STA257

Neil Montgomery | HTML is official | PDF versions good for in-class use only
Last edited: 2016-11-22 13:26

$E(g(X))$ and extensions

Motivation: suppose $X \sim N(\mu, \sigma^2)$. What is $E(X)$? The answer is μ. Lots of ways to figure this out.

Using the density is tedious but do-able. Or we could use the fact that $X = \mu + \sigma Z$ with $Z \sim N(0, 1)$.

Theorem: Given X and $E(X)$ exists, consider $g(x) = a + bx$. Then $E(g(X)) = E(a + bX) = a + bE(X)$.

Proof: …

$E(g(X))$

A theorem which is too difficult to prove generally is: given X, any* g, and $Y = g(X)$, then:

$$E(Y) = E(g(X)) = \begin{cases} \sum g(x)p(x) & : X \text{ discrete} \\ \int g(x)f(x)\,dx & : X \text{ continuous} \end{cases}$$

in both cases provided the sum/integral congerges "absolutely" (i.e. with $|g(x)|$.)

Example: Average volume of sphere with radius $R \sim \text{Exp}(1)\ldots$

$E(g(X_1, \ldots, X_n))$

Some typical applications:

$$E(X_1 \cdot X_2)$$

$$E\left(X_1 + X_2\right)$$

$$E(X_1 + \cdots + X_n)$$

$$E\left(\overline{X}\right) = E\left(\frac{X_1 + \cdots + X_n}{n}\right)$$

Theorem (continuous version): $X_1, \ldots, X_n$ have joint density $f(x_1, \ldots, x_n)$ and $Y = g(X_1, \ldots, X_n)$. Then:

$$E(Y) = \int \ldots \int g(x_1, \ldots, x_n) f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n$$

examples

Suppose $X_1 \perp X_2$. Consider $E(X_1 \cdot X_2)\ldots$

Exercise: $X_1 \perp X_2$. Consider $E(g(X_1)\, h(X_2))$

Now suppose $X_1, \ldots, X_n$ are i.i.d. with $E(X_i) = \mu$. Consider:

$$E\left(\overline{X}\right) = E\left(\frac{X_1 + \cdots + X_n}{n}\right)\ldots$$

$X \sim \text{NegBin}(r, p)\ldots$

putting a number on variation

Expected value is a measure of "location", but random variables with the same "location" can be quite different.

Consider the coin tossing game with $E(Y) = 0$:

$$P(Y = y) = \begin{cases} 0.5 & : y = 100 \\ 0.5 & : y = -100 \end{cases}$$

One thing leads to another. Family trees are compared and contrasted, and after more than a few things get interesting:

$$P(Y_2 = y) = \begin{cases} 0.5 & : y = 1000 \\ 0.5 & : y = -1000 \end{cases}$$

Still, $E(Y_2) = 0$. But the values of Y_2 are more spread out.

variance

One way to measure spread is to use the of X, defined as: $\mathrm{Var}(X) = E\left[(X - E(X))^2\right]$.

This is a use of $E(g(X))$ with $g(x) = (x - E(X))^2$.

Very useful:

$$\begin{aligned}
\mathrm{Var}(X) &= E\left(X^2 - 2XE(X) + E(X)^2\right) \\
&= E\left(X^2\right) - 2E(X)E(X) + E(X)^2 \\
&= E\left(X^2\right) - E(X)^2.
\end{aligned}$$

examples

$X \sim \text{Bernoulli}(p)\ldots$

$Z \sim N(0,1)\ldots$

$X \sim \text{Poisson}(\lambda)\ldots$(uses a trick!)

Variance of $X = a$ constant.

Basic examples for exercise: Exponential, Gamma, Geometric (trick: differentiate power series twice), Binomial (use Poisson trick).

$\mathrm{Var}(a + bX)$, $\mathrm{Var}(X + Y)$ (independent case)

$\mathrm{Var}(a + bX) = b^2\mathrm{Var}(X)$. Proof…

Example: $X \sim N(\mu, \sigma^2)$

When $X \perp Y$, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. Proof…

Actually independence is stronger than necessary. Only needed $E(XY) = E(X)E(Y)$; to be revisited.

variance of the "sample average"

This is a "grand" example of particular importance.

Suppose again $X_1, \ldots, X_n$ is i.i.d. with $E(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$.. We already know $E\left(\overline{X}\right) = \mu$.

What about $\mathrm{Var}\left(\overline{X}\right)$?