المحاضرة الرابعة: توزيعات المعاينة (يتبع)

3-2-3. توزيع المعاينة للفرق $(\overline{X}_1 - \overline{X}_2)$ عندما يكون تبايني المجتمعين مجهولين والعينتان مستقلتان وصغيرتا الحجم:

إن قيمة التباينات في المجتمعات غالبا تكون مجهولة، وعليه عند تقدير الانحراف المعياري للفرق $(\overline{X}_1 - \overline{X}_2)$ علينا أن نستخدم تبايني العينتين، بالإضافة إلى ذلك إذا كانت العينتين مستقلتين وصغيرتا الحجم فإن توزيع المعاينة لـ $(\overline{X}_1 - \overline{X}_2)$ ليس له توزيع طبيعي ،بل له توزيع **ستيودنت** (t) بدرجات حرية $\boldsymbol{\upsilon}$.

وبالتالي، عند دراسة توزيع المعاينة للفرق $(\overline{X}_1 - \overline{X}_2)$ لما يكون تبايني المجتمعين مجهولين وحجم العينتين صغير لدينا حالتين هما:

1- تبايني المجتمعين متساويين.
2- تبايني المجتمعين غير متساويين.

1-حالة تبايني المجتمعين متساويين ومجهولين

ليكن σ_1^2 هو التباين الخاص بالمجتمع الأول وهو مجهول القيمة، وكان σ_2^2 هو التباين الخاص بالمجتمع الثاني وهو مجهول القيمة أيضا، وكان $\boldsymbol{\sigma_1^2} = \boldsymbol{\sigma_2^2}$، فإن التباين σ^2 يمثل التباين المشترك لهما، وبالتالي فهو مجهول أيضا، حيث نضع:

$$\sigma^2 = \sigma_1^2 = \sigma_2^2$$

ومنه تكون صيغة الانحراف المعياري للفرق $(\overline{X}_1 - \overline{X}_2)$ على الشكل التالي:

$$\sigma_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} = \sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad .... (21)$$

وبما أن التباين المشترك σ^2 مجهول القيمة فإننا نستطيع تقديره باستخدام تبايني العينيتين $(S_1^2 ، S_2^2)$ كما في الصيغة الموالية، حيث يكون تقدير التباين هو متوسط مرجع للقيم S_1^2

و s_2^2 وتكون الترجيحات مبنية على أساس حجم العينات. ولكي يكون تقدير التباين تقديرا غير متحيز لـ σ^2 فإننا يجب أن نستخدم درجات الحرية (n_1-1) و (n_2-1) كترجيحات بدلا من إستخدام العينيتين: n_1 ، n_2 بشكل مباشر.

وبناء على ذلك فإن مقدر التباين المشترك (σ^2) هو s_p^2 ويُحسب بالصيغة التالية:

$$s_p^2 = \frac{(n_1-1)\,s_1^2+(n_2-1)\,s_2^2}{(n_1-1)+(n_2-1)}$$

ومنه يكون:

$$s_p^2 = \frac{(n_1-1)\,s_1^2+(n_2-1)\,s_2^2}{n_1+n_2-2}$$

حيث يسمى الإحصاء s_p^2 **بتباين العينة التجميعي**"pooled sample variance" وذلك لأنه مكون من تجميع المعلومات عن العينتين معا.

وتأسيسا على ما سبق، يكون تقدير الخطأ أو الانحراف المعياري للفرق$(\bar{X}_1-\bar{X}_2)$ على الشكل التالي:

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \quad \ldots\ldots\ldots\ldots\ldots\ldots (22)$$

نظرية (06): إذا كان $\bar{X}_1$ هو الوسط الحسابي لعينة عشوائية حجمهاn_1 سحبت من مجتمع وسطه μ_1 وتباينه σ_1^2، وكان $\bar{X}_2$ هو الوسط الحسابي لعينة عشوائية حجمهاn_2 سحبت من مجتمع آخر مستقل عن المجتمع الأول وسطه μ_2 وتباينه σ_2^2، وكان حجم العينتين صغير وتبايني المجتمعين مجهولين ومتساويين، فإن المتغير:

$$[(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)] \Big/ \sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \quad \ldots\ldots(23)$$

يخضع لتوزيع **ستيودنت (t)** بدرجة حرية: $\upsilon = n_1+n_2-2$

مثال (07): سحبت عينة عشوائية حجمها 16 وحدة من مجتمع طبيعي وسطه 30 وتباينه σ_1^2 مجهول، وسحبت أيضا عينة عشوائية حجمها 25 وحـدة من مجتمـع طبيـعي آخـر مستقـل عن الأول وسطــه 28 وتباينـــه σ_2^2

مجهول أيضا. وكان $\overline{X}_1$ و $\overline{X}_2$ هما الوسطين الحسابيين للعينتين الأولى والثانية على الترتيب، وتباين العينة الأولى هو 4، وتباين العينة الثانية هو 7.

المطلوب: إذا كان تبايني المجتمعين متساويين، أوجد احتمال أن الفرق بين متوسطي العينتين يكون أقل من 3.

الحل:

لدينا:

$$n_1 = 16 \text{ ،} \; n_2 = 25 \text{،} \; s_1^2 = 4 \text{ ،} \; s_2^2 = 7$$

$$\mu_1 = 30 \text{ ،} \; \mu_2 = 28$$

بما أن تبايني المجتمعين مجهولين ومتساويين فإننا نستخدم توزيع ستيودنت t بدرجة حرية $(n_1 + n_2 - 2)$ وذلك لإيجاد الاحتمال التالي:

$$P\left(\bar{X}_1 - \bar{X}_2 < 3\right)$$

لدينا:

$$P\left(\bar{X}_1 - \bar{X}_2 < 3\right) = P\left(\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} < \frac{(3) - (30 - 28)}{\sqrt{s_p^2\left(\frac{1}{16} + \frac{1}{25}\right)}}\right)$$

الآن نقوم بحساب s_p^2 ، حيث:

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}$$

$$s_p^2 = \frac{(16-1)4 + (25-1)7}{16+25-2} = 5.84$$

ومنه يكون:

$$\boldsymbol{P}\left(\bar{X}_1 - \bar{X}_2 < 3\right) = P\left(T < \frac{1}{\sqrt{5.84\left(\frac{1}{16} + \frac{1}{25}\right)}}\right) = P(T < 1.30)$$

عند درجات الحرية: $(n_1+_2-2) \;=16+25-2= 39$

نجد: $\boldsymbol{P}\left(\bar{X}_{\mathbf{1}}-\bar{X}_{\mathbf{2}}<3\right)=0.9$

2- حالة تبايني المجتمعين غير متساويين ومجهولين:

نظرية (7): إذا كان $\overline{\mathrm{X}}_1$ و $\overline{\mathrm{X}}_2$ هما الوسطين الحسابين لعينتين مستقلتين صغيرتا الحجم تم سحبهما من مجتمعين ذو متوسطين μ_1 و μ_2، وذو تباينين مجهولين وغير متساويين فإننا نستطيع تقدير الانحراف المعياري للفرق $(\overline{\mathrm{X}}_1-\overline{\mathrm{X}}_2)$ باستخدام تبايني العينتين مباشرة $(s_1^2 \;\text{و}\; s_2^2)$ وذلك كما يلي:

$$\sigma_{\overline{\mathrm{X}}_1-\overline{\mathrm{X}}_2}=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \;.......... (24)$$

وعليه فإن المتغير:

$$\left(\frac{(\overline{\mathrm{X}}_1-\overline{\mathrm{X}}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}\right) .... (25)$$

يخضع تقريبا لتوزيع **ستيودنت** بدرجة حرية لها الصيغة المركبة التالية :

$$\upsilon=\left(\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1-1)}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2-1)}}\right) .... (26)$$

مثال(08): إذا كانت نقاط طلبة السنة ثانية (ل.م.د) إدارة أعمال في مقياس الإحصاء لديها وسط حسابي قدره 15، وكانت نقاط طلبة السنة ثانية (ل.م.د) محاسبة في نفس المقياس لديها وسط حسابي قدره 10. وقمنا بسحب عينة من طلبة السنة ثانية (ل.م.د) إدارة أعمال حجمها 25 طالب،

وقمنا أيضابسحب عينة من طلبة السنة ثانية (ل.م.د) محاسبة حجمها 20 طالبا. فإذا كان تباين العينة الأولى هو 6 وتباين العينة الثانية هو 4.

المطلوب: أوجد احتمال أن يكون الفرق بين متوسطي العينتين أكثر من 6، هذا إذا كان تبايني المجتمعين مجهولين وغير متساويين.

الحل:

لدينا:

$$n_1 = 25 \text{ ، } n_2 = 20 \text{، } s_1^2 = 6 \text{ ، } s_2^2 = 4$$

$$\mu_1 = 15 \text{ ، } \mu_2 = 10$$

بما أن تبايني المجتمعين مجهولين وغير متساويين فإننا نستخدم توزيع t بدرجة حرية لها الصيغة المركبة الموضحة في العلاقة (26) وذلك لإيجاد الاحتمال التالي:

$$\boldsymbol{P}\left(\bar{X}_1 - \bar{X}_2 > 6\right)$$

ومنه يكون:

$$\boldsymbol{P}\left(\bar{X}_1 - \bar{X}_2 > 6\right) = P\left(\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} > \frac{(6) - (15 - 10)}{\sqrt{\frac{6}{25} + \frac{4}{20}}}\right)$$

$$= P(T > 1.51)$$

عند درجات الحرية:

$$\boldsymbol{\upsilon} = \left(\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1 - 1)} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2 - 1)}}\right) = \left(\frac{\left(\frac{6}{25} + \frac{4}{20}\right)^2}{\frac{\left(\frac{6}{25}\right)^2}{(25-1)} + \frac{\left(\frac{4}{20}\right)^2}{(20-1)}}\right) = 43$$

نجد:

$$P(\bar{X}_1-\bar{X}_2>6)=0.05$$

3-3. توزيع المعاينة لتباين العينة S^2

نعلم جيدا أن S^2 تقيس الاختلافات ومن ثم فهي تدل على التشتت بين القيم في عينة عشوائية، حيث أن التباين يعتبر من أهم المقاييس الهامة مثله في ذلك مثل مقاييس النزعة المركزية، وبالتالي فإن أهمية S^2 للاستدلال عن σ^2 تُضاهي أهمية $\bar{X}$ عند الاستدلال عن µ. وفيما يلي سنحدد الوسط الحسابي والانحراف المعياري للتباين S^2، ثم نبحث عن توزيع المعاينة لـ S^2.

3-3-1. الوسط الحسابي والانحراف المعياري لتباين العينة

بناء على ما سبق فإن تباين العينة (S^2) والتي حجمها n هو:

$$S^2=\frac{\Sigma(x_i-\bar{X})^2}{n-1}$$

والقيمة المتوقعة لتباين العينة (متوسط القيم لجميع العينات الممكنة ذات الحجم n) هو:

$$E(S^2)=\sigma^2$$

أما الانحراف المعياري لتباين العينة فيعطى بالصيغة التالية:

$$\sigma_{S^2}=\sigma^2.\sqrt{\frac{2}{n-1}} \quad ....\ (27)$$

مثال (09): إذا كان لدينا مجتمع كبير جدا وسطه الحسابي هو 50 وانحرافه المعياري هو 0.5

المطلوب: أحسب القيمة المتوقعة والانحراف المعياري للتباين S^2 لجميع العينات العشوائية الممكنة ذات الحجم 10 والتي يمكن سحبها من هذا المجتمع.

الحل:

$$E(S^2)=(0.5)^2=0.25$$

$$\sigma_{S^2}=\sigma^2.\sqrt{\frac{2}{n-1}}=0.25.\sqrt{\frac{2}{10-1}}=0.1178$$

2-3-3. توزيع المعاينة لـ S^2: مقدمة لتوزيع المعاينة للمتغير χ^2 (كآي تربيع)

نظرية (8): إذا سحبنا عينة عشوائية $(x_1,x_2......,x_n)$ من مجتمع له توزيع طبيعي (هذا شرط أساسي) وسطه μ وتباينه σ^2، وكانت S^2 تمثل تباين العينة، فإن المتغير:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \quad .... (28)$$

له توزيع **كآي تربيع** (Chi-square statistic) بدرجات حرية: $\upsilon = n-1$

المتغير السابق (χ^2) هو دالة في S^2. والرمز χ هو حرف يوناني **كآي**. أما درجات الحرية $(n-1)$ المقترنة بالإحصاء كآي تربيع تعكس حقيقة أن هناك $(n-1)$ من درجات الحرية مقترنة بتباين العينة S^2.

وما يجب الإشارة إليه هو أن **متوسط** توزيع كآي تربيع هو: $(n-1)$، أي:

$$E(\chi^2) = n-1 \quad .... (29)$$

وبخصوص منحنى توزيع χ^2 فإنه غير متماثل ، وهو منحنى ملتو الى اليمين (موجب الالتواء) ويقل هذا الالتواء كلما زادت قيمة درجة الحرية. بعبارة أخرى كلما زادت قيمة درجة الحرية كلما اقترب منحنى χ^2 من التوزيع الطبيعي. إضافة إلى ذلك فإن قيم المتغير العشوائي χ^2 لا تكون سالبة، حيث يبدأ منحنى التوزيع من الصفر على المحور الأفقي ويمتد إلى اليمين، وعند القيم الكبيرة يقترب من المحور الأفقي ولكن لا يلاقيه.

تحسب قيمة المتغير χ^2 لأي عينة باستخدام الصيغة (28) ويكون احتمال أن تكون χ^2 **أكبر** من أي عدد يساوي **المساحة الواقعة تحت منحنى** χ^2 على يمين ذلك العدد. وعادة يُرمز لقيمة χ^2 التي تكون المساحة الواقعة على يمينها مساوية α والمناظرة لدرجات حرية υ بالرمز: $\chi^2_{(,\upsilon)}$

مثال(10): إذا كانت S^2 هو تباين عينة عشوائية ذات الحجم 4 وحدات مسحوبة من مجتمع له توزيع طبيعي تباينه25.

المطلوب:

1- أوجد احتمال أن تباين العينة سوف يكون 2.5 أو أقل.
2- أوجد احتمال أن تباين العينة سوف يتعدى 66.

الحل:

1- بتطبيق الصيغة (28) نجد:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{3S^2}{25}$$

لها توزيع كآي تربيع بدرجات حرية: $\upsilon = n-1 = 4-1 = 3$ ،

ومنه يكون:

$$\boldsymbol{P}(S^2 \leq 2.5) = 1 - P(S^2 > 2.5)$$

لدينا:

$$\boldsymbol{P}(S^2 > 2.5) = \boldsymbol{P}\left(\frac{(n-1)S^2}{\sigma^2} > \frac{3(2.5)}{25}\right)$$

$$= \boldsymbol{P}(\chi^2 > 0.3)$$

وباستخدام جدول توزيع **كآي تربيع** نجد:

$$\boldsymbol{P}(\chi^2 > 0.3) \cong 0.95$$

أي أن:

$$P(S^2 > 2.5) = P(\chi^2 > 0.3) \cong 0.95$$

وفي الأخير نجد:

$$P(S^2 \leq 2.5) = 1 - P(S^2 > 2.5)$$

$$= 1 - 0.95 = 0.05$$

2- إيجاد الاحتمال التالي:

$$P(S^2 > 66) = P\left(\frac{(n-1)S^2}{\sigma^2} > \frac{3(66)}{25}\right) = P(\chi^2 > 7.92)$$

عند درجات الحرية: $\upsilon = 3$ ، نجد: $\boldsymbol{P}(\chi^2 > 7.92) \cong 0.05$

أي أن:

$$\boldsymbol{P}(S^2 > 66) = \boldsymbol{P}(\chi^2 > 7.92) \cong 0.05$$