إعداد د: محمود الدريني

تابع تحليل التباين الأحادي

ONE-WAY ANALYSIS OF VARANCE

3-4 اختبار تجانس التباينات

يستند تحليل التباين الأحادي على افتراض أن تباينات المعالجات متساوية، أي أن $\sigma_1^2 = \sigma_2^2 = ... = \sigma_t^2 = \sigma^2$، وللتحقق من صحة هذا الافتراض هناك العديد من الطرق الإحصائية لاختبار تجانس تباينات المعالجات، ومن أهمها اختبار بارتلت Bartelt ، واختبار Levene، وسوف يتم عرض خطوات اختبار Levene لسهولة وبساطة استخدامه.

- صياغة الفرض العدم والبديل:

الفرض العدم: تباينات المعالجات متجانسة

$$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_t^2 = \sigma^2$$

الفرض البديل : أحد التباينات على الأقل مختلف

$$H_1 : at\, least\, one\, of\; \sigma_i^2 \neq \sigma^2$$

- حساب إحصائية الاختبار:

- حساب الفروق المطلقة z_{ij}، حيث أن

$$z_{ij} = |\, y_{ij} - \bar{y}_i . |$$

- حساب إحصائية الإختبار بنفس الطريقة المتبعة في حساب إحصائية الاختبار F^* لاختبار تساوي متوسطات المعالجات، ولكن مع التطبيق على البينات الجديدة Z_{ij}، أي أن:

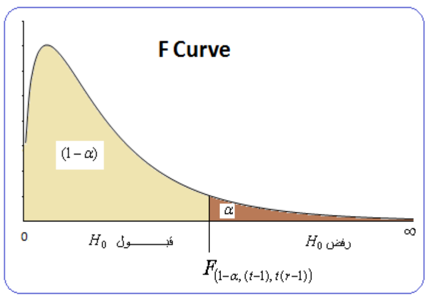
$$L^* = \frac{MSTr_{(z)}}{MSE_{(z)}} = \frac{r\sum_{i=1}^{t}(\bar{z}_i . - \bar{z}..)^2/(t-1)}{\sum_{i=1}^{t}\sum_{j=1}^{r}(z_{ij} - \bar{z}_i .)^2/(t(r-1))}$$

$$= \frac{\left(\frac{1}{r}\sum_{i=1}^{t} z_i .^2 - \frac{z..^2}{rt}\right)/(t-1)}{\left(\sum_{i=1}^{t}\sum_{j=1}^{r} z_{ij}^2 - \frac{1}{r}\sum_{i=1}^{t} z_i .^2\right)/(t(r-1))}$$

وهذه الإحصائية أيضا لها توزيع F بدرجات حرية بسط $(t-1)$، ودرجات حرية مقام $t(r-1)$

- تحديد مستوى المعنوية α.
- الكشف عن القيمة الحرجة أو القيمة الجدولية.

وهي قيمة تستخرج من جدول رقم (4) الخاص بتوزيع مئويات F عند درجة حرية بسط $(t-1)$ ، ودرجة حرية مقام $t(r-1)$ ، واحتمال تجميعي $(1-\alpha)$ أي يقع على يمينها مساحة منطقة الرفض α، ويرمز لهذه القيمة بالرمز $F_{(\alpha,\, t-1,\, t(r-1))}$ ، ويبين الشكل التالي مناطق رفض الفرض العدم وقبوله.



- القرار: إذا كانت القيمة المحسوبة L^* تزيد عن القيمة الجدولية $F_{(\alpha,\, t-1,\, t(r-1))}$ ، أي تقع في منطقة الرفض، فإنه لا يمكن قبول الفرض العدم $H_0: \sigma_1^2 = \sigma_2^2 = ... = \sigma_t^2 = \sigma^2$ ، ويقبل الفرض البديل $H_1: at\, least\, one\, of\, \sigma_i^2 \neq \sigma^2$، وفي هذه الحالة يكون افتراض تجانس التباين غير محقق ومن ثم يتم استخدام التحويلات الرياضية للوصول إلى الحالة التي يتحقق فيها تجانس التباينات.

تطبيق (1)

في التطبيق السابق (المحاضرة الرابعة) الخاص بنتائج إجراء تجربة لدراسة تأثيرات استخدام الكبريت في تخفيض مرض جرب البطاطس. تم الحصول على البيانات التالية:

	كمية الكبريت بالباوند				
	T_0	T_{300}	T_{600}	T_{1200}	
1	12	9	16	10	
2	10	9	10	4	
3	24	16	18	4	
4	29	4	18	5	
Total	75	38	62	23	198

إعداد د: محمود الدريني

تحقق من فرض تجانس نسب الإصابة بالنسبة للمعالجات الأربعة، $\alpha = 0.05$.

مناقشة التطبيق:

في هذا التطبيق يلاحظ أن: عدد المعالجات: $(t = 4)$، عدد المكررات: $(r_1 = r_2 = r_3 = r_4 = r = 4)$

وللتحقق من فرض تجانس التباينات للمعالجات الأربعة نقوم أولا بحساب الفروق المطلقة:

	كمية الكبريت بالباوند			
	T_0	T_{300}	T_{600}	T_{1200}
1	12	9	16	10
2	10	9	10	4
3	24	16	18	4
4	29	4	18	5
Total	75	38	62	23
Mean	**18.75**	**9.5**	**15.5**	**5.75**

	$z_{ij} = \lvert y_{ij} - \bar{y}_{i\cdot} \rvert$			
	T_0	T_{300}	T_{600}	T_{1200}
1	6.75	0.5	0.5	4.25
2	8.75	0.5	5.5	1.75
3	5.25	6.5	2.5	1.75
4	10.25	5.5	2.5	0.75
Total	**31**	**13**	**11**	**8.5**

ونحسب إحصائية الاختبار F^* كالتالي:

حساب مجموع المربعات بتطبيق المعادلات (10.15) ، (10.16) ، (10.17) كما هو مبين في المحاضرة السابقة

مجموع المربعات الكلي لبيانات $z_{ij} = \lvert y_{ij} - \bar{y}_{i\cdot} \rvert$:

$$SSTo_z = \sum_{i=1}^{t}\sum_{j=1}^{r} z_{ij}^2 - CF_z = \sum_{i=1}^{4}\sum_{j=1}^{4} z_{ij}^2 - CF_z$$

$$\sum_{i=1}^{4}\sum_{j=1}^{4} z_{ij}^2 = (6.75)^2 + \ldots + (0.75)^2 = 395.5$$

إعداد د: محمود الدريني

$$CF=\frac{z..^2}{tr}=\frac{(63.5)^2}{(4\times 4)}=\frac{4032.25}{16}=252.01563$$

$$SSTo_z=395.5-252.01563=143.484$$

مجموع مربعات المعالجات لبيانات $z_{ij}=|y_{ij}-\bar{y}_i.|$:

$$SSTr_z=\sum_{i=1}^{t}\frac{z_i.^2}{r}-CF_z$$

$$\sum_{i=1}^{t}\frac{z_i.^2}{r}=\frac{1}{r}\sum_{i=1}^{4}z_i.^2=\frac{(31)^2+(13)^2+(11)^2+(8.5)^2}{4}=\frac{1323.25}{4}=330.81$$

$$SSTr_z=330.81-252.01563=78.797$$

مجموع مربعات الأخطاء:

$$SSE_z=SSTo_z-SSTr_z=143.484-78.797=64.687$$

- درجات الحرية هي:

$$df_{SSTo}=tr-1=16-1=15,\qquad df_{SSTr}=(t-1)=(4-1)=3,$$
$$df_{SSE}=tr-t=16-4=12$$

وفيما يلي خطوات الاختبار.

-الفرض العدم: تباينات المعالجات متجانسة

$$H_0:\sigma_1^2=\sigma_2^2=\sigma_3^2=\sigma_4^2=\sigma^2$$

الفرض البديل : أحد التباينات على الأقل مختلف

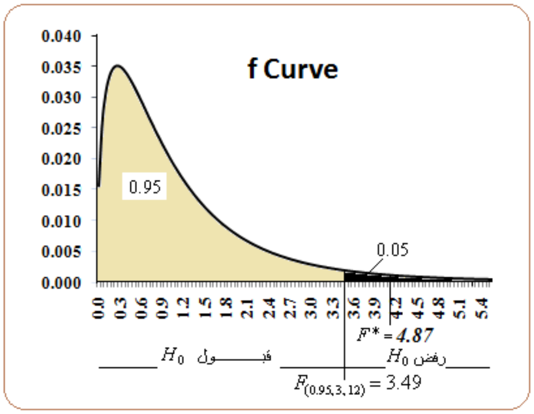
$$H_1: at\ least\ one\ of\ \sigma_i^2\neq\sigma^2$$

- حساب إحصائية الاختبار:

$$L^*=\frac{MSTr_{(z)}}{MSE_{(z)}}=\frac{SSTr_z/df_{tr}}{SSE_z/df_{err}}$$
$$=\frac{78.797/3}{64.687/12}=4.87$$

- مستوى المعنوية $\alpha=0.05$.
- القيمة الحرجة أو القيمة الجدولية. $F_{(\alpha,\ t-1,\ t(r-1))}=F_{(0.05,\ 4,\ 12)}=3.49$
- القرار: بما أن القيمة المحسوبة $L^*=4.87$ تزيد عن القيمة الجدولية $F_{(0.05,\ 4,\ 12)}=3.49$ ، أي تقع في منطقة الرفض

إعداد د: محمود الدريني



إذا لا يمكن قبول الفرض العدم $H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_t^2 = \sigma^2$ ، ويقبل الفرض البديل $H_1 : at\, least\, one\, of\ \sigma_i^2 \neq \sigma^2$ وفي هذه الحالة يكون افتراض تجانس التباين غير محقق ومن ثم يتم استخدام التحويلات الرياضية المناسبة لنسب الإصابة بجرب البطاطس. للوصول إلى الحالة التي يتحقق فيها تجانس التباينات.

4-4 التحويلات الرياضية.

في كثير من النواحي التطبيقية تكون تباينات المعالجات غير متجانسة، وينشأ ذلك في بعض الحالات من أن البيانات غير طبيعية non normality ، وأن الاختلافات داخل المعالجات مرتبطة بمتوسطات المعالجات، أي أن$\left[\sigma^2 = f(\mu)\right]$. ومن ثم إذا كان التوزيع الأصلي للبيانات معلوم، يمكن حساب العلاقة بين متوسطات المعالجات وتباينات المعالجات. كما يمكن تحويل البيانات أو قياسها بمعيار آخر، وعندئذ يقرب توزيع البيانات المحولة من التوزيع الطبيعي. مثل هذه التحويلات يكون الهدف منها جعل المتوسطات مستقلة عن التباينات، وتصبح التباينات الناتجة متجانسة. هذه النتيجة ليست دائما محققة. عندما يكون من غير الممكن إيجاد التحويلة التي تجعل المتوسطات والتباينات مستقلة والتباين ثابت أو مستقر، و من ثم هناك بعض طرق التحليل، مثل التحليلات المرجحة يمكن استخدامها.

من التحويلات الأكثر استخداما، الجذر التربيعي، اللوغاريثم، التحويل الزاوي، وسوف تناقش هذه التحويلات.

- **تحويلة الجذر التربيعي** $\sqrt{y}$

يستخدم الجذر التربيعي للبيانات في الحالة التي تتكون فيها البيانات من أعداد صحيحة صغيرة، ومثال على ذلك أعدد المستعمرات البكتيرية في عدد الصفائح، وعدد النباتات أو الحشرات من أنواع محددة في منطقة محددة، وهذه الأعداد تتبع التوزيع البواسوني، وهو توزيع خاص بالحوادث النادرة الوقوع، والذي يتصف بأن تباينه يساوي متوسطه، أي أن $\left(\sigma^2 = \mu\right)$، وخلاصة القول، يستخدم

تحويلة الجذر التربيعي للبيانات إذا كانت تتبع توزيع بواسوني، كما يلاحظ أنه إذا احتوت البيانات على قيم صغيرة جدا (أقل من 10 أو حتي 15) ، يتجه الجذر التربيعي $\sqrt{y}$ إلى أعلى من الصحيح، وخاصة إذا كان هناك أصفار، لذا يقترح استخدام $\sqrt{y+0.5}$ بدلا من $\sqrt{y}$.

- **التحويلة اللوغاريثمية** $Log_{10}(y)$

إذا كان التباين يتناسب مع مربع متوسط المعالجة، أو أن الانحراف المعياري يتناسب مع متوسط المعالجة تستخدم التحويلة اللوغاريتمية $Log_{10}(y)$، وإذا كانت القيم صغيرة أو أصفار يستخدم التحويلة $Log_{10}(y+1)$.

- **التحويل الزاوي** $\arcsin\left(\sqrt{y}\right)=\sin^{-1}\left(\sqrt{y}\right)$

إذا كانت البيانات على شكل نسب مئوية يستخدم التحويل الزاوي.