Tutorial letter 201/2/2018

Basic Statistics

STA1510

Semester 2

Department of Statistics

SOLUTIONS TO ASSIGNMENT 01

CHAPTER 1

QUESTION 1

Scales of measurement
Rating availability of parking space implies an ordinal variable and; listing occupation is a nominal variable.

Option 2

QUESTION 2

Types of data
Number of viewers of a movie is a countable numeric value. It will be classified as: Quantitative, discrete data.

Option 1

CHAPTER 2

QUESTION 3

Summarising numeric data.
From the histogram only 7 shoppers out of 30 spent R800 or less.
This means that $\frac{7}{30} \times 100 = 23.33\%$

Option 5

QUESTION 4

Numeric data
From the histogram only 26, i.e. $(7 + 14 + 5)$ shoppers spent R1600 or less

Option 1

QUESTION 5

Summarising categorical data
From the table only 7 out of 30, i.e. 23.33% of all shoppers surveyed are males who prefer to shop at Pick and Pay.

Option 1

CHAPTER 3

QUESTION 6

Quartiles
Start by ordering the values as follows:

$$-4,\ -3,\ -3,\ -1,\ 1,\ 2,\ 4,\ 5,\ 6,\ 10,\ 10$$
$$n = 11$$

First quartile (lower quartile) is found in position: $\frac{n+1}{4} = \frac{11+1}{4} = 3^{\text{rd}}$ position of an ordered list.

$\therefore$ First quartile $(Q_1) = -3$

Third quartile (upper quartile) is found in position:

$$\frac{3\,(n+1)}{4} = \frac{3\,(12)}{4} = 9^{\text{th}} \text{ position}$$

$\therefore$ Third quartile $(Q_3) = 6$

$$\begin{aligned} \text{Interquartle range } (IQR) &= Q_3 - Q_1 \\ &= 6 - (-3) \\ &= 6 + 3 \\ &= 9 \end{aligned}$$

Option 2

QUESTION 7

Numerical descriptive measures we have

0	0	0	$n = 1$
1	1	1	Mode $= 0$ and 1
2	2		Median $= 1$
4	4		Mean $= 1.82$
5			Standard deviation $= 1.78$

Thus, mean is greater than the median, telling us that the distribution of this data set is positively skewed or skewed to the right.

Option 1

QUESTION 8

Measure of variation or dispersion
Coefficient of variation $(CV) = \frac{S}{\bar{X}} \times 100$

From Question 7, mean $= 1.82$ and standard deviation $= 1.78$

$\therefore$ CV $= \dfrac{1.78}{1.82} \times 100 = 97.80\%$

The standard deviation is 97.80% of the size of the mean.

Option 5

QUESTION 9

Descriptive measures

4	1	5	8				
5	0	2	2	5	9		
6	1	2	5	5	6	6	7
7	0	3					

$n = 17$

Range $= 73 - 41 = 32$

Mode $= 52,\ 65,\ 66$

The numbers $52,\ 65,\ 66$ appears twice each. Therefore, Option 2 becomes an incorrect statement. We have a trimodal case in the data set.

Option 2

CHAPTER 4

QUESTION 10

Probability

$P(A) = 0.30,\ P(B) = 0.20$ A & B mutually exclusive means $P(A \cap B) = 0.$

Option 5 is the only incorrect statement, because $P(A|B) \neq P(A)$ which is required for A and B to be independent.

$P(A|B) = 0$ and $P(A) = 0.30.$ All other statements are correct.

Option 5

QUESTION 11

Probability (intersection of events)

$P\left(\text{White farmer}\right) = \dfrac{5}{138} = 0.0362$

Option 5

QUESTION 12

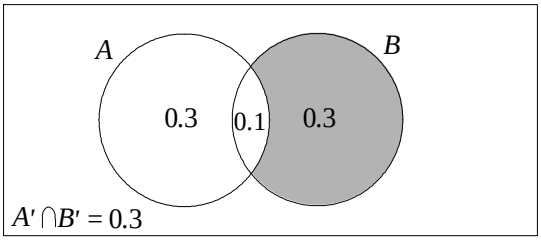
Addition rule

P (Neither Indian nor Teacher) $= \dfrac{74}{138} = 0.5362$

Option 1

QUESTION 13

Probability



From the Venn diagram, the $P\left(A' \cap B'\right)$ is equal to 0.3.

Option 4

CHAPTER 5

QUESTION 14

Poisson
$\lambda = 3$ per night

$$P\left(X = 5\right) = \frac{3^5 \times e^{-3}}{5!} = 0.1008$$

or just use the Poisson probability table, using the fact that $\lambda = 3$ with $X = 5$.

Option 3

QUESTION 15

Binomial
$\pi = 0.8$ within 12 hours
$n = 0.2$ after 12 hours (or late)
$n = 7$
Only one parcel means $P\left(X = 1\right)$

$$\begin{aligned} P\left(X = 1\right) &= nC_X \pi^X \left(1 - \pi\right)^{n-X} \\ &= 7C_1 \left(0.2\right)^1 \left(0.8\right)^{7-1} \\ &= 7 \times 0.2 \times 0.2621 \\ &= 0.3670 \end{aligned}$$

or just use the binomial probability table, using the fact that $\pi = 0.2,\ n = 7$ and $X = 1$.

Option 1

QUESTION 16

Binomial
$\pi = 0.8$ within 12 hours
$n = 7$
Only one parcel means $P(X = 1)$

$$\begin{aligned} P(X=1) &= {}_nC_X \pi^X (1-\pi)^{n-X} \\ &= {}_7C_1 (0.8)^1 (0.2)^{7-1} \\ &= 7(0.8)(0.0001) \\ &= 0.0004 \end{aligned}$$

or just use the binomial probability table, using the fact that $\pi = 0.8,\ n = 7$ and $X = 1.$

Option 1

QUESTION 17

Discrete random variable.
Option 3 is the correct statement as follows:

$$\begin{aligned} P(2 < X \leq 4) &= P(X \text{ is greater than } 2 \text{ but less or equal to } 4) \\ &= P(X=3) + P(X=4) \\ &= 0.299 + 0.138 \\ &= 0.4370 \end{aligned}$$

Option 3

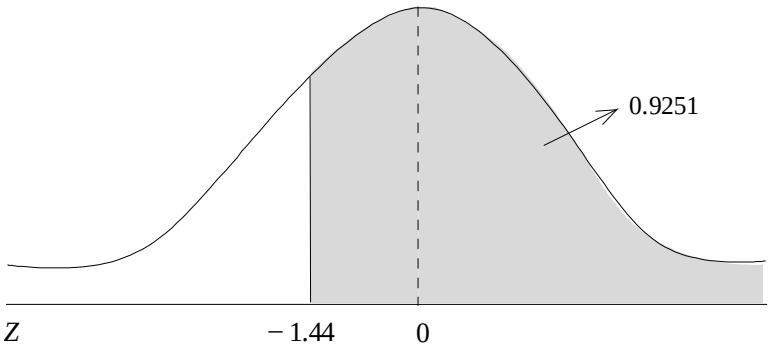
CHAPTER 6

QUESTION 18

Standard normal distribution

$$P(Z \underset{\downarrow}{>} -1.44) = 0.9251$$

Means greater than

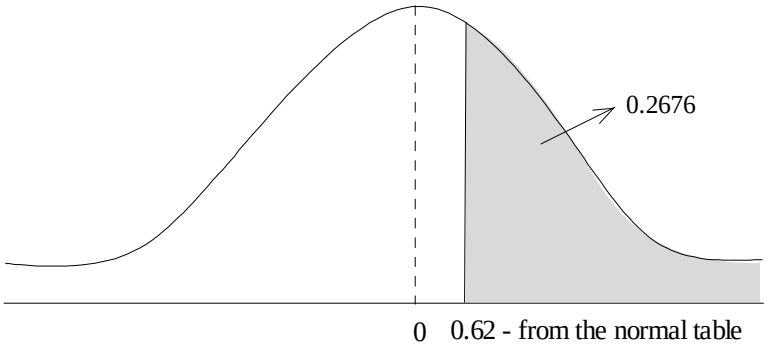


All other options are incorrect.

Option 1

QUESTION 19

Normal distribution
$\mu = 25$ and $\sigma = 2.5$
$P(X > K) = 0.2676$



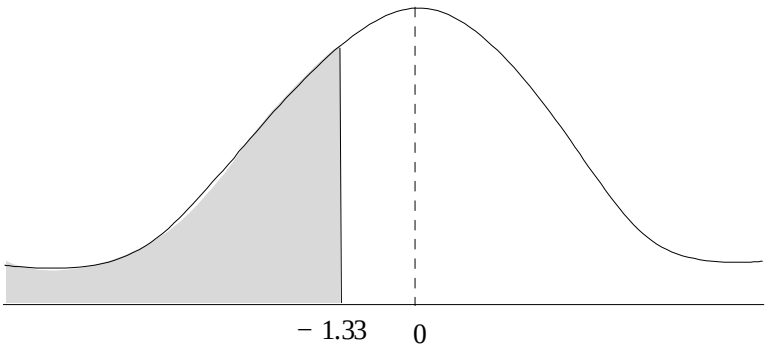
$$\begin{aligned} \therefore \quad X &= \mu + z\sigma \\ &= 25 + (0.62)(2.5) \\ &= 26.55 \\ \therefore \quad P(X > 26.55) &= 0.2676 \end{aligned}$$

Option 1

QUESTION 20

Normal distribution

$$\begin{aligned} \mu &= 5 \quad \text{and} \quad \sigma = 0.75 \\ & P(X \leq 4) \text{ means at most four} \\ & P\left(Z \leq \frac{4-5}{0.75}\right) \\ & P(Z \leq -1.33) \\ &= 0.0918 \end{aligned}$$



Option 1