

Προσομοίωση - Εξέταση 12/7/2022 ΠΜΣ ``Βιοστατιστική''

ΘΕΜΑ 1. Έστω X, Y συνεχείς τυχαίες μεταβλητές, όπου η X έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x) = x^2/9, \quad 0 \leq x \leq 3$$

και η δεσμευμένη κατανομή της Y δοθέντος ότι $X = x$ είναι κανονική με μέση τιμή $3x$ και τυπική απόκλιση $2\sqrt{x}$.
Να εκτιμήσετε την πιθανότητα $P(X > 2, Y > 5)$ και την μέση υπέρβαση $E(\max(Y - 5, 0))$, μαζί με διαστήματα εμπιστοσύνης 95% για κάθε περίπτωση, χρησιμοποιώντας 1000 σενάρια προσομοίωσης.

ΘΕΜΑ 2. Έστω ο παρακάτω κανόνας για τη διάγνωση Covid μέσω συνδυασμού rapid test και PCR test: Ένα άτομο κάνει 3 rapid test ταυτόχρονα και ανεξάρτητα μεταξύ τους. Αν προκύψουν 2 ή 3 θετικά αποτελέσματα γίνεται διάγνωση ότι το άτομο έχει Covid. Αν προκύψει ακριβώς 1 θετικό αποτέλεσμα, το άτομο κάνει PCR τεστ και η διάγνωση είναι το αποτέλεσμα του PCR. Αν και τα τρία rapid test είναι αρνητικά, τότε γίνεται διάγνωση ότι το άτομο δεν έχει Covid. Υποθέτουμε ότι η θετικότητα του ιού στον πληθυσμό είναι p , δηλαδή η πιθανότητα ένα τυχαίο άτομο να έχει Covid είναι ίση με p . Επιπλέον το rapid test έχει πιθανότητα p_1 να βγει θετικό αν το άτομο είναι μολυσμένο και πιθανότητα p_0 να βγει θετικό αν το άτομο είναι υγιές. Το PCR υποθέτουμε ότι είναι 100% ακριβές.

(α) Να δημιουργήσετε μια συνάρτηση που παίρνει ως ορίσματα τις πιθανότητες p, p_0, p_1 και ένα αριθμό N , και προσομοιώνει την παραπάνω διαδικασία σε N άτομα αυτού του πληθυσμού. Η συνάρτηση επιστρέφει μια σημειακή εκτίμηση για την πιθανότητα λάθος διάγνωσης (οποιοδήποτε τύπου).

(β) Έστω $N = 10000, p_0 = 0.05, p_1 = 0.80$. Να εκτιμήσετε την πιθανότητα λάθος διάγνωσης συναρτήσει του ποσοστού θετικότητας p , για τιμές του $p = 0.01, 0.02, \dots, 0.20$ και να δημιουργήσετε ένα γράφημα με την εκτίμηση της πιθανότητας λάθους συναρτήσει του p .

ΘΕΜΑ 3. Έστω X τυχαία μεταβλητή που ακολουθεί ομοιόμορφη κατανομή στο διάστημα $[0, a]$ με άγνωστο a . Για την εκτίμηση του a λαμβάνεται δείγμα $n = 10$ παρατηρήσεων και υπολογίζεται η εκτιμήτρια $\hat{a} = \max(x_1, \dots, x_n)$. Να εκτιμηθεί η διασπορά και το μέσο τετραγωνικό σφάλμα αυτής της εκτιμήτριας $MSE = E[(\hat{a} - a)^2]$ αν η αληθινή τιμή της παραμέτρου a είναι ίση με 5, χρησιμοποιώντας 1000 σενάρια προσομοίωσης.

Σημείωση: Σε κάθε σενάριο θα πρέπει να δημιουργηθεί δείγμα 10 παρατηρήσεων από την πραγματική κατανομή, να υπολογιστεί η εκτιμήτρια και η απόκλιση της από την αληθή τιμή.