Πανεπιστημιο Κρητης, Τμ. Μαθηματικων και Εφαρμοσμενων Μαθηματικων
Παραμετρικη Στατιστικη, Φθινοπωρο 2014-15, Μιχ. Κολουντζακης

Τελικό Διαγώνισμα

Πέμπτη 15 Ιαν. 2015

Πρόβλημα 1. Η ΤΜ (X, Y) είναι ομοιόμορφα κατανεμημένη στο εσωτερικό της κλειστής πολυγωνικής γραμμής που έχει ως κορυφές τα σημεία του επιπέδου

$$(0,0),(2a,0),(2a,2a),(a,2a),(a,a),(0,a),(0,0)$$

(με τη σειρά που δίνονται). Βρείτε μια συνάρτηση των X, Y που να είναι αμερόληπτη εκτιμήτρια της παραμέτρου a.
Λύση: Η πυκνότητα του ζεύγους $f(x, y)$ είναι σταθερή μέσα στο πολύγωνο και 0 εκτός αυτού. Η σταθερά ισούται με $1/(3a^2)$ αφού το εμβαδό του πολυγώνου είναι $3a^2$. Θα υπολογίσουμε τη $\mathbb{E}[X]$. Γι' αυτό χρειαζόμαστε την (περιθώρια) πυκνότητα της X που δίνεται από το ολοκλήρωμα

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy.$$

Για $x \in \mathbb{R} \setminus (0, 2a)$ το ολοκλήρωμα αυτό είναι 0 ενώ για $0 \le x < a$ ισούται με $1/(3a)$ και για $a \le x \le 2a$ ισούται με $2/(3a)$. Άρα

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^a x/(3a)\,dx + \int_a^{2a} 2x/(3a)\,dx = 7a/6.$$

Άρα η ΤΜ $\frac{6}{7}X$ έχει μέση τιμή a και είναι συνεπώς αμερόληπτη εκτιμήτρια της παραμέτρου a.

Πρόβλημα 2. Οι ΤΜ $X_1, X_2, \ldots, X_N$ είναι ανεξάρτητες και ακολουθούν την κατανομή Poisson(λ). Βρείτε τη συνάρτηση πιθανοφάνειας του λ δεδομένων των μετρήσεων $X_1, \ldots, X_N$ και βρείτε ποια τιμή του λ μεγιστοποιεί αυτή τη συνάρτηση.
Λύση: Η πυκνότητα της διακριτής κατανομής Poisson(λ) είναι η $f(n) = e^{-\lambda}\lambda^n/n!$ (για $n = 0, 1, 2, \ldots$) και άρα η συνάρτηση πιθανοφάνειας του λ είναι η

$$L(\lambda) = \prod_{i=1}^{N} f(X_i) = \prod_{i=1}^{n} e^{-\lambda}\lambda^{X_i}/X_i! = e^{-N\lambda}\lambda^{\sum_{i=1}^N X_i}\prod_{i=1}^{N}\frac{1}{X_i!}.$$

Για να τη μεγιστοποιήσουμε αρκεί να μεγιστοποιήσουμε το λογάριθμό της

$$\log L(\lambda) = -N\lambda + \sum_{i=1}^{N} X_i \log\lambda + \log\prod_{i=1}^{N}\frac{1}{X_i!}.$$

Παίρνοντας παράγωγο έχουμε

$$(\log L(\lambda))' = -N + \sum_{i=1}^{N} X_i \frac{1}{\lambda}$$

που μηδενίζεται όταν $\lambda = \frac{1}{N}\sum_{i=1}^N X_i$.

Πρόβλημα 3. Οι ΤΜ $X_1, X_2, \ldots, X_N$ είναι ανεξάρτητες και ακολουθούν κανονική κατανομή $\mathcal{N}(\mu, \sigma = 10)$. Αν ο δειγματικός μέσος $\overline{X}$ των $X_1, X_2, \ldots, X_N$ προκύψει να είναι ίσος με 1000, βρείτε ένα διάστημα εμπιστοσύνης 90% για το μ. Στην απάντησή σας, εκτός από το N, μπορείτε να χρησιμοποιήσετε και τη συνάρτηση $\Phi(\cdot)$ (συνάρτηση κατανομής της τυπικής κανονικής $\mathcal{N}(0,1)$) και την αντίστροφή της $\Phi^{-1}(\cdot)$.
Λύση: Η ΤΜ $\overline{X} = (X_1 + \cdots + X_N)/N$ (δειγματικός μέσος) έχει κατανομή $\mathcal{N}(\mu, \sigma^2/N)$ και άρα η ΤΜ

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{N}}$$

ακολουθεί τυπική κανονική κατανομή και άρα για κάθε $L > 0$ έχουμε

$$\mathbb{P}\left[\left|\overline{X}-\mu\right| > L\right] = \mathbb{P}\left[\frac{|\overline{X}-\mu|}{\sigma/\sqrt{N}} > \frac{L}{\sigma/\sqrt{N}}\right] = \Phi(-\frac{L}{\sigma/\sqrt{N}}) + 1 - \Phi(\frac{L}{\sigma/\sqrt{N}}).$$

Για να είναι αυτή η πιθανότητα ίση με $a = 1 - 0.90 = 0.10$ αρκεί

$$\Phi(-\frac{L}{\sigma/\sqrt{N}}) = 1 - \Phi(\frac{L}{\sigma/\sqrt{N}}) = a/2 = 0.05,$$

ισότητα που συνεπάγεται (παίρνοντας το Φ^{-1} και των δύο μελών) ότι

$$\frac{L}{\sigma/\sqrt{N}} = -\Phi^{-1}(0.05),$$

το οποίο μας προσδιορίζει $L = -\Phi^{-1}(0.05)\frac{\sigma}{\sqrt{N}} = -\Phi^{-1}(0.05)\frac{10}{\sqrt{N}}$. Το διάστημα εμπιστοσύνης είναι συνεπώς το

$$(1000 + \Phi^{-1}(0.05)\frac{10}{\sqrt{N}}, 1000 - \Phi^{-1}(0.05)\frac{10}{\sqrt{N}}).$$

(Προσέξτε ότι η ποσότητα $\Phi^{-1}(0.05)$ είναι αρνητική.)

Πρόβλημα 4. Δύο φυσικά μεγέθη X και Y ικανοποιούν ένα φυσικό νόμο της μορφής $Y = \lambda X^2$, όπου λ είναι μια άγνωστη σταθερά. Κάνουμε N μετρήσεις των φυσικών αυτών μεγεθών από τις οποίες προκύπτουν τα ζεύγη τιμών $(x_1, y_1), \ldots, (x_N, y_N)$ (οι μετρήσεις έχουν και σφάλματα και άρα οι τιμές αυτές δεν ικανοποιούν ακριβώς τις εξισώσεις $y_i = \lambda x_i^2$). Δείξτε πώς από τις μετρήσεις αυτές θα βρείτε μια εκτίμηση για το λ η οποία να ελαχιστοποιεί το άθροισμα των τετραγώνων των σφαλμάτων $\epsilon_i = y_i - \lambda x_i^2$.
Λύση: Το προς ελαχιστοποίηση άθροισμα τετραγώνων είναι η συνάρτηση του λ

$$\sum_{i=1}^{N}(y_i - \lambda x_i^2)^2$$

της οποίας η παράγωγος ως προς λ είναι η

$$\sum_{i=1}^{N} -2x_i^2(y_i - \lambda x_i^2) = 2(\sum_{i=1}^{N} x_i^4\lambda - \sum_{i=1}^{N} y_i x_i^2).$$

Αυτή η μηδενίζεται όταν $\lambda = \frac{\sum_{i=1}^N y_i x_i^2}{\sum_{i=1}^N x_i^4}$ που είναι η ζητούμενη τιμή.

Πρόβλημα 5. (α) Ορίστε την κατανομή χ_n^2.
(β) Αποδείξτε ότι αν οι ΤΜ X, Y είναι ανεξάρτητες και $X \sim \chi_m^2$, $Y \sim \chi_n^2$ τότε $X + Y \sim \chi_{m+n}^2$.
Λύση: (α) Η κατανομή χ_n^2 ορίζεται να είναι η κατανομή του αθροίσματος των τετραγώνων n ανεξαρτήτων τυπικών κανονικών ΤΜ.

(β) Η κατανομή της $X + Y$ εξαρτάται μόνο από το ποιες είναι οι κατανομές των X και Y και το γεγονός ότι αυτές είναι ανεξάρτητες. Με άλλα λόγια, μπορούμε να επιλέξουμε όποιες ΤΜ X και Y θέλουμε αρκεί αυτές να ικανοποιούν τις προϋποθέσεις του (β). Ας είναι λοιπόν

$$w_1, w_2, \ldots, w_m, w_{m+1}, w_{m+2}, \ldots, w_{m+n}$$

ανεξάρτητες τυπικές κανονικές ΤΜ και ορίζουμε

$$X = w_1^2 + w_2^2 + \cdots + w_m^2, \quad Y = w_{m+1}^2 + \cdots + w_{m+n}^2.$$

Προκύπτει ότι $X \sim \chi_m^2$ και $Y \sim \chi_n^2$ και ότι οι δύο ΤΜ X, Y είναι ανεξάρτητες αφού εξαρτώνται από διαφορετικές w_i η κάθε μία. Ικανοποιούνται λοιπόν οι προϋποθέσεις του (β) και άρα αρκεί να δείξουμε ότι η $X + Y$ ακολουθεί κατανομή χ_{m+n}^2. Όμως

$$X + Y = w_1^2 + w_2^2 + \cdots + w_{m+n}^2$$

και άρα από το (α) η $X + Y$ ακολουθεί χ_{m+n}^2.

Το (β) αποδεικνύεται επίσης και με χρήση ροπογεννητριών συναρτήσεων.

Όλες οι σημειώσεις πρέπει να είναι κλειστές. Οι αιτιολογήσεις σας να είναι πλήρεις και καθαρές. Απαντήσεις χωρίς πλήρη αιτιολόγηση δε παίρνουν μονάδες. Δεν επιτρέπεται η χρήση υπολογιστικών μέσων. **Διάρκεια 2 ώρες.**