Ferhat Abbas-Sétif 1 University
Faculty of Sciences
Department Informatics

Solved exercises

Exercise 1 Classify each random variable as either discrete or continuous.
1- The number of arrivals at an emergency room between midnight and 6 :00 a.m.
2- The weight of a box of cereal labelled "18 ounces."
3- The duration of the next outgoing telephone call from a business office.
4- The number of kernels of popcorn in a 1-pound container.
5- The number of applicants for a job.

Solution
1- discrete
2- continuous
3- continuous
4- discrete
5- discrete

Exercise 2 I roll two dice and observe two numbers X and Y.
a- Find S_X and S_Y and the PMFs of X and Y.
b- Compute $P(X = 2, Y = 6)$
c- Compute $P(X > 3|Y = 2)$
d- Let $Z = X + Y$. Find the range and PMF of Z
e- $P(X = 4|Z = 8)$.
Solution
a- $S_X = S_Y = \{1, 2, 3, 4, 5, 6\}$

$$P_X(k) = \begin{cases} \frac{1}{6} & \text{for } k = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}$$

Similarly,

$$P_Y(k) = \begin{cases} \frac{1}{6} & \text{for } k = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}$$

b- Since X and Y are independent random variables, then

$$\begin{aligned} P(X = 2, Y = 6) &= P(X = 2)P(Y = 6) \\ &= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} \end{aligned}$$

c- Since X and Y are independent, then

$$\begin{aligned} P(X > 3|Y = 2) &= \frac{P(X > 3, Y = 2)}{P(Y = 2)} \\ &= \frac{P(X > 3)P(Y = 2)}{P(Y = 2)} \\ &= P_X(4) + P_X(5) + P_X(6) \\ &== \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}. \end{aligned}$$

d- Since we have $S_Z = \{2, 3, 4, ..., 12\}$. Thus, PMF of Z is as follows

$$\begin{aligned} P_Z(2) &= P(Z = 2) = P(X = 1, Y = 1) \\ &= P(X = 1)P(Y = 1) \text{ (since } X \text{ and } Y \text{ are independent)} \\ P_Z(3) &= P(Z = 3) = P(X = 1, Y = 2) + P(X = 2, Y = 1) \\ &= P(X = 1)P(Y = 2) + P(X = 2)P(Y = 1) \\ &= \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{18} \end{aligned}$$

We continuous this process we get $P_Z(4) = 3 \cdot \frac{1}{36} = \frac{1}{12}, P_Z(5) = 1/9, P_Z(6) = 5/36, P_Z(7) = 1/6, P_Z(8) = 5/36, P_Z(9)1/9, P_Z(10) = 1/12, P_Z(11) = 1/18, P_Z(12) = 1/36$
d- Since $Z = X + Y$, it is clear that Z depends on the values of X. So,

$$\begin{aligned} P(X = 4|Z = 8) &= \frac{P(X = 4, Z = 8)}{P(Z = 8)} \\ &= \frac{P(X = 4, Y = 4)}{P(Z = 8)} \\ &= \frac{P(X = 4)P(Y = 4)}{P(Z = 8)} \text{ (since } X \text{ and } Y \text{ are independent)} \\ &= \frac{\frac{1}{6} \cdot \frac{1}{6}}{\frac{5}{36}} = \frac{1}{5} \end{aligned}$$

Exercise 3 Let X be a continuous random variable with PDF

$$f_X(x) = \begin{cases} 4x^3 & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Compute $P(X \leq \frac{2}{3}|X > \frac{1}{3})$.
Solution

$$\begin{aligned} P(X \leq \frac{2}{3}|X > \frac{1}{3}) &= \frac{P(\frac{1}{3} < X \leq \frac{2}{3})}{P(X > \frac{1}{3})} \\ &= \frac{\int_{\frac{1}{3}}^{\frac{2}{3}} 4x^3 dx}{\int_{\frac{1}{3}}^{1} 4x^3 dx} = \frac{3}{16}. \end{aligned}$$

Exercise 4 Suppose that the length of a phone call in minutes is an exponential random variable with average length 10 minutes.
a- What is the probability of your phone call being more than 10 minutes ?
b- What is the probability of your phone call being Between 10 and 20 minutes ?
Solution
a- We have Here $\lambda = \frac{1}{10}$, thus

$$P(X > 10) = e^{-0.1 \times 10} = e^{-1} = 0.368.$$

b-

$$P(10 < X < 20) = F_X(20) - F_X(10) = e^{-1} - e^{-2} = 0.233.$$

Exercise 5 Suppose the life of a smart-phone has exponential distribution with mean life of 4 years. Let X denote the life of a phone (or time until it dies). Given that the phone has lasted 3 years, what is the probability that it will 5 more years.
Solution
We have $\lambda = 4$, then

$$\begin{aligned} P(X > 5 + 3|X > 3) &= \frac{P(X > 8)}{P(X > 3)} \\ &= \frac{e^{-4 \times 8}}{e^{-4 \times 3}} = e^{-4 \times 5} = P(X > 5)) \end{aligned}$$

Exercise 6 Suppose that the time (in minutes) required to check out a book at the library can be represented by an exponentially distributed random variable with parameter $\lambda = \frac{2}{11}$.
a- What is the probability that it will take at least 5 minutes to check out a book ?
b- What is the probability that it will take at least 11 minutes to check out a book given that you have

already waited for 6 minutes ?
Solution
a-

$$P(X > 5) = e^{-\frac{10}{11}}$$

b- Since the exponential distribution is a memoryless one then

$$\begin{aligned} P(X > 11|X > 6) &= P(X > 6 + 5|X > 6) \\ &= P(X > 5) = e^{-\frac{10}{11}} \end{aligned}$$

Exercise 6 Let's say that the time (measured in minutes) needed to complete the checkout process for a book at the library follows an exponential distribution with parameter $\lambda = 2/11$.
1- What is the probability that it will take at least 5 minutes to check out a book ?
2- What is the probability that it will take at least 11 minutes to check out a book given that you have already waited for 6 minutes ?
Solution
1-

$$\mathbb{P}(X > 5) = e^{-\frac{10}{11}}$$

2- Using the memoryless propriety

$$\begin{aligned} \mathbb{P}(X > 11 \mid X > 6) &= \mathbb{P}(X > 6 + 5 \mid X > 6) \\ &= \mathbb{P}(X > 5) = e^{-\frac{10}{11}} \end{aligned}$$

Exercise 7 An insurance company insures a large number of homes. The insured value, denoted as X, of any randomly chosen home, is presumed to follow to a distribution characterized by its density function.

$$f_X(x) = \begin{cases} \frac{8}{x^3} & x > 2 \\ 0 & \text{otherwise} \end{cases}$$

1- Knowing that a randomly selected home is insured for at most 4, find the probability that it is insured for less than 3.
2- Knowing that a randomly selected home is insured for at least 3, find the probability that it is insured for less than 4.
Solution
1- Since we have

$$\mathbb{P}(X < 4) = \int_2^4 \frac{8}{x^3} dx = -\left.\frac{4}{x^2}\right|_2^4 = -\frac{1}{4} + 1 = \frac{3}{4}$$

and

$$\mathbb{P}(X < 3) = \mathbb{P}(2 < X < 3) = \int_2^3 \frac{8}{x^3} dx = -\left.\frac{4}{x^2}\right|_2^3 = -\frac{4}{9} + 1 = \frac{5}{9}.$$

Using the Conditional probability definition, we get

$$\mathbb{P}(X < 3 \mid X < 4) = \frac{\mathbb{P}(X < 3)}{\mathbb{P}(X < 4)} = \frac{\frac{5}{9}}{\frac{3}{4}} = \frac{20}{27} \approx 0.74074074.$$

2- Since we have

$$\begin{aligned} \mathbb{P}(X > 3) &= \int_3^\infty \frac{8}{x^3} dx = -\left.\frac{4}{x^2}\right|_3^\infty = \frac{4}{9} \\ \mathbb{P}(3 < X < 4) &= \int_3^4 \frac{8}{x^3} dx = -\left.\frac{4}{x^2}\right|_3^4 = \frac{7}{36}. \end{aligned}$$

Then, by the conditional probability definition, it follows that

$$\mathbb{P}(X < 4 \mid X > 3) = \frac{\mathbb{P}(3 < X < 4)}{\mathbb{P}(X > 3)} = \frac{\frac{7}{36}}{\frac{4}{9}} = \frac{7}{16} \approx 0.4375$$

Exercise 8 A company sets prices for its hurricane insurance based on the following assumptions :
1- At most one hurricane can occur in any given calendar year.
2- The probability of a hurricane in any given calendar year is 0.05.
3- The occurrences of hurricanes in different calendar years are mutually independent.
Using these assumptions, calculate the probability that there are at most 2 hurricanes in a 20-year period.
Solution
From the assumptions it seems that $X \sim Binomial(20, 0.05)$. Therefore,

$$\begin{aligned}\mathbb{P}(X \leqslant 2) &= C(20,0)(0.05)^0(0.95)^{20} + C(20,1)(0.05)^1(0.95)^{19} + C(20,2)(0.05)^2(0.95)^{18} \\ &= 0.9245.\end{aligned}$$

Exercise 9 The expected number of of typographical errors on a page of the new Harry Potter book is 0.2. Explain what assumptions you used to find the probability that the next page you read contains : :
1- 0 typographical errors ?
2- At least 2 typographical errors ?
Solution

Considering that each word has a small probability of being a typo, the number of typos should be approximately distributed according to a Poisson distribution. Hence
1- $e^{-0.2}$
2- $1 - e^{-0.2} - 0.2e^{-0.2} = 1 - 1.2e^{-0.2}$.

Exercise 10 A certain type of storage battery lasts, on average, 3 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, calculate the probability that a given battery will last less than 2.3 years.
Solution
We have $Z = \frac{2.3-3}{0.5} = -1.4$, therefore

$$P(X < 2.3) = P(Z < -1.4) = 0.0808$$

Exercise 11 In an industrial process, the diameter of a ball bearing is an important measurement. The buyer sets specifications for the diameter to be 3.0 ± 0.01 cm. The implication is that no part falling outside these specifications will be accepted. It is known that in the process the diameter of a ball bearing has a normal distribution with mean $\mu = 3$ and standard deviation $\sigma = 0.005$. On average, how many manufactured ball bearings will be scrapped ?
Solution
We have $Z_1 = \frac{2.99-3}{0.005} = -2$ and $Z_2 = \frac{3.01-3}{0.005} = 2$. Hence

$$\begin{aligned}P(2.99 < X < 3.01) &= P(-2 < Z < 2) \\ &= P(Z < 2) - P(Z < -2) = 2P(Z < 2) - 1 = 2(0.9772) - 1 = 0.9544\end{aligned}$$

Consequently, it is expected that approximately 4.56% of manufactured ball bearings will be scrapped on average.

Exercise 12 Gauges are utilized to reject all components for which a certain dimension does not fall within the specification of $1.50 \pm d$. It is established that this measurement follows a normal distribution with a mean of 1.5 and a standard deviation of 0.2. Find the value d such that the specifications "cover" 95% of the measurements.
Solution
From standard normal distribution table we have

$$P(-1.96 < Z < 1.96) = 0.95.$$

It follows that

$$1.96 = \frac{(1.50 + d) - 1.50}{0.2}$$

from which we obtain,

$$1.96 = \frac{(1.50 + d) - 1.50}{0.2}$$

Exercise 13 A specific machine manufactures electrical resistors with a mean resistance of 40 ohms and a standard deviation of 2 ohms. It is assumed that the resistance follows a normal distribution and can be measured with any degree of precision, what percentage of resistors will have a resistance exceeding 43 ohms ?

Solution

$$z = \frac{43 - 40}{2} = 1.5$$

Then,

$$P(X > 43) = P(Z > 1.5) = 1 - P(Z < 1.5) = 1 - 0.9332 = 0.0668.$$

Therefore, 6.68% of the resistors will have a resistance that exceed 43 ohms.

Exercise 14 This week, a financial regulator from the Federal Reserve (FED) will assess two banks. For each evaluation, the regulator will choose with equal probability between two different stress tests. Failing the first stress test incurs a penalty of 10000$ for a bank, while failing the second test results in a 5000$ penalty. The probability of the first bank failing either test is 0.4, while independently, the second bank has a 0.5 probability of failing either test. Let X represent the total fees collected by the regulator after evaluating both banks. Determine the cumulative distribution function of X.

Solution

The random variable X can take the values $0, 5000, 10000, 15000$ and 20000 depending on which test was applied to each bank, and if the bank fails the evaluation or not. Denote by B_i the event that the i-th bank fails and by T_i the event that test i applied. Then

$$\begin{aligned}
&\mathbb{P}\left(T_1\right)=\mathbb{P}\left(T_2\right)=0.5, \mathbb{P}\left(B_1\right)=\mathbb{P}\left(B_1 \mid T_1\right)=\mathbb{P}\left(B_1 \mid T_2\right)=0.4 \\
&\mathbb{P}\left(B_2\right)=\mathbb{P}\left(B_2 \mid T_1\right)=\mathbb{P}\left(B_2 \mid T_2\right)=0.5
\end{aligned}$$

Since banks and tests are independent we have

$$\begin{aligned}
&\mathbb{P}(X=0)=\mathbb{P}\left(B_1^c \cap B_2^c\right)=\mathbb{P}\left(B_1^c\right) \cdot \mathbb{P}\left(B_2^c\right)=0.6 \cdot 0.5=0.3, \\
&\mathbb{P}(X=5000)=\mathbb{P}\left(B_1\right) \mathbb{P}\left(T_2\right) \mathbb{P}\left(B_2^c\right)+\mathbb{P}\left(B_1^c\right) \mathbb{P}\left(B_2\right) \mathbb{P}\left(T_2\right)=0.25, \\
&\mathbb{P}(X=10000)=\mathbb{P}\left(B_1\right) \mathbb{P}\left(T_1\right) \mathbb{P}\left(B_2^c\right)+\mathbb{P}\left(B_1\right) \mathbb{P}\left(T_2\right) \mathbb{P}\left(B_2\right) \mathbb{P}\left(T_2\right)+\mathbb{P}\left(B_1^c\right) \mathbb{P}\left(B_2\right) \mathbb{P}\left(T_1\right)=\frac{3}{10} \\
&\mathbb{P}(X=15000)=\mathbb{P}\left(B_1\right) \mathbb{P}\left(T_1\right) \mathbb{P}\left(B_2\right) \mathbb{P}\left(T_2\right)+\mathbb{P}\left(B_1\right) \mathbb{P}\left(T_2\right) \mathbb{P}\left(B_2\right) \mathbb{P}\left(T_1\right)=0.1 \\
&\mathbb{P}(X=20000)=\mathbb{P}\left(B_1\right) \mathbb{P}\left(T_1\right) \mathbb{P}\left(B_2\right) \mathbb{P}\left(T_1\right)=0.05.
\end{aligned}$$

Therefore, the CDF is given as follows

$$F_X(x)= \begin{cases}0 & x<0 \\ 0.3 & 0 \leqslant x<5000 \\ 0.55 & 5000 \leqslant x<10000 \\ 0.85 & 10000 \leqslant x<15000 \\ 0.95 & 15000 \leqslant x<20000 \\ 1 & x \geqslant 20000\end{cases}$$