Expectation and variance

Math 477

October 2, 2014

1 Expectation

1.1 Definition

Definition 1.1. *Let X be a discrete random variable. The expectation of X, denoted as $E(X)$ is defined as*

$$E(X) = \sum_k kP(X = k).$$

Remark: $E(X)$ is the average value of X. It is a constant. It has the interpretation that if we see many independent realizations of X, then the average of these realizations will be close to $E(X)$. This is the so called Law of Large Number, which we will discuss later.

Example 1.2. *Let X be a RV with distribution $P(X = 3) = 1/2$ and $P(X = -2) = 1/2$. Then $E(X) = 3/2 - 1 = 1/2$.*

Example 1.3. *Let X be the RV representing the result of a die toss. Then*

$$E(X) = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5.$$

Example 1.4. *Let X be the RV representing the sum of the 2 independent dice. Then*

$$E(X) = \frac{(2 + 12) + (3 + 11)2 + (4 + 10)3 + (5 + 9)4 + (6 + 8)5 + (7)6}{36} = 7.$$

1.2 Expectation of a function of X

Theorem 1.5. *Let $g : \mathbb{R} \to \mathbb{R}$ and X a random variable. Then*

$$E(g(X)) = \sum_k g(k)P(X = k).$$

Note: The above is NOT a definition. Indeed by definition we have

$$E(g(X)) = \sum_k kP(g(X) = k).$$

Converting from the definition to the statement in the theorem is what we need to show. *Proof.* From the discussion of function of random variables, we already have

$$P(g(X) = k) = \sum_{j:g(j)=k} P(X = j).$$

Therefore,

$$\begin{aligned} E(g(X)) &= \sum_k kP(g(X) = k) = \sum_k k \sum_{j:g(j)=k} P(X = j) \\ &= \sum_k \sum_{j:g(j)=k} g(j)P(X = j) \\ &= \sum_k g(k)P(X = k). \end{aligned}$$

Example 1.6. *Let X be the result of a die toss. Then*

$$E(X^2) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = 91/6.$$

Example 1.7. *Let X be a RV with distribution $P(X = 1) = 1/3, P(X = -1) = 1/6, P(X = 2) = 1/2$. Then*

$$E(X^3) = 1/3 - 1/6 + 8/2 = 25/6.$$

Note: It is NOT true that $E(g(X)) = g(E(X))$ unless g is a linear function of X (that is $g(x) = ax+c$, see below). For example for the die toss example, $E(X^2) = 91/6$ is NOT equal to $(E(X))^2 = 49/4$.

1.3 Properties

For any RVs X, Y and constants a, c we have

$$E(aX + c) = aE(X) + c.$$

In particular, $E(c) = c$ for any constant c.

Proof. We have

$$E(aX) = \sum_k (ak + c)P(X = k) = a \sum_k P(X = k) + c \sum_{P(X=k)} = aE(X) + c.$$

Example 1.8. *Let X be the result of a die toss. Then*

$$\begin{aligned} E(2X) &= 2E(X) = 7 \\ E(6X^2) &= 6E(X^2) = 91. \end{aligned}$$

2 Variance

2.1 Definition

Definition 2.1. *Let X be a random variable. The variance of X, denoted as $Var(X)$ is defined as*

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

Remark: $Var(X)$ is a *non negative* number, and a positive number if X is not a constant. It measures the "spreadness" of X: the larger the variance, the more spread out the random variable is (in terms of its value).

There is an alternative version for the $Var(X)$ as followed.

Lemma 2.2. *Let X be a random variable. Then $Var(X) = E(X^2) - E^2(X)$.*

Proof.

$$\begin{aligned} Var(X) &= E[(X - E(X))^2] = E(X^2 - 2XE(X) + E^2(X)) \\ &= E(X^2) - 2E(X)E(X) + E^2(X) \\ &= E(X^2) - E^2(X). \end{aligned}$$

Example 2.3. *Let X be the random variable representing the result of the die toss. Then*

$$Var(X) = 91/6 - 49/4.$$

Remark: Because $Var(X) \geq 0$, we can also deduce that for any random variable X: $E(X^2) \geq E^2(X)$.

2.2 Properties

Let X be a RV and a, c constants. Then

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

In particular, $Var(c) = 0$ for a constant c. Note that the constant a factors out of Var as a^2, NOT a.

Proof.

$$Var(aX + c) = E([aX + c - E(aX + c)]^2) = E(a^2(X - E(X))^2) = a^2 Var(X).$$