Unbiasedness of LS Ests.

① $E(\hat{\beta}_0) = \beta_0$

② $E(\hat{\beta}_1) = \beta_1$

Variance of LS Ests.

② $\underline{Var}(\hat{\beta}_1) = \sigma^2 / SSX$

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

① $Var(\hat{\beta}_0) = \sigma^2 \left(\frac{1}{n} + \bar{x}^2 / SSX \right)$

Strategy:

You have to find c_i s.t. equality holds.

Ⓐ $\hat{\beta}_0 = \sum_{i=1}^{n} c_i y_i$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \boxed{\bar{y} - \left(\sum_{j=1}^{n} k_j y_j \right) \bar{x}}$

$$= \sum_{j=1}^{n} \frac{1}{n} y_j - \bar{x} \sum_{j=1}^{n} k_j y_j$$

$$= \sum_{j=1}^{n} \underbrace{\left(\frac{1}{n} - \bar{x} k_j \right)}_{c_j} y_j$$

|

(B)

$Var(\hat{\beta}_0) = Var(\sum_{i=1}^{n} c_i y_i)$

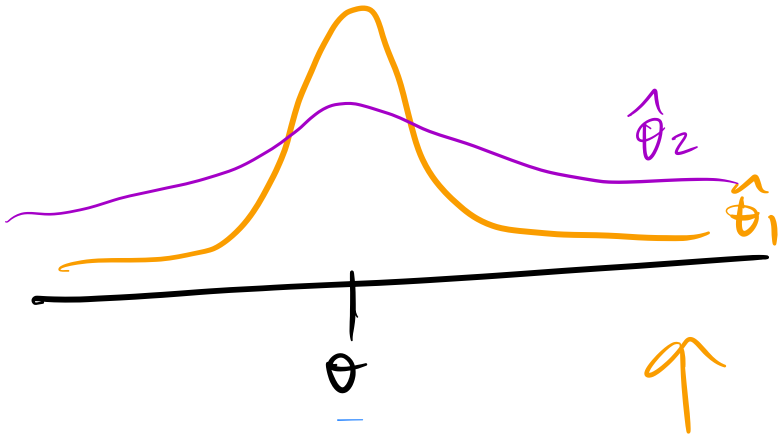
$= \sum_{i=1}^{n} c_i^2 \, Var(y_i) = \sigma^2 \sum_i c_i^2$

(C) Calculate $\sum_{i=1}^{n} c_i^2$.

<u>Minimum Variance Properties of the LS estimators</u>

For a general parameter θ, assume we have 2 possible estimators to consider: $\hat{\theta}_1$ & $\hat{\theta}_2$.

If both estimators are unbiased and we know $Var(\hat{\theta}_1) \leq Var(\hat{\theta}_2)$ for all values of θ, then this implies that $\hat{\theta}_1$ does a better job estimating θ than $\hat{\theta}_2$.



<u>Def</u>: An estimator $\hat{\theta}$ is the <u>b</u>est <u>l</u>inear <u>u</u>nbiased <u>e</u>stimator (BLUE) if & only if $\hat{\theta}$ has the smallest variance among <u>all</u> linear estimators which are unbiased for θ, $\forall \theta \in \Omega$.

all possible values of θ

i.e.

$$Var(\hat{\theta}) \le Var(\tilde{\theta}) \quad \forall \theta \in \Omega,$$

where $\tilde{\theta}$ is any other linear unbiased est.

<u>Theorem</u>: <u>Gauss - Markov</u> Thm:

Under the assumptions of the SLR model, the least squares estimators are <u>BLUE</u>.

<u>Pf Idea</u>:

$$\hat{\beta}_1 = \sum_{i=1}^{n} k_i y_i \quad \leftarrow \text{LS Est}$$

$$\tilde{\beta}_1 = \sum_{i=1}^{n} \tilde{k}_i y_i$$

Let's say $\tilde{k}_i = k_i + d_i$

$$Var(\tilde{\beta}_1) = Var\left(\sum_{i=1}^{n} (k_i + d_i) y_i\right)$$
$$\leq Var(\hat{\beta}_1)$$

Distribution of the LS Ests.

Thm: Under the classical SLR assumptions, the dists of the LSEs are:

i. $\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2\left[\frac{1}{n} + \frac{\bar{x}^2}{SSX}\right]\right)$

ii. $\hat{\beta}_1 \sim N\left(\beta_1, \sigma^2 / SSX\right)$

How do I argue for Normality?

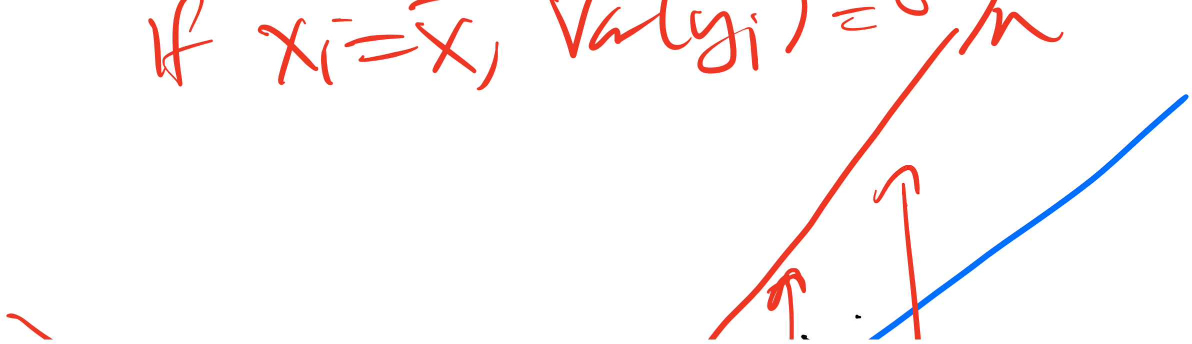
Fitted Values:

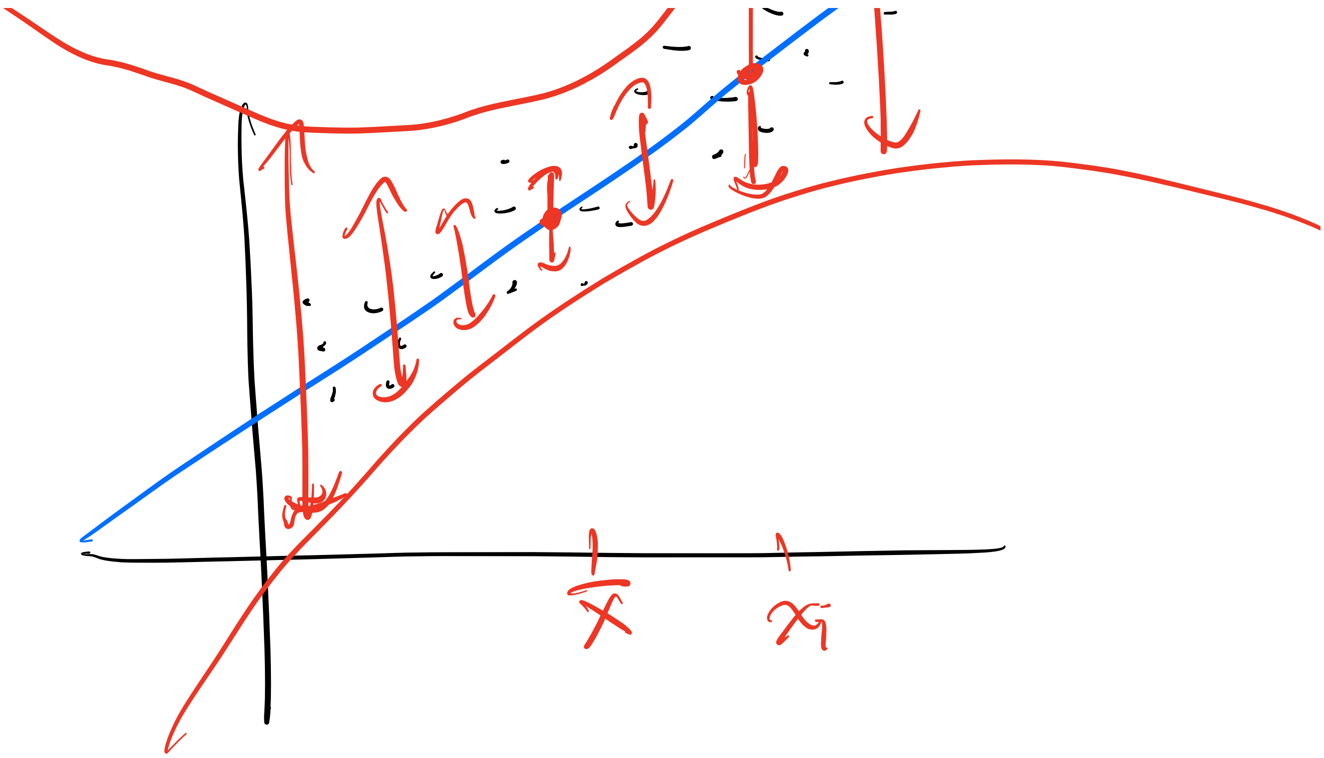
$\hat{y}_i$ are called the "fitted values".

Thm: Under the classical SLR assumptions, we have

$$\hat{y}_i \overset{1}{\sim} N\left(\overset{2}{\beta_0+\beta_1 x_i},\ \overset{3}{\sigma^2\left(\frac{1}{n}+\frac{(x_i-\bar{x})^2}{SSX}\right)}\right)$$

If $x_i=\bar{x}$, $Var(\hat{y}_i)=\sigma^2/n$





Pf:

① Why is the dist Normal?

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Lin. Comb of Normal RVs.

②

$$E(\hat{y}_i) = E(\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$$= E(\hat{\beta}_0) + E(\hat{\beta}_1 x_i)$$

$$= \beta_0 + x_i E(\hat{\beta}_1)$$

$$= \beta_0 + x_i \beta_1 = \beta_0 + \beta_1 x_i$$

③

$Var(\hat{y}_i)$

---

watch out

one way

$$Var(\hat{y}_i) = Var(\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$$= \underline{Var(\hat{\beta}_0)} + \underline{Var(\hat{\beta}_1 x_i)} + \underline{2Cov(\hat{\beta}_0, \hat{\beta}_1)}$$

$Var(\hat{y}_i) =$

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = \sum_{j=1}^{n} c_j y_j + x_i \sum_{j=1}^{n} k_j y_j$$

$$k_j = \frac{x_j - \bar{x}}{SSX}$$

$$= \sum_{j=1}^{n} (c_j + x_i k_j) y_j$$

$$= \sum_{j=1}^{n} \left(\frac{1}{n} - \bar{x} k_j + x_i k_j\right) y_j$$

$$= \sum_{j=1}^{n} \left(\frac{1}{n} + (x_i - \bar{x}) k_j\right) y_j$$

$$Var(\hat{y}_i) = Var\left(\sum_{j=1}^{n} \left(\frac{1}{n} + (x_i - \bar{x}) k_j\right) y_j\right)$$

$$\perp$$

$$= \sum_{j=1}^{n} Var\left(\left(\frac{1}{n} + (x_i - \bar{x}) k_j\right) y_j\right)$$

$$= \sum_{j=1}^{n} \left(\frac{1}{n} + (x_i - \bar{x}) k_j\right)^2 Var(y_j)$$

$$= \sum_{j=1}^{n} \left(\frac{1}{n} + (x_i - \bar{x}) k_j\right)^2 \sigma^2$$

$$= \sigma^2 \sum_{j=1}^{n} \left(\frac{1}{n} + (x_i - \bar{x}) k_j\right)^2$$

$$= \sigma^2 \sum_{j=1}^{n} \left(\frac{1}{n^2} + \frac{2(x_i - \bar{x}) k_j}{n} + \left[(x_i - \bar{x}) k_j\right]^2\right)$$

$$0$$

$$= \sigma^2 \left[\frac{1}{n} + \boxed{\cancel{\frac{2(x_i - \bar{x})}{n} \sum_{j=1}^{n} k_j}} + (x_i - \bar{x})^2 \sum_{j=1}^{n} k_j^2\right]$$

① $$\sum_{j=1}^{n} k_j = \sum_{j=1}^{n} \frac{x_j - \bar{x}}{SSX}$$

$$= \frac{1}{SSX} \sum_{j=1}^{n} (x_j - \bar{x})$$

$$= \frac{1}{SSX}\left(n\bar{x} - n\bar{x}\right) = 0$$

② $$\sum_{j=1}^{n} k_j^2 = \sum_{j=1}^{n} \left(\frac{(x_j - \bar{x})}{SSX}\right)^2$$

$$= \frac{SSX}{(SSX)^2} = 1/SSX$$

$$= \sigma^2\left[\frac{1}{n} + \frac{(x_i - \bar{x})^2}{SSX}\right]$$