Tutorial letter 202/2/2019

Basic Statistics

STA1510

Semester 2

Department of Statistics

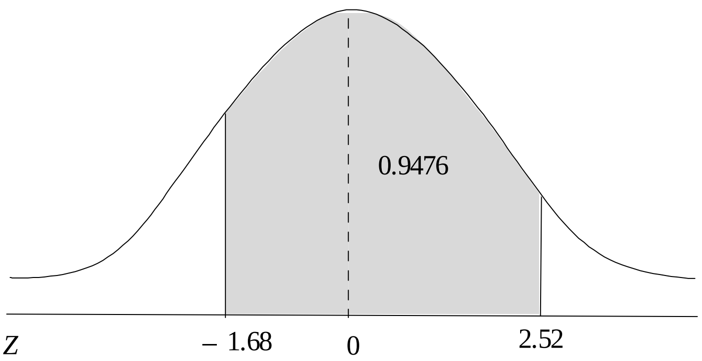
SOLUTIONS TO ASSIGNMENT 02

QUESTION 1

Sampling distribution of the mean

$$\mu = 40\ 000 \text{ km} \qquad \sigma_{\bar{X}} = \frac{400}{\sqrt{45}}$$
$$\sigma = 4\ 000 \text{ km} \qquad = 596.2848$$

$$P\left(39\ 000 < \bar{X} < 41\ 500\right)$$
$$P\left(\frac{39\ 000 - 40\ 000}{596.2848} < Z < \frac{41\ 500 - 40\ 000}{596.2848}\right)$$
$$P\left(-1.68 < Z < 2.52\right)$$
$$= 0.9941 - 0.0465$$
$$= 0.9476$$



Option 1

QUESTION 2

Confidence interval for the proportion.

$$p = \frac{30}{100} = 0.3$$

99% confidence interval for π

$$p \pm Z\sqrt{\frac{p\left(1-p\right)}{n}}$$
$$0.3 \pm 2.58\sqrt{\frac{0.3 \times 0.7}{100}}$$
$$0.3 \pm 0.1182$$
$$\left[0.1818 \leq \pi \leq 0.4182\right]$$

We are 99% confident that the proportion of students who pay their own fees will be between 18.18% and 41.82%.

Option 5

QUESTION 3

Confidence interval for μ (σ unknown).

$$\begin{aligned} n &= 10 \\ \bar{X} &= 9.25 \\ S &= 2.21 \end{aligned}$$

99% confidence interval for μ.

$$\bar{X} \pm t_{\frac{s}{\sqrt{n}}}$$
$$9.25 \pm (3.2498)\,(0.7)$$
$$[6.9751 \leq \mu \leq 11.5249]$$

We are 99% confident that the mean number of hours the battery will last in an iPod will be between 6.975 hours and 11.52 hours.

Option 2

QUESTION 4

Confidence interval for the μ (σ known)

$$\begin{aligned} n &= 50 \\ \bar{X} &= \text{R}2\,500 \\ \sigma &= \text{R}500 \end{aligned}$$

95% confidence interval for μ

$$\bar{X} \pm Z\frac{\sigma}{\sqrt{n}}$$
$$2500 \pm 1.96\,(70.7107)$$
$$[2361.41 \leq \mu \leq 2638.59]$$

We are 95% confident that the mean monthly rent will be between R2361.41 and R2638.59.

Option 4

QUESTION 5

Hypothesis test for proportion, Z_{STAT}

$$\begin{aligned} p &= \frac{36}{200} = 0.18 \qquad \alpha = 0.01 \\ H_0 &: \pi \geq 0.24 \\ H_1 &: \pi < 0.24 \rightarrow \text{ less than} \\ \text{Reject } H_0 \text{ if } Z_{STAT} &< -2.33 \\ \sigma_p &= 0.0302 \\ \therefore \quad Z_{STAT} &= \frac{0.18 - 0.24}{0.0302} = -1.99 \end{aligned}$$

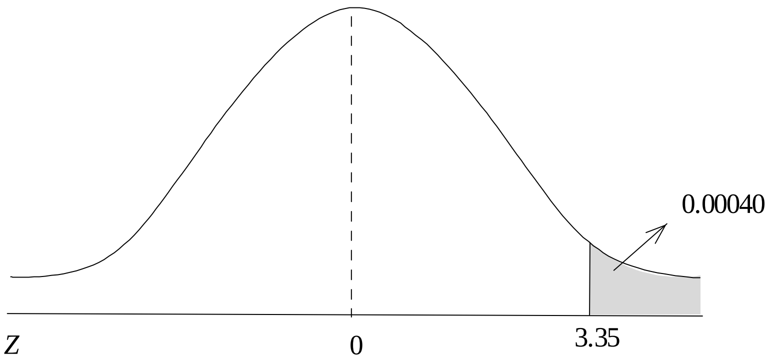
Option 5

QUESTION 6

Hypothesis test for μ the p–value

$$\begin{aligned} n &= 45 \quad \sigma = 4 \text{ years} \\ \bar{X} &= 30 \text{ years} \\ \alpha &: 0.01 \\ H_0 &: \pi \leq 28 \\ H_1 &: \mu > 28 \rightarrow \text{ upper tail test} \\ Z_{\text{stat}} &= \frac{30-28}{0.5963} \\ &= 3.35 \end{aligned}$$

So, $P(Z > 3.35) = 0.00040$ the p–value.



Option 4

QUESTION 7

Hypothesis testing: Rejection region using information in Question 6.

$$\begin{aligned} \alpha &= 0.01 \\ H_0 &: \mu \leq 28 \\ H_1 &: \mu > 28 \rightarrow \text{upper tail test} \\ & \quad \sigma \text{ known.} \\ \therefore & \quad Z_{\text{critical}} = Z_{0.01} = 2.33 \end{aligned}$$

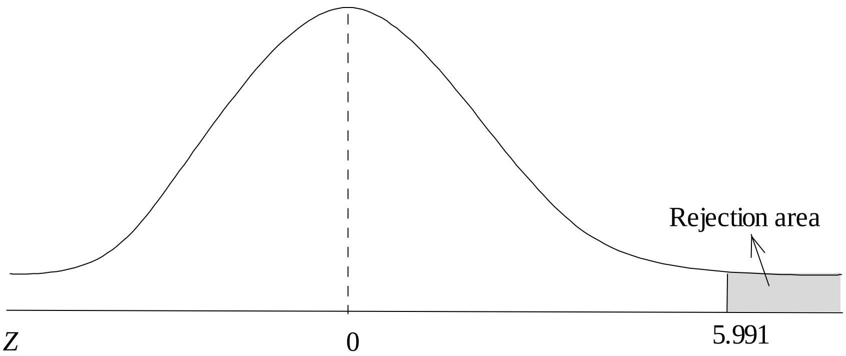
Reject H_0 if Z_{stat} is > 2.33.

Option 1

QUESTION 8

χ^2 hypothesis testing

$$\begin{aligned} \chi^2_{2;0.05} &= 5.991 \\ Z_{\text{stat}} &= 28.70 \end{aligned}$$



28.70 is greater than 5.991 therefore, reject H_0 at 5% level.

Option 3

QUESTION 9

Simple linear regression analysis

$$\begin{aligned} b_0 &= 854.10 \\ b_1 &= -4.33 \\ \therefore \quad \hat{y} &= 854.10 - 4.33x \end{aligned}$$

When $x = 100$ then $\hat{y} = 854.1 - 4.33\ (100) = 421.10$

When $x = 140$ then $\hat{y} = 854.1 - 4.33 = (140) = 247.90$

Option 4

QUESTION 10

Correlation analysis
Referring to Question 9, since

$$\begin{aligned} r &= -0.87 \quad \text{or} \quad -0.8663 \\ r^2 &= 75.05\% \end{aligned}$$

$r > 0$ inverse relationship..

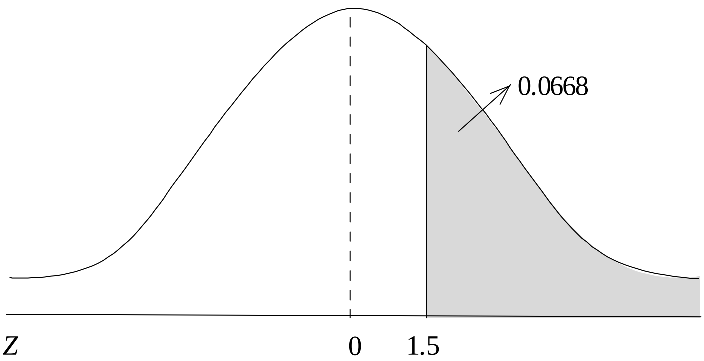
Option 1

QUESTION 11

Sample distribution of the mean μ.

$$\begin{aligned} \mu &= 75 \\ \sigma &= 12 \\ n &= 36 \end{aligned}$$

$$
\begin{aligned}
P\left(\bar{X} > 78\right) &= P\left(Z < \frac{78-75}{2}\right) \\
&= P\left(Z > 1.5\right)
\end{aligned}
$$



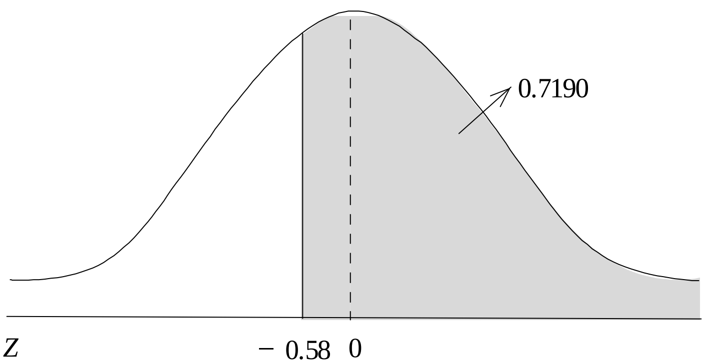
Option 4

QUESTION 12

Sampling distribution of the proportion.

$$
\begin{aligned}
n &= 200 \\
\pi &= 0.06
\end{aligned}
$$

$$
\begin{aligned}
& P\left(P > 0.58\right) \\
&= P\left(Z > \frac{0.58-0.6}{\sqrt{\frac{0.6(0.4)}{200}}}\right) \\
&= P\left(Z > -0.578\right) \\
&= 0.7190
\end{aligned}
$$



Option 5

QUESTION 13

χ^2 hypothesis testing. Conclusion

$$\begin{aligned} \alpha &= 0.10 \\ df &= (2-1)(2-1) = 1 \\ \therefore \quad \chi^2_{\text{critical}} &= \chi^2_{1;\ 0.10} = 2.706 \end{aligned}$$

	Yes	No	
Cold	10 (6.3)	110 (113.7)	120
Warm	11 (14.7)	269 (265.3)	280
	21	379	400

$$\begin{aligned} \chi^2_{\text{stat}} &= 2.173 + 0.1204 + 0.9313 + 0.0516 \\ &= 3.2763 \end{aligned}$$

since $\chi^2_{\text{stat}} > \chi^2_{\text{critical}}$ $\quad \therefore$ Reject H_0 ate 10% level.

Option 3

QUESTION 14

Simple linear regression and correlation analysis. Interpretation of r^2.

$$r^2 = 0.82 \quad \text{or} \quad 82\%$$

The interpretation is that 82% of the variation in the dependent variable can be explained by the variation in the independent variable.

Option 2

QUESTION 15

Correlation analysis. Determine r

$$\begin{aligned} SSR &= b_0 \sum Y_i + b_1 \sum X_i Y_i - \frac{\left(\sum Y_i\right)^2}{n} \\ &= -0.3517\,(59.97) + 0.1156\,(1496.69) - \frac{(59.97)^2}{30} \\ &= -21.0914 + 173.0174 - 119.88 \\ &= 32.046 \\ SST &= \sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n} = 155.3025 - \frac{(59.97)^2}{30} \\ &= 35.4225 \\ \therefore \quad & r^2 = \frac{SSR}{SST} = \frac{32.046}{35.4225} = 0.9047 \end{aligned}$$

Since b_1 is positive, r will be the positive square root of 0.9047.

$$r = 0.9512$$

Option 2