الفصل الرابع

الاستدلال الإحصائي و اختبار الفرضيات

Statistical Inference and Hypothesis Testing

1.4) عزوم العينة و دوالها :

1.1.4) تمهيد : إن هدف الإحصاء كعلم هو القيام باستقراء حول خصائص الجمهرة (المجتمع) اعتماداً على عينة مأخوذة من هذا المجتمع، وكل مسألة إحصائية تبدأ بعينة من القياسات أو الملاحظات ، و لا يخفى أن لطريقة اختيار العينة أثراً حاسماً في الدراسة ، و عندما لا يكون هناك أرجحية لانتقاء عينة دون أخرى ، أي عندما يكون لجميع العينات ذات الحجم نفسه الإمكانات نفسها في السحب ، فمن شأن ذلك أن يدفعنا إلى عد القيم المميزة للعينة هي القيمة المميزة للمجتمع (للجمهرة) الذي أخذت منه العينة ، و يجب الانتباه إلى إن ما يوضع في الحسبان ليس عينة واحدة ، سحبناها و عيّنا خصائصها ، و لكن مجمل العَّينات التي يمكن أن نحصل عليها من المجتمع المدروس إلا أننا لا نعتمد على تلك المعلومات ككيان معزول قائم بذاته ، و إنما نعتمد عليها في سياق سلسلة متكاملة تتضمن العينة المدروسة وغيرها من العينات الممكنة . من أجل ذلك علينا أن نصيغ مفهوم العينة بشكل آخر و يتناسب مع هذا التصور .

لقد لاحظنا أن المتغيرات العشوائية و قوانين توزيعها ما هي إلا نماذج رياضية لدراسة المجتمعات الإحصائية ، و لذلك فإن دراسة المجتمع إحصائياً تعود لدراسة المتغير العشوائي X الذي يصف هذا المجتمع ، وإذا كان لـ X التوزيع $F(x)$

فيمكن القول إن للمجتمع التوزيع $F(x)$ ، و هكذا يمكن التعبير عن القيم المميزة للمجتمع مثل متوسطه μ و تباينه σ^2 بدلالة القيم المميزة للمتغير العشوائي X الذي يمثل هذا المجتمع كما يأتي :

$$\sigma^2 = V(X) \quad , \quad \mu = E(X)$$

و بملاحظة إنَّ عناصر العينة تتغير من تكرار إلى آخر، فيمكن القول إنها قيم لمتغيرات عشوائية مستقلة تمثل المجتمع المدروس .

تعريف (1.4) : نقول عن مجموعة من المتغيرات العشوائية $X_1, X_2 \ , \dots , X_n$ إنها عينة عشوائية من الحجم n للمتغير العشوائي X إذا كانت مستقلة و لها جميعاً قانون X.

و بعبارة أخرى تكون المتغيرات العشوائية $X_1, X_2 \ , \dots , X_n$ عينة عشوائية من الحجم n للمتغير العشوائي X إذا وفقط إذا كان :

$$f\,(X_1, X_2 \ , \dots , X_n) = \prod_{i=1}^{n} f_x\,(X_i\,) \qquad (1.4)$$

ومن أجل متغيرات متقطعة أو مستمرة.

2.1.4) الإحصاءات :

تعريف (2.4): ندعو كل دالة في عينة عشوائية لا تتعلق بوسطاء مجهولة إحصاءً (إحصائية)، فإذا كانت $X_1, X_2 \ , \dots , X_n$ عينة عشوائية من X فإن الدوال:

$$\sum_{i=1}^{n} x_i \ ، \ \frac{1}{n}\sum_{i=1}^{n} x_i \ ، \ X_3 + X_4 ، \prod_{i=1}^{n} x_i$$

تكون جميعها إحصاءات .

بينما الدوال :

$$\sum_{i=1}^{n} x_i^2 - \mu\sigma \text{ ،} \prod_{i=1}^{n} x_i - \sigma^2 \text{ ،} \sum_{i=1}^{n} (x_i - \mu)^2$$

ليست إحصاءات لأنها تحوي وسطاء مجهولة مثل μ و σ^2 و σ و.....

وتصبح إحصاءات عندما تكون الوسطاء معلومة.

ومن الواضح أَن الإحصاء الذي هو دالة في العينة العشوائية ما هو إلا متغير عشوائي تتغير قيمته بتغير العينات ، ومِنْ ثَمّ فله توزيع احتمالي يتعين بوساطة دالة التوزيع للعينة العشوائية الذي هو دالة لها .

تعريف(3.4) متوسط العينة: إذا كانت $X_1, X_2, \dots, X_n$ عينة عشوائية من الحجم n للمتغير العشوائي X، عندئذ متوسط العينة هو الإحصاء:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (2.4)$$

تعريف (4.4) تباين العينة: إذا كانت $X_1, X_2, \dots, X_n$ عينة عشوائية من الحجم n للمتغير العشوائي X ، عندئذ تباين العينة يعرّف بالإحصاء :

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = \frac{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)} \qquad (3-4)$$

2.7) توزيع كاي– مربع وتوزيع ستيودنت :

إن لتوزيع كاي– مربع وتوزيع ستيودنت أهمية كبيرة في مجال الإحصاء التطبيقي، ومِنْ ثَمّ لا بد من استعراض هذين التوزيعين ، وسنبدأ بتعريف دالة غاما ذات العلاقة بهما.

تعريف(5.7): الدالة

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} \cdot e^{-x} \, dx \qquad (4.4)$$

تدعى بدالة غاما ،حيث لها الخواص الآتية :

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1) \;\; ; \;\; \Gamma(1) = 1 \;\; ; \;\; \Gamma(n) = (n-1)!$$

$$\Gamma(2) = 1 \quad ; \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

1.2.4) : تعريف (6.4): توزيع (كاي– مربع):

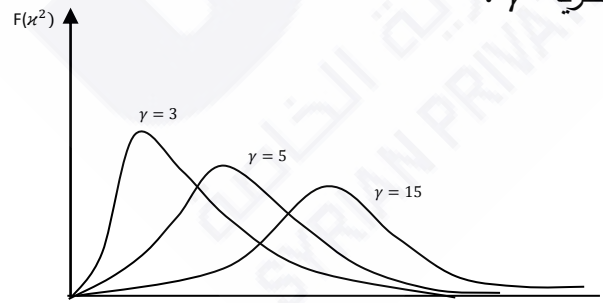
نقول إن للمتغير العشوائي المستمر X التوزيع $\varkappa^2$ بـ γ (نيو) درجة من الحرية اذا كانت دالة كثافته الاحتمالية معطاة بالعلاقة الآتية :

$$f_X(x) = \frac{1}{2^{\gamma/2} \cdot \Gamma(\gamma/2)} \cdot x^{\frac{\gamma}{2}-1} \cdot e^{-\frac{x}{2}} \qquad ; x > 0 \quad (5.4)$$

ونكتب $X \sim \varkappa^2(\gamma)$

و درجة الحرية هنا تعبير خاص يدل على الوسيط لهذا التوزيع .

و الشكل(1.4) يعطي التمثيل البياني لدالة كثافة التوزيع x^2 من أجل قيم مختلفة لدرجة الحرية γ .



الشكل (1.4)

1) بعض الخواص الهامة لتوزيع كاي– تربيع ($\varkappa^2$) :

1 . إذا كان المتغير العشوائي Z طبيعياً معيارياً فإن Z^2 له التوزيع $\varkappa^2$ بدرجة واحدة من الحرية .

2. إذا كانت المتغيرات العشوائية $X_1, X_2 , \dots , X_n$ مستقلة و لكل منها توزيع $N\ (0,1)$ فإن المتغير العشوائي $Y = \sum_{i=1}^{n} \varkappa_i^2$ له توزيع كاي– تربيع بـ $\gamma = n$ درجة من الحرية .

3. إذا كان X_1, X_2 متغيرين عشوائيين مستقلين و لكل منهما توزيع $\varkappa^2$ بـ γ_1, γ_2 درجة من الحرية على الترتيب ، فإنه يكون للمتغير العشوائي $X_1 + X_2$ توزيع كاي– تربيع بـ $\gamma = \gamma_1 + \gamma_2$ درجة من الحرية .

4. إذا كان X_1, X_2 متغيرين عشوائيين مستقلين و كان لـ X_1 توزيع $\varkappa^2$ بـ γ_1 درجة من الحرية و كان لمجموعهما $X_1 + X_2$ توزيع كاي– تربيع بـ $\gamma >$ γ_1 درجة من الحرية، فإنه يكون للمتغير العشوائي X_2 توزيع $\varkappa^2$ ب $\gamma_2 =$ $\gamma - \gamma_1$ درجة من الحرية .

5. إذا كان X متغيراً عشوائياً له التوزيع $\varkappa^2$ بـ γ درجة من الحرية فإن :

$$V\ (X) = 2\,\gamma \quad ، \quad E\ (X) = \gamma$$

و جدول كاي– تربيع للتوزيع الاحتمالي معطى في ملحق الجداول ، و هو يعطي القيمة x^2 التي تقع على يسارها α من المساحة الكلية ، تحت منحني الكثافة ، و ذلك من أجل قيم مختلفة ل α و γ و تكون $x_\alpha^2(\gamma)$ قيمة المتغير x^2 في الجدول الواقعة عند تلاقي السطر γ و العمود الموافق لقيمة α .

مثلاً :

$$x^2_{0.025}(19) = 8.91 ، x^2_{0.975}(10) = 20.5$$

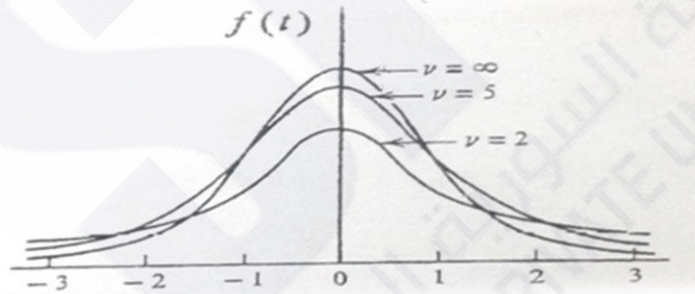
2.2.4) : توزيع ستيودنت t :

تعريف (7.4) : نقول إن للمتغير العشوائي X التوزيع t − ستيودنت ب γ درجة من الحرية إذا كانت دالة كثافته معطاة بالعلاقة الآتية :

$$f_X(x) = \frac{\Gamma\left(\frac{\gamma+1}{2}\right)}{\Gamma\left(\gamma/_2\right).\sqrt{2\pi}}\left(1+\frac{x^2}{\gamma}\right)^{-\frac{\gamma+1}{2}} \qquad -\infty < x < +\infty$$

حيث لها عدد صحيح موجب يدعى بدرجة الحرية .

و الشكل (2.4) يعطي أمثلة من منحنيات الكثافة لهذا التوزيع ، فهو متناظر حول محور التراتيب شأنه شأن التوزيع الطبيعي المعياري ، و يعتمد التوزيع على الوسيط γ (درجة من الحرية).



الشكل (2.4)

و جدول توزيع ستيودنت معطى بملحق الجداول ، و هو يعطي القيم الموجبة لـ t التي تقع على يسارها α من المساحة الكلية ، تحت منحني الكثافة لتوزيع t− ستيودنت ، و ذلك من أجل قيم مختلفة لـ α و γ و سنرمز $t_\alpha(\gamma)$ للدلالة على

قيمة المتغير t من الجدول الواقعة عند تلاقي السطر γ و العمود الموافق لقيمة α .

مثلاً: $t_{0.60}(15) = 0.258$

مبرهنة (1.4) : إذا كان Z متغيراً عشوائياً له توزيع $N\ (0\ ,1)$ و كان X متغيراً عشوائياً له توزيع $\varkappa^2$ بـ γ درجة من الحرية و كان X،Z مستقلين عشوائياً ، فإنه يكون للمتغير العشوائي

$$T = \frac{Z}{\sqrt{X/_V}}$$

توزيع t – ستيودنت بـ γ درجة من الحرية .

3.4) : توزيعات بعض الإحصاءات :

1. توزيع مجموع متغيرات عينة عشوائية $\sum_{i=1}^{n} X_i$:

إذا كانت $X_1, X_2\ , \dots, X_n$ عينة عشوائية من مجتمع طبيعي متوسطه μ و تباينه σ^2 فإن يكون للمتغير $Y = \sum_{i=1}^{n} x_i$ التوزيع الطبيعي بمتوسط $\mu_y = n\,\mu$ و تباين $\sigma_y^2 = n\,\sigma^2$ (أي $y \sim N\ (n\mu, n\,\sigma^2\ )$).

أما إذا كانت $X_1, X_2\ , \dots, X_n$ عينة عشوائية من مجتمع متوسطه μ و تباينه σ^2 فإنه حسب مبرهنة النهاية المركزية ، ومن أجل n كبيرة كبراً كافياً ($n \geq 30$) فإن :

$y = \sum_{i=1}^{n} x_i \sim N\ (n\mu, n\,\sigma^2\ )$ (تقريباً طبيعي) .

2. توزيع متوسط العينة العشوائية $\bar{x}$:

إذا كانت $X_1, X_2, \dots, X_n$ عينة عشوائية لمتغير عشوائي X له التوزيع الطبيعي بمتوسط μ و تباين σ^2 ، فإن يكون لمتوسط العينة $\bar{x}$ التوزيع الطبيعي بمتوسط μ و تباين $\frac{\sigma^2}{n}$ أي $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

أما إذا كانت العينة من متغير عشوائي X متوسط μ و تباين σ^2 ، فحسب مبرهنة النهاية المركزية يكون $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ (تقريباً طبيعي) و ذلك من أجل $(n \geq 30)$.

مثال (4-1) :

إذا كانت $X_1, X_2, \dots, X_n$ عينة عشوائية لمتغير X يتوزيع وفق N (3 ، 16)

أوجد $p[0.5 < \bar{x} < 6]$

الحل :

بما أن العينة طبيعية ، سيكون عندئذ للمتغير $\bar{x}$ توزيع طبيعي بمتوسط $\mu_{\bar{x}} = 3$، وتباين $\sigma^2 = \frac{\sigma^2}{n} = \frac{16}{9}$ و منه :

$$p[0.5 < \bar{x} < 6] = p\left[\frac{0.5-3}{4/3} < \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{6-3}{4/3}\right]$$

$$= p[-1.88 < z < 2.25]$$

$$= p[z < 2.25] - p[z < -1.88] =$$

$$0.9878 - 0.0301 = 0.9577$$

مثال (2.4) :

إذا كان متوسط الدخل الأسبوعي لمجموعة كبيرة من العمال المهرة هو s.p 15000 بانحراف معياري s.p 120.

فاحسب احتمال أن يكون متوسط دخل عينة حجمها n= 64 أكبر من 14900 s.p في الأسبوع .

الحل :

بما أن حجم العينة $n = 64 \geq 30$ ، فيمكن تطبيق مبرهنة النهاية المركزية و يكون للمتوسط $\bar{X}$ تقريباً التوزيع الطبيعي $\bar{X} \sim N\left(15000, \frac{14400}{64}\right)$ و منه:

$$p\,[\bar{x} > 14900] = p\left[\frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} > \frac{14900-15000}{15}\right] = p[z > -6.67\,] =$$

$$1 - p[z < -6.67\,] = 1 - 0 = 1$$

بحيث : $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{120}{8}$

مثال (3-4): يتوزع وزن مجموعة من المرضى وفق التوزيع الطبيعي بمتوسط 65 k.g و بانحراف معياري 15 k.g من أجل عينة عشوائية من هؤلاء المرضى من الحجم 100 مريض، ما احتمال أن يتجاوز الوزن الكلي للمرضى 7000k.g .

الحل :

إن أوزان المرضى هي عينة عشوائية لمتغير عشوائي طبيعي متوسطه $\mu = 65$ و تباينه $\sigma^2 = (15)^2 = 225$ ، و منه يكون للمجموع $Y = \sum_{i=1}^{n} x_i$ التوزيع الطبيعي بمتوسط $\mu_y = n\,\mu = 6500$ و بتباين $\sigma_y^2 = n\,\sigma^2 =$ 22500 و منه:

$$p[\,y > 7000] = p\,[\frac{y-\mu_y}{\sigma_y} > \frac{7000-6500}{\sqrt{225500}} = p\,[\,z > 3.33] =$$

$$1 - p\,[\,z < 3.33\,] = 1 - 0.9996 = 0.0004$$

3 . توزيع الإحصاء $\frac{(n-1)S^2}{\sigma^2}$:

إذا كان $\bar{x}$ و S^2 متوسط و تباين عينة عشوائية من الحجم n للتوزيع الطبيعي $N\,(\,\mu, \sigma^2\,)$ فإن :

1. $\bar{x}$ و S^2 متغيران عشوائيان مستقلان .
2. $\frac{(n-1)S^2}{\sigma^2}$ متغير عشوائي له توزيع كاي – مربع بـ $\gamma = n - 1$ درجة من الحرية .

4. توزيع المتغير العشوائي $T = \frac{\bar{x}-\mu}{s/\sqrt{n}}$:

إذا كان $\bar{x}$ و S^2 متوسط و تباين عينة عشوائية من الحجم n للتوزيع الطبيعي $N\,(\,\mu, \sigma^2\,)$ فإنه يكون للمتغير العشوائي $T = \frac{\bar{x}-\mu}{s/\sqrt{n}}$ توزيع ستيودنت بـ $\gamma =$ $n - 1$ درجة من الحرية .

4.4) التقدير النقطي :

رأينا أن المجتمع الإحصائي يمثل بمتغير عشوائي X ، و معرفة توزيع المتغير العشوائي X يجعلنا قادرين على دراسته احتمالياً، و عادة يكون توزيع المجتمع أو التوزيع الاحتمالي لـ X يتبع وسطاء ، و عندما تكون هذه الوسطاء مجهولة نلجأ لتقديرها اعتماداً على عينة عشوائية حجمها n من X ، فإذا كان θ وسيطاً مجهولاً لتوزيع X فإننا نقوم بتقدير θ بوساطة دالة مثل T =

$T(X_1, X_2, \dots, X_n)$ و قيمة T و لتكن t مقدراً لـ θ و نرمز لمقدر θ عاده بـ $\hat{\theta}$ و من خلال طرق التقدير النقطية سيكون:

1- متوسط العينة $\bar{x}$ مقدراً لمتوسط المجتمع μ .

2- و تباين العينة S^2 مقدراً لتباين المجتمع σ^2 .

3- و الانحراف المعياري S مقدراً للانحراف المعياري للمجتمع σ .

4- إذا كان لدينا مجتمع إحصائي و أردنا أن تقدير نسبة العناصر من المجتمع التي تحقق صفة معينة ، فيمكن أن نمثل هذا المجتمع بمتغير عشوائي برنولي X يأخذ القيمة 0 إذا لم يحقق الصفة و القيمة 1 إذا كان يحققها . عندئذ إذا كانت $X_1, X_2, \dots, X_n$ عينة عشوائية من X الذي يتوزع وفق برنولي بالوسيط p (احتمال حدث النجاح) فإن $\hat{p} = \bar{x} = \frac{y}{n}$ هو مقدر لـ p حيث يدل y على عدد العناصر التي تحقق الصفة في العينة .

و نريد من هذه المقدرات أن تحقق معايير جودة المقدرات (مقدر غير منحاز (منصف) – مقدر فعال – مقدر متماسك)

تعريف (4-8): إذا كان $\hat{\theta}$ مقدراً للوسيط θ فإننا ندعوه بالمقدر غير المنحاز لـ θ إذا حقق الشرط الآتي : $E(\hat{\theta}) = \theta$.

تعريف (4-9) : إذا كان $\hat{\theta}$ مقدراً للوسيط θ و كان له أصغر تباين من بين كل المقدرات الأخرى لـ θ عندئذ نقول إن المقدر $\hat{\theta}$ هو مقدر فعال لـ θ .

و إذا كان $\hat{\theta}$ مقدراً منصفاً (غير منحاز) لـ θ و كان $V(\hat{\theta}) = \min$ فإننا ندعوه عندئذ بالمقدر المنصف ذي التشتت الأصغري (التباين الأصغري).

تعريف (4-10) : يكون $\hat{\theta}$ مقدراً متماسكاً (متسق) للوسيط θ إذا كان :

$$\lim_{n\to\infty} V(\hat{\theta}) = 0 \quad ، \quad E(\hat{\theta}) = \theta$$

و بناء عليه فإن متوسط العينة $\bar{x}$ هو مقدر غير منحاز لمتوسط المجتمع μ و كذلك $\hat{p} = \frac{y}{n}$ هو مقدر غير منحاز لنسبة تحقق صفة معينة في المجتمع P و لكن $\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ هو مقدر غير منصف (منحاز) لتباين المجتمع σ^2 بينما يكون $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ هو مقدر منصف لـ σ^2.

5.4) : مجالات الثقة (التقدير المجالي) :

رأينا في الفقرة السابقة مسألة التقدير النقطي حيث عينا بالمقدر النقطي القيمة التي يأخذها المقدر من أجل عينه معينة من قيم المتغير X فإذا فرضنا أن X متغير عشوائي توزيعه الاحتمالي يتبع وسيطاً مجهولاً θ و أن $\hat{\theta}$ هو مقدر للوسيط θ يحقق معايير الجودة في الإنصاف و الفعالية .

فإننا سنتوقع من أجل عينة معينة من قيم X مثل $X_1, X_2, \dots, X_n$ بأن قيمة المقدر $\hat{\theta}$ لن تختلف كثيراً عن القيمة الحقيقية للوسيط θ و أن الخطأ المرتكب عندما نفترض أن قيمة المقدر مساوية قيمة الوسيط المجهول θ لن يكون كبيراً، و مع ذلك فإن قيمة واحدة للمقدر لا تحمل أي دلالة على الدقة الحاصلة في ذلك التقدير ، لذلك يكون من المرغوب فيه إنشاء مجال بحيث يغطي هذا المجال القيمة الحقيقية للوسيط المجهول θ باحتمال مفروض .

تعريف (4-11) : ليكن X متغيراً عشوائياً توزيعه الاحتمالي يتبع وسيطاً مجهولاً θ و لنفترض أن $X_1, X_2, \dots, X_n$ عينة عشوائية للمتغير X و أن L_1, L_2 إحصاءان على أساس العينة المفروضة .

إننا نقول عن المجال $[L_1, L_2]$ إنه مجال ثقة لـ θ بمستوى من الثقة

$100(1-\alpha)\%$. حيث α يدعى بمستوى المعنوية و $(1-\alpha)$ معامل الثقة .

إذا كان $P[L_1 \leq \theta \leq L_2] = 1 - \alpha$

و بملاحظة أن طرفي المجال $[L_1, L_2]$ هما متغيران عشوائيان يتغيران من عينة لأخرى ، فيمكن أن نسمي هذا المجال بالمجال العشوائي و من أجل كل عينة من قيم Xنستطيع حساب كل من L_1, L_2 و الحصول على مجال يكون على ثقة $100(1-\alpha)\%$ من أنه يحوي على القيمة الحقيقية للوسيط المجهول θ، و سنواجه حالات يكون فيها هناك أكثر من حل ، و سنختار منها تلك الحالة التي يكون فيها طول مجال الثقة أصغرياً قدر الإمكان .

1.5.4) : مجال الثقة لمتوسط متغير عشوائي طبيعي تباينه معلوم

ليكن $X \sim N(\mu, \sigma^2)$ فيه μ مجهول و σ^2 معلوم .

و لتكن $X_1, X_2, \dots, X_n$ عينة عشوائية لـ X ، رأينا أن:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

و المطلوب تعيين مجال ثقة حول μ بمستوى $(1-\alpha)$ من الثقة أي من

$$p[z_1 \leq z \leq z_2] = 1 - \alpha \qquad (7\text{-}4)$$

و من خواص الكثافة الطبيعية المعيارية المتناظرة بالنسبة لمحور التراتيب سنجد

$$p\left[-z_{1-\alpha/2} \leq z \leq z_{1-\alpha/2}\right] = 1 - \alpha \Rightarrow$$

$$p\left[-z_{1-\alpha/2} \le \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \le z_{1-\alpha/2}\right] = 1-\alpha \qquad \Leftrightarrow$$

$$p\left[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = 1-\alpha$$

أي إن المجال $[\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$

هو $100(1-\alpha)$ مجال ثقة للمتوسط μ

و نقول إننا واثقون باحتمال $(1-\alpha)$ من أن μ لن يقل عن $\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$ و لن يزيد على $\bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$.

ملاحظة: بفضل مبرهنة النهاية المركزية و من أجل العيّنات العشوائية من مجتمعات غير طبيعية ذات تباين معلوم σ^2 و من الحجم $n \ge 30$ يكون لدينا نفس مجال الثقة حول μ بمستوى $(1-\alpha)$.

و نسمي المقدار $\varepsilon = \left|\mp z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right|$ بالخطأ الأعظمي المطلق و المرتكب في تقدير μ عن طريق $\bar{x}$ وبثقة $(1-\alpha)$ ، و هذا الخطأ يتناقص بازدياد حجم العينة n، لذلك يمكن التحكم بالخطأ الأعظمي بوساطة حجم العينة .

و إذا أردنا تعيين حجم العينة التي ينبغي أخذها بحيث لا يتجاوز الخطأ في تقديرنا لـ μ ب $\bar{X}$ المقدار ε و بثقة $(1-\alpha)$ من المقدار

$$\left|\mp z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right| \le \varepsilon$$

و بتربيع الطرفين و نقل n للطرف الآخر نجد :

$$n \geq \left[\frac{z_{1-\alpha/2}\sigma}{\varepsilon}\right]^2$$

مثال (4-4) :

أجريت معايرة كمية الخضاب في الدم لعينة مؤلفة من 36 طفلاً ، فكان متوسط الخضاب لديهم 11.3 g ، فإذا كانت العينة مختارة من مجتمع فيه الانحراف المعياري لكمية الخضاب 2.5 g فالمطلوب :

1. أوجد 98% مجال ثقة حول المتوسط الحقيقي μ لكمية الخضاب لمجتمع الأطفال الذي أخذت من العينة .

2. كم ينبغي أن يكون حجم العينة بحيث لا يتجاوز الخطأ في تقدير μ بثقة 98% المقدار $\varepsilon = 0.80$؟

الحل :

1. بما ان $n = 36 \geq 30$ ، فيمكن وضع مجال ثقة تقريبي لـ μ متوسط كمية خضاب الدم في المجتمع و ذلك على الرغم من أن كمية الخضاب ليس لها التوزيع الطبيعي ، و بملاحظة أن :

$1 - \alpha = 0.98 \Leftarrow 1 - \alpha/2 = 0.99$ و من جدول التوزيع الطبيعي المعياري $z_{0.99} = 2.33$

و يكون المجال :

$$\left[\bar{x} - z_{1-\alpha\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

أي

$$[\,11.3-(\,2.33)\frac{(2.5)}{6}\ \le\ \mu \le 11.3+(\,2.33)\frac{(2.5)}{\sigma}\,]$$

فيكون : $[\,10.329\ \le\ \mu \le 12.271\,]$

أي نكون واثقين 98% من أن المتوسط كمية الخضاب الدم عند الأطفال المجتمع المدروس لن تقل عن 10.329 g و لن تزيد على 12.271 g .

2. لتعيين n بحيث يكون : $n\ \ge \left[\frac{z_{1-\alpha/2}\sigma}{\varepsilon}\right]^2 = \left[\frac{(2.33\,)(\,2.5)}{0.80}\right]^2$

أي يجب أن يكون حجم العينة $n\ \ge 54$

<u>ملاحظة:</u> في الحالة التي يكون فيها تباين المجتمع σ^2 مجهولاً و من أجل العيّنات ذات الحجم $n\ \ge 30$ يكون تباين العينة S^2 مقدراً جيداً لـ σ^2.

ومِنْ ثَمّ يمكن استبدال بـ σ ، S الانحراف المعياري للعينة . و هكذا يمكننا تعيين مجال ثقة تقريبي لمتوسط المجتمع μ بمستوى $(1-\alpha)$ من الثقة

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

و يكون الخطأ المطلق الأعظمي في تقدير μ و بثقة $(1-\alpha)$ هو

$$\left|\mp\, z_{1-\alpha/2}\frac{S}{\sqrt{n}}\right|$$

أما حجم العينة التي ينبغي سحبها بحيث لا يتجاوز الخطأ المرتكب في تقدير μ المقدار ε بثقة $(1-\alpha)$ ينتج من العلاقة $n\ \ge \left[\frac{z_{1-\alpha/2}\cdot\ S}{\varepsilon}\right]^2$

حيث S هو الانحراف المعياري للعينة العشوائية المسحوبة من مجتمع الدراسة حجمها $n\ \ge 30$

مثال (4-5): في دراسة احصائية حول زمن تخثر الدم عند الأشخاص الذين يقطنون في إحدى المدن من البلدان الحارة ، تم دراسة عينة عشوائية من سكانها من الحجم 100 ، و وجد أن متوسط زمن تخثر الدم في العينة هو 12 ثانية و بانحراف معياري 2 ثانية . و المطلوب :

1. أوجد 95% مجال ثقة حول μ المتوسط الحقيقي لزمن تخثر الدم عند سكان المدينة المدروسة .

2. ما حجم العينة المناسب لتقدير μ بثقة 95% و بخطأ أعظمي لا يتجاوز 0.25 ثانية ؟

الحل :

a) إن $1-\alpha=0.95$ مجال ثقة لـ μ يكون من الشكل :

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

(لأن $n = 100 \ge 30$ و الانحراف المعياري S هو مقدر لـ σ)

$$\left[12 - (1.96).\left(\frac{2}{10}\right) \le \mu \le 12 + (1.96).\left(\frac{2}{10}\right)\right]$$

(حيث $1-\alpha = 0.95 \leftarrow 1-\alpha/2 = 0.975$ و من جدول $Z \leftarrow z_{0.975} = 1.96$)

و منه $[12 - 0.392 \le \mu \le 12 + 0.392]$

أي $[11.608 \le \mu \le 12.392]$

أي نكون واثقين 0.95% من أن متوسط زمن تخثر الدم عند سكان المدينة لن يقل عن 11.608 و لن يزيد على 12.392 ثانية.

2- إن حجم العينة المناسب لتقدير μ بثقة 0.95% و بخطأ لا يتجاوز 0.25 ثانية واحدة يتعين

$$n \geq \left[\frac{z_{1-\alpha/2} \cdot S}{\varepsilon}\right]^2 = \left[\frac{(1.96).(2)}{0.25}\right]^2 = 245.86$$

أي يجب أن يكون $n \geq 246$.

2.5.4) : مجال الثقة لمتوسط متغير عشوائي طبيعي تباينه مجهول:

لقد لاحظنا في الدراسة السابقة بأن معرفتنا لتباين المجتمع σ^2 ضروريّةٌ لتعيين مجال الثقة حول μ ، و في الحالة التي يكون فيها σ^2 مجهولاً فإننا لا نستطيع دوماً استبدال تباين المجتمع σ^2 تباين العينة s^2 ، و ذلك لأن الإحصاء $\frac{\bar{x}-\mu}{S/\sqrt{n}}$ يحوي متغيرين عشوائيين هما $\bar{x}$ ، S .

و عندما يكون حجم العينة $n \geq 30$ فإن تغيرات s^2 من عينة لأخرى تكون معدومة، لذلك يمكن استبدال لـ σ , S الانحراف المعياري للعينة . و هذا ما أشرنا إليه في الفقرة السابقة حيث اعتبرنا $\frac{\bar{x}-\mu}{S/\sqrt{n}} \sim N(0,1)$ ، و لكن في الحالة التي يكون فيها $n \leq 30$ فإن تغيرات s^2 تغدو مؤثرة في شكل توزيع الاحصاء $\frac{\bar{x}-\mu}{S/\sqrt{n}}$ ، و يصبح توزيع هذا الإحصاء مختلفاً عن التوزيع الطبيعي المعياري ، و رأينا مسبقاً في بداية هذا الفصل أن للمتغير العشوائي

$T = \frac{\bar{x}-\mu}{S/\sqrt{n}}$ التوزيع t– ستيودنت بـ $\gamma = n-1$ درجة من الحرية

ومِنْ ثَمّ من أجل بناء $(1-\alpha)$ مجال ثقة حول μ نأخذ العلاقة

$$P\left[-t_{1-\alpha/2}(\gamma) \le T \le t_{1-\alpha/2}(\gamma)\right] = 1-\alpha$$

و باستبدال T بما يساويه نجد :

$$P\left[-t_{1-\alpha/2}(\gamma) \le \frac{\bar{x}-\mu}{S/\sqrt{n}} \le t_{1-\alpha/2}(\gamma)\right] = 1-\alpha \quad \leftrightarrow$$

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

أي إن المجال $\left[\bar{x}-t_{1-\alpha/2}(\gamma)\frac{S}{\sqrt{n}} \le \mu \le \bar{x}+t_{1-\alpha/2}(\gamma)\frac{S}{\sqrt{n}}\right]$ هو

$(1-\alpha)$ مجال ثقة للمتوسط μ .

و يكون $e = \left|\mp t_{1-\alpha/2}(\gamma)\frac{S}{\sqrt{n}}\right|$ ، $n \ge \left[\frac{t_{1-\alpha/2}(\gamma)\ .\ S}{\varepsilon}\right]^2$

كما رأينا في الفقرة السابقة .

مثال (6-4) :

تم قياس ارتفاع 15 نبتة من نوع معين بعد فترة زمنية من زرعها، فكان متوسط الارتفاع 83 c.m بانحراف معياري 5.8 c.m و المطلوب :

1. أوجد 0.95% مجال ثقة حول μ المتوسط الحقيقي لارتفاع النبته في مجتمع الدراسة بافتراض أن ارتفاع النبتة في المجتمع يخضع للتوزيع الطبيعي .

2. ما حجم العينة المناسب لتقدير μ بثقة 0.95% و بخطأ لا يتجاوز 2c.m؟

الحل :

1- إن 0.95% مجال ثقة حول μ حيث $n = 15 < 30$ يكون من الشكل

$$\bar{x} - t_{1-\alpha/2}(\gamma)\frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{1-\alpha/2}(\gamma)\frac{S}{\sqrt{n}}$$

حيث:

$\bar{x} = 83$، $S = 58$ و $1 - \alpha = 0.95 \leftarrow 1 - \alpha/2 = 0.975$

كما أن $\gamma = n - 1 = 15 - 1 = 14$ و منه من جدول t – ستيودنت نجد $t_{0.975}(14) = 2.145$

و منه :

$$\left[83 - (2.145)\frac{(5.8)}{\sqrt{15}} \leq \mu \leq 83 + (2.145)\frac{(5.8)}{\sqrt{15}}\right]$$

فيكون : $[79.788 \leq \mu \leq 86.212]$

أي نكون واثقين 95% من أن متوسط ارتفاع النبتة في مجتمع الدراسة لن يقل عن 79.788 و لن يزيد على 86.212.

2- $$n \geq \left[\frac{t_{1-\alpha/2}(\gamma) \cdot S}{\varepsilon}\right]^2 = \left[\frac{(2.145)(5.8)}{2}\right]^2 = 38.7$$

أي حجم العينة المناسب لتقدير μ بثقة 0.95% و بخطأ أعظمي لا يتجاوز 2c.m ينبغي أن يكون $n \geq 39$.

ملاحظة : كلما ازداد حجم العينة أصبح S تقديراً أفضل لـ σ ومِنْ ثَمّ اقتربت قيم t_α من قيم Z_α في التوزيع الطبيعي المعياري ، و هذا ما نلاحظه في جدول t – ستيودنت من خلال سطوره الأخيرة و مقارنة مع جدول التوزيع الطبيعي المعياري.

3.5.4) : مجال الثقة للنسبة في المجتمع :

من المسائل المهمة التي كثيراً ما نصادفها تقدير نسبة العناصر من المجتمع المحققة لصفة معينة مثل نسبة المؤيدين لبرنامج اقتصادي معين ، نسبة نجاح حملة صحية معينة ، نسبة المصابين بمرض معين ، نسبة فعالية دواء معين لعلاج أحد الأمراض ، نسبة نجاح حملة تلقيح ضد مرض معين .....

فإذا كانت نسبة العناصر المحققة للصفة A في مجتمع مدروس هي P فإن احتمال أن نختار عنصراً يحقق هذه الصفة يساوي P ، ومِنْ ثَمّ يمكن أن نمثل المجتمع المدروس بمتغير عشوائي X له توزيع برنولي بوسيط P هي النسبة في المجتمع ، و لقد رأينا أن $\hat{P} = \bar{X} = \frac{y}{n}$ هو مقدر للنسبة P .
علماً بأن $\bar{x}$ هو متوسط عينة حجمها n لمتغير عشوائي برنولي X وسيطه P، و أن y يدل على عدد عناصر العينة المحققة للصفة A (عدد النجاحات) ،

و أن $\hat{p}$ هو مقدر غير منحاز لـ P و من أجل عينة كبيرة كبراً كافياً ($n \geq 30$) فإن $\hat{p}$ حسب مبرهنة النهاية المركزية له تقريباً توزيع طبيعي بمتوسط $P = \mu$ و تباين $\sigma^2 = \frac{pq}{n}$ (أي $\hat{p} \approx N\left(p, \frac{pq}{n}\right)$)

ومِنْ ثَمّ $$Z = \frac{\hat{p} - P}{\sqrt{\frac{pq}{n}}} \approx N(0,1)$$

و من أجل $(1 - \alpha)$ مجال ثقة حول P نأخذ

$$P\left[-z_{1-\alpha/2} \le z \le z_{1-\alpha/2}\right] = 1 - \alpha$$

نعوض فيه Z :

$$P\left[-z_{1-\alpha/2} \le \frac{\hat{p} - P}{\sqrt{\frac{pq}{n}}} \le z_{1-\alpha/2}\right] = 1 - \alpha$$

و منه نجد : ($1 - \alpha$) مجال ثقة حول P من الشكل :

$$\hat{p} - z_{1-\alpha/2}\sqrt{\frac{pq}{n}} \le P \le \hat{p} + z_{1-\alpha/2}\sqrt{\frac{pq}{n}}$$

لكن طرفي المجال السابق ينبعان الوسيط P و باعتبار أن تغيرات $\sqrt{\frac{pq}{n}}$ تكون بطيئة جداً ، فيمكن أن نستبدل بـ p ، q ، $\hat{p}$ ، $\hat{q}$ حيث أن $\hat{q} = 1 - \hat{p}$

و هكذا فإن ($1 - \alpha$) مجال ثقة حول p هو تقريباً المجال

$$\left[\hat{p} - z_{1-\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}},\ \hat{p} + z_{1-\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}}\right]$$

و منه يلحظ أن الخطأ المطلق الأعظمي المرتكب في تقدير p وبثقة $(1 - \alpha)$ يساوي $\left|\mp z_{1-\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}}\right|$ و أن حجم العينة الواجب أخذها لنقدر p بثقة ($1 - \alpha$) و بخطأ لا يتجاوز ε هو $n \ge \frac{z_{1-\alpha/2}^{2}.\hat{p}.\hat{q}}{e^2}$.

مثال (4–7): لدى تخدير 100 شخص من المرضى الكبار بالسن تبين أن 36 منهم حصلت لهم مضاعفات جراء تخديرهم . و المطلوب :

1. أوجد 99% مجال ثقة لنسبة الذين يعانون من مضاعفات جراء التخدير من بين المرضى الكبار بالسن .

2. عين الخطأ المطلق المرتكب عندما نفترض أن $\hat{p} = p$ بثقة 99%

3. ما حجم العينة التي ينبغي دراستها لتقدير p بثقة 99% و بخطأ لا يتجاوز 0.04؟

الحل :

1. لدينا $n = 100$ و $y = 36$ و منه

$\hat{p} = \frac{y}{n} = \frac{36}{100} = 0.36$ و من أجل

$1 - \alpha = 0.99 \rightarrow 1 - \alpha/2 = 0.995$ و يكون

$z_{0.995} = 2.58$ و منه 99% مجال ثقة حول p نسبة الذين يعانون مضاعفات جراء التخدير من بين المرضى الكبار في السن

$$\hat{p} - z_{0.995}\sqrt{\frac{\hat{p}.\hat{q}}{n}} \leq P \leq \hat{p} + z_{0.995}\sqrt{\frac{\hat{p}.\hat{q}}{n}}$$

$$(0.36) - (2.58)\sqrt{\frac{(0.36).(0.64)}{100}} \leq P \leq (0.36) + (2.58)\sqrt{\frac{(0.36).(0.64)}{100}}$$

$$0.36 - 0.124 \leq P \leq 0.36 + 0.124$$

$$0.236 \leq P \leq 0.484$$

و منه نكون واثقين 99% من أن هذه النسبة P لن تقل عن 0.236 و لن تزيد على 0.484

.2

$$\varepsilon = \left| \mp z_{1-\alpha/2} \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} \right| = 0.124$$

.3

$$n \geq \frac{z_{1-\alpha/2}^2 \cdot \hat{p} \cdot \hat{q}}{\varepsilon^2} = \frac{(2.58)^2 .(0.36).(0.64)}{(0.04)^2} = 958.52$$

أي حجم العينة يجب أن يكون $n \geq 959$

مثال (4-8) : ادعى باحث أن 10% من الأشخاص عسراويون ، و لاختبار هذا الادعاء ، تم اختيار 400 شخص ، و وجد 48 منهم عسراويين . فهل يمكننا قبول هذا الادعاء بمستوى 99% من الثقة .

الحل : لنشكل 99% مجال ثقة لـ p نسبة الأشخاص العسراويين و بملاحظة أن لدينا $n = 400$ و $x = 48$ و منه $\hat{p} = \frac{x}{n} = \frac{48}{400} = 0.12$ ، ومن أجل $1 - \alpha = 0.99 \Rightarrow 1 - \alpha/2 = 0.995$

و يكون: $z_{0.995} = 2.58$ (و هذا ينتج من جدول التوزيع الطبيعي المعياري)، ومِنْ ثَمّ فمجال ثقة المطلوب :

$$\hat{p} - z_{1-\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}} \leq P \leq \hat{p} + z_{1-\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}} \Leftrightarrow$$

$$(0.12) - (2.58)\sqrt{\frac{(0.12).(0.88)}{400}} \leq P \leq (0.12) + (2.58)\sqrt{\frac{(0.12).(0.88)}{400}}$$

$$0.12 - 0.042 \leq P \leq 0.12 + 0.042 \Leftrightarrow 0.078 \leq P \leq 0.162$$

و بما أن النسبة المدعاة 10% تقع ضمن مجال الثقة ، ومِنْ ثَمّ يمكن قبول ادعاء الباحث عند مستوى 99% من الثقة .

4.5.4): مجال الثقة لتباين متغير عشوائي طبيعي وسيطاه مجهولان:

إذا كان $\bar{x}$ و S^2 هما متوسط و تباين عينة حجمها n لمتغير عشوائي له التوزيع الطبيعي $N\ (\mu, \sigma^2)$ حيث رأينا أن المتغير العشوائي $\varkappa^2 = \frac{(n-1)S^2}{\sigma^2}$ له التوزيع كاي– مربع بـ $\gamma = n - 1$ درجة من الحرية .

و بتبديل بـ x^2 بما يساويها في العلاقة الاحتمالية الصحيحة الآتية :

$$\mathrm{p}\left[\varkappa^2{}_{\alpha/2}(n-1) \le \varkappa^2 \le \varkappa^2{}_{1-\alpha/2}(n-1)\right] = 1 - \alpha$$

أي :

$$\mathrm{p}\left[\varkappa^2{}_{\alpha/2}(n-1) \le \frac{(n-1)S^2}{\sigma^2} \le \varkappa^2{}_{1-\alpha/2}(n-1)\right] = 1 - \alpha$$

و بإصلاح هذه العلاقة نجد أن :

$$\mathrm{p}\left[\frac{(n-1)S^2}{\varkappa^2{}_{1-\alpha/2}(n-1)} \le \sigma^2 \le \frac{(n-1)S^2}{\varkappa^2{}_{\alpha/2}(n-1)}\right] = 1 - \alpha$$

و منه نكون قد عينا $(1 - \alpha)$ مجال ثقة حول تباين متغير عشوائي طبيعي σ^2 من الشكل : $\left[\frac{(n-1)S^2}{\varkappa^2{}_{1-\alpha/2}(n-1)} \quad , \quad \frac{(n-1)S^2}{\varkappa^2{}_{\alpha/2}(n-1)}\right]$

مثال (4-9) : أخذت عينة من 20 شخصاً و قيست أوزانهم ، فوجد أن انحرافها المعياري 9 K.G ، فإذا كان للأوزان التوزيع الطبيعي فأوجد 95% مجال ثقة لتباين الأوزان في المجتمع .

الحل : لدينا $1 - \alpha = 0.95$ منه $\alpha/2 = 0.025$ و يكون

$1 - {}^{\alpha}/_{2} = 0.975$ و من جدول توزيع كاي – تربيع نجد :

$$\varkappa^2{}_{\alpha/2}(\gamma = n-1) = \varkappa^2{}_{0.025}(19) = 8.907$$

$$\varkappa^2{}_{1-\alpha/2}(\gamma = n-1) = \varkappa^2{}_{0.975}(19) = 38.582$$

إذا $1 - \alpha = 0.95$ مجال ثقة حول σ^2 (تباين الأوزان في المجتمع المدروس)

$$\left[\frac{(n-1)S^2}{\varkappa^2{}_{1-\alpha/2}(n-1)} \quad , \quad \frac{(n-1)S^2}{\varkappa^2{}_{\alpha/2}(n-1)}\right] = \left[\frac{(19)(81)}{38.582} \quad , \quad \frac{(19)(81)}{8.907}\right]$$

$$= [\,39.89\,,172.785\,]$$

و يكون 95% مجال ثقة حول الانحراف المعياري σ لأوزان الأشخاص في المجتمع المدروس من الشكل :

$$\left[\sqrt{\frac{(n-1)S^2}{\varkappa^2{}_{1-\alpha/2}(n-1)}} \quad , \sqrt{\frac{(n-1)S^2}{\varkappa^2{}_{\alpha/2}(n-1)}}\right] = \left[\sqrt{39.89}\,,\sqrt{172.785}\right]$$

$$= [\,6.32\,,13.15\,]$$

6.4) : اختبار الفرضيات :

1.6.4: **تمهيد** : للإحصاء أهمية بالغة في اتخاذ القرار في المواقف التي تخضع للمصادفة ، و يحتاج الأمر لاتخاذ قرار عقلاني مع تقدير كمي لحجم المخاطرة ، و بذلك فإن الإحصاء هو الفن في اتخاذ القرارات الحاسمة في المواقف الصعبة ، و تحتل نظرية التقدير و نظرية اختبار الفرضيات مكانة الصداره، و رأينا في الفقرات السابقة التقدير النقطي و التقدير المجالي و في هذه الفقرة سندرس اختبار الفرضيات .

تعريف (4-12) : الفرضية الإحصائية : هي أيّ مقولة أو إفادة أو تخمين تتعلق بوسطاء المجتمع الإحصائي أو بشكل توزيعه و تحتمل الصحة أو الخطأ .

إذ إن صحة الفرضية أو خطأها لا يمكن معرفته بدقة إلا إذا تناولت الدراسة جميع عناصر المجتمع ، و هذا في معظم الحالات غير عملي ، لذا نختار عينة عشوائية من هذا المجتمع ، و اعتماداً على المعلومات التي تحويها العينة يتم القرار إذا كانت الفرضية الإحصائية صحيحة أم خاطئة .

و من الواضح أن العينة التي تتناقض مع الفرضية تقودنا إلى رفض الفرضية ، بينما إذا دعمت هذه المعطيات الفرضية قبلناها ، و يجب أن نشير هنا إلى أن قبول الفرضية الإحصائية ليس إلا نتيجة لعدم كفاية رفضها ، و لا يعني أنها بالضرورة صحيحة . ومِنْ ثَمّ إن رفض الفرضية معناه أن نقرر بأنها خاطئة ، بينما قبولنا الفرضية يعني أننا لم نجد الأسباب الكافية لرفضها ، و هذا يتطلب من الإحصائي أن يضع الفرضية المضادة لتلك التي يعتقد صحتها على أمل أن تقود طرائق الاختبار الإحصائي لها إلى رفضها .

و الفرضية التي يصوغُها الإحصائي بأمل رفضها تدعى الفرضية الابتدائية (الأساسية – العدم) H_0 .

إن رفض H_0 يقود إلى قبول فرضية بديلة يرمز لها بـ H_1 .

فمثلاً إذا رغبنا أن نقرر أن معدل طلاب مدرسة معينة μ_1 أعلى من معدل طلاب مدرسة مجاورة μ_2 ، فيمكن أن نصوغ الفرضية H_0 الآتية :

$H_0 : \mu_1 - \mu_2 = 0$ أي أن H_0 هي فرضية عدم وجود فرق بين المعدلين، وبشكل عام فإن بدائل الفرضية $H_0 : \mu_1 = \mu_2$ هي:

$H_1 : \mu_1 < \mu_2$ ، $H_1 : \mu_1 > \mu_2$ ، $H_1 : \mu_1 \neq \mu_2$.

كذلك الأمر إذا أردنا أن نظهر بأن نسبة الإصابة بمرض السكري في مدينة معينة مختلفة عن P_0 ، فيمكن أن نصوغ الفرضية الأساسية $H_0 : P = P_0$ مقابل الفرضية البديلة $H_0 : P \neq P_0$.

2.6.4) : اختبار الفرضيات الإحصائية :

من الفقرة السابقة يُلحظ بأنه لاختبار صحة الفرضية H_0 مقابل الفرضية البديلة H_1 تخص كل منها متغيراً عشوائياً x ، لذلك علينا أن نختار دالة سندعوها دالة الاختبار ، و هي إحصاء للعينة السابقة تساعدنا على اتخاذ القرار، و بما أن دالة الاختبار إحصاء فهي متغير عشوائي تتغير قيمتها من عينة لأخرى ، لهذا يجب علينا معرفة دالة توزيعها الاحتمالي ، و ذلك لكي نتمكن من اتخاذ قرار بشأن الفرضية الإحصائية ، و هنا سنستخدم توزيعات الإحصاءات التي تعرفنا عليها في بداية هذا الفصل أو الفصل السابق ، ثم نقوم بتجزئة مجال تحولات دالة الاختبار إلى منطقتين ، تسمى إحدهما منطقة الرفض للفرضية H_0 أو تدعى المنطقة الحرجة ، و تسمى المنطقة الأخرى منطقة القبول ، و عادة تحدد منطقة الرفض بتلك المنطقة التي تتألف من قيم دالة الاختبار قليلة الاحتمال عندما تكون H_0 صحيحة ، و تحدد منطقة القبول بالمنطقة التي تتألف من قيم دالة الاختبار كثيرة الحدوث إذا كانت H_0 صحيحة . و يكون القرار رفض الفرضية H_0 و قبول H_1 إذا وقعت قيمة دالة الاختبار في منطقة الرفض ، و يكون القرار بعدم رفض الفرضية H_0 إذا وقعت قيمة دالة الاختبار في منطقة القبول ، و بذلك عند اتخاذ القرار برفض أو قبول H_0 يمكن أن يرتكب خطأ ، و هذا الخطأ على نوعين :

a. **الخطأ من النوع الأول** : و هو الخطأ الناجم عن رفض الفرضية H_0 فيما هي صحيحة و نرمز لاحتمال الوقوع في الخطأ من النوع الأول بـ α ، و يدعى α عندئذ بالخطأ من النوع الأول أو حجم منطقة الرفض أو مستوى أهمية الاختبار أو مستوى المعنوية أو مستوى الدلالة الإحصائية .

b. **الخطأ من النوع الثاني** : و هو الخطأ الناجم عن قبول H_0 و هي خاطئة و يرمز عادة لاحتمال الوقوع في الخطأ من النوع الثاني بـ β ، و يدعى β هذا بالخطأ من النوع الثاني ، و يدعى $\rho = 1 - \beta$ بقوة اختبار الفرضية H_0 مقابل H_1 الموافق لمنطقة الرفض المحددة بالحجم α.

ملاحظة : نرمز لمنطقة الرفض بـ C و لمنطقة القبول بـ $\grave{C}$

و نحاول أن نجد أفضل اختبار ، و هو الذي تكون فيه α، β أصغر ما يمكن ، و لكن وجد أنه لا يمكن تصغير الخطأيْنِ α ، β في آن واحد إلا إذا كبرنا حجم العينة ، و هذا غير ممكن دوماً . لذلك اعتمد الإحصائيون إستراتيجية تحديد الخطأ من النوع الأول α ، ثم البحث عن منطقة الرفض التي تجعل من الخطأ من النوع الثاني أصغر ما يمكن .

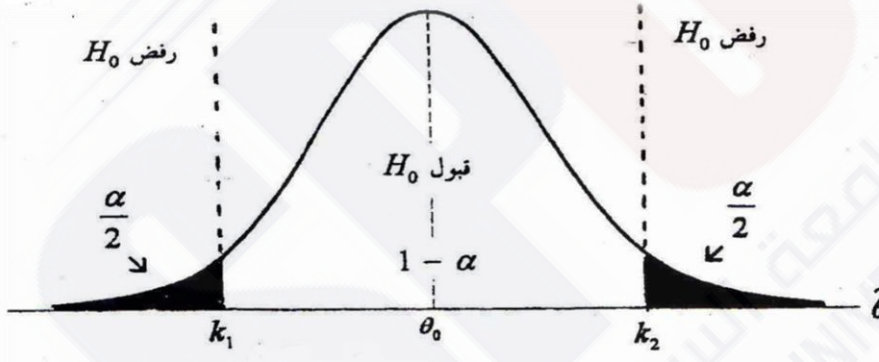
3.6.4) : الاختبارات الوسيطية :

هنا سيتم اختبار الفرضيات الإحصائية المتعلقة بوسطاء المجتمعات الإحصائية علماً أن طبيعة توزيعات تلك المجتمعات معروفه ، فإذا كان θ وسيطاً لمجتمع إحصائي ، و كانت لدينا الفرضية موضع الاختبار : $H_0 : \theta = \theta_0$ فإن الفرضية البديلة H_1 تكون على ثلاثة أنواع :

a. الفرضية البديلة ذات الطرفين (من جانبيين) : أي $H_1 : \theta \neq \theta_0$ و هنا من أجل α مستوى المعنوية و كانت دالة الاختبار ، فإننا نضع نصف α في كل طرف من طرفي توزيع دالة الاختبار .

فمثلاً إذا كان الشكل (3-4) يمثل منحني الكثافة لدالة الاختبار حول θ_0 فإننا نحدد K_1 ، K_2 بحيث يكون

$$P\,[\theta_0 < K_1\,] = P\,[\theta_0 > K_2\,] = {}^{\alpha}/_{2}$$



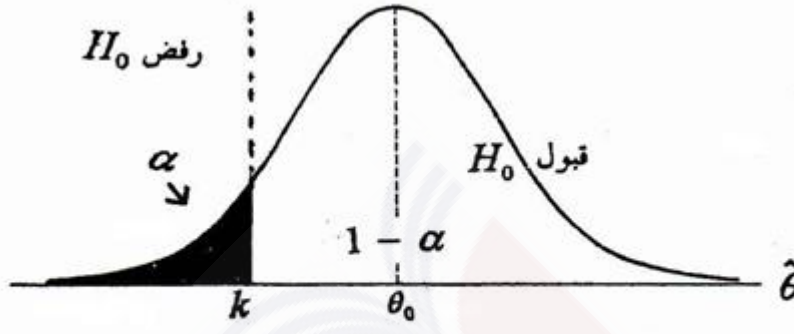
الشكل(3.4)

المنطقة الحرجة لـ $H_1 : \theta \neq \theta_0$

ونقبل الفرضيةH_0 إذا وقعت قيمة دالة الاختبار θ_0 المحسوبة من العينة بين K_1 ، K_2 و نرفض H_0 إذا وقعت قيمة θ_0 في المنطقة المظللة التي تمثل منطقة الرفض.

b. الفرضية البديلة من الطرف الأيسر الأدنى : أي $H_1 : \theta < \theta_0$

و هنا نضع α في الطرف الأيسر من توزيع دالة الاختبار θ_0 كما في الشكل (4-4)



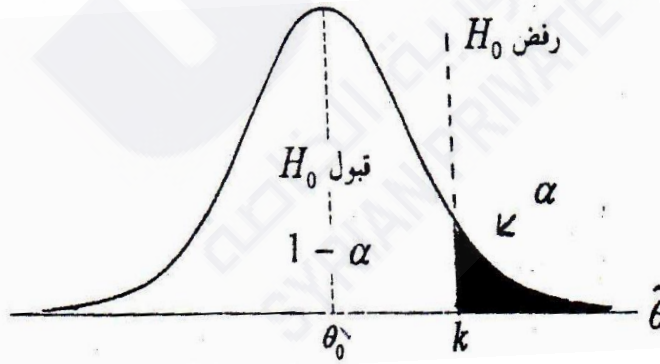
الشكل 4.4

المنطقة الحرجة لـ $H_1 : \theta < \theta_0$

حيث تحدد K بحيث يكون $P\,[\theta_0 < K\,] = \alpha$ و نقبل H_0 إذا كانت قيمة إحصاء الاختبار θ_0 أكبر من K ، و نرفض H_0 إذا كانت القيمة أصغر من K.

c. الفرضية البديلة ذات الطرف الأيمن (الذيل الأيمن) الأعلى أي:

$H_1 : \theta > \theta_0$ و هنا نضع α في الطرف الأيمن من توزيع دالة الاختبار θ_0 كما في الشكل (5-4) الذي يمثل مثلاً منحني دالة الكثافة لإحصاء الاختبار θ_0



الشكل 5.4

المنطقة الحرجة لـ $H_1 : \theta > \theta_0$

حيث تحدد K بحيث يكون $P[\theta > K] = \alpha$ و نقبل H_0 إذا كانت قيمة إحصاء الاختبار θ_0 أصغر من K و نرفض H_0 إذا كانت القيمة دالة الاختبار أكبر من K .

4.6.4) : اختبارات حول المتوسط :

1.4.6.4) :اختبارات حول متوسط مجتمع طبيعي تباينه σ^2 معلوم:

ليكن $X \sim N(\mu, \sigma^2)$ و σ^2 معلوم و لتكن $X_1, X_2, \dots, X_n$ عينة عشوائية لمتغير عشوائي X متوسطها $\overline{X}$ و يكون $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

حيث $\overline{X}$ يصلح أن يكون إحصاء الاختبار لفرضيات تتعلق بالمتوسط μ :

وسوف نميّز ثلاث حالات :

1- من أجل اختبار $H_0 : \theta = \theta_0$ ضد الفرضية $H_1 : \theta \neq \theta_0$ و عند مستوى المعنوية α و الاختبار من الطرفين ستكون قيمة إحصائية الاختبار تحت صحة H_0 : $Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

و من أجل α مستوى المعنوية و الاختبار من الطرفين و التوزيع طبيعي معياري لإحصائية الاختبار ، ستكون منطقة القبول لـ H_0 هي :

$$\left[Z_{\alpha/2}, Z_{1-\alpha/2}\right]$$

و منطقة الرفض لـ H_0 : $]-\infty, Z_{\alpha/2}[\vee]Z_{1-\alpha/2}, +\infty[$

(حيث $\vee$ تعني أو).

مثال (4-10): إذا علمنا أن وزن قطعة غذاء من نوع معين لها التوزيع الطبيعي بانحراف معياري 5 G ، و لدى معاينة عينة تحوي 16 قطعة غذاء من هذا النوع ، تبين أن متوسط وزنها هو 244 G و المطلوب: عند مستوى المعنوية $\alpha = 0.05$ اختبر صحة الفرضية $H_0 : \mu = 250$ ضد الفرضية $H_1 : \mu \neq 250$

الحل : هنا الاختبار من الطرفين و يكون من أجل

$$\alpha = 0.05 \Rightarrow 1 - \alpha/2 = 0.975 \Rightarrow Z_{0.975} = 1.96$$

و قيمة إحصائية الاختبار هنا تحت صحة H_0

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{244 - 250}{5/\sqrt{16}} = -4.8$$

ومن أجل $\alpha = 0.05$ و الاختبار من الطرفين و التوزيع طبيعي معياري لإحصائية الاختبار ، ستكون منطقة القبول للفرضية H_0 :

$$\left[Z_{\alpha/2}, Z_{1-\alpha/2}\right] = [-1.96 \ , \ 1.96]$$

و منطقة الرفض : $]-\infty, -1.96[\ V \ ]1.96, +\infty[$

و بمقارنة $Z_0 = -4.8$ مع مناطق الرفض و القبول نجد أن Z_0 تنتمي لمنطقة الرفض من جهة اليسار و منه نرفض H_0 و نقبل H_1 أي $\mu \neq 250$ و كون الرفض من جهة اليسار فإن $\mu < 250$.

2- لاختبار $H_0 : \mu = \mu_0$ ضد الفرضية $H_1 : \mu < \mu_0$ عند مستوى المعنوية α، فإننا نرفض H_0 عندما تكون قيمة إحصائية الاختبار Z_0 تنتمي

للمنطقة $]-\infty , Z_{\alpha}[$ و نقبل H_0 إذا كانت Z_0 تنتمي للمنطقة $]Z_{\alpha} , +\infty[$

3- لاختبار $H_0 : \mu = \mu_0$ ضد الفرضية $H_1 : \mu > \mu_0$ عند مستوى المعنوية α ، فإننا نرفض H_0 عندما تكون إحصائية الاختبار Z_0 تنتمي للمنطقة $]Z_{1-\alpha} , +\infty[$ و نقبل H_0 إذا كانت Z_0 تنتمي للمنطقة $]-\infty , Z_{1-\alpha}[$.

مثال (4-11): تدل الدراسات أن وزن الطفل عند الولادة يخضع للتوزيع الطبيعي بمتوسط 3.2 K.G و انحراف معياري 0.6 K.G و عند معاينة 16 طفلاً أمهاتهم خضعن لنظام غذائي جديد تبين أن متوسط الوزن بهذه العينة 3.5 K.G هل يمكننا القول : إنَّ النظام الغذائي الجديد قد أسهم بتحسين متوسط وزن الطفل عند مستوى الأهمية $\alpha = 0.025$.

الحل : هنا نختبر $H_0 : \mu = 3.2$ مقابل الفرضية $H_1 : \mu > 3.2$

و عند $\alpha = 0.025$ و الاختبار من اليمين

إن إحصائية الاختبار هنا $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0 , 1)$

و قيمة إحصائية الاختبار تحت صحة H_0 ستكون :

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{3.5 - 3.2}{0.6/\sqrt{16}} = 2$$

و عند مستوى معنوية $\alpha = 0.025$ و الاختبار من الطرف الأيمن و التوزيع طبيعي معياري لإحصائية الاختبار ، ستكون منطقة القبول H_0 من الشكل :

$]-\infty, Z_{1-\alpha}] =]-\infty, Z_{0.975}] = [-\infty, 1.96]$ ومنطقة الرفض H_1 من الشكل $]1.96, +\infty[$

و بمقارنة $Z_0 = 2$ مع مناطق الرفض و القبول نجد أن Z_0 تنتمي لمنطقة رفض H_0 أي إننا نقبل H_1 أي $\mu > 3.2$ ، و هذا يعني أن النظام الغذائي الجديد قد أسهم في تحسين وزن الطفل عند ولادته .

ملاحظات:

1. إذا كان المجتمع غير طبيعي و $n \geq 30$ نطبق مبرهنة النهاية المركزية ليكون الإحصاء : $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \approx N(0, 1)$

2. في الحالة التي يكون فيها تباين المجتمع σ^2 مجهولاً (سواء كان المجتمع طبيعياً أم غير طبيعي) ، و من أجل ($n \geq 30$) فإن تباين العينة S^2 هو تقدير جيد لـ σ^2 ، و لذلك يمكن الاستعاضة عن σ^2 بقيمة S^2 من العينة و يجرى اختبار الفرضية حول المتوسط تماماً كما سبق .

مثال (4-12) : إذا كان متوسط المدة اللازمة لاستجابة المريض لدواء مهدئ هو 5 دقائق ، اقترح دواءٌ جديدٌ ، و جُرِّبَ على 64 مريضاً ، فكان متوسط المدة اللازمة لاستجابة المريض هو 4.7 دقيقة و بانحراف معياري قدره 1.2 دقيقة. فهل الدواء الجديد المقترح أفضل من القديم عند مستوى الأهمية $\alpha = 0.01$ ؟

الحل:

هنا نختبر صحة الفرضية $H_0 : \mu = 5$ مقابل الفرضية البديلة

$H_1 : \mu < 5$ و من أجل $\alpha = 0.01$ يكون $1 - \alpha = 0.99$

حيث $Z_{0.99} = 2.33$ و إحصائية الاختبار : $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N\,(0,1)$

و قيمة إحصائية الاختبار تحت صحة H_0 : $Z_0 = \frac{4.7 - 5}{1.2/8} = -2$

و منطقة قبول H_0 : $[\ Z_\alpha, +\infty\ [= [\ -2.33\ , +\infty\ [$

وبملاحظة أن $Z_0 \in [\ -2.33\ , +\infty\ [$ فإننا نقبل H_0 ، ومِنْ ثَمّ فالنظام الجديد ليس أفضل من النظام القديم للتهدئة .

مثال (4-13) :

تبين من عينة عشوائية من الحجم 100 متوفى أن متوسط العمر لهؤلاء كان 71.8 سنة بانحراف معياري 8.9 سنة . فهل يشير هذا إلى أن متوسط العمر الآن أكبر من 70 سنة بمستوى الأهمية $\alpha = 0.05$؟

الحل : لدينا الفرضية الابتدائية (الأساسية أو فرضية العدم) : $H_0 : \mu = 70$ و الفرضية البديلة $H_1 : \mu > 70$ و مستوى الأهمية (الدلالة الإحصائية أو مستوى المعنوية) : $\alpha = 0.05$

إن إحصائية الاختبار هنا $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N\,(0,1)$ و قيمة إحصائية الاختبار تحت صحة H_0: $Z_0 = \frac{71.8 - 70}{8.7/\sqrt{100}} = 2.02$

و عند مستوى الأهمية $\alpha = 0.05$ و الاختبار من جهة اليمين و التوزيع طبيعي معياري لإحصائية الاختبار ستكون منطقة قبول H_0 هي :

$$]\ -\infty\ ,\ Z_{1-\alpha/2}[\ =\ ]-\infty\ ,Z_{0.95}\ ]\ =\ ]-\infty\ ,1.65\ ]$$

و بمقارنة $Z_0 = 2.02$ مع منطقة قبول H_0 نجد أن Z_0 خارج هذه المنطقة .

أي تقع في منطقة الرفض ، و منه نرفض H_0 و نقبل H_1 ، أي إن متوسط العمر الآن أكبر فعلاً من 70 سنة و بثقة 95% .

2.4.6.4) : اختبارات حول متوسط مجتمع طبيعي تباينه σ^2 مجهول:

رأينا سابقاً أنه إذا كانت $X_1, X_2\ , \dots , X_n$ عينة عشوائية من X حيث $X \sim N\ (\mu, \sigma^2)$ متوسطها $\bar{x}$ و تباينها S^2 , σ^2 مجهول أن الإحصاء ستيودنت $T = \frac{\bar{x} - \mu}{S/\sqrt{n}} \sim T_{(\gamma = n-1)}$; $(n < 30)$

و هذا الإحصاء يستخدم لاختبار فرضيات حول μ متوسط مجتمع طبيعي تباينه σ^2 مجهولاً ، و يعتمد كإحصاء للاختبار وفق :

1- إذا كانت $H_0 : \mu = \mu_0$ مقابل الفرضية البديلة $H_1 : \mu \neq \mu_0$ عند مستوى المعنوية α

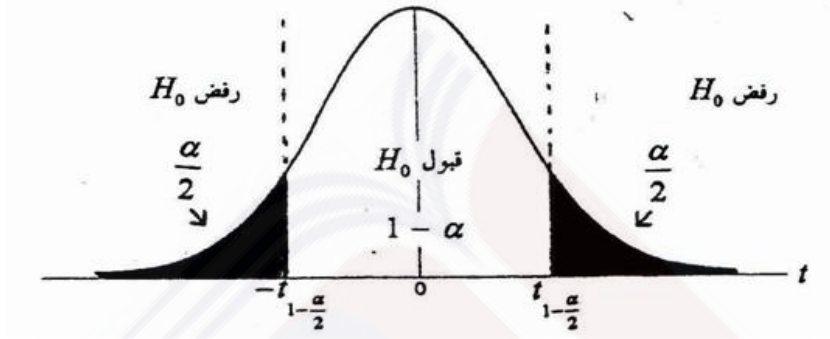
إن قيمة إحصائية الاختبار تحت صحة H_0 تكون $T_0 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$.

و من أجل مستوى المعنوية α و الاختبار من الطرفين و التوزيع لإحصائية الاختبار لستيودنت بـ $\gamma = n - 1$ درجة من الحرية

تكون منطقة قبول H_0 : $[\ t_{\alpha/2}(\gamma), t_{1-\alpha/2}(\gamma)\ ]$

و تكون منطقة رفض H_0 : $]-\infty\ , t_{\alpha/2}(\gamma)[\ \vee\ ]t_{1-\alpha/2}(\gamma)\ \ ,+\infty\ [$

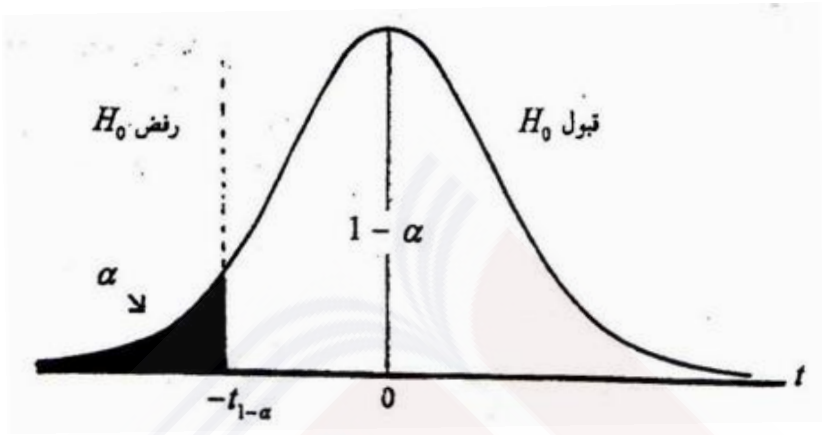
ثم نقارن T_0 مع مناطق الرفض و القبول لاتخاذ القرار المناسب كما رأينا في الفقرة السابقة و كما في الشكل (4-6)



الشكل (6.4)

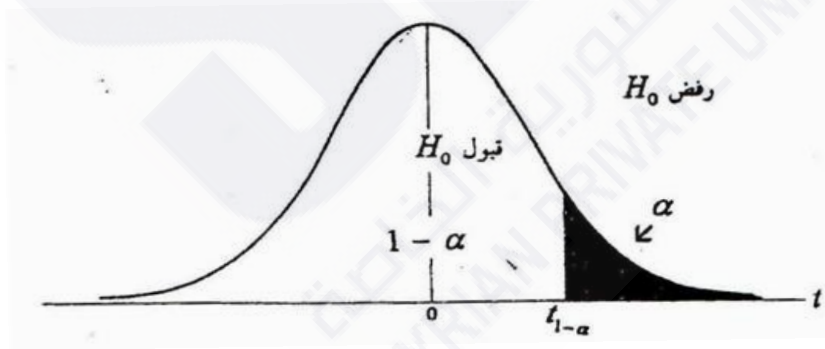
2- إذا كانت $H_0 : \mu = \mu_0$ مقابل الفرضية البديلة $H_1 : \mu < \mu_0$ عند مستوى المعنوية α والاختبار من الطرف الأيسر إن منطقة الرفض H_0 هي $]-\infty, t_\alpha(\gamma)[$ و منطقة القبول $[t_{1-\alpha}(\gamma), +\infty[$ كما في الشكل (7-4) حيث:

$$t_\alpha(\gamma) = -t_{1-\alpha}(\gamma)$$



الشكل (7.4)

3- إذا كانت $H_0 : \mu = \mu_0$ مقابل الفرضية البديلة $H_1 : \mu > \mu_0$ عند مستوى المعنوية α و الاختبار من الطرف الأيمن إن منطقة الرفض H_0 هي $[\ t_{1-\alpha}(\gamma)\ ,+\infty[$ ومنطقة القبول $]-\infty\ ,t_{1-\alpha}(\gamma)[$ كما في الشكل (8-4)



الشكل (8.4)

مثال(4-14): تدعي شركة لإنتاج البطاريات التي تستخدم في الأجهزة الطبية بأن عمر البطارية من إنتاجها له التوزيع الطبيعي بمتوسط 3 سنوات أخذت عينة عشوائية من إنتاج هذه الشركة تحوي 6 بطاريات ، فكانت أعمارها بالسنوات كما يأتي :

3.8 1.9 2.9 0.9 4.0 3.5

هل نستنتج بأن الشركة تبالغ في ادعائها بالنسبة لمتوسط عمر البطاريات التي تنتجها عند مستوى المعنوية $\alpha = 0.01$ ؟.

الحل: هنا سنختبر $H_0 : \mu = 3$ مقابل الفرضية البديلة $H_1 : \mu < 3$ عند مستوى المعنوية $\alpha = 0.01$ و الاختبار من الطرف الأيسر

$\gamma = n - 1 = 6 - 1 = 5$, $1 - \alpha = 0.99$, $\alpha = 0.01$(من جدول ستيودنت)

$$t_{0.01}(5) = -t_{0.99}(5) = -3.365$$

و بحساب $\bar{x}$ ، S^2 من العينة المعطاة نجد أن

$$\bar{x} = 2.833 \quad , \quad S = 1.213$$

و إحصائية الاختبار هنا ستيودنت $T = \frac{\bar{x} - \mu}{S/\sqrt{n}} \sim T_{(\gamma = n-1)}$

و تحت صحة H_0 تكون قيمة إحصائية الاختبار

$$T_0 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{2.833 - 3}{1.213/\sqrt{6}} = -0.337$$

و من أجل مستوى المعنوية $\alpha = 0.01$ و الاختبار من الطرف الأيسر و التوزيع لإحصائية الاختبار هو ستيودنت ب $\gamma = 5$ درجة من الحرية تكون

منطقة قبول H_0 : $[\, t_\alpha(\gamma)\, , +\infty\, [\, = [\, -3.365\, , +\infty[$

و منطقة رفض H_0 : $]-\infty\, , t_\alpha(\gamma)[\, =]-\infty\, , -3.365\, [$

وبمقارنة $T_0 = -0.337$ مع مناطق الرفض و القبول نجد أن T_0 تقع في منطقة القبول ، ومِنْ ثَمّ نقبل H_0 ، و هذا يعني أن الشركة لم تبالغ في ادعائها.

مثال (4-15) : يمثل البيان الآتي إنتاج عشر شجيرات لصنف جديد من الخضار مقيساً بالكيلوغرام :

2.2	2.1	3.0	4.1	3.9	2.7	3.1	2.3	4.2	1.9

فإذا علمنا أن قياسات الإنتاج في مجتمع شجيرات هذا الصنف من الخضار له توزيع طبيعي بمتوسط $\mu = 3K.G$ فهل نستنتج بأن إنتاج الصنف الجديد أفضل من القديم بمستوى أهمية $\alpha = 0.05$.

الحل: هنا سنختبر $H_0 : \mu = 3$ مقابل الفرضية البديلة $H_1 : \mu > 3$ عند مستوى المعنوية $\alpha = 0.05$ فإن $1 - \alpha = 0.995$ و الاختبار من الطرف الأيمن $\gamma = n - 1 = 10 - 1 = 9$ من جدول ستيودنت

$$t_{1-\alpha}(\gamma) = t_{0.995}(9) = 3.250$$

و بحساب $\bar{x}$ ، S^2 من البيان المفروض نجد أن $S = 0.862$, $\bar{x} = 2.95$

و قيمة إحصائية الاختبار تحت صحة H_0 تكون:

$$T_0 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} = \frac{2.95 - 3}{0.862/\sqrt{10}} = -0.183$$

و من أجل مستوى المعنوية $\alpha = 0.05$ و الاختبار من الطرف الأيمن و التوزيع لإحصائية الاختبار هو ستيودنت بـ $\gamma = 9$ درجة من الحرية تكون منطقة رفض H_0 من الشكل:

$$\left[t_{1-\alpha}(\gamma)\,, +\infty\,[=]t_{0.995}(9)\,, +\infty\right] =]\,3.250\,, +\infty[$$

و منطقة قبول H_0: $]-\infty\,, t_{0.995}(9)\,] =]-\infty\,,\ 3.250\,[$

وبمقارنة $T_0 = -0.183$ مع مناطق الرفض و القبول نجد أنها تنتمي لمنطقة القبول H_0 أي $\mu = 3$ ، و هذا يعني أن الصنف الجديد ليس أفضل من الصنف القديم في الإنتاج .

5.6.4) اختبارات حول النسبة في المجتمع (وسيط متغير برنولي P):

إن مثل هذه الاختبارات مرغوبة في العديد من المجالات ، فمثلاً رجال السياسة يهتمون بمعرفة نسبة المقترعين لصالح مرشح معين ، و الطبيب يهتم بمعرفة نسبة نجاح حملة تلقيح معينة أو طريقة علاج مرض معين ... إلخ . ولقد رأينا مسبقاً أن $\hat{p} = \bar{x} = \frac{y}{n}$ مقدر منصف لـ p . حيث y عدد النجاحات و $\bar{x}$ متوسط العينة المسحوبة من مجتمع برنولي و من أجل $n \geq 30$ يكون للمتغير العشوائي

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}} = \frac{y - np}{\sqrt{n\hat{p}\,(1-\hat{p})}} \approx N\,(0\,,1)$$

و هو إحصاء الاختبار ، لاختبار فرضيات حول p أي $H_0 : \mathrm{p} = \mathrm{p}_0$

و تكون قيمة إحصاء الاختبار تحت صحة H_0 :

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{y - np_0}{\sqrt{np_0(1-p_0)}}$$

و يمكن أن نميز ثلاث حالات (حسب نوعيه الاختبار البديل) :

1. من أجل $H_0 : p = p_0$ مقابل الفرضية البديلة $H_1 : p \neq p_0$ عند مستوى المعنوية α ، هنا الاختبار من الطرفين ستكون منطقة قبول H_0 من الشكل : $\left[Z_{\alpha/2} \ , \ Z_{1-\alpha/2}\right]$ و منطقة الرفض لـ H_0 من الشكل :

$$]-\infty, \ Z_{\alpha/2}[\ V \ ]Z_{1-\alpha/2}, +\infty[$$

و نقارن Z_0 مع مناطق الرفض و القبول لاتخاذ القرار المناسب كما رأينا مسبقاً.

2. من أجل $H_0 : p = p_0$ مقابل الفرضية البديلة $H_1 : p < p_0$ عند مستوى المعنوية α ، هنا الاختبار من الطرف الأيسر ستكون منطقة قبول لـ H_0 من الشكل : $[Z_\alpha \ , +\infty[$ و منطقة الرفض لـ H_0 من الشكل :

$$]-\infty, \ Z_\alpha[$$

و نقارن Z_0 مع مناطق الرفض و القبول لاتخاذ القرار المناسب .

3. من أجل $H_0 : p = p_0$ مقابل الفرضية البديلة $H_1 : p > p_0$ عند مستوى المعنوية α، هنا الاختبار من الطرف الأيمن ستكون منطقة قبول لـ H_0 من الشكل : $]-\infty \ , Z_{1-\alpha}]$ و منطقة الرفض لـ H_0 من الشكل : $]Z_{1-\alpha}, +\infty[$

و نقارن Z_0 مع مناطق الرفض و القبول لاتخاذ القرار المناسب .

مثال (4-16): تدعي شركة لصناعة الأدوية ، بأن أحد أدويتها الخاصة بمعالجة أحد الأمراض، تحدث استجابة خلال فترة قصيرة لـ 0.80 من الأشخاص المصابين بالمرض المدروس، و لاختبار صحة هذا الادعاء أخذت عينة عشوائية من 150 مريضاً بهذا المرض و جرب عليهم هذا الدواء ، فوجد أن 110 من المرضى استجابوا للمعالجة خلال الفترة المفروضة ، فهل هذه النتائج تدعم صحة ادعاء الشركة بمستوى $\alpha = 0.01$ من الأهمية ؟

الحل: هنا نختبر الفرضية $H_0 : p = 0.80$ مقابل $H_0 : p \neq 0.80$ و عند مستوى المعنوية

$$\alpha = 0.01 \Rightarrow \alpha/2 = 0.005 \Rightarrow 1 - \alpha/2 = 0.995 \Rightarrow$$

$$Z_{0.995} = 2.58 \quad , Z_{0.005} = -2.58$$

و إحصائية الاختبار : $Z = \frac{y-np}{\sqrt{np(1-p)}} \approx N(0,1)$

و تحت صحة H_0 تكون قيمة إحصائية الاختبار

$$Z_0 = \frac{110-(150).(0.80)}{\sqrt{(150).(0.80).(0.20)}} = -2.04$$

و من أجل $\alpha = 0.01$ و الاختبار من الطرفين و التوزيع لإحصائية الاختبار هو طبيعي معياري :

ستكون منطقة رفض H_0 : $]-\infty, -2.58[\; V \;]2.58, +\infty[$

و تكون منطقة قبول H_0 : $[-2.58 \; , \; 2.58]$

و بمقارنة $Z_0 = -2.04$ مع مناطق الرفض و القبول نجد أن Z_0 تنتمي لمنطقة القبول ، و منه نقبل H_0 و نرفض H_1 ، و هذا يعني أن ادعاء الشركة صحيحٌ.

مثال (4-17) : إذا كان 54 % من إجمالي السكان يفضلون السكن داخل المدينة ، و نظراً للظروف البيئية من تلوث و ضجيج، يعتقد أن تغيراً قد طرأ على هذه النسبة ، و لاختبار هذا التغيير ، تم سؤال عينة عشوائية من 1000 شخص من سكان هذه المدينة فكان منهم 480 ممن يفضلون السكن في المدينة و 520 يفضلون السكن في ريف المدينة ، و المطلوب عند مستوى المعنوية $\alpha = 0.05$ ، هل نسبة من يفضلون السكن في المدينة أصبحت أقل مما كانت عليه في البداية ؟

الحل : هنا ستختبر $H_0 : p = 0.54$ مقابل $H_1 : p < 0.54$ عند مستوى المعنوية $\alpha = 0.05$ و الاختبار من الطرف الأيسر حيث $1 - \alpha = 0.95$ و إحصائية الاختبار

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p\,q}{n}}} \sim N(0,1) \qquad , \ (q = 1 - p) \ \text{حيث}$$

و قيمة إحصائية الاختبار تحت صحة H_0 :

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.48 - 0.54}{\sqrt{\frac{(0.54).(0.46)}{1000}}} = -3.82$$

و من أجل مستوى الأهمية $\alpha = 0.05$ و الاختبار من الطرف الأيسر و التوزيع لإحصائية الاختبار هو طبيعي معياري، ستكون منطقة رفض H_0 من الشكل : $]-\infty, Z_\alpha[=]-\infty, Z_{0.05}[=]-\infty, -1.65[$

و منطقة قبول H_0 من الشكل : $[\,Z_\alpha\,,+\infty\,[\,=\,[\,-1.65\,,+\infty\,[$

و بمقارنة $Z_0 = -3.82$ مع مناطق الرفض و القبول نجد أن Z_0 تنتمي لمنطقة رفض H_0 ، ومِنْ ثَمّ نرفض H_0 أي $p = 0.54$ ونقبل H_1 أي $p <$ 0.54 أي هناك تراجع بنسبة من يفضلون السكن في المدينة .

6.6.4) اختبارات حول تباين مجتمع طبيعي وسيطاه مجهولان :

رأينا في بداية هذا الفصل أنه من أجل عينة عشوائية $X_1, X_2\ , \dots, X_n$ من X حيث $X \sim N\ (\mu, \sigma^2\ )$ ومتوسط العينة $\overline{X}$ وتباينها S^2 سيكون للإحصاء

$$\varkappa^2 = \frac{(n-1)S^2}{\sigma^2}$$

توزيع كاي– مربع بـ $\gamma = n - 1$ درجة من الحرية ، و لذلك سنستخدم هذا الاحصاء كدالة اختبار لفرضيات تتعلق بـσ^2 و سنميز الحالات الآتية :

1. من أجل اختبار $H_0 : \sigma^2 = {\sigma_0}^2$ مقابل الفرضية البديلة $H_1 : \sigma^2 \neq {\sigma_0}^2$ عند مستوى الأهمية α ، و هنا الاختبار من الطرفين ستكون إحصائية الاختبار

$$\varkappa^2 = \frac{(n-1)S^2}{\sigma^2} \sim \varkappa^2_{\ (\gamma=n-1)}$$

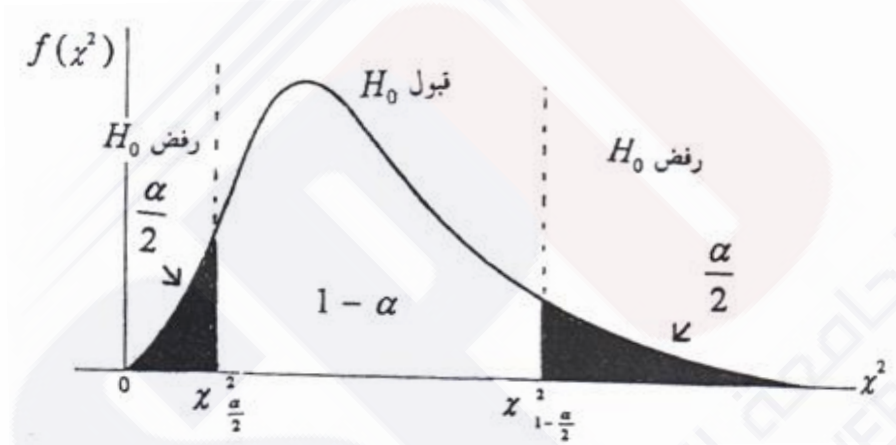
و ستكون قيمة إحصائية الاختبار تحت صحة H_0 : $\varkappa_0{}^2 = \frac{(n-1)S^2}{{\sigma_0}^2}$

و من أجل مستوى الأهمية α و الاختبار من الطرفين و التوزيع لإحصائية الاختبار هو كاي– مربع بـ $\gamma = n - 1$ درجة من الحرية ستكون منطقة الرفض لـ H_0 :

$$]\,0\,,\varkappa^2_{\alpha/2}(n-1)[\ \ V\ \ ]\varkappa^2_{1-\alpha/2}(n-1)\,,+\infty\,[$$

و منطقة قبول الفرضية H_0 : $\left[\varkappa^2_{\alpha/2}(n-1)\ \ ,\ \varkappa^2_{1-\alpha/2}(n-1)\right]$

ثم نقارن قيمة إحصائية الاختبار الناتجة ${\varkappa_0}^2$ مع مناطق الرفض و القبول لـ H_0 لاتخاذ القرار المناسب لقبول أو رفض H_0 كما في الشكل (4-9) .



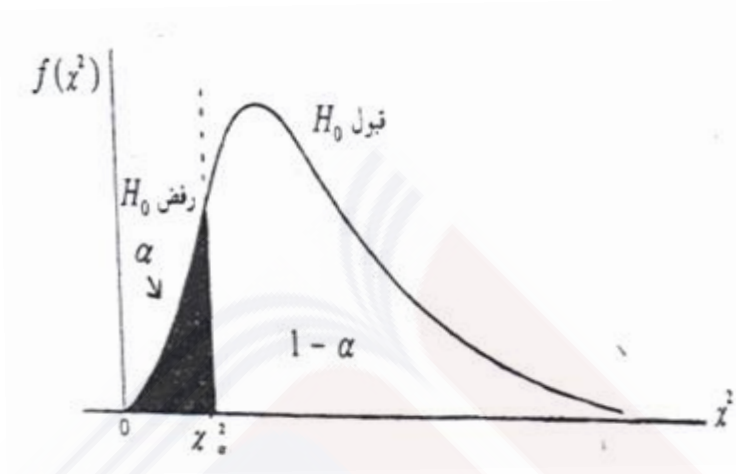
الشكل (9.4)

2. و لاختبار ${\sigma_0}^2 = \sigma^2 : H_0$ مقابل الفرضية البديلة ${\sigma_0}^2 > \sigma^2 : H_1$ عند مستوى الأهمية α و هنا الاختبار من اليسار

ستكون منطقة الرفض لـ H_0 : $]\ 0\ ,x^2_\alpha(n-1)\,[$

و منطقة قبول الفرضية H_0 : $[\ x^2_\alpha(n-1)\ ,+\infty\,[$

كما في الشكل (4- 10) .

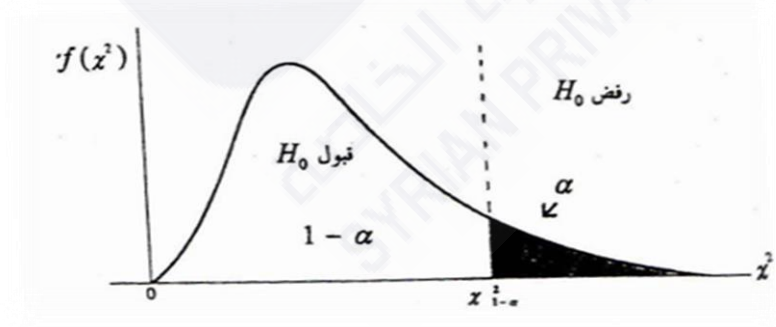


الشكل (10.4)

3. و لاختبار $H_0 : \sigma^2 = \sigma_0{}^2$ مقابل الفرضية البديلة $H_1 : \sigma^2 > \sigma_0{}^2$ عند مستوى الأهمية α و هنا الاختبار من الطرف الأيمن ستكون منطقة الرفض لـ H_0: $[\ \varkappa^2_{1-\alpha}(n-1)\ ,+\infty\ [$

و منطقة قبول الفرضية H_0 : $]\ 0\ ,\varkappa^2_{1-\alpha}(n-1)\ [$

كما في الشكل (4- 11) .



الشكل (11.4)

مثال (4-18) :

ينتج معمل الأدوية نوعاً من العلاج يحتوي على مادة فعالة يجب أن تكون محددة بشكل دقيق . و لدراسة مدى دقة المصنع في إضافة كمية المادة الفعالة إلى كل حبة من حبوب هذا العلاج ، قام المسؤولون في المصنع بتحليل عينة من 30 حبة ، فوجدوا أن الانحراف المعياري لكمية هذه المادة في الحبوب يساوي 1.30M.G، استخدم هذه المعلومات لاختبار صحة الفرضية $H_0 : \sigma^2 = 1.50$ مقابل الفرضية $H_1 : \sigma^2 < 1.50$ عند مستوى الأهمية (المعنوية) $\alpha = 0.05$ ؟

الحل:

بما أن $\alpha = 0.05$ و $\gamma = n - 1 = 29$ و الاختبار من الطرف الأيسر عندئذ:

$$\varkappa_{\alpha}^{2}(n-1) = \varkappa_{0.05}^{2}(29) = 17.7$$

(من جدول كاي- مربع عند $\gamma = 29$ درجة من الحرية) .

و إحصائية الاختبار ستكون $x^2 = \frac{(n-1)S^2}{\sigma^2} \sim x^2_{(\gamma=n-1)}$

وقيمة إحصائية الاختبار تحت صحة H_0 ستكون :

$$x_0{}^2 = \frac{(n-1)S^2}{\sigma_0{}^2} = \frac{(29)\,.\,(1.3)^2}{1.5} = 32.673$$

و من أجل مستوى معنوية $\alpha = 0.05$ و الاختبار من الطرف الأيسر و التوزيع لإحصائية الاختبار هو كاي- مربع بـ $\gamma = 29$ درجة من الحرية ستكون منطقة

رفض H_0 من الشكل: $]0\ ,17.7\ ]$ $[=]0\ ,\varkappa_{\alpha}^{2}(n-1)$ و منطقة قبول H_0 من الشكل : $[\ \varkappa_{\alpha}^{2}(n-1)\ ,+\infty[\ =\ [17.7\ ,+\infty[$

و بمقارنة إحصائية الاختبار الناتجة $x_0{}^2 = 32.673$ مع مناطق الرفض و القبول لـ H_0 نجد أن $x_0{}^2$ تقع في منطقة قبول H_0 ، ومِنْ ثَمّ نقبل H_0 و نرفض H_1 أي $\sigma^2 = 1.50$.

7.4) : تمارين غير محلولة (للقسم العملي) :

1. يمثل البيان التالي إنتاج 10 شجيرات من نوع معين من الخضار مقيساً بالكيلوغرام:

4.0 , 4.3 , 2.9 , 3.7 , 2.9 , 5.0 , 4.1 , 3.2 , 3.6 , 3.3

إذا علمنا أن قياسات الإنتاج في مجتمع (جمهرة) الشجيرات تتوزع طبيعياً بتباين 0.40 فالمطلوب :

أ- أوجد 95 % مجال ثقة حول متوسط الإنتاج الحقيقي μ للشجيرة .

ب- ما حجم العينة المناسب لتقدير متوسط إنتاج الشجيرة μ بحيث لا يتجاوز الخطأ الأعظمي المرتكب لتقدير μ ما مقداره 0.2 K و بثقة 95 % .

2. خضعت عينة من 12 فأراً تجريبياً لنظام غذائي معين خلال الأشهر الثلاثة الأولى من حياتها ، و قيست الزيادة في وزن كل فأر بالغرام و كانت كما يأتي:

61 , 62 , 67 , 59 , 62 , 60 , 64 , 65 , 58 , 54 , 62 , 55

و المطلوب :

أ- عين 90 % مجال ثقة لمتوسط الزيادة في الوزن خلال الأشهر الثلاثة الأولى من حياة الفئران في مجتمع الدراسة الذي جاءت منه العينة ، علماً بأن الزيادة بالوزن تتبع التوزيع الطبيعي .

ب- ما حجم العينة المناسب لتقدير المتوسط أعلاه بثقة 90 % و بخطأ لا يتجاوز 2 غرام .

3. من عينة عشوائية مؤلفة من 36 مريضاً مصاباً بمرض الإيدز ، كان متوسط العمر الذي قضوه بعد الإصابة بالمرض هو 2.6 سنة و بانحراف معياري 0.3 سنة و المطلوب :

أ- أوجد 99 % مجال ثقة حول μ متوسط العمر الحقيقي الذي يقضيه مريض الإيدز بعد الإصابة بالمرض .

ب- ما حجم العينة المناسب لتقدير μ بثقة 99 % و بخطأ لا يتجاوز 0.05 ؟

4. في دراسة حول التخلف العقلي لدى حديثي الولادة و الناتج من عدوى وراثية عن طريق الصبغيات الأنثوية وجد أن هناك 4 حالات من عينة مؤلفة من 150 طفل حديثي الولادة يعانون هذا التخلف العقلي . و المطلوب :

أ- أوجد 95 % مجال ثقة للنسبة الحقيقية للمتخلفين عقلياً و الناتج من العدوى الوراثية، و فسر الناتج.

ب- ما حجم العينة الملائم لتقدير هذه النسبة بثقة 95 % و بخطأ لا يتجاوز 0.001 ؟

5. لتقدير تباين كمية النحاس المركز في نوع معين من النباتات الموجودة على ضفاف أحد الأنهار ، اخترنا عشوائياً عينة مؤلفة من 16 نبتة ، و حرقناها، ثم حللنا الرماد الحاصل لكل نبتة ، فوجدنا كمية النحاس المركز كما يأتي (حسب وحدة قياس معينة):

50 , 8 , 14 , 27 , 18 , 34 , 3 , 5
19 , 60 , 25 , 70 , 20 , 35 , 43 , 38

و بفرض أن المجتمع المدروس يتوزع طبيعياً ، فالمطلوب :

أوجد $90\ \%$ مجال ثقة حول التباين الحقيقي σ^2 الدال على كمية النحاس المتوجود في هذا النوع من النباتات .

6. شركة لصناعة الأدوية المنظمة للضغط المرتفع ، تدعي بأن خلال فترة قصيرة يمكن أن ينتظم الضغط بنسبة $95\ \%$ ، و لاختبار هذا الادعاء تم تجريب الدواء على عينة من 200 مريض يعاني ارتفاع الضغط الشرياني، و لوحظ انتظام الضغط عند 185 منهم ضمن الفترة القصيرة المستخدمة لهذا الادعاء و المطلوب: هل نقبل بصحة ادعاء الشركة عند مستوى الأهمية $\alpha = 0.05$ من الأهمية ؟

7. من عينة عشوائية مؤلفة من 500 مدخن ، تم إيجاد 160 منهم مصابين باختلاطات رئوية سببها التدخين ، و المطلوب : أوجد $99\ \%$ مجال ثقة حول النسبة الحقيقية للمصابين باختلاطات رئوية في مجتمع المدخنين ، و فسر الناتج .

8. يخصص لدواء معين $30\ \%$ من الأسبرين لكل حبة ، أخذت عينة عشوائية من 16 حبة ، و جرى تحليلها ، فتبين أن متوسط كمية الأسبرين فيها 0.304 بانحراف معياري 0.008 و المطلوب: هل يتفق هذا الدواء مع المواصفات المطلوبة بمستوى 0.01 من الأهمية .

9. قام باحث طبي باختيار عقار جديد مضاد لأحد الجراثيم، فأعطي هذا الدواء لـ 10 أشخاص مصابين بنفس الجرثومة المدروسة، وقد اختيروا عشوائياً من مجتمع المصابين بهذا الجرثومة . و الهدف من ذلك معرفة متوسط عدد الأيام اللازمة للشفاء و تباينها . و فيما يأتي الجدول الذي يبين عدد الأيام

اللازمة للشفاء للأشخاص الذين جرب عليهم هذا الدواء :

5 , 7 , 10 , 10 , 8 , 6 , 6 , 5 , 7 , 9

و بفرض أن المجتمع المدروس يتوزع طبيعياً و المطلوب :

أ- أوجد 95 % مجال ثقة حول متوسط عدد الأيام اللازمة للشفاء لمجتمع مستخدمي هذا العقار ، و فسر الناتج .

ب- أوجد 95 % مجال ثقة حول تباين عدد الأيام اللازمة للشفاء لمجتمع مستخدمي هذا العقار ، و فسر الناتج .

10. ادعى باحث أن نصف الأشخاص الكبار في السن الذين يتم تخديرهم استعداداً للعمليات الجراحية ، يعانون مضاعفات . و لاختبار هذا الادعاء، أخذت عينة من 100 شخص خضعوا للتخدير من أجل العمل الجراحي ، فوجد أن 36 منهم حصلت لهم مضاعفات . و المطلوب : هل يمكننا قبول هذا الادعاء بمستوى 95 % من الأهمية.

11. تبين من عينة عشوائية من الحجم $n = 100$ متوفى كانوا مريضين بسرطان الرئة أن متوسط العمر لهؤلاء هو 52.8 سنة بانحراف معياري قدره 6.2 سنة، فهل يشير ذلك إلى أن متوسط عمر المصابين بسرطان الرئة هو أكبر من 50 سنة ، و ذلك عند مستوى المعنوية (الأهمية) $\alpha = 0.05$؟

12. تدعي شركة لصناعة الأدوية أن أحد أدويتها الخاصة بمعالجة التشنجات العضلية تحدث استجابة خلال فترة قصيرة لـ 80 % من المرضى ؛ و لاختبار هذا الادعاء ، أخذت عينة عشوائية من 160 مريضاً ، فوجد أن 100

منهم قد حدثت لهم استجابة فعلاً خلال الفترة الزمنية المفروضة لدى استعمالهم الدواء ، و المطلوب عند مستوى الدلالة الإحصائية (المعنوية) $\alpha = 0.01$ اختبار الفرضية الآتية : قبول صحة ادعاء الشركة .

13. بهدف تقدير متوسط كمية الخضاب في دم الأطفال في بلد معين ، تم أخذ عينة عشوائية من 100 طفل من هذا البلد ، فوجد أن متوسط كمية خضاب الدم هو 12.5G وبانحراف معياري قدره 1.5 G و المطلوب: أوجد 95 % مجال ثقة حول متوسط μ الحقيقي لكمية خضاب الدم عند أطفال البلد المدروس ، و فسر الناتج ؟

14. أجري اختبار في دار التوليد بدمشق ، لمعايرة التروكسين لدى 49 مولوداً ذكراً ، فكان المتوسط الحسابي لهذه العينة $\bar{x} = 9.8$ بانحراف معياري 3.10 و المطلوب: أوجد 99 % مجال ثقة حول المتوسط الحقيقي μ لكمية التروكسين في مجتمع المواليد ، و فسر الناتج .

15. في تجربة نفسية ممكن أن يكون رد فعل مريض جراء تناوله أحد أنواع المنشطات، أحد الشكلين A أو B ، و رغب الباحث في تقدير النسبة P التي تدل على رد الفعل A ، حيث جرب على 100 مريض و كان 60 منهم قد أبدوا رد فعل A ، و المطلوب :

أ- أوجد 90 % مجال ثقة حول P ، و فسر الناتج .

ب- ما حجم العينة الملائم لتقدير P بثقة 90 % وبخطأ لا يتجاوز 0.04؟

16. ادعى باحث اجتماعي أن 60 % من سكان المدينة يفضلون السكن في الريف ، ولاختبار هذا الادعاء أخذت عينة من 200 شخص ، فوجد

110 منهم يفضلون السكن في الريف . فهل يمكنا قبول ادعاء الباحث بثقة 99 % .

17. عند القيام بمهمة اختبار انعدام الوزن عند رواد الفضاء ، لوحظ أن معدل خفقان القلب ل 12 رائد فضاء يزداد بمعدل 27.33 ضربة بالدقيقة بانحراف معياري 4.28 ضربة بالدقيقة ، و المطلوب : أوجد 99 % مجال ثقة حول المتوسط الحقيقي لازدياد معدل خفقان القلب ، و ماذا نستنتج ؟ علماً بأن المجتمع المدروس طبيعيٌّ .

18. لدى مصنع لأدوية المضاد الحيوي رغبة كبيرة في معرفة فعالية الدواء و جودة إنتاجه . و من أجل ذلك تم تجريب الدواء على 12 مريضاً و كانت مدة الفعالية لكل مريض كالآتي :

8.9, 9.0, 9.1, 8.9, 9.1, 9.0, 9.0, 9.0, 8.9, 8.8, 9.1, 9.2

أوجد 90 % مجال ثقة حول الانحراف المعياري الحقيقي لقياس مدة الفعالية، و فسر الناتج .

19. آلة لتعبئة زجاجات الحليب المعقم ، أخذت عينة عشوائية تتضمن 36 زجاجة من إنتاج هذه الآلة ، فكان متوسط وزن الحليب في الزجاجة 495G بانحراف معياري 4G و المطلوب :

أ- أوجد 95 % مجال ثقة حول المتوسط الحقيقي μ لوزن الحليب الذي تفرغه الآلة في الزجاجة الواحدة، و فسر الناتج .

ب- ما حجم العينة المناسب لتقدير μ بثقة 99 % و بخطأ لا يتجاوز 2G ؟

20. يرغب مشفى بتقدير عدد الأيام التي يحتاج إليها علاج مريض من مرض معين ، و من أجل ذلك تم دراسة عينة عشوائية من 300 شخص مصابين

بالمرض المدروس ، و وجد أن متوسط عدد الأيام اللازمة للعلاج يساوي 5.8 يوماً ، و بانحراف معياري 1.5 يوماً، و المطلوب أوجد % 99 مجال ثقة حول المتوسط الحقيقي μ لعدد أيام العلاج ، و فسر الناتج.

21. تبين عينة عشوائية من الحجم n = 100 متوفًّى أن متوسط العمر لهم عند الوفاة كان 72 عاماً بانحراف معياري قدره 9 أعوام ، فهل هذا يشير إلى أن متوسط العمر في المجتمع أكبر من 71 عاماً ،و ذلك عند مستوى المعنوية $\alpha = 0.05$.