

Quiz 22

written by Alvin Wan . alvinwan.com/cs70

Thursday, April 14, 2016

This quiz does not count towards your grade. It exists to simply gauge your understanding. Treat this as though it were a portion of your midterm or final exam. Today, we will walk through conditional expectation with problems of increasing difficulty.

1 Linear Regression

1. We are pulling marbles at random from a bag and the placing the marble back in before sampling again. Let X be the number times we sample before we get a red marble. Let Y be the number of times we sample before we get a blue marble. Assume there are a red, b blue, and c randomly-colored marbles in the bag. In this problem, we will walk through computing $L[Y|X]$.

Note: This problem is largely to walk you through computing these values. Linear regression is useful when you have only *data* and need to find a model to fit it.

- (a) Compute $E[X]$.
- (b) Compute $E[Y]$.
- (c) Compute $var(X)$
- (d) Compute $cov(X, Y)$. (Note that this gives us a fairly boring result.)
- (e) Finally, compute $L[Y|X]$. Remember that the formula for linear regression is the following:

$$L[Y|X] = E[Y] + \frac{cov(X, Y)}{var(X)}(X - E[X])$$

2. You are given a set of sample points, taken uniformly at random from $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Ω	1	2	3
$X(\Omega)$	2	4	6
$Y(\Omega)$	2	1	3

- (a) Compute $E[X]$.
- (b) Compute $E[Y]$.
- (c) Compute $var(X)$.
- (d) Compute $cov(X, Y)$.
- (e) Compute $L[Y|X]$.
- (f) Estimate Y , when $X = 4$. (It's not 1!)