المحاضرة التاسعة:
الفصل الثاني " التقدير الإحصائي"
يتبع ....

5-فترة الثقة لتباين المجتمع σ^2

من المعلوم أنه إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا بتباين معلومσ^2 ، وسحبنا منه عينة عشوائية حجمها n ، فإن المتغير العشوائيχ^2 ، يساوي:

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

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

من جدول **"توزيع كاي تربيع"** نستطيع الحصول على قيمتين للمتغيرχ^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: $\chi^2_{(\frac{\alpha}{2})}$، 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:

$$(1-\alpha)$$

ونعبر عن هذه المساحة باستخدام الإحتمال كما يلي:

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

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 فإن فترة الثقة$100(1-\alpha)\%$ لتباين المجتمع σ^2 تأخذ الشكل التالي:

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

مثال (12): 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:

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

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، فإن فترة الثقة لتباين 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: $1-\alpha = 0.95$ ، $\alpha = 0.05$ ، $\frac{\alpha}{2} = 0.025$

$$\upsilon = n-1 = 10-1 = 9$$

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

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

وبالتالي فإن 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 فإن فترة الثقة لتباين المجتمع عند مستوى الثقة 95% هي:

$$[90.21 \quad ، \quad 635.56]$$

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 فإن فترة الثقة للانحراف المعياري لدرجات كل الطلبة المشتركين في هذا الامتحان عند مستوى الثقة 95% هي:

$$[9.50 \quad ، \quad 25.21]$$

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، فإن المتغير F يمكن تعريفه كما يلي :

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

من جدول **"توزيع فيشر"** نستطيع الحصول عل قيمتين للمتغير F عند درجتي الحرية υ_1 و υ_2 ، حيث υ_1 تمثل درجة حرية البسط، و υ_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- α) . ونعبر عن هذه المساحة باستخدام الاحتمال كما يلي:

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

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 فإن فترة الثقة (1- α100%) للنسبة بين تبايني المجتمعين $\frac{\sigma_1^2}{\sigma_2^2}$ تأخذ الشكل التالي:

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

مثال: 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 مستقلتين، إذن فترة الثقة المطلوبة هي:

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

لدينا:

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

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

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

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

وبالتالي فإن الحد الأدنى لفترة الثقة هو:

$$\frac{s_1^2/s_2^2}{F_{(\frac{\alpha}{2},\upsilon_1,\upsilon_2)}} = \frac{64/36}{2.40} = 0.74$$

والحد 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 فإن 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: [0.74 ; 4.13] ، 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 p يقترب من التوزيع الطبيعي عندما يكون حجم العينة كبيرا، وذلك بوسط حسابي وتباين قدرهما على التوالي كما يلي:

$$-\ \mu_{\bar{p}}=P \qquad -\ \ \sigma_{\bar{p}}^2=\frac{Pq}{n}$$

وبالتالي فإن المتغير العشوائي Z ، حيث:

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

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 $\bar{p}$.

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، فإن فترة الثقة $100(1-\alpha)\%$ التقريبية لنسبة المجتمع P (أي نسبة النجاح في المجتمع) هي كما يلي:

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

وبالطبع إذا كان السحب تم مع عدم الإرجاع و $\frac{n}{N} \geq 0.05$ ، يجب استعمال معامل التصحيح.

مثال: 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، فإن فترة الثقة المطلوبة تأخذ الشكل التالي:

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

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

$$\bar{p} = \frac{180}{500} = 0.36$$

و $1-\alpha = 0.99$ فإن : $\frac{\alpha}{2} = 0.005$ ، وبالتالي من جدول التوزيع الطبيعي المعياري يكون:

$$z_{\alpha/2} = z_{0.005} = 2.58$$

وعند التعويض في فترة الثقة السابقة نحصل على:

الحد الأدنى لفترة الثقة:

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

الحد الأعلى لفترة الثقة:

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

وعليه يمكن القول بثقة 99% بأن نسبة الطلبة الذين يملكون هواتف نقالة في جامعة بسكرة تقع بين النسبتين 30.46% و 41.54% .

- **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 فكرة عن قيمة P فإننا نأخذ بمبدأ أسوأ الأوضاع وهو أن تكون قيمة P= ½ لأن ذلك يؤدي إلى أكبر خطأ معياري للمقدر.

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

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

إذا كانت P معلومة من دراسات سابقة. أما إذا كانت P مجهولة بالكامل فإن:

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

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

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

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

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

$$n \geq 542.89 \Rightarrow n = 543$$

2- لدينا $P = 0.3$, $E = 0.05$

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

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

$$n \geq 456.0276 \Rightarrow n = 456$$

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

بنفس الطريقة السابقة، فإن إيجاد فترة الثقة $100(1-\alpha)\%$ 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، فإن فترة الثقة $100(1-\alpha)\%$ للفرق بين نسبتي المجتمعين $(P_1 - P_2)$ هي:

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