Утверждение. Если для парной регрессии выполнены условия теоремы Гаусса-Маркова, то $\hat{\sigma}_\varepsilon^2 = \frac{RSS}{n-2}$ является несмещённой оценкой параметра дисперсии ошибок регрессии σ_ε^2.

Доказательство

$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$; $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

$e_i = Y_i - \hat{Y}_i = (Y_i - \bar{Y}) - (\hat{Y}_i - \bar{Y}) = y_i - [(\hat{\beta}_0 + \hat{\beta}_1 X_i) - (\hat{\beta}_0 + \hat{\beta}_1 \bar{X})]$
$= y_i - \hat{\beta}_1 x_i$

$y_i = Y_i - \bar{Y} = (\beta_0 + \beta_1 X_i + \varepsilon_i) - (\beta_0 + \beta_1 \bar{X} + \bar{\varepsilon}) = (\varepsilon_i - \bar{\varepsilon}) + \beta_1 x_i$

$\Rightarrow e_i = (\varepsilon_i - \bar{\varepsilon}) - (\hat{\beta}_1 - \beta_1) x_i$

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2 + (\hat{\beta}_1 - \beta_1)^2 \sum_{i=1}^n x_i^2 - 2(\hat{\beta}_1 - \beta_1)\sum_{i=1}^n x_i(\varepsilon_i - \bar{\varepsilon})$$

$$E\left[\sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2\right] = E\left(\sum_{i=1}^n \varepsilon_i^2 - 2\bar{\varepsilon}\underbrace{\sum_{i=1}^n \varepsilon_i}_{=n\bar{\varepsilon}} + n\bar{\varepsilon}^2\right) = E\left(\sum_{i=1}^n \varepsilon_i^2 - n\bar{\varepsilon}^2\right) =$$
$$= \sum_{i=1}^n E(\varepsilon_i^2) - nE(\bar{\varepsilon}^2) = n\sigma_\varepsilon^2 - n\cdot\frac{1}{n}\sigma_\varepsilon^2 = (n-1)\sigma_\varepsilon^2$$

$$E\left[(\hat{\beta}_1 - \beta_1)^2 \sum_{i=1}^n x_i^2\right] = \sum_{i=1}^n x_i^2 \operatorname{var}(\hat{\beta}_1) = \sum x_i^2 \cdot \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n x_i^2} = \sigma_\varepsilon^2$$

$$\hat{\beta}_1 = \sum_{i=1}^n w_i \underbrace{y_i}_{=Y_i - \bar{Y}} = \sum_{i=1}^n w_i Y_i - \bar{Y}\sum_{i=1}^n w_i = \sum_{i=1}^n w_i Y_i =$$
$$= \sum_{i=1}^n w_i(\beta_0 + \beta_1 X_i + \varepsilon_i) = \beta_0 \underbrace{\sum_{i=1}^n w_i}_{=0} + \beta_1 \underbrace{\sum_{i=1}^n w_i X_i}_{=1} + \sum_{i=1}^n w_i \varepsilon_i \Rightarrow$$

$\hat{\beta}_1 - \beta_1 = \sum_{i=1}^n w_i \varepsilon_i$

$$E\left[(\hat{\beta}_1 - \beta_1)\sum_{i=1}^n x_i(\varepsilon_i - \bar{\varepsilon})\right] = E\left[\sum_{i=1}^n \underbrace{w_i}_{=\frac{x_i}{\sum x_j^2}} \varepsilon_i \cdot \left(\sum_{i=1}^n x_i \varepsilon_i - \bar{\varepsilon}\underbrace{\sum_{i=1}^n x_i}_{=0}\right)\right] =$$
$$= E\left[\frac{\left(\sum_{i=1}^n x_i \varepsilon_i\right)^2}{\sum_{j=1}^n x_j^2}\right] = \frac{1}{\sum_{j=1}^n x_j^2}\cdot\sum_{i=1}^n x_i^2 \sigma_\varepsilon^2 = \sigma_\varepsilon^2$$

$$\Rightarrow E\left(\sum_{i=1}^{n} e_i^2\right) = (n-1)\sigma_\varepsilon^2 + \sigma_\varepsilon^2 - 2\sigma_\varepsilon^2 = (n-2)\sigma_\varepsilon^2.$$

$$\Rightarrow E\left(\hat{\sigma}_\varepsilon^2\right) = E\left(\frac{RSS}{n-2}\right) = E\left(\frac{\sum_{i=1}^{n} e_i^2}{n-2}\right) = \sigma_\varepsilon^2.$$