

# 課題の解答

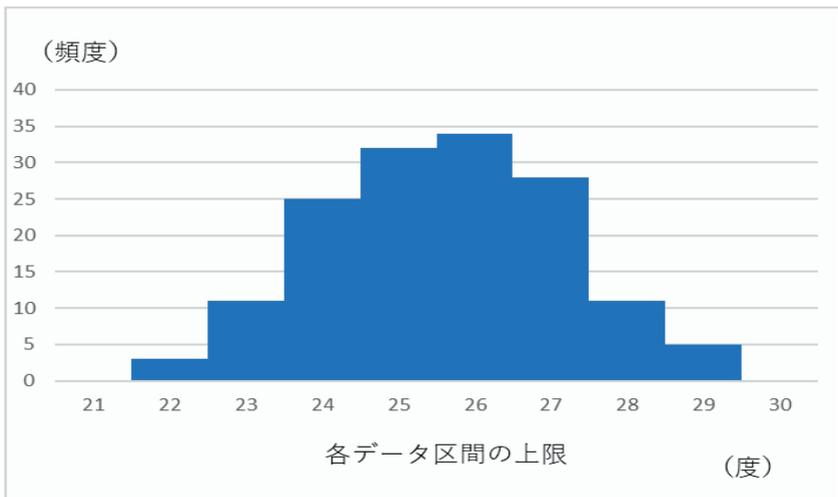
課題で使う Excel のデータは、オーム社の Web サイト内にある本書紹介ページにて提供しています。

## 第 1 章

### 課題 1

### ヒストグラム

7月の気温のヒストグラムです。



## 課題 2

## 平均値と中央値

年収のデータの基本統計量です。

年収	
平均	564.2857
標準誤差	101.8883
中央値（メジアン）	440
最頻値（モード）	440
標準偏差	466.9108
分散	218005.7
尖度	3.926996
歪度	1.866877
範囲	1910
最小	90
最大	2000
合計	11850
データの個数	21

## 課題 3

## 相関係数

実質 GDP とその構成項目の相関係数です。

	実質GDP	実質民間 最終消費	実質民間 設備投資	実質財貨・ サービスの輸出
実質GDP	1.00			
実質民間最終消費	0.96	1.00		
実質民間設備投資	0.97	1.00	1.00	
実質財貨・サービスの輸出	0.85	0.69	0.73	1.00

## 第 2 章

### 課題 1

### 東京の7月の気温の平均、標準偏差

7月の気温に関するデータです。

平均値 25.1 度

標準偏差 1.58 度

Z 値 2.26

下位から 98.8%

### 課題 2

### 出生率低下の要因を探れ

合計特殊出生率の推定結果です。

被説明変数：合計特殊出生率

	係数	標準誤差	t	P-値
切片	1.462267	0.141852	10.30837	2.6E-13
女性未婚率	-0.02818	0.002841	-9.91716	8.63E-13
有配偶出生率	0.013817	0.000905	15.26972	3.38E-19
重決定 R2	0.913643	F値	232.756	
補正 R2	0.909718	P値 (F値)	3.97E-24	
観測数		47		

### 課題 3

## 消費関数と投資関数

実質民間最終消費と実質民間企業設備投資の推定結果です。

被説明変数：実質民間最終消費

	係数	標準誤差	t	P-値
切片	30480.44	17826.55	1.709834	0.099207
実質GDP	0.497165	0.034986	14.21053	9.07E-14
重決定 R2	0.885934	F値	201.9391	
補正 R2	0.881547	P値 (F値)	9.07E-14	
観測数	28			

被説明変数：実質民間企業設備投資

	係数	標準誤差	t	P-値
切片	92341.83	2218.111	41.63084	2.51E-25
国内銀行貸出約定平均金利	-7047.86	1211.056	-5.8196	3.93E-06
重決定 R2	0.56571	F値	33.86776	
補正 R2	0.549006	P値 (F値)	3.93E-06	
観測数	28			

# 第 3 章

## 課題 1

## 生産関数

実質 GDP の生産関数の推定結果です。Y は実質 GDP、K は資本ストック、L は労働投入量です。

被説明変数： $\log(Y) - k\log(L)$

	係数	標準誤差	t	P-値
切片	-0.43303	0.01359582	-31.8504	9.28E-22
$\log(K) - \log(L)$	0.464581	0.09054576	5.130892	2.66E-05
trend	0.007026	0.00088807	7.912042	2.87E-08
重決定 R2	0.970118	F値	405.8106	
補正 R2	0.967727	P値 (F値)	8.76E-20	
観測数		28		

## 課題 2

## 対数線形

イカとタコの所得弾力性、価格弾力性は以下の通りです。

	イカ	タコ
所得弾力性	4.9	3.0
価格弾力性	- 2.3	- 1.6

推定結果は次ページに掲載しています。

被説明変数：イカの購入数量

	係数	標準誤差	t	P-値
切片	-47.2128	14.94779	-3.15852	0.004942
実質実収入	4.847757	1.162515	4.17006	0.000473
イカ価格掛	-2.2563	0.170201	-13.2567	2.29E-11
重決定 R2	0.910931	F値	102.2725	
補正 R2	0.902024	P値 (F値)	3.14E-11	
観測数	23			

被説明変数：タコの購入数量

	係数	標準誤差	t	P-値
切片	-27.0836	7.935791	-3.41284	0.002758
実質実収入	3.071497	0.620912	4.946754	7.77E-05
タコ価格掛	-1.60862	0.123207	-13.0562	3.02E-11
重決定 R2	0.898712	F値	88.72827	
補正 R2	0.888583	P値 (F値)	1.14E-10	
観測数	23			

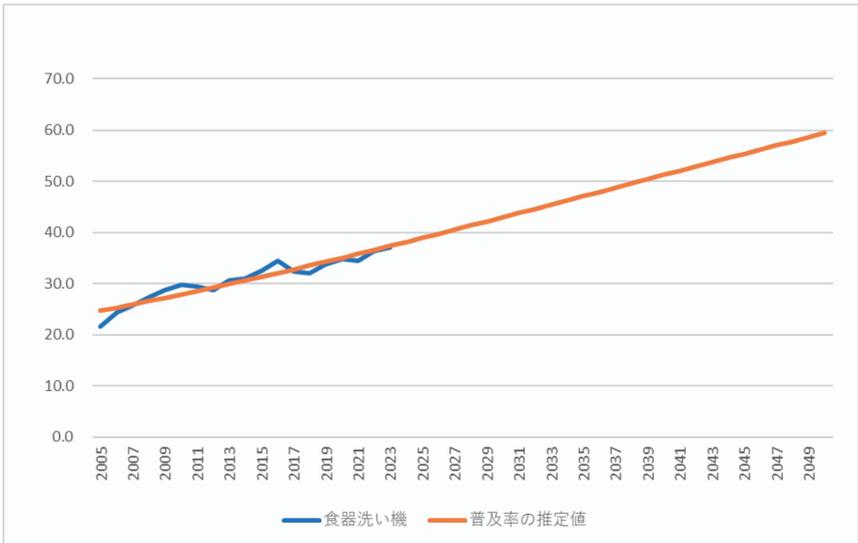
# 課題 3

## ロジスティック曲線

食器洗い機の普及率の推定結果と 2050 年までの予測値です。

被説明変数：食器洗い機普及率 (%)

	係数	標準誤差	t	P-値
切片	-67.9925	5.597389	-12.1472	8.35E-10
西暦	0.033354	0.002779	12.00129	1E-09
重決定 R2	0.89443	144.031		
補正 R2	0.88822	1E-09		
観測数	19			



# 第 4 章

## 課題 1

### 多重共線性

(1) 同じ説明変数が 2 つあるので、 $\beta_1$  と  $\beta_2$  の組み合わせは無限にあり、1 つに決まりません。ただ、Excel では  $\beta_1$  が 0、 $\beta_2$  が  $X_1$  のみで回帰したときの係数が計算されます。

$$\beta_1 \quad 0 \qquad \beta_2 \quad 0.54$$

(2) 本来あるべき  $X_2$  が説明変数にないので、 $\beta_1$  が約 2 倍の値になっています。

$$\beta_1 \quad 0.96 \qquad \beta_2 \quad 0.50$$

(3) 本来あるべき  $X_1$  が説明変数にないので、 $\beta_1$  が約 2 倍の値になっています。 $X_1$  と  $X_2$  の相関が高いので、(2) と (3) の推定結果は似ています。

$$\beta_1 \quad 0.96 \qquad \beta_2 \quad 0.50$$

(4) 相関の高い 2 つの変数をまとめたため、 $\beta_1$  は真の値 (0.5) に近い値が推定されています。

$$\beta_1 \quad 0.49 \qquad \beta_2 \quad 0.50$$

## 課題 2

### 賃金の推定

所定内給与の推定結果を、次ページに掲載しています。

被説明変数：所定内給与

	係数	標準誤差	t	P-値
切片	3.65601	10.55124	0.346501	0.729405
age	9.235369	0.285482	32.35015	1.4E-73
female	68.10906	10.55124	6.455079	1.14E-09
d1000	-70.9183	12.92257	-5.48794	1.5E-07
d100	-25.093	12.92257	-1.9418	0.053855
age*female	-3.30044	0.285482	-11.561	4.71E-23
age*d1000	3.69931	0.349642	10.58028	2.55E-20
age*d100	1.21197	0.349642	3.466318	0.000672
重決定 R2	0.96992	F値	764.6555	
補正 R2	0.968651	P値 (F値)	9.2E-123	
観測数	174			

### 課題3

## 男女賃金に差はあるか？

【1】女性ダミーをみると、有意に女性の賃金の水準が68.1万円高く、年齢×女性ダミーをみると、年間3.3万円昇給のテンポに違いがあります。ゼロ歳時点では女性が68.1万円分速いですが、21歳時点で男性の賃金が高くなり、その後格差が広がっていくことを表しています。

【2】1000万円以上の企業は、ゼロ歳段階では70.9万円安いですが20歳時点で10人～99人の賃金より高くなり、その後年間3.7万円分、10人～99人の企業より昇給テンポが速いです。100人～999人の企業は、ゼロ歳段階では25.1万円10人～99人の企業より安いですが、5%水準で有意ではありません。昇給テンポは、10人～99人に比べて1.2万円分早いです。

# 第 5 章

## 課題 1

### 為替レートの予測

対ドル円レート (階差) の 1 期前の係数は有意ではなく、自由度修正済み決定係数も有意ではないので、予測には使えません。

被説明変数：対ドル円レート(階差)

	係数	標準誤差	t	P-値
切片	0.359447	0.2929617	1.226943	0.222394
X 値 1	0.030755	0.0944545	0.325606	0.745325
重決定 R2	0.000937	F値	0.106019	
補正 R2	-0.0079	P値 (F値)	0.745325	
観測数		115		

## 課題 2

### 日本の株価は米国の株価に影響を受けているか？

日本の株価の推定では米国の株価 1 期前が有意で、米国の株価の推定では日本の株価が有意なので、お互いに影響を受けていることがわかります。

#### ▼ 日本の株価の推定結果

被説明変数：日本の株価

	係数	標準誤差	t	P-値
切片	0.00021322	0.00033319	0.63993442	0.52234041
米国の株価1期前	0.40506953	0.02448569	16.5431154	2.0987E-55
日本株価1期前	-0.11523462	0.0271463	-4.24494702	2.3594E-05
重決定 R2	0.18945464	F値	136.853155	
補正 R2	0.18807027	P値	3.8826E-54	
観測数		1174		

▼米国の株価の推定結果

被説明変数：米国の株価

	係数	標準誤差	t	P-値
切片	0.00037928	0.00037433	1.0132216	0.31116363
米国の株価1期前	-0.36028293	0.02939144	-12.2580898	1.3743E-32
日本の株価	0.44001599	0.03258324	13.5043664	9.9203E-39
重決定 R2	0.16709662	F値	117.462686	
補正 R2	0.16567407	P値 (F値)	3.2222E-47	
観測数	1174			



### 課題3 ランダムウォーク系列を作って回帰してみよう

Excelで2つのランダムウォーク系列を作り、1つを被説明変数、もう1つを説明変数として、回帰分析をします。下記は1つの例です。乱数を発生させているので、同じ結果にはなりません。

被説明変数：Y

	係数	標準誤差	t	P-値
切片	26178.65	210.3888	124.4298	5.84E-40
X	146.8979	12.2136	12.0274	1.41E-12
重決定 R2	0.83783	F値	144.6583	
補正 R2	0.832038	P値(F値)	1.41E-12	
観測数	30			

# 第 6 章

## 課題 1

### ロジスティック回帰

67%なので合格できます。

## 課題 2

### ベイズによる分析

事後確率は以下の通りです。

	1	2	3	4
事後確率	60%	100%	60%	50%

## 課題 3

### 働く力の作成

固有値は1を超えていれば意味があるので、第1主成分を使えば働く力の指標として使えます。固有ベクトルは、変数のウエートを表します。所定内給与が高く、昼夜間人口比率が高く、完全失業率が低いほど働く力が高い自治体といえます。主成分得点をみると働く力の大きいのは東京、愛知県、大阪府となりました。

# 第 7 章

## 課題 1

## 英語のテストと研修の効果 (RCT)

A 社に関しては、全員研修を受けているので、一対の標本による平均値の差の検定を行います。対立仮説を「研修後の方が点数が高い」として、片側検定で検定します。 $P$  値は 11% なので、5% 水準で棄却できません。つまり、研修に効果があったとは言えないことになります。

t-検定: 一対の標本による平均の検定ツール

	研修前	研修後
平均	78.6	88
分散	397.3778	63.11111
観測数	10	10
ピアソン相関	-0.19505	
仮説平均との差異	0	
自由度	9	
t	-1.30071	
$P(T \leq t)$ 片側	0.112836	
t 境界値 片側	1.833113	
$P(T \leq t)$ 両側	0.225672	
t 境界値 両側	2.262157	

B 社に関しては、研修を受けてない人に関するデータもあるので、より正確に効果が測定できます。両者のデータを分散が等しくないと仮定した 2 標本による検定をします。 $P$  値は 38% なので、研修に効果があったという仮説は棄却できません。

t-検定: 分散が等しくないと仮定した2標本による検定

---

平均	13.7	11.1
分散	87.78889	617.2111
観測数	10	10
仮説平均との差異	0	
自由度	12	
t	0.309655	
P(T<=t) 片側	0.381067	
t 境界値 片側	1.782288	
P(T<=t) 両側	0.762134	
t 境界値 両側	2.178813	

---



## 課題 2

## 少子化対策 (傾向スコアマッチング)

---

自治体のタイプをそろえて、一対の標本による平均値の差の検定を行います。帰無仮説は「少子化対策の効果の大きさがゼロ」です。P値は片側検定で  $8.42 \times e^{-6}$  非常に小さいので、帰無仮説が棄却できます。

t-検定: 一対の標本による平均の検定ツール

---

	変数 1	変数 2
平均	11.42857	9.142857
分散	18.95238	16.14286
観測数	7	7
ピアソン相関	0.996414	
仮説平均との差異	0	
自由度	6	
t	12.39355	
P(T<=t) 片側	8.42E-06	
t 境界値 片側	1.94318	
P(T<=t) 両側	1.68E-05	
t 境界値 両側	2.446912	

---

### 課題 3

## 就学年数と年収（操作変数法）

操作変数を使うと就学年数の係数 (0.29) が正しく推定できています。

(1) 現実には不可能な、個人の能力のデータが入手できた場合の推定

$$\text{年収} = \alpha + \beta_1 \times \text{修学年数} + \beta_2 \times \text{個人の能力}$$

被説明変数：年収

	係数	標準誤差	t	P-値
切片	103.6494	24.12896	4.295644	0.003586
就学年数	0.28664	0.233994	1.224991	0.260196
本人の能力	4.685165	0.3284	14.26663	1.98E-06
重決定 R2	0.985234	F値	233.535	
補正 R2	0.981015	P値 (F値)	3.91E-07	
観測数	10			

(2) 操作変数を使った 1 段階目の推定

$$\text{就学年数} = \alpha + \beta \times \text{大学までの距離}$$

被説明変数：就学年数

	係数	標準誤差	t	P-値
切片	104.5	2.38157	43.87862	8.03E-11
大学までの距離	2	0.71807	2.785242	0.023732
重決定 R2	0.492308	F値	7.757576	
補正 R2	0.428846	P値 (F値)	0.023732	
観測数	10			

(3) 操作変数を使った2段階目の推定

$$\text{年収} = \alpha + \beta \times \text{修学年数の推定値}$$

被説明変数：年収

	係数	標準誤差	t	P-値
切片	593.2491	198.7907	2.98429	0.017486
就学年数（推定値）	0.28664	1.798422	0.159384	0.877317
重決定 R2	0.003165	F値	0.025403	
補正 R2	-0.12144	P値（F値）	0.877317	
観測数		10		



課題の解答はこれで終わりです  
お疲れさまでした！

