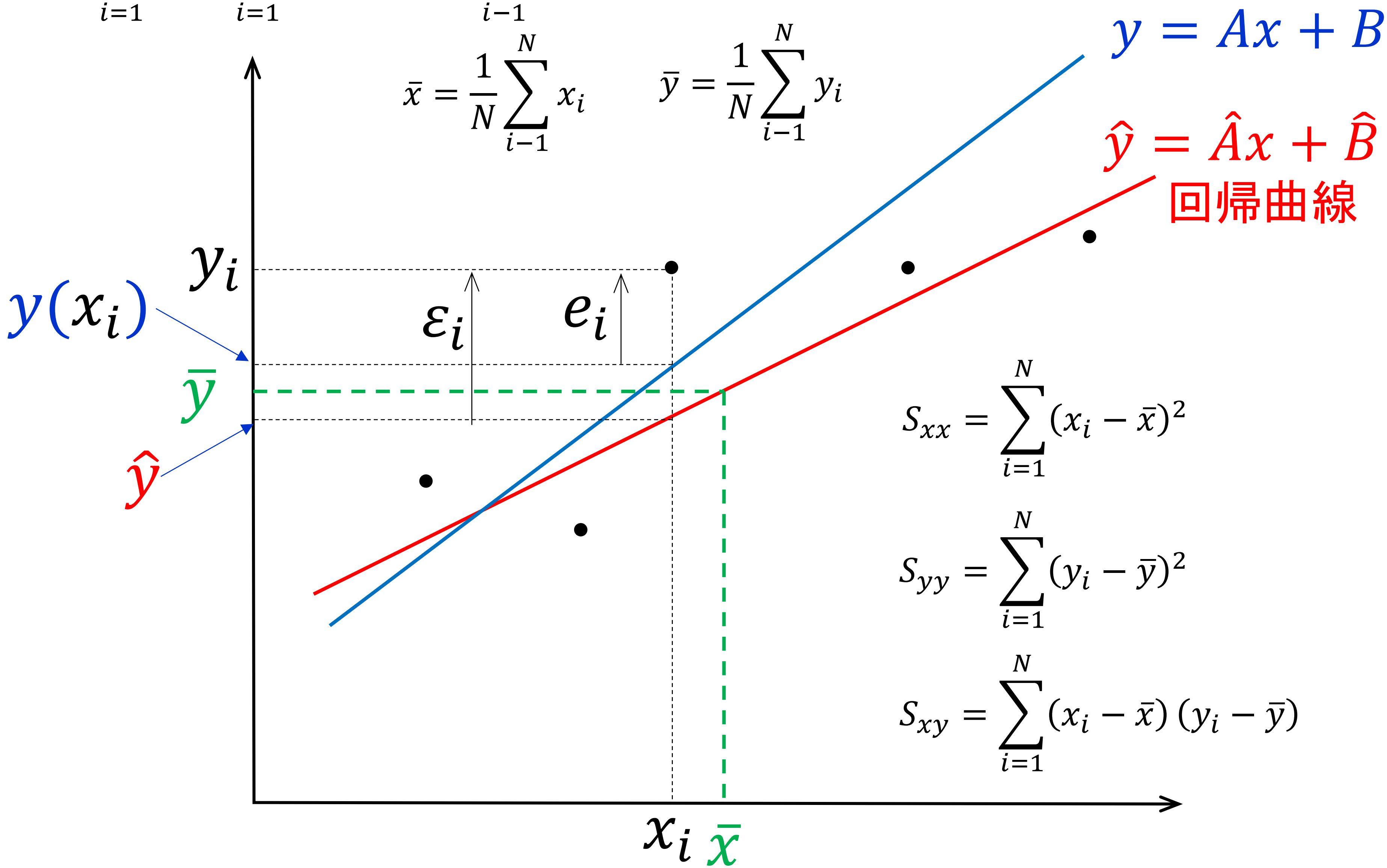
回帰曲線の信頼･予測区間を算出

$$S_{\varepsilon\epsilon} = \sum_{i=1}^{N} \varepsilon_i^2 = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = \sum_{i-1}^{N} \left(y_i - \hat{B} - \hat{A}x_i\right)^2$$

$$\bar{x} = \frac{1}{N}\sum_{i-1}^{N} x_i \qquad \bar{y} = \frac{1}{N}\sum_{i-1}^{N} y_i$$



$$S_{xx} = \sum_{i=1}^{N} (x_i - \bar{x})^2$$

$$S_{yy} = \sum_{i=1}^{N} (y_i - \bar{y})^2$$

$$S_{xy} = \sum_{i=1}^{N} (x_i - \bar{x})\,(y_i - \bar{y})$$

$$\overline{x^2} = \frac{1}{N}\sum_{i-1}^{N} x_i^2$$

$$S_{xx} = \sum_{i=1}^{N}(x_i - \bar{x})^2 = \sum_{i-1}^{N} x_i^2 - 2\sum_{i=1}^{N} x_i\bar{x} + \sum_{i=1}^{N} \bar{x}^2 = N\,\overline{x^2} - 2N\bar{x}^2 + N\bar{x}^2$$
$$= N(\overline{x^2} - \bar{x}^2)$$

$$\overline{y^2} = \frac{1}{N}\sum_{i-1}^{N} y_i^2$$

$$S_{yy} = \sum_{i=1}^{N}(y_i - \bar{y})^2 = \sum_{i-1}^{N} y_i^2 - 2\sum_{i=1}^{N} y_i\bar{y} + \sum_{i=1}^{N} \bar{y}^2 = N\,\overline{y^2} - 2N\bar{y}^2 + N\bar{y}^2$$
$$= N(\overline{y^2} - \bar{y}^2)$$

$$S_{xy} = \sum_{i=1}^{N}(x_i - \bar{x})\,(y_i - \bar{y}) = \sum_{i=1}^{N}(x_iy_i - \bar{x}y_i - x_i\bar{y} + \bar{x}\bar{y})$$
$$= N\overline{xy} - 2N\bar{x}\bar{y} + N\bar{x}\bar{y}$$
$$= N(\overline{xy} - \bar{x}\bar{y})$$

回帰曲線$\hat{y}=\hat{A}x+\hat{B}$　の係数$\hat{A}$及び$\hat{B}$を最小二乗法で算出する

$$S_{\varepsilon\epsilon}=\sum_{i=1}^{N}\varepsilon_i^2=\sum_{i=1}^{N}(y_i-\hat{y}_i)^2=\sum_{i-1}^{N}\left(y_i-\hat{B}-\hat{A}x_i\right)^2$$

$$\frac{\partial s_{\varepsilon\varepsilon}}{\partial\hat{B}}=-2\sum_{i=1}^{\mathrm{N}}\left(y_i-\hat{B}-\hat{A}x_i\right)=0 \longrightarrow \hat{\boldsymbol{B}}+\hat{\boldsymbol{A}}\overline{\boldsymbol{x}}=\overline{\boldsymbol{y}}$$

$$\frac{\partial s_{\varepsilon\varepsilon}}{\partial\hat{A}}=-2\sum_{i=1}^{\mathrm{N}}\left(y_i-\hat{B}-\hat{A}x_i\right)x_i=0$$

$$\sum_{i-1}^{N}\left(y_i-\hat{B}-\hat{A}x_i\right)x_i=\sum_{i=1}^{N}\left(x_iy_i-\hat{B}x_i-\hat{A}{x_i}^2\right)=N\overline{xy}-N\hat{B}\bar{x}-N\hat{A}\overline{x_2}$$

$$=0$$

$$S_{xx}=N(\overline{x^2}-\bar{x}^2)\quad S_{yy}=N(\overline{y^2}-\bar{y}^2)\quad S_{xy}=N(\overline{xy}-\bar{x}\bar{y})$$

$$\boldsymbol{N\hat{B}\overline{x}}+\left(\boldsymbol{S_{xx}}+\boldsymbol{N\overline{x^2}}\right)\boldsymbol{\hat{A}}=\boldsymbol{S_{xy}}+\boldsymbol{N\overline{xy}}$$

$$\hat{\boldsymbol{B}}=\overline{\boldsymbol{y}}-\frac{\boldsymbol{S_{xy}}}{\boldsymbol{S_{xx}}}\overline{\boldsymbol{x}}\qquad\hat{\boldsymbol{A}}=\frac{\boldsymbol{S_{xy}}}{\boldsymbol{S_{xx}}}\qquad\hat{\boldsymbol{A}}=\frac{\boldsymbol{S_{xy}}}{\sqrt{\boldsymbol{S_{xx}S_{yy}}}}\sqrt{\frac{\boldsymbol{S_{yy}}}{\boldsymbol{S_{xx}}}}=\boldsymbol{r}\sqrt{\frac{\boldsymbol{S_{yy}}}{\boldsymbol{S_{xx}}}}$$

$$\hat{\boldsymbol{A}} = \frac{\boldsymbol{S_{xy}}}{\boldsymbol{S_{xx}}} = \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y}) \frac{1}{S_{xx}} = \sum_{i=1}^{N} \frac{(x_i - \bar{\boldsymbol{x}})}{S_{xx}} (y_i - \bar{\boldsymbol{y}})$$

$y_i - \bar{\boldsymbol{y}}$の分散が$\boldsymbol{\sigma}^2$とする

$$V_{\hat{A}} = \sum_{i=1}^{N} \left(\frac{x_i - \bar{\boldsymbol{x}}}{S_{xx}}\right)^2 \sigma^2 = \frac{\sigma^2}{S_{xx}}$$

$\hat{\boldsymbol{B}} = \bar{\boldsymbol{y}} - \frac{\boldsymbol{S_{xy}}}{\boldsymbol{S_{xx}}} \bar{\boldsymbol{x}}$　の分散は

$$V_{\hat{B}} = V_{\bar{y}} + \bar{x}^2 V_{\bar{A}} = \left(\frac{\sigma}{\sqrt{N}}\right)^2 + \bar{x}^2 \frac{\sigma^2}{S_{xx}} = \left(\frac{1}{N} + \frac{\bar{x}^2}{S_{xx}}\right)\sigma^2$$

↑標本平均の分散

$\hat{y} = \hat{A}(x - \bar{x}) + \bar{y}$

$$V_{\hat{y}} = V_{\hat{A}}(x - \bar{x})^2 + \left(\frac{1}{N}\right)\sigma^2 = \frac{\sigma^2}{S_{xx}}(x - \bar{x})^2 + \left(\frac{1}{N}\right)\sigma^2 = \left(\frac{1}{N} + \frac{(x - \bar{x})^2}{S_{xx}}\right)\sigma^2$$

σ^2を自由度N-2の不偏分散$\hat{V}_E = S_{\varepsilon\epsilon}/(N-2)$に置き換えて

$$V_{\hat{y}} = V_{\hat{A}}(x - \bar{x})^2 + \left(\frac{1}{N}\right)\sigma^2 = \frac{\sigma^2}{S_{xx}}(x - \bar{x})^2 + \left(\frac{1}{N}\right)\sigma^2 = \left(\frac{1}{N} + \frac{(x - \bar{x})^2}{S_{xx}}\right)\hat{V}_e$$

回帰曲線$\hat{\boldsymbol{y}} = \hat{\boldsymbol{A}}\boldsymbol{x} + \hat{\boldsymbol{B}}$の信頼度$1-\alpha$の信頼区間は、

$$\hat{A}x + \hat{B} \pm t(N-2, \alpha)\sqrt{\left(\frac{1}{N} + \frac{(x - \bar{x})^2}{S_{xx}}\right)\hat{V}_e}$$

回帰曲線$\hat{\boldsymbol{y}} = \hat{\boldsymbol{A}}\boldsymbol{x} + \hat{\boldsymbol{B}}$の信頼度$1-\alpha$の信頼区間は、

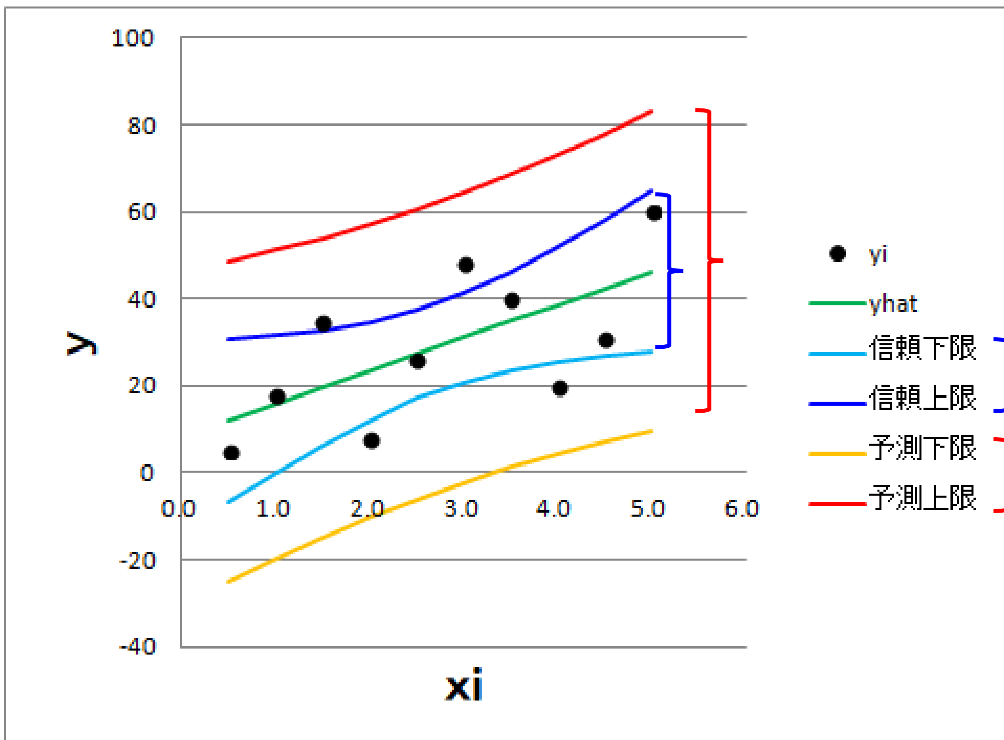
信頼区間

$$\hat{A}x + \hat{B} \pm t(N-2, \alpha)\sqrt{\left(\frac{1}{N} + \frac{(x-\bar{x})^2}{S_{xx}}\right)\hat{V}_e}$$

回帰曲線$\hat{\boldsymbol{y}} = \hat{\boldsymbol{A}}\boldsymbol{x} + \hat{\boldsymbol{B}}$の信頼度$1-\alpha$の予測区間は、

予測区間

$$\hat{A}x + \hat{B} \pm t(N-2, \alpha)\sqrt{\left(1 + \frac{1}{N} + \frac{(x-\bar{x})^2}{S_{xx}}\right)\hat{V}_e}$$



回帰曲線$\hat{\boldsymbol{y}} = \hat{\boldsymbol{A}}\boldsymbol{x} + \hat{\boldsymbol{B}}$は信頼度$1-\alpha$でこの区間に入る

観測値y_iは信頼度$1-\alpha$でこの区間に入る

Excelで回帰曲線の信頼・予測区間を算出

$$\boldsymbol{y} = \hat{A}x_i + \hat{B} \pm t(n-2,\alpha)\sqrt{\left\{\frac{1}{N} + \frac{(x_i-\bar{x})^2}{S_{xx}}\right\}\widehat{V_e}}$$

$$\boldsymbol{y} = \hat{A}x_i + \hat{B} \pm t(n-2,\alpha)\sqrt{\left\{1 + \frac{1}{N} + \frac{(x_i-\bar{x})^2}{S_{xx}}\right\}\widehat{V_e}}$$

$x_i - \bar{x}$ $y_i - \bar{y}$

	xi	yi	xi−xbar	yi−xbar	yhati	ei	信頼下限	信頼上限	予測下限	予測上限
1	0.5	5	−2.3	−24.1	11.9	−6.9	−6.69	30.58	−24.83	48.72
2	1.0	18	−1.8	−11.1	15.8	2.2	−0.05	31.56	−19.67	51.18
3	1.5	35	−1.3	5.9	19.6	15.4	6.28	32.86	−14.81	53.95
4	2.0	8	−0.8	−21.1	23.4	−15.4	12.07	34.69	−10.28	57.04
5	2.5	26	−0.3	−3.1	27.2	−1.2	17.02	37.37	−6.10	60.49
6	3.0	48	0.3	18.9	31.0	17.0	20.83	41.18	−2.29	64.30
7	3.5	40	0.8	10.9	34.8	5.2	23.51	46.13	1.16	68.48
8	4.0	20	1.3	−9.1	38.6	−18.6	25.34	51.92	4.25	73.01
9	4.5	31	1.8	1.9	42.4	−11.4	26.64	58.25	7.02	77.87
10	5.0	60	2.3	30.9	46.3	13.7	27.62	64.89	9.48	83.03
平均値	2.75	29.1								
σ	1.44	16.46								

=SUMPRODUCT(D2:D11,E2:E11)

$$S_e = \sum_{i=1}^{N} e_i$$

Ahat	7.62	Sxy	157.25
Bhat	8.13	Sxx	20.63

Se	1511.99	t(n−2,α)	2.31
n	10		
Ve	189.00		

=SUMSQ(D2:D11)

=TINV(0.05,8)

$$\hat{A} = \frac{S_{xy}}{S_{xx}}$$

$$\hat{B} = \bar{y} - \frac{S_{xy}}{S_{xx}}\bar{x}$$

$$V_e = \frac{S_e}{N-2}$$

α5%

N-2＝10－2