Case Q→Q: Hypotheses to Conclusion

Learning Objective: In a given context, carry out the appropriate inferential method for comparing relationships and draw the appropriate conclusions.

Learning Objective: Specify the null and alternative hypotheses for comparing relationships.

Let's introduce our leading example, which was actually our leading example in the Exploratory Data Analysis section as well.

Example

In a study of the legibility and visibility of highway signs, a Pennsylvania research firm determined the maximum distance at which each of 30 drivers could read a newly designed sign. The 30 participants in the study ranged in age from 18 to 82 years old. The government agency that funded the research hoped to improve highway safety for older drivers and wanted to examine the relationship between age and sign legibility distance. (Data adopted with permission from Utts and Heckard, Mind on Statistics).

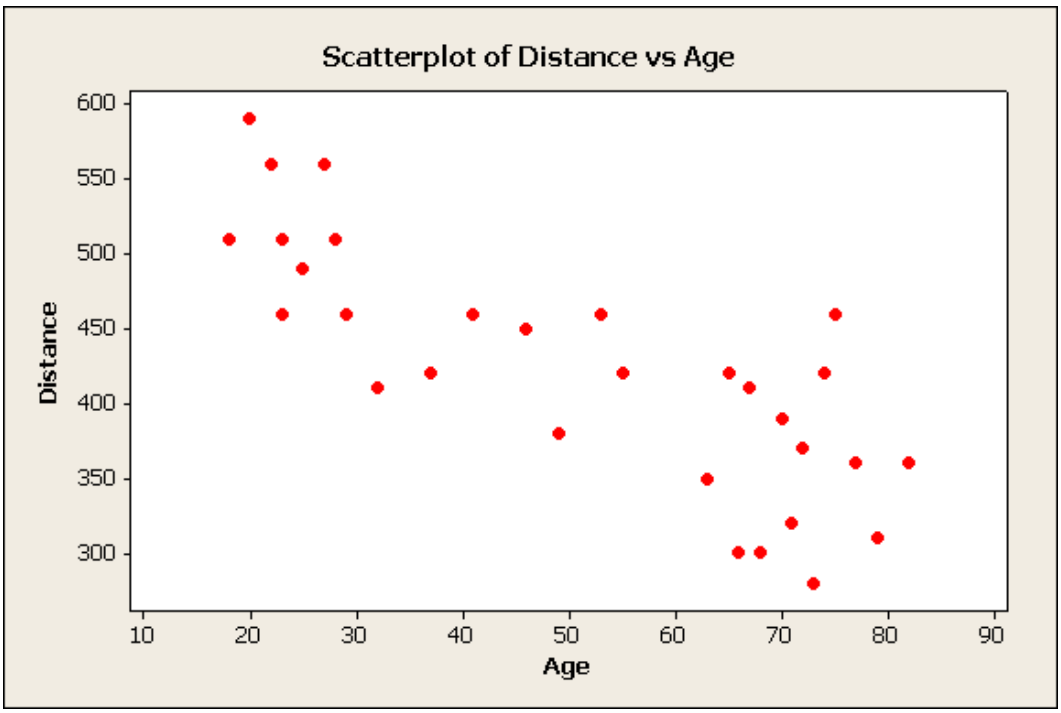
Let's go through the entire process (outlined on the previous page) for this example.

Starting point: The researchers wanted to examine the relationship between age and sign legibility distance in the population of drivers. The researchers collected data from a random sample of 30 drivers:

Age	Distance
18	510
20	590
22	560
23	510
23	460
25	490
27	560
.	.
.	.
.	.
77	360
79	310
82	360

Exploratory Analysis:

The researchers display the data on a scatterplot:



and observe a negative linear relationship in the data. In order to quantify the strength of that linear relationship, the researchers supplement the scatterplot with a numerical measure—the correlation coefficient (r), which turns out to be r = -0.801, confirming the researchers' visual assessment of a negative, fairly strong linear relationship between age and legibility distance.

Scenario: Infant Vocalization and IQ

The purpose of this activity is to introduce the example that you will work on in the activities of this section and to have you get a first feel for the data via exploratory data analysis.

Background: A method for predicting IQ as soon as possible after birth could be important for early intervention in cases such as brain abnormalities or learning disabilities. It has been thought that greater infant vocalization (for instance, more crying) is associated with higher IQ. In 1964, a study was undertaken to see if IQ at 3 years of age is associated with amount of crying at newborn age (Karelitz, et al., "Relation of Crying Activity in Early Infancy to Speech and Intellectual Development at Age Three Years," *Child Development*, 1964 (35), p. 769–777. Click here to read the research paper that describes this study and its results).

In the study, 38 newborns were made to cry after being tapped on the foot, and the number of distinct cry vocalizations within 20 seconds was counted. The subjects were followed up at 3 years of age and their IQs were measured.

Here is what the data look like:

	Cry Count	IQ
1	10	87
2	20	90
3	17	94
4	12	94
5	12	97
6	15	100
7	19	103
8	12	103
9	14	103
10	23	103
11	15	104
12	14	106
13	13	106
14	27	108
15	17	109
16	12	109
17	18	109
18	15	112

19	15	112
20	23	113
21	16	114
22	21	114
23	16	118
24	12	119
25	9	119
26	13	120
27	19	120
28	18	124
29	19	132
30	16	133
31	22	135
32	31	135
33	16	136
34	17	141
35	26	155
36	21	157
37	27	159
38	13	162

As you can see, our data consist of two variables: cry count and IQ.

Learn By Doing (1/1 point)

Which is the explanatory variable and which is the response?

Your Answer:

explanatory: cry count
response: IQ

Our Answer:
Since we want to examine how the vocalization of newborns can explain IQ at age three, the explanatory variable is cry count and the response variable is IQ.

Resubmit Reset

To get a feel for the data, we conduct exploratory analysis by creating a scatterplot of the data, and supplement it with the correlation coefficient r.

Here is the scatterplot:

Using statistics software we find the correlation value to be 0.4018185.

Learn By Doing (1/1 point)

Based on the analysis above, comment on the direction, form, and strength of the relationship.

Your Answer:

No distinct, actually, based on what it looks; but the correlation value means there's a mild positively correlated linear relationship

Our Answer:
From the scatterplot we see that there is a general trend of a positive relationship between cry count and IQ, and that this relationship seems to be linear but quite weak. In order to quantify the strength of the linear relationship, we supplement the scatterplot with r, which turns out to be roughly 0.4. The value of r tells us that the linear relationship between cry count and IQ is weak to moderate.

Resubmit Reset

Inference:

The researchers would now like to see whether the observed linear relationship between age and legibility distance can be generalized to the entire population of drivers. In other words, the researchers want to check whether the observed linearity is due to true linearity in the population, or a pattern that could have happened just by chance.

The test that the researchers are going to carry out is a t-test (most commonly known as the "t-test for the slope" for reasons that we are not going to get into) which is testing the following two hypotheses (step 1):

H_0: There is no linear relationship between age and distance.

H_a: There is a linear relationship between age and distance.

and in general,

H_0: There is no linear relationship between X and Y.

H_a: There is a linear relationship between X and Y.

Did I Get This

1/1 point (graded)

We would like to see if the observed linear relationship between the damage caused by a fire and the distance the fire was from a fire station can be generalized to all fires. To test this, which of the following sets of hypotheses would you use?

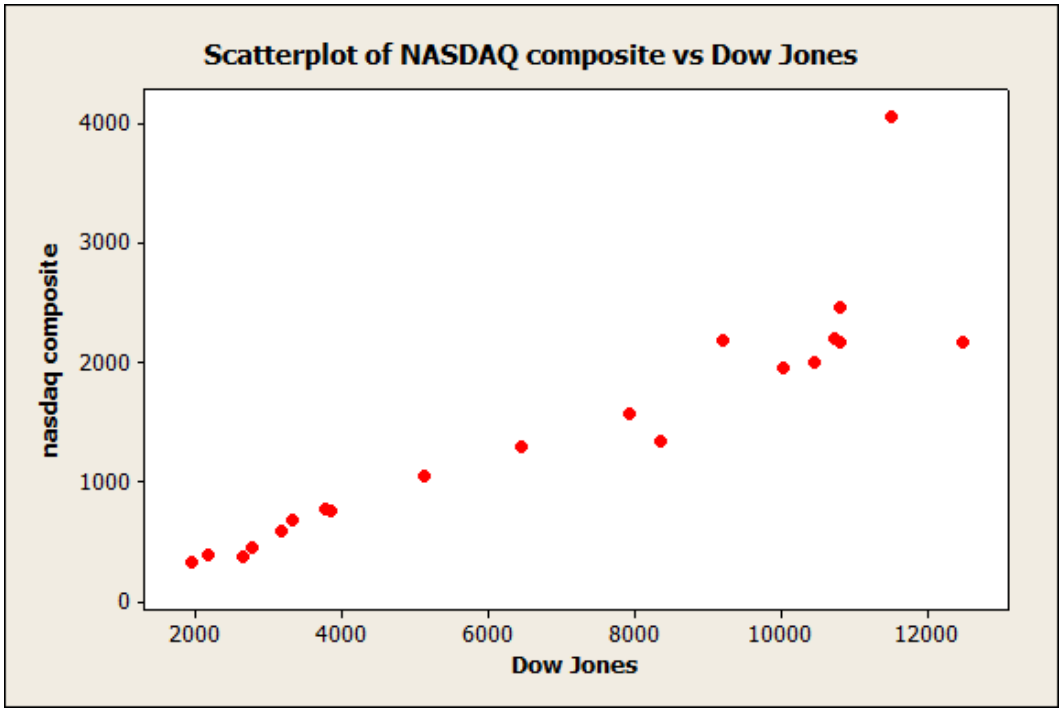
- H_0: there is a relationship between distance and damage
  H_a: there is no relationship between distance and damage
- H_0: there is no relationship between distance and damage
  H_a: there is a relationship between distance and damage
- H_0: there is a linear relationship between distance and damage
  H_a: there is no linear relationship between distance and damage
- **(selected)** H_0: there is no linear relationship between distance and damage
  H_a: there is a linear relationship between distance and damage
  ✔
- H_0: distance and damage are independent
  H_a: distance and damage are dependent

Answer
Correct:
These hypotheses correctly describe the type of relationship, i.e. linear. Also, the null hypothesis correctly indicates no relationship and the alternative hypothesis correctly indicates a relationship.

Submit

Did I Get This

1/1 point (graded)
Using the scatterplot below, which of the following hypotheses would you use to test for the observed linear relationship?



- H_0: the Dow Jones and NASDAQ indexes are independent

  H_a: the Dow Jones and NASDAQ indexes are dependent

- H_0: there is a linear relationship between the Dow Jones and NASDAQ indexes

  H_a: there is not a linear relationship between the Dow Jones and NASDAQ indexes

- H_0: there is a relationship between the Dow Jones and NASDAQ indexes

  H_a: there is not a relationship between the Dow Jones and NASDAQ indexes

- H_0: there is not a relationship between the Dow Jones and NASDAQ indexes

  H_a: there is a relationship between the Dow Jones and NASDAQ indexes

- (●) H_0: there is not a linear relationship between the Dow Jones and NASDAQ indexes

  H_a: there is a linear relationship between the Dow Jones and NASDAQ indexes

  ✔

Answer
Correct:
These hypotheses correctly describe the type of relationship, i.e. linear. Also, the null hypothesis correctly indicates no relationship and the alternative hypothesis correctly indicates a relationship.

Submit

Comments

1. As we mentioned earlier, we are going to keep this discussion on the qualitative side and in particular we will not go very deeply into **step 2** of the hypothesis test. As for the **test-statistic** in this case, we'll just say that the test is a t-test, which, as we know, means that the null distribution of its test statistic (under which the p-values are calculated) is some t distribution.

2. We are also going to focus on only some of the **conditions** that allow us to safely use this t-test. They are:

- the observed data indeed look linear (otherwise it would not make sense to try and generalize them)
- the observations are independent
- there are no extreme outliers in the data
- the sample size is fairly large

Note that in our example all these conditions are met; the data definitely look linear, the observations (drivers) are independent of each other (since they were randomly chosen), there are no extreme observations in the data, and a sample size of n = 30 is fairly large.

For **step 3,** the researchers use statistical software to find a test statistic value of -7.09, and a p-value that is so small that it is essentially 0. This means that it would be extremely unlikely (actually, quite impossible) to get data like those observed if age and legibility distance were not linearly related. In other words, it would be extremely unlikely to get data like those observed just by chance.

The researchers conclude (**step 4**) that since the p-value is so small, the data provide extremely strong evidence to reject H_0 and conclude that age and legibility distance are linearly related.

Did I Get This

1/1 point (graded)

H_0: There is not a linear relationship between the response time and the damage in a fire.

H_a: There is a linear relationship between the response time and the damage in a fire.

Our conclusion is that there is not sufficient evidence of a significant linear relationship between response time and fire damage.

For the p-value below, would it support this conclusion?

p = 0.004

- ○ would support conclusion
- ◉ would *not* support conclusion ✔

Answer

Correct: With a p-value < 0.05, we would reject the null hypothesis of no linear relationship.

Submit

Did I Get This

1/1 point (graded)

H_0: There is not a linear relationship between the response time and the damage in a fire.

H_a: There is a linear relationship between the response time and the damage in a fire.

Our conclusion is that there is not sufficient evidence of a significant linear relationship between response time and fire damage.

For the p-value below, would it support this conclusion?

p = 0.020

- ○ would support conclusion
- ◉ would *not* support conclusion ✔

Answer

Correct: With a p-value < 0.05, we would reject the null hypothesis of no linear relationship.

Submit

Did I Get This

1/1 point (graded)

H_0: There is not a linear relationship between the response time and the damage in a fire.

H_a: There is a linear relationship between the response time and the damage in a fire.

Our conclusion is that there is not sufficient evidence of a significant linear relationship between response time and fire damage.

For the p-value below, would it support this conclusion?

p = 0.072

- (•) would support conclusion ✔
- ( ) would *not* support conclusion

Answer

Correct:

When the p-value > 0.05 we cannot reject the null hypothesis (we do not have sufficient evidence to reject the hypothesis of no significant linear relationship).

Submit

Did I Get This

1/1 point (graded)

H_0: There is not a linear relationship between the response time and the damage in a fire.

H_a: There is a linear relationship between the response time and the damage in a fire.

Our conclusion is that there is not sufficient evidence of a significant linear relationship between response time and fire damage.

For the p-value below, would it support this conclusion?

p = 0.355

- (●) would support conclusion ✔
- ( ) would *not* support conclusion

Answer
Correct:
When the p-value > 0.05 we cannot reject the null hypothesis (we do not have sufficient evidence to reject the hypothesis of no significant linear relationship).

Submit

Scenario: Infant Vocalization and IQ

The purpose of this activity is to give you guided practice in testing whether the data provide evidence of a significant linear relationship, and in verifying that the basic conditions under which the results of such a test are reliable are met.

Recall the example from the previous activity:

A method for predicting IQ as soon as possible after birth could be important for early intervention in cases such as brain abnormalities or learning disabilities. It has been thought that greater infant vocalization (for instance, more crying) is associated with higher IQ. In 1964, a study was undertaken to see if IQ at 3 years of age is associated with amount of crying at newborn age. In the study, 38 newborns were made to cry after being tapped on the foot and the number of distinct cry vocalizations within 20 seconds was counted. The subjects were followed up at 3 years of age and their IQs were measured.

We would now like to test whether the observed (weak-to-moderate) linear relationship between cry count and IQ is significant (in other words, we would like to carry out the "t-test for the slope" for this example).

Learn By Doing

1/1 point (graded)
True or false? This is the appropriate alternative hypothesis (for the t-test for the slope):

Ha: There is no linear relationship between cry count and IQ.

- ○ True
- ◉ False ✔

Answer
Correct:
This is the appropriate *null* hypothesis, Ho. The appropriate alternative hypothesis is Ha: There is a significant linear relationship between cry count and IQ.

Submit

Verify that the general conditions that allow you to reliably use the test are met.

Comment: Note that not all the information you need is given in the problem description. In particular, no information is given about how the babies were selected for the study, which is important for assessing whether we can assume that the observations are independent. To get this information, click here and read the first couple of pages of the paper that describes the study.

Learn By Doing

1/1 point (graded)
Which of the following conditions must be met in order to test the linear relationship between cry count and IQ? Check all that apply.

- ☑ Independent observations
- ☑ The relationship is linear
- ☑ There are no extreme outliers
- ☑ The sample size is reasonably large

✔

Answer

Correct:
Independence is one of the conditions that must be met. According to the paper, the 38 infants were a sample of infants born at the Long Island Jewish Hospital at the time of the study. Given this information, we can assume that the observations are independent (even though it doesn't explicitly say that the sample was random). The only obvious way that the independence assumption would be violated is if twins (or triplets ...) were included in the study, which we can assume was not the case. Also, the relationship between the explanatory and response variable should be linear. In this example, the relationship seems linear (even though it is moderately weak). Additionally, there should be no extreme outliers. Lastly, the sample size should be reasonably large. In this example, sample size meets this condition (n = 38).

Submit

Learn By Doing

1/1 point (graded)

Here is the output from the simple linear regression test:

Simple linear regression results:
Dependent Variable: IQ
Independent Variable: cry count
IQ = 90.75499 + 1.5363518 cry count
Sample size: 38
R (correlation coefficient) = 0.4018
R-sq = 0.1614581
Estimate of error standard deviation: 17.99418

Parameter estimates:

Parameter	Estimate	Std. Err.	Alternative	DF	T-Stat	P-Value
Intercept	90.75499	10.47342	≠ 0	36	8.665267	<0.0001
Slope	1.5363518	0.5835417	≠ 0	36	2.6328056	**0.0124**

Do the data provide enough evidence for us to conclude that there is a significant linear relationship among infants between vocalization right after birth (as measured by cry count) and IQ at age 3?

- Yes because the p-value is greater than 0.05.
- Yes because the p-value is less than 0.05. ✔
- No because the p-value is greater than 0.05.
- No because the p-value is less than 0.05.

Answer

Correct:
The p-value of the test is 0.012. The small p-value (in particular, smaller than .05) tells us that it would be quite unlikely (1.2% chance) to get data like those observed just by chance. The data, therefore provide enough evidence for us to conclude that there is a significant linear relationship among infants between vocalization right after birth (as measured by cry count) and IQ at age 3.

Submit

Note: It is important to distinguish between the information provided by r and by the p-value. The correlation coefficient **r informs us about the strength of the linear relationship in the data**: close to +1 or -1 for a strong linear relationship, close to 0 for a weak linear relationship. In contrast, the regression **p-value informs us about the strength of evidence** that there is a linear relationship in the population from which the data were obtained.

In our example, since the p-value is 0.000 and r = -0.801, we would say that we have extremely strong evidence of a fairly strong relationship between age and distance in the population of drivers.

Did I Get This

1/1 point (graded)
In our cry count of newborns and IQ at age three example, we have found the following:

The correlation coefficient of our observed data is r = 0.4.

The p-value of the test:

- H_o: There is no linear relationship between cry count and IQ.
- H_a: There is a significant linear relationship between cry count and IQ.

is p = 0.012.

Which of the following statements is true?

- The data provide moderately weak evidence of a fairly strong linear relationship between cry count (a few days after birth) and IQ (at age 3) among infants.
- The data provide fairly strong evidence of a fairly strong linear relationship between cry count (a few days after birth) and IQ (at age 3) among infants.
- The data provide moderately weak evidence of a moderately weak linear relationship between cry count (a few days after birth) and IQ (at age 3) among infants.

- The data provide fairly strong evidence of a moderately weak linear relationship between cry count (a few days after birth) and IQ (at age 3) among infants. ✔

Answer

Correct:

Indeed, r = 0.4 indicates that there is a moderately weak linear relationship between cry count and IQ in our data. In addition, the small p-value (0.012) tells us that these data provide fairly strong evidence that this linear relationship can be generalized to infants in the population.

Submit