



Bookmarks

- ▶ [Module 1: The Basics of R and Introduction to the Course](#)
- ▶ [Entrance Survey](#)
- ▶ [Module 2: Fundamentals of Probability, Random Variables, Distributions, and Joint Distributions](#)
- ▼ [Module 3: Gathering and Collecting Data, Ethics, and Kernel Density Estimates](#)

[Gathering and Collecting Data](#)[Finger Exercises due Oct 17, 2016 05:00 IST](#)

Module 3: Gathering and Collecting Data, Ethics, and Kernel Density Estimates > Summarizing and Describing Data > Kernel Density Estimation, Part I - Quiz

## Kernel Density Estimation, Part I - Quiz

🔖 Bookmark this page

### Question 1

1 point possible (graded)

The choice of the boundaries for a kernel density plot is dictated by a trade off between \_\_\_\_\_ caused by too small of a bandwidth and \_\_\_\_\_ as a result of too large of a bandwidth.

- ☒ a. variance, bias
- ☐ b. bias, variance
- ☐ c. bias, smoothness
- ☐ d. variance, smoothness

### Explanation

If your bandwidth is smaller than optimal, your function will look very jagged. As Professor Dufo mentions, the smaller the bandwidth, the bigger the variance. When you pick a bandwidth that is too large for your data, the function will try to find data that does not exist. This will cause the

**Summarizing and Describing Data**

Finger Exercises due Oct 17, 2016 05:00 IST

**Module 3: Homework**

Homework due Oct 10, 2016 05:00 IST



- ▶ Module 4: Joint, Marginal, and Conditional Distributions & Functions of Random Variable
- ▶ Module 5: Moments of a Random Variable, Applications to Auctions, & Intro to Regression
- ▶ Module 6: Special Distributions, the Sample Mean, the Central Limit Theorem, and Estimation

result kernel density function to be biased, meaning that there is an important characteristics of the data that you are missing as a result of selecting a non-optimal bandwidth.

Submit

You have used 0 of 2 attempts

**Question 2**

1 point possible (graded)

For which of the following scenarios would it be better to use a histogram over a kernel density plot?

- ☐ a. To show a continuous distribution function of student's math grades
- ☐ b. To demonstrate the impact of an intervention on school attendance rates
- ☐ c. To show the difference in the distribution of male and female height
- ☒ d. To display the proportion of the U.S. households that fall within certain income brackets (e.g. between 0 and \$15,000; between \$15,000 and \$30,000; etc.)

Explanation

- ▶ [Module 7: Assessing and Deriving Estimators - Confidence Intervals, and Hypothesis Testing](#)
- ▶ [Module 8: Causality, Analyzing Randomized Experiments, & Nonparametric Regression](#)
- ▶ [Module 9: Single and Multivariate Linear Models](#)
- ▶ [Module 10: Practical Issues in Running Regressions, and Omitted Variable Bias](#)
- ▶ [Module 11: Intro to Machine Learning and Data Visualization](#)
- ▶ [Module 12: Endogeneity](#)

As discussed in this segment, a kernel density plot provides a continuous probability density function of a random variable. In the last segment, we learned that a histogram is also an estimate of a probability distribution for a continuous variable, but, unlike the kernel density function, it is not a continuous function. Instead, a histogram demonstrates the number of observations or the proportion of observations that fit into a certain “bin.” From this, a histogram would be preferable to a kernel density function for cases in which we want to demonstrate the proportion of observations that fall into certain bins, such as when we want to show the proportion of households that fall within specified income brackets.

You have used 0 of 2 attempts

### Discussion

**Topic:** Module 3 / Kernel Density Estimation, Part I - Quiz

Instrumental Variables,  
and Experimental  
Design

- ▶ Exit Survey
- ▶ Final Exam

© All Rights Reserved



© 2016 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY  
OPENedX®

