

**ΕΡΩΤΗΣΕΙΣ ΕΠΑΝΑΛΗΨΗΣ ΥΛΗΣ
ΓΙΑ ΤΗΝ ΕΞΕΤΑΣΤΙΚΗ ΙΟΥΝΙΟΥ 2008**

ΠΡΟΣΟΧΗ! Οι ερωτήσεις αυτές δίνονται για να σας βοηθήσουν στην επανάληψη του βασικού μέρους της ύλης που καλύφθηκε στην τάξη κατά την διάρκεια του εξαμήνου. Δεν αντιστοιχούν σε ερωτήσεις της εξέτασης και δεν καλύπτουν πιθανές ασκήσεις της εξέτασης!

1. Υπάρχουν τρία βασικά είδη στατιστικών υποθέσεων. Ονομάστε το κάθε είδος και εξηγήστε με ποιο τρόπο το χρησιμοποιούμε στην διαδικασία ενός στατιστικού ελέγχου.
2. Εξηγήστε την έννοια της στατιστικής ελέγχου και τον τρόπο με τον οποίο την χρησιμοποιούμε.
3. Κατά την διαδικασία της επαγωγής μέσω ενός στατιστικού ελέγχου υπάρχει πάντα η πιθανότητα να κάνουμε σφάλμα. Εξηγήστε την έννοια και περιγράψτε τα είδη των στατιστικών σφαλμάτων που μπορούμε να κάνουμε καθώς και τις αντίστοιχες πιθανότητες για το κάθε σφάλμα. Παρουσιάστε τον πίνακα (2×2) που συνδέει μια αληθή/ψευδή μηδενική υπόθεση με το αποτέλεσμα ενός στατιστικού ελέγχου. Τι θα θέλαμε ιδανικά να είχαμε (σχετικά με τα σφάλματα) κατά την διαδικασία ενός στατιστικού ελέγχου;
4. Έστω ότι έχετε ένα τυχαίο δείγμα $\{X_i\}_{i=1}^N$ όπου η κάθε τυχαία μεταβλητή ακολουθεί την κανονική κατανομή $X_i \sim N(\mu, \sigma^2)$ όπου ο μέσος μ της κατανομής είναι άγνωστος και η διακύμανση της κατανομής σ^2 είναι γνωστή. Υποθέστε ότι το δείγμα είναι μεγάλο $N \rightarrow \infty$. Εξηγήστε με λεπτομέρειες πως θα ελέγξετε την μηδενική υπόθεση $H_0 : \mu = \mu_0$ όπου μ_0 είναι μια γνωστή τιμή. Απαραίτητα στοιχεία στην απάντησή σας είναι να αναφέρετε/εξηγήσετε: τις διατηρήσιμες υποθέσεις, την εναλλακτική υπόθεση, την στατιστική ελέγχου και την δειγματική κατανομή της στατιστικής ελέγχου και τον τρόπο με τον οποίο παίρνετε την απόφαση για απόρριψη ή μη της μηδενικής υπόθεσης.
5. Με βάση την απάντησή σας στην ανωτέρω ερώτηση εξηγήστε με λεπτομέρειες την έννοια της δύναμης ενός στατιστικού ελέγχου και (α) δείξτε πως σχετίζεται με το επίπεδο σημαντικότητας του ελέγχου (β) δείξτε πως δεν μπορούμε να ελαχιστοποιήσουμε την πιθανότητα και των δύο ειδών σφαλμάτων ταυτόχρονα (γ) δείξτε πως σχετίζεται η δύναμη του ελέγχου με το μέγεθος του δείγματος.
6. Έστω ότι έχετε ένα τυχαίο δείγμα $\{X_i\}_{i=1}^N$ όπου η κάθε τυχαία μεταβλητή ακολουθεί την κανονική κατανομή $X_i \sim N(\mu, \sigma^2)$ όπου ο μέσος μ της κατανομής είναι άγνωστος και η διακύμανση της κατανομής σ^2 είναι επίσης άγνωστη. Υποθέστε ότι το δείγμα είναι μικρό $N \ll \infty$. Εξηγήστε με λεπτομέρειες πως θα ελέγξετε την μηδενική υπόθεση $H_0 : \mu = \mu_0$ όπου μ_0 είναι

μια γνωστή τιμή. Απαραίτητα στοιχεία στην απάντησή σας είναι να αναφέρετε/εξηγήσετε: τις διατηρήσιμες υποθέσεις, την εναλλακτική υπόθεση, την στατιστική ελέγχου και την δειγματική κατανομή της στατιστικής ελέγχου και τον τρόπο με τον οποίο παίρνετε την απόφαση για απόρριψη ή μη της μηδενικής υπόθεσης.

7. Έστω ότι έχετε ένα τυχαίο δείγμα $\{X_i\}_{i=1}^N$ όπου η κάθε τυχαία μεταβλητή ακολουθεί την κατανομή $X_i \sim \text{Bernoulli}(p)$ όπου η παράμετρος p της κατανομής είναι άγνωστη. Πρώτα υποθέστε ότι το δείγμα είναι μεγάλο $N \rightarrow \infty$ και μετά ότι το δείγμα είναι μικρό $N \ll$. Εξηγήστε με λεπτομέρειες πως θα ελέγξετε την μηδενική υπόθεση $H_0 : p = p_0$ όπου p_0 είναι μια γνωστή τιμή. Απαραίτητα στοιχεία στην απάντησή σας είναι να αναφέρετε/εξηγήσετε: τις διατηρήσιμες υποθέσεις, την εναλλακτική υπόθεση, την στατιστική ελέγχου και την δειγματική κατανομή της στατιστικής ελέγχου και τον τρόπο με τον οποίο παίρνετε την απόφαση για απόρριψη ή μη της μηδενικής υπόθεσης.
8. Έστω ότι έχετε δύο ανεξάρτητα τυχαία δείγματα $\{X_i\}_{i=1}^N$ και $\{Y_j\}_{j=1}^N$ που ακολουθούν κανονικές κατανομές $X_i \sim N(\mu_X, \sigma_X^2)$ και $Y_j \sim N(\mu_Y, \sigma_Y^2)$ με άγνωστους μέσους και άγνωστες διακυμάνσεις. Υποθέστε ότι το δείγμα είναι μεγάλο $N \rightarrow \infty$ και ότι οι διακυμάνσεις είναι διαφορετικές $\sigma_X^2 \neq \sigma_Y^2$. Εξηγήστε με λεπτομέρειες πως θα ελέγξετε την μηδενική υπόθεση $H_0 : \mu_X = \mu_Y$. Απαραίτητα στοιχεία στην απάντησή σας είναι να αναφέρετε/εξηγήσετε: τις διατηρήσιμες υποθέσεις, την εναλλακτική υπόθεση, την στατιστική ελέγχου και την δειγματική κατανομή της στατιστικής ελέγχου και τον τρόπο με τον οποίο παίρνετε την απόφαση για απόρριψη ή μη της μηδενικής υπόθεσης.
9. Σε συνέχεια της ανωτέρω ερώτησης υποθέστε τώρα ότι επιθυμείτε να ελέγξετε την ισότητα των διακυμάνσεων, δηλαδή την υπόθεση $H_0 : \sigma_X^2 = \sigma_Y^2$. Εξηγήστε.
10. Σε συνέχεια και των δύο παραπάνω ερωτήσεων υποθέστε τώρα ότι έχετε αποδεχθεί την υπόθεση των ίσων διακυμάνσεων και επιθυμείτε να ελέγξετε ξανά την υπόθεση της ισότητας των δύο μέσων $H_0 : \mu_X = \mu_Y$. Τώρα όμως υποθέστε ότι το δείγμα σας είναι μικρό $N \ll$. Εξηγήστε, δίνοντας έμφαση στις διαφορές της απάντησή σας σε σχέση με την ερώτηση 8.
11. Η χρήση της κανονικής κατανομής στην διαδικασία στατιστικών ελέγχων είναι σημαντική, ιδιαίτερα σε μικρά δείγματα. Είναι επομένως χρήσιμο να έχουμε τρόπους να ελέγχουμε εάν τα δεδομένα μας είναι συμβατά με την κανονική κατανομή. Εξηγήστε με λεπτομέρειες την χρήση των δειγματικών ροπών ασυμμετρίας και κύρτωσης για τον έλεγχο της υπόθεσης της κανονικότητας των δεδομένων σας.
12. Υποθέστε ότι έχετε χρονολογικά δεδομένα και επιθυμείτε να ελέγξετε την υπόθεση της μηδε-

νικής αυτοσυσχέτισης των δεδομένων σας. Για τον σκοπό αυτό μπορείτε να χρησιμοποιήσετε αποτελέσματα που αναφέρονται στην συνάρτηση αυτοσυσχέτισης. Εξηγήστε με λεπτομέρειες (α) την έννοια και τύπο της συνάρτησης αυτοσυσχέτισης (β) την δειγματική κατανομή της συνάρτησης αυτοσυσχέτισης κάτω από την υπόθεση της μηδενικής αυτοσυσχέτισης (γ) τις μηδενικές υποθέσεις, τις στατιστικές ελέγχου που γνωρίζετε και τις αντίστοιχες δειγματικές κατανομές τους.

13. Έστω ότι έχετε ένα τυχαίο δείγμα από πολυμεταβλητά δεδομένα $\{\mathbf{X}_i\}_{i=1}^N$, όπου \mathbf{X}_i είναι ένα διάνυσμα μεγέθους k δηλαδή $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{ik})$, που ακολουθούν την πολυμεταβλητή κανονική κατανομή $\mathbf{X}_i \sim MN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ με διάνυσμα μέσων $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k)$ και πίνακα διακυμάνσεων-συνδιακυμάνσεων $\boldsymbol{\Sigma}$. Εξηγήστε πως θα εκτιμήσετε το διάνυσμα των μέσων και των πίνακα διακυμάνσεων-συνδιακυμάνσεων. Μετά υποθέστε ότι $k = 2$ και έστω ότι επιθυμείτε να ελέγξετε την υπόθεση της ισότητας των μέσων $H_0 : \mu_1 = \mu_2$. Εξηγήστε με λεπτομέρειες τον τρόπο που θα κάνετε τον έλεγχο αυτό και σχολιάστε τις διαφορές με τον αντίστοιχο έλεγχο των ερωτήσεων 8 και 10.
14. Έστω ότι έχετε ένα τυχαίο δείγμα για τις μεταβλητές X και Y , $\{y_i, x_i\}_{i=1}^N$ και το υπόδειγμα της απλής παλινδρόμησης $y_i = \alpha + \beta x_i + \epsilon_i$ και έστω ότι επιθυμείτε να ελέγξετε την υπόθεση $H_0 : \beta = 0$. Εξηγήστε με λεπτομέρειες πως θα κάνετε αυτόν τον έλεγχο στηριζόμενοι στις υποθέσεις του υποδείγματος σχετικά με τον όρο του σφάλματος ϵ_i και τους αντίστοιχους εκτιμητές των παραμέτρων $\hat{\alpha}$ και $\hat{\beta}$.