ΕΚΤΙΜΗΤΙΚΗ ΚΑΙ ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ

ΦΡΟΝΤΙΣΤΗΡΙΟ #6

#1 Το περιοδικό Journal of Fish Biology (Αύγουστος 1990) δημοσίευσε μία μελέτη που έκανε σύγκριση των παράσιτων που βρέθηκαν στα είδη ψαριών στη Μεσόγειο και στον Ατλαντικό. Στη Μεσόγειο από τα 588 ψάρια που πιάστηκαν και εξετάστηκαν βρέθηκαν μολυσμένα από παράσιτα τα 211. Στον Ατλαντικό Ωκεανό, από 123 που εξετάστηκαν βρέθηκαν μολυσμένα τα 26. Συγκρίνετε την αναλογία των παράσιτων στις δύο θάλασσες χρησιμοποιώντας ένα 90% δ.ε. Ερμηνεύστε το διάστημα.

Λύση Η εκτιμώμενη αναλογία των μολυσμένων ψαριών στη Μεσόγειο είναι $\hat{p}_1 = \frac{211}{588} = 0.36$ ενώ στον Ατλαντικό είναι $\hat{p}_2 = \frac{26}{123} = 0.211$.

Ένα 90% δ.ε. για τη διαφορά των αναλογιών $p_1 - p_2$ δίνεται από τον τύπο:

$$\hat{p}_1 - \hat{p}_2 \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$ όπου $n_1 = 588$,

$n_2 = 123$, $z_{1-\alpha/2} = z_{1-0.05} = z_{0.95} = 1.64$

Άρα,

$$0.36 - 0.211 \pm 1.64\sqrt{\frac{0.36 \cdot 0.64}{588} + \frac{0.211 \cdot 0.789}{123}}$$

$$= 0.149 \pm 0.069 \Rightarrow (0.08, 0.218)$$

Το δ.ε. αναφέρεται στη διαφορά των πραγματικών αναλογιών p_1 και p_2 και παρατηρούμε ότι η διαφορά αυτή είναι πάντα θετική. Αυτό σημαίνει ότι σε κάθε περίπτωση, η αναλογία p_1 είναι μεγαλύτερη της p_2 δηλ. η Μεσόγειος Θάλασσα έχει μεγαλύτερη αναλογία μολυσμένων ψαριών από τον Ατλαντικό Ωκεανό.

#2 Από προηγούμενη εμπειρία, είναι γνωστό ότι η τυπική απόκλιση τιμών pH για ένα υπό έλεγχο χημικό είναι 0.19. Μία παρτίδα 10 δειγμάτων από έναν καινούριο προμηθευτή δίνει τις ακόλουθες τιμές pH

7.6 8.4 8.3 8.0 8.2 7.4 7.8 8.2 8.1 7.7

(α) Βρείτε τη μέση τιμή και τη διακύμανση του δείγματος

(β) Βρείτε ένα 95% δ.ε. για τη μέση τιμή βασιζόμενοι στην γνωστή από προηγούμενη εμπειρία τυπική απόκλιση.

(γ) Δώστε ένα προσεγγιστικό 95% δ.ε. για την νέα τυπική απόκλιση.

(δ) Είναι η υπόθεση της τιμής 0.19 για την τυπική απόκλιση βάσιμη;

(ε) Βρείτε ένα 95% δ.ε. για τη μέση τιμή χρησιμοποιώντας τη δειγματική τυπική απόκλιση.

Λύση (α) Είναι $\overline{X}_{10} = 8$, $S^* = 0.36515$, $n = 10$

(β) Ένα 95% δ.ε. για τη μέση τιμή μ (αφού $\sigma = 0.19$) είναι $\overline{X}_{10} \pm z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow 8 \pm z_{0.975}\frac{0.19}{\sqrt{10}}$

$$\Rightarrow 8 \pm 1.96\,\frac{0.19}{\sqrt{10}} \Rightarrow (7.88,\ 8.12)$$

(γ) Ένα 95% δ.ε. για την νέα τυπική απόκλιση σ' είναι:

$$\frac{\sqrt{n-1}\ S^*}{\chi^2_{9,0.975}} \le \sigma' \le \frac{\sqrt{n-1}\ S^*}{\chi^2_{9,0.025}}$$

Από τους πίνακες:

$$\chi^2_{9;0.975} = 19.02, \quad \chi^2_{9;0.025} = 2.7$$

Με αντικατάσταση:

$$0.05759 \leq \sigma' \leq 0.4057$$

(δ) Η υπόθεση της τιμής 0.19 για την τυπική απόκλιση έχει βάση διότι σύμφωνα με το 95% δ.ε. του ερωτήματος (γ) 95 στα 100 διαστήματα εμπιστοσύνης για την τυπική απόκλιση περιέχουν ως πιθανή τιμή της τυπικής απόκλισης την τιμή 0.19.

(ε) Θεωρώντας ότι η τυπική απόκλιση σ είναι άγνωστη ένα 95% δ.ε. είναι: $\bar{X}_{10} \pm t_{9;1-\frac{0.05}{2}} \cdot \frac{s^*}{\sqrt{n}}$

Με αντικατάσταση: $8 \pm t_{9;0.975} \frac{0.365}{\sqrt{10}}$

$$= 8 \pm 2.26 \cdot \frac{0.365}{\sqrt{10}}$$

$$\Rightarrow (7.739, 8.261).$$

#3 Σε τ.δ. $n=25$ κουτιών παστεριωμένου γάλακτος υπολογίσαμε τυπική απόκλιση του βάρους $s^* = 1.6$ gr. Ζητούνται:

(α) Να βρεθεί ένα 95% δ.ε. για τη διακύμανση του καθαρού βάρους σε όλα τα κουτιά αν είναι γνωστό ότι η κατανομή του βάρους είναι κανονική.

(β) Αν η ίδια τυπική απόκλιση έχει υπολογιστεί σε τ.δ. $n=250$ κουτιών, το εύρος του διαστήματος εμπιστοσύνης θα ήταν μεγαλύτερο ή μικρότερο απ' ότι στο (α);

Λύση

(α) Επειδή η κατανομή του πληθυσμού είναι κανονική, η τ.μ.

$$\chi^2 = \frac{(n-1)S^{*2}}{\sigma^2} \sim \chi^2_{24}$$

Ένα 95% δ.ε. για τη σ^2 είναι: $(n=25)$

$$\frac{(n-1)S^{*2}}{\chi^2_{24;1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)S^{*2}}{\chi^2_{24;\alpha/2}}, \quad \alpha = 0.05$$

Με αντικατάσταση από τους πίνακες:

$$\frac{24\cdot(1.6)^2}{39.36} \le \sigma^2 \le \frac{24\cdot(1.6)^2}{12.4}$$

$$\Leftrightarrow \underline{1.56 \le \sigma^2 \le 4.95}$$

(β) Όσο αυξάνει το μέγεθος το μέγεθος του δείγματος, τόσο αυξάνει η ακρίβεια της εκτίμησης. Επομένως, σε μέγεθος δείγματος $n=250$ κουτιών αντιστοιχεί πολύ "μικρότερο" διάστημα εμπιστοσύνης

#4 Από μία μεγάλη παρτίδα κονσέρβες ροδάκινου λάβαμε τ.δ. μεγέθους $n=50$ και βρήκαμε $X=5$ ελαττωματικές.

(α) Να βρεθεί ένα 99% δ.ε. για την αναλογία των ελαττωματικών στο σύνολο της παρτίδας.

(β) Αν τετραπλασιαστεί το μέγεθος του δείγματος n τι γίνεται με το εύρος d του διαστήματος;

Λύση (α) Η δειγματική αναλογία ισούται με $\hat{p}_X = \frac{5}{50} = 0.10$

Επειδή το μέγεθος δείγματος είναι μεγάλο, το ζητούμενο δ.ε. είναι: $\hat{p}_X - z_{1-\alpha/2}\sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n}} \le p \le \hat{p}_X + z_{1-\alpha/2}\sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n}}$

Για $1-\alpha = 0.99 \Rightarrow \alpha/2 = 0.005 \Rightarrow z_{1-\alpha/2} = z_{0.995} = 2.58$

Άρα το 99% δ.ε. είναι:

$$0.10 - 2.58\sqrt{\frac{(0.10)(0.9)}{50}} \le p \le 0.10 + 2.58\sqrt{\frac{0.10 \cdot 0.9}{50}}$$

$\Leftrightarrow 0.10 - 0.1095 \le p \le 0.10 + 0.1095$

Όμως πρέπει $p \ge 0$

$\Rightarrow$ Συνεπώς

$$0 \le p \le 0.2095$$

$\Rightarrow p \in [0, 0.2095]$.

(β) Το εύρος d του διαστήματος είναι:

$$d = \hat{p}_x + z_{1-\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n}} - \hat{p}_x - z_{1-\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n}}$$

$$= 2 z_{1-\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n}}$$

Αν $n^* = 4n$ τότε

$$d^* = 2 z_{1-\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{4n}} = z_{1-\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n}}$$

$$= \frac{d}{2}$$

δηλ. αν τετραπλασιαστεί το μέγεθος του δείγματος τότε υποδιπλασιάζεται το εύρος του διαστήματος.

Γενικά, αποδεικνύεται ότι: Για να μειωθεί το εύρος d του διαστήματος από d στο $\frac{d}{\delta}$, $\delta > 1$, το μέγεθος του δείγματος θα πρέπει να πολλαπλασιαστεί επί δ^2, δηλ. να γίνει $\delta^2 n$.