ΣΤΑΤΙΣΤΙΚΗ

Διάλεξη 20η

23/05/2018

(*) Το 2ο κεφάλαιο δεν θα εξεταστεί στο τέλος. (περιγραφική στατιστική)

Δειγματικές μονάδες	Δοκιμασίες 1	2	3	4	y_{ij}
1	12	14	19	24	
2	18	12	17	30	
3		13	21		
Σύνολο δειγματικών	30	35	57	54	$180 = y_{00}$
Μέσοι	15	13	19	27	$18 = \bar{y}_{00}$
Μέγεθος δείγματος	2	3	3	2	$10 = n$

$\boxed{y_{ij} = \mu_j + \varepsilon_{ij}}$, $i = 1, \dots, n$ και $j = 1, \dots, r$, $n = \sum_{j=1}^{r} n_j$

$\varepsilon_{ij} \sim N(0, \sigma^2)$, $y_{ij} \sim N(\mu_j, \sigma^2)$

$\boxed{\varepsilon_{ij} = y_{ij} - \mu_j}$, $\boxed{\hat{\mu}_j = \bar{y}_{0j},\ j = 1, \dots, r}$

$\boxed{y_{0j} = \sum_{i=1}^{n_j} y_i}$, $\boxed{\bar{y}_{0j} = y_{0j} / n_j}$

$\boxed{y_{00} = \sum_j \sum_i y_{ij}}$, $\boxed{\bar{y}_{00} = \frac{y_{00}}{n}}$

$H_0: \mu_1 = \dots = \mu_r = \mu$

H_a: όχι όλα τα μέσα υποβλητών

$\boxed{e_{ij} = y_{ij} - \hat{\mu}_j = y_{ij} - \bar{y}_{0j}}$ || $\sum_{i=1}^{n_j} e_{ij} = 0$, $j = 1, \dots, r$

ΣΘΕΛ 209 (Βιβλίο)

Προέλευση μεταβλητότητας	Αθροίσματα τετραγώνων (SS)	Βαθμοί ελευθερ.	Μέσα τετράγωνα (MS)
Διαφορές μεταξύ των δοκιμ.	258 $SS_{tre}=\sum_{j=1}^{r} n_j(\bar{Y}_{\circ j}-\bar{Y}_{\circ\circ})^2$	3 $r-1$	86 $MS_{tre}=\dfrac{SS_{tre}}{r-1}$
Υπόλοιπο ή εντός των δοκιμ.	46 $SS_{res}=\sum_{j=1}^{r}\sum_{i=1}^{n_j}(Y_{ij}-\bar{Y}_{\circ j})^2$	6 $n-r$	7.67 $MS_{res}=\dfrac{SS_{res}}{n-r}$
Ολική μεταβλητότητα	$SS_{tot}=\sum_{j=1}^{r}\sum_{i=1}^{n_j}(Y_{ij}-\bar{Y}_{\circ\circ})^2$ 304	9 $n-1$.	

(*) αποτελέσματα του παραδείγματος σελ 211.
(παράδ. 6.4. - δεδομένα ΑΒΓΔ)

~o ~o ~o ~o

$\hat{\sigma}^2 = MS_{res}$ και $E(MS_{res}) = \dfrac{1}{n-r}\sum_{j=1}^{r}(n_j-1)E(S_j^2)=\sigma^2$.

πρέπει αμερόληπτο

$$E(MS_{res}) = E\left[\frac{1}{n-r}SS_{res}\right] = E\left[\frac{1}{n-r}\sum_{j=1}^{r}\sum_{i=1}^{n_j}(Y_{ij}-\bar{Y}_{\circ j})^2\right]$$

$$= E\left[\frac{1}{n-r}\sum_{j=1}^{r}(n_j-1)\sum_{i=1}^{n_j}(Y_{ij}-\bar{Y}_{\circ j})^2\right]$$

$$= E\left[\frac{1}{n-r}\sum_{j=1}^{r}(n_j-1)S_j^2\right]$$

$$= \frac{1}{n-r}\sum_{j=1}^{r}(n_j-1)E(S_j^2) =$$

$$= \frac{1}{n-r}\sum_{j=1}^{r}(n_j-1)\sigma^2 = \sigma^2\cdot\frac{n-r}{n-r} = \sigma^2 \quad //$$

$\hat{\sigma}^2 = MS_{res}$ και $E(MS_{res}) = \sigma^2$ (παντα αμερόληπτος)

$$E(MS_{tre}) = \sigma^2 + \frac{1}{r-1}\sum_{j=1}^{r} n_j(\mu_j-\mu)^2, \quad \mu = \sum_{j=1}^{r} n_j \frac{\mu_j}{n} \quad (\text{Ασκ. } 6.1)$$

$$F = \frac{MS_{tre}}{MS_{res}} \left(= \frac{\frac{SS_{tre}}{\sigma^2} \Big| (r-1)}{\frac{SS_{res}}{\sigma^2} \Big| (n-r)} \equiv \frac{X^2_{r-1}|(r-1)}{X^2_{n-r}|(n-r)} \right) \sim F_{r-1,n-r}$$

$\frac{SS_{tre}}{\sigma^2} \sim X^2_{r-1}$, οταν H_0: αληθής

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

κρ. περιοχη: $F \geq F_{a,r-1,n-r}$

ΑΒΓΕΖ: $F = \frac{86}{7.67} = 11.2 > 4.76$

αρα απορ. η H_0

Σχετικά με μ_j: $\bar{Y}_{0j} \sim N(\mu_j, \sigma^2/n_j) \rightarrow \frac{\bar{Y}_{0j} - \mu_j}{\sqrt{MS_{res}/n_j}} \sim t_{n-r}$

$H_0: \mu_\kappa - \mu_\nu = 0$ v $H_a: \mu_\kappa - \mu_\nu \neq 0$

$$\frac{\bar{Y}_{0\kappa} - \bar{Y}_{0\nu} - (\mu_\kappa - \mu_\nu)}{\sqrt{MS_{res}}\sqrt{\frac{1}{n_\kappa} + \frac{1}{n_\nu}}} \sim t_{n-r} \quad \text{και} \quad t = \frac{\bar{Y}_{0\kappa} - \bar{Y}_{0\nu} - 0}{\sqrt{MS_{res}}\sqrt{\frac{1}{n_\kappa} + \frac{1}{n_\nu}}}$$

($\sim t_{n-r}$, οταν H_0 αληθής)

$H_0: \mu_3 - \mu_4 = 0$ v $H_a: \mu_3 - \mu_4 \neq 0$

$$t = \frac{\bar{Y}_{03} - \bar{Y}_{04}}{\sqrt{MS_{res}}\sqrt{\frac{1}{n_3} + \frac{1}{n_4}}} = \frac{19-27}{\sqrt{7.67}\sqrt{\frac{1}{3}+\frac{1}{2}}} = -3.164 \rightarrow 3.164 > 2.447, \text{ απορ. } H_0$$

κρ. περιοχη: $|t| \geq t_{\frac{a}{2}, n-r} (= t_{0.025,6} = 2.447$

$$MS_{tre} = \frac{SS_{tre}}{r-1}$$

$$SS_{tre} = \sum_{j=1}^{r} \frac{Y_{\cdot j}^2}{n_j} - \frac{Y_{\cdot\cdot}^2}{n}$$

$$E(SS_{tre}) = \sum_{j=1}^{r} \frac{E(Y_{\cdot j}^2)}{n_j} - \frac{E(Y_{\cdot\cdot}^2)}{n} = (*)$$

Ισχύει λοιπόν ο τύπος $Y_{ij} = \mu_j + \varepsilon_{ij}$

$$E(X^2) = \operatorname{Var} X + (EX)^2$$

$$Y_{\cdot j} = \sum_{i=1}^{n_j} Y_{ij}, \quad E(Y_{\cdot j}) = \sum_{i=1}^{n_j} E(Y_{ij}) = \sum_{i=1}^{n_j} \mu_j = n_j \mu_j$$

$$\operatorname{Var}(Y_{\cdot j}) = \sum_{i=1}^{n_j} \operatorname{Var}(Y_{ij}) = \sum_{i=1}^{n_j} \sigma^2 = n_j \sigma^2$$

$$E(Y_{\cdot\cdot}) = \sum_{j}\sum_{i} E(Y_{ij}) = \sum_{j}\sum_{i=1}^{n_j} \mu_j = \sum_{j=1}^{r} n_j \mu_j = n\mu$$

$$\operatorname{Var}(Y_{\cdot\cdot}) = \sum_{j}\sum_{i} \operatorname{Var}(Y_{ij}) = \sum_{j}\sum_{i} \sigma^2 = n\sigma^2$$

$$= (*) = \sum_{j=1}^{r} \frac{n_j\sigma^2 + n_j^2\mu_j^2}{n_j} + \frac{n\sigma^2 + n^2\mu^2}{n} = \underbrace{\sum_{j=1}^{r} \sigma^2 - \sigma^2}_{(r-1)\sigma^2} + \sum_{j=1}^{r} n_j \mu_j^2 - n\mu^2)$$

Άρα $E(MS_{tre}) = \frac{1}{r-1}(r-1)\sigma^2 + \frac{1}{r-1}\left[\sum_{j=1}^{r} n_j \mu_j^2 - n\mu^2\right]$

$$= \sigma^2 + \frac{1}{r-1}\sum_{j=1}^{r} n_j(\mu_j - \mu)^2$$