Statistics

Expectation and Variance of RVs

Discrete Random Variables:

- **Probability Mass Function (PMF):** Let X be a discrete random variable with possible values $x_1, x_2, x_3, \ldots$ and corresponding probabilities $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, $p_3 = P(X = x_3)$, ... (where $\Sigma p_i = 1$).
- **Expectation (Mean):**
  $E[X] = \mu = \Sigma [x_i * p_i]$ (sum over all possible values of x)
- **Variance:**
  $Var(X) = E[(X - \mu)^2] = \Sigma [(x_i - \mu)^2 * p_i]$
  $Var(X) = E[X^2] - (E[X])^2$ (a computationally useful form)
- **Expectation of a Function of X:**
  $E[g(X)] = \Sigma [g(x_i) * p_i]$

Expectation and Variance of RVs

General Properties of Expectation:

- **Linearity:** $E[aX + bY] = aE[X] + bE[Y]$, where a and b are constants and X and Y are random variables.
- **Constant:** $E[c] = c$, where c is a constant.

General Properties of Variance:

- **Constant:** $Var(c) = 0$, where c is a constant.
- **Scaling:** $Var(aX) = a^2 Var(X)$, where a is a constant.
- **Linear Transformation:** $Var(aX + b) = a^2 Var(X)$, where a and b are constants.
- **Independence:** If X and Y are independent random variables, then $Var(X + Y) = Var(X) + Var(Y)$. (This does *not* generally hold if X and Y are dependent).

Expectation and Variance of RVs

Q1. Let Z be a random variable with the following probability distribution:

$P(Z = -1) = 0.2$ $P(Z = 0) = 0.5$ $P(Z = 1) = 0.3$

Define a new random variable $W = Z^2$.

- Find the expected value of W, E[W].
- Find the variance of W, Var(W).

First, we need to find the probability distribution of W.

- If $Z = -1$, then $W = (-1)^2 = 1$. $P(W = 1) = P(Z = -1) = 0.2$
- If $Z = 0$, then $W = (0)^2 = 0$. $P(W = 0) = P(Z = 0) = 0.5$
- If $Z = 1$, then $W = (1)^2 = 1$. $P(W = 1) = P(Z = 1) = 0.3$

Notice that W can only take the values 0 and 1. The probability distribution of W is:

- $P(W = 0) = 0.5$
- $P(W = 1) = 0.2 + 0.3 = 0.5$

Now we can calculate E[W]:

$$
\begin{aligned}
E[W] &= \Sigma\, [w * P(W = w)] \\
&= (0 * 0.5) + (1 * 0.5) \\
&= 0 + 0.5 = 0.5
\end{aligned}
$$

$Var(W) = E[W^2] - (E[W])^2$

Since W can only be 0 or 1, W^2 will also only be 0 or 1. In fact, $W^2 = W$ in this case. This is because $0^2=0$ and $1^2=1$. So, $E[W^2] = E[W] = 0.5$

$$
\begin{aligned}
Var(W) &= E[W^2] - (E[W])^2 \\
&= 0.5 - (0.5)^2 \\
&= 0.5 - 0.25 \\
&= 0.25
\end{aligned}
$$

Expectation and Variance of RVs

Q2. Let X be a random variable with E[X] = 5 and Var(X) = 2. Let Y = 3X - 4.

- Find E[Y].
- Find Var(Y).

We can use the linearity of expectation, which states that $E[aX + b] = aE[X] + b$, where 'a' and 'b' are constants.

$E[Y] = E[3X - 4]$

$= 3E[X] - 4$ (using linearity of expectation)

$= 3(5) - 4$ (substituting $E[X] = 5$)

$= 15 - 4$

$= 11$

We can use the property of variance that states $Var(aX + b) = a^2 Var(X)$, where 'a' and 'b' are constants. Notice that the constant term 'b' does not affect the variance.

$Var(Y) = Var(3X - 4)$

$= 3^2 Var(X)$ (using the property of variance)

$= 9(2)$ (substituting $Var(X) = 2$)

$= 18$

Co-Variance of RVs

Q3. Let U and V be two independent standard normal random variables, i.e., U,V ~ N(0,1). Define the new random variables:

$R = 5 + 2U - 3UV$

$S = 2 - U + V$

Find cov(R, S)

The covariance between two random variables R and S is defined as:

$\mathrm{cov}(R, S) = E[(R - E[R])(S - E[S])] = E[RS] - E[R]E[S]$

Co-Variance of RVs

Covariance:

- **Definition:** $Cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$
- **Relationship to Variance:** $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$ (This holds in general, whether or not X and Y are independent.)
- **Independence:** If X and Y are independent, $Cov(X, Y) = 0$. (The converse is not necessarily true.)

Standard Deviation:

- The standard deviation of X, denoted σ or SD(X), is the square root of the variance: $\sigma = \sqrt{Var(X)}$. It provides a measure of the spread of the distribution in the original units of the random variable.

Co-Variance of RVs

Q1. Let U and V be two independent standard normal random variables, i.e., U,V ~ N(0,1). Define the new random variables: $R = 5 + 2U - 3UV$ and $S = 2 - U + V$. Find cov(R, S)

The covariance between two random variables R and S is defined as:

$\text{cov}(R, S) = E[(R - E[R])(S - E[S])] = E[RS] - E[R]E[S]$

First, let's find the expected values of R and S:

- $E[R] = E[5 + 2U - 3UV] = 5 + 2E[U] - 3E[UV]$

Since U and V are independent, $E[UV] = E[U]E[V]$. Also, $E[U] = E[V] = 0$, as U and V are standard normal.

Therefore, $E[R] = 5 + 2(0) - 3(0)(0) = 5$

- $E[S] = E[2 - U + V] = 2 - E[U] + E[V] = 2 - 0 + 0 = 2$

Now, let's find E[RS]:

$E[RS] = E[(5 + 2U - 3UV)(2 - U + V)] = E[10 - 5U + 5V + 4U - 2U^2 + 2UV - 6UV + 3U^2V - 3UV^2]$

$\quad = 10 - 5E[U] + 5E[V] + 4E[U] - 2E[U^2] + 2E[UV] - 6E[UV] + 3E[U^2V] - 3E[UV^2]$

Since U and V are standard normal, $E[U] = E[V] = 0$ and $E[U^2] = E[V^2] = 1$.

Also, since U and V are independent, $E[UV] = E[U]E[V] = 0$, $E[U^2V] = E[U^2]E[V] = 1 * 0 = 0$, and $E[UV^2] = E[U]E[V^2] = 0 * 1 = 0$.

Therefore, $E[RS] = 10 - 5(0) + 5(0) + 4(0) - 2(1) + 2(0) - 6(0) + 3(0) - 3(0) = 10 - 2 = 8$

Finally, we can find the covariance: $\text{cov}(R, S) = E[RS] - E[R]E[S] = 8 - (5)(2) = 8 - 10 = -2$