

# 第 5 回：記述統計表の作成

北村 友宏

2020 年 10 月 30 日

# 本日の内容

## 1. 記述統計表の作成

# 実習 1

前回出力した，中古マンション価格，最寄り駅までの所要時間，築年数，面積，1K1R ダミーの記述統計を再び出力する．

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. setagayaapartment.gdt を選択し，「開く」をクリック．

4. Ctrl キーを押しながら「minutes」, 「area」, 「onekr」, 「price\_10th」, 「age」の5つを左クリックして選択し, その上で右クリック→「基本統計量」と操作.
5. 「**全ての統計量を表示する**」を選んで状態で「OK」をクリックすると, 選択した変数の記述統計12種類が表示される.
  - ▶ **最新バージョン (2020年8月6日版) では, この表示が日本語化されている.**

	平均	中央値	最小値	最大値
minutes	8.9845	8.0000	0.00000	29.000
area	53.531	50.000	10.000	280.00
onekr	0.19072	0.00000	0.00000	1.0000
price_10th	3762.6	3600.0	500.00	19000
age	15.201	11.500	1.0000	43.000

  

	標準偏差	変動係数	歪度	過剰尖度
minutes	5.4130	0.60248	0.79320	0.37659
area	29.115	0.54390	2.5473	17.388
onekr	0.39389	2.0652	1.5745	0.47891
price_10th	2151.0	0.57167	1.9703	11.370
age	11.406	0.75035	0.77453	-0.48386

  

	5百分位数	95百分位数	10幅	欠損値数
minutes	1.7500	19.000	7.0000	0
area	15.000	95.000	35.000	0
onekr	0.00000	1.0000	0.00000	0
price_10th	820.00	7000.0	2525.0	0
age	1.0000	38.000	18.250	0

このような画面が表示されれば成功。

Mac の PC では、小数点以下の表示桁数が異なっている場合がある。

最新バージョン（2020年8月6日版）では、上の画像のように統計量名が全て日本語で表示される。

- ▶ 統計量の名前の位置がズレていて見づらいが、各変数について出力された数字は、1 段目は左から平均，中央値，最小値，最大値の順。2 段目は左から標準偏差，変動係数，歪度，過剰尖度の順。3 段目は左から 5 百分位数，95 百分位数，IQ 幅（四分位範囲），欠損値数の順。

# レポート・論文用の記述統計表の作成

見やすく，理解しやすい表を載せるには，

- ▶ 表番号と表のタイトルをつける。
- ▶ 変数名は統計解析ソフトでの変数名そのままではなく，分かりやすいように書き直す。
- ▶ 単位がある変数は単位を明記する。
- ▶ 小数の数値はあまり細かく表示せず，小数第1～4位程度まで示せば十分。縦方向に見たとき，小数点の位置が揃うようにする。
- ▶ 観測値数が全変数について同じ場合，表の下に「注（Note）」として「観測値数は\*\*\*である」などと注記してもよい。

この授業では Word を用いた記述統計表の作成方法を説明する。

## 実習 3

1. Word を起動し，summary20201030.docx という名前で 2020microdatag フォルダに保存.
2. 「挿入」→「表」と操作して 6 行 5 列の表を作る.
3. 表全体をドラッグし，「参考資料」（Mac 版では「参照設定」）→「図表番号の挿入」と操作.
4. ラベルを「表」に，位置を「選択した項目の上」にして OK をクリックすると，表のすぐ上の行に「表 1」と入力される.
  - ▶ ラベルに「表」がなければ，「新しいラベル...」をクリックして出てくるダイアログボックスの入力ボックスに表と入力して OK をクリック.
5. 「表 1」の後に全角スペースを入れて記述統計と入力し，中央揃えにする.



6. 表の1行2列目に平均値，1行3列目に標準偏差，1行4列目に最小値，1行5列目に最大値と入力。
  - ▶ 今回は，教科書『新しい計量経済学』p.13に載っている表1.3に合わせ，中央値はレポートしない。
7. 表の1行2列目から1行5列目までをドラッグし，「レイアウト」タブ（右端の，色が濃いほう）から「配置」→「中央揃え」と操作。
8. 表の2行1列目に価格（万円），3行1列目に最寄り駅所要時間（分），4行1列目に築年数（年），5行1列目に面積（平方メートル），6行1列目にワンルームダミーと入力。

9. 表の2行1列目から6行1列目までをドラッグし、「レイアウト」タブ（右端の、色が濃いほう）から「配置」→「中央揃え」と操作.
10. gretl に出力されていた表の数値を、Word で作成した表の対応するセルにコピー・貼り付けする. 数値をドラッグして選択し、右クリック→「コピー」と操作すればコピーできる.
  - ▶ price\_10th は価格（万円）、minutes は最寄り駅所要時間（分）、age は築年数（年）、area は面積（平方メートル）、onekr はワンルームダミー、平均は平均値.

11. 平均値と標準偏差の数値を全て、小数第2位までにする（小数第3位を四捨五入）。例えば、最寄り駅所要時間（分）の平均値と築年数（年）の標準偏差は、
- ▶ 8.9845⇒8.98
  - ▶ 29.115⇒29.12
  - ▶ こうすることで、各列で小数点以下の表示桁数を統一し、縦方向に見たとき、小数点の位置が揃うようにする。
13. 表の2行2列目から6行5列目まで（数値の入っているセル）をドラッグし、「レイアウト」タブ（右端の、色が濃いほう）から「配置」→「中央揃え（右）」と操作。
14. Wordで作成した表のすぐ下の行に、（注）観測値数は194である。と入力して上書き保存。
- ▶ 観測値数は、gretlの画面下部の「日付なし：全範囲 1 - 194」と書かれている部分を見れば分かる。

The image shows a Microsoft Word document with a table titled "表 1 記述統計". The table contains five rows of data with five columns: a variable name, mean, standard deviation, minimum, and maximum. Below the table is a note: "(注) 観測値数は 194 である。".

	平均値	標準偏差	最小値	最大値
価格 (万円)	3762.60	2151.00	500	19000
最寄り駅所要時間 (分)	8.98	5.41	0	29
築年数 (年)	15.20	11.41	1	43
面積 (平方メートル)	53.53	29.12	10	280
ワンルームダミー	0.19	0.39	0	1

(注) 観測値数は 194 である。

このような表を作成できればよい。

# 補足

- ▶ ダミー変数を分析に用いる場合，そのダミー変数の定義をレポートや論文の本文中などで説明しておく。
  - ▶ e.g., ワンルームダミーはマンションの種類が1Kまたは1Rであれば1，そうでなければ0をとるダミー変数である。

本日の作業はここまで。