Homework - 3

GOUTAMI PADMANABHAN
UTA ID: 1001669338

9.7 GIVEN: $n = 100$, $\bar{x} = 37.7$, $\sigma = 9.2$,

a) 90% confidence interval for expectation of the number of concurrent users is

$$\bar{x} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$1-\alpha = 0.9$, $\alpha = 0.1$, $\frac{\alpha}{2} = \frac{0.1}{2} = 0.05$,

$z_{0.05} = 1.645$

$$\bar{x} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} = 37.7 \pm 1.645 \times \frac{9.2}{\sqrt{100}}$$

$$= 37.7 \pm 1.5134$$

$$\cancel{= [36.2134, 39.}$$

$$= [36.1866, 39.2134]$$

b) 1% significance level $\Rightarrow \alpha = 0.01$

$H_0: \mu = 35$

$H_A: \mu > 35$

Step 1: Test statistic is $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

$$z = \frac{37.7 - 35}{9.2/\sqrt{100}} = 2.9348$$

Step 2: $z_\alpha = z_{0.01} = 2.326$

Step 3: Here $z > z_\alpha$, So we reject H_0 and accept H_A

Yes, there is sufficient evidence

9.8 a) $\sigma = 5$ minutes

$n = 64$

$\bar{x} = 42$ minutes

95% confidence interval $\Rightarrow 1-\alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$Z_{\alpha/2} = Z_{0.025} = 1.960$

95% confidence interval for population mean is

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 42 \pm \left[1.960 \times \frac{5}{\sqrt{64}}\right]$$

$$= 42 \pm 1.2275$$

$$= [40.7725, 43.2275]$$

b) $P(40.7725 < X < 43.2275)$

$$= P\left(\frac{40.7725-40}{5} < Z < \frac{43.2275-40}{5}\right)$$

$$= P(0.151 < Z < 0.645)$$

$$= 0.7405 - 0.5600$$

$$= 0.1805$$

40 minutes lies in the interval.

The null hypothesis H_0 is accepted.

Scanned by CamScanner

9.10 $n = 200$, defective $= 24$

$\hat{p} = \frac{24}{200} = 0.12$

a) 96% confidence interval $\Rightarrow 1-\alpha = 0.96$,

$\alpha = 0.04$, $\frac{\alpha}{2} = \frac{0.04}{2} = 0.02$

$Z_{0.02} = 2.0537$

96% confidence interval for proportion is

$\hat{p} \pm Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$= 0.12 \pm 2.0537 \sqrt{\frac{0.12(1-0.12)}{200}}$

$= 0.12 \pm 0.04719$

$= [0.0728, 0.16719] = [0.073, 0.167]$

b) $p_D = \frac{1}{10} = 0.1$

$H_0 : p_D = 0.1$

$H_A : p_D > 0.1$ $\quad z = (\hat{p} - p_D) \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Test statistic is $Z = (0.12 - 0.1) \sqrt{\frac{0.12(1-0.12)}{200}}$

$= 0.8704$

The critical values from table of Normal distribution are

$Z_{0.04} = 1.75$, $Z_{0.15} = 1.04$, $Z_{0.20} = 0.84$

Therefore, we do not have significant evidence at 4% and 15% levels to disprove the manufacturer's claim.

9.11) $\hat{p}_A = \frac{24}{200} = 0.12$, $\hat{p}_B = \frac{13}{150} = 0.087$

$H_0 : \hat{p}_A = \hat{p}_B$

$H_A : \hat{p}_A < \hat{p}_B$

The pooled sample proportion is

$$\hat{p} = \frac{\hat{p}_A \times n + \hat{p}_B m}{n+m} = \frac{0.12 \times 200 + 0.087 \times 150}{200 + 150}$$

$\hat{p} = 0.1059$

$$Var(\hat{p}_A - \hat{p}_B) = \hat{p}(1-\hat{p})\left(\frac{1}{n}+\frac{1}{m}\right)$$

$$\cancel{= 0.1059 \times (1-0.1059)}$$

$$= (0.1059) \times (1-0.1059)\left(\frac{1}{200}+\frac{1}{150}\right)$$

$$= 0.001104$$

$$Z = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{Var}} = \frac{0.12 - 0.087}{0.0332} = 0.99$$

P-value $= P(Z < 0.99) = 0.8389$

At 5% significance level, p-value > significance level.

We should not reject H_0. So, there is no sufficient evidence to ~~[illegible]~~ conclude that the quality of items produced by new supplier is higher.

9.17) $\hat{p}_1 = \frac{45}{100}$

$\hat{p}_2 = \frac{35}{100}$

$p_1 - p_2 = \frac{45}{100} - \frac{35}{100} = \frac{10}{100}$

$n = 900$

$(1-\alpha)100 = 95$, $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$, $Z_{\alpha/2} = Z_{0.025} = 1.96$

Margin of error for p_1,

$$Z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n}} = 1.96\sqrt{\frac{0.45 \times 0.55}{900}}$$

$$= 0.0325$$

Margin of error for p_2,

$$Z_{0.025}\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n}} = 1.96\sqrt{\frac{0.35 \times 0.65}{900}} = 0.0312$$

Margin of error for $p_1 - p_2$

$$Z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

$$= 1.96\sqrt{\frac{0.45 \times 0.55}{900} + \frac{0.35 \times 0.65}{900}}$$

$$= 0.045$$

9.18) Sample before change of firewall settings

56, 47, 49, 37, 38, 60, 50, 43, 43, 59, 50, 56, 54, 58

Sample after change of firewall settings

53, 21, 32, 49, 45, 38, 44, 33, 32, 43, 53, 46, 36, 48, 39, 35, 37, 36, 39, 45

$1-\alpha = 0.95$, $\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$, $t_{0.025} = 2.306$

a) $\bar{x} = \frac{700}{14} = 50$, $n = 14$, $\bar{y} = \frac{804}{20} = 40.2$, $m = 20$

$$s_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = 7.6158$$

$$s_y^2 = \frac{\sum_{i=1}^{m} (y_i - \bar{y})^2}{m-1} = 7.9578$$

$$s_p^2 = \frac{(n-1)\, s_x^2 + (m-1)\, s_y^2}{n+m-2} = \frac{13 \times 7.6158^2 + 19 \times 7.9578^2}{14+20-2} = \frac{1957.24}{32}$$

$s_p^2 = 61.1628$, $s_p = 7.8206$

Confidence interval for difference of mean with equal variances is

$$\bar{x} - \bar{y} \pm t_{\alpha/2}\, s_p \sqrt{\frac{1}{n} + \frac{1}{m}} = 50 - 40.2 \pm 2.306 \times 7.8206 \sqrt{\frac{1}{14} + \frac{1}{20}}$$

$$= \underline{[4.249, 15.351]}$$

b) Since the value '0' is not included in the CI range, we can state that there is a significant difference between the before and after firewall installation.

Scanned by CamScanner

For equal variance:

Test statistic is $t = \dfrac{\bar{x} - \bar{y} - D}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$

$$t = \frac{50 - 40.2}{7.8206\sqrt{\frac{1}{14} + \frac{1}{20}}} = 3.6$$

$H_0 : \theta = \theta_0$

$H_A : \theta < \theta_0$

$P = P\{t \le t_{obs}\} = F_v(t_{obs}), \quad = P(t > 3.6)$

$P =$ Between 0.0005 and 0.001

P-value is very small (< 0.01). So reject H_0 the null hypothesis and accept H_A

For unequal variance:

Test statistic $t = \dfrac{\bar{x} - \bar{y} - D}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$

$$t = \frac{50 - 40.2}{\sqrt{\frac{7.6158^2}{14} + \frac{7.9578^2}{20}}} = 3.6$$

P-value

$H_0 : \theta = \theta_0$

$H_A : \theta < \theta_0$

$P = P\{t \le t_{obs}\} = F_v(t_{obs})$

$P = 0.0005$

Changing to unequal variance has no effect on results

P-value is very small (< 0.01), So reject H_0 and accept H_A

Scanned by CamScanner

9.20) Level of significance $\alpha = 0.02$

$\sigma = 5,$

$s = 6.2$

$n = 40$

$H_0 : \sigma = 5$

$H_A : \sigma \neq 5$

Test statistic is $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$

$= \dfrac{(40-1)(6.2)^2}{5^2}$

$= 59.96$

$d.f = n-1 = 40-1 = 39$

By Two-sided tail test, critical value is

$\chi^2_{0.01} = 72.1$

$\chi^2_{0.999} = 17.3$

Here, χ^2_{obs} lies between the critical values

So we accept H_0 the null hypothesis

At 2% level of significance, there is significant evidence that standard deviation is 5 minutes

9.23 a) $\bar{x} = 85$, $n = 6$, $s_x^2 = 162.8176$

$\bar{y} = 80$, $m = 6$, $s_y^2 = 10.3684$

$s_x = 12.76$ $n-1 = 5$

$m-1 = 5$

$s_y = 3.22$

$H_0: \mu_x = \mu_y$

$H_A: \mu_x > \mu_y$

comparing variances,

Test statistic is $F_{obs} = \frac{s_x^2}{s_y^2} = 15.65$

$P = 2 \min (P(F \geq F_{obs}), P(F \leq F_{obs}))$

= in between 0.02 and 0.01

$\sigma_x \neq \sigma_y$

Using satterthwaite approximation,

$$t_{obs} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = \frac{85 - 80}{\sqrt{\frac{12.76^2}{6} + \frac{3.22^2}{6}}}$$

$t_{obs} = 0.93$

Degrees of freedom $\nu = \frac{\left(\frac{s_x^2}{m} + \frac{s_y^2}{n}\right)^2}{\frac{s_x^4}{n^2(n-1)} + \frac{s_y^4}{m^2(m-1}}$

Scanned by CamScanner

$$\nu = \frac{\frac{12.76^2}{6} + \frac{3.22^2}{6}}{\frac{12.76^4}{180} + \frac{3.22^4}{180}}$$

$\nu = 5.64$

From table A_5, p-value for $t_{obs} = 0.93$

$$P = P(t > t_{obs}) > 0.10$$

There is no evidence that Anthony's average grade is higher than Eric's grade

b) $H_0 : \sigma_x = \sigma_y$

$H_A : \sigma_x > \sigma_y$

$F_{obs} = 15.65$ (already solved)

From table A7, with 5 degrees of freedom, we get

p-value $= P(F \geq F_{obs})$. it is in the range $(0.01, 0.05)$

There is significant evidence that $\sigma_x > \sigma_y$ and Eric's claim that his grades are stable. Eric's claim is accepted.

Q9) First dice - unbiased.

(i) Average amount of information (in bits) encoded by each dice:

$H = -p \log_2 p$

First dice : $p = \frac{1}{6}$, unbiased fair dice

$\Rightarrow H = -6 \times \frac{1}{6} \log_2(1/6) = -(-2.584)$

$= \underline{2.584 \text{ bits}} \simeq 3 \text{ bits}$

Second dice:

$H = -\frac{1}{9}\log_2\left(\frac{1}{9}\right) - \frac{1}{9}\log_2(1/9) - \frac{1}{9}\log_2(1/9) - \frac{2}{9}\log_2(2/9)$

$- \frac{2}{9}\log_2(2/9) - \frac{2}{9}\cdot\log_2(2/9)$

$= -\frac{3}{9}\log_2(1/9) - 3 \times \frac{2}{9}\log_2(2/9)$

$= -\frac{3}{9}(-3.17) - \frac{6}{9}(-2.17)$

$= 1.0566 + 1.4466$

$= \underline{2.5032 \text{ bits}} \simeq 3 \text{ bits}$

Third dice:

$H = -\frac{1}{9}\log_2\left(\frac{1}{9}\right) - \frac{1}{9}\log_2\left(\frac{1}{9}\right) - \frac{2}{9}\log_2\left(\frac{2}{9}\right) - \frac{2}{9}\log_2\left(\frac{2}{9}\right) - \frac{2}{9}\log_2\left(\frac{2}{9}\right)$

$- \frac{1}{9}\log_2\left(\frac{1}{9}\right)$

$= -\frac{3}{9}\log_2\left(\frac{1}{9}\right) - \frac{6}{9}\log_2\left(\frac{2}{9}\right)$

$= 2.5032 \text{ bits} \simeq 3 \text{ bits}$

(ii) Relative entropy: $D(p||q) = \sum_x p(x) \log_b (p(x)/q(x))$

a) second dice w.r.t first dice:

$$D(\text{second} || \text{first}) = \frac{1}{9}\log_2\left(\frac{1/9}{1/6}\right) + \frac{1}{9}\log_2\left(\frac{1/9}{1/6}\right) + \frac{1}{9}\log_2\left(\frac{1/9}{1/6}\right)$$

$$+ \frac{2}{9}\log_2\left(\frac{2/9}{1/6}\right) + \frac{2}{9}\log_2\left(\frac{2/9}{1/6}\right) + \frac{2}{9}\log_2\left(\frac{2/9}{1/6}\right)$$

$$= \frac{3}{9}\log\left(\frac{2}{3}\right) + \frac{6}{9}\log\left(\frac{4}{3}\right)$$

$$= \frac{3}{9}(-0.5851) + \frac{6}{9}(0.415)$$

$$= 0.0816.$$

$$\cancel{D(\text{first} || \text{second}) = \frac{3}{9}\log\left(\frac{3}{2}\right) + \frac{6}{9}\log\left(\frac{3}{4}\right) = \frac{3}{9}(0.5849) + \frac{6}{9}(0.415)}$$

b) Third w.r.t first dice

$$D(\text{Third} || \text{first}) = \frac{1}{9}\log_2\left(\frac{1/9}{1/6}\right) + \frac{1}{9}\log_2\left(\frac{1/9}{1/6}\right) + \frac{2}{9}\log_2\left(\frac{2/9}{1/6}\right) +$$

$$\frac{2}{9}\log_2\left(\frac{2/9}{1/6}\right) + \frac{2}{9}\log_2\left(\frac{2/9}{1/6}\right) + \frac{1}{9}\log\left(\frac{1/9}{1/6}\right)$$

$$= \frac{3}{9}\log\left(\frac{2}{3}\right) + \frac{6}{9}\log\left(\frac{4}{3}\right)$$

$$= 0.0816$$

c) second dice w.r.t third dice

$$D(\text{second} || \text{third}) = \frac{1}{9}\log\left(\frac{1/9}{1/9}\right) + \frac{1}{9}\log_2\left(\frac{1/9}{1/9}\right) + \frac{1}{9}\log_2\left(\frac{1/9}{2/9}\right) + \frac{2}{9}\log_2\left(\frac{2/9}{2/9}\right)$$

$$+ \frac{2}{9}\log_2\left(\frac{2/9}{2/9}\right) + \frac{2}{9}\log\left(\frac{2/9}{1/9}\right) = -\frac{1}{9} + \frac{2}{9} = \frac{1}{9}$$

$$D(\text{second} || \text{third}) = 0.1111$$

Scanned by CamScanner