Subject: PSDA Assignment 4 Date:

TEE HUI YOU (A19EC0170)
HAM JING YI (A19EC0048)
KHOR YONG XIN (A19EC0061)

Question 1

a. i) $n = 1100$

$\hat{p} = \frac{990}{1100}$

$= 0.9$

ii) $1-\alpha = 0.99$

$\alpha = 0.01$

$\frac{\alpha}{2} = 0.005$

$Z_{\alpha/2} = 2.575$

99% confidence interval, $p = 0.9 \pm 2.575 \left(\sqrt{\frac{0.9 \times 0.1}{1100}}\right)$

$= 0.9 \pm 2.575 (9.045 \times 10^{-3})$

$= 0.9 \pm 0.02329$

$= (0.8767, 0.9233)$

∴ It is 99% confident that the proportion of all drivers who have engaged in careless or aggressive driving in the previous 6 months is between 0.8767 and 0.9233.

b) $n = 16$, $\bar{x} = 18.6s$, $s = 9.486$, $df = 16-1 = 15$, confidence level = 80%

$t_{0.2, 15} = 1.341$

80% confidence interval $= \bar{x} \pm \left(t_{\text{critical value}} \sqrt{\frac{s^2}{n}}\right)$

$= 18.6 \pm 1.341 \sqrt{\frac{9.486^2}{16}}$

$= 18.6 \pm 1.341 (2.3715)$

$= 18.6 \pm 3.18$

$= (15.42, 21.78)s$

∴ It is 80% confident that mean time of men's 100-metre run of secondary school students lies between 15.42s and 21.78s.

Question 2

a) Hypothesis test: At least 30% of public is allergic to cheese products.

$H_0: p = 0.3$

$H_1: p > 0.3$

Type I error: The null hypothesis is rejected when it is actually true.

Subject : Date:

b) $\mu = 35$ minutes , $n = 20$, $\bar{x} = 33.1$, $s = 4.3$, σ is unknown

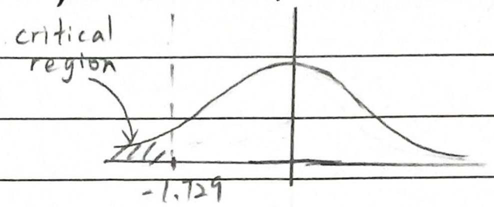
i) $H_0 : \mu = 35$

$H_1 : \mu < 35$

ii) $\alpha = 0.05$, $df = 19$

$t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ $t_{0.05,19} = -1.729$ (critical value)

$= \dfrac{33.1 - 35}{4.3/\sqrt{20}}$



$= -1.976$

$\therefore$ Since $t <$ critical value, t lies in the critical region. H_0 is rejected.
Hence, in 0.05 significance level, there is sufficient evidence to conclude that the average time taken for students in a university to complete a standardize test is less than 35 minutes.

c) $\bar{x} = \dfrac{3.73 + 4.37 + 3.73 + 4.33 + 3.39 + 2.47 + 4.47 + 3.43 + 3.68 + 4.68 + 3.52 + 3.02 + 4.09 + 4.13 + 3.22 + 2.54}{16}$

$= 3.675$

$s^2 = \dfrac{2(3.73 - 3.675)^2 + (4.37 - 3.675)^2 + (4.33 - 3.675)^2 + (3.39 - 3.675)^2 + (2.47 - 3.675)^2 + (4.47 - 3.675)^2 + (3.43 - 3.675)^2 + (3.68 - 3.675)^2 + (4.68 - 3.675)^2 + (3.52 - 3.675)^2 + (3.02 - 3.675)^2 + (4.09 - 3.675)^2 + (4.13 - 3.675)^2 + (3.22 - 3.675)^2 + (2.54 - 3.675)^2}{16 - 1}$

$= 0.432$

$\sigma = 0.470$, $\alpha = 0.05$

i) $H_0 : \sigma = 0.470$

$H_1 : \sigma \neq 0.470$

ii) $x^2 = \dfrac{(n-1)s^2}{\sigma^2}$

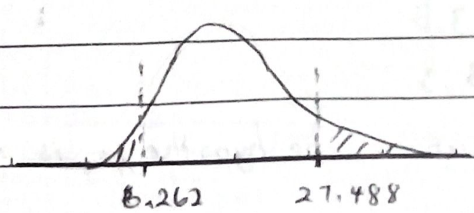
$= \dfrac{(16-1)(0.432)}{0.47^2}$

$= 29.33$

iii) $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$, $df = 16 - 1 = 15$

$x^2_{0.025,15} = 27.488$ } critical value

$x^2_{1-0.025,15} = 6.262$



Subject : Date:

iv) The $x^2 > 27.488$, x^2 lies in the critical region.
Hence, Ho is rejected.
Conclusion:
There is sufficient evidence to show that vitamin supplement appear to affect the variation among birth weights.

Question 3

a) $\hat{p} = \frac{130}{350}$
$= 0.3714$

$1-\alpha = 0.95$
$\alpha = 0.05$
$\frac{\alpha}{2} = 0.025$
$z_{\alpha/2} = 1.96$

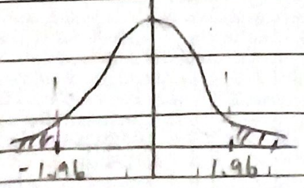
95 % confidence interval $= 0.3714 \pm 1.96 \left(\sqrt{\frac{0.3714(1-0.3714)}{350}}\right)$
$= 0.3714 \pm 1.96 (0.0258)$
$= 0.3714 \pm 0.0506$
$= (0.3208, 0.4220)$

∴ Hence, we are 95% confident that proportion of all food shops with special offers lies between 0.3208 and 0.4220.

b) margin error $= 0.04$, $n = ?$
$0.04 = 1.96 \sqrt{\frac{0.3714(1-0.3714)}{n}}$
$n = 560.54$
$n \approx 561$

Question 4

i) $\bar{x} = 58392$, $\sigma = 648$, $n = 6$, variance known
$H_0 : \mu = 58000$
$H_1 : \mu \neq 58000$
$\alpha = 0.05$
$\frac{\alpha}{2} = 0.025$
$z_{\alpha/2} = 1.96$ (critical value)



$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
$= \frac{58392 - 58000}{\frac{648}{\sqrt{6}}}$
$= 1.482$ $(-1.96 < z < 1.96)$

Test statistic does not lies in critical region
Null hypothesis is fail to be rejected.
Conclusion: There is sufficient evidence to conclude that average compressive strength of steel is 58000 psi.

Subject : Date:

b) $n = 101$, $\sigma^2 = 0.18$, $s^2 = 0.13$

$H_0 : \sigma^2 = 0.18$

$H_1 : \sigma^2 \neq 0.18$

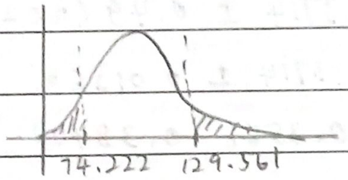
$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$df = 100$

$\chi^2_{0.025, 100} = 129.561$

$\chi^2_{1-0.025, 100} = 74.222$



test statistic, χ^2

$= \frac{(n-1)s^2}{\sigma^2}$

$= \frac{(100)(0.13)}{0.18}$

$= 72.22$ (< 74.222 ...)

Test statistic lies in the critical region

Hence, null hypothesis is rejected.

Conclusion:

There is sufficient evidence to conclude that the inspector is not making satisfactory measurements.

Question 5

a) $\hat{p} = \frac{280}{418}$

$= 0.6699$

$1-\alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$z_{\alpha/2} = 1.96$

95% confidence interval $= 0.6699 \pm 1.96 \sqrt{\frac{0.6699(1-0.6699)}{418}}$

$= 0.6699 \pm 0.04508$

$= (0.625, 0.715)$

∴ It is 95% confident that the proportion of all insured patients who blame rising insurance rates on large court settlements against doctors lies between 0.625 and 0.715.

b) $n = 126$

$\bar{x} = 29.2$

$\mu = 18.2$

$\sigma = 7.5$

$1-\alpha = 0.9$

$\frac{\alpha}{2} = 0.05$

$z_{\alpha/2} = 1.645$

Since varience is known,

90% confidence interval $= 29.2 \pm 1.645 \left(\frac{7.5}{\sqrt{126}}\right)$

$= 29.2 \pm 1.1$

$= (28.1, 30.3)$

∴ It is 90% confident that the average blood lead level is between 28.1 ug/dl and 30.3 ug/dl.

Subject :

Date:

Question 6

a) $n = 64$, $\bar{x} = 38$, $s = 5.8$, σ unknown

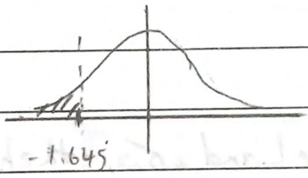
$H_0 : \mu = 40$

$H_1 : \mu < 40$

$\alpha = 0.05$

Since $n > 30$, critical value:

$z_\alpha = 1.645$



$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= \frac{38 - 40}{5.8/\sqrt{64}}$$

$$= -2.759$$

z < critical value, z lies in critical region.

Null hypothesis is rejected.

Conclusion:

There is sufficient evidence to believe that $\mu < 40$.

b) $n = 10$

$$\bar{x} = \frac{10.2 + 9.7 + 10.1 + 10.3 + 10.1 + 9.8 + 9.9 + 10.4 + 10.3 + 9.8}{10}$$

$= 10.06$

$$s = \sqrt{\frac{(10.2-10.06)^2 + (9.7-10.06)^2 + 2(10.1-10.06)^2 + 2(10.3-10.06)^2 + 2(9.8-10.06)^2 + (9.9-10.06)^2 + (10.4-10.06)^2}{10-1}}$$

$= 0.2459$

$H_0 : \mu = 10$

$H_1 : \mu \neq 10$

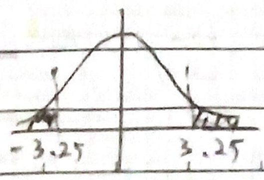
$\alpha = 0.01$

$\frac{\alpha}{2} = 0.005$

Since $n < 30$, critical value:

$df = 9$

$t_{0.005, 9} = 3.25$



$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= \frac{10.06 - 10}{0.2459/\sqrt{10}}$$

$= 0.7716$ $(-3.25 < t < 3.25)$

t does not lies in critical region.

Hence, null hypothesis is fail to reject

Conclusion:

There is sufficient evidence to conclude that the average content of particular lubricant is 10 litres.

Subject :   Date:

Question 7

a) Sample 1 : (50.3, 52.7) } 95% confidence, $s_1 = s_2$, variance unknown
Sample 2 : (49.8, 50.6)

margin error of sample 1 $= \frac{52.7 - 50.3}{2}$
$= 1.2$

margin error of sample 2 $= \frac{50.6 - 49.8}{2}$
$= 0.4$

From formulae, $\downarrow$margin error $= \downarrow t_{critical\ value} \left(\frac{s}{\sqrt{n}\uparrow}\right)$

When n increase, $t_{critical\ value}$ and $\frac{s}{\sqrt{n}}$ decrease. This will causes the margin error to decrease.

Hence, we can conclude that sample 2 has larger sample size compared to sample 1 as it has smaller margin error.

b) I will recommend Statistic II.

First, Statistic II and III are better compared to Statistic I as the true value of proportion characteristic lies at the centre of the curve. However, the datas of Statistic II are more concentrated at the centre of curve as shown by height of centre part of curve. This shows that Statistic II would have smaller area of critical region at the same significance level compared to Statistic III. Besides that, Statistic II would have smaller standard deviation compared to Statistic III which means Statistic II would have smaller estimation error and standard error. Statistic II would be less likely to be rejected if the statistic is being tested.

c) $n = 500$
$\hat{p} = 0.15$
$\alpha = 0.1$
$z_{\alpha/2} = 1.645$

90% confidence interval $= 0.15 \pm 1.645\sqrt{\frac{0.15(1-0.15}{500}}$
$= 0.15 \pm 0.0263$
$= (0.1237, 0.1763)$

$\therefore$ It is 90% confident that the market share for its new brand is between 0.1237 and 0.1763.