24/10/16 (1)

ΕΠΑΝΑΛΗΨΗ

ΑΓΠ $Y_i = \beta_0 + \beta_1 x_i + e$, $i=1,\dots,n$

- $\hat{\beta}_1 = \dfrac{\sum (x_i-\bar{x})(Y_i-\bar{Y})}{\sum (x_i-\bar{x})^2}$

- $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}$

- $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad i=1,\dots,n \qquad e_i = Y_i - \hat{Y}_i \quad i=1,\dots,n$

όσο μικρότερο το e τόσο καλύτερο μοντέλο

ANOVA $SS_{tot} = SS_{reg} + SS_{res}$

1. $SS_{res} << SS_{reg}$ καταλληλότητα μοντέλου

2. $R^2 = \dfrac{SS_{reg}}{SS_{tot}} = 1 - \dfrac{SS_{res}}{SS_{tot}}$ με $0 \le R^2 \le 1$ ↓ κατάλληλο μοντέλο

του μοντέλου

• ΥΠΟΘΕΣΕΙΣ ΣΦΑΛΜΑΤΩΝ

① $E(\varepsilon_i)=0 \quad i=1,\ldots,n$

② $Var(\varepsilon_i)=\sigma^2 \,,\ i=1,\ldots,n$, αυτό σημαίνει στον διακύμανση τους δεν εξαρτάται από τα x_i

③ Τα ε_i ασυσχέτιστα ανά δύο

$Cov(\varepsilon_i,\varepsilon_j)=0 \quad \forall i,j=1,\ldots,n \,,\ i\neq j$

④ $\varepsilon_i \sim N(0,\sigma^2)$

από τις υποθέσεις αυτές υπάρχουν κάποιες συνέπειες

• ΣΥΝΕΠΕΙΕΣ στ Y_i

$(Y_i=B_0+B_1X_i+\varepsilon_i)$

① $E(Y_i)=B_0+B_1X_i \,,\ i=1,\ldots,n$

② $Var(Y_i)=\sigma^2 \,,\ i=1,\ldots,n$

③ $Cov(Y_i,Y_j)=0$

④ $Y_i \sim N(B_0+B_1X_i,\sigma^2) \,,\ i=1,\ldots,n$

Θεώρημα: Αν οι υποθέσεις για τα σφάλματα ικανοποιούνται τότε

α. $\hat{B}_1 \sim N\left(B_1, \frac{\sigma^2}{\Sigma(x_i-\bar{x})^2}\right)$

β. $\hat{B}_0 \sim N\left(B_0, \frac{\Sigma x_i^2 \cdot \sigma^2}{n\Sigma(x_i-\bar{x})^2} = \sigma^2\left\{\frac{1}{n}+\frac{\bar{x}^2}{\Sigma(x_i-\bar{x})^2}\right\}\right)$

Παρατήρηση: οι εκτιμητές $\hat{B}_0$, $\hat{B}_1$ είναι αμερόληπτοι

<u>Αποδειξη</u>

α. $\hat{B}_1 = \dfrac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\Sigma(x_i-\bar{x})^2} = \Sigma \dfrac{x_i-\bar{x}}{\Sigma(x_i-\bar{x})^2} y_i$

το $\hat{B}_1$ γραμμικος συνδυασμό των ανεξαρτητων Y_i

ανεξαρτητων διοτί $Y_i \sim N$ και είναι ασυσχετιστα

Άρα $\hat{B}_1 \sim$ Normal

οπότε αρκεί να αποδείξουμε πόσο είναι η μ.τ και η διακύμανση

- $E(\hat{B}_1) = E\left[\Sigma \dfrac{(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2} y_i\right] = \Sigma \dfrac{x_i-\bar{x}}{\Sigma(x_i-\bar{x})^2} E(y_i)$

$$= \Sigma \frac{(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2} E(B_0 + B_1 x_i + \varepsilon_i)$$

$$= \Sigma \frac{(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2} (B_0 + B_1 x_i + \cancel{E(\varepsilon_i)}^{0})$$

$$= B_0 \cancel{\Sigma \frac{(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2}}^{0} + B_1 \Sigma \frac{(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2} \cdot x_i$$

$$= B_1 \frac{\Sigma(x_i^2 - \bar{x}x_i)}{\Sigma(x_i^2 - 2\bar{x}x_i + \bar{x}^2)}$$

$$= B_1 \frac{\Sigma x_i^2 - \bar{x}\Sigma x_i}{\Sigma x_i^2 - 2\bar{x}\Sigma x_i + n\bar{x}^2}$$

$$= B_1 \frac{\Sigma x_i^2 - n\bar{x}^2}{\Sigma x_i^2 - 2n\bar{x}^2 + n\bar{x}^2} = B_1 \frac{\Sigma x_i^2 - n\bar{x}^2}{\Sigma x_i^2 - n\bar{x}^2} = B_1$$

Οπότε:

Αν $\omega_1, \dots, \omega_n$ τ.μ τότε

$$Var\left(\sum_{i=1}^{n} a_i \omega_i\right) = \sum_{i=1}^{n} a_i^2 Var(\omega_i) + \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(\omega_i, \omega_j)$$

Αν όμως $\omega_1, \dots, \omega_n$ ασυσχετ. τότε

$$Var\left(\sum_{i=1}^{n} a_i \omega_i\right) = \sum_{i=1}^{n} a_i^2 Var(\omega_i)$$

οπότε χρησιμοποιώντας αυτό τον τύπο βρίσκω την διακύμανση

$$Var(\hat{\beta}_1) = Var\left(\sum \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} Y_i\right)$$

$$= \sum \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} Var(Y_i)$$

$$= \sigma^2 \sum \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = \sigma^2 \frac{\sum (x_i - \bar{x})}{[\sum (x_i - \bar{x})^2]^2}$$

$$= \frac{\sigma^2}{\sum (x_i - \bar{x})^2}$$

Παρατ: Το σ^2 άγνωστο. Άρα για να αξιοποιηθούν οι κατανομές αυτές το σ^2 πρέπει να εκτιμηθεί

Θεώρημα: Από τις υποθέσεις για τα σφάλματα

$$E(MS_{res}) = \sigma^2$$

Παρατήρηση: Το $\hat{\sigma}^2 = MS_{res}$ είναι αμερόληπτος εκτιμητής της σ^2

Θεώρημα: Υπο τις υποθέσεις στα κατάλοιπα

$$\frac{SS_{res}}{\sigma^2} \sim \chi^2_{n-2}$$

Τα $Y_i \sim N(B_0 + B_1 x_i, \sigma^2)$, $i = 1, \dots, n$

$$\Rightarrow \frac{Y_i - (B_0 + B_1 x_i)}{\sigma} \sim N(0,1)$$

$$\Rightarrow \frac{[Y_i - (B_0 + B_1 x_i)]^2}{\sigma^2} \sim \chi_1^2$$

$$\Rightarrow \sum_{i=1}^{n} \frac{[Y_i - (B_0 + B_1 x_i)]^2}{\sigma^2} \sim \sum_{i=1}^{n} \chi_1^2 = \chi^2_{\sum_{i=1}^{n} 1} \equiv \chi_1^2$$

Για τα υπόλοιπα έχω

$$SS_{res} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y})^2 = \sum_{i=1}^{n} (Y_i - (\hat{B}_0 + \hat{B}_1 x_i))^2$$

$$\Rightarrow \sum_{i=1}^{n} \frac{[Y_i - (\hat{B}_0 + \hat{B}_1 x_i)]^2}{\sigma^2} \sim \chi^2_{n-2}$$

Για να υπολογίσω αυτό το άθροισμα έχω n-πληροφορίες όμως οι σχέσεις μεταξύ των $\hat{B}_0, \hat{B}_1$ μου τις περιορίζουν σε n-2

$$\Rightarrow \frac{SS_{res}}{\sigma^2} \sim \chi^2_{n-2}$$

Θεώρημα: Τα $\hat{B}_0, \hat{B}_1$ είναι ανεξάρτητοι του MSres και άρα ανεξάρτητοι του SSres

Αναγκαιότητα κατασκευής test για B_0, B_1

Έχει νόημα: $H_0\ B_1 = B_1^*$ (B_1^* γνωστό)

$H_0\ B_1 = 0$

Ναί διότι $H_0\ B_1 = B_1^*$ ελέγχει αν η μεταβολή της Y για μοναδιαία μεταβολή της X είναι ίση με B_1^*

ή αν $H_0\ B_1 = 0$ τότε

δεν υπάρχει σχέση μεταξύ X, Y

Τεστ για έλεγχο $H_0: B_1 = B_1^*$ (B_1^* γνωστό) έναντι $H_a: B_1 \neq B_1^*$

Βασική ιδέα
- Τεστ Wald

Έστω τδ ω_1, ω_n από $N(\mu, \sigma^2)$

$H_0: \mu = \mu_0$ (μ_0 = γνωστό)

σ^2 γνωστό → z-test

$$z = \frac{\bar{\omega} - \mu_0}{\sigma/\sqrt{n}}$$

σ^2 άγνωστο → t-test

$$t = \frac{\bar{\omega} - \mu}{S/\sqrt{n}}, \quad S = \frac{1}{n-1}\Sigma(\omega_i - \bar{\omega})^2$$

Άρα για μια στατιστική συνάρτηση έχω

$$\Sigma\Sigma T = \frac{\text{εκτίμησης} - E(\text{εκτιμητή})}{\text{τυπική απόκλιση(εκτιμητή)}}$$

1. Κατανομή της ΣΣΤ υπό H_0
2. κρίσιμο σημείο
   $a = P(\text{απορ } H_0 / H_0 \text{ αληθ})$

του $\hat{B}_1 \sim N(B_1, \frac{\sigma^2}{\Sigma(x_i-\bar{x})^2})$

Υπό την συνθήκη $H_0: B_1 = B_1^*$

$$\hat{B}_1 \sim N(B_1^*, \frac{\sigma^2}{\Sigma(x_i-\bar{x})^2})$$

$$\frac{\hat{B}_1 - B_1^*}{\sqrt{\frac{\sigma^2}{\Sigma(x_i-\bar{x})^2}}} \sim N(0,1) \text{ υπο την } H_0: B_1 = B_1^*$$

όμως δεν γνωρίζω σ^2, επομένως το λογικό είναι να το στρέψω σε t-κατανομή. Ξέρω ότι

$t_v = \frac{N(0,1)}{\sqrt{\chi_n^2/n}}$. Άρα πρέπει να ~~...~~ διαιρέσω με την κατάλληλη ποσότητα ώστε το σ^2 να εξαφανιστεί και να οδηγηθώ στην t-κατανομή. Έχω λοιπόν την ποσότητα $\frac{SS_{res}}{\sigma^2} \sim \chi^2_{n-2}$

$$\frac{\frac{\hat{B}_1 - B_1^*}{\sqrt{\frac{\sigma^2}{\Sigma(x_i-\bar{x})^2}}}}{\sqrt{\frac{SS_{res}}{\sigma^2}/(n-2)}} = \frac{\hat{B}_1 - B_1^*}{1/\sqrt{\Sigma(x_i-\bar{x})^2}\sqrt{\frac{SS_{res}}{n-2}}} = \frac{\hat{B}_1 - B_1^*}{1/\sqrt{\Sigma(x_i-\bar{x})^2}\sqrt{MS_{res}}}$$

$$\Rightarrow t = \frac{\hat{B}_1 - B_1^*}{\sqrt{\frac{MS_{res}}{\Sigma(x_i-\bar{x})^2}}} \equiv \frac{N(0,1)}{\sqrt{\chi^2_{n-2}/(n-2)}} \sim t_{n-2}$$

και εφόσον έχω ανεξαρτησία άρα έχω t-κατανομή!

Υπο την $H_0: B_1 = B_1^*$ η

$$t = \frac{\hat{B_1} - B_1^*}{\sqrt{\frac{USres}{\Sigma(x_i-\bar{x})^2}}} \sim t_{n-2}$$

Οπότε αν το t μεγάλες τιμές τότε

$\hat{B_1}$ πολύ διαφορετικό απο το B_1^* δηλαδή και

τότε B_1 πολύ διαφορετικό απο το B_1^* άρα

Θα πρέπει να απορρίψω το H_0

Τελικά μεγάλες τιμές του t οδηγούν σε απορ H_0

Δηλαδή η κρίσιμη περιοχή αποτελείται από μεγάλες τιμές

του t, δηλ $|t| \geq c$

▶ Μορφή κ.π $|t| \geq c$, αρκεί να προσδιορίσω το κρίσιμο σημείο c

το οποίο θα το κάνω $P(\text{απο}H_0 \mid H_0 \text{ αληθ}) = \alpha$

▶ Παρατηρώντας ότι $\sqrt{\frac{USres}{\Sigma(x_i-\bar{x})^2}} = \sqrt{\frac{\hat{\sigma}^2}{\Sigma(x_i-\bar{x})^2}} = \sqrt{\hat{Var}(\hat{B_1})}$

Δένας εκτιμητής $\hat{B_1}$

Οπότε Για τον έλεγχο της $H_0: B_1 = B_1^*$ έναντι $H_a: B_1 \neq B_1^*$

η ΣΣΤ είναι $t = \frac{\hat{B_1} - B_1^*}{\sqrt{\hat{Var}(\hat{B_1})}}$ με κατανομή t_{n-2} υπο H_0

και κ.π μεγέθους α, την $|t| \geq t_{n-2, \frac{\alpha}{2}}$

όπου $\hat{Var}(\hat{B_1}) = \frac{USres}{\Sigma(x_i-\bar{x})^2}$

Επειδή $t = \frac{\hat{B}_1 - B_1}{\sqrt{\hat{var}(\hat{B}_1)}} \sim t_{n-2}$ είναι αναστρέψιμη

δηλαδή $\exists q_1, q_2$ τ.ω

$$1-a = P(q_1 < t < q_2) = P\left(q_1 < \frac{\hat{B}_1 - B_1}{\sqrt{\hat{var}(\hat{B}_1)}} < q_2\right)$$

$$= P\left(\hat{B}_1 - q_2\sqrt{\hat{var}(\hat{B}_1)} \leq B_1 \leq \hat{B}_1 - q_1\sqrt{\hat{var}(\hat{B}_1)}\right)$$

Άρα $100(1-a)\%$ δ.ε για την B_1 είναι

$$\left(\hat{B}_1 - q_2\sqrt{\hat{var}(\hat{B}_1)}, \ \hat{B}_1 - q_1\sqrt{\hat{var}(\hat{B}_1)}\right)$$

Το δ.ε ίσων ουρών προκύπτει για $q_2 = -q_1$ και

$q_1 = -t_{n-2, a/2}$

$q_2 = t_{n-2, a/2}$

Οπότε το $100(1-a)\%$ δ.ε για B_1 είναι

$$\left(\hat{B}_1 - t_{n-2, \frac{a}{2}}\sqrt{\hat{var}(\hat{B}_1)}, \ \hat{B}_1 + t_{n-2, a/2}\sqrt{\hat{var}(\hat{B}_1)}\right)$$

Ανάλογα. Για τον έλεγχο $H_0: B_0 = B_0^*$ (B_0^* γνωστό) έναντι $H_a: B_0 \neq B_0^*$

η ΣΣΤ είναι $t = \frac{\hat{B}_0 - B_0^*}{\sqrt{\hat{var}(\hat{B}_0)}}$ με κατανομή t_{n-2} υπό H_0

και κ.π μεγέθους a την $|t| \geq t_{n-2, a/2}$

όπου $\hat{var}(\hat{B}_0) = \sqrt{\frac{\Sigma (x_i)^2}{n \Sigma (x_i - \bar{x})^2} MS_{res}}$ και $100(1-a)\%$ δ.ε

είναι το $\left(\hat{B}_0 - t_{n-2, \frac{a}{2}}\sqrt{\hat{var}(\hat{B}_0)}, \ \hat{B}_0 + t_{n-2, a/2}\sqrt{\hat{var}(\hat{B}_0)}\right)$

Το μοντέλο της α.δ.π χρησιμοποιείται για προβλέψεις.

Ειδικότερα το εκτιμώμενο μοντέλο χρησιμοποιείται

για προβλέψεις

Έτσι η πρόβλεψη της Y όταν $x = x_k$

είναι $\hat{Y}_k = \hat{B}_0 + \hat{B}_1 x_k$

Όμως το θέμα είναι πόσο καλή είναι η πρόβλεψη $\hat{Y}_k$;

Η πρόβλεψη $\hat{Y}_k$ είναι καλή όταν $Var(\hat{Y}_k)$ είναι μικρή.

▶ Υπολογισμός $Var(\hat{Y}_k)$

$$\hat{Y}_k = \hat{B}_0 + \hat{B}_1 x_k = \bar{Y} - \hat{B}_1 \bar{x} + \hat{B}_1 x_k$$

$$= \bar{Y} - \hat{B}_1 (\bar{x} - x_k)$$

Επομένως λοιπόν

$$Var(\hat{Y}_k) = Var(\bar{Y} - \hat{B}(\bar{x} - x_k))$$

για να το υπολογίσω θυμίζω ότι:

$$Var\left(\sum_{i=1}^{n} a_i w_i\right) = \sum a_i^2 Var(w_i) + \sum_i \sum_j a_i a_j Cov(w_i, w_j)$$

για $i, j = 1, \ldots, n$

$i \neq j$

οπότε

$$Var(\hat{Y}_k) = Var(\bar{Y}) + (\bar{x} - x_k)^2 Var(\hat{B}_1) - (\bar{x} - x_k) Cov(\bar{Y}, \hat{B}_1) \quad \underline{(1)}$$

Τώρα θα η συνδιακύμανση είναι:

$$Cov(\bar{Y}, \hat{B}_1) = Cov\left(\frac{1}{n}\sum Y_i, \sum \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} Y_i\right) = Cov\left(\sum \frac{1}{n} Y_i, \sum \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2} Y_i\right)$$

(2)

Άρα τελικά

από (2),(3) η συνδιακύμανση είναι

$$Cov(\bar{Y},\hat{B}_1) = Cov\left(\sum_{i=1}^{n} \frac{1}{n} Y_i , \sum_{i=1}^{n} \frac{(x_i-\bar{x})}{\Sigma(x_i-\bar{x})^2} Y_i\right)$$

$$= \sum_{i=1}^{n} \frac{1}{n} \frac{x_i-\bar{x}}{\Sigma(x_i-\bar{x})^2} Var(Y_i)$$

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

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

$$= \frac{\sigma^2}{n} \frac{\cancelto{=0}{\Sigma(x_i-\bar{x})}}{(\Sigma(x_i-\bar{x}))^2} \quad (4)$$

Δηλαδή από (1),(4)

$$\boxed{Var(\hat{Y}_k) = Var(\bar{Y}) + (\bar{x} - x_k)^2 Var(\hat{B}_1)}$$

Άρα τι πρέπει να συμβαίνει όταν να είναι η $Var(\hat{Y}_k)$ όσο πιο μικρή γίνεται

αρκεί λοιπόν $(\bar{x} - x_k)$ να είναι όσο μικρό γίνεται

δηλαδή αρκεί το x_k να είναι κοντά στο $\bar{x}$!

Λεπτομερής απόδειξη:

$$E(MS_{res}) = \sigma^2$$

Είναι $MS_{res} = \dfrac{SS_{res}}{n-2} = \dfrac{1}{n-2}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$

Ισχύει $SS_{res} = SS_{tot} - SS_{reg}$

$$E(SS_{tot}) = E\left\{\sum_{i=1}^{n}(y_i-\bar{y})^2\right\} = \sum_{i=1}^{n}E(y_i-\bar{y})^2 = \sum_{i=1}^{n}E\Big\{b_0+b_1x_i+\varepsilon_i -$$
$$-\frac{1}{n}\sum_{i=1}^{n}(b_0+b_1x_i+\varepsilon_i)\Big\}^2 = \sum_{i=1}^{n}E\{\cancel{b_0}+b_1x_i+\varepsilon_i-\cancel{b_0}-b_1\bar{x}-\bar{\varepsilon}\}^2.$$

Μετά τις πράξεις και λαμβάνοντας υπόψη ότι:

$$E(\varepsilon_i-\bar{\varepsilon}) = E(\varepsilon_i) - \frac{1}{n}\sum_{i=1}^{n}E(\varepsilon_i) = 0-0=0$$

παίρνουμε:

$$E(SS_{tot}) = b_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{i=1}^{n}E(\varepsilon_i-\bar{\varepsilon})^2. \qquad (1)$$

Αλλά:

$$\sum_{i=1}^{n}E(\varepsilon_i-\bar{\varepsilon})^2 = \sum_{i=1}^{n}\left(E(\varepsilon_i^2)-2E(\varepsilon_i\bar{\varepsilon})+E(\bar{\varepsilon}^2)\right) = \sum_{i=1}^{n}E(\varepsilon_i^2)-2E\Big(\sum_{i=1}^{n}\varepsilon_i\Big)\bar{\varepsilon}+$$
$$+nE(\bar{\varepsilon}^2) = \sum_{i=1}^{n}E(\varepsilon_i^2)-2E(n\bar{\varepsilon}\cdot\bar{\varepsilon})+nE(\bar{\varepsilon}^2) = \sum_{i=1}^{n}E(\varepsilon_i^2)-nE(\bar{\varepsilon}^2)$$
$$= \sum_{i=1}^{n}\left[Var(\varepsilon_i)+[\cancelto{0}{E(\varepsilon_i)}]^2\right]-n\left[Var(\bar{\varepsilon})+[\cancelto{0}{E(\bar{\varepsilon})}]^2\right]$$
$$= \sum_{i=1}^{n}\sigma^2 - nVar\Big(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\Big) = n\sigma^2 - n\frac{1}{n^2}\sum_{i=1}^{n}Var(\varepsilon_i) = n\sigma^2-\frac{1}{n}\sum_{i=1}^{n}\sigma^2$$

και τελικά: $\sum_{i=1}^{n}E(\varepsilon_i-\bar{\varepsilon})^2 = (n-1)\sigma^2$ (2)

Από τις (1) και (2): (3)

$$E(SS_{tot}) = b_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2 + (n-1)\sigma^2$$

Επίσης $SS_{reg} = \sum_{i=1}^{n}(y_i-\hat{y}_i)^2 = \hat{b}_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2$

και $E(SS_{reg}) = \left\{\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}E(\hat{b}_1^2)$

ή

$$E(SS_{reg}) = \left\{\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}\left[Var(\hat{b}_1) + [E(\hat{b}_1)]^2\right].$$

Λαμβάνοντας υπόψη ότι

$$Var(\hat{b}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad \text{και} \quad E(\hat{b}_1) = b_1 \,,$$

μετά από πράξεις προκύπτει:

$$E(SS_{reg}) = \sigma^2 + b_1^2 \sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad (4)$$

Έτσι από τις (3) και (4) έχουμε:

$$E(MS_{res}) = \frac{1}{n-2}E(SS_{res}) = \frac{1}{n-2}\left[E(SS_{tot}) - E(SS_{reg})\right]$$

$$= \frac{1}{n-2}\left\{b_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2 + (n-1)\sigma^2 - \sigma^2 - b_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}$$

$$= \frac{1}{n-2}(n-2)\sigma^2$$

$$= \sigma^2$$

▲