STQD 6214 Assignment 1

1)

a) Frequency distribution table

Trust on the Internet	Frequency	Relative frequency	Percentage %
A	4	$4/40 = 1/10$	10
M	28	$28/40 = 7/10$	70
H	6	$6/40 = 3/20$	15
S	2	$2/40 = 1/20$	5
Total	40	1	100

b) The mode for this data is M as it appears more frequently (28 times) than any other survey.

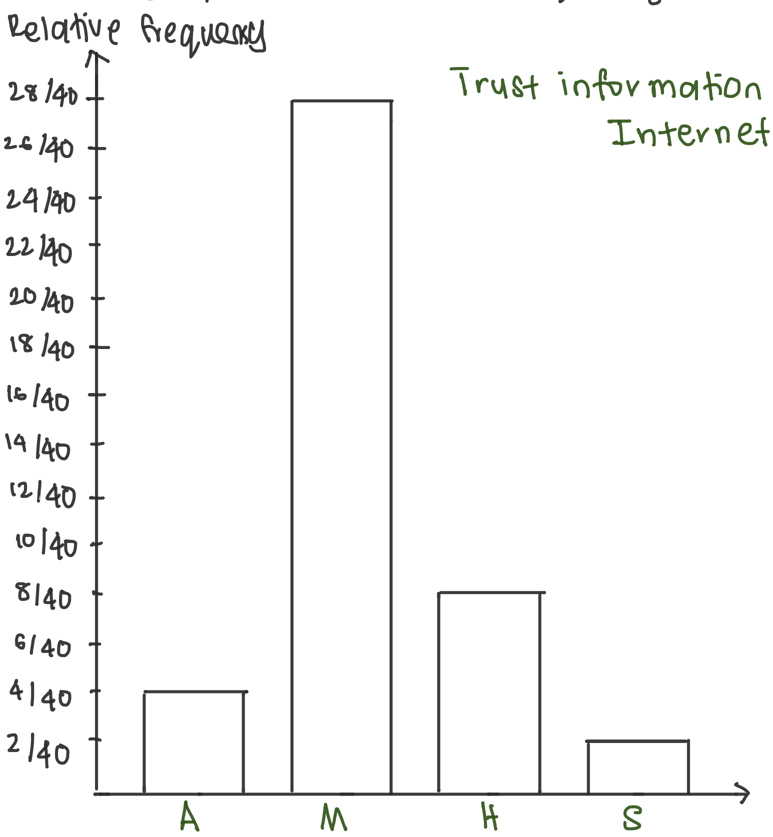
c) Percentages of the respondent trust more than one-half of what they read.

A = trust in all that they read = 10%
M = trust in most of what they read = 70%

$\underline{\underline{80\%}}$

∴ 80% of the respondent trust more than one-half of what they read.

d) Bar graph for relative frequency



② This is population data

62, 67, 72, 73, 75, 77, 81, 83, 84, 85, 90, 100, 107, 116, 135

a) Mean, $\mu = \frac{\Sigma x}{N} = \frac{1307}{15} = 87.13$

median = 83

variance $\sigma^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}{N}$

$= \frac{119361 - \frac{(1307)^2}{15}}{15}$

$= \frac{119361 - 113883.27}{15}$

$= 365.182$

std dev $\sigma = \sqrt{365.182}$

$= 19.11$

b)

[62, 67, 72, 73, 75, 77, 81], 83, [84, 85, 90, 100, 107, 116, 135]

Q1 — Q2 — Q3

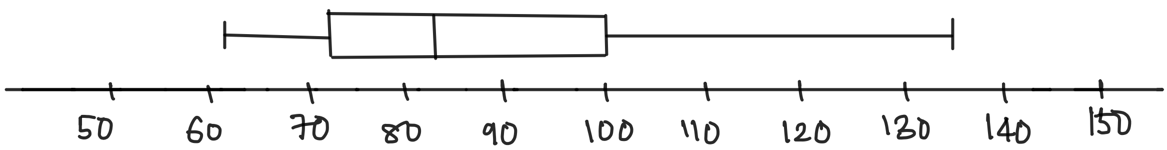
$Q_1 = 73$
$Q_2 = 83$
$Q_3 = 100$

$IQR = Q_3 - Q_1 = 100 - 73 = 27$

$$\begin{aligned}\text{Lower Inner fence} &= Q_1 - 1.5\ IQR \\ &= 73 - 1.5(27) \\ &= 32.5\end{aligned}$$

$$\begin{aligned}\text{Upper Inner fence} &= Q_3 + 1.5\ IQR \\ &= 100 + 1.5(27) \\ &= 140.5\end{aligned}$$

Smallest within inner fence = 62
Largest within inner fence = 135



From the box and whisker plot, there is no outlier.

c) The data is skewed to the right as it seems longer whisker on the right.

d) For skewed or data with outlier, the mean may not be accurately represent the center of the data tendency.

As for median, it represents the middle value when data is ordered, making it good measure of centrel tendency for skewed tendency.

∴ For this case, median is more preferable to be used to measure the central tendency.

③

	Male	Female	
Bachelor	721	982	1703
Master	292	439	731
Doctoral	80	84	164
	1093	1505	2598

a) $P(\text{Bachelor's Degree}) = \frac{1703}{2598} = 0.6555$

b) $P(\text{Male and Master}) = \frac{292}{2598} = 0.1124$

c) $P(\text{Doctoral} \mid \text{Female}) = \frac{84}{1505} = 0.0558$

d) $P(\text{Bachelor or master}) = \frac{721 + 982 + 292 + 439}{2598} = \frac{2434}{2598} = \frac{1217}{1299} = 0.937$

e) $P(\text{male and master}) = P(\text{male}) \times P(\text{master})$

$$\frac{292}{2598} = \frac{1093}{2598} \times \frac{731}{2598}$$

$$0.1124 \neq 0.1184$$

∴ Event are not independent. The occurences of 1 event does not affect the occurences of one another.

④ Blue, B = 7
Red, R = 3
Green, G = 6
Yellow, Y = 4
Total = 20

a) $^{n}C_{r} = \frac{n!}{(n-r)!r!} = {}^{20}C_{4} = \frac{20!}{(16)!4!} = 4845$, There are 4 8 4 5 total combinations.

b) The order is important, hence it is permutation. $^{n}P_{r} = \frac{n!}{(n-r)!}$

$^{20}P_{4} = 116\,280$ or $\underline{20}\ \underline{19}\ \underline{18}\ \underline{17} = 116\,280$.

c) $= \frac{\text{4 favorable outcomes}}{\text{total number of possible outcome}}$

i- $P(\text{all 4 are B}) = \frac{^{7}C_{4}}{4845} = \frac{35}{4845} = 7.224 \times 10^{-3}$

ii- $P(2R, 1B, 1G) = \frac{^{3}C_{2} \times {}^{7}C_{1} \times {}^{6}C_{1}}{4845} = \frac{126}{4845} = 0.0260$

iii- $P(\text{4 different colors}) = \frac{^{3}C_{1} \times {}^{7}C_{1} \times {}^{6}C_{1} \times {}^{4}C_{1}}{4845} = \frac{504}{4845} = 0.1040$

iv- P(all are R) = Selecting 4 items out of 3 is impossible, hence the number of ways to satisfy the condition is 0.

$= \frac{0}{4845}$

$= 0$ #

5. For probability question in discrete random variable, i'm constructing the probability distribution of x first.

$P(X=-1) = 1/3 - 0 = 1/3$
$P(X=1) = 1/2 - 1/3 = 1/6$
$P(X=3) = 3/4 - 1/2 = 1/4$
$P(X=5) = 1 - 3/4 = 1/4$

x	-1	1	3	5
$P(X=x)$	$1/3$	$1/6$	$1/4$	$1/4$

a) $P(X \leq 3) = F(-1) + F(1) + F(3)$
$= 1/3 + 1/6 + 1/4$
$= 3/4$

b) $P(X=5) = F(5)$
$= 1/4$

c) $P(X \geq 1) = F(1) + F(3) + F(5)$
$= 1/6 + 1/4 + 1/4$
$= 2/3$

d) $P(-0.4 < X < 4) = F(1) + F(3)$
$= 1/4 + 1/6$
$= 5/12$

⑥

a- Let X be the number of household that never purchased organic food.
X has Binomial distribution with $n = 10$, $p = 0.7$, $q = 0.3$

possible values of $X = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$

b- mean, $\mu = np = (10)(0.7) = 7$

variance, $\sigma^2 = npq = (10)(0.7)(0.3) = 2.1$

std-dev, $\sigma = \sqrt{2.1} = 1.4491$

c- Probability of :

i- $P(X = 10) = {}^{10}C_{10}\,(0.7)^{10}(0.3)^{0} = 0.0282$

ii- $P(X = 6) = {}^{10}C_{6}\,(0.7)^{6}(0.3)^{4} = 0.2001$

iii- $P(X \geq 8) = P(8) + P(9) + P(10)$
$= {}^{10}C_{8}\,(0.7)^{8}(0.3)^{2} + {}^{10}C_{9}\,(0.7)^{9}(0.3)^{1} + {}^{10}C_{10}\,(0.7)^{10}(0.3)^{0}$
$= 0.2335 + 0.1211 + 0.0282$
$= 0.3828$

7

a) Let X be the number of purchases a person make until the person purchase a bottle with a winning bottle cap.
X has a Geometric distribution and each trial is independent.

$p = 1/6$

mean, $\mu = \frac{1}{p} = \frac{1}{1/6} = 6$

variance, $\sigma^2 = \frac{1-p}{p^2} = \frac{1-(1/6)}{(1/6)^2} = \frac{5/6}{1/36} = 30$

b) Probability of :

i- $P(x=3) = \frac{5}{6} \times \frac{5}{6} \times \frac{1}{6}$

$= 0.1157$

ii- $P(X \leq 3) = \frac{1}{6} + \left(\frac{5}{6} \times \frac{1}{6}\right) + \left(\frac{5}{6} \times \frac{5}{6} \times \frac{1}{6}\right)$

$= 0.4213$

iii - P(Not winning on 1 purchase) $= \left(\frac{5}{6}\right)^5$

$= 0.4019$

(8)

a) cumulative distribution function, cdf

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$= \int_{-\infty}^{x} \frac{1}{288}(36-t^2)\,dt \qquad \text{for } -6 < x < 6$$

$$= \frac{1}{288}\int_{-6}^{x}(36-t^2)\,dt$$

$$= \frac{1}{288}\left[36t - \frac{t^3}{3}\right]_{-6}^{x}$$

$$= \frac{1}{288}\left[\left(36x - \frac{x^3}{3}\right) - (-216-(-72))\right]$$

$$= \frac{1}{288}\left(36x - \frac{x^3}{3} + 144\right)$$

$$= \frac{x}{8} - \frac{x^3}{864} + \frac{1}{2}$$

$$F(x) = \begin{cases} 0 & \text{for } x \leq -6 \\ \frac{x}{8} - \frac{x^3}{864} + \frac{1}{2} & \text{for } -6 < x < 6 \\ 1 & \text{for } x \geq 6 \end{cases}$$

(8)

b) Probabilities one of these flights :

i- $P(x \le -2) = F(-2) = \frac{(-2)}{8} - \frac{(-2)^3}{864} + \frac{1}{2} = 0.2593$

ii- $P(x \ge 1) = 1 - P(x < 1)$

$= 1 - F(1)$

$= 1 - \left(\frac{1}{8} - \frac{13}{864} + \frac{1}{2}\right)$

$= 0.3762$

iii- $P(1 \le x \le 3) = F(3) - F(1)$

$= \left[\left(\frac{3}{8} - \frac{(3)^3}{864} + \frac{1}{2}\right) - \left(\frac{1}{8} - \frac{(1)^3}{864} + \frac{1}{2}\right)\right]$

$= \left[\frac{27}{32} - \frac{539}{864}\right]$

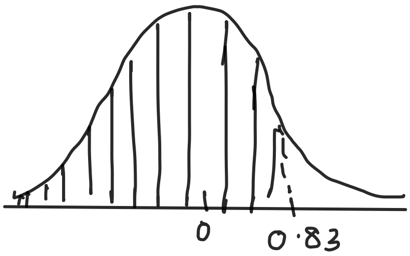
$= 0.2199$

iv- $P(x = 5) = 0$, Because area under the curve of 5 is zero.

(9) Let X be electricity consumption for all household.

$X \sim N(1250, 300^2)$

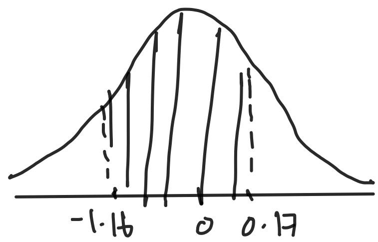
a) $P(X < 1500) = P\left(Z < \frac{1500 - 1250}{300}\right)$



$= P(Z < 0.83)$
$= 1 - P(Z < 0.83)$
$= 1 - 0.2033$
$= 0.7967$

b) $P(900 < X < 300) = P\left(\frac{900 - 1250}{300} < X < \frac{1300 - 1250}{300}\right)$

$= P(-1.16 < Z < 0.17)$



$= P(Z < 0.17) - P(Z < -1.16)$
$= P(Z > 0.17) - P(Z > 1.16)$
$= 0.4325 - 0.1230$
$= 0.3096$
$= 30.96\%$

c) $P(X > x) = 0.15$
$P(Z > z) = 0.1492$
$P(Z > 1.04) = 0.1492$ gives the closest probability

$x = \mu + z\sigma$
$= 1250 + (1.04)(300)$
$= 1562$

d) 5 household are selected. Let X be the mean of the monthly electricity.

i- CLT cannot be used here as the value of n is not larger than 30.
$\mu = 1250$, $\sigma = 300$, $n = 5$

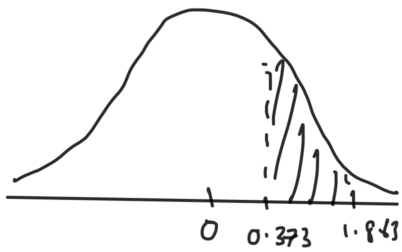
$\mu_{\bar{x}} = \mu = 1250$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{300}{\sqrt{5}} = 134.16$

$\bar{X} \sim N(1250, 134.16^2)$

ii. $P(1300 < \bar{x} < 1500) = P\left(\frac{1300-1250}{134.16} < z < \frac{1500-1250}{134.16}\right)$

$= P(0.373 < z < 1.863)$



$= P(z > 0.373) - P(z > 1.863)$

$= 0.3557 - 0.0314$

$= 0.3243$

(10)

a) Hypothesis :

$H_0 : \mu = 15$
$H_1 : \mu \neq 15$

two tailed test

b) Test statistic :

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{15.75 - 15.00}{2.4 / \sqrt{36}} = \frac{0.75}{0.4} = 1.875$$

conclusion :

$$\text{p-value} = 2 \times P(z > |1.875|)$$
$$= 2(0.0301)$$
$$= 0.0602$$

p-value $> \alpha = 0.05$

∴ Therefore, we do not have enough evidence to reject the restaurant's claim that it serve food to its customers, on average, within 15 minutes after the order is placed.

c) Hypotheses :

$H_0 : \mu_B = 15$ min
$H_1 : \mu_B < 15$ min

Test - statistic :

$$z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{15 - 15}{3 / \sqrt{50}} = \frac{0}{0.4243} = 0$$

p-value $= P(Z < z) = P(Z < 0) = 0.5$
Reject H_0 if p-value $< \alpha = 0.05$
In this case, p-value $> \alpha$.
We do not have enough evidence to reject H_0.
There is not enough evidence to support Restaurant B's claim that it serves food to customer faster than 15 minutes on average.