EE5110: Probability Foundations for Electrical Engineers

July - Nov 2024

1 References

- See Chapter 2: Discrete-type Random Variables, Bruce Hajek for a review of discrete random variables.
- See Chapter 2: Discrete random variables, Bertsekas and Tsitsiklis, for a review of discrete random variables.

2 Solved Problems

1. If $X \geq a$, then $\mathsf{E}[X] \geq a$.

   Let X be a discrete random variable with support $\Omega_X = \{x_1, x_2, \cdots\}$ and p.m.f. $\{p_X(x_i)\}$. Let us assume the following.

   (a) $X(\omega) \geq a$ for all $\omega \in \Omega$, or,

   (b) $\mathsf{P}(\{\omega : X(\omega) \geq a\}) = 1$

   This would imply that $x_i \geq a$ for all i such that $p_X(x_i) > 0$. Hence, without loss of generality, we can assume that $\{x_1, x_2, \cdots\}$ are such that $x_i \geq a$ for all i, and with probability mass function $\{p_X(x_i)\}$.

   The expected value of the random variable is defined as

   $$\mathsf{E}[X] = \sum_i x_i \; p_X(x_i)$$

   But $x_i \geq a$ for all i, and $p_X(x_i) \geq 0$ always. Hence,

   $$\mathsf{E}[X] = \sum_i x_i \; p_X(x_i) \geq \sum_i a \; p_X(x_i) = a \sum_i p_X(x_i) = a$$

   In particular, if $X \geq 0$, then $\mathsf{E}[X] \geq 0$.

   □

2. (Fundamental Theorem of Expectation): Let X be a discrete random variable and $g : \mathbb{R} \to \mathbb{R}$ a measurable function.Then,

   $$\mathsf{E}[g(X)] = \sum_i g(x_i) \; p_X(x_i)$$

Let X be a discrete random variable with support $\Omega_X = \{x_1, x_2, \cdots\}$ and p.m.f. $\{p_X(x_i)\}$. Let g be a measurable function from $\mathbb{R}$ to $\mathbb{R}$.

Define $Y = g(X)$ or $Y(\omega) = g(X(\omega))$. Then, Y is a discrete random variable as well. The expected value of the random variable Y is defined as

$$\mathsf{E}[Y] = \sum_j y_j \; p_Y(y_j)$$

Substituting for $p_Y(y_j) = \sum_i \mathbb{1}_{\{g(x_i)=y_j\}} \; p_X(x_i)$ in the above expression, we get,

$$\begin{aligned}
\mathsf{E}[Y] = \sum_j y_j \; p_Y(y_j) &= \sum_j y_j \sum_i \mathbb{1}_{\{g(x_i)=y_j\}} \; p_X(x_i) \\
&= \sum_j \sum_i y_j \; \mathbb{1}_{\{g(x_i)=y_j\}} \; p_X(x_i) \\
&= \sum_i \sum_j y_j \; \mathbb{1}_{\{g(x_i)=y_j\}} \; p_X(x_i) \\
&= \sum_i p_X(x_i) \sum_j y_j \; \mathbb{1}_{\{g(x_i)=y_j\}} \\
&= \sum_i p_X(x_i) \; g(x_i)
\end{aligned}$$

where we will assume that interchanging the order of summations is permitted.

$\square$

3. (Total Expectation Theorem): Let $A_1, A_2, \cdots A_n$ denote a partition of Ω such that $\mathsf{P}(A_i) > 0$ for all i. Let X be a discrete random variable with p.m.f. $\{p_X(x_i)\}$. Then,

$$\mathsf{E}[X] = \sum_j \mathsf{P}(A_j) \; \mathsf{E}[X|A_j]$$

The expected value of X is defined as

$$\mathsf{E}[X] = \sum_i x_i \; p_X(x_i)$$

Using total probability theorem, we have, for all i,

$$p_X(x_i) = \sum_j \mathsf{P}(A_j) \; p_{X|A_j}(x_i)$$

Substituting the above expression in the definition of expectation of X, we get,

$$
\begin{aligned}
\mathsf{E}[X] = \sum_i x_i \; p_X(x_i) &= \sum_i x_i \; \sum_j \mathsf{P}(A_j) \; p_{X|A_j}(x_i) \\
&= \sum_i \sum_j x_i \; \mathsf{P}(A_j) \; p_{X|A_j}(x_i) \\
&= \sum_j \sum_i x_i \; \mathsf{P}(A_j) \; p_{X|A_j}(x_i) \\
&= \sum_j \mathsf{P}(A_j) \sum_i x_i \; p_{X|A_j}(x_i) \\
&= \sum_j \mathsf{P}(A_j) \; \mathsf{E}[X|A_j]
\end{aligned}
$$

□

3 Practice Problems

1. Compute the expected value and variance of the Binomial random variable with parameters n, p.

2. Compute the expected value and variance of the Geometric random variable with parameter $0 < p < 1$.

3. Compute the expected value and variance of the Poisson random variable with parameter $\lambda > 0$.

4. Give an example of a discrete random variable (support and probability mass function) with infinite mean. Give an example of a discrete random variable (support and probability mass function) with finite mean and infinite variance.

5. (Markov's inequality) Let X be a non-negative, discrete random variable with p.m.f. $\{p_X(x_i)\}$. Show that for $a > 0$

$$\mathsf{P}(X \geq a) \leq \frac{\mathsf{E}[X]}{a}$$

6. Show that $\text{var}(X) = \mathsf{E}[X^2] - \mathsf{E}[X]^2$. Now, prove that $\mathsf{E}[X^2] \geq \mathsf{E}[X]^2$.

7. (Chebychev's Inequality) Let X be a random variable with mean μ_X and variance σ_X^2. Using Markov's inequality, show that

$$\mathsf{P}(|X - \mu_X| \geq a) \leq \frac{\sigma_X^2}{a^2}$$

8. (Bertsekas, Solved exercise) Average Speed vs Average Time: If the weather is good (which happens with probability 0.6), Alice walks the 2 miles to class at a speed of $V = 5$ miles per hour, and otherwise rides her motorcycle at a speed of $V = 30$ miles per hour. What is the mean of the time T to get to class? (Answer: 4/15 hours)

9. (Ross) For a nonnegative integer-valued random variable N, show that

$$\mathsf{E}[N] = \sum_{i=1}^{\infty} \mathsf{P}(N \geq i)$$

10. (Ross) Four buses carrying 148 students from the same school arrive at a football stadium. The buses carry, respectively, 40, 33, 25, and 50 students. One of the students is randomly selected. Let X denote the number of students that were on the bus carrying the randomly selected student. One of the 4 bus drivers is also randomly selected. Let Y denote the number of students on her bus. Which of $\mathsf{E}[X]$ or $\mathsf{E}[Y]$ do you think is larger? Why?

11. (Ross) A sample of 3 items is selected at random from a box containing 20 items of which 4 are defective. Find the expected number of defective items in the sample.

3.1 Associated Problems

1. Define mean, median and mode of a collection of real numbers.

2. (Jensen's inequality) Let X be a non-negative, discrete random variable with p.m.f. $\{p_X(x_i)\}$. Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function. Then, show that
$$f(\mathsf{E}[X]) \leq \mathsf{E}[f(X)]$$

3. (Bertsekas, Solved exercise) Compute the expected value and variance of a geometric random variable using the total expectation theorem.

4. Prove the law of total variance.

5. Let $A_1, A_2, \cdots, A_n$ be a partition of Ω such that $\mathsf{P}(A_i) > 0$ for all i. Furthermore, for any B such that $\mathsf{P}(A_i B) > 0$, show that

$$\mathsf{E}[X|B] = \sum_i \mathsf{P}(A_i|B)\ \mathsf{E}[X|A_i B]$$

6. A coin that has probability of heads equal to p is tossed successively and independently until a head comes twice in a row or a tail comes twice in a row. Find the p.m.f. and the expected value of the number of tosses.

7. Describe negative binomial and hyper geometric random variables.

8. (Hajek) Suppose each corner of a three dimensional cube burns out with probability 0.001, independently of the other corners. Find an upper bound, using the union bound, on the probability that there exist two neighboring corners that both burn out.

9. (Bertsekas) Prove the inclusion-exclusion principle for a general collection of events $A_1, A_2, \cdots, A_n$ using indicator functions and expectations. (Hint: $\mathsf{P}(A_1 \cup \cdots \cup A_n) = \mathsf{E}[\mathbb{1}_{\{A_1 \cup \cdots \cup A_n\}}]$)

10. Read about St. Petersburg Paradox.