otes on Expectation, Moment Generating Functions, Variance, Covarian

For a random variable X, the *expected value* of X (or the *mean* or the *average* of X) is given by

$$E(X) = \begin{cases} \sum_{x\in A} x \cdot p(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x f(x) dx & \text{if } X \text{ is continuous} \end{cases}$$

We can also compute expectations of functions of the random variable X, using the "Law of the Unconscious Statistician":

$$E(g(X)) = \begin{cases} \sum_{x\in A} g(x) \cdot p(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} g(x) \cdot f(x) dx & \text{if } X \text{ is continuous} \end{cases}$$

WARNING: Do **NOT** use the Law of the Unemployed Statistician, which would say that $E(g(X)) = g(E(X))$. (*Example*: It is NOT true, in general, that $E(X^2) = [E(X)]^2$, since this means that $var(X) = 0$.)

Example: If $X \sim exp(\lambda)$, then

$$E(X^3 \cos X) = \int_0^{\infty} x^3 \cos x \cdot \lambda e^{-\lambda x} dx$$

.

More generally, for two random variables X and Y,

$$E(g(X,Y)) = \begin{cases} \sum_x \sum_y g(x,y) \cdot p(x,y) & \text{if } X, Y \text{ are discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \cdot f(x,y) dy dx & \text{if } X \text{ and } Y \text{ are continuous} \end{cases}$$

Examples: If X and Y are uniform over the 2-by-2 square centered on the origin, then

$$E(e^X Y^3) = \int_{-1}^{1} \int_{-1}^{1} e^x y^3 \cdot \frac{1}{4} dy dx$$

and

$$E(X) = \int_{-1}^{1} \int_{-1}^{1} x \cdot \frac{1}{4} dy dx$$

Def. The nth *moment* of X is $E(X^n)$.

Def. The *moment generating function* of X is $M_X(t) = E(e^{tX})$, provided that this expectation exists (is finite) for values of t in some interval $(-\delta, \delta)$ that contains $t = 0$. Moment generating functions are useful for generating the moments of X (hence, the name!): to compute the nth moment of X, we simply take the nth derivative (with respect to t) of $M_X(t)$, and then plug in $t = 0$:

$$E(X) = M_X'(0), E(X^2) = M_X''(0), E(X^n) = M_X^{(n)}(0)$$

Example: If $X \sim exp(\lambda)$, then

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} \cdot \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}$$

(we assume that $t < \lambda$, or else the integral blows up to infinity). Thus, we can compute $M_X'(t) = \lambda/(\lambda - t)^2$, $M_X''(t) = (2\lambda)/(\lambda - t)^3$. This gives us the first and second moments: $E(X) = M_X'(0) = \lambda/(\lambda - 0)^2 = 1/\lambda$, and $E(X^2) = M_X''(0) = (2\lambda)/(\lambda - 0)^3 = 2/\lambda^2$.

Facts About Expectations

(1). $E(aX + b) = aE(X) + b$
(2). $E(X + Y) = E(X) + E(Y)$, *even* if X and Y are NOT independent. More generally, $E(\sum_i a_i X_i) = \sum_i a_i E(X_i)$.

Def. The *variance* of X is $var(X) = E((X - \mu)^2) = E(X^2) - \mu^2$, where $\mu = E(X)$ is the mean of X.

Def. The *covariance* of X and Y is $cov(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$.

Def. The *correlation* of X and Y is

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

Facts About Variance and Covariance

(1). $var(aX + b) = a^2 \cdot var(X)$; in particular, putting $a = 0$ gives: $var(b) = 0$, for a constant b.
(2). If X and Y are independent, then $E[g(X)h(Y)] = E(g(X)) \cdot E(h(Y))$. (NOTE: the converse is FALSE – if $E[g(X)h(Y)] = E(g(X)) \cdot E(h(Y))$, then X and Y need not be independent!)
(3). If X and Y are independent, then $cov(X, Y) = 0$. (follows immediately from (2), and the definition of covariance) (NOTE: again, if $cov(X, Y) = 0$, X and Y need not be independent! For an example, let $X = -1, 0, 1$, each with probability 1/3. Let Y be 0 if $X \neq 0$, and Y is 1 if $X = 0$. Then Y certainly depends on X; the rv's are NOT independent. But you can compute that $cov(X, Y) = 0$.)
(4).

$$var(X + Y) = var(X) + var(Y) + 2cov(X, Y)$$
$$var(\sum_i X_i) = \sum_i var(X_i) + 2 \sum_{i<j}\sum cov(X_i, X_j)$$

In particular, IF X and Y are independent, then $var(X + Y) = var(X) + var(Y)$.

(5). $-1 \leq \rho(X,Y) \leq 1$
(6).

$$cov(a,X) = 0$$
$$cov(a+bX, c+dY) = bd \cdot cov(X,Y)$$
$$cov(X+Y,Z) = cov(X,Z) + cov(Y,Z)$$
$$cov(\sum_i X_i, \sum_j Y_j) = \sum_i \sum_j cov(X_i, Y_j)$$

Examples:

1. $cov(X, 5X+3) = 1 \cdot 5 \cdot cov(X,X) = 5 \cdot var(X)$

2.

$$\rho(X, 5X+3) = \frac{5var(X)}{\sqrt{var(X)var(5X+3)}} = 1$$

3. Roll 2 dice. Let W_1 and W_2 be the values shown. Let $X = W_1 + W_2$ and $Y = W_1 - W_2$. Then,

$$\begin{aligned} cov(X,Y) &= cov(W_1+W_2, W_1-W_2) = cov(W_1,W_1) - cov(W_2-W_2) + cov(W_1,W_2) - cov(W_1,W_2) \\ &= var(W_1) - var(W_2) = 0 \end{aligned}$$

where we have used the fact that W_1 and W_2 are independent to conclude that $cov(W_1, W_2) = 0$. BUT, note that even though $cov(X,Y) = 0$, X and Y are NOT indpendent. (Why? Well, if we know something about X, does this affect the distribution of Y? YES: If we know that $X = 12$, then we know that $Y = 0$.)

4. If X and Y have joint density

$$f(x,y) = \begin{cases} \frac{1}{3}(x+y) & 0 < x < 1, 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

then we can compute

$$var(2X - 3Y + 8) = 4 \cdot var(X) + 9 \cdot var(Y) + 2 \cdot 2 \cdot (-3) \cdot cov(X,Y)$$

using $var(X) = E(X^2) - [E(X)]^2$, $cov(X,Y) = E(XY) - E(X)E(Y)$, $E(X) = \int_0^1 \int_0^2 x\frac{1}{3}(x+y)dydx$, $E(X^2) = \int_0^1 \int_0^2 x^2\frac{1}{3}(x+y)dydx$, $E(Y) = \int_0^1 \int_0^2 y\frac{1}{3}(x+y)dydx$, $E(Y^2) = \int_0^1 \int_0^2 y^2\frac{1}{3}(x+y)dydx$, $E(XY) = \int_0^1 \int_0^2 xy\frac{1}{3}(x+y)dydx$.