Lecture 8

Andrei Antonenko

Feb 09, 2005

1 More on expectation and variance

Let's compute the expectation and variance of several other random variables.

Consider a toss of a fair die. Let the random variable X be a number obtained. The possible values of X are $1, 2, 3, 4, 5$, and 6, and each of them occurs with the probability $1/6$. Such random variables which take integer values from 1 to n each with probability $1/n$ (or, in general, all integer number from a to b each with probability $1/(b-a+1)$) are called **uniform discrete random variables**. Let's compute the expectation of the discrete uniform random variable.

$$\mathbf{E}[X] = \sum_{i=1}^{n} \frac{1}{n} \cdot i = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}. \tag{1}$$

For example, in case of a fair die there are 6 choices, thus the expectation is $\frac{6+1}{2} = 3/5$. By the similar arguments, the expectation of discrete uniform random variable taking values $a \le X \le b$

$$\mathbf{E}[X] = \frac{b+a}{2}.$$

Now we can compute the variance of X. In case it takes values from 1 to n, we have:

$$\mathbf{E}\left[X^2\right] = \sum_{i=1}^{n} \frac{1}{n} i^2 = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}; \tag{2}$$

$$\mathbf{var}(X) = \mathbf{E}\left[X^2\right] - (\mathbf{E}[X])^2 = \frac{1}{6}(n+1)(2n+1) - \frac{1}{4}(n+1)^2 = \frac{n^2-1}{12}. \tag{3}$$

Since the general discrete uniform random variable is equal to the random variable from 1 to $b-a$, plus a, we have

$$\mathbf{var}(X) = \frac{(b-a)(b-a+2)}{12}. \tag{4}$$

Now let's compute the expectation of the Poisson random variable. It's PMF is $p_X(k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$. Therefore

$$\mathbf{E}[X] = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \underbrace{\sum_{m=0}^{\infty} e^{-\lambda} \frac{\lambda^m}{m!}}_{=1} = \lambda, \tag{5}$$

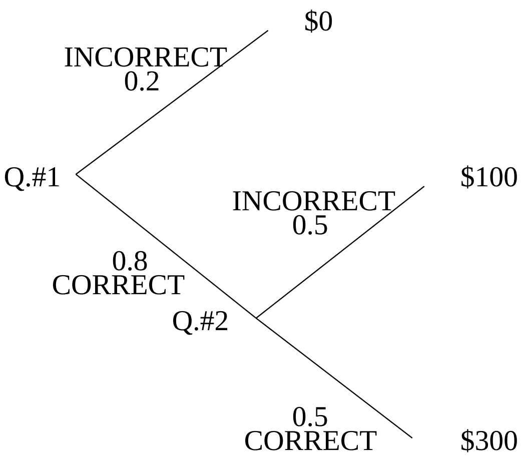
where the last equality is true, since the sum is equal to 1 because it is the sum of all probabilities of Poisson random variable. It is also possible to compute the variance of the Poisson random variable:

$$\mathbf{var}\,(X) = \lambda. \tag{6}$$

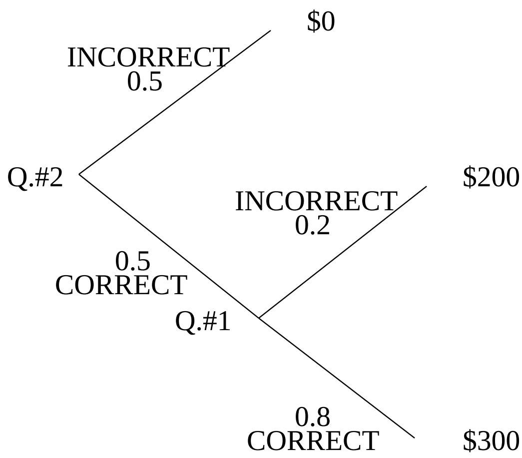
Let's consider several problems now.

Example 1.1 (Quiz Problem). Assume one is supposed to take a quiz with two questions. The probability of answering the Question # 1 correctly is 0.8, and a correct answer leads to a prize of \$100. Question # 2 is answered correctly with probability of 0.5, and the prize for a correct answer is \$500. A person has a choice to attempt the Q.# 1 first, and than if it is answered correctly, the Q.# 2 is asked. Anoher option is to attempt Q.# 2 first, and if it is answered correctly, the Q.# 1 is asked. What choice of questions is better to maximize the expected prize?

Schematically, there are two following scenarios:



Question #1 is attempted first



Question #2 is attempted first

In the first case, we have:

$$\begin{aligned} p_X(0) &= 0.2 \\ p_X(100) &= 0.8 \cdot 0.5 \\ p_X(300) &= 0.8 \cdot 0.5 \\ \mathbf{E}\,[X] &= \$160 \end{aligned}$$

and in the second case,

$$\begin{aligned} p_X(0) &= 0.5 \\ p_X(200) &= 0.5 \cdot 0.2 \\ p_X(300) &= 0.5 \cdot 0.58 \\ \mathbf{E}\,[X] &= \$140 \end{aligned}$$

Therefore, it is better to attempt Question # 1 first.

Let's generalize the situation. Assume p_1 and p_2 are probabilities of answering questions # 1 and # 2 correctly, and v_1 and v_2 are values of prizes for correct answers. In case the Question # 1 is attempted first, we have:

$$\mathbf{E}[X] = p_1(1-p_2)v_1 + p_1p_2(v_1+v_2) = p_1v_1 + p_1p_2v_2.$$

In case the Question # 2 is attempted first, we have:

$$\mathbf{E}[X] = p_2(1-p_1)v_2 + p_2p_1(v_2+v_1) = p_2v_2 + p_1p_2v_1.$$

First strategy is preferred if

$$p_1v_1 + p_1p_2v_2 \geq p_2v_2 + p_1p_2v_1,$$

or

$$\frac{p_1v_1}{1-p_1} \geq \frac{p_2v_2}{1-p_2}. \tag{7}$$

Therefore, the questions should be ordered by decreasing values of $\frac{pv}{1-p}$. This general result can be generalized to the case of more than 2 questions.

Our next example will show that $\mathbf{E}[g(X)] \neq g(\mathbf{E}[X])$.

Example 1.2. If the weather is good (which happens with probability 0.6), Alison walks to school for 2 miles with the speed $V = 5$ mph. If the weather is bad, Alison takes her motorcycle to school and drives 2 miles to school with the speed $V = 30$ mph. What is the expected time to get to school?

Let T be the time to get to school. T is equal to $2/5$ with probability 0.6, and $2/30$ with probability 0.4. Therefore, $\mathbf{E}[T] = 0.6 \cdot (2/5) + 0.4 \cdot (2/30) = (4/15)$ hours.

However, computing the average speed, and dividing 2 mph by it is wrong: $\mathbf{E}[V] = 0.6 \cdot 5 + 0.4 \cdot 30 = 15$, and $\frac{2}{\mathbf{E}[V]} = \frac{2}{15}$. So, we see, that

$$\mathbf{E}[T] = \mathbf{E}\left[\frac{2}{V}\right] \neq \frac{2}{\mathbf{E}[V]}. \tag{8}$$

2 Joint PMF of multiple random variables

Often in the experiment we can observe several numerical characteristics, possible dependent on each other, but different. In this chapter we will deal with such cases and develop machinery to work with such situation.

Assume X and Y are two discrete random variables. We can define their **joint probability mass function** as

$$p_{X,Y}(x,y) = P(\{X=x\} \cap \{Y=y\}) = P(X=x, Y=y), \tag{9}$$

i.e. $p_{X,Y}(x,y)$ is a probability that variable X takes value x and variable Y takes value y. If we want to consider event A, which consists of all pairs (x,y) with certain properties, we can compute the probability of the event as

$$P(A) = \sum_{(x,y)\in A} p_{X,Y}(x,y). \tag{10}$$

Moreover, PMFs of random variables X and Y can be computed as

$$p_X(x) = \sum_y p_{X,Y}(x,y), \tag{11}$$

$$p_Y(y) = \sum_x p_{X,Y}(x,y). \tag{12}$$

Let's prove the first of these equalities:

$$p_X(x) = P(X = x) = \sum_y P(X = x, Y = y) = \sum_y p_{X,Y}(x,y).$$

These individual PMFs of X and Y are normally referred to as **marginal PMFs**.

Joint PMFs for discrete random variables are normally given by the tables. On the next table we give value of the PMF, and values of marginal PMFs for X and Y. In the leftmost column and in the bottom row we are given the values of Y and X respectively. Values of $p_X(x)$ are obtained by adding numbers in each column, and are given in the upper row. The values for $p_Y(y)$ are obtained by adding values in each line, and are given in the last column.

	3/20	6/20	8/20	3/20	p_X/p_Y
4	0	1/20	1/20	1/20	3/20
3	1/20	2/20	3/20	1/20	7/20
2	1/20	2/20	3/20	1/20	7/20
1	1/20	1/20	1/20	0	3/20
y/x	1	2	3	4	

We also may consider functions of two random variables $Z = g(X,Y)$. In this case,

$$p_Z(z) = \sum_{\{(x,y)|g(x,y)=z\}} p_{X,Y}(x,y). \tag{13}$$

Also we may compute the expectation of $g(X,Y)$:

$$\mathbf{E}\left[g(X,Y)\right] = \sum_{x,y} g(x,y) p_{X,Y}(x,y). \tag{14}$$

In special case when $g(X,Y) = aX + bY + x$, we have:

$$\mathbf{E}\left[aX + bX + c\right] = a\mathbf{E}\left[X\right] + b\mathbf{E}\left[Y\right] + c \tag{15}$$

Example 2.1. Assume the join PMF of X and Y is as in the previous example. Let $Z = X+Y$. Let's compute the $\mathbf{E}\left[Z\right] = \mathbf{E}\left[X + Y\right]$. In order to do it we should go over all possible pair of values of X and Y:

$$\begin{aligned}
\mathbf{E}\left[Z\right] = \mathbf{E}\left[X+Y\right] &= (1+1)p_{X,Y}(1,1) + (1+2)p_{X,Y}(1,2) + (1+3)p_{X,Y}(1,3) + (1+4)p_{X,Y}(1,4) + \\
&+ (2+1)p_{X,Y}(2,1) + (2+2)p_{X,Y}(2,2) + (2+3)p_{X,Y}(2,3) + (2+4)p_{X,Y}(2,4) + \\
&+ (3+1)p_{X,Y}(3,1) + (3+2)p_{X,Y}(3,2) + (3+3)p_{X,Y}(3,3) + (3+4)p_{X,Y}(3,4) + \\
&+ (4+1)p_{X,Y}(4,1) + (4+2)p_{X,Y}(4,2) + (4+3)p_{X,Y}(4,3) + (4+4)p_{X,Y}(4,4) = \\
&= 2 \cdot (1/20) + 3 \cdot (1/20) + 4 \cdot (1/20) + 5 \cdot 0 + \\
&+ 3 \cdot (1/20) + 4 \cdot (2/20) + 5 \cdot (2/20) + 6 \cdot (1/20) + \\
&+ 4 \cdot (1/20) + 5 \cdot (3/20) + 6 \cdot (3/20) + 7 \cdot (1/20) + \\
&+ 5 \cdot 0 + 6 \cdot (1/20) + 7 \cdot (1/20) + 8 \cdot (1/20) = \\
&= \frac{101}{20}.
\end{aligned}$$

In case of more than two random variables, we can consider the joint PMF:

$$p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n). \tag{16}$$

Also,

$$p_{X_1}(x_1) = \sum_{x_2}\sum_{x_3}\cdots\sum_{x_n} p_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n), \tag{17}$$

and

$$\mathbf{E}\left[a_1X_1 + a_2X_2 + \cdots + a_nX_n + b\right] = a_1\mathbf{E}\left[X_1\right] + \cdots + a_n\mathbf{E}\left[X_n\right] + b. \tag{18}$$

These technics can be described to obtain the expectation of the binomial random variable. Binomial random variable with parameters n and p can be represented as a sum of n Bernoulli random variables with parameter p:

$$X = X_1 + X_2 + \cdots + X_n, \tag{19}$$

where

$$p_{X_i}(x) = \begin{cases} p, & \text{if } x = 1 \\ 1 - p, & \text{if } x = 0. \end{cases} \tag{20}$$

Since X_i is equal to 1 in case the i-th trial was successful, $X = X_1 + \cdots + X_n$ gives a total number of successes in n consecutive trials, and therefore a binomial random variable. Thus,

$$\mathbf{E}\left[X\right] = \mathbf{E}\left[X_1 + X_2 + \cdots + X_n\right] = \mathbf{E}\left[X_1\right] + \mathbf{E}\left[X_2\right] + \cdots + \mathbf{E}\left[X_n\right] = np, \tag{21}$$

since the expectation of each of X_i's is equal to p.