Question Bank with Solutions (Midterm)

MATH 394: Probability I

Instructor
Arman Jahangiri

Term
Summer 2026

University of Washington

Department of Mathematics

Problem 1

How many different seven-place license plates are possible if the first two places are for letters and the other five are for digits?

(a) Repetition is allowed.

(b) No letter or digit can be repeated on a single license plate.

Solution

There are 26 choices for each letter and 10 choices for each digit.

(a) With repetition allowed,
$$26^2\,10^5 = 67{,}600{,}000.$$

(b) Without repetition, the two letter positions can be filled in $26 \cdot 25$ ways and the five digit positions in $10 \cdot 9 \cdot 8 \cdot 7 \cdot 6$ ways. Thus
$$26 \cdot 25 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = 196{,}560{,}000.$$

Problem 2

When all letters are used, how many different letter arrangements can be made from the letters in each word?

(a) FLUKE

(b) PROPOSE

(c) MISSISSIPPI

(d) ARRANGE

Solution

For n letters with repeated-letter multiplicities $n_1, \ldots, n_r$, the number of distinct arrangements is

$$\frac{n!}{n_1! \cdots n_r!}.$$

(a) FLUKE has five distinct letters:

$$5! = 120.$$

(b) PROPOSE has seven letters, with P repeated twice and O repeated twice:

$$\frac{7!}{2!\,2!} = 1260.$$

(c) MISSISSIPPI has 11 letters, with I repeated four times, S four times, P twice, and M once:

$$\frac{11!}{4!\,4!\,2!} = 34{,}650.$$

(d) ARRANGE has seven letters, with A repeated twice and R repeated twice:

$$\frac{7!}{2!\,2!} = 1260.$$

Problem 3

(a) Show that
$$\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n.$$

(b) Simplify
$$\sum_{k=0}^{n} \binom{n}{k} x^k.$$

Solution

The binomial theorem states
$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

(a) Set $a = 1$ and $b = 2$:
$$\sum_{k=0}^{n} \binom{n}{k} 2^k = (1+2)^n = 3^n.$$

(b) Set $a = 1$ and $b = x$:
$$\sum_{k=0}^{n} \binom{n}{k} x^k = (1+x)^n.$$

Problem 4

If 12 people are to be divided into three committees of respective sizes 3, 4, and 5, how many divisions are possible?

Solution

Choose 3 of the 12 people for the first committee, then 4 of the remaining 9 for the second; the final 5 form the third:

$$\binom{12}{3}\binom{9}{4}\binom{5}{5} = \frac{12!}{3!\,4!\,5!} = 27{,}720.$$

Problem 5

1. Expand $(3x^2 + y)^5$.
2. Without considering part (a), i.e., without using the expansion, write down the third term.

Solution

(1) By the binomial theorem,

$$(3x^2 + y)^5 = \sum_{k=0}^{5} \binom{5}{k} (3x^2)^{5-k} y^k.$$

Therefore,

$$(3x^2 + y)^5 = 243x^{10} + 405x^8 y + 270x^6 y^2 + 90x^4 y^3 + 15x^2 y^4 + y^5.$$

(2) The $(k+1)^{th}$ term is given by $\binom{n}{k} a^k b^{n-k}$. Thus, the third term is given by $\binom{5}{2}(y)^2(3x^2)^{5-2}$.

Problem 6

Prove that
$$\binom{n+m}{r} = \binom{n}{0}\binom{m}{r} + \binom{n}{1}\binom{m}{r-1} + \cdots + \binom{n}{r}\binom{m}{0}.$$
Hint: Consider a group of n men and m women and count groups of size r.

Solution

Count the number of ways to form a group of r people from n men and m women.
Directly, there are
$$\binom{n+m}{r}$$
such groups.
Alternatively, classify the group according to the number k of men selected. If exactly k men are selected, then $r-k$ women are selected, giving
$$\binom{n}{k}\binom{m}{r-k}$$
possibilities. Summing over all feasible values of k gives
$$\binom{n+m}{r} = \sum_{k=0}^{r} \binom{n}{k}\binom{m}{r-k}.$$
Terms with impossible choices are interpreted as zero.

Problem 7

(a) Give a combinatorial proof of
$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}.$$
Interpret both sides as the number of ways to select a nonempty committee from n people and choose its chairperson.

(b) Verify for $n = 1, 2, 3, 4, 5$, and then prove combinatorially, that
$$\sum_{k=1}^{n} \binom{n}{k} k^2 = 2^{n-2} n(n+1).$$
Interpret both sides as the number of ways to select a committee, its chairperson, and its secretary, where the chairperson and secretary may be the same person.

(c) Prove that
$$\sum_{k=1}^{n} \binom{n}{k} k^3 = 2^{n-3} n^2(n+3).$$

Solution

(a) Count pairs consisting of a nonempty committee and its chairperson.

If the committee has size k, choose its members in $\binom{n}{k}$ ways and its chairperson in k ways. This gives
$$\sum_{k=1}^{n} k\binom{n}{k}.$$

Alternatively, first choose the chairperson in n ways. Each of the other $n-1$ people may independently be included or excluded, giving 2^{n-1} choices. Hence
$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}.$$

Count selections of a committee, a chairperson, and a secretary, where the two officers may coincide. A committee of size k contributes $k^2\binom{n}{k}$, so the total is the left-hand side.

Split into two cases.

If the chairperson and secretary are the same, choose that person in n ways and choose any subset of the remaining $n-1$ people:
$$n2^{n-1}.$$

If they are different, choose the chairperson and secretary as an ordered pair in $n(n-1)$ ways, and choose any subset of the remaining $n-2$ people:
$$n(n-1)2^{n-2}.$$

Therefore,
$$\sum_{k=1}^{n} \binom{n}{k} k^2 = n2^{n-1} + n(n-1)2^{n-2} = 2^{n-2} n(n+1).$$

(b) Count selections of a committee together with three ordered offices, where a person may hold more than one office. A committee of size k contributes $k^3\binom{n}{k}$.

Classify according to the number of distinct officeholders.

One distinct officeholder:

$$n2^{n-1}.$$

Exactly two distinct officeholders: choose which one of the three offices is held by the person who holds only one office (3 choices), then choose the two distinct people in order ($n(n-1)$ choices), and choose any subset of the remaining $n-2$ people:

$$3n(n-1)2^{n-2}.$$

Three distinct officeholders:

$$n(n-1)(n-2)2^{n-3}.$$

Adding,

$$\begin{aligned}\sum_{k=1}^{n}\binom{n}{k}k^3 &= n2^{n-1} + 3n(n-1)2^{n-2} + n(n-1)(n-2)2^{n-3} \\ &= 2^{n-3}n^2(n+3).\end{aligned}$$

Problem 8

Show that, for $n > 0$,

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i} = 0.$$

Solution

Apply the binomial theorem to $(1-1)^n$:

$$(1-1)^n = \sum_{i=0}^{n}\binom{n}{i}1^{n-i}(-1)^i = \sum_{i=0}^{n}(-1)^i\binom{n}{i}.$$

For $n > 0$, the left-hand side is $0^n = 0$. Therefore,

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i} = 0.$$

Problem 9

A total of 28% of American males smoke cigarettes, 7% smoke cigars, and 5% smoke both cigars and cigarettes.

(a) What percentage smoke neither cigars nor cigarettes?

(b) What percentage smoke cigars but not cigarettes?

Solution

Let C be the event of smoking cigarettes and G the event of smoking cigars. Then

$$P(C) = 0.28, \qquad P(G) = 0.07, \qquad P(C \cap G) = 0.05.$$

(a)

$$P(C \cup G) = 0.28 + 0.07 - 0.05 = 0.30.$$

Thus

$$P((C \cup G)^c) = 1 - 0.30 = 0.70.$$

The answer is 70%.

(b)

$$P(G \cap C^c) = P(G) - P(G \cap C) = 0.07 - 0.05 = 0.02.$$

The answer is 2%.

Problem 10

Two cards are chosen uniformly at random without replacement from a standard deck of 52 playing cards. What is the probability that they

(a) are both aces?

(b) have the same rank?

Solution

There are $\binom{52}{2}$ unordered pairs of cards.

(a) There are $\binom{4}{2}$ pairs of aces:

$$P(\text{both aces}) = \frac{\binom{4}{2}}{\binom{52}{2}} = \frac{6}{1326} = \frac{1}{221}.$$

(b) For each of the 13 ranks, choose 2 of the 4 suits:

$$P(\text{same rank}) = \frac{13\binom{4}{2}}{\binom{52}{2}} = \frac{78}{1326} = \frac{1}{17}.$$

Problem 11

Let E, F, and G be three events. Write expressions for the events that

(a) only E occurs;

(b) both E and G, but not F, occur;

(c) at least one of the events occurs;

(d) at least two of the events occur;

(e) all three events occur;

(f) none of the events occurs;

(g) at most one of the events occurs;

(h) at most two of the events occur;

(i) exactly two of the events occur;

(j) at most three of the events occur.

Solution

(a) Only E: $E \cap F^c \cap G^c$.

(b) E and G, but not F: $E \cap F^c \cap G$.

(c) At least one: $E \cup F \cup G$.

(d) At least two: $(E \cap F) \cup (E \cap G) \cup (F \cap G)$.

(e) All three: $E \cap F \cap G$.

(f) None: $E^c \cap F^c \cap G^c = (E \cup F \cup G)^c$.

(g) At most one:

$$(E^c \cap F^c \cap G^c) \cup (E \cap F^c \cap G^c) \cup (E^c \cap F \cap G^c) \cup (E^c \cap F^c \cap G).$$

Equivalently,

$$\big[(E \cap F) \cup (E \cap G) \cup (F \cap G)\big]^c.$$

(h) At most two:

$$(E \cap F \cap G)^c.$$

(i) Exactly two:

$$(E \cap F \cap G^c) \cup (E \cap F^c \cap G) \cup (E^c \cap F \cap G).$$

(j) At most three:

$$\Omega,$$

because there are only three events.

Problem 12

If $P(E) = 0.9$ and $P(F) = 0.8$, show that $P(E \cap F) \geq 0.7$. More generally, prove

$$P(E \cap F) \geq P(E) + P(F) - 1.$$

Solution

The addition rule gives

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

Because $P(E \cup F) \leq 1$,

$$P(E) + P(F) - P(E \cap F) \leq 1,$$

and therefore

$$P(E \cap F) \geq P(E) + P(F) - 1.$$

For $P(E) = 0.9$ and $P(F) = 0.8$,

$$P(E \cap F) \geq 0.9 + 0.8 - 1 = 0.7.$$

Problem 13

Prove that

$$P(E \cap F^c) = P(E) - P(E \cap F).$$

Solution

The event E is the disjoint union

$$E = (E \cap F) \cup (E \cap F^c).$$

Hence, by additivity,

$$P(E) = P(E \cap F) + P(E \cap F^c).$$

Rearranging gives

$$P(E \cap F^c) = P(E) - P(E \cap F).$$

Problem 14

A laboratory blood test has sensitivity 0.95: when a person has a certain disease, the test is positive with probability 0.95. The false-positive probability is 0.01: when a person is healthy, the test is positive with probability 0.01. If 0.5% of the population has the disease, find the probability that a person has the disease given that the test result is positive.

Solution

Let D be the event that the person has the disease and $+$ the event that the test is positive. Then

$$P(D) = 0.005, \quad P(D^c) = 0.995, \quad P(+ \mid D) = 0.95, \quad P(+ \mid D^c) = 0.01.$$

By Bayes' rule,

$$\begin{aligned} P(D \mid +) &= \frac{P(+ \mid D)P(D)}{P(+ \mid D)P(D) + P(+ \mid D^c)P(D^c)} \\ &= \frac{(0.95)(0.005)}{(0.95)(0.005) + (0.01)(0.995)} \\ &= \frac{0.00475}{0.01470} = \frac{95}{294} \approx 0.3231. \end{aligned}$$

Thus, given a positive result, the probability of having the disease is about 32.3%.

Problem 15

Let $A \subseteq B$. Express the following probabilities as simply as possible:

$$P(A \mid B), \qquad P(A \mid B^c), \qquad P(B \mid A), \qquad P(B \mid A^c).$$

State any conditions needed for the conditional probabilities to be defined.

Solution

Because $A \subseteq B$, we have $A \cap B = A$, $A \cap B^c = \varnothing$, and $B \cap A = A$.
Provided $P(B) > 0$,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}.$$

Provided $P(B^c) > 0$,

$$P(A \mid B^c) = \frac{P(A \cap B^c)}{P(B^c)} = 0.$$

Provided $P(A) > 0$,

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = 1.$$

Provided $P(A^c) > 0$,

$$P(B \mid A^c) = \frac{P(B \cap A^c)}{P(A^c)} = \frac{P(B) - P(A)}{1 - P(A)}.$$

Problem 16

Two fair dice are rolled, and X is the product of the two outcomes. Compute $P(X = i)$ for $i = 1, \ldots, 36$.

Solution

The 36 ordered outcomes are equally likely. The nonzero probabilities are

i	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
$36P(X = i)$	1	2	2	3	2	4	2	1	2	4	2	1	2	2	2	1	2	1

Thus

$$P(X = i) = \begin{cases} \dfrac{1}{36}, & i \in \{1, 9, 16, 25, 36\}, \\ \dfrac{1}{18}, & i \in \{2, 3, 5, 8, 10, 15, 18, 20, 24, 30\}, \\ \dfrac{1}{12}, & i = 4, \\ \dfrac{1}{9}, & i \in \{6, 12\}, \\ 0, & \text{otherwise.} \end{cases}$$

Problem 17

Suppose the distribution function of X is

$$F(b) = \begin{cases} 0, & b < 0, \\ \dfrac{b}{4}, & 0 \le b < 1, \\ \dfrac{1}{2} + \dfrac{b-1}{4}, & 1 \le b < 2, \\ \dfrac{11}{12}, & 2 \le b < 3, \\ 1, & b \ge 3. \end{cases}$$

(a) Find $P(X = i)$ for $i = 1, 2, 3$.

(b) Find

$$P\left(\frac{1}{2} < X < \frac{3}{2}\right).$$

Solution

For any point a,

$$P(X = a) = F(a) - F(a^-),$$

the size of the jump of the CDF at a.

(a) At 1,

$$P(X = 1) = F(1) - F(1^-) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

At 2,

$$P(X = 2) = F(2) - F(2^-) = \frac{11}{12} - \frac{3}{4} = \frac{1}{6}.$$

At 3,

$$P(X = 3) = F(3) - F(3^-) = 1 - \frac{11}{12} = \frac{1}{12}.$$

(b) Since there is no point mass at $1/2$ or $3/2$,

$$\begin{aligned} P\left(\frac{1}{2} < X < \frac{3}{2}\right) &= F\left(\frac{3}{2}\right) - F\left(\frac{1}{2}\right) \\ &= \left(\frac{1}{2} + \frac{(3/2) - 1}{4}\right) - \frac{1/2}{4} \\ &= \frac{5}{8} - \frac{1}{8} = \frac{1}{2}. \end{aligned}$$

Problem 18

Suppose

$$F(b) = \begin{cases} 0, & b < 0, \\ \frac{1}{2}, & 0 \le b < 1, \\ \frac{3}{5}, & 1 \le b < 2, \\ \frac{4}{5}, & 2 \le b < 3, \\ \frac{9}{10}, & 3 \le b < 3.5, \\ 1, & b \ge 3.5. \end{cases}$$

Calculate the probability mass function of X.

Solution

The probability at each support point is the jump of the CDF:

$$p_X(x) = P(X = x) = F(x) - F(x^-).$$

Therefore,

$$P(X = 0) = \frac{1}{2},$$

$$P(X = 1) = \frac{3}{5} - \frac{1}{2} = \frac{1}{10},$$

$$P(X = 2) = \frac{4}{5} - \frac{3}{5} = \frac{1}{5},$$

$$P(X = 3) = \frac{9}{10} - \frac{4}{5} = \frac{1}{10},$$

and

$$P(X = 3.5) = 1 - \frac{9}{10} = \frac{1}{10}.$$

Hence

$$p_X(x) = \begin{cases} \frac{1}{2}, & x = 0, \\ \frac{1}{10}, & x = 1, \\ \frac{1}{5}, & x = 2, \\ \frac{1}{10}, & x = 3, \\ \frac{1}{10}, & x = 3.5, \\ 0, & \text{otherwise}. \end{cases}$$

The probabilities sum to

$$\frac{1}{2} + \frac{1}{10} + \frac{1}{5} + \frac{1}{10} + \frac{1}{10} = 1.$$

Problem 19

The expected number of typographical errors on a page of a certain magazine is 0.2. Assuming that the number of errors on a page has a Poisson distribution, find the probability that the next page contains

(a) no typographical errors;

(b) two or more typographical errors.

Solution

Let X denote the number of typographical errors on the next page. Under the Poisson assumption,

$$X \sim \text{Pois}(0.2).$$

Thus,

$$P(X = k) = e^{-0.2}\frac{(0.2)^k}{k!}.$$

(a)
$$P(X = 0) = e^{-0.2} \approx 0.8187.$$

(b) Using the complement,

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - e^{-0.2} - (0.2)e^{-0.2} \\ &= 1 - 1.2e^{-0.2} \\ &\approx 0.0175. \end{aligned}$$

The expected value alone does not uniquely determine these probabilities. The calculation requires the additional assumption that the number of errors follows a Poisson distribution.

Problem 20

The monthly worldwide average number of airplane crashes involving commercial airlines is 3.5. Assuming a Poisson model, find the probability that there will be

(a) at least two such accidents in the next month;

(b) at most one such accident in the next month.

Solution

Let X denote the number of accidents during the next month. Under the Poisson assumption,

$$X \sim \text{Pois}(3.5).$$

Therefore,

$$P(X = k) = e^{-3.5}\frac{3.5^k}{k!}.$$

(a)

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - e^{-3.5} - 3.5e^{-3.5} \\ &= 1 - 4.5e^{-3.5} \\ &\approx 0.8641. \end{aligned}$$

(b)

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) \\ &= e^{-3.5} + 3.5e^{-3.5} \\ &= 4.5e^{-3.5} \\ &\approx 0.1359. \end{aligned}$$

Notice that the answers in parts (a) and (b) are complements.

Problem 21

Approximately 80,000 marriages took place in the state of New York last year. Estimate the probability that, for at least one of these couples,

(a) both partners were born on April 30;

(b) both partners celebrate their birthdays on the same day of the year.

State the assumptions used.

Solution

Assume that:

- birthdays are uniformly distributed among the 365 days of the year;
- leap days are ignored;
- the birthdays of the two partners are independent;
- different couples may be treated independently.

Let $n = 80{,}000$.

(a) For a particular couple, the probability that both partners were born on April 30 is

$$p = \left(\frac{1}{365}\right)^2.$$

Thus, the exact probability that at least one couple has this property is

$$\begin{aligned} P(\text{at least one}) &= 1 - (1-p)^{80{,}000} \\ &= 1 - \left(1 - \frac{1}{365^2}\right)^{80{,}000}. \end{aligned}$$

Because n is large and p is small, a Poisson approximation with

$$\lambda = np = \frac{80{,}000}{365^2} \approx 0.60049$$

gives

$$P(\text{at least one}) \approx 1 - e^{-0.60049} \approx 0.4515.$$

(b) For a particular couple, condition on the birthday of the first partner. The probability that the second partner has the same birthday is

$$p = \frac{1}{365}.$$

Therefore,

$$P(\text{at least one matching couple}) = 1 - \left(1 - \frac{1}{365}\right)^{80{,}000}.$$

Using a Poisson approximation,

$$\lambda = \frac{80{,}000}{365} \approx 219.18,$$

so

$$P(\text{at least one matching couple}) \approx 1 - e^{-219.18} \approx 1.$$

Problem 22

Compare the Poisson approximation with the exact binomial probability in each of the following cases:

(a) $P(X = 2)$, where $n = 8$ and $p = 0.1$;

(b) $P(X = 9)$, where $n = 10$ and $p = 0.95$;

(c) $P(X = 0)$, where $n = 10$ and $p = 0.1$;

(d) $P(X = 4)$, where $n = 9$ and $p = 0.2$.

Solution

If $X \sim \text{Bin}(n, p)$, then the exact probability is

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

The usual Poisson approximation uses

$$Y \sim \text{Pois}(\lambda), \qquad \lambda = np,$$

and

$$P(X = k) \approx P(Y = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

(a) Here $n = 8$, $p = 0.1$, $k = 2$, and $\lambda = 0.8$. The exact probability is

$$\begin{aligned} P(X = 2) &= \binom{8}{2} (0.1)^2 (0.9)^6 \\ &\approx 0.1488. \end{aligned}$$

The Poisson approximation is reasonably accurate:

$$\begin{aligned} P(Y = 2) &= e^{-0.8} \frac{0.8^2}{2!} \\ &\approx 0.1438. \end{aligned}$$

(b) Here $n = 10$, $p = 0.95$, $k = 9$, and $\lambda = 9.5$. The exact probability is

$$\begin{aligned} P(X = 9) &= \binom{10}{9} (0.95)^9 (0.05) \\ &\approx 0.3151. \end{aligned}$$

The direct Poisson approximation is poor because $p = 0.95$ is not small.

$$\begin{aligned} P(Y = 9) &= e^{-9.5} \frac{9.5^9}{9!} \\ &\approx 0.1300. \end{aligned}$$

A better approach is to let $Z = 10 - X$, the number of failures. Then $Z \sim \text{Bin}(10, 0.05)$, and $X = 9$ is equivalent to $Z = 1$. Approximating Z by Pois(0.5) gives the following, which is much closer to the exact value:

$$P(X = 9) = P(Z = 1) \approx e^{-0.5}(0.5) \approx 0.3033,$$

(c) Here $n = 10$, $p = 0.1$, $k = 0$, and $\lambda = 1$. The exact probability is

$$P(X = 0) = (0.9)^{10} \approx 0.3487.$$

The Poisson approximation is reasonably accurate: $P(Y = 0) = e^{-1} \approx 0.3679$.

(d) Here $n = 9$, $p = 0.2$, $k = 4$, and $\lambda = 1.8$. The exact probability is

$$\begin{aligned} P(X = 4) &= \binom{9}{4}(0.2)^4(0.8)^5 \\ &\approx 0.0661. \end{aligned}$$

The Poisson approximation is

$$\begin{aligned} P(Y = 4) &= e^{-1.8}\frac{1.8^4}{4!} \\ &\approx 0.0723. \end{aligned}$$

Problem 23

The rate of a certain event in a state is one event per 100,000 inhabitants per month.

(a) Find the probability that a city of 400,000 inhabitants experiences eight or more events in a given month.

(b) Find the probability that at least two months during a twelve-month year have eight or more events.

(c) Counting the present month as month 1, find the probability that the first month having eight or more events is month i, where $i \geq 1$.

State the assumptions used.

Solution

Assume that monthly event counts are independent and identically distributed Poisson random variables and that the rate remains constant over time.
For a city of 400,000 inhabitants, the expected monthly number is

$$\lambda = 400{,}000\left(\frac{1}{100{,}000}\right) = 4.$$

Thus, if X is the number of events in a given month,

$$X \sim \text{Pois}(4).$$

(a) Let $q = P(X \geq 8)$. Then

$$\begin{aligned} q &= 1 - P(X \leq 7) \\ &= 1 - \sum_{k=0}^{7} e^{-4}\frac{4^k}{k!} \\ &\approx 0.05113. \end{aligned}$$

(b) Let Y be the number of months, among the next 12 months, that have at least eight events. Under the independence assumption,

$$Y \sim \text{Bin}(12, q),$$

where $q \approx 0.05113$.

Therefore,

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y = 0) - P(Y = 1) \\ &= 1 - (1-q)^{12} - 12q(1-q)^{11} \\ &\approx 0.1229. \end{aligned}$$

(c) Let T denote the first month with at least eight events. Each month is a success with probability q, independently of other months. Therefore,

$$T \sim \text{Geom}(q),$$

using the convention that T is the trial number of the first success.

Hence,

$$P(T = i) = (1 - q)^{i-1} q, \qquad i = 1, 2, \ldots,$$

where

$$q = 1 - \sum_{k=0}^{7} e^{-4} \frac{4^k}{k!} \approx 0.05113.$$

Numerically, $P(T = i) \approx (0.94887)^{i-1}(0.05113)$.

Problem 24

The expected number of typographical errors on a page of a certain magazine is 0.2. Assuming that the number of errors on a page has a Poisson distribution, find the probability that the next page contains

(a) no typographical errors;

(b) two or more typographical errors.

Solution

Let X denote the number of typographical errors on the next page. Under the Poisson assumption,

$$X \sim \text{Pois}(0.2).$$

Thus,

$$P(X = k) = e^{-0.2}\frac{(0.2)^k}{k!}.$$

(a)

$$P(X = 0) = e^{-0.2} \approx 0.8187.$$

(b) Using the complement,

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - e^{-0.2} - (0.2)e^{-0.2} \\ &= 1 - 1.2e^{-0.2} \\ &\approx 0.0175. \end{aligned}$$

The expected value alone does not uniquely determine these probabilities. The calculation requires the additional assumption that the number of errors follows a Poisson distribution.

Problem 25

A random variable X is said to have the Yule–Simon distribution if

$$P(X = n) = \frac{4}{n(n+1)(n+2)}, \qquad n = 1, 2, \ldots$$

(a) Show that this is a probability mass function.

(b) Show that $E[X] = 2$.

(c) Show that $E[X^2] = \infty$.

The proposed probabilities are nonnegative. It remains to verify that they sum to one.

(a) First observe that

$$\frac{1}{n(n+1)(n+2)} = \frac{1}{2}\left(\frac{1}{n(n+1)} - \frac{1}{(n+1)(n+2)}\right).$$

Therefore,

$$\frac{4}{n(n+1)(n+2)} = 2\left(\frac{1}{n(n+1)} - \frac{1}{(n+1)(n+2)}\right).$$

For $N \geq 1$,

$$\begin{aligned}\sum_{n=1}^{N} \frac{4}{n(n+1)(n+2)} &= 2\sum_{n=1}^{N}\left(\frac{1}{n(n+1)} - \frac{1}{(n+1)(n+2)}\right) \\ &= 2\left(\frac{1}{1 \cdot 2} - \frac{1}{(N+1)(N+2)}\right).\end{aligned}$$

Letting $N \to \infty$,

$$\sum_{n=1}^{\infty} \frac{4}{n(n+1)(n+2)} = 2\left(\frac{1}{2}\right) = 1.$$

Hence the given function is a probability mass function.

(b)

$$\begin{aligned}E[X] &= \sum_{n=1}^{\infty} n \frac{4}{n(n+1)(n+2)} \\ &= 4\sum_{n=1}^{\infty} \frac{1}{(n+1)(n+2)}.\end{aligned}$$

Using

$$\frac{1}{(n+1)(n+2)} = \frac{1}{n+1} - \frac{1}{n+2},$$

we obtain

$$\begin{aligned}E[X] &= 4\sum_{n=1}^{\infty}\left(\frac{1}{n+1}-\frac{1}{n+2}\right)\\ &= 4\left(\frac{1}{2}\right)\\ &= 2.\end{aligned}$$

(c)

$$\begin{aligned}E[X^2] &= \sum_{n=1}^{\infty} n^2 \frac{4}{n(n+1)(n+2)}\\ &= 4\sum_{n=1}^{\infty}\frac{n}{(n+1)(n+2)}.\end{aligned}$$

For $n \geq 2$,

$$n+1 \leq 2n, \qquad n+2 \leq 2n,$$

and hence

$$\frac{n}{(n+1)(n+2)} \geq \frac{n}{(2n)(2n)} = \frac{1}{4n}.$$

Therefore,

$$E[X^2] \geq 4\sum_{n=2}^{\infty}\frac{1}{4n} = \sum_{n=2}^{\infty}\frac{1}{n} = \infty.$$

Thus,

$$E[X^2] = \infty.$$

In particular, X has a finite mean but does not have a finite variance.

Problem 26

Let X be a random variable with

$$E[X] = \mu \qquad \text{and} \qquad \mathrm{Var}(X) = \sigma^2,$$

where $\sigma > 0$. Define

$$Y = \frac{X - \mu}{\sigma}.$$

Find $E[Y]$ and $\mathrm{Var}(Y)$.

Solution

Using linearity of expectation,

$$\begin{aligned} E[Y] &= E\left[\frac{X-\mu}{\sigma}\right] \\ &= \frac{1}{\sigma}\left(E[X] - \mu\right) \\ &= \frac{1}{\sigma}(\mu - \mu) \\ &= 0. \end{aligned}$$

For the variance, use

$$\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X).$$

Thus,

$$\begin{aligned} \mathrm{Var}(Y) &= \mathrm{Var}\left(\frac{X-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma^2}\,\mathrm{Var}(X) \\ &= \frac{\sigma^2}{\sigma^2} \\ &= 1. \end{aligned}$$

Therefore, the standardized random variable Y has mean 0 and variance 1.

Problem 27

Show how the binomial probability formula

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, \qquad i = 0, \ldots, n,$$

leads to a proof of the binomial theorem

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$$

when x and y are nonnegative.

Solution

First suppose that $x + y > 0$, and define

$$p = \frac{x}{x+y}.$$

Then

$$1 - p = \frac{y}{x+y}.$$

If $X \sim \mathrm{Bin}(n, p)$, the probabilities of all possible values of X sum to one:

$$\sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = 1.$$

Substituting the expressions for p and $1 - p$,

$$\sum_{i=0}^{n} \binom{n}{i} \left(\frac{x}{x+y}\right)^i \left(\frac{y}{x+y}\right)^{n-i} = 1.$$

Because every term has denominator $(x+y)^n$,

$$\frac{1}{(x+y)^n} \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i} = 1.$$

Multiplying by $(x+y)^n$ gives

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

If $x = y = 0$, the identity also holds directly for $n > 0$, because both sides equal zero.

Problem 28

Let X be a Poisson random variable with parameter $\lambda > 0$. Show that $P(X = i)$ first increases and then decreases as i increases. Determine where the maximum occurs.

Solution

The Poisson probability mass function is

$$p_i = P(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}, \qquad i = 0, 1, 2, \ldots$$

For $i \geq 1$, consider the ratio of consecutive probabilities:

$$\begin{aligned}\frac{p_i}{p_{i-1}} &= \frac{e^{-\lambda}\lambda^i/i!}{e^{-\lambda}\lambda^{i-1}/(i-1)!} \\ &= \frac{\lambda}{i}.\end{aligned}$$

Therefore,

$$p_i > p_{i-1} \quad \Longleftrightarrow \quad i < \lambda,$$

and

$$p_i < p_{i-1} \quad \Longleftrightarrow \quad i > \lambda.$$

Thus the probabilities increase while $i < \lambda$ and decrease once $i > \lambda$.
If λ is not an integer, the unique mode is

$$\lfloor \lambda \rfloor.$$

If $\lambda = m$ is a positive integer, then

$$\frac{p_m}{p_{m-1}} = \frac{m}{m} = 1,$$

so

$$p_{m-1} = p_m.$$

Hence there are two modes:

$$m - 1 \text{ and } m.$$

In particular, $\lfloor \lambda \rfloor$ is always a mode, but when λ is an integer, it is not the only mode.

Problem 29

Let X be a geometric random variable with success probability p, where X is the trial number on which the first success occurs. Show analytically that

$$P(X = n + k \mid X > n) = P(X = k), \qquad n \geq 0, \quad k \geq 1.$$

Also give a verbal explanation.

Solution

Because X is geometric,

$$P(X = j) = (1 - p)^{j-1}p, \qquad j = 1, 2, \ldots$$

Also,

$$P(X > n) = (1 - p)^n.$$

Since the event $\{X = n + k\}$ is contained in the event $\{X > n\}$,

$$\begin{aligned}
P(X = n + k \mid X > n) &= \frac{P(X = n + k)}{P(X > n)} \\
&= \frac{(1 - p)^{n+k-1}p}{(1 - p)^n} \\
&= (1 - p)^{k-1}p \\
&= P(X = k).
\end{aligned}$$

Therefore,

$$P(X = n + k \mid X > n) = P(X = k).$$

For the verbal argument, the condition $X > n$ means that the first n trials were failures. Because the trials are independent, those failures do not affect the outcomes of future trials. Starting after trial n, the waiting time until the first success has the same geometric distribution as it had at the beginning. This property is called the *memoryless property*.

Problem 30

Compute $E[X]$ for each of the following density functions.

(a)

$$f(x) = \begin{cases} \frac{1}{4}xe^{-x/2}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

(b)

$$f(x) = \begin{cases} c(1-x^2), & -1 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(c)

$$f(x) = \begin{cases} \frac{5}{x^2}, & x > 5, \\ 0, & x \leq 5. \end{cases}$$

Solution

(a) By definition,

$$E[X] = \int_0^\infty x f(x)\, dx.$$

Therefore,

$$\begin{aligned} E[X] &= \int_0^\infty x \left(\frac{1}{4}xe^{-x/2}\right) dx \\ &= \frac{1}{4}\int_0^\infty x^2 e^{-x/2}\, dx. \end{aligned}$$

Using integration by parts, we have

$$\int_0^\infty x^2 e^{-x/2}\, dx = \frac{2!}{(1/2)^3} = 16.$$

Equivalently, X has a gamma distribution with shape 2 and scale 2, whose mean is $2 \cdot 2 = 4$.

(b) First determine c by requiring the density to integrate to one:

$$\begin{aligned} 1 &= \int_{-1}^1 c(1-x^2)\, dx \\ &= c\left[x - \frac{x^3}{3}\right]_{-1}^1 \\ &= c\left(\frac{4}{3}\right) \Rightarrow c = \frac{3}{4} \end{aligned}$$

Now,

$$E[X] = \int_{-1}^{1} x \frac{3}{4}(1 - x^2)\, dx = 0 \quad \text{(symmetric interval around 0 of an odd function is zero).}$$

(c) First verify that f is a density:

$$\int_5^\infty \frac{5}{x^2}\, dx = 5 \left[-\frac{1}{x} \right]_5^\infty = 1.$$

The expected value is

$$\begin{aligned} E[X] &= \int_5^\infty x \frac{5}{x^2}\, dx \\ &= 5 \int_5^\infty \frac{1}{x}\, dx. \end{aligned}$$

The integral diverges. Hence, $E[X] = \infty$.

Problem 31

Suppose X has density

$$f(x) = \begin{cases} a + bx^2, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

If

$$E[X] = \frac{3}{5},$$

find a and b.

Solution

Because f is a probability density function,

$$\int_0^1 (a + bx^2)\,dx = 1.$$

Thus,

$$a + \frac{b}{3} = 1.$$

Also,

$$\begin{aligned} E[X] &= \int_0^1 x(a + bx^2)\,dx \\ &= \frac{a}{2} + \frac{b}{4}. \end{aligned}$$

Since $E[X] = \frac{3}{5}$,

$$\frac{a}{2} + \frac{b}{4} = \frac{3}{5}.$$

We therefore solve

$$\begin{cases} a + \dfrac{b}{3} = 1, \\ \dfrac{a}{2} + \dfrac{b}{4} = \dfrac{3}{5}. \end{cases}$$

From the first equation,

$$a = 1 - \frac{b}{3}.$$

Substituting into the second equation gives

$$\begin{aligned} \frac{1}{2}\left(1 - \frac{b}{3}\right) + \frac{b}{4} &= \frac{3}{5}, \\ \frac{1}{2} + \frac{b}{12} &= \frac{3}{5}, \\ \frac{b}{12} &= \frac{1}{10}. \end{aligned}$$

Hence,

$$b = \frac{6}{5}.$$

Then

$$a = 1 - \frac{1}{3}\left(\frac{6}{5}\right) = 1 - \frac{2}{5} = \frac{3}{5}.$$

Therefore,

$$a = \frac{3}{5}, \qquad b = \frac{6}{5}.$$

The resulting density is nonnegative on $[0, 1]$, as required.

Problem 32

The lifetime X, measured in hours, of an electronic tube has density

$$f(x) = xe^{-x}, \qquad x \geq 0.$$

Compute the expected lifetime.

Solution

The expected lifetime is

$$\begin{aligned} E[X] &= \int_0^\infty x f(x)\, dx \\ &= \int_0^\infty x^2 e^{-x}\, dx. \end{aligned}$$

Using

$$\int_0^\infty x^n e^{-x}\, dx = n!,$$

we obtain

$$E[X] = 2! = 2.$$

Thus the expected lifetime is

$$2 \text{ hours}.$$

Problem 33

* A point is chosen at random on a line segment of length L.

(a) Give a mathematical interpretation of the statement that the point is chosen at random.

(b) Find the probability that the ratio of the shorter resulting segment to the longer resulting segment is less than $\frac{1}{4}$.

Solution

Let X be the distance from the left endpoint of the segment to the randomly selected point. Choosing the point uniformly at random means that

$$X \sim \text{Unif}(0, L),$$

with density

$$f_X(x) = \frac{1}{L}, \qquad 0 < x < L.$$

The two resulting segment lengths are

$$X \qquad \text{and} \qquad L - X.$$

By symmetry, first consider $0 < X \le L/2$. In this case, X is the shorter segment and $L - X$ is the longer segment. The desired condition is

$$\frac{X}{L - X} < \frac{1}{4}.$$

Solving,

$$\begin{aligned} 4X &< L - X, \\ 5X &< L, \\ X &< \frac{L}{5}. \end{aligned}$$

By symmetry, the same condition occurs near the right endpoint when

$$X > \frac{4L}{5}.$$

Thus the favorable set is

$$\left(0, \frac{L}{5}\right) \cup \left(\frac{4L}{5}, L\right).$$

Its total length is

$$\frac{L}{5} + \frac{L}{5} = \frac{2L}{5}.$$

Therefore,

$$P\left(\frac{\text{shorter segment}}{\text{longer segment}} < \frac{1}{4}\right) = \frac{2L/5}{L} = \frac{2}{5}.$$

Problem 34

Let

$$X \sim N(10, 36).$$

Compute the following probabilities:

(a) $P(X > 5)$;

(b) $P(4 < X < 16)$;

(c) $P(X < 8)$;

(d) $P(X < 20)$;

Solution

Since

$$X \sim N(10, 36),$$

the mean is

$$\mu = 10$$

and the standard deviation is

$$\sigma = 6.$$

Define the standardized random variable

$$Z = \frac{X - 10}{6},$$

so that $Z \sim N(0, 1)$. Let Φ denote the standard normal CDF.

(a)

$$\begin{aligned} P(X > 5) &= P\left(Z > \frac{5 - 10}{6}\right) \\ &= P\left(Z > -\frac{5}{6}\right) \\ &= \Phi\left(\frac{5}{6}\right) \\ &\approx 0.7977. \end{aligned}$$

(b)

$$\begin{aligned} P(4 < X < 16) &= P\left(\frac{4 - 10}{6} < Z < \frac{16 - 10}{6}\right) \\ &= P(-1 < Z < 1) \\ &= \Phi(1) - \Phi(-1) \\ &\approx 0.6827. \end{aligned}$$

(c)

$$\begin{aligned}
P(X < 8) &= P\left(Z < \frac{8-10}{6}\right) \\
&= P\left(Z < -\frac{1}{3}\right) \\
&= \Phi\left(-\frac{1}{3}\right) \\
&\approx 0.3694.
\end{aligned}$$

(d)

$$\begin{aligned}
P(X < 20) &= P\left(Z < \frac{20-10}{6}\right) \\
&= P\left(Z < \frac{5}{3}\right) \\
&= \Phi\left(\frac{5}{3}\right) \approx 0.9522.
\end{aligned}$$

(e)

$$\begin{aligned}
P(X > 16) &= P\left(Z > \frac{16-10}{6}\right) \\
&= P(Z > 1) \\
&= 1 - \Phi(1) \\
&\approx 0.1587.
\end{aligned}$$

Problem 35

The salaries of physicians in a certain specialty are approximately normally distributed. Suppose that

$$P(X < 180{,}000) = 0.25$$

and

$$P(X > 320{,}000) = 0.25.$$

Approximately what fraction of physicians earn

(a) less than \$200,000?

(b) between \$280,000 and \$320,000?

Solution

The lower and upper quartiles are

$$Q_1 = 180{,}000, \qquad Q_3 = 320{,}000.$$

Because the normal distribution is symmetric, its mean is the midpoint:

$$\mu = \frac{180{,}000 + 320{,}000}{2} = 250{,}000.$$

For a standard normal random variable,

$$\Phi^{-1}(0.75) \approx 0.67449.$$

Thus,

$$\frac{320{,}000 - 250{,}000}{\sigma} = 0.67449.$$

Therefore,

$$\sigma = \frac{70{,}000}{0.67449} \approx 103{,}782.$$

(a) Standardize 200,000:

$$z = \frac{200{,}000 - 250{,}000}{103{,}782} \approx -0.4818.$$

Hence,

$$\begin{aligned} P(X < 200{,}000) &= \Phi(-0.4818) \\ &\approx 0.3150. \end{aligned}$$

Thus approximately 31.5% of physicians earn less than \$200,000.

(b) Standardize 280,000:

$$z = \frac{280{,}000 - 250{,}000}{103{,}782} \approx 0.2891.$$

Also, because 320,000 is the upper quartile,

$$P(X < 320{,}000) = 0.75.$$

Therefore,

$$\begin{aligned} P(280{,}000 < X < 320{,}000) &= P(X < 320{,}000) - P(X < 280{,}000) \\ &= 0.75 - \Phi(0.2891) \\ &\approx 0.75 - 0.6137 \\ &= 0.1363. \end{aligned}$$

Thus approximately 13.6% of physicians earn between \$280,000 and \$320,000.

Problem 36

* Let Y be a continuous random variable with density f_Y, and assume that $E[|Y|] < \infty$. Show that

$$E[Y] = \int_0^\infty P(Y > y)\,dy - \int_0^\infty P(Y < -y)\,dy.$$

Solution

Write $Y = Y^+ - Y^-$, where $Y^+ = \max(Y, 0)$, and $Y^- = \max(-Y, 0)$. Then $E[Y] = E[Y^+] - E[Y^-]$. We first consider the positive part. Since

$$P(Y > y) = \int_y^\infty f_Y(x)\,dx,$$

Tonelli's theorem gives

$$\begin{aligned}
\int_0^\infty P(Y > y)\,dy &= \int_0^\infty \int_y^\infty f_Y(x)\,dx\,dy \\
&= \int_0^\infty \int_0^x dy\, f_Y(x)\,dx \\
&= \int_0^\infty x f_Y(x)\,dx \\
&= E[Y^+].
\end{aligned}$$

Similarly,

$$P(Y < -y) = \int_{-\infty}^{-y} f_Y(x)\,dx.$$

Hence,

$$\int_0^\infty P(Y < -y)\,dy = \int_0^\infty \int_{-\infty}^{-y} f_Y(x)\,dx\,dy.$$

For a fixed $x < 0$, the condition $x < -y$ is equivalent to

$$0 < y < -x.$$

Therefore, changing the order of integration gives

$$\begin{aligned}
\int_0^\infty P(Y < -y)\,dy &= \int_{-\infty}^0 \int_0^{-x} dy\, f_Y(x)\,dx \\
&= \int_{-\infty}^0 (-x) f_Y(x)\,dx \\
&= -\int_{-\infty}^0 x f_Y(x)\,dx \\
&= E[Y^-].
\end{aligned}$$

Combining the two identities,

$$\begin{aligned} E[Y] &= E[Y^+] - E[Y^-] \\ &= \int_0^\infty P(Y > y)\, dy - \int_0^\infty P(Y < -y)\, dy. \end{aligned}$$

Thus,

$$E[Y] = \int_0^\infty P(Y > y)\, dy - \int_0^\infty P(Y < -y)\, dy.$$

Problem 37

Let $Z \sim N(0,1)$. For $x > 0$, show that

(a)
$$P(Z > x) = P(Z < -x);$$

(b)
$$P(|Z| > x) = 2P(Z > x);$$

(c)
$$P(|Z| < x) = 2P(Z < x) - 1.$$

Solution

The standard normal density is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

Because

$$\phi(-z) = \phi(z),$$

the density is symmetric about zero.

(a)
$$P(Z < -x) = \int_{-\infty}^{-x} \phi(z)\, dz.$$

Using the substitution $u = -z$,

$$\begin{aligned} P(Z < -x) &= \int_x^{\infty} \phi(-u)\, du \\ &= \int_x^{\infty} \phi(u)\, du \\ &= P(Z > x). \end{aligned}$$

Therefore,

$$P(Z > x) = P(Z < -x).$$

(b) The events $\{Z > x\}$ and $\{Z < -x\}$ are disjoint, so

$$\begin{aligned} P(|Z| > x) &= P(Z > x) + P(Z < -x) \\ &= 2P(Z > x). \end{aligned}$$

Thus,

$$P(|Z| > x) = 2P(Z > x).$$

(c) By symmetry,

$$P(Z \leq -x) = P(Z \geq x) = 1 - \Phi(x).$$

Therefore,

$$\begin{aligned} P(|Z| < x) &= P(-x < Z < x) \\ &= \Phi(x) - \Phi(-x) \\ &= \Phi(x) - \big(1 - \Phi(x)\big) \\ &= 2\Phi(x) - 1. \end{aligned}$$

Since $\Phi(x) = P(Z < x)$, $P(|Z| < x) = 2P(Z < x) - 1$.

Problem 38

Let

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

be the density of a normal random variable with mean μ and variance σ^2. Show that

$$x = \mu - \sigma \qquad \text{and} \qquad x = \mu + \sigma$$

are inflection points of f.

Solution

Differentiate f. By the chain rule,

$$\begin{aligned} f'(x) &= f(x)\left(-\frac{x-\mu}{\sigma^2}\right) \\ &= -\frac{x-\mu}{\sigma^2} f(x). \end{aligned}$$

Differentiating again,

$$\begin{aligned} f''(x) &= -\frac{1}{\sigma^2} f(x) - \frac{x-\mu}{\sigma^2} f'(x) \\ &= -\frac{1}{\sigma^2} f(x) + \frac{(x-\mu)^2}{\sigma^4} f(x) \\ &= \frac{(x-\mu)^2 - \sigma^2}{\sigma^4} f(x). \end{aligned}$$

Since $f(x) > 0$ for every x,

$$f''(x) = 0$$

if and only if

$$(x-\mu)^2 - \sigma^2 = 0.$$

Thus,

$$(x-\mu)^2 = \sigma^2,$$

which gives

$$x - \mu = \pm\sigma.$$

Therefore,

$$x = \mu - \sigma \quad \text{or} \quad x = \mu + \sigma.$$

Moreover,

$$f''(x) > 0$$

when $|x - \mu| > \sigma$, while

$$f''(x) < 0$$

when $|x - \mu| < \sigma$. Hence the concavity changes at both points, so they are indeed inflection points.

Problem 39

The median of a continuous random variable with CDF F is a value m satisfying

$$F(m) = \frac{1}{2}.$$

Find the median when X is

(a) uniformly distributed on (a, b);

(b) normally distributed with parameters μ and σ^2;

(c) exponentially distributed with rate λ.

Solution

(a) If $X \sim \text{Unif}(a, b)$, then for $a \leq x \leq b$,

$$F(x) = \frac{x-a}{b-a}.$$

Set $F(m) = 1/2$:

$$\begin{aligned} \frac{m-a}{b-a} &= \frac{1}{2}, \\ m - a &= \frac{b-a}{2}, \\ m &= \frac{a+b}{2}. \end{aligned}$$

Thus,

$$m = \frac{a+b}{2}.$$

(b) A normal distribution is symmetric about its mean μ. Hence

$$P(X < \mu) = \frac{1}{2}.$$

Therefore,

$$m = \mu.$$

(c) If $X \sim \text{Exp}(\lambda)$, then for $x \geq 0$,

$$F(x) = 1 - e^{-\lambda x}.$$

Set $F(m) = 1/2$:

$$\begin{aligned} 1 - e^{-\lambda m} &= \frac{1}{2}, \\ e^{-\lambda m} &= \frac{1}{2}, \\ -\lambda m &= -\log 2. \end{aligned}$$

Therefore,

$$m = \frac{\log 2}{\lambda}.$$

Problem 40

Let X be a continuous random variable with continuous cumulative distribution function F. Define

$$Y = F(X).$$

Show that Y is uniformly distributed on $(0, 1)$.

Solution

Let $0 < y < 1$, and let

$$F^{-1}(y) = \inf\{x : F(x) \geq y\}$$

denote the generalized inverse of F.
Because F is continuous,

$$F(F^{-1}(y)) = y.$$

Now,

$$\begin{aligned} P(Y \leq y) &= P(F(X) \leq y) \\ &= P(X \leq F^{-1}(y)) \\ &= F(F^{-1}(y)) \\ &= y. \end{aligned}$$

Therefore, the CDF of Y is

$$F_Y(y) = \begin{cases} 0, & y \leq 0, \\ y, & 0 < y < 1, \\ 1, & y \geq 1. \end{cases}$$

This is the CDF of a uniform random variable on $(0, 1)$. Hence,

$$F(X) \sim \text{Unif}(0, 1).$$

Problem 41

Let

$$\Phi(x) = P(Z \le x),$$

where $Z \sim N(0,1)$. Determine which of the following statements are true:

(a)
$$\Phi(-x) = \Phi(x);$$

(b)
$$\Phi(x) + \Phi(-x) = 1;$$

(c)
$$\Phi(-x) = \frac{1}{\Phi(x)}.$$

Solution

By symmetry of the standard normal distribution,

$$P(Z \le -x) = P(Z \ge x).$$

Since the normal distribution is continuous,

$$P(Z \ge x) = 1 - P(Z \le x) = 1 - \Phi(x).$$

Therefore,

$$\Phi(-x) = 1 - \Phi(x).$$

It follows that:

(a) The statement

$$\Phi(-x) = \Phi(x)$$

is generally false. It holds only when $x = 0$.

(b)
$$\Phi(x) + \Phi(-x) = \Phi(x) + 1 - \Phi(x) = 1.$$

Thus this statement is

$$\text{true.}$$

(c) The statement

$$\Phi(-x) = \frac{1}{\Phi(x)}$$

is false. In fact,

$$\Phi(-x) = 1 - \Phi(x).$$

Hence, only statement (b) is true for every real x.

Problem 42

Let

$$f(x) = \begin{cases} \frac{1}{3}e^{x}, & x < 0, \\ \frac{1}{3}, & 0 \le x < 1, \\ \frac{1}{3}e^{-(x-1)}, & x \ge 1. \end{cases}$$

(a) Show that f is a probability density function.

(b) If X has density f, find $E[X]$.

Solution

(a) Clearly,

$$f(x) \ge 0$$

for every x. Also,

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \frac{1}{3}\int_{-\infty}^{0} e^{x}\,dx + \frac{1}{3}\int_{0}^{1} 1\,dx + \frac{1}{3}\int_{1}^{\infty} e^{-(x-1)}\,dx \\ &= \frac{1}{3}(1) + \frac{1}{3}(1) + \frac{1}{3}(1) \\ &= 1. \end{aligned}$$

Therefore, f is a probability density function.

(b)

$$E[X] = \frac{1}{3}\int_{-\infty}^{0} xe^{x}\,dx + \frac{1}{3}\int_{0}^{1} x\,dx + \frac{1}{3}\int_{1}^{\infty} xe^{-(x-1)}\,dx.$$

For the first integral,

$$\int xe^{x}\,dx = (x-1)e^{x},$$

so

$$\int_{-\infty}^{0} xe^{x}\,dx = -1.$$

For the second integral,

$$\int_{0}^{1} x\,dx = \frac{1}{2}.$$

For the third integral, let $u = x - 1$. Then $x = u + 1$, so

$$\begin{aligned}
\int_1^\infty xe^{-(x-1)}\,dx &= \int_0^\infty (u+1)e^{-u}\,du \\
&= \int_0^\infty ue^{-u}\,du + \int_0^\infty e^{-u}\,du \\
&= 1+1 \\
&= 2.
\end{aligned}$$

Therefore,

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

Problem 43

Let

$$f(x) = \frac{\theta^2}{1+\theta}(1+x)e^{-\theta x}, \qquad x > 0,$$

where $\theta > 0$, and let $f(x) = 0$ for $x \leq 0$.

(a) Show that f is a probability density function.

(b) Find $E[X]$.

(c) Find $\mathrm{Var}(X)$.

Solution

Since $\theta > 0$, $f(x) \geq 0$. We will use the gamma-integral identity

$$\int_0^\infty x^k e^{-\theta x}\, dx = \frac{k!}{\theta^{k+1}}, \qquad k = 0, 1, 2, \ldots$$

(a)

$$\begin{aligned}\int_0^\infty f(x)\, dx &= \frac{\theta^2}{1+\theta}\int_0^\infty (1+x)e^{-\theta x}\, dx \\ &= \frac{\theta^2}{1+\theta}\left(\frac{1}{\theta} + \frac{1}{\theta^2}\right) \\ &= \frac{\theta^2}{1+\theta}\left(\frac{\theta+1}{\theta^2}\right) \\ &= 1.\end{aligned}$$

(b)

$$\begin{aligned}E[X] &= \int_0^\infty x f(x)\, dx \\ &= \frac{\theta^2}{1+\theta}\int_0^\infty x(1+x)e^{-\theta x}\, dx \\ &= \frac{\theta^2}{1+\theta}\left(\int_0^\infty x e^{-\theta x}\, dx + \int_0^\infty x^2 e^{-\theta x}\, dx\right) \\ &= \frac{\theta^2}{1+\theta}\left(\frac{1}{\theta^2} + \frac{2}{\theta^3}\right) = \frac{\theta+2}{\theta(\theta+1)}.\end{aligned}$$

(c) First compute the second moment:

$$
\begin{aligned}
E[X^2] &= \int_0^\infty x^2 f(x)\,dx \\
&= \frac{\theta^2}{1+\theta}\int_0^\infty x^2(1+x)e^{-\theta x}\,dx \\
&= \frac{\theta^2}{1+\theta}\left(\frac{2}{\theta^3}+\frac{6}{\theta^4}\right) \\
&= \frac{2(\theta+3)}{\theta^2(\theta+1)}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathrm{Var}(X) &= E[X^2]-E[X]^2 \\
&= \frac{2(\theta+3)}{\theta^2(\theta+1)}-\left(\frac{\theta+2}{\theta(\theta+1)}\right)^2 \\
&= \frac{2(\theta+3)(\theta+1)-(\theta+2)^2}{\theta^2(\theta+1)^2} \\
&= \frac{2\theta^2+8\theta+6-(\theta^2+4\theta+4)}{\theta^2(\theta+1)^2} = \frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2}.
\end{aligned}
$$

Problem 44

* Let $A_1, A_2, \ldots$ be a sequence of events. Prove that

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n).$$

Solution

Define disjoint events

$$B_1 = A_1$$

and, for $n \geq 2$,

$$B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j.$$

Then the events $B_1, B_2, \ldots$ are pairwise disjoint and

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n.$$

Also,

$$B_n \subseteq A_n,$$

so

$$P(B_n) \leq P(A_n).$$

By countable additivity,

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P\left(\bigcup_{n=1}^{\infty} B_n\right) \\ &= \sum_{n=1}^{\infty} P(B_n) \\ &\leq \sum_{n=1}^{\infty} P(A_n). \end{aligned}$$

Thus,

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n).$$

Problem 45

* Let $A_1, A_2, \ldots$ be a sequence of events. Prove that

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) \geq 1 - \sum_{n=1}^{\infty} P(A_n^c).$$

Solution

By De Morgan's law,

$$\left(\bigcap_{n=1}^{\infty} A_n\right)^c = \bigcup_{n=1}^{\infty} A_n^c.$$

Therefore,

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - P\left(\bigcup_{n=1}^{\infty} A_n^c\right).$$

By Boole's inequality,

$$P\left(\bigcup_{n=1}^{\infty} A_n^c\right) \leq \sum_{n=1}^{\infty} P(A_n^c).$$

Hence,

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) \geq 1 - \sum_{n=1}^{\infty} P(A_n^c).$$

Problem 46

Determine whether each statement is true. If it is true, prove it. If it is false, give a counterexample.

(a) If $P(A) = 1$, then $A = \Omega$.

(b) If $P(B) = 0$, then $B = \varnothing$.

Solution

Both statements are false in general.

(a) Consider

$$\Omega = [0, 1]$$

with the uniform probability distribution, and let

$$A = (0, 1].$$

Then

$$A \neq \Omega,$$

because $0 \notin A$, but

$$P(A) = 1.$$

Thus, an event may have probability 1 without being the entire sample space.

(b) In the same probability space, let

$$B = \left\{\frac{1}{2}\right\}.$$

Then

$$B \neq \varnothing,$$

but a singleton has probability zero under the uniform distribution:

$$P(B) = 0.$$

Thus, a nonempty event may have probability zero.

The statements would be true in certain finite probability spaces in which every individual outcome has positive probability, but they are not true for arbitrary probability spaces.

Problem 47

Let A and B be events satisfying

$$P(A) = P(B) = 1.$$

Show that

$$P(A \cap B) = 1.$$

Solution

By Bonferroni's inequality,

$$P(A \cap B) \geq P(A) + P(B) - 1.$$

Therefore,

$$P(A \cap B) \geq 1 + 1 - 1 = 1.$$

Since no probability can exceed 1,

$$P(A \cap B) = 1.$$

Alternatively,

$$(A \cap B)^c = A^c \cup B^c.$$

Since

$$P(A^c) = P(B^c) = 0,$$

Boole's inequality gives

$$P((A \cap B)^c) \leq P(A^c) + P(B^c) = 0.$$

Hence $P(A \cap B) = 1$.

Problem 48

Is it possible to define a probability measure on a countably infinite sample space so that all outcomes are equally probable?

Solution

Let

$$\Omega = \{\omega_1, \omega_2, \dots\}$$

be countably infinite, and suppose every outcome has the same probability p.
If $p > 0$, then by countable additivity,

$$P(\Omega) = \sum_{n=1}^{\infty} P(\{\omega_n\}) = \sum_{n=1}^{\infty} p = \infty,$$

which contradicts $P(\Omega) = 1$.
If $p = 0$, then

$$P(\Omega) = \sum_{n=1}^{\infty} 0 = 0,$$

which again contradicts $P(\Omega) = 1$.
Therefore,
No countably infinite sample space can have equally probable outcomes under a countably additive probability measure

Problem 49

Let $A_1, \ldots, A_n$ be events satisfying

$$P(A_1) = \cdots = P(A_n) = 1.$$

Show that

$$P(A_1 \cap \cdots \cap A_n) = 1.$$

Solution

By De Morgan's law,

$$(A_1 \cap \cdots \cap A_n)^c = A_1^c \cup \cdots \cup A_n^c.$$

Since $P(A_i) = 1$,

$$P(A_i^c) = 0$$

for every i.
By Boole's inequality,

$$P(A_1^c \cup \cdots \cup A_n^c) \leq \sum_{i=1}^{n} P(A_i^c) = 0.$$

Thus,

$$P((A_1 \cap \cdots \cap A_n)^c) = 0,$$

and consequently,

$$P(A_1 \cap \cdots \cap A_n) = 1.$$

Problem 50

How many $n \times m$ matrices with entries in $\{0, 1\}$ are there?

Solution

An $n \times m$ matrix contains

$$nm$$

entries. Each entry can independently be chosen to be either 0 or 1, giving two choices per entry. By the multiplication rule, the number of matrices is

$$2^{nm}.$$

Problem 51

How many four-digit numbers can be formed using only the digits

$$2, 4, 6, 8, 9?$$

How many of these numbers have at least one repeated digit?

Solution

There are five choices for each of the four digit positions. Since none of the available digits is zero, every resulting string is a four-digit number.
Thus, the total number is

$$5^4 = 625.$$

To count numbers with at least one repeated digit, first count those with all four digits distinct.
There are

$$5 \cdot 4 \cdot 3 \cdot 2 = 120$$

such numbers.
Therefore, the number with at least one repeated digit is

$$625 - 120 = 505.$$

Hence,

$$625 \text{ total numbers}, \qquad 505 \text{ with repetition}.$$

Problem 52

Six fair dice are tossed. Find the probability that at least two of them show the same face.

Solution

Let A be the event that at least two dice show the same face.
Its complement A^c is the event that all six dice show different faces.
There are

$$6^6$$

equally likely ordered outcomes.
For all six faces to be different, every face $1, \ldots, 6$ must appear exactly once. There are

$$6!$$

such outcomes.
Therefore,

$$\begin{aligned} P(A) &= 1 - P(A^c) \\ &= 1 - \frac{6!}{6^6} \\ &= 1 - \frac{720}{46656} \\ &= \frac{899}{900}. \end{aligned}$$

Problem 53

Show that if

$$P(A) = 1,$$

then

$$P(B \mid A) = P(B).$$

Solution

Since $P(A) = 1$,

$$P(A^c) = 0.$$

Now,

$$B = (B \cap A) \cup (B \cap A^c),$$

where the union is disjoint. Therefore,

$$P(B) = P(B \cap A) + P(B \cap A^c).$$

Since

$$B \cap A^c \subseteq A^c,$$

we have

$$P(B \cap A^c) = 0.$$

Thus,

$$P(B \cap A) = P(B).$$

Because $P(A) = 1 > 0$,

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B)}{1} = P(B).$$

Problem 54

* An integer is selected uniformly at random from

$$\{1, 2, \ldots, 10{,}000\}$$

and is observed to be odd. Find the conditional probability that it is

(a) divisible by 3;

(b) divisible by neither 3 nor 5.

Solution

There are exactly

$$5000$$

odd integers between 1 and 10,000.

(a) An odd integer divisible by 3 is an odd multiple of 3.

There are

$$\left\lfloor \frac{10{,}000}{3} \right\rfloor = 3333$$

multiples of 3. Among these, the odd multiples correspond to odd multipliers $1, 3, \ldots, 3333$, of which there are

$$1667.$$

Therefore,

$$P(3 \mid \text{odd}) = \frac{1667}{5000}.$$

(b) Among the 5000 odd integers, count those divisible by 3 or 5.

The number of odd multiples of 3 is 1667.

There are

$$\left\lfloor \frac{10{,}000}{5} \right\rfloor = 2000$$

multiples of 5, of which 1000 are odd.

The numbers divisible by both 3 and 5 are multiples of 15. There are

$$\left\lfloor \frac{10{,}000}{15} \right\rfloor = 666$$

multiples of 15, of which 333 are odd.

By inclusion–exclusion, the number of odd integers divisible by 3 or 5 is

$$1667 + 1000 - 333 = 2334.$$

Thus, the number divisible by neither is

$$5000 - 2334 = 2666.$$

Therefore,

$$P(\text{neither 3 nor 5} \mid \text{odd}) = \frac{2666}{5000} = \frac{1333}{2500}.$$

Problem 55

A judge initially assigns probability 0.65 to the event that Susan is guilty. Robert knows whether Susan is guilty or innocent. If Susan is guilty, Robert lies with probability 0.25. If Susan is innocent, Robert tells the truth. Find the probability that Robert commits perjury.

Solution

Let G be the event that Susan is guilty and L the event that Robert lies.
We are given

$$P(G) = 0.65,$$

$$P(L \mid G) = 0.25,$$

and

$$P(L \mid G^c) = 0.$$

By the law of total probability,

$$\begin{aligned} P(L) &= P(L \mid G)P(G) + P(L \mid G^c)P(G^c) \\ &= (0.25)(0.65) + (0)(0.35) \\ &= 0.1625. \end{aligned}$$

Thus, the probability that Robert commits perjury is

$$16.25\%.$$

Problem 56

* Let X denote the time until a new car breaks down, and define

$$Y = \begin{cases} X, & X \le 5, \\ 5, & X > 5. \end{cases}$$

Let F be the cumulative distribution function of X. Find the cumulative distribution function of Y in terms of F.

Solution

The random variable Y can be written as

$$Y = \min(X, 5).$$

Let

$$F_Y(y) = P(Y \le y).$$

If $y < 5$, then

$$\{Y \le y\} = \{X \le y\},$$

because whenever $X > 5$, we have $Y = 5 > y$. Therefore,

$$F_Y(y) = F(y), \qquad y < 5.$$

If $y \ge 5$, then $Y \le 5 \le y$ for every outcome, so

$$F_Y(y) = 1.$$

Hence,

$$F_Y(y) = \begin{cases} F(y), & y < 5, \\ 1, & y \ge 5. \end{cases}$$

Notice that Y has a point mass at 5 of size

$$P(Y = 5) = 1 - F(5^-).$$

This mass includes both the event $X = 5$ and the event $X > 5$.

Problem 57

Suppose the cumulative distribution function of X is

$$F(x) = \begin{cases} 0, & x < -2, \\ \frac{1}{2}, & -2 \le x < 2, \\ \frac{3}{5}, & 2 \le x < 4, \\ \frac{8}{9}, & 4 \le x < 6, \\ 1, & x \ge 6. \end{cases}$$

Determine the probability mass function of X and describe its graph.

Solution

For a discrete random variable, the probability at a point is the size of the jump of the CDF:

$$P(X = x) = F(x) - F(x^-).$$

At $x = -2$,

$$P(X = -2) = \frac{1}{2} - 0 = \frac{1}{2}.$$

At $x = 2$,

$$P(X = 2) = \frac{3}{5} - \frac{1}{2} = \frac{1}{10}.$$

At $x = 4$,

$$P(X = 4) = \frac{8}{9} - \frac{3}{5} = \frac{40 - 27}{45} = \frac{13}{45}.$$

At $x = 6$,

$$P(X = 6) = 1 - \frac{8}{9} = \frac{1}{9}.$$

Therefore,

$$p_X(x) = \begin{cases} \frac{1}{2}, & x = -2, \\ \frac{1}{10}, & x = 2, \\ \frac{13}{45}, & x = 4, \\ \frac{1}{9}, & x = 6, \\ 0, & \text{otherwise}. \end{cases}$$

The probabilities sum to

$$\frac{1}{2} + \frac{1}{10} + \frac{13}{45} + \frac{1}{9} = 1.$$

Problem 58

For each of the following functions, determine the value of k for which p is a probability mass function.

(a)
$$p(x) = kx, \qquad x = 1, 2, 3, 4, 5.$$

(b)
$$p(x) = k(1 + x)^2, \qquad x = -2, 0, 1, 2.$$

(c)
$$p(x) = k\left(\frac{1}{9}\right)^x, \qquad x = 1, 2, 3, \ldots$$

(d)
$$p(x) = kx, \qquad x = 1, 2, \ldots, n.$$

(e)
$$p(x) = kx^2, \qquad x = 1, 2, \ldots, n.$$

Solution

(a)
$$1 = k\sum_{x=1}^{5} x = k(1 + 2 + 3 + 4 + 5) = 15k \Rightarrow k = \frac{1}{15}.$$

(b) The values of $(1 + x)^2$ are

$$1, \quad 1, \quad 4, \quad 9$$

for $x = -2, 0, 1, 2$, respectively. Hence,

$$1 = k(1 + 1 + 4 + 9) = 15k \Rightarrow k = \frac{1}{15}.$$

(c) Using the geometric-series formula,

$$\begin{aligned} 1 &= k\sum_{x=1}^{\infty}\left(\frac{1}{9}\right)^x \\ &= k\frac{1/9}{1 - 1/9} \\ &= \frac{k}{8} \Rightarrow k = 8 \end{aligned}$$

(d) Since

$$\sum_{x=1}^{n} x = \frac{n(n + 1)}{2},$$

we require

$$1 = k\frac{n(n+1)}{2} \Rightarrow k = \frac{2}{n(n+1)}.$$

(e) Since

$$\sum_{x=1}^{n} x^2 = \frac{n(n+1)(2n+1)}{6},$$

we require

$$1 = k\frac{n(n+1)(2n+1)}{6}.$$

Hence,

$$k = \frac{6}{n(n+1)(2n+1)}.$$

Problem 59

Let X have probability mass function

$$p(x) = \begin{cases} \dfrac{|x-3|+1}{28}, & x = -3, -2, -1, 0, 1, 2, 3, \\ 0, & \text{otherwise.} \end{cases}$$

Find $\text{Var}(X)$.

Solution

The probabilities are

x	-3	-2	-1	0	1	2	3
$28p(x)$	7	6	5	4	3	2	1.

First compute the mean:

$$\begin{aligned} E[X] &= \frac{1}{28}\left[(-3)(7) + (-2)(6) + (-1)(5) + 0(4) + 1(3) + 2(2) + 3(1)\right] \\ &= \frac{-28}{28} \\ &= -1. \end{aligned}$$

Next,

$$\begin{aligned} E[X^2] &= \frac{1}{28}\left[9(7) + 4(6) + 1(5) + 0(4) + 1(3) + 4(2) + 9(1)\right] \\ &= \frac{112}{28} \\ &= 4. \end{aligned}$$

Therefore,

$$\begin{aligned} \text{Var}(X) &= E[X^2] - E[X]^2 \\ &= 4 - (-1)^2 \\ &= 3. \end{aligned}$$

Problem 60

Let X have cumulative distribution function

$$F(x) = \begin{cases} 0, & x < -3, \\ \frac{3}{8}, & -3 \le x < 0, \\ \frac{3}{4}, & 0 \le x < 6, \\ 1, & x \ge 6. \end{cases}$$

Find the variance and standard deviation of X.

Solution

The probability masses are the jumps of F:

$$P(X = -3) = \frac{3}{8},$$

$$P(X = 0) = \frac{3}{4} - \frac{3}{8} = \frac{3}{8},$$

and

$$P(X = 6) = 1 - \frac{3}{4} = \frac{1}{4}.$$

Thus,

$$\begin{aligned} E[X] &= (-3)\frac{3}{8} + 0\frac{3}{8} + 6\frac{1}{4} \\ &= -\frac{9}{8} + \frac{12}{8} \\ &= \frac{3}{8}. \end{aligned}$$

Also,

$$\begin{aligned} E[X^2] &= 9\frac{3}{8} + 0 + 36\frac{1}{4} \\ &= \frac{27}{8} + 9 \\ &= \frac{99}{8}. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{Var}(X) &= E[X^2] - E[X]^2 \\ &= \frac{99}{8} - \left(\frac{3}{8}\right)^2 \\ &= \frac{792}{64} - \frac{9}{64} \\ &= \frac{783}{64}. \end{aligned}$$

The standard deviation is

$$\sigma_X = \sqrt{\frac{783}{64}} = \frac{\sqrt{783}}{8} \approx 3.4978.$$

Problem 61

Suppose X is a discrete random variable satisfying

$$E[X] = 1$$

and

$$E[X(X-2)] = 3.$$

Find

$$\mathrm{Var}(-3X+5).$$

Solution

Since

$$X(X-2) = X^2 - 2X,$$

we have

$$E[X^2] - 2E[X] = 3.$$

Using $E[X] = 1$,

$$E[X^2] - 2 = 3,$$

so

$$E[X^2] = 5.$$

Therefore,

$$\mathrm{Var}(X) = E[X^2] - E[X]^2 = 5 - 1 = 4.$$

Now use

$$\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X).$$

Thus,

$$\mathrm{Var}(-3X+5) = (-3)^2(4) = 36.$$

Problem 62

Let X be the amount, in fluid ounces, in a randomly selected bottle from company A, and let Y be the amount in a randomly selected bottle from company B. Their distributions are

x	15.85	15.9	16	16.1	16.2
$P(X=x)$	0.15	0.21	0.35	0.15	0.14
$P(Y=x)$	0.14	0.05	0.64	0.08	0.09

Find $E[X]$, $E[Y]$, $\mathrm{Var}(X)$, and $\mathrm{Var}(Y)$, and interpret the results.

Solution

For X,

$$\begin{aligned} E[X] &= 15.85(0.15) + 15.9(0.21) + 16(0.35) \\ &\quad + 16.1(0.15) + 16.2(0.14) \\ &= 15.9995. \end{aligned}$$

Similarly,

$$\begin{aligned} E[Y] &= 15.85(0.14) + 15.9(0.05) + 16(0.64) \\ &\quad + 16.1(0.08) + 16.2(0.09) \\ &= 16.0000. \end{aligned}$$

The second moments are

$$E[X^2] = \sum_x x^2 P(X=x) = 256. - \text{computed through the distribution,}$$

and

$$E[Y^2] = \sum_x x^2 P(Y=x).$$

Using

$$\mathrm{Var}(X) = E[X^2] - E[X]^2,$$

the resulting variances are

$$\mathrm{Var}(X) = 0.01257475$$

and

$$\mathrm{Var}(Y) = 0.00805.$$

Thus, the standard deviations are approximately

$$\sigma_X \approx 0.1121$$

and

$$\sigma_Y \approx 0.0897.$$

Both companies fill their bottles with an average of approximately 16 fluid ounces. Company B, however, has the smaller variance and standard deviation, so its bottle amounts are more tightly concentrated around 16 ounces. Company B is therefore more consistent.

Problem 63

A couple wants to have at least a 95% probability of having at least one boy and at least one girl. Assume that the sexes of the children are independent and that a child is equally likely to be a boy or a girl.

What is the minimum number of children they should plan to have?

Solution

For n children, the complement of having at least one boy and at least one girl is the event that all children have the same sex.
Thus,

$$\begin{aligned} P(\text{at least one boy and one girl}) &= 1 - P(\text{all boys or all girls}) \\ &= 1 - \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^n \\ &= 1 - 2^{1-n}. \end{aligned}$$

We require

$$1 - 2^{1-n} \geq 0.95.$$

Therefore,

$$2^{1-n} \leq 0.05.$$

Testing integers,

$$n = 5: \qquad 1 - 2^{1-5} = 1 - \frac{1}{16} = 0.9375 < 0.95,$$

whereas

$$n = 6: \qquad 1 - 2^{1-6} = 1 - \frac{1}{32} = 0.96875 > 0.95.$$

Hence, the minimum number is

$$6.$$

Problem 64

A rare blood type occurs in 0.05% of the population. A group of 3000 people is selected independently from the population. Find the probability that at least two people in the group have the rare blood type.

Solution

Let X be the number of people with the rare blood type. Then

$$X \sim \text{Bin}(3000, 0.0005).$$

The desired probability is

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - (1 - 0.0005)^{3000} \\ &\qquad - 3000(0.0005)(1 - 0.0005)^{2999}. \end{aligned}$$

Thus,

$$P(X \geq 2) \approx 0.4422.$$

Because n is large and p is small, a Poisson approximation is also appropriate. Here,

$$\lambda = np = 3000(0.0005) = 1.5.$$

Thus,

$$\begin{aligned} P(X \geq 2) &\approx 1 - P(Y = 0) - P(Y = 1) \\ &= 1 - e^{-1.5}(1 + 1.5) \\ &\approx 0.4422, \end{aligned}$$

where $Y \sim \text{Pois}(1.5)$.

Problem 65

Suppose X is a Poisson random variable satisfying

$$P(X = 1) = P(X = 3).$$

Find $P(X = 5)$.

Solution

Let

$$X \sim \text{Pois}(\lambda).$$

Then

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

The condition gives

$$e^{-\lambda}\lambda = e^{-\lambda}\frac{\lambda^3}{3!}.$$

For $\lambda > 0$, canceling $e^{-\lambda}\lambda$ yields

$$1 = \frac{\lambda^2}{6}.$$

Therefore,

$$\lambda^2 = 6$$

and

$$\lambda = \sqrt{6}.$$

Consequently,

$$\begin{aligned} P(X = 5) &= e^{-\sqrt{6}}\frac{(\sqrt{6})^5}{5!} \\ &= e^{-\sqrt{6}}\frac{36\sqrt{6}}{120} \\ &= \frac{3\sqrt{6}}{10}e^{-\sqrt{6}}. \end{aligned}$$

Problem 66

The duration X of a certain soap opera, measured in tens of hours, has cumulative distribution function

$$F(x) = \begin{cases} 0, & x < 4, \\ 1 - \dfrac{16}{x^2}, & x \geq 4. \end{cases}$$

(a) Find the probability density function of X.

(b) Describe the graphs of F and f.

(c) Find the probability that the soap opera lasts:

(i) at most 50 hours;

(ii) at least 60 hours;

(iii) between 50 and 70 hours;

(iv) between 10 and 35 hours.

Solution

(a) For $x > 4$,

$$f(x) = F'(x) = \frac{32}{x^3}.$$

Thus,

$$f(x) = \begin{cases} \dfrac{32}{x^3}, & x \geq 4, \\ 0, & x < 4. \end{cases}$$

There is no point mass at 4, because

$$F(4) = 1 - \frac{16}{16} = 0.$$

(b) The CDF is zero for $x < 4$, begins at zero when $x = 4$, and increases continuously toward 1.

The density is zero for $x < 4$, has value

$$f(4) = \frac{32}{64} = \frac{1}{2},$$

and decreases toward zero as $x \to \infty$.

(c) Since X is measured in tens of hours, 50 hours corresponds to $x = 5$, 60 hours to $x = 6$, and so forth.

(i)

$$P(X \leq 5) = F(5) = 1 - \frac{16}{25} = \frac{9}{25}.$$

(ii) Since the distribution is continuous,

$$\begin{aligned} P(X \geq 6) &= 1 - F(6) \\ &= \frac{16}{36} \\ &= \frac{4}{9}. \end{aligned}$$

(iii)

$$\begin{aligned} P(5 < X < 7) &= F(7) - F(5) \\ &= \left(1 - \frac{16}{49}\right) - \left(1 - \frac{16}{25}\right) \\ &= \frac{16}{25} - \frac{16}{49} \\ &= \frac{384}{1225}. \end{aligned}$$

(iv) The support of X is $[4, \infty)$, corresponding to durations of at least 40 hours. Hence,

$$P(1 < X < 3.5) = 0.$$

Problem 67

The lifetime of a randomly selected used tire is $10{,}000X$ miles, where X has density

$$f(x) = \begin{cases} \dfrac{2}{x^2}, & 1 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

(a) What percentage of the tires last fewer than 15,000 miles?

(b) Among the tires lasting fewer than 15,000 miles, what percentage last between 10,000 and 12,500 miles?

Solution

(a) A lifetime below 15,000 miles corresponds to

$$X < 1.5.$$

Therefore,

$$\begin{aligned} P(X < 1.5) &= \int_1^{3/2} \frac{2}{x^2}\,dx \\ &= \left[-\frac{2}{x}\right]_1^{3/2} \\ &= 2 - \frac{4}{3} \\ &= \frac{2}{3}. \end{aligned}$$

Thus, $66\frac{2}{3}\%$ of the tires last fewer than 15,000 miles.

(b) The interval 10,000 to 12,500 miles corresponds to

$$1 < X < 1.25 = \frac{5}{4}.$$

The required conditional probability is

$$P\left(1 < X < \frac{5}{4} \,\middle|\, X < \frac{3}{2}\right).$$

Since $\{1 < X < 5/4\} \subseteq \{X < 3/2\}$,

$$P\left(1 < X < \frac{5}{4} \,\middle|\, X < \frac{3}{2}\right) = \frac{P(1 < X < 5/4)}{P(X < 3/2)}.$$

Now,

$$
\begin{aligned}
P\left(1 < X < \frac{5}{4}\right) &= \int_1^{5/4} \frac{2}{x^2}\,dx \\
&= 2 - \frac{8}{5} \\
&= \frac{2}{5}.
\end{aligned}
$$

Therefore, $\frac{2/5}{2/3} = \frac{3}{5}$. Thus, 60% of the tires lasting fewer than 15,000 miles last between 10,000 and 12,500 miles.

Problem 68

Let X have CDF

$$F(x) = \begin{cases} 0, & x < 4, \\ 1 - \dfrac{16}{x^2}, & x \geq 4. \end{cases}$$

(a) Find $E[X]$.

(b) Show that $\mathrm{Var}(X)$ does not exist as a finite number.

Solution

The density is

$$f(x) = \frac{32}{x^3}, \qquad x \geq 4.$$

(a)

$$\begin{aligned} E[X] &= \int_4^\infty x \frac{32}{x^3}\, dx \\ &= 32 \int_4^\infty \frac{1}{x^2}\, dx \\ &= 32 \left[-\frac{1}{x} \right]_4^\infty \\ &= 32 \left(\frac{1}{4} \right) \\ &= 8. \end{aligned}$$

Because X is measured in tens of hours, the expected duration is

80 hours.

(b)

$$\begin{aligned} E[X^2] &= \int_4^\infty x^2 \frac{32}{x^3}\, dx \\ &= 32 \int_4^\infty \frac{1}{x}\, dx. \end{aligned}$$

The integral diverges, so

$$E[X^2] = \infty.$$

Therefore,

$\mathrm{Var}(X)$ is not finite.

Although $E[X]$ exists, the second moment and variance do not.

Problem 69

The time X, in hours, required for a student to finish an aptitude test has density

$$f(x) = \begin{cases} 6(x-1)(2-x), & 1 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

Find the mean and standard deviation of X.

Solution

The density is symmetric about $x = 3/2$, so we expect the mean to be $3/2$. Directly,

$$\begin{aligned} E[X] &= \int_1^2 x\, 6(x-1)(2-x)\, dx \\ &= \frac{3}{2}. \end{aligned}$$

Next,

$$\begin{aligned} E[X^2] &= \int_1^2 x^2\, 6(x-1)(2-x)\, dx \\ &= \frac{23}{10}. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{Var}(X) &= E[X^2] - E[X]^2 \\ &= \frac{23}{10} - \left(\frac{3}{2}\right)^2 \\ &= \frac{23}{10} - \frac{9}{4} \\ &= \frac{46-45}{20} \\ &= \frac{1}{20}. \end{aligned}$$

Thus, the standard deviation is

$$\sigma_X = \sqrt{\frac{1}{20}} = \frac{1}{2\sqrt{5}} \approx 0.2236 \text{ hours.}$$

The mean is

$$E[X] = 1.5 \text{ hours.}$$

Problem 70

Let X be a continuous random variable with density

$$f(x) = \begin{cases} \dfrac{2}{x^2}, & 1 < x < 2, \\ 0, & \text{otherwise.} \end{cases}$$

Find $E[\ln X]$.

Solution

By definition,

$$E[\ln X] = \int_1^2 \ln(x) \frac{2}{x^2}\, dx.$$

Integrate by parts. Let

$$u = \ln x, \qquad dv = \frac{2}{x^2}\, dx.$$

Then

$$du = \frac{1}{x}\, dx, \qquad v = -\frac{2}{x}.$$

Therefore,

$$\begin{aligned} E[\ln X] &= \left[-\frac{2\ln x}{x}\right]_1^2 + 2\int_1^2 \frac{1}{x^2}\, dx \\ &= -\ln 2 + 2\left[-\frac{1}{x}\right]_1^2 \\ &= -\ln 2 + 2\left(1 - \frac{1}{2}\right) \\ &= 1 - \ln 2. \end{aligned}$$

Numerically,

$$E[\ln X] \approx 0.3069.$$

Problem 71

Let X have density

$$f(x) = \frac{1}{2}e^{-|x|}, \qquad -\infty < x < \infty.$$

Calculate $\mathrm{Var}(X)$.

Solution

The density is symmetric about zero, so

$$E[X] = 0.$$

Therefore,

$$\mathrm{Var}(X) = E[X^2].$$

Since $x^2e^{-|x|}$ is an even function,

$$\begin{aligned} E[X^2] &= \frac{1}{2}\int_{-\infty}^{\infty} x^2e^{-|x|}\,dx \\ &= \int_0^{\infty} x^2e^{-x}\,dx. \end{aligned}$$

Using

$$\int_0^{\infty} x^ne^{-x}\,dx = n!,$$

we obtain

$$E[X^2] = 2! = 2.$$

Hence,

$$\mathrm{Var}(X) = 2.$$

Problem 72

Let X have a gamma distribution with shape parameter r and rate parameter λ, so that

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, \qquad x > 0.$$

Assume $r > 1$. Show that the density has a unique maximum at $\frac{r-1}{\lambda}$.

Solution

Because the multiplicative constant

$$\frac{\lambda^r}{\Gamma(r)}$$

is positive, it is enough to maximize

$$g(x) = x^{r-1} e^{-\lambda x}.$$

Take logarithms:

$$\log g(x) = (r-1)\log x - \lambda x.$$

Differentiate:

$$\frac{d}{dx} \log g(x) = \frac{r-1}{x} - \lambda.$$

The critical point satisfies

$$\frac{r-1}{x} - \lambda = 0,$$

so

$$x = \frac{r-1}{\lambda}.$$

Moreover,

$$\frac{d^2}{dx^2} \log g(x) = -\frac{r-1}{x^2} < 0$$

for $x > 0$. Thus $\log g$, and hence g, is strictly concave at the critical point.
Also,

$$\frac{r-1}{x} - \lambda > 0$$

when $x < (r-1)/\lambda$, and it is negative when $x > (r-1)/\lambda$. Therefore, the density increases before this point and decreases afterward.
Hence the unique mode is

$$\frac{r-1}{\lambda}.$$

Problem 73

* Let X have a gamma distribution with shape parameter r and rate parameter λ. Let $c > 0$, and define

$$Y = cX.$$

Find the distribution function and density of Y.

Solution

For $y \le 0$,

$$F_Y(y) = 0.$$

For $y > 0$,

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(cX \le y) \\ &= P\left(X \le \frac{y}{c}\right) \\ &= F_X\left(\frac{y}{c}\right). \end{aligned}$$

Thus,

$$F_Y(y) = \begin{cases} 0, & y \le 0, \\ F_X\left(\frac{y}{c}\right), & y > 0. \end{cases}$$

Using the change-of-variables formula,

$$\begin{aligned} f_Y(y) &= \frac{1}{c} f_X\left(\frac{y}{c}\right) \\ &= \frac{1}{c}\frac{\lambda^r}{\Gamma(r)}\left(\frac{y}{c}\right)^{r-1} e^{-\lambda y/c} \\ &= \frac{(\lambda/c)^r}{\Gamma(r)} y^{r-1} e^{-(\lambda/c)y}, \qquad y > 0. \end{aligned}$$

Therefore,

$$Y \sim \text{Gamma}\left(r, \frac{\lambda}{c}\right)$$

under the shape–rate parameterization.

Problem 74

Let

$$f(x) = \begin{cases} \dfrac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x > 0, \\ 0, & x \le 0, \end{cases}$$

where $r > 0$ and $\lambda > 0$. Prove that

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Solution

Because $f(x) = 0$ for $x \le 0$,

$$\int_{-\infty}^{\infty} f(x)\,dx = \frac{\lambda^r}{\Gamma(r)} \int_0^{\infty} x^{r-1} e^{-\lambda x}\,dx.$$

Use the substitution

$$u = \lambda x.$$

Then

$$x = \frac{u}{\lambda}, \qquad dx = \frac{du}{\lambda}.$$

Therefore,

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \frac{\lambda^r}{\Gamma(r)} \int_0^{\infty} \left(\frac{u}{\lambda}\right)^{r-1} e^{-u} \frac{du}{\lambda} \\ &= \frac{\lambda^r}{\Gamma(r)} \frac{1}{\lambda^r} \int_0^{\infty} u^{r-1} e^{-u}\,du \\ &= \frac{1}{\Gamma(r)} \Gamma(r) \\ &= 1. \end{aligned}$$

Thus,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Problem 75

* Consider

$$f(x) = \begin{cases} 12x(1-x)^2, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Determine whether f is the density of a beta random variable. If so, find $E[X]$ and $\mathrm{Var}(X)$.

Solution

A beta density with parameters $\alpha, \beta > 0$ has the form

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1.$$

Here,

$$x(1-x)^2 = x^{2-1}(1-x)^{3-1},$$

so the possible parameters are

$$\alpha = 2, \qquad \beta = 3.$$

The beta normalizing constant is

$$\frac{1}{B(2,3)} = \frac{\Gamma(5)}{\Gamma(2)\Gamma(3)} = \frac{4!}{1!\,2!} = 12.$$

Thus f is exactly the density of a

$$\mathrm{Beta}(2,3)$$

random variable.
For a beta random variable,

$$E[X] = \frac{\alpha}{\alpha+\beta},$$

so

$$E[X] = \frac{2}{5}.$$

Also,

$$\mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Therefore,

$$\begin{aligned} \mathrm{Var}(X) &= \frac{(2)(3)}{5^2(6)} \\ &= \frac{1}{25}. \end{aligned}$$

Problem 76

* Determine whether

$$f(x) = \begin{cases} 120x^2(1-x)^4, & 0 < x < 1, \\ 0, & \text{otherwise} \end{cases}$$

is a probability density function.

Solution

The function is nonnegative. It remains to check whether it integrates to one.
The kernel

$$x^2(1-x)^4 = x^{3-1}(1-x)^{5-1}$$

corresponds to a beta distribution with parameters

$$\alpha = 3, \qquad \beta = 5.$$

Its required normalizing constant is

$$\begin{aligned} \frac{1}{B(3,5)} &= \frac{\Gamma(8)}{\Gamma(3)\Gamma(5)} \\ &= \frac{7!}{2!\,4!} \\ &= 105. \end{aligned}$$

The proposed coefficient is 120, not 105.
Indeed,

$$\begin{aligned} \int_0^1 120x^2(1-x)^4\,dx &= 120B(3,5) \\ &= \frac{120}{105} \\ &= \frac{8}{7}. \end{aligned}$$

Since

$$\frac{8}{7} \neq 1,$$

the proposed function is not a probability density function.
Thus,

$$f \text{ is not a probability density function.}$$

Replacing 120 by 105 would produce the density of a $\text{Beta}(3,5)$ random variable.

Problem 77

Suppose that

$$X \sim \text{Bin}(100, 0.4).$$

Use a Normal approximation, with an appropriate continuity correction, to approximate

$$\mathbb{P}(35 \leq X \leq 45).$$

Clearly verify that the conditions for the Normal approximation are satisfied and state the approximating Normal distribution.

Solution

We have

$$X \sim \text{Bin}(100, 0.4).$$

To check whether the Normal approximation is appropriate, we compute

$$np = 100(0.4) = 40$$

and

$$n(1-p) = 100(0.6) = 60.$$

Since both quantities are greater than 10, the Normal approximation is appropriate.
The mean and variance are

$$\mu = np = 40$$

and

$$\sigma^2 = np(1-p) = 100(0.4)(0.6) = 24.$$

Therefore, we approximate X by

$$Y \sim N(40, 24).$$

Using the continuity correction,

$$\mathbb{P}(35 \leq X \leq 45) \approx \mathbb{P}(34.5 < Y < 45.5).$$

Standardizing,

$$\begin{aligned}\mathbb{P}(34.5 < Y < 45.5) &= \mathbb{P}\left(\frac{34.5 - 40}{\sqrt{24}} < Z < \frac{45.5 - 40}{\sqrt{24}}\right) \\ &= \mathbb{P}(-1.123 < Z < 1.123).\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{P}(35 \leq X \leq 45) &\approx \Phi(1.123) - \Phi(-1.123) \\ &= 2\Phi(1.123) - 1 \\ &\approx 0.739.\end{aligned}$$

Problem 78

The odds are 1 to 5000 in favor of a customer buying a particular fiction bestseller.

Suppose that 800 customers enter a bookstore each day and that a month has 30 days.

How many copies of the book should the bookstore stock each month so that, with probability greater than 98%, it does not run out?

Determine whether a Poisson approximation or a Normal approximation is more appropriate, and verify the required conditions before solving the problem.

Solution

Let
$$X = \text{the number of customers who buy the book during the month.}$$

The total number of customers during the month is

$$n = 30(800) = 24000.$$

Because the odds in favor of buying the book are 1 to 5000, the probability that a customer buys the book is
$$p = \frac{1}{1+5000} = \frac{1}{5001}.$$

Therefore,
$$X \sim \text{Bin}\left(24000, \frac{1}{5001}\right).$$

We first determine the appropriate approximation.
The expected number of successes is

$$np = 24000\left(\frac{1}{5001}\right) \approx 4.799.$$

Although n is large,
$$np \approx 4.799 < 10,$$

so the usual Normal approximation is not appropriate.
However, $n = 24000$ is large and
$$p = \frac{1}{5001}$$
is very small. Therefore, the Poisson approximation is appropriate, with

$$\lambda = np = \frac{24000}{5001} \approx 4.799.$$

Thus,
$$X \approx W, \qquad W \sim \text{Pois}(4.799).$$

If the bookstore stocks m copies, then it does not run out when

$$X \leq m.$$

We seek the smallest integer m such that

$$\mathbb{P}(X \leq m) > 0.98.$$

Using the Poisson approximation,

$$\mathbb{P}(X \leq 9) \approx \mathbb{P}(W \leq 9) = \sum_{k=0}^{9} e^{-4.799} \frac{4.799^k}{k!} \approx 0.9749.$$

Since

$$0.9749 < 0.98,$$

stocking 9 copies is not sufficient.
Next,

$$\mathbb{P}(X \leq 10) \approx \mathbb{P}(W \leq 10) = \sum_{k=0}^{10} e^{-4.799} \frac{4.799^k}{k!} \approx 0.9896.$$

Since

$$0.9896 > 0.98,$$

the smallest satisfactory value is $m = 10$.
Therefore, the bookstore should stock

$$10 \text{ copies.}$$

Problem 79

Every day, a factory produces 5000 light bulbs, of which 2500 are Type I and 2500 are Type II.

A sample of 40 light bulbs is selected uniformly at random without replacement.

Approximate the probability that the sample contains at least 18 light bulbs of each type.

Clearly explain and justify every approximation that you use.

Solution

Let
$$X = \text{the number of Type I bulbs in the sample.}$$
Because the sample is selected without replacement, the exact distribution is
$$X \sim \text{Hypergeo}(5000, 2500, 40).$$
The sampling fraction is
$$\frac{40}{5000} = 0.008.$$
Since the sample size is very small relative to the population size, sampling without replacement is approximately equivalent to sampling with replacement. Therefore,
$$X \approx \text{Bin}\left(40, \frac{2500}{5000}\right) = \text{Bin}\left(40, \frac{1}{2}\right).$$
For this Binomial distribution,
$$np = 40\left(\frac{1}{2}\right) = 20$$
and
$$n(1-p) = 40\left(\frac{1}{2}\right) = 20.$$
Since both quantities are greater than 10, the Normal approximation is appropriate.
The mean and variance are
$$\mu = np = 20$$
and
$$\sigma^2 = np(1-p) = 40\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 10.$$
Thus, we further approximate X by
$$Y \sim N(20, 10).$$
If there are X Type I bulbs, then there are
$$40 - X$$
Type II bulbs.
The sample contains at least 18 bulbs of each type when
$$X \geq 18$$
and
$$40 - X \geq 18.$$

The second inequality is equivalent to
$$X \leq 22.$$
Therefore, the desired event is
$$18 \leq X \leq 22.$$
Using the continuity correction,
$$\mathbb{P}(18 \leq X \leq 22) \approx \mathbb{P}(17.5 < Y < 22.5).$$
Standardizing,
$$\begin{aligned}\mathbb{P}(17.5 < Y < 22.5) &= \mathbb{P}\left(\frac{17.5-20}{\sqrt{10}} < Z < \frac{22.5-20}{\sqrt{10}}\right)\\ &= \mathbb{P}(-0.791 < Z < 0.791).\end{aligned}$$
Therefore,
$$\begin{aligned}\mathbb{P}(18 \leq X \leq 22) &\approx \Phi(0.791) - \Phi(-0.791)\\ &= 2\Phi(0.791) - 1\\ &\approx 2(0.785) - 1\\ &= 0.570.\end{aligned}$$
Thus, the approximate probability that the sample contains at least 18 bulbs of each type is 0.570.

Problem 80

Suppose that the lifetimes of light bulbs produced by a company are Normally distributed with mean 1000 hours and standard deviation 100 hours.

The company claims that 95% of its light bulbs last at least 900 hours.

Is the company's claim correct? Justify your answer mathematically.

Solution

Let

$$X = \text{the lifetime of a randomly selected light bulb.}$$

Then

$$X \sim N(1000, 100^2).$$

The proportion of bulbs that last at least 900 hours is

$$\mathbb{P}(X \geq 900).$$

Standardizing,

$$\begin{aligned}\mathbb{P}(X \geq 900) &= \mathbb{P}\left(Z \geq \frac{900 - 1000}{100}\right)\\ &= \mathbb{P}(Z \geq -1).\end{aligned}$$

By symmetry of the standard Normal distribution,

$$\mathbb{P}(Z \geq -1) = \mathbb{P}(Z \leq 1) = \Phi(1).$$

Using the standard Normal table,

$$\Phi(1) \approx 0.8413.$$

Therefore,

$$\mathbb{P}(X \geq 900) \approx 0.8413.$$

Thus, approximately 84.13%, rather than 95%, of the bulbs last at least 900 hours. Since

$$0.8413 < 0.95,$$

the company's claim is incorrect.

No, the company's claim is not correct.

Problem 81

The lifetime of a bulb produced by the first company is Normally distributed with mean 1000 hours and standard deviation 100 hours.

The lifetime of a bulb produced by the second company is Normally distributed with mean 900 hours and standard deviation 150 hours.

Howard buys one bulb from each company. Assuming that the two bulb lifetimes are independent, what is the probability that at least one of the bulbs lasts 980 hours or more?

Solution

Let

$$X = \text{the lifetime of the bulb from the first company}$$

and

$$Y = \text{the lifetime of the bulb from the second company.}$$

Then

$$X \sim N(1000, 100^2)$$

and

$$Y \sim N(900, 150^2).$$

We want

$$\mathbb{P}(X \geq 980 \text{ or } Y \geq 980).$$

It is easier to use the complementary event:

$$\mathbb{P}(X \geq 980 \text{ or } Y \geq 980) = 1 - \mathbb{P}(X < 980,\ Y < 980).$$

Because X and Y are independent,

$$\mathbb{P}(X < 980,\ Y < 980) = \mathbb{P}(X < 980)\mathbb{P}(Y < 980).$$

For the first bulb,

$$\begin{aligned}\mathbb{P}(X < 980) &= \mathbb{P}\left(Z < \frac{980 - 1000}{100}\right) \\ &= \mathbb{P}(Z < -0.20) \\ &= \Phi(-0.20) \\ &\approx 0.4207.\end{aligned}$$

For the second bulb,

$$\begin{aligned}\mathbb{P}(Y < 980) &= \mathbb{P}\left(Z < \frac{980 - 900}{150}\right) \\ &= \mathbb{P}(Z < 0.533) \\ &= \Phi(0.533) \\ &\approx 0.7031.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{P}(X < 980,\ Y < 980) &\approx (0.4207)(0.7031) \\ &\approx 0.2958.\end{aligned}$$

Consequently,

$$\begin{aligned}\mathbb{P}(X \geq 980 \text{ or } Y \geq 980) &\approx 1 - 0.2958 \\ &= 0.7042.\end{aligned}$$

Thus, the probability that at least one of the two bulbs lasts 980 hours or more is approximately 70.42%.

References

1. Sheldon Ross, *A First Course in Probability*, 10th ed., Pearson, 2019.
2. Saeed Ghahramani, *Fundamentals of Probability with Stochastic Processes*, 4th ed., Pearson, 2018.