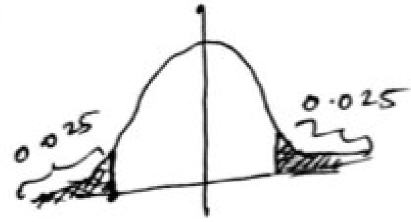
ASSIGNMENT 1 : STATISTICS : SOLUTIONS

3.1 $n = 6$; $s = 2.5\%$; $\bar{x} = 14.1\%$; $CI = 95\%$

Problem says "assume normal distribution".



$\alpha = 5\%$ ie: 0.05

$\alpha/2 = 0.025$ & $(1-\alpha/2) = 0.975$

$Z_{0.025} = \text{qnorm}(0.025) = -1.96$

$Z_{0.975} = \text{qnorm}(0.975) = 1.96$

$z = \dfrac{(x-\bar{x})}{\sigma/\sqrt{n}}$ $\therefore$ $\boxed{x = \bar{x} + z \cdot \dfrac{\sigma}{\sqrt{n}}}$

$\therefore x_L = 14.1 - 1.96 \times \dfrac{2.5}{\sqrt{6}} = 14.1 - 2 = 12.1$

$x_H = 14.1 + 1.96 \times \dfrac{2.5}{\sqrt{6}} = 14.1 + 2 = 16.1$

$\therefore$ Interval for % of copper (@ 95% CI) = [12.1, 16.1]

3.2 $n = 25$; $\bar{x} = 300007$; $S = 10$; C.I 95%

dof $= (25-1) = 24$ (since $n < 30$, t-distribution assumed)

$x_L = 300007 + t_{0.025,24} \times \dfrac{10}{\sqrt{25}} = 300007 - 2.0639 \times 2$

$x_H = 300007 + t_{0.975,24} \times \dfrac{10}{\sqrt{25}} = 300007 + 2.0639 \times 2$

$x_L = 300002.8 \sim 300003$

$x_H = 300011.13 \sim 300011$

$\therefore$ 95% CI for speed of light [300003, 300011]

3.3 $n = 25$; $\bar{x} = 0.910835$; $s = 0.000045$

dof $= (25-1) = 24$ (t-distribution)

$x_L = 0.910835 + t_{0.025,24} \times \dfrac{45 \times 10^{-6}}{\sqrt{25}} = \cancel{0.910835}$

$x_H = 0.910835 + t_{0.975,24} \times \dfrac{45 \times 10^{-6}}{\sqrt{25}}$

$x_L = 0.910835 - 2.0639 \times 9 \times 10^{-6}$

$x_H = 0.910835 + 2.0639 \times 9 \times 10^{-6}$

C. Interval @ 95% = [0.9108164, 0.9108536]

3.4 $n = 10$; $dof = 9$; $\bar{x} = 1.0002$; $s = 0.0001$.

90% CI required for exact length.

$\alpha = 10\% = 0.1$; $\alpha/2 = 0.05$; $(1-\alpha/2) = 0.95$

$\therefore x_L = 1.0002 + t_{0.05,9} \times \frac{0.0001}{\sqrt{10}} = 1.0002 - 1.833 \times 3.16 \times 10^{-5}$

$x_H = 1.0002 + t_{0.95,9} \times \frac{0.0001}{\sqrt{10}} = 1.0002 + 1.833 \times 3.16 \times 10^{-5}$

$x_L = 1.0002 - 5.8 \times 10^{-5}$

$x_H = 1.0002 + 5.8 \times 10^{-5}$

Interval $= [1.00014, 1.00026]$

3.5 $x = (9.8, 10.2, 10.4, 9.8, 10.0, 10.2, 9.6)$

CI 85%; assume normal dist

$\bar{x} = 10$; $s = 0.2828427$;

$\alpha = 15\% = 0.15$; $\alpha/2 = 0.075$; $(1-\alpha/2) = 0.925$

$x_L = 10 + z_{0.075} \times \frac{0.283}{\sqrt{7}} = 10 - 1.44 \times 0.107 = 10 - 0.154$

$x_H = 10 + z_{0.925} \times \frac{0.283}{\sqrt{7}} = 10 + 1.44 \times 0.107 = 10 + 0.154$

Interval $= [9.846, 10.154]$

3.6 Difference between sample and true mean needs to be ~~Between~~ 15s.

$\therefore (x - \mu) = 15$

$\sigma = 40$; CI 95%; $\alpha = 0.025$

$z = \frac{x-\mu}{s}$. $\therefore s = \frac{x-\mu}{z} = \frac{15}{1.96} = 7.653$

Now $s = \frac{\sigma}{\sqrt{n}}$ $\therefore n = \left(\frac{\sigma}{s}\right)^2 = \left(\frac{40}{7.653}\right)^2 = 27.32 \sim 28$ (rounded)

Sample size should be at least 28.

3.7 $n = 8$; $\bar{x} = 1806$; $s = 2.4$; CI 99%.

$\alpha = 1\%$; $\alpha/2 = 0.005$; $(1-\alpha/2) = 0.995$

dof $= (8-1) = 7$

$$x_L = 1806 + t_{0.005,7} \times \frac{2.4}{\sqrt{8}} = 1806 - 3.5 \times 0.848$$

$$x_H = 1806 + t_{0.995,7} \times \frac{2.4}{\sqrt{8}} = 1806 + 3.5 \times 0.848$$

$x_L = 1806 - 2.968 = 1803$

$x_H = 1806 + 2.968 = 1808.968 \sim 1809$

Interval $= [1803, 1809]$

3.8 $n = 100$; $\bar{x} = 150$; $s = 40$; CI 95%

$\alpha = 5\% = 0.05$; $\alpha/2 = 0.025$; $1-\alpha/2 = 0.975$

Distribution : NORMAL.

$$x_1 = 150 + Z_{0.025} \times \frac{40}{\sqrt{100}} = 150 - 1.96 \times 4 = 142.16.$$

$$x_2 = 150 + Z_{0.975} \times \frac{40}{\sqrt{100}} = 150 + 1.96 \times 4 = 157.84$$

3.9 Chance $= 1/20 \Rightarrow 0.05$, 5% error. acceptable

$x - \bar{x} = 50$ $\quad \sigma = 40$ $\quad n = ?$

$$Z_{0.975} = \frac{x - \bar{x}}{\sigma_{\bar{x}}} = \frac{x - \bar{x}}{\sigma/\sqrt{n}} = \frac{50}{40/\sqrt{n}}$$

$$\therefore \sqrt{n} = \frac{Z_{0.975} \times 40}{50} = \frac{1.96 \times 40}{50} = 1.568$$

$\therefore n = 2.46 \rightarrow$ rounded off to 3

$\therefore \boxed{n = 3}$

3.10 $x = \{0.8, 1.3, 1.5, 1.7, 1.7, 1.8, 2.0, 2.0, 2.2\}$

$n = 9$; $dof = (9-1) = 8$; $95\%\ CI \Rightarrow \alpha/2 = 0.025$; $1-\alpha/2 = 0.975$

$\bar{x} = 1.667 \quad s = 0.4243$

$\bar{x}_L = 1.667 + t_{0.025,8} \times \frac{0.4243}{\sqrt{9}} \qquad t_{0.025,8} = -2.306$

$\bar{x}_H = 1.667 + t_{0.975,8} \times \frac{0.4243}{\sqrt{9}} \qquad t_{0.975,8} = 2.306$

Interval $[1.3408, 1.993]$

3.11 $n = 20$; $s^2 = 53$; normal distribution; 99% CI

$\alpha = 1\% = 0.01$; $\alpha/2 = 0.005$; $(1-\alpha/2) = 0.995$

The test statistic to be used for variances is : $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

$\therefore$ the interval is determined by calculating σ^2 for $\chi^2_{\alpha/2}$ and $\chi^2_{1-\alpha/2}$

$\sigma^2_{0.005} = \frac{(20-1)\times 53}{\chi^2_{0.005}} = \frac{(20-1)\times 53}{\text{qchisq}(0.005, 19)} = \frac{19\times 53}{6.844} = 147.136$

$\sigma^2_{0.995} = \frac{19\times 53}{\chi^2_{0.995}} = \frac{19\times 53}{38.58226} = 26.1$

$\therefore \quad 26.1 < \sigma^2 < 147.136$

~~304~~ NOTE: In case of χ^2 statistic, 'α' is usually taken from the right side, in which case, α and $1-\alpha$ will correspond with $(1-\alpha)$ and α in "R".

3.14 $n = 16$; $X = \{4.260, \ldots\}$; 95% CI $\rightarrow$ $\alpha = 5\% = 0.05$, $\alpha/2 = 0.025$, $1-\alpha/2 = 0.975$

$\bar{x} = 4.371 \quad s^2 = 0.1175$

$\sigma^2_{0.025} = \frac{(16-1)\times(0.1175)}{\chi^2_{0.025,15}} = \frac{15\times 0.1175}{6.262}$

$\sigma^2_{0.975} = \frac{(16-1)\times(0.1175)}{\chi^2_{0.975,15}} = \frac{15\times 0.1175}{27.488}$

Interval $[0.064, 0.281]$

3.15

$n = 10$; 95% CI ; $x = \{16.9, \ldots\}$

dof $= (10-1) = 9$; $\bar{x} = 16.12$; $s^2 = 0.2862$

$\alpha = 5\% = 0.05$; $\alpha/2 = 0.025$; $(1-\alpha) = 0.975$

$$\cancel{\chi^2_{0.975,9}} \quad \frac{(10-1) \times 0.2862}{\chi^2_{0.975,9}} < \sigma^2 < \frac{(10-1) \times 0.2862}{\chi^2_{0.025,9}}$$

$$\frac{9 \times 0.2862}{19.023} < \sigma^2 < \frac{9 \times 0.2862}{2.7}$$

$$\underline{0.135 < \sigma^2 < 0.954}$$

3.16

$n = 17$; $x = \{16, 22, \ldots\}$; 90% CI

$\bar{x} = 20$; $s^2 = 15.875$

dof $= (17-1) = 16$; $\alpha = 10\% = 0.1$; $\alpha/2 = 0.05$; $(1-\alpha/2) = 0.95$

$$\therefore \frac{(17-1) \times 15.875}{\chi^2_{0.95,16}} < \sigma^2 < \frac{(17-1) \times 15.875}{\chi^2_{0.05,16}}$$

$$\frac{16 \times 15.875}{26.296} < \sigma^2 < \frac{16 \times 15.875}{7.961}$$

$$\underline{9.66 < \sigma^2 < 31.90}$$

3.17

$n = 31$; $s = 10.2$; $s^2 = (10.2)^2$; 98% CI

$\alpha = 2\% = 0.02$; $\alpha/2 = 0.01$; $(1-\alpha/2) = 0.99$

dof $= 30$;

$$\therefore \frac{(31-1) \times (10.2)^2}{\chi^2_{0.99,30}} < \sigma^2 < \frac{(31-1) \times (10.2)^2}{\chi^2_{0.01,30}}$$

$$\frac{30 \times (10.2)^2}{50.892} < \sigma^2 < \frac{30 \times (10.2)^2}{14.953}$$

$$\therefore \sqrt{61.33} < \sigma < \sqrt{208.734}$$

$$\underline{7.83 < \sigma < 14.44}$$

3.18 $p = 5/50 = 0.1$ (ie: 10% of terminals give incorrect response)

a) If the firm has 800 terminals, estimated terminal with incorrect response will be 10% of 800 ie: $0.1 \times 800 = \underline{80}$

b) In case of proportions, the distribution is NORMAL and the test statistic is given by $z = \dfrac{\hat{p} - p}{\sqrt{p(1-p)/n}}$

and the Confidence interval is given by:

$$\hat{p} - z\sqrt{p(1-p)/n} < p < \hat{p} - z\cdot\sqrt{p(1-p)/n}$$

In our case, we are assuming $\hat{p}$ to be same as $p = 0.1$
Given this, the confidence interval will be calculated as follows

~~ad~~ Assume C.I $= 95\%$

$\therefore \alpha = 0.05$; $\alpha/2 = 0.025$; $(1-\alpha/2) = \cancel{0.95}\ 0.975$ | $z_{0.025} = -1.96$, $z_{0.975} = 1.96$.

$$\therefore\ 0.1 - 1.96\sqrt{0.1\times0.9/800} < p < 0.1 + 1.96\sqrt{0.1\times0.9/800}$$

$$0.1 - 0.0201 < p < 0.1 + 0.0201$$

$$0.0799 < p < \cancel{12.01}\ 0.1201$$

$$\sim \quad \underline{8\% < p < 12\%}$$

3.19 what is n?
Given: Need to estimate probability within $\underline{0.04}$ $(\hat{p}-p)$ & with 95% conf.

a) No idea of percent defective
→ If we have no idea, then we have to assume worst case situation that good and bad cases are equally likely. ie: $p = 50\% = \underline{0.5}$

For 95% confidence, $z = 1.96$

$$\therefore\ z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \Rightarrow \left[\frac{z}{\hat{p}-p}\right]^2 = \frac{n}{p(1-p)}$$

$$\therefore\ n = p(1-p)\cdot\left[\frac{z}{\hat{p}-p}\right]^2 = 0.5\times0.5\times\left[\frac{1.96}{0.04}\right]^2$$

$\boxed{n = 600.25 \sim 601}$

b) Percent defective max $= 6\%$
Since max defects are 6%,
$p = 0.06$

$$n = p(1-p)\cdot\left[\frac{z}{\hat{p}-p}\right]^2$$

$$= 0.06\times0.94\cdot\left[\frac{1.96}{0.04}\right]^2$$

$$= 135.41$$

$\boxed{n \sim 136}$

3.20 Problem: Estimate books with publication date 1970 or earlier.

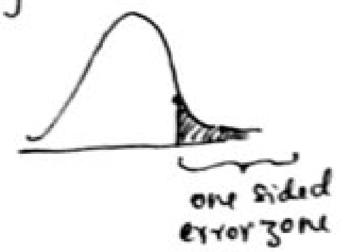
what should be 'n' for 90% confidence such that,

$$(\hat{p} - p) = 0.05$$

Since we are concerned with "1970 or earlier", the region of interest is "one-sided".

$\therefore$ We have to apply our "permissible error" to only one side

$\therefore$ $\alpha = 10\% = 0.1$ is applicable only on the right side.



$\Rightarrow$ permissible area $= (1-\alpha) = 0.90$

$$Z_{0.90} = 1.281$$

$$\therefore \quad 1.281 = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}}$$

$$= \frac{0.05}{\sqrt{0.5 \times (1-0.5)/n}}$$

Since we have to assume that every book sampled has an equal chance of being before 1970, as after 1970 $\Rightarrow p = 0.5$

$$\therefore \quad \left[\frac{1.281}{0.05}\right]^2 \times \sqrt{0.5 \times 0.5} = n$$

$$\therefore \quad n = 164.1 \sim 165$$

$\therefore$ 165 books have to be sampled.