Worksheet - Probability and random variables

Year: University Year 1 · 20

Show your working.

Question 1

Question 1 – Events and the axioms *(Section A · Skills practice)*

Events A and B satisfy $P(A) = 0.6$, $P(B) = 0.5$ and $P(A \cap B) = 0.3$.

a. Find $P(A \cup B)$ and $P(A^c \cap B^c)$. [3 marks]
b. Find $P(A^c \cap B)$. [2 marks]
c. Find $P(A \mid B)$ and decide whether A and B are independent. [3 marks]

Question 2

Question 2 – Two urns and Bayes' theorem *(Section A · Skills practice)*

Urn 1 contains 3 red and 2 blue balls; urn 2 contains 1 red and 4 blue balls. A fair coin decides which urn is used, then one ball is drawn.

a. Draw a tree diagram and find $P(\text{red})$. [3 marks]
b. Given that the ball is red, find the probability that urn 1 was used. [3 marks]
c. Given that the ball is blue, find the probability that urn 1 was used. [2 marks]

Question 3

Question 3 – Expectation and variance *(Section A · Skills practice)*

A discrete random variable X has the distribution below.

x	0	1	2	3
$P(X = x)$	0.1	0.3	0.4	0.2

a. Check that this is a valid distribution and find $E[X]$. [2 marks]
b. Find $E[X^2]$ and $\text{Var}(X)$. [3 marks]
c. Find $E[2X - 1]$ and $\text{Var}(2X - 1)$. [3 marks]

Question 4

Question 4 – The binomial distribution *(Section A · Skills practice)*

Let $X \sim B(12,\, 0.25)$.

a. Find $P(X = 3)$. [2 marks]

b. Find $P(X \leq 1)$. [3 marks]
c. Find the mean and the standard deviation of X. [2 marks]

Question 5

Question 5 – The Poisson distribution *(Section A · Skills practice)*

Let $X \sim \text{Po}(2.5)$.

a. Find $P(X = 2)$. [2 marks]
b. Find $P(X \leq 2)$ and $P(X \geq 3)$. [3 marks]
c. State $E[X]$ and $\text{Var}(X)$. [1 mark]

Question 6

Question 6 – Normal and continuous distributions *(Section A · Skills practice)*

Let $X \sim N(50,\ 8^2)$, and let Y have density $f(y) = 3y^2$ on $[0, 1]$.

a. Find $P(X$ and $P(40$. [4 marks]
b. Find the 90th percentile of X. [2 marks]
c. Find $E[Y]$ and $\text{Var}(Y)$. [3 marks]

Question 7

Question 7 – A screening test *(Section B · Real-life practice)*

A disease affects 2% of a population. A screening test detects it in 90% of people who have it (sensitivity) and gives a negative result for 95% of people who do not (specificity).

Per 10 000 people	Test positive	Test negative	Total
Ill	?	?	200
Healthy	?	?	9 800

a. Complete the table.
b. Find the probability that a randomly chosen person tests positive.
c. Find the probability that a person with a positive result is ill. Comment.
d. A person who tested positive takes a second, independent test, which is also positive. Find the new probability of being ill.

Question 8

Question 8 – Faulty phone chargers *(Section B · Real-life practice)*

A shop buys phone chargers from three factories.

Factory	Share of supply	Proportion faulty
A	50%	2%
B	30%	3%
C	20%	5%

a. Find the probability that a randomly chosen charger is faulty.
b. A charger is faulty. Find the probability that it came from factory C.
c. Which factory is most likely to have produced a faulty charger?
d. The shop sells 2 000 chargers a month. How many faulty ones should it expect, and how many of those from factory A?

Question 9

Question 9 – Calls to a helpline *(Section B · Real-life practice)*

A student helpline receives calls at random at an average rate of 4 per 10 minutes. Model the number of calls in 10 minutes by $X \sim \mathrm{Po}(4)$.

a. State two conditions needed for the Poisson model.
b. Find the probability of no calls in a 10-minute period.
c. Find the probability of more than 6 calls in 10 minutes.
d. Find the probability of at most one call in 5 minutes.
e. State the distribution of the number of calls in half an hour, with its mean and variance.

Question 10

Question 10 – Overbooking a flight *(Section B · Real-life practice)*

An airline has 180 seats on a flight and sells 186 tickets. Each passenger independently fails to turn up with probability 0.05. Let Y be the number of no-shows.

a. State the distribution of Y and its mean.
b. The flight is overbooked if fewer than 6 passengers fail to turn up. Write this as a probability involving Y.
c. Calculate it with the binomial distribution.
d. Approximate it with a Poisson distribution and compare.
e. Find the expected number of passengers who turn up.

Question 11

Question 11 – A dice game at the school fair *(Section B · Real-life practice)*

At a school fair it costs £2 to roll two fair dice. A double six wins £10, any other double wins £4, and anything else wins nothing. Let W be the prize.

Prize w (£)	10	4	0
$P(W = w)$	?	?	?

a. Complete the table.
b. Find $E[W]$ and the expected profit per game for the school.
c. Find $\mathrm{Var}(W)$.
d. The school expects 300 games. Estimate the total profit.

Question 12

Question 12 – Heights of first-year students *(Section B · Real-life practice)*

The heights of first-year students at a university are modelled by $H \sim N(170,\, 7^2)$ cm.

a. The doors of an old lecture hall are 185 cm high. Find the proportion of students taller than this.
b. Find the proportion between 160 cm and 180 cm.
c. Find the height exceeded by only 5% of students.
d. In a group of 500 students, how many would you expect to be taller than 185 cm?

Question 13

Question 13 – Waiting for a bus *(Section B · Real-life practice)*

Buses on a route leave every 12 minutes. A student who arrives at a random time waits T minutes, with $T \sim U(0, 12)$. On another route buses arrive at random, and the waiting time S is exponential with mean 5 minutes, density $f(s) = \frac{1}{5}e^{-s/5}$.

a. Find $E[T]$, $\mathrm{Var}(T)$ and $P(T > 9)$.
b. Find $P(S > 8)$.
c. Show that $P(S > s + t \mid S > s) = P(S > t)$ and explain what this means for a student who has already waited.
d. Which route gives the longer average wait?

Question 14

Question 14 – Filling cereal boxes *(Section B · Real-life practice)*

The mass of cereal in a box is $C \sim N(500,\, 4^2)$ g and the mass of the empty box is $B \sim N(20,\, 1^2)$ g, independently.

a. Find the distribution of the total mass $M = C + B$.
b. Find $P(M$.
c. Two full boxes are chosen at random. Find the distribution of the difference of their masses and the probability that they differ by more than 10 g.
d. Explain why variances add even when we subtract.

Question 15

Question 15 – The birthday problem *(Section C · Problem solving)*

How many people must be in a room before it is more likely than not that two share a birthday? Ignore 29 February and assume all 365 days are equally likely. Let q_n be the probability that n people all have different birthdays.

n	10	20	23	30	50
$1 - q_n$	0.117	0.411	?	0.706	0.970

a. Explain why $q_n = \dfrac{365}{365} \times \dfrac{364}{365} \times \cdots \times \dfrac{365 - n + 1}{365}$.
b. Compute $1 - q_{23}$ and complete the table.
c. Using $1 - x \approx e^{-x}$, show that $q_n \approx e^{-n(n-1)/730}$ and solve $q_n = \frac{1}{2}$ approximately.
d. A class of 30 students: find the probability that at least two share a birthday and explain why it surprises most people.

Question 16

Question 16 – Pricing home insurance *(Section C · Problem solving)*

An insurer models the number of claims on one home policy in a year by $N \sim \mathrm{Po}(0.3)$; each claim costs £1 500. The company holds 1 000 independent policies.

Claims	0	1	2	3 or more
Probability	?	?	?	?

a. Complete the table to 4 decimal places.
b. Find the expected cost per policy and its standard deviation.
c. State the distribution of the total number of claims Y on 1 000 policies.
d. Use a normal approximation with continuity correction to estimate $P(Y > 330)$.
e. The annual premium is £480 per policy. Discuss whether this is enough to cover claims in most years.

Question 17

Question 17 – The Monty Hall problem *(Section C · Problem solving)*

In a TV game show a car is behind one of three doors and goats are behind the other two. You choose door 1. The host, who knows where the car is, always opens a different door showing a goat, and offers you the chance to switch. A class simulated the game 300 times.

Strategy	Games	Wins
Stick	150	49
Switch	150	101

a. Find the experimental probability of winning for each strategy.
b. Let C_i be "the car is behind door i" and H_3 "the host opens door 3". Find $P(H_3 \mid C_1)$, $P(H_3 \mid C_2)$, $P(H_3 \mid C_3)$.
c. Use Bayes' theorem to find $P(C_2 \mid H_3)$.
d. Explain the result in words, and say whether the simulation supports it.

Question 18

Question 18 – An opinion poll *(Section C · Problem solving)*

In a large city, 45% of voters support a new cycle lane. A newspaper asks a random sample of 400 voters. Let X be the number who support it.

a. State the exact distribution of X and its mean and variance.
b. Explain why a normal approximation is reasonable.
c. Estimate $P(X \geq 200)$, that is, the probability that the poll wrongly shows majority support, using a continuity correction.
d. Find the interval $180 \pm 2\sigma$ and interpret it as a proportion.
e. How large a sample would make the standard deviation of the sample proportion $\hat{p} = X/n$ at most 0.01?

Question 19

Question 19 – Cakes and coffee: a joint distribution *(Section C · Problem solving)*

A café records, for each customer, the number of cakes bought X and whether they buy a coffee ($Y = 1$) or not ($Y = 0$).

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	0.10	0.20	0.10
$Y = 1$	0.15	0.25	0.20

a. Find the marginal distributions of X and Y.
b. Find $E[X]$, $E[Y]$ and $E[XY]$.
c. Find $\mathrm{Cov}(X, Y)$.
d. Are X and Y independent? Explain carefully.
e. Coffee costs £3 and a cake £2.50. Find the expected spend per customer.

Question 20

Question 20 – Challenge: collecting all the toys *(Section D · Challenge)*

Each cereal box contains one of 6 different toy figures, equally likely and independent. A child buys boxes until the whole set is collected.

a. When the child has k different figures, show that the number of extra boxes needed for a new one is geometric with $p = \dfrac{6-k}{6}$.

b. Prove that a geometric random variable with $P(G=g)=(1-p)^{g-1}p$ has $E[G]=\dfrac{1}{p}$, using $\sum_{g\ge 1} g\,r^{g-1}=\dfrac{1}{(1-r)^2}$.
c. Deduce that the expected number of boxes is $6\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}\right)$ and evaluate it.
d. For n figures the answer is about $n\ln n$. Compare with the exact value for $n=6$ and comment.