

MATH 1052H - S62 - Lecture 04

Inference for a Single Mean

Parameter and Point Estimation

Parameter of Interest: Mean of **all** [whatever your data is]

μ : a population proportion

Point Estimate: Average of **sampled** [whatever your data is]

\bar{x} : a sample mean

Formula: SE of a Point Estimate \bar{x}

When we have a **sample mean**, the standard error has a known formula:

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}}$$

What are μ and n ?

1. n is the number of samples (it's a **sample mean**)

2. μ

is the true underlying population mean ...

Sample Means are Almost Normally Distributed

Remember the Central Limit Theorem (CLT).

Sample means will be nearly normally distributed with mean equal to the population mean, μ , and standard error equal to $SE_{\bar{x}}$ from the last slide. We can write this formally.

$$\bar{x} \sim \mathcal{N} \left(\text{mean} = \mu, SE = \frac{\sigma}{\sqrt{n}} \right)$$

But, of course, this is only true under certain conditions ... same ones as usual!

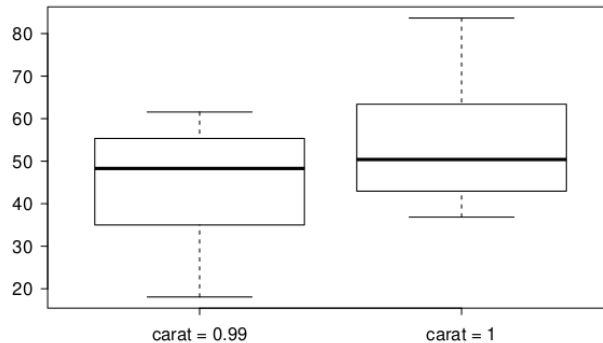
Difference of Two Means

Example: Diamonds

- Weights of diamonds are measured in carats.
- 1 carat = 100 points, 0.99 carats = 99 points, etc.
- The difference between the size of a 0.99 carat diamond and a 1 carat diamond is undetectable to the naked human eye, but does the price of a 1 carat diamond tend to be higher than the price of a 0.99 diamond?
- We are going to test to see if there is a difference between the average prices of 0.99 and 1 carat diamonds.
- In order to be able to compare equivalent units, we divide the prices of 0.99 carat diamonds by 99 and 1 carat diamonds by 100, and compare the average point prices.



Data



	0.99 carat	1 carat
\bar{x}	44.50	53.43
s	13.32	12.22
n	23	30

(variables are **pt99** and **pt100**, respectively)

Data (Source)

These data are a random sample from the **diamonds** data set in **ggplot2** R package. If you'd like to explore it,

```
library("ggplot2")  
data(diamonds)
```

Parameter and point estimate

- **Parameter of interest:** Average difference between the point prices of **all** 0.99 carat and 1 carat diamonds.

$$\mu_{pt99} - \mu_{pt100}$$

- **Point estimate:** Average difference between the point prices of **sampled** 0.99 carat and 1 carat diamonds.

$$\bar{x}_{pt99} - \bar{x}_{pt100}$$

Hypotheses

Which of the following is the correct set of hypotheses for testing if the average point price of 1 carat diamonds (μ_{pt100}) is higher than the average point price of 0.99 carat diamonds (μ_{pt99})?

- $H_0 : \mu_{pt99} = \mu_{pt100}$ versus $H_A : \mu_{pt99} \neq \mu_{pt100}$
- $H_0 : \mu_{pt99} = \mu_{pt100}$ versus $H_A : \mu_{pt99} > \mu_{pt100}$
- $H_0 : \mu_{pt99} = \mu_{pt100}$ versus $H_A : \mu_{pt99} < \mu_{pt100}$
- $H_0 : \bar{x}_{pt99} = \bar{x}_{pt100}$ versus $H_A : \bar{x}_{pt99} < \bar{x}_{pt100}$

Hypotheses

Which of the following is the correct set of hypotheses for testing if the average point price of 1 carat diamonds (μ_{pt100}) is higher than the average point price of 0.99 carat diamonds (μ_{pt99})?

- $H_0 : \mu_{pt99} = \mu_{pt100}$ versus $H_A : \mu_{pt99} \neq \mu_{pt100}$
- $H_0 : \mu_{pt99} = \mu_{pt100}$ versus $H_A : \mu_{pt99} > \mu_{pt100}$
- $H_0 : \mu_{pt99} = \mu_{pt100}$ versus $H_A : \mu_{pt99} < \mu_{pt100}$
- $H_0 : \bar{x}_{pt99} = \bar{x}_{pt100}$ versus $H_A : \bar{x}_{pt99} < \bar{x}_{pt100}$

Conditions

Which of the following does **not** need to be satisfied in order to conduct this hypothesis test using theoretical methods?

- Point price of one 0.99 carat diamond in the sample should be independent of another, and the point price of one 1 carat diamond should independent of another as well.
- Point prices of 0.99 carat and 1 carat diamonds in the sample should be independent.
- Distributions of point prices of 0.99 and 1 carat diamonds should not be extremely skewed.
- Both sample sizes should be at least 30.

Conditions

Which of the following does **not** need to be satisfied in order to conduct this hypothesis test using theoretical methods?

- Point price of one 0.99 carat diamond in the sample should be independent of another, and the point price of one 1 carat diamond should independent of another as well.
- Point prices of 0.99 carat and 1 carat diamonds in the sample should be independent.
- Distributions of point prices of 0.99 and 1 carat diamonds should not be extremely skewed.
- *Both sample sizes should be at least 30.*

Sampling distribution for the difference of two means

Test statistic

Test statistic for inference on the difference of two small sample means:

The test statistic for inference on the difference of two means where σ_1 and σ_2 are unknown is the T statistic:

$$T_{df} = \frac{\text{point estimate} - \text{null value}}{SE}$$

where

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad \text{and} \quad df = \min(n_1 - 1, n_2 - 1)$$

The calculation of the df is actually **much** more complicated. For simplicity we'll use the above formula to **estimate** the true df when conducting the analysis by hand. See SLIDE BLAH FOR MORE.

Test statistic (cont.)

	0.99 carat	1 carat
\bar{x}	44.50	53.43
s	13.32	12.22
n	23	30

in context...

Test statistic (cont.)

	0.99 carat	1 carat
\bar{x}	44.50	53.43
s	13.32	12.22
n	23	30

in context...

$$T = \frac{\text{point estimate} - \text{null value}}{SE} = \frac{(44.50 - 53.43) - 0}{\sqrt{\frac{13.32^2}{23} + \frac{12.22^2}{30}}}$$

Doing It

```
s1 <- 13.32
n1 <- 23
s2 <- 12.22
n2 <- 30
xbar1 <- 44.50
xbar2 <- 53.43
SE <- sqrt( s1^2/n1 + s2^2/n2 )
h0 <- 0
T <- ((xbar1 - xbar2) - h0) / SE
T
```

```
## [1] -2.506644
```

Test statistic (cont.)

Which of the following is the correct df for this hypothesis test?

- 22
- 23
- 30
- 29
- 52

Test statistic (cont.)

Which of the following is the correct df for this hypothesis test?

- 22
- 23
- 30
- 29
- 52

$$df = \min(n_{pt99} - 1, n_{pt100} - 1) = \min(23 - 1, 30 - 1) = 22.$$

p-value

We know the p-value and the test statistic: let's do this!

$$T = -2.508 \quad df = 22$$

```
pt(q = -2.508, df = 22, lower.tail = TRUE)
```

```
## [1] 0.0100071
```

We use `lower.tail = TRUE` because the alternative was $H_A : \mu_{pt99} < \mu_{pt100}$.

Synthesis

What is the conclusion of the hypothesis test? How (if at all) would this conclusion change your behavior if you went diamond shopping?

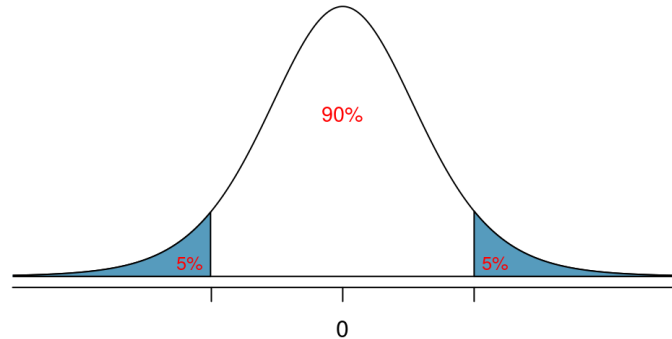
- p-value is small so reject H_0 . The data provide convincing evidence to suggest that the point price of 0.99 carat diamonds is lower than the point price of 1 carat diamonds.
- Maybe buy a 0.99 carat diamond? It looks like a 1 carat, but is significantly cheaper.

Confidence intervals for the difference of two means

Equivalent confidence level

What is the equivalent confidence level for a one-sided hypothesis test at $\alpha = 0.05$?

- 90%
- 92.5%
- 95%
- 97.5%



Equivalent confidence level

What is the equivalent confidence level for a one-sided hypothesis test at $\alpha = 0.05$?

- 90%
- 92.5%
- 95%
- 97.5%

One-sided tests with specific alpha aren't identical to the corresponding confidence levels!

Critical value

What is the appropriate t^* for a confidence interval for the average difference between the point prices of 0.99 and 1 carat diamonds?

- 1.32
- 1.72
- 2.07
- 2.82

```
qt(p = (1 - 0.90) / 2, df = 22, lower.tail = FALSE)
```

```
## [1] 1.717144
```

Confidence interval

Calculate the interval, and interpret it in context.

$$\text{point estimate} \pm ME$$

Filled in, this is:

$$(\bar{x}_{pt99} - \bar{x}_{pt1}) \pm t_{df}^* \times SE = (44.50 - 53.43) \pm 1.72 \times 3.56$$

```
(44.50 - 53.43) + c(-1, 1) * 1.717144 * 3.56
```

```
## [1] -15.043033 -2.816967
```

Thus, we are 90% confident that the average point price of a 0.99 carat diamond is \$15.04 to \$2.82 lower than the average point price of a 1 carat diamond.

Recap: Inference using difference of two small sample means

- If σ_1 or σ_2 is unknown, difference between the sample means follow a t -distribution with $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$.
- Conditions:
 - independence within groups (often verified by a random sample, and if sampling without replacement, $n < 10\%$ of population) and between groups
 - no extreme skew in either group

Recap: Inference using difference of two small sample means

- Hypothesis testing:

$$T_{df} = \frac{\text{point estimate} - \text{null value}}{SE}, \text{ where } df = \min(n_1 - 1, n_2 - 1)$$

- Confidence interval:

$$\text{point estimate} \pm t_{df}^* \times SE$$

Advanced: Using R, Skipping the df Approximation

Advanced Option

We used an approximation for the degrees-of-freedom in the earlier discussion, because we needed a convenient way to approximate the df without using a really complicated formula. However, when you're using the built-in functions in R, it does it for you!

So, in this example, we'll show you how this might shake out. The only things that get influenced are p-values (for hypothesis tests) and the critical (star) values (for confidence intervals).

Some Data

In a spreadsheet called `grades.csv`, I have saved a random sample of final (completely anonymous!) grades from MATH 1051H (Fall, 2019) and MATH 1051H (Summer, 2019). Just grades, no student info. I'm interested in whether the overall average performance is the same between Fall and Summer. Let's do a test!

```
grades <- read.csv("grades.csv", header = TRUE)
str(grades)
```

```
## 'data.frame':   40 obs. of  2 variables:
## $ Fall   : int  93 87 57 75 96 70 50 76 100 87 ...
## $ Summer: int  54 70 69 79 91 46 66 88 84 66 ...
```

Manually

I assume:

$$H_0 : \mu_1 - \mu_2 = 0 \quad \text{versus} \quad H_A : \mu_1 - \mu_2 \neq 0$$

```
xbar1 <- mean(grades$Fall, na.rm = TRUE)
s1 <- sd(grades$Fall, na.rm = TRUE)
n1 <- length(which(!is.na(grades$Fall)))
xbar2 <- mean(grades$Summer, na.rm = TRUE)
s2 <- sd(grades$Summer, na.rm = TRUE)
n2 <- length(which(!is.na(grades$Summer)))
SE <- sqrt( s1^2/n1 + s2^2/n2 )
h0 <- 0
T <- ((xbar1 - xbar2) - h0) / SE
T
```

```
## [1] -0.5169846
```

Conclusion (Manual)

So we compute the p-value:

```
df <- min(n1 - 1, n2 - 1)
pt(T, df = df, lower.tail = TRUE) * 2
```

```
## [1] 0.6085139
```

So this p-value is large, and we fail to reject the null: we do not have evidence of a difference in the average final grade for MATH 1051H for Fall versus Summer.

Automatic

```
t.test(x = grades$Fall, y = grades$Summer,  
       alternative = "two.sided", mu = 0)
```

```
##  
## Welch Two Sample t-test  
##  
## data:  grades$Fall and grades$Summer  
## t = -0.51698, df = 72.951, p-value = 0.6067  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
##  -8.773833  5.159547  
## sample estimates:  
## mean of x mean of y  
##  72.25000  74.05714
```

See the output? We get exactly the same test statistic, but the df is estimated to be 72.951 instead of the smaller of $n_1 - 1$ and $n_2 - 1$. This changes the p-value from 0.6085 to 0.6067.

Conclusion

So be careful and use the method indicated by the problem you're doing. Mostly, we will use the simple approximation, but look for WeBWork problems that tell you not to do this!