

『計量経済学の第一歩』

田中 隆一 (著)

統計ソフトウェアのチュートリアル

発行所 株式会社有斐閣

2015年12月20日 初版第1刷発行

ISBN 978-4-641-15028-7

©2015, Ryuichi Tanaka, Printed in Japan

1. Stata (ステータ)

Stata は、StataCorp 社の販売している統計ソフトウェアで、計量経済学においてもっともよく使われています。最新の計量経済学的手法の論文を執筆する際に、Stata による推定方法（コマンド）も同時に発表されることがよくあり、基本的な分析からより進んだ最新の手法まで幅広くカバーされています。

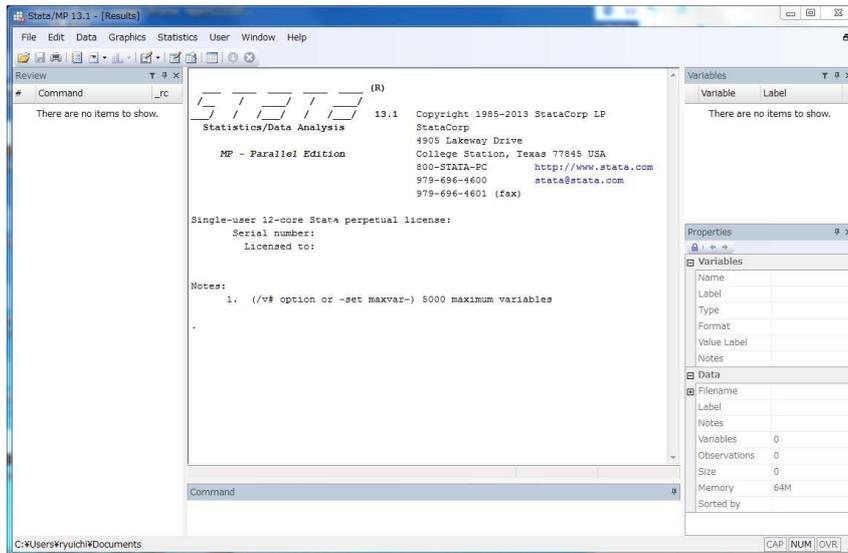
Stata は簡単な単回帰モデルや重回帰モデルの推定はもちろんのこと、本書で紹介したすべての推定方法についてコマンドが用意されています。さらに、複数のデータセットを融合したり、変数を加工したりする上では他の統計ソフトウェアに比べて格段に扱いやすくなっている点も特徴的です。最新版（version 14）では、日本語の変数名を持ったデータをそのまま読み込んでくれる点も便利です。

このチュートリアルでは、第 5 章の実証分析問題 5-A を使って、Windows での Stata の使い方を簡単に説明します。ここでは (1) コマンドを選びながら 1 つずつ進める方法、(2) ダウンロードした do ファイルのコマンドをコマンドウィンドウにコピーしながら進める方法、(3) do ファイルを使って一気に進める手順の 3 つを紹介します。

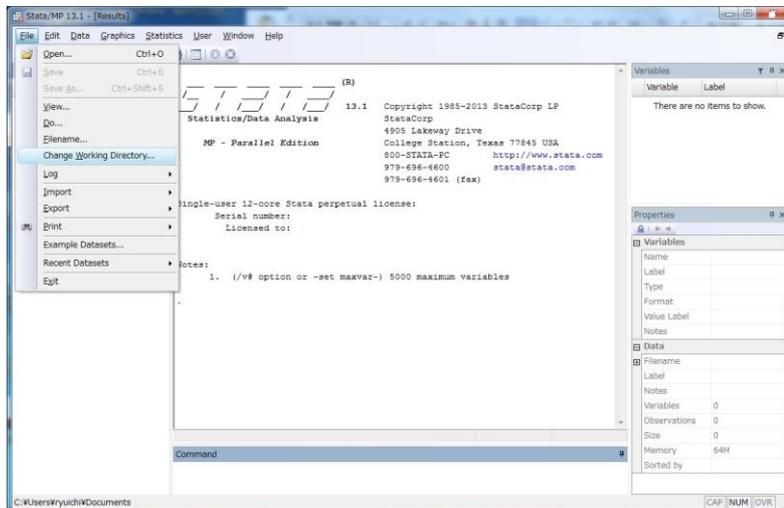
共通の準備

- ▼ データと do ファイルの準備：教科書のウェブサポートページにアクセスして、データと do ファイルをダウンロードしましょう。第 5 章の実証分析問題 5-A の Stata 用データファイルは、「5_1_income.dta」、do ファイルは「5_empirical_A.do」です。これらのファイルの保存先として、デスクトップに新しいフォルダ（ファイルの保存先と名前はなんでも構いませんが、ここでは保存先としてデスクトップを、フォルダ名として“data”としておきます）を作り、そこに 2 つのファイルを保存します。

- ▼ Stata の起動：デスクトップにアイコンがあれば、それをダブルクリックすることで Stata が起動して、次の画面が現れます（もしアイコンがない場合には、スタートメニューから Stata を立ち上げてください）。Stata13, Windows 版の場合は、次の画面が表示されます。



- ▼ 作業フォルダの指定：データファイルと do ファイルを保存した「data」フォルダを作業用フォルダとして指定します。「File」→「Change Working Directory...」と移動します。



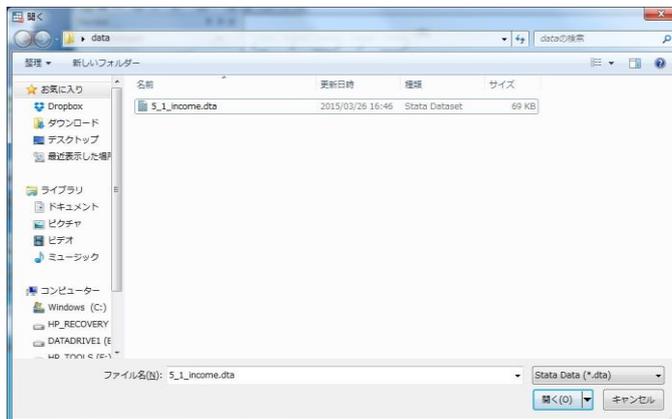
- ▼ 「Change Working Directory...」をクリックすると、次の画面が現れますので、先ほど作成した「data」フォルダを選んで「OK」をクリックします。



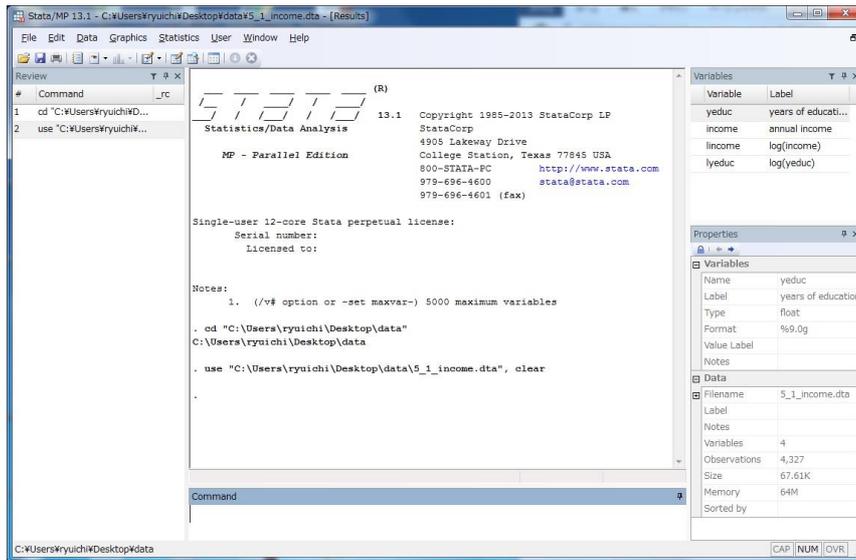
これで準備は整いました。

コマンド選択による方法

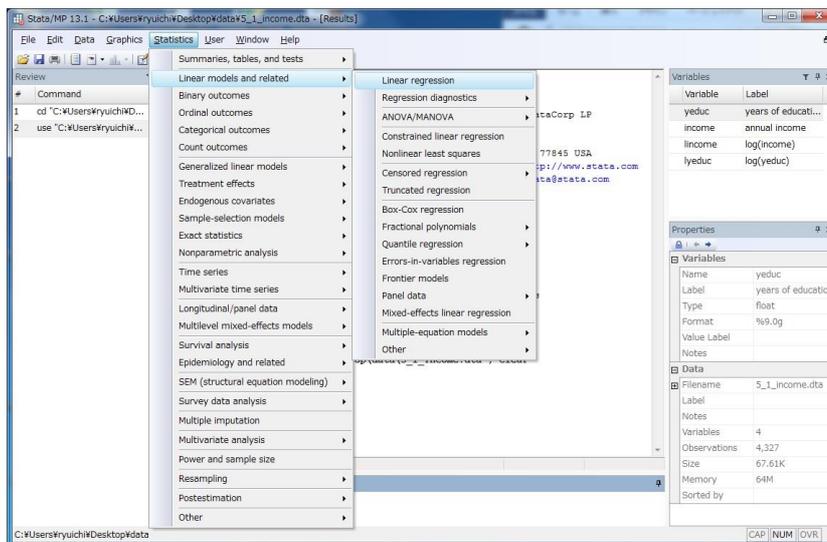
- ▼ データの読み込み：データファイルを開くためには、まず「File」→「Open」（または「File」の下にあるフォルダアイコン  をクリック）すると、フォルダウィンドウが開きますので、デスクトップに作成した「data」フォルダを選択してください。そこに先ほど保存した「5_1_income.dta」が表示されていれば、正しく選択できています。



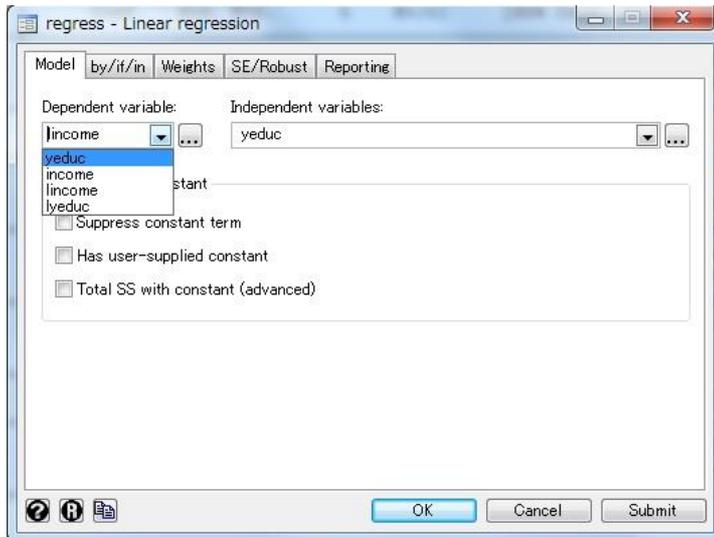
- ▼ そこで「5_1_income.dta」を選択して、「開く (O)」をクリックすると、次の画面が表示され、右上の「Variable」ウィンドウに読み込んだデータの変数名（ここでは yeduc, income, lincome, lyeduc の4変数）が表示されています。



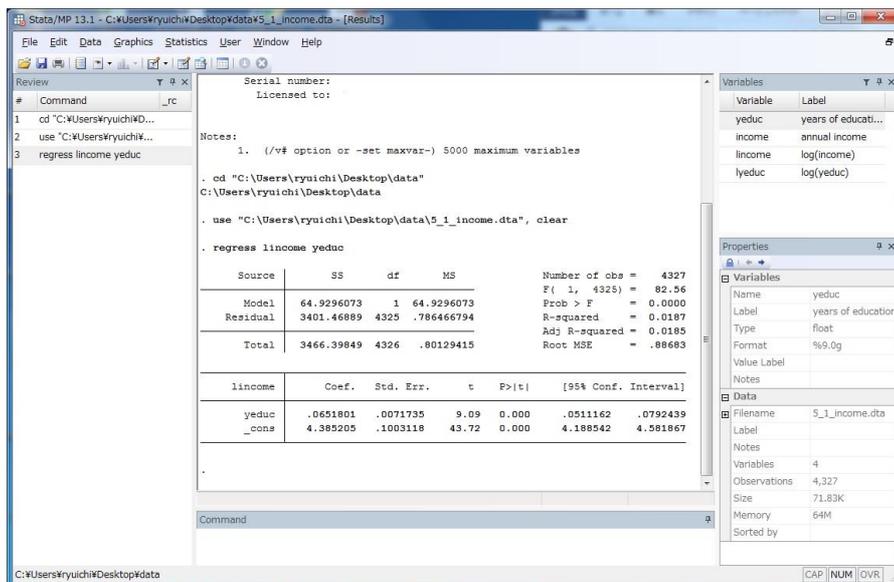
- ▼ 回帰モデルを推定する：「Statistics」 → 「Linear models and related」 → 「Linear regression」と移動します。



- ▼ 「Linear regression」をクリックすると、次の画面が現れるので、「Dependent variable:」に被説明変数を、「Independent variables:」に説明変数を選択します。ここでは、被説明変数に「lnincome」を、説明変数に「yeduc」を選びます。それぞれのボックスの右隣にある下向きの三角のアイコンをクリックすると、読み込んだ変数のリストが表示されますので、その中から選んでください。

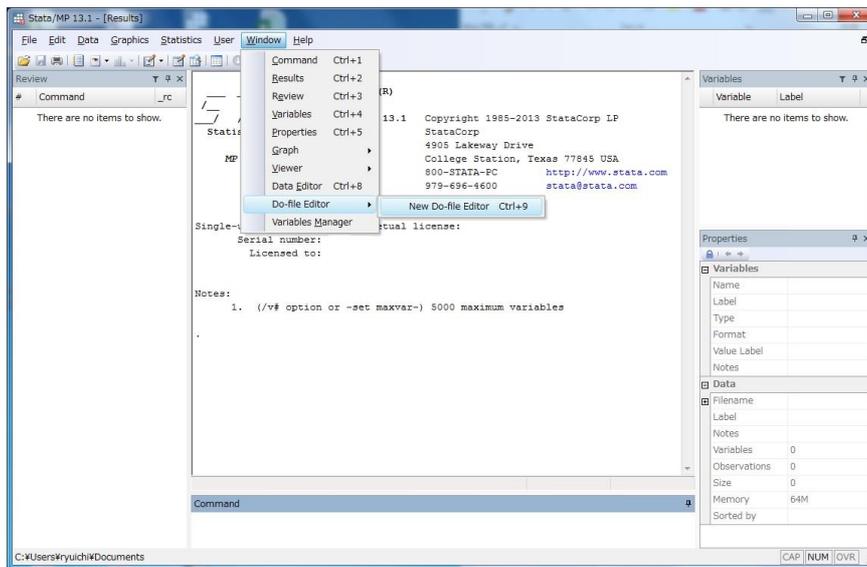


- ▼ 説明変数と被説明変数の選択がすんだら、最後に「OK」をクリックしてみましょう。次の画面が表示されれば、推定がうまくできたことになります。

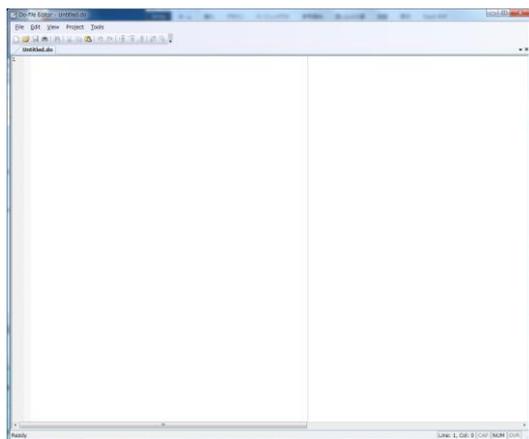


do ファイルのコマンドを Command ウィンドウにコピーしながら進める方法

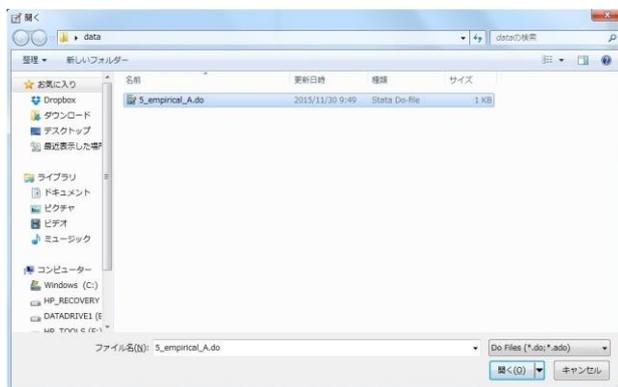
- ▼ do ファイルを開く : 「Window」 → 「Do-file Editor」 → 「New Do-file Editor」と移動します。



- ▼ 「New Do-file Editor」をクリックすると、「Do-file Editor」という次の画面が開きます。



- ▼ 「Do-file Editor」の「File」→「Open」→「File...」（または「File」の下にあるフォルダアイコン)をクリックすると、次のフォルダウィンドウが開きます。



- ▼ 先ほどダウンロードした「5_empirical_A.do」を選択して、「開く (O)」をクリックすると、「5_empirical_A.do」の内容が Do-file Editor に表示されます。

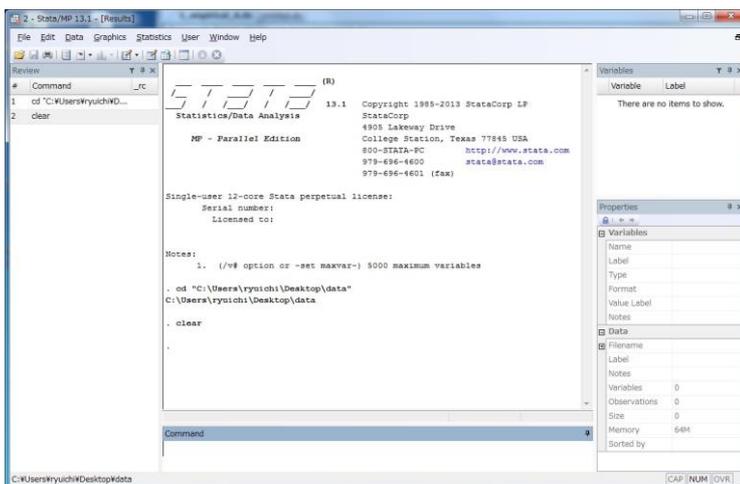
```

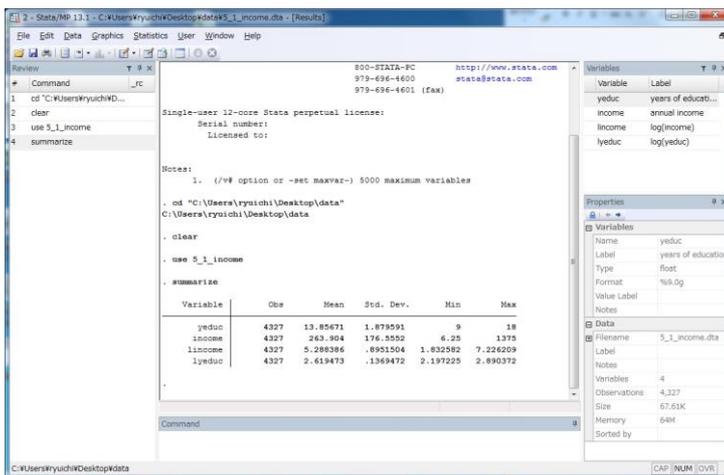
1 /* First steps in econometrics */
2 /* Chapter 5: 5-2 */
3 /* Ryuichi Tanaka (University of Tokyo) */
4
5 clear
6
7 use 5_1_income
8
9 summarize
10
11 regress income yeduc
12 regress lincome yeduc
13 regress income lyeduc
14 regress lincome lyeduc
15
16
17
18
19
20
21
    
```

- ▼ それぞれの行を実行するためには、do ファイルの行をコピーして、それを真ん中下にある「Command ウィンドウ」に貼り付けたうえで「enter」を押すだけでよいのですが、この do ファイルの内容をもう少し詳しく解説しましょう。

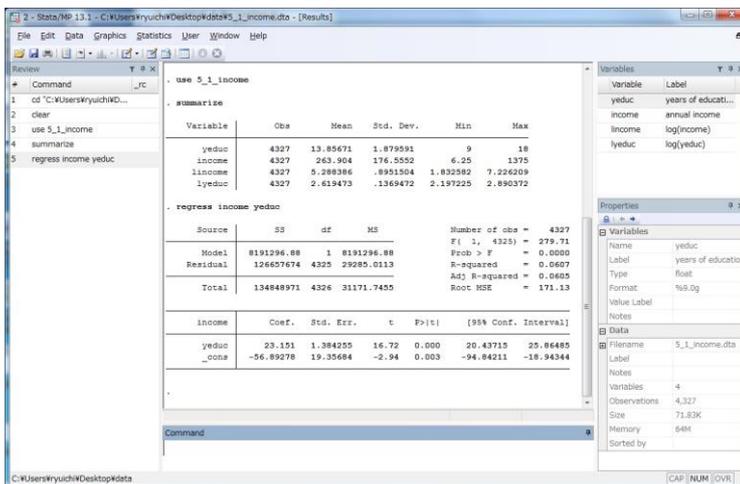
まず、最初の 3 行の「/* */」で囲まれた部分は単なるコメントなので、Stata は実行しません。

5 行目の「clear」は、それまでに読み込んでいたデータをメモリから消去するコマンドです。



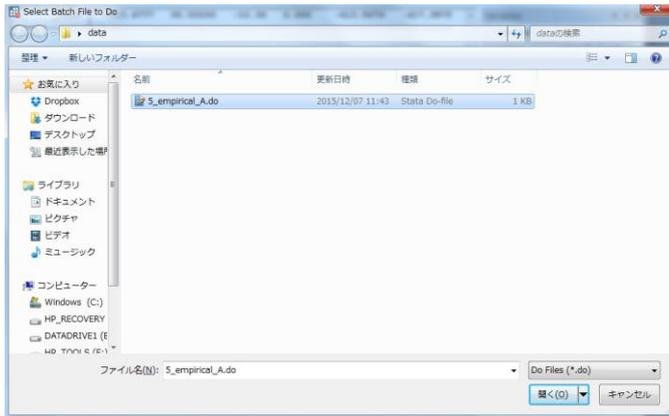


- ▼ 11行目の「regress income yeduc」は、「incomeをyeducに回帰する回帰モデル」を推定するためのコマンドです。

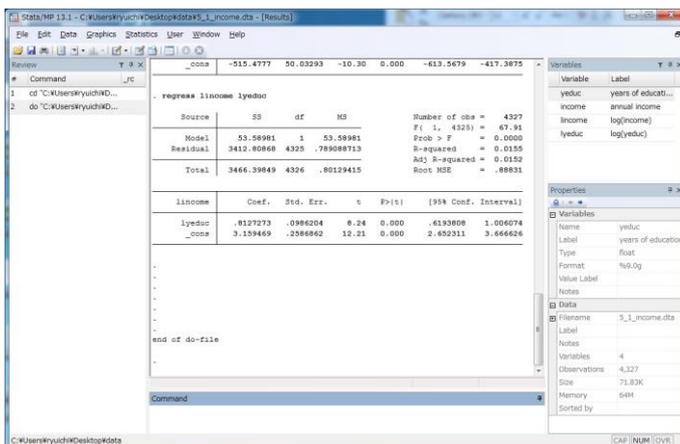


do ファイルを一気に実行する方法

- ▼ 「File」 → 「Do...」をクリックすると、次の画面が表示されます。



- ▼ 「5_empirical_A.do」を選んで「開く (O)」をクリックしてください。次の画面が表示されれば、do ファイルに書いてあるコマンドをすべて実行したことになります。画面を上にもスクロールすれば、最初のほうの結果を見ることができます。



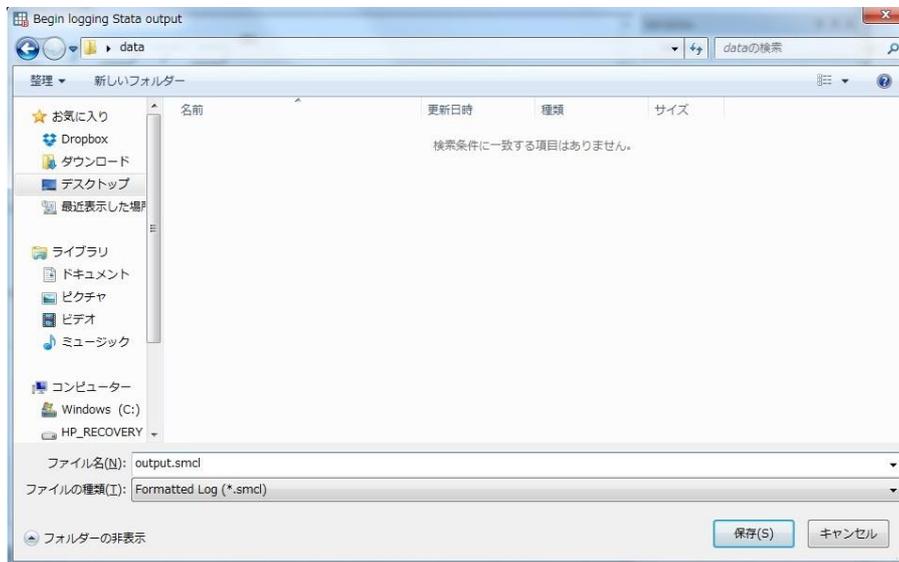
出力結果の保存法

推定結果の表が長くて、一度に画面上に出力出来ない場合には、出力結果の途中で「-more-」と表示されて止まることがあります。その時には、スペースキーをおせば、続きが表示されます。

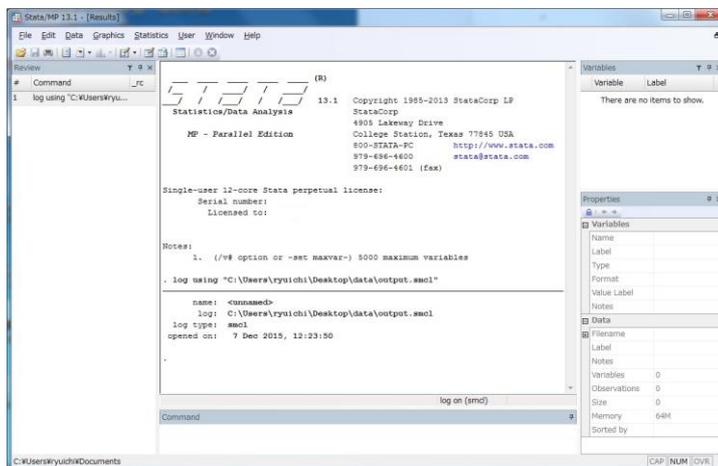
出力結果を保存したい場合には、出力結果が表示されるウィンドウで結果を選択して、他のファイルにペーストすることができます。もうひとつ別の出力結果の保存法として、ログファイルを作成することができます。

- ▼ 「File」→「Log」→「Begin」と移動し、「Begin」をクリックします（またはのアイコンをクリックします）。するとログファイルの保存先と名前をきかれますので、適

宜場所と名前を指定してください（ここでは先ほどデータを保存したフォルダ「data」に“output”という名前のログファイルを作ってみましょう）。

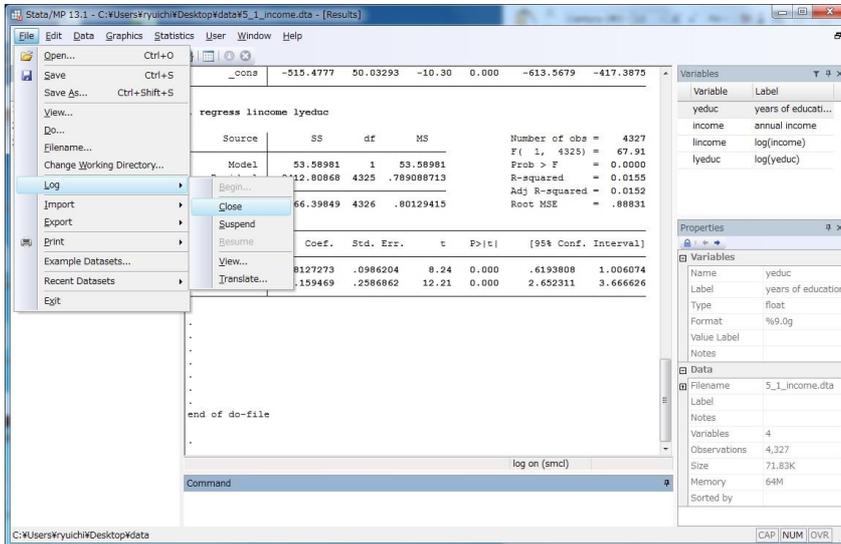


- ▼ メインのウィンドウに次のように表示されていれば、ログファイルの作成が始まっています。

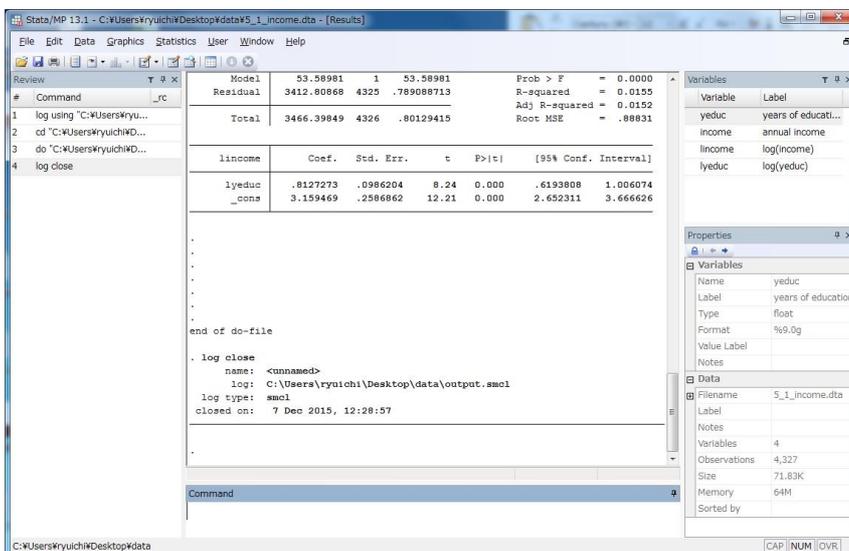


- ▼ 共通の準備も含めて、いずれかの方法で分析を進めます。

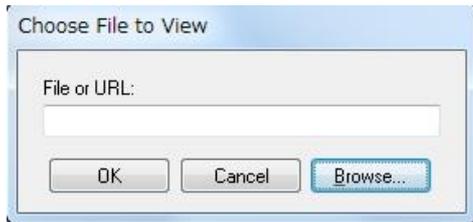
- ▼ 分析が終了したら、「File」→「Log」→「Close」と移動し、「Close」をクリックします（またはのアイコンをクリックします）。



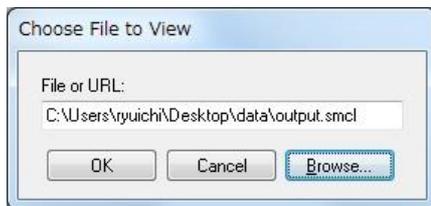
- ▼ 次の画面が表示されれば、ログファイルの作成が終了しています。



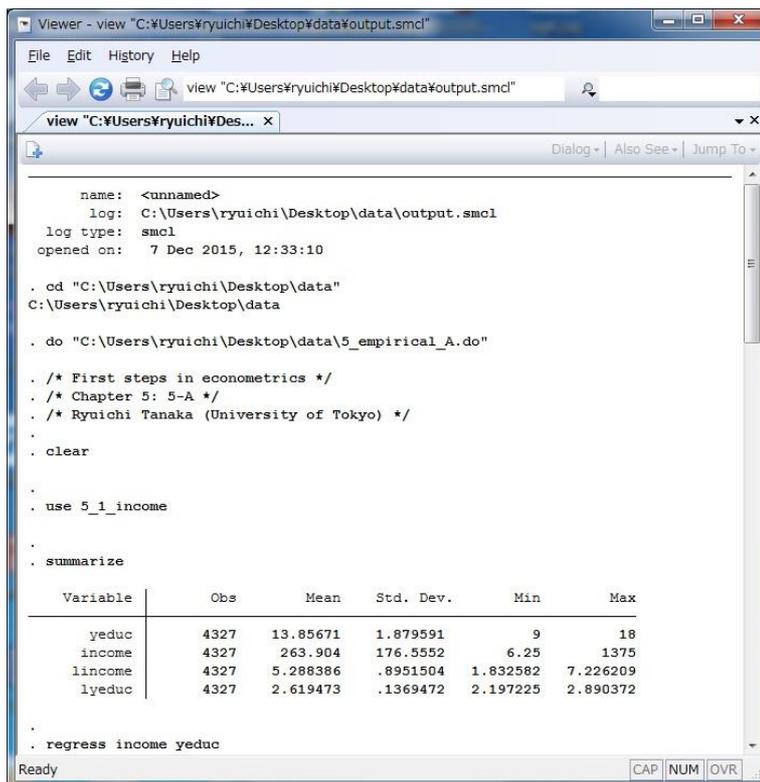
- ▼ ログファイルを開いて、結果がきちんと保存されているか確認してください。ログファイルを確認するためには、「File」→「Log」→「View」と選択して、先ほど作成したログファイルを選択します。ログファイルを選択するためには、次に開くウィンドウ



で「Browse...」をクリックして、別のウィンドウからログファイルのあるフォルダを参照して探します。ログファイルを見つけたら、「開く (O)」とすると次のようになります。



- ▼ ここで「OK」をクリックするとログファイルを見ることができます。



参考までに、Stata の入手先のウェブページは以下の通りです。

© StataCorp 社 (英語)

<http://www.stata.com>

© ライトストーン社

<http://www.lightstone.co.jp/stata/>

その他の Stata の使い方については、まずは google などを使って、「stata 入門」などと検索をしてみてください。すると、Stata の使い方を紹介した日本語の文書や資料がダウンロードできるページがたくさん出てきます。そのような中でも、日本語による詳細な入門的解説に優れている愛知学院大学の野村友和先生の資料を紹介しておきます。

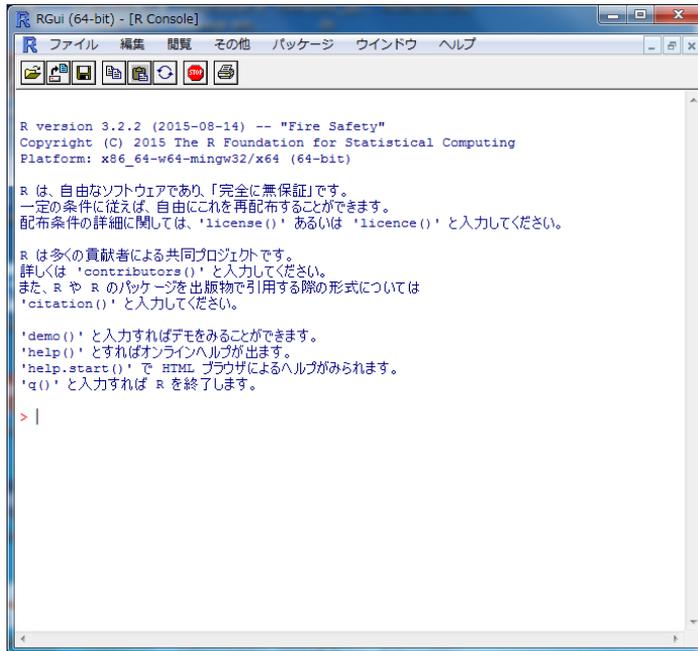
<http://www.agu.ac.jp/~nomura/lecture/archive/stata.pdf>

2. R (アール)

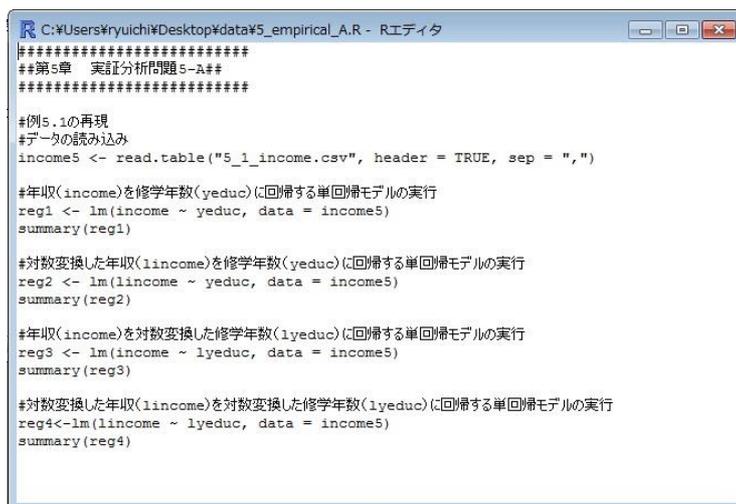
R は無料でダウンロードできる統計分析用のソフトウェアです。自分でプログラムを書く量が Stata よりは多いですが、分析用のパッケージがたくさん用意されていて、これらのパッケージを使うと様々な計量経済学的分析が可能になります。R の入手法やインストールについては、ウェブで「R ダウンロード」と検索すれば簡単に見つけることができます。

このチュートリアルでは、第 5 章の実証分析問題 5-A を使って、Windows での R の使い方を簡単に説明します。なお Windows の場合は、本書で用意しているスクリプト・ファイルのうち、文字コードが Shift-JIS となっているほうを利用してください。

- ▼ データとスクリプトファイルの準備：教科書のホームページにアクセスして、データ「5_1_income.csv」と R のスクリプト「5_empirical_A.R」をダウンロードして、保存します。保存先と名前はなんでも構いませんが、ここではデスクトップに新しいフォルダ“data”を作っておきます。
- ▼ R の起動：デスクトップに R のアイコンがあれば、それをダブルクリックすることで R が起動して、次の画面が現れます（もしアイコンがない場合には、スタートメニューから R を立ち上げてください）。R が起動すると、次の画面が表示されます。



- ▼ ディレクトリの変更：「ファイル」→「ディレクトリの変更」とすると、「スクリプトを開く」ウィンドウが開きますので、データとスクリプトのファイルを保存したデスクトップ上のフォルダ「data」を選択してください。
- ▼ スクリプトを開く：「ファイル」→「スクリプトを開く...」をクリックすると、「フォルダーの参照」ウィンドウが開きますので、「5_empirical_A.R」を選んで、「開く(O)」をクリックしてください。すると「R エディタ」にスクリプトの内容が表示されます。



- ▼ スクリプトを実行する：「編集」→「全て実行」をクリックすると、スクリプト内の全てのコマンドが実行され、分析結果が「R Console」ウィンドウに表示されます。

```

R Console

lm(formula = income ~ lyeduc, data = income5)

Residuals:
    Min       1Q   Median       3Q      Max
-321.25 -119.73  -9.46   90.54 1177.02

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -515.48     50.03  -10.3  <2e-16 ***
lyeduc         297.53     19.07   15.6  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 171.8 on 4325 degrees of freedom
Multiple R-squared:  0.05326, Adjusted R-squared:  0.05304
F-statistic: 243.3 on 1 and 4325 DF, p-value: < 2.2e-16

>
> #対数変換した年収(lincome)を対数変換した修学年数(lyeduc)に回帰する単回帰
> reg4<-lm(lincome ~ lyeduc, data = income5)
> summary(reg4)

Call:
lm(formula = lincome ~ lyeduc, data = income5)

Residuals:
    Min       1Q   Median       3Q      Max
-3.6295 -0.3400  0.2172  0.5786  2.1179

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.15946     0.25869  12.213 < 2e-16 ***
lyeduc         0.81273     0.09862   8.241 2.24e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8883 on 4325 degrees of freedom
Multiple R-squared:  0.01546, Adjusted R-squared:  0.01523
F-statistic: 67.91 on 1 and 4325 DF, p-value: 2.241e-16

> |

```

◎Rのダウンロードサイト

なお、Rは以下のサイトからダウンロードできます。

<https://cran.r-project.org/>

◎パッケージのインストール

第7章（重回帰分析の応用）や第8章（操作変数法）、第9章（パネル・データ分析）のスクリプトを実行するには、拡張機能を実行するためのパッケージを新たにインストール

する必要があります(第7章では“lmtest”, “sandwich”, 第8章では“AER”, 第9章では“plm”など)。本書のスクリプトファイルでは `install.packages()` 関数で上記パッケージをインストールして実行するように記載されていますが, 詳しくは R について解説した参考書やウェブサイトを参照してみてください。

◎より便利に R を使うソフト：RStudio

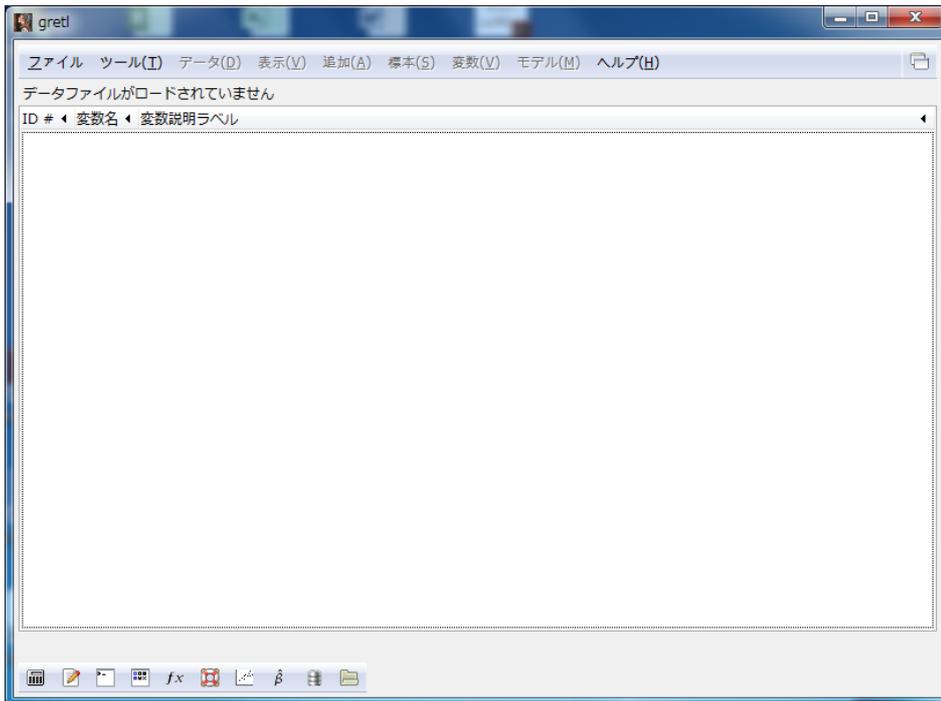
R は, R を実行するための統合開発環境である RStudio というソフトを活用すると, より便利に使うことができます。R をインストールしたうえで, 以下のサイトからダウンロードすると RStudio で R を実行できます。RStudio を用いた解説書や解説サイトもたくさんありますので, 詳しくはそちらを参照してみてください。

<https://www.rstudio.com/>

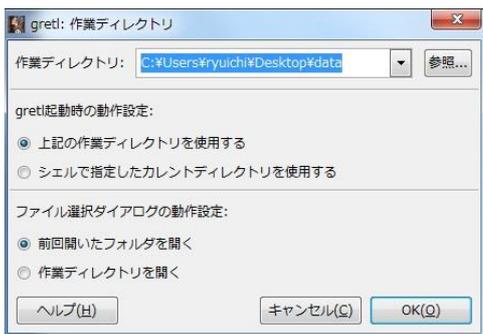
3. gretl (グレーテル)

gretl は無料でダウンロードできる統計分析用のソフトウェアです。gretl を使って, コマンドを選択しながら実証分析問題を解く方法については, 詳細な解答例を別に用意してありますので, そちらを参照してください。このチュートリアルでは, 実証分析問題 5-A を, スクリプトファイルを使って分析する方法を簡単に紹介します。

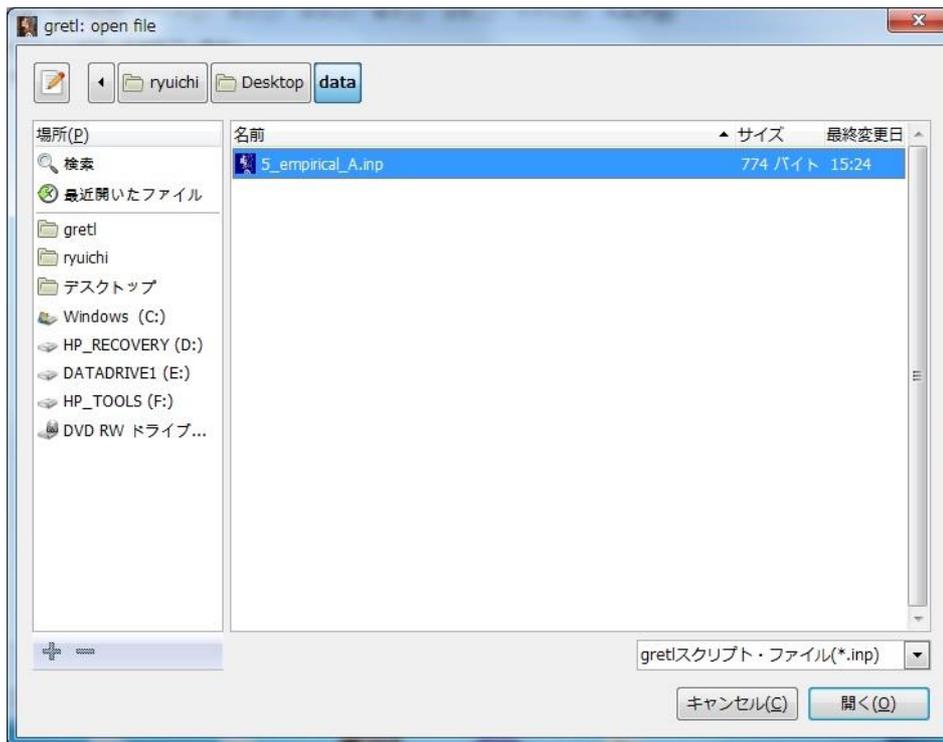
- ▼ データとスクリプトファイルの準備: 教科書のウェブサポートページにアクセスして, データ (拡張子が「.csv」のファイル) と gretl のスクリプト (拡張子「.inp」) のファイルをダウンロードしましょう。これらのファイルの保存先として, デスクトップに新しいフォルダ (名前はなんでも構いませんが, ここでは“data”という名前にしておきます) を作っておきましょう。
- ▼ gretl の起動: デスクトップにアイコンがあれば, それをダブルクリックすることで gretl が起動して, 次の画面が現れます (もしアイコンがない場合には, スタートメニューから gretl を立ち上げてください)。



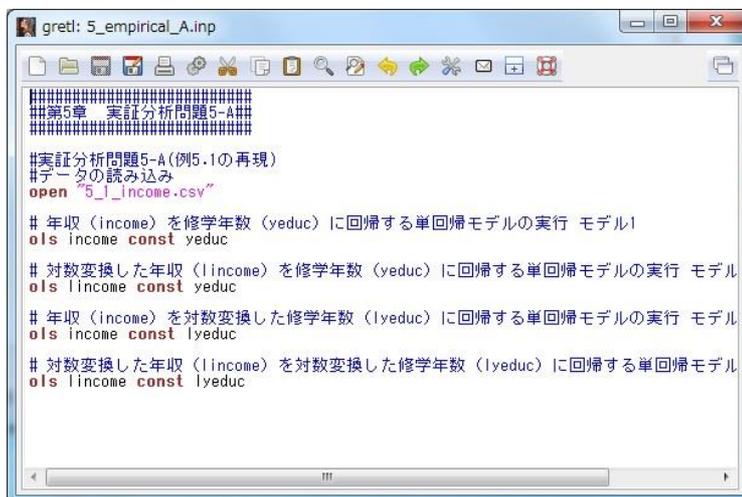
- ▼ 作業ディレクトリの指定: データとスクリプトのファイルを保存したフォルダを作業ディレクトリとして指定するためには、「ファイル」→「作業ディレクトリ (W...)」をクリックします。すると次の画面が表示されるので、作業ディレクトリのところにデータとスクリプトファイルが保存されているフォルダになっているかを確認してから「OK(O)」をクリックしてください。



- ▼ スクリプトの実行: 「ファイル」→「スクリプト・ファイル (S)」→「ユーザー・ファイル (U) ...」をクリックすると、作業ディレクトリにあるスクリプトファイルの一覧を表示します。



- ▼ 実行したいファイルを選んで、「開く (O)」をクリックすると、そのスクリプトファイルに書かれているコマンドが別ウィンドウで表示されます。



- ▼ コマンドを実行するためには、コマンドの表示されているウィンドウにある実行アイコンをクリックします。次の「gretl:スクリプト実行結果」の画面が表示され、すべて

の分析結果が出力されていることを確認してください。

```

gretlバージョン 2015d
現在のセッション: 2015-12-07 15:39

#####
###第5章 実証分析問題5-A###
#####
#実証分析問題5-A(例5.1の再現)
#データの読み込み
? open "5_1_income.csv"
処理中: C:\Users\Ryuichi\Desktop\data\5_1_income.csv...
区切り記号として','を使用します
最長の行には、28 文字あります
第1フィールド: 'yeduc'
列の数: 4
変数の数: 4
行の数(空行を除く): 4328
変数名を探しています...
行: yeduc,income,lincome,lyeduc
行ラベルと日付を探しています...
これらのデータを日付なしデータとして扱います

5 個の変数をリストします:
0) const 1) yeduc 2) income 3) lincome 4) lyeduc

# 年収 (income) を修学年数 (yeduc) に回帰する単回帰モデルの実行 モデル1
? ols income const yeduc

モデル 1: 最小二乗法(OLS), 観測: 1-4327
従属変数: income

```

◎gretl のダウンロードサイト

なお、gretl は以下のサイトからダウンロードできます。

<http://gretl.sourceforge.net/>

4. Excel (エクセル)

Excel を使って分析してみたいという方は、Excel にアドインとして備わっている「分析ツール」を使えば第 8 章の操作変数法以外の分析はできます。操作変数法による推定も、少々面倒ではありますが手動でできます。Excel の使い方については、インターネットで「excel 回帰分析」とでも検索すれば、詳しく紹介したウェブページが大量に見つかりますので、自分に合ったものを探してみてください。本書のウェブページからダウンロードできるデータのファイルで、拡張子が「.csv」のものは全てエクセルで開くことができます。