Interval Estimation

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with pdf $f(x,\theta)$, $\theta \in \Theta$.

(a) A random interval is an interval whose end points are random variables.

(b) A confidence interval for θ with confidence-coefficient $(1-\alpha)$; $0<\alpha<1$ is a random interval whose end points are statistics, say

$$L(X_1, X_2, \ldots, X_n) \text{ and } U(X_1, X_2, \ldots, X_n)$$

Such that $L(\underline{X}) \leq U(\underline{X})$, where

$$\underline{X} = (X_1, X_2, \ldots, X_n) \text{ and}$$

$$P(L(\underline{X}) \leq \theta \leq U(\underline{X})) = 1-\alpha.$$

Then $[L(\underline{X}), U(\underline{X})]$ is called $100(1-\alpha)\%$ confidence interval for θ.

Selected questions: 1, 4, 8, 11, 20, 21, 22

1) Confidence Interval for μ of the Normal distribution with known σ^2.

Step I. Choose a confidence level γ (95%, 99% or the like)

Step II. Determine the corresponding 'c'

γ	0.90	0.95	0.99	0.999
c	1.645	1.960	2.576	3.291

Step III Compute the mean $\bar{x}$ of the sample $x_1, x_2, \dots, x_n$

Step IV Compute $k = \frac{c\sigma}{\sqrt{n}}$. The confidence interval for μ is

$$CONF_{\gamma} \{ \bar{x} - k \leq \mu \leq \bar{x} + k \}.$$

Ex Determine a 95% confidence interval for the mean of a normal dist. with variance $\sigma^2 = 9$, using a sample of $n = 100$ values with mean $\bar{x} = 5$.

Soln. Step I $\gamma = 0.95$

Step II $c = 1.960$

Step III $\bar{x} = 5$ (given)

Step IV $k = \frac{c\sigma}{\sqrt{n}} = \frac{1.960 \times 3}{\sqrt{100}} = 0.588$.

Hence $\bar{x} - k = 5 - 0.588 = 4.412$

$\bar{x} + k = 5 + 0.588 = 5.588$

Thus, 95% confidence interval for μ is

$$CONF_{95} \{ 4.412 \leq \mu \leq 5.588 \}.$$

Ex Find a 95% confidence interval for μ of a normal ~~distribution~~ population with standard deviation 4, from the sample 30, 42, 40, 34, 48, 50.

Soln: step I $\gamma = 0.95$

step II $c = 1.960$

step III $\bar{x} = \frac{30+42+40+34+48+50}{6} = 40.66$

step IV $k = \frac{c\sigma}{\sqrt{n}} = \frac{1.96 \times 4}{\sqrt{6}} = 3.2006.$

Now, $(\bar{x}-k, \bar{x}+k)$ will be 95% confidence interval for μ, that is, $CONF_{95}\{37.459 \le \mu \le 43.8607\}$.

II Confidence Interval for μ of the Normal Distribution with unknown σ^2.

Step I. Choose a confidence Level γ (95%, 99% or the like)

Step II. Determine the solution 'c' of the equation

$$F(c) = \frac{1}{2}(1+\gamma)$$

from the table of t-distribution with $(n-1)$-degrees of freedom.

(Table A9 in Appendix ; where n = sample size)

Step III. Compute the mean $\bar{x}$ and the sample variance s^2 of the sample $x_1, x_2, \ldots, x_n$.

Step IV. Compute $k = \frac{cs}{\sqrt{n}}$.

The confidence interval is.

$$CONF_y \{ \bar{x} - k \leq \mu \leq \bar{x} + k \}.$$

Chi-Square Distribution

If $X \sim N(0,1)$, then $X^2 \sim \chi^2_{(1)}$.

(Chi-square with 1 degree of freedom)

Degree of Freedom:

The degree of freedom of distribution is sum of square of standard normal distribution.

E.g. ① If $X \sim N(0,1)$ then $X^2 \sim \chi^2_{(1)\cdot d.f.}$

② If $X_1, X_2 \overset{iid}{\sim} N(0,1)$ then $X_1^2 + X_2^2 \sim \chi^2_{(2)\cdot d.f}$

⋮

③ If $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(0,1)$ then $\sum_{i=1}^{n} X_i^2 \sim \chi^2_{(n)}$.

The pdf of chi-square distribution is given as

$$f_X(x) = \frac{x^{\frac{n}{2}-1} \, e^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right) \, 2^{\frac{n}{2}}} \; ; \; x > 0 \; ; \text{ This is } \chi^2_{(n)}.$$

Student's t-distribution.

If $X \sim N(0,1)$ and $Y \sim \chi^2_{(n)}$ and X and Y are independent, then

$$\frac{X}{\sqrt{Y/n}} \sim t_n.$$

Eg. Find independent measurements of the point of inflammation of Diesel oil gave the values 144 147 146 142 144. Assuming normality, determine a 99% confidence interval for the mean.

Solⁿ: step I. $\gamma = 0.99$

step II. $F(c) = \frac{1}{2}(1+\gamma)$

$$= \frac{1}{2}(1+0.99) = 0.995.$$

Here $n = 5$, so the value of 'c' for which $F(c) = 0.995$ with degree of freedom $(5-1) = 4$ is 4.60.

step III
$$\bar{x} = \frac{144 + 147 + 146 + 142 + 144}{5}$$

$$= 144.6$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$= \frac{1}{5-1}\Big[(144-144.6)^2 + (147-144.6)^2 + (146-144.6)^2 + (142-144.6)^2 + (144-144.6)^2\Big]$$

$$= 3.8$$

step IV
$$k = \frac{4.6 \times \sqrt{3.8}}{\sqrt{5}} = 4.01$$

The confidence interval is

$$CONF_{0.99}\{140.5 \leq \mu \leq 148.7\}.$$

III. Determination of a confidence interval for the variance σ^2 of a normal distribution, whose mean need not be known.

Step I. Choose a confidence level γ (95, 99% or like this)

Step II. Determine sol^ns. c_1 and c_2 of the equations

$$F(c_1) = \frac{(1-\gamma)}{2}, \quad F(c_2) = \frac{(1+\gamma)}{2}$$

from the table of the chi-square distribution with $n-1$ degrees of freedom (Table 10 in App. 5)

Step III. Compute $k_1 = \frac{(n-1)s^2}{c_1}$ and $k_2 = \frac{(n-1)s^2}{c_2}$.

The confidence interval is

$$\text{CONF}_\gamma\{k_2 \le \sigma^2 \le k_1\},$$

Eg. Determine a 95% confidence interval for the variance with sample

89, 84, 87, 81, 89, 86, 91, 90, 78, 89, 87, 99, 83, 89.

Sol^n step I. $\gamma = 0.95$

step II. For $n-1 = 13$, we have:

$$F(c_1) = \frac{1}{2}(1-0.95) = 0.025 \;\therefore\; \Rightarrow c_1 = F^{-1}(0.025) = 5.01$$

$$F(c_2) = \frac{1}{2}(1+0.95) = 0.975 \Rightarrow c_2 = F^{-1}(0.975) = 24.74$$

(From χ^2 table with 13 degree of freedom)

step III. $k_1 = \frac{13\,s^2}{c_1}$ and $k_2 = \frac{13\,s^2}{c_2}$.

where $s^2 = \frac{1}{13}\sum_{i=1}^{14}(x_i - \bar{x})^2$ and $\bar{x} = \frac{89+84+\cdots+89}{14}$

20\. If the weight X of bags of cement is normally distributed with a mean of 40 kg and a SD of 2 kg, how many bags can a delivery truck carry so that the probability of the total load exceeding 2000 kg will be 5%?

Sol: Given $E(X) = 40$ and $\sigma = 2 \Rightarrow V = \sigma^2 = 4$

To find $P(X > 2000) = 0.05$

$P(X \leq 2000) = 1 - 0.05 = 0.95$

Using the standardization process, we get

$$P\left(\frac{X - E(X)}{\sqrt{V(X)}} \leq \frac{2000 - 40n}{2\sqrt{n}}\right) = 0.95$$

(Since, there are 'n' bags, $E(X) = 40n$ and $V(X) = 4n$)

$$\Rightarrow P\left(Z \leq \frac{2000 - 40n}{2\sqrt{n}}\right) = 0.95$$

$$\Rightarrow \Phi\left(\frac{2000 - 40n}{2\sqrt{n}}\right) = 0.95$$

$$\Rightarrow \frac{2000 - 40n}{2\sqrt{n}} \geq 1.45$$

$$\Rightarrow 2000 - 40n - 2.9\sqrt{n} \geq 1$$

$$\Rightarrow \sqrt{n} = -7.11, 7.03$$

$$\Rightarrow \sqrt{n} = 7.03$$

$$\therefore n \approx 49$$

$\therefore$ We need 49 bags loaded on a truck, in order for the load to exceed ~~2000kg~~ 2000 kg.