Dummy Variables

ASST.PROF.DR.AHMED.H.AL-HUSAYNY

المتغيرات الصورية
(Dummy Variables)

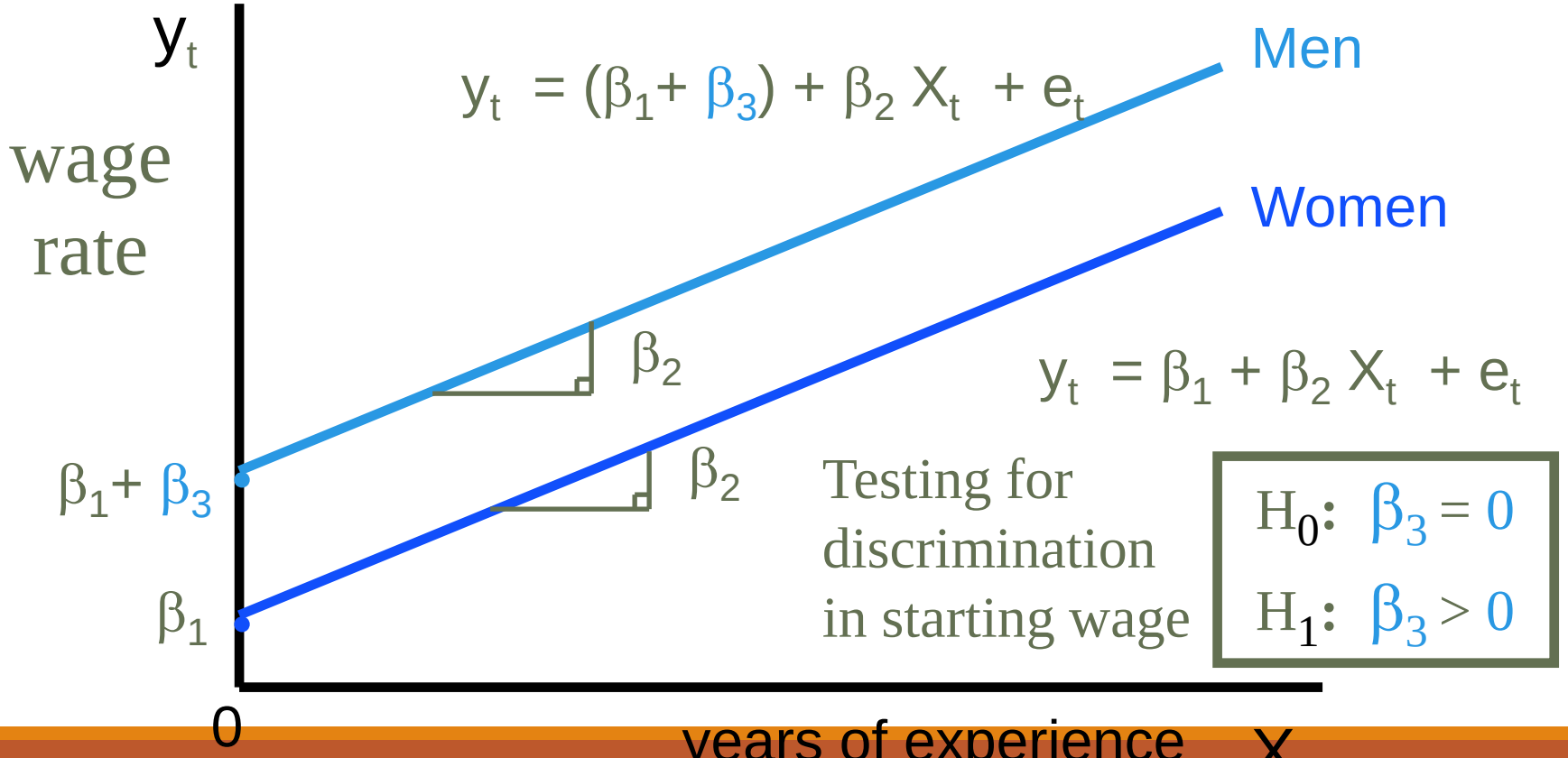
المتغيرات الصورية هي تلك المتغيرات التي تعبر عن صفات معينة مثل اللون ، والديانة ، والجنسية ، والجنس أو النوع ، والحروب ، والفقر ، والزلازل.

تسمى المتغيرات الصورية بالمتغيرات الكيفية Qualitative Variables .

تستخدم القيمة واحد صحيح (1) للدلالة على وجود صفة معينة و القيمة صفر (0) للدلالة على عدم وجود هذه الصفة.

$$y_t = \beta_1 + \beta_2 X_t + \beta_3 D_t + e_t$$

For men: $D_t = 1$.
For women: $D_t = 0$.



تقدير معاملات الانحدار في حالة اشتمال النموذج على متغيرات مستقلة صورية

1. تقدير معاملات الانحدار في حالة اشتمال النموذج على متغير مستقل صوري واحد:

مثال تطبيقي:

بفرض أن نموذج الانحدار الخطي المتعدد المراد تقديره كان كما يلي:

$$Y_i = \alpha_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \alpha_1 D_{1i} + \varepsilon_i$$

$$i = 1, 2, \ldots, N$$

حيث أن:

Y = دخل الفرد

X_1 = عدد السكان

X_2 = عدد سنوات التعليم

D_1 = النوع أو الجنس (1 إذا كان النوع أنثى ، 0 إذا كان النوع ذكر)

α_0 = الحد الثابت

β_1 = ميل العلاقة بين X_1 و Y وهو يقيس أثر تغير السكان بوحدة واحدة على دخل الفرد

β_2 = ميل العلاقة بين X_2 و Y وهو يقيس أثر تغير سنوات التعليم بوحدة واحدة على دخل الفرد

α_1 = الحد الثابت التفاضلي وهو يقيس أثر النوع أو الجنس على دخل الفرد

Y_i	X_{1i}	X_{2i}	D_{1i}
5000	80	9	1
6000	95	8	1
7000	100	10	1
8000	101	10	1
9000	103	11	1
10000	115	14	1
11000	105	15	1
12000	116	13	0
13000	120	16	0
14000	110	17	0

شكل دالة الدخل المقدرة في حالة إذا كان النوع ذكراً يكون كما يلي:

أما شكل دالة الدخل ال $\hat{Y}_i = \hat{\alpha}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + e_i$

وبتطبيق طريقة المربعات الصغرى العادية نحصل على شكل دالة الدخل التالية:

$$\hat{Y}_i = (\hat{\alpha}_0 + \hat{\alpha}_1) + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + e_i$$

$$\hat{Y}_i = \hat{\alpha}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \hat{\alpha}_1 D_{1i} + e_i$$

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.98
R Square	0.97
Adjusted R Square	0.95
Standard Error	674.94
Observations	10

ANOVA					
	df	*SS*	*MS*	*F*	*Sig F*
Regression	**3**	**79766713.89**	**26588904.63**	**58.37**	**7.8555 3E-05**
Residual	**6**	**2733286.11**	**455547.68**		
Total	**9**	**82500000**			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	**-3327.01**	**2908.33**	**-1.14**	**0.30**	**-10443.43**	**3789.41**
X1i	**66.12**	**30.83**	**2.14**	**0.08**	**-9.32**	**141.56**
X2i	**567.46**	**119.85**	**4.73**	**0.00**	**274.20**	**860.73**
D1i	**-1517.70**	**649.46**	**-2.34**	**0.06**	**-3106.87**	**71.47**

معادلة الدخل المقدرة

$$\hat{Y}_i = -3327.53 + 66.12X_{1i} + 567.46X_{2i} - 1517.61D_{1i}$$

معادلة الدخل المقدرة في حالة إذا كان النوع ذكراً ($D_{1i}=0$)

معادلة الدخل المقدرة

$$\hat{Y}_i = -3327.53 + 66.12X_{1i} + 567.46X_{2i}$$

$$\hat{Y}_i = -4845.14 + 66.12X_{1i} + 567.46X_{2i}$$

تقدير معاملات الانحدار في حالة اشتمال النموذج على متغيرات مستقلة صورية

2. تقدير معاملات الإنحدار في حالة اشتمال النموذج على متغيرين مستقلين صوريين :

- بفرض أن نموذج الإنحدار الخطي المتعدد المراد تقديره كان كما يلي :

$$Y_i = \alpha_0 + \alpha_1 D_{1i} + \alpha_2 D_{2i} + \varepsilon_i$$

$$i = 1, 2, \ldots, N$$

- حيث أن:

Y = دخل الفرد

D_1 = درجة الدكتوراة (1 درجة الدكتوراة ، 0 درجة الماجستير ، 0 درجة البكالوريوس)

D_2 = درجة الماجستير (1 درجة الماجستير ، 0 درجة الدكتوراة ، 0 درجة البكالوريوس)

α_0 = الحد الثابت

α_1 = حد ثابت تفاضلي يقيس أثر الحصول على درجة الدكتوراة على متوسط الدخل

α_2 = حد ثابت تفاضلي يقيس أثر الحصول على درجة الماجستير على متوسط الدخل

المطلوب تقدير المعادلة مع إيجاد متوسط دخل الفرد بعد الحصول على كل درجة علمية على حدة

Y_i	D_{1i}	D_{i2}
5000	0	0
6000	0	0
7000	0	0
8000	0	1
9000	0	1
10000	0	1
11000	0	0
12000	1	0
13000	1	0
14000	1	0

$\hat{Y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 D_{1i} + \hat{\alpha}_2 D_{2i} + e_i$

- وبتطبيق طريقة المربعات الصغرى العادية نحصل على شكل دالة الدخل التالية:

$$\hat{Y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 D_{1i} + \hat{\alpha}_2 D_{2i} + e_i$$

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.84
R Square	0.70
Adjusted R Square	0.61
Standard Error	1880.35
Observations	10

ANOVA					
	df	*SS*	*MS*	*F*	*Sig. F*
Regression	2	57750000	28875000	8.17	0.015
Residual	7	24750000	3535714		
Total	9	82500000			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	7250	940.17	7.71	0.00	5026.84	9473.16
D_{1i}	5750	1436.14	4.00	0.01	2354.07	9145.93
D_{2i}	1750	1436.14	1.22	0.26	-1645.93	5145.93

- بتطبيق طريقة المربعات الصغرى ينتج ما يلي :

$$\hat{Y}_i = 7250 + 5750D_{1i} + 1750D_{2i}$$

- ومن ثم يمكن تحديد ما يلي :
- متوسط دخل الحاصلين على درجة البكالوريوس $(D_{1i} = D_{2i} = 0)$

$$\hat{Y}_i = \hat{\alpha}_0 = 7250$$

- متوسط دخل الحاصلين على درجة الماجستير $(D_{1i} = 0)$

$$\hat{Y}_i = \hat{\alpha}_0 + \hat{\alpha}_2 = 7250 + 1750 = 9000$$

- متوسط دخل الحاصلين على درجة الدكتوراة $(D_{2i} = 0)$

$$\hat{Y}_i = \hat{\alpha}_0 + \hat{\alpha}_1 = 7250 + 5750 = 13000$$

تقدير معاملات الانحدار في حالة اشتمال النموذج على متغيرات مستقلة صورية

3. تقدير معاملات الإنحدار في حالة اشتمال النموذج على ثلاثة متغيرات (مستقلة , تابعة) صورية :

- بفرض أن نموذج الإنحدار المراد تقديره كان كما يلي :

$$D_{1i} = \alpha_0 + \alpha_1 D_{2i} + \alpha_2 D_{3i} + \varepsilon_i$$

$$i = 1, 2, \ldots, N$$

- حيث أن:

D_1 = الفقر (1 للفقر ، 0 خلاف ذلك)

D_2 = الجنس أو النوع (1 للأنثى ، 0 للذكر)

D_3 = اللون (1 للأسود ، 0 للأبيض)

α_0 = الحد الثابت

α_1 = معامل D_{2i} ، و يقيس أثر النوع على إحتمال حدوث الفقر

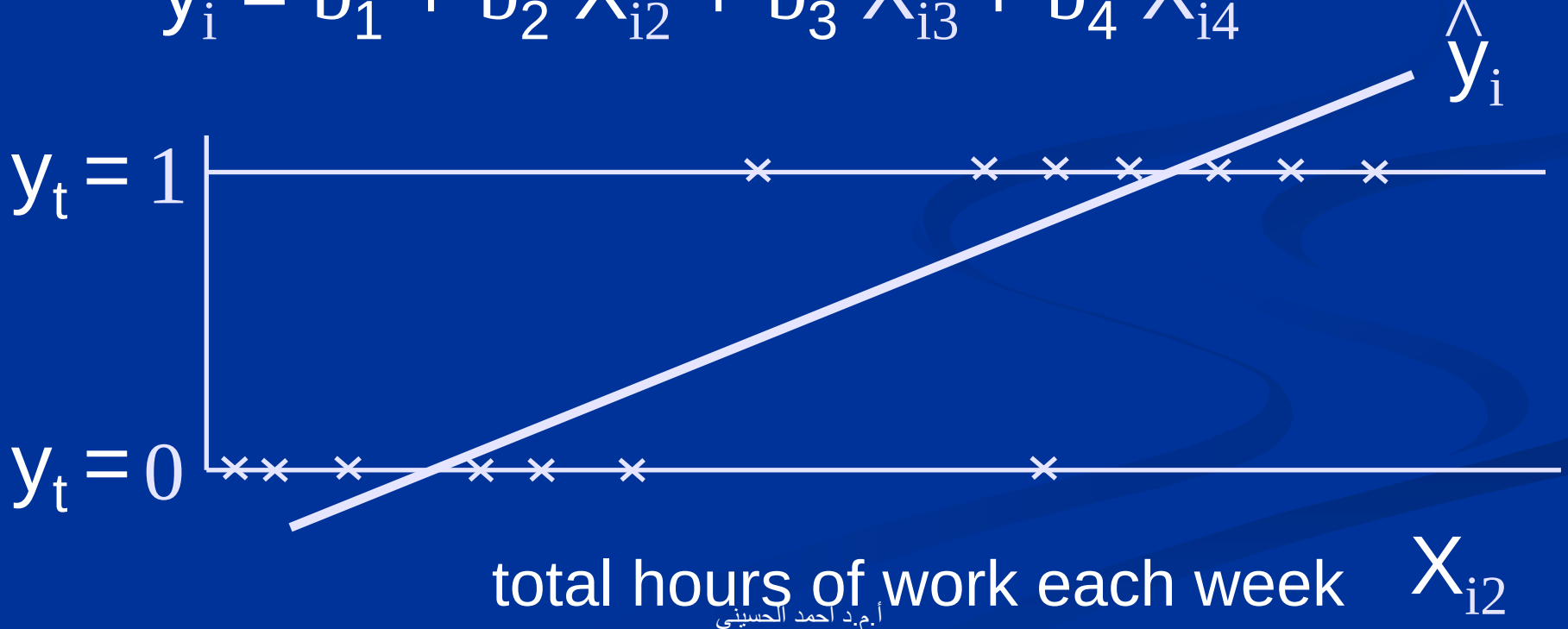
α_2 = معامل D_{3i} ، و يقيس أثر اللون على إحتمال حدوث الفقر

Linear Probability Model

$$y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + e_i$$

Read predicted values of y_i off the regression line:

$$\hat{y}_i = b_1 + b_2 X_{i2} + b_3 X_{i3} + b_4 X_{i4}$$



المطلوب

1. تقدير معاملات الإنحدار للمعادلة
2. إيضاح أثر كل من النوع و اللون على إحتمال تحقق الفقر

D_{1i}	D_{2i}	D_{3i}
1	1	0
1	1	1
1	1	1
0	0	0
0	0	0
0	0	0
0	0	0
1	1	1
0	0	1
0	0	1
0	1	0
0	0	0
1	1	1
1	1	1
0	0	0
0	0	0

D_{1i}	D_{2i}	D_{3i}
0	0	1
0	0	0
0	1	0
0	0	0
1	1	1
1	0	1
0	0	0
1	1	1
1	1	0
1	1	1
1	0	1
0	0	1
1	0	0
0	0	0

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.74
R Square	0.55
Adjusted R Square	0.51
Standard Error	0.35
Observations	30

ANOVA					
	df	*SS*	*MS*	*F*	*Sig. F*
Regression	2	4.02	2.01	16.19	0.000
Residual	27	3.35	0.12		
Total	29	7.37			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0.05	0.09	0.53	0.60	-0.14	0.24
D2i	0.55	0.14	3.96	0.00	0.26	0.84
D3i	0.35	0.14	2.57	0.02	0.07	0.63

1. تقدير معاملات الإنحدار للمعادلة:

- بتطبيق طريقة المربعات الصغرى العادية ينتج ما يلي :

$$\hat{D}_{1i} = 0.05 + 0.55D_{2i} + 0.35D_{3i}$$

- ويطلق على هذه المعادلة إصطلاح نموذج الإحتمال الخطي
- ويتضح من هذه المعادلة ما يلي :

- إن معامل D_{2i} يساوي 0.55 ، ويعني هذا إن إحتمال تحقق الفقر للسيدات أكبر من إحتمال تحقق الفقر للرجال بحوالي 0.55 .

- إن معامل D_{3i} يساوي 0.35 ، ويعني هذا إن إحتمال تحقق الفقر للفرد الأسود أكبر من إحتمال تحقق الفقر للفرد الأبيض بحوالي 0.35 .

2. إيضاح أثر كل من النوع واللون على إحتمال تحقق الفقر :

- إحتمال تحقق الفقر لرب العائلة ذو اللون الأبيض :

$$D_{1i} = 1, D_{2i} = D_{3i} = 0$$

$$\hat{D}_{1i} = \hat{\alpha}_0 = 0.05$$

- إحتمال تحقق الفقر لربة العائلة ذات اللون الأبيض :

$$D_{1i} = 1, D_{2i} = 1, D_{3i} = 0$$

$$\hat{D}_{1i} = \hat{\alpha}_0 + \hat{\alpha}_1 = 0.05 + 0.55 = 0.60$$

- إحتمال تحقق الفقر لربة العائلة ذات اللون الأسود :

$$\hat{D}_{1i} = 1, D_{2i} = 1, D_{3i} = 1$$

$$\hat{D}_{1i} = \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_3 = 0.05 + 0.55 + 0.35 = 0.95$$