

# Z-tests and One sample tests

Evan Misshula

2017-04-01

# Outline

- 1 Learning objectives
- 2 Introduction
- 3 Steps in Hypothesis Testing
- 4 Steps in Hypothesis Testing
- 5 Steps In Hypotheses Testing
- 6 There are two basic kinds of tests:
- 7 Different tests are appropriate under different circumstances.
- 8 Z test Example 2
- 9 Z test Example 2
- 10 Z test Example 2
- 11 Finding Critical Vals (Exercise)
- 12 Research
- 13 Step 1: State the Hypotheses
- 14 Step 2: Locate the critical region and critical t value(s)
- 15 Step 3: Compute the Test Statistic (obtained value)
- 16 Step 4: Make a Decision

# Learning Objectives

- Steps in hypothesis testing
- Discuss the assumptions for a Z test
- Calculate and interpret the Z test analysis
- Discuss the assumptions for a one-sample t-test
- Calculate and interpret the one-sample t-test

Hypothesis testing: develop hypothesis and then research to find out if it is true

## Steps in Hypothesis testing

- 1 Develop Research hypothesis
- 2 Develop Null hypothesis
- 3 Draw samples
- 4 Select direction of test
- 5 Determine the test and calculate obtained value
- 6 Determine the significance and critical value
- 7 Make a decision

# Step 6: Determining significance and critical regions

## Critical value

- Level of certainty (probability of making Type I error), degrees of freedom, one- or two-tailed test
- Level of significance

## Step 7: Making a decision

Value obtained from the statistical test, *test value*, is compared to the critical value

- If test  $>$  critical, reject null hypothesis
- If test  $\leq$  critical, fail to reject null hypothesis

- The *critical or rejection* region is the range of values of the test value that indicates that there is a significant difference and that the null hypothesis should be rejected.
- The *noncritical or non-rejection* region is the range of values of the test value that indicates that the difference was probably due to chance and that the null hypothesis should not be rejected.

- make the assumption that the population is normal.



# Nonparametric tests

- do not make the assumption that the population is normal.

# Which tests

- 1 Z tests are used for comparing a large sample to a population or two large samples.
- 2 The t-test is used for comparing a small sample to a population or two small samples.
- 3 The Chi-square Test of Independence is used for nonparametric data.

# The Z Test

- The Z test examines whether a sample could have come from a known population or whether two samples come from the same population.
- Draws directly from the Normal Curve, Central Limit Theorem, and Sampling distributions.
- Tests whether two distributions are similar enough that they could be from the same population.

# The Z Test

- This is a parametric procedure.
- Assumptions:
  - The dependent variable must be at least interval level.
  - The population must be normally distributed.
  - The variances of the two samples should be equal.
  - The means from the two samples drawn are equal (although this is not an assumption that is the goal of the research).

# The Z test formula

- The formula for calculating the z test is the same formula we used when calculating the z score for sample means.
- Now though, instead of looking at the percentage above or below our z score, we will be comparing it to a critical value.

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{N}}}$$

# Z-Test Critical Values

- When conducting z-tests, we have to determine what our critical region will be.
- This is based on whether the test is a one-tailed or two-tailed test and our level of confidence.
- For a 95% confidence level, we can use two critical values
  - One-tailed test: 1.65 or -1.65
  - Two-tailed test: 1.96 & -1.96

# Example

- A researcher is interested in determining if juvenile prisoners have a shorter attention span than the general population of juveniles. The average length of the attention span of juveniles is 84 minutes with a standard deviation of 16.
- The researcher draws a random sample of 125 juvenile delinquents and administers an attention span questionnaire. The average attention span for the juveniles was 78.
- Is there a statistically significant difference between juvenile prisoners and the general juvenile population when it comes to attention span?

# User R to solve

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{N}}}$$

```
xBar <- 78
N <- 125
mu <- 84
sdev <- 16
testValue <- round((xBar-mu)/(sdev/sqrt(N)),2)
probChance<-round(pnorm(testValue,0,1,lower.tail=T),2)
paste0("testVal= ",testValue," prob chance = ",probChance)

[1] "testVal= -4.19 prob chance = 0"
```



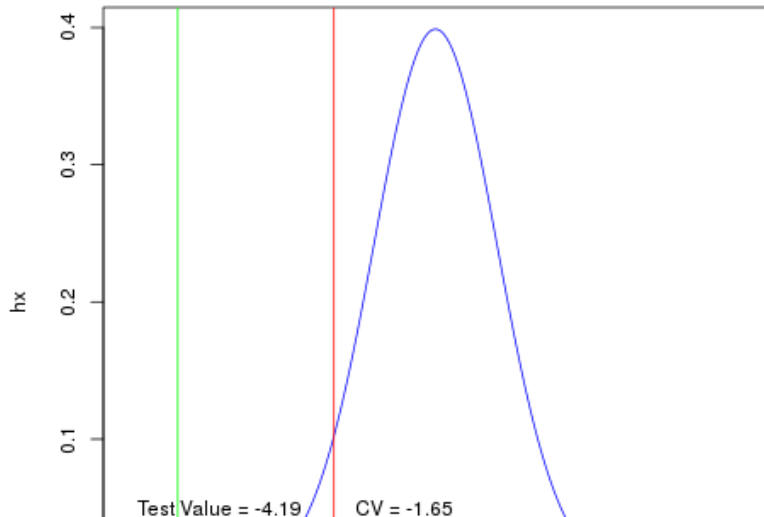
## Step 6: Obtaining the Critical Value.

- Since this is a one-tailed test, use the z-score that corresponds to .0500 in the tail.
- If you had a two-tailed test, you would look for the z score that corresponds to .025 in Each tail, or 1.96 and -1.96.
- We see that this corresponds to the Z critical -1.65.

## Step 7: Making a decision

- Because the obtained value ( $-4.19$ ) is further out than the critical value ( $-1.65$ ), we can reject the null hypothesis.

# Graphic Representation



# Interpretation and Application: Known Probability of Error

- Interpretation of a Z test relies on the known probability of error to determine whether the two distributions came from the same population.
- The known probability of error is tied to the Normal Curve.
- The Normal Curve is consistent and determinable.
- The population or one of the samples is used as the baseline for comparison with the other sample.

# A researcher is interested in determining if criminals have a different IQ

score than the general population. The population mean is 100 with a standard deviation of 10. The researcher draws a random sample of 137 inmates at Rikers Island and administers an IQ test. The average IQ for the inmates was 87. Is there a statistically significant difference between inmates and the general public when it comes to IQ?

# Follow the Steps.

- 1 Research Hypothesis
- 2 Null Hypothesis
- 3 Select the test
- 4 Calculate the obtained value
- 5 Calculate the critical value
- 6 Make the decision

# Research Hypothesis

- H1: Criminals will have different IQ scores on average than the general population.
- H1:  $\mu \neq 100$  \*\* Null Hypothesis
  - H0: Criminals will have the same IQ scores on average as the general population
  - H0:  $\mu = 100$
- \*\* Select the test
  - Two-tailed test because we are only concerned with if IQ scores are different. No direction is specified

# Calculating the Obtained Value.

\*\*\*\*What do we know?

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{N}}}$$

```
xBar <- 87
```

```
N <- 137
```

```
mu <- 100
```

```
sdev <- 10
```

```
testValue <- round((xBar-mu)/(sdev/sqrt(N)),2)
```

```
probChance<-round(pnorm(testValue,0,1,lower.tail=T),2)
```

```
paste0("testVal= ",testValue," prob chance = ",probChance)
```

```
[1] "testVal= -15.22 prob chance = 0"
```



# Z score of Sample Means

- Must know:
  - Sample Mean ( $\bar{x}$ )
  - Population Mean ( $\mu$ )
  - Sample Size (N)
  - Population Std. deviation ( $\sigma$ )

What if we do not know the population std. deviation?

# The T Statistic

- If the researcher does not know the std. deviation of the population  $\sigma$ , but knows  $\mu$ ,  $n$ , &
- Then, standard deviation of the Sample ( $s$ ) is used.
- Standard Deviation of sample denoted as “ $s$ ”

# Sampling Error

- We are still using samples to infer about the population,
- Therefore, sampling error will still occur
- Now we use “s” instead of “ $\sigma$ ”
- When using s to calculate our expected amount of error it is called the Estimated Standard Error.

# T-test is used over a Z-test when:

- 1 The t-statistic is specifically designed to test hypotheses about mean differences in situations where the population s.d. ( $\sigma$ ) is unknown
- 2 Uses the sample s.d. ( $S$ ) to estimate the population s.d. ( $\sigma$ )
- 3 For small sample sizes ( $n < 120$ ), the sample s.d. can be a biased estimate of the population s.d.

## T-test is used over a Z-test when (part 2):

- 1 t-statistic can also be applied to situations where the sample size is small ( $n < 120$ )
- 2 Because Z-test or Z-distribution (the normal distribution) based on the Central Limit Theorem does not work well with samples of smaller sizes ( $n < 30$ ) unless the values in the whole population were normal

# Degrees of freedom

- “s” is used instead of “ $\sigma$ ”
- “N-1” is used instead of “N” in the denominator
- The critical value is adjusted for sample size
- In essence, t is an estimated z-score: It is the value we estimate the value of z would be if we knew the population variability.

Formula is very close to our old one

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{N-1}}}$$

# T-Distribution

- Because we are using the standard deviation of the sample and not population,  $t$  values will have more variability than  $Z$  values, which results in a different shape of the distribution.
- Why?
- Because every time a  $Z$  score is calculated only one thing changes:
  - The sample Mean
- Every time a  $t$  value is calculated two things change:
  - The Sample Mean
  - The standard deviation.



# Characteristics of the t-distribution (sampling distribution of t-statistic)

- The shape of the t-distribution changes with sample size (d.f.)
- There is no one fixed/ particular shape for the t-distribution: different samples with a different size (d.f) will have a different shape
- All t- distributions are symmetrical around the standardized mean of zero (similar to the normal Z-distribution)

# T-distribution

- Converges to Gaussian when the samples (df) have a size of at least 120
- T values are expected to fall further away from 0 when d.f. is  $< 120$ .
- The distribution tends to have thicker tails and less of a peak than Z, but approaches the Z distribution as df increases and is normal at a df of 120

# T distribution

- T values are expected to be more extreme than Z values.
- Therefore, when conducting a hypothesis test with t, Critical values have to be further away from zero than Z at all common alpha levels
- But as degree of freedom increases, the distribution becomes more normal, therefore critical values can become less extreme. . . meaning closer to zero.

# T numeric values

- Two sided examples

```
side <- 2
regVal <- .05
myRegZtest <- qnorm(regVal/side,0,1,lower.tail = T)
my3DegTtest <- qt(regVal/side,3,lower.tail = T)
my30DegTtest <- qt(regVal/side,30,lower.tail = T)
my120DegTtest <- qt(regVal/side,120,lower.tail = T)
c(myRegZtest,my3DegTtest,my30DegTtest,my120DegTtest)

[1] -1.959964 -3.182446 -2.042272 -1.979930
```

# Using R to find critical values

- We can use R to find critical values for a given  $\alpha$  level and test type

```
side <- 2
alphaTrad <- .05
alphaLoose <- .10
alphaTight <- .01
oneSidedZtest <- qnorm(alphaTrad,0,1,lower.tail = T)
myCritValueF <- function(x){
  return(qnorm(x,0,1,lower.tail=T))}
alphas <- c(alphaTight,alphaTrad,alphaLoose)
myCritValues1 <- sapply(alphas,myCritValueF)
myCritValues1

[1] -2.326348 -1.644854 -1.281552
```

# Two sided Z tests

- Here are two sided Z-tests

```
myCritValueF2 <- function(x){  
  return(qnorm(x/2,0,1,lower.tail=T))}
```

```
myCritValues2 <- sapply(alphas,myCritValueF2)  
myCritValues2
```

```
[1] -2.575829 -1.959964 -1.644854
```

# Critical Values of T tests

- depend on sample size as well

```
my3DegTtest <- qt(alphaTrad/side,3,lower.tail = T)
my30DegTtest <- qt(alphaTrad/side,30,lower.tail = T)
my120DegTtest <- qt(alphaTrad/side,120,lower.tail = T)
c(my3DegTtest,my30DegTtest,my120DegTtest)
```

```
[1] -3.182446 -2.042272 -1.979930
```

# T Distribution Table

$t(df, \alpha/2)$  = two tailed  $t(df, \alpha)$  = one tailed



# In class exercise

- Find the following:
  - $t(20, .05/2) =$
  - $t(30, .025) =$
  - $t(121, .05/2) =$

# Research Question

- A researcher wants to evaluate whether group therapy using Cognitive Behavioral Therapy (CBT) effects the depression level of patients.
- Assume that the average depression level of the population (depressive patients) is  $\mu = 15$ .
- In order to evaluate the effect of CBT group therapy, the researcher randomly selected a group of 9 patients ( $n=9$ ) from the population to attend the group therapy.
- Assume that after the group therapy, the sample mean = 10; the sample standard deviation = 3.428. Based on this result, is it reasonable to conclude that the CBT group therapy has an treatment effect?

# Null Hypothesis

- CBT group therapy has no treatment effect on depressive patients
- Patients who receive CBT group therapy will have a mean depression level equal to the population  $H_0 : \mu = 15$
- Research Hypothesis
- CBT group therapy has a treatment effect on depressive patients
- Patients who receive CBT group therapy will have a different mean depression level than the population  $H_1 : \mu \neq 15$

## Find the critical value (Critical t-value)

- by using the function `qt` as t-distribution table
- Information needed for finding the critical value
  - $\alpha = .05$  for a two-tailed test
- Degrees of freedom ( $d.f = n-1$ ):  $9-1 = 8$  ( $d.f = 8$ )
- The critical values for a two-tailed t-test with  $d.f = 8$  are  $\pm 2.306$

```
my8DegTtest <- qt(alphaTrad/side,8,lower.tail = T)
c(-my8DegTtest,my8DegTtest)
```

```
[1] 2.306004 -2.306004
```

# In order to see if our sample mean is in the critical region

- we have to convert it into a t-score by using the following formula

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{N-1}}}$$

## Step 3: Compute the Test Statistic in R

```
xBar<-10  
mu<-15  
s<-3.428  
N<-9  
tStat<-(xBar-mu)/(s/sqrt(N-1))  
tStat  
  
[1] -4.125477
```

# Compare the obtained t-value and critical-value

- The obtained t-value: -4.13
- The critical t-value: -2.306
- Since the obtained t-value, -4.13, is less than the critical t-value of, -2.306, we reject the null hypothesis and conclude that the CBT group therapy has an effect on the depression level of the patients.

# A researcher wants to compare the mean number of prior arrests of individuals

- at Riker's Island to the population of the entire United States. Using Census data, we are confident that the mean number of prior arrests for everyone in the United States population is  $\mu = 4$ .

$$H_0 = \mu = 4.$$

$H_1 = \mu \neq 4$ . Open the dataset "Rikers 1989\_\_sex recode.SAV", use the variable "prior", and run a one sample t-test.

```
library(foreign)
mu <- 4
df <- read.spss(file = "../stats1/rikers1989.sav", to.data.frame = TRUE)
Xbar <- mean(df$prior)
N <- length(df$prior)
sampleSD <- function(x){
```



# Question: Is average age of homicide victims is different

- than the average age of the population of the entire United States. Your professor knows that the most recent statistics show that the average age of everyone in the United States population is  $\mu = 35.3$  years.
- $H_0 = \mu = 35.3$ .
- $H_1 = \mu \neq 35.3$ .
- Open the dataset “NYC Homicide 2005”, use the variable “Age”, and run a one sample t-test.

- The common role of hypothesis testing in a study is to detect significant differences
- Significant differences: differences that are not caused by chance/ sampling error.
- Or, differences that are caused by systematic influences (e.g., treatment effect)

# Review of Z-tests

- One-sample Z-tests: Used to compare a (large) sample mean to a population mean
- E.g. the mean  $\neq$  arrests of people in jail vs. the general public
- E.g. the average age of homicide victims vs. the general public