Assignment 9 of CISC 1006

ZHANG HUAKANG
DB92760
Computer Science
Faculty of Science and Technology

April 30, 2021

1

1.1

$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}$$

1.1.1

When $n = 64$

$$\sigma_{\overline{X}1}^2 = \frac{\sigma^2}{64}$$

When $n = 196$

$$\sigma_{\overline{X}2}^2 = \frac{\sigma^2}{169}$$

$$\begin{aligned}\frac{\sigma_{\overline{X}2}}{\sigma_{\overline{X}1}} =& \frac{\frac{\sigma^2}{169}}{\frac{\sigma^2}{64}} \\ =& \frac{64}{169} < 1 \\ \sigma_{\overline{X}2} - \sigma_{\overline{X}1} =& \frac{\sigma^2}{169} - \frac{\sigma^2}{64} \\ =& -\frac{105\sigma^2}{10816} \\ =& -\frac{1029}{3380} < 0\end{aligned}$$

The value becomes smaller.

1.1.2

When $n = 64$

$$\sigma_{\overline{X}3}^2 = \frac{\sigma^2}{784}$$

When $n = 196$

$$\sigma_{\overline{X}4}^2 = \frac{\sigma^2}{49}$$

$$
\begin{aligned}
\frac{\sigma_{\overline{X}4}}{\sigma_{\overline{X}3}} =& \frac{\frac{\sigma^2}{49}}{\frac{\sigma^2}{784}} \\
=& \frac{784}{49} > 1 \\
\sigma_{\overline{X}4} - \sigma_{\overline{X}3} =& \frac{\sigma^2}{49} - \frac{\sigma^2}{784} \\
=& -\frac{15\sigma^2}{784} \\
=& \frac{3}{5} > 0
\end{aligned}
$$

The value becomes greater.

1.2

1.2.1

$$
\begin{aligned}
\sigma_{\overline{X}}^2 =& \frac{\sigma^2}{n} \\
2^2 =& \frac{\sigma^2}{36} \\
\sigma =& 12
\end{aligned}
$$

1.2.2

$$
\begin{aligned}
\sigma_{\overline{X}}^2 =& \frac{\sigma^2}{n} \\
1.2^2 =& \frac{12^2}{n} \\
n =& 100
\end{aligned}
$$

2

2.1

Assume $X \sim Normal(250, 5)$

Calculate by *Excel*
=NORM.DIST(245,250,5/6,TRUE)

$$P(\overline{X} < 250) \approx 0.1587$$

2.2

$$
\begin{aligned}
n =& 36 \\
\mu_{\overline{X}} =& \mu \\
=& 250 \\
\sigma_{\overline{X}} =& \frac{\sigma}{\sqrt{n}} \\
=& \frac{5}{6}
\end{aligned}
$$

2.3

$\overline{X} \sim Normal(250, \frac{5}{6})$

Calculate by *Excel*
=NORM.DIST(245,250,5/6,TRUE)

$$P(\overline{X} < 250) \approx 9.8659 \times 10^{-10}$$

2.4

All of them are very close to 250

3

Calculate by *Excel*
=1-NORM.DIST(0.23,0.2,1/5000,TRUE)

Assume that $\mu = 0.2$, base on the ***Central Limit Theorem***, the probability that the mean $\mu_{\overline{X}}$ of $n = 50$ samples is equal to 0.23 or greater than 0.23 is

$$\begin{aligned}\overline{X} \sim& Normal(\mu_{\overline{X}}, \sigma_{\overline{X}})\\ \mu_{\overline{X}} =& \mu\\ =& 0.2\\ \sigma^2_{\overline{X}} =& \frac{\sigma^2}{n}\\ =& \frac{1}{5000}\\ P(X \geq 0.23) \approx& 0.0000\\ 2P(X \geq 0.23) \approx& 0.0000\end{aligned}$$

It is impossible. Thus, $\mu \neq 0.2$

4

4.1

$$\overline{X}_A - \overline{X}_B \sim Normal(\mu_{\overline{X}_A - \overline{X}}, \sigma_{\overline{X}_A - \overline{X}})$$

where

$$\begin{aligned}\mu_{\overline{X}_A - \overline{X}} =& \mu - \mu\\ =& 0\\ \sigma^2_{\overline{X}_A - \overline{X}_B} =& \frac{\sigma^2}{n} + \frac{\sigma^2}{n}\\ =& \frac{1}{18}\end{aligned}$$

Calculate by *Excel*
=NORM.DIST(-0.2,0,1/18,TRUE)*2

$$\begin{aligned}P(|\overline{X}_A - \overline{X}_B| \geq 0.2) =& 2P(\overline{X}_A - \overline{X}_B \leq -0.2)\\ \approx& 0.0003\end{aligned}$$

4.2

Yes, if the two machines are same, $|\overline{X}_A - \overline{X}_B| \geq 0.2$ is impossible, beacuse of $P(|\overline{X}_A - \overline{X}_B| \geq 0.2)$ is only 0.0003

5

$$MOE = z_\gamma \sqrt{\frac{\sigma^2}{n}}$$
$$n = \frac{\sigma^2 z_\gamma^2}{MOE^2}$$

With 95% confidence, $z_\gamma = 1.96$.

$$n \approx 177.6356$$

Since $n \in \mathbb{N}^+$

$$n = 178$$

With 99% confidence, $z_\gamma = 2.58$.

$$n \approx 307.7919$$

Since $n \in \mathbb{N}^+$

$$n = 308$$

6

$$MOE = z_\gamma \sqrt{\frac{\sigma^2}{n}}$$
$$n = \frac{\sigma^2 z_\gamma^2}{MOE^2}$$

With 95% confidence, $z_\gamma = 1.96$.

$$n \approx 164.8142$$

Since $n \in \mathbb{N}^+$

$$n = 165$$

With 99% confidence, $z_\gamma = 2.58$.

$$n \approx 285.5762$$

Since $n \in \mathbb{N}^+$

$$n = 286$$

7

Calculate by *Excel*
=AVEDEV(15, 7, 8, 95, 19, 12, 8, 22,14)
= 16.1728
=MEDIAN(15, 7, 8, 95, 19, 12, 8, 22,14)
= 14
=MODE.SNGL(15, 7, 8, 95, 19, 12, 8, 22,14)
= 8

=STDEV.P(15, 7, 8, 95, 19, 12, 8, 22,14)
= 26.1780 Since variance is big, average is not good, and since the number of data is only 9, too few, mode is not good too.
Thus, median is the best in the three of them.

8

8.1

8.1.1

Theorem Let $\mathbb{E}[X] = \mu$ and c be a constant. Then

$$\mathbb{E}[X + c] = \mathbb{E}[X] + c$$

Proof.

$$\begin{aligned}\mathbb{E}[X + c] =& \frac{1}{n}\sum_x (x + c) \\ =& \frac{1}{n}(\sum_x x + nc) \\ =& \frac{1}{n}\sum_x x + c \\ =& \mathbb{E}[X] + c\end{aligned}$$

□

Proof. Let $var(X) = \sigma$ and $\mathbb{E}[X] = \mu$. Let c be a constant.

$$\begin{aligned}var(X + c) =& \mathbb{E}[(X + c)^2] - \mathbb{E}[X + c]^2 \\ =& \mathbb{E}[X^2 + 2Xc + c^2] - (\mathbb{E}[X] + c)^2 \\ =& \mathbb{E}[X^2] + 2c\mathbb{E}[X] + c^2 - (\mathbb{E}[X]^2 + 2c\mathbb{E}[X] + c^2) \\ =& \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ =& var(X)\end{aligned}$$

If $c = -d$, then

$$var(X - d) = var(X)$$

□

8.1.2

Theorem Let $\mathbb{E}[X] = \mu$ and c be a constant. Then

$$\mathbb{E}[cX] = c\mathbb{E}[X]$$

Proof.

$$\begin{aligned}\mathbb{E}[cX] =& \frac{1}{n}\sum_x (cx) \\ =& c\frac{1}{n}(\sum_x x) \\ =& c\mathbb{E}[X]\end{aligned}$$

□

Proof. Let $var(X) = \sigma$ and $\mathbb{E}[X] = \mu$. Let c be a constant.

$$\begin{aligned}var(cX) =& \mathbb{E}[(cX)^2] - \mathbb{E}[cX]^2 \\ =& \mathbb{E}[c^2X^2] - (c\mathbb{E}[X])^2 \\ =& c^2\mathbb{E}[X^2] - c^2\mathbb{E}[X]^2 \\ =& c^2(\mathbb{E}[X^2] - \mathbb{E}[X]^2) \\ =& c^2 var(X)\end{aligned}$$

$\square$

8.2

8.2.1

$var(X)$
Calculate by *Excel*
=VAR.S(4, 9, 3, 6, 4,7)
$= 5.1$

8.2.2

$var(3X)$
Calculate by *Excel*
=VAR.S(12, 27, 9, 18, 12, 21)
$= 45.9 = 5.1 \times 3^2$

8.2.3

$var(X + 5)$
Calculate by *Excel*
=VAR.S(9, 14, 8, 11, 9, 12)
$= 5.1 = 5.1$

9

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi(n-1)$$

where $n = 25$

9.1

$$\begin{aligned} P(S^2 > 9.1) =& P(\frac{24 \times S^2}{6} > \frac{182}{5}) \\ \approx& 0.0502 \end{aligned}$$

9.2

$$\begin{aligned} P(3.462 \leq S^2 \leq 10.745) =& P(\frac{1731}{125} \leq S^2 \leq \frac{2149}{50}) \\ =& 0.0400 \end{aligned}$$