Hypothesis Tests for the Correlation Coefficient *r*

Step 1: State the Null Hypothesis (H_0) and Alternative Hypothesis (H_1)

The null hypothesis is a statement assumed to be true. For a hypothesis test about the correlation coefficient *r*, the null hypothesis is

H_0: There is no linear relation between the two variables.

If the evidence against the null hypothesis is strong enough, we will reject the null hypothesis and support the alternative hypothesis. For a hypothesis test about the correlation coefficient *r*, the alternative hypothesis is

H_1: A linear relation exists between the two variables.

So, we are assuming that there is no linear relation and we are trying to prove that there is a linear relation.

Step 2: State the Level of Significance α (alpha)

The level of significance is the boundary line that indicates the evidence is considered to be strong enough to reject the null hypothesis.

Assuming that the null hypothesis is true (there is no linear relation), we find the chance of having a sample with a correlation coefficient at least as extreme as our sample. If this chance is less that the level of significance then we consider this to be strong evidence against the null hypothesis.

Strong evidence → Reject the null hypothesis → Support the alternative hypothesis

Lack of Strong Evidence → Fail to Reject the null hypothesis → Fail to Support the alternative hypothesis

For this section, we shall always choose 0.05 for the level of significance.

Step 2: $\alpha = 0.05$

Step 3: State the Test You Are Performing

We will cover many different hypothesis tests throughout the semester. We label the test so the reader knows exactly what test we are performing.

Step 3: Test for Linear Correlation

Step 4: Compute the Test Statistic and the *P*-value

The test statistic is the value of the correlation coefficient *r*.

The *P*-value is the chance of observing a correlation as extreme as the one in our sample if no linear relation exists between the two variables.

Both of these computations are handled by StatCrunch.

1. Stat > Summary Stats > Correlation
2. Select the two columns containing the data.
3. Check the box for a Two-sided P-value and click Compute.
4. Round your correlation coefficient *r* to 3 decimal places.
   The *P*-value will be contained in parentheses.

Step 5: Make a Decision about the Null Hypothesis and a Conclusion about the Alternative Hypothesis

If the *P*-value is less than the level of significance, then

- We Reject the null hypothesis (we say it is false).
- We Support the alternative hypothesis (we say it is true).

Otherwise, if the *P*-value is not less than the level of significance, then

- We Fail to Reject the null hypothesis (we cannot say it is false).
- We Fail to Support the alternative hypothesis (we cannot say it is true).

The wording and its interpretation is very important. If we lack strong evidence against the null hypothesis then we cannot say it is false. That is not the same as saying it is true – we just cannot say that it is false. In a legal trial we assume a person is innocent, and if there is a lack of strong evidence then we do not say the person is innocent, instead we say that we were not able to prove that the person was guilty. We will cover this with great emphasis as the semester progresses.

Example 1

Hours	Score
7	70
8	76
4	57
9	77
13	91
5	66
9	82
6	64
16	96
3	50

A random sample of 10 students was asked how many hours they studied to prepare for a final exam. The number of hours, as well as each student's score, is listed in the table.

Use the 0.05 level of significance to test whether there is a linear relation between "hours studied" and "score".

Step 1 H_0:

H_1:

Step 2 $\alpha =$

Step 3 Test for Linear Correlation

Step 4 $r =$ *P*-value

Step 5 We Reject / Fail to Reject H_0.

There is / is not sufficient evidence to support that there is a linear relation between "hours studied" and "score".

Example 2

Height	Weight
73	201
69	170
72	180
70	200
72	190
66	175
72	205
72	185
74	186

Here are the heights (in inches) and weights (in pounds) of nine players on the COS baseball team.
Use the 0.05 level of significance to test whether there is a linear relation between “height” and “weight”.

Example 3

Income	Home
38.7	94
47.2	133
43.1	128
34.3	98
51.9	158
37.5	113
37.2	109
36.9	122
44.3	154
65.3	235
51.3	189
52.1	210
52.5	208
48	192
82.6	355
72.4	407

Here is the median income (in $1000s) and median home price (in $1000s) of sixteen randomly selected cities.
Use the 0.05 level of significance to test whether there is a linear relation between “income” and “home price”.