المحاضرة الثالثة عشرة

2.5: اختبار الفرضيات للمقارنة بين مجتمعين:

في هذه الفقرة سنقارن من خلال اختبار الفرضيات الوسيطية ما بين متوسطين وفي حالات مختلفة وما بين نسبتين وما بين تباينين تماماً كما رأينا في الفقرة السابقة ببناء مجالات ثقة حولها.

1.2.5:اختبـارات حـول الفـرق بـين متوسـطي مجتمعـين طبيعيـين تبايناهمـا معلومان:

لـيكن $X\sim N(\mu_1,{\sigma_1}^2)$ حيـث μ_1 مجهـول و σ_1^2 معلـوم ولـتكن لـدينا عينـة عشوائية من X من الحجم n_1 متوسطها $\overline{X}$. وليكن $Y\sim N(\mu_2,{\sigma_2}^2)$ حيث μ_2 مجهـول و σ_1^2 معلـوم ولـتكن لـدينا عينـة عشـوائية مـن Y مـن الحجـم n_2 متوسطها $\overline{Y}$. والعينتان مستقلتان.

فإن الإحصاء:

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

وهذا الإحصاء يستخدم لاختبار فرضيات من الشكل:

1) $H_0: \ \mu_1=\mu_2$ مقابل الفرضية البديلة $H_1: \ \mu_1\neq\mu_2$

عنـد مسـتوى المعنويـة α ، وهنـا الاختبـار مـن الطـرفين نحسـب قيمـة إحصـاء الاختبار تحت صحة H_0 أي :

$$Z_0 = \frac{(\overline{X}-\overline{Y})}{\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}}$$

ومن أجل α مستوى المعنوية والاختبار من الطرفين و التوزيع طبيعي معياري لإحصائية الاختبار . ستكون منطقة رفض H_0 من الشكل:

$$\left]-\infty, Z_{\frac{\alpha}{2}}\right[\ \vee \ \left]Z_{1-\frac{\alpha}{2}}, +\infty\right[$$

ومنطقة قبول H_0 من الشكل: $\left[Z_{\frac{\alpha}{2}}, Z_{1-\frac{\alpha}{2}}\right]$.

ثم نقارن Z_0 الناتجة مع مناطق الرفض والقبول لـ H_0 لاتخاذ القرار المناسب في رفض أو قبول لـ H_0 كما رأينا في الفصل السابق.

مثال (5 - 7):

من مجتمع الإناث في مدينة معينة، تبيّن أن الأوزان لهم تتوزع طبيعياً وفق $N(\mu_1, 36)$ ومن أجل عينة مسحوبة من الحجم $n_1 = 16$ تبيّن أن متوسط الوزن فيها $\overline{X} = 58\ K.G$ ومن مجتمع الإناث في مدينة أخرى من نفس البلد تبيّن أن الأوزان لهم تتوزع طبيعياً وفق $N(\mu_2, 64)$ ومن أجل عينة مسحوبة من الحجم $n_2 = 25$ تبيّن أن متوسط الوزن فيها $\overline{Y} = 55\ K.G$. فهل يمكن القول أن متوسطي الوزن في المجتمعين متساويان عند مستوى المعنوية $\alpha = 0.01$ ؟

الحل: لدينا من هذه الدراسة:

X إناث المدينة (1)	$n_1 = 16$	$\overline{X} = 58$	μ_1	${\sigma_1}^2 = 36$
Y إناث المدينة (2)	$n_2 = 25$	$\overline{Y} = 55$	μ_2	${\sigma_2}^2 = 64$

يراد اختبار: $H_0 : \mu_1 = \mu_2$ مقابل $H_1 : \mu_1 \neq \mu_2$

وعند مستوى المعنوية $\alpha = 0.01$ حيث يكون $\frac{\alpha}{2} = 0.005$ و $1 - \frac{\alpha}{2} = 0.995$ ومن جدول Z :

$$Z\alpha_{/2} = -2.58 \quad و \quad Z_{1-\frac{\alpha}{2}} = Z_{0.995} = 2.58$$

وتكون إحصائية الاختبار:

$$Z = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma^2{}_1}{n_1}+\frac{\sigma^2{}_2}{n_2}}} \sim N(0,1)$$

وقيمة إحصائية الاختبار تحت صحة H_0 تكون من الشكل:

$$Z_0 = \frac{(\overline{X}-\overline{Y})}{\sqrt{\frac{\sigma^2{}_1}{n_1}+\frac{\sigma^2{}_2}{n_2}}} = \frac{58-55}{\sqrt{\frac{36}{16}+\frac{64}{25}}} = 1.38$$

ومن أجل $\alpha = 0.01$ مستوي من الأهمية والاختبار من الطرفين والتوزيع طبيعي معياري لإحصائية الاختبار ستكون منطقة رفض H_0 من الشكل:

$$\left]-\infty, Z_{\frac{\alpha}{2}}\right[\ \vee \ \left]Z_{1-\frac{\alpha}{2}}, +\infty\right[\Rightarrow \ ]-\infty, -2.58[\ \vee \ ]2.58, +\infty[$$

ومنطقة قبول H_0 من الشكل : $[-2.58\ , 2.58\ ]$

وبمقارنة $Z_0 = 1.38$ مع مناطق الرفض والقبول لـ H_0 نجد أن Z_0 تنتمي لمنطقة القبول ومنه نقبل H_0 ونرفض H_1 أي أن متوسطي الوزن عند الإناث في المدينتين متساويين.

2) لاختبار صحة الفرضية $H_0: \mu_1 = \mu_2$ مقابل الفرضية البديلة $H_1: \mu_1 < \mu_2$ وعند مستوى المعنوية α، هنا الاختبار من الطرف الأيسر ستكون منطقة الرفض ل H_0 من الشكل: $]-\infty\ , Z_\alpha\ [$.

ومنطقة القبول ل H_0 من الشكل: $]-\infty, Z_\alpha[$.

3) لاختبـــار صـــحة الفرضـــية: $H_0 : \mu_1 = \mu_2$ مقابـــل الفرضـــية البديلـــة $H_1 : \mu_1 > \mu_2$ وعند مستوى المعنوية α، هنا الاختبار من الطرف الأيمن.

ستكون منطقة الرفض لـ H_0 من الشكل: $]Z_{1-\alpha}, +\infty[$

ومنطقة القبول لـ H_0 من الشكل: $]-\infty, Z_{1-\alpha}]$.

مثال (5 - 8):

إذا كانت الزيادة في الوزن خلال أسبوع لطفل يتبع نظام التغذية A تخضع للتوزيع الطبيعي بمتوسط μ_1 وتباين 150 G. وكانت الزيادة في الوزن خلال أسبوع لطفل يتبع نظام التغذية B تخضع للتوزيع الطبيعي بمتوسط μ_2 وتباين 275 G. وعندما طبقنا نظام التغذية A على عينة من 16 طفلاً كان متوسط الزيادة في وزنهم خلال أسبوع 450 G. وطبقنا نظام التغذية B على 25 طفلاً كان متوسط الزيادة في الوزن خلال أسبوع 435 G. فهل هذه النتائج تدل على أن نظام التغذية A يسبب زيادة أكبر في وزن الأطفال؟ وذلك عند مستوى المعنوية $\alpha = 0.005$.

الحل:

لدينا من هذه الدراسة:

النظام الغذائي A (1)	$n_1 = 16$	$\bar{X} = 450$	μ_1	$\sigma_1{}^2 = 150$
النظام الغذائي B (2)	$n_2 = 25$	$\bar{Y} = 435$	μ_2	$\sigma_2{}^2 = 275$

يراد اختبار: $H_0 : \mu_1 = \mu_2$ مقابل $H_1 : \mu_1 > \mu_2$

وعند مستوى المعنوية $\alpha = 0.005$ ويكون $1 - \alpha = 0.995$ والاختبار من الطرف الأيمن.

ومن جدول Z : $Z_{1-\alpha} = Z_{0.995} = 2.58$

وستكون قيمة إحصائية الاختبار تحت صحة H_0:

$$Z_0 = \frac{(\overline{X}-\overline{Y})}{\sqrt{\frac{\sigma^2{}_1}{n_1}+\frac{\sigma^2{}_2}{n_2}}} = \frac{450-435}{\sqrt{\frac{150}{16}+\frac{275}{25}}} = 3.323$$

وعند مستوى المعنوية $\alpha = 0.005$ والاختبار من الطرف الأيمن والتوزيع طبيعي معياري لإحصائية الاختبار ستكون منطقة الرفض لـ H_0 من الشكل:

$$]Z_{1-\alpha}\,, +\infty[\; =]2.58\,, +\infty[$$

ومنطقة القبول لـ H_0 من الشكل: $]-\infty,\ Z_{1-\alpha}] =]-\infty,\ 2.58]$.

وبمقارنة $Z_0 = 3.323$ مع مناطق الرفض والقبول لـ H_0 نجد أن Z_0 تنتمي إلى منطقة الرفض. ومنه نرفض H_0 ونقبل H_1 ، أي إن متوسط الزيادة في الوزن باتباع النظام A أكبر من الوزن باتباع النظام B.

ملاحظة (1): في الحالة التي لا يكون فيها للمجتمعين التوزيع الطبيعي وعندما يكون $n_1, n_2 \geq 30$ وبفضل مبرهنة النهاية المركزية يكون

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

ومِنْ ثَمّ يمكن إجراء الاختبارات حول $(\mu_1 - \mu_2)$ بالأسلوب السابق نفسه .

ملاحظة (2): إذا كان σ_1^2 ، σ_2^2 مجهولين وكان $n_1, n_2 \geq 30$ فيمكن استبدال σ_1^2 ، σ_2^2 بـ S_1^2 ، S_2^2 حيث S_1^2 تباين العينة من المجتمع الأول و ${S_2}^2$ تباين العينة من المجتمع الثاني. ثم نجري الاختبارات حول $(\mu_1 - \mu_2)$ بالأسلوب السابق نفسه وحيث يكون إحصاء الاختبار :

$$Z = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim N(0,1)$$

مثال (5 - 9):

في اختبار بمقرر التشريح المرضي، تقدم 75 طالباً و 50 طالبة فكان متوسط درجات الطلاب 82 بانحراف معياري 8 درجات، بينما كان متوسط درجات الطالبات 76 بانحراف معياري 6 درجات.

والمطلوب اختبار إذا كان الطلاب والطالبات يعملون بالمستوى نفسه في مقرر التشريح المرضي ، وذلك عند مستوى المعنوية $\alpha = 0.05$.

الحل:

يراد اختبار: $H_0 : \mu_1 = \mu_2$ مقابل $H_1 : \mu_1 \neq \mu_2$

وعند مستوى المعنوية $\alpha = 0.05$ والاختبار من الطرفين.

إن إحصائية الاختبار هي :

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

وقيمة إحصائية الاختبار تحت صحة H_0:

$$Z_0 = \frac{(\overline{X}-\overline{Y})}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = \frac{82-76}{\sqrt{\frac{64}{75}+\frac{36}{50}}} = 4.78$$

واعتماداً على مستوى المعنوية $\alpha = 0.05$ والاختبار من الطرفين والتوزيع طبيعي معياري لإحصائية الاختبار ستكون منطقة رفض H_0 من الشكل:

$$\left]-\infty, Z_{\frac{\alpha}{2}}\right[\ \vee \ \left]Z_{1-\frac{\alpha}{2}}, +\infty\right[\ \Rightarrow \ ]-\infty, -1.96[\ \vee \ ]1.96, +\infty[$$

ومنطقة قبول H_0 من الشكل :

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

وبمقارنة $Z_0 = 4.78$ مع مناطق رفض وقبول لـ H_0 نجد أن Z_0 تنتمي لمنطقة الرفض من جهة اليمين ، ومِنْ ثَمّ نرفض H_0 ونقبل H_1 ، أي إن هناك فرقاً حقيقياً في مستوى الأداء في مقرر التشريح المرضي ، وبكون الرفض من اليمين فهذا يعني أن $\mu_1 - \mu_2 > 0$ أي $\mu_1 > \mu_2$ ، ومِنْ ثَمّ مستوى درجات الطلاب أفضل من مستوى درجات الطالبات وبثقة 0.95 وخطأ 0.05.

2.2.5 اختبارات حول الفرق بين متوسطي مجتمعين طبيعيين تبايناهما مجهولان:

رأينا في مسألة كون $n_1, n_2 \geq 30$ تتحقق $\widehat{\sigma_1^2} = S_1^2$ و $\widehat{\sigma_2^2} = S_2^2$ ونتابع الاختبار كما رأينا في الفقرة السابقة.

لكن في حالة كون $n_1, n_2 < 30$ فإننا سندرس الحالة التي يكون فيها للمجتمعين التباين نفسه

$\sigma_1^2 = \sigma_2^2 = \sigma^2$. إذ رأينا في فقرة سابقة أن المتغير:

$$T = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_c \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{(\gamma=n_1+n_2-2)}$$

حيث $S_c^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ هو التباين المشترك

حيث تكون معطيات هذه الحالة:

X طبيعي	n_1	$\overline{X}$	S_1	μ_1	σ^2
Y طبيعي	n_2	$\overline{Y}$	S_2	μ_2	σ^2

1) لاختبار صحة الفرضية : $H_0 : \mu_1 = \mu_2$ مقابل $H_1 : \mu_1 \neq \mu_2$

وعند مستوى المعنوية α والاختبار من الطرفين.

نحسب قيمة إحصائية الاختبار تحت صحة H_0:

$$T_0 = \frac{(\overline{X}-\overline{Y})}{S_c \ . \ \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

ومن أجل α مستوى الأهمية والاختبار من الطرفين والتوزيع لستيودنت بـ

$\gamma = n_1 + n_2 - 2$ لإحصائية الاختبار.

ستكون منطقة رفض H_0 من الشكل: $\left]-\infty \, , t_{\alpha/2}\right[\text{V} \left]t_{1-\alpha/2} \, , +\infty\right[$

ومنطقة قبول H_0 من الشكل: $\left[t_{\alpha/2} \, , t_{1-\alpha/2}\right]$

(حيث يكون $t_{\alpha/2} = -t_{1-\alpha/2}$).

ثم نقارن T_0 الناتجة مع مناطق رفض وقبول لـ H_0 لاتخاذ القرار المناسب.

مثال (5 - 10):

ليكن X المتغير الدال على عدد الأيام اللازمة للشفاء من مرض معين باستخدام الطريقة A في العلاج حيث $X \sim N(\mu_1, \sigma_1^2)$. وليكن Y المتغير الدال على عدد الأيام اللازمة للشفاء من نفس المرض باستخدام الطريقة B في العلاج حيث

$Y \sim N(\mu_2\,, \sigma_2^2)$. ولمقارنة الطريقتين من حيث عدد الأيام اللازمة للشفاء تم أخذ عينة عشوائية من الحجم $n_1 = 15$ طبق عليها الطريقة A في العلاج ، فكان متوسط عدد الأيام اللازمة للشفاء 6.8 وتباينها $S_1^2 = 10.3$ وتم أخذ عينة عشوائية من الحجم $n_2 = 12$ من المرضى بنفس المرض كما في العينة الأولى وطبق عليها الطريقة B في العلاج فكان متوسط عدد الأيام اللازمة للشفاء 9.3 وتباينها $S_2^2 = 15.7$ وبفرض أن العينتين مستقلتان فهل هناك فرق حقيقي بين متوسطي عدد الأيام اللازمة للشفاء ما بين A و B والمجتمعان متجانسان $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$ عند مستوى الأهمية $\alpha = 0.01$.

الحل:

لدينا النموذج المدروس:

σ^2	μ_1	$S_1^2 = 10.3$	$\bar{X} = 6.8$	$n_1 = 15$	X والطريقة A
σ^2	μ_2	$S_2^2 = 15.7$	$\bar{Y} = 9.3$	$n_2 = 12$	Y والطريقة B

ويراد اختبار: $H_0 : \mu_1 = \mu_2$ مقابل الفرضية البديلة $H_1 : \mu_1 \neq \mu_2$

وعند مستوى المعنوية $\alpha = 0.01$ والاختبار من الطرفين.

$$\alpha = 0.01 \Rightarrow \frac{\alpha}{2} = 0.005\,; 1 - \frac{\alpha}{2} = 1 - 0.005 = 0.995$$

ودرجة الحرية: $\gamma = n_1 + n_2 - 2 = 15 + 12 - 2 = 25$ ويكون من جدول ستيودنت :

$$t_{1-\alpha/2}(\gamma) = t_{0.995}(25) = 2.78\,;\ t_{\alpha/2}(\gamma) = t_{0.005}(25)$$

$$= -t_{1-\alpha/2}(\gamma) = -2.787$$

وستكون قيمة إحصائية الاختبار تحت صحة H_0 :

$$T_0 = \frac{(\overline{X}-\overline{Y})}{S_c \ . \ \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$S_c = \sqrt{\frac{(n_1-1){S_1}^2+(n_2-1){S_2}^2}{n_1+n_2-2}} = \sqrt{\frac{(14)(10.3)+(11)(15.7)}{25}} = 3.56$$

$$T_0 = \frac{(6.3-9.3)}{3.56 \ . \ \sqrt{\frac{1}{15}+\frac{1}{12}}} = -1.813$$

ومن أجل مستوى الأهمية $\alpha = 0.01$ والاختبار من الطرفين والتوزيع لإحصائية الاختبار هو ستيودنت $\gamma = 25$ درجة من الحرية ستكون منطقة الرفض لـ H_0 من الشكل:

$$\left]-\infty, -t_{1-\alpha/_2}(\gamma)\right[=]-\infty, -2.787[\vee \left]t_{1-\alpha/_2}(\gamma)\ , +\infty\right[=]2.787, +\infty[$$

ومنطقة القبول لـ H_0 من الشكل :

$$\left[-t_{1-\alpha/_2}(\gamma), t_{1-\alpha/_2}(\gamma)\ \right] = [-2.787, 2.787]$$

وبمقارنة $T_0 = -1.813$ مع مناطق الرفض والقبول نجد أن T_0 تقع في منطقة القبول، ومِنْ ثَمّ تقبل H_0 ونرفض H_1 ، أي إن $\mu_1 = \mu_2$ (لا فرق حقيقياً بين طريقتي العلاج) .

1) لاختبار صحة الفرضية $H_0: \mu_1 = \mu_2$ مقابل الفرضية البديلة : H_1 $\mu_1 > \mu_2$

عند مستوى الأهمية α ، والاختبار هنا من اليمين، ودرجة الحرية $\gamma = n_1 + n_2 - 2$ ستكون منطقة رفض H_0 من الشكل: $]t_{1-\alpha}(\gamma), +\infty[$ ومنطقة القبول

لـ H_0 من الشكل: $]-\infty, t_{1-\alpha}(\gamma)\,[$ ، ومن ثم نقارن T_0 مع مناطق الرفض والقبول لـ H_0 لاتخاذ القرار المناسب.

2) لاختبار صحة الفرضية $H_0: \mu_1 = \mu_2$ مقابل الفرضية البديلة $H_1 : \mu_1 < \mu_2$ و عند مستوى الأهمية α ، و الاختبار هنا من اليسار ودرجة الحرية $\gamma = n_1 + n_2 - 2$ ستكون منطقة رفض H_0 من الشكل: $]-\infty, t_\alpha(\gamma)[$ أي $]-\infty, -t_{1-\alpha}(\gamma)[$ ومنطقة القبول لـ H_0 من الشكل: $]t_\alpha(\gamma), +\infty\,[$ ، ومن ثم نقارن T_0 مع مناطق الرفض والقبول لـ H_0 لاتخاذ القرار المناسب.

مثال (5 - 11):

شارك 12 شخصاً في برنامج لتخفيف الوزن ، والجدول الآتي يعطي مستوى الكولسترول عندهم قبل تطبيق البرنامج وبعده :

رقم الشخص	1	2	3	4	5	6	7	8	9	10	11	12
قبل التجربة X	201	221	231	261	230	237	240	233	270	248	201	200
بعد التجربة Y	200	215	233	234	226	215	195	295	245	220	210	208

وإذا علمنا مستوى الكولسترول لدى الأشخاص يخضع للتوزيع الطبيعي، فهل نستنتج من هذه البيانات أن هذا البرنامج فعّال في إنقاص مستوى الكولسترول عند مستوى المعنوية $\alpha = 0.01$.

الحل:

لدينا من فرضيات المسألة:

X قبل التجربة	$n_1 = 12$	$\overline{X} = 239.58$	$S_1 = 37.684$	μ_1	σ^2
Y بعد التجربة	$n_2 = 12$	$\overline{Y} = 224.83$	$S_2 = 26.536$	μ_2	σ^2

ويراد اختبار: $H_0 : \mu_1 = \mu_2$ مقابل الفرضية البديلة $H_1 : \mu_1 > \mu_2$

وعند مستوى المعنوية $\alpha = 0.01$ والاختبار هنا من الطرف الأيمن.

$$\alpha = 0.01 \Rightarrow 1 - \alpha = 0.99; \gamma = n_1 + n_2 - 2 = 12 + 12 - 2 = 22$$

ومن جدول توزيع ستيودنت:

$$t_{1-\alpha}(\gamma) = t_{0.99}(22) = 2.51$$

والانحراف المعياري المشترك:

$$S_c = \sqrt{\frac{(n_1-1){S_1}^2+(n_2-1){S_2}^2}{n_1+n_2-2}} = \sqrt{\frac{(11)(37.684)+(11)(26.536)}{22}} = 32.59$$

وقيمة إحصائية الاختبار تحت صحة H_0 :

$$T_0 = \frac{(\overline{X}-\overline{Y})}{\delta_c.\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{239.58-224.83}{(32.59)\sqrt{\frac{1}{12}+\frac{1}{12}}} = 1.11$$

ومن أجل مستوى المعنوية $\alpha = 0.01$ والاختبار من اليمين والتوزيع لإحصائية الاختبار هو ستيودنت بدرجة من الحرية $\gamma = 22$.

تكون منطقة رفض لـ H_0: $]t_{1-\alpha}(\gamma)\,, +\infty[=]2.51\,, +\infty[$

تكون منطقة قبول لـ H_0: $]-\infty, t_{1-\alpha}(\gamma)\,] =]-\infty, 2.51]$

وبمقارنة $T_0 = 1.11$ مع مناطق رفض أو قبول H_0 نجد أن : T_0 تنتمي إلى منطقة القبول ، ومنه تقبل T_0 أي $\mu_1 = \mu_2$ وهذا يعني أن البيانات السابقة لا تدل على أن هذا البرنامج فعال في إنقاص مستوى الكولسترول.

3.2.5 : اختبارات حول الفرق بين نسبتي مجتمعي $(\mathbf{P_1 - P_2})$:

هنا نختبر $H_0: P_1 = P_2$ (أي تساوي وسيطي مجتمعين لكل منهما توزيع برنولي) مقابل الفرضية البديلة H_1 حيث:

$$\mathrm{H_1 : p_1 > p_2} \quad \text{أو} \quad \mathrm{H_1 : p_1 < p_2} \quad \text{أو} \quad \mathrm{H_1 : p_1 \neq p_2}$$

ومن أجل $n_1, n_2 \geq 30$ سيكون:

$$\mathrm{Z} = \frac{(\widehat{\mathrm{p_1}} - \widehat{\mathrm{p_2}}) - (\mathrm{p_1} - \mathrm{p_2})}{\sqrt{\frac{\mathrm{p_1 q_1}}{\mathrm{n_1}} + \frac{\mathrm{p_2 q_2}}{\mathrm{n_2}}}} \approx \mathrm{N}(0\,,1)$$

وتحت صحةH_0 نستطيع في عبارة Z أن نضع $\mathrm{p_1 = p_2 = p}$ فتصبح على الشكل الآتي:

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

وإذا استبدلنا بـ p ، $\hat{\mathrm{p}}$ المقدّر المنصف لـp المعين بالعلاقة $\hat{\mathrm{p}} = \frac{X+Y}{n_1+n_2}$ حيث X، Y عدد النجاحات في العينة الأولى وعدد النجاحات في العينة الثانية على الترتيب فإننا سنحصل على الإحصاء الآتي:

(قيمة إحصاء الاختبار تحت صحة H_0) :
$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

1) لاختبار $\mathrm{H_0: p_1 = p_2}$ مقابل الفرضية البديلة $\mathrm{H_1 : p_1 \neq p_2}$ وعند مستوى المعنوية α والاختبار من الطرفين والتوزيع طبيعي معياري لإحصائية الاختبار . ستكون منطقة رفض لـ H_0 من الشكل:

: الشكل من H_0 قبول منطقة و $\left]-\infty,-Z_{1-\alpha/2}\right[\vee \left]Z_{1-\alpha/2},+\infty\right[$
$\left[-Z_{1-\alpha/2}, Z_{1-\alpha/2}\right]$

ثم نقارن Z_0 مع مناطق الرفض والقبول لاتخاذ القرار المناسب في قبول أو رفض H_0.

2) لاختبار $H_0: p_1 = p_2$ مقابل الفرضية البديلة $H_1 : p_1 < p_2$ وعند مستوى المعنوية α والاختبار من الطرف الأيسر.

ستكون منطقة رفض لـ H_0 من الشكل : $\left]-\infty, Z_\alpha\right[= \left]-\infty,-Z_{1-\alpha}\right[$

و منطقة قبول H_0 من الشكل : $\left[Z_\alpha,+\infty\right[= \left[-Z_{1-\alpha},+\infty\right[$

3) لاختبار $H_0: p_1 = p_2$ مقابل الفرضية البديلة $H_1 : p_1 > p_2$ وعند مستوى المعنوية α والاختبار من الطرف الأيسر. ستكون منطقة رفض لـ H_0 من الشكل: $\left]Z_{1-\alpha},+\infty\right[$ و منطقة قبول H_0 من الشكل : $\left]-\infty, Z_{1-\alpha}\right]$.

مثال (5 - 12):

تبين من سجلات مشفى أن من بين 1000 رجل دخلوا المشفى كان من بينهم 46 رجلاً يعانون مرض القلب ، ومن بين 600 امرأة دخلت المشفى كان من بينهن 18 امرأة تعاني مرض القلب، هل تقدم هذه المعلومات الإحصائية دلاله كافيه على أن نسبة الإصابة بمرض القلب عند الرجال تساوي نسبة الإصابة بمرض القلب عند النساء بمستوى $\alpha = 0.05$ من المعنوية ؟

الحل:

لدينا من معطيات المسألة:

مجتمع الرجال	$n_1 = 1000$	$X = 46$	$\widehat{p_1} = 0.046 = \frac{46}{1000}$	p_1
مجتمع النساء	$n_2 = 600$	$Y = 18$	$\hat{p}_2 = 0.03 = \frac{18}{1000}$	p_2

وسنختبر $H_0: p_1 = p_2$ مقابل الفرضية البديلة $H_1 : p_1 \neq p_2$

وعند مستوى المعنوية $\alpha = 0.05$ والاختبار هنا من الطرفين.

$$\alpha = 0.05 \Rightarrow 1 - \alpha = 0.95 \;; 1 - \frac{\alpha}{2} = 0.975 \;; \quad Z_{0.975} = 1.96$$

وقيمة إحصائية الاختبار تحت صحة H_0 تكون:

$$Z_0 = \frac{\widehat{P}_1 - \widehat{P}_2}{\sqrt{\widehat{P}(1-\widehat{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \;; \quad \hat{p} = \frac{X+Y}{n_1+n_2} = \frac{46+18}{1000+600} = 0.04$$

$$Z_0 = \frac{0.046-0.03}{\sqrt{(0.04)(0.96)\left(\frac{1}{1000}+\frac{1}{600}\right)}} = 1.58$$

وعند مستوى المعنوية $\alpha = 0.05$ والاختبار من الطرفين والتوزيع طبيعي معياري لإحصائية الاختبار. ستكون منطقة رفض H_0 من الشكل :

$$\left]-\infty, -Z_{1-\alpha/2}\right[\vee \left]Z_{1-\alpha/2}, +\infty\right[\Rightarrow \;]-\infty, -1.96[\vee]1.96, +\infty[$$

و منطقة قبول H_0 من الشكل : $\left[-Z_{1-\alpha/2}, Z_{1-\alpha/2}\right] = [-1.96\,,1.96]$

ثم نقارن $Z_0 = 1.58$ مع مناطق الرفض والقبول ، فنجد أن Z_0 تنتمي لمنطقة القبول، ومنه تقبل H_0 ونرفض H_1 ، أي $p_1 = p_2$ أي تساوي نسبة الإصابة بمرض القلب عند الرجال والنساء.

مثال (5 - 13):

أعطي نوعان من الأدوية بهدف تخفيف الألم الحادث بعد العمليات الجراحية. فمن أصل 100 مريض أعطي لهم الدواء A ادعى 38 منهم أنه خفف الألم ، بينما من أصل 120 مريض أعطي لهم الدواء B ادعى 56 منهم أنه خفف الألم، والمطلوب: هل ثمة دلالة إحصائية على وجود فرق معنوي بين الدواءَينِ بفعالية تخفيف الألم بعد العمليات الجراحية عند مستوى المعنوية $\alpha = 0.05$.

الحل:

إن فرضية العدم هي $\mathrm{H_0: p_1 = p_2}$

والفرضية البديلة هي $\mathrm{H_1 : p_1 \neq p_2}$

وعند مستوى المعنوية $\alpha = 0.05$ والاختبار من الطرفين

إن إحصائية الاختبار

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim N(0\,,1)$$

$$\mathrm{\hat{p}_1 = \frac{X}{n_1} = \frac{38}{100} = 0.38 \quad ; \quad \hat{p}_2 = \frac{Y}{n_2} = \frac{56}{120} = 0.467}$$

$$\mathrm{\hat{p} = \frac{X+Y}{n_1+n_2} = \frac{38+56}{100+120} = \frac{94}{220} = 0.427}$$

وتحت صحة H_0 تكون قيمة إحصائية الاختبار

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.38-0.467}{\sqrt{(0.427)(1-0.427)\left(\frac{1}{100}+\frac{1}{120}\right)}} = -1.30$$

وعند مستوى المعنوية $\alpha = 0.05$ والاختبار من الطرفين والتوزيع طبيعي معياري لإحصائية الاختبار.

ستكون منطقة قبول H_0 من الشكل :

$$\left[Z\alpha_{/2}, Z_{1-\alpha/2}\right] = [Z_{0.025}, Z_{0.975}] = [-1.96, 1.96]$$

وبالمقارنة مع $Z_0 = -1.30$ فنجد أنها تنتمي لمنطقة القبول. ومنه تقبل H_0 ونرفض H_1 أي لا فرق بين فعالية الدواءَيْنِ لتخفيف الألم بعد العمليات الجراحية.