الفصل III: إختبار الفرضيات

- يشكل إختبار الفرض أحد المواضيع الرئيسية للاستدلال الإحصائي، ويستهدف الوصول إلى القرار بشأن معلمة المجتمع من خلال قبول أو رفض تقدير المعتمد على معطيات العينة المسحوبة من ذلك المجتمع، ففي غالب الأحيان يحاول الباحث اتخاذ قرار لمشكلة محددة بشأن قواص توزيع ما (كالمتوسط أو نسبة) لعينة عشوائية تم سحبها من المجتمع، معتمدا على قرار إحصائي لذا يتوجب عليه وضع فرض عن قواص المجتمع، ومن هنا يختبر مدى صحة الفرض، لذا سنتناول في متن هذا الفصل مختلف هذه التفاصيل، وكذا مراحل اختبارات الفروض.

أولا: مفاهيم أساسية حول إختبار الفرض:

1- الفرض العدم والفرض البديل:

أ- الفرض العدم (الفرضية الصفرية): وهي التي تصاغ بشكل نفي وجود علاقة بين متغيرين أو فروق بين مجموعتين أو أكثر لمتغير مستقل على آخر تابع، وهي الفرضية التي يجري عليها الإختبار الإحصائي ويبنى القرار على عدم صحتها، ويرمز لها بالرمز H_0.

مثال: - لا توجد فروق ذات دلالة إحصائية في استخدام تكنولوجيا المعلومات تعزى لحجم المؤسسة.
- لا توجد علاقة بين توزيع مبيعات المؤسسة والارتفاع المبيعات.

ب- الفرض البديل: وهي المعاكسة للفرضية الصفرية ويرمز لها بالرمز H_1.

فعدم تحقق الفرضية الصفرية H_0 يؤدي إلى رفضها وبالتالي قبول الفرضية البديلة H_1.

مثال حول الفرضية الصفرية والبديلة:

إذا كان معدل الطلبة في مادة الإحصاء هو 10 وأردنا اختيار معدل عينة من أولئك الطلبة من أجل اتخاذ قرار يتعلق بمعدلاتهم في مادة الإحصاء، فإن الفرضيات تصاغ كما يلي:

الفرضية الصفرية: $H_0: \mu = 10$

أما الفرضية البديلة فإنها يمكن أن تأخذ ثلاثة أشكال وذلك حسب ما يقترحه الباحث:

$$H_1: \begin{cases} \mu \neq 10 \\ \mu > 10 \\ \mu < 10 \end{cases}$$

ففي الفرضية الصفرية نتيح الباحث عدم وجود فروق بين معدل العينة في مادة الإحصاء والمعدل العام للطلبة في هذه المادة، بينما الفرضية البديلة تتضمن وجود فروق (اختلاف) بين معدل الطلبة في العينة والمعدل العام، أو أن المعدل الأكبر من 10، أو أقل من 10.

2- مستوى المعنوية (مستوى الدلالة): هو احتمال رفض فرضية صفرية وهو صحيح ويرمز له بالرمز α وقد استخدمنا مستوى معنوية 5% فإن ذلك يعني أن الخطأ المسموح ما تكابد هو 5%، وأننا سوف نكون واثقين بنسبة 95% في أننا سنتخذ القرار السليم

3- إحصاء الاختبار: هو إحصاء يتم حساب قيمته من خلال بيانات العينة مثل (الإحصائي Z، والإحصائي t)

وتستخدم قيمته للوصول إلى قرار رفض أو قبول الفرضية الصفرية H_0، واختيار إحدى الإختبارات بعد دراسة شكل التوزيع والفرضيات المفروضة.

4- أخطاء إختبار الفرضيات وأنواعها:

إن عملية اتخاذ أي قرار إحصائي ينطوي على أخطاء بنسب معينة، حيث أنه من المحتمل أن نرفض فرضية معينة في حين أنها صحيحة، والعكس صحيح، وعموما يمكن القول أنه هناك نوعان من الأخطاء الإحصائية وهي:

- الخطأ من النوع الأول: هو الخطأ الذي نقع فيه عندما نرفض فرضية العدم H_0 عندما تكون صحيحة ويرمز لاحتمال وقوع هذا الخطأ بالرمز α وتسمية مستوى معنوية إختبار

- الخطأ من النوع الثاني: هو الخطأ الذي نقع فيه عندما نقبل الفرضية الصفرية بالرغم من عدم صحتها ويرمز لاحتمال هذا الخطأ بالرمز β

جدول رقم 1: الخطأ من النوع الأول α والثاني β.

حالة القرار	H_0 صحيحة	H_0 خاطئة
قبول H_0	قرار سليم $1-\alpha$	β
رفض H_0	α	قرار سليم $1-\beta$

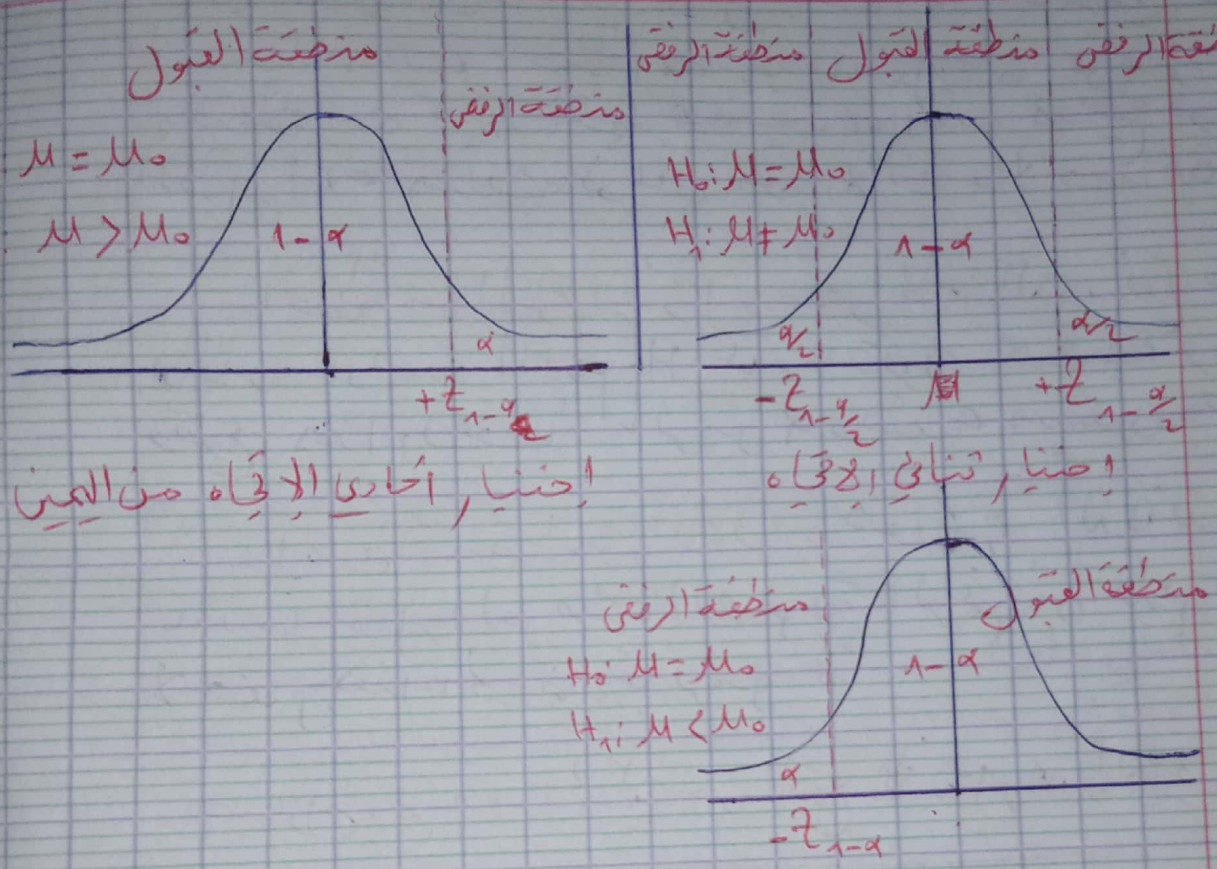
5- منطقة اتخاذ القرار (تحديد قاعدة القرار):

تقسم منطقة اتخاذ القرار إلى منطقتين احداهما تحوي القيم التي تدعم الفرضية البديلة وتسمى منطقة الرفض والمنطقة الأخرى تحتوي على القيم التي تدعم الفرضية الصفرية وتسمى منطقة القبول

فمنطقة الرفض إذن تتضمن قبول الفرضية البديلة

والتي يعبر عنها إما باتجاه واحد على اليمين أو على اليسار، أو باتجاهين على جانبي المنحنى الحرجين.

والشكل التالي يوضح منطقتي الرفض والقبول للفرضية H_0 في الحالات الثلاث:



إختبار أحادي الإتجاه من اليمين

إختبار ثنائي الإتجاه

إختبار أحادي الإتجاه من اليسار

ثانيا: إختبار الفرضيات المتعلقة بالمتوسط:

1. الإختبار ثنائي الإتجاه للمتوسط:

لتوضيح هذا النوع من الإختبار والكيفية التي يتم بها تنفيذ خطوات نضع المثال الموالي:

مثال:

وجد في دراسة سابقة أن معدل حضور الطلبة للمحاضرات بإحدى كليات جامعة عنابة هو 70%

طالبا، وأن توزيعها يقترب من التوزيع الطبيعي. نريد اختبار فرضية أن معدل الحضور قد تغير خلال هذا الموسم، فتم اختيار 81 طالبا عشوائيا، فوجد أن معدل الحضور هو 73 طالبا بانحراف معياري قدره 10 طلاب.
- باستعمال مستوى دلالة 5% هل تؤيد تلك الفرضية؟

الحل:

- الخطوة الأولى: صياغة الفرضية الصفرية والفرضية البديلة.

$$H_0 : \mu = 70$$

$$H_1 : \mu \neq 70$$

- الخطوة الثانية: تحديد مستوى الدلالة (مناطق القبول والرفض لفرضية العدم).

مستوى الدلالة هو 5%، بما أن الاختبار ثنائي الاتجاه فإن منطقة الرفض ستكون على جانبين ~~متساويين~~

- الخطوة الثالثة: تحديد قاعدة القرار:
بما أن توزيع المجتمع طبيعي فإننا سنستخدم قاعدة القرار بناء على الإحصاءة z_c، وكون الاختبار ثنائي الاتجاه فإن قاعدة القرار تكون كما يلي:

$$\begin{cases} -z_{1-\alpha/2} \leq z_c \leq +z_{1-\alpha/2} \Rightarrow H_0 \text{ (قبول) عدم رفض} \\ z_c < -z_{1-\alpha/2} \text{ أو } z_c > +z_{1-\alpha/2} \Rightarrow H_0 \text{ رفض} \end{cases}$$

- الخطوة الرابعة: حساب القيمة الجدولية والفعلية للإحصاءة الاختبار.

أولا: القيمة الجدولية:

$$\alpha = 0,05 \Rightarrow z_{1-\frac{0,05}{2}} = z_{0,975} = 1,96$$

ثانيا: القيمة المحسوبة:

$$z_c = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{73 - 70}{\frac{10}{\sqrt{81}}} = 2,7$$

- الخطوة الخامسة: المقارنة، واتخاذ القرار

نلاحظ أن: $z_c > z_{1-\frac{\alpha}{2}}$ أي أن: $2,7 > 1,96$

وهذا يعني رفض الفرضية الصفرية وقبول الفرضية البديلة مع أتمالى أن متوسط الحضور قد تغير هذا الموسم.

مثال 2: نريد اختبار الفرضية

$H_0 : \mu = 25$

$H_1 : \mu \neq 25$

عند مستوى معنوية 0,05 وباستعمال الوسط الحسابي لعينة حجمها 100 أخذت من مجتمع انحرافه المعياري 12

المطلوب:

- إيجاد احتمال الوقوع في الخطأ من النوع β عندما يكون $\mu = 25,1$

الحل: المطلوب هنا:

$\beta = p$ (قبول H_0 / H_1 صحيحة)

الخطأ المعياري لمتوسط العينة $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{100}} = 1,2$

بما أن حجم العينة أكبر من 30 فإنه يمكن الاعتماد على نظرية النهاية المركزية وبالتالي استعمال z.

وعند المستقبل مستوى معنوية $\alpha = 0,05$ نرفض H_0 عندما
$25 \mp 1,96 \cdot 1,2$ لقيم $\bar{x}$ خارج المجال

أي أننا نرفض H_0 لما: $\begin{cases} \bar{x} > 22,648 \\ \bar{x} < 27,352 \end{cases}$

إذن:

$$\beta = P(\text{نقبل } H_0 / H_1 \text{ صحيحة}) \Rightarrow$$

$$\beta = P(22,648 < \bar{x} < 27,352)$$

$$\beta = P\left(\frac{22,648-26,6}{1,2} < Z < \frac{27,352-26,6}{1,2}\right) \Rightarrow$$

$$\beta = P(-2,46 < Z < 1,46) = 0,4931 + 0,4279 = 0,921$$

نلاحظ أن قيمة احتمال الوقوع في الخطأ الثاني β مرتفعة جدا لأن القيمة 26,6 في الفرضية البديلة قريبة جدا من القيمة 25 في الفرضية الصفرية.

2- اختبار أحادي الاتجاه من اليمين للمتوسط:

مثال: بلغ معدل تسجيل متوسط البكالوريا في كلية العلوم الاقتصادية لجامعة قسنطينة 11,5 وفي السنوات الأخيرة زاد الطلب على التسجيل في هذه الكلية فرغبت الإدارة معدل القبول للدراسة بها. ثم اختيار عينة عشوائية من الطلبة المسجلين بالكلية هذا العام حجمها 49 طالبا فأعطى معدل التسجيل 11,80 بانحراف معياري 1,4.
- قم باختبار الفرضية التي تقول بأن معدل القبول قد ارتفع عند مستوى معنوية 5%.

الحل: صياغة الفرضية:

$H_0: \mu = 11,5$

$H_1: \mu > 11,5$

الخطوة الثانية: تحديد مستوى الدلالة: $\alpha = 0,05$

الخطوة الثالثة: تحديد قاعدة القرار:

بما أن حجم العينة أكبر من 30 فإنه يتم التقريب للتوزيع الطبيعي، ونستعمل قاعدة القرار بناءً على إذا حساب Z_c وكون الاختبار أحادي الجانب فإن قاعدة القرار تكتب كالتالي:

$$\begin{cases} Z_c \leq +Z_{1-\alpha} \Rightarrow H_0 \text{ قبول} \\ Z_c > +Z_{1-\alpha} \Rightarrow H_0 \text{ رفض} \end{cases}$$

الخطوة الرابعة: حساب القيمة الحدودية والعملية لحساب الاختبار

أولا: القيمة الحدودية:

$\alpha = 0,05 \Rightarrow 1 - 0,05 = 0,95$

$\Rightarrow Z_{0,95} = 1,65$

ثانيا: القيمة العملية:

$$Z_c = \frac{\bar{x} - \mu}{\delta_{\bar{x}}} = \frac{11,8 - 11,5}{\frac{1,5}{\sqrt{49}}} = 1,4$$

الخطوة الخامسة: اتخاذ القرار:

نلاحظ أن: $Z_c < +Z_{0,95}$

أي أن: $1,4 < 1,65 \Leftarrow$ قبول الفرضية H_0 ورفض H_1 أي أنه يمكن أن معدل المستهل بالليرة لم يرتفع.

3- اختبار أحادي الاتجاه من اليسار للمتوسط:

مثال: لاحظت إدارة محطة بنزين أن معدل تعبئة

انتظار السيارة للتزود بالوقود هو 4 دقائق، وأراد أن تقلص من هذه الفترة فقامت بادارة تنظيم صفوف الانتظار، فأخذت عينة حجمها 9 سيارات، وبعد أن دونت فترات الانتظار، وجدت أن معدل فترات الانتظار هو 3 دقائق بانحراف معياري 30 ثانية

المطلوب: إختبار قرار إدارة المحطة لتخفيض فترات الانتظار عند مستوى معنوية 0,01.

الحل:

الخطوة الأولى: صياغة الفرضيات.

تخفيض فترة الانتظار ⇐

$H_0: M = 4$

$H_1: M < 4$

الخطوة الثانية: مستوى المعنوية: $\alpha = 0,01$

الخطوة الثالثة: تحديد قاعدة القرار:

$n < 30$، تباين المجتمع مجهول وبالتالي فإن العينة تتبع توزيع ستيودنت وستحدد قاعدة القرار بناءا على الإحصاء t كالتالي:

$t_c \geq -t_{1-\alpha, n-1} \Rightarrow$ نقبل H_0

$t_c < -t_{1-\alpha, n-1} \Rightarrow$ نرفض H_0

الخطوة الرابعة: حساب t_c و t_{tab}.

حساب t_{tab}:

$\alpha = 0,01 \Rightarrow 1 - 0,01 = 0,99$

$\Rightarrow -t_{0,99,9} = -2,821$

ثانيا: حساب t_c

$$t_c = \frac{\bar{x} - \mu}{\delta_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{3-4}{\frac{0,5}{\sqrt{9}}} = -6,02$$

الخطوة الخامسة: المقارنة واتخاذ القرار:

نلاحظ أن:

$t_c < -t_{tab}$ أي $-6,02 < -2,821$

$\Leftarrow$ رفض H_0 وقبول H_1 أي نستنتج أن القرار بإعادة تنظيم طبقوف إذا انتظار أدى إلى انخفاض قرار انتظار السيارات للفوز بالبترين.

ثالثا: اختبار الفرضيات المتعلقة بالفرق بين وسطي مجتمعين $(\mu_1 - \mu_2)$:

إذا كان $\bar{x}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 تم سحبها من مجتمع له توزيع طبيعي متوسطه μ_1 وتباينه σ_1^2 معلوم، وكان $\bar{x}_2$ هو الوسط الحسابي لعينة عشوائية حجمها n_2 مستقلة عن الأولى تم سحبها من مجتمع له توزيع طبيعي متوسطه μ_2 وتباينه معلوم σ_2^2

إذا أردنا اختبار الفرضية الصفرية: $H_0: \mu_1 - \mu_2 = 0$

مقابل الفرضية البديلة:

$H_1: \mu_1 - \mu_2 \neq 0$ فإننا نرفض H_0 عند مستوى دلالة α إذا كانت $z_c > z_{1-\alpha/2}$ أو $z_c < -z_{1-\alpha/2}$

$H_1: \mu_1 - \mu_2 > 0$ فإننا نرفض H_0 عند α إذا كانت $z_c > z_{1-\alpha}$

$H_1: \mu_1 - \mu_2 < 0$ فإننا نرفض H_0 عند α إذا كانت $z_c < -z_{1-\alpha}$

وبالتالي:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

مثال: إذا علمت أن $\bar{X} \sim N(\mu_1, 36)$، سحبت منه عينة حجمها 16، وسطها الحسابي 30، و $\bar{X} \sim N(\mu_2, 25)$، سحبت منه عينة حجمها 12، وسطها الحسابي 33.

- اختبر وجود فرق معنوي بين متوسطي المجتمعين عند $\alpha = 10\%$.

الحل:

$n_1 = 16, n_2 = 12, \bar{X}_1 = 30, \bar{X}_2 = 33$

الفرضيات:

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 - \mu_2 \neq 0$

القيمة الجدولية:

$$Z_{1-\alpha/2} = Z_{1-\frac{0,1}{2}} = Z_{0,95} = 1,64$$

القيمة المحسوبة:

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{30 - 33}{\sqrt{\frac{36}{16} + \frac{25}{12}}} = -1,44$$

بما أن $-1,4 < 1,64$ أي $Z_c < Z_{tab}$

$\Rightarrow$ نقبل H_0 وبالتالي المجتمعين لهما نفس الوسط الحسابي

ثانيا: اختبار الفرضية حول السنة في مجتمع:

- إن الخطوات المتبعة في اختبار الفرضيات المتعلقة بالوسط هي نفسها التي سيتم استخدامها في اختبار الفرضيات

المتعلقة بالنسبة حيث نقارن بين نسبة العينة p' ونسبة المجتمع p

$$Z = \frac{p' - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

أما الفرضية الصفرية: $H_0 : p = p_0$

في حين أن الفرضية البديلة تأخذ الحالات الثلاث كما رأينا في اختيار الفرضيات المختلفة بالتوسع

مثال: تقدر الدوائر الرسمية نسبة المتزوجين الحاصلين على شهادة الأولى التي تليها خريجو الجامعيين الذين يحصلون على عمل في السنة الواحدة بـ 70٪. وجد دراسة أجريت على عينة من 900 طالب أن نسبة الحصول على عمل 67٪. كيف يمكن اختبار هذا المستوى معنوية 5٪ إذا كانت النسبة الرسمية صحيحة أم مبالغ فيها؟

$$H_0 : p = 0,7$$
$$H_1 : p < 0,7$$

$$Z_c = \frac{p' - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0,67 - 0,7}{\sqrt{\frac{0,7 \cdot 0,3}{900}}} = -196,34 < -Z_{1-0,05} = -1,64$$

$\Leftarrow$ رفض H_0 (RH_0) أي أن النسبة مبالغ فيها.