

WST 211

Practical 2

Date: 15 February 2016

Due: 22 February 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:

IQ is known to be approximately normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$. Let

$X =$ IQ of randomly selected person

then $X \sim N(100, 15^2)$.

1. Create DATASET A in SAS with the following **theoretical values**.

- (a) Calculate the quantiles

P_{10}	P_{25}	P_{50}	P_{75}	P_{90}
----------	----------	----------	----------	----------

of X by making use of the QUANTILE function in SAS:

$$= \text{QUANTILE}('NORMAL', p, \mu, \sigma)$$

Note: The QUANTILE function for the normal distribution returns the $p \times 100^{th}$ quantile for the normal distribution, with mean μ and standard deviation σ .

- (b) Calculate the following probabilities:

- i. $P(X > 110)$
- ii. $P(90 \leq X \leq 110)$

by making use of the CDF function in SAS:

$$= \text{CDF}('NORMAL', x, \mu, \sigma)$$

Note: The CDF function for the normal distribution returns the probability that an observation from the normal distribution, with mean μ and standard deviation σ is less than or equal to x .

2. Create **DATASET B** in SAS with $n = 1000$ observations generated from a $N(100, 15^2)$ population with the RANNOR function in SAS:

$$= \mu + \text{RANNOR}(\text{seed}) * \sigma$$

NB: Use a seed of 15 and name the variable IQ.

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

- (a) Make use of PROC UNIVARIATE to investigate the distribution of IQ.

- i. Give the empirical values for μ and σ . Compare with the theoretical values.
- ii. Obtain the empirical values of the quantiles

P_{10}	P_{25}	P_{50}	P_{75}	P_{90}
----------	----------	----------	----------	----------

for the variable IQ. Compare with the theoretical values.

- iii. Draw a histogram with the normal density functions for the theoretical and empirical distributions by making use of the NORMAL option in the histogram statement:

$$\text{histogram / normal (color=(red blue) mu}=\mu \text{ est sigma}=\sigma \text{ est);}$$

Compare the empirical (blue) and theoretical (red) distributions.

(b) Create the variable IQgrp in DATASET B with the following categories:

IQgrp		IQ
1	:	< 90
2	:	$90 - 110$
3	:	> 110

Make use of PROC FREQ and the variable IQgrp to obtain the following probabilities:

- i. $P(X > 110)$
- ii. $P(90 \leq X \leq 110)$

Compare these empirical values with the theoretical values.