١

تحليل التغاير

Analysis of Co_Variance

مقدمـــة

يهتم تحليل التغاير بدراسة تحليل أثر متغيرات مختلفة النوعية (وصفية، وكمية) ، فهو يجمع بين تحليلين هما: تحليل التباين، وتحليل الانحدار وفيما يلي الفرق بين كل تحليل:

Analysis	Response Variable	Independent Variable	Model
Analysis of Variance تحليل التباين	y	Qualitative Variables (Treatments)	$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$
Regression Analysis تحليل الانحدار	y	Quantitative Variables (Weights , Volumes)	$y_i = \alpha_0 + \alpha_1 x + \varepsilon_i$
Analysis of Covariance تحليل التغاير	y	Quantitative and Qualitative	$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \overline{X}..) + \varepsilon_{ij}$

والغرض من هذا التحليل في مجال تصميم وتحليل التجارب، هو دراسة، وتحليل أثر المعالجات بعد استبعاد أثر المتغير المستقل الكمي، فإذا نفذت التجربة وفقا لتصميم معين ، وقيست الاستجابة y من ناحية ومتغيرات مستقلة Covariates من ناحية أخرى مثل الظواهر التي يصعب على الباحث التحكم

فيها، ويرغب الباحث في التخلص من تأثيرها على الاستجابة y ، يجب عليه استخدام تحليل تحليل التغاير لعزل التأثيرات غير المرغوب فيها، ومثال على ذلك عن دراسة أثر برنامج معين للتغذية على الزيادة في وزن الجسم، فمن المعلوم قبل التجربة أن الزيادة في وزن الجسم تتأثر بوزن الجسم نفسه، ولدراسة تأثير البرنامج على الزيادة في الوزن لابد من فصل أثر الوزن على الزيادة، وإذا لم يستخدم تحليل التغاير، وتم استخدام تحليل التباين، تكون النتيجة هو زيادة الخطأ التجريبي ويصبح من الصعب اكتشاف الأثر الحقيقي للمعالجات.

نموذج تحليل التغاير وافتراضاته في حالة التصميم تام التعشية

إذا تم إجراء تجربة وفقا للتصميم التام التعشية، وقيست الاستجابة y_{ij}، وهي قيمة المشاهدة للوحدة التجريبية رقم j التي استلمت المعالجة رقم i، كما قيس أيضا المشاهدة x_{ij}، وتعبر عن قيمة المتغير المستقل للوحدة التجريبية رقم j التي استلمت المعالجة رقم i.

فإن التغير في الاستجابة y_{ij} يمكن التعبير عنه بنموذج رياضي خطي يأخذ الصورة التالية:

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \overline{X}..) + \varepsilon_{ij} \quad , \quad i = 1,2,...,t \quad , \quad j = 1,2,...r \qquad (1)$$

حيث أن :

μ **: متوسط عام** τ_i **: أثر المعالجة رقم** i

β **: أثر المتغير المستقل على المتغير y_{ij}، ويطلق عليه أحيانا بمعامل الانحدار regression coefficient أو الميل slop.**

ε_{ij} **: هو الخطأ العشوائي ، ويفترض أن :** $\varepsilon_{ij} \sim N(0,\sigma^2)$

والنموذج *(1)* هو نموذج خطي للتصميم تام التعشية، ويستند على الافتراضات التالية:

١- مجموع تأثيرات المعالجات يساوي صفر. $\sum_{i=1}^{t}\tau_i = 0$

٢- قيم المتغير المستقل محددة وثابتة **fixed** تقاس بدون خطأ، وهي مستقلة عن المعالجات.

٣- أن العلاقة بين x, y خطية.

٤- أن معامل الانحدار β متساوية لكل المعالجات.

الافتراض رقم (٢) من أهم الافتراضات في تحليل التغاير، لأنه إذا كان هناك علاقة بين المعالجات والمتغير المستقل لاختلفت β من معالجة لأخرى. ، ويعاد صياغة النموذج (١) في شكل آخر.

ومن الافتراضات السابقة، يمكن استنتاج الآتي:

$$\mu_{ij} = \mu + \tau_i + \beta(x_{ij} - \overline{X}_{..})$$

$$y_{ij} \sim N(\mu_{ij}, \sigma^2)$$

تحليل التغاير كتعديل لتحليل التباين في حالة تساوي التكرارات

في هذه الحالة يتم إتباع الخطوات التالية:

<u>الخطوة الأولى</u>: حساب مجموع المربعات لكل من (XY, X , Y)

<u>مجموع المربعات المجموع الكلي</u>

١- مجموع مربعات كلي خاص بـY

$$S_{yy} = \sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij} - \overline{Y}..)^2 = \sum_{i=1}^{t}\sum_{j=1}^{r} y_{ij}^2 - \frac{Y..^2}{tr}$$

٢- مجموع مربعات كلي خاص بـX

$$S_{xx}=\sum_{i=1}^{t}\sum_{j=1}^{r}(x_{ij}-\overline{X}..)^2=\sum_{i=1}^{t}\sum_{j=1}^{r}x_{ij}^2-\frac{X..^2}{tr}$$

٢- مجموع حاصل ضرب خاص بـXY

$$S_{xy}=\sum_{i=1}^{t}\sum_{j=1}^{r}(x_{ij}-\overline{X}..)(y_{ij}-\overline{Y}..)=\sum_{i=1}^{t}\sum_{j=1}^{r}x_{ij}y_{ij}-\frac{X..\,Y..}{tr}$$

<u>مجموع مربعات المعالجات</u>

١- مجموع مربعات المعالجات الخاص بـY

$$T_{yy}=\sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{y}_i.-\overline{Y}..)^2=\sum_{i=1}^{t}\frac{Y_{i.}^2}{r}-\frac{Y..^2}{tr}$$

٢- مجموع مربعات المعالجات الخاص بـX

$$T_{xx}=\sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{x}_i.-Y\overline{X}..)^2=\sum_{i=1}^{t}\frac{X_{i.}^2}{r}-\frac{X..^2}{tr}$$

٣- مجموع حاصل ضرب خاص بـXY

$$T_{xy}=\sum_{i=1}^{t}\sum_{j=1}^{r}(\overline{X}_{i.}-\overline{X}..)(\overline{Y}_{i.}-\overline{Y}..)=\sum_{i=1}^{t}\frac{X_{i.}Y_{i.}}{r}-\frac{X..\,Y..}{tr}$$

<u>مجموع مربعات الأخطاء</u>

١- مجموع مربعات الأخطاء الخاص بـY

$$E_{yy}=\sum_{i=1}^{t}\sum_{j=1}^{r}(y_{ij}-\overline{Y}_{i}.)^2=S_{yy}-T_{yy}$$

٢- مجموع مربعات الأخطاء الخاص بـX

$$E_{xx}=\sum_{i=1}^{t}\sum_{j=1}^{r}(x_{ij}-\overline{X}_{i}.)^2=S_{xx}-T_{xx}$$

٣-مجموع حاصل ضرب خاص بـXY

$$E_{xy}=\sum_{i=1}^{t}\sum_{j=1}^{r}(x_{ij}-\overline{X}_{i}.)(y_{ij}-\overline{Y}_{i}.)=S_{xy}-T_{xy}$$

<u>الخطوة الثانية</u> : تلخيص النتائج في جدول مكون من جزئين يسمى بجدول تحليل التغاير وهو:

جدول تحليل التغاير

S.O.V	df	Y	X	XY
Treatment	(t-1)	T_{yy}	T_{xx}	T_{xy}
Error	t(r-1)	E_{yy}	E_{xx}	E_{xy}
Total	tr-1	S_{yy}	S_{xx}	S_{xy}
	Adj. df	Adj. SS	Adj. MS	F
Treatment	t-1	Adj.SST	Adj.MST	F^*=Adj.MST/Adj.MSE
Error	t(r-1)-1	Adj.SSE	Adj.MSE	
Total	tr-2	Adj.SSTo		

ويتم حساب المجاميع ودرجات الحرية المعدلة علي النحو التالي:

Adj. SSTo

$$Adj.\ SSTo = S_{yy} - \frac{S_{xy}^2}{S_{xx}} \quad , \quad Adj.\ df = (tr-1)-1 = tr-2$$

Adj. SSE

$$Adj.\ SSE = E_{yy} - \frac{E_{xy}^2}{E_{xx}} \quad , \quad Adj.\ df = t(r-1)-1$$

Adj. SST

$$Adj.\ SST = Adj.\ SSTo - Adj.\ SSE \quad , \quad Adj.\ df = t-1$$

ويلاحظ أنه تم حذف درجة حرية واحدة لكل من الخطأ التجريبي ومجموع المربعات الكلي بسبب تقدير معامل الانحدار β ، كما يعبر S_{xy}^2 / S_{xx} عن مجموع مربعات الانحدار في النموذج $y_{ij} = \mu + \beta(x_{ij} - \bar{X}..) + \varepsilon_{ij}$ ، وهو ما تم استبعاد حتي يمكن معرفة مجموع المربعات الذي يعزي إلي المعالجات:

الاستدلال الإحصائي

<u>أولا</u>:- اختبارات الفروض

الفروض التي يمكن اختبارها في هذا التحليل هي:

١- اختبار فرض تساوي آثار المعالجات، وهو ما يهم الباحث في هذا التحليل، حيث يأخذ الفرض العدم والبديل الصورة التالية:

$$H_o: \tau_1 = \tau_2 = \dots = \tau_t = 0$$
$$H_a: at\ Least\ Two\ of\ \tau_i\ different$$

وإحصائية الاختبار المستخدمة في الاختبار هي:

$$F^* = \frac{Adj.MST}{Adj.MSE} \sim F_{\left((t-1),(t(r-1)-1\right)}$$

٢- اختبار فرض حول معامل الانحدار، حيث يأخذ الفرض العدم والبديل الصورة التالية:

$$H_o: \beta = 0 \ , \ H_a: \beta \neq 0$$

وإحصائية الاختبار المستخدمة في الاختبار هي:

$$F_\beta^* = \frac{\left(E_{xy}^2 / E_{xx}\right)}{Adj.MSE} \sim F_{\left(1,(t(r-1)-1\right)}$$

ثانيا:- تقدير فترات الثقة.

لتقدير فترة ثقة لمتوسط المعالجة أو الفرق بين الوسطين يجب أولا إيجاد الوسط الحسابي المعدل، وهو:

$$\overline{Y}_{i.Adj} = \overline{Y}_{i.} - \hat{\beta}(\overline{X}_{i.} - \overline{X}_{..})$$
$$\hat{\beta} = E_{xy} / E_{xx}$$

ويكون هذا التقدير غير متحيز للمتوسط

$$E(\bar{Y}_{i.}) = \mu + \tau_i + \beta(\bar{X}_{i.} - \bar{X}_{..})$$

كما أن الخطأ القياسي للمتوسط المعدل هو:

$$S^2_{\bar{Y}_{i.Adj}} = S^2_{\bar{Y}_{i.}} + S^2_{\hat{\beta}}(\bar{X}_{i.} - \bar{X}_{..})^2$$

$$= Adj.MSE\left(\frac{1}{r} + \frac{(\bar{X}_{i.} - \bar{X}_{..})^2}{E_{xx}}\right)$$

١- فترة ثقة لمتوسط المعالجة i هي:

$$\bar{Y}_{i.Adj} \pm t_{\left((1-\alpha/2),\ (t(r-1)-1)\right)}\sqrt{S^2_{\bar{Y}_{i.Adj}}}$$

٢- فترة ثقة للفرق بين متوسطي معالجتين هما:

$$(\bar{Y}_{i.Adj} - \bar{Y}_{i'.Adj}) \pm t_{\left((1-\alpha/2),\ (t(r-1)-1)\right)}\sqrt{S^2_{\bar{Y}_{i.Adj}-\bar{Y}_{i.Adj}}} \quad ,$$

$$S^2_{\bar{Y}_{i.Adj}-\bar{Y}_{i'.Adj}} = Adj.MSE\left(\frac{2}{r} + \frac{(\bar{X}_{i.} - \bar{X}_{i'.})^2}{E_{xx}}\right)$$

تطبيـــق

فيما يلي بيانات تجربة وفقا لتصميم تام التعشية، بخصوص برنامج لزيادة وزن الجسم، حيث تم استخدام 3 طرق للتغذية ، وقبل تنفيذ البرنامج تم قياس الأوزان، حيث جربت كل طريقة على 5 أشخاص، وبعد

تنفيذ البرنـامـج تم تسجيل الزيادة في الوزن وكانت البيانات كالتالي:

i	طرق التغذية					
	(1)		(2)		(3)	
	الزيادة في الوزن	الوزن قبل	الزيادة في الوزن	الوزن قبل	الزيادة في الوزن	الوزن قبل
1	10.4	47.84	14.0	55.00	11.5	57.00
2	10.0	44.00	15.0	62.1	12.3	56.58
3	12.5	57.5.0	17.6	80.96	10.8	49.68
4	13.3	61.18	13.0	51.98	10.5	48.30
5	9.7	44.62	15.7	72.22	9.5	43.70

والمطلوب

١- تكوين جدول تحليل تباين معدل.

٢- بعد استبعاد أثر الوزن هل يوجد فرق معنوي بين آثار الطرق؟

٣- اختبر معنوية معامل الانحدار.

٤- قدر فترة ثقة 95% للفرق بين متوسط الطريقة الأولي والثانية.

الحـــل

١- تكوين جدول تحليل التغاير

أولا:- تكوين جدول لحساب مجموع المربعات لكل من xy , x , y

$$S_{yy}=\sum_{i=1}^{3}\sum_{j=1}^{5}y_{ij}^2-\frac{Y..^2}{tr}=2382.32-\frac{185.8^2}{15}=2382.32-2301.44$$
$$=80.88$$

	y	x	xy
T1	**10.4**	**47.84**	**497.536**
	10.0	**44.00**	**440.000**
	12.5	**57.5.0**	**718.750**
	13.3	**61.18**	**813.694**
	9.7	**44.62**	**432.814**
sum	$Y_1.$	$X_1.$	
	55.9	**255.14**	
T2	**14.0**	**55.00**	**770.000**
	15.0	**62.10**	**931.500**
	17.6	**80.96**	**1424.896**
	13.0	**51.98**	**675.740**
	15.7	**72.22**	**1133.854**
Sum	$Y_2.$	$X_2.$	
	75.3	**322.26**	
T3	**11.5**	**57.00**	**655.500**
	12.3	**56.58**	**695.934**
	10.8	**49.68**	**536.544**
	10.5	**48.30**	**507.150**
	9.5	**43.70**	**415.150**
Sum	$Y_3.$	$X_3.$	
	54.6	**255.26**	

	$\sum_{i=1}^{3}\sum_{j=1}^{5} y_{ij}^2$	$\sum_{i=1}^{3}\sum_{j=1}^{5} x_{ij}^2$	$\sum_{i=1}^{3}\sum_{j=1}^{5} x_{ij}y_{ij}$
	2382.32	**47779.4116**	**10649.062**

$$S_{xx} = \sum_{i=1}^{3}\sum_{j=1}^{5} x_{ij}^2 - \frac{X_{..}^2}{tr} = 47779.41 - \frac{(832.66)^2}{15}$$
$$= 47779.41 - 46221.51 = 1557.90$$

$$S_{xy} = \sum_{i=1}^{3}\sum_{j=1}^{5} x_{ij}y_{ij} - \frac{X_{..}Y_{..}}{tr} = 10649.06 - \frac{(832.66)(185.8)}{15} = 335.18$$

$$T_{yy} = \sum_{i=1}^{t}\frac{Y_{i.}^2}{r} - \frac{Y_{..}^2}{tr} = \frac{1}{5}\left(55.92^2 + 75.3^2 + 54.6^2\right) - 2301.44 = 54.22$$

$$T_{xx} = \sum_{i=1}^{t}\frac{X_{i.}^2}{r} - \frac{X_{..}^2}{tr} = \frac{1}{5}\left(255.142^2 + 322.26^2 + 255.26^2\right) - 46221.15$$
$$= 600.17$$

$$T_{xy} = \sum_{i=1}^{t}\frac{X_{i.}Y_{i.}}{r} - \frac{X_{..}\,Y_{..}}{tr} = \frac{1}{5}\left(55.9\times 255.14 + 75.3\times 322.26 + 54.6\times 255.26\right)$$
$$- \frac{832.66\times 185.8}{15} = 10493.14 - 10313.88 = 179.26$$

$$E_{yy} = S_{yy} - T_{yy} = 80.88 - 54.22 = 26.66$$

$$E_{xx} = S_{xx} - T_{xx} = 1557.9 - 600.17 = 957.73$$

$$E_{xy} = S_{xy} - T_{xy} = 335.18 - 179.26 = 155.92$$

٢- حساب المجاميع المعدلة

$$Adj.\ SSTo = S_{yy} - \frac{S_{xy}^2}{S_{xx}} = 80.88 - \frac{335.18^2}{1557.9} = 8.77$$

$$Adj.\ df = (tr-1)-1 = 15-2 = 13$$

$$Adj.\ SSE = E_{yy} - \frac{E_{xy}^2}{E_{xx}} = 26.66 - \frac{155.92^2}{957.73} = 1.276$$

$$Adj.\ df = t(r-1)-1 = 12-1 = 11$$

$$Adj.\ SST = Adj.\ SSTo - Adj.\ SSE = 8.77 - 1.276 = 7.494$$

$$Adj.\ df = t-1 = 2$$

٣- تكوين الجدول المعدل

S.O.V	df	Y	X	XY
Treatment	2	54.22	600.17	179.26
Error	12	26.66	957.73	155.92
Total	14	80.88	1557.9	335.18
	Adj. df	Adj. SS	Adj. MS	F
Treatment	2	7.494	3.747	$F^{*}=32.302$
Error	11	1.276	0.116	
Total	13	8.770		