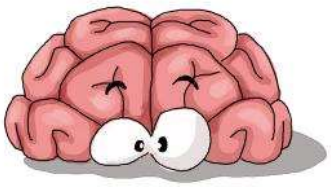
WST211
Practical 7



Date: 22 April 2016
Due: 3 May 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. Use the PDF function to draw the density function of a WEI (θ,β) distribution with β fixed at 3 and θ assuming four different values; 1, 1.2, 1.4 and 1.6. (**see Section 3.4 of WST211 notes**). Use *x* values from 0 to 3 by 0.1. Plot the four density functions on the same graph.

SAS Code for graph:

```
goptions reset=all i=join;
symbol1 colour=blue line=1 width=2;
symbol2 colour=purple line=3 width=2;
symbol3 colour=green line=33 width=2;
symbol4 colour=red line=5 width=2;

proc gplot;
plot (fx1 fx2 fx3 fx4)*x/vaxis=0 to 3 overlay legend;
run;
```

What is the effect of the parameter θ on the density function?

2. Write a SAS program which will perform the following:
   a) Generate 40 random numbers from an WEI (θ,β) distribution with θ = 2 and β = 0.9 by using the RAND ('WEIBULL', a, b) function.
   Note that in the syntax used in SAS a = β and b = θ.
   b) Calculate the empirical distribution function of the 40 values, that is, calculate $F_{40}(x)$.
   c) Calculate the theoretical values, F(x), for the distribution function for the same *x* values as in a) by using the CDF ('WEIBULL', x, a, λ) function.
   Note that in the syntax used in SAS a = β and λ = θ.

d) On the same set of axes draw a step graph to represent b) and a continuous graph to represent c). The code for the graph is given below.

Note on calculating the empirical distribution function:
Consider the dataset, datx, consisting of the following four values:

Obs	x
1	0.04454
2	0.56508
3	0.06747
4	0.93391

The following SAS code will create the dataset ordx which has the same data as datx but with the observations ordered and numbered from ii=1 to 4. This dataset can be used to calculate the empirical distribution function.

```
proc sort data=datx;
by x;

data ordx;
set datx;
ii=_n_;

proc print data=ordx;
var ii x; run;
```

Obs	ii	x
1	1	0.04454
2	2	0.06747
3	3	0.56508
4	4	0.93391

SAS Code for graph:

```
goptions reset=global;
symbol1 color=blue interpol=step;
symbol2 color=black interpol=join;

proc gplot data=wei;
plot Fn*x F*x/overlay;
run;
```

3. Repeat Question 2 (a) – (d), but generate 400 random numbers. In which case is the empirical distribution function closer to the theoretical distribution function? Why?