Εργαστήριο Μαθηματικών & Στατιστικής

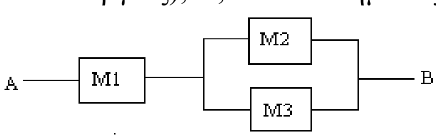
1η Πρόοδος στο Μάθημα <u>Στατιστική</u> **3/12/2010**

<u>Α΄ ΣΕΙΡΑ ΘΕΜΑΤΩΝ</u>

1ο Θέμα [25] Το τμήμα ποιοτικού ελέγχου ενός εργοστασίου παραγωγής κομπόστας ροδάκινου, ελέγχει, μεταξύ άλλων, το μήκος της διαμέτρου των ροδάκινων που προμηθεύεται από τους παραγωγούς. Έστω X τυχαία μεταβλητή που εκφράζει το μήκος της διαμέτρου των ροδάκινων που προμηθεύεται το εργοστάσιο από τον παραγωγό *Α*. Το τμήμα ποιοτικού ελέγχου θεωρεί (από εμπειρία ετών) ότι η μέση τιμή της X είναι $60mm$ και η τυπική απόκλισή της $10mm$. **(Ι)** Το τμήμα ποιοτικού ελέγχου παίρνει ένα τυχαίο δείγμα 30 ροδάκινων από την παραγωγή του προμηθευτή *Α* και μετράει το μήκος της διαμέτρου τους. **α)** Ποια είναι η πιθανότητα, η μέση τιμή του μήκους της διαμέτρου των 30 ροδάκινων να ξεπερνά τα $65mm$. **β)** Αν η μέση τιμή του μήκους της διαμέτρου των 30 ροδάκινων βρεθεί ίση με $66mm$, μήπως το τμήμα ποιοτικού ελέγχου πρέπει πλέον να αμφιβάλλει για κάτι; **(ΙΙ)** Μπορούμε να απαντήσουμε στο ερώτημα, *«ποια είναι η πιθανότητα, ένα τυχαία επιλεγμένο ροδάκινο από την παραγωγή του προμηθευτή Α να έχει μήκος διαμέτρου που βρίσκεται μεταξύ 52 και 61mm»*; Εξηγείστε.

2ο Θέμα [40] Μια βιομηχανία τροφίμων παράγει, μεταξύ άλλων, ελαφρά συμπυκνωμένο χυμό τομάτας σε συσκευασίες των $200gr$. Σύμφωνα με τις προδιαγραφές παραγωγής, η ποσότητα λυκοπενίου (πρόκειται για φυτοχημική ουσία, ευεργετική για την υγεία) που περιέχεται σε κάθε τέτοια συσκευασία είναι τυχαία μεταβλητή, έστω X, και ακολουθεί κανονική κατανομή με μέση τιμή $24mg$ και τυπική απόκλιση $2mg$. **α)** Επιλέγουμε τυχαία από την παραγωγή του εργοστασίου μια τέτοια συσκευασία. Ποια είναι η πιθανότητα, η ποσότητα λυκοπενίου που περιέχει να βρίσκεται μεταξύ 23 και $25mg$; **β)** Επιλέγουμε τυχαία από την παραγωγή του εργοστασίου 5 τέτοιες συσκευασίες. Ποια είναι η πιθανότητα, το πολύ μία συσκευασία να περιέχει ποσότητα λυκοπενίου που βρίσκεται μεταξύ 23 και $25mg$; **γ)** Βρείτε εκείνη την ποσότητα λυκοπενίου για την οποία ισχύει ότι: μεγαλύτερη ποσότητα λυκοπενίου από αυτή, περιέχει μόνο το 2% της παραγωγής του εργοστασίου. **δ)** Βρείτε ένα διάστημα συμμετρικό ως προς τη μέση ποσότητα λυκοπενίου που περιέχεται στις συσκευασίες ελαφρά συμπυκνωμένου χυμού του εργοστασίου, τέτοιο ώστε, το 99% της παραγωγής να περιέχει ποσότητα λυκοπενίου που βρίσκεται εντός αυτού του διαστήματος.

3ο Θέμα [35] α) Οι 3 μονάδες Μ1, Μ2, Μ3 του παρακάτω συστήματος, λειτουργούν ανεξάρτητα η μια από την άλλη και έχουν αξιοπιστία (πιθανότητα λειτουργίας) 0.9, 0.6 και 0.8 αντίστοιχα. Το σύστημα λειτουργεί αν και μόνο αν υπάρχει ένα τουλάχιστον «μονοπάτι» από το Α στο Β μέσω λειτουργούντων μονάδων. Να βρεθεί η αξιοπιστία (η πιθανότητα λειτουργίας), R, του συστήματος.



β) Αν *Α, Β* είναι δύο ενδεχόμενα του δειγματικού χώρου *Ω* ενός πειράματος τύχης με $P(A)=0.42$ και $P(B)=0.38$, τότε, για την τιμή της πιθανότητας $P(AB)$ ποια από τις παρακάτω απαντήσεις είναι η σωστή: i) 0.80 ii) 0.1596 iii) οι πληροφορίες που δόθηκαν, δεν αρκούν για να υπολογισθεί. **γ)** Είναι λογικό η μέση τιμή της κατανομής *Poisson* να είναι ίση με τη διασπορά της; Εξηγείστε. **δ)** Μπορεί η συνάρτηση κατανομής F μιας συνεχούς τυχαίας μεταβλητής X, να πάρει τιμές μεγαλύτερες του 1; Η συνάρτηση πυκνότητας f; Εξηγείστε. **ε)** Να αναφέρετε τρεις πηγές ***μη δειγματοληπτικών*** *σφαλμάτων*.

<u>Ενδεικτικές απαντήσεις (Α΄ Σειρά)</u>

1ο Θέμα: (I). α) Έστω X_i το μήκος της διαμέτρου του ροδάκινου i ($i = 1,2,...,30$). Προφανώς ζητάμε την πιθανότητα $P(\overline{X} > 65)$ όπου $\overline{X} = \frac{X_1 + X_2 + ... + X_{30}}{30}$. Οι τυχαίες μεταβλητές X_i είναι ανεξάρτητες και ισόνομες με μέση τιμή $\mu_i = E(X_i) = 60$ και διασπορά $\sigma_i^2 = 10^2$ και επομένως από το Κ.Ο.Θ. προκύπτει ότι η $\overline{X}$ προσεγγίζεται ικανοποιητικά από την κανονική κατανομή $N(60, \frac{10^2}{30})$ ή $N(60, 1.83^2)$. Επομένως η ζητούμενη πιθανότητα είναι:

$$P(\overline{X} > 65) = P(Z > \frac{65-60}{1.83}) = P(Z > 2.73) = 1 - \Phi(2.73) = 0.0032.$$

β) Επειδή η πιθανότητα να εμφανισθεί η τιμή αυτή στο δείγμα είναι (με βάση την απάντηση στο (α) ερώτημα) πολύ μικρή και παρόλα αυτά εμφανίσθηκε, είναι πολύ πιθανό, η μέση τιμή της διαμέτρου των ροδάκινων στον πληθυσμό να μην είναι 60mm αλλά μεγαλύτερη. **γ)** Όχι γιατί δεν γνωρίζουμε την κατανομή της X (γνωρίζουμε μόνο τη μέση τιμή και τη διασπορά της).

2ο Θέμα: $X \sim N(24, 2^2)$ άρα:

α) $P(23 < X < 25) = P(\frac{23-24}{2} < \frac{X-24}{2} < \frac{25-24}{2}) = P(-0.5 < Z < 0.5) =$
$2\Phi(0.5) - 1 = 0.383$.

β) Έστω Y ο αριθμός των συσκευασιών από τις πέντε που επελέγησαν, που η κάθε μία περιέχει ποσότητα λυκοπενίου μεταξύ 23 και 25mg. Προφανώς, $Y \sim B(5,\ 0.383)$ άρα

$$P(Y \le 1) = P(Y=0) + P(Y=1) = \binom{5}{0} \cdot (0.383)^0 \cdot (0.617)^5 + \binom{5}{1} \cdot (0.383)^1 \cdot (0.617)^4 = ...$$

γ) Πρέπει: $P(X > \xi) = 0.02 \Leftrightarrow P(z > \frac{\xi - 24}{2}) = 0.02 \Leftrightarrow \Phi(\frac{\xi-24}{2}) = 0.98 \Leftrightarrow$
$\Leftrightarrow \frac{\xi - 24}{2} = 2.06 \Leftrightarrow \xi = 28.12$.

δ) Πρέπει: $P(24 - c < X < 24 + c) = 0.99 \Leftrightarrow P(\frac{-c}{2} < Z < \frac{c}{2}) = 0.99 \Leftrightarrow$
$\Leftrightarrow 2 \cdot \Phi(\frac{c}{2}) - 1 = 0.99 \Leftrightarrow \Phi(\frac{c}{2}) = 0.995 \Leftrightarrow \frac{c}{2} = 2.58 \Leftrightarrow c = 5.16$ άρα το ζητούμενο διάστημα είναι: (18.84, 29.16)

3ο Θέμα: α) Έστω M_i το ενδεχόμενο: η μονάδα i λειτουργεί ($i = 1,2,3$). Δίνονται οι πιθανότητες: $P(M_1) = 0.9$, $P(M_2) = 0.6$ και $P(M_3) = 0.8$. Από τη συνθήκη λειτουργίας του συστήματος και επειδή οι μονάδες λειτουργούν ανεξάρτητα η μία από την άλλη έχουμε:

$$R = P(M_1) \cdot P(M_2 \cup M_3) = P(M_1) \cdot [1 - P[(M_2 \cup M_3)']] =$$
$$= P(M_1) \cdot [1 - P(M_2' \cdot M_3')] = P(M_1) \cdot [1 - P(M_2') \cdot P(M_3')] = 0.9 \cdot [1 - 0.4 \cdot 0.2] = 0.828.$$

β) η απάντηση *(iii)*.

γ), δ) και **ε).** Υπάρχουν στις απαντήσεις άλλων εξεταστικών περιόδων.

Εργαστήριο Μαθηματικών & Στατιστικής

1η Πρόοδος στο Μάθημα <u>Στατιστική</u>

3/12/2010

<u>Β΄ ΣΕΙΡΑ ΘΕΜΑΤΩΝ</u>

1ο Θέμα [25] Μια βιομηχανία τροφίμων παράγει, μεταξύ άλλων, ελαφρά συμπυκνωμένο χυμό τομάτας σε συσκευασίες των 200*gr*. Σύμφωνα με τις προδιαγραφές παραγωγής, η ποσότητα λυκοπενίου (πρόκειται για φυτοχημική ουσία, ευεργετική για την υγεία) που περιέχεται σε κάθε τέτοια συσκευασία είναι τυχαία μεταβλητή, έστω *X*, με μέση τιμή 24*mg* και τυπική απόκλιση 2*mg*. **(Ι)** Ένας φοιτητής του Γ.Π.Α., στο πλαίσιο μιας εργαστηριακής άσκησης, μέτρησε την ποσότητα λυκοπενίου σε 30 τέτοιες, τυχαία επιλεγμένες, συσκευασίες. **α)** Ποια είναι η πιθανότητα, η μέση ποσότητα λυκοπενίου στις 30 συσκευασίες να είναι μικρότερη από 22*mg*; **β)** Αν η μέση ποσότητα λυκοπενίου των 30 συσκευασιών βρεθεί ίση με 21*mg,* μήπως ο φοιτητής πρέπει να αμφιβάλλει για κάτι; **(ΙΙ)** Μπορούμε να απαντήσουμε στο ερώτημα, *«ποια είναι η πιθανότητα, μια τυχαία επιλεγμένη συσκευασία να περιέχει ποσότητα λυκοπενίου μεταξύ 23 και 25mg»*; Εξηγείστε.

2ο Θέμα [40] Το τμήμα ποιοτικού ελέγχου ενός εργοστασίου παραγωγής κομπόστας ροδάκινου, ελέγχει, μεταξύ άλλων, τη διάμετρο των ροδάκινων που προμηθεύεται από τους παραγωγούς. Από εμπειρικά δεδομένα προηγούμενων ετών, έχει επαληθευθεί ότι, η τυχαία μεταβλητή *X* που εκφράζει το μήκος της διαμέτρου των ροδάκινων που προμηθεύεται το εργοστάσιο από τον παραγωγό *A,* ακολουθεί κανονική κατανομή με μέση τιμή 60*mm* και τυπική απόκλιση 10*mm.* **α)** Επιλέγουμε τυχαία ένα ροδάκινο από την παραγωγή του προμηθευτή *A*. Ποια είναι η πιθανότητα, το μήκος της διαμέτρου του να βρίσκεται μεταξύ 59 και 61*mm*; **β)** Επιλέγουμε τυχαία, 5 ροδάκινα από την παραγωγή του προμηθευτή *A*. Ποια είναι η πιθανότητα, το πολύ ένα από αυτά να έχει διάμετρο με μήκος μεταξύ 59 και 61*mm*; **γ)** Βρείτε εκείνο το μήκος διαμέτρου για το οποίο ισχύει ότι: μεγαλύτερο μήκος διαμέτρου από αυτό έχει μόνο το 2% των ροδάκινων του προμηθευτή *A*. **δ)** Βρείτε ένα διάστημα συμμετρικό ως προς το μέσο μήκος διαμέτρου των ροδάκινων του παραγωγού *A*, τέτοιο ώστε, το 99% των ροδάκινων του παραγωγού *A* να έχει μήκος διαμέτρου που βρίσκεται εντός αυτού του διαστήματος.

3ο Θέμα [35] α) Μια αιματολογική εξέταση ανιχνεύει έλλειψη σιδήρου στο 95% των περιπτώσεων που πράγματι υπάρχει έλλειψη σιδήρου. Όμως, ανιχνεύει έλλειψη σιδήρου και στο 5% των περιπτώσεων που υπάρχει επάρκεια σιδήρου. Αν το ποσοστό του πληθυσμού στο οποίο υπάρχει έλλειψη σιδήρου είναι 10%, ποια είναι η πιθανότητα ένα άτομο στο οποίο έγινε διάγνωση έλλειψης σιδήρου να έχει πράγματι έλλειψη σιδήρου.

β) Έστω *X* και *Y* τυχαίες μεταβλητές με $X \sim B(25, 0.01)$ και $Y \sim B(25, 0.45)$. Σε ποια από τις δύο περιπτώσεις η κανονική προσέγγιση είναι ικανοποιητική και γιατί.

γ) Σε μια στατιστική έρευνα, τα ***δειγματοληπτικά*** *σφάλματα* μπορούν να αποφευχθούν; Εξηγείστε.

δ) Η πρόταση που ακολουθεί είναι σωστή ή λάθος και γιατί: *«Αν Α, Β είναι δύο ξένα ενδεχόμενα του δειγματικού χώρου Ω ενός πειράματος τύχης με* $P(A) = 0.1$ *και* $P(B) = 0.7$ *τότε,* $P(AB) = 0.8$ ».

ε) Μπορεί η συνάρτηση κατανομής *F* μιας συνεχούς τυχαίας μεταβλητής *X*, να πάρει τιμές μεγαλύτερες του 1; Η συνάρτηση πυκνότητας *f*; Εξηγείστε.

Ενδεικτικές απαντήσεις (Β΄ Σειρά)

1ο Θέμα: (Ι). α) Έστω X_i η ποσότητα λυκοπενίου που περιέχεται στη συσκευασία i ($i = 1,2,...,30$). Προφανώς ζητάμε την πιθανότητα $P(\overline{X} < 22)$ όπου $\overline{X} = \frac{X_1 + X_2 + ... + X_{30}}{30}$. Οι τυχαίες μεταβλητές X_i είναι ανεξάρτητες και ισόνομες με μέση τιμή $\mu_i = E(X_i) = 24$ και διασπορά $\sigma_i^2 = 2^2$ και επομένως από το Κ.Ο.Θ. προκύπτει ότι η $\overline{X}$ προσεγγίζεται ικανοποιητικά από την κανονική κατανομή $N(24, \frac{2^2}{30})$ ή $N(24, 0.365^2)$. Επομένως η ζητούμενη πιθανότητα είναι:

$$P(\overline{X} < 22) = P(Z < \frac{22-24}{0.365}) = P(Z < -5.48) = \Phi(-5.47) = 1 - \Phi(5.48) \approx 0 .$$

β) Επειδή η πιθανότητα να εμφανισθεί η τιμή αυτή στο δείγμα είναι (με βάση την απάντηση στο (α) ερώτημα) πολύ μικρή και παρόλα αυτά εμφανίσθηκε, είναι πολύ πιθανό, η μέση τιμή του λυκοπενίου που περιέχεται στις συσκευασίες των 200gr να μην είναι 24mg αλλά μικρότερη. **γ)** Όχι γιατί δεν γνωρίζουμε την κατανομή της X (γνωρίζουμε μόνο τη μέση τιμή και τη διασπορά της).

2ο Θέμα: $X \sim N(60, 10^2)$ άρα:

α) $P(59 < X < 61) = P(\frac{59-61}{10} < \frac{X-60}{10} < \frac{61-60}{10}) = P(-0.1 < Z < 0.1) = 2\Phi(0.1) - 1 = 0.0796$.

β) Έστω Y ο αριθμός των ροδάκινων από τα πέντε που επελέγησαν, που το καθένα έχει μήκος διαμέτρου μεταξύ 59 και 61mm. Προφανώς, $Y \sim B(5, 0.0796)$ άρα

$$P(Y \le 1) = P(Y = 0) + P(Y = 1) = \binom{5}{0} \cdot (0.0796)^0 \cdot (0.9204)^5 + \binom{5}{1} \cdot (0.0796)^1 \cdot (0.9204)^4 = ...$$

γ) Πρέπει: $P(X > \xi) = 0.02 \Leftrightarrow P(z > \frac{\xi - 60}{10}) = 0.02 \Leftrightarrow \Phi(\frac{\xi - 60}{10}) = 0.98 \Leftrightarrow$

$\Leftrightarrow \frac{\xi - 60}{10} = 2.06 \Leftrightarrow \xi = 80.6$.

δ) Πρέπει: $P(60 - c < X < 60 + c) = 0.99 \Leftrightarrow P(\frac{-c}{10} < Z < \frac{c}{10}) = 0.99 \Leftrightarrow$

$\Leftrightarrow 2 \cdot \Phi(\frac{c}{10}) - 1 = 0.99 \Leftrightarrow \Phi(\frac{c}{10}) = 0.995 \Leftrightarrow \frac{c}{10} = 2.58 \Leftrightarrow c = 25.8$ άρα το ζητούμενο διάστημα είναι: $(34.2, \ 85.8)$.

3ο Θέμα: α) Έστω τα ενδεχόμενα, *Ε: υπάρχει έλλειψη σιδήρου* και *Θ: το τεστ είναι θετικό*. Δίνονται οι πιθανότητες: $P(E) = 0.10$, $P(\Theta / E) = 0.95$ και $P(\Theta / E') = 0.05$.

Από τον τύπο του Bayes έχουμε: $P(E / \Theta) = \frac{P(\Theta / E) P(E)}{P(\Theta)} = 0.67857$.

Η πιθανότητα $P(\Theta)$ υπολογίσθηκε από το θεώρημα ολικής πιθανότητας:

$$P(\Theta) = P(\Theta / E) P(E) + P(\Theta / E') P(E') = 0.95 \cdot 0.10 + 0.05 \cdot 0.90 = 0.14 .$$

β) Στη δεύτερη περίπτωση γιατί στην πρώτη περίπτωση η διωνυμική παρουσιάζει θετική ασυμμετρία ενώ στη δεύτερη όχι. (δεκτή είναι επίσης η εξής απάντηση: στην πρώτη περίπτωση δεν ισχύει η συνθήκη $n \cdot p > 5$). **γ)** Όχι. Είναι αναπόφευκτα σφάλματα που συνδέονται με τη μεταβλητότητα του δείγματος, με το μέγεθος του δείγματος και με την επιλογή σχεδίου δειγματοληψίας. **δ)** Είναι λάθος γιατί εφόσον είναι ξένα, $P(AB) = P(\emptyset) = 0$ **ε)** Υπάρχει στις απαντήσεις άλλων εξεταστικών περιόδων.