# Lecture 13

## 1. Inequalities

Let us start with an inequality.

**Lemma 13.1.** *If* $h$ *is a nonnegative function, then for all* $\lambda > 0$*,*

$$P\{h(X) \geqslant \lambda\} \leqslant \frac{E[h(X)]}{\lambda}.$$

**Proof.** We know already that

$$E[h(X)] = \sum_x h(x)f(x) \geqslant \sum_{x:\ h(x)\geqslant\lambda} h(x)f(x).$$

If $x$ is such that $h(x) \geqslant \lambda$, then $h(x)f(x) \geqslant \lambda f(x)$, obviously. Therefore,

$$E[h(X)] \geqslant \lambda \sum_{x:\ h(x)\geqslant\lambda} f(x) = \lambda P\{h(X) \geqslant \lambda\}.$$

Divide by $\lambda$ to finish. □

Thus, for example,

$$P\{|X| \geqslant \lambda\} \leqslant \frac{E(|X|)}{\lambda} \qquad \text{“Markov's inequality.”}$$

$$P\{|X - EX| \geqslant \lambda\} \leqslant \frac{\mathrm{Var}(X)}{\lambda^2} \qquad \text{“Chebyshev's inequality.”}$$

To get Markov's inequality, apply Lemma 13.1 with $h(x) = |x|$. To get Chebyshev's inequality, first note that $|X-EX| \geqslant \lambda$ if and only if $|X-EX|^2 \geqslant \lambda^2$. Then, apply Lemma 13.1 to find that

$$P\{|X - EX| \geqslant \lambda\} \leqslant \frac{E\left(|X - EX|^2\right)}{\lambda^2}.$$

Then, recall that the numerator is $\mathrm{Var}(X)$.

In words:

- If $E(|X|) < \infty$, then the probability that $|X|$ is large is small.
- If $\mathrm{Var}(X)$ is small, then with high probability $X \approx EX$.

## 2. Conditional distributions

If $X$ is a random variable with mass function $f$, then $\{X = x\}$ is an event. Therefore, if $B$ is also an event, and if $P(B) > 0$, then

$$P(X = x \,|\, B) = \frac{P(\{X = x\} \cap B)}{P(B)}.$$

As we vary the variable $x$, we note that $\{X = x\} \cap B$ are disjoint. Therefore,

$$\sum_x P(X = x \,|\, B) = \frac{\sum P(\{X = x\} \cap B)}{P(B)} = \frac{P\left(\cup_x \{X = x\} \cap B\right)}{P(B)} = 1.$$

Thus,

$$f(x \,|\, B) = P(X = x \,|\, B)$$

defines a mass function also. This is called the *conditional mass function of* $X$ *given* $B$.

**Example 13.2.** Let $X$ be distributed uniformly on $\{1, \dots, n\}$, where $n$ is a fixed positive integer. Recall that this means that

$$f(x) = \begin{cases} \frac{1}{n} & \text{if } x = 1, \dots, n, \\ 0 & \text{otherwise.} \end{cases}$$

Choose and fix two integers $a$ and $b$ such that $1 \leqslant a \leqslant b \leqslant n$. Then,

$$P\{a \leqslant X \leqslant b\} = \sum_{x=a}^{b} \frac{1}{n} = \frac{b - a + 1}{n}.$$

Therefore,

$$f(x \,|\, a \leqslant X \leqslant b) = \begin{cases} \frac{1}{b - a + 1} & \text{if } x = a, \dots, b, \\ 0 & \text{otherwise.} \end{cases}$$

## 3. Conditional expectations

Once we have a (conditional) mass function, we have also a conditional expectation at no cost. Thus,

$$E(X\,|\,B) = \sum_x x f(x\,|\,B).$$

**Example 13.3** (Example 13.2, continued)**.** In Example 13.2,

$$E(X\,|\,a \leqslant X \leqslant b) = \sum_{k=a}^{b} \frac{k}{b-a+1}.$$

Now,

$$\begin{aligned}\sum_{k=a}^{b} k &= \sum_{k=1}^{b} k - \sum_{k=1}^{a-1} k \\ &= \frac{b(b+1)}{2} - \frac{(a-1)a}{2} \\ &= \frac{b^2+b-a^2+a}{2}.\end{aligned}$$

Write $b^2 - a^2 = (b-a)(b+a)$ and factor $b+a$ to get

$$\sum_{k=a}^{b} k = \frac{b+a}{2}(b-a+1).$$

Therefore,

$$E(X\,|\,a \leqslant X \leqslant b) = \frac{b+a}{2}.$$

This should not come as a surprise. Example 13.2 actually shows that given $B = \{a \leqslant X \leqslant b\}$, the conditional distribution of $X$ given $B$ is uniform on $\{a, \ldots, b\}$. Therefore, the conditional expectation is the expectation of a uniform random variable on $\{a, \ldots, b\}$.

**Theorem 13.4** (Bayes's formula for conditional expectations)**.** *If* $P(B) > 0$*, then*

$$EX = E(X\,|\,B)P(B) + E(X\,|\,B^c)P(B^c).$$

**Proof.** We know from the ordinary Bayes's formula that

$$f(x) = f(x\,|\,B)P(B) + f(x\,|\,B^c)P(B^c).$$

Multiply both sides by $x$ and add over all $x$ to finish. □

**Remark 13.5.** The more general version of Bayes's formula works too here: Suppose $B_1, B_2, \ldots$ are disjoint and $\cup_{i=1}^{\infty} B_i = \Omega$; i.e., "one of the $B_i$'s happens." Then,

$$EX = \sum_{i=1}^{\infty} E(X \mid B_i) P(B_i).$$

**Example 13.6.** Suppose you play a fair game repeatedly. At time 0, before you start playing the game, your fortune is zero. In each play, you win or lose with probability 1/2. Let $T_1$ be the first time your fortune becomes +1. Compute $E(T_1)$.

More generally, let $T_x$ denote the first time to win $x$ dollars, where $T_0 = 0$.

Let $W$ denote the event that you win the first round. Then, $P(W) = P(W^c) = 1/2$, and so

$$E(T_x) = \frac{1}{2} E(T_x \mid W) + \frac{1}{2} E(T_x \mid W^c). \tag{11}$$

Suppose $x \neq 0$. Given $W$, $T_x$ is one plus the first time to make $x - 1$ more dollars. Given $W^c$, $T_x$ is one plus the first time to make $x+1$ more dollars. Therefore,

$$\begin{aligned} E(T_x) &= \frac{1}{2}\Big[1 + E(T_{x-1})\Big] + \frac{1}{2}\Big[1 + E(T_{x+1})\Big] \\ &= 1 + \frac{E(T_{x-1}) + E(T_{x+1})}{2}. \end{aligned}$$

Also $E(T_0) = 0$.

Let $g(x) = E(T_x)$. This shows that $g(0) = 0$ and

$$g(x) = 1 + \frac{g(x+1) + g(x-1)}{2} \qquad \text{for } x = \pm 1, \pm 2, \ldots.$$

Because $g(x) = (g(x) + g(x))/2$,

$$g(x) + g(x) = 2 + g(x+1) + g(x-1) \qquad \text{for } x = \pm 1, \pm 2, \ldots.$$

Solve to find that for all integers $x \geqslant 1$,

$$g(x+1) - g(x) = -2 + g(x) - g(x-1).$$