A First Guide to Hypothesis Testing in Linear Regression Models

A Generic Linear Regression Model: Scalar Formulation of Population Regression Equation

For the i-th population (or sample) observation, the **scalar formulation of the PRE,** is written as:

$$
\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + u_i && \forall\, i \\
&= \beta_0 + \sum_{j=1}^{j=k} \beta_j X_{ij} + u_i && \forall\, i \qquad \textbf{(1)} \\
&= \sum_{j=0}^{j=k} \beta_j X_{ij} + u_i\,, \quad X_{i0} = 1\ \forall\, i && \forall\, i
\end{aligned}
$$

where

$Y_i \equiv$ the i-th population value of the regressand, or dependent variable;

$X_{ij} \equiv$ the i-th population value of the j-th regressor, or independent variable;

$\beta_j \equiv$ the partial regression coefficient of X_{ij};

$u_i \equiv$ the i-th population value of the *unobservable* random error term.

Note:

(1) **Lower case "k"** denotes the **number of *slope* coefficients** in the PRF (population regression function).

(2) **Upper case "K"** denotes the ***total* number of regression coefficients** in the PRF.

(3) Therefore: **K = k + 1**.

Scalar Formulation of the Ordinary Least Squares (OLS) Sample Regression Equation

For the i-th sample observation, the **scalar formulation of the OLS SRE** is written as:

$$
\begin{aligned}
Y_i &= \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \cdots + \hat{\beta}_k X_{ik} + \hat{u}_i && \forall\ i = 1, \ldots, N \\
&= \hat{\beta}_0 + \sum_{j=1}^{j=k} \hat{\beta}_j X_{ij} + \hat{u}_i && \forall\ i = 1, \ldots, N \qquad \textbf{(2)} \\
&= \hat{Y}_i + \hat{u}_i && \forall\ i = 1, \ldots, N
\end{aligned}
$$

where

$Y_i \equiv$ the i-th sample value of the regressand, or dependent variable;

$X_{ij} \equiv$ the i-th sample value of the j-th regressor, or independent variable;

$\hat{\beta}_j \equiv$ the **OLS estimator or estimate** of the regression coefficient β_j;

$\hat{u}_i \equiv$ the i-th **OLS sample residual**;

$\hat{Y}_i = \hat{\beta}_0 + \sum_{j=1}^{j=k} \hat{\beta}_j X_{ij} \equiv$ the **OLS sample regression function (SRF)** for the i-th sample observation.

1. Tests of *One* Coefficient Restriction: *One* Restriction on *One* Coefficient

H_0 specifies ***only one equality*** **restriction** on ***one coefficient***.

❑ *Two-Tail* Tests of *One* Restriction on *One* Coefficient

- ***Example:*** H_0: $\beta_j = b_j$ versus H_1: $\beta_j \neq b_j$ where b_j is a specified constant.
- ***Use:*** either a ***two-tail*** **t-test** or an **F-test**.
- ***Test Statistics:***

⇨ $t(\hat{\beta}_j) = \dfrac{\hat{\beta}_j - \beta_j}{\hat{se}(\hat{\beta}_j)} \sim t[N-K]$ $\Rightarrow$ ***sample value*** $= t_0(\hat{\beta}_j) = \dfrac{\hat{\beta}_j - b_j}{\hat{se}(\hat{\beta}_j)}$.

⇨ $F(\hat{\beta}_j) = \dfrac{\left(\hat{\beta}_j - \beta_j\right)^2}{\hat{Var}(\hat{\beta}_j)} \sim F[1, N-K]$ $\Rightarrow$ ***sample value*** $= F_0(\hat{\beta}_j) = \dfrac{\left(\hat{\beta}_j - b_j\right)^2}{\hat{Var}(\hat{\beta}_j)}$.

Note: $\left[t(\hat{\beta}_j)\right]^2 = F(\hat{\beta}_j)$ or $t(\hat{\beta}_j) = \sqrt{F(\hat{\beta}_j)}$ and $\left[t_{\alpha/2}[N-K]\right]^2 = F_\alpha[1, N-K]$.

- ***Decision Rules – Two-Tail Tests of One Coefficient Restriction on One Coefficient:***

Reject **H_0** if $|t_0| > t_{\alpha/2}[N-K]$ *or* ***two-tail p-value*** for $t_0 = \Pr\left(|t| > |t_0| \mid H_0 \text{ is true}\right) < \alpha$;

$F_0 > F_{\alpha}[1, N-K]$ *or* ***p-value*** for $F_0 = \Pr(F > F_0 \mid H_0 \text{ is true}) < \alpha$.

Retain **H_0** if $|t_0| \le t_{\alpha/2}[N-K]$ *or* ***two-tail p-value*** for $t_0 = \Pr\left(|t| > |t_0| \mid H_0 \text{ is true}\right) \ge \alpha$;

$F_0 \le F_{\alpha}[1, N-K]$ *or* ***p-value*** for $F_0 = \Pr(F > F_0 \mid H_0 \text{ is true}) \ge \alpha$.

<u>Two-tail t test</u>: rejection and non-rejection regions

Two-tail critical values of t[50] distribution at 5% significance level (α = 0.05, α/2 = 0.025)

- ***upper*** **0.025 critical value of t[50] = $t_{\alpha/2}[50] = t_{0.025}[50] = 2.0086$**
- ***lower*** **0.025 critical value of t[50] = $-t_{\alpha/2}[50] = -t_{0.025}[50] = -2.0086$**

Stata **commands to compute** ***upper*** **0.025 critical value of t[50] = $t_{0.025}[50]$**

```
. display invttail(50, 0.025)
2.0085591

. display ttail(50, 2.0086)
.02499776

. display 2*ttail(50, 2.0086)
.04999551
```

two-tail rejection **region is $|t| > t_{\alpha/2}$:** $\Pr(|t| > t_{\alpha/2} \mid H_0 \text{ is true}) = \alpha = 0.05$

right-tail rejection **region is $t > t_{\alpha/2}$:** $\Pr(t > t_{\alpha/2} \mid H_0 \text{ is true}) = \alpha/2 = 0.05/2 = 0.025$

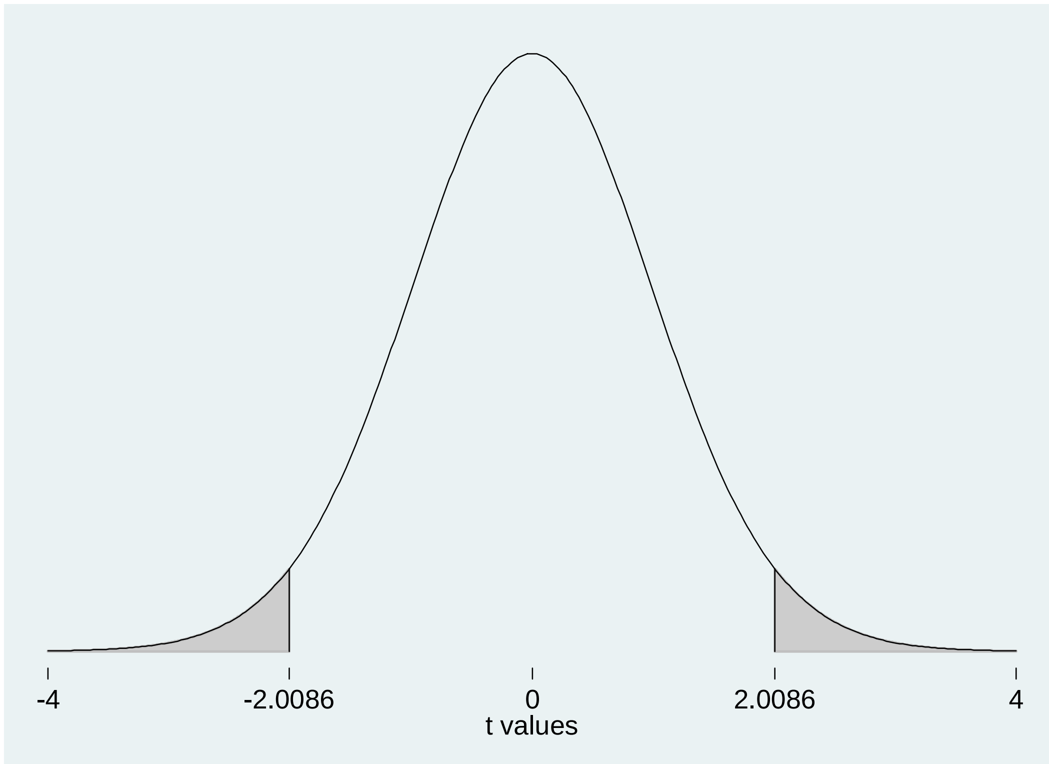
left-tail rejection **region is $t < -t_{\alpha/2}$:** $\Pr(t < -t_{\alpha/2} \mid H_0 \text{ is true}) = \alpha/2 = 0.05/2 = 0.025$

non-rejection **region is $-t_{\alpha/2} \le t \le t_{\alpha/2}$:**

$$\Pr(-t_{\alpha/2} \le t \le t_{\alpha/2} \mid H_0 \text{ is true}) = \Pr(|t| \le t_{\alpha/2} \mid H_0 \text{ is true}) = 1 - \alpha = 1 - 0.05 = 0.95$$

Two-tail t test: rejection and non-rejection regions (continued)

At 5% significance level, $\alpha = 0.05$ and $\alpha/2 = 0.025$: $t_{0.025}[50] = 2.0086$; $-t_{0.025}[50] = -2.0086$



$-t_{\alpha/2}[50]$ $t_{\alpha/2}[50]$

left-tail rejection region | non-rejection region | right-tail rejection region

area = $\alpha/2$ = 0.025 area = $1 - \alpha$ = 0.95 area = $\alpha/2$ = 0.025

<u>F test</u>: rejection and non-rejection regions

Critical value of F[1,50] at 5% significance level (α = 0.05) = F_α[1,50] = $F_{0.05}$[1,50] = 4.0343

- **0.05 critical value of F[1,50] = F_α[1,50] = $F_{0.05}$[1,50] = 4.0343**

  ***Stata* commands to compute 0.05 critical value of F[1,50] = $F_{0.05}$[1,50]**

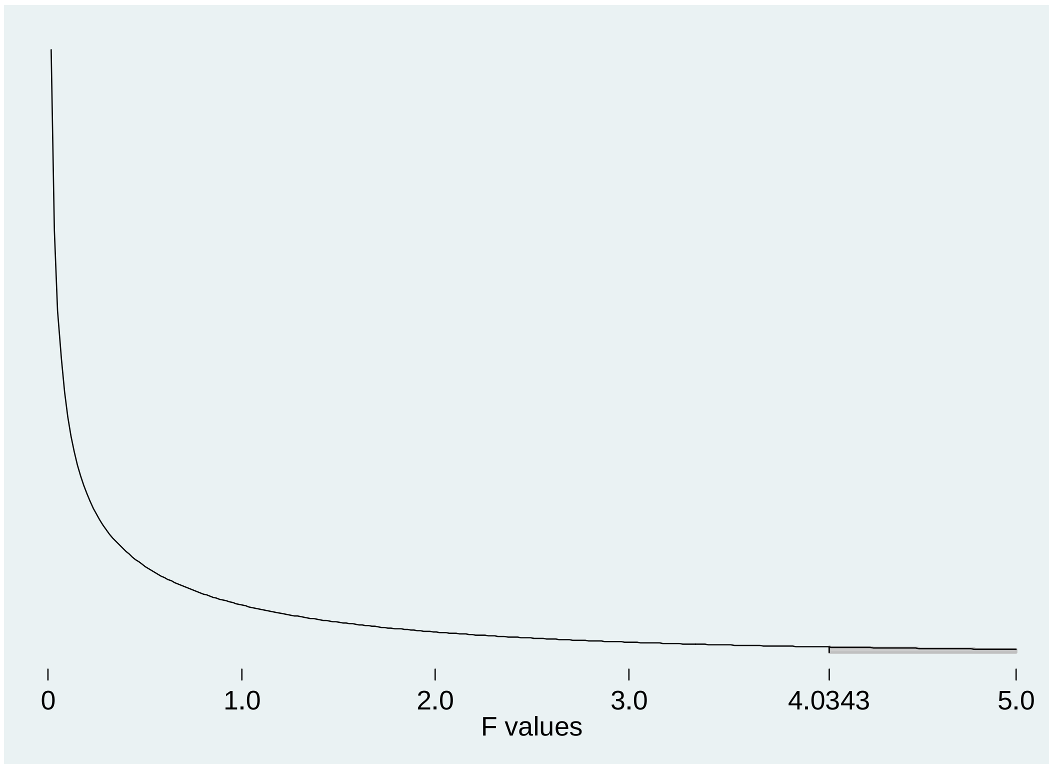
```
. display invFtail(1, 50, 0.05)
4.0343097

. display Ftail(1, 50, 4.0343)
.05000027
```

***rejection* region is $F > F_\alpha$:** $\Pr(F > F_\alpha \mid H_0 \text{ is true}) = \alpha = 0.05$

***non-rejection* region is $F \le F_\alpha$:** $\Pr(F \le F_\alpha \mid H_0 \text{ is true}) = 1 - \alpha = 1 - 0.05 = 0.95$

F test: rejection and non-rejection regions (continued)

At 5% significance level, $\alpha = 0.05$: 0.05 (5%) critical value of $F[1,50] = F_{0.05}[1,50] = 4.0343$



$F_{\alpha}[1,50]$

non-rejection region | rejection region

area = $1 - \alpha = 0.95$ area = $\alpha = 0.05$

❑ ***One-Tail*** **Tests of *One* Restriction on *One* Coefficient**

- ***Examples:*** H_0: $\beta_j = b_j$ (or $\beta_j \ge b_j$) versus H_1: $\beta_j < b_j$ a ***left-tail* test**
  H_0: $\beta_j = b_j$ (or $\beta_j \le b_j$) versus H_1: $\beta_j > b_j$ a ***right-tail* test**

- ***Use:*** a ***one-tail* t-test**.

- ***Test Statistic:***

  ⇨ $t(\hat{\beta}_j) = \dfrac{\hat{\beta}_j - \beta_j}{\hat{se}(\hat{\beta}_j)} \sim t[N-K]$ $\Rightarrow$ ***sample value*** $= t_0(\hat{\beta}_j) = \dfrac{\hat{\beta}_j - b_j}{\hat{se}(\hat{\beta}_j)}$.

- ***Decision Rules -- left-tail t-test:***

  Reject* H₀** if $t_0 < -t_\alpha[N-K]$ *or* ***left-tail p-value for $t_0 = \Pr(t < t_0 \mid H_0 \text{ is true}) < \alpha$;

  Retain* H₀** if $t_0 \ge -t_\alpha[N-K]$ *or* ***left-tail p-value for $t_0 = \Pr(t < t_0 \mid H_0 \text{ is true}) \ge \alpha$.

- ***Decision Rules -- right-tail t-test:***

  Reject* H₀** if $t_0 > t_\alpha[N-K]$ *or* ***right-tail p-value for $t_0 = \Pr(t > t_0 \mid H_0 \text{ is true}) < \alpha$;

  Retain* H₀** if $t_0 \le t_\alpha[N-K]$ *or* ***right-tail p-value for $t_0 = \Pr(t > t_0 \mid H_0 \text{ is true}) \ge \alpha$.

<u>Left-tail t test</u>: rejection and non-rejection regions

Left-tail critical value of t[50] distribution at 5% significance level (α = 0.05)

- ***lower*** **(left-tail) 0.05 critical value of t[50] = $-t_{\alpha}$[50] = $-t_{0.05}$[50] = −1.6759**

Stata **commands to compute *lower* 0.05 critical value of t[50] = $-t_{0.05}$[50]**

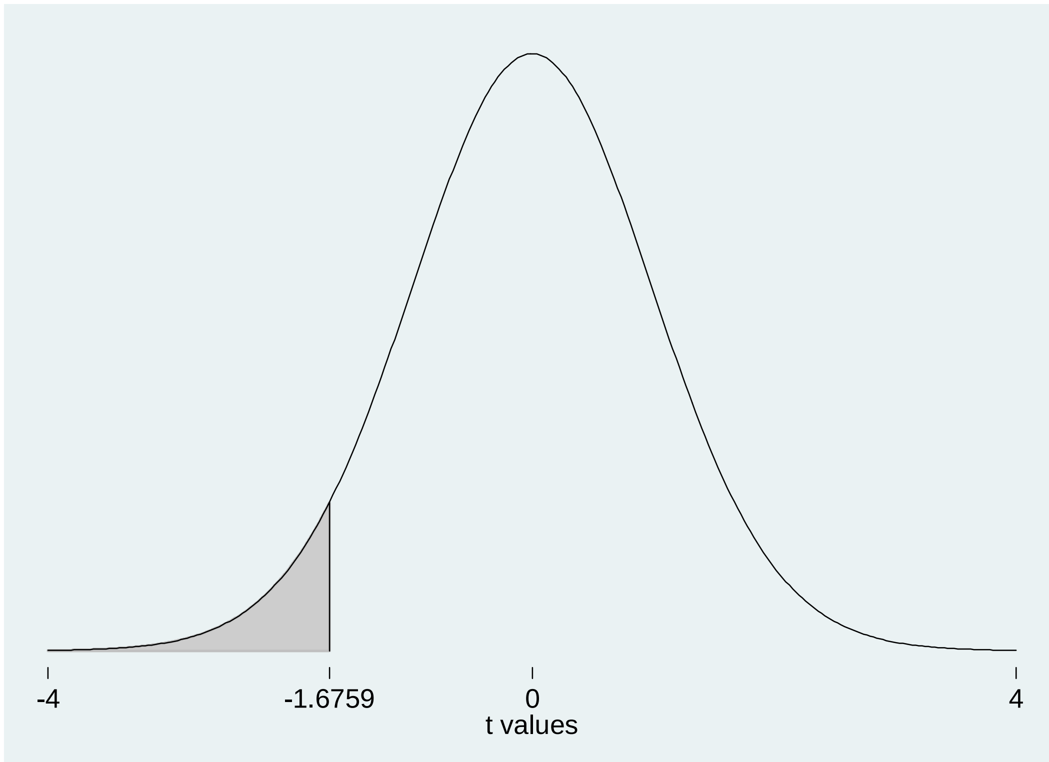
```
. display -1*invttail(50, 0.05)
-1.675905

. display ttail(50, -1.6759)
.9499995

. display 1 - ttail(50, -1.6759)
.0500005
```

rejection **region is $\mathbf{t < -t_{\alpha}}$:** $\Pr(t < -t_{\alpha} \mid H_0 \text{ is true}) = \alpha = 0.05$

non-rejection **region is $\mathbf{t \geq -t_{\alpha}}$:** $\Pr(t \geq -t_{\alpha} \mid H_0 \text{ is true}) = 1 - \alpha = 1 - 0.05 = 0.95$

Left-tail t test: rejection and non-rejection regions (continued)

At 5% significance level, $\alpha = 0.05$: $-t_{0.05}[50] = -1.6759$



$-t_{\alpha}[50]$

rejection region | non-rejection region

area = α = 0.05 area = $1 - \alpha$ = 0.95

<u>Right-tail t test</u>: rejection and non-rejection regions

Right-tail critical value of t[50] distribution at 5% significance level ($\alpha = 0.05$)

- ***upper*** **(right-tail) 0.05 critical value of t[50] = $t_{\alpha}[50] = t_{0.05}[50]$ = 1.6759**

  Stata **commands to compute** ***upper*** **0.05 critical value of t[50] = $t_{0.05}[50]$**

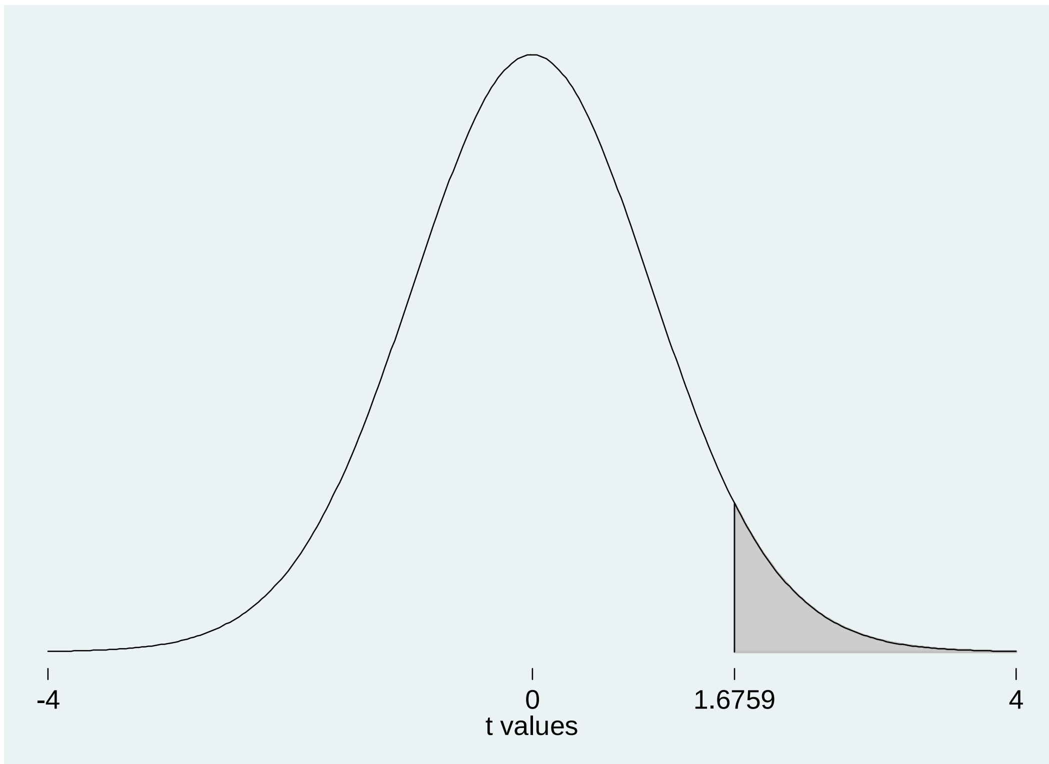
```
. display invttail(50, 0.05)
1.675905

. display ttail(50, 1.6759)
.0500005
```

rejection **region is $t > t_{\alpha}$:** $\Pr(t > t_{\alpha} \mid H_0 \text{ is true}) = \alpha = 0.05$

non-rejection **region is $t \le t_{\alpha}$:** $\Pr(t \le t_{\alpha} \mid H_0 \text{ is true}) = 1 - \alpha = 1 - 0.05 = 0.95$

<u>Right-tail t test</u>: rejection and non-rejection regions (continued)

At 5% significance level, $\alpha = 0.05$: $t_{0.05}[50] = 1.6759$



$t_{\alpha}[50]$

non-rejection region | rejection region

area = $1 - \alpha = 0.95$ area = $\alpha = 0.05$

2. Tests of *One* Linear Restriction on *Two or More* Coefficients

H_0 specifies ***only one*** linear restriction on ***two or more*** regression coefficients.

❑ *Two-Tail* Tests of *One* Linear Restriction on *Two* Coefficients

- ***Example:*** H_0: $c_j\beta_j + c_h\beta_h = c_0$ versus H_1: $c_j\beta_j + c_h\beta_h \neq c_0$.
- ***Use:*** either a ***two-tail*** **t-test** or an **F-test**.
- ***Test Statistics:***

t-statistic

$$\Rightarrow \quad t\left(c_j\hat{\beta}_j + c_h\hat{\beta}_h\right) = \frac{(c_j\hat{\beta}_j + c_h\hat{\beta}_h) - (c_j\beta_j + c_h\beta_h)}{\hat{se}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)} \sim t[N-K] = t[N-K_1]$$

where: $\hat{se}(c_j\hat{\beta}_j + c_h\hat{\beta}_h) = \sqrt{\hat{Var}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)}$

$$\hat{Var}(c_j\hat{\beta}_j + c_h\hat{\beta}_h) = c_j^2\hat{Var}(\hat{\beta}_j) + c_h^2\hat{Var}(\hat{\beta}_h) + 2c_jc_h\hat{Cov}(\hat{\beta}_j, \hat{\beta}_h).$$

sample value $= t_0\left(c_j\hat{\beta}_j + c_h\hat{\beta}_h\right) = \dfrac{(c_j\hat{\beta}_j + c_h\hat{\beta}_h) - c_0}{\hat{se}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)}$.

❑ ***Two-Tail* Tests of *One* Linear Restriction on *Two* Coefficients (continued)**

F-statistic

$$\Rightarrow \; F\left(c_j\hat{\beta}_j + c_h\hat{\beta}_h\right) = \frac{\left[(c_j\hat{\beta}_j + c_h\hat{\beta}_h) - (c_j\beta_j + c_h\beta_h)\right]^2}{\hat{Var}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)} \sim F[1, N-K] = F[1, N-K_1]$$

where: $\hat{Var}(c_j\hat{\beta}_j + c_h\hat{\beta}_h) = c_j^2\hat{Var}(\hat{\beta}_j) + c_h^2\hat{Var}(\hat{\beta}_h) + 2c_jc_h\hat{Cov}(\hat{\beta}_j, \hat{\beta}_h)$.

sample value $= F_0\left(c_j\hat{\beta}_j + c_h\hat{\beta}_h\right) = \dfrac{\left[(c_j\hat{\beta}_j + c_h\hat{\beta}_h) - c_0\right]^2}{\hat{Var}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)}$

- ***Decision Rules – Two-Tail Tests of One Coefficient Restriction on Two Coefficients:***

Reject* H₀** if $|t_0| > t_{\alpha/2}[N-K]$ *or* ***two-tail p-value for $t_0 = \Pr\left(|t| > |t_0| \,\middle|\, H_0 \text{ is true}\right) < \alpha$;

$F_0 > F_\alpha[1, N-K]$ *or* ***p-value*** for $F_0 = \Pr(F > F_0 \mid H_0 \text{ is true}) < \alpha$.

Retain* H₀** if $|t_0| \le t_{\alpha/2}[N-K]$ *or* ***two-tail p-value for $t_0 = \Pr\left(|t| > |t_0| \,\middle|\, H_0 \text{ is true}\right) \ge \alpha$;

$F_0 \le F_\alpha[1, N-K]$ *or* ***p-value*** for $F_0 = \Pr(F > F_0 \mid H_0 \text{ is true}) \ge \alpha$.

♦ Example of Tests of *One* Linear Restriction on *Two* Coefficients: *Stata* Commands

- Consider the ***linear regression model*** given by the population regression equation, or PRE, (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \qquad (1)$$

- Estimate the PRE (1) by **OLS (Ordinary Least Squares)** using the following **Stata *regress*** command:

```
regress y x1 x2 x3 x4 x5
```

- Perform **individual two-tail t-test** of H_0: $\beta_2 = \beta_4$ versus H_1: $\beta_2 \neq \beta_4$ using the **Stata *lincom*** command:

```
lincom x2 - x4
```

or

```
lincom _b[x2] - _b[x4]
```

- Perform **individual F-test** of H_0: $\beta_2 = \beta_4$ versus H_1: $\beta_2 \neq \beta_4$ using the **Stata *test*** command:

```
test x2 = x4
```

or

```
test x2 - x4 = 0
```

<u>Example of Tests of *One* Linear Restriction on *Two* Coefficients: *Stata* Commands</u> (continued)

- How is $\mathbf{t_0}$ **statistic** produced by ***lincom*** command related to $\mathbf{F_0}$ **statistic** produced by ***test*** command?

  lincom command computes:

$$t_0\left(\hat{\beta}_2 - \hat{\beta}_4\right) = \frac{(\hat{\beta}_2 - \hat{\beta}_4) - 0}{\hat{se}(\hat{\beta}_2 - \hat{\beta}_4)} = \frac{(\hat{\beta}_2 - \hat{\beta}_4)}{\hat{se}(\hat{\beta}_2 - \hat{\beta}_4)} \sim t[N-6] \text{ under } H_0$$

  test command computes:

$$F_0\left(\hat{\beta}_2 - \hat{\beta}_4\right) = \frac{\left[(\hat{\beta}_2 - \hat{\beta}_4) - 0\right]^2}{\hat{Var}(\hat{\beta}_2 - \hat{\beta}_4)} = \frac{(\hat{\beta}_2 - \hat{\beta}_4)^2}{\hat{Var}(\hat{\beta}_2 - \hat{\beta}_4)} \sim F[1, N-6] \text{ under } H_0$$

- **Relationship between t and F tests**: they **yield *identical* inferences** because

$$F_0\left(\hat{\beta}_2 - \hat{\beta}_4\right) = \left(t_0(\hat{\beta}_2 - \hat{\beta}_4)\right)^2 \qquad \text{and} \qquad F_\alpha[1, N-6] = \left(t_{\alpha/2}[N-6]\right)^2$$

  p-value for $\mathbf{F_0}$ = ***two-tail p-value*** for $\mathbf{t_0}$

❑ ***One-Tail*** **Tests of** ***One*** **Linear Restriction on** ***Two*** **Coefficients**

- ***Examples:*** H_0: $c_j\beta_j + c_h\beta_h = c_0$ versus H_1: $c_j\beta_j + c_h\beta_h < c_0$ a ***left-tail*** **test**
  H_0: $c_j\beta_j + c_h\beta_h = c_0$ versus H_1: $c_j\beta_j + c_h\beta_h > c_0$ a ***right-tail*** **test**

- ***Use:*** a ***one-tail*** **t-test**.

- ***Test Statistic:***

$$\Rightarrow \quad t\left(c_j\hat{\beta}_j + c_h\hat{\beta}_h\right) = \frac{(c_j\hat{\beta}_j + c_h\hat{\beta}_h) - (c_j\beta_j + c_h\beta_h)}{\hat{se}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)} \sim t[N-K] = t[N-K_1]$$

where $\hat{se}(c_j\hat{\beta}_j + c_h\hat{\beta}_h) = \sqrt{\hat{Var}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)}$

$$\hat{Var}(c_j\hat{\beta}_j + c_h\hat{\beta}_h) = c_j^2\hat{Var}(\hat{\beta}_j) + c_h^2\hat{Var}(\hat{\beta}_h) + 2c_jc_h\hat{Cov}(\hat{\beta}_j, \hat{\beta}_h).$$

sample value $= t_0\left(c_j\hat{\beta}_j + c_h\hat{\beta}_h\right) = \dfrac{(c_j\hat{\beta}_j + c_h\hat{\beta}_h) - c_0}{\hat{se}(c_j\hat{\beta}_j + c_h\hat{\beta}_h)}.$

❑ ***One-Tail* Tests of *One* Linear Restriction on *Two* Coefficients** (continued)

- ***Decision Rules -- left-tail t-test:***

Reject* H₀** if $t_0 < -t_\alpha[N-K]$ *or* ***left-tail p-value for $t_0 = \Pr(\, t < t_0 \mid H_0 \text{ is true}) < \alpha$;

Retain* H₀** if $t_0 \geq -t_\alpha[N-K]$ *or* ***left-tail p-value for $t_0 = \Pr(\, t < t_0 \mid H_0 \text{ is true}) \geq \alpha$.

- ***Decision Rules -- right-tail t-test:***

Reject* H₀** if $t_0 > t_\alpha[N-K]$ *or* ***right-tail p-value for $t_0 = \Pr(\, t > t_0 \mid H_0 \text{ is true}) < \alpha$;

Retain* H₀** if $t_0 \leq t_\alpha[N-K]$ *or* ***right-tail p-value for $t_0 = \Pr(\, t > t_0 \mid H_0 \text{ is true}) \geq \alpha$.

3. ***Tests* of *Two or More* Linear Coefficient Restrictions**

H_0 specifies ***two or more* linear coefficient restrictions** on the PRE given by (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \qquad (1)$$

We consider two different sets of two or more coefficient restrictions on PRE (1).

- ***Example 1:*** H_0: $\beta_2 = \beta_4$ *and* $\beta_3 = \beta_5$ (number of restrictions = 2)
  H_1: $\beta_2 \neq \beta_4$ *and/or* $\beta_3 \neq \beta_5$

- ***Example 2:*** H_0: $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$ (number of restrictions = 3)
  H_1: $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_5 \neq 0$

- ***Use:*** a **general F-test**; only an F-test can be used to **test jointly *two or more*** coefficient restrictions.

❑ ***Tests* of *Two or More* Linear Coefficient Restrictions (continued)**

- ***Test Statistics:*** Either of the following two forms of the **general F-statistic**.

⇨ $$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} = \frac{(RSS_0 - RSS_1)/(K - K_0)}{RSS_1/(N - K)}.$$

⇨ $$F = \frac{(R_U^2 - R_R^2)/(df_0 - df_1)}{(1 - R_U^2)/df_1} = \frac{(R_U^2 - R_R^2)/(K - K_0)}{(1 - R_U^2)/(N - K)}.$$

RSS_0 = the ***restricted*** **OLS residual sum of squares** for the ***restricted*** **model** implied by $\mathbf{H_0}$;

RSS_1 = the ***unrestricted*** **OLS residual sum of squares** for the ***unrestricted*** **model** implied by $\mathbf{H_1}$;

$df_0 = N - K_0$ = df for RSS_0; $df_1 = N - K$ = df for RSS_1;

R_R^2 = the R^2 value for the ***restricted*** **model** implied by $\mathbf{H_0}$;

R_U^2 = the R^2 value for the ***unrestricted*** **model** implied by $\mathbf{H_1}$.

- ***Null distribution:*** $F \sim F[df_0 - df_1, df_1] = F[K - K_0, N - K]$.

- ***Sample value*** of F-statistic under $H_0 = F_0$.

❑ ***Tests* of *Two or More* Linear Coefficient Restrictions** (continued)

- ***Decision Rules:***

Reject $\mathbf{H_0}$ if $F_0 > F_{\alpha}[df_0 - df_1, df_1] = F_{\alpha}[K - K_0, N - K]$ *or*

p-value for $F_0 = \Pr(F > F_0 \mid H_0 \text{ is true}) < \alpha$.

Retain $\mathbf{H_0}$ if $F_0 \le F_{\alpha}[df_0 - df_1, df_1] = F_{\alpha}[K - K_0, N - K]$ *or*

p-value for $F_0 = \Pr(F > F_0 \mid H_0 \text{ is true}) \ge \alpha$.

❑ ***Tests* of *Two* Linear Coefficient Restrictions: Example 1**

- ***Example 1:*** H_0: $\beta_2 = \beta_4$ *and* $\beta_3 = \beta_5$ (number of restrictions = 2)
  H_1: $\beta_2 \neq \beta_4$ *and/or* $\beta_3 \neq \beta_5$

- ***Unrestricted model*** is given by PRE (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \quad (1)$$

- ***Restricted model*** is given by PRE (2): set $\beta_4 = \beta_2$ *and* $\beta_5 = \beta_3$ in PRE (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \quad (1)$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_2 X_{i4} + \beta_3 X_{i5} + u_i$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 (X_{i2} + X_{i4}) + \beta_3 (X_{i3} + X_{i5}) + u_i \quad (2)$$

❑ ***Tests* of *Two* Linear Coefficient Restrictions: Example 1** **(continued)**

- OLS estimation of (1) yields the ***unrestricted SRE*** (1*):

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \hat{\beta}_3 X_{i3} + \hat{\beta}_4 X_{i4} + \hat{\beta}_5 X_{i5} + \hat{u}_i \qquad (1^*)$$

$$RSS_1 = \sum_{i=1}^{N} \hat{u}_i^2 = \hat{u}^T\hat{u} \quad \text{with} \quad df_1 = N - K = N - 6 \quad \text{since } K = 6..$$

- OLS estimation of (2) yields the ***restricted SRE*** (2*):

$$Y_i = \tilde{\beta}_0 + \tilde{\beta}_1 X_{i1} + \tilde{\beta}_2 (X_{i2} + X_{i4}) + \tilde{\beta}_3 (X_{i3} + X_{i5}) + \tilde{u}_i \qquad (2^*)$$

$$\tilde{\beta}_4 = \tilde{\beta}_2 \quad \text{and} \quad \tilde{\beta}_5 = \tilde{\beta}_3$$

$$RSS_0 = \sum_{i=1}^{N} \tilde{u}_i^2 = \tilde{u}^T\tilde{u} \quad \text{with} \quad df_0 = N - K_0 = N - 4 \quad \text{since } K_0 = 4.$$

- Substitute values of RSS_1, RSS_0, df_1 and df_0 into formula for general F-statistic:

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} = \frac{(RSS_0 - RSS_1)/(K - K_0)}{RSS_1/(N-K)}$$

- **Null distribution of F:** $F \sim F[df_0 - df_1, df_1] = F[K - K_0, N - K] = F[2, N - K]$.

- Apply decision rule.

❑ ***Tests* of *Two* Linear Coefficient Restrictions: Example 1**

- ***Example 1:*** H_0: $\beta_2 = \beta_4$ *and* $\beta_3 = \beta_5$ (number of restrictions = 2)
  H_1: $\beta_2 \neq \beta_4$ *and/or* $\beta_3 \neq \beta_5$

F test: rejection and non-rejection regions

Critical value of F[2,1000] at 5% significance level (α = 0.05) = F_α[2,1000]
= $F_{0.05}$[2,1000] = 3.0047

- **0.05 critical value of F[2,1000] = F_α[2, 1000] = $F_{0.05}$[2, 1000] = 3.0047**

  ***Stata* commands to compute 0.05 critical value of F[2,1000] = $F_{0.05}$[2,1000]**

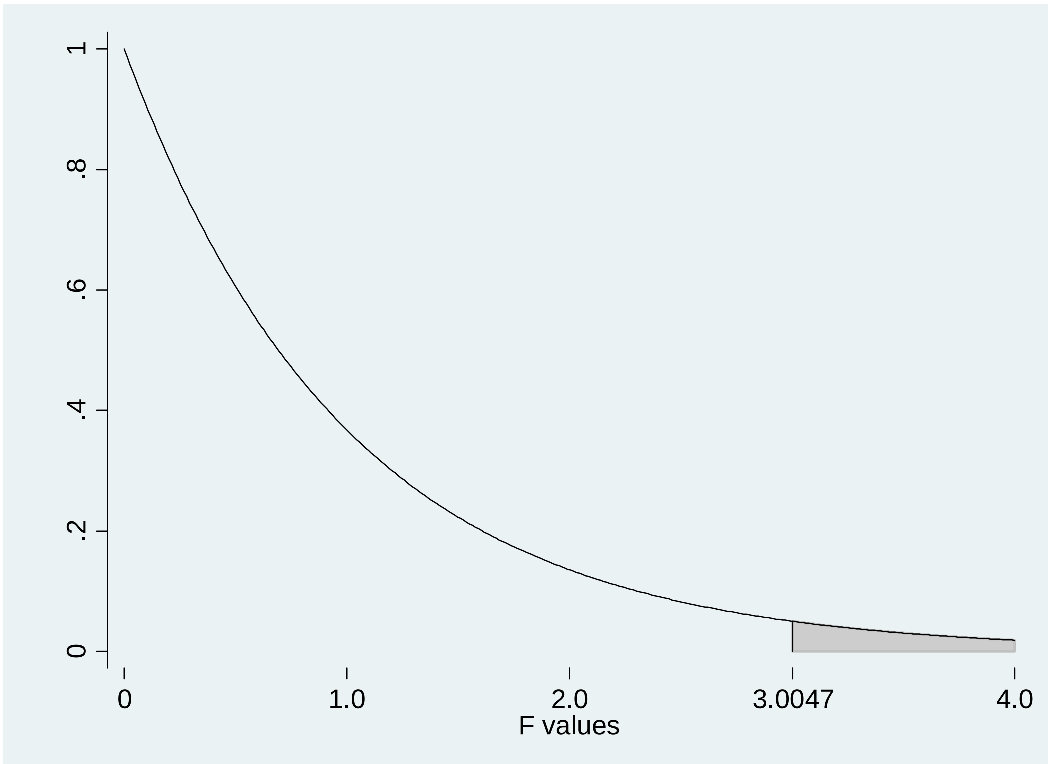
```
. display invFtail(2, 1000, 0.05)
3.0047246

. display Ftail(2, 1000, 3.0047)
.05000122
```

***rejection* region is $F > F_\alpha$:** $\Pr(F > F_\alpha \mid H_0 \text{ is true}) = \alpha = 0.05$

***non-rejection* region is $F \leq F_\alpha$:** $\Pr(F \leq F_\alpha \mid H_0 \text{ is true}) = 1 - \alpha = 1 - 0.05 = 0.95$

F test: rejection and non-rejection regions: Example 1 (continued)

At 5% significance level, $\alpha = 0.05$: 0.05 (5%) critical value of F[2,1000] = $F_{0.05}[2,1000]$ = 3.0047



$F_{\alpha}[2,1000]$

non-rejection region | rejection region

area = $1 - \alpha = 0.95$ area = $\alpha = 0.05$

♦ Example 1 -- Tests of *Two* Linear Coefficient Restrictions: *Stata* Commands

- ***Example 1:*** H_0: $\beta_2 = \beta_4$ *and* $\beta_3 = \beta_5$ (number of restrictions = 2)
  H_1: $\beta_2 \neq \beta_4$ *and/or* $\beta_3 \neq \beta_5$

- ***Unrestricted model*** is given by PRE (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \qquad (1)$$

- Estimate ***unrestricted model*** given by PRE (1) by OLS using the **Stata *regress*** command:

```
regress y x1 x2 x3 x4 x5
```

- Perform **joint F-test** of H_0 versus H_1 using the following **Stata *test*** commands:

```
test x2 = x4, notest
test x3 = x5, accumulate
```

or

```
test x2 - x4 = 0, notest
test x3 - x5 = 0, accumulate
```

❑ ***Tests* of *Three* Linear Coefficient Restrictions: Example 2**

- ***Example 2:*** H_0: $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$ (number of restrictions = 3)
  H_1: $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_5 \neq 0$

- ***Unrestricted model*** is given by PRE (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \qquad (1)$$

- ***Restricted model*** is given by PRE (3): set $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$ in PRE (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \qquad (1)$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + 0X_{i3} + 0X_{i4} + 0X_{i5} + u_i$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i \qquad (3)$$

❑ ***Tests* of *Three* Linear Coefficient Restrictions: Example 2 (continued)**

- OLS estimation of (1) yields the ***unrestricted SRE*** (1*):

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \hat{\beta}_3 X_{i3} + \hat{\beta}_4 X_{i4} + \hat{\beta}_5 X_{i5} + \hat{u}_i \qquad (1^*)$$

$$RSS_1 = \sum_{i=1}^{N} \hat{u}_i^2 = \hat{u}^T \hat{u} \quad \text{with} \quad df_1 = N - K = N - 6 \quad \text{since } K = 6.$$

- OLS estimation of (3) yields the ***restricted SRE*** (3*):

$$Y_i = \tilde{\beta}_0 + \tilde{\beta}_1 X_{i1} + \tilde{\beta}_2 X_{i2} + \tilde{u}_i \qquad (3^*)$$

$$\tilde{\beta}_3 = 0 \quad \text{and} \quad \tilde{\beta}_4 = 0 \quad \text{and} \quad \tilde{\beta}_5 = 0$$

$$RSS_0 = \sum_{i=1}^{N} \tilde{u}_i^2 = \tilde{u}^T \tilde{u} \quad \text{with} \quad df_0 = N - K_0 = N - 3 \quad \text{since } K_0 = 3.$$

- Substitute values of RSS_1, RSS_0, df_1 and df_0 into formula for general F-statistic:

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} = \frac{(RSS_0 - RSS_1)/(K - K_0)}{RSS_1/(N - K)}$$

- **Null distribution of F:** $F \sim F[df_0 - df_1, df_1] = F[K - K_0, N - K] = F[3, N - K]$.

- Apply decision rule.

❑ ***Tests* of *Three* Linear Coefficient Restrictions: Example 2**

- ***Example 2:*** H_0: $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$ (number of restrictions = 3)
  H_1: $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_5 \neq 0$

F test: rejection and non-rejection regions

Critical value of F[3,1000] at 5% significance level ($\alpha = 0.05$) = $F_{\alpha}[3,1000]$
= $F_{0.05}[3,1000]$ = 2.6138

- **0.05 critical value of F[3,1000] = $F_{\alpha}[3, 1000]$ = $F_{0.05}[3, 1000]$ = 2.6138**

  ***Stata* commands to compute 0.05 critical value of F[3,1000] = $F_{0.05}[3,1000]$**

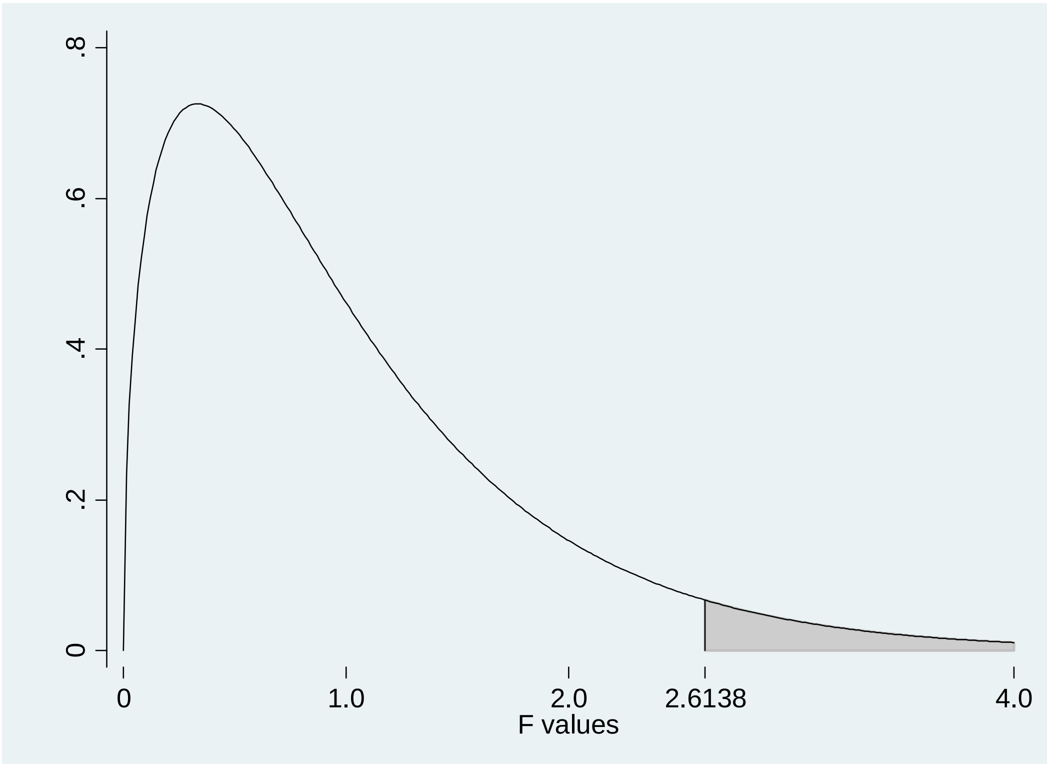
```
. display invFtail(3, 1000, 0.05)
2.6138036

. display Ftail(3, 1000, 2.6138)
.05000024
```

***rejection* region is $\mathbf{F > F_{\alpha}}$:** $\Pr(F > F_{\alpha} \mid H_0 \text{ is true}) = \alpha = 0.05$

***non-rejection* region is $\mathbf{F \leq F_{\alpha}}$:** $\Pr(F \leq F_{\alpha} \mid H_0 \text{ is true}) = 1 - \alpha = 1 - 0.05 = 0.95$

<u>F test</u>: rejection and non-rejection regions: Example 2 (continued)

At 5% significance level, α = 0.05: 0.05 (5%) critical value of F[3,1000] = $F_{0.05}$[3,1000] = 2.6138



F_{α}[3,1000]

non-rejection region | rejection region

area = 1 – α = 0.95 area = α = 0.05

♦ **Example 2 -- Tests of *Three* Linear Coefficient Restrictions: *Stata* Commands**

- ***Example 2:*** H_0: $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$ (number of restrictions = 3)
  H_1: $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_5 \neq 0$

- ***Unrestricted model*** is given by PRE (1):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + u_i \qquad (1)$$

- Estimate ***unrestricted model*** given by PRE (1) by OLS using the following **Stata *regress*** command:

```
regress y x1 x2 x3 x4 x5
```

- Perform **joint F-test** of H_0 versus H_1 using the following **Stata *test*** command(s):

```
test x3 x4 x5
```

or (the long way)

```
test x3 = 0, notest
test x4 = 0, accumulate notest
test x5 = 0, accumulate
```