Topic 8

The Expected Value

Functions of Random Variables

Outline

Names for $Eg(X)$

- If $g(x) = x$, then $\mu = EX$ is called variously the distributional mean, and the first moment.
- If $g(x) = x^k$, then EX^k is called the k-th moment. These names were made in analogy to a similar concept in physics. The second moment in physics is associated to the moment of inertia.
- If $g(x) = (x - \mu)^k$, then $E(X - \mu)^k$ is called the k-th central moment.
- The most frequently used central moment is the second central moment $\sigma^2 = E(X - \mu)^2$ commonly called the (distributional) variance. We often write $\text{Var}(X)$ for the variance.

Variance and Standard Deviation

$$\begin{aligned}\sigma^2 &= \text{Var}(X) = E(X-\mu)^2 = EX^2 - 2\mu EX + \mu^2 \\ &= EX^2 - 2\mu^2 + \mu^2 = EX^2 - \mu^2.\end{aligned}$$

In analogy with the corresponding concept with quantitative data, we call σ the standard deviation.

If we subtract the mean and divide by the standard deviation, the resulting random variable

$$Z = \frac{X-\mu}{\sigma}$$

has mean 0 and variance 1. Z is called the standardized version of X.

Variance and Standard Deviation

Exercise. Compute the variance and standard deviation for

1. a single Bernoulli trial,
2. a fair dice
3. the distance on a dart board.
4. $Y = aX + b$ for constants a and b. Give the answer in terms of the variance or standard deviation of X.

Names for $Eg(X)$

- The third moment of the standardized random variable

$$E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

  is called the skewness. Random variables with positive skewness have a more pronounced tail to the density on the right. Random variables with negative skewness have a more pronounced tail to the density on the left.

- The fourth moment of the standard normal random variable is 3. The kurtosis compares the fourth moment of the standardized random variable to this value

$$E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]-3.$$

  Random variables with a negative kurtosis are called leptokurtic. Lepto means slender. Those with a positive kurtosis are called platykurtic. Platy means broad.

Names for $Eg(X)$.

Expected values in the case of more than one random variable is based on the same concepts as for a single random variable.

For example, for two discrete random variables X_1 and X_2 and for a real valued function g, we have

$$Eg(X_1, X_2) = \sum_{x_1} \sum_{x_2} g(x_1, x_2) f_{X_1,X_2}(x_1, x_2)$$

the expected value is based on the joint mass function $f_{X_1,X_2}(x_1, x_2)$.

Independence

For the case in which the random variables are independent. Here, we have the factorization identity $f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$ for the joint mass function.

Apply the formulate for $Eg(X_1,X_2)$ to the product of functions $g(x_1,x_2) = g_1(x_1)g_2(x_2)$ to find that

$$\begin{aligned}
E[g_1(X_1)g_2(X_2)] &= \sum_{x_1}\sum_{x_2} g_1(x_1)g_2(x_2)f_{X_1,X_2}(x_1,x_2) \\
&= \sum_{x_1}\sum_{x_2} g_1(x_1)g_2(x_2)f_{X_1}(x_1)f_{X_2}(x_2) \\
&= \sum_{x_1}\sum_{x_2} g_1(x_1)g_2(x_2)f_{X_1}(x_1)f_{X_2}(x_2) \\
&= \left(\sum_{x_1} g_1(x_1)f_{X_1}(x_1)\right)\left(\sum_{x_2} g_2(x_2)f_{X_2}(x_2)\right) = E[g_1(X_1)] \cdot E[g_2(X_2)]
\end{aligned}$$

Covariance and Correlation

Take X_1 and X_2 random variables with respective means μ_1 and μ_2. Then the variance of their sum

$$\begin{aligned}
\mathrm{Var}(X_1 + X_2) &= E[((X_1 + X_2) - (\mu_1 + \mu_2))^2] = \\
E[((X_1 + X_2) - (\mu_1 + \mu_2))^2] & \\
&= E[((X_1 - \mu_1) + (X_2 - \mu_2))^2] \\
&= E[(X_1 - \mu_1)^2] + 2E[(X_1 - \mu_1)(X_2 - \mu_2)] + E[(X_2 - \mu_2)^2] \\
&= \mathrm{Var}(X_1) + 2\mathrm{Cov}(X_1, X_2) + \mathrm{Var}(X_2).
\end{aligned}$$

where the covariance $\mathrm{Cov}(X_1, X_2) = E[(X_1 - \mu_1)(X_2 - \mu_2)]$

Covariance and Correlation

The definition of covariance is analogous to that for a sample covariance. The analogy continues for the correlation, ρ, defined for random variables X_1 and X_2, as

$$\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)}\sqrt{\text{Var}(X_2)}}.$$

We can also modify the argument used for sample covariance to see that

$$-1 \leq \rho(X_1, X_2) \leq 1.$$

Correlation ± 1 occurs only when X and Y have a perfect linear association.

Covariance and Correlation

If X_1 and X_2 are independent, then

$$\begin{aligned}\mathrm{Cov}(X_1, X_2) &= E[(X_1 - \mu_1)(X_2 - \mu_2)] \\ &= E[X_1 - \mu_1] \cdot E[X_2 - \mu_2] = 0\end{aligned}$$

and $\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2)$.

We can generalize the Pythagorean identity to independent $X_1, \ldots, X_n$ and constants $c_1, \ldots, c_n$.

$$\begin{aligned}&\mathrm{Var}(c_1 X_1 + \cdots c_n X_n) \\ = \; &c_1^2 \mathrm{Var}(X_1) + \cdots + c_n^2 \mathrm{Var}(X_n).\end{aligned}$$

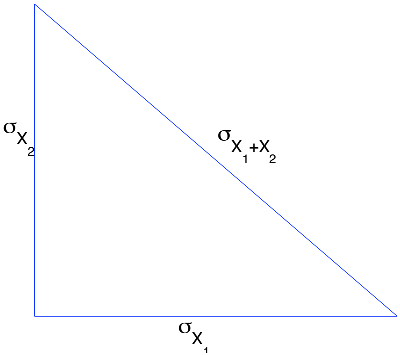

Figure: Pythagorean identity for the variance of independent random variables

Covariance and Correlation

Exercise.

1. Let X and Z be independent random variables mean 0, variance 1. Define $Y = \rho_0 X + \sqrt{1-\rho_0^2}\, Z$.
    - Show that Y has mean 0, variance 1.
    - Show that X and Y have correlation ρ_0.
2. Find the variance of a binomial random variable based on n trials with success parameter p.