

- An **interval estimator** involves statements like

$$P[\theta_1 < \theta < \theta_2] = 1 - \alpha, \alpha > 0 \quad (\text{A.1})$$

where θ is the unknown parameter,

$1 - \alpha$ is a given constant, called the **confidence coefficient**, or **degree of confidence**, and α is called the confidence level

θ_1, θ_2 are called **confidence limits**, and depend on the data or samples taken from the population and are thus R.V's.

For given α , the interval $\theta_1 < \theta < \theta_2$ is called a $(1 - \alpha) \times 100\%$ **confidence interval** computed from the selected sample.

Of course, the aim is a short interval with a high degree of confidence.

[e.g., it is better to have 95% confidence that the average life of a certain TV transistor is between 6 and 7 years -- then to be 99% confident that it is between 3 and 10 years].

The objective of interval estimation is to find the θ_1, θ_2 s.t. the length of the confidence interval $\theta_2 - \theta_1$ is minimized subject to the constraint (A.1). In general, this is not an easy task.

Suppose that we do not know the mean or the variance - and we wish to find their maximum likelihood estimators. We find that the estimator for the mean

$$\hat{m}_x = \frac{1}{n} \sum_{i=1}^n X_i = \text{sample mean} \quad (\text{A.2})$$

while the unbiased estimator for the variance (with unknown mean)

$$\hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \hat{m}_x)^2 \quad (\text{A.3})$$

The *maximum likelihood estimator of the variance* can be shown to be

$$[\hat{\sigma}_x^2]_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{m}_x)^2 \quad (\text{A.4})$$

which turns out to be a **biased** estimator of the variance since one can show

$$E [[\hat{\sigma}_x^2]_{\text{MLE}}] = \frac{(n-1) \sigma_x^2}{n} \quad (\text{A.5})$$

example The following 10 samples are drawn from a population that is known to be Gaussian :

7.31, 10.8 , 11.27 , 11.91 , 5.51 , 8.0 , 9.03 , 14.42 , 10.24 , 10.91

We can readily estimate the sample mean and sample variance :

$$\hat{m}_x = 9.94 \quad , \quad \hat{\sigma}_x^2 = 6.51$$

Note that these estimates for the mean and variance are R.V's.

If we draw another 10 samples, we will obtain a different mean and variance.

In fact, the 10 samples were actually drawn from a Gaussian with true mean = 10, true variance = 4 [i.e., the pdf $G(10,4)$]

Drawing different series of 10 samples from $G(10,4)$ we obtain the following estimates of the sample mean and variance :

Batch #	Sample Mean	Sample Variance
1	10.24	9.42
2	9.32	4.90
3	11.44	4.16
4	9.65	5.23
5	10.98	4.85
6	10.03	8.30
7	9.72	4.36
8	10.01	3.89
9	10.39	5.28
10	9.84	2.70

Can we say anything about the pdf for the mean and variance estimators?

(a) pdf for mean estimate

Since the mean estimate is just the arithmetic average of the samples drawn (which have Gaussian pdf's)

\Rightarrow mean estimate has also a Gaussian pdf,

with the expectation value of the sample mean $E[\bar{x}] = m_x$

and the variance of the sample mean $\text{var}[\bar{x}] = \frac{\text{Var}(X_i)}{n}$,

where
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i .$$

(A.6)

Hence we can now write down the pdf of the sample mean.

For a 10-sample batch [with true mean = 10 and true variance = 4], the variance of the sample mean = 4/10 so that the pdf:

$$f_{\bar{x}}(\alpha) = \frac{1}{(0.8 \pi)^{1/2}} \exp \left[- \frac{(\alpha - 10)^2}{0.8} \right], \text{ for 10-sample batch} \tag{A.7}$$

Now as the number of samples increases, the variance of the mean decreases, see (A.6), so that the pdf of the sample mean get narrower and narrower. Thus for a 100-sample batch, the variance of the sample mean = 4/100

$$f_{\bar{x}}(\alpha) = \frac{1}{(0.08 \pi)^{1/2}} \exp \left[- \frac{(\alpha - 10)^2}{0.08} \right], \text{ for 100-sample batch} \tag{A.8}$$

and these pdf's for the sample mean are shown in following **Transparency Plot** .

Finding Interval Estimates for the MEAN and the VARIANCE

1. Confidence Intervals for the Sample Mean : variance known

Reconsider the sketches for the pdf of the sample mean - assuming Gaussian samples. The confidence limits of (A.1) are shown as 2 vertical GREEN lines and have been placed symmetrically about the true mean.

Consider the R.V

$$Z = \frac{n^{1/2} (\bar{x} - m_x)}{\sigma_x} \tag{A.9}$$

Now $E[Z] = 0$, and the variance of $Z : \text{var}(Z) = 1$

- this is because the variance of the sample mean is the variance of each sample/n (see (A.6)).

Since the sample means are Gaussian, so must be Z . Thus

$$P \left[- z_c < Z = \frac{n^{1/2} (\bar{x} - m_x)}{\sigma_x} < z_c \right] = 1 - \alpha \tag{A.10}$$

gives the end points $\pm z_c$ of the confidence interval once the degree of confidence, $1 - \alpha$, is specified.

$$\text{But } P\left[-z_c < Z = \frac{n^{1/2} (\bar{x} - m_x)}{\sigma_X} < z_c\right] = \int_{-z_c}^{z_c} du \frac{\exp(-u^2/2)}{(2\pi)^{1/2}} = 1 - 2Q(z_c)$$

(A.11)

so that the $(1-\alpha)$ degree of confidence for the sample mean:

$$\begin{aligned} P\left[-z_c < Z = \frac{n^{1/2} (\bar{x} - m_x)}{\sigma_X} < z_c\right] &= P\left[-\frac{\sigma_X z_c}{n^{1/2}} < \hat{m}_x - m_x < \frac{\sigma_X z_c}{n^{1/2}}\right] \\ &= P\left[\hat{m}_x - \frac{\sigma_X z_c}{n^{1/2}} < m_x < \hat{m}_x + \frac{\sigma_X z_c}{n^{1/2}}\right] = 1 - \alpha = 1 - 2Q(z_c) \end{aligned}$$

(A.12)

Here m_x and σ_x are the true mean and variance, and n is the number of samples.

e.g., suppose we want a degree of confidence of 95 % = $1 - \alpha$ (%). [i.e., $1 - \alpha = 0.95$].

We can then determine the endpoint z_c from the tables of the Q-function:
 $\Rightarrow z_c = 1.96$.

- We can thus construct the following degree of confidence vs. endpoint table

1- α (in %)	endpoint z_c
90%	1.645
93%	1.812
95%	1.96
98%	2.326
99%	3.09

Note that if we want to get a higher level of confidence, then we must INCREASE the allowed probability interval for the R.V Z.

[• not unlike the uncertainty principle in Quantum Mechanics!]

example suppose we want to determine the confidence limit for a 99% degree of confidence for a 10-sample set with sample mean = 9.94.

Now for $1 - \alpha = 0.99 \Rightarrow z_c = 3.09$, so that from (A.12) with true variance of 4 (so $\sigma_x = 2$) we have

$$P \left[\hat{m}_x - \frac{\sigma_x z_c}{n^{1/2}} < m_x < \hat{m}_x + \frac{\sigma_x z_c}{n^{1/2}} \right] = 1 - \alpha = 0.99$$

$$= P \left[9.94 - \frac{2 \times 3.09}{10^{1/2}} < m_x < 9.94 + \frac{2 \times 3.09}{10^{1/2}} \right]$$

i.e., $P [7.986 < m_x < 11.894] = 0.99$ (A.13)

i.e., now suppose we had taken a 100-sample set whose sample mean = 9.96 [expect it to be different from the 10-sample set!].

Since n has increase, we will get a tighter (i.e., smaller) confidence interval :

$$P [9.342 < m_x < 10.578] = 0.99$$

Confidence limits for the Variance : Mean known

We first must determine the pdf for the sample variance, with the mean known. Normalize the maximum likelihood estimator for the variance (A.4) :

$$V \equiv \frac{n \hat{\sigma}_x^2}{\sigma_x^2} = \sum_{i=1}^n \left(\frac{X_i - m_x}{\sigma_x} \right)^2$$

(A.14)

This sum is composed of squares of zero-mean, unit-variance Gaussian R.V's

[It is true we do NOT know the variance σ_x^2 - but that will work out o.k in the end].

• How do we construct the pdf for V?

Well, let us first construct the pdf for the sum of squares of 2 indep. Gaussian R.V. with zero mean.

Consider the transformation $y = x^2$ for a Gaussian R.V X with zero mean, variance s^2 and determine the pdf $f_Y(y)$.

Now for given $y > 0$, there are 2 roots : $x_1 = y^{1/2}$ and $x_2 = -y^{1/2}$.

Moreover, $| dx_i/dy | = 1/2 y^{-1/2}$ so that $f_Y(y) = 0$, for $y \leq 0$; and for $y > 0$

$$f_Y(y) = \frac{1}{2 y^{1/2}} \left[\frac{\exp(-x^2/2\sigma^2)}{(2\pi\sigma^2)} \right]_{x_1=y^{1/2}} + \frac{1}{2 y^{1/2}} \left[\frac{\exp(-x^2/2\sigma^2)}{(2\pi\sigma^2)} \right]_{x_1=-y^{1/2}}$$

$$= \frac{\exp(-y/2\sigma^2)}{(2\pi\sigma^2 y)^{1/2}}$$

(A.15)

Now from (4.37) we know that the pdf of the sum of 2 indep. R.V is the convolution of the individual pdf's.

Hence for the sum of squares of 2 indep. Gaussian R.V. with zero mean,

$$\begin{aligned} f_{Z_2}(z) &= \int_{-\infty}^{\infty} dy f_Y(y) f_Y(z-y) \\ &= \int_0^z dy \frac{\exp[-y/2\sigma^2]}{(2\pi\sigma^2 y)^{1/2}} \cdot \frac{\exp[-\{z-y\}/2\sigma^2]}{(2\pi\sigma^2 \{z-y\})^{1/2}} \\ &= \frac{\exp(-z/2\sigma^2)}{2\pi\sigma^2} \int_0^z \frac{dy}{[y(z-y)]^{1/2}}, \quad z > 0 \end{aligned}$$

This integral can be transformed into the beta integral by the transformation $y \rightarrow u$, where $y = zu$:

$$\int_0^z \frac{dy}{[y(z-y)]^{1/2}} = \int_0^1 \frac{du}{[u(1-u)]^{1/2}} = \Gamma^2\left(\frac{1}{2}\right)$$

so that

$$f_{Z_2}(z) = \frac{\exp\left(-\frac{z}{2\sigma^2}\right)}{2\sigma^2}, \quad z > 0 \quad \text{----- an exponential pdf} \quad (\text{A.16})$$

The mean of this pdf is $2\sigma^2$, and the variance = $4\sigma^4$.

• We can now proceed to the sum of 3 indep. Gaussian R.V's

For simplicity, suppose the variance $s^2 = 1$. Then convoluting (A.15) with (A.16) we obtain

$$\begin{aligned} f_{Z_3}(z) &= \int_0^z du \frac{\exp(-u/2)}{(2\pi u)^{1/2}} \cdot \frac{\exp(-\{z-u\}/2)}{2} = \frac{e^{-z/2}}{(8\pi)^{1/2}} \int_0^z \frac{du}{u^{1/2}} \\ &= \frac{z^{1/2} \exp\left(-\frac{z}{2}\right)}{(2\pi)^{1/2}}, \quad z > 0 \end{aligned} \quad (\text{A.17})$$

which is the χ^2 -pdf of order 3 (i.e., with 3 degrees of freedom).

- By induction, one could deduce that the pdf for the sum of m squares of indep. Gaussian R.V's with zero mean and unit variance is a χ^2 -pdf of order m

$$f_{Z_m}(z) = \frac{z^{m/2 - 1} \exp\left(-\frac{z}{2}\right)}{\Gamma\left(\frac{m}{2}\right) 2^{m/2}}, \quad z > 0 \quad ; \quad \text{and } m = 1, 2, 3, \dots$$

(A.18)

- Now back to our original problem: how to estimate the variance from a population of Gaussian R.V's with known mean

$$V \equiv \frac{n \hat{\sigma}_x^2}{\sigma_x^2} = \sum_{i=1}^n \left(\frac{X_i - m_x}{\sigma_x} \right)^2$$

The confidence interval for the variance is defined by

$$P \left[\gamma < V = \frac{n \hat{\sigma}_x^2}{\sigma_x^2} < \delta \right] = 1 - \alpha$$

(A.19)

-- see **PLOT #2**, a plot of the χ^2 pdf, showing the tails to be integrated in finding the confidence intervals with the mean known -- where the parameters γ , α , δ are so chosen that the 2 tail areas (shaded) are equal and both equal = $\alpha/2$

(A.19) can also be written:

$$P \left[\frac{n \hat{\sigma}_x^2}{\delta} < \sigma_x^2 < \frac{n \hat{\sigma}_x^2}{\gamma} \right] = 1 - \alpha$$

(A.20)

example Let us reconsider that 10-sample batch drawn from a Gaussian with known mean = 10. We want to determine the 95% degree of confidence intervals for estimating the variance of the samples

7.31, 10.8, 11.27, 11.91, 5.51, 8.0, 9.03, 14.42, 10.24, 10.91
with known mean = 10.

solution : Now the most likelihood estimator for the variance (given mean)

$$\hat{\sigma}_x^2 = \frac{1}{10} \sum_{i=1}^{10} (X_i - 10)^2 = 5.87$$

For the degree of confidence $1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$

This is split between the two tails of the χ^2 -pdf for the sample variance.

From the Tables of the χ^2 -pdf, (with $nu = 10$)

$$P[V > \delta] = \alpha/2 = 0.025 \Rightarrow \delta = 20.483$$

$$P[V > \gamma] = 1 - \alpha + \alpha/2 = 1 - \alpha/2 = 0.975 \Rightarrow \gamma = 3.247$$

Thus, from (A.20)

$$P \left[\frac{10 \times 5.87}{20.483} < \sigma_x^2 < \frac{10 \times 5.87}{3.247} \right] = 0.95$$

$$\text{i.e., } P [2.87 < \sigma_x^2 < 18.08] = 0.95$$

(•)

Confidence Limits for the Variance: mean unknown

The unbiased estimator for the variance with mean unknown is given by

$$\hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \hat{m}_x)^2$$

so that on normalizing it, we consider the R.V W

$$W = \frac{(n-1) \hat{\sigma}_x^2}{\sigma_x^2} = \sum_{i=1}^n \left(\frac{X_i - m_x}{\sigma_x} \right)^2 \quad (\text{A.21})$$

If the R.V's X_i are Gaussian \Rightarrow the pdf for W is χ^2 - with $(n-1)$ degrees of freedom. [Proof is messy, and omitted]

- The only change from the case of mean known, (A.20), is to replace
 mean known \rightarrow mean unknown
 n \rightarrow n - 1

example: find the 95% degree of confidence intervals for the above example -- BUT with the mean unknown.

solution: Now we look up the tables of the χ^2 -pdf, but for $nu = 10 - 1 = 9$

Also the sample variance is now 6.51. (A.3) on using sample mean and $n - 1 = 9$

$$P[V > \delta] = 0.025 \Rightarrow \delta = 19.023$$

$$P[V > \gamma] = 0.975 \Rightarrow \gamma = 2.700$$

$$\text{Thus } P \left[\frac{9 \times 6.51}{19.023} < \sigma_x^2 < \frac{9 \times 6.51}{2.700} \right] = 0.95$$

$$\text{i.e., } P [3.08 < \sigma_x^2 < 21.70] = 0.95$$

(•)

Confidence Limits for the Mean : variance unknown

To determine the confidence intervals for the mean with the variance unknown, consider the R.V

$$T = \frac{n^{1/2} (\hat{m}_x - m_x)}{\hat{\sigma}_x} \quad (\text{A.22})$$

i.e., T is the ratio of 2 R.V's, and can be rewritten in the form

$$T = \frac{Z}{[W/(n-1)]^{1/2}} \quad (\text{A.23})$$

where Z is defined in (A.9) and W is defined in (A.21):

Z is a Gaussian R.V, with zero mean and unit variance ,

W is a χ^2 R.V

It can be shown that Z and W are statistically indep. and that T is distributed according to a Student's t pdf

$$f_T(t) = \frac{\Gamma\left[\frac{r+1}{2}\right]}{(\pi r)^{1/2} \Gamma\left[\frac{r}{2}\right]} \cdot \frac{1}{\left(1 + \frac{t^2}{r}\right)^{(r+1)/2}} \quad , \quad r = n - 1 \quad (\text{an integer}) \quad (\text{A.24})$$

The confidence interval is

$$P\left[\hat{m}_x - \frac{t_c \hat{\sigma}_x}{n^{1/2}} < m_x < \hat{m}_x + \frac{t_c \hat{\sigma}_x}{n^{1/2}}\right] = 1 - \alpha \quad (\text{A.25})$$

Question: Find 95% confidence intervals for the mean of the 10-sample 7.31, 10.8, 11.27, 11.91, 5.51, 8.0, 9.03, 14.42, 10.24, 10.91 if the variance is unknown.

Solution : To determine t_c , we use the Student t pdf for $r = n - 1 = 9$,

$$P[T > t_c] = P[T < -t_c] = \alpha/2 = 0.025 \quad \text{for the 95\% confidence interval} \\ \Rightarrow t_c = 2.263$$

For these 10-sample data, the sample mean and sample variance has been found:

$$\hat{m}_x = 9.94 \quad \text{and} \quad \hat{\sigma}_x = (6.51)^{1/2} = 2.55$$

so that

$$P\left[9.94 - \frac{2.263 \times 2.55}{10^{1/2}} < m_x < 9.94 + \frac{2.263 \times 2.55}{10^{1/2}}\right] = 0.95 \quad (\text{A.26})$$

$$\text{i.e., } P[8.11 < m_x < 11.77] = 0.95 \quad (\text{A.27})$$