WST 211

Practical 3

Date: 7 March 2016

Due: 14 March 2016

Instructions:

- Answer all questions.
- Hand in a typed document with your answers which includes the following:
- Table of contents
- Answers
- SAS program
- SAS output
- Attach a copy of the SAS program in an appendix.
- Attach the relevant outputs of the SAS output.
- According to the question, make interpretations about the SAS output.
- Round the answers to 3 decimal places.
- Hand in a typed document with your answers and include the SAS programs.

Questions:

1. Create DATASET D in SAS with $n = 100000$ observations generated from a $BIN\,(9, 0.7)$ population with the RANBIN function in SAS:

   $$=\text{RANBIN}(\text{seed}, n, p)$$

   NB: Use a seed of 7896 and name the variable WINS.

   Calculate the following **empirical values** from DATASET D.

   (a) Make use of PROC CHART to obtain a bar chart of the variable WINS. Compare the shape of the empirical distribution with that of the theoretical distribution.

(b) Use PROC FREQ and the variable WINS to calculate the empirical values for probability mass function of X. Compare with 1(a).

(c) Use the empirical distribution in 2(a) to calculate the following: (Compare with theoretical values.)

i. $P(X \leq 5)$

ii. $P(3 \leq X \leq 7)$

iii. Expected number of 'Wins' by making use of the formula

$$\mu = \sum xp(x)$$

(d) Answer the previous question (2(c)) with PROC MEANS.
Hint: Use the variable WINS and create the classification variables GrpA and GrpB in DATASET D.

GrpA		WINS
1	:	$\{0, 1, 2, 3, 4, 5\}$
0	:	otherwise

GrpB		WINS
1	:	$\{3, 4, 5, 6, 7\}$
0	:	otherwise

2.

Allocation of probabilities:

The classical definition of probability assumes, with other assumptions, that we know the number of elements in the sample space to allocate probabilities.

Example:

Probability to obtain head if we flip a balanced coin; 1 of 2 possible outcomes.

How will you allocate a probability for the event rain tomorrow? In this case the structure of the sample space is more complex and we need other methods to allocate the probabilities. We can do this by using the frequency limit definition of probability.

Frequency limit definition of probability:

The probability of an event A is given by the limit of the ratio of the number of times event A happens divided by the number of times the random experiment is performed. Mathematically we write

$$\mathbf{P(A)} = \lim \frac{\mathbf{number\ of\ A}}{\mathbf{k}}$$

where k is the number of times the experiment is performed.

Thus, if we flip a balanced coin k times, and count the number of times that we obtained head, the ratio will tend to $\frac{1}{2}$, a value which we have a natural feeling for.

By increasing the number of rolls to 50, 100, 1000, etc. the ratio will close in on 1/6. Since it is not feasible to do this by hand, we will perform a computer simulation. For this purpose, study the WST211 notes about Simulations in SAS.

a. Generate 25 random numbers from a uniform distribution over the interval [0,1], using a seed value of 72. Assign 1 as the outcome if the number generated is in the interval [0, 1/6];

2 if in the interval (1/6, 2/6]; 3 if in (2/6, 3/6]; 4 if in (3/6, 4/6]; 5 if in (4/6, 5/6] and 6 if in (5/6, 1].

Calculate the ratio: $\frac{\textbf{number of 5's}}{\textbf{k}}$, where k is the number of times the experiment is repeated. Repeat the above experiment 200 times and accumulate the number of 5's. Output the accumulated ratio after every 25 rolls. Thus, the first ratio will be $\frac{\textbf{number of 5's after 25 rolls}}{\textbf{25}}$, the second ratio will be $\frac{\textbf{number of 5's after 50 rolls}}{\textbf{50}}$ and the last ratio in the sequence will be $\frac{\textbf{number of 5's after 5000 rolls}}{\textbf{5000}}$.

b. Draw a graph of the ratios in (a) with the number of rolls on the x-axes and the ratio on the y-axes. Use PROC GPLOT instead of PROC PLOT.

Does the graph support the frequency limit definition? Motivate.