1-4-15

Συμπερασματολογία για τη διαφορά δύο μέσων τιμών.

Ομάδα 1: 100, 102, 96, 110, 110, 120, 112, 90, 106, 112 | $n_1=10$
Ομάδα 2: 104, 88, 100, 102, 96, 100, 96, 92, 98, 96. $n_2=10$

Υπάρχει διαφορά μεταξύ μ_1 και μ_2;

treatments.

Γενικά: Έστω $x_{11},\dots,x_{1n_1}$, τδ από πληθ. (μ_1,σ_1^2) και
$x_{21},\dots,x_{2n_2}$, τδ από άλλο ανεξ. πληθ. (μ_2,σ_2^2)

Ενδιαφέρει να εκτιμηθεί η διαφορά $\mu_1-\mu_2$ και ο έλεγχος της $H_0: \mu_1=\mu_2$.

Έστω $\bar{x}_1=\frac{1}{n_1}\cdot\sum_{i=1}^{n_1}x_{1i}$, $s_1^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_{1i}-\bar{x}_1)^2$ και

$\bar{x}_2=\frac{1}{n_2}\cdot\sum_{i=1}^{n_2}x_{2i}$, $s_2^2=\frac{1}{n_2-1}\cdot\sum_{i=1}^{n_2}(x_{2i}-\bar{x}_2)^2$.

Δείγμα	Μέγεθος	Βαθμοί ελευθερίας	Μέσες Τιμές	Αθροίσματα τετραγώνων (SS)
1	n_1	n_1-1	$\bar{x}_1$	$\sum_{i=1}^{n_1}(x_{1i}-\bar{x}_1)^2$
2	n_2	n_2-1	$\bar{x}_2$	$\sum_{i=1}^{n_2}(x_{2i}-\bar{x}_2)^2$
Συνολικά		n_1+n_2-2		Συνολικό S.S

$\widehat{\mu_1-\mu_2}=\bar{x}_1-\bar{x}_2$ κ' $Var(\bar{x}_1-\bar{x}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ κ' $\hat{\sigma}_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$

① Κανονικοί πληθυσμοί, γνωστές διακυμάνσεις $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$

Επειδή $\bar{x}_1\sim N(\mu_1,\frac{\sigma_1^2}{n_1})$ κ' $\bar{x}_2\sim N(\mu_2,\frac{\sigma_2^2}{n_2})$ είναι:

$\bar{x}_1-\bar{x}_2\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$

Άρα $$\frac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

έτσι: $(1-\alpha)100\%$ ΔΕ για $\mu_1-\mu_2$: $\overline{x_1}-\overline{x_2} = \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

Για τον έλεγχο των υποθέσεων:

* (i) $H_0: \mu_1-\mu_2 \leq \delta_0$ v $H_\alpha: \mu_1-\mu_2 > \delta_0$ i) $Z \geq Z_\alpha$
(ii) $H_0: \mu_1-\mu_2 \geq \delta_0$ v $H_\alpha: \mu_1-\mu_2 < \delta_0$ ii) $Z \leq -Z_\alpha$
(iii) $H_0: \mu_1-\mu_2 = \delta_0$ v $H_\alpha: \mu_1-\mu_2 \neq \delta_0$ iii) $|Z| \geq Z_{\alpha/2}$

Χρησιμοποιούμε το στατιστικό:

$$Z = \frac{\overline{x_1}-\overline{x_2}-\delta_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \overset{H_0}{\sim} N(0,1)$$

κ' για επιπ. σημαντικότητας α
α κρισιμ. περιοχές.

② Κανονικοί πληθυσμοί, άγνωστες ή ίσες διακυμάνσεις
$N(\mu_1,\sigma^2)$, $N(\mu_2,\sigma^2)$ σ^2: άγνωστο.

$\widehat{\mu_1-\mu_2} = \overline{x_1}-\overline{x_2}$, $Var(\overline{x_1}-\overline{x_2}) = \sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$, $\hat{\sigma}_{\overline{x_1}-\overline{x_2}} = S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$

όπου $S_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} = \frac{\text{Συνολικό SS}}{\text{Συνολικό Β.Ε}}$

Επειδή $\frac{\overline{x_1}-\overline{x_2}-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$

είναι $(1-\alpha)100\%$ ΔΕ για το $\mu_1-\mu_2 = \overline{x_1}-\overline{x_2} \pm t_{\alpha/2,\,n_1+n_2-2} \cdot \left(S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$

Για τον έλεγχο των υποθέσεων

i) * τα ίδια με πάνω i) $t \geq t_{\alpha,\,n_1+n_2-2}$
ii) >> ii) $t \leq -t_{\alpha,\,n_1+n_2-2}$
iii) >> iii) $|t| \geq t_{\alpha/2,\,n_1+n_2-2}$

Χρησιμοπ. το στατιστικό $t = \frac{\bar{x}_1 - \bar{x}_2 - \delta_0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \overset{H_0}{\sim} t_{n_1+n_2-2}$ κ' κρίσιμες περιοχές μεγέθους α.

ΚΟΘ μεγάλα δείγματα

$\bar{x}_1 \overset{\text{προσδ}}{\sim} N\left(\mu_1, \frac{S_1^2}{n_1}\right)$

$\bar{x}_2 \overset{\text{προσδ}}{\sim} N\left(\mu_2, \frac{S_2^2}{n_2}\right)$

$\left.\right\} \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \overset{\text{προσδ.}}{\sim} N(0,1)$

ΔΕ $\mu_1-\mu_2 : \bar{x}_1 - \bar{x}_2 \pm z_{a/2}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$

$Z = \frac{\bar{x}_1 - \bar{x}_2 - \delta_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \underset{\text{προσδ.}}{\overset{H_0}{\sim}} N(0,1)$ κ' κρισ. περιοχή.

Συγκρίσεις κατά ζεύγη - Ho τέστ στις διαφορές

δείγμα	1	2	3
πριν (x_{1i}) :	6.88	7.05	8.24
μετά (x_{2i}) :	6.85	6.84	7.07

διαφορές ($x_{1i}-x_{2i}$): 0.03, 0.21, 1.17, 0.15, −0.02, −0.04,
0.15, 0.52, 0.22, −0.01.

Είναι η επεξεργασία θερμότητας αποτελεσματική (α=0,05)

$H_0 : \mu_1-\mu_2 = 0$ ν $H_a : \mu_1-\mu_2 \neq 0 \ (\delta_0 = 0)$

$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{105,8 - 97,2}{7,11 \cdot \sqrt{1/10 + 1/10}} = 2.70$

$|t| \geq t_{0.025, 18} (=2.101)$

$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{\text{Συνολ. SS}}{\text{Συνολ. βε}}$

$= \frac{909,2}{18} = 50,51$ κ' $S_p = 7,11$ $\quad 2,70 > 2,101$ απορ. Ho

Παραδ.

n ζεύγη : (x_{1i}, x_{2i}) $i=1,\dots,n$

$D_i = x_{1i} - x_{2i}$, $i=1,\dots n$.

Σκοπός: συμπερασματολογία για $\delta = \mu_1 - \mu_2$

Αν $D_i \sim N(\delta, \sigma^2)$

$\hat{\delta} = \bar{D}$

$\hat{\sigma}_{\bar{D}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{S_D}{\sqrt{n}}$ και $\frac{\bar{D}-\delta}{S_D/\sqrt{n}} \sim t_{n-1}$

$(1-a)100\%$ ΔΕ για δ: $\bar{D} \pm t_{a/2,\, n-1} \frac{S_D}{\sqrt{n}}$

και για τον έλεγχο τ } $H_0: \delta = \delta_0$ v $H_a: \delta \neq \delta_0$

$t = \frac{\bar{D} - \delta_0}{S_D/\sqrt{n}} \overset{H_0}{\sim} t_{n-1}$ και κρ.περ $|t| \geq t_{a/2,\, n-1}$

$H_0: \delta (= \mu_1 - \mu_2) = 0$ v $H_a: \delta > 0$

$t = \frac{\bar{D} - 0}{S_D/\sqrt{n}} = 2.0312$

κρ.περ. $t > t_{a,\, n-1} (= t_{0.05,\, 9} = 1.8331)$ απορ. H_0.

$\bar{D} = 0.236$, $S_D^2 = 0.1352$, $S_D = 0.368$