e.g) Error bounds

① $\theta = \mu \qquad \hat{\theta} = \bar{Y}$

2 s.e. bound $= \pm 2 \cdot \frac{\sigma}{\sqrt{n}} \approx \pm 2 \cdot \frac{S}{\sqrt{n}}$

② $\theta = p \qquad \hat{\theta} = \frac{Y}{n}$

2 · s.e. bound $= \pm 2 \cdot \sqrt{\frac{p(1-p)}{n}} \approx \pm 2 \sqrt{\frac{\frac{Y}{n}(1-\frac{Y}{n})}{n}}$

8.37) $n = 10 \qquad Y_1 \cdots Y_{10} \sim \text{Exp}(\theta)$

$\bar{Y} = 1020$ hrs $= \hat{\theta}$

$S.e.(\hat{\theta}) = \sqrt{Var(\hat{\theta})} = \sqrt{Var(\bar{X})} = \sqrt{\frac{Var(Y)}{n}} = \sqrt{\frac{\theta^2}{n}}$

$= \frac{\theta}{\sqrt{n}}$

2 · S.E bound $= \pm 2 \frac{\hat{\theta}}{\sqrt{n}} = \pm 2 \cdot \frac{1020}{\sqrt{10}}$

; not a 95% bound (rather 75%)

Confidence Intervals

99, 95, → about procedure of obtaing C.I. not about any single C.I.

$100(1-\alpha)\,\%$ confidence interval → Random Variable !!

- Interval of values that is likely to contain the true value of the parameter θ 99% C.I. $(3.5,\ 6.9)$ ~~$P(\theta \in (3.5, 6.9)) = .99$~~?
- The long-run relative frequency over many repetitions of sampling → random variable
- $P(\hat{\theta}_L \le \theta \le \hat{\theta}_U) = 1-\alpha$: two-sided confidence interval
- $P(\hat{\theta}_L \le \theta) = 1-\alpha$: one-sided confidence ~~interval~~ bound
  $P(\theta \le \hat{\theta}_U) = 1-\alpha$: one-sided confidence ~~interval~~ bound

Pivotal method to construct C.I.

Find a pivotal quantity that possesses two characteristics

- a function of the sample measurements and the unknown parameter θ, where θ is the only unknown quantity.
  statistic
- Second, Its probability distribution does not depend on the parameter θ.

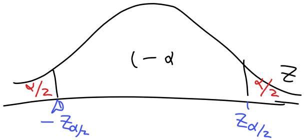
Construct Interval for pivotal quantity

$\Longrightarrow$ pivot to make C. I. for θ.

e.g

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

pivotal



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

pivot

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

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

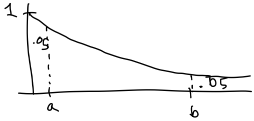
$$\bar{Y} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Example Y single obs $\sim Exp(\theta)$

$X = \frac{Y}{\theta} \sim Exp(1)$ ← check!

: pivotal.

90% C.I.



$a = .051 \qquad b = 2.996$

$$P\left(0.051 \le \frac{Y}{\theta} \le 2.996\right) = .90$$

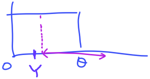
$$\Rightarrow P\left(\frac{Y}{2.996} \le \theta \le \frac{Y}{.051}\right) = .90$$

90%

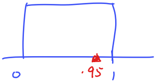
$$\left[\frac{Y}{2.996}, \frac{Y}{.051}\right] : \text{C.I.}$$

$\boxed{\text{Eg 8.5}}$ $Y \sim U(0, \theta)$

95% lower confidence bound of θ?



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



$$P\left(\frac{Y}{\theta} < .95\right) = .95$$

$$\Rightarrow P\left(\theta > \frac{Y}{.95}\right) = .95$$

$$\left[\frac{Y}{.95}, \infty\right)$$

Ex 8.48 $\quad Y_1 \cdots Y_n \sim \text{Gamma}(2, \beta)$

$\Sigma Y_i \sim \text{Gamma}(2n, \beta)$

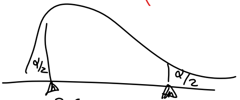
$\frac{2}{\beta} \Sigma Y_i \sim \text{Gamma}\left(2n, \beta \cdot \frac{2}{\beta}\right)$

$= \text{Gamma}(2n, 2)$

$= \chi^2(4n)$

a) $\dfrac{2\Sigma Y_i}{\beta} \sim \chi^2(4n)$

b)



$\chi^2_{1-\alpha/2}(4n) = a \qquad \chi^2_{\alpha/2}(4n) = b$ from χ^2-table

$$P\left(a \le \frac{2n\bar{Y}}{\beta} \le b\right) = 1-\alpha$$

$$\Rightarrow P\left(\frac{2n\bar{Y}}{b} \le \beta \le \frac{2n\bar{Y}}{a}\right) = 1-\alpha .$$

c) $n = 5$. $\bar{y} = 5.39$

find C.I. $\therefore \quad (\qquad , \qquad) \rightarrow$ exact

try this! If n is large, approx. to Normal . . . . .

. . . . . . $(\qquad , \qquad) \rightarrow$ approximate

Large-Sample Confidence Intervals

Wald C.I.

- For large samples,

$$Z = \frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}}$$

  possesses approximately a standard normal distribution.
- Z forms a pivotal quantity
- $100(1-\alpha)$ % confidence interval (with normal approximation)
  (point estimate) $\pm z_{\alpha/2}$(standard error)
  $= \hat{\theta} \pm z_{\alpha/2}\sigma_{\hat{\theta}}$
- As the confidence level $1-\alpha$ increases, the width of C.I. becomes larger.
- As the sample size increases, the width of C.I. becomes smaller.

Width of C.I. depend on

n, $1-\alpha$, "underlying uncertainty" (σ)

cannot control.

$\bar{Y} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

① $\theta = \mu$

$\bar{X} \pm z_{\alpha/2} \frac{S}{\sqrt{n}}$: approximate C.I.

② $\theta = \mu_1 - \mu_2$

$\bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$

$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ is pivotal.

③ $\theta = p$

$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

In practice

$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\frac{1}{2}(1-\frac{1}{2})}{n}}$

to be conservative.

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

$$\left(\frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n}}{1 + \frac{z_{\alpha/2}^2}{2n}} \pm \sqrt{\bigcirc} \right)$$

better coverage,

less utility.

④ $\theta = p_1 - p_2$

$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Sample Size Determination

How large n should be to achieve a certain degree of utility with high confidence?

Given $1-\alpha$, desired error bound, d, we calculate n.

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \longrightarrow d = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow n = \left(\frac{z_{\alpha/2}\,\sigma}{d}\right)^2$$

given

given

σ : unknown $\left(\text{or use } \sigma \approx \dfrac{\max - \min}{4} \right)$