CENG 222

Assignment 2

Deadline: May 13, 23:59

Full Name: Adil Kaan Akan
Id Number: 2171155

Answer 9.16

a

Since there are a lot sample, we can use the values in the question directly.
$\hat{a} = \frac{10}{250} = 0.04$
$\hat{b} = \frac{18}{300} = 0.06$
$z_{\alpha/2} = 2.326$
If we use the formula for difference,

$$
\begin{aligned}
&= (\hat{p}_A - \hat{p}_B) \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}} \\
&= (0.04 - 0.06) \pm 2.326\sqrt{\frac{0.04(1-0.04)}{250} + \frac{0.06(1-0.06)}{300}} \\
&= -0.02 \pm 0.043 \\
&= (-0.063, 0.023)
\end{aligned}
$$

b

The level of significance, $\alpha = 0.02$

$$
\begin{aligned}
Z &= \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}} \\
&= \frac{0.04 - 0.06}{\sqrt{\frac{0.04(1-0.04)}{250} + \frac{0.06(1-0.06)}{300}}} \\
&= -1.06
\end{aligned}
$$

We should find the p-value for those two lots.

$$
\begin{aligned}
\textit{p-value} &= 2P(Z < Z_0) \\
&= 2xP(Z \leq -1.06) \\
&= 2x0.14457 \\
&= 0.28915
\end{aligned}
$$

Since we found a p-value that is greater than our significance level after doing our calculations, we can say that there is no significant difference between the qualites of two lots.

Answer 10.2

When we calculate the mean of the 64 observations of the system, we get $\bar{X} = 5.0$
The exponential distribution says that the Cumulative distribution function F(X),

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= 1 - e^{-\lambda x} \ \textit{where } 0 \leq x \leq \infty \\ \lambda &= \frac{1}{\bar{X}} \\ &= \frac{1}{5.0} \\ &= 0.2 \end{aligned}$$

$$F(x) = 1 - e^{-0.2x} \ \textit{where } 0 \leq x \leq \infty$$

By using that function we can calculate the expected fequencies of j^{th} class $(a_j - b_j)$ by using
$e_j = F(b_j) - F(a_j)$
After calculating it, we should test the assumption at 5% significance level whether the asssumption of the Exponentiality supported by these data.
We can test it by using the following function,
$X^2 = \sum(\frac{(o_j - e_j)^2}{e_j})$ where
o_j is the number of frequencies in the intervals(0-2,2-4,4-6,....,14-16)
$e_j = F(b_j) - F(a_j)$

Class Interval	o_j	e_j	(o_j - e_j)^2 / e_j
0-2	13	21.10	3.11
2-4	16	14.14	0.24
4-6	15	9.48	3.21
6-8	7	6.36	0.07
8-10	5	4.26	0.13
10-12	5	2.86	1.61
12-14	2	1.91	0.00
14-16	1	3.89	2.15
Total	64	64	10.52

Since there are values that are less than 5 in e_j column, we should merge them.

Class Interval	o_j	e_j	(o_j - e_j)^2 / e_j
0-2	13	21.10	3.11
2-4	16	14.14	0.24
4-6	15	9.48	3.21
6-8	7	6.36	0.07
8-12	10	7.12	1.16
12-16	3	5.80	1.35
Total	64	64	9.14

$$X^2 = \sum\left(\frac{(o_j - e_j)^2}{e_j}\right)$$
$$= 9.14$$

The degrees of freedom in this case is

$$\begin{aligned} df &= n - 1 - 1 \\ &= 6 - 1 - 1 \\ &= 4 \end{aligned}$$

The conclusion:
If we look to the p-value for the test at 4 df, our value is 9.14 and the α value is in between 0.05 and 0.1. In generally α values are in that interval. We can say that there is not sufficient evidence to reject the assumption. So, we should accept the assumption of Exponentiality is not supported by these data.

Answer 10.3

a

When we calculate the mean and the standard deviation of the 100 observations, we get $\bar{X} = -0.058$ and $\sigma = 1.058$.

Class Size	o_j	e_j	(o_j-e_j)^2/e_j
below -1.5	8	6.68	0.26
-1.5 to -1.0	15	9.19	3.67
-1.0 to -0.5	9	14.98	2.39
-0.5 to 0.0	22	19.15	0.42
0.0 to 0.5	15	22.21	2.34
0.5 to 1.0	12	13.97	0.28
1.0 to 1.5	11	8.2	0.96
1.5 and above	8	5.59	1.04
Total	100	100	11.36

We should test the assumption 5 % significance level whether the data follows the normal distribution.

$$X^2 = \sum \frac{(o_j - e_j)^2}{e_j}$$
$$= 11.36$$

Degrees of freedom

$$df = n - 1$$
$$= 8 - 1$$
$$= 7$$

If we look to the p-value for the test at 7 df, our value is 11.36 and the α value is in between 0.1 and 0.2. Since our value is greater than general α interval(0.05, 0.1), we should accept the assumption whether the data follows the standard normal distribution.

b

The pdf of the uniform distribution is
$f(x) = \frac{1}{b-a}$ where $a \leq x \leq b$
For this question, a is -3 and b is 3.

Class Size	o_j	e_j	(o_j-e_j)^2/e_j
below -1.5	8	25	11.56
-1.5 to -1.0	15	8.33	5.34
-1.0 to -0.5	9	8.33	0.05
-0.5 to 0.0	22	8.33	22.43
0.0 to 0.5	15	8.33	5.34
0.5 to 1.0	12	8.33	1.62
1.0 to 1.5	11	8.33	0.86
1.5 and above	8	25	11.56
Total	100	100	58.76

We should test the assumption 5 % significance level whether the data follows the uniform distribution.

$$X^2 = \sum \frac{(o_j - e_j)^2}{e_j}$$
$$= 58.76$$

Degrees of freedom

$$df = n - 1$$
$$= 8 - 1$$
$$= 7$$

If we look to the p-value for the test at 7 df, our value is 58.76 and the α value is far less than 0.001. Since our value is far less than general α interval(0.05, 0.1), we should reject the assumption whether the data follows the uniform distribution.

c

Since they are both distributions and possibilities, the chi square test can accept both of them simultaneously.