Dr Oliver Mathematics
Advance Level Further Mathematics
Statistics 2: Calculator
1 hour 30 minutes

The total number of marks available is 75.
You must write down all the stages in your working.

1. In a call centre, the number of telephone calls, X, received during any 10-minute period follows a Poisson distribution with mean 9.

   (a) Find (4)

   (i) $\mathrm{P}(X > 5)$,

   Solution
   Let $X \sim \mathrm{Po}(9)$. Then

   $$\begin{aligned} \mathrm{P}(X > 5) &= \mathrm{P}(X \geqslant 6) \\ &= 1 - \mathrm{P}(X \leqslant 5) \\ &= 1 - 0.1157 \text{ (from tables)} \\ &= \underline{\underline{0.8843}}. \end{aligned}$$

   (ii) $\mathrm{P}(4 \leqslant X < 10)$.

   Solution

   $$\begin{aligned} \mathrm{P}(4 \leqslant X < 10) &= \mathrm{P}(X \leqslant 9) - \mathrm{P}(X \leqslant 3) \\ &= 0.5874 - 0.0212 \\ &= \underline{\underline{0.5662}}. \end{aligned}$$

   The length of a working day is 7 hours.

   (b) Using a suitable approximation, find the probability that there are fewer than 370 telephone calls in a randomly selected working day. (5)

   Solution
   The amount of calls received is

   $$9 \times 6 \times 7 = 378$$

and let $Y \sim \text{Po}(378)$ and $W \approx\sim \text{N}(378, 378)$. Now,

$$\begin{aligned}
\text{P}(Y < 370) &= \text{P}(Y \leqslant 369.5) \\
&\approx \text{P}(W \leqslant 369.5) \\
&= \text{P}\left(Z \leqslant \frac{369.5 - 378}{\sqrt{378}}\right) \\
&= \text{P}(Z \leqslant -0.44) \text{ (2 dp)} \\
&= \text{P}(Z \geqslant 0.44) \\
&= 1 - \text{P}(Z \leqslant 0.44) \\
&= 1 - 0.6700 \\
&= \underline{\underline{0.3300}}.
\end{aligned}$$

A week, consisting of 5 working days, is selected at random.

(c) Find the probability that in this week at least 4 working days have fewer than 370 telephone calls. (3)

Solution

Let $W \sim B(5, 0.3300)$. Then

$$\begin{aligned}
\text{P}(\text{at least 4 working days}) &= \text{P}(W = 4) + \text{P}(W = 5) \\
&= \binom{5}{4}(0.3300)^4(0.6700) + (0.3300)^5 \\
&= 0.043\,641\,892\,8 \text{ (FCD)} \\
&= \underline{\underline{0.043\,6}}.
\end{aligned}$$

2. A fair coin is spun 6 times and the random variable T represents the number of tails obtained.

(a) Give two reasons why a binomial model would be a suitable distribution for modelling T. (2)

Solution

There is a fixed number of trials, each trial should be success or failure, the trials are independent, and the probability of success is constant.

(b) Find $\text{P}(T = 5)$. (2)

Solution

Let $T \sim \mathrm{B}(6, 0.5)$. Now,

$$\begin{aligned} \mathrm{P}(T = 5) &= \binom{6}{5}(0.5)^5(0.5) \\ &= \underline{\underline{\tfrac{3}{32}}}. \end{aligned}$$

(c) Find the probability of obtaining more tails than heads. (2)

Solution

$$\begin{aligned} \mathrm{P}(\text{more tails than heads}) &= \mathrm{P}(4T, 2H) + \mathrm{P}(5T, 1H) + \mathrm{P}(6T) \\ &= \binom{6}{4}(0.5)^4(0.5)^2 + \binom{6}{5}(0.5)^5(0.5)^1 + (0.5)^6 \\ &= \tfrac{1}{64}(15 + 6 + 1) \\ &= \underline{\underline{\tfrac{11}{32}}}. \end{aligned}$$

A second coin is biased such that the probability of obtaining a head is $\frac{1}{4}$.

This second coin is spun 6 times.

(d) Find the probability that, for the second coin, the number of heads obtained is greater than or equal to the number of tails obtained. (3)

Solution

Let $S \sim \mathrm{B}(6, 0.25)$. Then

$$\begin{aligned} & \mathrm{P}(S \geqslant 3) \\ = \; & \mathrm{P}(3H, 3T) + \mathrm{P}(4H, 2T) + \mathrm{P}(5H, 1T) + \mathrm{P}(6H) \\ = \; & \binom{6}{3}(0.25)^3(0.75)^3 + \binom{6}{4}(0.25)^4(0.75)^2 + \binom{6}{5}(0.25)^5(0.75)^1 + (0.25)^6 \\ = \; & \tfrac{135}{1\,024} + \tfrac{135}{4\,096} + \tfrac{9}{2\,048} + \tfrac{1}{4\,096} \\ = \; & \underline{\underline{\tfrac{347}{2\,048}}}. \end{aligned}$$

3. The length of time, T, minutes, spent completing a particular task has probability density

function

$$f(t) = \begin{cases} \frac{1}{2}(t-1) & 1 < t \leqslant 2, \\ \frac{1}{16}(14t - 3t^2 - 8) & 2 < t \leqslant 4, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Use algebraic integration to find $E(T)$. (4)

Solution

$$\begin{aligned} E(T) &= \int_1^2 t \cdot \tfrac{1}{2}(t-1)\,dt + \int_2^4 t \cdot \tfrac{1}{16}(14t - 3t^2 - 8)\,dt \\ &= \int_1^2 \tfrac{1}{2}(t^2 - t)\,dt + \int_2^4 \tfrac{1}{16}(14t^2 - 3t^3 - 8t)\,dt \\ &= \tfrac{1}{2}\left[\tfrac{1}{3}t^3 - \tfrac{1}{2}t^2\right]_{t=1}^2 + \tfrac{1}{16}\left[\tfrac{14}{3}t^3 - \tfrac{3}{4}t^4 - 4t^2\right]_{t=2}^4 \\ &= \tfrac{1}{2}\left[(\tfrac{8}{3} - 2) - (\tfrac{1}{3} - \tfrac{1}{2})\right] + \tfrac{1}{16}\left[(\tfrac{896}{3} - 192 - 64) - (\tfrac{112}{3} - 12 - 16)\right] \\ &= \tfrac{5}{12} + 2\tfrac{1}{12} \\ &= \underline{\underline{2\tfrac{1}{2}}}. \end{aligned}$$

Given that $E(T^2) = \frac{267}{40}$,

(b) find $\text{Var}(T)$. (2)

Solution

$$\begin{aligned} \text{Var}(T) &= E(T^2) - E(T)^2 \\ &= \tfrac{267}{40} - (2\tfrac{1}{2})^2 \\ &= \underline{\underline{\tfrac{17}{40}}}. \end{aligned}$$

(c) Find the cumulative distribution function $F(t)$. (5)

Solution

$\underline{1 < t \leqslant 2}$:

$$\begin{aligned} \mathrm{F}(t) &= \int_1^t \tfrac{1}{2}(x-1)\,\mathrm{d}x \\ &= \tfrac{1}{2}\left[\tfrac{1}{2}x^2 - x\right]_{x=1}^t \\ &= \tfrac{1}{2}\left[(\tfrac{1}{2}t^2 - t) - (\tfrac{1}{2} - 1)\right] \\ &= \tfrac{1}{2}(\tfrac{1}{2}t^2 - t + \tfrac{1}{2}) \\ &= \tfrac{1}{4}(t^2 - 2t + 1) \\ &= \tfrac{1}{4}(t-1)^2. \end{aligned}$$

and

$$\mathrm{F}(2) = \tfrac{1}{4}(2-1)^2 = \tfrac{1}{4}.$$

$\underline{2 < t \leqslant 4}$:

$$\begin{aligned} \mathrm{F}(t) &= \tfrac{1}{4} + \int_2^t \tfrac{1}{16}(14x - 3x^2 - 8)\,\mathrm{d}x \\ &= \tfrac{1}{4} + \tfrac{1}{16}\left[7x^2 - x^3 - 8x\right]_{x=2}^t \\ &= \tfrac{1}{4} + \tfrac{1}{16}(7t^2 - t^3 - 8t - 4) \\ &= \tfrac{1}{16}(7t^2 - t^3 - 8t). \end{aligned}$$

Hence,

$$\underline{\underline{\mathrm{F}(t) = \begin{cases} 0 & t \leqslant 1, \\ \tfrac{1}{4}(t-1)^2 & 1 < t \leqslant 2, \\ \tfrac{1}{16}(7t^2 - t^3 - 8t) & 2 < t \leqslant 4, \\ 1 & t > 4. \end{cases}}}$$

(d) Find the 20th percentile of the time taken to complete the task. (3)

Solution

$$\begin{aligned} \tfrac{1}{4}(t-1)^2 = 0.2 &\Rightarrow (t-1)^2 = 0.8 \\ &\Rightarrow t - 1 = \tfrac{2\sqrt{5}}{5} \\ &\Rightarrow \underline{\underline{t = \tfrac{5+2\sqrt{5}}{5} \text{ or } 1.89 \text{ minutes (3 sf)}}}. \end{aligned}$$

(e) Find the probability that the time spent completing the task is more than 1.5 minutes. (2)

Solution

$$\begin{aligned} F(1.5) &= \tfrac{1}{4}(1.5-1)^2 \\ &= \tfrac{1}{16} \end{aligned}$$

and the probability that the time spent completing the task is more than 1.5 minutes is

$$\begin{aligned} 1 - F(1.5) &= 1 - \tfrac{1}{16} \\ &= \underline{\underline{\tfrac{15}{16}}}. \end{aligned}$$

Given that a person has already spent 1.5 minutes on the task,

(f) find the probability that this person takes more than 3 minutes to complete the task. (2)

Solution

$$\begin{aligned} F(3) &= \tfrac{1}{16}(7 \times 3^2 - 3^3 - 8 \times 3) \\ &= \tfrac{3}{4} \end{aligned}$$

and the probability that the time spent completing the task is more than 3 minutes is

$$\begin{aligned} 1 - F(3) &= 1 - \tfrac{3}{4} \\ &= \tfrac{1}{4}. \end{aligned}$$

Hence, given that a person has already spent 1.5 minutes on the task, the probability that the time spent completing the task is more than 3 minutes is

$$\frac{\frac{1}{4}}{\frac{15}{16}} = \underline{\underline{\tfrac{4}{15}}}.$$

4. David aims to catch the train to work each morning. The scheduled departure time of the train is 0830.

The number of minutes after 0830 that the train departs may be modelled by the random variable X. Given that X has a continuous uniform distribution over $[\alpha, \beta]$ and that $E(X) = 4$ and $\mathrm{Var}(X) = 12$,

(a) find the value of α and the value of β. (5)

Solution

$$\mathrm{E}(X) = 4 \Rightarrow \frac{\alpha+\beta}{2} = 4$$

and

$$\mathrm{Var}(X) = 12 \Rightarrow \frac{(\beta-\alpha)^2}{12} = 12.$$

Now,

$$\begin{aligned}\frac{\alpha+\beta}{2} = 4 &\Rightarrow \alpha+\beta = 8\\ &\Rightarrow \alpha = 8-\beta\end{aligned}$$

and

$$\begin{aligned}\frac{(\beta-\alpha)^2}{12} = 12 &\Rightarrow (2\beta-8)^2 = 144\\ &\Rightarrow 2\beta - 8 = 12\\ &\Rightarrow 2\beta = 20\\ &\Rightarrow \underline{\underline{\beta = 10}}\\ &\Rightarrow \underline{\underline{\alpha = -2}}.\end{aligned}$$

Each morning, the probability that David oversleeps is 0.05.

If David oversleeps he will be late for work.

If he does not oversleep he will be in time to catch the train, but will be late for work if the train departs after 0835.

(b) Find the probability that David will be late for work. (3)

Solution

The probability that the train will arrive late is $\frac{5}{12}$ and

$$\begin{aligned}\mathrm{P}(\text{David will arrive late}) &= 0.05 + 0.95 \times \tfrac{5}{12}\\ &= \underline{\underline{\tfrac{107}{240}}}.\end{aligned}$$

Given that David is late for work,

(c) find the probability that he overslept. (2)

> **Solution**
>
> $$\begin{aligned} \mathrm{P}(\text{missed train}|\text{late}) &= \frac{\mathrm{P}(\text{overslept})}{\mathrm{P}(\text{late})} \\ &= \frac{0.05}{\frac{107}{240}} \\ &= \underline{\underline{\tfrac{12}{107}}}. \end{aligned}$$

5. Past records show that the proportion of customers buying organic vegetables from *Tesson* supermarket is 0.35.

During a particular day, a random sample of 40 customers from *Tesson* supermarket was taken and 18 of them bought organic vegetables.

(a) Test, at the 5% level of significance, whether or not this provides evidence that the proportion of customers who bought organic vegetables has increased. State your hypotheses clearly. (5)

> **Solution**
>
> H_0: the rate of organic vegetables is unchanged, $p = 0.35$.
> H_1: the rate of organic vegetables is higher, $p > 0.35$.
> Level of significance: 5%.
> Let $V \sim \mathrm{B}(40, 0.35)$.
> Then
>
> $$\begin{aligned} \mathrm{P}(V \geqslant 18) &= 1 - \mathrm{P}(V \leqslant 17) \\ &= 1 - 0.8761 \\ &= 0.1239 \\ &= 12.39\%. \end{aligned}$$
>
> Now, $12.39\% > 5\%$ so there is not significant enough evidence to reject H_0.
> Hence, at the 5% level, there is insufficient evidence that the proportion of customers buying organic vegetables from *Tesson* supermarket has increased.

The manager of *Tesson* supermarket claims that the proportion of customers buying organic eggs is different from the proportion of those buying organic vegetables. To test this claim the manager decides to take a random sample of 50 customers.

(b) Using a 5% level of significance, find the critical region to enable the *Tesson* supermarket manager to test her claim. The probability for each tail of the region (3) should be as close as possible to 2.5%.

Solution

H_0: the rate of organic eggs is unchanged, $p = 0.35$.
H_1: the rate of organic eggs is higher, $p \neq 0.35$.
Level of significance: 5%.
Let $E \sim \mathrm{B}(50, 0.35)$.
Then,

$$\begin{aligned}\mathrm{P}(E \leqslant 10) &= 0.0160,\\ \mathrm{P}(E \leqslant 11) &= 0.0342,\\ \mathrm{P}(E \geqslant 24) &= 1 - \mathrm{P}(E \leqslant 23) = 1 - 0.9604 = 0.0396,\\ \mathrm{P}(E \geqslant 25) &= 1 - \mathrm{P}(E \leqslant 24) = 1 - 0.9793 = 0.0207.\end{aligned}$$

Hence, $\underline{\underline{E \leqslant 10 \text{ and } E \geqslant 25}}$.

During a particular day, a random sample of 50 customers from *Tesson* supermarket is taken and 8 of them bought organic eggs.

(c) Using your answer to part (b), state whether or not this sample supports the manager's claim. Use a 5% level of significance. (1)

Solution

The manager's claim is supported; there is sufficient evidence that the proportion of customers buying organic eggs is different from those buying organic vegetables.

(d) State the actual significance level of this test. (1)

Solution

$$0.0160 + 0.0207 = 0.0367;$$

hence, $\underline{\underline{3.67\%}}$.

The proportion of customers who buy organic fruit from *Tesson* supermarket is 0.2.

During a particular day, a random sample of 200 customers from *Tesson* supermarket is taken. Using a suitable approximation, the probability that fewer than n of these customers bought organic fruit is 0.0465, correct to 4 decimal places.

(e) Find the value of n. (6)

Solution

H_0: the rate of organic fruit is unchanged, $p = 0.2$.
H_1: the rate of organic fruit is higher, $p > 0.2$.
Level of significance: 5%.
Let $F \sim \mathrm{B}(200, 0.2)$.
Then,

$$\mathrm{E}(F) = 200 \times 0.2 = 40$$
$$\mathrm{Var}(F) = 200 \times 0.2 \times 0.8 = 32.$$

and so we can use the Normal approximation: $F \approx\sim \mathrm{N}(40, 32)$. Now,

$$\begin{aligned} \mathrm{P}(F < n) &= \mathrm{P}(F \leqslant n - 1) \\ &= \mathrm{P}\left(F \leqslant \frac{n - \frac{1}{2} - 40}{\sqrt{32}}\right). \end{aligned}$$

Now,

$$\Phi(1.68) = 0.9535 \Rightarrow 1 - 0.9535 = 0.0465.$$

Finally,

$$\begin{aligned} \frac{n - \frac{1}{2} - 40}{\sqrt{32}} = -1.68 &\Rightarrow n - \tfrac{1}{2} - 40 = -9.503\,515\,139 \text{ (FCD)} \\ &\Rightarrow n = 30.996\,484\,86 \text{ (FCD)}; \end{aligned}$$

hence, $\underline{\underline{n = 31 \text{ (nearest integer)}}}$.

6. The continuous random variable X has the following cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0 & x \leqslant 1, \\ \frac{4}{15}(x-1) & 1 < x \leqslant 2, \\ k\left(\frac{ax^3}{3} - \frac{x^4}{4}\right) + b & 2 < x \leqslant 4, \\ 1 & x > 4, \end{cases}$$

where k, a, and b are constants.

Given that the mode of X is $\frac{8}{3}$,

(a) show that $a = 4$. (4)

Solution

The probability density function (pdf) is

$$\mathrm{f}(x) = \begin{cases} 0 & x \leqslant 1, \\ \frac{4}{15} & 1 < x \leqslant 2, \\ k\left(ax^2 - x^3\right) & 2 < x \leqslant 4, \\ 0 & x > 4. \end{cases}$$

Now,

$$\begin{aligned} \mathrm{f}'(x) = 0 &\Rightarrow k\left(2ax - 3x^2\right) = 0 \\ &\Rightarrow kx(2a - 3x) = 0 \\ &\Rightarrow x = 0 \text{ or } x = \frac{2a}{3}. \end{aligned}$$

Now, the mode of X is $\frac{8}{3}$ which means

$$a = \tfrac{3}{2} \times \tfrac{8}{3} = \underline{\underline{4}},$$

as required.

(b) Find $\mathrm{P}(X < 2.5)$, giving your answer to 3 significant figures. (6)

Solution

For $x = 2$,

$$\begin{aligned} \mathrm{F}(2) = \tfrac{4}{15} &\Rightarrow \tfrac{4}{15}(2-1) = k\left(\frac{4 \times 2^3}{3} - \frac{2^4}{4}\right) + b \\ &\Rightarrow \tfrac{4}{15} = \tfrac{20}{3}k + b \quad (1). \end{aligned}$$

For $x = 4$,

$$\begin{aligned} \mathrm{F}(4) = 1 &\Rightarrow k\left(\frac{4 \times 4^3}{3} - \frac{4^4}{4}\right) + b = 1 \\ &\Rightarrow \tfrac{64}{3}k + b = 1 \quad (2). \end{aligned}$$

Now, $(2) - (1)$,

$$\begin{aligned} \tfrac{44}{3}k = \tfrac{11}{15} &\Rightarrow k = \tfrac{1}{20} \\ &\Rightarrow b = -\tfrac{1}{15}. \end{aligned}$$

Finally,

$$\begin{aligned} \mathrm{P}(X < 2.5) &= \mathrm{F}(2.5) \\ &= \tfrac{1}{20}\left(\frac{4 \times 2.5^3}{3} - \frac{2.5^4}{4}\right) - \tfrac{1}{15} \\ &= \underline{\underline{\tfrac{623}{1\,280}}}. \end{aligned}$$