Lecture 20

Andrei Antonenko

March 11, 2005

1 Conditional variance as a random variable

The conditional variance can be defined in the similar way as conditional expectation:

$$\mathbf{var}(X|Y=y) = \mathbf{E}\left[(X - \mathbf{E}[X|Y=y])^2 | Y=y\right]. \tag{1}$$

The following formula is the analog of the law of iterated expectations and called **a law of conditional variances:**

$$\mathbf{var}(X) = \mathbf{E}[\mathbf{var}(X|Y)] + \mathbf{var}(\mathbf{E}[X|Y]). \tag{2}$$

Proof. We have:

$$X - \mathbf{E}[X] = (X - \mathbf{E}[X|Y]) + (\mathbf{E}[X|Y] - \mathbf{E}[X]).$$

Squaring both parts, and taking expectations, we have:

$$\begin{aligned}\mathbf{var}(X) &= \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] \\ &= \mathbf{E}\left[(X - \mathbf{E}[X|Y])^2\right] + \mathbf{E}\left[(\mathbf{E}[X|Y] - \mathbf{E}[X])^2\right] \\ &+ 2\mathbf{E}[(X - \mathbf{E}[X|Y])(\mathbf{E}[X|Y] - \mathbf{E}[X])]\end{aligned}$$

Using the law of iterated expectations, the first term in the right-hand side of the above equation can be written as

$$\mathbf{E}\left[\mathbf{E}\left[(X - \mathbf{E}[X|Y])^2 | Y\right]\right] = \mathbf{E}[\mathbf{var}(X|Y)].$$

The second term is equal to $\mathbf{var}(\mathbf{E}[X|Y])$ since $\mathbf{E}[X]$ is the expectation of $\mathbf{E}[X|Y]$. Finally, the third term is zero, as we now show. Indeed, if we define $h(Y) = 2(\mathbf{E}[X|Y] - \mathbf{E}[X])$, the third term is:

$$\begin{aligned}\mathbf{E}[(X - \mathbf{E}[X|Y])h(Y)] &= \mathbf{E}[Xh(Y)] - \mathbf{E}[\mathbf{E}[X|Y]\,h(Y)] \\ &= \mathbf{E}[Xh(Y)] - \mathbf{E}[\mathbf{E}[Xh(Y)|Y]] \\ &= \mathbf{E}[Xh(Y)] - \mathbf{E}[Xh(Y)] \\ &= 0.\end{aligned}$$

□

Now let's look at the examples.

Example 1.1. Recall the stick example from the previous lecture. We found out that

$$\mathbf{E}[X|Y] = \frac{Y}{2},$$

and Y is a uniform random variable of $[0, l]$, therefore $\mathbf{var}(Y) = l^2/12$. Now we have:

$$\mathbf{var}(\mathbf{E}[X|Y]) = \mathbf{var}\left(\frac{Y}{2}\right) = \frac{1}{4}\mathbf{var}(Y) = \frac{l^2}{48}.$$

Also, X is uniformly distributed on $[0, Y]$, and therefore,

$$\mathbf{var}\,(X|Y) = \frac{Y^2}{12}.$$

Moreover,

$$\mathbf{E}\left[\mathbf{var}\,(X|Y)\right] = \mathbf{E}\left[\frac{Y^2}{12}\right] = \frac{1}{12}\int_0^l y^2 \frac{1}{l}\, dy = \frac{1}{12}\cdot\frac{1}{3l}\, y^3\Big|_0^l = \frac{l^2}{36}.$$

Now, using the law of conditional variances,

$$\mathbf{var}\,(X) = \frac{l^2}{48} + \frac{l^2}{36} = \frac{7l^2}{144}.$$

Example 1.2. Let X be the random variable with PDF

$$f_X(x) = \begin{cases} 1/3, & 0 \le x \le 1 \\ 2/3, & 1 < x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

Now, let's Y be the random variable such that

$$Y = \begin{cases} 1, & x \le 1 \\ 2, & x \ge 1 \end{cases}$$

Now, if Y is equal to 1 (which happens with probability $1/3$) X is uniform on $[0, 1]$, and if $Y = 2$ (which happens with probability $2/3$) X is uniform on $[1, 2]$. Therefore,

$$\mathbf{E}\,[X|Y] = \begin{cases} 1/2 & \text{with probability } 1/3 \\ 3/2 & \text{with probability } 2/3 \end{cases}$$

From here,

$$\mathbf{E}\,[\mathbf{E}\,[X|Y]] = \frac{1}{2}\times\frac{1}{3} + \frac{3}{2}\times\frac{2}{3} = \frac{7}{6}$$

$$\mathbf{E}\,[\mathbf{E}\,[X|Y]] = \left(\frac{1}{2}\right)^2 \times \frac{1}{3} + \left(\frac{3}{2}\right)^2 \times \frac{2}{3} = \frac{19}{12}$$

$$\mathbf{var}\,(\mathbf{E}\,[X|Y]) = \frac{19}{12} - \left(\frac{7}{6}\right)^2 = \frac{2}{9}.$$

Now we have:

$$\mathbf{var}\,(X|Y = y) = \frac{1}{12} \quad \text{for any } y.$$

Therefore,

$$\mathbf{E}\,[\mathbf{var}\,(X|Y)] = \frac{1}{12}.$$

Finally,

$$\mathbf{var}\,(X) = \mathbf{E}\,[\mathbf{var}\,(X|Y)] + \mathbf{var}\,(\mathbf{E}\,[X|Y]) = \frac{1}{12} + \frac{2}{9} = \frac{11}{36}.$$

2 Sum of random number of random variables

Let's assume that N is a discrete random variable, which takes only nonnegative integer values. Let $X_1, X_2, \ldots$ be random variables. Assume that all above mentioned random variables are independent. Suppose that random variable N took some value. We want to find the characteristics of the sum of N random variables Y:

$$Y = X_1 + X_2 + \cdots + X_N.$$

Let's notice that substituting N by its expectation does not give us correct results, as we will demonstrate in the next example.

Example 2.1. Assume that X_i are uniform random variables on $[0, 1]$, and N takes values 1 or 3 with equal probabilities 1/2. Now, we can notice that

$$\sum_{i=1}^{N} X_i = X_1 + \cdots + X_N \in [0, 3],$$

but if we take the expectation of N, which is $\mathbf{E}[N] = 2$, we will have:

$$\sum_{i=1}^{\mathbf{E}[N]} X_i = X_1 + X_2 \in [0, 2].$$

Therefore, therefore the distribution of the sum of N random variables is not equal to the distribution of the sum of $\mathbf{E}[N]$ random variables.

Now, let's assume that all X_i's have the same expectation μ and the same variance σ^2 (they need not necessarily be normal!):

$$\mathbf{E}[X_i] = \mu, \quad \mathbf{var}(X_i) = \sigma^2 \quad \forall i. \tag{3}$$

We can compute the expectation of $Y = X_1 + \cdots + X_N$ conditioned on the value of $N = n$:

$$\begin{aligned}
\mathbf{E}[Y|N = n] &= \mathbf{E}[X_1 + X_2 + \cdots + X_N | N = n] \\
&= \mathbf{E}[X_1 + X_2 + \cdots + X_n | N = n] \\
&= \mathbf{E}[X_1 + X_2 + \cdots + X_n] \qquad \text{because } X_i\text{'s and } N \text{ are independent} \\
&= \mathbf{E}[X_1] + \mathbf{E}[X_2] + \cdots + \mathbf{E}[X_n] \\
&= n\mu.
\end{aligned}$$

Therefore,

$$\mathbf{E}[Y|N] = N\mu.$$

Now, we can use the law of iterated expectations

$$\mathbf{E}[\mathbf{E}[Y|N]] = \mathbf{E}[Y]. \tag{4}$$

We have:

$$\begin{aligned}
\mathbf{E}[Y] &= \mathbf{E}[\mathbf{E}[Y|N]] \\
&= \mathbf{E}[N\mu] \\
&= \mu\mathbf{E}[N] \\
&= \mathbf{E}[X_i]\,\mathbf{E}[N].
\end{aligned}$$

Therefore, the formula for the expectation of the random number N of identically distributed random variables X_i $Y = X_1 + \cdots + X_N$ is

$$\mathbf{E}[Y] = \mathbf{E}[X_i]\,\mathbf{E}[N]. \tag{5}$$

Now, let's find the formula for the variance. Again, conditioning on the value of N, we have:

$$\begin{aligned}
\mathbf{var}(Y|N = n) &= \mathbf{var}(X_1 + X_2 + \cdots + X_N | N = n) \\
&= \mathbf{var}(X_1 + X_2 + \cdots + X_n | N = n) \\
&= \mathbf{var}(X_1 + X_2 + \cdots + X_n) \\
&= n\sigma^2,
\end{aligned}$$

and therefore,

$$\mathbf{var}(Y|N) = N\sigma^2.$$

Now using the formula for the variance

$$\mathbf{var}(Y) = \mathbf{E}[\mathbf{var}(Y|N)] + \mathbf{var}(\mathbf{E}[Y|N]),$$

we have:

$$\begin{aligned}\mathbf{var}(Y) &= \mathbf{E}[\mathbf{var}(Y|N)] + \mathbf{var}(\mathbf{E}[Y|N]) \\ &= \mathbf{E}\left[N\sigma^2\right] + \mathbf{var}(N\mu) \\ &= \sigma^2\mathbf{E}[N] + \mu^2\mathbf{var}(N) \\ &= \mathbf{E}[N]\,\mathbf{var}(X_i) + \mathbf{E}[X_i]^2\,\mathbf{var}(N).\end{aligned}$$

So, the final formula is

$$\mathbf{var}(Y) = \mathbf{E}[N]\,\mathbf{var}(X_i) + \mathbf{E}[X_i]^2\,\mathbf{var}(N). \tag{6}$$

Now let's consider some applications of this theory.

Example 2.2. Assume in the village there are 3 gas stations, and each of them is opened with probability 1/3 independent of the others. Assume that the amount of gas in each of them is uniformly distributed from 0 to 1000 gallons. What is the expectation of the total amount of available gas? What is its variance?

The random variable N equal to the number of open gas stations has the following PMF:

$$p_N(n) = \begin{cases} (1/2)^3 = 1/8, & n = 0 \\ \binom{3}{1}(1/2)(1/2)^2 = 3/8, & n = 1 \\ \binom{3}{2}(1/2)^2(1/2) = 3/8, & n = 2 \\ (1/2)^3 = 1/8, & n = 3 \end{cases}$$

Therefore,

$$\begin{aligned}\mathbf{E}[N] &= 0\frac{1}{8} + 1\frac{3}{8} + 2\frac{3}{8} + 3\frac{1}{8} = \frac{3}{2} \\ \mathbf{E}\left[N^2\right] &= 0^2\frac{1}{8} + 1^2\frac{3}{8} + 2^2\frac{3}{8} + 3^2\frac{1}{8} = 3 \\ \mathbf{var}(N) &= \mathbf{E}\left[N^2\right] - (\mathbf{E}[N])^2 = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}.\end{aligned}$$

Now, let X_i be the amount of gas available on i-th gas station. Since X_i is uniform on $[0, 1000]$, we have:

$$\begin{aligned}\mathbf{E}[X_i] &= \frac{1000}{2} = 500 \\ \mathbf{var}(X_i) &= \frac{1000^2}{12}.\end{aligned}$$

Now, let $Y = X_1 + \cdots + X_N$ – amount of available gas. We have:

$$\begin{aligned}\mathbf{E}[Y] &= \mathbf{E}[N]\,\mathbf{E}[X_i] = \frac{3}{2} \times 500 = 750 \\ \mathbf{var}(Y) &= \mathbf{E}[N]\,\mathbf{var}(X_i) + (\mathbf{E}[X]_i)^2\mathbf{var}(N) = \frac{3}{2} \times \frac{1000^2}{12} + 500^2 \times \frac{3}{4} = 312,500.\end{aligned}$$