Probability and Statistics
Solutions

Probability

1. Draw six diagrams similar to figure 1 and shade the following sets

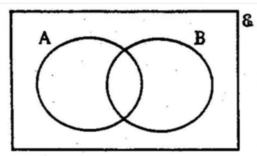
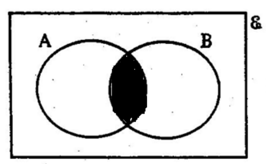
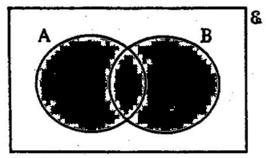


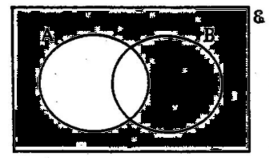
Figure 1

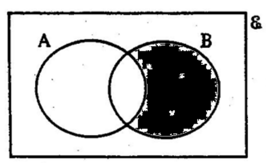
   (a) $A \cap B$
   (b) $A \cup B$
   (c) A'
   (d) $A' \cap B$
   (e) $B' \cap A$
   (f) $(B \cup A)'$

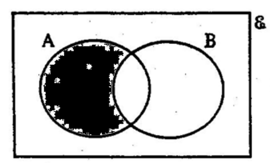
Solution 1

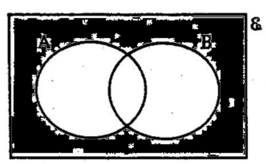












2. There are 176 students at a college following a general course in computing. Students on this course can choose to take up to three extra options.

   112 take systems support

   78 take developing software

   81 take networking

41 take developing software and systems support

34 take networking and developing software

43 take systems support and networking

8 take all three extra options

(a) Draw a Venn Diagram to represent this information.

A student from the course is chosen at random. Find the probability that the student takes

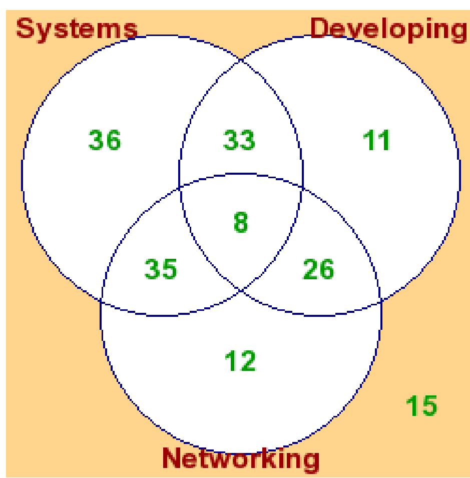
(b) none of the three extra options

(c) networking only.

Students who want to become technicians take systems support or networking. Given that a randomly chosen student wants to become a technician,

(d) find the probability that this student takes all three extra options

Solution 2

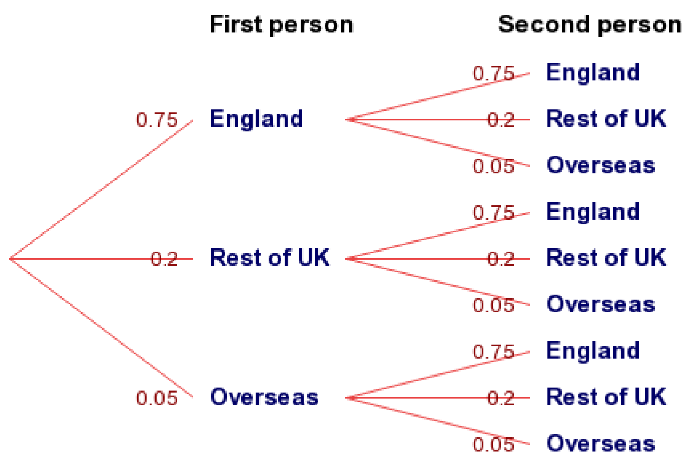


(b) P = $\frac{15}{176}$

(c) P = $\frac{12}{176}$

(d) P(All 3 | Technician) = $\frac{\frac{8}{176}}{\frac{150}{176}} = \frac{8}{150}$

3. In a large town 75% of the population were born in England, 20% in the rest of the UK and 5% abroad. Two people are selected at random.



You may use the above tree diagram in answering this question. Find the probability that

(a) both these people were born in the rest of the UK.

(b) at least one of these people was born in England.

(c) neither of these people was born overseas.

(d) Find the probability that both these people were born in the rest of the UK given that neither was born overseas.

(e) 6 people are selected at random. Find the probability that at least one of them was not born in England.

(f) An interviewer selects n people at random. The interviewer wishes to ensure that the probability that at least one of them was not born in England is more than 80%. Find the least possible value of n.

Solution 3

(a) P(both in rest of UK) $= 0.2^2 = 0.04$

(b) P(at least 1 England) $= 0.75 + (0.2 \times 0.75) + (0.05 \times 0.75)$

(c) P(neither overseas) $= (1 - 0.05)^2 = 0.9025$

(d)

A: both in rest of UK
B: neither overseas

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.04}{0.9025} = 0.0443$$

(e) P $= 1 - 0.75^6 = 0.822$

(f)

$$\begin{aligned}
1-0.75^n &> 0.8 \\
0.75^n &< 0.2 \\
n &> \frac{\log 0.2}{\log 0.75} \\
&= 5.5945 \\
\therefore n &= 6
\end{aligned}$$

4. The events A and B are independent such that $P(A) = 0.14$ and $P(B) = 0.23$. Find

   (a) $P(A \cap B)$

   (b) $P(A \cup B)$

   (c) $P(A|B')$

Solution 4

(a) $P(\cap B) = P(A) \times P(B) = 0.14 \times 0.23 = 0.0322$

(b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.14 + 0.23 - 0.0322 = 0.3378$

(c) $P(A|B') = \frac{P(A \cap B')}{P(B')} = \frac{0.14 \times (1-0.23)}{1-0.23} = 0.14$

5. The discrete random variable X can take only the values $8, 9$ or 10. For these values the cumulative distribution function is defined by

$$F(x) = \frac{(x+k)^2}{144} \qquad \text{for} \quad x = 8, 9, 10$$

where k is a positive integer

   (a) Find k.

   (b) Find the probability distribution of X.

Solution 5

(a)

$$\begin{aligned}
F(10) &= 1 \\
\frac{(10+k)^2}{144} &= 1 \\
(10+k)^2 &= 144 \\
\therefore k &= 2
\end{aligned}$$

(b)

$$
\begin{aligned}
P(X=8) &= F(8) = \frac{100}{144} \\
P(X=9) &= F(9) - F(8) = \frac{121}{144} - \frac{100}{144} = \frac{21}{144} \\
P(X=10) &= F(10) - F(9) = \frac{144}{144} - \frac{121}{144} = \frac{23}{144}
\end{aligned}
$$

x	8	9	10
$P(X=x)$	$\frac{100}{144}$	$\frac{21}{144}$	$\frac{23}{144}$

6. The discrete random variable X has the probability function $f(x)$ defined by

$$f(x) = kx^2 \qquad x = 2, 3, 4, 5, 6$$

(a) Construct a table showing the probability distribution of the random variable X.

(b) Find the value of k.

(c) Find $E(X)$ and $Var(X)$.

(d) Find the mean and variance of the random variable Y where $Y = 3X - 7$

Solution 6

(a)

x	2	3	4	5	6
$f(x)$	4k	9k	16k	25k	36k

(b)

$$
\begin{aligned}
\sum f(x) &= 1 \\
(4 + 9 + 16 + 25 + 36)k &= 1 \\
k &= \frac{1}{90}
\end{aligned}
$$

(c)

$$
\begin{aligned}
E(X) &= (8+27+64+125+216)k \\
&= 4.89 \\
E(X^2) &= (16+81+256+625+1296)k \\
&= 25.27 \\
Var(X) &= E(X^2) - E^2(X) = 1.37
\end{aligned}
$$

(d)

$$
\begin{aligned}
E(3X-7) &= 3E(X) - 7 = 7.67 \\
Var(3X-7) &= 9Var(X) = 12.29
\end{aligned}
$$

7. Bob plays 12 squash games. In each game he either wins or loses.

   (a) State, in this context, two conditions needed for a binomial distribution to arise.

   (b) Assuming these conditions are satisfied, define a variable in this context which has a binomial distribution.

   (c) The random variable X has the distribution $B(24, p)$ where $0 < p < 1$. Given that $P(X = 7) = P(X = 6)$ find the value of p.

Solution 7

(a)

- Results are independent
- Probability of winning is constant

(b) Variable: Number of wins

(c)

$$
\begin{aligned}
P(X=7) &= P(X=6) \\
{}^{24}C_7 p^7 (1-p)^{17} &= {}^{24}C_6 p^6 (1-p)^{18} \\
\frac{24!}{7! \times 17!} p &= \frac{24!}{6! \times 18!}(1-p) \\
\frac{1}{7} p &= \frac{1}{18}(1-p) \\
18p &= 7(1-p) \\
25p &= 7 \\
p &= \frac{7}{25}
\end{aligned}
$$

8. The continuous random variable X has probability density function $f(x)$ given by

$$f(x) = \begin{cases} kx^2(7-x) & 0 \le x \le 7 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $k = \frac{12}{2401}$

(b) Find $E(X)$

(c) Find $P(X < 4)$

Solution 8

(a)

$$\begin{aligned} \int_0^7 f(x)dx &= 1 \\ \int_0^7 kx^2(7-x)dx &= 1 \\ k\left[\frac{7}{3}x^3 - \frac{1}{4}x^4\right]_0^7 &= 1 \\ k\left[\frac{2401}{3} - \frac{2401}{4}\right] &= 1 \\ \therefore k &= \frac{12}{2401} \end{aligned}$$

(b)

$$\begin{aligned} E(X) &= \int_0^7 x.kx^2(7-x)dx \\ &= k\int_0^7 (7x^3 - x^4)fx \\ &= k\left[\frac{7}{4}x^4 - \frac{1}{5}x^5\right]_0^7 \\ &= \frac{21}{5} \end{aligned}$$

(c)

$$P(X < 4) = \frac{12}{2041}\left[\frac{7}{3}x^3 - \frac{1}{4}x^4\right]_0^4 = 0.426$$

9. The length of a telephone call made to a company is denoted by the continuous random variable T. It is modelled by the probability density function

$$f(x) = \begin{cases} kt & 0 \leq t \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that the value of k is $\frac{1}{18}$
(b) Find $P(T > 1)$
(c) Calculate an exact value for $E(T)$ and for $Var(T)$

Solution 9

(a)

$$\begin{aligned} \int ktdx &= 1 \\ \int_0^6 ktdx &= 1 \\ \left[\frac{1}{2}kt^2\right]_0^6 &= 1 \\ 8k &= 1 \\ k &= \frac{1}{18} \end{aligned}$$

(b)

$$P(T > 1) = \int_1^6 ktdx = \left[\frac{1}{2}kt^2\right]_1^6 = \frac{35}{36}$$

(c)

$$\begin{aligned} E(T) &= \int_0^6 kt^2dx = \left[\frac{1}{3}kt^3\right]_0^6 = \frac{4}{1} \\ E(T^2) &= \int_0^6 kt^3dx = \left[\frac{1}{4}kt^4\right] = \frac{!8}{1} \\ Var(T) &= \frac{18}{1} - \left(\frac{4}{1}\right)^2 = \frac{2}{1} \end{aligned}$$

10. The continuous random variable X has cumulative distribution function $F(x)$ given by

$$F(x)) = \begin{cases} 0 & x < 0 \\ \frac{1}{4}x^2(5 - x^2) & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

(a) Find $P(X > 0.6)$

(b) Find the probability density function $f(x)$ of X

(c) Calculate $E(X)$ and show that, to 3 decimal places, $Var(X) = 0.057$

Solution 10

(a)

$$\begin{aligned} P(X > 0.6) &= 1 - P(X < 0.6) \\ &= 1 - \frac{1}{4}(0.6)^2(5 - 0.6^2) \\ &= 0.582 \end{aligned}$$

(b)

$$f(x) = \frac{dF(x)}{dx} = \frac{10}{4}x - x^3$$

(c)

$$\begin{aligned} E(X) &= \int_0^1 x\left(\frac{10}{4}x - x^3\right)dx \\ &= \int_0^1 \frac{10}{4}x^2 - x^4 dx \\ &= \left[\frac{10}{12}x^3 - \frac{4}{20}x^5\right]_0^1 \\ &= 0.633 \\ E(X^2) &= \int_0^1 x^2\left(\frac{10}{4}x - x^3\right)dx \\ &= \int_0^1 \frac{10}{4}x^3 - x^5 dx \\ &= \left[\frac{10}{16}x^4 - \frac{4}{24}x^6\right]_0^1 = 0.458 \\ Var(X) &= E(X^2) - [E(X)]^2 = 0.458 - 0.633^2 = 0.057 \end{aligned}$$

For the following questions use Standard Normal Distribution tables which can be found at the back of any statistics or probability text book

11. The lifetimes of bulbs used in a lamp are normally distributed. A company X sells bulbs with a mean lifetime of 856 hours and a standard deviation of 58 hours.

(a) Find the probability of a bulb, from company X, having a lifetime of less than 833 hours

(b) In a box of 400 bulbs, from company X, find the expected number having a lifetime of less than 833 hours.

A rival company Y sells bulbs with a mean lifetime of 882 hours and 19% of these bulbs have a lifetime of less than 830 hours.

(c) Find the standard deviation of the lifetimes of bulbs from company Y.

Both companies sell bulbs for the same price.

(d) State which company you would recommend

Solution 11

(a)

$$\begin{aligned} P(X < 833) &= \Phi\left(\frac{833-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{833-856}{58}\right) \\ &= \Phi(-0.397) = 0.345 \end{aligned}$$

(b) Expected number $= 400 \times 0.345 = 138$

(c)

$$\begin{aligned} P(Y < 830) &= 0.19 \\ \Phi\left(\frac{830-882}{\sigma}\right) &= 0.19 \\ \frac{830-882}{\sigma} &= -0.8778 \\ \therefore \sigma &= 59.2 \end{aligned}$$

(d) Choose company Y because higher mean but smaller standard deviation

12. In large-scale tree-felling operations, a machine cuts down trees, strips off the branches and then cuts the trunks into logs of length X metres for transporting to a sawmill. It may be assumed that values of X are normally distributed with mean 4.3 and standard deviation 0.17.

(a) Determine $P(X < 4.5)$

(b) Determine $P(X > 4)$

(c) Determine $P(4 < X < 4.5)$

Solution 12

(a)

$$\begin{aligned} S_{xx} &= \sum x^2 - \frac{(\sum x)^2}{n} = 211 - \frac{(45)^2}{10} = 8.5 \\ S_{xy} &= \sum xy - \frac{(\sum x)(\sum y)}{n} = 1460 - \frac{45 \times 321}{10} = 15.5 \end{aligned}$$

(b)

$$\begin{aligned} b &= \frac{S_{xy}}{S_{xx}} = \frac{15.5}{8.5} = 1.824 \\ a &= \bar{y} - b\bar{x} = \frac{321}{10} - 1.824\frac{45}{10} = 23.892 \\ y &= 23.89 + 1.82x \end{aligned}$$

(c) Comment: A typical car will travel 1820 miles every year

(d)

$$x = 4.5 \qquad y = 23.89 + 1.82(4.5) = 32.08$$

mileage = 32080

13. A second hand car dealer has 10 cars for sale. She decides to investigate the link between the age of the cars, x years, and the mileage, y thousand miles. The data collected from the cars are shown in the table below.

Age (years) x	2.5	3.5	5.5	4.5	5.5	5.5	5	4.5	4	4.5
Mileage (1000) y	33	24	39	31	34	32	22	45	21	40

You may assume that

$$\sum x = 45; \quad \sum y = 321; \quad \sum x^2 = 221; \quad \sum xy = 1460;$$

(a) Find S_{xx} and S_{xy}

(b) Find the equation of the regression line of $yonx$ in the form $y = a+bx$. Give the values of a and b to 2 decimal places

(c) Give a practical interpretation of the slope b

(d) Using your answer to part (b), find the mileage predicted by the regression line for a 45 year old car

Solution 13

(a)

$$\begin{aligned} P(x < 4.5) &= \Phi\left(\frac{4.5-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{4.5-4.3}{0.17}\right) \\ &= \Phi(1.176) = 0.881 \end{aligned}$$

(b)

$$\begin{aligned} P(X > 4) &= 1 - P(X < 4) \\ &= 1 - \Phi\left(\frac{4-\mu}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{4-4.3}{0.17}\right) \\ &= 1 - \Phi(-1.765) \\ &= 1 - 0.0392 \\ &= 0.961 \end{aligned}$$

(c)

$$\begin{aligned} P(4 < X < 4.5) &= 0.881 - (1 - 0.961) \\ &= 0.842 \end{aligned}$$

14. A sample of bivariate data was taken and the results were summarised as follows:

$$n = 10; \quad \sum x = 771; \quad \sum x^2 = 60379; \quad \sum y = 723;$$

$$\sum y^2 = 53125; \quad \sum xy = 55905;$$

Show that the value of the the product moment correlation coefficient is 0181, correct to 3 significant figures

Solution 14

$$
\begin{aligned}
S_{xx} &= \sum x^2 - \frac{(\sum x)^2}{n} = 60379 - \frac{(771)^2}{10} = 934.9 \\
S_{yy} &= \sum y^2 - \frac{(\sum y)^2}{n} = 53125 - \frac{(723)^2}{10} = 852.1 \\
S_{xy} &= \sum xy - \frac{(\sum x)(\sum y)}{n} = 55905 - \frac{771 \times 723}{10} = 161.7
\end{aligned}
$$

$$
r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{161.7}{\sqrt{934.9 \times 852.1}} = 0.181
$$