

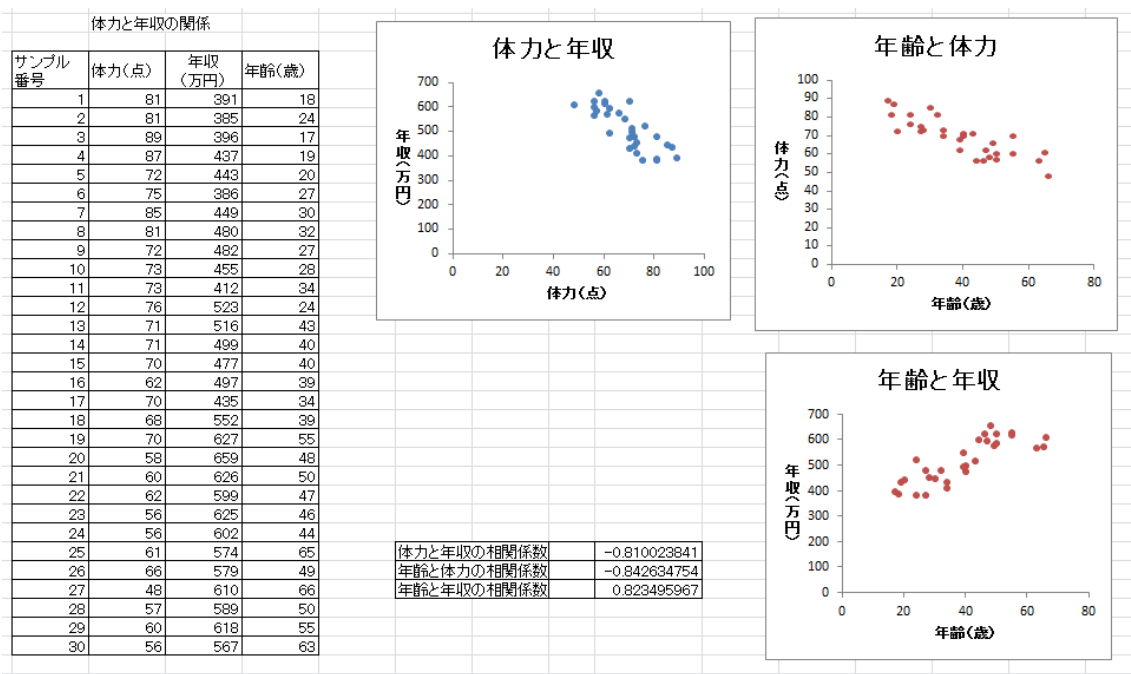
統計基礎 第13回 (2018.1.15)
 テーマ：クロス集計、相関分析の演習

1 はじめに

テスト 第15回目(1/29)にテスト(総合演習)を行います。PCを使った問題と筆記型の試験問題の併用型です。資料(教科書、配布資料等の紙類)の持込みは可。ケータイの持込みは不可。冒頭に座席の並び替えを行います。持込み可です。冒頭に、座席を指定しますのでご協力下さい。

前回の演習4の解答例

- (1) 以下の30人のデータから、体力と年収について、散布図を描くとともに、相関係数を求めてください。



- (2) 体力と年収の相関係数は? → -0.81
- (3) ①体力と年収の相関係数=-0.81
 ②年齢と体力の相関係数=-0.84
 ③年齢と年収の相関係数=0.82
- 体力と年収の因果関係はない。「年齢が高くなると体力が落ち」、「年齢が高くなると年収が上がる」ことから、「体力と年収」は相関(疑似相関)が見られる。

(注)「年収の高い者は、デスクワークに従事する人が多いために、体力が低い」という因果関係も考えられるが、「年齢と体力」、「年齢と年収」の散布図から、年齢による影響が大きいと考えられる。

ビデオ NHK女神ビジュアル「徹底検証！男性を魅了するモテ仕草」
(2013. 3. 13、12分)

(ポイント)

- ① 統計では、できるだけ調べたい項目以外の条件を一定にする ⇒ 双子の調査
- ② アンケートは人の意思が介在する。脳の状態を測る方が正確。
(ただし、脳は未解明の部分多し)
- ③ たくさんの人で調べれば傾向は見えてくる。

2 量的データと質的データ

- ・ **量的データ**：四則演算ができるデータ
(例) テストの得点、金額、温度、速度
- ・ **質的データ**：四則演算ができないデータ
(例) 5段階評価 (とても好き、好き、ふつう、嫌い、とても嫌い)
所在地 (都道府県)

演習 1 次のデータは量的データですか、質的データですか？

- (1) 身長
- (2) 血液型
- (3) アンケートの回答 (「そう思わない=1」「どちらかといえばそう思わない=2」
「どちらともいえない=3」「どちらかといえばそう思う=4」「そう思う=5」)

3 クロス集計とは何ですか？

① 2つの項目の関係を調べる方法です。

- 「量的データ」 → 相関分析（散布図、相関係数等）
- 「質的データ」（例：アンケート調査） → クロス集計

② データから、クロス集計表（度数）を作成します。

③ 必要に応じ、クロス集計表（度数）から、クロス集計表（割合）を作成します。

- 作成方法には、3つの方法があります。
どの方法を選択するかは、「何を知りたいのか」により決まります。

4 クロス集計表の作り方

演習2 20人の性別（男、女）と朝食の種別（パン、ご飯）に関するアンケート結果（架空のデータ）です。

整理番号	性別	朝食の種別
1	男	パン
2	男	ご飯
3	女	パン
4	女	パン
5	男	パン
6	女	ご飯
7	男	ご飯
8	女	パン
9	男	ご飯
10	女	ご飯

整理番号	性別	朝食の種別
11	女	パン
12	男	パン
13	女	パン
14	女	ご飯
15	男	ご飯
16	男	パン
17	女	パン
18	男	ご飯
19	男	ご飯
20	男	ご飯

- (1) クロス集計表（度数）を完成させてください。

表1 クロス集計表（度数、単位：人）

	パン	ご飯	合計
男性			
女性			
合計			

- (2) (1)で作成したクロス集計表（度数）から、クロス集計表（割合）を3種類作成して下さい。

表2 男性と女性の差異に関するクロス集計表（横方向の割合）

（目的：男性（や女性）が何を食べているかを知りたい）

	パン	ご飯	合計
男性	%	%	100%
女性	%	%	100%
全体	%	%	100%

表3 パンとご飯の差異に関するクロス集計表（縦方向の割合）

（目的：パン（やご飯）をどの性別が食べているのかを知りたい）

	パン	ご飯	全体
男性	%	%	%
女性	%	%	%
合計	100%	100%	100%

表4 それぞれの組合せに関するクロス集計表（各セルの割合）

（目的：両方の条件に当てはまる人の全体に占める割合を知りたい）

	パン	ご飯	合計
男性	%	%	%
女性	%	%	%
合計	%	%	100%

図1 男女の主食の違い

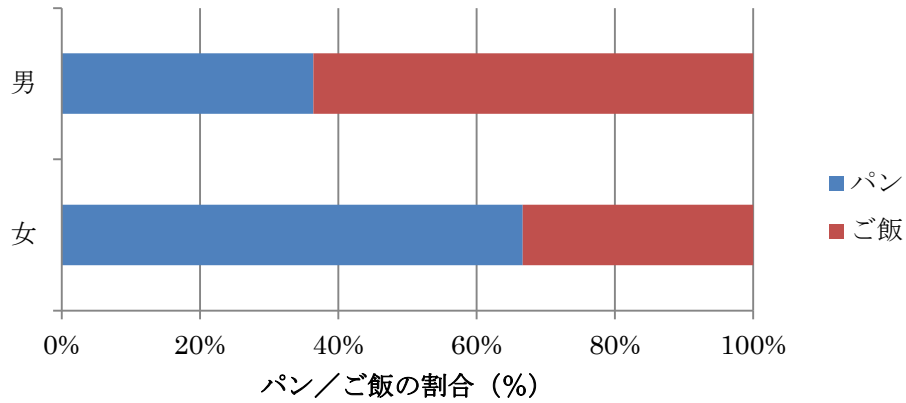


図2 主食に占める男女の割合

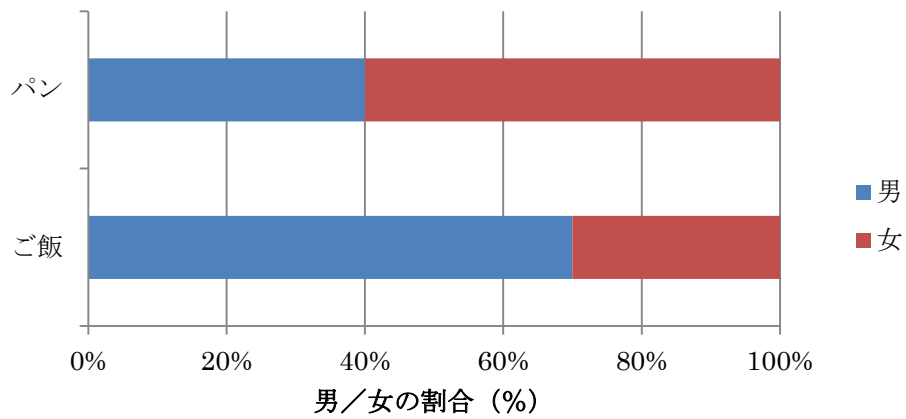
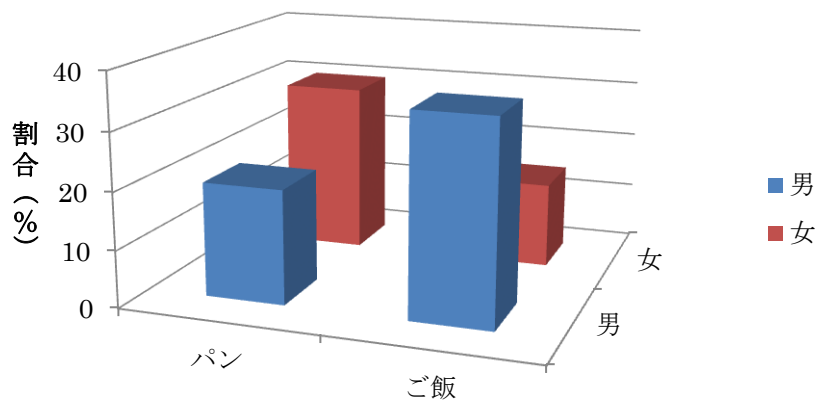


図3 性別・主食別の割合



演習3 次の表は、社会マネジメント学科学生へのアンケート調査の結果（出典：相模女子大学ウェブマガジン bloomy Vol. 3, 2011. 12）（社マネ2年有志が実施・分析したもの）の一部（筆者が一部加工）です。分析して下さい。

表：「アルバイトの有無」と「彼氏の有無」に関するクロス集計

（単位：人）

	彼氏がいる	彼氏がない	合計
バイトをしている	37	68	105
バイトをしていない	6	37	43
合計	43	105	148

【解答欄】

表：「アルバイトの有無」と「彼氏の有無」に関するクロス集計

（単位：%）

	彼氏がいる	彼氏がない	合計
バイトをしている			100%
バイトをしていない			100%
合計			100%

（事実）

（考察）

演習 4 次の表は、各都道府県の人口、面積、警察施設数です。ここで警察施設とは、警察署・派出所・駐在所を指します。

(1) 散布図を作成して下さい。

- 都道府県の人口と警察施設数
- 都道府県の面積と警察施設数

(2) 相関係数を求めて下さい。

- 都道府県の人口と警察施設数
- 都道府県の面積と警察施設数
- 北海道を除く都道府県の面積と警察施設数
- (注) 北海道は特殊な都道府県(面積が極端に大きく人口が少ない)(=外れ値)なので除外するとよいです。全体の傾向を知りたい場合は、特殊な例を除外することは問題なく、むしろ望ましいです。

(3) 相関はあると判断しますか? ないと判断しますか?

- 都道府県の人口と警察施設数
- 都道府県の面積と警察施設数

都道府県の人口、面積、警察施設数				
都道府 県番号	都道府 県	人口 (千人) 2013. 10. 1現在	面積 (千 k m ²)	警察施設 数 (警察 署・派出 所・駐在 所数、 2012年)
1	北海道	5,431	83.46	791
2	青森県	1,335	9.64	171
3	岩手県	1,295	15.28	215
4	宮城県	2,328	6.86	248
5	秋田県	1,050	11.64	162
6	山形県	1,141	6.65	171
7	福島県	1,946	13.78	237
8	茨城県	2,931	6.10	268
9	栃木県	1,986	6.41	286
10	群馬県	1,984	6.36	205
11	埼玉県	7,222	3.77	411
12	千葉県	6,192	5.08	525
13	東京都	13,300	2.10	1186
14	神奈川県	9,079	2.42	668
15	新潟県	2,330	10.36	366
16	富山県	1,076	2.05	162
17	石川県	1,159	4.19	187
18	福井県	795	4.19	149
19	山梨県	847	4.20	159
20	長野県	2,122	13.10	269
21	岐阜県	2,051	9.77	246
22	静岡県	3,723	7.26	381
23	愛知県	7,443	5.12	602
24	三重県	1,833	5.76	218
25	滋賀県	1,416	3.77	174
26	京都府	2,617	4.61	314
27	大阪府	8,849	1.90	710
28	兵庫県	5,558	8.40	747
29	奈良県	1,383	3.69	192
30	和歌山県	979	4.73	182
31	鳥取県	578	3.51	112
32	島根県	702	6.71	170
33	岡山県	1,930	7.01	321
34	広島県	2,840	8.48	323
35	山口県	1,420	6.11	261
36	徳島県	770	4.15	146
37	香川県	985	1.86	158
38	愛媛県	1,405	5.68	233
39	高知県	745	7.11	121
40	福岡県	5,090	4.85	363
41	佐賀県	840	2.44	150
42	長崎県	1,397	4.11	223
43	熊本県	1,801	7.27	197
44	大分県	1,178	5.10	160
45	宮崎県	1,120	6.79	174
46	鹿児島県	1,680	9.04	252
47	沖縄県	1,415	2.28	162

5 Excel の基本操作 (前回の資料と同じです)

(1) 離れた2つのセル範囲を選択する方法 (復習)

- ① 1つ目のセル範囲を選択
- ② ctrl キーを押したまま2つ目のセル範囲を選択

(2) 散布図のX軸 (横軸) とY軸 (縦軸) を交換する方法

(注) 散布図では、原因となるものをX軸とし、結果となるものをY軸とします。

- ・ グラフツール→「デザイン」タブ→「データの選択」
- ・ 「データソースの選択」のダイアログボックスで「編集」ボタンをクリック
 - 系列名: グラフのタイトルを入力 (又はタイトルのあるセルを選択)
 - 系列Xの値の欄: セルの範囲を設定 (項目名は含めない)
 - 系列Yの値の欄: セルの範囲を設定 (項目名は含めない)

(3) X軸ラベル、Y軸ラベルに素早く入力する方法

- ・ (グラフの余白をクリック) →グラフツール→デザイン→グラフのレイアウト
→ (グラフのX軸ラベルに「軸ラベル」、グラフのY軸ラベルに「軸ラベル」という文字が入力されているレイアウトを選択)

(4) セル内で改行する方法

- ・ Alt キーを押したまま Enter キーを押す