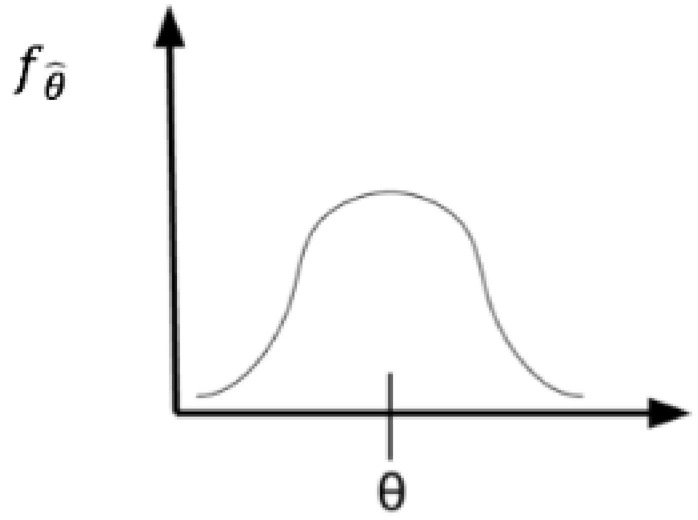
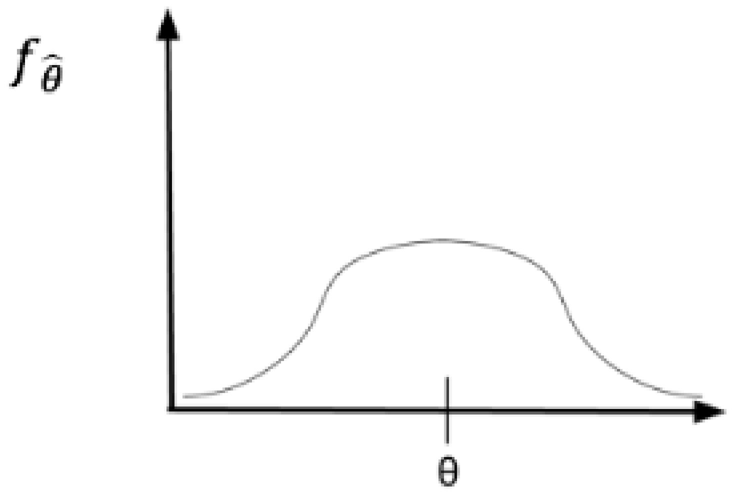
Question 1

0.0/1.0 point (graded)

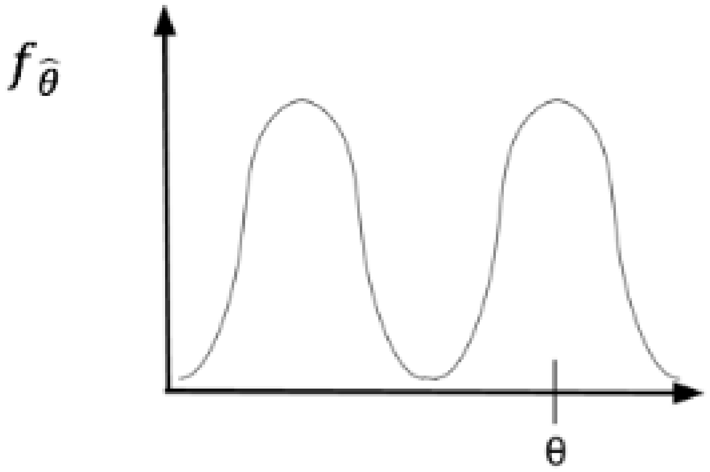
Which estimators $\hat{\theta}$ are unbiased? These graphs show PDFs of $\hat{\theta}$. (Select all that apply.)

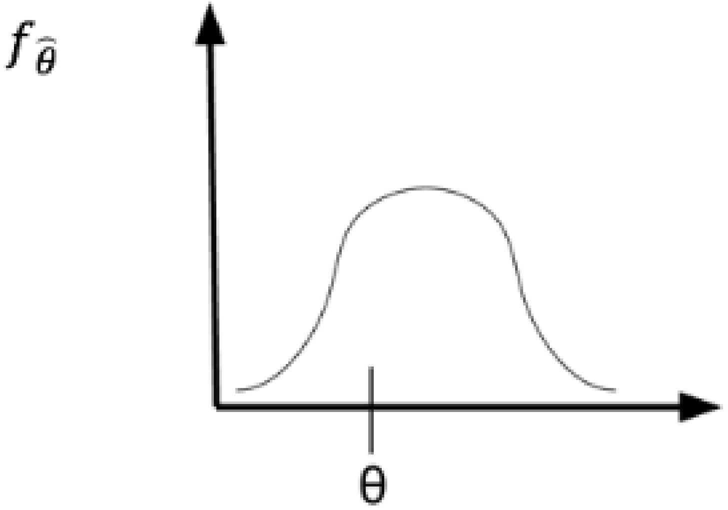
- [ ] 

✔

- [ ] 

✔





Explanation

Recall the definition of an unbiased estimator: an estimator $\hat{\theta}$ is unbiased if, in expectation, it is equal to the parameter it is trying to estimate. (In other words, an estimator is unbiased for θ if $E[\hat{\theta}] = \theta$ for all θ in Φ.) Therefore, the estimators where the expectation of $\hat{\theta}$ is the center of the distribution are correct.

Question 2

0.0/1.0 point (graded)

Suppose we are trying to estimate the mean μ of a $N(\mu, \sigma^2)$ distribution. Which of the following estimators would be unbiased? (Select all that apply)

- [ ] $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$ ✔
- [ ] $\hat{\mu} = \frac{1}{2}\max\{X_1, X_2, \ldots X_n\}$
- [ ] $\hat{\mu} = n\min\{X_1, X_2, \ldots, X_n\}$
- [ ] $\hat{\mu} = \frac{1}{n-1} \sum (X_i - \bar{X}_n)^2$

Explanation

We can calculate the expectation of any of these estimators to determine whether they are biased. The estimator in (a) is unbiased: $E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{n\mu}{n} = \mu$. $\hat{\mu} = \frac{1}{2}\max\{X_1, X_2, \ldots X_n\}$ is biased because, the n^{th} order statistic follows a distribution not centered at μ. Intuitively, since the normal distribution takes on values on the entire real line, the estimator will tend to be larger than μ. Similarly the estimator in $\hat{\mu} = n\min\{X_1, X_2, \ldots, X_n\}$ can be shown to be a biased estimator for μ. $\hat{\mu} = \frac{1}{n-1}\sum (X_i - \bar{X}_n)^2$ would estimate σ^2, not μ.

Show answer

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

Question 1

0.0/1.0 point (graded)

Which of the following estimators could be unbiased for an i.i.d. sample? (Select all that apply.)

- [ ] $\frac{n}{2}\sum_{i=1}^{n} X_i$ for the population mean
- [ ] $\frac{1}{2}\max\{X_1, X_2, \ldots X_n\}$ for the population mean
- [ ] $\frac{1}{n}\sum\left(X_i - \bar{X}_n\right)^2$ for the population variance
- [ ] $\frac{1}{n-1}\sum\left(X_i - \bar{X}_n\right)^2$ for the population variance ✔

Explanation

In this lecture segment, we learnt that the sample mean and sample variance are unbiased estimators of the population mean and population variance, respectively, for an i.i.d. sample. Neither $\frac{n}{2}\sum_{i=1}^{n} X_i$ nor $\frac{1}{2}\max\{X_1, X_2, \ldots X_n\}$ are the correct formulas for sample mean. $\frac{1}{n-1}\sum\left(X_i - \bar{X}_n\right)^2$ is an unbiased estimator of population variance. Note that both $\frac{1}{n}\sum\left(X_i - \bar{X}_n\right)^2$ and $\frac{1}{n-1}\sum\left(X_i - \bar{X}_n\right)^2$ are consistent estimators of the population variance, meaning that with a large enough sample, both of these approach the true population variance asymptotically. But in a finite sample, $\frac{1}{n}\sum\left(X_i - \bar{X}_n\right)^2$ is biased estimator (in particular biased downwards) on average. For an explanation for why this is the case, see here.

Efficient Estimators - Quiz

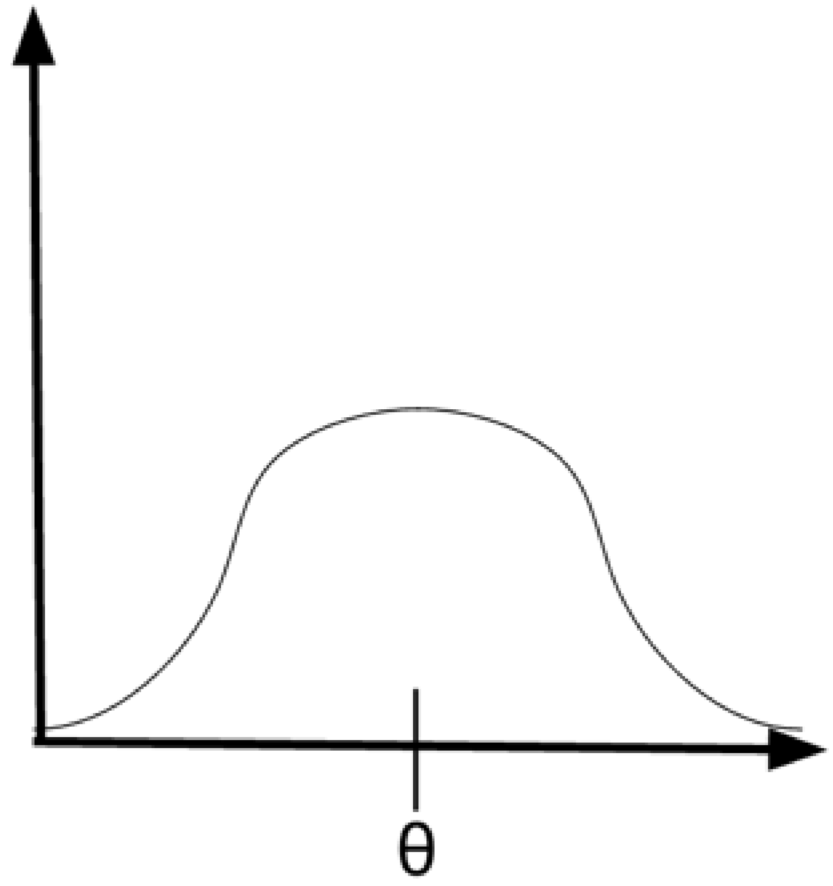
Bookmark this page

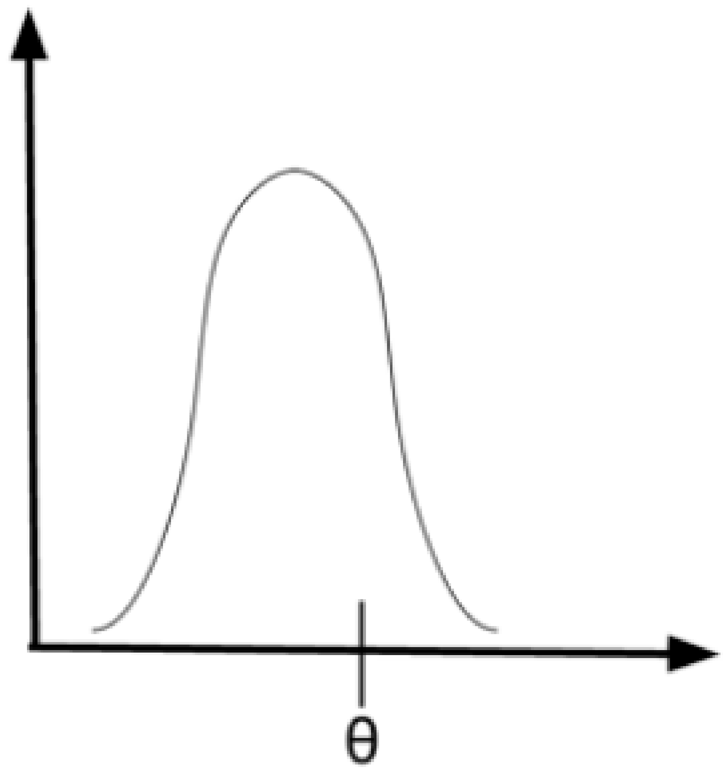
Finger Exercises due Oct 19, 2020 19:30 EDT Past Due

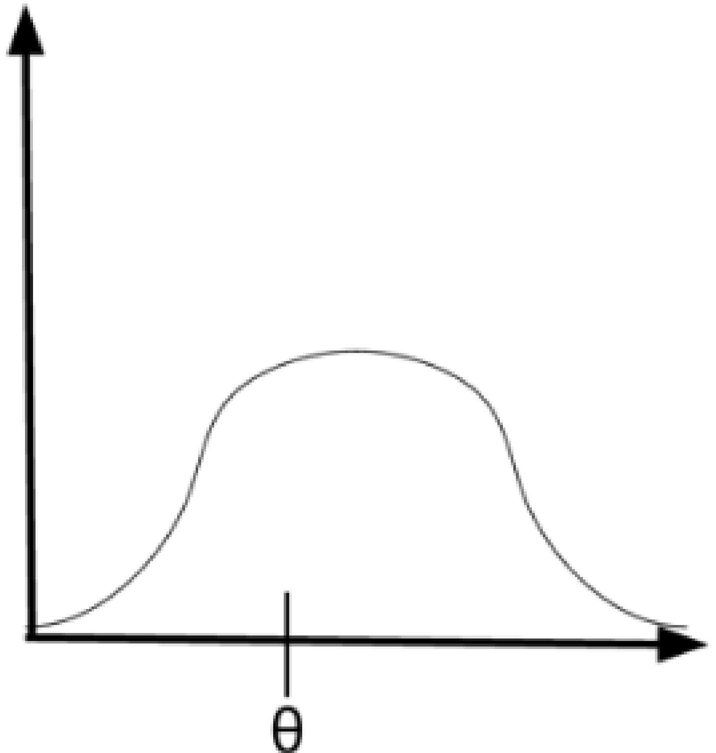
Question 1

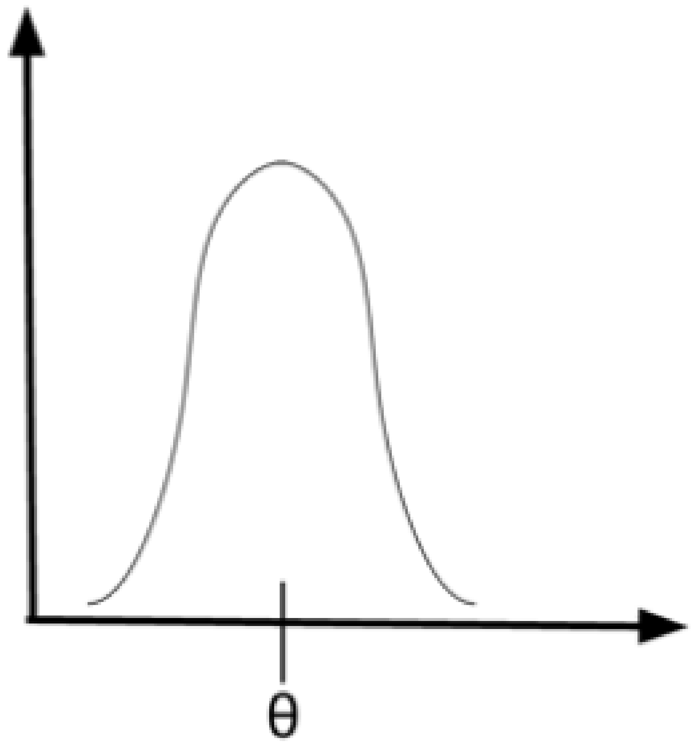
0.0/1.0 point (graded)

Assuming that the scale of these axes is consistent, select the unbiased estimator $\hat{\theta}$ that is **most efficient**. Recall that we have only defined efficiency for unbiased estimators. These graphs show PDFs of $\hat{\theta}$.




θ


θ



Explanation

The estimators that are unbiased are the graphs where θ is at the center of the distribution. Of the two, the one that is more tightly distributed is more efficient.

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

How might we choose an estimator that gives us the best trade off between bias and efficiency?

- Minimize the median squared error.
- Minimize the mean squared error. ✔
- Maximize the median squared error.
- Maximize the mean squared error.

Explanation

The mean squared error $MSE[\hat{\theta}] = \mathrm{Var}(\hat{\theta}) + (E[\hat{\theta}] - \theta)^2$ can be thought of as the sum of the estimator's variance and the square of the estimator's bias. In order to pick an estimator that has a low variance and bias, we want to minimize this sum.

Show answer

Submit You have used 0 of 2 attempts

Question 3

0.0/1.0 point (graded)

Which of the following is true about consistent estimators? (Select all that apply.)

- [ ] They are always unbiased.
- [ ] $\lim_{n\to\infty} P\left(|\theta - \hat{\theta}_n| < \delta\right) = 0$
- [ ] $\lim_{n\to\infty} P\left(|\theta - \hat{\theta}_n| < \delta\right) = 1$ ✔
- [ ] The distribution of the estimator collapses to a single point as n goes to infinity. ✔

Explanation

The definition of a consistent estimator is given by $\lim_{n\to\infty} P\left(|\theta - \hat{\theta}_n| < \delta\right) = 1$. As n goes to infinity, the distribution becomes more and more concentrated, collapsing to a single point. $\lim_{n\to\infty} P\left(|\theta - \hat{\theta}_n| < \delta\right) = 1$ can be true even if the estimator is biased.

Show answer

Submit You have used 0 of 2 attempts

Question 4

0.0/1.0 point (graded)

Which of the following is true about estimating the parameter θ of a $U\left[0,\theta\right]$ distribution?

- Estimating θ by doubling the sample mean results in a unbiased estimator that is more efficient than estimating θ using the n^{th} order statistic.
- Estimating θ by doubling the sample mean results in an unbiased estimator, but this method is less efficient than estimating θ using the n^{th} order statistic. ✔
- Estimating θ by doubling the sample mean results in a biased estimator, but we might still use this method because it is more efficient than estimating θ using the n^{th} order statistic.
- Estimating θ by doubling the sample mean is both more biased and more efficient than estimating θ using the n^{th} order statistic.

Explanation

Doubling the sample mean is an unbiased estimator of θ in a $U\left[0,\theta\right]$ distribution, eliminating two of the answer options. In this lecture segment, we learned that the n^{th} order statistic is more tightly distributed (i.e. more efficient) than the sample mean as an estimator, giving us the correct answer.

Robust Estimators - 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 criteria might one take into account when choosing an estimator? (Select all that apply.)

- [ ] Ease of computing the estimator ✔
- [ ] Robustness ✔
- [ ] Bias ✔
- [ ] Efficiency ✔

Explanation

All of these criteria are perfectly valid to take into consideration when choosing an estimator. Some estimators are far easier to compute than others, which will matter more or less depending on the application. We learnt about robustness, whether the estimator will still do a good job even if our assumptions about the underlying distribution are wrong, in this lecture segment. We learnt about bias and efficiency in the previous lecture segments.

Question 2

0.0/1.0 point (graded)

Which of the following are examples of a given estimator $\hat{\theta}_1$ being more robust than some other estimator $\hat{\theta}_2$? (Select all that apply.)

- [ ] $\hat{\theta}_1$ is more biased than $\hat{\theta}_2$ if we've misspecified the tail probabilities of the underlying distribution.
- [ ] $\hat{\theta}_1$ is less biased than $\hat{\theta}_2$ if we've misspecified the tail probabilities of the underlying distribution. ✔
- [ ] $\hat{\theta}_1$ is more biased than $\hat{\theta}_2$ if our assumed underlying distribution is shifted from the true distribution.
- [ ] $\hat{\theta}_1$ is less biased than $\hat{\theta}_2$ if our assumed underlying distribution is shifted from the true distribution. ✔

Explanation

An estimator is robust if it does a good job of estimator the parameter even if we've made a mistake in our assumptions about the underlying distribution. It is possible that an estimator is robust to one type of mistaken assumption but not another.

Show answer

Submit   You have used 0 of 2 attempts

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

Question 1

0.0/1.0 point (graded)

In this lecture segment, we hear 3 ways to derive estimators: 1) the method of moments, 2) maximum likelihood estimation, and 3) dreaming them up. Which of the following estimators for θ from a $U\,[0,\theta]$ distribution is derived using the method of moments?

- $(N-1)^{th}$ order statistic
- 2 times the sample mean ✔
- N^{th} order statistic
- Random sample

Explanation

As shown in the lecture segment, estimating θ using the sample mean is a method derived using the method of moments. We equate the first population moment with the first sample moment and solve for the parameter. On the other hand, the n^{th} order statistic is an estimator derived using maximum likelihood estimation.

Show answer

Question 2

0.0/1.0 point (graded)

What is the first population moment of a $U\left[0, \theta\right]$ distribution?

- $\frac{2}{n}\sum_{i=1}^{n} X_i$
- $\frac{\theta}{2}$ ✔
- $\frac{1}{n}\sum_{i=1}^{n} X_i$
- θ

Explanation

The first population moment is $E\left[X\right]$, which is $\frac{\theta}{2}$ for a $U\left[0, \theta\right]$ distribution. $\frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean, which is the first sample moment.

Show answer

Submit You have used 0 of 2 attempts

Question 3

0.0/1.0 point (graded)

Which of the following is the second sample moment?

- ○ $\frac{1}{2}(\sum_{i=1}^{n} X_i)^2$
- ○ $\frac{1}{n^2}\sum_{i=1}^{n} X_i$
- ○ $\frac{1}{n}\sum_{i=1}^{n} {X_i}^2$ ✔
- ○ $\frac{2}{n}\sum_{i=1}^{n} X_i$

Explanation

The sample moments are defined by $\frac{1}{n}\sum_{i=1}^{n} X_i,\ \frac{1}{n}\sum_{i=1}^{n} X_i^2,\ \frac{1}{n}\sum_{i=1}^{n} X_i^3,\ \ldots$ The population moments, on the other hand, are defined by expectations and can be expressed as functions of the parameters $E[X], E[X^2], E[X^3], \ldots$.

Show answer

Submit You have used 0 of 2 attempts

Question 4

0.0/1.0 point (graded)

True or False: The method of moments can be used if we wish to estimate more than one parameter.

- True ✔
- False

Explanation

Yes, we can use the method of moments even if we have more than one parameter to estimate. We just use as many sample and population moments as necessary. If we have k parameters to estimate, then we equate the 1^{st} through k^{th} population moments with the 1^{st} through k^{th} sample moments. Then we have k equations and k unknown parameters to solve for.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Maximum Likelihood Estimation - Quiz

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

Question 1

0.0/1.0 point (graded)

Which of the following is true about the maximum likelihood estimator?

- To find the maximum likelihood estimator, we equate population and sample moments.
- The maximum likelihood estimator is always unbiased.
- The maximum likelihood estimator is the value of the parameter which corresponds to the distribution that most likely produced the observed data. ✔

Explanation
The maximum likelihood estimator is the value of the parameter associated with the member of the family of distributions we are examining that “best fits” the observed data. We learned previously that the nth order statistic, which is biased, is a maximum likelihood estimator. Equating population and sample moments is how we find estimators using method of moments.

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

We find the maximum likelihood estimator by:

- ◯ Maximizing the likelihood function, $L(x|\theta)$, over the parameter x.
- ◯ Maximizing the likelihood function, $L(x|\theta)$, over the parameter θ.
- ◯ Maximizing the likelihood function, $L(\theta|x)$, over the parameter x.
- ◯ Maximizing the likelihood function, $L(\theta|x)$, over the parameter θ. ✔

Explanation

The likelihood function tells us the likelihood that the parameter of the underlying distribution is θ given our observations x -- that is, it is function of our parameter θ conditional on x (hence the $\theta|x$).

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

Question 1

0.0/1.0 point (graded)

For an i.i.d. random variable, the likelihood function is simply equal to... (Select all that apply)

- [ ] The joint PDF of the data ✔
- [ ] $\Pi_i f(\theta|x)$
- [ ] $\Pi_i f(x|\theta)$ ✔
- [ ] $f(x|\theta)$
- [ ] $f(\theta|x)$

Explanation

The likelihood function is a reinterpretation of the joint PDF of our data or random sample. The likelihood function is the same as the joint PDF of the data, which for i.i.d. random variables is equal to $\Pi_i f(x|\theta)$

Question 2

0.0/1.0 point (graded)

Sometimes the likelihood function is computationally difficult to maximize. In this case, what could we maximize instead? Select all that apply.

- [ ] The log of the likelihood function. ✔
- [ ] The inverse of the likelihood function.
- [ ] The sine of the likelihood function.
- [ ] Any monotone (increasing) transformation of the likelihood function. ✔

Explanation

Maximizing any monotonic transformation of the likelihood function will give us the same parameters as maximizing the original likelihood function. Taking the log is a monotonic transformation, so the log of the likelihood function is correct. In practice, we often maximize the log-likelihood rather than the likelihood because it is often computationally easier to maximize the log-likelihood.

Show answer

Examples of Maximum Likelihood Estimation, Part II - Quiz

Bookmark this page

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

Question 1

0.0/1.0 point (graded)

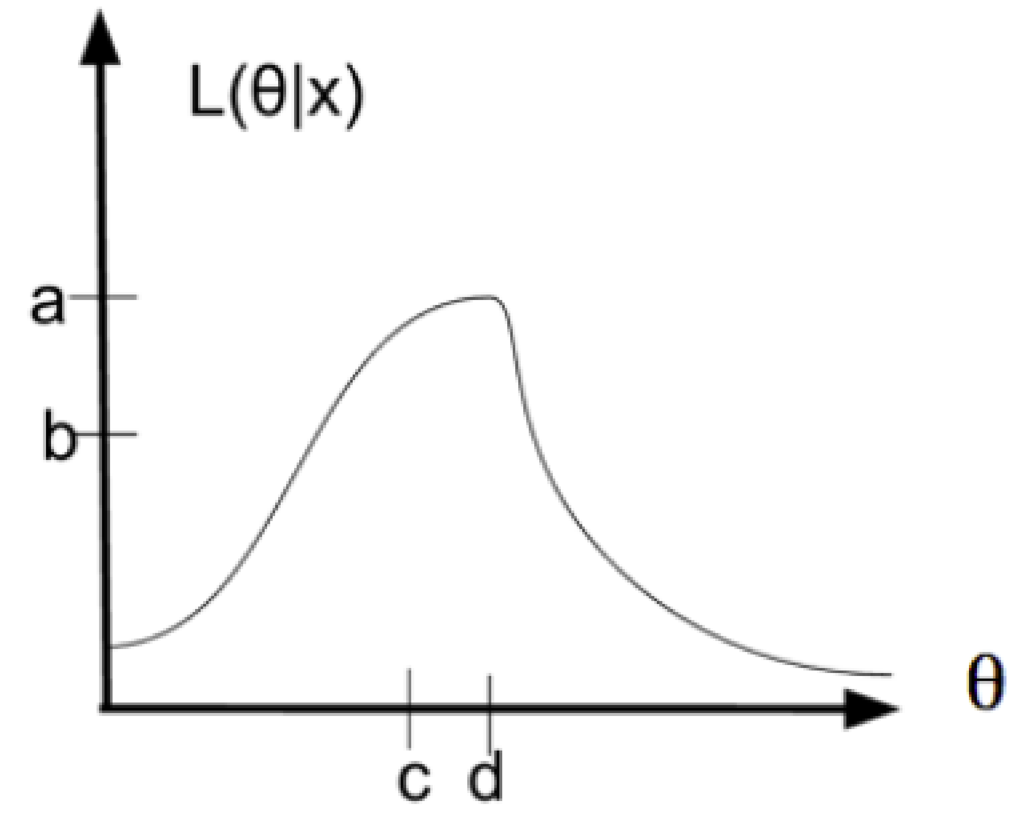
The example in this segment looks at estimating θ in a (uniform) $U\left[0,\theta\right]$ distribution. We see that the n^{th} order statistic is: (Select all that apply)

- [ ] A lower bound for θ ✔
- [ ] An upper bound for θ
- [ ] The maximum likelihood estimator. ✔
- [ ] The maximum value in our observation sample. ✔

Explanation

In this lecture segment, we see that maximizing the likelihood function results in choosing the n^{th} order statistic for our estimator. By definition, the n^{th} order statistic is the maximum value we observe. We know that must be greater than or equal to the nth order statistic because observing a value greater than θ is a zero probability event in a $U\left[0,\theta\right]$ distribution.

According to the following graph, which of the following is the maximum likelihood estimator?



- a
- b
- c
- d

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

Question 1

0.0/1.0 point (graded)

In this example we see a uniform distribution $U\left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right]$. Which of the following is true about the n^{th} order statistic? (Select all that apply)

- [ ] θ must be at least $\frac{1}{2}$ below the n^{th} order statistic.
- [ ] θ can be no more than $\frac{1}{2}$ below the n^{th} order statistic. ✔
- [ ] The n^{th} order statistic can be no more than $\frac{1}{2}$ away from the first order statistic.
- [ ] The n^{th} order statistic can be no more than 1 away from the first order statistic. ✔

Explanation

In this example, the distribution has known length equal to 1. Therefore the 1^{st} and n^{th} order statistics cannot be separated by a length of more than 1. Also, the n^{th} order statistic cannot be more than $\frac{1}{2}$ above θ because $\theta + \frac{1}{2}$ is an upper bound on the distribution.

Show answer

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

Question 1

0.0/1.0 point (graded)

Which of the following is true about maximum likelihood estimators? (Select all that apply.)

- [ ] Under certain regularity conditions, maximum likelihood estimators will have asymptotically normal distributions. ✔
- [ ] In general, maximum likelihood estimators are more robust to mistakes in assumptions about the underlying distribution than method of moment estimators.
- [ ] If there is an efficient estimator in a class of consistent estimators, maximum likelihood estimation will produce it. ✔
- [ ] If there is an unbiased estimator in a class of estimators, maximum likelihood estimation will produce it.

Explanation

One reason MLEs are nice is this result that they will have asymptotically normal distributions under certain assumptions. MLEs are also nice because we know that they will produce an efficient estimator if it exists in that class of estimators. However, MLEs are in general more sensitive to mistakes about the underlying distribution, because they rely on these assumptions more than estimators derived using method of moments. We know that MLEs do not always produce unbiased estimators; for example, the n^{th} order statistic is a biased estimator for θ in $U[0,\theta$ distributions.

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

These questions give a taste of each topic we have covered so far. They are by no means comprehensive.

Question 1

0.0/1.0 point (graded)

Which of the following is always equal to $P(X|Y)$? Note: X and Y are not assumed to be independent. (Select all that apply.)

- [ ] $P(Y|X) * P(X) / P(Y)$ ✔
- [ ] $P(X,Y) / P(Y)$ ✔
- [ ] $P(X) * P(Y) / P(Y)$
- [ ] $P(X|Y) * P(Y) / P(X)$

Explanation

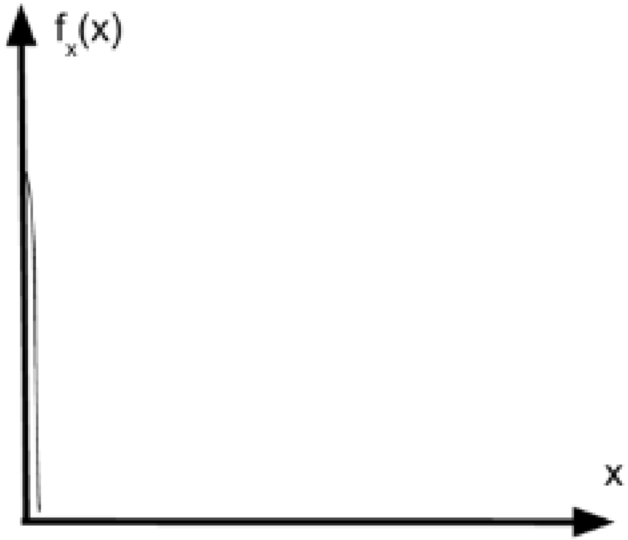
This question reviews conditional probability and Bayes' Rule. $P(Y|X) * P(X) / P(Y)$ and $P(X,Y) / P(Y)$ are different ways of stating Bayes' Rule and the definition of conditional probability. $P(X) * P(Y) / P(Y)$ is only true in the case that X and Y are independent.

Question 2

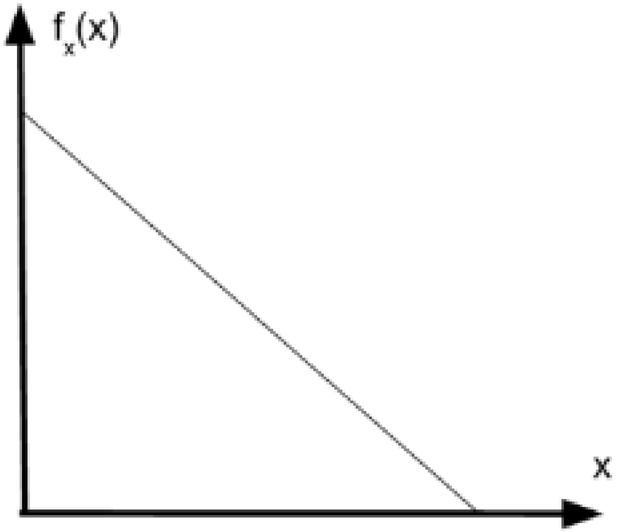
0.0/1.0 point (graded)

PDFs:

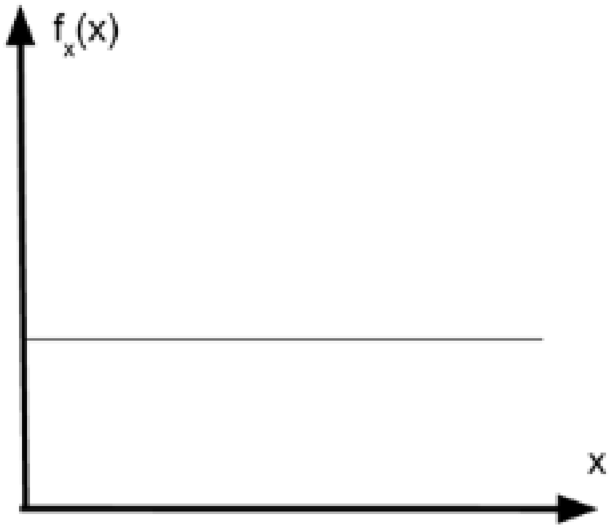
PDF A:



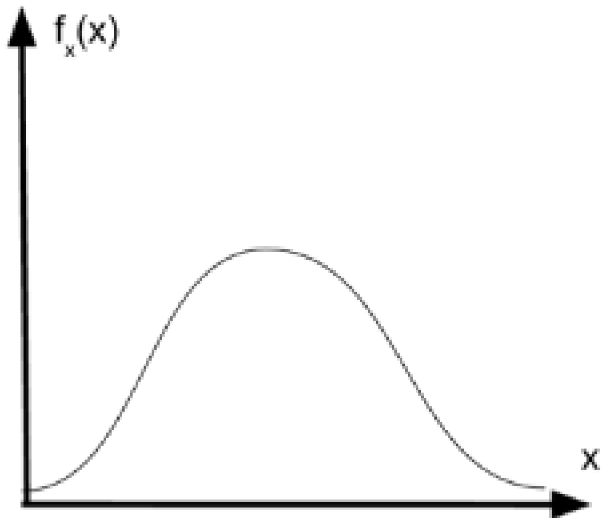
PDF B:



PDF C:

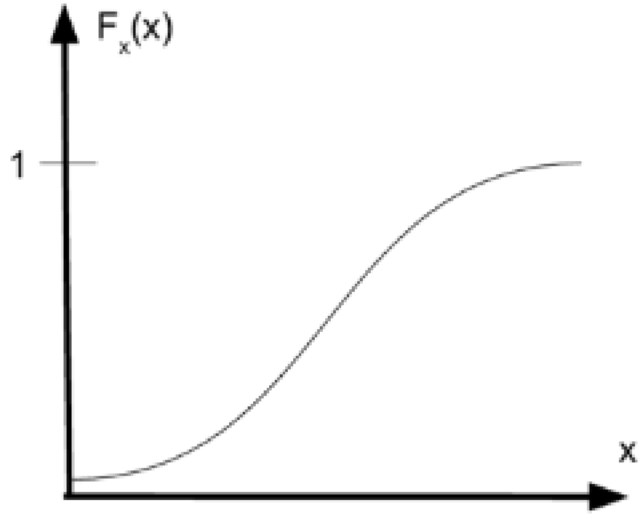


PDF D:

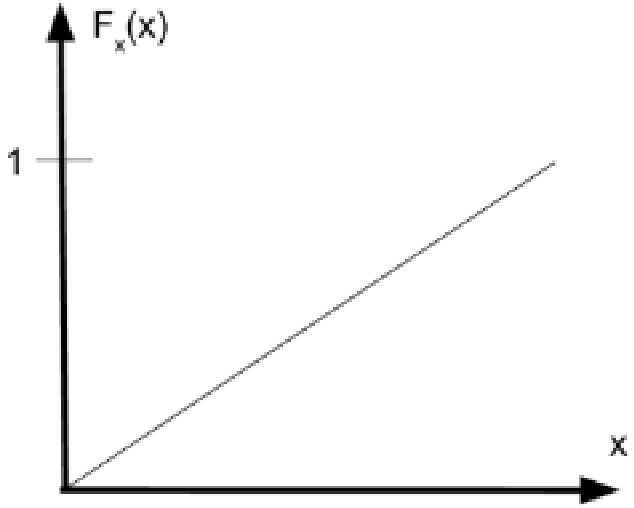


CDFs:

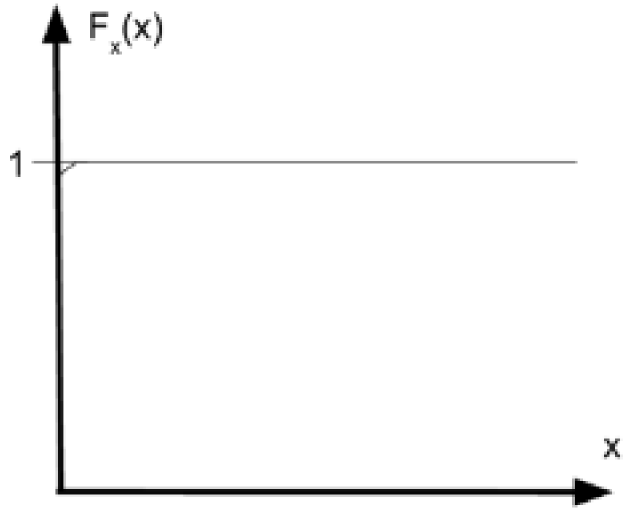
CDF E:



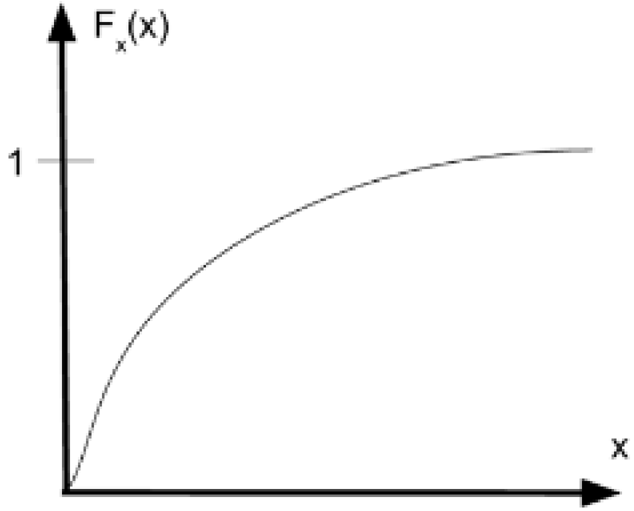
CDF F:



CDF G:



CDF H:



Match the PDFs (probability distribution functions) above with their corresponding CDFs (cumulative distribution functions):

Note: The corresponding label for each graph is at the bottom left

PDF A

Select an option **Answer:** g

PDF B

Select an option **Answer:** h

PDF C

Select an option **Answer:** f

PDF D

Select an option **Answer:** e

Explanation

The cumulative distribution function represents the probability that the random variable is less than or equal to the argument to the function. It is equal to the area under the probability distribution function. We can integrate the probability distribution function to get the cumulative distribution function.

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 3

0.0/1.0 point (graded)

Suppose a coin comes up heads with probability 2/5. What is the mean and variance of the distribution that expresses the number of times the coin comes up heads in 5 tosses?

Mean:

Answer: 2

Variance:

Answer: 6/5

Explanation

Obtaining heads on a coin flip follows a binomial distribution, where $H|N = n \sim B(n, 2/5)$ with H referring to the number "obtaining heads," and N refers to the the number of coin flips.

The expectation for this binomial distribution after 5 tosses is given by: $E[H|N]$, where N =the number of coin flips. $E[H|N=5] = np = 5(2/5) = 2$. The variance for a binomial distribution is calculated by: $np(1-p)$. For 5 tosses, this will be given as: $5\left(\frac{2}{5}\right)\left(1-\frac{2}{5}\right) = 2\left(\frac{3}{5}\right) = \frac{6}{5}$.

Question 3

0.0/1.0 point (graded)

Suppose a coin comes up heads with probability 2/5. What is the mean and variance of the distribution that expresses the number of times the coin comes up heads in 5 tosses?

Mean:

Answer: 2

Variance:

Answer: 6/5

Explanation

Obtaining heads on a coin flip follows a binomial distribution, where $H|N = n \sim B(n, 2/5)$ with H referring to the number "obtaining heads," and N refers to the the number of coin flips.

The expectation for this binomial distribution after 5 tosses is given by: $E[H|N]$, where N =the number of coin flips. $E[H|N = 5] = np = 5(2/5) = 2$. The variance for a binomial distribution is calculated by: $np(1-p)$. For 5 tosses, this will be given as: $5\left(\frac{2}{5}\right)\left(1-\frac{2}{5}\right) = 2\left(\frac{3}{5}\right) = \frac{6}{5}$.

Question 4

0.0/1.0 point (graded)

Which of the following is equal to $E\left[aX + bY + c\right]$? (Select all that apply.)

- [ ] $aE\left[X\right] + bE\left[Y\right] + c$ ✔
- [ ] $a^2E\left[X\right] + b^2E\left[Y\right] + c$
- [ ] $E\left[aX\right] + E\left[bY\right] + c$ ✔
- [ ] $\left(a + b + c\right)E\left[X + Y\right]$

Explanation
This question reviews properties of expectation. By linearity of expectation, we can add together the expectation of each random variable separately.

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 5

0.0/1.0 point (graded)

Which of the following is equal to Var$(aX+bY+c)$?

- $a\mathrm{Var}\left(X\right)+b\mathrm{Var}\left(Y\right)$
- $a^2\mathrm{Var}\left(X\right)+b^2\mathrm{Var}\left(Y\right)+2ab\mathrm{Cov}\left(X,Y\right)$ ✔
- $\mathrm{Var}\left(aX\right)+\mathrm{Var}\left(bY\right)+\mathrm{Var}\left(c\right)$
- $\left(a+b\right)\mathrm{Var}\left(X+Y\right)$

Explanation

This question reviews properties of variance. The variance of a constant is equal to zero.

Show answer