

<家計に関する統計>

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 4】 経済統計の講義用 HP から、家計調査の分析のファイルをダウンロードしてみよう。

📖 手順

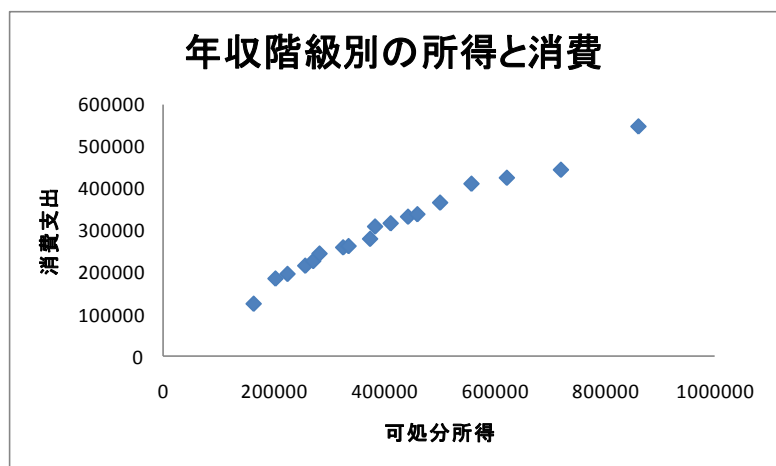
- ① 徳山大学のトップページ (<http://www.tokuyama-u.ac.jp>) で、「在学生の皆様へ」のメニューから、「教員紹介」を選び、河田の教員紹介を開く。
- ② そのページの下の方にある、「個人サイト」のところをクリックすると、「河田研究室」のページにジャンプする。(ここまでの手順は、<http://www2.tokuyama-u.ac.jp/kawada> とアドレスを直接入力してもよい)
- ③ 「統計学基礎」をクリックし、「第16回 11月29日(月)」の配布資料にある、「家計調査の分析」をクリックし、自分の使いやすい場所に保存する。

2. 散布図の描画

【課題 5】 年間収入階級別データの、可処分所得を横軸に、消費支出を縦軸にとり、散布図を描いてみよう。

📖 手順

- ① 2-6 表の可処分所得(K25:AB25)を「消費関数」のシートの D2:D19 セルに、消費支出(K14:AB14)を「消費関数」のシートの E2:E19 セルにそれぞれコピーする。貼り付ける際に、「形式を選択して貼り付け」を選び、「行列を入れ替える」にチェックをいれる。
- ② D2:E19 を範囲指定し、リボン内にグラフのグループにある、散布図 のボタンをクリックする。(挿入タブをクリックすることで表示される。)
- ③ 散布図グラフのフォーマット (型式) メニューにおいて、「散布図(マーカーのみ)」(左上) をクリックし、リボンの中のレイアウト 1(左端) 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。
- ④ 右の図のようにタイトル、軸ラベルを入力し、目盛線、凡例を非表示にする。



3. 統計関数による回帰直線の導出

Excelが備えている関数を用いて、回帰直線の導出を行ってみることにする。

(1) 回帰係数の導出（関数 SLOPE とINTERCEPT¹⁾）

$Y=a+bX$ という回帰モデルにおいて、**SLOPE** は回帰直線の傾き **b** を、**INTERCEPT** は切片 **a** を求める関数である。どちらも引数は2個あるが、最初の引数が従属変数(=の左側)の範囲で、2個目の引数が独立変数(=の右側)の範囲をとる。引数の順序に注意しなければならない。

(2) 予測値の導出（関数 FORECAST²⁾）

予測値 \hat{Y} として、関数**FORECAST** がある。FORECAST(X_i , 回帰のY範囲, 回帰のX範囲)として用いる。なお、この予測値は戻り値は $a + bx_i$ として求められた数値1個である。

関数**SLOPE**と**INTERCEPT**を用いて求めたa, bから、 $a + bx_i$ として求めた値と等しくなる。

【課題 6】消費支出を Y、可処分所得を X とした回帰分析 $Y=a+bX$ をおこない、予測値を求め、散布図を描きいれよう。

📖 手順

- ① 「消費関数」のシートの E21 セルに a の推定値を、E22 セルに b の推定値を求める。(D21 セルと D22 セルに a, b と記入しておこう) E21 セルに =INTERCEPT(E2:E19,D2:D19) と入力、E22 セルに =SLOPE(E2:E19, D2:D19) と入力する。
- ② F2 セルに =FORECAST(D2,\$E\$2:\$E\$19,\$D\$2:\$D\$19) と入力し、F3:F19 セルにコピーする。(F1 セルには 予測値 と記入しておこう)
- ③ **グラフをアクティブ**にした状態で、リボンの中の「データの選択」ボタン(出ていない場合には上部の「グラフツール」をクリックする)をクリックし、凡例項目(系列)の「追加」ボタンをクリックする。そして、「系列Xの値」を**D2:D19**、「系列Yの値」を**F2:F19**とする。
- ④ ③の操作で散布図上に赤色のマーカーが現れたはずである。これを直線で結ぶ。赤色のマーカーのひとつを右クリックして、「データ系列の書式設定」を選ぶ。そこで「マーカーのオプション」をクリックし、マーカーの種類として「なし」をチェックし、「線の色」をクリックし、線の色として「線 (単色)」をチェックする。グラフエリアの外をクリックすると回帰直線が引けたことがわかるはずである。

4. 分析ツールの利用

Excel には統計分析を行うためのいくつかの分析ツールが付属している。これらのツールを使えば一度に詳細な分析結果を得ることができる。分析ツールを最初に使用する場合には、アドイン(有効にすること)しなくてはならない。分析ツールのアドインは次のようにおこなう。

- ① 左上の **Office ボタン** を押し、下にある「Excel のオプション」のボタンをクリックする。
- ② 「分析ツール」にチェックをつけ、**OK ボタン** をクリックする。

すると、データタブの中に「データ分析」のボタンが出てくるので、下のほうにある、回帰分析を選べばよい。

¹ 回帰直線の傾きと切片を求める関数には、**LINEST** という関数がある。この関数は傾きと切片以外に分析結果に関する多くの情報量を与えてくれる、非常に便利な関数である反面、使用法および結果の解釈の仕方が難しい。LINEST 関数の説明はここでは省略する。

² FORECAST 関数と同様に予測値を求める関数として、**TREND** という関数があるが、ここでは説明は省略する

5. 統計関数をもちいたエンゲル関数の導出

【課題 7】 食料、住居、… といった十大費目について、各費目を E、消費支出を C とした回帰分析 $E=a+bC$ をおこない、エンゲル関数を求めよう。

📖 手順

- ① 2-6 表の消費支出(K14:AB14)を「エンゲル関数」のシートの D2:D19 セルに、各費目別の支出 (K15:AB24)を「エンゲル関数」のシートの E2:N19セルにそれぞれコピーする。貼り付ける際に、「形式を選択して貼り付け」を選び、「行列を入れ替える」にチェックをいれる。
- ② 最初に、食料のエンゲル関数について考える。「エンゲル関数」のシートの H22セルに a の推定値を、J22セルに b の推定値をそれぞれ求める。H22セルに $=\text{INTERCEPT}(E2:E19,D2:D19)$ と入力、J22セルに $=\text{SLOPE}(E2:E19, D2:D19)$ と入力する。
- ③ 住居についてのエンゲル関数も同様であるが、E2:E19の部分が F2:F19になる。このようにして十大費目のエンゲル関数の表を完成させよう。

6. ローレンツ曲線の描画

【課題 8】 2009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 2-6表の集計世帯数(K13:AB13)のデータを「ローレンツ曲線・ジニ係数」のシートのD3:D20セルにコピーする。貼り付ける際に、「形式を選択して貼り付け」を選び、「行列を入れ替える」にチェックをいれる。
- ② 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。
- ③ 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。
- ④ 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³(すると $=E3/DS21$ 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**：2-6表のデータは、各階級の平均年収である（たとえば、K26セルの 158 は、年収200万未満の36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

³ 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。

年収の合計を求め、累積年収、累積年収比率を求めていく。

📖 手順

- ① 2-6表の年間収入(K26:AB26)のデータを「ローレンツ曲線・ジニ係数」のシートのG3:G20セルにコピーする。貼り付ける際に、「形式を選択して貼り付け」を選び、「行列を入れ替える」にチェックをいれる。
- ② 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)
- ③ 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/\$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	36	36	0.00849	158	5688	5688	0.00192
4	200	-	250	97	133	0.03136	226	21922	27610	0.00933
5	250	-	300	113	246	0.05801	275	31075	58685	0.01983
6	300	-	350	194	440	0.10375	324	62856	121541	0.04107
7	350	-	400	248	688	0.16223	373	92504	214045	0.07232
8	400	-	450	287	975	0.22990	422	121114	335159	0.11325
9	450	-	500	315	1290	0.30417	474	149310	484469	0.16370
10	500	-	550	312	1602	0.37774	521	162552	647021	0.21862
11	550	-	600	293	1895	0.44683	573	167889	814910	0.27535
12	600	-	650	301	2196	0.51780	621	186921	1001831	0.33851
13	650	-	700	265	2461	0.58029	673	178345	1180176	0.39877
14	700	-	750	244	2705	0.63782	721	175924	1356100	0.45821
15	750	-	800	219	2924	0.68946	772	169068	1525168	0.51534
16	800	-	900	376	3300	0.77812	844	317344	1842512	0.62256
17	900	-	1000	270	3570	0.84178	944	254880	2097392	0.70868
18	1000	-	1250	387	3957	0.93303	1095	423765	2521157	0.85187
19	1250	-	1500	159	4116	0.97053	1350	214650	2735807	0.92440
20	2000	-		125	4241	1.00000	1790	223750	2959557	1.00000
21			計	4241				2959557		

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、線で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。

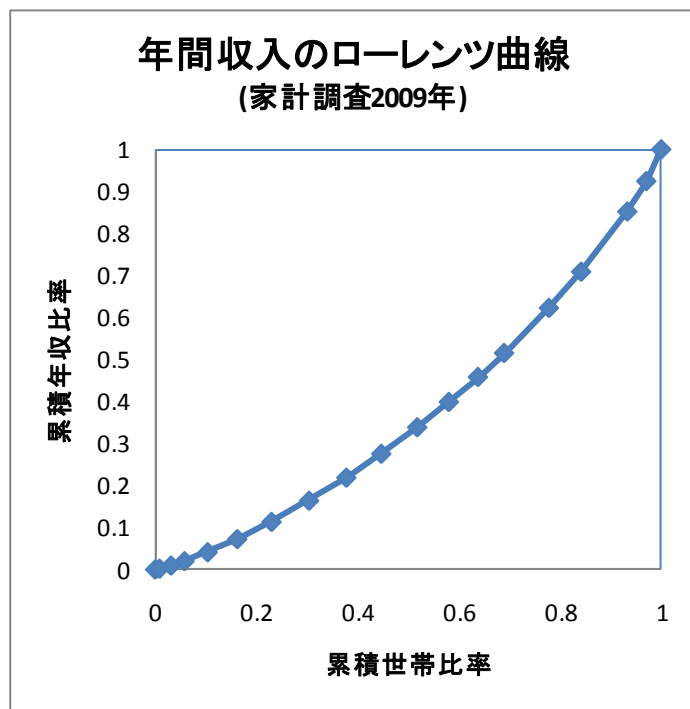
- ④ この時点でグラフのサンプルが自動的に描かれている。これを修正していく。まずレイアウトを変更する。リボンの中の「グラフのレイアウト」を展開し、レイアウト1（左上）をクリックし、タイトルや軸ラベルなどが書き込めるようにする。そして、

タイトル：「年間収入のローレンツ曲線（家計調査2009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 9】2009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。

☒ 本日実習したファイルは、河田まで提出すること。(講義時間中にできなかったものもおこない、完成させること)

提出は 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)おこなうこと。

締め切りは 12月6日(月)13:00 とする。

なお、ファイル名は「家計調査の分析 E39-○○○」のように、学籍番号をつけること。