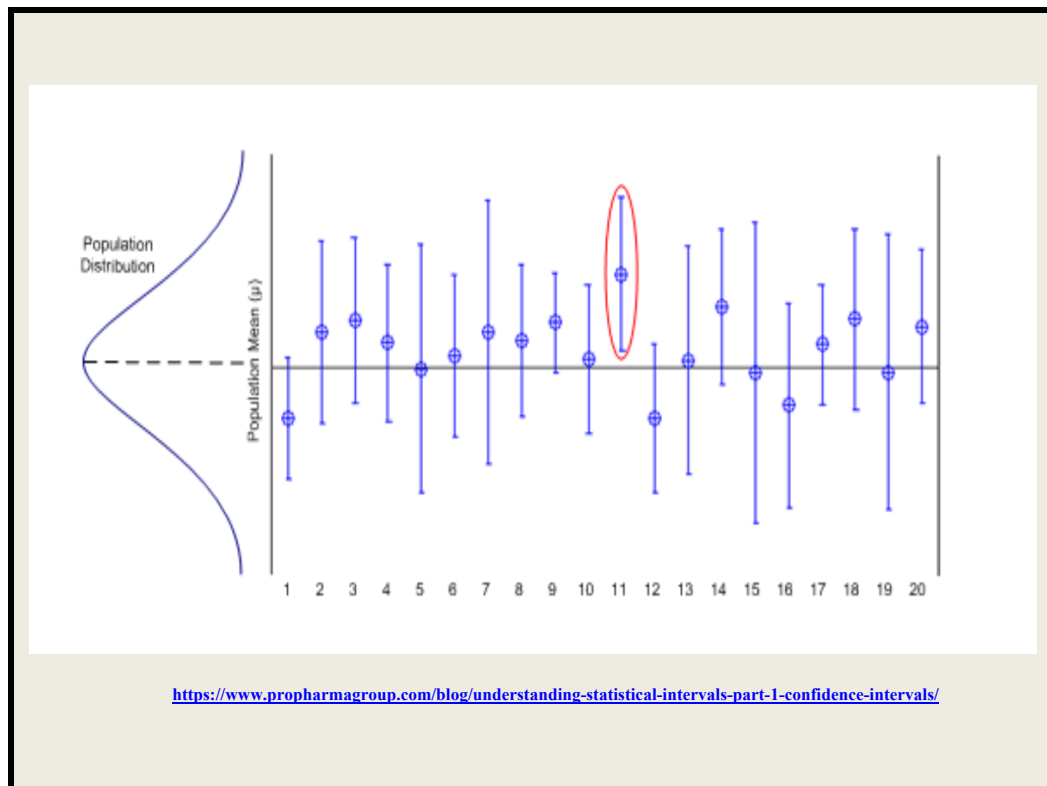


Unit 9. ONE Sample Inference

“Use at least twelve observations in constructing a confidence interval”

- Gerald van Belle



Cheers!



“ If you use a bigger net (99% confidence interval), you will have higher confidence that your net has the fish. But your boss won’t be happy to receive a 100 feet long fish net for a dead fish. Point: You get accuracy with high confidence interval but you lose precision”



“ If you use a small fish net (90% confidence interval), it will be easy for you to carry it and hand it over to your boss, but you won’t have that high confidence of being successful in your mission. Point: You get precision but you lose accuracy with low confidence interval”

(Source: Michael Hochster <https://www.quora.com/What-is-a-confidence-interval-in-laymans-terms>)

Nature ——— Population/ Sample ——— Observation/ Data ——— Relationships/ Modeling ——— Analysis/ Synthesis

Table of Contents

Topic		
	1. Learning Objectives	4
	2. One Sample - Normal Distribution: μ	5
	a. Confidence Interval for μ , σ^2 Known	5
	b. Hypothesis Test for μ , σ^2 Known	10
	c. Hypothesis Test for μ , σ^2 Known – <i>Critical Region Approach</i> ..	12
	d. Confidence Interval for μ , σ^2 Unknown	17
	e. Hypothesis Test for μ , σ^2 Unknown	21
	3. One Sample - Normal Distribution: σ^2	25
	a. Confidence Interval for σ^2	25
	b. Hypothesis Test for σ^2	30
	4. One Sample <i>Paired Data</i> - Normal Distribution: $\mu_{\text{DIFFERENCE}}$	34
	a. Introduction to Paired Data	34
	b. Confidence Interval for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ Known	36
	c. Hypothesis Test for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ Known	39
	d. Confidence Interval for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ Unknown	42
	e. Hypothesis Test for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ Unknown	46
	f. Confidence Interval for $\sigma^2_{\text{DIFFERENCE}}$	51
	5. One Sample Binomial Distribution:	55
	a. Confidence Interval for π , normal approximation approach	55
	b. Hypothesis Test for π – Exact	60
	c. Hypothesis Test for π – Normal Approximation	65
	Appendices	
	i. Derivation of Confidence Interval for μ – Single Normal with σ^2 Known	68
	ii. Derivation of Confidence Interval for σ^2 – Single Normal	71
	iii. Standard Error (SE) of a Binomial Proportion	73

1. Learning Objectives

When you have finished this unit, you should be able to:

- Calculate point and confidence interval estimates of the mean and variance of a single Normal distribution.
- Calculate point and confidence interval estimates of the mean and variance of a single Normal distribution in the paired data setting.
- Calculate point and confidence interval estimates of the π (event probability) parameter of a single binomial distribution.
- Perform and interpret the statistical hypothesis tests described in the one sample settings described in these notes.

2. One Sample – Normal Distribution: μ

2a. Confidence Interval for μ (σ^2 Known)

Tip -

For a single sample from a normal distribution with KNOWN σ^2 , a confidence interval for the mean utilizes percentiles of the **standard normal distribution**.

We want to compute a confidence interval estimate of μ for a population distribution that is normal with σ known. Available are data from a random sample of size= n .

- These pages show you how to construct a confidence interval.
- Appendix 1 gives the [statistical theory](#) underlying this calculation

1. The Point Estimate of μ is the Sample Mean \bar{X}_n

Recall that, for a sample of size= n , the sample mean is calculated as

$$\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$$

Features:

1. Under simple random sampling, the sample mean (\bar{X}_n) is an unbiased estimator of the population mean parameter μ , regardless of the underlying probability distribution.
2. When the underlying probability distribution is normal, the sample mean (\bar{X}_n) also satisfies the criterion of being minimum variance unbiased.

2. The Standard Error of \bar{X}_n is σ/\sqrt{n}

The precision of \bar{X}_n as an estimate of the unknown population mean parameter μ is reflected in its standard error. Recall:

$$SE(\bar{X}_n) = \sqrt{\text{variance}(\bar{X}_n)} = \frac{\sigma}{\sqrt{n}}$$

- ♣ SE is smaller for smaller σ (measurement error)
- ♣ SE is smaller for larger n (study design)

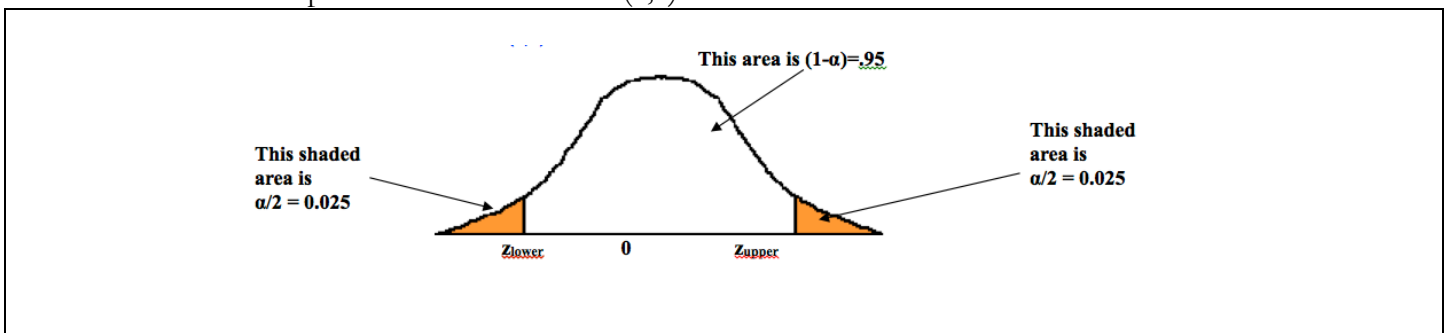
Introduction to the notation α and $(1 - \alpha)$

α = Probability the confidence interval will NOT include the population parameter
 $(1 - \alpha)$ = Probability the confidence interval WILL include the population parameter

3. The Confidence Coefficient

The **confidence coefficient** for a 95% confidence interval is the number needed to insure 95% coverage “in the long run” (in probability).

- ♣ 95% coverage *leaves 5% in the tails*. This is split evenly in the 2 tails. Thus, for a 95% confidence interval (5% in tails total / 2 tails) = **2.5% in each tail**. For a 95% confidence interval, this number will be the 97.5th percentile of the Normal (0,1) distribution = 1.96.



- ♣ For a $(1-\alpha)100\%$ confidence interval, this number will be the $(1-\alpha/2)100^{\text{th}}$ percentile of the Normal (0,1) distribution.
- ♣ The table below gives some of these values in the setting of constructing a confidence interval estimate of μ when data are from a Normal distribution with σ^2 known.

Confidence Level	Confidence Coefficient = Percentile Value from Normal (0,1)	
	Percentile	
.50	75 th	0.674
.75	87.5 th	1.15
.80	90 th	1.282
.90	95 th	1.645
.95	97.5 th	1.96
.99	99.5 th	2.576
$(1 - \alpha)$	$(1 - \alpha/2)100^{\text{th}}$	-

Nature ——— Population/ Sample ——— Observation/ Data ——— Relationships/ Modeling ——— Analysis/ Synthesis

Example -

The following data are the weights (micrograms) of drug inside each of 30 capsules, after subtracting the capsule weight. Suppose it is known that $\sigma^2 = 0.25$. Under the assumption of normality, calculate a 95% confidence interval estimate of μ .

0.6	0.3	0.1	0.3	0.3
0.2	0.6	1.4	0.1	0.0
0.4	0.5	0.6	0.7	0.6
0.0	0.0	0.2	1.6	-0.2
1.6	0.0	0.7	0.2	1.4
1.0	0.2	0.6	1.0	0.3

- ♣ The data are simple random sample of size $n=30$ from a Normal distribution with mean = μ and variance = σ^2 .
- ♣ The population variance is known and has value $\sigma^2 = 0.25$
- ♣ **Remark – In real life, we will rarely know σ^2 !** This example is for illustration only.

The solution for the confidence interval is point estimate \pm safety net:

$$\begin{aligned}\text{Lower limit} &= (\text{point estimate}) - (\text{multiple}) (\text{SE of point estimate}) \\ \text{Upper limit} &= (\text{point estimate}) + (\text{multiple}) (\text{SE of point estimate})\end{aligned}$$

Obtain the Point Estimate of μ = the Sample Mean $\bar{X}_{n=30}$

$$\bar{X}_{n=30} = \frac{\sum_{i=1}^n X_i}{n=30} = 0.51$$

Obtain the Standard Error of the Point Estimate = the Standard Error of $\bar{X}_{n=30}$ is σ/\sqrt{n}

$$SE(\bar{X}_{n=30}) = \sqrt{\text{variance}(\bar{X}_{n=30})} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{0.25}}{\sqrt{30}} = 0.0913$$

Obtain the Confidence Coefficient

For desired confidence = .95:

$$(1 - \alpha) = .95$$

$$\alpha = .05$$

$$\alpha/2 = .025 \rightarrow \text{Solve for the 2.5}^{\text{th}} \text{ percentile and the 97.5}^{\text{th}} \text{ percentile so that each tail area} = .025.$$

Using Art of Stat and its utility for the Normal distribution, we obtain 2.5th percentile = -1.96 and 97.5th percentile = +1.96

Desired Confidence Level	Value of Confidence Coefficients
.95	± 1.96

Putting this all together –

$$\begin{aligned} \text{Lower limit} &= (\text{point estimate}) - (\text{multiple}) (\text{SE of point estimate}) \\ &= 0.51 - (1.96) (0.0913) \\ &= 0.33 \end{aligned}$$

$$\begin{aligned} \text{Upper limit} &= (\text{point estimate}) + (\text{multiple}) (\text{SE of point estimate}) \\ &= 0.51 + (1.96) (0.0913) \\ &= 0.69 \end{aligned}$$

Thus, we have the following general formula for a $(1 - \alpha)100\%$ confidence interval -

$$\bar{X}_n \pm [(1 - \alpha/2)100^{\text{th}} \text{ percentile of Normal}(0,1)] \text{SE}(\bar{X}_n)$$

Art of Stat Illustration

Sorry! Art of Stat does not have a utility for a one sample confidence interval for a Normal distribution mean in the setting where the standard deviation is known.

However, there is a utility for a one sample confidence interval for a Normal distribution mean in the setting where the standard deviation is NOT known. It utilizes the Student t distribution. Stay tuned.

R Illustration

```
# EXAMPLE
ywt <- c(0.6,0.2,0.4,0.0,1.6,1.0,0.3,0.6,0.5,0.0,0.0,0.2,0.1,1.4,0.6,0.2,0.7,0.6,0.3,0.1,0.7,1.6,0.2,1.0,0.3,0.0,0.6,
0.2,1.4,0.3)

yn <- length(ywt)                                # sample size using length( )
ymean <- mean(ywt,na.rm=TRUE)                    # sample mean using mean( )
ysd <- sqrt(0.25)                                # known sd = sqrt(known variance)
ysem <- ysd/sqrt(yn)                              # standard deviation of the mean
multiplier <- qnorm(.975, mean=0, sd=1)          # for 95% CI, multiplier = 97.5th percentile of Normal(0,1)
CI_lower <- ymean - multiplier*ysem               # lower limit of 95% CI
CI_upper <- ymean + multiplier*ysem              # upper limit of 95% CI

table1 <- cbind(ymean,yn,ysd,CI_lower,CI_upper)
colnames(table1) <- c("mean", "n", "sd", "lower 95% CI", "upper 95% CI")
table1
```

```
      mean  n  sd lower 95% CI upper 95% CI
[1,] 0.51 30 0.5    0.3310806    0.6889194
```

HOMEWORK DUE Monday November 21, 2022

Question #1 of 8

Suppose we assume that the results of IQ tests are distributed normal. Suppose that in 2016, the distribution of IQ test scores for persons aged 18-35 years has a variance $\sigma^2 = 225$. A simple random sample of 9 persons take the IQ test. The sample mean score is 115. Calculate the 50%, 75%, 90% and 95% confidence interval estimates of the unknown population mean IQ score.

HOMEWORK DUE Monday November 21, 2022

Question #2 of 8

This exercise is asking you to think about, and compare, two aspects of the concept of a confidence interval: (1) its width, and (2) the level of confidence that we attach to the interval we are reporting.

Hint – precision versus confidence...

What trade-offs are involved in reporting one interval estimate over another?

2b. Hypothesis Test for μ (σ^2 Known)

Tip -

For a single sample from a normal distribution with KNOWN σ^2 , a hypothesis test about the mean utilizes a p-value calculation from the **standard normal distribution**.

An example of a test for μ , when data are from a normal distribution with σ^2 known has been presented previously.

- Therefore, an abbreviated presentation is given here (so that these notes are easy to read!)
- For full details, see again Unit 8 notes, pp 13-18.

Example –

With standard care, cancer patients are expected to survive a mean duration of time equal to 38.3 months. Hypothesized is that a new therapy will improve survival. In this study, the new therapy is administered to 100 cancer patients. Their average survival time is 46.9 months. Suppose σ^2 known = 43.3² months squared; this means that the standard deviation $\sigma = 43.3$ months. Is this statistically significant evidence of *improved* survival?

Nature ——— Population/ Sample ——— Observation/ Data ——— Relationships/ Modeling ——— Analysis/ Synthesis

Null Hypothesis Probability Model Assumptions.

X_1, X_2, \dots, X_{100} is a simple random sample from a Normal($\mu, \sigma^2 = 43.3^2$); the standard deviation $\sigma = 43.3$ months.

Specify the null (H_0) and alternative (H_A) hypotheses

$H_0: \mu_{true} = \mu_0 \leq 38.3$ months

$H_A: \mu_{true} = \mu_A > 38.3$ months one sided. This is because the investigator is researching “improvement”

Reason “proof by contradiction”

IF: we assume that the null hypothesis model is true, so that $\mu_{true} = \mu_0 = 38.3$

THEN: what are the chances of observing a sample mean (average) survival time as great as 46.9 months, or greater, relative to the null hypothesis expected mean survival time of 38.3 months?

Specify the p-value calculation. This is the “proof by contradiction” reasoning.

Statistically, the null hypothesis applied to the observed data will lead to an inconsistent conclusion if there is at most a small chance of a mean of 100 survival times being 46.9 or greater when the expected value is 38.3. We calculate the p-value as

$$\Pr[\bar{X}_{n=100} \geq 46.9 | \mu_{true} = \mu_0 = 38.3]$$

The test statistic is a Z-Score

Under the assumption that the null hypothesis is true:

- X_1, X_2, \dots, X_{100} is each distributed Normal($\mu, \sigma^2 = 43.3^2$). From there, we get
- $\bar{X}_{n=100}$ is distributed Normal ($\mu = 38.3, \sigma^2 = 43.3^2/100$)
- We’ll use as our test statistic a *particular z-score standardization* of $\bar{X}_{n=100}$. The standardization used is the one in which the population mean value subtracted is the one corresponding to the null hypothesis.

$$\text{Test Statistic} = \text{Z-score} = \frac{\bar{X}_{n=100} - \mu_{null}}{\text{SE}(\bar{X}_{n=100})} = \frac{\bar{X}_{n=100} - \mu_{null}}{\sqrt{\sigma^2 / n}} = \frac{46.9 - 38.3}{\sqrt{43.3^2 / 100}} = 1.99$$

Nature ——— Population/ Sample ——— Observation/ Data ——— Relationships/ Modeling ——— Analysis/ Synthesis

P-value calculation

$$\begin{aligned}\text{p-value} &= \Pr[\bar{X}_{n=100} \geq 46.9 \mid \mu_{\text{true}} = \mu_{\text{null}} = 38.3] \\ &= \Pr[Z\text{-score} \geq 1.99] = .02\end{aligned}$$

“Evaluate”.

IF the new therapy yields no improvement in survival so that the survival experience under the new therapy is identical to that experienced with receipt of standard care,

THEN there was a 2% chance of observing an average survival time as great or greater than the observed average survival time of 46.9 months.

Interpret.

The assumption of the null hypothesis and its application to our observations have led us to a very unlikely event (it had a 2% chance of occurrence) when considered in light of the observed data. This suggests abandoning the null hypothesis model. →. **Reject the null hypothesis.** Conclude that, compared to standard care, there is statistically significant evidence of longer survival times on the new treatment.

2c. Hypothesis Test for μ (σ^2 Known) – Critical Region Approach

Tip -

For a single sample from a normal distribution with KNOWN σ^2 , a hypothesis test about the mean utilizes a p-value calculation from the **standard normal distribution**.

See again Unit 8 for detail on the distinction between significance level testing and critical region testing.

Previously (pp 10-11), we utilized the **significance level approach** and we were asking the question:

- Under the assumption that the null hypothesis is true, what were my chances of obtaining a test statistic as extreme or more extreme?

Now we will utilize the **critical region approach**. Recall the thinking:

- **If** I assume that the null hypothesis is true,
- **And if** I agree that I will reject the null hypothesis under certain extreme conditions,
- **Then** what values of my test statistic will lead to rejection of the null hypothesis if I want my **type I error** to be a certain value?

Nature ——— Population/ Sample ——— Observation/ Data ——— Relationships/ Modeling ——— Analysis/ Synthesis

And recall the steps.

- We agree *in advance (prior to collecting data)* that we will honor a *threshold test statistic value (this is what we mean by critical value)*, beyond which we will reject the null hypothesis ... no matter what. Thus, it is theoretically possible for the null hypothesis to be true and we reject it anyway.
- Thus, when we do have our test statistic in hand (this might be a **Z-test**):
 - IF the **Z-test_{value}** is beyond the **critical value**;
 - AND IF the null hypothesis is actually true,
 - THEN we will have made a mistake. Specifically, we will have *incorrectly rejected a true null hypothesis*. This is called a *type I error*.

Example – again –

With standard care, cancer patients are expected to survive a mean duration of time equal to 38.3 months. It is hypothesized that a new therapy will improve survival. In this study, the new therapy is administered to 100 cancer patients. Their average survival time is 46.9 months (on the face of it, this is good news!). Suppose σ^2 known = 43.3² months squared. This means that the standard deviation $\sigma = 43.3$ months. The “critical region approach” question is: Is this statistically significant evidence of improved survival *at the 0.05 level*?

Null Hypothesis Probability Model Assumptions.

X_1, X_2, \dots, X_{100} is a simple random sample from a Normal($\mu, \sigma^2 = 43.3^2$); thus, standard deviation $\sigma = 43.3$ months

Specify the null and alternative hypotheses

$H_0: \mu_{true} = \mu_0 \leq 38.3$ months

$H_A: \mu_{true} = \mu_A > 38.3$ months *one sided*.

The appropriate Test Statistic is a Z-Score

The null hypothesis gives us the following:

- X_1, X_2, \dots, X_{100} is each distributed $\text{Normal}(\mu, \sigma^2 = 43.3^2)$
- $\bar{X}_{n=100}$ is distributed $\text{Normal}(\mu = 38.3, \sigma^2 = 43.3^2/100)$
- Again, we'll use as our test statistic the z-score standardization of $\bar{X}_{n=100}$, obtained under the assumption that the null hypothesis is correct.

$$\text{Test Statistic} = \text{z-score} = \frac{\bar{X}_{n=100} - \mu_{\text{null}}}{\text{SE}(\bar{X}_{n=100})}$$

Using the direction of the alternative, obtain the 0.05 critical region

Step 1: Consider the direction of the alternative, relative to the null hypothesis

(eg – if the new treatment is better, survival will be longer):

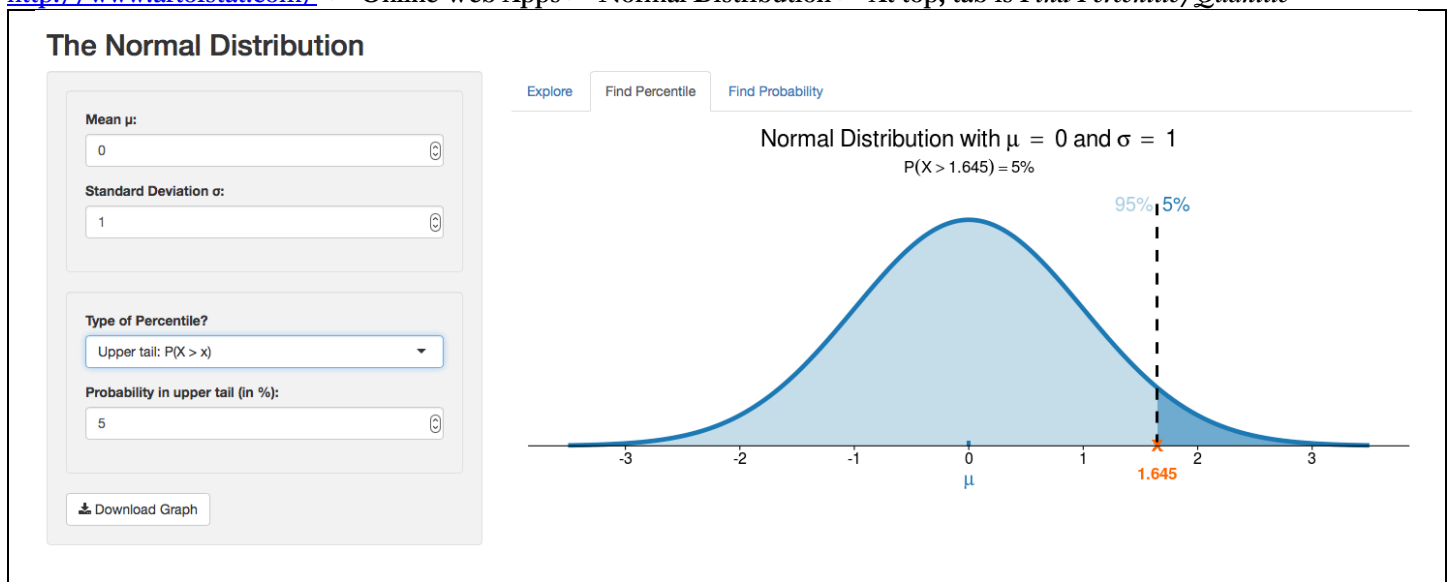
In this example, the alternative hypothesis model is **one sided**. This means that we are looking to reject the null hypothesis in favor of the alternative (improved survival) if the test statistic is extremely **large**.

Step 2: Solve for the threshold and critical region of the test statistic

Here is where our pre-specified type I error of 5% chances (see previous page) get used.

- (provisionally): Assume the null hypothesis is true
- The **Z-statistic** (see bottom of page 33) is then distributed Normal(0,1)
- Go the Normal(0,1) curve.
- Now solve for the threshold value of our Z-statistic, beyond which all values will prompt rejection of the null hypothesis, *even when the null hypothesis is true (we won't know that of course)*.
- Now you have your future rule:
When the null hypothesis is true, then our chances of making this mistake are **5% chances**; **"beyond which"** translates to **right tail area = .05**.

<http://www.artofstat.com/> > Online Web Apps > Normal Distribution > At top, tab is *Find Percentile/Quantile*



Key:

- 1) Our test statistic is the Z-statistic which is distributed Normal(0,1) when the null is true. →
At left, set **mean=0** and **standard deviation=1**
- 2) The direction of the alternative hypothesis is longer survival → **Type of Percentile = Upper Tail: P(X > x)**
- 3) Pre-set type I error that you've decided you'll live with is .05 → Set **Probability in Upper tail (in %) = 5**

The calculator then returns **critical Z-score value = 1.645**

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

Step 3: Using the critical region of the Z-statistic, solve for the critical region of \bar{X} :

How? We do this by setting the formula for the z-score equal to the value of the critical value for the z-score that was obtained in step 2, namely 1.6449.

$$\begin{aligned}
 \text{z-score} &\geq 1.6449 \rightarrow \\
 \frac{\bar{X}_{n=100} - \mu_{\text{null}}}{\text{SE}(\bar{X}_{n=100})} &\geq 1.6449 \rightarrow \\
 \bar{X}_{n=100} - \mu_{\text{null}} &\geq (1.6449) * \text{SE}(\bar{X}_{n=100}) \rightarrow \\
 \bar{X}_{n=100} &\geq [(1.6449) * \text{SE}(\bar{X}_{n=100})] + \mu_{\text{null}} \rightarrow \\
 \bar{X}_{n=100} &\geq [(1.6449) * (4.33)] + 38.3 \rightarrow \\
 \bar{X}_{n=100} &\geq 45.42
 \end{aligned}$$

We have now translated the critical Z-statistic value to its equivalent critical $\bar{X}_{n=100} \geq 45.42$

Step 4: Interpret:

- In step 2, you solved for your threshold value of Z-Score
- In step 3, you translated your threshold value of Z-score to a threshold value of $\bar{X}_{n=100}$
- So now you have another version of your future rule:
 - If your future observed $\bar{X}_{n=100}$ is at or beyond the threshold value of 45.42, then you will reject the null (and by extension infer the alternative hypothesis), even though such extreme values are theoretically possible when the null is actually true.
- That is – “this critical region one sided .05 test of the null versus alternative hypotheses has been defined to reject the null hypothesis for any future $\bar{X}_{n=100} \geq 45.42$.

Examine the observed to see if it is in the critical region

Now collect your data and compute your sample mean. In this example, the sample mean $\bar{X}_{n=100} = 46.9$. Because it exceeds the threshold value of 45.42, it falls in the critical region.

Interpret.

Because the observed $\bar{X}_{n=100} = 46.9$ *exceeds the value of the threshold 45.42* and is, therefore, *in the critical region*, in critical region parlance we say “*it is significant at the 0.05 level*”. → **reject the null hypothesis**. The conclusion is the same: these data provide statistically significant evidence that, compared to standard care, survival times on the new treatment are longer.

2d. Confidence Interval for μ (σ^2 UNknown)

Tip -

For a single sample from a normal distribution with UNknown σ^2 , a confidence interval for the mean utilizes percentiles of the **student t-distribution**. The degrees of freedom is **df = (n-1)**.

- In section 2a, we assumed that σ^2 is known and obtained a confidence interval for μ of the form

$$\text{lower limit} = \bar{X} - z_{(1-\alpha/2)100} (\sigma / \sqrt{n})$$

$$\text{upper limit} = \bar{X} + z_{(1-\alpha/2)100} (\sigma / \sqrt{n})$$

- The required confidence coefficient ($z_{1-\alpha/2}$) was obtained as a percentile from the standard normal, $N(0,1)$, distribution. (e.g. for a 95% CI, we used the 97.5th percentile)
- **More realistically, however, σ^2 will not be known. Now what?** Reasonably, we might replace σ with “s”. Recall that s is the sample standard deviation and we get it as follows:

$$s = \sqrt{s^2} \text{ where } s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

- So far so good. But there is a problem.

Whereas $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$ IS distributed Normal (0,1)

$\frac{\bar{X} - \mu}{s / \sqrt{n}}$ is **NOT** distributed Normal (0,1).

- We have to modify our “machinery” (specifically the SE piece of our machinery) to accommodate the unknown-ness of σ^2 .

Whereas we previously used when σ^2 was known	With σ^2 unknown we now use
z-score	t-score
Percentile from Normal(0,1)	Percentile from Student's t

Under simple random sampling from a normal distribution, the confidence interval for an unknown mean μ , the confidence interval will be of the following form

$$\text{lower limit} = \bar{X} - t_{\text{DF}; (1-\alpha/2)100} \left(s / \sqrt{n} \right)$$

$$\text{upper limit} = \bar{X} + t_{\text{DF}; (1-\alpha/2)100} \left(s / \sqrt{n} \right)$$

When σ^2 is not known, the computation of a confidence interval for the mean μ is not altered much.

- We simply replace the confidence coefficient from the Normal(0,1) with one from the appropriate Student's t-Distribution (the one with df = n-1)
- We replace the (now unknown) standard error with its estimate. The latter looks nearly identical except that it utilizes “s” in place of “ σ ”
- Recall

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}}$$

- Thus,

Confidence Interval for μ in two settings of a sample from a Normal Distribution	
σ^2 is KNOWN	σ^2 is NOT Known
$\bar{X} \pm [Z_{1-\alpha/2}] \cdot \left(\frac{\sigma}{\sqrt{n}} \right)$	$\bar{X} \pm [t_{n-1; 1-\alpha/2}] \cdot \left(\frac{s}{\sqrt{n}} \right)$ <p>Notice the additional subscript “(n-1)”. This is keeping track of the particular t distribution we are using, namely the one with degrees of freedom df = (n-1)</p>

Example

A random sample of size $n=20$ durations (minutes) of cardiac bypass surgeries has a mean duration of $\bar{X}_n = 267$ minutes, and variance $s^2 = 36,700$ minutes². Assuming the underlying distribution is normal with unknown variance, construct a 90% CI estimate of the unknown true mean, μ .

Answer: (192.9, 341.1) minutes

Solution:

Step 1 - Point Estimate of μ is the Sample Mean \bar{X}_n

$$\bar{X}_{n=20} = \frac{\sum_{i=1}^n X_i}{n=20} = 267 \text{ minutes.}$$

Step 2 – The Estimated Standard Error of \bar{X}_n is s/\sqrt{n}

$$\widehat{SE}(\bar{X}_{n=20}) = \sqrt{\widehat{\text{Variance}}(\bar{X}_{n=20})} = \sqrt{\frac{S^2}{n}} = \frac{\sqrt{36,700}}{\sqrt{20}} = 42.8369$$

minutes

Step 3 - The Confidence Coefficient

See again page 6 for how to think about this. In a nutshell:

For a 90% confidence interval, think coverage = 90%, then think NON-coverage=10%, then think splitting non-coverage equally in the two tails, 5% on each side... and you arrive at wanting the 95th percentile of the Student's t-Distribution that has degrees of freedom = $df = (n-1) = 19$. This value is 1.729133.

Putting this all together –

$$\begin{aligned} \text{Lower limit} &= (\text{point estimate}) - (\text{confidence coefficient.}) (\text{SE of point estimate}) \\ &= 267 - (1.729133) (42.8369) \\ &= 192.9293 \end{aligned}$$

$$\begin{aligned} \text{Upper limit} &= (\text{point estimate}) + (\text{confidence coefficient}) (\text{SE of point estimate}) \\ &= 267 + (1.729133) (42.8369) \\ &= 341.0707 \end{aligned}$$

Thus, a 90% confidence interval for the true mean duration of surgery is **(192.9, 341.1) minutes**.

Art of Stat Illustration

Solution:

Solve for sample standard deviation = $S = \sqrt{S^2} = \sqrt{36,700} = 191.57244$

In Art of Stat: Online Web Apps > Inference for a Population Mean

At left, make the following selections:

Enter Data: **Summary Statistics**

Name of Variable: **Duration Cardiac Bypass Surgery (minutes)** ... or whatever name you like ...

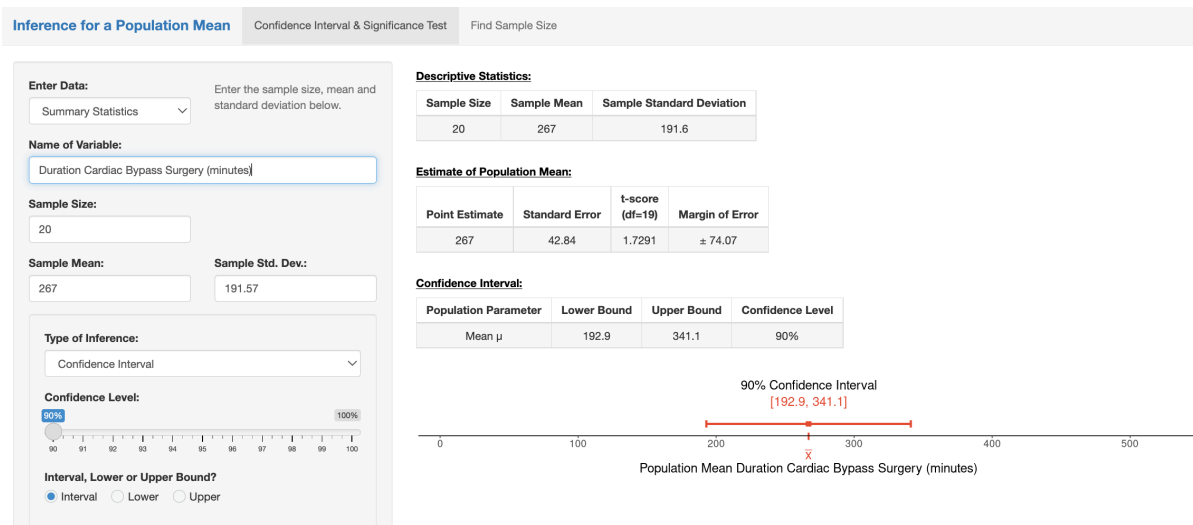
Sample Size: **20**

Sample Mean: **267**

Sample Std. Dev.: **191.57**

Type of Inference: **Confidence Interval**

Confidence Level: **90%**



R Illustration

```
my_n <- 20 # User supplies sample size, n
my_mean <- 267 # User supplies sample mean
my_var <- 36700 # User supplies sample variance
my_conf <- .90 # User supplies desired confidence level
```

```
my_sd <- sqrt(my_var) # standard deviation
sem <- my_sd / sqrt(my_n) # standard error of the mean
crit_percentile <- 1 - (1 - my_conf)/2 # percentile for desired confidence
t_crit <- qt(p=crit_percentile, df=my_n-1) # confidence coefficient multiplier
lowerci <- my_mean - t_crit*sem # lower confidence interval limit
upperci <- my_mean + t_crit*sem # upper confidence interval limit
```

```
table1 <- cbind(my_mean, my_n, my_sd, lowerci, upperci)
colnames(table1) <- c("mean", "n", "sd", "lower 95% CI", "upper 95% CI")
table1
```

```
mean n sd lower 95% CI upper 95% CI
[1,] 267 20 191.5724 192.9293 341.0707
```

2e. Hypothesis Test for μ (σ^2 UNknown)

Tip -

For a single sample from a normal distribution with UNknown σ^2 , a hypothesis test about the mean utilizes a p-value calculation from the **student t distribution**. The degrees of freedom is **df = (n-1)**.

Hypothesis testing in the setting of a sample from a single normal distribution with σ^2 not known is quite similar to that when the data are from a distribution with σ^2 known. The test statistic is a **t-score** instead of a z-score.

Same example –

With standard care, cancer patients are expected to survive a mean duration of time equal to 38.3 months. Hypothesized is that a new therapy will improve survival. In this study, the new therapy is administered to 100 cancer patients. Their average survival time is 46.9 months. Suppose σ^2 is not known. Suppose instead that what is available is the sample variance of survival times $S^2 = 43.3^2$ months squared. This means that the sample standard deviation is $S = 43.3$ months. Do these data provide statistically significant evidence of improved survival?

Null Hypothesis Probability Model.

X_1, X_2, \dots, X_{100} is a simple random sample from a $\text{Normal}(\mu, \sigma^2)$
 σ^2 is **NOT** known.

Null and alternative hypotheses

$H_0: \mu_{true} = \mu_0 \leq 38.3$ months
 $H_A: \mu_{true} = \mu_A > 38.3$ months one sided.

Reason “proof by contradiction” and use this to define the p-value calculation.

Statistically, the null hypothesis, when examined in light of the value of the test statistic, will have led to an unlikely outcome (an inconsistency or “contradiction”) if the null hypothesis model probability of that test statistic value (a sample mean being as extreme as 46.9 months or larger) is small. The chances of this (the p-value) is the following probability calculation:

$$\Pr[\bar{X}_{n=100} \geq 46.9 \mid \mu_{true} = \mu_0 = 38.3]$$

Nature ——— Population/ Sample ——— Observation/ Data ——— Relationships/ Modeling ——— Analysis/ Synthesis

The appropriate Test Statistic is now a **Student T-Test or T-Score**

Under the assumption of the null hypothesis:

- X_1, X_2, \dots, X_{100} is a simple random sample from a Normal($\mu = 38.3, \sigma^2$).
- $\bar{X}_{n=100}$ is distributed Normal ($\mu = 38.3, \sigma_{\bar{X}}^2 = \sigma^2/100$)
- Here, we'll use as our test statistic the **t-score standardization** of $\bar{X}_{n=100}$, obtained under the assumption that the null hypothesis is correct.

$$\text{Test Statistic} = \text{t-score} = \frac{\bar{X}_{n=100} - \mu_{\text{null}}}{\hat{SE}(\bar{X}_{n=100})} = \frac{\bar{X}_{n=100} - \mu_{\text{null}}}{\sqrt{S^2 / n}} = \frac{46.9 - 38.3}{\sqrt{43.3^2 / 100}} = 1.99$$

- Notice that the standard error in the denominator, previously a known standard error when computing a Z-score is now an estimate of the unknown SE because we do not know the value of the population σ .

P-value calculation

$$\text{p-value} = \Pr[\bar{X}_{n=100} \geq 46.9 \mid \mu_{\text{true}} = \mu_{\text{null}} = 38.3]$$

$$= \Pr[\text{t-score}_{\text{degrees of freedom}=99} \geq 1.99] = .02467 \quad \text{how about that, close to .02 obtained on p 11.}$$

“Evaluate”.

IF we assume that the null hypothesis is true, meaning that the new therapy elicits **no improvement** in survival so that the survival experience under the new therapy is **the same as** that experienced with receipt of standard care,

THEN its application to the data yielded an **estimated 2.4%** (p-value = .024) chance that we would have obtained our observed average survival time of 46.9 months **or greater**.

Interpret.

The assumption that the null hypothesis is true, *when examined in light of the observed statistic's value*, has led to an unlikely conclusion. Abandon the null hypothesis. → **Reject the null hypothesis**. The conclusion is again the same: these data provide statistically significant evidence that, compared to standard care, survival times on the new treatment are longer.

Art of Stat Illustration

Solution:

Solve for sample standard deviation = $S = \sqrt{S^2} = \sqrt{36,700} = 191.57244$

In Art of Stat: Online Web Apps > scroll down quite a ways to get to: **Inference for a Population Mean**

At left, make the following selections:

Enter Data: **Summary Statistics**

Name of Variable: **Survival, months ... or whatever name you like ...**

Sample Size: **100**

Sample Mean: **46.9**

Sample Std. Dev.: **43.3**

Type of Inference: **Significance Test**

Null value: **38.3**

Alternative: **Greater**

Inference for a Population Mean

Confidence Interval & Significance Test

Find Sample Size

Enter Data:

Summary Statistics

Enter the sample size, mean and standard deviation below.

Name of Variable:

Survival, months

Sample Size:

100

Sample Mean:

46.9

Sample Std. Dev.:

43.3

Type of Inference:

Significance Test

Null value μ_0 :

38.3

Alternative:

Greater

Options:

☐ Show Degrees of Freedom

☐ Decision Reject/Not Reject

Descriptive Statistics:

Sample Size	Sample Mean	Sample Standard Deviation
100	46.9	43.3

Test Statistic:

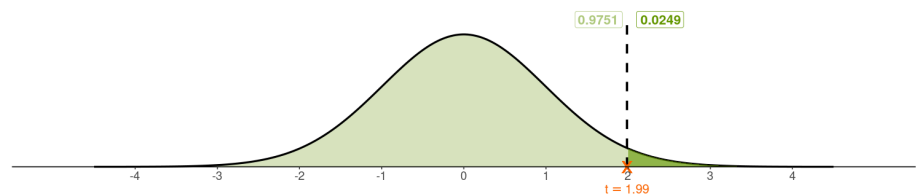
Null Value	Standard Error	Test Statistic t
38.3	4.33	1.986

Hypothesis Test:

Population Parameter	Null Hypothesis	Alternative Hypothesis	Test Statistic t	P-value
Mean μ	$\mu = 38.3$	$\mu > 38.3$	1.9861	0.0249

t Distribution with df = 99

$H_0: \mu = 38.3$, $H_a: \mu > 38.3$. Test Statistic: $t = 1.986$, P-value = 0.0249



R Illustration

```

yn <- 100 # User supplies sample size
ymean <- 46.9 # User supplies sample mean
ysd <- 43.3 # User supplies sample standard deviation
my_null <- 38.3 # User supplies null mean
my_alt <- "greater than" # Choose alternative: "greater than", "less than", "two sided"

ysem <- ysd/sqrt(yn) # Sample SEM = standard error of the mean
my_df <- (yn - 1) # degrees of freedom

if (my_alt == "two sided") {
  my_tscore <- abs((ymean - my_null)/ysem) # For two-sided alternative, use absolute value of t-score
} else {
  my_tscore <- (ymean - my_null)/ysem # Otherwise, use t-score as calculated
}

if (my_alt == "two sided") {
  my_pvalue <- 2*pt(q=my_tscore, df=my_df,lower.tail=FALSE) # 2 sided p-value
} else if (my_alt == "greater than") {
  my_pvalue <- pt(q=my_tscore, df=my_df,lower.tail=FALSE) # 1 sided p-value (greater)
} else{
  my_pvalue <- pt(q=my_tscore, df=my_df,lower.tail=TRUE) # 1 sided p-value (less than)
}

my_tscore <- round(my_tscore, digits=4)
my_pvalue <- round(my_pvalue, digits=4)

table1 <- cbind(my_null, my_alt, yn, ymean, ysd,my_tscore, my_pvalue)
colnames(table1) <- c("Null", "Alternative", "n", "Mean", "Std Dev.", "t-statistic", "p-value")
table1

```

	Null	Alternative	n	Mean	Std Dev.	t-statistic	p-value
[1,]	"38.3"	"greater than"	"100"	"46.9"	"43.3"	"1.9861"	"0.0249"

3. One Sample – Normal Distribution: σ^2

3a. Confidence Interval for σ^2

Tip -

For a single sample from a normal distribution, a confidence interval for the variance σ^2 utilizes percentiles of the **chi square distribution**. The degrees of freedom are **df = (n-1)**.

- The following are some settings where our interest lies in estimation of the variance, σ^2
 - Standardization of equipment – repeated measurement of a standard should have small variability
 - Evaluation of technicians – are the results from a particular technician “too variable”
 - Comparison of measurement techniques – are the measurements obtained using a new technique too variable compared to the precision of the old technique?
- We have a point estimator the population variance σ^2 . It is the sample variance S^2 .
- How do we get a confidence interval? The answer will utilize a new standardized variable, based on the way in which S^2 is computed. It is a **chi square** random variable.

The definition of the chi square distribution gives us what we need to construct a confidence interval estimate of σ^2 when data are a simple random sample from a normal probability distribution. The approach here is similar to that for estimating the mean μ .

- The table below shows how to construct a confidence interval.
For the interested reader, Appendix ii is the **derivation** behind the calculation.

(1- α)100% Confidence Interval for σ^2 Setting – Normal Distribution	
Lower limit =	$\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}$
Upper limit =	$\frac{(n-1)S^2}{\chi^2_{\alpha/2}}$

Nature _____ Population/ Sample _____ Observation/ Data _____ Relationships/ Modeling _____ Analysis/ Synthesis

Example

A precision instrument is guaranteed to read accurately to within ± 2 units. A sample of 4 readings on the same object yield 353, 351, 351, and 355. Find a 95% confidence interval estimate of the population variance σ^2 and also for the population standard deviation σ .

Answer: 95% confidence interval estimate of $\sigma^2 = (1.18, 51.0)$ units squared

Solution:

1. Obtain the point estimate of σ^2 . It is the sample variance S^2

To get the sample variance S^2 , we will need to compute the sample mean first.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = 352.5 \quad \text{and} \quad S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = 3.67$$

2. Determine the correct chi square distribution to use.

It has degrees of freedom, $df = (n-1) = (4-1) = 3$.

3. Obtain the correct multipliers.

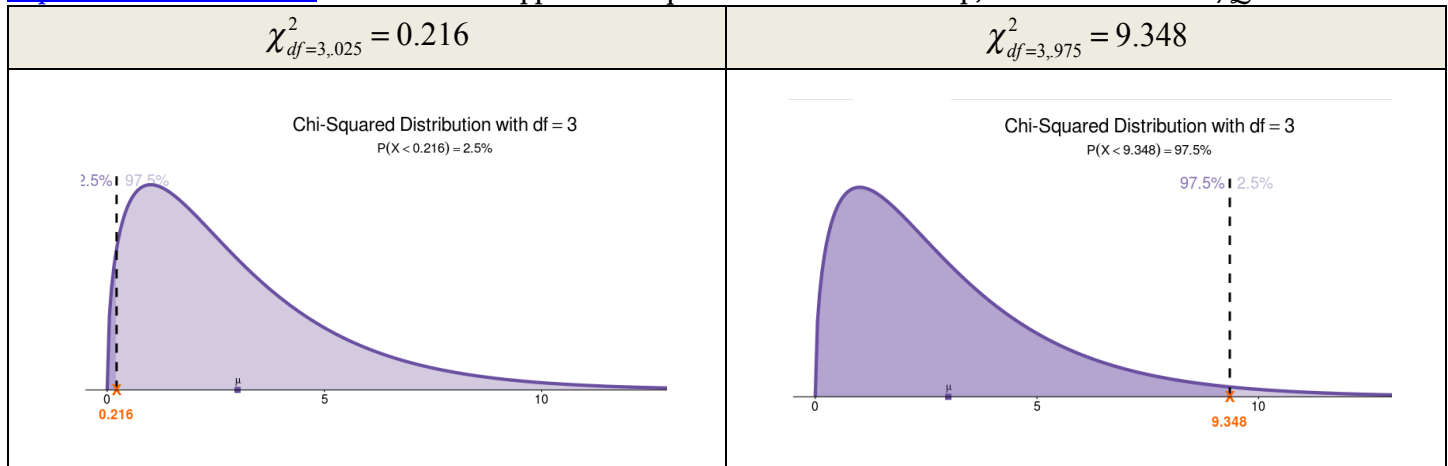
Because the desired confidence level is 0.95, we set $0.95 = (1-\alpha)$. Thus $\alpha = .05$

For a 95% confidence level, the percentiles we want are

- (i) $(\alpha/2)100^{\text{th}} = 2.5^{\text{th}}$ percentile
- (ii) $(1 - \alpha/2)100^{\text{th}} = 97.5^{\text{th}}$ percentile

Obtain percentiles for chi square distribution with degrees of freedom = 3

<http://www.artofstat.com/> > Online Web Apps > Chi Square Distribution > At top, tab is *Find Percentile/Quantile*



ArtofStat link, here: <https://stats.shinyapps.io/ChisqDist/>

4. Thus,

$$(i) \text{ Lower limit} = \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}} = \frac{(3)(3.67)}{9.348} = 1.178$$

$$(ii) \text{ Upper limit} = \frac{(n-1)S^2}{\chi^2_{\alpha/2}} = \frac{(3)(3.67)}{0.216} = 50.97$$

Illustration Using GraphPad (Sorry, Art of Stat does not have a utility for this)
 (Search on “GraphPad Online confidence interval of a SD)

<https://www.graphpad.com/quickcalcs/CIstd1/?Format=50>

graphpad.com/quickcalcs/CIstd1/?Format=50

1. Select category 2. Choose calculator 3. Enter data 4. View results

Confidence interval of a SD

It is straightforward to calculate the standard deviation from a bunch of values. But how accurate is the standard deviation? Just by chance you may have happened to obtain data that are closely bunched together, making the SD low. Or you may have happened to obtain data that are far more scattered than the overall population, making the SD high. If you assume that your data are randomly sampled from a population that follows a Gaussian distribution, This calculator can compute a 95% confidence interval for the standard deviation.

1. Choose data entry format

Caution: Changing format will erase your data.

☒ Enter up to 50 rows.

☐ Enter or paste up to 10000 rows.

☐ Enter SEM and N.

☐ Enter SD and N.

2. Enter data

1: 353

2: 351

3: 351

4: 355

5:

Scroll down to click on: **CALCULATE NOW**

GraphPad by Dotmatics

Prism Customers Resources Support Pricing Cart Sign In

1. Select category 2. Choose calculator 3. Enter data 4. View results

Confidence interval of a SD

Parameter	Value
SD	1.91
SEM	0.96
N	4
90% CI of the SD	1.19 to 5.59
95% CI of the SD	1.08 to 7.14
99% CI of the SD	0.93 to 12.38

These results assume that you have randomly sampled data from a population that is distributed according to a Gaussian distribution. Note that the confidence intervals are not symmetric around the SD. This makes sense. SD values must be greater than zero, so the uncertainty for the upper confidence limit extends further than the lower limit.

GraphPad returns 95% CI for SD $\sigma = 1.08 - 7.14$
 Squaring these values yields 95% CI for Variance $\sigma^2 = 1.18 - 50.97$

Nature ——— Population/ Sample ——— Observation/ Data ——— Relationships/ Modeling ——— Analysis/ Synthesis

R Illustration

```
# EXAMPLE
my_data <- c(353,351,351,355) # User supplies data

my_var <- var(my_data, na.rm=TRUE) # sample variance
my_sd <- sqrt(my_var) # sample standard deviation
my_df <- length(my_data) - 1 # degrees of freedom = (n-1)
chisq_upper <- qchisq(.975,df=my_df) # 97.5th percentile of Chi Square with df = (n-1)
chisq_lower <- qchisq(.025,df=my_df) # 2.5th percentile of Chi Square with df = (n-1)
lowerci_var <- my_df*my_var/chisq_upper # lower limit of 95% CI for variance
upperci_var <- my_df*my_var/chisq_lower # upper limit of 95% CI for variance
lowerci_sd <- sqrt(lowerci_var) # lower limit of 95% CI for SD
upperci_sd <- sqrt(upperci_var) # upper limit of 95% CI for SD

table2 <- cbind(my_var,lowerci_var,upperci_var)
colnames(table2) <- c("Sample variance", "lower 95% CI", "upper 95% CI")
table2

table3 <- cbind(my_sd,lowerci_sd,upperci_sd)
colnames(table3) <- c("Sample Standard Deviation", "lower 95% CI", "upper 95% CI")
table3
```

```
Sample variance lower 95% CI upper 95% CI
[1,] 3.666667 1.176671 50.97424

Sample Standard Deviation lower 95% CI upper 95% CI
[1,] 1.914854 1.084745 7.139624
```

Obtain a Confidence Interval for the Population Standard Deviation σ

Answer: 95% confidence interval estimate of $\sigma = (1.09, 7.14)$ units

Solution:

Step 1 – Obtain a confidence interval for σ^2

(1.178, 51.02)

Step 2 – The associated confidence interval for σ is obtained by taking the square root of each of the lower and upper limits

$$\text{- 95\% Confidence Interval} = (\sqrt{1.178}, \sqrt{50.97}) = (1.09, 7.14)$$

$$\text{- Point estimate} = \sqrt{3.67} = 1.92$$

Remarks on the Confidence Interval for σ^2

- It is **NOT** symmetric about the point estimate; the “safety net” on each side of the point estimate is of different lengths.
- These intervals tend to be wide. Thus, large sample sizes are required to obtain reasonably narrow confidence interval estimates for the variance and standard deviation parameters.

Nature _____ Population/
Sample _____ Observation/
Data _____ Relationships/
Modeling _____ Analysis/
Synthesis

HOMEWORK DUE Monday November 21, 2022

Question #3 of 8

The objectives of a study by Kennedy and Bhambhani (1991) were to use physiological measurements to determine the test-retest reliability of the Baltimore Therapeutic Equipment Work Simulator during three simulated tasks performed at light, medium, and heavy work intensities, and to examine the criterion validity of these tasks by comparing them to real tasks performed in a controlled laboratory setting. Subjects were 30 healthy men between the ages of 18 and 35. The investigators reported a standard deviation of $s=0.57$ for the variable peak oxygen consumption (l/min) during one of the procedures.

Assuming normality, compute a 95% confidence interval for the population variance for the oxygen consumption variable.

3b. Hypothesis Test of σ^2

Tip -

For a single sample from a normal distribution, a hypothesis test for the variance σ^2 utilizes a p-value calculation from the **chi square distribution**.

Example -

Recall - Studies of laboratory performance (reproducibility, precision) are studies of variance.

In drug manufacturing it is important, not only that the amount of drug in the capsules be a particular value on the average, but also that the variation around that value be very small. The drug company will consider its machine accurate enough if the capsules are filled within 1 SD = .5 mg of the desired amount of the drug (2.5 mg). Data is collected for n=20 capsules. The observed sample standard deviation is S= 0.787. Is this variability statistically significantly greater than what the company will tolerate? Test whether the drug company should adjust its machines again. The company will only adjust the machine if the variance is too large.

Research Question:

Is the variance of drug in the capsules greater than $(.5)^2 = 0.25 \text{ mg}^2$?

Null Hypothesis Assumptions:

The data are a simple random sample from a normal distribution.

Specify Hypotheses:

$$H_0: \sigma^2 \leq 0.25$$

$$H_a: \sigma^2 > 0.25 \text{ one-sided}$$

Reason “proof by contradiction” and use it to define the p-value calculation.

Statistically, the null hypothesis, when examined *in light of the observed data*, lead to an inconsistency or “contradiction” if the null hypothesis probability is small that the observed sample SD among n=20 capsules is 0.787 or larger. Thus, the required p-value calculation is:

$$\text{P-value} = \Pr[S \geq 0.787 | \sigma_{\text{true}} = \sigma_o = 0.5]$$

The Test Statistic is a Chi Square:

S, as is, cannot be our test statistic. Its “behavior” isn’t convenient. Thus, we do not calculate $\Pr [S > 0.787]$ directly.

Instead we work with a related random variable that does have convenient properties. It is called Y and is obtained under the assumption that the null hypothesis is true. Note – This is analogous working with a Z-score standardization of the sample mean.

In particular, under the assumption that the null hypothesis is true

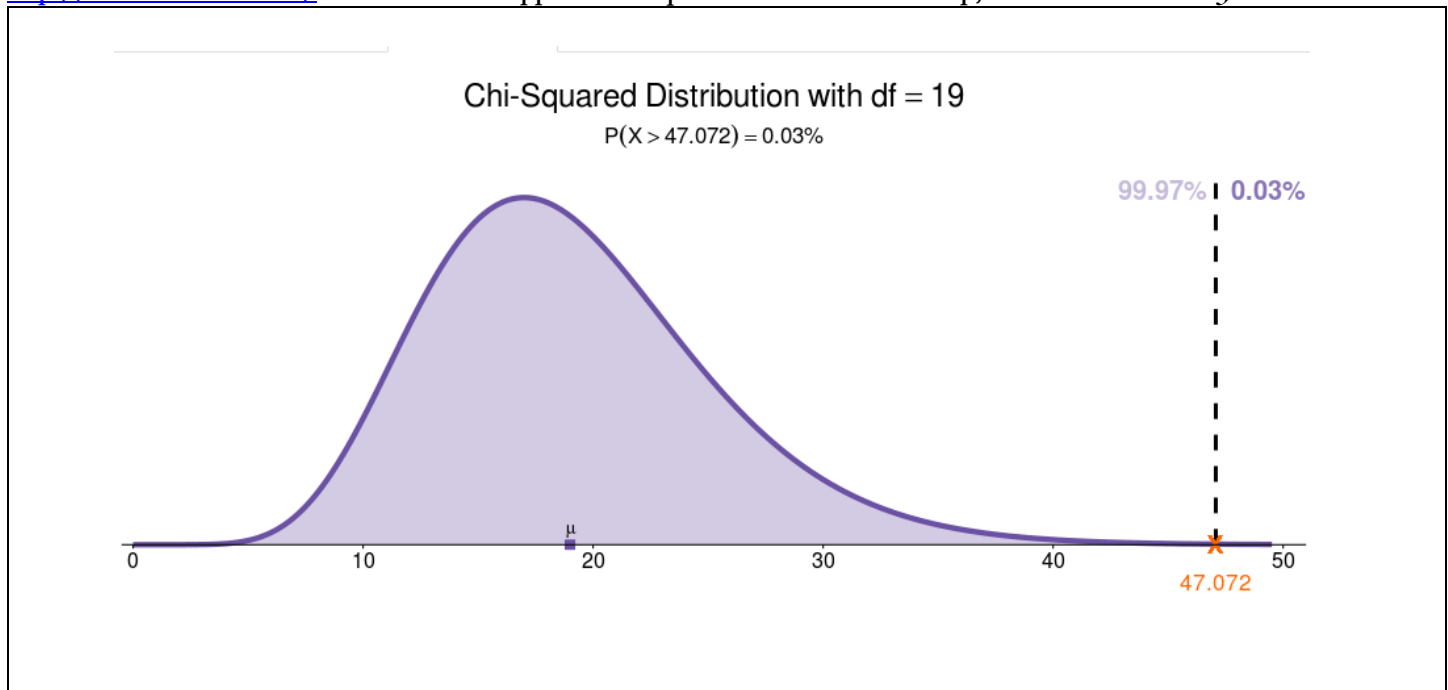
$$Y = \frac{(n-1)S^2}{\sigma_{\text{NULL}}^2} \text{ is distributed chi square with degrees of freedom } = (n-1)$$

P-value = the probability of that a chi square random variable is as extreme or more extreme (translation: larger) than the value that we obtained from our data. As before, it is calculated under the assumption that the null hypothesis is true.

$$Y = \frac{(n-1)S^2}{\sigma_{\text{NULL}}^2} = \frac{(19)(0.787)^2}{0.25} = 47.072$$

$$\text{p-value} = \Pr [\text{Chi Square}_{\text{DF}=19} \geq 47.072] = 0.0003 \quad (\text{the area under the curve is tiny!})$$

<http://www.artofstat.com/> > Online Web Apps > Chi Square Distribution > At top, tab is Find Probability



ArtofStat link, here: <https://istats.shinyapps.io/ChisqDist/>

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

“Evaluate”.

Assumption of the null hypothesis as being true and its application to the observed data (evaluating it in light of the data), has led to an inconsistency/contradiction. Under the null hypothesis model, the probability was an **estimated 0.03%** chance that our sample variance would be 0.787^2 or larger.

Interpret.

The null hypothesis is abandoned. → Reject the null hypothesis. Conclude that these data provide statistically significant evidence that the variability in capsule drug content is greater than $.25 \text{ mg}^2$.

Illustration Using Statistics Kingdom (Here, too. Sorry. Art of Stat does not have a utility for this)

(Search on “Staistics Kingdom One Sample Variance Test Calculator)

<https://www.statskingdom.com/210VarCHI1.html>

Home > Variance > χ^2 test

One Sample Variance Test Calculator

Chi-square test of a single variance, checks if the population variance (σ^2) is different from the expected value (σ_0^2).

Tails: Right ($H_1: \sigma > \sigma_0$)

Significance level (α): 0.05

Effect: Medium

Effect size: σ_1^2/σ_0^2 : 1.3

Expected standard deviation (σ_0): 0.5

Outliers: Included

☒ Enter summarized data (n, S)
☐ Enter raw data directly
☐ Enter raw data from excel

Group name: Group-1

Sample size (n): 20

Sample Standard deviation (S): 0.787

When entering raw data, the tool will run the Shapiro-Wilk normality test, calculate outliers and generate R code, as part of the test calculation.

Alternative is RIGHT sided

Null population SD = 0.5

Sample size is n = 20

Sample SD = 0.787

Scroll down to click on: CALCULATE TEST

Chi-square test for variance, using χ^2 distribution (df=19) (right-tailed) (validation)

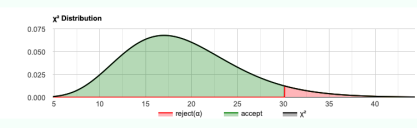
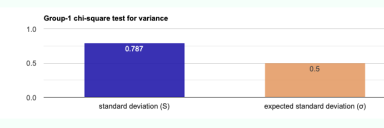
1. H_0 hypothesis
 Since p-value < α , H_0 is rejected.
 The standard deviation (S) of Group-1's population is considered to be **greater than** the expected standard deviation (σ).

2. P-value
 The p-value equals **0.0003488**, ($p(\chi^2) = 0.9997$). It means that the chance of type I error (rejecting a correct H_0) is small: 0.0003488 (0.035%). The smaller the p-value the more it supports H_1 .

3. The statistics
 The test statistic χ^2 equals **47.072**, which is not in the 95% region of acceptance: $[-\infty : 30.1435]$. $S=0.79$, is not in the 95% region of acceptance: $[-\infty : 0.6298]$.

4. Effect size
 The observed effect size phi is **large, 1.53**. This indicates that the magnitude of the difference between the standard deviation (S) and expected standard deviation (σ) is large.

If you like the page, please share or like. Questions, comments and suggestions are appreciated. (statskingdom@gmail.com)

p-value < .003

Chi Square Test Statistic = 47.072

R Illustration

```
my_n <- 20 # User supplies sample size
my_sd <- 0.787 # User supplies sample standard deviation
my_nullSD <- 0.5 # User supplies null mean
my_alt <- "greater than" # Choose: "greater than", "less than", "two sided"

my_var <- my_sd^2 # Sample variance
my_df <- my_n - 1 # Degrees of freedom
my_nullVAR <- my_nullSD^2 # Null hypothesis variance

my_chisquare <- (my_df * my_var) / my_nullVAR # Chi Square Statistic

if (my_alt == "two sided") {
  my_pvalue <- 2*pchisq(q=my_chisquare, df=my_df,lower.tail=FALSE) # 2 sided p-value
} else if (my_alt == "greater than") {
  my_pvalue <- pchisq(q=my_chisquare, df=my_df,lower.tail=FALSE) # 1 sided p-value (greater)
} else{
  my_pvalue <- pchisq(q=my_chisquare, df=my_df,lower.tail=TRUE) # 1 sided p-value (less than)
}

my_chisquare <- round(my_chisquare, digits=4)
my_pvalue <- round(my_pvalue, digits=4)

table1 <- cbind(my_nullSD, my_alt, my_n, my_sd, my_chisquare, my_df, my_pvalue)
colnames(table1) <- c("Null SD", "Alternative", "n", "Sample SD", "Chi Square Statistic", "DF", "p-value")
table1
```

	Null SD	Alternative	n	Sample SD	Chi Square Statistic	DF	p-value
[1,]	"0.5"	"greater than"	"20"	"0.787"	"47.072"	"19"	"0.0003"

4. One Sample *Paired Data* – Normal Distribution

4a. Introduction to Paired Data

Bottom line-

IF $\sigma^2_{\text{DIFFERENCE}}$ is KNOWN,

THEN a confidence interval for the mean difference $\mu_{\text{DIFFERENCE}}$ utilizes percentiles of the **standard normal distribution**.

BUT IF $\sigma^2_{\text{DIFFERENCE}}$ is UNknown,

THEN a confidence interval for the mean difference $\mu_{\text{DIFFERENCE}}$ utilizes percentiles of the **student t distribution**.

- Paired data occur when each individual (more specifically, each unit of measurement) in a sample is measured twice.
- Paired data are familiar: "pre/post", "before/after", "right/left", "parent/child", etc.
- **Examples** -
 - 1) Blood pressure prior to and following treatment,
 - 2) Number of cigarettes smoked per week measured prior to and following participation in a smoking cessation program,
 - 3) Number of sex partners in the month prior to and in the month following an HIV education campaign.
- In each of these examples, the two occasions of measurement are linked by the fact that the two measurements being made on the same individual (*they are paired*).

We are interested in comparing the two paired outcomes.

When the paired data are **continuous**, the comparison focuses on the **difference** between the two paired measurements. *Looking ahead – We'll see later that when the data are discrete, an analysis of paired data might focus on the ratio (eg. relative risk) of the two measurements rather than on the difference.*

Examples -

- 1) Blood pressure prior to and following treatment. Interest is $d = \text{pre-post}$.
Large (pre-post) differences are evidence of blood pressure lowering associated w treatment.
- 2) Number of cigarettes smoked per week measured prior to and following participation in a smoking cessation program. Interest is $d = \text{pre-post}$.
Large (pre-post) differences “d” are evidence of smoking reduction.
- 3) Number of sex partners in the month prior to and in the month following an HIV education campaign. Interest is $d = \text{pre} - \text{post}$. Large (pre-post) differences are evidence of safer sex behaviors.

Confidence Interval for $\mu_{\text{DIFFERENCE}}$

- Suppose paired measurements are made of the same phenomenon (eg. blood pressure, # cigarettes/week, etc) on each individual in a sample. Call them X and Y. If X and Y are each normally distributed, then their difference is also distributed normal. (see again, section 4d)
- Thus, the setting is our focus on the difference D and the following assumptions
 - (1) $D = (X - Y)$ is distributed Normal with
 - (2) Mean of $D = \mu_{\text{difference}}$. Let's write this as μ_d
 - (3) Variance of $D = \sigma_{\text{DIFFERENCE}}^2$ Let's write this as σ_d^2
- In this setting, estimation for paired data is a special case of selected methods already presented. Attention is restricted to the single random variable defined as the difference between the two measurements. The methods already presented that we can use here are

- (1) Confidence Interval for μ_d - Normal Distribution σ_d^2 unknown
- (2) Confidence Interval for σ_d^2 - Normal Distribution

4b. Confidence Interval for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ is KNOWN

When the $\sigma^2_{\text{DIFFERENCE}}$ is **known**, a confidence interval for the mean difference $\mu_{\text{DIFFERENCE}}$ utilizes percentiles of the **standard normal distribution**.

Example

source: Anderson TW and Sclove SL. *Introductory Statistical Analysis*. Boston: Houghton Mifflin, 1974. page 339

A researcher is interested assessing the improvement in reading skills upon completion of the second grade (Y) in comparison to those prior to the second grade (X). For each child, his or her improvement is measured using the difference “d” which is defined $d = Y - X$. A sample of $n=30$ children are studied

ID	PRE(X)	POST(Y)	d=(Y-X)
1	1.1	1.7	0.6
2	1.5	1.7	0.2
3	1.5	1.9	0.4
4	2.0	2.0	0.0
5	1.9	3.5	1.6
6	1.4	2.4	1.0
7	1.5	1.8	0.3
8	1.4	2.0	0.6
9	1.8	2.3	0.5
10	1.7	1.7	0.0
11	1.2	1.2	0.0
12	1.5	1.7	0.2
13	1.6	1.7	0.1
14	1.7	3.1	1.4
15	1.2	1.8	0.6
16	1.5	1.7	0.2
17	1.0	1.7	0.7
18	2.3	2.9	0.6
19	1.3	1.6	0.3
20	1.5	1.6	0.1
21	1.8	2.5	0.7
22	1.4	3.0	1.6
23	1.6	1.8	0.2
24	1.6	2.6	1.0
25	1.1	1.4	0.3
26	1.4	1.4	0.0
27	1.4	2.0	0.6
28	1.5	1.3	-0.2
29	1.7	3.1	1.4
30	1.6	1.9	0.3

Suppose it is known that $\sigma^2_d = 0.20$

Note – Did you notice that, here, we are considering differences defined as (post-pre) instead of as (pre-post)? This is fine! You just need to keep track of which approach you are using so that you then interpret your results accordingly.

Nature _____ Population/ Sample _____ Observation/ Data _____ Relationships/ Modeling _____ Analysis/ Synthesis

Solution for a 99% Confidence Interval for μ_d when σ_d^2 is known

Step 1 – Point Estimate of μ_d is the Sample Mean $\bar{d}_{n=30}$

$$\bar{d}_{n=30} = \frac{\sum_{i=1}^n d_i}{n=30} = 0.51$$

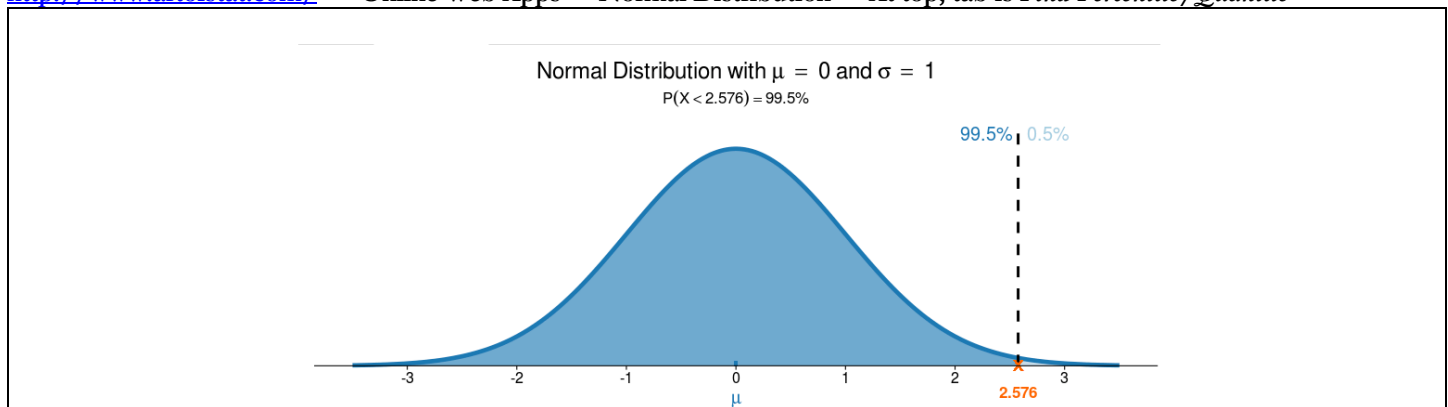
Step 2 – The Standard Error of \bar{d}_n is σ_d / \sqrt{n}

$$SE(\bar{d}_{n=30}) = \frac{\sigma_d}{\sqrt{n}} = \frac{\sqrt{0.20}}{\sqrt{30}} = 0.0816$$

Step 3 – The Confidence Coefficient

For a 99% confidence interval, this number will be the 99.5th percentile of the Normal(0,1) distribution. This value is **2.576**.

<http://www.artofstat.com/> > Online Web Apps > Normal Distribution > At top, tab is *Find Percentile/Quantile*



ArtofStat link, here: <https://istats.shinyapps.io/NormalDist/>

Step 4 – Substitute into the formula for a confidence interval

$$\begin{aligned} \text{Lower limit} &= (\text{point estimate}) - (\text{confidence coefficient}) (\text{SE of point estimate}) \\ &= 0.51 - (2.576) (0.0816) \\ &= 0.2998 \end{aligned}$$

$$\begin{aligned} \text{Upper limit} &= (\text{point estimate}) + (\text{confidence coefficient.}) (\text{SE of point estimate}) \\ &= 0.51 + (2.576) (0.0816) \\ &= 0.7202 \end{aligned}$$

Thus, a 99% confidence interval for the mean change in reading skills (Post – Pre) = **(0.2998, 0.7202)**

HOMEWORK DUE Monday November 21, 2022

Question #4 of 8

An investigation by Alahuhta et al (1991) evaluated the effects of extradural block for elective caesarian section simultaneously on several maternal and fetal hemodynamic variables and to determine if the block modified fetal myocardial function. The study subjects were n=8 healthy parturient in gestational weeks 38-42 with uncomplicated singleton pregnancies undergoing elective caesarian section under extradural anesthesia. Among the measurements taken were maternal diastolic arterial pressure during two stages of the study. The following are the lowest values of this variable at the two stages.

Patient ID	1	2	3	4	5	6	7	8
Stage 1	70	87	72	70	73	66	63	57
Stage 2	79	87	73	77	80	64	64	60

Source: Alahuhta S., Rasanen J, Jouppila R, Jouppila P., and Kangas-Saarela T., and Hoomen AI (1991) "Uteroplacental and fetal hemodynamics during extradural anesthesia for caesarian section", British Journal of Anesthesia, 66: 319-323. Cited in Daniel (p 248, 7.4.3) Copyright 1999 by John Wiley & Sons, Inc. By permission of John Wiley.

Compute a 95% confidence interval for the difference in diastolic blood pressure between the two stages.

4c. Hypothesis Test for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ is KNOWN

When the $\sigma^2_{\text{DIFFERENCE}}$ is known, a hypothesis test for the mean difference $\mu_{\text{DIFFERENCE}}$ utilizes a p-value calculation from the standard normal distribution.

Example –

(Note: These data are hypothetical.)

Twelve patients (n=12) in a needle exchange trial who were randomized to the control condition defined as pharmacy sales alone provided hair samples that were positive for cocaine at the baseline interview. Follow-up hair samples were obtained from these same n=12 at the 6 month visit, yielding paired data.

So what? Well, there's good reason to be interested in these paired data. In a study of risky behavior, quite plausibly, participation *itself* might be associated with reduced hair content of cocaine? **If it occurs, then at least two issues arise: (1) there might be selection bias which makes generalization to a larger population difficult; and/or (2) it is difficult to attribute reduced cocaine use to the intervention and/or (3) ???**

Research Question.

In the absence of an effect of study participation, it is expected that cocaine use would be stable over time. Accordingly, the hair content of cocaine would be expected to be the same at the baseline and follow-up visits. Does participation alone in an intervention study reduce cocaine use?

- * Let the 12 pairs of cocaine measurements be denoted $(X_1, Y_1) \cdots (X_{12}, Y_{12})$.
- * Focus is on the 12 differences because these represent change over 6 months:

$$\begin{aligned} d_1 &= (Y_1 - X_1) \\ &\vdots \\ d_{12} &= (Y_{12} - X_{12}) \end{aligned}$$

- * Among n=12 participants, we observe $\bar{d}_{n=12} = -20.17$.

Null Hypothesis Probability Model.

The observed 12 differences in hair cocaine content is a sample, $d_1 \dots d_{12}$, from a Normal population with unknown mean μ_d but known standard deviation $\sigma_d = 23.15$

 H_0 and H_A .

$$H_0 : \mu_d = 0$$

$$H_A : \mu_d < 0 \quad \text{one sided} \quad \text{The investigator is hopeful for a reduction in use}$$

The test statistic is a Z-Score when the Variance is Known.

$$z_{score} = \left[\frac{\bar{d} - E(\bar{d} | H_0 \text{ true})}{SE(\bar{d} | H_0 \text{ true})} \right]$$

Proof by Contradiction and the definition of the p-value calculation .

The likelihood of these findings or ones more extreme when we assume that the null hypothesis H_0 is true and we apply it to our observed data is

$$\text{p-value} = \Pr[\bar{d}_{n=12} \leq -20.17 | \mu_d = 0].$$

P-Value Calculation.

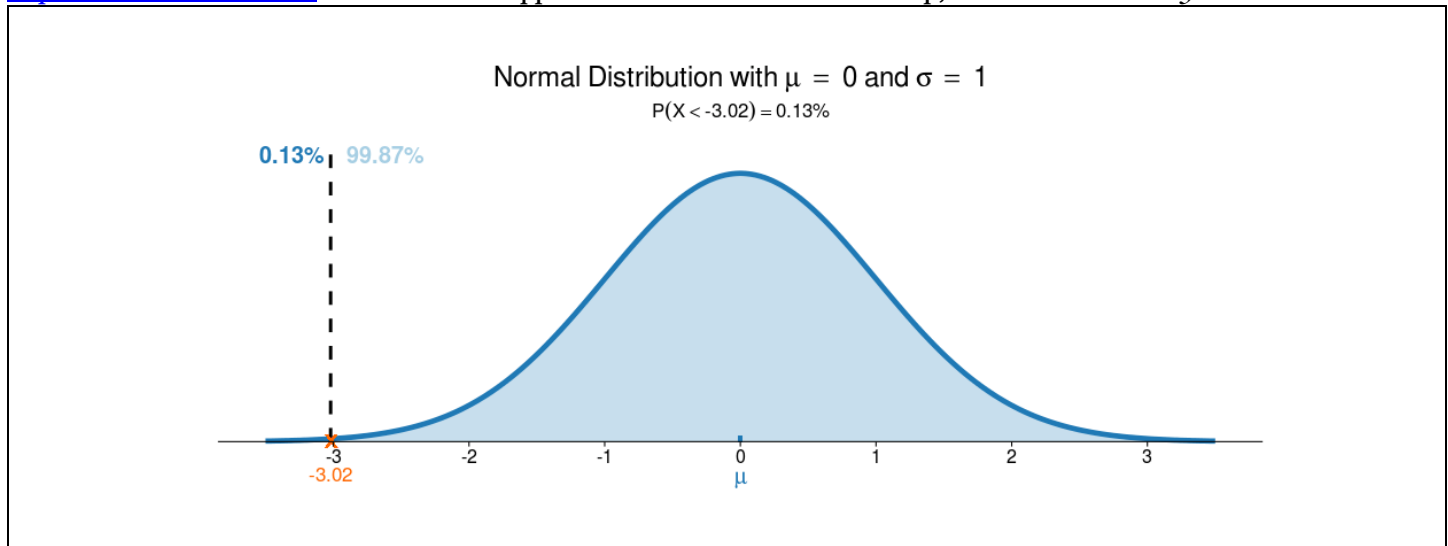
When the null hypothesis is true, the $d_1 \dots d_{12}$ are a sample from a Normal ($\mu_d = 0$, $\sigma_d^2 = 23.15^2$) distribution.

Therefore, when the null is true, $\bar{d}_{n=12}$ is distributed Normal ($\mu = 0$, $\sigma_{\bar{d}}^2 = \left[\frac{23.15^2}{12} \right]$)

p-value =

$$\begin{aligned} \Pr[\bar{d}_{n=12} \leq -20.17] &= \Pr\left[\left(\frac{\bar{d}_{12} - 0}{\sigma_d / \sqrt{n}}\right) \leq \left(\frac{-20.17}{23.15 / \sqrt{12}}\right)\right] \\ &= \Pr[\text{Normal}(0,1) \leq -3.02] = 0.0013 \end{aligned}$$

<http://www.artofstat.com/> > Online Web Apps > Normal Distribution > At top, tab is *Find Probability*



ArtofStat link, here: <https://istats.shinyapps.io/NormalDist/>

“Evaluate”.

IF we assume that the null hypothesis model is true, then participation in the needle exchange trial in the pharmacy sales condition has no effect on cocaine use,

THEN there was a null hypothesis probability of .1% ($0.0013 = 0.1\%$ chance) of obtaining an observed sample mean change in hair content of -20.17 or greater reduction among 12 participants.

Interpret.

The null hypothesis, when assumed to be true and applied to the observed data, has led to an unlikely outcome, a 0.1% chance event. Abandon the null hypothesis. We conclude that there is statistically significant evidence of a “participation effect” on 6-month change over time in hair content of cocaine.

So. Having rejected the null hypothesis, now what? Consider that there is more than one possible explanation for what we have observed, including at least the following:

- Trial participation results in less use of cocaine; *and/or*
- Trial participation results in less detection of use of cocaine.

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

4d Confidence Interval for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ is Unknown

When the $\sigma^2_{\text{DIFFERENCE}}$ is **Unknown**, a confidence interval for the mean difference $\mu_{\text{DIFFERENCE}}$ utilizes percentiles of the **student-t distribution**.

Consider again the example of paired data introduced in 4b, page 30

source: Anderson TW and Sclove SL. *Introductory Statistical Analysis*. Boston: Houghton Mifflin, 1974. page 339

A researcher is interested assessing the improvement in reading skills upon completion of the second grade (Y) in comparison to those prior to the second grade (X). For each child, his or her improvement is measured using the difference “d” which is defined $d = Y - X$. A sample of $n=30$ children are studied.

ID	PRE(X)	POST(Y)	d=(Y-X)
1	1.1	1.7	0.6
2	1.5	1.7	0.2
3	1.5	1.9	0.4
4	2.0	2.0	0.0
5	1.9	3.5	1.6
6	1.4	2.4	1.0
7	1.5	1.8	0.3
8	1.4	2.0	0.6
9	1.8	2.3	0.5
10	1.7	1.7	0.0
11	1.2	1.2	0.0
12	1.5	1.7	0.2
13	1.6	1.7	0.1
14	1.7	3.1	1.4
15	1.2	1.8	0.6
16	1.5	1.7	0.2
17	1.0	1.7	0.7
18	2.3	2.9	0.6
19	1.3	1.6	0.3
20	1.5	1.6	0.1
21	1.8	2.5	0.7
22	1.4	3.0	1.6
23	1.6	1.8	0.2
24	1.6	2.6	1.0
25	1.1	1.4	0.3
26	1.4	1.4	0.0
27	1.4	2.0	0.6
28	1.5	1.3	-0.2
29	1.7	3.1	1.4
30	1.6	1.9	0.3

Suppose $\sigma^2_d = \text{UNKNOWN}$

Solution for a 99% Confidence Interval for μ_d when σ_d^2 is UNknown

Step 1 – Point Estimate of μ_d is the Sample Mean $\bar{d}_{n=30}$

$$\bar{d}_{n=30} = \frac{\sum_{i=1}^n d_i}{n=30} = 0.51$$

Step 2 – The unknown standard Error of \bar{d}_n is now estimated by S_d/\sqrt{n}

Calculate the sample variance of the individual differences:

$$S_d^2 = \frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1} = 0.2416$$

The estimated variance of the sample mean of the differences is therefore

$$\text{estimated variance}(\bar{d}_{n=30}) = \frac{S_d^2}{n} = \frac{0.2416}{30}$$

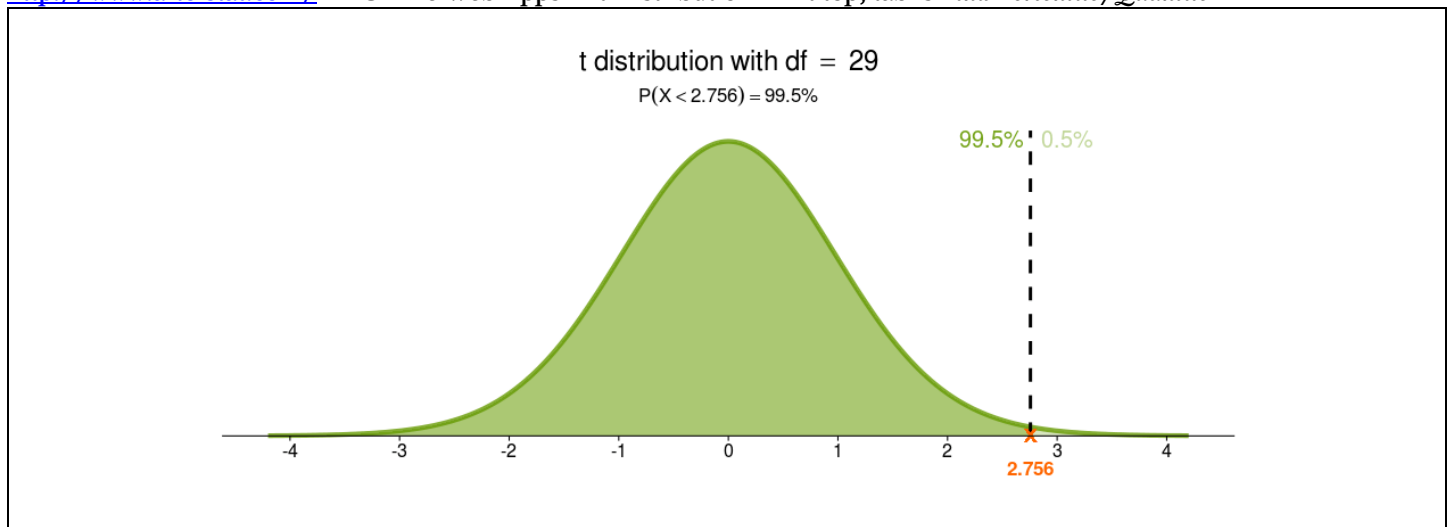
Thus,

$$SE(\bar{d}_{n=30}) = \sqrt{\text{variance}(\bar{d}_{n=30})} = \frac{S_d}{\sqrt{n}} = \frac{\sqrt{0.2416}}{\sqrt{30}} = 0.0897$$

Step 3 – The Confidence Coefficient

For a 99% confidence interval, this number will be the 99.5th percentile of the Student's t-Distribution that has degrees of freedom = (n-1) = 29. This value is **2.756**.

<http://www.artofstat.com/> > Online Web Apps > t Distribution > At top, tab is *Find Percentile/Quantile*



ArtofStat link, here: <https://istats.shinyapps.io/tdist/>

Step 4 – Substitute into the formula for a confidence interval

$$\begin{aligned}\text{Lower limit} &= (\text{point estimate}) - (\text{confidence coefficient}) (\text{SE of point estimate}) \\ &= 0.51 - (2.756) (0.0897) \\ &= 0.2627\end{aligned}$$

$$\begin{aligned}\text{Upper limit} &= (\text{point estimate}) + (\text{confidence coefficient.}) (\text{SE of point estimate}) \\ &= 0.51 + (2.756) (0.0897) \\ &= 0.7573\end{aligned}$$

Thus, a 99% confidence interval for the mean change in reading skills (Post – Pre) is = **(0.2627, 0.7573)**

Art of Stat Illustration

Solution:

KEY! This is solved as a **SINGLE sample** confidence interval for the mean using the differences $d = (\text{post} - \text{pre})$

In Art of Stat: Online Web Apps > Inference for a Population Mean

At left, make the following selections:

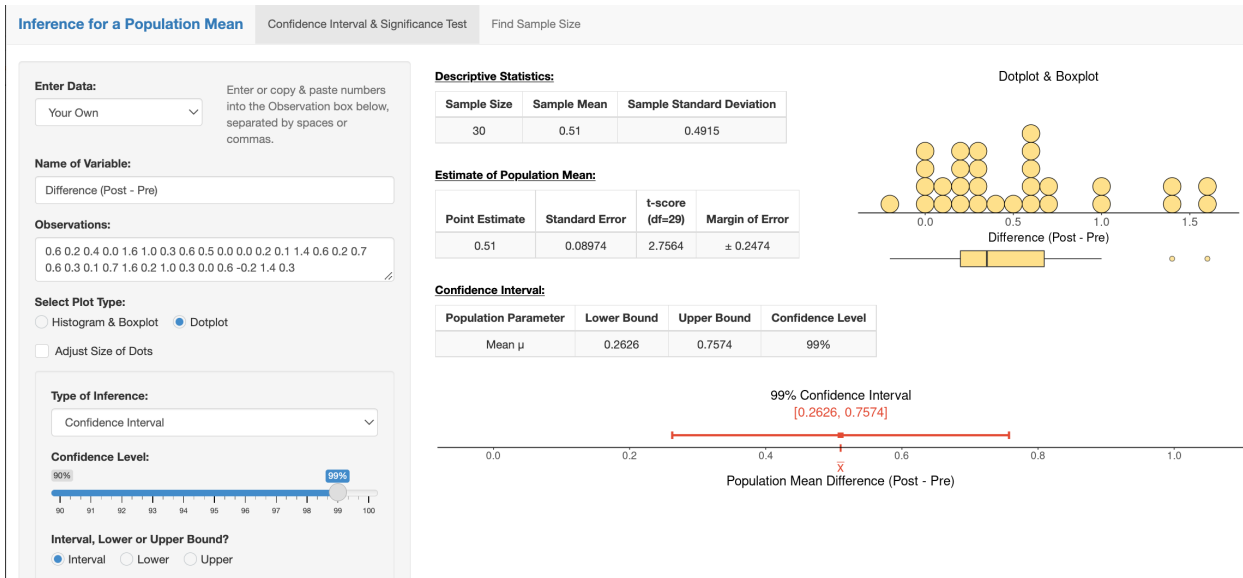
Enter Data: **Your Owns**

Name of Variable: **Difference (Post – Pre) ... or whatever name you like ...**

Observations: **Enter your differences = (post – pre) separated by spaces**

Type of Inference: **Confidence Interval**

Confidence Level: **99%**



R Illustration

```
# EXAMPLE
datatable=read.table(text="
id      pre      post
1       1.1     1.7
2       1.5     1.7
....  rows omitted ...
29      1.7     3.1
30      1.6     1.9",header=TRUE)

data4d <- as.data.frame.matrix(datatable)
data4d$d <- data4d$post - data4d$pre

yn <- length(data4d$d)           # sample size
ydf <- yn-1                     # degrees of freedom = (n-1)
ymean <- mean(data4d$d,na.rm=TRUE) # sample mean of differences (post-pre)
ysd <- sd(data4d$d, na.rm=TRUE)  # sample sd of differences (post-pre)
ysem <- ysd/sqrt(yn)             # standard deviation of the mean of differences
multiplier <- qt(.995, df=ydf)  # for 99% CI, multiplier = 99.5th percentile of Student-t, df=(n-1)
CI_lower <- ymean - multiplier*ysem # lower limit of 95% CI
CI_upper <- ymean + multiplier*ysem # upper limit of 95% CI

table4d<- cbind(ymean,yn,ysd,CI_lower,CI_upper)
colnames(table1) <- c("mean (post-pre)", "n", "sd", "lower 95% CI", "upper 95% CI")
table4d
```

	mean (post-pre)	n	sd	lower 95% CI	upper 95% CI
[1,]	0.51	30	0.4915493	0.2626303	0.7573697

4e. Hypothesis Test for $\mu_{\text{DIFFERENCE}}$, $\sigma^2_{\text{DIFFERENCE}}$ is Unknown

When the $\sigma^2_{\text{DIFFERENCE}}$ is **Unknown**, a hypothesis test for the mean difference $\mu_{\text{DIFFERENCE}}$ utilizes a p-value calculation from the **student-t distribution**.

Consider again the example of 4c, page 33

(Note: These data are hypothetical.)

Twelve patients (n=12) in a needle exchange trial who were randomized to the pharmacy sales alone condition provided hair samples that were positive for cocaine at the baseline interview. Follow-up hair samples were obtained from these same n=12 at the 6 month visit, yielding paired data. The question is: Is participation *itself* associated with reduced hair content of cocaine?

Research Question.

In the absence of an effect of study participation, it is expected that cocaine use would be stable over time. Accordingly, the hair content of cocaine would be expected to be the same at the baseline and follow-up visits. Does participation alone in an intervention study reduce cocaine use?

- * Let the 12 pairs of cocaine measurements be denoted $(X_1, Y_1) \dots (X_{12}, Y_{12})$.
- * Focus is on the 12 differences because these represent change over 6 months:

$$\begin{aligned} d_1 &= (Y_1 - X_1) \\ &\vdots \\ d_{12} &= (Y_{12} - X_{12}) \end{aligned}$$

- * Among n=12 participants, we observe $\bar{d}_{n=12} = -20.17$.

Suppose now that the sample standard deviation of the differences is $s_d = 23.15$

Null Hypothesis Probability Model.

The observed 12 differences in hair cocaine content is a sample, $d_1 \dots d_{12}$, from a Normal population with unknown mean μ_d with *unknown variance*.

H_0 and H_A .

$$H_0 : \mu_d = 0$$

$$H_A : \mu_d < 0 \quad \text{one sided} \quad \text{The investigator is hopeful for a reduction in use}$$

Test statistic is a **T-Score** when the Variance is **UNKnown**.

$$t_{\text{score}} = \left[\frac{\bar{d} - E(\bar{d} \mid H_0 \text{ true})}{\hat{SE}(\bar{d} \mid H_0 \text{ true})} \right]$$

Solution for the unknown standard error of \bar{d}_n is now estimated by S_d / \sqrt{n}

$$\hat{SE}(\bar{d}_{n=30}) = \sqrt{\text{variance}(\bar{d}_{n=30})} = \frac{S_d}{\sqrt{n}} = \frac{\sqrt{23.15}}{\sqrt{12}} = 1.3889$$

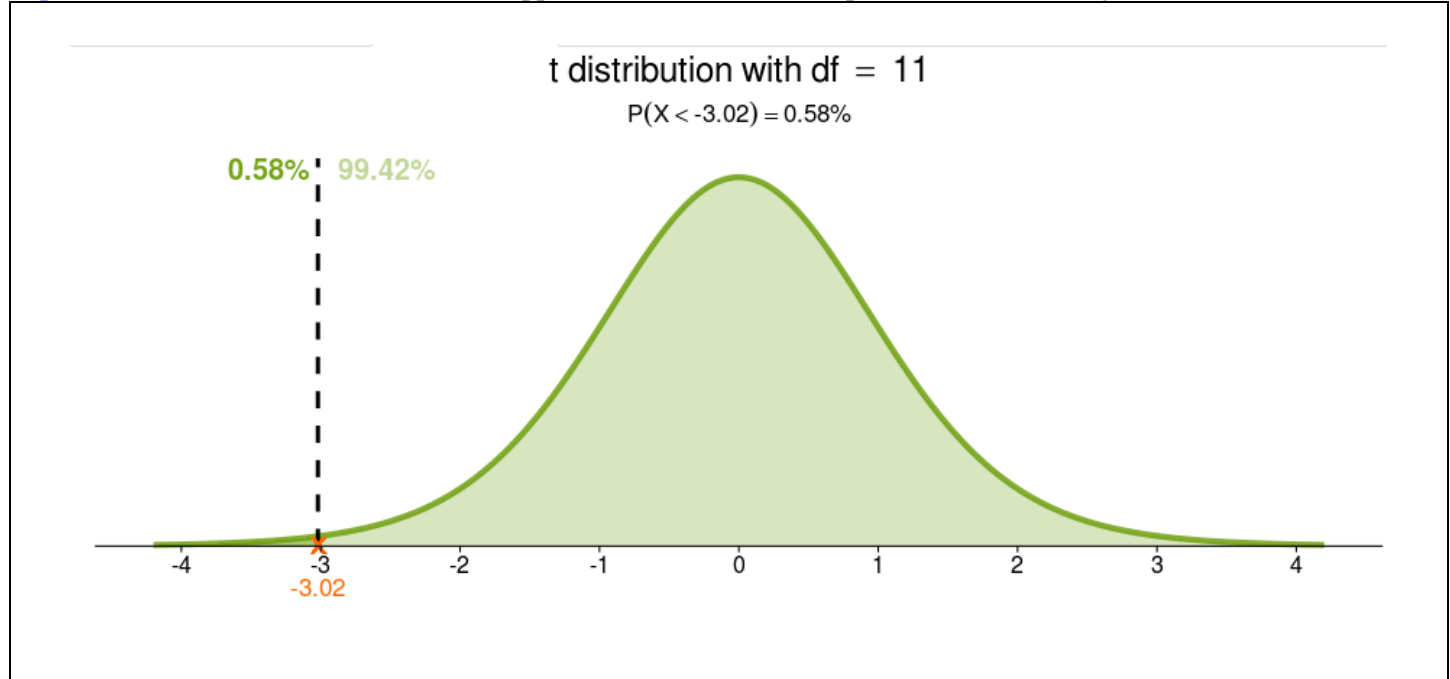
P-Value Calculation.

p-value

$$= \text{pr}[\bar{d}_{n=12} \leq -20.17] = \text{pr}\left[\left(\frac{\bar{d}_{12} - 0}{S_d / \sqrt{n}}\right) \leq \left(\frac{-20.17}{23.15 / \sqrt{12}}\right)\right]$$

$$= \text{pr}[\text{Student's } t_{DF=11} \leq -3.02] = 0.00583 \quad \text{notice – this is bigger than the .0013 on page 40(Okay, I admit; we're talking a tiny tiny amount!).}$$

<http://www.artofstat.com/> > Online Web Apps > t Distribution > At top, tab is *Find Probability*



ArtofStat link, here: <https://istats.shinyapps.io/tdist/>

Interpret.

The conclusion is the same as before.

The null hypothesis, in light of the data, has led to an unlikely outcome, a 0.1% chance event. Abandon the null hypothesis. We conclude that there is statistically significant evidence of a “participation effect” on 6-month change over time in hair content of cocaine.

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

HOMEWORK DUE Monday November 21, 2022

Question #5 of 8

Halcion is a sleeping pill that is relatively rapidly metabolized by the body and therefore has fewer hangover effects the next morning, compared to other sleeping pills. Opponents of Halcion argue that, because this agent is so rapidly metabolized by the body, patients do not sleep as long with this drug as with Dalmane. Data on 10 insomniacs, each of whom took Dalmane on one occasion and Halcion on a second, is collected. The variable measured is number of hours of sleep:

Patient	Number of Hours Sleep with	
	Dalmane	Halcion
1	4.58	3.97
2	5.19	4.88
3	3.94	4.09
4	6.32	5.87
5	7.68	6.93
6	3.48	4.00
7	5.72	5.08
8	7.04	6.95
9	5.27	4.96
10	5.84	5.13

Do these data suggest that Halcion is not as effective as Dalmane with respect to number of hours of sleep? Carry out an appropriate statistical test and interpret your findings. You may assume that the measurements of sleep are continuous, distributed normal.

HOMEWORK DUE Monday November 21, 2022

Question #6 of 8

For the Halcion versus Dalmane data in Question 5:

Construct a 99% confidence interval estimate of discrepancy in the efficacies of the two drugs.

Compare this to the acceptance region that would have been obtained had you constructed a statistical test with type I error pre-specified at 0.01.

4f. Confidence Interval for $\sigma^2_{\text{DIFFERENCE}}$

Tip -

For a single sample of paired data from a normal distribution, a confidence interval for $\sigma^2_{\text{DIFFERENCE}}$ utilizes percentiles from the chi square distribution.

Nice! The formula here is the same as what we saw before (see again, page 21) because we are working with a single sample from a normal distribution.

(1- α)100% Confidence Interval for $\sigma^2_{\text{DIFFERENCE}}$ Setting – Normal Distribution	
Lower limit =	$\frac{(n-1)S^2_{\text{DIFFERENCE}}}{\chi^2_{1-\alpha/2}}$
Upper limit =	$\frac{(n-1)S^2_{\text{DIFFERENCE}}}{\chi^2_{\alpha/2}}$

Consider again the example of 4b, page 36

source: Anderson TW and Sclove SL. *Introductory Statistical Analysis*. Boston: Houghton Mifflin, 1974. page 339. A researcher is interested assessing the improvement in reading skills upon completion of the second grade (Y) in comparison to those prior to the second grade (X). For each child, their improvement over the course of the school year is measured using the difference “d” which is defined $d = Y - X$. A sample of $n=30$ children are studied

Obtain a 90% Confidence Interval for σ_d^2 .

Step 1 - Obtain the point estimate of σ_d^2 .

$$S_d^2 = \frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1} = 0.2416$$

Step 2 - Determine the correct chi square distribution to use.

It has $df = (30-1) = 29$.

Step 3 - Obtain the correct confidence coefficient multipliers.

Because the desired confidence level is 0.90 we set $0.90 = (1-\alpha)$. Thus $\alpha = .10$

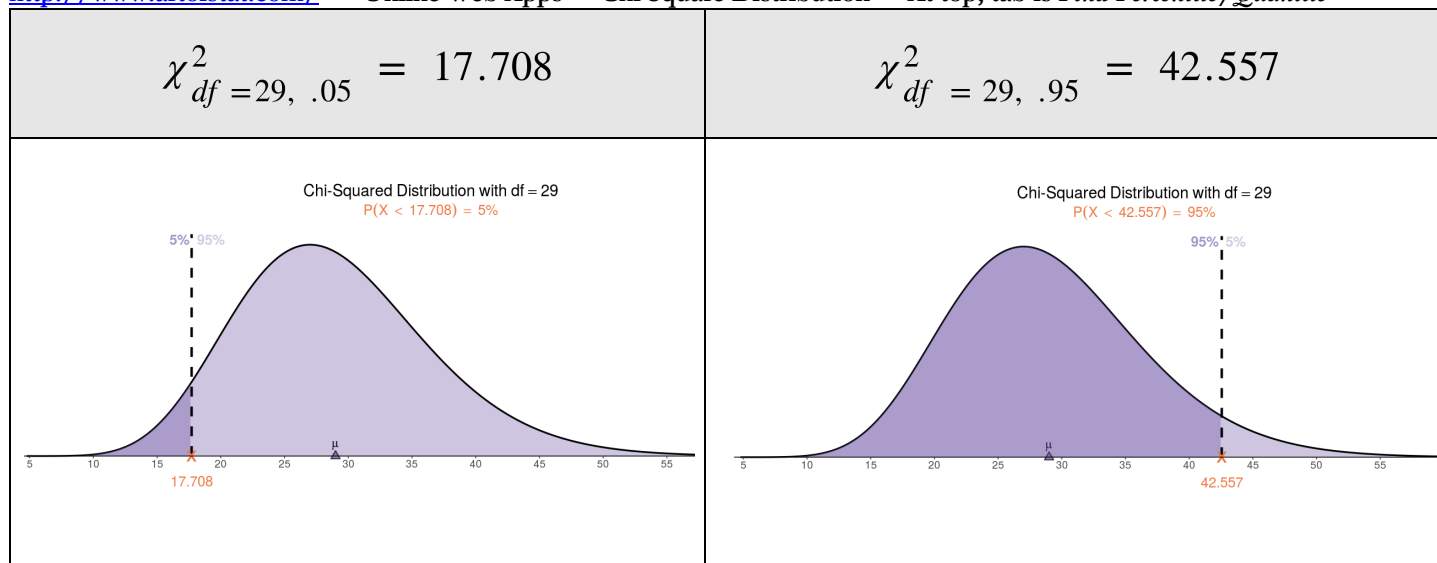
For a 90% confidence level, the percentiles we want are

- (i) $(\alpha/2)100^{\text{th}} = 5^{\text{th}}$ percentile
- (ii) $(1 - \alpha/2)100^{\text{th}} = 95^{\text{th}}$ percentile

Obtain percentiles for chi square distribution with degrees of freedom = 29

- (i) $\chi_{df=29, .05}^2 = 17.708$
- (ii) $\chi_{df=29, .95}^2 = 42.557$

<http://www.artofstat.com/> > Online Web Apps > Chi Square Distribution > At top, tab is *Find Percentile/Quantile*



ArtofStat link, here: <https://istats.shinyapps.io/ChisqDist/>

Step 4 – Substitute into the formula for the confidence interval

$$(i) \text{ Lower limit} = \frac{(n-1)S_d^2}{\chi^2_{1-\alpha/2}} = \frac{(29)(0.2416)}{42.557} = 0.1646357$$

$$(ii) \text{ Upper limit} = \frac{(n-1)S_d^2}{\chi^2_{\alpha/2}} = \frac{(29)(0.2416)}{17.708} = 0.395663$$

Thus, an 80% confidence interval for the unknown variance is = **(0.16, 0.40)**

Illustration Using GraphPad (Sorry, Art of Stat does not have a utility for this)

(Search on “GraphPad Online confidence interval of a SD)

<https://www.graphpad.com/quickcalcs/CIstd1/?Format=50>

Confidence interval of a SD

It is straightforward to calculate the standard deviation from a bunch of values. But how accurate is the standard deviation? Just by chance you may have happened to obtain data that are closely bunched together, making the SD low. Or you may have happened to obtain data that are far more scattered than the overall population, making the SD high. If you assume that your data are randomly sampled from a population that follows a Gaussian distribution, This calculator can compute a 95% confidence interval for the standard deviation.

1. Choose data entry format

Caution: Changing format will erase your data.

☐ Enter up to 50 rows.

☐ Enter or paste up to 10000 rows.

☐ Enter SEM and N.

☒ Enter SD and N.

2. Enter data

SD:

N:

Scroll down to click on: **CALCULATE NOW**

Confidence interval of a SD

Parameter	Value
SD	0.491500
SEM	0.089735
N	30
90% CI of the SD	0.405730 to 0.628976
95% CI of the SD	0.391433 to 0.660731
99% CI of the SD	0.365867 to 0.730695

GraphPad returns 90% CI for SD of differences $\sigma_d = 0.41 - 0.63$
 Squaring these values yields 90% CI for Variance $\sigma_d^2 = 0.16 - 0.40$

R Illustration

```

my_vardiff <- 0.2416 # User supplies sample variance of differences
my_n <- 30 # User supplies number of pairs

my_sddiff <- sqrt(my_vardiff) # sample standard deviation of differences
my_df <- my_n - 1 # degrees of freedom = (n-1)
chisq_upper <- qchisq(.95,df=my_df) # 95th percentile of Chi Square with df = (n-1)
chisq_lower <- qchisq(.05,df=my_df) # 5th percentile of Chi Square with df = (n-1)
lowerci_var <- my_df*my_vardiff/chisq_upper # lower limit of 90% CI for variance differences
upperci_var <- my_df*my_vardiff/chisq_lower # upper limit of 90% CI for variance differences
lowerci_sd <- sqrt(lowerci_var) # lower limit of 90% CI for SD differences
upperci_sd <- sqrt(upperci_var) # upper limit of 90% CI for SD differences

table2 <- cbind(my_vardiff,lowerci_var,upperci_var)
colnames(table2) <- c("Sample variance of differences", "lower 90% CI", "upper 90% CI")
table2

table3 <- cbind(my_sddiff,lowerci_sd,upperci_sd)
colnames(table3) <- c("Sample Standard Deviation of differences", "lower 90% CI", "upper 90% CI")
table3

```

	Sample variance of differences	lower 90% CI	upper 90% CI
[1,]	0.2416	0.1646358	0.3956548

	Sample Standard Deviation of differences	lower 90% CI	upper 90% CI
[1,]	0.4915282	0.4057534	0.629011

Nature ————— Population/ Sample ————— Observation/ Data ————— Relationships/ Modeling ————— Analysis/ Synthesis

5. One Sample – Binomial Distribution: π

5a. Confidence Interval for π , normal approximation approach

Tip -

For a single sample from a binomial, a confidence interval for the probability of event π *always* utilizes percentiles of the **standard normal distribution**. There is no separate solution for a “variance unknown” scenario because the variance is $\pi(1-\pi)$

Recall – The **Binomial** Distribution was introduced in Unit 6 - *Bernoulli and Binomial Distributions*

- We have **n independent Bernoulli trials**. To be general, let’s call the two possible outcomes “event” and “non-event”.
- “**Event/non-event**” might refer to: “alive/dead”, “tumor/remission”, “success/failure”, “heads/tails”, etc.
- In each Bernoulli trial, the outcome of “event” occurs with the same probability = π
- Suppose that “event” occurs in x of the n trials. An estimate of π is the proportion of trials in which event occurred. This is equal to x/n .
- The number of occurrences of event, X, in n independent Bernoulli trials is distributed **Binomial(n, π)**.

(With apology) There are a variety of notations for representing an estimate of π

The most clear is $\hat{\pi}$. The **caret** on the top is an indication that this is a guess.

- Another notation for $\hat{\pi}$ is p for “proportion”. This is awkward because some texts use the notation “p” is used for the population parameter π itself. Therefore, I recommend against using this to represent the estimate of π .
- A better choice for a notation for $\hat{\pi}$ is to write it as \hat{p} because it has the caret on top.
- Still another notation for $\hat{\pi}$ is X/n . This is nice because you can recognize it as the observed proportion.
- Still another is \bar{X} . This also makes sense since it is the sum n Bernoulli outcomes coded as of 0’s (non-events) and 1’s (events), divided by n, the number of trials. *Putting these all together ...*
- **Summarizing: $\hat{\pi}$ is equivalently written: $\hat{\pi} = \hat{p} = \bar{X} = X/n$. Notice, I left off the notation “p”.**

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

In constructing a confidence interval for π of a Binomial distribution - just as we did for the mean parameter μ of a Normal distribution – we need:

1. Point estimate
2. SE of the point estimate
3. Confidence coefficient

Example –

(Source: Daniel WW. *Biostatistics: A Foundation for Analysis in the Health Sciences*)

Suppose we are interested in estimating the proportion of individuals who obtain a dental check up twice a year in a certain urban population. In a simple random sample of $n=300$ persons, $X=123$ reported having had 2 dental check ups in the last year. Construct a 95% confidence interval for π , the unknown true proportion.

1. The Point Estimate of π is the Sample Mean $\hat{\pi} = \bar{X}$

$$\bar{X} = \frac{X}{n} = \frac{123}{300} = 0.41$$

2. The Standard Error of $\hat{\pi} = \bar{X}$ is estimated using $SE(\hat{\pi}) = \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$

This formula makes sense for two reasons:

- If X is distributed Binomial(n, π) Then $\text{Variance}(X) = n \pi (1-\pi)$
- $\text{Variance}[(\text{constant})X] = (\text{constant})^2 \text{Variance}(X)$
- For the interested: Appendix 3 is the solution for this SE formula.

$$SE(\hat{\pi}) = \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} = \sqrt{\frac{0.41(0.59)}{300}} = 0.028$$

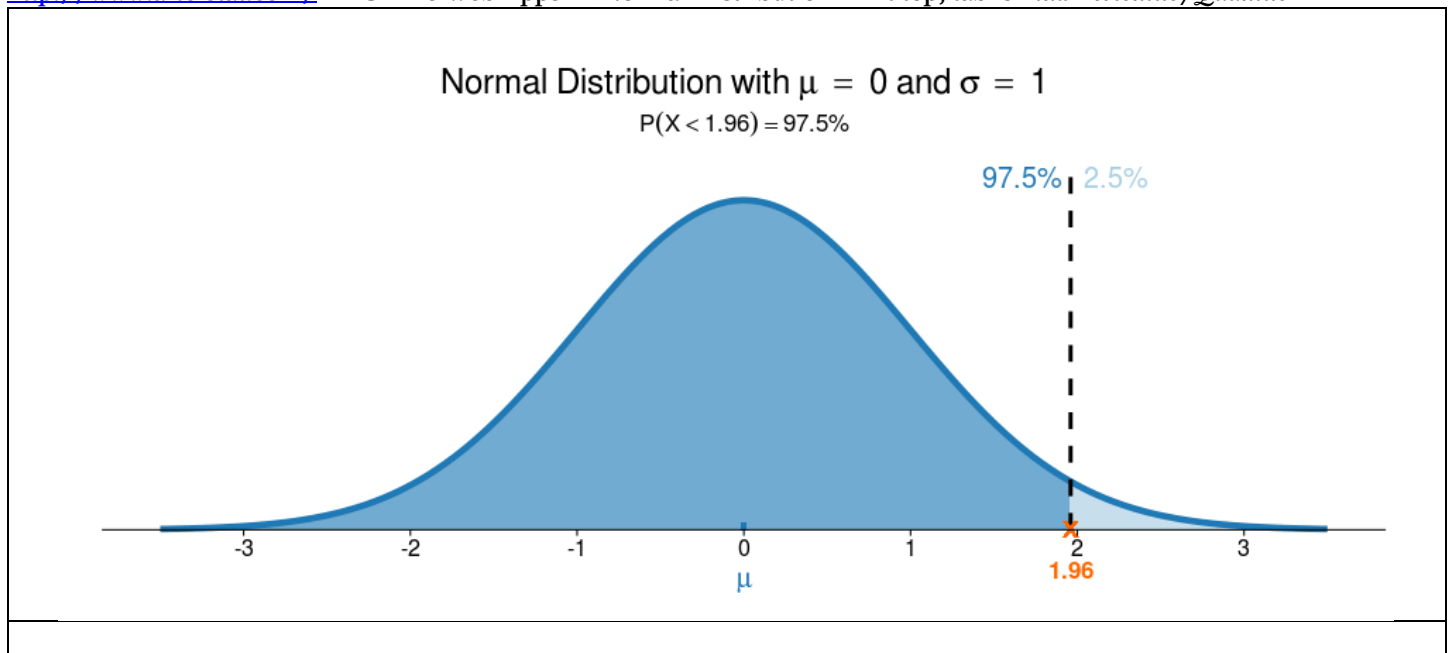
3. The Confidence Coefficient is a Percentile from the Normal(0,1) Distribution

Reminder!!! Do NOT use percentiles from the Student t-distribution here! The correct percentile is one from the Normal(0,1) for reasons having to do with the central limit theorem.

- ♣ As we saw before - For a 95% confidence interval, this number will be the 97.5th percentile of the Normal (0,1) distribution.
- ♣ And in general - For a $(1-\alpha)100\%$ confidence interval, this number will be the $(1-\alpha/2)100^{\text{th}}$ percentile of the Normal (0,1) distribution.

$$z_{.975} = 1.96$$

<http://www.artofstat.com/> > Online Web Apps > Normal Distribution > At top, tab is *Find Percentile/Quantile*



ArtofStat link, here: <https://istats.shinyapps.io/NormalDist/>

4. Putting it all together.

Lower limit of confidence interval

$$\begin{aligned}
 &= (\text{point estimate}) - (\text{confidence coefficient multiplier}) (\text{SE of estimate}) \\
 &= 0.41 - (1.96)(0.028) \\
 &= 0.36
 \end{aligned}$$

Upper limit of confidence interval

$$\begin{aligned}
 &= (\text{point estimate}) + (\text{confidence coefficient multiplier}) (\text{SE of estimate}) \\
 &= 0.41 + (1.96)(0.028) \\
 &= 0.46
 \end{aligned}$$

Thus, a 95% confidence interval for the binomial probability of event is = **(0.36, 0.46)**

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

Art of Stat Illustration

Solution:

In Art of Stat: Online Web Apps > *Inference for a Proportion* > Choose tab: *Confidence Interval & Significance Test*

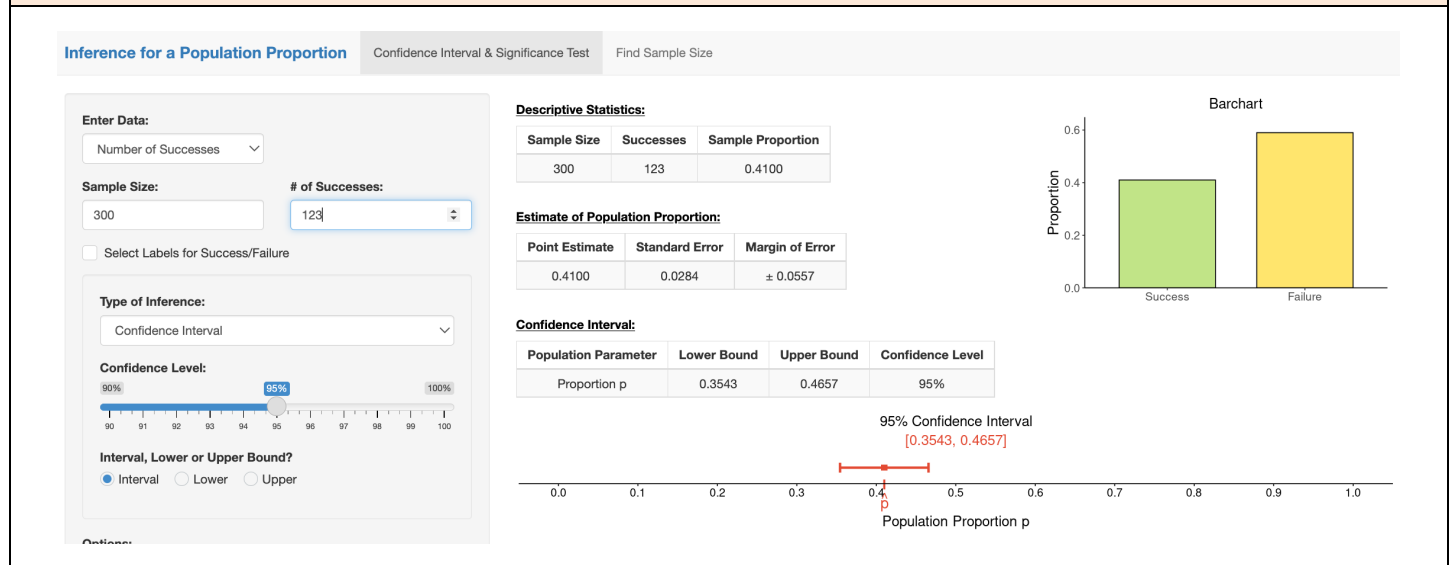
At left, make the following selections:

Enter Data: *Number of Successes*

Sample Size: **300** # Successes: **123**

Type of Inference: *Confidence Interval*

Confidence Level: **95%**



R Illustration

```
x <- 123 # User supplies X = # successes
ntrials <- 300 # User supplies ntrials
my_conf <- .95 # User supplies desired confidence level

xbar <- x/ntrials # sample proportion of success
sem <- sqrt(xbar*(1-xbar)/ntrials) # standard error of sample proportion
crit_percentile <- 1 - (1 - my_conf)/2 # percentile for desired confidence
z_crit <- qnorm(p=crit_percentile) # confidence coefficient multiplier
lowerci <- xbar - z_crit*sem # lower confidence interval limit
upperci <- xbar + z_crit*sem # upper confidence interval limit

table1 <- cbind(x, ntrials, xbar, lowerci, upperci)
colnames(table1) <- c("Successes", "Trials", "Proportion Success", "lower 95% CI", "upper 95% CI")
table1

binom.test(x=xevents, n=ntrials, conf.level=my_conf)$conf.int # Alternative method using binom.test( )

      Successes Trials Proportion Success lower 95% CI upper 95% CI
[1,]      123    300      0.41    0.3543448    0.4656552

[1] 0.3538120 0.4679837 # Solution using binom.test( )
attr(,"conf.level")
[1] 0.95
```

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

Confidence Interval for a binomial event probability (proportion) π , normal approximation
Single sample from a Binomial (n, π) Distribution

$$\hat{\pi} \pm (z_{1-\alpha/2}) \hat{SE}(\hat{\pi})$$

where the required calculations are

(1) $\bar{X} = \frac{X}{n}$ the observed proportion of events in the n trials

(2) $\hat{\pi} = \bar{X}$

(3) $\hat{SE} = \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$

(4) For small number of trials ($n \leq 30$ or so) use $\hat{SE} = \sqrt{\frac{0.5(0.5)}{n}}$

Why? For small number of trials n (say $n \leq 30$), it may be desirable to compute a more conservative (wider) confidence interval by using a slightly different SE calculation.

- A closer look at the SE calculation $\hat{SE} = \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$ reveals that it attains its largest value when $\bar{X} = 0.50$

5b. Hypothesis Test for π - Exact Test

Tip -

For a single sample from a binomial, an “exact” hypothesis test for the probability of event π utilizes a p-value calculation using the **binomial distribution**.

The exact test is used when the sample size is small

See page 56 for conditions for using a normal approximation test

Research Question:

In an ICU study, data were collected on 20 consecutive patients. Four (4) of the patients died in the hospital. Is there evidence that the mortality rate at Baystate Medical Center is different than 25%?

Null Hypothesis Probability Assumptions

- The data is the outcome of some Binomial random variable with number of trials $n=20$: Binomial ($n=20, \pi$).
- The number of events (mortality) observed in this sample of 20 trials is $X=4$

H_0 and H_A

$$H_0 : \pi = 0.25$$

$$H_A : \pi \neq 0.25 \quad \text{two sided}$$

Tip – Calculate the null hypothesis model value of the expected number of events of mortality. This is $E[X | H_0 : \pi = 0.25]$. You will need this to calculate your p-value.

$$E[X | H_0 : \pi = 0.25] = [n][0.25] = [20][0.25] = 5$$

The Test Statistic is X.

X = number of events of mortality

Proof by Contradiction Reasoning and the definition of the p-value calculation.

Because the alternative hypothesis is two sided, the p-value calculation we want to do here answers the question:

“If H_0 is true,

Then what are the chances of obtaining a number of events of death as far away from $E[X | H_0 : \pi = 0.25] = 5$ in either direction?”

How to determine the two tails in a two sided p-value calculation for a Binomial model

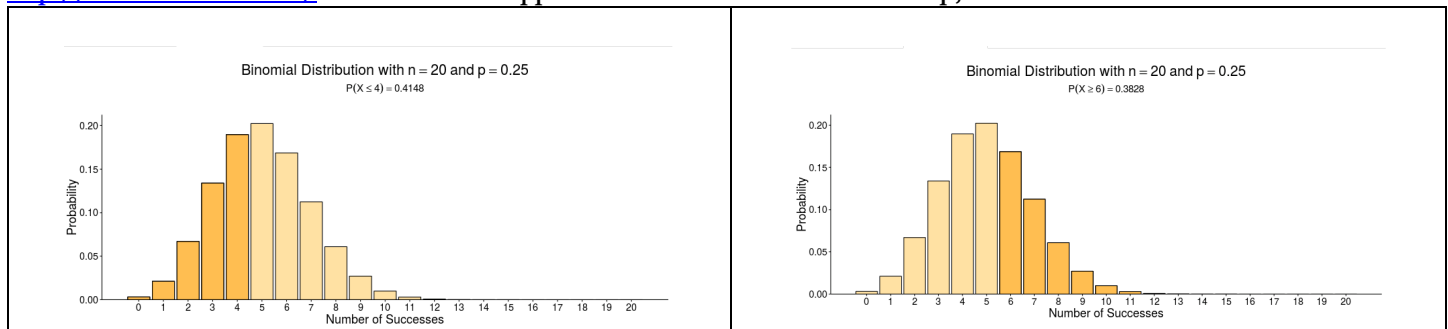
In doing p-value calculations utilizing the Normal(0,1) or the Student-t distributions, the null hypothesis expected value was zero. This made it easy to understand the two tails in p-value calculations. In contrast, for the Binomial, the null hypothesis expected value is NOT 0.

But we still want our p-value calculation to satisfy “*as far away from the expected value as what was observed in either direction.*” In this example, the observed value of 4 deaths is 1 death different from the mean 5 in the left direction. 1 death different from the mean 5 in the other direction is 6 deaths. Thus,

P-Value Calculation, two-sided.

$$\begin{aligned} \text{p-value, two-sided} &= \text{Prob} [\text{Binomial}(20, .25) \geq 6] + \text{Prob} [\text{Binomial}(20, .25) \leq 4] \\ &= 0.4148 + 0.3828 \\ &= 0.7976 \end{aligned}$$

<http://www.artofstat.com/> > Online Web Apps > Binomial Distribution > At top, tab is *Find Probabilities*



ArtofStat link, here: <https://istats.shinyapps.io/BinomialDist/>

“Evaluate”.

Under the null hypothesis H_0 , that the mortality rate at Baystate is 0.25, the likelihood of an observed mortality rate as small or smaller than 4/20 OR as large or larger than 6/20 is approximately 80 chances in 100. **Aside: this is obviously too small a study to be useful!**

Interpret.

The assumption of the null hypothesis model has not led to an unlikely outcome. Do NOT reject the null hypothesis. Conclude that these data do not provide statistically significant evidence for its rejection.

Nature _____ Population/ Sample _____ Observation/ Data _____ Relationships/ Modeling _____ Analysis/ Synthesis

Statistics Kingdom Illustration

Google: Statistics Kingdom One Sample Proportion Test

https://www.statskingdom.com/111proportion_normal1.html

Solution:

Make the following selections:

Tails: **Two**

Proportion or total number (x): **4**

Expected Proportion: **.25**

Sample size (n): **20**

At bottom: **Calculate binomial test**

One Sample Proportion Test Proportion Z-test and Binomial test

[Video](#) [Two sample proportion calculator](#)

Tails Two ($H_1: p \neq p_0$)		Digits 6	
Significance level (α): 0.05		Continuity True	
h effect size 0.5		Calculate the expected h effect size Calculate h	
Name Group 1		Expected proportion (P_0) .25	
Proportion (\hat{p}) or total number (x) 4		Sample size (n) 20	

Calculate binomial test Calculate z test Clear

One sample proportion binomial test, using binomial distribution (two-tailed) (validation)

1. H_0 hypothesis

Since $p\text{-value} > \alpha$, H_0 cannot be rejected.

The proportion (\hat{p}) of **Group 1's** population is assumed to be **equal to** the expected proportion (P_0).

In other words, the difference between the proportion (\hat{p}) of **Group 1** and the expected proportion (P_0) is not big enough to be statistically significant.

2. P-value

The p-value equals **0.797669**, ($p(X \leq 0) = 0$). It means that the chance of type I error, rejecting a correct H_0 , is too high: 0.7977 (79.77%).

The larger the p-value the more it supports H_0 .

3. The statistics

The test statistic X equals **4**, which is in the 95% region of acceptance: [2 : 9].

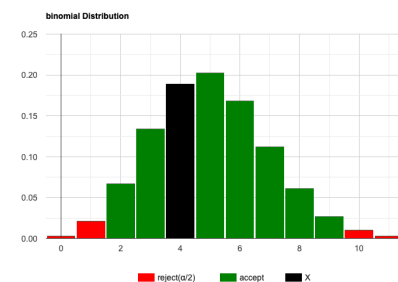
$\hat{p}_1 = 0.2$, is in the 95% region of acceptance: [0.1 : 0.45].

The standard deviation of the difference, S' equals 0.0968, is used to calculate the statistic.

4. Effect size

The observed effect size h is **small**, **0.12**. This indicates that the magnitude of the difference between the proportion (\hat{p}) and expected proportion (P_0) is small.

If you like the page, please share or like. Questions, comments and suggestions are appreciated. (statskingdom@gmail.com)



R Illustration

```

xevents <- 4
ntrials <- 20
nullp <- .25
myalt <- "two.sided"

# User supplies X = # successes
# User supplies ntrials
# User supplies null hypothesis Pr["success"]
# User Chooses "two.sided", "less", "greater"

binom.test(x=xevents,n=ntrials,p=nullp,alternative=myalt)
Exact binomial test

data: xevents and ntrials
number of successes = 4, number of trials = 20, p-value = 0.7977
alternative hypothesis: true probability of success is not equal to 0.25
95 percent confidence interval:
 0.057334 0.436614
sample estimates:
probability of success
      0.2
    
```

HOMEWORK DUE Monday November 21, 2022

Question #7 of 8

An entomologist samples a field for egg masses of a harmful insect by placing a yard-square frame at random locations and carefully examining the ground within the frame. A simple random sample of 75 locations selected from a county's pasture land found egg masses in 13 locations.

Compute a 95% confidence interval estimate of all possible locations that are infested.

HOMEWORK DUE Monday November 21, 2022

Question #8 of 8

OPTIONAL (for the brave!)

Dear BIOSTATS 540 class - This question is NOT a mimicking of the lecture notes. It is asking you to start your thinking from the width of a confidence interval and then reason your solution from there - cb.

Alzheimers' disease has a poorer prognosis when it is diagnosed at a relatively young age. Suppose we want to estimate the age at which the disease was first diagnosed using a 90% confidence interval. Under the assumption that the distribution of age at diagnosis is normal, if the population variance is $\sigma^2=85$, how large a sample size is required if we want a confidence interval that is 10 years wide?

Hint. Confidence Interval Width = [Upper Limit] - [Lower Limit]

5c. Hypothesis Test for π - Normal Approximation

Tip -

For a single sample from a binomial, a “normal approximation” hypothesis test for the probability of event π always utilizes a p-value calculation using the **standard normal distribution**.

When Can I do a “Normal Approximation” Test? Guideline and Definition of the Normal Approximation Test

Guidelines: Use a Z-score test approximation when the sample size is moderate or large. We can do this because the central limit theorem applies (proof not shown). As a rough rule of thumb, you can use the following normal approximation test when the following holds:

$$(n)(\pi_{\text{null}})(1 - \pi_{\text{null}}) \geq 5$$

Definition: \bar{X} is distributed Normal($\pi, \frac{\pi(1-\pi)}{N}$) approximately

Research Question:

In an ICU study, data was collected on 200 consecutive patients. 40 of the patients died in the hospital. Is there evidence that the mortality rate at Baystate Medical Center is different (in either direction) than 25%?

Null Hypothesis Probability Assumptions

- Data are a random sample of patients (over time), and the outcome of mortality, X =(# patients among the 200 who die in hospital) has exact distribution that is Binomial ($N=200, \pi$).
- The observed number of events of mortality is $X=40$
- A quick check of the guidelines at the top of this page yields $(n)(\pi)(1 - \pi) = (200)(.25)(.75) = 37.5$. Thus, we have plenty of sample size for using a normal Z-score approximation test. In this setting, application of the central limit theorem to a binomial random variable outcome gives the following:

\bar{X} is distributed Normal($\pi, \frac{\pi(1-\pi)}{N}$) approximately

- The observed proportion (the observed % who died) is $\bar{X}=40/200=0.20$

Nature _____ Population/
Sample _____ Observation/
Data _____ Relationships/
Modeling _____ Analysis/
Synthesis

H_0 and H_A .

$$H_0 : \pi = 0.25$$

$$H_A : \pi \neq 0.25 \quad \text{two sided}$$

The Test statistic is a z-score.

$$Z\text{-score} = \frac{\bar{X} - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{N}}} = \frac{\bar{X} - 0.25}{\sqrt{\frac{0.25(0.75)}{200}}}$$

“Proof by Contradiction Reasoning and the definition of the p-value calculation.

Under the assumption that the null hypothesis is true, the probability of obtaining an observed sample mortality **as different** from the expected value of 25% as the observed 20% is a two sided calculation:

$$p\text{-value} = (2) \Pr \left[\text{Normal}(0,1) \leq \left(\frac{\bar{X} - 0.25}{\sqrt{\frac{0.25(0.75)}{200}}} \right) \right].$$

P-Value Calculation.

$$p\text{-value} = (2) \Pr \left[\text{Normal}(0,1) \leq \left(\frac{0.20 - 0.25}{\sqrt{\frac{0.25(0.75)}{200}}} \right) \right] = (2) \Pr [\text{Normal}(0,1) \leq -1.63]$$

$$= (2) (0.051)$$

$$= 0.102$$

“Evaluate”.

Under the null hypothesis H_0 , that the mortality rate at Baystate is 0.25, the probability of an observed mortality rate as far away (in either direction) as 20% is .102, or approximately 10 chances in 100. The null hypothesis, when applied to the data, has not led to an unlikely outcome.

Interpret.

Do NOT reject the null hypothesis. Conclude that, in this sample, the observed mortality rate of 20% is consistent with the hypothesized rate of 25%.

Art of Stat Illustration

Solution:

In Art of Stat: Online Web Apps > **Inference for a Proportion** > Choose tab: **Confidence Interval & Significance Test**

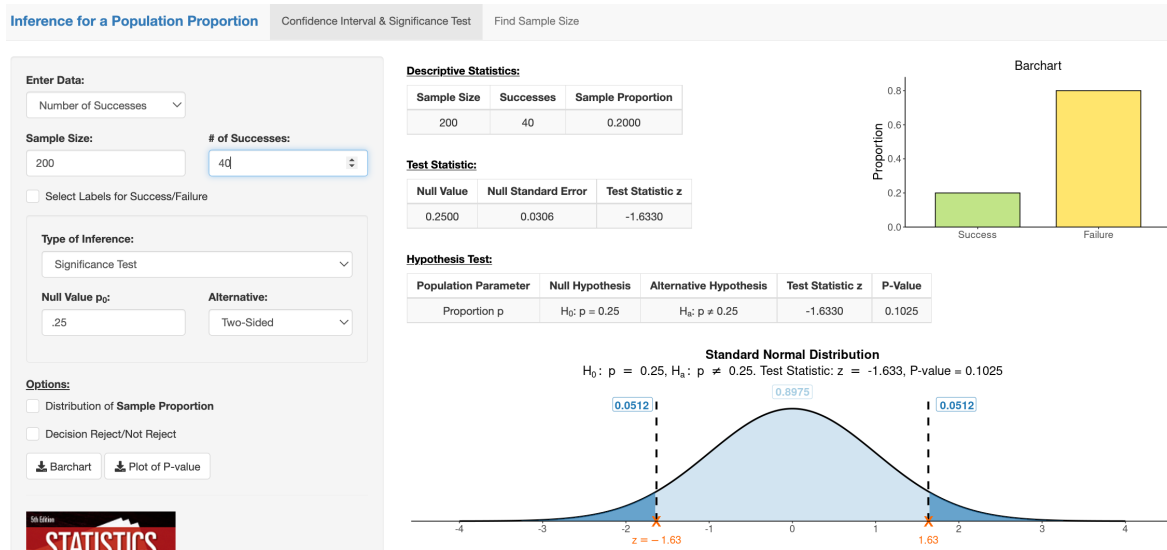
At left, make the following selections:

Enter Data: **Number of Successes**

Sample Size: **200** # Successes: **40**

Type of Inference: **Significance Test**

Null value: **.25** Alternative: **Two-Sided**



R Illustration

```

xevents <- 40
ntrials <- 200
nullp <- .25
myalt <- "two.sided"

# User supplies X = # successes
# User supplies ntrials
# User supplies desired confidence level
# CHOOSE "two.sided", "less", "greater"

prop.test(x=40,n=200,p=.25,alternative="two.sided")
    
```

1-sample proportions test with continuity correction

```

data: 40 out of 200, null probability 0.25
X-squared = 2.4067, df = 1, p-value = 0.1208 # p-value differs from Art of Stat because of continuity correction
alternative hypothesis: true p is not equal to 0.25
95 percent confidence interval:
 0.1482520 0.2635577
sample estimates:
 p
0.2
    
```

Nature — Population/ Sample — Observation/ Data — Relationships/ Modeling — Analysis/ Synthesis

Appendices

A. Derivation of Confidence Interval for μ : Single Sample from Normal, σ^2 known

The setting is the example in Section 2a (Confidence Interval for μ , σ^2 known).

Recall that we were given the weight in micrograms of drug inside each of 30 capsules, after subtracting the capsule weight.

0.6	0.3	0.1	0.3	0.3
0.2	0.6	1.4	0.1	0.0
0.4	0.5	0.6	0.7	0.6
0.0	0.0	0.2	1.6	-0.2
1.6	0.0	0.7	0.2	1.4
1.0	0.2	0.6	1.0	0.3

We're told that $\sigma^2 = 0.25$

Step 1 – Obtain a point estimate \bar{X}

$$\bar{X} = 0.51$$

$n = 30$

Step 2 – Obtain the SE of the point estimate \bar{X} by recalling that $SE(\bar{X}) = \sigma / \sqrt{n}$

$$SE(\bar{X}) = \sigma / \sqrt{n} = 0.5 / \sqrt{30}$$

$$= 0.0913$$

Step 3 – Select desired confidence = $(1 - \alpha)$

Suppose we want a 95% confidence interval.

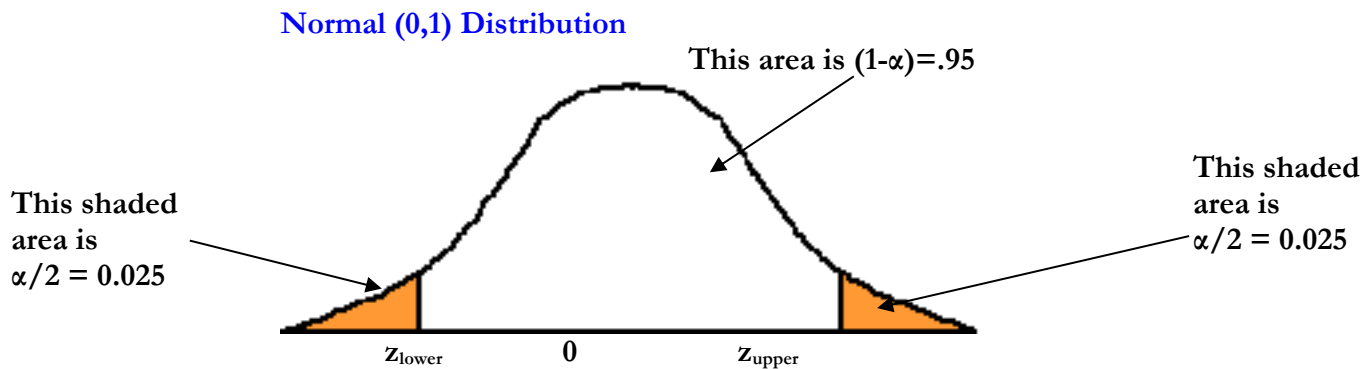
Then $(1 - \alpha) = 0.95$.

This means that $\alpha = 0.05$

The $\alpha = 0.05$ is the probability of error

Step 4 – Using a calculator for the Normal (0,1) distribution, obtain symmetric values of a standard normal deviate Z (call these z_{lower} and z_{upper}) such that

$$\text{Probability} [z_{\text{lower}} \leq Z \leq z_{\text{upper}}] = 0.95$$



$$\text{Probability} [-1.96 \leq Z \leq +1.96] = 0.95 \text{ so that}$$

$$z_{\text{lower}} = -1.96$$

$$z_{\text{upper}} = +1.96$$

This expression, $\text{Probability} [-1.96 \leq Z \leq +1.96] = 0.95$ in this example.

More generally, it is $\text{Probability} [z_{\text{lower}} \leq Z \leq z_{\text{upper}}] = (1 - \alpha)$

To get to this expression, standardize \bar{X}

$\text{Probability} [-1.96 \leq Z \leq +1.96] = 0.95$ in this example is actually

$$\text{Probability} [z_{\text{lower}} \leq Z \leq z_{\text{upper}}] = (1 - \alpha) \rightarrow$$

note #1 - Because the Normal(0,1) distribution is symmetric about the value 0

$$z_{\text{lower}} = (-1) z_{\text{upper}}$$

So let's call z_{upper} simply z

This allows us to simplify the above expression with two convenient substitutions

$$z_{\text{upper}} = z$$

$$z_{\text{lower}} = -z$$

Probability $[-z \leq Z \leq z] = (1 - \alpha) \rightarrow$

note #2 - Now we'll insert another convenient substitution

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

Probability $[-z \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq z] = (1 - \alpha) \rightarrow$

note #3 - All that remains is to do the algebra necessary to “isolate” μ

Probability $\left[\left(\frac{\sigma}{\sqrt{n}} \right) -z \leq \bar{X} - \mu \leq \left(\frac{\sigma}{\sqrt{n}} \right) z \right] = (1 - \alpha) \rightarrow$

With confidence $(1 - \alpha)100\%$, $\left[\bar{X} - \left(\frac{\sigma}{\sqrt{n}} \right) z \leq \mu \leq \bar{X} + \left(\frac{\sigma}{\sqrt{n}} \right) z \right]$ **which matches.**

B. Derivation of Confidence Interval for σ^2 : Single Sample from a Normal

The setting here is the example in Section 3a.

A precision instrument is guaranteed to read accurately to within ± 2 units. A sample of 4 readings on the same object yield 353, 351, 351, and 355. Calculate a 95% confidence interval estimate of the population variance σ^2 .

Step 1 – Obtain a point estimate S^2 and its associated degrees of freedom

$$S^2 = 3.67$$

$$\text{df} = 3$$

Step 2 – Recalling the material from section 2b, define the appropriate chi square random variable

$$Y = \frac{(n-1)S^2}{\sigma^2} \text{ is distributed Chi Square with degrees of freedom } = (n-1)$$

Step 3 – Select desired confidence = $(1 - \alpha)$

For desired confidence = .95, $(1 - \alpha) = 0.95$.

Step 4 – Substitute for χ^2 in the middle of the “area under the curve” calculation for a chi square random variable as follows.

$$\text{Probability} \left[\chi_{\text{df}; \alpha/2}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{\text{df}; (1-\alpha/2)}^2 \right] = (1 - \alpha)$$

Step 5 – Do the algebra to obtain an expression that is the confidence interval for σ^2 .

$$\text{Probability} \left[\chi_{\text{df}, \alpha/2}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{\text{df}, (1-\alpha/2)}^2 \right] = (1-\alpha) \rightarrow$$

$$\text{Probability} \left[\frac{1}{\chi_{\text{df}, (1-\alpha/2)}^2} \leq \frac{\sigma^2}{(n-1)S^2} \leq \frac{1}{\chi_{\text{df}, \alpha/2}^2} \right] = (1-\alpha) \rightarrow$$

$$\text{With confidence } (1-\alpha)100\%, \left[\frac{(n-1)S^2}{\chi_{\text{df}, (1-\alpha/2)}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{\text{df}, \alpha/2}^2} \right] \text{ which matches.}$$

C. The Standard Error of $\hat{\pi} = \bar{X}$ is estimated using $SE(\hat{\pi}) = \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}$

We can take advantage of two results -

- If X is distributed Binomial(n,π) Then Variance(X) = n π (1-π)
- Variance[(constant)X] = (constant)² Variance (X)

Proof

$$SE(\bar{X}) = \sqrt{\text{Variance}(\bar{X})}$$

$$= \sqrt{\text{Variance}\left(\frac{X}{n}\right)}$$

$$= \sqrt{\left(\frac{1}{n^2}\right) (\text{Variance}[X])}$$

$$= \sqrt{\left(\frac{1}{n^2}\right) (n\pi[1-\pi])}$$

$$= \sqrt{\frac{\pi[1-\pi]}{n}}$$

The problem now is that π is not known. So it is replaced by its estimate

$$\approx \sqrt{\frac{\bar{X}[1-\bar{X}]}{n}}$$

which matches