Regression II
Vanderbilt University
Prof. Sean P. Corcoran

Expected Values, Variance, Covariance, and Correlation

Expected value
For a discrete random variable X with probability density function $f(x) = Pr(X = x)$:

$$\begin{aligned}\mu = E(X) &= x_1 f(x_1) + x_2 f(x_2) + ... + x_k f(x_k) \\ &= \sum_{j=1}^{k} x_j f(x_j)\end{aligned}$$

For a continuous random variable X and PDF $f(X)$:

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

For continuous function $g(X)$, the expected value of $g(X)$ for discrete and continuous X is:

$$E[g(X)] = \sum_{j=1}^{k} g(x_j) f(x_j)$$
$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Useful note: if $g(X)$ is nonlinear, $E[g(X)] \neq g(E(X))$.

Properties of expected values:

- $E(c) = c$ for any constant c
- $E(aX + b) = aE(X) + b$ for any constants a and b
- For n random variables $X_1, ..., X_n$ and constants $a_1, ..., a_n$:

$$E(a_1 X_1 + a_2 X_2 + ... + a_n X_n) = a_1 E(X_1) + a_2 E(X_2) + ... + a_n E(X_n)$$
$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)$$

Variance
The variance of X is the expected (squared) deviation of X from its mean:

$$\sigma^2 = Var(X) = E(X-\mu)^2$$
or
$$= E(X^2) - \mu^2$$

Properties of variance:

- The standard deviation is the square root of the variance: $sd(X) = \sigma$
- $Var(aX+b) = a^2 Var(X)$ for any constants a and b
- $sd(aX+b) = |a| sd(X)$ for any constant a
- $Z = (X-\mu)/\sigma$ is a standardized version of X with $E(Z) = 0$ and $Var(Z) = 1$

Covariance
The covariance between two random variables X and Y is a measure of linear dependence:

$$\sigma_{XY} = Cov(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$
or
$$= E(XY) - \mu_X \mu_Y$$

Correlation
The correlation between two random variables X and Y is:

$$\rho_{XY} = Corr(X,Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Properties of covariance and correlation:

- $-1 \leq \rho_{XY} \leq +1$
- If X and Y are independent, then $\sigma_{XY} = \rho_{XY} = 0$, but the latter does not imply the former.
- $Cov(a_1 X + b_1, a_2 Y + b_2) = a_1 a_2 Cov(X,Y)$ for any constants a_1, b_1, a_2, b_2
- $Corr(a_1 X + b_1, a_2 X + b_2) = Corr(X,Y)$ if $a_1 a_2 > 0$ and $-Corr(X,Y)$ if $a_1 a_2 < 0$
- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab Cov(X,Y)$
- $Var\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 Var(X_i) + \sum_{i=1}^n \sum_{j=i+1}^n 2 a_i a_j Cov(X_i, X_j)$.
  Note the summation notation in the last term: this is a sum of all pairwise covariances, where each pair is counted only once.

Conditional Expectation For a discrete random variable Y with conditional probability density function $f(Y|X) = Pr(Y = y|X = x)$:

$$E(Y|x) = \sum_{j=1}^{m} y_j f_{Y|X}(y_j|x)$$

Properties of conditional expectations

- $E[c(X)|X] = c(X)$ for any function $c(X)$
- $E(g(x)Y|X) = g(x)E(y|x)$: "when you condition on X you can effectively treat it like a constant"
- $E[a(X)Y + b(X)|X] = a(X)E(Y|X) + b(X)$ for any functions $a(X)$ and $b(X)$
- If X and Y are independent then $E(Y|X) = E(Y)$
- Law of iterated expectations (I): $E[E(Y|X)] = E(Y)$
- Law of iterated expectations (II): $E(Y|X) = E[E(Y|X, Z)|X]$
- If $E(Y|X) = E(Y)$ then $Cov(X, Y) = 0$. In fact, every function of X will be uncorrelated with Y