Random Variables, Expectation, Variance 8

Computer Science Mentors 70

October 22 - 26, 2018

1 Expectation of Random Variables

1.1 Introduction

Random variable: a function $X : \omega \to R$ that assigns a real number to every outcome ω in the probability space.

Expectation: The expectation of a random variable X is defined as

$$\mathrm{E}(X) = \sum_{\alpha \in A} a * \mathrm{P}[X = a]$$

where the sum is over all possible values taken by the random variable. Expectation is usually denoted with the symbol μ.

Linearity of Expectation: For any random variables $X_1, X_2, ...X_n$, expectation is linear, i.e.:

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

This is true even when these random variables aren't independent.

1.2 Questions

1. You win the lottery with probability .01. If you win, you get $10000. You lose with probability .99 and get no money. Define a random variable to represent this.

Solution: We define a random variable X that has value 10000 with probability .01 (the event that you win the lottery) and value 0 with probability .99 (the event that you do not).

2. Given the random variable X defined as taking on the value 1 with probability 0.25, 2 with probability 0.5, and 20 with probability 0.25, what is the expectation of X?

Solution: $\mathrm{E}(X) = 0.25 * 1 + 0.5 * 2 + 0.25 * 20 = 6.25$

3. Show that $E(aX + b) = aE[X] + b$ where X is any random variable.

Solution: First, for the expectation:

$$\mathrm{E}[aX + b] = \sum_{x \in X} (aX + b)P(X = x) = a \sum_{x \in X} X \cdot P(X = x) + \sum_{x \in X} bP(X = x)$$

$$= a\mathrm{E}[X] + b \sum_{x \in X} P(X = x) = a\mathrm{E}[X] + b$$

.

4. Suppose X is a random variable. Does X always have to take on the value $E(X)$ at some point?

Solution: No. Consider X with uniform probability space $\{0, 1\}$. The expectation of X is $\frac{1}{2}$, but X never takes on the value $\frac{1}{2}$.

5. An urn contains n balls numbered $1, 2, \ldots, n$. We remove k balls at random (without replacement) and add up their numbers. Find the mean of the total.

Solution: The required total is $T = \sum_{i=1}^{k} X_i$, where X_i is the number shown on the ith ball. Hence $\mathrm{E}(T) = k * \mathrm{E}(X_1) = \frac{1}{2} * k * (n + 1)$.

6. A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence ”book” appears?

Solution: There are $1000000 - 4 + 1 = 999997$ places where "book" can appear, each with a (nonindependent) probability of $\frac{1}{26}^4$ of happening. If A is the random variable that tells how many times "book" appears, and A_i is the indicator variable that is 1 if "book" appears starting at the ith letter, then

$$E[A] = E[A_1 + ... + A_9 99997] = E[A_1] + ... + E[A_9 99997]$$
$$= \frac{999997}{26^4}$$

7. In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $\frac{1}{3}$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $\frac{1}{5}$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

Solution: Let A_i be the indicator you win the ith time you play game A and B_i be the same for game B. The expected value of A_i and B_i are:

$$E[A_i] = 1 * \frac{1}{3} + 0 * \frac{2}{3} = \frac{1}{3}$$
$$E[B_i] = 1 * \frac{1}{5} + 0 * \frac{4}{5} = \frac{1}{5}$$

Let T_A be the random variable for the number of tickets you win in game A, and T_B be the number of tickets you win in game B.

$$E[T_A + T_B] = 3E[A_1] + ... + 3E[A_{10}] + 4E[B_1] + ... + 4E[B_{20}]$$
$$= 10 * (3 * \frac{1}{3}) + 20 * (4 * \frac{1}{5}) = 26$$

2 Variance

2.1 Introduction

Variance: The variance of a random variable X is defined as

$$\text{Var}(X) = \text{E}((X - \text{E}(X))^2) \ = \text{E}(X^2) - \text{E}(X)^2$$

The latter version of variance is the one we usually use in computations.

The square root of Var(X) is called the standard deviation of X. It is usually denoted with the variable σ.
Important property of variance: for some constant c,

$$\text{Var}(cX) = c^2 * \text{Var}(X)$$

.

2.2 Questions

1. **True or False?**
   Assume that X is a discrete random variable. If $Var(X) = 0$, then X is a constant.

   Solution: TRUE. Let $\mu = \text{E}[X]$. By definition,

   $$0 = \text{Var}(X) = \text{E}[(X - \mu)^2] = \sum \omega \text{P}[\omega](X(\omega) - \mu)^2$$

   The RHS is the sum of non-negative numbers, so if the sum is 0, each term must be 0. So

   $$\text{P}[\omega] > 0 \to (X(\omega) - \mu)^2 = 0 = X(\omega) = \mu.$$

   Therefore X is constant (equal to $\mu = \text{E}[X]$).

2. Show that $\text{Var}(X) = \text{E}((X - \mu)^2) = \text{E}(X^2) - \mu^2$ where X is any random variable and $\mu = \text{E}[X]$.

   Solution:

   $$\text{Var}(X) = \text{E}((X - \mu^2) = \text{E}[X^2 - 2\mu X + \mu^2]$$
   $$= \text{E}[X^2] - 2\mu\text{E}[X] + \mu^2 \text{ (due to linearity of expectation)}$$
   $$= \text{E}[X^2] - \mu^2$$

3. Show that $Var(aX + b) = a^2 Var(X)$ where X is any random variable.

Solution:

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

4. Let's consider the classic problems of flipping coins and rolling dice. Let X be a random variable for the number of coins that land on heads and Y be the value of the die roll.

   (a) What is the expected value of X after flipping 3 coins? What is the variance of X?

   Solution:

   $$\mathrm{E}(X) = 0*\frac{1}{8}+1*\frac{3}{8}+2*\frac{3}{8}+3*\frac{1}{8}=\frac{3}{2}$$
   $$\mathrm{E}(X^2) = 0^2*\frac{1}{8}+1^2*\frac{3}{8}+2^2*\frac{3}{8}+3^2*\frac{1}{8}=\frac{24}{8}=3$$
   $$\mathrm{E}(X)^2=\frac{9}{4}$$
   $$\text{Var}(X)=3-\frac{9}{4}=\frac{3}{4}$$

   (b) Let Y be the sum of rolling a dice 1 time. What is the expected value of Y?

   Solution: $\mathrm{E}(Y)=\frac{1}{6}*(1+2+3+4+5+6)=\frac{7}{2}$

   (c) What is the variance of Y?

   Solution: $\mathrm{E}(Y^2)=[\frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2)]=\frac{91}{6}$ $\text{Var}(Y)=\mathrm{E}(Y^2)-(\mathrm{E}(Y))^2=\frac{91}{6}\frac{7}{2}^2=\frac{35}{12}$

5. You are at a party with n people where you have prepared a red solo cup labeled with their name. Before handing red cups to your friends, you pick up each cup and put a sticker on it with probability $\frac{1}{2}$ (independently of the other cups). Then you hand back the cups according to a uniformly random permutation. Let X be the number of people who get their own cup back AND it has a sticker on it.

(a) Compute the expectation $\mathrm{E}(X)$.

Solution: Define $X_i = 1$ if the i-th person gets their own cup back and it has a sticker on it and $X_i = 0$ otherwise. Hence $\mathrm{E}(X) = \mathrm{E}(\sum_{i=1}^{n}(X_i)) = \sum_{i=1}^{n}\mathrm{E}(X_i)$ $\mathrm{E}(X_i) = \mathrm{P}[X_i = 1] = \frac{1}{2n}$ since the i-th student will get his/her cup with probability $\frac{1}{n}$ and has a sticker on it with probability $\frac{1}{2}$ and stickers are put independently. Hence $\mathrm{E}(X) = n \cdot \frac{1}{2n} = \frac{1}{2}$.

(b) Compute the variance $\mathrm{Var}(X)$

Solution: To calculate $\mathrm{Var}(X)$, we need to know $\mathrm{E}(X^2)$

$$\mathrm{E}(X^2) = \mathrm{E}(X_1 + X_2 + \ldots + X_n)^2 = \mathrm{E}(\sum_{i,j}(X_i * X_j)) = \sum_{i,j}(\mathrm{E}(X_i * X_j)$$

(by linearity of expectation)

Then we consider two cases, either $i = j$ or $i \neq j$. Hence

$$\sum_{i,j}\mathrm{E}(X_i * X_j) = \sum_{i}\mathrm{E}(X_i^2) + \sum_{i \neq j}\mathrm{E}(X_i * X_j)$$

$E(X_i^2) = \frac{1}{2n}$ for all i. To find $\mathrm{E}(X_i * X_j)$, we need to calculate $\mathrm{P}[X_iX_j = 1]$. $\mathrm{P}[X_i * X_j = 1] = \mathrm{P}[X_i = 1]\mathrm{P}[X_j = 1|X_i = 1] = \frac{1}{2n} * \frac{1}{2*(n-1)}$ since if student i has received his/her own cup, student j has $n-1$ choices left. Hence

$$\mathrm{E}(X^2) = n * \frac{1}{2n} + n * (n-1) * \frac{1}{2n} * \frac{1}{2 * (n-1)} = \frac{3}{4}$$

$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$.