

## V. Extensions of Multiple Regression

### A. Dummy Variables

- 1. Definition:** A *binary variable* that indicates a quality, condition, characteristic, etc. exists.
- 2. Examples:** Cross-Section; Time-Series.
- 3. Uses for Dummy Variables**
  - a. Shifting the Intercept
  - b. Shifting the Slope
- 4. Interpretation**

---

---

---

---

---

---

---

---

Regression Analysis: wage versus yrsed, exp, f, fexp

The regression equation is

$$\text{wage} = -4.66 + 1.49\text{yrsed} + 0.155\text{exp} - 1.03\text{f} - 0.048\text{fexp}$$

Predictor	Coef	SE Coef	T	P
Constant	-4.657	2.184	-2.13	0.033
yrsed	1.4850	0.1499	9.91	0.000
exp	0.15522	0.03771	4.12	0.000
f	-1.031	1.311	-0.79	0.432
fexp	-0.04778	0.05348	-0.89	0.372

Two regressions:

If  $f=0$ :

$$\text{wage} = -4.66 + 1.49\text{yrsed} + 0.155\text{exp}$$

If  $f=1$ :

$$\text{wage} = (-4.66 - 1.03) + 1.49\text{yrsed} + (0.155 - 0.048)\text{exp}$$

---

---

---

---

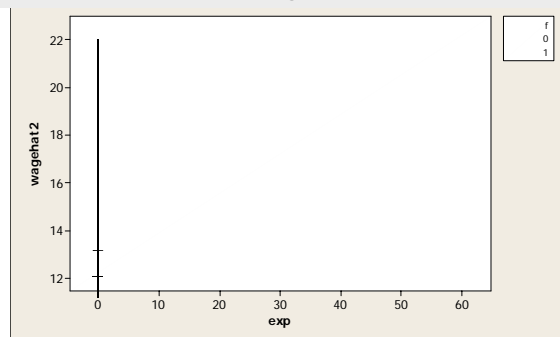
---

---

---

---

### Shifting the Intercept and Slope: Male vs. Female Wages



---

---

---

---

---

---

---

---

#### 4. Interpretations

- Model:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \delta D_i + \gamma (X_{1i} D_i) + u_i$$

- The parameter  $\delta$  :

---

---

---

---

---

---

---

---

#### Regression Analysis: wage versus yrsed, exp, f, fexp

The regression equation is

$$\text{wage} = -4.66 + 1.49\text{yrsed} + 0.155\text{exp} - 1.03\text{f} - 0.048\text{fexp}$$

Predictor	Coef	SE Coef	T	P
Constant	-4.657	2.184	-2.13	0.033
yrsed	1.4850	0.1499	9.91	0.000
exp	0.15522	0.03771	4.12	0.000
f	-1.031	1.311	-0.79	0.432
fexp	-0.04778	0.05348	-0.89	0.372

---

---

---

---

---

---

---

---

#### 4. Interpretations

- Model:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \delta D_i + \gamma (X_{1i} D_i) + u_i$$

- The parameter  $\gamma$  :

---

---

---

---

---

---

---

---

Interpretations – are the groups different?

- Are the intercepts different? **Statistically different?**
  - If  $\delta = 0$ :
  - Statistical test:
- Are the slopes different? **Statistically different?**
  - If  $\gamma = 0$ :
  - Statistical test:

---

---

---

---

---

---

---

---

The regression equation is  
sales = 17377 - 3865 prose + 2243 pearn - 32.3 dinc + 492 d2 +  
331 prosed2

Predictor	Coef	SE Coef	T	P
Constant	17377	5295	3.28	0.008
prose	-3864.6	530.5	-7.28	0.000
pearnc	2242.5	817.8	2.74	0.021
dinc	-32.26	24.68	-1.31	0.221
d2	492	4837	0.10	0.921
prosed2	331	1443	0.23	0.823

S = 846.3      R-Sq = 88.6%      R-Sq(adj) = 82.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	55433947	11086789	15.48	0.000
Residual Error	10	7162409	716241		
Total	15	62596356			

---

---

---

---

---

---

---

---

F-Test: Are the PRFs for the two groups different?

- The hypothesis:  
 $H_o : \delta = \gamma = 0$   
 $H_a : \text{either } \delta \text{ or } \gamma, \text{ or both, not zero.}$
- Calculate the F-statistic:

$$F_{calc} = \frac{[ESS_{(2)} - ESS_{(1)}] / (K_{(2)} - K_{(1)})}{RSS_{(2)} / (n - K_{(2)} - 1)}$$

---

---

---

---

---

---

---

---

**Regression Analysis: sales versus prose, pcarn, dinc**

The regression equation is  
sales = 13355 - 3628 prose + 2634 pcarn - 19.3 dinc

Predictor	Coef	SE Coef	T	P
Constant	13355	6485	2.06	0.062
prose	-3628.2	635.6	-5.71	0.000
pcarn	2634	1013	2.60	0.023
dinc	-19.25	30.69	-0.63	0.542

S = 1076.29 R-Sq = 77.8% R-Sq(adj) = 72.2%

**Analysis of Variance**

Source	DF	SS	MS	F	P
Regression	3	48695526	16231842	14.01	0.000
Residual Error	12	13900830	1158402		
Total	15	62596356			

---

---

---

---

---

---

---

---



**PRS 8: What is your conclusion?**

1. The variables  $\delta$  and  $\gamma$  explain a statistically significant portion of the variation in sales.
2. Sales are not significantly different in quarter 2 than quarters 1,3 and 4.
3. The variables d2 and prosed2 do explain a significant portion of the variation in sales.
4. Reject the null hypothesis, these two variables are not important.

---

---

---

---

---

---

---

---