

University of Colorado
Department of Civil, Environmental and Architectural Engineering
Quantitative Methods CVEN-5454

Midterm

Date :04/03/2003

15 minutes

10 points

1. What are the key differences between parametric and nonparametric approaches to functional estimation? (list atleast two) [2]

Parametric: (i) The fit is global (i.e. One equation is fit to the entire data)
(ii) Underlying functional form is assumed (e.g. Assume Gamma or Normal or Exponential distribution and fit it; or assume a linear relationship for fitting a straight line etc.)

Nonparametric: (i) The fit is local (i.e. Estimate at a point depends on small number of neighboring points)
(ii) No assumption regarding the underlying functional form

2. For any function to be a probability density function what are the two required conditions? [1]

$f(x) \geq 0$ and integral of $f(x)$ over the entire domain of X should equal 1.
For discrete data the probability mass function $P(x)$ should sum to 1.

3. Define Bootstrap and Monte Carlo – (few words to a sentence). List their utility. [1,1,1]

Bootstrap: Resample the data with replacements to generate new samples

Monte Carlo: Simulate values from a fitted CDF.

Utility: To estimate confidence intervals for any model parameters; confidence intervals for the basic statistics (e.g. Mean, variance, skew, etc.); To test differences between groups (i.e. Difference in Means, Variance etc.) and last but not least, for testing various hypotheses.

4. In kernel density estimators what is the key parameter and what is one of the main draw back of such estimators? [1]

Bandwidth (h) is the key parameters

The bias due to boundary (i.e. The boundary problem) is one of the main draw backs. It gets worse in higher dimensions.

5. If conditional pdf $P(X | Y)$ is equal to unconditional pdf $P(X)$ then what can you say about the random variables X and Y ? [1]

X and Y are independent

6. When do you fit a probability mass function? Give an example [1]

For Discrete data. E.g., # of wet days in a month, # of days of water quality violations in a year etc .

7. In performing the tests for means and variance (i.e. t-test, f-test etc.) what is the basic assumption? [1]

The data is assumed to be normally distributed.

Closed part