● **一致推定量（Consistent Estimator）：**　$\hat{\alpha}$，$\hat{\beta}$ **が** α，β **の一致推定量であることをそれぞれ示す。**

(* 復習) 望ましい推定量について：

n **個の確率変数** X_1，X_2，$\cdots$，X_n **は互いに独立で，**

それぞれは母数 θ **に依存するものとする（例えば，**$\theta = (\mu,\ \sigma^2)$ **である）。**

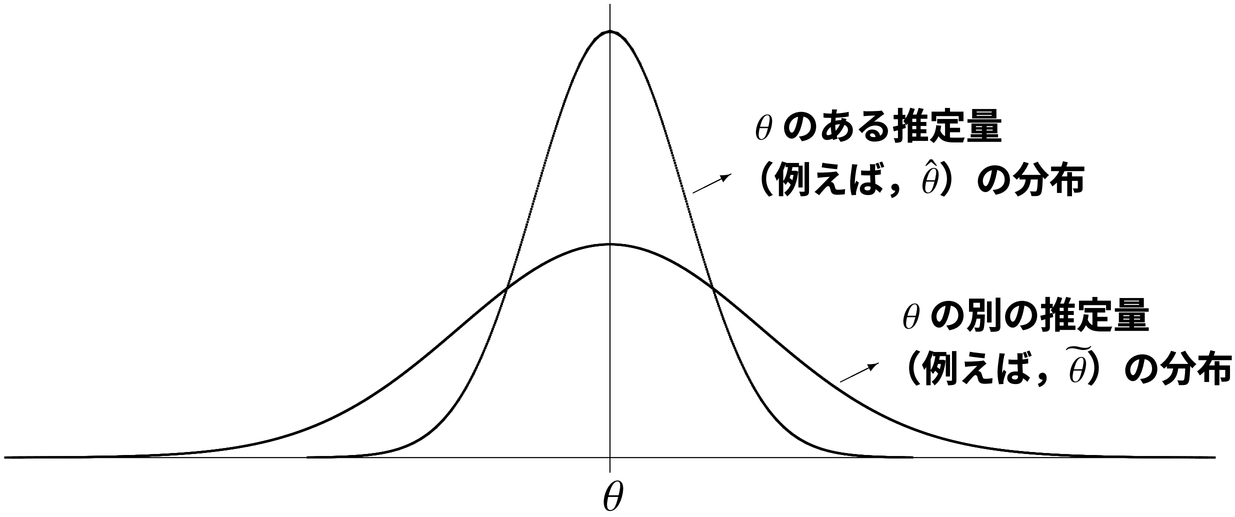
θ **の推定量を** $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ **とする。**

(1) 不偏推定量：　$\mathbf{E}(\hat{\theta}) = \theta$

(2) 有効推定量：　**不偏推定量の中で最も分散の小さい推定量**

(3) 一致推定量：　$\hat{\theta}$ **を** $\hat{\theta}_n$ **と表記する。「**$n \longrightarrow \infty$ **のとき（**n **が大きくなるとき），**

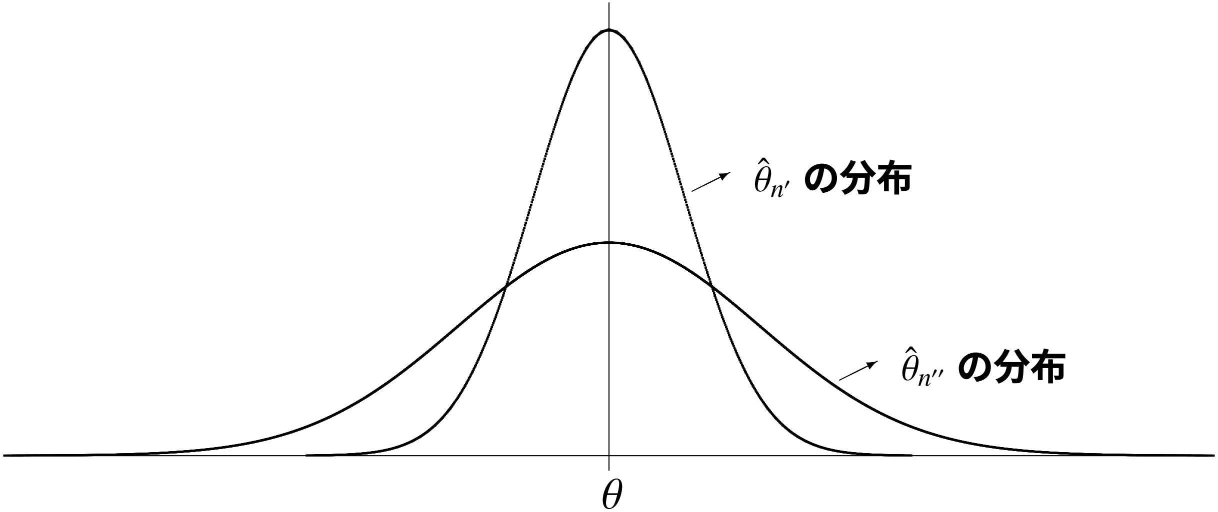
$\hat{\theta}_n \longrightarrow \theta$ **となる（**$\hat{\theta}_n$ **が** θ **に収束する）」とき，**$\hat{\theta}_n$ **は** θ **の一致推定量であるという。**



目的： θ の値を知りたい。

分散の小さい推定量の方が θ の範囲を絞ることができる。

$\mathbf{V}(\hat{\theta}) < \mathbf{V}(\widetilde{\theta})$ なので，$\hat{\theta}$ が $\widetilde{\theta}$ より良い推定量と言える。



$n' > n''$ のとき，$\mathbf{V}(\hat{\theta}_{n'}) < \mathbf{V}(\hat{\theta}_{n''})$ となる。

$n \longrightarrow \infty$ のとき，$\hat{\theta}$ の分布は θ に一点集中。

(* 復習) $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ は μ の最良線形不変推定量か？:

$X_1, X_2, \cdots, X_n$ は互いに独立で，X_i の平均 μ，分散 σ^2 とする。

μ の別の線形推定量を $\widetilde{X} = \sum_{i=1}^{n} c_i X_i$ として考える（c_i は定数）。

$\widetilde{X}$ が μ の不偏推定量となるためには，$\mathbf{E}(\widetilde{X}) = \mathbf{E}(\sum_{i=1}^{n} c_i X_i) = \sum_{i=1}^{n} \mathbf{E}(c_i X_i) = \sum_{i=1}^{n} c_i \mathbf{E}(X_i) = \mu \sum_{i=1}^{n} c_i$

なので，$\sum_{i=1}^{n} c_i = 1$ とならなければならない。

$\widetilde{X}$ の分散は $\mathbf{V}(\widetilde{X}) = \mathbf{V}(\sum_{i=1}^{n} c_i X_i) = \sum_{i=1}^{n} \mathbf{V}(c_i X_i) = \sum_{i=1}^{n} c_i^2 \mathbf{V}(X_i) = \sigma^2 \sum_{i=1}^{n} c_i^2$ となる。

$\sum_{i=1}^{n} c_i = 1$ のもとで，$\sum_{i=1}^{n} c_i^2$ を最小になる $c_1, c_2, \cdots, c_n$ を求める。

ラグランジェ未定乗数法により，$c_1, c_2, \cdots, c_n$ を求める。

ラグランジェ関数 $L = \sum_{i=1}^{n} c_i^2 - 2\lambda(\sum_{i=1}^{n} c_i - 1)$ を c_i と λ（ラグランジェ乗数）について微分する。

$$\frac{\partial L}{\partial c_i} = 0 \qquad \Longrightarrow \qquad c_i = \lambda \qquad i = 1, 2, \cdots, n$$

$$\frac{\partial L}{\partial \lambda} = 0 \qquad \Longrightarrow \qquad \sum_{i=1}^{n} c_i = 1$$

解くと，$i = 1, 2, \cdots, n$ について，$c_i = \dfrac{1}{n}$ となる。

すなわち，$\widetilde{X} = \overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ が μ の最良（最小分散）線形不偏推定量となる。

(* 復習) $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ は μ の一致推定量か ?:

$X_1, X_2, \cdots, X_n$ は互いに独立で，X_i の平均 μ，分散 σ^2 とする。

μ の推定量を $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ とする。

$$\mathbf{E}(\overline{X}) = \mathbf{E}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(X_i) = \mu$$

$$\mathbf{V}(\overline{X}) = \mathbf{V}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{V}(X_i) = \frac{\sigma^2}{n}$$

すなわち，$\overline{X}$ は μ の周りに分布していて，n が大きくなるにつれて $\mathbf{V}(\overline{X})$ はゼロに収束する。

$n \longrightarrow \infty$ のとき，$\overline{X} \longrightarrow \mu$ となる。

$\overline{X}$ は μ の一致推定量である。

● $\hat{\beta}$ の一致性について：　$\mathrm{E}(\hat{\beta}) = \beta$ となることは既に証明した。

n が大きくなると，$\mathrm{V}(\hat{\beta})$ はゼロに近づくかどうかを調べる。

$$\mathrm{V}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \overline{X})^2}$$

$n \longrightarrow \infty$ のとき，$\mathrm{V}(\hat{\beta}) \longrightarrow 0$ となれば，$\hat{\beta}$ は β の一致推定量となる。

最小二乗法の仮定の一つに，「$n \longrightarrow \infty$ のとき，$\sum_{i=1}^n (X_i - \overline{X})^2 \longrightarrow \infty$」というものがあることを思い出して欲しい。

この仮定は，n が大きくなると，2 乗したものを次々に足し合わせていくことなので，現実的な仮定と言える。

この仮定が，「$n \longrightarrow \infty$ のとき，$\mathrm{V}(\hat{\beta}) \longrightarrow 0$」を保証する。

したがって，$\hat{\beta}$ は β の一致推定量となる。

● $\hat{\alpha}$ の一致性について： $\hat{\alpha}$ についても，同様に，$\mathrm{E}(\hat{\alpha}) = \alpha$ であることは分かっている。

$$\mathbf{V}(\hat{\alpha}) = \sigma^2\Big(\frac{1}{n} + \frac{\overline{X}^2}{\sum_{i=1}^n (X_i - \overline{X})^2}\Big) = \frac{\sigma^2 \sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \overline{X})^2}$$

となる。

$\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ はデータの平均値なので，n が大きくなると $\overline{X}$ は何らかの値に収束すると考えるのが自然である。

「$n \longrightarrow \infty$ のとき，$\mathrm{V}(\hat{\alpha}) \longrightarrow 0$」となるので，$\hat{\alpha}$ も α の一致推定量であると言える。

● まとめ：　α，β の最小二乗推定量 $\hat{\alpha}$，$\hat{\beta}$ は，

1. 不偏推定量（Unbiased Estimator）

2. 最良線形不偏推定量（Best Linear Unbiased Estimator）　$\Longrightarrow$ 有効推定量（または，最良不偏推定量）ではない

3. 一致推定量（Consistent Estimator）

である。

望ましい推定量の持つべき性質を概ね持っている。

5.3.7 誤差項 (または，攪乱項) u_i の分散 σ^2 について

(* 復習) 分散の推定量について：

n 個の確率変数 X_1，X_2，$\cdots$，X_n は互いに独立で，

X_i は平均 μ，分散 σ^2 の分布（代表的な分布は正規分布）に従う。

- μ の推定量は $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$
- σ^2 の推定量は $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2 \quad \Longrightarrow \quad n-1$ は自由度

S^2 の特徴： $\mathrm{E}(S^2) = \sigma^2 \qquad \Longrightarrow \quad S^2$ は標本不偏分散と呼ばれる。

X_i の分布に正規分布を仮定すれば，$\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} \sim \chi^2(n-1)$

$\Longrightarrow$ 自由度 $n-1$ のカイ二乗分布

回帰モデル：

$$Y_i = \alpha + \beta X_i + u_i$$

誤差項 (または，攪乱項) の仮定：$u_i \sim N(0, \sigma^2)$

$$Y_i = \hat{\alpha} + \hat{\beta} X_i + \hat{u}_i$$

u_i **の分散** σ^2 **の不偏推定量** s^2**：**

$$s^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{\textbf{自由度}}$$

単回帰の場合は,「自由度 $=$ **標本サイズ（**n**）**$-$ **パラメータ数（**2**）**$= n - 2$**」である。**

誤差項 (または，攪乱項) の母分散 σ^2 の不偏推定量 s^2 は，

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}X_i)^2$$

によって与えられる。

● s^2 の不偏性の証明：　u_i の平均ゼロ・分散 σ^2 を仮定する（正規分布の仮定は必要なし）。

まず，次のように書き直す。

$$u_i = Y_i - \alpha - \beta X_i = (\hat{\alpha} + \hat{\beta}X_i + \hat{u}_i) - \alpha - \beta X_i = (\hat{\alpha} - \alpha) + (\hat{\beta} - \beta)X_i + \hat{u}_i$$

2 つ目の等式では，$Y_i = \hat{\alpha} + \hat{\beta}X_i + \hat{u}_i$ が代入されている。

両辺を二乗する。

$$u_i^2 = (\hat{\alpha} - \alpha)^2 + (\hat{\beta} - \beta)^2 X_i^2 + \hat{u}_i^2 + 2(\hat{\alpha} - \alpha)(\hat{\beta} - \beta)X_i$$
$$+2(\hat{\alpha} - \alpha)\hat{u}_i + 2(\hat{\beta} - \beta)X_i\hat{u}_i$$

次に，両辺について総和をとる。

$$\begin{aligned}\sum_{i=1}^{n} u_i^2 = n(\hat{\alpha} - \alpha)^2 + (\hat{\beta} - \beta)^2 \sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n} \hat{u}_i^2 + 2(\hat{\alpha} - \alpha)(\hat{\beta} - \beta) \sum_{i=1}^{n} X_i \\ +2(\hat{\alpha} - \alpha) \sum_{i=1}^{n} \hat{u}_i + 2(\hat{\beta} - \beta) \sum_{i=1}^{n} X_i\hat{u}_i \\ = n(\hat{\alpha} - \alpha)^2 + (\hat{\beta} - \beta)^2 \sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n} \hat{u}_i^2 + 2n(\hat{\alpha} - \alpha)(\hat{\beta} - \beta)\overline{X}\end{aligned}$$

2 つ目の等式になるためには，$\sum_{i=1}^{n} \hat{u}_i = 0$，$\sum_{i=1}^{n} X_i\hat{u}_i = 0$ が利用されている。この 2 つの式は

残差平方和の最小化問題を解く際に得られたことを思い起こせ。

さらに，両辺について期待値をとる。

$$\mathbf{E}\Big(\sum_{i=1}^{n} u_i^2\Big) = n\mathbf{E}\Big((\hat{\alpha}-\alpha)^2\Big) + \mathbf{E}\Big((\hat{\beta}-\beta)^2\Big)\sum_{i=1}^{n} X_i^2 + \mathbf{E}\Big(\sum_{i=1}^{n} \hat{u}_i^2\Big) + 2n\mathbf{E}\Big((\hat{\alpha}-\alpha)(\hat{\beta}-\beta)\Big)\overline{X}$$

右辺第 3 項を除いて，それぞれの期待値は，

$$\mathbf{E}\Big(\sum_{i=1}^{n} u_i^2\Big) = \sum_{i=1}^{n} \mathbf{E}(u_i^2) = \sum_{i=1}^{n} \sigma^2 = n\sigma^2$$

$$\mathbf{E}\Big((\hat{\alpha}-\alpha)^2\Big) = \mathbf{V}(\hat{\alpha}) = \frac{\sigma^2 \sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

$$\mathbf{E}\Big((\hat{\beta}-\beta)^2\Big) = \mathbf{V}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

$$\mathbf{E}\Big((\hat{\alpha}-\alpha)(\hat{\beta}-\beta)\Big) = \mathbf{Cov}(\hat{\alpha},\,\hat{\beta}) = \frac{-\sigma^2\overline{X}}{\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

となるので，それぞれを代入すると，

$$
\begin{aligned}
n\sigma^2 &= \frac{\sigma^2 \sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i - \overline{X})^2} + \frac{\sigma^2 \sum_{i=1}^n X_i^2}{\sum_{i=1}^n (X_i - \overline{X})^2} + \mathbf{E}(\sum_{i=1}^n \hat{u}_i^2) - \frac{2n\sigma^2 \overline{X}^2}{\sum_{i=1}^n (X_i - \overline{X})^2} \\
&= 2\sigma^2 \left(\frac{\sum_{i=1}^n X_i^2 - n\overline{X}^2}{\sum_{i=1}^n (X_i - \overline{X})^2} \right) + \mathbf{E}(\sum_{i=1}^n \hat{u}_i^2) = 2\sigma^2 + \mathbf{E}(\sum_{i=1}^n \hat{u}_i^2)
\end{aligned}
$$

となる。最後の等式では，$\sum_{i=1}^n X_i^2 - n\overline{X}^2 = \sum_{i=1}^n (X_i - \overline{X})^2$ が使われている。

すなわち，

$$
\mathbf{E}(\sum_{i=1}^n \hat{u}_i^2) = (n-2)\sigma^2
$$

となる。

よって，

$$\mathbf{E}(s^2) = \mathbf{E}(\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}) = \frac{1}{n-2}\mathbf{E}(\sum_{i=1}^{n} \hat{u}_i^2) = \sigma^2$$

を得る。

すなわち，s^2 は σ^2 の不偏推定量である。$\Longrightarrow$ u_i に正規分布を仮定する必要なし

● s^2 の分布（u_i に正規分布を仮定）：

(* 復習) カイ二乗分布について：

- $Z \sim N(0,1)$，$U = Z^2$ のとき，$U \sim \chi^2(1)$
- $Z_1, Z_2, \cdots, Z_n$ は互いに独立で，$Z_i \sim N(0,1)$，$U = \sum_{i=1}^{n} Z_i^2$ のとき，$U \sim \chi^2(n)$
- $X_1, X_2, \cdots, X_n$ は互いに独立で，$X_i \sim N(\mu, \sigma^2)$，$U = \sum_{i=1}^{n} \left(\dfrac{X_i - \mu}{\sigma}\right)^2$ のとき，$U \sim \chi^2(n)$
- $X_1, X_2, \cdots, X_n$ は互いに独立で，$X_i \sim N(\mu, \sigma^2)$，$U = \sum_{i=1}^{n} \left(\dfrac{X_i - \overline{X}}{\sigma}\right)^2$ のとき，$U \sim \chi^2(n-1)$

    ただし，$\overline{X} = \dfrac{1}{n}\sum_{i}^{n} X_i$

回帰分析に当てはめる。

回帰モデルは，

$$Y_i = \alpha + \beta X_i + u_i \qquad i = 1, 2, \cdots, n$$

ただし，$u_1, u_2, \cdots, u_n$ はそれぞれ互いに独立で，

すべての $i = 1, 2, \cdots, n$ について，$u_i \sim N(0, \sigma^2)$ を仮定する。

である。

$u_i \sim N(0, \sigma^2)$ なので，$\dfrac{u_i}{\sigma} \sim N(0, 1)$ となる。

さらに，$\left(\dfrac{u_i}{\sigma}\right)^2 \sim \chi^2(1)$ となる。

$u_1, u_2, \cdots, u_n$ はそれぞれ互いに独立であると仮定しているので，$\displaystyle\sum_{i=1}^{n}\left(\frac{u_i}{\sigma}\right)^2 \sim \chi^2(n)$ となる。

$u_i = Y_i - \alpha - \beta X_i$ を代入すると，

$$\frac{\sum_{i=1}^{n} u_i^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(Y_i - \alpha - \beta X_i)^2}{\sigma^2} \sim \chi^2(n)$$

となる。

α，β を推定量 $\hat{\alpha}$，$\hat{\beta}$ に置き換えると，

$$\frac{\sum_{i=1}^{n} \hat{u}_i^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta} X_i)^2}{\sigma^2} \sim \chi^2(n-2)$$

となる。自由度は「標本サイズ (n) − パラメータ数 $(2) = n - 2$」

さらに，

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

なので，

$$\frac{\sum_{i=1}^{n}\hat{u}_i^2}{\sigma^2} = \frac{(n-2)s^2}{\sigma^2} \sim \chi^2(n-2)$$

を得る。

● s^2 の不偏性の証明（別解）：

(* 復習) カイ二乗分布の平均・分散について：

$U \sim \chi^2(k)$ のとき，$\mathbf{E}(U) = k$，$\mathbf{V}(U) = 2k$ となる。

$u_1, u_2, \cdots, u_n$ は互いに独立で，u_i に正規分布を仮定すると，

$$U = \frac{\sum_{i=1}^n \hat{u}_i^2}{\sigma^2} = \frac{(n-2)s^2}{\sigma^2} \sim \chi^2(n-2)$$

となるので，

$$\mathbf{E}\Big(\frac{(n-2)s^2}{\sigma^2}\Big) = n-2$$

となる。よって，

$$\mathbf{E}(s^2) = \frac{\sigma^2}{n-2}(n-2) = \sigma^2$$

を得る。s^2 は σ^2 の不偏推定量となっている。

前述の証明より非常に簡単な証明になっているが，この場合は，$u_i \sim N(0, \sigma^2)$ の仮定（正規分布の仮定）が必要となる。

● s^2 の一致性の証明：　簡単化のために，$u_i \sim N(0, \sigma^2)$ を仮定（正規分布の仮定）する。

s^2 は

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}X_i)^2$$

と定義する。

このとき，$\dfrac{(n-2)s^2}{\sigma^2} \sim \chi^2(n-2)$ なので，

$$\mathbf{V}\left(\frac{(n-2)s^2}{\sigma^2}\right) = 2(n-2)$$

となる。

さらに，書き直すと，

$$\frac{(n-2)^2}{\sigma^4}\mathbf{V}(s^2) = 2(n-2) \qquad \mathbf{V}(s^2) = \frac{2\sigma^4}{n-2}$$

を得る。

「$\mathbf{E}(s^2) = \sigma^2$ で，しかも，$n \longrightarrow \infty$ のとき $\mathrm{V}(s^2) \longrightarrow 0$」が言えるので，$s^2$ は σ^2 の一致推定量である。

● まとめ：　σ^2 の推定量 $s^2 = \dfrac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2 = \dfrac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}X_i)^2$ について，

1. 不偏推定量（Unbiased Estimator）

2. 一致推定量（Consistent Estimator）

となっている。

　有効推定量（または，最良不偏推定量）ではないが，他の推定量としての望ましい性質を持っている。

● **標準誤差（Standard Error）について：　標準誤差 = 不偏分散の平方根**

誤差項 (または，攪乱項) の標準誤差 s

$$s = \sqrt{\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}}$$