ICMA 151 Statistics for Science I

Academic Year 2023-2024 Trimester 3

Quiz No. 3 (7.5 %)

(Due July 4th by 11.59 PM)

Name Ye Zow Aung ID 6681078

Each problem is worth 20 points. (Show the calculation of all problems. Any answer without calculation details will not be graded.)

Problem 1. A random sample of n = 1000 observations from a binomial population produced x = 728 successes.

1.1 (1 point) Estimate the binomial proportion p.

binomial proportion $p = \frac{x}{n} = \frac{728}{1000} = 0.728.$

1.2 (2 points) Calculate the margin of error for the 95% confidence interval for the true proportion

$M_{95} = Z_{\frac{\alpha}{2}} \times SE$

$= Z_{\frac{0.05}{2}} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96 \times \sqrt{\frac{(0.728)(0.272)}{1000}} = 0.02758$

1.3 (2 points) Calculate the margin of error for the 99% confidence interval for the true proportion

$M_{99} = 2.576 \sqrt{\frac{(0.728)(0.272)}{1000}} = 0.03624$

1.4 (2 points) Find the 90% confidence interval for the true proportion by direct calculation

$M_{90} = 1.645 \sqrt{\frac{(0.728)(0.272)}{1000}} = \cancel{0.036249}\ 0.023148.$

$CI = \hat{p} \pm M_{90} = 0.728 \pm \overset{0.023148}{\cancel{0.036249}}$ (0.7048519 , 0.7511481)

1.5 (2 points) Find the 90% confidence interval for the true proportion by using prop.test .

prop.test (728, 1000, conf.level = 0.9 , correct = F)

CI = (0.7042618, 0.7505078)

1.6 (4 points) Test the null hypothesis $H_0: p = 0.6$ against the alternative $H_0: p \neq 0.6$ at the significance level 0.01 by direct calculation

1.6.1 The calculated test statistic is $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.728 - 0.6}{\sqrt{\frac{0.6(0.4)}{1000}}} = 8.262364$.

1.6.2 The p-value is

2x pnorm (abs(8.262364) , lower = F) = 1.42 8148e-16

1.6.3 The critical value(s) is (are)

1.6.4 Conclusion is Since p-value is lesser than significance value, null hypothesis is rejected.

1.7 (3 points) Test the null hypothesis $H_0: p = 0.6$ against the alternative $H_0: p \neq 0.6$ at the significance level 0.01 by prop.test

1.7.1 Rcode prop.test(728, 1000, conf level = 0.99, correct = F)

1.7.2 Result pvalue < 2.2e-16.

1.7.3 Conclusion Since pvalue is lesser than significance value, null hypothesis is rejected.

1.8 (2 points) Test the null hypothesis $H_0: p = 0.6$ against the alternative $H_0: p > 0.6$ at the significance level 0.01

1.8.1 The p-value is

pnorm (8.262364 , lower = F) = 7.140771e-17

1.8.2 Conclusion is

Since p-value is smaller than the significance value, null hypothesis is rejected.

1.9 (2 points) Test the null hypothesis $H_0: p = 0.6$ against the alternative $H_0: p < 0.6$ at the significance level 0.01

1.9.1 The p-value is

pnorm (8.262364) = 1

1.9.2 Conclusion is

Since p-value is greater than the significance value, null hypothesis is failed to rejected.

$P_1 = \frac{80}{3000} = 0.0267$ $P_2 = \frac{36}{2000} = 0.018.$

Problem 2 A certain change in the process for manufacturing component parts is being considered. Samples are taken under both the existing and the new process to determine if the new process results in an improvement. If 80 of 3000 items from the existing process are found to be defective and 36 of 2000 items from the new process are found to be defective

2.1 (1 point) Find the point estimate of $p_1 - p_2$

$$P_e = \hat{P}_1 - \hat{P}_2 = \frac{80}{3000} - \frac{36}{3000} = 0.0267 - 0.018 = 0.0087$$

2.2 (2 points) Calculate the margin of error for the 95% confidence interval for $p_1 - p_2$

$$ME_{95} = Z\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}} = 1.96\sqrt{\frac{(0.0267)(1-0.0267)}{3000} + \frac{(0.018)(1-0.018)}{2000}}.$$

$$= 0.008196891$$

2.3 (2 points) Calculate the margin of error for the 90% confidence interval for $p_1 - p_2$

$$ME_{90} = 1.645\sqrt{\frac{(0.0267)(1-0.0267)}{3000} + \frac{(0.018)(1-0.018)}{2000}},$$

$$= 0.006879533$$

2.4 (2 points) Find a 99% confidence interval for the true difference in the proportion of defectives between the existing and the new process by direct calculation.

$$ME_{99} = 2.576\sqrt{\frac{(0.0267)(1-0.0267)}{3000} + \frac{(0.018)(1-0.018)}{2000}}. = 0.01077306$$

$$CI = \hat{P}_1 - \hat{P}_2 \pm ME = (-0.00207306,\ 0.01947306)$$

2.5 (2 points) Find a 99% confidence interval for the true difference in the proportion of defectives between the existing and the new process by using prop.test

prop.test (c(80,36), c(3000,2000), conf.level = 0.99, correct = F)

(-0.002105676, 0.01943900 9).

2.6 (4 points) Test the null hypothesis $H_0: p_1 = p_2$ against the alternative $H_0: p_1 \neq p_2$ at the significance level 0.01 by direct calculation:

$$P_{pool} = \frac{80 + 36}{3000 + 2000} = 0.0232.$$

2.6.1 The calculated test statistic is

$$Z = \frac{P_e - P_0}{\sqrt{P_{pool}(1-P_{pool})(\frac{1}{n_1}+\frac{1}{n_2})}} = \frac{0.0087 - 0}{\sqrt{0.0232(0.9768)(\frac{1}{3000}+\frac{1}{2000})}} = 2.001995.$$

2× pnorm (abs(2.001995), lower=F).

$\Downarrow$

2.6.2 The p-value is 0.04528523.

2.6.3 Conclusion is

Since. ~~signifi~~ p-value is greater than significance value, null hypothesis is failed to rejected.

2.7 (3 points) Test the null hypothesis $H_0: p_1 = p_2$ against the alternative $H_0: p_1 \neq p_2$ at the significance level 0.01 by using prop.test

2.7.1 Rcode

```{r}
prop.test(c(80, 36), c(3000, 2000), conf.level = 0.99, correct=F)
```

2.7.2 Result

```
	2-sample test for equality of proportions without continuity correction

data:  c(80, 36) out of c(3000, 2000)
X-squared = 3.9773, df = 1, p-value = 0.04612
alternative hypothesis: two.sided
99 percent confidence interval:
 -0.002105676  0.019439009
sample estimates:
    prop 1     prop 2 
0.02666667 0.01800000 
```

2.7.3 Conclusion Since p-value is larger than the significance level, the null hypothesis is failed to rejected

2.8 (2 points) Test the null hypothesis $H_0: p_1 = p_2$ against the alternative $H_0: p_1 < p_2$ at the significance level 0.01

2.8.1 The p-value is pnorm (2.001995) = 0.9773574.

2.8.2 Conclusion is
Since p-value is greater than significance level, null hypothesis is failed to rejected.

2.9 (2 points) Test the null hypothesis $H_0: p_1 = p_2$ against the alternative $H_0: p_1 > p_2$ at the significance level 0.01

2.9.1 The p-value is pnorm (2.001995, lower = F) = 0.0226426 2.

2.9.2 Conclusion is
Since p-value is greater than significance level, null hypothes
is filed to reject.

Problem 3 Let χ_v^2 represent the chi-square distribution with the degree of freedom v

3.1 (2 points) Find $P(\chi_5^2 < 10)$

3.1.1 Rcode

pchisq (10, 5)

3.1.2 Result

0.9247648

3.2 (2 points) Find $P(\chi_6^2 > 9)$

3.2.1 Rcode

pchisq (9, 6, lower =F).

3.2.2 Result

0.1735781

3.3 (2 points) Find $(8 \leq \chi_{12}^2 < 14)$

3.3.1 Rcode

pchisq (14, 12) - pchisq (8, 12).

3.3.2 Result

0.4844221,

3.4 (2 points) Find the 75 percentiles of the chi-square distribution with degree of freedom 6

3.4.1 Rcode

qchisq (0.75, 6).

3.4.2 Result

7.840804.

3.5 (2 points) Find the value a such that $P(\chi_8^2 > a) = 0.36$

3.5.1 Rcode

qchisq (0.36, 8, lower = F).

3.5.2 Result

8.793667

3.6 (10 points) The grades in a statistics course for a particular semester were as follows

Grade	A	B	C	D	F
Frequencies	14	18	32	20	16

Test the hypothesis, at the 0.05 level of significance, that the distribution of grades is uniform.

3.6.1 (1 point) The expected frequencies of A, B, C, D, and F are

$$E = 14 \times 0.2 + 18 \times 0.2 + 32 \times 0.2 + 20 \times 0.2 + 16 \times 0.2$$
$$= 20.$$

3.6.2 (2 points) The calculated chi-square statistic is

$$x^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$
$$= \frac{(14-20)^2}{20} + \frac{(18-20)^2}{20} + \frac{(32-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(16-20)^2}{20}$$
$$= 10.$$

3.6.3 (2 points) The p-value is

pchisq (10, 4, lower = F). # 0.04042768

3.6.4 (1 point) The critical value at the significance level 0.05 is

qchisq (0.95, 4). [1] 9.487729

3.6.5 (1 point) Conclusion is

As P-value is less than the significance level, the null hypothesis is rejected. ~~that null~~

∴ The distribution of grades is not uniform.

3.6.6 (3 points) Repeat the hypothesis test by using chisq.test

chisq.test (c(14, 18, 32, 20, 16), p=rep (0.2, 5))

```{r}
chisq.test(c(14,18,32,20,16), p = rep(0.2, 5))
```

```
	Chi-squared test for given probabilities

data:  c(14, 18, 32, 20, 16)
X-squared = 10, df = 4, p-value = 0.04043
```