STA2001 Assignment 1 Solution

(2.1-3) For each of the following determine the constant c so that $f(x)$ satisfies the conditions of being a pmf for a random variable X, and then depict each pmf as a line graph:

(a) $f(x) = x/c, \quad x = 1, 2, 3, 4.$

(b) $f(x) = cx, \quad x = 1, 2, 3, \ldots, 10.$

(c) $f(x) = c(1/4)^x, \quad x = 1, 2, 3, \ldots.$

(d) $f(x) = c(x+1)^2, \quad x = 0, 1, 2, 3.$

(e) $f(x) = x/c, \quad x = 1, 2, 3, \ldots, n.$

(f) $f(x) = \frac{c}{(x+1)(x+2)}, \quad x = 0, 1, 2, 3, \ldots$

Hint: In part(f), write $f(x) = 1/(x+1) - 1/(x+2)$

Solution:

Using the condition $\sum_{x \in S} f(s) = 1$, it is easy to determine the value of c (you should work out this in details by yourself).

(a) $c = 10$

(b) $c = 1/55$

(c) $c = 3$

(d) $c = 1/30$

(e) $c = 2n(n+1)$

(f) $c = 1$

The line graphs are omitted.

(2.2-1) Find $E(X)$ for each of the distributions given in Exercise 2.1-3.

Solution:

(a) $E(X) = \sum_{x=1}^{4} xf(x) = \sum_{x=1}^{4} x \times \frac{x}{10} = \frac{1}{10} \times 1 + \frac{2}{10} \times 2 + \frac{3}{10} \times 3 + \frac{4}{10} \times 4 = 3$

(b) $E(X) = \sum_{x=1}^{10} xf(x) = \sum_{x=1}^{10} x \times \frac{x}{55} = \frac{1}{55} \times 1 + \frac{2}{55} \times 2 + \cdots + \frac{10}{55} \times 10 = 7$

(c)

$$E(X) = \sum_{x=1}^{\infty} xf(x) = \sum_{x=1}^{\infty} 3x \times \frac{1}{4^x} = \frac{4}{3}$$

The following are detailed steps in calculating the expectation (i.e. the sum of this series). First, we define the partial sum

$$S_n = \sum_{x=1}^{n} 3x \left(\frac{1}{4}\right)^x = 3\left[\left(\frac{1}{4}\right) + 2\left(\frac{1}{4}\right)^2 + \cdots + n\left(\frac{1}{4}\right)^n\right]$$

Multiply $\frac{1}{4}$ then we have

$$\frac{1}{4}S_n = 3\left[\left(\frac{1}{4}\right)^2 + 2\left(\frac{1}{4}\right)^3 + \cdots + n\left(\frac{1}{4}\right)^{n+1}\right]$$

Then,

$$S_n - \frac{1}{4}S_n = \frac{3}{4}S_n = 3\left[\left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)^2 + \cdots + \left(\frac{1}{4}\right)^n - n\left(\frac{1}{4}\right)^{n+1}\right] = 3\left[\frac{1}{4} \cdot \frac{1-(1/4)^n}{1-1/4}\right] - n\left(\frac{1}{4}\right)^{n+1}$$

Note that we obtain the result of $1/4S_n$ by using the sum formula of geometric progression.

Now we have an equation with respect to S_n, and solve for S_n we obtain

$$S_n = \frac{4}{3} - \frac{n+4/3}{4^n}$$

So the expectation (i.e. the sum of series) is given by

$$E(x) = \lim_{n\to\infty} S_n = \frac{4}{3}$$

(d) $E(X) = \sum_{x=0}^{3} xf(x) = \sum_{x=0}^{3} \frac{1}{30}x(x+1)^2 = \frac{1}{30} \times 0 \times 1 + \frac{1}{30} \times 1 \times 2^2 + \frac{1}{30} \times 2 \times 3^2 + \frac{1}{30} \times 3 \times 4^2 = \frac{7}{3}$

(e) $E(X) = \sum_{x=1}^{\infty} xf(x) = \sum_{x=1}^{\infty} x\frac{2x}{n(n+1)} = \frac{2(1^2+2^2+3^2+\cdots+n^2)}{n(n+1)} = \frac{(2n+1)}{3}$

Note that we have used the formula $\sum_{x=1}^{n} x^2 = \frac{1}{6}n(n+1)(2n+1)$ in deriving the result.

(f)

$$\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} xf(x) \\
&= \sum_{x=0}^{\infty} \frac{x}{(x+1)(x+2)} \\
&= \sum_{x=0}^{\infty} \frac{x}{x+1} - \frac{x}{x+2} \\
&= 0 - 0 + \frac{1}{2} - \frac{1}{3} + \frac{2}{3} - \frac{2}{4} + \cdots \\
&= \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots\right) - \left(\lim_{x\to\infty} \frac{x}{x+2}\right) \\
&= +\infty
\end{aligned}$$

where we have used the fact that a harmonic series is divergent and goes to infinity.

Therefore, for this random variable X, its mean $E(X)$ does not exist.

(2.2-5) In Example 2.2-1 let $Z = u(X) = X^3$.

(a) Find the pmf of Z, say $h(z)$.

(b) Find $E(Z)$.

(c) How much, on average, can the young man expect to win on each play if he charges 10 dollars per play?

Solution:

(a) From example 2.2-1, we know that $S_X = \{x : x = 1, 2, 3\}$, and thus $S_Z = \{z : z = u(x), x \in S_X\} = \{1, 8, 27\}$.

As the function x^3 is a bijective function, the pmf can be easily obtained by finding the inverse function

$$h(z) = P(Z = z) = P\left(X^3 = z\right) = P\left(X = z^{\frac{1}{3}}\right) = \frac{4 - z^{\frac{1}{3}}}{6}, \quad \forall z \in S_Z$$

(b) $E(Z) = \sum_{z \in S_Z} zh(z) = 1 \cdot h(1) + 8 \cdot h(8) + 27 \cdot h(27) = \frac{23}{3}$

(c) $10 - E(Z) = \frac{7}{3}$ dollars.

(2.3-2) For each of the following distributions, find $\mu = E(X)$, $E[X(X-1)]$, and $\sigma^2 = E[X(X-1)] + E(X) - \mu^2$

(a) $f(x) = \frac{3!}{x!(3-x)!}\left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, \quad x = 0,1,2,3.$

(b) $f(x) = \frac{4!}{x!(4-x)!}\left(\frac{1}{2}\right)^4, \quad x = 0,1,2,3,4.$

Solution:

(a)

$$\begin{aligned}
\mu &= E(X) \\
&= \sum_{x=0}^{3} x \frac{3!}{x!(3-x)!}\left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x} \\
&= 3\sum_{x=1}^{3} \frac{2!}{(x-1)!(3-x)!}\left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x} \\
&= 3\sum_{k=0}^{2} \frac{2!}{(k)!(2-k)!}\left(\frac{1}{4}\right)^{k+1} \left(\frac{3}{4}\right)^{2-k} \\
&= \frac{3}{4}\sum_{k=0}^{2} \frac{2!}{(k)!(2-k)!}\left(\frac{1}{4}\right)^{k} \left(\frac{3}{4}\right)^{2-k} \\
&= \frac{3}{4} \times \left(\frac{1}{4} + \frac{3}{4}\right)^2 \\
&= \frac{3}{4}
\end{aligned}$$

where we have applied binomial theorem in the 6th equality.

$$\begin{aligned}
E[X(X-1)] &= \sum_{x=0}^{3} x(x-1)\frac{3!}{x!(3-x)!}\left(\frac{1}{3}\right)^x \left(\frac{3}{4}\right)^{3-x} \\
&= \sum_{x=2}^{3} x(x-1)\frac{3!}{x!(3-x)!}\left(\frac{1}{3}\right)^x \left(\frac{3}{4}\right)^{3-x} \\
&= 2 \times 1 \times \frac{3!}{2!1!}\left(\frac{1}{3}\right)^2 \left(\frac{3}{4}\right) + 3 \times 2 \times \frac{3!}{3!0!}\left(\frac{1}{3}\right)^3 \left(\frac{3}{4}\right)^0 \\
&= \frac{3}{8}
\end{aligned}$$

Finally,

$$\sigma^2 = E[X(X-1)] + E(X) - \mu^2 = \frac{9}{16}$$

(b) Note that we could write

$$f(x) = \frac{4!}{x!(4-x)!}\left(\frac{1}{2}\right)^4 = \frac{4!}{x!(4-x)!}\left(\frac{1}{\sqrt{2}}\right)^x \left(\frac{1}{\sqrt{2}}\right)^{4-x}$$

So we could compute all required quantities in a similar way (same as (a)), and the results are

$$\mu = 2, \quad E[X(X-1)] = 3, \quad \sigma^2 = 1$$

(2.3-4) Let μ and σ^2 denote the mean and variance of the random variable X. Determine $E[(X-\mu)/\sigma]$ and $E\left\{[(X-\mu)/\sigma]^2\right\}$.

Solution:

$$E[(X-\mu)/\sigma] = (1/\sigma)[E(X)-\mu] = (1/\sigma)(\mu-\mu) = 0$$
$$E\{[(X-\mu)/\sigma]^2\} = 1/\sigma^2 \cdot E\left[(X-\mu)^2\right] = 1/\sigma^2 \cdot \sigma^2 = 1$$

(2.4-5). In a lab experiment involving inorganic syntheses of molecular precursors to organometallic ceramics, the final step of a five-step reaction involves the formation of a metalCmetal bond. The probability of such a bond forming is $p = 0.20$. Let X equal the number of successful reactions out of $n = 25$ such experiments.

(a) Find the probability that X is at most 4.
(b) Find the probability that X is at least 5.
(c) Find the probability that X is equal to 6.
(d) Give the mean, variance, and standard deviation of X.

Solution:

X follows a binomial distribution with parameter $p = 0.2$, that is, $X \sim B(25, 0.2)$. Therefore we know its pdf is given by

$$P(X = x) = \binom{25}{x} 0.2^x 0.8^{25-x}$$

(a) $P(X \le 4) = \sum_{x=0}^{4} \binom{25}{x} 0.2^x 0.8^{25-x} = 0.42$

(b) $P(X \ge 5) = 1 - P(X \le 4) = 0.58$

(c) $P(X = 6) = \binom{25}{6}(0.2)^6(0.8)^{19} = 0.16$

(d) Using the property of a binomial distribution, we have

$$E(x) = np = 25 \times 0.2 = 5$$

$$\sigma^2 = np(1-p) = 25 \times 0.2 \times 0.8 = 4$$

$$\sigma = \sqrt{4} = 2$$

(2.4-7). Suppose that 2000 points are selected independently and at random from the unit square $\{(x, y) : 0 \leq x < 1, 0 \leq y < 1\}$. Let W equal the number of points that fall into $A = \{(x, y) : x^2 + y^2 < 1\}$.
(a) How is W distributed?
(b) Give the mean, variance, and standard deviation of W.
(c) What is the expected value of $W/500$?

Solution:

(a) Let $S = \{(x, y) : 0 \leq x < 1, 0 \leq y < 1\}$. We know that the region $A \subset S$, and both regions A and S are defined on the first quadrant. So probability that one of the 2000 points fall in the region A, is $\pi/4$, as the region A is essentially a quarter-circle (i.e. 1/4 of a unit circle). As each point is generated independently, so the random variable W follows a binomial distribution, that is,

$$W \sim b(2000, \pi/4)$$

(b) Using the properties of a binomial distribution, we have

$$E(W) = np = 2000 \times \frac{\pi}{4} = 500\pi$$

$$\text{Var}(W) = np(1 - p) = 2000 \times \frac{\pi}{4}\left(1 - \frac{\pi}{4}\right) \approx 337$$

$$\sigma = \sqrt{337} = 18.36$$

(c) Note that expectation is a linear operator, we have

$$E\left(\frac{W}{500}\right) = \frac{1}{500}E(W) = \pi$$

(2.4-20). (i) Give the name of the distribution of X (if it has a name), (ii) find the values of μ and σ^2, and (iii) calculate $P(1 \le X \le 2)$ when the moment-generating function of X is given by

(a) $M(t) = (0.3 + 0.7e^t)^5$
(b) $M(t) = \frac{0.3e^t}{1-0.7e^t}, \quad t < -\ln(0.7)$
(c) $M(t) = 0.45 + 0.55e^t$
(d) $M(t) = 0.3e^t + 0.4e^{2t} + 0.2e^{3t} + 0.1e^{4t}$
(e) $M(t) = \sum_{x=1}^{10}(0.1)e^{tx}$

Solution:

In the following, we identity the distribution if it's possible and then use the corresponding properties to compute the mean and variance. However, you should be able to derive these quantities by using the mgf only, or using the definition only (i.e. using more fundamental approach).

(a) It's a binomial distribution, that is, $X \sim b(5, 0.7)$

$$\begin{aligned} \mu &= np = 5 \times 0.7 = 3.5 \\ \sigma^2 &= np(1-p) = 3.5 \times 0.3 = 1.05 \\ P(1 \le X \le 2) = P(X=1) + P(X=2) &= \binom{5}{1} \times 0.7 \times 0.3^4 + \binom{5}{2} \times 0.7^2 \times 0.3^3 = 0.1607 \end{aligned}$$

(b) It's a geometric distribution with parameter $p = 0.3$.

$$\begin{aligned} \mu &= \frac{1}{p} = \frac{10}{3} \\ \sigma^2 &= \frac{1-p}{p^2} = \frac{0.7}{0.3^2} = \frac{70}{9} \\ P(1 \le X \le 2) &= P(X=1) + P(X=2) = 0.3 + 0.3 \times 0.7 = 0.51 \end{aligned}$$

(c) It's a Bernoulli distribution with parameter $p = 0.55$.

$$\begin{aligned} \mu &= p = 0.55 \\ \sigma^2 &= p(1-p) = 0.55 \times 0.45 = 0.2475 \\ P(1 \le X \le 2) &= P(X=1) = 0.55 \end{aligned}$$

(d) This discrete random variable doesn't have a name.

The pmf of this random variable can be obtained by matching coefficients of the given moment generating function (compare the coefficients of the closed form and its definition), as it uniquely determines a distribution. Once we obtain the pmf $f(x)$, we have

$$\begin{aligned} \mu &= \sum_{x=1}^{4} x \cdot f(x) = 1 \times 0.3 + 2 \times 0.4 + 3 \times 0.2 + 4 \times 0.1 = 2.1 \\ \sigma^2 &= E\left[(X-\mu)^2\right] = 1.1^2 \times 0.3 + 0.1^2 \times 0.4 + 0.9^2 \times 0.2 + 1.9^2 \times 0.1 = 0.89 \\ P(1 \le X \le 2) &= P(X=1) + P(X=2) = 0.3 + 0.4 = 0.7 \end{aligned}$$

(e) It's a discrete uniform distribution on the set $\{1, 2, \ldots, 10\}$. Similar with question (d), we compare the coefficients and obtain the pmf. Then we could compute

$$\begin{aligned} \mu &= \sum_{x=1}^{10} x \cdot 0.1 = 5.5 \\ \sigma^2 &= E\left[(X-\mu)^2\right] = 0.1 \times \sum_{x=1}^{10}(x-5.5)^2 = 8.25 \\ P(1 \le X \le 2) &= 0.1 + 0.1 = 0.2 \end{aligned}$$

(2.5-10). In 2012, Red Rose tea randomly began placing 1 of 12 English porcelain miniature figurines in a 100-bag box of the tea, selecting from 12 nautical figurines.

(a) On the average, how many boxes of tea must be purchased by a customer to obtain a complete collection consisting of the 12 nautical figurines?
(b) If the customer uses one tea bag per day, how long can a customer expect to take, on the average, to obtain a complete collection?

Solution:

(a) Let X_k be the number of box of tea we need to purchase in order to have a nautical figurine that is different from the k nautical figurines we already got.

Now, X_k can be viewed as a geometric random variable with parameter $p_k = \frac{12-k}{12}$, for $k = 0, 1, \cdots, 11$. Note that in this case $X_0 = 1$ with probability 1, so it's also a degenerate random variable.

Let X be the total number of box we need to purchase to get the whole collection, so the mean is given by

$$E(X) = E\left(\sum_{k=0}^{11} X_k\right) = \sum_{k=0}^{11} E\left(X_k\right) = \sum_{k=0}^{11} \frac{1}{p_k} = \sum_{k=0}^{11} \frac{12}{12-k} = \frac{86021}{2310} = 37.2385$$

(b) As each box contains 100 bags of tea, so we have

$$100 \times E(X) \approx 3724 \text{ days}$$

which is approximately equal to 10.2 years.

(2.6-11). An airline always overbooks if possible. A particular plane has 95 seats on a flight in which a ticket sells for \$300. The airline sells 100 such tickets for this flight.
(a) If the probability of an individual not showing up is 0.05, assuming independence, what is the probability that the airline can accommodate all the passengers who do show up?
(b) If the airline must return the \$300 price plus a penalty of \$400 to each passenger that cannot get on the flight, what is the expected payout (penalty plus ticket refund) that the airline will pay?

Solution:

(a) Let X be the number of passengers not showing up, so $X \sim b(100, 0.05)$. Therefore,

$$\begin{aligned}
&P(\text{the airline can accommodate all the passengers who do show up}) \\
&= P(X \geq 5) \\
&= 1 - P(X \leq 4) \\
&= 1 - \sum_{k=0}^{4} \binom{100}{k} 0.05^k 0.95^{100-k} \\
&= 0.564
\end{aligned}$$

(b) Note that the company must pay penalty and refund if $X = 0, 1, 2, 4$, that is, there are some extra passengers showing up.

By the definition of expectation, we could compute

$$\begin{aligned}
E(\text{payout}) &= 700 \times [1 \times P(X = 4) + 2 \times P(X = 3) + 3 \times P(X = 2) + 4 \times P(X = 1) + 5 \times P(X = 0)] \\
&= 598.56
\end{aligned}$$

Note: it may easier to compute the results of (a) and (b) if we use Poisson distribution to approximate this binomial distribution, and the approximated results are actually quite close to the true results, that is, 0.56 for (a) and 613.9 for (b).