MAT2379, Introduction to biostatistics

Assignment 2

Due date: Friday October 21, 2016 at 3:00 p.m.

Total = 100 marks

Part (I) (50 marks) Solve the following problems from the textbook using a *Faculty-standard calculator* [1] and a *Standard Normal Table*:

6.1 (10 marks), 6.5 (10 marks), 7.8 (15 marks), 9.4* (15 marks)

* In 9.4, add question (c) Find the sample mean and the sample standard deviation.

Part (II) (50 marks) Use `R` to solve the following problems.

1. (5 marks) Let X be a random variable following a binomial distribution with $n = 12$ trials and probability of success $p = 0.3$. Use `R` to find the following probabilities:
   a) $P(X \geq 3)$; **b)** $P(4 \leq X < 8)$.

2. (15 marks) Let X be a random variable following a normal distribution with mean $\mu = 8$ and standard deviation $\sigma = 2.5$.
   Use `R` to find the following probabilities **a)** $P(X \leq 9)$; **b)** $P(1.8 < X < 7.8)$.
   Use `R` to find a value t such that **c)** $P(X > t) = 0.25$; **d)** $P(|X - 8| < t) = 0.8$

3. (20 marks) The following data give the length (in cm) of 12 gentoo penguins and 12 chinstrap penguins.

Gentoo	76.9,	84.2,	78.0,	78.1,	80.8,	78.2,	81.8,	82.1,	80.8,	75.7,	83.5,	78.6
Chinstrap	76.2,	76.3,	75.5,	75.9,	76.5,	73.1,	73.8,	76.3,	74.9,	76.3,	76.8,	75.7

   a) Input this dataset into `R` as two vectors. Find the five number summary for each group.
   b) Construct the histogram for each group (using at least 6 bins). What do you observe on the shape of the data distribution in each group?
   c) Construct the side-by-side boxplots for the above two groups. For each group, describe the data distribution in shape, central tendency, and dispersion. Are there any outliers?

4. (10 marks) The Kelvin is the primary unit of temperature measurement in the physical sciences. It is on the same scale as degrees Celsius. However, 0 Kelvin represents an absolute zero, since it is the temperature at which all thermal motion ceases. The summary of descriptive statistics for temperatures in degrees Kelvin for 15 infants is given in the following:

[1] Only the following calculators are allowed during Faculty of Science examinations: Texas Instruments TI-30 and TI-34, Casio FX-260 and Casio FX-300 (scientific and non-programmable calculators).

```
> summary(K)
   Min.    1st Qu.  Median    Mean    3rd Qu.    Max.
   309.1   310.2    310.7    310.8    311.5    312.3
> sd(K)
[1] 1.04
```

Find all summary of descriptive statistics (i.e. x_{min}, x_{max}, Q_1, Q_2, Q_3, IQR, sample mean and sd) for temperatures in degrees Fahrenheit. The formula for converting Kelvin to Fahrenheit is:

$$F = K \times 9/5 - 459.67.$$

Please write down your answers (including necessary steps, if applicable) to all the questions in Parts I and II, and attach the printout of your R workspace for the questions in Part II. Note that only attaching R output is not an answer to a question. All assignments must be handed in to the Assignment Box in the lobby of KED building (585 King Edward Ave) by the deadline.