Expected value of discrete random variables

Example 1 *The expected value of the degenerate distribution is* $E(X) = c \cdot P(X = c) = c$.

Example 2 *The expected value of a Bernoulli random variable* X *is* $E(X) = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = p$. *Remember that* X *is the indicator of event* A, *which has a probability of* p, *and* X *equals to 1, if* A *occurs, and 0 if not.*

Example 3 *Let* X *be a random variable, which has a uniform distribution on numbers* $x_1, \ldots, x_n$, *i.e., they have the same probability* $\frac{1}{n}$. $E(X) = \frac{1}{n} \cdot \sum\limits_{i=1}^{n} x_i$, *the mean value.*

Example 4 *If* $X \sim$ *Binomial*(n, p), *the expected value is the product of parameters:*

$$E(X) = \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} np \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} = np.$$

Example 5 *If* $X \sim$ *Poisson*(λ)*:*

$$E(X) = \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda.$$

Remark 1 *Not all of the random variables have finite expected value. For example, if* $P(X = 2^k) = \frac{1}{2^k}$ $(k = 1, 2, \ldots)$, $E(X) = 1 + 1 + 1 + \ldots$. *The St. Petersburg paradox: a casino offers a game in which a single fair coin is tossed at each stage. The pot starts at 1 dollar, and doubles every time a head appears. The first time a tail appears, the game ends and the player wins the pot. It means, if the first tail appears on the* k*th toss, the player wins* 2^{k-1} *dollars. On the other hand, the probability of winning exactly* 2^{k-1} *dollars is* $\frac{1}{2^k}$. *Thus the expected value of the pot is* $E(X) = \sum\limits_{k=1}^{\infty} \frac{1}{2^k} 2^{k-1} = \infty$.

The variance of random variables

The following definition is designed to measure the fluctuation of random quantities, or how far a set of numbers is spread out.

Definition 1 *The variance of a random variable is the expected value of squared difference between the variable and the variable's expected value, i.e.,* $Var(X) = E\left((X - E(X))^2\right)$. *Sometimes* $Var(X)$ *is denoted by* $D^2(X)$.

Remark 2 $Var(X) = E(X^2 - 2XE(X) + E^2(X)) = E(X^2) - 2E(X)E(X) + E^2(X)$. *Thus* $Var(X) = E(X^2) - E^2(X)$, *and in practice we use this formula.*

Properties:

1. $Var(X) \geq 0$, because it is the expected value of a nonnegative random variable.
2. Let a, b be real numbers. $Var(aX + b) = a^2 Var(X)$, because $Var(aX + b) = E((aX + b) - E(aX + b))^2) = E((aX + b - aE(X) - b)^2) = a^2 E\left((X - E(X))^2\right)$.
3. If $E(X) < \infty$, it doesn't necessarily follows that $Var(X) < \infty$. As an example, let X be a discrete random variable with distribution $P(X = k) = \frac{c}{k^3}$ (c can be easily calculated). In this case, $E(X)$ is finite, but $E(X^2) = c(1 + \frac{1}{2} + \ldots + \frac{1}{k} + \ldots)$, which is infinite.

Example 6 *The variance of degenerate distribution is* $Var(X) = E(X^2) - E^2(X) = c^2 - c^2 = 0$. *Conversely, if the variance of a random variable is 0, the variable is constant with probability 1, because* $(X - E(X))^2 \geq 0$, *and it's expected value is 0, thus* $(X - E(X)) = 0$ *with probability 1.*

Example 7 *The variance of a Bernoulli*(p) *variable is* $Var(X) = p - p^2 = p(1 - p)$.

Example 8 *The variance of dice throw:* $Var(X) = E(X^2) - E^2(X) = \frac{1+4+\ldots+36}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$.

Example 9 *Let X be a Poisson(λ) random variable.* $E(X^2) = \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} (k-1+1) \frac{\lambda^{k-1}}{(k-1)!} = \lambda^2 + \lambda$, *and* $Var(X) = E(X^2) - E^2(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.

Definition 2 *The standard deviation of a random variable is the squareroot of the variance, i.e.,* $D(X) = \sqrt{Var(X)}$.

The variance of sum of two random variables is $Var(X+Y) = E((X+Y-E(X+Y))^2) = E((X-E(X)+Y-E(Y))^2) = E((X-E(X))^2) + E((Y-E(Y))^2) + 2E((X-E(X))(Y-E(Y))) = Var(X) + Var(Y) + 2E((X-E(X))(Y-E(Y)))$. For example, if $X = Y$, $Var(X+Y) = 4Var(X)$.

Proposition 1 *Suppose that X and Y are independent discrete random variables, and $E(X)$, $E(Y)$ are finite values. Then $E(XY) = E(X)E(Y)$.*

PROOF: Let the range of X and Y be $(x_1, x_2, \ldots)$ and $(y_1, y_2, \ldots)$. $E(XY) = \sum_{k,l} x_k y_m P(X = x_k, Y = y_l)$. Because of independence, this is equal to $\sum_k x_k P(X = x_k) \sum_l y_l P(Y = y_l) = E(X)E(Y)$. □

Remark 3 *If X and Y are independent, $Var(X+Y) = Var(X) + Var(Y)$. Moreover, it can be proved by induction that for pairwise independent random variables $X_1, \ldots, X_n$, the variance of the sum equals to the sum of variances: $Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n)$. It follows that the variance of a binomial random variable is $np(1-p)$, i.e., the sum of n Bernoulli variances.*

Distribution functions

Definition 3 (distribution function) *Let $F_X(z) : \mathbb{R} \to \mathbb{R}$ denote the* ***(cumulative) distribution function*** *of random variable X, defined as $F_X(z) := P(X < z)$.*

Proposition 2 *The following properties hold for distribution functions.*

1. $0 \leq F_X(z) \leq 1$
2. *monotonically increasing (non-decreasing)*
3. $\lim_{z \to -\infty} F_X(z) = 0$ *and* $\lim_{z \to \infty} F_X(z) = 1$
4. *left-continuous and the right hand limits exist*

Example 10 *The distribution function of degenerate distribution is* $F(z) = \begin{cases} 0 & if\ z \leq c \\ 1 & if\ z > c \end{cases}$. *(Remember: in this case $P(X = c) = 1$.)*

Example 11 *The d.f. of a Bernoulli random variable is* $F(z) = \begin{cases} 0 & z \leq 0 \\ 1-p & 0 < z \leq 1 \\ 1 & z > 1 \end{cases}$.

Definition 4 *If the distribution function of a random variable X is continuous, the variable is called* ***continuous****.*

Remark 4 *The range of continuous random variables is uncountably infinite.*

Example 12 *Uniform distribution on interval $[a, b]$.* $F(z) = \begin{cases} 0 & z \leq a \\ \frac{z-a}{b-a} & a < z \leq b \\ 1 & z > b \end{cases}$. *Notation: $X \sim U(a, b)$.*

Example 13 *The exponential distribution with parameter $\lambda > 0$ is defined by d.f.* $F(z) = \begin{cases} 0 & z \leq 0 \\ 1 - e^{-\lambda z} & 0 < z \end{cases}$.

Remark 5 $P(a \leq X < b) = F(b) - F(a)$, $P(a < X < b) = F(b) - F(a+0)$, $P(a \leq X \leq b) = F(b+0) - F(a)$, $P(X = a) = F(a+0) - F(a)$ *(last equality means, if F is continuous, each point has probability 0).*