التقدير الإحصائي

د. مصطفى بابكر

التقدير الاحصائي واختبار الفرضيات

التقدير الاحصائي

- هنالك نوعان من التقدير الإحصائي هما التقدير باستخدام القيم الممثلة للعينة (Point Estimation) والتقدير باستخدام حدود أو فترات الثقة (Confidence Interval Estimation). ولقد تناولنا بعض خصائص النوع الأول في استعراضنا لتوزيع المعاينة في الجزء السابق وعليه سيختص هذا الجزء بالنوع الثاني من التقدير الإحصائي.

- لأي مقدر $\hat{\theta}$ لمعلمة المجتمع θ هنالك طريقة للتعبير عن دقة المقدر إذا علم توزيع المعاينة للمقدر وذلك ببناء ما يعرف بفترات الثقة (Confidence Interval) حول المقدر. حيث نستخدم قيمة المقدر $\hat{\theta}$ لتحديد قيمتين $\hat{\theta}_1$ و $\hat{\theta}_2$ بحيث تحتوي الفترة $(\hat{\theta}_1, \hat{\theta}_2)$ قيمة المعلمة θ باحتمال معين. ويسمى الاحتمال معبرا عنه بالنسبة المئوية معامل الثقة أو مستوى الثقة للقيمة التقديرية للمقدر.

- إذا افترضنا أن للمقدر $\hat{\theta}$ توزيع طبيعي يمكن استخدام خصائص التوزيع الطبيعي المعياري لحساب احتمال وقوع θ في الفترة $(\hat{\theta}_1,\hat{\theta}_2)$ كالآتي :

$$P\left(\hat{\theta}_1 < \theta < \hat{\theta}_2\right) = P\left(\frac{\hat{\theta}_1 - \theta}{Se(\hat{\theta})} < Z < \frac{\hat{\theta}_2 - \theta}{Se(\hat{\theta})}\right) = P\left(Z_1 < Z < Z_2\right)$$

حيث $Se(\hat{\theta})$ هو الخطأ المعياري للمقدر. $\hat{\theta}$

- لأي عدد α ينتمي للفترة (0,1 :

$$P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right) = 1-\alpha$$

حيث $P\left(Z > Z_{\alpha/2}\right) = \frac{\alpha}{2}$ و $P\left(Z < -Z_{\alpha/2}\right) = \frac{\alpha}{2}$.

وتسمى α بمستوى المغزوية (Level of Significance) في حالة التعبير بالنسب المئويـة وتسمى $(1-\alpha)$ بمستوى الثقـة (Confidence Level) في حالة التعبير المئوي وتعرف بمعامل الثقة (Confidence Coefficient) في حالة التعبير الكسري. ويستخدم الإحصائيون التطبيقيون في بناء فترات الثقة واختبار الفرضيات مستوى المغزوية 5% و1%. وبالرجوع لجداول التوزيع الطبيعي المعياري نجد أن قيمة $Z_{\alpha/2}$ المقابلة لـ 5% هي 1.96 وقيمة $Z_{\alpha/2}$ المقابلة لـ 1% هي 2.58. عليه:

$$P(-1.96 < Z < 1.96) = 0.95$$

$$P(-2.58 < Z < 2.58) = 0.99$$

- بتعميم ما سبق فإن القانون العام لحساب فترات الثقة يتم بتعويض قيمة مقدرات العينة كالآتي :

$$P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2} < \frac{\hat{\theta}-\theta}{Se\,(\hat{\theta})} < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2}\,Se\,(\hat{\theta}) < \hat{\theta}-\theta < Z_{\alpha/2}\,Se\,(\hat{\theta}\right) = 1-\alpha$$

$$P\left(\hat{\theta}-Z_{\alpha/2}\,Se\,(\hat{\theta}) < \theta < \hat{\theta}+Z_{\alpha/2}\,Se\,(\hat{\theta}\right) = 1-\alpha$$

- أي أن فترة الثقة للمعلمة θ والمقدر $\hat{\theta}$ ذو التوزيع الطبيعي هي:

$$\left(\hat{\theta} - Z_{\alpha/2} Se(\hat{\theta}), \hat{\theta} + Z_{\alpha/2} Se(\hat{\theta})\right)$$

مثلا فترة الثقة 95% للمعلمة θ والمقدر $\hat{\theta}$ هي:

$$\left(\hat{\theta} - 1.96 Se(\hat{\theta}), \hat{\theta} + 1.96 Se(\hat{\theta})\right)$$

- بالتطبيق على الوسط الحسابي للعينة $(\bar{x})$ في حالة معرفة تباين المجتمع (σ^2) تكون فترة الثقة $(1-\alpha)$:

$$\left[\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

مثلا فترة الثقة 95% للوسط الحسابي للمجتمع هي:

$$\left[\bar{x}-1.96\frac{\sigma}{\sqrt{n}},\bar{x}+1.96\frac{\sigma}{\sqrt{n}}\right]$$

مثال لذلك إذا كان $\bar{x} = 400$، $\sigma^2 = 25$ ، $n = 100$ فإن فترة الثقة 95% هي (399.02, 400.98) أما في حالة عدم معرفة σ وصغر حجم العينة يستعاض بالتوزيع تي والانحراف المعياري للعينة s وتكون فترة الثقة $(1-\alpha)$ هي:

$$\left[\bar{x} - t_{\alpha/2,\nu}\frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2,\nu}\frac{s}{\sqrt{n}}\right]$$

حيث ν هي درجات الحرية وتحسب $t_{\alpha/2,\nu}$ من جداول توزيع تي حيث

$$P\left(t > t_{\alpha/2,\nu}\right) = \frac{\alpha}{2}$$

مثال:

ضمن خطتها لإصلاح حركة المرور في المدينة قامت بلدية كازاكا بإجراء مسح ميداني لتحديد حجم الحركة عبر تقاطع رئيسي في المدينة خلال فترات الصباح. باختيار يوم السبت لثمانية أسابيع متوالية تم عِدِ المركبات التي تمر عبر التقاطع بين الساعة 7:00 والساعة 9:00 صباحا ووجد أن متوسط عدد المركبات للعينة يساوي 1500 والانحراف المعياري للعينة يساوي 300. المطلوب حساب فترة الثقة 99% لمتوسط عدد المركبات في مجتمع الحركة عبر التقاطع باعتبار أن توزيع المجتمع طبيعي .

الحل:

نسبة لصغر حجم العينة وعدم معرفة الانحراف المعياري للمجتمع نستخدم التوزيع تي لبناء فترة الثقة. من جداول تي نجد أن 3.499 = $t_{0.005,7}$ = $t_{\alpha/2,\nu}$ وعلي فإن فترة الثقة 99% لمتوسط الحركة في المجتمع هي:

$$\left[1500-3.499\times\frac{300}{\sqrt{7}},1500+3.499\times\frac{300}{\sqrt{7}}\right]$$

أو

$$[1103.25, 1896.75]$$

- بالتطبيق على النسبة في العينة $\hat{P}$ يمكن حساب فترة الثقة $(1-\alpha)$ باستخدام التوزيع الطبيعي لتقريب توزيع النسب كالآتي:

$$\left[\hat{P}-Z_{\alpha/2}\sqrt{\hat{P}\hat{q}/n},\hat{P}+Z_{\alpha/2}\sqrt{\hat{P}\hat{q}/n}\right]$$

مثال:

إذا كان عدد الأسر ذات الدخل المحدود لعينة من 500 أسرة هو 200 يمكن حساب فترة الثقة 95% لنسبة الأسر ذات الدخل المحدود في المجتمع كالآتي :

$$\left[0.4-1.96\sqrt{0.4\times0.6/500},0.4+1.96\sqrt{0.4\times0.6/500}\right]$$

أو

$$[0.357, 0.443]$$

- يمكن تحديد حجم العينة المطلوب للحصول على دقة معينة في تقدير معلمة المجتمع باستخدام فترات الثقة. مثلاً إذا حددنا مسبقاً مقدار خطأ العينة $\left|\hat{\theta}-\theta\right|$ بالا يتجاوز D نحصل من قانون فترة الثقة للتوزيع الطبيعي على التعبير :

$$D = Z_{\alpha/2} Se(\hat{\theta})$$

وبالتطبيق على الوسط الحسابي للعينة :

$$D = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

أو

$$\sqrt{n} = \frac{Z_{\alpha/2}\sigma}{D}$$

أو

$$n = \frac{Z^2_{\alpha/2}\sigma^2}{D^2}$$

مثلاً لـ $\alpha = 0.05$ ، $\sigma = 4000$ و $D = 500$ يكون حجم العينة المطلوب للحصول على هذه الدقة:

$$n = \frac{(1.96)^2 (4000)^2}{(500)^2} = 245.86$$

أي أننا نحتاج لعينة تحتوي على 246 مشاهدة.

- أيضاً يمكن بناء فترات الثقة للفرق في الوسط الحسابي أو النسبة بين عينتين لتقدير الفرق بين المجتمعين في حالة التوزيع الطبيعي كالآتي:

بالنسبة للفرق $(\bar{x}_1 - \bar{x}_2)$ لتقدير $(\mu_1 - \mu_2)$:

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

وللفرق في النسبة $(\hat{P}_1 - \hat{P}_2)$ لتقدير $(P_1 - P_2)$:

$$\left[(\hat{P}_1-\hat{P}_2)-Z_{\alpha/2}\sqrt{\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}},(\hat{P}_1-\hat{P})+Z_{\alpha/2}\sqrt{\frac{P_1q_1}{n_1}+\frac{P_2q_2}{n_2}}\right]$$

مثال:

لتقدير التفرقة في الأجور بين الإناث والذكور وجد لعينتين من 80 ذكر و 60 أنثى أن متوسطي الدخول يساويان 20,600 بانحراف معياري 3000 للذكور و 19,700 بانحراف معياري 2500 للإناث. المطلوب حساب فترة الثقة 95% للفرق في الدخول بافتراض التوزيع الطبيعي .

الحل:

نسبة لكبر حجم العينتين يمكن استخدام التوزيع الطبيعي بالرغم من عدم معرفتنا للانحراف المعياري للمجتمع. وبتطبيق التوزيع الطبيعي فإن فترة الثقة 95% تساوي :

$$\left[900-1.96\sqrt{\frac{3000^2}{80}+\frac{2500^2}{60}},900+1.96\sqrt{\frac{3000^2}{80}+\frac{2500^2}{60}}\right]$$

أو

$$[134.6, 1665.4]$$

وإذا استبدلنا في المثال الأوساط الحسابية بالنسب في العينة مثلاً أن نسبة الذكور ذوي الدخول أكثر من 10,000 تساوي 40% وللإناث تساوي 30% يمكن حساب فترة الثقة 95% للفرق في النسب:

$$\left[0.1-1.96\sqrt{\frac{0.4\times0.6}{80}+\frac{0.3\times0.7}{60}},0.1+1.96\sqrt{\frac{0.4\times0.6}{80}+\frac{0.3\times0.7}{60}}\right]$$

اختبار الفرضيات(Test of Hypotheses)

- بالمقارنة مع فترات الثقة يستخدم اختيار الفرضيات عندما تكون هنالك بعض المعلومات متوفرة عن المجتمع ولكن يود التأكد من دقتها . فقد تكون لدى الدارس مثلاً معرفة بمتوسط المجتمع من دراسات سابقة ولكن بعض التغييرات التي حدثت للمجتمع في وقت لاحق قد غيرت من قيمة المتوسط ونسبة لهذا الشك فقد يكون الطريق الوحيد للتأكد من حدوث التغير أو عدمه هو اختبار قيمة المتوسط في المرحلة الجديدة لمعرفة إن كان هو نفس المتوسط أم أنه قد تغير .

- لتوضيح ما يقصد باختبار الفرضيات وعملية اتخاذ القرار خذ المثال التالي: "تود إدارة المناهج بوزارة التربية إدخال طريقة جديدة لتدريس الرياضيات الحديثة. ولكن قبل اتخاذ هذا القرار تود الإدارة أن تتأكد من أن مستوى الطلاب يتحسن عند التدريس بالطريقة الجديدة. ولقياس هذا التحسن أو عدمه ستحتكم إلى الدرجات التي يحصل الطلاب عليها في الامتحان. ولنفترض أن متوسط الدرجات للرياضيات باستعمال الطريقة القديمة كان $\mu = 80$ وعليه فإن الطريقة الجديدة سيحكم عليها بالنجاح إذا اتضح أن متوسط الدرجات أعلى من 80 ولن يحكم عليها بالنجاح إذا كان المتوسط أقل من ذلك".

- العوامل التي تؤثر على اتخاذ القرار :

* بما أن إدارة المناهج غير مستعدة لقبول الطريقة الحديثة وتعميمها إلا إذا أثبتت نجاحها، فليس أمامها سوى أن تجرِبها على عدد محدود من الطلاب ثم بعد ذلك الحكم عليها . إذا فإن اللجنة ستحتكم إلى عينة من المجتمع للوصول إلى نتيجة محدودة حول متوسط العينة للوصول إلى قرار حول متوسط المجتمع فهي تقبل الطريقة الجديدة إذا كان هنالك ما يشير إلى أن المتوسط أكبر من 80 وترفضها إذا لم يكن هنالك ما يشير لذلك .

* هنالك احتمال باتخاذ القرار الصحيح وهو قبول الطريقة عندما يكون فعلاً المتوسط أكبر من 80 أو رفضها عندما يكون المتوسط 80 أو أقل فعلاً. ولكن في نفس الوقت هنالك احتمال باتخاذ القرار الخاطئ وهو في هذه الحالة قبول الطريقة عندما تكون حقيقة الأمر أن المتوسط 80 أو أقل أو رفض الطريقة عندما يكون المتوسط أكبر من 80 حقيقة. والمتسبب في القرار الخاطئ هو أن نتائج العينة قد تبعد عن طريق الصدفة عن قيم المجتمع .

من هنا يتضح أن الوصول إلى قرار معين بناء على نتيجة العينة يكتنفه عنصر المخاطرة وعلينا أن نحدد ما مقدار المجازفة التي سنقبل بها وهذا المقدار تحدده مدى خطورة أو عدم خطورة اتخاذ القرار الخاطئ. فكلما زادت الخطورة وجب علينا تقليل عنصر المخاطرة أما إذا كان مثل هذا القرار ليس خطير فيمكننا التساهل في مقدار المخاطرة .

- قاعدة اتخاذ القرار:

1. بالرجوع إلى المثال أعلاه. ذكرنا أنه عند احتكامنا إلى نتيجة العينةٍ سنصل إلِى قرار نهائي حول الطريقة الحديثة وسيكون القرار إما قبولا أو رفضا وتحديد القاعدة التي يتم بمقتضاها القبول أو الرفض سيتفاوت بين افتراضين :

* الافتراض الأول: هو القولِ بأن الطريقة الحديثة ليست أفضل من الطريقة القديمة (يعرف أيضا بافتراض عدم الاختلاف .(
* الافتراض الثاني: هو القول بأن الطريقة الحديثة أفضل من الطريقة القديمة .

ويسمى الافتراض الأول فرضية العدم (Null Hypothesis) ويرمز لـه بـ H_0 . أما الافتراض الثانـي فيسمى فرضية البديـل (Alternative Hypothesis) ويرمز له بـ H_1 . وعند تحديد هاتين الفرضيتين لابد من تحديد مستوى المخاطر التي يمكن القبول بها عند اتخاذ القرار الخاطئ وهو مستوى المغزوية (أو المعنوية) α الذي أشرنا إليه في الحديث عن فترات الثقة. ويقل مستوى المغزوية كلما كنا حريصين على عدم قبول الفرضية الخاطئة.

2. إن مستوى المغزوية هو في حقيقة الأمر درجة الاحتمال الذي نرفض به فرضية العدم H_0 عندما تكون هي الصحيحة وهو ما يعرف باحتمال الوقوع في الخطأ من النوع الأول وذلك لأن طريقة اتخاذ القرار الخاص بالفرضيات قد يقودنا إلى الوقوع في نوعين من الخطأ:

* الخطأ من النوع الأول (Type I Error): ويحدث عند رفض فرضية العدم عندما تكون هي صحيحة .
* الخطأ من النوع الثاني (Type II Error): ويحدث عند قبول فرضية العدم عندما تكون الفرضية البديلة هي الصحيحة. ويمكن توضيح ذلك في الجدول التالي :

القرار / الحقيقة	قبول	رفض
صحيحة	قرار صحيح	الخطأ من النوع الأول
غير صحيحة	الخطأ من النوع الثاني	قرار صحيح

وعليه فإن مستوى المغزوية α هو درجة احتمال الوقوع في الخطأ من النوع الأول ويسمى احتمال رفض H_0 عندما تكون فعلاً خاطئة بقوة الاختبار (Power of the Test). وقد يتبادر إلى الذهن السؤال لماذا لا نقلل الخطأين معاً لأقل حد ممكن بحيث نحصل على قوة اختبار عالية والإجابة هي أن هنالك علاقة عكسية بين حجم الخطأ في النوع الأول وحجم الخطأ في النوع الثاني والحل الوحيد لتقليل كليهما هو زيادة حجم العينة.

3. إن القاعدة التي تستعمل للحكم على نتائج العينة سواء بقبول فرضية العدم أو البديل عند مستوى معزوية معين يعتمد على التوزيع الاحتمالي للمتغير قيد الدراسة .

- اختبار الفرضيات للوسط الحسابي μ :

  * عرفنا من توزيع المعاينة أن الوسط الحسابي للعينة $(\bar{x})$ له توزيع طبيعي إذا كانت العينة ماخوذة من توزيع طبيعي أو عندما يكون حجم العينة كبيرا، هذا إذا كان الانحراف المعياري للمجتمع (σ) معروفا أما إذا كنا سنستخدم S لتقدير σ وحجم العينة متغير فإن التوزيع الاحتمالي في هذه الحالة هو تي.

* الطريقة المتبعة لاختبار الفرضيات هو تحديد فرضية العدم H_0 والبديل H_1 ومستوى المغزوية α ثم حساب قيمة Σ إذا كان التوزيع طبيعي أو t إذا كان التوزيع تي. ونحسب قيمتا Z و t بناءً على فرضية H_0 أي أننا نعوض لقيمة المعلمة في حسابهما بقيمة المعلمة في ظل صحة الفرضية H_0. ثم نقارن قيمتا Z و t بالقيم المحسوبة من الجداول أي $-Z_{\alpha/2}$ و $Z_{\alpha/2}$ في حالة Z و $-t_{\alpha/2,v}$ و $t_{\alpha/2,v}$ في حالة t. فإذا وقعت القيمة في الفترة $[-Z_{\alpha/2}, Z_{\alpha/2}]$ أو $[-t_{\alpha/2,v}, t_{\alpha/2,v}]$ عنى ذلك أنه ليس هنالك ما يشير إلى رفض فرضية العدم أما إذا وقعت قيمتها خارج الفترتين فمعنى ذلك أن هناك ما يشير إلى رفض فرضية العدم ويكون الميل إلى قبول فرضية البديل.

ويسمى مثل هذا الاختبار اختبار ذو طرفين (ذيلين) "Two Sided (Tailed) Test" وذلك لأننا لم نحدد اتجاها محددا لفرضية البديل وتصاغ هاتين الفرضيتين كما يلي :

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

حيث μ_0 قيمة محددة يراد اختبارها وتسمى القيم $[-z_{\alpha/2}, z_{\alpha/2}]$ أو $[-t_{\alpha/2,v}, t_{\alpha/2,v}]$ بالقيم الحرجة (Critical Values) حيث يتم قبول أو رفض فرضية العدم بناءً على موقع z أو t بالنسبة لهذه القيم .

* هنالك بعض الحالات حيث يمكننا تحديد اتجاه فرضية البديل فقد تكون فرضية البديل $\mu > \mu_0$ أو $\mu < \mu_0$ وفي هاتين الحالتين يسمى الاختبار اختبار ذو طرف (ذيل) واحد. وتبعا لذلك يختلف حساب القيم الحرجة فمثلاً لمستوى المغزوية 5% فإن قيمة $Z_{\alpha/2}$ تساوي 1.96 في حالة الاختبار ذو الطرفين في حين أن قيمة Z_α تساوي 1.65 في حالة الاختبار ذو الطرف الواحد. وفي حالة مستوى المغزوية 1% (أي $\alpha = 0.01$) تكون قيمة $Z_{\alpha/2}$ تساوي 2.58 وقيمة Z_α تساوي 2.33. أما قيم $t_{\alpha/2,\nu}$ و $t_{\alpha,\nu}$ فتعتمد بالإضافة إلى α على درجات الحرية ν ويمكن إيجادها لأي زوج (α, ν) بالنظر إلى جداول التوزيع تي.

مثال: (اختبار ذو طرفين)

أجريت دراسة في مجتمع معين لتحديد ما إذا كان متوسط مصروف الأسرة في الأسبوع يختلف عن متوسط المنصرف القومي الذي يساوي 52 دينار. أخذت عينة عشوائية من 100 أسرة من هذا المجتمع ووجد متوسطها وانحرافها المعياري للمنصرف الأسبوعي يساوي 56 و 10.28 على التوالي. هل هذه البيانات تعطي أدلة كافية بأن متوسط المنصرف الأسبوعي لهذه الأسرة يختلف عن المنصرف القومي (استعمل $\alpha = 0.05$).

الحل:

(أ) الفرضيات:

$$H_0 : \mu = 52$$

$$H_1 : \mu \neq 52$$

(ب) التوزيع الاحتمالي والقيم الحرجة:

بالرغم من عدم معرفة σ إلا أن حجم العينة 100 يعتبر كافياً لاستخدام التوزيع الطبيعي. لمستوى المغزوية 5% وحيث H_1 لها الصيغة $\neq$ فإن $[-Z_{\alpha/2}, Z_{\alpha/2}]$ تساوي $(-1.96, 1.96)$.

الحل:

(ج) استخدام بيانات العينة لحساب Z في ظل الفرضية H_0:

$$Z = \frac{\bar{x} - \mu}{Se(\bar{x})} = \frac{56 - 52}{10.28/\sqrt{100}} > Z_{\alpha/2} = 1.96$$

(د) القرار:

بما أن قيمة Z المحسوبة من العينة أكبر من القيمة الحرجة نرفض فرضية العدم أي أن هنالك ما يشير إلى أن الوسط الحسابي لهذا المجتمع يختلف عن الوسط الحسابي للمنصرف القومي.

مثال: (اختبار ذو طرف واحد(

متوسط قوة مقاومة حبال للقطع من إنتاج أحد المصانع هو 1.8 طن. باستخدام طريقة جديدة للتصنيع ادعي أحد المنتجين أن قوة مقاومة الحبال ستزداد. لاختبار هذا الادعاء أخذت عينة من 25 حبلاً ووجد أن متوسط مقاومتها للقطع هو 1.85 طن وانحرافها المعياري 0.10 طن. هل يمكن تأييد هذا الادعاء عند مستوى المغزوية 0.01؟

الحل:

(أ) الفرضيات:

$$H_0 : \mu = 1.80$$

$$H_1 : \mu > 1.80$$

(ب) نموذج الاختبار:

نسبة لعدم معرفة σ ولصغر حجم العينة يستخدم التوزيع الاحتمالي تي. وحيث أن H_1 على الصيغة $>$ تكون القيمة الحرجة $t_{0.01,24}$ والتي تساوي بالرجوع لجدول تي 2.485 .

(ج) حساب t من العينة:

$$t = \frac{1.85 - 1.8}{0.1/5} = 2.5 > 2.485$$

(د) القرار:

ترفض الفرضية H_0 ونميل إلى قبول الادعاء بأن الطريقة الجديدة تزيد من مقاومة الحبال للقطع.

- اختبار فرضيات للنسبة P

* يتبع نفس الأسلوب في اختبار فرضيات الوسط الحسابي باستبدال النسبة مكان الوسط الحسابي .

مثال:

كانت نسبة النجاح في الشهادة الثانوية العامة لإحدى الدول 72% ثم أحدثت بعض التغييرات في المناهج مع إضافة وحذف بعض المواد. وكانت النتيجة بعد التعديلات أن نجح ثمانية آلاف طالب من مجموع عشرة آلاف طالب جلسوا لذلك الامتحان. هل تحسنت نسبة النجاح بعد التعديلات؟ (استعمل مستوى مغزوية 5%).

الحل:

أ) الفرضيات :

$$H_0 : P = 0.72$$

$$H_1 : P > 0.72$$

ب) النموذج:

لكبر حجم العينة يستخدم التوزيع الطبيعي. وحيث H_1 على الصيغة > يكون الاختبار طرف واحد وقيمة $Z_{0.05}$ تساوي 1.65 .

(ج) حساب Z من العينة:

$$Z = \frac{0.8 - 0.72}{\sqrt{\frac{0.8 \times 0.2}{10000}}} = 20 > 1.65$$

(د) القرار:

ترفض فرضية العدم ويكون القرار أن نسبة النجاح قد تحسنت بعد التعديلات.

- اختبار الفرضيات للفرق $\mu_1 - \mu_2$ بدلالة: $\bar{x}_1 - \bar{x}_2$

* لاختبار الفرق في الوسط الحسابي لمجتمعين يتبع نفس الطريقة لاختبار فرضيات الوسط الحسابي لمجتمع واحد وتتلخص الخطوات في الآتي :

1. تحديد فرضية العدم والبديل .
2. تحديد توزيع $\bar{x}_1 - \bar{x}_2$.
3. تحديد مستوى المغزوية والقيم الحرجة .
4. حساب قيم Z أو t بناءً على فرضية العدم .
5. الوصول إلى قرار حول قبول أو رفض فرضية العدم ومن ثم فرضية البديل .

مثال:

أجريت التجربة التالية لمقارنة نوعين من الأعلاف. خصص لكل علف 10 أبقار وكان معدل الزيادة في أوزان الأبقار كما يلي:

* علف رقم (1): 31، 34، 29، 26، 32، 35، 38، 34، 30، 32
* علف رقم (2): 28، 26، 24، 28، 29، 30، 29، 26، 31، 29

هل هناك فرق مغزوي بين نوعي العلف وذلك بمستوى مغزوية 0.05؟

الحل:

أ) الفرضيات :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu$$

ب) النموذج:

لصغر حجم العينتين وعدم معرفة الانحراف المعياري للمجتمعين يستخدم التوزيع تي وحيث أن H_1 على الصيغة $\neq$ فإن القيم الحرجة $t_{\alpha/2,v}$ تساوي 2.11 حيث . $v = n_1 + n_2 - 2 = 18$

(ج) حساب t من العينة:

من البيانات نجد أن $\bar{x}_1 = 32.1$ ، $\bar{x}_2 = 28$ ، $Se(\bar{x}_1 - \bar{x}_2) = 1.33$

عليه:

$$t = \frac{(32.1 - 28) - 0}{1.33} = 3.02 > 2.11$$

(د) القرار:

ترفض فرضية العدم ونقرر أنه يوجد اختلاف معزوي بين نوعي العلف.