المحـــور الثــــاني:

التقدير الإحصائي

1– مقدمة:

تعتمد التوزيعات الإحصائية على معالم محددة ،فمثلا يعتمد التوزيع **ذي الحدين** على المعلمة P (نسبة النجاح في التجربة) ، ويعتمد توزيع **بواسون** على معلمة محددة تسمى بالوسط الحسابي وهي λ، ويعتمد التوزيع **الطبيعي** على المعلمتين μ ، σ^2 (أي الوسط الحسابي و التباين) . وفي الكثير من الحالات تكون هاته المعالم مجهولة ، وهدف أي إحصائي معرفتها ، وليس من السهل معرفة قيمة هاته المعالم بالضبط ، لهذا نضطر إلى تقديرها ، حيث أن هناك طريقتين للتقدير هما :

أ– **التقدير بنقطة " Point Estimation "**: والذي يعني اختيار قيمة واحدة كتقدير لمعلمة مجهولة كأن يستخدم متوسط الدخل الشهري للأسرة المحسوب من عينة عشوائية من الأسر المسحوبة من مجتمع معين **كتقدير** لمتوسط الدخل الشهري للأسرة في ذلك المجتمع.

ب– **التقدير بفترة" Interval Estimation "** : من الطبيعي أن التقدير بنقطة لأي معلمة لا نتوقع فيه أن يقدر تلك المعلمة بدون خطأ ؛ أي لا نتوقع أن يكون تقدير أي معلمة مطابقا تماما لقيمة المعلمة المطلوب تقديرها. فمثلا لا نتوقع أن يقدر الوسط الحسابي للعينة $(\overline{X})$ متوسط المجتمع المسحوب منه العينة (μ) بدون أي خطأ. بعبارة أخرى لا نتوقع أن يكون $\overline{X}$ مطابقا تماما لـ μ

و لذلك فإنه قد يكون من المرغوب فيه تحديد فترة يتوقع أن تقع قيمة المعلمة داخلها و هذه الفترة تسمى **بفترة الثقة** ، حيث تقع المعلمة داخل حدود تلك الفترة بدرجة ثقة (أو احتمال) يحدد ذلك.

فمثلا بدلا من استخدام $\overline{X}=300.00$ دج كتقدير بنقطة لمتوسط دخل الأسرة الشهري بالدينار في مجتمع معين (أي كتقدير لـ μ) ، تستخدم فترة ثقة لذلك كأن يقال بأن متوسط دخل الأسرة الشهري بالدينار (μ) يقع في الفترة من 270.00 إلى 330.00 دينار بدرجة ثقة 95%

أي أن : $270.00 \leq \mu \leq 330.00$ باحتمال قدره : 0.95

2– التقدير بنقطة Point Estimation:

إن أحد أهداف النظرية الإحصائية هو إيجاد الطرق التي تمكننا من الاستدلال على قيمة معالم المجتمع، و ذلك عن طريق النتائج التي نحصل عليها من عينة عشوائية نختارها من بين مفردات المجتمع.

فإذا كنا نرغب في تقدر أحد معالم المجتمع و ليكن θ من خلال عينة من المشاهدات $(x_1,x_2, \ldots\ldots,x_n)$ مسحوبة من مجتمع . فإن القيمة التي يتم حسابها للمعلمة θ من واقع هذه المشاهدات تسمى **تقديرا** (Estimate) ، بينما الدالة أو الصيغة الرياضية الإحصائية التي تستخدم

للوصول إلى هذا التقدير تسمى **مقــــدرا** (Estimator) .و المقــــدر هو إحصاء أو دالة تعتمـــد على المشاهدات ، بينما التقدير هو قيمة هذه الدالة عند التعويض بقيم المشاهدات فيها ، و لهذا فإن التقدير يختلف من عينة لأخرى رغم استخدام نفس المقدر، و هذا أمر طبيعي ،حيث هناك اختلاف بين قيم المشاهدات من عينة لأخرى رغم أن المقدر له نفس الصيغة التي يتم التعويض فيها.

وعلى العموم يمكن تقدير معلمة المجتمع بقيمة واحدة و هي قيمة المقدر التي نحصل عليها من بيانات العينة حيث :

- الوسط الحسابي لعينة عشوائية ($\overline{X}$) هو مقدر لمتوسط المجتمع (μ).
- تباين العينة (S^2) هو مقدر لتباين المجتمع (σ^2).
- نسبة صفة معينة في العينة $\overline{p}$ هي تقدير نقطي للنسبة الحقيقية لنفس الصفة في المجتمع P

إذن كل مقدر من هذه المقدرات المذكورة يسمى تقديرا بنقطة لأن كل منهم عدد حقيقي واحد أو نقطة واحدة على خط الأعداد الحقيقة.

خصائص المقدر الجيد:

من أهم الخصائص التي يجب أن تتوفر في المقدر الجيد هي :

أ- عدم التحيز (Unbiasedness)

ب- الكفاءة (Efficiency)

ج- الاتساق (Consistency)

وفيما يلي سنتطرق للخاصية الأولى والثانية بشيء من الإسهاب، بينما الخاصية الثالثة سنكتفي بعرضها بصورة مبسطة، لأن معالجتها تحتاج إلى مستوى عال من الرياضيات.

أ- خاصية عدم التحيز :

يقال بأن الاحصاءة $\hat{\theta}$ مقدرا غير متحيز للمعلمة المجهولة θ إذا كانت القيمة المتوقعة لــ $\hat{\theta}$ تساوي المعلمة المجهولة θ ، أي : : $\mathbf{E(\hat{\theta}) = \theta}$ ....(01)

ونعلم مما سبق أن الوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة هو نفسه الوسط الحسابي للمجتمع الأصلي، أي أن:

$$\mu_{\overline{X}} = \mu \quad \text{أي :} \quad \mathbf{E(\overline{X}) = \mu}$$

وبالتالي ، فان الاحصاءة $\overline{X}$ هي مقدر غير متحيز للمعلمة المجهولة μ ، ويمكن إثبات ذلك في النظرية التالية.

نظرية (01) : إذا كان $\overline{X}$ الوسط الحسابي لعينة عشوائية $(x_1,x_2,\dots,x_n)$ مسحوبة من مجتمع وسطه الحسابي μ فان : $E(\overline{X})=\mu$

البرهان :

$$
\begin{aligned}
E(\overline{X}) &= E\left(\frac{x_1+x_2+\cdots+x_n}{n}\right) \\
&= \frac{1}{n}[\,E(x_1+x_2+\cdots+x_n)\,] \\
&= \frac{1}{n}[\,E(x_1)+E(x_2)+\cdots+E(x_n)\,] \\
&= \frac{1}{n}[\,\mu+\mu+\cdots+\mu] \\
&= \frac{1}{n}[\,n\,\mu] \\
&= \mu
\end{aligned}
$$

كذلك نستطيع أن نثبت أن تباين العينة S^2 هو مقدر غير متحيز لتباين المجتمع σ^2 في حالة سحب العينة العشوائية مع الإرجاع.

أما في حالة السحب دون إرجاع، فان المقدر غير المتحيز لتباين المجتمع هو $(\frac{N-1}{N}S^2)$

وفي المثال (01) الموالي سوف نثبت عدديا أن تباين العينة هو مقدر غير متحيز للمعلمة المجهولة σ^2 في حالة السحب مع الإرجاع، أما الإثبات النظري فسنوضحه في النظرية التالية.

نظرية (02): إذا كان S^2 تباين عينة عشوائية $(x_1,x_2,\dots,x_n)$ مسحوبة من مجتمع تباينه σ^2 فان : $E(S^2)=\sigma^2$

البرهان:

$$
\begin{aligned}
E(S^2) &= E\left(\frac{\Sigma(x_i-\overline{X})^2}{n-1}\right) \\
&= \frac{1}{n-1}E[\,\Sigma\,(x_i^2-2x_i\overline{X}+\overline{X}^2)\,] \\
&= \frac{1}{n-1}E\,(\Sigma x_i^2-2\overline{X}\,\Sigma x_i+n\,\overline{X}^2) \\
&= \frac{1}{n-1}E\,(\Sigma x_i^2-n\,\overline{X}^2) \\
&= \frac{1}{n-1}[\Sigma E(x_i^2)-nE(\overline{X}^2)\,]
\end{aligned}
$$

وذلك لان $\Sigma x_i = n\,\overline{X}$ ، وبما أن تباين المجتمع يمكن حسابه كما يلي :

$$\sigma^2 = E(x)^2 - \mu^2$$

وكذلك تباين توزيع المعاينة للوسط الحسابي للعينة $\sigma^2_{\overline{X}}$ يمكن حسابه كما يلي:

$$\sigma^2_{\overline{X}} = E(\overline{X})^2 - \mu^2$$

وبالتعويض عن $E(x)^2$ و $E(\overline{X})^2$ نجد أن :

$$\mathbf{E(S^2)} = \frac{1}{n-1}\left[\Sigma(\sigma^2 + \mu^2) - n(\sigma^2_{\overline{X}} + \mu^2)\right]$$

$$= \frac{1}{n-1}(n\sigma^2 + n\mu^2 - n\sigma^2_{\overline{X}} - n\mu^2)$$

$$= \frac{1}{n-1}(n\sigma^2 - n\sigma^2_{\overline{X}})$$

وبما أن السحب تم مع الإرجاع فان $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$

إذن :

$$E(S^2) = \frac{1}{n-1}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right)$$

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

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

$$= \boldsymbol{\sigma^2}$$

وبالتالي فان S^2 هو مقدر غير متحيز لتباين المجتمع σ^2

مثال (01) : بفرض انه لدينا مجتمعا يتكون من مفردتين قيمتهما 3 ، 6 .فان التوزيع الاحتمالي لهذا المجتمع موضح في الجدول التالي :

الجدول (01) التوزيع الاحتمالي للمجتمع

x_i	3	6
$P(x_i)$	1/2	1/2

المطلوب: اثبت أن تباين العينة هو مقدر غير متحيز لتباين المجتمع في حالة السب مع الإرجاع.

الحل: من الجدول السابق نقوم بحساب الوسط الحسابي وتباين المجتمع على التوالي كما يلي :

$$\mu = \Sigma\, x_i P(x_i) = 3(1/2) + 6(1/2) = \frac{9}{2}$$

$$\sigma^2 = E(x)^2 - \mu^2 = \Sigma x_i^2 P(x_i) - \mu^2$$
$$= 3^2(1/2) + 6^2(1/2) - (\frac{9}{2})^2$$
$$= \frac{9}{4}$$

إذا سحبنا من هذا المجتمع مع الإرجاع كل العينات الممكنة ذات الحجم 3 ، وحسبنا لكل عينة وسطها الحسابي ، فسوف نتحصل على النتائج الموضحة في الجدول(02) .

الجدول(02) العينات ذات الحجم 3 الممكن سحبها من المجتمع مع الإرجاع مقرونة بمتوسطاتها الحسابية

العينة	$\overline{x}$	العينة	$\overline{x}$
(3, 3, 3)	3	(6, 3, 3)	4
(3, 3, 6)	4	(6, 3, 6)	5
(3, 6, 3)	4	(6, 6, 3)	5
(3, 6, 6)	5	(6, 6, 6)	6

الآن نقوم بحساب تباين كل عينة S^2 ، ونلخص ذلك في الجدول (03) ، حيث :

$$S^2 = \Sigma(x_i - \overline{x})^2 / n-1$$

الجدول (03) العينات ذات الحجم 3 الممكن سحبها من المجتمع مع الإرجاع مقرونة بتبايناتها

العينة	S^2	العينة	S^2
(3, 3, 3)	0	(6, 3, 3)	3
(3, 3, 6)	3	(6, 3, 6)	3
(3, 6, 3)	3	(6, 6, 3)	3
(3, 6, 6)	3	(6, 6, 6)	0

ومن خلال هذا الجدول، فان توزيع المعاينة لتباين العينة يمكن توضيحه في الجدول الموالي:

الجدول (04) توزيع المعاينة لتباين العينة

S^2	0	3
$P(S^2)$	2/8	6/8

ومن هنا نجد أن القيمة المتوقعة لتوزيع المعاينة للتباينات أي $E(S^2)$ هي كما يلي :

$$E(S^2) = \Sigma S^2 P(S^2) = 0(2/8) + 3(6/8)$$
$$= \frac{9}{4}$$

وفي الأخير نجد أن: $\mathbf{E(S^2) = \sigma^2}$

وبالتالي ، فان تباين العينة هو مقدر غير متحيز لتباين المجتمع في حالة السحب مع الإرجاع.

ب- خاصية الكفاءة :

قد يوجد للمعلمة الواحدة أكثر من مقدر ، وبالتالي يمكننا المقارنة بين هذه المقدرات من خلال المقارنة بين تبايناتهم ،حيث نعتبر أن المقدر **الأقل تباينا** هو المقدر **الأكثر كفاءة** (More Efficient) .

فإذا كان لدينا المقدران : $\hat{\theta}_1$ ، $\hat{\theta}_2$ للمعلمة θ. وكان تباين المقدر الأول هو $\sigma^2_{\hat{\theta}_1}$ و تباين المقدر الثاني هو $\sigma^2_{\hat{\theta}_2}$ ، فإن المقدر $\hat{\theta}_1$ يكون أكثر كفاءة من المقدر $\hat{\theta}_2$ إذا كان $\sigma^2_{\hat{\theta}_1} < \sigma^2_{\hat{\theta}_2}$

على افتراض أن : $\hat{\theta}_1$ و $\hat{\theta}_2$ هي مقدرات غير متحيزة .

مثال (02) : : بالنسبة للمجتمعات الطبيعية ، فإن الوسط الحسابي و الوسيط يمثلان تقديرين غير متحيزين لمتوسط المجتمع μ ، و لكن تباين الوسط أقل من تباين الوسيط ، وبالتالي فإن الوسط الحسابي يكون مقدرا أكثر كفاءة لمتوسط المجتمع μ. و للتأكد من ذلك لدينا :

تباين الوسط هو $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$

أما تباين الوسط فهو يساوي تقريبا $\sigma^2_{Me} = \frac{\pi\sigma^2}{2n}$

و تقاس الكفاءة النسبية بين التباينين بقسمة أحدهما على الآخر ، و ذلك كما يلي :

$$0.64 = \frac{2}{\pi} = \frac{\frac{\sigma^2}{n}}{\frac{\pi\sigma^2}{2n}} = \frac{\text{تباين الوسط}}{\text{تباين الوسيط}}$$

وعليه يتضح لنا انه عند نفس حجم العينة فان تباين الوسط أقل من تباين الوسيط ، وبالتالي فان الوسط الحسابي يعتبر مقدرا أكثر كفاءة من الوسيط .

ج- خاصية الاتساق :

يقال بأن الاحصاءة $\hat{\theta}$ هي مقدر متسق للمعلمة θ إذا كانت $\hat{\theta}$ تقترب من θ كلما زاد حجم العينة واقترب من المالانهاية .

و عليه يمكن إثبات أن $\hat{\theta}$ تكون مقدر متسق لــ θ ، إذا كانت $\hat{\theta}$ مقدر غير متحيز لــ θ ، و كان تباين $\hat{\theta}$ يقترب من الصفر كلما اقتربت n من المالانهاية.

فمثلا إذا كان $\overline{X}$ هو الوسط الحسابي لعينة حجمها n مسحوبة من مجتمع وسطه μ و تباينه σ^2 يكون $\overline{X}$ مقدرا متسقا لــ μ لأن :

$$E(\overline{X}) = \mu \ , \ \sigma^2_{\overline{X}} = \frac{\sigma^2}{n} \to 0 \quad \text{As} \quad n \to \omega$$

فالشرط الأول سبق التأكد منه ، أما الشرط الثاني فالتأكد منه بسيط جدا حيث ،كلما زاد حجم العينة (و قرب من مالا نهاية) فإن تباين المتوسط $\sigma^2_{\overline{X}}$ يقترب من الصفــــر .

وعليه يمكن القول بأن S^2 تقدير متسق لــ σ^2

3- تقدير متوسط المجتمع بفترة أو فترة الثقة لمتوسط المجتمع μ:

أن تقدير وسط المجتمع بفترة هو عبارة عن أيجاد تقدير نقطي لوسط المجتمع أولا ومن ثم استعمال هذا المقدار لإيجاد قيمتين تعتمدان على التوزيع الاحتمالي لهذا المقدر وعلى معامل الثقة ، وبالتالي تمثل هاتان القيمتان الحد الأدنى والحد الأعلى لفترة الثقة المطلوبة .

فإذا كان وسط المجتمع μ غير معلوم وأردنا إيجاد فترة ثقة للمعلمة μ فإننا نستعمل وسط العينة $\overline{X}$ كتقدير نقطي للمعلمة، ونستعمل توزيع المعاينة للإحصاء $\overline{X}$ لنحدد الحد الأدنى والحد الأعلى لفترة الثقة المطلوبة.

3– 1.فترة الثقة لمتوسط المجتمع μ عندما يكون تباين المجتمع σ^2 معلوما:

إذا كانت العينة المدروسة كبيرة الحجم وتم سحبها من مجتمع تباينه معلوم، فان توزيع المعاينة للإحصاء $\overline{X}$ هو التوزيع الطبيعي أو تقريبا التوزيع الطبيعي ذي الوسط μ والتباين $\frac{\sigma^2}{n}$ ، وفي مثل هذه الحالات نستطيع إيجاد فترة الثقة للوسط μ من العبارة الاحتمالية التالية :

$$\mathbf{P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1- \alpha}$$

حيث: $$\mathbf{Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}$$

ولدينا $z_{\alpha/2}$ هي النقطة على المحور الافقي لمنحنى التوزيع الطبيعي المعياري الذي يقع على يمينها $\alpha/2$ من المساحة انظر الشكل (01)

الشكل (01)

ص9 من كتاب الاحصاء التطبيقي ات/ 2667

ومن المعادلة السابقة نجد فترة الثقة ذات معامل ثقة **(1- α)** للوسط μ ، ونسميها فترة ثقـــة (1- α) 100% كما يلي:

$$\overline{X} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$$

وبهذا نصل إلى النظرية التالية :

نظرية (03): إذا كانت ($x_1,x_2......,x_n$) عينة عشوائية مسحوبة من مجتمع طبيعي وسطه μ تباينه σ^2 معلوم ، فان فترة ثقـــة (1- α) 100% للمعلمة μ هي :

$$\left[\overline{X} - \boldsymbol{z_{\alpha/2}}.\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}} \ , \overline{X} + \boldsymbol{z_{\alpha/2}}.\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}\right] \quad .... (02)$$

حيث $z_{\alpha/2}$ هي النقطة على المحور الافقي لمنحنى التوزيع الطبيعي المعياري الذي يقع إلى يمينها مساحة $\alpha/2$

مثال (03): أوجد فترة ثقة 95% للوسط μ في مجتمع طبيعي تباينه 64 ، هذا إذا اخترت عينة عشوائية حجمها 9 ، وكان وسطها الحسابي 52 .

الحل :

بما أن : $1-\alpha = 0.95$ فان : $\frac{\alpha}{2} = 0.025$

وبالتالي يكون : $z_{\alpha/2} = 1.96$ من جدول التوزيع الطبيعي في الملحق رقم (01).

اذن بالتعويض في فترة الثقة نجد ان فترة ثقة 95% للوسط μ هي :

$$52 - 1.96\cdot\frac{8}{\sqrt{9}} \le \mu \le 52 + 1.96\cdot\frac{8}{\sqrt{9}}$$

$$52 - 5.227 \le \mu \le 52 + 5.227$$

$$46.773 \le \mu \le 57.227$$

وهي فترة ثقة 95% للوسط μ

انطلاق مما سبق ، فان الهيكل العام لفترة الثقة لتوزيعات المعاينة المتماثلة يأخذ الشكل التالي :

الخطأ المعياري للتقدير بنقطة. قيمة جزئية $\pm$ التقدير بنقطة

هامش خطأ المعاينة $\pm$ التقدير بنقطة

– اختيار حجم العينة المناسب لتقدير الوسط μ :

إن التحكم في خطأ المعاينة هو احد الرؤى الأساسية عند التخطيط لدراسة إحصائية معينة ، ويمكن تحديد هامش خطأ معاينة مرغوب فيه (الخطأ في تقدير μ) وذلك باختيار حجم العينة المناسب.

حيث يسمى المقدار $z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}$ حد $100(1-\alpha)\%$ للخطأ في تقدير μ . ولهذا إذا تم تحديد المقدار الأكبر المسموح به لهذا الخطأ في التقدير أمكن حساب حجم العينة المناسب لتحقيق ذلك الحد عن طريق حل المتباينة التالية:

$$z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}} \le E$$

حيث أن E هو حد المقدار الأكبر المسموح به لهذا الخطأ، وينتج من ذلك أن:

$$n \ge \left[\frac{z_{\alpha/2}\cdot\sigma}{E}\right]^2 \quad \text{..........} \quad (03)$$

مثال (04): أراد عميد إحدى الكليات أن يحدد حجم العينة اللازمة لحساب متوسط الزمن اللازم لانتقال الطلبة من مدرج إلى آخر، وذلك ألا يزيد الخطأ في هذا المتوسط عن 0.3 دقيقة، وهذا بدرجة ثقة 95% على أساس أن الانحراف المعياري σ من دراسات مماثلة يساوي1.5 دقيقة.

الحل:

لتحديد حجم العينة اللازمة، فإننا سوف نطبق العلاقة التالية:

$$n \geq \left[\frac{z_{\alpha/2}.\sigma}{\mathrm{E}}\right]^2 \Leftrightarrow n \geq \left[\frac{1.96 \times 1.5}{0.3}\right]^2$$

$$\Leftrightarrow n \geq 96.04$$

$$\Rightarrow n \approx 97 \text{ مفردة}$$

مثال (05): قامت مجموعة من الخبراء بتقدير متوسط الزمن اللازم لتركيب آلة معينة في صناعة كبيرة وذلك استنادا على عينة عشوائية حجمها 150 ، ومن المعلومات المتوفرة لدى تلك المجموعة ان الانحراف المعياري σ يساوي 6.2 دقيقة

المطلوب : احسب الحد الأقصى للخطأ لهذا التقدير بدرجة ثقة 99% .

الحل:

$$\text{الخطأ} = z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$$

$$= 2.575.\frac{6.2}{\sqrt{150}}$$

$$\approx 1.30$$

وبذلك يمكن القول بأن الخطأ الأقصى هو 1.3 بدرجة ثقة 99% .

3- 2.فترة الثقة لمتوسط المجتمع μ عندما يكون تباين المجتمع σ^2 مجهول:

نظرية (04): إذا كان $\overline{X}$ هو الوسط الحسابي لعينة عشوائية حجمها كبير بدرجة كافية مسحوبة من مجتمع طبيعي وسطه μ وتباينـــه σ^2 مجهول ، فان فتـــرة ثقـــة $(1-\alpha)$ 100% للمعلمة μ هي :

$$\left[\overline{X} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \quad , \overline{X} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\right] \quad .... (03)$$

حيث $z_{\alpha/2}$ هي نقطة على المحور الأفقي لمنحنى التوزيع الطبيعي المعياري التي يقع إلى يمينـــها مساحة $\alpha/2$

مثال (06): إذا كان الوسط الحسابي والانحراف المعياري لمعدلات عينة عشوائية مؤلفة من 36 طالبا في إحدى الجامعات هما على التوالي: 66.3 ، 12 .

المطلوب: أوجد فترة ثقة 95% لمعـــدل جميع الطلبة الجامعة إذا افترضنا أن علامات الطلبـــة تخضـــع للتوزيع الطبيعي.

الحل:

نلاحظ أن تباين المجتمع مجهول ، وحجم العينة كبير بدرجة كافية فان فترة الثقة المطلوبة هي:

$$\overline{X} - z_{\alpha/2}.\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + z_{\alpha/2}.\frac{S}{\sqrt{n}}$$

بما أن : $1-\alpha = 0.95$ فان : $\frac{\alpha}{2} = 0.025$

وبالتالي يكون : $z_{\alpha/2} = 1.96$ من جدول التوزيع الطبيعي في الملحق رقم (01).

أذن بالتعويض في فترة الثقة نجد أن فترة ثقة 95% للوسط μ هي :

$$66.3 - 1.96.\frac{12}{\sqrt{36}} \le \mu \le 66.3 + 1.96.\frac{12}{\sqrt{36}}$$

$$66.3 - 3.92 \le \mu \le 66.3 + 3.92$$

$$62.38 \le \mu \le 70.22$$

وهي فترة ثقة 95% لمعدل جميع طلبة الجامعة.

نظرية (05): إذا كان $\overline{X}$ هو الوسط الحسابي لعينة عشوائية صغيرة الحجم تم سحبها من مجتمع طبيعي وسطه μ وتباينــه σ^2 مجهول ، فان فتــرة ثقــة $(1-\alpha)$ 100% للمعلمة μ هي :

$$\left[\overline{X} - t_{(\frac{\alpha}{2},\upsilon)}.\frac{S}{\sqrt{n}} \quad , \overline{X} + t_{(\frac{\alpha}{2},\upsilon)}.\frac{S}{\sqrt{n}}\right] \quad .... (04)$$

حيث $t_{(\frac{\alpha}{2},\upsilon)}$ هي نقطة على المحور الأفقي لتوزيع ستيودنت t ذي درجات الحرية $(n-1)$ والتي يقع إلى يمينــها مساحة $\alpha/2$

مثال (07): أخذت عينة عشوائية حجمها 7 من مجتمع طبيعي ، فكان وسطها الحسابي 11.2 ، وانحرافها المعياري 0.4

المطلوب: أوجد فترة ثقة 95% لمعـــدل المجتمع μ

الحل:

بتطبيق النظرية الخامسة نجد فترة الثقة المطلوبة هي :

$$\left[\overline{X} - t_{(\frac{\alpha}{2},\upsilon)}.\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + t_{(\frac{\alpha}{2},\upsilon)}.\frac{S}{\sqrt{n}}\right]$$

بما أن : $1-\alpha = 0.95$ فان : $\frac{\alpha}{2} = 0.025$

ومن جدول توزيع t بدرجات حرية 6 الوارد في الملحق رقم (02) نجد أن :

$$t_{(0.025,6)} = 2.447$$

إذن فترة الثقة 95% للمعـــدل μ هي:

$$11.2 - 2.447.\frac{0.4}{\sqrt{7}} \le \mu \le 11.2 + 2.447.\frac{0.4}{\sqrt{7}}$$

$$11.2 - 0.37 \le \mu \le 11.2 + 0.37$$

$$10.83 \le \mu \le 11.57$$

وهذا بدرجة ثقة 95%

3- 3.فترة الثقة لمتوسط مجتمع غير طبيعي:

نظرية (06): إذا أخذت عينة عشوائية حجمها n من مجتمع غير طبيعي، فانه بإمكاننا استعمال نظرية النهــاية المركــزية لمعرفة توزيع الوســط إذا كان حجم العينة **كبيــر** ، ومن ثم تكون فتــرة ثقــة $(1-\alpha)$ 100% للمعدل μ هي :

$$\left[\overline{X} - \boldsymbol{z_{\alpha/2}}.\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}} \quad , \overline{X} + \boldsymbol{z_{\alpha/2}}.\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}\right] \quad .... (05)$$

وفي حالة عدم معرفة σ نستعمل S بدلا منها ، شريطة أن تكون n كبيرة.

مثال (08): تم دراسة عينة عشوائية حجمها 81 سيجارة لتقدير متوسط النيكوتين في السجائر، فوجد أن متوسط العينة هو 3.6 ملغم بانحراف معياري قدره 0.9 ملغم .

المطلوب: أوجد فترة ثقة 99% لمتوسط النيكوتين في السجائر.

الحل:

لدينا: $1-\alpha = 0.99$ فان : $\frac{\alpha}{2} = 0.005$

وبالتالي ، من جدول التوزيع الطبيعي في الملحق رقم (01) يكون : $z_{\alpha/2} = 2.57$

إذن، باستعمال النظرية السادسة فان فترة الثقة المطلوبة هي على الشكل التالي:

$$\bar{x} - z_{\alpha/2}.\frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}.\frac{S}{\sqrt{n}}$$

وبالتطبيق نجد:

$$3.6 - 2.57.\frac{0.9}{\sqrt{81}} \leq \mu \leq 3.6 + 2.57.\frac{0.9}{\sqrt{81}}$$

أي أن :

$$3.6 - 0.257 \leq \mu \leq 3.6 + 0.257$$

$$3.343 \leq \mu \leq 3.85$$

وهي فترة الثقة المطلوبة.

4– فترة الثقة للفرق بين متوسطي مجتمعين $(\mu_1 - \mu_2)$:

4–1 .عندما يكون تبايني المجتمعين σ_1^2 و σ_2^2 معلومين:

نحتاج في بعض الدراسات الإحصائية إلى مقارنة متوسطين حسابيين لمجتمعين ومعرفة الفرق بينهما $(\mu_1 - \mu_2)$ ، فإذا كان المجتمع الأول يتوزع توزيعا طبيعيا بمتوسط μ_1 وتباين σ_1^2 معلوم ، والمجتمع الثاني يتوزع توزيعا طبيعيا بمتوسط μ_2 وتبـــاين σ_2^2 معلوم أيضا ، وسحبنا من المجتمـــع الأول عينـــة عشوائية حجمها n_1 ، ومن المجتمع الثاني عينة عشوائية حجمها n_2 ، وكانت العينـــتان مستقلتـــين فقد علمنا من دراستـــنا السابقة انه في هذه الحالة يكون توزيع المعاينة للـــفرق بين متوسطي العينتـــين $(\bar{X}_1 - \bar{X}_2)$ توزيعا طبيعيا بمتوسط وتباين قدرهما عل التوالي:

$$\sigma^2_{\bar{X}_1-\bar{X}_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2} \quad ، \quad \mu_{\bar{X}_1-\bar{X}_2}=\mu_1-\mu_2$$

وبالتالي، فان المتغير العشوائي Z حيث :

$$Z=[(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)]/\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

هو متغير عشوائي يتبع توزيعا طبيعيا معياريا.

وبهذا نصل إلى النظرية التالية:

نظرية (07): إذا كان $\bar{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 تم سحبها من مجتمع له توزيع طبيعي متوسطه μ_1 وتباينه **معلوم** σ_1^2، وكان $\bar{X}_2$ هو الوسط الحسابي لعينة عشوائية حجمها n_2 مستقلة عن الأولى تم سحبها من مجتمع له توزيع طبيعي متوسطه μ_2 وتباينه σ_2^2 **معلوم** أيضا ، فان فترة ثقة $(1-\alpha)$ 100% للفرق بين متوسطي المجتمعين $(\mu_1-\mu_2)$ تعطى حسب العلاقة التالية:

$$\left[(\bar{X}_1-\bar{X}_2)-z_{\alpha/2}\cdot\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ ,(\bar{X}_1-\bar{X}_2)+z_{\alpha/2}\cdot\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ \right] \quad \ldots\ldots (06)$$

مثال (09): لمقارنة متوسط طول نوع معين من الأنابيب المنتجة من المصنع (أ) بطول الأنابيب المنتجة من المصنع (ب)، سحبنا عينة عشوائية من المصنع (أ)تحتوي على 20 أنبوبة ، وكان الوسط الحسابي لأطوالها 3.8 سم ، وسحبنا عينة عشوائية مستقلة عن العينة الأولى من المصنع (ب)تحتوي على 25 أنبوبة ، وكان الوسط الحسابي لأطوالها 3.2 سم، فإذا كان طول الأنابيب المنتجة في المصنع (أ) يتبع توزيعا طبيعيا بتباين 0.82 سم ، وطول الأنابيب المنتجة في المصنع (ب) يتبع توزيعا طبيعيا بتباين 0.64 سم

المطلوب : أوجد فترة ثقة 95% للفرق بين متوسطي الطول في المصنعين $(\mu_1-\mu_2)$.

الحل:

بما ان المجتمعين يتوزعان توزيعا طبيعيا ، والعينتين مستقلتان ، وتبايني المجتمعين معلومين ، فان فترة الثقة المطلوبة هي على الشكل التالي:

$$\left[(\bar{X}_1-\bar{X}_2)-z_{\alpha/2}.\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ ,(\bar{X}_1-\bar{X}_2)+z_{\alpha/2}.\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]$$

انطلاقا مما سبق لدينا: $z_{\alpha/2}=1.96$ ، وبالتعويض في هذه الفترة بالبيانات المتوافرة لدينا نتحصل على:

الحد الأدنى لفترة الثقة:

$$(\bar{X}_1-\bar{X}_2)-z_{\alpha/2}.\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}=(3.8-3.2)-1.96.\sqrt{\frac{0.82}{20}+\frac{0.64}{25}}$$
$$=0.09$$

الحد الأعلى لفترة الثقة:

$$(\bar{X}_1-\bar{X}_2)+z_{\alpha/2}.\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}=(3.8-3.2)+1.96.\sqrt{\frac{0.82}{20}+\frac{0.64}{25}}$$
$$=1.11$$

إذن **فترة الثقة للفــــرق** بين الوسط الحسابي لأطوال أنابيب المصنع (أ) والوسط الحسابي لأطوال أنابيب المصنع (ب) عند مستوى الثقة 95% هي: [0.09 ، 1.11]

أي نستطيع القول بثقة قدرها 95% أن الفرق الحقيقي بين الوسط الحسابي لأطوال كل أنابيب المصنع (أ) ، والوسط الحسابي لأطوال كل أنابيب المصنع (ب)، يقع بين القيمتين 0.09 سم و1.11 سم.

وبمـــا أن **حدي فتـــرة الثقـــة موجبـــان**، نستنتج أن الوسط الحسابي لأطوال الأنابيب المنتجة من المصنع (أ) اكبر من الوسط الحسابي لأطوال الأنابيب المنتجة من المصنع (ب).

4-2 .عندما يكون تبايني المجتمعين σ_1^2 و σ_2^2 مجهولين والعينتان مستقلتان وصغيرتا الحجم:

وفي هذا العنصر لدينا حالتين:

- حالة تبايني المجتمعين مجهولين ومتساويين.
- حالة تبايني المجتمعين مجهولين وغير متساويين.

ويمكن توضيح كل هذا في النظريتين **الثامنة والتاسعة** التاليتين على الترتيب.

نظرية (08): إذا كان $\overline{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 سحبت من **مجتمع طبيعي** وسطه μ_1 وتباينه σ_1^2 ، وكان $\overline{X}_2$ هو الوسط الحسابي لعينة عشوائية حجمها n_2 سحبت من **مجتمع طبيعي** آخر مستقل عن المجتمع الأول وسطه μ_2 وتباينه σ_2^2 ، وكان **حجم العينتين صغير وتبايني المجتمعين مجهولين ومتساويين** ، فإن فترة ثقة $(1-\alpha)100\%$ للفرق بين متوسطي المجتمعين $(\mu_1 - \mu_2)$ تعطى حسب العلاقة التالية:

$$\left[(\overline{X}_1 - \overline{X}_2) - t_{(\frac{\alpha}{2},\upsilon)} \cdot \sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}\ ,(\overline{X}_1 - \overline{X}_2) + t_{(\frac{\alpha}{2},\upsilon)} \cdot \sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}\right] \quad ..... (07)$$

حيث درجات الحرية تأخذ الشكل التالي: $\upsilon = n_1+n_2-2$

مثال (10): استخدمت طريقتين لإنتاج سلعة معينة، وتوضح البيانات التالية الوقت بالدقائق المستغرق في إنتاج الوحدة بالنسبة لست وحدات أنتجت بالطريقة الأولى ، وخمس وحدات أنتجت بالطريقة الثانية ، حيث X_1 يمثل الوقت المستغرق عند استعمال الطريقة الأولى، و X_2 يمثل الوقت المستغرق عند استعمال الطريقة الثانية:

X_1 : 40 40 50 60 60 50

X_2 : 40 45 55 58 62

فإذا علمت أن الوقت المستغرق في الإنتاج بالطريقتين يتوزع توزيعا طبيعيا بتباينين متساويين

المطلوب : أوجد فترة ثقة 95% للفرق بين متوسطي الوقت المستغرق بالطريقتين $(\mu_1 - \mu_2)$.

الحل:

بما أن الوقت المستغرق في الإنتاج بالطريقتين يتوزع توزيعا طبيعيا **بتباينين مجهولين ومتساويين** ، والعينتين صغيرتان ومستقلتان لان كل عينة خاصة بآلة مستقلة عب الآلة الأخرى ، إذن فترة الثقة المناسبة في هذه الحالة هي :

$$\left[(\overline{X}_1 - \overline{X}_2) - t_{(\frac{\alpha}{2},\upsilon)} \cdot \sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}\ ,(\overline{X}_1 - \overline{X}_2) + t_{(\frac{\alpha}{2},\upsilon)} \cdot \sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}\right]$$

ولإيجاد هذه الفترة يجب إيجاد ما يلي: $\overline{X}_1$ ، $\overline{X}_2$ ، s_1^2 ، s_2^2 ، s_p^2 ، وذلك كما يلي :

$$\overline{X}_1 = \frac{\sum X_{1i}}{n_1} = \frac{300}{6} = 50$$

$$\overline{X}_2 = \frac{\sum X_{2i}}{n_2} = \frac{260}{5} = 52$$

$$s_1^2 = \frac{\Sigma\left(X_{1i}-\overline{X}_1\right)^2}{n_1-1} = \frac{400}{5} = 80$$

$$s_2^2 = \frac{\Sigma\left(X_{2i}-\overline{X}_2\right)^2}{n_2-1} = \frac{338}{4} = 84.5$$

إذن يكون :

$$s_p^2 = \frac{(n_1-1)\, s_1^2+(n_2-1)\, s_2^2}{n_1+n_2-2} = \frac{(6-1)80+(5-1)84.5}{6+5-2} = 82$$

وبما أن : $1-\alpha = 0.95$ فان : $\frac{\alpha}{2} = 0.025$

ولدينا: $\upsilon = n_1+n_2-2 = 6+5-2 = 9$

فيكون : $t_{(\frac{\alpha}{2},\upsilon)} = t_{(0.025,9)} = 2.262$

إذن بالتعويض في فترة الثقة السابقة نحصل على :

الحد الأدنى لفترة الثقة:

$$(\overline{X}_1 - \overline{X}_2) - t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})} = (50-52) - 2.626.\sqrt{82(\frac{1}{6}+\frac{1}{5})}$$
$$= -14.40$$

الحد الأعلى لفترة الثقة:

$$(\overline{X}_1 - \overline{X}_2) + t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})} = (50-52) + 2.626.\sqrt{82(\frac{1}{6}+\frac{1}{5})}$$
$$= 10.40$$

وبالتالي ،فان فترة ثقة للفرق بين متوسطي الوقت المستغرق بالطريقتين$(\mu_1 - \mu_2)$ عند مستوى ثقة 95 % هي: [-14.40 ، 10.40]

أي نستطيع القول بثقة قدرها 95 % بأن الفرق الحقيقي بين متوسط الوقت المستغرق لإنتاج الوحدة بالطريقة الأولى ، ومتوسط الوقت المستغرق لإنتاج الوحدة بالطريقة الثانية ، يقع بين القيمتين -14.40 دقيقة و 10.40 دقيقة.

نظرية (09): إذا كان $\overline{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمها n_1 سحبت من **مجتمع طبيعي** وسطه μ_1 وتباينه σ_1^2 ، وكان $\overline{X}_2$ هو الوسط الحسابي لعينة عشوائية حجمها n_2 سحبت من **مجتمع طبيعي** آخر مستقل عن المجتمع الأول وسطه μ_2 وتباينه σ_2^2 ، وكان **حجم العينتين صغير وتبايني المجتمعين مجهولين** ، فإن فترة ثقة $(1-\alpha)100\%$ للفرق بين متوسطي المجتمعين $(\mu_1-\mu_2)$ تعطى حسب العلاقة التالية:

$$\left[(\overline{X}_1-\overline{X}_2)-t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2})}\ ,(\overline{X}_1-\overline{X}_2)+\ t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2})}\right]$$

..... (08)

حيث أن درجات الحرية في هذه الحالة لها الصيغة المركبة التالية:

$$\upsilon=\left(\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1-1)}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2-1)}}\right)$$

وعندما يكون **حجما العينتين كبيرين** بدرجة كافية أي حجم كل عينة يكون اكبر أو يساوي 30 ، فان فترة الثقة للفرق بين متوسطي المجتمعين $(\mu_1-\mu_2)$ عندما يكون تبايني المجتمعين مجهولين وغير متساويين ، وعند مستوى ثقة $(1-\alpha)100\%$ تعطى حسب العلاقة التالية:

$$\left[(\overline{X}_1-\overline{X}_2)-z_{\alpha/2}.\sqrt{(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2})}\ ,(\overline{X}_1-\overline{X}_2)+\ z_{\alpha/2}.\sqrt{(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2})}\right]$$

..... (09)

وهذا حتى لو لم يكن المجتمعان خاضعان للتوزيع الطبيعي، وذلك وفقا لنظرية النهاية المركزية.

مثال (10): البيانات التالية خاصة بعينتين مستقلتين مسحوبتين من مجتمعين يتوزعان توزيعا طبيعيا بتباينين غير متساويين:

العينة الأولى:

$$n_1 = 4$$
$$\overline{X}_1 = 1$$
$$s_1^2 = 0.0547$$

العينة الثانية:

$$n_2 = 5$$
$$\overline{X}_2 = 1.2$$
$$s_2^2 = 0.0991$$

المطلوب : أوجد فترة ثقة 90% للفرق بين متوسطـــي المجتمعين.

الحل:

بنا أن المجتمعين يتوزعان توزيعا طبيعيا بتباينين غير متساويين ، والعينتين مستقلتان ، فان فترة الثقة المطلوبة في هذه الحالة هي:

$$\left[(\overline{X}_1 - \overline{X}_2) - t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}\ ,(\overline{X}_1 - \overline{X}_2) + \ t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}\right]$$

ولدينا : $1\text{-}\ \alpha = 0.90$ فان : $\frac{\alpha}{2} = 0.05$

$$\upsilon = \left(\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1-1)}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2-1)}}\right) = \left(\frac{\left(\frac{0.0547}{4}+\frac{0.0991}{5}\right)^2}{\frac{\left(\frac{0.0547}{4}\right)^2}{(4-1)}+\frac{\left(\frac{0.0991}{5}\right)^2}{(5-1)}}\right) = 6.99 \approx 7$$

وعليه يكون :

$$t_{(\frac{\alpha}{2},\upsilon)} = \ t_{(0.05,7)} = 1.895$$

وعند التعويض في فترة الثقة المطلوبة نحصل على:

الحد الأدنى لفترة الثقة:

$$(\overline{X}_1 - \overline{X}_2) - t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})} = (1 - 1.2) - 1.895.\sqrt{(\frac{0.0547}{4} + \frac{0.0991}{5})}$$
$$= -\ 0.5468$$

الحد الأعلى لفترة الثقة:

$$(\overline{X}_1 - \overline{X}_2) + \ t_{(\frac{\alpha}{2},\upsilon)}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})} = (1 - 1.2) + 1.895.\sqrt{(\frac{0.0547}{4} + \frac{0.0991}{5})}$$
$$= 0.1468$$

إذن فترة ثقة للفرق بين متوسطي المجتمعين $(\mu_1 - \mu_2)$ عند مستوى الثقة 90% هي بالتقريب:

[-0.55 ،0.15]

مثال (11): أجري امتحان في مادة الإحصاء لمجموعتين مستقلتين من الطلبة، الأولى تشمل 75 طالبا، والثانية تشمل 50 طالبة، وكانت نتائج هذا الامتحان للمجموعتين كما يلي:

مجموعة الطلبة:	مجموعة الطالبات:
$\overline{X}_1 = 72$	$\overline{X}_2 = 66$
$s_1^2 = 64$	$s_2^2 = 36$

بافتراض أن تباين درجات الطلبة لا يساوي تباين درجات الطالبات (أي تبايـني المجتمعين غيـر متساويين).

المطلوب : أوجد فترة الثقة للفرق بين متوسطـي درجات الطلبة ودرجات الطالبات$(\mu_1 - \mu_2)$ باستخدام مستوى ثقة 90%.

الحل:

بما أن تبايني المجتمعين مجهولين وغير متساويين، وحجم العينتين كبير بدرجة كافية، فان فترة الثقة المطلوبة في هذه الحالة هي:

$$\left[(\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}\ ,(\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}\right]$$

ولدينا: $1\text{-}\ \alpha = 0.96$ فان : $\frac{\alpha}{2} = 0.02$ ، وبالتالي من جدول التوزيع الطبيعي في الملحق رقم (01) يكون : $z_{\alpha/2} = z_{0.02} = 2.055$

وعند التعويض في فترة الثقة السابقة نحصل على:

الحد الأدنى لفترة الثقة:

$$(\overline{X}_1 - \overline{X}_2) - z_{\alpha/2}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})} = (72 - 66) - 2.055.\sqrt{(\frac{64}{75} + \frac{36}{50})}$$
$$= 3.42$$

الحد الأعلى لفترة الثقة:

$$(\overline{X}_1 - \overline{X}_2) + z_{\alpha/2}.\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})} = (72 - 66) + 2.055.\sqrt{(\frac{64}{75} + \frac{36}{50})}$$
$$= 8.58$$

إذن فترة ثقة للفرق بين متوسطي درجات كل الطلبة ، ومتوسط درجات كل الطالبات $(\mu_1 - \mu_2)$ عند مستوى الثقة 96% هي:

[3.42 ،8.58]

4- 3 .فترة الثقة للفرق بين متوسطي مجتمعين $(\mu_1 - \mu_2)$ باستعمال عينتين مزدوجتين أو مقرونتين :

في الكثير من الدراسات عند مقارنة متوسطي مجتمعين ، نجد أن العينتين غير مستقلتين ، وتكون كل قيمة في العينة الأولى مقرونة بقيمة في العينة الثانية ، أي تكون القيمة الأولى في العينة الأولى والقيمة الأولى في العينة الثانية تابعتين لنفس المفردة ، والقيمة الثانية في العينة الأولى والقيمة الثانية في العينة الثانية تابعتين لنفس المفردة وهكذا ،...، أي نجد أن القيم المشاهدة في العينتين تكون في شكل أزواج من القيم ولذلك سميت العينتان في هذه الحالة بالعينتين **المزدوجتين أو المقرونتين** (paired sample) ، وبالطبع في هذه الحالة، يجب أن تكون العينتان متساويتين في الحجم.

ومن الحالات التي تستخدم فيها العينتين المزدوجتين ما يلي:

- دراسة خاصة بالفرق بين طول اليد اليمنى وطول اليد اليسرى للأشخاص ، ففي هذه الحالة نختار عينة تشمل n من الأشخاص ، ولكل شخص نقيس طول يده اليمنى وطول يده اليسرى ، فستكون لدينا **عينتان** ، العينة الأولى تشمل أطوال اليد اليمنى ، والعينة الثانية تشمل أطوال اليد اليسرى ، وسنجد أن القيمة الأولى في العينة الأولى تمثل طول اليد اليمنى للشخص الأول، والقيمة الأولى في العينة الثانية تمثل طول اليد اليسرى لنفس الشخص ، أي أن القيمة الأولى في العينة الأولى والقيمة الأولى في العينة الثانية **تابعتان** لنفس الشخص ، أي قيمتان **مقرونتان تمثل زوجا من القيم** ، وهكذا بالنسبة لبقية الأشخاص ، وبالتالي يكون لدينا n زوج من القيم ، وتكون العينتان مزدوجتين.
- دراسة خاصة بتأثير الإعلان على الكميات المباعة من سلعة معينة ، فنختار عشوائيا عينة تشمل n من المحلات التي تبيع هذه السلعة ، ونجمع بيانات عن الكمية المباعة في هذه المحلات خلال فترة معينة ، ثم نقوم بحملة إعلانية عن هذه السلعة ، وبعد هذه الحملة الإعلانية نقوم بتجميع البيانات عن الكمية المباعة في نفس المحلات السابقة لنفس المدة ، وبالتالي سيكون لدينا أمام كل محل قيمتان ،الأولى خاصة بالكمية المباعة قبل الإعلان ، والثانية خاصة بالكمية المباعة بعد الإعلان ، أي سيكون لدينا n زوج من القيم ، وتكون العينتان مزدوجتين.

وفي اغلب الحالات التي تكون فيها العينتان مزدوجتين ، يكون الهدف هو **المقارنة بين وسطين حسابيين لمجتمعين قبل وبعد إجراء معين.**

وإذا رمزنا لقيم المجتمع الأول بالرمز X_1 ، ولقيم المجتمع الثاني بالرمز X_2 ، وكان حجم كل عينة يساوي n ، فستكون بيانات العينتين المزدوجتين كما يلي :

رقم المفردة (i)	1	2	………..	n
العينة الأولى	X_{11}	X_{12}	………..	X_{1n}
العينة الثانية	X_{21}	X_{22}	………..	X_{2n}

في هذه الحالة يكون لدينا زوج من القيم خاص بكل مفردة ، فالزوج (X_{21}، X_{11}) خاص بالمفردة الأولى،والزوج (X_{22}،X_{12}) خاص بالمفردة الثانية، وهكذا ،...، إلى الزوج (X_{2n}،X_{1n}) أي تتكون البيانات من n زوج من المشاهدات. وبما أن الغرض من الدراسة هو تقدير الفرق بين الوسط الحسابي للمجتمع الأول الذي يمثل الظاهرة الأولى X_1 ، والوسط الحسابي للمجتمع الثاني الذي يمثل الظاهرة الثانية X_2 ، أي تقدير $(\mu_1 - \mu_2)$ ، ففي هذه الحالة نحسب **الفرق بين قيم الظاهرتين التابعتين لكل مفردة** ، فنحصل على **عينة من الفروق** ، وبذلك نكون قد اختزلنا هذه المشكلة من مشكلة عينتين إلى مشكلة عينة واحدة وهي عينة الفروق ، فإذا رمزنا للفرق بين **قيمتي الظاهرتين** بالرمز $\boldsymbol{d}$ ، فستكون قيم عينة الفروق كما يلي :

عينة الفروق:

$$d_1 = X_{11} - X_{21}$$

$$d_2 = X_{12} - X_{22}$$

....

$$d_n = X_{1n} - X_{2n}$$

وبافتراض أن المجتمع الأول يتبع توزيعا طبيعيا والمجتمع الثاني يتبع توزيعا طبيعيا أيضا، فيمكن إثبات أن **توزيع الفروق سيتبع هو الآخر التوزيع الطبيعي**، أي أن مجتمع الفروق سيتوزع **توزيعا طبيعيا** بوسط حسابي قدره $\boldsymbol{\mu_d}$ ، وتباين قدره $\boldsymbol{\sigma_d^2}$ ، حيث:

$$\boldsymbol{\mu_d} = E(d) = E(X_1 - X_2) = E(X_1) - E(X_2) = \boldsymbol{\mu_1 - \mu_2}$$

وأفضل مقدر للوسط الحسابي لمجتمع الفروق ($\mu_d = \mu_1 - \mu_2$) هو الوسط الحسابي لعينة الفروق $\bar{d}$. وبما ان تباين مجتمع الفروق مجهول، فاننا سوف نقدره بتباين عينة الفروق S_d^2 ، ونجد أن المتغير العشوائي التالي:

$$\frac{\bar{d}-\mu_d}{\sqrt{\frac{S_d^2}{n}}} \quad .... (10)$$

يتبع توزيع ستيودنت بدرجات حرية: $\upsilon = n - 1$ ، حيث نجد بأن :

$$\bar{d} = \frac{\sum d_i}{n} \quad .... (11)$$

$$S_d^2 = \frac{\Sigma(d_i-\bar{d})^2}{n-1} \quad .... (12)$$

وبإجراء نفس الخطوات التي أجريناها في فترات الثقة السابقة، نتوصل إلى النظرية التالية:

نظرية (10): إذا كان لدينا عينة عشوائية من الفروق ($d_1, d_2, ..., d_n$) تم سحبها من مجتمع الفروق وسطه μ_d وتباينه σ_d^2 مجهول ، وكان مجتمع الفروق يتبع التوزيع الطبيعي، فان فترة الثقة $(1-\alpha)100\%$ للفرق بين متوسطي مجتمعين في حالة عينتين مزدوجتين ($\mu_d = \mu_1 - \mu_2$) تعطى بالصيغة التالية:

$$\left[\bar{d} - t_{(\frac{\alpha}{2},\upsilon)}.\frac{S_d}{\sqrt{n}} \ , \ \bar{d} + t_{(\frac{\alpha}{2},\upsilon)}.\frac{S_d}{\sqrt{n}}\right] \quad .... (13)$$

حيث: $\upsilon = n - 1$

مثال (12): توضح البيانات التالية عدد الأخطاء المطبعية لخمس طباعات قبل وبعد الاشتراك في برنامج تدريب خاص بالطباعة:

الطباعة	1	2	3	4	5
قبل التدريب	8	7	4	10	4
بعد التدريب	7	5	6	7	3

المطلوب: أوجد فترة ثقة 95% للفرق بين الوسط الحسابي للأخطاء قبل وبعد الاشتراك في برنامج التدريب وذلك بافتراض أن المجتمعين يتوزعان توزيعا طبيعيا.

الحل:

نقوم أولا بإيجاد عينة الفروق ثم نحسب وسطها الحسابي $\bar{d}$ وتباينها s_d^2 وذلك كما يلي:

الطباعة	X_1	X_2	$d = X_1 - X_2$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$
1	8	7	1	0	0
2	7	5	2	1	1
3	4	6	-2	-3	9
4	10	7	3	2	4
5	4	3	1	0	0

حيث :

$$\bar{d} = \frac{\sum d_i}{n} = \frac{5}{5} = 1$$

$$s_d^2 = \frac{\sum (d_i - \bar{d})^2}{n-1} = \frac{14}{4} = 3.50$$

بما ان العينتين مزدوجتين ، والمجتمعين يتوزعان طبيعيا ، فان فترة ثقة 95% للفرق بين الوسط الحسابي للأخطاء قبل وبعد الاشتراك في برنامج التدريب تحسب كما يلي:

$$\left[\bar{d} - t_{(\frac{\alpha}{2},\upsilon)} \cdot \frac{s_d}{\sqrt{n}} \quad , \quad \bar{d} + t_{(\frac{\alpha}{2},\upsilon)} \cdot \frac{s_d}{\sqrt{n}}\right]$$

بما أن: $1-\alpha = 0.95$ فان : $\frac{\alpha}{2} = 0.025$

ولدينا: $\upsilon = n - 1 = 5 - 1 = 4$

فيكون : $t_{(\frac{\alpha}{2},\upsilon)} = t_{(0.025,4)} = 2.776$

وعند التعويض في فترة الثقة السابقة نحصل على:

الحد الأدنى لفترة الثقة:

$$\bar{d} - t_{(\frac{\alpha}{2},\upsilon)} \cdot \frac{s_d}{\sqrt{n}} = 1 - (2.776) \cdot \frac{\sqrt{3.50}}{\sqrt{5}} = -1.32$$

الحد الأعلى لفترة الثقة:

$$\bar{d} + t_{(\frac{\alpha}{2},\upsilon)} \cdot \frac{s_d}{\sqrt{n}} = 1 + (2.776) \cdot \frac{\sqrt{3.50}}{\sqrt{5}} = 3.32$$

إذن فترة ثقة 95 % للوسط الحسابي للفروق ، أي للفرق بين الوسط الحسابي للأخطاء قبل وبعد الاشتراك في برنامج التدريب هي: [-1.32 ، 3.32]

5- فترة الثقة لتباين المجتمع σ^2 :

علمنا انه إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بتباين معلوم σ^2 ، وسحبنا منه عينة عشوائية حجمها n ، فان المتغير العشوائي χ^2 ، يساوي:

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

والتوزيع الاحتمالي لهذا المتغير يسمى **توزيع كأي تربيع** بدرجة حرية: $\upsilon = n-1$

من الجدول الموضح في الملحق (3) نستطيع الحصول على قيمتين للمتغير χ^2 عند درجة حرية معينة ، بحيث تكون المساحة على يمين القيمة الكبرى تساوي المساحة على يسار القيمة الصغرى $\alpha/2$ ،وإذا رمزنا لكل قيمة باستعمال المساحة التي على يمينها سنرمز للقيمة الكبرى بالرمز: $\chi^2_{\alpha/2}$ ، وللقيمة الصغرى بالرمز: $\chi^2_{1-\frac{\alpha}{2}}$ ، وحيث أن المساحة الكلية تحت منحنى توزيع χ^2 تساوي **الواحد** الصحيح ، إذن المساحة بين هاتين القيمتين تساوي $(1-\alpha)$ ، وذلك كما هو مبين في الشكل(02) ونعبر عن هذه المساحة باستخدام الاحتمال كما يلي:

$$P[\chi^2_{(1-\frac{\alpha}{2},\upsilon)} \leq \chi^2 \leq \chi^2_{(\frac{\alpha}{2},\upsilon)}] = 1-\alpha$$

الشكل (02)

من كتاب الاستنتاج الاحصائي ص185

وبهذا يمكن توضيح فترة الثقة لتباين المجتمع في النظرية التالية:

نظرية (11): إذا كان S^2 هو تباين عينة عشوائية تم سحبها من مجتمع يتوزع توزيعا طبيعيا وسطه μ وتباينه σ^2 فان فترة الثقة $(1-\alpha)100\%$ لتباين المجتمع σ^2 تأخذ الشكل التالي:

$$\left[\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},\upsilon)}} \text{ ، } \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},\upsilon)}} \right] \ldots\ldots(14)$$

مثال (13): إذا علمت أن درجات طلبة المرحلة الابتدائية في مادة الرياضيات تتوزع توزيعا طبيعيا وسحبنا من طلبة هذه المرحلة عينة عشوائية تحتوي على 10 طلاب ، وكانت درجاتهم كما يلي:

69 ، 67 ، 49 ، 81 ، 38 ، 55 ، 65 ، 50 ، 72 ، 44 .

المطلوب: قدر التباين والانحراف المعياري لدرجات كل الطلبة المشتركين في هذا الامتحان، وذلك باستخدام فترة الثقة عند مستوى ثقة 95 % .

الحل:

نلاحظ أن المجتمع محل الدراسة يتكون من درجات كل الطلبة المشتركين في هذا الامتحان، وبما أن هذا المجتمع يتبع التوزيع الطبيعي ، فان فترة الثقـــة لتبـــاين درجات كل الطلبة المشتركين في هذا الامتحان ستكون كما يلي :

$$\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},\upsilon)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},\upsilon)}}$$

نقوم أولا بحساب الوسط الحسابي وتباين العينة:

$$\overline{X} = \frac{\Sigma Xi}{n} = \frac{590}{10} = 59$$

$$S^2 = \frac{\Sigma(x_i - \overline{X})^2}{n-1} = \frac{(69-59)^2+\cdots+(44-59)^2}{10-1}$$
$$= 190.67$$

ولدينا: $1-\alpha = 0.95$ ، $\alpha = 0.05$ ، $\frac{\alpha}{2} = 0.025$ ، $\upsilon = n-1 = 10-1 = 9$

ويكـــــــون: $\chi^2_{(\frac{\alpha}{2},\upsilon)} = \chi^2_{(0.025,9)} = 19.023$

$$\chi^2_{(1-\frac{\alpha}{2},\upsilon)} = \chi^2_{(0.975,9)} = 2.70$$

وبالتالي فان الحد الأدنى لهذه الفترة هو: $\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},\upsilon)}} = \frac{(10-1)190.67}{19.023} = 90.21$

و الحد الأعلى هو: $\frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},\upsilon)}} = \frac{(10-1)190.67}{2.70} = 635.56$

وعليه فان فترة الثقة لتباين المجتمع عند مستوى الثقة 95% هي: [90.21 ، 635.56]

أي نستطيع القول بثقة قدرها 95% بأن تباين درجات كل الطلبة المشتركين في هذا الامتحان يقع بين القيمتين 90.21 و 635.56 .

وبأخذ الجذر التربيعي للحدين نحصل على حدي فترة الثقة للانحراف المعياري ، وبالتالي فان فترة الثقة للانحراف المعياري لدرجات كل الطلبة المشتركين في هذا الامتحان عند مستوى الثقة 95% هي:

[9.50 ، 25.21]

6- فترة الثقة للنسبة بين تبايني المجتمعين $\frac{\sigma_1^2}{\sigma_2^2}$:

إذا كان لدينا مجتمعان يتوزعان طبيعيا، وسحبنا من المجتمع الأول الذي تباينه σ_1^2 كل العينات العشوائية الممكنة ذات الحجم n_1 ، وحسبنا منها قيم المتغير s_1^2 ، ثم سحبنا من المجتمع الثاني الذي تباينه σ_2^2 كل العينات العشوائية الممكنة ذات الحجم n_2 ، وحسبنا منها قيم المتغير s_2^2 وكانت العينات المسحوبة من المجتمع الأول مستقلة عن العينات المسحوبة من المجتمع الثاني، فان المتغير F يمكن تعريفه كما يلي :

$$\mathrm{F} = \frac{s_1^2}{s_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}$$

ومن الجدول الموضح في الملحق رقم (4) نستطيع الحصول عل قيمتين للمتغيرF عند درجتي الحرية υ_1 و υ_2 ، حيث υ_1 تمثل درجة حرية البسط وυ_2 تمثل درجة حرية المقام ، بحيث تكون المساحة الواقعة على يمين القيمة الكبرى تساوي المساحة الواقعة عل يسار القيمة الصغرى وتساوي $\alpha/2$ ، وإذا رمزنا لكل قيمة باستعمال المساحة التي على يمينها سنرمز للقيمة الكبرى بالرمز $F_{(\frac{\alpha}{2})}$ ، وللقيمة الصغرى بالرمز $F_{(1-\frac{\alpha}{2})}$ ، وبما أن المساحة الكلية تحت منحنى توزيع F تساوي **الواحد** الصحيح ، فان المساحة بين هاتين القيمتين تساوي $(1-\alpha)$ ، وذلك كما هو مبين في الشكل(03) الموالي ، ونعبر عن هذه المساحة باستخدام الاحتمال كما يلي:

$$\mathrm{P}[F_{(1-\frac{\alpha}{2},\upsilon_1,\upsilon_2)} \leq \mathrm{F} \leq F_{(\frac{\alpha}{2},\upsilon_1,\upsilon_2)}] = 1-\alpha$$

الشكل (03)

من كتاب الاستنتاج الاحصائي ص188

وبهذا يمكن توضيح فترة الثقة للنسبة بين تبايني المجتمعين في النظرية التالية:

نظرية (12): إذا كان s_1^2 هو تباين عينة عشوائية حجمها n_1 تم سحبها من مجتمع يتوزع توزيعا طبيعيا ذو تباين σ_1^2 ، وكان s_2^2 هو تباين عينة عشوائية حجمها n_2 مستقلة عن الأولى تم سحبها من مجتمع يتوزع توزيعا طبيعيا ذو تباين σ_2^2 فان فتـــرة الثقـــة $(1-\alpha)100\%$ للنسبة بين تبايـــي المجتمعين $\frac{\sigma_1^2}{\sigma_2^2}$ تأخذ الشكل التالي:

$$\left[\ \frac{s_1^2/\,s_2^2}{F_{(\frac{\alpha}{2},\upsilon_1,\upsilon_2)}}\ ،\ \frac{s_1^2/\,s_2^2}{F_{(1-\frac{\alpha}{2},\upsilon_1,\upsilon_2)}}\ \right] ....(15)$$

مثال (14): إذا علمت أن مجتمع طول الطالبات ومجتمع طول الطلبة في جامعة محمد خيضر ببسكرة يتبع التوزيع الطبيعي ، سحبنا من مجتمع الطالبات عينة عشوائية تشمل 25 طالبة ، ومن الطلبة عينة عشوائية تشمل 21 طالبا وكانت العينتان مستقلتين ، ووجدنا أن تباين الطول لعينة الطالبات يساوي 64 ، وتباين الطول لعينة الطلبة يساوي 36 .

المطلوب :أوجد فترة الثقة لنسبة تباين مجتمع طول الطالبات إلى تباين مجتمع طول الطلبة ، وذلك باستخدام مستوى ثقة 95% .

الحل:

بما أن المجتمعين يتوزعان توزيعا طبيعيا ، والعينتين مستقلتان ، إذن فترة الثقة المطلوبة هي :

$$\frac{s_1^2/\,s_2^2}{F_{(\frac{\alpha}{2},\upsilon_1,\upsilon_2)}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2/\,s_2^2}{F_{(1-\frac{\alpha}{2},\upsilon_1,\upsilon_2)}}$$

لدينا: $\upsilon_1 = n_1 - 1 = 25 - 1 = 24\ ,\ \upsilon_2 = n_2 - 1 = 21 - 1 = 20$

وبما أن : $1-\alpha = 0.95$ ، $\alpha = 0.05$ ، $\frac{\alpha}{2} = 0.025$ ، يكـــــون :

$$F_{(\frac{\alpha}{2},\upsilon_1,\upsilon_2)} = F_{(0.025,24,20)} = 2.40$$

$$F_{(1-\frac{\alpha}{2},\upsilon_1,\upsilon_2)} = F_{(0.975,24,20)} = 0.43$$

وبالتالي فان الحد الأدنى لفترة الثقة هو:
$$\frac{s_1^2/\ s_2^2}{F_{(\frac{\alpha}{2},\upsilon_1,\upsilon_2)}} = \frac{64/36}{2.40} = 0.74$$

و الحد الأعلى هو:
$$\frac{s_1^2/\ s_2^2}{F_{(1-\frac{\alpha}{2},\upsilon_1,\upsilon_2)}} = \frac{64/36}{0.43} = 4.13$$

وعليه فان فترة الثقة لنسبة تباين مجتمع طول الطالبات إلى تباين مجتمع طول الطلبة $\frac{\sigma_1^2}{\sigma_2^2}$ في جامعة بسكرة عند مستوى الثقة 95% هي: [4.13 ، 0.74]

أي نستطيع القول بثقة قدرها 95% بأن النسبة بين تباين طول كل الطالبات إلى تباين طول كل الطلبة في جامعة بسكرة يقع بين القيمتين 0.74 و 4.13 .

7- فترة الثقة لنسبة المجتمع P :

من خلال دراستنا لتوزيعات المعاينة وجدنا أن توزيع المعاينة لنسبة العينة $\bar{p}$ يقترب من التوزيع الطبيعي عندما يكون حجم العينة كبيرا ، وذلك بوسط حسابي وتباين قدرهما على التوالي كما يلي:

$$\sigma_{\bar{p}}^2 = \frac{\mathbf{Pq}}{\mathbf{n}} \quad ، \quad \mu_{\bar{p}} = P$$

وبالتالي فان المتغير العشوائي Z ، حيث :

$$\mathbf{Z} = \frac{\bar{p} - \mu_{\bar{p}}}{\sqrt{\sigma_{\bar{P}}^2}}$$

سيتبع تقريبا التوزيع الطبيعي المعياري.

بما أن نسبة المجتمع P مجهولة، وهي التي نرغب في تقديرها بإيجاد فترة الثقة لها، فلا نستطيع حساب تباين توزيع المعاينة لنسبة العينة $\sigma_{\bar{p}}^2 = \frac{Pq}{n}$ ، ولكننا سنقدره باستخدام أفضل مقدر لنسبة المجتمع المجهولة P ، وهو نسبة العينة $\bar{p}$ ، وعند استخدام مقدر التباين وهو :

$$\widehat{\sigma_{\bar{P}}^2} = \frac{\bar{p}\bar{q}}{n} \quad \text{حيث} \quad \bar{p} = 1 - \bar{q}$$

في صيغة Z السابقة سيتوزع المتغير الجديد توزيعا قريبا من التوزيع الطبيعي المعياري ، عندما يكون حجم العينة كبير ، وبهذا يمكن توضيح فترة الثقة لنسبة المجتمع في النظرية التالية:

نظرية (13): إذا كانت $\bar{p}$ تمثل نسبة النجاح في عينة عشوائية حجمها n كبير بدرجة كافية سحبت من توزيع ذي الحدين، فان فترة الثقـــة$(1-\alpha)100\%$ التقريبية لنسبة المجتمع P (أي نسبة النجاح في المجتمع) هي كما يلي:

$$\left[\bar{p} - z_{\alpha/2}.\sqrt{\left(\frac{\bar{p}\bar{q}}{n}\right)} \quad , \quad \bar{p} + z_{\alpha/2}.\sqrt{\left(\frac{\bar{p}\bar{q}}{n}\right)}\right] \quad \ldots\ldots (16)$$

وبالطبع إذا كان السحب تم مع عـــدم الإرجـــاع و $\frac{n}{N} \geq 0.05$ ، يجب استعمال معامـــل التصحيح.

مثال (15): إذا اخترنا عشوائيا 500 طالب من طلبة جامعة بسكرة، ووجدنا أن 180 منهم يملكون هواتف نقالة، فقدر نسبة الطلبة الذين يملكون هواتف نقالة في جامعة بسكرة ككل باستخدام مستوى ثقة 99% .

الحل:

بما أن حجم العينة كبير، فان فترة الثقة المطلوبة تأخذ الشكل التالي:

$$\left[\bar{p} - z_{\alpha/2}.\sqrt{\left(\frac{\bar{p}\bar{q}}{n}\right)} \quad , \quad \bar{p} + z_{\alpha/2}.\sqrt{\left(\frac{\bar{p}\bar{q}}{n}\right)}\right]$$

وفي هذه الحالة لا نستخدم معامل التصحيح، لأن حجم المجتمع كبير ، ولدينا كذلك :

$$n = 500 \quad , \quad \bar{p} = \frac{180}{500} = 0.36$$

و $1-\alpha = 0.99$ فان : $\frac{\alpha}{2} = 0.005$ ، وبالتالي من جدول التوزيع الطبيعي المعياري الموضح في الملحق رقم (01) يكـــون : $z_{\alpha/2} = z_{0.005} = 2.58$

وعند التعويض في فترة الثقة السابقة نحصل على:

الحد الأدنى لفترة الثقة:

$$\left[\bar{p} - z_{\alpha/2}.\sqrt{\left(\frac{\bar{p}\bar{q}}{n}\right)}\right] = 0.36 - 2.58\sqrt{\frac{(0.36)(0.64)}{500}}$$
$$= 0.3046$$

الحد الأعلى لفترة الثقة:

$$\left[\bar{p}+ z_{\alpha/2}\cdot\sqrt{(\frac{\bar{p}\bar{q}}{n})}\right] = 0.36 +2.58\sqrt{\frac{(0.36)(0.64)}{500}}$$
$$= 0.4154$$

وعليه يمكن القول بثقة 99% بأن نسبة الطلبة الذين يملكون هواتف نقالة في جامعة بسكرة تقع بين النسبتين 30.46% و 41.46% .

- **تحديد حجم العينة المناسب لتقدير نسبة المجتمع:**

من الجدير بالذكر أن P غالبا تكون مجهولة ، وإذا كانت هناك معلومات سابقة من دراسات مماثلة فبالإمكان استخدام قيمة P المعروفة من تلك الدراسات السابقة ، أما إذا لم تكن هنالك أية فكـــرة عن قيمة P فإننا نأخذ بمبدأ أسوأ الأوضاع وهو أن تكون قيمة P= ½ لأن ذلك يؤدي إلى أكبر خطـــأ معياري للمقدر .

ولهذا إذا تم تحديد المقدار الأكبر المسموح به للخطأ في تقدير P ، أمكن حساب حجم العينة اللازمة لتحقيق ذلك الحد ، حيث تكون :

$$n \geq \left(\frac{z_{\alpha/2}}{E}\right)^2 .P(1-P).....(17)$$

إذا كانت P معلومة من دراسات سابقة .

أما إذا كانت P مجهولة بالكامل فان :

$$n \geq \frac{1}{4}\left(\frac{z_{\alpha/2}}{E}\right)^2 ....(18)$$

مثال (16): نريد القيام بدراسة طبية لتقدير نسبة المواطنين الذين يعانون من مشاكل في النظر، كم شخصا يجب فحصهم كي نكون واثقين بنسبة 98% أن الخطأ في تقدير هذه النسبة لا يزيد عن 0.05 في الحالتين التاليتين :

1- إذا لم يكن لدينا أية معلومات سابقة عن هذه النسبة.

2- إذا كنا نعلم من دراسات سابقة أن هذه النسبة قد تكون حوالي 0.3 .

الحل:

لدينا: $1-\alpha = 0.98$ فان : $\frac{\alpha}{2} = 0.01$ ، ومنه $z_{\alpha/2} = z_{0.01} = 2.33$

-1 $E = 0.05$, $P = \frac{1}{2}$

فتكون:
$$n \geq \frac{1}{4}\left(\frac{z_{\alpha/2}}{E}\right)^2$$

$$n \geq \frac{1}{4}\left(\frac{2.33}{0.05}\right)^2$$
$$n \geq 542.89$$

$$n = 543$$

-2 $P = 0.3 \quad , \quad E = 0.05$

فتكون:
$$n \geq \left(\frac{z_{\alpha/2}}{E}\right)^2 .P(1-P)$$

$$n \geq \left(\frac{2.33}{0.05}\right)^2 .(0.3)(0.7)$$

$$n \geq 456.0276$$

$$n = 456$$

8- فترة الثقة للفرق بين نسبتي مجتمعين $P_1 - P_2$:

بنفس الطريقة السابقة، فان إيجاد فترة الثقة $(1-\alpha)100\%$ للفرق بين نسبتي المجتمعين $P_1 - P_2$ يعتمد على النظرية التالية :

نظرية (14): إذا أخذت عينة عشوائية حجمها n_1 من توزيع ذي الحدين $B(1, P_1)$ ، ثم أخذت عينة عشوائية أخرى مستقلة عن الأولى حجمها n_2 من توزيع آخر ذي الحدين $B(1, P_2)$ وكان حجم العينتين كبير ، فان فترة الثقة $(1-\alpha)100\%$ للفـــرق بين نسبـــتي المجتمعين $P_1 - P_2$ هي:

$$(\bar{p}_1 - \bar{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\bar{p}_1\,\bar{q}_1}{n_1} + \frac{\bar{p}_2\,\bar{q}_2}{n_2}} \quad ....(19)$$

ولتوضيح ذلك نأخذ المثال التالي :

مثال (17):سحبت عينتان عشوائيتان مستقلتان مع الإرجاع، الأولى تحتوي على 120 وحدة منتجة بالآلـــة (أ) ووجـــدنا بها 6 وحدات معيبة، والثانية تحتوي على 200 وحدة منتـــجة بالآلـــة (ب)

ووجدنا بها 9 وحدات معيبة ، إذن قدر الفرق بين نسبة الوحدات المعيبة في الإنتاج الكلي للآلة (أ) و نسبة الوحدات المعيبة في الإنتاج الكلي للآلة (ب) ، وذلك باستخدام مستوى ثقة 95% .

الحل:

بما أن العينتين العشوائيتين مستقلتين ، وحجم العينتين كبير ، فان فترة الثقة المطلوبة هي :

$$(\bar{p}_1 - \bar{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\bar{p}_1\,\bar{q}_1}{n_1} + \frac{\bar{p}_2\,\bar{q}_2}{n_2}}$$

ومن البيانات المتوافرة لدينا نستطيع حساب ما يلي:

نسبة الوحدات المعيبة لعينة إنتاج الآلــة (أ) هي: $\bar{p}_1 = \frac{6}{120} = 0.05$

نسبة الوحدات المعيبة لعينة إنتاج الآلــة (ب) هي: $\bar{p}_2 = \frac{9}{200} = 0.045$

ولدينا: $1-\alpha = 0.95$ فان : $\frac{\alpha}{2} = 0.025$ ، ومنه $z_{\alpha/2} = z_{0.025} = 1.96$

وعند التعويض في فترة الثقة السابقة نحصل على:

$$(\bar{p}_1 - \bar{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\bar{p}_1\,\bar{q}_1}{n_1} + \frac{\bar{p}_2\,\bar{q}_2}{n_2}} = (0.005) \pm 1.96 \sqrt{\frac{(0.05)(0.95)}{120} + \frac{(0.045)(0.955)}{200}}$$

$$\Rightarrow (P_1 - P_2) \in [-0.0434 \ , \ 0.0534]$$

أي نستطيع القول بثقة قدرها 95% أن الفرق بين نسبة الوحدات المعيبة في الإنتاج الكلي للآلة (أ) و نسبة الوحدات المعيبة في الإنتاج الكلي للآلة (ب) ، يقع بين النسبتين -4.34% و 5.34% .