Φ. ΚΟΛΥΒΑ - ΜΑΧΑΙΡΑ
Επικ. Καθηγήτρια Α.Π.Θ.

ΜΑΘΗΜΑΤΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

ΤΟΜΟΣ Ι
ΕΚΤΙΜΗΤΙΚΗ

Θεωρία και Λυμένες ασκήσεις

Θεσσαλονίκη

Εκδόσεις Ζήτη

Πρόλογος

Το βιβλίο αυτό απευθύνεται στους τεταρτοετείς φοιτητές των Μαθηματικών τμημάτων και σε όσους ασχολούνται με τη Στατιστική και έχουν καλό υπόβαθρο στη Μαθηματική Ανάλυση και στη θεωρία των Πιθανοτήτων. Το βιβλίο αυτό αντικαθιστά το «Μαθηματική Στατιστική τόμος I» των Φ. Κολυβά - Μαχαίρα και Κ. Μπαγιάτη. Η θεωρία αποτελεί βελτίωση της θεωρίας του προηγούμενου βιβλίου, ενώ οι ασκήσεις είναι ανανεωμένες ως προς το πνεύμα και ως προς το ύφος και φυσικά υπάρχει πληθώρα νέων ασκήσεων. Η ανάγκη βελτίωσης προέκυψε από την δεκαετή διδασκαλία του μαθήματος από τη συγγραφέα. Σ' αυτή την έκδοση διατηρήθηκε η δομή του προηγούμενου βιβλίου, για να μην αισθάνονται σε ξένο περιβάλλον οι φοιτητές παλαιοτέρων ετών που θα ξαναδιαβάσουν το μάθημα.

Το βιβλίο περιέχει τέσσερα κεφάλαια και δύο παραρτήματα.

Στο πρώτο κεφάλαιο μελετάται η κατανομή του τυχαίου δείγματος καθώς επίσης και οι κατανομές των κλασσικών στατιστικών συναρτήσεων.

Στο δεύτερο κεφάλαιο μελετώνται οι ιδιότητες των εκτιμητών και ο τρόπος κατασκευής εκτιμητών που έχουν κάποιες επιθυμητές ιδιότητες.

Στο τρίτο κεφάλαιο δίνονται μέθοδοι εύρεσης εκτιμητών με αναλυτικό τρόπο, ασχέτως ιδιοτήτων, και συγκρίνονται οι μέθοδοι ως προς την ποιότητα των εκτιμητών που παρέχουν.

Στο τέταρτο κεφάλαιο δίνονται μέθοδοι κατασκευής διαστημάτων εμπιστοσύνης.

Το παράρτημα Α είναι ένα λεξικό πιθανοτήτων με ορισμούς και τα κυριότερα θεωρήματα (χωρίς απόδειξη) που θεωρούνται απαραίτητα εργαλεία για τη Μαθηματική Στατιστική.

Το παράρτημα Β είναι οι πίνακες των κυριοτέρων κατανομών που χρησιμοποιούνται στα διαστήματα εμπιστοσύνης και είναι απαραίτητοι σε κάθε βιβλίο Στατιστικής.

Κάθε ένα από τα τέσσερα κεφάλαια περιέχει πολλά παραδείγματα για την κατανόηση των εννοιών και λυμένες ασκήσεις (συνολικά 100 ασκήσεις) στο τέλος κάθε κεφαλαίου. Σε κάθε άσκηση μελε-

τάται ένα πρόβλημα ολοκληρωμένα ενώ με τα παραδείγματα αποσαφηνίζονται οι έννοιες κάθε φορά που ορίζονται. Μερικές ασκήσεις και θεωρήματα ξεφεύγουν από τις απαιτήσεις των εξετάσεων, μπήκαν όμως για να κεντρίσουν το ενδιαφέρον αυτών που τους ενδιαφέρει η Στατιστική σαν αντικείμενο περαιτέρω έρευνας.

Τελειώνοντας θα ήθελα να ευχαριστήσω θερμά τους μεταπτυχιακούς φοιτητές Γιάννη Ανδρεάδη και Δέσποινα Δάσιου για τη βοήθειά τους στη διόρθωση των δοκιμίων.

Θεσσαλονίκη, Δεκέμβριος 1993

Φ. Κολυβά - Μαχαίρα

Περιεχόμενα

4. ΕΚΤΙΜΗΣΗ ΣΕ ΔΙΑΣΤΗΜΑ

Παράρτημα Α. ΣΤΟΙΧΕΙΑ ΘΕΩΡΙΑΣ ΠΙΘΑΝΟΤΗΤΩΝ

Συντομογραφίες - Συμβολισμοί

α.ε.ο.ε.δ.	αμερόληπτος εκτιμητής ομοιόμορφα ελάχιστης διασποράς
α.ο.ε.δ.	αμερόληπτος ομοιόμορφα ελάχιστης διασποράς
β.ε.	βαθμοί ελευθερίας
δ.ε.	διάστημα εμπιστοσύνης
Ε.Ε.Τ.	εκτιμητής ελαχίστων τετραγώνων
Ε.Μ.Π.	εκτιμητής μέγιστης πιθανοφάνειας
Κ.Ο.Θ.	κεντρικό οριακό θεώρημα
σ.α.κ.	συνάρτηση αθροιστικής κατανομής
σ.κ.	συνάρτηση κατανομής
σ.π.	συνάρτηση πιθανότητας
σ.π.π.	συνάρτηση πυκνότητας πιθανότητας
στ.σ.	στατιστική συνάρτηση
τ.δ.	τυχαίο δείγμα
τ.μ.	τυχαία μεταβλητή
Σ.Ε.Ε.Τ.	σταθμισμένοι εκτιμητές ελαχίστων τετραγώνων
χ.σ.	χαρακτηριστική συνάρτηση
C - R	Cramér - Rao
$\xrightarrow{P}$	συγκλίνει κατά πιθανότητα
$\xrightarrow{\sigma.\beta.}$	συγκλίνει σχεδόν βέβαια
$\xrightarrow{K.N.}$	συγκλίνει κατά νόμο
A'	ανάστροφος του πίνακα A
A^{-1}	αντίστροφος του πίνακα A
$\det A$	ορίζουσα του πίνακα A
$\underset{\sim}{x}$	διάνυσμα στήλη
$\underset{\sim}{x}'$	διάνυσμα γραμμή
$\lvert a \rvert$	απόλυτη τιμή του α

Κεφάλαιο 1

ΚΑΤΑΝΟΜΗ ΚΑΙ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΥΧΑΙΟΥ ΔΕΙΓΜΑΤΟΣ

1.1 Εισαγωγή

Τα φαινόμενα ή οι πληθυσμοί περιγράφονται με τη βοήθεια των τυχαίων μεταβλητών και των κατανομών τους, που τις περισσότερες φορές δεν είναι τέλεια γνωστές. Η μελέτη μιας κατανομής επιτυγχάνεται με τη μελέτη **όλων** των στοιχείων του πληθυσμού ή του φαινομένου που περιγράφει, πράγμα που είναι τις περισσότερες φορές πρακτικά αδύνατο είτε λόγω κόστους είτε λόγω έλλειψης χρόνου. Για να καταλήξουμε λοιπόν σε συμπεράσματα για την κατανομή, χρησιμοποιούμε έναν αριθμό ανεξάρτητων παρατηρήσεων όσο το δυνατόν αντιπροσωπευτικότερων του πληθυσμού ή του φαινομένου που μελετάμε.

Ορισμός 1.1.1

Τυχαίο δείγμα (τ.δ.) $\underset{\sim}{X}$ είναι ένα πεπερασμένο σύνολο ανεξάρτητων δοκιμών $(X_1, X_2, \ldots, X_n)$ της ίδιας τ.μ. Χ. Ο αριθμός n λέγεται **μέγεθος** του δείγματος. Τα αποτελέσματα των n δοκιμών σημειώνονται με $\underset{\sim}{x}' = (x_1, x_2, \ldots, x_n)$ ***και δεν είναι τ.μ.*** ενώ το τ.δ. $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$ **είναι τ.μ.**

Ο τρόπος με τον οποίο λαμβάνεται το δείγμα λέγεται **δειγματοληψία** και γίνεται είτε με επανάθεση είτε χωρίς επανάθεση.

Ο πληθυσμός από τον οποίο προέρχεται το δείγμα μπορεί να είναι άπειρου πλήθους, πεπερασμένου ή το πολύ αριθμήσιμου πλήθους.

Αν ο πληθυσμός είναι άπειρος τότε οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες και ισχύει:

$$f_{\underset{\sim}{X}}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)\, f_{X_2}(x_2) \ldots f_{X_n}(x_n)$$

όπου $f_{\underset{\sim}{X}}(x_1, x_2, \ldots, x_n)$ είναι η κοινή κατανομή του τ.δ. $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$ και $f_{X_i}(x_i)$, $i=1, 2, \ldots, n$ είναι η κατανομή της τ.μ. X_i.

Αν ο πληθυσμός είναι πεπερασμένος και η δειγματοληψία γίνεται **χωρίς επανάθεση** τότε οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι εξαρτημένες και ισχύει:

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{N}\,\frac{1}{(N-1)}\,\cdots\,\frac{1}{(N-n+1)} = \frac{1}{(N)_n}$$

όπου N είναι το μέγεθος του πληθυσμού ενώ n είναι το μέγεθος του δείγματος.

Αν ο πληθυσμός είναι πεπερασμένος αλλά η δειγματοληψία γίνεται **με επανάθεση** τότε οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες και ισχύει:

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{N^n} .$$

Σημείωση: *Αν δεν υπάρχει ιδιαίτερη ένδειξη υποθέτουμε ότι η δειγματοληψία γίνεται με επανάθεση οπότε το δείγμα είναι τυχαίο.*

Στο εξής θα υποθέτουμε πάντοτε ότι το δείγμα μας είναι τυχαίο δηλαδή οι τ.μ. $X_1, X_2, \ldots, X_n$ θα είναι ανεξάρτητες και ισόνομες και συνεπώς θα ισχύει:

$$\mathbf{f_{\underset{\sim}{X}}(\underset{\sim}{x}) = \prod_{i=1}^{n} f_{X_i}(x_i)} \quad .$$

1.2 Μελέτη της κατανομής του δείγματος

1.2.1 Περίπτωση διακριτού αρχικού πληθυσμού

Έστω X μια διακριτή τ.μ. που παίρνει τις τιμές $\alpha_1, \alpha_2, \ldots, \alpha_N, \ldots$ με κατανομή

$$P(X=\alpha_i) = p_i \ , \quad \sum_i p_i = 1 \ .$$

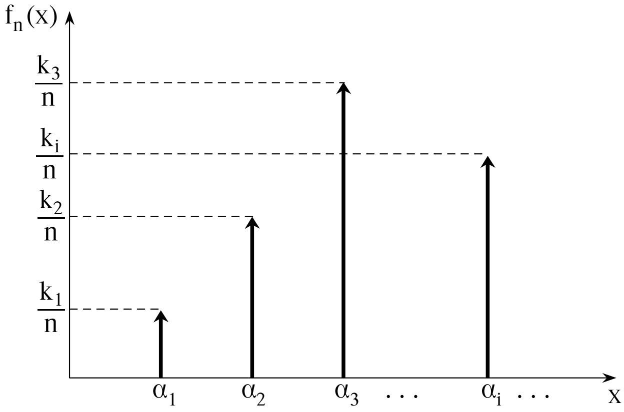
Στη Στατιστική τα p_i θεωρούνται άγνωστες παράμετροι που προσπαθούμε να προσδιορίσουμε με τη βοήθεια του δείγματος. Κάνουμε n δοκιμές, παίρνουμε ένα δείγμα μεγέθους n και ας υποθέσουμε χωρίς περιορισμό της γενικότητας ότι:

η τιμή	α_1	εμφανίζεται	k_1	φορές
»	α_2	»	k_2	»
	$\vdots$		$\vdots$	
»	α_i	»	k_i	»
	$\vdots$		$\vdots$	

έτσι ώστε $k_i \geq 0$, $\sum_i k_i = n$ $\quad \sum_i \frac{k_i}{n} = 1$.

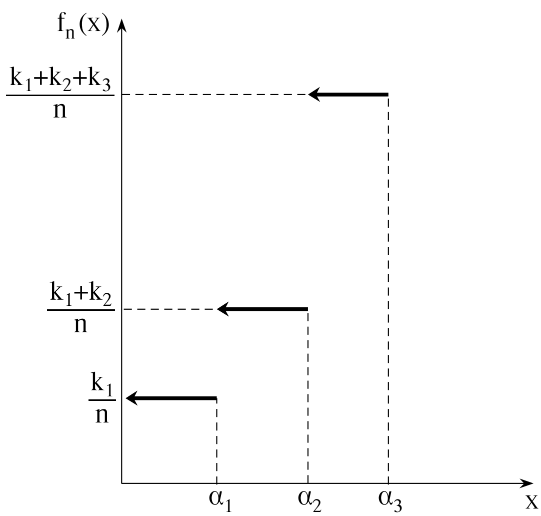
Παρατηρούμε λοιπόν ότι αντιστοιχώντας σε κάθε τιμή της τ.μ. X τον αριθμό $\frac{k_i}{n}$ δημιουργείται μια καινούργια κατανομή που καλείται **εμπειρική κατανομή**.

Γραφικά η εμπειρική κατανομή παριστάνεται με το **ραβδόγραμμα συχνοτήτων** (Σχήμα 1.2.1).



Σχήμα 1.2.1

Η αθροιστική συνάρτηση κατανομής $F_n(x)$ που αντιστοιχεί στον παραπάνω εμπειρικό νόμο, είναι το ποσοστό των τιμών που είναι μικρότερες ή ίσες του x και ορίζεται έτσι ώστε να είναι συνεχής από αριστερά. Γραφικά η σ.κ. $F_n(x)$ παριστάνεται με το παρακάτω σχήμα 1.2.2.



Σχήμα 1.2.2

1.2.2 Περίπτωση συνεχούς αρχικής κατανομής

Έστω ότι η τ.μ. X που περιγράφει τον πληθυσμό είναι συνεχής· τότε υπάρχει κάποια (συνήθως άγνωστη) σ.π.π. $f(x)$ για την οποία ισχύει:

$$f(x) \geq 0 \quad \forall\, x \in R \quad \text{και} \quad \int_R f(x)\,dx = 1 \;.$$

Κάνουμε n δοκιμές και παίρνουμε:

στην	1η	δοκιμή	την	τιμή	x_1
»	2η	»	»	»	x_2
	⋮				
»	n-στή	»	»	»	x_n

Προφανώς ισχύει:

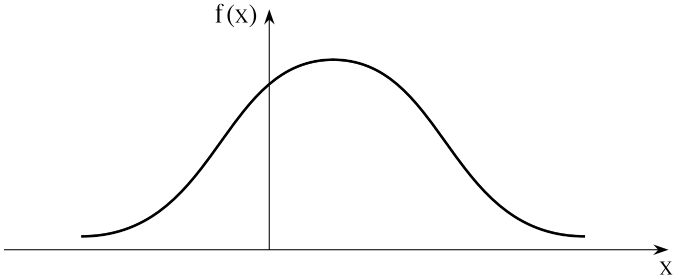
$$P(X=x_1) = 0\,,\quad P(X=x_2) = 0,\, \ldots,\, P(X=x_n) = 0 \;.$$

Επίσης είναι πολύ πιθανόν όλα τα x_i να είναι διαφορετικά μεταξύ τους. Σε μια συνεχή κατανομή ο πληθυσμός είναι άπειρος και συνεπώς όλες οι τιμές x_i, $i=1, 2, \ldots, n$ είναι ανεξάρτητες μεταξύ τους. Έτσι αντιστοιχώντας σε κάθε τιμή x_i πιθανότητα ίση με $\frac{1}{n}$ δημιουργούμε μια καινούργια κατανομή, μια διακριτή εμπειρική κατανομή. Στο παρακάτω σχήμα 1.2.3 έχουμε την γραφική παράσταση της κατανομής του αρχικού πληθυσμού και το ραβδόγραμμα που αντιστοιχεί στην εμπειρική κατανομή.

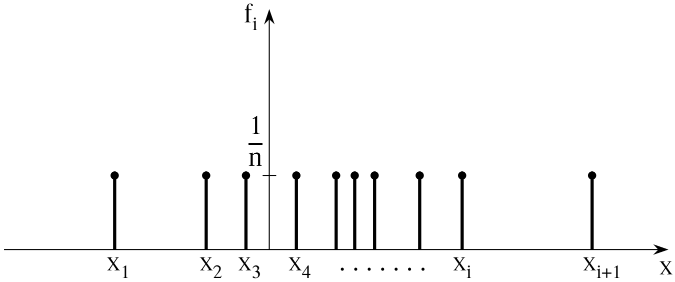
Παρατηρούμε ότι, το ραβδόγραμμα συχνοτήτων μας δίνει πολύ μικρή πληροφορία για τον αρχικό πληθυσμό. Σ' αυτήν την περίπτωση σκεφτόμαστε ως εξής: Αφού ο αρχικός πληθυσμός έχει συνεχή κατανομή δεν μπορούμε να αντιστοιχίσουμε θετική πιθανότητα σε σημείο, αλλά μόνο σε διάστημα. Έτσι χωρίζουμε το πεδίο ορισμού της τ.μ. σε διαστήματα που χαρακτηρίζονται και αντιπροσωπεύονται από τα μέσα τους και σαν πιθανότητα θεωρούμε την πυκνότητα των δοκιμών σ' αυτό το διάστημα· δηλαδή σε κάθε διάστημα αντιστοιχούμε τον αριθμό $\frac{k_i}{n}$ όπου k_i είναι το πλήθος των δοκιμών που ανήκουν σ' αυτό.

Έτσι λοιπόν η γραφική παράσταση της εμπειρικής κατανομής στην περίπτωση συνεχούς αρχικής κατανομής είναι το **ιστόγραμμα συχνοτήτων**, που είναι ένα σύνολο παραλληλογράμμων με βάση το διάστημα (x_i, x_{i+1}), $i=1, \ldots, n-1$ και εμβαδόν ίσο με $\frac{k_i}{n}$ ούτως ώστε το άθροισμα των εμβαδών να είναι ίσο με τη μονάδα.

Παρατηρούμε ότι το ιστόγραμμα συχνοτήτων μοιάζει με την σ.π.π. και μάλιστα, όσο το πλήθος των παρατηρήσεων και των διαστημάτων μεγαλώνει τότε το σχήμα 1.2.4 I τείνει να ταυτισθεί με το σχήμα 1.2.4 II.

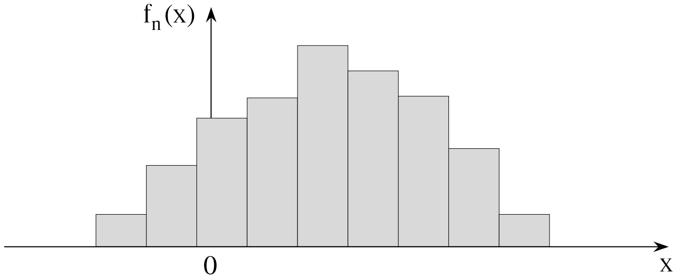


I. Αρχικός πληθυσμός με συνεχή κατανομή

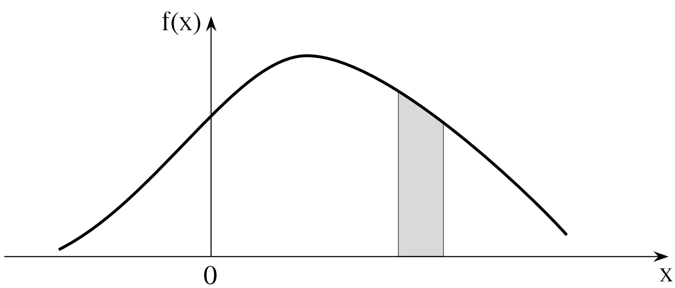


II. Διακριτή εμπειρική κατανομή

Σχήμα 1.2.3



I. Ιστόγραμμα συχνοτήτων



II. σ.π.π. συνεχούς αρχικής κατανομής

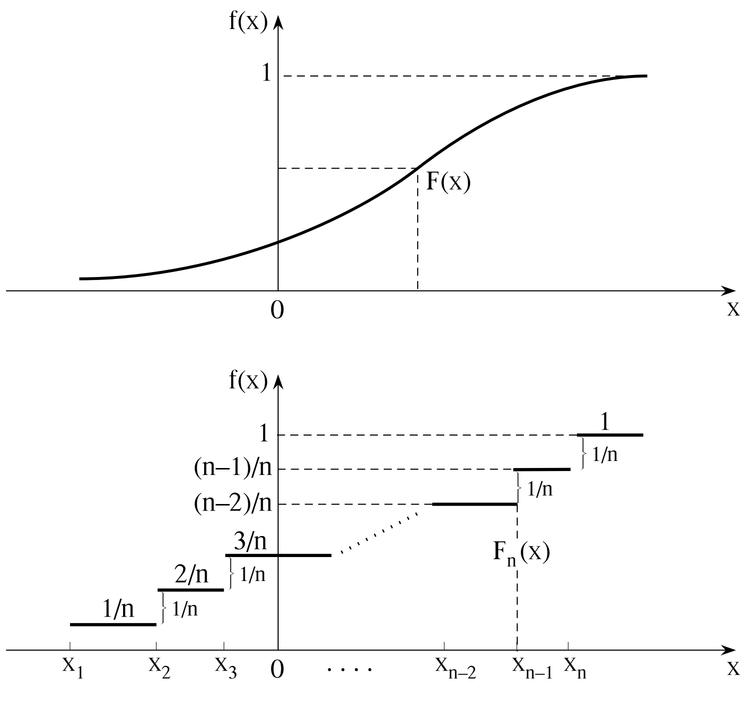
Σχήμα 1.2.4

1.2.3 Μελέτη της σ.κ. της εμπειρικής κατανομής

Ας θεωρήσουμε την τ.μ. X που έχει κάποια σ.κ. F(x). Κάνουμε n ανεξάρτητες δοκιμές, παίρνουμε ένα τ.δ. $\underset{\sim}{X}'= (X_1, X_2, \ldots, X_n)$ και σε κάθε x_i αντιστοιχούμε πιθανότητα $\frac{1}{n}$. Κατ' αυτόν τον τρόπο δημιουργούμε την εμπειρική κατανομή. Επίσης θεωρούμε μια συνάρτηση $F_n(x)$ που παριστάνει το κλάσμα, ή ποσοστό, των σημείων (δοκιμών) που βρίσκονται αριστερά του x. Για παράδειγμα, αν οι παρατηρήσεις είναι

$$x_1 \leq x_2 \leq \ldots x_k \leq x < x_{k+1} \leq \ldots \leq x_n \quad \text{θα είναι} \quad F_n(x) = \frac{k}{n}.$$

Έτσι για κάθε $x \in R$ η $F_n(x)$ ορίζει μια τ.μ. που την ονομάζουμε **εμπειρική συνάρτηση κατανομής** του δείγματος $\underset{\sim}{X}'= (X_1, X_2, \ldots, X_n)$. Η εμπειρική σ.κ. είναι διακριτή με σημεία ασυνέχειας τα σημεία $x_1, x_2, \ldots, x_n$ με βήμα $\frac{1}{n}$ αν όλα τα σημεία είναι διαφορετικά (σχήμα 1.2.5).



Σχήμα 1.2.5

Στη συνέχεια για κάποιο σταθερό x θεωρούμε την τ.μ. $R_n(x)$ που ορίζε-

ται ως εξής: $R_n(x)$ είναι το πλήθος των στοιχείων του δείγματος που βρίσκονται πριν από το x.
Αν:

$$\{R_n(x)=k\} \Rightarrow \begin{cases} k & \text{από τα } x_i \text{ είναι } <x \\ n-k & \text{»} \quad \text{»} \quad x_i \quad \text{»} \quad \geq x \end{cases}$$

τότε:

$$P(R_n(x)=k) = \binom{n}{k} F(x)^k (1-F(x))^{n-k} \quad , \quad k=0, 1, 2, \ldots, n \; .$$

Παρατηρούμε ότι η τ.μ. $R_n(x)$ ακολουθεί διωνυμική κατανομή $B(n, p=F(x))$. Έτσι:

$$E(R_n(x)) = nF(x)$$

$$E\left(\frac{R_n(x)}{n}\right) = F(x)$$

$$\text{var}(R_n(x)) = nF(x)(1-F(x)) \Rightarrow \text{var}\left(\frac{R_n(x)}{n}\right) = \frac{F(x)(1-F(x))}{n}$$

Όμως από τον ορισμό της εμπειρικής σ.κ. $F_n(x)$ και της τ.μ. $R_n(x)$ έχουμε ότι:

$$F_n(x) = \frac{R_n(x)}{n} \; ,$$

άρα: $$E(F_n(x)) = F(x) \quad \text{και} \quad \text{var}(F_n(x)) = \frac{F(x)(1-F(x))}{n} \; .$$

Εφαρμόζοντας την ανισότητα του Chebychev έχουμε:

$$P(|F_n(x)-F(x)| > \varepsilon) \leq \frac{F(x)(1-F(x))}{n\varepsilon^2} \; .$$

Όμως $\dfrac{F(x)(1-F(x))}{n\varepsilon^2} \to 0$ όταν $n\to\infty$ $\forall\, \varepsilon>0$.

Άρα $$F_n(x) \xrightarrow{p} F(x) \; .$$

Σύμφωνα με τον ισχυρό νόμο των μεγάλων αριθμών κάτω από τις ίδιες συνθήκες ισχύει:

$$F_n(x) \xrightarrow{\sigma.\beta.} F(x)$$

που σημαίνει ότι η κλιμακωτή συνάρτηση $F_n(x)$ συγκλίνει προς τη συνεχή συνάρτηση $F(x)$, όταν το μέγεθος του δείγματος μεγαλώνει ($n\to\infty$).

Πράγματι:

Αν για κάθε $x \in R$ ορίσουμε τις τ.μ. $Y_j(x)$, $j=1, 2, \ldots, n$ από τη σχέση:

$$Y_j(x) = \begin{cases} 1 & \text{αν} \quad X_j \le x \\ 0 & \text{αν} \quad X_j > x \end{cases} \qquad j=1, 2, \ldots, n$$

τότε οι τ.μ. $Y_j(x)$ είναι ανεξάρτητες με κατανομή $B(1,\ p=F(x))$.

Σύμφωνα με τον ισχυρό νόμο των μεγάλων αριθμών

$$\frac{1}{n}\sum_{j=1}^{n} Y_j(x) \xrightarrow{\sigma.\beta.} F(x)$$

όμως
$$\frac{1}{n}\sum_{j=1}^{n} Y_j(x) = F_n(x) \qquad \text{ο.ε.δ.}$$

Η χρησιμοποίηση της $F_n(x)$ σαν εκτιμήτριας της θεωρητικής κατανομής $F(x)$ δικαιολογείται και από την παρακάτω πολύ ισχυρή ιδιότητα.

Θεώρημα 1.2.1 (Clivenko - Cantelli)

Η εμπειρική συνάρτηση κατανομής $F_n(x)$ συγκλίνει ομοιόμορφα προς την θεωρητική συνάρτηση κατανομής $F(x)$ με πιθανότητα 1, ή

$$\sup_x |F_n(x) - F(x)| \xrightarrow[n\to\infty]{\sigma.\beta.} 0 .$$

Απόδειξη

Για δοσμένο $\varepsilon > 0$ θεωρούμε τα σημεία $c_1, c_2, \ldots, c_\varkappa$ τέτοια ώστε

$$F(c_i) - F(c_{i-1}) < \frac{\varepsilon}{2} , \quad i=1, 2, \ldots, \varkappa+1$$

με $c_0 = -\infty$, $c_{\varkappa+1} = \infty$.

Αφού ισχύει $F_n(x) \xrightarrow{\sigma.\beta.} F(x)$ τότε για $\delta > 0$ υπάρχει N_i έτσι ώστε

$$P\left(|F_n(c_i) - F(c_i)| < \frac{\varepsilon}{2} ,\ \forall\, n > N_i\right) > 1 - \frac{\delta}{\varkappa+1} \qquad (1.1)$$

Επαγωγικά μπορεί να δειχθεί ότι:

$$P\left(\bigcap_{i=1}^{\varkappa+1} A_i\right) \ge \sum_{i=1}^{\varkappa+1} P(A_i) - \varkappa \qquad (1.2)$$

Διαλέγοντας σαν $N = \max\{N_1, \ldots, N_{\varkappa+1}\}$ και εφαρμόζοντας την (1.2) για το ενδεχόμενο $A_i = \left\{ |F_n(c_i) - F(c_i)| < \frac{\varepsilon}{2} \text{ για } n > N \right\}$ η (1.1) γίνεται:

$$P\left(|F_n(c_i)-F(c_i)|<\frac{\varepsilon}{2}\quad \forall\, n>N,\quad i=1,2,\ldots,\varkappa+1\right)>(\varkappa+1)\left(1-\frac{\delta}{\varkappa+1}\right)-\varkappa=1-\delta \quad (1.3)$$

Οι σ.κ. F_n και F είναι μη φθίνουσες συναρτήσεις από τον ορισμό τους και έτσι όταν συμβαίνει το γεγονός $\left\{|F_n(c_i)-F(c_i)|<\frac{\varepsilon}{2},\ \forall\, n>N,\ i=1,2,\ldots,\varkappa+1\right\}$ τότε για $c_{i-1}<x<c_i$ ισχύουν:

$$F_n(x)-F(x)<F_n(c_i)-F(c_{i-1})=F_n(c_i)-F(c_i)+F(c_{i-1})<\varepsilon$$

$$F(x)-F_n(x)<F(c_i)-F_n(c_{i-1})=F(c_i)-F(c_{i-1})-\big(F_n(c_{i-1})-F(c_{i-1})\big)<\varepsilon\,.$$

δηλ. $|F_n(x)-F(x)|<\varepsilon,\ \forall\, x$ άρα $\sup\limits_x |F_n(x)-F(x)|<\varepsilon$ λόγω της (1.3)

$$P\big(\sup_x |F_n(x)-F(x)|<\varepsilon\,,\ \forall\, n>N\big)<1-\delta \qquad \text{ο.ε.δ.}$$

Για πολύ μεγάλα δείγματα ισχύει το αποτέλεσμα

$$\operatorname*{Sup}_x |F_n(x)-F(x)|<\frac{c}{\sqrt{n}}\,.$$

1.3 Χαρακτηριστικά δείγματος

1.3.1 Διατεταγμένο δείγμα

Θεωρούμε ένα τ.δ. $\underset{\sim}{X}'=(X_1, X_2, \ldots, X_n)$ από μια τ.μ. X και το διατάσσουμε κατ' αύξουσα σειρά μεγέθους. Το δείγμα που προκύπτει λέγεται **διατεταγμένο δείγμα**, και είναι:

$$\min\{X_1, X_2, \ldots, X_n\}=X_{(1)}\le X_{(2)}\le \ldots \le X_{(n)}=\max\{X_1, X_2, \ldots, X_n\}\,.$$

Οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες και ισόνομες και μας ενδιαφέρει η κοινή κατανομή των τ.μ. $Y_1=X_{(1)}$, $Y_2=X_{(2)}$, …, $Y_n=X_{(n)}$.

Θεώρημα 1.3.1

Αν οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι συνεχείς ανεξάρτητες και ισόνομες με σ.κ. $F_X(x)=P(X\le x)$ και σ.π.π. $f_X(x)$ τότε:

(i) Η σ.π.π. $f_\varkappa(y)$ της τ.μ. $Y_\varkappa=X_{(\varkappa)}$ είναι:

$$f_\varkappa(y)=\frac{n!}{(\varkappa-1)!\,(n-\varkappa)!}\,(F_X(y))^{\varkappa-1}\,(1-F_X(y))^{n-\varkappa}\,f_X(y)\,,\quad \varkappa=1,2,\ldots,n.$$

Για $\varkappa=1$: $f_1(y)=n\big(1-F_X(y)\big)^{n-1}f_X(y)$

Για $\varkappa=n$: $f_n(y)=n\big(F_X(y)\big)^{n-1}f_X(y)$

(ii) Η κοινή κατανομή των τ.μ. $Z=X_{(\varkappa)}$, $Y=X_{(s)}$, $1 \le \varkappa \le s \le n$ είναι:

$$f_{\varkappa s}(z, y) = \begin{cases} c\,(F_X(z))^{\varkappa-1}\,(F_X(y)-F_X(z))^{s-\varkappa-1}\,(1-F_X(y))^{n-s}\, f_X(z)\, f_X(y), & z<y \\ 0 & , z \ge y \end{cases}$$

όπου $c=\dfrac{n!}{(\varkappa-1)!\,(s-\varkappa-1)!\,(n-s)!}$

(iii) Η κοινή πυκνότητα των $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ είναι:

$$f_{1,2,\ldots,n}(y_1, y_2, \ldots, y_n) = \begin{cases} n!\, f_X(y_1)\, f_X(y_2) \ldots f_X(y_n) & , \ y_1<y_2< \ldots <y_n \\ 0 & , \ \text{αλλού} \end{cases}$$

Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση που οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι ένα τ.δ. από ομοιόμορφη κατανομή $U(0, 1)$ τότε:

(i) Η κατανομή της τ.μ. $Y_\varkappa=X_{(\varkappa)}$ είναι:

$$f_\varkappa(y) = \frac{n!}{(\varkappa-1)!\,(n-\varkappa)!}\, y^{\varkappa-1}\,(1-y)^{n-\varkappa}$$

και $$E(X_{(\varkappa)}) = \frac{\varkappa}{n+1} \quad , \quad \mathrm{var}X_{(\varkappa)} = \frac{\varkappa(n-\varkappa+1)}{(n+1)^2\,(n+2)} \quad , \quad \varkappa=1, 2, \ldots, n$$

(ii) Η κοινή κατανομή των $X_{(1)}, X_{(n)}$ είναι:

$$f_{1n}(z, y) = \begin{cases} n(n-1)\,(y-z)^{n-2} & , \ 0 \le z < y \le 1 \\ 0 & , \ \text{αλλού} \end{cases}$$

με $\mathrm{cov}\,(X_{(1)}, X_{(n)}) = \dfrac{1}{(n+1)^2\,(n+2)}$.

Ορισμός 1.3.1

Αν $X_1, X_2, \ldots, X_n$ είναι ένα τ.δ. από κάποια κατανομή τότε είναι:

(i) **εύρος ή πλάτος του δείγματος** η ποσότητα $R= X_{(n)}-X_{(1)}$.

(ii) **διάμεσος ή διχοτόμος του δείγματος** η ποσότητα

$$d= \begin{cases} X_{(\varkappa)} & \text{αν } \ n=2\varkappa+1 \\ \dfrac{X_{(\varkappa)}+X_{(\varkappa+1)}}{2} & \text{αν } \ n=2\varkappa \end{cases}$$

(iii) **μέσο του πλάτους του δείγματος** η ποσότητα $\dfrac{X_{(1)}+X_{(n)}}{2}$.

Η κατανομή της τ.μ. $R= X_{(n)}-X_{(1)}$, όταν $X_i \sim U(0, 1)$, είναι

$$f_R(r) = n(n-1)r^{n-1}(1-r) \quad , \quad 0<r<1$$

με $ER=\dfrac{n-1}{n+1}$ και $varR=\dfrac{2(n-1)}{(n+1)^2(n+2)}$.

1.3.2 Στατιστικά δείγματος

Ορισμός 1.3.2

Έστω τ.δ. $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$ από τ.μ. X. Κάθε μετρήσιμη συνάρτηση του δείγματος $U(X_1, X_2, \ldots, X_n)$ που δεν περιέχει άγνωστες παραμέτρους καλείται **στατιστική συνάρτηση (στ.σ.)**.

Με τη βοήθεια των στ.σ. μπορούμε να ορίσουμε τα στατιστικά του δείγματος σε αναλογία με τις παραμέτρους του πληθυσμού από τον οποίο προέρχεται. Τα στατιστικά αυτά είναι:

i) **Ο δειγματικός μέσος:**

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

ii) **Η δειγματική ροπή r τάξης**

$$m'_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \,, \quad r=2, 3, \ldots$$

Η δειγματική ροπή πρώτης τάξης είναι ο δειγματικός μέσος.

iii) **Η δειγματική κεντρική ροπή r τάξης**

$$m_r = \frac{1}{n}\sum_{i=1}^{n} (X_i-\bar{X})^r \,, \quad r= 2, 3, \ldots$$

Για r=2 έχουμε **τη δειγματική διασπορά** που συμβολίζεται:

$$S'^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i-\bar{X})^2 \,.$$

Πολλές φορές όμως και για λόγους που θα εξηγήσουμε παρακάτω σα δειγματική διασπορά παίρνουμε την ποσότητα

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i-\bar{X})^2 \,.$$

Η ποσότητα S ή S' λέγεται **δειγματική τυπική απόκλιση**.

Αν $\underset{\sim}{Y}' = (Y_1, Y_2, \dots, Y_n)$ είναι ένα τ.δ. από τ.μ. Y και

$\underset{\sim}{X}' = (X_1, X_2, \dots, X_n)$ είναι ένα τ.δ. από τ.μ. X τότε:

iv) **Δειγματική ή εμπειρική συνδιασπορά** είναι:

$$S'_{xy} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) \quad ή \quad S_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) .$$

v) **Δειγματικός ή εμπειρικός συντελεστής συσχέτισης** είναι:

$$r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{\sum_i (X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_i (X_i-\bar{X})^2 \sum_i (Y_i-\bar{Y})^2}} .$$

1.3.3 Μελέτη της κατανομής της μέσης τιμής και της διασποράς

Οι στ.σ. είναι τ.μ. και συνεπώς έχουν κάποια κατανομή. Στη συνέχεια θα μελετήσουμε την κατανομή της μέσης τιμής $\bar{X}$ και της διασποράς S^2 ενός δείγματος που προέρχεται από κανονική κατανομή.

Θεώρημα 1.3.2

Έστω $\underset{\sim}{X}' = (X_1, X_2, \dots, X_n)$ ένα τ.δ. από κανονική κατανομή $N(\mu, \sigma^2)$. Τότε η κατανομή της δειγματικής τιμής $\bar{X}$ είναι κανονική $N\left(\mu, \frac{\sigma^2}{n}\right)$.

Απόδειξη

Η χαρακτηριστική συνάρτηση (χ.σ.) της μέσης τιμής $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ είναι:

$$\varphi_{\bar{X}}(t) = Ee^{it\bar{X}} = Ee^{i\frac{t}{n}\sum_{i=1}^{n} X_i} = E\prod_{i=1}^{n} e^{\frac{it}{n}X_i} = \prod_{i=1}^{n} Ee^{\frac{it}{n}X_i} = \prod_{i=1}^{n}\varphi_{X_i}\left(\frac{t}{n}\right) \tag{1.4}$$

Όμως $X_i \sim N(\mu, \sigma^2)$ έτσι η (1.4) γίνεται:

$$\varphi_{\bar{X}}(t) = \prod_{i=1}^{n} e^{i\frac{t}{n}\mu - \frac{\sigma^2 t^2}{2n^2}} = e^{it\mu - \frac{\sigma^2 t}{2n}} \tag{1.5}$$

Η σχέση (1.5) δεν είναι άλλη από την χ.σ. της κατανομής $N\left(\mu, \frac{\sigma^2}{n}\right)$ ο.ε.δ.

Θεώρημα 1.3.3

Έστω $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$ τ.δ. από κατανομή $F(x)$ με $E(X_i)=\mu$ και $\mathrm{var}X_i = \sigma^2 < \infty$ τότε:

$$\frac{\bar{X}-\mu}{\sigma}\sqrt{n} \xrightarrow{K.N} N(0, 1)\ .$$

Απόδειξη

Το θεώρημα αυτό είναι άμεση απόδειξη του θεωρήματος των Lindeberg - Lévy (Κ.Ο.Θ.) και η σύγκλιση είναι καλή για $n \geq 30$.

Θεώρημα 1.3.4

Έστω $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$ τ.δ. από κανονική κατανομή $N(\mu, \sigma^2)$. Τότε οι τ.μ. $\bar{X}$ και S^2 είναι ανεξάρτητες η τ.μ. $\frac{\bar{X}-\mu}{S}\sqrt{n}$ ακολουθεί κατανομή t_{n-1} και η τ.μ. $\frac{(n-1)S^2}{\sigma^2}$ ακολουθεί κατανομή χ^2_{n-1} .

Απόδειξη

Η δειγματική διασπορά S^2 γράφεται

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}(X_i-\mu)^2 - n(\bar{X}-\mu)^2\right) \Rightarrow$$

$$\Rightarrow \frac{(n-1)S^2}{\sigma^2} + \frac{n(\bar{X}-\mu)^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \tag{1.6}$$

Θεωρούμε τον μετασχηματισμό $Y_i = \frac{X_i-\mu}{\sigma}$ οπότε $Y_i \sim N(0, 1)$ και οι τ.μ. Y_i είναι ανεξάρτητες $i=1, 2, \ldots, n$.

Με την παρατήρηση ότι:

$$\sum_{i=1}^{n} Y_i = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma} = n\,\frac{\bar{X}-\mu}{\sigma} \Rightarrow \frac{\bar{X}-\mu}{\sigma} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

η (1.6) γίνεται:

$$\frac{(n-1)S^2}{\sigma^2} + \frac{1}{n}\left(\sum_{i=1}^{n} Y_i\right)^2 = \sum_{i=1}^{n} Y_i^2 \tag{1.7}$$

Στη συνέχεια θεωρούμε ακόμη ένα γραμμικό μετασχηματισμό:

$$\underset{\sim}{Z} = \Omega \underset{\sim}{Y} \qquad \underset{\sim}{Y} = \Omega' \underset{\sim}{Z}$$

όπου Ω είναι ένας ορθογώνιος πίνακας (δηλ. $\Omega'=\Omega^{-1}$) του οποίου η πρώτη γραμμή είναι η

$$\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right).$$

Η σ.π.π. της τ.μ. $\underset{\sim}{Z}$ είναι:

$$g(z_1, z_2, \ldots, z_n) = f\big(y_1(z_1, z_2, \ldots, z_n), \ldots, y_n(z_1, z_2, \ldots, z_n)\big)\cdot|\det\Omega| =$$

$$= 1\cdot\left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\underset{\sim}{Y}'\underset{\sim}{Y}\right\} = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\underset{\sim}{Z}'\Omega\Omega'\underset{\sim}{Z}\right\} =$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\underset{\sim}{Z}'\underset{\sim}{Z}\right\}.$$

($\det \Omega$ είναι η ορίζουσα του πίνακα Ω και ισχύει $\det\Omega = 1$).

Παρατηρούμε λοιπόν ότι οι τ.μ. Z_i ακολουθούν κατανομή $N(0, 1)$, $i = 1, 2, \ldots, n$ και είναι ανεξάρτητες μεταξύ τους. Επίσης

$$\sum_{i=1}^{n} Z_i^2 = \underset{\sim}{Z}'\underset{\sim}{Z} = \underset{\sim}{Y}'\Omega'\Omega\underset{\sim}{Y} = \underset{\sim}{Y}'\underset{\sim}{Y} = \sum_{i=1}^{n} Y_i^2 \tag{1.8}$$

και
$$Z_1 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i = \frac{1}{\sqrt{n}}\frac{\bar{X}-\mu}{\sigma} = \frac{1}{n}\frac{\bar{X}-\mu}{\sigma}\sqrt{n} \quad \text{και} \quad Z_1^2 = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i\right)^2 \tag{1.9}$$

Με τη βοήθεια των (1.8) και (1.9) η σχέση (1.7) γίνεται:

$$\frac{(n-1)S^2}{\sigma^2} + Z_1^2 = \sum_{i=1}^{n} Z_i^2 \qquad \frac{(n-1)S^2}{\sigma^2} = \sum_{i=2}^{n} Z_i^2 \tag{1.10}$$

Συμπέρασμα

i) Από τη σχέση (1.9) διαπιστώνουμε ότι η τ.μ. $\bar{X}$ είναι συνάρτηση μόνο του Z_1 και από τη σχέση (1.10) ότι το S^2 είναι συνάρτηση μόνο των $Z_2, Z_3, \ldots, Z_n$. Άρα οι τ.μ. $\bar{X}$ και S^2 είναι ανεξάρτητες.

ii) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ αφού $Z_i \sim N(0, 1)$

Έτσι: $ES^2=\sigma^2$ και $\operatorname{var}S^2 = \frac{\sigma^4}{(n-1)^2}2(n-1) = \frac{2\sigma^4}{n-1}$.

iii) Από το θεώρημα 1.3.2 ισχύει: $\frac{\bar{X}-\mu}{\sigma}\sqrt{n} \sim N(0, 1)$.

Άρα σύμφωνα με το θεώρημα Α. 3.9 έχουμε ότι:

$$\frac{\frac{\bar{X}-\mu}{\sigma}\sqrt{n}}{\sqrt{\frac{\frac{(n-1)S^2}{\sigma^2}}{n-1}}} = \frac{\bar{X}-\mu}{S}\sqrt{n} \sim t_{n-1} \qquad \text{ο.ε.δ.}$$

Το αποτέλεσμα αυτό είναι πολύ χρήσιμο στη δοκιμασία των υποθέσεων διότι συνήθως στην πράξη η θεωρητική διασπορά είναι άγνωστη οπότε την αντικαθιστούμε με τη δειγματική διασπορά. Αυτό όμως έχει σαν αποτέλεσμα να αλλάζει η κατανομή του $\bar{X}$, από κανονική σε Student.

Θεώρημα 1.4.5

Αν $\underset{\sim}{X}$ είναι ένα τ.δ. μεγέθους n από κατανομή $N(\mu_1, \sigma_1^2)$ και $\underset{\sim}{Y}$ είναι ένα τ.δ. μεγέθους m από κατανομή $N(\mu_2, \sigma_2^2)$ και τα δείγματα είναι ανεξάρτητα τότε

$$\bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2\,,\ \frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}\right).$$

Απόδειξη

Η απόδειξη είναι άμεση συνέπεια των θεωρημάτων Α.3.1. και 1.3.2.

Θεώρημα 1.4.6

Αν $\underset{\sim}{X}$ είναι τ.δ. μεγέθους n από κατανομή $N(\mu_1, \sigma_1^2)$ με δειγματική διασπορά S_1^2 και $\underset{\sim}{Y}$ είναι τ.δ. μεγέθους m από κατανομή $N(\mu_2, \sigma_2^2)$ με δειγματική διασπορά S_2^2 και τα δύο δείγματα είναι ανεξάρτητα τότε:

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F_{n-1,\, m-1}\,.$$

Απόδειξη

Η απόδειξη είναι άμεση συνέπεια των θεωρημάτων Α.3.7 και 1.3.3.

Θεώρημα 1.4.7

Αν $\underset{\sim}{X}$ και $\underset{\sim}{Y}$ είναι δύο ανεξάρτητα τ.δ. από πληθυσμούς $N(\mu_1, \sigma^2)$ και $N(\mu_2, \sigma^2)$ αντίστοιχα τότε η τ.μ.

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

ακολουθεί κατανομή t_{n+m-2} όπου:

$\bar{X}$ είναι η μέση τιμή, S_1^2 είναι η διασπορά και n το μέγεθος του δείγματος $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$.

$\bar{Y}$ είναι η μέση τιμή, S_2^2 είναι η διασπορά και m το μέγεθος του δείγματος $\underset{\sim}{Y}' = (Y_1, Y_2, \ldots, Y_m)$.

Απόδειξη

Σύμφωνα με τα θεωρήματα 1.3.2 και A.3.1 έχουμε

$$\left.\begin{aligned}\bar{X} &\sim N\left(\mu_1, \frac{\sigma^2}{n}\right) \\ \bar{Y} &\sim N\left(\mu_2, \frac{\sigma^2}{m}\right)\end{aligned}\right. \qquad \bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2\,,\ \sigma^2\left(\frac{1}{n}+\frac{1}{m}\right)\right)$$

Επίσης σύμφωνα με τα θεωρήματα 1.3.3 και A.3.4.

$$\left.\begin{aligned}\frac{(n-1)\,S_1^2}{\sigma^2} &\sim \chi^2_{n-1} \\ \frac{(m-1)\,S_2^2}{\sigma^2} &\sim \chi^2_{m-1}\end{aligned}\right\} \Rightarrow \frac{(n-1)\,S_1^2+(m-2)\,S_2^2}{\sigma^2} \sim \chi^2_{n+m-2} \quad .$$

Έτσι σύμφωνα με το θεώρημα A.3.6 η τ.μ.

$$\frac{\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\dfrac{\frac{(n-1)\,S_1^2+(m-1)\,S_2^2}{\sigma^2}}{n+m-2}}} = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}\left(\frac{1}{n}+\frac{1}{m}\right)}}$$

ακολουθεί την t_{n+m-2} – κατανομή. ο.ε.δ.

ΛΥΜΕΝΕΣ ΑΣΚΗΣΕΙΣ

1.1. Έστω $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ τ.δ. από κατανομή $f_X(x)$. Να βρεθεί η κατανομή της στ.σ. $T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ όταν η κατανομή είναι:

α) Γάμμα $G(\alpha, \beta)$ β) Γεωμετρική
γ) Bernoulli $B(1, p)$ δ) Poisson $P(\lambda)$.

Λύση

Οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες και ισόνομες σαν στοιχεία του τ.δ. $\underset{\sim}{X}$. Συνεπώς η ροπογεννήτρια του αθροίσματος $T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ ισούται με το γινόμενο των ροπογεννητριών των τ.μ. X_i δηλαδή:

$$M_{\Sigma X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) \ .$$

α) Σύμφωνα με τον πίνακα ΑΙΙ η ροπογεννήτρια της κατανομής $G(\alpha, \beta)$ είναι: $M_X(t) = \left(1 - \frac{t}{\beta}\right)^{-\alpha}$. Έτσι:

$$M_{\Sigma X_i}(t) = \prod_{i=1}^{n} \left(1 - \frac{t}{\beta}\right)^{-\alpha} = \left(1 - \frac{t}{\beta}\right)^{-n\alpha} \ .$$

Άρα η στ.σ. $T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ ακολουθεί την κατανομή $G(n\alpha, \beta)$.

β) Σύμφωνα με τον πίνακα ΑΙ , η ροπογεννήτρια της Γεωμετρικής κατανομής είναι $M_X(t) = \frac{pe^t}{1-(1-p)e^t}$. Έτσι:

$$M_{\Sigma X_i}(t) = \prod_{i=1}^{n} \frac{pe^t}{1-(1-p)e^t} = \left(\frac{pe^t}{1-(1-p)e^t}\right)^n \ .$$

Άρα η στ.σ. $T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ ακολουθεί την αρνητική διωνυμική κατανομή, σύμφωνα με τον πίνακα ΑΙ.

γ) Σύμφωνα με τον πίνακα AI η ροπογεννήτρια της κατανομής Bernoulli $B(1, p)$ είναι: $M_X(t) = (1-p+pe^t)$. Έτσι:

$$M_{\Sigma X_i}(t) = \prod_{i=1}^{n} (1-p+pe^t) = (1-p+pe^t)^n$$

Αυτή η ροπογεννήτρια, δεν είναι άλλη από την ροπογεννήτρια της διωνυμικής κατανομής $B(n, p)$.

δ) Η ροπογεννήτρια της κατανομής Poisson $P(\lambda)$ είναι: $M_X(t) = \exp\{\lambda(e^t-1)\}$. Έτσι:

$$M_{\Sigma X_i}(t) = \prod_{i=1}^{n} \exp\{\lambda(e^t-1)\} = \exp\{n\lambda(e^t-1)\}$$

δηλαδή η τ.μ. $T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ ακολουθεί την κατανομή Poisson $P(n\lambda)$.

1.2. Έστω $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)'$ τ.δ. από ομοιόμορφη κατανομή $U(\alpha, \beta)$. Αν $x_{(1)} = \min\{x_i, i=1, 2, \ldots, n\}$ και $x_{(n)} = \max\{x_i, i=1, 2, \ldots, n\}$, να βρεθεί η κατανομή των παρακάτω στ.σ.:

α) $T_1(\underset{\sim}{x}) = X_{(1)}$ **β)** $T_2(\underset{\sim}{x}) = X_{(n)}$

γ) $T_3(\underset{\sim}{x}) = \frac{1}{2}(X_{(1)}+X_{(n)})$ **δ)** $T_3(\underset{\sim}{x}) = X_{(n)}-X_{(1)}$

Λύση

Η σ.π.π και η σ.κ. της ομοιόμορφης κατανομής $U(\alpha, \beta)$ είναι αντίστοιχα:

$$f_X(x) = \frac{1}{\beta-\alpha} \quad \text{και} \quad F_X(x) = \frac{x-\alpha}{\beta-\alpha}, \quad x\in[\alpha, \beta].$$

Σύμφωνα με το θεώρημα (1.3.1) η σ.π.π. των τ.μ. $Y_1=X_{(1)}$ και $Y_n=X_{(n)}$ είναι αντίστοιχα:

$$f_1(y) = n\big(1-F_X(y)\big)^{n-1} f_X(y)$$
$$f_n(y) = n\big(F_X(y)\big)^{n-1} f_X(y)$$

Άρα:

α) $f_1(y) = n\left(1-\frac{y-\alpha}{\beta-\alpha}\right)^{n-1} \frac{1}{\beta-\alpha}$

β) $f_n(y) = n\,\frac{(y-\alpha)^{n-1}}{(\beta-\alpha)^n}$.

Η κοινή σ.π.π. των τ.μ. Y_1, Y_n είναι:

$$f_{1n}(y_1, y_n) = n(n-1)\frac{(y_n-y_1)^{n-2}}{(\beta-\alpha)^n} \qquad \alpha \le y_1 < y_n \le \beta$$

γ) Θεωρούμε δύο νέες τ.μ. Z και W έτσι ώστε:

$$\left.\begin{array}{l} Z=Y_1 \\ W=\frac{1}{2}(Y_1+Y_n) \end{array}\right\} \Rightarrow \left\{\begin{array}{l} Y_1=Z \\ Y_n=2W-Z \end{array}\right. , \quad \alpha \le z < 2w-z \le \beta$$

$$J=\frac{D(Y_1, Y_n)}{D(Z, W)} = \begin{vmatrix} \frac{\partial y_1}{\partial z} & \frac{\partial y_1}{\partial w} \\ \frac{\partial y_n}{\partial z} & \frac{\partial y_n}{\partial w} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = 2$$

$$f_{Z,W}(z, w) = f_{Y_1,Y_n}(z, 2w-z)\,|J| = 2n(n-1)\frac{(2w-2z)^{n-2}}{(\beta-\alpha)^n}$$

$$f_W(w) = \int_\alpha^w f_{Z,W}(z, w)\,dz = 2^{n-1}\frac{n(n-1)}{(\beta-\alpha)^n}\int_\alpha^w (w-z)^{n-2}\,dz = \frac{2^{n-1}n(w-\alpha)^{n-1}}{(\beta-\alpha)^n} \ .$$

Συνεπώς η σ.π.π. της στ.σ. $T_3(\underset{\sim}{X}) = \frac{1}{2}(X_{(1)}+X_{(n)})$ είναι η:

$$f_W(w) = 2^{n-1}\, n\,\frac{(w-\alpha)^{n-1}}{(\beta-\alpha)^n} \qquad \alpha < w < \beta \ .$$

δ) Για να βρούμε τη σ.π.π. της στ.σ. $T_4(\underset{\sim}{X}) = X_{(n)}-X_{(1)}$, θεωρούμε δύο νέες τ.μ. Z και R έτσι ώστε:

$$\left.\begin{array}{l} Z=Y_1 \\ R=Y_n-Y_1 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} Y_1=Z \\ Y_n=R+Z \end{array}\right. , \quad \alpha \le z < z+r \le \beta \ .$$

$$J = \frac{D(Y_1, Y_n)}{D(Z, R)} = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1$$

$$f_{Z,R}(z, r) = f_{Y_1,Y_n}(z, r+z)\,|J| = n(n-1)\frac{r^{n-2}}{(\beta-\alpha)^n}$$

$$f_R(r) = \int_\alpha^\beta f_{Z,R}(z, r)\,dz = \int_\alpha^{\beta-r} n(n-1)\frac{r^{n-2}}{(\beta-\alpha)^n}\,dz = \frac{n(n-1)r^{n-2}(\beta-\alpha-r)}{(\beta-\alpha)^n} \ .$$

Άρα η σ.π.π. της στ.σ. $T_4(\underset{\sim}{x}) = X_{(n)}-X_{(1)}$ είναι η:

$$f_R(r) = \frac{n(n-1)r^{n-2}(\beta-\alpha-r)}{(\beta-\alpha)^n} \qquad 0<r<\beta-\alpha\,.$$

1.3. Έστω $\underset{\sim}{X} = (X_1, X_2, ..., X_n)$ τ.δ. από κατανομή $G(\alpha, \beta)$, $\alpha\in \mathbb{N}$, $\beta>0$. Να δειχθεί ότι η τ.μ. $Z= 2\beta \sum_{i=1}^{n} X_i$ ακολουθεί την $X^2_{2\alpha n}$ - κατανομή.

Λύση

Σύμφωνα με την άσκηση 1.1. η τ.μ. $W= \sum_{i=1}^{n} X_i$ ακολουθεί κατανομή $G(n\alpha, \beta)$. Θεωρούμε τον μετασχηματισμό $Z=2\beta W$. Η ροπογεννήτρια της τ.μ. Z είναι:

$$M_Z(t) = E(e^{tZ}) = E(e^{2\beta t W}) = E(e^{t'W}) = \left(1-\frac{t'}{\beta}\right)^{-n\alpha} =$$

$$= \left(1-\frac{2\beta t}{\beta}\right)^{-n\alpha} = (1-2t)^{-\frac{2n\alpha}{2}}\,.$$

Σύμφωνα με τον πίνακα AII η $M_Z(t) = (1-2t)^{-\frac{2n\alpha}{2}}$ είναι η ροπογεννήτρια της $X^2_{2\alpha n}$ - κατανομής. ο.ε.δ.

1.4. Θεωρούμε δύο ανεξάρτητα τ.δ. $\underset{\sim}{X}' = (X_1, X_2, ..., X_n)$ και $\underset{\sim}{Y}' = (Y_1, Y_2, ..., Y_m)$ από κατανομές $G(1, \beta_1)$ και $G(1, \beta_2)$ αντίστοιχα. Να δειχθεί ότι η τ.μ. $W= \dfrac{\beta_1 \bar{X}}{\beta_2 \bar{Y}}$, όπου $\bar{X}= \dfrac{1}{n} \sum_{i=1}^{n} X_i$ και $\bar{Y} = \dfrac{1}{n} \sum_{j=1}^{m} Y_j$, ακολουθεί την $F_{2n,\, 2m}$ - κατανομή.

Λύση

Σύμφωνα με την άσκηση 1.3 η τ.μ. $Z_1= 2\beta_1 \sum_{i=1}^{n} X_i$ ακολουθεί την X^2_{2n} - κατανομή ενώ η $Z_2= 2\beta_2 \sum_{j=1}^{m} Y_j$ ακολουθεί την X^2_{2m} - κατανομή. Σύμφωνα με το θεώρημα A.3.10 η τ.μ. $W= \dfrac{\frac{Z_1}{2n}}{\frac{Z_2}{2m}}$ ακολουθεί την $F_{2n,\, 2m}$ - κατανομή.

Όμως
$$\frac{\frac{Z_1}{2n}}{\frac{Z_2}{2m}} = \frac{\frac{2\beta_1 \sum_{i=1}^{n} X_i}{2n}}{\frac{2\beta_2 \sum_{j=1}^{m} Y_j}{2m}} = \frac{\beta_1 \overline{X}}{\beta_2 \overline{Y}} \; .$$

Άρα η τ.μ. $W = \dfrac{\beta_1 \overline{X}}{\beta_2 \overline{Y}}$ ακολουθεί την $F_{2n,\,2m}$ - κατανομή ο.ε.δ.

1.5. Να υπολογιστούν τα ολοκληρώματα:

α) $\displaystyle\int_0^{\infty} x^{\alpha-1}\, e^{-\beta x}\, dx \quad , \quad \alpha>0, \quad \beta>0.$

β) $\displaystyle\int_0^{1} x^{\alpha-1}\, (1-x)^{\beta-1}\, dx \quad , \quad \alpha>0\,, \quad \beta>0.$

Λύση

α) Θεωρούμε τη σ.π.π. της Γάμμα κατανομής $G(\alpha, \beta)$.

$$f_X(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta x} \qquad x>0, \quad \alpha>0, \quad \beta>0.$$

Γνωρίζουμε ότι: $\displaystyle\int_0^{\infty} f_X(x)\, dx = 1.$

Άρα
$$\int_0^{\infty} \frac{\beta^{\alpha}}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta x}\, dx = 1 \Rightarrow \int_0^{\infty} x^{\alpha-1}\, e^{-\beta x}\, dx = \frac{\Gamma(\alpha)}{\beta^{\alpha}} \;,$$

όπου η συνάρτηση $\Gamma(\alpha)$ έχει τις εξής ιδιότητες:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1}\, e^{-x}\, dx \qquad \alpha>0$$

$$\Gamma(\alpha) = (\alpha-1)\,\Gamma(\alpha-1)$$

$$\Gamma(\alpha) = (\alpha-1)! \quad \text{αν } \alpha \in \mathbb{N}$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \; .$$

β) Η σ.π.π. της Βήτα κατανομής $\beta(\alpha, \beta)$ είναι:

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, x^{\alpha-1}\,(1-x)^{\beta-1} \quad , \quad x\in(0,\,1)\;, \quad \alpha>0,\, \beta>0.$$

Έτσι: $$\int_0^1 x^{\alpha-1}\,(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)} \qquad \text{ο.ε.δ.}$$

1.6. Αν X_1, X_2 είναι ανεξάρτητες, τ.μ. από κατανομές $G(\alpha_1,\,\beta)$, $G(\alpha_2,\,\beta)$ αντίστοιχα τότε οι τ.μ. $Y_1 = X_1+X_2$ και $Y_2 = \dfrac{X_1}{(X_1+X_2)}$ είναι επίσης ανεξάρτητες και έχουν κατανομή $\Gamma(\alpha_1+\alpha_2,\,\beta)$ και $\beta(\alpha_1,\,\alpha_2)$ αντίστοιχα.

Λύση

Η κοινή κατανομή των τ.μ. $(X_1,\,X_2)$ είναι λόγω της ανεξαρτησίας των X_1 και X_2:

$$f_{X_1,\,X_2}(x_1,\,x_2) = f_{X_1}(x_1)\,f_{X_2}(x_2) = \frac{\beta^{\alpha_1}}{\Gamma(\alpha_1)}\,x_1^{\alpha_1-1}\,e^{-\beta x_1}\,\frac{\beta^{\alpha_2}}{\Gamma(\alpha_2)}\,x_2^{\alpha_2-1}\,e^{-\beta x_2}$$

θεωρούμε τον μετασχηματισμό:

$$\left.\begin{aligned} Y_1 &= X_1+X_2 \\ Y_2 &= \frac{X_1}{X_1+X_2} \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} X_1 &= Y_1\,Y_2 \\ X_2 &= Y_1-Y_1Y_2 \end{aligned}\right.$$

Η Ιακωβιανή είναι:

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\[2ex] \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ 1-y_2 & -y_1 \end{vmatrix} = -y_1\ .$$

Τότε

$$\begin{aligned} g_{Y_1,\,Y_2}(y_1,\,y_2) &= |J|\,f_{X_1,\,X_2}(y_1\,y_2,\,y_1(1-y_2)) = \\ &= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\,(y_1\,y_2)^{\alpha_1-1}\,(y_1(1-y_2))^{\alpha_2-1}\,e^{-\beta y_1}\,y_1 = \\ &= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}\,y_1^{\alpha_1+\alpha_2-1}\,e^{-\beta y_1}\cdot\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\,y_2^{\alpha_1-1}\,(1-y_2)^{\alpha_2-1} = \\ &= g_{Y_1}(y_1)\cdot g_{Y_2}(y_2) \quad , \quad y_1>0\;,\; 0<y_2<1\;. \end{aligned}$$

δηλαδή οι τ.μ. Y_1 και Y_2 είναι ανεξάρτητες και ακολουθούν κατανομές $\Gamma(\alpha_1+\alpha_2,\,\beta)$ και $\beta(\alpha_1,\,\alpha_2)$ αντίστοιχα. ο.ε.δ.

1.7. Αν η τ.μ. $Z=\dfrac{mS^2}{\sigma^2}$ ακολουθεί την X^2_m - κατανομή και η δεσμευμένη κατανομή της τ.μ. Y όταν γνωρίζουμε το S^2 είναι η $N\left(0,\dfrac{\sigma^2}{S^2}\right)$, τότε η τ.μ. Y ακολουθεί την t_m- κατανομή του Student.

Λύση

Από τα δεδομένα του προβλήματος έχουμε ότι:

$$f_Z(z) = \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} z^{\frac{m}{2}-1} \exp\left\{-\frac{z}{2}\right\} \quad , \; z>0 .$$

Η σ.π.π. της τ.μ. $S^2=\dfrac{\sigma^2 Z}{m}$ είναι:

$$f_{S^2}(s^2) = f_Z(s^2)\left|\frac{dz}{ds^2}\right| = \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\left(\frac{ms^2}{\sigma^2}\right)^{\frac{m}{2}-1} \exp\left\{-\frac{ms^2}{2\sigma^2}\right\}\frac{m}{\sigma^2} \quad , \; s^2>0 .$$

Η δεσμευμένη σ.π.π. της τ.μ. Y όταν δίνεται το S^2 είναι:

$$f_{y/s^2}(y/s^2) = \frac{S}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{s^2y^2}{2\sigma^2}\right\} \quad , \; y\in R.$$

Όμως:

$$f_{y/s^2}(y/s^2) = \frac{f_{Y,S^2}(y,s^2)}{f_{S^2}(s^2)} \qquad \text{και} \quad f_{Y,S^2}(y,s^2) = f_{y/s^2}(y/s^2)\cdot f_{S^2}(s^2) .$$

Η κατανομή της τ.μ. Y θα δίνεται από τον τύπο:

$$f_Y(y) = \int_0^\infty f_{Y,S^2}(y,s^2)\,ds^2 = \int_0^\infty f_{y/s^2}(y/s^2)\,f_{S^2}(s^2)\,ds^2 =$$

$$= \int_0^\infty \frac{m^{\frac{m+1}{2}}}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)}\left(\frac{s^2}{2\sigma^2}\right)^{\frac{m+1}{2}-1}\exp\left\{-\frac{s^2}{2\sigma^2}(y^2+m)\right\} d\,\frac{s^2}{2\sigma^2} =$$

$$= \frac{m^{\frac{m+1}{2}}}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)}\int_0^\infty \omega^{\frac{m+1}{2}-1}\exp\{-(y^2+m)\omega\}\,d\omega =$$

$$= \frac{1}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)}\left(\frac{m}{m+y^2}\right)^{\frac{m+1}{2}}\Gamma\left(\frac{m+1}{2}\right) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)}\left(1+\frac{y^2}{m}\right)^{-\frac{m+1}{2}}, \quad y\in R$$

ο.ε.δ.

1.8. Έστω $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$ τ.δ. από κατανομή με σ.π.π.

$$f(x;\theta) = \exp\{-(x-\theta)\} \quad x\geq\theta>0.$$

Θεωρούμε το διατεταγμένο δείγμα $\underset{\sim}{Y}' = (Y_1, Y_2, \ldots, Y_n)$ όπου $Y_i=X_{(i)}$ $i=1, 2, \ldots, n$. Να δειχθεί ότι οι τ.μ. Y_s-Y_1, $s=2, 3, \ldots, n$ και Y_1 είναι ανεξάρτητες, και να υπολογισθεί η κατανομή της τ.μ. Y_s-Y_1.

Λύση

Η κοινή κατανομή των τ.μ. Y_s και Y_1 δίνεται από τον τύπο:

$$f_{1s}(x, y) = \frac{n!}{(s-2)!\,(n-s)!}\,(F(y)-F(x))^{s-2}\,(1-F(y))^{n-s}\,f(x)\,f(y) \qquad x<y \quad .$$

Στο πρόβλημά μας $F(x) = \int_\theta^x \exp\{-(t-\theta)\}\,dt = 1-\exp\{-(x-\theta)\}$.

Οπότε:

$$f_{1s}(x, y) = \frac{n!}{(s-2)!\,(n-s)!}\,(\exp\{-(x-\theta)\}-\exp\{-(y-\theta)\})^{s-2}\exp\{-(y-\theta)\}^{n-s+1}$$

$$\cdot\exp\{-(x-\theta)\}\ .$$

Στη συνέχεια θεωρούμε τον μετασχηματισμό:

$$\left.\begin{array}{l} Z=Y_1 \\ W= Y_S-Y_1 \end{array}\right\} \Rightarrow \left.\begin{array}{l} Y_1=Z \\ Y_s= Z+W \end{array}\right\} \Rightarrow J = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1.$$

$$f_{Z,W}(z, w) = f_{1s}(z, z+w)\,|J| =$$

$$= \frac{n!}{(s-2)!\,(n-s)!}\exp\{-(z-\theta)\}^n\,(1-\exp\{-w\})^{s-2}\exp\{-w\}^{n-s+1} =$$

$$= n\exp\{-n(z-\theta)\}\,\frac{(n-1)!}{(s-2)!\,(n-s)!}\,(1-\exp\{-w\})^{s-2}\exp\{-w\}^{n-s+1}\ .$$

$$f_W(w)= \int_\theta^\infty f_{Z,W}(z, w)\,dz = \frac{(n-1)!}{(s-2)!\,(n-s)!}\,(1-\exp\{-w\})^{s-2}\exp\{-w\}^{n-s+1}, \quad 0\leq w<\infty$$

$$f_Z(z) = \int_\theta^\infty f_{Z,W}(z, w)\,dw = n\exp\{-(z-\theta)\}^n, \quad z\geq\theta\ .$$

Παρατηρούμε ότι $f_{Z,W}(z, w) = f_Z(z) f_W(w)$, συνεπώς οι τ.μ. Y_1 και $Y_s - Y_1$, $s=2, \ldots, n$ είναι ανεξάρτητες. ο.ε.δ.

1.9. Έστω $\underset{\sim}{X}' = (X_1, X_2, \ldots, X_n)$ τ.δ. από εκθετική κατανομή με σ.π.π.:

$$f(x, \underset{\sim}{\theta}) = \frac{1}{\theta_1} \exp\left\{-\frac{x-\theta_2}{\theta_1}\right\} I_{(\theta_2, \infty)}(x) .$$

Να δειχθεί ότι οι στ.σ. $T_1 = \sum_{i=1}^{n} (X_i - X_{(1)})$ και $T_2 = nX_{(1)}$ όπου $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$ είναι ανεξάρτητες και ακολουθούν κατανομές: $G\left(n-1, \frac{1}{\theta_1}\right)$, και εκθετική με σ.π.π.

$$f(z_1, \underset{\sim}{\theta}) = \frac{1}{\theta_1} \exp\left\{-\frac{z_1 - n\theta_2}{\theta_1}\right\} I_{(n\theta_2, \infty)}(z_1) \text{ αντίστοιχα.}$$

Λύση

Κατ' αρχήν παρατηρούμε ότι $\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} X_{(i)}$ όπου $X_{(i)}$ είναι η i παρατήρηση στο διατεταγμένο δείγμα. Συνεπώς

$$T_1 = \sum_{i=1}^{n} X_i - nX_{(1)} = \sum_{i=1}^{n} X_{(i)} - nX_{(1)} ,$$

γι' αυτό θα εργασθούμε για τις στ.σ. $T_1 = \sum_{i=1}^{n} X_{(i)} - nX_{(1)}$ και $T_2 = nX_{(1)}$. Χάριν απλότητας βάζουμε $X_{(i)} = Y_i$ $i=1, \ldots, n$. Η κοινή κατανομή του διατεταγμένου δείγματος $\underset{\sim}{Y}' = (Y_1, Y_2, \ldots, Y_n)$ είναι:

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = n! f_{\underset{\sim}{X}}(\underset{\sim}{y}) = n! \frac{1}{\theta_1^n} \exp\left\{-\frac{1}{\theta_1} \sum_{i=1}^{n} (y_i - \theta_2)\right\} , \quad y_i > \theta_2 , \; i=1, 2, \ldots, n .$$

Θεωρούμε τον μετασχηματισμό: $\underset{\sim}{Z} = A\underset{\sim}{Y}$ έτσι ώστε:

$$\left.\begin{aligned} Z_1 &= nY_1 \\ Z_2 &= (n-1)(Y_2 - Y_1) \\ Z_3 &= (n-2)(Y_3 - Y_2) \\ &\vdots \\ Z_n &= Y_n - Y_{n-1} \end{aligned}\right\} \Rightarrow \sum_{i=1}^{n} Z_i = \sum_{i=1}^{n} Y_i .$$

Ο πίνακας του μετασχηματισμού είναι:

$$A=\begin{pmatrix} n & 0 & 0 & \dots & 0 \\ -(n-1) & n-1 & 0 & \dots & 0 \\ 0 & -(n-2) & n-2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \Rightarrow |A|=n!$$

Η κοινή κατανομή της τ.μ. $\underset{\sim}{Z}' = (Z_1, Z_2, \dots, Z_n)$ είναι:

$$f_{\underset{\sim}{Z}}(\underset{\sim}{z}) = f_{\underset{\sim}{Y}}(\underset{\sim}{z})\,|A^{-1}| = \frac{1}{\theta_1^n}\exp\left\{-\frac{1}{\theta_1}\sum_{i=1}^{n}(z_i-n\theta_2)\right\} =$$

$$= \frac{1}{\theta_1}\exp\left\{-\frac{z_1-n\theta_2}{\theta_1}\right\}\prod_{i=2}^{n}\frac{1}{\theta_1}\exp\left\{-\frac{z_i}{\theta_1}\right\}, \quad z_i\geq 0,\ i=2,\dots,n, \quad z_1\geq n\theta_2 .$$

Οι κατανομές των τ.μ. $Z_1, Z_2, \dots, Z_n$ βρίσκονται ως εξής:

$$f_{Z_1}(z_1) = \frac{1}{\theta_1}\exp\left\{-\frac{z_1-n\theta_2}{\theta_1}\right\}\int_0^{\infty}\dots\int_0^{\infty}\frac{1}{\theta_1^{n-1}}\exp\left\{-\frac{1}{\theta_1}\sum_{i=2}^{n}z_i\right\}dz_2\dots dz_n =$$

$$= \frac{1}{\theta_1}\exp\left\{-\frac{z_1-n\theta_2}{\theta_1}\right\}, \quad z_1\geq n\theta_2 .$$

$$f_{Z_i}(z_i) = \int_{n\theta_2}^{\infty}\int_0^{\infty} f_{\underset{\sim}{Z}}(\underset{\sim}{z})\,dz_1\dots dz_{i-1}\,dz_{i+1}\dots dz_n = \frac{1}{\theta_1}\exp\left\{-\frac{z_i}{\theta_1}\right\}, z_i\geq 0,\ i=2,\dots,n.$$

Εύκολα διαπιστώνεται ότι οι τ.μ. $Z_1, Z_2, \dots, Z_n$ είναι ανεξάρτητες διότι:

$$f_{\underset{\sim}{Z}}(\underset{\sim}{z}) = \prod_{i=1}^{n} f_{Z_i}(z_i) .$$

Η κατανομή των τ.μ. $Z_2, \dots, Z_n$ είναι η $G\left(1, \frac{1}{\theta_1}\right)$. Άρα η τ.μ. $\sum_{i=2}^{n} Z_i$ ακολουθεί κατανομή $G\left(n-1, \frac{1}{\theta_1}\right)$.

Όμως:

$$\sum_{i=2}^{n} Z_i = \sum_{i=1}^{n} Z_i - Z_1 = \sum_{i=1}^{n} Y_i - nY_1 = \sum_{i=1}^{n} X_{(i)} - nX_{(1)} = T_1$$

$$Z_1 = nY_1 = nX_{(1)} = T_2 \quad .$$

Συνεπώς οι στ.σ. T_1 και T_2 είναι ανεξάρτητες αφού η T_2 είναι συνάρτηση της τ.μ. Z_1 και η T_2 είναι συνάρτηση των τ.μ. $Z_2, \ldots, Z_n$ και η T_2 ακολουθεί κατανομή Γάμμα $G\left(n-1, \frac{1}{\theta_1}\right)$ και η T_1 ακολουθεί εκθετική κατανομή με σ.π.π. $f(z_1, \underset{\sim}{\theta}) = \frac{1}{\theta_1} \exp\left\{-\frac{z_1 - n\theta_2}{\theta_1}\right\}$, $z_1 \geq n\theta_2$. ο.ε.δ.

1.10. Έστω $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ τ.δ. από διδιάστατη κανονική κατανομή $N(\underset{\sim}{\mu}, \Sigma)$ όπου $\underset{\sim}{\mu}' = (\mu_X, \mu_Y)$ και

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} \text{με } \sigma_{XY} = \varrho\sigma_X\sigma_Y, \ (\mu_X, \mu_Y) \in \mathbf{R}^2,$$

$(\sigma_X^2, \sigma_Y^2) \in (\mathbf{R}^+ \times \mathbf{R}^+)$, $\varrho \in [-1, 1]$. Να δειχθεί ότι η τ.μ. $(\bar{X}, \bar{Y})$ ακολουθεί διδιάστατη κανονική κατανομή $N\left(\underset{\sim}{\mu}, \frac{1}{n}\Sigma\right)$.

Λύση

Η ροπογεννήτρια της κατανομής $N(\underset{\sim}{\mu}, \Sigma)$ είναι:

$$M_{X,Y}(t_1, t_2) = Ee^{t_1X + t_2Y} = \exp\left\{\mu_X t_1 + \mu_Y t_2 + \frac{1}{2}(\sigma_X^2 t_1^2 + 2\varrho\sigma_X\sigma_Y t_1 t_2 + \sigma_Y^2 t_2^2)\right\} .$$

Η ροπογεννήτρια της τ.μ. $(\bar{X}, \bar{Y})$ θα είναι:

$$M_{\bar{X},\bar{Y}}(t_1, t_2) = Ee^{t_1\bar{X} + t_2\bar{Y}} = E\exp\left\{\frac{t_1}{n}\sum_{i=1}^{n} X_i + \frac{t_2}{n}\sum_{i=1}^{n} Y_i\right\} =$$

$$= E\exp\left\{\sum_{i=1}^{n}\left(\frac{t_1}{n}X_i + \frac{t_2}{n}Y_i\right)\right\} .$$

Λόγω της ανεξαρτησίας των τ.μ. $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ ισχύει:

$$M_{\bar{X},\bar{Y}}(t_1, t_2) = E\exp\left\{\sum_{i=1}^{n}\left(\frac{t_1}{n}X_i + \frac{t_2}{n}Y_i\right)\right\} = \prod_{i=1}^{n} E\exp\left\{\frac{t_1}{n}X_i + \frac{t_2}{n}Y_i\right\} =$$

$$= \prod_{i=1}^{n} \exp\left\{\mu_X\frac{t_1}{n} + \mu_Y\frac{t_2}{n} + \frac{1}{2}\left(\sigma_X^2\left(\frac{t_1}{n}\right)^2 + 2\varrho\sigma_X\sigma_Y\frac{t_1t_2}{n^2} + \sigma_Y^2\left(\frac{t_2}{n}\right)^2\right)\right\} =$$

$$= \exp\left\{\mu_X t_1 + \mu_Y t_2 + \frac{1}{2}\left(\frac{\sigma_X^2}{n} t_1 + 2\varrho\,\frac{\sigma_X\sigma_Y}{n} t_1 t_2 + \frac{\sigma_Y^2}{n} t_2^2\right)\right\} .$$

Άρα η τ.μ. $(\bar{X},\ \bar{Y})$ ακολουθεί τη διδιάστατη κανονική κατανομή $N\left(\underset{\sim}{\mu},\ \frac{1}{n}\Sigma\right)$ δηλαδή:

$$\mu_X=\mu_{\bar{X}}\,, \qquad \sigma^2_{\bar{X}}=\frac{1}{n}\sigma^2_X\,, \qquad \sigma_{\bar{X}\bar{Y}}=\frac{1}{n}\sigma_X\sigma_Y$$

$$\mu_Y=\mu_{\bar{Y}}\,, \qquad \sigma^2_{\bar{Y}}=\frac{1}{n}\sigma^2_Y$$