ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ & ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

Εφαρμοσμένη Στατιστική- 28 Αυγούστου 2012

Άσκηση 1 (25 μονάδες)

α) Αν ρίξουμε μια φορά ένα τίμιο ζάρι ποια είναι η αναμενόμενη τιμή μ και ποια η διασπορά σ^2 του αποτελέσματος;
β) Ρίχνουμε $n = 105$ φορές ένα τίμιο ζάρι. Αν S_n είναι το άθροισμα όλων των ζαριών, υπολογίστε προσεγγιστικά την πιθανότητα p του ενδεχομένου $\{S_n \leq 350\}$.
γ) Πώς πρέπει να διαλέξουμε το $k \in \mathbb{N}$ ώστε $\mathbb{P}[S_n \leq k] \approx 0,95$;
δ) Αν επαναλάβουμε $m = 5$ φορές το πείραμα του ερωτήματος (β), υπολογίστε προσεγγιστικά την πιθανότητα του ενδεχομένου σε τουλάχιστον μια από τις 5 προσπάθειες το άθροισμα των ζαριών να μην ξεπεράσει το 350.

Άσκηση 2 (25 μονάδες)

Δίνεται η από κοινού σ.π.π. των θετικών τ.μ. X, Y

$$f(x,y) = C(\lambda)x^3 y e^{-\lambda x(1+y)}, \qquad x, y > 0,$$

όπου λ είναι μια άγνωστη θετική παράμετρος και $C(\lambda)$ μια θετική σταθερά.
α) Δείξτε ότι $C(\lambda) = \lambda^4$.
β) Εκτιμήστε την άγνωστη παράμετρο λ από ένα τυχαίο δείγμα $\{(x_1, y_1), (x_2, y_2), \ldots, (x_\nu, y_\nu)\}$ των (X, Y), με τη μέθοδο της μέγιστης πιθανοφάνειας.
γ) Υπολογίστε την από κοινού κατανομή των X, Z όπου $Z = XY$ και συμπεράνετε ότι οι X, Z είναι ανεξάρτητες και ισόνομες.

Άσκηση 3 (30 μονάδες)

Ο χρόνος αναμονής T (σε sec) μέχρι τη διάσπαση ενός ραδιενεργού πυρήνα ακολουθεί εκθετική κατανομή $\mathcal{E}xp(\lambda)$, όπου λ είναι μια άγνωστη θετική παράμετρος.
α) Υπολογίστε σαν συνάρτηση του λ την πιθανότητα p του ενδεχομένου ένας τέτοιος πυρήνας να μην έχει διασπαστεί μετά από t sec, δηλ. $p = p(\lambda) = \mathbb{P}[T > t]$.
β) Αρχικά έχουμε N τέτοιους πυρήνες που διασπώνται ανεξάρτητα και ορίζουμε την τ.μ. $Q_N(t)$ ως το κλάσμα των πυρήνων που δεν έχουν διασπαστεί μετά από t sec, δηλ.

$$Q_N(t) = \frac{1}{N}\sum_{i=1}^{N} X_i,$$

όπου $X_i = 1$ στο ενδεχόμενο $\{T_i > t\}$ και 0 διαφορετικά, δείξτε ότι η $Q_N(t)$ είναι η ΕΜΠ της παραμέτρου p.
γ) Υπολογίστε την ΕΜΠ $\hat{\lambda}_N$ της άγνωστης παραμέτρου λ.
δ) Δείξτε ότι με πιθανότητα 1

$$\lim_{N\to\infty} Q_N(t) = e^{-\lambda t}.$$

ε) Με τη βοήθεια του κεντρικού οριακού θεωρήματος υπολογίστε την οριακή κατανομή της τ.μ. $\sqrt{N}\big(Q_N(t) - e^{-\lambda t}\big)$ και χρησιμοποιήστε τα παραπάνω αποτελέσματα για να βρείτε ένα προσεγγιστικό 0,99-δ.ε. για την άγνωστη παράμετρο λ. Υπολογίστε προσεγγιστικά το εύρος αυτού του διαστήματος αν $N = 10^{22}$ και $Q_N(1) = 1/2$.

Άσκηση 4 (20 μονάδες)

Ένα εργαστήριο χρησιμοποιεί δύο μεθόδους για τη χρονολόγηση οστών. Για το σφάλμα κάθε μεθόδου υποθέτουμε ότι ακολουθεί κανονική κατανομή με μέση τιμή 0 και διασπορά σ_i^2, $i = 1, 2$. Έτσι, αν ένα οστό έχει πραγματική ηλικία ℓ, τότε αν μετρήσουμε την ηλικία του με την μέθοδο $i \in \{1, 2\}$ θα πάρουμε σαν αποτέλεσμα $L_i = \ell + X_i$, $X_i \sim \mathcal{N}(0, \sigma_i^2)$. Υποθέτουμε επίσης ότι τα σφάλματα σε διαφορετικές μετρήσεις είναι ανεξάρτητα. Ένας ερευνητής θέλει να ελέγξει την υπόθεση ότι οι δύο μέθοδοι είναι εξίσου ακριβείς, δηλαδή $H_0 : \sigma_1^2 = \sigma_2^2$, έναντι της εναλλακτικής $H_1 : \sigma_1^2 \neq \sigma_2^2$, με επίπεδο σημαντικότητας $\alpha = 0,05$. Για το σκοπό αυτό τεμαχίζει ένα οστό σε 12 μέρη και χρονολογεί 6 από αυτά με τη μία μέθοδο και 6 με την άλλη. Τα αποτελέσματα των μετρήσεων του σε έτη είναι

Μέθοδος 1	1224	1295	1264	1283	1276	1314
Μέθοδος 2	1057	1426	1284	1360	1338	1233

Ποιο θα είναι το αποτέλεσμα του ελέγχου βάσει των παραπάνω μετρήσεων;

ΤΥΠΟΛΟΓΙΟ

Συνάρτηση Γάμμα: $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt, \quad x > 0.$

Κατανομή Γάμμα $G(\alpha, p)$: σ.π.π. $f(x) = \frac{\alpha^p}{\Gamma(p)}x^{p-1}e^{-\alpha x}, \quad x > 0.$

Κανονική κατανομή $\mathcal{N}(\mu, \sigma^2)$: σ.π.π. $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, σ.κ.π. $F(x) = \int_{-\infty}^x f(x)\,dx$

Διωνυμική κατανομή $b(n, q)$: σ.μ.π. $p(k) = \binom{n}{k}q^k(1-q)^{n-k}$.

Χρήσιμες ιδιότητες

- $\Gamma(x+1) = x\Gamma(x), \quad \Gamma(1) = 1, \ \Gamma(\frac{1}{2}) = \sqrt{\pi}.$
- Αν $X_i \sim G(\alpha, p_i)$ ανεξάρτητες, τότε $\sum X_i \sim G(\alpha, \sum p_i)$.
- $\mathcal{E}xp(\alpha) = G(\alpha, 1)$.
- Αν $X_i \sim \mathcal{N}(0,1)$ ανεξάρτητες, τότε $\displaystyle\sum_{i=1}^n X_i^2 \sim \chi^2(n) = G(\frac{1}{2}, \frac{n}{2})$.
- Αν $X_i \sim \mathcal{N}(\mu, \sigma^2)$ ανεξάρτητες, τότε $\sigma^{-2}(n-1)S^2 = \sigma^{-2}\sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1)$ και ανεξάρτητη της $\bar{X}$.
- Αν $X \sim \mathcal{N}(0,1)$, $nY \sim \chi^2(n)$ και X, Y ανεξάρτητες τότε $\frac{X}{\sqrt{Y}} \sim t(n)$ (Student).
- Αν $nX \sim \chi^2(n)$, $mY \sim \chi^2(m)$ και X, Y ανεξάρτητες τότε $\frac{X}{Y} \sim F(n, m)$ (Snedecor.)

Στατιστικά για τον έλεγχο μέσου κανονικών πληθυσμών: Γνωστή διασπορά $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ Άγνωστη διασπορά $\frac{\bar{X}-\mu}{S/\sqrt{n}}$

Σημεία x όπου η σ.κ.π. φτάνει μια κρίσιμη τιμή

Κατανομή	$F(x) = 0,95$	$F(x) = 0,975$	$F(x) = 0,99$	$F(x) = 0,995$
$\chi^2(11)$	19,675	21,92	24,725	26,757
$\chi^2(12)$	21,026	23,337	26,217	28,3
$F(5,5)$	5,0503	7,1464	10,967	14,94
$F(6,6)$	4,2839	5,8198	8,4661	11,073
$\mathcal{N}(0,1)$	1,6449	1,96	2,3263	2,5758

Τιμές $F(x)$ της τυπικής κανονικής σ.κ.π. για διάφορες τιμές του x:

x	0	0,9714	0,9857	1	1,96	2	3	3,4157	17,5
$F(x)$	0,5	0,8343	0,8379	0,8413	0,975	0,9772	0,9887	0,9997	≈ 1

Διάρκεια Εξέτασης 2 ώρες και 30 λεπτά

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!