Μαθημα 8ο — 23/3/2017

Στατιστικη

Συμπερασματα

		Αποδοχη H_0	Απορριψη H_0
Πραγματικοτητα	Αληθευει H_0	0 πιθ = $1-\alpha$	Σφαλμα Ι πιθ = α
	Αληθευει H_a	Σφαλμα ΙΙ πιθ = β	0 πιθ = $1-\beta$

$\alpha = P$(Σφαλμα Ι) $= P$(απορ H_0 | Αληθευει H_0)

$\beta = P$(Σφαλμα ΙΙ) $= P$(αποδ H_0 | Αληθευει H_a)

Ελεγχος Στατιστικων Υποθεσεων - Στατιστικο Τεστ

παραδειγμα

Δοκιμαστες κρασιων


Στα δυο ποτηρια ιδιο κρασι στο ενα αλλο

$p = P$(πιθανοτητα σωστης διακρισης)

$p = \frac{1}{3}$ η πιθανοτητα να το βρει απο τα 3 ποτηρια χωρις να εχει την ικανοτητα διακρισης

$p > \frac{1}{3}$ η πιθανοτητα να το βρει απο τα 3 ποτηρια αν εχει την ικανοτητα διακρισης

Μηδενικη Υποθεση : H_0 $\quad p = \frac{1}{3}$ Null Hypothesis

Εναλλακτικη Υποθεση : H_a $\quad p > \frac{1}{3}$ Alternative Hypothesis

Θα κανουμε n δοκιμες

Εστω X = αριθμος σωστων διακρισεων (Αριθμος Επιτυχιων)

Αρα εχω οτι $X \sim B(n,p)$

αν εκτελεσω το πειραμα τοτε το $x = 0, 1, .., n$ γνωστο

$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, .., n$

Αποδοσιζουμε το $n = 10$

Κρισιμη περιοχη (κ.π): Η περιοχη απορριψης της H_0, κ.π ή C (Critical regions) ειναι η περιοχη τιμων για την οποια αν το στατιστικο (ειδων x) παρει τιμες που να ανηκουν στην κ.π θα απορριψω την H_0

Εστω $C_5 = \{5, 6, 7, 8, 9, 10\}$

$C_8 = \{8, 9, 10\}$

για το παραδειγμα μας θα υπολογισουμε

$a(C_5) = P(\text{απορ } H_0 \mid \text{Αληθ. } H_0) = P(X \geq 5 \mid X \sim B(n=10, p=\frac{1}{3})) =$

$= \sum_{x=5}^{10} \binom{10}{x} (\frac{1}{3})^x (\frac{2}{3})^{n-x} = \ldots = 0,2131$

$a(C_8) = P(X \geq 8 \mid X \sim B(n=10, p=\frac{1}{3})) = \sum_{x=8}^{10} \binom{10}{x} (\frac{1}{3})^x (\frac{2}{3})^{10-x} = \ldots =$

$= 0,0034$

$B(C_5) = P(\text{δεχ } H_0 \mid \text{αληθ } H_a) = P(X \leq 4 \mid p > \frac{1}{3}) =$

$= P(X \leq 4 \mid X \sim B(n=10, p=0.7)) =$

$= \sum_{x=0}^{4} \binom{10}{x} 0.7^x \, 0.3^{10-x} = \ldots = 0,6171$

Τα στοιχεία του Τεστ

① Τη μηδενική H_0 και εναλλακτική H_a

② Το προκαθορισμένο επίπεδο σημαντικότητας α

③ Τη σ.σ του τεστ και την κρίσιμη περιοχή (κ.π)

④ Τη τιμή της σ.σ που προκύπτει από τα δεδομένα

⑤ Το συμπέρασμα: αν η τιμή της σ.σ ανήκει στην κ.π, απορρίπτουμε την H_0

→ Με βάση το α αυτό που επιλέγω την κ.π που έχει $a \leq 0.05$ και έτσι διαλέγω και το σ

Στο συγκεκριμένο παράδειγμα $C_7 = \{7, 8, 9, 10\}$ $\forall\, p > \frac{1}{3}$ έχει το καλύτερο $a \leq 0.05$ και σ

⊛ Υποθέσεις:

Απλή : $p = \frac{1}{3}$ $\qquad$ $H_0 : p = \frac{1}{3}$ αν $p \leq \frac{1}{3}$

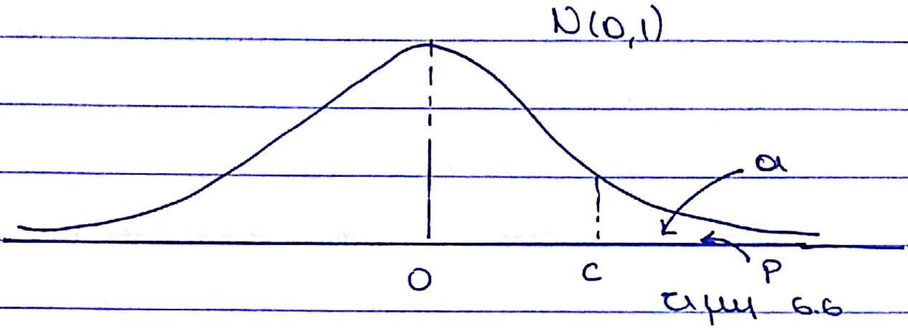
Μονοπλευρες : $p > \frac{1}{3}$ $\qquad$ $H_a : p > \frac{1}{3}$

Δεξιοπλευρη ή Αριστεροπλευρη : $p < \frac{1}{3}$, $C = [c, \infty)$, $C = (-\infty, c]$

Διπλευρη : $p \neq \frac{1}{3}$ $\qquad$ $(\underset{\uparrow\, \alpha/2}{-\infty, -c}] \cup [\underset{\uparrow\, \alpha/2}{c, +\infty})$

Τιμή P: Είναι η πιθανότητα η σ.σ να πάρει τιμές πιο ακραίες από αυτές που παρατηρήθηκε αν αληθεύει η H_0

Κανόνας : αν $P \leq a$ α απορ. την H_0



Παράδειγμα

72, 69, 82, 75, 103, 121, 114, 100, 85, 99

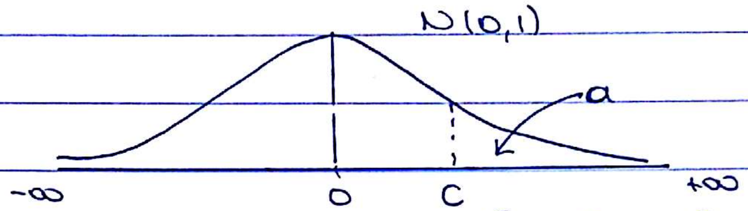
$n=10$, $N(\mu, \sigma^2 = 49)$, $\alpha = 0.05$, μ άγνωστο

$H_0: \mu = 88$ ή $H_a: \mu > 88$

$$\hat{\mu} = \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \equiv N\left(\mu, \frac{49}{10}\right)$$

$$\Rightarrow \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \Rightarrow \frac{\bar{x}-\mu}{\sqrt{49/10}} \sim N(0,1)$$

$$\Rightarrow \frac{\bar{x}-\mu}{7/\sqrt{10}} \sim N(0,1) \quad \text{με } z = \frac{\bar{x}-\mu}{7/\sqrt{10}}$$

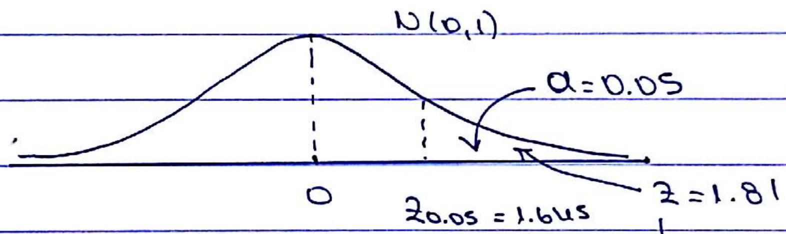


$= C_a = Z_a = Z_{0.05} = 1.645$

$Z = \dfrac{\bar{x}-88}{7/\sqrt{10}} \sim N(0,1)$ οταν H_0 αληθευει

κρ. περιοχη $Z \geq Z_{0.05} = 1.645$, $\bar{x} = 92$ και $Z = 1.81$

Επειδη $1.81 > 1.645$ απορ το H_0



P τιμη $= P(Z \geq 1.81) =$

$= 0.035 \leq 0.05$

απορ H_0