

『わかる統計学』練習問題の解答

1 章

1 学年：順序尺度 学籍番号：名義尺度 性別：名義尺度 出生年：間隔尺度
血液型：名義尺度 健康状態：順序尺度 身長：比尺度 体重：比尺度 体温：
間隔尺度

2 章

1 ①× ②× ③○ ④×

2 58 kg

3 150～159 mmHg

4 ③

5

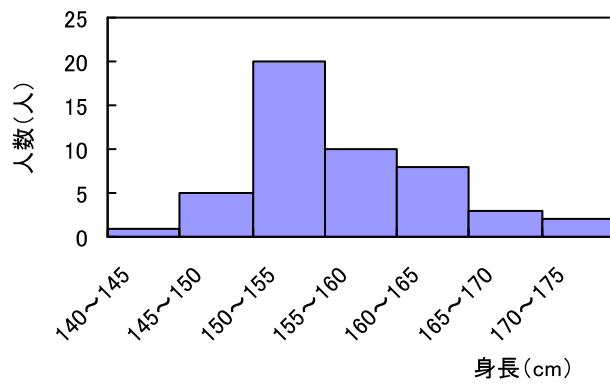
身長の度数分布

	度数 (%)	累積度数 (%)
140～145	1 (2.0)	1 (2.0)
145～150	5 (10.2)	6 (12.2)
150～155	20 (40.8)	26 (53.1)
155～160	10 (20.4)	36 (73.5)
160～165	8 (16.3)	44 (89.9)
165～170	3 (6.1)	47 (95.2)
170～175	2 (4.1)	49 (100.0)
計	49 (100.0)	

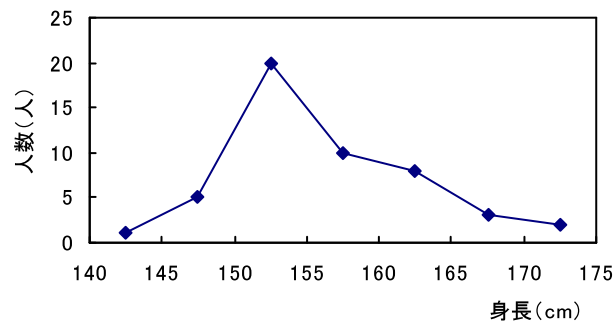
身長の幹葉表示（少数部を四捨五入したもの）

```

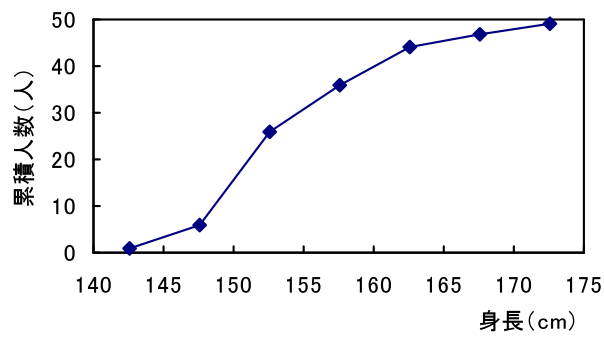
14* | 0
14# | 6 6 8 9 9
15* | 0 0 0 1 1 1 1 1 2 2 2 3 3 3 3 4 4
15# | 5 5 5 5 5 6 7 7 7 8 8
16* | 0 0 0 1 2 2 3 3 4
16# | 5 7 8 8
17* | 1 1
  
```



身長のヒストグラム



身長の度数多角形



身長の累積度数曲線

身長

25 パーセンタイル : 151.05 cm

50 パーセンタイル : 154.70 cm

75 パーセンタイル : 160.75 cm

体重の度数分布

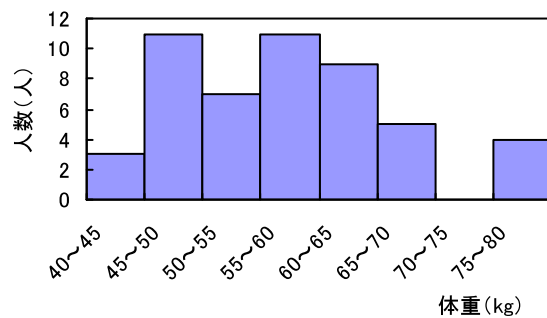
	度数 (%)	累積度数 (%)
40～45	3 (6.0)	3 (6.0)
45～50	11 (22.0)	14 (28.0)
50～55	7 (14.0)	21 (42.0)
55～60	11 (22.0)	32 (64.0)
60～65	9 (18.0)	41 (82.0)
65～70	5 (10.0)	46 (92.0)
70～75	0 (0.0)	46 (92.0)
75～80	4 (8.0)	50 (100.0)
計	50 (100.0)	

体重の幹葉表示（少数部を四捨五入したもの）

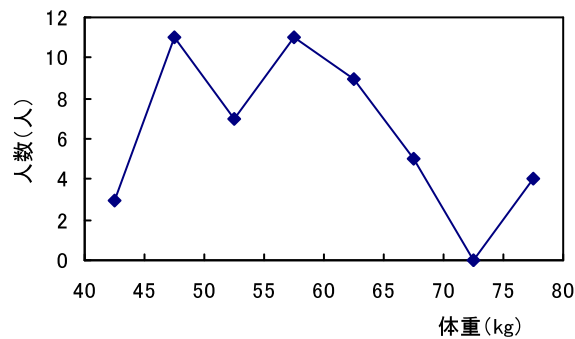
```

4* | 2 3
4# | 5 6 7 7 7 7 7 8 9 9 9
5* | 3 3 3 3 4 4
5# | 5 5 5 6 6 7 7 7 8 8 9 9
6* | 1 2 2 2 3 3 3 3 4
6# | 6 6 7 7 7
7* |

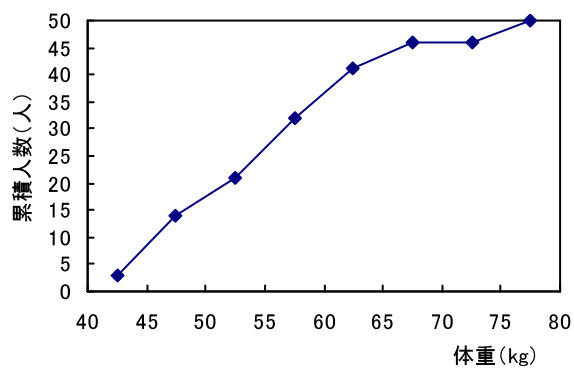
```



体重のヒストグラム



体重の度数多角形



体重の累積度数曲線

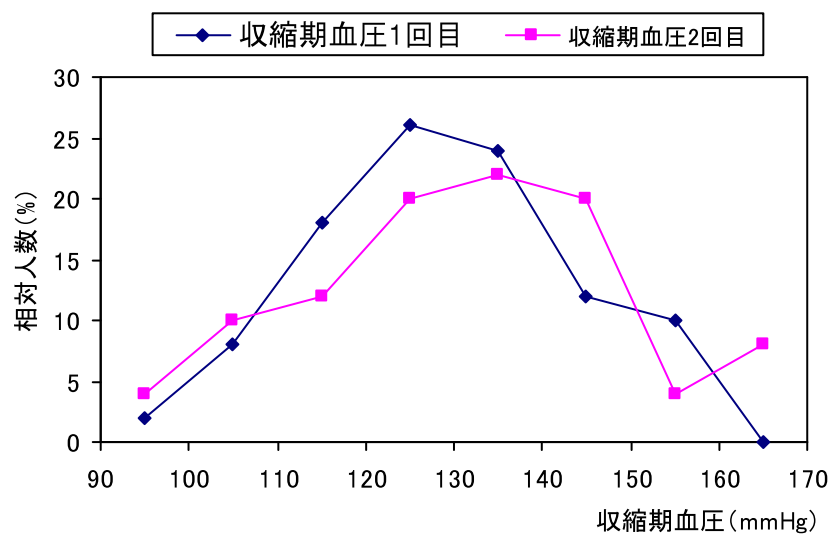
体重

25 パーセンタイル : 48.75 kg

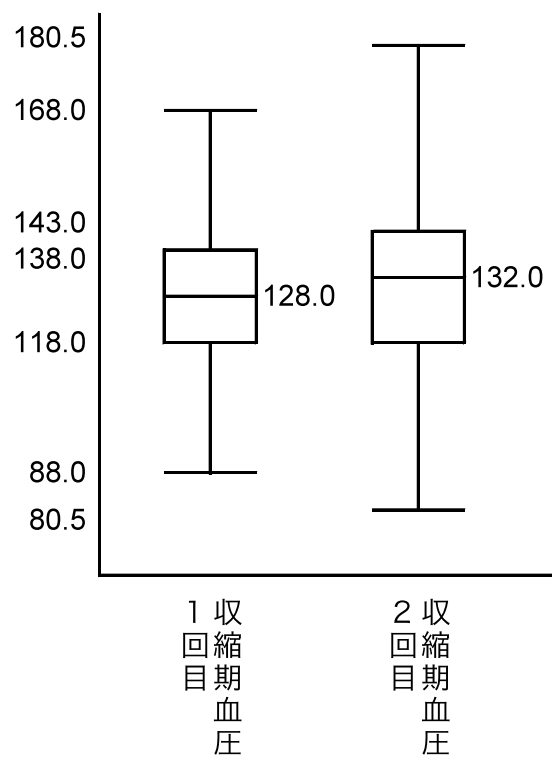
50 パーセンタイル : 56.25 kg

75 パーセンタイル : 63.00 kg

6



収縮期血圧の度数多角形



収縮期血圧の箱ヒゲ図

7 ③

3 章

1 ①

2 ①

3 ④

4 ③

5

	平均値	中央値	最頻値	分散	不偏分散
年齢	57.7	58.0	61.0	82.6	84.3
身長	155.9	154.7	154.7	45.3	46.2
体重	57.0	56.3	47.0	82.3	84.0
収縮期血圧1回目	127.6	128.0	126.0	220.8	225.3
収縮期血圧2回目	130.8	132.0	136.0	297.0	303.0
拡張期血圧1回目	80.4	81.0	78.0	92.9	94.8

6

- ・元のデータを a 倍した場合：平均値は a 倍，分散は a^2 倍になる．
- ・元のデータに b を加えた場合：平均値は b だけ増えるが，分散は変わらない．

7 比尺度のみ

8 ①

9 ① ④

4 章

1 ③

2 ③

3 ① ③

4 ①

5 ①

6 ③

7

	疾病A既往状況		計
	既往あり	既往なし	
男	7 30.4	16 69.6	23 100.0
女	2 7.4	25 92.6	27 100.0
計	9 18.0	41 82.0	50 100.0

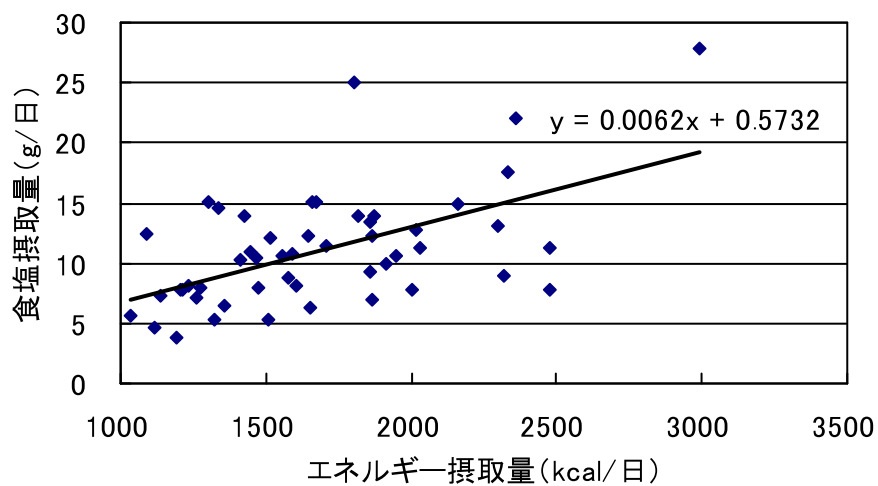
上段：人数

下段：相対度数 (%)

クラメールの関連係数：0.2987

ファイ係数：0.2987

8



エネルギー摂取量と食塩摂取量の相関図および回帰直線

各変数間の相関係数

	食塩摂取量	食塩摂取量の期待値	残 差	エネルギー摂取量
食塩摂取量	1.000			
食塩摂取量の期待値	0.556	1.000		
残 差	-0.833	0.000	1.000	
エネルギー摂取量	0.556	1.000	0.000	1.000

5 章

- 1 ④
- 2 ④
- 3 ③
- 4 二項分布に従うので, $p = 0.4$, $q = 0.6$, $n = 10$, $x = 8$ として, 式 (5.1) を用いる.

$${}_{10}C_8 \times 0.4^8 \times 0.6^2 = 0.0106$$

- 5 ① 1587 人
- ② $\text{NORMSDIST}(-1.5) = 0.0668$
 $10,000 \times 0.0668 = 668$ (人)
- ③ 6826 人

6 章

- 1 ②
- 2 ③
- 3 ④
- 4 ①
- 5 医師 10 人, 看護師 20 人, 薬剤師 2 人, 栄養士 1 人を無作為に抽出する.

7 章

- 1 ⑤
- 2 式 (7.3) を用いる.

$$30.2 \pm 1.96 \times \sqrt{\frac{0.5}{20}}$$

したがって, その日製造された特定保健用食品の関与成分量の平均値は, 29.9 g から 30.5 g の間にあるといえる.

8 章

1

① 帰無仮説（一等が 10,000 本に 1 本）のもとで、まれなことが起こったのであるから、帰無仮説を棄却することになり、一等が当たる確率がもっと大きいと考えられる。もしかすると全部当たりくじ？

② 貨幣を 20 回投げて、表が 15 回以上出るか、表が 5 回以下しか出ない確率は、あわせて 4.2%。したがって、5% の危険率で、表が出る確率は 1/2 という仮説は棄却される。

③ その町にはきっと、多くの日本人が訪れているからである。

④ やはり 11 日目も雨が降るであろう。

2 ④

3 ②

9 章

1 ④

2 独立性の検定（イエーツの連続性補正をした χ^2 検定）を用いる。

【A食品】

	発 病	非発病	計
食べた	16	12	28
食べない	14	18	32
計	30	30	60

検定統計量 χ_{0C}^2 : 0.603

自由度 : 1

有意確率 : 0.4376

【B食品】

	発 病	非発病	計
食べた	17	14	31
食べない	13	16	29
計	30	30	60

検定統計量 χ_{0C}^2 : 0.267

自由度 : 1

有意確率 : 0.6054

【C食品】

	発 病	非発病	計
食べた	22	8	30
食べない	8	22	30
計	30	30	60

検定統計量 χ_{0C}^2 : 11.267

自由度 : 1

有意確率 : 0.0008

【D食品】

	発 病	非発病	計
食べた	9	12	21
食べない	21	18	39
計	30	30	60

検定統計量 χ_{0C}^2 : 0.293

自由度 : 1

有意確率 : 0.5883

【E食品】				検定統計量 χ^2_{0c} : 0.000 自由度: 1 有意確率: 1.000
	発 病	非発病	計	
食べた	10	11	21	
食べない	20	19	39	
計	30	30	60	

食品の摂取と食中毒発症との間に有意な関連が認められたのは、C 食品のみであり、そのほかの食品では関連が認められなかった。すなわち、C 食品が食中毒の原因食品である可能性が高い。

10 章

1 母比率の検定を用いる.

帰無仮説 H_0 : 「広報活動後の認知度は 20% に等しい ($p = p_0$)」であり、対立仮説 H_1 : 「広報活動後の認知度は 20% より高い ($p > p_0$)」である。よって片側検定を行う。標本サイズ $n = 50$ (人), 標本比率 $p = 20/50 = 0.4$, 既知の値 = 0.2 であり, 50 (標本サイズ) > 50 であるから, 正規分布で近似できる。したがって, 式 (10.3) より $Z_{0c} = 3.359 > 1.64 = Z(0.05)$ となり, 帰無仮説が棄却される。したがって, 有意水準 0.05 で, 広報活動後の認知度は 20% より高いといえ, 広報活動はそれなりに知名度を高める効果があったといえる。

2 2 群の比率の差の検定であり, 独立性の検定と同等とみなす方法を適用する.

式 (9.5) から

検定統計量 χ^2_{0c} : 55.663

自由度: 3

有意確率: 0.00000

男女間で, 来診者の年齢構成は有意に異なる。

3 以下の表にして, 独立性の検定を適用する.

	罹患者	非罹患者	計
接 種 群	15	106	121
非接種群	49	48	97
計	64	154	218

検定統計量 χ^2_{0c} : 35.906

自由度: 1

有意確率: 0.00000

予防接種を受けた群のほうが, 受けなかった群より有意に罹患割合が低い。

正規分布による方法（式 10.14）を用いることもでき，同じ結果となる．

11 章

1 ②

2 ④

3 検定統計量は自由度 48 の t 分布に従う．

タンパク質摂取量	相関係数	0.768			
	検定統計量 t_0	8.309			
	有意確率	0.00000			
脂肪摂取量	相関係数	0.802	0.587		
	検定統計量 t_0	9.317	5.019		
	有意確率	0.00000	0.00001		
カルシウム摂取量	相関係数	0.664	0.610	0.483	
	検定統計量 t_0	6.149	5.330	3.818	
	有意確率	0.00000	0.00000	0.00039	
食塩摂取量	相関係数	0.556	0.583	0.307	0.337
	検定統計量 t_0	4.632	4.973	2.233	2.478
	有意確率	0.00003	0.00001	0.03022	0.01678
	エネルギー 摂取量	タンパク質 摂取量	脂肪 摂取量	カルシウム 摂取量	

信頼区間を求めるには式（11.4）～（11.6）を用いればよい．

0.768 : 0.750～0.785

0.802 : 0.787～0.816

0.587 : 0.559～0.614

0.664 : 0.640～0.687

0.610 : 0.583～0.636

0.483 : 0.450～0.514

0.556 : 0.527～0.584

0.583 : 0.555～0.610

0.307 : 0.269～0.344

0.337 : 0.300～0.373

4

・回帰式

カルシウム摂取量 = $0.3560 \times$ エネルギー摂取量 $- 132.75$

・回帰係数に関する検定結果

検定統計量 t_0 : 4.598

自由度 : 48

有意確率 : 0.00003

回帰係数は 0 ではないと判定される.

12 章

1 ②

2 ③

3 ②

4 ④

5 ②

6 2 群の平均値の差の検定を行う.

両群の母分散は未知であるから, まず等分散の検定を行う.

	標本サイズ	平均値	不偏分散	不偏標準偏差
M部屋	10	33.2	65.060	8.066
N部屋	10	19.6	286.253	16.919
全 体	20	26.4	215.092	14.666

検定統計量 $K_0 = \frac{286.253}{65.060} = 4.400$: 自由度 (9, 9) の F 分布に従う.

有意確率 P は 0.0379 であり, 母分散は異なることから, ウェルチの検定を用いる.

検定統計量 t_0 : 2.294

自由度 : 12.89

有意確率 : 0.0406

2 部屋の落下細菌数には有意な差が認められる.

7 対応のある 2 標本の平均値の差の検定を用いる.

A 施設と B 施設の評価点の差の平均値 : -3.400

差の標準偏差 : 4.326

検定統計量 t_0 : 2.486

自由度 : 9

有意確率 : 0.0347

二つの給食施設の評価点には有意な差が認められる.

8 2群の平均値の差の検定を行う.

両群の母分散は未知であるから, まず等分散の検定を行う.

検定統計量 F_0 : 1.380

自由度: (4, 3)

有意確率: 0.8240

等分散であるとしてよいから, スチューデントの t 検定を用いる.

検定統計量 t_0 : 0.942

自由度: 7

有意確率: 0.3776

培地によってコロニーの大きさに差があるとはいえない.

13 章

1 一元配置分散分析を用いる.

【収縮期血圧 1 回目】

	人数	平均値	不偏標準偏差
30~49歳	7	129.857	10.040
50~59歳	13	118.846	13.637
69歳~	7	133.714	14.209
全 体	27	125.556	14.181

分散分析表

変動の 要因	偏差 平方和 (変動)	自由度	平均平方 (不偏分散)	F値	P
級 間	1180.689	2	590.344	3.500	0.0464
級 内	4047.978	24	168.666		
全 体	5228.667	26	201.193		

1 回目の収縮期血圧に有意な年齢間差が認められる.

【収縮期血圧 2 回目】

	人数	平均値	不偏標準偏差
30~49歳	7	135.714	9.552
50~59歳	13	121.538	14.564
69歳~	7	141.429	23.316
全 体	27	130.370	18.000

分散分析表

変動の 要因	偏差 平方和 (変動)	自由度	平均平方 (不偏分散)	F値	P
級 間	2069.923	2	1034.961	3.909	0.0339
級 内	6354.374	24	264.766		
全 体	8424.296	26	324.011		

2 回目の収縮期血圧に有意な年齢間差が認められる.

【拡張期血圧】

	人数	平均値	不偏標準偏差
30～49歳	7	78.000	11.464
50～59歳	13	76.923	9.368
69歳～	7	86.000	2.619
全 体	27	79.556	9.635

分散分析表

変動の 要因	偏差 平方和 (変動)	自由度	平均平方 (不偏分散)	F 値	P
級 間	397.744	2	198.872	2.263	0.1258
級 内	2108.923	24	87.872		
全 体	2506.667	26	96.410		

拡張期血圧に有意な年齢間差は認められない.

2 一元配置分散分析を用いる.

【肝重量】

	人数	平均値	不偏標準偏差
A餌	4	3.753	0.236
B餌	5	3.420	0.166
C餌	3	3.757	0.139
全 体	12	3.615	0.242

分散分析表

変動の 要因	偏差 平方和 (変動)	自由度	平均平方 (不偏分散)	F 値	P
級 間	0.326	2	0.163	4.637	0.0413
級 内	0.316	9	0.035		
全 体	0.642	11	0.058		

餌の種類によって肝重量に有意な差が認められる.

3 クラスカル・ウォリス検定を用いる.

講義担当者	A	B	C
順	1	4	5
	2	7	9
	3	8	10
位	6		11
順位和	12	19	35
人 数	4	3	4
平均順位	3.000	6.333	8.750

検定統計量 $K_0 = 6.053$ (自由度 2 の χ^2 分布に従う)

$P = 0.0485$

担当講師によって試験成績が有意に異なる.