

## &lt;ローレンツ曲線とジニ係数&gt;

**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、年収階級分布のファイルをダウンロードしてみよう。

## 📖 手順

- ① 徳山大学のトップページ (<http://www.tokuyama-u.ac.jp>) で、「在学生の皆様へ」のメニューから、「教員紹介」を選び、河田の教員紹介を開く。
- ② そのページの下の方にある、「個人サイト」のところをクリックすると、「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)
- ③ 「統計学基礎」をクリックし、「第15回 11月21日(月)」の配布資料にある、「年収階級分布」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】 2010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<sup>1</sup>(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158は、年収200万未満の218世帯の平均が158万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。

<sup>1</sup> 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列に累積世帯比率を計算する。J3セルに  $=I3/H21$  と入力し、**F4**キーを押す(すると  $=I3/SH$21$  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	218	218	0.02788	158	34444	34444	0.00725
4	200	-	250	344	562	0.07189	226	77744	112188	0.02361
5	250	-	300	485	1047	0.13392	275	133375	245563	0.05169
6	300	-	350	683	1730	0.22128	323	220609	466172	0.09812
7	350	-	400	679	2409	0.30814	373	253267	719439	0.15143
8	400	-	450	685	3094	0.39575	423	289755	1009194	0.21242
9	450	-	500	588	3682	0.47096	473	278124	1287318	0.27096
10	500	-	550	528	4210	0.53850	524	276672	1563990	0.32920
11	550	-	600	484	4694	0.60041	573	277332	1841322	0.38757
12	600	-	650	438	5132	0.65643	621	271998	2113320	0.44482
13	650	-	700	372	5504	0.70402	672	249984	2363304	0.49744
14	700	-	750	316	5820	0.74444	721	227836	2591140	0.54539
15	750	-	800	269	6089	0.77884	773	207937	2799077	0.58916
16	800	-	900	491	6580	0.84165	844	414404	3213481	0.67639
17	900	-	1000	345	6925	0.88578	945	326025	3539506	0.74501
18	1000	-	1250	496	7421	0.94922	1101	546096	4085602	0.85996
19	1250	-	1500	196	7617	0.97429	1360	266560	4352162	0.91606
20	1500	-		201	7818	1.00000	1984	398784	4750946	1.00000
21			計	7818				4750946		

累積世帯比率(F列)を横軸に、累積年収比率(H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

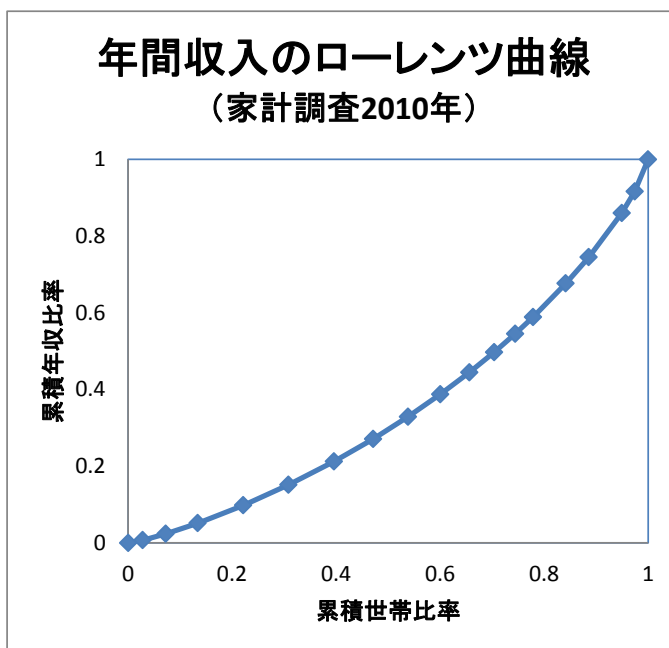
- ① 最初にグラフに描く範囲を範囲指定する。F2:F20をドラッグし、**Ctrl**キーを押しながらJ2:J20をドラッグして範囲指定する。
- ② グラフを作成するには、**挿入タブ**をクリックすることで、リボン内にグラフのグループが表示される。ここでは、散布図のボタンをクリックする。
- ③ すると散布図グラフのフォーマット(型式)メニューが出るので、左下の散布図(直線とマーカー)をクリックする。
- ④ この時点でグラフのサンプルが自動的に描かれている。これを修正していく。まずレイアウトを変更する。リボンの中の「グラフのレイアウト」を展開し、レイアウト1(左上)をクリックし、タイト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2010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】 2010 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 28 日(月)9:00 とする。

なお、ファイル名は「年収階級分布 E41-○○○」のように、学籍番号をつけること。