# Assignment 3 (Week 5 - 6) Solution

STAT 2601 - Business Statistics
School of Mathematics and Statistics, Carleton University

Total Marks: 34

## Q1: [10] Canadian Household Debt

$$X \sim N(\mu = \$41,500, \sigma = \$3,900)$$

(a)

$$\begin{aligned}
P(X > \$50,000)[0.5] &= P(\frac{X-\mu}{\sigma} > \frac{\$50,000-\mu}{\sigma}) \\
&= P(Z > \frac{\$50,000 - \$41,500}{\$3,900})[0.5] \\
&= P(Z > 2.18)[0.5] \\
&= 1 - 0.9854[0.5] \\
&= 0.0146.[0.5]
\end{aligned}$$

(b)

$$\begin{aligned}
P(\$35,000 < X < \$45,000)[0.5] &= P(\frac{\$35,000-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{\$45,000-\mu}{\sigma}) \\
&= P(\frac{\$35,000 - \$41,500}{\$3,900} < Z < \frac{\$45,000 - \$41,500}{\$3,900})[0.5] \\
&= P(-1.67 < Z < 0.90)[0.5] \\
&= 0.8159 - 0.0475[0.5] \\
&= 0.7684[0.5]
\end{aligned}$$

(c) Let $U$ represent the household debt of top 20% Canadian households in 2023.

$$\begin{aligned}
P(X > U) &= 0.20[0.5] \\
P(\frac{X-\mu}{\sigma} > \frac{U-\mu}{\sigma}) &= 0.20 \\
P(Z > \frac{U - \$41,500}{\$3,900}) &= 0.20[0.5] \\
P(Z > 0.84) &\approx 0.20 \,\text{(Using Standard Normal Table)}[0.5] \\
\therefore 0.84 &= \frac{U - \$41,500}{\$3,900}[0.5] \\
U &= \$44,776[0.5]
\end{aligned}$$

Alternatively,

$$
\begin{aligned}
P(X < U) &= 0.80\textcolor{red}{[0.5]} \\
P(\frac{X-\mu}{\sigma} < \frac{U-\mu}{\sigma}) &= 0.80 \\
P(Z < \frac{U - \$41,500}{\$3,900}) &= 0.80\textcolor{red}{[0.5]} \\
P(Z < 0.84) &\approx 0.80\,(\text{Using Standard Normal Table})\textcolor{red}{[0.5]} \\
\therefore 0.84 &= \frac{U - \$41,500}{\$3,900}\textcolor{red}{[0.5]} \\
U &= \$44,776\textcolor{red}{[0.5]}
\end{aligned}
$$

(d) Let $X_{65}$ represent the debt of a Canadian household excluding mortgages in 2023 that corresponds to $65^{th}$ percentile.

$$
\begin{aligned}
P(X < X_{65}) &= 0.65\textcolor{red}{[0.5]} \\
P(\frac{X-\mu}{\sigma} < \frac{X_{65}-\mu}{\sigma}) &= 0.65 \\
P(Z < \frac{X_{65} - \$41,500}{\$3,900}) &= 0.65\textcolor{red}{[0.5]} \\
P(Z < 0.39) &\approx 0.65\,(\text{Using Standard Normal Table})\textcolor{red}{[0.5]} \\
\therefore 0.39 &= \frac{X_{65} - \$41,500}{\$3,900}\textcolor{red}{[0.5]} \\
X_{65} &= \$43,021\textcolor{red}{[0.5]}
\end{aligned}
$$

Alternatively,

$$
\begin{aligned}
P(X > X_{65}) &= 0.35\textcolor{red}{[0.5]} \\
P(\frac{X-\mu}{\sigma} > \frac{X_{65}-\mu}{\sigma}) &= 0.35 \\
P(Z > \frac{X_{65} - \$41,500}{\$3,900}) &= 0.35\textcolor{red}{[0.5]} \\
P(Z > 0.39) &\approx 0.35\,(\text{Using Standard Normal Table})\textcolor{red}{[0.5]} \\
\therefore 0.39 &= \frac{X_{65} - \$41,500}{\$3,900}\textcolor{red}{[0.5]} \\
X_{65} &= \$43,021\textcolor{red}{[0.5]}
\end{aligned}
$$

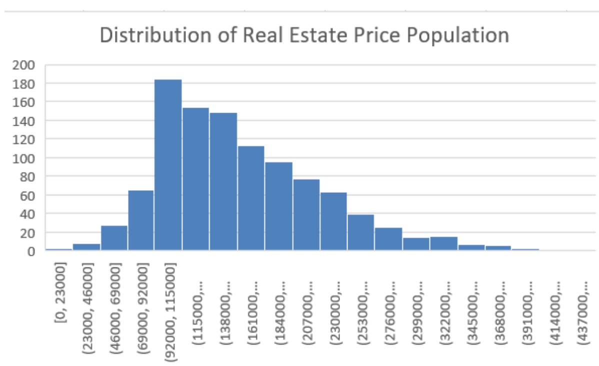
## Q2: [6] Real Estate

(a) The mean of real estate price population is $163,862.1251. [0.5]
EXCEL Output: [0.5]

| ummary Statistics of Real Estate Price | |
|---|---|
| | |
| Mean | 163862.1251 |
| Standard Error | 2090.761373 |
| Median | 151917 |
| Mode | 139079 |
| Standard Deviation | 67651.55892 |
| Sample Variance | 4576733424 |
| Kurtosis | 0.75980745 |
| Skewness | 0.876159911 |
| Range | 429578 |
| Minimum | 16858 |
| Maximum | 446436 |
| Sum | 171563645 |
| Count | 1047 |

(b) The distribution of real estate price population is right-skewed (or positively skewed [0.5] i.e., some house prices are exorbitant which are far away from typical house prices).

EXCEL Output: [0.5]




(c) Confidence Interval and Coverage of $\mu$:

| **Samples** | **Confidence Interval for $\mu$** [1] | **Covered $\mu$ (Y/N)?** [1] |
|---|---|---|
| Sample 1 | (141993.3, 191532.7) | Yes |
| Sample 2 | (127262.3, 179892) | Yes |
| Sample 3 | (112376.4, 152313.1) | No |
| Sample 4 | (128889.4, 166349.9) | Yes |
| Sample 5 | (151335.3, 215423.9) | Yes |
| Sample 6 | (125600.9, 170466) | Yes |
| Sample 7 | (152101.2, 201584.8) | Yes |
| Sample 8 | (129588.2, 176450.8) | Yes |
| Sample 9 | (154607.4, 199911) | Yes |
| Sample 10 | (119264.1, 157312.8) | No |
| Sample 11 | (126856.1, 173962.5) | Yes |
| Sample 12 | (133059.9, 173437) | Yes |
| Sample 13 | (146988, 197049) | Yes |
| Sample 14 | (143364.5, 196313.9) | Yes |
| Sample 15 | (129977.4, 179963.7) | Yes |
| Sample 16 | (123009, 173281.1) | Yes |
| Sample 17 | (130355.1, 181304.5) | Yes |
| Sample 18 | (141436.2, 200323.8) | Yes |
| Sample 19 | (138918.2, 188729.9) | Yes |
| Sample 20 | (139736.9, 196520.3) | Yes |

**Note:** The results will be different for each individual as the drawn samples are randomly chosen from real estate price population.

(d) [0.5] 18 out of 20 i.e., 90% confidence intervals contain population mean price of real estate $163,862.1251.[0.5] The empirical percentage (90%) does not match with confidence level (95%) specified in part (b) [0.5] because theoretically ($C_{30}^{1047}$) samples are possible at "without replacement" scheme and the underlying samples are only 20 of them [0.5]. In case of repeated sampling (i.e., in the long run), it is expected that 95% of the confidence intervals constructed in such manner will contain population mean price.

# Q3: [5] Online Ordering

Let

(i) Sample proportion of customers who favoured the online ordering system, $\hat{p} = \frac{x}{n} = \frac{85}{120} = 0.71$.

(ii) Sample proportion of customers who did not favour the online ordering system, $\hat{q} = 1 - \hat{p} = 1 - 0.71 = 0.29$.

(a) Determination of Sample Size:

$$\begin{aligned} n &= (\frac{z_{\alpha/2}\sqrt{\hat{p}\hat{q}}}{e})^2 [0.5] \\ &= (\frac{1.645\sqrt{(0.5)(0.5)}}{.02})^2 [0.5] \\ &= 1,691.27 [0.5] \\ &\approx 1,692 [0.5]. \end{aligned}$$

(b) 98% Confidence Interval for $p$:

$$\begin{aligned} &\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} [0.5] \\ &= 0.71 \pm 2.33\sqrt{\frac{(0.71)(0.29)}{120}} [0.5] \\ &= 0.71 \pm 0.10 \\ &= (0.61, 0.81) [0.5] \end{aligned}$$

(c) Assumptions and Conditions:

Assumption: The sample of 120 customers is a random sample. [0.5]

Conditions:

(i) $n\hat{p} = 120 \times 0.71 = 85 > 5$ [0.5]

(ii) $n\hat{q} = 120 \times 0.29 = 35 > 5$ [0.5]

$\hat{p} \sim N(p, \sqrt{\frac{pq}{n}})$

# Q4: [8] Airport Wait Time for Check In

(a) The sampling distribution of $\bar{X}$ follows Normal distribution [0.5] with mean, $\mu_{\bar{X}} = \mu = 5$ [0.5] and standard deviation, $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{35}}$ [0.5].

$$\bar{X} \sim N(5, \frac{2}{\sqrt{35}})$$

(b)

$$\begin{aligned} P(4 < \bar{X} < 5.5)[0.5] &= P(\frac{4-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{5.5-\mu}{\frac{\sigma}{\sqrt{n}}}) \\ &= P(\frac{4-5}{\frac{2}{\sqrt{35}}} < Z < \frac{5.5-5}{\frac{2}{\sqrt{35}}})[0.5] \\ &= P(-2.96 < Z < 1.48)[0.5] \\ &= 0.9306 - 0.0015[0.5] \\ &= 0.9291[0.5] \end{aligned}$$

(c) If the distribution of wait time is positively skewed, the answer in part (b) is still valid [0.5] because the distribution of sample means of wait time is approximately normally distributed [0.5] with mean $\mu = 5$ minutes and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{35}}$ minutes [0.5] using *Central Limit Theorem* [0.5] given the sample size 35 is large (i.e. $> 30$) [0.5].

(d) As the distribution of wait time is positively skewed, the sampling distribution of the average wait time of 15 randomly passengers cannot be determined [0.5] as the sample size 15 is less than 30 [0.5] and hence Central Limit Theorem (CLT) cannot be applied [0.5]. Therefore, the probability cannot be calculated.

# Q5: [5] Air Canada Flight Delay

(a) Determination of Sample Size:

$$\begin{aligned} n &= (\frac{z_{\alpha/2}\sigma}{e})^2[0.5] \\ &= (\frac{1.96 \times 10}{2})^2[0.5] \\ &= 96.04[0.5] \\ &\approx 97[0.5]. \end{aligned}$$

(b) 95% Confidence Interval for $\mu$:

$$\begin{aligned} &\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}[0.5] \\ &= 25 \pm 1.96\frac{10}{\sqrt{15}}[0.5] \\ &= 25 \pm 5.06 \\ &= (19.94, 30.06)[0.5] \end{aligned}$$

(c) Assumptions and Conditions:

Assumption: The sample of Air Canada flights in Ottawa is a random sample. [0.5]

Conditions: As the delay time of Air Canada flights follows a Normal distribution [0.5],

$$\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})[0.5]$$