

<家計に関する統計>

1. 準備

今回の実習には、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

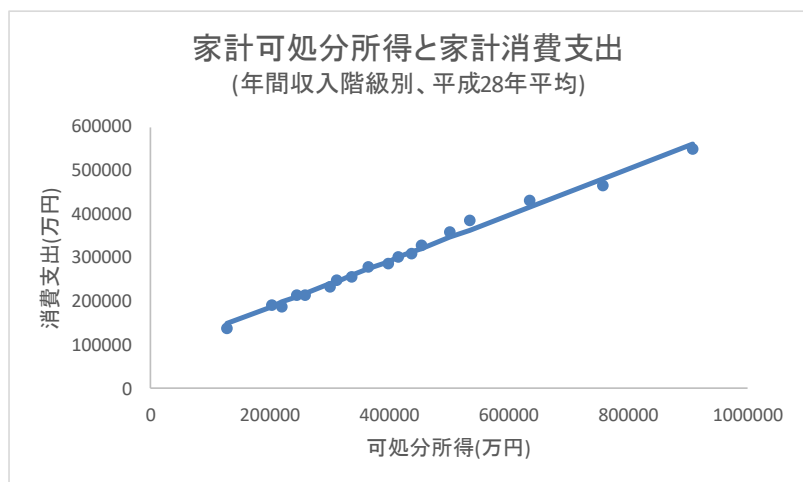
- ① 検索エンジンで、「河田研究室」と入力し検索すると、「河田研究室」のページにジャンプする。(ここまでの手順は、<http://www2.tokuyama-u.ac.jp/kawada> とアドレスを直接入力してもよい)
- ② 「経済統計」をクリックし、「第13回 11月15日(水)」の配布資料にある、「家計調査の分析 H27」をクリックし、自分の使いやすい場所に保存する。

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(Q240:AH240)を「消費関数」のシートの D2:D19 セルに、消費支出(Q72:AH72)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) をクリックする。そして、リボンの中の「グラフのレイアウト」にある、クイックレイアウト 1 (左上のもの) 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INTERCEPT<sup>1)</sup>)

$Y=a+bX$  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 FORECAST<sup>2)</sup>)

**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「塗りつぶしと線」のマークをクリックし、線の色として「**線 (単色)**」をチェックし、「マーカー」-「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 「ファイル」のタブをクリックし、下にある「オプション」のボタンをクリックする。
- ② 「Excel のオプション」のウインドウが開くので、左側の「アドイン」をクリックする。
- ③ 一番下に表示される「Excel アドイン」の右の**設定**ボタンを押す。
- ④ 「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。

<sup>1</sup> 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。

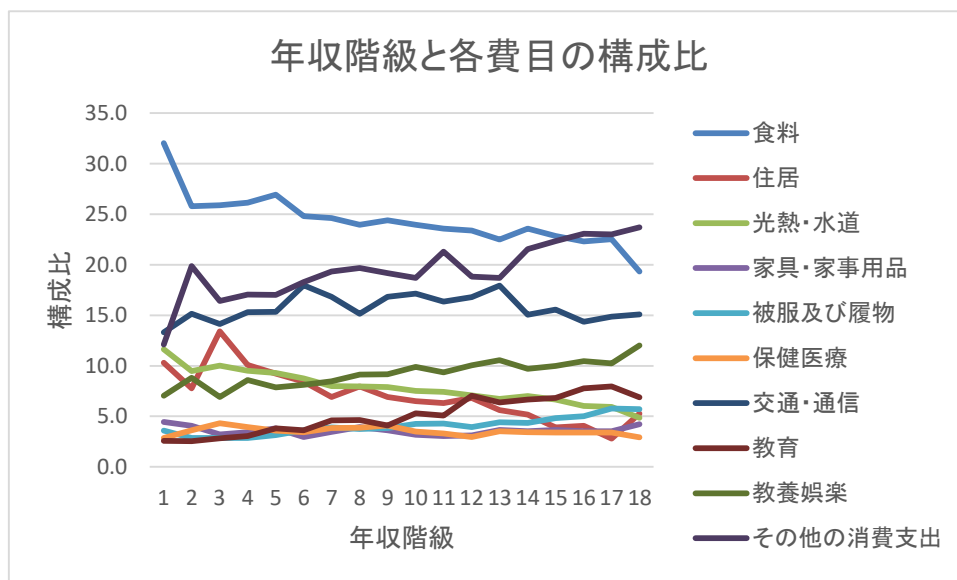
<sup>2</sup> FORECAST 関数と同様に予測値を求める関数として、**TREND** という関数があるが、ここでは説明は省略する

## 5. 構成比の導出とグラフの描画

【課題 7】 食料、住居、… といった十大費目について、各費目の構成比(消費支出に占めるそれぞれの費目の割合)を求めよう。そして、それを折れ線グラフに描こう。

### 📖 手順

- ① 2-6 表の消費支出(Q72:AH72)を「エンゲル関数」のシートの D2:D19 セルに、各費目別の支出(食料なら Q73:AH73。以下 10 大費目を順次コピーする)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、食料の消費支出に占める割合を求める。P2 セルに  $=E2/\$D2*100$  と入力する。
- ③ P2 セルを P2:Y19 にコピーすれば、十大費目の構成比が算出される。(P1:Y1 に各費目の名称(E1:N1)をコピーしよう。)
- ④ P1:Y19 を範囲指定し、折れ線グラフを描く。「クイックレイアウト」でレイアウト 1 を選び、「グラフ要素を追加」で「軸ラベル」-「第 1 横軸」を付け加えれば、下のようなグラフとなる。



## 6. 統計関数をもちいたエンゲル関数の導出

【課題 8】 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 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。
- ② 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。

## 7. 弾性値の導出

【課題 9】食料、住居、…といった十大費目について、各費目の支出弾力性の値(弾性値)を求めよう。

支出弾力性の値は、各費目と消費支出をともに対数を取り、回帰分析  $\log E = a + b \log C$  をおこなった場合の  $b$  の値である。

### 📖 手順

- ① 消費支出と各費目別の支出の対数を求める。ここでは、年収 200 万未満の世帯の消費支出を D35 セルに求めることにしよう。D35 セルに `=LOG(D2)` と入力する。これを D35:N52 セルにコピーする。
- ② 最初に、食料の対数をとったものを、消費支出の対数に回帰する。M22 セルに両対数をとったときの、 $b$  の推定値を求める。M22 セルに `=SLOPE(E35:E52,D35:D52)` と入力する。
- ③ 住居の弾性値は、E35:E52 の部分が F35:F52 になる。このようにして十大費目の弾性値をすべて求め、表を完成させよう。

## 8. ローレンツ曲線の描画

【課題 10】 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 2-6表の集計世帯数(Q16:AH16)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=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<sup>3</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**：2-6表のデータは、各階級の平均年収である(たとえば、Q271セルの 158 は、年収200万未満の39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。

### 📖 手順

- ① 2-6表の年間収入(Q271:AH271)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。
- ④ 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。

<sup>3</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。

(1)、(2)により次ページの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	39	39	0.00999	158	6162	6162	0.00229
4	200	-	250	78	117	0.02997	225	17550	23712	0.00880
5	250	-	300	107	224	0.05738	271	28997	52709	0.01956
6	300	-	350	158	382	0.09785	324	51192	103901	0.03857
7	350	-	400	220	602	0.15420	374	82280	186181	0.06911
8	400	-	450	280	882	0.22592	425	119000	305181	0.11328
9	450	-	500	275	1157	0.29636	474	130350	435531	0.16166
10	500	-	550	300	1457	0.37321	522	156600	592131	0.21979
11	550	-	600	270	1727	0.44237	573	154710	746841	0.27721
12	600	-	650	284	2011	0.51511	621	176364	923205	0.34267
13	650	-	700	266	2277	0.58325	673	179018	1102223	0.40912
14	700	-	750	252	2529	0.64780	723	182196	1284419	0.47675
15	750	-	800	210	2739	0.70159	772	162120	1446539	0.53693
16	800	-	900	361	3100	0.79406	843	304323	1750862	0.64988
17	900	-	1000	250	3350	0.85809	943	235750	1986612	0.73739
18	1000	-	1250	326	3676	0.94160	1096	357296	2343908	0.87001
19	1250	-	1500	136	3812	0.97643	1352	183872	2527780	0.93826
20	1500	-		92	3904	1.00000	1808	166336	2694116	1.00000
21			計	3904				2694116		

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

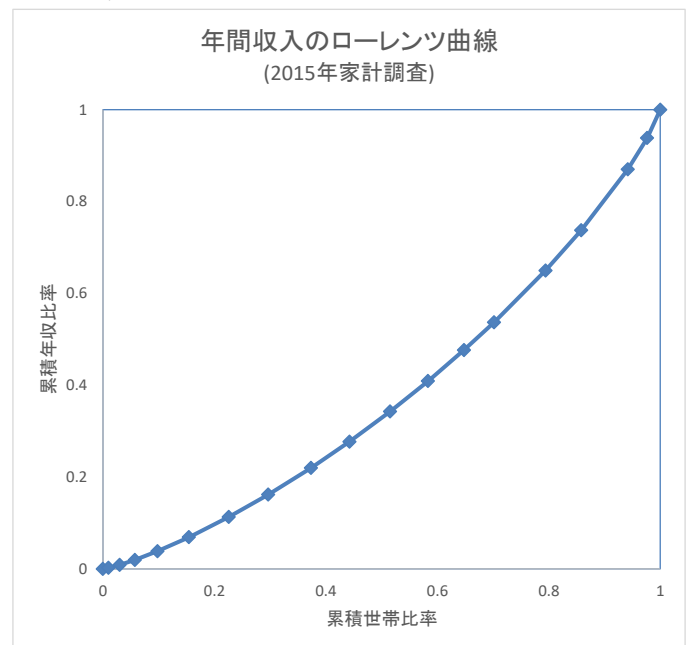
- ① 最初にグラフに描く範囲を範囲指定する。F2:F20をドラッグし、**[Ctrl]**キーを押しながらJ2:J20をドラッグして範囲指定する。
- ② グラフを作成するには、**挿入タブ**をクリックすることで、リボン内にグラフのグループが表示される。ここでは、**散布図**のボタンをクリックする。
- ③ すると散布図グラフのフォーマット(型式)メニューが出るので、左下の**散布図(直線とマーカー)**をクリックする。
- ④ この時点でグラフのサンプルが自動的に描かれている。これを修正していく。まずレイアウトを変更する。リボン中の「**クイックレイアウト**」をクリックし、レイアウト1(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(2016年家計調査)」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)



## 9. ジニ係数の導出

【課題 11】 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。

☒ 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つたものもおこない、完成させること)。提出は

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 **12月4日(月)10:55** とする。

なお、ファイル名は「家計調査の分析 H28 E46-○○○」のように、**学籍番号**をつけること。