Hypothesis Tests for Means using Confidence Intervals

Donna Dietz

American University

dietz@american.edu

STAT 202 - Spring 2020

What we know

From our last discussion, we know that we can set up an interval that surrounds the proportion we found in a sample, and that's an estimator for where we expect the proportion of the population to actually be.

$$\hat{p} - z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

What we want

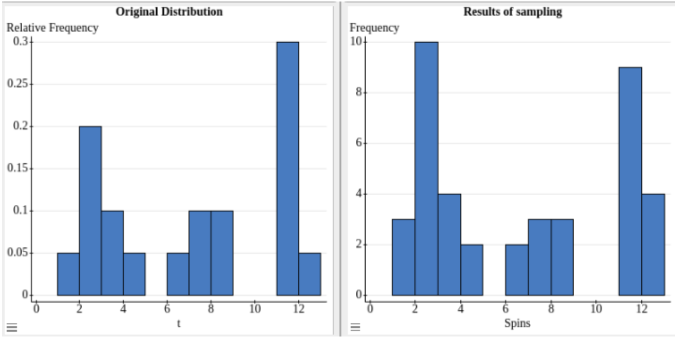
We would like to generalize this for any underlying distribution.
Proportions are very convenient because

- The distribution is simple: Two bars
- The population proportion must be between 0 and 1
- The properties of the Binomial problem help us analyse it

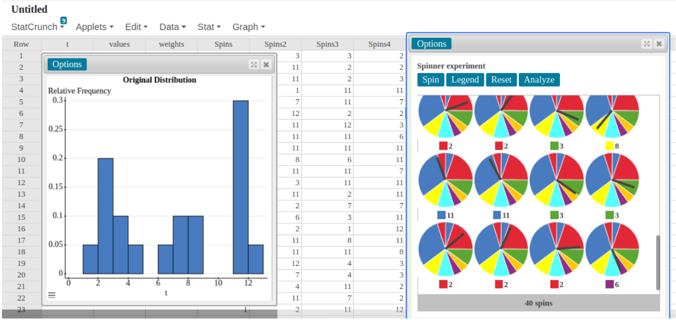
We don't have those benefits generally, but we do get lucky anyhow.

Let's consider a different distribution

Maybe what you want is to recover the entire shape of the distribution. Here is a quick example of what that could look like. The more samples you take, the closer your sampling results will match the original distribution.

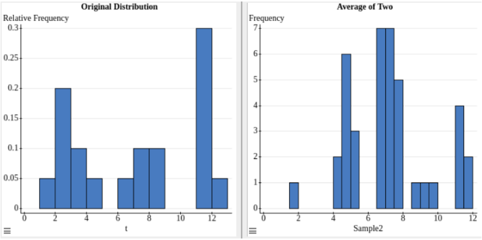


StatCrunch has a Spinner tool



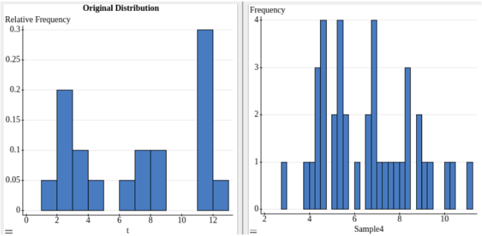
Starting with an original distribution, turning it into a Spinner, then running the Spinner 40 times, this is what StatCrunch looks like.

Average of two



This is one example of taking forty samples of two. The second plot shows the average of two spins.

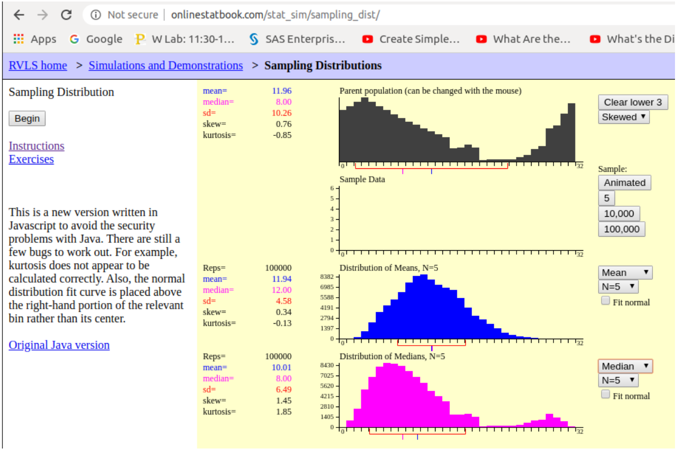
Average of four



This is what happened when I took forty samples of four. Notice that the gaps are filling in, the middle is becoming fuller, and tails are starting to form.

Online Simulator

There is a cool online simulator at:
http://onlinestatbook.com/stat_sim/sampling_dist



You can pick a starting distrubution or use your mouse to mess up one of theirs. Then, ask it to repeatedly draw from the distribution!

CLT = Central Limit Theorem

One of the biggest theorems in Statisics is that pretty much any distribution you may wish to start with, at least in the real world, can do this. Pick your favorite distribution to start out with. Then take samples of increasing size. So, if $n = 100$, you will probably already find that your sample means form a nice bell-like curve. The higher n goes, the closer to normal you will be.

Bags of candy

I like to imagine bags of candy. The distrubtion of weights for your original candy population can be just about anything, but then if you're making bags of candy, maybe 100 pieces per bag, randomly selected, then you record the total weight of each bag, it's that total (or average) for the **bags** which will end up looking normal.

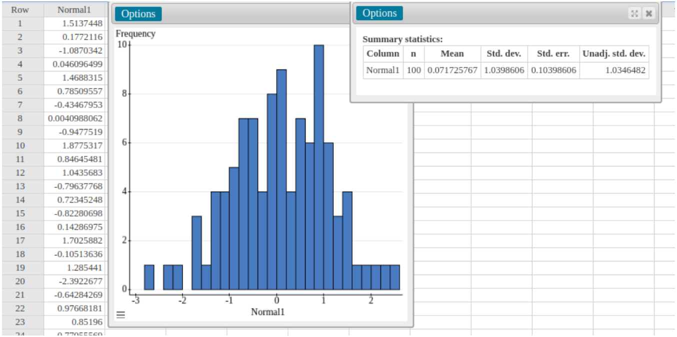
So we can just be Normal, OK?

So, let's transition a bit. Since any sampling distribution we might wish to use will end up as a Normal distribution, we might as well start with one that's Normal.

Comparing σ and s

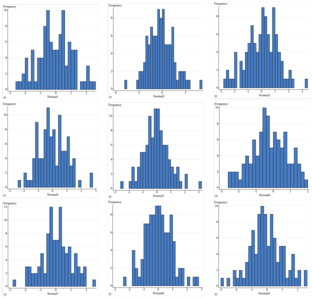
If you recall, we have two standard deviations. We have the population standard deviation or σ, and we have the (so-called) sample standard deviation or s. How different are they anyhow? (Not too different!) Let's see!

Sampling from $N(0,1)$



In StatCrunch, I randomly drew from $N(0,1)$ 100 times. This histogram shows the results of those 100 draws using bins of 0.2. This is to give you a sense of how close (or not) we got to having a mean of zero and standard deviation of 1. Notice that we are fairly close, and also that σ and s are very close.

Let's do it nine more times!



This is to give you a sense for what these things might look like, and how different or similar various draws of 100 might be!

Statistics for all ten samples of 100

Summary statistics:

Column	n	Mean	Std. dev.	Std. err.	Unadj. std. dev.
Normal0	100	0.071725767	1.0398606	0.10398606	1.0346482
Normal1	100	-0.054873605	1.0946809	0.10946809	1.0891938
Normal2	100	-0.10401209	1.060599	0.1060599	1.0552827
Normal3	100	0.014887071	1.1552807	0.11552807	1.1494897
Normal4	100	0.01947448	0.96685056	0.096685056	0.96200416
Normal5	100	-0.041099674	0.94130858	0.094130858	0.93659021
Normal6	100	-0.21841773	1.0409649	0.10409649	1.0357469
Normal7	100	0.040032095	0.97712309	0.097712309	0.9722252
Normal8	100	0.037875974	0.92813455	0.092813455	0.92348221
Normal9	100	-0.027868925	1.0715811	0.10715811	1.0662098

You can see that in all cases, we got very close. You can also see that σ (Unadj. std. dev.) and s (Std. dev.) are quite similar. Half the means are positive and half are negative. For s, 6 are too high and 4 are too low. For σ, it's the same. They are off by around 1%, which is expected with $n = 100$.

For samples of size 10

Summary statistics:
Where: Row<11

Column	n	Mean	Std. dev.	Unadj. std. dev.
Normal0	10	0.34031449	1.0400117	0.98664173
Normal1	10	0.076473072	0.89930813	0.85315861
Normal2	10	0.16705125	1.1887154	1.1277144
Normal3	10	0.3334111	1.3495785	1.2803226
Normal4	10	0.052336462	0.67778541	0.6430037
Normal5	10	-0.20202891	0.93244359	0.88459366
Normal6	10	0.13135013	0.54122432	0.51345047
Normal7	10	-0.378194	1.1908115	1.129703
Normal8	10	-0.11353782	0.81508627	0.77325873
Normal9	10	-0.22883429	0.97730044	0.92714861

When we lower the sample sizes, we find σ and s are farther apart, but the sampling error is way more important than the difference between σ and s.

Assumptions for using z-scores

So, before we proceed to the exercises, I need to inform you that we are only going to use z-scores in one of two special situations. Otherwise, we'll use t-scores. (We'll talk about this soon!)

The two cases which permit us to use z-scores are: If we are sampling from a Normal Distribution, or if our sample size is *large enough* to warrant it. This is dependent on your needed level of accuracy, and *large enough* is probably somewhere between 100 and 1000.

Formulas

Samples

If you take a sample of size n from a distribution with $N(\mu, \sigma)$, the sum of those items looks like $N(n\mu, \sigma\sqrt{n})$ and the mean of the collection looks like $N(\mu, \sigma/\sqrt{n})$.

Confidence Interval

For a distribution with $N(\mu, \sigma)$, for a sample of size n, 95% of the samples should fall within ± 1.96 standard deviations of the mean.

$$\mu - 1.96(\sigma/\sqrt{n}) < \bar{x} < \mu + 1.96(\sigma/\sqrt{n})$$

Flipping the confidence interval!

Confidence Interval

For a sample of size n, 95% of the actual population mean is expected to fall within ± 1.96 sample standard deviations of the sample mean.

$$\bar{x} - 1.96(s/\sqrt{n}) < \mu < \bar{x} + 1.96(s/\sqrt{n})$$

For a 99% confidence interval, use 2.576 instead of 1.96.

Example:

If you have a population which is known to be normally distributed and have a mean of 100 and a standard deviation of 18, and you draw 36 items in a sample, what is the 95% confidence interval for where you would expect the sample mean to end up? (Feel free to use $z^* = 2$ instead of 1.96.)

Example:

If you have a population which is known to be normally distributed and have a mean of 100 and a standard deviation of 18, and you draw 36 items in a sample, what is the 95% confidence interval for where you would expect the sample mean to end up?

$$100 - 2 \cdot 18/6 = 94 < \bar{x} < 106 = 100 + 2 \cdot 18/6$$

Example:

This time, presume you *took* a sample of size 49 and got a mean of 100 and a sample standard deviation of 14. What is your 95% confidence interval for where you think the population mean would be?

Example:

This time, presume you *took* a sample of size 49 and got a mean of 100 and a sample standard deviation of 14. What is your 95% confidence interval for where you think the population mean would be?

$$100 - 2 \cdot 14/7 = 96 < \mu < 104 = 100 + 2 \cdot 14/7$$

Hypotheses

Just as before when we did hypothesis testing with confidence intervals, we need H_0 to have an equal sign in it, and the H_A to negate it.

Example

A dental school student makes a claim that the average person spends 45 seconds each day brushing their teeth. You get 100 students to agree to secretly time their roommates brushing their teeth and report the times back to you. You determine the average to be 50 seconds with a standard deviation of 15 seconds.

In this example, H_0 is $\mu = 45\ sec$ and H_A is $\mu \neq 45\ sec$. For this exercise, you can create a confidence interval using the sample mean and sample standard deviation to see if the hypothesized mean is inside.

Answer:

$$50 - 2 \cdot 15/10 < \mu < 50 + 2 \cdot 15/10$$

$$47 < \mu < 53$$

Since 45 is smaller than the lower edge of the interval, you can reject the null hypothesis. Since you rejected it, you may now talk about direction. A good way to rephrase the results would be to say, "Based on our study, students here spend more than 45 seconds per day brushing their teeth."

This is a great example of something that's statistically significant but not of any practical use, since the ADA recommends 2 minutes per day of brushing, and 50 seconds is still far below that recommended guideline.

Wikipedia Says...

William Sealy Gosset (13 June 1876 – 16 October 1937) was an English statistician, chemist and brewer who served as Head Brewer of Guinness and Head Experimental Brewer of Guinness and was a pioneer of modern statistics. He pioneered small sample experimental design and analysis with an economic approach to the logic of uncertainty. Gosset published under the pen name Student and developed most famously Student's t-distribution – originally called Student's "z" – and "Student's test of statistical significance".

William Sealy Gosset

When z-scores aren't enough

In roughly 1908, William Sealy Gosset figured out that when you have small data sets, the Normal distribution doesn't do a really good job at making confidence intervals. The intervals you get by using z-scores are too narrow. So, 95% confidence intervals would work far less than 95% of the time, for example.

Gosset (under a pen name of *Student*) created new tables with more conservative t^* values. We call them t-scores and we call the table *Student's Table*.

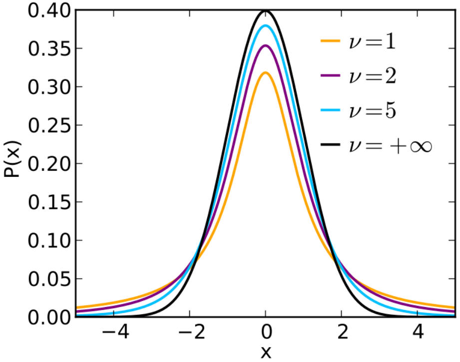
Degress of Freedom

The concept of *degrees of freedom* comes up repeatedly in math and science. For this table, however, for the way we use it in this course, you need only know that *df* means *degrees of freedom* and you simply subtract one from the sample size. When you land on the table between two lines, you fall back to the lower *df* value. The final line of Students Tables reflects z-score values.

$$df = n - 1$$

Calculations done by computer will result in narrower confidence intervals, as Student's Table will give a more conservative or wider interval when calculated by hand.

Student Family of Curves



Note that the Student curves have thicker tails and don't have quite as high of a peak in the center. However, as sample size increases, this family of curves converges to the Normal curve.

Find the t^* values

Just like you have z^* values for creating confidence intervals, you also have t^* values. They work the same way as z^* values.

Example: Find the 95% confidence interval based on a sample of size 27 that has a mean of 73 and a sample standard deviation of 16.

Answer

Example: Find the 95% confidence interval based on a sample of size 27 that has a mean of 73 and a sample standard deviation of 16.

The t^* value can be found in the Table in the row marked *df* 26 and the column with a 95% at the bottom. That value is $t^* = 2.056$. The standard deviation of the sampling distribution is approximated by $16/\sqrt{27}$. So the confidence interval would be:

$$73 - 2.056 \cdot 16/\sqrt{27} < \mu < 73 - 2.056 \cdot 16/\sqrt{27}$$

$$73 - 2.056 \cdot 16/\sqrt{27} < \mu < 73 - 2.056 \cdot 16/\sqrt{27}$$

$$66.67 < \mu < 79.33$$

On StatCrunch

You can look up the t^* value on StatCrunch using:
Stat $>$ Calculators $>$ T
Enter df and cut off the top 2.5% to find t^*.

Paper vs StatCrunch

Table entry for p and C is the critical value t^* with probability p lying to its right and probability C lying between $-t^*$ and t^*.

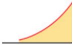
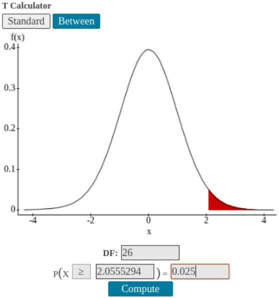


TABLE D

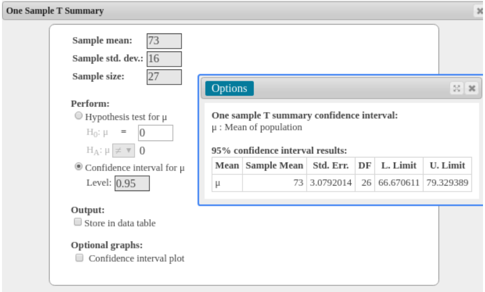
t distribution critical values

	Upper-tail					
df	.25	.20	.15	.10	.05	.025
1	1.000	1.376	1.963	3.078	6.314	12.71
2	0.816	1.061	1.386	1.886	2.920	4.303
3	0.765	0.978	1.250	1.638	2.353	3.182
4	0.741	0.941	1.190	1.533	2.132	2.776
5	0.727	0.920	1.156	1.476	2.015	2.571
6	0.718	0.906	1.134	1.440	1.943	2.447
7	0.711	0.896	1.119	1.415	1.895	2.365
8	0.706	0.889	1.108	1.397	1.860	2.306
9	0.703	0.883	1.100	1.383	1.833	2.262
10	0.700	0.879	1.093	1.372	1.812	2.228
11	0.697	0.876	1.088	1.363	1.796	2.201
12	0.695	0.873	1.083	1.356	1.782	2.179
13	0.694	0.870	1.079	1.350	1.771	2.160
14	0.692	0.868	1.076	1.345	1.761	2.145
15	0.691	0.866	1.074	1.341	1.753	2.131
16	0.690	0.865	1.071	1.337	1.746	2.120
17	0.689	0.863	1.069	1.333	1.740	2.110
18	0.688	0.862	1.067	1.330	1.734	2.101
19	0.688	0.861	1.066	1.328	1.729	2.093
20	0.687	0.860	1.064	1.325	1.725	2.086
21	0.686	0.859	1.063	1.323	1.721	2.080
22	0.686	0.858	1.061	1.321	1.717	2.074
23	0.685	0.858	1.060	1.319	1.714	2.069
24	0.685	0.857	1.059	1.318	1.711	2.064
25	0.684	0.856	1.058	1.316	1.708	2.060
26	0.684	0.856	1.058	1.315	1.706	2.056



Whole process on StatCrunch

Stat > T Stats > One Sample > With Summary



Sample Size

As before, we can determine the sample size needed for the confidence interval based on the desired margin of error and confidence level.

$$n \geq \frac{z^2 s^2}{(ME)^2}$$

$$n \geq \frac{t^2 s^2}{(ME)^2}$$

From data

To calculate a confidence interval based on data, place it in a single column in StatCrunch.

Stat > T Stats > One Sample > With Data

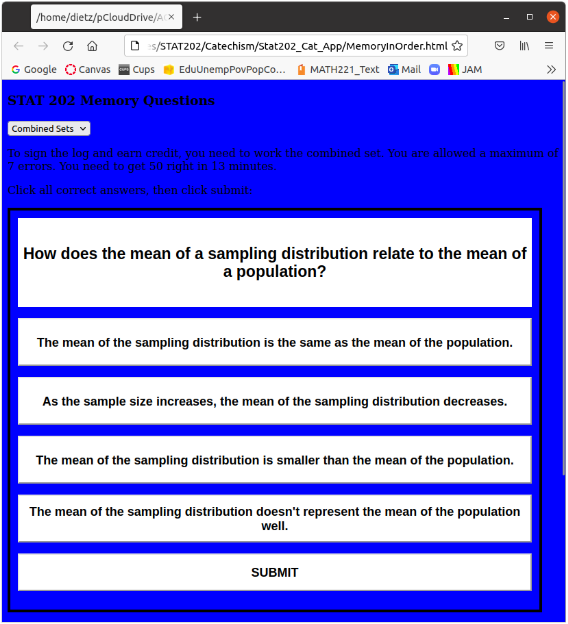
var1
530
570
950
1550
500
950
1200
1570
625

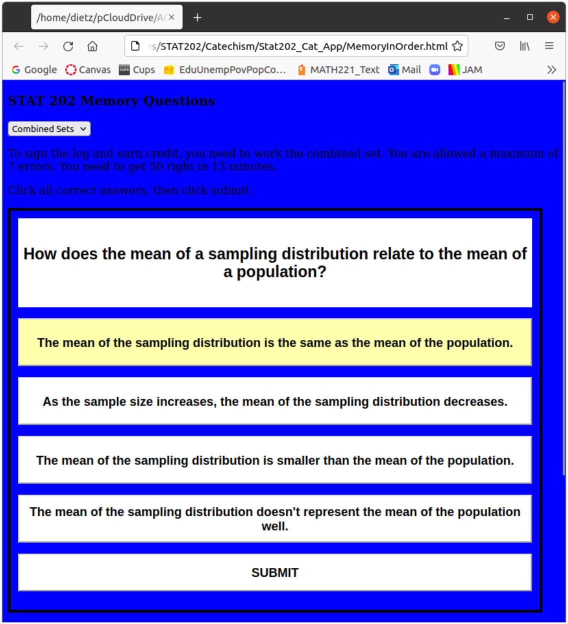
One sample T confidence interval:
μ : Mean of variable

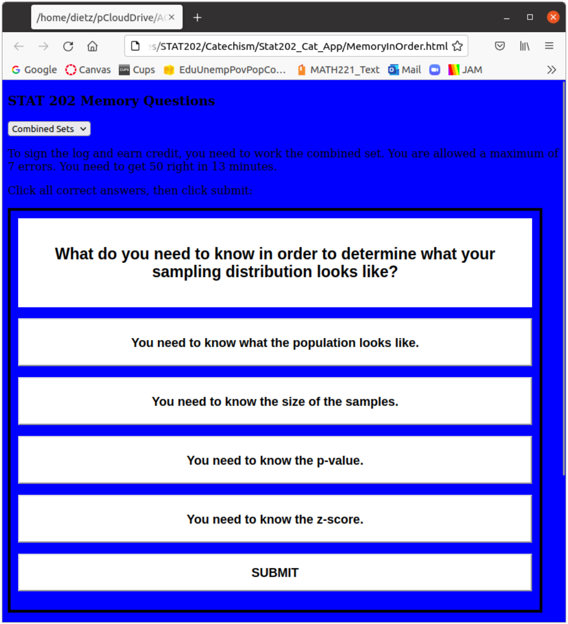
95% confidence interval results:

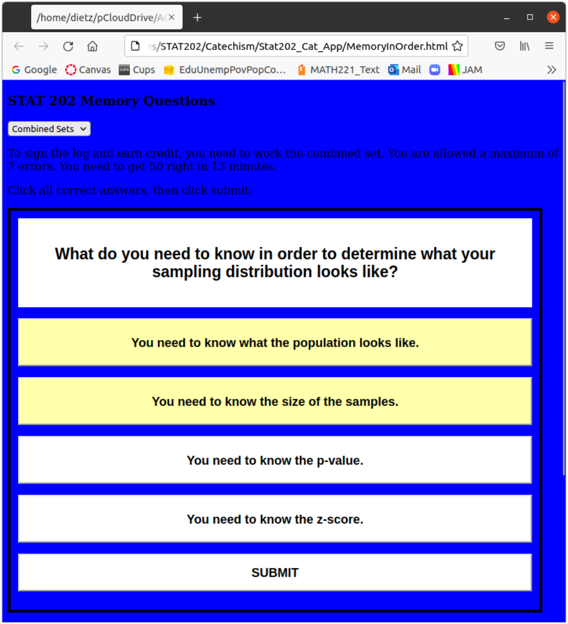
Variable	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
var1	938.33333	141.15594	8	612.82715	1263.8395

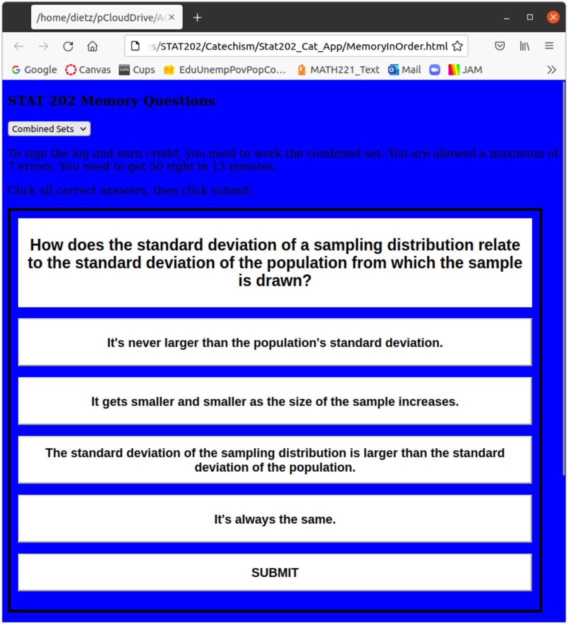
6 MEMORY QUESTIONS

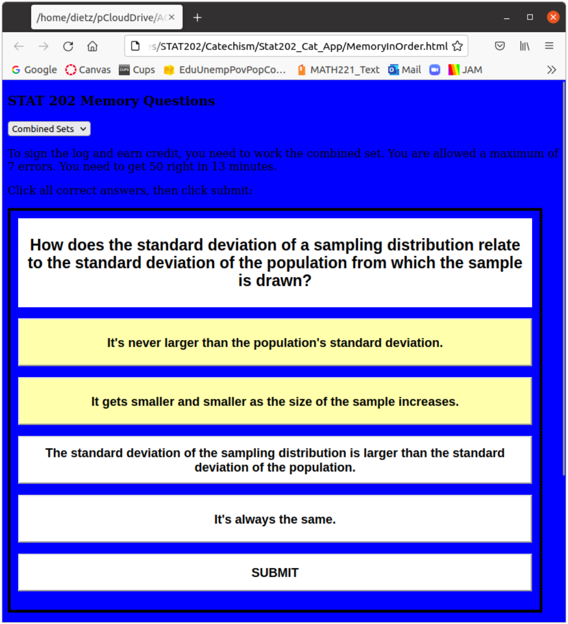

/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
How does the mean of a sampling distribution relate to the mean of a population?
The mean of the sampling distribution is the same as the mean of the population.
As the sample size increases, the mean of the sampling distribution decreases.
The mean of the sampling distribution is smaller than the mean of the population.
The mean of the sampling distribution doesn't represent the mean of the population well.
SUBMIT

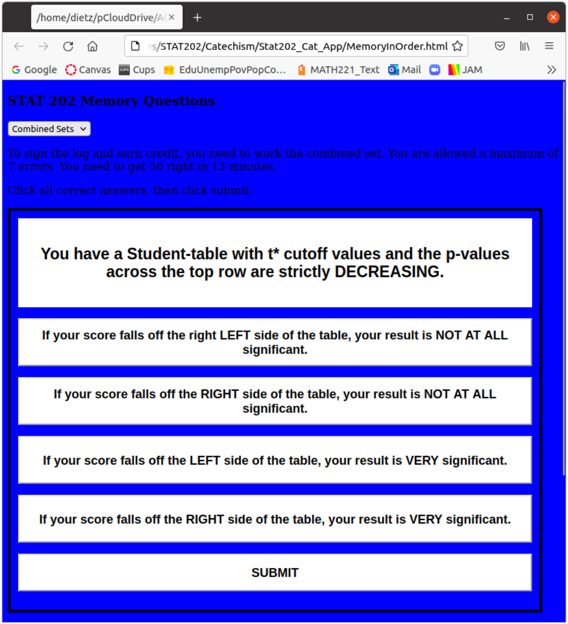

/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
How does the mean of a sampling distribution relate to the mean of a population?
The mean of the sampling distribution is the same as the mean of the population.
As the sample size increases, the mean of the sampling distribution decreases.
The mean of the sampling distribution is smaller than the mean of the population.
The mean of the sampling distribution doesn't represent the mean of the population well.
SUBMIT

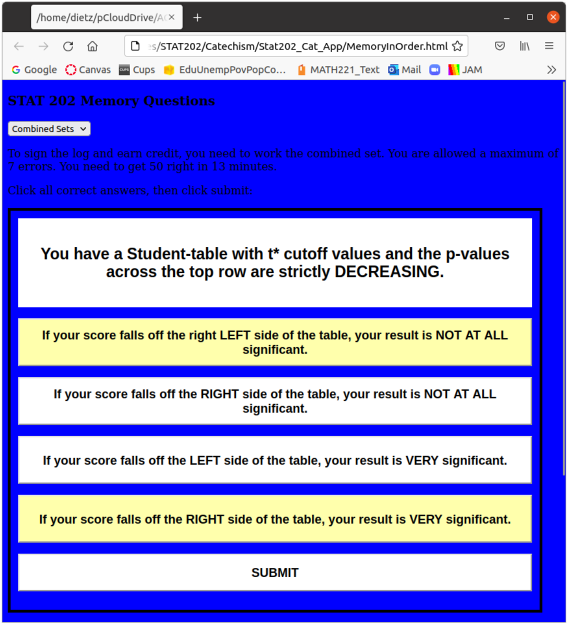

/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
What do you need to know in order to determine what your sampling distribution looks like?
You need to know what the population looks like.
You need to know the size of the samples.
You need to know the p-value.
You need to know the z-score.
SUBMIT

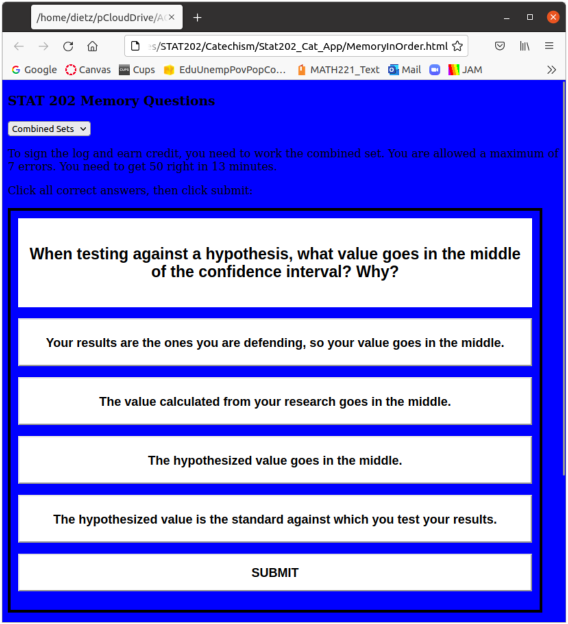

/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
What do you need to know in order to determine what your sampling distribution looks like?
You need to know what the population looks like.
You need to know the size of the samples.
You need to know the p-value.
You need to know the z-score.
SUBMIT

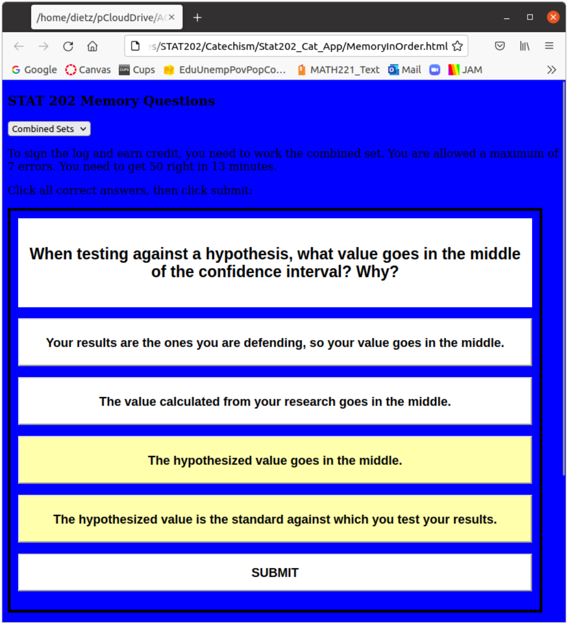

/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google Canvas Cups EduUnempPovPopCo... MATH221_Text Mail JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
How does the standard deviation of a sampling distribution relate to the standard deviation of the population from which the sample is drawn?
It's never larger than the population's standard deviation.
It gets smaller and smaller as the size of the sample increases.
The standard deviation of the sampling distribution is larger than the standard deviation of the population.
It's always the same.
SUBMIT

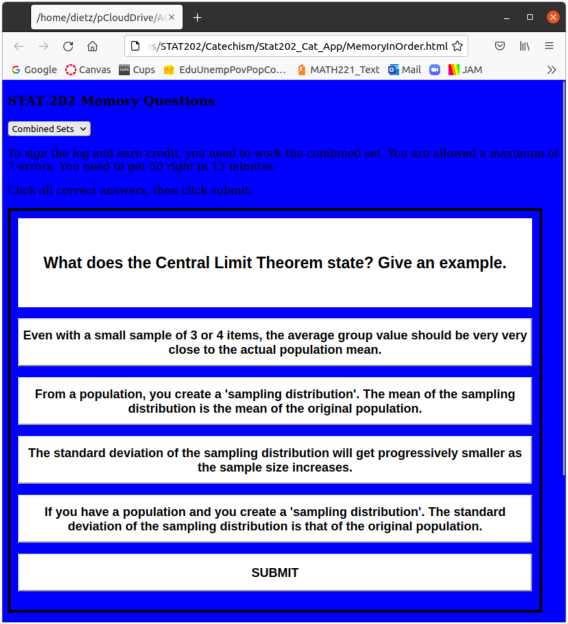

/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
How does the standard deviation of a sampling distribution relate to the standard deviation of the population from which the sample is drawn?
It's never larger than the population's standard deviation.
It gets smaller and smaller as the size of the sample increases.
The standard deviation of the sampling distribution is larger than the standard deviation of the population.
It's always the same.
SUBMIT

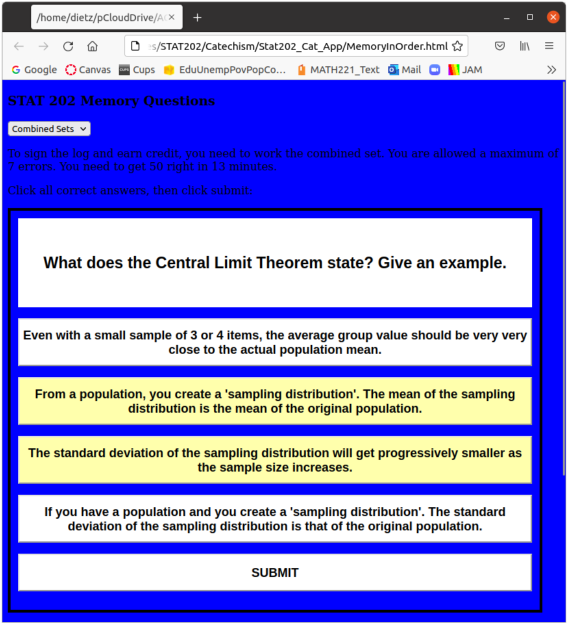

/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
You have a Student-table with t* cutoff values and the p-values across the top row are strictly DECREASING.
If your score falls off the right LEFT side of the table, your result is NOT AT ALL significant.
If your score falls off the RIGHT side of the table, your result is NOT AT ALL significant.
If your score falls off the LEFT side of the table, your result is VERY significant.
If your score falls off the RIGHT side of the table, your result is VERY significant.
SUBMIT


/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
You have a Student-table with t* cutoff values and the p-values across the top row are strictly DECREASING.
If your score falls off the right LEFT side of the table, your result is NOT AT ALL significant.
If your score falls off the RIGHT side of the table, your result is NOT AT ALL significant.
If your score falls off the LEFT side of the table, your result is VERY significant.
If your score falls off the RIGHT side of the table, your result is VERY significant.
SUBMIT


/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google Canvas Cups EduUnempPovPopCo... MATH221_Text Mail JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
When testing against a hypothesis, what value goes in the middle of the confidence interval? Why?
Your results are the ones you are defending, so your value goes in the middle.
The value calculated from your research goes in the middle.
The hypothesized value goes in the middle.
The hypothesized value is the standard against which you test your results.
SUBMIT


/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
When testing against a hypothesis, what value goes in the middle of the confidence interval? Why?
Your results are the ones you are defending, so your value goes in the middle.
The value calculated from your research goes in the middle.
The hypothesized value goes in the middle.
The hypothesized value is the standard against which you test your results.
SUBMIT


/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google Canvas Cups EduUnempPovPopCo... MATH221_Text Mail JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
What does the Central Limit Theorem state? Give an example.
Even with a small sample of 3 or 4 items, the average group value should be very very close to the actual population mean.
From a population, you create a 'sampling distribution'. The mean of the sampling distribution is the mean of the original population.
The standard deviation of the sampling distribution will get progressively smaller as the sample size increases.
If you have a population and you create a 'sampling distribution'. The standard deviation of the sampling distribution is that of the original population.
SUBMIT


/home/dietz/pCloudDrive/A
s/STAT202/Catechism/Stat202_Cat_App/MemoryInOrder.html
Google
Canvas
Cups
EduUnempPovPopCo...
MATH221_Text
Mail
JAM
STAT 202 Memory Questions
Combined Sets
To sign the log and earn credit, you need to work the combined set. You are allowed a maximum of 7 errors. You need to get 50 right in 13 minutes.
Click all correct answers, then click submit:
What does the Central Limit Theorem state? Give an example.
Even with a small sample of 3 or 4 items, the average group value should be very very close to the actual population mean.
From a population, you create a 'sampling distribution'. The mean of the sampling distribution is the mean of the original population.
The standard deviation of the sampling distribution will get progressively smaller as the sample size increases.
If you have a population and you create a 'sampling distribution'. The standard deviation of the sampling distribution is that of the original population.
SUBMIT