

---

# CENG 222

## Statistical Methods for Computer Engineering

Spring '2018-2019

Homework 4

---

### Student Information

Full Name : Yavuz Selim Yesilyurt

Id Number : 2259166

### Answer 9.10

The given information describes that the from a sample of 200 items 24 defective items were found. We can calculate the sample proportion of defective items as follows:

$$\begin{aligned}\hat{p} &= \frac{x}{n} \\ &= \frac{24}{200} \\ &= 0.12\end{aligned}$$

Also, for the 96% confidence interval we will have a significance level of  $\alpha = 0.04$  and from standard normal table we can check its value which is 2.054.

a)

With the information we obtained above, let us construct a 96% confidence interval for the proportion of defective items as follows:

$$\begin{aligned}96\% \text{ Confidence Interval} &= \hat{p} \pm Z_{critical} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \\ &= 0.12 \pm 2.054 \sqrt{\frac{0.12(1 - 0.12)}{200}} \\ &= 0.12 \pm 0.047 \\ &= (0.073, 0.167)\end{aligned}$$

b)

First let us state our hypotheses:

$$H_0 : p \leq 0.1 \text{ and } H_A : p > 0.1$$

And as level of significance we have  $\alpha = 0.04$  and  $\alpha = 0.15$  values. Let us calculate the test statistic and then find the  $p$  value to determine whether to accept or reject the Null hypothesis ( $H_0$ ):

$$\begin{aligned} Z &= \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \\ &= \frac{0.12 - 0.1}{\sqrt{\frac{0.1(1-0.1)}{200}}} \\ &= \frac{0.02}{0.0212} \\ &= 0.943 \end{aligned}$$

Find the  $p$  value:

$$\begin{aligned} p &= P(Z > Z_0) \\ &= 1 - P(Z \leq 0.943) \\ &= 1 - 0.82716 \text{ (From Std Norm Table)} \\ &= 0.17284 \end{aligned}$$

In Conclusion, we see that the  $p$  value (0.17284) is greater than the given significance levels (0.04 and 0.15), so we fail to reject the null hypothesis and conclude that there would be sufficient evidence to say that the claim is incorrect.

## Answer 9.12

We are given:

$$\bar{x} = 0.62, \sigma = 0.2 \text{ and } n = 52.$$

a)

To construct a 95% confidence interval for the population mean resistance first let us find our level of significance at 95% which is  $\alpha = 0.05$  and from Standard Normal table it has value of 1.96. So:

$$\begin{aligned} 95\% \text{ Confidence Interval} &= \bar{x} \pm Z_{critical} \left( \frac{\sigma}{\sqrt{n}} \right) \\ &= 0.62 \pm 1.96 \left( \frac{0.2}{\sqrt{52}} \right) \\ &= 0.62 \pm 0.05436 \\ &= (0.56564 < \mu < 0.67436) \end{aligned}$$

b)

The probability that resistance is 0.62 ohms or higher can be found as follows:

$$\begin{aligned} P(X > 0.62) &= 1 - P(Z \leq 0.62) \\ &= 1 - P\left(Z \leq \frac{0.62 - 0.6}{0.2/\sqrt{52}}\right) \\ &= 1 - P(Z \leq 0.721) \\ &= 1 - 0.764545 \text{ (From Std Norm Table)} \\ &= 0.235455 \end{aligned}$$

## Answer 10.3

From the given random number generator data with  $N = 100$ , after arranging data into increasing order, let us form a continuous type frequency table of the data which is given below:

Class interval	Frequency
below -2.0	4
-2.0 to -1.5	4
-1.5 to -1.0	15
-1.0 to -0.5	9
-0.5 to 0	22
0 to 0.5	15
0.5 to 1.0	12
1.0 to 1.5	11
1.5 to 2.0	7
2.0 and above	1
Total	100

a)

Using the table of normal distribution; the table of expected, observed frequencies and other required columns comes out to be:

Class interval	$obs_i$	$exp_i$	$\chi^2$
below -2.0	4	3.32	0.14
-2.0 to -1.5	4	5.32	0.33
-1.5 to -1.0	15	10.02	2.48
-1.0 to -0.5	9	15.14	2.49
-0.5 to 0	22	18.38	0.71
0 to 0.5	15	17.92	0.48
0.5 to 1.0	12	14.03	0.29
1.0 to 1.5	11	8.82	0.54
1.5 to 2.0	7	4.46	1.45
2.0 and above	1	2.59	0.97
Total	100	100	9.88332

Now to test the hypothesis at 5% significance level whether the given data follows a normal distribution, we need to calculate the test statistic, which can be calculated as follows:

$$\begin{aligned}\chi^2 &= \sum \frac{(obs_i - exp_i)^2}{exp_i} \\ &= 9.88\end{aligned}$$

And degrees of freedom in this case is:

$$\begin{aligned} v &= n - 1 \\ &= 10 - 1 \\ &= 9 \end{aligned}$$

The  $p$  value for the above value of test statistic at 9 degrees of freedom can be found from table A6 and founded to be between 0.2 and 0.8, which is more than the significance level 0.05. We conclude that there is no evidence against the data follows a normal distribution.

**b)**

The pdf of Uniform distribution is given by:

$$f(x) = \frac{1}{b-a} \quad a \leq x \leq b$$

In this problem we have our  $a = -3$ ,  $b = 3$ . We need to calculate the table of expected, observed frequencies and other required columns again for Uniform distribution, which comes out to be:

Class interval	$obs_i$	$exp_i$	$\chi^2$
below -2.0	4	16.67	9.63
-2.0 to -1.5	4	8.33	2.25
-1.5 to -1.0	15	8.33	5.33
-1.0 to -0.5	9	8.33	0.05
-0.5 to 0	22	8.33	22.41
0 to 0.5	15	8.33	5.33
0.5 to 1.0	12	8.33	1.61
1.0 to 1.5	11	8.33	0.85
1.5 to 2.0	7	8.33	0.21
2.0 and above	1	16.67	14.73
Total	100	100	62.42

Now to test the hypothesis at 5% significance level whether the given data follows a Uniform distribution, we need to calculate the test statistic, which can be calculated as follows:

$$\begin{aligned} \chi^2 &= \sum \frac{(obs_i - exp_i)^2}{exp_i} \\ &= 62.42 \end{aligned}$$

And degrees of freedom in this case is:

$$\begin{aligned} v &= n - 1 \\ &= 10 - 1 \\ &= 9 \end{aligned}$$

The  $p$  value for the above value of test statistic at 9 degrees of freedom can be found from table A6 and founded to be  $< 0.001$ , which is much lower than the significance level 0.05. We conclude that there is strong evidence against the data follows a Uniform distribution.

c)

According to the central limit theorem, theoretically it is possible that a data follows Normal and Uniform distributions simultaneously for a large sample.

## Answer 10.9

Let us first state our Null and Alternative hypotheses:

$$H_0 : \text{all three section's performance is equal}$$

$$H_A : \text{all three section's performance is not equal}$$

We have our level of significance as  $\alpha = 0.05$ . Now let us again create a table of expected, observed frequencies and other required columns and perform a Chi-Square test. Note that this time we have 3 samples and we will draw the table in a slightly different fashion. Each column shows a section and each row shows a grade. There are three numbers in each entry which corresponds to:

$$\begin{array}{c} \text{Observed} \\ \text{Observed - Expected} \\ \chi^2 \end{array}$$

Grade	S01	S02	S03	Total
A	40	20	20	80
	5.71	-8.57	2.86	0
	0.95	2.57	0.48	4
B	50	40	20	110
	2.86	0.71	-3.57	0
	0.17	0.01	0.54	0.73
C	20	25	15	60
	-5.71	3.57	2.14	0
	1.27	0.6	0.36	2.22
D	2	5	2	9
	-1.86	1.79	0.07	0
	0.89	0.99	0	1.89
F	8	10	3	21
	-1	2.5	-1.5	0
	0.11	0.83	0.5	1.44
Total	120	100	60	280
	0	0	0	0
	3.4	5.01	1.88	10.28

As can be seen from the last row, last column of our table,  $\chi^2$  of the total is equal to 10.28 and we have a  $v = 8$  degrees of freedom.

The  $p$  value for the above value of test statistic at 8 degrees of freedom can be found from table A6 and founded to be between 0.8 and 0.2, which is greater than the significance level 0.05.

So in conclusion, Since, the  $p$  value greater than the given significance level 0.05, so we fail to reject the null hypothesis and conclude that the three sections performance is equal.