STEP Support Programme

Statistics STEP Questions: Solutions

2010 S1 Q12

1 Preparation

(i) **(a)** The sum of the probabilities is 1, so we have $k + 2k + 3k + 4k = 1 \implies k = \frac{1}{10}$.
(b) $P(X \geqslant 3) = P(X = 3) + P(X = 4) = \frac{7}{10}$.
(c) $E(X) = \frac{1}{10}(1 \times 1 + 2 \times 2 + 3 \times 3 + 4 \times 4) = 3$.

(ii) **(a)** This is the probability that I will roll "not 6, not 6, 6" which is equal to $\left(\frac{5}{6}\right)^2 \times \left(\frac{1}{6}\right)$.
(b) If I need more than 5 rolls, the first 5 must all be "not 6", i.e. probability is $\left(\frac{5}{6}\right)^5$.

(iii) This is a Geometric Progression with $a = 2$ and $r = 0.1$. The sum is $\frac{2}{1-0.1} = \frac{20}{9}$.

(iv) You could either differentiate, or complete the square to get $y = 4 - (x-2)^2$. The maximum value is $y = 4$ (which occurs when $x = 2$).

2 The Probability question

It is easy to skip the requests in the second sentence if you do not read the question carefully!

$E(X) = 1 \times P(X = 1) + 2 \times P(X = 2) + 3 \times P(X = 3) + \ldots$. To show the second part you can write $E(X)$ as:

$$\begin{array}{rccccccccc}
E(X) & = & P(X=1) & + & 2\times P(X=2) & + & 3\times P(X=3) & + & 4\times P(X=4) & + \cdots \\
 & = & P(X=1) & + & P(X=2) & + & P(X=3) & + & P(X=4) & + \cdots \\
 & & & + & P(X=2) & + & P(X=3) & + & P(X=4) & + \cdots \\
 & & & & & + & P(X=3) & + & P(X=4) & + \cdots \\
 & & & & & & & + & P(X=4) & + \cdots
\end{array}$$

Looking along the rows, we have $E(X) = P(X \geqslant 1) + P(X \geqslant 2) + P(X \geqslant 3) + \ldots$, i.e. $E(X) = \displaystyle\sum_{n=1}^{\infty} P(X \geqslant n)$.

$P(X \geqslant 4)$ means that I have to open at least 4 boxes to get one of each type of penguin. This means that the first 3 boxes are either all mummy penguins (probability q^3) or are all daddy penguins (probability p^3). Hence $P(X \geqslant 4) = p^3 + q^3$.

Similarly to above, we have $P(X \geqslant n) = p^{n-1} + q^{n-1}$ for $n \geqslant 2$. Note that $P(X \geqslant 1) = 1$ (as we **have** to open at least 1 box) which is not equal to $p^{1-1} + q^{1-1} = 2$. However $P(X \geqslant 2) = 1 = p^{2-1} + q^{2-1}$, which is why we can say that the above result holds for $n \geqslant 2$.

Using $\mathrm{E}(X) = \displaystyle\sum_{n=1}^{\infty} \mathrm{P}(X \geqslant n)$ gives us:

$$\begin{aligned}
\mathrm{E}(X) &= 1 + \sum_{n=2}^{\infty} \left(p^{n-1} + q^{n-1}\right) \\
&= 1 + p\left(1 + p + p^2 + p^3 + \ldots\right) + q\left(1 + q + q^2 + q^3 + \ldots\right) \\
&= 1 + \frac{p}{1-p} + \frac{q}{1-q} \\
&= 1 + \frac{p}{q} + \frac{q}{p} \\
&= 1 + \frac{p^2 + q^2}{pq} \\
&= 1 + \frac{(p+q)^2 - 2pq}{pq} \\
&= 1 + \frac{(p+q)^2}{pq} - 2 \\
&= \frac{1}{pq} - 1 \qquad (\text{using } p + q = 1)
\end{aligned}$$

Using $q = 1 - p$ we have $\mathrm{E}(X) = \frac{1}{p(1-p)} - 1$. $\mathrm{E}(X)$ will be minimised when $p(1-p)$ is maximised, and as $p(1-p) = \frac{1}{4} - \left(p - \frac{1}{2}\right)^2$, the maximum value of $p(1-p)$ is $\frac{1}{4}$ and so the minimum value of $\mathrm{E}(X)$ is $4 - 1 = 3$. Hence $\mathrm{E}(X) \geqslant 3$.

2009 S1 Q13

3 Preparation

(i) $6!(= 720)$

(ii) $10!$

(iii) **(a)** $4!$

(b) $7!$

(c) $7! \times 4!$

(d) The probability is:

$$\begin{aligned}\frac{7! \times 4!}{10!} &= \frac{4!}{10 \times 9 \times 8} \\ &= \frac{\cancel{4} \times 3 \times \cancel{2} \times 1}{10 \times 9 \times \cancel{8}} \\ &= \frac{3}{90} = \frac{1}{30}\end{aligned}$$

(iv) The probability is given by:

$$\begin{aligned}\frac{5! \times 6!}{10!} &= \frac{5!}{10 \times 9 \times 8 \times 7} \\ &= \frac{\cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times 1}{\cancel{10}_2 \times \cancel{9}_3 \times \cancel{8} \times 7} \\ &= \frac{1}{42}\end{aligned}$$

Note that this is smaller than the answer to **(iii)(d)** — you are less likely to get all 6 women standing together than the 4 men standing together.

(v) Here there are $6!$ ways to arrange the women inside their "rope" and $4!$ ways to arrange the men inside their "rope". There are then 2 ways to arrange the two "ropes": [men], [women] or [women], [men]. The probability is given by:

$$\begin{aligned}\frac{2 \times 6! \times 4!}{10!} &= \frac{2 \times 4!}{10 \times 9 \times 8 \times 7} \\ &= \frac{2 \times \cancel{4} \times 3 \times \cancel{2} \times 1}{10 \times 9 \times \cancel{8} \times 7} \\ &= \frac{\cancel{2} \times \cancel{3} \times 1}{\cancel{10}_5 \times \cancel{9}_3 \times 7} \\ &= \frac{1}{105}\end{aligned}$$

Note that this is smaller than the two previous answers. Using a bit of "common sense" is often a good way to check your answers.

(vi) **(a)** 6! ways

(b) There are $7 \times 6 \times 5 \times 4$ ways to arrange the 4 men into the 7 gaps.

(c) The probability is:

$$\begin{aligned}\frac{6! \times 7 \times 6 \times 5 \times 4}{10!} &= \frac{\not{7} \times 6 \times 5 \times 4}{10 \times 9 \times 8 \times \not{7}} \\ &= \frac{\not{6}_2 \times \not{5} \times \not{4}}{\not{10}_2 \times \not{9}_3 \times \not{8}_2} \\ &= \frac{1}{6}\end{aligned}$$

(vii) **(a)** There are 6 options for choosing a women at one end of the line and then 5 options for choosing a woman at the other end of the line, so there are 6×5 ways to choose two women for either end of the line.

(b) There are 8! ways to arrange the rest of the people.

(c) The probability is given by:

$$\begin{aligned}\frac{6 \times 5 \times 8!}{10!} &= \frac{6 \times 5}{10 \times 9} \\ &= \frac{1}{3}\end{aligned}$$

4 The Probability question

(i) $K = 3$ means that all three girls are together. There are 3! ways to arrange the girls in a "rope" and then $(n+1)!$ ways to arrange the "rope" and the n boys. There are a total of $(n+3)!$ ways of arranging the $n+3$ children.

$$\begin{aligned}\mathrm{P}(K = 3) &= \frac{3! \times (n+1)!}{(n+3)!} \\ &= \frac{6}{(n+3)(n+2)}\end{aligned}$$

(ii) $K = 1$ means that the girls are all separated. There are $n!$ ways to arrange the n boys. Then there are $n+1$ "slots" for the first girl, n slots for the second girl and $n-1$ slots for the third girl. The probability is:

$$\begin{aligned}\mathrm{P}(K = 1) &= \frac{n! \times (n+1) \times (n) \times (n-1)}{(n+3)!} \\ &= \frac{(n+1)(n)(n-1)}{(n+3)(n+2)(n+1)} \\ &= \frac{n(n-1)}{(n+3)(n+2)}\end{aligned}$$

(iii) The only possible values of K are $1, 2$ and 3. To find $\mathrm{E}(K)$ we need to find $\mathrm{P}(K=2)$. We can do this by using the fact that the probabilities sum to 1.

$$\begin{aligned}\mathrm{P}(K=2) &= 1 - \mathrm{P}(K=1) - \mathrm{P}(K=3)\\ &= 1 - \frac{n(n-1)}{(n+3)(n+2)} - \frac{6}{(n+3)(n+2)}\\ &= \frac{(n+3)(n+2) - n(n-1) - 6}{(n+3)(n+2)}\\ &= \frac{n^2+5n+6-n^2+n-6}{(n+3)(n+2)}\\ &= \frac{6n}{(n+3)(n+2)}\end{aligned}$$

The expectation is given by:

$$\begin{aligned}\mathrm{E}(K) &= 1 \times \frac{n(n-1)}{(n+3)(n+2)}\\ &\qquad + 2 \times \frac{6n}{(n+3)(n+2)} + 3 \times \frac{6}{(n+3)(n+2)}\\ &= \frac{n(n-1)+12n+18}{(n+3)(n+2)}\\ &= \frac{n^2+11n+18}{(n+3)(n+2)}\\ &= \frac{(n+9)(n+2)}{(n+3)(n+2)}\\ &= \frac{n+9}{n+3}\end{aligned}$$

This answer will be more than 1 and less than 3, which is what we expect!

5 **1995 S1 Q12** First thing to note is that n is the **total** number of students here. This means that the number of non-hockey players is $n-r$.

(i) There are r ways of choosing a hockey player for one end of the row and then $r-1$ ways of choosing a hockey player for the other end of the row. There are then $(n-2)!$ ways of arranging the rest of the students. The probability is:

$$\frac{r(r-1)\times(n-2)!}{n!} = \frac{r(r-1)}{n(n-1)}$$

(ii) There are $r!$ ways to arrange the hockey players inside a "rope" and then $(n-r+1)!$ ways to arrange the $n-r$ non-hockey players and the "rope". The probability is:

$$\frac{r!\times(n-r+1)!}{n!}$$

No need to "simplify" this, but if you used a different approach you may have the probability as $\dfrac{n-r+1}{\binom{n}{r}}$. These two answers are equivalent, either is fine!

(iii) If we arrange the $n-r$ non-hockey players first then there are $n-r+1$ gaps. We will not be able to separate the hockey players if $r > n-r+1$, i.e. if $r > \frac{1}{2}(n+1)$. It might be a good idea to try a few values of r and n to convince yourself that this is correct.

Assuming that $r \leqslant \frac{1}{2}(n+1)$, then there are $(n-r)!$ ways of arranging the non-hockey players. There are $n-r+1$ slots for the first hockey player, $n-r$ slots for the second hockey player and so on until we get to $(n-r+1)-(r-1) = n-2r+2$ slots for the r^{th} hockey player.

The probability is then:

$$
\begin{aligned}
&\frac{(n-r+1)(n-r)\dots(n-2r+2)\times(n-r)!}{n!}\\
=&\frac{(n-r+1)(n-r)\dots(n-2r+2)\times(n-2r+1)(n-2r)\dots1}{(n-2r+1)(n-2r)\dots1}\times\frac{(n-r)!}{n!}\\
=&\frac{(n-r+1)!}{(n-2r+1)!}\frac{(n-r)!}{n!}
\end{aligned}
$$

The combined answer for the probability is

$$
\text{probability} = \begin{cases} \dfrac{(n-r+1)!}{(n-2r+1)!}\dfrac{(n-r)!}{n!} & \text{for} \quad r \leqslant \frac{1}{2}(n+1) \\ 0 & \text{for} \quad r > \frac{1}{2}(n+1) \end{cases}
$$

If you like, you can write the non-zero probability as:

$$
\begin{aligned}
\frac{(n-r+1)!}{(n-2r+1)!}\frac{(n-r)!}{n!} &= \frac{(n-r+1)!}{(n-2r+1)!r!}\times\frac{(n-r)!r!}{n!}\\
&= \frac{\binom{n-r+1}{r}}{\binom{n}{r}}
\end{aligned}
$$

Here $\binom{n}{r}$ is the number of ways of "choosing" r students out of n to be hockey players and $\binom{n-r+1}{r}$ is the number of ways of choosing r slots out of the $n-r+1$ in total to be filled with hockey players.

1999 S2 Q12

6 Preparation

(i) When $x = 0$, $y = 0$ and there is a vertical asymptote at $x = \frac{5}{2}$. The derivative is $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{15}{(5-2x)^2}$, so the gradient of the graph is always positive.

To find where the horizontal asymptote is, you can write y as $y = \dfrac{3}{\frac{5}{x} - 2}$, and so as $x \to \pm\infty$ we have $y \to -\frac{3}{2}$. Another way of doing this[1] is to divide $3x$ by $5 - 2x$ to get $y = -\frac{3}{2} + \dfrac{15}{2(5-2x)}$. This is not the way I would choose for this particular question!

Once you have sketched your graph, try plotting it using Desmos or something similar.

Using your sketch, you can see that in the interval $-1 \leqslant x \leqslant 2$, the maximum occurs when $x = 2$ (and is equal to 6) and the minimum occurs when $x = -1$ (and is equal to $-\frac{3}{7}$).

(ii) This graph crosses the y axis at $\left(0, \frac{1}{5}\right)$ and crosses the x axis at $(1, 0)$. It has a vertical asymptote at $x = \frac{5}{2}$ and a horizontal asymptote at $y = \frac{1}{2}$. The derivative is $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{-3}{(5-2x)^2}$, so the gradient of the graph is always negative.

In the range $-1 \leqslant x \leqslant 2$, the maximum occurs when $x = -1$ (and is equal to $\frac{2}{7}$) and the minimum occurs when $x = 2$ (and is equal to -1).

(iii) **(a)** $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) = 0.3 + 0.5 - 0.15 = 0.65$.

$\mathrm{P}(A|B) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \dfrac{0.15}{0.5} = 0.3$. Note that $\mathrm{P}(A|B) = \mathrm{P}(A)$ — this is what A being independent of B means, whether B happens or not has no affect on the probability of A happening!

Similarly $\mathrm{P}(B|A) = \mathrm{P}(B) = 0.5$.

(b) Using $\mathrm{P}(A \cap B) = \mathrm{P}(A) \times \mathrm{P}(B|A)$[2] gives $\mathrm{P}(A \cap B) = 0.6 \times 0.4 = 0.24$.
We then have $\mathrm{P}(A|B) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \dfrac{0.24}{0.5} = 0.48$.
$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) = 0.6 + 0.5 - 0.24 = 0.86$.

(c) The probability that a Martian who has tested positive to tentacle rot actually has tentacle rot is $\dfrac{8}{107} \approx 7.5\%$.

[1] This is the approach you will probably use if there is an oblique (slanted) asymptote.

[2] This relationship says "The probability that A and B both happen is equal to the probability that A happens and then that B happens given that A has happened.

7 The Probability question

We have:

$$\begin{aligned}
\mathrm{P}(A|S) &= \frac{\mathrm{P}(A\cap S)}{\mathrm{P}(S)}\\
&= \frac{2p\times 0.7}{2p\times 0.7 + p\times 0.8 + (1-3p)\times 0.9}\\
&= \frac{1.4p}{0.9-0.5p}\\
&= \frac{14p}{9-5p}
\end{aligned}$$

and:

$$\begin{aligned}
\mathrm{P}(C|S) &= \frac{\mathrm{P}(C\cap S)}{\mathrm{P}(S)}\\
&= \frac{(1-3p)\times 0.9}{2p\times 0.7 + p\times 0.8 + (1-3p)\times 0.9}\\
&= \frac{0.9-2.7p}{0.9-0.5p}\\
&= \frac{9-27p}{9-5p}
\end{aligned}$$

The first inspector want to maximise $\frac{14p}{9-5p}$, subject to $0 \leqslant p \leqslant \frac{1}{3}$. In this range $\frac{14p}{9-5p}$ is increasing (the derivative with respect to p is $\frac{14\times 9}{(9-5p)^2}$) as p increases and we do not cross a vertical asymptote in this range, so the value of p that maximises $\frac{14p}{9-5p}$ is $p = \frac{1}{3}$.

The second inspector wants to maximise $\frac{9-27p}{9-5p}$, subject to $0 \leqslant p \leqslant \frac{1}{3}$. Here the function is decreasing (the derivative is $\frac{-22\times 9}{(9-5p)^2}$) and we do not cross a vertical asymptote, so the value of p that maximises $\frac{9-27p}{9-5p}$ is $p = 0$.

In the first case, the chip has come from A, and the value of p maximises the probability that a chip is produced by A. In the second case the chip comes from C and the value of p maximises the probability that the chip comes from C. This is what you would expect if you only look at one chip — a bigger sample would be necessary to get a more meaningful estimate!

2006 S2 Q13

8 Preparation

(i) **(a)** The probability that the child who is first in the queue is the 2 year old is $\frac{1}{5}$.

(b) Given that the first child is the youngest, the probability that the second child is the 10 year old is $\frac{1}{4}$.

(ii) **(a)** The number of numbers is $4 \times 3 \times 2 \times 1 = 24$.

(b) The last two digits have to be divisible by 4, so they can be 12, 24 or 32. There are six possible numbers (for each of these pairs of last digits the other two digits can go either way around).

(c) Anything starting with a 3 or 4 will be bigger than 2413, and there are 6 of each of these. The only other number bigger than 2413 is 2431, so the probability of creating a number bigger than 2413 is $\frac{13}{24}$.

(d) If we know that the first digit is 1, then we know that the number is one of six possibilities: $1234, 1243, 1324, 1342, 1423, 1432$. Of these six, two of them are divisible by 4, so the probability is $\frac{1}{3}$.

9 The Probability question

(i) $P_4(1)$ is the probability that I choose the biggest ice-cream if there are 4 in total. There are various ways of doing this, but given that we need to find $P_4(2)$, $P_4(3)$ and $P_4(4)$ the most efficient way is probably to list all the permutations of the four sizes that could occur and then identify which ice-cream I would pick in each case. Below are all the combinations of the 4 ice-creams which are possible (1 being the biggest), with the one I pick circled.

1	2	3	④
1	2	4	③
1	3	2	④
1	3	4	②
1	4	2	③
1	4	3	②
2	①	3	4
2	①	4	3
2	3	①	4
2	3	4	①
2	4	①	3
2	4	3	①
3	①	2	4
3	①	4	2
3	②	1	4
3	②	4	1
3	4	①	2
3	4	②	1
4	①	2	3
4	①	3	2
4	②	1	3
4	②	3	1
4	③	1	2
4	③	2	1

From this we have $P_4(1) = \frac{11}{24}$, $P_4(2) = \frac{7}{24}$, $P_4(3) = \frac{4}{24}$ and $P_4(4) = \frac{2}{24}$.

(ii) $P_n(1)$ is the probability that I pick the biggest ice-cream if there are n in total.

- If the biggest ice-cream is first, then the probability I pick it is 0.
- If the biggest ice-cream is second, then the probability I pick it is 1.
- If the biggest ice-cream is third, then the probability I pick it is $\frac{1}{2}$ as it depends on which of the first two is the bigger. If the second ice-cream is bigger than the first ice cream then I will pick the second ice-cream, whereas if the second ice-cream is smaller than the first then I will pick the third one (which is the biggest overall).
- If the biggest ice-cream is fourth, then the probability I pick it is $\frac{1}{3}$, as I will pick it as long as out of the three previous ice-creams the biggest one is first.

- If the biggest ice cream is last, then the probability I pick it is $\frac{1}{n-1}$, as I will pick it as long as the biggest ice-cream out of the $n-1$ previous ones is first.

The biggest ice-cream has a probability of $\frac{1}{n}$ of being in each position, so we have:

$$\mathrm{P}_n(1) = \frac{1}{n}\left(0 + 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1}\right)$$

It is a good idea to check your answer! Taking $n = 4$ we have:

$$\begin{aligned}\mathrm{P}_4(1) &= \frac{1}{4}\left(0 + 1 + \frac{1}{2} + \frac{1}{3}\right)\\ &= \frac{1}{4}\left(\frac{11}{6}\right)\\ &= \frac{11}{24}\end{aligned}$$

which is the same as our answer in part **(i)**.

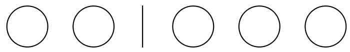
Another way of tackling this part is to consider the cases when the biggest ice-cream is first, the second biggest ice-cream is first, the third biggest ice-cream is first etc.

2008 S2 Q13

10 Preparation

(i) There are a couple of ways of thinking about this. Firstly we can consider each bun separately. The probability that the first one has the cherry is $\frac{1}{5}$ and the probability that the second has the cherry is $\frac{4}{5} \times \frac{1}{4} = \frac{1}{5}$ (in this case the first one has to not be a cherry bun). The probability that I have the cherry bun if therefore $\frac{2}{5}$.

Otherwise, you can think of partitioning the 5 buns into a group of 2 and a group of 3, something like:



We then apply the cherry at random and the probability that it lands on one of the two that we have picked is $\frac{2}{5}$.

(ii) Using the partition idea (like above) we have the probability $\frac{r}{n}$.

(iii) The total number of ways I can pick the team is ${}^{10}C_5$, and the number of ways of choosing a team that includes the 2 males is ${}^{8}C_3$ (as we have to choose 3 more from the 8 female swimmers). The probability is then:

$$\begin{aligned}\frac{{}^{8}C_3}{{}^{10}C_5} &= \frac{\frac{8!}{3!5!}}{\frac{10!}{5!5!}} \\ &= \frac{8!}{3!5!} \times \frac{5!5!}{10!} \\ &= \frac{8! \times 5!}{3! \times 10!} \\ &= \frac{5 \times 4}{10 \times 9} \\ &= \frac{2}{9}\end{aligned}$$

(iv) The number of ways of picking r buns from the n in total is ${}^{n}C_r$. If one of the buns you pick is the one with a cherry then there are ${}^{n-1}C_{r-1}$ ways of picking the rest. The probability of picking the one with the cherry on top is therefore:

$$\begin{aligned}\frac{{}^{n-1}C_{r-1}}{{}^{n}C_r} &= \frac{\frac{(n-1)!}{(r-1)![(n-1)-(r-1)]!}}{\frac{n!}{r!(n-r)!}} \\ &= \frac{(n-1)!}{(r-1)!(n-r)!} \times \frac{r!(n-r)!}{n!} \\ &= \frac{(n-1)! \times r!}{(r-1)! \times n!} \\ &= \frac{r}{n}\end{aligned}$$

(v) Here the probability is:

$$\begin{aligned}
\frac{{}^{n-2}C_r}{{}^{n}C_r} &= \frac{\frac{(n-2)!}{r![(n-2)-r]!}}{\frac{n!}{r!(n-r)!}} \\
&= \frac{(n-2)!}{r!(n-r-2)!} \times \frac{r!(n-r)!}{n!} \\
&= \frac{(n-r)(n-r-1)}{n(n-1)}
\end{aligned}$$

(vi) Starting on the RHS we have:

$$\begin{aligned}
{}^{n-1}C_r +^{n-1} C_{r-1} &= \frac{(n-1)!}{r!((n-1)-r)!} + \frac{(n-1)!}{(r-1)!((n-1)-(r-1))!} \\
&= \frac{(n-1)!}{r!(n-r-1)!} + \frac{(n-1)!}{(r-1)!(n-r)!} \\
&= \frac{(n-1)!}{(r-1)!(n-r-1)!}\left[\frac{1}{r} + \frac{1}{n-r}\right] \\
&= \frac{(n-1)!}{(r-1)!(n-r-1)!}\left[\frac{n-r}{r(n-r)} + \frac{r}{r(n-r)}\right] \\
&= \frac{(n-1)!}{(r-1)!(n-r-1)!}\left[\frac{n}{r(n-r)}\right] \\
&= \frac{n \times (n-1)!}{[r \times (r-1)!] \times [(n-r) \times (n-r-1)!]} \\
&= \frac{n!}{r!(n-r)!} \\
&=^n C_r
\end{aligned}$$

Alternatively, ${}^nC_r$ is the number of ways of picking r objects from n. Consider what happens with the first object, there are two cases:

- we pick the first object, then need to pick $r-1$ more from the $n-1$ left, and there are ${}^{n-1}C_{r-1}$ ways of doing this
- we don't pick the first object, so need to pick r from the remaining $n-1$ objects, and there are ${}^{n-1}C_r$ ways of doing this

Therefore we have ${}^nC_r =^{n-1} C_{r-1} +^{n-1} C_r$.

(vii) $a_i/a_{i-1} > 1$ means that $a_i > a_{i-1}$, so for $i \leqslant k$ the sequence is increasing. After this, the sequence starts decreasing so the largest term is a_k.

11 The Probability question

(i) Either the black counter is moved from P to Q and then back from Q to P (where it is a case of picking the black one from $n+k$ in total), or it is not picked from P in the first place. The probability is:

$$\begin{aligned}\frac{k}{n}\times\frac{k}{n+k}+\frac{n-k}{n} &= \frac{k^2+(n-k)(n+k)}{n(n+k)}\\ &= \frac{k^2+n^2-k^2}{n(n+k)}\\ &= \frac{n}{n+k}\end{aligned}$$

This probability is maximised when $k=0$ (which makes sense — if we take no counters out of P the black counter will certainly stay in P).

(ii) Here the options are:

- the black counter is picked from P then neither black counter is picked from Q
- the black counter is picked from P then both black counters are picked from Q
- the black counter is not picked from P then the black counter is picked from Q

The probability of them ending in the same bag is:

$$\begin{aligned}&\frac{k}{n}\times\frac{{}^{n+k-2}C_k}{{}^{n+k}C_k}+\frac{k}{n}\times\frac{{}^{n+k-2}C_{k-2}}{{}^{n+k}C_k}+\frac{n-k}{n}\times\frac{k}{n+k}\\
=&\frac{k}{n}\times\frac{(n+k-2)!}{k!(n-2)!}\times\frac{k!n!}{(n+k)!}+\frac{k}{n}\times\frac{(n+k-2)!}{(k-2)!n!}\times\frac{k!n!}{(n+k)!}+\frac{n-k}{n}\times\frac{k}{n+k}\\
=&\frac{k}{n}\times\frac{n(n-1)}{(n+k)(n+k-1)}+\frac{k}{n}\times\frac{k(k-1)}{(n+k)(n+k-1)}+\frac{k(n-k)}{n(n+k)}\\
=&\frac{k}{n(n+k)}\left[\frac{n(n-1)}{(n+k-1)}+\frac{k(k-1)}{(n+k-1)}+(n-k)\right]\\
=&\frac{k}{n(n+k)}\left[\frac{n^2-n+k^2-k+(n-k)(n+k-1)}{(n+k-1)}\right]\\
=&\frac{k}{n(n+k)}\left[\frac{2n^2-2n}{(n+k-1)}\right]\\
=&\frac{2k(n-1)}{(n+k)(n+k-1)}\end{aligned}$$

There are a couple of ways of maximising this. You can differentiate with respect to k and so maximise the probability. Or you can look at the ratio of probabilities as k increases.

Let the probability equal p_k. We have:

$$\begin{aligned}\frac{p_k}{p_{k-1}} &= \frac{2k(n-1)}{(n+k)(n+k-1)}\times\frac{(n+k-1)(n+k-2)}{2(k-1)(n-1)}\\ &= \frac{k(n+k-2)}{(n+k)(k-1)}\end{aligned}$$

When this ratio is greater than 1 we have:

$$\begin{aligned}
\frac{p_k}{p_{k-1}} &> 1 \\
\frac{k(n+k-2)}{(n+k)(k-1)} &> 1 \\
k(n+k-2) &> (n+k)(k-1) \\
nk + k^2 - 2k &> nk + k^2 - n - k \\
n &> k
\end{aligned}$$

Hence we have $p_k > p_{k-1}$ when $k < n$. So up to, and including, $k = n - 1$ the probabilities are increasing. We don't know what happens between $k = n - 1$ and $k = n$, but k cannot be greater than n, and it's not too much of an effort to work out the probabilities for both $k = n - 1$ and $k = n$.

We have:

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

and:

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

which are the same, and hence the values of k that maximise the probability are $k = n - 1$ and $k = n$.

If you go down the differentiation route, you will find that the probability is maximised when $k^2 = n(n-1)$, i.e. $k = \sqrt{n(n-1)}$. This is not an integer, so you would need to find the probabilities at the integer values of k either side of this, which are $k = n - 1$ and $k = n$. Like above, you should find that these both maximise the probability.

2015 S2 Q12

12 Preparation

(i) **(a)** $\frac{1}{4}$

(b) Here the first 4 tosses are TTHH, so probability is $\frac{1}{16}$

(c) We can choose P by tossing (HH), (TT)(HH), (TT)(TT)(HH), etc. The probability that we choose P is:

$$\frac{1}{4}\left(1+\frac{1}{4}+\left(\frac{1}{4}\right)^2+\dots\right)=\frac{\frac{1}{4}}{1-\frac{1}{4}}=\frac{1}{3}$$

A similar result hold for Q and R, so the three results are equally likely. This is a way you can use a coin to choose fairly between 3 objects.

(ii) If the first toss is a Head, then on the next toss we will either get another Head an choose P or a Tail and choose Q.

If we get a Tail first then either we get a Head and choose R or we get another tail and keep going. However, if we get a Tail first the only choose we can make is R. We have:

$$\mathrm{P}(P)=\frac{1}{4},\ \mathrm{P}(Q)=\frac{1}{4},\ \mathrm{P}(R)=\frac{1}{2}$$

This method does not choose fairly between the three objects.

13 The Probability question

(i) A can only win, and will win, if the first two tosses are Heads (if one of the first two tosses is a Tail then THH must occur before HHT, and if the first two are heads then HHT must occur before THH). Hence the probability that A wins is equal to the probability that the first two tosses are heads, i.e. it is $\frac{1}{4}$.

(ii) If all four players play, then A will win if (and only if) the first two tosses are heads, and C will win if (and only if) the first two tosses are tails.

Player B and Player D will have the same probability of winning as their winning sequences are symmetrical (swap the H's for T's and vice-verse and they are identical).

The only drawn sequences are the infinite sequences "HTHTHTH..." and "THTHTHTH..." both of which have probability 0 (in the limit as the number of tosses tends to infinity). Hence $\mathrm{P}(B) = \mathrm{P}(D) = \frac{1}{4}$.

(iii) If the first two tosses are Tails then C will win (C needs one Head to win whereas B would need two). So $\mathrm{P}(C|\mathrm{TT}) = 1$.

We are given:

- $\mathrm{P}(C|\mathrm{HT}) = p$
- $\mathrm{P}(C|\mathrm{TH}) = q$
- $\mathrm{P}(C|\mathrm{TT}) = r$

Consider starting with "HT". Then either the next coin toss is "T", and then C will win (as the last two tosses in the sequence are "TT") or the next toss is a "H" and C will win with probability q (as the last two tosses in the sequence are "TH"). Hence $p = \frac{1}{2} \times 1 + \frac{1}{2} \times q$, i.e. $p = \frac{1}{2} + \frac{1}{2}q$.

Similarly, considering the starting sequence "TH" we have $q = \frac{1}{2} \times 0 + \frac{1}{2} \times p = \frac{1}{2}p$ (note that if the first two tosses are "TH" then if we toss another "H" we have "THH" and B wins).

With the starting sequence "HH" we have $r = \frac{1}{2}r + \frac{1}{2}p$.

We now have three equations in three unknowns:

- $p = \frac{1}{2} + \frac{1}{2}q$
- $q = \frac{1}{2}p$
- $r = \frac{1}{2}r + \frac{1}{2}p$

This gives $p = \frac{2}{3}$, $q = \frac{1}{3}$ and $r = \frac{2}{3}$. This gives the probability that C wins as:

$$\frac{1}{4}\left(1 + \frac{2}{3} + \frac{1}{3} + \frac{2}{3}\right) = \frac{2}{3}$$