M1S : EXERCISE SHEET 5 : SOLUTIONS

ASSESSED

(a) Using a binary sequence argument; the set of all possible sequences of selections is identical to the set of all binary sequences of length $r+b$ that contain r 1s. If the event E is the event that the first red ball is selected on the $(k+1)$st draw, then E essentially corresponds to the subset of binary sequences in which the first k elements are all 0s, the $(k+1)$st element is 1, and then the remaining $r+b-(k+1)$ elements comprise $r-1$ 1s and $b-k$ 0s in some order. Hence

$$\mathrm{P}(\ E\ ) = \frac{n_E}{n} = \frac{\begin{pmatrix} r+b-k-1 \\ r-1 \end{pmatrix}}{\begin{pmatrix} r+b \\ b \end{pmatrix}}$$

[*5 MARKS*]

For the probability that the last ball drawn is red, consider the corresponding binary sequences. The number of binary sequences of length $r+b$ containing r 1s that end in 1 is

$$\begin{pmatrix} r+b-1 \\ r-1 \end{pmatrix} \Longrightarrow \text{ probability required is } \frac{\begin{pmatrix} r+b-1 \\ r-1 \end{pmatrix}}{\begin{pmatrix} r+b \\ b \end{pmatrix}} = \frac{(r+b-1)!}{(r-1)!b!}\frac{r!b!}{(r+b)!} = \frac{r}{r+b}$$

[*3 MARKS*]

(b) The denominator in the probability calculation is equal to the number of ways of allocating the $r=8$ indistinguishable balls to the $n=6$ boxes; the formula for this given in lectures is

$$\begin{pmatrix} n+r-1 \\ r \end{pmatrix}$$

Now, for the event of interest, consider cases when boxes one and two contain $k=0,1,2,3,4$ balls in total, so that the other four boxes contain $8-k$ balls. The number of ways of allocating k balls to two boxes is

$$\begin{pmatrix} 2+k-1 \\ k \end{pmatrix}$$

and the number of ways of allocating $8-k$ balls to four boxes is

$$\begin{pmatrix} 4+(8-k)-1 \\ 8-k \end{pmatrix}$$

and hence the number total number of ways of carrying out such an allocation is

$$\begin{pmatrix} 2+k-1 \\ k \end{pmatrix}\begin{pmatrix} 4+(8-k)-1 \\ 8-k \end{pmatrix}.$$

Thus the probability required is

$$\frac{\displaystyle\sum_{k=0}^{4}\begin{pmatrix} 2+k-1 \\ k \end{pmatrix}\begin{pmatrix} 4+(8-k)-1 \\ 8-k \end{pmatrix}}{\begin{pmatrix} n+r-1 \\ r \end{pmatrix}} = \frac{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 11 \\ 8 \end{pmatrix}+\begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 10 \\ 7 \end{pmatrix}+\begin{pmatrix} 3 \\ 2 \end{pmatrix}\begin{pmatrix} 9 \\ 6 \end{pmatrix}+\begin{pmatrix} 4 \\ 3 \end{pmatrix}\begin{pmatrix} 8 \\ 5 \end{pmatrix}+\begin{pmatrix} 5 \\ 4 \end{pmatrix}\begin{pmatrix} 7 \\ 4 \end{pmatrix}}{\begin{pmatrix} 13 \\ 8 \end{pmatrix}}$$

which gives prob. equal to 0.821.

[*6 MARKS*]

(c) This is a Hypergeometric problem with $N = 200$, $R = 120$, $n = 5$ and $r = 5$, thus the probability required is

$$\frac{\binom{5}{5}\binom{200-5}{120-5}}{\binom{200}{120}} = \frac{120 \times 119 \times 118 \times 117 \times 116}{200 \times 199 \times 198 \times 197 \times 196} = 0.075$$

which may also be calculated using a conditional probability argument.

Thus the probability that an all-male committee is selected, if all selections of five from two hundred are equally likely, is 0.075. This is small, but typically we consider a probability of 0.05 as providing sufficient evidence against a hypothesized model, so perhaps this is not conclusive evidence of sex-bias.

[6 MARKS]

EXERCISES

1. Hypergeometric probabilities with $N = 49$, $R = 6$, $n = 6$, $r = 0$.

$$\text{(i)P(No winning numbers)} = \frac{\binom{n}{r}\binom{N-n}{R-r}}{\binom{N}{R}} = \frac{\binom{43}{6}}{\binom{49}{6}} \qquad \text{(ii)P(No winning numbers for } x \text{ weeks)} = \left\{\frac{\binom{43}{6}}{\binom{49}{6}}\right\}^x$$

2. Hypergeometric probabilities with $N = 80$, $R = 20$, $n = 5, 10, 15$, $r = 0$.

P(No winning numbers) given by

$$\text{(i)}\ \frac{\binom{5}{0}\binom{75}{20}}{\binom{80}{20}} = \frac{\binom{20}{0}\binom{60}{5}}{\binom{80}{5}} \quad \text{(ii)}\ \frac{\binom{10}{0}\binom{70}{20}}{\binom{80}{20}} = \frac{\binom{20}{0}\binom{60}{10}}{\binom{80}{10}} \quad \text{(iii)}\ \frac{\binom{15}{0}\binom{65}{20}}{\binom{80}{20}} = \frac{\binom{20}{0}\binom{60}{15}}{\binom{80}{15}}$$

3. Total number of possible allocations is $n \times n \times ... \times n = n^n$ (MULTIPLICATION RULE). For n_E, consider selecting two different boxes from n to be the empty box and the box containing two balls; after this selection has been made, can arrange the balls in the boxes in $n!$ ways; hence $n_E = \binom{n}{2} \times n!$, and

$$\mathrm{P}(E) = \frac{\binom{n}{2} \times n!}{n^n}$$

4. (i) Total number of arrangements of $2n$ objects is $(2n)!$. For n_E, there are n *pairs* of *matched* short and long parts, which can be arranged in $n!$ ways in the sequence; each pair can be arranged in 2 ways. Hence $n_E = 2^n n!$, and hence

$$\mathrm{P}(E) = \frac{2^n n!}{(2n)!}$$

(ii) Here, for n_E, consider arranging a sequence of n short parts and a parallel sequence of n long parts to match with them; there are $n!$ ways of arranging the short parts and $n!$ ways of arranging the long parts. Once the parallel sequences have been arranged individually and paired, each pair can be arranged in 2 ways (long/short or short/long). Hence $n_E = 2^n n! n!$, and

$$\mathrm{P}(E) = \frac{2^n n! n!}{(2n)!} = \frac{2^n}{\binom{2n}{n}}$$

5. Total number of allocations of birthdays to months is 12^{30}. For n_E, select 6 months from 12 as the "two-birthday" months, and then *partition* the 30 birthdays into 12 subgroups, six which contain 2 birthdays, and six which contain three. Hence, using the MULTINOMIAL FORMULA for partitions,

$$n_E = \binom{12}{6}\frac{30!}{(2!)^6(3!)^6} \qquad \Longrightarrow \qquad \mathrm{P}(E) = \binom{12}{6}\frac{30!}{12^{30}(2!)^6(3!)^6}$$

6. Hypergeometric: $2n$ females (TYPE I) and $2n$ males (TYPE II), sample of size $2n$ without replacement from population of size $4n$. Hence probability of even split in each group is

$$\frac{\binom{2n}{n}\binom{2n}{n}}{\binom{4n}{2n}}$$

7. Multistage Hypergeometric: select the first hand with n then the second hand with m all without replacement. Hence by multiplication theorem, probability is

$$\frac{\binom{13}{n}\binom{39}{13-n}}{\binom{52}{13}} \times \frac{\binom{13-n}{m}\binom{26+n}{13-m}}{\binom{39}{13}}$$

By extension, need to select third hand in similar fashion, but once first three hands are selected, final hand is fixed. Hence probability is

$$\frac{\binom{13}{r_1}\binom{39}{13-r_1}}{\binom{52}{13}} \times \frac{\binom{13-r_1}{r_2}\binom{26+r_1}{13-r_2}}{\binom{39}{13}} \times \frac{\binom{13-r_1-r_2}{r_3}\binom{13+r_1+r_2}{13-r_3}}{\binom{26}{13}}$$

TUTORIAL

1. There are a total of 10^n selections of digits. For n_E, can have $r = 1, 2, ..., n$ distinct digits; consider selecting the r distinct digits, then partitioning the n digits into r subgroups (using the multinomial formula).

$n = 3$	r	n_E	*Probability*	
	1	$\binom{10}{1} \times \frac{3!}{3!}$	0.01	
	2	$\binom{10}{1}\binom{9}{1} \times \frac{3!}{2!1!}$	0.27	consider all 2/1 splits
	3	$\binom{10}{3} \times \frac{3!}{1!1!1!}$	0.72	consider all 1/1/1/ splits

$n = 4$	r	n_E	*Probability*	
	1	$\binom{10}{1} \times \frac{4!}{4!}$	0.001	
	2	$\binom{10}{1}\binom{9}{1} \times \frac{4!}{3!1!} + \binom{10}{2} \times \frac{4!}{2!2!}$	0.063	consider all 3/1 splits and all 2/2 splits
	3	$\binom{10}{1}\binom{9}{2} \times \frac{4!}{2!1!1!}$	0.432	consider all 2/1/1 splits
	4	$\binom{10}{4} \times \frac{4!}{1!1!1!1!}$	0.504	consider all 1/1/1/1 splits

For a general formula, we first consider how to split the r distinct digits; essentially we must find all solutions to $n_1 + n_2 ... + n_r = n$ to give the $n_1/n_2/.../n_r$ splits; if we then arrange $n_1, n_2, ...n_r$ into *descending* order we can then correctly identify the numbers of ways of selecting the r digits involved. For example, if $n = 6$ and $r = 3$, one possible choice for (n_1, n_2, n_3) is $(2, 2, 2)$, for which the correct number of ways is selecting the three digits is

$$\binom{10}{3} \quad \textbf{not} \binom{10}{1}\binom{9}{1}\binom{8}{1}$$

essentially because the possible results $(1, 1, 5, 5, 8, 8)$ and $(5, 5, 8, 8, 1, 1)$ are regarded as identical. The general formula follows by summing these quantities multiplied by the partition formula

$$\frac{n!}{n_1!...n_r!}$$

over all possible values of $n_1, n_2, ..., n_r$.

2. Method A: Consider choosing items in the population to label Type I and Type II, broken down by whether they are IN the sample or OUT of the sample. The denominator is the total number of ways of choosing R from N items to label Type I. The numerator is the number of ways of choosing r from n items IN the sample to be Type I items, multiplied by the number of ways of choosing the $R - r$ from $N - R$ items OUT of the sample to be Type II items.

Method B: Consider items to be IN the sample, broken down by Type. The denominator is the total number of ways of choosing n items from N to be IN the sample. The numerator is the number of ways of choosing r Type I items from R to be IN the sample, multiplied by the number of ways of choosing the $n - r$ Type II items from $N - R$ to be IN the sample.