Introducing the Standard Errors - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT *Past Due*

Question 1

0.0/1.0 point (graded)

True or False: As the sample size increases, the expected standard error of the sample mean as the estimator of population mean will decrease.

- True ✔
- False

Explanation

From the formula Professor Ellison showed in class, we know that $SE\left(\bar{X}_n\right) = \frac{\sigma}{\sqrt{n}}$. So as your sample size increases, the standard error decreases. Intuitively, the larger your sample size, the more precise your estimates will be.

Show answer

Submit You have used 0 of 1 attempt

Question 2

0.0/1.0 point (graded)

What is the relationship between sample size and confidence intervals? (Select all that apply)

- [ ] the larger your sample size, the narrower your confidence interval. ✔
- [ ] the smaller your sample size, the narrower your confidence interval.
- [ ] the larger your sample size, the wider your confidence interval.
- [ ] the smaller your sample size, the wider your confidence interval. ✔

Explanation

Remember, just like standard errors, confidence intervals are just another way to represent the same information about the tightness of the distribution of the estimator. So the larger your sample size, the smaller your standard errors, and so the tighter the distribution of the estimator, and the narrower your interval. Similarly, the smaller your sample size, the larger your standard errors, and so the more dispersed the distribution of your estimator, and the wider your interval.

Show answer

Submit You have used 0 of 2 attempts

Confidence Intervals - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT *Past Due*

Question 1

0.0/1.0 point (graded)

True or False: The confidence level of your confidence interval depends on your underlying distribution of your data.

- ◯ True
- ◯ False ✔

Explanation

The confidence level of your confidence interval is whatever you want it to be. It represents your desired "degree of confidence". Another way of saying this is if we compute the interval in repeated samples, then α-percent of the time, the interval will bracket the true mean. Note that the interval is random by virtue of random sampling; but the true mean is fixed.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Question 2

0 points possible (ungraded)

True or False: The bounds on the confidence interval you obtain are random variables.

- True ✔
- False

Explanation
Since your sample is random, and your interval is obtained by evaluating some function at your particular realization, the bounds are necessarily random.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

The Chi Squared Distribution - Quiz

Bookmark this page

Question 1

0 points possible (ungraded)

True or False: For i.i.d. samples $X_1, \ldots, X_n$ drawn from a normal distribution, the sample variance also follows a χ^2 distribution (up to a multiplicative constant).

- True ✔
- False

Explanation

This is true. The proof and derivation of that are slightly complicated and outside the scope of this course. However, it is important that you understand that this is **not** the case in general: the sample variance follows a χ^2_{n-1} if and only if $X_1, \ldots, X_n$ are i.i.d samples from a standard normal distribution.

Show answer

Submit You have used 0 of 1 attempt

Suppose we are sampling from a $N(\mu, \sigma^2)$ distribution. Which of the following statements is true? (Select all that apply)

- [ ] $\frac{(n-1)s^2}{\sigma^2} \sim N(0,1)$
- [ ] $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_n$
- [ ] $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$ ✔
- [ ] $\frac{(\bar{X}-\mu)}{\sqrt{\left(\frac{\sigma^2}{n}\right)}} \sim \chi^2_{n-1}$
- [ ] $\frac{(\bar{X}-\mu)}{\sqrt{\left(\frac{\sigma^2}{n}\right)}} \sim \chi^2_n$
- [ ] $\frac{(\bar{X}-\mu)}{\sqrt{\left(\frac{\sigma^2}{n}\right)}} \sim N(0,1)$ ✔
- [ ] $\frac{(\bar{X}-\mu)}{\sqrt{\left(\frac{\sigma^2}{n}\right)}} \sim N(\mu, \sigma^2)$

Question 2

0.0/1.0 point (graded)

Student's distribution says that $\frac{x-\mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$. As the sample size increases (i.e. $n \to \infty$), $s \to \sigma$, i.e. the sample variance converges to the population variance by the law of large numbers. By Slutsky's theorem (which roughly just says that we can "plug" this convergence into the previous one), it follows that:

$$\frac{x-\mu}{\frac{s}{\sqrt{n}}} \to \frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} \text{ as } n \to \infty$$

What follows as a consequence?

- [ ] The student's t-distribution converges to the normal distribution as n increases ✔
- [ ] The normal distribution converges to the student's t-distribution

Explanation

We know that $\frac{x-\mu}{\frac{\sigma}{\sqrt{(n)}}} \sim N(0,1)$ follows a normal distribution. Because $\frac{x-\mu}{\frac{s}{\sqrt{(n)}}} \to \frac{x-\mu}{\frac{\sigma}{\sqrt{(n)}}}$, it follows that $t_{n-1} \to N(0,1)$.

Show answer

Submit You have used 0 of 1 attempt

More on the t-distribution - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT Past Due

Question 1

0.0/1.0 point (graded)

The t-distribution was formulated by a brewer in:

- Elgin
- Edinburgh
- Dublin ✔
- London

Explanation

As Professor Ellison explained in class, the t-distribution was formulated by William Sealy Gosset in his job as Chief Brewer in the Guinness Brewery in Dublin, Ireland. He published it under the pseudonym "student" in 1908.

The F-distribution - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT Past Due

Question 1

0.0/1.0 point (graded)

True or False: Suppose you draw two random samples of size n_1 and n_2 from two different populations. $X_1 \sim N\left(\mu_1, \sigma_1^2\right)$ and $X_2 \sim N\left(\mu_2, \sigma_2^2\right)$. You estimate the sample variances, s_1^2 and s_2^2. The ratio of your estimates of the sample variances, $\frac{s_1^2}{s_2^2}$, follows an F distribution with $n_1 - 1, n_2 - 1$ degrees of freedom.

- True
- False ✔

Explanation

For the statement to be true, it must be the case that σ_1 and σ_2 are the same.

Show answer

Submit You have used 0 of 1 attempt

Which distribution is useful for testing the following hypotheses:

Scenario A. You think the true variance is σ^2, and you want to use your data to make inferences about σ^2.

Select an option **Answer:** chi^2

Scenario B. You think the true mean is 0, and you want to use your data to test whether $\mu = 0$. You don't know the true variance.

Select an option **Answer:** t

Scenario C: You want to compare the variances in two independent population

Select an option **Answer:** F

Explanation

For **Scenario A**, since you know the that $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$, you can construct your test-statistic, and compare that to the χ^2 with $n-1$ degrees of freedom.

For **Scenario B**, as Professor Ellison explained in class, when the population variance is unknown, the standardized sample mean follows a t-distribution with $n-1$ degrees of freedom. So by comparing your test statistic to the t-distribution you can make inferences about the sample mean.

For **Scenario C**, as explained in question 1, we know that the ratio of the sample variances divided by the their degrees of freedom follows an F distribution if the variances are equal. So if you compare the test statistic you obtained with the F distribution, you can test whether the variances are in fact equal.

More on Hypothesis Testing - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT Past Due

Question 1

0.0/1.0 point (graded)

Suppose you are interested in testing $H_0 : \mu < 0$ versus $H_1 : \mu > 0$. This is an example of:

- testing a composite hypothesis vs a composite hypothesis. ✔
- testing a composite hypothesis vs a simple hypothesis.
- testing a simple hypothesis vs a composite hypothesis.
- testing a simple hypothesis vs a simple hypothesis.

Explanation

Recall that a simple hypothesis is one that is characterized by a single point, whereas a composite hypothesis is one that is characterized by multiple points. Since the stated hypotheses are about a range of values, they are both composite hypotheses.

Type I and Type II Errors - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT *Past Due*

Question 1

0.0/1.0 point (graded)

Which of the following statements is correct? (Select all that apply)

- [ ] A type I error is when we falsely reject the null hypothesis. ✔
- [ ] A type II error is when we falsely reject the null hypothesis.
- [ ] A type I error is when we falsely accept the null hypothesis.
- [ ] A type II error is when we falsely accept the null hypothesis. ✔

Explanation

This is just a taxonomy to help analyze and control errors. A type I error is when we falsely reject the null hypothesis and a type II error is when we falsely accept the null hypothesis.

Show answer

For fixed null and alternate hypotheses H_0 and H_A, consider two different hypothesis tests:

Test 1 has significance level 0.05 and operating characteristic 0.05
Test 2 has significance level 0.01 and operating characteristic 0.10

What follows from this? (Select all that apply)

- [ ] As compared to Test 1, Test 2 incurs a lower Type I error but incurs a higher Type II error. ✔
- [ ] As compared to Test 1, Test 2 incurs a higher Type I error but incurs a lower Type II error.
- [ ] As compared to Test 1, Test 2 has a higher confidence level but a lower power. ✔
- [ ] As compared to Test 1, Test 2 has a lower confidence level but a higher power.

Explanation
Recall the following:

(1) The probability of a type I error is equal to the significance level α.
(2) The confidence level is equal to $1-\alpha$.
(3) The probability of a type II error is equal to the operating characteristic β.
(4) Power is equal to $1-\beta$.

From (1) and (2): the probability of a type I error is increasing in the significance level, and decreasing in the confidence level of your test.

Finger Exercises due Oct 19, 2020 19:30 EDT **Past Due**

Question 1

0.0/1.0 point (graded)

Based on the example in the video and assuming that your sample size is fixed, which of the following statements is are true? (Select all that apply)

- [ ] Increasing the critical value k, means your operating characteristic will decrease, and your significance level will increase.
- [ ] Increasing the critical value k, means β will increase, and α will decrease. ✔
- [ ] Decreasing the critical value k means the probability of a type II error increases, while the probability of a type I error decreases.
- [ ] Increasing the critical value k, means the probability of a type II error increases, while the probability of a type I error decreases. ✔

Explanation

Recall the relationships Professor Ellison went through in class. Increasing k means α decreases and β increases. Similarly, decreasing k will cause these quantities to move in the opposite direction.

Show answer

Hypothesis Testing Example, Continued - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT **Past Due**

Question 1

0.0/1.0 point (graded)

True or False: If you reduce your probability of a Type I error, it must be true that the probability of a Type II error goes up.

- True
- False ✔

Explanation

You can reduce your probability of a Type I error by either increasing k, or increasing your sample size n. The trade-off Professor Ellison discussed in class is true assuming n is fixed. However, if you increase n, you can do better on both α and β.

Show answer

Submit You have used 0 of 1 attempt

Power Calculations in Experimental Design - Quiz

Bookmark this page

Finger Exercises due Oct 19, 2020 19:30 EDT *Past Due*

Question 1

1 point possible (graded)

Suppose that you are designing an experiment. At the outset, which of the following are things that you should consider when deciding on a sample size? (Select all that apply.)

- [ ] Decreasing the sample size reduces variation of the sample population, and may allow you to get a more accurate treatment effect
- [ ] Increasing the sample size allows you to estimate the treatment effect with greater precision ✔
- [ ] Increasing the sample size will increase the cost of the experiment ✔

Explanation

One of the key tradeoffs in experimental design is that a greater sample size will result in lower variance and hence allow you to estimate the treatment effect with greater precision. On the other hand, increasing the sample size will usually increase the cost of running the experiment, and you may have to limit the size of your sample to stay within your budget.

Show answer

Question 2

2 points possible (graded)

In hypothesis testing, which one of these definitions matches the correct error?

Not finding a treatment effect when one exists (failing to reject the null when it is in fact false)

- ◯ Type I error
- ◯ Type II error ✔

Finding a treatment effect that actually does not exist (rejecting the null when it is in fact true)

- ◯ Type I error ✔
- ◯ Type II error

Explanation
Type I error refers to the possibility of detecting a treatment effect when there is none, while type II error refers to the possibility of not detecting a treatment effect that does exist. You may sometimes hear Type I error referred to as a false positive, and Type II error referred to as a false negative.

One way to proceed with power calculations is to use $\alpha, \beta, \tau, \sigma$, and γ, and using these inputs, calculate the sample size needed to detect a significant treatment effect. α refers to the significance level of the test, and $1-\beta$ is power. What do τ, σ, and γ refer to?

τ refers to ____________

- ◯ the target average treatment effect
- ◯ the actual measured standard deviation
- ◯ the fraction of the sample in the treatment group
- ◯ the actual measured average treatment effect

σ refers to ____________

- ◯ the desired level of significance of the test
- ◯ the actual measured standard deviation
- ◯ the fraction of the sample in the treatment group

- an estimate of standard deviation of the outcome ✔

γ refers to ______________

- a representative estimate of the variance
- the fraction of the sample in the treatment group ✔
- the desired level of significance of the test
- the fraction of the sample in the control group

Explanation

τ refers to the targeted average treatment effect, or the minimum average treatment effect that, if measured, would be both statistically different from zero and of a meaningful size (for example, if an intervention improves some outcome by 0.01, but is only cost-effective if it improves the outcome by 0.05, you may want to set τ to be 0.05 or more. σ refers to an estimate of the standard deviation of the outcome of interest. Typically researchers will use an estimate from previously-collected data on a relevant or similar population in the power calculations leading up to the design of their experiments. γ refers to the fraction of the population allocated to the treatment. Power calculations are tricky in practice because they necessarily use ingredients that you do not know before conducting your experiment.

Question 2

0.0/1.0 point (graded)

A power calculation is a function of a treatment effect τ, which is supplied by the researcher. What are some sensible ways to decide on a τ? (Select all that apply)

- [ ] A τ that would be interesting and actionable to policy-makers ✔
- [ ] A τ that reflects previously known estimates in the literature, if the particular research was for the purpose of replication/searching for generalizability (external validity) ✔
- [ ] A τ that is extremely large, to be conservative.

Explanation

The first two choices are sensible ways to decide on a τ. For the last choice, if we use a very large number for τ, we will only require a small sample, but then we will fail to reject the null even if the true treatment effect was positive, but smaller than what we targeted.

Show answer

Submit You have used 0 of 2 attempts

Calculating the test statistic for determining power and sample size relies on the assumption that ________________.

- ◯ The assumption that the standardized difference in means is normally-distributed ✔
- ◯ The critical value is smaller than 1.96
- ◯ The significance level is 0.05
- ◯ The assumption that the standardized difference in means is uniformly-distributed

Explanation
As described in the derivation of the formula for determining sample size given a desired level of power, we use the central limit theorem to assume that the difference in treatment and control means, scaled by the standard error of that difference is normally-distributed with N(0,1). This is depicted in mathematical terms as follows:

$$\frac{\overline{Y_t^{obs}} - \overline{Y_c^{obs}}}{\sqrt{\frac{\sigma^2}{N_t} + \frac{\sigma^2}{N_c}}} \sim N(0,1)$$