

---

## CENG 222

### Assignment 2

Deadline: May 13, 23:59

Full Name: Beyazıt Yalçınkaya

Id Number: 2172138

---

### Answer 9.16

**a**

We have  $n_1 = 250$ ,  $n_2 = 300$ ,  $\hat{p}_1 = 10/250$ ,  $\hat{p}_2 = 18/300$ ,  $\alpha = 1 - 0.98 = 0.02$ , and  $z_{\alpha/2} = z_{0.01} = 2.33$ . For the confidence interval,

$$\begin{aligned}\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}} &= 0.04 - 0.06 \pm 2.33 \sqrt{\frac{0.04(1 - 0.04)}{250} + \frac{0.06(1 - 0.06)}{300}} \\ &= -0.02 \pm 0.043\end{aligned}\quad (1)$$

Then the required confidence interval is as follows,

$$[-0.063, 0.023] \quad (2)$$

**b**

In part **a**, we have computed a 98% confidence interval (which is the same thing as a significance level  $\alpha = 0.02$  since  $(1 - \alpha)100\% = 98\%$ ) for the difference of proportions of defective items:  $[-0.063, 0.023]$ . This interval contains 0, therefore, the test of

$$H_0 : p_1 = p_2 \text{ vs } H_A : p_1 \neq p_2$$

accepts the null hypothesis at the 2% level. Thus, there is no evidence of a significant difference between the quality of the two lots.

### Answer 10.2

For the distribution  $F$  of the service times, let us test

$$H_0 : F = \text{Exponential}(\lambda) \text{ for some } \lambda \text{ vs } H_A : F \neq \text{Exponential}(\lambda) \text{ for any } \lambda.$$

Maximum likelihood estimator of  $\lambda$  is as follows (Probability and Statistics for Computer Scientists, page 244, Example 9.8).

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n X_i} = \frac{64}{320} = 0.2 \quad (3)$$

Splitting the support into bins as follows. Note that the unit of the intervals is minutes.

$$\begin{aligned} B_1 &= (-\infty, 1.2), B_2 = [1.2, 2.3), B_3 = [2.3, 3.3), B_4 = [3.3, 4.4), \\ B_5 &= [4.4, 5.3), B_6 = [5.3, 7.1), B_7 = [7.1, 10.2), B_8 = [10.2, \infty). \end{aligned} \quad (4)$$

Calculating each  $p_k$ .

$$\begin{aligned} p_1 &= 1 - e^{(-0.2)(1.2)} = 0.213372 \\ p_2 &= (1 - e^{(-0.2)(2.3)}) - (1 - e^{(-0.2)(1.2)}) = 0.155344 \\ p_3 &= (1 - e^{(-0.2)(3.3)}) - (1 - e^{(-0.2)(2.3)}) = 0.114432 \\ p_4 &= (1 - e^{(-0.2)(4.4)}) - (1 - e^{(-0.2)(3.3)}) = 0.102068 \\ p_5 &= (1 - e^{(-0.2)(5.3)}) - (1 - e^{(-0.2)(4.4)}) = 0.068328 \\ p_6 &= (1 - e^{(-0.2)(7.1)}) - (1 - e^{(-0.2)(5.3)}) = 0.104742 \\ p_7 &= (1 - e^{(-0.2)(10.2)}) - (1 - e^{(-0.2)(7.1)}) = 0.111685 \\ p_8 &= 1 - (1 - e^{(-0.2)(10.2)}) = 0.130029 \end{aligned} \quad (5)$$

After counting each  $Obs(k)$  from the given data and calculating each  $Exp(k) = np_k$ , we get the following table.

$k$	1	2	3	4	5	6	7	8
$B_k$	$(-\infty, 1.2)$	$[1.2, 2.3)$	$[2.3, 3.3)$	$[3.3, 4.4)$	$[4.4, 5.3)$	$[5.3, 7.1)$	$[7.1, 10.2)$	$[10.2, \infty)$
$p_k$	0.213372	0.155344	0.114432	0.102068	0.068328	0.104742	0.111685	0.130029
$Exp(k)$	13.655808	9.942016	7.323648	6.532352	4.372992	6.703488	7.14784	8.321856
$Obs(k)$	8	8	8	8	8	8	8	8

From the table, the chi-square statistic is as follows.

$$\chi_{obs}^2 = \sum_{k=1}^8 \frac{\{Obs(k) - Exp(k)\}^2}{Exp(k)} = 6.487084 \quad (6)$$

Comparing it against the Chi-square distribution with  $8 - 1 - 1 = 6$  d.f. in the table (since we have one estimated parameter  $\hat{\lambda}$ , we subtracted an extra one). We find the following P-value.

$$P = P\{\chi^2 \geq 6.487084\} > 0.2 \quad (7)$$

Hence, there is no evidence against an Exponential Distribution of the service times. (Note that, for testing  $H_0$  with a P-value, the practically used values are adopted, for details see Probability and Statistics for Computer Scientists, page 282, Testing hypotheses with a P-value chapter.)

## Answer 10.3

**a**

For the distribution  $F$  of the given data, let us test

$$H_0 : F = StandardNormal \text{ vs } H_A : F \neq StandardNormal.$$

Splitting the support into bins as follows.

$$B_1 = (-\infty, -1.052), B_2 = [-1.052, -0.373), B_3 = [-0.373, 0.251), B_4 = [0.251, 0.978), B_5 = [0.978, \infty) \quad (8)$$

Calculating each  $p_k$ , by using the Standard Normal Distrubution Table.

$$\begin{aligned} p_1 &= \Phi(-1.052) = 0.1464 \\ p_2 &= \Phi(-0.373) - \Phi(-1.052) = 0.3557 - 0.1469 = 0.2082 \\ p_3 &= \Phi(0.251) - \Phi(-0.373) = 0.5987 - 0.3557 = 0.2445 \\ p_4 &= \Phi(0.978) - \Phi(0.251) = 0.8365 - 0.5987 = 0.2369 \\ p_5 &= 1 - \Phi(0.978) = 1 - 0.8365 = 0.1640 \end{aligned} \quad (9)$$

After counting each  $Obs(k)$  from the given data and calculating each  $Exp(k) = np_k$ , we get the following table.

$k$	1	2	3	4	5
$B_k$	$(-\infty, -1.052)$	$[-1.052, -0.373)$	$[0.373, 0.251)$	$[0.251, 0.978)$	$[0.978, \infty)$
$p_k$	0.1464	0.2082	0.2445	0.2369	0.1640
$Exp(k)$	14.64	20.82	24.45	23.69	16.40
$Obs(k)$	20	20	20	20	20

From the table, the chi-square statistic is as follows.

$$\chi_{obs}^2 = \sum_{k=1}^5 \frac{\{Obs(k) - Exp(k)\}^2}{Exp(k)} = 4.1674 \quad (10)$$

Comparing it against the Chi-square distribution with  $5 - 1 = 4$  d.f. in the table. We find the following P-value.

$$P = \mathbf{P}\{\chi^2 \geq 4.1674\} > 0.2 \quad (11)$$

Hence, there is no evidence against a Standard Normal Distribution of the given data. (Note that, for testing  $H_0$  with a P-value, the practically used values are adopted, for details see Probability and Statistics for Computer Scientists, page 282, Testing hypotheses with a P-value chapter.)

## b

For the distribution  $F$  of the given data, let us test

$$H_0 : F = Uniform(-3, 3) \text{ vs } H_A : F \neq Uniform(-3, 3).$$

Splitting the support into bins as follows.

$$B_1 = (-3, -1.052), B_2 = [-1.052, -0.373), B_3 = [-0.373, 0.251), B_4 = [0.251, 0.978), B_5 = [0.978, 3) \quad (12)$$

Calculating each  $p_k$ , by using the cdf function of the Uniform Distribution.

$$\begin{aligned}
p_1 &= \int_{-3}^{-1.052} \frac{1}{6} dx = \frac{-1.052 + 3}{6} = 0.3247 \\
p_2 &= \int_{-1.052}^{-0.373} \frac{1}{6} dx = \frac{0.373 + 1.052}{6} = 0.1132 \\
p_3 &= \int_{-0.373}^{0.251} \frac{1}{6} dx = \frac{0.251 + 0.373}{6} = 0.1040 \\
p_4 &= \int_{0.251}^{0.978} \frac{1}{6} dx = \frac{0.978 - 0.251}{6} = 0.1211 \\
p_5 &= \int_{0.978}^3 \frac{1}{6} dx = \frac{3 - 0.978}{6} = 0.3370
\end{aligned} \tag{13}$$

After counting each  $Obs(k)$  from the given data and calculating each  $Exp(k) = np_k$ , we get the following table.

$k$	1	2	3	4	5
$B_k$	$(-3, -1.052)$	$[-1.052, -0.373)$	$[0.373, 0.251)$	$[0.251, 0.978)$	$[0.978, 3)$
$p_k$	0.3247	0.1132	0.1040	0.1211	0.3370
$Exp(k)$	32.47	11.32	10.40	12.11	33.70
$Obs(k)$	20	20	20	20	20

From the table, the chi-square statistic is as follows.

$$\chi_{obs}^2 = \sum_{k=1}^5 \frac{\{Obs(k) - Exp(k)\}^2}{Exp(k)} = 31.01 \tag{14}$$

Comparing it against the Chi-square distribution with  $5 - 1 = 4$  d.f. in the table. We find the following P-value.

$$P = \mathbf{P}\{\chi^2 \geq 31.01\} < 0.001 \tag{15}$$

Hence, there is significant evidence against a Uniform(-3, 3) Distribution of the given data. (Note that, for testing  $H_0$  with a P-value, the practically used values are adopted, for details see Probability and Statistics for Computer Scientists, page 282, Testing hypotheses with a P-value chapter.)

### c

Yes, it is possible to accept both. Rather than using conventional values for level of significance of the P-test, we can adjust the level of significance so that both null hypothesis are accepted by the P-test. This makes sense because of the central limit theorem. By central limit theorem, it is known that for big amount of data, uniform distribution can be converged into standard normal distribution. Thus, it is theoretically possible and it makes sense.