CENG 222

Assignment 2
Deadline: May 13, 23:59

Full Name:Alper KOCAMAN
Id Number:2169589

Answer 9.16

a

Sample proportion $\widehat{p} = \frac{number\ of\ sampled\ items}{n}$.

Also,
$E(\widehat{p}) = p$ and $\mathrm{Var}(\widehat{p}) = \frac{p \times (1-p)}{n}$.

The paramater of interest $\theta = p_1 - p_2$ and it is estimated by $\widehat{\theta} = \widehat{p_1} - \widehat{p_2}$,this will be center.

The standard error is estimated by $s(\widehat{\theta}) = \sqrt{\frac{\widehat{p_1} \times (1-\widehat{p_1})}{n_1} + \frac{\widehat{p_2} \times (1-\widehat{p_2})}{n_2}}$,this will be margin.

Thus.confidence interval formula for difference of proportions is:

$$\widehat{\theta} \mp z_{0.02/2} \times s(\widehat{\theta}) = (\widehat{p_1} - \widehat{p_2}) \mp z_{0.02/2} \times \sqrt{\frac{\widehat{p_1} \times (1-\widehat{p_1})}{n_1} + \frac{\widehat{p_2} \times (1-\widehat{p_2})}{n_2}}.$$

In this problem, $n_1 = 250$ and $n_2 = 300$.
So,$\widehat{p_1} = 10/250 = 0.04$ and $\widehat{p_2} = 18/300 = 0.06$.

Center $= 0.04 - 0.06 = -0.02$.
$z_{0.02/2} = 2.326$
Margin $= \sqrt{\frac{0.04 \times 0.96}{250} + \frac{0.06 \times 0.94}{300}}$.

Thus,the answer is $0.02 \mp 0.043 = (-0.063, 0.023)$.

b

Null Hypothesis - H_0: There is no significant difference between two lots' quality.

Alternative Hypothesis - H_A: There is a significant difference between two lots' quality.

Step 1 : Test Statistic.

For these Bernoulli data, the variance depends on the unknown parameters p_1 and p_2 which are estimated by the sample proportions $\widehat{p_1}$ and $\widehat{p_2}$.

$$\frac{\widehat{p_1}-\widehat{p_2}}{\sqrt{\frac{\widehat{p_1}\times(1-\widehat{p_1})}{n_1}+\frac{\widehat{p_2}\times(1-\widehat{p_2})}{n_2}}} = \frac{0.04-0.06}{\sqrt{\frac{0.04\times0.96}{250}+\frac{0.06\times0.94}{300}}} = -1.06.$$

From part a) $z_{0.01} = 2.326$, since $|-1.06| < |2.326|$,we accept hypothesis Null (H_0).
There is no significant difference between two lots' quality.

Answer 10.2

There are 64 values in question and the mean value of values given in the question is $\bar{X} = 5$.

Cumulative distribution function is $1 - e^{-\lambda x}$ where
$\lambda = 1/\bar{X} = 1/5 = 0.2$

So,the equation becomes $F(x) = 1 - e^{-0.2x}$.

Since there are 64 values , these data should be groupped.The bins should contain at least 5 element and number of groups should be between 5-8(preferably).
A suitable groupping can be $B_0 = [0, 2]$ minutes , $B_1 = [2, 4]$ minutes , $B_2 = [4, 6]$ minutes, $B_3 = [6, 8]$ minutes ,$B_4 = [8, \infty]$ minutes.There are 5 groups.

By the aid of above function $F(x) = 1 - e^{-0.2x}$, the values of that partitons calculated as:

$x = 0 \rightarrow 1 - e^{-0.2x} = 0$

$x = 2 \rightarrow 1 - e^{-0.2x} = 0.329$

$x = 4 \rightarrow 1 - e^{-0.2x} = 0.55$

$x = 6 \rightarrow 1 - e^{-0.2x}$ =0.70

$x = 8 \rightarrow 1 - e^{-0.2x}$ =0.80

Then expexted number of elements in bins corresponding to their values :

$B_0 = [0, 2] = 64 \times 0,329 = 21.05$

$B_1 = [2, 4] = 64 \times (0.55 - 0,329) = 14.14$

$B_2 = [4, 6] = 64 \times (0.7 - 055) = 9.6$

$B_3 = [6, 8] = 64 \times (0.8 - 0.7) = 6.4$

$B_4 = [8, \infty] = 64 \times (1 - 0.8) = 12.8$

For each of bin , using formula $\frac{(obs-exp)^2}{exp}$:

$B_0 \rightarrow = 3.08$

$B_1 \rightarrow = 0.24$

$B_2 \rightarrow = 3.04$

$B_3 \rightarrow =$0.06

$B_4 \rightarrow =$0.003

Then,using formula $\chi^2 = \Sigma \frac{(obs-exp)^2}{exp}$:

$3.08 + 0.24 + 3.04 + 0.06 + 0.003 = 6.423$.

Degrees of Freedom is 4 and the corresponding value of that is 9.49 with a significance level of %5(from table of coursebook page 420). Found value is less than 9.49 so there is no evidence against an exponential distribution.

Answer 10.3

a

There are lots of data(100) and this data should be groupped.There can be 10 bins for these data. Mean of these data is $\bar{X} = -0.058$.

$B_0 = [-\infty, -2]$ observed data:4

$B_1 = [-2, -1.5]$ observed data:4

$B_2 = [-1.5, -1.0]$ observed data:15

$B_3 = [-1.0, -0.5]$ observed data:9

$B_4 = [-0.5, 0]$ observed data:22

$B_5 = [0, 0.5]$ observed data:15

$B_6 = [0.5, 1]$ observed data:12

$B_7 = [1, 1.5]$ observed data:11

$B_8 = [1.5, 2]$ observed data:7

$B_9 = [2, \infty]$ observed data:1

From the Standard normal distribution table from book(page 417) expected values calculated as:
$B_0 = [-\infty, -2]$ expexted data: $0.0228 \times 100 = 2.28$

$B_1 = [-2, -1.5]$ expexted data: $(0.0668 - 0.0228) \times 100 = 4.4$

$B_2 = [-1.5, -1.0]$ expexted data: $(0.1587 - 0.0668) \times 100 = 9.19$

$B_3 = [-1.0, -0.5]$ expexted data: $(0.3085 - 0.1587) \times 100 = 14.98$

$B_4 = [-0.5, 0]$ expexted data: $(0.5 - 0.3085) \times 100 = 19.15$

$B_5 = [0, 0.5]$ expexted data: $(0.6915 - 0.5) \times 100 = 19.15$

$B_6 = [0.5, 1]$ expexted data: $(0.8413 - 0.6915) \times 100 = 14.98$

$B_7 = [1, 1.5]$ expexted data: $(0.9332 - 0.8413) \times 100 = 9.19$

$B_8 = [1.5, 2]$ expexted data: $(0.9772 - 0.9332) \times 100 = 4.4$

$B_9 = [2, \infty]$ expexted data: $(1 - 0.9772) \times 100 = 2.28$

For each of bin , using formula $\frac{(obs-exp)^2}{exp}$:

$B_0 = [-\infty, -2]$ value $= 1.30$

$B_1 = [-2, -1.5]$ value $= 0.04$

$B_2 = [-1.5, -1.0]$ value $= 3.67$

$B_3 = [-1.0, -0.5]$ value $= 2.39$

$B_4 = [-0.5, 0]$ value $= 0.42$

$B_5 = [0, 0.5]$ value $= 0.90$

$B_6 = [0.5, 1]$ value $= 0.60$

$B_7 = [1, 1.5]$ value $= 0.36$

$B_8 = [1.5, 2]$ value $= 1.53$

$B_9 = [2, \infty]$ value $= 0.71$

Then,using formula $\chi^2 = \Sigma \frac{(obs-exp)^2}{exp} = 11.92$.

Degrees of Freedom is 9 and the corresponding value of that is 16.9 with a significance level of %5(from table of coursebook page 420). Found value is less than 16.9 so there is no evidence against Standard Normal distribution.

b

Pdf of Uniform distribution is $\frac{1}{b-a}$.(a= -3 , b= 3)

$B_0 = [-\infty, -2]$ observed data:4

$B_1 = [-2, -1.5]$ observed data:4

$B_2 = [-1.5, -1.0]$ observed data:15

$B_3 = [-1.0, -0.5]$ observed data:9

$B_4 = [-0.5, 0]$ observed data:22

$B_5 = [0, 0.5]$ observed data:15

$B_6 = [0.5, 1]$ observed data:12

$B_7 = [1, 1.5]$ observed data:11

$B_8 = [1.5, 2]$ observed data:7

$B_9 = [2, \infty]$ observed data:1

Expected values calculated as:

$B_0 = [-\infty, -2]$ expexted data = 16.67

$B_1 = [-2, -1.5]$ expexted data = 8.33

$B_2 = [-1.5, -1.0]$ expexted data = 8.33

$B_3 = [-1.0, -0.5]$ expexted data = 8.33

$B_4 = [-0.5, 0]$ expexted data = 8.33

$B_5 = [0, 0.5]$ expexted data = 8.33

$B_6 = [0.5, 1]$ expexted data = 8.33

$B_7 = [1, 1.5]$ expexted data $= 8.33$

$B_8 = [1.5, 2]$ expexted data $= 8.33$

$B_9 = [2, \infty]$ expexted data $= 16.67$

For each of bin , using formula $\frac{(obs-exp)^2}{exp}$:

$B_0 = [-\infty, -2]$ value $= 9.63$

$B_1 = [-2, -1.5]$ value $= 2.25$

$B_2 = [-1.5, -1.0]$ value $= 5.33$

$B_3 = [-1.0, -0.5]$ value $= 0.05$

$B_4 = [-0.5, 0]$ value $= 22.41$

$B_5 = [0, 0.5]$ value $= 5.33$

$B_6 = [0.5, 1]$ value $= 1.61$

$B_7 = [1, 1.5]$ value $= 0.85$

$B_8 = [1.5, 2]$ value $= 0.21$

$B_9 = [2, \infty]$ value $= 14.73$

Then,using formula $\chi^2 = \Sigma \frac{(obs-exp)^2}{exp} = 62.42$.

Degrees of Freedom is 9 and the corresponding value of that is 16.9 with a significance level of %5(from table of coursebook page 420). Found value is far more than 16.9 so there is a strong evidence against Uniform distribution.

c

Since data sample is large ($n \geq 30$) ,central limit theorem can be applied.So , although this problem is a counterexample , a large data sample can follow both Uniform and Standard distribution at the same time theoretically.