

## OCR Statistics 1 Solutions - June 2007

- 1 To find  $E[X]$ , multiply the values of  $x$  by the probabilities:  
 $E[X] = (0 \times 0.1) + (1 \times 0.2) + (2 \times 0.3) + (3 \times 0.4) = 2.$

To find  $\text{Var}[X]$ , we use the formula  $\text{Var}[X] = E[X^2] - \{E[X]\}^2$ .  
 We can find  $E[X^2]$  by drawing up a table showing values of  $x^2$ :

$x^2$	0	1	4	9
$P(X = x)$	0.1	0.2	0.3	0.4

Therefore,  $E[X^2] = (0 \times 0.1) + (1 \times 0.2) + (4 \times 0.3) + (9 \times 0.4) = 5$

So,  $\text{Var}[X] = 5 - 2^2 = 1.$

- 2 To find Spearman's rank correlation coefficient, we need to display the table in a different way:

	UK	France	Russia	Poland	Canada
Rank Judge 1	1	2	3	4	5
Rank Judge 2	4	3	1	5	2
d	-3	-1	2	-1	3
d <sup>2</sup>	9	1	4	1	9

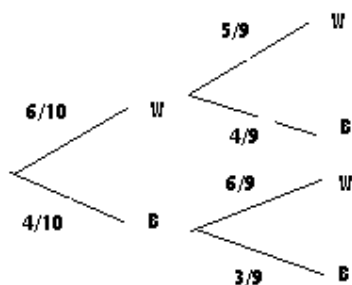
So,  $\sum d^2 = 24.$

Therefore,  $r_s = 1 - \frac{6 \times \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \times 24}{5(5^2 - 1)} = -0.2$

- 3 (i) Number of teams =  ${}^{15}C_7 = 6435$  (combinations as order is not important)

- (ii) Number of ways of choosing 3 forwards out of 6 =  ${}^6C_3 = 20$   
 Number of ways of choosing 4 defenders out of 9 =  ${}^9C_4 = 126$   
 So number of ways of choosing 3 forwards AND 4 defenders =  $20 \times 126 = 2520$

- 4 (i) We can draw a tree diagram:

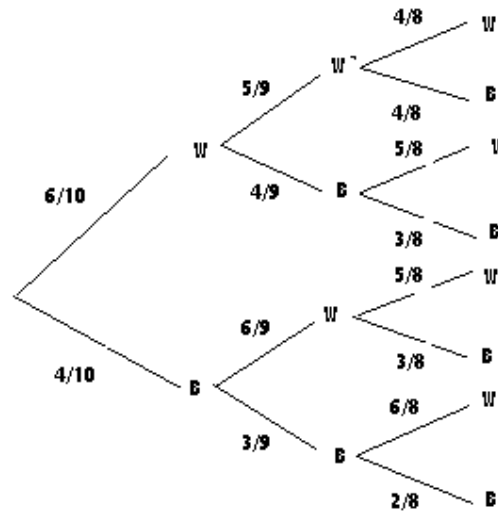


- a) From the tree diagram, the probability that the second disc is blue, given that the first disc was blue is  $5/9.$

b) The probability that the 2<sup>nd</sup> disc is blue is:

$$\left(\frac{6}{10} \times \frac{4}{9}\right) + \left(\frac{4}{10} \times \frac{3}{9}\right) = \frac{36}{90} = \frac{2}{5}$$

c) We can extend the tree diagram to include the 3<sup>rd</sup> disc selected:



$$P(3^{\text{rd}} \text{ disc blue} \mid 1^{\text{st}} \text{ disc blue}) = \frac{P(3^{\text{rd}} \text{ disc blue and } 1^{\text{st}} \text{ disc blue})}{P(1^{\text{st}} \text{ disc blue})}$$

$$\begin{aligned} \text{From the tree diagram, } P(3^{\text{rd}} \text{ disc blue and } 1^{\text{st}} \text{ disc blue}) &= P(\text{BBB}) + P(\text{BWB}) \\ &= \left(\frac{4}{10} \times \frac{3}{9} \times \frac{2}{8}\right) + \left(\frac{4}{10} \times \frac{6}{9} \times \frac{3}{8}\right) \\ &= \frac{24}{720} + \frac{72}{720} \\ &= \frac{96}{720} \end{aligned}$$

$$\text{Therefore, } P(3^{\text{rd}} \text{ disc blue} \mid 1^{\text{st}} \text{ disc blue}) = \frac{96/720}{4/10} = \frac{1}{3}$$

(ii) X does not have a geometric distribution. This is because the probabilities are not constant (or because the events are not independent).

5 (i) There were more births in 1991.

In 1991, there were about 695000 births. In 2001, there were about 590000 births. So there were about 105000 extra births in 1991.

(ii) a) The completed table is

Year	Median age (years)	IQR (years)	Proportion of mothers giving birth below 25	Proportion of mothers giving birth aged 35 or above
1991	27.5	7.3	33%	9%
2001	29.5	33.5 - 25.5 = 8	150/595 × 100 = 25%	18%

b) Overall the median ages show that women in 2001 tended to be slightly older than those in 1991. This is also shown by the proportion of mothers that are aged 35 or over (up from 9% to 18%) and the proportion of mothers below 25 (down from 33% to 25%).

6 (i) a) The formula for the PMCC is:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

Here

$$S_{xy} = 767 - \frac{60 \times 72}{8} = 227$$

$$S_{xx} = 1148 - \frac{60^2}{8} = 698$$

$$S_{yy} = 810 - \frac{72^2}{8} = 162$$

So,

$$r = \frac{227}{\sqrt{698 \times 162}} = 0.675 \text{ (3 SF)}$$

b) Spearman's rank correlation coefficient = 1 (as points are increasing)

(ii) a)  $r$  would increase as points would lie closer to a straight line.

b) Spearman's rank will still be 1 as the points will still be increasing.

(iii) When  $x = 29$ ,  $y$  will be about 13.9

(iv) The estimate found in (iii) will be the most accurate. The regression line is not appropriate as there is not a linear relationship between the variables.

7 (i) Assumption 1: The probability that a packet contains a voucher is constant.  
Assumption 2: The packets of soup are independent, i.e. if one packet is found to contain a voucher it does not alter the chances of another packet containing a voucher.

Note: These assumptions must be stated in the context of the question.

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

Using tables,  $P(X \leq 6) = 0.9857$ .

(iii)  $P(X \geq 7) = 1 - P(X \leq 6) = 1 - 0.9857 = 0.0143$

(iv) If Kim is to claim her free gift in the 12<sup>th</sup> week but not before then she must get 6 vouchers in the first 11 weeks and then get a 7<sup>th</sup> voucher in the 12<sup>th</sup> week.

The probability of this is:  $({}^{11}C_6 \times 0.25^6 \times 0.75^5) \times 0.25 = 0.00669$

8 (i) Let  $p = \text{Prob}(\text{heads})$ .

We are given that the probability of 2 heads is 0.04, i.e.  $p^2 = 0.04$ .

Consequently  $p = 0.2$ .

So the probability of tails is 0.8. So the probability of tails both times =  $0.8^2 = 0.64$ .

(ii) If  $p = \text{Prob}(\text{heads})$ , then the probability of getting exactly one head in two throws is):

$$P(H, T) + P(T, H) = p(1 - p) + (1 - p)p = 2p(1 - p).$$

We can therefore form an equation:

$$2p(1 - p) = 0.42$$

or  $p(1 - p) = 0.21.$

If we expand out the brackets and rearrange, we end up with the quadratic equation:

$$p^2 - p + 0.21 = 0.$$

This can be solved using the quadratic formula. The solutions are  $p = 0.7$  or  $p = 0.3.$

9 (i) a) For a Geometric distribution,  $E(X) = 1/p.$

So here,  $E(X) = 5.$

b)  $P(X = 4) = \left(\frac{4}{5}\right)^3 \left(\frac{1}{5}\right) = 0.102$  (3 SF)

c)  $P(X > 4) = q^4 = \left(\frac{4}{5}\right)^4 = 0.410$  (3 SF).

Note: in part (c), we are using the useful result:  $P(X > x) = q^x.$

(ii) a)  $P(Y = 1) = p$   
 $P(Y = 3) = q^2 p$   
 $P(Y = 5) = q^4 p$   
 $P(Y = 7) = q^6 p$   
 Etc

Therefore,  $P(Y \text{ is odd}) = p + q^2 p + q^4 p + q^6 p + \dots$

b) The probabilities form a geometric progression with  $a = p$  and  $r = q^2.$

So  $P(Y \text{ is odd}) = \frac{a}{1 - r} = \frac{p}{1 - q^2} = \frac{p}{(1 - q)(1 + q)}$

However,  $p = 1 - q,$  so  $P(Y \text{ is odd}) = \frac{1 - q}{(1 - q)(1 + q)} = \frac{1}{1 + q}$