

السلسلة السابعة معالجة البيانات

تمرين 1

تم وزن عينة بشكل متكرر ، فتم الحصول على النتائج التالية: 29.8 ، 30.2 ، 28.6 ، 29.7 . احسب المتوسط ، الانحراف المعياري .

تمرين 2

تم تحليل عينة ما لمعرفة تركيز الزئبق فيها ، وذلك 5 مرات متكررة باستخدام طريقة قياسية و 5 مرات أخرى باستخدام طريقة جديدة مقترحة، والجدول التالي يبين ملخص النتائج:

الطريقة الجديدة (ppm)	الطريقة القياسية (ppm)
10.5	10.1
9.9	10.3
10.4	10.2
11.2	10.3
10.5	10.4

والمطلوب حساب. sp

تمرين 3

في تحليل للذهب باستخدام طريقة جديدة وأخرى قياسية ، إذا كانت الـ *variances* للطريقتين 2.34 و 3.62 بالترتيب . فإذا علمت أنه تم الحصول على الـ *variance* للطريقة الجديدة عن طريق إجراء 7 تحاليل مكررة ، بينما تم إجراء 5 تحاليل مكررة باستخدام الطريقة القياسية ، هل هناك فرق إحصائي جوهري بين الـ *variances* للطريقتين عند مستوى ثقة 95% ؟

تمرين 4

عندما تم تحليل 5 عينات متكررة لخام الحديد فإن نتائج التحليل أشارت إلى أن نسبة الحديد كانت 14.1 ، 14.6 ، 15.5 ، 13.9 ، 15.4 فإذا علمت أن النتيجة الصحيحة للتحليل كانت % 14.3 ، هل هناك فرق إحصائي جوهري بين النتائج في الطريقتين عند مستوى ثقة 95% ؟

تمرين 5

في معايرة ما حصلنا على النتائج التالية 12.13 ، 12.25 ، 12.40 ، 12.24 ، و 15.25 مل . ويبدو واضحاً أن القيمة 15.25 تبدو شاذة ، فهل يتم الاحتفاظ بها أم استبعادها عند مستوى ثقة 95% .

تمرين 6

في تحليل نسبة الفضة في عينة ما حصلنا على النتائج التالية:
41.3 ، 45.0 ، 42.5 ، 42.4 ، 43.1 ، 42.2 ، 41.9 % ، ويبدو واضحاً أن القيمة 45.0 تبدو شاذة ، فهل يتم الاحتفاظ بها أم استبعادها عند مستوى ثقة 95% .

حلول السلسلة السابعة معالجة البيانات

تمرين 1

أولاً: لحساب المتوسط يتم جمع القيم والقسمة على عددها ، وذلك كما يلي:

$$\bar{X} = \frac{29.8 + 30.2 + 28.6 + 29.7}{4} = 29.6$$

تذكر أن العدد 4 لا يتدخل في تحديد عدد الأرقام المعنوية في الجواب لأنه عدد مضبوط يعبر عن عدد التجارب. وحيث أن عدم التأكد في جميع القيم في الجزء العشري فيجب أن يكون الجواب محتوياً جزءاً عشرياً لا أكثر ولا أقل
ثانياً: لحساب الانحراف المعياري نكون الجدول التالي

X_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	عدد الأرقام ذات الدلالة
29.8	0.2	0.04	1
30.2	0.6	0.36	1
28.6	1.0	1.00	2
29.7	0.1	0.01	1
		$\Sigma = 1.41$	2

$$s = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}}$$

$$s = \sqrt{\frac{1.41}{4 - 1}} = 0.69 \text{ mg}$$

لحساب s_p يجب بداية حساب متوسط نتائج المجموعة الأولى والثانية:

$$\bar{x}_1 = \frac{10.5 + 9.9 + 10.4 + 11.2 + 10.5}{5} = 10.5$$

$$\bar{x}_2 = \frac{10.1 + 10.3 + 10.2 + 10.3 + 10.4}{5} = 10.3$$

$$\begin{aligned} \sum_{i=1}^{N1} (x_{i1} - \bar{x}_1)^2 &= (10.5 - 10.5)^2 + (9.9 - 10.5)^2 \\ &+ (10.4 - 10.5)^2 + (11.2 - 10.5)^2 + (10.5 \\ &- 10.5)^2 = 0.86 \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^{N2} (x_{i2} - \bar{x}_2)^2 &= (10.1 - 10.3)^2 + (10.3 - 10.3)^2 \\ &+ (10.2 - 10.3)^2 + (10.3 - 10.3)^2 + (10.4 \\ &- 10.3)^2 = 0.06 \end{aligned}$$

$$\sum_{i=1}^{N1} (x_{i1} - \bar{x}_1)^2 + \sum_{i=1}^{N2} (x_{i2} - \bar{x}_2)^2 = 0.86 + 0.06 = 0.92$$

$$s_p = \sqrt{\frac{0.92}{5 + 5 - 2}} = 0.34 \text{ ppm}$$

تمرين 3

في البداية نوجد قيمة F بحيث تكون أكبر من 1 ، وذلك من العلاقة:

$$F = \frac{s_1^2}{s_2^2} = \frac{3.62}{2.34} = 1.547$$

حيث أن ال $variance$ الأكبر نتج عن 5 تجارب (درجات تحرر) بينما نتج ال $variance$ الأصغر عن 7 تجارب (درجات تحرر) فيمكن معرفة قيمة F من الجدول السابق ، وتساوي 4.53 ، وبمقارنة القيمة المحسوبة مع القيمة الجدولة نجد أن قيمة F المحسوبة أقل من قيمتها الجدولة $F_{calc} < F_{tab}$ وعليه نستنتج أنه لا يوجد فرق إحصائي جوهري بين ال ($variances$) للطريقتين ، بمعنى أن الطريقة الجديدة تعطي $precision$ مقارنة لتلك التي حصلنا عليها من الطريقة القياسية.

حيث أننا نريد مقارنة النتائج ، وبما أن النتيجة الصحيحة معلومة ، فإننا نستخدم العلاقة أدناه ، لكن في البداية نحن بحاجة لمعرفة المتوسط والانحراف المعياري:

$$\bar{x} = \frac{14.1 + 14.6 + 15.5 + 13.9 + 15.4}{5} = 14.7$$

أما الانحراف المعياري فيمكن حسابه ، حيث نحصل على قيمة (s = 0.73) ، والآن نحسب قيمة t من العلاقة:

$$\pm t = (\bar{x} - \mu) \frac{\sqrt{N}}{s}$$

$$\pm t = (14.7 - 14.3) \frac{\sqrt{5}}{0.73}$$

$$\pm t = 1.22$$

وبمقارنة $t_{calculated}$ مع $t_{tabulated}$ نجد أن المحسوبة (1.22) أقل من المجدولة (2.776) ، أربع درجات تحرر ومستوى ثقة (95 %) ، لذلك نستنتج أنه لا يوجد فرق إحصائي جوهري بين نتائج الطريقتين ، وإذا نجح إختبار F فإنه يمكن استخدام الطريقة الجديدة بدلاً من القياسية ، واعتبارهما متكافئتان.

تمرين 5

في البداية نقوم بترتيب النتائج إما تصاعدياً أو تنازلياً ، ومن ثم نستخدم العلاقة المذكورة أعلاه لحساب قيمة Q

$$\begin{array}{|c|} \hline a \\ \hline 15.25 , 12.40 , 12.25 , 12.24 , 12.13 \\ \hline w \\ \hline \end{array}$$

$$Q = a/w$$

$$Q = (15.25 - 12.40) / (15.25 - 12.13) = 0.913$$

لكن قيمة $Q_{tab} = 0.71$ عند مستوى ثقة 95 % لخمس نتائج ، وبما أن $Q_{calc} > Q_{tab}$ فهذا يعني أن القيمة 15.25 يجب استبعادها ، لأن السبب في قيمتها الشاذة عن القيم الأخرى لا يمكن أن يكون إحصائياً عند مستوى الثقة المطلوب.

تمرين 6

في البداية نقوم بترتيب النتائج إما تصاعدياً أو تنازلياً ، ومن ثم نستخدم العلاقة المذكورة أعلاه لحساب قيمة Q

45.0 % ، 43.1 ، 42.5 ، 42.4 ، 42.2 ، 41.9 ، 41.3

$$Q=aw$$

$$Q=(45.0-43.1)/(45.0-41.3)=0.514$$

لكن قيمة $Q_{tab}=0.568$ عند مستوى ثقة % 95 ، ومن الواضح أن $Q_{calc}<Q_{tab}$ فهذا يعني أنه يجب الاحتفاظ بالقيمة 45.0 ، لأن السبب في قيمتها الشاذة عن القيم الأخرى من الممكن أن يكون إحصائياً عند مستوى الثقة المطلوب.