Math 525: Lecture 7

January 25, 2018

1 Expectation (general case)

In a previous lecture, we define the expectation of a discrete random variable. Armed now with an intuitive understanding of expectations, let's extend our definition to the general case.

Definition 1.1. Let X be a random variable and n be a nonnegative integer. Let $\{E_n^k\}_{k\in\mathbb{Z}}$ be the partition of $\mathbb{R}$ defined by

$$E_n^k = \left(\frac{k}{2^n}, \frac{k+1}{2^n}\right].$$

Then, the n-th *lower and upper dyadic approximations* of X are the random variables $\underline{X}_n$ and $\overline{X}_n$ defined by

$$\underline{X}_n(\omega) = \sum_{k\in\mathbb{Z}} \frac{k}{2^n} I_{E_n^k}(X(\omega)) \qquad \text{and} \qquad \overline{X}_n(\omega) = \sum_{k\in\mathbb{Z}} \frac{k+1}{2^n} I_{E_n^k}(X(\omega)).$$

Note that $I_{E_n^k} \circ X$ is a random variable ($I_{E_n^k}$ is Borel measurable because $E_n^k \in \mathcal{B}(\mathbb{R})$). It is obvious from the definition that the lower and upper dyadic approximations of X are discrete random variables and hence we know how to take their expectations (assuming they are integrable). It is also useful to note that

$$\frac{k}{2^n} = \inf E_n^k \qquad \text{and} \qquad \frac{k+1}{2^n} = \max E_n^k.$$

Proposition 1.2. *Let X be a random variable. For each nonnegative integer n,*

1. $\overline{X}_n - \underline{X}_n = 1/2^n$.
2. $\underline{X}_n \leq \underline{X}_{n+1} < X \leq \overline{X}_{n+1} \leq \overline{X}_n$.
3. *The dyadic approximations $\{\underline{X}_0, \overline{X}^0, \underline{X}_1, \overline{X}^1 \ldots\}$ are either all integrable or all non-integrable.*

Proof. Let n be a nonnegative integer. The first claim follows from

$$\overline{X}_n(\omega) - \underline{X}_n(\omega) = \sum_{k\in\mathbb{Z}} \frac{k+1}{2^n} I_{E_n^k}(X(\omega)) - \sum_{k\in\mathbb{Z}} \frac{k}{2^n} I_{E_n^k}(X(\omega)) = \frac{1}{2^n} \sum_{k\in\mathbb{Z}} I_{E_n^k}(X(\omega)) = \frac{1}{2^n}.$$

Suppose now that $X(\omega) \in E_n^k = (k/2^n, (k+1)/2^n]$ for some particular $k \in \mathbb{Z}$. Then, $\underline{X}_n(\omega) = k/2^n$ and $\overline{X}_n(\omega) = (k+1)/2^n$. This establishes $\underline{X}_n(\omega) < X_n(\omega) \leq \overline{X}_n(\omega)$. Moreover, since $E_n^k = E_{n+1}^{2k} \cup E_{n+1}^{2k+1}$, it follows that $\underline{X}_{n+1}(\omega) \geq 2k/2^{n+1} = k/2^n$ and $\overline{X}_n(\omega) \leq (2k+2)/2^{n+1} = (k+1)/2^n$, establishing the remaining inequalities.

The last claim is established by noting that

$$X(\omega) - \underline{X}_n(\omega) = \sum_{k\in\mathbb{Z}} \left(X(\omega) - \frac{k}{2^n} \right) I_{E_n^k}(X(\omega)) = \frac{1}{2^n} \sum_{k\in\mathbb{Z}} I_{E_n^k}(X(\omega)) = \frac{1}{2^n}$$

and similarly,

$$X(\omega) - \underline{X}_n(\omega) = \frac{1}{2^n}.$$

Therefore, any two dyadic approximations differ by at most 1. □

The above result implies that if any particular dyadic approximation (e.g., $\overline{X}_0$) is integrable, then

$$\mathbb{E}\underline{X}_0 \leq \mathbb{E}\underline{X}_1 \leq \mathbb{E}\underline{X}_2 \leq \cdots \leq \mathbb{E}\overline{X}^2 \leq \mathbb{E}\overline{X}^1 \leq \mathbb{E}\overline{X}^0.$$

Moreover, since $\mathbb{E}\overline{X}_n - \mathbb{E}\underline{X}_n = 1/2^n$, it follows that if the dyadic approximations are integrable, $\mathbb{E}\underline{X}_n$ and $\mathbb{E}\overline{X}_n$ converge from below and above, respectively, to the same point as $n \to \infty$. This observation motivates the following definition:

Definition 1.3. A random variable X is *integrable* if $\overline{X}_0$ is integrable. In this case, we define the expectation of X as

$$\mathbb{E}X = \lim_{n\to\infty} \mathbb{E}\overline{X}_n = \lim_{n\to\infty} \mathbb{E}\underline{X}_n.$$

Next, we will "reprove" some properties of expectations that we already encountered in the discrete case. In order to make matters simpler, we first introduce a bit of notation:

Definition 1.4. We say an event occurs *almost surely* (a.s.) if it occurs with probability one.

Example 1.5. Let X and Y be random variables. We say $X = Y$ a.s. if

$$\mathbb{P}(\{X = Y\}) = \mathbb{P}(\{\omega \in \Omega \colon X(\omega) = Y(\omega)\} = 1.$$

In other words, there could be outcomes $\omega \in \Omega$ for which $X(\omega) \neq Y(\omega)$, but the set of all outcomes has probability zero. For all intents and purposes, X and Y are the "same". In this case, for any set $B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}(\{X \in B\}) = \mathbb{P}(\{Y \in B\}).$$

To see why, note that

$$\mathbb{P}(\{X \in B\}) = \mathbb{P}(\{X \in B\} \cap \{X = Y\}) = \mathbb{P}(\{Y \in B\} \cap \{X = Y\}) = \mathbb{P}(\{Y \in B\}).$$

We are now ready to prove some properties of expectations:

Proposition 1.6. *Let X and Y be random variables and let $a \in \mathbb{R}$. Then,*

1. *If $X = Y$ a.s., then Y is integrable if and only if X is, and if so, $\mathbb{E}X = \mathbb{E}Y$.*
2. *If X and Y are integrable, so is $aX + bY$ and*
$$\mathbb{E}\left[aX + bY\right] = a\mathbb{E}X + b\mathbb{E}Y.$$
3. *If $|X| \leq |Y|$ and Y is integrable, then X is integrable.*
4. *If X and Y are integrable and $X \leq Y$, then $\mathbb{E}X \leq \mathbb{E}Y$.*
5. *If X is integrable, $\mathbb{E}X \leq \mathbb{E}|X|$.*

The properties above are exactly those in the discrete case, but now they hold for general random variables! We do have to do some work to prove them, however.

Proof. Recall that $\overline{X}_n - \underline{X}_n = 1/2^n$. Since $\underline{X}_n \leq X \leq \overline{X}_n$, this implies $|\overline{X}_n - X| \leq 1/2^n$, which in turn implies
$$\left|\overline{X}_n\right| - 1/2^n \leq |X| \leq \left|\overline{X}_n\right| + 1/2^n.$$

1. First, note that the statement is true in the discrete case. Now, if $X = Y$ a.s., then $\overline{X}_n = \overline{Y}_n$ a.s., from which the desired result follows.
2. We will use (3) to prove this. Note that
$$|aX + bY| \leq |a|\,|X| + |b|\,|Y| \leq |a|\left(\left|\overline{X}_0\right| + 1\right) + |b|\left(\left|\overline{Y}_0\right| + 1\right),$$
and hence $aX + bY$ is integrable. Now, let $Z = aX + bY$. Then,
$$\begin{aligned}\left|\overline{Z}_n - \left(a\overline{X}_n + b\overline{Y}_n\right)\right| &\leq \left|\overline{Z}_n - Z\right| + \left|Z - \left(a\overline{X}_n + b\overline{Y}_n\right)\right| \\ &\leq 1/2^n + |a|\left|X - \overline{X}_n\right| + |b|\left|Y - \overline{Y}_n\right| \\ &\leq 1/2^n\left(1 + |a| + |b|\right) \\ &\to 0.\end{aligned}$$
Therefore,
$$\lim_{n\to\infty} \mathbb{E}\left[\overline{Z}_n\right] = \lim_{n\to\infty} \mathbb{E}\left[a\overline{X}_n + b\overline{Y}_n\right] = \lim_{n\to\infty} a\mathbb{E}\left[\overline{X}_n\right] + b\mathbb{E}\left[\overline{Y}_n\right] = a\mathbb{E}X + b\mathbb{E}Y.$$
3. First, note that $|\overline{X}_0| \leq |\overline{Y}_0| + 2$ since
$$\left|\overline{X}_0\right| - 1 \leq |X| \leq |Y| \leq \left|\overline{Y}_0\right| + 1.$$
Now, if Y is integrable, $\overline{Y}_0$ is integrable by definition. In this case, $\overline{X}_0$ is integrable, and hence X is integrable by definition.
4. Exercise.

5. Take $Y = |X|$ in (4).

$\square$

Example 1.7. Let X be **any** random variable and define its moment generating function (MGF) by $M(\theta) = \mathbb{E}[e^{\theta X}]$. Note that the discussions of the previous lecture, albeit based in discrete random variables, only used properties of the expectation. Therefore, we can still talk about the moment generating function for an arbitrary random variable, and we still have the nice result $M^{(k)}(0) = \mathbb{E}[X^k]$ (assuming, of course, that M is differentiable in a neighbourhood of zero).

Proposition 1.8. *Let X and Y be independent and integrable random variables. Then, XY is integrable and*

$$\mathbb{E}[XY] = \mathbb{E}X\mathbb{E}Y.$$

Proof. Consider first the discrete case:

$$\begin{aligned}
\mathbb{E}[|XY|] &= \sum_{i,j} |x_i|\,|y_j|\,\mathbb{P}(\{X = x_i\} \cap \{Y = y_j\}) \\
&= \sum_{i,j} |x_i|\,|y_j|\,\mathbb{P}(\{X = x_i\})\mathbb{P}(\{Y = y_j\}) \\
&= \left(\sum_i |x_i|\,\mathbb{P}(\{X = x_i\})\right)\left(\sum_j |y_j|\,\mathbb{P}(\{Y = y_j\})\right) \\
&= \mathbb{E}[|X|]\,\mathbb{E}[|Y|].
\end{aligned}$$

This establishes that XY are integrable whenever X and Y. Repeating the same argument now without the absolute value signs establishes $\mathbb{E}[XY] = \mathbb{E}X\mathbb{E}Y$.

Now, let's move on to the general case. If X and Y are integrable,

$$\begin{aligned}
\mathbb{E}\left[\left|\overline{X}_n\overline{Y}_n - XY\right|\right] &= \mathbb{E}\left[\left|\overline{X}_n\overline{Y}_n - \overline{X}_nY + \overline{X}_nY - XY\right|\right] \\
&\le \mathbb{E}\left[\left|\overline{X}_n\right|\left|\overline{Y}_n - Y\right|\right] + \mathbb{E}\left[|Y|\left|\overline{X}_n - X\right|\right] \\
&\le 1/2^n \cdot \left(\mathbb{E}\left[\left|\overline{X}_n\right|\right] + \mathbb{E}[|Y|]\right) \\
&\to 0.
\end{aligned}$$

Since

$$\mathbb{E}\left[|XY| - \left|\overline{X}_n\overline{Y}_n\right|\right] \le \mathbb{E}\left[\left|\overline{X}_n\overline{Y}_n - XY\right|\right],$$

this implies that XY is integrable. Moreover, if X and Y are independent, so too are $\overline{X}_n$ and $\overline{Y}_n$. This implies

$$\mathbb{E}[XY] = \lim_{n\to\infty} \mathbb{E}\left[\overline{X}_n\overline{Y}_n\right] = \lim_{n\to\infty} \mathbb{E}\left[\overline{X}_n\right]\mathbb{E}\left[\overline{Y}_n\right] = \mathbb{E}X\mathbb{E}Y$$

as desired. $\square$

Example 1.9. Consider n independent random variables $X_1, \ldots, X_n$ with MGFs $M_{X_i}(\theta) = \mathbb{E}[e^{\theta X_i}]$. Then,

$$M_{X_1+\cdots+X_n}(\theta) = \mathbb{E}\left[e^{\theta(X_1+\cdots+X_n)}\right] = \mathbb{E}\left[e^{\theta X_1}\right]\cdots\mathbb{E}\left[e^{\theta X_n}\right] = M_{X_1}(\theta)\cdots M_{X_n}(\theta).$$

In other words, the MGF of the sum is the product of the MGFs.

2 Moment inequalities

There are a few useful inequalities concerning moments of random variables we should cover. We start with Chebyshev's inequality.

2.1 Chebyshev's inequality

Proposition 2.1 (Chebyshev's inequality)**.** *Let $p > 0$, $\lambda > 0$, and X be a random variable with X^p integrable. Then,*

$$\mathbb{P}(\{|X| \geq \lambda\}) \leq \frac{1}{\lambda^p}\mathbb{E}\left[|X|^p\right].$$

Proof. First, note that

$$\mathbb{P}(\{|X| \geq \lambda\}) = \mathbb{P}(\{|X|^p \geq \lambda^p\}) = \mathbb{E}\left[I_{\{|X|^p \geq \lambda^p\}}\right].$$

But if $|X(\omega)|^p \geq \lambda^p$, then $1 \leq \lambda^p / |X(\omega)|^p$. Therefore,

$$\mathbb{E}\left[I_{\{|X|^p \geq \lambda^p\}}\right] \leq \mathbb{E}\left[\frac{|X|^p}{\lambda^p} I_{\{|X|^p \geq \lambda^p\}}\right] \leq \frac{1}{\lambda^p}\mathbb{E}\left[|X|^p I_{\{|X|^p \geq \lambda^p\}}\right] \leq \frac{1}{\lambda^p}\mathbb{E}\left[|X|^p\right]. \qquad \square$$

Corollary 2.2. *Let $\lambda > 0$ and Y be a square integrable (i.e., Y^2 is integrable) random variable. Then,*

$$\mathbb{P}(\{|Y - \mathbb{E}Y| \geq \lambda\}) \leq \frac{1}{\lambda^2}\operatorname{Var}(Y).$$

Proof. Take $p = 2$ and $X = Y - \mathbb{E}Y$ in Chebyshev's inequality. $\square$

2.2 Cauchy-Schwarz(-Buniakovski) inequality

Proposition 2.3 (Cauchy-Schwarz(-Buniakovski) inequality)**.** *Let X and Y be square integrable random variables. Then,*

$$\mathbb{E}\left[XY\right] \leq \sqrt{\mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]}.$$

Proof. If either X or Y is zero a.s., then the inequality is trivial. Therefore, suppose that neither is zero a.s. Let $\lambda \geq 0$. Then,

$$0 \leq \mathbb{E}\left[(X - \lambda Y)^2\right] = \mathbb{E}\left[X^2\right] - 2\lambda\mathbb{E}\left[XY\right] + \lambda^2\mathbb{E}\left[Y^2\right]$$

and hence

$$\mathbb{E}\left[XY\right] \leq \frac{1}{2}\left(\frac{1}{\lambda}\mathbb{E}\left[X^2\right] + \lambda\mathbb{E}\left[Y^2\right]\right).$$

Letting $\lambda = \sqrt{\mathbb{E}\left[X^2\right]}/\sqrt{\mathbb{E}\left[Y^2\right]}$ yields

$$\begin{aligned}
\mathbb{E}\left[XY\right] &\leq \frac{1}{2}\left(\frac{\sqrt{\mathbb{E}\left[Y^2\right]}}{\sqrt{\mathbb{E}\left[X^2\right]}}\mathbb{E}\left[X^2\right] + \frac{\sqrt{\mathbb{E}\left[X^2\right]}}{\sqrt{\mathbb{E}\left[Y^2\right]}}\mathbb{E}\left[Y^2\right]\right) \\
&= \frac{\mathbb{E}\left[Y^2\right]\mathbb{E}\left[X^2\right]}{\sqrt{\mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]}} \\
&= \sqrt{\mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]}. \qquad \square
\end{aligned}$$

Example 2.4. Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Bernoulli}(p)$ be random variables. Then, by the Cauchy-Schwarz inequality,

$$\mathbb{E}\left[XY\right] \leq \sqrt{\mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]} = \sqrt{\lambda\left(\lambda+1\right)p}.$$

If the two random variables are independent, then

$$\mathbb{E}\left[XY\right] = \mathbb{E}X\mathbb{E}Y = \lambda p.$$

Indeed, you can check that for all $0 \leq p \leq 1$ and $\lambda \geq 0$,

$$\lambda p \leq \sqrt{\lambda\left(\lambda+1\right)p}.$$