

Stat 401 B – Lecture 13

Multiple Regression

- A single numerical response variable, Y .
- Multiple numerical explanatory variables, X_1, X_2, \dots, X_k

1

Multiple Regression

$$Y = \mu_{Y|x_1, x_2, \dots, x_k} + \varepsilon$$
$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$


2

Example

- Y , Response – Effectiveness score based on experienced teachers' evaluations.
- Explanatory – Test 1, Test 2, Test 3, Test 4.

3


Stat 401 B – Lecture 13



Test of Model Utility

- Is there any explanatory variable in the model that is helping to explain significant amounts of variation in the response?


4



Conclusion

- At least one of the tests is providing statistically significant information about the evaluation score.
- The model is useful. Maybe not the best, but useful.

5




Individual Slope Parameters

- In order to see what tests for the various parameters in multiple regression mean, we need to go back to simple linear regression.


6

Stat 401 B – Lecture 13

 **SLR – EVAL on Test 1**


- Predicted Eval = $329.23 + 1.424 \cdot \text{Test1}$
- For each additional point scored on Test 1, the Evaluation score increases by 1.424 points, on average.

7

 **Explained Variation**

- $R^2 = 0.295$, only 29.5% of the variation in Evaluation is explained by the linear relationship with Test 1.

8

 **Inference on β_1**

- t-Ratio = 2.97
- P-value = 0.0074
- Reject the null hypothesis that $\beta_1 = 0$, because the P-value is so small.
- There is a statistically significant linear relationship between EVAL and Test 1.

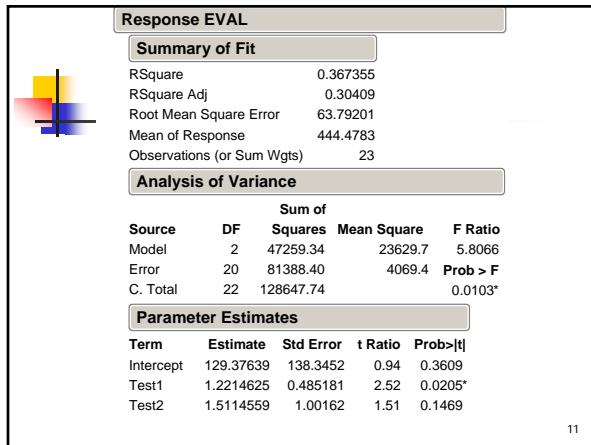
9

Stat 401 B – Lecture 13

Model with Test 1

- If Test 1 is the only explanatory variable in the model, then the other tests are ignored by this model.
- What happens if we add Test 2 to the model with Test 1?

10



Response EVAL

Summary of Fit

RSquare	0.367355
RSquare Adj	0.30409
Root Mean Square Error	63.79201
Mean of Response	444.4783
Observations (or Sum Wgts)	23

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	47259.34	23629.7	5.8066
Error	20	81388.40	4069.4	Prob > F
C. Total	22	128647.74		0.0103*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	129.37639	138.3452	0.94	0.3609
Test1	1.2214625	0.485181	2.52	0.0205*
Test2	1.5114559	1.00162	1.51	0.1469


11

Model with Test 1, Test 2

- Predicted EVAL = $129.38 + 1.221 \cdot \text{Test1} + 1.511 \cdot \text{Test2}$
- For each additional point on Test 1, **while holding Test 2 constant**, the Evaluation score increases by 1.221 points, on average.


12

Stat 401 B – Lecture 13

 **Model with Test 1, Test 2**


- Predicted EVAL = $129.38 + 1.221 \cdot \text{Test1} + 1.511 \cdot \text{Test2}$
- For each additional point on Test 2, **while holding Test 1 constant**, the Evaluation score increases by 1.511 points, on average.

13

 **Explained Variation**

- $R^2 = 0.367$, 36.7% of the variation in Evaluation is explained by the linear relationship with **Test 1 and Test 2**.


14

 **Explained Variation**

- $R^2 = 0.367$ – Test 1 and Test 2.
- $R^2 = 0.295$ – Test 1 alone.
- $0.367 - 0.295 = 0.072$, 7.2% of the variation in EVAL is explained by **the addition of Test 2 to Test 1**.

15


Stat 401 B – Lecture 13



Parameter Estimates – Test 2

- Has Test 2 added significantly to the relationship between Test 1 and Evaluation?
- Note that this is different from asking if Test 2 is linearly related to Evaluation!


16



Parameter Estimates – Test 2

- t-Ratio = 1.51
- P-value = 0.1469
- Because the P-value is not small, Test 2's addition to the model with Test 1 is not statistically significant.

17




Parameter Estimates – Test 2

- Although R^2 has increased by adding Test 2, that increase could have happened just by chance. The increase is not large enough to be deemed statistically significant.

18


Stat 401 B – Lecture 13



Parameter Estimates – Test 1

- Does Test 1 add significantly to the relationship between Test 2 and Evaluation?
- Note that this is different from asking if Test 1 is linearly related to Evaluation!


19



Parameter Estimates – Test 1

- t-Ratio = 2.52
- P-value = 0.0205
- Because the P-value is small, Test 1's addition to the model with Test 2 is statistically significant.

20



Parameter Estimates – Test 1

- If we had started with a model relating Test 2 to EVAL, adding Test 1 would result in an increase in R^2 . That increase is large enough to be deemed statistically significant.

21
