Inference for σ^2

Inference for σ^2

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

Theorem 5 provides the basis for inference for σ^2.

Suppose $Y_1, Y_2, \ldots, Y_n$ are i.i.d. $N(\mu, \sigma^2)$. Let c_1 and c_2 be such that

$$P(c_1 < X < c_2) = 1 - \alpha$$

where $\boxed{X \sim \chi^2_{n-1}}$.

Then

$$\boxed{\left(\frac{(n-1)S^2}{c_2}, \frac{(n-1)S^2}{c_1}\right)}$$

is the $100(1-\alpha)\%$ confidence interval for σ^2.

Proof of CI for σ^2

wts $P\left(\frac{(n-1)S^2}{c_2} < \sigma^2 < \frac{(n-1)S^2}{c_1} \right) = 1-\alpha$

Know $1-\alpha = P(c_1 < X < c_2)$, where $X \sim \chi^2_{n-1}$

and $X = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

swap

$$\underbrace{P\left(c_1 < \boxed{\frac{(n-1)S^2}{\sigma^2}} < c_2 \right)}_{\substack{\shortparallel \\ 1-\alpha}} = P\left(\frac{c_1}{(n-1)S^2} < \frac{1}{\sigma^2} < \frac{c_2}{(n-1)S^2} \right)$$

by definition of c_1 & c_2

$$= P\left(\frac{(n-1)S^2}{c_1} > \sigma^2 > \frac{(n-1)S^2}{c_2} \right)$$

$$= P\left(\underbrace{\frac{(n-1)S^2}{c_2}}_{L} < \sigma^2 < \underbrace{\frac{(n-1)S^2}{c_1}}_{U} \right)$$

QED

Proof of CI for σ^2

Choosing c_1 and c_2

c_1 and c_2 can be chosen in various ways consistent with

$$P(c_1 < X < c_2) = 1 - \alpha$$

where $X \sim \chi^2_{n-1}$.

$P(X < c_1) = \beta < \alpha$

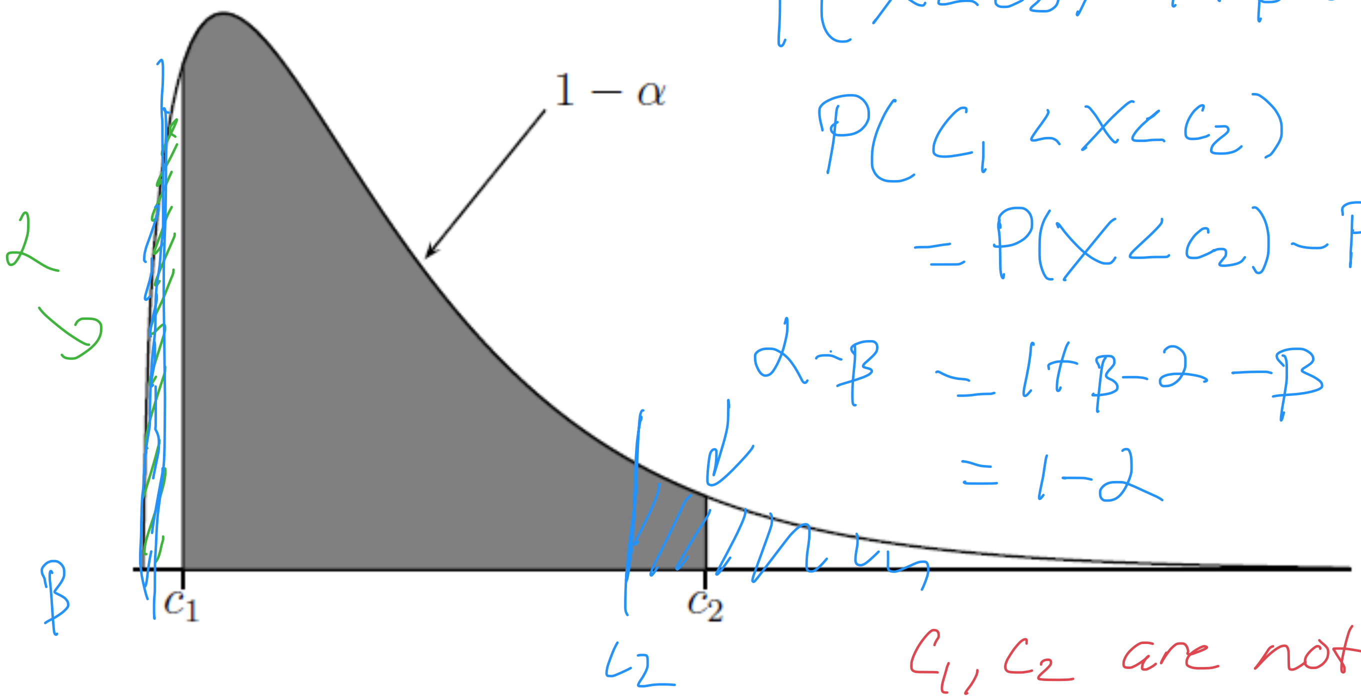
$P(X < c_2) = 1 + \beta - \alpha$

$P(c_1 < X < c_2)$

$= P(X < c_2) - P(X < c_1)$

$\alpha - \beta$ $= 1 + \beta - \alpha - \beta$

$= 1 - \alpha$

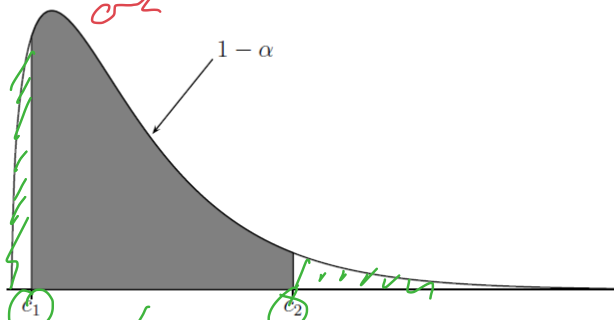


c_1, c_2 are not unique!

Choosing c_1 and c_2

Recall we are talking about σ^2

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

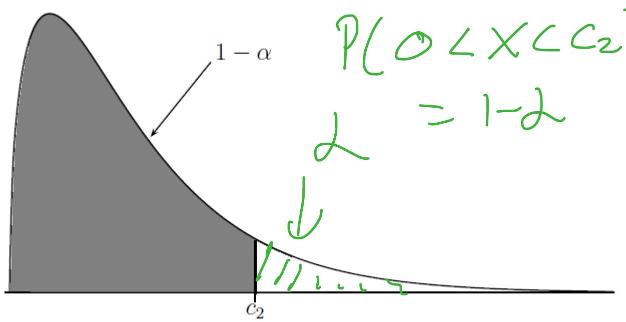


$\frac{\alpha}{2}$

Symmetric:

$$P(X < c_1) = \alpha/2$$
$$P(X > c_2) = \alpha/2$$

C.I. σ^2: $\left(\frac{(n-1)S^2}{c_2}, \frac{(n-1)S^2}{c_1}\right)$



$P(0 < X < c_2) = 1-\alpha$

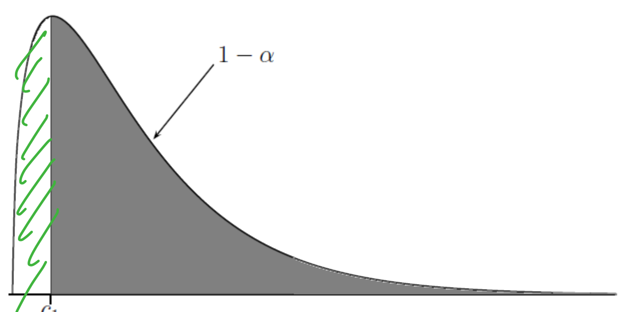
α

One-sided lower CI:

$$\boxed{P(X > c_2) = \alpha}$$
$$c_1 = 0$$

σ^2: $\left(\frac{(n-1)S^2}{c_2}, \infty\right)$

α



One-sided upper CI:

$$\boxed{c_2 = \infty}$$
$$P(X < c_1) = \alpha$$

σ^2: $\left(0, \frac{(n-1)S^2}{c_1}\right)$

Hypothesis test for σ^2

The confidence intervals described in the previous slide can also be used to derived one-sided and two-sided hypothesis tests for σ^2.

① $H_0: \sigma^2 = \sigma_0^2$ vs $H_a: \sigma^2 \neq \sigma_0^2$

② have n, S^2, choose $\alpha = 0.05$

③ calculate $\left(\frac{(n-1)S^2}{c_2}, \frac{(n-1)S^2}{c_1} \right) = CI$

④ ask is $\sigma_0^2 \in CI$?

if yes, retain, if no, reject

Example 2.9

A juice company wants to find out the variation, as measured by variance, of the amount of juice in 500mL bottles. The company statistician took a random of 25 bottles from the production line and compute the sample variance $s^2 = 2\ mL^2$. Find the 95% upper one-sided confidence interval for the variance σ^2.

$1-\alpha = 0.95$, $\boxed{\alpha = 0.05}$ where $P(X < c_1) = \alpha$, $X \sim \chi^2_{n-1}$

know $\left(0, \frac{(n-1)S^2}{c_1}\right)$

our 95% CI:

$$\left(0, \frac{(n-1)s^2}{c_1}\right) = \left(0, \frac{(25-1)2}{13.848}\right)$$

$$= \left(0, \frac{48}{13.848}\right)$$

$$\approx \left(0, \boxed{3.466}\right)$$

$n = 25$

$s^2 = 2$

$c_1 = \text{qchisq}(\alpha, n-1)$

≈ 13.848

So we are 95% confident that the true variation in 500mL bottles of juice lies between 0 and 3.466 mL^2.

Hypothesis test for $\sigma_1^2 = \sigma_2^2$

$Y_{11}, Y_{12}, \ldots, Y_{1n_1} \sim N(\mu_1, \sigma_1^2)$, $Y_{21}, Y_{22}, \ldots, Y_{2n_2} \sim N(\mu_2, \sigma_2^2)$

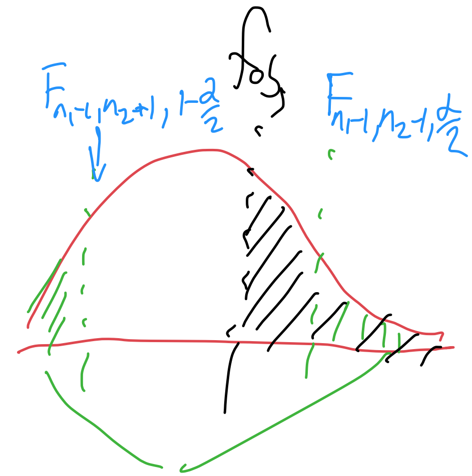
In the case of two samples, we want to test

$$\boxed{\begin{aligned} H_0&: \sigma_1^2 - \sigma_2^2 = 0, \\ H_a&: \sigma_1^2 - \sigma_2^2 \neq 0 \end{aligned}}$$

or $H_0: \sigma_1^2 = \sigma_2^2$

Recall from Example 2.4 that

$$\boxed{F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}} \sim F_{n_1-1, n_2-1}$$



critical region
reject H_0

Under H_0:

$$F = \frac{S_1^2/\cancel{\sigma_1^2}}{S_2^2/\cancel{\sigma_2^2}} = \boxed{\frac{S_1^2}{S_2^2}} \sim F_{n_1-1, n_2-1}$$

good test stat.

The critical region is $F > \boxed{F_{n_1-1, n_2-1, \frac{\alpha}{2}}}$ or $F < \boxed{F_{n_1-1, n_2-1, 1-\frac{\alpha}{2}}}$.

p-value: $\boxed{2 \min \{ P(F > f_{obs}), P(F < f_{obs}) \}}$ $\quad f_{obs} = \frac{S_1^2}{S_2^2}$

Example 2.10

An experiment to explore the pain threshold to electrical shocks for males and females resulted in the data summary given in the following table. Do the data represent sufficient evidence to suggest a difference in the variability of pain thresholds for between men and women? Use $\alpha = 0.10$. Also find the P-value.

Men	Women
$n_1 = 14$	$n_2 = 10$
$\bar{y}_1 = 16.2$	$\bar{y}_2 = 14.9$
$s_1^2 = 12.7$	$s_2^2 = 26.4$

Example 2.10 Solution

$\alpha = 0.1$
$n_1 = 14$
$n_2 = 10$
$S_1^2 = 12.7$
$S_2^2 = 26.4$

Let σ_1^2 be the true population variance of pain threshold for males
" σ_2^2 " " " " " " " " " females

$H_0: \sigma_1^2 = \sigma_2^2$ vs $H_a: \sigma_1^2 \neq \sigma_2^2$

Test stat: $F = \frac{S_1^2}{S_2^2} = \frac{12.7}{26.4}$

$f_{obs} \approx \boxed{0.4811} \sim F_{14-1,10-1}$ under H_0
$= F_{13,9}$

$\alpha = 0.1$, $\frac{\alpha}{2} = 0.05$, $1-\frac{\alpha}{2} = 0.95$

Critical Region:

$F_{13,9,1-\frac{\alpha}{2}} = F_{13,9,0.95} = qf(0.05, 13, 9) \approx 0.3684$

$1 - 0.95$

$F_{13,9,\frac{\alpha}{2}} = F_{13,9,0.05} = qf(0.95; 13, 9) \approx 3.0475$

$1 - 0.05$

Critical Region $\{F < 0.3684, F > 3.0475\}$

Well, since $f_{obs} = 0.4811$ is not in our critical region, there is insufficient evidence to reject H_0.

Example 2.10 Solution

$$P\text{-value} = 2\min\{\boxed{P(F<0.4811)}, P(F>0.4811)\}$$

$$F \sim F_{13,9} \qquad (pf)$$

$$p = 2 \times pf(0.4811, 13, 9) \approx 0.2237 > \alpha \ (0.1)$$