Financial Econometrics: A refresher class

Gaëlle LE FOL

Master 203 & Master 104

September 2020

Outline

Introduction
- Definition
- Basics in statistics and mathematics
- The data
- Limits, critics & model construction

The classical linear regression model
- Presentation of the model
- Assumptions
- Properties of the OLS estimator
- Precision and standard errors
- Goodness of fit

Introduction to statistical inference
- The idea
- Distribution of the estimated parameters
- Significativity test
- Confidence interval approach
- The level of significance : choosing α
- The exact level of significance : the p-value

Simple linear regression with Gretl or R
- Presentation
- Choosing a software
- Getting started with Gretl
- Financial application

The multiple regression model
- Matrix form of the model
- OLS estimators
- Estimation of the variance of the errors I

Statistical inference in the multiple regression model
- **Distribution of the estimated parameters**
- **Testing individually the estimated parameters**
- **Testing simultaneously the estimated parameters**

Financial application : The APT model
- Constructing factors and excess returns
- Regression model

Heteroskedasticity and serial correlation
- The generalised regression model
- The GLS estimators
- Heteroskedasticity of the errors
- Dealing with Heteroskedasticity
- Autocorrelation of the errors
- Dealing with Autocorrelation

Financial application : The APT model (Ctd)

Other assumptions violation and diagnostic tests
- Stochastic regressors and exogeneity
- Normality of the errors
- Multicollinearity
- Model selection and diagnostic tests
- Selection criteria

Alternative to OLS
- Two stage least squares
- Maximum likelihood estimation
- Generalized Least Squares
- Quantile regression

Appendix & References

- Suppose that the errors are distributed as a normal $\varepsilon \sim N(0, \sigma_\varepsilon^2 \mathbf{I})$ and following equation (42) :

$$\hat{\beta} = \beta + (X'X)^{-1}X'\varepsilon,$$

  $\hat{\beta}$ is linear in ε and is normally distributed as a normal $\hat{\beta} \sim N(\beta, \Omega_{\hat{\beta}})$, where $\Omega_{\hat{\beta}} = \sigma_\varepsilon^2 (X'X)^{-1}$ (see equation 45).

- For any $\hat{\beta}_i$ associated to i^{st} explanatory variable x_{it}, we have :

$$\hat{\beta}_i \sim N(\beta_i, \sigma_\varepsilon^2 a_{i+1,i+1}),$$

  where $a_{i+1,i+1}$ is the $(i+1)^{st}$ diagonal element of $(X'X)^{-1}$. We have :

$$\frac{\hat{\beta}_i - \beta_i}{\sigma_\varepsilon \sqrt{a_{i+1,i+1}}} \sim N(0,1).$$

Distribution of the estimated parameters II

From equation (46), we know :

$$\widehat{\sigma}_\varepsilon^2 = \frac{e'e}{T-k-1}.$$

Recall that if $w \sim N(0, \sigma_w^2 \mathbf{I})$ then $\frac{w'w}{\sigma_w^2}$ is distributed as a χ^2 and

$$\frac{e'e}{\sigma_\varepsilon^2} \sim \chi^2_{T-k-1}$$

$$\frac{(T-k-1)\widehat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} \sim \chi^2_{T-k-1}$$

Recall that for any $z \sim N(0,1)$ and $\nu \sim \chi^2_r$, $\frac{Z\sqrt{r}}{\sqrt{\nu}} \sim t(r)$. We then have :

$$\frac{\frac{\hat{\beta}_i - \beta_i}{\sigma_\varepsilon \sqrt{a_{i+1,i+1}}} \sqrt{T-k-1}}{\sqrt{\frac{(T-k-1)\widehat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2}}}$$

Finally, we have :

$$\frac{\hat{\beta}_i - \beta_i}{\widehat{\sigma}_\varepsilon \sqrt{a_{i+1,i+1}}} \sim t(T-k-1) \tag{47}$$

Testing individually the estimated parameters I

- Testing the null hypothesis that β_i equals the particular value of β^* :

$$\begin{cases} H_0 : \beta_i = \beta^* \\ H_a : \beta_i \neq \beta^* \end{cases}$$

Under H_0,

$$\frac{\hat{\beta}_i - \beta^*}{\widehat{\sigma}_\varepsilon \sqrt{a_{i+1,i+1}}} \sim t(T-k-1)$$

and the decision rule gives :

- If $\left| \frac{\hat{\beta}_i - \beta^*}{\widehat{\sigma}_\varepsilon \sqrt{a_{i+1,i+1}}} \right| \leq t_{p/2}$ we do not reject H_0 that $\beta_i = \beta^*$ at a $100\alpha\%$ significance level.
- If $\left| \frac{\hat{\beta}_i - \beta^*}{\widehat{\sigma}_\varepsilon \sqrt{a_{i+1,i+1}}} \right| > t_{p/2}$ we reject H_0 and $\beta_i \neq \beta^*$ at a $100\alpha\%$ significance level.

Testing individually the estimated parameters II

The most common test is the significance of the parameters :

- Testing the null hypothesis that $\beta_i = 0$:

$$\begin{cases} H_0 : \beta_i = 0 \\ H_a : \beta_i \neq 0 \end{cases}$$

Under H_0,

$$t_{\hat{\beta}_i} = \frac{\hat{\beta}_i}{\widehat{\sigma}_{\hat{\beta}_i}} \sim t(T-k-1) \text{ where } \widehat{\sigma}_{\hat{\beta}_i} = \widehat{\sigma}_\varepsilon \sqrt{a_{i+1,i+1}}$$

- If $\left|\frac{\hat{\beta}_i}{\widehat{\sigma}_{\hat{\beta}_i}}\right| \leq t_{p/2}$ we do not reject H_0 that $\beta_i = 0$ at a $100\alpha\%$ significance level $\Longleftrightarrow$ the variable X_{it} is not significant and does not influence Y_t.
- If $\left|\frac{\hat{\beta}_i}{\widehat{\sigma}_{\hat{\beta}_i}}\right| > t_{p/2}$ we reject H_0 and $\beta_i \neq 0$ at a $100\alpha\%$ significance level $\Longleftrightarrow$ the variable X_{it} is significant and does influence Y_t.

- Testing the significance of an isolated parameter corresponds to a Student t-test.
- Testing the significance of several (all) the estimated parameters at the same time corresponds to a Fisher test. Suppose that the parameters are subject to r constraints :

$$R\beta = r,$$

where R is a $(q, k+1)$ matrix and r a vector of size q. We suppose $q \leq k+1$ and R to be full rank $\Longleftrightarrow$ the constraints are linearly independent.

  - The case where $R = [0 \dots 010 \dots 0]$ and $r = 0$ corresponds to the previous Student-t test.
  - The case where $R = [01 - 1 \dots 0]$ and $r = 0$ corresponds to the test where $\beta_1 = \beta_2$.

- The case where

$$R = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \text{ and } r = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

corresponds to the test of global significance of the parameters or the model.

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\iff \beta_1 = \beta_2 = \ldots = \beta_k = 0.$

Exercise 10 : Test examples

Considering the model

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + \varepsilon_t,\ t = 1, \ldots, T.$$

give R and r in the following cases :

1. $H_0 : \beta_3 = 1\,;\ H_a : \beta_3 \neq 1.$
2. $H_0 : \beta_0 + \beta_1 = 1\,;\ H_a : \beta_0 + \beta_1 \neq 1.$
3. $H_0 : \beta_2 \times \beta_1 = 0\,;\ H_a : \beta_2 \times \beta_1 \neq 0.$

- Aiming at implementing the previous test, we need to consider two models :
  - The unconstrained model $\mathcal{M}_{uc}$ corresponds to the regression of the dependent variable on the set of all explanatory variables without imposing any contraint on the parameters.
  - The constrained model $\mathcal{M}_c$ corresponds to the regression of the dependent variable on a sub-set of explanatory variables that account for the linear contraints on the parameters. **Note that**, we may have to change variables in order to isolate the estimated parameters at the right side of the equal sign for example.
  - The constrained model $\mathcal{M}_c$ is a special case of theunconstrained model $\mathcal{M}_{uc}$.
  - The constrained $\mathcal{M}_c$ and unconstrained $\mathcal{M}_{uc}$ models must be estimated to compute the statistic of the test F.

- The statistic of the test is :

$$F = \frac{(ESS_c - ESS_{uc})/q}{ESS_{uc}/(T-k-1)} \sim F(q, T-k-1) \quad (48)$$

- The Fisher test is testing that the relative gap between the two ESS is significant $H_0 : ESS_c - ESS_{uc} = 0$ and thus, that the constaint is statistically justified.

What is the alternative hypothesis of that test ?

- If $F \leq F(q, T-k-1)$, we do not reject the null hypothesis H_0. In the contrary case, on reject the null.
- The Student-t test of the significativity of β_i can be seen as a Fisher test, where the constrained model resums to the regression of the dependent variable on a set of explanatory variables including all the explanatory variables but X_{it}. In that case, the Student-t statistic, t-stat is such that $t^2(T-k-1) = F(1, T-k-1)$.
- The null hypothesis of the overall significance test of a multiple regression, is $H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0$.
- The unconstrained model is :

$$\mathcal{M}_{uc} : y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + u_t,\ t = 1, \ldots, T,$$

- The constrained model is $\mathcal{M}_c : y_t = \beta_0 + u_t$.

- We have $\hat{y}_t = \hat{\beta}_0$, $\bar{\hat{y}} = \bar{y} = \hat{\beta}_0$ and $q = k$. Finaly,

$$ESS_c = \sum_{t=1}^{T}(y_t - \hat{y}_t)^2 = \sum_{t=1}^{T}(y_t - \bar{y})^2 = TSS$$

- Replacing, ESS_c and ESS_{uc} in

$$F = \frac{(ESS_c - ESS_{uc})/q}{ESS_{uc}/(T-k-1)} \sim F(q, T-k-1), \tag{49}$$

we get,

$$\begin{aligned} F &= \frac{(TSS - ESS)/k}{ESS/(T-k-1)} \\ &= \frac{RSS/k}{ESS/(T-k-1)} \sim F(k, T-k-1) \\ &= \frac{R^2/k}{(1-R^2)/(T-k-1)} \sim F(k, T-k-1) \end{aligned}$$

where RSS is the regression sum of squares.

- The Student test of the significance of β_i can be interpreted as a Fisher test, where the constraint model resumes to regressing the dependant variable over the set of explanatory variables but x_{it}.

Exercise 11 : Constrained vs unconstrained model

Consider the model

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + u_t,\ t = 1, \ldots, T$$

Describe the method (set of regressions included) to test $\{\beta_1 = 1; \beta_3 = -1\}$.

- To perform multiple tests,
  - we estimate the two models $\mathcal{M}_c$ and $\mathcal{M}_{nc}$ and each time, we extract the ESS,
  - we compute the statistic of the test,
  - we compare it to a critical value of the corresponding distribution,
  - we reject H_0 whenever the statistic is greater than the critical value (tabulated value).
- Equivalently,
  - we can calculate the risk of rejecting the true hypothesis (type-I error),
  - compare it to a risk tolerence level we are willing to bear (α),
  - we reject H_0 if the p-value is lower than the risk we are willing to bear.

Outline

Introduction
Definition
Basics in statistics and mathematics
The data
Limits, critics & model construction
The classical linear regression model
Presentation of the model
Assumptions
Properties of the OLS estimator
Precision and standard errors
Goodness of fit
Introduction to statistical inference
The idea
Distribution of the estimated parameters
Significativity test
Confidence interval approach
The level of significance : choosing α
The exact level of significance : the p-value
Simple linear regression with Gretl or R
Presentation
Choosing a software
Getting started with Gretl
Financial application
The multiple regression model
Matrix form of the model
OLS estimators
Estimation of the variance of the errors I
Statistical inference in the multiple regression model
Distribution of the estimated parameters
Testing individually the estimated parameters
Testing simultaneously the estimated parameters
Financial application : The APT model
Constructing factors and excess returns
Regression model
Heteroskedasticity and serial correlation
The generalised regression model
The GLS estimators
Heteroskedasticity of the errors
Dealing with Heteroskedasticity
Autocorrelation of the errors
Dealing with Autocorrelation
Financial application : The APT model (Ctd)
Other assumptions violation and diagnostic tests
Stochastic regressors and exogeneity
Normality of the errors
Multicollinearity
Model selection and diagnostic tests
Selection criteria
Alternative to OLS
Two stage least squares
Maximum likelihood estimation
Generalized Least Squares
Quantile regression
Appendix & References

- The classical APT model explains single stock excess returns by the variations of a market index excess return and other factors. Brooks tries to explain Microsoft monthly returns by unexpected changes of macroeconomics and financial variables.
- File called macro.xls : 325 monthly observations from March 1986 to April 2013. Variables are : Dates, Microsoft stock price, $S\&P500$ index value, the consummer price index, an industrial production index, Tresury bill yields with maturities 3 months, 6months, 1 year, 3 years, 5 years and 10 years, a measure of money supply, a consumer credit series, an a credit spread series.

- Load the file under gretl and construct the returns for the microsoft stock and the $S\&P500$ index.
```
genr rmsoft = log(Microsoft/Microsoft(-1))
```
- Assuming naive expectations, unexpected changes are obtained as the first difference in any variable of interest. Construct the interesting unexpected changes in the macro and financial variables.
```
genr dcredit = CONSUMER_CREDIT - CONSUMER_CREDIT(-1)
```
- Construct the slope of the yield curve of interest rate
```
genr term = USTBY10 - USTB3M
```
- Construct the risk free rate as the monthly 3 months treasury bill yields.
```
genr mustb3m = USTB3M/12
```
- Construct the excess returns for Microsoft and the $S\&P500$.
```
genr ermsoft = rmsoft - mustb3m
```
- Run the appropriate regression.
```
Model APT <- ols ermsoft const ersandp dprod dcredit dinflation
dmoney dspread dterm
```
- Interpret the output.

The APT-style model of Brooks III

Model APT : OLS, using observations 1986 :05–2007 :04 ($T = 252$)
Dependent variable : ermsoft

	Coefficient	Std. Error	t-ratio	p-value
const	−0.587603	1.45790	−0.4030	0.6873
ersandp	1.48943	0.203276	7.3271	0.0000
dprod	0.289322	0.500919	0.5776	0.5641
dcredit	$-5.58389\text{e}{-005}$	0.000160491	−0.3479	0.7282
dinflation	4.24781	2.97734	1.4267	0.1549
dmoney	−1.16153	0.713974	−1.6268	0.1051
dspread	12.1578	13.5510	0.8972	0.3705
rterm	6.06761	3.32136	1.8268	0.0689

Mean dependent var	−0.420803	S.D. dependent var	15.41135
Sum squared resid	47480.62	S.E. of regression	13.94965
R^2	0.203545	Adjusted R^2	0.180696
$F(7, 244)$	8.908218	P-value(F)	$9.08\text{e}{-10}$
Log-likelihood	−1017.642	Akaike criterion	2051.284
Schwarz criterion	2079.520	Hannan–Quinn	2062.646
$\hat{\rho}$	−0.078198	Durbin–Watson	2.156221