

## TOPIC: HYPOTHESIS TEST FOR CORRELATION COEFFICIENT

### Hypothesis Test for Correlation Coefficient Using TI-84

- ◆ Recall: The corr. coeff. measures strength of correlation. A hyp. test can determine if the correlation is *significant*.
  - ▶ To run a hypothesis test for the population corr. coeff. ( $\rho$ ), use the **F:LInRegTTest** function.
  - ▶ If testing "no correlation",  $\rho$  \_\_\_ 0; "positive correlation",  $\rho$  \_\_\_ 0; "negative correlation",  $\rho$  \_\_\_ 0.

#### EXAMPLE

A city's public health dept. is investigating whether poor air quality (high AQI) is associated with asthma-related ER visits. Use the data below to find the corr. coeff. & test whether the correlation is statistically significant using  $\alpha = 0.01$ .

$$H_0: \rho = \underline{\hspace{2cm}}$$

$$H_a: \rho [ \neq | < | > ] \underline{\hspace{2cm}}$$

$$r = \underline{\hspace{2cm}}$$

$$P\text{-value} = \underline{\hspace{2cm}}$$

| Air Quality Index vs Asthma ER Visits |    |    |    |    |    |    |    |    |    |    |
|---------------------------------------|----|----|----|----|----|----|----|----|----|----|
| AQI                                   | 42 | 58 | 67 | 35 | 80 | 90 | 55 | 73 | 40 | 62 |
| ER Visits                             | 8  | 10 | 13 | 6  | 17 | 20 | 11 | 15 | 7  | 12 |

Because  $P\text{-value}$  [ < | > ]  $\alpha$ , we [ **REJECT** | **FAIL TO REJECT** ]  $H_0$ ,  
So there is [ **ENOUGH** | **NOT ENOUGH** ] evidence to suggest there is a \_\_\_\_\_  
linear correlation between air quality & asthma related ER visits.



#### HOW TO: Hyp Test for Correl. Coeff.

1) **STAT**, 1: **EDIT...**

Enter data in **L1** & **L2**

2) **STAT** > **TESTS**

**▼** **F:LInRegTTest**

3) **Xlist:** **L1**

**Ylist:** **L2**

**Freq:** 1

**$\beta$  &  $\rho$ :**  **$\neq 0$**  **< 0** **> 0**

**Calculate**

## TOPIC: HYPOTHESIS TEST FOR CORRELATION COEFFICIENT

### PRACTICE

An economist wonders if the inflation rate is linearly correlated with the unemployment rate and is looking to use the results of their analysis for further study. They take a random sample of recent months and record the unemployment rate and inflation rate. They find  $r = 0.23$  and run a hypothesis test, getting a  $P$ -value of 0.35. Interpret the value of  $r$  and the results of the test.

### EXAMPLE

A communications manager at a call center is interested in whether shorter call lengths lead to increases in customer satisfaction, so they record the length (in minutes) of 15 randomly selected calls and the satisfaction rating (out of 10) in the table below.

| Call Length vs Customer Satisfaction |   |    |   |    |   |    |   |   |   |    |   |    |   |   |   |
|--------------------------------------|---|----|---|----|---|----|---|---|---|----|---|----|---|---|---|
| Call Length                          | 6 | 26 | 8 | 11 | 7 | 13 | 8 | 7 | 9 | 10 | 6 | 3  | 8 | 9 | 6 |
| CS Rating                            | 8 | 2  | 7 | 4  | 7 | 3  | 6 | 6 | 5 | 4  | 8 | 10 | 9 | 6 | 9 |

(A) Find  $r$  and interpret.

(B) Perform a hypothesis test for  $\rho$  with a significance level of 0.05.

$H_0$ : \_\_\_\_\_  $H_a$ : \_\_\_\_\_  $P$ -value: \_\_\_\_\_

Because  $P$ -value [  $<$  |  $>$  ]  $\alpha$ , we [ REJECT | FAIL TO REJECT ]  $H_0$ .

So there is [ ENOUGH | NOT ENOUGH ] evidence to suggest there is a \_\_\_\_\_ linear correlation between call length and customer satisfaction level.

 **HOW TO: Hyp Test for Correl. Coeff.**

-  , 1: EDIT...  
Enter data in L1 & L2
-   **TESTS**  
 **F:LinRegTTest**
- Xlist: L1  
Ylist: L2  
Freq: 1  
 $\beta$  &  $\rho$ :  $\neq 0$    $> 0$   
**Calculate**

## TOPIC: HYPOTHESIS TEST FOR CORRELATION COEFFICIENT

### EXAMPLE

An HR manager is interested in whether the length of training (in days) an employee receives impacts the score they receive on their first performance review (out of 10). They collect the following data:

| Length of Training vs Employee Score |   |    |   |   |   |    |   |   |   |   |   |   |   |   |   |
|--------------------------------------|---|----|---|---|---|----|---|---|---|---|---|---|---|---|---|
| Training                             | 9 | 10 | 3 | 4 | 7 | 10 | 8 | 7 | 6 | 6 | 5 | 4 | 8 | 9 | 6 |
| Score                                | 7 | 8  | 8 | 5 | 6 | 10 | 9 | 8 | 8 | 7 | 6 | 7 | 9 | 8 | 9 |

(A) Create a scatterplot. Does it seem possible that length of training and employee score are linearly correlated?

(B) Find  $r$  and interpret.

(C) Perform a hypothesis test for  $\rho$  with a significance level of 0.05.

$H_0$ : \_\_\_\_\_  $H_a$ : \_\_\_\_\_  $P$ -value: \_\_\_\_\_

Because  $P$ -value [ < | > ]  $\alpha$ , we [ REJECT | FAIL TO REJECT ]  $H_0$ .

So there is [ ENOUGH | NOT ENOUGH ] evidence to suggest there is a \_\_\_\_\_ linear correlation between time spent on training and employee score.



### HOW TO: Hyp Test for Correl. Coeff.

1) **STAT**, 1: **EDIT...**

Enter data in **L1** & **L2**

2) **STAT** > **TESTS**

**▼** **F:LinRegTTest**

3) **Xlist:L1**

**Ylist:L2**

**Freq:1**

**$\beta$  &  $\rho$ :  $\neq 0$  < 0 > 0**

**Calculate**