**ST2334 (2021/2022 Semester 2) Solutions to Questions in Tutorial 10**

Question 1

If $\mu = 20$, then $\Pr(\bar{X} > 24) = \Pr\left(\frac{\bar{X}-20}{4.1/\sqrt{9}} > \frac{24-20}{4.1/\sqrt{9}}\right) = \Pr(T_8 > 2.9268) < 0.01$ since $\Pr(T_8 > 2.9268) < \Pr(T_8 > 2.8965)\ (= 0.01)$. Note $\Pr(T_8 > 2.9268) = 0.00955$ (from Excel: "=1-t.dist((24-20)/(4.1/sqrt(9)),8,TRUE)"; R: "1-pt((24-20)/(4.1/sqrt(9)),8)").

We conclude that $\mu > 20$. $\Pr(\bar{X} > 24 | \mu = 20)$ being small shows that it is very unlikely to get a mean of 24 if the population mean is really 20.

Question 2

(a) $\Pr(\bar{X}_B - \bar{X}_A \geq 0.2) = \Pr\left(Z > \frac{0.2}{\sqrt{1/36+1/36}}\right) = \Pr(Z > 0.8485) = 0.19807$ (from Excel: "=1-norm.dist(0.2/sqrt(2/36),0,1,TRUE)"; R: "1-pnorm(0.2/sqrt(2/36),0,1)").

Note: $\Pr\left(Z > \frac{0.2}{\sqrt{1/36+1/36}}\right) = \Pr(Z > 0.85) = 0.1977$ (from the normal table).

(b) Since the probability in part (a) is not small, therefore it is not unlikely to observe $\bar{X}_B - \bar{X}_A \geq 0.2$ when $\mu_A = \mu_B$.

Having the difference, $\bar{X}_B - \bar{X}_A$, more extreme than what we observed ($= 0.2$) is not unlikely ($\Pr(\bar{X}_B - \bar{X}_A \geq 0.2) = 0.19807$). Hence, we do not reject the statement $\mu_A = \mu_B$ or we believe that the conjecture that $\mu_A \neq \mu_B$ is likely not true.

Question 3

(a) $\Pr(S^2 > 9.1) = \Pr\left(\frac{(n-1)S^2}{\sigma^2} > \frac{(24)(9.1)}{6}\right) = \Pr(\chi^2_{24} > 36.4) = 0.05017$ (from Excel: "=1-chisq.dist(24*9.1/6,24,TRUE)"; R: "1-pchisq(24*9.1/6,24)")

Note: $\Pr(\chi^2_{24} > 36.4) \approx 0.05$ (from the $\chi^2$-table).

(b) $\Pr(3.462 < S^2 < 10.745) = \Pr\left(\frac{(24)(3.462)}{6} < \frac{(n-1)S^2}{\sigma^2} < \frac{(24)(10.745)}{6}\right) = \Pr(13.848 < \chi^2_{24} < 42.98) = 0.9900 - 0.0500 = 0.94$

Question 4

Since $\sigma_1^2$ and $\sigma_2^2$ are equal and the underlying distributions are normal, therefore $S_1^2/S_2^2$ follows an $F$ distribution with (7, 11) degrees of freedom. Hence $\Pr(S_1^2/S_2^2 < 4.89) = 0.99003$ (from Excel: "=f.dist(4.89,7,11,TRUE)"; R: "pf(4.89,7,11)")

Note: $\Pr(S_1^2/S_2^2 < 4.89) = 0.99$ (From $F$-table).

Question 5

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Mine 1: | 8260 | 8130 | 8350 | 8070 | 8340 | | $S_1^2 = 15750$ |
| Mine 2: | 7950 | 7890 | 7900 | 8140 | 7920 | 7840 | $S_2^2 = 10920$ |

$\Pr\left(\frac{S_1^2}{S_2^2} > \frac{15750}{10920}\right) = \Pr\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} > \frac{15750}{10920}\right) = \Pr(F_{4,5} > 1.4423) = 0.3436$, where $F_{4,5} \sim F(4,5)$.

We also find the probability that the ratio of two sample variances is less than what we observed ($= 1.4423$). That is, $\Pr\left(\frac{S_1^2}{S_2^2} < 1.4423\right) = 0.6564$.

Since under equal variances assumption, the probabilities of the ratio of the two sample variances is more extreme (larger or smaller) than what we observed ($s_1^2/s_2^2 = 1.4423$) are not small (i.e. 0.3436 and 0.6564), therefore, the equal variances assumption seems to be plausible. Hence, the two variances may be considered as equal.

[Remark: $\Pr(F_{4,5} > 1.4423) = 0.3436$ (from Excel: "=1-f.dist(15750/10920,4,5,TRUE)"; R: "1-pf(15750/10920,4,5)")].

Question 6

(a) $E(U) = E(X)/n = np/n = p$. Since $E(U) = p$, therefore $U$ is an unbiased estimator of $p$.

(b) $E(V) = \frac{E(X+n/2)}{3n/2} = \frac{np+n/2}{3n/2} = \frac{p+1/2}{3/2} = \frac{2p+1}{3} \neq p$ unless $p = 1$. Since $E(V) \neq p$, therefore $V$ is a biased estimator of $p$.

Question 7

$Y$ = helium porosity of a coal sample. $Y \sim N(\mu, \sigma^2)$.

(a) It is given that $\sigma = 0.75, n = 20$ and $\bar{y} = 4.85$. Hence a 95% confidence interval for $\mu$ is given by $\bar{Y} \pm z_{0.025}\frac{\sigma}{\sqrt{n}} = 4.85 \pm 1.96\frac{0.75}{\sqrt{20}} = 4.85 \pm 0.3287 = (4.5213, 5.1787)$.

(b) The length of a 95% confidence interval is $2\ z_{0.025}\frac{\sigma}{\sqrt{n}}$. Hence the length of 95% CI being 0.4 implies that $2(1.96)\frac{0.75}{\sqrt{n}} = 0.4$. Therefore $n = 54.0205 \approx 54$

(c) It is given that $S = 0.75,\ n = 20$ and $\bar{y} = 4.85$. Hence a 95% confidence interval for $\mu$ is given by $\bar{Y} \pm t_{19;\,0.025}\frac{S}{\sqrt{n}} = 4.85 \pm 2.093\frac{0.75}{\sqrt{20}} = 4.85 \pm 0.351 =$ $(4.4990, 5.2010)$.

Question 8

(a) 95% confidence interval for $\mu$ is given by $\bar{X} \pm z_{0.025}\frac{\sigma}{\sqrt{n}} = 0.310 \pm (1.96)\frac{0.0015}{\sqrt{75}} =$ $0.310 \pm 0.00034 = (0.30966, 0.31034)$

(b) $n \geq \left(\frac{z_{0.025}\,\sigma}{e}\right)^2 = \left(\frac{1.96\ x\ 0.0015}{0.0005}\right)^2 = (5.88)^2 = 34.573$. Take $n = 35$.

Question 9

A 90% confidence interval for $\mu$ is given by $\bar{X} \pm t_{11;0.05}\frac{s}{\sqrt{n}} = 48.50 \pm (1.7959)\frac{1.5}{\sqrt{12}} =$ $48.50 \pm 0.7776 = (47.7224, 49.2776)$

Question 10

A 94% confidence interval for $\mu_1 - \mu_2$ is given by $(\bar{X}_1 - \bar{X}_2) \pm z_{0.03}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = (80 - 75) \pm$ $(1.8807)\sqrt{\frac{5^2}{25} + \frac{3^2}{36}} = 5 \pm 2.1028 = (2.8977, 7.1028)$

Question 11

98% confidence interval for $\mu_1 - \mu_2$ is given by $(\bar{X}_1 - \bar{X}_2) \pm z_{0.01}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} =$

$(12.2 - 9.1) \pm (2.3263)\sqrt{\frac{1.1^2}{100} + \frac{0.9^2}{200}} = 3.1 \pm 0.2956 = (2.8044, 3.3956)$

Since the 98% confidence interval does not cover 0 and is in the positive range, the treatment appears to reduce the mean amount of metal removed.

Question 12

Since $E(Z^{2k+1}) = E[(-Z)^{2k+1}] = E[-Z^{2k+1}] = -E[Z^{2k+1}]$, therefore $E(Z^{2k+1}) = 0$.