

<ローレンツ曲線とジニ係数>

1. 準備

今回も、実習にはあらかじめ河田が作成したファイルを用いる。

【課題 16】 統計学基礎の講義用 HP から、年収階級分布 2016 のファイルをダウンロードしてみよう。

📖 手順

- ① 検索エンジンで、「河田研究室」と入力し検索すると、「河田研究室」のページにジャンプする。(ここまでの手順は、<http://www2.tokuyama-u.ac.jp/kawada> とアドレスを直接入力してもよい)
- ② 「統計学基礎」をクリックし、「第14回 11月16日(木)」の配布資料にある、「年収階級分布 2016」をクリックし、自分の使いやすい場所に保存する。

2. ローレンツ曲線の描画

【課題 17】 2016 年の年収階級分布について、ローレンツ曲線を描いてみよう。

ローレンツ曲線の横軸は累積世帯比率、縦軸は累積年収比率となるので、最初にそれらを計算する。

- (1) 累積世帯比率の導出：累積世帯比率は、累積世帯数を求め、世帯数の合計で割ることによって求められる。

📖 手順

① まずD3セルからD20セルに入力されている、世帯数のデータの合計をD21セルに求める $=\text{sum}(D3:D20)$ とすれば、世帯数の合計を計算できる。

② 次にE列に累積世帯数を求める。

i. E3セルに $=D3$ と入力する。

ii. E4セルに $=E3+D4$ と入力し、それをE5セルからE20セルまでにコピーする。

これで、累積世帯数が計算される。

③ 最後にF列に累積世帯比率を計算する。F3セルに $=E3/D21$ と入力し、**F4** キーを押す¹(すると $=E3/\$D\21 となるはず)。これをF4:F20セルにコピーする。

- (2) 累積年収比率の計算：この表にあるデータは、各階級の平均年収である（たとえば、G3セルの 153 は、年収200万未満の229世帯の平均が153万円であることを意味している）。最初に各階級の年収の合計を求め、累積年収、累積年収比率を求めていく。

¹ これはセルを絶対参照にする操作である。Excel は通常そのセルとの相対的な位置関係に基づく相対参照となっており、コピーした場合などは、セルの中身が変化していく。しかし絶対参照では変化しない。私の家の2軒右隣りに酒屋さんがあるとする。これを他人に説明する場合、「私の家の2軒右隣り」というのが相対参照で、酒屋の住所(周南市城ヶ丘〇丁目〇番地)というのが絶対参照である。

📖 手順

- ① H列に各階級の年収の合計を求める。各階級の年収の合計は「世帯数」×「平均年収」で求められるので、H3セルに $=D3*G3$ と入力し、これをH4:H20セルにコピーする。そして、全階級の年収の合計をH21セルに求める。(sum関数を用いる)
- ② 次にI列に累積年収を求める。これは累積世帯数の場合と同様であり、
 - i. I3セルに $=H3$ と入力する。
 - ii. I4セルに $=I3+H4$ と入力し、それをI5:I20セルにコピーする。
 これで、累積年収が計算される。
- ③ 最後にJ列に累積年収比率を計算する。J3セルに $=I3/H21$ と入力し、**F4**キーを押す(すると $=I3/\$H\21 となるはず)。これをJ4:J20セルにコピーする。

(1)、(2)により下図のように計算される。

	A	B	C	D	E	F	G	H	I	J
1	階		級	集計世帯数	累積世帯数	累積世帯比率	年間収入	年収総額	累積年収総額	累積年収比率
2						0				0
3		-	200	229	229	0.02970	153	35037	35037	0.00758
4	200	-	250	372	601	0.07795	226	84072	119109	0.02578
5	250	-	300	550	1151	0.14929	276	151800	270909	0.05863
6	300	-	350	765	1916	0.24851	324	247860	518769	0.11226
7	350	-	400	691	2607	0.33813	373	257743	776512	0.16804
8	400	-	450	619	3226	0.41842	422	261218	1037730	0.22457
9	450	-	500	557	3783	0.49066	474	264018	1301748	0.28170
10	500	-	550	504	4287	0.55603	523	263592	1565340	0.33874
11	550	-	600	440	4727	0.61310	572	251680	1817020	0.39321
12	600	-	650	409	5136	0.66615	621	253989	2071009	0.44817
13	650	-	700	345	5481	0.71089	673	232185	2303194	0.49842
14	700	-	750	339	5820	0.75486	722	244758	2547952	0.55138
15	750	-	800	288	6108	0.79222	772	222336	2770288	0.59950
16	800	-	900	460	6568	0.85188	843	387780	3158068	0.68341
17	900	-	1000	316	6884	0.89287	944	298304	3456372	0.74797
18	1000	-	1250	440	7324	0.94994	1099	483560	3939932	0.85261
19	1250	-	1500	172	7496	0.97224	1357	233404	4173336	0.90312
20	1500	-		214	7710	1.00000	2092	447688	4621024	1.00000
21			計	7710				4621024		

累積世帯比率(F列)を横軸に、累積年収比率(J列)を縦軸にとった散布図を描き、線をつないだものがローレンツ曲線となる。なお、列の先頭に0をいれてある。

📖 手順

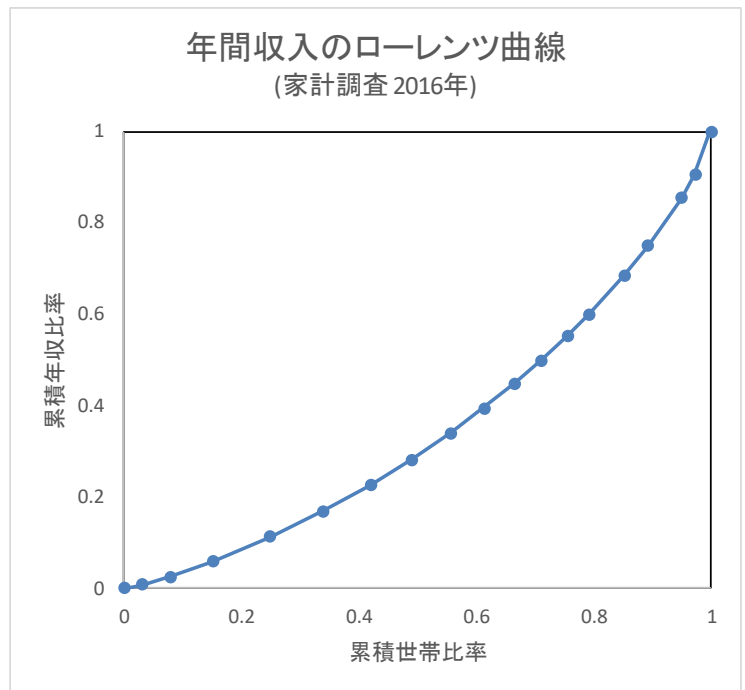
- ① 最初にグラフに描く範囲を範囲指定する。F2:F20をドラッグし、**Ctrl**キーを押しながらJ2:J20をドラッグして範囲指定する。
- ② グラフを作成するには、**挿入タブ**をクリックすることで、リボン内にグラフのグループが表示される。ここでは、散布図のボタンをクリックする。
- ③ すると散布図グラフのフォーマット(型式)メニューが出るので、左下の散布図(直線とマーカー)をクリックする。
- ④ この時点でグラフのサンプルが自動的に描かれている。これを修正していく。まずレイアウトを変更する。リボンの中の「グラフのレイアウト」を展開し、レイアウト1(左上)をクリックし、タイトルや軸ラベルなどが書き込めるようにする。そして、

タイトル:「年間収入のローレンツ曲線 (家計調査2016年)」と記入する。

X 軸:「累積世帯比率」と記入する。

Y 軸:「累積年収比率」と記入する。

- ⑤ さらにいくつかの細かい修正を加えたものが右図である。右図のようにするには、
- i. 凡例の消去
 - ii. 「軸の書式設定」において、最大値の変更
 - iii. 目盛線の消去
 - iv. 「プロットエリアの書式設定」において、「枠線」を単色にする。
 - v. タイトルのフォント変更
 - vi. グラフの大きさ変更 (プロットエリアを正方形に近づける)



3. ジニ係数の導出

【課題 18】 2016 年の年収階級分布について、ジニ係数を求めてみよう。

ジニ係数を計算するには、既に述べたように台形の面積を足し合わせたものを 2 倍して、正方形の面積 1 から引けばよい。

ここで、台形の面積は $(\text{上底} + \text{下底}) \times \text{高さ} \div 2$ という公式で求まるが、

上底 … 1つ前の階級までの累積年収比率

下底 … その階級までの累積年収比率

高さ … その階級までの累積世帯比率から 1つ前の階級までの累積世帯比率を引いたものとなる。

K 列に、台形の面積を求め、その 2 倍の合計を 1 から引くことによって、ジニ係数を求める。

📖 手順

- ① K3セルに、 $= (J2+J3) * (F3-F2) / 2$ とする(上底がJ2、下底がJ3、高さがF3-F2である)。
- ② K3をK4:K20セルにコピーし、K21セルに合計を求める。
- ③ K23セルに、 $= 1 - K21 * 2$ と入力する。これがジニ係数である。

📁 演習問題

1970 年のワークシートも同様に、ローレンツ曲線を描き、ジニ係数を求めてみよう。

📁 本日実習したファイルは、河田まで提出すること。(1970年のものもおこなうこと)

1. Webclass 経由

2. E-mail に添付ファイルとして(送付先アドレスは kawada@tokuyama-u.ac.jp)のいずれでもかまわない。

締め切りは 11 月 27 日(月)9:15 とする。

なお、ファイル名は「年収階級分布 2016 E47-○○○」のように、学籍番号をつけること。