الباب الثامن: التقدير واختبارات الفروض

Chapter 8: Estimation & Hypotheses Testing

مقدمة

نظرا للصعوبات التي تواجه الباحثين في الحصول علي بيانات المجتمع ككل واللجوء إلى أسلوب العينة في جمع البيانات ، أصبحت أساليب الاستدلال الإحصائي هي الوسيلة لاتخاذ القرارات الإحصائية بل وأصبح الإحصاء الوصفي- الفرع الأول من فروع علم الإحصاء- هو مرحلة من مراحل البحث الإحصائي التي يتم على أساسها تحديد أسلوب الاستدلال الإحصائي المناسب الذي يمثل الهدف الأساسي من دراسة الإحصاء.

ومن التعريفات الهامة التي يقوم عليها الاستدلال الإحصائي تعريفي المعلمة والإحصاء.

- **المعلمة** (Parameter)هي عبارة عن خاصية أو مقياس يتم حسابها من المجتمع محل الدراسة.

مثل : - نسبة البطالة في السعودية

- متوسط العمر الافتراضي لجهاز معين.

أي ان **المعالم** هي مقاييس تحدد خصائص المجتمع (التوزيع)

- **الاحصاءة** (Statistic) هي عبارة عن خاصية أو مقياس يتم حسابها من العينة المسحوبة من المجتمع محل الدراسة .

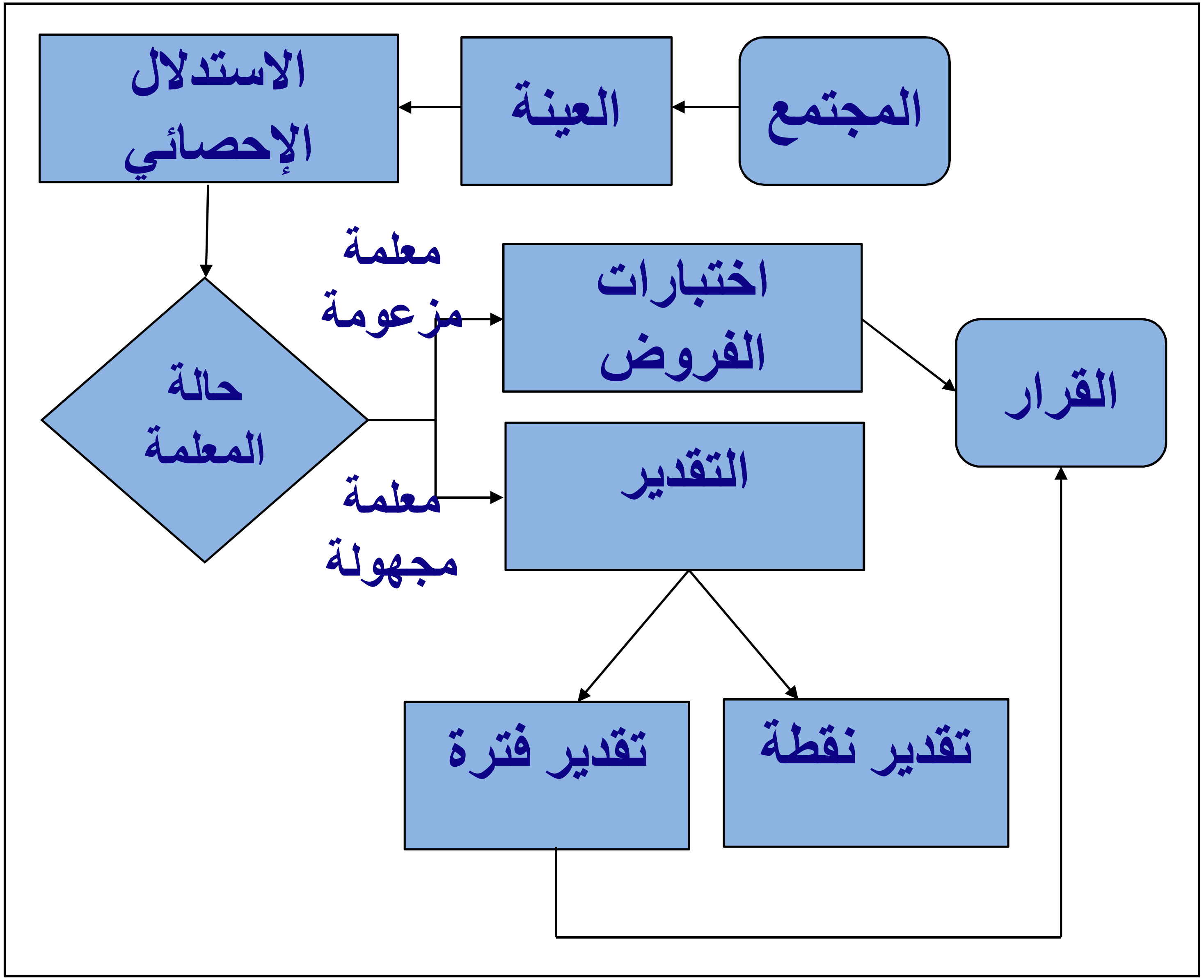
أي أن **الاحصاءة** هي دالة في بيانات العينه .

و ينقسم الإحصاء الاستدلالي أو (الاستدلال الإحصائي) إلى موضوعين رئيسيين:

(1)تقدير معالم المجتمع (Estimation).

(2)اختبارات فروض بشأن صحة قيم معالم المجتمع (Testing of Hypotheses).

ويتم ذلك عن طريق سحب عينة أو عينات من المجتمع أو المجتمعات المراد تقدير معالمها أو إجراء اختبارات فروض بشأنها،
كما يتضح من نموذج الاستدلال الإحصائي التالي:


المجتمع
العينة
الاستدلال الإحصائي
حالة المعلمة
معلمة مزعومة
معلمة مجهولة
اختبارات الفروض
التقدير
القرار
تقدير نقطة
تقدير فترة

يستخدم أسلوب التقدير لتقدير معالم المجتمع إذا كان الهدف هو تحديد قيمة معلمة مجهولة (**Unknown Parameters**)

.

أما اختبارات الفروض فتستخدم بهدف الوصول إلى قرار بشأن رفض أو عدم رفض فرض إحصائي عن معلمة مزعومة (**Hypothesized Parameter**).

تقدير معالم المجتمع

التقدير (Estimation)هو أسلوب إحصائي مبني على نظريات إحصائية، يستخدم لتقدير معلمة ما محل الاهتمام عن طريق استخدام مقاييس العينة.

هناك أسلوبان لتقدير معلمة المجتمع المجهولة وهما:

التقدير بنقطة:
تستخدم بيانات العينة لتقدير معلمة المجتمع المجهولة بنقطة واحدة فقط، أي بقيمة واحدة فقط.

التقدير بفترة:
تستخدم بيانات العينة لتقدير معلمة المجتمع المجهولة بفترة (مجموعة من القيم).

من أهم معالم المجتمع متوسط المجتمع μ
وسوف نرى كيف نقوم بتقديره بنقطة وبفترة

– نستخدم المتوسط (الوسط الحسابي) $\bar{x}$ للعينة كتقدير نقطة (**Point Estimation**) لمتوسط المجتمع μ .

–ويمتاز الوسط الحسابي للعينة بأربع خصائص أساسية من المفترض توفرها في أي تقدير وأهم هذه الخصائص، عدم التحيز.

–ومن المتوقع أن لا تكون قيمة متوسط العينة مساوية لقيمة متوسط المجتمع، ولكن قريب منها. ويسمى الفرق المطلق بين قيمة الإحصاءة وقيمة المعلمة الفعلية المراد تقديرها باسم الخطأ المعياري.

$$\left|\hat{\theta} - \theta\right| \Rightarrow estimated\ error$$

كما ان لكل إحصاءة من الإحصاءات التي نحسبها من العينة خطأ معياري خاص بها :

١:فهناك الخطأ المعياري لمتوسط العينة

٢:الخطأ المعياري للانحراف المعياري للعينة

٣:الخطأ المعياري لمعامل الارتباط بين بيانات العينة وغيرها.

وهذا يدخل تحت موضوع الاستدلال الإحصائي.

وسوف نركز فقط على تعريف الخطأ المعياري لمتوسط العينة.

الخطأ المعياري (Standard Error) للمتوسط هو الانحراف المعياري لتوزيع مجتمع متوسطات العينات.
بمعنى انحراف متوسطات العينات عن متوسط مجتمعها.
ويرمز له بالرمز $\sigma_{\bar{X}}$ ، حيث:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

غالباً ما تكون قيمة الانحراف المعياري للمجتمع (σ) غير معلومة ، لذا نقدرها بقيمة الانحراف المعياري للعينة(S) .
فيصبح لدينا تقدير للخطأ المعياري للمتوسط باستخدام الانحراف المعياري للعينة بالصيغة التالية:

$$\hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}}$$

مثال :

أوجد الخطأ المعياري للوسط الحسابي لعينة حجمها (49) تم سحبها من مجتمع له انحراف معياري يساوي (14).

- الحل: الخطأ المعياري لمتوسط العينة يساوي

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{14}{\sqrt{49}} = 2$$

مثال :

تم قياس مستوى الأداء على مهارة معينة لأفراد عينة حجمها (36)، فكان الوسط الحسابي والانحراف المعياري للعينة 30 و9 على الترتيب. أوجد الخطأ المعياري لمتوسط العينة.

الحل:

لأن قيمة الانحراف المعياري للمجتمع (σ) مجهولة، لذا نقدرها بقيمة الانحراف المعياري للعينة(S) . فيصبح لدينا تقدير للخطأ المعياري للوسط الحسابي للعينة يساوي:

$$\hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}} = \frac{9}{\sqrt{36}} = \mathbf{1.5}$$

تقدير الفترة لمتوسط المجتمع

وجدنا فيما سبق أن تقدير النقطة نادراً ما يساوي المعلمة التي نرغب في تقديرها، لذلك فإننا نحدد **فترة** تحتوي على مجموعة من القيم تتضمن فيما بينها قيمة معلمة المجتمع، وتسمى هذه الفترة بتقدير **(فترة الثقة)**، واحتمال وقوع المعلمة في هذه الفترة يسمى **درجة الثقة**.

ونرمز لدرجة الثقة بالرمز $(1 - \alpha)$، ومكمل هذه القيمة يسمى **مستوى المعنوية** ويرمز له بالرمز α	فمثلاً إذا كانت درجة الثقة **95%** ومكمل هذه القيمة يكون **مستوى المعنوية** يساوي **0.05**

ولتقدير فترة ثقة لمتوسط المجتمع μ نستخدم متوسط العينة $\bar{x}$ وفي حالة حجم العينة الكبير $(n \geq 30)$ نستخدم العلاقة الآتية:

$$\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

حيث تتحدد قيمة $Z_{\alpha/2}$ حسب درجة الثقة (أو مستوى المعنوية) كما يلي:

إذا كانت درجة الثقة **90%** فإن $Z_{\alpha/2} = 1.65$

وإذا كانت درجة الثقة **95%** فإن $Z_{\alpha/2} = 1.96$

وإذا كانت درجة الثقة **99%** فإن $Z_{\alpha/2} = 2.58$

عندما نكوّن فترة ثقة بنسبة $100\ (1 - \alpha)\ \%$ حول μ فإننا عادة ما نعبر عن هذه الفترة بعبارة مثل: «لدينا ثقة بنسبة $100\ (1 - \alpha)\ \%$ أن المتوسط μ للمجتمع يقع في هذه الفترة»، حيث يتم تحديد هذه الفترة في التطبيقات العملية، والعبارة تعكس ثقتنا في عملية التقدير أكثر من القيمة المحسوبة من بيانات العينة.

درجة الثقة	قيمة α (مستوى المعنوية) المناظر
0.99	0.01
0.95	0.05
0.90	0.1

ملاحظات هامة:

- إذا كانت قيمة المعلمة σ مجهولةٍ نستخدم قيمة الانحراف المعياري للعينة S بدلاً عنها كتقدير لها.

- قيم $Z_{\alpha/2}$ ستستخدم في كثير من العلاقات تحت موضوع التقدير واختبارات الفروض.

مثال:

مصنع لإنتاج المصابيح الكهربائية، اختير من إنتاجه عينة حجمها (100) مصباح لتقييم جودة الإنتاج، حيث حدد مدير المصنع معيار الجودة أن يتراوح عمر المصباح بين (1000) و(1100). إذا كان الوسط الحسابي لعمر المصباح في العينة المختارة (1200) ساعة وانحرافه المعياري (250) ساعة. قدر بدرجة ثقة 95% متوسط عمر المصباح من إنتاج المصنع كله.

الحل:

حيث أن الانحراف المعياري للمجتمع σ مجهول نقدره باستخدام الانحراف المعياري للعينة.

$$n = 100, \quad \bar{x} = 1200, \quad S = 250, \quad 1-\alpha = 0.95 \Rightarrow Z_{\alpha/2} = 1.96$$

$$P\left[\bar{x} - Z_{\alpha/2}\frac{S}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\alpha/2}\frac{S}{\sqrt{n}}\right] = 1-\alpha$$

$$P\left[1200 - (1.96)\frac{250}{\sqrt{100}} \le \mu \le 1200 + (1.96)\frac{250}{\sqrt{100}}\right] = 0.95$$

$$P\left[1200 - 49 \le \mu \le 1200 + 49\right] = 0.95$$

$$P\left[1151 \le \mu \le 1249\right] = 0.95$$

متوسط عمر المصباح من إنتاج المصنع كله يتراوح بين 1151 و1249 ساعة بدرجة ثقة 95%. وتدل هذا التقدير أن جودة المصابيح عالية من هذا المصنع.

مثال :

تسبب المقاعد الخالية لشركات الطيران في خسارة لمصدر الدخل، بفرض أن إحدى شركات الطيران الكبرى أرادت «تقدير» عدد المقاعد الخالية لكل رحلة خلال العالم الماضي. ولهذا الغرض، تم اختيار عشوائي لعدد 225رحلة طيران وتسجيل عدد المقاعد الخالية في كل رحلة من هذه العينة. وكان الوسط الحسابي والانحراف المعياري لعدد المقاعد الخالية في هذه العينة هما: $\bar{x} = 11.6$ مقعد، $S = 4.1$ مقعد. قدر μ (عدد المقاعد الخالية للرحلة خلال العام الماضي) باستخدام فترة ثقة 90%.

$\bar{x} = 11.6 \quad , s = 4.1$

$$n = 225, \quad 1-\alpha = 0.90 \Rightarrow Z_{\alpha/2} = 1.65$$

$$P\left[\bar{x} - Z_{\alpha/2}\frac{S}{\sqrt{n}} \le \mu \le \bar{x} + Z_{\alpha/2}\frac{S}{\sqrt{n}}\right] = 1-\alpha$$

$$P\left[11.6 - (1.65)\frac{4.1}{\sqrt{225}} \le \mu \le 11.6 + (1.65)\frac{4.1}{\sqrt{225}}\right] = 0.90$$

$$P\left[11.6 - 0.45 \le \mu \le 11.6 + 0.45\right] = 0.90$$

$$P\left[11.15 \le \mu \le 12.05\right] = 0.90$$

أي ان متوسط عدد المقاعد الخالية خلال العام الماضي يقع داخل الفترة (12.05, 11.15)، وذلك بدرجة ثقة 90%.

اختبار الفروض حول متوسط المجتمع

مقدمة:

لماذا نحتاج إلى اختبار فرض؟ تصور أن إحدى شركات إنتاج المعلبات تكتب على المنتج أن وزنه الصافي 120جرام ثم تم أخذ عينة (Sample)حجمها 100من هذه المعلبات لاختبارها، ووجد أن الوسط الحسابي أو المتوسط لأوزان هذه العينة هو 118.9جرام، فهل يمكن الحكم على هذه الشركة بأنها تقوم بعملية غش أو كذب على المستهلك بناءً على هذه العينة؟ وأن جميع منتجات هذه الشركة من هذه المعلبات أقل من 120 جرام؟ الإجابة بالطبع لا؛ حتى نقوم بعمل إجراء إحصائي للتحقق من صحة ما كتب على المنتج وهذا الإجراء هو ما يسمى باختبار فرض (Test of hypothesis)حول متوسط المجتمع.

السبب أن المتوسط $\bar{x} = 118.9$ جرام قد حصلنا عليه من عينة؛ وأن الفرق بين الــ 120جرام (المتوسط المطلوب للوزن في المجتمع (Population)والــ 118.9جرام (المتوسط الذي تم الحصول عليه من العينة) قد يكون نتيجة خطأ في اختيار العينة (Sample error)وتصور أنه أخذت عينة أخرى من نفس منتجات الشركة حجمها أيضًا 100 وحدة؛وأن متوسط الوزن لهذه العينة الأخرى كان 120.4وحدة؛ وأن متوسط الوزن لهذه العينة الأخرى كان 120.4جرام.

لهذا فإننا قد نقوم بإجراء اختبار فرض لنرى حجم الفرق بين الــ 120جرام والــ 118.9جرام ولنقرر بالنفي أو الإيجاب وقوع هذا الفرق بالصدفة وحدها.

ولكن إذا كان الوزن 118.9جرام هو متوسط وزن « جميع » المعلبات وليس فقط للعينة فلسنا في حاجة لإجراء اختبار فرض في هذه الحالة ونقرر مباشرة متوسط وزن إنتاج هذه الشركة من المعلبات أقل من 120 جرام.ولذلك فإننا: نقوم بإجراء اختبار فرض فقط عند البحث عن قرار بشأن معلمة للمجتمع بناءً على قيمة الإحصاء في العينة.


المجتمع
عينة
إحصاءة
اختبار فرض
قرار

خطوات إجراء أي اختبار للفروض الإحصائية بشكل عام كما يلي:

- صياغة فرضان يسميان **فرض العدم** و**الفرض البديل** حول معلمة (أو خاصية) في مجتمع الدراسة.
- حساب بعض الإحصاءات كالمتوسط، والانحراف المعياري...الخ.
- نحسب من قيم الاحصاءات **إحصاء الاختبار**.
- نتخذ القرار برفض أو عدم رفض فرض العدم.

فرض العدم H_0 **(Null Hypothesis)** هو ادعاء عن معلمة مجتمع يفترض صحته حتى يثبت عكس ذلك.

الفرض البديل H_1 **(Alternative Hypothesis)** هو ادعاء عن معلمة مجتمع سوف يكون صحيحًا إذا كان فرض العدم غير صحيح.

إحصاء الاختبار **(Test Statistics)** هو أسلوب أو طريقة لتحديد قاعدة ترفض فرض العدم.

اختبار فرض حول متوسط المجتمع من جانبين في حالة العينات الكبيرة $(n \geq 30)$

- فرض العدم والفرض البديل:

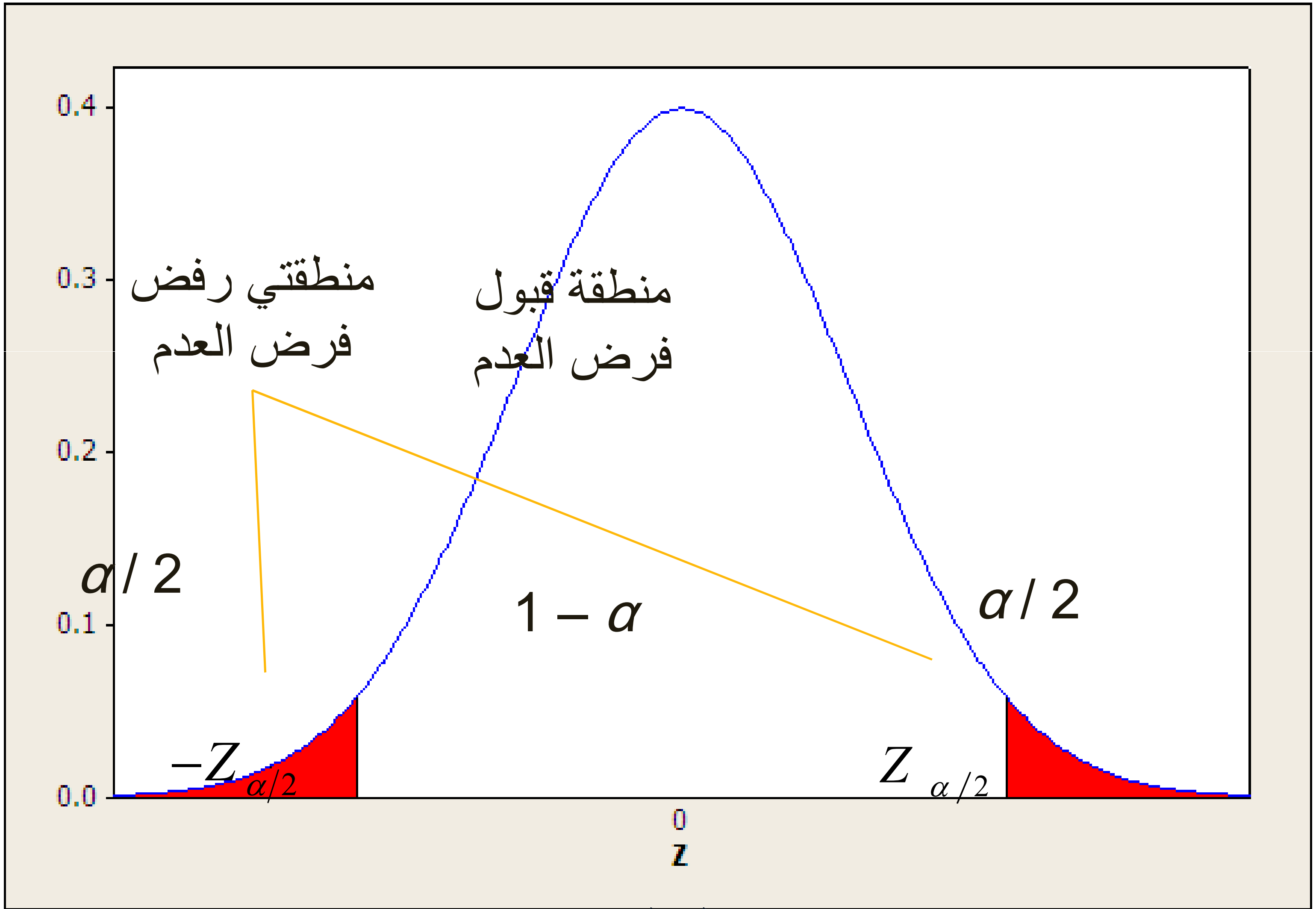
$$H_0 : \mu = \mu_0 \quad (\text{vs}) \quad H_1 : \mu \neq \mu_0$$

- إحصاء الاختبار (**Z**):

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

نحدد مكان وقوع قيمة إحصاء الاختبار (Z)في شكل (8 - 1).

إذا وقعت قيمة (Z)في منطقة القبول ($1 - \alpha$) نقبل فرض العدم . وإذا وقعت قيمة (Z)في منطقة الرفض نرفض فرض العدم ونقبل الفرض البديل .



شكل (8 - 1)

ملاحظات هامة:

- إذا كانت قيمة المعلمة σ مجهولة نستخدم قيمة الانحراف المعياري S للعينة بدلاً عنها كتقدير لها.
- قيم $Z_{\alpha/2}$ نفسها المستخدمة في التقدير.

مثال :

إذا كان متوسط الزيادة في أجور العاملين في إحدى المؤسسات عام 1998 هو (36) ريال ، وفي عام 2001م أخذت عينة من (64) فرداً من العاملين في هذه المؤسسة، فوجد أن الوسط الحسابي للزيادة في أجورهم (40) ريال والانحراف المعياري (8) ريال . هل يدل ذلك على أن متوسط الزيادة في أجور العاملين في المؤسسة عام 2001م قد اختلف عن متوسط الزيادة في الأجور عام 1998م؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل:

فرض العدم والفرض البديل:

$$H_0 : \mu = 36 \quad (\text{vs}) \quad H_1 : \mu \neq 36$$

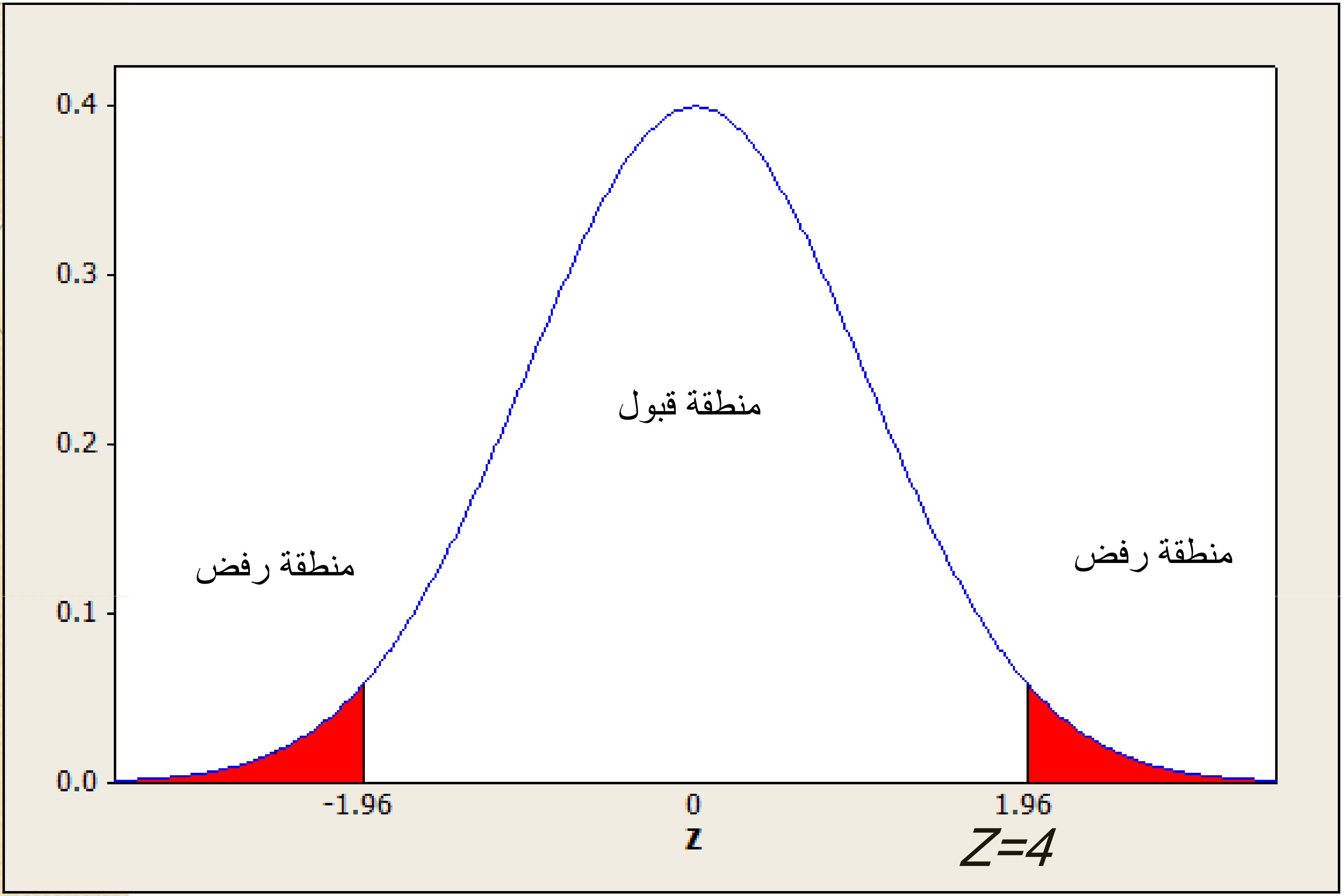
معطيات:

$$n = 64, \quad \bar{x} = 40, \quad S = 8, \quad \alpha = 0.05 \Leftrightarrow Z_{\alpha/2} = 1.96$$

إحصاء الاختبار (**Z**):

$$Z = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{40 - 36}{8/\sqrt{64}} = 4$$

القرار:



نرفض فرض العدم H_0 ونقبل الفرض البديل H_1 ؛ أي أن متوسط الزيادة في الأجور عام 2001م قد اختلف عن متوسط الزيادة في الأجور عام 1998م. وذلك عند مستوى معنوية $\alpha = 0.05$.

مثال :

شركة لإنتاج وتعبئة منتج غذائي أرادت أن تختبر كفاءة إحدى ماكينات التعبئة الخاصة بالشركة، والماكينة مصممة لملء عبوة بما مقداره $\mu = 120$ جرام، وأرادت الشركة أن تتأكد من سلامة الماكينة، وعملية الجودة هذه دفعت الشركة إلى سحب عينة حجمها 100عبوة للتأكد من سلامة الماكينة، ووجد أن الوسط الحسابي للعينة ، والانحراف المعياري 118.5 , 5 على التوالي .

اختبر الفرض القائل بأن متوسط التعبئة يختلف عن (120) جرام، وذلك عند مستوى معنوية $0.01 = \alpha$.

الحل:

فرض العدم والفرض البديل:

$$H_0 : \mu = 120 \quad (\text{vs}) \quad H_1 : \mu \neq 120$$

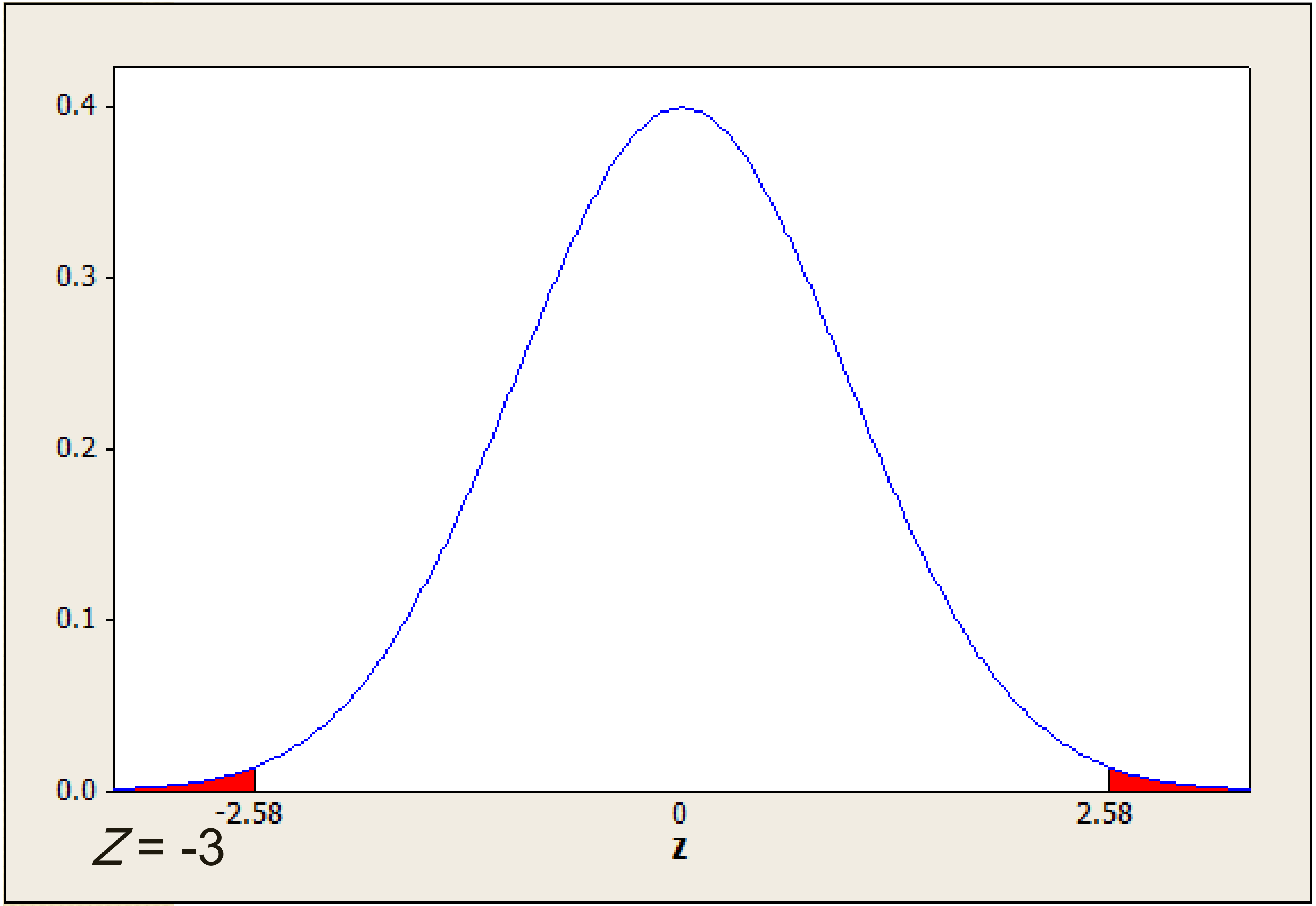
معطيات:

$$n = 100, \quad \bar{x} = 118.5, \quad S = 5, \quad \alpha = 0.01 \Leftrightarrow Z_{\alpha/2} = 2.58$$

إحصاء الاختبار (**Z**):

$$Z = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{118.5 - 120}{5/\sqrt{100}} = -3$$

القرار:



نرفض فرض العدم H_0 ونقبل الفرض البديل H_1 ؛ أي أن متوسط التعبئة قد اختلف عن 120 جرام. وذلك عند مستوى معنوية $\alpha = 0.01$.

استخدام قيمة (P) لاختبار الفروض الإحصائية

- قيمة (**P-Value**)هي أصغر قيمة لمستوى المعنوية (α) يمكن عندها رفض فرض العدم.
- حيث إذا كانت _ $\alpha \leq$ **P-Value** فإننا نقبل فرض العدم.
- حيث إذا كانت _ $\alpha >$ **P-Value** فإننا نرفض فرض العدم.

وأغلب البرامج الإحصائية توفر هذه القيمة مباشرة عند إجراء الاستدلال الإحصائي على مجموعة من البيانات، وتكون تحت مسمى **"P-Value"** أو **"P"** أو **"Sig."**.

مثال:

أدعى أحد الباحثين في دراسة له حول متوسط الاستهلاك لأحد المواد الغذائية الأساسية أنه يساوي (1000) كيلوجرام في إحدى الأحياءٍ بمدينة جدة. وبعد جمع وإدخال البيانات في الحاسب الآلي ومعالجتها إحصائياً وباستخدام برنامج Excel، حصل على النتيجة التالية:

P-Value = 0.600492

ما هو القرار حول متوسط الاستهلاك، هل يختلف عن (1000) كيلوجرام؟ اختبر هذا الفرض عند مستوى معنوية 0.05.

الحل:

من المعطيات نجد أن فرض العدم والفرض البديل هما على النحو الآتي:

$$H_0 : \mu = 1000 \quad (\text{vs}) \quad H_1 : \mu \neq 1000$$

حيث أن قيمة P-Value > α فإننا لا نرفض فرض العدم، أي ان متوسط الاستهلاك للمادة الغذائية الاساسية يساوي (1000) كيلوجرام في إحدى الأحياء بمدينة جدة، وذلك عند مستوى معنوية 0.05.

مثال :

أدعى أحد الباحثين أن متوسط سرعة القراءة لدى طلاب إحدى الجامعات تساوي (200) كلمة فيٍ الدقيقة. وبعد جمع وإدخال البيانات في الحاسب الآلي ومعالجتها إحصائياً وباستخدام برنامج Excel، حصل على النتيجة التالية:

P-Value = 0.000010

هل متوسط سرعة القراءة لدى هؤلاء الطلاب يختلف عن (200) كلمة في الدقيقة، عند مستوى معنوية 0.10.

الحل:

من المعطيات نجد أن فرض العدم والفرض البديل هي على النحو الآتـي:

$$H_0 : \mu = 200 \quad (\text{vs}) \quad H_1 : \mu \neq 200$$

حيث أن قيمة α > P-Value فإننا نرفض فرض العدم ونقبل الفرض البديل، أي أن متوسط سرعة القراءة لا يساوي (200) كلمة في الدقيقة لدى هؤلاء الطلاب، وذلك عند مستوى معنوية 0.10.

اختبار مربع كاي

يستخدم هذا الاختبار في الحالات التالية :

- جودة التوفيق
- الاستقلال
- التجانس

وفكرة تطبيق اختبار مربع كاي للاستقلال لمعرفة إذا كان هناك علاقة بين صفتين من صفات المجتمع أم لا .

مثل : هل هناك علاقة بين مستوى الدخل و مستوى التعليم

هل هناك علاقة بين التدخين و الإصابة بسرطان الرئة .

ملاحظة : يكون مقياس البيانات اسمي او ترتيبي

خطوات إجراء اختبار مربع كاي للاستقلال :

١)- صياغة الفروض الإحصائية :

فرض العدم H_0 : لا توجد علاقة بين الصفتين .

الفرض البديل H_1 : توجد علاقة بين الصفتين .

٢)- تكوين جدول التوافق وتحديد التكرارات المشاهدة : وهو يحتوي على التكرارات المشاهدة O_{ij} لكل خلية . حيث تمثل O_{ij} القيمة المشاهدة للخلية التي تقع في الصف i والعمود j .

٣)- حساب التكرار المتوقع : E_{ij} ويحسب لكل خلية بواسطة العلاقة التالية :

$$E_{ij} = \frac{\text{مجموع الصف الذي به الخلية } (i) \times \text{ مجموع العمود الذي به الخلية } (j)}{\text{مجموع التكرارات (حجم العينة)}}$$

٤)- حساب إحصاء الاختبار : نوجد χ^2 المحسوبة (الفعلية) من العلاقة التالية :

$$\chi^2 = \sum_i \sum_j \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

وتوجد صيغة أخرى لحساب إحصاء الاختبار وهي

$$\chi^2 = \sum_i \sum_j \frac{O_{ij}^2}{E_{ij}} - n$$

٥)- تحديد قاعدة الرفض : نوجد χ^2 النظرية (الجدولية) بدرجات حرية مناسبة

$$\chi^2_{[(r-1)(c-1),\alpha]}$$

حيث : r : عدد الصفوف c : عدد الأعمدة

α : مستوى المعنوية

٦)- القرار :

- إذا كانت قيمة χ^2 المحسوبة (الفعلية) أقل من قيمة χ^2 النظرية
(الجدولية) بمعنى إذا وقعت χ^2 المحسوبة في منطقة القبول
إذا **نقبل** H_0 ونرفض H_1 **أي أنه لا توجد علاقة بين الصفتين**
.

- إذا كانت قيمة χ^2 المحسوبة (الفعلية) أكبر من قيمة χ^2 النظرية
(الجدولية) بمعنى إذا وقعت χ^2 المحسوبة في منطقة الرفض
إذا **نرفض** H_0 ونقبل H_1 **أي أنه توجد علاقة بين الصفتين** .

مثال :

شركة تأمين ترغب في دراسة العلاقة بين عمر السائق وحوادث السيارات ، فقامت باختيار عينة مكونة من 140 شخص ، وتم تقسيمهم حسب الأعمار إلى ثلاث فئات ، وسئل كل واحد منهم هل حصل له حادث عام 1429 هـ ، فكانت النتائج موضحة في الجدول التالي :

عمر السائق	هل حصل لك حادث عام 1429 هـ	
	نعم	لا
أقل من 25 سنة	34	26
بين 25 و 35 سنة	14	36
أكبر من 35 سنة	12	18

هل تدل النتائج على وجود علاقة بين عمر السائق وحوادث السيارات ؟ أختبر عند مستوى معنوية 0.05 ؟

بحساب التكرارات المتوقعة لكل خلية نحصل على الجدول التالي:

عمر السائق	هل حصل لك حادث عام 1429 هـ		$\sum$
	نعم	لا	
أقل من 25 سنة	25.71	34.29	60
بين 25 و 35 سنة	21.43	28.57	50
أكبر من 35 سنة	12.86	17.14	30
$\sum$	60	80	140

جدول توزيع مربع كاي

درجات الحرية	0.2	0.1	0.05	0.025	0.01
1	1.6424	2.7055	3.8415	5.0239	6.6349
2	3.2189	4.6052	5.9915	7.3778	9.2103
3	4.6416	6.2514	7.8147	9.3484	11.3449
4	5.9886	7.7794	9.4877	11.1433	13.2767
5	7.2893	9.2364	11.0705	12.8325	15.0863
6	8.5581	10.6446	12.5916	14.4494	16.8119
7	9.8032	12.0170	14.0671	16.0128	18.4753
8	11.0301	13.3616	15.5073	17.5345	20.0902
9	12.2421	14.6837	16.9190	19.0228	21.6660
10	13.4420	15.9872	18.3070	20.4832	23.2093
11	14.6314	17.2750	19.6751	21.9200	24.7250
12	15.8120	18.5493	21.0261	23.3367	26.2170
13	16.9848	19.8119	22.3620	24.7356	27.6882
14	18.1508	21.0641	23.6848	26.1189	29.1412
15	19.3107	22.3071	24.9958	27.4884	30.5779
16	20.4651	23.5418	26.2962	28.8454	31.9999
17	21.6146	24.7690	27.5871	30.1910	33.4087
18	22.7595	25.9894	28.8693	31.5264	34.8053
19	23.9004	27.2036	30.1435	32.8523	36.1909
20	25.0375	28.4120	31.4104	34.1696	37.5662
21	26.1711	29.6151	32.6706	35.4789	38.9322
22	27.3015	30.8133	33.9244	36.7807	40.2894
23	28.4288	32.0069	35.1725	38.0756	41.6384
24	29.5533	33.1962	36.4150	39.3641	42.9798
25	30.6752	34.3816	37.6525	40.6465	44.3141
26	31.7946	35.5632	38.8851	41.9232	45.6417
27	32.9117	36.7412	40.1133	43.1945	46.9629
28	34.0266	37.9159	41.3371	44.4608	48.2782
29	35.1394	39.0875	42.5570	45.7223	49.5879
30	36.2502	40.2560	43.7730	46.9792	50.8922

١)- صياغة الفروض الإحصائية :

فرض العدم H_0: لا توجد علاقة بين الصفتين .

الفرض البديل H_1: توجد علاقة بين الصفتين .

٢)- حساب إحصاء الاختبار :

$$\chi^2 = \sum_i \sum_j \frac{O_{ij}^2}{E_{ij}} - n$$

$$= \frac{34^2}{25.71} + \frac{26^2}{34.29} + \frac{14^2}{21.34} + \frac{36^2}{28.57} + \frac{12^2}{12.86} + \frac{18^2}{17.14} - 140$$

$$= 44.96 + 19.71 + 9.15 + 45.36 + 11.20 + 18.90 - 140$$

$$= 149.28 - 140 = 9.28$$

٣)– تحديد قاعدة الرفض :

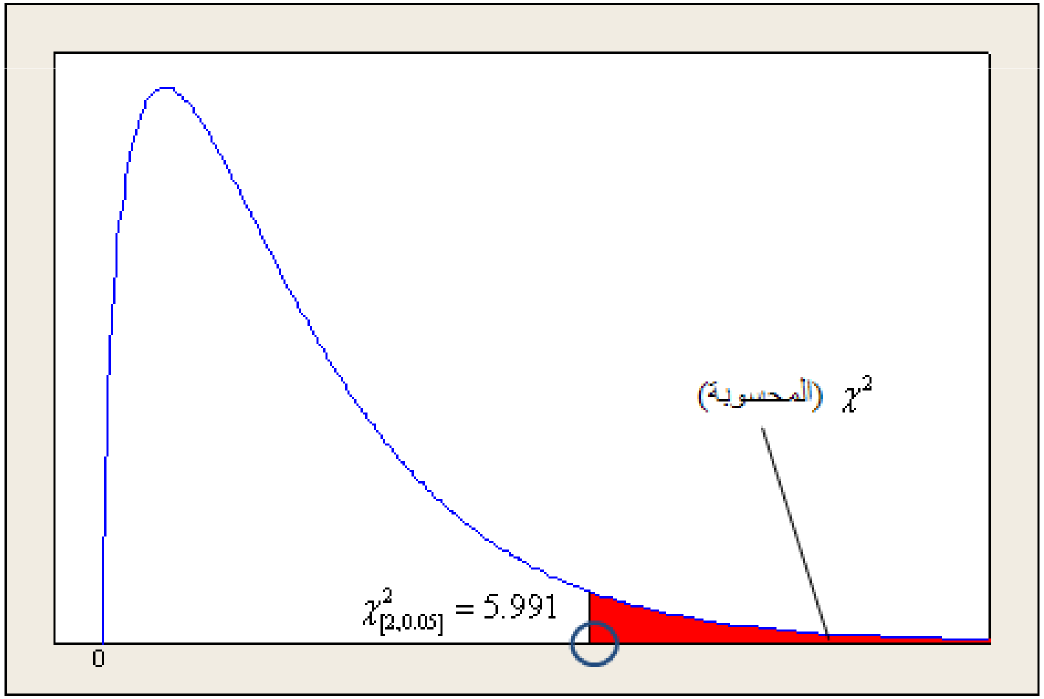
$$\chi^2_{[(r-1)(c-1),\alpha]} = \chi^2_{[(3-1)(2-1),0.05]} = \chi^2_{[2\times 1,0.05]}$$

$$= \chi^2_{[2,0.05]} = 5.991$$

٤)- القرار :

بما أن قيمة χ^2 المحسوبة = 9.28 أكبر من قيمة χ^2 الجدولية = 5.991 إذا **نرفض** فرض العدم H_0 وبالتالي نقبل الفرض H_1 البديل

أي أنه توجد علاقة بين عمر السائق وحوادث السيارات .



اختبار ذاتي.

إذا كان من المعروف أن متوسط الوقت اللازم من قبل العامل لإنجاز عمل معين بإحدى الشركات هو (12) دقيقة. اختيرت عينة من (100) عامل في تلك الشركة، وأجرى لهم برنامج تدريبي على أداء ذلك العمل. وبعد إتمام التدريب سجل الوقت اللازم من كل منهم لإنجاز ذلك العمل، فكان الوسط الحسابي والانحراف المعياري للعينة هما(10) دقائق ، (1.5) دقيقة على التوالي. أدعى مدير الشركة أن برنامج التدريب كان له تأثير إيجابي على متوسط الوقت اللازم لإنجاز ذلك العمل ؟

١:الفرض البديل هو

A.	B.	C.	D.
$H_0: \mu = 12$	$H_1: \mu = 12$	$H_0: \mu \neq 12$	$H_1: \mu \neq 12$

٢:قيمة (Z) المحسوبة

A.	B.	C.	D.
-1.3	-13.3	-23.3	0

٣: القرار عند مستوى معنوية α=0.05 مع العلم ان (P-value=0.000)

A.	B.	C.	D.
رفض فرض العدم	رفض الفرض البديل	قبول فرض العدم	الفقرتين (B ,C)

في دراسة لتقدير متوسط الإنفاق الشهري لدى الأسر بإحدى المدن الكبرى، أخذت عينة من (5625) أسره ، ووجد أن الوسط الحسابي يساوي (6000) ريال والانحراف المعياري يساوي (1500) ريال.

٤: الخطأ المعياري للمتوسط يساوي

A. 10	B. 20	C. 30	D. 40

٥: تقدير الفترة لمتوسط المجتمع بدرجة ثقة (95)% يتراوح بين

A. (5060.0,6039.2)	B. (5960.8, 6000.2)	C. (5000.8,6000.2)	D. (5960.8,6039.2)

٨-٦-٣: أدعى أحد الباحثين الاجتماعين ان عدد ساعات النوم له علاقة بمستوى أداء الطــلاب بالمرحلــة الثانوية في الاختبار النهائي، وبسؤال عينة من الطلاب بمدارس مختلفة عن عدد ساعات النوم خلال أيــام الاختبارات النهائية تم تقسيمهم حسب إجاباتهم الى فئتين: فئة عدد ساعات النوم مثالي والأخرى غير مثالي وتم تصنيفهم حسب مستوى الأداء في الاختبارات فحصلنا على النتائج التالية:

مستوى الأداء	عدد ساعات النوم		Σ
	غير مثالي	مثالي	
ضعيف	60	700	760
متوسط	40	430	470
عالي	70	200	270
Σ	170	1330	؟

٦: فرض العدم هو

D.	C.	B.	A.
H_0 : توجد علاقة بين عدد ساعات النوم ومستوى أداء الطلاب بالمرحلة الثانوية في الاختبار النهائي.	H_0: لاتوجد علاقة بين عدد ساعات النوم ومستوى أداء الطلاب بالمرحلة الثانوية في الاختبار النهائي.	H_0: لاتوجد علاقة بين مستوى أداء الطلاب بالمرحلة الثانوية في الاختبار النهائي ووضعهم الاجتماعي.	H_0 : توجد علاقة بين مستوى أداء الطلاب بالمرحلة الثانوية في الاختبار النهائي ووضعهم الاجتماعي.

٧:عدد الطلاب في عينة البحث

A. 1000	B. 1500	C. 2000	D. 2500

٨:التكرار المتوقع لعدد الطلاب الذين عدد ساعات نومهم مثالي ومستوى أدائهم عالي يساوي

A. 30.6	B. 70.0	C. 200.0	D. 239.4

٩: قيمة χ^2 النظرية (الجدولية) عند مستوى المعنوية α=0.01

A. 5.991	B. 0.0	C. 10.251	D. 9.210

١٠:إذا كانت قيمة إحصاء الاختبار تساوي $\chi^2 = 69.88$ ، فإن القرار هو

A. قبول فرض العدم	B. قبول الفرض البديل	C. رفض فرض العدم	D. الفقرتين B,C