1.2 **تمهيد**:

في هذا الفصل سوف نتناول جميع الاساليب والمقاييس الاحصائية التي تخص قوة الاختبار الاحصائي والتي يتم تناولها في الجانب التطبيقي .

ينتقسم علم الاحصاء الي قسمين :-

2.2 **الاحصاء الوصفي**[1]

وهو يهتم بطرق جمع البيانات وتبويبها وعرضها في صورة جداول ورسوم بيانية ثم تلخيصها بالطرق الرياضية السليمة فالاحصاء الوصفي يبين لنا كيف يتم توزيع هذه البيانات وما اذا كانت تتمركز حول قيمة معينة ام انها متباينة.

3.2 **الاحصاء الاستدلالي**[2]

هو عبارة عن الحصول على استنتاج عن مجتمع او ظاهرة من عينة عشوائية مختارة من هذا المجتمع فاذا اراد باحث اقتصادي دراسة ظاهرة الفقر فانه ربما يركز على تجمع سكاني معين يدرس فيه هذه الظاهرة ويستخدم نتائج دراسته للتعميم على التجمعات السكانية الاخرى ويستند هذا الاسلوب من الأساليب الإحصائية إلى مجموعة من النظريات الإحصائية لعل أهمها نظرية الاحتمالات ونظرية العينات اللتان تمثلان حلقة الوصل بين الإحصاء الوصفى والاستدلالي . ويسعى هذا النوع من الأساليب الإحصائية إلى الوصول إلى تقديرات لمعالم وخصائص مجتمعات الدراسة من خلال ما هو متوفر من معلومات عن العينات المختارة . من تلك المجتمعات ، فضلا عن اختبار الفروض الإحصائية عن مجتمع البحث على أساس البيانات المتاحة عن عينات الدراسة . ويطلق على هذا النوع من الأساليب أكثر من تسمية تؤدى جميعها إلى نفس المعنى فأحيانا يسمى بالإحصاء الاستدلالي ، أو الاستنباطي Inductive أو التعميمي Generalizing حيث يهدف إلى الوصول إلى تعميمات عن مجمع

[1] طرق الاحصاء – تطبيقات اقتصادية وادارية باستخدام spss –أ.د شفيق العتوم الجامعة الاردنية ص 21
[2] طرق الاحصاء – تطبيقات اقتصادية وادارية باستخدام spss –أ.د شفيق العتوم الجامعة الاردنية ص 21

الدراسة من خلال العينة المسحوبة من هذا المجتمع . ويشمل هذا النوع من الأساليب الإحصائية ، الاحتمالات ، العينات ، اختبار الفروض ، الاستدلال من خلال عينة واحدة أو أكثر وما يتضمنه ذلك من اختيارات مختلفة مثل كا2 chi^2 اختبار جاما gamma ، فاى phi ... الخ ويقصد بوظيفة الاستدلال اشتقاق النتائج من دراسة وفحص المقدمات والبيانات المتوافرة عن ظاهرة معينة. ولهذا يطلق علي عملية الإحصائية التي تستخدم الاستدلالي علي أساس المنطق الاستدلالي المبني علي نظرية الاحتمالات الرياضية فمن عينة محددة من أعمال أحد المصانع وباستخدام الأسلوب الإحصاء الاستدلالي يكون من الم مكن التنبؤ بمعدلات الزيادة في الإنتاج ومقدار التغير في نسبة الغياب وفي هذه الحالة نجد أن الدقة في التنبؤ تعتمد علي عوامل كثيرة من أهمها ملائمة الأدوات الإحصائية المستخدمة وحجم العينة محل الدارسة والإجراءات الإحصائية اتخذت عند اختيارها .

وتعتبر وظيفة الاستدلال أو الاستقراء من الأهمية بمكان فى البحث العلمى فمثلا : إذا كانت الظاهرة موضوع الدراسة والتحليل ممثلة للمجتمع الذى تنتمى إليه فانهـ يمكن الحصول على نتائج معنوية عن المجتمع بتحليل بيانات هذه الظاهرة وهو ما يعرف بالاستدلال ويعتمد هذا الأسلوب فى البحث على الشروط التى يجب توافرها حتى يكون هذا الاستدلال سليما – وبما أن الاستدلال لا يمكن أن يكون مؤكد فان لغة الاحتمال تستخدم عند عرض النتائج وتعتبر وظيفة الاستقراء لها أهمية كبيرة – فهى تمكن الباحث من الوصول إلى تعميمات عن المجتمع على أساس المعلومات المتاحة من عينة منه . وفى هذ هـالحالة فان أساليب ومقاييس الوصف يقتصر وصفها على ذلك الجزء (العينة) فقط من المجتمع – ومن هنا تأتى أهمية وظيفة الاستقراء – فهى تمكننا من وصف المجتمع (التعميم) باستخدام بيانات العينة .

إن القوانين فى العلوم الطبيعية والاجتماعية تجد برهانها عند الوقائع والحقائق الإحصائية ولذا يعد الاستقراء الاحصائى (Statistical Inference) أساسا لتطور المعرفة العلمية باعتباره البرهان لهذه القوانين . ووظيفة الاستقراء تحقق مطلبين أساسيين فى البحث : الأول تقدير خواص المجتمع والثانى اختبارات الفروض حول هذ هالخواص . ولا تقتصر هذ هالوظيفة على مجرد الاستقراء بل تقدم لنا تقييما عن مدى دقة هذا الاستقراء وأكثر من ذلك فهى تمكننا من التحكم فى مستوى الدقة وذلك بعدة طرق منها استخدام الأسلوب المناسب للمعاينة والحجم المناسب للعينة . وباختصار فان هذ هالوظيفة للإحصاء تمدنا بالاستقراء المنطقى وتختلف الأساليب المتبعة فى الاستقراء حسب طبيعة محل الاستقراء.

4.2 **ينقسم الاستدلال الإحصائي الي** :

1. التقدير الاحصائي
2. اختبارات الفروض الاحصائية

2-4-1 التقدير بنقطة[3]:

تقدر معالم المجتمع من بيانات العينة لانه من غير المعقول وغير عملي وبسبب محدودية الموارد والوقت والامكانيات الفنية , استخدام جميع مفردات المجتمع للحصول علي القيمة الفعلية للمعلمة او المعالم , كما ان الحصر الشامل او الفحص الكامل لجميع مفردات المجتمع قد يؤدي في العديد من الحالات الي تلافيها فاذا كان المطلوب مثلاً تقدير عمر المصباح الكهربائي الذي ينتجه مصنعت معين فانه من غير المعقول استخدام عدد كبير من المصابيح للوصول الي هذه القيمة والتقدير النقطي يعطي قيمة واحدة كتقدير لمعلمة المجتمع التي نهدف الي تقديرها ولابد من التميز بين بعض الكلمات المستخدمة في عملية

[3] زين العابدين عبدالرحيم- الاستدلال الاحصائي، السعودية - الرياض - 1993- ص57

التقدير فكلمة مقدر [4] estimate عبارة عن الصيغة الجبرية للقياس الاحصائي التي يتم التعويض فيها للحصول علي القيمة التي تقدر بها المعلمة والتقدير هو عبارة عن القيمة الرقمية التي نحصل عليها بالتعويض من بيانات العينة في صيغة المقدر . وعملية التقدير عبارة عن مجمل الاجراءات التي تستخدم للحصول علي قيمة التقدير .فهنالك طرق لتقدي النقطي ومن اهمها طريقة العزوم وطريقة الامكان الاكبر وطريقة المربعات الصغري .

2.4.2 **التقدير بفترة**:

تعريف فترة الثقة اذا كان المتغير المتصل x له دالة كثافة احتمال $f(x;\theta)$ *بمعلمة واحدة* θ وقدرنا هذه المعلمة بالمقدر $\hat{\theta}$ (مقدر العزوم او مقدر الامكان الاكبر او مقدر المربعات الصغري) فانه ليس من السهل الحكم علي جودة هذا المقدر او حساب الخطا في التقدير كما انه يصعب إقران هذا المقدر بدرجة ثقة معينة لان قيمة هذا المقدر تعتمد علي مشاهدات العينة والتي تتغير من عينة الي اخري علي الرغم من استخدام نفس المقدر . لقد وجد انه من المنطقي والمفيد ترك مجال حول المقدر $\hat{\theta}$ بمدي معين بدلاً من تقديرها بنقطة ويسمي هذا المدي فترة ثقة confidence interval ونهدف في هذا الاسلوب الي ان تحتوي الفترة علي قيمة θ غير المعلومة وان تكون هذه الفترة اقصر ما يمكن . وتجدر الاشارة الي ان الطريقة التي يتم فيها حساب فترة الثقة يجعلنا قادرين على ان نقرنها بدرجة ثقة معينة تسمي درجة الثقة degree of confidence كما ان الحد الادني والاعلي لهذه الفترة هما دالتان في مشاهدات العينة العشوائية ويتغيران بتغير هذه المشاهدات وبالتالي فان طول الفترة ومكان المعلمة داخل هذه الفترة هما عشوائيان ايضاً . ومن المهم ان نذكر ان حدود فترة الثقة تحسب بالاعتماد على المقدر النقطي غير المتحيز $\hat{\theta}$ كنقطة بداية وتوزيع المعاينة لهذا المقدر النقطي.

---

[4] جلال الصياد – الاستدلال الاحصائي – الرياض السعودية

5ـ2 **إختبارات الفروض** :ـ

يرتبط مفهوم اختبارات الفرضيات الاحصائية باحد فروع علم الاحصاء المتمثل بالاستدلال الاحصائي الذي يقسم من موجهة نظر ذوي الاختصاص في علم الاحصاء هو التقدير واختبارت الفروض

1ـ5ـ2 **تعريف المقصود باختبارات الفروض**[5]:

اذا كان المتغير x له دالة كثافة احتمالية $f(x:\Theta)$ فان اختبار الفرض يعني بشكل اساسي باختبار الفروض او التكهنات او الادعاءات المتعلقة بالمعلمة Θ وبشكل عام فان الحياة العملية تزخر بالادعاءات والافتراضات الشخصية بشأن العديد من الظواهر والمؤشرات او القيم غير المعلومة ولايمكن باي حال قبول مثل هذه الادعاءات او التكهنات واستخدامها او الاعتماد عليها في اتخاذ القرارات الادارية او الاقتصادية او الصحية او غيرها وتمثل الافتراضات والادعاءات حالة عدم التأكد التي تسود في مختلف المجالات ويتم الخروج من هذه الحالة بالاعتماد على دليل العينة المختارة عشوائياً من المجتمع محل الادعاء او الافتراض ومن هنا جاءت تسمية الادعاءات بالفروض الاحصائية.

ان معظم الفروض الاحصائية تتعلق بمعالم او ثوابت المجتمعات او التوزيعات مثل متوسط المجتمع وتباينه ونسبة المفردات التي لها خاية معينة والفروق بين المتوسطات والنسب والعلاقات بين المتغيرات والظواهر المختلفة كما انها تتعلق احياناً بانواع او طبيعة هذه التوزيعات ومدى ملائمة النماذج التي تمثل العلاقات بين الظواهر المختلفة.

بطريقة اخرى اختبار الفرض إحصائياً هو عبارة عن طريقة او اسلوب للتاكد من مدى صحة الفرض او الادعاء وعندما يعتمد هذا الاسلوب على بيانات الحصر الشامل فان الاستنتاج الذي يتم التواصل اليه بشان الفرض اما ان يكون نعم او لا, ولكن اذا كان الاختبار يعتمد على دليل

---

[5] طرق الاحصاء – تطبيقات اقتصادية وادارية باستخدام spss –أ.د شفيق العتوم الجامعة الاردنية ص 385

العينة فان القرار بشأن الفرض يكون القبول او الرفض باحتمالات معينة بمعنى ان الاسلوب المستخدم يترك مجالاً لاحتمال الخطأ في اتخاذ القرار.

2ـ5ـ2 **الفرضية**: *Hypothesis*:

هي ادعاء حول صحة شيء ما. وتنقسم إلى فرضية مبدئية (فرضية العدم H_0) والفرضية البديلة H_1.

3ـ5ـ2 **فرض العدم** H_0 **(Null Hypothesis)**:

هي الفرضية حول معلمة المجتمع التي نجري اختبار عليه ا باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما. وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها، وخلاف ذلك نقبلها وتعني كلمة *(Null)* انه لا يوجد فرق بين معلمة المجتمع والقيمة المدعاة (إحصائية العينة).

4ـ5ـ2 **الفرض البديل** **(**H_1**) Alternative Hypothesis**:

هي الفرضية التي يضعها الباحث كبديل عن فرضية العدم و نقبلها عندما نرفض فرضية العدم باعتبارها ليست صحيحة بناء على المعلومات المستقاة من العينة.

5ـ5ـ2 **أنواع اختبارات الفروض**:

الفروض Hypotheses هى علاقات متوقعة بين متغيرين أو أكثر ، أو هى توقعات الباحث لنتائج دراسته ، وتعد الفروض حلولاً محتملة للمشكلة موضع الدراسة ، وتعتمد صياغة الفروض على النظريات أو البحوث السابقة أو كليهما، كما أنها تستخدم المصطلحات والمتغيرات التى حددها الباحث ، والفرض هو حل للمشكلة تؤيده بعض المعلومات أو الحقائق أو الأدلة النظرية أو الدراسات السابقة ، ولكن صحته تعتمد على مدى تأييد الأدلة والشواهد والبيانات الفعلية للفرض .

6ـ5ـ2 الفرض البحثى[6] Research Hypothesis :

يشتق الفرض البحثى عادة اشتقاقاً مباشراً من إطار نظرى معين ، وهو يربط بين الظاهرة المراد تفسيرها وبين المتغير

7ـ5ـ2 الفرض الإحصائى Statistical Hypothesis :

عندما نعبر عن الفروض البحثية والصفرية بصيغة رمزية وعددية ، فإنها تسمى عادة الفروض الإحصائية ، فالفرض الإحصائى الصفرى يعد بمثابة قضية تتعلق بحدث مستقبلى أو بحدث نواتجه غير معلومة حين التنبؤ ، ولكنه يصاغ صياغة رمزية تسمح بإمكانية رفضه ، وهو ما نلجأ بالفعل إلى اختباره بالأساليب الإحصائية .

وقد يكون الفرض الإحصائى "فرض موجه Directed " وهو صياغة للفرض مع تحديد اتجاه العلاقة " موجبة أوسالبة " ، أو تحديد اتجاه للفروق بين المجموعات فى المتغير التابع

وقد يكون الفرض الإحصائى " فرض غير موجه " وهو صياغة للفرض دون تحديد اتجاه للعلاقة أو الفروق ، ومن أمثلته : توجد علاقة بين درجات التحصيل والابتكار لدى طلبة كلية التربية الأساسية . يوجد فرق دال إحصائياً بين متوسطى درجات المجموعتين التجريبية والضابطة فى التحصيل الدراسى .

عندما نقبل رفض العدم فإننا نقبله بنسبة دقة 90% أو 95% أو 99% أو غير ذلك وتسمى مستويات الثقة Significance Levels أي يوجد نسبة خطأ معين في قبولنا للفرضية المبدئية بمعنى أننا نقبل صحة الفرضية المبدئية وهي خاطئة وهذا الخطأ هو الخطأ α ويسمى مستوى المعنوية، أي إذا كان مستوى الثقة 95% $(1-\alpha)$ فان مستوى المعنوية α تساوي 5% وهي عبارة عن مساحة منطقة تحت منحنى التوزيع تمثل منطقة الرفض وتكون أما على صورة ذيل

[6] عادل بن احمد باطين – مباحث في الدلالة الاحصائية – ام القري - 1422

واحد جهة اليمين أو اليسار أو ذيلين متساويين في المساحة واحد جهة اليمين والثاني جهة اليسار.

8.5.2 اختبار الفروض من طرف واحد:

هو الاختبار الذي تبين فيه الفروض البديلة أن المعلمة للمجتمع اكبر أو اصغر من إحصائية العينة، فهناك تحديد للاتجاه.

9.5.2 اختبار الفروض من طرفين(ذيلين):

هو الاختبار الذي لا تبين فيه الفرضية البديلة أن معلمة المجتمع أكبر أو أصغر من إحصائية العينة، بل مجرد أنها تختلف .

6.2 اختبار الفروض:

اختبار الفرض هو قاعدة او اجراء تؤدي الي رفض او عدم رفض فرض العدم , ويتم اختبار الفرض بتقسيم فضاء العينة لكل النتائج الممكنة للتجربة العشوائية لجزئين غير متداخلين : احداهما النتائج التي اذا حدثت لا نرفض فرض العدم , ويسمى هذان القسمان منطقة القبول (acceptance region) ومنطقة الرفض (rejection region) بالترتيب ويطلق علي القيمة او القيم التي تفصل بين منطقة القبول ومنطقة الرفض القيمة او القيم الحرجة , ويكون القرار برفض فرض العدم اذا كانت النتيجة المشاهدة للتجربة تقع في منطقة الرفض وعدم رفضة اذا كانت في منطقة القبول .

وقد يحدث ان تؤدي النتيجة لرفض فرض العدم بينما هو في الواقع صحيح ونسمي هذا خطأ من النوع الاول (type 1 error) كذلك قد يحدث الا نرفض فرض العدم وهو في الحقيقة خطأ ونسمي هذا خطأ من النوع الثاني (type 11 error) ونرمز لاحتمال (او حجم) الخطأ من النوع الاول والثاني بـ $\propto$ و β بالترتيب , وكذلك اذا كان Ω فضاء عينة ذي n بعد وعرفنا فية منطقة الرفض c ومنطقة القبول c^* حيث c^* متممه c في فضاء العينة Ω .

واذا كانت $(x_1,x_2,\ldots\ldots,x_n)$ تمثل نتيجة العينة العشوائية التي تجري مشاهداتها فان حجم الخطأ من النوع الاول يمكن كتابته رمزياً كما يلي :

$$Prob\ (c/H_0) = prob\ [(x_1 , x_2 , \ldots.. , x_n)\in c/H_0] = \propto$$

كذلك ان حجم الخطأ من النوع الثاني ياخذ الشكل :

$$Prob\ (c/H_1) = prob\ [(x_1 , x_2 , \ldots.. , x_n)\in c/H_1] = \beta$$

وفي العادة يجري الاختبار علي اساس مشاهدة قيمة لإحصائية مناسبة μ تسمي إحصائية الاختبار (test statistic) واذا كانت V مجموعة من قيم μ التي تناظر c فاننا نرفض H_0 اذا كانت $\mu \in v$ وبالتالي فان :

$$Prob\ (\mu \in v) = \propto$$

وهناك نوعان من أنواع الخطأ error يمكن ارتكابهما في محاولة الوصول إلى القرار الصحيح بشأن الفرضية الأساسية وهما كالتالي:

- خطأ النوع الأول Type I error: ويتمثل هذا الخطأ في أننا قد نرفض الفرضية الأساسية في الوقت الذي تكون فيه صحيحة وكان يجب علينا عدم رفضها. ويدعى هذا الخطأ أيضاً "خطأ ألفا" α-error.
- خطأ النوع الثاني Type II error: ويتمثل هذا الخطأ في أننا قد لا نرفض الفرضية الأساسية في الوقت الذي كان يجب علينا رفضها لأنها غير صحيحة. ويدعى هذا الخطأ أيضاً "خطأ بيتا" β-error.

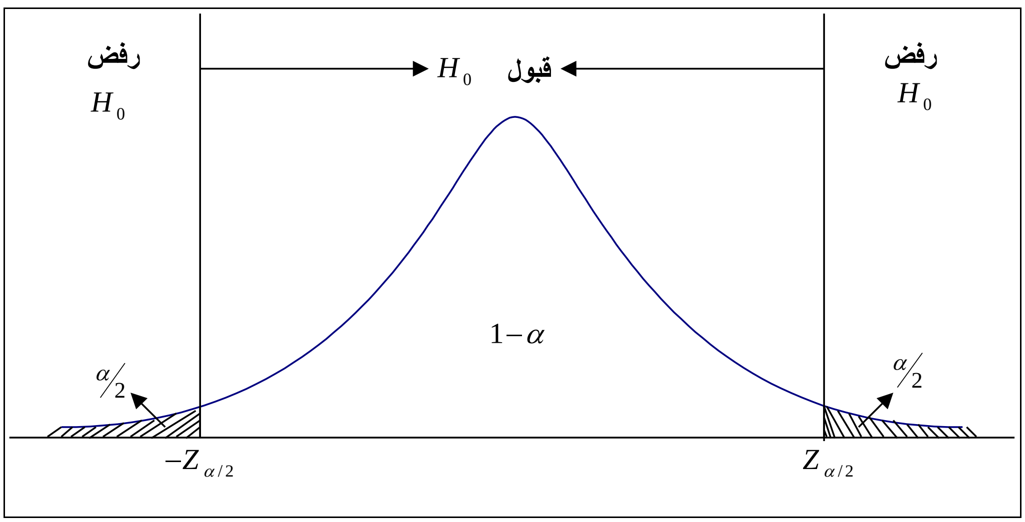
وينتج الخطأ الأول لأسباب عديدة وأهمها سبب رئيسي واحد وهو استخدام الاختبار الإحصائي الخاطئ لتحليل البيانات ومعرفة الفروقات بين المجموعات. أما الخطأ الثاني فإنه ينتج عن استخدام حجم عينة sample size صغير. كما أن بعض الاختبارات الإحصائية أقوى من غيرها في التعرف على الفروقات بين المجموعات. ويمكننا تعريف "قوة الاختبار الإحصائي

Statistical power″ على أنها "قدرة الاختبار الإحصائي على رفض الفرضية الأساسية عندما تكون غير صحيحة ويجب رفضها". ويمكن تعريف "القوة" أيضاً بالمعادلة التالية:

$$\text{Power} = 1-\beta$$

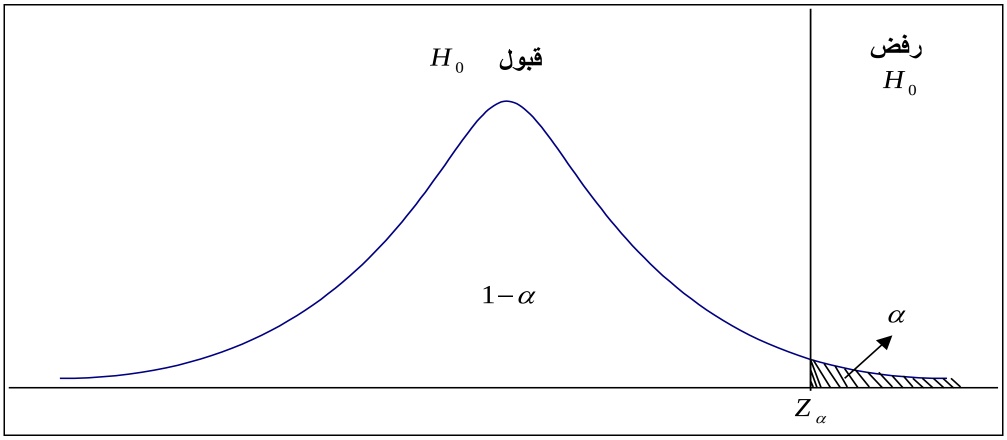
حيث أن β كما أسلفنا هي مقدار خطأ النوع الثاني وتبلغ في العادة (*0.8*). وكقاعدة عامة، فإن قوة الاختبار الإحصائي تزداد بزيادة حجم العينة *sample size*.

الشكل (2−1) يمثل المنحني الطبيعي



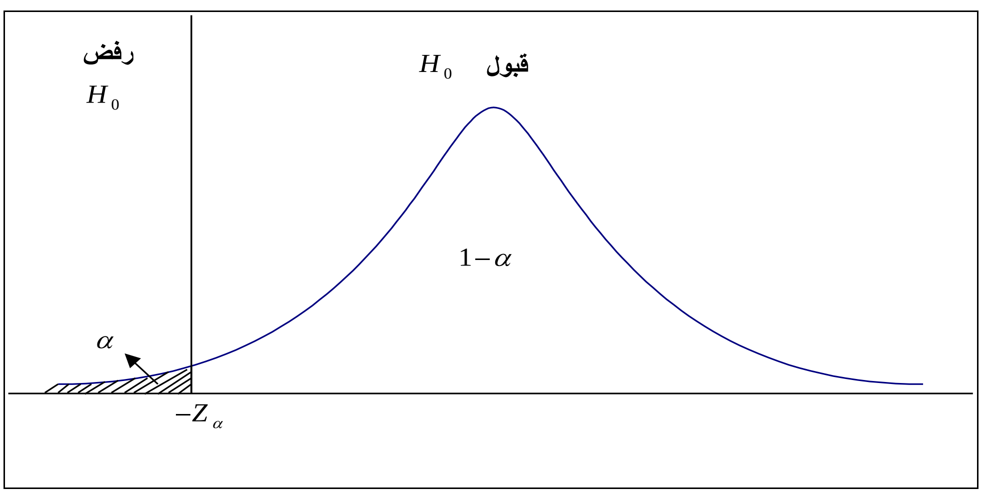
المصدر : اعداد الباحث

الشكل (2-2) يمثل المنحني الطبيعي رفض من جهه اليمين



المصدر : اعداد الباحث

الشكل (2-3) يمثل المنحني الطبيعي رفض فرض العدم



المصدر: اعداد الباحث

الجدول (2−1) تفسير النتائج المحتملة من التحليل الإحصائي:

القرار / حقيقة الفرض	الفرض العدمى H_0	الفرض البديل H_1
قبول الفرض العدمى	وحقيقة الفرض H_0 صحيح يدل على قوة الاختبار $1-\beta$	بالرغم من ان H_0 غير صحيح ينتج خطأ من النوع β
رفض الفرض العدمى	بالرغم من ان H_0 صحيح يرجع خطأ من النوع α	وحقيقة الفرض H_0 غير صحيح يدل على درجة الثقة فى الاختبار $1-\alpha$

المصدر: اعداد الباحث

ولاشك ان الاختبارالذي يقلل احتمالات الخطأ من النوعين في الوقت نفسه ولاقل قدر ممكن هو الاختبارالذي نفضله علي سواء , ولكن لحجم عينة ثابت , ولايمكن تخفيض احتمال احد احد الخطأين دون زيادة حجم الاخر والطريقة الوحيدة لتخفيضهما معاً هي زيادة حجم العينة.

7.2 **اختبار المعنوية**[7]:

يرتبط اختبار المعنوية باحد فروع علم الاحصاء المتمثل بالاستدلال الاحصائي والذي يقسم من وجهة نظر ذوي الاختصاص في علم الاحصاء الي قسمسن رئيسين هما التقدير واختبارات الفروض الاحصائية وان كافة الاجراءات التي تتدرج تحت موضوع اختبار المعنوية يطلق عليها احياناً باختبار الفرضيات ويعد موضوع اختبار المعنوية من المواضيع الاساسية للتطبيقات الاحصائية في مختلف المجالات العلمية , ويهتم هذا النوع من الاختبارات باتخاذ القرارات الاحصائية المناسبة.

[7] الاختبارات الاحصائية اسس وتطبيقات (2010) – الدكتور حسن ياسين طعمة – ص 15

وفي ضوء ماتقدم تعد الاجراءات المتعلقة باختبارات المعنوية اداة اساسية من الادوات التحليلية الاحصائية ويعود الفضل الاول في اقتراح هذه الاجراءات لكل من نيمان وبيرسون عام 1930م , وان الاجراء المتبع في إختبار الفرضيات ينطوي علي ماياتي:

صياغة فرضية معينة

اختبار الفرضية

اتخاذ قرار بشان الفرضية كنتيجة للاختبار

وعادة ماتكون الفرضية التي يتم صياغتها عن حقيقة حالة الطبيعة فعلى سبيل المثال قد تكون الفرضية تمثل نسبة ظهور احد اوجه زهر النرد مساو الي $\frac{1}{6}$ عند رمي زهرة نرد متجانسة , والمشكلة التي تظهر هنا من الناحية الاحصائية هي تحديد ما اذا كانت نتائج عينة من رميات زهرة النرد منسجمة ومتسقة مع الفرضية من خلال اتباع بعض الاجراءات حول اختبار موضوعية ان اجراءات الاختبار تنظر الي نتائج عينة الرميات على انها منسجمة وتسقة مع مع الفرضية في حالة اذا كانت الفروق لا تتعدى اختلاف المعاينة.

وفي ضوء ماتقدم يتضح بان هذه الاجراءات تفترض اختلافاَ في المجتمع الذي تنتمي اليه الفرضية , وفي حالة عدم وجود اختلاف في المجتمع فان الحاجة تنتفي من اجراء الاختبار الاحصائي , لانه قد لايوجد اختلاف في المجتمع ففي هذه الحالة قد يكون من المتوقع ان تختلف نتائج العينة , وفي كل الاحوال اذا كان هنالك اختلاف في المجتمع ففي هذه الحالة قد يكون من المتوقع ان تختلف نتائج العينة وهذا بدورة ان يؤدي الي ظهور حالة عدم التأكد.

ان حالة عد التاكد تتمثل بالفروض او بالافتراضات والتكهنات الشخصية بشان العديد من المؤشرات او القيم المجهولة (غير المعلومة) ولا يمكن باي حال من الاحوال قبول هذه الفرضيات او التكهنات او الاعتماد عليها من قبل متخذي القرارات في المجالات الادارية والاقتصادية او الطبية او الهندسية او غيرها من المجالات الاخرى في الحياة العملية , للخروج

ن حالة عدم التاكد يتم ذلك بالاعتماد على دليل العينة التي يتم اختيار مفرداتها بشكل عشوائي من بين مفرادات المجتمع الاصلي محل الافتراض (الفرضية) ومن هنا جاءات تسمية هذه الافتراضات بالفرضيات الاحصائية.

وتاسيساً على ما تقدم يمكن القول بان اغلب الفرضيات الاحصائية تتعلق بمعلمات المجتمع الاحصائي او التوزيعات الاحتمالية مثال ذلك متوسط المجتمع وتباينه والفروق بين لمتوسطات والنسب والعلاقات بين المتغيرات والظواهر المختلفة كما انها تتعلق بطبيعة هذه التوزيعات ونوعها.

وفي ضوء ما تقدم يعرف اختبار الفرضية احصائياً بانه اسلوب احصائي يراد به التاكد من صحة الفرضية او عدم صحتها , وعندما يعتمد هذا الاسلوب على جميع مفردات المجتمع الاحصائي فان الاستدلال الذي يتم التوصل اليه بشان الفرضية يكون اما نعم اة لا , اما اذا كان الاسلوب (الاختبار) يعتمد على مفردات العينة التي يتم اختيارها من المجتمع الاحصائي فان القرار الذي يتم التوصل اليه بشان الفرضية يكون اما بالقبول او الرفض باحتمالات معينة , وهذا يعني ان الاسلوب الذي يعتمد على دليل العينة يترك مجالا لاحتمال الوقوع في الخطأ في اتخاذ القرارات.

8.2 مفهوم اختبار المعنوية[8] :

يعد موضوع اختبارت المعنوية من المواضيع الاساسية في مجال اتخاذ القرارات ويستخدم هذا الاسلوب للكشف عن الاختلافات بين احدى المجموعات ومجموعة اخرى مثلاً وتنتفى الحاجة الي هذه الاختبارات عندما يتم دراسة جميع مفردات للمجتمع وهذا لايحدث دائماً بل يتم اللجؤ في اغلب الاحيان الي اختيار عينات بطريقة عشوائية ممثلة للمجتمعات ويعد موضوع اختبارات المعنوية (اختبار الفرضيات) احد فروع الاحصاء الاستدلالي الذي يتنقسم الي

[8] الاختبارات الاحصائية اسس وتطبيقات (2010) – الدكتور حسن ياسين طعمة – ص 17

قسمين رئيسين كما ذكرنا سابقاً هما التقديرات واختبارات الفروض , عليه فاننا في اغلب الاحيان لانكتفي بتقدير معلمة المجتمع μ بان تحدد لها قيمة معينة ولتكن μ_0 او نبنى لها فترة ثقة معينة , بل نحتاج الي اتخاذ قرار بشان صحة الفرضية المعينة او عدم صحتها , بمعنى اخر نحتاج الي اختبار الفرضيات المتعلقة بمعلمات المجتمع الاحصائي.

وفي ضوء ماتقدم يعد موضوع اختبار الفرضيات من المواضيع الاساسية في مجال اتخاذ القرارات ويستخدم هذا الاسلوب (الاختبار) للكشف عن الاختلافات بين احدى المجموعت ومجموعة اخرى مثلا , وتنتفي الحاجة لهذه الاختبارات عندما يتم دراسة جميع مفردات المجتمع وهذا لايحدث دائماً بل يتم اللجوء في اغلب الاحيان الي اختيار عينات بطريقة عشوائية ممثلة للمجتمعات.

9.2 الدلالة الإحصائية

عبر كارفر 1987 (المشار إليه في هوستن 1993:4) بقوله ان الدلالة الإحصائية ببساطة : تعني الندرة الإحصائية، فالنتائج قد تبدو ذات دلالة من وجهة النظر الإحصائية لأنها تظهر بشكل نادر في العينات العشوائية تحت شروط الفرض الصفري، وبالتالي فالدلالة الإحصائية تعني القليل أو لا شيء" أي أن الدلالة الإحصائية تهتم بمستوي الثقة التي نوليها للنتائج فنقول ما دام الفرق دالا عند مستوى (5%) فهذا يعني أن الفرق بين المجموعتين حقيقي وأن مجتمع المجموعة الأولي يختلف عن مجتمع المجموعة الثانية، وأننا نثق في هذا الحكم بنسبة(95%) منصور ، 1997م

وأشار البهي (1997) إلى أن الدلالة الإحصائية هي درجة اقتراب قيمة مقاييس العينات من مقاييس المجتمع الأصل. ويمكن اكتشافها بواسطة الاختبار الإحصائي وهو مجموعة من القواعد تمكن الباحث من رفض أو قبول الفرض الإحصائي .وبموجبه يمكن الحكم على الفرض الإحصائي (الصياد ، 1990)

كما أشار كاشن وجيجر *(Cashen & Geiger) 2004* إلى أن الاستدلال الإحصائي الكلاسيكي يختبر الفرض الصفري أي أنه لا توجد أي فروق بين المعالم قيد الاختبار أو أن الفروق يمكن أن تهمل، والتي يقارنها الباحثون بالفرض البديل وهي تعني ان الظاهرة قيد التحقيق حاضرة او انه توجد فروق فروق بين المعالم قيد الاختبار، ولأن الباحثون عاد ة ما يأملون في رفض الفرض العدمي، فإنهم عادة ما يرصدون الاحتمالات المرتبطة بترجيح أن هذا الفرض خاطئ

وأيد ذلك قول الصياد، بأنه تعد الطريقة الإحصائية أكثر الطرق شيوعا من حيث كونها طريقة من طرق صناعة القرار تحت شرط حالة عدم التأكد، أي شرط الاحتمالية، لذا يصيغ الباحث في ظل استخدام الطريقة الإحصائية الكلاسيكية نوعين من الفروض الإحصائية كترجمة لغرض البحث تحت الدراسة .هذان النوعان من الفروض هما الفرض العدمي والفرض البديل وعادة ما يكون هدف الباحث هو الحصول على قرار من خلال رفض الفرض العدمي ، أي قبول البديل وذلك عند مستوى دقة معينة وهذه الدقة تشتمل ضمن ما تشتمل ثلاثة مكونات رئيسية وهي الخطأ من النوع الأول والذي يطلق عليه جوازا مستوى الدلالة باعتباره حدا أقصى يرتضيه الباحث level of significance الإحصائية للخطأ من النوع الأول وقوة العلاقة بين المتغيرات التابعة والمتغيرات المستقلة

ومما سبق يلاحظ أن الدلالة الإحصائية يتم قياسها عن طريق العديد من اختبارات الفروض التي تقوم على اختبار الفرض العدمي ضد الفرض البديل والتي يهم الباحث فيها برفض الفرض العدمي حيث أن ذلك دلالة على صحة توجهه وذلك طبقاً لمستوى الدلالة وتوزيع المعاينة المناسب وتحديد القيمة الحرجة وعلى ذلك التوزيع الاحتمالي والتي تفصل بين منطقتي الرفض والقبول، ثم بمقارنة القيمة الحرجة مع القيمة المحسوبة من الاختبار الإحصائي نأخذ القرار إما برفض الفرض العدمي وبالتالي قبول الفرض البديل

10.2 مفهوم القوة الاحصائية[9]

تحليل القوة الإحصائية هو أحد الاختبارات المكملة لاختبارات الدلالة الإحصائية ، وهو يقدر احتمالية الخطأ من النوع *(II)* عندما تكون H_o (عدم رفض الفرضية الصفرية عندما تكون خاطئة) *(Wilkinson, 1992)* وقد عرفها هـ وستون بأنها (احتمال ان يرفض الاختبار الفرضية الصفرية بطريقة صحيحة)

وقد لاحظ كوهين *(1988)* وشيفر *(1980)* أنّ هذا التوجه نادراً ما يستخدم من قبل الباحثين ، وهو تحليل يتضمن التعامل مع $\propto$ و β ، وحجم العينة n و وقد لاحظ ثومبسون *(1987)* أنّ القوة إذا كانت عالية، فإنّ النتائج غير الدالة إحصائياً تسهم في أساس المعرفة، وكما يرى لاي *(1973)*بأنّ أي باحث يقبل الفرضية الصفرية بدون أن يعرف قوة الاختبار الإحصائي فإنه يكون مسؤولاً عن حدوث ، نسبة خطأ كبيرة من النوع *(II)* وأنّ الباحثين الذين يؤيدون وبطريقة غير صحيحة قبول الفرضية الصفرية إذا لم يلاحظوا أي دلالة إحصائية دون اعتبار للقوة الإحصائية يسيئون قيادة القراءة *(Wilkinson 1992)*. ويرى نيكس وبارنيت *(1998)* بأنه إذا كان لدينا اختبار منخفض للقوة فإنه لن يمكننا تحسس والتقاط الدلالة الإحصائية ، بينما إذا كان لدينا اختبار عالي القوة فإنّ حجم الفروق بغض النظر عن صغرها سوف تكون دالة إحصائياً) .

ويعتقد الكثيرون، أنّ التركيز على التحكم في الخطأ من النوع *(I)* المستعمل في الأساليب الشائعة، ربما يكون السبب وراء رؤية العديد من الدراسات الإحصائية التي تعاني من نقص في القوة وعلى العموم فإنّ التحكم في الخطأ من النوع *(I)* يجعل الباحثين يدفعون الثمن وهو العجز في التحكم في الخطأ من النوع *(II)* أي النقص في قوة الاختبار، مما يعني الفرصة الأقل للحصول على نتائج دالة إحصائياً

[9] عادل بن احمد باطين - مباحث في الدلالة الاحصائية - ام القري - 1422

11_2 قوة الاختبار الاحصائي[10]:

تعرف قوة الاختبار بانها (احتمال رفض فرض العدم عندما تكون هذه الفرضية خاطئة) اي عند اختبار فرض ما هنالك احتمال $\propto$ للوقوع في الخطأ من النوع الاول , واحتمال β للوقوع في الخطأ من النوع الثاني . وانه لايمكن تصغير الخطأين معاً الا عن طريق زيادة حجم العينة , ولاشك ان اي مقارنة بين اختبارين لمعرفة ايهما افضل يجب ان تضع في الاعتبار قيم $\propto$ و β لكل من الاختبارين ويتم ذلك عادة من خلال مفهوم قوة الاختبار .

12_2 دالة قوة الاختبار:

في حالة الفروض البسيطة توجد قيمة واحدة لـ $\propto$ وقيمة واحدة لـ β لكل اختبار , ولهذا فان معرفة هاتين القيمتين تكفي في كثير من الاحيان لتقويم الاختبار ومغارنته بغيره من الاختبارات مباشرة . اما في حالة الفروض المركبة فإنه ليس هنالك $\propto$ واحدة اذ هنالك اكثر من قيمة للمعلم يحددها فرض العدم كما انه ليس هنالك قيمة واحدة لـ β لان هنالك اكثر من قيمة يحددها الفرض البديل . فمثلاص عند اختبار H_0 ضد H_1 حيث :

$$H_0 : \theta \in \omega \, , \ H_1 : \theta \in \omega'$$

وحيث ω مجموعة محددة من قيم θ و ω' مجموعة قيم θ المتممه لـ ω في فضاء المعلم θ , فإن قيم θ التي يحددها H_0 هي مجموعة القيم التي تشملها المجموعة ω كما ان قيمة θ التي يحددها H_1 هي مجموعة القيم التي تشملها ω' . في هذه الحالة لابد لتقويم الاختبار من معرفة حجم الخطأ من النوع الاول $\propto$ لكل قيم من قيم θ التي يحددها فرض العدم , كما انه لابد من معرفة حجم الخطأ من النوع الثاني β لكل قيمة من قيم θ التي يحددها الفرض البديل بمعني اخر لابد من معرفة

$$\alpha(\theta) \qquad \theta \in \omega$$
$$\beta(\theta) \qquad \theta \in \omega'$$

[10] زين العبدين عبدالرحيم – الاستدلال الاحصائي – السعودية – الرياض - 1993

ويبرز الترميز حقيقة كون كل من الخطأ من النوع الاول والنوع الثاني دالة في θ في حالة الفروض المركبة , وتعطي الدالة $\alpha(\theta)$ حجم الخطأ من النوع الاوللكل قيم θ في ω (اي التي يحددها H_0) بينما تعطي $\beta(\theta)$ حجم الخطأ من النوع الثاني لكل قيم θ في ω' (اي التي يحددها H_1) .

وبما ان $1-\beta(\theta)$ هي احتمال رفض فرض العدم عندما تكون $\theta \in \omega'$ فإن الدالة $\pi(\theta)$ حيث :

$$\pi(\theta) = \begin{cases} \alpha\theta(\ ) & : \ \theta \in \omega \\ 1-\beta(\theta) & : \ \theta \in \omega' \end{cases}$$

تعطي احتمال رفض فرض العدم لكل قيم θ , وتسمي $\pi(\theta)$ دالة قوة الاختبار (The power function of the test) اي ان دالة قوة الاختبار هي دالة قيمتها عبارة عن احتمالات رفض فرض العدم للقيم المختلفة للمعلم θ .

كما ذكرنا انفاً ان قوة الاختبار الاحصائي هي (احتمال رفض فرض العدم عندما تكون هذه الفرضية خاطئة)[11] اي ان:

P.O.T = P(Rejecting (H_0)/(H_0) is false)
= 1- P(Accepting (H_0)/(H_0) is False)
= 1 - B

يتضح من المعادلة اعلاه بان (B) تمثل احتمال الوقوع في الخطأ من النوع الثاني اي ان :

B = P(Accepting (H_0)/(H_0) is false)

[11] حسن ياسين طعمة , ايمان حسين حنوش – الاحصاء الاستدلالي 2012

وفي ضوء ماتقدم يتضح بانه كلما كان احتمال الوقوع في الخطأ من النوع الثاني (B) قليل كلما ادى ذلك الي زيادة قوة الاختبار (B– 1) وهذا يعني زيادة شدة رفض الفرضية العدمية عندما تكون هذه الفرضية خاطئة.

وتاسيساً على ماتقدم يمكن توضيح العلاقة بين الخطأين من النوع الاول والنوع الثاني على النحو التالي:

1. ان انخفاض احدى الخطأين يؤدي الي زيادة الخطأ الاخر
2. ان زيادة حجم العينة (n) يقلل من احتمال الوقوع في كل من الخطأين وبالتالي زيادة درجة الثقة
3. يحسب احتمال الوقوع من الخطا من النوع الاول على اساس الفرضية العدمية في حين يحسب احتمال الوقوع في الخطأ من النوع الثاني على اساس الفرضية البديلة

13ـ2 معامل الالتواء:[12]

الالتوء هو درجة نماثل او بعد عن التماثل لتوزيع , اذا كان المنحني التكراري لتورزيع له ذيل اكبر الي يمين نقطة النهاية العظمي عنه الي اليسارها ويسمي التوزيع بانه ملتو الي اليمين او موجب الالتواء , اما اذا كان العكس صحيحاً فيقال انه ملتو الي اليسار او سالب الالتواء . في التوزيعات الملتوية يقع الوسط علي نفس جانب المنموال وذلك علي نفس جانب الطرف الاطول وكمقياس للتماثل ناقذ الفرق (الوسط – المنوال) وهذا المقياس يمكن تخليصه من الوحدات بقسمته علي مقاييس التشتت مثل الانحراف المعياري مما يؤدي الي التعريف التالي

$$Skewness = \frac{\bar{x} - mode}{\sigma} \;\ldots..\;(4)$$

ولتحاشي استخدام المنوال من الممكن استخدام الصيغة الاعتبارية

[12] العتوم – شفيق – طرق الاحصاء الجامعة الاردنية

$$Skewness = \frac{3(\bar{x} - median)}{\sigma} \ldots\ldots (5)$$

والمقياسان السابقان يسميان علي الترتيب معامل بيرسون الاول للالتواء ومعامل بيرسون الثاني للالتواء .

ان خاصية تماثل المنحنيات خول نقطة التماثل وكذلك حول محور التماثل والذي يهم الباحثين في دراسة التماثل هو تماثل منحنيات التوزيعات التكرارية حول وسطها الحسابي[13] .

ان صفة التماثل وصفة الانتشار وصفة التمركز من الصفات التي تحدد صورة المجتمع الاحصائي بصورة واضحة وتبين خصائصه وكل خاصية اخرى مضافة تحدد جانباً اخراً من صفات المجتمع الاحصائي وتوضح صورة اكثر وضوحاً ودقة.

ان مقاييس التشتت ومقاييس التمركز ومقاييس التماثل والتي تطرقنا اليها فانها تعطي الباحث ملامح واضحة عن طبيعة تلك المجتمعات في هذه الفقرة سوف نتطرق الي صفة الالتواء ومقاييسها وهي الصفة المناقضة الي صفة التماثل فالمنحنى غير المتماثل يكون ملتوياً والمنحنى الملتوى يكون غير متماثلاً.

كي يكون المنحنى متماثلاً حول الوسط الحسابي , نتصور وضع مراة مستوية في موقع قيمة الوسط بوضع عمودي على محور السينات فاذا كانت صورة الجزء الواقع امام المراة من المدرج التكراري في المراة منطبقة على صورة الجزء الواقع على من هذا المدرج خلفها , كان التوزيع متماثلاً حول الوسط الحسابي وبخلاف ذلك كان التوزيع ملتوياً وكلما ذات الاختلاف بين واقع التوزيع وحالة تطابق الصورة في المراة زاد التوزيع التواءً.

14.2 **العينات** :

من أهم الأسباب التي تدفعنا إلى نظرية العينات هي رغبتنا في الحصول على معلومات عن المجتمع المطلوب دراسته في الوقت المناسب، ورقابة أخطاء المعاينة بوضعها في حدود

[13] محمد صلاح الدين صدقي – جامعة القاهرة - اساليب كمية – ص 237

مقبولة حتى لا تؤدي النتائج غير معقولة إلى اتخاذ قرارات خاطئة أو غير حكيمة. ولذلك فان معرفتنا لنظرية العينات نستطيع أن نكتشف إجراءات المعاينة التي تعطي الدقة المطلوبة وعن طريق بعض المعالجات الاحصائية .علما بان الإحصاء بمفهومه الحديث، علم قائم بذاته له قواعده وأسسه الرياضية الخاصة به، فيه يستخدم على نطاق واسع في دراسة مختلفة القضايا التربوية والنفسية والاجتماعية والاقتصادية ...بالإضافة إلى العلوم الطبيعية والتطبيقية الأخرى . فلمى سبيل المثال" أن استخدام الإحصاء في التشخيص الطبي والدرسات الطبية الحيوية فان أي خطأ فيه أو عدم الدقة في استخدامه سيشكل كارثة طبية على حياة أفراد المجتمع، فضلا عن الصحة الاجتماعية والعقلية قد تكون مهدده بالخطر عندما يتجاوز الإحصائي المعايير الأخلاقية في الدراسات الاجتماعية والعقلية والنفسية وهذا ينطبق على جميع ميادين الحياة ومجالاتيا الأخرى مثل الزراعة والهندسية والتربوية[14]

وبعد الحصول على المعمومات الأولية أو الخام *Data Raw* لابد لمباحث أن يقوم بعرضها وتبويبيا بطريقة ما يجعل من السهل قرائتها وبالتالي معالجتيا رياضيا واستخلاص النتائج، وهذا يعني تفسير نتائج البيانات ومن ثم عمل استنتاجات إحصائية أو إصدار أحكام إحصائية على مجتمع الدراسة الأصلي. وهنا لابد أن نذكر" إن هذه العملية خطوة هامة وحساسة لابد للباحث أن يكون ملما بجوانب المشكلة المدروسة، وخاصة وان عملية التفسير هذه ليست عملية ذات طبيعة إحصائية بحتة بل تتوقف أيضا على المعرفة الأكاديمية المتخصصة التي يعمل فيها الباحث[15] وبناء عليه، ولذلك يمكننا تعريف أخطاء المعاينة بأنها الفرق بين النتائج التي تحصل عليها من العينة والنتائج التي يمكن الحصول عليها تحت نفس الظروف التي سحبت فيها العينة من إجراءات تعداد شامل لهذا المجتمع .ويتوقف حجم هذه الأخطاء على حجم العينة وتباين المجتمع وطريقة اختيار العينة وحساب النتائج .فإذا كبر حجم العينة كلما قلت أخطاء المعاينة وا

[14] ابويوسف محمد 1992 ص 19
[15] أحمد 1998 ص 67

زدت ثقتنا في التقديرات التي نحصل عليها من العينة ومن المعروف إن أساس المعاينة الجيدة هو اختيار الطريقة المناسبة التي ستسحب بها العينة لقياس خواص هذا المجتمع " لذا عمل الكثير من العلماء في تطوير نظرية العينات، و ممن لهم الفضل في ذلك (بيروني) و(بواسون) و(لابراس) وفى عام *1991* صدرت أعمال(ستيودنت) التي لعبت دوا ر كبيراُ فى تطوير نظرية العينات، وخاصة ما أصبح يسمى بالعينات الصغيرة، وخلال الحرب العالمية الثانية وبهدف ضبط اقتصاد الدول المتحاربة والإحاطة باتجاهات تطورها، تطورت نظرية العينات تطوراً سريعاً نظرياً وعملياً، واستمر ذلك حتى الآن .حيث أصبحت هذه النظرية تستخدم على نطاق واسع لد ا رسة مختلف الجوانب[16] وبعد تحديد المشكلة وما يرتبط بها من فروض و أسئلة ، تأتي خطوة جمع البيانات ،وذلك لكي يتأكد من فروضه التي وضعها، وأول خطوة من خطوات جمع البيانات هو اختيار المجتمع الذي ستطبق الدراسة عليه ، والذي ستنعكس النتائج عليه ولكن ليس في الإمكان في اغلب الأحيان في مجال البحث العلمي بشكل عام ، والبحث في مجال العلوم الإنسانية بشكل خاص ، تناول المجتمع بكل فئاته بشكل عام، وخاصة إذا كان مجتمع البحث كبيراً جداً و منتشراً في بقعة جغرافية ممتدة في مساحات شاسعة، لهذا يلجأ الباحث الي دراسة الظاهرة عند عدد محدود من الأفراد الذين يمثلون تمثيلاً كاملاً مجتمع البحث الأصلي و تعرف هذه الطريقة ((بالعينة)) ثم يستدل من نتائجها على الميزات الأساسية للمجتمع الكبير الذي اشتقت منه، "ويشغل موضوع العينات حيزاً هاماً في البحث التربوي و القياس نظراً لان البحوث التربوية تعتمد في اغلب الحالات، إن لم يكن في جميع الحالات ، على اختيار عينة بطريقة ما. وتتبع أهمية موضوع العينات من انه يدخل مباشرة في نطاق الاستدلال الإحصائي ويقوم على استخلاص الخواص الإحصائية للأصل من الخواص الإحصائية لإحدى أو بعض عيناته ، أي انه يستنتج صفات الكل من الجزء أو الاجزاء (التي تنطوي تحت إطاره[17] لذا اصبح اختيار العينة من

[16] القرشي 1999 ص 87
[17] البهي السيد 1978 ص 304

المجالات المهمة في اسلوب اختيارها وما هو الحجم الذي يمثل المجتمع الذي سحبت منه ؟وهل بالإمكان تعميم نتائجيا على المجتمع ؟وهنا برزت اهمية البحث والتي تستعرض أهم المعايير لتحديد واختيار حجم العينة المناسبة من مجتمع الدراسة، وتقدم لنا من الاستراتيجيات والطرق العلمية لتحديد حجم العينة بشكل دقيق بحيث تمثل المجتمع بشكل صحيح، وتخدم أغراض الدراسة والبحث بما يجعلنا نثق في النتائج البحثية كما نثق في واقعية تطبيقيا وتعميمها على المجتمع، والتي من خلالها يمكن التعرف على الاسس المعتمدة في اشتقاق العينة من مجتمع البحث

إن استعمال العينات لدراسة ظاهرة ما دراسة علمية أصبح شائعة في مجال البحث العلم أي أن الباحث يجد نفسه لا يستطيع القيام بدراسة شاملة لجميع مقررات البحث فلا يجد غير وسيلة بديلة يستطيع الاعتماد عليها وهي الاكتفاء بعدد قليل من هذه المقررات،ولكن حتى يكون ذلك ممكنا ودقيقا في تمثيل المجتمع،يجب أن يكون التصميم العيني وتطويره منسجما مع المبادئ والمنهجية المقترحة في الصفحات الموالية التي نهدف من ورائها توفير نص أساسي حول تقنيات المعاينة في العلوم الاجتماعية.

15.2 **مفهوم العينة**[18] :

العينة هي عبارة عن جزء من مجتمع الدراسة وعملية المعاينة هي عن مجموعة خطوات او اجراءات لاختبار هذا الجزء من اجل الحصول علي استنتاجات تتعلق بمجتمع الدراسة .
واسلوب المعاينة من الادوات التي يلجأ اليها معظم الناس للحصول علي فكرة مبدئية او انطباع اولي عن بعض امور الحياة اليومية .

---

[18] العتوم – شفيق – طرق الاحصاء الجامعة الاردنية – 2005 – ص 27

16ـ2 **الاطار** :

هو عبارة عن طريقة للوصل الي كل مفردة من مفردات المجتمع وقد يكون الاطار قائمة بجميع عناصر المجتمع التي يتم الاختيار منها ويسمي هذا الاطار في بعض الاحيان مجتمع الدراسة لانه يزود الباحث بجميع عناصر المجتمع التي يتم الاختيار منها ويجب ان يكون الاطار شاملاً لجميع وحدات المعاينة وفي حالة عدم احتوائة علي بعض العناصر فانه يكون اطار ناقصاً غير ممثل لمجتمع الدراسة ويكون الباحث في مواجهة خطأ المعاينة .

17ـ2 **خطوات اختيار العينة**[919] :

إذا كان الباحث بصدد اختيار العينة ، فإن عليه أن يعى تماماً أن هناك شرطاً رئيسياً يحكم قدرته على تعميم نتائجه على المجتمع الأصلى ، إنه التمثيل ، ويتطلب هذا توفر الشروط التالية :

أ. توافر كل صفات وخصائص المجتمع الأصلى فى العينة ، بحيث تكون نموذجاً مصغراً لهذا المجتمع ، وآنذاك نستطيع أن نقول إن ما يصدق على هذا النموذج يصدق على المجتمع الأصلى الذى اشتق منه.

ب. التناسب بين عدد أفراد العينة ، وعدد الأفراد الذين يشكلون المجتمع الأصلى ، فلا يكون المجتمع الأصلى طلاب المرحلة الثانوية مثلاً ، ويتخذ الباحث عينة عبارة عن فصل دراسى من إحدى المدارس الثانوية مكون من عشرين طالباً.

ت. منح جميع أفراد المجتمع الأصلى فرصة متكافئة لأن يتم اختيارهم للانضمام للعينة ، بمعنى آخر موضعية الاختيار وعدم التحيز لفرد معين أو فئة معينة دون غيرها .

---

[9] العتوم – شفيق – طرق الاحصاء الجامعة الاردنية – 2005 – ص 27

18ـ2 **انواع العينات**:

تنقسم العينات بشكل عام الي عينات احتمالية وعينات غير احتمالية:

19ـ2 **العينات الاحتمالية**:

في المعاينة الاحتمالية يرتبط اختيار اية وحدة من وحدات المجتمع باحتمال معين ويمكن فيها قياس خطأ المعاينة العشوائي ومن هطه العينات

20ـ2 **العينة العشوائية البسيطة** *Simple Random Sample* :

يتم اختيار عينة عشوائية بسيطة إذا كان المجتمع الإحصائي صغيراً ومتجانساً ونقصد بالتجانس أنه لا توجد اختلافات كبيرة بين مفرداته أو أنه لا يتكون من مجموعات (أو طبقات أو فئات أو أقسام) مختلفة في هذه الحالات يقال أن المجتمع متجانس. فمثلاً طلاب قسم الاحصاء يمثلون مجتمعاً متجانساً لأنهم يدرسون المقررات نفسها ويحصلون على المؤهل نفسه " بينما . مثلاً . طلاب كلية الدراسات التجارية بجامعة السودان للعلوم والتكنولوجيا لا يمثلون مجتمعاً متجانساً لأن بالكلية سبعة أقسام مختلفة، وكل قسم هنا يمثل طبقة أو مجموعة مختلفة. وبالتالي فإن مجتمع طلاب الكلية . أو الجامعة . مجتمع غير متجانس (أي مجتمع طبقي).

فإذا كان المجتمع متجانساً ويتكون من عدد صغير (أو محدود نسبياً) من المفردات فإن العينة العشوائية البسيطة تكون هي العينة المناسبة في هذه الحالة. ويتم اختيارها إما بالكيس المثالي أو باستخدام جداول الأرقام العشوائية أو بالحاسب الآلي حيث تتاح الفرصة نفسها (أو الاحتمال نفسه) لكل مفردات المجتمع للاختيار في العينة (أي يطبق هنا مبدأ العشوائية البسيطة على كل مفردات المجتمع). وهي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده فيها ، حيث يتم الاختيار باستخدام

أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار. والعينات العشوائية إذا ما تم اختيارها بالطريقة العلمية السليمة والمناسبة يمكن أن تكفل درجه عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة .. من أهم أنواع العينات العشوائية مايلي.

21.2 **العينة العشوائية الطبقية**:[20]

اذا كان مجتمع الدراسة غير متجانس فانه يتم تقسيم المجتمع الي مجموعات جزئية تسمى طبقات ومفردها طبقة بحيث تكون هذه الطبقات متجانسة في داخلها ومختلفة فيما بينها.

22.2 **العينة العشوائية المنتظمة**:

هي عبارة عن طريقة يتم فيها اختيار الوحدة الاولى عشوائياً واختيار هذه الوحدة يحدد اختيار بقية الوحدات حسب فترة المعاينة او مايسمى ايضاً فترة الانتظام وتعتبر هذه الطريقة من اسهل اساليب العينات الاحتمالية.

23.2 **العينة العشوائية متعددة المراحل**:

تشير هذه الطريقة في المعاينة الي اكثر من مرحلة في عملية الاختيار واذا تم هذا الاختيار على مرحلتين فان العينة تكون ثنائية المراحل ومن انوع العينة العشوائية متعددة المراحل , المكانية فاذا اراد باحث اجراء دراسة عن ميزانية الاسر في مدينة ما من اجل تقدير نسبة ما تنفقه الاسر على السلع والبنود الضرورية من مجموع الانفاق الاستهلاكي فانه يتسطيع تنفيذ عملية اختيار الاسر باختيار عينة من المناطق التي تقسم اليها المدينة في المرحلة الاولى وعينة من احياء المناطق المختارة في المرحلة الثانية ثم عينة من شوارع

[20] العتوم – شفيق – طرق الاحصاء الجامعة الاردنية – 2005 – ص 30

المناطق المختارة في الحمرحلة الثالثة واخيراً عينة من مساكن هذخ الشوارع المختارة في المرحلة الرابعة.

24ـ2 **العينة العشوائية العنقودية:**[21]

تعتبر العينة العشوائية العنقودية معاينة ذات مرحلتين والهدف منها هو تقليل التكاليف مع الاحتفاظ بخصائص ومميزات المعاينة الاحتمالية

25ـ2 **التوزيع الطبيعي**[22]:

يعتبر التوزيع الطبيعي هو الاداة الاحصائية التي يمكن عن طريق صفاته تحليل بيانات المتغيرات المتصلة والمتغير المتصل هو المتغير الذي يمكن ان يأخذ جميع القيم بما فيها القيمة ذات الكسور داخل المسافة التي يتحرك فيها وهذه المتغيرات ترتبط اكثر ماترتبط بالظواهر الطبيعية كالاعمار والاطوال والاوزان ....الخ , وصفات وخواص التوزيع الطبيعي هي اساس النظرية الاحصائية وتطبيقاتها والمشروعات الصناعية والمنشأت التجارات ومنشات الخدمات.

يعد التوزيع الطبيعي من اهم التوزيعات النظرية لان كثير من الظواهر الطبيعية تتبع هذا التوزيع كما ان متوسط العينة لاي مجتمع من المجتمعات يتبع التوزيع الطبيعي مهما كان المجتمع المأخوذ منه العينة ومهما كان نوع التوزيع وسوا كان توزيع ذو الحدين او اي توزيع اخر وبعنى اخر فان التوزيع الخاص بمتوسط العينات الممكن الحصول عليها من اي مجتمع من المجتمعات بطريقة تكافؤ الفرص يقترب من التوزيع الطبيعي.

التوزيع الطبيعي هو اكثر التوزيعات الاحتمالية اهمية واستخداماً في مختلف المجالات الاقتصادية والصحية والادارية والتربوية وغيرها وقد اكتشف هذا التوزيع الي العالم الرياضي

---

[21] العتوم – شفيق – طرق الاحصاء الجامعة الاردنية – 2005 – ص 34

[22] محمد صلاح الدين – اساليب كمية 2 جامعة القاهرة – كلية التجارة

الانكليزي (دي مويفر في العام 1733) وكان اول من استخدم التوزيع الطبيعي في دراسة الاخطاء المحتملة في القياس كل من العالمين الرياضيين لابلاس و كاوس في العام 1809 حيث لاحظ هؤلاء العلماء درجة عالية من الانتظام في اخطأ القياس واطلق عليه حينه اسم التوزيع الطبيعي للاخطأ *normal curve of error* ويمكن عرض هذا التوزيع بشكله العام والقياسي ويعد التوزيع الطبيعي من اهم التوزيعات الاحتمالية المتصلة من الناحيتين النظرية والتطبيقية اذ انه يستخدم علي نطاق واسع في وصف عدد كبير من الظواهر الطبيعية منها علي سبيل المثال لا الحصر وصف متغيرات الاوزان والاطوال وضبط الجودة ....الخ .

وبافتراض ان لدينا متغير عشوائي متصل وليكن X يتوزع وفق التوزيع الطبيعي بوسط حسابي μ وتباين σ^2 فان دالة التوزيع الاحتمالي له تاخذ الشكل الاتي :

$$F(x) = \{\frac{1}{\sqrt{2\pi\sigma^2}}\ e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

حيث ان : μ و σ^2 : تمثل معلمات التوزيع الطبيعي وان $[\sigma^2 > 0, -\infty < \mu < \infty]$

π : تمثل النسبة التقريبية الثابتة اذ ان $[\pi = 3.14159]$

وغالباً مايعبر عن التوزيع الطبيعي اختصاراً بالاصطلاح الاتي : $X \sim N\ (\mu, \sigma^2)$

هذا يعني ان المتغير العشوائي المتصل (X) يتوزع وفق التوزيع الطبيعي بالمعلمتين μ , σ^2 .

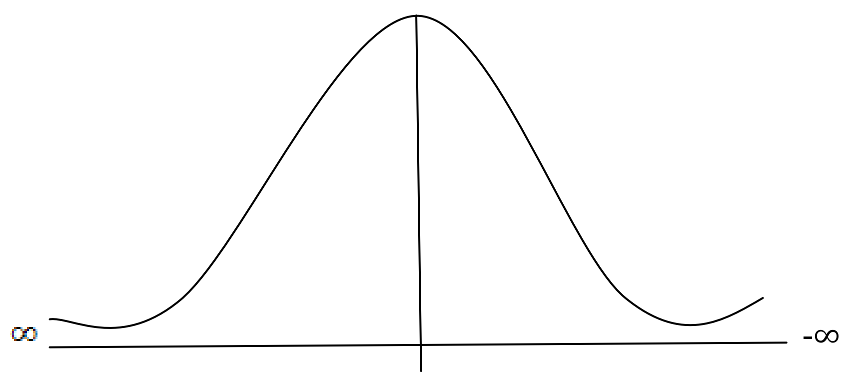
ان الوسط الحسابي (μ_x) والتباين (σ_x^2) والانحراف المعياري (σ_x) للتوزيع الطبيعي تكتب علي النحو التالي :

$$\mu_x = \mu$$
$$\sigma_x^2 = \sigma^2$$
$$\sigma_x = \sqrt{\sigma_x^2}$$

ويتصف التوزيع الطبيعي بالخصائص الاتية :

ان منحني دالة الكثافة الاحتمالية *[f(x)]* للتوزيع الطبيعي يشبة شكل الناقوس (الجرس) ويكون متماثلاً حول المحور الراسي المار بنقطة الاصل $(x = \mu)$ والشكل التالي يوضح ذلك :

الشكل (2–7) يوضح التوزيع الطبيعي



يتقارب طرفا منحني دالة الكثافة الاحتمالية *[f(x)]* من الصفر اي ان :

$$lim_{x\sim\infty} \quad f(x) = \lim_{x\sim-\infty} f(x) = zero$$

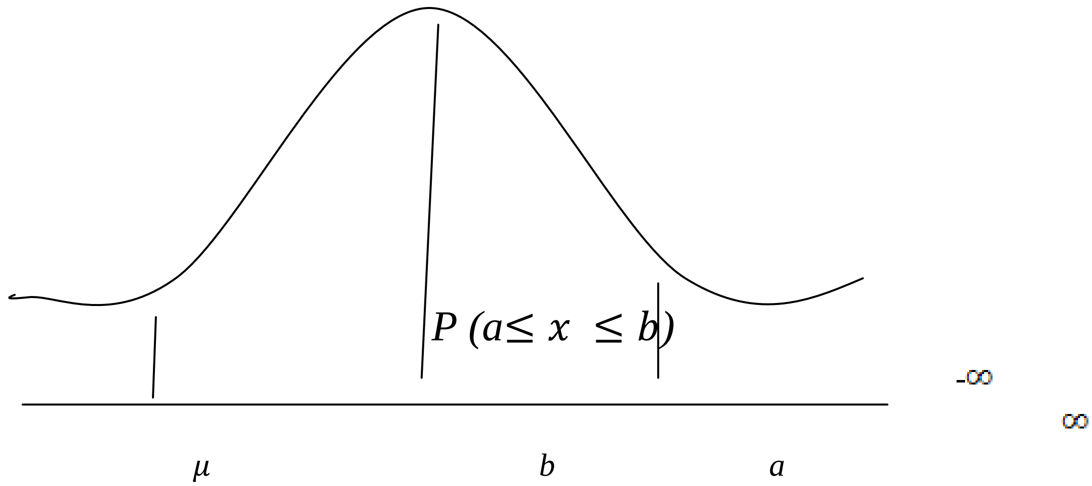
ان المساحة تحت منحني التوزيع الطبيعي تساوي الواحد صحيح اي ان :

$$P(-\infty < x < \infty) = \int_{-\infty}^{\infty} f(x)\ dx = 1$$

اذ يمكن حساب المساحة تحت منحني التوزيع الطبيعي بين النقطتين *[b , a]* مثلاً علي النحو الاتي :

$$P(a \le x \le b) = \int_a^b f(x)\,dx = p(x \le b) - p(x \le a)$$

والشكل (2–8) يوضح المسافة تحت المنحني بين النقطتين *[a , b]*



1. ان الوسط الحسابي والوسيط والمنوال للتوزيع الطبيعي متساوية دائماً

2. يتحدد شكل منحني دالة التوزيع الطبيعي من خلال الوسط الحسابي μ والتباين σ^2

فاذا كان المطلوب هو حساب احتمال ان يكون المتغير الطبيعي العام اقل من قيمة محددة a او اكبر من قيمة b او بين القيمتين a,b فانه يمكن حساب هذه الاحتمالات باستخدام حساب التكامل وذلك علي النحو التالي :

$$Pr\,(x \le a) = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\ dx \ \ldots\ldots.(8)$$

$$Pr(x \ge b\,) = \int_{b}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\ dx \ \ldots\ldots(9)$$

$$Pr\,(\,a \le x \ \ge b\,) = \int_{a}^{b} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\ dx \ \ldots\ldots\ (10)$$

ولصعوبة هذه التكاملات وعدم توفر جداول لقيم μ $(\infty > \mu < \infty-)$ وجميع قيم σ $(\sigma > 0)$ يتم تحويل المتغير الطبيعي العام الي متغير طبيعي قياسي , يرمز له بالرمز z باستخدامات الصيغة التالية:

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

ويمكن بسهولة اثبات ان :

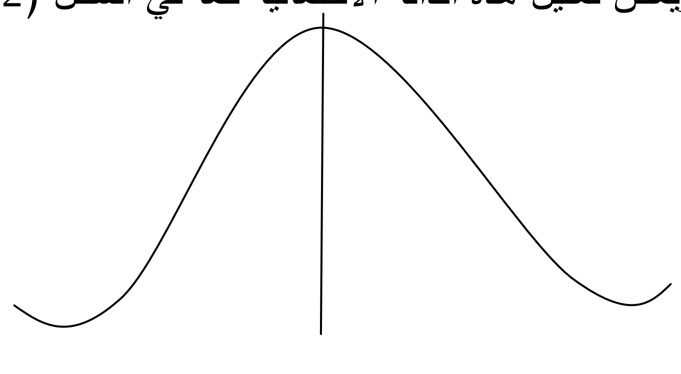
$$E(z) = 0$$

$$Var\ (z) = 1$$

ودالة كثافة احتمال المتغير الطبيعي القياسي تكون علي الشكل التالي:

$$F(z) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{z^2}{2}}\ , \quad -\infty < z < \infty$$

ويمكن تمثيل هذه الدالة الاحتمالية كما في الشكل (2−9):



الشكل (2−9) يوضح التوزيع الطبيعي القياسي

ومن الواضح ان دالة كثافة الاحتمال للتوزيع الطبيعي القياسي لا تعتمد علي اية معالم غير معلومة وبالتالي فان شكل التوزيع المبين في الشكل لا يتغير من حيث التوسط او التشتت ومن ناحية اخري فان الاحتمالات المطلوبة والتي تم التعبير عنها في المعادلات السابق ذكرها يمكن حسابها من دالة كثافة الاحتمال للمتغير الطبيعي القياسي وذلك باستخدام تحويلة المعادلة علي النحو التالي:

$$Pr(x \le a) = pr\left(z \le \frac{a-\mu}{\sigma}\right)$$

$$Pr(x \ge b) = pr\left(z \ge \frac{b-\mu}{\sigma}\right)$$

$$Pr(a \le x \le b) = pr\left(\frac{a-\mu}{\sigma} \le z \le \frac{b-\mu}{\sigma}\right)$$

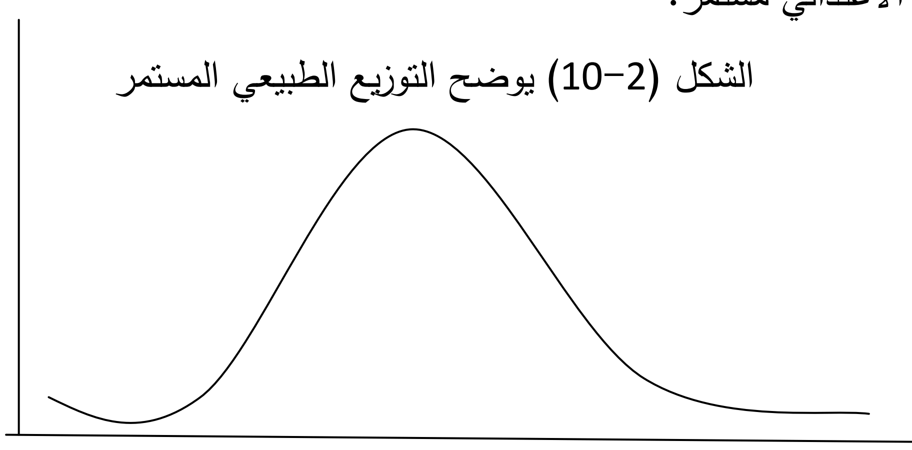
ويمكن ايجاد هذه الاحتمالات مباشرة من الجدول الخاص بهذا التوزيع .

26.2 اعتدالية التوزيع التكراري :

المقصود بالاعتدالية هي مدى تحرر التوزيع التكراري من الالتواء ، والالتواء قد يكون سالباً أو موجباً ، في حين أن التوزيع الاعتدالي لا التواء فيه ، ويمتد معامل الالتواء من (−3) إلى (+3) وكلما اقترب معامل الالتواء من الصفر كان التوزيع اعتدالياً ، ففي التوزيع الاعتدالي يكون المتوسط الحسابي = الوسيط = المنوال . والتوزيع الاعتدالي هو توزيع تكراري متماثل حيث يكون أحد نصفيه مطابقاً للنصف الآخر ويجب الانتباه إلي أن كل توزيع متماثل ليس بالضرورة أن يكون اعتدالياً حيث يطلق اسم التوزيع الاعتدالي فقط علي التوزيعات التي تتصف بالخصائص التالية:

1. المنحنى ناقوس الشكل ومتماثل حول المتوسط وذو منوال واحد.
2. تتساوي أسفل قيمته قيم المتوسط والمنوال والوسيط.
3. يمتد طرفي المنحني نظرياً إلي ما لا نهاية تدريجياً ولا يمس طرفي المنحني مع الإحداثي الأفقي.
4. تتحدد نسبة الجزء تحت المنحنى بين أي قيمتين تحديداً دقيقاً بمعرفة متوسط التوزيع $(\bar{X})$ وانحرافه المعياري (σ)، وذلك لأن المساحة المحصورة تحت المنحنى والمحور السيني تكون وحدة مربعة وهذا يمكننا من إيجاد احتمال تحقق أي حدث.
5. التوزيع الاعتدالي مستمر.

الشكل (2−10) يوضح التوزيع الطبيعي المستمر



28ـ2 **كثافة الاحتمال للتوزيع الطبيعي**:

ان التوزيع الطبيعي لايمكن تحديد دالة احتمالية مباشرة له وذلك راجع الي ان هذا التوزيع هو توزيع متصل بمعنى ان المتغير فيه يأخذ جميع القيم داخل مسافة معينة بما فيها القيم ذات الكسور بما يؤدى الي ان عدد القيم التي يأخذها متغير التوزيع الطبيعي كبيراً جداً بل هو لانهائياً ويتوجب على ذلك ان يكون الاحتمال بالنسبة للتوزيع الطبيعي منصرفاً الي وقوع المتغير داخل مسافة معينة ولايمكن باي حال ان يكون الاحتمال قاصراً على قيمة معينة.

29ـ2 **اختبار** t : **مقدمة** [23]

يعتبر اختبار t احد الاختبارات الاحصائية الشائعة الاستخدام في ابحاث ودراسات العلوم الانسانية وهو اختبار احصائي يستخدم للكشف عن دلالة الفروق بين متوسطي عينتين ولفهم الدور الذي يمكن ان يقو به هذا النوع من الاختبارت فانه يجب الاخذ في الاعتبار ان الاحصاء امر والرياضيات امر اخر قد يكون الاحصاء احد فروع الرياضيات ولكنهما شيئان مختلفان حتي وان كنا نستخدم في كليهما نفس العمليات الرياضية وللتمييز بينهما نذكر المثال التالي : نفرض أننا نريد المقارنة بين متوسطي أعمار الذكور وأعمار الإناث في مجتمع معين ، نجد أن كلاً من المتخصصين في الرياضيات وكذلك في الإحصاء سوف يتجهون لنفس الطريقة وهي حساب أعمار الذكور بجمعها ثم قسمتها على عددهم للحصول على متوسط أعمار الذكور (وليكن مثلاً : 25) وبالمثل يُحسب متوسط أعمار الإناث (وليكن : 23) ، إذاً أين الاختلاف ؟ الاختلاف يكمن في نظرة كل منهما للمتوسط ، فالمتخصصون في الرياضيات سيقولون أن متوسط أعمار الإناث أقل من متوسط أعمار الذكور بسنتين ، بمعنى أن الذكور

---

[23] العتوم –شفيق – طرق الاحصاء – الجامعة الاردنية

أكبر أعماراً من الإناث في هذا المجتمع ، في حين أن المتخصصين في الإحصاء يتكون نظرتهم وتفسيرهم أوسع وأبعد من ذلك ، فلابد من حساب دلالة الفرق بين المتوسطين هل هذا الفرق دال إحصائياً ؟ هل هو فرق جوهري أم أنه فرق ظاهري وبسيط ولايمكن الأخذ به ؟ ومن هذا المنطلق يأتي دور اختبار (ت) للتعرف علي الدلالة.

30ـ2 مجالات استخدامه:

من أهم المجالات التي يستخدم فيها هذا الاختبار الكشف عن الفروق بين تحصيل الذكور والإناث في مادة دراسية معينة وذلك عن طريق حساب دلالة فرق متوسط تحصيل الذكور بالنسبة لتحصيل الإناث ، أو من أجل المفاضلة بين طريقتين من طرق التدريس ، ومعرفة مدى مايحدث من تغير في سلوك المتعلمين نتيجة تعرضهم لمؤثر معين وبناً علية فان اختبار t يستخدم لقياس دلالات الفرق بين المتوسطات غير المرتبطة والمرتبطة لدي العينات المتساوية وغير المتساوية .

31ـ2 شروط اختبار(t)[24]

لابد للباحث قبل استخدامة لاختبار t ان يدرس خصائص متغيرات بحثه من النواحي التالية:

- حجم العينة : ان الاصل في هذا الاختبار أنه من مقاييس دلالة العينات الصغيرة ، ولكن هذا لايحول دون استخدامه لدى العينات الكبيرة، واستخدامه للعينات الصغيرة جداً (التي يقل عدد أفرادها عن 30 مفردة) أمر مشكوك فيه إذ يميل فيها التوزيع إلى أن يكون مدبباً ، أما العينات الكبيرة فهي التي يزيد عدد أفرادها عن 30 مفردة وفيها يميل التوزيع إلى أن يكون اعتدالياً طبيعياً .

[24] جولى بالانت (2006) : التحليل الإحصائى بإستخدام برامج*spss* ، دار الفاروق للنشر والتوزيع ، القاهرة ، ط2 .الترجمة د / خالد العامرى

-أن تكون العينتان (المجموعتان) مستقلتين وعشوائيتين .

-أن يكون مستوى قياس المتغير التابع كميا (فتريا أو نسبيا).

-أن يكون توزيع المتغير التابع معتدلا (طبيعيا). ويمكن التغاضي عن هذا الشرط إذا كان حجم العينة كبيراً.

-تساوي التباين وفي حالة عدم تساوي التباين يستخدم اختبار تي للفرق بين متوسطي عينتين مستقلتين مع عدم افتراض تساوي التباين

32ـ2 **تحليل التباين**: **مقدمة**:

ذكرنا عند الحديث عن اختيار الفرق بين وسطين حسابيين أن الباحث قد يرغب في اختبار ما إذا كان متوسط الدخل في إحدى الدول يساوي متوسط الدخل في دولة أخرى. أو إجراء اختبار عما إذا كان متوسط عمر الناخب في إحدى المناطق يساوي متوسط عمر الناخب في منطقة أخرى. أي أن الباحث قد يرغب في إجراء اختبار ما إذا كان متوسط مجتمع يساوي متوسط مجتمع آخر.

ويكون الفرض العدمي في هذه الحالات هو :

$$HO: \mu_1 = \mu_2$$

ولكن في كثير من الحالات قد يرغب الباحث في إجراء اختبار عن تساوي ثلاث متوسطات أو أكثر. فقد يرغب في اختبار ما إذا كانت متوسطات الدخول في أربع دول متساوية أم لا، أو ما إذا كانت متوسطات أعمار الناخبين في ست مناطق متساوية أم أن هناك اختلافات بينها.

ففي المثال الأول يكون الفرض العدمي هو :

$$HO: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

مقابل الفرض البديل بعدم تساوي بعض هذه المتوسطات (اثنان على الأقل غير متساويين)

وفي المثال الثاني يكون الفرض العدمي هو :

$$HO: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$$

مقابل الفرض البديل بعدم تساوي بعض هذه المتوسطات (اثنان على الأقل غير متساويين).
في مثل هذه الحالات لا نأخذ كل اثنين من المتوسطات على حدة ونجري اختبار الفرق بين وسطين لهما لأن هذا سوف يستغرق وقتاً أطول ومجهوداً أكبر، والأهم من ذلك أن احتمال أخذ قرار صحيح سوف يقل أو ينخفض، ويزيد بالتالي – كثيراً – احتمال الخطأ أو احتمال اتخاذ قرار غير سليم، ولتوضيح ذلك نأخذ المثال الأول الخاص بمقارنة المتوسط لأربع دول :فإذا أخذنا كل دولتين على حده فإنه يلزم إجراء هذا الاختبار ست مرات، فإذا كان مستوى المعنوية المستخدم في كل اختبار هو 0.05 (أو أن درجة الثقة هي 0.95) فإن احتمال الوصول إلى القرار الصحيح للاختبارات الستة معاً يساوي $(0.95)^6$ أي 0.95 مضروبة في نفسها ست مرات أي يساوي 0.745 ومعنى ذلك أن احتمال اتخاذ القرار الصحيح سوف ينخفض من 0.95 إلى 0.735 .فقط وبمعنى آخر فإن احتمال الخطأ في اتخاذ القرار الصحيح سوف يرتفع من مجرد 0.05 فقط إلى 0.265 والذي يساوي ($0.735 - 1$) وهو احتمال كبير للخطأ عند اتخاذ القرار.

وكلما زاد عدد المتوسطات كلما زاد احتمال الخطأ وقل احتمال اتخاذ قرار صحيح ففي المثال الثاني الخاص بمقارنة متوسطات أعمار الناخبين في ست مناطق فإنه يلزم إجراء الاختبار 15 مرة وبالتالي سوف ينخفض احتمال اتخاذ قرار صحيح في الخمسة عشر اختبار معاً من 0.95 إلى $(0.95)^{15}$ أي إلى 0.46 فقط وبالتالي يرتفع احتمال الخطأ في اتخاذ القرار الصحيح من مجرد 0.05 إلى 0.54 والذي يساوي ($1-0.46$) وهو احتمال كبير جداً للخطأ في اتخاذ القرار.

لذلك كان لابد من التفكير في أسلوب آخر بديل يوفر الوقت والمجهود وفي الوقت نفسه لا يقلل احتمال اتخاذ القرار الصحيح أو يكبر احتمال الخطأ في اتخاذ القرار، هذا الأسلوب هو

الذي يسمى " تحليل التباين " والذي يختبر ما إذا كانت المتوسطات كلها متساوية مرة واحدة دون أخذهم اثنين اثنين ودون أن ينخفض احتمال اتخاذ قرار صحيح أو يزيد احتمال الخطأ عند اتخاذه. وهو الذي يسمى اختصاراً *ANOVA* وهو اختصار للمصطلح الإنجليزي *Analysis of Variance*.

ويعتمد هذا الأسلوب من أساليب التحليل الإحصائي على ما يعرف باختبار F والذي يعتمد أساساً على تحليل التباين. ونعلم من الفصول الأولى من الكتاب أن التباين ما هو إلا متوسط مربعات انحرافات القيم عن وسطها الحسابي. أي أن التباين يعتمد أساساً على مجموع مربعات ثم القسمة على عدد المشاهدات. ويعتمد أسلوب تحليل التباين على تقسيم مجموع المربعات الكلي إلى أقسام فيمثل كل منهما أو يقيس أحد مصادر التغير أو الاختلاف *Source of Variation* يمثل أحدها – مثلاً – التغير بسبب المعاملات (أو المجتمعات) المختلفة ، ويمثل آخر التغير بسبب الأخطاء ثم تعرف الإحصائية (أو الاختبار) F بأنها خارج قسمة التباين بسبب المعاملات على التباين بسبب الأخطاء. وهكذا. أي أنه يتم حساب التباين بسبب المعاملات، والتباين بسبب الأخطاء فيحصل على قيمة F المحسوبة وبمقارنة هذه القيمة بالقيمة الجدولية F نصل إلى قرار إما بقبول الفرض العدمي أو عدم قبوله عند مستوى المعنوية المطلوب. ولتحليل التباين تطبيقات كثيرة في مختلف المجالات :

وسوف نتناول هنا أبسط حالة لتحليل التباين وهي التي تسمى التصنيف الأحادي *One - Way Classification* مع العلم بأن هناك حالات أخرى كثيرة لتحليل التباين منها على سبيل المثال التصنيف الأحادي في حالة اختلاف أحجام العينات وتحليل التباين الثنائي *Two-way Analysis* لكننا لن نتعرض في هذا الكتاب سوى لأبسط حالة وهي حالة التصنيف الأحادي بافتراض تساوي أحجام العينات

33.2 **اختبار مربع كاي** χ^2 (Chi-square Distribution)

مقدمة:

يعتبر اختبار مربع كاي من أشهر وأهم الأدوات الإحصائية المستخدمة في تحليل الظواهر الاجتماعية سواءً الوصفية منها أو غير الوصفية. لذا فغالباً ما لا تخلو الدراسات والأبحاث السياسية التي تنتهج الأسلوب الكمي أو السلوكي من تطبيق أو استخدام هذا الأسلوب في التحليل الإحصائي.

والفكرة الأساسية من استخدام هذا الأسلوب هي مقارنة التكرارات الفعلية أو المشاهدة بالمتوقعة. فمثلاً عند رمي قطعة نقود (وليكن مائة مرة) بافتراض أن هذه القطعة سليمة، فإنه من المتوقع أن نصف هذه الرميات صورة والنصف الآخر كتابة. إلا أنه من الناحية العملية من النادر أن نحصل على النصف (تماماً) صورة والنصف الآخر كتابة. وبالتالي كلما قل الفرق بينهما فإن ذلك يعني أن هناك تجانساً بين الصورة والكتابة. والعكس صحيح. فكلما كبر الفرق بينهما كلما كان ذلك مؤشراً على أن وجهي القطعة غير متجانسين، وللتأكد من ذلك يستخدم اختبار يسمى " باختبار التجانس ".

فعندما تتوفر بيانات عن الظاهرة محل الدراسة في شكل تكرارات (تسمى التكرارات المشاهدة *Observed Frequencies*) فإن مقارنة هذه التكرارات بما هو متوقع يمكننا من التوصل إلى بعض خصائص المجتمع محل الدراسة. كذلك فعند دراسة العلاقة بين مستوى التعليم في المجتمع ودرجة الوعي السياسي مثلاً، فإنه بمقارنة التكرارات الفعلية بالتكرارات المتوقعة يمكننا من التوصل إلى نتيجة فيما إذا كانت هاتان الظاهرتان مستقلتين أم أن هناك علاقة أو ارتباط فيما بينهما.

وهكذا فقد تكون البيانات مصنفة تبعاً لمتغير واحد أو لمتغيرين أو ظاهرتين ففي مثل هذه الحالات – التي تعتمد على تحليل التكرارات المشاهدة ومقارنتها بالتكرارات المتوقعة فإن التوزيع

الذي يسمى أو " مربع كاي " *(Chi-square dist. – dist)*. هو أداة التحليل الإحصائي المناسبة لمثل هذه الحالات.

ومن مزايا هذا الأسلوب في التحليل – أنه اختبار غير معلمي *Non-Parametric* بمعنى أنه لا يعتمد على طبيعة التوزيعات التي تتبعها البيانات علاوة على وجود جداول إحصائية للتوزيع تعطي القيم المختلفة عند مختلف درجات الحرية. ولمزيد من الإيضاح نسوق المثال التالي :

افترض أننا نريد قياس درجة الوعي السياسي للناخبين، ثم قمنا بإجراء هذه الدراسة على مجموعة (عددها n) من الناخبين، وكانت التكرارات المشاهدة لدرجة الوعي السياسي (بافتراض أن الوعي السياسي مقسم إلى عدد k من الدرجات) هي : $O_1 , O_2 ,..., O_K$

والتكرارات المتوقعة هي : $e_1 , e_2 ,..., e_K$

فإنه كلما كانت الفروق بين التكرارات المشاهدة والتكرارات المتوقعة طفيفة كلما كنا نميل إلى قبول الفرض بأن التكرارات المشاهدة تتفق مع التكرارات المتوقعة. والعكس صحيح، أي كلما كانت الفروق بين التكرارات المشاهدة والمتوقعة كبيرة كان هذا دليلاً على ضعف التطابق وبالتالي رفض النموذج أو التوزيع المفترض. والمقياس الذي يحدد إلى أي مدى تتفق التكرارات المشاهدة مع التكرارات المتوقعة هو الذي يسمى كا2 أو $X2$ يأخذ الشكل التالي :

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i} \qquad .... (12)$$

بدرجات حرية تساوي $K - I$ حيث O_i ترمز للتكرار المشاهد في الخلية (أو الفئة) رقم i، ei ترمز للتكرار المتوقع في الخلية (أو الفئة) رقم i , k ترمز لعدد الخلايا (أو الفئات أو التقسيمات).

فلو تم على سبيل المثال تقسيم الوعي السياسي إلى أربعة أقسام هي (عالي جداً – عال – متوسط – منخفض) فإن K في هذه الحالة = 4.

ونلاحظ أن X^2 تؤول إلى الصفر إذا ما كان هناك تطابق كامل بين التكرارات المشاهدة والتكرارات المتوقعة. وأن قيمة $X2$ تكبر كلما كبرت الفروق بينهما، بمعنى أن زيادة قيمة χ^2 تعني ضعف التطابق بين التكرارات المشاهدة والمتوقعة وبالتالي يكون احتمال رفض الفرض العدمي كبيراً في هذه الحالات والعكس صحيح بطبيعة الحال. ويشترط ألا يقل أي تكرار متوقع عن 5 تكرارات في كل خلية حتى يكون استخدام هذا النوع من التحليل مناسباً. ومن التطبيقات المعروفة لتوزيع :

1 – اختبار التجانس (أو التماثل).

2 – اختبار جودة التوفيق.

3 – اختبار الاستقلال.

وفيما يلي نتناول بالتفصيل استخدام χ^2 لاختبار التجانس.

وفي الفصل التالي – إن شاء الله – نتناول بالتفصيل اختبار الاستقلال. مع ملاحظة أنه يمكن اعتبار اختبار التجانس حالة خاصة من اختبار جودة التوفيق.

34.2 **اختبار التجانس** :

يعتبر اختبار التجانس – أو التماثل – كما سبق وأن أشرنا أحد التطبيقات المهمة لتوزيع χ^2 . فلو أراد الباحث اختبار ما إذا كانت لمجموعة من البرامج الإذاعية الأفضلية نفسها أو (التجانس أو التساوي في الأفضلية) لدى المستمعين فإن المطلوب في مثل هذه الحالات هو اختبار فرض التجانس (أو التماثل) بين الفئات أو الخلايا أو التقسيمات المختلفة (البرامج الإذاعية في هذا المثال) وذلك مقابل الفرض البديل أنها غير متجانسة.

وتكون خطوات اختبار التجانس كما يلي : –

1 – الفرض العدمي : هو فرض التجانس (أو التماثل).

2 – الفرض البديل : هو عدم التجانس.

3 – الإحصائية : وتأخذ الإحصائية الشكل التالي :

$$\chi^2 = \sum_{I}^{K} \frac{(O-e)^2}{e}$$

والتي لها توزيع مربع كاي بدرجات حرية $K-I$ حيث K هي عدد الخلايا أو الأقسام أو الفئات أو البرامج الإذاعية...، O ترمز للتكرارات المشاهدة، e ترمز للتكرارات المتوقعة، مع ملاحظة أن مجموع التكرارات المشاهدة يساوي مجموع التكرارات المتوقعة. أي أن : $\sum O = \sum e$

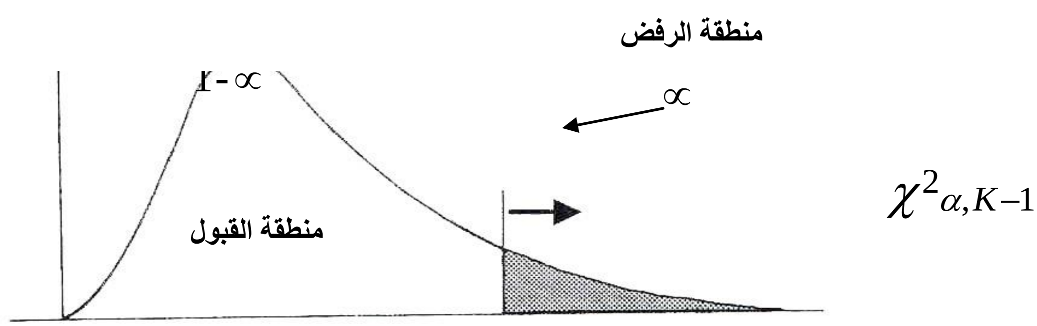
ويتم حساب التكرارات المتوقعة بقسمة مجموع التكرارات المشاهدة على عدد الخلايا K، فإذا كان عدد البرامج الإذاعية هي 4، وعدد المستمعين هم 100 مستمع، فإن التكرارات المتوقعة لكل البرامج هي $\frac{100}{4} = 25$ أي نعيد توزيع مجموع التكرارات المشاهدة بالتساوي (أو بالتماثل) بين الفئات أو الخلايا المختلفة. ثم نحسب الفروق بين التكرارات المشاهدة والمتوقعة، ثم نربع هذه الفروق ونقسم هذه المربعات على التكرارات المتوقعة فنحصل على قيمة الإحصائية.

ويمكن تنظيم الحل أو الحصول على قيمة الإحصائية كما في الجدول التالي :

الخلايا أو الأقسام	التكرارات المشاهدة O	التكرارات المتوقعة e	الفرق بين التكرارات المشاهدة والتكرارات المتوقعة $(o-e)$	مربعات الفروق $(o-e)2$	$\frac{(O-e)^2}{e}$ قسم المربعات على التكرارات المتوقعة
1	$O1$	$e1$	$(O_1 - e_1)$	$(O_1 - e_1)^2$	$(O_1 - e_1)^2 / e_1$
2	O_2	$e2$	$(O_2 - e_2)$	$(O_2 - e_2)^2$	$(O_2 - e_2)^2 / e_2$
K	$O\,k$	ek	$(O_k - e_k)$	$(O_k - e_k)^2$	
المجموع	$\sum O$	$\sum e$	*Zero*		$X^2 = \sum \frac{(O-e)^2}{e}$

ومجموع العمود الأخير يمثل قيمة الإحصائية χ^2.

4 – حدود منطقتي القبول والرفض : يستخدم هنا اختبار الطرف الأيمن لتوزيع كا 2 بدرجات حرية $K - 1$ عند مستوى المعنوية $\propto$ كما في الشكل التالي :



ونلاحظ ما يلي :

أ – أن توزيع χ^2 توزيع غير متماثل، بل هو توزيع ملتوٍ ناحية اليمين (موجب الالتواء).

ب – أنه غير معّرف في المنطقة السالبة، أو بمعنى آخر فإنه معّرف فقط في المنطقة الموجبة والتي تبدأ من الصفر.

ج – أن منطقة القبول تبدأ من الصفر وحتى القيمة $X^2_{\alpha,K-1}$ والتي نحصل عليها من جدول كا 2 عند مستوى معنوية يساوي $\propto$ ودرجات حرية تساوي $K-I$ (الجدول رقم 3).

د – أن منطقة الرفض تشمل القيم التي أكبر من $X^2_{\alpha,K-1}$ (اختبار الطرف الأيمن).

35.2 المقارنة والقرار :

حيث تتم مقارنة قيمة الإحصائية (المحسوبة من الخطوة الثالثة) بحدود منطقتي القبول والرفض، فإذا وقعت قيمة الإحصائية في منطقة القبول فإن القرار يكون قبول الفرض العدمي (أي قبول فرض التجانس)، والعكس إذا وقعت قيمة الإحصائية في منطقة الرفض يكون القرار هو رفض الفرض العدمي وقبول الفرض البديل. أي قبول فرض عدم التجانس وذلك بمستوى معنوية يساوي $\propto$.

2-36 توزيع بواسون *Poisson Distribution*

فى الحياة العملية احياناّ ما نقابل بعض الظواهر التى ينطبق عليها شروط توزيع ذى الحدين و لكن هذة الحوادث تكون نادرة الوقوع و هذا يعنى أن احتمال النجاح يكون صغير جداّ يقترب من الصفر و علية فأنة يمكن القول أن

$np=\lambda$ حيث λ هى مقدار ثابت و بذلك يكون احتمال الفشل كبير أى أنة يقترب من الواحد. و لكى نراقب بعض حالات النجاح فأننا سنجد أن n سوف تكون كبيرة جداّ فمثلاّ لو اردنا حساب احتمال خروج القطار من على الشريط المحفصى لة " القضبان" فأننا سنقوم بمراقبة القطارات او عدد كبير جداّ منها و نحسب عدد مرات خروج القطار من على الشريط أى حالات النجاح (التى حققت فيها الحادثة) حتى نستطيع أن نحسب الاحتمال.

وبذلك تكون شروط هذا التوزيع كالآتى:-

1- أن تكون احتمال النجاح ثابت و كذلك احتمال الفشل فى كل محاولة و يرمز لهما بالرمز q, p على المتوالى.

2- أن يكون احتمال النجاح صغيراّ و يقترب من الصفر و احتمال الفشل يقترب من الواحد الصحيح.

3- أن تكون عدد المحاولات كبيراّ جداّ حيث أن $=np=\lambda$ مقدار ثابت.

و يعتبر توزيع بولسون من التوزيعات الاحتمالية المتقطعة التى تقف متغيرات عشوائية منقطعة لها الشروط

السابقة. و سمى هذا التوزيع بهذا الأسم نسبة الى أحد مكتشفة و هو بواسون و يعتبر من اهم التوزيعات فى السائل المتعلقة بالمكالمات التليفونية و حركة المرور، بعض الظواهر

النادرة مثل الزلزال، و الحرائق، الحوادث على أحدى الطرق، عدد الاخطاء المطبعية فى صفحة ما من كتاب و غير ذلك. و دالة كثافة الاحتمال لتوزيع

بواسون هى :ـ

$$f(x) = \begin{cases} \dfrac{\lambda^x e^{-\lambda}}{x!} & , \quad x = 0,1,2,3,\ldots\ldots\ldots \\ 0 & , \end{cases} \qquad \ldots\ldots(13)$$

خلاف ذلك

حيث $e = 2.718$ تمثل مقدار ثابت.

$$\lambda = n\ p$$

و نأخذ X قيماً صحيحة موجة اعتباراً من الصفر الى مالانهاية.

اثبات انها دالة كثافة احتمال :ـ

يلاحظ أن المتسـلسـلة

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\ldots\ldots = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \qquad \ldots\ldots\ldots(14)$$

حيث $\lambda \rangle 0$

و حيث أن $\lambda \rangle 0$ فى $\therefore f(x)$

$$f(x) \geq 0$$

$$\therefore \sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda}\ e^{\lambda} = e = 1$$

يكثر استخدام هذا التوزيع في الحالات التي تقع فيها الأحداث وفقا لمعدلات زمنية، وكذلك في حالة الأحداث نادرة الوقوع، ومن أمثلة ذلك:

عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر. $X:\{x=0,1,2,\ldots\}$

عدد مرات ري نوع معين من المحاصيل الزراعية خلال الموسم. $X:\{x=0,1,2,\ldots\}$

عدد العملاء الذين يتم خدمتهم البنكية كل *10* دقائق. $X:\{x=0,1,2,\ldots\}$

عدد مرات زيارة المريض للطبيب كل سنة. $X:\{x=0,1,2,\ldots\}$

عدد مرات تناول الأسرة للحوم الحمراء خلال الأسبوع. $X:\{x=0,1,2,\ldots\}$

عدد أخطاء الطباعة لكل صفحة من صفحات الكتاب. $X:\{x=0,1,2,\ldots\}$

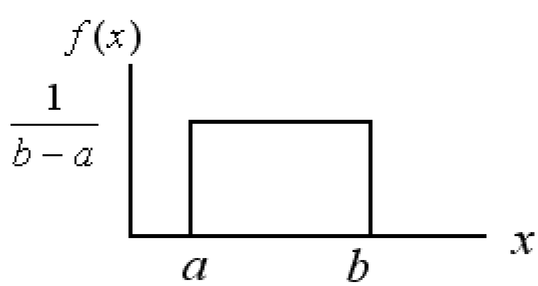
37-2 التوزيع المنتظم *Uniform distribution*

شكل دالة كثافة الاحتمال *p.d.f*

هو توزيع له دالة احتمال ثابتة، ويستخدم في حالة الظواهر التي يمكن أن تحدث بشكل منتظم، فإذا كان المتغير x متغير عشوائي له توزيع منتظم *Uniform*، مداه هو $a<x<b$ فإن دالة كثافة احتماله هي:

$$\boxed{f(x)=\frac{1}{b-a}\ ,\ a<x<b} \qquad (١٥-٨)$$

ويمكن تمثيل هذه الدالة بيانيا كما يلي:



معالم هذا التوزيع

توجد معلمتان لهذا التوزيع هما (a,b)، ولذا يكتب رمز لهذا التوزيع الصورة

X~U(a,b)

خصائص التوزيع المستطيل

الوسط الحسابي μ ، والتباين σ^2 لهذا المتغير هما :

$$\mu = E(x) = \frac{a+b}{2} \ , \ \ \sigma^2 = \frac{(b-a)^2}{12} \quad \ldots\ldots(14)$$

دالة التوزيع التجميعي $C.D.F$

تأخذ دالة **التوزيع** التجميعي $F(x)$ الشكل الآتي

$$\boxed{\begin{aligned} F(x) &= p(X \le x) = \int_a^x f(x)dx = \frac{1}{b-a}\int_a^x dx \\ &= \frac{x-a}{b-a} \end{aligned}} \qquad (١٦-٨)$$

2-38 **التوزيع الأسي** *Distribution exponentially*

عادة ما يستخدم التوزيع الأسي في مسائل متعلقة بقياس الزمن. من ذلك مدة خدمة شباك البريد، مدة مكالمة هاتفية، مدة تفريغ باخرة شحن، مدة تصليح آلة، مدة انتظار زبون قبل الحصول على الخدمة...في العلوم الدقيقة يستخدم التوزيع الأسي لتمثيل مدة حياة الذرات المشعة *(atoms radioactive)* قبل أن تتفكك، حيث يعبر الوسيط عن اللحظة التي يبقى فيها نصف المجتمع الأصلي25.

من الضروري فهم الآتي: كقاعدة عامة يستخدم التوزيع الأسي لتمثيل مدة حياة ظاهرة ما إذا كان لها متوسط ثابت $1/\lambda$ وكانت هذه الظاهرة لا تخضع للتقادم *(vieillissement)* أي أن مدة حياة الظاهرة بعد لحظة ما T لا تتبع اللحظة T؛ أي لا تتأثر بالمدة التي دامتها

[25] راجع موقع موسوعة Wikipédia .

الظاهرة من قبل. مثلا قد نستبعد استخدام التوزيع الأسي لتمثيل مدة حياة آلة عاملة قبل تعطلها لأن احتمال تعطلها في لحظة ليس مستقلا عن المدة التي عملتها الآلة من قبل، كذلك الأمر بالنسبة لمدة حياة الإنسان. عمليا، نتحقق من دقة تمثيل التوزيع الأسي , أو أي توزيع آخر, لظاهرة ما من خلال تقنيات اختبارات الفروض، وبالتحديد اختبار التجانس و التعديل.

نشير أخيرا إلى أن للتوزيع الأسي علاقة بالتوزيع بواسون، فإذا كان وقوع أحداث ما يتبع هذا التوزيع، فإن المدة بين وقوع حدثين تتبع التوزيع الأسي؛ كمثال على ذلك، إذا كان وصول الزبائن إلى مركز خدمة ما يتبع التوزيع بواسون فإن المدة الزمنية بين وصول زبون "أ" والزبون الموالي تتبع التوزيع الأسي. تتبين هذه العلاقة عند استنتاج صيغة القانون الأسي.

صيغة القانون الأسي أو دالة الكثافة و الدالة التجميعية للتوزيع.

بينت دراسة أن عدد حوادث العمل في معمل معين تتبع توزيع بواسون بمعدل λ حادث يوميا.

أوجد احتمال أن يسجل حادث على الأقل (حادث أو أكثر) في مدة t يوم.

$P(X \geq 1) = 1- P(0) = 1-[\lambda 0t * e- \lambda t/0!] \quad => \quad P(X \geq 1) = 1-e-\lambda t$

لنرمز ب T للزمن (باليوم) بين حادثين إذن سيكون لدينا $f(t)$ دالة الكثافة للزمن بين حادثين، و $F(t) = P(T \leq t)$ دالة التوزيع ل T.

لنحسب احتمال P أن يكون الزمن بين حادثين يوم أو أقل:

لدينا $P = P(T \leq t = 1)$ إذن:

$$P = F(t = 1)$$

لاحظ من ناحية أخرى أن P هو معادل لاحتمال أن يسجل على الأقل حادث في يوم معين:

$$P = P(X \geq 1) = 1\text{-}e^{-\lambda t}$$

من (1) و (2) نستنتج أن $\quad F(t) = 1 - e^{-\lambda t}$

و منه $f(t) = F(t)' = (1 - e^{-\lambda t})'$

إذن $f(t) = 1 - e^{-\lambda t}$

قاعدة: إذا كان حدث عشوائي ما يتكرر في الزمن وفق توزيع بواسون:

$$\boxed{p_{\tau}(x) = \frac{(\lambda \tau)^{x} e^{-\lambda \tau}}{x!}}$$

فإن الزمن T بين حادثين يتبع التوزيع التالي:

$$f(\tau) = \begin{cases} \lambda e^{-\lambda \tau} & , \quad \tau > 0 \\ 0 & , \quad \tau \leq 0 \end{cases}$$

حيث λ عدد حقيقي موجب.

و يسمى هذا التوزيع التوزيع الأسي ويسمى أيضا التوزيع الأسي السالب لعلاقته بتوزيع بواسون.