Confidence Intervals

Author: CHEA MAKARA Academy Year: 2022-2023
Got Inspired from Professors and Senior

1). (a): Suppose we construct a 99% confidence interval. What are we 90% confident about?
(b) : Which of the confidence interval in wider, 90% or 99%
(c) : In computing a confidence interval, when do you use the $t-$distribution and when do you use z, with normal approximation
(d) : How does the sample size affect the width of a confidence interval?

Solution

(a) : It means that $100(1-\alpha)\% = 99\%$ of the intervals constructed using that procedure will contain the true parameter.you have a only 1 percent chance of being wrong.
(b) : 99% is wider than 90%.
(c) : Using $t-$distribution and z, with the normal approximation.
$t-$distribution we use for find the confidence intervals for population mean when μ is known and variance σ^2 is unknown.
and $z-$distribution also $z-$score is computed for the population mean when we know the variance σ^2.
Above 30 degrees of freedom, the t-distribution roughly matches the z-distribution.
(d) : Increasing the sample size decreases the width of confidence intervals, because it decreases the standard error.

---

(2). Consider the probability statement .

$$P(-2.81 < Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 2.75) = k$$

where $\bar{X}$ is the mean of a random sample of size n from $N(\mu, \sigma^2)$ distribution with known σ^2
a). Find k.
we have $P(-2.81 < Z < 2.75) = k \iff \Phi(2.75) - \Phi(-2.81) = k$
$\implies 0.997 - 0.0025 = k \implies k = 0.9945$
Therefore $k = 0.9945$ ■
b). use this statement to find a confidence interval for μ
we have $P(-2.81 < \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 2.75) = 0.9945$
$\implies P(-2.81\dfrac{\sigma}{\sqrt{n}} < \bar{x}-\mu < 2.75\dfrac{\sigma}{\sqrt{n}} = 0.9945$
$\implies P(-2.81\dfrac{\sigma}{\sqrt{n}} - \bar{x} < -\mu < -\bar{x} + 2.75\dfrac{\sigma}{\sqrt{n}} = 0.9945$
so $\bar{x} - 2.75\dfrac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 2.81\dfrac{\sigma}{\sqrt{n}}) = 0.9945$
Therefore confidence interval for population μ is $[\bar{x} - 2.75\dfrac{\sigma}{\sqrt{n}}, \bar{x} + 2.81\dfrac{\sigma}{\sqrt{n}}]$ ■
c). What is the confidence level of this confidence interval?
The confidence level of this confidence interval is $P = 0.9945 = 99.45\%$ ■
d). Find the symmetric confidence for μ
Since $Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(\mu, \sigma^2) \implies Z$ is pivotal quantity
By symmetric principle $P(-z_{\alpha/2} < z < z_{\alpha/2}) = 1-\alpha = k$
$\implies P(\bar{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}})$
$\iff CI = [\bar{x} - z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}]$
we have $1-\alpha = k = 0,9945 \implies \alpha = 1 - 0.9945 = 0.0055$
then $\dfrac{alpha}{2} = 0.00275$ so $z-\alpha/2 = \Phi^{-1}(1-0.00275) = \Phi^{-1}(0.99725) = 2.77$
Therefore $CI = [\bar{x} - 2.77\dfrac{\sigma}{\sqrt{n}}, \bar{x} + 2.77\dfrac{\sigma}{\sqrt{n}}]$■

---

3). Let $X_1, X_2, \cdots, X_n$ be a random sample from an $N(\mu, \sigma^2)$, where the value of σ^2 is unknown .
a). Construct a $100(1-\alpha)\%$ confidence interval for μ when the value of σ^2 is known.

when we know the variance σ^2
we get CI for μ is given by $CI = \left[\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$
b). Construct a $100(1-\alpha)\%$ confidence interval for μ when the value of σ^2 is unknown.
we get CI for μ is given by $CI = \left[\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right]$ ■

---

4). A random sample of size 50 from a particular brand of $16 - ounce$ tea packets produced a mean weight of 15.65 ounces. Assume that the weight of these brands of tea packets are normally distributed with standard deviation of 0.59 ounce . Find a 95% confidence interval for the true mean μ.
Find 95% for the true mean μ
we have $X_1, X_2, \cdots, X_n \sim^{iid} N(\mu, \sigma^2)$
we obtain $CI = \left[\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$
For 95% of size 50 we have $1-\alpha = 0.95 \Longrightarrow \alpha = 1-0.95 = 0.05 \Longrightarrow \frac{\alpha}{2} = 0.025$
then $z_{\alpha/2} = \phi^{-1}(1-0.025) = \phi^{-1}(0.975) = 1.96$ and $\bar{x} = 15.65$, also $\sigma = 0.59$
Thus $CI = \left[15.65 \pm 1.96\frac{0.59}{\sqrt{50}}\right] = \left[15.48, 15.81\right]$

---

5). A researcher wishes to estimate within \$25 the average cost of postage a community college spends in one year. If she wishes to be 90% confident. how large of a sample will be necessary if the population standard deviation is \$80
Find the size of sample
We know that $n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$
Since $E = 25, \sigma = 80$and $1-\alpha = 0.90 \Longrightarrow \alpha = 1-0.90 = 0.10$
then $\alpha/2 = 0.05 \Longleftrightarrow z_{0.05} = \phi^{-1}(1-0.05) = \phi^{-1}(0.95) = 1.65$
$\Longrightarrow n = \left(\frac{1.65 \times 80}{25}\right)^2 = \left(\frac{132}{25}\right)^2 = 27.878 \approx 28$
Therefore $n = 28$ ■

---

6). A university dean wishes to estimate the average number of hours that freshmen study each week. The standard deviation from a previous study is 2.6 hours. How large a sample must be selected if he wants to be 99% confident of finding whether the true mean differs from the sample mean by 0.5 hour?
Find the large sample
We have $n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$
Since $E = 0.5, \sigma = 2.6, \alpha = 0.01 \Longrightarrow z_{\alpha/2} = z_{0.005} = \phi^{-1}(0.995) = 2.58$
Thus $n = \left(\frac{2.6 \times 2.58}{0.5}\right)^2 = 179.98 \approx 180$
Therefore $n = 180$ ■

---

7). In a large university, the following are the ages of 20 randomly chosen employees:

24	31	28	43	28	56	48	39	52	32
38	49	51	49	62	33	41	58	63	56

Assuming that the data come from a normal population, construct a 95% confidence interval for the population mean μ of the ages of the employees of this university. Interpret your answer.
Find the confidence interval for population μ
We have $CI = \left[\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right]$

Since $\bar{x} = \sum_{i=1}^{20} x_i = 40.5$ and $s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{20}(x_i - \bar{x})^2} = 21.5$
For 95% $\Longrightarrow \alpha = 0.05 \Longleftrightarrow \alpha/2 = 0.025$
$\Longrightarrow t_{0.025,19} = 2.0930$
$\Longrightarrow CI = \left[40.5 \pm 20930 \times \frac{21.5}{\sqrt{20}}\right] = \left[30.433, 50.567\right]$

Therefore we have 95% of the employees are the ages on $\Big[30.433, 50.567\Big]$ ■

---

8). A random sample of size 26 is drawn from a population having a normal distribution. The sample mean and the sample standard deviation from the data are given respectively, as $\bar{x} = -2.22$ and $s = 1.67$. Construct a 98% confidence interval for the population mean μ and interpret the result.
Construct the 98% confidence interval for population μ
We have $X_1, X_2, \cdots, X_{26} \sim^{iid} N(\mu, \sigma^2)$
We have $CI = \left[\bar{x} \pm t_{\alpha/2,n-1} \times \frac{s}{\sqrt{n}}\right]$
Since $\bar{x} = -2.22 \& s = 1.67$
For 98% $\Longrightarrow \alpha = 1 - 0.98 = 0.02 \Longleftrightarrow \alpha/2 = 0.01$
Then $t_{0.01,25} = 2.4851$
$\Longrightarrow CI = \left[-2.22 \pm 2.4851 \times \frac{1.67}{\sqrt{26}}\right] = \left[-3.033, -1.406\right]$
Therefore for random sample size 26 has 98% on the confidence interval $CI = \Big[-3.033, -1.406\Big]$ ■

---

9). A random sample from normal population yields the following 25 values:

90	87	121	96	106	107	89	107	83	92
117	93	98	120	97	109	78	87	99	79
104	85	91	107	89					

(a). Calculate an unbiased estimate μ of the population mean.
By previous exercise $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$
The unbiased estimator for μ is $\hat{\mu} = \bar{x} = \frac{1}{25}\sum_{i=1}^{25} x_i = 97.24$
Therefore the unbiased estimator $\mu = \bar{x} = 97.24$ ■
(b). Give approximation 99% confidence interval for the population mean .
Assume that σ^2 is unknown .
For 99% $\Longrightarrow \alpha = 1 - 0.99 = 0.01$
Then $\alpha/2 = 0.005 \Longrightarrow t_{0.005,24} = 2.7969$
Find sample standard deviation s, such that
$s = \sqrt{S^2}$ we know that $S^2 = \frac{1}{25-1}\sum_{i=1}^{25}(x_i - \bar{x})^2 = \frac{1}{24}(\sum_{i=1}^{25} x_i^2) - \frac{1}{24}(25\bar{x}^2) = 148.44$
$\Longleftrightarrow s = \sqrt{S^2} = \sqrt{148.44} = 12.18$
so $CI = \left[\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right] = \left[97.24 \pm 2.7969 \times \frac{12.18}{\sqrt{25}}\right]$
Therefore the 99% confidence interval is $\Big[90.42, 97.05]\Big]$ ■

---

10). The following data represent the rates (micrometers per hour) at which a razor cut made in the skin of anesthetized newts is closed by new cells.

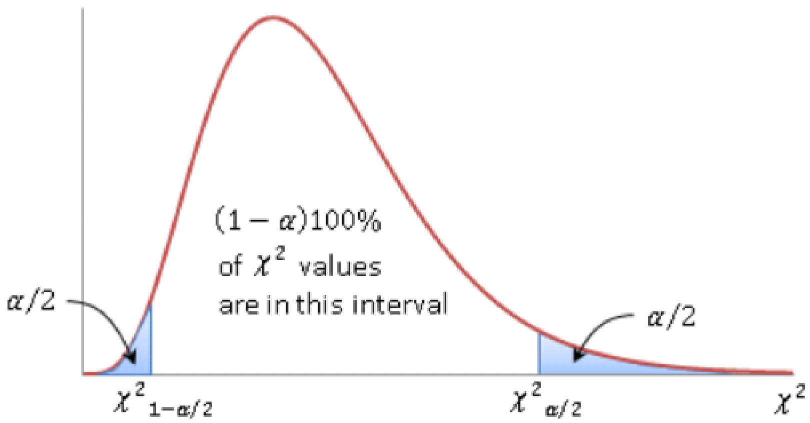
28	20	21	39	32	28	18	31	14	23
18	22	28	24	33	12	23	21	25	25

(a). Can we say that the data are approximately normal distribution?
We can say that the data are approximately normal distribution (Because its graph is approximately symmetric).
(b). Find a 95% confidence interval for population mean rate μ for the new cells to close a razor in the skin of anesthetized.
Assume its variance is unknown , so
$CI = \left[\bar{x} \pm t_{\alpha/2,n-1} \times \frac{s}{\sqrt{n}}\right]$
We have $\bar{x} = \sum_{i=1}^{20} x_i = 24$ and $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sqrt{\frac{1}{19}\sum_{i=1}^{20}(x_i - \bar{x})^2} = 6.52$

for 95% $\Longrightarrow \alpha = 1 - 0.95 = 0.05 \Longrightarrow \alpha/2 = 0.025$
Then $t_{0.025,19} = 2.0930$
we obtain $\mathrm{CI} = \left[24 \pm 2.0930 \times \dfrac{6.52}{\sqrt{20}}\right]$
Therefore $\mathrm{CI} = \left[20.9, 27.05\right]$ ■
(c). Find a 99% confidence interval for μ
For 99% confidence interval we get $\alpha = 1 - 0.99 = 0.01$
Then $t_{\alpha/2,n-1} = t_{0.005,19} = 2.8609$
so $\mathrm{CI} = \left[24 \pm 2.8609 \times \dfrac{6.52}{\sqrt{20}}\right]$
Therefore $\mathrm{CI} = \left[19.829, 28.17\right]$ ■
(d). Is the 95% CI wider or narrow than the 99%CI? Briefly explain why?
The 95% is narrow than 99% because 95% has 5% that can be wrong and 99% has only 1% can be wrong for confidence interval for population mean.■

---

11). Let $X_1, X_2, \cdots, X_n$ be a random sample from a normal distribution $N(\mu, \sigma^2)$, where the values of μ and σ^2 are unknown.
(a). Construct a $100(1-\alpha)\%$ confidence interval for σ^2 , choosing an appropriate pivot. Interpret its meaning.
Let $Q = \dfrac{(n-1)s^2}{\sigma^2} = \displaystyle\sum_{i=1}^{n}(\frac{x_i - \bar{x}}{\sigma})^2 \sim \chi^2(n-1)$
So, Q is pivotal quantity.
We have $1 - \alpha = P(a \le Q \le b) = P(a \le \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(x_i - \bar{x})^2 \le b)$
$\Longrightarrow 1 - \alpha = P\Big(\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{b} \le \sigma^2 \le \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{a}\Big)$



we have $a = \chi^2_{1-\alpha/2,n-1}$ and $b = \chi^2_{\alpha/2,n-1}$
then we obtain a$100(1-\alpha)\%$CI for σ^2 is
$\mathrm{CI} = \left[\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\chi^2_{\alpha/2,n-1}}, \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\chi^2_{1-\alpha/2,n-1}}\right] = \left[\dfrac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right]$ ■
(b). Suppose a random sample from a normal distribution gives the following summary statistics: $n = 21, \bar{x} = 44.3$ and $s = 3.96$. Using part (a), find a 90% confidence interval for σ^2. Interpret its meaning.
Find 90% confidence interval
we have $n = 21, s = 3.96$
and $\alpha = 1 - 0.90 = 0.1 \Longrightarrow \alpha/2 = 0.05$
For $\chi^2_{0.05,20} = 28.412$ and $chi^2_{1-0.005,20} = \chi^2_{0.95,20} = 10.851$
Then $\mathrm{CI} = \left[\dfrac{20(3.96)^2}{28.412}, \dfrac{(20)(3.96)^2}{10.851}\right] = \left[11.03, 28.90\right]$ ■

---

12). A random sample of 20 automobiles has a population by-product release standard deviation of 2.3 ounces when 1 gallon of gasoline is used. Find the 90% confidence interval of the population standard

deviation. Assume the variable is normally distributed.
Find 90% confidence interval for population standard deviation.
Assume that the population variance and mean are unknown.
We know that for confident interval of population varaince is

$$CI = 1-\alpha = P\Big(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{1-\alpha/2,n-1}}\Big)$$

$$\Longrightarrow CI = P\Big(\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{\alpha/2,n-1}}} \leq \sigma \leq \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{1-\alpha/2,n-1}}}\Big)$$

$$\Longrightarrow CI = \Big(\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}}, \sqrt{\frac{(n-1)s2}{\chi^2_{1-\alpha/2,n-1}}}\Big)\text{(for population standard deviation)}$$

we have $s = 2.3, n = 20\&\alpha = 0.1$
then $\chi^2_{0.05,19} = 30.144\&$
$\chi^2_{0.95,19} = 10.117$

we obtain $CI = \Big[\sqrt{\frac{19(2.3)^2}{30.144}}, \sqrt{\frac{19(2.3)^2}{10.117}}\Big] = \Big[1.826, 3.151\Big]$ ■

---

13). A random sample from a normal population yields the following 25 values:

90	87	121	96	106	107	89	107	83	92
117	93	98	120	97	109	78	87	99	79
104	85	91	107	89					

(a). Calculate an unbiased estimate of the population varaince.

By previous exercise we have $\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ is an unbiased estimator for population variance.

Since $\hat{\sigma^2} = \frac{1}{n}\Big[\sum_{i=1}^{n}x_i^2 - (n\bar{x}^2)\Big]$, we have

$\bar{x} = \frac{1}{25}\sum_{i=1}^{n}x_i = 97.21$ and $\hat{\sigma^2} = \frac{1}{25}(\sum_{i=1}^{25}x_i^2 - 25\bar{x}^2) = 142.40$

Therefore $\hat{\sigma^2} = 142.40$ ■

(b). Give approximation 99% confidence interval for the population variance.

We have $CI = \Big[\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{\alpha/2,n-1}}, \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{1-\alpha/2,n}}\Big] = \Big[\frac{(n)\sigma^2}{\chi^2_{\alpha/2,n}}, \frac{(n)\sigma^2}{\chi^2_{1-\alpha/2,n}}\Big]$

We have $S^2 = \hat{\sigma^2} = 142.40$ and $\alpha = 0.01 \Longrightarrow \alpha/2 = 0.005$ Then $\chi^2_{0.005,25} = 46.928$ and $\chi^2_{0.995,24} = 10.5$

so $CI = \Big[\frac{25(142.40)}{46.928}, \frac{25(142.40)}{10.5}\Big] = \Big[75.86, 339.04\Big]$ ■

(c). Interpret your results and state any assumptions you made in order to solve the problem.

---

14). In a random sample of 50 college seniors, 18 indicated that they planning to pursue a graduate degree. Find a 98% confidence interval for the true proportion of all college seniors planning to purse a graduate degree, and interpret the result, and state any assumption you have made.

Find the 98% confidence interval

Let p the true indicate that planning to pursue a graduate degree

then $Ci = \Big[\hat{p} \pm z_{\alpha/2}.\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\Big]$

we know that $\hat{p} = \frac{x}{n}$ (x number of succes)

$\Longrightarrow \hat{p} = \frac{18}{50} = 0.36$

for 98% $\Longrightarrow \alpha = 0.02 \Longrightarrow z_{\alpha/2} = z_{0.01} = \phi^{-1}(0.99) = 2.33$

so $CI = \Big[0.36 \pm 2.33\sqrt{\frac{0.36(1-0.36)}{50}}\Big] = \Big[0.201, 0.518\Big]$

Therefore 98% CI for p is $\Big[0.201, 0.518\Big]$ ■

---

15). It is believed that slightly over 40% of Cambodians own pets. How large a sample is necessary to

estimate the true proportion within 0.02 with 90% confidence?
Find the size of a sample
we have $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 \tilde{p}(1-\tilde{p})$ (round up to the whole number)
since $E = 0.02$, $\tilde{p} = 40\% = 0.4$ and $\alpha = 0.1 \Longrightarrow z_{\alpha/2} = z_{0.05} = \phi^{-1}(0.95) = 1.65$
Then we obtain $n = \left(\frac{1.65}{0.02}\right)^2 \times 0.4(0.6) = 1633.5 \approx 1634$
Therefore $n = 1634$■

---

16).In a random sample of 500 items from a large lot of manufactured items, there were 40 defectives.
(a). Find a 90% confidence interval for the true population of defectives in the lot.
Let p is the true population in the lot.
Then $CI = \left[\hat{p} \pm z_{\alpha/2}\sqrt{\frac{(\hat{p}(1-\hat{p})}{n}}\right]$
we have $\hat{p} = \frac{x}{n} = \frac{40}{500} = 0.08$
for 90% we get $\alpha = 1 - 0.90 = 0.1 \Longrightarrow z_{\alpha/2} = z_{0.05} = \phi^{-1}(0.95) = 1.65$
Hence , $CI = \left[0.08 \pm 1.65\sqrt{\frac{0.08(1-0.08)}{500}}\right] = \left[0.059, 0.100\right]$
Therefore a 90% confidence interval of p is $\left[0.059, 0.100\right]$ ■
(b). Is the assumption of normal approximation valid?
we can assume that it's approximately normal because it's a large sample and $np > 10, nq > 10$
(c). Suppose we suspect the another lot has the same proportion of defectives as in the first lot. What should be the sample size if we want to estimate the true proportion within 0.01 with 90% confidence.
how large of sample
We have $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 \tilde{p}(1-\tilde{p})$
we have $\alpha = 0.1 \Longrightarrow z_{0.05} = \phi^{-1}(0.95) = 1.65$ and $E = 0.01$
so $n = \left(\frac{1.65}{0.01}\right)^2 0.08(1-0.08) = 2003.76 \approx 2004$
Therefore for get a 90% CI with Error 0.01 the large sample is $n = 2004$ ■

---

17). A study found that 73% of randomly selected prekindergarten ages 3 to 5 whose mothers had a bachelor's degree of higher were enrolled in center-based early childhood care and education programs. How large a sample is needed to estimate the true proportion within 3 percentage points with 95% confidence? How large a sample is needed if you had no prior knowledge of the proportion?
Find large of sample
We have $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 \tilde{p}(1-\tilde{p})$
Since $E = 0.3$ and $\alpha = 0.05 \Longrightarrow z_{0.025} = \phi^{-1}(0.975) = 1.96$
$\tilde{p} = 0.73 \Longrightarrow (1-\tilde{p}) = 0.27$
Then $n = \left(\frac{1.96}{0.3}\right)^2 0.73(0.27) = 8.41 \approx 9$ So $n = 9$ for sample true proportion p ■

---

18). Let $X_1, X_2, \cdots, X_n$ be a random from a log-normal distribution $Log(m, \sigma^2)$, where m and σ^2 are parameters.
(a). Find MLE $\hat{m}$ for m. Is $\hat{m}$ efficient.
Find MLE $\hat{m}$ for m.
We have $X_1, X_2, \cdots, X_n \sim Log(m, \sigma^2)$
Then $f(x; \mu, \sigma^2) = \frac{1}{\sigma x_i\sqrt{2\pi}} e^{-\frac{(\ln x_i - m)^2}{2\sigma^2}}$
the likelihood function is $L(x; m; \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma x_i\sqrt{2\pi}} e^{-\frac{(\ln x_i - m)^2}{2\sigma^2}}$
The log likelihood function $\ln(L(x; m; \sigma^2)) = \sum_{i=1}^{n} \ln\left[\frac{1}{\sigma x_i\sqrt{2\pi}} e^{-\frac{(\ln x_i - m)^2}{2\sigma^2}}\right]$

$\Longrightarrow n\sum_{i=1}^{n}\ln\left[\frac{1}{\sigma x_i\sqrt{2\pi}}\right]-\sum_{i=1}^{n}\left(\frac{(\ln x_i-m)^2}{2\sigma^2}\right)$

Then $\frac{\partial \ln L(x;m;\sigma^2)}{\partial m}=\frac{\sum_{i=1}^{n}\ln x_i-nm}{\sigma^2}$

Set $\frac{\partial \ln(L(x;m;\sigma^2))}{\partial m}=0\Longleftrightarrow m=\frac{1}{n}\sum_{i=1}^{n}\ln x_i$

Therefore $\hat{m}=\frac{1}{n}\sum_{i=1}^{n}\ln x_i$ ■

Is it efficient ?

We have $V(\hat{m})=V(\frac{1}{n}\sum_{i=1}^{n}\ln x_i)=\frac{1}{n^2}\sum_{i=1}^{n}V(\ln x_i)$

We know that $V(\ln x)=E[(\ln x)^2]-[E(\ln x)]^2$

We have $E(\ln x_i)=\int_{-\infty}^{\infty}\ln x\frac{1}{\sigma x\sqrt{2\pi}}\exp(-\frac{(\ln x-m)^2}{2\sigma})dx$

Let $u=\ln x\Longrightarrow x=e^u$ and $du=\frac{1}{e^u}dx\Longleftrightarrow e^u du=dx$

we obtain $E(\ln x)=\int_{-\infty}^{\infty}u\frac{1}{e^u\sigma\sqrt{2\pi}}\exp(-\frac{(u-m)^2}{2\sigma})e^u du=\int_{-\infty}^{\infty}\frac{u}{\sigma\sqrt{2\pi}}e^{\frac{-(u-m)^2}{2\sigma}}du$

Suppose $m=0$

$E(\ln x)=\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}-(-\frac{u^2}{2\sigma})'e^{-\frac{u^2}{2\sigma}}du=0$

$E(\ln x)^2=\int_{0}^{\infty}\frac{(\ln x)^2}{\sigma x\sqrt{2\pi}}e^{\frac{-(\ln x-m)^2}{2\sigma}}dx=\int_{0}^{\infty}\frac{u^2}{\sigma\sqrt{2\pi}}e^{-\frac{(u-m)^2}{2\sigma^2}}du=\sigma^2$ (when $m=0$)

Then $V(\ln x)=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2=\frac{\sigma^2}{n}$

we also have $E(\frac{\partial^2\ln L(x;m;\sigma^2)}{\partial m^2})=-\frac{n}{\sigma^2}\Longrightarrow -\frac{1}{E(\frac{\partial^2\ln L(x;m;\sigma^2)}{\partial m^2})}=\frac{\sigma^2}{n}$

Since $V(\hat{m})=\frac{1}{I_n(m)}$

Thus $\hat{m}$ is an efficient. ■ (For second method we can transform it to normal distribution)

(b). Construct a 95% CI for m when $\sigma=1$ and $\sum_{i=1}^{25}\ln x_i=54.95$

We have 95% CI for $\hat{m}$ is defined by $\left[\hat{m}\pm z_{alpha/2}\times\frac{\sigma}{\sqrt{n}}\right]$

We have $\hat{m}=\frac{1}{25}\sum_{i=1}^{25}\ln x_i=54.95/25=2.198$, $\alpha=0.05\Longrightarrow z_{0.025}=\phi^{-1}(0.975)=1.96$

Hence $CI=\left[2.198\pm 1.96\frac{1}{\sqrt{25}}\right]=\left[1.806, 2.59\right]$

Therefore a 95% CI for $\hat{m}$ is $\left[1.806, 2.59\right]$ ■

---

19). Let $X_1, X_2, \cdots, X_n$ be a random sample from a population X with pdf

$$f(x;\theta)=\begin{cases}\frac{2x}{\theta^2} & \text{if} 0\le x\le\theta\\ 0 & \text{otherwise}\end{cases}$$

where $\theta>0$ is an unknown parameter.

(a). Determine an estimator T_n for θ, using the moment method. Is T_n efficient?.

By using moment method we have $E(x)=\frac{1}{n}\sum_{i=1}^{n}x_i$

we have $E(x)=\int_{0}^{\theta}x\frac{2x}{\theta^2}dx=2\int_{0}^{\theta}\frac{x^2}{\theta^2}dx=\frac{2}{3}\left[\frac{x^3}{\theta^2}\right]_0^{\theta}=\frac{2\theta}{3}$

Then we got $\frac{2\theta}{3}=\frac{1}{n}\sum_{i=1}^{n}x_i\Longrightarrow\hat{\theta}=\frac{3}{2\theta}\sum_{i=1}^{n}x_i$

Therefore $T_n = \hat{\theta} = \dfrac{3}{2n}\sum_{i=1}^{n} x_i$ ■

Is it efficient?

We have $V(T_n) = V(\dfrac{3}{2n}\sum_{i=1}^{n} x_i) = (\dfrac{3}{2n})^2 \sum_{i=1}^{n} V(x_i)$

We have $V(x_i) = E(x)^2 - [E(x)]^2$

Find $E(x^2) = \displaystyle\int_0^\theta \frac{2x^3}{\theta^2}dx = \Big[\frac{x^4}{2\theta^2}\Big]_0^\theta = \frac{\theta^2}{2}$

Then $V(x) = \dfrac{\theta^2}{2} - (\dfrac{2\theta}{3})^2 = \dfrac{\theta^2}{2} - \dfrac{4\theta^2}{9} = \dfrac{\theta^2}{18} \Longleftrightarrow V(T_n) = \dfrac{9}{4n} \times \dfrac{\theta^2}{18} = \dfrac{\theta^2}{8n}$

and $E(\dfrac{\partial^2 L(\theta)}{\partial\theta^2}) = \dfrac{\partial^2\Big(\sum_{i=1}^{n}\ln(\frac{2x}{\theta^2})\Big)}{\partial\theta^2} = \dfrac{2n}{\theta^2}$

$\Longrightarrow I_n(\theta) = -E(\dfrac{\partial^2 L(\theta)}{\partial\theta^2}) = -\dfrac{2n}{\theta^2}$

Since $V(T_n) > \dfrac{1}{I_n(\theta)}$

Therefore it is not an efficient.

(b). Find an unbiased estimator $\hat{\theta}_n$ for θ. Which estimator between T_n and $\hat{\theta}_n$ is more efficient?

We have log -likelihood function

$\ln L(\theta) = \displaystyle\sum_{i=1}^{n} \ln\frac{2x}{\theta^2} \Longrightarrow \frac{\partial^2 \ln L(\theta)}{\partial\theta^2} = \frac{2n}{\theta^2} > 0.$

so, $\hat{\theta} = max(x_i)$

Set $E(\hat{\theta}) = \theta \Longleftrightarrow max(x_i) = \theta$

Thus $\hat{\theta} = max(x_i) = \theta$ is an unbiased for θ

We have $V(\hat{\theta}) = V(max(x_i)) = 0$ $(max(x_i) = cte)$

Since $V(\hat{\theta}) < V(T_n)$

Therefore $\hat{\theta}$ is more efficient that T_n ■

(c). Find a 95% confidence interval for θ when $max(x_1, \cdots, x_{20}) = 5$

20). Let $X_1, X_2, \cdots, X_n$ be a random sample from a population X with pdf

$$f(x;\theta) = \begin{cases} \dfrac{x_0^{\frac{1}{\theta}}}{\theta x^{1+1\frac{1}{\theta}}} & \text{if} x > x_0 \\ 0 & x \leq x_0 \end{cases}$$

where $\theta > 0$ is an unknown parameter and $x_0 > 0$

(a). Find the MLE $\hat{\theta}_n$ for θ. Is $\hat{\theta}_n$ efficient?

(b). Find a 95% CI for θ when $\displaystyle\prod_{i=1}^{14} x_i = 2565^{14}$ and $x_0 = 1900$.

21). Let $X_1, X_2, \cdots, X_n$ be a random sample from a population X with pdf

$$f(x;\theta) = \begin{cases} \dfrac{1}{2\theta\sqrt{x}}e^{-\frac{\sqrt{x}}{\theta}} & \text{if} x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\theta > 0$ is an unknown parameter.

(a). Find the MLE $\hat{\theta}_n$ for θ. Is $\hat{\theta}_n$ efficient.

(b). Find a 90% CI for θ when $\displaystyle\sum_{i=1}^{20}\sqrt{x_i} = 47.4$