Introductory Statistics Lectures

# Testing a claim about a population mean

One sample hypothesis test of the mean

Anthony Tanbakuchi
Department of Mathematics
Pima Community College



(Compile date: Tue May 19 14:50:42 2009)

## Contents



## 1 Testing a claim about a population mean

### 1.1 Introduction

*Example* 1. The Carolina Tobacco Company advertised that its best selling non-filtered cigarettes contain at most 40 mg of nicotine, but Consumer Advocate magazine ran tests of 10 randomly selected cigarettes and found the amounts (in mg):

47.3, 39.3, 40.3, 38.3, 46.3, 43.3, 42.3, 49.3, 40.3, 46.3

It's a serious matter to charge that the company advertising is wrong, so the magazine editor wants to be extremely careful not to make a false claim. Can the editor prove that the mean nicotine level is greater than 40 mg?

*Question* 1. What would $H_0$ and $H_a$ be?

*Question* 2. What $\alpha$ should be used?

REVIEW

**What is a hypothesis test?**
A hypothesis test calculates the probability of observing our sample data **assuming the null hypotheses $H_0$ is true**. If the probability is low enough ($p$-value $\leq \alpha$), we reject $H_0$ and have evidence to support our alternative hypothesis $H_a$.

**Eight simple steps**

0. Write down what is **known**.
1. Determine which type of hypothesis **test** to use.
2. Check the test's **requirements**.
3. Formulate the **hypothesis**: $H_0$, $H_a$
4. Determine the **significance level** $\alpha$.
5. Find the $p$**-value**.
6. Make the **decision**.
7. State the final **conclusion**.

K-T-R-H-S-P-D-C: "Know The Right Hypothesis So People Don't Complain"

**What is the $p$-value?**
The $p$-value represents the probability of observing our sample data **assuming** $H_0$ is true. We use the sampling distribution to determine if sampling error could explain our observed sample statistic's deviation from $H_0$'s claim about the parameter. If we decide to reject $H_0$, **the $p$-value is the actual Type I error** for the study data.[1]

The $p$-value is calculated using the test statistic and it's corresponding distribution.

**Common form of a test statistic**

$$\text{test statistic} = \frac{(\text{sample statistic}) - (\text{null hypothesis of parameter})}{(\text{standard deviation of sample statistic})} \tag{1}$$

**What are the types of errors?**

**Type I error $\alpha$ / $p$-value** occurs when we **reject** $H_0$ but $H_0$ **is actually true**. The $p$-value is the Type I error. (ex. Convicting a innocent person — when $H_0$ is innocence.)

**Type II error $\beta$** occurs when we **fail to reject** $H_0$ but $H_0$ **is actually false**. (ex. Letting a criminal go free — when $H_0$ is innocence.)

## 1.2 Testing a claims when $\sigma$ is known

USE

**Often used to help answer:**

[1] The $\alpha$ we use to make our decision is just the **maximum** Type I error we will accept. It is the significance level, not the actual Type I error.

1. Is the mean of a population equal to $\mu_0$?
2. Is the mean of a population different than $\mu_0$?

## COMPUTATION

ONE SAMPLE HYPOTHESIS TEST, $\sigma$ KNOWN. DEFINITION 1.1

**requirements** (1) simple random sample, (2) $\sigma$ known, (3) C.L.T. applies.

**null hypothesis** $H_0 : \mu = \mu_0$

**alternative hypothesis** (1) $H_a : \mu \neq \mu_0$, (2) $H_a : \mu < \mu_0$, (3) $H_a : \mu > \mu_0$

**test statistic** : described by the $z$ distribution

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \tag{2}$$

*Question* 3. What does the test statistic $z$ represent?

**Finding the $p$-value "manually"**

Since there is no R function to do hypothesis tests when $\sigma$ is known, we need to manually calculate the $p$-value. To find a $p$-value for any hypothesis test:

1. Calculate the test statistic.
2. Using the distribution of the test statistic:
   - If $H_a$ contains $<$: $p$-value is the **area in the lower tail** bounded by the test statistic.
   - If $H_a$ contains $>$: $p$-value is the **area in the upper tail** bounded by the test statistic.
   - If $H_a$ contains $\neq$: $p$-value is **double** the tail area bounded by the test statistic. If the test statistic is negative use lower tail area. If the test statistic is positive use upper tail area.

*Question* 4. What distribution are we using when finding a $p$-value?

*Question* 5. What is the distribution of the test statistic $z$?

## A COMPLETE EXAMPLE

Back to our original example:

*Example* 2. The Carolina Tobacco Company advertised that its best selling non-filtered cigarettes contain at most 40 mg of nicotine, but Consumer Advocate magazine ran tests of 10 randomly selected cigarettes and found the amounts (in mg):

47.3, 39.3, 40.3, 38.3, 46.3, 43.3, 42.3, 49.3, 40.3, 46.3

It's a serious matter to charge that the company advertising is wrong, so the magazine editor chooses a significance level of $\alpha = 0.01$. Can the editor prove that the mean nicotine level is greater than 40 mg?

Assume nicotine levels are normally distributed and $\sigma = 3.8$.

**Solving the problem: step-by-step**

**Step 0** known information:

```
R: x = c(47.3, 39.3, 40.3, 38.3, 46.3, 43.3, 42.3,
+      49.3, 40.3, 46.3)
R: sigma = 3.8
```

**Step 1** Test: single sample mean test with **sigma known**.

**Step 2** Requirements: (1) simple random sample, (2) $\sigma$ known, (3) CLT. **All ok!**

**Step 3** Hypothesis: $H_0 : \mu = 40mg$, $H_a : \mu > 40mg$

```
R: mu0 = 40
```

**Step 4** Significance: $\alpha = 0.01$

**Step 5** $p$-value: must do manually find (1) test statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

```
R: x.bar = mean(x)
R: x.bar
[1] 43.3
R: n = length(x)
R: n
[1] 10
R: z = (x.bar - mu0)/(sigma/sqrt(n))
R: z
[1] 2.7462
```

(2) find $p$-value: this is one-tailed with $H_a :>$, find area in upper tail:

```
R: p.val = 1 - pnorm(z)
R: p.val
[1] 0.0030146
```

**Step 6** Decision: since $p$-val $\leq \alpha$ **reject** $H_0$. $(0.00301 \leq 0.01)$

**Step 7** Conclusion: "The sample data supports the claim that the mean level of nicotine is greater than 40 mg. The probability of observing our data by chance if the mean level was 40 mg was true is only 0.00301."

*Question* 6. What is the probability that we made a Type I Error (rejected the null hypothesis when it is actually true)?

## 1.3 Testing a claim when $\sigma$ is unknown

ONE SAMPLE HYPOTHESIS TEST, $\sigma$ NOT KNOWN. DEFINITION 1.2

**requirements** (1) simple random sample, (2) $\sigma$ not known, (3) C.L.T. applies.

**null hypothesis** $H_0 : \mu = \mu_0$

**alternative hypothesis** (1) $H_a : \mu \neq \mu_0$, (2) $H_a : \mu < \mu_0$, (3) $H_a : \mu > \mu_0$

**test statistic** : described by the $t$-distribution

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \qquad df = n - 1 \tag{3}$$

**If you're not given the raw data, you must use the test statistic.**

*Question* 7. What distribution do you use to calculate the $p$-value?

*Question* 8. A study was conducted. From the 35 random samples, $\bar{x} = 50.8$, $s = 2.8$. Find the test statistic and $p$-value if $H_a : \mu > 50$. (Check: $t = 1.69$, $p$-value=0.0501)

If you have the raw data, then use R:

ONE SAMPLE HYPOTHESIS TEST ($\sigma$ UNKNOWN): R COMMAND

```
t.test(x, mu=0, alternative="two.sided")
```

**x vector** of sample data

**mu** represents the null hypothesis value of $\mu_0$ (default 0).

**alternative** $H_a$ $\neq$:`"two.sided"`, $<$:`"less"`, $>$:`"greater"`

A COMPLETE EXAMPLE

Back to our original example

*Example* 3. The Carolina Tobacco Company advertised that its best selling non-filtered cigarettes contain at most 40 mg of nicotine, but Consumer Advocate magazine ran tests of 10 randomly selected cigarettes and found the amounts (in mg):

$$47.3, 39.3, 40.3, 38.3, 46.3, 43.3, 42.3, 49.3, 40.3, 46.3$$

It's a serious matter to charge that the company advertising is wrong, so the magazine editor chooses a significance level of $\alpha = 0.01$. Can the editor prove that the mean nicotine level is greater than 40 mg?

Now $\sigma$ is **not** known!
Assume nicotine levels are normally distributed.

**Solving the problem: step-by-step**
**Step 0** known information:

```
R: x = c(47.3, 39.3, 40.3, 38.3, 46.3, 43.3, 42.3,
+      49.3, 40.3, 46.3)
```

**Step 1** Test: single sample mean test with **sigma NOT known**.
**Step 2** Requirements: (1) simple random sample, (2)$\sigma$ not known, (3) CLT **all ok!**
**Step 3** Hypothesis: $H_0 : \mu = 40mg$, $H_a : \mu > 40mg$

```
R: mu0 = 40
```

**Step 4** Significance: $\alpha = 0.01$

**Step 5** $p$-value: can use R this time

```
R: res = t.test(x, mu = mu0, alternative = "greater")
R: res
        One Sample t-test

data:  x
t = 2.7458, df = 9, p-value = 0.01132
alternative hypothesis: true mean is greater than 40
95 percent confidence interval:
 41.097     Inf
sample estimates:
mean of x
     43.3
```

**Step 6** Decision: since $p$-val $\not\leq \alpha$ **fail to reject** $H_0$. $(0.0113 \not\leq 0.01)$

**Step 7** Conclusion: (**different than last time!**) "The sample data does not contradict the claim that the mean level of nicotine is 40 mg at the 0.01 significance level. The likelihood of observing our sample data by chance assuming the mean level is 40 mg was 0.0113."

*Question* 9. Why did we have a different conclusion this time?

DETERMINING REQUIRED SAMPLE SIZE

**Before** you conduct a hypothesis test (post-hoc), you **must** determine the necessary sample size to detect the smallest effect you are interested in.

**One sample $t$-test power**
Given the following parameters, you can determine the required $n$:
$\alpha$ significance level, typically 0.05.
**Power=**$1-\beta$ desired power, typically 0.8 or 0.9.
$h=\mu-\mu_0$ smallest difference in actual mean $\mu$ from the null hypothesis mean $\mu_0$ you want to detect. (Effect size.)
$\sigma$ population standard deviation (or best estimate).

$t$-TEST REQUIRED SAMPLE SIZE:
`power.t.test(delta=`$h$`, sd =`$\sigma$`, sig.level=`$\alpha$`, power=1 − `$\beta$`, type ="one.sample", alternative="two.sided")`

R COMMAND

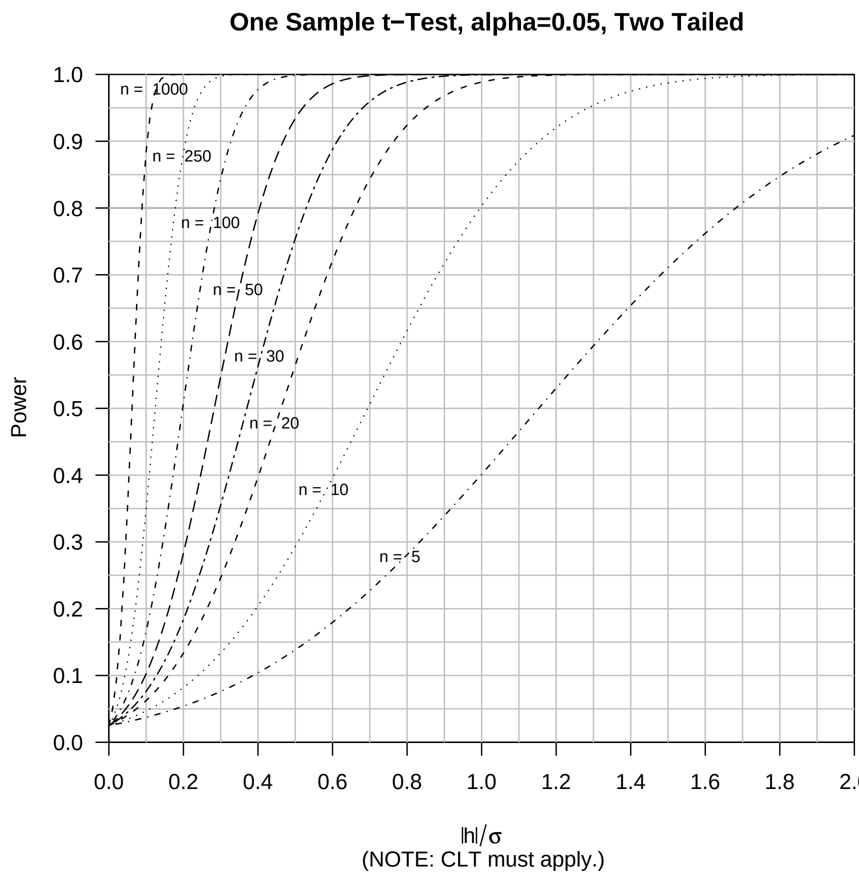
- **delta** minimum effect size of interest
- **sd** estimate of population standard deviation
- **power** desired power
- **alternative** set to either "one.sided" or "two.sided"

*Example* 4. Consumer Advocate magazine believes that the nicotine content in the Carolina Tobacco Company's best selling non-filtered cigarettes contain more than the advertised 40 mg of nicotine per cigarette. A reporter wants to design a study to to determine if the claim is false, how many cigarettes should be randomly sampled?. The editor won't run the article unless the cigarets actually have 43 mg or more of nicotine. It's a serious matter to charge that the company advertising is wrong, so the magazine editor chooses a significance level of $\alpha = 0.01$. Prior study data indicates that $\sigma = 3.8$ mg for these cigarettes. What is the appropriate sample size?

Known Info: $H_0 : \mu = 40$, $H_a : \mu > 40$ (one sided), minimum effect interested in detecting $h = 43 - 40 = 3$ mg, $\sigma = 3.8$ mg, $\alpha = 0.01$, choose Power$= 0.8$.

```
R: power.t.test(delta = 3, sd = 3.8, sig.level = 0.01,
+       power = 0.8, type = "one.sample", alternative = "one.sided"↩
    )
     One-sample t test power calculation

              n = 18.897
          delta = 3
             sd = 3.8
```




```
 sig.level = 0.01
     power = 0.8
alternative = one.sided
```

Thus, required sample size is $n = 19$.

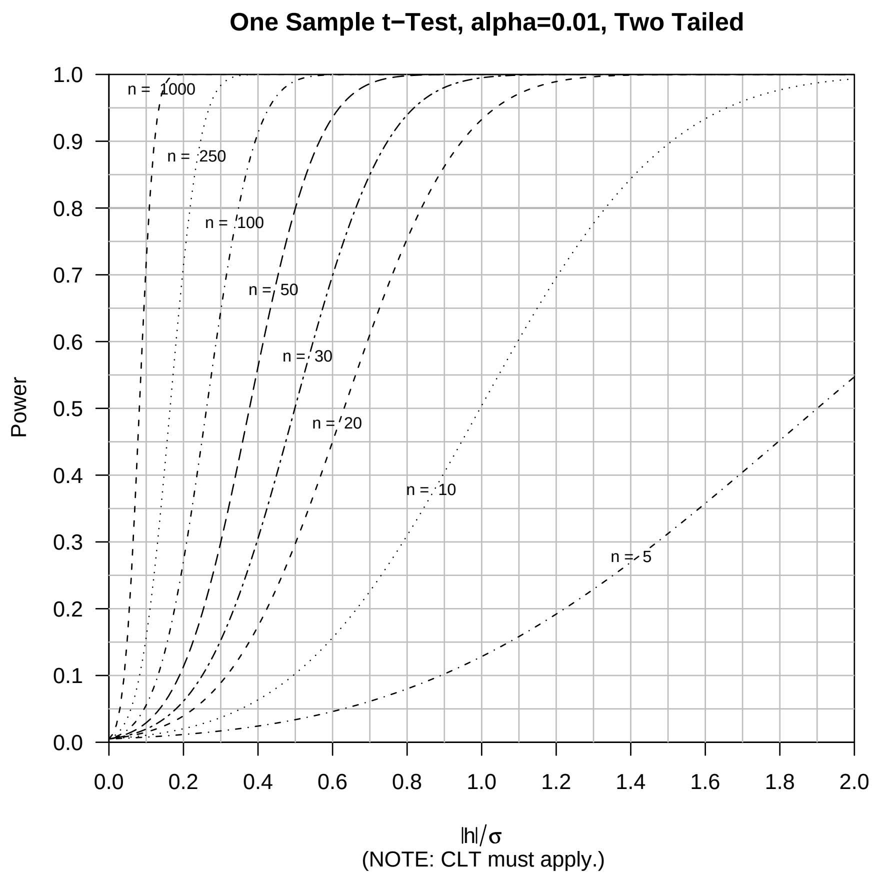
Since our original sample only used $n = 10$, our power was much less than 0.8.

## OC CURVES FOR $T$-TEST

**Definition 1.3** Operating Characteristic Curves.

OC curves for various hypothesis tests provide a simple method for quickly estimating the required sample size (ad-hoc) necessary to achieve a desired power for a given hypothesis test. They are also used to estimate the actual Type II error (pos-hoc) once a study has been conducted.

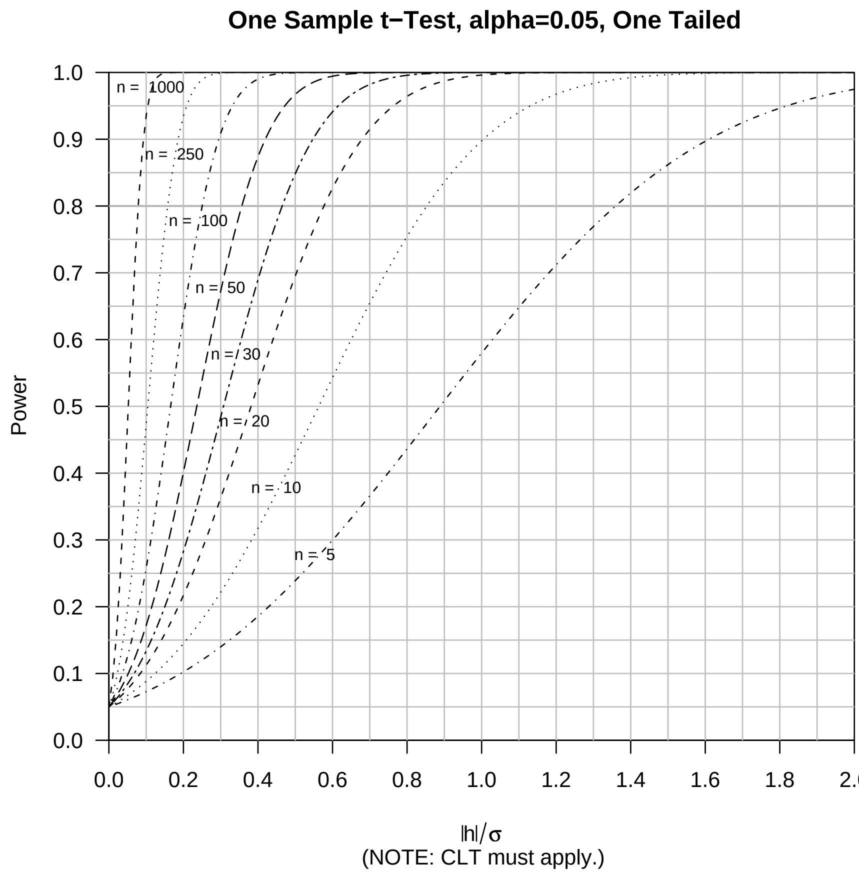
Now we will repeat our previous example using OC curves.



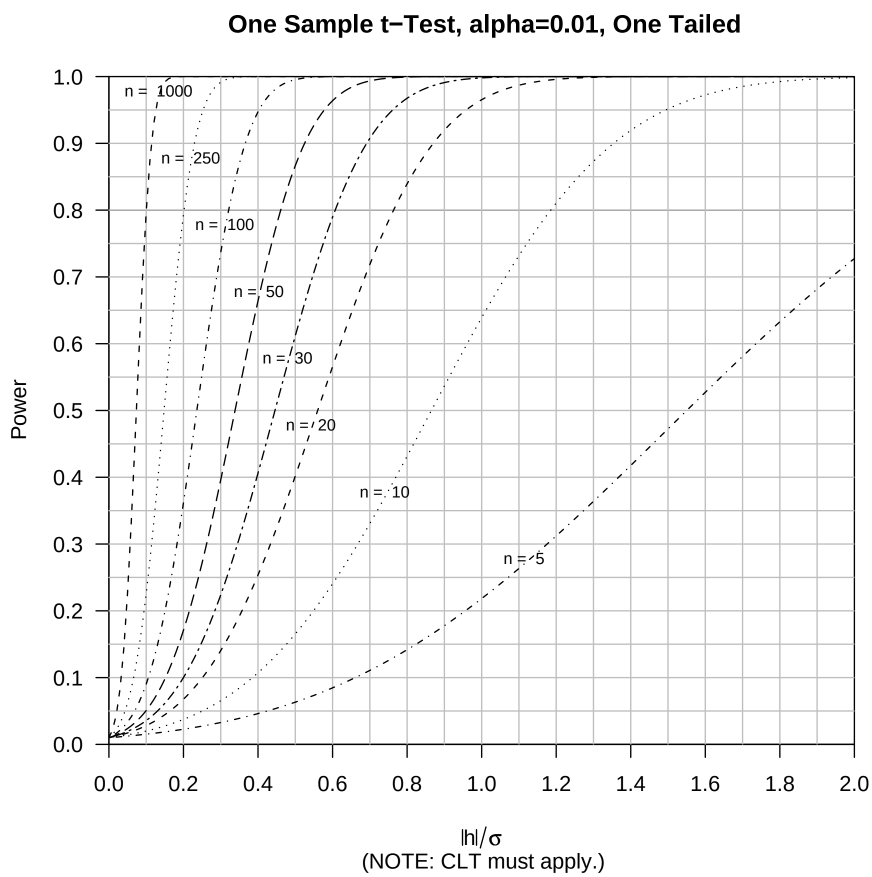

*Example* 5. Consumer Advocate magazine believes that the nicotine content in the Carolina Tobacco Company's best selling non-filtered cigarettes contain more than the advertised 40 mg of nicotine per cigarette. A reporter wants to design a study to to determine if the claim is false, how many cigarettes should be randomly sampled?. The editor won't run the article unless the cigarets actually have 43 mg or more of nicotine. It's a serious matter to charge that the company advertising is wrong, so the magazine editor chooses a significance level of $\alpha = 0.01$. Prior study data indicates that $\sigma = 3.8$ mg for these cigarettes. What is the appropriate sample size?

Known Info: $H_0 : \mu = 40$, $H_a : \mu > 40$ (one sided), minimum effect interested in detecting $h = 43 - 40 = 3$ mg, $\sigma = 3.8$ mg, $\alpha = 0.01$, choose Power$= 0.8$.

*Question* 10. Use the following OC curves to estimate the required sample size.




*Question* 11. The study was conducted with a sample size of only 10, the sample standard deviation for the data was 3.8 mg. Use the OC curves to estimate the actual power and Type II error.


One Sample t−Test, alpha=0.01, One Tailed
Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
n = 1000
n = 250
n = 100
n = 50
n = 30
n = 20
n = 10
n = 5
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
|h|/σ
(NOTE: CLT must apply.)

## 1.4 Summary

**Hypothesis test for** $\mu$
**requirements** (1) simple random sample, (2) C.L.T. applies.
**null hypothesis** $H_0 : \mu = \mu_0$
**alternative hypothesis** (1) $H_a : \mu \neq \mu_0$, (2) $H_a : \mu < \mu_0$, (3) $H_a : \mu > \mu_0$
**test** when $\sigma$ **known**:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

See page 3 for steps to find $p$-value manually.
**test** when $\sigma$ **unknown**:
`t.test(x, mu=`$\mu_0$`, alternative="two.sided")`
**x vector** of sample data
**mu** represents the null hypothesis value of $\mu_0$ (default 0).
**alternative** $H_a$ $\neq$:`"two.sided"`, $<$:`"less"`, $>$:`"greater"`
If your not given the raw data, you must use the test statistic on page 5.

## 1.5 Additional Examples

*Example* 6. A cereal lobbyist claims that the mean sugar content in cereal (grams of sugar per gram of cereal) is less than -.35 g for all cereals. You believe it is higher. If the mean content is at least 0.45 g you want to counter their claim.

*Question* 12. What is $H_0$ and $H_a$?

*Question* 13. What sample size should you use to test your claim (Previous data indicates $\sigma = 0.15$ g, desired power is 0.8)?

*Example* 7. Different cereals are randomly selected, and the sugar content (grams of sugar per gram of cereal) is obtained for each cereal, with the results given below. Use a 0.05 significance level to test the claim of a cereal lobbyist that the mean for all cereals is less than 0.35 g.

0.03, 0.24, 0.30, 0.47, 0.43, 0.07, 0.47, 0.13, 0.44, 0.39, 0.48, 0.17, 0.13, 0.09, 0.45, 0.43

*Question* 14. Run the hypothesis test assuming $\sigma = 0.15$.[2]

[2] CHECK: $z = -1.47$, $p$-value $= 0.0712$

*Question* 15. For the previous question, would would the $p$-value be if $H_a: \mu \neq 0.35$ g?

*Question* 16. Run the hypothesis test assuming $\sigma$ is unknown.[3]

[3] CHECK: $p$-value $= 0.105$