المحاضرة الثانية عشرة

الفصل الخامس

مقارنة المتوسطات والنسب

Comparing Means And Proportion

في هذا الفصل سوف ندرس مقارنة متوسطي مجتمعين إحصائيين أو نسبتي مجتمعين إحصائيين من خلال بناء مجالات ثقة حول الفرق بين المتوسطين أو الفرق بين النسبتين ، وسوف نختبر فرضيات حول تساوي المتوسطين أو النسبتين وذلك باتباع أسلوب مشابه لما رأيناه في الفصل السابق.

1.5 : مجال الثقة حول الفرق بين متوسطي متغيرين عشوائيين:

في كثير من الأحيان نرغب في مقارنة متوسطي مجتمعين، كمقارنة متوسطي الدخل أو العمر في بلدين مختلفين أو مقارنة جودة الإنتاج لمصنعين أو مقارنة طريقتين في العلاج لمرض معين أو دواءَيْنِ لمعالجة مرض معين، ومن أجل ذلك يلزم تعيين مجال ثقة للفرق بين متوسطي المجتمعين أو مجال ثقة للفرق بين متوسطي المتغيرين العشوائيين الممثلين للمجتمعين المدروسين.

1.1.5: مجال الثقة للفرق بين متوسطي متغيرين عشوائيين طبيعيين تبايناهما معلومان:

إذا كان X متغيراً عشوائياً طبيعياً متوسطه μ_1 وتباينه $\sigma_1{}^2$ ، وكان Y متغيراً عشوائياً طبيعياً متوسطه μ_2 وتباينه $\sigma_2{}^2$ ، ولتكن $X(n_1)$ عينة عشوائية من

X، ولتكن $Y(n_2)$ عينة عشوائية من Y، والعينات مستقلة. وليكن $\overline{X}$ متوسط العينة من X و $\overline{Y}$ متوسط العينة من Y و ${\sigma_1}^2$ و ${\sigma_2}^2$ معلومان.

عندئذٍ $\overline{X}$ و $\overline{Y}$ متغيران عشوائيان مستقلان وأن:

$$\overline{X} \sim N\left(\mu_1 , \frac{{\sigma_1}^2}{n_1}\right) \quad ; \quad \overline{Y} \sim N\left(\mu_2, \frac{{\sigma_2}^2}{n_2}\right)$$

ومِنْ ثَمّ

$$\overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2, \frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}\right)$$

وأيضاً

$$Z = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}} \sim N(0,1)$$

ومن العلاقة الاحتمالية $P\left[-Z_{1-\alpha/2} \leq Z \leq Z_{1-\alpha/2}\right] = 1 - \alpha$

وتعويض Z بالعلاقة أعلاه نجد أن:

$$P\left[-Z_{1-\alpha/2} \leq \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}} \leq Z_{1-\alpha/2}\right] = 1 - \alpha$$

ومنه يكون $(1-\alpha)$ مجال ثقة حول $(\mu_1 - \mu_2)$ من الشكل:

$$(\overline{X} - \overline{Y}) - Z_{1-\alpha/2} . \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} \leq (\mu_1 - \mu_2) \leq$$

$$(\overline{X} - \overline{Y}) + Z_{1-\alpha/2} . \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$$

ملاحظة (1): إذا كان الحد الأدنى لمجال الثقة له القيمة a وكان للحد الأعلى للمجال القيمة b أي

$$a \leq (\mu_1 - \mu_2) \leq b$$

هنا نميز ثلاث حالات:

1) إذا كان $a, b > 0$ فإن $\mu_1 - \mu_2 > 0$ ومنه $\mu_1 > \mu_2$ ، ومِنْ ثَمّ نفسر الناتج وفق هذه النتيجة والثقة المفروضة.

2) إذا كان $a, b < 0$ فإن $\mu_1 - \mu_2 < 0$ ، ومنه $\mu_1 < \mu_2$ ، ومِنْ ثَمّ نفسر الناتج وفق هذه النتيجة والثقة المفروضة.

3) إذا كان $a < 0$ و $b > 0$ عندئذٍ $(1 - \alpha)$ يحوي المجال القيمة صفر، ، ومِنْ ثَمّ باحتمال $(1 - \alpha)$ يمكن لـ $\mu_1 - \mu_2 = 0$ أي $\mu_1 = \mu_2$ ، ومِنْ ثَمّ لا فرق بين المتوسطين.

ملاحظة (2): في الحالة التي لا يكون فيها للمتغيرين العشوائيين التوزيع الطبيعي، وعندما يكون $n_1, n_2 \geq 30$ وبفضل مبرهنة النهاية المركزية يكون أيضاً للفرق بين المتوسطين $(\mu_1 - \mu_2)$ مجال ثقة تقريبي كالمجال الذي توصلنا إليه حول $(\mu_1 - \mu_2)$ أعلاه.

ملاحظة (3): في الحالة التي يكون فيها ${\sigma_1}^2$ و ${\sigma_2}^2$ مجهولين ، وعندما يكون $n_1, n_2 \geq 30$ فيمكن أن نستبدل بـ ${\sigma_2}^2, {\sigma_1}^2$ ، ${S_2}^2, {S_1}^2$ حيث ${S_1}^2$ تباين للعينة العشوائية من X و ${S_2}^2$ تباين للعينة العشوائية من Y. ويصبح $(1 - \alpha)$ مجال ثقة حول $(\mu_1 - \mu_2)$ من الشكل:

$$\left[(\overline{X}-\overline{Y})-Z_{1-\alpha/2}.\sqrt{\frac{{S_1}^2}{n_1}+\frac{{S_2}^2}{n_2}}\quad ,(\overline{X}-\overline{Y})+Z_{1-\alpha/2}.\sqrt{\frac{{S_1}^2}{n_1}+\frac{{S_2}^2}{n_2}}\right]$$

مثال (5 - 1):

أجريت دراسة في إحدى الجامعات للمقارنة بين المعدلات التي حصلت عليها مجموعتان من الطلبة، الأولى من المتزوجين والأخرى من غير المتزوجين ولهذه الغاية أخذت عينتان، واحدة من كل منهما، وبعد إجراء الامتحان كان لدينا النتائج الآتية:

μ_1	$S_1 = 4$	$\overline{X} = 28.5$	$n_1 = 100$	عينة المتزوجين
μ_2	$S_2 = 3$	$\overline{Y} = 27.3$	$n_2 = 100$	عينة غير المتزوجين

والمطلوب:

1) ما تقدير الفرق بين معدلي المجتمعين اللذين أخذت منهما العينتان؟

2) ما الخطأ الأعظمي المرتكب في هذا التقدير بثقة 95%؟

3) أوجد 95% مجال ثقة حول الفرق الحقيقي بين المتوسطين، وفسر الناتج.

الحل:

1) $\hat{\mu}_1 - \hat{\mu}_2 = \overline{X} - \overline{Y} = 28.5 - 27.5 = 1.2$

2) والخطأ الأعظمي المرتكب في تقدير $(\mu_1 - \mu_2)$ وبثقة 95% هو :

$$e = \left| \pm Z_{1-\alpha/2}.\sqrt{\frac{{S_1}^2}{n_1}+\frac{{S_2}^2}{n_2}} \right| = (1.96).\sqrt{\frac{16}{100}+\frac{9}{100}} = 0.98$$

3) إن $1-\alpha = 95\%$ مجال ثقة حول $(\mu_1 - \mu_2)$ وبثقة 95% هو:

$$(\bar{X}_1-\bar{X}_2)-Z_{1-\alpha/2}.\sqrt{\frac{s_1{}^2}{n_1}+\frac{s_2{}^2}{n_2}}\le(\mu_1-\mu_2)\le(\bar{X}_1-\bar{X}_2)+Z_{1-\alpha/2}.\sqrt{\frac{s_1{}^2}{n_1}+\frac{s_2{}^2}{n_2}}$$

$$1.2-0.98\le(\mu_1-\mu_2)\le1.2+0.98$$

$$0.22\le(\mu_1-\mu_2)\le2.18$$

وبما أن طرفي المجال موجبان عندئذٍ بثقة 95% يكون $\mu_1-\mu_2>0$ أي $\mu_1>\mu_2$ ، وهذا يدل على أن الطلبة المتزوجين يحققون معدلات بالنتيجة أفضل من الطلبة غير المتزوجين.

مثال (5- 2):

في اختبار تجريبي في مقرر الإحصاء الحيوي تقدم 75 طالباً و 50 طالبة، فكان متوسط درجات الطلاب 82 درجة بانحراف معياري قدره 8 درجات، بينما كان متوسط درجات الطالبات 76 درجة بانحراف معياري قدره 6 درجات.أوجد 96% مجال ثقة حول الفرق الحقيقي بين متوسطي درجات مجتمع الطلاب ودرجات مجتمع الطالبات، وفسر الناتج.

الحل:

X درجات الطلاب	$n_1=75$	$\bar{X}=82$	$S_1=8$	μ_1
Y درجات الطالبات	$n_2=50$	$\bar{Y}=76$	$S_2=6$	μ_2

وكون العينات أكبر من 30 ومن $1-\alpha=0.96$ فإن $1-\frac{\alpha}{2}=0.98$ حيث من جدول التوزيع الطبيعي المعياري نجد : $z_{1-\alpha/2}=z_{0.98}=2.05$

ومِنْ ثَمّ مجال الثقة حول $(\mu_1-\mu_2)$ وبمستوى 96% يكون من الشكل:

$$(\bar{X}-\bar{Y})-Z_{0.98}.\sqrt{\frac{s_1{}^2}{n_1}+\frac{s^2}{n_2}}\le(\mu_1-\mu_2)\le(\bar{X}-\bar{Y})+Z_{0.98}.\sqrt{\frac{s_1{}^2}{n_1}+\frac{s_2{}^2}{n_2}}$$

$$(82-76)-(2.05).\sqrt{\frac{64}{75}+\frac{36}{50}} \le (\mu_1-\mu_2) \le (82-76)+(2.05).\sqrt{\frac{64}{75}+\frac{36}{50}}$$

$$(3.43) \le (\mu_1-\mu_2) \le (8.75)$$

وبما أن طرفي مجال الثقة موجبان ، فهذا يعني أن $\mu_1-\mu_2 > 0$ أي $\mu_1 > \mu_2$ أي بثقة 96% يكون مستوى الطلاب في الاختبار أفضل من مستوى الطالبات.

2.1.5: مجال الثقة للفرق بين متوسطين متغيرين عشوائيين طبيعيين تبايناهما مجهولان:

في الحالة التي يكون فيها حجما العينتين $n_1, n_2 \ge 30$ ، فيمكن التعويض عن $\sigma_2{}^2, \sigma_1{}^2$ ب $S_2{}^2, S_1{}^2$ كما رأينا في الفقرة السابقة.

أما في الحالة التي يكون فيها $n_1, n_2 < 30$ فإن مجال الثقة حول الفرق $(\mu_1-\mu_2)$ غير واضح باستثناء الحالة التي يكون فيها المتغيران العشوائيان Y , X متجانسين أي التباينات متساوية أي $\sigma_1{}^2=\sigma_2{}^2=\sigma^2$

عندئذٍ يكون للمتغير العشوائي

$$Z=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim N(0,1)$$

وبسهولة نلحظ أن:

$$S_c{}^2=\frac{(n_1-1).S_1{}^2+(n_2-1).S_2{}^2}{n_1+n_2-2}$$

هو مقدّر منصف (غير منحاز) للتباين σ^2 التباين المشترك لـ X ، Y.

وبملاحظة أن المتغيرين العشوائيين : $\frac{(n_1-1)S_1{}^2}{\sigma^2}$ ، $\frac{(n_2-1)S_2{}^2}{\sigma^2}$ مستقلان ،

وأن لهما توزيع كاي_مربع بدرجات من الحرية $(n_1 - 1)$ ، $(n_2 - 1)$ على الترتيب، فيكون لمجموعهما:

$$x^2 = \frac{(n_1-1){S_1}^2}{\sigma^2} + \frac{(n_2-1){S_2}^2}{\sigma^2} = \frac{(n_1+n_2-2){S_c}^2}{\sigma^2}$$

التوزيع كاي _ مربع $\gamma = n_1 + n_2 - 2$ درجة من الحرية ، ويمكن إثبات أن المتغيرين العشوائيين x^2, Z مستقلان وحسب تعريف المتغير T يكون:

$$T = \frac{Z}{\sqrt{x^2/_{n_1+n_2-2}}} = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{Sc\ .\ \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{(\gamma=n_1+n_2-2)}$$

وباستبدال T بما يساويه في العلاقة الاحتمالية الآتية:

$$P\left[-t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2) \le T \le t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)\right] = 1 - \alpha$$

نجد أن $(1 - \alpha)$مجال ثقة حول$(\mu_1 - \mu_2)$من الشكل:

$$(\overline{X} - \overline{Y}) - t_{1-\frac{\alpha}{2}}(\gamma).S_c.\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le (\ \mu_1 - \mu_2) \le$$
$$(\overline{X} - \overline{Y}) + t_{1-\frac{\alpha}{2}}(\gamma).S_c.\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

حيث $(\gamma = n_1 + n_2 - 2)$

مثال (5 - 3):

أجريت على نوعين من الأدوية منتجَيْنِ من قبل شركتين مختلفتين في البلاد ، ولكن بامتياز من نفس الشركة الأم ، وذلك لعلاج مرض معين ، والهدف من الدراسة هو معرفة إذا كان تركيز المادة الفعالة في الدواء هو نفسه، فمن عينة من

الدواء الأول ومن الحجم 10 تبين أن متوسط كمية المادة الفعالة في الحبة يساوي M.G 3.1 بانحراف معياري MG 0.5 وأن متوسط كمية المادة الفعالة في الحبة من الدواء الثاني يساوي MG 2.7 وبانحراف معياري MG 0.7.

فإذا علمنا أن لكمية المادة الفعالة في الدواء لكلا النوعين من الدواء التوزيع الطبيعي وأن لهما التباين نفسه. فأوجد 95% مجال الثقة حول الفرق $(\mu_1 - \mu_2)$ وفسر الناتج، حيث μ_1 هو متوسط كمية المادة الفعالة في مجتمع الدواء الأول و μ_2 هو متوسط كمية المادة الفعالة في مجتمع الدواء الثاني.

الحل: لدينا من فرضيات المسالة:

σ^2	μ_1	$S_1 = 0.5$	$\overline{X} = 3.1$	$n_1 = 10$	X الدواء الأول
σ^2	μ_2	$S_2 = 0.7$	$\overline{Y} = 2.7$	$n_2 = 8$	Y الدواء الثاني

ويكون التباين المشترك:

$$S_c = \sqrt{\frac{(n_1-1){S_1}^2+(n_2-1){S_2}^2}{n_1+n_2-2}} = \sqrt{\frac{(10-1)(0.25)+(8-1)(0.49)}{10+8-2}} = 0.596$$

وأيضاً

$$\overline{X} - \overline{Y} = 3.1 - 2.7 = 0.4$$

$$1 - \alpha = 0.95 \Rightarrow 1 - \frac{\alpha}{2} = 0.975 \;;\; n_1 + n_2 - 2 = 16$$

ومن جدول ستيودنت بـ $\gamma = 16$ درجة من الحرية يكون:

$$t_{1-\frac{\alpha}{2}}(\gamma) = t_{0.975}(16) = 2.12$$

ويكون $1 - \alpha = 0.95$ مجال ثقة حول $(\mu_1 - \mu_2)$ من الشكل:

$$\left(\overline{X}-\overline{Y}\right)-t_{1-\frac{\alpha}{2}}(\gamma).S_c.\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\leq(\mu_1-\mu_2)\leq$$

$$\left(\overline{X}-\overline{Y}\right)+t_{1-\frac{\alpha}{2}}(\gamma).S_c.\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$0.4-(2.12).(0.596).\sqrt{\frac{1}{10}+\frac{1}{8}}\leq(\mu_1-\mu_2)\leq$$

$$0.4+(2.12).(0.596).\sqrt{\frac{1}{10}+\frac{1}{8}}$$

$$-0.2\leq(\mu_1-\mu_2)\leq 1.0$$

وبما أن الحد الأدنى للمجال سالب والحد الأعلى للمجال موجب، عندئذٍ بثقة 95% يكون $\mu_1-\mu_2=0$ أي $\mu_1=\mu_2$.

ومنه لا فرق بين متوسطي كمية المادة الفعالة في الدواءَيْنِ.

3.1.5 : مجال الثقة للفرق بين نسبتي مجتمعين (P_1-P_2):

من المسائل المهمة التي عادة ما تحدث في العمل الإحصائي مقارنة النسب في المجتمعات بالنسبة لصفة معينة، كأن نقارن نسبة الإناث من بين طلاب كلية الصيدلة مع نسبة الإناث في كلية الطب البشري أو نقارن نسبة المصابين بسرطان الرئة من بين المدخنين مع نسبة المصابين بسرطان الرئة من بين غير المدخنين.

وتصبح المسألة تقدير الفرق بين هاتين النسبتين، وإذا ما تذكرنا من أن النسبة في المجتمع توافق وسيط متغير عشوائي برنولي يمثل المجتمع، فإن الدراسة تؤول إلى

تقدير الفرق بين وسيطي متغيرين عشوائيين X_2, X_1 لكل منهما توزيع برنولي بوسيطين P_2 , P_1 على الترتيب.

فإذا أخذنا عينتين عشوائيتين للمتغيرين المستقلين X_2, X_1 حجم الأولى n_1 وحجم الثانية n_2 وكان Y_1 عدد مرات النجاح في العينة الأولى و Y_2 عدد مرات النجاح في العينة الثانية، فإن تقديري P_1, P_2 سيكون:

$$\hat{P}_1 = \frac{Y_1}{n_1} \quad ; \quad \hat{P}_2 = \frac{Y_2}{n_2}$$

ويكون $\left(\widehat{P_1} - \hat{P}_2\right)$مقدراً منصفاً للفرق $\left(P_1 - P_2\right)$.

ومن أجل n_1, n_2 كبيرتين كبراً كافياً $(n_1, n_2 \geq 30)$ وبملاحظة أن X_1, X_2 مستقلة ، وأن لكل منهما تقريباً التوزيع الطبيعي ، وبالاعتماد على مبرهنة النهاية المركزية ، سيكون للمتغير العشوائي $\left(\widehat{P}_1 - \widehat{P}_2\right)$ التوزيع الطبيعي بمتوسط $\mu = P_1 - P_2$ وبتباين: $\sigma^2 = V\left(\widehat{P}_1 - \widehat{P}_2\right) = \frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}$

أي إن :

$$\left(\widehat{P}_1 - \widehat{P}_2\right) \approx N\left(\mu = P_1 - P_2 \quad ; \sigma^2 = \frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}\right)$$

ومنه

$$Z = \frac{\left(\widehat{P}_1 - \widehat{P}_2\right) - \left(P_1 - P_2\right)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} \approx N(0,1)$$

وبالاستفادة من العلاقة الاحتمالية:

$$P\left[-Z_{1-\alpha/2} \leq Z \leq Z_{1-\alpha/2}\right] = 1 - \alpha$$

نحصل على العلاقة المكافئة:

$$P\left[-Z_{1-\alpha/2} \le \frac{(\widehat{P_1}-\hat{P}_2)-(P_1-P_2)}{\sqrt{\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}}} \le Z_{1-\alpha/2}\right] = 1-\alpha$$

ومنه يكون $(1-\alpha)$ مجال ثقة حول (P_1-P_2) من الشكل:

$$(\hat{P}_1-\ \hat{P}_2)-Z_{1-\alpha/2}.\sqrt{\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}} \le (P_1-P_2) \le$$
$$(\widehat{P_1}-\ \hat{P}_2)+Z_{1-\alpha/2}.\sqrt{\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}}$$

ولكن يلحظ أن طرفي المجال تابعان لـ P_1 , P_2 وأن تغيرات التباين $\sigma^2=\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}$ بطيئة جداً، فيمكن أن نستبدل P_1 بـ$\hat{P}_1$ و P_2 بـ $\hat{P}_2$.
ومنه يكون مجال الثقة حول (P_1-P_2) بمستوى $(1-\alpha)$ من الثقة من الشكل:

$$\left[(\widehat{P_1}-\ \hat{P}_2)-Z_{1-\alpha/2}.\sqrt{\frac{\widehat{P_1}\widehat{q_1}}{n_1}+\frac{\widehat{P_2}\hat{q}_2}{n_2}}\ \ ;\ \ (\widehat{P_1}-\ \hat{P}_2)+Z_{1-\alpha/2}.\sqrt{\frac{\widehat{P_1}\widehat{q_1}}{n_1}+\frac{\widehat{P_2}\widehat{q_2}}{n_2}}\ \right]$$

مثال (5 - 4):

طُبِّقَتْ طريقتان A ، B لمعالجة مرض معين، وتم أخذ عينتين من المرضى، حيث تم تطبيق الطريقة A على العينة الأولى، وتطبيق الطريقة B على العينة الثانية. فإذا كان حجم العينة الأولى $n_1=42$ مريضاً، شفي منهم 18، وكان حجم العينة الثانية $n_2=38$ مريضاً، شفي منهم 15.

والمطلوب: أوجد 99% مجال ثقة حول الفرق الحقيقي بين نسبتي الذين تم شفاؤهم من المرضى، وماذا تستنتج؟

الحل:

بفرض P_1 نسبة الذين سيتم شفاؤهم من الذين خضعوا للطريقة A في العلاج.

بفرض P_2 نسبة الذين سيتم شفاؤهم من الذين خضعوا للطريقة B في العلاج.

عندئذٍ سيكون لدينا:

$$\hat{P}_1 - \hat{P}_2 = \frac{X_1}{n_1} - \frac{X_2}{n_2} = \frac{18}{42} - \frac{15}{38} = 0.034$$

ومن أجل $1 - \alpha = 0.99$ مستوى للثقة سيكون $Z_{1-\alpha/2} = 2.58$.

وحجم الخطأ الأعظمي المرتكب في تقدير الفرق بين النسبتين وبثقة 99% سيكون:

$$\varepsilon = \left| \pm Z_{1-\alpha/2} . \sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}} \right| = (2.58). \sqrt{\frac{(0.43)(0.57)}{42} + \frac{(0.39)(0.61)}{18}} = 0.28$$

وسيكون مجال الثقة حول $(P_1 - P_2)$ وبثقة 99% من الشكل:

$$\left(\hat{P}_1 - \hat{P}_2\right) - \varepsilon \leq (P_1 - P_2) \leq \left(\hat{P}_1 - \hat{P}_2\right) + e$$

$$0.034 - 0.28 \leq (P_1 - P_2) \leq 0.034 + 0.28$$

$$-0.246 \leq (P_1 - P_2) \leq 0.314$$

وبما أن طرفي المجال مختلفان بالإشارة هذا يعني أنه بثقة 99% سيكون $P_1 - P_2 = 0$ أي $P_1 = P_2$ ومنه لا فرق بين طريقتي العلاج A ، B من جهة نسبة الشفاء.

مثال (5 - 5):

أجري استفتاء لسكان المدينة وريفها المحيط بها لمعرفة رأيهم حول اقتراح إنشاء مركز صحي عند أطراف المدينة. فمن عينة من 5000 مواطن من المدينة أيد منهم 2400 لصالح المشروع. ومن عينة من 2000 مواطن من ريف المدينة أيد منهم 1200 لصالح المشروع. والمطلوب:

أوجد 90% مجال ثقة حول الفرق الحقيقي للنسبتين الدالتين على تأييد المشروع. وماذا تستنتج؟

الحل:

لدينا النموذج المدروس:

من المدينة	$n_1 = 5000$	$\mathrm{X} = 2400$	$\widehat{\mathrm{P}}_1 = \frac{\mathrm{X}}{\mathrm{n}_1} = 0.48$	P_1
من ريف المدينة	$n_2 = 2000$	$\mathrm{Y} = 1200$	$\widehat{\mathrm{P}}_2 = \frac{\mathrm{Y}}{\mathrm{n}_2} = 0.60$	P_2

وسيكون $1-\alpha = 0.90$ مجال ثقة حول $(\mathrm{P}_1 - \mathrm{P}_2)$ من الشكل:

$$\left(\widehat{\mathrm{P}}_1 - \widehat{\mathrm{P}}_2\right) - \mathrm{Z}_{1-\alpha/2} . \sqrt{\frac{\widehat{\mathrm{P}}_1 \hat{\mathrm{q}}_1}{\mathrm{n}_1} + \frac{\widehat{\mathrm{P}}_2 \hat{\mathrm{q}}_2}{\mathrm{n}_2}} \leq (\mathrm{P}_1 - \mathrm{P}_2) \leq$$

$$\left(\widehat{\mathrm{P}}_1 - \widehat{\mathrm{P}}_2\right) + \mathrm{Z}_{1-\alpha/2} . \sqrt{\frac{\widehat{\mathrm{P}}_1 \hat{\mathrm{q}}_1}{\mathrm{n}_1} + \frac{\widehat{\mathrm{P}}_2 \hat{\mathrm{q}}_2}{\mathrm{n}_2}}$$

$$\varepsilon = \left| \pm Z_{1-\alpha/2} . \sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\widehat{P_2} \widehat{q_2}}{n_2}} \right| =$$

$$(1.65) . \sqrt{\frac{(0.48)(0.52)}{5000} + \frac{(0.60)(0.40)}{2000}}$$

$$= 0.02$$

$$(0.48 - 0.60) - 0.02 \le (P_1 - P_2) \le$$

$$(0.48 - 0.60) + 0.02 \quad -0.14 \le (P_1 - P_2) \le -0.10$$

وبما أن طرفي المجال سالبان فهذا يعني أنه بثقة 90% أن $P_1 - P_2 < 0$ أي $P_1 < P_2$ ، أي إن سكان الريف يفضلون هذا المشروع أكثر من سكان المدينة.

4.1.5: التقدير المجالي للنسبة بين تباينين:

قد يكون أحياناً من المرغوب به مقارنة دقة جهاز قياس بدقة جهاز قياس آخر أو مقارنة كل من خطين لإنتاج دواء معين أو مقارنة دقة فعالية طريقتين بالعلاج لمرض معين وأمثلة عديدة في ذلك المجال فإن ذلك يمكن إجراؤه في دراسة المقارنة بين تبايني مجتمعين بوساطة النسبه بينهما.

فإذا كان $X \sim N(\mu_1, {\sigma_1}^2)$ وكان لدينا عينة عشوائية من الحجم n_1 من X تباينها ${S_1}^2$.

وكان $Y \sim N(\mu_2, {\sigma_2}^2)$ وكان لدينا عينة عشوائية من الحجم n_2 من Y تباينها ${S_2}^2$. وهذه العينة مستقلة عن العينة الأولى من X .

رأينا أن الإحصاء ${\varkappa_1}^2 = \frac{(n_1 - 1){S_1}^2}{{\sigma_1}^2}$ يتوزع وفق كاي_مربع بدرجة من الحرية: $\gamma_1 = n_1 - 1$

و رأينا أن الإحصاء ${\varkappa_2}^2 = \frac{(n_2 - 1){S_2}^2}{{\sigma_2}^2}$ يتوزع وفق كاي _ مربع بدرجة من الحرية: $\gamma_2 = n_2 - 1$

وحسب تعريف متغير فيشر ومن أجل مجتمعين مستقلين يكون الإحصاء:

$$F=\frac{\frac{(n_1-1)S_1{}^2}{\sigma_1{}^2}\Big/_{(n_1-1)}}{\frac{(n_2-1)S_2{}^2}{\sigma_2{}^2}\Big/_{(n_2-1)}}\sim f_{(\gamma_1=n_1-1,\gamma_2=n_2-1)}$$

يتوزع وفق فيشر بدرجتي حرية γ_2,γ_1 حيث $\gamma_1=n_1-1$ درجة حرية البسط، و $\gamma_2=n_2-1$ درجة حرية المقام ومنه : $F=\frac{S_1{}^2}{S_2{}^2}.\frac{\sigma_2{}^2}{\sigma_1{}^2}\sim f_{(\gamma_1,\gamma_2)}$

وحسب العلاقة الاحتمالية في توزيع فيشر:

$$\mathrm{P}\left[f_{\alpha/2}(\gamma_1,\gamma_2)\le F\le f_{1-\frac{\alpha}{2}}(\gamma_1,\gamma_2)\right]=1-\alpha$$

$$\mathrm{P}\left[f_{\alpha/2}(\gamma_1,\gamma_2)\le \frac{S_1{}^2}{S_2{}^2}.\frac{\sigma_2{}^2}{\sigma_1{}^2}\le f_{1-\frac{\alpha}{2}}(\gamma_1,\gamma_2)\right]=1-\alpha$$

$$\mathrm{P}\left[\frac{S_2{}^2}{S_1{}^2}.f_{\alpha/2}(\gamma_1,\gamma_2)\le \frac{\sigma_2{}^2}{\sigma_1{}^2}\le \frac{S_2{}^2}{S_1{}^2}.f_{1-\frac{\alpha}{2}}(\gamma_1,\gamma_2)\right]=1-\alpha$$

ومنه يكون $(1-\alpha)$ مجال ثقة حول $\frac{\sigma_2{}^2}{\sigma_1{}^2}$ من الشكل:

$$\left[\frac{S_2{}^2}{S_1{}^2}.f_{\alpha/2}(\gamma_1,\gamma_2)\quad ,\frac{S_2{}^2}{S_1{}^2}.f_{1-\frac{\alpha}{2}}(\gamma_1,\gamma_2)\right]$$

وهنا نميز ثلاث حالات:

1) إذا كان $\frac{\sigma_2{}^2}{\sigma_1{}^2}>1$ أي $\sigma_2{}^2>\sigma_1{}^2$.

2) إذا كان $\frac{\sigma_2{}^2}{\sigma_1{}^2}<1$ أي $\sigma_2{}^2<\sigma_1{}^2$.

3) إذا كان $\frac{\sigma_2{}^2}{\sigma_1{}^2}=1$ (أي $\frac{\sigma_2{}^2}{\sigma_1{}^2}$ يقع بين قيمتين أكبر من الواحد وأقل من الواحد).

عندئذٍ من الممكن أن تكون ${\sigma_2}^2 = {\sigma_1}^2$ بثقة $(1 - \alpha)$.

وبناء على نتيجة النسبة نتخذ القرار وتفسر الناتج حسب طبيعة الدراسة.

مثال (5 - 6):

في دراسة أجريت للمقارنة بين كمية المادة الفعالة في صنفين من الأدوية لعلاج مرض معين ، وهما من إنتاج شركتين مختلفتين ، ولكن بامتياز من نفس الشركة الأم. تبين بالنسبة للدواء الأول أن معدل المادة الفعالة هو 3.1 وبانحراف معياري 0.5 MG ومن أجل الدواء الثاني كان معدل المادة الفعالة هو 2.7 وبانحراف معياري 0.7 MG وذلك من أجل عينة من 10 حبّات من الدواء الأول وعينة من 8 حبّات من الدواء الثاني. وبفرض أن مجتمعي العينتين يتوزعان طبيعياً بتباين مختلف. أوجد 98% مجال ثقة للنسبة الحقيقية للتباين في مجتمعي العينتين، وماذا تستنتج من هذه الدراسة؟

الحل:

ليكن σ_1^2 تباين معدل المادة الفعالة في الدواء الأول و σ_2^2 تباين معدل المادة الفعالة في الدواء الثاني.

إن $1 - \alpha = 0.98$ مجال ثقة حول $\frac{\sigma_2^2}{\sigma_1^2}$ يكون من الشكل:

$$\left[\begin{array}{c}\frac{{S_2}^2}{{S_1}^2} . f_{\alpha/2}(\gamma_1 = n_1 - 1, \gamma_2 = n_2 - 1) \leq \frac{\sigma_2^2}{\sigma_1^2} \leq \\ \frac{{S_2}^2}{{S_1}^2} . f_{1-\frac{\alpha}{2}}(\gamma_1 = n_1 - 1, \gamma_2 = n_2 - 1)\end{array}\right]$$

$$1 - \alpha = 0.98 \Rightarrow \alpha = 0.02 ; \frac{\alpha}{2} = 0.01 ; 1 - \frac{\alpha}{2} = 0.99$$

$$\gamma_1 = n_1 - 1 = 10 - 1 = 9 \; ; \; \gamma_2 = n_2 - 1 = 8 - 1 = 7$$

ومن جدول فيشر نجد:

$$f_{1-\frac{\alpha}{2}}(\gamma_1,\gamma_2) = f_{0.99}(9,7) = 6.72$$

ومن خواص جدول فيشر :

$$f_{\alpha/2}(\gamma_1,\gamma_2) = \frac{1}{f_{1-\frac{\alpha}{2}}(\gamma_2\,,\gamma_1)} = \frac{1}{f_{0.99}(7,9)} = \frac{1}{5.61} = 0.1783$$

ومنه يصبح المجال من الشكل:

$$\left(\frac{0.49}{0.25}\right).(0.1783) \leq \frac{\sigma_2^2}{\sigma_1^2} \leq \left(\frac{0.49}{0.25}\right).(6.72)$$

$$(0.395) \leq \frac{\sigma_2^2}{\sigma_1^2} \leq (13.17)$$

يُلحظ أن الحد الأدنى أصغر من الواحد والحد الأعلى أكبر من الواحد ، ومِنْ ثَمّ بثقة 98% يمكن أن يكون $\frac{\sigma_2^2}{\sigma_1^2} = 1$ أي $\sigma_2^2 = \sigma_1^2$ أي لا فرق حقيقي بين تباينين المادة الفعالة في الدواءَينِ.