الفصل السابع

التقدير الإحصائي لعينتين

Statistical Estimation: Two Samples

في بعض الأحيان نحتاج لمقارنة منتجين أو طريقتين لتقييم المهمات أو الخدمات، وعليه فإنه يجب علينا جمع بيانات من عينتين مختلفتين والتعامل مع متغيرين عشوائيين. ولكنّه لا يمكننا تقدير تلك المتغيرين؛ نظراً لعدم وجود توزيعات احتماليه يمكن استخدامها في هذه الحالة. لذلك يمكن إجراء تعديل بسيط في المتغيرات مما يمكننا من استخدام التوزيعات الاحتمالية المشار لها في الفصل السادس. فعلى سبيل المثال، بدلاً من التعامل مع متغيرين عشوائيين $\overline{X}_1$ ، $\overline{X}_2$ واستخدامها بصورة منفصلة لتقدير متوسط المجتمعتين موضوع الدراسة يمكن تحويلها لمتغير عشوائي واحد بطرح أحدهما من الآخر. ويصبح المتغير العشوائي الجديد هو $(\overline{X}_1 - \overline{X}_2)$ والذي ينتمي لتوزيع احتمالي طبيعي ذو متغير واحد متوسطه $(\mu_1 - \mu_2)$ وتباينه $(\sigma_1^2/n_1 + \sigma_2^2/n_2)$. والاختلاف الوحيد في هذه الحالة هو التباين حيث إنه لا يمكن أن يكون سالباً وعليه تم جمع تباين المتغيرين العشوائيين.

ويمكن ببساطة الحصول على تقديرات النقطة بحساب قيم المتغيرات العشوائية لكل عينة. وسيتم عرض فترات الثقة في الأجزاء القادمة لمجموعة من المتغيرات العشوائية لكل منها توزيع معاينة مختلف.

فترات الثقة للفرق بين متوسطي مجتمعين $(\mu_1 - \mu_2)$ لها انحرافات معيارية معلومة

Confidence Intervals for $(\mu_1 - \mu_2)$ with σ_1 σ_2 Known

نفترض أن المنتج يرغب في مقارنة متوسط عدد أيام الشبق لمجموعتين مختلفة من منتجات الهرمونات المعطاة لقطيع الماشية. قام المنتج بأخذ عينة عشوائية من المجتمع الأول والذي يساوي متوسطة μ_1 وانحرافه المعياري σ_1 حيث تمت معاملة الحيوانات منها بالجرعة الأولى وحسب متوسط العينة $\overline{X}_1$ ، ثم أخذت عينة عشوائية ثانية مستقلة من المجتمع الثاني والذي متوسطه μ_2 وانحرافه المعياري σ_2 حيث تمت معاملة الحيوانات في هذا المجتمع بالجرعة الثانية ثم حسب متوسط العينة $\overline{X}_2$.

فإذا قام المنتج بأخذ عينات متكررة من المجموعين بحجم n_1 ، n_2 وقام بحساب المقدر $(\overline{X}_1 - \overline{X}_2)$ لكل تجربة فإنه سيحصل على توزيع معاينة طبيعي بمتوسط يساوي الفرق بين متوسطات المجتمع وتباين يساوي مجموع تباين المجتمعين. وعليه فإنه يمكن حساب فترة الثقة للفرق بين متوسطي المجتمع $(\mu_1 - \mu_2)$ بدرجة ثقة تساوي $100(1-\alpha)$:

$$P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) = (1-\alpha) \qquad (7.1)$$

وبالتعويض في المعادلة أعلاه بقيمه z والتي تساوي :

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_2 - \mu_2)}{\sqrt{\frac{\sigma^2{}_1}{n_1} + \frac{\sigma^2{}_2}{n_2}}} \qquad (7.2)$$

حيث المقام يعبّر عن الخطأ المعياري للفرق بين المتوسطات ويرمز له بالرمز $\sigma_{(\bar{x}_1 - \bar{x}_2)}$. وهذا الحد يدخل ضمن المعادلة التي نحسب منها فترة الثقة بدرجة ثقة

$(1-\alpha)$ 100 والتي يمكن التعبير عنها بالمعادلة التالية :

$$(\bar{X}_1-\bar{X}_2)-Z_{\alpha/2}.\sigma(\bar{x}_1-\bar{x}_2)<(\mu_1-\mu_2)<(\bar{X}_1-\bar{X}_2)+Z_{\alpha/2}.(\sigma_{\bar{x}_1-\bar{x}_2}) \quad (7.3)$$

والتي تمثل فتره الثقة للفرق بين متوسطات المجتمع عندما يكون الانحراف المعياري لمجتمع الدراسة معلوم. ويمكن إيضاح تطبيق ذلك من خلال المثال التالي :

قامت إحدى المختصات بدراسة التربة بإجراء اختبار معملي لمحتويات العنب الأحمر من السكر فوجدت أن الانحراف المعياري له يساوي خمسة $(\sigma_1=5)$ عند زراعته في تربة رملية بينما يكون الانحراف المعياري له يساوي ٧,٧٥ $(\sigma_2=7.75)$ عند زراعته في تربه رمليه طينية. وحيث إن الهدف من ذلك هو تقدير متوسط الفرق في محتويات السكر لأنواع التربة ؛ بهدف تحديد شروط الزراعة بشكل تجاري. لذا ، قامت الباحثة بأخذ عينة عشوائية من حقول العنب المزروعة في التربة الرملية حجمها ٢٥ $(n_1=25)$ فوجدت أن متوسط السكر فيها يساوي ٢١ $(\bar{X}_1=21)$ وأخذت عينه عشوائية أخرى من حقول العنب المزروع في التربة الرملية الطينية حجمها ٣٠ $(n_2=30)$ فوجدت أن متوسط السكر فيها يساوي ٢٦ $(\bar{X}_2=26)$. والمطلوب حساب فتره الثقة بدرجه ٩٨٪ لفرق المتوسط للسكر في العنب الأحمر المنتج بشكل تجاري في هذه التربة.

لكتابة فتره الثقة يجب أولاً أن نحسب قيمة $Z_{\alpha/2}$ من الملحق جدول (٧)، والتي تساوي ±٢,٣٣ ثم نحسب الخطأ المعياري للفرق في المتوسطات كالتالي :

$$\sigma_{(\bar{x}_1-\bar{x}_2)}=\sqrt{\sigma_1{}^2/\mathrm{n}_1+\sigma_2{}^2/\mathrm{n}_2)}=\sqrt{(5^2/25+7.75^2/30)}=\sqrt{3.002}=1.73$$

بعد ذلك يتم التعويض في المعادلة رقم (7.3) بالقيم المتحصل عليها آنفاً لنحصل على فتره الثقة بدرجة ٩٨٪ والتي تساوي :

$$(\bar{X}_1 - \bar{X}_2) - Z_{\alpha/2} \cdot \sigma_{(\bar{x}_1 - \bar{x}_2)} < (\mu_1 - \mu_2) < (\bar{X}_1 - \bar{X}_2) + Z_{\alpha/2} \cdot \sigma_{(\bar{x}_1 - \bar{x}_2)}$$
$$(21 - 26) - 2.33 \cdot (1.73) < (\mu_1 - \mu_2) < (21 - 26) + 2.33 \cdot (1.73)$$
$$-5 - 4.03 < (\mu_1 - \mu_2) < -5 + 4.03$$
$$-9.03 < (\mu_1 - \mu_2) < -0.97.$$

وعليه، فإنه يمكن القول وبدرجه ثقة ٩٨٪ إن متوسط السكر للعنب الأحمر المنتج تجارياً في التربة الرملية الطينية أكبر بمعدل يتراوح بين ٠,٩٧ و ٩,٠٣ وحدة عند مقارنته بما يتم إنتاجه في التربة الرملية.

فترات الثقة للفرق بين متوسطي مجتمعين $(\mu_1 - \mu_2)$ لها انحرافات معيارية غير معلومة

Confidence Intervals for $(\mu_1 - \mu_2)$ with σ_1 σ_2 Unknown

لتقدير الفرق بين متوسطي مجتمعين عندما يكون الانحراف المعياري لهما غير معلوم نستخدم توزيع t ولكن في الحقيقة يعتمد استخدامه على حجم العينات. فإذا كانت العينتان كبيرتين، أي أن حجمهما أكبر من ٣٠ $(n_1 , n_2 > 30)$، فإنه يمكن استخدام تباين العينات كمقدّر لتباين المجتمع، في حين أنه لا يمكن عمل ذلك في حالة العينات الصغيرة.

العينات الكبيرة Large Samples

في العينات الكبيرة يستخدم التوزيع الطبيعي كتقريب لتوزيع t لتقدير الفرق بين المتوسطات ويمكن إيجاد فتره الثقة بطريقة مشابهة تقريباً للجزء السابق ماعدا أنه يتم حساب Z باستخدام الصيغة الرياضية الموضحة في المعادلة رقم (7.4):

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2_1}{n_1} + \frac{S^2_2}{n_2}}} \qquad (7.4)$$

حيث يتم التعويض بها في المعادلة التالية :

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

ومن ثم حلّها رياضياً للحصول على فتره الثقة المطلوبة والموضحة في المعادلة رقم (7.5) :

$$(\overline{X}_1 - \overline{X}_2) - Z_{\alpha/2} \cdot S_{(\overline{x}_1 - \overline{x}_2)} < (\mu_1 - \mu_2) < (\overline{X}_1 - \overline{X}_2) + Z_{\alpha/2} \cdot S_{(\overline{x}_1 - \overline{x}_2)} \qquad (7.5)$$

حيث إن $S_{(\overline{x}_1 - \overline{x}_2)} = \sqrt{S_1^2/n_1 + S_2^2/n_2}$ عبارة عن تقدير للخطأ المعياري للمجتمع للفرق بين المتوسطات بناء على تباين العينة. ويمكن عرض ذلك من خلال المثال التالي :

يمكن صناعة جزء من المحراث باستخدام عمليات البثق أو يمكن صنعه باستخدام آلة القوالب ثم تلحيمه ويرغب مدير المشروع في معرفة قوة ذلك الجزء المصنوع بالطريقتين حيث تتوفر لديه مجموعة من الأجزاء مصنوعة وقام بحساب الإحصاءات الموضحة في الجدول رقم (٧،١). والمطلوب إيجاد فتره الثقة بدرجة ثقة ٩٩٪ لمتوسط الفرق لقوة الجزء المصنوع بكلا الطريقتين.

ولإيجاد فتره الثقة المطلوبة فإننا أولاً نوجد $Z_{\alpha/2}$ من الجدول رقم (٧) بالملحق عند مستوى معنوية $\alpha = 0.01$ حيث إن القيمة تساوي ± ٢،٥٧.

وبالتعويض في معادلة الانحراف المعياري بالقيم المذكورة أعلاه نستطيع حساب قيمة الانحراف المعياري كالتالي :

$$S_{(\bar{x}_1-\bar{x}_2)} = \sqrt{S_1^{\,2}/n_1 + S_2^{\,2}/n_2} = \sqrt{(256/64+171.5/49)} = \sqrt{7.5} = 2.74$$

الجدول رقم (٧.١). بيانات العينة المتحصل عليها لطريقي البثق و آله القوالب لصناعة جزء المحراث.

البند	طريقه البثق	طريقه الطبع والتلحيم
حجم العينة	٦٤	٤٩
متوسط القوة	٥١٠	٤٩٠
تباين العينة	٢٥٦	١٧١,٥

ومن ثم يتم التعويض بالقيم في المعادلة رقم (7.5) كالتالي :

$$(\overline{X}_1-\overline{X}_2)-Z_{\alpha/2}.S_{\bar{x}_1-\bar{x}_2} < (\mu_1-\mu_2) < (\overline{X}_1-\overline{X}_2)+Z_{\alpha/2}.S_{\bar{x}_1-\bar{x}_2}$$
$$(510-490)-2.57.\ 2.74 < (\mu_1-\mu_2) < (510-490)+2.57.2.74$$
$$20-7.04, < (\mu_1-\mu_2) < 20+7.04$$
$$12.96 < (\mu_1-\mu_2) < 27.04$$

وبناء على ذلك فإنه وبدرجة ثقة ٩٩٪ تكون أجزاء المحراث المصنوعة بطريقه البثق أقوى في المتوسط بمقدار يتراوح بين ١٢,٩٦ و ٢٧,٠٤ مقارنة بطريقة استخدام آلة القوالب والتلحيم.

العينات الصغيرة المستقلة Small, Independent Samples

عندما تكون العينات موضع الدراسة صغيره ويتم سحبها بطريقه مستقلة عن بعضها فإن التوزيع المناسب في هذه الحالة هو توزيع t والذي يستخدم لدراسة المتغير العشوائي الذي هو عبارة عن الفرق بين متوسطات المجتمع.

وفي هذه الحالة فإن العينات الصغيرة لا تعطي تقدير ثابت للتباين وعليه فإننا

نستخدم في هذه الحالة جميع البيانات المتحصل عليها لتقدير التباين حيث يتم استخدام الصيغة الرياضية الموضحة في المعادلة رقم (7.6) بحيث يتم ترجيح التباين باستخدام درجات الحرية[١] :

$$S^2 = \frac{(n_1 - 1)S_1^{\ 2} + (n_2 - 1)S_2^{\ 2}}{n_1 + n_2 - 2} \tag{7.6}$$

وللحصول على الانحراف المعياري لفرق المتوسطات $S_{\overline{X}_1 - \overline{X}_2}$ يتم قسمة التباين المتحصل عليه من المعادلة رقم (7.6) باستخدام بيانات كل عينه ثم جمع الناتج وإيجاد الجذر التربيعي كما في المعادلة رقم (7.7) .

$$S_{\overline{X}_1 - \overline{X}_2} = \sqrt{S^2 / n_1 + S^2 / n_2}\,. \tag{7.7}$$

وللحصول على فتره الثقة بدرجه حرية $(1-\alpha)100$ باستخدام توزيع t يمكن عرض ذلك كالتالي :

$$\begin{gathered} P(-t_{(\alpha/2,\, v)} < t < t_{(\alpha/2,\, v)}) = (1-\alpha) \\ t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S^2}{n_1} + \dfrac{S^2}{n_2}}} \end{gathered} \tag{7.8}$$

وبالتعويض عن قيمة t في معادلة فتره الثقة وإعادة ترتيب المعادلة فإن فتره الثقة لفرق متوسطات المجتمع $(\mu_1 - \mu_2)$ والتي نرغب في إيجادها تكون كما في المعادلة رقم (7.9) التالية :

$$(\overline{X}_1 - \overline{X}_2) - t_{(\alpha/2,\, v)} \cdot S_{\overline{x}_1 - \overline{x}_2} < (\mu_1 - \mu_2) < (\overline{X}_1 - \overline{X}_2) + t_{(\alpha/2,\, v)} \cdot S_{\overline{x}_1 - \overline{x}_2} \tag{7.9}$$

ويمكننا إيضاح كيفية تطبيق ذلك من خلال المثال التالي :

قام أحد مربي الحيوانات بدراسة لمعرفة تأثير إضافة فيتامين لمياه الشرب على الوزن المكتسب للحيوان فاختار عشرة حيوانات ووضع الفيتامين في مياه الشرب وأختار مجموعه أخرى حجمها ٩ حيوانات كمجموعة تحكم دون أن يضيف لها فيتامين. وقد لاحظ المربي بأن متوسط الوزن المكتسب من الولادة إلى الفطام للحيوانات المضاف لمياه الشرب الخاصة بها فيتامين يساوي ٢٥ رطلاً بانحراف معياري ٨ رطل بينما متوسط الوزن المكتسب لمجموعة التحكم التي لم يضاف الفيتامين للماء يساوي ٢٠ رطلاً بانحراف معياري ١١ رطلاً.

المطلوب إيجاد فتره الثقة لمتوسط الفرق في الوزن المكتسب للمجموعتين بدرجه ثقة ٩٥٪.

لإيجاد فتره الثقة نحسب أولاً درجه الحرية كالتالي :

$$\nu = (n_1 + n_2 - 2) = (10 + 9 - 2) = 17$$

وحيث إن مستوى الثقة المطلوب ٩٥٪ فإن $\alpha = 0.05$ ومن الجدول رقم (٨) بالملحق نجد أن قيم t هي ± ٢,١١ وعليه يمكن حساب قيمة التباين التجميعي كما يلي.

$$S^2 = [(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2] / (n_1 + n_2 - 2)$$
$$S^2 = [(10 - 1)8^2 + (9 - 1)11^2]/(10 + 9 - 2)$$
$$S^2 = [576 + 968]/17 = 90.82$$

وباستخدام القيمة المتحصل عليها أعلاه يمكن حساب الخطأ المعياري لفرق المتوسطات.

$$S_{\bar{x}_1-\bar{x}_2} = \sqrt{S^2/n_1 + S^2/n_2}$$
$$S_{\bar{x}_1-\bar{x}_2} = \sqrt{90.82/10 + 90.82/9}$$
$$S_{\bar{x}_1-\bar{x}_2} = \sqrt{9.082 + 10.091} = \sqrt{19.173}$$
$$S_{\bar{x}_1-\bar{x}_2} = 4.38$$

وأخيراً يمكن التعبير عن فتره الثقة للفرق بين المتوسطات باستخدام المعادلة رقم (7.9) والتعويض بالقيم المتحصل عليها أعلاه في ذلك كالتالي :

$$(\bar{X}_1 - \bar{X}_2) - t_{(\alpha/2,v)} \cdot S_{\bar{x}_1-\bar{x}_2} < (\mu_1 - \mu_2) < (\bar{X}_1 - \bar{X}_2) + t_{(\alpha/2,v)} \cdot S_{\bar{x}_1-\bar{x}_2}$$
$$(25-20) - 2.11 \cdot 4.38 < (\mu_1 - \mu_2) < (25-20) + 2.11 \cdot 4.38$$
$$5 - 9.24 < (\mu_1 - \mu_2) < 5 + 9.24$$
$$-4.24 < (\mu_1 - \mu_2) < 14.24$$

ونلاحظ أن قيمة فترة الثقة تشمل قيمة الصفر حيث إنها تمتد من -٤,٢٤ إلى ١٤,٢٤ وهذا يعني أنه قد لا يكون هنالك فرق في الوزن المكتسب للحيوانات التي تم تغذيتها بالفيتامين مقارنة بحيوانات التحكم ؛ نظراً لأن $(\mu_1 - \mu_2) = 0$ مشمولة بفترة الثقة.

وبناء على هذه النتائج للتجربة ، فإن منتج الحيوانات قد لا يكون لديه الرغبة في شراء الفيتامين لحيواناته. وهذه الفترة كبيرة نسبياً ويحدث ذلك في حالة أن حجم العينة صغير والتباين كبير. ويمكن إجراء بعض الخطوات لتقليل التباين وإعادة التجربة إذا كان ذلك ممكناً.

العينات الصغيرة المزدوجة (غير المستقلة) Small, Paired Samples

إذا كان لدينا عينات صغيرة الحجم وبيانات مترابطة ، مثلاً البيانات غير مستقلة ؛ نظراً لوجود بعض العوامل التي تجعلها مترابطة ، في هذه الحالة نستخدم

توزيع t كقاعدة لحساب فترة الثقة بدرجه ثقة $100(1-\alpha)$ لفرق متوسطات المجتمع $(\mu_1 - \mu_2)$، وهذا يعني أننا ننشئ فتره الثقة باستخدام البيانات التي تمثل الفرق بين المشاهدات للعينتين بدلاً من استخدام البيانات الأصلية نفسها. ويحدث اقتران البيانات بعضها؛ نظراً لنوع التجربة التي تم إجراؤها. في هذه الحالة، فإن التجربة تشتمل على مشاهدات في نفس المجال قبل وبعد إجراء التجربة، أو قد تتأثر المشاهدات في العينتين ببعض العوامل التي لا يمكن التحكم فيها في حالة التصميم العشوائي، مثل نوع التربة في حالة تجربة إنتاجية القطن. وفي هذه الحالة فإن حجم العينات يجب أن يكون متساوياً حتى نستطيع ربطهما وعليه إيجاد الفرق، ويمكن التغلب على العامل المسبب لارتباط العينات ببعضها عن طريق طرح قيم العينات من بعضها فمثلاً الوراثة تؤثر على قدرة الحيوانات على اكتساب الوزن. وعليه فإنه عند إجراء الدراسات المتعلقة بالوزن المكتسب يفضل استخدام الحيوانات التي بينها ترابط مثل التوائم وإجراء اختبار t على النتائج.

ويمكن اشتقاق فتره الثقة باستخدام الصيغة الاحتمالية التالية:

$$P(-t_{(\alpha/2,v)} < t < t_{(\alpha/2,v)} = (1-\alpha)$$

ولكن في هذه الحالة فإن الصيغة الرياضية المستخدمة لحساب قيمة t مختلفة. ويمكن توضيح ذلك بتعريف المتغير العشوائي d والذي يمثل الفرق بين بيانات العينة، أي $d_1 = x_{11} - x_{21}$ ، $d_2 = x_{12} - x_{22}$ .... ، $d_n = x_{1n} - x_{2n}$ لجميع بيانات العينة، و $\bar{d}$ يعبّر عن متوسط الفروق، و S_d^2 يعبّر عن تباين الفروق. ويمكن حساب هذه المتغيرات مثل أي حساب لعينه مفرده فمثلاً $\bar{d} = \sum d/n$ ، إلى آخره. ويمكن كتابة الصيغة الرياضية لـ t كالتالي:

$$t = \frac{\bar{d} - (\mu_1 - \mu_2)}{\frac{S_d}{\sqrt{n}}} \tag{7.10}$$

وبالتعويض بالطرف الأيمن من معادلة t السابقة في صيغة فترة الثقة السابقة وإجراء بعض التعديلات نحصل على فترة الثقة لـ $(\mu_1 - \mu_2)$ والموضحة في المعادلة رقم (7.11) كالتالي :

$$\bar{d} - t_{(\alpha/2, v)} \cdot S_d / \sqrt{n} < (\mu_1 - \mu_2) < \bar{d} + t_{(\alpha/2, v)} \cdot S_d / \sqrt{n} \tag{7.11}$$

ويمكن إيضاح كيفية تطبيق ذلك من خلال المثال التالي :

لاختبار هجينين جديدين من حبوب الذرة المزروعة تحت ظروف زراعة أولية ، قامت شركة البذور باختيار عدد ٩ مزارع عشوائياً وتم إعطاؤها للمزارع لزراعتها في أحواض التجربة. وتم قياس الإنتاجية بالقنطار / أيكر للمواقع التسعة والموضحة في الجدول رقم (٧,٢).

الجدول رقم (٧,٢). إنتاجية حبوب الذرة بالقنطار/أيكر لعدد تسع مزارع مختلفة.

المزرعة	هجين ١	هجين ٢
١	٨٥	٧٩
٢	٨٨	٨٠
٣	٥٨	٦٠
٤	٩٤	٩٢
٥	٨٥	٧٨
٦	٩٣	٨٧
٧	٧٤	٧٥
٨	٨٠	٧٦
٩	١٠١	٩٥

المطلوب إيجاد فترة الثقة لمتوسط فرق الإنتاجية للهجينين بدرجة ثقة ٩٥٪.

للحصول على فتره الثقة بدرجة ٩٥٪ نحسب أولاً درجة الحرية كالتالي :

$$v = (n-1) = (9-1) = 8$$

ثم نوجد قيم t عند مستوى معنوية $\alpha = 0.05$ من الجدول رقم (٨) بالملحق والتي تساوي ±٢,٣٠٦ . ثم نحسب قيمة $\bar{d}$ و S_d بحساب d أولاً وإجراء الحسابات اللازمة باستخدام الصيغة الرياضية الخاصة بها والموضحة في الجدول رقم (٧,٣) وهي :

$$\bar{d} = \Sigma d/n = 36/9 = 4$$
$$S_d^{\,2} = \Sigma(d-\bar{d})^2/(n-1) = 102/(9-1) = 12.75$$
$$S_d = \sqrt{S_d^{\,2}} = \sqrt{12.75} = 3.57$$

الجدول رقم (٣.٧). حسابات متوسط وتباين الفرق d لبيانات إنتاجية الذرة الهجين.

المزرعة	هجين ١	هجين ٢	d	$d-\bar{d}$	$(d-\bar{d})^2$
١	٨٥	٧٩	٦	٦- ٤=٢	٤
٢	٨٨	٨٠	٨	٨- ٤=٢	١٦
٣	٥٨	٦٠	- ٢	- ٢- ٤=- ٦	٣٦
٤	٩٤	٩٢	٢	٢- ٤=- ٢	٤
٥	٨٥	٧٨	٧	٧- ٤=٣	٩
٦	٩٣	٨٧	٦	٦- ٤=٢	٤
٧	٧٤	٧٥	- ١	- ١- ٤=- ٥	٢٥
٨	٨٠	٧٦	٤	٤- ٤=٠	٠
٩	١٠١	٩٥	٦	٦- ٤=٢	٤
الإجمالي			٣٦	٠	١٠٢

وبالتعويض عن القيم نحصل على فترة الثقة، كما في الخطوات التالية :

$$\bar{d} - t_{(\alpha/2,v)} + S_d/\sqrt{n} < (\mu_1 - \mu_2) < \bar{d} + t_{(\alpha/2,v)} \cdot S_d/\sqrt{n}$$

$$4 - 2.306 \cdot 3.57/\sqrt{9} < (\mu_1 - \mu_2) < 4 + 2.306 \cdot 3.57/\sqrt{9}$$

$$4 - 2.74 < (\mu_1 - \mu_2) < 4\ 2.74$$

$$1.26 < (\mu_1 - \mu_2) < 6.74$$

وتشير فترة الثقة المتحصل عليها أنه بدرجة ثقة ٩٥٪ فإن متوسط الإنتاجية للهجين الأول أكبر من الهجين الثاني بمقدار يتراوح بين ١,٢٦ و ٦,٧٤.وباستخدام هذه الطريقة تم إلغاء أثر الفروقات الفردية من مزرعة إلى مزرعة والتي تؤثر على مستوى الإنتاجية مثل سقوط الأمطار، اختلاف خصوبة التربة، وكذلك المقدرة الإدارية للملاّك.

فترات الثقة للفرق بين نسبتين $p_1 - p_2$

Confidence intervals for $p_1 - p_2$

يمكن إيجاد فترة الثقة بدرجة ثقة $100\,(1-\alpha)$ للفرق بين النسب للعينات باستخدام تقريب التوزيع الطبيعي لتوزيع المعاينة ذي الحدين للمتغير العشوائي $(\hat{P}_1 - \hat{P}_2)$ كمقدّر للفرق بين نسب المجتمع $(P_1 - P_2)$ ،حيث $\hat{P}_1$ تعرف على أنها r_1/n_1 و $\hat{P}_2$ تعرف على أنها تساوي r_2/n_2. وتكون صيغة فترة الثقة ممثلة بالتالي :

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

و Z معرّفه بالمعادلة رقم (7.12) التالية :

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{S_{p_1 - p_2}} \qquad (7.12)$$

يمكن الحصول على الثقة $(P_1 - P_2)$ بالتعويض عن الطرف الأيمن من معادلة Z آنفاً في صياغة فترة الثقة وإعادة ترتيب الحدود لنحصل على المعادلة رقم (7.13) والتي تعبّر عن فترة الثقة.

$$(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2} \cdot S_{p_1-p_2} < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2} \cdot S_{p_1-p_2} \qquad (7.13)$$

ويمكن إيضاح تطبيق الصيغة الرياضية أعلاه من خلال المثال التالي :

ترغب إحدى المختصات بالإنتاج النباتي في تقدير الفرق في نسبة نمو عقل العنب في البيوت المحمية. قامت بمعاملة مجموعة من النباتات عددها ١٦٠ بمنشط للجذور ولاحظت أن ١٤٠ نباتاً تبقى حية في نهاية التجربة، بينما تركت مجموعة أخرى من النباتات عددها ١٢٠ بدون معاملة ولاحظت أن ٩٦ منها تبقى حية. المطلوب إيجاد فترة الثقة بدرجة ثقة ٩٥٪ لنسبة الفرق في النباتات الحية لهذه ا لتجربة.

لإيجاد فترة الثقة فإننا أولاً يجب أن نحسب قيمة Z من الجدول رقم (٧) بالملحق عند قيمه $\alpha = 0.05$ حيث نجد أن $Z_{\alpha/2} = \pm 1.96$. ثم نوجد الحسابات التالية :

$$\hat{p}_1 = r_1/n_1 = 140/160 = 0.875;$$
$$\hat{p}_2 = r_2/n_2 = 96/120 = 0.800$$
$$S_{p_1-p_2} = \sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}$$
$$S_{p_1-p_2} = \sqrt{(0.875)(1-0.875)/160 + (0.800)(1-0.800)/120}$$
$$S_{p_1-p_2} = \sqrt{0.1094/160 + 0.1600/120}$$
$$S_{p_1-p_2} = \sqrt{0.00068 + 0.00133}$$
$$S_{p_1-p_2} = \sqrt{0.002} = 0.04$$

وعليه فإن لدينا الآن معلومات كافيه لإيجاد فترة الثقة باستخدام المعادلة رقم (7.13) كالتالي :

$$(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2} \cdot S_{p_1 - p_2} < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + Z_{\alpha/2} \cdot S_{p_1 - p_2}$$

$$(0.875 - 0.800) - 1.96\,(0.04) < (p_1 - p_2) < (0.875 - 0.800) + 1.96 \cdot (0.04)$$

$$0.075 - 0.078 < (p_1 - p_2) < 0.075 + 0.078$$

$$-0.003 < (P_1 - P_2) < 0.153$$

وحيث إن فترة الثقة تحتوي على الصفر فإن المختصة ربما تحتاج لمعلومات إضافية قبل استخدام المنشط الجذري. وتشير النتائج آنفاً إلى أن الفرق في نسبة نمو عقل العنب المعاملة بالمنشط الجذري يتراوح بين -٠,٠٠٣ و ٠,١٥٣ مقارنة بالنباتات التي لم تعامل بالمنشط الجذري.

خاتمة Endnote

(١) يمكن عرض صيغة إجمالي درجات الحرية كالتالي:

$(n_2 - 1) + (n_1 - 1)$ يمكن كتابتها على النحو $(n_1 + n_2 - 2)$

تمارين Exercises

١- في عينة عشوائية بسيطة لعدد ١٠٠ حيوان من سلالات مختلطة وجد أن متوسط وزن البيع يساوي ٢٢٠ رطلاً بانحراف معياري يساوي ٢٠ رطلاً بينما في عينه لعدد ٤٠٠ حيوان من سلالات نقيه وجد أن متوسط وزن البيع يساوي ٢١٠ رطل بانحراف معياري يساوي ٣٠ رطل. المطلوب إيجاد فترة الثقة بدرجه ثقة ٩٥٪ للفرق الحقيقي في متوسط وزن البيع للحيوانات في المجموعتين.

٢- يرغب أحد المزارعين في شراء أسلاك بالات لمزرعته فقام بأخذ عينة من أسلاك ربط متوسطة القوه حجمها ٥٠ قطعة وعينة أخرى من أسلاك ربط قويه حجمها ٦٠ قطعة وأختبرها في أحد المعامل القريبة لمعرفة قوة الشد للأسلاك

فكان تقرير المعمل كالتالي:

أ) متوسط قوة الشد للأسلاك متوسطة القوه يساوي ٩٥,٨ رطلاً بانحراف معياري يساوي ٥ رطل.

ب) ٣٤٪ من الأسلاك متوسطة القوه لها قوة شد تساوي ١٠٠ رطل أو أكثر.

ج) متوسط قوة الشد للأسلاك القوية تساوي ٩٩ رطلاً بانحراف معياري يساوي ٦ رطل.

د) ٣٠٪ من الأسلاك القوية لها قوة شد تساوي ١٠٠ رطل أو أكثر

والمطلوب:

- أوجد فتره الثقة بدرجه ثقة ٩٠٪ لمتوسط الفرق في قوة الشد لكلا النوعين من الأسلاك.
- أوجد فترة الثقة بدرجه ثقة ٩٥٪ للفرق الفعلي في نسبة السلك والتي لها قوة شد ١٠٠ أو أكثر لكلا النوعين.
- فسّر النتائج المتحصل عليها.

٣- في دراسة لتقييم معدل الإنبات لبذور غار الجبل تم اختيار ١٠٠ بذره وكشطها ثم زراعتها واختيار ٧٠ بذرة أخرى ومعاملتها بالحامض ثم زراعتها. وكانت النتائج أن ٥٠ بذرة من البذور المكشوطة نبتت بينما نبت ٤٨ بذرة من البذور المعاملة بالحامض. أوجد فتره الثقة للفرق في نسبة الإنبات لنوعي المعاملة التي تمت على البذور.

٤- في عينة عشوائية لعدد ١٠ مزارع بطيخ وجد أن متوسط أجر الساعة لعمال الحصاد يساوي ٧,٢٥ دولار بانحراف معياري ١,٢٥ دولار ، بينما متوسط أجر الساعة للعمال في ١٢ مزرعة فواكه يساوي ٨ دولارات بانحراف معياري يساوي ١

دولار. أوجد فتره الثقة بدرجه ثقة ٩٠٪ لمتوسط الفرق في أجر الساعة المدفوع لعمال الحصاد لكلا نوعي المزارع.

٥- في عينة عشوائية لعدد ٨ من مزارعي القمح وجد أن متوسط تكاليف البيع تساوي ٠,١دولار/ بوشل بانحراف معياري يساوي ٠,٠٥ دولار/بوشل ، بينما وجد في عينة عشوائية لعدد ١٥ من مزارعي الشعير أن متوسط تكاليف البيع تساوي ٠,٠٧ دولار/بوشل بانحراف معياري يساوي ٠,٠٣ دولار/بوشل. أوجد فتره الثقة بدرجه ثقة ٩٥٪ لمتوسط الفرق في تكاليف البيع للبوشل.

٦- تمت تغذية عدد ٦ مجموعات من توأم أبقار الحليب مع صغارهن بنفس الحصص ماعدا أحد التوأم أعطي هرمون BST والذي يزيد من إنتاج الحليب. المطلوب إيجاد فتره الثقة بدرجه ثقة ٩٥٪ للفرق في إنتاج الحليب بالرطل للحيوان للبيانات التالية :

التوأم	كمية الحليب بالرطل بدون BST	كمية الحليب بالرطل بـ BST
١	٤٥	٤٨
٢	٥٠	٥١
٣	٦٠	٦٤
٤	٥٥	٥٦
٥	٥٣	٥٧
٦	٤٩	٥٢

٧- البيانات التالية توضح الإنتاجية المشتركة لصنفين من الشعير. كل صنف تمت زراعته في مزرعة مختلفة. المطلوب تقدير الفرق في الإنتاجية بدرجه ثقة ٩٥٪ لفترة الثقة.

الصنف	إنتاجيه عاليه	صنف أوكالا ٩٥
١	٣٧	٤٢
٢	٤١	٤٥
٣	٣٥	٣٣
٤	٥٢	٥٦
٥	٤٤	٥٠
٦	٣٠	٣٢
٧	٣٨	٣٥
٨	٥٤	٥٥