

# Hypothesis Testing with Excel

## Lecture 52

Robb T. Koether

Hampden-Sydney College

Mon, Apr 28, 2008

# Outline

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- 1 Introduction
- 2 Testing the Significance of the Regression Line
- 3 The Goodness of Fit Test
- 4 The Test of Independence

# Introduction

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- It is possible to perform statistical tests in Excel.
- Typically this requires some preliminary calculations.
- Then one may apply any of the functions
  - `ZTEST ()`
  - `TTEST ()`
  - `CHITEST ()`
  - `FTEST ()`

to get the  $p$ -value.

# Excel and the Regression Line

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- We can obtain the regression coefficients  $a$  and  $b$  within cells by using statistical functions.
  - `SLOPE ()` computes the slope of the regression line.
  - `INTERCEPT ()` computes the  $y$ -intercept of the regression line.
- The function `TREND ()` will use the regression line to predict  $y$  for a given value of  $x$ .

# Testing the Significance of $\beta$ and $\rho$

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- By doing the necessary calculations, we can perform the linear regression  $t$  test to determine whether  $\beta$  and  $\rho$  are statistically significant.
- To do this, we need one more function.
  - `TDIST()` computes the area of a tail (or two tails) of the  $t$  distribution with the given number of degrees of freedom.

# The Goodness of Fit Test

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- For the goodness of fit test, we need
  - A set of observed counts.
  - A set of hypothetical proportions.
- We may place the observed counts in one row and the hypothetical proportions in another row.

# The Goodness of Fit Test

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- Get the total of the observed counts. This is the sample size.
- Multiply the hypothetical proportions by the sample size to get the expected counts.

# The Goodness of Fit Test

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- Enter the formula for  $\frac{(O-E)^2}{E}$  in the leftmost cell and fill right
- Get the total of the  $\frac{(O-E)^2}{E}$  values. This is the value of  $\chi^2$ .
- Use the function `CHIDIST()` to find the  $p$ -value.

# The Goodness of Fit Test

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- Or we can use the `CHITEST()` function.
- Get the total of the observed counts, as before.
- Create a new row of the sample size times the hypothetical proportions, as before.
- Then use the function `CHITEST()` to find the  $p$ -value.

# Test of Independence

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- To perform the test of independence, we must calculate the expected counts ourselves.
- Create the array of observed counts.
  - Enter the observed counts in an array of rows and columns.
  - Compute the row totals, the columns totals, and the grand total.
  - Suppose the row total is in cell  $F2$  and the column total is in cell  $B6$ .
  - Then the grand total should be in cell  $F6$ .

# Test of Independence

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- Create the array of expected counts.
  - Enter in the upper-left cell of the array of expected counts the formula for

$$\frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}$$

- Enter the formula `=F$2*$B6/$F$6`.
- The dollar signs will prevent the row or column label from being incremented when the formula is pasted into the other cells.
- Select the array of cells for the expected counts and fill right and down.

# Test of Independence

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- Create the array of values of  $\frac{(O-E)^2}{E}$ .
  - Enter in the upper-left cell the formula for  $\frac{(O-E)^2}{E}$  for that cell.
  - Select the entire array reserved for these values.
  - Fill right and down.
- In a new cell, find the sum of this array. That is the value of  $\chi^2$ .
- Use `CHIDIST ()` to find the  $p$ -value.

# Test of Independence

Hypothesis  
Testing with  
Excel

Robb T.  
Koether

Introduction

Testing the  
Significance  
of the  
Regression  
Line

The  
Goodness of  
Fit Test

The Test of  
Independence

- Or, instead of computing  $\frac{(O-E)^2}{E}$ , use `CHITEST()` and give it the array of observed counts and the array of expected counts.