Chapter 8. Expectation for continuous random variables.

We will now generalize the definition of expectation to the case of continuous r.v.'s. Since the sample space of a continuous r.v. is necessarily uncountably infinite the same technicalities that arose in our discussion of expectation for a discrete r.v. with countably infinite sample space need to be considered before we provide a general definition of the expected value of a continuous r.v. As before the technicality is the requirement of absolute convergence but now in association with an integral instead of a sum. The integral \( \int_{-\infty}^{\infty} g(x) \, dx \) is said to be absolutely convergent if \( \int_{-\infty}^{\infty} |g(x)| \, dx \) exists.

**Definition.** If \( X \) is a continuous r.v. with p.d.f. \( f_X \) and if the integral \( \int_{-\infty}^{\infty} x f_X(x) \, dx \) is absolutely convergent, then the expected value of \( X \) is defined by

\[
E(X) = \int_{-\infty}^{\infty} x f_X(x) \, dx.
\]

If the integral \( \int_{-\infty}^{\infty} x f_X(x) \, dx \) is not absolutely convergent, then the expected value of \( X \) does not exist.

The definition of the expected value of \( X \) can be extended to functions of \( X \). If \( X \) is a continuous r.v. with p.d.f. \( f_X \), \( g \) is a real valued function for which \( Y = g(X) \) is a continuous r.v. with p.d.f. \( f_Y \), and if \( E(Y) \) exists, then it is easy to see that

\[
E(Y) = \int_{-\infty}^{\infty} y f_Y(y) \, dy = \int_{-\infty}^{\infty} g(x) f_X(x) \, dx = E(g(X)).
\]

We will now consider the variance. Recall that the aim here is to define a measure of variability and, assuming \( E(X) = \mu_X \) exists, that this is done by letting \( g(X) = (X - \mu_X)^2 \) and using the expected value of this function as a measure of variability.

**Definition.** If \( X \) is a continuous r.v. with p.d.f. \( f_X \) and mean \( \mu_X = E(X) \), then the variance of \( X \) is defined by

\[
\text{var}(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) \, dx,
\]

provided this series converges. If the series does not converge, then the variance of \( X \) does not exist. The principal square root of \( \text{var}(X) \) is know as the standard deviation of \( X \).

**Theorem.** If \( a \) is a constant, \( X \) and \( Y \) are continuous r.v.'s, and \( E(X) \) and \( E(Y) \) exist, then

1) \( E(a) = a \).
2) If \( \Pr(X \geq a) = 1 \), then \( E(X) \geq a \). Similarly, if \( \Pr(X \leq a) = 1 \), then \( E(X) \leq a \).
3) \( E(aX) = aE(X) \).
4) \( E(a + X) = a + E(X) \).
5) \( E(X + Y) = E(X) + E(Y) \).
6) if \( \text{var}(X) \) exists, then \( \text{var}(X) = E(X^2) - [E(X)]^2 \).

Proof. The proof, with integrals replacing sums, is completely analogous to that for discrete r.v.'s. □

**Theorem.** If \( a \) is a constant, \( X \) is a continuous r.v., and \( \text{var}(X) \) exists, then

1) \( \text{var}(a) = 0 \).
2) \( \text{var}(aX) = a^2 \text{var}(X) \).
3) \( \text{var}(a + X) = \text{var}(X) \).
4) \( \text{var}(X) = 0 \) if and only if there is a constant \( a \) such that \( \text{Pr}(X = a) = 1 \).

Proof. The proof, with integrals replacing sums, is completely analogous to that for discrete r.v.'s. □

**Covariance.**

The definitions and theorems regarding covariances we provided for discrete r.v.'s are unchanged when we extend them to continuous r.v.'s. We summarize them in this section for ease of reference.

**Definition.** If \( X \) and \( Y \) are discrete or continuous r.v.'s for which \( \text{var}(X) \) and \( \text{var}(Y) \) exist, then the covariance of \( X \) and \( Y \) is

\[
\text{cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X\mu_Y,
\]

where \( \mu_X = E(X) \) and \( \mu_Y = E(Y) \).

**Theorem.** If \( X \) and \( Y \) are independent discrete or continuous r.v.'s for which \( \text{var}(X) \) and \( \text{var}(Y) \) exist, then \( \text{cov}(X, Y) = 0 \).

It is important to note that the converse of this theorem is not true, i.e., in general \( \text{cov}(X, Y) = 0 \) does not imply that \( X \) and \( Y \) are independent.

**Theorem.** If \( X \) and \( Y \) are discrete or continuous r.v.'s for which \( \text{var}(X) \) and \( \text{var}(Y) \) exist, then

\[
\text{var}(X + Y) = \text{var}(X) + 2\text{cov}(X, Y) + \text{var}(Y).
\]

**Theorem.** If \( X_1, \ldots, X_n \) are discrete or continuous r.v.'s with variances \( \sigma_1^2, \ldots, \sigma_n^2 \), then

\[
\text{var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \sigma_i^2 + 2 \sum_{i<j} \text{cov}(X_i, X_j).
\]
Corollary. If $X_1, \ldots, X_n$ are independent discrete or continuous r.v.’s with variances $\sigma_1^2, \ldots, \sigma_n^2$, then

$$\text{var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \sigma_i^2.$$ 

The Schwarz inequality. If $X$ and $Y$ are discrete or continuous r.v.’s for which $E(X^2)$ and $E(Y^2)$ exist, then

$$[E(XY)]^2 \leq E(X^2)E(Y^2)$$

with equality if and only if there is a constant $c$ such that $\Pr(Y = cX) = 1$.

Theorem. If $X$ and $Y$ are discrete or continuous r.v.’s for which $\text{var}(X)$ and $\text{var}(Y)$ exist, then

$$|\text{cov}(X, Y)| \leq \sqrt{\text{var}(X)\text{var}(Y)}.$$ 

If we standardize the r.v.’s $X$ and $Y$ to have mean zero and variance one (by subtracting the mean and dividing by the standard deviation) and then compute the covariance between these standardized r.v.’s we obtain the correlation of $X$ and $Y$

$$\rho(X, Y) = E \left[ \left( \frac{X - \mu_X}{\sigma_X} \right) \left( \frac{Y - \mu_Y}{\sigma_Y} \right) \right].$$

Note that $-1 \leq \rho(X, Y) \leq 1$. The r.v.’s $X$ and $Y$ are said to be uncorrelated when $\rho(X, Y) = 0$, which is equivalent to $\text{cov}(X, Y) = 0$. When $\rho(X, Y) = 1$, $X$ and $Y$ are said to be perfectly positively correlated and when $\rho(X, Y) = -1$, $X$ and $Y$ are said to be perfectly negatively correlated.