الفصل الحادي عشر

توزيع المعاينة والتقدير واختبارات الفروض

(١١ ـ ١) مقدمـــة

لقد سبق لنا في الفصل الأول تعريف المجتمع الإحصائي والعينة الإحصائية والأخيرة عبارة عن جزء من المجتمع. تعرضنا كذلك إلى طرق لأخذ العينات مثل العينة العشوائية البسيطة، والعينة الطبقية وغيرها والأسباب التي أدت إلى أخذ العينات. والمجتمع الإحصائي عادة يحتوي على معالم تكون غير معلومة مثل المتوسط والتباين. . . إلخ. ويرغب الباحثون عادة في تقدير مثل هذه المعالم من البيانات المأخوذة من العينات، وذلك بحساب متوسط العينة كتقدير لمتوسط المجتمع، وكذلك تباين العينة كتقدير لتباين المجتمع. . . إلخ. وتسمى التقديرات المحسوبة من العينة الإحصائيات. وتستخدم الإحصائيات هذه لتقدير معالم المجتمع، وذلك بأخذ أحد أسلوبين هما التقدير بنقطة أو التقدير بفترة، والتي تسمى عادة فترة الثقة. وتستخدم الإحصائيات في تحديد ما إذا كانت الاختلافات المشاهدة بين عينتين ترجع إلى تقلبات الصدفة أو إلى اختلافات معنوية فعلاً وهذا يتطلب دراسة ما يسمى باختبارات المعنوية، والفروض، والتي تساعد في اتخاذ القرارات. ولدراسة تقدير المعالم بفترة الثقة واختبارات الفروض لا بد لنا أولا من دراسة توزيعات المعاينة.

(١١ ـ ٢) توزيع المعاينة للأوساط

إذا كان لدينا مجتمع محدود عدد مفرداته ن وأخذنا كل العينات الممكنة ذات الحجم المتساوي ن، وحسبنا لكل منها الوسط الحسابي $\bar{س}$، ثم وضعنا هذه الأوساط

في جدول تكراري فإننا نحصل على ما يسمى توزيع المعاينة للأوساط. وهذا التوزيع يكون قريبًا جدًّا من التوزيع المعتدل، ومتوسطه يساوي متوسط المجتمع (μ)، وتباينه يكون أقل من تباين المجتمع (σ^2). فإذا رمزنا لمتوسط توزيع المعاينة للأوساط بالرمز $\mu(\bar{س})$ وتباين توزيع المعاينة للأوساط بالرمز تبا ($\bar{س}$) فإننا نحصل على

ا) $\mu(\bar{س}) = \mu$ .... (١)

ب) $\sigma^2(\bar{س}) = \begin{cases} \dfrac{\sigma^2}{ن}\left(\dfrac{ن - ن}{ن - ١}\right) & \text{إذا كان حجم المجتمع صغيراً} \\ \dfrac{\sigma^2}{ن} & \text{إذا كان حجم المجتمع كبيرًا} \end{cases}$ .... (٢)

ملاحظـة

الجذر التربيعي لتباين توزيع المعاينة للأوساط يسمى الخطأ المعياري وسوف نرمز له بالرمز σ ($\bar{س}$).

ويمكن التحقق من العلاقتين (١) ، (٢) السابقتين بالمثال التالي.

مثـال (١)

مجتمع يحتوي على أوزان خمسة أطفال في سن الرضاعة كالتالي:

٢ ، ٥ ، ٣ ، ٤ ، ٦ كجم

احسب متوسط التوزيع العيني للأوساط وكذلك الخطأ المعياري لكل العينات الممكنة المكونة من طفلين وتحقق من العلاقتين (١) ، (٢)

الحـــل

يمكن تلخيص الحل كالتالي :

العينات والأوساط الحسابية المناظرة لها

جمع العينات الممكنة	الوسط الحسابي $\bar{س}$ لكل عينة
(٢ ، ٥)	٣,٥
(٢ ، ٣)	٢,٥
(٢ ، ٤)	٣
(٢ ، ٦)	٤
(٥ ، ٣)	٤
(٥ ، ٤)	٤,٥
(٥ ، ٦)	٥,٥
(٣ ، ٤)	٣,٥
(٣ ، ٦)	٤,٥
(٤ ، ٦)	٥
المجمـــوع	٤٠,٠

وتكون القيمة المتوقعة للوسط هي :

$$\mu\ (\bar{س}) = \mu\,(\bar{س}) = \frac{٤٠}{١٠} = ٤ \text{ كجم}$$

$$\text{متوسط المجتمع } \mu = \frac{١}{٥}\ (٢ + ٥ + ٣ + ٤ + ٦) = ٤ \text{ كجم}$$

$$\therefore\ \mu\,(\bar{س}) = \mu$$

أي أن العلاقة (١) تكون صحيحة

ولحساب الخطأ المعياري σ ($\bar{س}$) نكوّن الجدول التالي :

$\bar{س}$	ك	ك $\bar{س}$	ك $\bar{س}^2$
٢,٥	١	٢,٥	٦,٢٥
٣	١	٣,٠	٩,٠٠
٣,٥	٢	٧,٠	٢٤,٥
٤	٢	٨,٠	٣٢,٠٠
٤,٥	٢	٩,٠	٤٠,٥
٥	١	٥,٠	٢٥,٠٠
٥,٥	١	٥,٥	٣٠,٢٥
المجموع	١٠	٤٠,٠	١٦٧,٥

$$\sigma^2(\bar{\text{س}}) = \frac{١}{\text{ن}}\ \text{مجـ ك}\ \bar{\text{س}}^2 - \left(\frac{\text{مجـ ك}\ \bar{\text{س}}}{\text{ن}}\right)^2$$

$$= \frac{١}{١٠}(١٦٧,٥) - \left(\frac{٤٠}{١٠}\right)^2$$

ومن ذلك يكون

$$\sigma^2(\bar{\text{س}}) = ٠,٧٥$$

$$\sigma(\bar{\text{س}}) = \sqrt{٠,٧٥} = ٠,٨٦٦$$

$$\sigma^2 = \frac{١}{\text{ن}}\ \text{مجـ}\ (\bar{\text{س}} - \mu)^2$$

$$= \frac{١}{٥}\left[(٢-٤)^2 + (٥-٤)^2 + (٣-٤)^2 + (٤-٤)^2 + (٦-٤)^2\right]$$

$$= \frac{١}{٥}(٤ + ١ + ١ + \text{صفر} + ٤) = ٢$$

$\therefore\ \sigma^2(\bar{\text{س}}) \neq \sigma^2$

وأن $\sigma^2(\bar{\text{س}}) < \sigma^2$

نحسب المقدار $\frac{\sigma^2}{\text{ن}}\left(\frac{\text{ن} - ٢}{\text{ن} - ١}\right) = \frac{٢}{٢}\left(\frac{٥ - ٢}{٥ - ١}\right) = ٠,٧٥$

وهو يساوي $\sigma^2(\bar{\text{س}})$ أي أن العلاقة (٢) تكون محققة .

(١١ - ٣) توزيع المعاينة للفرق بين متوسطين

إذا كان لدينا مجتمعان مستقلان متوسط كل منهما $\mu_١$ ، $\mu_٢$ وتباين كل منهما تبا$_١$ ، $\sigma^2_٢$ على الترتيب، وقد يكون المجتمع الأول، على سبيل المثال، أوزان طلاب جامعة الملك سعود أو أطوالهم، والمجتمع الثاني أوزان طلاب جامعة الملك فهد للبترول والمعادن أو أطوالهم. فإذا أخذنا عينة من المجتمع الأول حجمها ن$_١$ ، وكان متوسطها $\bar{\text{س}}_١$ وعينة من المجتمع الثاني حجمها ن$_٢$ وكان متوسطها $\bar{\text{س}}_٢$ فإن

ا) $\mu\,(\bar{\text{س}}_١ - \bar{\text{س}}_٢) = \mu\,(\bar{\text{س}}_١ - \bar{\text{س}}_٢) = \mu_١ - \mu_٢$ .... (٣)

ب) تباين $(\bar{\text{س}}_١ - \bar{\text{س}}_٢) = \sigma^2(\bar{\text{س}}_١ - \bar{\text{س}}_٢) = \frac{\sigma^2_١}{\text{ن}_١} + \frac{\sigma^2_٢}{\text{ن}_٢}$ .... (٤)

بفرض أن حجم كل من المجتمعين كبير.

مثال (٢)

إذا كان متوسط أوزان طلاب جامعة الملك سعود هو ٦٢كجم والانحراف المعياري لأوزان الطلاب هو ٤ كجم. وكان متوسط أوزان طلاب جامعة الملك فهد للبترول والمعادن هو ٦٥كجم وانحراف معياري قدره ٥كجم أخذت عينة من جامعة الملك سعود حجمها ٤٠ طالبا. ثم أخذت عينة من طلاب جامعة الملك فهد للبترول والمعادن حجمها ٣٠ طالبا فاحسب التوقع والخطأ المعياري للفرق بين متوسطي العينتين.

الحـل

نفرض أن متوسط العينة الأولى $\bar{س}_1$ ومتوسط العينة الثانية $\bar{س}_2$ ومتوسط المجتمع الأول μ_1 = ٦٢كجم، ومتوسط المجتمع الثاني μ_2 = ٦٥كجم وحيث إن:

$$\mu\,(\bar{س}_1 - \bar{س}_2) = \mu\,(\bar{س}_1 - \bar{س}_2) = \mu_1 - \mu_2$$

$$= ٦٢ - ٦٥ = -٣ \text{ كجم}$$

نفرض أن تباين المجتمع الأول $\sigma^2_1 = (٤)^2 = ١٦$

وتباين المجتمع الثاني $\sigma^2_2 = (٥)^2 = ٢٥$

عندئذ يكون:

$$\text{تباين}\,(\bar{س}_1 - \bar{س}_2) = \sigma^2(\bar{س}_1 - \bar{س}_2) = \frac{\sigma^2_1}{ن_1} + \frac{\sigma^2_2}{ن_2} = \frac{١٦}{٤٠} + \frac{٢٥}{٣٠}$$

ومن ذلك نجد أن:

$$\sigma^2(\bar{س}_1 - \bar{س}_2) = ٠,٤ + ٠,٨٣ = ١,٢٣$$

ويكون الخطأ المعياري

$$\sigma\,(\bar{س}_1 - \bar{س}_2) = \sqrt{١,٢٣} = ١,١٠٩$$

(١١ - ٤) توزيع المعاينة للنسبة

سندرس فيما يلي حالتين من توزيع المعاينة للنسبة.

(١١ - ٤ - ١) توزيع المعاينة للنسبة (ح ر)

أحيانا يكون المجتمع الإحصائي ذا صفتين فقط. فمثلاً عند دراسة ظاهرة التدخين فإن المجتمع ينقسم إلى قسمين: أشخاص يدخنون، وآخرون لا يدخنون. وكذلك عند دراسة إنتاج مصنع معين فإن وحدات الإنتاج تنقسم إلى نوعين: وحدات سليمة (صالحة للاستخدام)، وأخرى معيبة (غير صالحة للاستخدام)، .... الخ فإذا كان حجم المجتمع محل الدراسة هو **ن** وعدد العناصر التي لها الخاصية الأولى في

هذا المجتمع هو ا فيكون عدد العناصر التي لها الخاصية الأخرى هو **ن - ١**. وتكون نسبة عناصر المجموعة التي لها الخاصية الأولى من المجتمع تساوي $\frac{١}{ن}$، وسوف نرمز لها بالرمز ح، فإذا أخذنا كل العينات الممكنة ذات الحجم المتساوي ن، وحسبنا لكل عينة النسبة $ح_د$ فإن التوزيع العيني لـ $ح_د$ يقترب من التوزيع الطبيعي ويكون له توقع $\mu(ح_د)$ وتباين $\sigma^2(ح_د)$ يعطى كالتالي

$$\mu\,(ح_د) = \mu(ح_د) = \frac{١}{ن} = ح \qquad \ldots\ldots\ldots\ (٥)$$

$$\text{وتباين}\ (ح_د) = \sigma^2(ح_د) = \frac{ح\,(١ - ح)}{ن} \qquad \ldots\ldots\ldots\ (٦)$$

مثـال (٣)

إذا كانت نسبة المعيب في انتاج إحدى الماكينات هو ٢٠٪ أخذت عينة مكونة من ٣٠ وحدة فاحسب قيمة الاحتمال أن يكون بها نسبة معيب قدرها ١٦٪

الحـل

$$\text{نسبة المعيب في المجتمع}\ ح = \frac{٢٠}{١٠٠} = ٠,٢$$

$$\mu\ (ح_د) = ح = ٠,٢$$

$$\text{تباين}\ (ح_د) = \sigma^2(ح_د) = \frac{ح\,(١ - ح)}{ن} = \frac{٠,٢\ (١ - ٠,٢)}{٣٠}$$

$$= \frac{٠,٢ \times ٠,٨}{٣٠} = ٠,٠٠٥٣$$

ومن ذلك يكون الخطأ المعياري

$$\sigma\ (ح_د) = \sqrt{٠,٠٠٥٣} = ٠,٠٧٣$$

والإحصائية هي:

$$ص = \frac{ح_د - ح}{\sigma^2(ح_د)}$$

تتبع التوزيع الطبيعي القياسي ق (ص)

ومن ذلك نجد أن :

$$\text{الاحتمال المطلوب ح}\,(\text{ح}_{\text{د}} \geqslant ٠,١٦) = \text{ح}\,\left(\text{ص} \geqslant \frac{٠,١٦ - ٠,٢}{٠,٠٧٣}\right)$$

ومنه نجد أن :

$$\text{ح}\,(\text{ص} \geqslant -\,٠,٥٥) = \text{ق}\,(-٠,٥٥) = ٠,٢٩١٢$$

أي أن الاحتمال المطلوب = ٠,٢٩ تقريبـا .

(١١ - ٤ - ٢) توزيع المعاينة لفروق النسب $(\text{ح}_{\text{د}_١} - \text{ح}_{\text{د}_٢})$

نفرض أن لدينا مجتمعين مستقلين وكان حجم المجتمع الأول $\text{ن}_١$ وعدد العناصر التي تتميز بالخاصية الأولى $\text{ا}_١$. وحجم المجتمع الثاني $\text{ن}_٢$ وعدد العناصر التي تتميز بالخـاصية الأولى أيضـا $\text{ا}_٢$. فإنه يكون لدينا نسبة الخاصية الأولى للمجتمع الأول $\text{ح}_١ = \frac{\text{ا}_١}{\text{ن}_١}$. وكذلك نسبة الخاصية الأولى للمجتمع الثاني $\text{ح}_٢ = \frac{\text{ا}_٢}{\text{ن}_٢}$ يلاحظ بأن توزيع المعاينة للفرق بين النسبتين $(\text{ح}_{\text{د}_١} - \text{ح}_{\text{د}_٢})$ يقترب من التوزيع الطبيعي بتوقع وتباين كما هو موضح فيما يلي :

$$\text{أ)}\quad \mu\,(\text{ح}_{\text{د}_١} - \text{ح}_{\text{د}_٢}) = \mu\,(\text{ح}_{\text{د}_١} - \text{ح}_{\text{د}_٢}) = \text{ح}_١ - \text{ح}_٢$$

$$\text{ب)}\quad \sigma^2\,(\text{ح}_{\text{د}_١} - \text{ح}_{\text{د}_٢}) = \frac{\text{ح}_١\,(١ - \text{ح}_١)}{\text{ن}_١} + \frac{\text{ح}_٢\,(١ - \text{ح}_٢)}{\text{ن}_٢}$$

حيث إن $\text{ن}_١$ ، $\text{ن}_٢$ هما حجم كل من العينة الأولى والثانية على الترتيب .

مثـال (٤)

إذا كان لدينا انتاج آلتين، وكانت نسبة المعيب للآلة الأولى هو ١٨٪ والمعيب للآلة الثانية هو ١٤٪ سحبت عينتان من انتاج الآلتين حجمهما ٤٠ ، ٦٠ وحدة على الترتيب . فإذا كانت نسبة المعيب للعينة الأولى هوح $_{\text{د}_١}$ ونسبة المعيب في العينة الثانية هوح $_{\text{د}_٢}$.

فاحسب قيمة الاحتمال ح $(\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢} \geq ٠,١)$.

الحـــل

ولحل المثال نجد أولاً :

أن توقع نسبة المعيب للعينة الأولى هو:

$$\text{مت}(\text{ح}_{\text{د}١}) = \mu(\text{ح}_{\text{د}١}) = ٠,١٨$$

وأن توقع نسبة المعيب للعينة الثانية هو:

$$\text{مت}(\text{ح}_{\text{د}٢}) = \mu(\text{ح}_{\text{د}٢}) = ٠,١٤$$

$$\text{تباين}(\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢}) = \sigma^2(\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢})$$

$$= \frac{\text{ح}_١(١ - \text{ح}_١)}{\text{ن}_١} + \frac{\text{ح}_٢(١ - \text{ح}_٢)}{\text{ن}_٢}$$

$$= \frac{٠,١٨\,(١ - ٠,١٨)}{٤٠} + \frac{٠,١٤\,(١ - ٠,١٤)}{٦٠}$$

$$= ٠,٠٠٥٧$$

ومن ذلك يكون الخطأ المعياري هو:

$$\sigma(\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢}) = \sqrt{٠,٠٠٥٧} = ٠,٠٧٥$$

والاحتمال المطلوب ح $(\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢} \geq ٠,١)$

ومن ذلك تكون الإحصائية المراد حسابها هي :

$$\text{ص} = \frac{(\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢}) - (\text{ح}_١ - \text{ح}_٢)}{\sigma^2(\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢})}$$

تتبع توزيعًا طبيعيًا قياسيًا ق (ص)

$$\text{ح}((\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢}) \geq ٠,١) = \text{ح}\left(\text{ص} \geq \frac{٠,١ - ٠,٠٤}{٠,٠٧٥}\right)$$

$$= \text{ح}(\text{ص} \geq ٠,٨)$$

$$= \text{ق}(٠,٨) = ٠,٧٨٨١$$

أي أن :

الاحتمال المطلوب = ٠,٧٩ تقريبًا

(١١ ـ ٥) التقديـر الإحصائي

سبق لنا دراسة بعض التوزيعات الإحصائية مثل توزيع ذي الحدين، وتوزيع بواسون والتوزيع المعتدل. ولاحظنا بعض المعالم المجهولة في هذه التوزيعات مثل ح في توزيع ذي الحدين، م في توزيع بواسون، وμ، σ في التوزيع المعتدل. فإذا كان المجتمع الإحصائي يتبع توزيعا معينا من هذه التوزيعات، فإننا نرغب في تقدير معالم هذا التوزيع من العينة المأخوذة من هذا المجتمع. يوجد نوعان من التقدير هما التقدير بنقـطة إذا قدِّرت معلمـة المجتمـع برقم واحد. والتقدير بفترة وهو أن تكون معلمة المجتمـع واقعـة بين رقمـين. وتسمى هذه الفترة فترة الثقة، وتعتبر فترة الثقة أفضل إحصـائيًا من التقدير بنقطة لأنها تكون مصحوبة بمعنوية أو دقة التقدير. ولتوضيح معنى التقدير نأخذ المثال التالي على سبيل المثال.

إذا كان متوسط أوزان عينة من طلاب جامعة الملك سعود هو $\bar{س}$ = ٦٥كجم فإن هذا يعني التقدير بنقطة لمتوسط مجتمع طلاب جامعة الملك سعود (μ). أما إذا كان متوسط أوزان طلاب الجامعة (μ) يقع بين القيمتين ٦٥ ± ٣ أي أن (μ) تقع بين الوزنين ٦٢ ، ٦٨ فإن هذا التقدير يسمى التقدير بفترة. وعادة ما تكون الفترة التي يقع داخلهـا المتوسط (μ) ذات احتمال مثل ٠,٩ ، ٠,٩٥ ، ٠,٩٩ . . . إلخ ، ونكتب الاحتمال في المثال السابق ح (٦٢ $\leqslant \mu \leqslant$ ٦٨) = ٠,٩ أو ٠,٩٥ . . . إلخ .

وسوف نستعرض فيما يلي تقدير فترة الثقة لكل من الأوساط والنسب والفرق بين متوسطين والفرق بين نسبتين.

(١١ ـ ٥ ـ ١) تقدير فترة الثقة للأوساط الحسابية (μ)

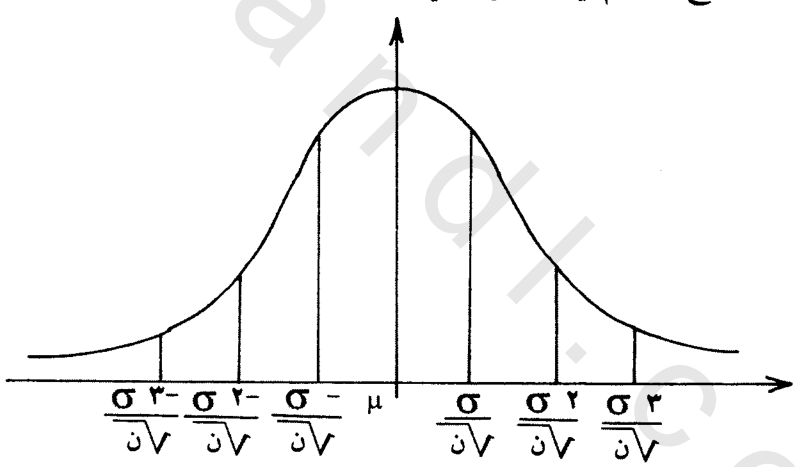
لقد سبق أن وجدنا أن توزيع المعاينة للأوساط يقترب من التوزيع المعتدل بتوقع $\mu(\bar{س}) = \mu$، وخطأ معياري يساوي $\frac{\sigma}{\sqrt{ن}}$. وأن نسبة عدد المتوسطات $\bar{س}$ على جانبي محور التماثل المار بمتوسط التوزيع للمعاينة $\mu(\bar{س}) = \mu$ تتعين كالتالي:

٦٨,٢٧٪ تقريبا تقع في الفترة $(\mu - \frac{\sigma}{\sqrt{ن}} ، \mu + \frac{\sigma}{\sqrt{ن}})$ ،

٩٥,٤٥٪ تقريبا تقع في الفترة $(\mu - ٢\frac{\sigma}{\sqrt{ن}} ، \mu + ٢\frac{\sigma}{\sqrt{ن}})$ ،

٩٩,٧٣٪ تقريبا تقع في الفترة $(\mu - ٣\frac{\sigma}{\sqrt{ن}} ، \mu + ٣\frac{\sigma}{\sqrt{ن}})$ ،

كما هو موضح بالرسم في الشكل التالي:



شكل (١١ ـ ١): نسب الإحتمالات حول محور التماثل تو

والنقطتان اللتان تبعدان عن μ بمقدار ٢ خطأ معياري تكون

$\mu - ٢\frac{\sigma}{\sqrt{ن}}$ ، $\mu + ٢\frac{\sigma}{\sqrt{ن}}$. ولذلك فإن الفترة $(\mu - ٢\frac{\sigma}{\sqrt{ن}} ، \mu + ٢\frac{\sigma}{\sqrt{ن}})$ تحتوي على ٩٥٪ من الأوساط الحسابية $\bar{س}$. وعندما يكون الوسط μ غير معلوم فإننا نستخدم الفترة التالية $(\bar{س} - ٢\frac{\sigma}{\sqrt{ن}} ، \bar{س} + ٢\frac{\sigma}{\sqrt{ن}})$ كتقدير بفترة لمعلمة

المجتمع (μ) بدرجة ثقة ٩٥٪ بأن يقع الوسط (μ) داخل هذه الفترة ويكتب احتمال التقدير بفترة الثقة عند احتمال يساوي ٠,٩٥ كالتالي:

$$\text{ح}\left(\bar{\text{س}} - ٢\frac{\sigma}{\sqrt{\text{ن}}} \leqslant \mu \leqslant \bar{\text{س}} + \frac{\sigma}{\sqrt{\text{ن}}}\right) = ٠,٩٥$$

ويكتب تقدير حدود فترة الثقة بصفة عامة كالأوساط مثالاً كالتالي:

$$\bar{\text{س}} \pm \text{ص}_{\frac{\text{أ}}{٢}} \frac{\sigma}{\sqrt{\text{ن}}}$$

حيث احتمال التقدير بفترة في هذه الحالة أكبر أو يساوي ١ - أ ويوضح كالتالي:

$$\text{ح}\left(\bar{\text{س}} - \text{ص}_{\frac{\text{أ}}{٢}} \frac{\sigma}{\sqrt{\text{ن}}} \leqslant \mu \leqslant \bar{\text{س}} + \text{ص}_{\frac{\text{أ}}{٢}} \frac{\sigma}{\sqrt{\text{ن}}}\right) \geqslant ١ - \text{أ}$$

وتكون درجة الثقة مساوية ١٠٠ (١ - أ)٪ لمتوسط المجتمع تو واقع داخل هذه الفترة. وتكون أ عبارة عن قيم موجبة صغيرة مثل ٠,٠١، ٠,٠٥ .... وهكذا وتسمى أ بمستوى المعنوية لفترة الثقة. والقيم ص هي قيم معيارية تحدد من الجداول الإحصائية للتوزيع المعتدل المعياري (القياسي) لكل قيمة من قيم أ ويكتب تقدير فترة الثقة للأوساط المعاينة بصفة عامة كالتالي:

$$\bar{\text{س}} \pm \text{ص}_{\frac{\text{أ}}{٢}} \frac{\sigma}{\sqrt{\text{ن}}}$$

وذلك عندما يكون الانحراف المعياري للمجتمع (σ) معلوم ونوضح طريقة الحساب بالمثال التالي.

مثـال (٥)

إذا كان متوسط طول عينة مكونة من ٤٠٠ ورقة من نبات الغار هو ١٤٢ملم. وكان معلومًا أن الانحراف المعياري لأوراق الغار هو ١٦ملم فأوجد ٩٥٪، ٩٩٪ حدود ثقة لمتوسط طول أوراق نبات الغار.

الحـــل

يمكن حساب ٩٥٪ حدود ثقة لمتوسط أطوال أوراق نبات الغار كالتالي:

حيث أن $\bar{س}$= ١٤٢ ملم ، ن = ٤٠٠ ورقة ، σ= ١٦ ملم، أ = ٠,٠٥ ،

$ص_{٠,٠٢٥}$ = ١,٩٦ إذًا.

$$\bar{س} \pm ص_{\frac{أ}{٢}} \frac{\sigma}{\sqrt{ن}}$$

$$= ١٤٢ \pm ١,٩٦ \frac{١٦}{\sqrt{٤٠٠}}$$

$$= ١٤٢ \pm ١,٥٧$$

أي أن

تقدير الحد الأعلى = ١٤٣,٥٧ ملم

تقدير الحد الأدنى = ١٤٠,٤٣ ملم

ولإيجاد ٩٩٪ حدود ثقة لمتوسط طول أوراق نبات الغار نكتب ما يلي

أ = ٠,٠١ ، $ص_{\frac{أ}{٢}}$ = ٢,٥٨

إذًا $$\bar{س} \pm ٢,٥٨ \frac{\sigma}{\sqrt{ن}}$$

$$= ١٤٢ \pm ٢;٥٨ \frac{١٦}{\sqrt{٤٠٠}}$$

$$= ١٤٢ \pm ٢,٠٦$$

أي أن

تقدير الحد الأعلى = ١٤٤,٠٦ ملم

تقدير الحد الأدنى = ١٣٩,٩٤ ملم

(١١ ـ ٥ ـ ٢) تقدير فترة الثقة للنسبة ح

سبق أن ذكرنا أن توزيع المعاينة للنسب يقترب من التوزيع المعتدل بتوقع $\mu(\text{ح}_{\text{د}}) = \text{ح}$ والخطأ المعياري $\sigma_{\text{ح}_{\text{د}}}$ يكون أقل من الانحراف المعياري للمجتمع σ ، وبذلك تكون تقدير فترة الثقة للنسبة ح عند مستوى معنوية أ بدرجة ثقة قدرها ١٠٠(١ ـ أ)٪ تعطى كالتالي:

$$\text{تقدير فترة الثقة للنسبة ح} = \text{ح}_{\text{د}} \pm \text{ص}_{\frac{\text{أ}}{2}}\ \sigma(\text{ح}_{\text{د}})$$

$$\because \ \sigma(\text{ح}_{\text{د}}) = \sqrt{\frac{\text{ح}_{\text{د}}(1-\text{ح}_{\text{د}})}{\text{ن}}}$$

∴ تقدير فترة الثقة للنسبة ح هو:

$$\text{ح}_{\text{د}} \pm \text{ص}_{\frac{\text{أ}}{2}} \sqrt{\frac{\text{ح}_{\text{د}}(1-\text{ح}_{\text{د}})}{\text{ن}}}$$

$$\text{ح}_{\text{د}} - \text{ص}_{\frac{\text{أ}}{2}} \sqrt{\frac{\text{ح}_{\text{د}}(1-\text{ح}_{\text{د}})}{\text{ن}}} \leqslant \text{ح} \leqslant \text{ح}_{\text{د}} + \text{ص}_{\frac{\text{أ}}{2}} \sqrt{\frac{\text{ح}_{\text{د}}(1-\text{ح}_{\text{د}})}{\text{ن}}}$$

ونوضح طريقة حساب تقدير فترة الثقة للنسب بالمثال التالي.

مثال (٦)

في دراسة بالعينة لأعداد المدخنين في إحدى الجامعات. أخذت عينة مكونة من ١٠٠ طالب، وجد أن ١٨٪ منهم يدخنون.

احسب تقدير حدود الثقة لنسبة المدخنين من طلاب الجامعة عند درجة ثقة تساوي ٩٥٪، ٩٩٪.

الحل

لحساب ٩٥٪ حدود ثقة لنسبة المدخنين من طلاب الجامعة نجد أن

$$\text{ح}_{\text{د}} = ٠,١٨ \quad ، \quad \text{ص}_{٠,٠٢٥} = ١,٩٦ \quad ، \quad \text{ن} = ١٠٠ \ \text{طالب}$$

ومن ذلك نكتب

$$\text{ح}_{\text{د}} \pm ١,٩٦ \sqrt{\frac{\text{ح}_{\text{د}}(١-\text{ح}_{\text{د}})}{\text{ن}}}$$

$$= ٠,١٨ \pm ١,٩٦ \sqrt{\frac{٠,١٨ \times ٠,٨٢}{١٠٠}}$$

$$= ٠,١٨ \pm ١,٩٦ \times ٠,٠٣٨$$

$$= ٠,١٨ \pm ٠,٠٧$$

ومن ذلك يكون

تقدير الحد الأعلى = ٠,٢٥ أي ٢٥٪ من الطلاب

تقدير الحد الأدنى = ٠,١١ أي ١١٪ من الطلاب

ولحساب ٩٩٪ حدوداً لنسبة المدخنين من طلاب الجامعة حيث $\text{ص}_{٠,٠٠٥} = ٢,٥٨$ تكون

$$\text{ح}_{\text{د}} \pm ٢,٥٨ \sqrt{\frac{\text{ح}_{\text{د}}(١-\text{ح}_{\text{د}})}{\text{ن}}}$$

$$= ٠,١٨ \pm ٢,٥٨ \sqrt{\frac{٠,١٨ \times ٠,٨٢}{١٠٠}}$$

$$= ٠,١٨ \pm ٢,٥٨ \times ٠,٠٣٨$$

$$= ٠,١٨ \pm ٠,٠٩٨$$

أي أن

تقدير الحد الأعلى = ٠,١٨ + ٠,٠٩٨ = ٠,٢٧٨ أي ٢٧,٨٪ من الطلاب

تقدير الحد الأعلى = ٠,١٨ − ٠,٠٩٨ = ٠,٠٨٢ أي ٨,٢٪ من الطلاب

(١١ ـ ٥ ـ ٣) : تقدير فترة الثقة بين متوسطين ($\mu_1 - \mu_2$)

سبق أن تكلمنا عن توزيع المعاينة للفرق بين متوسطين ووجدنا أنه يقترب من التـوزيع المعتدل، وأن التوقع μ ($\bar{س}_1 - \bar{س}_2$) = $\mu_1 - \mu_2$، وأن الخطأ المعياري له

$$\sigma\,(\bar{س}_1 - \bar{س}_2) = \sqrt{\frac{\sigma_1^2}{ن_1} + \frac{\sigma_2^2}{ن_2}}$$

وبذلك يمكن تقدير حدود الثقة للفرق بين متوسطين عند مستوى معنوية أ بدرجة ثقة ١٠٠ (١ - أ)٪ كالتالي :

$$(\bar{س}_1 - \bar{س}_2) \pm ص_{\frac{أ}{2}} \sqrt{\frac{\sigma_1^2}{ن_1} + \frac{\sigma_2^2}{ن_2}}$$

ولتوضيح ذلك نورد المثال التالي.

مثـال (٧)

لدراسـة إنتاجية الأرض للقمح في المناطق المختلفة بالمملكة. أخذنا منطقتين مختلفتين فوجدنا أن متوسط انتاج الفدان الواحد من القمح هو ١٤٠٠ صاع للمنطقة الأولى لعينة مكونة من ١٥٠ فدانًا. وأن متوسط إنتاج الفدان من القمح هو ١٢٠٠ صاع للمنطقة الثانية لعينة مكونة من ١٠٠ فدان بانحراف معياري قدره ٩٠ صاعًا للمنطقة الأولى، وانحراف معياري قدره ٥٠ صاعًا للمنطقة الثانية.

احسب ٩٥٪ حدود ثقة للفرق بين متوسطي انتاج القمح للمنطقتين.

الحـــل

لحساب ٩٥٪ حدود ثقة للفرق بين متوسطي الإنتاج من القمح تعطى كالتالي

حدود الثقة = $(\bar{\text{س}}_{1} - \bar{\text{س}}_{2}) \pm \text{١,٩٦}\sqrt{\frac{\sigma^2}{\text{ن}_{1}} + \frac{\sigma^2}{\text{ن}_{2}}}$

حيث إن

$\bar{\text{س}}_{1}$ = ١٤٠٠ صاع ، $\bar{\text{س}}_{2}$ = ١٢٠٠ صاع

σ = ٩٠ صاعا ، σ = ٥٠ صاعا

ن_{1} = ١٥٠ فدانا ، ن_{2} = ١٠٠ فدان

ومن ذلك نجد أن :

حدود الثقة = $(\text{١٤٠٠} - \text{١٢٠٠}) \pm \text{١,٩٦}\sqrt{\frac{(\text{٩٠})^2}{\text{١٥٠}} + \frac{(\text{٥٠})^2}{\text{١٠٠}}}$

$= \text{٢٠٠} \pm \text{١,٩٦}\sqrt{\text{٥٤} + \text{٢٥}}$

$= \text{٢٠٠} \pm \text{١,٩٦} \times \text{٨,٨٩}$

$= \text{٢٠٠} \pm \text{١٧,٤٢}$

وبذلك يكون

تقدير الحد الأعلى لفرق متوسطي الإنتاج = ٢١٧,٤٢ صاعا

تقدير الحد الأدنى لفرق متوسطي الإنتاج = ١٨٢,٥٨ صاعا

(١١ ـ ٥ ـ ٤): تقدير فترة الثقة للفرق بين نسبتين ($\text{ح}_{1} - \text{ح}_{2}$)

سبق أن علمنا أن توزيع المعاينة للفرق بين نسبتين يقترب من التوزيع المعتدل بمتوسط ($\text{ح}_{1} - \text{ح}_{2}$) وخطأ معياري يساوي

$$\sigma(\text{ح}_{\text{د}1} - \text{ح}_{\text{د}2}) = \sqrt{\frac{\text{ح}_{\text{د}1}(1-\text{ح}_{\text{د}1})}{\text{ن}_{1}} + \frac{\text{ح}_{\text{د}2}(1-\text{ح}_{\text{د}2})}{\text{ن}_{2}}}$$

وبذلك يمكن تقدير فترة الثقة للفرق بين نسبتين ($\text{ح}_{1} - \text{ح}_{2}$) عند مستوى معنوية أ وبدرجة ثقة قدرها ١٠٠ (١ - أ)٪ كالتالي :

$$\text{فترة الثقة} = (\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢}) \pm \text{ص}_{\frac{أ}{٢}} \sqrt{\frac{\text{ح}_{\text{د}١}(١-\text{ح}_{\text{د}١})}{\text{ن}_١} + \frac{\text{ح}_{\text{د}٢}(١-\text{ح}_{\text{د}٢})}{\text{ن}_٢}}$$

ويمكن توضيح طريقة الحساب لهذه الفترة بالمثال التالي.

مثال (٨)

في دراسة لمعرفة الفرق بين نسبتين لوجود عقار «البنادول» في الوصفات التي تعطى في مستشفيين في منطقتين مختلفتين بالمملكة وجد في عينة مكونة من ٤٠٠ وصفة للمستشفى الأول ٢٠٠ وصفة تحتوى على عقار «البنادول». كما وجد في عينة مكونة من ٥٠٠ وصفة للمستشفى الثاني ١٠٠ وصفة تحتوي على عقار «البنادول».

احسب ٩٩٪ حدود الثقة للفرق بين نسبتي «البنادول» في وصفات كل من المستشفيين.

الحـــل

لحساب ٩٩٪ حدود الثقة للفرق بين نسبتي «البنادول» كالتالي:

$$\text{حدود الثقة} = (\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢}) \pm ٢,٥٨ \sqrt{\frac{\text{ح}_{\text{د}١}(١-\text{ح}_{\text{د}١})}{\text{ن}_١} + \frac{\text{ح}_{\text{د}٢}(١-\text{ح}_{\text{د}٢})}{\text{ن}_٢}}$$

حيث إن

$$\text{ح}_{\text{د}١} = \frac{٢٠٠}{٤٠٠} = ٠,٥ \quad ، \quad \text{ن}_١ = ٤٠٠$$

$$\text{ح}_{\text{د}٢} = \frac{١٠٠}{٥٠٠} = ٠,٢ \quad ، \quad \text{ن}_٢ = ٥٠٠$$

$$\therefore \text{حدود الثقة} = (٠,٥ - ٠,٢) \pm ٢,٥٨ \sqrt{\frac{٠,٥(١-٠,٥)}{٤٠٠} + \frac{٠,٢(١-٠,٢)}{٥٠٠}}$$

$$= ٠,٣ \pm ٢,٥٨ \sqrt{٠,٠٠٠٦ + ٠,٠٠٠٣}$$

$$= ٠,٣ \pm ٢,٥٨ \times ٠,٠٣$$

$$= ٠,٣ \pm ٠,٠٧٧$$

وبذلك نجد أن

تقدير الحد الأعلى للفرق بين النسبتين = ٣٧٧, ٠ أي ٣٧,٧٪

تقدير الحد الأدنى للفرق بين النسبتين = ٢٢٣, ٠ أي ٢٢,٣٪

(١١ ـ ٦) اختبـارات الفـروض

بعد أن أوضحنا توزيع المعاينة وتقدير حدود الثقة للأوساط والنسبة والفروق بين الأوسـاط والفروق بين النسب فقد حان الوقت لدراسة موضوع اختبارات الفروض الذي يعتبر الموضوع الأساسي الثاني في الإحصاء.

وتعتبر اختبارات الفروض محاولة إلى الوصول لقرار معين سواء كان بالرفض أو القبول لغرض معين متعلق بإحدى معالم المجتمع مثل النسبة ح في حالة ما إذا كان المجتمع يتبع توزيع ذي الحدين أو μ، σ إذا ما كان المجتمع يتبع التوزيع المعتدل. ويمكن مقارنة القيمة المفروضة لمعلمة المجتمع بالقيمة المقدرة لهذه المعلمة من العينة، فإذا ما كانت الفروق صغيرة (مقارنة بقيم القراءات) فإنها تعزى إلى الصدفة وتسمى فروق غير معنوية، وإذا ما كانت الفروق كبيرة (مقارنة بقيم القراءات) فإنه يمكن أن تسمى فروقًا حقيقية أو معنوية. وبذلك يمكن تسمية اختبارات الفروض بالاختبارات المعنوية. فعلى سبيل المثال إذا أخذنا عينة مكونة من ١٦ طالبًا من طلاب جامعة الملك سعود، وحسبنا متوسط الوزن للعينة $\bar{س}$، وكانت قيمة $\bar{س}$ تساوي ٦٥ كجم وإذا افترض الباحث أن متوسط الوزن لمجتمع الطلاب في الجامعة تو يساوي ٦٨ كجم، وأن الانحـراف المعياري نحـر لوزن مجتمـع الـطلاب معـروف ويساوي ٦ كجم. والمطلوب معرفة صحة اختبار الفرض بأن $\mu = ٦٨$ أو لا، وسوف نرمز له فر، ويسمى الفرضية الأولية أو فرض العدم. حيث نفترض عدم وجود فرق حقيقي بين متوسط المجتمـع الحقيقي والقيمـة المفـروضـة وأن الفروق المشاهدة في العينة إنما تعزى إلى الصدفة. ولاختبار ذلك نكوّن إحصائية يكون توزيعها معروفًا حتى نتمكن من اتخاذ قرار بشأن الفرض فر من حيث القبول أو الرفض. والإحصائية التالية

$$ص = \frac{\bar{س} - \mu}{\sigma(\bar{س})} \quad \ldots\ldots\ldots\ldots\ldots \quad (٨)$$

بالنسبة لأوزان عينة من طلاب الجامعة التي سبق ذكرها. نعلم مما سبق أن هذه الإحصائية تتبع التوزيع المعتدل القياسي، ويمكن أن نقسم مجال تغير هذه الإحصائية (ص) السابقة إلى منطقتين، المنطقة الأولى تسمى منطقة القبول وهي التي يكون فيها مقدار احتمال حدوث الإحصائية (ص) كبيراً عندما يكون اختبار الفرض (فر$_{٠}$) صحيحًا، والمنطقة الثانية تسمى منطقة الرفض للفرض فر$_{٠}$، وذلك عندما يكون احتمال حدوث الإحصائية (ص) قليلاً أو نادر الحدوث عندما يكون الفرض فر$_{٠}$ صحيحًا وأي فرض آخر يختلف عن فر$_{٠}$ يسمى الفرض البديل، ويرمز له بالرمز فر$_{١}$ وفي حالة أوزان طلاب الجامعة يمكن أن يكون الفرض البديل (فر$_{١}$) أحد الفروض التالية

$\mu \neq ٦٨$ أو $٦٨ > \mu$ أو $٦٨ < \mu$.

(١١ ـ ٦ ـ ١): الخطأ من النوع الأول والخطأ من النوع الثاني

ينتج عن أي قرار إحصائي نوعان من الخطأ. فإذا رفضنا الفرضية الأولية أو فرض العدم (فر$_{٠}$). وكان صحيحًا نكون قد ارتكبنا خطأ من النوع الأول باحتمال قدره أ وأحيانًا تسمى أ مستوى المعنوية وتأخذ قيمًا صغيرة مثل ٠,٠١ أو ٠,٠٥ أو ٠,١ .... وبناء على قيمة أ وتوزيع الإحصائية ص يمكن أن نحدد منطقتي الرفض والقبول لفرض العدم (فر$_{٠}$). أما الخطأ من النوع الثاني فهو أن نقبل فرض العدم (فر$_{٠}$) عندما يكون غير صحيح، ونرمز لهذا النوع من الخطأ بالرمز ب.

يمكن توضيح منطقتي الرفض والقبول لفرض العدم (فر$_{٠}$) حسب نوع الفرض البديل (فر$_{١}$) ويمكن توضيح ذلك باستخدام المتوسط (تو) كالتالي:

أ ـ الفرض البديل بطرفين

عندما يكون فرض العدم فر$_{٠}$ والفرض البديل فر$_{١}$ على الصورة التالية

فر$_{٠}$: $\mu = \mu$ ،

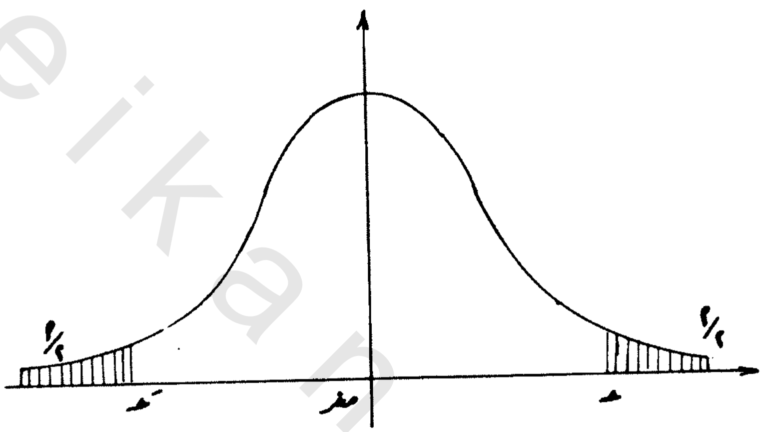
فر$_{١}$: $\mu \neq \mu$

وفي مثال أوزان الطلاب يكون

فر$_0$: μ = ٦٨ ،

فر$_1$: $\mu \neq$ ٦٨

فإنه يكون نصف الخطأ من النوع الأول على طرفي التوزيع للإحصائية ص كما هو موضح بالجزء المظلل في الرسم التالي



شكل (١١ ـ ٢) : منطقة الرفض لفرض العدم من الطرفين

ويكون الاحتمال ح (ص ≤ − جـ) = ح (ص > جـ) = $\frac{أ}{٢}$

ب ـ الفرض البديل ذو الطرف الأعلى

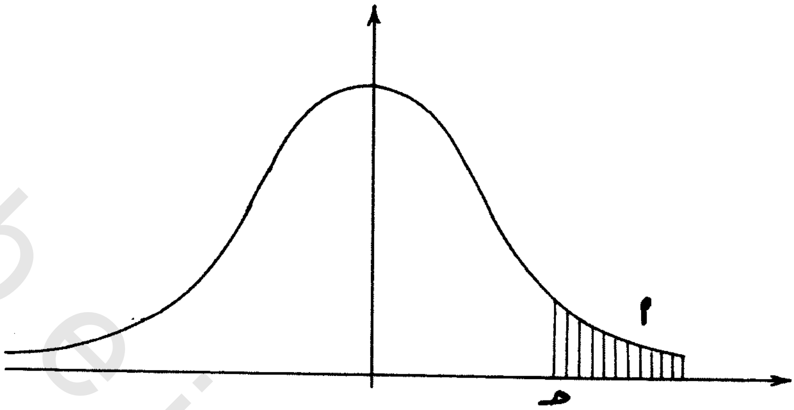
وهو عندما يكون فر$_0$ ، فر$_1$ على الصورة التالية :

فر$_0$: $\mu = \mu_0$، فر$_1$: $\mu > \mu_0$

وفي مثال أوزان الطلاب

فر$_0$: μ = ٦٨ ، فر$_1$: $\mu >$ ٦٨

فإن الخطأ من النوع الأول أ يكون على الطرف الأعلى لتوزيع الإحصائية ص كما هو موضح بالجزء المظلل في الرسم التالي



شكل (١١ ـ ٣): منطقة الرفض لفرض العدم من الطرف الأيمن

ويكون الاحتمال ح (ص < جـ) = أ

جـ ـ الفرض البديل بالطرف الأدنى

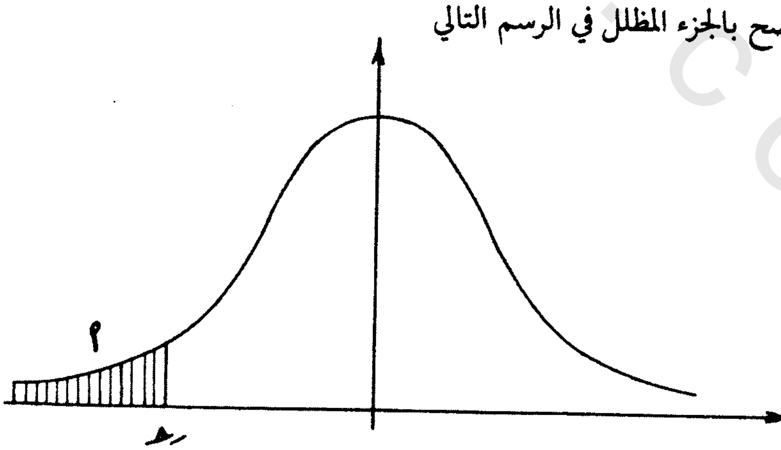
وهو عندما يكون فر$_0$ ، فر$_1$ على الصورة التالية:

فر$_0$: $\mu = \mu_0$ ، فر$_1$: $\mu < \mu_0$

وفي مثال أوزان الطلاب يكون

فر$_0$: μ = ٦٨ ، فر$_1$: μ < ٦٨

ويكون الخطأ من النوع الأول أ عند الطرف الأدنى لتوزيع الإحصائية ص كما هو موضح بالجزء المظلل في الرسم التالي



شكل (١١ ـ ٤): منطقة الرفض لفرض العدم من الطرف الأيسر

ويكون الاحتمال ح (ص < -جـ) = أ

والقيم جـ ، -جـ تسمى الحدود الحرجة التي تفصل بين منطقتي الرفض والقبول لفرض العدم فر. عندما يكون صحيحًا. وعندما أ = ٠,٠٥ فإن القيم الحرجة للإحصائية ص عندما يكون الفرض البديل ذا طرفين تكون جـ = ١,٩٦ ، -جـ = -١,٩٦ وعندما يكون الفرض البديل ذا طرف واحد أعلى أو أدنى فإن قيمة جـ = ١,٦٤٥ أو -جـ = -١,٦٤٥ وحيث إن الإحصائية (٨) تتبع التوزيع الطبيعي القياسي، وتقع قيم ص المحسوبة من (٨) خارج منطقة القبول أي خارج الحدود ١,٩٦ ، -١,٩٦ . عندما يكون الفرض البديل فر١ من طرفين فإننا نرفض فر. عند مستوى معنوية ٠,٠٥ أو ٥٪ أي أنه يكون هناك ٥ فرص في كل ١٠٠ فرصة، إننا سوف نرفض الفرض فر. عندما يكون صحيحًا. وإننا سنكون واثقين بنسبة ٩٥٪ أننا سنتخذ القرار الصحيح . وتسمى النسبة ٩٥٪ بدرجة الثقة في اتخاذ القرار الصحيح . وعندما يكون الفرض البديل ذا طرف واحد هو الطرف الأعلى وتقع قيم ص المحسوبة من (٨) خارج منطقة القبول أي ص > ١,٦٤٥ فإننا نرفض فر. بمستوى معنوية قدره ٠,٠٥ وعندما يكون الفرض البديل ذا طرف واحد هو الطرف الأدنى وتقع قيم ص المحسوبة من (٨) خارج منطقة القبول أي أن ص < -١,٦٤٥ فإننا نرفض فر. بمستوى معنوية قدره ٠,٠٥ . وبالمثل عندما تكون أ = ٠,٠١ من طرفين أو من طرف واحد، والجدول التالي يبين قيم ص الحرجة لبعض قيم (أ) أي بعض مستويات المعنوية الأخرى، ويمكن الحصول عليها من الجداول الإحصائية للتوزيع الطبيعي القياسي في نهاية الكتاب [جدول رقم (٢)].

المستويات المعنوية وقيم ص الحرجة للفرض البديل من طرف واحد وطرفين

مستوى المعنوية أ	٠,١	٠,٠٥	٠,٠١	٠,٠٠٥	٠,٠٠٢
قيم ص الحرجة للفرض البديل من طرف واحد	± ١,٢٨	± ١,٦٤٥	± ٢,٣٣	± ٢,٥٨	± ٢,٨٨
قيم ص الحرجة للفرض البديل من طرفين	± ١,٦٤٥	± ١,٩٦	± ٢,٥٨	± ٢,٨١	± ٣,٠٨

ولاختبار الفرض فر. : $\mu = ٦٨$ لمتوسط أوزان طلاب جامعة الملك سعود ضد الفرض البديل فر١ : $\mu \neq ٦٨$ عند مستوى معنوية أ = ٠,٠٥ ، أ = ٠,٠١ فإننا نحسب الإحصائية ص = $\frac{\bar{\text{س}} - \mu}{\sigma\,(\bar{\text{س}})}$ كالتالي :

$$\bar{\text{س}} = ٦٥ \quad ، \quad \mu = ٦٨ \quad ، \quad \sigma\,(\bar{\text{س}}) = \frac{\sigma}{\sqrt{\text{ن}}} = \frac{٦}{\sqrt{١٦}} = ١,٥$$

$$\text{ص} = \frac{٦٥ - ٦٨}{١,٥} = \frac{-٣}{١,٥} = -٢$$

إذا كانت منطقة الرفض للفرض فر. عند مستوى معنوية ٠,٠٥ خارج حدود ± ١,٩٦ وحيث إن قيمة ص المحسوبة = -٢ تقع في منطقة الرفض فإننا نرفض فر. عند مستوى معنوية ٠,٠٥ أو درجة ثقة قدرها ٩٥٪ وبالتالي يكون قرار الرفض صحيح .

وإذا كانت منطقة الرفض لـ فر. عند مستوى معنوية ٠,٠١ خارج حدود ± ٢,٥٨ حيث إن القيمة ص المحسوبة = -٢ تقع داخل منطقة القبول للفرض فر. . أي أننا لا نستطيع رفض فر. عند مستوى معنوية ٠,٠١ وبعبارة أخرى يمكن القول إن قرار عدم الرفض صحيح بدرجة ثقة قدرها ٩٩٪ .

وعندما تقع قيمة ص المحسوبة بين حدي الرفض لمستوى معنوية ٠,٠٥ ، ٠,٠١ يقال لهذه القيمة إنها محتملة المعنوية ، وعندما تقع قيمة ص المحسوبة خارج حدي الرفض لمستوى المعنوية ٠,٠٥ ، ٠,٠١ يقال لهذه القيمة إنها مرتفعة المعنوية .

وسوف نوضح فيما يلي اختبارات الفروض للإحصائية ص عندما تمثل الأوساط أو النسب أو الفروق بين الأوساط أو الفروق بين النسب وذلك عندما يكون تباين المجتمعات معلوما .

(١١ ـ ٦ ـ ٢) اختبارات الفروض للأوساط

عندما يكون تباين المجتمع (σ^2) معلومًا فإن الإحصائية $\dfrac{\bar{س} - \mu}{\sigma\,(\bar{س})}$ تقترب من التوزيع الطبيعي القياسي بمتوسط = صفر وانحراف معياري = ١ ونرمز لهذه القيم بالرمز ص حيث ص = $\dfrac{\bar{س} - \mu}{\sigma_{\bar{س}}}$

ونبني اختبار الفروض فر$_0$ ، فر$_1$ كالتالي:

فر$_0$: $\mu = \mu_0$ ، فر$_1$: $\mu \neq \mu_0$ عندما يكون الفرض البديل ذا طرفين

فر$_0$: $\mu = \mu_0$ ، فر$_1$: $\mu \gtrless \mu_0$ عندما يكون الفرض البديل ذا طرف واحد

ونختار مستوى المعنوية للاختبار (أ) ونختبر ما إذا كانت قيمة الإحصائية ص المحسوبة واقعة داخل منطقة الرفض أو خارجها فإنه يمكن أن نقرر رفضها أو عدم رفضها، ونوضح ذلك بالمثال التالي.

مثـال (٩)

أخذت عينة مكونة من ٣٦ عجلاً من مزرعة لتسمين الماشية فوجد أن متوسط الوزن للعينة $\bar{س}$ = ١٩٠ كجم.

اختبر الفرض القائل: إن متوسط العجول بالمزرعة μ = ٢٠٠ كجم، إذا علم أن الانحراف المعياري لأوزان العجول بالمزرعة يساوي ١٨ كجم.

الحـل

ولحل المثال نكون فرض العدم (فر$_0$) والفرض البديل (فر$_1$) على الصورة التالية

فر$_0$: μ = ٢٠٠ كجم ، فر$_1$: $\mu \neq$ ٢٠٠ كجم

أي أن الفرض البديل ذو طرفين وتكون الحدود الحرجة للاختبار عند مستوى المعنوية ٠,٠٥ ، ٠,٠١ هي (١,٩٦ ، −١,٩٦)، (٢,٥٨ ، −٢,٥٨) على الترتيب.

$\because \ \bar{\text{س}} = ١٩٠$ كجم ، $\mu = ٢٠٠$ كجم ، ن = ٣٦ عجلاً

$$\therefore \ \text{ص} = \frac{\bar{\text{س}} - \mu}{\sigma_{\bar{\text{س}}}} = \frac{١٩٠ - ٢٠٠}{\frac{١٨}{\sqrt{٣٦}}} = \frac{-١٠}{\frac{١٨}{٦}}$$

$$= \frac{-١٠}{٣} = -٣,٣٣$$

ومما سبق نلاحظ أن قيمة ص المحسوبة = -٣,٣٣ واقعة خارج حدود القيم الحرجة لكل من مستويي المعنوية ٠,٠٥ ، ٠,٠١ أي أننا نرفض الفرض فر. عند مستوى المعنوية ٠,٠٥ ، ٠,٠١ أي أن قيمة ص = -٣,٣٣ عالية المعنوية .

(١١ ـ ٦ ـ ٣) اختبارات الفروض للنسبة

الإحصائية ص $= \frac{\text{ح}_{\text{د}} - \text{ح}_{.}}{\sigma\,(\text{ح}_{\text{د}})}$ تقترب من التوزيع الطبيعي القياسي وبذلك يمكن أن نكوّن اختبار الفروض لكل من فر. ، فر١ كالتالي

فر. : ح = ح. ،

فر١ : ح $\neq$ ح. عندما يكون الفرض البديل ذا طرفين ،

فر. : ح = ح. ،

فر١ : ح $\gtrless$ ح. عندما يكون الفرض البديل ذا طرف واحد ،

ونختار مستوى المعنوية أ ونختبر الفرض . إذا كانت قيمة الإحصائية المحسوبة ص واقعة داخل منطقة رفض فر. أو خارجها ، ويمكن توضيح ذلك بالمثال التالي .

مثـال (١٠)

إذا أخذت عينة مكونة من ٢٠٠ وصفة طبية بإحدى المستشفيات وجد منها ٨٠ وصفة تحتوي على عقار «البنادول» . فاختبر الفرض القائل : إن نسبة الوصفات التي بها عقار «البنادول» هي ٥٠٪ . وذلك باستخدام ٠,٠١ مستوى معنوية .

الحـــل

نكوّن أولاً اختبار الفروض فر. ، فر١ كما يلي:

$$\text{فر}_. : ح = ٠,٥ ، \text{فر}_١ : ح \neq ٠,٥$$

أي أن الفرض البديل بطرفين، فتكون الإحصائية ص كالتالي

$$ص = \frac{ح_د - ح_.}{\sigma (ح_.)}$$

$$\text{حيث } \sigma (ح_.) = \sqrt{\frac{ح_. (١ - ح_.)}{ن}}$$

أي أن

$$ح_د = \frac{٨٠}{٢٠٠} = ٠,٤ ، ح_. = ٠,٥$$

ومن ذلك نجد أن

$$ص = \frac{٠,٤ - ٠,٥}{\sqrt{\frac{٠,٥ (١ - ٠,٥)}{٢٠٠}}} = \frac{-٠,١}{\sqrt{\frac{٠,٢٥}{٢٠٠}}}$$

$$= \frac{-٠,١}{٠,٠٣٥} = -٢٨,٦$$

منطقة الرفض للفرض عند مستوى معنوية ٠,٠١ أي خارج الحدود الحرجة ٢,٥٨ ، -٢,٥٨ وقيمة ص المحسوبة = -٢٨,٦ واقعة في منطقة رفض فر. .

مما سبق نرفض الفرض القائل إن نسبة الوصفات الطبية التي بها عقار «البنادول» هي ٥٠٪ وذلك عند مستوى معنوية ٠,٠١ .

(١١ ـ ٦ ـ ٤) اختبارات الفروض بين الأوساط

سبق أن أوضحنا أن توزيع المعاينة للفروض بين الأوساط ($\bar{س}_١ - \bar{س}_٢$) يقترب

من التوزيع المعتدل بتوقع μ (س̄١ - س̄٢) = $\mu_1 - \mu_2$ ، وخطأ معياري

σ (س̄١ - س̄٢) = $\sqrt{\frac{\sigma^2}{ن_1} + \frac{\sigma^2}{ن_2}}$ وبذلك تكون الاحصائية

ص = $\frac{(س̄_1 - س̄_2) - (\mu_1 - \mu_2)}{\sigma\,(س̄_1 - س̄_2)}$ لها توزيع معتدل قياسي بمتوسط = صفر

وتباين = ١ وإننا نرغب في تكوين اختبار الفروض فر. ، فر١ عند مستوى معنوية أ كالتالي

فر. : $\mu_1 = \mu_2$ ، μ_1 : $\mu_1 \neq \mu_2$ الفرض البديل ذو طرفين

فر. : $\mu_1 = \mu_2$ ، μ_1 : $\mu_1 \gtrless \mu_2$ الفرض البديل ذو طرف واحد

وباعتبار صحة الفرض فر. فإن القيمة $\mu_1 - \mu_2$ = صفر أي أن العينتين اللتين متوسطهما س̄١ ، س̄٢ مسحوبتان من مجتمعتين لهما نفس الوسط الحسابي، ويمكن كتابة الإحصائية ص السابقة كالتالي:

$$ص = \frac{س̄_1 - س̄_2}{\sqrt{\frac{\sigma^2}{ن_1} + \frac{\sigma^2}{ن_2}}}$$

وبذلك يمكن اختبار فرض العدم (فر.) ضد الفرض البديل (فر١)، وذلك عند مستوى معنوية مناسب ولتوضيح كيفية ذلك ندرس حل المثال التالي.

مثال (١١)

أعطي اختبار لفصلين يتكون الفصل الأول من ٥٠ طالبًا ويتكون الفصل الثاني من ٦٠ طالبًا، وكان متوسط الدرجات للفصل الأول ٦٤ درجة بانحراف معياري قدره ٦ درجات. بينما كان متوسط الدرجات للفصل الثاني ٦٦ درجة بانحراف معياري قدره ٥ درجات. اختبر الفرض القائل: إنه لا يوجد اختلاف معنوي في أداء الفصلين، وذلك عند مستوى المعنوية ٠,٠٥ .

الحل

نكوّن صيغة الفروض لـ فر. ، فر١ كالتالي :

فر. : $\mu_1 = \mu_2$ ،

فر١ : $\mu_1 \neq \mu_2$ الفرض البديل ذو طرفين

وباعتبار صحة الفرض فر. فإن الإحصائية ص تعطى كالتالي :

$$ص = \frac{\bar{س}_1 - \bar{س}_2}{\sqrt{\frac{\sigma_1^2}{ن_1} + \frac{\sigma_2^2}{ن_2}}}$$

حيث

$\bar{س}_1$ = ٦٤ درجة ، $\bar{س}_2$ = ٦٦ درجة

σ_1 = ٦ درجات ، σ_2 = ٥ درجات

$ن_1$ = ٥٠ طالبًا ، $ن_2$ = ٦٠ طالبًا

$$\therefore ص = \frac{٦٤ - ٦٦}{\sqrt{\frac{(٦)^2}{٥٠} + \frac{(٥)^2}{٦٠}}} = \frac{-٢}{\sqrt{\frac{٣٦}{٥٠} + \frac{٢٥}{٦٠}}} = \frac{-٢}{\sqrt{٠,٧٢ + ٠,٤٢}}$$

$$= \frac{-٢}{\sqrt{١,١٤}} = \frac{-٢}{١,٠٧} = -١,٨٧$$

مما سبق نجد أن منطقة الرفض للفرض فر. عند مستوى معنوية ٠,٠٥ خارج الحدود الحرجة ١,٩٦ ، -١,٩٦ ونجد أن قيمة ص المحسوبة هي ص = -١,٨٧ عند مستوى معنوية ٠,٠٥ أي لا نستطيع رفض فر. عند معنوية ٠,٠٥ أي لا يوجد فروق معنوية بين أداء الفصلين.

(١١ ـ ٦ ـ ٥) اختبار الفروض للفروق بين النسب

سبق أن تكلمنا عن توزيع المعاينة للفرق بين النسب (ح د١ - ح د٢) بأنه يقترب من التوزيع المعتدل بتوقع μ (ح د١ - ح د٢) = ح١ - ح٢ وخطأ معياري

σ (ح $_{د_١}$ - ح $_{د_٢}$) يساوي $\sqrt{\frac{ح_١(١-ح_١)}{ن_١} + \frac{ح_٢(١-ح_٢)}{ن_٢}}$ فإن الإحصائية

ص = $\frac{(ح_{د_١} - ح_{د_٢}) - (ح_١ - ح_٢)}{\sigma (ح_{د_١} - ح_{د_٢})}$ يكون لها توزيع معتدل قياسي بمتوسط = صفر وتباين = ١ .

والآن نقوم بتكوين اختبار الفروض لكل من فر$_٠$ ، فر$_١$ عند مستوى معنوية أ كالتالي.

فر$_٠$: ح$_١$ = ح$_٢$ ، فر$_١$: ح$_١$ $\neq$ ح$_٢$ الفرض البديل ذو طرفين،

فر$_٠$: ح$_١$ = ح$_٢$ ، فر$_١$: ح$_١$ $\gtrless$ ح$_٢$ الفرض البديل ذو طرف واحد

وباعتبار صحة الفرض فر$_٠$ فإن القيمة ح$_١$ - ح$_٢$ = صفر أي أن العينتين اللتين نسبتهما ح $_{د_١}$ ، ح $_{د_٢}$ مسحوبتان من مجتمعين لهما نفس النسبة، ويمكن كتابة الإحصائية ص السابقة كالتالي:

$$ص = \frac{ح_{د_١} - ح_{د_٢}}{\sqrt{\frac{ح(١-ح)}{ن_١} + \frac{ح(١-ح)}{ن_٢}}}$$

$$حيث\ ح = \frac{ن_١ ح_{د_١} + ن_٢ ح_{د_٢}}{ن_١ + ن_٢}$$

وبذلك يمكن اختبار فرض العدم (فر$_٠$) ضد الفرض البديل (فر$_١$) عند مستوى معنوية مناسب ونوضح ذلك بالمثال التالي.

مثال (١٢)

مجموعتان و ، ب تتكون المجموعة و من ٨٠ مريضًا بمرض معين، والمجموعة ب تتكون من ١٢٠ مريضًا أعطيت المجموعة و مصلاً فشفي منها ٥٠ شخصًا. ولم تعطي المجموعة ب أي مصل فشفي منها ٦٠ شخصًا. اختبر الفرض القائل: إن المصل لا يساعد على الشفاء عند مستوى معنوية ٠,٠٥ .

الحـــل

ولدراسة هذا المثال نكوّن أولاً الصيغ المناسبة للفروض فر. ، فر١ كما يلي :

$\text{فر}_{٠} : \text{ح}_{١} = \text{ح}_{٢}$ ، $\text{فر}_{١} : \text{ح}_{١} > \text{ح}_{٢}$ الاختبار البديل من طرف واحد،

وباعتبار صحة الفرض فر. فإن الإحصائية ص تعطى كالتالي :

$$\text{ص} = \frac{\text{ح}_{\text{د}١} - \text{ح}_{\text{د}٢}}{\sqrt{\frac{\text{ح}(١-\text{ح})}{\text{ن}_{١}} + \frac{\text{ح}(١-\text{ح})}{\text{ن}_{٢}}}} \qquad \text{حيث ح} = \frac{\text{ن}_{١}\,\text{ح}_{\text{د}١} + \text{ن}_{٢}\,\text{ح}_{\text{د}٢}}{\text{ن}_{١} + \text{ن}_{٢}}$$

$$\text{ح}_{\text{د}١} = \frac{٥٠}{٨٠} = \frac{٥}{٨} \quad ، \quad \text{ح}_{\text{د}٢} = \frac{٦٠}{٢٠} = ٠,٥$$

$$\text{ن}_{١} = ٨٠ \quad ، \quad \text{ن}_{٢} = ١٢٠$$

$$\therefore \text{ح} = \frac{٨٠ \times \frac{٥}{٨} + ١٢٠ \times \frac{١}{٢}}{٨٠ + ١٢٠} = \frac{١١٠}{٢٠٠} = ٠,٥٥$$

$$\text{ص} = \frac{\frac{٥}{٨} - \frac{١}{٢}}{\sqrt{\frac{٠,٥٥ \times ٠,٤٥}{٨٠} + \frac{٠,٥٥ \times ٠,٤٥}{١٢٠}}} = \frac{٠,١٢٥}{\sqrt{٠,٠٠٣١ + ٠,٠٠٢١}}$$

$$= \frac{٠,١٢٥}{\sqrt{٠,٠٠٥٢}} = \frac{٠,١٢٥}{٠,٠٧٢} = ١,٧٤$$

مما سبق نجد أن منطقة الرفض للفرض فر. عند مستوى معنوية ٠,٠٥ خارج الحدود الحرجة لمنطقة القبول ١,٩٦ ، -١,٩٦ ، ونجد قيمة ص المحسوبة = ١,٧٤ داخل الحدود الحرجة السابقة أي أننا لا نستطيع رفض الفرض فر. عند مستوى معنوية ٠,٠٥ أي أن الفروق المشاهدة للنسبتين $\text{ح}_{\text{د}١}$ ، $\text{ح}_{\text{د}٢}$ ترجع إلى الصدفة . أي أن المصل غير فعّال .

(١١ ـ ٦ ـ ٦) توزيع الأوساط والفروق بين الأوساط للعينات الصغيرة

افترضنا فيما سبق عند تقدير فترات الثقة واختبارات الفروض لكل من الأوساط والفروق بين الأوساط أن تباين المجتمع (σ^2) معلوم. وإذا لم يكن معلومًا فلابد أن يكون حجم العينات كبيرًا ولا يقل حجم العينة ن عن ٣٠ مفردة حتى يمكننا أن نقدر التبـاين (σ^2) للمجتمع من العينة وليكن (ع^2)، ونستخدم التباين المقدر (ع^2) في الإحصائية ص لكل من الأوساط والفروق بين الأوساط بدلاً من تبا فيقترب توزيع الإحصائية ص من التوزيع المعتدل القياسي. وبذلك نتمكن من حساب حدود الثقة للتقدير بفترة وكذلك إجراء اختبارات الفروض كما سبق.

ولكن عندما يكون الانحراف المعياري للمجتمع (σ) غير معلوم لنا وحجم العينات صغيرًا (أي أن حجم العينة ن يقل عن ٣٠ مفردة) فإن الإحصائيتين السابقتين

$$\frac{\bar{\text{س}} - \mu}{\sigma_{\bar{\text{س}}}} \quad ، \quad \frac{(\bar{\text{س}}_1 - \bar{\text{س}}_2) - (\mu_1 - \mu_2)}{\sigma\,(\bar{\text{س}}_1 - \bar{\text{س}}_2)} \quad ،$$

لا تقتربا من التوزيع المعتدل القياسي ولا نستطيع استخدام الإحصائية ص السابقة لها. وفي هذه الحـالـة فإن الإحصائيات السابقة تتبع توزيعًا آخر يسمى توزيع تي. وتستخدم في هذه الحالة الإحصائية تي التي تقابل الإحصائية ص فيما سبق. وذلك في حالة بيانات العينة مأخوذة من مجتمع طبيعي.

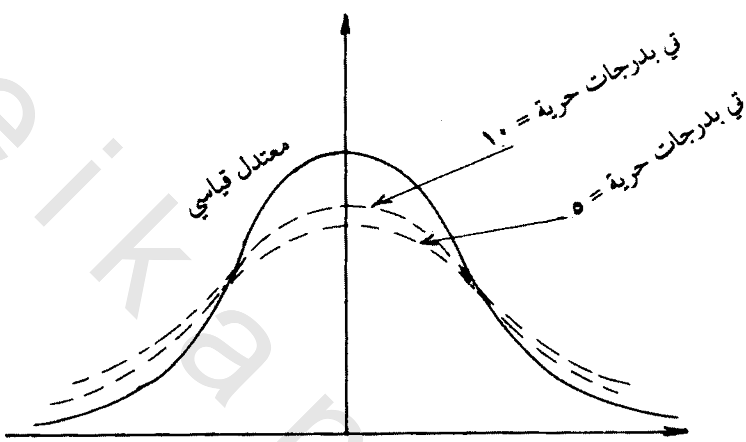
توزيع تـي

تعطى دالة كثافة الاحتمال ح (تي) لتوزيع تي كالتالي:

$$\text{ح}(\text{تي}) = \text{ك}\left[1 + \frac{\text{تي}^2}{\text{ن}-1}\right]^{-\frac{\text{ن}}{2}} \qquad -\infty < \text{تي} < \infty$$

ويسمى المقدار ن ـ ١ درجات الحرية. ولقد تبين أن هذا التوزيع لا يعتمد إلا على قيمة درجات الحرية (ن ـ ١) بشرط أن المتغير العشوائي الأساس س يتوزع توزيعًا معتدلاً. ويقترب توزيع تي من التوزيع المعتدل القياسي كلما زادت درجات الحرية فعندما تصل كمية درجات الحرية ٣٠ درجة فينطبق توزيع تي على التوزيع الطبيعي القياسي

وتستخدم الإحصائية ص في هذه الحالة بدلاً من الإحصائية تي. ويعرف توزيع تي أيضًا **بتوزيع طالب**، وهو اسم مستعار لمكتشفه جوست (Gosset) في أوائل القرن العشرين. والشكل التالي يوضح المقارنة بين منحنى توزيع تي عند درجات حرية ٥، ١٠ ومنحنى توزيع المعتدل القياسي.



شكل (١١ ـ ٥): منحنى توزيع تي مقارنًا بمنحنى التوزيع الطبيعي

ويمكن إيجاد فترات الثقة واختبارات الفروض باستخدام الإحصائية تي لكل من الأوساط والفروق بين الأوساط كما سيتضح من الأمثلة التالية.

مثال (١٣)

إذا كان لدينا عينة مكونة من أوزان عشرة طلاب بالكجم كالتالي:

٧٩، ٦٣، ٦٥، ٧٤، ٦٩، ٧٠، ٧٣، ٦٨، ٧٦، ٦٣

١) اختبر ما إذا كانت العينة مأخوذة من مجتمع متوسطة ٦٦ كجم.

٢) احسب تقديرًا لحدود الثقة لمتوسط المجتمع عند مستوى معنوية ٠,٠٥

الحـــل

ولحل المثال نحسب أولاً الوسط الحسابي ($\bar{س}$) للعينة من العلاقة $\bar{س} = \frac{مجـ\ س}{ن}$

$$\bar{س} = \frac{٧٩ + ٦٣ + ٦٥ + ٧٤ + ٦٩ + ٧٠ + ٧٣ + ٦٨ + ٧٦ + ٦٣}{١٠}$$

$$\bar{س} = \frac{٧٠٠}{٧٠} = ٧٠ \text{ كجم}$$

بعد ذلك نحسب التباين للعينة ($ع^٢$) كتقدير لتباين المجتمع (σ^2) ومنه نحسب الانحراف المعياري (ع) كالتالي:

$$ع^٢ = \frac{١}{ن - ١} \text{مجـ} (س - \bar{س})^٢$$

$$\therefore ع^٢ = \frac{١}{٩} [(٩)^٢ + (-٧)^٢ + (-٥)^٢ + (٤)^٢ + (-١)^٢ + (٠)^٢ + (٣)^٢ + (-٢)^٢ + (٦)^٢ + (-٧)^٢]$$

$$= \frac{١}{٩} [٨١ + ٤٩ + ٢٥ + ١٦ + ١ + ٠ + ٩ + ٤ + ٣٦ + ٤٩]$$

$$= \frac{٢٧٠}{٩} = ٣٠$$

ومنه نجد أن الانحراف المعياري للعينة هو:

$$ع = ٥,٤٨$$

$$ع_{\bar{س}} = \frac{ع}{\sqrt{ن}} = \frac{٥,٤٨}{\sqrt{١٠}} = \frac{٥,٤٨}{٣,١٦} = ١,٧٣$$

نكوّن الاختبار لفرض العدم ($فر_٠$) والفرض البديل ($فر_١$) وليكن عند مستوى معنوية أ = ٠,٠٥ مثلاً. ثم نحسب الإحصائية تي باستخدام مشاهدات العينة والفرض $فر_٠$ كالتالي

$فر_٠ : \mu = ٦٦$ ،

$فر_١ : \mu \neq ٦٦$ الفرض البديل ذو طرفين

$$\therefore تي = \frac{\bar{س} - \mu}{ع_{(\bar{س})}} = \frac{٧٠ - ٦٦}{١,٧٣} = \frac{٤}{١,٧٣} = ٢,٣١$$

نقارن قيمة تي المحسوبة ٢,٣١ بالقيمة الموجودة بجداول تي (جدول رقم ٣) الموجودة بآخر الكتاب أمام درجات حرية = ن - ١ = ١٠ - ١ = ٩ وتحت مستوى المعنوية ٠,٠٥.

قيمة تي من الجدول: تي (أ، ن - ١) = تي (٠,٠٥، ٩) = ٢,٢٦٢.

مما سبق نجد أن قيمة تي = ٢,٣١ واقعة في منطقة الرفض للفرض فر٠ عندما يكون صحيحًا وهي خارج الحدود الحرجة لمنطقة القبول فر٠ وتكون كالتالي (٢,٢٦٢ ، -٢,٢٦٢).

وبذلك نرفض الفرض فر٠ القائل: إن العينة مأخوذة من مجتمع متوسطه ٦٦ كجم وبهذا نكون قد توصلنا إلى حل الفقرة الأولى من المثال.

ولإيجاد تقدير حدود الثقة لمتوسط المجتمع (μ) عند مستوى معنوية ٠,٠٥ نكوّن فترة الثقة كالتالي:

س̄ - تي (٠,٠٥ ، ٩) ع/√ن ≤ μ ≤ س̄ + تي (٠,٠٥ ، ٩) ع/√ن

٧٠ - ٢,٢٦٢ × ١,٧٣ ≤ μ ≤ ٧٠ + ٢,٢٦٢ × ١,٧٣

٧٠ - ٣,٩١ ≤ μ ≤ ٧٠ + ٣,٩١

٦٦,٠٩ ≤ μ ≤ ٧٣,٩١

ومن ذلك نجد أن تقديري حديّ فترة الثقة هما:

الحد الأعلى للوزن = ٧٣,٩١ كجم

والحد الأدنى للوزن = ٦٦,٠٩ كجم

مثـال (١٤)

في أحـد مراكز البحوث الخاصة بالزراعة، كان المطلوب اختبار متوسط انتاج نوعين من القمح و، ى.

فاختير لهذا الغرض ١٨ قطعة من الأرض تتساوي في المساحة والظروف المتشابهة من ناحية الخصوبة والرى والتسميد، وزرع عشر قطع بالقمح من النوع و، وزرعت ثمان القطع الباقية بالقمح من النوع ى، فكان متوسط المحصول لفدان القمح من النوع و ٩٨٠ صاعًا بانحراف معياري هو ٤٠ صاعًا ومتوسط المحصول لفدان القمح من النوع ى ٨٤٠ صاعًا بانحراف معياري هو ٣٠ صاعًا والمطلوب إيجاد ما يلي:

١) اختبر ما إذا كان متوسط الإنتاج للفدان لنوع القمح ويساوي متوسط انتاج الفدان للقمح من النوع ى.

٢) احسب تقديرًا لحدود الثقة للفرق بين متوسطي الإنتاج للقمح من النوع و والنوع ى عند مستوى معنوية ٠,٠١.

الحـــل

لحل الفقرة الأولى من المثال نكوّن أولاً صيغ فرض العدم فر_{0} والفرض البديل فر_{1}، وليكن عند مستوى معنوية ٠,٠١ مثلاً، ثم نحسب الإحصائية تي باستخدام المشاهدات بالعينتين والفرض فر_{0} كالتالي:

$\text{فر}_{0} : \mu_1 = \mu_2$ ، $\text{فر}_{1} : \mu_1 \neq \mu_2$ (أي أن الفرض البديل بطرفين)

$$\text{تي} = \frac{\bar{\text{س}}_1 - \bar{\text{س}}_2}{\text{ع}\sqrt{\frac{1}{\text{ن}_1} + \frac{1}{\text{ن}_2}}} \quad ،$$

$$\text{ع}^2 = \frac{(\text{ن}_1 - 1)\,\text{ع}^2 + (\text{ن}_2 - 1)\,\text{ع}^2}{(\text{ن}_1 + \text{ن}_2 - 2)}$$

وحيث $\bar{\text{س}}_{١}$ = ٩٨٠ صاع ، $\bar{\text{س}}_{٢}$ = ٨٤٠ صاع

$\text{ن}_{١}$ = ١٠ قطع ، $\text{ن}_{٢}$ = ٨ قطع

$\text{ع}_{١}$ = ٤٠ صاعًا ، $\text{ع}_{٢}$ = ٣٠ صاعا ،

$\text{ع}^{٢} = (\text{ع})^{٢}$

$$\therefore \quad \text{ع}^{٢} = \frac{٩ \times ١٦٠٠ + ٧ \times ٩٠٠}{١٠ + ٨ - ٢} = ١٢٩٤,٧٥$$

$$\text{ع} = \sqrt{\text{ع}^{٢}} = \sqrt{١٢٩٤,٧٥}$$

ع = ٣٦ صاعًا تقريبًا

$$\therefore \quad \text{تي} = \frac{٩٨٠ - ٨٤٠}{٣٦\sqrt{\frac{١}{١٠} + \frac{١}{٨}}} = \frac{١٤٠}{٣٦ \times ٠,٤٧} = \frac{١٤٠}{١٦,٩٢} = ٨,٢٧$$

نقارن قيمة تي المحسوبة ٨,٢٧ بالقيمة الموجودة بجداول تي أمام درجات الحرية $\text{ن}_{١} + \text{ن}_{٢} - ٢ = ١٠ + ٨ - ٢ = ١٦$ وتحت مستوى معنوية ٠,٠١ وتكون قيمة تي من الجداول كالتالي:

$$\text{تي}\ (٠,٠١ ،\ ١٦) = ٢,٩٢$$

مما سبق نجد أن قيمة تي المحسوبة = ٨,٢٧ واقعة في منطقة رفض فر وهي خارج الحدود الحرجة لمنطقة القبول لـ فر وهي (٢,٩٢، -٢,٩٢) وبذلك نرفض الفرض فر القائل إنه لا توجد فروق معنوية بين متوسطي إنتاج القمح من النوعين.

ولحل الفقرة الثانية أي لحساب تقدير حدود الثقة للفرق بين متوسطي الإنتاج لنوعي القمح $(\mu_{١} - \mu_{٢})$ عند مستوى معنوية الثقة كالتالي:

$$(\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}) - \text{تي}\ (٠,٠١ ،\ ١٦)\ \text{ع}\ (\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}) \leqslant (\mu_{١} - \mu_{٢})$$

$$\leqslant (\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}) + \text{تي}\ (٠,٠١ ،\ ١٦)\ \text{ع}\ (\bar{\text{س}}_{١} - \bar{\text{س}}_{٢})$$

$$١٤٠ - ٢,٩٢ \times ١٦,٩٢ \leqslant \mu_{١} - \mu_{٢} \leqslant ١٤٠ + ٢,٩٢ \times ١٦,٩٢$$

$$١٤٠ - ٤٩,٤١ \leqslant \mu_{١} - \mu_{٢} \leqslant ١٤٠ + ٤٩,٤١$$

$$٩٠,٥٩ \leqslant \mu_{١} - \mu_{٢} \leqslant ١٨٩,٤١$$

وبذلك يكون تقديري حديَّ فترة الثقة هما:

الحد الأعلى لفرق المتوسطين هو: ١٨٩,٤١ صاع

الحد الأدنى لفرق المتوسطين هو: ٩٠,٠٩ صاعًا

(١١ ـ ٦ ـ ٧) اختبار الفرق بين متوسطي عينتين غير مستقلتين

استخدمنا في الاختبارات السابقة اختبار الفرق بين عينتين مستقلتين بمعنى أن مفردات العينة الأولى مستقلة عن مفردات العينة الثانية. ولكن قد يحدث في الحياة العملية أن المشاهدات تكون على شكل أزواج مرتبة (س، ص). فمثلًا إذا أخذنا عينة وحصلنا على مشاهدات لها في المرة الأولى كالتالي $س_1$، $س_2$، . . . . .، $س_ن$ ثم نضع هذه العينة تحت تأثير مؤثر ثم نعود مرة ثانية، ونحصل على مشاهدات لها مرة أخرى كالتالي $ص_1$، $ص_2$، . . . . .، $ص_ن$ وبإجراء اختبارات للمقارنة بين مجموعتي القراءتين لنفس المفردات في العينة في المرة الأولى والثانية يمكننا استنتاج تأثير المؤثر على مفردات العينة وفي هذه الحالة ندرس الفرق بين مقدارين غير مستقلين عن بعضهما وهما متوسط القراءات الأولى $س_1$، $س_2$، . . . . .، $س_ن$ ومتوسط القراءات الثانية $ص_1$، $ص_2$، . . . . .، $ص_ن$.

ويمكن استخدام اختبار تي في هذه الحالة كما يلي:

١ ـ نوجد الفرق $ف_ر = س_ر - ص_ر$ للازدواج القيم (س، ص).

٢ ـ نحسب الوسط الحسابي $\bar{ف}$ لهذه الفروق.

٣ ـ نوجد الانحراف المعياري $ع_{\bar{ف}}$ لهذه الفروق.

٤ ـ نكوّن الاحصائية تي = $\dfrac{\bar{ف}}{ع_{\bar{ف}} / \sqrt{ن}}$

والإحصائية تي المحسوبة تخضع لاختبار تي (أ، ن ـ ١)، ونوضح ذلك بالمثال التالي.

مثال (١٥)

إذا كانت درجات مجموعة مكونة من ٧ طلاب في مادتي الإحصاء والرياضيات كما في الجدول:

درجات سبعة طلاب في مادتي الإحصاء والرياضيات

رقـم الطالـب	١	٢	٣	٤	٥	٦	٧
درجـة الإحصـاء (س)	٦٢	٨٢	٧٧	٥٧	٦٢	٩٠	٨٢
درجـة الرياضيـات (ص)	٥٣	٧٥	٦٥	٥٥	٦٧	٨٥	٧٩

اوجد كلاً من:

١) اختبر ما إذا كان هناك فروق معنوية بين متوسطي درجات الإحصاء والرياضيات.

٢) اوجد تقدير لحدود الثقة لمتوسط فرق الدرجات للإحصاء والرياضيات بمستوى معنوية ٠٫٠٥.

الحـل

لحساب المتوسط فَ والانحراف المعياري لها نجع نكوّن الجدول التالي:

رقـم الطالـب	درجـة الإحصاء (س)	درجـة الرياضيات (ص)	ف = س ـ ص	ف ـ فَ	(ف ـ فَ)^{2}
١	٦٢	٥٣	٩	٤	١٦
٢	٨٤	٧٥	٩	٤	١٦
٣	٧٧	٦٥	١٢	٧	٤٩
٤	٥٧	٥٥	٢	٣-	٩
٥	٦٢	٦٧	٥-	١٠-	١٠٠
٦	٩٠	٨٥	٥	صفر	صفر
٧	٨٢	٧٩	٣	٢-	٤
المجموع	ـ	ـ	٣٥		١٩٤

$$\bar{\text{ف}} = \frac{\text{مجـ ف}}{\text{ن}} = \frac{٣٥}{٧} = ٥$$

$$ع^2_{ت} = \frac{\text{مجـ}(ف - \bar{ف})^2}{ن - ١} = \frac{١٩٤}{٧ - ١} = \frac{١٩٤}{٦} = ٣٢,٣$$

ومن ذلك نجد أن:

$$ع_{ت} = ٥,٦٨٦$$

نكوّن الاختبار لفرض العدم $ف_{ر}$ والفرض البديل $ف_{ر}$ وليكن عند مستوى معنوية ا = ٠,٠٥ مثلاً. ثم نحسب الإحصائية تي من مشاهدات العينة والاختبار $ف_{ر}$ كالتالي

$ف_{ر}$: $\bar{ف}$ = صفر ، $ف_{ر}$: $\bar{ف} \neq$ صفر (أي أن الفرض البديل ذو طرفين)

$$\text{تي المحسوبة} = \frac{ف}{\frac{ع_{ت}}{\sqrt{ن}}}$$

$$\text{تي المحسوبة} = \frac{٥}{\frac{٥,٦٨٦}{\sqrt{٧}}}$$

$$= ٢,٣٢٧$$

نقارن تي المحسوبة وتساوي ٢,٣٢٧ بالقيمة الموجودة بجداول تي أمام ن − ١ درجات حرية = ٧ − ١ = ٦ ومستوى معنوية أ = ٠,٠٥ فتكون تي (٠,٠٥ ، ٦) = ٢,٤٤٧ فنجد أن تي المحسوبة وتساوي ٢,٣٢٧ ليست واقعة في منطقة الرفض للفرض $ف_{ر}$ ، وهي داخل الحدود الحرجة لمنطقة القبول لـ $ف_{ر}$ وهي (٢,٤٤٧ ، −٢,٤٤٧). أي لا نستطيع رفض فرض العدم $ف_{ر}$ ، القائل: إنه لا توجد فروق معنوية بين درجات الطلاب في مادتي الإحصاء والرياضيات.

لإيجاد تقدير حدود الثقة للفرق بين متوسطي درجات الإحصاء والرياضيات $\bar{ف}$ عند مستوى معنوية ٠,٠٥ نكوّن فترة الثقة كالتالي

$$\bar{\text{ف}} - \text{تي}(\text{٠,٠٥ ، ٦})\frac{\text{ع}_{\text{ت}}}{\sqrt{\text{ن}}} \leqslant (\mu_1 - \mu_2) \leqslant \bar{\text{ف}} + \text{تي}(\text{٠,٠٥ ، ٦})\frac{\text{ع}_{\text{ت}}}{\sqrt{\text{ن}}}$$

ومن ذلك نجد أن

$$\text{٥} - \text{٢,٤٤٧} \times \text{٢,١٤٩} \leqslant \mu_1 - \mu_2 \leqslant \text{٥} + \text{٢,٤٤٧} \times \text{٢,١٤٩}$$

$$\text{٥} - \text{٥,٢٦} \leqslant \mu_1 - \mu_2 \leqslant \text{٥} + \text{٥,٢٦}$$

$$-\text{٠,٢٦} \leqslant \mu_1 - \mu_2 \leqslant \text{١٠,٢٦}$$

(١١ ـ ٧) تمارين

١ ـ أخذت عينة من ٣٦ طفلاً فكان متوسط الوزن لهذه العينة ٨ كجم وكانت أوزان الأطفال تتبع توزيعًا طبيعيًا بمتوسط μ وانحرافًا معياريًا ١،٥ كجم. اوجد فترة الثقة للمتوسط μ عند مستوى معنوية أ = ٠,٠٥ ، أ = ٠,٠١ .

٢ ـ لمقارنة متوسط الدخل للأسر في مدينتين مختلفتين أخذت عينة من المدينة الأولى حجمها ٥٠ أسرة فكان متوسط الدخل لها ٤٥٠٠ ريال وعينة من المدينة الثانية حجمها ٨٠ أسرة فكان متوسط الدخل لها ٥٠٠٠ ريال. فإذا علم أن المجتمعين الإحصائيين يخضعان لتوزيعين معتدلين متوسط الأول μ_1 وانحرافه المعياري ٤٠ ريالاً، ومتوسط الثاني μ_2 وانحرافه المعياري ٥٠ ريالاً.

أوجد فترة الثقة للفرق بين المتوسطين عند مستوى معنوية أ إذا كانت:

(١) ا = ٠,٠٥ ، (٢) ا = ٠,٠١

٣ ـ إذا كانت نسبة المدخنين في إحدى المدن هي ح، أخذت عينة مكونة من ٣٠٠ من سكان هذه المدينة، فوجد من بينهم ٩٠ مدخنًا، فأوجد فترة الثقفة لنسبة التدخين ح عند مستوى معنوية ا إذا كانت:

(١) ا = ٠,٠٥ ، (٢) ا = ٠,٠١

٤ ـ عينة مكونة من ٣٠٠ شخص من البالغين و ٤٠٠ شخص من المراهقين الذين شاهدوا برنامجًا تلفزيونيًا معينًا، فإذا علم أن ٨٠ من البالغين، و ٢٠٠ من المراهقين يفضلون هذا البرنامج.

فاوجد تقديرًا لحدود الثقة للفرق بين نسبة كل من البالغين ونسبة كل المراهقين الذين شاهدوا هذا البرنامج، وذلك عند مستوى معنوية ا كالتالي:

(١) ا = ٠,٠٥ ، (٢) ا = ٠,٠١

٥ ـ أخذت عينتان من توزيعين معتدلين لهما نفس التباين، وجد أن حجم العينة الأولى $ن_{١} = ١٢$ ومتوسطها $\bar{س}_{١} = ٤٨$ وتباينها $ع^{٢}_{١} = ٩٠$، وحجم العينة الثانية $ن_{٢} = ٨$ ومتوسطها $\bar{س}_{٢} = ٥٢$ وتباينها $ع^{٢}_{٢} = ١٢٠$.

اوجد تقديرًا لحدود الثقة للفرق بين المتوسطين $\mu_{١} - \mu_{٢}$ عند مستوى معنوية ا كالتالي:

(١) ا = ٠,٠٥ ، (٢) ا = ٠,٠١

٦ ـ أخذت عينة مكونة من ٣٠٠ طالب من طلاب الجامعة، وجد من بينهم ٥٠ طالبًا يستخدمون أيديهم اليسرى في الكتابة.

كوّن تقديرًا لحدود الثقة لنسبة الطلاب الذين يستخدمون أيديهم اليسرى في هذه الجامعة عند مستوى معنوية ا = ٠,٠٥.

٧ ـ أخذت عينتان حجماهما $ن_{١} = ٩٠$، $ن_{٢} = ١٢٠$ من توزيعين وسطاهما $\mu_{١}$، $\mu_{٢}$ ووجد على الترتيب أن:

$\bar{س}_{١} = ٥٥$ ، $ع^{٢}_{١} = ٩٦$

$\bar{س}_{٢} = ٦٥$ ، $ع^{٢}_{٢} = ٨٤$

اختبر الفروض التالية عند مستوى معنوية ٠,٠٥

(١) $فر_{٠} : \mu_{١} = ٦٠$ ، $فر_{١} : ٦٠ > \mu$

(٢) $فر_{٠} : \mu = ٦٠$ ، $فر_{١} : \mu_{١} \neq ٦٠$

(٣) $فر_{٠} : \mu_{١} = \mu_{٢}$ ، $فر_{١} : \mu_{١} \neq \mu_{٢}$

(٤) $فر_{٠} : \mu_{١} = \mu_{٢}$ ، $فر_{١} : \mu > \mu_{٢}$

٨ ـ قيس الزمن الذي يستغرقه جنديان في فك قطعه من السلاح في ٣٦ حالة لكل منهما، فإذا كانت قياسات كل منهما تخضع لتوزيع طبيعي تباينه ١٤ ثانية وكان الوسط الحسابي لقياسات الجندي الأول ١٣٠ ثانية وللجندي الثاني ١٢٠ ثانية فهل توجد فروق جوهرية بين متوسطيْ كفاءتيهما؟

٩ ـ مصنع للأدوية يدَّعي أن دواءً من انتاجه له فاعلية بنسبة ٨٥٪ في شفاء مرض معين. أخذت عينة مكونة من ١٥٠ شخصًا مصابين بهذا المرض. أدى الدواء الى شفاء ١٢٠ شخصًا منهم. اختبر ما إذا كان ادعاء المصنع صحيحًا؟

١٠ـ متوسط العمر الإنتاجي لعينة من ١٢٠ مصباحًا كهربائيًا من انتاج أحد المصانع هو ١٥٠٠ ساعة وانحرافها ٩٦ ساعة. إذا كان متوسط العمر الإنتاجي لجميع المصابيح المنتجة من المصنع هو μ فاختبر الفرض فر. = ١٦٠٠ ساعة من الفرض البديل فر١ ≠ ١٦٠٠ ساعة مستخدمًا مستوى المعنوية ا إذا كان:

(١) ا = ٠,٠٥ (٢) ا = ٠,٠١

١١ـ عينة من ١٢ قياسًا لأقطارة كرة أعطت متوسط $\bar{س}$ = ٤,٣ ملم وانحراف معياري نجع = ٠,٠٥ ملم. اوجد ما يأتي:

ا ـ اختبر افرض القائل إن العينة مأخوذة من مجتمع متوسطه ٤,٥ ملم.

ب ـ اوجد تقدير حدود الثقة لمتوسط المجتمع μ عند مستوى معنوية ا = ٠,٠٥.

١٢ـ اختبرت ٨ حبال من انتاج أحد المصانع لمعرفة قوة مقاومتها للقطع فأظهرت مقاومتها مقدار ٧٥٠٠ ثقل كجم بانحراف معياري قدره ١٢٠ ثقل كجم. بينما يدعي المصنع المنتج أن قوة المقاومة للقطع لإنتاجه من الحبال هي ٧٨٠٠ ثقل كجم. هل يمكن تأييد ادعاء المصنع عند مستوى المعنوية إذا كان:

(١) ا = ٠,٠٥ (٢) ا = ٠,٠١

١٣ـ إذا كانت نسبة الذكاء لعينة من ١٢ طالبًا في أحد المناطق متوسطها ٩٩ وحدة بانحراف معياري ٨ وحدات. بينما نسبة الذكاء لعينة من ١٤ طالبًا في منطقة أخرى كان متوسطها ١٠٨ وحدات بانحراف معياري ١٢ وحدة فهل هناك اختلاف معنوي بين نسب الذكاء في المجموعتين؟ عند مستوى المعنوية:

(١) ا = ٠,٠٥ (٢) ا = ٠,٠١

١٤ـ إذا أخذنا عينة مكونة من ٨ أشخاص وقرأنا ضغط الدم س لكل واحد من العينة ثم أعطينا كل شخص دواءً معينًا لمدة معينة ثم أعدنا قراءة الضغط بعد الدواء ص لكل واحد في العينة فكانت النتائج كما في الجدول التالي:

قراءات الضغط لثمانية أشخاص قبل وبعد تناول دواء معين

رقم الفرد	١	٢	٣	٤	٥	٦	٧	٨
القراءة س	١٧٠	١٦٠	١٥٠	١٨٠	١٨٥	١٩٠	١٩٥	٢٠٠
القراءة ص	١٦٠	١٥٥	١٤٠	١٦٥	١٧٠	١٧٥	١٨٠	١٨٠

اوجد كلًا من:

١) اختبر الفرض القائل إن الفروق بين متوسطيْ القراءتين غير معنوي عند مستوى معنوية ا إذا كانت ا = ٠,٠٥ و ا = ٠,٠١

٢) اوجد تقديرًا لحدود الثقة للفروق بين متوسطيْ القراءتين عند مستوى معنوية ا إذا كانت ا = ٠,٠٥ و ا = ٠,٠١.

١٥ـ في حالة محاكمة قضائية لشخص متهم بالغش والتزوير فأي من نوعيْ الخطأ في الحكم يحتمل ظهوره وأي من نوعيْ الخطأ أهم بالنسبة للمجتمع.

١٦ـ تبين من الامتحانات السابقة في أول فصل للدورة المكثفة في اللغة الانجليزية أن متوسط الدرجات هو ٧٥ بانحراف معياري ١٠ وقد حصل ٦٥ طالبًا من خريجي إحدى المدارس الثانوية بمدينة الرياض على متوسط درجات قدره ٧٩ فهل يمكن القول: إن خريجي هذه المدرسة الثانوية أحسن مستوى في اللغة الإنجليزية من بقية الطلاب؟

١٧ـ من عينة عشوائية حجمها ١٩٦ شخصًا مأخوذة من أحد أحياء مدينة ما وجد أن عدد النساء ٤٠ فهل يمكن اختبار الفرض القائل: إن نسبة النساء في هذا الحي ٠,٣ باحتمال ٩٥٪؟

١٨ـ في إحدى التجارب التي قام بها طلاب قسم الحيوان لمعرفة تأثير غذاء معين على زيادة الوزن، أخذت عينة مكونة من عشرة فئران وأعطيت الغذاء، وكانت أوزانها بعد التغذية بفترة مناسبة هي:

٢٦٠، ٣٢٠، ٤٨٠، ٢٨٠، ٣٤٠، ٣٠٠، ٤٦٠، ٤٢٠، ٣٠٠، ٣٦٠.

فهل نستطيع أن نحكم على أن هذه العينة مأخوذة من مجتمع متوسط الوزن فيه ٣٨٠ وذلك باعتبار أن مستوى المعنوية ا = ٠,٠٥؟

١٩- اشرح عمليًا لماذا لا يمكن الجزم بأن قطعة نقدية متزنة إذا رميت ألف مرة حصلنا على ٥٥٠ صورة؟