सङ्ख्यानम्

STATISTICS

STATISTICS

30-10-2024

$\mathcal{S}$ - population (all data points); theoretically $\exists$ a r.v. X corresponding to which at most all possible outcomes are in $\mathcal{S}$.

The distribution of $\mathcal{S}$ is the distribution of X.

$(x_{11}, x_{12}, \dots, x_{1n})$ is a sub-collection from $\mathcal{S}$.

$$\text{avg.} = \frac{\sum_{i=1}^{n} x_{1i}}{n}$$

$$= \hat{\mu}_1$$

$(x_{21}, x_{22}, \dots, x_{2n})$

$$\hat{\mu}_2 = \frac{\sum_{i=1}^{n} x_{2i}}{n}$$

$(X_1, X_2, \dots, \underbrace{X_n)}_{\text{iid}}$ such that X_i are copies of X, is called a random sample and is also called sample.

The observation of the given sample will be denoted by $(x_1, x_2, \dots, x_n)$.

<u>Aim:</u> To study $(X_1, \dots, X_n)$ to get information about the distribution of $\mathcal{S}$.

To be specific to know the population mean μ and variance σ^2.

<u>Method:</u>

① Form functions of $(X_1, \dots, X_n)$ such that studying these functions, we get to know about the parameters of the population like μ, σ^2;

i.e., in general, any parameter θ of $\mathcal{S}$. ~~the~~

We shall construct a function Θ of $(X_1, \dots, X_n)$ such that Θ is such a good f^n that an observation of $\hat{\Theta}$, say $\hat{\theta}$, will be an approx to the original parameter θ.

(We need to have corr. theory);

e.g., $\bar{x}$ is an observation of $\bar{X}$ will be an approx. to μ. (why?)

such $\hat{\Theta}$ is called an estimate, or a point estimator of the parameter θ and an observation $\hat{\theta}$ of $\hat{\Theta}$ will be called an estimation of θ.

② On getting such approximation, we shall find the error and approximation for which we need to know the distribution of $\hat{\Theta}$.

③ Define what should be called "good".

— Good properties of estimators.

$\bar{X}$ has properties:

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

and $\operatorname{var}(\bar{X}) = \frac{\sigma^2}{n} \to 0$

as $n \to \infty$.

Eg. For $\epsilon > 0$, what can we say about $P(|(\bar{X} - \mu)| \geq \epsilon)$ as $n \to \infty$.

↳ will be called convergence in probability

↳ weak convergence

• $P(\lim_{n \to 0} \bar{X} = \mu)$

↳ Almost sure convergence

↳ strong convergence

Markov's Inequality

X is non-negative.

Then $P(X \geq a) \leq \frac{E(X)}{a}$ $(a > 0)$

Chebyshev's Inequality

Suppose X be a r.v., with mean $\mu = E(X)$ and variance $\sigma^2 = var(X)$. ($\epsilon > 0$ is given)

Then,

① $P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \Rightarrow \boxed{P(|X - E(X)| \geq \epsilon) \leq \frac{Var(X)}{\epsilon^2}}$

$|X-\mu| \geq \epsilon$

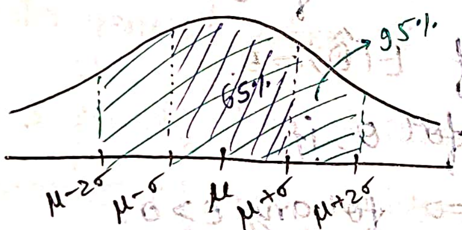
$\equiv (X-\mu)^2 \geq \epsilon^2$

So, $P(|X-\mu| \geq \epsilon)$

$= P((X-\mu)^2 \geq \epsilon^2)$

Apply Markov's, $\leq \frac{E(X-\mu)^2}{\epsilon^2} = \frac{var(X)}{\epsilon^2}$.

#



$[K=2]$

② For any $k > 0$. [Not useful for $K < 1$]

$P(|X-\mu| \geq K\sigma) \leq \frac{\sigma^2}{K^2\sigma^2} = \frac{1}{K^2} \Rightarrow \boxed{P(|X-E(X)| \geq K \cdot sd(X)) \leq \frac{1}{K^2}}$

Eg Replace x by $\bar{X}$ and see what happens when $n \to \infty$.

(01-11-2024)

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ called sample mean.

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

$var(\bar{X}) = \frac{1}{n}\sigma^2$

$\left[\begin{array}{c} E(X_i) = \mu \\ Var(X_i) = \sigma^2 \end{array}\right]$

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

Replace x by $\bar{x}$.

$E(\bar{x}) = \mu$, $Var(\bar{x}) = \frac{1}{n}\sigma^2$

Plug in Chebyshev's,

$$P(|\bar{x}-\mu| \geq \epsilon) \leq \frac{1}{n}\sigma^2/\epsilon^2$$

$\longrightarrow 0$ as $n \to \infty$

<u>Weak Law of Large Numbers:</u>

$$\lim_{n\to\infty} P(|\bar{x}-\mu| \geq \epsilon) = 0$$

(From Chebyshev) [Convergence in Probability]

<u>Strong Law of Large Numbers:</u>

$$P\left(\lim_{n\to\infty} \bar{x} = \mu\right) = 1$$

$\hookrightarrow$ Almost sure convergence

Defn: Suppose $\hat{\Theta}$ be an estimator of Θ.

① $\hat{\Theta}$ is called unbiased if $E(\hat{\Theta}) = \theta$.

② $\hat{\Theta}$ is called consistent for θ, if

$$\lim_{n\to\infty} P(|\hat{\Theta}-\theta| \geq \epsilon) = 0 \text{ for any } \epsilon > 0.$$

$\rightarrow$ ① and ② will be two of the "good" properties of an estimator.

Eg. $\bar{x}_n = \dfrac{\sum_{i=1}^{n} x}{n-1}$ $(n \geq 2)$

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

So, $\bar{x} = \dfrac{n-1}{n}\bar{x}_n$

$$E(\bar{x}_n) = E\left(\frac{n}{n-1}\bar{x}\right)$$

$$= \frac{n}{n-1} E(\bar{x})$$

$$= \frac{n}{n-1}\cdot\mu \quad \to \mu \text{ as } n\to\infty$$
$$\left[\because \frac{n}{n-1} \to 1 \text{ as } n\to\infty\right]$$

So, $E(\overline{X}_n) \neq \mu$, but
$E(\overline{X}_n) \to \mu$ as $n\to\infty$,

i.e., $\overline{X}_n$ is not unbiased estimator of μ, however, since $\lim_{n\to\infty} E(\overline{X}_n) = \mu$, we can say that $\overline{X}_n$ is asymptotically unbiased.

Is $\overline{X}_n$ consistent for μ? ✓

$\hookrightarrow \lim_{n\to\infty} P(|\overline{X}_n - \mu| \geq \epsilon) = 0$ [Exercise]

$var(\overline{X}_n) = var\left(\frac{n}{n-1}\overline{X}\right)$

$$= \left(\frac{n}{n-1}\right)^2 var(\overline{X})$$
$$= \frac{n^2}{(n-1)^2}\cdot\frac{1}{n}\cdot\sigma^2$$
$$= \frac{n}{(n-1)^2}\sigma^2 \to 0 \text{ as } n\to\infty.$$

→ In general, when var → 0 as $n\to\infty$, it will be consistent.

Eg. In a pond, it is desired to measure the average amount of magnesium/cc. A sample of size 10 is taken and the values are found to be

3.5, 4.7, 2.8, 4, 3.9, 4.5, 2.5, 4.5, 3.8, 2.1.

Find an estimator of the avg. amount of mg/cc in the pond. (The measurements are in mg/cc).

Soln: Let μ be the avg. Mg content.
Since, we know that $\overline{X}$ is a good estimator for μ, $\overline{x}$, an obs. of $\overline{X}$, will be an estimate of μ.

So, $\overline{x} = \frac{(\,)+(\,)+\dots+(\,)}{10}$ is an estimate of μ.

Central Limit Theorem (CLT)

X_i be an iid.

$E(X_i) = \mu$, $var(X_i) = \sigma^2$

then, $\boxed{Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \to N(0,1) \text{ as } n \to \infty}$

06-11-2024

$E(\bar{X}) = \mu$ and $var(\bar{X}) = \frac{\sigma^2}{n}$ $\left[\bar{X} = \sum_{i=1}^{n} X_i/n\right]$

$\Rightarrow s.d.(\bar{X}) = \frac{\sigma}{\sqrt{n}}$

i.e., if ϕ_n is the c.d.f. of $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$, then

$$\lim_{n \to \infty} \phi_n(x) = \Phi(x)$$

where Φ is the c.d.f. of $N(0,1)$.

i.e., if $Y_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$,

then $P(Y_n \le x) = \phi_n(x)$,

and if $Z \sim N(0,1)$

such that $P(Z \le x) = \Phi(x)$

then $\lim_{n \to \infty} P(Y_n \le x) = P(Z \le x)$

i.e., $\lim_{n \to \infty} \phi(x) = \Phi(x)$.

① Convergence in probability
(weak law of ~~convergence~~ large no.s)

② Almost sure convergence
(Strong law of large no.s)

③ Convergence in distribution.

Practical Use:

CLT ⇒ If n is large enough then $\frac{X-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ approx,

i.e., if n is large no., then $\bar{X} \sim N(\mu, \sigma^2/n)$ approx.

Thumb rule:

If $\boxed{n \geq 30}$, we can use this approximation to get good result.

Binomial probability table:

n	r	p: 0.1	0.2	0.3
1	0			
	1			
2	0			
	1			
	2			
⋮				

$\sum_{x=0}^{r} b(x; n, p)$

For $n=10$, $r=5$, $p=0.25$, binomial probability sum $=0.9803$

Approximations in Probability

Poisson Approximation to Binomial:

(Theorem): Let $X_n \sim b(n,p)$

Such that $np=\lambda$ is constant, and $n \to \infty$ and $p \to 0$.

Then, $\boxed{X_n \sim P(\lambda)}$ as $n\to\infty$.

Proof: $f_{X_n}(x) = {}^nC_x p^x q^{n-x}$

$= \frac{n(n-1)\ldots}{x(x-1)\ldots 1} p^x (1-p)^{n-x}$

Replace $p=\lambda/n$ and calculate

$\lim_{n\to\infty} f_{X_n}(x) = \ldots$

$= e^{-\lambda}\frac{\lambda^x}{x!}$

Alt: $M_{X_n}(t) = (pe^t + q)^n$

$= (pe^t + (1-p))^n$

Replace $p = \frac{\lambda}{n}$

$M_{X_n}(t) = \left(\frac{\lambda}{n} e^t + 1 - \frac{\lambda}{n}\right)^n$

$\lim_{n\to\infty} M_{X_n}(t) = \left(\frac{\lambda}{n} e^t + 1 - \frac{\lambda}{n}\right)^n = \lim_{n\to\infty}\left[1 + \frac{\lambda}{n}(e^t - 1)\right]^n$

$= e^{\lambda(e^t-1)}$

$\left[\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a\right]$

Practical use of Poisson approx. to Binomial:

When n is large enough and p is small enough such that np is moderate $(0 < np < 7)$

then $X_n \sim P(\lambda)$ approx, $\lambda = np$.

Thumb Rule:

① $n \geq 20$ and $p = 0.05$
$(\Rightarrow np \leq 1)$

② $n \geq 50$ & $np \leq 5$

③ $n \geq 100$ & $np \leq 10$.

Normal Approximation to Binary:

Let $X_n \sim b(n,p)$, p is fixed.

Then $\frac{X_n - np}{\sqrt{npq}} = Y_n \sim N(0,1)$ as $n \to \infty$.

(D'Moivre-Laplace limit theorem)

Follows from CLT:

Proof: $\lim_{n\to\infty} M_{X_n}(t) = (pe^t + q)^n$

$\lim_{n\to\infty} M_{Y_n}(t) =$

$M_{X+a}(t) = e^a M_X(t)$.

$M_{aX}(t) = M_X(at)$.

07-11-2024

$X_n \sim b(n, p)$, p is fixed.

$$\frac{X_n - np}{\sqrt{npq}} \sim N(0,1) \text{ as } n \to \infty$$

Thumbs rule:

1. $p \approx 0.5$ and $n \geq 10$
2. $np \geq 5$ & $npq \geq 5$
3. $np \geq 10$ & $npq \geq 10$.

If $np \geq 5$ & $npq \geq 5$.

Then, $\dfrac{X_n - np}{\sqrt{npq}} \sim N(0,1)$ approx.

i.e., $X_n \sim N(np, npq)$

CLT: X_is are iid, mean μ and s.d. σ.

Then $\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$ as $n \to \infty$.

CLT $\Rightarrow$ D'Moivre-Laplace

Take $X_i \sim b(1,p)$, $i = 1, \dots, n$, iid

$$X_n = \sum_{i=1}^{n} X_i \sim b(n,p)$$

$$\overline{X} = \sum_{i=1}^{n} X_i/n = X_n/n.$$

Plug in CLT & get D'Moivre result.

Eg. $X \sim b(n_1, p) \Rightarrow M_X(t) = (pe^t + q)^{n_1}$,

$Y \sim b(n_2, p) \Rightarrow M_Y(t) = (pe^t + q)^{n_2}$

$M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$ if X_1 & X_2 are independant.

$= (pe^t + q)^{n_1+n_2}$

$\Rightarrow X_1 + X_2 \sim b(n_1 + n_2, p)$

Eg. 5% of christmas trees light bulbs manufactured by a company are defective. The company's quality control manager is quiet concerned, therefore samples 100 bulbs from the factory. Let X denotes the no. of bulbs in the sample that are defective. What is the probability that samples contains atmost 3 defective bulbs.

Soln: $(X_1, \dots, X_{100})$

$X_i \sim b(1, 0.05)$

Let X denote the no. of faulty bulbs in the sample.

$X = \sum_{i=1}^{100} X_i \sim b(100, p)$.

$P(X \le 3)$

$= P(X=0 \text{ or } X=1 \text{ or } X=2 \text{ or } X=3)$

$= P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$= \sum_{i=1}^{3} P(X=i)$

${}^{100}C_i \cdot p^i \cdot q^{100-i}$

$n = 100, \ p = 0.05$

$np = 5$

$npq = 4.75$

Thumb Rule of Poisson Approximation:

1. $n \geq 20$ & $p \leq 0.5$
2. $n \geq 50$ & $np \leq 5$
3. $n \geq 100$ & $np \leq 10$
4. $n \geq 20$ & $np \leq 10$

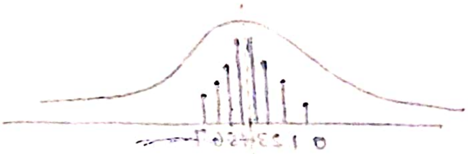
Since $n = 100 \geq 50$ & $np \leq 5$, we apply Poisson approximation to Binomial to calculate the probability.

So, $P(X \leq 3) \approx P(Y \leq 3)$ when $Y \sim p(np)$

$= 0.2650$ (from table) i.e., $Y \sim p(5)$.

Poisson Probability Sums $\sum_{x=0}^{r} p(x, \mu)$

r \ μ	0.1	0.2	...	5.0
0				
1				
2				
3				0.2650
4				
⋮				



Eg. About 10% of the population is left-handed.

(i) In a class of 150 students, what would be the probability that atleast 25 students are left handed.

(ii) In a class of 150 students, what would that no. of no. of left-handed students is in b/w 15 and 20.

Sol^n: Given that X, the count of left-handed students in a class of 150, follows $b(150, p)$, $p = 0.1$.

To find $P(X \geq 25)$

$= 1 - P(X \leq 24)$, where $X \sim b(150, 0.1)$

$n = 150$, $p = 0.1$, $np = 15$, $npq = 13.5$

Since $np \geq 10$ & $npq \geq 10$,

we say that $X \sim N(np, npq)$ approx, by D'Moivre result.

$P(X \le 24)$

$= P\left(\frac{X-np}{\sqrt{npq}} \le \frac{24-np}{\sqrt{npq}}\right)$

$\approx P\left(Z \le \frac{24-np}{\sqrt{npq}}\right)$ where $Z \sim N(0,1)$

$\vdots$ $\times$

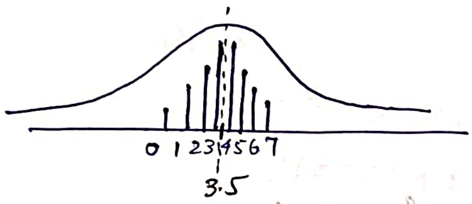
08-11-2024

Continuity Correction

$X \sim b(41, 0.5)$

$n = 41,\ p = 1/2$

$x = 0, 1, \ldots, 41$ (even)



$X \sim b(7, 0.5)$

$E(X) = np = 7 \times 0.5 = 3.5$

D'Moivre:

$X_n \sim b(n, p)$

$E(X_n) = np$, s.d. $(X_n) = \sqrt{npq}$

$\frac{X_n - np}{\sqrt{npq}} \sim N(0,1)$ approx. when n is large.

$\Rightarrow X_n \sim N(np, (\sqrt{npq})^2)$

Originally:

$P(X = i) = P(X = 7-i)\ \forall\ i = 0, \ldots, 7$

$\Rightarrow P(X \le 3) = P(X \ge 4)$

$= 1/2$

& $P(X \le 3) + P(X \ge 4) = 1$

$P(X=3) \ne 0 \ne P(X=4)$

After approximation:

$X \sim N(np, npq)$

$P(X = i) = 0$

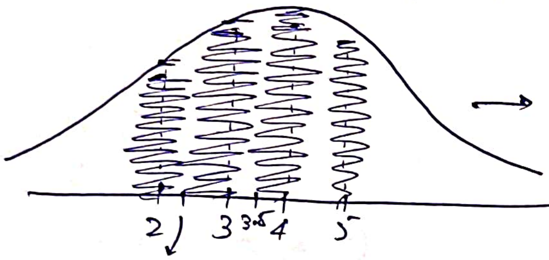
$P(X \leq 3.5) = P(X \geq 3.5)$

$= 1/2$

& $P(X \leq 3) \neq P(X \leq 3.5) = \frac{1}{2}$

$P(X = 3) \neq 0$ (before approx.)

≈ 0 (after approx.)



→ spreading [on approximation]

2.5 ⇝ junction b/w 2 & 3

When the approx. is good

then $\underset{\text{original}}{P(X=3)} \approx \underset{\text{(after approx.)}}{P(2.5 \leq X \leq 3.5)}$

and $[2.5, 3.5]$ is called the class of 3.

$\underset{\text{originally}}{P(X \leq 3)} \approx \underset{\text{after approx.}}{P(X \leq 3.5)}$

$\underset{\text{original}}{P(X < 2)} \approx \underset{\text{after approx.}}{P(X < 1.5)}$

↳ 2 is not included

$P(X=2) \leq P(1.5 \leq X \leq 2)$

$X \sim b(150, 0.1)$

$np = 15, \quad \sqrt{npq} = \sqrt{13.5} \doteq 3.674$

$P(X \geq 25) = 1 - P(X \leq 24)$

To find : $P(X \leq 24)$

$\approx P(X \leq 24.5)$ where

$X \sim N(np, (\sqrt{npq})^2)$

$X \sim N(15, (3.674)^2)$ approx.

$$= P\left(\frac{X-15}{3.67} \le \frac{24.5-15}{3.674}\right)$$

$$= P(Z \le 2.585) \quad \text{where } Z \sim N(0,1).$$

$$\approx P(Z \le 2.59) = 0.9952 \quad [\text{Approximation to } Z \le 2.59]$$

OR

[2 : 8

2.58 — 2.582 — 2.59]

1 : 1

[2.58 — 2.585 — 2.59]

$$= \frac{P(Z \le 2.58) + P(Z \le 2.59)}{2}$$

$$= \frac{0.9951 + 0.9952}{2} = 0.99515$$

Eg. Civil engineers believe that W, the amount of weight measured in 1000 pounds, that a certain span of bridge can withstand without structural damage, is normal distribution with mean 400 and S.d. 40. Suppose that weight of a car is a R.V. with mean 3 and s.d. 0.3.

Approximately how many cars would have to be on the bridge span for the probability of structural damage to exceed 0.1.

Solⁿ. Given that $W \sim N(400, (40)^2)$

Let n be the no. of cars on the bridge span and W_i be the weight of i^{th} car.

Given that W_i has mean 3 & sd. 0.3

Can assume that W_i are iid.

To discuss, ~~that wil connected~~ damage in the bridge, we must have on an average more than 100 cars on the bridge.

$n \times 3 = 400$
$\Rightarrow n \geq 100$

and hence can apply CLT to approximate the dist. of sum of

$$\sum_{i=1}^{n} w_i/n = \tilde{W} \sim N(0,1) \text{ approx.}$$

by CLT as $n \geq 30$.

13-11-2024

W_i, $i = 1, 2, \ldots, n$
(Weight of cars)

$E(W_i) = 3$

$s.d.(W_i) = 0.3$

$W \sim N(400, (40)^2)$
(withstanding wt. on the bridge span)

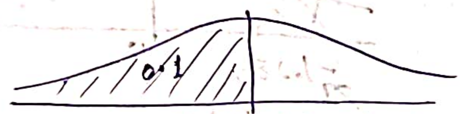
Total weight of cars,

$$X = \sum_{i=1}^{n} W_i$$

To find n s.t.

$P(X \geq W) \geq 0.1$

$\Rightarrow P(W - X \leq 0) \geq 0.1 \quad$ —①



$P(X - W \geq 0) = 0.1$

Since, to have possible damage in the bridge, should have 100 cars.

$3 \cdot n = 400$
$n \geq 100$

So, we can apply CLT to have approx. dist. of X.

$$\text{CLT} \Rightarrow \frac{\overline{W} - E(\overline{W})}{s.d.(\overline{W})} \sim N(0,1), \text{ approx. } X.$$

i.e, $\overline{w} \sim N\left(E(\overline{w}), (sd(\overline{w}))^2\right)$

$$X = \sum_{i=1}^{n} w_i$$

$$\frac{X}{n} = \overline{w} \Rightarrow X = n\overline{w}$$

$$n\overline{w} \sim N\left(nE(\overline{w}),\ n^2(s.d.(\overline{w}))^2\right)$$

$$\Rightarrow X \sim N\left(3n,\ n^2\left(\frac{0.3}{n}\right)^2\right)$$

So,

$$-X \sim N(-3n, n(0.3)^2)$$

$$\Rightarrow W + (-X) \sim N(400 - 3n,\ (40)^2 + n(0.3)^2)$$

So, we shall solve the eqn ① for n,

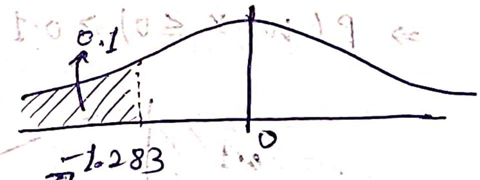
i.e., $P(W - X \leq 0) \approx 0.1$

write the eqn in $N(0,1)$ variate

$$P\left(\frac{(W-X) - (400-3n)}{\sqrt{40^2 + n(0.3)^2}} \leq \frac{3n - 400}{\sqrt{40^2 + n(0.3)^2}}\right) \approx 0.1$$

$$P\left(Z \leq \frac{3n - 400}{\sqrt{40^2 + n(0.3)^2}}\right) \approx 0.1$$

Find n, from the table.



from $N \sim (0,1)$ table

$$P(Z \leq -1.283) \approx 0.1$$

So,

$$\frac{3n - 400}{\sqrt{40^2 - n(0.3)^2}} = -1.283$$

$$\Rightarrow \frac{(3n - 400)^2}{40^2 - n(0.3)} = (1.283)^2$$

$$\Rightarrow n = 116, 150$$

For damage,

$$\boxed{\begin{aligned} &X \geq W \\ \Rightarrow\ &E(X) \geq E(W) \end{aligned}}$$

$$\Rightarrow n(3) \geq 400$$

$$\Rightarrow n \geq 133$$

$$\therefore n = 150.$$

<u>Discussion:</u>

(i) $X \geqslant 0$

$E(X) \geqslant 0$

When we have two R.Vs

(ii) $X \geqslant Y$

$$\Rightarrow X - Y \geqslant 0$$

$$\Rightarrow E(X - Y) \geqslant 0$$

$$\Rightarrow \iint (x-y) f_{XY}(x,y)\, dx\, dy \geqslant 0$$

$$\Rightarrow \iint_{\mathbb{R}^2} x f_{XY}(x,y)\, dx\, dy - \iint_{\mathbb{R}^2} y f_{XY}(x,y)\, dx\, dy \geqslant 0$$

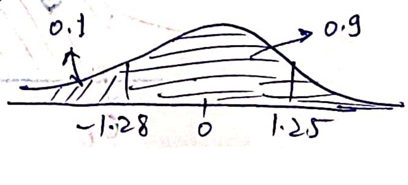
<u>Recall:</u> $\underbrace{f_X(x)}_{\text{marginal}} = \int_{\mathbb{R}} f_{XY}(x,y)\, dy$

$$= \int_{\mathbb{R}} x \left(\int_{\mathbb{R}} f_{XY}(x,y)\, dy \right) dx - \int_{\mathbb{R}} y \left(\int f_{XY}(x,y)\, dx \right) dy$$

$$= \int_{\mathbb{R}} x f_X(x)\, dx - \int_{\mathbb{R}} y f_Y(y)\, dy$$

$$= E(X) - E(Y)$$

For $n = 150$, $\dfrac{50}{\sqrt{(40)^2 - (0.09 \times 150)}} = 1.25 \longrightarrow P = 0.9$



14-11-2024

Let S be a population with parameter θ & $\hat{\theta}$ is an estimator of θ (point estimator).

Defn: ① Standard error in estimating θ by $\hat{\theta}$ is given by $Var(\hat{\Theta})$.

② Bias of $\hat{\theta}$ in estimating θ is denoted by $Bias(\hat{\Theta})$ and is given by

$$Bias(\hat{\Theta}) = E(\hat{\Theta}-\theta)$$

Ideally we want: (i) $E(\hat{\Theta}) = 0$

(ii) $Var(\hat{\Theta}-\theta) = 0$ → Not possible.

At least $var(\hat{\Theta}-\theta) \to 0$ as $n \to \infty$.

[$Var(\hat{\Theta}-\theta) = Var(\hat{\Theta}) \to 0$ as $n \to \infty$]

[$E(\hat{\Theta}) = \theta$ and $Var(\hat{\Theta}) \to 0$ as $n \to \infty \Rightarrow$ consistency!]

Point estimator is not good!

Need better estimators which shall cover a region for estimation (interval).

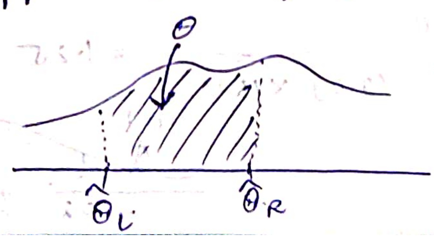
We need: $\hat{\Theta}_L$ and $\hat{\Theta}_R$

such that "$[\hat{\Theta}_L, \hat{\Theta}_R]$"

∴ random interval's observation shall give possible values of θ with definite probability,

[$Bias(\hat{\Theta}) = 0 \Leftrightarrow \hat{\Theta}$ is unbiased]

so that any observation of the random interval is an approximation to θ.



We want: $\theta \in [\hat{\Theta}_L, \hat{\Theta}_R] = 1-\alpha$

shall say $[\hat{\Theta}_L, \hat{\Theta}_R]$ is an $(1-\alpha)\times 100\%$ random confidence interval for θ.

An observation of $[\hat{\Theta}_L, \hat{\Theta}_R]$, say $[\hat{\theta}_L, \hat{\theta}_R]$, is called $(1-\alpha)100\%$ confidence interval for θ.

How to achieve:

Get point estimator $\bar{X}$ of μ (good one) such that when $X_i \sim N(\mu, \sigma^2)$

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

n is large such that we can apply CLT and get

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

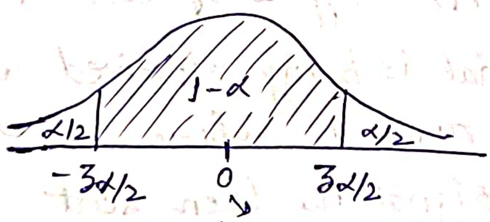
In those two cases we know that

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

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

$$\Rightarrow \bar{X} \sim N(\mu, (\sigma/\sqrt{n})^2)$$

$Z \sim N(0,1)$



translate to μ

Defn of z_β:

$Z \sim N(0,1)$, then

$P(Z \geq z_\beta) = \beta$

$\Rightarrow z_\beta$ is upper $100\beta\%$ point.

Translating the diagram,

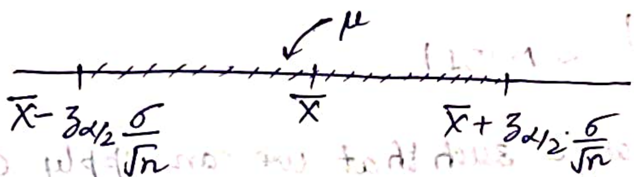
$$P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) = 1-\alpha$$

Here, $Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

So, $P\left(-z_{\alpha/2} \le \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right) = 1-\alpha$

$\Rightarrow P\left(\bar{X} - z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}}\right) = 1-\alpha$

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



So, an $(1-\alpha)100\%$ random C.I. for μ is

$$\left[\bar{X} - z_{\alpha/2}\cdot\sigma/\sqrt{n},\ \bar{X} + z_{\alpha/2}\cdot\sigma/\sqrt{n}\right].$$

When approximating μ by $\bar{X}$, the max. error shall happen when μ lies at the end points.

So, $e_{max} = \frac{1}{2}\cdot$ length of the interval

$= z_{\alpha/2}\cdot\sigma/\sqrt{n}$ with ~~points~~ prob. $(1-\alpha)$.

Eg. Suppose from a place A, a signal of constant value μ is being sent and the signal is being received at place B with added random noise, N, where $N \sim N(0, 3^2)$. Suppose from A, the signal has been sent 9 times and at B, the recorded values are -

5, 8.5, 12, 15, 7, 9, 7.5, 6.5, 10.5

① Give an estimate of μ.

② Find ~~the~~ a 95% confidence interval for μ.

③ Find the no. of times the signal has to be recorded at B s.t. the error does not exceed 0.2 with 90% confidence.

Soln:

$X_i \sim N(0+\mu, 3^2)$

$(X_1, \ldots, X_9)$

Observations: $x_1, \ldots, x_9$

$\bar{X}$ is an estimator of μ.

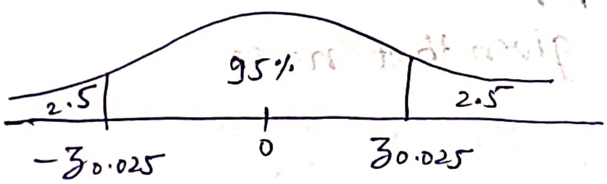
$[\because E(\bar{X}) = \mu]$

Count: $\bar{x} = \sum x_i/9 = 9$

Distribution of $\bar{X}$:

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

$\alpha = 0.025 \times 2$



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

$\Rightarrow (1-\alpha)100\%$ random CI in $\left[\bar{X} - z_{\alpha/2}\cdot\sigma/\sqrt{n}, \bar{X} + z_{\alpha/2}\cdot\sigma/\sqrt{n}\right]$.

So, a $100(1-\alpha)\%$ C.I. for μ is

$$\left[\bar{x} - z_{\alpha/2}\cdot\sigma/\sqrt{n},\ \bar{x} + z_{\alpha/2}\cdot\sigma/\sqrt{n}\right]$$

where $\alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

$\sigma = 3,\ n = 9$

$z_{\alpha/2} = z_{0.025} = -1.96$

③ $e_{max} = 0.2$

$= \frac{\sigma}{\sqrt{n}}\cdot z_{\alpha/2} = \frac{3}{\sqrt{n}}(-1.64)$

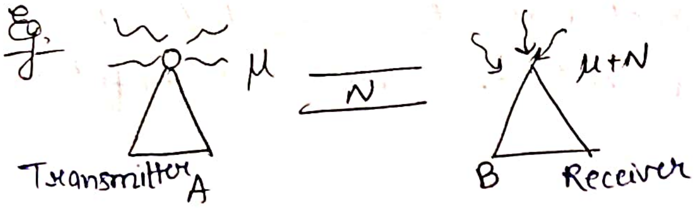
$\alpha = 0.1$

$\alpha/2 = 0.05$

$\Rightarrow n = \left[\frac{3(-1.64)}{0.2}\right]^2 \approx 605$

20-11-2024

Eg.



$E(N)=0,\quad Var(N)=3^2$

Distribution of N is unknown.

A sample of size 49 is taken at B and found that sample mean is 8.5.

Find a 95% CI for μ.

Soln: $E(N)=0$

$\Rightarrow E(\mu+N)=\mu$

$\Rightarrow$ μ is the mean of population.

For sample $(x_1,\dots,x_n)$, we know that a good point estimator of μ (pop. mean) is

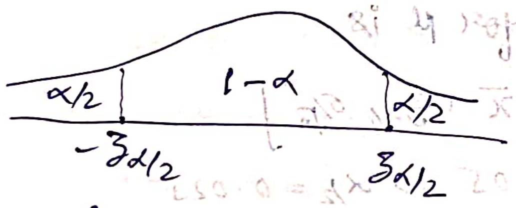
$$\overline{x}=\sum_{i=1}^{n} x_i \quad \text{given that } n=49.$$

Dist. of $\overline{x}$:

$$\frac{\overline{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ approx.}$$

By CLT as $n=49 \geq 30$

$1-\alpha = 9.5$



$$P\left(-z_{\alpha/2} \leq \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\overline{x}-z_{\alpha/2}\cdot\sigma/\sqrt{n} \leq \mu \leq \overline{x}+z_{\alpha/2}\cdot\sigma/\sqrt{n}\right) = 1-\alpha$$

$$\left[\begin{array}{l|l} \overline{x}=8.5 & \alpha/2 = 0.025 \\ \sigma=3 & \Rightarrow z_{\alpha/2}=1.96 \\ n=49 & \end{array}\right]$$

Case-3: σ is unknown & n is small.

Case-4: σ is unknown & n is large.

Aim: To find C.I. for μ (pop. mean) for case-3 & case-4.

Aim: Finding dist. of $\bar{X}$ or possibly the dist. of

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}.$$

Q) can we replace σ by something else s.t. we know the distribution.

Estimation of σ^2:

$$\sigma^2 = var(X)$$
$$= E((X - E(X))^2)$$
$$\sum_{i=1}^{n} x_i = \bar{x} \approx E(x)$$
$$\approx E((x-\bar{x})^2)$$
$$\approx \sum_{i=1}^{n} (x_i-\bar{x})^2/_{n-1}$$
$$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$$

is called sample var and is an estimation of σ^2.

Earlier it was believed that $\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim N(0,1)$.

Later it was found that

$$\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(n)$$ [Student's t-distribution with n d.f.]

which tends to $N(0,1)$ as $n \to \infty$.

χ^2 - distribution:

Let $Z \sim N(0,1)$,

then the distribution of Z^2 is called χ^2-distribution with 1 degree of freedom.

[Chi-squared]

Show that [Exercise]

$$M_{Z^2}(t) = (1-2t)^{-1/2}$$

$$\parallel$$

$$E(e^{tZ^2}) = \int_{-\infty}^{\infty} e^{tz^2} f(z).dz$$

So, X is following a χ^2-distribution with 1 d.f.

$[X \sim \chi^2(1)]$ iff $M_X(t) = (1-2t)^{-1/2}$

General defn:

Let X be a R.V. s.t. $M_X(t) = (1-2t)^{-n/2}$, then X is said to be a χ^2 r.v. with n d.f.

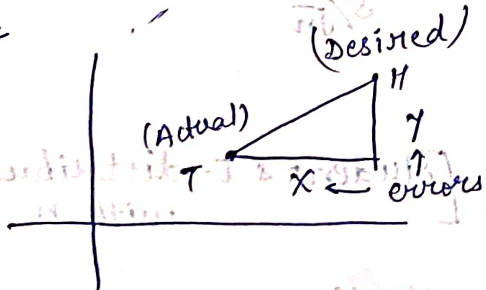
Exercise: Let $X_i \sim N(0,1)$, iid. (independent) $\forall\ i = 1, \dots, n$.

Let $V = X_1^2 + X_2^2 + \dots + X_n^2$,

then show that $V \sim \chi^2_V(n)$.

[Hint: $M_{X_1^2+X_2^2+\dots+X_n^2}(t) = \prod_{i=1}^{n} M_{X_i^2}(t)$]

#



$X \sim N(0, 2^2)$

$Y \sim N(0, 2^2)$

$P(D^2 \le 8^2)$

$D^2 = X^2 + Y^2$

21-11-2024

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

Sample variance :

χ^2—distribution with n d.f.

$X \sim \chi^2(n)$

$\Leftrightarrow M_X(t) = (1-2t)^{-n/2}$

Q) Is S^2 a good point estimator of σ^2.

① Unbiasedness ($E(\hat{\theta}) = \theta$)

$E(S^2) = \sigma^2$ (Exercise)

② $Var(S) = \frac{2\sigma^4}{n-1}$ (Exercise)

$\rightarrow 0$ as $n \rightarrow \infty$.

③ Consistency: ($\forall \epsilon > 0$, $P(|\hat{\theta} - \theta| \geq \epsilon) \rightarrow 0$ as $n \rightarrow \infty$.

Show that (Exercise)

$\forall \epsilon > 0$

$\lim_{n\to\infty} P(|S^2 - \sigma^2| \geq \epsilon) = 0.$

Use ① Chebychev's inequality

② Unbiasedness

③ $Var(S^2) \rightarrow 0$ as $n \rightarrow \infty$.

So, S^2 is unbiased and consistent for σ^2.

t-distribution

Let $Z \sim N(0,1)$

$V \sim \chi^2(k)$

where Z and V are independent.

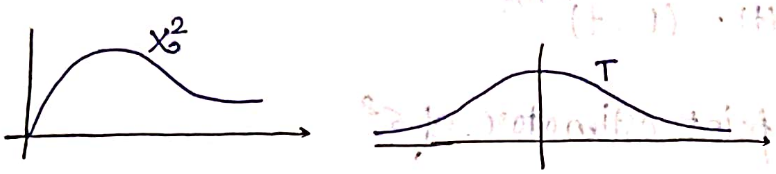
Then, $T = \frac{Z}{\sqrt{V/k}}$ has t-distribution with k d.f.

Properties of χ^2 and t-distribution

Let $V \sim \chi^2(K)$ and $T \sim t(K)$.
Then,

① $E(V) = K$, $E(T) = 0$

② $Var(V) = 2K$, $Var(T) = \frac{K}{K-2}$.

③ 

Show that :

① $$s^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right],$$ for calculation

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

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

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

Theory: Let X and Y are independent such that $X+Y \sim \chi^2(m)$ and $Y \sim \chi^2(n)$, $m > n$.
Then, $X \sim \chi^2(m-n)$.

Proof:

$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$ [$\because$ Independent]

$\Rightarrow (1-2t)^{-m/2} = M_X(t)(1-2t)^{-n/2}$

$\Rightarrow M_X(t) = (1-2t)^{-(m-n)/2}$

$\Rightarrow X \sim \chi^2(m-n)$

Eg. Let X and Y are independent such that

$$X+Y \sim N(\mu_1, \sigma_1^2)$$
$$Y \sim N(\mu_2, \sigma_2^2)$$

Find the distribution of X.

Soln: $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

$\Rightarrow e^{\left(\mu_1 t + \frac{\sigma_1^2 t^2}{2}\right)} = M_X(t) \cdot e^{\left(\mu_2 t + \frac{\sigma_2^2 t^2}{2}\right)}$

$\Rightarrow M_X(t) = \exp\left[(\mu_1 - \mu_2)t + \frac{(\sigma_1^2 - \sigma_2^2)t^2}{2}\right]$

$X \sim N(\mu_1 - \mu_2, \sigma_1^2 - \sigma_2^2)$

Theorem: Let S be a population such that $S \sim N(\mu, \sigma^2)$.

Let $(X_1, \ldots, X_n)$ be a (random) sample then $\bar{X}$ and $(\bar{X} - X_1, \ldots, \bar{X} - X_n)$ are independent.

Corr: Let $S \sim N(\mu, \sigma^2)$, then

S^2 and $\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2$ are independent.

Proof: since S^2 is a f^n of $(\bar{X} - X_1, \ldots, \bar{X} - X_n)$.

we have S^2 and $\bar{X}$ are independent.

$\Rightarrow S^2$ and $\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)$ are independent. [All f^ns of $\bar{X}$ are indep.]

$$X = \left(\frac{n-1}{\sigma^2}\right) S^2$$

$$X + Y = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma/\sqrt{n}}\right)^2$$

Corrolary: $S \sim N(\mu, \sigma^2)$

Then, $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

Since, $S \sim N(\mu, \sigma^2)$

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

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

$\Rightarrow \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(n)$ and $\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$.

Since the equality ② holds when s^2 and $\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$ are independent by cancellation property.

χ^2-distribution, we see that $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$.

<u>Theorem</u>: Let $S \sim N(\mu, \sigma^2)$, where σ^2 is possibly not known, then

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1).$$

<u>Proof</u>: Since $S \sim N(\mu, \sigma^2)$,

we have $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$

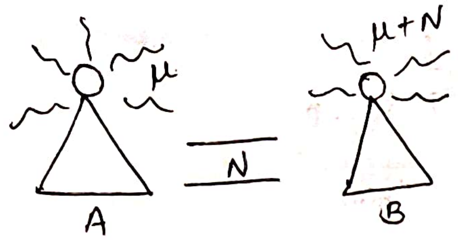
and $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

Now, since $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ and $\frac{(n-1)s^2}{\sigma^2}$ are independent,

we see that

$$\frac{\frac{\bar{X}-\mu}{\cancel{\sigma}/\sqrt{n}}}{\sqrt{\frac{\cancel{(n-1)}s^2}{\cancel{\sigma^2}}/\cancel{(n-1)}}} = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)$$

Eg.



It is known that N has mean 0. ($sd(N)$ is unknown)

At B, 10 times signal has been recorded:

5, 8.5, 12, 15, 7, 9, 7.5, 6.5, 10.5, 9

① Estimate μ.

② Find a 95% CI for μ.

③ Find the error in estimating μ (by $\overline{X}$) with 95% confidence.

Soln: ① $\overline{X}$ is an estimator for pop. mean, which is here μ.

$$E(\mu+N) = \mu + 0 = \mu.$$

So, $\overline{X}$ is an estimator of μ where $\overline{x} = \sum_{i=1}^{10} x_i/10 = 9$.

② To find CI for μ, we need to know the distribution of $\overline{X}$.

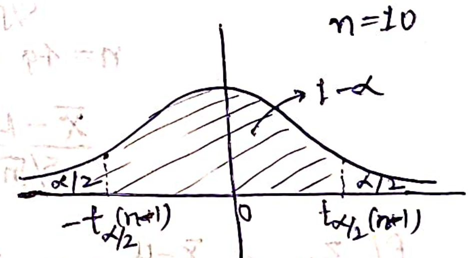
Distribution of $\overline{X}$:

$$T = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1), \text{ where } n=10.$$

$$1-\alpha = 0.95$$
$$\Rightarrow \alpha = 0.05$$
$$\Rightarrow \alpha/2 = 0.025$$

$$P\left(-t_{\alpha/2} \le \frac{\overline{X}-\mu}{S/\sqrt{n}} \le t_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\overline{X} - t_{\alpha/2}\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$



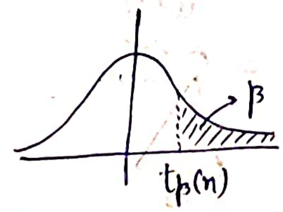
So, $(1-\alpha)100\%$ random CI for μ is

$$\left[\overline{X} - t_{\alpha/2}\frac{S}{\sqrt{n}}, \ \overline{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}\right]$$

and hence a CI is $\left[\overline{x} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \ \overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right]$.

$\overline{x} = 9$, $t_{\alpha/2} = t_{0.025}$

$P(T \geq t_\beta(n)) = \beta$



$\nu \longleftrightarrow n$

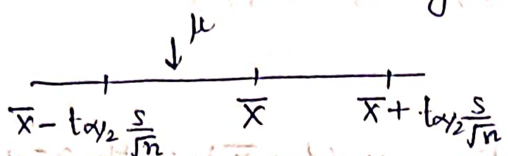
$\alpha/2 \longleftrightarrow \beta$

From the table, $\nu = 9, \alpha = 0.025$

$t_{0.025}(9) = 2.262$ → (from table)

$\bar{x} = 9, \quad s = 2.906$ (~~by calculation~~). → (By calculation) → $\left[s^2 = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}\right]$

③



$e_{max} = t_{\alpha/2}\dfrac{s}{\sqrt{n}}$ with 95% CI with $\alpha/2 = 0.025$

$= 2.262 \times \dfrac{2.906}{\sqrt{10}} = 2.078.$

Eg. The recorded no. of datapoints at B is 49, and found that 9 is the sample mean.

① Find a 99% CI for μ.

② Find the error with 95% CI.

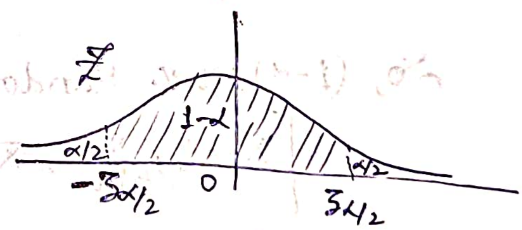
Aim: To find $(1-\alpha)100\%$ CI for μ, we need to know the distribution of $\bar{X}$.

Dist. of $\bar{X}$: $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

$n = 49 \Rightarrow$ By CLT

$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ approx.

$P\left(-z_{\alpha/2} \leq \dfrac{\bar{X}-\mu}{S/\sqrt{n}} \leq z_{\alpha/2}\right) = 1-\alpha.$



So, $(1-\alpha)100\%$ random CI for μ is

$\left[\bar{X} - z_{\alpha/2}\dfrac{S}{\sqrt{n}}, \ \bar{X} + z_{\alpha/2}\dfrac{S}{\sqrt{n}}\right],$

and hence a $(1-\alpha)100\%$ CI is

$\left[\bar{x} - z_{\alpha/2}\dfrac{s}{\sqrt{n}}, \ \bar{x} + z_{\alpha/2}\dfrac{s}{\sqrt{n}}\right]$

where $\bar{x} = 9$, $z_{\alpha/2}$ = from table

$s = 2.5$ (given)

$n = 49 \Rightarrow \sqrt{n} = 7, \quad \alpha = 0.005.$

Eg. Let $W, X, Y \sim N(\mu, \sigma^2)$ are independent. Show that

$$\frac{\sqrt{2}\,(W-\mu)}{\sqrt{X^2+Y^2+2\mu^2-2\mu(X+Y)}} \sim t(2).$$

Soln:

$$\frac{\sqrt{2}(W-\mu)/\sigma}{\sqrt{X^2+Y^2+2\mu^2-2\mu(X+Y)}\,/\sigma} \sim t(2).$$

$$\left(\frac{X-\mu}{\sigma}\right)^2 + \left(\frac{Y-\mu}{\sigma}\right)^2.$$

$$Z = \frac{W-\mu}{\sigma}, \quad X' = \frac{X-\mu}{\sigma}, \quad Y' = \frac{Y-\mu}{\sigma}$$

$$T = \frac{Z}{\sqrt{\frac{V}{K}}}$$

Let $V = X'^2 + Y'^2$

As $T \sim t(2)$.

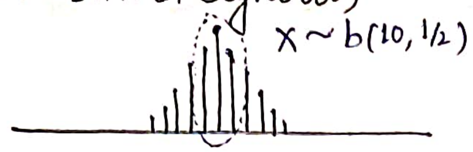
$$\Rightarrow \frac{\frac{W-\mu}{\sigma}}{\sqrt{\frac{\left(\frac{X-\mu}{\sigma}\right)^2+\left(\frac{Y-\mu}{\sigma}\right)^2}{2}}} \sim t(2)$$

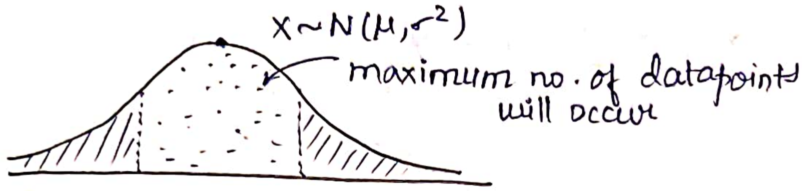
$$\Rightarrow \frac{\sqrt{2}\,(W-\mu)}{\sqrt{X^2+\mu^2-2\mu X+Y^2+\mu^2-2\mu Y}} \sim t(2)$$

$$\Rightarrow \frac{\sqrt{2}\,(W-\mu)}{\sqrt{X^2+Y^2+2\mu^2-2\mu(X+Y)}} \sim t(2).$$

$\bar{X}$ estimates population mean μ.
S^2 estimates population variance σ^2.

ML- Estimator
(Maximum likelyhood)





→ This maximum value captures some properties from n, p or μ, σ^2 to be at that place.

Process:

Let $S \sim f(\theta_1, \theta_2, \ldots, \theta_n)$, assume that $\theta_2, \ldots, \theta_n$ are ~~called~~ known and θ_1 is unknown.

$S \sim f(\theta_1)$.

To get as estimator of θ_1,

take a sample $(X_1, \ldots, X_n)$ from S

$\Rightarrow X_i \sim f(\theta_1)$,

and the joint distribution of $(X_1, X_2, \ldots, X_n)$ is given by

$L(\theta_1) = f_{X_1}^{x_1}(\theta_1) \cdot f_{X_2}^{x_2}(\theta_1) \cdots f_{X_n}^{x_n}(\theta_1)$ as

$X_1, \ldots, X_n$ are iid.

Assume that we have an observation $(\hat{x}_1, \ldots, \hat{x}_n)$ specific value, then

$L(\theta) = f_{X_1}(\hat{x}_1, \theta_1) \ldots f_{X_n}(\hat{x}_n, \theta_1)$ is called likelyhood function.

Our aim here is to maximize $L(\theta)$ given $(\hat{x}_1, \ldots, \hat{x}_n)$.

For $\hat{\theta} = \theta(\hat{x}_1, \ldots, \hat{x}_n)$,

$L(\theta)$ is maximum.

Then, $\hat{\theta}$ is an estimate of θ called ML estimate,

and $\hat{\theta} = \theta(X_1, \ldots, X_n)$ is called the ML estimator of θ.

Eg. Let $S \sim N(\mu, \sigma^2)$.

Aim: To find an ML-estimator of μ.

Let $(X_1, X_2, \ldots, X_n)$ be a random sample.

$\Rightarrow X_i \sim N(\mu, \sigma^2)$.

The likelyhood f^n for μ is

$$L(\mu) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

To maximize $L(\mu)$ w.r.t. μ,

we consider $\log_e L(\mu)$ and maximize

$$\ln(L(\mu)) = \sum_{i=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\frac{d}{d\mu}(\ln(L(\mu))) = 0 + \frac{1}{2\sigma^2} \sum_{i=1}^{n} 2(x_i - \mu).$$

To maximize, $\frac{d}{d\mu}(\ln(L(\mu))) = 0$.

$$\Rightarrow \sum_{i=1}^{n} (x_i - \mu) = 0 \Rightarrow \mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\downarrow$$

$$\left[\sum_{i=1}^{n} x_i - n\mu = 0\right]$$

So, $$\hat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n}$$

is an ML-estimator of μ.

So, in this case, $\hat{\mu} = \overline{X}$.

Eg. Do it for σ^2 (where μ is replaced by its known value, likely $\hat{\mu} = \bar{x}$) $\hat{\mu} = \overline{X}$.

Eg. Find the ML-estimator of θ where the $S \sim f(x,\theta)$,

$$f(x,\theta) = \theta e^{-\theta x}, \; x \geq 0.$$

If an observation gives values (2.5, 3.1, 1.5, 3.6, 4.1), find an estimate of θ.

Sol^n: Likelyhood f^n:

$$L(\theta) = \prod_{i=1}^{n} f_{x_i}(x_i, \theta), \text{ where } x_i \sim f(x_i, \theta).$$

and $x_1, \dots, x_n$ is a random sample.

$$= \prod_{i=1}^{n} \theta e^{-\theta x_i}$$

Consider $\ln(L(\theta)) = \sum_{i=1}^{n} \ln(\theta . e^{-\theta x_i})$

$$\Rightarrow \ln(L(\theta)) = \sum_{i=1}^{n} \ln(\theta) - \sum_{i=1}^{n} \theta x_i$$

To maximize $L(\theta)$, we maximize $\ln(L(\theta))$.

$$\frac{d}{d\theta}(\ln(L(\theta))) = 0 \Rightarrow \sum_{i=1}^{n} \frac{1}{\theta} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \frac{n}{\theta} = \sum_{i=1}^{n} x_i$$

$$\Rightarrow \theta = \frac{n}{\sum_{i=1}^{n} x_i}.$$

So, the ML estimator of θ is

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}$$

So, an ML estimator of θ is

$$\hat{\theta} = \frac{5}{14.8} = 0.337.$$

Show that $\boxed{\left.\frac{d^2}{d\theta^2}(\ln(L(\theta)))\right|_{\theta = \frac{n}{\sum_{i=1}^{n} x_i}} < 0}$. ⇝ for maxima.