



独立行政法人

統計センター

一般用ミクロデータ 就業構造基本調査版

～擬似標本データによる統計演習例～

統計センター 伊原 一
北林 三就

総務省統計局 滝澤 有美

信頼に応じて作る統計表

平成30年9月12日
2018年度統計関連学会連合大会

<はじめに>

統計データには数量データと、性別や就業状態などのカテゴリーデータがあり、数量データについて2変量間の相関係数を求めるためには個票データを用いる必要がある。

一方で、性別や就業状態などのカテゴリーデータについて相関係数を求める場合は個票データを利用しなくても四分点分割表から相関係数を計算することができる。

本稿では、カテゴリーデータとして統計センターが提供を行っている一般用ミクロデータ就業構造基本調査版を用いた相関分析を擬似標本データによる統計演習例として紹介する。

1. 女性と有業者の相関分析

一般用ミクロデータ就業構造基本調査版(2012年)を利用して、女性と有業者の相関係数による相関分析を行った。

本稿の相関分析表において、 (a_w, b_w, c_w, d_w) は、それぞれ四分点分割表の加重度数を示す。

四分点分割表の加重度数

	無業(0)	有業(1)
男性(0)	a_w	b_w
女性(1)	c_w	d_w

1. 女性と有業者の相関分析

(1) 女性と有業者の相関係数

	加重度数 Σw		公表値ベース		標本度数 Σ	
	無業者(0)	有業者(1)	無業者(0)	有業者(1)	無業者(0)	有業者(1)
男(0)	16,669,600	36,743,300	16,668,700	36,744,500	33,507	73,257
女(1)	29,725,100	27,675,400	29,725,700	27,676,200	59,282	55,584
相関係数	-0.208406		-0.208424		-0.204848	

注) 擬似標本は丸め誤差を含む結果表を基に作成されているため、加重度数は公表値と完全には一致しない。
(以下、同様。)

加重相関係数と検定値

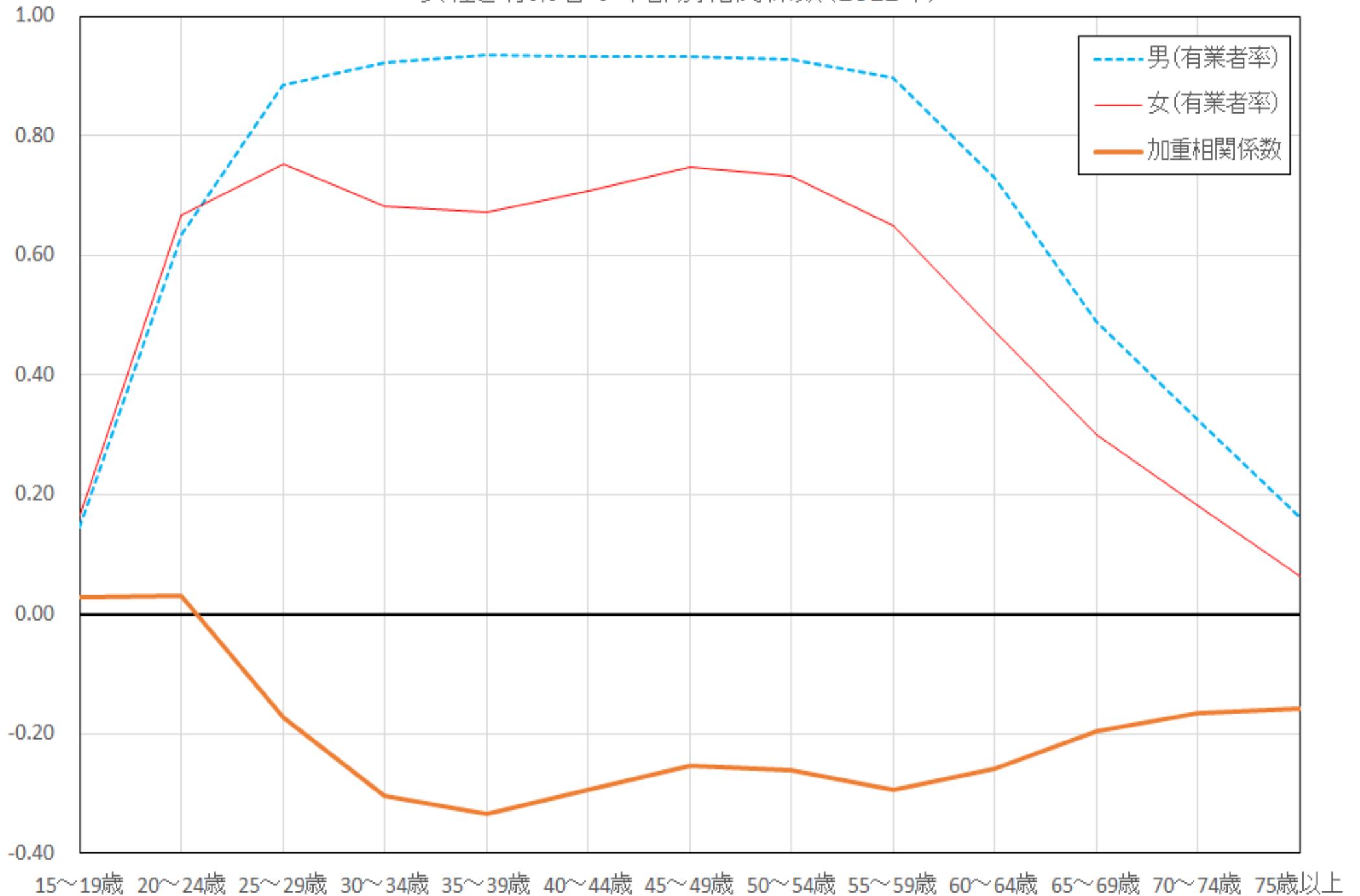
母数N	a_w	b_w	c_w	d_w	標本 n	相関係数 ϕ_w	カイ ² 乗 χ^2	危険率 p 値
110813400	16669600	36743300	29725100	27675400	221630	-0.2084	9626.15	0.00000

1. 女性と有業者の相関分析

(2) 女性と有業者の年齢別相関分析表

年齢	母数N	a _w	b _w	c _w	d _w	標本n	相関係数φ _w	カイ2乗χ ²	危険率p値
15～19歳	6041800	2642700	451100	2460200	487800	12061	+0.0271	8.87	0.00144
20～24歳	6240700	1156700	2026000	1022200	2035800	12613	+0.0305	11.79	0.00029
25～29歳	7018500	409100	3154100	854500	2600800	14160	-0.1723	420.82	0.00000
30～34歳	7813900	304700	3649700	1228700	2630800	15666	-0.3037	1445.60	0.00000
35～39歳	9402900	308900	4451300	1528900	3113800	18759	-0.3333	2084.88	0.00000
40～44歳	9454400	322100	4450600	1373300	3308400	18850	-0.2943	1633.14	0.00000
45～49歳	8193900	278600	3837600	1034200	3043500	16376	-0.2534	1051.97	0.00000
50～54歳	7671000	277500	3556300	1030300	2806900	15352	-0.2607	1043.85	0.00000
55～59歳	7948200	405300	3538200	1402000	2602700	15935	-0.2950	1386.85	0.00000
60～64歳	10241500	1369500	3651200	2751700	2469100	20389	-0.2592	1370.16	0.00000
65～69歳	8200900	2004300	1927400	2995300	1273900	16450	-0.1964	634.82	0.00000
70～74歳	7394700	2326000	1113300	3243100	712300	14806	-0.1661	408.61	0.00000
75歳以上	15191000	4864200	936500	8800700	589600	30213	-0.1594	768.11	0.00000

女性と有業者の年齢別相関係数(2012年)



1. 女性と有業者の相関分析

<結論>

(1) 分析結果

年齢別の相関分析の結果から、年齢合計については女性と有業者の相関係数は有意に負値を示すが、年齢階級別に見てみると、15～19歳及び20～24歳の年齢については相関係数が有意に正値を示すことがわかる。

(2) 計算方法

性別や就業状態などのカテゴリー変数について相関分析を行う場合は変数に0と1の値を付与し、二項分布の計算により相関係数を求めるが、相関係数は二項分布の性質を利用することで、四分点分割表から計算することができる。

2. 一般用マイクロデータ（就業構造基本調査版）

統計センターでは、統計演習など教育用に利用可能な一般用マイクロデータを提供しており、平成30年9月から一般用マイクロデータ就業構造基本調査版の提供を開始している。

本稿では、擬似標本による個票データとして、就業構造基本調査に基づく一般用マイクロデータを用いている。

一般用マイクロデータ（統計センター）

<<https://www.nstac.go.jp/services/ippan-microdata.html>>

2. 一般用マイクロデータ（就業構造基本調査版）

(1) 一般用マイクロデータの特徴

就業構造基本調査に基づく一般用マイクロデータ（PUMESS: Public Use Microdata of Employment Status Survey）は、就業構造基本調査の公表結果表を母集団情報として用い、系統抽出と同様の方法により抽出率500分の1の擬似標本データを作成している。

このため、PUMESS データには母集団の分布特性が反映されており、母集団フレームからの系統抽出による擬似的な無作為標本とみなすことで、収録項目の範囲内で人口分析や統計演習等に利用することが可能である。

2. 一般用マイクロデータ（就業構造基本調査版）

(1) 一般用マイクロデータの特徴

PUMESS データには公表結果表をベンチマーク人口（就業状態、都道府県、男女、年齢別の人口）として比推定補正を行った復元ウェイトを収録しており、この復元ウェイトを用いて集計を行うことで集計結果の標本誤差を抑える工夫をしている。

利用にあたっては、擬似標本の標本規模が実際の調査の5分の1程度で標本誤差が大きくなっており、集計結果は公表結果表と完全には一致しないこと、また、丸め誤差を含む結果表から作成しているため、標本誤差に加えて丸め誤差が含まれていることに留意する必要がある。

2. 一般用マイクロデータ（就業構造基本調査版）

(2) 一般用マイクロデータの収録項目

本稿では、有業者と無業者の共通項目として、下記の3項目を用いている。

- ① 性別（男／女）、
- ② 年齢（5歳階級）、
- ③ 就業状態（有業者／無業者）

一般用マイクロデータ就業構造基本調査版には、下表の項目が収録されている。

一般用ミクロデータ就業構造基本調査版の収録項目

		収録項目 (項目名)		2012	2007	2002	1997	1992
共通項目	データ名称	DataSource	○	○	○	○	○	○
	調査年	Year	○	○	○	○	○	○
	都道府県	Prefecture	○	○	○	○	○	○
	政令指定市	City	○	○	○	○	○	○
	市部	Urban	○	○	○	×	×	×
	性別	Gender	○	○	○	○	○	○
	年齢	Age	○	○	○	○	○	○
	就業状態	WorkStatus	○	○	○	○	○	○
有業者	雇用者	WorkEmploy	○	○	○	○	○	○
	正規就業者	WorkRegular	○	○	×	×	×	×
	産業	WorkIndustry	○	○	○	○	○	○
無業者	就業希望	NoworkWish	○	○	○	○	○	○
	求職	NoworkApply	○	○	○	○	○	○
	非就業者年齢	NoworkAge	○	○	○	○	○	○
	配偶関係	NoworkMarriage	○	○	×	○	○	○
符号ラベル (テキスト) T_Prefecture、T_City、T_Urban、T_Gender、T_Age、T_WorkStatus、 T_WorkEmploy、T_WorkRegular、T_WorkIndustry、 T_NoworkWish、T_NoworkApply、T_NoworkAge、T_NoworkMarriage								
復元ウェイト (Weight)								

ippan_2012shugyou_dataset

※一般用マイクロデータは、公表済みの結果表から作成したマイクロデータ形式のデータセットです。
 このため、一般用マイクロデータから導かれた分析結果は実証研究の結果とみなすことはできません。
 ※注意事項など、一般用マイクロデータの最新情報は、統計センターHP(<http://www.nstac.go.jp/services/ippan-microdata.html>)において公開しています。
 ご利用、いただく前にお確認ください。

DataSource	Year	Prefecture	City	Urban	Gender	Age	WorkStatus	WorkEmplo	WorkRegul	WorkIndust	NoworkWis	NoworkApp	NoworkAge	NoworkMar	T_Prefectur	T_City	T_Urban	T_Gender	T_Age	T_WorkStat	T_WorkEmp	T_WorkReg	T_WorkIndu	T_NoworkW	T_NoworkA	T_NoworkM	Weight
PUM_ESS	2012	01	100	001	001	200	001	007	044	360					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	建設業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	361					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	製造業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	365					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	卸売業, 小売業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	365					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	卸売業, 小売業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	365					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	卸売業, 小売業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	365					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	卸売業, 小売業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	367					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	不動産業, 物品賃貸業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	369					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	宿泊業, 飲食サービス業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	369					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	宿泊業, 飲食サービス業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	370					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	生活関連サービス業, 娯楽業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	370					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	生活関連サービス業, 娯楽業				486.842105
PUM_ESS	2012	01	100	001	001	200	001	007	044	374					北海道	札幌市	市部	男	15~19歳	有業者	雇業者	非正規の職	サービス業(他に分類されないもの)				486.842105
PUM_ESS	2012	01	100	001	001	201	001	007	011	360					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	建設業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	360					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	建設業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	361					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	建設業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	363					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	製造業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	363					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	情報通信業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	363					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	情報通信業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	369					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	宿泊業, 飲食サービス業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	369					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	宿泊業, 飲食サービス業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	370					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	生活関連サービス業, 娯楽業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	371					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	教育, 学習支援業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	372					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	医療, 福祉				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	372					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	医療, 福祉				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	372					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	医療, 福祉				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	374					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	サービス業(他に分類されないもの)				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	374					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	サービス業(他に分類されないもの)				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	374					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	サービス業(他に分類されないもの)				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	375					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	公務(他に分類されるものを除く)				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	011	376					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	正規の職	分類不能の産業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	361					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	製造業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	363					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	情報通信業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	364					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	運輸業, 郵便業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	364					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	運輸業, 郵便業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878
PUM_ESS	2012	01	100	001	001	201	001	007	044	365					北海道	札幌市	市部	男	20~24歳	有業者	雇業者	非正規の職	卸売業, 小売業				487.804878

2. 一般用マイクロデータ（就業構造基本調査版）

(2) 一般用マイクロデータの収録項目

一般用マイクロデータは、データ作成の際に政府統計総合窓口(e-Stat)のAPI機能を利用しており、符号についてはAPIに用いられているものを使用している。

<参考>

一般用マイクロデータ就業構造基本調査版の概要
～系統抽出による擬似標本データ～

<http://www.nstac.go.jp/services/society_paper/30_06_03.pdf>

3. 加重相関係数の一般式 (例題)

カテゴリーデータの加重相関係数 ϕ_w については、相関係数の一般式から求めることができる。

$$\begin{aligned}\phi_w &= \frac{\sigma_{wxy}}{\sigma_{wx} \sigma_{wy}} \\ &= \frac{\sum wxy / \sum w - \mu_{wx} \mu_{wy}}{\sqrt{(\sum wx^2 / \sum w - \mu_{wx}^2) (\sum wy^2 / \sum w - \mu_{wy}^2)}}\end{aligned}$$

x	加重平均	$\mu_{wx} = \sum wx / \sum w$
y	加重平均	$\mu_{wy} = \sum wy / \sum w$
x^2	加重平均	$\sum wx^2 / \sum w$
y^2	加重平均	$\sum wy^2 / \sum w$
xy	加重平均	$\sum wxy / \sum w$

3. 加重相関係数の一般式（例題）

例) 一般式による計算

無作為標本15件のウェイト付個票データについて、女性と有業者の加重相関係数を求めて無相関検定を行う。

性別	就業状態	Weight
男	有業者	7348660
男	無業者	8334800
男	無業者	8334800
女	有業者	6918850
女	無業者	7431275

3. 加重相関係数の一般式 (例題)

補正乗率 w' を求める。

$$w_1' = 7348660 * 15 / 110813400 = 0.994734$$

$$w_2' = 7348660 * 15 / 110813400 = 0.994734$$

$$w_3' = 7348660 * 15 / 110813400 = 0.994734$$

$$w_4' = 7348660 * 15 / 110813400 = 0.994734$$

$$w_5' = 7348660 * 15 / 110813400 = 0.994734$$

$$w_6' = 8334800 * 15 / 110813400 = 1.128220$$

$$w_7' = 8334800 * 15 / 110813400 = 1.128220$$

$$w_8' = 6918850 * 15 / 110813400 = 0.936554$$

$$w_9' = 6918850 * 15 / 110813400 = 0.936554$$

$$w_{10}' = 6918850 * 15 / 110813400 = 0.936554$$

3. 加重相関係数の一般式（例題）

補正乗率 w' を求める。

$$w_{11}' = 6918850 * 15 / 110813400 = 0.936554$$

$$w_{12}' = 7431275 * 15 / 110813400 = 1.005917$$

$$w_{13}' = 7431275 * 15 / 110813400 = 1.005917$$

$$w_{14}' = 7431275 * 15 / 110813400 = 1.005917$$

$$w_{15}' = 7431275 * 15 / 110813400 = 1.005917$$

補正のポイント

加重相関係数の値は、補正乗率で計算しても変わらない。
ウェイト付データは、ウェイトを積算すると母集団度数の合計 N になるため、合計が n になるように補正することで n を用いた加重相関係数の検定が可能となる。

3. 加重相関係数の一般式 (例題)

単純平均、2乗平均、 x y 積平均を求める。

性別	就業状態	$\Sigma w = N$	$\Sigma w' = n$	x	y	x^2	y^2	$x y$
男	有業者	7348660	0.9947	0	1	0	1	0
男	有業者	7348660	0.9947	0	1	0	1	0
男	有業者	7348660	0.9947	0	1	0	1	0
男	有業者	7348660	0.9947	0	1	0	1	0
男	有業者	7348660	0.9947	0	1	0	1	0
男	無業者	8334800	1.1282	0	0	0	0	0
男	無業者	8334800	1.1282	0	0	0	0	0
女	有業者	6918850	0.9365	1	1	1	1	1
女	有業者	6918850	0.9365	1	1	1	1	1
女	有業者	6918850	0.9365	1	1	1	1	1
女	有業者	6918850	0.9365	1	1	1	1	1
女	無業者	7431275	1.0059	1	0	1	0	0
女	無業者	7431275	1.0059	1	0	1	0	0
女	無業者	7431275	1.0059	1	0	1	0	0
女	無業者	7431275	1.0059	1	0	1	0	0

3. 加重相関係数の一般式 (例題)

$$w \text{ の合計} \quad \Sigma w = 110813400$$

$$\begin{aligned} x \text{ 加重平均} \quad \mu_{w x} &= \Sigma w x / \Sigma w \\ &= 57400500 / 110813400 = 0.517992409 \end{aligned}$$

$$\begin{aligned} y \text{ 加重平均} \quad \mu_{w y} &= \Sigma w y / \Sigma w \\ &= 64418700 / 110813400 = 0.581325905 \end{aligned}$$

$$\begin{aligned} x^2 \text{ 加重平均} \quad \Sigma w x^2 / \Sigma w \\ &= 57400500 / 110813400 = 0.517992409 \end{aligned}$$

$$\begin{aligned} y^2 \text{ 加重平均} \quad \Sigma w y^2 / \Sigma w \\ &= 64418700 / 110813400 = 0.581325905 \end{aligned}$$

$$\begin{aligned} x y \text{ 加重平均} \quad \Sigma w x y / \Sigma w \\ &= 27675400 / 110813400 = 0.249747774 \end{aligned}$$

3. 加重相関係数の一般式 (例題)

加重相関係数

$$\begin{aligned}\phi_w &= \frac{\sigma_{wxy}}{\sigma_{wx} \sigma_{wy}} \\ &= \frac{\sum wxy / \sum w - \mu_{wx} \mu_{wy}}{\sqrt{(\sum wx^2 / \sum w - \mu_{wx}^2) (\sum wy^2 / \sum w - \mu_{wy}^2)}} \\ &= \{0.249747 - 0.517992 * 0.581325\} \\ &\quad / \{ (0.517992 - 0.517992^2) \\ &\quad * (0.581325 - 0.581325^2) \}^{1/2} \\ &= -0.208406\end{aligned}$$

3. 加重相関係数の一般式 (例題)

$$w' \text{の合計} \quad \Sigma w' = 15$$

$$\begin{aligned} x \text{ 加重平均} \quad \mu_{w'x} &= \Sigma w'x / \Sigma w' \\ &= 7.769884 / 15 = 0.517992409 \end{aligned}$$

$$\begin{aligned} y \text{ 加重平均} \quad \mu_{w'y} &= \Sigma w'y / \Sigma w' \\ &= 8.719886 / 15 = 0.581325905 \end{aligned}$$

$$\begin{aligned} x^2 \text{ 加重平均} \quad \Sigma w'x^2 / \Sigma w' & \\ &= 7.769884 / 15 = 0.517992409 \end{aligned}$$

$$\begin{aligned} y^2 \text{ 加重平均} \quad \Sigma w'y^2 / \Sigma w' & \\ &= 8.719886 / 15 = 0.581325905 \end{aligned}$$

$$\begin{aligned} x y \text{ 加重平均} \quad \Sigma w'xy / \Sigma w' & \\ &= 3.746216 / 15 = 0.249747774 \end{aligned}$$

3. 加重相関係数の一般式 (例題)

加重相関係数 (補正乗率)

$$\begin{aligned}\phi_{w'} &= \frac{\sigma_{w'x} \sigma_{w'y}}{\sigma_{w'x} \sigma_{w'y}} \\ &= \frac{\sum w' x y / \sum w' - \mu_{w'x} \mu_{w'y}}{\sqrt{(\sum w' x^2 / \sum w' - \mu_{w'x}^2) (\sum w' y^2 / \sum w' - \mu_{w'y}^2)}} \\ &= \{0.249747 - 0.517992 * 0.581325\} \\ &\quad / \{ (0.517992 - 0.517992^2) \\ &\quad (0.581325 - 0.581325^2) \}^{1/2} \\ &= -0.208406\end{aligned}$$

3. 加重相関係数の一般式 (例題)

$$\begin{aligned}\chi^2 &= n \phi_w^2 = n \phi'^2 \\ &= 15 * (-0.208406)^2 \\ &= 0.651501\end{aligned}$$

$$\begin{aligned}p \text{ 値} &= \text{CHIDIST} (0.651501, 1) / 2 \\ &= 0.209788\end{aligned}$$

∴ 加重相関係数 $\phi_w = -0.208406$ は、危険率 $p = 0.209788$ ($> 2.5\%$) から信頼水準95% (片側97.5%) で有意であるとはいえない。

4. 相関係数の簡易式

カテゴリーデータについては、四分点分割表を用いて簡易式により相関係数 ϕ_w を計算できる。また、カイ 2 乗値 χ^2 により無相関検定を行うことができる。

加重相関係数 (簡易式)

$$\phi_w = \frac{a_w d_w - b_w c_w}{\sqrt{(a_w + b_w)(c_w + d_w)(a_w + c_w)(b_w + d_w)}}$$

カイ 2 乗値

$$\chi^2 = n \phi_w^2$$

4. 加重相関係数の簡易式（例題）

例) 簡易式による計算

無作為標本15件のウェイト付個票データについて、女性と有業者の加重相関係数を求めて無相関検定を行う。

性別	就業状態	Weight
男	有業者	7348660
男	無業者	8334800
男	無業者	8334800
女	有業者	6918850
女	無業者	7431275

4. 加重相関係数の簡易式（例題）

補正乗率 w' ($= n w / N$) を求めて、加重度数表を集計する。

加重度数表

	無業(Y0)	有業(Y1)
男(X0)	16,669,600	36,743,300
女(X1)	29,725,100	27,675,400

加重度数表（補正乗率 w' ）

	無業(Y0)	有業(Y1)
男(X0)	2.256441	4.973671
女(X1)	4.023669	3.746216

4. 加重相関係数の簡易式 (例題)

加重相関係数

$$\begin{aligned}\phi_w &= \frac{a_w d_w - b_w c_w}{\sqrt{(a_w + b_w)(c_w + d_w)(a_w + c_w)(b_w + d_w)}} \\ &= (16,669,600 * 27,675,400 - 36,743,300 * 29,725,100) \\ &\quad / \{ (16,669,600 + 36,743,300) \\ &\quad * (29,725,100 + 27,675,400) \\ &\quad * (16,669,600 + 29,725,100) \\ &\quad * (36,743,300 + 27,675,400) \}^{1/2} \\ &= -0.208406\end{aligned}$$

4. 加重相関係数の簡易式 (例題)

加重相関係数 (補正乗率 w')

$$\begin{aligned}\phi_{w'} &= \frac{a_{w'}d_{w'} - b_{w'}c_{w'}}{\sqrt{(a_{w'}+b_{w'})(c_{w'}+d_{w'})(a_{w'}+c_{w'})(b_{w'}+d_{w'})}} \\ &= (2.256*3.746 - 4.973*4.023) \\ &\quad - \left\{ (2.256 + 4.973)(4.023 + 3.746) \right. \\ &\quad \left. * (2.256 + 4.023)(4.973 + 3.746) \right\}^{1/2} \\ &= -0.208406\end{aligned}$$

4. 加重相関係数の簡易式 (例題)

$$\begin{aligned}\chi^2 &= n \phi_w^2 = n \phi_w'^2 \\ &= 15 * (-0.208406)^2 \\ &= 0.651501\end{aligned}$$

$$\begin{aligned}p \text{ 値} &= \text{CHIDIST} (0.651501, 1) / 2 \\ &= 0.209788\end{aligned}$$

∴加重相関係数 $\phi_w = -0.208406$ は、危険率 $p = 0.209788$ ($> 2.5\%$) から信頼水準95% (片側97.5%) で有意であるとはいえない。

<まとめ>

一般用ミクロデータ就業構造基本調査版は、擬似標本データとして公表結果表を基に作成しており、擬似的な無作為標本として扱うことで収録項目の範囲内で個票データと同様の統計演習等に利用することができる。

①加重集計

擬似標本データを用いることで、ウェイト付個票データの集計演習に利用することができる。

②相関分析

擬似標本データを用いることで、個票データによる相関分析の確認計算に利用することができる。

四分点相関係数の簡易式とカイ 2 乗値の証明等については、下記を参照されたい。

<参考>

四分点相関係数による相関分析手法
(エストレーラ 9 月号, 2018年, 伊原)

本稿の意見等については、筆者の個人的な見解によるものです。