HOÀN THÀNH BẢNG SAU VỀ ƯỚC LƯỢNG KHOẢNG

Parameter	Conditions (if any)	The formula of Confident Interval		
		2 sided	Upper	Lower
Mean	$\sigma know$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=\frac{1-\alpha}{2}\to z_{\frac{\alpha}{2}}\Rightarrow\mathcal{E}=z_{\frac{\alpha}{2}}.\frac{\sigma}{\sqrt{n}}\Rightarrow(\bar{x}-\varepsilon;\bar{x}+\varepsilon)$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=0,5-\alpha\to z_{\alpha}\Rightarrow\mathcal{E}=z_{\alpha}.\frac{\sigma}{\sqrt{n}}\Rightarrow(\bar{x}-\varepsilon;+\infty)$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=0,5-\alpha\to z_{\alpha}\Rightarrow\mathcal{E}=z_{\alpha}.\frac{\sigma}{\sqrt{n}}\Rightarrow(-\infty;\bar{x}+\varepsilon)$
Mean	σ unknow; $n\geq 30$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=\frac{1-\alpha}{2}\to z_{\frac{\alpha}{2}}\Rightarrow\mathcal{E}=z_{\frac{\alpha}{2}}.\frac{s}{\sqrt{n}}\Rightarrow(\bar{x}-\varepsilon;\bar{x}+\varepsilon)$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=0,5-\alpha\to z_{\alpha}\Rightarrow\mathcal{E}=z_{\alpha}.\frac{s}{\sqrt{n}}\Rightarrow(\bar{x}-\varepsilon;+\infty)$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=0,5-\alpha\to z_{\alpha}\Rightarrow\mathcal{E}=z_{\alpha}.\frac{s}{\sqrt{n}}\Rightarrow(-\infty;\bar{x}+\varepsilon)$
Proportion		$\varphi\left(z_{\frac{\alpha}{2}}\right)=\frac{1-\alpha}{2}\to z_{\frac{\alpha}{2}}\Rightarrow\mathcal{E}=z_{\frac{\alpha}{2}}.\frac{\sqrt{f(1-f)}}{\sqrt{n}}\Rightarrow(f-\varepsilon;f+\varepsilon)$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=0,5-\alpha\to z_{\alpha}\Rightarrow\mathcal{E}=z_{\alpha}.\frac{\sqrt{f(1-f)}}{\sqrt{n}}\Rightarrow(f-\varepsilon;0)$	$\varphi\left(z_{\frac{\alpha}{2}}\right)=0,5-\alpha\to z_{\alpha}\Rightarrow\mathcal{E}=z_{\alpha}.\frac{\sqrt{f(1-f)}}{\sqrt{n}}\Rightarrow(0;f+\varepsilon)$
Variance	μ Unknow	$\alpha\to X_1^2=X^2_{(n-1;1-\frac{\alpha}{2})},X_2^2=X^2_{(n-1;\frac{\alpha}{2})}\Rightarrow(\frac{(n-1)s^2}{X_2^2};\frac{(n-1)s^2}{X_1^2})$	$\alpha\to X_1^2=X^2_{(n-1;\alpha)}\Rightarrow(\frac{(n-1)s^2}{X_2^2};+\infty)$	$\alpha\to X_1^2=X^2_{(n-1;1-\alpha)}\Rightarrow(0;\frac{(n-1)s^2}{X_1^2})$
2 MEANS	$\sigma know$	$(\overline{x_1}-\overline{x_2})\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$		
2 MEANS	σ unknow	$(\overline{x_1}-\overline{x_2})\pm t_{\frac{\alpha}{2},n_1+n_2-2}\sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}}$		
2 PROPORTIONS	$\widehat{p_1}-\widehat{p_2}\neq 0;1$	$(\widehat{p_1}-\widehat{p_2})\pm z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p_1}(1-\widehat{p_1})}{n_1}+\frac{\widehat{p_2}(1-\widehat{p_2})}{n_2}}$		
2 VARIANCE	σ unknow $n\geq 30$	$s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$	$s+z_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{2n}}$	$s-z_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{2n}}$

HOÀN THÀNH BẢNG SAU VỀ KIỂM ĐỊNH

Parameter	Conditions (if any)	Test statistic	Critical values		
			2Tails	Right Tail	Left tail
Mean	σ known	$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$	$\pm Z_{\frac{\alpha}{2}}$	Z_α	$-Z_\alpha$
Mean	σ unknown	$T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$	$\pm T_{\frac{\alpha}{2}}$	T_α	$-T_\alpha$
Proportion		$Z=\frac{\hat{P}-P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$	$\pm\frac{Z_\alpha}{2}$	Z_α	$-Z_\alpha$
Variance		$\chi^2=\frac{(n-1)S^2}{\sigma^2}$	$\pm\chi^2_{\frac{\alpha}{2}}$	χ^2_α	$-\chi^2_\alpha$
2 MEANS	σ known	$Z=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)$	$\pm Z_{\frac{\alpha}{2}}$	Z_α	$-Z_\alpha$
2 MEANS	σ unknown	$t=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{S_1^2/n_1+S_2^2/n_2}}\sim$	$\pm Z_{\frac{\alpha}{2}}$	Z_α	$-Z_\alpha$
2 PROPORTIONS		$Z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}}\sim N(0,1)$	$\pm Z_{\frac{\alpha}{2}}$	Z_α	$-Z_\alpha$
2 VARIANCE	Giả sử $\sigma_1^2=\sigma_2^2$	$T=\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ ~~$\sim t(n-1)$~~ $\sim t(n_1+n_2-2)$	$\pm T_{\frac{\alpha}{2}}$	T_α	$-T_\alpha$

HOÀN THÀNH BẢNG SAU VỀ REG.

Parameter	Conditions (if any)	The formula of confident interval		
		2 slides	upper	lower
Beta_0		$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \le \beta_1 \le \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$		
Beta_1		$\hat{\beta}_0 - t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right]} \le \beta_0 \le \hat{\beta}_0 + t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right]}$		
Mean reponse		$\hat{\mu}_{Y\|x_0} - t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]} \le \mu_{Y\|x_0} \le \hat{\mu}_{Y\|x_0} + t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]}$		
Future observation		$\hat{y}_0 - t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]} \le Y_0 \le \hat{y}_0 + t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]}$		

Parameter	Condition (if any)	The formula of CI		
		2 sided	Upper	Lower
Mean	σ ~~un~~known	$(\bar{x}-\varepsilon ; \bar{x}+\varepsilon)$ $\varepsilon = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$	$(-\infty ; \bar{x}+\varepsilon)$ $\varepsilon = z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$(\bar{x}-\varepsilon ; +\infty)$ $\varepsilon = z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}$
Mean	σ unknown	$(\bar{x}-\varepsilon ; \bar{x}+\varepsilon)$ $\varepsilon = t_{\frac{\alpha}{2};n-1} \cdot \frac{s}{\sqrt{n}}$	$(-\infty ; \bar{x}+\varepsilon)$ $\varepsilon = t_{\alpha;n-1} \cdot \frac{s}{\sqrt{n}}$	$(\bar{x}-\varepsilon ; +\infty)$ $\varepsilon = t_{\alpha;n-1} \cdot \frac{s}{\sqrt{n}}$
Proportion	n - large $n\hat{p} > 5$ $n(1-\hat{p}) > 5$	$(\hat{p}-\varepsilon ; \hat{p}+\varepsilon)$ $\varepsilon = z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$		
Variance		$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2};n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2};n-1}}$	$\sigma^2 \le \frac{(n-1)s^2}{\chi_{1-\alpha;n-1}}$	$\sigma^2 \ge \frac{(n-1)s^2}{\chi_{\alpha;n-1}}$

BÀI TẬP B

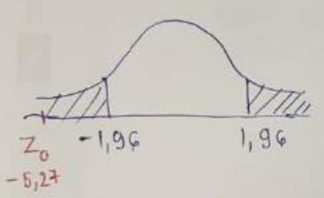
9.92

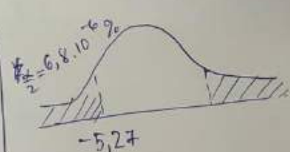
9.92 a) $\begin{cases} H_0 : p = 0.9 \\ H_1 : p \neq 0.9 \end{cases}$ $(\alpha = 0.05)$

$\hat{p} = \frac{850}{1000} = 0,85$

• $Z_0 = \frac{X - \mu}{\sigma} = \frac{X - n.p}{\sqrt{n.p(1-p)}} = \frac{850 - 1000.0,9}{\sqrt{1000.0,9.0,1}} = -5,27$

$P\text{value} = \min \alpha = 1,36.10^{-5}\,\%$





• $Z_0 = -5,72 < Z_{0,025} = -1,96$

$\Rightarrow$ Reject H_0

b) 95% CI for 2sides

$\hat{p} - Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$\Leftrightarrow 0,85 - 1,96 \cdot \sqrt{\frac{0,85(1-0,85)}{1000}} \leq p \leq 0,85 + 1,96 \cdot \sqrt{\frac{0,85(1-0,85)}{1000}}$

$\Leftrightarrow 0,828 \leq p \leq 0,872$

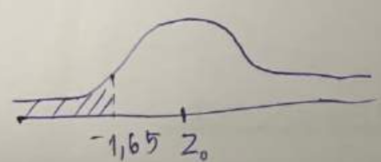
~~value~~ $p = 0,9 \notin CI\,(0,828 ; 0,872) \Rightarrow$ reject H_0

9.93

9.93 a) $\begin{cases} H_0: p = 0{,}03 \\ H_1: p < 0{,}03 \end{cases}$ $(\alpha = 0{,}05)$

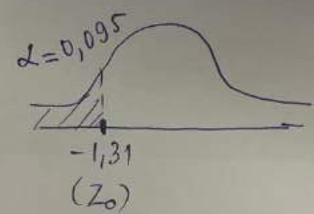
$\hat{p} = \frac{10}{500} = 0{,}02$

$$Z_0 = \frac{10 - 500 . 0{,}03}{\sqrt{500 . 0{,}03 . (1 - 0{,}03)}} = -1{,}31$$



• $Z_0 = -1{,}31 > Z_{0{,}05} = -1{,}65$

$\Rightarrow$ Fail to reject H_0



P-value = min α = 0,095

b) ~~CI~~ 95 % for one-side

$p < \hat{p} + Z_\alpha . \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$p < 0{,}02 + 1{,}65 . \sqrt{\frac{0{,}02(1-0{,}02)}{500}}$

$p < 0{,}0303$

• $p = 0{,}03 \in (-\infty ; 0{,}0303)$

$\Rightarrow$ Fail to reject H_0

10.32

a. Let X : weight of a paper sheet 1.

Y: weights of a paper sheet 2.

Using shapiro. test we can test the normality of the variabies X & Y in r software.

Null & Alternative hypothesis:

H_0 : The data is not mornally distributed. Vs :

H_a : The data is normally distributed.

R-code :

```
rm (list = ls ())
x = c(3,481,3.448,3.485,3.475, 3.472, 3.472,3.
    464.3.470, 3. 470, 3.477, 3.473, 3. 474);
y = c(3.268, 3.254, 256,3. 249, 3.241, 3.254,3.
    247,3.257,3.239,3.250, 3.258,3.239,3.
    245, 3.240, 3.254)
qqnorm (x)
qqnorm (y)
shapiro. test (x)
shapiro. test (y)
```

R- output

```
rm ( list = ls())
```

a. Let X : weight of a paper sheet 1.

Y: weights of a paper sheet 2.

Using shapiro.test we can test the normality of the variables X & Y in r software.

Null & Alternative hypothesis:

Ho : The data is not normally distributed. Vs :

Ha : The data is normally distributed.

R-code :

```
rm (list = ls ())
x = c(3.481,3.448,3.485,3.475,3.472,3.472,3.
    464,3.470,3.470,3.477,3.473,3.474);
y = c(3.268,3.254,256,3.249,3.241,3.254,3.
    247,3.257,3.239,3.250,3.258,3.239,3.
    245,3.240,3.254)
qqnorm (x)
qqnorm (y)
shapiro. test (x)
shapiro. test (y)
```

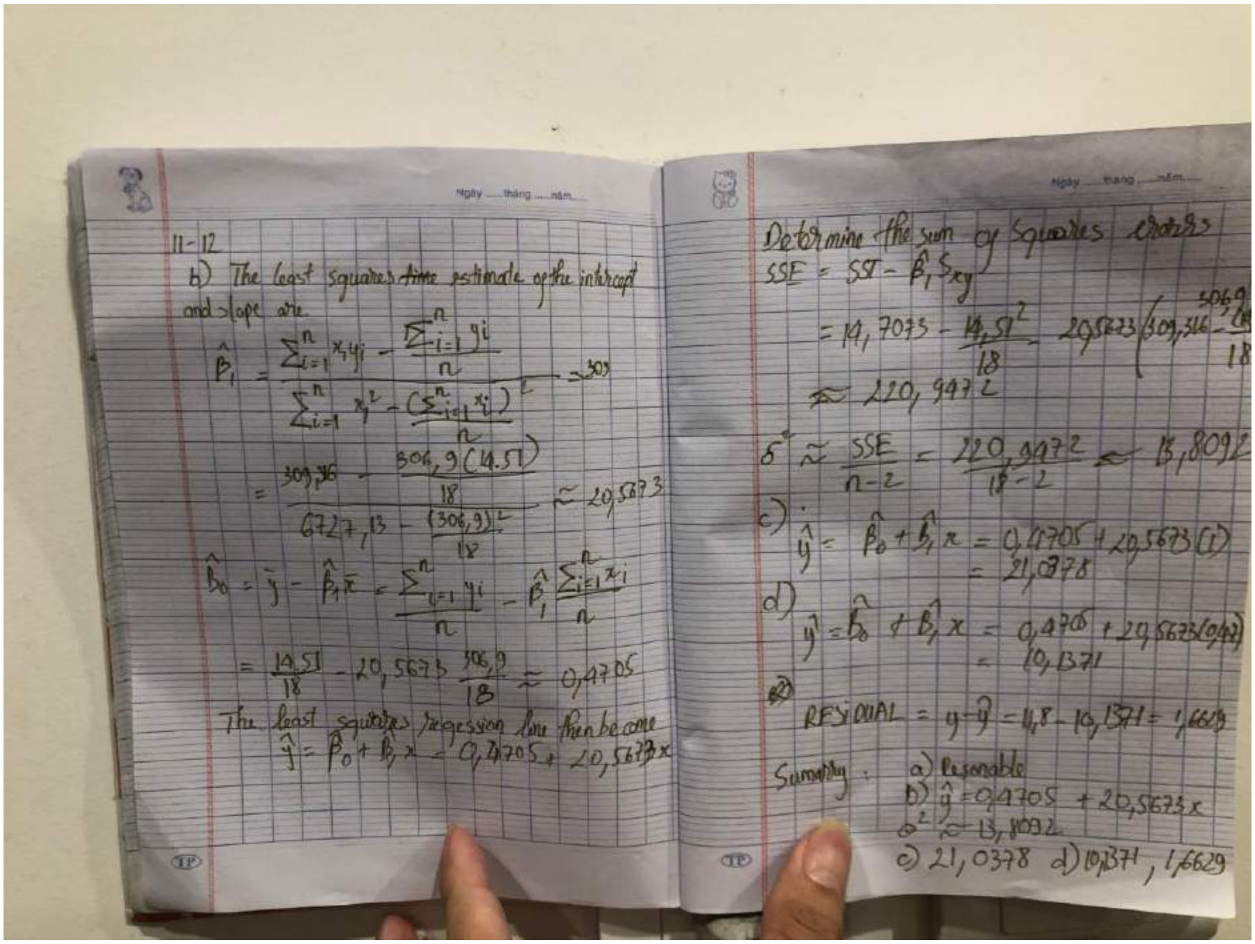
R- output

```
rm (list = ls())
```

```
>x=c(3.481,3.488,3.485,3.475,3.472,3.477,3.472,
      3.464, 3.472, 3.470, 3.470, 3.477, 3.473,
      3.474);
  y=c(3.258,3.254,3.256,3.249,3.241,3.254,
     3.247, 3.257, 3.239, 3.250, 3.258, 3.239,
     3.245, 3.240, 3.254);
>qqplot(x,y)
>rm(list=ls())
>x=c(3.481 3.488, 3.485, 3.475, 3.472, 3.477, 3.472,
      3.464, 3.472, 3.470, 3.470, 3.477, 3.473,
      3.474)
>y=c(3.258, 3.254, 3.256, 3.249, 3.241, 3.254,
      3.247, 3.257, 3.239, 3.250, 3.258, 3.239,
      3.245, 3.240, 3.254);
>qqnorm(x)
>qqnorm(y)
>shapiro test (x)
   Shapiro- Wilk normally test
```

KOKUYO

11.12



11-12

b) The least squares ~~time~~ estimate of the intercept and slope are

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} y_i}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}$$

$$= \frac{309{,}36 - \frac{306{,}9(14.51)}{18}}{6727{,}13 - \frac{(306{,}9)^2}{18}} \approx 20{,}5673$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \frac{\sum_{i=1}^{n} y_i}{n} - \hat{\beta}_1 \frac{\sum_{i=1}^{n} x_i}{n}$$

$$= \frac{14{,}51}{18} - 20{,}5673 \frac{306{,}9}{18} \approx 0{,}4705$$

The least squares regression line then become

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 0{,}4705 + 20{,}5673x$$

Determine the sum of squares errors

$$SSE = SST - \hat{\beta}_1 S_{xy}$$

$$= 14{,}7073 - \frac{14{,}51^2}{18} - 20{,}5673\left(309{,}316 - \frac{306{,}9 \ldots}{18}\right)$$

$$\approx 220{,}9472$$

$$\hat{\sigma}^2 \approx \frac{SSE}{n-2} = \frac{220{,}9472}{18-2} \approx 13{,}8092$$

c) $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 0{,}4705 + 20{,}5673(1) = 21{,}0378$

d) $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 0{,}4705 + 20{,}5673(0{,}47) = 10{,}1371$

RESIDUAL $= y - \hat{y} = 11{,}8 - 10{,}1371 = 1{,}6629$

Summary: a) Resonable

b) $\hat{y} = 0{,}4705 + 20{,}5673x$

$\hat{\sigma}^2 \approx 13{,}8092$

c) 21,0378 d) 10,1371 , 1,6629

8.40

a)

In the previously described way we construct a normal probability plot based on observations 8.24 8.25 8.20 8.23 8.24 8.21 8.26 8.26 8.20 8.25 8.23 8.23 8.19 8.28 8.24

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance σ^2, $\textbf{ a $100(1-\alpha)\%$ confidence interval on μ}$ is given by

$$\bar{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}, \tag{1}$$

where $t_{\frac{\alpha}{2},n-1}$ is the upper $100\frac{\alpha}{2}$ percentage point of the t distribution with $n-1$ degrees of freedom.

b)

The simple mean is

$$\bar{x} = \frac{1}{15}\sum_{i=1}^{15} x_i = 8.234,$$

and standard deviation is

$$s = \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}\right)^{\frac{1}{2}} = \left(\frac{\sum_{i=1}^{15}(x_i-8.234)^2}{14}\right)^{\frac{1}{2}} = 0.0253.$$

Now, we need to find a 95% CI on the population mean, then

$$\frac{\alpha}{2} = \frac{1-0.95}{2} = 0.025 \quad \Rightarrow \quad t_{\frac{\alpha}{2},n-1} = t_{0.025,14} = 2.145. \tag{2}$$

From Equations (1) and (2) we get

$$\begin{gathered} 8.234 - 2.145\frac{0.0253}{\sqrt{15}} \le \mu \le 8.234 + 2.145\frac{0.0253}{\sqrt{15}} \\ 8.234 - 2.145 \times 0.0065 \le \mu \le 8.234 + 2.145 \times 0.0065. \end{gathered}$$

Therefore, a 95% CI on the population mean is

$$\boxed{8.219 \le \mu \le 8.248.}$$

c.

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance σ^2,a $\textbf{$100(1-\alpha)\%$ upper confidence bound for μ is given by}$

$$\mu \le \bar{x} + t_{\alpha,n-1}\frac{s}{\sqrt{n}}. \tag{3}$$

Now, we need to find a 95% upper confidence bound on the population mean, then

$$\alpha = 1 - 0.95 = 0.05 \quad \Rightarrow \quad t_{\alpha,n-1} = t_{0.05,14} = 1.761. \tag{4}$$

From Equations (1) and (2) we get

$$\mu \le 8.234 + 1.761\frac{0.0253}{\sqrt{15}}$$
$$\mu \le 8.234 + 1.761 \times 0.0065.$$

Therefore, a 95% upper confidence bound on the population mean is

$$\boxed{\mu \le 8.2455.}$$

From subsection $b)$ and $c)$ we conclude that 95% two-sided confidence interval on the population mean is higher then 95% upper confidence bound on the population mean, because

$$t_{0.05,14} = 1.761 < 2.145 = t_{0.025,14}.$$