Consider a linear estimator for β of the form $\sum_{i=1}^{n} d_i Y_i$ ← constants

It is unbiased if

$$E\left(\sum_{i=1}^{n} d_i Y_i\right) = \beta \quad \Rightarrow \beta = \sum_{i=1}^{n} d_i(\alpha+\beta x_i) = \alpha\sum_{i=1}^{n} d_i + \beta\sum_{i=1}^{n} d_i x_i$$

is true for any β. Thus

$$\sum_{i=1}^{n} d_i = 0 \text{ and } \sum_{i=1}^{n} d_i x_i = 1$$

What about the best linear unbiased estimator (BLUE)?

The BLUE is the one with the smallest variance:

$$Var\left(\sum_{i=1}^{n} d_i Y_i\right) = \sum_{i=1}^{n} d_i^2 Var(Y_i) = \sigma^2 \sum_{i=1}^{n} d_i^2$$

$\Rightarrow$ We have to minimize $\sum_{i=1}^{n} d_i^2$ "distance" from origin under the condition

$\sum_{i=1}^{n} d_i = 0$ and $\sum_{i=1}^{n} d_i x_i = 1$ intersection of 2 "planes"

The solution: $d_i = \dfrac{(x_i-\bar{x})}{S_{xx}}$

$$\Rightarrow Var\left(\sum_{i=1}^{n} d_i Y_i\right) = \frac{1}{S_{xx}^2}\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})^2}_{S_{xx}}\sigma^2 = \frac{\sigma^2}{S_{xx}}$$

Then we get (after an experiment)

$$b := \sum_{i=1}^{n}\frac{(x_i-\bar{x})}{S_{xx}} y_i \underset{\Uparrow}{=} \sum_{i=1}^{n}\frac{(x_i-\bar{x})(y_i-\bar{y})}{S_{xx}} = \frac{S_{xy}}{S_{xx}}$$

$$\sum (x_i-\bar{x})\bar{y} = \bar{y}\sum(x_i-\bar{x}) = 0$$

This coefficient is the one obtained by data fitting.

We can do the same for α and one gets the BLUE for α gives $\alpha = \bar{y} - b\bar{x}$.

If we do further analysis, we need to impose more on the distribution of ε_i.

We consider the normal model: $\varepsilon_i \sim n(0, \sigma^2)$ ↑ ↑ imposed

In this context, $Y_i \sim n(\alpha+\beta x_i, \sigma^2)$

Lemma: Assume $Y_i \sim n(\alpha+\beta x_i, \sigma^2)$ then ($\beta = \sum d_i Y_i$ for β)

the sample distribution for $\hat{\alpha}$ and $\hat{\beta}$ given by the BLUE are given by

$$\hat{\beta} \sim n\left(\beta, \frac{\sigma^2}{S_{xx}}\right) \text{ and } \hat{\alpha} \sim n\left(\alpha, \frac{\sigma^2}{nS_{xx}}\sum_{i=1}^{n} x_i^2\right)$$

The sample variance S^2 given by

$$S^2 = \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}x_i)^2$$

The factor $1/(n-2)$ is imposed by the requirement that S^2 is unbiased.

satisfies $\frac{n-2}{\sigma^2} S^2 \sim \chi^2_{n-2}$

In addition, S^2 is independent of $\hat{\alpha}$ and $\hat{\beta}$,

but $\hat{\alpha}$ and $\hat{\beta}$ are not independent. [Thm. 11.3.3]

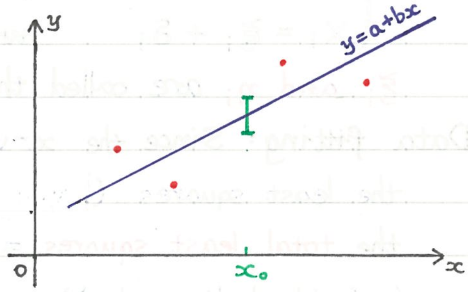
Corollory: $\dfrac{(\hat{\beta}-\beta)/\sqrt{\sigma^2/S_{xx}}}{\sqrt{S^2/\sigma^2}} = \dfrac{\hat{\beta}-\beta}{S/\sqrt{S_{xx}}} \sim t_{n-2}$ student t dist. ⊛

$$\frac{(\hat{\alpha}-\alpha)\sqrt{\frac{\sigma^2}{nS_{xx}}\sum x_i^2}}{\sqrt{S^2/\sigma^2}} = \frac{\hat{\alpha}-\alpha}{S\sqrt{\sum x_i^2/nS_{xx}}} \sim t_{n-2}$$

From ⊛ with the hypothesis $H_0 : \hat{\beta} = \beta$

we get the confidence interval with confident coef. $(1-\alpha)$ as

$$\left[\hat{\beta} - t_{n-2,\alpha/2}\frac{S}{\sqrt{S_{xx}}},\ \hat{\beta} + t_{n-2,\alpha/2}\frac{S}{\sqrt{S_{xx}}}\right]$$



What about estimating Y for a given x_0?

If we fix x_0, the point estimator for the corresponding Y is $Y = \hat{\alpha} + \hat{\beta}x_0$. One has

$$E(Y) = E(\hat{\alpha}) + x_0 E(\hat{\beta}) = \alpha + \beta x_0$$

$$\mathrm{Var}(Y) = \mathrm{Var}(\hat{\alpha}+\hat{\beta}x_0) = \mathrm{Var}(\hat{\alpha}) + x_0^2\,\mathrm{Var}(\hat{\beta}) + 2x_0\,\mathrm{Cov}(\hat{\alpha},\hat{\beta})$$

$$= \cdots = \sigma^2\left(\frac{1}{n} + \frac{(x_0-\bar{x})}{S_{xx}}\right)$$

As a consequence

$$Y = \hat{\alpha} + \hat{\beta}x_0 \sim n\left(\alpha+\beta x_0,\ \sigma^2\left(\frac{1}{n}+\frac{(x_0-\bar{x})}{S_{xx}}\right)\right)$$

In order to estimate the parameter σ^2 (which is usually unknown)

we use the sample variance S^2, which is independent of $\hat{\alpha}$ and $\hat{\beta}$, and get

$$\frac{\hat{\alpha}+\hat{\beta}x_0-(\alpha+\beta x_0)}{S\sqrt{\frac{1}{n}+\frac{(x_0-\bar{x})}{S_{xx}}}} \sim t_{n-2}$$

The related confidence interval with confident coeff. $1-\alpha$ is given by

$$\left[\hat{\alpha}+\hat{\beta}x_0 - t_{n-2,\alpha/2}\,S\sqrt{\frac{1}{n}+\frac{(x_0-\bar{x})}{S_{xx}}},\ \hat{\alpha}+\hat{\beta}x_0 + t_{n-2,\alpha/2}\,S\sqrt{\frac{1}{n}+\frac{(x_0-\bar{x})}{S_{xx}}}\right]$$

Remark: we have a smaller interval either by taking n large,

or by fixing $x_1, \cdots, x_n$ s.t. $x_0 \cong \bar{x}$.

We can also maximize S_{xx}.