

付録 C 解答例

C.1 確率問題 解答例

A.2 (誕生日問題)

```
(1) Plot[x/365, {x, 1, 365}]
(2) f[N_] := 365!/(365 - N)!
    Plot[1 - f[x]/(365^x), {x, 1, 60}]
    Plot[1 - f[x]/(365^x), {x, 22, 23}]
```

A.3 (サイコロをふる)

```
(4) tmax = 1000;
    t1 = Table[Random[Integer, {1, 6}], {tmax}];
    t2 = Table[Sum[t1[[i]], {i, 1, j}]/j, {j, tmax}];
    plot1 = ListPlot[t2, PlotJoined -> True, PlotRange -> {3, 4}]
    N[t2[[tmax]]]
```

A.4 (2 項分布)

```
(1) binomialdist[n_, p_, k_] := ( n! / (n-k)! / k!) p^k (1-p)^(n-k)
(2) Do[Print[binomialdist[10, 1/2, k]], {k, 0, 10}]
(3) Plot[binomialdist[10, 1/2, k], {k, 0, 10}]
```

A.5 (標準正規分布)

```
(1) normaldist[x_] := Exp[-x^2/2]/Sqrt[2 Pi]
    Plot[normaldist[x], {x, -5, 5}]
(2) bunpu[a_, b_] := NIntegrate[normaldist[x], {x, a, b}]
(3) bunpu[(125 - 100)/15, 1000]
(4) bunpu[(200 - 100)/15, 1000]
```

C.2 統計問題 解答例（2022 年版データ）

B.1 平均 $\bar{x} = \frac{1}{234} \sum_{i=1}^{234} x_i$, 分散 $S_x^2 = \frac{1}{234} \sum_{i=1}^{234} (x_i - \bar{x})^2$, 標準偏差 $\sigma_x = \sqrt{S_x^2}$ などから求める.

	標本平均	標本分散	標準偏差
父親	$\bar{x} = 170.75$	$S_x^2 = 32.08$	$\sigma_x = 5.66$
母親	$\bar{y} = 158.57$	$S_y^2 = 31.14$	$\sigma_y = 5.58$
息子	$\bar{z} = 170.75$	$S_z^2 = 35.26$	$\sigma_z = 5.94$
娘	$\bar{w} = 159.32$	$S_w^2 = 39.36$	$\sigma_w = 6.27$

B.2 例えば父親と母親の相関係数 r_{xy} は,

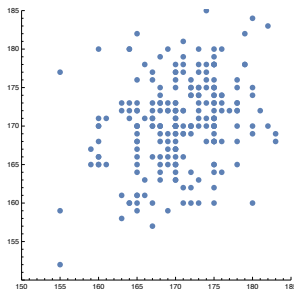
$$S_{xx} = \sum_{i=1}^{234} (x_i - \bar{x})^2 = 7511.25, \quad S_{yy} = \sum_{i=1}^{234} (y_i - \bar{y})^2 = 7287.81,$$

$$S_{xy} = \sum_{i=1}^{234} (x_i - \bar{x})(y_i - \bar{y}) = 1400.95 \text{ より, } r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = 0.189.$$

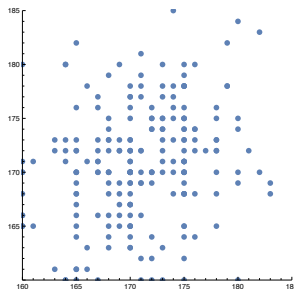
表形式にして相関係数をまとめると, 次のようになる.

	父親 x	母親 y	息子 z	娘 w
父親 x	—	0.189	0.346	0.377
母親 y		—	0.359	0.359
息子 z			—	0.441
娘 w				—

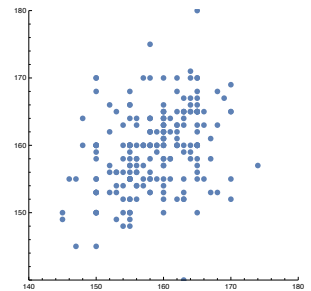
父親と母親のデータ分布図



父親と息子のデータ分布図



母親と娘のデータ分布図



B.3 (1) 例えば, 父親 x と息子 z の身長データの回帰直線 $z = ax + b$ は, (付録 B.3) より, $a = S_{xz}/S_{xx} = 0.363$, $b = \bar{z} - a\bar{x} = 108.81$ となる.

回帰直線は, 次のようになる.

$$(\text{父親 } x, \text{ 母親 } y) \quad y = +0.187x + 126.72$$

$$(\text{父親 } x, \text{ 息子 } z) \quad z = +0.363x + 108.81 \quad \text{図参照}$$

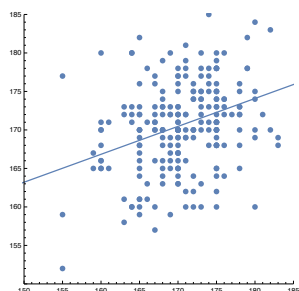
$$(\text{父親 } x, \text{ 娘 } w) \quad w = +0.417x + 88.18$$

$$(\text{母親 } y, \text{ 息子 } z) \quad z = +0.382y + 110.16$$

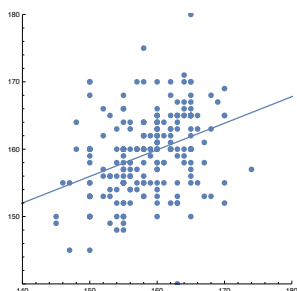
$$(\text{母親 } y, \text{ 娘 } w) \quad w = +0.395y + 96.76 \quad \text{図参照}$$

$$(\text{息子 } z, \text{ 娘 } w) \quad w = +0.465z + 79.93 \quad \text{図参照}$$

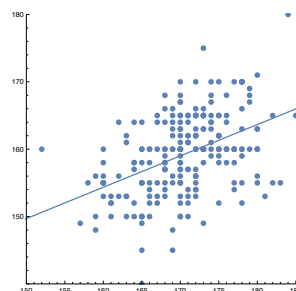
父親と息子のデータ分布図



母親と娘のデータ分布図



息子と娘のデータ分布図



- (2) 父親 x と母親 y の身長に対する息子の身長 z の回帰方程式は, (付録 B.4) にならって, $z = a_1x + a_2y + b$ とおける. (付録 B.5) より, 係数を定めると,

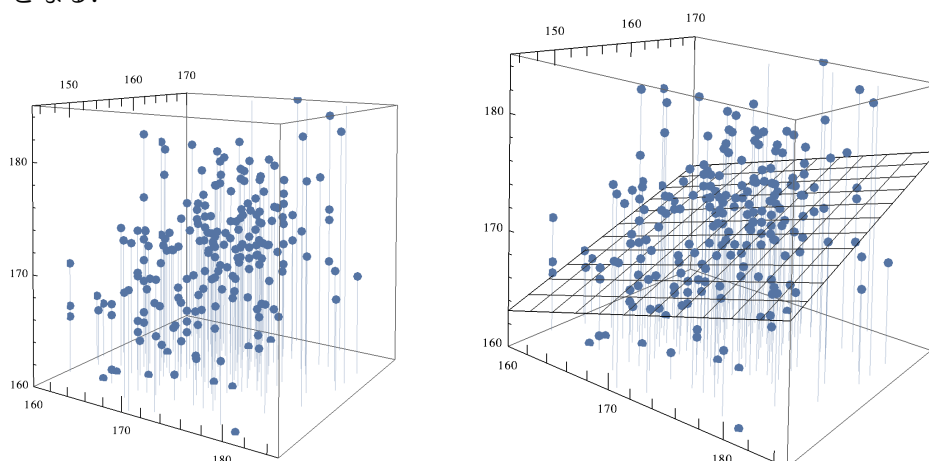
$$z = 0.302x + 0.324y + 67.78 \dots\dots\dots (\#)$$

となる. (x_i, y_i, z_i) の 3 次元データプロットを下図に示す. 得られた回帰方程式は, (x, y, z) 空間内の平面の方程式である. データの点に平面を重ねた図も示す. データのちらばり具合を代表する平面になっていることがわかる.

父親 x と母親 y の身長に対する娘の身長 w の回帰方程式は, 同様に

$$w = 0.356x + 0.326y + 46.84 \dots\dots\dots (b)$$

となる.



- (2) で得られた (#) は父親の身長 x と母親の身長 y から息子の身長 z を予測する式, (b) は娘の身長 w を予測する式であるといえる. サンプル数が少ないので誤差は多いと考えられるが, このデータからは, 息子も娘も父親の影響を強く受けて身長が決まるといった結果になった.
- ちなみに, web で「身長 予測 Height Predictor」などで検索すると, 次のような式が見つかる.

$$\begin{cases} \text{息子の身長} &= (\text{父親の身長} + \text{母親の身長}) \times 1.08 / 2 \\ \text{娘の身長} &= (\text{父親の身長} \times 0.923 + \text{母親の身長}) / 2 \end{cases}$$

$$\begin{cases} \text{息子の身長} &= (\text{父親の身長} + \text{母親の身長} + 13) / 2 \\ \text{娘の身長} &= (\text{父親の身長} + \text{母親の身長} - 13) / 2 \end{cases}$$

はてさて真偽は??