Recap:

~~Intro~~

① UMPU test.

② UMPU test for a normal distribution boils down to the two sided normal test.

Likelihood ratio test

$X_1, \ldots, X_n$ have a joint likelihood $f_\theta(\underline{X})$

$H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_1$

the LRT test uses a test statistic

$$\lambda = \frac{\sup_{\theta \in \Theta_0} f_\theta(\underline{X})}{\sup_{\theta \in \Theta_0 \cup \Theta_1} f_\theta(\underline{X})}$$

if λ is small $\Rightarrow$ Θ_0 is not very likely

hence we reject when $\lambda < c$, c is chosen based on the level requirement.

Example: $X_1, \ldots, X_n$ iid location shifted exponential with location parameter θ.

$$f_\theta(x) = \begin{cases} e^{-(x-\theta)} & , x \geq \theta \\ 0 & \text{o.w.} \end{cases}$$

$$f_\theta(\underline{X}) = \begin{cases} \prod_{i=1}^{n} \exp\{-(x_i - \theta)\} & , \quad x_1, \ldots, x_n \geq \theta \\ 0 & \text{o.w.} \end{cases}$$

$$\Leftrightarrow f_\theta(\underline{X}) = \begin{cases} \exp\{-\sum_{i=1}^{n} x_i + n\theta\} & \text{if } x_{(1)} \geq \theta \\ 0 & \text{o.w.} \end{cases}$$

Consider testing the hypothesis

$H_0: \theta \le \theta_0$ vs. $H_1: \theta > \theta_0$.

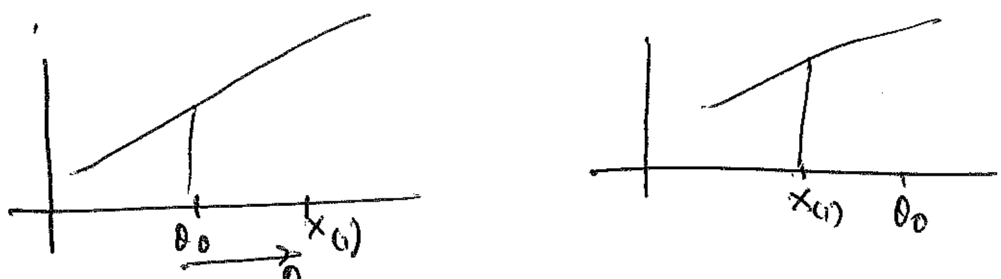
$$\sup_{\theta \in \Theta_0} f_\theta(\underline{x}) = \sup_{\theta \le \theta_0} f_\theta(\underline{x})$$

$f_\theta(\underline{x})$ is increasing in θ,

when $\theta_0 \le x_{(1)} \Rightarrow f_\theta(\underline{x})$ is maximized at θ_0 in the region $\theta \le \theta_0$

when $\theta_0 > x_{(1)} \Rightarrow f_\theta(\underline{x})$ is maximized at $x_{(1)}$ in the region $\theta \le \theta_0$

$$\sup_{\theta \le \theta_0} f_\theta(\underline{x}) = \begin{cases} \exp\left\{-\sum_{i=1}^n x_i + n\theta_0\right\}, & \theta_0 \le x_{(1)} \\ \exp\left\{-\sum_{i=1}^n x_i + n x_{(1)}\right\}, & \theta_0 > x_{(1)} \end{cases}$$



$$\sup_{\theta \in \mathbb{R}} f_\theta(\underline{x}) = \exp\left\{-\sum_{i=1}^n x_i + n x_{(1)}\right\}$$

$$\Rightarrow \lambda = \frac{\sup_{\theta \le \theta_0} f_\theta(\underline{x})}{\sup_{\theta \in \mathbb{R}} f_\theta(\underline{x})} = \begin{cases} \exp\left\{n(\theta_0 - x_{(1)})\right\}, & \theta_0 \le x_{(1)} \\ 1 & \theta_0 > x_{(1)} \end{cases}$$

LRT test reject H_0, if $\lambda < c$, where c is some constant depending on the level of the test

$\lambda < c \iff \exp\{n(\theta_0 - X_{(1)})\} < c$

$\iff X_{(1)} \geq \theta_0 - \frac{\log c}{n}$.

If the above holds for the observed data, the null hypothesis H_0 is rejected.

In this example LRT test dependent on the data only through a sufficient statistic.

By factorization theorem, if $T(\underline{X})$ is a sufficient statistic then

$$f_\theta(\underline{X}) = g_\theta(T(\underline{X}))\, h(\underline{X})$$

$$\lambda = \frac{\sup_{\theta \in \Theta_0} f_\theta(\underline{X})}{\sup_{\theta \in \Theta} f_\theta(\underline{X})} = \frac{\sup_{\theta \in \Theta_0} g_\theta(T(\underline{X}))\, \cancel{h(\underline{X})}}{\sup_{\theta \in \Theta} g_\theta(T(\underline{X}))\, \cancel{h(\underline{X})}}$$

$\Theta = \Theta_0 \cup \Theta_1$

$$= \frac{\sup_{\theta \in \Theta_0} g_\theta(T(\underline{X}))}{\sup_{\theta \in \Theta} g_\theta(T(\underline{X}))}$$

~~ce Intersection~~

Union-Intersection test and Intersection-Union test

lets say the null hypothesis can be expressed as $H_0 : \theta \in \bigcap_{\gamma \in \Gamma} \Theta_\gamma$

Γ is an indexing set.

$\{a_n : n \geq 1\}$ $\mathbb{N} \leftarrow$ countable indexing.

$\{a_1, a_2, a_3, a_4\}$ indexing set $\{1, 2, 3, 4\}$

Γ is an indexing set which is either finite, countably infinite or uncountably infinite.

$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$

$H_0 : \mu = \mu_0 = \{\mu \leq \mu_0\} \cap \{\mu \geq \mu_0\}$.

In this case, can we use individual tests on $H_0 : \theta \in \Theta_\gamma$ vs. $H_1 : \theta \in \Theta_\gamma^c$ to draw some inference on the full hypotheses, $H_0 : \theta \in \bigcap_{\gamma \in \Gamma} \Theta_\gamma$ vs. $H_1 : \theta \in \left(\bigcap_{\gamma \in \Gamma} \Theta_\gamma\right)^c$.

Lets try to set up some testing procedure for the above type of null hypothesis.

Let $R_\gamma = \{\underline{X} : T_\gamma(X) \in S_\gamma\}$ be the rejection region for the test

$H_{0\gamma} : \theta \in \Theta_\gamma$ vs. $H_{1\gamma} : \theta \in \Theta_\gamma^c$.

We define the rejection region for a test, known as the Union-Intersection test, by

$$\bigcup_{\gamma \in \Gamma} R_\gamma .$$

The rationale is that if any one of the $H_{0\gamma}$ is rejected, then H_0 is rejected.

Intersection union test is applied when H_0 is expressed as $H_0 : \theta \in \bigcup_{\gamma \in \Gamma} \Theta_\gamma$.

Suppose $R_\gamma = \{ \underline{X} : T_\gamma(\underline{X}) \in S_\gamma \}$ is the rejection region for $H_{0\gamma}$. Intersection-union test has a rejection region given by

$$\bigcap_{\gamma \in \Gamma} R_\gamma .$$

Intersection-Union	Union-Intersection
$H_0 : \theta \in \bigcup_{\gamma \in \Gamma} \Theta_\gamma$	$H_0 : \theta \in \bigcap_{\gamma \in \Gamma} \Theta_\gamma$
$\bigcap_{\gamma \in \Gamma} R_\gamma$	$\bigcup_{\gamma \in \Gamma} R_\gamma .$

$$H_0 : \mu = \mu_0 = \{\mu \le \mu_0\} \cap \{\mu \geq \mu_0\} .$$

$$H_0 : \mu \geq 0 = \{\mu \le 5\} \cup \{\mu \geq 5\}$$

Theorem: Consider a UIT for testing $H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_0^c$, where $\Theta_0 = \bigcap_{\gamma \in \Gamma} \Theta_\gamma$.

Let $\lambda_\gamma(\underline{X})$ be the LRT statistic for testing $H_{0\gamma}$ vs. $H_{1\gamma}$. Let $\lambda(\underline{X})$ be the LRT statistic for testing $H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_0^c$.

Define $T(\underline{X}) = \inf_{\gamma \in \Gamma} \lambda_\gamma(\underline{X})$ so that UIT rejection region is $\bigcup R_\gamma = \{\underline{x} : T(\underline{x}) < c\}$ where $R_\gamma = \{\underline{x} : \lambda_\gamma(\underline{x}) < c\}$. Then

1. $T(\underline{X}) \geq \lambda(\underline{X}) \quad \forall \underline{x}$
2. If $\beta_T(\theta)$ and $\beta_\lambda(\theta)$ are power functions based on the tests T and λ respectively, then $\beta_T(\theta) \leq \beta_\lambda(\theta) \quad \forall \theta$.

Pf: $\lambda(\underline{x})$ is the LRT test statistic for the entire region. $\qquad \Theta_0 = \bigcap_{\gamma \in \Gamma} \Theta_\gamma$

~~clearly the numerator~~

$$\lambda(\underline{x}) = \frac{\sup_{\theta \in \Theta_0} f_\theta(\underline{x})}{\sup_{\theta \in \Theta} f_\theta(\underline{x})} \leq \frac{\sup_{\theta \in \Theta_{0\gamma}} f_\theta(\underline{x})}{\sup_{\theta \in \Theta} f_\theta(\underline{x})} = \lambda_\gamma(\underline{x})$$

$$\Rightarrow \lambda(\underline{x}) \leq \lambda_\gamma(\underline{x}) \quad \forall \gamma \in \Gamma$$

$$\Rightarrow \lambda(\underline{x}) \leq \inf_{\gamma \in \Gamma} \lambda_\gamma(\underline{x}) = T(\underline{x})$$

$\beta_T(\theta) = P_\theta(T(\underline{X}) < c) \le P_\theta(\lambda(\underline{X}) < c) = \beta_\lambda(\theta)$

Hence the full LRT test is more powerful than the union intersection test.

Usefulness:

① By looking at various tests in UIT you might get additional information.

$P_\theta(T(\underline{X}) < c)$

$\{\underline{x} : \lambda(\underline{x}) < c\} \subset \{\underline{x} : T(\underline{x}) < c\}$

② Sometimes it is difficult to construct LRT for $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$, the reason being optimization over Θ_0 can be difficult.

Theorem: Let α_γ be the size of the test of $H_{0\gamma}$ with the rejection region R_γ. Then the Intersection-union test with the rejection region $\bigcap_{\gamma \in \Gamma} R_\gamma$ rejects H_0 at ~~test~~ level $\alpha = \sup_{\gamma \in \Gamma} \alpha_\gamma$.

Pf: Let $\theta \in \Theta_0$, then $\theta \in \Theta_\gamma$ for some γ.

and $P_\theta(\underline{X} \in \bigcap_{\gamma \in \Gamma} R_\gamma) \le P_\theta(\underline{X} \in R_\gamma, \gamma \in \Gamma)$

$= \alpha_\gamma \le \alpha$

$\Theta_0 = \bigcup_{\gamma \in \Gamma} \Theta_\gamma$

Hence for IUT test, it is easy to achieve the level requirement.

Ex: Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, μ, σ^2 both unknown.

Consider testing $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$.

$H_0: \{\mu = \mu_0\} = \{\mu \le \mu_0\} \cap \{\mu \ge \mu_0\}$.

Hence we can carry out LRT test for

$H_{01}: \mu \le \mu_0$ vs. $H_{11}: \mu > \mu_0$

and

$H_{02}: \mu \ge \mu_0$ vs. $H_{02}: \mu < \mu_0$

clearly $\Theta_1 = \{\mu \le \mu_0\}$ $\Theta_2 = \{\mu \ge \mu_0\}$

indexing set $\Gamma = \{1, 2\}$.

LRT test $H_{01}: \mu \le \mu_0$ vs. $H_{11}: \mu > \mu_0$ rejects H_{01} if $\sqrt{n}\left(\frac{\bar{X} - \mu_0}{S}\right) \ge t_1$, $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Another LRT test ~~that~~ rejects $H_{02}: \mu \ge \mu_0$ vs. $H_{12}: \mu < \mu_0$ if $\sqrt{n}\left(\frac{\bar{X} - \mu_0}{S}\right) \le t_2$.

Rejection region of Union-Intersection test is

$$\left\{\bar{X} : \sqrt{n}\left(\frac{\bar{X} - \mu_0}{S}\right) \le t_2\right\} \cup \left\{\bar{X} : \sqrt{n}\left(\frac{\bar{X} - \mu_0}{S}\right) \ge t_1\right\}.$$

If the level of the test is α and we divide it equally into both tails, we solve for t_1 and t_2 based on the following equations -

back (1)

$$P_{\mu_0}\left(\sqrt{n}\,\frac{(\bar{X}-\mu_0)}{S} \le t_2 \right) = \alpha/2$$

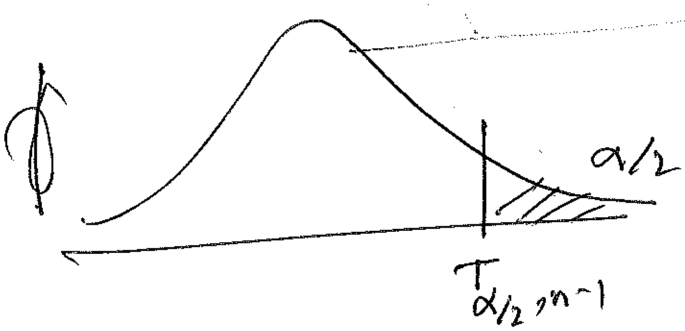
$$P_{\mu_0}\left(\sqrt{n}\,\frac{(\bar{X}-\mu_0)}{S} \ge t_1 \right) = \alpha/2$$

$$t_2 = -T_{\alpha/2,\,n-1} \qquad t_1 = T_{\alpha/2,\,n-1}$$

rejection region

$$\left\{ \bar{X} : \sqrt{n}\,\frac{(\bar{X}-\mu_0)}{S} \le -T_{\alpha/2,\,n-1} \right\} \cup \left\{ \bar{X} : \sqrt{n}\,\frac{(\bar{X}-\mu_0)}{S} \ge T_{\alpha/2,\,n-1} \right\}$$

$$\Rightarrow \text{reject when } \boxed{\left| \sqrt{n}\,\frac{(\bar{X}-\mu_0)}{S} \right| > T_{\alpha/2,\,n-1}}$$



Union-intersection test in a two sided t-test.

Bayesian testing :

$H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_1$

A Bayesian will look at $P(\theta \in \Theta_0 \mid \underline{X})$ and $P(\theta \in \Theta_1 \mid \underline{X})$.

Simplest Bayesian testing procedure will reject H_0 if $P(\theta \in \Theta_0 \mid \underline{X}) < 0.5$.

back ②

Monday 9:30

Recap:

We have seen (1) UMP test (2) UMPU test.

and point estimation

(3) Interval estimation.

Definition (Interval estimation) An interval estimate of a real valued parameter θ is defined through a pair of functions $L(X_1, \ldots, X_n)$ and $U(X_1, \ldots, X_n)$ [where $X_1, \ldots, X_n$ is the random sample from a distribution parametrized by θ] that satisfy

$L(X_1, \ldots, X_n) \le U(X_1, \ldots, X_n)$ for all $\underline{x} = (x_1, \ldots, x_n)$.

The random interval $[L(\underline{X}), U(\underline{X})]$ is called an interval estimator of θ.

Now, $L(\underline{X})$ can be $-\infty$ or $U(\underline{X})$ can be $+\infty$.

$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$ and the UMVUE is $\bar{X}$. $P_\mu(\mu = \bar{X}) = 0$ therefore the true value of μ will exactly be the point estimator with zero prob.

However, $P_\mu\left(\bar{X} - z_{\alpha/2}\frac{1}{\sqrt{n}} \le \mu \le \bar{X} + z_{\alpha/2}\frac{1}{\sqrt{n}}\right)$

$= 1-\alpha$. Since, $\sqrt{n}(\bar{X}-\mu) \sim N(0,1)$.

Thus $\left[\bar{X} - z_{\alpha/2}\frac{1}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{1}{\sqrt{n}}\right]$ has $100(1-\alpha)\%$ chance of including the true parameter μ.

This quantity $(1-\alpha)$ is called the coverage prob.

Def: For an interval estimator $[L(\underline{X}), U(\underline{X})]$ of a parameter θ, the coverage probability of $[L(\underline{X}), U(\underline{X})]$ is the probability that the random interval covers the true parameter θ. In symbol, it is denoted by $P_\theta(\theta \in [L(\underline{X}), U(\underline{X})])$

Suppose, $X_1, \ldots, X_n \overset{iid}{\sim} f_\theta$

For a random sample ~~$x_1, \ldots, x_n$~~ 5, 7, 1, 8, 13.

For this random sample calculated $L(\underline{x}) = 4$, $U(\underline{x}) = 9$. Your estimator of θ is $[4, 9]$.

$P_\theta(\theta \in [4, 9]) = 0.95$

It does not mean that θ has 95% chance of being included in the interval $[4, 9]$.

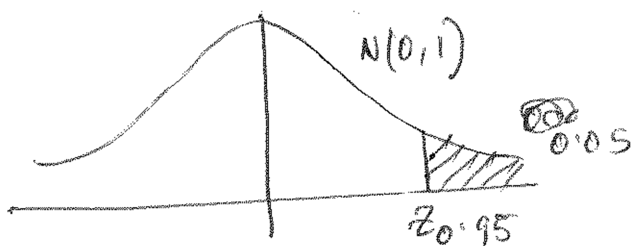
If you repeat this procedure of constructing $[L(\underline{X}), U(\underline{X})]$ for a large number of times, then 95% of these times θ will lie in those intervals constructed.

Thus frequentist coverage probabilities ~~are not~~ do not make a lot of sense for a single sample.

Def: For an interval estimator $[L(\underline{X}), U(\underline{X})]$ of a parameter θ, the confidence coefficient of $[L(\underline{X}), U(\underline{X})]$ is the infimum of the coverage probabilities, $\inf_{\theta} P_\theta\left[\theta \in [L(\underline{X}), U(\underline{X})]\right]$. Interval estimators together with the confidence coefficients is sometimes known as the confidence interval.

$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$. If we want 95% Interval estimator for μ.

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



$$P\left(-\infty < \sqrt{n}(\bar{X}-\mu) < Z_{0.95}\right) = 0.95$$

$$P\left(-Z_{0.95} < \sqrt{n}(\bar{X}-\mu) < \infty\right) = 0.95$$

$$P\left(-Z_{0.975} < \sqrt{n}(\bar{X}-\mu) < Z_{0.975}\right) = 0.95.$$

Interval estimators is not unique.

How to find interval estimators

We will discuss two different ways to find interval estimators.

① Inverting a test statistic

② Using a pivotal quantity.

Inverting a test statistic

$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, ~~$\sigma^2$ known and let~~ σ^2 known

~~$H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$.~~

Goal: Find $100(1-\alpha)\%$ Confidence interval for μ.

Lets think about the hypothesis testing

$H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$.

the H_0 is rejected if $\sqrt{n}\,|\bar{X} - \mu_0| \geq \sigma z_{\alpha/2}$.

or,

$$P_{\mu_0}\left(\bar{X} - \sigma z_{\alpha/2} \frac{1}{\sqrt{n}} \leq \mu_0 \leq \bar{X} + \sigma z_{\alpha/2} \frac{1}{\sqrt{n}}\right) = 1 - \alpha \quad \cdots ①$$

probability of not rejecting H_0 when it is true

① is true for any μ, i.e.

$$P_{\mu}\left(\bar{X} - \sigma z_{\alpha/2} \frac{1}{\sqrt{n}} \leq \mu \leq \bar{X} + \sigma z_{\alpha/2} \frac{1}{\sqrt{n}}\right) = 1 - \alpha.$$

Thus $\left[\bar{X} - \sigma z_{\alpha/2} \frac{1}{\sqrt{n}}, \bar{X} + \sigma z_{\alpha/2} \frac{1}{\sqrt{n}}\right]$ has a coverage prob. of $(1-\alpha)$.

Hence this is $100(1-\alpha)\%$ CI for μ.

Result: For each $\theta_0 \in \Theta$, let $A(\theta_0)$ be the acceptance region of a level α test of $H_0: \theta = \theta_0$.

For each $\underline{X} = (X_1, \ldots, X_n) \in \mathcal{X}$, define a set $C(\underline{X})$ in the parameter space by

$C(\underline{X}) = \{\theta_0 : \underline{X} \in A(\theta_0)\}$. Then the random set $C(\underline{X})$ is the $(1-\alpha)$ confidence set.

$$P_\theta(\theta \in C(\underline{X})) = P_\theta(\underline{X} \in A(\theta)) = P_{\theta_0}(\underline{X} \in A(\theta_0)) = 1-\alpha.$$

Inverting an LRT Statistic:

Let $X_1, \ldots, X_n \overset{iid}{\sim} Exp(\lambda)$.

Goal: $100(1-\alpha)\%$ CI for λ.

Test: $H_0: \lambda = \lambda_0$ vs. $H_1: \lambda \neq \lambda_0$.

LRT test statistic

$$L = \frac{\left(\frac{1}{\lambda_0}\right)^n \exp\left(-\sum_{i=1}^n x_i/\lambda_0\right)}{\sup_{\lambda > 0} \left(\frac{1}{\lambda}\right)^n \exp\left(-\sum_{i=1}^n x_i/\lambda\right)}$$

$$\arg\sup_{\lambda > 0} \left(\frac{1}{\lambda}\right)^n \exp\left(-\sum_{i=1}^n x_i/\lambda\right) = \frac{\sum_{i=1}^n x_i}{n}$$

$$\ell(\lambda) = -n\log\lambda - \sum_{i=1}^n \frac{x_i}{\lambda}$$

$$\ell'(\lambda) = 0 \Rightarrow -\frac{n}{\lambda} + \sum_{i=1}^n \frac{x_i}{\lambda^2} = 0$$

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

$$L = \frac{\left(\frac{1}{\lambda_0}\right)^n \exp\left(-\sum_{i=1}^{n} x_i/\lambda_0\right)}{\left(\frac{1}{\bar{x}}\right)^n \exp\left(-\sum_{i=1}^{n} x_i/\bar{x}\right)} = \frac{\left(\frac{1}{\lambda_0}\right)^n \exp\left(-\sum_{i=1}^{n} x_i/\lambda_0\right)}{\left(\frac{1}{\bar{x}}\right)^n \exp(-n)}$$

For a fixed λ_0, the acceptance region is given by

$$A(\lambda_0) = \left\{ \underline{x} : \left(\frac{\sum_{i=1}^{n} x_i}{\lambda_0}\right)^n e^{-\sum_{i=1}^{n} x_i/\lambda_0} \geq K^* \right\}$$

ignoring constants like $\exp(-n)$.

Inverting this LRT we will obtain the $100(1-\alpha)\%$ Confidence set as

$$C(\underline{x}) = \left\{ \lambda : \left(\frac{\sum_{i=1}^{n} x_i}{\lambda}\right)^n \exp\left(-\sum_{i=1}^{n} x_i/\lambda\right) \geq K^* \right\}$$

clearly the confidence set to be an interval $U(\underline{x})$ and $L(\underline{x})$ must be only functions of $\sum_{i=1}^{n} x_i$.

Thus $C\left(\sum_{i=1}^{n} x_i\right) = \left\{ \lambda : L\left(\sum_{i=1}^{n} x_i\right) \leq \lambda \leq U\left(\sum_{i=1}^{n} x_i\right) \right\}$.

Note that $\frac{\sum_{i=1}^{n} x_i}{\lambda} \sim \text{Gama}(n, 1)$.

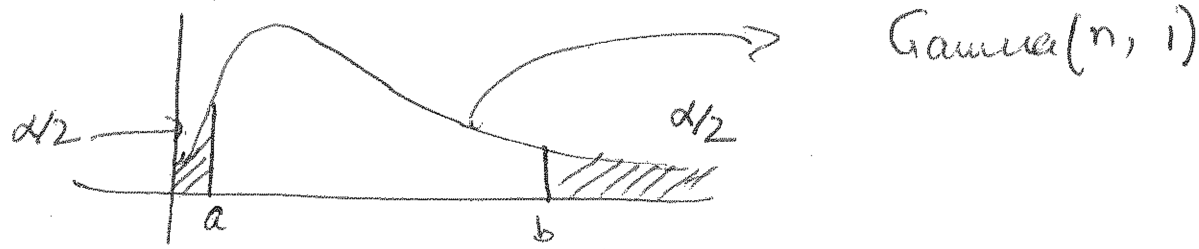
$X_1 \sim \text{Exp}(\lambda)$ $\quad X_2 \sim \text{Exp}(\lambda) \Rightarrow X_1 + X_2 \sim \text{Gama}(2, \lambda)$

Gama$(1, \lambda)$

$$\frac{e^{-x/\lambda} \cdot x^{1-1}}{\Gamma(1)} \frac{1}{\lambda}$$

back ①

$$\frac{\sum_{i=1}^{n} X_i}{\lambda} \sim \text{Gamma}(n, 1)$$



$$P_\lambda\left(a < \frac{\sum_{i=1}^{n} X_i}{\lambda} < b\right) = 1-\alpha$$

$$\cancel{\Rightarrow P_\lambda\left(\frac{b}{\sum_{i=1}^{n} X_i} < \lambda < \frac{a}{\sum_{i=1}^{n} X_i}\right) = 1-\alpha}$$

$$\Rightarrow P_\lambda\left(\frac{\sum_{i=1}^{n} X_i}{b} < \lambda < \frac{\sum_{i=1}^{n} X_i}{a}\right) = 1-\alpha$$

$\Rightarrow \left[\frac{\sum_{i=1}^{n} X_i}{b}, \frac{\sum_{i=1}^{n} X_i}{a}\right]$ is the $100(1-\alpha)\%$ CI for λ.

Interval estimation based on pivotal quantities

Def: A random variable $Q(\underline{X}, \theta) = Q(X_1, \ldots, X_n, \theta)$ is a pivotal quantity (or pivot) if the distribution of $Q(\underline{X}, \theta)$ is independent of the parameter θ.

Ex! $X_1, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\lambda)$

$$\frac{\sum_{i=1}^{n} X_i}{\lambda} \sim \text{Gamma}(n, 1)$$

$$\frac{X_1 + X_2}{\lambda} \sim \text{Gamma}(2, 1)$$

back ②

$X \sim U(0,\theta)$

$\frac{X}{\theta} \sim U(0,1)$

Pivotal quantities for different families of distributions

(i) Location family : when $X_i \sim f(x-\mu)$. Here $\bar{X}-\mu$ is a pivotal quantity.

$= \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)$ $\quad X_i - \mu \sim f(x)$

(ii) Scale family : when $X_i \sim \frac{1}{\sigma} f(\frac{x}{\sigma})$

Here $\frac{\bar{X}}{\sigma}$ is a pivotal quantity.

$\frac{\bar{X}}{\sigma} = \frac{1}{n} \sum_{i=1}^{n} \frac{X_i}{\sigma}$ and $\frac{X_i}{\sigma} \sim f(x)$.

(iii) location scale family : when $X_i \sim \frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$.

$\frac{\bar{X}-\mu}{\sigma}$ is a pivotal quantity.

$\frac{\bar{X}-\mu}{S}$ is also a pivotal quantity.

Example: $Exp(\lambda)$ is a scale family,

$N(\mu, \sigma^2)$ is a location scale family.

If we have a pivotal quantity $Q(\underline{X}, \theta)$, it is easy to ~~come~~ construct a confidence interval for θ.

back ③

Since $Q(\underline{X}, \theta)$ has a distribution free of θ.
We will be able to find a, b from this distribution (a, b free of θ) s.t.

$$P_\theta\left(a < Q(\underline{X}, \theta) < b\right) = 1 - \alpha.$$

Clearly ~~for each~~ $\cancel{\theta,\ A(\theta) = \{\underline{x} : a < Q(\underline{x}, \theta) < b\}}$
~~is the~~ ac $Q(\underline{X}, \theta)$ is monotone in θ

$$\Rightarrow \quad P_\theta\left(a < Q(\underline{X}, \theta) < b\right) = P_\theta\left(L(\underline{X}) < \theta < U(\underline{X})\right)$$

$$\left\{a < Q(\underline{X}, \theta) < b\right\}$$

$\Rightarrow$ Another set $\left\{L(\underline{X}) < \theta < U(\underline{X})\right\}$

such that $P_\theta\left(a < Q(\underline{X}, \theta) < b\right) = P_\theta\left(L(\underline{X}) < \theta < U(\theta)\right)$

$$P_\lambda\left(a < \frac{\sum_{i=1}^{n} X_i}{\lambda} < b\right) = P_\lambda\left(\frac{\sum_{i=1}^{n} X_i}{b} < \lambda < \frac{\sum_{i=1}^{n} X_i}{a}\right)$$

thus just by inverting the pivotal quantity a $100(1-\alpha)\%$ confidence interval can be constructed.