

第4回：度数分布とヒストグラムの 確認

北村 友宏

2020年5月29日

本日の内容

1. 度数分布とヒストグラム

度数とヒストグラム

- ▶ ある変数の、ある階級 (class) に属する観測値数を、その階級の度数 (frequency) という。
 - ▶ e.g., 1ヶ月間の可処分所得が20万円以上25万円未満の観測値数が10なら、その階級の度数は10.
 - ▶ gretlでは、「階級」は「区間」として表示される.
- ▶ 観測値の総数に占めるその階級の度数の割合を相対度数 (relative frequency) という。
 - ▶ e.g., 1ヶ月間の可処分所得の観測値の総数が100で、20万円以上25万円未満の観測値数が10なら、その階級の相対度数は0.1.
 - ▶ gretlでは、「階級」は「区間」として表示される.
- ▶ 各階級とその度数を表したグラフをヒストグラム (柱状図, histogram) という。
 - ▶ 横軸はその変数の値.

- ▶ gretl では、以下の項目を設定して度数分布とヒストグラムを出力できる。
 - ▶ 階級の数または
 - ▶ 左端の階級の下限值（変数の、グラフをスタートさせる値）と階級幅
- ▶ 初期状態ではこれらの項目が自動的に設定されているが、ほとんどの場合、「左端の階級の下限值と階級幅」が半端な値になる。



- ▶ 階級幅と最初の階級の最小値を設定すると、分かりやすいヒストグラムを描くことができる。
 - ▶ e.g., 最初の階級の最小値は 10 万円, 階級幅は 5 万円, 等.

- ▶ 実習として，可処分所得（千円単位）のヒストグラムを出力する．
- ▶ 前回の授業で記述統計を出力した際，最小値が約 14.5 万円であった．

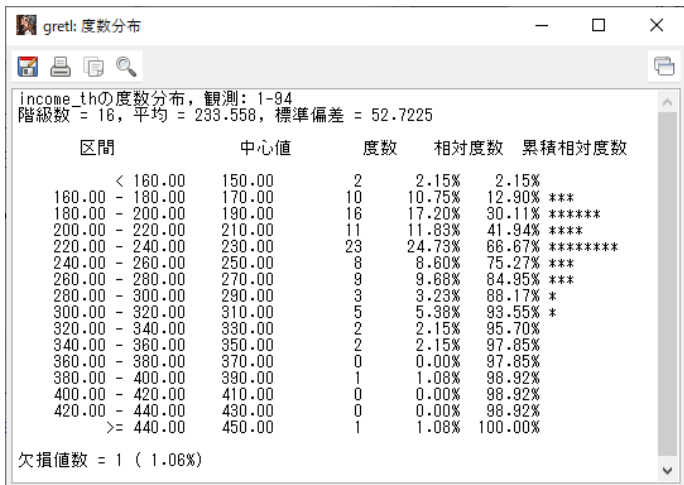


最初の階級の最小値を 14 万円とする．

実習 1

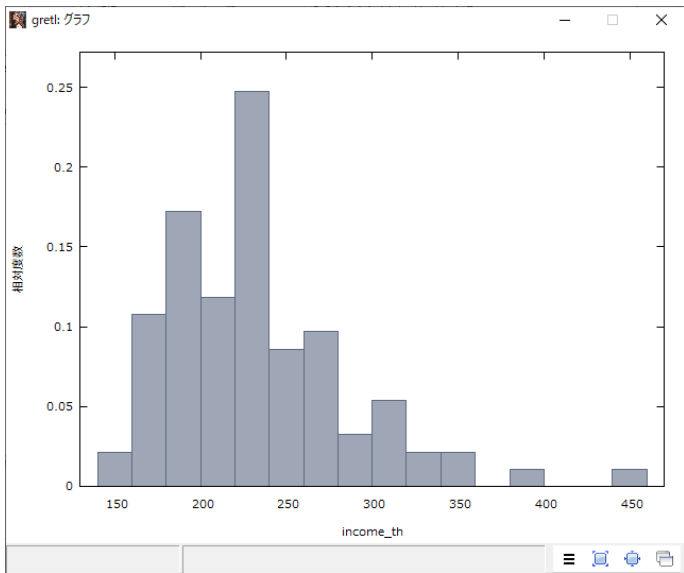
1. gretl を起動.
2. 「ファイル」 → 「データを開く」 → 「ユーザー・ファイル」と操作.
3. 消費 2009.gdt を選択し、「開く」をクリック.
4. 「income_th」を左クリックして選択し、その上で右クリック → 「度数分布」と操作.
5. 「gretl: 度数分布」ダイアログボックスの、「左端の階級の下限值:」をクリックして選択し、その右の入力ボックスに 140.000 と入力.
 - ▶ 最初の階級の最小値が 140 千円 (14 万円) となる.
6. 「階級幅:」の右の入力ボックスに 20.000 と入力.
 - ▶ 階級幅が 20 千円 (2 万円) となる.

7. 「グラフを表示」にチェックが入っていることを確認。入っていなければクリックしてチェックを入れる。
8. 「OK」をクリックすると、千円単位の可処分所得（変数名：income_th）の度数分布とヒストグラムが表示される。



度数分布については、このような画面が表示されれば成功。

まだ作業があるので、「gretl: 度数分布」のウィンドウは**まだ閉じない!**



ヒストグラムについては、このような画面が表示されれば成功。 **まだウィンドウを閉じない！**

9. 表示されている度数分布表（「gretl: 度数分布」のウィンドウ）の画面上で右クリック→「名前を付けて保存...」と操作.
10. 出てきたダイアログボックスの、「ファイルに保存」を選び、「OK」をクリック.
11. 可処分所得度数分布.txt という名前で「2020 ミクロデータ分析 1」フォルダに保存.
12. 表示されているヒストグラム（「gretl: グラフ」のウィンドウ）の画面上で右クリック→「PNGとして保存...」と操作.
13. 可処分所得ヒストグラム.png という名前で「2020 ミクロデータ分析 1」フォルダに保存.
 - ▶ ファイルの種類を png として保存すれば、レポートや卒業論文を Word 等で作成する際に、ヒストグラムを Word 等に貼り付けることができる.

出力した度数分布・ヒストグラムから分かること

2009年における単身勤労世帯の1ヶ月間の可処分所得の都道府県別・男女別平均は、

- ▶ 16万円から28万円にかけて観測値が集中している。
- ▶ 22万円から24万円の階級（区間）に属している観測値が最も多い（相対度数24.73%）。
- ▶ **大きく外れた観測値（40万円を超える）が存在する！**



次回の授業で、大きく外れているのはどの都道府県・性別かを突き止める。

本日の作業はここまで.

今回は gretl のデータセットに変更を加えていない
ので, **gretl のデータセット (消費 2009.gdt)** を上書き
保存する必要はない.