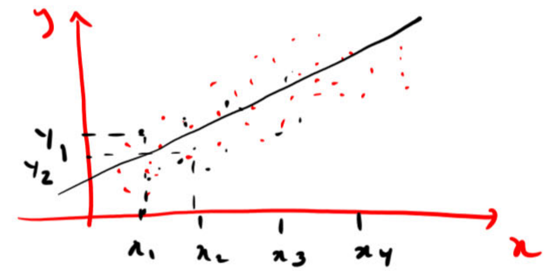
Math 4501 - Probability and Statistics II

6.5 - Regression ← we will employ MLE technique to determine the regression parameters

Simplest regression problem

$Y_i - \alpha_1 + \beta x_i = \varepsilon_i$



Given the data points

$$\rightarrow \quad (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

estimate the parameters α and β of the linear model

$$\boxed{\underline{E[Y|x]} = \underbrace{\alpha_1 + \beta x}_{\mu(x)} \,,}$$

$$\underbrace{Y - E[Y|x]}_{\alpha_1 + \beta x} = \varepsilon \sim N(0, \sigma^2)$$

unknown

that is, fit a straight line to the given set of data.

$$E[Y|x] = \mu(x) = \alpha_1 + \beta x$$

Assumptions:

- for each particular value of x the value of Y differs from its mean by a random amount ε.
- the distribution of ε is $N\left(0, \sigma^2\right)$.

$$\underbrace{Y = \alpha_1 + \beta x + \varepsilon}$$

σ^2 is another parameter to estimate

Consequence: For the linear model described above, we have

$$\rightsquigarrow \quad \boxed{Y_i = \overbrace{\alpha_1 + \beta x_i}^{\mu(x_i)} + \varepsilon_i \,,}$$

$\varepsilon_i \sim N(0, \sigma^2)$

where ε_i, $i = 1, 2, \ldots, n$, are independent $N\left(0, \sigma^2\right)$ random variables.

$$\Rightarrow \quad Y_i \sim N(\alpha_1 + \beta x_i, \sigma^2)$$

GOAL : Estimate

$\alpha_1, \beta, \sigma^2$

$\mu(x) = \alpha_1 + \beta x$

$y = mx + b$

$y = y_0 + m(x - x_0)$

- For convenience, we set

$$\boxed{\alpha_1 = \alpha - \beta\bar{x}\ ,}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean of the observations $x_1, \ldots, x_n$.

- For each $i = 1, 2, \ldots, n$, we have that

$$Y_i = \alpha_1 + \beta x + \varepsilon_i \xrightarrow{\alpha_1 = \alpha - \beta\bar{x}} \boxed{Y_i = \alpha + \beta(x_i - \bar{x}) + \varepsilon_i}$$

is equal to a nonrandom quantity $\alpha + \beta(x_i - \bar{x})$ plus a mean-zero normal random variable ε_i.

- The random variables $Y_1, Y_2, \ldots, Y_n$ are mutually independent normal variables with respective means

$$\alpha + \beta(x_i - \bar{x})\ , \quad i = 1, 2, \ldots, n$$

and unknown variance σ^2.

estimate α, β, σ^2

$Y_i \sim N(\alpha + \beta(x_i - \bar{x}), \sigma^2)$

Proposition

Under the conditions described above, the maximum likelihood estimators of α, β *and* σ^2 *are given by:*

$$\boxed{\widehat{\alpha} = \bar{Y}}$$

$$\boxed{\widehat{\beta} = \frac{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)\left(x_i - \bar{x}\right)}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}}$$

$$\widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} \left[Y_i - \widehat{\alpha} - \widehat{\beta}\left(x_i - \bar{x}\right) \right]^2$$

Interpretation for :

$$\hat{\beta} = \frac{\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} = \text{ratio of sample covariance to "sample" variance}$$

$$\overbrace{\hat{Y}_i}^{\text{estimates for } Y_i} = \hat{\alpha} - \hat{\beta}(x_i - \bar{x})$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left[Y_i - \hat{Y}_i \right]^2$$

average of the squares of the deviations between actual values and estimated values!

Proof (yet another example of MLE):

θ stands for σ^2

Recall that $Y_i \sim N(\alpha + \beta(x_i - \bar{x}), \theta)$ independent!

pdf of $X \sim N(\mu, \sigma^2)$ is $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Define the likelihood function:

$$L(\alpha, \beta, \underset{\sigma^2}{\theta}) = \prod_{i=1}^{n} f(y_i; \alpha, \beta, \sigma^2) =$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta}} \exp\left(- \frac{(y_i - \overbrace{(\alpha + \beta(x_i - \bar{x}))}^{\mu(x)})^2}{2\theta} \right)$$

θ is σ^2 $\qquad$ θ is σ^2

$$= (2\pi\theta)^{-n/2} \exp\left(-\frac{1}{2\theta} \sum_{i=1}^{n} (y_i - (\alpha + \beta(x_i - \bar{x})))^2 \right)$$

$e^{a_1} e^{a_2} \cdots e^{a_n} = e^{a_1 + a_2 + \cdots + a_n}$

Apply natural log:

$$\ln\left(L(\alpha,\beta,\overset{\theta}{\sigma^2})\right) = \overbrace{-\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\theta - \frac{1}{2\theta}\underbrace{\sum_{i=1}^{n}\left(y_i-\alpha-\beta(x_i-\bar{x})\right)^2}_{H(\alpha,\beta)}}^{\text{maximize}}$$

$\theta = \sigma^2 > 0$

To maximize $\ln\left(L(\alpha,\beta,\overset{\theta}{\sigma^2})\right)$, we need to make $H(\alpha,\beta)$ as small as possible!

We start by minimizing $H(\alpha,\beta) = \sum_{i=1}^{n}\left(y_i-\alpha-\beta(x_i-\bar{x})\right)^2$

First order conditions are:

$$\begin{cases} \dfrac{\partial H}{\partial \alpha} = 0 \\[2ex] \dfrac{\partial H}{\partial \beta} = 0 \end{cases} \iff \begin{cases} -2\sum_{i=1}^{n}\left(y_i-\alpha-\beta(x_i-\bar{x})\right) = 0 \\[2ex] -2\sum_{i=1}^{n}(x_i-\bar{x})\left(y_i-\alpha-\beta(x_i-\bar{x})\right) = 0 \end{cases}$$

Divide both sides of each equation by -2:

$$\begin{cases} \sum_{i=1}^{m}\Big(y_i - \alpha - \beta(x_i-\bar{x})\Big) = 0 \\ \sum_{i=1}^{m}(x_i-\bar{x})\Big(y_i - \alpha - \beta(x_i-\bar{x})\Big) = 0 \end{cases}$$

1^{st} equation
$$\sum_{i=1}^{m} y_i - \sum_{i=1}^{m}\alpha - \sum_{i=1}^{m}\beta(x_i-\bar{x}) = 0$$

$$\Rightarrow \sum_{i=1}^{m} y_i - m\alpha - \beta\underbrace{\sum_{i=1}^{m}(x_i-\bar{x})}_{=0} = 0$$

$$\sum_{i=1}^{m} x_i - \sum_{i=1}^{m}\bar{x} = m\bar{x} - m\bar{x} = 0$$

$$\Rightarrow \sum_{i=1}^{m} y_i - m\alpha = 0 \Rightarrow \boxed{\alpha = \frac{1}{m}\sum_{i=1}^{m} y_i = \overline{Y}}$$

$\leftarrow$ Same as in the statement!

2^{nd} equation: (replacing α by $\bar{y}$)

$$\sum_{i=1}^{n} (x_i - \bar{x}) \left((y_i - \bar{y}) - \beta (x_i - \bar{x}) \right) = 0$$

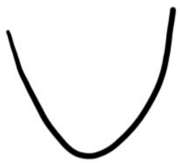
$$\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) - \beta (x_i - \bar{x})^2 = 0$$

$$\sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y}) - \beta \sum_{i=1}^{m} (x_i - \bar{x})^2 = 0$$

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

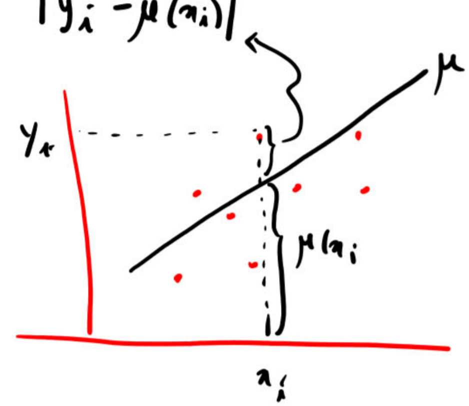
As given in the statement

Relation with method of least squares



- The parameters α and β minimize the quantity

$$H(\alpha, \beta) = \sum_{i=1}^{n} \left[y_i \overbrace{- \alpha - \beta \left(x_i - \bar{x} \right)}^{\mu(x_i)} \right]^2 .$$



- Since

$$\left| y_i - \alpha - \beta \left(x_i - \bar{x} \right) \right| = \left| y_i - \mu \left(x_i \right) \right|$$

is the vertical distance from the point (x_i, y_i) to the line $y = \mu(x)$, then $H(\alpha, \beta)$ represents the sum of the squares of those distances.

- Selecting α and β so that the sum of the squares is minimized means that we are fitting the straight line to the data by the *method of least squares*.

- Thus, the maximum likelihood estimates of α and β are also called *least squares estimates*.

Let us finish our computation:

We already proved that $\alpha = \bar{y}$ and $\beta = \dfrac{\sum_{i=1}^{m}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{m}(x_i - \bar{x})^2}$

are critical pts of $\boxed{H(\alpha, \beta) = \sum_{i=1}^{m}\left[y_i - \alpha - \beta(x_i - \bar{x})\right]^2}$

We still need to check that the critical pt obtained actually minimizes H.

We will use the 2^{nd} derivative test from Multivariable Calculus.

$$\frac{\partial^2 H}{\partial \alpha^2} = \frac{\partial}{\partial \alpha}\left[\frac{\partial H}{\partial \alpha}\right] = \frac{\partial}{\partial \alpha}\left[-2\sum_{i=1}^{m}\left(y_i - \alpha - \beta(x_i - \bar{x})\right)\right] = -2\overbrace{\sum_{i=1}^{m} -1}^{-m} = 2m$$

$$\frac{\partial^2 H}{\partial \alpha \partial \beta} = \frac{\partial^2 H}{\partial \beta \partial \alpha} = \frac{\partial}{\partial \beta}\left[\frac{\partial H}{\partial \alpha}\right] = \frac{\partial}{\partial \beta}\left[-2\sum_{i=1}^{m}\left(y_i - \alpha - \beta(x_i - \bar{x})\right)\right] = -2\sum_{i=1}^{m} -(x_i - \bar{x})$$

$$= 2\sum_{i=1}^{m}(x_i - \bar{x}) = 0$$

because $\sum_{i=1}^{m} x_i - \bar{x} = \sum_{i=1}^{m} x_i - m\bar{x} = m\bar{x} - m\bar{x} = 0$

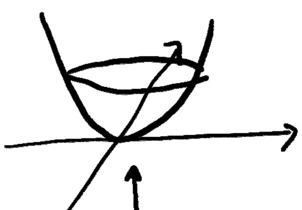
$$\frac{\partial^2 H}{\partial \beta^2} = \frac{\partial}{\partial \beta}\left[\frac{\partial H}{\partial \beta}\right] = \frac{\partial}{\partial \beta}\left[-2\sum_{i=1}^{n}(x_i-\bar{x})\left(y_i-\alpha-\beta(x_i-\bar{x})\right)\right]$$

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

$\underbrace{\text{The matrix of } 2^{nd} \text{ derivatives is}}_{\text{Hessian matrix of } H}$

$$D^2H = \begin{pmatrix} 2n & 0 \\ 0 & 2\sum_{i=1}^{n}(x_i-\bar{x})^2 \end{pmatrix} \qquad \sum x_i^2 - n(\bar{x})^2$$

Two alternative ways to show the critical pt is a minimum:



① D^2H has two positive eigenvalues $\Rightarrow$ critical pt is a minimum

② det of D^2H is positive and the component on the 1^{st} row, 1^{st} column is positive

close to critical pt graph of H has the approximate shape of a paraboloid openning up

CONCLUSION (so far): $\hat{\alpha} = \overline{Y}$

$$\hat{\beta} = \frac{\sum_{i=1}^{m} (Y_i - \overline{Y})(x_i - \overline{x})}{\sum_{i=1}^{m} (x_i - \overline{x})^2}$$

are the MLE of α, β.

One parameter left to estimate $\underline{\underline{\theta = \sigma^2}}$. Recall that

$$\overbrace{\sum_{i=1}^{m} \left(y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x})\right)^2}^{H(\hat{\alpha}, \hat{\beta})}$$

$$\ln\left(L\left(\hat{\alpha}, \hat{\beta}, \overset{\sigma^2}{\theta}\right)\right) = -\frac{m}{2}\ln 2\pi - \frac{m}{2}\ln\theta - \frac{1}{2\theta} H(\hat{\alpha}, \hat{\beta})$$

to maximize $\ln(L(\alpha, \beta, \theta))$ w.r.t θ, we take the 1st derivative w.r.t θ:

$$\frac{\partial}{\partial\theta} \ln\left(L(\hat{\alpha}, \hat{\beta}, \theta)\right) = -\frac{m}{2}\frac{1}{\theta} + \frac{1}{2\theta^2} H(\hat{\alpha}, \hat{\beta})$$

FOC: $\frac{\partial}{\partial\theta} \ln\left(L(\hat{\alpha}, \hat{\beta}, \theta)\right) = 0 \iff -\frac{m}{2}\frac{1}{\theta} + \frac{1}{2\theta^2} H(\hat{\alpha}, \hat{\beta}) = 0 \iff \boxed{\theta = \frac{1}{m} H(\hat{\alpha}, \hat{\beta})}$

The critical point (for the problem concerning θ only) is

$$\theta = \frac{1}{m} \sum_{i=1}^{m} \left(y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x}) \right)^2 \leftarrow \frac{1}{m} H(\hat{\alpha}, \hat{\beta})$$

To see that ~~the~~ this critical pt is indeed a maximizer of $\ln(L(\hat{\alpha}, \hat{\beta}, \theta))$, we need to study the 2nd derivative (w.r.t. θ):

$$\frac{\partial^2}{\partial \theta^2} \ln(L(\hat{\alpha}, \hat{\beta}, \theta)) = \frac{\partial}{\partial \theta} \left[-\frac{m}{2} \frac{1}{\theta} + \frac{1}{2\theta^2} H(\hat{\alpha}, \hat{\beta}) \right]$$

$$= \frac{m}{2} \frac{1}{\theta^2} - \frac{1}{\theta^3} H(\hat{\alpha}, \hat{\beta})$$

Evaluate the 2nd derivative at the critical pt:

$$\frac{\partial^2}{\partial \theta^2} \ln(L(\hat{\alpha}, \hat{\beta}, \theta)) \Big|_{\theta = \frac{1}{m} H(\hat{\alpha}, \hat{\beta})} = \frac{m}{2} \frac{1}{\left(\frac{1}{m} H(\hat{\alpha}, \hat{\beta})\right)^2} - \frac{1}{\left(\frac{1}{m} H(\hat{\alpha}, \hat{\beta})\right)^3} \cdot H(\hat{\alpha}, \hat{\beta})$$

$$= \frac{n^3}{2\left(H(\hat{\alpha},\hat{\beta})\right)^2} - \frac{n^3}{\left(H(\hat{\alpha},\hat{\beta})\right)^2} = -\frac{1}{2}\frac{n^3}{\left(H(\hat{\alpha},\hat{\beta})\right)^2} < 0$$

$$\Rightarrow \quad \hat{\theta} = \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x})\right]^2$$

is a maximizer of $\ln\left(L(\hat{\alpha},\hat{\beta},\theta)\right)$

$\Rightarrow$ $\widehat{\sigma^2}$ is the MLE of σ^2 we were seeking

Evaluation of $\widehat{\beta}$

- To evaluate the maximum likelihood estimator of β,

from the statement of the theorem →

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)\left(x_i - \bar{x}\right)}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2},$$

we may resort to the following identities:

$$\sum_{i=1}^{n} \bar{Y}(x_i - \bar{x}) = \bar{Y}\overbrace{\sum_{i=1}^{n} x_i - \bar{x}}^{0} = \bar{Y}\cdot 0$$

zero!!

x_i y_i x_i^2 y_i^2 x_iy_i

$\sum x_i$ $\sum y_i$ $\sum x_i^2$ $\sum y_i^2$ $\sum x_iy_i$

$$\begin{aligned}
\widehat{\beta} &= \frac{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)\left(x_i - \bar{x}\right)}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} = \frac{\sum_{i=1}^{n} Y_i(x_i - \bar{x}) \overbrace{- \sum_{i=1}^{n} \bar{Y}(x_i - \bar{x})}^{=0}}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\sum_{i=1}^{n} Y_i\left(x_i - \bar{x}\right)}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \\
&= \frac{\sum_{i=1}^{n} x_i Y_i - (1/n)\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} Y_i\right)}{\sum_{i=1}^{n} x_i^2 - (1/n)\left(\sum_{i=1}^{n} x_i\right)^2}.
\end{aligned}$$

??

To get to the final expression:

Numerator:

$$\sum_{i=1}^{m} y_i(x_i - \bar{x}) = \sum_{i=1}^{m} y_i x_i - \sum_{i=1}^{m} y_i \cdot \bar{x} =$$

$$= \sum_{i=1}^{m} y_i x_i - \underbrace{\bar{x}}_{\frac{1}{m}\sum x_i} \sum_{i=1}^{m} y_i = \sum_{i=1}^{m} y_i x_i - \left(\frac{1}{m}\sum_{i=1}^{m} x_i\right)\left(\sum_{i=1}^{m} y_i\right)$$

Denominator:

$$\sum_{i=1}^{m}(x_i - \bar{x})^2 = \sum_{i=1}^{m}\left(x_i^2 - 2x_i\bar{x} + \bar{x}^2\right) = \sum_{i=1}^{m} x_i^2 - 2\bar{x}\overbrace{\sum_{i=1}^{m} x_i}^{m\bar{x}} + m\bar{x}^2$$

$$= \sum_{i=1}^{m} x_i^2 \underbrace{- 2\bar{x}\cdot m\bar{x} + m\bar{x}^2}_{-2m\bar{x} + m\bar{x}^2 = -m\bar{x}^2} = \sum_{i=1}^{m} x_i^2 - m\bar{x}^2$$

$$= \sum_{i=1}^{m} x_i^2 - m\left(\frac{1}{m}\sum_{i=1}^{m} x_i\right)^2 = \sum_{i=1}^{m} x_i^2 - \frac{1}{m}\left(\sum_{i=1}^{m} x_i\right)^2$$

Evaluation of $\widehat{\sigma^2}$

- To evaluate the maximum likelihood estimator of σ^2,

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \widehat{\alpha} - \widehat{\beta}\,(x_i - \bar{x})\right]^2 \,,$$

from the statement

???

we may resort to the following identity for $n\widehat{\sigma^2}$:

$$n\widehat{\sigma^2} = \sum_{i=1}^{n} Y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} Y_i\right)^2 - \widehat{\beta}\sum_{i=1}^{n} x_i Y_i + \widehat{\beta}\left(\frac{1}{n}\right)\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} Y_i\right) .$$

for use in practical examples when we're given $\sum x_i$, $\sum x_i^2$, $\sum y_i$, etc...

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{\alpha} - \hat{\beta}(x_i - \bar{x})\right)^2 \implies n\hat{\sigma}^2 = \sum_{i=1}^{n}\left[(y_i - \bar{y}) - \hat{\beta}(x_i - \bar{x})\right]^2$$

$$\hat{\alpha} = \bar{y}$$

$$n\hat{\sigma}^2 = \sum_{i=1}^{n}\left[(y_i - \bar{y})^2 - 2\hat{\beta}(y_i - \bar{y})(x_i - \bar{x}) + \hat{\beta}^2(x_i - \bar{x})^2\right]$$

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

$$= \sum_{i=1}^{n}(y_i - \bar{y})^2 + \hat{\beta}\left[-2\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x}) + \hat{\beta}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]$$

$$= \sum_{i=1}^{n}(y_i - \bar{y})^2 + \hat{\beta}\left[-2\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x}) + \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\cdot\sum(x_i - \bar{x})^2\right]$$

$$m\widehat{\sigma}^2 = \boxed{\sum_{i=1}^{m}(y_i-\bar{y})^2} - \hat{\beta}\boxed{\sum_{i=1}^{m}(y_i-\bar{y})(x_i-\bar{x})}$$

$$= \sum_{i=1}^{m} y_i^2 - \frac{1}{m}\left(\sum_{i=1}^{m} y_i\right)^2 - \hat{\beta}\left(\sum_{i=1}^{m} y_i x_i - \frac{1}{m}\left(\sum_{i=1}^{m} x_i\right)\left(\sum_{i=1}^{m} y_i\right)\right)$$

$$m\widehat{\sigma}^2 = \sum_{i=1}^{m} y_i^2 - \frac{1}{m}\left(\sum_{i=1}^{m} y_i\right)^2 - \hat{\beta}\sum_{i=1}^{m} y_i x_i + \hat{\beta}\cdot\frac{1}{m}\sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i$$

Alternative approach:

- Let $\widehat{Y}_i$ denote the estimated mean value of Y_i:

$$\boxed{\widehat{Y}_i = \widehat{\alpha} + \widehat{\beta}\,(x_i - \bar{x})\ .}$$

- The difference

$$Y_i - \widehat{Y}_i = Y_i - \widehat{\alpha} - \widehat{\beta}\,(x_i - \bar{x})$$

is called the *ith residual*, $i = 1, 2, \ldots, n$.

- The maximum likelihood estimator of σ^2

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \widehat{\alpha} - \widehat{\beta}\,(x_i - \bar{x})\right]^2 = \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \widehat{Y}_i\right]^2 ,$$

is the arithmetic average of the squares of the residuals.

useful when

Y_i $\hat{Y}_i$ $Y_i - \hat{Y}_i$

Remark

The sum of the residuals is zero:

$$\sum_{i=1}^{n} \left[Y_i - \widehat{Y}_i \right] = 0 \; .$$

Homework: prove this formula!

Example

$n = 10$

The table contains ten pairs of test scores of ten students in a certain class, with x being the score on a preliminary test and y the score on the final examination. Also included in the table are the sums needed to calculate estimates for a simple regression model.

Data set

	x	y	x^2	xy	y^2	
x_1	70	77	4,900	5,390	5,929	
x_2	74	94	5,476	6,956	8,836	
	72	88	5,184	6,336	7,744	
	68	80	4,624	5,440	6,400	
	58	71	3,364	4,118	5,041	
	54	76	2,916	4,104	5,776	
	82	88	6,724	7,216	7,744	
	64	80	4,096	5,120	6,400	
	80	90	6,400	7,200	8,100	
x_{10}	61	69	3,721	4,209	4,761	
	683	813	47,405	56,089	66,731	Sum of values above!

Example (continued)

$[n = 10]$

Determine:

a) the least squares regression line for these data;

b) the estimate for $\widehat{\sigma^2}$.

a) The least squares regression line is

$$y = \hat{\alpha} + \hat{\beta}(x - \bar{x}) = 81.3 + 0.742\,(x - 68.3)$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}\,683 = 68.3$$

$$\hat{\alpha} = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{10}\,813 = 81.3$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} = \frac{56{,}089 - \frac{1}{10}(683)\cdot(813)}{47{,}405 - \frac{1}{10}(683)^2} \approx 0.742$$

To get $\hat{\sigma}^2$; we may proceed as follows:

for each x_i, $i = 1, 2, \dots, 10$, evaluate $\hat{Y}_i = \hat{\alpha} + \hat{\beta}(x_i - \bar{x})$

$$\hat{Y}_i = 813 + 0.742(x_i - 68.3), \quad i = 1, \dots, 10$$

evaluate residuals $Y_i - \hat{Y}_i$, $i = 1, \dots, 10$

evaluate residuals squared $[Y_i - \hat{Y}_i]^2$, $i = 1, \dots, 10$

$$\widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} [Y_i - \hat{Y}_i]^2$$

residuals

x	y	x^2	xy	y^2	$\widehat{y}$	$y - \widehat{y}$	$(y - \widehat{y})^2$
70	77	4,900	5,390	5,929	82.561566	–5.561566	30.931016
74	94	5,476	6,956	8,836	85.529956	8.470044	71.741645
72	88	5,184	6,336	7,744	84.045761	3.954239	15.636006
68	80	4,624	5,440	6,400	81.077371	–1.077371	1.160728
58	71	3,364	4,118	5,041	73.656395	–2.656395	7.056434
54	76	2,916	4,104	5,776	70.688004	5.311996	28.217302
82	88	6,724	7,216	7,744	91.466737	–3.466737	12.018265
64	80	4,096	5,120	6,400	78.108980	1.891020	3.575957
80	90	6,400	7,200	8,100	89.982542	0.017458	0.000305
61	69	3,721	4,209	4,761	75.882687	–6.882687	47.371380
683	813	47,405	56,089	66,731		0.000001	217.709038

$$\widehat{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \frac{1}{10} 217.71 = 21.771$$