Probability and Statistics Exercise 2 - Discrete Random Variables

Solution 1

a) The sample space of this experiment is defined as S below:

$$S = \{\{H, H\}, \{H, T\}, \{T, H\}, \{T, T\}\}$$

b) We know that for a single unbiased coin, $P(\{H\}) = \frac{1}{2}$ and $P(\{T\}) = \frac{1}{2}$. Then the probability mass function of X is:

$$p_X(x) = P(X = x) = \begin{cases} \frac{1}{4} & x = 0 \\ \frac{1}{2} & x = 1 \\ \frac{1}{4} & x = 2 \end{cases}$$

c) The probability mass function of Y is:

$$p_Y(x) = P(Y = x) = \begin{cases} \frac{1}{4} & x = 1 \\ \frac{3}{4} & x = 3 \end{cases}$$

Solution 2

The probability mass function $p(x)$ is shown below:

x	p(x)
2	$\frac{1}{36}$
3	$\frac{1}{18}$
4	$\frac{1}{12}$
5	$\frac{1}{9}$
6	$\frac{5}{36}$
7	$\frac{1}{6}$
8	$\frac{5}{36}$
9	$\frac{1}{9}$
10	$\frac{1}{12}$
11	$\frac{1}{18}$
12	$\frac{1}{36}$

Solution 3

a) $P(\text{four heads}) = \frac{1}{16}$

b) $P(\text{three heads}) = \frac{1}{4}$

c) $P(\text{at least two heads}) = 1 - P(\text{one head} \cup \text{no heads}) = \frac{11}{16}$

d) $P(\text{not more than one head}) = \frac{5}{16}$

Solution 4

a) i. The probability mass function of X is:

$$p_X(x) = \begin{cases} \frac{3}{28} & x = 0 \\ \frac{15}{28} & x = 1 \\ \frac{5}{14} & x = 2 \end{cases}$$

ii. TBD

b) If the marble are drwan with replacement, then:

$$p_X(x) = \begin{cases} \frac{9}{64} & x = 0 \\ \frac{15}{32} & x = 1 \\ \frac{25}{64} & x = 2 \end{cases}$$

Solution 5

a) $P(\text{none pass}) = (1 - 0.4)^5 = 0.07776$

b) $P(\text{one passes}) = 5 \cdot (1 - 0.4)^4 \cdot 0.4 = 0.2592$

c) $P(\text{at least one passes}) = 1 - P(\text{none pass}) = 0.92224$

Solution 6

a) Let X be the random variable representing the number of passes in total. Then we can have a map $X : S \to \mathbb{R}$ where $X(s)$ denotes the number of people passing the exam. (e.g. $X(\text{no people pass} = 0)$, $X(\text{one person passes}) = 1$, $X(\text{two people pass}) = 2)$

i. The expected number of passes is $E(X) = np = 110 \cdot 0.8 = 88$

ii. The standard deviation is $\sigma^2 = np(1 - p) = 110 \cdot 0.8 \cdot 0.2 = 17.6$

b) i. The expected number of graduates is $E(X) = np = 11000 \cdot 0.8 = 8800$

ii. The standard deviation of the number of graduates is $\sigma^2 = np(1 - p) = 11000 \cdot 0.8 \cdot 0.2 = 1760$

Solution 7

This problem involves binomial distribution of the random variable X which denotes the number of companies that will make a claim. The probability mass function is $p_X(x) = \binom{5}{x} 0.2^x \cdot 0.8^{5-x}$

a) $P(\text{all companies will claim}) = p_X(5) = 0.2^5 = \frac{1}{3125}$

b) $P(\text{at least three companies will claim}) = p_X(3) + p_X(4) + p_X(5) = \frac{181}{3125}$

c) $P(\text{only two will claim}) = p_X(2) = \frac{128}{625}$

d) $P(\text{at least one will not claim}) = 1 - P(\text{all companies will claim}) = 1 - 0.2^5 = \frac{3124}{3125}$

Solution 8

a) $E(X) = np = 100 \cdot 0.9 = 90$, $\sigma = \sqrt{np(1-p)} = \sqrt{100 \cdot 0.9 \cdot 0.1} = 3$, $\gamma = \frac{1-2p}{\sqrt{np(1-p)}} = -\frac{4}{15}$

b) $E(X) = np = 100 \cdot 0.7 = 70$, $\sigma = \sqrt{np(1-p)} = \sqrt{100 \cdot 0.7 \cdot 0.3} = \sqrt{21}$, $\gamma = \frac{1-2p}{\sqrt{np(1-p)}} = -\frac{2}{5\sqrt{21}}$

c) $E(X) = np = 100 \cdot 0.5 = 50$, $\sigma = \sqrt{np(1-p)} = \sqrt{100 \cdot 0.5 \cdot 0.5} = 5$, $\gamma = \frac{1-2p}{\sqrt{np(1-p)}} = 0.$

d) $E(X) = np = 1000 \cdot 0.9 = 900$, $\sigma = \sqrt{np(1-p)} = \sqrt{1000 \cdot 0.9 \cdot 0.1} = 3\sqrt{10}$, $\gamma = \frac{1-2p}{\sqrt{np(1-p)}} = -\frac{4}{15\sqrt{10}}$

e) $E(X) = np = 1000 \cdot 0.7 = 700$, $\sigma = \sqrt{np(1-p)} = \sqrt{1000 \cdot 0.7 \cdot 0.3} = \sqrt{210}$, $\gamma = \frac{1-2p}{\sqrt{np(1-p)}} = -\frac{2}{5\sqrt{210}}$

f) $E(X) = np = 1000 \cdot 0.5 = 500$, $\sigma = \sqrt{np(1-p)} = \sqrt{100 \cdot 0.5 \cdot 0.5} = 5$, $\gamma = \frac{1-2p}{\sqrt{np(1-p)}} = 0$

The standard deviation and skewness will change square-root-proportionally along with n. When the probability of the binomial distribution becomes 0.5, the skewness is 0, meaning that the distribution is centered. When the skewness becomes negative, the graph is left-skewed; when the skewness value becomes positive, the graph is right-skewed.

Solution 9

The mean and standard deviation of the number of adequate batteries in the box can be calculated using the three sub-means and sub-sd, all of which have binomial distribution:

$$E(X_1) = np = 300 \cdot 0.9 = 270$$
$$E(X_2) = np = 150 \cdot 0.5 = 75$$
$$E(X_3) = np = 50 \cdot 0.4 = 20$$

Therefore, $E(X) = E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 270 + 75 + 20 = 365$

Standard deviation follows that
$\sigma_X = \sqrt{Var(X)} = \sqrt{Var(X_1 + X_2 + X_3)} = \sqrt{Var(X_1) + Var(X_2) + Var(X_3)} = \sqrt{27 + 37.5 + 12} = \sqrt{76.5}$

Solution 10

$E(X) = 2 \cdot 0.4 + 4 \cdot 0.6 + 5 \cdot 0.7 + 7 \cdot 0.8 + 2 \cdot 0.9 = 14.1$

$\sigma_X = \sqrt{Var(X)} = \sqrt{0.48 + 0.96 + 1.05 + 1.12 + 0.18} = \sqrt{3.79}$

Solution 11

a) This problem forms a geometric distribution such that $p_X(x) = p(1-p)^{x-1}$. Therefore, the average number of times he will have to try to use a machine until success is the mean of this distribution $E(X) = \frac{1}{p} = 2.5$.

b) $p_X(1) = 0.4$

c) $P(\text{success on three different occasions}) = 0.4^3 = 0.064$

Solution 12

a) This will basically form a binomial distribution such that $E(X) = np$ and $Var(X) = np(1-p)$.

b) If they have different parameters, then $E(X) = p_1 + p_2 + p_3 + \ldots + p_n$ and
$Var(X) = Var(X_1) + Var(X_2) + Var(X_3) + \ldots + Var(X_n) = p_1(1-p_1) + p_2(1-p_2) + p_3(1-p_3) + \ldots + p_n(1-p_n)$

c) If they are not independent, then the expectation(mean) and variance will not be the same as presented in b).

Solution 13

a) $G(z) = e^{-\lambda(1-z)}$

b) $G(z) = \frac{z^1+z^2+z^3+\ldots+z^N}{N}$

c) answer delayed

Solution 14

a) Let $E = \text{random carton selected that Molly will like}$, and $F_1 = \text{carton being produced at Lancashire}$, $F_2 = \text{carton being produced at Derbyshire}$, and $F_3 = \text{carton being produced at Yorkshire}$. Then by the partition rule, we have:

$$\begin{aligned} P(E) &= P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + P(E|F_3)P(F_3) \\ &= 0.95 \cdot 0.5 + 0.4 \cdot 0.2 + 0.25 \cdot 0.3 \\ &= 0.63 \end{aligned}$$

b) Using Bayes Theorem, the probabilities that it was produced at each of the plants are as follows:

$$\begin{aligned} P(F_1|\overline{E}) &= \frac{P(F_1)P(\overline{E}|F_1)}{P(\overline{E})} = \frac{P(F_1)(1-P(E|F_1))}{1-P(E)} = \frac{0.5 \cdot (1-0.95)}{1-0.63} = 0.0676 \\ P(F_2|\overline{E}) &= \frac{P(F_2)P(\overline{E}|F_2)}{P(\overline{E})} = \frac{P(F_2)(1-P(E|F_2))}{1-P(E)} = \frac{0.2 \cdot (1-0.4)}{1-0.63} = 0.3243 \\ P(F_3|\overline{E}) &= \frac{P(F_3)P(\overline{E}|F_3)}{P(\overline{E})} = \frac{P(F_3)(1-P(E|F_3))}{1-P(E)} = \frac{0.3 \cdot (1-0.25)}{1-0.63} = 0.6081 \end{aligned}$$

c) Let $X = \text{number of cartons in the box that Molly will like}$ and
$X_1 = \text{number of cartons produced at Lancashire that Molly will like}$, and
$X_2 = \text{number of cartons produced at Derbyshire that Molly will like}$, and
$X_3 = \text{number of cartons produced at Yorkshire that Molly will like}$. We know that
$E(X) = E(X_1 + X_2 + X_3)$.

Since the numbers of cartons produced at each plant are exactly in proportion to the production percentages given, we can now obtain the mean and standard deviation of X_1, X_2, and X_3. All of these random variables follow the binomial distribution pattern where there are many identical independent Bernoulli trials.

$$\begin{aligned} E(X_1) &= 250 \cdot 0.95 = 237.5 \\ E(X_2) &= 100 \cdot 0.4 = 40 \\ E(X_3) &= 150 \cdot 0.25 = 37.5 \end{aligned}$$

Therefore, $E(X) = E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 237.5 + 40 + 37.5 = 315$. This means that the expectation of the number of cartons in the box that Molly will like is 315.

The same follows standard deviation calculation:

$$\sigma_X = \sqrt{\sigma_{X_1}^2 + \sigma_{X_2}^2 + \sigma_{X_3}^2} = \sqrt{(250 \cdot 0.95 \cdot 0.05) + (100 \cdot 0.4 \cdot 0.6) + (150 \cdot 0.25 \cdot 0.75)} = \sqrt{64} = 8$$