Stats 130
Day 16 Notes

Elijah Hantman

Variance

The variance is a measure of the dispersion of the distribution of a random variable. It is a measure of how "wide" a distribution is.

Definition:

$$var(X) = E(X-\mu)^2, \qquad \mu = EX$$
$$var(x) = E(X-\mu)^2 = E(X^2 - 2X\mu + \mu^2) = EX^2 - 2\mu EX + \mu^2$$
$$= EX^2 - 2\mu^2 + \mu^2 = EX^2 - \mu^2$$

Usually the last line is the easiest way to calculate the variance.

Properties of the Variance

- $var(X) \geq 0$
  Since the variance is the square of a real value, it is definitionally non-negative.
- $var(X) = 0 \iff Pr(X = c) = 1$
- For constants a and b, $var(aX + b) = a^2 var(X)$
  The variance of a constant is zero, and the variance is proportional to the square of what is inside. So constant shifts don't affect the variance, and coefficients become squared when pulled outside.
- $X_1, ..., X_n$ independent, then $var(\sum_{i=1}^n X_i) = \sum_{i=1}^n var(X_i)$

Bernoulli Trials

$X \sim Ber(p)$ then

$$var(X) = EX^2 - (EX)^2 = p - p^2 = pq$$

$Y \sim Bin(n, p)$

$$Y = \sum var(X)$$

Poisson

$X \sim Pois(\lambda)$

$$var(X) = EX^2 - (EX)^2 = \lambda^2 + \lambda - \lambda^2 \tag{1}$$
$$= \lambda \tag{2}$$

Geometric

$X \sim Geo(p)$

$$var(X) = E(X(X-1)) + EX - (EX)^2 = 2\frac{q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{q}{p^2} \tag{3}$$

$Y \sim NegBin(r, p)$

$$var(Y) = \sum_{i=1}^{r} var(X_i) = r\frac{q}{p^2}$$

Covariance

The mean and the variance provide a characterization of the average behavior of a single random variable. In order to measure the association between two random variables we use the covariance.

$$\mu_X = EX \tag{4}$$
$$\mu_Y = EY \tag{5}$$
$$cov(X,Y) = E((X-\mu_X)(Y-\mu_Y)) \tag{6}$$
$$= E(XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y) \tag{7}$$
$$= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X \mu_Y \tag{8}$$
$$= E(XY) - 2\mu_X \mu_Y + \mu_X \mu_Y \tag{9}$$
$$= E(XY) - \mu_X \mu_Y \tag{10}$$

Order doesn't matter due to the multiplication. If X and Y are independent $cov(X,Y) = 0$.

Something to note is that $cov(X,Y) = 0$ does not imply that X and Y must be independent. Also, $cov(X,X) = var(X)$

Examples

Consider two random variables with joint density

$$f(x,y) = \begin{cases} x+y & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Calculate covariance.

$$\mu_X = \int_0^1 \int_0^1 x(x+y)dydx = \frac{7}{12} = \mu_Y \tag{11}$$
$$EXY = \int_0^1 \int_0^1 xy(x+y)dydx = \frac{1}{3} \tag{12}$$

then

$$cov(X,Y) = EXY - \mu_X \mu_Y = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}$$

Interpretation

The covariance measures how random variables tend to move together. It answers the question, if X is big, what is Y likely to be.
The Covariance can range from negative to positive infinity, and it is dependent on the scale used. This makes the Covariance difficult to directly interpret as it carries whatever units you are using.

Correlation

One way is to use a standardized covariant. We use the Cauchy-Schwarz inequality.

Theorem: For any two random variables U and V with finite variance and mean, then:

$$E(UV)^2 \leq EU^2EV^2$$

The correlation coefficient for X and Y is defined as:

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$$

By taking $U = X - \mu_X$ and $V = Y - \mu_Y$ we have that

$$(cov(X,Y))^2 \leq var(X)var(Y)$$

This means that $|\rho(X,Y)| \leq 1$

Properties

- If X and Y are independent then $\rho(X,Y) = 0$
- if $Y = aX$ for a real number a, then

$$\rho(X,Y) = \begin{cases} 1 & a > 0 \\ -1 & a < 0 \end{cases}$$

$$\rho(X,Y) = \frac{cov(X,aX)}{\sqrt{var(X)a^2var(X)}} = \frac{acov(X,Y)}{|a|var(X)} = \pm 1$$

- $var(X+Y) = varX + varY + 2cov(X,Y)$
  Consequence of $var(X+Y) = E(X+Y)^2 - (\mu_X + \mu_Y)^2$
- $$var\sum_i X_i = \sum_i varX_i + 2\sum\sum_{i<j} cov(X_i, X_j)$$

Conditional Expectation

For two random variables X and Y, then conditional distribution of one given the other takes a specific value is a distribution we can calculate an expectation for.

The conditional expectation of X, given $Y = y$ is

$$E(X|Y = y) = \begin{cases} \sum_x x g_X(x|y) & \text{discrete} \\ \int_{\mathbb{R}} x g_X(x|y) dx & \text{continuous} \end{cases}$$

Notice that $h(y) = E(X|Y = y)$ defines a function of y.

Consider the random variable $h(y)$.

The conditional expectation of X, given Y is $E(X|Y) = h(Y)$. This is a random variable.

Theorem: $E(E(X|Y)) = E(X)$

$$E(E(X|Y)) = E(h(Y)) = \int h(y) f_Y(y) dy$$

$$= \int \int x g_X(x|y) f_Y(y) dy dx$$

$$= \int \int x f(x, y) dy dx = EX$$

Examples

Consider a clinical trial where patients have two possible outcomes. Suppose the probability of success is not known, and you denote it as a random variable Θ. Then, for patient i, $Pr(X_i = i|\Theta = \theta) = \theta$. The total number of successes is $Y = \sum_{i=1}^{n} X_i$ and $E(Y|\Theta = \theta) = n\theta$

Suppose that we assume $\Theta \sim Unif(0, 1)$, then $E(Y|\Theta) = n\Theta$. We have to assume $E\Theta = 1/2$ then

$$EY = E(E(Y|\Theta)) = E(n\Theta) = \frac{n}{2}$$