

1. 自前ファイルの分析

今回は、gretl に付属した sample ファイルを OLS (最小二乗法) で分析した。これは、データに最も当てはまる直線を求めるものだった。同様な分析を自分で Web や調査から入手したデータに対して行ってみよう。それができれば、回帰分析を駆使して様々なレポートや卒論を作成することができる。

1.1 csv データの読み込み

まず、私が用意したファイルを開こう。

- 1) デスクトップの「統計解析論データ」のフォルダーをダブルクリックしてあげ、その中の、sample.csv をダブルクリックする。
- 2) エクセルが立ち上がる。
- 3) 保存するときは、「ファイル」「名前をつけて保存」を左クリックで選択。出てきたウィンドウの「ファイルの種類」のメニューリストから、「CSV(カンマ区切り)」を選ぶ。後のメッセージには「はい」を必要なだけ左クリックする。excel は終了させる。
- 4) gretl を立ち上げ、メニューバーの「ファイル」「データを開く」「import」「CSV」を左クリックで選ぶ。
- 5) 出てきたウィンドウの「comma()」のラジオボタンを左クリックして、「OK」ボタンを左クリックする。
- 6) 出てきたウィンドウの「デスクトップ」をクリックして、デスクトップ上の「統計解析論データ」をダブルクリックして、その中の「sample.csv」をダブルクリックする。すると、gretl に sample データを読み込むことができる。

皆さんは、自前のファイルを CSV 形式でどこかにセーブして、それを 4)以降の手順で gretl に読み込ませればよい。例えば、Web 上のデータには excel 形式のものがあるが、いったんダウンロードして、3)の手順で、CSV 形式のファイルを作成し、4)以降の手順で対応すればよい。また、データの入力から始める場合は、excel を立ち上げて、データを入力し、3)の手順で CSV 形式のファイルを作成する。以下では、そのうち excel 形式のデータを CSV に変換する方法を練習しよう。

- 7) デスクトップの「統計解析論データ」のフォルダー内の、exceldata.xls をダブルクリックする。
- 8) 「ファイル」「名前をつけて保存」を左クリックで選択。出てきたウィンドウの「ファイルの種類」のメニューリストから、「CSV(カンマ区切り)」を選ぶ。

9) excel は終了させる。

1.2 時系列データの場合のおまじない

データファイルがクロスセクションデータの場合は、1.1の手順のみでよいが、時系列データの場合は、時系列データあることを gretl に認識させなければならない。クロスセクションデータとは、ある一時点のデータをあつめたもので、例えば、2005年のある業種の会社全ての財務データなどである。それに対して、時系列データとは、同一のデータを継続的に取得したものである。

- 1) メニューバーから「データ」「Data Structure」を左クリックで選択
- 2) でてきたウィンドウの「Time Series」のラジオボタンを左クリックして選択
- 3) 「進む」ボタンを左クリック
- 4) 次のウィンドウで「Monthly」のラジオボタンを左クリックして選択(このデータは月次データなので。もし、日次データなら Daily, 四半期次なら Quarterly, 年次データなら Annual を選択)
- 5) 「Starting Observation」のウィンドウの中を「1990:08」(1990年8月)にする。その際、ウィンドウの横の 印を利用する。
- 6) 「進む」ボタンを左クリック。
- 7) 出てきたウィンドウの「OK」ボタンを左クリック。
- 8) gretl メインウィンドウで「month」の行を左クリックした後、右クリックし、ポップアップしたメニューの中の「Delete」を左クリックする。
- 9) 新たにでてくるウィンドウの「OK」ボタンと左クリックする。

2. gretl データファイルの作成と読み込み

データを次に gretl で利用するために保存したいだろう。

- 1) 持ってきた USB メモリを PC の後の USB ポートに差し込む。この時に、USB メモリの向きに注意する。合わないときには、無理に押し込もうとせず、上下を入れ替えて力をかけず差し込む。
- 2) Ctrl-S (Ctrl キーと S キーを同時におす)。
- 3) 出てきたウィンドウの左から「マイ コンピュータ」を左クリックする。
- 4) ウィンドウ内に表示されたなかで、USB メモリに対応する「リムーバブル ディスク」を左クリックする。
- 5) 「保存」ボタンを左クリックする。

次にこのデータを gretl に読み込んでみよう。

- 6) gretl メインウィンドウにもどり、メニューバーから「データを開く」「user file」を左クリック
- 7) 出てきたウィンドウの左から「マイ コンピュータ」を左クリックする。

10) ウィンドウ内に表示されたなかで、USBメモリに対応する「リムーバブル ディスク」を左クリックする。

11) 先ほど作ったファイル「XXX.gdt」を左クリックし、「開く」ボタンを左クリック。

12) 次に出てきたウィンドウの「はい」ボタンを左クリック。

USBメモリを抜くためには、必ず画面右下の「ハードウェアの取り外し」を左クリックし、「USB大量記憶装置デバイス」を左クリックするが、いまはしない。

3. gretl セッションファイルの活用

作業の途中でPCを終了させなければならないときのために、活用すべきはセッションファイル(session file)である。現在のgretlの状態をすべてファイルに保存してくれ、次にPCで作業をするときには、このセッションファイルを開くことで、作業を再開することができる。

1) メニューから「ファイル」「Session files」「save session」をマウスでたどり、左クリック。

2) 出てきたウィンドウの左から「マイ コンピュータ」を左クリックする。

3) ウィンドウ内に表示されたなかで、USBメモリに対応する「リムーバブル ディスク」を左クリックする。

4) セッションの名前を「ファイル名」の入力ウィンドウにタイプ入力する。例えば、「070518」などといれる。

5) 「保存」ボタンを左クリックする。

さて、セッションファイルのありがたみを知ろう。

6) gretl メインウィンドウにもどり、右上の×を左クリックし、gretlを終了させる。

7) 先ほど保存したセッションファイル「XXXX.gretl」をダブルクリックする。

8) gretl メインウィンドウの他に出るウィンドウは実は使い道がかなりあるのだが、とりあえず無視しよう。

以上で、先ほどの状態に復帰することができる。

4. データの export

データを他のファイルに書き出したいことがある。その結果を今度は、excelなどで読み込ませたいといったものだ。そのデータで、グラフを書き直すなどの理由が考えられる。

1) メニューから「ファイル」「export data」「CSV」をマウスでたどり、左クリックする。

2) 新たにでてきたウィンドウの「OK」ボタンを押す。

3) 次のウィンドウでどの変数を書き出すのかを指定する。ここでは、すべての変数を選択したいので、「All->」ボタンを左クリックし、「OK」ボタンを押す。

- 4) 出てきたウィンドウの左から「マイ コンピュータ」を左クリックする。
- 5) ウィンドウ内に表示されたなかで、USB メモリに対応する「リムーバブル ディスク」を左クリックする。
- 6) 保存したい名前を「ファイル名」の入力ウィンドウにタイプ入力する。例えば、「070518」などといれる。
- 7) 「保存」ボタンを左クリックする。

家でこのデータを excel で開いてみよう。

5. ワードソフトとの連携 レポートを作ろう

gretl はもちろん統計分析、計量経済学分析を行うツールだが、そのツールが出力した様々な図を文書の中で利用できなければ、その価値もおおいに減退してしまう。そこで、gretl の出力したグラフをワードソフトに貼り込むやり方を解説する。

そのまえに、各変数の意味を説明しておこう。

Nikkei: 東京証券取引所日経 225 種平均

Nreturn: 日経 225 種平均の前月との変化率 (月次収益率)

SP500: ニューヨーク証券取引所 Stock&Poors500 種平均

Spreturn: Stock&Poors500 種平均の月次収益率

- 1) gretl メインウィンドウにもどり、メニューバーから「モデル」 「Ordinary least squares」を選択する。(このやり方は前回の回帰と同じ)
- 2) 被説明変数 (dependent variable) は Nreturn, 説明変数 (independent variables) は Spreturn とする。指定の仕方は前回参照。指定後 OK をおす。
- 3) 次に、このモデルのあてはまりを、グラフ化してみよう。まず、実際の値と予測値のグラフを書いてみよう。最小二乗法 (Ordinary least squares) の結果を出力しているウィンドウのメニューバーから、「Graphs」 「Fitted actual plot」 「Against time」をマウスでたどり、左クリックで選択する。
- 4) 次に、説明変数と被説明変数の散布図と回帰直線のグラフを書く。最小二乗法 (Ordinary least squares) の結果を出力しているウィンドウのメニューバーから、「Graphs」 「Fitted actual plot」 「Against Spreturn」をマウスでたどり、左クリックで選択する。
- 5) これで 2 つグラフが書けた。これを、ワードに張り込んでみよう。まず、word を立ち上げる。
- 6) グラフの書いてあるウィンドウ上で左クリックして「Copy to clipboard」 「color」とマウスでたどり左クリックする。もし、単色のグラフでよければ、「Copy to clipboard」 「monochrome」とマウスでたどり左クリックする。
- 7) word 上で Ctrl-V (Ctrl キーと V キーを同時に押す) <ペーストすると呼ぶ>

- 8) グラフの大きさを変更したい場合、ペーストしたグラフの右下にマウスマウスカーソルを持ってくると、斜め両矢印になるところがあるので、そこでマウスの左ボタンを押したままマウスを動かす。好みの大きさになったところで、左ボタンを押し続けるのをやめると、大きさが変更される。
- 9) もう一つのグラフも 8) の手順で word の画面上にペーストする。
- 10) 両方のグラフが同じ大きさで横に並ぶように大きさを調整する。

6. 基礎統計学を gretl で実行 (教科書第 1 章)

6.1 平均と分散, 標準偏差

教科書 p.13 参照

- 1) p.13 の国名(日本語ではなく英語で), G と L のデータを excel を立ち上げて, 入力。(gretl は日本語のデータを扱うことができない)
- 2) これを CSV 形式で USB メモリ上に書き出す。名前は適当につけよ。
- 3) gretl を起動
- 4) メニューバーの「ファイル」 「データを開く」 「import」 「CSV」を左クリックで選ぶ。
- 5) 出てきたウィンドウの「comma()」のラジオボタンを左クリックして、「OK」ボタンを左クリックする。
- 6) 出てきたウィンドウの「マイ コンピュータ」をクリックして、ウィンドウ内に表示されたなかで、USB メモリに対応する「リムーバブル ディスク」を左クリックする。その中の先ほど作ったファイルを左クリックして指定する。これで gretl にデータを読み込むことができる。
- 7) メニューバーの「データ」 「Select All」で全ての変数を選択し、メニューバーの「View」 「Summary Statistics」で G,L 等のデータの平均, 分散, 標準偏差などの基本統計量が表示される。

6.2 相関係数

- 1) gretl のウィンドウにもどり G の上にマウスマウスカーソルをおいて、マウスの左ボタンを押しながら、下の L の上にマウスマウスカーソルを持って行ってボタンを放す。
- 2) メニューバーの「View」 「Correlation Matrix」を選ぶと、両者の相関が与えられる。

6.3 平均の検定

- 1) 先ほどの月次収益率のデータを使用する。従って、1.1 の手順をもう一度行う。
- 2) まず、Nreturn(日経 225 の月次収益率)の平均が 0 かどうかを検定する。帰無

仮説は平均が 0 , 対立仮説は平均が 0 でないである .

- 3) gretl メインウィンドウにもどり , ツールバーの「Tools」 「Test Statistics Calculator」を選択する
- 4) 新しく出てきたウィンドウの「Use variable from dataset」の左のチェックボックスを左クリックし , その右のボックスは値のリストをプルダウンして Nreturn を左クリックする .
- 5) 「H0: mean=」の右のボックスに帰無仮説での値 0 をキーボードから入力する .
- 6) 「OK」ボタンを押す .
- 7) 出てきた文字だけのウィンドウを見る .

Null hypothesis: population mean = 0

Sample size: n = 77

Sample mean = -0.00452889, std. deviation = 0.0738912

Test statistic: $t(76) = (-0.00452889 - 0)/0.00842068 = -0.53783$

Two-tailed p-value = 0.5923

(one-tailed = 0.2961)

判定は Two-tailed p-value を 0.05 と比較する . これより小さければ , 収益率の平均は 0 とことなり , そうでなければ , 0 とかわらない .

6 . 4 平均差の検定

- 1) gretl メインウィンドウにもどり , , メニューバーの「モデル」 「Bivariate Tests」 「Difference of means」 「Assuming unequal variances」と選択する . (両者の平均は等しいかが帰無仮説 , 対立仮説は等しくない)
- 2) 新しくでてきたウィンドウ上で , Nreturn と Spreturn を選択し , 「OK」ボタンを押す .
- 3) 出力結果の判定

Equality of means test (assuming unequal variances)

Nreturn: Number of observations = 77

Spreturn: Number of observations = 77

Difference between sample means = $-0.00452889 - 0.0115698 = -0.0160986$

Null hypothesis: The two population means are the same.

Estimated standard error = 0.00908736

Test statistic: $t(152) = -1.77154$

p-value (two-tailed) = 0.0784743

基本的には p-value を見る . 0.05 より小さければ , 異なる . そうでなければ , 同じと判定される .

もし , 対立仮説が SP500 の収益率の方が大きいなら , p-value は半分 , すなわち , 0.039 になるので , 帰無仮説は棄却され , SP500 の月次収益率の平均の方が Nikkei225 の月次収益率の平均よりも高かったことになる .

6 . 5 分散比の検定

- 1) 収益率の分散 = ヴォラティリティの検定を行う . 帰無仮説は両方のヴォラティリティが等しい , 対立仮説は両方のヴォラティリティがことなるである .
gretl ウィンドウに戻り , 「モデル」 「Bivariate Tests」 「Difference of variances」を選択 .
- 2) 新しくでてきたウィンドウ上で , Nreturn と Spreturn を選択し , 「OK」ボタンを押す .
- 3) 以下の出力を読む

Equality of variances test

Nreturn: Number of observations = 77

Spreturn: Number of observations = 77

Ratio of sample variances = 6.07493

Null hypothesis: The two population variances are equal

Test statistic: F(76,76) = 6.07493

p-value (two-tailed) = 7.45111e-014

p-value が 0.05 より小さければ , 両者の分散は異なり , そうでなければ , ことならない .

終了するときには , セッションファイルを作って USB メモリーに格納しておこう . 作り方は , 3 を参照のこと .

問題

- 1 . SP500 月次収益率の基本統計量の値をもとめよ .
 - 2 . Nikkei225 月次収益率の基本統計量の値を求めよ .
- なお , 1 , 2 とともに統計用語は邦訳すること .
- 3 . SP500 月次収益率と Nikkei225 月次収益率の相関係数を求めよ

- 4 . 私の Web 上の Nikkeitopix.xls は日次の株価データだが，これを使って，Nikkei225 の日次収益率の時系列グラフを書け．(本来休日など欠損値があるが，無視して日次データとして解析せよ)
- 5 . 上記データを用いて，以下の検定を行え，
 - (1) Nikkei225 の日次収益率平均は 0 であるのか 0 ではないかの検定
 - (2) Nikkei225 の日次収益率平均と円ドルレートの変動率平均が同じか異なるのかの検定
 - (3) Nikkei225 の日次収益率と円ドルレートの変動率の分散が等しいか等しくないのかの検定