DATA 100: Vitamin 7 Solutions

April 9, 2019

1 Statistical Inference

Probability theory is used in statistical inference because it allows us to characterize randomness and quantify ____ due to sampling.

- ☐ observations

- ☑ uncertainty

- ☐ probability

- ☐ error

2 Random Variables

Which of the following are true of discrete random variables?

- ☑ They can be added to other random variables

- ☑ They can be multiplied by other random variables

- ☑ They are denoted by capital letters

- ☑ They have an associated expectation and variance

3 Expectation and Variance

Which of the following are true?

- ☑ $E[X + Y] = E[X] + E[Y]$ for all random variables X and Y.

- ☑ $E[X + Y] = E[X] + E[Y]$ for independent random variables X and Y.

- ☐ $Var[X + Y] = Var[X] + Var[Y]$ for all random variables X and Y.

- ☑ $Var[X + Y] = Var[X] + Var[Y]$ for independent random variables X and Y.

Explanation: Option three is incorrect because it fails to take into account the covariance between X and Y. When X and Y are independent, their covariance is 0, and therefore option 4 is correct.

4   Estimators

An "estimator" is a function ...

☐ of a statistic that computes a parameter.

☐ of a distribution that computes a statistic.

☐ of a random variable that computes its expectation.

☑ of a sample that computes an estimate of a parameter.

5   Error

The "mean squared error" of an estimator X with respect to a parameter θ is defined as:

☐ $(X - \theta)^2$

☐ $E[X - \theta]^2$

☑ $E[(X - \theta)^2]$

☐ $(E[X] - \theta)^2$

Explanation: The L_2 loss function for random variable X and parameter θ is defined as $L_2(X, \theta) = (X - \theta)^2$. The risk of this loss function is defined as $E[(X - \theta)^2]$, which is commonly known as the mean squared error.