1 Expected Value and Variance

1.1 Conditional Expectation

Recall that if X is a discrete random variable with PMF p, and B is an event with $\mathbb{P}(B) > 0$, then

$$p(x|B) = \frac{p(x)}{\mathbb{P}(B)}$$

is a probability distribution on B. This is the PMF of the random variable X, given B.

Definition 1.1

Let X be a random variable with PMF p. Suppose that $\mathbb{P}(B) > 0$. Then, the conditional expectation of X given B is

$$\mathbb{E}(X|B) = \sum_x xp(x|B).$$

Remark: The situation is similar in the continuous case, but we instead have a conditional PDF

$$f(x|B) = \frac{f(x)}{\mathbb{P}(B)}$$

and the conditional expectation is given by

$$\mathbb{E}(X|B) = \int_{-\infty}^{\infty} xp(x|B)dx.$$

1.1.1 Law of Total Expectation

Just like how there is a Law of Total Probability, there is also a Law of Total Expectation.

Theorem 1.1: Law of Total Expectation

Let X be a random variable on sample space Ω. Suppose that $B_1, \ldots, B_n$ is a partition Ω. Then,

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X|B_i)\mathbb{P}(B_i).$$

This is useful because, often, $\mathbb{E}(X)$ is sometimes difficult to find directly. However, if we condition on a well-chosen B_i, then it becomes manageable.

(Example.) In the gambling game "craps," a player makes a bet and then rolls a pair of dice. If the sum is 7 or 11, the player wins. If it is 2, 3, or 12, the player loses. If the sum is any other number s, the player continues to roll until either another s (they win) or 7 (they lose) occurs (7 is lucky the first time). Now, let R be the number of rolls in a single game of craps.

1. Find $\mathbb{E}(R)$.

1. By the Law of Total Expectation, we have

$$\mathbb{E}(R) = \sum_{x=2}^{12} \mathbb{E}(R|X = x)\mathbb{P}(X = x),$$

where X is the initial sum. Note that if

$$x \in \{2, 3, 7, 11, 12\},$$

then

$$\mathbb{E}(R|X = x) = 1$$

since the game is immediately over if we get one of those numbers. In particular,

- There is 1 way to get a 2 (11).
- There are 2 ways to get a 3 (12, 21).
- There are 6 ways to get a 7 (16, 61, 25, 52, 43, 34).
- There are 2 ways to get a 11 (56, 65).
- There is 1 way to get a 12 (66).

Hence,

$$\sum_{x\in\{2,3,7,11,12\}} \mathbb{E}(R|X = x)\mathbb{P}(X = x) = \frac{12}{36}.$$

Now, if

$$x \in \{4, 5, 6, 8, 9, 10\},$$

then we can use a similar argument to the one above. For example, when $x = 4$, we have 3 ways to get 4 (13, 31, 22). This gives us

$$\mathbb{P}(X = 4) = \frac{3}{36}.$$

There are also 6 ways to get 7 (16, 61, 25, 52, 43, 34). Therefore, the number of rolls R, given that the initial sum is $X = 4$, is distributed as $1 + G$, where G is a geometric random variable (where the success is defined by rolling a 4 or 7) with success probability $p = \frac{9}{36}$. Note that the 1 is there because of the initial roll of the 4, but then we have to keep rolling. Thus, we get

$$\mathbb{E}(R|X = 4)\mathbb{P}(X = 4) = \left(1 + \frac{36}{9}\right)\frac{3}{36}.$$

To see how we got $\left(1 + \frac{36}{9}\right)$, recall that $E(X) = \frac{1}{p}$ if X is Geometric(p). So, by Linearity of Expectation, we get that the expected value of 1 (so it's 1) plus the expected value of the geometric (so it's 36/9).

So, by similar reasoning, we get

$$\mathbb{E}(R) = \frac{12}{36} + \left(1 + \frac{36}{9}\right)\frac{3}{36} + \left(1 + \frac{36}{10}\right)\frac{4}{36} + \left(1 + \frac{36}{11}\right)\frac{5}{36} + \left(1 + \frac{36}{11}\right)\frac{5}{36} + \left(1 + \frac{36}{10}\right)\frac{4}{36} + \left(1 + \frac{36}{9}\right)\frac{3}{36} = \frac{557}{165} \approx 3.375$$

So, on average, we expect a little bit more than 3 and 1/3 rolls of the dice in each of te game of craps.

1.1.2 Martingales

Informally, we can think of martingales as random processes that encapsulate the idea of a *fair game*.

Definition 1.2

Let $(X_0, X_1, X_2, \dots)$ be a sequence of random variables and ϕ a function. Let $M_n = \phi(X_n)$. The sequence $(M_0, M_1, M_2, \dots)$ is called a **martingale** (MG) with respect to $(X_0, X_1, X_2, \dots)$ if, for any n and any $x_0, x_1, \dots, x_n$, we have that

$$\mathbb{E}(M_{n+1} - M_n | X_n = x_n, \dots, X_0 = x_0) = 0.$$

Remark: If we think of M_n as the total winnings after n bets by a gambler, then $(M_0, M_1, M_2, \dots)$ is a "fair game" in the sense that neither the gambler nor the "house" has an advantage. In other words, after the $(n+1)$th game, we would expect no additional gain (hence why this is equal to 0).

Theorem 1.2

For any random variables X and Y, we have that

$$\mathbb{E}[\mathbb{E}(X|Y)] = \mathbb{E}(X).$$

Remarks:

- Recall that $\mathbb{E}(M_{n+1}|X_n, \dots, X_0) = M_n$ for a MG. Then, taking expectations on both sides, it follows that $\mathbb{E}(M_{n+1}) = \mathbb{E}(M_n)$ for all n. Hence, for all (deterministic) times $n \geq 0$, we have that $\mathbb{E}(M_n) = \mathbb{E}(M_0)$.
- These are a powerful tool, which can lead to quick or slick proofs of things that would be computationally challenging otherwise.

(Example.) Recall, from Lecture 1, that Peter and Paul keeps flipping a coin. If it is "Heads," then Peter wins \$1. Otherwise, Peter loses \$1. Let X_i be Peter's winning after the ith bet. Note that $X_0 = 0$.

To see that this is a MG, note that the series of flips are independent. Hence, for any n and any $x_1, \dots, x_n$, we have

$$\mathbb{E}(X_{n+1} - X_n | X_n = x_n, \dots, X_1 = x_1) = (1)\frac{1}{2} + (-1)\frac{1}{2} = 0.$$

Recall that $\mathbb{E}(M_n) = \mathbb{E}(M_0)$ for all (deterministic) times $n \geq 0$. However, this is not necessarily true for *random* T. Hence, we restrict to a special class of random times.

Definition 1.3

A time T is a **stopping time** (ST) if and only if, for any n, to know if $T = n$, we only need to know the values of $X_0, X_1, \dots, X_n$. That is, we do not need any information about the future after time n.

For example, the first time we visit 0 is a ST, but the last time we visit 0 is not (in order for us to know if time n is 0, we need to know the full history).

Theorem 1.3: The Optional Stopping Theorem (OST)

If $(M_0, M_1, \dots)$ is a MG and T is a ST, then $\mathbb{E}(M_T) = \mathbb{E}(M_0)$ if the following conditions are satisfied:

1. M_n is bounded until time T, and
2. $\mathbb{P}(T < \infty) = 1$.

(Example: Gambler's Ruin). Suppose that Peter currently has \$1. Furthermore, suppose that they play instead with a coin that comes up Heads with probability $p \neq 1/2$. What is the probability $\mathbb{P}(J)$ that Peter wins the "Jackpot" (\$$N$) before going "bust" (\$0)?

For this biased RW, we know that

$$M_n = (q/p)^{X_n}$$

is a MG, where X_n is Peter's winnings. Note that here

$$\phi(x) = (q/p)^x.$$

Then,

$$\begin{aligned}\mathbb{E}(M_{n+1} - M_n | X_n = x_n, \ldots, X_1 = x_1) &= (q/p)^{x_n}(q/p - 1)p + (q/p)^{x_n}(p/q - 1)q \\ &= (q/p)^{x_n}[(q-p) + (p-q)] \\ &= 0.\end{aligned}$$

Now, let T be the first time that $X_n \in \{0, N\}$. By the OST, we know that

$$q/p = \mathbb{E}(M_T) = 1 \cdot \mathbb{P}(J^C) + (q/p)^N \cdot \mathbb{P}(J).$$

Hence,

$$\mathbb{P}(J) = \frac{(q/p) - q}{(q/p)^N - 1}.$$