2008年度 松山大学経営学部 開講科目

経営工学概論

檀 裕也

(dan@cc.matsuyama-u.ac.jp)

http://www.cc.matsuyama-u.ac.jp/~dan/

前回の課題

- 最小自乗法について説明せよ。

（解答例）
最小自乗法とは、回帰分析において、誤差 ε の自乗和が最小になるように回帰式
　y＝ax＋b
の係数（aとb）を求めることである。

今回の内容

- 回帰分析による予測
  - 最小二乗法（一般の線形回帰の場合）
  - 寄与率 R^2

定理

- 表のデータについて回帰分析すると、回帰直線の方程式は

$$y = \alpha x + \beta$$

である。ただし、

$$\alpha = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

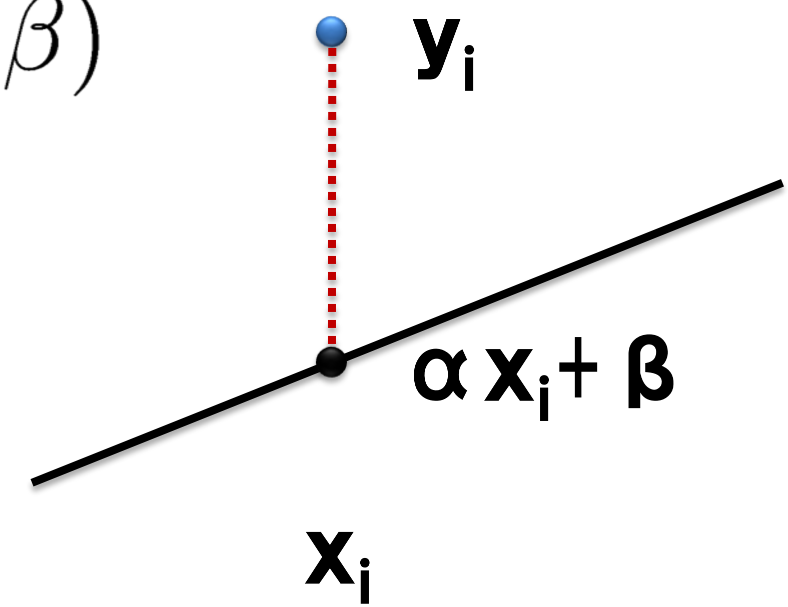
$$\beta = \bar{y} - \alpha \bar{x}$$

x	y
x_1	y_1
x_2	y_2
:	:
:	:
x_n	y_n
$\bar{x}$	$\bar{y}$

定理の証明 (1/4)

- 回帰直線の方程式を$y = \alpha x + \beta$と仮定すると、データ（x_i、y_i）の回帰式からの誤差 ε_i は

$$\varepsilon_i = y_i - (\alpha x_i + \beta)$$



と表せる。そこで、

$$\sum_{i=1}^{n} \varepsilon_i^2$$

の最小化を目指せばよい。

定理の証明 (2/4)

- 誤差の自乗和をβについて整理すると、

$$
\begin{aligned}
\sum_{i=1}^{n} \varepsilon_i^2 &= \sum_{i=1}^{n} \{y_i - (\alpha x_i + \beta)\}^2 \\
&= \sum_{i=1}^{n} \{(y_i - \alpha x_i) - \beta)\}^2 \\
&= n\beta^2 - 2\sum_{i=1}^{n}(y_i - \alpha x_i)\cdot\beta + \sum_{i=1}^{n}(y_i - \alpha x_i)^2
\end{aligned}
$$

これはβに関する2次関数となる。

定理の証明 (3/4)

- ２次関数の最小値に関する補題を適用すると、

$$\beta = \frac{1}{n}\sum_{i=1}^{n}(y_i - \alpha x_i)$$ のとき、最小値

$$-\frac{1}{n}\left\{\sum_{i=1}^{n}(y_i - \alpha x_i)\right\}^2 + \sum_{i=1}^{n}(y_i - \alpha x_i)^2$$

を取ることが分かる。

定理の証明 (4/4)

- さらに、この値は

$$\sum_{i=1}^{n}(x_i-\bar{x})^2\alpha^2-2\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\alpha+\sum_{i=1}^{n}(y_i-\bar{y})^2$$

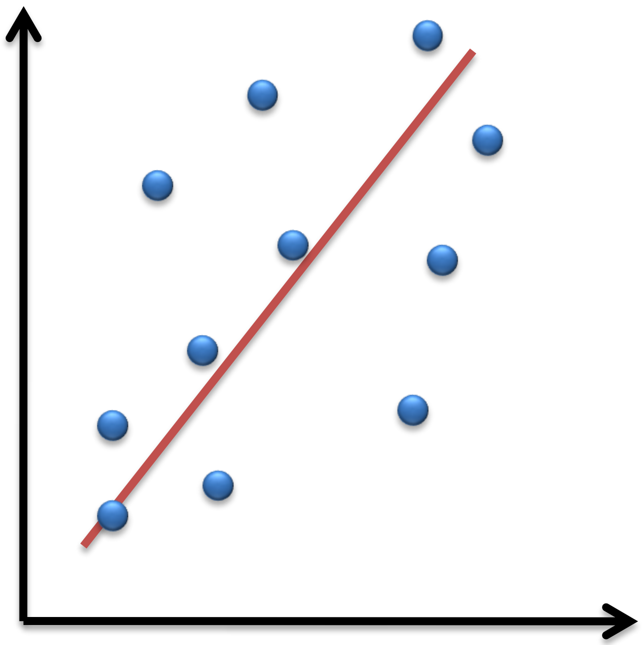
と変形できる。再び補題を適用すると、上式は

$$\alpha=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

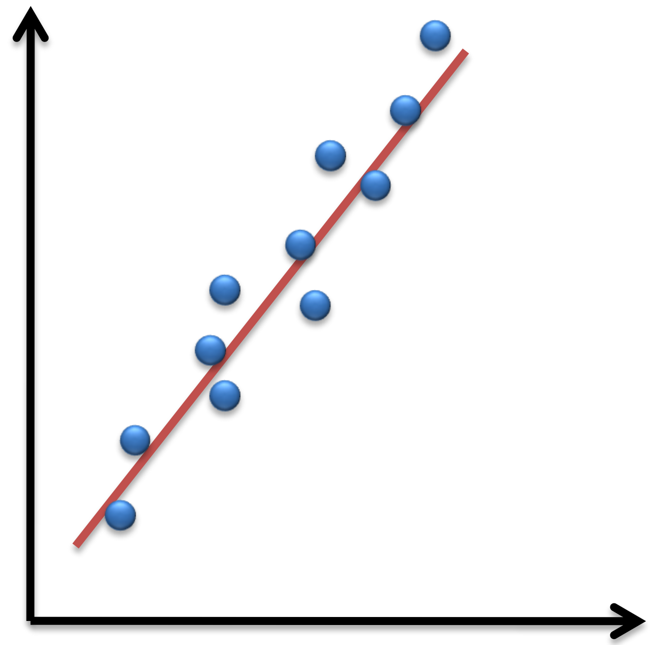
のとき最小値を取る。

- β は $\beta=\bar{y}-\alpha\bar{x}$ によって与えればよい。（終）

相関の強さ



弱い相関



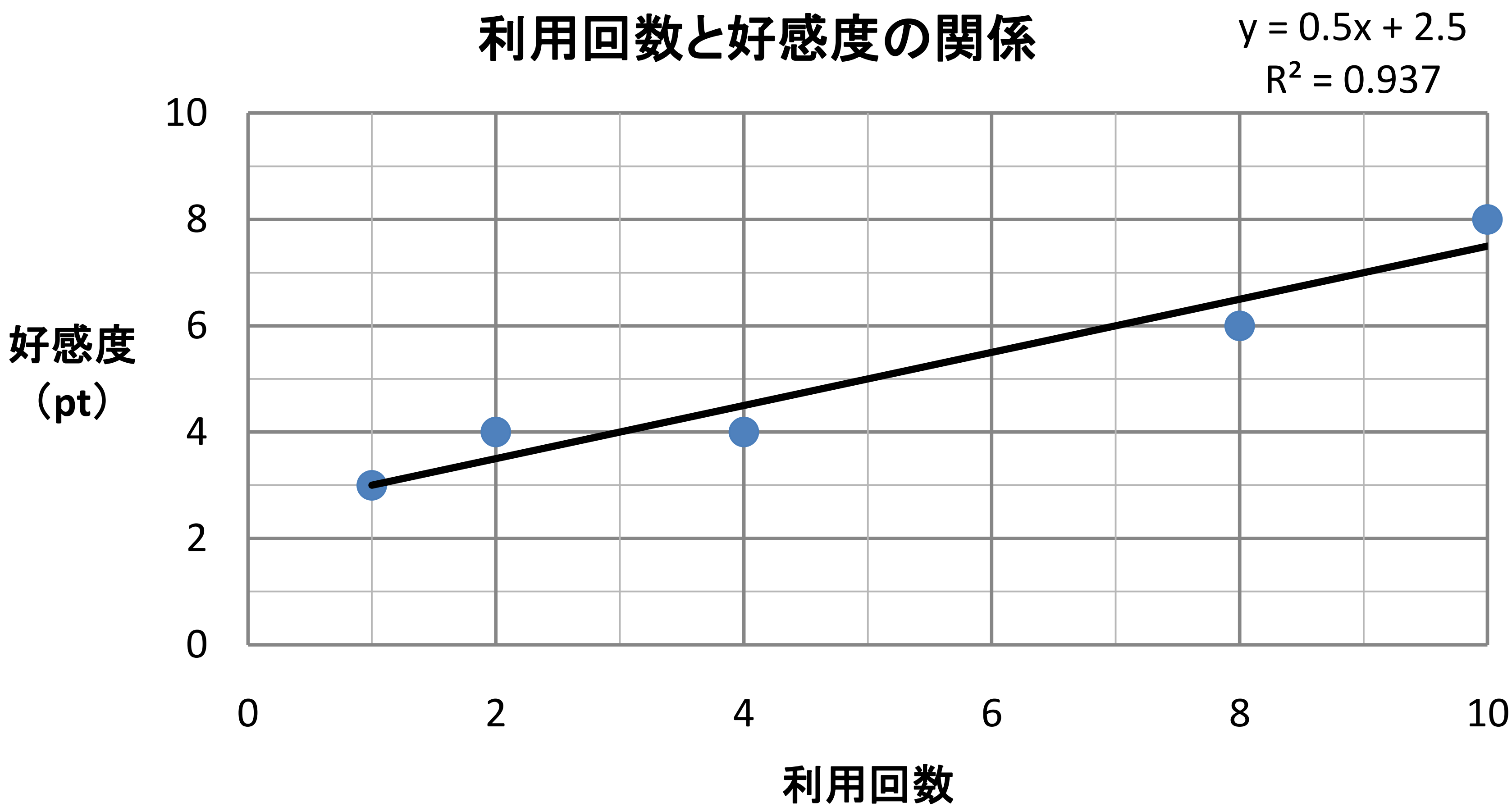
強い相関

練習（再掲）

- ある店舗における利用回数と好感度の関係について、下の表に示したデータがある。このとき、回帰分析によって両者の関係を調べよ。

利用回数	好感度（pt）
1	3
2	4
4	4
8	6
10	8

散布図に回帰直線を重ねる



回帰直線の当てはまりの良さ

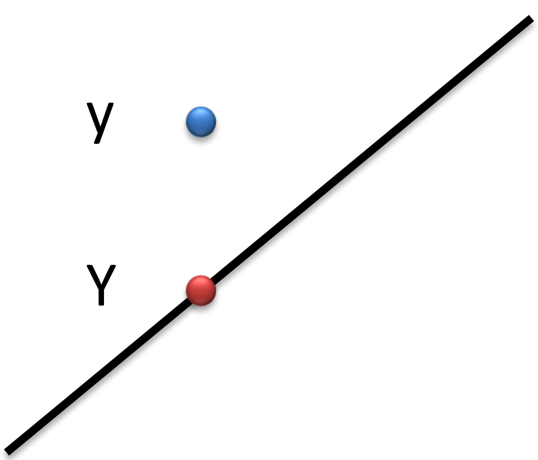
- 回帰直線$y = \frac{1}{2}x + \frac{5}{2}$によって予測される値

x	y	Y	$(y_i - \bar{y})^2$	$(Y_i - \bar{y})^2$	$(y_i - Y_i)^2$
1	3	3	4	4	0
2	4	3.5	1	2.25	0.25
4	4	4.5	1	0.25	0.25
8	6	6.5	1	2.25	0.25
10	8	7.5	9	6.25	0.25
5	5		16	15	1

寄与率

- **寄与率または決定係数 R^2**

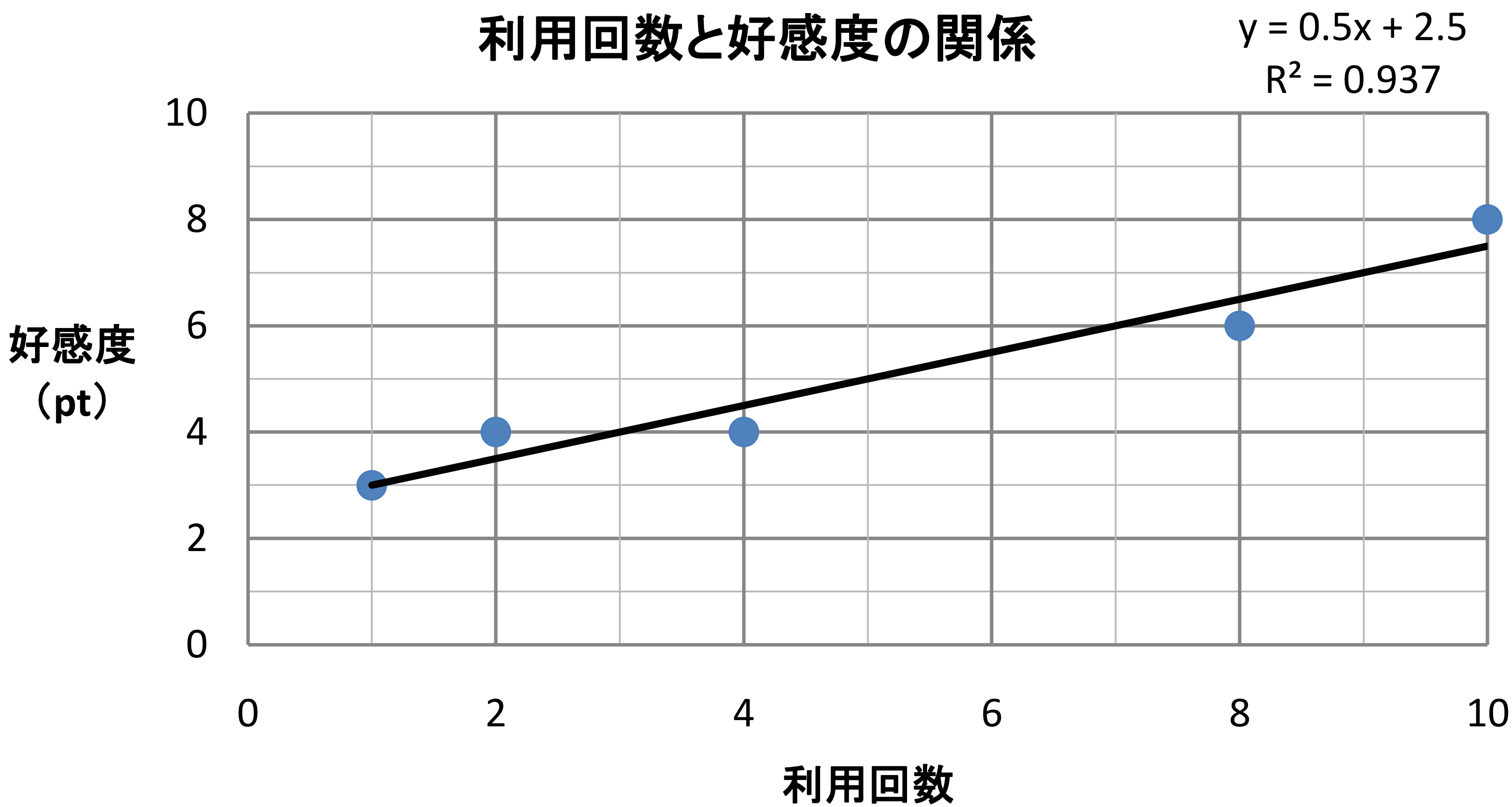
$$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(Y_i-\bar{y})^2+\sum_{i=1}^{n}(y_i-Y_i)^2$$



$$1=\boxed{\frac{\sum_{i=1}^{n}(Y_i-\bar{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}}+\frac{\sum_{i=1}^{n}(y_i-Y_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$$

$$\frac{15}{16}=0.9375$$

散布図に回帰直線を重ねる



寄与率の性質

- 回帰分析の当てはまりの良さを表す0～1の値

$$R^2 = \frac{\displaystyle\sum_{i=1}^{n}(Y_i - \bar{y})^2}{\displaystyle\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

  - 1に近いほど当てはまりが良い
  - 0に近いほど当てはまりが悪い
  - 寄与率の平方根は「相関係数」に等しい

まとめ

- 回帰分析による予測
  - 最小二乗法（一般の線形回帰の場合）
  - 寄与率 R^2

松山大学 Since1923

提出課題

- 寄与率（決定係数）について説明せよ。

次回の予定

- 第8回　線形計画法
  - 日時：2008年 6月13日(金) 4時限目
  - 場所：845教室