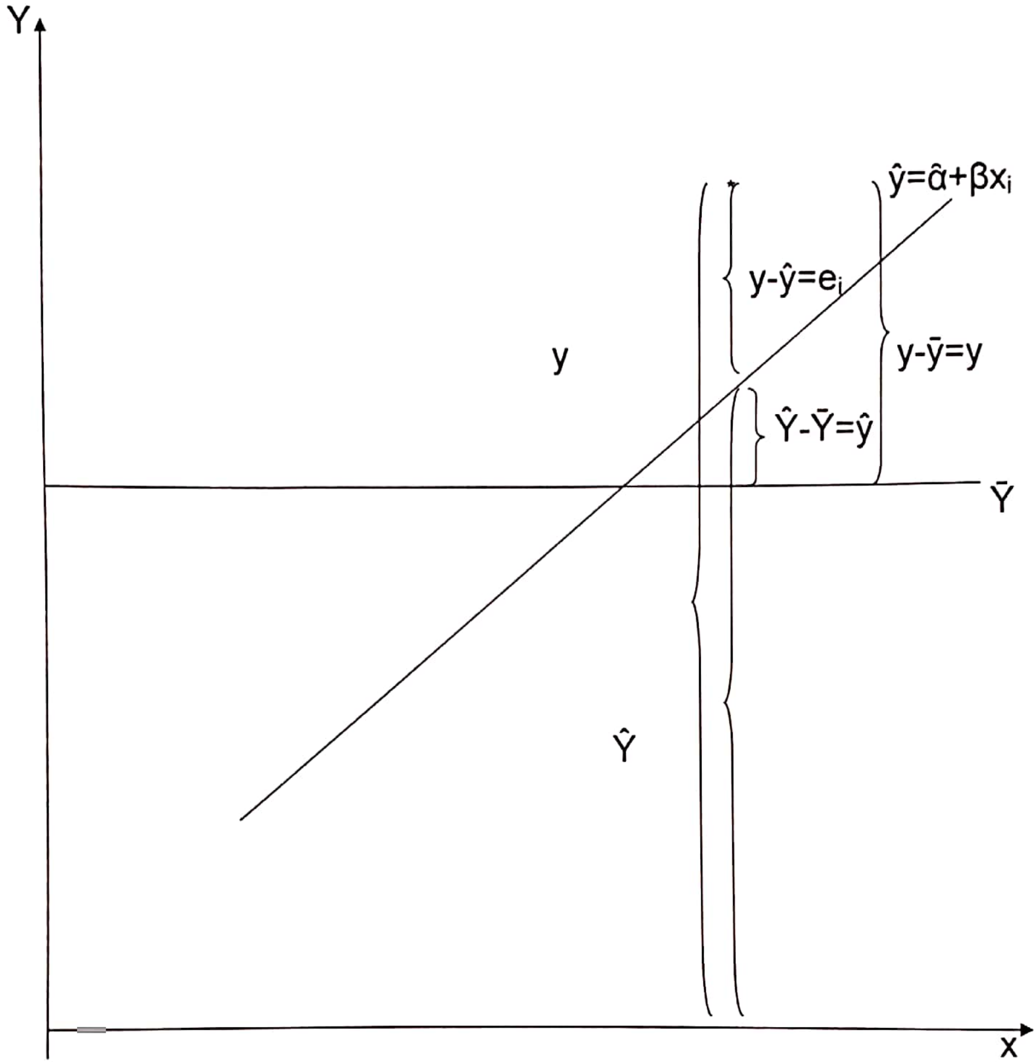
المحاضرة السادسة

تحليل التباين

و هذه العلاقة الاخيرة تسمى بمعادلة تحليل التباين و هي الاساس في معالجة تحليل التباين مجاله متغيرين اثنين.



ملاحظة: نقول ان $\sum y_i^2$ يمثل التباين الاجمالي لقيم y حول وسط العينة و يمكن ان يجزا الى جزئين.

الجزء الاول: $\sum \hat{y}_i^2=\sum(\hat{y}-\bar{y})$ هو تباين قيم $\hat{y}$ حول وسطه و هو غالبا ما يفسر او يشار اليه على انه مجموع المربعات المفسرة من قبل التاثير الخطي لـ x على y.

و كلما كانت $\sum y_i^2$ كبيرة كلما دل على ان العلاقة قوية بين x و y.

الجزء الثاني: Σe_i^2 هو التباين المتبقي او غير المفسر لقيم y حول مستقيم المربعات الصغرى اي غير المفسر للتاثير الخطي لـx على y. و كلما كان هذا الجزء كبير كلما دل على عدم وجود علاقة بين x و y.

ـو من معادلة تحليل التباين

$$\frac{\Sigma\hat{y}^2}{\Sigma y^2} = \frac{\Sigma(\hat{\beta}x_i)^2}{\Sigma y^2} = \hat{\beta}^{2\frac{\Sigma x_i^2}{\Sigma y^2}} = r^2$$

لا سبق بينا

$$r^2=\beta\frac{\Sigma xy}{\Sigma y^2}. \qquad \left(\frac{\Sigma xy}{\Sigma x^2}\right)\left(\frac{\Sigma xy}{\cancel{\Sigma x^2}}\right)\frac{\cancel{\Sigma x^2}}{\Sigma y^2} = \hat{\beta} = \frac{\Sigma xy}{\Sigma y^2}$$

اي: $r^2=\frac{\Sigma\hat{y}^2}{\Sigma y}$

تسمى هذه النسبة بمعامل التحديد coefficient de détermination

و هي عبارة عن مربع معامل الارتباط

لان معامل الارتباط $\text{л}=\beta\frac{sx}{sy}$

$$\text{л}^2=\beta\frac{\Sigma x^2/n}{\Sigma y^2/n} = \beta^2\frac{\Sigma {x_i}^2}{\Sigma y^2}$$

و يمكن صياغة معامل التحديد على شكل اخر.

$$\text{л}^2=\frac{\Sigma\hat{y}^2}{\Sigma y^2} = \frac{\Sigma y^2-\Sigma e_i^2}{\Sigma y^2} = -1\,\frac{\Sigma e_i^2}{\Sigma y^2}$$

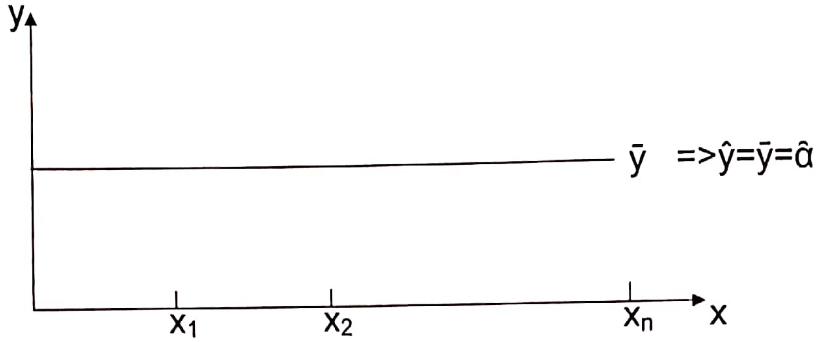
نتيجة: $max : r^2=1 \Rightarrow max : r=\pm 1$

ملاحظة: عندما تاخذ л^2 القيمة العظمى و هي الواحد الصحيح فان هذا يؤدي الى ان $\Sigma e_i^2=0$ و هذا يعني $\Sigma(y_i-\hat{y})^2=0$ و هذا يعني يجب ان يكون $y_i=\hat{y}$ و هذا يعني $y_i-\hat{y}=0$ و هذا يعني $e_i=0$ و هذا يعني ان جميع النقاط واقعة تماما على مستقيم المربعات الصغرى و هذا يعني ان العلاقة بين x و y هي علاقة تامة.

*عندما تاخذ r^2 القيمة الدنيا و هي الصفر و هذا لا يكون الا اذا كانت النسبة $\frac{\Sigma e_i^2}{\Sigma y^2}=1$

اي $\Sigma e_i^2=\Sigma y^2$

و هذا يعني ان التباين الاجمالي بكامله لا يمكن تفسيره بواسطة التاثير الخطي لـ$\bar{x}$على y اي ان التباين الاجمالي مصدره ليست العلاقة بين x وy و بالتالي ليست هناك اية علاقة بين xوy و في هذه الحالة مستقيم المربعات الصغرى ينطبق على المستقيم $\bar{y}$



اختبار تحليل التباين: test de l'analyse de la variance:-

لدينا تعريفا:-اذا كان γ_1^2 متحول عشوائي يخضع لتوزيع x^2 بـ r_1 درجة حرية و كان γ_2^2 متحول عشوائي اخر مستقل عن الاول و يخضع لتوزيع x^2 بـr_2 درجة حرية فان النسبة F:

$$F=\frac{\gamma_1^2/r_1}{\gamma_2^2/r_2}$$

يخضع لتوزيع فيشر بـ(r_1,r_2) درجة حرية اي بـ r_1 درجة للبسط و r_2 درجة حرية للمقام.

لدينا: $\beta \sim N(\beta,\frac{\delta_u^2}{\Sigma x^2})$

و منه: $\frac{\hat{\beta}-\beta}{\delta_u/\sqrt{\Sigma x^2}} \sim N(0,1)$

و منه: $\frac{(\hat{\beta}-\beta)^2}{{\delta_u}^2/\Sigma x^2} \sim x_1^2$ اي بواحد درجة حرية.

و لدينا كذلك: $\frac{\Sigma e_i^2}{\delta_u^2} \sim \chi^2_{n-2}$

و منه:

$$F=\frac{\frac{(\hat{\beta}-\beta)^2}{\delta u^2/\Sigma x^2}/n_1}{\frac{\Sigma e_i^2}{\delta_u^2}/n_2}=\frac{\frac{(\hat{\beta}-\beta)^2}{\delta u^2/\Sigma x^2}}{\frac{\Sigma e_i^2}{\delta_u^2}/n-2}\sim F(1,n-2)$$

و منه:

$$F=\frac{\frac{(\hat{\beta}-\beta)^2\Sigma x^2}{\delta_u^2}}{\frac{\Sigma e_i^2}{\delta_u^2(n-2)}}$$

$$=\frac{(\hat{\beta}-\beta)^2\Sigma x^2}{\frac{\Sigma e_i^2}{n-2}}$$

$$F=\frac{(\hat{\beta}-\beta)^2\Sigma x^2}{\frac{\Sigma e_i^2}{n-2}}\sim F(1,n-2)$$

حيث F متحول عشوائي يخضع لتوزيع فيشر بـ (1,n-2) درجة حرية للبسط و المقام على التوالي.

ملاحظة: اذا كانت $\beta=0$(و هي فرضية اللعلاقة بين xو y) فان النسبة F تتقلص الى:

$$F=\frac{\hat{\beta}^2\Sigma x^2}{\frac{\Sigma e_i^2}{n-2}}$$

اذن الصورة تساوي التباين المفسر لان :-

$$\hat{y}=\beta x$$

$$\hat{y}^2=\beta^2x^2$$

$$\Sigma\hat{y}^2=\beta^2\Sigma x^2$$

و نرمز له بـ SCR (somme des carres résiduelle)او (somme des carres non-expliquée)و منه:-

$$F=\frac{SCE/1}{SCR/n-2}$$

ملاحظة: هذه الصيغة ليست صحيحة الا في حالة اللعلاقة بين xوy.

*و هذا الاختبار هو اختبار فيشر في حالة فرضية اللاعلاقة (H_0 :β=0)

اذا كان Fc>Ft نرفض فرضية اللاعلاقة و بالتالي هو علاقة بينxوy ذات دلالة.اما اذا كان Fc<Ft فاننا نقبل فرضية العدم اي عدم وجود علاقة بين xوy. و يتم عادة حساب F في جدول يسمى جدول تحليل التباين.

ملاحظة: قبل الجدول يجب ذكر كيفية حساب SCE و SCR

مصدر التباين	مجموع المربعات	درجات الحرية	وسطي المربع
X المتبقي	$SCE=\Sigma\hat{y}^2=\beta^2\Sigma x^2$ $=\beta\Sigma xy$	1	SCE/1
	$SCR=\Sigma e_i^2$	n-2	SCR/n-2
المجموع	$SCT=\Sigma y^2$	n-1	

و قيمة F المناسبة هي نسبة الاوساط المربعة في العمود الاخير من هذا الجدول.

اولا: حساب التباين المفسر:- يمكن حساب هذا التباين بطرق مختلفة

$$SCE=\beta^2\Sigma x^2=\pi^2\Sigma y^2=\beta\Sigma xy$$

لدينا $\beta=\frac{\Sigma xy}{\Sigma x^2}$

و منه: $\beta\Sigma x^2=\frac{(\Sigma xy)^2}{(\Sigma x^2)^2}.\Sigma x^2$

$$=\frac{(\Sigma xy)^2}{\Sigma x^2}$$

و منه: $=(\frac{\Sigma xy}{\Sigma x^2})\Sigma xy$

$$=\beta\Sigma xy$$

ثانيا: حساب التباين الغير مفسر: اما التباين غير المفسر فيتم حسابه كما يلي:-

$SCR=\Sigma y^2-SCE.$

اختبار معامل الارتباط:-

لدينا: $SCE=r^2SCT=r^2\Sigma y^2$

و لدينا: $SCR=(1-r^2)SCT$

حيث: $F=\frac{SCE/1}{SCR/n-2}$

و منه:- $F=\frac{r^2/1}{(1-r^2)n-2}$

كما يمكننا ان نكتب $t=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

لانه لدينا سابقا $t=\frac{\beta\sqrt{\Sigma x^2}}{\delta_u}$

و لدينا سابقا $r^2=\beta^2\frac{sx^2}{sy^2}$

و منه: $r=\beta\frac{sx}{sy}$

$$r=\beta\frac{\sqrt{\Sigma x^2/n}}{\sqrt{\Sigma y^2/n}}$$

$$r=\beta\frac{\sqrt{\Sigma x^2}}{\sqrt{\Sigma y^2}} => \hat{\beta}\sqrt{\Sigma x^2} = r\sqrt{\Sigma y^2} - (1)$$

و لدينا من العلاقة $SCR=(1-r^2)SCT=(1-r^2)\Sigma y^2$

و منه:- $(1-r^2)=\frac{\Sigma e_i^2}{\Sigma y^2}$

و منه:- $\sqrt{1-r^2} = \sqrt{\frac{\Sigma e_i^2}{\Sigma y^2}} = \frac{\sqrt{\Sigma e_i^2}}{\sqrt{\Sigma y^2}} => \sqrt{\Sigma e_i^2} = \sqrt{1-r^2}.\sqrt{\Sigma y^2}$

و لدينا: $\delta_u=\sqrt{\frac{\Sigma e_i^2}{n-2}}=\frac{\sqrt{\Sigma e_i^2}}{\sqrt{n-2}}--(1)$

و بالتعويض بـ (1) في δ_u نجد:-

$$\delta_u=\frac{\sqrt{1-r^2}\sqrt{\Sigma y^2}}{\sqrt{n-2}}$$

و بالتعويض بقيمة δ_u و قيمة $r\sqrt{\Sigma y^2}=\beta\sqrt{\Sigma x^2}$ في t نجد:-

$$t=\frac{\beta\sqrt{\Sigma x^2}}{\delta_u}=\frac{r\sqrt{\Sigma y^2}}{\frac{\sqrt{1-r^2}\sqrt{\Sigma y^2}}{\sqrt{n-2}}}$$

نتيجة:نلاحظ ان التابع الاختباري الخاص بمعامل الارتباط هو تماما المقدار t الذي يمكن ان نختبر به الفرضية القائلة $H_0:\beta=0$

$$t=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

ملاحظة: تحديد العلاقة بين tوF

لدينا: $t=\frac{\beta\sqrt{\Sigma x^2}}{\delta_u}$

و بتربيع الطرفين $t^2=\beta^2\frac{\Sigma x^2}{\delta_u^2}$

حيث: $\beta^2\Sigma x^2=SCE$

$$\delta_u^2=\frac{\Sigma e_i^1}{n-2}$$

$$\delta_u^2=\frac{SCR}{n-2}$$

و منه و بالتعويض:- $t^2=\frac{SCE}{SCR}=F$

النتائج المسجلة: العلاقة بين الاختبارات الثلاثة استيودنت(معامل الانحدار)، فيشر(تحليل التباين)، معامل الارتباط.