Cochran's Theorem: ~~$Y = X_1 + X_2$~~ , ~~$X_1 \perp X_2$~~ ,

~~where $X_1 \sim \chi^2(r_1)$, $Y \sim \chi^2(r)$, then~~

decomposition of χ^2-distribution:

$$Y = X_1 + X_2$$

① $\left.\begin{array}{l} X_1 \sim \chi^2(r_1), \\ Y \sim \chi^2(r) \\ X_1 \perp X_2 \end{array}\right\} \Rightarrow X_2 \sim \chi^2(r - r_1)$

② $\left.\begin{array}{l} X_1, X_2 \sim \chi^2(r_1), \chi^2(r_2) \\ Y \sim \chi^2(r) \\ r_1 + r_2 = r \end{array}\right\} \Rightarrow X_1 \perp X_2$.

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2, \quad \frac{SST}{\sigma^2} = \frac{(n-1) S_y^2}{\sigma^2} \sim \chi^2(n-1)$$

$$SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}}_i)^2 = \sum_{i=1}^{n} (\hat{\beta}_0 + \hat{\beta}_1 x_i - (\hat{\beta}_0 + \hat{\beta}_1 \bar{x}))^2 = \hat{\beta}_1^2 \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2 = \hat{\beta}_1^2 \cdot S_{xx}$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = \sum_{i=1}^{n} c_i \cdot (y_i - \bar{y}),$$

$$= \sum_{i=1}^{n} c_i y_i \sim Normal\left(\sum_{i=1}^{n} c_i (\beta_0 + \beta_1 x_i), \sum_{i=1}^{n} c_i^2 \sigma^2\right)$$

$$\sim N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right)$$

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

$\sum_{i=1}^{n} c_i = 0$

$\sum_{i=1}^{n} c_i^2 = \frac{1}{S_{xx}}$

$\sum_{i=1}^{n} c_i x_i = 1$

$$\frac{SSR}{\sigma^2} = \frac{\hat{\beta}_1^2 \cdot S_{xx}}{\sigma^2} \overset{H_0: \beta_1 = 0}{\sim} \chi^2(1)$$

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \quad \frac{SSE}{\sigma^2} \sim \chi^2(r) \quad \sim \chi^2(n-2)$$

$$= \sum_{i=1}^{n} e_i^2 \qquad \left\{\begin{array}{l} \sum_i e_i = 0 \\ \sum_i e_i x_i = 0 \end{array}\right. \qquad e \perp \mathcal{L}\{1, x\}$$

$(\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i)$ $\qquad SSE \perp SSR$

① $\hat{\beta}_1 \sim N(\beta_1, \frac{\sigma^2}{S_{xx}})$ $\quad S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$

$$E(\hat{\beta}_1) = \beta_1 \qquad Var(\hat{\beta}_1) = Var(\sum_{i=1}^{n} c_i y_i) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}, \qquad c_i = \frac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$SE(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\frac{MSE}{S_{xx}}}$$

② Sampling Distribution:

$$\frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1)} = \frac{\frac{\hat{\beta}_1 - \beta_1}{\sqrt{Var(\hat{\beta}_1)}}}{\sqrt{\frac{SE^2(\hat{\beta}_1)}{Var(\hat{\beta}_1)}}} = \frac{Z \sim N(0,1)}{\sqrt{\frac{MSE}{S_{xx}} \Big/ \frac{\sigma^2}{S_{xx}}}}$$

$$= \frac{Z}{\sqrt{\frac{MSE}{\sigma^2}}} = \frac{Z}{\sqrt{\frac{SSE}{\sigma^2} / (n-2)}} \sim t(n-2)$$

Chochran's Theorem: $SST = SSE + SSR$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{SST}{\sigma^2} \sim \chi^2(n-1) \qquad \frac{SSR}{\sigma^2} = \hat{\beta}_1^2 \cdot \sum_{i=1}^{n}(x_i-\bar{x})^2 \Big/ \sigma^2 \sim \chi^2(1)$$

and $SSE \perp SSR$ $\quad (e \perp L\{1, x\})$

$$\Rightarrow \frac{SSE}{\sigma^2} \sim \chi^2(n-2) \qquad \Rightarrow \begin{cases} SSE \perp \hat{\beta}_1^2 \\ \therefore SSE \perp Z \end{cases}$$

1. Confidence Interval of $\underbrace{\text{Response}}_{}$ (Mean): $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0,1)$

L.S. Estimation: $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \quad \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

Fitted value: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i, \quad i = 1 \cdots n$

~~C.I. for~~ mean response $\underbrace{\beta_0 + \beta_1 x_i}_{E(Y_i) =}$, $i = 1 \cdots n$

Note that $E(\hat{Y}_i) = E(\hat{\beta}_0 + \hat{\beta}_1 x_i) = \beta_0 + \beta_1 x_i$, unbiased estimator of $\beta_0 + \beta_1 x_i$.

$$\begin{aligned}
Var(\hat{Y}_i) &= Var(\hat{\beta}_0 + \hat{\beta}_1 x_i) \\
&= Var(\hat{\beta}_0) + Var(\hat{\beta}_1) \cdot x_i^2 + 2x_i \cdot Cov(\hat{\beta}_0, \hat{\beta}_1) \\
&= \frac{\sigma^2 \sum_i x_i^2}{n \cdot S_{xx}} + \frac{\sigma^2 \cdot x_i^2}{S_{xx}} + 2x_i \cdot \left[- \frac{\sigma^2 \sum_i x_i}{n \cdot S_{xx}} \right]
\end{aligned}$$

where
$$\begin{aligned}
Cov(\hat{\beta}_0, \hat{\beta}_1) &= E(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1) \\
&= E\left[\sum_{i=1}^n k_i (Y_i - EY_i) \cdot \sum_{j=1}^n c_j (Y_j - EY_j) \right] \\
&= E\left[\sum_{i=1}^n k_i \varepsilon_i \cdot \sum_{j=1}^n c_j \varepsilon_j \right] = \sum_{i=1}^n k_i c_i \sigma^2 + \sum_{i \neq j} k_i c_j \underbrace{E(\varepsilon_i) \cdot E(\varepsilon_j)}_{=0} \\
&= \sum_{i=1}^n k_i c_i \sigma^2 = \sum_{i=1}^n \left(\frac{1}{n} - \bar{x} c_i\right) \cdot \sigma^2 \cdot c_i \\
&= -\bar{x} \cdot \frac{\sigma^2}{S_{xx}} = -\bar{x} \cdot Var(\hat{\beta}_1)
\end{aligned}$$

$\hat{\beta}_0 = \sum_i k_i Y_i \quad k_i = \frac{1}{n} - \bar{x} c_i$

$\hat{\beta}_1 = \sum_i c_i Y_i, \quad c_i = \frac{x_i - \bar{x}}{S_{xx}}$

$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

$\sum_i c_i = 0 \quad \sum_i c_i^2 = \frac{1}{S_{xx}}$

or
$$\begin{aligned}
Cov(\hat{\beta}_0, \hat{\beta}_1) &= Cov(\bar{Y} - \hat{\beta}_1 \bar{x}, \hat{\beta}_1) = -\bar{x} Var(\hat{\beta}_1) \\
&= Cov(\bar{Y}, \hat{\beta}_1) - \bar{x} \cdot Var(\hat{\beta}_1) \\
&= Cov(\bar{Y}, \sum_{i=1}^n c_i Y_i) - \bar{x} Var(\hat{\beta}_1)
\end{aligned}$$

~~$Cov(\bar{Y}, \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}))$~~

$$\begin{aligned}
&= Cov\left(\frac{1}{n} \sum_{i=1}^n Y_i, \sum_{i=1}^n c_i Y_i\right) \\
&= \underbrace{\sum_{i=1}^n \frac{1}{n} c_i \cdot Var(Y_i)}_{=0} + \sum_{i \neq j} \frac{c_j}{n} \underbrace{Cov(Y_i, Y_j)}_{=0}
\end{aligned}$$

$Y_i \perp Y_j$

$\sum_{i=1}^n c_i = 0$

$$\therefore Var(\hat{Y}_i) = \frac{\sigma^2}{nS_{xx}}\left[\sum_{i=1}^{n} x_i^2 + n x_i - 2x_i \cdot \sum_{i=1}^{n} x_i\right]$$

$$= \frac{\sigma^2}{nS_{xx}}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(x_i-\bar{x})^2\right] = \sigma^2\left\{\frac{1}{n} + \frac{(x_i-\bar{x})^2}{S_{xx}}\right\}$$

Note that when $x_i = \bar{x}$, $Var(\hat{Y}_i) = \frac{\sigma^2}{n}$.

$$\Rightarrow \hat{Y}_i \sim N\left(\beta_0+\beta_1 x_i, \sigma^2\left(\frac{1}{n}+\frac{(x_i-\bar{x})^2}{S_{xx}}\right)\right)$$

$$Y_i \sim N(\beta_0+\beta_1 x_i, \sigma^2)$$

Standard error of $\hat{Y}_i$: $Se(\hat{Y}_i)$

$$Se^2(\hat{Y}_i) = MSE\left\{\frac{1}{n} + \frac{(x_i-\bar{x})^2}{S_{xx}}\right\}, \quad MSE = \frac{SSE}{n-2}$$

Sampling Distribution:

$$\frac{\left[\hat{Y}_i - (\beta_0+\beta_1 x_i)\right] / \sqrt{Var(\hat{Y}_i)}}{\sqrt{\frac{SSE}{\sigma^2} / (n-2)}} \sim t(n-2)$$

$100(1-\alpha)\%$ C.I. for $(\beta_0+\beta_1 x_i)$ is

$$\hat{Y}_i \pm t_{\frac{\alpha}{2}}(n-2) \cdot Se(\hat{Y}_i)$$

2. Prediction Interval of Y_k at x_k (time-related data)

(x_k, Y_k) is independant of $\{(x_1, y_1), \cdots, (x_n, y_n)\}$

Predicted value $\tilde{Y}_k = \hat{\beta}_0 + \hat{\beta}_1 \cdot x_k$, where $\hat{\beta}_0, \hat{\beta}_1$ relies on $\{(x_1, y_1), \cdots, (x_n, y_n)\}$.

$\therefore \; Y_k \perp\!\!\!\perp \tilde{Y}_k = \hat{\beta}_0 + \hat{\beta}_1 x_k$

Prediction error $\quad w = Y_k - \tilde{Y}_k$

$$
\begin{aligned}
Var(w) &= Var(Y_k - \tilde{Y}_k) \\
&= Var(Y_k) + Var(\tilde{Y}_k) \qquad Y_k \perp\!\!\!\perp \tilde{Y}_k \\
&= \sigma^2 + \sigma^2 \left\{ \frac{1}{n} + \frac{(x_k - \bar{x})^2}{S_{xx}} \right\}, \qquad S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 \\
&= \sigma^2 \left\{ 1 + \frac{1}{n} + \frac{(x_k - \bar{x})^2}{S_{xx}} \right\}
\end{aligned}
$$

Standard prediction error at x_k: $S_{pred}(x_k)$

$$
S^2_{pred}(x_k) = MSE \left\{ 1 + \frac{1}{n} + \frac{(x_k - \bar{x})^2}{S_{xx}} \right\}
$$

$100(1-\alpha)\%$ prediction interval of Y_k is

$$
(\hat{\beta}_0 + \hat{\beta}_1 x_k) \pm t_{\frac{\alpha}{2}}(n-2) \cdot S_{pred}(x_k)
$$