Post-Midterm 1
Regression Review

Brooks (4th edition): Chapters 3, 4 & 5

1

Review: CLM & OLS

- *Classical linear regression model* (CLM) - Assumptions:
  - (**A1**) DGP: $\boldsymbol{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ is correctly specified (& linear!).
  - (**A2**) $E[\boldsymbol{\varepsilon}|\mathbf{X}] = 0$
  - (**A3**) $Var[\boldsymbol{\varepsilon}|\mathbf{X}] = \sigma^2\,\mathbf{I}_T$
  - (**A4**) $\mathbf{X}$ has full column rank –rank($\mathbf{X}$) = k, where $T \geq k$.

Objective function: $S(\boldsymbol{x};\boldsymbol{\beta}) = \sum_{i=1}^{T}\varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta})'\,(\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta})$

$\Rightarrow \mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\,\boldsymbol{y}$ (kx1) vector

- $\mathbf{b}$ is an estimate of the marginal effect (first derivative) on (**A1**).

Review: Properties of OLS b

- $\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'}\,\boldsymbol{y}$ $\Rightarrow$ **b** is a (linear) function of the data.

- Under the typical assumptions, we can establish properties for **b**.

1) $E[\mathbf{b}|\mathbf{X}] = \boldsymbol{\beta}$ –**b** is *unbiased.* (**b** is a *k*x1 matrix)

2) $\text{Var}[\mathbf{b}|\mathbf{X}] = \sigma^2 (\mathbf{X'X})^{-1}$ (a kxk matrix)

3) **Gauss-Markov Theorem**: **b** is BLUE (*Best Linear Unbiased Estimator*). No other linear & unbiased estimator has a lower variance.

4) If (**A5**) $\boldsymbol{\varepsilon}|\mathbf{X} \sim$ *i.i.d.* $N(\mathbf{0}, \sigma^2\mathbf{I}_T)$ $\Rightarrow \mathbf{b}|\mathbf{X} \sim$ *i.i.d.* $N(\boldsymbol{\beta}, \sigma^2 (\mathbf{X'}\,\mathbf{X})^{-1})$

$b_k|\mathbf{X} \sim N(\beta_k, v_k^2)$

$SD[b_k|\mathbf{X}] = \sqrt{\{[\sigma^2(\mathbf{X'X})^{-1}]_{kk}\}}$

Note: We use the distribution of $\mathbf{b}|\mathbf{X}$ to derive the distribution of tests (t, F, and Wald) to draw inferences.

Review: Properties of OLS b

5) If (**A5**) is not assumed, we still can obtain a (limiting) distribution for **b**. Under additional assumptions –mainly, the matrix **X'X** does not explode as *T* becomes large–, as $T \rightarrow \infty$,

(i) $\mathbf{b} \xrightarrow{p} \boldsymbol{\beta}$ (**b** is consistent)

(ii) $\mathbf{b} \xrightarrow{a} N(\boldsymbol{\beta}, \sigma^2 (\mathbf{X'}\,\mathbf{X})^{-1})$ (**b** is asymptotically normal)

- Properties (1)-(4) are called *finite* (or *small)* sample properties.

- Properties (5.i) and (5.ii) are called *asymptotic* properties, they only hold when *T* is large (actually, as *T* tends to ∞). We use (5.ii) to draw inferences.

Note: If not sure about the applicability of the *asymptotic* distribution, use bootstrap to draw inferences.

Review: Fitted Values, Residuals & s^2

- OLS estimates $\boldsymbol{\beta}$ with **b.** Now, we define *fitted values* as:

  $\hat{\boldsymbol{y}} = \mathbf{X}\,\mathbf{b}$

Now we define the estimated error, **e** (also called *residuals*):

$\boldsymbol{e} = \boldsymbol{y} - \hat{\boldsymbol{y}}$

It can be shown that **e** is uncorrelated with **X** $\Rightarrow \boldsymbol{e} \perp \mathbf{X}$

- Using $\boldsymbol{e}$, we define a measure of unexplained variation:

  Residual Sum of Squares (RSS) = $\boldsymbol{e}'\boldsymbol{e} = \sum_i e_i^2$

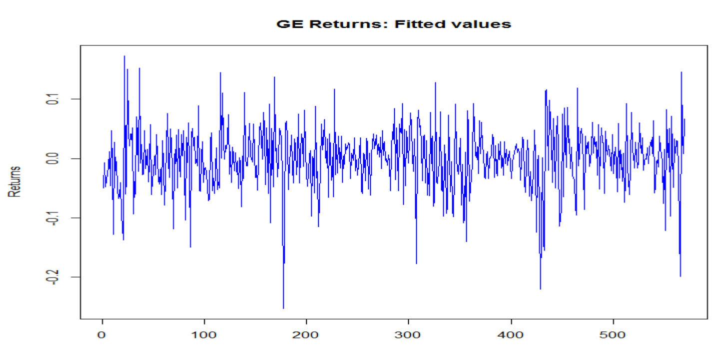
- We use RSS to calculate s^2, the unbiased estimator of σ^2:

  s^2 = RSS/ $/(T-k) = \sum_i e_i^2\,/(T-k)] = \boldsymbol{e}'\boldsymbol{e}/(T-k)$

- Then, the estimator of Var[$\mathbf{b}$|$\mathbf{X}$] = $s^2\,(\mathbf{X'X})^{-1}$

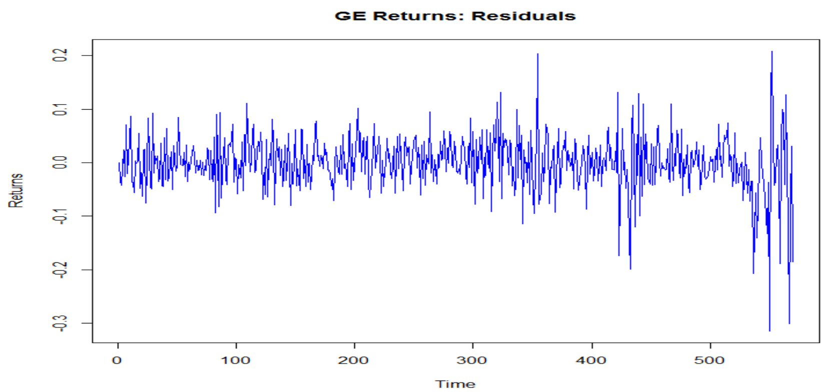
Review: Fitted Values, Residuals & s^2

```
fit_ge <- lm(ge_x ~ Mkt_RF + SMB + HML)
y_ge <- fit_ge$fitted                                    # Extract fitted values from lm
plot(y_ge, type = "l", col = "blue",                     # Plot GE fitted value returns
main = « GE Returns: Fitted Values", ylab = "Returns", xlab = "Time")
```



Review: Fitted Values, Residuals & s^2

```
fit_ge <- lm(ge_x ~ Mkt_RF + SMB + HML)
e_ge <- fit_ge$residuals                          # Extract residuals from lm
plot(e_ge, type = "l", col = "blue",              # Plot GE residual returns
main = « GE Returns: Residuals", ylab = "Returns", xlab = "Time")
```



Review: Goodness of Fit – R^2 & Adjusted R^2

• We use RSS to measure how much the model explains the variation of y. We define variation of y as TSS:

$$\text{TSS} = \sum_i (y_i - \bar{y})^2$$

• Decomposition of total variation (assume $\mathbf{X}_1 = \mathbf{\imath}$ –a constant.)

TSS = SSR + RSS (SSR: Regression Sum of Squares)

• R-squared (R^2)

R^2 = SSR/TSS = Regression variation/Total variation

R^2 = 1 – RSS/ TSS

With a constant in the model, R^2 lies between 0 and 1. It measures how much of total variation of y is explained by the regression (SSR).

Review: Goodness of Fit – R^2 & Adjusted R^2

• Main problem with R^2: R^2 never falls when regressors (say **z**) are added to the regression. This occurs because RSS decreases with more information.

Solution: Incorporate a penalty for number of parameters in R^2. This is what *Adjusted*-R^2 does:

$$\overline{R}^2 = 1 - \frac{(T-1)}{(T-k)}(1 - R^2) = 1 - \frac{s^2}{TSS/(T-1)}$$

$\Rightarrow$ maximizing $\overline{R}^2$ <=> minimizing $[RSS/(T-k)] = s^2$

We can use $\overline{R}^2$ to compare models. There are other popular goodness of fit measures with penalties for number of parameters: AIC & BIC.

Review: Testing Only One Parameter

• We are interested in testing a hypothesis about one parameter in our linear model: $\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

1. Set H_0 and H_1 (about only one parameter): H_0: $\beta_k = \beta_k^0$
   H_1: $\beta_k \neq \beta_k^0$
2. Appropriate $T(X)$: *t-statistic*:

$$t_k = \frac{b_k - \beta_k^0}{s_{b,k}} \sim t_{T-k}.$$

3. Compute t_k, $\hat{t}$, using b_k, β_k^0, s, and $(\mathbf{X'X})^{-1}$. Get *p-value*($\hat{t}$).

4. Rule: Set an α level. If *p-value*($\hat{t}$) $< \alpha$ $\Rightarrow$ Reject H_0: $\beta_k = \beta_k^0$
   Alternatively, if $|\hat{t}| > t_{T-k,\alpha/2}$ $\Rightarrow$ Reject H_0: $\beta_k = \beta_k^0$

Review: Testing Only One Parameter

- Special case: H_0: $\beta_k = 0$
  H_1: $\beta_k \neq 0$.

Then,

$$t_k = \frac{b_k}{s_{b,k}} \sim t_{T-k}$$

This special case of t_k is called the *t-value* or *t-ratio* (also refer as the "t-stats").

- Usually, α = 5%, then if | $\hat{t}_k$ |> **1.96** ≈ 2, we say the coefficient b_k is "*significant*."

OLS Estimation – Testing the CAPM

Example: We test the CAPM for GE. Recall that the CAPM states:

$$E[r_{i=GE,t} - r_f] = \beta_{i=GE}\, E[(r_{m,t} - r_f)].$$

According to the CAPM, equilibrium excess returns are only determined by excess market returns –i.e., the CAPM is a one factor model. There is no constant or extra factors besides the market.

A linear data generating process (DGP) consistent with the CAPM is:

$$(r_{GE,t} - r_f) = \alpha_{GE} + \beta_{GE}\,(r_{m,t} - r_f) + \varepsilon_{GE,t}, \qquad t = 1, \ldots, T$$

Thus, we test the CAPM by testing H_0 (CAPM holds): $\alpha_{GE} = 0$
H_1 (CAPM rejected): $\alpha_{GE} \neq 0$.

```
SFX_da <-
read.csv("http://www.bauer.uh.edu/rsusmel/4397/Stocks_FX_1973.csv",head=TRUE,sep=",")
x_ge <- SFX_da$GE                   # Extract IBM price data
x_Mkt_RF <- SFX_da$Mkt_RF           # Extract Market excess returns (in %)
x_RF <- SFX_da$RF                   # Extract risk free rate (in %)
```

OLS Estimation – Testing the CAPM

Example (continuation):

```
T <- length(x_ge)                       # Sample size
 lr_ge <- log(x_ge[-1]/x_ge[-T])        #  Log returns for IBM (lost one observation)
 Mkt_RF <- x_Mkt_RF[-1]/100             # Adjust size (take one observation out )
RF <- x_RF[-1]/100
ge_x <- lr_ge – RF                      # Define excess returns for IBM

fit_ge_capm <- lm(ge_x ~ Mkt_RF)        # OLS estimation with lm package in R
> summary(fit_ge_capm)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.007338   0.002275  -3.225  0.00133 **      ⇒ |t0| > 1.96   ⇒ Reject  (one-factor) CAPM.
xMkt_RF    1.129255    0.049291   22.910  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Q: Is the intercept (α_{GE}) equal to 0 (H_0: α_{GE} = 0)? Use the t-value:

$\hat{t}_k = b_k/\text{Est. SE}[b_k] = -3.225$ $\Rightarrow |\hat{t}_0| > 1.96 \Rightarrow$ Reject H_0

OLS Estimation – Testing the CAPM

Example (continuation):

$\Rightarrow |\hat{t}_\alpha| > 1.96$ $\Rightarrow$ Reject H_0 (CAPM) at 5% level

Conclusion: The CAPM is rejected for IBM at the 5% level.

Note: You can also reject H_0 by looking at the *p-value* of intercept.

Interpretation: Given that the intercept is significant (& negative). GE *underperformed* relative to what the CAPM expected:

- GE excess returns: mean(ge_x) = **-0.0009589826**

- GE excess returns (CAPM) = **1.129255** * mean(Mkt_RF)

= **1.129255** * 0.0056489 = **0.006378998**

- Ex-post difference: **-0.000959** - **0.006379** = **-0.007338** ($\approx \alpha_{GE}$)

OLS Estimation – The 3-Factor F-F Model

• The CAPM is routinely rejected. A popular alternative is the empirically derived 3-Factor Fama-French Model (1993) with:

a) *Size* factor (SMB) measured as returns of small (size portfolio) minus returns of big (size portfolio)

b) *Value* factor or book-to-market factor (HML), measured as returns of high (B/M portfolio) minus returns of low (B/M portfolio).

• Then, a linear DGP generating this model is:

$$(r_{i,t} - r_f) = \beta_0 + \beta_1 (r_{m,t} - r_f) + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_t.$$

• Under this model, the main drivers of expected returns are sensitivity to the market, sensitivity to size, and sensitivity to value stocks, as measured by the book-to-market ratio.

OLS Estimation – The 3-Factor F-F Model

• The 3-factor FF model produces expected excess returns:

$$E[r_{i,t} - r_f] = \beta_1 E[r_{m,t} - r_f] + \beta_2 E[SMB_t] + \beta_3 E[HML_t].$$

A significant constant would be evidence against this model: something is missing in the model.

• In 2014, Fama and French added two additional factors to their 3-factor model: RMW & CMA.
- RMW measures the return of the portfolio of most profitable firms ("robust") minus the portfolio least profitable ("weak").
- CMA measures the return of a portfolio of firms that invest conservatively minus a portfolio of firms that invest aggressively.

• Again, the 5-factor FF model produces expected excess returns:

$$E[r_{i,t} - r_f] = \beta_1 E[r_{m,t} - r_f] + \beta_2 E[SMB_t] + \beta_3 E[HML_t] + \beta_4 E[RMW_t] + \beta_5 E[CMA_t]$$

Review: Is GE's Beta equal to 1?

Example: For the 3-Factor Fama-French Model for GE returns we want to test if the 3 F-F factors are significant. The model:

$$(r_{GE,t} - r_f) = \beta_0 + \beta_1 (r_{m,t} - r_f) + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_t.$$

Before testing H_0: $\beta_1 = 1$, we check the adequacy of the model:
- Check R^2 and interpret it
- Goodness of Fit test and interpret it
- Signs of coefficients and interpret them.

Then, we test

H_0: $\beta_1 = 1$

H_1: $\beta_1 \neq 1$.

Review: Is GE's Beta equal to 1?

Example (continuation): using lm function in R

```
fit_ge_ff3 <- lm(ge_x ~ Mkt_RF + SMB + HML) # Regress ge_x against 3 F-F factors
> summary(fit_ge_ff3)

Coefficients:
             Estimate   Std. Error t value Pr(>|t|)
(Intercept) -0.008239   0.002219  -3.712 0.000226 ***   ⇒ |t0|> 1.96   ⇒ Reject 3-factor FF model?
Mkt_RF       1.236430   0.050783  24.348  < 2e-16 ***   ⇒ |t1|> 1.96   ⇒ Mkt_RF significant
SMB         -0.318929   0.075303  -4.235 2.67e-05 ***   ⇒ |t2|> 1.96   ⇒ Mkt_RF significant
HML          0.358122   0.075389   4.750 2.58e-06 ***   ⇒ |t3|> 1.96   ⇒ Mkt_RF significant
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05219 on 565 degrees of freedom
Multiple R-squared:  0.5143, Adjusted R-squared:  0.5117
F-statistic: 199.4 on 3 and 565 DF,  p-value: < 2.2e-16
```

Interpretation of β_1: A 1% increase in Mkt_RF increases GE excess returns by **1.24%**.

Review: Is GE's Beta equal to 1?

Example (continuation): using lm function in R

Interpretation of R^2: The 3 F-F factors explain **51%** of the variability of GE returns.

Interpretation of F-test (Goodness of Fit Test):

F-statistic: **199.4** on 3 and 565 DF, p-value: < **2.2e-16**

$\Rightarrow$ Very low *p-value*. That is, strong rejection of H_0: (No joint significance of 3 F-F factors).

The t-stats point out that the 3 F-F factors are significant drivers of GE excess returns.

Interpretation of constant (α_{GE}): The significant constant signals that something is missing from the model. It constant, α_{GE}, is also negative: GE underperformed relative to the 3-factor F-F model.

Review: Is GE's Beta equal to 1?

Example (continuation):

• Q: Is GE's market beta (β_1) equal to 1? That is,

H_0: $\beta_1 = 1$ vs.

H_1: $\beta_1 \neq 1$

$\Rightarrow \hat{t}_k = (b_k - \beta_k^0)/\text{Est. SE}(b_k)$

$\hat{t}_1 = (\mathbf{1.28643} - \mathbf{1})/\mathbf{0.050783} = \mathbf{4.655733}$

Decision Rule:

$|\hat{t}_1 = \mathbf{4.6557}| > \mathbf{1.96}$ $\Rightarrow$ Reject H_0: $\beta_1 = 1$ at 5% level.

Conclusion: GE systematic market risk is greater than the market.

Note: $\hat{t}_1$ can be calculated using *summary(fit_ge)$coef*, which gets the whole lm matrix.

```
> t_b_1 <- (summary(fit_ge_ff3)$coef[2,1] - 1)/summary(fit_ge)$coef[2,2]
> t_b_1
[1] 4.655733
```

Review: Is GE's Beta equal to 1?

Example (continuation):

• 95% CI for GE's market beta (β_k):

[$\mathbf{b}_k - t_{T-k,\alpha/2}$ * **Estimated SE($\mathbf{b}_k$)**, $\mathbf{b}_k + t_{T-k,\alpha/2}$ * **Estimated SE($\mathbf{b}_k$)**]

$\Rightarrow$ [**1.28643** – **1.96** * **0.050783** , **1.28643** + **1.96** * **0.050783**] =

$\beta_1 \in$ [**1.186895, 1.385965**] with 95% confidence

Clearly, $\beta_1 = 1$ is outside the range $\Rightarrow$ GE is riskier than the market.

Review: General Linear Hypothesis – H_0: Rβ = q

• Suppose we are interested in testing *J* joint hypotheses.

Example: We want to test that in the 3 FF factor model that the SMB and HML factors have the same coefficients, $\beta_{SMB} = \beta_{HML} = \beta^0$.

We can write linear restrictions as H_0: $\mathbf{R}\boldsymbol{\beta} - \mathbf{q} = \mathbf{0}$,
where **R** is a *J*x*k* matrix and **q** a *J*x1 vector.

In the above example (*J*=2), we write:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} \beta_1 \\ \beta_{Mkt} \\ \beta_{SMB} \\ \beta_{HML} \end{bmatrix} = \begin{bmatrix} \beta^0 \\ \beta^0 \end{bmatrix}$$

Review: General Linear Hypothesis – H_0: $\mathbf{R}\beta = \mathbf{q}$

- Q: Is **Rb** – **q** close to **0**? Two different approaches to this questions.

Approach (1). Wald test.

We base the answer on the discrepancy vector:

$$\mathbf{m} = \mathbf{Rb} - \mathbf{q}.$$

Then, we construct a Wald statistic:

$$W = \mathbf{m}' \,(\text{Var}[\mathbf{m}\,|\,\mathbf{X}])^{-1}\, \mathbf{m}$$

to test if **m** is different from 0.

$$W^* = (\mathbf{Rb} - \mathbf{q})' \,\{\mathbf{R}[s^2(\mathbf{X'X})^{-1}]\mathbf{R}\}^{-1}\, (\mathbf{Rb} - \mathbf{q})$$

- If (**A5**) is assumed: $\quad F = W^*/J \sim F_{J,T-k}$

- If (**A5**) is not assumed, results are only asymptotic: $J * F \xrightarrow{d} \chi^2_J$ 23

Review: Wald Test Statistic for H_0: $\mathbf{R}\beta - \mathbf{q} = \mathbf{0}$

Example: In the 3 FF factor model for GE (T=571), we test:

H_0: $\beta_{Mkt} = 1$, $\beta_{SMB} = -0.1$ and $\beta_{HML} = 0.3$.

H_1: $\beta_{Mkt} \neq 1$ and/or $\beta_{SMB} \neq -0.1$ and/or $\beta_{HML} \neq 0.3$. $\quad \Rightarrow J = 3$

```
library(car)
linearHypothesis(fit_ge_ff3, c("Mkt_RF = 1","SMB = -0.1", "HML= 0.3"), test="F")  # exact test

Hypothesis:
Mkt_RF = 1
SMB = - 0.1
HML = 0.3

Model 1: restricted model
Model 2: ge_x ~ Mkt_RF + SMB + HML

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1    568 1.6067
2    565 1.5389  3  0.067761 8.2927 2.094e-05 ***
```

Review: General Linear Hypothesis – H_0: $R\beta = q$

- Q: Is **Rb** – **q** close to **0**?

Approach (2). F test.

We base the answer on a model loss of fit when restrictions are imposed: RSS must increase and R^2 must go down.

We construct an F test to check if the unrestricted RSS (RSS_U) is different from the restricted RSS (RSS_R).

$$F = \frac{(RSS_R - RSS_U)/J}{RSS_U/(T-k_u)} \sim F_{J,T\text{-}K}.$$

25

Review: F Test – Are SMB and HML Priced?

Example: We want to test if the additional FF factors (SMB, HML) are significant for GE (T=570).

Unrestricted Model:

(U) $(r_{GE,t} - r_f) = \beta_0 + \beta_1\,(r_{m,t} - r_f) + \beta_2\, SMB_t + \beta_3\, HML_t + \varepsilon_t$

Hypothesis: H_0: $\beta_2 = \beta_3 = 0$

H_1: $\beta_2 \neq 0$ and/or $\beta_3 \neq 0$

Then, the Restricted Model:

(R) $(r_{GE,t} - r_f) = \beta_0 + \beta_1\,(r_{m,t} - r_f) + \varepsilon_t$

Test: $F = \frac{(RSS_R - RSS_U)/J}{RSS_U/(T-k_u)} \sim F_{J,T-k}$, $J = (k_U - k_R) = 4 - 2 = 2$

Review: F Test – Are SMB and HML Priced?

Example (continuation):

```
fit_ge_ff3 <- lm(ge_x ~ Mkt_RF + SMB + HML)      # U Model
e_ge3 <- fit_ge_ff3$residuals                     # Unrestricted residuals (e_U)
RSS_u <- sum(e_ge3^2)                             # Unrestricted RSS (RSS_U)
b_ge3 <- fit_ge_ff3$coefficients
k_u <- length(b_ge3)                              # k_U

fit_ge_r <- lm(ge_x ~ Mkt_RF)                     # R Model
e_ge_r <- fit_ge_r$residuals                      # Restricted residuals (e_R)
RSS_r <- sum(e_ge_r^2)                            # Restricted RSS (RSS_R)
b_ge_r <- fit_ge_r$coefficients
k_r <- length(b_ge_r)                             # k_R

J <- k_u – k_r                                    # J = df of numerator
F_test <- ((RSS_r – RSS_u)/J)/(RSS_u/(T – k_u))
> F_test
[1] 19.5149
```

Review: F Test – Are SMB and HML Priced?

Example (continuation):

```
> F_test
[1] 19.5149> qf(.95, df1=J, df2=(T-k))               # F_{2,566,.05} value (≈ 3)
[1] 3.011672                                     ⇒ Reject H0.
> p_val <- 1 - pf(F_test, df1=J, df2=(T-k))        # p-value of F_test
> p_val
[1] 0.005913161                                  ⇒ p-value is very small (0) ⇒ Reject H0.
```

Conclusion: Yes, the low *p-value* rejects H_0. That is, SMB and HML are priced factors for GE.