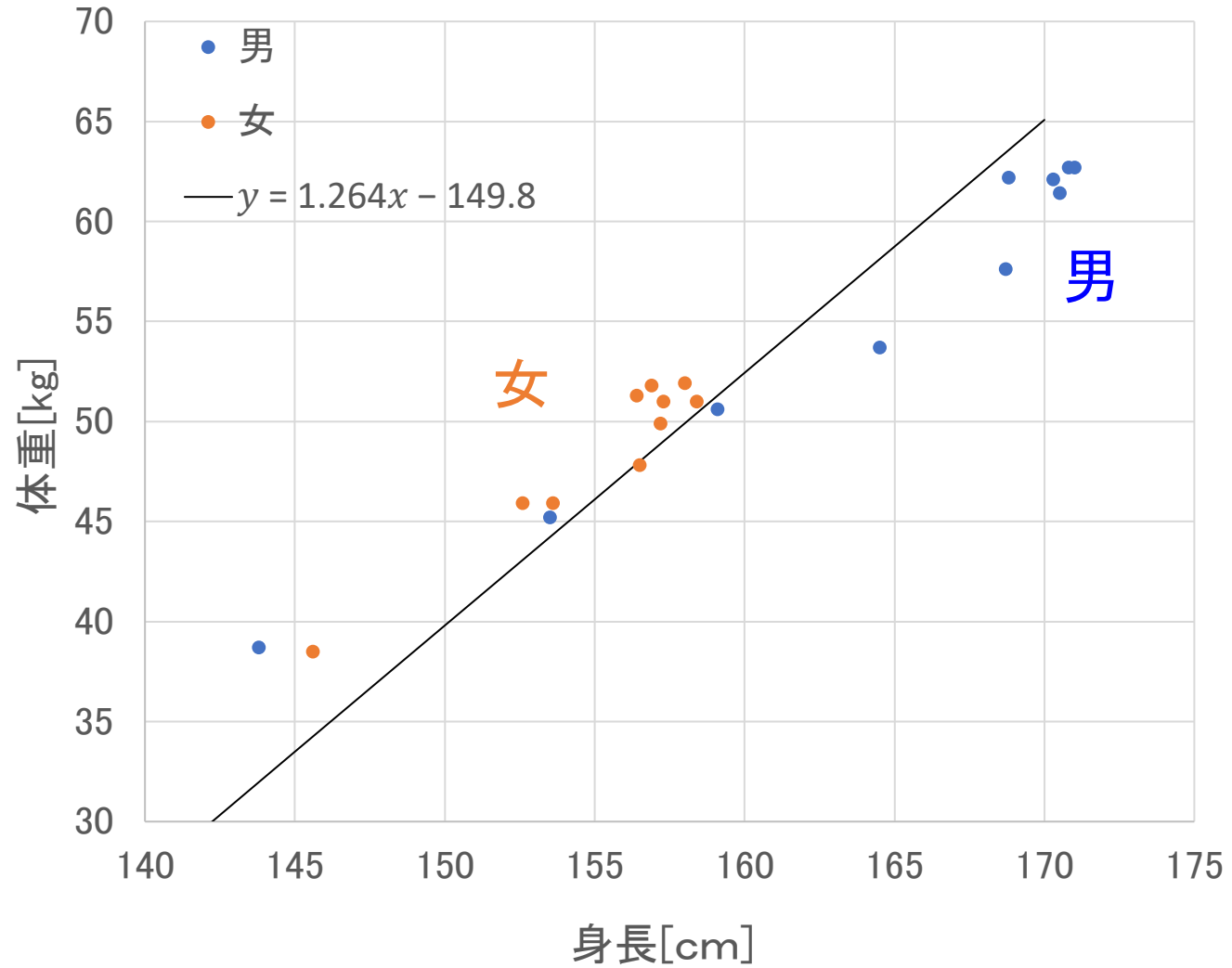
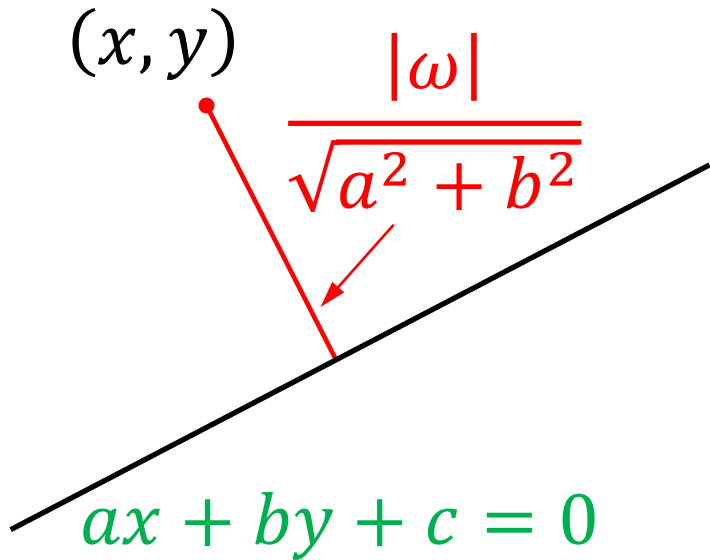


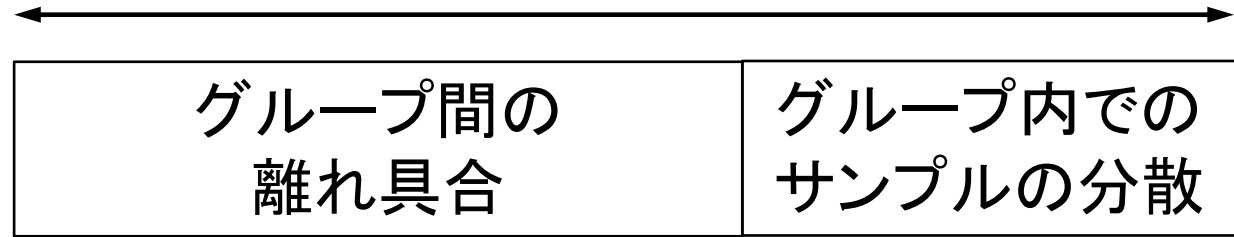
判別分析

男女を判別する関数を求める





全体の分散



$$F = \frac{\text{グループ間の離れ具合}}{\text{全体の分散}} = \frac{n_A(\bar{\omega}_A - \bar{\omega})^2 + n_B(\bar{\omega}_B - \bar{\omega})^2}{\sum_{i=1}^n (\omega_i - \bar{\omega})^2}$$

$$\begin{aligned}
 s^2 &= \frac{1}{n-1} \left\{ \sum_{i=1}^n (\omega_i - \bar{\omega})^2 \right\} \\
 &= \frac{1}{n-1} \left\{ \sum_{i=A}^n (\omega_i - \bar{\omega}_A + \bar{\omega}_A + \bar{\omega})^2 + \sum_{i=B}^n (\omega_i - \bar{\omega}_B + \bar{\omega}_B + \bar{\omega})^2 \right\} \\
 &= \frac{1}{n-1} \left\{ [n_A(\bar{\omega}_A - \bar{\omega})^2 + n_B(\bar{\omega}_B - \bar{\omega})^2] + [\sum_{i=A}^n (\omega_i - \bar{\omega}_A)^2 + \sum_{i=B}^n (\omega_i - \bar{\omega}_B)^2] \right\}
 \end{aligned}$$

AグループとBグループの離れ具合

各グループにおけるサンプルの分散

$$F = \frac{\text{グループ間の離れ具合}}{\text{全体の分散}} = \frac{n_A(\bar{\omega}_A - \bar{\omega})^2 + n_B(\bar{\omega}_B - \bar{\omega})^2}{\sum_{i=1}^n (\omega_i - \bar{\omega})^2}$$

$$= \frac{n_A\{a(\bar{x}_A - \bar{x}) + b(\bar{y}_A - \bar{y})\}^2 + n_B\{a(\bar{x}_B - \bar{x}) + b(\bar{y}_B - \bar{y})\}^2}{\sum_{i=1}^n \{a^2(x_i - \bar{x})^2 + 2ab(x_i - \bar{x})(y_i - \bar{y}) + a^2(y_i - \bar{y})^2\}}$$

$$= \frac{n_A\{a(\bar{x}_A - \bar{x}) + b(\bar{y}_A - \bar{y})\}^2 + n_B\{a(\bar{x}_B - \bar{x}) + b(\bar{y}_B - \bar{y})\}^2}{(n-1)(a^2s_x^2 + 2abs_{xy} + b^2s_y^2)}$$

$$= \frac{n_A(aX_A + bY_A)^2 + n_B(aX_B + bY_B)^2}{(n-1)(a^2s_x^2 + 2abs_{xy} + b^2s_y^2)}$$

$$= \frac{(n_A X_A^2 + n_B X_B^2)a^2 + 2(n_A X_A Y_A + n_B X_B Y_B)ab + (n_A X_A^2 + n_B X_B^2)b^2}{(n-1)(a^2s_x^2 + 2abs_{xy} + b^2s_y^2)}$$

$$= \frac{(n_A X_A^2 + n_B X_B^2)t^2 + 2(n_A X_A Y_A + n_B X_B Y_B)t + (n_A Y_A^2 + n_B Y_B^2)}{(n-1)s_x^2 t^2 + 2(n-1)s_{xy}t + (n-1)s_y^2} = \frac{A_1 t^2 + A_2 t + A_3}{B_1 t^2 + B_2 t + B_3}$$

$\frac{a}{b} = t$ とおく

$$\frac{\partial F}{\partial t} = \frac{\{(2A_1 t + A_2)(B_1 t^2 + B_2 t + B_3) - (A_1 t^2 + A_2 t + A_3)(2B_1 t + B_2)\}}{(B_1 t^2 + B_2 t + B_3)^2}$$

$$= \frac{(A_1 B_2 - A_2 B_1)t^2 + 2(A_1 B_3 - A_3 B_1)t + (A_2 B_3 - A_3 B_2)}{(B_1 t^2 + B_2 t + B_3)^2}$$

$\frac{\partial F}{\partial t} = 0$ のとき

$$\omega = tx + y + \frac{c}{b}$$

直線がグループA,Bの中心を通る

$$\frac{c}{b} = -t \frac{\bar{x}_A + \bar{x}_B}{2} - \frac{\bar{y}_A + \bar{y}_B}{2}$$

$$n = 20 \quad n_A = 10 \quad n_B = 10$$

$$A_1B_2 - A_2B_1 = 80005.069$$

$$A_2B_3 - A_3B_2 = 82176.28$$

$$\bar{x} = 159.68 \quad \bar{y} = 52.095$$

$$\bar{x}_A = 164.1 \quad \bar{y}_A = 55.69$$

$$\bar{x}_B = 155.25 \quad \bar{y}_B = 48.5$$

$$S^2_x = 67.912 \quad S^2_y = 55.99 \quad S_{xy} = 60.55$$

$$X_A = \bar{x}_A - \bar{x} = 4.425 \quad Y_A = \bar{y}_A - \bar{y} = 3.595$$

$$X_B = \bar{x}_B - \bar{x} = -4.425 \quad Y_B = \bar{y}_B - \bar{y} = -3.595$$

$$A_1 = n_A X_A^2 + n_B X_B^2 = 391.61 \quad A_2 = 2(n_A X_A + n_B X_B) = 636.31 \quad A_3 = n_A Y_A^2 + n_B Y_B^2 = 258.48$$

$$B_1 = (n - 1)S^2_x = 1290.3 \quad B_2 = 2(n - 1)S_{xy} = 2300.9 \quad B_3 = (n - 1)S^2_y = 1063.8$$

$$\frac{(A_1B_2 - A_2B_1)t^2 + 2(A_1B_3 - A_3B_1)t + (A_2B_3 - A_3B_2)}{(B_1t^2 + B_2t + B_3)^2} = 0$$

$$\frac{80006t^2 + 166148t + 82176}{(B_1t^2 + B_2t + B_3)^2} = 0 \text{ において } (B_1t^2 + B_2t + B_3)^2 \neq 0 \text{ のとき } 80006t^2 + 166148t + 82176 = 0$$

2次方程式の解の公式を用いて、tを求める

$$t_1 = -0.812429$$

$$t_2 = -1.264266$$

$$F(t_1) = 0$$

$$F(t_2) = 0.3679664$$

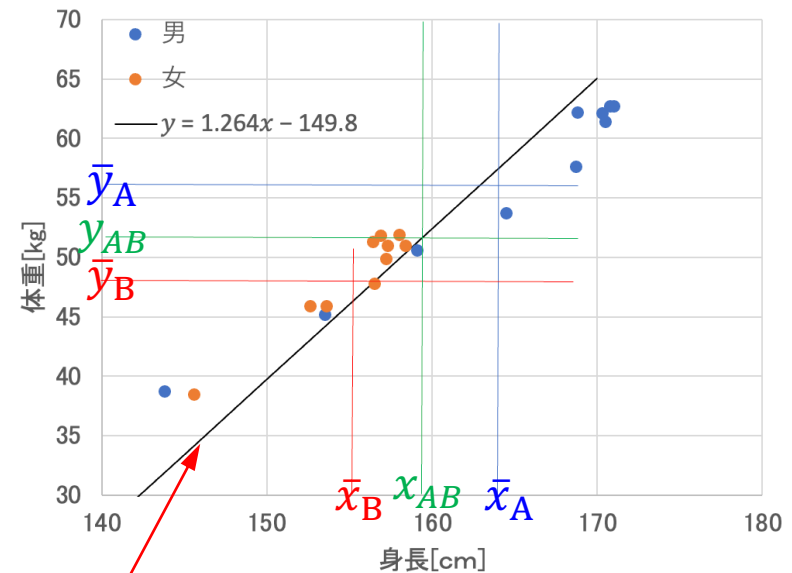
最大値になるのは、 $t = \frac{a}{b} = -1.264266$ (次ページ参照)

$$\omega = \frac{a}{b}x + y + \frac{c}{b} = -1.264x + y + \frac{c}{b}$$

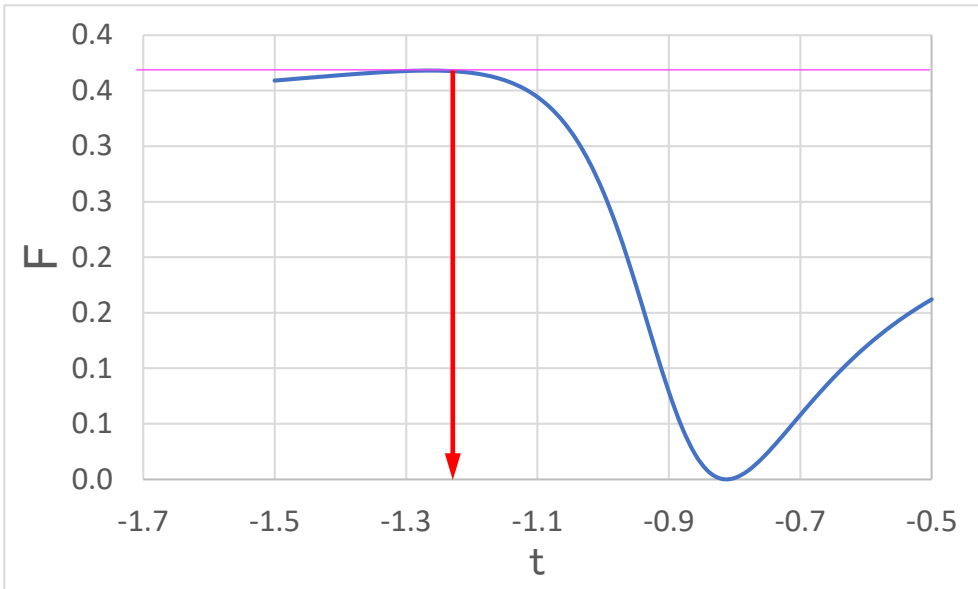
分割する平面がグループA,Bの中間地点 $(x_{AB}, y_{AB}) = \frac{1}{2}\{(\bar{x}_A, \bar{y}_A) + (\bar{x}_B, \bar{y}_B)\} = (159.7, 52.1)$ を通るとすると

$$-1.264 \times 159.7 + 52.1 + \frac{c}{b} = 0 \text{ より } \frac{c}{b} = 149.8$$

$$\omega = -1.264x + y + 149.8 \quad \omega = 0 \text{ のとき } y = 1.264x - 149.8$$



$$\omega = -1.264x + y + 149.8$$



最大値になるのは、 $t = \frac{a}{b} = -1.264266$

性別	身長	体重	w	性別	身長	体重	w
男	143.8	38.7	6.7368	女	145.6	38.5	4.2616
男	153.5	45.2	0.976	女	152.6	45.9	2.8136
男	159.1	50.6	-0.7024	女	153.6	45.9	1.5496
男	164.5	53.7	-4.428	女	156.5	47.8	-0.216
男	168.7	57.6	-5.8368	女	156.9	51.8	3.2784
男	170.5	61.4	-4.312	女	158.4	51.0	0.5824
男	170.3	62.1	-3.3592	女	157.2	49.9	0.9992
男	168.8	62.2	-1.3632	女	158.0	51.9	1.988
男	171.0	62.7	-3.644	女	157.3	51.0	1.9728
男	170.8	62.7	-3.3912	女	156.4	51.3	3.4104

$\omega < 0$ 男
 $\omega > 0$ 女

$$\text{判別的中率} = \frac{\text{正しく判別したサンプル数}}{\text{サンプル数}} = \frac{8+9}{20} = 0.85$$