرض أن المسافة التي يقطعها الطلبة للوصول إلى مدارسهم تتبع التوزيع الطبيعي .

2 ــ أوجد فترة ثقة 0.99 للمتوسط الحقيقي في المجتمع .

تمرين (5) :

يدعى أحد الطلاب بأن متوسط عدد ساعات الدراسة الأسبوعية للطلاب في الأسابيع الأخيرة قبل بدء الامتحان هو (20) ساعة. أخذت عينة عشوائية من (16) طالب، فوجد أن $\bar{x}=18$ساعة و S=4 ساعات، **المطلوب :** هل تؤدي هذه البيانات إلى قبول الإدعاء السابق عند مستوى معنوية ($\alpha=0.01$) بفرض أن ساعات الدراسة تتبع التوزيع الطبيعي .

تمرين (6):

تقوم إحدى الفضائيات العربية بتدريب مذيعي نشرات الأخبار الجدد باستخدام برنامجين مختلفين للتدريب، حيث يلتحق كل مذيع جديد بأحد البرنامجين عشوائياً سحبت عينة من (25) من مذيعي النشرات الإخبارية الذين تم تدريبهم باستخدام البرنامج (A) كما سحبت عينة أخرى مستقلة من (16) من مذيعي النشرات الذين تم تدريبهم باستخدام برنامج التدريب (B) والبيانات التالية تبين النتائج خلال (30) يوم من انتهاء التدريب .

(A)	(B)
$\overline{x}_2 = 80$	$\overline{x}_1 = 90$
$S_2^2 = 36$	$S_1^2 = 25$

والمطلوب : هل تدل هذه البيانات على أن برنامج التدريب (A) يكافئ برنامج التدريب (B) عند مستوى معنوية $\alpha = 0.05$، بافتراض أن درجات الاختبار تتبع التوزيع الطبيعي والتباينات متساوية $\delta_1^2 = \delta_2^2$.

تمرين (7) :

سحبت عينة عشوائية من (30) طالب من إحدى الجامعات السورية فوجد أن تباين إنفاقهم السنوي بمئات الوحدات النقدية مساوياً $S^2=100$، وإذا علمت أن الإنفاق السنوي لجميع طلبة الجامعات يتبع التوزيع الطبيعي بتباين قدره $S^2=90$ بمئات الوحدات النقدية .

المطلوب : هل تدل هذه البيانات على أن تباين إنفاق طلبة هذه الجامعة أكبر معنوياً من تباين إنفاق طلبة كل الجامعات وذلك عند $(\alpha = 0.05)$.

تمرين (9) :

قام طلاب كلية الطب في جامعة دمشق بدراسة ضغط الدم لمجموعة من لاعبي كرة القدم لمجموعتين من الطلاب في الجامعة ومن نفس العمر وقد تم

اختيار (11) لاعباً كعينة من المجتمع الأول فكان الانحراف المعياري (S=4)، ومن المجتمع الثاني اختبرت عينة من (21) لاعباً فكان الانحراف المعياري (S=5).

المطلوب : هل تعتقد أن التباين يختلف بشكل معنوي بين العينتين عند مستوى معنوية (α=0.05) .

تمرين (10) :

بفرض لدينا عينتين عشوائيتين مستقلتين لمرتبات العمال والعاملات في إحدى الشركات، فكانت النتائج الآتية .

العاملات	العمال
$n_2=8$	$n_1=5$
$\bar{x}_2 = 10491$	$\bar{x}_1 = 10750$
$S_2 = 200$	$S_1=150$

وبافتراض أن مرتبات العمال والعاملات تتبع التوزيع الطبيعي وإن التباينات متساوية.

المطلوب : اختبر الفرض القائل بأنه ليس هناك من فارق معنوي بين مرتبات العمال والعاملات هذه الشركة وذلك عند مستوى معنوية (α=0.05) .

الفصل السابع

الإحصاء اللا معلمي

Nonparametric

ـ الاختبارات الإحصائية اللامعلمية (لكاي مربع، الإستقلال لجداول التوافق، تصحيح ياتس)

ـ الاختبارات الإحصائية اللامعلمية للمقارنة بين عينتين (اختبار الإشارة، اختبار مان ويتني، اختبار ولكوكسون)

ـ الاختبارات الإحصائية اللامعلمية للمقارنة بين عدة عينات (اختبار كروسكال ـ واليس)

ـ الاختبارات الإحصائية اللامعلمية بين عينتين للبيانات الاسمية والرتبية (اختبار كولموجروف ـ سميرنوف ، اختبار ثنائي الحدين ، اختبار فيشر، اختبار الوسيط)

ـ بعض معاملات الارتباط للبيانات الاسمية والرتبية (معامل ارتباط سبيرمان، معامل ارتباط كاندل ـ تاو ـ للرتب، معامل الوفاق، معامل الاقتران الرباعي، معامل ارتباط فاي، معامل التوافق، معامل كرامير، معامل تشيبرو.

ـ تمارين غير محلولة.

1-7- مقدمة :

إن معظم أساليب اختبار الفروض التي ناقشناها سابقاً تعرف بالطرائق المعلمية Parametric methods، وذلك لأنها تهتم بمعالم (ثوابت) المجتمع مثل $(\bar{x}, \delta^2, \bar{P})$ ، وعلى الرغم من ذلك يوجد العديد من المسائل التي لانستطيع فيها تطبيق الطرق المعلمية، ولاسيما عندما تؤخذ المقاييس أو المشاهدات في صور رتب يصبح من الأفضل استخدام الطرائق اللامعلمية Non parametric، أضف إلى ذلك أن الطرائق المعلمية تعتمد على إفتراض معرفة التوزيع الاحتمالي للمجتمع، فعند إجراء اختبارات الفروض الإحصائية للمتوسط والتباين أو للفرق بين وسطين حسابيين والنسبة بين تباينين، افترضنا أن التوزيع توزيعاً طبيعياً أو قريب من الطبيعي للمجتمعات الأصلية التي سحبت منها العينات، أما بالنسبة للطرائق اللامعلمية فإنها لاتفترض شكل معين لتوزيع المجتمع الأصلي، أي أنها حرة التوزيع free distribution، وإن العبارتين لامعلمي وحرة التوزيع ليس لهما نفس المعنى تماماً، فعبارة لامعلمي تستخدم لوصف اختبارات الفروض التي لاتتضمن أية معلمة من معالم المجتمع، وحرة التوزيع (توزع حر) تستخدم لوصف اختبارات الفروض التي لاتفترض شكلاً معيناً لتوزيع المجتمع الأصلي.

لكن في حال تحقق شرط تبعية البيانات للتوزيع الطبيعي ينصح باستخدام الطرائق المعلمية والتي تؤدي في هذه الحالة إلى نتائج أدق، أما الطرائق اللامعلمية تستخدم فقط عندما يتأكد الباحث من عدم إمكانية استخدام الطرق المعلمية، علماً أن الطرائق اللامعلمية تتطلب عدداً أقل من الفروض وسهلة الشرح والتطبيق، ومن هذه الاختبارات المتعلقة بالإحصاء اللامعلمي نذكر:

2-7- الاختبارات الاحصائية اللامعلمية (لكاي مربع، الاستقلال لجداول التوافق، تصحيح باتس)

إختبار حسن المطابقة :

لقد بات معروفاً أن الدراسات الإحصائية تعتمد على فرض نموذج إحصائي تخضع له البيانات المتوفرة. فعلى سبيل المثال (درجات الطلاب تخضع للتوزيع الطبيعي، وأعداد حوادث السير تخضع لتوزيع بواسون وعدد المصابيح غير الصالحة في مصنع تخضع لتوزيع ثنائي الحدين وهكذا..).

لكن السؤال الذي يطرح نفسه كيف يمكن التأكد من أن البيانات المتوفرة تخضع لهذا النموذج الإحصائي المعطى ؟ إن الإجابة على هذا السؤال يتم من خلال اختبار كاي مربع χ^2 لحسن المطابقة، وتقوم فلسفة هذا الاختبار على حساب مقياس يعبر عن مدى التطابق بين المشاهدات الحقيقية (التكرارات الحقيقية) مع المشاهدات المتوقعة (التكرارات المتوقعة) فيما إذا كان النموذج الإحصائي صحيحاً، بمعنى آخر إن اختبار حسن المطابقة هو عبارة عن اختبار فيما إذا كان توزع المفردات للعينة يوافق أو يطابق ما هو مفروض أو متوقع أن يكون توزعها في المجتمع. فإذا رمزنا بـ E للتكرارات المتوقعة و Q للتكرارات الحقيقية المعروفة من خلال العينة، فإنه يمكن حساب الاختلاف أو مدى التطابق بين هذه التكرارات من خلال العلاقة الآتية :

$$\chi^2 = \sum_{i=1}^{r} \left(\frac{(Q-E)^2}{E} \right)$$

والاختبار سيكون على النحو التالي:

1 ـ صياغة الفرضية :

H_0 : تنص على أن التكرارات الحقيقية Q تتطابق مع التكرارات المتوقعة E

H_1 : تنص على أن التكرارات الحقيقية Q لا تتطابق مع التكرارات المتوقعة E.

2 ـ الاختبار الإحصائي :

$$\chi^2 = \sum_{i=1}^{r} \left(\frac{(Q-E)^2}{E} \right)$$

3 ـ إتخاذ القرار :

من جداول χ^2 نوجد قيمة χ^2 الجدولية $\chi^2(\alpha,v)$ عند مستوى معنوية α وعدد درجات حرية $v=r-1$.

من خلال المقارنة سيكون لدينا حالتين :

إما أن تكون (الجدولية) $\chi^2 < \chi^2$ (α , v) (المحسوبة)

وبالتالي يمكن القول أن التكرارات الحقيقية والمتوقعة متطابقة (نقبل H_0)

أو أن تكون (الجدولية $\chi^2 > \chi^2$ (α , v) (المحسوبة) وبالتالي يمكن القول أن التكرارات الحقيقية والمتوقعة غير متطابقة (نرفض الفرضية H_0)، ومن خلال التمارين التالية نوضح هذا الاختبار .

تمرين (1) :

رغبت إحدى المؤسسات الاجتماعية معرفة رأي طلبة الجامعة بالخصصة، فقامت المؤسسة بأخذ عينة عشوائية من (50) طالب، فوجد أن (32) منهم يؤيدون الفكرة و (18) منهم يعارضون، هل يمكن القبول بأن الآراء تنقسم بشكل متساوٍ حول هذا الموضوع وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل :

1 ـ صياغة الفرضية :

$$H_0 = P_1 = P_2 = 1/2$$

$$H_1 : P_1 \neq P_2 \neq 1/2$$

2 ــ الاختبار الإحصائي : $\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$

لحساب المجموع السابق نضع الجدول التالي :

التكرارات / الآراء	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
مؤيد	32	25	7	49	1.96
معارض	18	25	-7	49	1.96
$\sum$	50	50			χ^2=3.92

$$\chi^2 = 1.96 + 1.96 = 3.92$$

3 ــ إتخاذ القرار : من جداول χ^2 وعند $\alpha = 0.05$ و $V = 2-1=1$ نجد أن:

$$\chi^2(\alpha , v) = \chi^2 (0.05,1) = 3.84$$

ومن خلال المقارنة نجد أن :

$$X^2 = 3.92 > X^2_{(\alpha , v)} = 3.841 \text{ (المحسوبة)}$$

وبالتالي نرفض فرضية العدم H_0 ونقول أن الآراء في هذا المجتمع لاتنقسم بالتساوي.

تمرين (2) :

يهتم الباحثون في إحدى الدول بمعرفة الحالة التعليمية لجميع من لهم الحق في الانتخاب، وقد تم من خلال خبرة سابقة تصنيف جميع من لهم حق الإنتخاب حسب حالتهم التعليمية إلى خمسة أقسام: إبتدائي، إعدادي، ثانوي، جامعي، دراسات عليا . وإن النسب المقابلة لهذه الأقسام على النحو التالي :

الحالة التعليمية	ابتدائي	اعدادي	ثانوي	جامعي	دراسات عليا
النسبة المئوية	30%	35%	20%	10%	5%

أخذت عينة عشوائية من (1000) ناخب فوجد أن (270) إبتدائي، (315) إعدادي، (230) ثانوي، (125) جامعي، (60) دراسات عليا.

والمطلوب :

اختبار فيما إذا كان التوزيع الحالي للحالة التعليمية لمن لهم حق الانتخاب في هذه الدولة مازال هو نفس التوزيع السابق الذي تم الحصول عليه من خلال خبرة سابقة، وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل :

1 ــ صياغة الفرضية :

$$H_0 : P_1 = P_2 = P_3 = P_4 = P_5$$

علماً أن الحالة التعليمية تتوزع حسب النسب الآتية

$$(5\% , 10\% , 20\% , 35\% , 30\%)$$

$$H_1 : P_1 \neq P_2 \neq P_3 \neq P_4 \neq P_5$$

2 ـ الإختبار الإحصائي :

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

نحسب التكرارات المتوقعة E :

$E_1 = \frac{(30)(1000)}{100} = 300$ (الابتدائي)

$E_2 = \frac{(35)(1000)}{100} = 350$ (اعدادي)

$E_3 = \frac{(20)(1000)}{100} = 200$ (ثانوي)

$E_4 = \frac{(10)(1000)}{100} = 100$ (جامعي)

$E_5 = \frac{(5)(1000)}{100} = 50$ (دراسات عليا)

جدول تحديد التكرارات الحقيقية والمتوقعة :

الحالة التعليمية (الخلايا)	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
إبتدائي	270	300	-30	900	3.00
إعدادي	315	350	-35	1225	3.50
ثانوي	230	200	30	900	4.50
جامعي	125	100	25	625	6.25

دراسات عليا	60	50	10	100	2.00
$\sum$	1000	1000			$\chi^2=19.25$

3 ــ إتخاذ القرار : من جداول χ^2 نجد أن :

$$\chi^2_{(0,05,4)} = 9.48$$

و من خلال المقارنة نجد :

$$\chi^2 = 19.25 > \chi^2(\alpha, v) = 9.48$$

وبالتالي نرفض فرضية العدم H_0 ونقول بأن التوزيع للحالة التعليمية قد تغير عما كان عليه في السابق.

تمرين (3) :

رغب أحد الباحثين معرفة آراء المواطنين بشأن الإتجاه نحو عمل المرأة وقد تم من خلال خبرة سابقة تصنيف جميع الآراء والنسب المقابلة لها على النحو التالي :

الآراء	مؤيد بشدة	مؤيد	محايد	معارض	معارض بشدة	$\sum$
النسب المئوية	30%	20%	10%	25%	15%	100%

قام الباحث باختبار عينة عشوائية تتكون من (200) مواطن لاستطلاع آرائهم بشأن الاتجاه حول عمل المرأة من خلال اختيار إجابة واحدة للسؤال الموجه إليهم، وهو هل تؤيد عمل المرأة خارج البيت في بعض الأعمال الموجودة في المجتمع ؟ فكانت النتائج على النحو التالي:

الآراء	مؤيد بشدة	مؤيد	محايد	معارض	معارض بشدة	Σ
عدد المواطنين	75	55	15	45	10	100

المطلوب : اختبر الفرض القائل بأن هناك تشابه بين نسب إجابات العينة مع النسب في المجتمع، من خلال الخبرة السابقة عند مستوى (α=0.05) .

الحل :

1 ــ صياغة الفرضية :

$H_0 : P_1 = P_2 = P_3 = P_4 = P_5$

تتوزع الحالة (الآراء) على النحو التالي :

(15%, 25%, 10%, 20%, 30%)

$H_1 : P_1 \neq P_2 \neq P_3 \neq P_4 \neq P_5$

(أي لا تتوافق النسب في هذه الحالة مع النسب السابقة)

2 ــ الاختبار الإحصائي :

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

نحسب التكرارات المتوقعة E :

$$E_1 = \frac{(30)(200)}{100} = 60$$

(مؤيد) $E_2 = \frac{(20)(200)}{100} = 40$

(مؤيد بشدة)

$$E_3 = \frac{(10)(200)}{100} = 20$$ (محايد)

$$E_4 = \frac{(25)(200)}{100} = 50$$ (معارض)

$$E_5 = \frac{(15)(200)}{100} = 30$$

(معارض بشدة)

جدول التكرارات الحقيقية والمتوقعة :

الآراء / الخلايا	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
1	75	60	15	225	3.750
2	55	40	15	225	5.625
3	15	20	-5	25	1.250
4	45	50	-5	25	0.5
5	10	30	-20	400	13.333
$\sum$	200	200			$\chi^2=24.548$

3 ــ إتخاذ القرار : من جداول توزيع χ^2 نجد أن :

$$\chi^2_{(\alpha,v)} = \chi^2_{(0.05,4)} = 9.487$$

ومن خلال المقارنة نجد أن :

$$\chi^2 = 24.548 > \chi^2_{(\alpha,v)} = 9.487$$

وبالتالي نرفض فرضية العدم H_0 أي أن النسب لم تعد كالسابق وليس هناك تشابه بين إجابات العينة والمجتمع .

تمرين (4) :

وجد نتيجة رمي قطعة نقدية معدنية (200) مرة، أن عدد المرات التي جاءت بها القطعة على الوجه الأول (112) مرة و (88) مرة الوجه الثاني

والمطلوب :

اختبر الفرضية القائلة بأن قطعة النقود كانت سوية (متزنة) عند مستوى معنوية (0.05)؟ .

الحل :

1 ـ صياغة الفرضية :

$H_0 : P_1 = P_2 = ½$

$H_1 : P_1 \neq P_2 \neq 1/2$

2 ـ الاختبار الإحصائي :

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

جدول التكرارات الحقيقية والمتوقعة :

التكرارات	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
الوجه الأول	112	100	12	144	1.44
الوجه الثاني	88	100	-12	144	1.44
$\sum$	200	200			$\chi^2=2.99$

3 ـ إتخاذ القرار : من جداول χ^2 نجد أن $\chi^2_{(0.05,1)} = 3.841$

ومن خلال المقارنة نجد أن $\chi^2_{(\alpha,v)} = 3.841 > \chi^2 = 2.88$ وبالتالي نقبل بفرضية العدم H_0 ونقول بأن القطعة كانت سوية .

ملاحظات حول اختبار حسن المطابقة :

1ـ شاهدنا في بعض التطبيقات أننا افترضنا أن نسبة الإجابات في المجتمع الأصلي كانت متساوية لكل الاختبارات (التوزيع بالتساوي) لكن هذا ليس هو الحال في جميع التجارب أو التطبيقات أو البحوث، فقد تكون التكرارات المتوقعة E_i تتوزع وفقاً لقواعد أو نسب معينة معروفة في ضوء خبرة سابقة ؛

2 ـ عند استخدام اختبار χ^2 لجودة المطابقة، ينبغي ألا تقل القيمة المتوقعة (التكرار المتوقع) عن (5) أي عن 20% من العدد الكلي، ويمكن في بعض الأحيان اللجوء إلى ضم الخلايا (أي التكرارات المتوقعة) التي أقل من (5) إلى بعضها البعض، أما في حالة كانت هذه الخلية الوحيدة التي تقل عن (5) فلا تؤخذ بعين الاعتبار وتبقى كماهي،

3ـ هناك إمكانية دمج عدة قيم محسوبة لـ χ^2 في قيمة واحدة، بمعنى إذا كان لدينا قيم كاي مربع لعدة عينات خاصة بنفس الموضوع المدروس ومسحوبة في نفس المجتمع، فإن هذه القيم يمكن إضافتها إلى بعضها لإتخاذ قراراً أفضل من القرار الذي اتخذ على أساس البيانات الخاصة بعينة واحدة فقط، ويجب أن تكون درجات الحرية في العينة متساوية في كل التجارب وتكون درجة الحرية لاختبار χ^2 الكلية تساوي مجموع جميع درجات الحرية بالتجارب، وبعبارة أخرى فإن شرط دمج قيمة χ^2 هو :

1 ـ أن تكون الجداول في التجارب من نفس النمط.

2 ـ أن تكون التجارب تحمل نفس درجات الحرية. ونوضح ذلك من خلال التمرين التالي:

تمرين (5) :

بفرض أن البيانات التالية تمثل توزيع عدد الكتب المستعارة في المكتبتين A (كلية الاقتصاد) B (كلية العلوم السياسية) بجامعة دمشق في أسبوع معين .

المكتبة A (كلية الاقتصاد):

اليوم	السبت	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس	الجمعة	Σ
عدد الكتب المستعارة	24	28	12	10	7	8	11	100

المكتبة B (كلية العلوم السياسية):

اليوم	السبت	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس	الجمعة	Σ
عدد الكتب المستعارة	9	10	10	8	5	12	16	70

والمطلوب :

1 ـ اختبر الفرض القائل بأن نسبة الكتب المستعارة في كل يوم من يومي السبت والأحد تساوي 25% بينما تساوي نسبة الكتب المستعارة في كل يوم من أيام الأسبوع الباقية هي 10% وذلك عند مستوى معنوية $\alpha = 0.05$.

2 ـ إذا علمت أن قيمة χ^2 لتوزيع الكتب المستعارة في هذا الأسبوع في المكتبة C (كلية الآداب) تساوي (9.8) وفي المكتبة D (كلية التربية) تساوي

(15,52)، فما هو القرار الواجب إتخاذه على مستوى الجامعة ككل عند مستوى معنوية $\alpha = 0.05$

الحل :

1 ــ صياغة الفرضية :

$H_0 : P_1 = 25\% \quad , \quad P_2 = 10\%$

أي الكتب المستعارة في كلاً من A ,B تتوزع بنسبة 25% للسبت والأحد، أما الباقي 10%

$H_1 : P_1 \neq 25\% \ , \ P_2 \neq 10\%$

الكتب المستعارة لاتتوزع بالنسب السابقة .

2 ــ الاختبار الإحصائي :

$$\chi^2 = \sum \left(\frac{(Q - E)^2}{E} \right)$$

جدول التكرارات الحقيقية والمتوقعة في المكتبة A (كلية الاقتصاد):

أيام الأسبوع	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
السبت	24	25	1	1	0.04
الأحد	28	25	3	9	0.36
الاثنين	12	10	2	4	0.4
الثلاثاء	10	10	0	0	0
الأربعاء	7	10	3	9	0.9
الخميس	8	10	2	4	0.4
الجمعة	11	10	1	1	0.1
$\sum$	100	100			$\chi^2=2.2$

3 ـــ المقارنة وإتخاذ القرار : من جداول χ^2 نجد أن $\chi^2_{(0.05,\,6)} = 12.6$

وبالمقارنة

$$\chi^2 = 2.2 < \chi^2_{(\alpha,\,v)} = 12.6$$

وبالتالي نقبل فرضية العدم H_0 وتتوزع الكتب المستعارة حسب النسب المذكورة.

أما جدول التكرارات الحقيقية والمتوقعة في المكتبة B (كلية العلوم السياسية):

أيام الأسبوع	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
السبت	9	17.5	-8.5	72.25	4.13
الأحد	10	17.5	-7.5	56.25	3.21
الاثنين	10	7	3	9	1.28
الثلاثاء	8	7	1	1	0.14
الأربعاء	5	7	2	4	0.57
الخميس	12	7	5	25	3.57
الجمعة	16	7	9	81	11.57
$\sum$	70	70			$\chi^2=24.48$

3 ـــ المقارنة واتخاذ القرار :

من جداول χ^2 نجد أن : $\chi^2_{(0.05\,,\,6)} = 12.6$

$$\chi^2 = 24.48 > \chi^2_{(\alpha,\,v)} = 12.6$$

وبالتالي نرفض فرضية العدم H_0 ونقول أن نسبة الكتب لا تتوزع حسب النسب المذكورة.

II ــ بعد مناقشة قبول ورفض فرضية العدم لكل عينة على حدة من خـــلال الجدول التالي، نقوم بدمج قيم X^2 :

النتيجة الاولية	الكلية	χ^2	v
قبول H_0	A (الاقتصاد)	2.2	6
رفض H_0	B (العلوم السياسية)	24.48	6
قبول H_0	G (الآداب)	9.8	6
قبول H_0	D (التربية)	15.52	6
Σ		52	24

نحسب χ^2 الجدولية عند عدد من درجات الحرية المساوي لـــ (24) ومستوى معنوية $\alpha=0.05$ فنجد:

$$\chi^2_{(\alpha\,,\,v)} = \chi^2_{(0.05,\,24)} = 36.4$$

بالمقارنة نجد أن :

$$\chi^2 = 52 > \chi^2_{(\alpha,\,v)} = 36.4$$

وبالتالي نرفض فرضية لعدم H_0 ونقول بأن الكتب المستعارة تتوزع على مكتبات الجامعة بغير النسب المفروضة .

- **اختبار الاستقلال لجداول التوافق (2 × 2) أو (3 × 2) أو (3 × 3)**

إن هذا الاختبار يقوم على اختبار التطابق النسبي بين التكـــرارات الحقيقيـــة والتكرارات المتوقعة لظاهرتين ومن ثم إتخاذ قرار بشأن استقلال متغير مـــا إحصائياً عن متغير آخر، أي تحديد فيما إذا كان هناك علاقة مابين مجموعتين

من الصفات في المجتمع، من خلال بيانات العينة. بمعنى أخر اختبار العلاقة بين متغيرات الدراسة موضع الاهتمام غير القائم على الارتباط الخطي، وإنما ينصب اهتمامنا على التأكد من جودة العلاقة موضع البحث وهل يعملان باستقلالية عن بعضهما البعض، ويتم إعداد جدول التوافق الذي يبحث في مدى التطابق بين (Q) و (E) من خلال بيانات العينة على اعتبار كل ظاهرة منقسمة إلى قسمين أو أكثر. ويتم الاختبار على النحو التالي:

1 ــ صياغة الفرضية : (فرضية الاستقلال) وتنص على أنه لاتوجد علاقة بين الظاهرتين الأولى والثانية، وكل ظاهرة مستقلة عن الأخرى.

2 ــ الاختبار الإحصائي (اختبار التوافق):

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

يتم إعداد جدول التوافق الذي يبحث في مدى التطابق بين Q ,E من خلال بيانات العينة، على اعتبار كل ظاهرة مقسمة إلى قسمين أو أكثر.

ويتم حساب التكرارات المتوقعة E على النحو التالي :

$$E\ (\text{المتوقع}) = \frac{(\Sigma\ \text{العمود})\ (\Sigma\ \text{السطر})}{(\Sigma\ \text{الكلي})}$$

ويتم إنشاء الجدول المتعلق بالتكرارات الحقيقية والتكرارات المتوقعة كمايلي:

الخلايا	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
1	Q_1	E_1	-	-	-
2	Q_2	E_2	-	-	-

3	Q_3	E_3	-	-	-
4	Q_4	E_4	-	-	-
Σ					χ^2

من خلال الجدول نلاحظ أن تقارب (Q) من (E) يؤدي إلى قيمة صغيرة لكاي مربع، وبالتالي كان التوافق والتجانس جيداً وبالعكس كلما كان كبيراً كان التوافق بعيداً.

3 ـ اتخاذ القرار : من خلال جداول χ^2 نوجد قيمة χ^2 الجدولية عند مستوى معنوية α وعدد درجات حرية $V = (C-1)(r-1)$ (حيث C, r عدد الأسطر والأعمدة) ثم نقارن بين χ^2 الجدولية و χ^2 المحسوبة فنجد أن : إما أن تكون $\chi^2 < \chi^2_{(\alpha, v)}$ (المحسوبة) (الجدولية) وبالتالي نقبل فرضية الاستقلال ونقول ليس هناك علاقة بين الظاهرتين أو $\chi^2 > \chi^2_{(\alpha, v)}$ (المحسوبة) وبالتالي نرفض فرضية الاستقلال ونقول أن هناك علاقة بين الظاهرتين ولتوضيح ذلك لدينا التمارين التالية :

تمرين (4)

قام أحد الباحثين باختيار عينة عشوائية من مشاهدي البرامج التلفزيونية واستطلع آرائهم لمعرفة رأيهم في البرامج التلفزيونية المقترحة على القناتين الأولى والثانية فكانت النتائج على النحو التالي :

القناة / النوع	الأولى	الثانية	Σ
ذكور	110	90	200
إناث	40	60	100

Σ	150	150	300

والمطلوب :

اختبر الفرض القائل باستقلال كل نوع (ذكور، إناث) واختيار القناة عند مستوى معنوية $\alpha=0.05$.

الحل :

1 ــ صياغة الفرضية: (فرضية الاستقلال) لاتوجد علاقة بين النوع (ذكور، إناث) وبين اختيار القناة، وكل ظاهرة مستقلة عن الأخرى.

2 ــ الاختبار الإحصائي : (اختبار التوافق):

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

تحديد التكرارات المتوقعة E :

$$E_1 = \frac{(200)(150)}{300} = 100 \qquad E_3 = \frac{(100)(150)}{300} = 50$$

$$E_2 = \frac{(200)(150)}{300} = 20 \qquad E_4 = \frac{(100)(150)}{300} = 50$$

نضع جدول تحديد التكرارات الحقيقية والتكرارات المتوقعة :

الخلايا	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
ذكور قناة (1)	110	100	10	100	1
ذكور قناة	90	100	-10	100	1

(2)					
إناث قناة (1)	40	50	-10	100	2
إناث قناة (2)	60	50	10	100	2
$\sum$	300	300			$\chi^2=6$

3 ــ إتخاذ القرار :

من جداول χ^2 عند مستوى دلالة $\alpha = 0.01$ وعدد من درجات الحرية $V = (2-1)(2-1)$

أي $v=1$ نجد أن : $\chi^2_{(0.01,1)} = 6.634$

ومن خلال المقارنة نجد $\chi^2 = 6 < \chi^2_{(\alpha,v)} = 6.634$ وبالتالي نقبل فرضية الاستقلال، ونقول ليس هناك علاقة بين الجنس واختيارهم للقناة.

تمرين (5) :

قام أحد الباحثين باختيار عينة عشوائية من (100) طالب وطالبة لدراسة علاقة الجنس (طالب، طالبة) واستخدام المكتبة الجامعية فحصل على الجدول التالي :

استخدام المكتبة / النوع	مستفيدون	غير مستفيدين	$\sum$
طلاب	15	15	30
طالبات	25	45	70
$\sum$	40	60	100

والمطلوب : اختبار فيما إذا كان هناك علاقة بين النوع (طالب، طالبة) واستخدام المكتبة الجامعية عند مستوى معنوية $\alpha = 0.05$

الحل :

1 ــ صياغة الفرضية : (فرضية الاستقلال) لاتوجد علاقــة بــين النــوع واستخدام المكتبة وكل ظاهرة مستقلة عن الأخرى .

2 ــ الاختبار الاحصائي : (اختبار التوافق)

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

نحسب التكرارات المتوقعة E :

$$E_1 = \frac{(30)(40)}{100} = 12 \qquad E_3 = \frac{(70)(40)}{100} = 28$$

$$E_2 = \frac{(30)(60)}{100} = 18 \qquad E_4 = \frac{(70)(60)}{100} = 42$$

جدول تحديد التكرارات الحقيقية والمتوقعة :

الخلايا	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
طلاب مستفيدين	15	12	3	9	0.75
طلاب غير مستفيدين	15	18	-3	9	0.5
طالبات مستفيدات	25	28	-3	9	0.32

طالبات غير مستفيدات	45	42	3	9	0.21
$\sum$	100	100			$\chi^2=1.78$

3 ــ إتخاذ القرار :

من خلال جداول χ^2 عند مستوى $\alpha=0.05$ ودرجات حريـــة $v = (2-1)(2-1)=1$

نجد $\chi^2_{(0.05,1)} = 3.841$

ومن خلال المقارنة $\chi^2 = 1.78 < \chi^2_{(\alpha,v)} = 3.841$

وبالتالي نقبل فرضية الاستقلال ونقول بأنه لاتوجد علاقة بين النـــوع (طالب، طالبة) وبين استخدام المكتبة والظاهرتين مستقلتين .

ــ جداول (2 × 3) :

تمرين (3) :

البيانات التالية توضح نتائج دراسة حول تسلم المرأة لمناصب سياسية :

الآراء / النوع	مؤيد	معارض	محايد	Σ
ذكور	80	40	40	160
إناث	100	80	60	240
Σ	180	120	100	400

والمطلوب : إختبار فيما إذا كان هناك علاقة بين الرأي والنوع عند مستوى معنوية $\alpha=0.05$.

الحل :

1 ــ صياغة الفرضية : (فرضية الاستقلال) لاتوجد علاقة بين الآراء والنوع وكل ظاهرة مستقلة عن الأخرى .

2 ــ الاختبار الإحصائي (اختبار التوافق) :

$$\chi^2 = \sum \left(\frac{(Q-E)^2}{E} \right)$$

نحسب التكرارات المتوقعة E :

$$E_1 = \frac{(180)(160)}{400} = 12 \qquad E_3 = 48 \qquad E_5 = 40$$

$$E_2 = \frac{(120)(160)}{400} = 108 \qquad E_4 = 72 \qquad E_6 = 60$$

جدول تحديد التكرارات الحقيقية والمتوقعة :

الخلايا	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
1	80	72	8	64	0.89
2	100	108	8	64	0.59
3	40	48	8	64	1.33
4	80	72	8	64	0.89
5	40	40	0	0	0
6	60	60	0	0	0
$\sum$	400	400			$\chi^2=3.7$

3 ـ اتخاذ القرار : من جداول χ^2 وعند $\alpha = 0.05$ ودرجات حرية $v=(3-1)(2-1)=2$

نجد $\chi^2_{(0.05,2)} = 5.99$ وبالتالي $\chi^2 = 3.7 < \chi^2_{(\alpha,v)}$ وبالتالي نقبل فرضية الاستقلال ونقول بأنه لاتوجد علاقة بين الآراء والنوع .

جداول (3 × 3) :

تمرين (7) :

قام عالم اجتماع بإجراء تجربة لتحديد ما إذا كان الإنجاز العلمي لدكاترة الجامعات يتأثر بحالتهم الاجتماعية، حيث تم تصنيف الحالة الزواجية إلى أعزب ومتزوج ومطلق، كما تم تصنيف الإنجاز العلمي للدكاترة إلى جيد، متوسط، سيء وكانت النتائج كمايلي :

الحالة الزواجية / الإنجاز العلمي	أعزب	متزوج	مطلق	Σ
جيد	56	71	12	139
متوسط	47	163	38	248
سيء	14	42	85	141
Σ	117	276	135	528

والمطلوب :

اختبر الفرض القائل بعدم وجود علاقة بين الإنجاز العلمي والحالة الزواجية عند مستوى معنوية $\alpha = 0.05$.

الحل :

1 ــ صياغة الفرضية : (فرضية الاستقلال) لاتوجد علاقة بين الإنجاز العلمي والحالة الزواجية، وكل ظاهرة مستقلة عن الأخرى .

2 ــ الاختبار الإحصائي : (اختبار التوافق):

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

نحسب التكرارات المتوقعة E :

$$E_1 = \frac{(139)(117)}{528} = 30.80$$

$E_4 = 54.95$ $E_3 = 35.54$ $E_2 = 72.66$

وكذلك : E_8, E_7, E_6, E_5

جدول تحديد التكرارات الحقيقية والمتوقعة التالي :

الخلايا	Q	E	(Q-E)	$(Q-E)^2$	$\left(\frac{(Q-E)^2}{E}\right)$
1	56	30.80	25.2	635.04	20.62
2	71	72.66	-1.66	2.75	0.04
3	12	35.54	-23.54	554.13	15.59
4	47	54.95	-7.95	63.20	1.15
5	163	129.64	33.36	1112.89	8.58
6	38	63.41	-25.41	645.67	10.18
7	14	31.23	-17.24	297.21	9.51
8	42	73.70	-31.7	1004.89	13.63
9	85	36.05	48.95	2396.10	66.47
$\sum$	528	528			$\chi^2 =$ 145.77

3 ــ إتخاذ القرار : من جداول χ^2 عند $\alpha = 0.05$ ودرجات حرية $V = (3-1)(3-1) = 4$ نجد $\chi^2_{(0.05,4)} = 9.487$

ومن خلال المقارنة $\chi^2 = 145.77 > \chi^2_{(\alpha,v)} = 9.487$

وبالتالي نرفض فرضية الاستقلال ونقول أن هناك علاقة بين الإنجاز العلمي والحالة الزواجية لدكاترة الجامعات .

3-7- تصحيح ياتس :

يستخدم هذا التصحيح في اختبارات الاستقلال لجداول التوافق (2 × 2) و (3 × 2) إذا كان أحد التكرارات المتوقعة أقل من (10) في إحدى الخلايا، ويقوم على طرح (1/2) من كل تكرار حقيقي أكبر من التكرار المتوقع E المقابل له، وإضافة (1/2) إلى كل تكرار حقيقي Q إذا كان أقل من التكرار المتوقع E المقابل له (وذلك لجميع التكرارات الحقيقية)، وينتج عن هذا التصحيح تقليل الفرق بين التكرارات الحقيقية والتكرارات المتوقعة باعتبار أن التكرارات الحقيقية تعتمد على قيم صحيحة، مع العلم أن التكرارات الحقيقية تختلف وتقفز قفزات ذات درجات منفصلة، بينما جدول χ^2 الذي يمثل توزيع درجات χ^2 يعتبر ذا قيم متصلة، لذلك يستخدم من أجل تخفيف هذه القفزات، مع ذلك عندما تكون التكرارات (في العينة) كبيرة الحجم لا داع لاستخدام هذا التصحيح ويصبح دون معنى.

تمرين (8) :

أجرى أحد الباحثين الاجتماعيين بحثاً عن دور الإتصال عن طريق معرفة إتجاهات الناس إزاء برامج الإذاعة، فسأل عينة من الأفراد عددها (43) شخصاً السؤال التالي :

هل تجد من الأسهل أن تستمتع إلى الأخبار بدلاً من قراءتها، فأجاب (بـنعم) (10) أشخاص من أصل (19) شخص من المدراء، وأيضاً أجـابوا (بـنعم) (20) شخصاً من أصل (24) شخص من العمال، فإذا علمـت أن الإجابـات تنحصر بـ نعم أو لا ، والمطلوب :

اختبر فيما إذا كان هناك علاقة بين نوع الإجابة والمهنة عند مستوى معنويـة $\alpha=0.05$ ؟.

الحل :

1 ـ صياغة الفرضية : (فرضية الاستقلال) لاتوجد علاقة بين نوع الإجابة (نعم، لا) والمهنة، وكلاً منهما مستقل عن الآخر.

2 ـ الاختبار الإحصائي :

$$\chi^2 = \sum\left(\frac{(Q-E)^2}{E}\right)$$

نضع جدول التكرارات النظرية :

المهنة / الإجابة	المدراء	عمال	Σ
نعم	10	20	30
لا	9	4	13
Σ	19	24	43

ثم نحسب التكرارات المتوقعة E :

$$E_1 = \frac{(30)(19)}{43} = 13.26 \qquad E_3 = 5.74$$

$$E_2 = \frac{(30)(24)}{43} = 16.74 \qquad E_4 = 7.26$$

نلاحظ أن بعض التكرارات المتوقعة أقل من 10، أي $E < 10$ والعينة صغيرة لذلك لابد من إجراء تصحيح ياتس .

ـ فإذا كان $E > Q$ نضيف ½ لـ Q

ـ وإذا كان $E < Q$ نطرح ½ من Q

كما هو واضح من خلال الجدول التالي :

الخلايا	Q	E	Q^1	Q^1-E	$(Q^1-E)^2$	$\frac{(Q1-E)^2}{E}$
1	10	13.26	10.5	-2.76	7.62	0.57
2	20	16.74	19.5	+2.76	7.62	0.45
3	9	5.74	8.5	+2.76	7.62	1.33
4	4	7.26	4.5	-2.76	7.62	1.05
$\sum$	43					χ^2=4.3

3 ـ إتخاذ القرار : من جداول $\chi 2$ عند α=0.05 و v=1 نجد χ^2=3.84 وبالمقارنة $\chi 2 = 3.4 < \chi 2_{(\alpha, v)} = 3.84$ وبالتالي نقبل فرضية الاستقلال، ونقول بأنه لايوجد علاقة بين نوع الإجابة والمهنة.

ملاحظات :

1 ـ في بعض الأحيان قد تعطى البيانات على شكل نسب مئوية في هذه الحالة لابد من تحويل النسب المئوية إلى أرقام مطلقة ومن ثم إجراء الاختبار. مثلاً: إذا كان لدينا مجتمع الدراسة (n=150) حيث أن

(80%) منهم نساء و (20%) رجال فمن الضروري تحويل هذه البيانات إلى (120) نساء و (30) رجال.

2 ـ نحن نعلم أن فكرة جداول الاستقلال (2 × 2) قائمة على فكرة الاستقلال بين الظاهرتين المدروستين أو عدم الاستقلال، لذلك في بعض الأحيان بعد أن نقوم بإجراء الاختبار للجداول (2 × 2) نرغب معرفة النسبة المئوية التي تحدد مدى الارتباط أو الاستقلال بين الظاهرتين عن طريق استخدام معامل الاقتران الذي هو عبارة عن مقياس يقيس شدة العلاقة الارتباطية بين ظاهرتين في الجداول (2 × 2) وتعطى قيمة بالعلاقة :

$$r_c = \frac{AD - BC}{AD + BC}$$

وكلما اقتربت قيمة هذا المعامل من الواحد كانت العلاقة قوية، بمعنى إذا كانت $r_c > 50\%$ دل ذلك على وجود علاقة بين المتغيرات وبالعكس، ويستخدم هذا المعامل أحياناً لدعم القرار المتخذ في اختبار (2 × 2)، فيما إذا كان قرارنا صحيحاً أم لا. فعلى سبيل المثال لو كان قرارنا قبول فرضية الاستقلال وقمنا بحساب معامل الاقتران وكان $r_c=0.29$ فإن هذا يدل على أن $r_c < 50\%$ فالعلاقة ضعيفة بين الظاهرتين، وهذا يؤكد قرار الاختبار الإحصائي بقبول فرضية الاستقلال، أي لا علاقة بين الظاهرتين المدروستين.

3 ـ بنفس الأسلوب السابق عندما يكون لدينا جداول (3 × 3) نستخدم معامل التوافق

$$r_A = \sqrt{\frac{G-1}{G}} \quad ; \quad G = \sum\left(\frac{fij^2}{f_i \cdot f_j}\right)$$

وكلما اقتربت قيمة r_A من الواحد كانت العلاقة قوية ومتوافقة (أي توافق قوي بين البيانات)، وأيضاً يستخدم هذا المعامل لدعم القرار المتخذ في إختبار (3 × 3)، فيما إذا كان قرارنا صحيحاً أم لا، فإذا كان $r_A > 0.50$ فالقرار المتعلق بالاختبار هو رفض فرضية الاستقلال، وبالعكس إذا كان $r_A < 0.05$ كان القرار هو قبول فرضية الاستقلال، وسيوضح تطبيق هذه العلاقة لاحقاً.

7-4- الاختبارات الاحصائية اللامعلمية للمقارنة بين عينتين (اختبار الإشارة، إختبار مان ويتني، اختبار ولكوكسون):

- **اختبار الإشارة Sign Test :**

لقد استخدمنا في السابق اختبار (t) لتحديد ما إذا كان متوسط المجتمعين متساويين وبفرض أن كل مجتمع من المجتمعين له توزيع طبيعي وأن التباينين مجهولين ومتساويين، لكنه وفي كثير من الأحيان لايتحقق أحد الشرطين أو كلاهما، وبالتالي لانستطيع استخدام اختبار (t)، ويصبح البديل هو اختبار الإشارة القائم على أن المجتمعين لهما نفس التوزيع ويقوم هذا الاختبار على اختبار صفة واحدة من صفتين مثلاً (الصح والخطأ) أو (النجاح والفشل) أو (اختيار الصفة الأكثر انتشاراً)، وسوف نميز ثلاث حالات هي :

أ ــ **حالة عينة واحدة** : يتم هذا الاختبار على النحو التالي : نختار رمز الإشارة الموجبة (+) للصفة التي تهمنا والإشارة السالبة (−) للفشل، ويمكن أن يكون بالعكس إذا كان الأمر الذي يثير اهتمامنا هو مقدار الفشل، وهكذا بالنسبة لأية صفة وعكسها.

وفرضية العدم (H_0) ستكون ليس هناك أي فرق بين الصفة وعكسها أو عدد الإشارات الموجبة (+) تعادل عدد الإشارات السالبة (−) ، أي ($H_0: P = 1/2$) والفرضية البديلة (H_1) ($H_1: P \neq 1/2$)، والاختبار الإحصائي يتم بتحديد

مستوى المعنوية (α) ثم نحصي عدد الإشارات الموجبة ثم ننظر في الجدول الخاص باختبار الإشارة (في آخر الكتاب) عن القيمة المقابلة لعدد الإشارات الموجبة أو بالعكس إذا كانت الاشارة السالبة هي التي تثير اهتمامنا، ومن ثم إتخاذ القرار من خلال المقارنة بين (α) – P و (P) الجدولية لقبول أو رفض فرضية العدم، مع العلم أنه يمكن الإستعاضة عن جداول اختبار الإشارة بجداول ثنائي الحدين، ونوضح ماتم شرحه من خلال التمرين التالي :

تمرين (9) :

قام أحد الباحثين باختيار عينة عشوائية من (10) مشاهدين وأستطلع آرائهم حول مشاهدة إحدى القنوات العربية، فوجد من بينهم اثنين فقط يشاهدان هذه القناة. المطلوب: اختبر الفرض القائل بأن مشاهدي هذه القناة أقل من مشاهدي القنوات الأخرى عند مستوى معنوية $\alpha=0.10$

ـــ صياغة الفرضية : $H_0: P = 1/2$

$H_1: P < 1/2$

ـــ الاختبار الإحصائي :

$$n = 10 \Rightarrow (-8), (+2)$$

$n = 10$

ـــ إتخاذ القرار : من جداول الإشارة نجد أن :

$(+2), (-8)$

$P = \alpha = 0.10 \qquad P_\infty = (0.05469)$

من خلال المقارنة نجد أن : $P_\infty = (0.05469) < P = (0.10)$

وبالتالي نرفض فرضية العدم (H_0) ونقبل الفرضية البديلة (H_1) ونقول أن مشاهدي هذه القناة أقل من مشاهدي القنوات الأخرى .

ب ـ حالة وسيط عينة :

إذا كان لدينا مجتمعاً مستمراً (أي لاتوجد فيه مفردتان لهما نفس القيمة) وأردنا اختبار أن وسيطه (Med) يزيد عن قيمة معينة أو ينقص عنها، فإننا نختار عينة عشوائية حجمها (n) منه وننظر كم من مفردات هذه العينة يتجاوز القيمة المعينة أي بإشارة موجبة (+) وكم يقل عن تلك القيمة أي بإشارة سالبة (−) أو العكس، ولتوضيح ذلك لدينا التمرين التالي:

تمرين (10) :

تم وضع امتحان في كلية العلوم السياسية لأحد المقررات بحيث أن وسيط توزيع الزمن الذي ينتهي فيه الامتحان والخروج من القاعة يزيد عن ساعتين، وللتأكد من ذلك أخذت عينة عشوائية من الطلاب الذين دخلوا ذلك الامتحان حجمها (25) طالباً، فوجد أن (16) طالباً منهم انتهوا من ذلك الامتحان وخرجوا بعد مرور أكثر من ساعتين على بدء الامتحان. والمطلوب:

إختبر الفرض القائل بأن الزمن اللازم الذي استغرقه الطلاب في هذا الامتحان يتوزع حول توزيع وسيطه (M) الذي يتجاوز ساعتين من الزمن اللازم، عند مستوى معنوية $\alpha = 0.10$

ـ صياغة الفرضية : $H_0: Med = 2$

$H_1: Med > 2$

ـ الاختبار الإحصائي :

$(-9), (+16) \quad \Rightarrow \quad n = 25$

ـ إتخاذ القرار : من جداول الإشارة نجد أن :

$(-9), (+16) \quad \Rightarrow \quad n = 25$

$P = 0.11476$

من خلال المقارنة نجد أن : $P=(0.11476) > P=(0.10)$ وبالتالي نرفض فرضية العدم (H_0) ونقبل الفرضية البديلة (H_1) أي أن الزمن اللازم يتجاوز الساعتين .

جـ ـ حالة عينتين غير مستقلتين :

يقصد بالعينتين الزوجتين (عينتان غير مستقلتين) أي أن أحدهما مرتبطة بالأخرى (صورة أزواج من القيم) كأن تكون عينة واحدة أخذ لها القياس مرتين مرة قبل التجربة، ثم مرة أخرى بعد التجربة، مثل إنتاج العمال في الساعة قبل الإضراب وإنتاجهم لساعة بعد الإضراب، ويتم القياس للبيانات التي في صورة أزواج من خلال الفرق بين مفردة معينة والمفردة المقابلة لها، وإذا كان الفرق سالباً يعطي إشارة (-) وإذا كان الفرق موجباً يعطي اشارة (+) ويعطي الفرق المساوي للصفر عند التساوي وفي هذه الحالة، القيمة تسقط من الحساب ويعتبر (n) عدد الفروق التي لها إشارة موجبة أو سالبة. ونوضح ذلك من خلال التمرين (11) التالي :

تمرين (11) :

اختبر (11) شخصاً بطريقة عشوائية لمعرفة فيما إذا كان لنظام تغذية معين أثر على تخفيض الوزن قبل وبعد استخدام نظام التغذية عند مستوى معنوية $\alpha=0.10$، فأعطت النتائج الآتية:

رقم الشخص	1	2	3	4	5	6	7	8	9	10	11
الوزن قبل (x)	75	63	40	80	75	56	60	75	81	85	61
الوزن بعد (y)	77	60	41	82	73	56	69	79	86	93	63

والمطلوب :

اختبر الفرض القائل بعدم فعالية نظام التغذية مقابل الفرض القائل بأن نظام التغذية أثرّ على تخفيض الوزن بعد استخدام النظام .

الحل :

ـ صياغة الفرضية : $H_0 : P = 1/2$

$H_1 : P < 1/2$

ـ الاختبار الاحصائي :

رقم الشخص	1	2	3	4	5	6	7	8	9	10	11
الوزن قبل (x)	75	63	40	80	75	56	60	75	81	85	61
الوزن بعد (y)	77	60	41	82	73	56	69	79	86	93	63
الاشارة	+	-	+	+	-	0	+	+	+	+	+

نضع إشارة (+) في حالة زيادة الوزن ونضع إشارة (−) في حالة انخفض الوزن ونضع (0) في حالة التساوي قبل وبعد، وتستبعد القراءة رقم (6) من الدراسة ونتعامل فقط مع (10) حالات .

إتخاذ القرار :

$$(+8), (-2) \quad \Rightarrow \quad n=10$$
$$P = 0.0569$$

من خلال المقارنة نجد أن : $P=(0.0569) < P= (0.10)$

وبالتالي نرفض فرضية العدم (H_0) ونقبل الفرضية البديلة (H_1) أي يوجد تأثير لهذا النظام في التغذية على تخفيض الوزن .

حالة خاصة : تقريب اختبار الإشارة من التوزيع الطبيعي المعياري :

عندما يكون حجم العينة المدروسة كبيراً فإنه يمكن الاستغناء عن القيمة الفعلية (P) المستمدة من جداول اختبار الإشارة ونستخدم جداول (Z) المستمدة من التوزيع الطبيعي المعياري وفي الأغلب يعتبر حجم العينة (n) في هذه الحالة كبيراً إذا لم تقل عن (10) وبالتالي توزيع عدد الإشارات الموجبة تابعاً للتوزيع الطبيعي بمتوسط (n/2) وبتباين (n/4)، وفكرة التقريب من التوزيع الطبيعي جاءت من أن اختبار الإشارة قائم على مفهوم النجاح والفشل وعدد مرات النجاح والفشل (P), (1-P) لذلك يتبع التوزيع الثنائي $C_k^n P^k q^{n-k}$ وبالتالي :

القيمة المتوقعة لمتوسط توزيع ثنائي الحدين = $E(x) = np = \mu$ المتوسط

الانحراف المعياري لثنائي الحدين $\sigma = \sqrt{npq}$

ونحن نعلم بالإمكان تقريب توزيع ثنائي الحدين (التوزيع المتقطع) إلى التوزيع الطبيعي (Z) (التوزيع المستمر) عند زيادة عدد مرات التجربة، وعليه يمكن الحصول على (Z) بطرح المتوسط والقسمة على الانحراف المعياري ثم إجراء المقارنة مع ($Z_{\alpha/2}$) من جدول التوزيع الطبيعي المعياري لتحديد قبول أو رفض فرضية العدم (H_0)، فإذا رمزنا لعدد الإشارات الموجبة

في العينة بالرمز (T) فيمكن إيجاد (Z) بافتراض أن عدد (n) كبيراً على النحو التالي:

$$Z = \frac{T - n/2}{\sqrt{n/4}}$$

تمرين (12) :

نعود إلى التمرين (11) السابق والمتعلق بنظام التغذية والأثر على تخفيض الوزن، ونستخدم التوزيع الطبيعي كتوزيع تقريبي لاختبار الإشارة لاختبار تأثير نظام التغذية على تخفيض الوزن عند مستوى معنوية (α=0.10) ؟

الاختبار الإحصائي : (+8), (-2) $\Rightarrow$ n=10

$$Z = \frac{8 - n/2}{\sqrt{10/4}} = \frac{8-5}{\sqrt{2.5}} = 1.897$$

إتخاذ القرار : من خلال جداول (Z) نجد أن :

$Z_{\alpha=0.10}$ = 1,28 الاختبار من طرف واحد

من خلال المقارنة نجد أن (1.28) < (1.897) Z

وبالتالي نرفض فرضية العدم (H_0) ونقول ليس لهذا النوع من نظام التغذية أي تأثير على تخفيض الوزن.

ملاحظة: في بعض الأحيان لزيادة الدقة في التقريب يستخدم ما يسمى بمعامل الاتصال (عند الانتقال من التوزيع الثنائي إلى التوزيع الطبيعي) وتصبح العلاقة :

$$Z = \frac{T - n/2 - 1/2}{\sqrt{n/4}} \qquad \text{عندما } T > n/2$$

$$Z = \frac{T - n/2 + 1/2}{\sqrt{n/4}} \qquad \text{عندما } T < n/2$$

لو عدنا للتمرين (11) السابق، المطلوب : اختبر باستخدام معامل الاتصال:

الاختبار الاحصائي:

$$Z = \frac{T - n/2 \pm 1/2}{\sqrt{n/4}} = \frac{8 - 5 - 1/2}{\sqrt{2.5}} = 1.581$$

إتخاذ القرار : Z (1.581) > Z (1.28) وبالتالي نفس القرار السابق نرفض (H_0)

$\alpha = 0.10$

تمرين (13) :

تم استخدام عقار جديد لعلاج مرض السكر لمعرفة فيما إذا كان له تأثير على ضربات القلب فأعطى العلاج لـــ (12) مريضاً وكان عدد ضربات القلب في الدقيقة لكل منهم على النحو التالي:

رقم المريض	1	2	3	4	5	6	7	8	9	10	11	12
قبل العلاج	76	80	91	75	81	77	79	82	88	81	78	85

بعد العلاج	78	81	92	74	84	77	78	83	83	80	79	85

والمطلوب : اختبر الفرض القائل بأن للعقار تأثير على ضربات القلب وذلك باستخدام التوزيع الطبيعي كتوزيع تقريبي لاختبار الإشارة ($\alpha = 0.01$)؟

ـ صياغة الفرضية : $H_0 : P_1 = P_2 = 1/2$

$H_1 : P_2 > P_1$

الاختبار الإحصائي :

رقم المريض	1	2	3	4	5	6	7	8	9	10	11	12
قبل العلاج	76	80	91	75	81	77	79	82	88	81	78	85
بعد العلاج	78	81	92	74	84	77	78	83	83	80	79	85
الاشارة	+	+	+	-	+	0	-	+	-	-	+	0

تم استبعاد قرائتين : n=10 وعدد الإشارات الموجبة 6 أي :

$$n=10 \quad \Rightarrow \quad (-4), (+6)$$

$$Z = \frac{T - n/2}{\sqrt{n/4}} = \frac{6-5}{10/4} = 0.632$$

إتخاذ القرار :

من جداول (Z) نجد أن : Z(2.33) (اتجاه واحد)

$\alpha = 0.01$

من خلال المقارنة نجد أن : Z (0.63) < Z (2.33) وبالتالي نقبل فرضية العدم (H_0) وليس للعقار تأثير .

وبناءً على ما سبق شرحه، نشير إلى الملاحظات الآتية :

1 ــ أكثر استخدامات اختبار الإشارة يكون في اختبار :

ــ الفرض القائل بعدم فعالية نظام معين أو فعالية نظام معين (التغذية مثلاً)؛

ــ الفرض القائل بعدم تأثير أو تغير نظام معين (إنتاجية، دواء) أو تأثير وتغير نظام معين؛

2 ــ معظم اختبارات الإشارة قائمة على عينتين مرتبطتين وغير مستقلتين (قبل وبعد) أي البيانات في صورة أزواج للقيم؛

3 ــ في حالة عدم توفر جداول الإشارة يمكن استخدام جداول ثنائي الحدين (في نهاية الكتاب).

- **اختبار مان ويتني Mann Whitney :**

يستخدم هذا الاختبار عند الرغبة في فحص الفرق بين عينتين مختارتين ومستقلتين عن بعضهما. ويعتبر هذا الاختبار البديل الآخر لاختبار (t) في حالة عدم معرفة التوزيع الاحتمالي الذي تتبعه الظاهرة المدروسة، ويقوم هذا الاختبار على تحديد الرتبة المناظرة لكل قراءة بنفس أسلوب الرتب عند حساب معامل ارتباط الرتب (سبيرمان)، ومن ثم المقارنة بين مجموعات الرتب والهدف من هذا الاختبار هو إجراء المقارنة لمعرفة فيما إذا كانت إحدى مجموعات الرتب تقل بصورة ملموسة عن المجموعة الأخرى، وسنتناول هنا، حالة عينتين مستقلتين:

ــ حالة عينتين مستقلتين :

لقد تناولنا اختبار مجتمع معين وأيضاً مجتمعين في صورة أزواج للقيم، أي مرتبطين وغير مستقلين، أما الآن فسوف نتناول كيفية اختبار مجتمعين أثنين مستقلين (أي لايؤثر أحدهما على الآخر) ويشترط في

المجتمعين محل الدراسة، أن يكونا متطابقين من حيث شكل التوزيع وحجمه، وبالتالي إمكانية اختبار النزعة المركزية للمجتمعين، أي فيما إذا كان لهما نفس التمركز وفي الغالب يؤخذ الوسيط وفي حالة التماثل ممكن استخدام الوسيط والوسط الحسابي وإجراء الاختبار يتم على النحو التالي :

إذا كان لدينا مجتمعان متماثلان إحدهما (x) ووسيطه (M_x) والآخر (y) ووسيطه (M_y) فإنه يمكن إختبار وسيطي المجتمعين (M_x) و (M_y) فيما إذا كان ($M_x > M_y$) أو ($M_x < M_y$) أي ($Mx \neq My$)، تؤخذ عينة عشوائية من كل مجتمع من المجتمعين، نختار الرمز (x) لأحد المجتمعين ويفضل أن يكون المجتمع الذي سحبت منه العينة الأصغر (n) و (y) المجتمع الآخر الذي سحبت منه العينة الأكبر (m)، والحجم الكلي للعينتين هو ($N=n+m$).

تؤخذ عينتا المجتمعين وتمزجان معاً مع الاحتفاظ بالمقدرة على معرفة مصدر كل مفردة في العينة الكبيرة التي حجمها (N) ومن ثم يتم ترتيب هذه العينة المشتركة. ثم نجمع الرتب الخاصة، أي نجمع الرتب لمفردات العينة القادمة من المجتمع (x) ويرمز لها (Tx) وأيضاً بنفس الأسلوب (Ty) أو من خلال : ($Ty=1/2\ N(N+1)-Tx$) بعد الحصول على قيمة Tx يتم الكشف عن قيمة (P) المقابلة لها في جداول مان ويتني من خلال (m,n) بحيث (Tx) على يمينها و (Ty) على يسارها ومن ثم المقارنة بين (P) ومستوى المعنوية (α) لقبول أو رفض الفرضية H_0 . مع ملاحظة أنه يتم استخدام جداول مان ويتني، عندما يكون حجم العينة ($(n, m)\leq 10$) ولتوضيح ذلك لدينا التمرين التالي:

تمرين (14) :

في دراسة لمعرفة التأثير الجانبي لدواء جديد مطوّر في رفع ضربات القلب عند المرضى الذين يعالجون به عن الدواء القديم الذي يستخدم في العلاج لحالات المرضى تلك، أخذت عينة عشوائية متجانسة لهؤلاء المرضى حجمها (10) مرضى ووزعت عشوائياً إلى مجموعتين في كل مجموعة (5) مرضى، عولجت كل واحدة من هاتين المجموعتين عشوائياً بأحد النوعين من هذا الدواء. حيث عولجت إحدى المجموعتين بالدواء القديم و الأخرى بالدواء المطوّر، ثم قيست الزيادة في ضربات القلب الناتجة عن العلاج بأحد الدوائين فكانت النتيجة كما يلي :

رقم المفردة	1	2	3	4	5
الزيادة في ضربات القلب (الدواء القديم)	15	22	23	11	27
الزيادة في ضربات القلب (الدواء الجديد)	26	25	28	18	29

المطلوب :

اختبر الفرض القائل بأن الدواء الجديد يؤدي إلى ارتفاع ضربات القلب لدى المرضى المعالجين به، عند مستوى معنوية ($\alpha = 0.10$) .

الحل :

ـ صياغة الفرضية :

$H_0 : Mx = My$
$H_1 : Mx < My$

ـ الاختبار الإحصائي :

رقم المفردة	1	2	3	4	5
الدواء القديم (x)	15	22	23	11	27
Tx (الرتب)	2	4	5	1	8

الدواء الجديد (y)	26	25	28	18	29
Ty (الرتب)	7	6	9	3	10

$\sum T_x = 2 + 4 + 5 + 1 + 8 = 20$ عينة المجتمع (x)

إتخاذ القرار :

من خلال جداول مان ويتني نجد أن :

N: n=5 , m=5 , P = 0.0754

من خلال المقارنة نجد أن : P = (0.0754) < P= (0.10) وبالتالي نقبل الفرضية البديلة (H_1) ونرفض فرضية العدم (H_0) أي الدواء الجديد يسبب ارتفاع في ضربات القلب.

ـ تقريب اختبار مان ويتني بالتوزيع الطبيعي المعياري :

ويتم ذلك عندما تكون العينات كبيرة (n, m=15)، فإنه يمكن تقريب هذا الاختبار باستخدام التوزيع الطبيعي المعياري للحصول على القيمة المعيارية (Z). وذلك على النحو التالي:

$$Z = \frac{U - \frac{(n_1)(n_2)}{2}}{\sqrt{\frac{(n_1)(n_2)(n_1 + n_2 + 1)}{12}}}$$

حيث أن :

$$U = (n_1)(n_2) + \frac{n_1(n_2 + 1)}{2} - \sum T_x$$

ولتوضيح كيفية إجراء هذا الاختبار نشرح التمرين (15) التالي :

تمرين (15):

رغب أحد الباحثين دراسة أثر مشاهدة البرامج التلفزيونية على كمية المعلومات العامة لدى طلبة الجامعة فقام باختيار مجموعتين عشوائيتين من طلبة الجامعة حيث أخضعت المجموعة الأولى (A) لمشاهدة برامج تلفزيونية لمدة شهر، والمجموعة الثانية (B) لم تخضع لمشاهدة أي برنامج تلفزيوني، وبعد الإنتهاء طُبق اختبار معلومات عامة على المجموعتين فكانت النتائج في الجدول الآتي :

رقم الطالب	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
المجموعة (A) (x)	60	62	47	54	60	48	64	50	54	49	-	-	-	-	-
المجموعة (B) (y)	65	62	59	62	67	51	56	68	69	45	66	64	55	60	70

المطلوب :

اختبر الفرض القائل بعدم وجود تأثير للبرامج التلفزيونية على كمية المعلومات العامة لدى الطلبة، عند مستوى معنوية (α=0.05) .

الحل :

ـ صياغة الفرضية : $H_0 : \sum_{Tx} = \sum_{Ty}$

$H_1 : \sum_{Tx} \neq \sum_{Ty}$

ـ الاختبار الإحصائي :

$$Z = \frac{U - \frac{(n_1)(n_2)}{2}}{\sqrt{\frac{(n_1)(n_2)(n_1 + n_2 + 1)}{12}}}$$

رقم الطالب	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Σ
المجموعة (A) (x)	60	62	47	54	60	48	64	50	54	49	-	-	-	-	-	89.5
الرتب Tx	13	16	2	7.5	13	5	18.5	5	7.5	4	-	-	-	-	-	-
المجموعة (B) (y)	65	62	59	62	67	51	56	66	69	45	66	64	55	60	70	
الرتب Ty	20	16	11	16	22	6	10	23	24	3	21	18.5	9	13	25	235.5

$$U = (n_1)(n_2) + \frac{n_1(n_2+1)}{2} - \sum T_x$$

$$\sum T_x = 89.5 \quad , \quad \sum T_y = 235.5$$

$$U = (10)(15) + \frac{10(15+1)}{2} - 89.5 = 115.5$$

$$Z = \frac{115.5 - \frac{(10)(15)}{2}}{\sqrt{\frac{(10)(15)(10+15+1)}{12}}} = 0.80$$

إتخاذ القرار :

من جداول (Z) نجد أن : $Z_{0.05} = 1.96$

من خلال المقارنة نجد أن : $Z(0.80) < Z(1.96)$ وبالتالي نقبل فرضية العدم (H_0) ونقول

$\alpha=0.05$

لايوجد تأثير للبرامج التلفزيونية على كمية المعلومات للمجموعتين .

أمثلة غير محلولة

1 ــ لمعرفة فيما إذا كان هناك فرق بين دخل فئتين من عمال البناء وعمال السباكة، أخذت عينة عشوائية من (11) عامل بناء و (10) عمال سباكة، وتم تسجيل دخل كل منهم، كما هو موضح في الجدول التالي (بآلاف الوحدات النقدية) :

رقم العامل	1	2	3	4	5	6	7	8	9	10	11
الدخل السنوي لعمال السباكة	20	20.5	21.5	21	95	22	27.5	30	15.5	15	-
الدخل السنوي لعمل البناء	8	8.5	10	10.5	19.5	19	18	17.5	17	14.5	14

المطلوب :

اختبر الفرض القائل بأنه ليس هناك فرق بين دخل عمال السباكة و دخل عمال البناء عند مستوى معنوية ($\alpha=0.05$) . (يترك الحل للطالب).

2ــ لدراسة الفروق بين أجور العمال لشركتين، اختار باحث عينة عشوائية من العاملين في كل منهما وحصل من هاتين العينتين على النتائج الآتية :

رقم العامل	1	2	3	4	5	6	7	8	9	10	11
الشركة (A)	60	150	90	210	79	86	120	115	132	142	-
الشركة (B)	165	164	92	205	96	84	145	140	160	175	182

المطلوب : اختبر الفرض القائل بأنه ليس هناك فرق بين الأجور في الشركة (A) وعنها في الشركة (B) وذلك عند مستوى معنوية ($\alpha = 0.01$) ؟ (يترك الحل للطالب).

- **إختبار ولكوكسون Wilcoxon :**

إن القصور الأساسي في اختبار الإشارة يتركز في تجاهله التام لمقدار الفرق في كل زوج من القيم، وقد توصل ولكوكسون إلى اختبار يأخذ بعين الاعتبار مقدار الفرق في كل زوج من القيم (اختبار الإشارة والرتب) ، حيث يتطلب هذا الاختبار ترتيب القيم المطلقة للفروق ترتيباً تصاعدياً وتهمل أزواج القيم ذات الفروق الصفرية، مع العلم أن الفروق ترتب دون النظر لإشارتها، وبعد ذلك يتم حساب مجموع الرتب السالبة ومجموع الرتب الموجبة، إذ يطبق هذا الاختبار على هذين المجموعين، لبيان فيما إذا كان يوجد فرق بين مجموعتين مرتبطتين من القراءات .

هذا ويمكن استخدام اختبار ولكوكسون لاختبار عينتين زوجيتين لمعرفة فيما إذا كان هناك بينهما ارتباط أم لا (البيانات بصورة أزواج من القيم)، وهذا الاختبار مشابه لاختبار الإشارة ولكن يأخذ الرتب بعين الاعتبار، وبناءً عليه يمكن تقريب إختبار ولكوكسون بالتوزيع الطبيعي المعياري حيث يتم ذلك، عندما تكون العينات كبيرة [$(n,n) > 10$ أزواج)] فإنه يمكن تقريب هذا الاختبار باستخدام التوزيع الطبيعي المعياري للحصول على القيمة المعيارية (Z) على النحو التالي:

$$Z = \frac{w - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

حيث أن : w عبارة عن مجموع الرتب للإشارة الأقل تكراراً

n : عدد الفروق غير الصفرية.

ونوضح ذلك من خلال التمرين (16) التالي :

تمرين (16) :

رغب أحد الباحثين دراسة مدى تأثير استخدام طريقة جديدة لتدريس الإحصاء لطلاب السنة الثانية، فأخذ عينة من (10) أزواج من الطلبة المتكافئين، وبعد ذلك قام باستخدام الطريقة المعتادة لتدريس الاحصاء على المجموعة (A) والطريقة الجديدة على المجموعة (B)، وبعدها أجرى امتحاناً لهؤلاء الطلبة فكانت النتائج كما في الجدول التالي :

رقم الزوج	1	2	3	4	5	6	7	8	9	10
الدرجة بالطريقة القديمة (y)	53	61	59	63	41	64	55	54	60	45
الدرجة بالطريقة الجديدة (x)	75	65	72	58	48	62	59	59	69	68

المطلوب : اختبر الفرض القائل بعدم وجود فرق معنوي بين الطريقتين لتدريس الاحصاء باستخدام اختبار الإشارة والرتب (ولكوكسون) ، وذلك عند مستوى معنوية (α=0.01).

الحل :

صياغة الفرضية : $H_0: \sum T\,(x-y) = \sum T\,(x-y)$

$H_1 : \sum T\,(x-y) \neq \sum T\,(x-y)$

ـ الاختبار الإحصائي :

$$Z = \frac{w - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

رقم الزوج	1	2	3	4	5	6	7	8	9	10
الدرجة بالطريقة القديمة (y)	53	61	59	63	41	64	55	54	60	45
الدرجة بالطريقة الجديدة (x)	75	65	72	58	48	62	59	59	69	68
(x-y)	+22	+4	+13	-5*	+7	-2*	+4	+20	+9	+23
الفرق بالقيمة المطلقة	22	4	13	5	7	2	4	20	9	23
رتب الفرق T (x-y)	9	2.5	7	4*	5	1*	2.5	8	6	10

$$\sum T\,(x-y) = +9 + 2.5 + 7 + 5 + 2 + 2.5 + 8 + 6 + 10 = 50$$

الموجبة

$$\sum T\,(x-y) = 4 + 1 = 5$$ السالبة

نختار الإشارة الأقل تكراراً نجد أنها السالبة ، فيكون :

$$Z = \frac{5 - \frac{10(11)}{4}}{\sqrt{\frac{10(11)(21)}{24}}} = |-2,29|$$

إتخاذ القرار : من جداول (Z) نجد أن : $Z\,(2.58)$

$\alpha = 0.01$

من خلال المقارنة نجد أن : $Z\,(2.29) < Z\,(2.58)$

$\alpha = 0.01$

وبالتالي نقبل فرضية العدم (H_0) ونقول لايوجد فرق معنوي بين الطريقتين .

5-7- الاختبارات الاحصائية اللامعلمية بين عدة عينات (اختبار كروسكال ــ واليس)

- **اختبار كروسكال ــ واليس Kruska - Wallis**

عند إجراء اختبارات الفروض المتعلقة بالفروق بين متوسطات ثلاثة مجتمعات أو أكثر افترضنا أن المجتمعات الأصلية تتبع التوزيع الطبيعي، وأن تبايناتها مجهولة ومتساوية فإذا تحققت هذه الافتراضات فإنه يمكننا استخدام

اختبار (F) أما إذا كانت الفروق المشاهدة لمعرفة بين متوسطات العينات راجعة للصدفة فإنها تشير إلى فروق معنوية بين المجتمعات المتناظرة. أما إذا لم تتحقق هذه الافتراضات، فيمكن استخدام طريقة لامعلمية ملائمة تسمى اختبار مجموع الرتب (اختبار كروسكال ــ واليس) .

ــ حالة عينات مستقلة : يستخدم هذا الاختبار لاختبار مدى تطابق توزيعات عدة مجتمعات ويرمز له بالرمز (H) ويتطلب هذا الاختبار افتراض أن توزيعات المجتمعات مستمرة ولايتطلب أي افتراض آخر، وهذا الاختبار يتبع توزيع (χ^2). والاختبار الاحصائي يتم من خلال :

$$H = \frac{12}{N(N+1)} \sum_{k=1}^{K} \frac{R_k^2}{nk} - 3(N+1)$$

حيث أن :

K : عدد العينات (المجموعات) (k=1,2, ... , k)

P_k : مجموع رتب كل عينة (مجموعة)

nk : حجم كل عينة (مجموعة)

N : عدد جميع المشاهدات في العينات .

ولتوضيح ذلك لدينا التمرين التالي :

تمرين (17) :

قامت إحدى الشركات الكبرى بتصنيف العاملين فيها إلى ثلاث مجموعات وفقاً لمدة خبرتهم فيها، وقد رغبت إدارة الشركة في معرفة مدى اختلاف الدافع المعنوي بين هذه المجموعات الثلاث فقامت باختيار عينة عشوائية من كل مجموعة مكونة من (10) موظفين حيث قسمت دوافعهم

المعنوية بمقياس ذي (40) نقطة، فحصلت الإدارة على البيانات الواردة في الجدول التالي:

المجموعة A	21	23	25	27	31	33	30	35	13	39
المجموعة B	20	22	26	28	30	32	19	19	36	38
المجموعة C	15	16	17	24	18	29	30	30	37	24

المطلوب : هل يمكننا استنتاج تساوي الدافع المعنوي بين هذه المجموعات الثلاث عند (α=0.01) .

الحل :

ـ صياغة الفرضية :

$$H_0 : \sum R_1 = \sum R_2 = \sum R_3$$
$$H_1 : \sum R_1 \neq \sum R_2 \neq \sum R_3$$

ـ الاختبار الإحصائي :

$$H = \frac{12}{N(N+1)} \sum_{k=1}^{K} \frac{R_k^2}{nk} - 3(N+1)$$

الرقم	1	2	3	4	5	6	7	8	9	10	$\sum$
المجموعة A	21	23	25	27	31	33	30	35	13	39	-
الرتب	9	11	14	16	22	24	20	26	1	30	173
المجموعة B	20	22	26	28	30	32	19	34	36	38	-
الرتب	8	10	15	17	20	23	7	25	27	29	181

المجموعة C	15	16	17	24	18	29	30	14	37	24	-
الرتب	3	4	5	12.5	6	18	20	2	28	12.5	111

$R_3=111$, $R_2 = 181$, $R_1 = 173$

$$H = \frac{12}{30(31)}\left(\frac{(173)^2}{10}+\frac{(181)^2}{10}+\frac{(111)^2}{10}\right)-3(30+1)=3.78$$

ـ **إتخاذ القرار :** من جداول χ^2 نجد أن :

$\chi^2_{(0.01\ ,2)} = 9,210$

$V = k - 1 = 3-1 = 2$

من خلال المقارنة نجد أن $\chi^2=(3.78) < \chi^2_{(0.01\ ,2)}=(9.210)$ وبالتالي نقبل فرضية العدم (H_0) وهذا يعني تساوي الدافع المعنوي بين هذه المجموعات الثلاث .

ملاحظات :

1 ـ اختبار مجموع الرتب (كروسكال ـ واليس) هو دائماً اختبار من اتجاه واحد ومن اليمين ؛

2 ـ فإننا نكون أمام حالتين :

أ ـ تركز الرتب الكبيرة في مجموعة واحدة أو في عدد قليل جداً من المجموعات وبالتالي تكون قيمة H المحسوبة صغيرة نسبياً، وهذا يدل على عدم اختلاف توزيعات المجتمعات؛

ب ــ تركز الرتب الصغيرة في المجموعات الأخرى، فيكون عندها قيمة H كبيرة نسبياً مما يعكس هذا اختلاف في توزيعات المجتمعات.

6-7- الاختبارات الاحصائية اللامعلمية للمقارنة بين عينتين للبيانات الاسمية والرتبية (اختبار كولموجروف، سميرنوف، اختبار ثنائي الحدين، اختبار فشر، اختبار الوسيط):

- **اختبار كولموجروف ــ سميرنوف Kolmogorov – Smironov**

ــ حالة عينة واحدة : يستعان بهذا الأسلوب في حالة البيانات الإسمية للتحقق من صحة فرضية العدم (H_0) القائلة بأن الفروق بين (Q و E) كانت نتيجة الحظ والصدف، ويعتمد هذا الاختبار على جداول (χ^2)، علماً بأنه أكثر دقة من اختبار χ^2 عند التعامل مع حجم عينة ($n \leq 30$) ويعطى الاختبار بالعلاقة التالية :

$$K.S = \left| \frac{f_Q}{n} - \frac{fe}{n} \right|$$

حيث إن :

n :حجم العينة

f_Q : التكرار المتجمع الحقيقي

fe :التكرار المتجمع المتوقع

E : التكرار الحقيقي

وستكون (k.S) عبارة عن أكبر فرق مطلق بين النسب المتجمعة الحقيقية والمتوقعة ولتوضيح هذا الاختبار لدينا التمرين التالي :

تمرين (18) :

قام أحدالباحثين بدراسة لمعرفة رغبة أطفال إحدى رياض في اختيارهم للعبة معينة، وبغية تحقيق ذلك وضعت اللعبة بألوان مختلفة وكانت النتائج التي حصل عليها في الجدول التالي :

اللون	أبيض	أصفر	أحمر	أزرق	أخضر	Σ
العدد	3	9	16	1	1	30

والمطلوب :

اختبر الفرض القائل بأن اختيار الطفل للعبة لاعلاقة له باللون عند مستوى معنوية (α=0.05)

الحل :

صياغة الفرضية : فرضية الاستقلال ــ ليس هناك علاقة بين اختيار اللعبة ولون اللعبة.

ــ الاختبار الإحصائي :

$$K.S = \left|\frac{f_o}{n} - \frac{fe}{n}\right|$$

حساب التكرارات المتوقعة (E) : 30/5 = 6

الجدول المساعد :

الألوان (الخلايا)	Q	ت.م لـ f_o	ت.م.ن لـ Q	E	ت.م لـ E	ت.م.ن لـ E	$\left\|\frac{f_o}{n} - \frac{fe}{n}\right\|$
الأبيض	3	3	3/30	6	6	6/30	3/30
الأصفر	9	12	12/30	6	12	12/30	0
الأحمر	16	28	28/30	6	18	18/30	10/30
الأزرق	1	29	29/30	6	24	24/30	5/30
الأخضر	1	30	30/30	6	30	30/30	0

من خلال الجدول نجد أن أكبر فرق مطلق هـو (10/30=0.33) أي أن : (k.s = 0.33)

ـ إتخاذ القرار : من جداول (k.S) نجد أن :

$K.s_{(\alpha,N)} = K_{(0.05\,.\,30)}=0,24$

من خلال المقارنة نجـد أن : K.s= 0.33 > K.s= 0.24 وبالتـالي نرفض فرضية الاستقلال وهناك فروق جوهرية (معنوية) في اختيار الأطفال للعب باختلاف اللون، بمعنى أن هناك علاقة بين اختيار الطفل للعبة ولونها.

ـ حالة عينتين مستقلتين : يستعان بهذا الأسلوب لاختبـار الفـروق بـين عينتين، عندما تكون البيانات الخاصة بين أحد المتغيرين إسمية والثانية رتبية، ويمكن استخدامه في حالة البيانات الإسمية لكلا المتغيرين، ويتم الاختبار على النحو التالي :

$$K = \left[\frac{f_{G1}}{n_1} - \frac{f_{G2}}{n_2}\right]\sqrt{\frac{(n_1)(n_2)}{n_1 + n_2}}$$

$$K = F\sqrt{\frac{(n_1)(n_2)}{n_1 + n_2}}$$

حيث أن :

f_{G1} : التكرار التجميعي الصاعد للمجموعة الأولى

f_{G2} : التكرار التجميعي الصاعد للمجموعة الثانية

n_1,n_2 : حجم العينات

F : أكبر فرق مطلق بين التكرارين المتجمعين النسبيين .

ولتوضيح ذلك لدينا التمرين التالي :

تمرين (19) :

بفرض أن كلية العلوم السياسية بإحدى الجامعات تقبل الحاصلين على الثانوية العامة (علمي، أدبي)، وبهدف معرفة معنوية الفروق بين نتائج الطلبة حملة الثانوية العامة العلمية وحملة الثانوية العامة الأدبية في مقرر الإحصاء كانت النتائج على النحو التالي :

التقدير / التخصص	ضعيف جداً	ضعيف	مقبول	جيد	جيد جداً	ممتاز	Σ
أدبي	12	19	15	7	6	2	61
علمي	4	4	9	11	16	20	64

المطلوب :

اختبر الفرض القائل بأنه ليس هناك فروقاً معنوية في تقديرات مقرر الإحصاء بين الطلاب (العلمي والأدبي) عند مستوى معنوية ($\alpha = 0.01$) .

الحل :

ـ **صياغة الفرضية :** $H_0 = G_1 = G_2$

$H_1 = G_1 \neq G_2$

ـ **الاختبار الاحصائي :**

$$K = \left[\frac{f_{G1}}{n_1} - \frac{f_{G2}}{n_2}\right]\sqrt{\frac{(n_1)(n_2)}{n_1 + n_2}}$$

الجدول المساعد :

التقدير	الأدبي Q_1	ت.م f_{G1}	ت.م.ن f_{G1}/n_1	العلمي f_{G2}	ت.م f_{G2}	ت.م.ن f_{G2}/n_2	الفرق المطلق (F)
ضعيف جداً	12	12	12/61	4	4	4/64	0.13
ضعيف	19	31	31/61	4	8	8/64	0.38

مقبول	15	46	46/61	9	17	17/64	0.49
جيد	7	53	53/61	11	28	28/64	0.43
جيد جداً	6	59	59/61	16	44	44/64	0.28
ممتاز	2	61	61/61	20	64	64/64	0

$$K = F\sqrt{\frac{(n_1)(n_2)}{n_1 + n_2}}$$

ـ **إتخاذ القرار** : من خلال الجدول نجد أن أكبر فرق مطلق هـو (0.49) أي أن :

$$F = 0.49$$

وبالتالي نجد أن :

$$K = 0.49\sqrt{\frac{(61)(64)}{61+64}} = 2.74$$

من جداول (k.S) نجد أن : $K_{0.01} = 1.22$ ، ومن اتجاهين :

من خلال المقارنة نجد أن : $K = 2.74 > K = 1.22$ وبالتالي نرفض فرضـية العدم ونقول أن هناك فروقاً معنوية بين الطلاب العلمي والأدبي في تقديراتهم في مقرر الإحصاء .

• **اختبار ثنائي الحدين Binomiale :**

حالة عينة واحدة (بيانات إسمية) : يستعان بهذا الاختبار عندما يكون لـدينا عينة واحدة وقد اختيرت عشوائياً وتم الحصول منها على بيانـات بصـورة إسمية وثنائية التقسيم مثل مجموعة جاءت إجاباتها على صـورة (نعـم، لا)، (موافق، غير موافق)، (حضر، لم يحضر)، (صح، خطأ)، ويتم الاختبار مـن خلال العلاقة التالية :

$$K = \frac{n_1 - NP_1}{\sqrt{NP_1P_2}}$$

n_1 : عدد الإجابات للبديل الأول

N : حجم العينة الكلية

P_1 : احتمال ظهور إجابات البديل الأول = 0.5

P_2 : احتمال ظهور إجابات البديل الثاني = 0.5، وهنا نميز حالتين:

الحالة (أ): عندما يكون n>20 (في هذه الحالة نستخدم جداول Z)، ولتوضيح ذلك لدينا التمرين التالي :

تمرين (21):

في دراسة على عينة من الأطفال (n=30) من عمر أربع سنوات، بهدف التفريق بين شكل المكعب وشكل الكرة، أظهرت التجربة أن (20) طفلاً استطاعوا التفريق بين الشكلين في حين فشل الباقي، المطلوب : اختبر الفرض القائل بأن هؤلاء الأطفال تمكنوا من التمييز بين الشكلين عند مستوى معنوية $\alpha = 0.05$ ؟.

الحل :

ـ صياغة الفرضية : $H_0 = P_1 = P_2 = 1/2$

$H_1 \neq P_2 \neq 1/2$

ـ الاختبار الإحصائي :

$$K = \frac{n_1 - NP_1}{\sqrt{NP_1P_2}}$$

$$K = \frac{20 - (30)(0.5)}{\sqrt{(30)(0.5)(0.5)}} = 1.82$$

ـ **إتخاذ القرار :** من جداول (Z) نجد أن :

$Z_{(0.05)}=1.65$ (اتجاه واحد)

من خلال المقارنة نجد أن :

$$Z(1.82) > Z(1.65)$$

وبالتالي يمكن القول أن هناك فرقاً معنوياً بين عدد من استطاع التمييز ومن لم يستطع، أي أن الأطفال في هذا العمر يمكنهم التفريق بين هذين الشكلين .

الحالة (ب) عندما يكون $n<20$ (في هذه الحالة نستخدم جداول ثنائي الحدين):

n_1 : عدد أفراد البديل الأول

N : حجم العينة

ولتوضيح كيفية إجراء الاختبار في هذه الحالة لدينا التمرين التالي :

تمرين (22):

تم توجيه سؤال لعينة عشوائية حجمها (n=16) طالباً من طلاب كلية العلوم السياسية في إحدى الجامعات حول المواعيد المفضلة لديهم لحضور محاضرات الإحصاء (8-10) صباحاً أم (11-1) ظهراً فجاءت النتائج تشير إلى تفضيل (9) طلاب للفترة (8-10) . **المطلوب :** اختبر الفرضية القائلة بأن الطلاب لايفضلون حضور محاضرات الإحصاء من (8-10) على الفترة (11-1) ظهراً عند مستوى ($\alpha=0.05$) .

الحل :

N=16 n=9

عدد البدائل = 2 ، احتمال البديل = 1/2

ـ **صياغة الفرضية :**

$$H_0: P_1 = P_2 = 1/2$$

$H_1 : P_1 \neq P_2 \neq 1/2$

ـ **الاختبار الإحصائي :**

$P = \alpha = 0.05$

ـ **إتخاذ القرار :** من خلال جداول ثنائي الحدين عندما : $n_1=9$ $N=16$
$P=0.5$

نجد أن القيمة المقابلة (0.1746)

من خلال المقارنة نجد أن : $P=0.1746>P= 0.05$ وبالتالي نرفض فرضية العدم (H_0) ونقول أن الطلاب لايفضلون وقتاً على الآخر لدراسة الإحصاء في ضوء بيانات العينة.

تمرين (23) :

في دراسة بإحدى المدارس الخاصة عن لون الزي (الملبس) المرغوب الارتداء كانت العينة (n=19) طالباً عرضت عليهم الألوان (بني، أبيض، أزرق) وأشارت النتائج إلى تفضيل (11) طالباً للون الأبيض، والمطلوب: اختبر الفرض القائل بأن إجابات الطلاب تمثل إجابات المجتمع الذي تم اختيارهم منه ، وذلك عند مستوى معنوية (α =0.01) ؟.

الحل :

$N = 19$ $n_1 = 11$ عدد البدائل = 3

احتمالي البديل 1/3 = 0.33

ـ **صياغة الفرضية :**

$H_0 : P_1 = P_2 = P_3 = 1/3$
$H_1 : R \neq P_2 \neq P_3 \neq 1/3$

ـ **الاختبار الاحصائي :**

ـ **إتخاذ القرار :** من خلال جداول ثنائي الحدين عندما :
$n_1=11$, $N=19$

احتمالات البدائل هي : 0.20 , 0.25 , 0.50

والاحتمال البديل الأقرب هو (0.25) ومن خلال الجدول نجد أن الاحتمال المقابل لها (0.0018) .

من خلال المقارنة نجد أن : P (0.0018) < P(0.01) وبالتالي نرفض فرضية العدم (H_0)

ونقبل (H_1) أي البدائل غير متساوية أي أن نسبة المفضلين للون الأبيض هم الأغلبية .

- **اختبار فيشر Fisher وحالة مجموعتين مستقلتين :**

يستخدم هذا الاختبار للمقارنة بين عينتين مستقلتين (ذكور، إناث) مثلاً، و متغير اسمي ينقسم إنقساماً ثنائياً (ناجح، راسب)، بمعنى أن البيانات الخاصة بالمتغيرين المستقل (النوع) والتابع (التحصيل) إسمية ثنائية التقسيم، ويتم الاختبار من خلال العلاقة التالية :

$$F = \frac{[(a.d - b.c) \pm n/2]}{\sqrt{\frac{(a+b)(c+d)(a+c)(b+d)}{n-1}}}$$

مع العلم أن : (+n/2) عندما يكون (a.d - b.c > 0) و (- n / 2) عندما يكون (a.d – b.c < 0)

ولتوضيح هذا الاختبار لدينا التمرين التالي :

تمرين (24) :

صممت تجربة على مجموعتين، استخدم في المجموعة الأولى نظام الحوافز ولم يستخدم في المجموعة الثانية، وذلك بهدف معرفة رأي العمال في تقديرهم لرؤسائهم نتيجة الحوافز فكانت النتائج الآتية :

الرأي	تقدير	عدم تقدير	$\sum$
A	18	6	24
B	10	11	21
$\sum$	28	17	45

المطلوب :

اختبر الفرض القائل بأن للحوافز أثراً على رأي العاملين أي تقدير رؤسائهم عند مستوى معنوية (α=0.05)

ـ الاختبار الاحصائي :

$a = 18$ ، $b=6$ ، $c=10$ ، $d=11$ نعلم أن

نطبق : $a.d - b.c > 0$

نجد أن : $(18)(11) - (6)(10) > 0$

أي أن :

$$F = \frac{[(a.d - b.c) + n/2]}{\sqrt{\frac{(a+b)(c+d)(a+c)(b+d)}{n-1}}}$$

$$= \frac{[(18)(11) - (10)(6) + 45/2]}{\sqrt{\frac{(28)(17)(24)(21)}{45-1}}} = 1.56$$

ـ إتخاذ القرار :

من جداول (Z) نجد أن قيمة F(1.65) (اتجاه واحد)

α=0.05

من خلال المقارنة نجد أن : F (1.56) < F (1.65) وبالتالي لايوجد فروق معنوية بين تقدير الرؤوساء وعدم تقديرهم عند استخدام نظام الحوافز أو عدم استخدامه أي أن نظام الحوافز ليس له تأثير على الرأي في تقدير الرؤوساء .

- **إختبار الوسيط Median :**

ـ حالة عينتين مستقلتين : يستخدم هذا الاختبار عندما يرغب الباحث التحقق من معنوية الفروق بين وسيطي عينتين، ويعتمد هذا الاختبار على إيجاد الوسيط للعينتين موضع الدراسة ثم يحدد عدد المفردات التي تقع قيمتهم فوق الوسيط وكذلك يحدد عدد المفردات التي تقع قيمتهم مساوية أو أقل. وتصنف المفردات في جدول مزدوج (2 × 2) كما هو موضح بالشكل :

العينة	أعلى من الوسيط	أقل أو يساوي الوسيط	Σ
I	a	b	a+b
II	c	d	C+d
Σ	a+c	b+d	

ويتم الاختبار وفق العلاقة التالية :

$$\chi^2 = \frac{n[a.d - b.c]^2}{(a+b)(c+d)(a+c)(b+d)}$$

وبدرجة حرية واحدة .

ولتوضيح كيفية إجراء هذا الاختبار لدينا التمرين التالي :

تمرين (25) :

رغب باحث إجراء بحث يتعلق بأوزان الطلاب، فقام باختيار عينة عشوائية من طلاب كلية العلوم السياسية وعينة أخرى عشوائية من بقية طلاب الكليات الأخرى، وكانت البيانات على النحو التالي :

طلاب العلوم السياسية :

60	65	62	70	75	67	61	80	66	65	63	70	70	69	64	61

طلاب الكليات الأخرى :

70	71	75	59	60	67	62	79	62	68	74	71	70	65	82

المطلوب :

اختبر الفرض القائل بعدم وجود فروق معنوية بين أوزان طلاب كلية العلوم السياسية وطلاب بقية الكليات عند مستوى معنوية $(\alpha = 0.01)$؟

الحل :

ـ الطريقة الأولى : نقوم بدمج المجموعتين معاً فتصبح العينة الكلية (31) طالباً ثم ترتيب هذه المفردات من الأصغر إلى الأكبر أو بالعكس، ثم يتم تحديد المفردة أو المفردتين التي يسبقها عدد من المفردات يساوي عدد المفردات التي تليها، وإذا كان هناك درجتين يجب جمعهما والقسمة على (2) .

ـ الطريقة الثانية : يتم استخدامها إذا كانت مسألة الترتيب تحتاج إلى وقت وذلك بإنشاء جدول تكراري .

الفئات	f	ت.ت.ص
58-62	8	8
63-67	**8**	**16**
68-72	9	25
73-77	3	28
78-82	3	31

$$\sum f/2 = \frac{31}{2} = 15.5$$

$$Med = L + \left(\frac{\Sigma f/2 - \sum fe}{f_{me}}\right).C$$

$$= 63 + \left(\frac{15.5 - 8}{8}\right).5 = 67.7$$

نقوم بترتيب الجدول (2 × 2)

العينة	أكبر من الوسيط		الوسيط فأقل		Σ
العلوم السياسية	a	6	b	10	16
الكليات الأخرى	c	10	d	5	15
Σ	16		15		31

ـ الاختبار الإحصائي :

$$\chi^2 = \frac{n[a.d - b.c]^2}{(a+b)(c+d)(a+c)(b+d)}$$

$$\chi^2 = \frac{31[(6)(5) - (10)(10)]^2}{(16)(15)(16)(15)} = 2.64$$

ـ **إتخاذ القرار** : من جداول χ^2 نجد أن :

$\chi^2(3.841)$

$\alpha = 0.01$

$V=1$

من خلال المقارنة نجد أن : $\chi^2(3.841) > \chi^2(2.64)$ وبالتالي نقبل فرضية العدم (H_0)

$\alpha=0.01$

$V=1$

ونقول لاتوجد فروقاً معنوية بين أوزان طلاب كلية العلوم السياسية وبقية الكليات في الجامعة.

ـ **حالة عدة عينات مستقلة** : لقد بينا في السابق استخدام الوسيط للمقارنة بين عينتين مستقلتين، وبنفس الأسلوب يمكن إجراء المقارنة لعدة عينات مستقلة، وذلك وفق الخطوات الآتية:

1 ـ يتم دمج مفردات العينات موضع الدراسة وكأنها مفردات عينة واحدة.

2 ـ يستخرج وسيط مفردات هذه العينة الواحدة، إما بوضع المفردات من الأصغر إلى الأكبر أو بالعكس، ثم استخراج الوسيط على اعتبار أنه تلك المفردة التي يسبقها عدد من المفردات يساوي عدد المفردات التي تليها وذلك بعد الترتيب لهذه المفردات تصاعدياً أو تنازلياً ومن ثم تطبيق القانون لحساب قيمة الوسيط .

3 ـ تصنيف مفردات كل عينة كما كانت قبل الدمج في ضوء قيمة الوسيط بحيث نضع عوضاً عن المفردة إشارة (+) إذا كانت قيمة هذه الدرجة أكبر من الوسيط ونضع عوضاً عن المفردة إشارة (ـ) إذا كانت قيمة هذه الدرجة أصغر أوتساوي قيمة الوسيط.

4 ـ يحدد عدد تكرارات الإشارة (الموجبة) وعدد التكرارات للإشارة (السالبة) في كل عينة من العينات موضوع المقارنة قبل الدمج وتصنف التكرارات في جدول كما هو موضح :

العينات	+	−	Σ
I	a	b	a+b
II	c	d	c+d
III	e	f	e+f
	a+c+e	b+d+f	N

ولبيان كيفية إجراء هذا الاختبار لدينا التمرين التالي :

تمرين (26):

قامت إحدى الشركات الكبرى بتصنيف العاملين فيها إلى ثلاث مجموعات وفقاً لمدة سنوات الخدمة فيها، وطبقت اختباراً معيناً على هذه المجموعات بهدف التعرف على مدى الاختلاف بين هذه المجموعات فكانت النتائج الآتية :

المجموعةA	7	9	15	4	3	2
المجموعة B	8	5	9	3	11	8
المجموعة C	6	9	14	8	13	12

المطلوب : اختبر الفرض القائل بأنه ليس هناك فروقاً معنوية بين المجموعات الثلاث بالنسبة للاختبار .

الحل :

ـ الاختبار الإحصائي :

نقوم بدمج العينات الثلاث للحصول على الوسيط على النحو التالي :

رقم المفردة	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
المفردات	2	3	3	4	5	6	7	8	8	8	9	9	9	11	12	13	14	15

ترتيب الوسيط

$$k = \frac{n+1}{2} = \frac{18+1}{2} = 9.5 \Rightarrow$$

$$\text{وسيط}\ Med = \frac{8+8}{2} = 8$$

تصنيف البيانات يتم على النحو التالي :

العينة	+		−		Σ
	أكبر من الوسيط		أقل أو يساوي الوسيط		
A	a	2	b	4	6
B	c	2	d	4	6
C	e	4	f	2	6
Σ	8		10		18=N

حساب التكرارات المتوقعة (E) :

$$a = \frac{(6)(8)}{18} = 2.67 \qquad b = \frac{(6)(10)}{18} = 3.33$$

وهكذا لبقية الخلايا .....

$$\chi^2 = \left(\sum \frac{(Q-E)^2}{E} \right)$$

$$\chi^2 = \frac{(2-2.67)^2}{2.67} + \frac{(4-3.33)^2}{3.33} + \ldots\ldots + \frac{(2-3.33)^2}{3.33} = 1.79$$

إتخاذ القرار : من خلال جداول (χ^2) نجد أن :

χ^2 **(9.21)**

$\alpha = 0.01$

$V = (2-1)(3-1)=2$

من خلال المقارنة نجد أن : $\chi^2(1.79) < \chi^2(9.21)$ وبالتالي نقبل الفرضية القائلة بأنه لايوجد فروق معنوية بين المجموعات بالنسبة للاختبار .

7-7- بعض معاملات الارتباط للبيانات الإسمية والرتبية (Some Correlation Coefficient's) :

يعتبر استخدام معاملات الارتباط في البحوث الاجتماعية ذا أهمية كبيرة لمعرفة طبيعة العلاقات الموجودة بين المتغيرات، وإن معرفة مقدار معامل الارتباط الموجود بين متغيرين (x) و (y) يساعد على تحديد متانة العلاقة الموجودة بينهما واتجاهها أيضاً.

إن قيمة معامل الارتباط تتراوح بين (-1 ، +1)، فكلما اقترب مقدار معامل الارتباط من الواحد، هذا يدل على قوة العلاقة الارتباطية بين المتغيرين وبالعكس، وإذا كانت إشارة معامل الارتباط موجبة، هذا يدل على أن العلاقة موجبة (طردية)، أما إذا كانت الإشارة سالبة فيدل على أن العلاقة سالبة (عكسية)، وسنتعرض هنا إلى أهم المعاملات، ومنها:

- **معامل الارتباط سبيرمان للرتب Speerman Pank :**

يعتبر معامل الارتباط سبيرمان من الطرائق الإحصائية الهامة في قياس العلاقة بين متغيرين رتبيين، وبخاصة عندما ($n \geq 30$) حيث يقوم معامل ارتباط سبيرمان للرتب على مدى الفروق الموجودة بين الرتبتين الخاصتين

بدرجات كل مفردة من مفردات العينة ويعطى بالعلاقة التالية :

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d^2}{n(n-1)}$$

حيث أن d : الفروق بين حالات الرتب للمتغيرين المدروسين.

n : حجم العينة

الاختبار الإحصائي :

ـ إذا كان حجم العينة ($n<10$) يمكن الحصول على القيم الجدولية لـ r_s باستخدام جدول معامل ارتباط الرتب .

ـ إذا كان حجم العينة ($n\geq10$) ففي هذه الحالة يتم إختيار فرضية العدم (H_0) القائلة بأنه لاتوجد علاقة ارتباط بين المتغيرين مقابل الفرضية البديلة (H_1) توجد علاقة ارتباط بين المتغيرين وذلك باستخدام الاختبار :

$$t = \frac{r_s}{\sqrt{(1-r_s^2)/(n-2)}}$$

ولتوضيح كيفية إجراء اختبار معامل ارتباط سبيرمان نأخذ التمرين التالي .

تمرين (27) :

قام أحد الباحثين الاجتماعيين بدراسة حالة (10) أسر في أحد أحياء مدينة دمشق، وسجل لكل أسرة الحالة التعليمية لرب الأسرة والمستوى الاقتصادي للأسرة نفسها فكانت النتائج كالآتي :

المطلوب : حساب معامل الارتباط الرتبي لسبيرمان، ثم اختبر الفرض القائل بأنه لاتوجد علاقة بين المتغيرين المدروسين، عند مستوى معنوية $\alpha = 0.05$

رقم الأسرة	1	2	3	4	5	6	7	8	9	10
الحالة التعليمية لرب الأسرة	ثانوية	أمي	يقرأ ويكتب	إجازة جامعية	أمي	أمي	يقرأ ويكتب	ثانوية	أمي	ثانوية
المستوى الاقتصادي للأسرة	فقيرة	معدمة	فقيرة	غنية	معدمة	متوسط الحال	فقيرة	متوسط الحال	متوسط الحال	غنية

الحل : ننشئ الجدول المساعد :

رقم الأسرة	الحالة التعليمية (x)	المستوى الاقتصادي (y)	رتب (x)	رتب (y)	d	d^2
1	ثانوية	فقيرة	3	4	-1	1
2	أمي	معدمة	8.5	1.5	7	49
3	يقرأ ويكتب	فقيرة	5.5	4	1.5	2.25
4	إجازة	غنية	1	9.5	-8.5	72.25
5	أمي	معدمة	8.5	1.5	7	49
6	أمي	متوسط الحال	8.5	7	1.5	2.25
7	يقرأ ويكتب	فقيرة	5.5	4	1.5	2.25
8	ثانوية	متوسط الحال	3	7	-4	16
9	أمي	متوسط الحال	8.5	7	1.5	2.25
10	ثانوية	غنية	3	9.5	-6.5	42.25
Σ						238.5

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$= 1 - \frac{6(238.5)}{10(100-1)} = -0.44$$ العلاقة ضعيفة جداً

نختبر معنوية معامل الارتباط :

الاختبار الاحصائي :

$$t = \frac{0.44}{\sqrt{(1-0.2)/(8)}} = 0.2$$

إتخاذ القرار : من جداول (t) نجد أن : t (2.31) (اتجاهين)

V = 8

من خلال المقارنة نجد أن : t (0.2) < t (2.31) وبالتالي نقبل فرضية العدم القائلة بأنه لاتوجد علاقة بين المتغيرين .

تمرين (28) :

قام أحد الباحثين الاجتماعيين بدراسة حالة (10) أفراد في إحدى المدن وسجل لكل فرد المستوى الثقافي والمستوى الاقتصادي للفرد نفسه فكانت النتائج في الجدول التالي :

الأفراد	1	2	3	4	5	6	7
المستوى الثقافي	مرتفع	متوسط	فوق المتوسط	مرتفع جداً	تحت الوسط	منخفض	منخفض جداً
المستوى الاقتصادي	جيد	ممتاز	جيد جداً	جيد	مقبول	جيد	جيد جداً

والمطلوب :

اختبر الفرض القائل بأنه لاتوجد علاقة بين المستوى الاقتصادي والمستوى الثقافي عند مستوى معنوية ($\alpha = 0.05$) ، (يترك للطالب).

ـ معامل ارتباط كاندل ـ تاو للرتب Kendall's (Tau) Rank :

يُعد هذا المعامل من الطرائق اللامعلمية في تحديد العلاقة بين متغيرين رتبيين، ويستخدم كبديل لمعامل ارتباط سبيرمان، ويقوم على ترتيب أحد المتغيرين، أما رتبة المتغير الآخر، فتكون كما هي بالنسبة لكل مفردة من مفردات العينة فهو يأخذ عدد الرتب المتشابهة لقياس درجة التطابق بين مجموعتين من الرتب وعليه نميز عند حساب معامل كاندل ـ تاو التالي :

ـ إذا كان $n < 10$:

$$T_a = \frac{nc - np}{\left[\frac{n(n-1)}{2}\right]}$$

حيث أن : nc : عدد أزواج الرتب المتطابقة .

nP : عدد أزواج الرتب غير المتطابقة

ومن أجل اتخاذ القرار يتم المقارنة بين القيمة المحسوبة لمعامل كاندل تاو مع القيمة الجدولية المقابلة لمستوى معنوية معين باستخدام جدول معامل الرتب (سبيرمان) الذي يعطي القيم لـ (T_a) ، المقابلة لمستوى معنوية 5% و 1% .

ـ إذا كان $n > 10$:

يستخدم التوزيع الطبيعي كتقريب لتوزيع (T_a) من أجل اختبار معنوية معامل كاندل تاو وذلك من خلال العلاقة التالية :

$$Z = \frac{T_a}{\left[\frac{2(2n+5)}{9n(n-1)}\right]}$$

ومن أجل إتخاذ القرار يتم المقارنة بين القيمة المحسوبة لـــ (Z) مع القيمة الجدولية المقابلة لستوى معنوية معين، ولتوضيح ذلك نأخذ التمرين التالي:

تمرين (28):

بفرض لدينا معطيات الرتب المتعلقة بالعلاقة بدرجات الطلاب في الرياضيات والاحصاء لعينة عشوائية مكونة من (14) طالباً في كلية الاقتصاد بإحدى الجامعات:

رقم الطالب	1	2	3	4	5	6	7	8	9	10	11	12	13	14
رتب (x)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
رتب (y)	1	2	5	4	6	3	9	7	10	8	13	11	14	12

المطلوب :

اختبر الفرض القائل بأن معامل الارتباط في المجتمع يساوي الصفر مقابل الفرض البديل أنه لايساوي الصفر عند مستوى معنوية $\alpha = 0.05$ ؟

أو (اختبر الفرض القائل بأنه لاتوجد علاقة بين درجات الرياضيات ودرجات الإحصاء عند مستوى معنوية $\alpha=0.05$).

الجدول المساعد :

رقم الطالب	رتب (x)	رتب (y)	nc	NP
1	1	1	13	0
2	2	2	12	0
3	3	5	9	2
4	4	4	9	1
5	5	6	8	1
6	6	3	8	0
7	7	9	5	2
8	8	7	6	0
9	9	10	4	1
10	10	8	4	0
11	11	13	1	2
12	12	11	2	0
13	13	14	0	0
14	14	12	0	0
Σ			81	10

لتوضيح كيفية استخراج (nc)، (np) عملياً، نلاحظ أن ترتيب الطالب الأول على المتغير (y) حيث ترتيبه (1) هذا الترتيب يسبق جميع الرتب التي بعده وعددها (13)، كما أن هذه الرتبة لايوجد بعده ماهو أعلى منها أي (صفر)، لذلك نضع الرقم (13) في حقل الاتفاقات (nc) وصفر في حقل (np) أما الطالب الثاني فيحتل الرتبة (2) على المتغير (y).

وبملاحظة الرتب التي تأتي بعده نجد أن جميعها تأتي بعده في الترتيب وعددها (12) ولايوجد بينها أية رتبة تسبقها (صفر) وكذلك نلاحظ أن الطالب الثالث يحتل الرتبة (5) على المتغير (y) ولكن عند ملاحظة الرتب المكتوبة بعده في العمود نجد أنه يسبقها جميعاً عدا الطالبان الرابع والسادس حيث أن رتبة كل منهما (4) و(3) على التوالي وهاتان الرتبتان تسبقان الرتبة (5) ولذلك فإن (np) عدد الأزواج هو (2) في حين أن عدد الأزواج (nc) كان (9) وبالتالي نجد أن:

nc=81 ، np = 10 ولأجل صحة البيانات المستخرجة لكل من (nc) و(np) نستخرج قيمة المعادلة $\left(\frac{n^2-n}{2}\right)$،

فإذا ظهر أنها تساوي حاصل (nc+nP) هذا يعني سلامة الإجراءات المتبعة، أما إذا ظهر عكس ذلك فينبغي إعادة الحساب.

الاختبار الإحصائي :

$$Z = \frac{T_a}{\sqrt{\frac{2(2n+5)}{9n(n-1)}}}$$

$$T_a = \frac{nc - np}{\left[\frac{n(n-1)}{2}\right]} = \frac{81-10}{\frac{28(28-1)}{2}} = 0.19$$

$$Z = \frac{0.78}{\sqrt{\frac{2(28+5)}{9[28(28-1)]}}} = \frac{0.78}{\sqrt{\frac{66}{6804}}} = 1,93$$

إتخاذ القرار : من جداول (Z) نجد أن :

$Z_{(0.05)}=1,96$

من خلال المقارنة نجد أن : Z (19.5) > Z (1.96) فإننا نرفض فرضية العدم (H_0)، أي أن هناك علاقة بين درجات الإحصاء ودرجات الرياضيات .

ـ معامل الوفاق Coefficiet of concordance :

يكون لدينا أحياناً أكثر من مجموعتين من الرتب، كأن نطلب من مجموعة من مشرفين تصنيف عدد من الأسر حسب حق الأولوية في المساعدة مثلاً، ففي هذه الحالة والحالات المماثلة، يمكننا حساب معاملات الارتباط الرتبية بين المجموعات المختلفة لبيان إتفاق المشرفين أو المحكمين في ذلك ولدراسة مثل هذه الحالات يوجد معامل اتفاق كندال، ويعطى بالصيغة التالية :

$$W = \frac{2}{m^2} \times \frac{6\sum f^2}{n(n^2-1)}$$

حيث أن :

m : عدد المحكمين ؛

n : عدد الحالات ؛

f : مربعات فروق مجموع رتب كل مفردة عن المتوسط العام لمجموع الترتيبات.

ويتراوح قيمة معامل إتفاق كندال $0 \le W \le +1$ وعندما W=1 تدل على الإتفاق التام بين المحكمين والقيمة W=0 تدل على عدم الإتفاق ، وبصورة عامة كلما اقتربت قيمة (w) من الواحد الصحيح دل ذلك على قوة الإتفاق .

الاختبار الإحصائي لمعامل إتفاق كندال :

$$F = \frac{w(m-1)}{1-w}$$

ولتوضيح كيفية حساب معامل إتفاق كندال وكيفية إجراء الاختبار نأخذ التمرين التالي:

تمرين (29):

قام (5) خبراء بترتيب (4) منتجات من حيث الجودة في الجدول المرفق :

الخبراء / المنتجات	1	2	3	4	5
A	4	4	4	1	4
B	2	1	3	2	3
C	3	3	1	4	2
D	1	2	2	3	1

المطلوب :

حساب معامل إتفاق كندال، ثم إختبر الفرض القائل بأنه هناك ارتباط بين ترتيبات الخبراء لترتيب المنتجات الأربع من حيث الجودة عند مستوى معنوية $\alpha=0.05$ ؟.

الحل :

الجدول المساعد :

الخبراء / المنتجات	1	2	3	4	5	مجموع رتب كل منتج	f	f^2
A	4	4	4	1	4	17	4.5	20.25
B	2	1	3	2	3	11	-1.5	2.25
C	3	3	1	4	2	13	0.5	0.25
D	1	2	2	3	1	9	-3.5	12.25
Σ						50		35

ـ متوسط مجموع رتب كل مفردة $\frac{50}{4} = 12.5$ ،

ـ للحصول على (f) نطرح مجموع رتب كل منتج من (12.5) كما هو موضح بالجدول أعلاه، وعليه نجد أن :

$$W = \frac{2}{m^2} \times \frac{6\sum f^2}{n(n^2-1)}$$

$$= \frac{2}{5^2} \times \frac{6(35)}{4(16-1)} = 0.28$$

الاختبار الاحصائي :

$$F = \frac{w(m-1)}{1-w}$$

$$= \frac{0.28(5-1)}{1-0.28} = 1.56$$

إتخاذ القرار : من خلال جداول (F) نجد أن :

$$F_{(0.05\ ,\ 4.3)} = 9,12$$

من خلال المقارنة نجد أن : F (1.56) < F (9.12) وبالتالي لاتوجد علاقة بين ترتيبات الخبراء، ولايمكن القول بأن ترتيباتهم للمنتجات وثيقة الصلة، بمعنى هناك اختلاف بين الخبراء من حيث ترتيبهم للمنتجات من حيث الجودة .

ـ معامل الاقتران الرباعي Association coefficient :

يستخدم هذا المعامل في حالة المتغيرات النوعية غير القابلة للقياس العددي التي يقسم كل منها إلى صفتين فقط مثل (شفي ، لم يشفى) ، (شارك، لم يشارك). فإذا كان لدينا المتغيرين (x)

و(y)، وكل منهما ينقسم إلى صفتين فترتب النتائج في جداول تسمى بجداول التوافق وفق التالي :

X \ Y	I	II
I	a	B
II	c	D

حيث أن (a,b,c,d) تمثل المشاهدات في صورة تكرارات ويحسب المعامل بالصيغة التالية :

$$r_c = \frac{a.d - c.b}{a.d + c.b}$$

الاختبار الإحصائي : يتم تحويل قيمة معامل الإقتران الحاصلة إلى قيمة (Z) وفق الصيغة التالية:

$$Z = \frac{2r_c}{1 - r_c^2} \times \frac{1}{\sqrt{1/a + 1/b + 1/c + 1/d}}$$

ويمكن توضيح كيفية إجراء الاختبار لمعامل الاقتران من خلال التمرين التالي:

تمرين (30) :

قام أحد الباحثين باختيار عينة عشوائية من عمال إحدى الشركات للتعرف على العلاقة بين زيادة الأجور والإنضباط في العمل فحصل على النتائج التالية :

الأجور / الإنضباط	زيادة	عدم زيادة
أفضل	16	9
كما هو	2	10

المطلوب :

حساب معامل الاقتران الرباعي، ثم اختبر الفرض القائل بأنـــه لـــيس هناك علاقة بين زيادة الأجور والانضباط في العمل عند مستوى معنوية (α = 0.05) .

الحل :

$$r_c = \frac{a.d - c.b}{a.d + c.b}$$

$$= \frac{(16)(10) - (2)(9)}{(16)(10) + (2)(9)} = 0.80$$

الاختبار الإحصائي :

$$Z = \frac{2r_c}{1 - r_c^2} \times \frac{1}{\sqrt{1/a + 1/b + 1/c + 1/d}}$$

$$= \frac{2(0.80)}{1 - (0.80)^2} \times \frac{1}{\sqrt{1/16 + 1/9 + 1/2 + 1/10}} = 5.38$$

إتخاذ القرار : من جداول (Z) نجد أن : Z (1.96)

α=0.05

من خلال المقارنـــة نجـــد أن Z (5.38) > Z (1.96) وعليـــه نـــرفض الفرضية ونقول بأن هناك علاقة بين زيادة الأجور والإنضباط في العمل .

تمرين (31) :

قام أحد الباحثين باختيار عينة عشوائية من عمال إحدى الشركات للتعرف على أهمية المشاركة في الدورات التدريبية على تحسين الأداء فكانت النتائج التالية :

المشاركة / الأداء	شارك	لم يشارك
تحسن	24	8
كما هو	4	12

المطلوب :

حساب معامل الاقتران الرباعي، ثم اختبر الفرض القائل بأنه ليس هناك علاقة بين المشاركة في الدورات التدريبية وتحسن مستوى الأداء في العمل عند مستوى معنوية ($\alpha = 0.05$) ، (يترك الحل للطالب).

معامل ارتباط (فاي) Phi coefficient :

يستخدم هذا المعامل عندما تكون البيانات في صورة متغيرين ينقسم كل منهما إلى صفتين مثل (نعم، لا) أو (ذكر ، أنثى) أو (صح ، خطأ). فإذا كانت لدينا الإجابة ثنائية (نعم ، لا) على سؤالين مختلفين في اختبار ما وكان المطلوب معرفة الارتباط بين الإجابات على هذين السؤالين فسيكون لدينا الجدول التالي

X / Y	نعم	لا	Σ
نعم	a	b	a+b

لا	c	d	c+d
∑	a+c	b+d	n

حيث أن (a,b,c,d) تمثل المشاهدات في صورة تكرارات ويتم حساب المعامل بالصيغة التالية:

$$\phi = \frac{a.d - b.c}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$$

الاختبار الإحصائي : يتم من خلال العلاقة التالية :

$$Z = \phi\sqrt{n}$$

ويمكن توضيح حساب معامل ارتباط ϕ واختباره الإحصائي من خلال التمرين التالي:

تمرين (32) :

قام أحد الباحثين باختيار عينة عشوائية من طلاب إحدى الكليات بالجامعة للتعرف على العلاقة بين اتجاه الطلبة الجامعيين نحو عمل المرأة واتجاههم نحو الاختلاط في الدراسة الجامعية، وجاءت بيانات الإجابة على سؤالين هما :

1 ــ هل تؤيد عمل المرأة في المهن المتوفرة في المجتمع ؟ (نعم ، لا)

2 ــ هل تؤيد اختلاط الجنسين في الدراسة الجامعية ؟ (نعم ، لا)

وحصل الباحث على البيانات التالية :

الإجابة	نعم	لا
نعم	45	10
لا	15	30

والمطلوب :

حساب الارتباط بين الاجابات على هذين السؤالين، ثم اختبر الفرض القائل بأنه يوجد علاقة بين الإجابات على السؤال الأول وإجاباتهم على السؤال الثاني.

أو (هل توجد علاقة بين اتجاه الطلبة نحو عمل المرأة وبين اتجاههم نحو الاختلاط في الدراسة الجامعية) عند مستوى معنوية (α=0.05)

الحل :

الجدول المساعد:

الاستجابات	نعم	لا	Σ
نعم	45	10	55
لا	15	30	45
Σ	60	40	100

$$\phi = \frac{(45)(30)-(10)(15)}{\sqrt{(55)(45)(40)(60)}} = \frac{1350-150}{\sqrt{5940000}}$$

الاختبار الإحصائي :

$$Z = \phi\sqrt{n}$$
$$= 0.49\sqrt{100} = 4,9$$

اتخاذ القرار : من جداول (Z) نجد أن : Z(1.96)

α=0.05

من خلال المقارنة نجد أن : Z (1.96)<Z (4.92) وبالتالي نرفض الفرضية ونقول بأن هناك علاقة بين المتغيرين (اتجاه الطلبة نحو عمل المرأة وبين اتجاههم نحو الاختلاط في الدراسة الجامعية) .

تمرين (33):

قام أحد الباحثين باختيار عينة عشوائية من طلاب كلية العلوم السياسية بإحدى الجامعات للتعرف على العلاقة بين اتجاه الطلبة الجامعيين نحو عمل المرأة في السياسة واتجاههم نحو تسلم المرأة لمناصب سياسية فجاءت بيانات الإجابة على سؤالين هما :

1 ــ هل تؤيد عمل المرأة في السياسة ؟ (نعم ، لا)

2 ــ هل تؤيد تسلم المرأة لمناصب سياسية ؟ (نعم ، لا) .

وحصل الباحث على البيانات التالية :

الإجابة	نعم	لا
نعم	15	19
لا	23	14

المطلوب :

حساب الارتباط بين الإجابات على هذين السؤالين، ثم اختبر الفرض القائل بأنه يوجد علاقة بين اتجاه الطلبة نحو عمل المرأة في السياسة وبين اتجاههم نحو تسلم المرأة لمناصب سياسية، عند مستوى معنوية ($\alpha = 0.05$) (يترك الحل للطالب).

ــ معامل التوافق Contingency coefficent

يستخدم هذا المعامل في حالة المتغيرات النوعية غير القابلة للقياس العددي التي يقسم كل منها إلى أكثر من صفتين مثل (متزوج، أعزب،

مطلق)، (ممتاز، جيد جداً، جيد)، لذا يعد هذا المعامل أعم من معامل الاقتران الرباعي حيث يمكن تطبيقه في الجدول (3 × 3) أو أكثر وسواء تساوي عدد خلايا الصفوف وعدد خلايا الأعمدة أم لم يتساو. ويتم حساب هذا المعامل بالصيغة التالية :

$$K = \sqrt{1 - 1/G}$$

حيث أن :

$$G = \frac{(fij)^2}{(\sum f_i)(\sum f_j)}$$

الاختبار الإحصائي : يتم من خلال استخدام توزيع (χ^2) ووفق الصيغة التالية:

$$\chi^2 = \frac{(n)(k)^2}{1 - k^2}$$

حيث أن :

V = (m-1) (n-1) درجة حرية

n : عدد أفراد العينة

k^2 : مربع معامل التوافق .

ولتوضيح ذلك لدينا التمرين التالي :

تمرين (34) :

قام أحد الباحثين باختيار عينة عشوائية من طلاب إحدى الكليات بهدف معرفة العلاقة بين التقديرات للمقرر الرياضيات ومقرر الإحصاء فكانت النتائج التالية:

الرياضيات / الإحصاء	ممتاز	جيد	مقبول
ممتاز	2	4	4
جيد	3	1	6
مقبول	5	2	3

والمطلوب :

حساب معامل التوافق، ثم اختبر الفرض القائل بأنه لايوجد هناك علاقة بين تقديرات الرياضيات وتقديرات الإحصاء للطلاب عند مستوى معنوية (α = 0.05)

الحل :

الجدول المساعد :

الرياضيات / الإحصاء	ممتاز	جيد	مقبول	$\sum$
ممتاز	2	4	4	10
جيد	3	1	6	10
مقبول	5	2	3	10
$\sum$	10	7	13	30

حساب قيمة (G) وفق التالي :

$$G = \frac{(2)^2}{(10)(10)} + \frac{(4)^2}{(7)(10)} + \frac{(4)^2}{(13)(10)} + \ldots\ldots \frac{(3)^2}{(13)(10)} = 1.15$$

$$K = \sqrt{1 - 1/G} = \sqrt{1 - 1/1.15} = 0.36$$

الاختبار الإحصائي :

$$\chi^2 = \frac{(n)(k)^2}{1-k^2}$$

$$= \frac{(30)(0.36)^2}{1-(0.36)^2} = 4.47$$

ـ **إتخاذ القرار** : من جداول χ^2 نجد أن :

χ^2 (9.49)
V=(3-1)(3-1)=4
$\alpha = 0.05$

من خلال المقارنة نجد أن $\chi^2(9.49) > \chi^2$ (4.49) وبالتالي نقبل الفرض القائل بعدم وجود علاقة بين التقديرات للرياضيات والإحصاء لهؤلاء الطلبة.

ملاحظات :

1 ـ عندما تكون (G = 1) فهذا يعني أن الظاهرتين مستقلتين تماماً وبالتالي (K=0) والعلاقة معدومة أما إذا كانت (G > 1) فإن معامل التوافق (K≠0) وهذا يعني وجود علاقة .

2 ـ تتراوح قيمة $0 \le k \le +1$. ويمكن الاستفادة منها في إيجاد العلاقة بين متغيرين أحدهما كمي والآخر نوعي مثال : نوع السلعة وسعرها .

تمرين (35):

أوجد العلاقة بين الصفات الوراثية لطول القامة لدى الآباء وطول القامة لدى الأبناء ثم اختبر الفرض القائل بعدم وجود علاقة بين الصفات الوراثية بين الآباء والأبناء للبيانات في الجدول التالي :

طول القامة (الأبناء) / طول القامة (الآباء)	طويل	متوسط	قصير
طويل	12	14	14
متوسط	13	11	16
قصير	10	12	13

(يترك الحل للطالب).

تمرين (36) :

أوجد العلاقة بين سعر السلعة ولونها، ثم اختبر الفرض القائل بعدم وجود علاقة بين سعر السلعة ولونها للبيانات في الجدول التالي :

اللون / السعر	أبيض	أسود	أزرق
متوسط	12	10	15
كبير	11	10	11
كبير جداً	13	10	11

(يترك الحل للطالب).

ـ معامل كرامير Cramer Coefficent :

عبارة عن الشكل الآخر لمعامل التوافق، ويستخدم أيضاً في حالة المتغيرات النوعية غير القابلة للقياس العددي ويعطى بالعلاقة التالية :

$$C = \sqrt{\frac{G-1}{A-1}}$$

حيث أن :

G : تحسب بنفس الأسلوب السابق لمعامل التوافق

A : العدد الأقل للصفوف أو الأعمدة

وفي حالة اشتمل الجدول المزدوج للبيانات على صفين فقط وعمود فقط فإن :

$$C = \sqrt{G - 1}$$

وهي مماثلة تماماً لمعامل فاي ϕ .

الاختبار الإحصائي : يتم من خلال استخدام توزيع (χ^2) بنفس أسلوب معامل التوافق

$$\chi^2 = \frac{(n)(k)^2}{1 - k^2}$$

ولتوضيح كيفية إجراء هذا الاختبار نأخذ التمرين التالي :

تمرين (37):

قام أحد الباحثين باختيار عينة عشوائية حجمها (229) من العاملين بإحدى الشركات بهدف دراسة دور برامج التدريب على مستوى الأداء لدى العاملين في هذه الشركة، وقد جمعت البيانات التالية :

البرامج / الأداء	A	B	C
تحسن	45	50	30
كما هو	29	20	31
انخفض	6	3	15

الحل : حساب (G) على النحو التالي :

$$G = \frac{(45)^2}{(125)(80)} + \frac{(50)^2}{(125)(73)} + \ldots\ldots \frac{(15)^2}{(76)(24)} = 1.07$$

ونلاحظ أن عدد الأعمدة يساوي عدد الصفوف أي أن (A = 3)

$$K = \sqrt{\frac{G-1}{A-1}} = \sqrt{\frac{1.07-1}{3-1}} = 0.19$$

الاختبار الإحصائي : (أنظر الجدول المساعد المرفق)

$$\chi^2 = \frac{(n)(k)^2}{1-k^2}$$

$$= \frac{(229)(0.19)^2}{1-(0.19)^2} = 8.58$$

الجدول المساعد:

البرامج / الأداء	A	B	C	Σ
تحسن	45	50	30	125
كما هو	29	20	31	80
انخفض	6	3	15	24
Σ	80	73	76	229

اتخاذ القرار : من جداول (χ^2) نجد أن :

$\chi^2(9.49)$

$V = (3-1)(3-1)=4$

$\alpha = 0.05$

ومن خلال المقارنة نجد أن : $\chi^2 (8.58) < \chi^2 (9.49)$ وبالتالي نقبل الفرضية ونقول ليس هناك علاقة بين برامج التدريب ومستوى الأداء للعاملين.

تمرين (38) :

قام أحد الباحثين باختيار عينة عشوائية في دكاترة إحدى الجامعات بهدف دراسة العلاقة بين الحالة الزواجية ومستوى الإنجاز العلمي لدى الدكاترة في هذه الجامعة، فكانت النتائج التالية:

الحالة الزواجية / الإنجاز العلمي	متزوج	أعزب	مطلق
جيد	28	11	5
سيء	7	14	20

والمطلوب :

قياس الارتباط، ثم اختبر الفرض القائل بأنه ليس هناك علاقة بين الحالة الزواجية ومستوى الإنجاز العلمي عند مستوى معنوية ($\alpha = 0.01$) .

(يترك الحل للطالب).

ـ معامل تشيبرو Tachuprou Coefficent

يستخدم هذا المعامل في الحالات التي يكون فيها أحد المتغيرين منقسماً إنقساماً مثل (نعم، لا، محايد) أو للبند الواحد مثلاً أربعة بدائل (موافق بشدة، موافق، أرفض بشدة، أرفض) ويعطى بالعلاقة التالية :

$$CT = \sqrt{\frac{K^2}{(1-K^2)\sqrt{(L-1)(m-1)}}}$$

حيث أن :

K : معامل التوافق

L : عدد البدائل أو الانقسامات للمتغير الأول

m: عدد البدائل أو الانقسامات للمتغير الثاني

نأخذ التمرين (38) لتوضيح كيفية استخدام هذا المعامل:

تمرين (38) : توصل أحد الباحثين من خلال دراسة حول المستوى الحضاري (ريف، حضر) والمستوى الاقتصادي (مرتفع، متوسط، منخفض) إلى النتائج الآتية :

$$m = 2 \quad , \quad L = 2 \quad , \quad K = 0.34$$

والمطلوب : قياس الارتباط بين المستوى الاقتصادي والمستوى الحضري

الحل :

$$C.T = \sqrt{\frac{K^2}{(1-K^2)\sqrt{(L-1)(m-1)}}}$$

$$= \sqrt{\frac{(0.34)^2}{(1-(0.34)^2\sqrt{(2-1)(2-1)}}} = 0.35$$

وبالتالي العلاقة ضعيفة بين المتغيرين .

تمرين : توصل أحد الباحثين من خلال دراسة حول مستوى الأداء والنوع (ذكر، أنثى) إلى البيانات التالية :

الأداء / النوع	مرتفع	متوسط	منخفض
ذكور	20	12	5
إناث	5	10	8

والمطلوب : أوجد معامل الارتباط، ثم اختبر الفرض القائل بأنه ليس هناك علاقة بين مستوى الأداء والنوع عند مستوى معنوية ($\alpha = 0.01$). (الحل يترك للطالب).

تمارين غير محلولة

تمرين (1) :

تهتم إدارة إحدى الكليات بمعرفة تقديرات النجاح لجميع المقررات، وقد تم من خلال خبرة سابقة تصنيف هذه التقديرات لأحد المقررات على النحو التالي :

التقديرات	ضعيف	مقبول	جيد	جيد جداً	ممتاز
النسبة	30%	30%	25%	10%	5%

أخذت عينة عشوائية من (1000) طالب، فوجد أن : (250) طالب تقدير ضعيف، (300) طالب تقدير مقبول، (200) طالب تقدير جيد، (150) طالب تقدير جيد جداً، (100) طالب تقدير ممتاز.

والمطلوب : اختبار فيما إذا كان التوزيع الحالي للتقديرات مازال نفس التوزيع السابق الذي تم الحصول عليه من خبرة سابقة .

تمرين (2) :

رغب أحد الباحثين معرفة رغبات طلبة الجامعات في نوع الكتب المفضلة لديهم، فاختار عينة عشوائية من (56) طالباً وطلب من كل واحد منهم أن يختار الكتب التي يفضلها أكثر من غيرها من بين أربعة أنواع منها هي : كتب تربوية، اجتماعية، اقتصادية وسياسية، وطلب منهم عدم اختيار أكثر من نوع واحد فوجد أن : (18) منهم يفضلون الكتب التربوية، (8) منهم يفضلون الكتب الاجتماعية، (19) منهم يفضلون الكتب الاقتصادية، و (11) منهم يفضلون الكتب السياسية **والمطلوب :** هل رغبات هؤلاء الطلاب تمثل رغبات جميع الطلاب المتساوية في الرغبات لكل نوع من أنواع الكتب.

تمرين (3) :

يعتقد أن مجتمع الموظفين في القطاع العام في إحدى الدول العربيـــة بحســـب خبره سابقة يتوزع حسب الفئات التالية :

الفئات	1	2	3	4	5
النسب المئوية	10%	15%	25%	30%	20%

أخذت عينة عشوائية تمثل هذا المجتمع حجمها (210) موظفـاً فوجـد أنهـا تتوزع على هذه الفئات على النحو التالي :

الفئات	1	2	3	4	5
عدد الموظفين	15	35	60	50	50

والمطلوب :

اختبر صحة هذا الاعتقاد عند مستوى معنوية $\alpha = 0.05$

تمرين (4) :

لدينا التوزيع الاحتمالي لمتغير معين X كما يلي :

Pi : 0.1777 0.2592 0.1604 0.0568
0.1202 0.2260

أخذت عينة عشوائية حجمها (200) فوجد أنها تتوزع كما يلي :

ni : 12 18 20 100 25
25

والمطلوب :

اختبر فيما إذا كان التوزيع الاحتمالي الحالي للمتغير X هو نفس التوزيع السابق عند $\alpha = 0.05$

تمرين (5) :

وجه أحد الباحثين سؤال لعينة عشوائية من الطلاب حجمها (500) طالب حول إجراء الامتحانات بطريقة الأتمتة (الخيارات المتعددة)، وجاءت التكرارات لبدائل خمسة على النحو:

الآراء	مؤيد بشدة	مؤيد	محايد	معارض	معارض بشدة	Σ
التكرارات	200	60	100	120	20	500

فإذا كانت النسب المتوقعة طبقاً للدراسة هي :

الآراء	مؤيد بشدة	مؤيد	محايد	معارض	معارض بشدة	Σ
النسب المئوية	50%	12%	18%	17%	3%	100%

والمطلوب :

اختبر الفرض القائل بأن هناك تشابه بين نسب إجابات طلاب العينة والنسب المتوقعة طبقاً للدراسات السابقة عند مستوى معنوية $\alpha = 0.05$.

تمرين (6) :

تتوزع فئات أحد المجتمعات على النحو التالي :

الفئات	A	B	C	D	E	Σ
النسبة%	15	30	12	33	10	100

أخذت عينة عشوائية من هذا المجتمع فجاء توزيع الفئات على النحو التالي :

الفئات	A	B	C	D	E	Σ
التكرارات	24	40	20	44	12	140

والمطلوب :

اختبر الفرض القائل بأن هناك تشابه بين نسب إجابات أفراد العينة والنسب المتوقعة عند مستوى معنوية $\alpha=0.05$.

تمرين (7) : البيانات التالية تمثل عدد الطلبة المسجلين في عشرة مجموعات تقوم بتدريس مقرر الإحصاء

رقم المجموعة	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Σ**
المسجلين	**29**	**31**	**20**	**15**	**30**	**20**	**23**	**28**	**25**	**29**	**250**

والمطلوب :

هل تدل هذه البيانات على عدم تفضيل الطلبة لأي مجموعة على المجموعات الأخرى، وذلك عند مستوى معنوية $\alpha = 0.01$.

تمرين (8) :

في عينة عشوائية من (100) عامل من عمال إحدى الشركات، وجد أن من بين (70) عامل متعلم (15) عامل مدخن ومن بين غير المتعلمين وجد (14) عامل لايدخنون **والمطلوب:** هل يمكن إرجاء صفة التدخين أصلاً إلى صفة التعلم عند مستوى معنوية $\alpha=0.05$.

تمرين (9) :

البيانات التالية عبارة عن نتائج دراسة تتعلق بالتعليم المفتوح في إحدى الجامعات :

الرأي / النوع	مؤيد	معارض	محايد	Σ
ذكور	70	15	35	120
إناث	55	10	15	80
Σ	125	25	50	200

والمطلوب:

اختبر فيما إذا كان هناك علاقة بين الرأي والنوع عند مستوى معنوية $\alpha=0.05$.

تمرين (10) :

في عينة عشوائية من (100) طالب وطالبة كانت أعداد الناجحين والراسبين في امتحان مادة الإحصاء كمايلي :

النتيجة / النوع	ناجح	راسب	Σ
طالب	40	15	55
طالبة	35	10	45
Σ	75	25	100

المطلوب :

اختبر فيما إذا كان هناك علاقة بين النوع الطلبة ونتائجهم في الامتحان عند مستوى معنوية $\alpha = 0.05$

تمرين (11) :

سحبت عينة عشوائية مكونة من (330) من الناخبين في إحدى الدول لمعرفة رأيهم في قانون الأحزاب الجديد حيث تم طبقاً لإنتماءاتهم الحزبية، والجدول التالي يمثل النتائج التي حصلنا عليها :

الرأي / الإنتماء الحزبي	مؤيد	معارض	محايد	Σ
الحزب (A)	62	52	24	138
الحزب (B)	98	68	26	192
Σ	160	120	50	330

والمطلوب :

اختبر الفرض القائل باستقلال كلاً من الانتماء الحزبي والرأي في قانون الأحزاب الجديد عند مستوى معنوية $\alpha=0.05$.

تمرين (12) :

في دراسة حول الزواج المدني في إحدى الدول، أخذت عينة عشوائية من طلبة الجامعات فكانت النتائج كالآتي :

الرأي / النوع	مؤيد	معارض	محايد	Σ
ذكور	40	32	18	90
إناث	60	40	20	120
Σ	100	72	38	210

المطلوب :

إختبار فيما إذا كان هناك علاقة بين النوع والرأي عند مستوى معنوية $\alpha=0.05$.

تمرين (13) :

بفرض أن البيانات التالية لعينة عشوائية من (1000) ناخب في إحدى الدول التي تم تصنيفهم حسب إنتمائهم الحزبي وحسب تأييدهم أو معارضتهم لمبدأ معين.

الانتماء الحزبي / الرأي	الحزب A	الحزب B	Σ
معارض للمبدأ	250	200	450
مؤيد للمبدأ	400	150	550
Σ	650	350	1000

والمطلوب :

اختبر الفرض القائل بعدم وجود علاقة بين الانتماء الحزبي وإبداءه للرأي (معارض أو مؤيد) عند مستوى معنوية $\alpha = 0.05$.

تمرين (14) :

يرغب البعض في معرفة فيما إذا لمرتبة عضو هيئة التدريس لها علاقة باستخدام المكتبة الجامعية، سحبت عينة عشوائية مكونة من (1000) عضو هيئة تدريس من إحدى الجامعات، حيث تم تصنيفهم حسب المرتبة العلمية إلى: أستاذ، أستاذ مساعد، مدرس وحسب استخدامهم للمكتبة الجامعية فكانت النتائج في الجدول التالي :

المرتبة / الاستخدام	أستاذ	أستاذ مساعد	مدرس	Σ
مستفيدون	90	150	160	400
غير مستفيدين	110	250	240	600
Σ	200	400	400	1000

والمطلوب :

اختبر الفرض القائل بعدم وجود علاقة بين مرتبة عضو هيئة التدريس وبين استخدام المكتبة الجامعية عند مستوى معنوية $\alpha = 0.05$.

تمرين (15) :

بفرض أننا نرغب في معرفة ما إذا كان ذكاء الأشخاص ومعدل أوزانهم مستقلان عن بعضهما سحبت عينة عشوائية مكونة من (90) طالب من طلاب إحدى الكليات، فحصلنا على التكرارات التالية :

معدل الوزن / الذكاء	ضعيف	متوسط	ثقيل	Σ
ذكاء طبيعي	10	20	10	40
ذكاء ضعيف	25	10	15	50
Σ	35	30	25	90

المطلوب :

اختبر الفرض القائل باستقلال المتغيرين، عند مستوى معنية $\alpha = 0.01$

تمرين (16) :

عين محاضرين (x ، y) لتدريس مقرر الإحصاء في إحدى الكليات بالجامعة، وتم توزيع التقديرات لكل من هؤلاء المحاضرين، فحصلنا على البيانات التالية:

التقديرات / المحاضر	ممتاز	جيد	مقبول	Σ
X	62	52	24	138
Y	98	68	26	192
Σ	160	120	50	330

تمرين (22) :

لتحديد فعالية برنامج أمن صناعي معين، جمعت بيانات عن الحوادث الصناعية ببعض المصانع فكانت البيانات المرفقة تمثل متوسط عدد ساعات العمل المفقودة شهرياً قبل وبعد تنفيذ برنامج الأمن الصناعي:

رقم المصنع	1	2	3	4	5	6	7	8	9	10
قبل البرنامج	28	37	8	65	43	14	15	6	28	115
بعد البرنامج	19	38	7	53	31	19	13	4	26	100

والمطلوب: استخدم التوزيع الطبيعي كتوزيع تقريبي لاختبار الإشارة حول فعالية نظام الأمن الصناعي.

تمرين (23) :

رغب أحد الباحثين دراسة مدى تأثير استخدام إحدى الوسائل التعليمية الحديثة في قدرة الأطفال على التمييز بين الأشكال (شكل المكعب، شكل الكرة)، فقام الباحث بأخذ عينة عشوائية تتألف من (12) طفلاً في إحدى رياض الأطفال وطبق عليهم أحد الاختبارات الخاصة بقياس قدرة الأطفال على التمييز بين الأشكال قبل استخدام الوسيلة التعليمية وبعد استخدامها فكانت النتائج كمايلي :

رقم الطفل	1	2	3	4	5	6	7	8	9	10	11	12

درجة الاختبار بالطريقة القديمة (y)	25	19	24	24	21	29	21	18	23	28	25	28
درجة الاختبار بالطريقة الجديدة (x)	20	18	22	28	17	26	25	24	27	21	19	22

المطلوب : اختبر الفرض القائل بعدم تأثير استخدام الوسيلة التعليمية الجديدة في القدرة على التمييز بين الأشكال باستخدام اختبار ولكوكسون، عند مستوى معنوية $(\alpha = 0.05)$.

تمرين (24) :

أخذت عينة عشوائية من الطلبة (16) طالباً أخضعت لتجربة معينة للتعرف على أثر منح الجوائز التقديرية على الزمن الذي يقطعه الطالب في سباق الجري لمسافة (500)م. وقد سجل أحد الباحثين عدد الثواني التي استغرقها كل طالب في قطع هذه المسافة. وبعد ذلك أخبر الباحث هؤلاء الطلبة بأن هناك جوائز تقديرية ستقدم للفائزين في هذا السابق، ثم كرر التجربة وسجل الوقت الذي استغرقه كل طالب في المرة الثانية. كما هو موضح بالجدول :

رقم الزوج	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
نتيجة السباق الأول	120	118	110	116	120	114	108	106	112	106	112	120	114	115	115
نتيجة السباق الثاني	110	115	109	117	114	114	106	112	100	113	105	110	119	113	110

المطلوب : اختبر الفرض القائل بعدم تأثير منح الجوائز التقديرية على الزمن الذي يقطعه الطالب في السباق باستخدام اختبار ولكوكسون عند مستوى معنوية ($\alpha=0.05$).

تمرين (25) :

اقترح أحد التربويين طريقة جديدة لتدريس مقرر اللغة الإنكليزية لطلاب إحدى الكليات ولفحص جدوى هذه الطريقة، أخذ عينة من (10) طلاب بصورة أزواج، بحيث تم القاء المحاضرة على الطلاب بالطريقة الجديدة ومن ثم بالطريقة المعتادة بعد ذلك. أجري امتحان لهؤلاء الطلاب فكانت النتائج في الجدول التالي :

رقم الزوج	1	2	3	4	5	6	7	8	9	10
درجات الطلاب بالطريقة المعتادة (y)	71	95	85	93	79	90	95	95	87	86
درجات الطلاب بالطريقة الجديدة (x)	70	90	81	91	79	88	92	89	94	96

المطلوب : اختبر الفرض القائل بأنه ليس هناك فرقاً معنوياً بين الطريقتين باستخدام اختبار ولكوكسون عند مستوى معنوية ($\alpha=0.05$) .

تمرين (26) :

تم استخدام ثلاثة طرائق مختلفة لتعليم اللغة الإنكليزية لثلاث مجموعات (A,B,C) من الطلاب قد اختيروا بطريقة عشوائية، وبعد فترة أعطي لهم اختباراً موّحداً، فحصلنا على النتائج التالية :

الرقم	1	2	3	4	5	6	7	8	9	10
المجموعة (A)	93	77	93	79	92	99	98	71	87	-
المجموعة (B)	89	90	85	76	84	95	82	72	73	68
المجموعة (c)	78	80	75	81	91	88	86	94	69	100

المطلوب : اختبر الفرض القائل بتساوي مجتمعات درجات اختبار طرائق التعليم الثلاثة في مقابل الفرض البديل القائل باختلافها عند مستوى معنوية $(\alpha=0.05)$.

تمرين (27) :

الجدول التالي يمثل الدرجات التي حصل عليها خمسة طلاب في الاختبار النهائي لكل من الحاسوب، الإحصاء، الرياضيات :

الحاسوب	98	86	78	84	87
الإحصاء	90	91	95	97	93
الرياضيات	99	89	80	88	96

المطلوب : اختبر الفرض القائل بتساوي مستوى الطلبة في الموضوعات الثلاث عند مستوى معنوية $(\alpha = 0.05)$.

تمرين (28) :

لدراسة أثر التحصيل العلمي في الاتجاه نحو عمل المرأة قام أحد الباحثين باختيار ثلاث عينات، العينة (A) من جملة الشهادة الإعدادية وحجمها (5) والعينة الثانية (B) من جملة الشهادة الثانوية وحجمها (4)، والعينة الثالثة (C) من جملة الشهادة الجامعية وحجمها (4).

طبق مقياس الاتجاه نحو عمل المرأة وهو عبارة عن قائمة تحتوي (50) سؤالاً يتطلب الإجابة عليها بـــ نعم أو لا ،وبعد تطبيق هذا المقياس دلت النتائج كما في الجدول التالي :

الرقم	1	2	3	4	5
العينة (A)	40	35	32	23	20
العينة (B)	42	25	24	18	-
العينة (C)	48	46	41	-	

المطلوب : اختبر الفرض القائل بتساوي مستوى الإجابات في مقياس الاتجاه نحو عمل المرأة. عند مستوى معنوية ($\alpha = 0.05$) .

تمرين (29) :

رغب أحد معلمي اللغة الفرنسية تجريب أربع طرائق مختلفة لتعليم مفردات هذه اللغة من أجل اختبار الطريقة التي يتعلم بها الطلبة المفردات أسرع من غيرها، فاختبر أربع عينات بصورة عشوائية تتكون كل منها من (5) طلاب، وطبق الطرائق الأربع على كل من هذه العينات وبعد إجراء الامتحان لقياس تعلم المفردات اللغوية كانت النتائج كما يلي :

الرقم	1	2	3	4	5
العينة (A)	10	30	30	15	35
العينة (c)	25	30	40	45	40
العينة (D)	30	45	45	50	40

المطلوب : اختبر الفرض القائل لتساوي الطرائق الأربع، أي ليس هناك فرق معنوي بين العينات الأربع عند مستوى معنوية ($\alpha = 0.01$) .

تمرين (30) ــ بفرض أن أحد الباحثين المهتمين بدراسات الأطفال أراد التعرف على رغبات أطفال الروضة في اختيارهم للعبة معينة ذات أربعة أحجام مختلفة وجاءت النتائج بالجدول التالي:

الحجم	A	B	C	D	Σ
العدد	2	6	15	5	28

المطلوب :

اختبر الفرض القائل بأن اختيار اللعبة ليس له علاقة بحجمها عند مستوى معنوية (α=0.05)

تمرين (31) ــ بفرض أن كلية الاقتصاد باحدى الجامعات تقبل الحاصلين على الثانوية العامة (علمي، أدبي) ولمعرفة معنوية الفروق بين نتائج الطلبة ذوي الخلفية العلمية في الدراسة الثانوية وبين نتائج الطلبة ذوي الخلفية الأدبية في مقرر الرياضيات ، فكانت النتائج على النحو التالي :

التقدير / التخصص	ممتاز	جيد جداً	جيد	مقبول	ضعيف
العلمي	14	44	6	3	3
الأدبي	10	65	4	1	0

المطلوب :

اختبر الفرض القائل بأنه لاتوجد فروق معنوية في تقديرات مقرر الرياضيات بين الطلاب العلمي والأدبي عند مستوى معنوية (α=0.05).

الفصل الثامن

تحليل التباين

Analysis of Variance

- مقدمة
- تحليل التباين الأحادي التصنيف
- تمارين غير محلولة

8-1- مقدمة :

إن تحليل التباين، عبارة عن أسلوب إحصائي يأخذ بعين الاعتبار جميع العينات لإجراء اختبارات الفروض المتعلقة بالفروق بين متوسطات ثلاثة مجتمعات أو أكثر، ويفترض أن تكون المجتمعات الأصلية تتبع التوزيع الطبيعي، وإن تبايناتها مجهولة ومتساوية، وعند تحقق هذه الشروط يمكن استخدام اختبار (F) لمعرفة معنوية الفروق المشاهدة بين متوسطات العينات العشوائية المأخوذة من هذه المجتمعات.

وبشكل عام يتضمن التحليل الإحصائي تشتت متوسطات العينات المأخوذة من المجتمعات المطلوب دراستها، ويسمى هذا التحليل بتحليل التباين أحادي التصنيف باعتباره يتضمن عاملاً واحداً فقط، وهناك تحليل التباين ثنائي التصنيف لأنه يتضمن عاملين، ونحن سوف نتناول التباين الأحادي التصنيف فقط في دراستنا .

8-2- تحليل التباين أحادي التصنيف : (One -way analysis of : varince)

إن الهدف من تحليل التباين أحادي التصنيف معرفة ما إذا كانت الفروق المشاهدة بين متوسطات ثلاث عينات أو أكثر تعكس فروقاً حقيقية بين متوسطات المجتمعات التي سحبت منها هذه العينات أم أنها فروقاً تعود للصدفة والحظ فقط. فمثلاً قد يرغب باحث ما معرفة ما إذا كان هناك فرقاً معنوياً في مقارنة كفاءة طرق عديدة لتعليم الاحصاء لطلاب الكليات في الجامعة أو رغبة أحد الباحثين في معرفة ما إذا كان هناك فرقاً معنوياً في معدل استهلاك الوقود لأنواع مختلفة من السيارات. وعليه يتم تحليل التباين الأحادي التصنيف وفق المراحل التالية:

1 ـ صياغة الفرضية :

$$H_0 = \mu_1 = \mu_2 = \mu_3$$

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$$

2 ـ الاختبار الإحصائي :

$$F = \frac{MSB}{MSE}$$

حيث إن :

MSB هي متوسط التباين بين المجموعات.

MSE هي متوسط التباين داخل المجموعات.

ونحن نعلم بأن التباين داخل المجموعات + التباين بين المجموعات = التباين الكلي حيث أن:

$$SST = SSB + SSE$$

$$SSB = n_1(\overline{x_1} - \overline{X})^2 + n_2(\overline{x_2} - \overline{X})^2 + n_3(\overline{x_3} - \overline{X})^2$$ (التباين بين المجموعات)

حيث أن : $(\overline{x_i})$ متوسط العينة

$(\overline{X})$ المتوسط العام

(في حالة حجوم العينات متساوية) $\overline{X} = \frac{\sum \overline{x}}{K}$ (المتوسط العام)

(في حالة حجوم العينات غير متساوية) يكون: $\overline{X} = \frac{n_1\overline{x_1} + n_2\overline{x_2} + n_3\overline{x_3}}{n_1 + n_2 + n_3}$

(المتوسط العام)

حيث أن : (K) عدد العينات .

$SSE = S_1^2(n_1 - 1) + S_2^2(n_2 - 1) + S_3^2(n_3 - 1)$ (التباين داخل المجموعات)

$MSE = \frac{SSE}{N-1}$ (متوسط التباين داخل المجموعات) ،

$MSB = \frac{SSB}{K-1}$ (متوسط التباين بين المجموعات)

N : الحجم الكلي للعينات .

ولبيان كيفية تطبيق تحليل التباين لدينا التمرين التالي :

تمرين (1) :

بفرض استخدام ثلاث مجموعات من الطلبة لتعليمهم الإحصاء بثلاث طرائق تعليم مختلفة وذلك لمعرفة ما إذا كان هناك فرقاً معنوياً بين نتائج هذه الطرائق. اعطي اختبار موحد في الإحصاء للمجموعات الثلاث فحصلنا على النتائج التالية:

الطريقة A	5	5	6	7	7
الطريقة B	1	2	5	5	7
الطريقة C	3	5	5	6	6

المطلوب : بفرض أن درجات الاختبار تتبع التوزيع الطبيعي، وأن العينات مستقلة، استخدم (α=0.05) لاختبار فرضية العدم القائلة بأن المتوسطات في المجموعات الثلاث متساوية .

الحل : ـــ صياغة الفرضية :

$H_0 = \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

ـ الاختبار الإحصائي :

$$F = \frac{MSB}{MSE}$$

الجدول المساعد :

المجموعة / المقياس	A	B	C
n	5	5	5
Σx	30	20	25
$\bar{x}$	6	4	5
S^2	1.2	2.18	1.2

$$SSB = n_1(\overline{x_1} - \overline{X})^2 + n_2(\overline{x_2} - \overline{X})^2 + n_3(\overline{x_3} - \overline{X})^2$$

$$\overline{X} = \frac{\overline{x_1} + \overline{x_2} + \overline{x_3}}{k} = \frac{6+4+5}{3}$$

$$SSB = 5(1)^2 + 5(-1)^2 + 5(0)^2 = 10$$

$$MSB = \frac{SSB}{K-1} = \frac{10}{3-1} = 5$$

$$SSE = S_1^2(n_1 - 1) + S_2^2(n_2 - 1) + S_3^2(n_3 - 1)$$

$$= 1.2\,(4) + 2.18\,(4) + 1.2(4) = 18.32$$

$$= \frac{18.32}{15-3} = 1.53$$

$$MSB = \frac{SSB}{N-K}$$

$$F = \frac{MSB}{MSE} = \frac{5}{1.53} = 3.27$$

ـ **إتخاذ القرار** : من جداول (F) نجد أن : F (3.885)

$\alpha = 0.05$

$V = K-1 = 2$

$V = N-K = 12$

من خلال المقارنة نجد أن : $F = 3,27 < F_{(0.05\,,\,2.12)}$ وبالتالي نقبل فرضية العدم ونقول أن متوسطات الطرق الثلاث متساوية ، وعليه يمكننا وضع نتائج تحليل التباين في الجدول الآتي:

جدول تحليل التباين (ANOVA):

مصدر التباين	مجموع المربعات	درجة الحرية	التباين	النسبة F
بين المجموعات	SSB = 10	2	MSB = 5	$\frac{5}{1.53} = 3.27$
داخل المجموعات	SSE = 18.32	12	MSE = 1.53	
الكلي	SST = 28.32	14		

تمرين :

أعطي اختبار موحد لثلاث مجموعات من المتقدمين لشغل ثلاث وظائف، بحيث تتقدم كل مجموعة منها لشغل وظيفة معينة، فيما يلي الدرجات التي حصل عليها المتقدمون في كل مجموعة من هذه المجموعات :

A	5	4	3	2	1
B	6	4	2	1	-
C	3	2	1	-	-

المطلوب :

بفرض أن الدرجات تتبع التوزيع الطبيعي، وأن العينات مستقلة، اختبر الفرض القائل بتساوي متوسطات الدرجات عند مستوى معنوية $(\alpha = 0.01)$

ـ صياغة الفرضية :

$$H_0 = \mu_1 = \mu_2 = \mu_3$$

$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$$

ـ الاختبار الاحصائي :

$$F = \frac{MSB}{MSE}$$

الجدول المساعد :

المجموعة / المقياس	A	B	C
n	5	4	3
$\sum x$	15	13	6
$\bar{x}$	3	3.25	2
S^2	2	3.69	0.66

$$SSB = n_1(\overline{x_1} - \overline{X})^2 + n_2(\overline{x_2} - \overline{X})^2 + n_3(\overline{x_3} - \overline{X})^2$$

$$\overline{X} = \frac{n_1\overline{x_1} + n_2\overline{x_2} + n_3\overline{x_3}}{n_1 + n_2 + n_3}$$ (العينات غير متساوية)

$$\overline{X} = \frac{(5)(3) + (4)(3.25) + (3)(2)}{5 + 4 + 3} = 2.83$$

$$SSB = 5(3-2.83)^2 + 4(3.25-2.83)^2 + 3(2-2.83)^2$$

$$= (0.14) + (0.70) + (2.06) = 2.90$$

$$MSB = \frac{SSB}{k-1} = \frac{2.90}{3-1} = 1.45$$

$$SSE = S_1^2(n_1 - 1) + S_2^2(n_2 - 1) + S_3^2(n_3 - 1)$$

$$= 2(5-1)+3.69(4-1)+0.66(3-1)=20.39$$

$$MSE = \frac{SSE}{N-K} = \frac{20.39}{12-3} = 2.64$$

$$F = \frac{MSB}{MSE} = \frac{1.45}{2.26} = 0.64$$

ـ **إتخاذ القرار** : من جداول (F) نجد أن : F (8.022)

$\alpha = 0.01$

$V = k-1 = 2$

$V = N - K = 12 - 3$

من خلال المقارنة نجد أن : $F = 0.63 < F_{(0.05\,,\,2,9)}$ وبالتالي نقبل فرضية العدم (H_0) ونقول ليس هناك فرقاً معنوياً بين متوسطات الدرجات لشغل الوظائف، وعليه يكون جدول تحليل التباين (ANOVA) الآتي:

مصدر التباين	مجموع المربعات	درجة الحرية	التباين	النسبة
بين المجموعات	SSB = 2.90	2	MSB = 1.45	$\frac{1.45}{2.26} = 0.64$
داخل المجموعات	SSE = 27.2	9	MSE = 2.26	
الكلي	SST = 29.92	11		

تمارين غير محلولة

تمرين (1) :

تم تخصيص ثلاثة عمال لإنجاز عمل معين حيث سجل إنتاجهم اليومي لمدة (5) أيام يبين الجدول التالي الانتاجية اليومية لهؤلاء العمال .

اليوم	1	2	3	4	5
العامل (A)	44	46	55	55	60
العامل (B)	50	56	60	64	65
العامل (C)	46	48	48	52	61

المطلوب :

بفرض أن الإنتاج اليومي يتبع التوزيع الطبيعي وأن العينات مستقلة اختبر الفرض القائل بتساوي متوسطات إنتاجية العمال الثلاثة عند مستوى معنوية $(\alpha = 0.01)$.

تمرين (2):

البيانات التالية تمثل عدد حوادث السيارات اليومية التي حدثت على أربعة طرق سريعة لها نفس الطول وذلك خلال (5) أيام :

اليوم	1	2	3	4	5
A	3	5	5	6	6
B	1	2	3	4	5
C	5	5	6	7	7
D	1	2	5	5	7

المطلوب : بفرض أن أعداد حوادث السيارات اليومية تتبع التوزيع الطبيعي وأن العينات مستقلة، اختبر الفرض القائل بتساوي متوسطات أعداد الحوادث اليومية على هذه الطرق الأربع وذلك عند مستوى معنوية $(\alpha = 0.05)$.

تمرين :

بفرض وجود ثلاثة أحياء يقطن في كل منها عدد متساوٍ تقريباً من السكان. الجدول التالي يمثل عدد الذين تم القبض عليهم شهرياً لإرتكاب جرائم معينة في كل حي من هذه الأحياء وذلك خلال فترة (6) أشهر:

الشهر	1	2	3	4	5	6
X	3	4	4	5	5	7
Y	2	3	4	5	-	-
Z	4	5	6	-	-	-

المطلوب : بفرض أن أعداد المقبوض عليهم في الأحياء الثلاث تتبع التوزيع الطبيعي وأن العينات مستقلة، اختبر الفرض القائل بتساوي متوسطات أعداد المقبوض عليهم عند مستوى معنوية ($\alpha = 0.05$) .

تمرين (3):

استخدمت ثلاثة أنواع من برامج التدريب المختلفة (A,B,C) في ثلاث مجموعات في العمال المتماثلين في سنوات الخدمة والكفاءة، فحصلنا على جدول يبين الدرجات التي حصل عليها هؤلاء العمال في كل مجموعة من هذه المجموعات كمايلي:

المجموعة (A)	78	84	86	88	94	96	92	83	75
المجموعة (B)	64	68	73	76	78	78	80	60	93
المجموعة (C)	50	73	77	79	79	79	60	63	90

المطلوب : بفرض أن الدرجات تتبع الطبيعي وأن المجموعات مستقلة، اختبر الفرض القائل بتساوي متوسطات الدرجات، عند مستوى معنوية ($\alpha = 0.01$).

الجداول الإحصائية

ملحق الجداول الإحصائية

Appendix Tables

Table I **Binomial Coefficients** معاملات التوزيع الثنائي

$$C_x^n = \binom{n}{x} = \frac{n!}{x!(n-x)!}, x = 0,1,2,\cdots,13$$

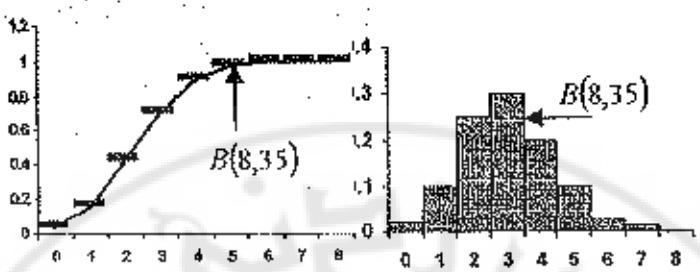
n	0	1	2	3	4	5	6	7	8
0	1								
1	1	1							
2	1	2	1						
3	1	3	3	1					
4	1	4	6	4	1				
5	1	5	10	10	5	1			
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
8	1	8	28	56	70	56	28	8	1
9	1	9	36	84	126	126	84	36	9
10	1	10	45	120	210	252	210	120	45
11	1	11	55	165	330	462	462	330	165
12	1	12	66	220	495	792	924	792	495
13	1	13	78	286	715	1287	1716	1716	1287
14	1	14	91	364	1001	2002	3003	3432	3003
15	1	15	105	455	1365	3003	5005	6435	6435
16	1	16	120	560	1820	4368	8008	11440	12870
17	1	17	136	680	2380	6188	12376	19448	24310
18	1	18	153	816	3060	8568	18564	31824	43758
19	1	19	171	969	3876	11628	27132	50388	75582
20	1	20	190	1140	4845	15504	38760	77520	125970
21	1	21	210	1330	5985	20349	54264	116280	203490
22	1	22	231	1540	7315	26334	74613	170544	319770
23	1	23	253	1771	8855	33649	100947	245157	490314
24	1	24	276	2024	10626	42504	134596	346104	735471
25	1	25	300	2300	12650	53130	177100	480700	1081575
26	1	26	325	2600	14950	65780	230230	657800	1562275

تابع الجدول (I) ..

$$C_x^n = \binom{n}{x} = \frac{n!}{x!(n-x)!}, x = 0,1,2,\cdots,13$$

n	9	10	11	12	13
0					
1					
2					
3					
4					
5					
6					
7					
8					
9	1				
10	10	1			
11	55	11	1		
12	220	66	12	1	
13	715	286	78	13	1
14	2002	1001	364	91	14
15	5005	3003	1365	455	105
16	11440	8008	4368	1820	560
17	24310	19448	12376	6188	2380
18	48620	43758	31824	18564	8568
19	92378	92378	75582	50388	27132
20	167960	184756	167960	125970	77520
21	293930	352716	352716	293930	203490
22	497420	646646	705432	646646	497420
23	817190	1144066	1352078	1352078	1144066
24	1307504	1961256	2496144	2704156	2496144
25	2042975	3268760	4457400	5200300	5200300
26	3124550	5311735	7726160	9657700	10400600

Table II The Binomial Distribution جدول التوزيع الثنائي



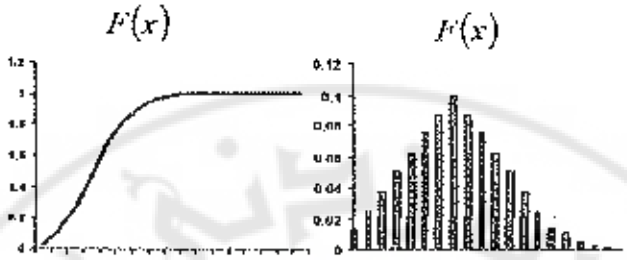
$$F(x) = P(X \le x)\sum_{k=0}^{x}\binom{n}{k}p^{k}(1-p)^{n-k}$$

		p								
n	x	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
n	x									
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664
	1	0.9928	0.9720	0.9393	0.8960	0.8438	0.7840	0.7183	0.6480	0.5748
	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360	0.9089
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
n	x									
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910
	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585
	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
n	x									
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562
	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931
	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688
	4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898	0.9815
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
n	x									
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277
	1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333	0.1636
	2	0.9978	0.9842	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443	0.4415
	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208	0.7447
	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590	0.9308
	5	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959	0.9917

	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
n	x									
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152
	1	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024
	2	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5323	0.4199	0.3164
	3	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8002	0.7102	0.6083
	4	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9444	0.9037	0.8471
	5	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9910	0.9812	0.9643
	6	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9984	0.9963
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
n	x									
8	0	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084
	1	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1691	0.1064	0.0632
	2	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.4278	0.3154	0.2201
	3	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.7064	0.5941	0.4770
	4	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.8939	0.8263	0.7396
	5	1.0000	1.0000	0.9998	0.9988	0.9958	0.9887	0.9747	0.9502	0.9115
	6	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9964	0.9915	0.9819
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9983
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
n	x									
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046
	1	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1211	0.0705	0.0385
	2	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.3373	0.2318	0.1495
	3	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.6089	0.4826	0.3614
	4	1.0000	0.9991	0.9944	0.9804	0.9511	0.9012	0.8283	0.7334	0.6214
	5	1.0000	0.9999	0.9994	0.9969	0.9900	0.9747	0.9464	0.9006	0.8342
	6	1.0000	1.0000	1.0000	0.9997	0.9987	0.9957	0.9888	0.9750	0.9502
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9986	0.9962	0.9909
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233
	2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996
	3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660
	4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044
	5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384
	6	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980
	7	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997

تابع للجدول II ..

Table III The Poisson Distribution جدول توزيع بواسون



$$F(x) = P(X \le x) = \sum_{k=0}^{x} \frac{\lambda^k e^{-\lambda}}{k!}$$

$\lambda = E(X)$

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	0.905	0.819	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	0.995	0.982	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2	1.000	0.999	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.996
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$\lambda = E(X)$

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
0	0.333	0.301	0.273	0.247	0.223	0.202	0.183	0.165	0.150	0.135
1	0.699	0.663	0.627	0.592	0.558	0.525	0.493	0.463	0.434	0.406
2	0.900	0.879	0.857	0.833	0.809	0.783	0.757	0.731	0.704	0.677
3	0.974	0.966	0.957	0.946	0.934	0.921	0.907	0.891	0.875	0.857
4	0.995	0.992	0.989	0.986	0.981	0.976	0.970	0.964	0.956	0.947
5	0.999	0.998	0.998	0.997	0.996	0.994	0.992	0.990	0.987	0.983
6	1.000	1.000	1.000	0.999	0.999	0.999	0.998	0.997	0.997	0.995
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$\lambda = E(X)$

x	2.2	2.4	2.6	2.8	3.0	3.2	3.4	3.6	3.8	4.0
0	0.111	0.091	0.074	0.061	0.050	0.041	0.033	0.027	0.022	0.018
1	0.355	0.308	0.267	0.231	0.199	0.171	0.147	0.126	0.107	0.092
2	0.623	0.570	0.518	0.469	0.423	0.380	0.340	0.303	0.269	0.238

تابع للجدول III ...

4	0.928	0.904	0.877	0.848	0.815	0.781	0.744	0.706	0.668	0.629
5	0.975	0.964	0.951	0.935	0.916	0.895	0.871	0.844	0.816	0.785
6	0.993	0.988	0.983	0.976	0.966	0.955	0.942	0.927	0.909	0.889
7	0.998	0.997	0.995	0.992	0.988	0.983	0.977	0.969	0.960	0.949
8	1.000	0.999	0.999	0.998	0.996	0.994	0.992	0.988	0.984	0.979
9	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.996	0.994	0.992
10	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$\lambda = E(X)$

x	4.2	4.4	4.6	4.8	5.0	5.2	5.4	5.6	5.8	6.0
0	0.015	0.012	0.010	0.008	0.007	0.006	0.005	0.004	0.003	0.002
1	0.078	0.066	0.056	0.048	0.040	0.034	0.029	0.024	0.021	0.092
2	0.210	0.185	0.163	0.143	0.125	0.109	0.095	0.082	0.072	0.238
3	0.395	0.359	0.326	0.294	0.265	0.238	0.213	0.191	0.170	0.433
4	0.590	0.551	0.513	0.476	0.440	0.406	0.373	0.342	0.313	0.629
5	0.753	0.720	0.686	0.651	0.616	0.581	0.546	0.512	0.478	0.785
6	0.867	0.844	0.818	0.791	0.762	0.732	0.702	0.670	0.638	0.889
7	0.936	0.921	0.905	0.887	0.867	0.845	0.822	0.797	0.771	0.949
8	0.972	0.964	0.955	0.944	0.932	0.918	0.903	0.886	0.867	0.979
9	0.989	0.985	0.980	0.975	0.968	0.960	0.951	0.941	0.929	0.992
10	0.996	0.994	0.992	0.990	0.986	0.982	0.977	0.972	0.965	0.997
11	0.999	0.998	0.997	0.996	0.995	0.993	0.990	0.988	0.984	0.999
12	1.000	0.999	0.999	0.999	0.998	0.997	0.996	0.995	0.993	1.000
13	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.998	0.997	1.000
14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	1.000
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$\lambda = E(X)$

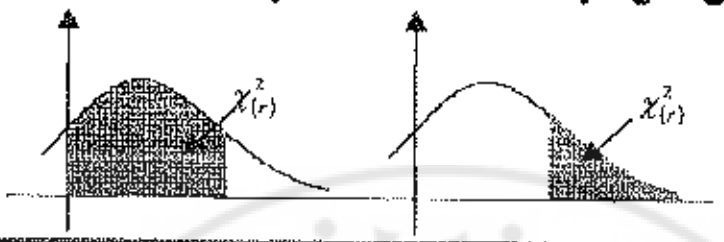
x	6.5	7	7.5	8	8.5	9	9.5	10	10.5	11.0
0	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.011	0.007	0.005	0.003	0.002	0.001	0.001	0.000	0.000	0.000
2	0.043	0.030	0.020	0.014	0.009	0.006	0.004	0.003	0.002	0.001
3	0.112	0.082	0.059	0.042	0.030	0.021	0.015	0.010	0.007	0.005
4	0.224	0.173	0.132	0.100	0.074	0.055	0.040	0.029	0.021	0.015
5	0.369	0.301	0.243	0.191	0.150	0.116	0.089	0.067	0.050	0.038
6	0.527	0.450	0.378	0.313	0.256	0.207	0.165	0.130	0.102	0.079
7	0.673	0.599	0.525	0.453	0.386	0.324	0.269	0.220	0.179	0.143
8	0.792	0.729	0.662	0.593	0.523	0.456	0.392	0.333	0.279	0.232
9	0.877	0.830	0.776	0.717	0.653	0.587	0.522	0.458	0.397	0.341
10	0.933	0.901	0.862	0.816	0.763	0.706	0.645	0.583	0.521	0.460
11	0.966	0.947	0.921	0.888	0.849	0.803	0.752	0.697	0.639	0.579
12	0.984	0.973	0.957	0.936	0.909	0.876	0.836	0.792	0.742	0.689
13	0.993	0.987	0.978	0.966	0.949	0.926	0.898	0.864	0.825	0.781
14	0.997	0.994	0.990	0.983	0.973	0.959	0.940	0.917	0.888	0.854

15	0.999	0.998	0.995	0.992	0.986	0.978	0.967	0.951	0.932	0.907
16	1.000	0.999	0.998	0.996	0.993	0.989	0.982	0.973	0.960	0.944
17	1.000	1.000	0.999	0.998	0.997	0.995	0.991	0.986	0.978	0.968
18	1.000	1.000	1.000	0.999	0.999	0.998	0.996	0.993	0.988	0.982
19	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.994	0.991
20	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.997	0.995
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

$\lambda = E(X)$

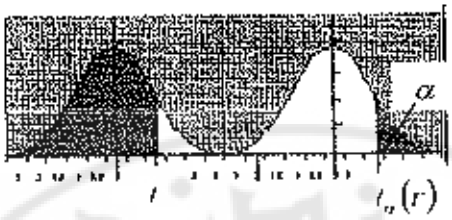
x	11.5	12	12.5	13	13.5	14	14.5	15	15.5	16
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.003	0.002	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000
4	0.011	0.008	0.005	0.004	0.003	0.002	0.001	0.001	0.001	0.000
5	0.028	0.020	0.015	0.011	0.008	0.006	0.004	0.003	0.002	0.001
6	0.060	0.046	0.035	0.026	0.019	0.014	0.010	0.008	0.006	0.004
7	0.114	0.090	0.070	0.054	0.041	0.032	0.024	0.018	0.013	0.010
8	0.191	0.155	0.125	0.100	0.079	0.062	0.048	0.037	0.029	0.022
9	0.289	0.242	0.201	0.166	0.135	0.109	0.088	0.070	0.055	0.043
10	0.402	0.347	0.297	0.252	0.211	0.176	0.145	0.118	0.096	0.077
11	0.520	0.462	0.406	0.353	0.304	0.260	0.220	0.185	0.154	0.127
12	0.633	0.576	0.519	0.463	0.409	0.358	0.311	0.268	0.228	0.193
13	0.733	0.682	0.628	0.573	0.518	0.464	0.413	0.363	0.317	0.275
14	0.815	0.772	0.725	0.675	0.623	0.570	0.518	0.466	0.415	0.368
15	0.878	0.844	0.806	0.764	0.718	0.669	0.619	0.568	0.517	0.467
16	0.924	0.899	0.869	0.835	0.798	0.756	0.711	0.664	0.615	0.566
17	0.954	0.937	0.916	0.890	0.861	0.827	0.790	0.749	0.705	0.659
18	0.974	0.963	0.948	0.930	0.908	0.883	0.853	0.819	0.782	0.742
19	0.986	0.979	0.969	0.957	0.942	0.923	0.901	0.875	0.846	0.812
20	0.992	0.988	0.983	0.975	0.965	0.952	0.936	0.917	0.894	0.868
21	0.996	0.994	0.991	0.986	0.980	0.971	0.960	0.947	0.930	0.911
22	0.998	0.997	0.995	0.992	0.989	0.983	0.976	0.967	0.956	0.942
23	0.999	0.999	0.998	0.996	0.994	0.991	0.986	0.981	0.973	0.963
24	1.000	0.999	0.999	0.998	0.997	0.995	0.992	0.989	0.984	0.978
25	1.000	1.000	0.999	0.999	0.998	0.997	0.996	0.994	0.991	0.987
26	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.995	0.993
27	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.997	0.996
28	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999	0.998
29	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
31	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
32	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
33	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
34	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table V The Chi-square Distribution توزيع مربع كاي



	0.01	0.025	0.05	0.100	0.900	0.950	0.975	0.990	0.995
r	$\chi^2_{0.99(r)}$	$\chi^2_{0.975(r)}$	$\chi^2_{0.95(r)}$	$\chi^2_{0.90(r)}$	$\chi^2_{0.10(r)}$	$\chi^2_{0.05(r)}$	$\chi^2_{0.025(r)}$	$\chi^2_{0.01(r)}$	$\chi^2_{0.005(r)}$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.086	16.750
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.647	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	4.107	5.009	5.892	7.041	19.812	22.362	24.736	27.688	29.819
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.403
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	12.878	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.335
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
60	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321

Table VI The t Distribution توزيع استيودنت

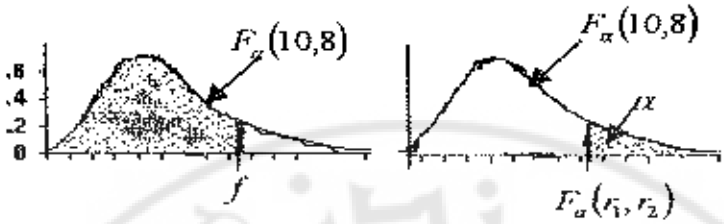


$$P(T \le t) = \int_{-\infty}^{t} \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)\left(1 + w^2/r\right)^{(r+1)/2}}\,dw$$

$$P[(T \le -t) = 1 - P(T \le -t)]$$

	0.60	0.75	0.90	0.95	0.975	0.99	0.995
r	$t_{0.40}(r)$	$t_{0.25}(r)$	$t_{0.10}(r)$	$t_{0.05}(r)$	$t_{0.025}(r)$	$t_{0.01}(r)$	$t_{0.005}(r)$
1	0.325	1.000	3.078	6.314	12.706	31.821	63.656
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576

Table VII The F Distribution توزيع فيشر



$$P(F \le f) = \int_0^f \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)(1+r_1 w/2)^{(r_1+r_2)/2}} dw$$

α%	r_2 ↓ / r_1 ▶	Numerator Degrees of Freedom r_1 درجات الحرية البسط								
		1	2	3	4	5	6	7	8	9
0.05	1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
0.025		647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.6	963.3
0.01		4052.2	4999.3	5403.3	5624.3	5764.0	5859.0	5928.3	5981.0	6022.4
0.05	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
0.025		38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
0.01		98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39
0.05	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
0.025		17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
0.01		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34
0.05	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
0.025		12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
0.01		21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
0.05	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
0.025		10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
0.01		16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
0.05	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
0.025		8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
0.01		13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
0.05	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
0.025		8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
0.01		12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
0.05	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
0.025		7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
0.01		11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
0.05	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
0.025		7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
0.01		10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
0.05	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
0.025		6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
0.01		10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
0.05	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
0.025		6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
0.01		9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
0.05		4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59

0.025	15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
0.01		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
0.05		4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
0.025	20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
0.01		8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
0.05		4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
0.025	24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
0.01		7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
0.05		4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
0.025	30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
0.01		7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
0.05		4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
0.025	40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
0.01		7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
0.05		4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
0.025	60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
0.01		7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
0.05		3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
0.025	120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
0.01		6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
0.05		3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
0.025	∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
0.01		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

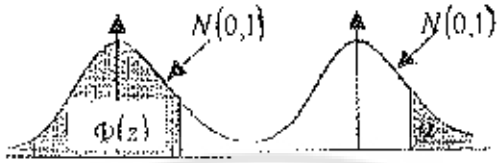
Table VII The F Distribution (Fisher) توزيع فيشر (يتبع)

α	r_2 \ r_1 →	Numerator Degrees of Freedom r_1 درجات الحرية للبسط							
		12	15	20	24	30	40	60	120
0.05		243.90	245.95	248.02	249.1	250.10	251.14	252.20	253.25
0.025	1	976.72	984.87	993.08	997.27	1001.4	1005.6	1009.8	1014.0
0.01		6107	6157	6209	6234	6260.4	6286.4	6313	6340
0.05		19.41	19.43	19.45	19.5	19.46	19.47	19.48	19.49
0.025	2	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49
0.01		99.42	99.43	99.45	99.46	99.47	99.48	99.48	99.49
0.05		8.74	8.70	8.66	8.6	8.62	8.59	8.57	8.55
0.025	3	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95
0.01		27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22
0.05		5.91	5.86	5.80	5.8	5.75	5.72	5.69	5.66
0.025	4	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31
0.01		14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56

Table VIII جدول الأرقام العشوائية في الفترة (0,1)

9588	0198	3820	0894	6705	2742	4569
4900	7466	7496	2909	4978	3722	3079
7521	2725	3902	4063	0949	8802	7023
5843	8325	0528	9800	2068	9029	9665
3949	5233	7119	8639	8131	2475	7942
8715	3299	6692	9803	4528	3506	4945
8401	7484	8384	0224	1242	3068	3077
9207	0149	8058	6870	8529	7588	6798
7736	7574	8588	1319	1872	8386	2689
6520	7841	7373	6333	0333	7685	6023
4165	7946	3533	1031	3257	7900	7313
4838	1276	1246	7801	6225	0457	4350
7947	9142	2410	4292	9920	7496	3669
0253	3400	6426	1920	4342	1146	5190
8348	7762	5936	2109	1036	0439	5598
6720	7203	2655	2512	6737	4896	0476
0628	4188	5390	4364	9341	5278	5372
3661	2808	6778	9171	3486	0070	8395
9784	8036	3352	4960	3833	1634	2902
6119	7492	6927	0885	5729	5241	3936
7482	6518	9499	2966	7333	4457	5776
5938	9499	8154	6110	2970	2117	8560
4725	7083	3201	7851	0122	0435	5401
4298	3395	5982	9762	3821	6014	4061
8745	3018	6243	7117	6795	7009	4831
2084	5778	1475	9227	6156	7372	7152
2578	7162	3083	1147	2882	2377	9183
0045	7556	2820	1923	6061	6569	1805
8138	7994	0808	5852	3319	8838	7975
4187	5360	9691	0736	2104	5497	3998
9239	6149	2790	8068	0064	5110	2736
6640	0233	9391	4707	6740	8089	4086
5314	1778	0038	4353	3054	3854	7375
1982	4704	4771	5948	5567	8559	6848
6178	7492	1538	8617	8481	2830	0787
2817	6883	8082	0589	6475	0674	8508
3270	8611	9218	7240	6695	9635	9953
2826	4639	7201	6681	2385	4960	2165
1734	1450	5790	0361	7877	7867	2580
4870	1693	8312	6607	5553	3707	1932
8207	3300	7158	1835	8846	8869	1195
8248	1060	7053	8547	9854	4010	3488
7596	8039	2547	3006	3395	4834	9408
4959	4082	2318	5127	578	4345	2857
3329	7550	9481	2988	4642	3842	4167

Table IX rd Normal Distribution التوزيع الطبيعي المعياري



$$P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw, \quad [\Phi(-z) = 1 - \Phi(z)]$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319
0.1	0.5398	0.5438	0.5478	0.5478	0.5557	0.5596	0.5636	0.5675	0.5714
0.2	0.5793	0.5832	0.5871	0.5871	0.5948	0.5987	0.6026	0.6064	0.6103
0.3	0.6179	0.6217	0.6255	0.6255	0.6331	0.6368	0.6406	0.6443	0.6480
0.4	0.6554	0.6591	0.6628	0.6628	0.6700	0.6736	0.6772	0.6808	0.6844
0.5	0.6915	0.6950	0.6985	0.6985	0.7054	0.7088	0.7123	0.7157	0.7190
0.6	0.7257	0.7291	0.7324	0.7324	0.7389	0.7422	0.7454	0.7486	0.7517
0.7	0.7580	0.7611	0.7642	0.7642	0.7704	0.7734	0.7764	0.7794	0.7823
0.8	0.7881	0.7910	0.7939	0.7939	0.7995	0.8023	0.8051	0.8078	0.8106
0.9	0.8159	0.8186	0.8212	0.8212	0.8264	0.8289	0.8315	0.8340	0.8365
1.0	0.8413	0.8438	0.8461	0.8461	0.8508	0.8531	0.8554	0.8577	0.8599
1.1	0.8643	0.8665	0.8686	0.8686	0.8729	0.8749	0.8770	0.8790	0.8810
1.2	0.8849	0.8869	0.8888	0.8888	0.8925	0.8944	0.8962	0.8980	0.8997
1.3	0.9032	0.9049	0.9066	0.9066	0.9099	0.9115	0.9131	0.9147	0.9162
1.4	0.9192	0.9207	0.9222	0.9222	0.9251	0.9265	0.9279	0.9292	0.9306
1.5	0.9332	0.9345	0.9357	0.9357	0.9382	0.9394	0.9406	0.9418	0.9429
1.6	0.9452	0.9463	0.9474	0.9474	0.9495	0.9505	0.9515	0.9525	0.9535
1.7	0.9554	0.9564	0.9573	0.9573	0.9591	0.9599	0.9608	0.9616	0.9625
1.8	0.9641	0.9649	0.9656	0.9656	0.9671	0.9678	0.9686	0.9693	0.9699
1.9	0.9713	0.9719	0.9726	0.9726	0.9738	0.9744	0.9750	0.9756	0.9761
2.0	0.9772	0.9778	0.9783	0.9783	0.9793	0.9798	0.9803	0.9808	0.9812
2.1	0.9821	0.9826	0.9830	0.9830	0.9838	0.9842	0.9846	0.9850	0.9854
2.2	0.9861	0.9864	0.9868	0.9868	0.9875	0.9878	0.9881	0.9884	0.9887
2.3	0.9893	0.9896	0.9898	0.9898	0.9904	0.9906	0.9909	0.9911	0.9913
2.4	0.9918	0.9920	0.9922	0.9922	0.9927	0.9929	0.9931	0.9932	0.9934
2.5	0.9938	0.9940	0.9941	0.9941	0.9945	0.9946	0.9948	0.9949	0.9951
2.6	0.9953	0.9955	0.9956	0.9956	0.9959	0.9960	0.9961	0.9962	0.9963
2.7	0.9965	0.9966	0.9967	0.9967	0.9969	0.9970	0.9971	0.9972	0.9973
2.8	0.9974	0.9975	0.9976	0.9976	0.9977	0.9978	0.9979	0.9979	0.9980
2.9	0.9981	0.9982	0.9982	0.9982	0.9984	0.9984	0.9985	0.9985	0.9986
3.0	0.9987	0.9987	0.9987	0.9987	0.9988	0.9989	0.9989	0.9989	0.9990

بعض القيم الهامة

α	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.005	0.001
z_α	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576	3.090
$z_{\alpha/2}$	0.842	1.036	1.282	1.645	1.960	2.240	2.576	2.807	3.291

جدول ر X

جدول إختبار الإشارة

n		p		n		p		n		p		n		p	
1	0	.50000	1	10	0	.00098	10	15	0	.00003	15	18	0	.00000	18
2	0	.25000	2		1	.01074	9		1	.00049	14		1	.00007	17
	1	.75000	1		2	.05469	8		2	.00369	13		2	.00066	16
3	0	.12500	3		3	.17188	7		3	.01758	12		3	.00377	15
	1	.50000	2		4	.37695	6		4	.05923	11		4	.01544	14
4	0	.06250	4		5	.62305	5		5	.15088	10		5	.04813	13
	1	.31250	3	11	0	.00049	11		6	.30362	9		6	.11894	12
	2	.68750	2		1	.00586	10		7	.50000	8		7	.24034	11
5	0	.03125	5		2	.03271	9	16	0	.00002	16		8	.40726	10
	1	.18750	4		3	.11328	8		1	.00026	15		9	.59274	9
	2	.50000	3		4	.27441	7		2	.00209	14	19	0	.00000	19
6	0	.01563	6		5	.50000	6		3	.01064	13		1	.00004	18
	1	.10938	5	12	0	.00024	12		4	.03841	12		2	.00036	17
	2	.34375	4		1	.00317	11		5	.10506	11		3	.00221	16
	3	.65625	3		2	.01929	10		6	.22725	10		4	.00961	15
7	0	.00781	7		3	.07300	9		7	.40181	9		5	.03178	14
	1	.06250	6		4	.19385	8		8	.59819	8		6	.08353	13
	2	.22656	5		5	.38721	7	17	0	.00001	17		7	.17964	12
	3	.50000	4		6	.61279	6		1	.00014	16		8	.32380	11
8	0	.00391	8	13	0	.00012	13		2	.00117	15		9	.50000	10
	1	.03516	7		1	.00171	12		3	.00636	14	20	0	.00000	20
	2	.14453	6		2	.01123	11		4	.02452	13		1	.00002	19
	3	.36328	5		3	.04614	10		5	.07173	12		2	.00020	18
	4	.63672	4		4	.13342	9		6	.16615	11		3	.00129	17
9	0	.00195	9		5	.29053	8		7	.31453	10		4	.00591	16
	1	.01953	8		6	.50000	7		8	.50000	9		5	.02069	15
	2	.08984	7	14	0	.00006	14						6	.05766	14
	3	.25391	6		1	.00092	13						7	.13159	13
	4	.50000	5		2	.00647	12						8	.25172	12
					3	.02869	11						9	.41190	11
					4	.08978	10						10	.58810	10
									5	.21198	9				
									6	.39526	8				
									7	.60474	7				

جدول رقم XI

Wilcoxon (Mann-Whitney) جدول ويلكوكسون (مان-وتني)

n = 1 (1) 5, m = 1 (1) 10

m	n=1			m	n=1			m	n=2			m	n=2		
1	1	0.50000	2	10	1	0.09091	11	7	3	0.02778	17	10	10	0.30303	15
					2	0.18182	10		4	0.05556	16		11	0.37879	14
2	1	0.33333	3		3	0.27273	9		5	0.11111	15		12	0.45455	13
					2	0.66667	2		4	0.36364	8		6	0.16667	14
					5	0.45455	7		7	0.25000	13	m		n=3	
3	1	0.25000	4		6	0.54545	6		8	0.33333	12	3	6	0.05	15
	2	0.50000	3						9	0.44444	11		7	0.10	14
				m		n=2			10	0.55556	10		8	0.20	13
4	1	0.20000	5		2	3	0.16667	7					9	0.35	12
	2	0.40000	4		4	0.33333	6		8	3	0.02222	19		10	0.50	11
					3	0.60000	3		5	0.66667	5		4	0.04444	18
									5	0.08889	17		4	6	0.02857	18
5	1	0.16667	6		3	3	0.10000	9		6	0.13333	16		7	0.05714	19
	2	0.33333	5		4	0.20000	8		7	0.20000	15		8	0.11429	18
	3	0.50000	4		5	0.40000	7		8	0.26667	14		9	0.20000	17
					6	0.60000	6		9	0.35556	13		10	0.31429	16
6	1	0.14286	7						10	0.44444	12		11	0.42857	15
	2	0.28571	6		4	3	0.06667	13	11	0.55556	11		12	0.77143	14
	3	0.42857	5		4	0.13333	10						13	0.68571	13
					4	0.57143	4		5	0.26667	9		9	3	0.01818	21
					6	0.40000	8		4	0.03636	20		5	6	0.01786	21
7	1	0.12500	8		7	0.60000	7		5	0.07273	19		7	0.03571	20
	2	0.25000	7						6	0.10909	18		8	0.07143	19
	3	0.37500	6		5	3	0.04762	13	7	0.16364	17		9	0.12500	18
	4	0.50000	5		4	0.09524	12		8	0.21818	16		10	0.19643	17
					5	0.19048	11		9	0.29091	15		11	0.28571	16
8	1	0.11111	9		6	0.28571	10		10	0.36364	14		12	0.39286	15
	2	0.22222	8		7	0.42857	9		11	0.45455	13		13	0.50000	14
					3	0.33333	7		8	0.57143	8		12	0.54545	12
	4	0.44444	6										6	6	0.01190	24
	5	0.55555	5		6	3	0.03571	15	10	3	0.01515	23		7	0.02381	23
					4	0.07143	14		4	0.03030	21		8	0.04762	22
9	1	0.10000	10		5	0.14286	13		5	0.06061	20		9	0.08333	21
	2	0.20000	9		6	0.21429	12		6	0.09091	19		10	0.13095	20
	3	0.30000	8		7	0.32143	11		7	0.13636	18		11	0.19048	19
	4	0.40000	7		8	0.42857	10		8	0.18182	17		12	0.27381	18
	5	0.50000	6		9	0.57143	9		9	0.24242	16		13	0.35714	17

جدول رقم XII

Wilcoxon (Mann-Whitney) جدول ويلكوكسون (مان-ويتني)

n = 1 (1) 5, m = 1 (1) 10

m	n=3			m	n=3			m	n=4			m	n=4		
6	14	0.45238	16	9	15	0.24091	24	5	15	0.14286	25	8	12	0.00808	40
	15	0.54762	15		16	0.30000	23		16	0.20635	24		13	0.01434	39
					17	0.36364	22		17	0.27778	23		14	0.02424	38
7	6	0.00833	27		18	0.43182	21		18	0.36508	22		15	0.03636	37
	7	0.01667	26		19	0.50000	20		19	0.45238	21		16	0.05455	36
	8	0.03333	25						20	0.54762	20		17	0.07677	35
	9	0.05833	24	10	6	0.00350	36						18	0.10707	34
	10	0.09167	23		7	0.00699	35	6	10	0.00476	34		19	0.14141	33
	11	0.13333	22		8	0.01399	34		11	0.00952	33		20	0.18384	32
	12	0.19167	21		9	0.02448	33		12	0.01905	32		21	0.23030	31
	13	0.25833	20		10	0.03846	32		13	0.03333	31		22	0.28485	30
	14	0.33333	19		11	0.05594	31		14	0.05714	30		23	0.34141	29
	13	0.41667	18		12	0.08042	30		15	0.08571	29		24	0.40404	28
	14	0.50000	17		13	0.10839	29		16	0.12857	28		25	0.46667	27
	15	0.58333	16		14	0.14336	28		17	0.17619	27		26	0.53333	26
									15	0.18531	27		18	0.23810	26
8	6	0.00606	30		16	0.23427	26		19	0.30476	25	9	10	0.00140	46
	7	0.01212	29		17	0.28671	25		20	0.38095	24		11	0.00280	45
	8	0.02424	28		18	0.34615	24		21	0.45714	23		12	0.00559	44
	9	0.04242	27		19	0.40559	23		22	0.54286	22		13	0.00979	43
	10	0.06667	26		20	0.46853	22						14	0.01678	42
	11	0.09697	25		21	0.53147	21	7	10	0.00303	38		15	0.02517	41
	12	0.13939	24						11	0.00606	37		16	0.03776	40
	13	0.18788	23	m		n=4			12	0.01212	36		17	0.05315	39
	14	0.24848	22	4	10	0.01429	26		13	0.02121	35		18	0.07413	38
	15	0.31515	21		11	0.02857	25		14	0.03636	34		19	0.09930	37
	16	0.38788	20		12	0.05714	24		15	0.05455	33		20	0.13007	36
	17	0.46061	19		13	0.10000	23		16	0.08182	32		21	0.16503	35
	18	0.53939	18		14	0.17143	22		17	0.11515	31		22	0.20699	34
					15	0.24286	21		18	0.15758	30		23	0.25175	33
9	6	0.00455	33		16	0.34286	20		19	0.20606	29		24	0.30210	32
	7	0.00909	32		17	0.44286	19		20	0.26364	28		25	0.35524	31
	8	0.01818	31		18	0.55714	18		21	0.32424	27		26	0.41259	30
	9	0.03182	30						22	0.39394	26		27	0.46993	29
	10	0.05000	29	5	10	0.00794	30		23	0.46364	25		28	0.53007	28
					11	0.07273	28		11	0.01587	29		24	0.53636	24
	12	0.10455	27		12	0.03175	28					10	10	0.00100	50
	13	0.14091	26		13	0.05556	27	8	10	0.00202	42		11	0.00200	49
	14	0.18636	25		14	0.09524	26		11	0.00404	41		12	0.00400	48

جدول رقم XIII

Wilcoxon (Mann-Whitney) جدول ويلكوكسون (مان-وتني)

n = 1 (1) 5, m = 1 (1) 10

m	n=4			m	n=5			m	n=5			m	n=5		
10	13	0.00699	47	6	18	0.01515	42	8	19	0.00932	51	9	34	0.34965	41
	14	0.01199	46		19	0.02597	41		20	0.01476	50		35	0.39860	40
	15	0.01798	45		20	0.04113	40		21	0.02253	49		36	0.44905	39
	16	0.02697	44		21	0.06277	39		22	0.03263	48		37	0.50000	38
					17	0.03796	43		22	0.08874	38		23	0.04662	47
	18	0.05295	42		23	0.12338	37		24	0.06371	46	10	15	0.00033	65
	19	0.07093	41		24	0.16450	36		25	0.08547	45		16	0.00067	64
	20	0.09391	40		25	0.21429	35		26	0.11111	44		17	0.00133	63
	21	0.11988	39		26	0.26840	34		27	0.14219	43		18	0.00233	62
	22	0.15185	38		27	0.33117	33		28	0.17716	42		19	0.00400	61
	23	0.18681	37		28	0.39610	32		29	0.21756	41		20	0.00633	60
	24	0.22677	36		29	0.46537	31		30	0.26185	40		21	0.00966	59
	25	0.26973	35		30	0.53463	30		31	0.31080	39		22	0.01399	58
	26	0.31768	34						32	0.36208	38		23	0.01998	57
	27	0.36663	33	7	15	0.00126	50		33	0.41647	37		24	0.02764	56
	28	0.41958	32		16	0.00253	49		34	0.47164	36		25	0.03763	55
	29	0.47253	31		17	0.00505	48		35	0.52836	35		26	0.04962	54
	30	0.52747	30		18	0.00884	47						27	0.06460	53
					19	0.01515	46	9	15	0.00050	60		28	0.08225	52
m	n=5				20	0.02399	45		16	0.00100	59		29	0.10323	51
5	15	0.00397	40		21	0.03662	44		17	0.00200	58		30	0.12721	50
	16	0.00794	39		22	0.05303	43		18	0.00350	57		31	0.15485	49
	17	0.01587	38		23	0.07449	42		19	0.00599	56		32	0.18548	48
	18	0.02778	37		24	0.10101	41		20	0.00949	55		33	0.21978	47
	19	0.04762	36		25	0.13384	40		21	0.01449	54		34	0.25674	46
	20	0.07540	35		26	0.17172	39		22	0.02098	53		35	0.29704	45
	21	0.11111	34		27	0.21591	38		23	0.02997	52		36	0.33933	44
	22	0.15476	33		28	0.26515	37		24	0.04146	51		37	0.38395	43
	23	0.21032	32		29	0.31944	36		25	0.05594	50		38	0.42957	42
	24	0.27381	31		30	0.37753	35		26	0.07343	49		39	0.47652	41
	25	0.34524	30		31	0.43813	34		27	0.09491	48		40	0.52348	40
					26	0.42063	29		32	0.50000	33		28	0.11988	47
					27	0.50000	28						29	0.14885	46
								8	15	0.00078	55		30	0.18182	45
				6	15	0.00216	45		16	0.00155	54		31	0.21878	44
					16	0.00433	44		17	0.00311	53		32	0.25924	43
					17	0.00866	43		18	0.00544	52		33	0.30320	42

جدول معامل ارتباط الرتب XIV

n	مستوى المعنوية (اختبار الجانب الواحد) .05	.01
4	1.000	
5	.900	1.000
6	.829	.943
7	.714	.893
8	.643	.833
9	.600	.783
10	.564	.746
12	.504	.701
14	.456	.645
16	.425	.601
18	.399	.564
20	.377	.534
22	.359	.508
24	.343	.485
26	.329	.465
28	.317	.448
30	.306	.432

المصطلحات الاحصائية

Statistical Terminology

إنكليزي – عربي

A

Absolute	مطلق
Absolute value	القيمة المطلقة
Accuracy	دقة
Adjustment data	تعديل البيانات
Aggregation	تجميع
Allowance Value	حدود السماح
Aggregate	كلي، شامل
Analysis of variance	تحليل التباين
Applied Statistics	الإحصاء التطبيقي
Arbitrary constant	ثابت اختياري
Area	مساحة
Area under the normal curve	مساحة تحت المنحني الطبيعي
Array	ترتيب
Association	اقتران
Alternative hypothesis	الفرض البديل

Asymmetry	عدم التماثل
Asymmetrical	غير متماثل
Asymptotic	تقريبي
Attribute	صفة
Average	متوسط
Axis	محور
Axis of abscissas	المحور السيني أو الأفقي
Axis of ordinates	المحور العمودي

B

Base	أساس أو قاعدة
Bayes theorem	نظرية بايز
Basic statistics	إحصاءات أساسية
Battery of test	مجموعة الاختبارات
Best fit	التوفيق الأفضل
Between Groups	بين المجموعات
Bias in surveys	التحيّز في المسوح
Binominal Distribution	توزيع ثنائي الحدين
Bivariate	ذو متغيرين
Breakdown	تقسيم

Business Cycles	دورات اقتصادية
Business indicators	مؤشرات اقتصادية

C

Calculation	حساب
Calculus of probability	حساب الاحتمالات
Census	تعداد
Census enumerator	عداد
Census schedule	استمارة التعداد
Causation	سببية
Central tendency	النزعة المركزية
Central limit theorem	نظرية الحد المركزي
Change	تغير
Chi-Squared test	اختبار كاي مربع
Classification	تصنيف
Class frequency	تكرار الفئة
Code	رمز
Coefficient	معامل
Coefficient of Correlation	معامل الارتباط
Coefficient of Detemination	معامل التحديد

Column	عمود
Confidence Intervals	فترات الثقة
Confidence limits	حدود الثقة
Consistency	اتساق
Constant	ثابت
Contingency	توافق
Contingency table	جدول التوافق
Continuous	متصل، مستمر
Control	مراقبة
Coordinate	إحداثي
Coordinate axis	محور الإحداثيات
Corrected rate	معدل مصحح
Correction factor	عامل التصحيح
Correlation	ارتباط
Correlation table	جدول ارتباطي
Correlation ratio	نسبة الارتباط
Classical probability	احتمال كلاسيكي
Conditional probability	احتمال شرطي
Continuous distribution	توزيع متصل (مستمر)

Critical region	منطقة حرجة
Curves	منحنيات

D

Data	بيانات
Deduction	استنباط
Degrees of freedom	درجات الحرية
Dependent variable	متغير تابع
Derivative function	دالة مشتقة
Design of experiments	تصميم التجارب
Determinant	محددة
Deviation	انحراف
Diagram (figure)	شكل بياني
Differential	تفاضلي
Discrete distribution	توزيع متقطع
Dispersion	تشتت
Distributions of probability	توزيعات احتمالية
Distribution function	دالة التوزيع
Double sampling	معاينة مزدوجة
Downward bias	تحيز نحو الأدنى

Downward trend	اتجاه هابط

E

Effect	تأثير
Efficiency of estimates	فعالية التقدير
Elimination	حذف
Enumerator	عداد
Equation	معادلة
Error	خطأ
Error of Estimate	خطأ التقدير
Error of Variance	تباين الأخطاء
Estimate / Estimation	تقدير
Estimation equation	معادلة التقدير
Event	واقعة / حادثة
Expected value	القيمة المتوقعة
Experimental error	الخطأ التجريبي
Experimental units	وحدات تجريبية
Exponential Curve	منحني آسي
Exponential equation	معادلة آسية

F

" F " Table	جدول " ف "
Factor	عامل
Factor analysis	تحليل عوامل
Factorial analysis	تحليل عاملي
Family budget	ميزانية الأسرة
Fluctuations	تقلبات
Forecasting	تنبؤ
Forms	استمارات
Frame	إطار
Frequency	تكرار
Frequency Curve	منحنى التكرار
Frequency density	كثافة التكرار
Frequency distribution	توزيع التكرار
Frequency polygon	مضلع التكرار
Frequency table	جدول التكرار
Function	دالة
Function relationship	علاقة دالة

G

Gaussian curve	منحنى غوس
Geometric mean	وسط هندسي
Graduation	تدريج
Graph	رسم بياني
Graphic analysis	تحليل بياني
Graphic presentation	عرض بياني
Goodness of fit test	اختبار جودة التوفيق
Grouping	تجميع / تصنيف
Grouping error	خطأ التجميع
Grouped data	بيانات مبوبة

H

Harmonic equation	دالة توافقية
Harmonic mean	وسط توافقي
Heteroscedasticity	اختلاف التباين
Histogram	المدرج التكراري
Homogeneous	متجانس
Hypothesis testing	اختبار الفروض

I

Ideal number	الرقم الأمثل
Identity	متطابقة
Independent variable	متغير مستقل
Index	دليل أو مقياس
Index number	الرقم القياسي
Infinite	لانهائي
Inferential statistics	احصاء استدلالي
Interclass correlation	ارتباط بين الفترات
Inersection of sets	تقاطع المجموعات
Interval estimation	فترة، مجال التقدير
Interview	مقابلة / استجواب
Inverse correlation	ارتباط عكسي
Inverse function	دالة عكسية

J

Joint distribution	توزيع مشترك
Joint regression	انحدار مشترك

K

Kurtosis	تفرطح

L

Large sample	عينات كبيرة
Law of large numbers	قانون الأعداد الكبيرة
Least squares	المربعات الصغرى
Least squares method	طريقة المربعات الصغرى
Level of significane	مستوى المعنوية
Likelihood	إمكان
Limit	نهاية أو حد
Linear	خطي / مستقيم
Linear correlation	ارتباط خطي
Linear equation	معادلة خطية
Linear interpolation	استكمال خطي
Linear regression	انحدار خطي
Linear trend	اتجاه مستقيم

M

Master Sample	عينة رئيسة
Matrix	مصفوفة
Miximum variation	أقصى اختلاف

Mean	وسط
Method of least squanes	طريقة المربعات الصغرى
Median	الوسيط
Model	نموذج
Moment	عزم
Muliple correlation	ارتباط متعدد
Multiple factor analysis	تحليل عاملي متعدد
Multipleication of probabilities	ضرب الاحتمالات
Mutually exclusive events	احداث متنافية

N

Negative correlation	ارتباط سالب
Non-Linear correlation	ارتباط غير خطي
Non-Statistics	إحصاء لا معلمي
Normal curve	منحنى طبيعي
Normal distribution	توزيع طبيعي
Normal equation	معادلة طبيعية
Null hupothesis	فرضية العدم

O

Observation	مشاهدة
One – tailed test	اختبار من طرف واحد

P

Parabola	قطع مكافئ
Parameter	معلمة، ثابت
Partial correlation	ارتباط جزئي
Partial regression	انحدار جزئي
Percentage	نسبة مئوية
Population (statistical)	مجتمع (إحصائي)
Positive correlation	ارتباط موجب
Prediction	تنبؤ
Point estimation	تقدير بنقطة
Poisson distribution	توزيع بواسون
Power of a test	قوة الاختبار
Probability	احتمال
Prior probability	الاحتمال المسبق
Probability density	كثافة الاحتمال
Probability sampling	المعاينة الاحتمالية

Q

Qualitative analysis	تحليل وصفي
Quality control	مراقبة الجودة
Quantitative analysis	تحليل كمي

Quanitity index	الرقم القياسي للكميات
Questionnaire	استبيان / استمارة
Quota sampling	عينة الحصص

R

Random distribution	توزيع عشوائي
Random error	خطأ عشوائي
Random numbers	أرقام عشوائية
Random sample	عينة عشوائية
Random selection	اختيار عشوائي
Random variable	متغير عشوائي
Range	مدى
Rank correlation	ارتباط الرتب
Rank correlation coefficent	معامل ارتباط الرتب
Rate	معدل
Ratio	نسبة
Ratio estimate	تقدير نسبي
Ratio test	اختبار النسب
Regression	انحدار
Regression coefficient	معامل الانحدار

Regression equation	معادلة الانحدار
Regression estimate	تقدير الانحدار
Regression line	خط الانحدار
Rejection region	منطقة الرفض

S

Sample	عينة
Sample census	مسح بالعينة
Sample design	تصميم العينة
Sampling	معاينة
Sampling error	خطأ المعاينة
Sampling method	طريقة المعاينة
Sampling unit	وحدة المعاينة
Sample size	حجم العينة
Sample space	فضاء العينة
Scatter diagram	شكل الانتشار
Second degree curve	منحنى من الدرجة الثانية
Set theory	نظرية المجموعات
Sensitivity	حساسية
Sequential analysis	تحليل تتابعي

Significance level	مستوى المعنوية
Simple correlation	ارتباط بسيط
Small samples	عينات صغيرة
Sources of error	مصدر الخطأ
Spurious correlation	ارتباط وهمي
Square	مربع
Standard normal distribution	توزيع طبيعي معياري
Standard deviation	انحراف معياري
Standard error	خطأ معياري
Standard error of estimate	خطأ معياري للتقدير
Standard error of the difference between means	خطأ معياري للفرق بين الوسطين
Statistical induction	استنتاج إحصائي
Statistical inference	استدلال إحصائي
Statistical method	الطريقة الإحصائية
Statistical probability	احتمال إحصائي
Statistical processing	التجهيز الإحصائي
Statistics	إحصاء
Survey	مسح (استقصاء)

Student's – distribution	توزيع (t) ستودنت

T

' T ' Table	جدول " ت " ستودنت
Table	جدول
Tabulation	جدول أو تبويب
Test	اختبار
Test of goodness of fit	اختبار جودة التوفيق
Test of homogeneity	اختبار التجانس
Test of significance	اختبار المعنوية
Third degree curve	منحنى من الدرجة الثالثة
Time series	سلسلة زمنية
Two – tailed test	اختبار من طرفين
Type I error	خطأ من النوع الأول
Type II error	خطأ من النوع الثاني

U

Union of sets	اتحاد المجموعات
Unbiased estimate	تقدير غير متحيز
Unit of measurement	وحدة القياس
Universe (statistical)	مجتمع (إحصائي)
Upward tend	اتجاه صاعد

V

Variable	متغير
Variance	تباين
Variance analysis	تحليل التباين
Variance ratio	نسبة التباين
Variation	اختلاف
Variation coefficent	معامل اختلاف
Venn diagram	شكل فن

W

Weight average	تثقيلات، أوزان
Weight average	متوسط مثقل أو مرجح

X

X – axis	محور السينات

Y

Y-axis	محور العينات
Yates continuity correction	تصحيح ياتس

Z

Z – test	اختبار (ز)

المراجع العربية والأجنبية

أولاً- المراجع العربية:

- أبو صالح، محمد صبحي؛ عوض، عدنان محمد ، 1990 ـ مقدمة في الإحصاء ـ مركز الكتب الأردني.
- أبو عمة، عبد الرحمن، راضي، الحسني عبد البر، الهندي، محمود إبراهيم، 1995، الإحصاء والاحتمالات، الرياض، جامعة الملك سعود.
- البلداوي، عبد الحميد، الإحصاء للعلوم الإدارية والتطبيقية، 1997، عمان، جامعة الإسراء.
- البشير، زين العابدين عبد الرحيم، 1997، الاستدلال الإحصائي، الرياض، جامعة الملك سعود .
- الجاعوني، فريد؛ إسماعيل، خبيب؛ غانم، عدنان، 1999، مبادئ الإحصاء، جامعة صنعاء.
- الجراد، خلف مطر، مجموعة محاضرات في الرياضيات الاقتصادية والإدارية، 1996 جامعة دمشق.
- الحلاق، عمر؛ عبيدو، أميرة، 1991، المدخل إلى الإحصاء الحيوي، منشورات جامعة حلب.
- العتوم، شفيق، 1982 ـ مقدمة في الأساليب الإحصائية، الجامعة الأردنية ـ كلية الاقتصاد والعلوم الإدارية .
- العلي، إبراهيم محمد، 1979، مدخل في نظرية العينات ـ منشورات جامعة حلب.
- العلي، إبراهيم محمد، 1980، الإحصاء الرياضي، جامعة حلب.
- العلي، إبراهيم محمد، 1985، نظرية الإحصاء، جامعة حلب.

- الأفندي، عبد القادر، 1975، نظرية الإحصاء، جامعة حلب.
- أوسيليفان، جورج، بانكروفت، جوردن، 1983، الرياضيات والإحصاء لدراسات المحاسبة والأعمال، دار ماكجرو هيل للنشر، ترجمة مقدسي، جمال سامي، ومراجعة السيد محمد الغزي.
- حميدان، عدنان؛ مخول، مطانيوس؛ الجاعوني، فريد؛ ناصر آغا، عمار، 2005، الإحصاء التطبيقي، منشورات جامعة دمشق.
- حميدان، عدنان؛ الجاعوني، فريد؛ ناصر آغا، عمار؛ العواد، منذر، 2003، مبادئ الإحصاء، منشورات جامعة دمشق.
- حميدان، عدنان؛ النعيمي، قاسم؛ ناصر آغا، عمار، 2003، مبادئ الإحصاء، منشورات جامعة دمشق، مركز التعليم المفتوح.
- حيدر، ناظم، 1982، مبادئ الإحصاء ــ منشورات جامعة دمشق.
- عبد العزيز، عمر عبد الجواد؛ بلعربي، عبد الحفيظ، 1999، مقدمة في الطرق الإحصائية مع تطبيقات تجارية ــ دار زهرات للنشر والتوزيع ــ الأردن.
- عبد المجيد، سمير، 1994، علم الإحصاء، الرياض، دار مرامر للطباعة والنشر.
- عربش، شفيق؛ مخول، مطانيوس وآخرون، محاضرات في مبادئ الإحصاء خلال السنوات الماضية في كلية الاقتصاد بجامعة دمشق.
- علي، إبراهيم، 1998، مبادئ عامة في العينات وبعض أنواع العينات ــ ورشة العمل حول مسح الهجرة الداخلية في الجمهورية العربية السورية .
- علي، نبيل، 1994، مقدمة في مبادئ الإحصاء، منشورات جامعة دمشق.

- شجيل، موراي ر.، 1972، سلسلة ملخصات شوم ـ نظريات ومسائل في الاحصاء ـ دار ماكجرو هيل للنشر.
- رمضان، محمد، 1984، الإحصاء الاجتماعي والعينات، جامعة دمشق.
- طيوب، محمود؛ غانم، عدنان، 2002، مبادئ الإحصاء، جامعة صنعاء.
- طيوب، محمود، 2001، محاضرات في الإحصاء اللامعلمي، جامعة صنعاء.
- كابوس، أمل، 1984، الإحصاء الحيوي ـ منشورات جامعة حلب.
- كاربنتر، ر أي ل؛ فاسو، إلين ستوري؛ ترجمة : حسب الله سيد؛ غندور، محمد جلال، سيد محمد، 1998، الإحصاء للمكتبين، دار المريخ للنشر، الرياض.
- كنجو، أنيس، 1995، تقنية المعاينة الإحصائية، جامعة الملك سعود، الرياض.
- مخول، مطانيوس، غانم، عدنان، مبادئ الإحصاء، جامعة دمشق 2006

ثانياً- المراجع الروسية :

1 ـ جرسيموفيتش، أ . ن، الإحصاء الرياضي، الطبعة الثانية، موسكو 1983.

2 ـ ايفتشينكو، ج،١ ، ميدفيديف، يو، ان ، نظرية الاحتمالات والاحصاء الرياضي، موسكو، 1984 .

3 ـ بوروفكوف، أ ، أ ، الاحصاء الرياضي، تقدير المعالم واختيار الفرضيات، موسكو، 1988 .

4 ــ كينينكا، ب، ف، نظرية الاحتمالات، موسكو ، 1994 .

ثالثاً- المراجع الإنكليزية:

1- Kolmogrove A.N, Theory Of Probability and Mathematical Statistics. Moscow " Seince', 1967.

2- Mitropolskyi. A.K., Statistical Calculations techiscs. Moscow, "Seince", 1978.

3- Druzhynin n.k , Sampling and Experiment. Gernal Logical Principles of Organization. Moscow "Statistica" 1977.

4- Fisher R.A., Statistical Methods for an analyst. Moscow 1958.

5- Cohen J. , Statistical Power Analysis. Hillsdale, New gersey, 1988.

6- Collyer c. and Enns J., analysis of Variance: The Basic Designs. Chicogo, 1976.

7- Kirk R. Experimental Dezign: Procedures for the Behavioral Sciences. California, 1982.

8- ferguson G.A, Statistical Analysis in Psychology and Education. London 1981: Mc graw HiLL.

9- Sprinthau R. C, Basic Statistical Analysis. Reading U.S.A 1982: Addison – Wesley.

10- Sche ffer W.C, Statistics for Biological Sciences. 2nd Ed. Reading. MA. U.S.A 1979: Addision -- Wesley. Bradley, James V.Distribution – Free Statistical Tests, Prentice-Hall, Inc., englewooed Cliffs, N.Y., 1968. De Groot, Morris H., Probability and Statictixs, Addision – Wesley, Inc., 1975.

11- Freedman, David, robert Pisani, and roger Purves, Statistics. W.W. Norton@Co., Inc., Inc., 1978.

12- Gibbons, Jean, Disckinson. Nonparametric Methods of quantitative Analysis. American Sciences Press, Inc., 1985, 2nd edition.

13- Groebner, David F., and Patrick W. Shannon. Business Statistics. Charles Merril Co. 1981.

14- Hogg, R.V., and A.T. Craig. Introduction to Mathematical Statistics, 1978, 4th edition.

15- Mendenhall, W. Introduction to probability and Statistics. PWS publishers, 1983, 6th edition.

PWS publishers, 1983, t6th edition.

16- Ott, L.,An Introcution to Statistical Methods and Data Anaysis. PWS-Kent Publishing Co. 1988, 3rd edition.

اللجنة العلمية :

- الدكتور امطانيوس مخول
- الدكتور محمود طيوب
- الدكتور فريد الجاعوني

المدقق اللغوي:

الدكتور ماجد أبو ماضي