

【データの相互関係】

- 実験やシミュレーション結果、実際の観察データ、種々の統計データなどからデータ間の関係を推定・整理する
- 原因-結果の相互関係が明確になり、モデルの作成や将来の予測が可能となる

【回帰分析】

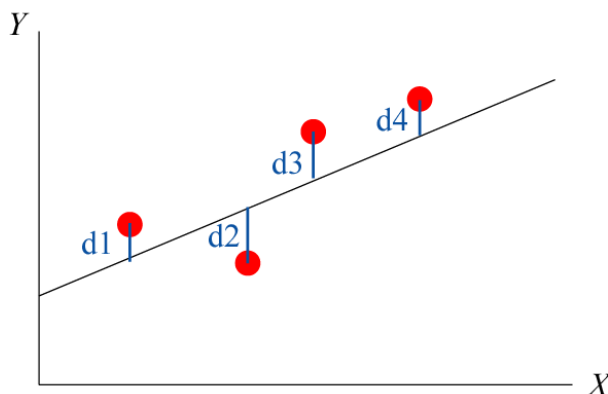
- 回帰分析(regression analysis)とは、**従属変数(目的変数)**と連続尺度の**独立変数(説明変数)**の間に式を当てはめ、従属変数が説明変数によってどれくらい説明できるのかを定量的に分析することである
(Wikipedia より引用)

【直線回帰(線形回帰)】

- 2つの変数間の関係を一次方程式で表すような回帰のことを直線回帰(線形回帰)という
- つまり、2つの変数 X, Y の関係を一次方程式 $Y = aX + b$ で表すこと

【最小二乗法】

- 回帰直線を求めるもっとも一般的な方法
- 直線とデータとの誤差の二乗和を最小にすることで、回帰直線を求める



直線とデータとの誤差 (d_1, d_2, d_3, d_4)
の二乗和 $d_1^2 + d_2^2 + d_3^2 + d_4^2$ を
最小にする

【具体的な計算方法】

- (1) 2 変数 X, Y の関係を $Y = aX + b$ (直線)と仮定する
- (2) データと直線との誤差を求め、二乗する
- (3) すべてのデータの組について誤差の二乗和 S を求める
- (4) S を最小にする a, b を求める(a, b で偏微分して 0 とおき、連立方程式を解く)

上記のデータが 4 組の場合については

$$\begin{cases} a = \frac{4D - CE}{4B - E^2} \\ b = \frac{BC - DE}{4B - E^2} \end{cases} \quad \text{ただし}$$

$$\begin{aligned} B &= (x_1^2 + x_2^2 + x_3^2 + x_4^2) \\ C &= (y_1 + y_2 + y_3 + y_4) \\ D &= (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) \\ E &= (x_1 + x_2 + x_3 + x_4) \end{aligned}$$

【データが4組の場合について a, b を導出してみよう】

【一般的な場合】

一般に、n組のデータ $(x_1, y_1), (x_2, y_2), (x_3, y_3) \wedge (x_n, y_n)$ がある場合の線形回帰式 $Y = aX + b$ の係数は

$$\left\{ \begin{array}{l} a = \frac{nD - CE}{nB - E^2} \\ b = \frac{BC - DE}{nB - E^2} \end{array} \right. \quad \text{ただし} \quad \begin{array}{l} B = \sum_{i=1}^n x_i^2, \quad C = \sum_{i=1}^n y_i \\ D = \sum_{i=1}^n x_i y_i, \quad E = \sum_{i=1}^n x_i \end{array}$$

と書くことができる。

【参考】

回帰直線の傾き a と切片 b は以下のように計算することもできる。

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

ただし \bar{x}, \bar{y} はそれぞれ x, y の平均値である。

【例: 需要関数の推定】

- あるパン屋で、同じパンを 100 円、150 円、200 円で販売したときの 1 日の売上はそれぞれ 12 個、10 個、2 個だった。
- 需要関数(価格と売上の関係)が一次方程式 $Y = aX + b$ で表せると仮定して線形回帰をおこなってみよう。
- たとえば 50 円でこのパンを売ったとすると、何個売れると予想されるだろうか？

回帰式(需要関数の推定値)

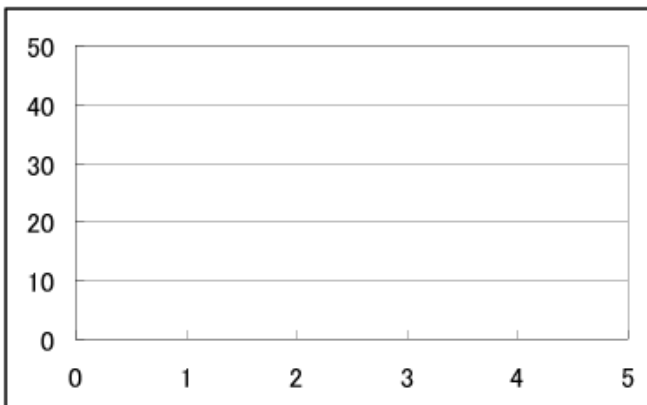
50 円で売った場合の予想売り上げ個数

個

【第12回のレポート】

- 4つの駅で、一日の乗降人数と売店の売上を調べたところ、以下の表のようになった。
乗降人数と売上の関係を最小二乗法を用いて線形回帰せよ。
- 新たに乗降人数が5千人の駅に売店を出すことにした。いくらの上が見込まれるか推定せよ。
※出席カードには回帰式と売上見込額を記入すること。

駅	乗降人数(千人)	売上(万円)
A駅	1	
B駅	2	
C駅	3	
D駅	4	



回帰式 :

新しい売店の売上見込 :

_____ 万円