Lecture 7

Andrei Antonenko

Feb 07, 2005

1 Commonly used discrete random variables

1.1 Poisson random variable

Consider the random variable with the following PMF:

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots \tag{1}$$

where λ is a given positive number. This random variable is called a **Poisson random variable** with parameter λ. It is easy to verify, that sum of the values of the PMF is equal to 1:

$$\sum_{k=0}^{\infty}\left(e^{-\lambda}\frac{\lambda^k}{k!}\right) = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \dots\right) = e^{-\lambda}e^{\lambda} = 1.$$

This random variable can be used as a good approximation for a binomial random variable with large n and small p. In order to do that we should assign $\lambda = np$. In this case,

$$e^{-\lambda}\frac{\lambda^k}{k!} \approx \frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n. \tag{2}$$

For example, if $n = 100$ and $p = 0.01$, $\lambda = np = 1$, and, say, for $k = 5$, we have:

$$\frac{100!}{95!5!}(0.01)^5(1-0.01)^{95} \approx 0.00290;$$
$$e^{-1}\frac{1}{5!} \approx 0.00306.$$

For example, the Poisson random variable can be used for a number of the accidents in the big city. In this case, n – the number of cars is large, and p – probability of an accident (for an individual car) is small. We will work with Poisson random variable a lot later in this course.

2 Functions of random variables

Consider the observations of the temperature. Assume, X is a temperature in Fahrenheit. The corresponding temperature in Celsius is given by the following formula:

$$Y = g(X) = (X - 32)\cdot\frac{5}{9}.$$

In this case Y is a linear function of X. In general, we can consider any function g, not necessarily linear. If X is a random variable, then $Y = g(X)$ is also a random variable, since it also provides a numeric value for the outcome of the experiment. If X is discrete with PMF p_X, then Y is also discrete, and its PMF p_Y can be calculated using the PMF of X. In particular, to obtain $p_Y(y)$ for any y, we add the probabilities of all values of x such that $g(x) = y$:

$$p_Y(y) = \sum_{\{x|g(x)=y\}} p_X(x).$$

Example 2.1. *Consider a random variable which takes integer values from* -2 *to* 2 *with equal probabilities* $1/5$. *The PMF of this r.v. is*

$$p_X(x) = \frac{1}{5}, \quad x = -2, -1, 0, 1, 2.$$

Now let $Y = |X|$. *The possible values for* Y *in this case are* $0, 1,$ *and* 2. *To compute* $p_Y(y)$ *for some given value* y *from this range, we must add* $p_X(x)$ *over all values* x *such that* $|x| = y$. *In particular, there is only one value of* X *that corresponds to* $y = 0$, *namely* $x = 0$. *Thus,* $p_Y(0) = p_X(0) = 1/5$. *Also, there are two values of* X *that correspond to each* $y = 1, 2$, *so for example,* $p_Y(1) = p_X(-1) + p_X(1) = 2/5$. *Thus, the PMF of* Y *is*

$$p_Y(y) = \begin{cases} 1/5, & \text{if } y = 0; \\ 2/5, & \text{if } y = 1; \\ 2/5, & \text{if } y = 2; \\ 0, & \text{otherwise.} \end{cases}$$

For another related example, let $Z = X^2$. *By applying the formula* $p_Z(z) = \sum_{\{x|x^2=z\}} p_X(x)$, *we obtain*

$$p_Z(z) = \begin{cases} 1/5, & \text{if } z = 0; \\ 2/5, & \text{if } z = 1; \\ 2/5, & \text{if } z = 4; \\ 0, & \text{otherwise.} \end{cases}$$

3 Expectation and variance

Assume you are spinning a wheel of fortune. This wheel has numbers $m_1, m_2, \dots m_n$ which occur with probabilities $p_1, p_2, \dots, p_n$. Each time you get the number m_i, you get m_i dollars. How much money do you expect to win on average?

Assume you spin the wheel some large number of times k, and out of these k spins, you get number m_i k_i times for each i. The amount of money you win during all these k spins is then $m_1k_1 + m_2k_2 + \cdots + m_nk_n$, and the average amount you win per spin is

$$\frac{m_1k_1 + m_2k_2 + \cdots + m_nk_n}{k} = m_1\frac{k_1}{k} + m_2\frac{k_2}{k} + \cdots + m_n\frac{k_n}{k}.$$

The numbers $\frac{k_i}{k}$ are close to the probabilities p_i of getting the number m_i:

$$p_i \approx \frac{k_i}{k}.$$

Thus, the average amount you win per spin is approximately equal to

$$m_1p_1 + m_2p_2 + \cdots + m_np_n.$$

This gives us the motivation to the following definitions:

Definition 3.1. *The* ***expectation*** *(or* ***expected value,*** *or* ***mean****) of the random variable X with PMF $p_X(x)$ is*

$$\mathbf{E}[X] = \sum_x x p_X(x). \tag{3}$$

Let's notice that the expectation might not be well defined, in case we are dealing with random values with infinite (but countable) range. For example, if the random variable takes values 2^k, $k = 1, 2, 3, \ldots$ with probabilities 2^{-k}, the formula for expectation gives

$$\sum_{k=1}^{\infty} 2^k \cdot 2^{-k} = \sum_{k=1}^{\infty} 1,$$

which is not convergent series.

Example 3.2. *Consider a sequence of 2 tosses of a biased coin with probability of head equal to p. If X is a random variable equal to the number of heads in the sequence, the PMF of X is*

$$p_X(x) = \begin{cases} p^2 & \text{if } x = 2 \\ 2 \cdot p(1-p) & \text{if } x = 1 \\ (1-p)^2 & \text{if } x = 0 \end{cases}$$

Therefore, the expected number of heads is

$$\mathbf{E}[X] = 2 \cdot p^2 + 1 \cdot 2p(1-p) + 0 \cdot (1-p)^2 = 2p^2 + 2p - 2p^2 = 2p.$$

For example, in case of a fair coin, $p = 1/2$, and the expected number of heads by the argument above will be equal to 1.

Along with computing the expected value of X, we might be interested in computing the expectation of X^2, and more generally, expectation of X^n for any n.

Definition 3.3. *The value $\mathbf{E}[X^n]$ is called* ***n-th moment*** *of X.*

The second most important characteristics of random variable is its variance.

Definition 3.4. *The* ***variance*** *of a random variable X is $\mathbf{var}(X) = \mathbf{E}[(X - \mathbf{E}[X])^2]$. The* ***standard deviation*** *of X is $\sigma_X = \sqrt{\mathbf{var}(X)}$.*

Example 3.5. *For a random variable from the previous example,*

$$\begin{aligned} &\mathbf{E}[X] = \frac{1}{5} \cdot (-2) + \frac{1}{5} \cdot (-1) + \frac{1}{5} \cdot 0 + \frac{1}{5} \cdot 1 + \frac{1}{5} \cdot 2 = 0; \\ &(X - \mathbf{E}[X])^2 = X^2; \\ &\mathbf{E}\left[X^2\right] = \frac{1}{5} \cdot 0 + \frac{2}{5} \cdot 1 + \frac{2}{5} \cdot 4 = 2, \end{aligned}$$

where in the last formula we used the PMF of $Z = X^2$, computed above.

There is a useful fact which can be used to compute the expectation of the function of a random variable:

$$\mathbf{E}\left[g(X)\right] = \sum_x g(x) p_X(x), \tag{4}$$

therefore,

$$\mathbf{E}\left[X^n\right] = \sum_x x^n p_X(x). \tag{5}$$

Example 3.6. *For a random variable from the previous example its second moment can be computed as*

$$\mathbf{E}\left[X^2\right] = \sum_x x^2 p_X(x) = \frac{1}{5} \cdot 4 + \frac{1}{5} \cdot 1 + \frac{1}{5} \cdot 0 + \frac{1}{5} \cdot 1 + \frac{1}{5} \cdot 4 = 2.$$

From the definition of variance one can see that it is always nonnegative. In case the variance is equal to 0, we have

$$(x - \mathbf{E}\left[X\right])^2 p_X(x) = 0$$

for every x, and thus $x = \mathbf{E}\left[X\right]$ for any x where $p_X(x) > 0$. It means that X is a random variable, which takes value $\mathbf{E}\left[X\right]$ with probability 1.

Next, we will obtain the formulae for the expectation and variance of $Y = aX + b$.

$$\begin{aligned}
\mathbf{E}\left[Y\right] &= \sum_x (ax+b) p_X(x) \\
&= a \sum_x x p_X(x) + b \sum_x p_X(x) \\
&= a\mathbf{E}\left[X\right] + b. \\
\mathbf{var}\left(Y\right) &= \sum_x (ax + b - \mathbf{E}\left[ax+b\right])^2 p_X(x) \\
&= \sum_x (ax + b - a\mathbf{E}\left[X\right] - b)^2 p_X(x) \\
&= a^2 \sum_x (x - \mathbf{E}\left[X\right])^2 p_X(x) \\
&= a^2 \mathbf{var}\left(X\right).
\end{aligned}$$

Therefore,

$$\mathbf{E}\left[aX + b\right] = a\mathbf{E}\left[X\right] + b, \tag{6}$$

and

$$\mathbf{var}\left(aX + b\right) = a^2 \mathbf{var}\left(X\right). \tag{7}$$

Now we will obtain important formula for the variance:

$$\begin{aligned}\mathbf{var}\,(X) &= \sum_x (x - \mathbf{E}\,[X])^2 p_X(x) \\ &= \sum_x (x^2 - 2x\mathbf{E}\,[X] + (\mathbf{E}\,[X])^2) p_X(x) \\ &= \sum_x x^2 p_X(x) - 2\mathbf{E}\,[X] \sum_x x p_X(x) + \sum_x (\mathbf{E}\,[X])^2 p_X(x) \\ &= \mathbf{E}\left[X^2\right] - 2(\mathbf{E}\,[X])^2 + (\mathbf{E}\,[X])^2 \\ &= \mathbf{E}\left[X^2\right] - (\mathbf{E}\,[X])^2.\end{aligned}$$

Thus, the formula is:

$$\mathbf{var}\,(X) = \mathbf{E}\left[X^2\right] - (\mathbf{E}\,[X])^2. \tag{8}$$

To see the use of this formula, let's compute the expectation and the variance of the Bernoulli random variable. If X is a Bernoulli random variable, we have

$$\begin{aligned}\mathbf{E}\,[X] &= 1 \cdot p + 0 \cdot (1-p) = p; \\ \mathbf{E}\left[X^2\right] &= 1^2 \cdot p + 0^2 \cdot (1-p) = p; \\ \mathbf{var}\,(X) &= \mathbf{E}\left[X^2\right] - (\mathbf{E}\,[X])^2 = p - p^2 = p(1-p).\end{aligned}$$