6/4/17

Σύγκριση δύο πιθανοτήτων χωρίς βάρη

100 → 50 → 34 (Ε) p_1 Φάρμακο, 14 (Α)

100 → 50 → 9 (Ε) p_2 Placebo, 41 (Α)

Ho: το φάρμακο δεν έχει κανένα αποτέλεσμα

Ho: $p_1 = p_2$ v Ha: $p_1 > p_2$

Γενικά: Έστω ότι έχουμε 2 ανεξ. πληθυσμούς, Πληθ1 κ' Πληθ2 με τα μέλη των δύο πληθυσμών ν' ανήκουν σε δύο κατηγορίες Α κ' Β.

π.χ. Α η κατηγορία το μέλος να έχει το γνώρισμα α κ' Β να μην έχει το γνώρισμα.

Έστω p_1 η πιθανότητα ένα μέλος του Πληθ.1 να είναι της κατηγορίας Α κ' p_2 η πιθανότητα ένα μέλος του Πληθ.2 να είναι της κατηγορίας Α. Ενδιαφέρει η $p_1 - p_2$

Έστω δείγμα μεγέθους n_1 από τον Πληθ1 και x_1 ο αριθμός των μελών Πληθ.2 κατηγορίας Α. Ανάλογα, δείγμα μεγέθους n_2 από Πληθ2 με x_2 μέλη να είναι της κατηγορίας Α.

$\hat{p}_1 = \dfrac{x_1}{n_1}$ με $X \sim B(n_1, p_1)$ κ' $\hat{p}_2 = \dfrac{x_2}{n_2}$ με $X_2 \sim B(n_2, p_2)$

$$\hat{p}_1 - \hat{p}_2 = \frac{x_1}{n_1} - \frac{x_2}{n_2}$$ (αμερόληπτος) γιατί $E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$

κ' $$\sigma^2_{\hat{p}_1 - \hat{p}_2} = Var(\hat{p}_1 - \hat{p}_2) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$

κ' $$\hat{\sigma}_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

$X_1 \sim B(n_1,p_1) \to X_1 \overset{\text{προσ}}{\sim} N(n_1p_1, n_1p_1(1-p_1)) \to \frac{X_1}{n_1} = \hat{p}_1 \overset{\text{προσ}}{\sim}$

$\overset{\text{προσ}}{\sim} N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right)$

και $\hat{p}_2 \overset{\text{προσ}}{\sim} N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$

$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \overset{\text{προσ}}{\sim} N(0,1)$$

Επειδή $\dfrac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \overset{\text{προσ}}{\sim} N(0,1)$

με το $(1-\alpha)\cdot 100\%$ Δ.Ε για $p_1 - p_2$:

$$L = \hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$U = \hat{p}_1 + \hat{p}_2 + z_{\alpha/2}\sqrt{\quad}$$

Για τον έλεγχο της $H_0: p_1 = p_2 = p$ vs $H_a: p_1 \neq p_2$
Χρησιμοποιούμε το στατιστικό

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{\text{προσ}}{\sim} N(0,1) \text{ όταν } H_0 \text{ αληθεύει}$$

με $\hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$ και για κρ. περιοχή μεγέθους α, $|Z| \geq z_{\alpha/2}$

$H_0: p_1 = p_2$ v $H_1: p_1 > p_2$ (το φάρμακο αποτελεσματικό)

$\hat{p}_1 = \frac{34}{50}$, $\hat{p}_2 = \frac{9}{50}$, $\hat{p} = \frac{34+9}{50+50} = \frac{43}{100}$

95% Δ.Ε. για $p_1 - p_2$: $\hat{p}_1 - \hat{p}_2 \pm \overset{1.96}{z_{0.025}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

$\rightsquigarrow [l, u] = [0.33, 0.67]$

από πίνακα τυπικής κανονικής

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

κρίσιμη περ. $Z \geq z_a (= z_{0.01} = 2.326)$

$\alpha = 0.01$

$$Z = \frac{\frac{34}{50} - \frac{9}{50}}{\sqrt{\frac{43}{100}\left(1 - \frac{43}{100}\right)\left(\frac{1}{50} + \frac{1}{50}\right)}} = 5,05$$

Επίσης $5.05 > 2.326$ απορ. H_0

Τιμή P: $P(Z \geq 5.05 \mid Z \sim N(0,1)) \approx 0.00 < 0,01$

απορρίπτεται H_0

(A) Μία Διακύμανση

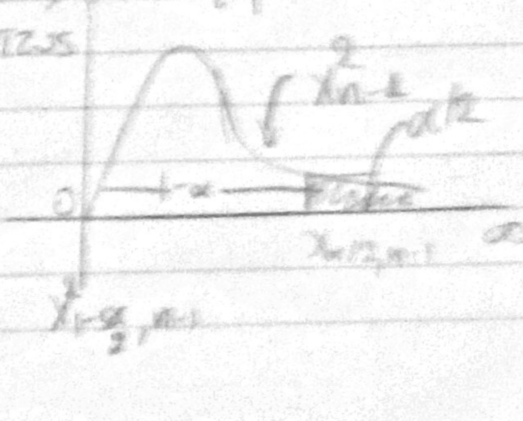
$4.2, 5.3, 10.3$ $(N(\mu, \sigma^2))$ $\quad s^2 = \frac{1}{n-1}\sum_{1}^{n}(x_i - \bar{x})^2$

$\sigma = 2;$ $(a = 0.05)$

Γενικά: $x_1, \ldots, x_n$ τ.δ. από $N(\mu, \sigma^2)$ Ενδιαφέρει το σ^2

$\hat{\sigma}^2 = s^2$, $E(s^2) = \sigma^2$ δηλ. αμερόληπτος

$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$



$$P\left(\chi^2_{1-\frac{\alpha}{2}, n-1} \leq \frac{(n-1)s^2}{\sigma^2} \leq \chi^2_{\frac{\alpha}{2}, n-1}\right) = 1 - \alpha$$

$$P\left(\frac{(n-1)S^2}{X^2_{\frac{a}{2},n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{X^2_{1-\frac{a}{2},n-1}}\right) = 1-\alpha$$

Για τον έλεγχο των:

(i) $H_0: \sigma^2 = \sigma_0^2$ v $H_\alpha: \sigma^2 > \sigma_0^2$ (σ_0^2 γνωστό)

(ii) $H_0: \sigma^2 = \sigma_0^2$ v $H_0: \sigma^2 < \sigma_0^2$

(iii) $H_0: \sigma^2 = \sigma_0^2$ v $H_\alpha: \sigma^2 \neq \sigma_0^2$

χρησιμοποιούμε το στατιστικό: $X^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim X^2_{n-1}$ όταν H_0 αληθής

(i) $X^2 \geq X^2_{\alpha, n-1}$

(ii) $X^2 \leq X^2_{1-\alpha, n-1}$

(iii) $X^2 \leq X^2_{1-\frac{a}{2}, n-1}$

$X^2 \geq X^2_{\frac{a}{2}, n-1}$

$H_0: \sigma = 2$ v $H_\alpha: \sigma \neq 2$

$$X^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{2 \cdot S^2}{4}$$

Κρ. περ. $X^2 \leq X^2_{0.0975,2} (= 0.0506)$ κ' $X^2 \geq X^2_{0.05,2} (= 7.378)$

(B) Σύγκριση δύο διασπορών (Κανονικοί Πληθυσμοί)

A: $n_1 = 10$, $S_1^2 = 1.25$ ($\alpha = 0.10$)

B: $n_2 = 8$, $S_2^2 = 0.28$

Υπάρχει διαφορά στην ποιότητα των A κ' B;

Γενικά: $x_{11}, \ldots, x_{1n}$ από $N(\mu_1, \sigma^2)$ τ.δ.

κ' $x_{21}, \ldots, x_{2n_2}$ τ.δ. από άλλο ανεξ. $N(\mu_2, \sigma_2^2)$

Ενδιαφέρει η σύγκριση των δύο διασπορών

(i) $H_0: \sigma_1^2 \leq \sigma_2^2$ v $H_\alpha: \sigma_1^2 > \sigma_2^2$

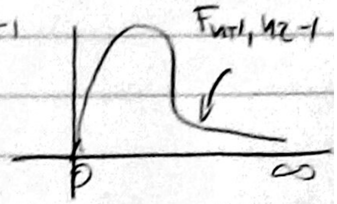
(ii) $H_0: \sigma_1^2 \geq \sigma_2^2$ v $H_\alpha: \sigma_1^2 < \sigma_2^2$

(iii) $H_0: \sigma_1^2 = \sigma_2^2$ v $H_\alpha: \sigma_1^2 \neq \sigma_2^2$

$$\left.\begin{aligned} \frac{(n_1-1)S_1^2}{\sigma_1^2} &\sim \chi^2_{n_1-1} \\ \frac{(n_2-1)S_2^2}{\sigma_2^2} &\sim \chi^2_{n_2-1} \end{aligned}\right\} \Rightarrow \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1,n_2-1}$$

$F = \frac{S_1^2}{S_2^2} \sim F_{n_1-1,n_2-1}$ όταν H_0 αληθεύει

και κρίσιμες περιοχές μεγέθους α.



(i) $H_0: \sigma_1^2 \leq \sigma_2^2$ v $H_\alpha: \sigma_1^2 > \sigma_2^2$ (i) $F \geq F_{\alpha, n_1-1, n_2-1}$

(ii) $H_0: \sigma_1^2 \geq \sigma_2^2$ v $H_\alpha: \sigma_1^2 < \sigma_2^2$ (ii) $F \leq F_{1-\alpha, n_1-1, n_2-1}$

(iii) $H_0: \sigma_1^2 = \sigma_2^2$ v $H_\alpha: \sigma_1^2 \neq \sigma_2^2$ (iii) $F \leq F_{1-\frac{\alpha}{2}, n_1-1, n_2-1}$ κ' $F \geq F_{\frac{\alpha}{2}, n_1-1, n_2-1}$

$H_0: \sigma_1^2 = \sigma_2^2$ v $H_\alpha: \sigma_1^2 \neq \sigma_2^2$

$F = S_1^2 / S_2^2$, κρ. περ. $F \leq F_{0.95, 9, 7} \left(= \frac{1}{F_{0.05,7,9}} = \frac{1}{3.29} = 0.30 \right)$

$F \geq F_{0.05, 9, 7} (= 3.68)$

$F_{\alpha, v_1, v_2} = \frac{1}{F_{1-\alpha, v_2, v_1}}$

$F = \frac{S_1^2}{S_2^2} = \frac{1.25}{0.28} = 4.46$

Επειδή $4.46 > 3.68$ απορ. H_0