

 Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit <http://online.stanford.edu>.

Course > Inference: Hypothesis Testing for the Population Mean > z-test for the Population Mean >
Statistics Package Exercise: Conducting a Hypothesis Test for the Population Mean: z test

 Bookmark this page

Statistics Package Exercise: Conducting a Hypothesis Test for the Population Mean: z test

Learning Objective: Carry out hypothesis testing for the population proportion and mean (when appropriate), and draw conclusions in context.

The purpose of this activity is to give you guided practice in going through the whole process of hypothesis testing for the population mean (assuming that σ is known).

Background:

The length of human pregnancy is known to have a mean of 266 days and a standard deviation of 16 days. Based on records from a large women's hospital, a random sample of 25 women who were smoking and/or drinking alcohol during their pregnancy and their pregnancy lengths are recorded in the datafile linked below. Do the data provide enough evidence to support the (well-known) fact that women who smoke and/or drink alcohol during their pregnancy have shorter pregnancies than women in general (in other words, are more likely to have premature labor)?

Comment:

It is reasonable to assume that the known standard deviation of 16 days applies also to women who smoke and/or drink during their pregnancy.

-  **StatCrunch** **TI Calculator** **Minitab** **Excel**

R Instructions

To open R with the data set preloaded, right-click here and choose "Save Target As" to download the file to your computer. Then find the downloaded file and double-click it to open it in R.

The data have been loaded into the data frame

pregnancy

. Enter the command

pregnancy

to see the data. The variable name in the data frame is

length

.

-  **StatCrunch** **TI Calculator** **Minitab** **Excel**

R Instructions

The R base package does not have a z-test command built in so we must either create our own z-test command function or use an add-on package. We will create our own function

z.test ()

. The parameters for

z.test ()

are:

- **x**

: a sample data set

- `sig`

: the population standard deviation

- `mu0`

: the mean hypothesized by the null hypothesis

- `alt`

: either "less" or "greater" or "two.sided", indicating the form of the alternative hypothesis $\mu < \mu_0$, $\mu > \mu_0$, or $\mu \neq \mu_0$ respectively. The default is "greater" meaning that if you do not identify alt it will automatically conduct the "greater than" hypothesis test.

Copy this function to R as a single block of code:

- `z.test = function(x, sig, mu0, alt="greater") {`
- `mu = mean(x);`
- `n = length(x);`
- `z = (mu-mu0)/(sig/sqrt(n));`
- `if (alt=="less"){p = pnorm(z)}`
- `else {`
- `if (alt=="two.sided"){p = 2*(1-pnorm(abs(z)))}`
- `else {p = 1-pnorm(z)}`
- `}`

- ```
paste("mean = ",mu,"n = ",n," z = ",z," p-value = ",round(p,5))
```
- ```
}
```

The function

```
z.test()
```

returns the sample mean, the sample size, the z-score for the sample and the p-value for the hypothesis test. To use the function

```
z.test()
```

on the data already entered, execute the command:

- ```
z.test(x=pregnancy$length, sig=16, mu0=266, alt="less")
```

Note: We use this same command for other "Learn By Doing" activities, so be aware that if you restart R before the next module, you must reenter the block of code that defines the

```
z.test()
```

function into R to use it again. Otherwise, you will get the following error:

- ```
Error: could not find function "z.test"
```

Learn By Doing (1/1 point)

State the test statistic, interpret its value, and show how it was found.

Your Answer:

test statistic $z = -1.975$

test statistic = measure of the sample mean's results in standard deviations!

We used $(\text{sample mean} - \text{population mean}) / (\text{stdev} / \sqrt{n})$

Our Answer:

RStatCrunch TI CalculatorMinitabExcel R Here is the R output: Based on the output, the sample mean pregnancy length of the 25 women is 259.68. In addition, we know that: $n = 25$, $\mu = 266$ and $\sigma = 16$. The test statistic is therefore: as given in the output. This means that the sample mean is almost 2 standard deviations below the null value. StatCrunch Here is the StatCrunch output: Based on the output, the sample mean pregnancy length of the 25 women is 259.68. In addition, we know that: $n = 25$, $\mu = 266$ and $\sigma = 16$. The test statistic is therefore: $= 259.68 - 266.1625 = -1.975$ as given in the output. This means that the sample mean is almost 2 standard deviations below the null value. TI Calculator Based on the output, the sample mean pregnancy length of the 25 women is 259.68. In addition, we know that: $n = 25$, $\mu = 266$ and $\sigma = 16$. The test statistic is therefore: as given in the output. This means that the sample mean is almost 2 standard deviations below the null value. Minitab Based on the output, the sample mean pregnancy length of the 25 women is 259.68. In addition we know that: $n = 25$, $\mu = 266$ and $\sigma = 16$. The test statistic is therefore: as given in the output. This means that the sample mean is almost 2 standard deviations below the null value. Excel Based on the dataset, the sample mean pregnancy length of the 25 women is 259.68. In addition, we know that: $n = 25$, $\mu = 266$ and $\sigma = 16$. The test statistic is therefore: This means that the sample mean is almost 2 standard deviations below the null value.

Resubmit

Reset

Learn By Doing (1/1 point)

State the p-value, and draw your conclusions in context.

Your Answer:

p-value is 0.02413. At 0.05 significance level, it's statistically significant. There's only probability 0.02 to encounter this value as the sample mean. So we can reject H_0 and accept H_a .

Our Answer:

RStatCrunch TI CalculatorMinitabExcel R The p-value is 0.024, which (using the 0.05 significance level) is small enough to indicate that the results are significant. In other words, the data provide enough evidence to reject H_0 and conclude that the mean pregnancy length of women who smoke and/or drink alcohol during pregnancy is smaller than the mean pregnancy length of women in general. StatCrunch The p-value is 0.024, which (using the 0.05 significance level) is small enough to indicate that the results are significant. In other words, the data provide enough evidence to reject H_0 and conclude that the mean pregnancy length of women who smoke and/or drink alcohol during pregnancy is smaller than the mean pregnancy length of women in general. TI Calculator The p-value is .024, which (using the .05 significance level) is small enough to indicate that the results are significant. In other words, the data provide enough evidence to reject H_0 and conclude that the mean pregnancy

length of women who smoke and/or drink alcohol during pregnancy is smaller than the mean pregnancy length of women in general. Minitab The p-value is .024, which (using the .05 significance level) is small enough to indicate that the results are significant. In other words, the data provide enough evidence to reject H_0 and conclude that the mean pregnancy length of women who smoke and/or drink alcohol during pregnancy is smaller than the mean pregnancy length of women in general. Excel Using the Excel function NORMSDIST(-1.97), we find that the p-value is .024. At the .05 significance level, this p-value is small enough to indicate that the results are significant. In other words, the data provide enough evidence to reject H_0 and conclude that the mean pregnancy length of women who smoke and/or drink alcohol during pregnancy is smaller than the mean pregnancy length of women in general.

[Resubmit](#)[Reset](#)

Open Learning Initiative [↗](#)



[↗](#) Unless otherwise noted this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License [↗](#).

© All Rights Reserved