Interval Estimation

· **Def**: An interval estimate for parameter θ is any pair of functions: $L(\vec{X}), U(\vec{X})$. $L(\vec{x}) \le U(\vec{x})\ \forall \vec{x} \in \Lambda$.

Random interval $[L(\vec{X}), U(\vec{X})]$ is interval estimator.

Remark: We sometimes consider $L(\vec{x}) = -\infty$ or $U(x) = +\infty$. or interval $[\ ,\ ), (\ ,\ )\ldots$

· **Def**: Coverage Prob: $P_\theta(\theta \in [L(\vec{X}), U(\vec{X})])$

Confidence coefficient of $[L(\vec{X}), U(\vec{X})]$ is $\inf_{\theta \in \Theta} P_\theta(\theta \in [L(\vec{X}), U(\vec{X})])$

Remark: i) The prob statement is for $\vec{X}$ r.v. not θ.

ii) Replace term interval estimator by confidence interval since it's always joint with confidence coefficience "$1-\alpha$"

(1) **Methods of finding Interval Estimators**:

① **Inverting a Test statistic**:

Thm: For each $\theta \in \Theta$, $A(\theta_0)$ is acceptance region of level α test of $H_0: \theta = \theta_0$. Def: $C(\vec{x}) = \{\theta_0 \mid \vec{x} \in A(\theta_0)\}$

Then the random set $C(\vec{X})$ is "$1-\alpha$" confidence set

Conversely, it's true!

Remark: Note that we carefully use "set" not "interval".

In most cases. one-side hypothesis give one-side intervals. two side hypothesis give two-side interval.

③ Pivotal Quantities:

· Def: r.v. $Q(\vec{X}, \theta)$ is a pivot if its dist indept of all parameters $\theta \in \Theta$.

i.e. $P_\theta(Q(\vec{X},\theta) \in A) = P(Q(\vec{X},\theta) \in A)$

Remark: i) For μ is unknown. then $Q(\vec{X}, \theta, \mu)$ isn't pivot for parameter θ.

e.g. $\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}$ $\times$ $\xrightarrow{\text{replace}}$ $\frac{\bar{X}-\mu}{\sqrt{S^2/n}}$

ii) By reverting: $\{\theta : Q(\vec{x}, \theta) \in A\}$. we obtain an estimator.

e.g.

Form	Type	Pivot
$f(x-\mu)$	Location	$\bar{X}-\mu$
$\frac{1}{\sigma} f(\frac{x}{\sigma})$	Scale	$\bar{X}/\sigma$
$\frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$	Location-Scale	$\frac{\bar{X}-\mu}{S}$ (For μ, σ may be unknown)

Procedure: $T \sim f(x|\theta)$. Obtain dist of $Q(\vec{X}, \theta)$:

let $x = Q(t, \theta)$. $\therefore f(x|\theta) = g(Q(t,\theta)) \left| \frac{\partial}{\partial t} Q(t,\theta) \right|$

$\Rightarrow$ Find a.b. st. $P_\theta(a \le Q(\vec{X},\theta) \le b) \ge 1-\alpha$.

Then $C(\vec{x}) = \{\theta_0 \mid a \le Q(\vec{x}, \theta_0) \le b\}$

③ Pivoting the CDF:

· For guaurantimg $C(\vec{x}) = \{\theta \mid a \le a(\vec{x},\theta) \le b\}$ is an interval. We need $a(\vec{X},\theta)$ is mono- of θ. $\forall \vec{x}$.

Then. note that for $T \sim F_T(t|\theta)$. (usually a S.S.)

$F_T(T|\theta) \sim$ Unsiform$(0,1)$.

By this pivot. with little assumption. We can guarantee $C_{F_T(T|\theta)}(\vec{x})$ is an interval.

Thm. (Conti. Case)

For $T \sim F_T(t|\theta)$. conti. cdf. Let $\alpha_1 + \alpha_2 = \alpha \in (0,1)$

suppose $\theta_L(t)$. $\theta_u(t)$ are defined as follow:

i) If $F_T(t|\theta)$ increase on θ for $\forall t$. Then

$F_T(t|\theta_L(t)) = \alpha_1$. $F_T(t|\theta_u(t)) = 1-\alpha_2$

ii) If $F_T(t|\theta)$ decrease on θ. for $\forall t$. Then

$F_T(t|\theta_L(t)) = 1-\alpha_2$. $F_T(t|\theta_u(t)) = \alpha_1$

Then $[\theta_L(T). \theta_u(T)]$ is "$1-\alpha$" confidence Interval. for θ.

Pf: Since T is conti. $\therefore F_T(T|\theta) \sim$ Unsiform$(0,1)$.

$\{t \mid \alpha_1 \le F_T(t|\theta) \le 1-\alpha_2\}$ is a "$1-\alpha$" acceptance Region.

Then fixed t. Convert to interval of $\theta_L(t)$. $\theta_u(t)$.

which're unique!

Remark: By using st. from equation in i). ii). Solve for $\theta_L(t)$, $\theta_u(t)$!

Thm. (Discrete Case)

$T \sim F_T(t|\theta) = P(T \le t|\theta)$. discrete. Let $\alpha_1 + \alpha_2 = \alpha \in (0,1)$

Suppose $\theta_L(t)$, $\theta_U(t)$ defined as follow:

i) $F_T(t|\theta)$ decrease on θ for each t.

$P(T \le t|\theta_U(t)) = \alpha_1$, $P(T \ge t|\theta_L(t)) = \alpha_2$

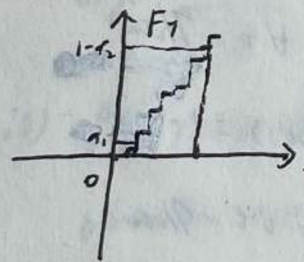
ii) Similar definition for increase case.

Then $[\theta_L(T), \theta_U(T)]$ is ASYPT-"$1-\alpha$" CI for θ.

Pf: i) Note that $[\theta_L(T), \theta_U(T)] =$

$\{\theta \mid \alpha_1 \le F_T(t|\theta(t)) \le 1-\alpha_2\}$

$\Rightarrow \{T \mid \alpha_1 \le F_T(T|\theta) \le 1-\alpha_2\} \triangleq A(T)$



$P_\theta(A(T)) \simeq 1-\alpha_2-\alpha_1 = 1-\alpha$

$$\begin{cases} P(F_T(T|\theta) \le 1-\alpha_2) \le 1-\alpha_2 \\ P(P_T(T \ge t|\theta) \ge \alpha_1) \ge \alpha_1 \end{cases}$$

Only when $P(P_T(T \ge t|\theta) \ge \alpha_1) \simeq \alpha_1$.

$A(T)$ is a "$1-\alpha$" confidence interval.

Remark: It's not easy to apply in discrete case!

④ Bayesian Intervals:

· Given $f(x|\theta)$. Suppose $\theta \sim \pi(\theta)$ we obtain: posterior dist of θ: $\pi(\theta|x)$. Then a "$1-\alpha$" credible set is $A(x)$. if $P(\theta \in A(x)|x) = \int_A \pi(\theta|x)d\theta = 1-\alpha$. (Differentiate "credible" and "confident".)

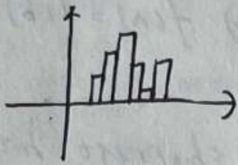
since the former is $p(\cdot|x)$, the latter is $p(\cdot)$).

Remark: Credible prob means coverage prob.

⑤ Bootstrap:

· We have n samples. n isn't large. then

Case one: Nonparametrized (unknow dist)

$\Rightarrow$ Use histogram:  (From $\{X_i\}_1^n$ samples)

We obtain an empirical cdf. Then resample a large amount of samples from this dist.

Case two: Parametric ($X \sim f(x|\theta)$)

since number of samples isn't large. the estimate $\hat{\theta}_0$ (From MLE) won't be accurate. Recognize $f(x|\hat{\theta}_0)$ as the "true" dist

Generate and resample $\{X_n^*\}$. $\Rightarrow$ estimate $\hat{\theta}^*$ from $\{X_n^*\}$.

$$\begin{array}{ccccl} \Delta^* = & \hat{\theta}^* & - & \hat{\theta}_0 & \Rightarrow \text{Interval } [\underline{\delta}_1^*, \bar{\delta}_2^*] \\ \downarrow & \downarrow & & \downarrow \text{ replace.} & \qquad\quad \wr \\ \Delta = & \hat{\theta} & - & \theta_0 \text{ (real param)} & \Rightarrow \text{Interval } [\underline{\delta}_1, \bar{\delta}_2]. \end{array}$$

Then the interval is: $P(\underline{\delta}_1^* \le \hat{\theta}^* - \hat{\theta}_0 \le \bar{\delta}_2^*) = 1-\alpha$

i.e. $[\underline{\delta}_1^* + \hat{\theta}_0, \bar{\delta}_2^* + \hat{\theta}_0]$. is "$1-\alpha$" CI.

(2) Method of evaluting

Interval estimators:

· Naturally. We want to obtain the set with small size but large coverage prob.

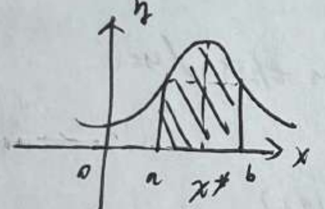
Firstly, we will restrict the confidence coefficient on "$1-\alpha$". Then find an interval with shortest length.

<u>Thm.</u> $f(x)$ is a unimodal pdf (i.e. $\exists x^* \in \mathbb{R}$, s.t. $f(x)\uparrow$ if $x \le x^*$, $f(x)\downarrow$ if $x \ge x^*$). $[a,b]$ satisfies:

i) $\int_a^b f(x) = 1-\alpha$ ii) $f(a) = f(b) > 0$. iii) $a \le x^* \le b$

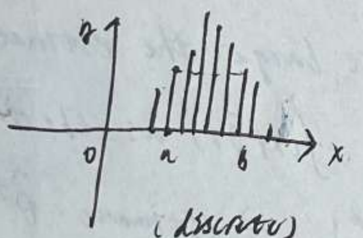
Then $[a,b]$ is the shortest interval of "$1-\alpha$" confi.

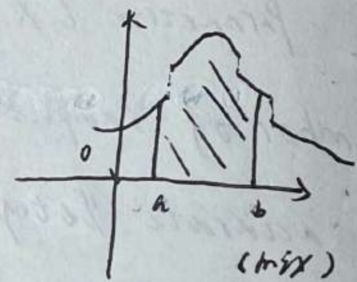
<u>Pf:</u>



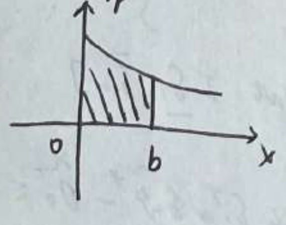
It's routine to check!

(By classified discussion)





<u>Remark:</u> Similarly, For $X \sim f(x|\theta)$, $\downarrow$, $x \ge 0$.



$[0,b]$ is the form of shortest interval with the same confident coefficien.

① <u>Test-Related Optimality:</u>

- From relation: Test of hypo $\longleftrightarrow$ Confidence set. So the optimality for them has some sense of corresponse.

Def: the prob. of false covering is

$P_\theta(\theta' \in C(X))$, $\theta' \neq \theta$, when $C(X) = [L(X), U(X)]$.

$P_\theta(\theta' \in C(X))$, $\theta' < \theta$, when $C(X) = [L(X), +\infty)$

$P_\theta(\theta' \in C(X))$, $\theta' > \theta$, when $C(X) = (-\infty, U(X)]$.

where θ is true parameter. θ' is what we want to cover.

$\Rightarrow$ A "$1-\alpha$" confidence set is called uniformly most accurate (UMA) if it minimized the prob of false coverage.

Remark: There's a relation between UMP test and UMA confidence set. Since the former often has the form of one-side interval, so may the latter is!

Thm. $X \sim f(\vec{x}|\theta)$, $\theta \in \mathbb{R}$. $\forall \theta_0$, $A^*(\theta_0)$ is UMP, level α AR of test $H_0: \theta = \theta_0$ v.s. $H_1: \theta > \theta_0$. Denote $C^*(x)$ is the inverting "$1-\alpha$" confidence set. Then:

$$P_\theta(\theta' \in C^*(X)) \le P_\theta(\theta' \in C(X)) \quad \forall\ \theta' < \theta$$

Pf: Set $H_0: \theta = \theta'$ v.s. $H_1: \theta > \theta'$

$$\therefore P_\theta(\theta' \in C^*(X)) = P_\theta(X \in A^*(\theta'))$$

$$\le P_\theta(X \in A(\theta')) = P_\theta(\theta' \in C(X))$$

For any confidence set $C(X)$.

Remark: i) If $C^*(X) = [L(X), +\infty)$. Then, it's UMA.

ii) Similar statement on $H_0: \theta = \theta_0$ v.s $H_1: \theta < \theta_0$.

Next. we deal with two-sided confidence set.

For simplification. we restrict on unbiased one:

· **Def**: A $1-\alpha$ confidence set $C(\vec{x})$ is unbiased if

$$P_\theta(\theta' \in C(\vec{x})) \le 1-\alpha. \ \forall \theta' \ne \theta.$$

Remark: Another perspective: $P_\theta(\theta' \in C(\vec{x})) \le P_\theta(\theta \in C(\vec{x}))$, the coverage prob $\ge$ the false coverage prob. $\forall \theta$.

For: $H_0: \theta=\theta'$. v.s. $H_1: \theta \ne \theta'$.

Then the power func. of $A(\theta')$ is unbiased!

Thm (Pratt): $X \sim f(x|\theta)$. $C(X)=[L(X). U(X)]$ is confidence interval for θ. If $L(x), U(x) \uparrow$ of x.

Then $\forall \theta^*$. $E_{\theta^*}(\text{Length}(C(X))) = \int_{\theta \ne \theta^*} P_{\theta^*}(\theta \in C(X))\, d\theta$.

Pf: $\int_{\theta \ne \theta^*} P_{\theta^*}(\theta \in C(X))\, d\theta = \int_{\theta \ne \theta^*} E_{\theta^*}(I(\theta \in C(X)))\, d\theta$.

$= E_{\theta^*}\left(\int_{\theta \ne \theta^*} I(\theta \in C(X))\, d\theta\right) = E_{\theta^*}(\text{Length}(C(X)))$

Remark: It claims the expected length is the integral of false coverage prob. So,

minimize the Length of CI $\iff$ minimize the prob of false cover.

But it doesn't work in one-sided case.

② Bayesian Optimality:

· The goal of obtaining the smallest confidence set with specific coverage prob. can also be attained by Bayesian Rule.

i.e. If we obtain $\pi(\theta|x)$ from $\pi(\theta)$. $f(x|\theta)$.

We want to find $C(\vec{x})$:

$$\begin{cases} \int_{C(x)} \pi(\theta|x) d\theta = 1-\alpha. \\ \text{Size}(C(x)) \le \text{Size}(C'(x)). \text{ for any other confidence} \end{cases}$$

set $C'(x)$. s.t. $\int_{C'(x)} \pi(\theta|x) d\theta \ge 1-\alpha$.

$\Rightarrow$ We can handle with the case:

when $\pi(\theta|x)$ is unimodle. then

$\{\theta \mid \pi(\theta|x) \ge k\}$ is the form. it's

called highest posterior density (HPD)

region.

③ Juola's Observation:

<u>Thm.</u> If pivot $Q(\vec{X},\theta) \sim f(t)$. $C = \{Q(\vec{X},\theta) \in A\}$ is confidence "$1-\alpha$" set. i.e. $P(Q \in A) = \int_A f dt = 1-\alpha$. If length$(C)$ has form $\int_A g(t)dt$. $\exists g$. Then the optimal solution s.t. $\min_A$ length(C) is $\{g \le \lambda f\}$. λ is for: $\int_{\{g \le \lambda f\}} f dt = 1-\alpha$.

<u>Pf:</u> For any other set A': $\int (I_A - I_{A'})(\lambda f - g) dt \ge 0$.

where $\int_{A'} f \ge 1-\alpha$ $\therefore$ $\int_A g(t) \ge \int_{A'} g(t) dt$.

<u>e.g.</u> $b-a = \int_a^b dx$.