Finger Exercises due Oct 12, 2020 19:30 EDT Past Due

Question 1

0.0/1.0 point (graded)

Which of the following is true about a sample mean? (Select all that apply)

- [ ] It can be described as the arithmetic average of n random variables from a random sample of size n. ✔
- [ ] It can be described as the arithmetic average of the realizations of n random variables. ✔
- [ ] It only applies to random variables from normal distributions
- [ ] It only applies to random variables from uniform distributions

Explanation

The sample mean is the arithmetic average of the random variables, but also describes the arithmetic average of the realization of those random variables. As Professor Ellison mentioned in the lecture, we have to keep both of these descriptions in mind. The underlying distribution of the random variables does not need to be normal or uniform in order to calculate the sample mean.

Show answer

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

When the sample mean is defined as the arithmetic average of n random variables from random sample of size n, the sample mean will also be a random variable.

- True ✔
- False

Explanation

A function of random variables must be a random variable. Being the arithmetic average of n random variables makes the sample mean a function of random variables, so the sample mean must be a random variable.

Show answer

Submit You have used 0 of 1 attempt

More on the Sample Mean - Quiz

Bookmark this page

Finger Exercises due Oct 12, 2020 19:30 EDT Past Due

Question 1

0.0/1.0 point (graded)

For an i.i.d. distribution, how does the expectation of sample mean of n random variables drawn from this distribution vary with n?

- No effect on the expectation of the sample mean ✔
- Expectation decreases wtih n
- Expectation increases with n

Explanation

The expected value of a sample mean is $\mu\left(E\left(X_i\right)\right)$. $E\left(Xi\right)$ will be determined by the shape of the distribution of the X's, which will not change as the sample size increases.

Show answer

For a sample of size n from an i.i.d distribution with variance σ^2, which of the following expressions is the variance of the sample mean?

- ( ) $\frac{\sigma^2}{n}$ ✔
- ( ) $\frac{\sigma}{n}$
- ( ) σ^2
- ( ) $\frac{\sigma^2}{n^2}$

Explanation

$\frac{\sigma^2}{n}$ is the correct answer. Professor Ellison derives this result in the lecture slides. Here is a quick summary of that derivation:

By the definition of the sample mean, $Var\left(\bar{X}\right) = Var\left(\frac{1}{n}\left(\sum X_i\right)\right)$

By properties of variance, $Var\left(\frac{1}{n}\left(\sum X_i\right)\right) = \frac{1}{n^2}\sum Var\left(X_i\right)$

Finally, it simplifies to $\frac{1}{n^2}\sum \sigma^2 = \frac{\sigma^2}{n}$

Finger Exercises due Oct 12, 2020 19:30 EDT **Past Due**

Question 1

0.0/1.0 point (graded)

The Central Limit Theorem (CLT) implies that were one to draw n samples $X_1, \dots X_n$ independently and identically, then for reasonably large n... (Select all that apply)

- [ ] Each X_i need not be approximately normally distributed, but the sample mean $\bar{X} = \sum_i X_i / n$ will be approximately normally distributed. ✔
- [ ] Each X_i will be approximately normally distributed, but the sample mean $\bar{X} = \sum_i X_i / n$ need not be approximately normally distributed.
- [ ] Both X_i and the sample mean $\bar{X} = \sum_i X_i / n$ will be approximately normally distributed
- [ ] Neither the X_i's nor the sample mean $\bar{X} = \sum_i X_i / n$ will be approximately normally distributed

Explanation

The CLT implies that the distribution of the sample mean for an i.i.d. random sample will be approximately normal if the sample size is large enough. The distribution of the population will be determined by the specific characteristics of that population and generally, will not be affected by changes in the sample size.

Show answer

Question 2

0.0/1.0 point (graded)

________ is the process of subtracting the mean of a distribution and dividing by the square root of its variance.

- ○ Standardization ✔
- ○ Realization
- ○ Randomization
- ○ Estimation

Explanation

As mentioned by Professor Ellison, standardization is the process of subtracting the mean of a distribution and dividing by the square root of its variance, which creates a random variable with 0 mean and variance 1. The other options are unrelated statistical concepts.

Show answer

Submit You have used 0 of 2 attempts

The central limit theorem says that:

$$\lim_{n\to\infty} P\left[\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \leq x\right] = \Phi\left(x\right)$$

$P\left[\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma} \leq x\right]$ represents:

- ◯ The cumulative distribution function (CDF) of the population
- ◯ The cumulative distribution function (CDF) of the standardized sample mean ✔
- ◯ The probability distribution function (PDF) of the population
- ◯ The probability distribution function (PDF) of the standardized sample mean

Explanation

We know that it is a CDF because the definition of a CDF is the probability that a random variable is less than or equal to some value of x. We know that it is the CDF of the standardized sample mean because $\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma}$ is the sample mean $\bar{x}$ after the process of subtracting the mean of the distribution of $\bar{x}$ (which is μ) and dividing by the square root of the variance (the square root of the variance is $\frac{\sigma}{\sqrt{n}}$)

Estimation - Quiz

Bookmark this page

Finger Exercises due Oct 12, 2020 19:30 EDT *Past Due*

Question 1

0.0/1.0 point (graded)

True or False: You can uniquely identify a given distribution if you know the family of distributions it is from (ex. Normal, uniform etc.) and the value of the relevant parameters for that family.

- True ✔
- False

Explanation

A parameter is a constant indexing a family of distributions. Indexing a family of distributions means that the parameters allow you to distinguish between the distributions in the given family. Thus, giving you the family restricts the distribution to that set and then the parameter allows you to uniquely identify a distribution in that set.

Show answer

Question 2

0.0/1.0 point (graded)

Which of the following are the typical notations used for parameters of the normal distribution?

- a and b
- μ and σ^2 ✔
- n and p
- λ and $\frac{1}{\lambda}$

Explanation

A normal distribution can be defined by its mean (μ) and its variance (σ^2). n and p are the parameters for the binomial distributions. a and b are the parameters for the uniform distributions. For λ and $\frac{1}{\lambda}$, λ alone is the parameter for the exponential distributions.

Show answer

Submit You have used 0 of 2 attempts

Finger Exercises due Oct 12, 2020 19:30 EDT **Past Due**

Question 1

0.0/1.0 point (graded)

For all families of distributions, estimation is trying to determine the specific ______ of a distribution.

- ◯ Maximum
- ◯ Variance
- ◯ Parameter ✔
- ◯ Mean

Explanation

Estimation is trying to determine the specific parameter of a distribution, because this will give us a lot information about the shape of the distribution. For example, for normal distributions, estimation will typically try to determine the mean and variance.

Show answer

Submit You have used 0 of 2 attempts

Question 2

0.0/1.0 point (graded)

True or False: Estimators are the realizations of applying estimates to random samples.

- True
- False ✔

Explanation

The function of a random sample is the estimator. The realization of the function of the random sample is the estimate, so the estimates are the realizations of applying the estimators to random samples.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Question 3

Question 3

0.0/1.0 point (graded)

Per the notational convention, what is the usual relationship between θ and $\hat{\theta}$?

- ◯ θ usually refers to the mean, while $\hat{\theta}$ refers to the variance.
- ◯ θ usually refers to the mean, while $\hat{\theta}$ refers to the sample mean.
- ◯ $\hat{\theta}$ usually refers to a parameter relevant to the underlying distribution, while θ refers to its estimation in a finite sample
- ◯ θ usually refers to a parameter relevant to the underlying distribution, while $\hat{\theta}$ refers to its estimation in a finite sample ✔

Explanation

The estimator and the estimate could both be represented by $\hat{\theta}$ and the parameter they are estimating could be represented by θ. The estimator and estimate are $\hat{\theta}$, because they are estimating the parameter.

Show answer

Submit You have used 0 of 2 attempts

Suppose $X \sim U\left[0, \theta\right]$

$$f_x\left(x\right) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta \\ 0, & \text{otherwise} \end{cases}$$

What are reasonable ways to estimate θ? (Select all that apply)

- [ ] Compute the minimum of the sample
- [ ] Compute the maximum (n^{th} order statistic) of the sample ✔
- [ ] Compute the sample mean and divide by 2
- [ ] Compute the sample mean and multiply by 2 ✔

Explanation

The maximum will be a reasonable estimate, because θ is the maximum value that the distribution can take on. Since the expectation of the sample mean in this distribution is $\frac{\theta}{2}$, multiplying the sample mean by 2 will yield a reasonable estimate θ. The minimum of the sample will be close to zero. Dividing the sample mean by two will give us an expected value of $\frac{\theta}{4}$, which would not be a reasonable estimate of θ.

Suppose $X \sim U[0, \theta]$

$$f_x(x) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta \\ 0, & \text{otherwise} \end{cases}$$

Bob wants to estimate θ by drawing a sample of size 10 i.i.d and computing the standard mean, i.e. estimating via $\hat{\theta} = \bar{X}_n$. How do you assess Bob's estimation strategy?

- The strategy is flawed primarily because the estimation is biased. Although $\hat{\theta}$ approaches θ, it is likely to underestimate θ on average
- The strategy is flawed primarily because the estimation is underpowered. That is, a sample of size 10 is too small. However, were Bob to using a very large sample, the realizations of $\hat{\theta}$ would be close to θ.
- The strategy is flawed primarily because this estimation is inconsistent. That is, regardless of how much Bob increases his sample size n, the realizations of $\hat{\theta}$ will never approach θ. ✔
- The estimation strategy is correct!

Explanation

Bob should compute *twice* the sample mean in his estimation. His current strategy of estimating $\hat{\theta} = \bar{X}$ using the sample

Criteria for Assessing Estimation - Quiz

Bookmark this page

Finger Exercises due Oct 12, 2020 19:30 EDT Past Due

Question 1

0.0/1.0 point (graded)

An estimator is unbiased if:

- ○ $\hat{\theta}$ has a uniform distribution
- ○ $E\left(\hat{\theta}\right) = \theta$ for all θ ✔
- ○ $E\left(\hat{\theta}\right) = \mu$ for all θ
- ○ $\hat{\theta}$ has a normal distribution

Explanation

As shown in the lecture slides, the expected value of the estimates from an unbiased estimator must be equal to the parameter of interest. The shape of the distribution of these estimates is irrelevant and having the expected value of the estimates be equal to μ will make the estimator biased when estimating a parameter other than μ.

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

True or False: Suppose you are able to generate random numbers from a normal distribution $\mathcal{N}\left(\mu, \sigma^2\right)$ of unknown mean μ. If $\hat{\mu}$ is the random variable whose realizations are the individual numbers generated, $\hat{\mu}$ is an unbiased estimator for μ.

- ◯ True ✔
- ◯ False

Explanation

The expected value of randomly drawing a number from a normal distribution would be the mean of that distribution, so $E\left(\hat{\theta}\right) = \theta$.

Show answer

Submit You have used 0 of 1 attempt

More on Criteria for Assessing Estimators - Quiz

Bookmark this page

Finger Exercises due Oct 12, 2020 19:30 EDT Past Due

Question 1

0.0/1.0 point (graded)

True or False: To prove an estimator is unbiased, you need to know the value of the parameter it is trying to estimate.

- True
- False ✔

Explanation

You can prove an estimator to be unbiased through a mathematical proof, without knowing the underlying value of the parameter. In Segment 4 of this lecture, Professor Ellison mathematically proved that the sample mean is an unbiased estimate of the mean.

Show answer

Submit You have used 0 of 1 attempt

For the example distribution in lecture, which of the following methods would result in an unbiased estimate of θ? (Select all that apply.)

- [ ] Compute the median of the sample and multiply by 2 ✔
- [ ] Compute the maximum (nth order statistic) of the sample
- [ ] Have R generate a random value from the underlying distribution.
- [ ] Compute the sample mean and multiply by 2 ✔

Explanation

If we consider repeated sampling of n data points from $U\left[0,\theta\right]$ and construct the statistics of 2*(sample mean) or 2*(sample median), they will trace a histogram that is in the limit (of infinite samples of n points each) symmetric and centered around θ. Thus is the intuition for why they are unbiased: more formally expectation of 2*(sample mean) or 2*(sample median) will be $theta$.

Computing the maximum of the n^{th} order statistic is "consistent", i.e. in the limit of increasing n gets closer and closer to $theta$. However, it is not unbiased in that fixed samples of n points will always give an n-order statistic less than $theta$, and thus the expectation of the n-order statistic over all samples of $npoints$ will be less than $theta$. More intuitively, perhaps, the n-order statistic can always underestimate $theta$ but has no chance of overestimating $theta$: this is the sense in which its biased downwards.

Having R generate a single random value of the underlying distribution will only regenerate that distribution in repeated sample. The expected value of this statistic would thus be the expected value of $U\left[0,\theta\right]$, i.e. $\frac{\theta}{2}$. (We could, however, multiply the random value by 2 to get an unbiased estimator of $theta$. However, this has convergence properties that are far worse than first taking a sample mean or median of several points and multiplying by 2).