CSL003P1M : Probability and Statistics
Lecture 12 (Variance Of A Discrete Random Variable)

Sumit Kumar Pandey

September 29, 2021

Expectation of a Function of a Random Variable

- Suppose that we are given a discrete random variable along with its probability mass function and that we want to compute the expected value of some function of X, say, $g(X)$. How can we accomplish this?
- One way is as follows: Since $g(X)$ is itself a discrete random variable, it has a probability mass function, which can be determined from the probability mass function of X.
- Once we have determined the probability mass function of $g(X)$, we can compute $E[g(X)]$ by using the definition of expected value.

Exercise-1

Let X denote a random variable that takes on any of the values $-1, 0$ and 1 with respective probabilities

$$P\{X = -1\} = 0.2, \quad P\{X = 0\} = 0.5, \quad P\{X = 1\} = 0.3$$

Compute $E[X^2]$.

Solution:

- Let $Y = X^2$. Then the probability mass function of Y is given by

$$P\{Y = 1\} = P\{X = -1\} + P\{X = 1\} = 0.5$$
$$P\{Y = 0\} = P\{X = 0\} = 0.5$$

- Hence,

$$E[X^2] = E[Y] = 1 \times (0.5) + 0 \times (0.5) = 0.5.$$

Exercise-1

- $$E[X] = (-1) \times (0.2) + 0 \times (0.5) + 1 \times (0.3) = 0.1$$

Note that
$$0.5 = E[X^2] \neq (E[X])^2 = 0.01.$$

Expectation of a Function of a Random Variable

Proposition

If X is a random variable that takes on one of the values x_i, $i \geq 1$, with respective probabilities $p(x_i)$, then, for any real-valued function g,

$$E[g(X)] = \sum_i g(x_i)p(x_i)$$

Let's check by solving Exercise-1,

- Given,

$$P\{X = -1\} = 0.2, \quad P\{X = 0\} = 0.5, \quad P\{X = 1\} = 0.3$$

-

$$\begin{aligned} E[X^2] &= (-1)^2 \times (0.2) + 0^2 \times (0.5) + 1^2 \times (0.3) \\ &= 1 \times (0.2 + 0.3) + 0 \times (0.5) \\ &= 0.5 \end{aligned}$$

Proof of the Proposition:

- The idea is that the proof proceeds by grouping together all the terms in $\sum_i g(x_i)p(x_i)$ having the same value of $g(x_i)$.
- Suppose $y_j, j \geq 1$, represent the different values of $g(x_i)$, $i \geq 1$.
- Then, grouping all the $g(x_i)$ having the same value gives

$$\begin{aligned}
\sum_i g(x_i)p(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \\
&= \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i) \\
&= \sum_j y_j \sum_{i:g(x_i)=y_j} p(x_i) \\
&= \sum_j y_j P\{g(X) = y_j\} \\
&= E[g(X)]
\end{aligned}$$

Corollary

If a and b are constants, then

$$E[aX+b] = aE[X]+b.$$

Proof:

- $$\begin{aligned} E[aX+b] &= \sum_{x:p(x)>0} (ax+b)p(x) \\ &= a\sum_{x:p(x)>0} xp(x) + b\sum_{x:p(x)>0} p(x) \\ &= aE[X]+b \end{aligned}$$

Moments of a Random Variable

- The expected value of a random variable X, $E[X]$, is also referred to as the **mean** or the **first moment** of X.

nth Moment of X

The quantity $E[X^n]$, $n \geq 1$ is called the nth moment of X. Note that,

$$E[X^n] = \sum_{x:p(x)>0} x^n p(x).$$

Example

Consider the following score:

Batsman	I-match	II-match	Average
Sachin	0	100	50
Dravid	40	60	50

What's your opinion about the consistency of these batsmen?

- The spread of runs of Sachin is higher than that of Dravid.

Variance of a Random Variable

Variance of a Random Variable

If X is a random variable with mean μ, then the variance of X, denoted by $Var(X)$, is defined by

$$Var(X) = E[(X - \mu)^2]$$

Alternate formula,

Variance of a Random Variable

If X is a random variable with mean μ, then the variance of X, denoted by $Var(X)$, is

$$Var(X) = E[X^2] - \mu^2 = E[X^2] - (E[X])^2$$

Proof: ???

Variance of a Random Variable

Proof:

- $$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= \sum_x (x-\mu)^2 p(x) \\ &= \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\ &= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\ &= E[X^2] - 2\mu^2 + \mu^2 \\ &= E[X^2] - \mu^2 \\ &= E[X^2] - (E[X])^2 \end{aligned}$$

Exercise-2

Calculate $Var(X)$ if X represents the outcome when a fair die is rolled.

Solution:

- $E[X] = 7/2$ (Try!!!).
-
$$\begin{aligned} E[X^2] &= 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right) + 5^2\left(\frac{1}{6}\right) + 6^2\left(\frac{1}{6}\right) \\ &= 91\left(\frac{1}{6}\right). \end{aligned}$$

- Thus,
$$Var(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

Exercise-3

Prove that for any constants a and b, if X is a discrete random variable, then

$$Var(aX + b) = a^2 Var(X).$$

Proof:

- Let $E[X] = \mu$ and $E[aX + b] = aE[X] + b = a\mu + b$
- $$\begin{aligned} Var(aX + b) &= E[\{(aX + b) - (a\mu + b)\}^2] \\ &= E[a^2(X - \mu)^2] \\ &= a^2 E[(X - \mu)^2] \\ &= a^2 Var(X) \end{aligned}$$

Standard Deviation of a Random Variable

Standard Deviation of a Random Variable

The square root of the $Var(X)$ is called the *standard deviation* of X, and we denote it by $SD(X)$. That is,

$$SD(X) = \sqrt{Var(X)}.$$

Exercise-4

Consider the following sequence of numbers:
9, 9, 1, 2, 3, 4, 2, 9, 8, 4, 1, 5, 7, 3, 0, 4, 0, 2, 1, 5, 3, 9, 0, 4, 3,
8, 6, 6, 4, 2, 0, 1, 1, 8, 3, 6, 5, 4, 8, 2, 9, 1, 7, 2, 8, 9, 1, 1, 9, 6
4, 5, 6, 2, 0, 1, 4, 2, 9, 8, 4, 5, 1, 3, 9, 1, 0, 7, 7, 2, 8, 3, 9, 1, 1,
2, 2, 6, 7, 9, 0, 7, 3, 8, 2, 9, 0, 1, 1, 6, 3, 8, 5, 6, 3, 2, 9, 3, 2, 1,

Calcuate the mean and the standard deviation.

Let f_i denote the frequency of number i and n be the total number in the above sequence. Here, $n = 100$.

Thank You