Στατιστική για Πολιτικούς Μηχανικούς
ΕΛΕΓΧΟΣ ΥΠΟΘΕΣΕΩΝ

Δημήτρης Κουγιουμτζής

18 Δεκεμβρίου 2012

Έλεγχος Υποθέσεων

1. Στατιστική υπόθεση
2. Στατιστική ελέγχου και περιοχή απόρριψης
3. Απόφαση ελέγχου

Στατιστική υπόθεση

μηδενική υπόθεση H_0: $\theta = \theta_0$
εναλλακτική υπόθεση H_1:

1. $H_1 : \theta \neq \theta_0$, *δίπλευρος έλεγχος*
2. $H_1 : \theta < \theta_0$, *μονόπλευρος έλεγχος*
   Σύνολο δυνατών τιμών της θ: θ_0 και $< \theta_0$
   Σωστότερα $H_0 : \theta \geq \theta_0$ και $H_1 : \theta < \theta_0$
3. $H_1 : \theta > \theta_0$, όπως (2)

Στατιστική ελέγχου, περιοχή απόρριψης

Δυνατές τιμές της θ χωρίζονται:

- Περιοχή αποδοχής της H_0
- Περιοχή απόρριψης R της H_0

Επίπεδο εμπιστοσύνης $(1-\alpha)$ για την απόφαση ελέγχου
α: *επίπεδο σημαντικότητας*

Η περιοχή αποδοχής και απόρριψης αλλάζει με το α

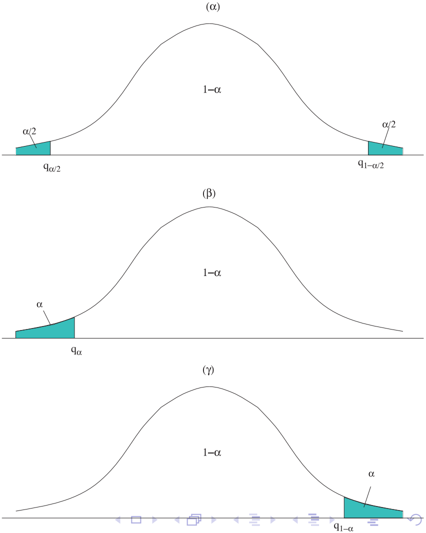
Κατανομή για την εκτιμήτρια $\hat{\theta}$ της θ;
- *παραμετρικός έλεγχος*: υποθέτουμε κατανομή για $\hat{\theta}$
- *μή–παραμετρικός έλεγχος*: δεν υποθέτουμε κατανομή για $\hat{\theta}$

Παραμετρικός έλεγχος

$\hat{\theta} \quad \rightarrow \quad q$ στατιστική ελέγχου

q έχει γνωστή κατανομή (είναι z, t ή χ^2)

$\alpha \quad \rightarrow$ κρίσιμη(ες) τιμή(ές) της $q \quad \rightarrow \quad R$



Απόφαση ελέγχου

Υπολογίζουμε από το δείγμα $\hat{\theta} \longrightarrow \tilde{q}$

$\tilde{q} \in R \rightarrow$ απόρριψη της H_0
$\tilde{q} \notin R \rightarrow$ μη απόρριψη της H_0

p-τιμή του ελέγχου
Για $\tilde{q}$ (γνωστή κατανομή της q όταν ισχύει η H_0)

1. $p = \mathrm{P}(q < -|\tilde{q}| \ \vee \ q > |\tilde{q}|)$
2. $p = \mathrm{P}(q < \tilde{q})$
3. $p = \mathrm{P}(q > \tilde{q})$

Η p-τιμή δηλώνει το μικρότερο επίπεδο σημαντικότητας για το οποίο μπορούμε να απορρίψουμε την H_0 με βάση το δείγμα.

Διαδικασία στατιστικού ελέγχου

1. Επιλογή επιπέδου σημαντικότητας α
2. Καθορισμός στατιστικής ελέγχου q
3. Υπολογισμός κρίσιμης (-ων) τιμής (-ών) της q από τον αντίστοιχο στατιστικό πίνακα και καθορισμός της περιοχής απόρριψης R
4. Υπολογισμός της στατιστικής ελέγχου $\tilde{q}$ από το δείγμα
5. Απόρριψη της H_0 αν $\tilde{q}$ ανήκει στην R.
   Μη απόρριψη της H_0 αν $\tilde{q}$ δεν ανήκει στην R.

Διαδικασία στατιστικού ελέγχου

Γενικά όταν θ είναι μ, p, $\mu_1 - \mu_2$, $p_1 - p_2$:

$\mathrm{H}_0 : \theta = \theta_0$

Στατιστική ελέγχου: $q = \frac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}}$

Κατανομή της q: $\begin{cases} \text{Τυπική κανονική } N(0,1) \\ \text{student με } \nu \text{ βαθμούς ελευθερίας } t_\nu \end{cases}$

Περιοχή απόρριψης (επίπεδο σημαντικότητας α):

1. $\mathrm{H}_1 : \theta \neq \theta_0 : R = \{q |\ q < q_{\alpha/2} \ \vee \ q > q_{1-\alpha/2}\}$
2. $\mathrm{H}_1 : \theta < \theta_0 : R = \{q |\ q < q_{\alpha}\}$
3. $\mathrm{H}_1 : \theta > \theta_0 : R = \{q |\ q > q_{1-\alpha}\}$

Διαδικασία στατιστικού ελέγχου (συνέχεια)

Δειγματική στατιστική ελέγχου: $\tilde{q} = \frac{\hat{\theta}-\theta_0}{\sigma_{\hat{\theta}}}$

$\tilde{q} \in R \quad \rightarrow \quad$ απόρριψη της H_0
$\tilde{q} \notin R \quad \rightarrow \quad$ μη απόρριψη της H_0

p-τιμή:

1. $p = P(q < \tilde{q} \ \vee \ q > \tilde{q})$
2. $p = P(q < \tilde{q})$
3. $p = P(q > \tilde{q})$

Παράδειγμα: Έλεγχος μέσης τιμής, γνωστό σ^2

Θέλουμε να ελέγξουμε σε επίπεδο σημαντικότητας $\alpha = 0.05$ αν η μέση αντοχή θραύσης σκυροδέματος μ είναι 6 ksi.
Η διασπορά είναι γνωστή, $\sigma^2 = 0.4\,\text{ksi}^2$.

$\text{H}_0 : \mu = 6$ και $\text{H}_1 : \mu \neq 6$ [δίπλευρος έλεγχος]

σ^2 γνωστή, $X \sim \text{N}(\mu, \sigma^2)$

$$q = z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim \text{N}(0, 1)$$

Κρίσιμη τιμή: $z_{0.975} = 1.96 \quad \Rightarrow \quad R = \{z |\; |z| > 1.96\}$

Παράδειγμα: Έλεγχος μέσης τιμής, γνωστό σ^2 (συνέχεια)

Στατιστική ελέγχου από το δείγμα ($\bar{x} = 5.67$, $n = 25$)

$$\tilde{z} = \frac{5.67 - 6}{\sqrt{0.4/25}} = -2.61$$

$\tilde{z} \in R \quad \rightarrow \quad$ απόρριψη της H_0 για $\alpha = 0.05$

p-τιμή για $\tilde{z} = -2.61$

$$\begin{aligned} p &= 2(1 - \mathrm{P}(z < |\tilde{z}|)) \\ &= 2(1 - \Phi(|\tilde{z}|)) \\ &= 2(1 - \Phi(2.61)) \\ &= 2(1 - 0.9955) = 0.009 \end{aligned}$$

Παράδειγμα: Έλεγχος μέσης τιμής, άγνωστό σ^2

Άγνωστη διασπορά αντοχής θραύσης σκυροδέματος.

$H_0 : \mu = 6$ και $H_1 : \mu \neq 6$ [δίπλευρος έλεγχος]

σ^2 άγνωστο, $n = 25 < 30$, $X \sim N(\mu, \sigma^2)$

$$q = t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$$

Κρίσιμη τιμή: $t_{24,0.975} = 2.06$

Περιοχή απόρριψης της H_0, $R = \{t| \ |t| > 2.06\}$

Παράδειγμα: Έλεγχος μέσης τιμής, άγνωστό σ^2 (συνέχεια)

Στατιστική ελέγχου από το δείγμα
$(\bar{x} = 5.67,\ n = 25,\ s^2 = 0.375)$

$$\tilde{t} = \frac{5.67 - 6}{\sqrt{0.375/25}} = -2.69$$

$\tilde{t} \in R \quad \rightarrow \quad$ απόρριψη της H_0 για $\alpha = 0.05$.

Διάστημα εμπιστοσύνης $[5.42,\ 5.92]$

p-τιμή για $\tilde{t} = -2.69$

$$\begin{aligned} p &= 2(1 - \mathrm{P}(t \leq |\tilde{t}|)) \\ &= 2(1 - \mathrm{P}(t \leq 2.69)) = 2(1 - 0.994) = 0.012 \end{aligned}$$

Έλεγχος υπόθεσης $H_0: \mu = \mu_0$ στις διάφορες περιπτώσεις

διασπορά	*κατανομή της X*	*n*	*κατανομή στατιστικής q*
γνωστή	κανονική		$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$
γνωστή	μη κανονική	μεγάλο	$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$
γνωστή	μη κανονική	μικρό	—
άγνωστη		μεγάλο	$z = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim N(0,1)$
άγνωστη	κανονική	μικρό	$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$
άγνωστη	μη κανονική	μικρό	—

Παράδειγμα: Έλεγχος διασποράς

Θέλουμε να ελέγξουμε αν θα μπορούσαμε να πούμε με μεγάλη σιγουριά (επίπεδο εμπιστοσύνης 99%) πως η τυπική απόκλιση σ της αντοχής θραύσης σκυροδέματος μπορεί να ξεπεράσει το επίπεδο του 0.5 ksi.

Έλεγχος για τη διασπορά σ^2:

$\mathrm{H}_0 : \sigma^2 \leq 0.25$, $\mathrm{H}_1 : \sigma^2 > 0.25$, [μονόπλευρος έλεγχος]

$$q = \chi^2 \equiv \frac{(n-1)s^2}{\sigma_0^2} \sim \mathcal{X}^2_{n-1},$$

Κρίσιμη τιμή: $\chi^2_{24,0.99} = 42.98 \quad \Rightarrow$
Περιοχή απόρριψης της H_0: $R = \{\chi^2 | \; \chi^2 > 42.98\}$

Παράδειγμα: Έλεγχος διασποράς (συνέχεια)

Στατιστική ελέγχου από το δείγμα
$(n = 25,\ s^2 = 0.375\ (\mathrm{ksi})^2)$

$$\tilde{\chi}^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{24 \cdot 0.375}{0.25} = 36.0$$

$\tilde{\chi}^2 \notin R \quad \rightarrow \quad$ μή απόρριψη της H_0 για $\alpha = 0.01$

p-τιμή για $\tilde{\chi}^2 = 36.0$

$$\begin{aligned} p &= 1 - \mathrm{P}(\chi^2 < \tilde{\chi}^2) \\ &= 1 - \mathrm{P}(\chi^2 \leq 36.0) = 1 - 0.945 = 0.055!!! \end{aligned}$$

Παράδειγμα: Έλεγχος αναλογίας

Μια εταιρεία ενδιαφέρεται να ελέγξει σε σημαντικότητα 0.05 αν η αναλογία των σκουριασμένων ραβδών χάλυβα στην αποθήκη της είναι 0.19.

Δείγμα: 12 σκουριασμένες σε σύνολο 100 ραβδών.

$\mathrm{H}_0 : p = 0.19 \quad \mathrm{H}_1 : p \neq 0.19$ [δίπλευρος έλεγχος]

n μεγάλο

$$q = z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim \mathrm{N}(0,1)$$

Κρίσιμη τιμή $z_{0.975} = 1.96$

Περιοχή απόρριψης $R = \{z \mid |z| > 1.96\}$

Παράδειγμα: Έλεγχος αναλογίας (συνέχεια)

Στατιστική ελέγχου από το δείγμα ($\hat{p} = 0.12$)

$$\tilde{z} = \frac{0.12 - 0.19}{\sqrt{\frac{0.19 \cdot 0.81}{100}}} = -1.784.$$

$\tilde{z} \notin R \;\; \rightarrow \;\;$ μη απόρριψη της H_0 για $\alpha = 0.05$

Διάστημα εμπιστοσύνης $[0.056, \; 0.184]$?

Υπολογίστηκε από $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Παράδειγμα: Έλεγχος αναλογίας (συνέχεια)

Η εταιρεία γνωρίζει ότι είναι απίθανο το ποσοστό να κυμαίνεται σε επίπεδα μεγαλύτερα του 19%

$\mathrm{H}_0 : p \geq 0.19, \quad \mathrm{H}_1 : p < 0.19$ [μονόπλευρος έλεγχος]

Κρίσιμη τιμή $z_{0.95} = 1.65$
Περιοχή απόρριψης $R = \{z | \; z < -1.65\}$

$\tilde{z} = -1.784$ (όπως πριν)

$\tilde{z} \in R \;\; \rightarrow$ απόρριψη της H_0 για $\alpha = 0.05$
p-τιμή

$$\begin{aligned} p &= \mathrm{P}(z < \tilde{z}) = \mathrm{P}(z < -1.784) = \Phi(-1.784) \\ &= 1 - \Phi(1.784) = 1 - 0.963 = 0.037 \end{aligned}$$

Ανώτατο επίπεδο εμπιστοσύνης που μπορούμε να απορρίψουμε την $\mathrm{H}_0 : p \geq 0.19$ είναι περίπου 96%.

Παράδειγμα: Έλεγχος διαφοράς μέσων τιμών, γνωστό σ^2

Θέλουμε να ελέγξουμε αν δύο τύποι A και B σκυροδέματος έχουν την ίδια μέση αντοχή θραύσης. Θεωρούμε γνωστή και κοινή διασπορά $\sigma^2 = 0.4\,(\text{ksi})^2$.

$\mathrm{H}_0 : \mu_1 = \mu_2, \;\; \mathrm{H}_1 : \mu_1 \neq \mu_2$ [δίπλευρος έλεγχος]

γνωστή και κοινή διασπορά, κανονικές κατανομές

$$q = z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \mathrm{N}(0,1)$$

Κρίσιμη τιμή: $z_{0.975} = 1.96$

Περιοχή απόρριψης της H_0, $R = \{z \mid |z| > 1.96\}$

Παράδειγμα: Έλεγχος διαφοράς μέσων τιμών, γνωστό σ^2 (συνέχεια)

Στατιστική ελέγχου από το δείγμα
$(\bar{x}_1 = 5.67, \;\; \bar{x}_2 = 5.38, \;\; n_1 = 25, \;\; n_2 = 20)$

$$\tilde{z} = \frac{5.67 - 5.38}{\sqrt{\frac{0.4}{25} + \frac{0.4}{20}}} = 1.53$$

$\tilde{z} \notin R \quad \rightarrow \quad$ μή απόρριψη της H_0 για $\alpha = 0.05$

p-τιμή για $\tilde{z} = 1.53$

$$\begin{aligned} p &= 2(1 - \mathrm{P}(z < |\tilde{z}|)) = 2(1 - \Phi(|\tilde{z}|)) \\ &= 2(1 - \Phi(1.53)) = 2(1 - 0.937) = 0.126 \end{aligned}$$

Παράδειγμα: Έλεγχος διαφοράς μέσων τιμών, άγνωστο σ^2

Διασπορά αντοχής θραύσης κοινή αλλά άγνωστη

$\mathrm{H}_0 : \mu_1 = \mu_2, \;\; \mathrm{H}_1 : \mu_1 \neq \mu_2$ [δίπλευρος έλεγχος]

άγνωστη και κοινή διασπορά, κανονικές κατανομές

$$q = t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim \mathrm{t}_{n_1+n_2-2}$$

Κρίσιμη τιμή: $t_{43,0.975} = 2.02$

Περιοχή απόρριψης της H_0, $R = \{t| \; |t| > 2.02\}$

Παράδειγμα: Έλεγχος διαφοράς μέσων τιμών, άγνωστο σ^2 (συνέχεια)

Στατιστική ελέγχου από το δείγμα
($\bar{x}_1 = 5.67,\ \ \bar{x}_2 = 5.38,\ \ n_1 = 25,\ \ n_2 = 20$
$s_1^2 = 0.375,\ s_2^2 = 0.326 \rightarrow s^2 = 0.353,\ s = 0.594$)

$$\tilde{t} = \frac{5.67 - 5.38}{0.594\sqrt{\frac{1}{25} + \frac{1}{20}}} = 1.63$$

$\tilde{t} \notin R \quad \rightarrow \quad$ μή απόρριψη της H_0 για $\alpha = 0.05$
Διάστημα εμπιστοσύνης $[-0.07,\ 0.65]$

p-τιμή για $\tilde{t} = 1.63$

$$\begin{aligned} p &= 2(1 - \mathrm{P}(t < |\tilde{t}|)) \\ &= 2(1 - \mathrm{P}(t \leq 1.63)) = 2(1 - 0.945) = 0.11 \end{aligned}$$

Θέμα 13

Καθορισμός μεγέθους δείγματος για έλεγχο διαφοράς μέσων τιμών από ανεξάρτητα δείγματα.

Παράδειγμα: Έλεγχος διαφοράς αναλογιών

Θέλουμε να ελέγξουμε αν οι αναλογίες σκουριασμένων ραβδών χάλυβα είναι ίδιες σε δύο αποθήκες.

Δεδομένα:

πρώτη αποθήκη, 12 στις 100 ράβδους σκουριασμένες,

δεύτερη αποθήκη, 26 στις 120 ράβδους σκουριασμένες.

$\mathrm{H}_0 : p_1 = p_2 \quad \mathrm{H}_1 : p_1 \neq p_2$ [δίπλευρος έλεγχος]

n μεγάλο

$$q = z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \sim \mathrm{N}(0,1)$$

Σύμφωνα με τη μηδενική υπόθεση είναι $p_1 = p_2 = p$

Η στατιστική ελέγχου γίνεται

$$q = z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim \mathrm{N}(0,1)$$

Παράδειγμα: Έλεγχος διαφοράς αναλογιών (συνέχεια)

Κρίσιμη τιμή $z_{0.975} = 1.96$

Περιοχή απόρριψης $R = \{z \mid |z| > 1.96\}$

$\hat{p}_1 = 0.12, \quad \hat{p}_2 = 0.217 \quad \rightarrow$ κοινή αναλογία

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{100 \cdot 0.12 + 120 \cdot 0.217}{100 + 120} = 0.173$$

Στατιστική ελέγχου από το δείγμα

$$\tilde{z} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.12 - 0.217}{\sqrt{0.173 \cdot (1-0.173) \cdot \left(\frac{1}{100}+\frac{1}{120}\right)}} = -1.89$$

Παράδειγμα: Έλεγχος διαφοράς αναλογιών (συνέχεια)

$\tilde{z} \notin R \;\; \rightarrow$ μή απόρριψη της H_0 για $\alpha = 0.05$

Διάστημα εμπιστοσύνης $[-0.198, \; 0.004]$

p-τιμή για $\tilde{z} = -1.89$

$$\begin{aligned} p &= 2(1 - \mathrm{P}(z < |\tilde{z}|)) = 2(1 - \Phi(|\tilde{z}|)) \\ &= 2(1 - \Phi(1.89)) = 2(1 - 0.97) = 0.06 \end{aligned}$$

Παράδειγμα: Έλεγχος διαφοράς αναλογιών (συνέχεια)

Έστω ότι γνωρίζουμε πως η δεύτερη αποθήκη έχει κατά κανόνα πιο παλιές ράβδους από την πρώτη
$\mathrm{H}_0 : p_1 \geq p_2, \;\; \mathrm{H}_1 : p_1 < p_2$ [μονόπλευρος έλεγχος]
Κρίσιμη τιμή $z_{0.95} = 1.65$
Περιοχή απόρριψης $R = \{z | \; z < -1.65\}$
$\tilde{z} = -1.89$ (όπως πριν)
$\tilde{z} \in R \;\; \rightarrow \;\;$ απόρριψη της H_0 για $\alpha = 0.05$
p-τιμή

$$\begin{aligned} p &= \mathrm{P}(z < \tilde{z}) = \mathrm{P}(z < -1.89) = \Phi(-1.89) \\ &= 1 - \Phi(1.89) = 1 - 0.97 = 0.03 \end{aligned}$$

Ανώτατο επίπεδο εμπιστοσύνης που μπορούμε να απορρίψουμε την $\mathrm{H}_0 : p_1 \geq p_2$ είναι περίπου 97%.

Χαρακτηριστικά Ελέγχου

1. *Σωστή απόφαση* (σωστή H_0) με πιθανότητα

$$\mathsf{P}(\text{αποδοχή της } \mathsf{H}_0 \mid \mathsf{H}_0 \text{ σωστή}) = 1 - \alpha$$

2. **Σφάλμα τύπου I** με πιθανότητα

$$\mathsf{P}(\text{απόρριψη της } \mathsf{H}_0 \mid \mathsf{H}_0 \text{ σωστή}) = \alpha.$$

3. **Σφάλμα τύπου II** με πιθανότητα

$$\mathsf{P}(\text{αποδοχή της } \mathsf{H}_0 \mid \mathsf{H}_0 \text{ λανθασμένη}) = \beta.$$

4. *Σωστή απόφαση* (σωστή H_1) με πιθανότητα

$$\mathsf{P}(\text{απόρριψη της } \mathsf{H}_0 \mid \mathsf{H}_0 \text{ λανθασμένη}) = 1 - \beta$$

[**δύναμη ελέγχου**]

Χαρακτηριστικά Ελέγχου (συνέχεια)