Orion

PAGE:
DATE: 28/03/17.

STATISTICS

Sample / Random sample

A collection of independent and identically distributed random variables $X_1, X_2, \ldots, X_n$ is a random sample.
Here, 'n' is called the size of the random sample.

$\underbrace{\{x_1, x_2, \ldots, x_n\}}_{\text{Random Sample}}$ ← Realization of the random variables $X_1, X_2, \ldots, X_n$

Sample Mean

$\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ → also a random variable.

$\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ $\{x_1, x_2, \ldots, x_n\}$

$\bar{x}$ is a realization of $\bar{X}$.

Median, **Mode**

Sample Variance:

$\boxed{\dfrac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}} \xrightarrow{\text{realization}} \boxed{\dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$

→ Functions of Random Sample (random variables) → Statistic

Distribution of $\bar{X}$

Let $X_1, X_2, \ldots, X_n$ be a random sample.
Then the sample mean $\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$

Let $X_1, X_2, \ldots, X_n$ have mean μ and variance σ^2.

$\boxed{\begin{aligned} \text{Mean of } \bar{X} &= \mu \\ \text{Variance of } \bar{X} &= \sigma^2/n \end{aligned}}$

* $E(\bar{X}) = E\left(\dfrac{\sum X_i}{n}\right) = \dfrac{E(\sum X_i)}{n} = \dfrac{\sum \mu}{n} = \dfrac{n\mu}{n} = \mu$

* $Var(\bar{X}) = Var\left(\dfrac{\sum X_i}{n}\right) = \dfrac{1}{n^2} Var(\sum X_i) = \dfrac{1}{n^2}(\overbrace{\sigma^2 + \cdots + \sigma^2}^{n \text{ times}}) = \dfrac{n\sigma^2}{n^2} = \sigma^2/n$

→ independent ⇒ No covariance.

Orion

PAGE:
DATE: / /

$$\bar{X} = \frac{X_1}{n} + \frac{X_2}{n} + \cdots + \frac{X_n}{n}$$

$$\bar{X} = \frac{X_1}{n} + \cdots + \frac{X_n}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$ approximately. (by CLT)

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$ (Approximately by CLT)

03/04/17.

$X_1, X_2, \ldots, X_n$ is a random sample.

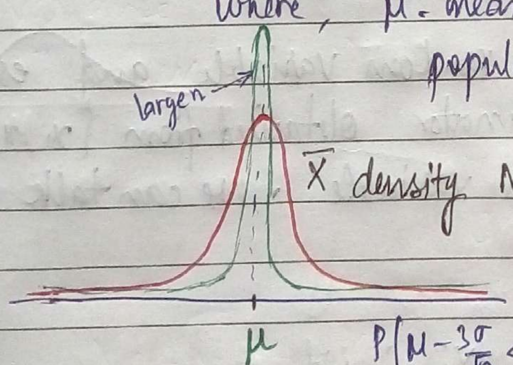
$\{x_1, x_2, \ldots, x_n\}$ ← random sample

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$ ← Samplemean

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$ ← Realization of samplemean.

CLT says, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ approximately

where, μ: mean & σ^2 = variance of X_i
population mean population variance.



n = very large, variance → very small
$n = 10,100$ $n = 10^{10}$.

$$P\left(\mu - \frac{3\sigma}{\sqrt{n}} < X < \mu + \frac{3\sigma}{\sqrt{n}}\right) = 99.99\%.$$

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$\bar{x}_1 = \frac{x_1' + \cdots + x_n'}{n}$$

$X_1, X_2, \ldots, X_n$

* $x_1, x_2, \ldots, x_n$

* $x_1', x_2', \ldots, x_n'$

$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 99.99\%.$$

$$\bar{X} = \frac{\Sigma X_i}{n}$$

$$E(\bar{X}) = \frac{\Sigma E(X_i)}{n} = \frac{n\mu}{n} = \mu$$

$$\boxed{E(\bar{X}) = \mu}$$

$\bar{X}$: estimator of μ
$\bar{x}$: estimate of μ
(realization of r.v. $\bar{X}$)

$\bar{X}$ is an unbiased estimator of μ.

Let $X_1, X_2, \ldots, X_n$ be a random sample.
($X_1\ X_2 \ldots X_n$ are all i.i.d with mean μ and variance σ^2).

To estimate the parameter μ. (Find μ.)
$\{x_1, x_2, \ldots, x_n\}$ is a realization of $X_1, \ldots, X_n$
by abuse of notation, we call $\{x_1\ x_2 \ldots x_n\}$ as a random sample

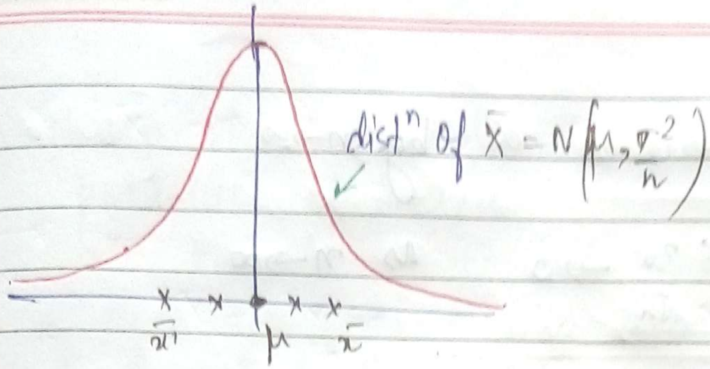
$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$ looks like a good choice to approximate μ.

$\bar{X}$ is called as an estimator of μ
$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$ is an estimate of μ

Note that estimator is a random variable and estimate is the realization of the estimator obtained from $\{x_1, x_2, \ldots, x_n\}$
Note that $\bar{X}$ is a random variable. We can talk of distribution of $\bar{X}$.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$ where X_i : iid
$E(X_i) = \mu$, $Var(X_i) = \sigma^2$

Using CLT,
$$\bar{X} = \frac{X_1}{n} + \ldots + \frac{X_n}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$ approximately.



$E(\bar{X}) = \mu$: unbiased property of $\bar{X}$.
(unbiased estimator)

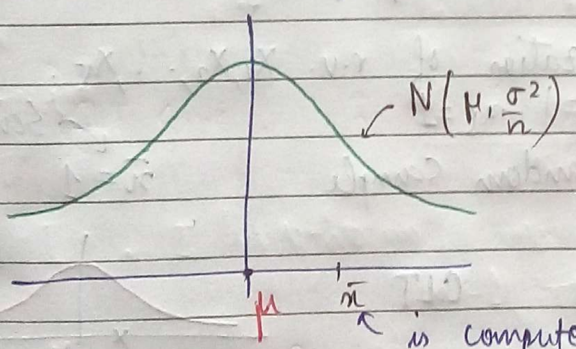
How to approximate/estimate μ??

1) Draw a random sample of size n from the population $x_1, x_2, \ldots, x_n$

2) Compute the estimate $\bar{x} = \dfrac{x_1 + x_2 + \ldots x_n}{n}$.

3) Declare $\bar{x}$ as a point estimate of μ.

From the theory of $\bar{X}$, we know that $\bar{x}$ is a realization of $\bar{X}$.



$\bar{x}$ is computed from the given sample.

Q How far away $\bar{x}$ is from μ??

$$P\left(\mu - \frac{3\sigma}{\sqrt{n}} \le \bar{X} \le \mu + \frac{3\sigma}{\sqrt{n}}\right) = 99.99\%$$

$\Rightarrow$ for 99.99% of the cases,

$$|\bar{x} - \mu| \le \frac{3\sigma}{\sqrt{n}}$$

We can take $\frac{3\sigma}{\sqrt{n}} \to 0$ by $n \to \infty$

$\Rightarrow |\bar{x} - \mu| \leq \frac{3\sigma}{\sqrt{n}} \to 0$ as $n \to \infty$

$\Rightarrow \boxed{\bar{x} \simeq \mu \quad \text{as } n \to \infty}$

Thus, what we have achieved by $n \to \infty$ is less variability in $\bar{X}$.

$Var(\bar{X}) = \frac{\sigma^2}{n} \to 0$ as $n \to \infty$.

$\Rightarrow$ Higher the sample size, "better" the estimate.

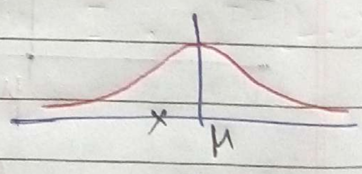
4/4/17.

$X_1, X_2, \ldots, X_5$ random sample
($X_1, X_2, \ldots, X_5$ are i.i.d)
$X_1, X_2, \ldots, X_5 \sim Bin(\underset{n}{3}, \underset{p.}{1/2})$ / $f(\theta)$ ← Normal(μ, σ^2)

$\{x_1, x_2, \ldots, x_5\}$: Realization of r.v. $X_1, X_2, \ldots, X_5$.
$x_1 = 1$ $x_2 = 0$ $x_3 = 2$ $x_4 = 1$ $x_5 = 1$
$\{1, 0, 2, 1, 1\}$ ← random sample

Sample Mean
$\bar{x} = 1$

$\bar{X} = \frac{\sum X_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ CLT



Problem: Estimator Problem.
Let $X_1 X_2 \ldots X_n$ be a random sample from an arbitrary probability distribution $f(\theta)$

$X_1, X_2, \ldots, X_n \overset{iid}{\sim} f(\theta)$

Here, the parameter(s) θ are to be estimated (guessed) from the random sample $X_1, X_2, \ldots, X_n$

How to do this ??

Form a suitable function of the random variables $X_1, X_2, \ldots X_n$
(This function is called as STATISTIC).
Eg - Mean.

This statistic is formed in such way that it "approximates" the parameter of interest.

— One criterion to check how good the approximation is to check if the statistic is unbiased.

Let T be a statistic formed from the sample $X_1, X_2, \ldots, X_n \sim f(\theta)$ to estimate the parameter θ. Then the statistic is called unbiased if $E(T) = \theta$.

(Another name for this statistic is 'Estimator')

A particular problem in estimation:

Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$ where the parameter μ is unknown and to be estimated and the parameter σ^2 is known.

Estimator:

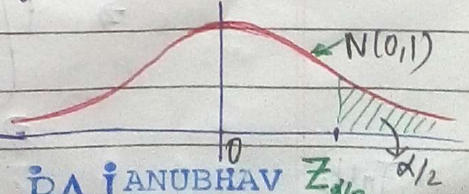
$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

(We have already proved using MGF)

$E(\bar{X}) = \mu$ (This is an unbiased estimator)

Note: This $\bar{X}$ is a point estimator (In other words, given $x_1, x_2, \ldots, x_n$ we get $\bar{x}$ and we say $\bar{x}$ is an estimate for the population mean).

Instead of giving one real number $(\bar{x})$ as a point estimate, we can compute an interval I from the sample such as $P(\mu \in I) = 95\%$. This strategy is called interval estimation and I is called as Confidence Interval for μ



$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

By standardization, $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

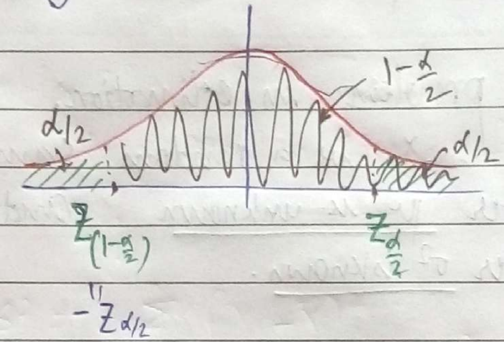
Consider a real number $\alpha \in [0,1]$, $\boxed{100(1-\alpha)\%}$

$\alpha = 0.01, 0.05, 0.1$

Define $Z_{\alpha/2} \in \mathbb{R}$ such that

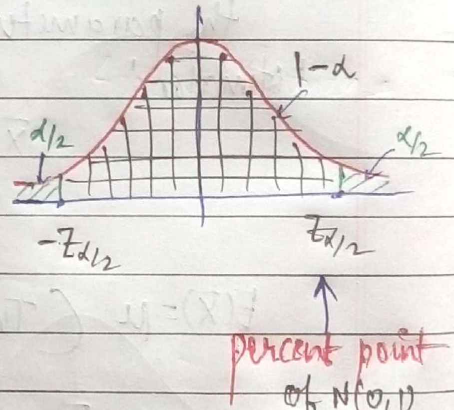
$$P(Z \geq Z_{\alpha/2}) = \alpha/2$$

If you are looking at CDF $N(0,1)$ table, $\Phi(Z_{\alpha/2}) = 1 - \frac{\alpha}{2}$



$Z \sim N(0,1)$

$$P\left(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}\right) = 1-\alpha$$



Note.

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

For a given value of α, construct an interval I, $I = [L, U]$
such that $P(\mu \in I) = 1-\alpha$

Since, $Z \sim N(0,1)$, we can write

$$P\left(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(-Z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$\Rightarrow P(\mu \in I) = 1-\alpha$ where $L = \bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$,

$U = \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Table

CDF Percent Point (Area Right to X) ~~CDF~~

Exercise! Let (−2.1, 1.9, 1.6, 0.3, 0.4, −0.2, −1.2, −0.03, 0.7, −0.6) be a random sample from $N(\mu, 1)$

- compute a point estimator of μ.
- Construct 95% confidence interval for μ.

* $\bar{x} = \dfrac{\Sigma x_i}{10} = \dfrac{0.77}{10} = 0.077$

$\bar{x}$ ← Point estimator; 0.077 ← point estimate

* 95% confidence interval ⇒ $100(1-\alpha)\% = 95\%$ ⇒ $\alpha = 0.05$

$\Phi(z_{\alpha/2}) = 0.975 \Rightarrow z_{\alpha/2} = 1.96$ ($z_{0.025}$)

$\sigma = 1$.

$\Rightarrow L = \bar{x} - z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}} = 0.077 - 1.96 \cdot \dfrac{1}{\sqrt{10}} = -0.5428$

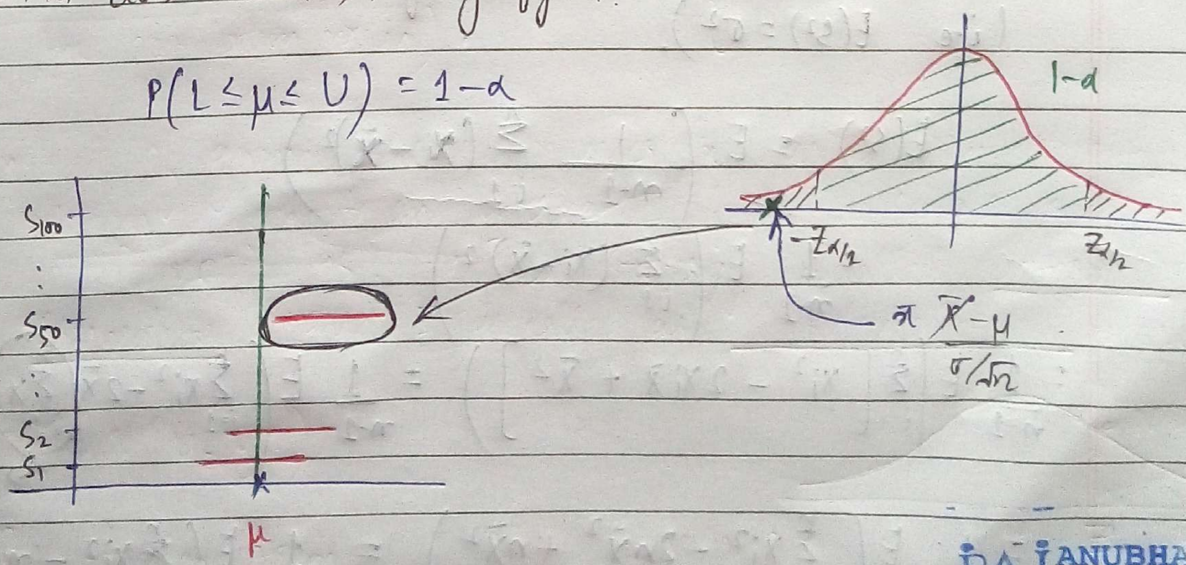
$\Rightarrow U = \bar{x} + z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}} = 0.077 + 1.96 \cdot \dfrac{1}{\sqrt{10}} = 0.6968$

$\boxed{I = [-0.5428, 0.6968]}$

Some observations on CI (confidence Interval)

1) What does $100(1-\alpha)\%$ signify ??

$P(L \le \mu \le U) = 1 - \alpha$



Orion

PAGE:
DATE: / /

2) Length of the confidence interval.

$$\boxed{U-L = 2 Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}}$$

For given σ & n, length depends on $Z_{\alpha/2}$ in turn on α.
Smaller the α, bigger the $Z_{\alpha/2}$.

Computer exercises. - large n.

Estimation Problem :
Let $X_1, X_2, \dots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Find an unbiased estimator of σ^2.

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1} \quad , \quad \text{Sample Variance}$$

$\{x_1, \dots, x_n\}$: data from $N(\mu, \sigma^2)$

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

S^2 is an unbiased estimator of σ^2.
(i.e $E(S^2) = \sigma^2$).

$$E(S^2) = E\left(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2\right)$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} (X_i - \bar{X})^2\right)$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n}\left[X_i^2 - 2X_i\bar{X} + \bar{X}^2\right]\right) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - 2\bar{X}\underbrace{\sum_{i=1}^{n} X_i}_{n\bar{X}} + \bar{X}^2 \sum_{i=1}^{n} 1\right)$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$$

$\sigma^2 = E(x_i^2) - \{E(x_i)\}^2 = E(x_i^2) - \mu^2$

$\Rightarrow E(x_i^2) = \sigma^2 + \mu^2$

Orion PAGE: DATE: / /

$$= \frac{1}{n-1}\left[E\left(\sum_{i=1}^{n} x_i^2\right) - n\,E\left(\bar{x}^2\right)\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} E(x_i^2) - n \cdot E\left(\bar{x}^2\right)\right]$$

$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\mu^2+\sigma^2\right) - n\left(Var(\bar{x}) + [E(\bar{x})]^2\right)\right]$$

$$= \frac{1}{n-1}\left[\cancel{n\mu^2} + n\sigma^2 - n\left[\frac{\sigma^2}{n} + \cancel{\mu^2}\right]\right]$$

$$= \frac{1}{n-1}\left[n\sigma^2 - \sigma^2\right] \Rightarrow \boxed{E(s^2) = \sigma^2}$$

$\bar{x} = 0.077$

x_i	-2.1	1.9	1.6	0.3	0.4	-0.2	-1.2	-0.03	0.7	-0.6
$(x_i-\bar{x})^2$	4.739	3.323	2.3195	0.0497	0.1043	0.0767	1.6307	0.01145	0.388	0.4583

$\frac{(x_i-\bar{x})^2}{9} = 1.4556 =$ Variance estimate.

ANUBHAV JAIN NOTES

Sample Variance

Let $X_1, X_2, \dots, X_n$ be a random sample from $N(\mu, \sigma^2)$

$$\left(X_1, X_2, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)\right)$$

$\{x_1, x_2, \dots, x_n\}$: realization of $\{X_1, X_2, \dots, X_n\}$

$$\left\{ \begin{aligned} &\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \text{ is an unbiased estimator of } \mu. \\ &S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \text{ is an unbiased estimator of } \sigma^2 \end{aligned} \right.$$

point estimators.

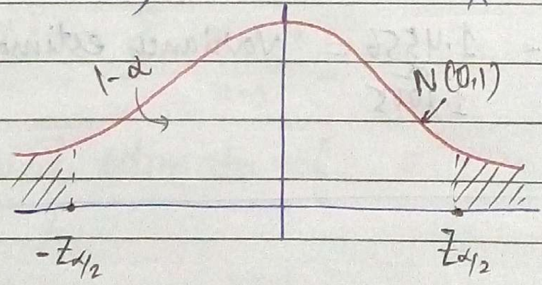
Interval estimation: Confidence Interval. for mean μ when variance σ^2 is known.

Estimate μ (C.I.) for a normal population when σ^2 is known.

true parameter to be estimated

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$



$$P\left(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\right) = 1 - \alpha$$

$$\Rightarrow P\left(-Z_{\alpha/2} \le \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \le Z_{\alpha/2}\right) = 1 - \alpha$$

$$\Rightarrow P\left(\underline{L} \le \mu \le \underline{U}\right) = 1 - \alpha$$

Estimation Problem 2:

Let $X_1, X_2, \ldots X_n$ be a random sample from $N(\mu, \sigma^2)$. Here the estimation problem involves estimation of σ^2.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$ is an unbiased estimator of σ^2.

(But this s^2 is a point estimate).

Next Step: Build a confidence interval for σ^2.

How to do this??

→ Construct a statistic involving s^2 & σ^2

→ Find the distribution of this statistic.

→ Find the percent points of this distribution.

→ Write the probability statement.

→ Algebraic manipulation to find the confidence interval.

Step 1

step 1 → Required statistic is $\frac{(n-1)s^2}{\sigma^2}$

step 2 → The distribution for this statistic is χ^2 (Chi squared) with $(n-1)$ degrees of freedom.

Sampling distribution

Statistic: $$\frac{(n-1)s^2}{\sigma^2} = \frac{\cancel{(n-1)}}{\sigma^2} \cdot \frac{1}{\cancel{n-1}} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$\boxed{\frac{(n-1)s^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X})^2}$$

χ^2 with k-degrees of freedom

Let $Z_1, Z_2, \ldots, Z_k$ be iid (Standard Normal) $N(0,1)$ (r.v.)

Define $Z_1^2 + Z_2^2 + \ldots + Z_k^2$ follows chi-squared distribution with k-degrees of freedom.

Typically χ^2_k denotes a chi-squared random variable with k dof.

If $\chi^2_{k_1}$ & $\chi^2_{k_2}$ are independent χ^2 r.v.s with k_1 and k_2 dof's respectively. then $\chi^2_{k_1} + \chi^2_{k_2} = \chi^2_{k_1+k_2}$

$$\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}\left(x_i-\bar{x}+\bar{x}-\mu\right)^2$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{i=1}^{n}(\bar{x}-\mu)^2 + 2\underbrace{\cancel{\sum_{i=1}^{n}(x_i-\bar{x})}}_{0}(\bar{x}-\mu)$$

$$\Rightarrow \boxed{\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{i=1}^{n}(\bar{x}-\mu)^2}$$

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

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

$$\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 = \frac{(n-1)s^2}{\sigma^2} + \frac{(\bar{x}-\mu)^2}{\sigma^2/n}$$

$\frac{x_i-\mu}{\sigma} \sim N(0,1)$

$\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 \sim \chi^2_n$ because of independence of x_i's

$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$\frac{(\bar{x}-\mu)^2}{\sigma^2/n} \sim \chi^2_1$

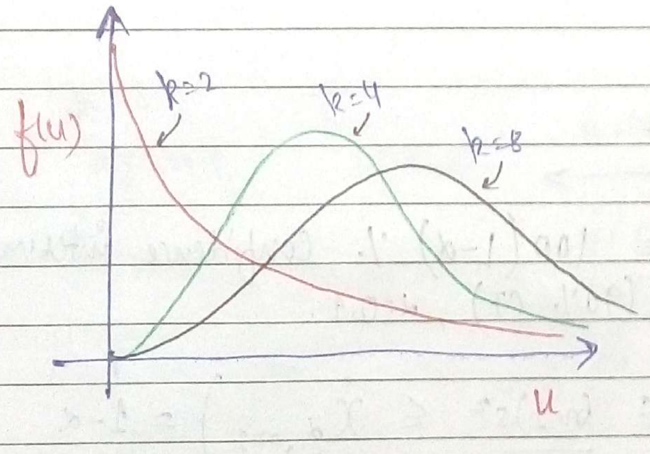
Then additive property of χ^2 says $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

Orion
PAGE:
DATE: 11 / 4 / 17

Let $Z_1, Z_2, \ldots, Z_k$ be independent std. normal r.v.
Then the random variables
$$\chi^2 = Z_1^2 + Z_2^2 + \ldots + Z_k^2$$
has the probability density function
$$f(u) = \frac{1}{2^{k/2}\,\Gamma[k/2]} u^{(k/2 - 1)} e^{-u/2} \qquad u > 0$$
$$= 0 \qquad \text{Otherwise}$$

The name of distribution is chi-squared distribution.
$\chi^2 \sim \chi^2_k$ (χ^2 follows chi-squared distribution with k degrees of freedom).
random variable
distribution

$\mu = k \qquad \sigma^2 = 2k \qquad$ for χ^2_k



Shape of pdfs for diff k.

Ex. $\dfrac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

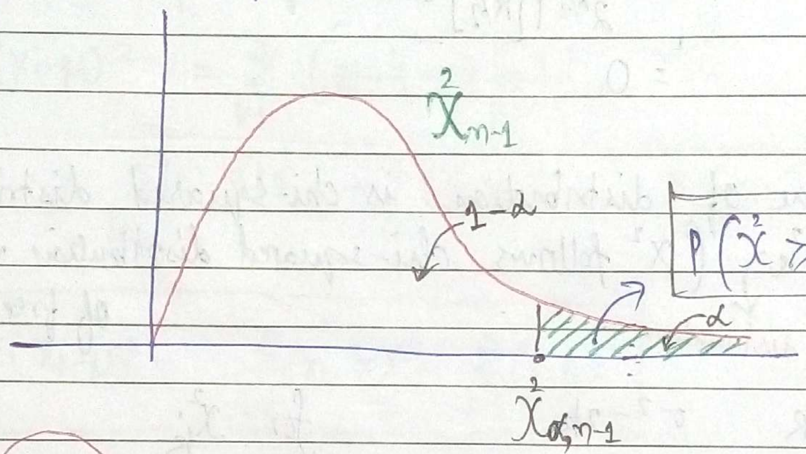
Estimation Problem:
Estimate σ^2 for a population $N(\mu, \sigma^2)$.
($X_1, X_2 \ldots X_n$ is a random sample from $N(\mu, \sigma^2)$).
Estimate σ^2.
$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \leftarrow \text{point estimate of } \sigma^2 \text{ (unbiased)}$$

Now to derive confidence interval for σ^2

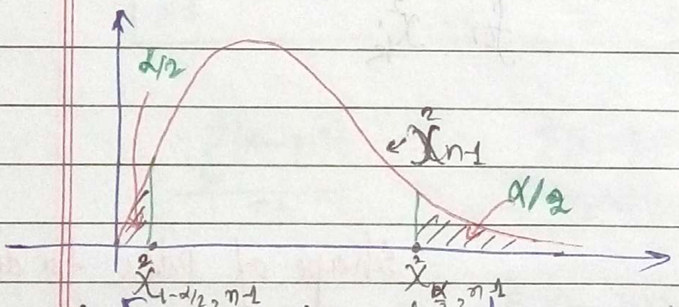
ANUBHAV JAIN NOTES

Scanned by CamScanner

1) Identify the statistic $\frac{(n-1)s^2}{\sigma^2}$
2) Distribution χ^2_{n-1}
3) Percent points
4) Probability statement
5) Algebraic manipulation to obtain L & U.



$$\boxed{P\left(\chi^2 \geq \chi^2_{\alpha, n-1}\right) = \alpha}$$



* For a given α, derive $100(1-\alpha)$ % confidence interval.
For (90% CI), $\alpha = 0.1$.

$$P\left(\chi^2_{1-\frac{\alpha}{2}, n-1} \leq \frac{(n-1)s^2}{\sigma^2} \leq \chi^2_{\frac{\alpha}{2}, n-1}\right) = 1-\alpha$$

$$\Rightarrow \boxed{P\left(\underbrace{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}, n-1}}}_{L} \leq \sigma^2 \leq \underbrace{\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}}_{U}\right) = 1-\alpha}$$

$\chi^2_{1-\frac{\alpha}{2}, n-1}$ $\qquad\qquad$ $\chi^2_{0.95, 15} = 7.26$ (dof)

$\chi^2_{\frac{\alpha}{2}, n-1}$ $\qquad\qquad$ $\chi^2_{0.05, 15} = 25$

where $\alpha = 0.1$, $n = 16$.

Consider the following random sample from $N(\mu, \sigma^2)$.

2.3, −1.2, 6.1, 0.02,
−0.6, −0.03, 0.32, 0.10
−1.2, −0.4, −0.16, 1.8
1.9, 2.2, −2.7, −1.5.

Compute a point estimate of σ^2 and 95% confidence interval for σ^2.

$$\bar{x} = \frac{6.95}{16} = 0.434375.$$

For 95% CI, $\alpha = 0.05$.

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = \frac{\overset{4.283 \times 15}{\cancel{2.0039}}}{15} = \boxed{\overset{4.283}{\cancel{0.1336}} = \text{Point estimate of } \sigma^2.}$$

$\Rightarrow S = \underset{2.0696}{\cancel{0.3655}}$ = Point estimate of σ.

$$\chi^2_{1-\frac{\alpha}{2}, n-1} = \chi^2_{0.975, 15} = 6.26$$

$$\chi^2_{\frac{\alpha}{2}, n-1} = \chi^2_{0.025, 15} = 27.5.$$

$$L = \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}} = \frac{\overset{4.283 \times 15}{\cancel{2.0039}}}{27.5} = \cancel{0.07287}.\ 2.336$$

$$U = \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}} = \frac{\overset{4.283 \times 15}{\cancel{2.0039}}}{6.26} = \cancel{0.32011}.\ 10.2636.$$

$$\hookrightarrow \boxed{P(L \leq \sigma^2 \leq U) = 0.95.}$$

Estimate Problem 3:

Given a random sample $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$ estimate μ when σ^2 is unknown.

Estimation of μ, σ^2 known	Estimation of μ, σ^2 unknown
$\bar{X}$: point estimator.	$\bar{X}$: point estimator
CI of μ statistic $\rightarrow$ $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	CI of μ $\dfrac{\bar{X}-\mu}{\cancel{\sigma}/\sqrt{n}} \sim t_{n-1}$ we want to learn

The t-distribution

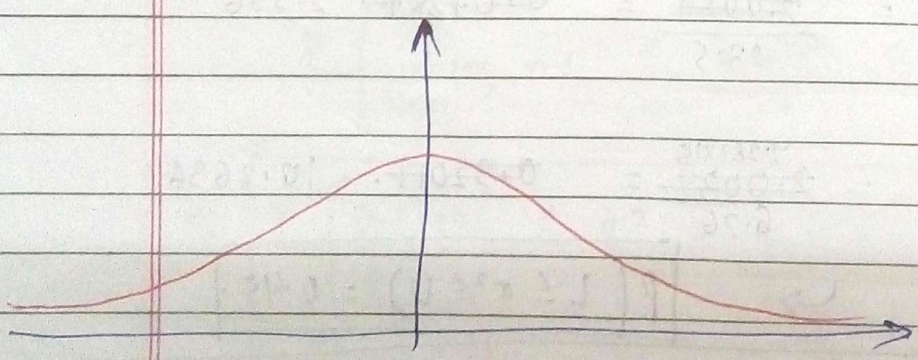
Let $Z \sim N(0,1)$ and $V \sim \chi^2_k$ and Z and V are independent.
Then the random variable

$$T = \frac{Z}{\sqrt{V/k}}$$

is said to follow t-distribution with k degrees of freedom with the following pdf.

$$f(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k}\,\Gamma(k/2)} \frac{1}{\left[\frac{t^2}{k}+1\right]^{\frac{k+1}{2}}} \qquad -\infty < t < \infty$$

k: dof of t



$$T = \frac{Z}{\sqrt{\frac{Z_1^2+\ldots+Z_k^2}{k}}} \sim t_k$$

$Z_1, Z_2 \ldots Z_k$ are independent

$k \to \infty$, $t_\infty \approx N(0,1)$

Given a random sample $X_1, X_2, \ldots, X_n$ from $N(\mu, \sigma^2)$ where both μ & σ^2 are unknown.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{and} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

Prove: $\dfrac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1}$

Proof: $\rightarrow \dfrac{(\bar{x}-\mu)/\sigma}{s/\sqrt{n}\cdot 1/\sigma} = \dfrac{(\bar{x}-\mu)/(\sigma/\sqrt{n})}{s/\sigma}$

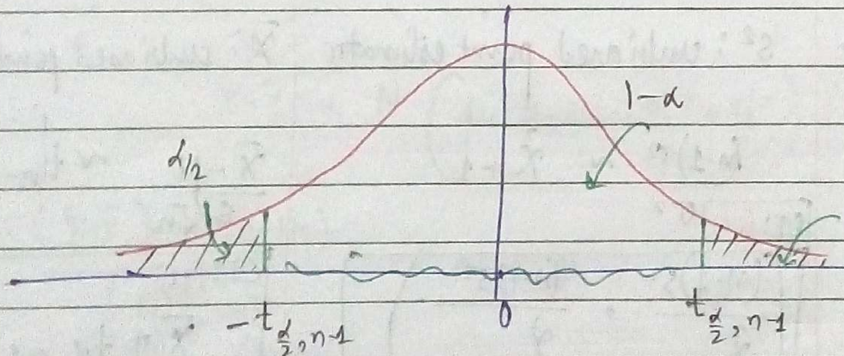
* $\dfrac{(\bar{x}-\mu)}{\sigma/\sqrt{n}} \sim N(0,1)$

* $\dfrac{s}{\sigma}$: $\dfrac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

$$\frac{s^2}{\sigma^2} \sim \frac{\chi^2_{n-1}}{n-1}$$

$$\frac{(\bar{x}-\mu)/(\sigma/\sqrt{n})}{s/\sigma} = \frac{N(0,1)}{\sqrt{\chi^2_{n-1}/n-1}} \sim t_{n-1}$$

Indep. [$\bar{x}$ and s^2 are indep.] ← difficult H.W.



For $100(1-\alpha)\%$ CI,

$$P\left(-t_{\frac{\alpha}{2},n-1} \le \frac{\bar{x}-\mu}{s/\sqrt{n}} \le t_{\alpha/2,n-1}\right) = 1-\alpha$$

$$\Rightarrow \boxed{P\left(\underbrace{\bar{x}-t_{\frac{\alpha}{2},n-1}\cdot\frac{s}{\sqrt{n}}}_{L} \le \mu \le \underbrace{\bar{x}+t_{\frac{\alpha}{2},n-1}\cdot\frac{s}{\sqrt{n}}}_{U}\right)=1-\alpha}$$

$\bar{x} = 0.434375$ $\qquad \mu = ?$ $\quad \sigma = ?$

$s^2 = 4.283$

$-t_{0.025,15} = -2.13$ $\qquad t_{0.025,15} = 2.13$

$L = \bar{x} - 2.13 \cdot \frac{\sqrt{4.283}}{\sqrt{16}} = -0.6677$

$U = \bar{x} + 2.13 \sqrt{\frac{4.283}{16}} = 1.53645$

$\Rightarrow \boxed{P(L \le \mu \le U) = 95\%}$

One population case!

Consider a population from $N(\mu, \sigma^2)$

Let $x_1, x_2, \ldots, x_n$ be a random sample from population.

Estimate Problem 1 Estimate μ (when known σ^2)	Estimate Problem 2 Estimate σ^2	Estimate Problem 3 Estimate μ (unknown σ^2)
$\bar{x}$: unbiased point estimator	s^2 : unbiased point estimator	$\bar{x}$: unbiased point estimator
$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$	$\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1}$
$\boxed{\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} : CI(\mu)}$	CI: (σ^2) $\boxed{\left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right)}$	CI: (μ) $\boxed{\bar{x} \pm t_{\frac{\alpha}{2},n-1} \frac{s}{\sqrt{n}}}$

Two population case:

Let $x_1, x_2, \ldots, x_n$ be a random sample from $N(\mu_1, \sigma^2)$.

and let $y_1, y_2, \ldots, y_n$ be a random sample from $N(\mu_2, \sigma^2)$

Assumption: σ^2 is known.

Estimation Problem 4: Estimate $\mu_1 - \mu_2$.

$\rightarrow P\left((\bar{x}-\bar{y}) - z_{\alpha/2}\sqrt{\frac{2\alpha}{n}} \le \mu_1 - \mu_2 \le (\bar{x}-\bar{y}) - z_{\alpha/2}\sqrt{\frac{2\alpha}{n}}\right) = 1-\alpha$

Orion PAGE: DATE: / /

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\left(\frac{\sigma_1}{\sqrt{n_1}}\right)^2+\left(\frac{\sigma_2}{\sqrt{n_2}}\right)^2}} \sim N(0,1)$$

$\bar{X} \sim N\left(\mu_1, \frac{\sigma^2}{n}\right)$

$\bar{Y} \sim N\left(\mu_2, \frac{\sigma^2}{n}\right)$

Point estimator : $\bar{X}-\bar{Y}$

let $X_1, X_2, \ldots, X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$
let $Y_1, Y_2, \ldots, Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$
Given: σ_1^2 and σ_2^2 are known.

These two population are indep.

Estimator problem: Estimate $\mu_1-\mu_2$.
Point Estimator: $\bar{X}-\bar{Y}$

For CI:

1) $\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$ $\quad \bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$

$$\bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2, \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$$

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

$$\boxed{P\left[\underbrace{(\bar{X}-\bar{Y})-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}_{L} \le \mu_1-\mu_2 \le \underbrace{(\bar{X}-\bar{Y})+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}_{U}\right] = 1-\alpha}$$

ANUBHAV JAIN NOTES

Ex. Let -2.1, 1.9, 1.6, 0.3, 0.4, -0.2, -1.2, -0.03, 0.7, -0.6 be a random sample from $N(\mu_1, 1)$

Further,

2.3, -1.2, 6.1, 0.02, -0.6, -0.03, 0.32, 0.1, -1.2, -0.4, -0.16, 1.8, 1.9, 2.2, -2.7, -1.5 be a random sample from $N(\mu_2, 2)$

Derive 90% CI for $\mu_1 - \mu_2$.

$n_1 = 10$ $\quad n_2 = 16$ $\quad$ 90% CI $\Rightarrow \alpha = 0.1$

$\sigma_1 = 1$ $\quad \sigma_2 = 2$

$\bar{x} = 0.077$ $\quad \bar{y} = 0.434375$

$$Z_{\alpha/2} = Z_{0.05} = \cancel{1.96}\ 0.5199$$

$$L = (\bar{x} - \bar{y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = -0.3573 - 0.3075 = -0.66495$$

$$U = (\bar{x} - \bar{y}) + Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = -0.3573 + 0.3076 = -0.0498$$

$$\Rightarrow P(L \le \mu_1 - \mu_2 \le U) = 0.9.$$

15/4/17

Estimation Problem:

Let $x_1, x_2, \ldots, x_{n_1}$ be a random sample from $N(\mu_1, \sigma^2)$

Let $y_1, y_2, \ldots, y_{n_2}$ be a random sample from $N(\mu_2, \sigma^2)$

(The two normal populations are independent).

* Construct the confidence interval on $\mu_1 - \mu_2$. (σ^2: unknown)

$$\bar{x} = \frac{\sum_{i=1}^{n_1} x_i}{n_1}, \quad \bar{y} = \frac{\sum_{i=1}^{n_2} y_i}{n_2}$$

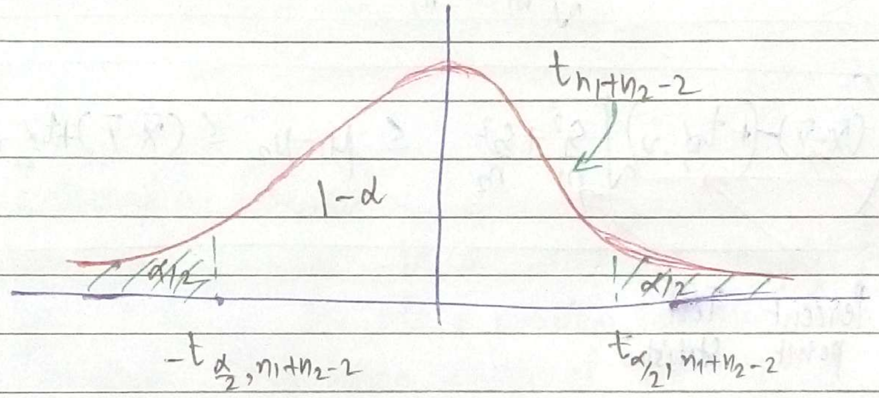
PAGE:
DATE: / /

Point estimator for $\mu_1 - \mu_2$ is $\bar{X} - \bar{Y}$

$$t = \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

where

$$S_p^2 = \frac{(n_1-1) s_1^2 + (n_2-1) s_2^2}{n_1+n_2-2} \rightarrow \text{pooled variance}$$



* Construct $100(1-\alpha)\%$ C.I.

Probability statement

$$\boxed{P\left(-t_{\alpha/2, n_1+n_2-2} \le \frac{(\bar{X}-\bar{Y}) - (\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \le t_{\frac{\alpha}{2}, n_1+n_2-2}\right) = 1-\alpha}$$

Estimation Problem:

$X \rightarrow N(\mu_1, \sigma_1^2)$

$Y \rightarrow N(\mu_2, \sigma_2^2)$

(σ_1 and σ_2 are unknown and unequal)

Define:-

$$t^* = \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_\nu$$

where

$$\nu = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{(s_1^2/n_1)^2}{n_1+1} + \dfrac{(s_2^2/n_2)^2}{n_2+1}} - 2$$

Probability Statement:-

$$\boxed{P\left(-t_{\frac{\alpha}{2},\nu} \le \frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}} \le t_{\frac{\alpha}{2},\nu}\right) = 1-\alpha}$$

$$\Rightarrow \boxed{P\left((\bar{x}-\bar{y}) - \left(+t_{\frac{\alpha}{2},\nu}\right)\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \le \mu_1-\mu_2 \le (\bar{x}-\bar{y}) + t_{\frac{\alpha}{2},\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right) = 1-\alpha}$$

Point estimator   Percent point   Test Statistic

Ex: Let $-2.1, 1.9, 1.6, 0.3, 0.4, -0.2, -1.2, -0.03, 0.7, -0.6$ be a random sample from $N(\mu_1, \sigma_1^2)$

Further, $2.3, -1.2, 6.1, 0.02, -0.6, -0.03, +0.32, 0.1, -1.2, -0.4, -0.16, 1.8, 1.9, 2.2, -2.7, -1.5$ be a random sample from $N(\mu_2, \sigma_2^2)$

Given that the normal populations are independent, construct 95% CI

$X: \quad n_1 = 10 \qquad \bar{X} = 0.077$

$Y: \quad n_2 = 16 \qquad \bar{Y} = 0.434375$

$100(1-\alpha) = 95 \Rightarrow \alpha = 0.05.$

$$S_1^2 = \frac{\sum_{i=1}^{n_1}(x_i-\bar{x})^2}{n_1-1} = 4.283.$$

$$S_2^2 = \frac{\sum_{i=1}^{n_2}(y_i-\bar{y})^2}{n_2-1} = \cancel{1.5456} \quad 1.2065$$

PAGE:
DATE: / /

$v = 25.8076 \approx 25.$

$t_{\alpha/2, 25} = t_{0.025, 25} = 2.060.$

$$L = (0.077 - 0.434375) - 2.060 \times \sqrt{\frac{1.9065}{10} + \frac{4.283}{16}} = \cancel{-0.14732}$$

$$= -2.689$$

$$U = \qquad\qquad = 3.54$$

$$\Rightarrow \boxed{P(-2.689 \le \mu_1 - \mu_2 \le 3.54) = 0.95}$$

Estimation Problem:
Let $X_1, X_2, \ldots X_{n_1}$ be a random sample from $N(\mu_1, \sigma_1^2)$
Let $Y_1, Y_2, \ldots Y_{n_2}$ be a random sample from $N(\mu_2, \sigma_2^2)$
Compute confidence interval for $\frac{\sigma_2^2}{\sigma_1^2}$

Point Estimator: $\frac{S_2^2}{S_1^2}$

Statistic: $\frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}$

Sampling distribution: $F_{m,n}$

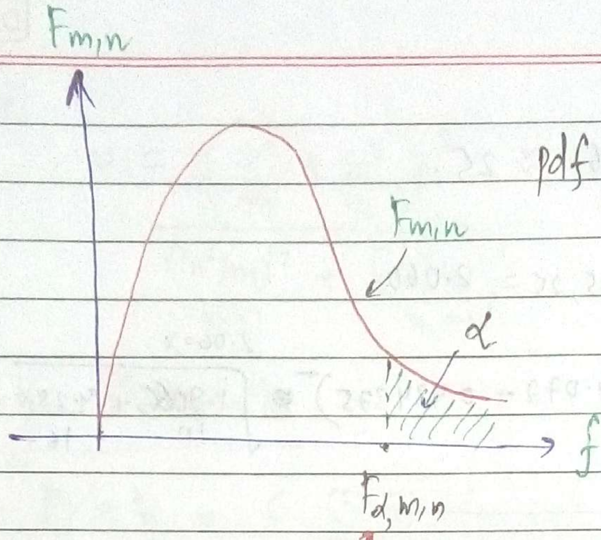
Let W and V be two independent χ^2_m & χ^2_n respectively. The the ratio

$$F = \frac{W/m}{V/n} \sim F_{m,n}$$

is said to follow an F distribution with m, n degrees of freedom. The pdf of $F_{m,n}$ is

$$h(f) = \frac{\Gamma\left(\frac{m+n}{2}\right)\left(\frac{m}{n}\right)^{m/2} f^{m/2 - 1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)\left(\frac{m}{n}f + 1\right)^{\frac{m+n}{2}}}, \quad f > 0$$

ANUBHAV JAIN NOTES



$$\boxed{P\left(F \geq F_{\alpha,m,n}\right) = \alpha}$$

$$\boxed{F_{1-\alpha,m,n} = \frac{1}{F_{\alpha,n,m}}}$$

Ex. $\boxed{\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}}$

$X_1, X_2, \ldots, X_{n_1} \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$

$Y_1, Y_2, \ldots, Y_{n_2} \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$

- S_1^2 is an unbiased estimator of σ_1^2
- S_2^2 is an unbiased estimator of σ_2^2

$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}$$

$$\frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$$

Estimation Problem:

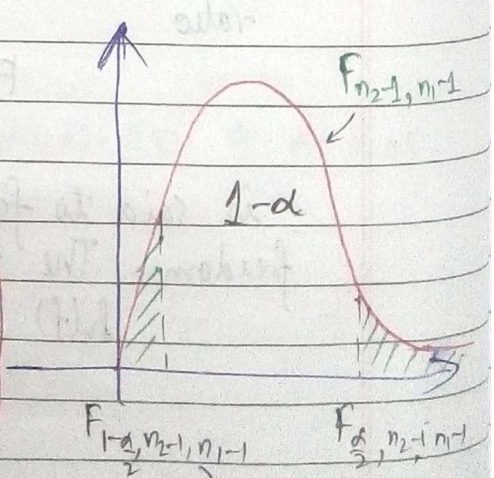
$X_1, X_2, \ldots, X_{n_1} \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$

$Y_1, Y_2, \ldots, Y_{n_2} \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$

Compute $100(1-\alpha)\%$ Confidence interval for $\dfrac{\sigma_2^2}{\sigma_1^2}$

Point estimator: $\dfrac{S_2^2}{S_1^2}$

Statistic: $\dfrac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} \sim F_{n_2-1,n_1-1}$



$$\boxed{P\left(F_{1-\frac{\alpha}{2},n_2-1,n_1-1} \leq \frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} \leq F_{\frac{\alpha}{2},n_2-1,n_1-1}\right) = 1-\alpha}$$

$$\Rightarrow P\left(F_{1-\frac{\alpha}{2},n_2-1,n_1-1} \cdot \frac{S_1^2}{S_2^2} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq F_{\frac{\alpha}{2},n_2-1,n_1-1} \cdot \frac{S_1^2}{S_2^2}\right) = 1-\alpha$$

$$\Rightarrow \boxed{P\left(\underbrace{F_{1-\frac{\alpha}{2}, n_1-1, n_2-1} \cdot \frac{s_2^2}{s_1^2}}_{L} \le \frac{\sigma_2^2}{\sigma_1^2} \le \underbrace{F_{\frac{\alpha}{2}, n_1-1, n_2-1} \cdot \frac{s_2^2}{s_1^2}}_{U}\right) = 1-\alpha}$$

Ex: Compute 90% Confidence Interval $\frac{\sigma_2^2}{\sigma_1^2}$.

$$F_{0.95, 9, 15} = \frac{1}{F_{0.05, 15, 9}} = 0.332.$$

$$F_{0.05, 9, 15} = 2.59$$

$$\Rightarrow L = 0.332 \times \frac{s_2^2}{s_1^2}$$

$$U = 2.59 \times \frac{s_2^2}{s_1^2}$$

Confidence interval on $\mu_1 - \mu_2$ for paired observations

Paired Observations

$(X_1, Y_1), (X_2, Y_2), \ldots (X_n, Y_n)$

(X, Y) such that X & Y are "independent" normally distributed.

$X \sim N$

$Y \sim N$

$D = Y - X$ is a point estimator of $\mu_2 - \mu_1$.

$E(D) = E(Y-X) = \mu_2 - \mu_1$

$Var(D) \rightarrow$ unknown.

D is normal random variable with mean $\mu_1 - \mu_2$ and variance σ^2 (σ^2 is unknown)

Sample: $D_1, D_2, \ldots, D_n$

$$\frac{\bar{D} - (\mu_2 - \mu_1)}{S_D / \sqrt{n}} \sim t_{n-1}$$

where $\bar{D} = \dfrac{D_1 + \dots + D_n}{n}$

$S_D = \sqrt{\dfrac{1}{n-1} \sum_{i=1}^{n} (D_i - \bar{D})^2}$

$$\boxed{P\left(\bar{D} - t_{\frac{\alpha}{2}, n-1} \frac{S_D}{\sqrt{n}} \le \mu_2 - \mu_1 \le \bar{D} + t_{\frac{\alpha}{2}, n-1} \frac{S_D}{\sqrt{n}}\right) = 1-\alpha}$$

Confidence interval on proportions:

$X \sim \text{Binomial}(n, p)$

$\hat{p} = \dfrac{X}{n} \rightarrow$ proportion

$E(\hat{p}) = p$

$\text{Var}(\hat{p}) = \dfrac{pq}{n}$

Using CLT,

$\hat{p} \sim N\left(p, \dfrac{pq}{n}\right)$ approximately

$\hat{p}$ is a point estimator of p.

$100(1-\alpha)\%$ CI on p:

$$\frac{\hat{p} - p}{\sqrt{pq/n}} \sim N(0,1)$$

$$P\left(-Z_{\alpha/2} \le \frac{\hat{p} - p}{\sqrt{pq/n}} \le Z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow \boxed{P\left(\underbrace{\hat{p} - Z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}}_{L} \le p \le \underbrace{\hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}}_{U}\right) = 1-\alpha}$$

Q Out of 75 students present today, 50 students used DCT+ before coming to class. Find 90% CI.

$n = 75$. $100(1-\alpha) = 90 \Rightarrow \alpha = 0.1$

$\hat{p} = \frac{50}{75} = \frac{2}{3}$. $\Rightarrow \hat{p} = \frac{2}{3}$ $\hat{q} = \frac{1}{3} \Rightarrow \frac{\hat{p}\hat{q}}{n} = \frac{\frac{2}{9}}{75} = \frac{2}{675}$

$Z_{\alpha/2} = Z_{0.05} = 1.65$

$L = \frac{2}{3} - 1.65\sqrt{\frac{2}{675}} = 0.57685.$

$U = \frac{2}{3} + 1.65\sqrt{\frac{2}{675}} = 0.75648.$

Sampling distributions

* $N(0,1)$
* χ^2_n: Sum of squares of n independent $N(0,1)$
* $t_n = \frac{N(0,1)}{\sqrt{\chi^2_n/n}}$
* $F_{m,n} = \frac{\chi^2_m/m}{\chi^2_n/n}$

Estimation Problems				$100(1-\alpha)\%$ CI.
1 population $N(\mu, \sigma^2)$ $X_1, X_2, \ldots, X_n$ random sample from $N(\mu, \sigma^2)$	Estimate μ (σ^2 known)	$\bar{x}$: unbiased pt estimator of μ.	$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\bar{x} \pm \left(Z_{\frac{\alpha}{2}}\right)\left(\frac{\sigma}{\sqrt{n}}\right)$
	Estimate μ (σ^2: unknown)	$\bar{x}$ unbiased estimator of μ	$\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1}$	$\bar{x} \pm t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$
	Estimate σ^2 (μ: unknown)	s^2 unbiased pt. estimator of σ^2	$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$	$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}, n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}$
1 population proportion $\hat{p} = X/n$ $X \sim Bin(n,p)$	Estimate p (proportion of desired event)	$\hat{p} = \frac{X}{n}$ is an unbiased estimator of p	$\frac{\hat{p}-p}{\sqrt{pq/n}} \sim N(0,1)$	$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$