CSL003P1M : Probability and Statistics Lecture 11 (Expectation Of A Discrete Random Variable)

Sumit Kumar Pandey

September 28, 2021

Discrete Random Variables

Discrete Random Variables

A random variable that can take on at most a countable number of possible values is said to be discrete. For a discrete random variable X, we define the **probability mass function p(a)** of X by

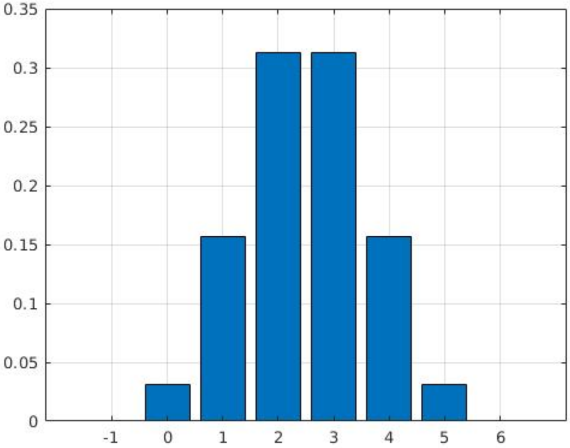
$$p(a) = P\{X = a\}$$

The probability mas function $p(a)$ is positive for at most a countable number of values of a. That is, if X assume one of the values $x_1, x_2 \ldots,$ then

$$\begin{array}{ll} p(x_i) \geq 0 & \text{for } i = 1, 2, \ldots \\ p(x) = 0 & \text{for all other values of } x \end{array}$$

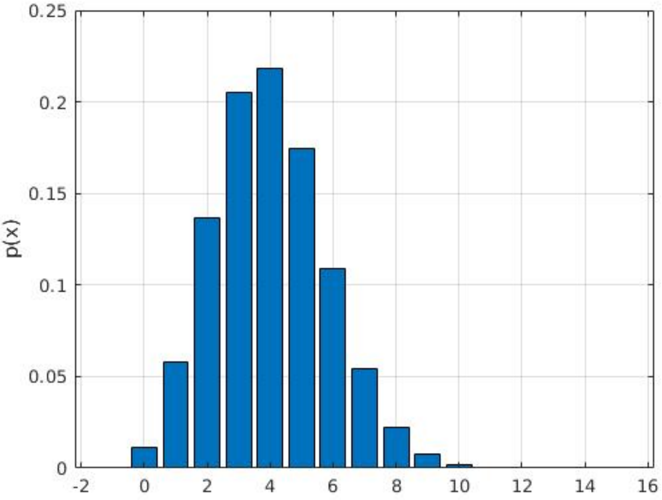
Since X must take on one of the values x_i, we have $\sum_{i=1}^{\infty} p(x_i) = 1$.

Example-1

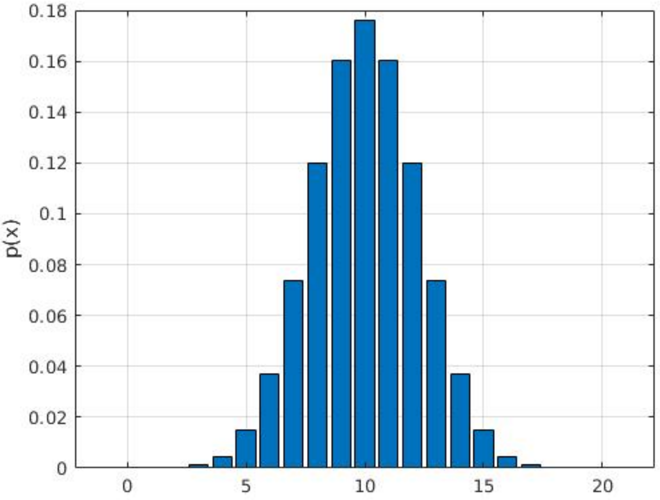


$$p(i), \;\; i = 0, 1, 2, 3, 4, 5$$

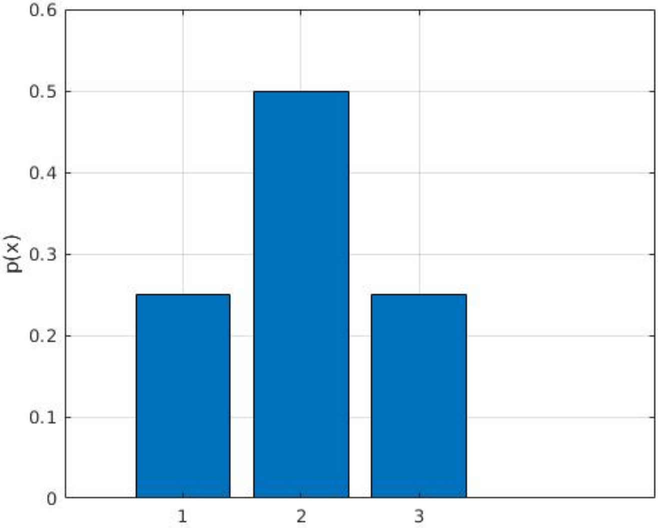
Example-2



Example-3



Example-4



Exercise-1

The probability mass function of a random variable X is given by $p(i) = c\lambda^i/i!$, $i = 0, 1, 2, \ldots,$ where λ is some positive value. Find

1. $P\{X = 0\}$ and
2. $P\{X > 2\}$.

Solution:

- Since $\sum_{i=0}^{\infty} p(i) = 1$, we have

$$c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = 1$$

- Because $e^x = \sum_{i=0}^{\infty} x^i/i!$, therefore

$$ce^{\lambda} = 1, \qquad \text{or } c = e^{-\lambda}.$$

Exercise-1

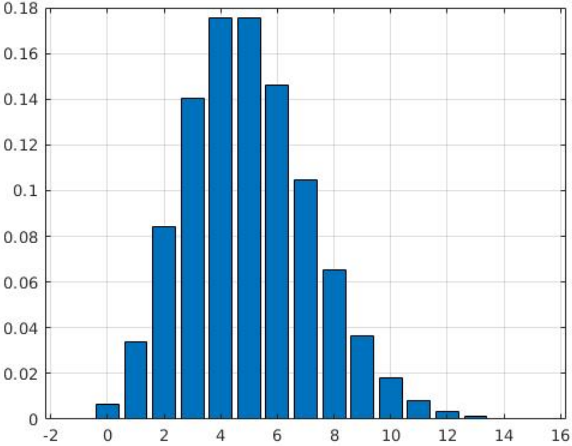
The probability mass function of a random variable X is given by $p(i) = c\lambda^i/i!$, $i = 0, 1, 2, \ldots$, where λ is some positive value. Find

1. $P\{X = 0\}$ and
2. $P\{X > 2\}$.

We have found $c = e^{-\lambda}$.

1. Hence,
$$P\{X = 0\} = e^{-\lambda}\lambda^0/0! = e^{-\lambda}.$$
2. $$\begin{aligned} P\{X > 2\} &= 1 - P\{X \leq 2\} \\ &= 1 - P\{X = 0\} - P\{X = 1\} - P\{X = 2\} \\ &= 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}. \end{aligned}$$

Exercise-1



$p(i)$ with parameter $\lambda = 5$.

Cumulative Distribution Function

The cumulative distribution function F can be expressed in terms of $p(a)$ by

$$F(a) = \sum_{\text{all } x \leq a} p(x)$$

If X is a discrete random variable whose possible values are $x_1, x_2, x_3, \ldots,$ where $x_1 < x_2 < x_3 < \cdots$, then the distribution function F of X is a **step function**.

That is, the value of F is constant in the intervals $\{x_{i-1}, x_i\}$ and then takes a step (or jump) of size $p(x_i)$ at x_i.

Example-5

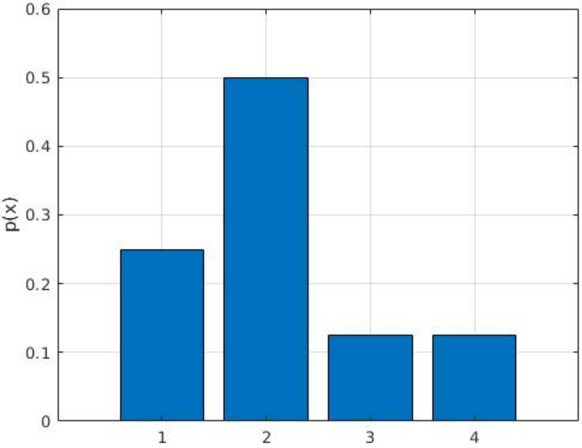
If X has a probability mass function given by

$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = \frac{1}{8}, \quad p(4) = \frac{1}{8}$$

then its cumulative distribution function is

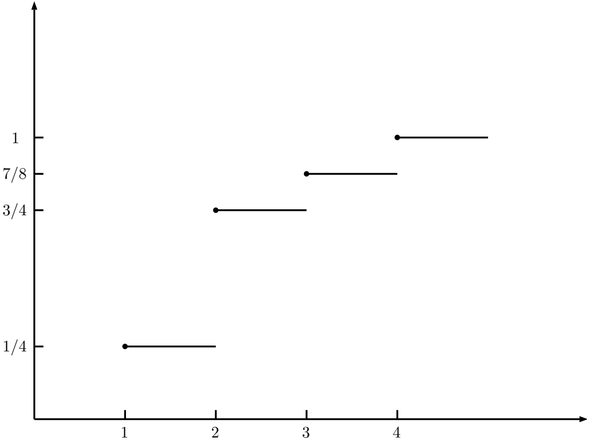
$$F(a) \quad = \quad \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \leq a \end{cases}$$

Example-5



$p(i)$ for $i = 1, 2, 3, 4.$

Example-5



$F(a)$.

Consider the following sequence of numbers:
9, 9, 1, 2, 3, 4, 2, 9, 8, 4, 1, 5, 7, 3, 0, 4, 0, 2, 1, 5, 3, 9, 0, 4, 3,
8, 6, 6, 4, 2, 0, 1, 1, 8, 3, 6, 5, 4, 8, 2, 9, 1, 7, 2, 8, 9, 1, 1, 9, 6
4, 5, 6, 2, 0, 1, 4, 2, 9, 8, 4, 5, 1, 3, 9, 1, 0, 7, 7, 2, 8, 3, 9, 1, 1,
2, 2, 6, 7, 9, 0, 7, 3, 8, 2, 9, 0, 1, 1, 6, 3, 8, 5, 6, 3, 2, 9, 3, 2, 1,

Calcuate the average.

Let "f_i" be the frequency of number "i", $i = 0, 1, 2, \ldots, 9$ and "n" be the total number of numbers. Let p_i be the probability that the number i occurs in the above sequence, then $p_i = f_i/n$.

The average is

$$
\begin{aligned}
& \frac{\sum_{i=0}^{9} i \cdot f_i}{n} \\
= & \sum_{i=0}^{9} i \cdot p_i.
\end{aligned}
$$

Expectation of a Random Variable

Expectation of a Random Variable

If X is a discrete random variable having a probability mass function $p(x)$, then the *expectation*, or the *expected value*, of X, denoted by $E[X]$, is defined by

$$E[X] = \sum_{x:p(x)>0} xp(x)$$

In words, the expected value of X is the weighted average of the possible values that X can take on, each value being weighted by the probability that X assumes it.

Example-6

Suppose the probability mass function of X is given by

$$p(0) = \frac{1}{2} = p(1)$$

then

$$E[X] = 0\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = \frac{1}{2}$$

is just the ordinary average of two possible values, 0 and 1, that X can assume.

Example-7

Suppose the probability mass function of X is given by

$$p(0) = \frac{1}{3}, \quad p(1) = \frac{2}{3}$$

then

$$E[X] = 0\left(\frac{1}{3}\right) + 1\left(\frac{2}{3}\right) = \frac{2}{3}$$

is a weighted average of two possible values, 0 and 1, where the value 1 is given twice as much weight as the value 0, since $p(1) = 2p(0)$.

Exercise-2

Find $E[X]$, where X is the outcome when we roll a fair die.

Solution:

- Since $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}$, we obtain

$$E[X] = 1\left(\frac{1}{6}\right)+2\left(\frac{1}{6}\right)+3\left(\frac{1}{6}\right)+4\left(\frac{1}{6}\right)+5\left(\frac{1}{6}\right)+6\left(\frac{1}{6}\right) = \frac{7}{2}.$$

Exercise-3

We say that I is an indicator variable for the event A if

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } \bar{A} \text{ occurs} \end{cases}$$

Find $E[I]$.

Solution:

- Since $p(1) = P(A)$, $p(0) = 1 - P(A)$, we have

$$E[I] = P(A).$$

  That is, the expected value of the indicator variable for the event A is equal to the probability that A occurs.

Exercise-4

A school class of 120 students is driven in 3 buses to a symphonic performance. There are 36 students in one of the buses, 40 in another, and 44 in the third bus. When the buses arrive, one of the 120 students is randomly chosen. Let X denote the number of students on the bus of that randomly chosen student, find $E[X]$.

Solution:

- Since the randomly chosen student is equally likely to be any of the 120 students, it follows that

$$P\{X = 36\} = \frac{36}{120}, \; P\{X = 40\} = \frac{40}{120}, \; P\{X = 44\} = \frac{44}{120}.$$

- Hence,

$$E[X] = 36\left(\frac{3}{10}\right) + 40\left(\frac{1}{3}\right) + 44\left(\frac{11}{30}\right) = \frac{1208}{30} \approx 40.2667.$$

Exercise-5

A contestant on a quiz show is presented with two questions, questions 1 and 2, which he is to attempt to answer in some order he chooses. If he decides to try question i first, then he will be allowed to go on to question j, $j \neq i$, only if his answer to question i is correct. If his initial answer is incorrect, he is not allowed to answer the other question. The contestant is to receive V_i dollars if he answers question i correctly, $i = 1, 2$. For instance he will receive $V_1 + V_2$ dollars if he answers both questions correctly. If the probability that he knows the answer to question i is P_i, $i = 1, 2$, which question should he attempt to answer first so as to maximize his expected winnings? Assume that the events E_i, $i = 1, 2$ that he knows the answer to question i are independent events.

Exercise-5

- If he attempts to answer question 1 first, then he will win

$$\begin{array}{ll} 0 & \text{with probability } 1 - P_1 \\ V_1 & \text{with probability } P_1(1 - P_2) \\ V_1 + V_2 & \text{with probability } P_1 P_2 \end{array}$$

- Hence, his expected winnings in this case will be

$$V_1 P_1(1 - P_2) + (V_1 + V_2) P_1 P_2.$$

- On the other hand, if he attempts to answer question 2 first, then he will win

$$\begin{array}{ll} 0 & \text{with probability } 1 - P_2 \\ V_2 & \text{with probability } P_2(1 - P_1) \\ V_1 + V_2 & \text{with probability } P_1 P_2 \end{array}$$

- Hence, his expected winnings will be

$$V_2 P_2(1 - P_1) + (V_1 + V_2) P_1 P_2.$$

Exercise-5

- Therefor, it is better to try question 1 first if

$$V_1P_1(1-P_1) \geq V_2P_2(1-P_1)$$

  or equivalently, if

$$\frac{V_1P_1}{1-P_1} \geq \frac{V_2P_2}{1-P_2}.$$

For example, if he is 60% percent of answering question 1, worth 200 dollars, correctly and he is 80% certain of answering question 2, worth 100 dollars, correctly, then he should attempt to answer question 2 first because

$$300 = \frac{(200)\cdot(0.6)}{0.4} \not\geq \frac{(100)\cdot(0.8)}{0.2} = 400.$$

Thank You