UNIT 31
A Hypothesis Test about Correlation and Slope with Regression

Objectives:

- To perform a hypothesis test concerning the slope of a least squares line
- To recognize that testing for a statistically significant slope in a linear regression of and testing for a statistically significant linear relationship (i.e., correlation) are the same test

Recall that we can use the Pearson correlation r and the least squares line to describe a linear relationship between two quantitative variables X and Y. We called a sample of observations on two such quantitative variables bivariate data, represented as (x_1,y_1) , (x_2,y_2) , $\cdots$, (x_n,y_n) . Typically, we let Y represent the response variable and let X represent the explanatory variable. From such a data set, we have previously seen how to obtain the Pearson correlation r and the least squares line. We now want a hypothesis test to decide whether a linear relationship is statistically significant.

For example, return to Table 10-2 where data is recorded on X = "the dosage of a certain drug (in grams)" and Y = "the reaction time to a particular stimulus (in seconds)" for each subject in a study. We have reproduced this data here as Table 31-1.

Table 31-1
Dosage and Reaction Time Data (from Table 10-2)

The dosage of a stimulant drug and the reaction time to a stimulus are recorded for each of several subjects injected with the drug.

Dosage (grams)	4	4	6	6	8	8	10	10
Reaction Time (seconds)	7.5	6.8	4.0	4.4	3.9	3.1	1.4	1.7

From the calculations in Table 10-3, you previously found the correlation between dosage and reaction time to be $r = -0.9628$, from these same calculations you also found te least squares line to be $rct = 10.225 - 0.875(dsg)$, where rct represents reaction time in seconds, and dsg represents dosage in grams. However, these are just descriptive statistics; alone, they can not tell us whether there is sufficient evidence of a linear relationship between the two quantitative variables.

When we study the prediction of a quantitative response variable Y from a quantitative explanatory variable X based on the equation for a straight line, we say we are studying the *linear regression of Y on X*. Deciding whether or not a linear relationship is significant is the same as deciding whether or not the slope in the linear regression of Y on X is different from zero. If the slope is zero, then Y does not tend to change as X changes; if the slope is not zero, then Y will tend to change as X changes.

We need a hypothesis test to decide if there is sufficient evidence that the slope in a simple linear regression is different from zero. Such a test could be one-sided or two-sided. If our focus was on finding evidence that the slope is greater than zero (i.e., finding evidence that the linear relationship was a positive one), then we would use a one-sided test. If our focus was on finding evidence that the slope is less than zero (i.e., finding evidence that the linear relationship was a negative one), then we would use a one-sided test. However, if our focus was on finding evidence that the slope is different from zero in either direction (i.e., finding evidence that the linear relationship was either positive or negative), then we would use a two-sided test.

With a random sample of bivariate data, the following t statistic with $n-2$ degrees of freedom is available for the desired hypothesis test:

$$t(n-2) \quad = \quad \frac{b - 0}{\left(\dfrac{s_{Y|X}}{\sqrt{\sum(x-\bar{x})^2}}\right)}$$

where $s_{Y|X}$ is called the *standard error of estimate*. The formula for this t statistic has a format similar to other t statistics we have encountered. The numerator is the difference between a sample estimate of the slope (b) and the hypothesized value of the slope (zero), and the denominator is an appropriate standard error. The calculation of this test statistic requires obtaining the slope of the least squares line b, the sum of the squared deviations of the sample x values from their mean $\sum (x - \bar{x})^2$, and the standard error of estimate $s_{Y|X}$. We have previously seen how to obtain b and $\sum (x - \bar{x})^2$, but we have never previously encountered the standard error of estimate $s_{Y|X}$.

You can think of the standard error of estimate $s_{Y|X}$ for bivariate data similar to the way you think of the standard deviation s for one sample of observations of a quantitative variable (which can be called *univariate data*). Recall that the standard deviation for a sample x values is obtained by dividing $\sum (x - \bar{x})^2$ by one less than the sample size, and taking the square root of the result. The standard error of estimate $s_{Y|X}$ is a measure of the dispersion of data points around the least squares line similar to the way that the standard deviation s is a measure of the dispersion of observations (around the sample mean) in a single sample. The standard error of estimate $s_{Y|X}$ is obtained by dividing the sum of the squared residuals by two less than the sample size, and taking the square root of the result. If you wish, you may do this calculation for the data of Table 31-1, after which you can calculate the test statistic $t(n-2)$. In general, a substantial amount of calculation is needed to obtain the test statistic $t(n-2)$. Since the ability to do these calculations easily is readily available with a lot of mathematics-related software and with many calculators, we shall not concern ourselves with these calculations in detail.

To illustrate the t test about the slope in a simple linear regression, let us consider using a 0.05 significance level to perform a hypothesis test to see if the data of Tables 10-2 and 31-1 provide evidence that the linear relationship between dosage and reaction time is significant, or in other words, evidence that the slope in the linear regression of reaction time from drug dosage is different from zero.

The first step is to state the null and alternative hypotheses, and choose a significance level. This test is two-sided, since we looking for a relationship which could be positive or negative (i.e., a slope which could be different from zero in either direction). We can complete the first step of the hypothesis test as follows:

H_0: slope = 0 vs. H_1: slope $\neq$ 0 ($\alpha = 0.05$, two-sided test)

The second step is to collect data and calculate the value of the test statistic. Whether you choose to do all the calculations yourself or to use appropriate software on a calculator or computer, you should find the test statistic to be $t(6) = -8.723$.

The third step is to define the rejection region, decide whether or not to reject the null hypothesis, and obtain the P-value of the test. Since we have chosen $\alpha = 0.05$ for a two-sided test, the rejection region is defined by the t-score with $df = 6$ above which 0.025 of the area lies; below the negative of this t-score will lie 0.025 of the area, making a total area of 0.05. From Table B.3, we find that $t(6;0.025) = 2.447$. We can then define our rejection region algebraically as

$$t(6) \geq +2.447 \text{ or } t(6) \leq -2.447 .$$

Since our test statistic, calculated in the second step, was found to be $t(6) = -8.723$, which is in the rejection region, our decision is to reject H_0: slope = 0; in other words, our data provides sufficient evidence to suggest that H_1: slope $\neq$ 0. From Table B.3, we find that 8.723 is greater than $t(6;0.0005) = 5.595$, which tells us that $P < 0.001$.

To complete the fourth step of the hypothesis test, we can summarize the results of the hypothesis test as follows:

Since $t(6) = -8.723$ and $t(6;0.025) = 2.447$, we have sufficient evidence to reject H_0. We conclude that the slope in the linear regression of reaction time on dosage is different from zero or in other words, the linear relationship between dosage and reaction time is significant ($P < 0.001$). The data suggest that the slope is less than zero, or in other words, that the linear relationship is negative.

Hypothesis Test for Self-Test Problem 31-1

Step 1:

H_0:

H_1:

α =

Step 2:

Step 3:

Step 4:

Self Test Problem 31-1. In Self-Test Problems 10-1 and 11-1, the relationship between age and grip strength among right-handed males is being investigated, with regard to the prediction of grip strength from age. The Age and Grip Strength Data, displayed in Table 10-4, was recorded for several right-handed males. For convenience, we have redisplayed the data here in Table 31-2.

(a) Verify that the least squares line (found in Self-Test Problem 11-1) is $grp = 26 + 2(age)$, and write a one sentence interpretation of the slope of the least squares line.

(b) A 0.05 significance level is chosen for a hypothesis test to see if there is any evidence that the linear relationship between age and grip strength among right-handed males is positive, or in other words, evidence that the slope in the linear regression of grip strength on age is positive. Complete the four steps of the hypothesis test by completing the table titled *Hypothesis Test for Self-Test Problem 31-1*.

(c) Considering the results of the hypothesis test, decide which of the Type I or Type II errors is possible, and describe this error.

(d) Decide whether H_0 would have been rejected or would not have been rejected with each of the following significance levels: (i) $\alpha = 0.01$, (ii) $\alpha = 0.10$.

Table 31-2
Age and Grip Strength Data

The age (years) and right-hand grip strength (pounds of force) are recorded for each of several right-handed males.

Age	15	17	19	11	16	22	17	25	12	14	25	23
Grip Strength	50	54	66	46	58	54	64	80	46	70	76	80

Answers to Self Test Problems

31-1 (a) The least squares line (found in Self Test Problem 11-1) is $grp = 26 + 2(age)$, where *grp* represents grip strength. With each increase of one year in age, grip strength increases on average by about 2 lbs.

(b) Step 1: H_0: slope = 0 vs. H_1: slope > 0 ($\alpha = 0.05$, one-sided test)

Step 2: $t(10) = +3.814$

Step 3: The rejection region is defined by $t(10) \geq +1.812$.
H_0 is rejected; $0.0005 < P < 0.005$.

Step 4: Since $t(10) = 3.814$ and $t(10;0.05) = 1.812$, we have sufficient evidence to reject H_0. We conclude that the slope in the regression of grip strength on age among right-handed males is greater than zero, or in other words, that the linear relationship between age and grip strength is positive ($0.0005 < P < 0.005$).

(c) Since H_0 is rejected, the Type I error is possible, which is concluding that slope > 0 when in reality slope = 0.

(d) H_0 would have been rejected with $\alpha = 0.01$ and with $\alpha = 0.10$.

Summary

When studying a possible linear relationship between two quantitative variables X and Y, we can obtain the Pearson correlation r and the least squares line from a sample consisting of bivariate data, represented as (x_1,y_1) , (x_2,y_2) , $\cdots$, (x_n,y_n) . With Y representing the response variable and X representing the explanatory variable, we say that we are studying the *linear regression of Y on X*. Deciding whether or not a linear relationship is significant is the same as deciding whether or not the slope in the linear regression of Y on X is different from zero. A test statistic which can be used in a hypothesis test to make this decision is the following t statistic with $n-2$ degrees of freedom:

$$t(n-2) = \frac{b - 0}{\left(\dfrac{s_{Y|X}}{\sqrt{\sum(x-\bar{x})^2}}\right)}$$

where $s_{Y|X}$ is called the *standard error of estimate*, calculated by dividing the sum of the squared residuals by two less than the sample size, and taking the square root of the result. The numerator of this t statistic is the difference between the slope of hte least squares line obtained from a random sample (b) and the hypothesized value of the slope (zero), and the denominator is an appropriate standard error. The ability to do the necessary calculations is readily available with a lot of mathematics-related software and with many calculators.

The null hypothesis in this hypothesis test could be stated as H_0: slope = 0, and the test could be one-sided or two-sided. If our focus was on finding evidence that the slope is greater than zero (i.e., finding evidence that the linear relationship was a positive one), then we would use a one-sided test. If our focus was on finding evidence that the slope is less than zero (i.e., finding evidence that the linear relationship was a negative one), then we would use a one-sided test. However, if our focus was on finding evidence that the slope is different from zero in either direction (i.e., finding evidence that the linear relationship was either positive or negative), then we would use a two-sided test.