WST211
Practical 4

WST211: 2016
Memorandum of Practical 4

Section A:

Question 1: SAS Program

```
data sum;
do i=1 to 10000;
dice=int(10*ranuni(1))+1;
head=int(2*ranuni(1));
x=dice+head;
x4=x*x*x*x;
if x>5 or x<1 then indb=1; else indb=0;
if 3<x<5 then indc=1; else indc=0;
output;
end;
proc freq;
tables x;
proc means mean;
var x4 indb indc;
run;
```

Question 1: SAS Output

a) The empirical probability mass function is given in the following table.

x	Empirical value for $P(X = x)$
1	0.0506
2	0.1023
3	0.0943
4	0.1008
5	0.0979
6	0.1026
7	0.0967
8	0.1035
9	0.0993
10	0.1012
11	0.0508

x	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	**506**	**5.06**	**506**	**5.06**
2	**1023**	**10.23**	**1529**	**15.29**
3	**943**	**9.43**	**2472**	**24.72**
4	**1008**	**10.08**	**3480**	**34.80**
5	**979**	**9.79**	**4459**	**44.59**
6	**1026**	**10.26**	**5485**	**54.85**
7	**967**	**9.67**	**6452**	**64.52**
8	**1035**	**10.35**	**7487**	**74.87**
9	**993**	**9.93**	**8480**	**84.80**
10	**1012**	**10.12**	**9492**	**94.92**
11	**508**	**5.08**	**10000**	**100.00**

Variable	Mean
x4	3292.67
indb	0.5541000
indc	0.1008000

b) $P(|X-3|>2) = P(X>5) + P(X<1) = 0.5541$
c) $P(|X-4|<1) = P(3<X<5) = 0.1008$
d) $E(X^4) = 3292.67$

Question 2:

2(a)

X ~ HYP(4,7,10) , P(X=2) = $\dfrac{\binom{7}{2}\binom{3}{2}}{\binom{10}{4}}$ **= 0.3**

(b)

X ~ HYP(4,7,10), P(X >= 3) = $\dfrac{\binom{7}{3}\binom{3}{1}}{\binom{10}{4}} + \dfrac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} = \dfrac{2}{3}$

(c)

SAS Program:

```
Data Gend;
       n = 2000;
       Do i = 1 to n ;
              female = 7;
              male = 3;
              count = 0;
              Do j = 1 to 4;
                     *Recalculates the probabilities after every step;
                     f= female/(female + male);
                     m= male/(female + male);
                    *Selects 1 (female) with prob f and 2 (male) with prob m;
                     x = rand('table',f,m);
                     If x = 1 then do;
                        *Counts the number of females selected;
                            count = count + 1;
                      *Readjusts number of females because there is no replacement;
                            female = female - 1;
                     end;
                     else male = male - 1;
                     *Readjusts number of males because there is no replacement;
              end;
              output;
      end;
proc freq data=gend;
 tables count;
run;
```

Students will get different empirical values. The main aim of the question is to compare the empirical values with the theoretical values obtained from part a and b.

Question 3:

(a)

$X \sim POI(3t)$ t in minutes $\quad X(1) \sim POI(3)$

$$
\begin{aligned}
P(X \geq 2) &= 1 - P(X < 2) \\
&= 1 - P(X \leq 1) \\
&= 1 - \frac{\sum_{x=0}^{1} e^{-3} 3^{x}}{x!} \\
&= 1 - \{0.1991\} \text{from table 2} \\
&= 0.8009
\end{aligned}
$$

(b)

$$
\begin{aligned}
&P(A \cap B) \\
&= P\left[\left(\text{at least 4 in first minute}\right) \cap \left(\text{at most 2 in second mimute}\right)\right] \\
&= P\left[\left(\text{at least 4 in first minute}\right)\right] \times P\left[\left(\text{at most 2 in second mimute}\right)\right] \\
&\text{Since events in disjoint intervals are independent} \\
&= \mathrm{P}\left[X \geq 4\right] \times \mathrm{P}\left[X \leq 2\right] \\
&= \left\{1 - P\left(X \leq 3\right)\right\} \times \mathrm{P}\left[X \leq 2\right] \\
&= \left\{1 - 0.6472\right\} \times 0.4232 \\
&= 0.1493
\end{aligned}
$$

(c)

SAS Program and Output

```
data cars;
n = 2000;
count = 0;
count1 = 0;
do i = 1 to n ;
      * generates a random value from a Poisson(3) distribution;
      x = ranpoi(0,3);
      y = ranpoi(0,3);
      *counts the number of times there were at least 2 cars in a given minute;

      if x >= 2 then count = count + 1;
*counts the number of times there were at least 4 cars in the first minute
      and at most 2 cars in the second minute;
      If x >= 4 and y <= 2 then count1 = count1 + 1;
end;
*calculates the probability (#event happened)/#iterations;
ProbA = count/n;
ProbB = count1/n;

Proc print data = cars;
var ProbA ProbB;
run;
```

Students will get different empirical values. The main aim of the question is to compare the empirical values with the theoretical values obtained from part a and b.

Section B

	Population exposed to risk			Number of deaths			Deaths per 100 exposed to risk (two decimal places)		
Economic Status	**Male**	**Female**	**Both**	**Male**	**Female**	**Both**	**Male**	**Female**	**Both**
I (high)	174	139	313	118	4	122	67.80	2.88	38.98
II	179	103	282	154	13	167	86.03	12.62	59.22
III	496	196	692	413	106	519	83.27	54.08	75.00
Other	**853**	**20**	**873**	**661**	**3**	**664**	**77.49**	**15.00**	**76.06**
Total	1702	458	2160	1346	126	1472	79.08	27.51	68.15

This table was completed from the output below:

Table 1 of Sex by Survived			
Controlling for Class=0			
Sex(Sex)	**Survived(Survived)**		
Frequency Percent Row Pct Col Pct	**0**	**1**	**Total**
0	3 0.34 15.00 0.45	17 1.95 85.00 8.13	20 2.29
1	661 75.72 77.49 99.55	192 21.99 22.51 91.87	853 97.71
Total	664 76.06	209 23.94	873 100.00

Table 2 of Sex by Survived			
Controlling for Class=1			
Sex(Sex)	**Survived(Survived)**		
Frequency Percent Row Pct Col Pct	**0**	**1**	**Total**
0	**4 1.28 2.88 3.28**	**135 43.13 97.12 70.68**	**139 44.41**
1	**118 37.70 67.82 96.72**	**56 17.89 32.18 29.32**	**174 55.59**
Total	**122 38.98**	**191 61.02**	**313 100.00**

Table 3 of Sex by Survived			
Controlling for Class=2			
Sex(Sex)	Survived(Survived)		
Frequency Percent Row Pct Col Pct	0	1	Total
0	13 4.61 12.62 7.78	90 31.91 87.38 78.26	103 36.52
1	154 54.61 86.03 92.22	25 8.87 13.97 21.74	179 63.48
Total	167 59.22	115 40.78	282 100.00

Table 4 of Sex by Survived			
Controlling for Class=3			
Sex(Sex)	Survived(Survived)		
Frequency Percent Row Pct Col Pct	0	1	Total
0	106 15.32 54.08 20.42	90 13.01 45.92 52.02	196 28.32
1	413 59.68 83.27 79.58	83 11.99 16.73 47.98	496 71.68
Total	519 75.00	173 25.00	692 100.00

SAS Program:

```
proc freq data=titanic;
tables class*sex*survived;
run;
```