Τμήμα Δασολογίας & Διαχείρισης Περιβάλλοντος & Φυσικών Πόρων
Εργαστήριο Διευθέτησης Ορεινών Υδάτων και Διαχείρισης Κινδύνου
Μεταπτυχιακό Πρόγραμμα Σπουδών

ΥΔΡΟΛΟΓΙΑ ΦΥΣΙΚΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ

Κεφάλαιο 2ο: Ανάλυση Υδρολογικής Πληροφορίας

Φ. Π. Μάρης
Αναπλ. Καθηγητής

2.1 ΕΙΣΑΓΩΓΗ

- Η ανάλυση συχνότητας της υδρολογικής πληροφορίας συμβάλλει ουσιαστικά στον ορθολογικό σχεδιασμό των έργων που έχουν άμεση ή έμμεση σχέση με το νερό. Παραδείγματα τέτοιων έργων που πρέπει να σχεδιάζονται με βάση τη σχέση μεγέθους - συχνότητας εμφάνισης του υδρολογικού φαινομένου είναι μεταξύ των άλλων οι εκχειλιστές ασφαλείας των φραγμάτων, τα δίκτυα αποχέτευσης ομβρίων νερών, τα στραγγιστικά δίκτυα, οι γέφυρες και οι αντιπλημμυρικές κατασκευές.

- Η στατιστική ανάλυση της υδρολογικής πληροφορίας δεν συνεισφέρει μόνο στο σχεδιασμό υδραυλικών κατασκευών με μικρή επικινδυνότητα λειτουργίας αλλά κυρίως στον αποδοτικό σχεδιασμό των έργων. Ένας πραγματικά αποδοτικός σχεδιασμός πρέπει να προέρχεται από συνδυασμό ανάλυσης συχνότητας εμφάνισης του υδρολογικού φαινομένου με την οικονομική ανάλυση (ανάλυση κόστους - ωφέλειας) και την ανάλυση των συνολικών επιπτώσεων από την κατασκευή του έργου (περιβαλλοντικών, κοινωνικών, αισθητικής τοπίου, κλπ).

- Η μεθοδολογία που παρουσιάζεται στις επόμενες παραγράφους έχει το πλεονέκτημα να συνδέει κάθε μέγεθος με τον αντίστοιχο βαθμό επικινδυνότητας πράγμα που δεν θα μπορούσε να δώσει οποιαδήποτε άλλη αιτιοκρατική εμπειρική ανάλυση.
- Η ανάλυση συχνότητας των υδρολογικών φαινομένων προϋποθέτει ότι οι παρατηρήσεις είναι τυχαίες με τη στατιστική έννοια του όρου και ομογενείς (χωρίς δηλαδή να έχει επέλθει καμιά μεταβολή κατά τη διάρκεια της περιόδου των παρατηρήσεων). Οι έλεγχοι που αφορούν στην τυχαιότητα και την ομογένεια των παρατηρήσεων παρουσιάζονται στο δεύτερο μέρος του κεφαλαίου αυτού.

2.2 ΧΡΟΝΟΣΕΙΡΕΣ ΥΔΡΟΛΟΓΙΚΩΝ ΦΑΙΝΟΜΕΝΩΝ

Ορισμός: Οι μετρήσεις ενός υδρολογικού φαινομένου διαταγμένες κατά χρονολογική σειρά αποτελούν μια χρονοσειρά (ή σειρά) του φαινομένου.

Οι παρατηρήσεις συνεχούς διάρκειας όπως τα ημερήσια ύψη βροχής, οι ημερήσιες παροχές, οι μηνιαίες τιμές της εξατμισοδιαπνοής αποτελούν τις *Πλήρεις Σειρές* των αντίστοιχων υδρολογικών μεγεθών. Οι πλήρεις σειρές επιδέχονται ανάλυση είτε στην αρχική τους μορφή ή με την επιλογή ορισμένων μεγεθών που έχουν σχέση με το πρόβλημα που εξετάζεται. Στην τελευταία περίπτωση έχουμε τις μερικές και τις ετήσιες σειρές. Οι *Μερικές Σειρές* αποτελούνται από τιμές του υδρολογικού φαινομένου πάνω ή κάτω από μια τιμή (μερικές σειρές υπερβάσεως ή μη υπερβάσεως αντίστοιχα) που ταιριάζει στο πρόβλημα, ενώ οι *ετήσιες σειρές* είναι οι σειρές με ετήσιες τιμές ή ετήσιους μέσους όρους ή προκύπτουν από την επιλογή των ετήσιων μεγίστων ή ελαχίστων. Στις μερικές και τις ετήσιες σειρές θεωρείται ότι επαληθεύεται η συνθήκη της τυχαιότητας των παρατηρήσεων. Ειδικά όμως στην περίπτωση των μερικών σειρών πρέπει να εξασφαλίζεται η ανεξαρτησία των παρατηρήσεων. Αν και οι ετήσιες σειρές περιέχουν συνήθως λιγότερη πληροφορία από ότι γενικά οι μερικές σειρές, χρησιμοποιούνται κυρίως λόγω της ευκολίας στη χρήση. Τέλος αν η τιμή που επιλέγεται για μια μερική σειρά είναι τέτοια ώστε ο αριθμός των παρατηρήσεων της μερικής σειράς να είναι ίσος με τον αριθμό των ετών των παρατηρήσεων τότε η μερική σειρά λέγεται *Μερική Σειρά Ετησίων Υπερβάσεων.* Στο Σχ. 2.1 φαίνεται η διαδικασία με την οποία από μια πλήρη σειρά (μηνιαίων) τιμών ενός υδρολογικού φαινομένου προκύπτει η ετήσια σειρά μεγίστων και η μερική σειρά ετησίων υπερβάσεων

2.3 ΠΕΡΙΟΔΟΣ ΕΠΑΝΑΦΟΡΑΣ

Έστω η ετήσια σειρά μεγίστων χ_i ,i = 1 (1) *N* διατάσσεται κατά φθίνουσα σειρά μεγέθους.

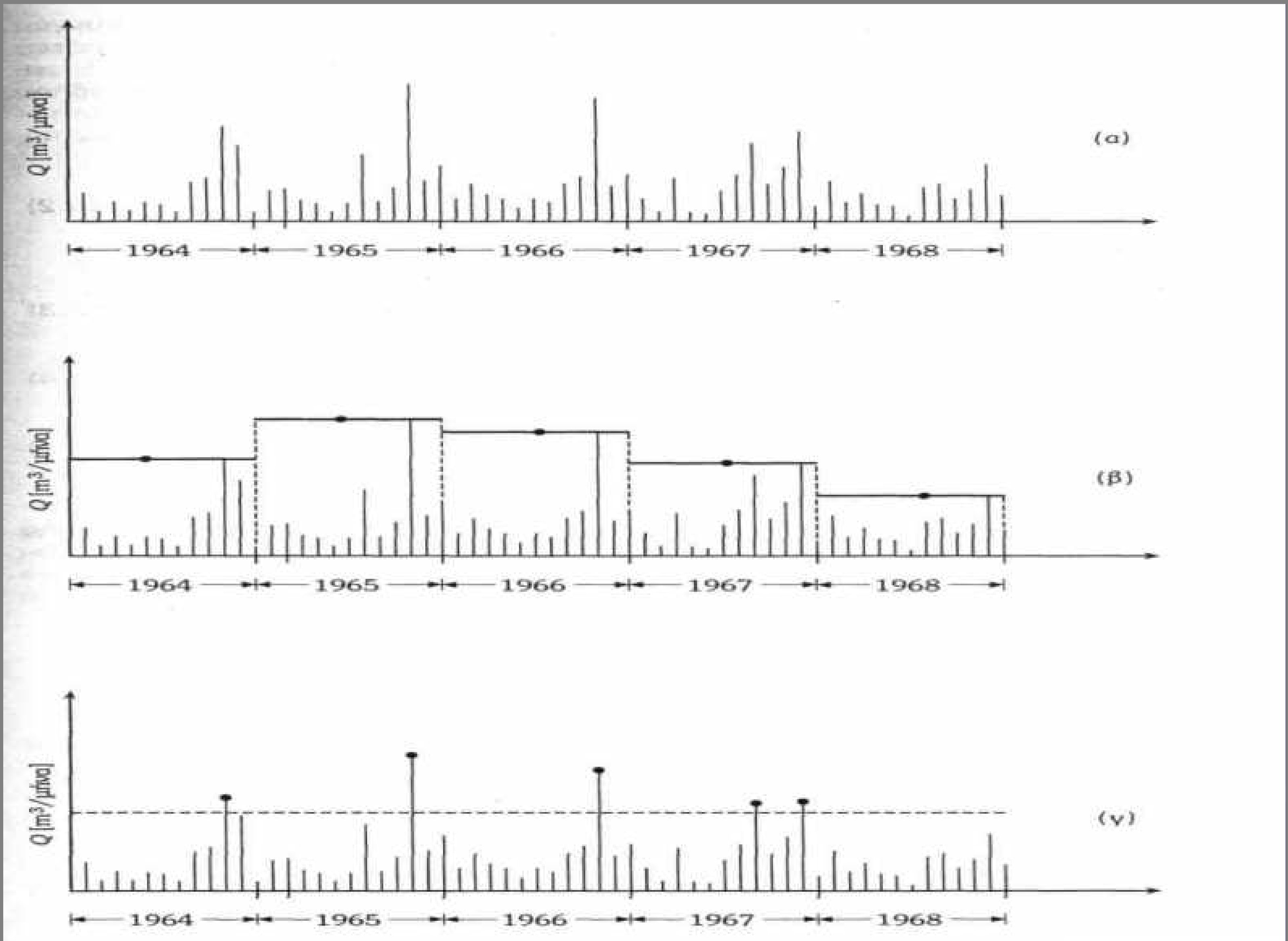
Δηλαδή $\chi_1 > \chi_2 > \chi_3 > \ldots > X_\nu$ Τότε η συχνότητα υπερβάσεως του μεγέθους *χ* είναι:

$$P(x \geq x) = \frac{m}{N} \qquad (2.1)$$

όπου *N* είναι ο αριθμός των παρατηρήσεων, *m* είναι η σειρά του μεγέθους χ στην φθίνουσα σειρά χ_χ, χ_2,...,χ_N. Η εξίσωση αυτή εφαρμόζεται ικανοποιητικά στις λεγόμενες κλειστές σειρές στις οποίες δεν υπάρχει πιθανότητα να εμφανισθεί κάποιο μέγεθος έξω από τα όρια της μεταβλητής του δείγματος σ' ένα άλλο δείγμα ή στο μέλλον. Όμως οι χρονοσειρές στην υδρολογία είναι συνήθως ανοικτές. Γιαυτό η εκτίμηση της πιθανότητας υπερβάσεως γίνεται με την ακόλουθη εξίσωση:

$$P(x \geq x) = \frac{m}{N+1} \qquad (2.2)$$

Weibull(1939)



Σχ. 2.1: *(α) Πλήρης σειρά, (β) Ετήσια σειρά μεγίστων, (γ) Μερική σειρά ετησίων υπερβάσεων.*

Άλλες εξισώσεις που επίσης χρησιμοποιούνται είναι:

$$P(x \geq x) = \frac{2m-1}{2N}$$ Hazen(1930) (2.3)

$$P(x \geq x) = \frac{m-0.3}{N+0.4}$$ Chegodayev(1955)(2.4)

$$P(x \geq x) = \frac{3m-1}{3N+1}$$ Tukey(1962)(2.5)

Παράδειγμα. Στην ακόλουθη σειρά, που είναι διατεταγμένη κατά φθίνουσα τάξη μεγέθους, να υπολογισθεί η πιθανότητα υπέρβασης της τιμής 324.

$$[x_i]_{i=1}^{i=21} = 741, 724, 666, \ldots, 341, 324$$

Σύμφωνα με την Εξ.2.1 $P(X \geq x_{21}) = \frac{21}{21} = 1$

δηλαδή η πιθανότητα να εμφανισθεί σ' άλλο δείγμα ή στο μέλλον τιμή μικρότερη από το 324 είναι μηδενική. Όμως σύμφωνα με την Εξ. 2.2

$$P(X \geq x_{21}) = \frac{21}{22} = 0.955$$

δηλαδή υπάρχει μια πιθανότητα 0.045 να εμφανισθεί τιμή μικρότερη από το 324.

Ορισμός: Περίοδος επαναφοράς δοθέντος ετήσιου μεγέθους ενός φαινομένου είναι το Μέσο χρονικό διάστημα *T* (έτη) μέσα στο οποίο το θεωρούμενο υδρολογικό φαινόμενο θα εμφανισθεί μια μόνο φορά με τιμή ίση ή μεγαλύτερη της δοθείσας. Αν η σειρά που εξετάζεται είναι σειρά ετησίων μεγίστων η περίοδος επαναφοράς αναφέρεται στο μέγεθος του φαινομένου ως ετήσιο μέγιστο, πράγμα που δεν συμβαίνει αν η σειρά είναι μερική.

Σύμφωνα με τον ορισμό προκειμένου για ανάλυση μεγίστων τιμών ενός φαινομένου:

$$P(X \geq x) = \frac{1}{T} \quad \text{ή} \quad T = \frac{1}{P(X \geq x)} \qquad (2.6)$$

Στην περίπτωση των σειρών ελαχίστων τιμών ενός φαινομένου ισχύει:

$$P(X \leq x) = \frac{1}{T} \quad \text{ή} \quad T = \frac{1}{P(X \leq x)} \qquad (2.7)$$

Αν και υπάρχει μια μικρή διαφορά στο νόημα της περιόδου επαναφοράς μερικής σειράς, (T_P), από την περίοδο επαναφοράς μιας ετήσιας σειράς (Γ), στην πράξη δεν γίνεται διάκριση μια που τα αποτελέσματα δεν διαφέρουν ουσιαστικά για περιόδους μεγαλύτερες των πέντε ετών. Γενικά υπάρχει η ακόλουθη σχέση μεταξύ των δύο περιόδων επαναφοράς:

$$T_p = \frac{1}{\ln T - \ln(T-1)} \qquad (2.8)$$

Όπως προκύπτει από τον ορισμό της περιόδου επαναφοράς για τα μέγιστα Εξ. 2.6, η πιθανότητα υπέρβασης σε ένα έτος είναι *1/ T*. Αν θεωρήσουμε ότι το γεγονός της υπέρβασης (ή μη υπέρβασης) είναι ανεξάρτητο για κάθε έτος τότε η πιθανότητα μη υπέρβασης του μεγέθους *χ* και στα *n* έτη ζωής του έργου είναι

$$P(X \le x)_n = (P(X \le x))^n = (1 - P(X > x))^n \qquad (2.9)$$

Η πιθανότητα υπέρβασης (έστω και μία φορά) στα *η* έτη λόγω της Εξ. 2.9 είναι

$$P(X \ge x)_n = 1 - P(X \le x)_n = 1 - (1 - P(X > x))^n \qquad (2.10)$$

Η Εξ.2.10 λόγω της εξ.2.6 γίνεται

$$P(X \ge x)_n = 1 - \left(1 - \frac{1}{T}\right)^n \qquad (2.11)$$

Από την τελευταία σχέση που εκφράζει την πιθανότητα υπέρβασης στην περίοδο *η* ετών μπορεί να προκύψει η περίοδος σχεδιασμού *η*

$$n = \frac{\log(1 - P(X \ge x)_n}{\log\left(1 - \frac{1}{T}\right)} \qquad (2.12)$$

Ορισμός: *Περίοδος Σχεδιασμού, η,* είναι η περίοδος σε έτη κατά την οποία πιθανότητα υπέρβασης του μεγέθους χ του υδρολογικού φαινομένου (με περίοδο επαναφοράς Τ) δεν υπερβαίνει τη δοθείσα τιμή πιθανότητας.

Παράδειγμα 1. Ποια είναι η πιθανότητα υπέρβασης σε ένα έτος μιας τιμής παροχής ενός ρεύματος που έχει περίοδο επαναφοράς 20 έτη;

$$P(X \geq x) = \frac{1}{T} = \frac{1}{20} = 0.05 \quad ή \quad 5\%$$

Παράδειγμα 2. Ποια είναι η πιθανότητα η πλημμυρική απορροή των 100 ετών να μην υπερβληθεί στα επόμενα δύο έτη;

$$P(X \leq x)_n = \left(1 - \frac{1}{T}\right)^n = \left(1 - \frac{1}{100}\right)^2 = 0.9801 \sim 98\%$$

Παράδειγμα 3. Ποια είναι η πιθανότητα υπέρβασης σε διάστημα 3 συνεχών συγκεκριμένων ετών της παροχής που έχει περίοδο επαναφοράς 20 έτη;

$$P(X \geq x)_n = \left(1 - \frac{1}{T}\right)^n = \left(1 - \frac{1}{20}\right)^3 = 0.143. ή 14,3\%$$

2.4 ΠΡΑΚΤΙΚΟΣ ΤΡΟΠΟΣ ΑΝΑΛΥΣΗΣ ΣΥΧΝΟΤΗΤΑΣ

Στη περίπτωση που το υπό μελέτη έργο είναι δευτερεύουσας σημασίας, η περίοδος επαναφοράς μικρή (T < 10) και μικρότερη από την περίοδο των παρατηρήσεων και η χρονοσειρά είναι σχετικά μικρή (αριθμός παρατηρήσεων < 30) είναι δυνατόν η ανάλυση συχνότητας του υδρολογικού φαινομένου να γίνει με την εξής απλή διαδικασία:

i. Επιλέγονται οι τιμές που θα αποτελέσουν το δείγμα ώστε να είναι σχετικές με το υπό μελέτη πρόβλημα.

ii. Διατάσσονται κατά φθίνουσα σειρά μεγέθους προκειμένου για μέγιστα και σε αύξουσα σειρά προκειμένου για ελάχιστα.

iii. Υπολογίζονται οι περίοδοι επαναφοράς και οι πιθανότητες υπέρβασης χρησιμοποιώντας τις σχέσεις 2.2 και 2.6 για μέγιστα και τις σχέσεις 2.2 και 2.7 για ελάχιστα.

Παρατήρηση: Αν δύο διαδοχικές παρατηρήσεις στη φθίνουσα (ή αύξουσα) σειρά είναι ίσες υπάρχουν δύο τρόποι να υπολογισθεί η πιθανότητα υπέρβασης της τιμής αυτής:

(α) να δοθεί σειρά m και $m + 1$ στις δύο ίσες παρατηρήσεις με αποτέλεσμα για το ίδιο μέγεθος να υπάρχουν δύο διαφορετικές πιθανότητες υπέρβασης.

(β) να δοθεί και στις δύο παρατηρήσεις σειρά $\frac{2m + 1}{2}$

Από τους παραπάνω δύο τρόπους ο δεύτερος είναι προτιμώτερος.

ίν. Οι εκτιμήσεις των σχέσεων μεγέθους - περιόδου επαναφοράς για την περίοδο που καλύπτεται από το δείγμα γίνεται με γραμμική παρεμβολή στα μεγέθη που έχουν προκύψει στο βήμα iii της διαδικασίας ή με εμπειρική προέκταση για περιόδους λίγο μεγαλύτερες του δείγματος.

Είναι προφανές ότι η μέθοδος δεν ενδείκνυται για προβλήματα που αναφέρονται στην εκτίμηση μεγέθους πολύ μεγαλύτερης περιόδου επαναφοράς σε σχέση με την περίοδο που καλύπτει το δείγμα.

2.5 ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΜΕ ΤΗ ΧΡΗΣΗ ΣΥΝΑΡΤΗΣΕΩΝ ΚΑΤΑΝΟΜΗΣ ΠΙΘΑΝΟΤΗΤΑΣ - ΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ

Το σύνηθες πρόβλημα στην ανάλυση συχνότητας των υδρολογικών φαινομένων είναι ότι η περίοδος επαναφοράς είναι μεγαλύτερη (αν όχι πολλαπλάσια) από το μέγεθος του δείγματος.

Επομένως η μέθοδος της προηγούμενης παραγράφου δεν μπορεί να εφαρμοσθεί. Σ' αυτή την περίπτωση καταφεύγουμε στις γνωστές συναρτήσεις κατανομής πιθανότητας. Όταν οι παράμετροι μιας τέτοιας συνάρτησης κατανομής προσδιορισθούν από τις υπάρχουσες παρατηρήσεις τότε είναι δυνατή η εκτίμηση της τιμής του φαινομένου για μεγάλες περιόδους επαναφοράς με τη χρησιμοποίηση της θεωρητικής κατανομής.

Υπάρχουν οι εξής αναλυτικές τεχνικές εκτιμήσεως των παραμέτρων των θεωρητικών Κατανομών πιθανότητας:

1. Η μέθοδος των Ροπών (method of moments)
2. Η μέθοδος του μέγιστου πιθανοφάνειας (method of maximum likelihood)
3. Η μέθοδος των ελαχίστων τετραγώνων (method of least squares)
4. Η μέθοδος του ελαχίστου χ^2 (minimum x^2 method)
5. Η μέθοδος των έξι διαιρέσεων (method of sextiles)
6. Η μέθοδος επιλογής ειδικών σημείων (method of matching selected points).

Εκτενής περιγραφή των μεθόδων αυτών μπορεί να βρεθεί στα βιβλία στατιστικής. Στην υδρολογία αντί των παραπάνω αναλυτικών μεθόδων με εξαίρεση ίσως τη μέθοδο των Ροπών, χρησιμοποιούνται κυρίως οι εξής μέθοδοι:

(i) Η γραφική μέθοδος - χαρτί πιθανότητας

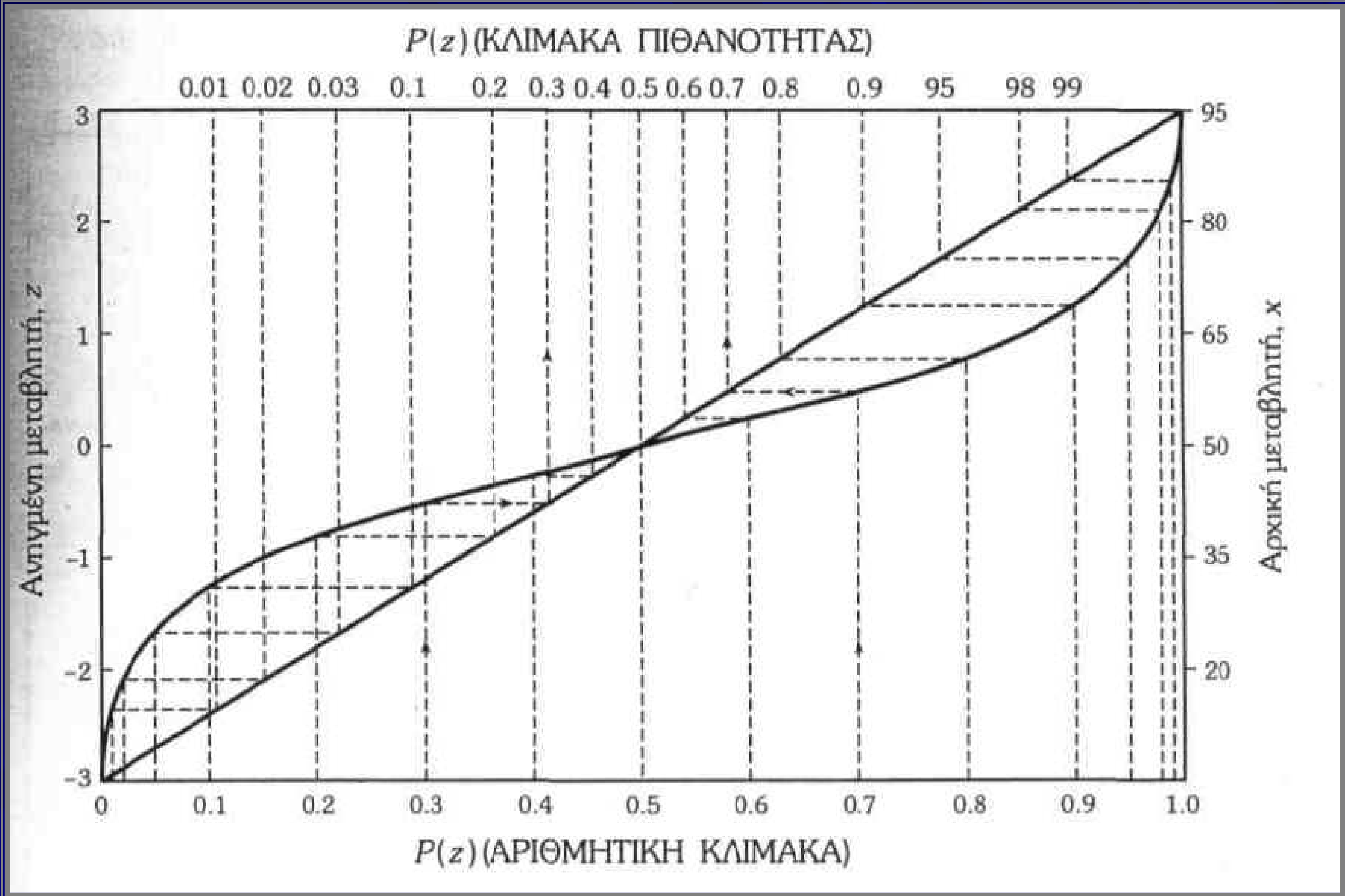
(ii) Η μέθοδος του παράγοντα συχνότητας.

Ακολουθεί μια συνοπτική παρουσίαση της γραφικής μεθόδου, της μεθόδου των Ροπών και του Παράγοντα συχνότητας.

Γραφική μέθοδος - Χαρτί πιθανότητας

Με το χαρτί πιθανότητας η καμπύλη αθροιστικής πιθανότητας γίνεται ευθεία γραμμή. Έτσι με την τοποθέτηση των σημείων που παριστάνουν την "αθροιστική πιθανότητα" στο χαρτί πιθανότητας μιας θεωρητικής κατανομής γίνεται ταυτόχρονα έλεγχος της καταλληλότητας της κατανομής και επέκταση του δείγματος με βάση τη θεωρητική κατανομή.

Στο Σχ. 2.2 φαίνεται η γραφική κατασκευή του χαρτιού πιθανότητας της κανονικής κατανομής. Η ανάλυση συχνότητας με το χαρτί πιθανότητας γίνεται ακολουθώντας την εξής διαδικασία (παράδειγμα μεγίστων τιμών):



Σχ. 2.2: *Γραφική κατασκευή του χαρτιού πιθανότητας της κανονικής κατανομής.*

i. Κατάταξη των δεδομένων σε φθίνουσα σειρά μεγέθους.

ii. Υπολογισμός της πιθανότητας υπέρβασης κάθε τιμής με τον τύπο

$$P(X \geq x) = \frac{m}{N+1}$$

iii. Επιλογή του τύπου του χαρτιού πιθανότητας που θα χρησιμοποιηθεί (π.χ. Κανονικής Κατανομής, Κατανομής Ακραίων Τιμών Τύπου Ι, κλπ).

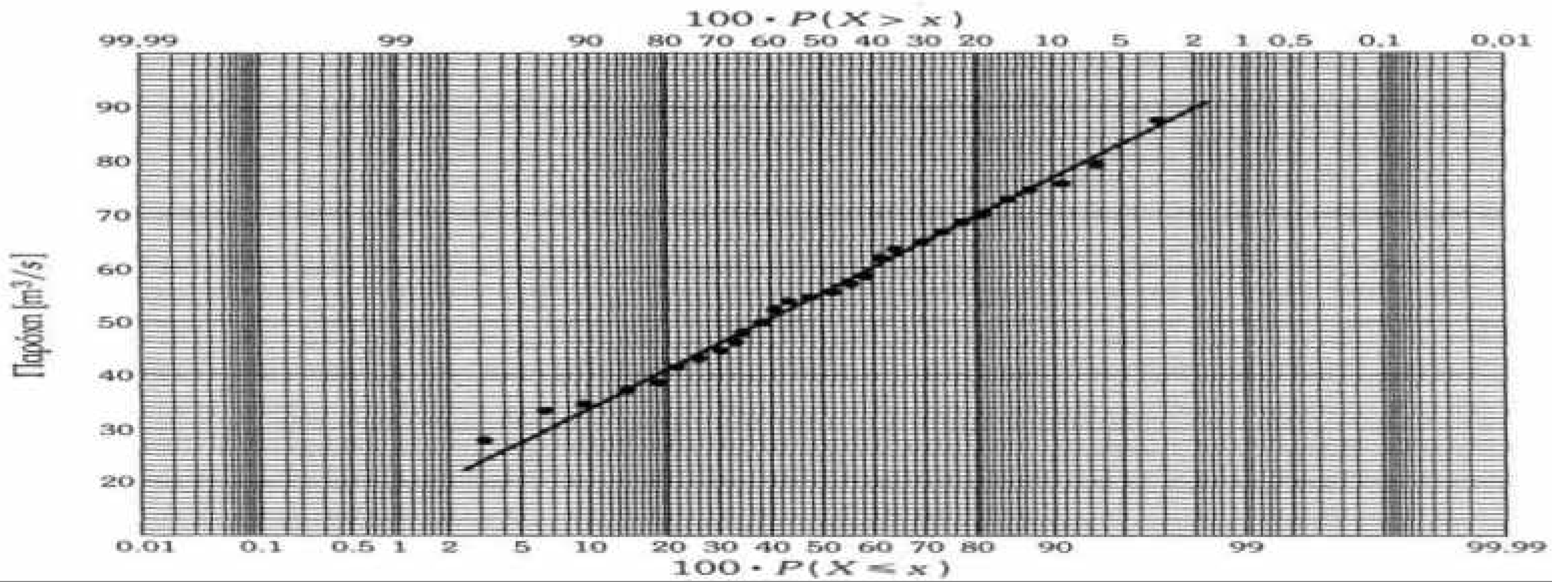
iv. Εμπειρική χάραξη της ευθείας που προσαρμόζεται καλύτερα στα δεδομένα.

Στο παράρτημα αυτού του βιβλίου δίνονται τα χαρτιά πιθανότητας για τις πιο συνήθεις Συναρτήσεις Κατανομών που χρησιμοποιούνται στην ανάλυση της υδρολογικής πληροφορίας.

Εφαρμογή. Στον Πίν. 2.1 βρίσκονται διατεταγμένες σε φθίνουσα σειρά μεγέθους οι μέσες ετήσιες παροχές ενός ρεύματος. Στην τρίτη στήλη του πίνακα βρίσκονται οι τιμές της πιθανότητας υπέρβασης. Οι τιμές αυτές με τα αντίστοιχα μεγέθη τοποθετούνται σε χαρτί πιθανότητας κανονικής κατανομής (Σχ. 2.3). Με τη χάραξη της ευθείας για κάθε μέγεθος παροχής αντιστοιχεί μια πιθανότητα υπέρβασης (και μια περίοδος επαναφοράς) και αντίστροφα.

Πίν. 2.1: *Υπολογισμός της πιθανότητας υπέρβασης των ημών μιας σειράς μέσων ετησίων παροχών ενός ρεύματος.*

Σειρά m	Παροχή [m³/s]	Πιθανότητα υπέρβασης (%) $100\frac{m}{N+1}$	Σειρά m	Παροχή [m³/s]	Πιθανότητα υπέρβασης (%) $100\frac{m}{N+1}$
1	88	3.1	16	55	51.6
2	79	6.3	17	55	51.6
3	75	9.4	18	54	56.3
4	74	12.5	19	53	59.3
5	72	15.6	20	50	62.5
6	70	18.8	21	48	65.6
7	68	21.9	22	46	68.8
8	67	25.0	23	44	71.9
9	64	29.7	24	43	75.0
10	64	29.7	25	42	78.1
11	63	34.4	26	38	81.3
12	62	37.5	27	37	85.9
13	58	40.6	28	37	85.9
14	57	43.8	29	35	90.6
15	56	46.9	30	34	93.8
			31	28	96.9



Όπως γίνεται φανερό από το Σχ. 2.3 στα δεδομένα του Πίν. 2.1 προσαρμόζεται ικανοποιητικά η κανονική κατανομή (τα δεδομένα ακολουθούν μια ευθεία γραμμή σε χαρτί πιθανότητας της Κανονικής Κατανομής).

Μέθοδος Ροπών Η μέθοδος των (στατιστικών) Ροπών στηρίζεται στο γεγονός ότι αν όλες οι ροπές της κατανομής είναι γνωστές όλα τα χαρακτηριστικά της κατανομής είναι γνωστά*.

Με τη μέθοδο των Ροπών εκτιμούνται από το δείγμα (αμερόληπτα) οι βασικές στατιστικές ποσότητες όπως ο μέσος όρος, μ, η τυπική απόκλιση, σ, και ο συντελεστής ασυμμετρίας,*g*,*ως εξής:*

$$\hat{«} = \sum_{§=1}^{N} \frac{i_§}{N} \qquad (2.13)$$

$$\hat{±} = (\hat{«}_2)^{\frac{1}{2}} = \left[\sum_{§=1}^{N} \frac{(i_§ - \bar{i})^2}{N-1} \right]^{1/2} \qquad (2.14)$$

$$\hat{g} = \frac{N}{(N-1)(N-2)}^{1/2} \sum_{§=1}^{N} \frac{(i_§ - \bar{i})^3}{(\hat{«}_2)^{3/2}} \qquad (2.15)$$

- Για τις διακριτές μεταβλητές η ροπή r τάξης ως προς χ = 0 ορίζεται:

$$v_r = \sum_1 f_r x_1^r \qquad \text{Για r=1} \qquad v_1 = \sum_1 f_1 x_1$$

Όπου κλάσεις του δείγματος και f_1 η σχετική συχνότητα της Ι κλάσης .Η ροπή r ως προς μια τιμή Α ορίζεται:

$$\text{«}_r = \sum_{i=1}^{k} f_r (x_i - A\text{¥})$$

Για Α=v_1 και r=2 =>μ_2(=διασπορά=$(\sigma)^2$)
Για Α=v_1 και r=3 =>μ_3

Ο συντελεστής ασυμμετρίας $g = \mu_3/\sigma^3$ (για συμμετρική κατανομή *g = 0)*. Αν οι παρατηρήσεις δεν έχουν ενταχθεί σε κλάσεις η σχετική συχνότητα αντικαθίσταται από το *1/N* όπου *N* ο αριθμός των παρατηρήσεων. Αμέσως μετά οι εκτιμήσεις εξισώνονται με τις στατιστικές ποσότητες της αντίστοιχης θεωρητικής Κατανομής πιθανότητας και με τη λύση του προκύπτοντος συστήματος γίνεται η εκτίμηση των παραμέτρων της Κατανομής πιθανότητας.

Μέθοδος του Παράγοντα Συχνότητας

Σύμφωνα με τον Chow (1951) για πολλές θεωρητικές κατανομές πιθανότητας μπορεί να γραφεί η ακόλουθη σχέση

$$x_T = \bar{x}(1 + c_v k_T) \qquad (2.16)$$

όπου:

χT :το μέγεθος του γεγονότος περιόδου επαναφοράς *T*,

χ :ο μέσος όρος

cv :ο συντελεστής διακύμανσης (ή μεταβλητότητας) (= σ/χ),

kT :ο παράγοντας συχνότητας που εξαρτάται από την περίοδο επαναφοράς *T* και τα χαρακτηριστικά της κατανομής.

Για τις συνήθεις κατανομές πιθανότητας έχουν προκύψει σχέσεις για τον υπολογισμό του παράγοντα συχνότητας, που παρουσιάζονται κατά την ανάπτυξη των επιμέρους κατανομών πιθανότητας.

2.6 ΘΕΩΡΗΤΙΚΕΣ ΚΑΤΑΝΟΜΕΣ ΠΙΘΑΝΟΤΗΤΑΣ

- Με την προϋπόθεση ότι οι τιμές του υδρολογικού φαινομένου είναι τυχαίες και ανεξάρτητες, η κατανομή πιθανότητας εμφανίσεως των μπορεί να περιγραφεί από τις θεωρητικές κατανομές πιθανότητας μερικές από τις οποίες περιγράφονται συνοπτικά στις επόμενες παραγράφους.

- Η ευκολία χρησιμοποιήσεως των θεωρητικών κατανομών ενέχει τον κίνδυνο να παρασύρει τους μελετητές σε ανάλυση παρατηρήσεων μη ανεξάρτητων μεταξύ τους. Για παράδειγμα οι ημερήσιες παροχές ενός ρεύματος δεν αποτελούν ανεξάρτητες παρατηρήσεις επειδή η παροχή κατά μία ημέρα έχει ισχυρή συσχέτιση με την παροχή της προηγούμενης ημέρας. Επίσης σχετικά με τις βροχοπτώσεις οι τιμές που αποτελούν το δείγμα θα πρέπει να προέρχονται από διαφορετικές ατμοσφαιρικές διαταράξεις. Βέβαια υπάρχουν περιπτώσεις που ο μελετητής με θεμελιωμένες παραδοχές μπορεί να παρεκλίνει από την αυστηρή απαίτηση των ανεξαρτήτων παρατηρήσεων με την ταυτόχρονη βέβαια γνώση ότι υπάρχει κίνδυνος εσφαλμένων αποτελεσμάτων.

Τις περισσότερες φορές κατά το σχεδιασμό των υδραυλικών έργων υπάρχει ιδιαίτερο ενδιαφέρον όχι για τις αρχικές παρατηρήσεις των υδρολογικών φαινομένων αλλά για τις ακραίες τιμές του υδρολογικού φαινομένου (μέγιστες ή ελάχιστες).

Είναι για παράδειγμα προφανές ότι η διαστασιολόγηση της διατομής μιας αποστραγγιστικής τάφρου ή ενός υπερχειλιστή ενός φράγματος θα πρέπει να γίνουν με τις μέγιστες παροχές επειδή το έργο πρέπει να ανταποκρίνεται στις δυσμενέστερες συνθήκες.

Αποδείχνεται ότι η ακραία τιμή από ένα δείγμα τυχαίων τιμών είναι επίσης τυχαία. Επίσης αποδείχνεται ότι η κατανομή πυκνότητας πιθανότητας της ακραίας τιμής εξαρτάται από την αρχική κατανομή πυκνότητας πιθανότητας και το μέγεθος του αρχικού δείγματος. Για μεγάλο αριθμό αρχικών παρατηρήσεων η κατανομή πυκνότητας πιθανότητας για τις ακραίες τιμές περιγράφεται από ασυμπτωτικές κατανομές (Gumbel 1954, 1958).

Τρεις τύποι ασυμπτωτικών κατανομών έχουν προταθεί:

- Τύπος I : Αρχική κατανομή χωρίς όριο προς την κατεύθυνση της ακραίας τιμής.
- Τύπος II : Αρχική κατανομή χωρίς όριο και προς τις δύο κατευθύνσεις.
- Τύπος III:Αρχική κατανομή με όριο προς την κατεύθυνση της ακραίας τιμής.

- Η κατανομή τύπου Ι (ή Gumbel) ενδείκνυται να εφαρμόζεται συνήθως για ανάλυση μεγίστων τιμών αν η αρχική κατανομή είναι η Κανονική, Λογαριθμοκανονική, Εκθετική ή Γάμμα και για ελάχιστα αν η αρχική κατανομή είναι κανονική.
- Η κατανομή Τύπου ΙΙ (ή Cauchy) δεν έχει σημαντική εφαρμογή στην Υδρολογία.
- Τέλος η κατανομή Τύπου ΙΙΙ ενδείκνυται να εφαρμόζεται για ανάλυση μεγίστων τιμών αν η αρχική κατανομή είναι Βήτα και για ελάχιστα αν η αρχική κατανομή είναι Βήτα, Λογαριθμοκανονική, Γάμμα ή Εκθετική.
- Όπως γίνεται φανερό προκειμένου για ανάλυση μεγίστων τιμών ενός υδρολογικού φαινομένου περισσότερο κατάλληλη φαίνεται η Κατανομή Τύπου Ι. Αυτό συμπεραίνεται από το γεγονός ότι το όριο των υδρολογικών μεταβλητών είναι συνήθως το μηδέν και επομένως δεν υπάρχει όριο προς την κατεύθυνση των μεγίστων (ακραίων) τιμών. Με όμοιο συλλογισμό συμπεραίνεται ότι η Κατανομή Τύπου ΙΙΙ είναι καταλληλότερη για ανάλυση ελάχιστων τιμών επειδή υπάρχει το όριο (δηλ. 0) προς την κατεύθυνση των (ελαχίστων) ακραίων τιμών. Μια τέτοια θεωρητική Κατανομή πιθανότητας, γνωστή ως Weibull, παρουσιάζεται στο Κεφάλαιο της Ξηρασίας σαυτό το βιβλίο.

2.6.1 Κανονική Κατανομή (Normal or Gaussian)

Συνάρτηση πυκνότητας πιθανότητας (p.d.f.)

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right] \qquad (2.17)$$

$-\infty < x < +\infty \quad -\infty < \mu < +\infty \quad$ ©~§ $\quad \sigma > 0$

Με αντικατάσταση

$$Z = \frac{x-\mu}{\sigma} \qquad (2.18)$$

Προκύπτει η μορφή:

$$p(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2} \right) \qquad -\infty < z < +\infty \qquad (2.19)$$

- Η συνάρτηση αθροιστικής πιθανότητας (c.d.f.)

$$P(z) = \int_{-\infty}^{z} \frac{1}{2\text{®}} \exp\left(-\frac{z^2}{2}\right) dz \qquad (2.20)$$

- Οι τιμές *P (ζ)* συναρτήσει της μεταβλητής *ζ* υπάρχουν σε πίνακες της τυπικής κανονικής κατανομής. Για ευκολία στον Πίν. 2.2 η τιμή της μεταβλητής *ζ* προκύπτει από την αθροιστική πιθανότητα σε ποσοστό (%).

- Επειδή οι παράμετροι της κανονικής κατανομής είναι ο μέσος όρος και η τυπική απόκλιση *(μ* και *σ)* σύμφωνα με τη μέθοδο των ροπών η εκτίμηση τους γίνεται από τα δεδομένα με βάση τις Εξ. 2.13 και 2.14.

Παράδειγμα

Μια σειρά μέσων ετησίων παροχών έχει μέσο όρο
χ = 60.5 m3/s και τυπική απόκλιση *σ* = 20.5 m3/s. Ζητείται να βρεθεί η ετήσια παροχή με περίοδο επαναφοράς 10 έτη με την παραδοχή ότι ισχύει η κανονική κατανομή.

Η αθροιστική πιθανότητα *P(ζ)* υπολογίζεται

$$P(z) = P(X \leq x) = 1 - \frac{1}{T} = 0.90$$

Η ανηγμένη μεταβλητή για *P(ζ)* = 0.90 προκύπτει από τον Πίν. 2.2 *ζ* = 1.282. Τέλος από την Εξ. 2.18 με την εκτίμηση του θεωρητικού μέσου όρου *μ* και της τυπικής απόκλισης *σ* από το μέσο όρο χ και την τυπική απόκλιση σ του δείγματος (μέγεθος των ροπών) υπολογίζεται η παροχή περιόδου επαναφοράς 10 ετών

Πίν. 2.2: *Τιμές του z για δοσμένες τιμές της αθροιστικής πιθανότητας P(z) της τυπικής Κανονικής κατανομής.*

P(z) %	z	P(z) %	z	P(z) %	z	P(z) %	z	P(z) %	z	P(z) %	z
1	−2.326	21	−0.806	41	−0.228	61	0.279	81	0.878	99.1	2.366
2	−2.054	22	−0.772	42	−0.202	62	0.305	82	0.915	99.2	2.409
3	−1.881	23	−0.739	43	−0.176	63	0.332	83	0.954	99.3	2.457
4	−1.751	24	−0.706	44	−0.151	64	0.358	84	0.994	99.4	2.512
5	−1.645	25	−0.674	45	0.126	65	0.385	85	1.036	99.5	2.576
6	−1.555	26	−0.643	46	−0.100	66	0.412	86	1.080	99.6	2.652
7	−1.476	27	−0.613	47	−0.075	67	0.440	87	1.126	99.7	2.748
8	1.405	28	−0.583	48	−0.050	68	0.467	88	1.175	99.8	2.878
9	−1.341	29	−0.553	49	−0.025	69	0.496	89	1.227	99.9	3.090
10	−1.382	30	−0.524	50	0.000	70	0.524	90	1.282		
11	−1.227	31	−0.496	51	0.025	71	0.553	91	1.341	99.91	3.121
12	−1.175	32	−0.468	52	0.050	72	0.583	92	1.405	99.92	3.156
13	−1.126	33	−0.440	53	0.075	73	0.613	93	1.476	99.93	3.195
14	−1.080	34	−0.412	54	0.100	74	0.643	94	1.555	99.94	3.239
15	−1.036	35	−0.385	55	0.126	75	0.674	95	1.645	99.95	3.291
16	−0.994	36	−0.358	56	0.151	76	0.706	96	1.751	99.96	3.353
17	−0.954	37	−0.332	57	0.176	77	0.739	97	1.881	99.97	3.432
18	−0.915	38	−0.305	58	0.202	78	0.772	98	2.054	99.98	3.540
19	−0.878	39	−0.279	59	0.228	79	0.806	99	2.326	99.99	3.719
20	−0.842	40	−0.253								

$$x_{10} = \mu + \sigma z = 60.5 + 20.3 \cdot 1.282 = 86.56 \ m^3/s$$

2.6.2 Κατανομή Ακραίων Τιμών Τύπου Ι (Gumbel)

Συνάρτηση πυκνότητας πιθανότητας (p.d.f.):

$$p(x) = \exp\left\{ \pm \frac{x -}{\sim} - \exp\left[\mp \frac{x -}{\sim} \right] \right\} \qquad (2.21)$$

$$-\infty < ¡ < +\infty, \quad -\infty < \quad < +\infty, \quad \sim > 0$$

όπου α είναι η παράμετρος κλίμακας και *β* είναι παράμετρος θέσεως.

Τα σύμβολα + και - είναι: (-) για μέγιστα και (+) για ελάχιστα.

Βασικές στατιστικές ποσότητες:

Προσδοκώμενη τιμή E (X) *β* + 0.577 *α* (μέγιστα)
(μέσος όρος) E(X) = *β*- 0.577 α (ελάχιστα)

Διασπορά Var(X) =1.645 $α^2$ (μέγιστα και ελάχιστα) (2.22)

Συντελεστής ασυμμετρίας *g*= 1.1396 (μέγιστα)
g = -1.1396 (ελάχιστα)

Με την αντικατάσταση:

$$y = \frac{x -}{\sim} \qquad (2.23)$$

$$p.d.f \qquad p(y) = \exp\left[\mp y - \exp(\mp y)\right] \qquad (2.24)$$

$$c.d.f \qquad P(y) = \int_{-\infty}^{y} \exp\left[(\mp y - \exp(\mp y))\right] dy \qquad (2.25)$$

Τελικά

$P(y) = 1 - \exp(-\exp y)$(£ª~¡§±²~) (2.26)

$P(y) = \exp(-\exp(-y))$(«£¥§±²~) (2.27)

Με βάση τις τελευταίες εξισώσεις έχει γίνει η πινακοποίηση της αθροιστικής πιθανότητας (U.S. National Bureau of Standards, 1953).
Με τη μέθοδο των ροπών η εκτίμηση των παραμέτρων γίνεται ως εξής:

$$\hat{\sim} = \frac{\hat{\pm}}{1.283} \quad (2.28)$$

©~§ ˆ = ¡̄ − 0.45±̂ («£¥§±²~) (2.29)

ˆ = ¡̄ + 0.45±̂ (£ª~¡§±²~) (2.30)

Παράδειγμα. Σειρά ετησίων μεγίστων υψών βροχής μιας ορισμένης διάρκειας έχει μέσο όρο χ = 14.19 mm και τυπική απόκλιση σ = 8.46 mm.
Χρησιμοποιώντας την κατανομή Gumbel να προσδιορισθεί το ύψος που αντιστοιχεί σε περίοδο επαναφοράς Τ= 10 έτη. Σύμφωνα με τη μέθοδο των ροπών α = σ/1.283 = 6.594, *β* = χ - 0.45 *σ* = = 10.38. Επίσης *P(y)* = 1 - (1/T) = 0.90 και επομένως exp[-exp(-y)] = 0.90 ή y = -ln[-1n 0.90] = 2.25.
Τέλος χ = *ay* +*β*=6.594 · 2.25 + 10.38 = 25.22 mm.

Με τη μέθοδο του παράγοντα συχνότητας για πολύ μεγάλο δείγμα $(N \to \infty)$ έχει προκύψει αναλυτικά η σχέση μεταξύ παράγοντα συχνότητας και περιόδου επαναφοράς

$$K_T = -0.7797\left\{0.5772 + \ln\left[\ln\left(T/(T-1)\right)\right]\right\} \qquad (2.31)$$

Για μικρά σχετικά δείγματα ($N < 100$) ο παράγοντας συχνότητας υπολογίζεται ως Εξής

$$K_T = \frac{-\left\{\ln\left[\ln\left(T/(T-1)\right)\right] + \overline{y_N}\right\}}{\pm_{\mathrm{N}}} \qquad (2.32)$$

όπου και σ_N (μέσος όρος και τυπική απόκλιση της ανηγμένης μεταβλητής y) βρίσκονται στον Πίν. 2.3 με βάση τον αριθμό των παρατηρήσεων του δείγματος. Αν στο προηγούμενο παράδειγμα εφαρμόσουμε τη μεθοδολογία Του Παράγοντα Συχνότητας (Εξ. 2.31 και 2.32) τότε προκύπτουν τα εξής αποτελέσματα:

Από την Εξ. 2.31

$$K_T = -0.7797\left[0.5772 + \ln\left(\ln\frac{10}{9}\right)\right] = 1.3046$$

Επίσης:

$$c_v = \frac{\hat{\sigma}}{\bar{x}} = \frac{8.46}{14.19} = 0.5962$$

Συνεπώς σύμφωνα με την Εξ.2.16

$$x_{10} = 14.19(1 + 0.5962 \bullet 1.2046) = 25.23mm$$

Αν θεωρηθεί δείγμα με αριθμό παρατηρήσεων 20 τότε από την Εξ. 2.32

$$K_T = -\frac{\left[\ln\left(\ln\frac{10}{9}\right) + 0.52355\right]}{1.06283} = 1.625$$

Και από την Εξ 2.16

$$x_{10} = 14.19(1 + 0.5962 \bullet 1.625) = 27.94$$

Δηλαδή αν το δείγμα είναι σχετικά μικρό υπάρχει σημαντική απόκλιση από την εκτίμηση που γίνεται με την υπόθεση ότι ο αριθμός παρατηρήσεων $N \to \infty$

Πίν. 2.3: *Τιμές των παραμέτρων* $\bar{y}_N$ *και* $\pm_N$ *για διάφορες τιμές του αριθμού των παρατηρήσεων N*

N	$\bar{y}_N$	σ_N	N	$\bar{y}_N$	σ_N	N	$\bar{y}_N$	σ_N
8	0.48430	0.90430	35	0.54034	1.12847	64	0.55330	1.17930
9	0.49020	0.92880	36	0.54100	1.13130	66	0.55380	1.18140
10	0.49520	0.94970	37	0.54180	1.13390	68	0.55430	1.18340
11	0.49960	0.96760	38	0.54240	1.13630	70	0.55477	1.18536
12	0.50350	0.98330	39	0.54300	1.13880	72	0.55520	1.18730
13	0.50700	0.99720	40	0.54362	1.14132	74	0.55570	1.18900
14	0.51000	1.00950	41	0.54420	1.14360	76	0.55610	1.19060
15	0.51280	1.02057	42	0.54480	1.14580	78	0.55650	1.19230
16	0.51570	1.03160	43	0.54530	1.14800	80	0.55688	1.19382
17	0.51810	1.04110	44	0.54580	1.14990	82	0.55720	1.19530
18	0.52020	1.04930	45	0.54630	1.15185	84	0.55760	1.19670
19	0.52200	1.05660	46	0.54680	1.15380	86	0.55800	1.19800
20	0.52355	1.06283	47	0.54730	1.15570	88	0.55830	1.19940
21	0.52520	1.06960	48	0.54770	1.15740	90	0.55860	1.20073
22	0.52680	1.07540	49	0.54810	1.15900	92	0.55890	1.20200
23	0.52830	1.08110	50	0.54854	1.16066	94	0.55920	1.20320
24	0.52960	1.08640	51	0.54890	1.16230	96	0.55950	1.20440
25	0.53086	1.09145	52	0.54930	1.16380	98	0.55980	1.20550
26	0.53200	1.09610	53	0.54970	1.16530	100	0.56002	1.20649
27	0.53320	1.10040	54	0.55010	1.16670	150	0.56461	1.22534
28	0.53430	1.10470	55	0.55040	1.16810	200	0.56715	1.23598
29	0.53530	1.10860	56	0.55080	1.16960	250	0.56878	1.23292
30	0.53622	1.11238	57	0.55110	1.17080	300	0.56993	1.24786
31	0.53710	1.11590	58	0.55150	1.17210	400	0.57144	1.25450
32	0.53800	1.11930	59	0.55180	1.17340	500	0.57240	1.25880
33	0.53880	1.12260	60	0.55208	1.17467	750	0.57377	1.26506
34	0.53960	1.11550	62	0.55270	1.17700	1000	0.57450	1.26851

Παρατήρηση

Η Κατανομή Ακραίων Τιμών Τύπου Ι χρησιμοποιείται . για την στατιστική επεξεργασία των μεγίστων τιμών ενός υδρολογικού μεγέθους όπως οι ραγδαίες βροχές, οι πλημμυρικές απορροές κλπ.

Όπως φαίνεται από την παραπάνω ανάπτυξη για θετικά πρόσημα η κατανομή μπορεί να χρησιμοποιηθεί και για ελάχιστες τιμές των υδρολογικών μεγεθών (π.χ. χαμηλές ροές - συνθήκες ξηρασίας). Στην τελευταία αυτή περίπτωση επίσης κατάλληλη είναι η Κατανομή Weibull που παρουσιάζεται στο Κεφάλαιο της Ξηρασίας.

2.6.3 Κατανομές Pearson III

Η κατανομή Pearson III έχει τρεις παραμέτρους που συμβολίζονται με χ0, *β* και *γ*. Στην ειδική περίπτωση που χ_0 = 0 προκύπτει η κατανομή Γάμμα. Επίσης στην περίπτωση *γ*= 1 προκύπτει η εκθετική κατανομή.

i. Εκθετική Κατανομή

Συνάρτηση πυκνότητας πιθανότητας (p.d.f.) είναι

$$p(x) = \frac{1}{} \exp\left[- \frac{x - x_0}{} \right] \qquad (2.33)$$

και συνάρτηση αθροιστικής πιθανότητας (c.d.f.) είναι

$$p(x) = 1 - \exp\left[- \frac{x - x_0}{} \right] \qquad (2.34)$$

όπου:

χ_0 : είναι παράμετρος θέσης και

β : είναι παράμετρος κλίμακας

Στα Σχήματα 2.4α και 2.4β φαίνονται οι γραφικές παραστάσεις των Εξ. 2.33 και 2.34 αντίστοιχα για χ_0 = 10. Βασικές στατιστικές ποσότητες:

$$E(X) = x_0 + \quad \text{(μέσος όρος)} \quad (2.35)$$

$$Var(X) = \quad ^2 \quad \text{(Διασπορά)}$$

$$g = 2 \quad \text{(συντελεστής ασυμμετρίας)}$$

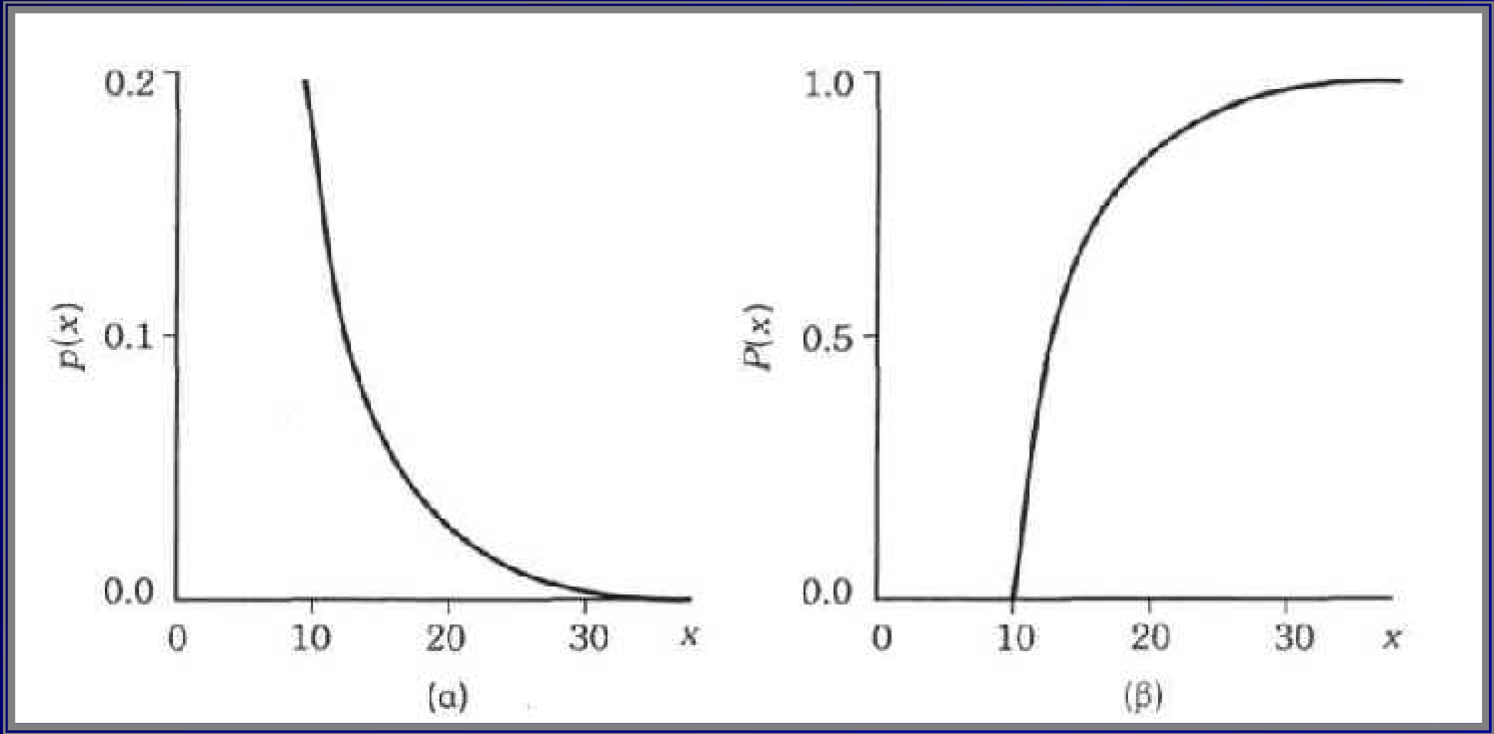
Με αντικατάσταση $y = (x - x_0)/$ προκύπτει:

$$p(y) = \exp(-y) \quad (p.d.f.) \quad (2.36)$$

Και

$$p(y) = 1 - \exp(-y) \quad (c.d.f.) \quad (2.37)$$

Η τελευταία συνάρτηση μπορεί να πινακοποιηθεί κατά τρεις τρόπους όπως φαίνεται στον Πίν. 2.4.



Σχ. 2.4: *Γραφική παράσταση εκθετικής κατανομής*

α. συνάρτηση πυκνότητας πιθανότητας για $\chi_0 = 10$

β. συνάρτηση αθροιστικής πιθανότητας για $\chi_0 = 10$.

(α)		(β)		(γ)	
$P(y)$	y	y	$P(y)$	T	y
0.00	0.000	0.00	0.000	2	0.69
0.10	0.105	0.25	0.221	5	1.61
0.20	0.223	0.50	0.393	6	1.79
0.30	0.357	0.75	0.528	7	1.95
0.40	0.511	1.00	0.632	8	2.08
0.50	0.693	1.25	0.713	9	2.20
0.60	0.916	1.50	0.777	10	2.30
0.70	1.204	1.75	0.826	20	3.00
0.80	1.609	2.00	0.865	25	3.22
0.90	2.303	2.25	0.895	30	3.40
0.95	2.996	2.50	0.918	40	3.69
0.99	4.605	2.75	0.936	50	3.91
		3.00	0.950	60	4.09
				70	4.25
				80	4.38
				90	4.50
				100	4.61
				1000	6.91

Πίν. 2.4: *Πινακοποίηση κατά τρεις τρόπους της Εξ. 2.37 της Εκθετικής κατανομής.*

ii. Κατανομή Γάμμα

Συνάρτηση πυκνότητας πιθανότητας (p.d.f.) είναι

$$p(x) = \frac{x^{\gamma-1}\exp(-x/\beta)}{\beta^{\gamma}\Gamma(\gamma)}dx \qquad \gamma > 0 \qquad (2.38)$$

και η συνάρτηση αθροιστικής πιθανότητας (c.d.f.) είναι

$$p(x) = \int_0^{\infty} \frac{x^{\gamma-1}\exp(-x/\beta)}{\beta^{\gamma}\Gamma(\gamma)}dx \qquad (2.39)$$

όπου:

Γ (γ): είναι η πλήρης συνάρτηση Γάμμα

β : είναι η παράμετρος κλίμακας και

γ : είναι η παράμετρος σχήματος

Βασικές στατιστικές ποσότητες:

$$E(X) = \quad ¥$$ (Μέσος όρος)

$$Var(X) = \quad ^{2}¥$$ (διασπορά) (2.40)

$$g = 2 / \sqrt{¥}$$ (συντελεστής ασυμμετρίας)

Με την αντικατάσταση y = *χ/β* προκύπτει

$$p(y) = \frac{y^{¥-1}\exp(-y)}{\Gamma(¥)} \qquad (p.d.f.) \qquad (2.41)$$

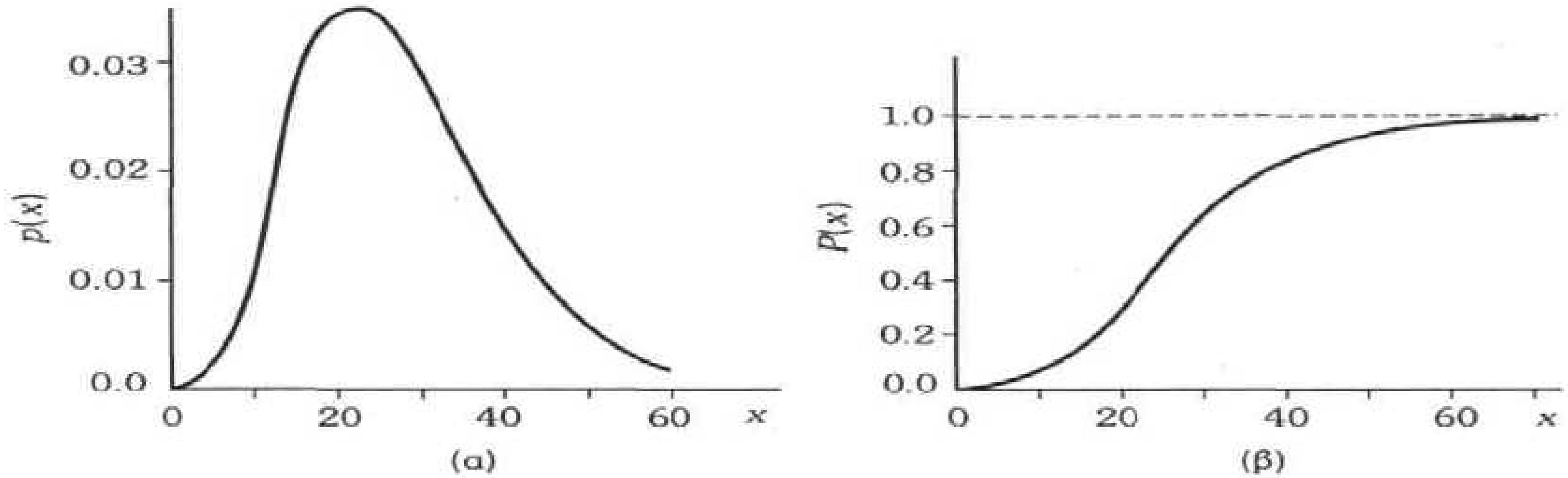
Και

$$P(y) = \int_0^y \frac{y^{¥-1}\exp(-y)}{\Gamma(¥)} \qquad (c.d.f.) \qquad (2.42)$$

Η τελευταία συνάρτηση μπορεί να πινακοποιηθεί για κάθε τιμή της παραμέτρου *γ* (Πίν.2.5).

Πίν. 2.5: *Τιμές της ανηγμένης μεταβλητής y της κατανομής Γάμμα*

$P(y)$	$\gamma = 1$	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$	$\gamma = 20$
0.10	0.105	0.532	2.433	6.221	14.53
0.20	0.223	0.824	3.090	7.289	16.17
0.30	0.357	1.097	3.634	8.133	17.44
0.40	0.511	1.376	4.148	8.904	18.57
0.50	0.693	1.678	4.671	9.669	19.67
0.60	0.916	2.022	5.237	10.476	20.81
0.70	1.204	2.439	5.890	11.387	22.08
0.80	1.609	2.994	6.721	12.519	23.63
0.90	2.303	3.890	7.994	14.206	25.90
0.95	2.996	4.744	9.154	15.705	27.88
0.99	4.605	6.638	11.605	18.783	31.85



Σχ. 2.5: *Γραφική παράσταση της κατανομής Γάμμα*
α. συνάρτηση πυκνότητας πιθανότητας
β. συνάρτηση αθροιστικής πιθανότητας.

Παράδειγμα

Σε μια σειρά μεγίστων υψών βροχής (συγκεκριμένης διάρκειας) προσαρμόζεται η κατανομή Γάμμα. Το δείγμα έχει μέσο όρο 40 mm και τυπική απόκλιση 9 mm. Ζητείται η τιμή με περίοδο επαναφοράς T = 10 έτη.

Από τα δεδομένα και τις Εξ 2.40 βγ=40 και $\beta^2\gamma=9^2$

Δηλαδή β=2.025 και γ=19.753.

Επίσης 1-P (y)=1/Τα=1/10 ή P (y)=0.90.

Από τον Πιν. 2.5 για γ=20 και P (y)=0.90 προκύπτει y=25.90 και

χ= βy=2.025·25.90=52.45 mm

Παρατήρηση. Η p.d.f. της κατανομής Γάμμα για *γ= 1* είναι η p.d.f. της εκθετικής με κάτω όριο χ_0 = 0. Επίσης για πολύ μεγάλη τιμή του *γ*

η ασυμμετρία γίνεται πολύ μικρή και η Γάμμα κατανομή προσεγγίζει την κανονική

iii. Κατανομή Pearson III

Η συνάρτηση πυκνότητας πιθανότητας και η συνάρτηση αθροιστικής πιθανότητας είναι αντίστοιχα

$$p(x) = \frac{(x - x_0)^{\yen - 1} \exp\left[-(x - x_0) / \quad \right]}{^{\yen}\Gamma(\yen)} \qquad (p.d.f) \qquad (2.43)$$

$$P(x) = \int_{x_0}^{x} \frac{(x - x_0)^{\yen - 1} \exp\left[-(x - x_0) / \quad \right]}{^{\yen}\Gamma(\yen)} dx \qquad (c.d.f) \qquad (2.44)$$

Με την αντικατάσταση $z = \chi - \chi_0$ προκύπτει η κατανομή Γάμμα με παραμέτρους *β* και *γ*.Δηλαδή η κατανομή Pearson III είναι η αντίστοιχη κατανομή Γάμμα αλλά με αρχικό όριο $\chi = \chi_0 \neq 0$.Βασικές στατιστικές ποσότητες:

$$E(X) = x_0 + \quad \yen \qquad \text{(μέσος όρος)}$$

$$Var(X) = \quad ^{2}\yen \qquad \text{(διασπορά)} \qquad (2.45)$$

$$g = 2 / \sqrt{\yen} \qquad \text{(συντελεστής ασυμμετρίας)}$$

Η τυπική μεταβλητή $y = (\chi - \chi_0)/\beta$ αυτής της κατανομής είναι η ίδια με της Γάμμα για την ίδια τιμή της παραμέτρου *γ*.

Παράδειγμα

Ένα δείγμα μεγίστων παροχών που κατανέμεται με την κατανομή Pearson III έχει μέσο όρο 50 m^3/s, τυπική απόκλιση 15 m^3/s και συντελεστή ασυμμετρίας 0.90.Ζητείται το μέγεθος με περίοδο επαναφοράς 10 έτη . Από τις σχέσεις $\chi_0+\beta\gamma=50$, $B^2\gamma=15^2$ και $2/\sqrt{\gamma}=0.9$ προκύπτουν γ=4.938, β=6.75, $\chi_0=16.67$.
Επίσης P(y)=1-1/T=0.90Από τον Πιν. 2.5 για γ=5(προσεγγιστικά) και P(y)=0.90 προκύπτει y=7.994 και $\chi=\chi_0+\beta y=16.67+6.75\cdot 7.994=70.63 m^3/s$

Η χρήση των κατανομών Pearson διευκολύνεται με την μέθοδο του παράγοντα συχνότητας .Ο παράγοντας συχνότητας δίνεται σε πίνακες Harter 1969) με βάση την αθροιστική πιθανότητα και το συντελεστή ασυμμετρίας όπως ο Πίν. 2.6. Για αρνητικές τιμές του συντελεστή ασυμμετρίας η τιμή του παράγοντα συχνότητας βρίσκεται ως εξής: Πρώτα επιλέγεται η τιμή που αντιστοιχεί στην αντίστοιχη θετική τιμή και στην αθροιστική πιθανότητα 1 - P(χ). Στην τιμή του παράγοντα συχνότητας που προκύπτει από τον Πίν. 2.6 αλλάζεται το πρόσημο.

Παράδειγμα

Σε μια σειρά ετησίων μέγιστων παροχών ενός ρεύματος έχουν υπολογιστεί

$\bar{x} = 6 \bullet 10 m^3 / s$ $\quad$ £ $= 1.45 m^3 / s$ $\quad$ $g = -0.40$.Ζητείται η παροχή με περίοδο επαναφοράς 20 έτη. Η αθροιστική πιθανότητα είναι P(x)=1-1/T=0.95.Επειδή ο συντελεστής ασυμμετρίας είναι αρνητικός από τον Πιν.2.6 για αθροιστική πιθανότητα 0.05 (αντί της 0.95) g=0.4 προκύπτει K_T=-1.52357 και με την αλλαγή προσήμου K_T=1.52357.Τελικά $x_T = \bar{x}(1 + c_v K_T) = 6.10(1 + 1.45 \bullet 1.52357) = 19.58 m^3 / s$

Ο Πίνακας 2.6 μπορεί να χρησιμοποιηθεί τόσο για την εκθετική κατανομή όσο και για την κατανομή Γάμμα. Για ευκολία εκτός του Πίνακα 2.6 παρατίθεται και ο λιγότερο λεπτομερής Πίν. 2.7 για τις αρνητικές τιμές του συντελεστή ασυμμετρίας. Λύνοντας τα παραδείγματα των κατανομών Pearson που παρουσιάσθηκαν λυμένα προηγούμενα με τη χρήση του παράγοντα συχνότητας προέκυψαν τα εξής αποτελέσματα:

Εκθετική : K_τ= 1.30259 και χ_τ= 63.03 mm (αντί 63 mm)

Γάμμα: K_Γ= 1.3199 και χ_τ= 51.88 mm (αντί 52.45 mm)

Pearson III : K_τ= 1.33889 και x_T=70m3/s (αντί 63 m^3/s)

Στις παραγράφους που προηγήθηκαν έχει γίνει σιωπηρώς η υπόθεση ότι η εκτίμηση των παραμέτρων των θεωρητικών κατανομών γίνεται με τη μέθοδο των ροπών και οι ποσότητες όπως προσδιορίζονται είναι στην ουσία εκτιμήσεις των παραμέτρων. Συνοπτικά για τις κατανομές Pearson οι εκτιμήσεις των παραμέτρων με τη μέθοδο των ροπών είναι:

Εκθετική: $\hat{\ } = \hat{£}$ ©~§ $\hat{x}_0 = \bar{x} - \hat{£}$

Γάμμα: $\hat{¥} = \bar{x}^2 / \hat{£}^2$ ©~§ $\hat{\ } = \hat{£}^2 / \bar{x}$

Pearson III : $\hat{¥} = 4 / \hat{g}^2$ ©~§ $\hat{\ } = \hat{£}\hat{g} / 2$ ©~§ $\hat{x}_0 = \bar{x} - 2\hat{£} / \hat{g}$

Πίν. 2.6: *Ο Παράγοντας Συχνότητας της Κατανομής Pearson III*

$P(x)$	$g = 0$	$g = 0.1$	$g = 0.2$	$g = 0.3$	$g = 0.4$	$g = 0.5$	$g = 0.6$	$g = 0.7$	$g = 0.8$	$g = 0.9$
0.0001	-3.71902	-3.50703	-3.29921	-3.09631	-2.89907	-2.70836	-2.52507	-2.35015	-2.18448	-2.02891
0.0005	-3.29053	-3.12767	-2.96698	-2.80889	-2.65390	-2.50257	-2.35549	-2.21328	-2.07661	-1.94611
0.0010	-3.09023	-2.94834	-2.80786	-2.66915	-2.53261	-2.39867	-2.26780	-2.14053	-2.01739	-1.89894
0.0020	-2.87816	-2.75706	-2.63672	-2.51741	-2.39942	-2.28311	-2.16884	-2.05701	-1.94806	-1.84244
0.0050	-2.57583	-2.48187	-2.38795	-2.29423	-2.20092	-2.10825	-2.01644	-1.92580	-1.83660	-1.74919
0.0100	-2.32635	-2.25258	-2.17840	-2.10394	-2.02933	-1.95472	-1.88029	-1.80621	-1.73271	-1.66001
0.0200	-2.05375	-1.99973	-1.94499	-1.88959	-1.83361	-1.77716	-1.72033	-1.66325	-1.60604	-1.54886
0.0400	-1.75069	-1.71580	-1.67999	-1.64329	-1.60574	-1.56740	-1.52830	-1.48852	-1.44813	-1.40720
0.0500	-1.64485	-1.61594	-1.58607	-1.55527	-1.52357	-1.49101	-1.45762	-1.42345	-1.38855	-1.35299
0.1000	-1.28155	-1.27037	-1.25824	-1.24516	-1.23114	-1.21618	-1.20028	-1.18347	-1.16574	-1.14712
0.2000	-0.84612	-0.84611	-0.84986	-0.85285	-0.85508	-0.85653	-0.85718	-0.85703	-0.85607	-0.85426
0.3000	-0.52440	-0.53624	-0.54757	-0.55839	-0.56867	-0.57840	-0.58757	-0.59615	-0.60412	-0.61146
0.4000	-0.25335	-0.26882	-0.28403	-0.29897	-0.31362	-0.32796	-0.34198	-0.35565	-0.36889	-0.38186
0.5000	0.00000	-0.01662	-0.03325	-0.04993	-0.06651	-0.08302	-0.09945	-0.11578	-0.13199	-0.14807
0.6000	0.25335	0.23763	0.22168	0.20552	0.18916	0.17261	0.15589	0.13901	0.12199	0.10486
0.7000	0.52440	0.51207	0.49927	0.48600	0.47228	0.45812	0.44352	0.42851	0.41309	0.39729
0.8000	0.84162	0.83639	0.83044	0.82377	0.81638	0.80829	0.79950	0.79002	0.77986	0.76902
0.9000	1.28155	1.29178	1.30105	1.30936	1.31671	1.32309	1.32850	1.33294	1.33640	1.33889
0.9500	1.64485	1.67279	1.69971	1.72562	1.75048	1.77428	1.79701	1.81864	1.83916	1.85856
0.9600	1.75069	1.78462	1.81756	1.84949	1.88039	1.91022	1.93896	1.96660	1.99311	2.01848
0.9750	1.95996	2.00688	2.05290	2.09795	2.14202	2.18505	2.22702	2.26790	2.30764	2.34623
0.9800	2.05375	2.10697	2.15935	2.21081	2.26133	2.31084	2.35931	2.40670	2.45298	2.49811
0.9900	2.32635	2.39961	2.47226	2.54421	2.61539	2.68572	2.75514	2.82359	2.89101	2.95735
0.9950	2.57583	2.66965	2.76321	2.85636	2.94900	3.04102	3.13232	3.22281	3.31243	3.40109
0.9980	2.87816	2.99978	3.12169	3.24371	3.36566	3.48737	3.60872	3.72957	3.84981	3.96932
0.9990	3.09023	3.23322	3.37703	3.52139	3.66608	3.81090	3.95567	4.10022	4.24439	4.38807
0.9995	3.29053	3.45513	3.62113	3.78820	3.95605	4.12443	4.29311	4.46189	4.63057	3.79899
0.9999	3.71902	3.93453	4.15301	4.37394	4.59687	4.82141	5.04718	5.27389	5.50124	5.72899

Πίν. 2.6: *Ο Παράγοντας Συχνότητας της Κατανομής Pearson III (Συνέχεια)*

$P(x)$	$g=1.0$	$g=1.1$	$g=1.2$	$g=1.3$	$g=1.4$	$g=1.5$	$g=1.6$	$g=1.7$	$g=1.8$	$g=1.9$
0.0001	-1.88410	-1.75053	-1.62838	-1.51752	-1.41753	-1.32774	-1.24728	-1.17520	-1.11054	-1.05239
0.0005	-1.82241	-1.70603	-1.59738	-1.49673	-1.40413	-1.31944	-1.24235	-1.17240	-1.10901	-1.05159
0.0010	-1.78572	-1.67825	-1.57695	-1.48216	-1.39408	-1.31275	-1.23805	-1.16974	-1.10743	-1.05068
0.0020	-1.74062	-1.64305	-1.55016	-1.46232	-1.37981	-1.30279	-1.23132	-1.16534	-1.10465	-1.04898
0.0050	-1.66390	-1.58110	-1.50114	-1.42439	-1.35114	-1.28167	-1.21618	-1.15477	-1.09749	-1.04427
0.0100	-1.58838	-1.51808	-1.44942	-1.38267	-1.31815	-1.25611	-1.19680	-1.14042	-1.08711	-1.03695
0.0200	-1.49188	-1.43529	-1.37929	-1.32412	-1.26999	-1.21716	-1.16584	-1.11628	-1.06864	-1.02311
0.0400	-1.36584	-1.32414	-1.28225	-1.24028	-1.19842	-1.15682	-1.11566	-1.07513	-1.03543	-0.99672
0.0500	-1.31684	-1.28019	-1.24313	-1.20578	-1.16827	-1.13075	-1.09338	-1.05631	-1.01973	-0.98381
0.1000	-1.12762	-1.10726	-1.08608	-1.06413	-1.04144	-1.01810	-0.99418	-0.96977	-0.94496	-0.91988
0.2000	-0.85161	-0.84809	-0.84369	-0.83842	-0.83223	-0.82516	-0.81720	-0.80837	-0.79868	-0.78816
0.3000	-0.61815	-0.62415	-0.62944	-0.63400	-0.63779	-0.64080	-0.64380	-0.64436	-0.64488	-0.64453
0.4000	-0.39434	-0.40638	-0.41794	-0.42899	-0.43949	-0.44942	-0.45873	-0.46739	-0.47538	-0.48265
0.5000	-0.16397	-0.17968	-0.19517	-0.21040	-0.22535	-0.23996	-0.25422	-0.26808	-0.28150	-0.29443
0.6000	0.08763	0.07032	0.05297	0.03560	0.01824	0.00092	-0.01631	-0.03344	-0.05040	-0.06718
0.7000	0.38111	0.36458	0.34772	0.33054	0.31307	0.29535	0.27740	0.25925	0.24094	0.22250
0.8000	0.75752	0.74537	0.73257	0.71915	0.70512	0.69050	0.67532	0.65959	0.64335	0.62662
0.9000	1.34039	1.34092	1.34047	1.33904	1.33665	1.33330	1.32900	1.32376	1.31760	1.31054
0.9500	1.87683	1.89395	1.90992	1.92472	1.93836	1.95083	1.96213	1.97227	1.98124	1.98906
0.9600	2.04269	2.06573	2.08758	2.10823	2.12768	2.14591	2.16293	2.17873	2.19332	2.20670
0.9750	2.38364	2.41984	2.45482	2.48855	2.52102	2.55222	2.58214	2.61076	2.63810	2.66413
0.9800	2.54206	2.58480	2.62631	2.66657	2.70556	2.74325	2.77964	2.81472	2.84848	2.88091
0.9900	3.02256	3.08660	3.14944	3.21103	3.27134	3.33035	3.38804	3.44438	3.49935	3.55295
0.9950	3.48874	3.57530	3.66073	3.74497	3.82798	3.90973	3.99016	4.06926	4.14700	4.22336
0.9980	4.08802	4.20582	4.32263	4.43839	4.55304	4.66651	4.77875	4.88971	4.99937	5.10768
0.9990	4.53112	4.67344	4.81492	4.95549	5.09505	5.23353	5.37087	5.50701	5.64190	5.77549
0.9995	4.96701	5.13449	5.30130	5.46735	5.63252	5.79673	5.95990	6.12196	6.28285	6.44251
0.9999	5.95691	6.18480	6.41249	6.63980	6.86661	7.09277	7.31818	7.54272	7.76632	7.98888

Πίν. 2.6: *Ο Παράγοντας Συχνότητας της Κατανομής Pearson III (Συνέχεια)*

$P(x)$	$g = 2.0$	$g = 2.1$	$g = 2.2$	$g = 2.3$	$g = 2.4$	$g = 2.5$	$g = 2.6$	$g = 2.7$	$g = 2.8$	$g = 2.9$
0.0001	-0.99990	-0.95234	-0.90908	-0.86956	-0.83333	-0.80000	-0.76923	-0.74074	-0.71429	-0.68966
0.0005	-0.99950	-0.95215	-0.90899	-0.86952	-0.83331	-0.79999	-0.76923	-0.74074	-0.71429	-0.68966
0.0010	-0.99900	-0.95188	-0.90885	-0.86945	-0.83328	-0.79998	-0.76922	-0.74074	-0.71428	-0.68965
0.0020	-0.99800	-0.95131	-0.90854	-0.86929	-0.83320	-0.79994	-0.76920	-0.74073	-0.71428	-0.68965
0.0050	-0.99499	-0.94945	-0.90742	-0.86863	-0.83283	-0.79973	-0.76909	-0.74067	-0.71425	-0.68964
0.0100	-0.98995	-0.94607	-0.90521	-0.86723	-0.83196	-0.79921	-0.76878	-0.74049	-0.71415	-0.68959
0.0200	-0.97980	-0.93878	-0.90009	-0.86371	-0.82959	-0.79765	-0.76779	-0.73987	-0.71377	-0.68935
0.0400	-0.95918	-0.92295	-0.88814	-0.85486	-0.82315	-0.79306	-0.76456	-0.73765	-0.71227	-0.68336
0.0500	-0.94871	-0.91458	-0.88156	-0.84976	-0.81927	-0.79015	-0.76242	-0.73610	-0.71116	-0.68759
0.1000	-0.89464	-0.86938	-0.84422	-0.81929	-0.79472	-0.77062	-0.74709	-0.72422	-0.70209	-0.68075
0.2000	-0.77686	-0.76482	-0.75211	-0.73880	-0.72495	-0.71067	-0.69602	-0.68111	-0.66603	-0.65086
0.3000	-0.64333	-0.64125	-0.63833	-0.63456	-0.62999	-0.62463	-0.61854	-0.61176	-0.60434	-0.59634
0.4000	-0.48917	-0.49494	-0.49991	-0.50409	-0.50744	-0.50999	-0.51171	-0.51263	-0.51276	-0.51212
0.5000	-0.30685	-0.31872	-0.32999	-0.34063	-0.35062	-0.35992	-0.36852	-0.37640	-0.38353	-0.38991
0.6000	-0.08371	-0.09997	-0.11590	-0.13148	-0.14665	-0.16138	-0.17564	-0.18939	-0.20259	-0.21523
0.7000	0.20397	0.18540	0.16682	0.14827	0.12979	0.11143	0.09323	0.07523	0.05746	0.03997
0.8000	0.60944	0.59183	0.57383	0.55549	0.53683	0.51789	0.49872	0.47934	0.45980	0.44015
0.9000	1.30259	1.29377	1.28412	1.27365	1.26240	1.25039	1.23766	1.22422	1.21013	1.19539
0.9500	1.99573	2.00128	2.00570	2.00903	2.01128	2.01247	2.01263	2.01177	2.00992	2.00710
0.9600	2.21888	2.22986	2.23967	2.24831	2.25581	2.26217	2.26743	2.27160	2.27470	2.27676
0.9750	2.68888	2.71234	2.73451	2.75541	2.77506	2.79345	2.81062	2.82658	2.84134	2.85492
0.9800	2.91202	2.94181	2.97028	2.99744	3.02330	3.04787	3.07116	3.09320	3.11399	3.13356
0.9900	3.60517	3.65600	3.70543	3.75347	3.80013	3.84540	3.88930	3.93183	3.97301	4.01286
0.9950	4.29832	4.37186	4.44398	4.51467	4.58393	4.65176	4.71815	4.78313	4.84669	4.90884
0.9980	5.21461	5.32014	5.42426	5.52694	5.62818	5.72796	5.82629	5.92316	6.01858	6.11254
0.9990	5.90776	6.03865	6.16816	6.29626	6.42292	6.54814	6.67191	6.79421	6.91505	7.03443
0.9995	6.60090	6.75798	6.91370	7.06804	7.22098	7.37250	7.52258	7.67121	7.81839	7.96411
0.9999	8.21034	8.43064	8.64971	8.86753	9.08404	9.29920	9.51301	9.72543	9.93643	10.14602

Πίν. 2.6: *Ο Παράγοντας Συχνότητας της Κατανομής PearsonIII (Συνέχεια)*

$P(x)$	$g = 3.0$	$g = 3.5$	$g = 4.0$	$g = 4.5$	$g = 5.0$	$g = 6.0$	$g = 7.0$	$g = 8.0$	$g = 9.0$
0.0001	−0.66667	−0.57143	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0005	−0.66667	−0.57143	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0010	−0.66667	−0.57143	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0020	−0.66667	−0.57143	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0050	−0.66666	−0.57143	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0100	−0.66663	−0.57143	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0200	−0.66649	−0.57142	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0400	−0.66585	−0.57136	−0.50000	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.0500	−0.66532	−0.57130	−0.49999	−0.44444	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.1000	−0.66023	−0.57035	−0.49986	−0.44443	−0.40000	−0.33333	−0.28571	−0.25000	−0.22222
0.2000	−0.63569	−0.56242	−0.49784	−0.44402	−0.39993	−0.33333	−0.28571	−0.25000	−0.22222
0.3000	−0.58783	−0.53993	−0.48902	−0.44114	−0.39914	−0.33330	−0.28571	−0.25000	−0.22222
0.4000	−0.51073	−0.49391	−0.46496	−0.43020	−0.39482	−0.33285	−0.28569	−0.25000	−0.22222
0.5000	−0.39554	−0.41253	−0.41265	−0.39985	−0.37901	−0.32974	−0.28528	−0.24996	−0.22222
0.6000	−0.22726	−0.27782	−0.31159	−0.32928	−0.33336	−0.31472	−0.28169	−0.24933	−0.22214
0.7000	0.02279	−0.05730	−0.12530	−0.17918	−0.21843	−0.25750	−0.25899	−0.24214	−0.22030
0.8000	0.42040	0.32171	0.22617	0.13737	0.05798	−0.06662	−0.14434	−0.18249	−0.19338
0.9000	1.18006	1.09552	1.00079	0.89964	0.79548	0.58933	0.40026	0.23929	0.11146
0.9500	2.00335	1.97147	1.92023	1.85300	1.77292	1.58541	1.37708	1.16295	0.95435
0.9600	2.27780	2.26862	2.23786	2.18874	2.12432	1.96048	1.76547	1.55444	1.33922
0.9750	2.86735	2.91299	2.93324	2.93105	2.90930	2.81743	2.67603	2.50001	2.30138
0.9800	3.15193	3.22641	3.27404	3.29767	3.30007	3.25128	3.14572	2.99810	2.82035
0.9900	4.05138	4.22473	4.36777	4.48303	4.57304	4.68680	4.72613	4.70514	4.63541
0.9950	4.96959	5.25291	5.50362	5.72400	5.91639	6.22616	6.44924	6.59931	6.68763
0.9980	6.20506	6.64627	7.05304	7.42733	7.77124	8.37634	8.88387	9.30709	9.65701
0.9990	7.15235	7.72024	8.25289	8.75202	9.21961	10.06812	10.81343	11.46855	12.04437
0.9995	8.10836	8.80779	9.47154	10.10110	10.69829	11.80316	12.80069	13.70366	14.52288
0.9999	10.35418	11.37334	12.35663	13.30504	14.22004	15.95660	17.57979	19.10191	20.53356

Πίν. 2.7: *Ο παράγοντας συχνότητας για αρνητικές τιμές του συντελεστή ασυμμετρίας στις κατανομές Pearson III και Γάμμα*

Συντελεστής ασυμμετρίας g	Αθροιστική πιθανότητα $P(x)$							
	0.01	0.50	0.80	0.90	0.96	0.98	0.99	0.995
−.0	−2.326	0	0.842	1.282	1.751	2.054	2.326	2.576
−.1	−2.400	0.017	0.846	1.270	1.716	2.000	2.252	2.482
−.2	−2.472	0.033	0.850	1.258	1.680	1.945	2.178	2.388
−.3	−2.544	0.050	0.853	1.245	1.643	1.890	2.104	2.294
−.4	−2.615	0.066	0.855	1.231	1.606	1.834	2.029	2.201
−.5	−2.686	0.083	0.856	1.216	1.567	1.777	1.955	2.108
−.6	−2.755	0.099	0.857	1.200	1.528	1.720	1.880	2.016
−.7	−2.824	0.116	0.857	1.183	1.488	1.663	1.806	1.926
−.8	−2.891	0.132	0.856	1.166	1.448	1.606	1.733	1.837
−.9	−2.957	0.148	0.854	1.147	1.407	1.549	1.660	1.749
−1.0	−3.022	0.164	0.852	1.128	1.366	1.492	1.588	1.664
−1.1	−3.087	0.180	0.848	1.107	1.324	1.435	1.518	1.581
−1.2	−3.149	0.195	0.844	1.086	1.282	1.379	1.449	1.501
−1.3	−3.211	0.210	0.838	1.064	1.240	1.324	1.383	1.424
−1.4	−3.271	0.225	0.832	1.041	1.198	1.270	1.318	1.351
−1.5	−3.330	0.240	0.825	1.018	1.157	1.217	1.256	1.282
−1.6	−3.388	0.254	0.817	0.994	1.116	1.166	1.197	1.216
−1.7	−3.444	0.268	0.808	0.970	1.075	1.116	1.140	1.155
−1.8	−3.499	0.282	0.799	0.945	1.035	1.069	1.087	1.097
−1.9	−3.553	0.294	0.788	0.920	0.996	1.023	1.037	1.044
−2.0	−3.605	0.307	0.777	0.895	0.959	0.980	0.990	0.995
−2.1	−3.656	0.319	0.765	0.869	0.923	0.939	0.946	0.949
−2.2	−3.705	0.330	0.752	0.844	0.888	0.900	0.905	0.907
−2.3	−3.753	0.341	0.739	0.819	0.855	0.864	0.867	0.869
−2.4	−3.800	0.351	0.725	0.795	0.823	0.830	0.832	0.833
−2.5	−3.845	0.360	0.711	0.771	0.793	0.798	0.799	0.800
−2.6	−3.889	0.368	0.696	0.747	0.764	0.768	0.769	0.769
−2.7	−3.932	0.376	0.681	0.724	0.738	0.740	0.740	0.741
−2.8	−3.973	0.384	0.666	0.702	0.712	0.714	0.714	0.714
−2.9	−4.013	0.390	0.651	0.681	0.683	0.689	0.690	0.690
−3.0	−4.051	0.396	0.636	0.660	0.666	0.666	0.667	0.667

Σημείωση:

Ειδικά για την κατανομή Γάμμα θεωρείται ότι καλύτερη εκτίμηση των παραμέτρων επιτυγχάνεται με τη μέθοδο του μεγίστου της πιθανοφάνειας. Σύμφωνα με την προσεγγιστική μέθοδο Thorn οι εκτιμήσεις των παραμέτρων δίνονται ως εξής:

$$\hat{¥} = \frac{1+\sqrt{1+4A/3}}{4A} - \Delta\hat{¥} \qquad \hat{} = \bar{¡}/\hat{¥}$$

ó®- 3

$$A = \ln \bar{¡} - \frac{1}{N}\sum_{1}^{N} \ln x_i$$

και η διόρθωση $\Delta \gamma$ προκύπτει ως συνάρτηση της αρχικής εκτίμησης 7 (χωρίς διόρθωση) από τον Πίν. 2.8

Πίν. 2.8: *Διόρθωση της παραμέτρου γ σύμφωνα με τη μέθοδο Thorn (1958)*

$\hat{\gamma}$	$\Delta\hat{\gamma}$	$\hat{\gamma}$	$\Delta\hat{\gamma}$	$\hat{\gamma}$	$\Delta\hat{\gamma}$	$\hat{\gamma}$	$\Delta\hat{\gamma}$
0.2	0.034	0.8	0.012	1.4	0.006	2.2	0.003
0.3	0.029	0.9	0.011	1.5	0.005	2.3	0.002
0.4	0.025	1.0	0.009	1.6	0.005	3.1	0.002
0.5	0.021	1.1	0.008	1.7	0.004	3.2	0.001
0.6	0.017	1.2	0.007	1.8	0.004	5.5	0.001
0.7	0.014	1.3	0.006	1.9	0.003	5.6	0.000

*Κατανομή Λογαριθμική Pearson **III** (Log Pearson **III**)*

- Λέγοντας ότι μια σειρά από παρατηρήσεις περιγράφεται ικανοποιητικά από την κατανομή Λογαριθμική Pearson **III** (Log Pearson **III**) σημαίνει ότι οι λογάριθμοι των παρατηρήσεων περιγράφονται από την κατανομή Pearson **III.** Επομένως η διαδικασία που ακολουθείται περιλαμβάνει:

i) Αντικατάσταση των τιμών των παρατηρήσεων χ_i από τους λογαρίθμους των y_i όπου $y_i = \log \chi_i$

ii) Υπολογισμό του μέσου όρου y, της τυπικής απόκλισης σ_v και του συντελεστή ασυμμετρίας *g*

iii) Υπολογισμό του μεγέθους y_T περιόδου επαναφοράς *T* ετών με τη μέθοδο του παράγοντα συχνότητας K_T

$$y_T = \bar{y} + \hat{\pm}_y K_T$$

Όπου K_T από τον Πιν.2.6 ή τον Πίν 2.7 της κατανομής Pearson III

iv) Υπολογισμός του μεγέθους χ_T=antilog y_T

Λογαριθμοκανονική Κατανομή (Κατανομή Galton)

Για την εφαρμογή της Λογαριθμοκανονική Κατανομής χρησιμοποιούνται οι λογάριθμοι των τιμών των παρατηρήσεων (συνήθως οι νεπέρειοι λογάριθμοι) και η κανονική κατανομή. Με τη χρησιμοποίηση του πίνακα της κανονικής κατανομής υπολογίζεται το μέγεθος y_T περιόδου επαναφοράς *T* από το οποίο

$$x_T = \exp y_T \qquad (2.46)$$

Η χρησιμοποίηση της Λογαριθμοκανονικής κατανομής γίνεται ευκολότερη με το αντίστοιχο χαρτί πιθανότητας.

Με τη χρήση του παράγοντα συχνότητας της κανονικής κατανομής $K_T = \zeta$ προκύπτει ότι:

$$y_T = \bar{y} + \hat{\pm}_Y \bullet z$$

2.7 ΕΛΕΓΧΟΣ ΚΑΤΑΛΛΗΛΟΤΗΤΑΣ ΤΗΣ ΘΕΩΡΗΤΙΚΗΣ ΚΑΤΑΝΟΜΗΣ ΠΙΘΑΝΟΤΗΤΑΣ

Οι μέθοδοι που αναπτύχθηκαν στις προηγούμενες παραγράφους προϋποθέτουν την επιλογή της θεωρητικής κατανομής πιθανότητας που προσαρμόζεται ικανοποιητικά στα δεδομένα του δείγματος. Υπάρχουν δύο πρακτικοί τρόποι επιλογής της κατανομής πιθανότητας:

α) με τη χρησιμοποίηση του χαρτιού πιθανότητας της αντίστοιχης θεωρητική κατανομής

β) με τη σύγκριση μεταξύ ιστογράμματος σχετικών συχνοτήτων και θεωρητικής καμπύλης πυκνότητας πιθανότητας.

Με χαρτί πιθανότητας εξετάζεται εμπειρικά αν τα σημεία που προκύπτουν από τις παρατηρήσεις μπορούν να παρασταθούν ικανοποιητικά με μια ευθεία γραμμή.

Τόσο ο πρώτος όσο και ο δεύτερος τρόπος είναι εμπειρικοί και ο έλεγχος είναι

ουσιαστικά υποκειμενικός. Περισσότερο αξιόπιστοι στατιστικοί έλεγχοι είναι ο έλεγχος χ^2 και ο έλεγχος Kolmogorov-Smirnov. Υπάρχουν και άλλοι έλεγχοι για τους οποίους ο αναγνώστης πρέπει να ανατρέξει στα βιβλία Στατιστικής.

Έλεγχος χ^2

- Αν οι παρατηρήσεις έχουν ενταχθεί σε *k* κλάσεις με O_i τον αριθμό των παρατηρήσεων στην κλάση *i* και E_i τον αριθμό των παρατηρήσεων που προκύπτει από την θεωρητική κατανομή για την ίδια κλάση (*Ei*= θεωρητική σχετική συχνότητα πολλαπλασιασμένη επί τον συνολικό αριθμό παρατηρήσεων), υπολογίζεται η ποσότητα

$$x_d^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad (2.47)$$

- Η ποσότητα αυτή συγκρίνεται με αυτήν που προκύπτει από την κατανομή χ^2 με *k* - *p* - 1 βαθμούς ελευθερίας όπου p ο αριθμός των παραμέτρων της θεωρητικής κατανομής που εκτιμώνται από τα δεδομένα. Η υπόθεση ότι οι παρατηρήσεις προέρχονται από ένα πληθυσμό της θεωρητικής κατανομής που εξετάζεται γίνεται δεκτή για ένα επίπεδο εμπιστοσύνης (1 - *α* που συνήθως εκλέγεται 0.75/0.99) αν

$$x_d^2 < x_{1-a,k-p-1}^2 \qquad (2.48)$$

- Η ποσότητα δεξιά της ανισότητας βρίσκεται στον πίνακα της κατανομής χ^2 για πιθανότητα υπέρβασης *α* και $v = k - \rho - 1$ βαθμούς ελευθερίας. Για ευκολία ένας τέτοιος πίνακας παρουσιάζεται στην παράγραφο αυτή (Πίν. 2.9).

- *Παρατήρηση*. Επειδή ο έλεγχος χ^2 είναι ευαίσθητος στις μικρές συχνότητες πρέπει σε κάθε κλάση να εξασφαλίζεται ελάχιστος αριθμός παρατηρήσεων ίσος με 3 (κατ' άλλους 5). Στη περίπτωση που η συνθήκη αυτή δεν επαληθεύεται προτείνεται να συμπυκνώνονται οι κλάσεις.

Έλεγχος Kolmogorov - Smirnov

- Κατά τον έλεγχο Kolmogorov - Smirnov υπολογίζεται για κάθε παρατήρηση του δείγματος η πραγματική και η θεωρητική αθροιστική πιθανότητα. Ο έλεγχος Kolmogorov-Smirnov είναι κατάλληλος για συνεχείς κατανομές. Η μεγαλύτερη διαφορά μεταξύ αυτών των δύο πιθανοτήτων Dmax συγκρίνεται με μια κρίσιμη τιμή *Dcr* για κάθε επίπεδο εμπιστοσύνης που υπάρχει στον αντίστοιχο πίνακα (Πίν. 2.10). Αν για το επίπεδο εμπιστοσύνης που επιλέγεται Dmax < *Dcr* τότε γίνεται δεκτή η υπόθεση ότι το δείγμα ανήκει σε πληθυσμό της αντίστοιχης κατανομής.

- Ο έλεγχος διευκολύνεται αν τα δεδομένα έχουν τοποθετηθεί σε χαρτί πιθανότητας. Σ' αυτή την περίπτωση (Σχ. 2.3) εκτιμάται η ποσότητα Dmax κατευθείαν πάνω στον άξονα της αθροιστικής πιθανότητας. Στο παράδειγμα του Σχ. 2.3 η Dmax εκτιμάται γραφικά ίση με 0.04 περίπου (τιμή με m = 19 του Πίν. 2.1). Για επίπεδο εμπιστοσύνης 1-α= 90% δηλαδή *α*= 0.10 και αριθμό παρατηρήσεων *N* = 31 η τιμή Dcr=0.22 (Πίν. 2.10). Επομένως γίνεται δεκτή η υπόθεση ότι το δείγμα ανήκει σε πληθυσμό της κανονικής κατανομής.

Πίν. 2.9: *Τιμές του χ2 για δοθείσες τιμές της πιθανότητας υπέρβασης α και του αριθμού των βαθμών ελευθερίας* ***ν (= k - ρ - 1)***

ν	α: .995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.798	19.768	39.087	42.557	45.772	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

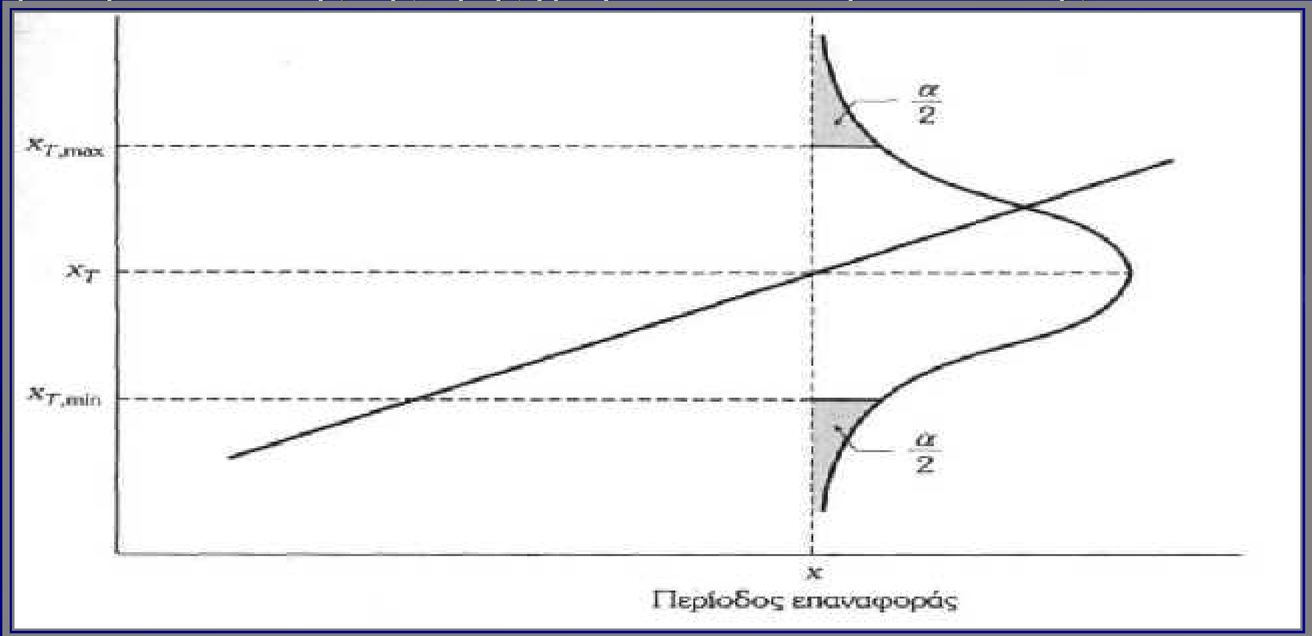
Πίν. 2.10: *Κρίσιμες τιμές Dcr του ελέγχου Kolmogorov-Smirnov*

Μέγεθος δείγματος N	α				
	0.20	0.15	0.10	0.05	0.01
1	0.90	0.93	0.95	0.98	0.99
2	0.68	0.73	0.78	0.84	0.93
3	0.57	0.60	0.64	0.71	0.83
4	0.49	0.53	0.56	0.62	0.73
5	0.45	0.47	0.51	0.56	0.67
6	0.41	0.44	0.47	0.52	0.62
7	0.38	0.41	0.44	0.49	0.58
8	0.36	0.38	0.41	0.46	0.54
9	0.34	0.36	0.39	0.43	0.51
10	0.32	0.34	0.37	0.41	0.49
11	0.31	0.33	0.35	0.39	0.47
12	0.30	0.31	0.34	0.38	0.45
13	0.28	0.30	0.33	0.36	0.43
14	0.27	0.29	0.31	0.35	0.42
15	0.27	0.28	0.30	0.34	0.40
16	0.26	0.27	0.30	0.33	0.39
17	0.25	0.27	0.29	0.32	0.38
18	0.24	0.26	0.28	0.31	0.37
19	0.24	0.25	0.27	0.30	0.36
20	0.23	0.25	0.26	0.29	0.35
25	0.21	0.22	0.24	0.26	0.32
30	0.19	0.20	0.22	0.24	0.29
35	0.18	0.19	0.21	0.23	0.27
40	0.17	0.18	0.19	0.21	0.25
45	0.16	0.17	0.18	0.20	0.24
50	0.15	0.16	0.17	0.19	0.23
$N > 50$	$\frac{1.07}{\sqrt{N}}$	$\frac{1.14}{\sqrt{N}}$	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

2.8 ΟΡΙΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

- Με τη χρησιμοποίηση των θεωρητικών κατανομών πιθανότητας η σχέση μεγέθους - περιόδου επαναφοράς είναι αρκετά αξιόπιστη για μεγέθη κοντά στο μέσο όρο. Με την επέκταση των τιμών για μεγάλες περιόδους επαναφοράς η πιθανότητα σφάλματος στην εκτίμηση του αντίστοιχου μεγέθους μεγαλώνει. Τα όρια εκατέρωθεν της τιμής χ_T που προκύπτει από την θεωρητική κατανομή για περίοδο επαναφοράς T βρίσκονται με τη θεώρηση ενός επιπέδου εμπιστοσύνης 1 - α. Σύμφωνα με ένα εμπειρικό τρόπο γίνεται η υπόθεση ότι η κατανομή πιθανότητας της χ_T είναι κανονική με αποτέλεσμα να χρησιμοποιούνται οι πίνακες της τυπικής κανονικής κατανομής. Δηλαδή επίπεδο εμπιστοσύνης 1 - α = 90 % σημαίνει ότι τα όρια επιλέγονται κατά τέτοιο τρόπο ώστε το εμβαδόν κάτω από την καμπύλη της συνάρτησης πυκνότητας πιθανότητας της κανονικής κατανομής να είναι 0.90 (Σχ. 2.6).

Σχ. 2.6: *Τα όρια εμπιστοσύνης της τιμής χT για επίπεδο εμπιστοσύνης 1-α.*



Τα όρια εμπιστοσύνης υπολογίζονται

$$X_{T,\max} = X_T + z_{1-a/2} S_T$$

$$X_{T,\min} = X_T + z_{1-a/2} S_T \qquad (2.49)$$

όπου $z_{1-a/2}$ η μεταβλητή της τυποποιημένης κανονικής κατανομής για αθροιστική πιθανότητα μεταξύ των ορίων 1 - *α*, και S_T η τυπική απόκλιση του χ_τ που είναι ίση:

$$S_T = \delta \frac{\hat{\sigma}}{\sqrt{N}} \qquad (2.50)$$

όπου:

$\hat{\sigma}$ = η τυπική απόκλιση του δείγματος μεγέθους N παρατηρήσεων και

δ = $1 + K_T^2 / 2)^{1/2}$ (για κανονική κατανομή)

(2.51)

δ = $\sqrt{1 + 1.3 K_T + 1.1 K_T^2}$ (για την κατανομή Gumbel)

Για την κατανομή Pearson III η τιμή της παραμέτρου *δ* δίνεται από πίνακα ως συνάρτηση της περιόδου επαναφοράς και του συντελεστή ασυμμετρίας (Πίν. 2.11).

Πίν. 2.11: *Οι τιμές της παραμέτρου δ στην Κατανομή Pearson III*

Συντελεστής ασυμμετρίας g	Περίοδος επαναφοράς T (έτη)					
	2	5	10	20	50	100
0.0	1.0801	1.1698	1.3748	1.6845	2.1988	2.6363
0.1	1.0808	1.2006	1.4367	1.7810	2.3425	2.8168
0.2	1.0830	1.2309	1.4989	1.8815	2.4986	3.0175
0.3	1.0866	1.2609	1.5610	1.9852	2.6656	3.2365
0.4	1.0918	1.2905	1.6227	2.0915	2.8423	3.4724
0.5	1.0987	1.3199	1.6838	2.1908	3.0277	3.7238
0.6	1.1073	1.3492	1.7441	2.3094	3.2209	3.9895
0.7	1.1179	1.3785	1.8032	2.4198	3.4208	4.2684
0.8	1.1304	1.4082	1.8609	2.5303	3.6266	4.5595
0.9	1.1449	1.4385	1.9170	2.6403	3.8374	4.8618
1.0	1.1614	1.4699	1.9714	2.7492	4.0522	5.1741
1.1	1.1799	1.5030	2.0240	2.8564	4.2699	5.4952
1.2	1.2003	1.5382	2.0747	2.9613	4.4896	5.8240
1.3	1.2223	1.5764	2.1237	3.0631	4.7100	6.1592
1.4	1.2457	1.6181	2.1711	3.1615	4.9301	6.4992
1.5	1.2701	1.6643	2.2173	3.2557	5.1486	6.8427
1.6	1.2952	1.7157	2.2627	3.3455	5.3644	7.1881
1.7	1.3204	1.7732	2.3081	3.4303	5.5761	7.5339
1.8	1.3452	1.8374	2.3541	3.5100	5.7827	7.8783
1.9	1.3690	1.9091	2.4018	3.5844	5.9829	8.2196
2.0	1.3913	1.9888	2.4525	3.6536	6.1755	8.5562

Λόγω της συμμετρίας της κανονικής κατανομής η διαδικασία υπολογισμού των ορίων εμπιστοσύνης διευκολύνεται με τη χρήση των πινάκων της αθροιστικής πιθανότητας για εμβαδόν από 0 έως z. Πα ευκολία δίνονται παρακάτω οι πιο συνηθισμένες τιμές του επιπέδου εμπιστοσύνης 1 - *α* και της αντίστοιχης μεταβλητής $z_{1-\alpha/2}$ (Πίν. 2.12).

Πίν. 2.12: *Η μεταβλητή της τυποποιημένης κανονικής κατανομής για αθροιστική πιθανότητα μεταξύ των ορίων* 1 - *α*

$1-\alpha$ (%)	$z_{1-\alpha/2}$
99	2.576
95	1.960
90	1.645
80	1.282
70	1.036

Εφαρμογή. Από την ανάλυση των ραγδαίων βροχών ενός Μετεωρολογικού Σταθμού προέκυψαν τα παρακάτω στοιχεία για το μέγιστο ύψος βροχής διάρκειας 15 min για την περίοδο ετών 1934-64 (Πίν. 2.13)

Πίν. 2.13: *Μέγιστα ύψη βροχής διάρκειας 15 min της περιόδου 1934-64*

Έτος	Ύψος Βροχής (mm)	Έτος	Ύψος Βροχής (mm)
1935	10.0	1950	6.0
1936	18.5	1951	5.2
1937	4.0	1952	7.0
1938	14.0	1953	8.0
1939	3.5	1954	26.1
1940	9.0	1955	4.5
1941	21.1	1956	19.1
1942	20.1	1957	4.0
1943	9.0	1958	12.2
1944	5.4	1959	13.5
1945	4.5	1960	3.8
1946	14.7	1961	14.3
1947	17.0	1962	6.0
1948	6.0	1963	5.0
1949	18.8	1964	13.0

1. Με τη χρησιμοποίηση της Κατανομής Ακραίων τιμών (π.χ. Gumbel) που προσαρμόζεται ικανοποιητικά στο δείγμα να υπολογισθεί το μέγιστο ύψος βροχής διάρκειας 15 min με περίοδο επαναφοράς 7=5, 10, 20, 50 και 100 έτη.
2. Να προσδιορισθούν τα όρια για επίπεδο εμπιστοσύνης 95% των μέγιστων υψών βροχής (διάρκειας 15 min) που αντιστοιχούν στις παραπάνω τιμές περιόδου επαναφοράς.
3. Να σχεδιασθεί η καμπύλη εντάσεως-συχνότητας της βροχής (διάρκειας 15 min) και η ζώνη εμπιστοσύνης για επίπεδο 95%.

Λύση. (α) Στον Πίνακα 2.14 γίνεται η μετατροπή των υψών βροχής σε εντάσεις και γίνεται κατάταξη των τιμών σε φθίνουσα σειρά από πλευράς μεγέθους. Υπολογίζεται κατόπιν η περίοδος επαναφοράς κάθε τιμής από την σχέση

$$T = \frac{N+1}{m}$$

όπου *N* ο αριθμός των παρατηρήσεων (*N* = 30) και *m* η τάξη μεγέθους στη φθίνουσα σειρά.

Προσδιορίζονται η μέση τιμή και η τυπική απόκλιση των εντάσεων *r* του δείγματος:

$$\overline{r} = \frac{\sum_{1}^{N} r}{N} = \frac{1293.2}{30} = 43.11 mm/hr$$

$$S = \sqrt{\frac{(r-\overline{r})^2}{N-1}} = \sqrt{\frac{18814{\bullet}2987}{29}} = 25.47 mm/hr$$

Η ένταση βροχής περιόδου επαναφοράς *T* υπολογίζεται σύμφωνα με τη μέθοδο του παράγοντα συχνότητας:

$$r_T = \overline{r} + \hat{\pm} K_T \qquad (2.52)$$

Όπου:

$\hat{\pm}$:τυπική απόκλιση δείγματος

$\overline{r}$:Μέση τιμή δείγματος

K_T :παράγοντας συχνότητας που είναι συνάρτηση του Τ.

Σ' αυτή την εφαρμογή ελέγχθηκε η προσαρμογή της Κατανομής Ακραίων Τιμών τύπου Ι (Gumbel) στα δεδομένα. Όπως φαίνεται στο Σχ. 2.7 όπου τα αποτελέσματα του Πίν. 2.14 έχουν τοποθετηθεί σε χαρτί πιθανότητας της κατανομής Gumbel, η κατανομή Gumbel θεωρείται ικανοποιητική.

Πίν. 2.14: *Υπολογισμός πιθανότητας υπερβάσεως*

m	Ύψος βροχής h_T (mm)	Ένταση βροχής r_T (mm/hr)	Περίοδος Επαναφοράς T (έτη)	Πιθαν. Υπέρβασης $1/T$
(1)	(2)	(3)	(4)	(5)
1	26.1	104.4	31.000	0.032
2	21.1	84.4	15.500	0.065
3	20.1	80.4	10.333	0.097
4	19.1	76.4	7.750	0.129
5	18.8	75.2	6.200	0.161
6	18.5	74.0	5.167	0.194
7	17.0	68.0	4.429	0.226
8	14.7	58.8	3.875	0.258
9	14.3	57.2	3.445	0.290
10	14.0	56.0	3.100	0.323
11	13.5	54.0	2.818	0.355
12	13.0	52.0	2.583	0.387
13	12.2	48.8	2.384	0.419
14	10.0	40.0	2.214	0.452
15	9.0	36.0	2.067	0.484
16	9.0	36.0	1.937	0.516
17	8.0	32.0	1.823	0.549
18	7.0	28.0	1.722	0.581
19	6.0	24.0	1.632	0.613
20	6.0	24.0	1.550	0.645
21	6.0	24.0	1.476	0.678
22	5.4	21.6	1.409	0.710
23	5.2	20.8	1.348	0.742
24	5.0	20.0	1.292	0.774
25	4.5	18.0	1.240	0.806
26	4.5	18.0	1.192	0.839
27	4.0	16.0	1.148	0.871
28	4.0	16.0	1.107	0.903
29	3.8	15.2	1.069	0.935
30	3.5	14.0	1.003	0.997

Για την Κατανομή Ακραίων Τιμών (Gumbel) που χρησιμοποιείται σ' αυτή την εφαρμογή ο παράγοντας συχνότητας υπολογίζεται:

$$K_T = -\frac{\left\{\ln\left[\ln\left(\frac{T}{T-1}\right)\right] + \bar{y}_n\right\}}{\pm_n} \qquad (2.53)$$

Με την βοήθεια του αντίστοιχου πίνακα για $N = 30$ προκύπτει

$\bar{y}_n = 0.53266$ $\qquad$ $\pm_n = 1.11238$

Από την Εξ. 2.52 και 2.53 σχηματίζεται ο Πίνακας 2.15

Πίν. 2.15: *Υπολογισμός μεγίστου ύψους βροχής ως συνάρτηση της περιόδου επαναφοράς*

T (έτη)	K_T	r_T (mm/hr)	h_T (mm)
(1)	(2)	(3)	(4)
5	0.8664	65.18	16.3
10	1.5409	82.36	20.6
20	2.1881	98.84	24.7
50	3.0257	120.17	30.0
100	3.6534	136.16	34.0

(β) Η τιμή της εντάσεως για μια ορισμένη περίοδο επαναφοράς που προκύπτει από την Εξ. 2.52 είναι η πιθανότερη τιμή r_T μιας σειράς τιμών που μπορεί να θεωρηθεί ότι προσαρμόζονται σε μια κανονική κατανομή με παραμέτρους:

Μέση τιμή: r_T (2.54)

Τυπική απόκλιση:

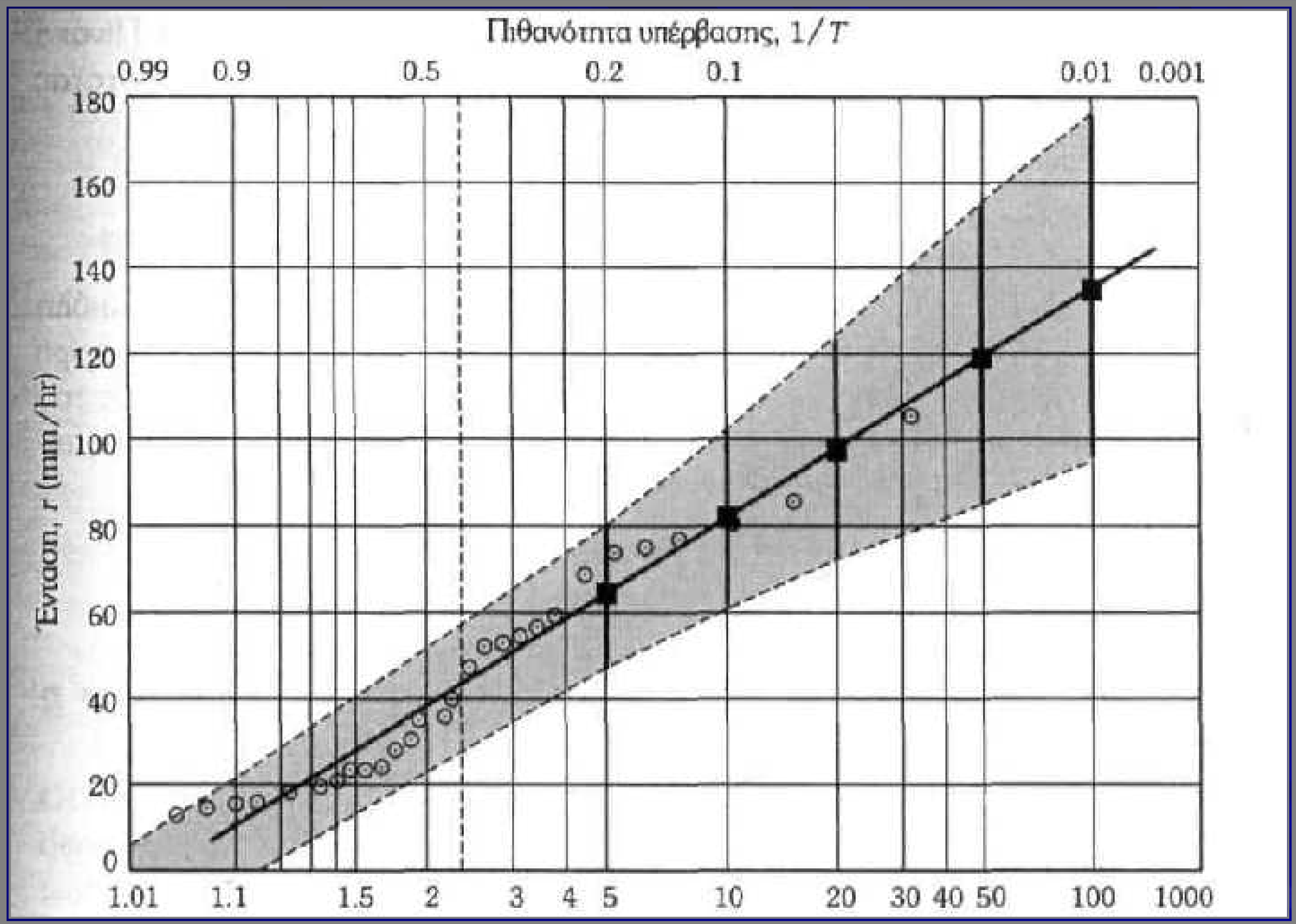
$$S_T = \left[\frac{\hat{\pounds}}{\sqrt{N}} \sqrt{(1 + bK_T + 1.10K_T^2)} \right] \qquad (2.55)$$

Όπου:

B=1.30(για N μικρό)

σ= τυπική απόκλιση δείγματος

K_T= παράγοντας συχνότητας



Περίοδος επαναφοράς, Τ(έτη)

2.7: *Καμπύλη εντάσεως-συχνότητος*

Το άνω και κάτω όριο εμπιστοσύνης δίνονται από τις εξισώσεις:

$$r_{T,\max} = r_T + zS_T \qquad\qquad r_{T,\min} = r_T + zS_T \qquad (2.56)$$

όπου ζ η ανηγμένη μεταβλητή της κανονικής κατανομής που αντιστοιχεί στο επιθυμητό όριο εμπιστοσύνης 1-α=0.95 δηλαδή (Πίν. 2.12) $z_{1-a/2} = 1.960$

Με βάση τον Πίνακα 2.15 και τις Εξισώσεις 2.55 και 2.56 προκύπτει ο Πίνακας 2.16.

Πίν. 2.16: *Υπολογισμός των ορίων εμπιστοσύνης* r_{Tmin}, r_{Tmax}, h_{Tmin} *και* $h_{T\ max}$ *ως συνάρτηση της περιόδου επαναφοράς*

T Έτη	K_T	r_T (mm/hr)	S_T (mm/hr)	$r_{T,min}$ (mm/hr)	$h_{T,min}$ (mm)	$r_{T,max}$ (mm/hr)	$h_{T,min}$ (mm)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5	0.8664	65.18	7.9897	49.52	12.4	80.84	20.2
10	1.5409	82.36	11.0190	60.76	15.2	103.96	26.0
20	2.1881	98.84	14.0363	71.33	17.8	126.35	31.6
50	3.0257	120.17	18.0123	84.87	21.2	155.47	38.9
100	3.6534	136.16	21.0193	94.96	23.7	177.36	44.3

Η τιμή του ζ στις Εξ. 2.56 μπορεί να εκτιμηθεί επίσης από τον Πίνακα 2.2 της τυπικής κατανομής για τιμή της αθροιστικής πιθανότητας 1 - α/2 = 97.5%: P (ζ) = 97.5% => (με γραμμική παρεμβολή), z = 1.96.

(γ) Στο Σχ. 2.7 σε χαρτί πιθανότητας Gumbel σχεδιάστηκε η καμπύλη έντασης - συχνότητας (ή έντασης - περιόδου επαναφοράς) που έχει μορφή ευθείας γραμμής, λόγω του ειδικού χαρτιού και της καταλληλότητας του σε ικανοποιητικό βαθμό. Στο ίδιο σχήμα παρουσιάζεται και η ζώνη εμπιστοσύνης για επίπεδο εμπιστοσύνης 95%.

Β' ΜΕΡΟΣ: ΣΤΟΙΧΕΙΑ ΣΤΟΧΑΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

2.9 ΧΡΟΝΟΣΕΙΡΕΣ ΥΔΡΟΛΟΓΙΚΩΝ ΜΕΓΕΘΩΝ

- Οι διαδοχικές παρατηρήσεις μιας υδρολογικής μεταβλητής (π.χ. παροχής) δεν είναι συνήθως ανεξάρτητες από τις προηγούμενες παρατηρήσεις όπως θεωρήθηκε στις προηγούμενες παραγράφους. Λόγω αυτής της έλλειψης ανεξαρτησίας, οι ιδιότητες των χρονοσειρών δεν μπορούν να αναλυθούν με βάση τη θεωρία των κατανομών πιθανότητας αλλά χρειάζονται πολυπλοκώτερα στατιστικά μοντέλα.

- Το πρώτο βήμα για να αντιληφθούμε την φύση των προβλημάτων που παρουσιάζει η ανάλυση χρονοσειρών είναι να εξετάσουμε οπτικά μερικές υδρολογικές και μετεωρολογικές χρονοσειρές.

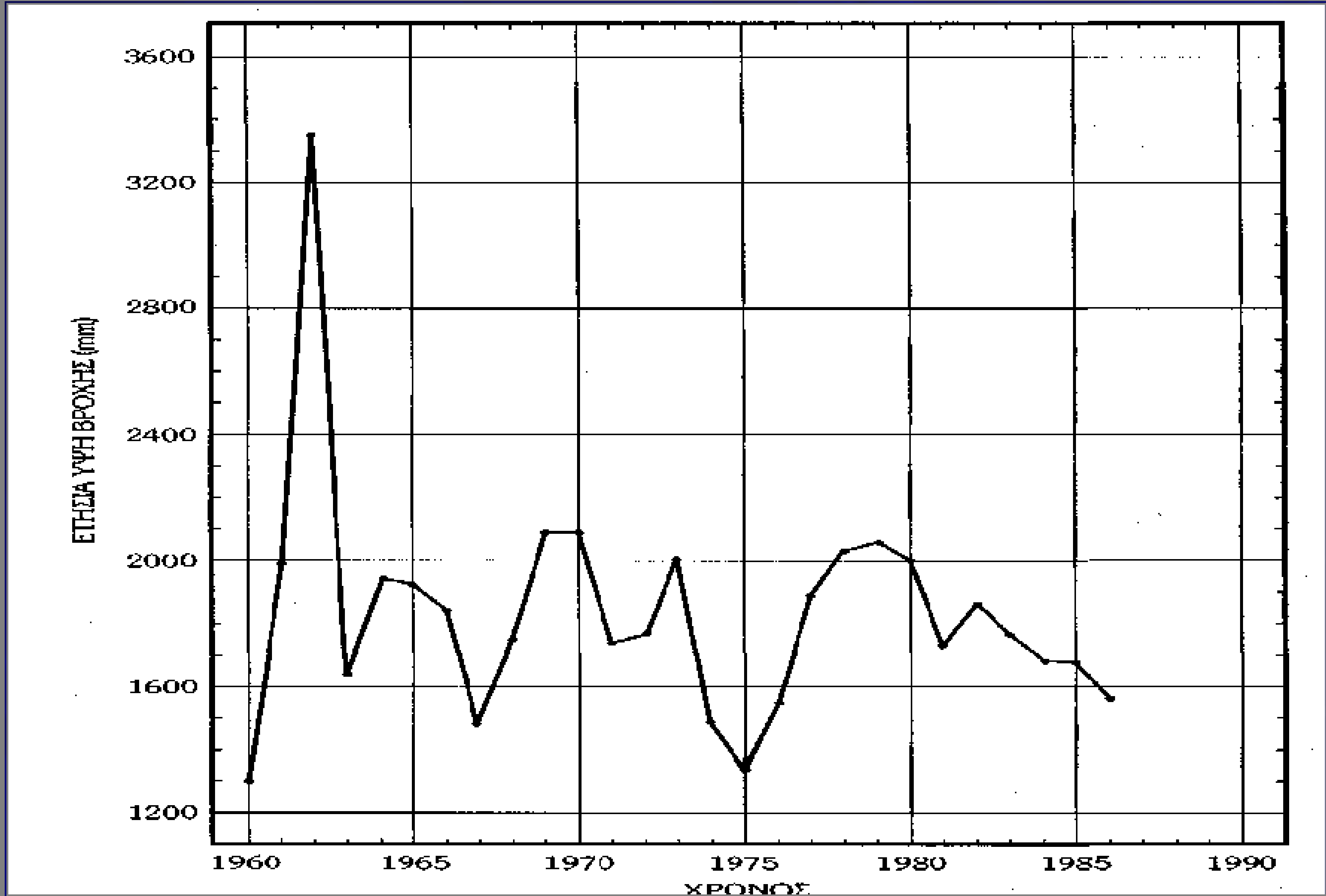
- Στο Σχ. 2.8 φαίνεται η χρονοσειρά ετησίων υψών βροχής του βαθμού Μεσοχώρας στη λεκάνη του Αχελώου ποταμού. Εξέταση της σειράς δείχνει ότι το μέγιστο του 1962 πρέπει να εξετασθεί ως ακραία τιμή. Τη σειρά αυτή θα τη χρησιμοποιήσουμε για την εφαρμογή των παρακάτω δοκιμών τυχαιότητας.

2.10 ΔΟΚΙΜΕΣ ΤΥΧΑΙΟΤΗΤΑΣ

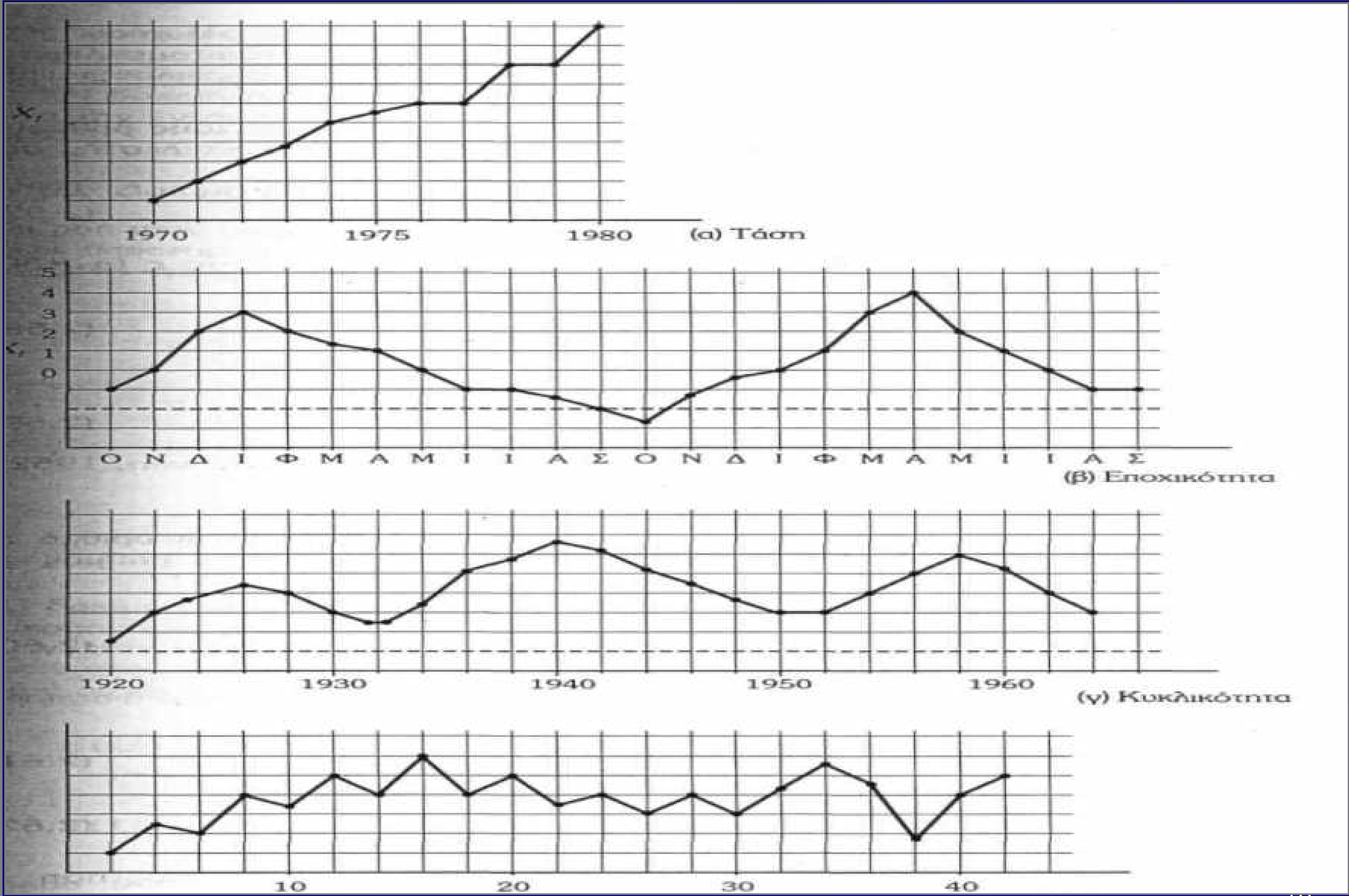
Μετά την εξέταση του γραφήματος μιας χρονοσειράς και πριν τη στατιστική ανάλυση είναι χρήσιμο να εξετασθεί αν η σειρά αποτελείται από ανεξάρτητες τυχαίες παρατηρήσεις. Στις περισσότερες περιπτώσεις ο μη τυχαίος χαρακτήρας της χρονοσειράς είναι προφανής αλλά υπάρχουν και υδρομετεωρολογικές σειρές (όπως π.χ. οι ετήσιες βροχοπτώσεις και οι ετήσιες παροχές) όπου χρειάζεται να ελεγχθεί η τυχαιότητα των παρατηρήσεων.

Η μηδενική υπόθεση είναι ότι οι παρατηρήσεις προέρχονται από την ίδια κατανομή και είναι ανεξάρτητες. Οι εναλλακτικές υποθέσεις είναι πολλές και περιλαμβάνουν όλες τις περιπτώσεις εξάρτησης των παρατηρήσεων π.χ. τάση, εποχιακότητα, κυκλικότητα και αυτοσυσχέτιση (Σχ. 2.9) και για το λόγο αυτό χρησιμοποιούνται συνήθως πολλές δοκιμές τυχαιότητας. Εδώ περιγράφονται μερικές μη παραμετρικές δοκιμές δηλαδή αυτές που δεν προϋποθέτουν παραδοχές σχετικά με την μορφή της κατανομής των παρατηρήσεων.

Σχ. 2.8: ***Ετήσια ύψη βροχής στο σταθμό Μεσοχώρας της λεκάνης του Αχελώου ποταμού.***



Σχ. 2.9: *Παραδείγματα χρονοσεφών με (α) Τάση, (β) Εποχικότητα, (γ) Κυκλικότητα, (δ) Αυτοσυσχέτιση.*



2.10.1 Δοκιμές Τάσης

(i) Συντελεστής Συσχέτισης Βαθμού Spearman

Η δοκιμή αυτή γίνεται ως εξής:

Οι παρατηρήσεις Χι,...,$χ_i$..., χn αντικαθίστανται από τους βαθμούς τους (ranks) y, κατά αυξανόμενη σειρά μεγέθους. Ο συντελεστής συσχέτισης βαθμού Spearman είναι

$$r_s = 1 - \frac{6}{n(n^2-1)} \sum (y_i - i)^2 \qquad (2.57)$$

Αν ισχύει η μηδενική υπόθεση η ασυμπτωτική κατανομή του r_s ($n > 30$) είναι κανονική με μέσο και διασπορά αντίστοιχα

$$E(r_s) = 0, \qquad V(r_s) = \frac{1}{n-1} \qquad (2.58)$$

Η δοκιμή γίνεται με την ανηγμένη μεταβλητή

$$Z = r_s \sqrt{n-1} \qquad (2.59)$$

Για $n < 30$ χρησιμοποιούμε πίνακες της κατανομής του r_s (Owen, 1962).

(ii)Συντελεστής Συσχέτισης Βαθμού Kendall

Για τη δοκιμή αυτή υπολογίζουμε για κάθε παρατήρηση χ, τον αριθμό *η*, των προηγουμένων παρατηρήσεων *xj* (j<i) των οποίων ο βαθμός yj είναι μικρότερος από το yi.
Η στατιστική συνάρτηση t της δοκιμής αυτής είναι

$$t = \sum_i n_i \qquad (2.60)$$

της οποίας η κατανομή είναι ασυμπτωτικά κανονική με μέσο και διασπορά αντίστοιχα

$$E(t) = \frac{n(n-1)}{4} \qquad (2.61) \qquad\qquad V(t) = \frac{n(n-1)(2n+5)}{72} \qquad (2.62)$$

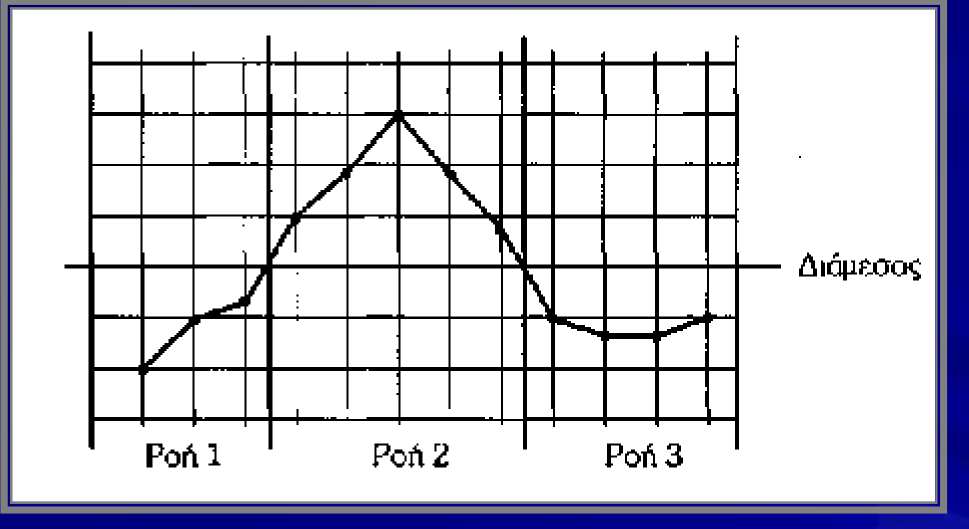
και συνεπώς η δοκιμή γίνεται με την ανηγμένη στατιστική συνάρτηση

$$u(t) = \frac{t - E(t)}{\sqrt{V(t)}} \qquad (2.63)$$

Οι δυο αυτές δοκιμές είναι εξίσου αποτελεσματικές και η ασυμπτωτική αποτελεσματικότητα τους είναι ίση με την αποτελεσματικότητα της γραμμικής παλινδρόμησης με το χρόνο.
Η δοκιμή Kendall έχει το πλεονέκτημα ότι μπορεί να χρησιμοποιηθεί για την εντόπιση του χρονικού σημείου εμφάνισης της τάσης (Sneyers, 1975)

(iii)Δοκιμή Ροών (Τάση και Εποχικότητα)

Η ροή μιας σειράς παρατηρήσεων είναι μια συνεχής σειρά αλληλοδιαδόχων παρατηρήσεων που όλες είναι μεγαλύτερες ή μικρότερες από ένα δεδομένο αριθμό, συνήθως τον διάμεσο των παρατηρήσεων (Σχ. 2.10).



Σχ. **2.10:** *Ορισμός ροών.*

Η δοκιμή ροών στηρίζεται στο ότι οι παρατηρήσεις είναι ανεξάρτητες και ακολουθούν την ίδια κατανομή και υπάρχουν α ροές με παρατηρήσεις μεγαλύτερες και β ροές με παρατηρήσεις μικρότερες από ένα δεδομένο αριθμό (συνήθως τον διάμεσο).

Ο μέσος και η τυπική απόκλιση του ολικού αριθμού ροών U είναι αντιστοίχως

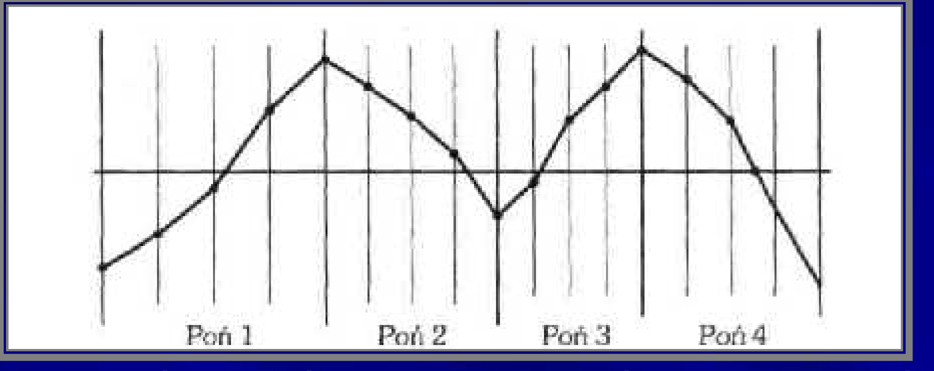
$$E(U) = 1 + \frac{2\sim}{n} \qquad (2.64)$$

$$S(U) = \sqrt{\frac{2\sim\ (2\sim\ - n)}{n^2(n^2 - 1)}} \qquad (2.65)$$

Όπου α+β=n

Για μικρά δείγματα χρησιμοποιούνται οι πίνακες κατανομής του *U* (Owen, 1962). Για μεγάλα δείγματα *(η* > 20) μπορούμε, να χρησιμοποιήσουμε τους πίνακες της ανηγμένης κανονικής μεταβλητής *(iv) Δοκιμή Ροών Ανερχομένων και Κατερχομένων Παρατηρήσεων*

Η ροή ανερχομένων παρατηρήσεων είναι μια σειρά αλληλοδιαδόχων παρατηρήσεων που η επόμενη είναι μεγαλύτερη της προηγούμενης. Ορίζουμε αντίστοιχα τη ροή κατερχομένων παρατηρήσεων.



Σχ. 2.11: *Ροές ανερχόμενων και κατερχόμενων παρατηρήσεων.*

Ο μέσος και η τυπική απόκλιση του ολικού αριθμού υ των ανερχομένων και των κατερχομένων ροών είναι αντίστοιχα

$$E(U) = \frac{2n-1}{n} \qquad (2.66)$$

$$S(U) = \sqrt{\frac{(16n-29)}{90}} \qquad (2.67)$$

(v) Δοκιμή Στασιμότητας της Διασποράς

Η στασιμότητα της διασποράς μιας σειράς μπορεί να ελεγχθεί με την εφαρμογή μιας δοκιμής τάσης στη σειρά των απολύτων τιμών των διαφορών των παρατηρήσεων από τον μέσο της σειράς.

2.10.2 Δοκιμή Αυτοσυσχέτισης

Ο συντελεστής αυτοσυσχέτισης *k* τάξης δίνεται από τον τύπο:

$$r_K = \frac{\sum_{i=1}^{n-k} (x_i - \overline{x})(x_{i+k} - \overline{x})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2}} \qquad (2.68)$$

Ο συντελεστής r_1 ακολουθεί ασυμπτωτικά την κανονική κατανομή με διασπορά 1/n και συνεπώς μπορεί να χρησιμοποιηθεί για τη δοκιμή τυχαιότητας μιας χρονολογικής σειράς.

2.11 ΣΤΟΧΑΣΤΙΚΕΣ ΑΝΕΛΙΞΕΙΣ

Η στατιστική ανάλυση ενός συνόλου παρατηρήσεων $(\chi_1,...,\chi_n)$ μπορεί να γίνει μόνο με βάση ένα μαθηματικό μοντέλο. Όταν υποθέτουμε ότι οι παρατηρήσεις αυτές είναι θεωρούμε εξαρτημένες, το μοντέλο της *στοχαστικής ανέλιξης*.

Οι παρατηρήσεις μιας χρονολογικής σειράς είναι συνήθως εξαρτημένες και συνεπώς για να προσδιορίσουμε τη δομή τους ή να προβλέψουμε τις μέλλουσες τιμές τους χρειαζόμαστε την υπόθεση ότι η σειρά αυτή παράγεται από μια στοχαστική ανέλιξη της οποίας υποθέτουμε γνωστή τη μορφή και εκτιμούμε τις παραμέτρους με βάση τις δεδομένες παρατηρήσεις. Η *στοχαστική ανέλιξη* είναι μια απειροπληθής οικογένεια συναρτήσεων *x(t)* τον χρόνο.

$$\left\{ x(t) t \pounds \mathrm{T} \right\} \qquad (2.69)$$

όπου Τ είναι το σύνολο των τιμών του γ. Στις περισσότερες εφαρμογές η ανέλιξη μπορεί να θεωρηθεί *διακριτή* (t = ±1, ±2,...). Σε άλλες περιπτώσεις η χρονολογική σειρά παράγεται από μία συνεχή στοχαστική ανέλιξη (tε Τ όπου *T* είναι διάστημα χρόνου).

Στην πιο γενική περίπτωση η ανέλιξη περιλαμβάνει προσδιοριστικές και τυχαίες συνιστώσες. Αυτό σημαίνει ότι δεν μπορούμε να προβλέψουμε ακριβώς την παρατήρηση σε χρόνο t αν γνωρίζουμε τις προηγούμενες τιμές. Για ένα δεδομένο Χρόνο *t* οι τιμές της *x(t)* ακολουθούν μια κατανομή πιθανότητας και συνεπώς το σύνολο $(\chi_1, \ldots, \chi n)$ των παρατηρήσεων είναι μια από τις άπειρες χρονολογικές σειρές ή *πραγματοποιήσεις* της στοχαστικής ανέλιξης (2.69). Ο ορισμός μιας στοχαστικής διαδικασίας μπορεί να βασισθεί στις από κοινού κατανομές για πεπερασμένα σύνολα τιμών του t (ti,...,t_k):

$$F_{t1}, \ldots . t_k(a_1, \ldots ., a_k) = P\left[x(t_1) \le a_1, \ldots, x(t_k) \le a_k\right] \qquad (2.70)$$

Για δεδομένο *k* η κοινή αυτή αθροιστική κατανομή εξαρτάται από τις *k* τιμές $t_{1;}, \ldots, t_k$. Ο Kolmogorov απέδειξε ότι κάτω από γενικές συνθήκες το σύστημα των κοινών αθροιστικών κατανομών (2.70) ορίζει τελείως την στοχαστική ανέλιξη (Κάκουλλος, 1978). Στις περισσότερες πρακτικές περιπτώσεις η μορφή των κοινών κατανομών (2.70) δεν είναι γνωστή. Στη γενική περίπτωση οι κατανομές αυτές εξαρτώνται από το χρόνο και συνεπώς το ίδιο ισχύει για τους μέσους και τις αυτοσυνδιασπορές.

$$E\left[x(t)\right] = \text{«}_x(t) \qquad (2.71)$$

$$E\left[(x(t) - \text{«}_x(t)), (x(t+r) - \text{«}_x(t+r))\right] = \text{¥}_{xx}(t, r) \qquad (2.72)$$

Οι παρατηρήσεις *(χ1,...,χn)* είναι μια μόνο πραγματοποίηση της στοχαστικής ανέλιξης (2.69) και συνεπώς για να προχωρήσει η ανάλυση χρειάζονται συμπληρωματικές υποθέσεις για τη δομή της στοχαστικής διαδικασίας. Αν υποθέσουμε ότι με δεδομένο το σύνολο *(t1,...,tk)* οι κοινές κατανομές (2.70) είναι οι ίδιες για οποιαδήποτε σταθερά μετατόπιση r των τιμών του t:

$$F_{t1,\dots tk}(a_1,\dots,a_k) = F_{t1+r},\dots,t_k + r(a_1,\dots,a_k) \qquad (2.73)$$

Οι στοχαστικές ανελίξεις που ικανοποιούν αυτή τη συνθήκη λέγονται *ισχυρώς στάσιμες* (strictly stationary) και για τις ανελίξεις αυτές αποδεικνύεται ότι:

(α) Η κατανομή πιθανότητας του *x(t)* είναι η ίδια για όλα τα t.

(β) Η αναμενόμενη τιμή του $\mu_\chi(\tau)$ είναι σταθερή και δεν εξαρτάται από το t.

(γ) Η συνάρτηση αυτοσυνδιασποράς $\gamma_{xx}(r)$ δεν εξαρτάται από το t.

Οι ιδιότητες (β) και (γ) είναι αναγκαίες αλλά όχι επαρκείς συνθήκες για να είναι μια ανέλιξη ισχυρώς στάσιμη. Μόνο στην περίπτωση που η ανέλιξη είναι κανονική δηλαδή όλες οι κατανομές (2.70) είναι κανονικές, οι ιδιότητες (β) και (γ) είναι επαρκείς συνθήκες για ισχυρώς στάσιμη ανέλιξη.

Μια στοχαστική ανέλιξη που ικανοποιεί τις συνθήκες (β) και (γ) χωρίς αναγκαστικά να ικανοποιεί τη συνθήκη (2.73) λέγεται *ασθενώς στάσιμη* (weakly stationary) ή στάσιμη κατά συσκέδαση (covariance stationary). Στην περίπτωση ασθενώς στάσιμων ανελίξεων ισχύει το *εργοδικό θεώρημα* που διατυπώνεται ως εξής: Αν μια ανέλιξη είναι ασθενώς στάσιμη ο μέσος μ_x και η αυτοδιασπορά $\gamma_{xx}(r)$ της ανέλιξης μπορούν να εκτιμηθούν από τον μέσο και την αυτοδιασπορά των παρατηρήσεων:

$$\hat{\mu}_x = \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t \qquad (2.74)$$

$$\hat{\gamma}_{xx}(r) = \frac{1}{n}\sum (x_t - \mu_x)(x_{t-r} - \mu_x) \qquad (2.75)$$

όπου n είναι ο αριθμός παρατηρήσεων και r η υστέρηση.
Σε μερικές περιπτώσεις οι στοχαστικές ανελίξεις των υδρολογικών χρονολογικών σειρών περιέχουν τάσεις και περιοδικές αυξομειώσεις και συνεπώς δεν ικανοποιούν τις συνθήκες των στάσιμων ανελίξεων.

Στις περιπτώσεις αυτές μπορούμε να χρησιμοποιήσουμε την ακόλουθη τροποποιημένη μορφή στοχαστικής ανέλιξης:

$$x(t) = \mu_x(t) + y(t) \qquad (2.76)$$

Όπου μ_y=0

$$\gamma_{xx}(r) = \gamma_{yy}(r) \qquad (2.77)$$

Στο μοντέλο αυτό $\mu_x(t)$ είναι προσδιοριστική συνάρτηση του χρόνου και *y(t)* ασθενώς στάσιμη ανέλιξη. Το εργοδικό θεώρημα ισχύει για την ανέλιξη *y(t)* και συνεπώς μπορούμε να εκτιμήσουμε την αυτοσυνδιασπορά $\gamma_{xx}(r)$ αν από τη χρονολογική σειρά αφαιρέσουμε την εκτιμημένη προσδιοριστική συνάρτηση $\mu_\chi(t)$. Η αφαίρεση αυτής της συνάρτησης που λέγεται τάση, εποχικότητα ή κυκλικότητα γίνεται με πολλές μεθόδους που θα μελετηθούν στη συνέχεια αυτού του κεφαλαίου.

2.12 ΜΟΝΤΕΛΑ ΧΡΟΝΟΣΕΙΡΩΝ

Η εμπειρική μέθοδος ανάλυσης χρονοσειρών βασίζεται στο μοντέλο:

$$X_i = T_i + C_i + S_i + E_i \qquad (2.78)$$

δηλαδή θεωρούμε ότι η χρονοσειρά είναι άθροισμα μιας τάσης *Ti* ενός κυκλικού όρου *Ci* ενός εποχιακού όρου Si και ενός τυχαίου όρου *Ei*. Στην περίπτωση που επικρατεί η εποχιακότητα με περίοδο p η ανάλυση γίνεται με τον ακόλουθο τρόπο (π.χ.Makridakis κ.ά., 1983):

1. Για κάθε παρατήρηση X, υπολογίζουμε τον κινητό μέσο m, με αριθμό όρων ίσο με την περίοδο. (Όταν p είναι άρτιος αριθμός, (π.χ. *p* = 12) ο κινητός μέσος mi υπολογίζεται ως μέσος των δύο μέσων)

$$\frac{xi - \frac{p}{2} + \ldots + xi + \frac{p}{2} - 1}{p} \quad \text{και} \quad \frac{xi - \frac{p}{2} + 1 + \ldots + xi + \frac{p}{2}}{p}$$

Κατά την υπόθεση μας η σειρά mi δεν περιλαμβάνει την εποχιακότητας Si και τον τυχαίο όρο Ei και συνεπώς mi= *Ti* + Ci δηλαδή ο κινητός μέσος mi με αριθμό όρων ίσο με ρ αντιπροσωπεύει το άθροισμα της τάσης και του κυκλικού όρου που αποτελείται από εναλασσόμενες θετικές και αρνητικές τιμές με μεταβλητή περίοδο.

2. Η σειρά (Xi- mi) = Si + Ei χρησιμοποιείται για να υπολογισθούν οι "εποχιακοί όροι" για κάθε περίοδο j (j = 1, 2,...,S) που κατά την υπόθεση είναι ίσοι με τον μέσο των ποσοτήτων (Si + Ei) για κάθε περίοδο j.

3. Εξετάζοντας τη σειρά mi = *Ti*+ C_i αποφασίζουμε ποια είναι η μαθηματική μορφή της τάσης Ti (π.χ. γραμμική, εκθετική, κτλ.) και υπολογίζουμε την εξίσωση *Ti*= *f(i)* με τη μέθοδο των ελαχίστων τετραγώνων.

4. Η σειρά *(mi- Ti)* κατά την υπόθεση μας είναι ίση με τον κυκλικό όρο Ci

5. Ο τυχαίος όρος Ei υπολογίζεται από την εξίσωση

$$X_i = T_i + C_i + S_i + E_i \qquad (2.79)$$

Στην περίπτωση που επικρατεί η τάση αρχίζουμε με την εκτίμηση της με τη μέθοδοτων ελαχίστων τετραγώνων (Στάδιο 3) και συνεχίζουμε με την εκτίμηση της εποχιακότητας (Στάδια 1 και 2). Η ανάλυση της χρονοσειράς Xi σε τάση, εποχιακότητα και κυκλικότητα μπορεί να γίνει και με πολλαπλή παλινδρόμηση όπου οι ανεξάρτητες μεταβλητές είναι μια συνάρτηση f(t) του χρόνου (τάση και κυκλικότητα) και ρ - 1 εικονικές μεταβλητές (εποχιακότητα). Η εξίσωση (για την περίπτωση που οι "εποχές" είναι μήνες) είναι:

$$X_i = \sim + \quad f(i) + \sum_{j=1}^{11} C_j D_j + E_i \qquad (2.80)$$

όπου f(i) = προσδιοριστική συνάρτηση του χρόνου και οι εικονικές μεταβλητές *Dj* είναι:

D_1 =1 για παρατηρήσεις Ιανουαρίου

= 0 για τους άλλους μήνες

D_2 = 1 για παρατηρήσεις Φεβρουαρίου

= 0 για τους άλλους μήνες

D_u = 1 για παρατηρήσεις Νοεμβρίου

= 0 για τους άλλους μήνες

εικονική μεταβλητή για τον Δεκέμβριο είναι ίση με $\left(1 - \sum_{j=1}^{11} D_j\right)$ και συνεπώς δεν χρειάζεται.

2.12.2 Μοντέλα ARIMA

Τα μοντέλα ARIMA βασίζονται στη θεωρία των στοχαστικών ανελίξεων και υποθέτουν ότι η παρατήρηση X, είναι άθροισμα τριών συνιστωσών:

(α) Μιας γραμμικής συνάρτησης των προηγουμένων παρατηρήσεων $\mathbf{X_{i-1},\ldots, X_{1-p}}$.

(β) Μιας γραμμικής συνάρτησης των τυχαίων ανεξαρτήτων μεταβλητών $\alpha_{\iota-1},..,\alpha_{i-q}$

(γ) Μιας τυχαίας μεταβλητής a_i

Τα μοντέλα ARIMA υποδιαιρούνται στις εξής περιπτώσεις:

(i) *Αυτοπαλινδρομικά μοντέλα (AR) τάξης p*

$$X_i = \Phi_1 X_{i-1} + \Phi_2 X_{i-2} + \ldots + \Phi_p X_{i-p} + a_i \qquad (2.81)$$

Όπου:

Xi: αρχικές ή μετασχηματισμένες παρατηρήσεις μετά την αφαίρεση του μέσου $X_i = x_i - \overline{x}$

$\Phi_1,\ldots,\Phi_p$: Παράμετροι

α_ι: ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την ίδια κατανομή.

(ii) Μοντέλα κινητού μέσου (MA) τάξης q

$$X_i = \sim_i - \text{¯}_1 \sim_{\S-1} - \ldots - \text{¯}_q a_{i-q} \qquad (2.82)$$

Όπου:

θ1,…,θq: είναι οι παράμετροι

ai,a_{i-1},…,a_{i-q}: τυχαίες μεταβλητές που ακολουθούν την ίδια κατανομή

(iii) Μικτά (ARMA) μοντέλα τάξης (ρ, q)

$$X_i = \text{¤}_1 X_{i-1} + \ldots + \text{¤}_p X_{i-p} + a_i - \text{¯}_1 a_{i-1} - \ldots - \text{¯}_q \sim_{i-p} \qquad (2.83)$$

Λεπτομερέστερη περιγραφή των μοντέλων ARIMA δίνεται στα βιβλία Μιμίκου (1990), Ξενάκης (1991), Makridakis κ.ά. (1983) και στο κλασικό βιβλίο των Box και Jenkins (1970).

Όταν υπάρχει τάση στη χρονοσειρά μπορούμε να χρησιμοποιήσουμε τις πρώτες, δεύτερες ή ανώτερης τάξης διαφορές των αρχικών τιμών και να υποθέσουμε ότι οι διαφορές αυτές ακολουθούν ανελίξεις αυτοσυσχέτισης, κινητού μέσου ή μικτές. Στην περίπτωση χρήσης διαφορών τα μοντέλα λέγονται ARIMA (I= integrated).
Όταν υπάρχει εποχιακότητα με περίοδο ρ μπορούμε να χρησιμοποιήσουμε τις εποχιακές διαφορές (Xi - *Xi-ρ)* των αρχικών παρατηρήσεων. Ο τρόπος αυτός να ληφθεί υπόψη η εποχικότητα δεν είναι συνήθως ικανοποιητικός για υδρολογικές σειρές γιατί οι διασπορές των μηνιαίων παρατηρήσεων δεν είναι ίσες. Ο συνήθης τρόπος αφαίρεσης της εποχικότητας στην υδρολογία είναι ο μετασχηματισμός

$y_i = \frac{x_i - \bar{x}_j}{s_j}$ όπου $\bar{x}_j$ και s_j είναι αντίστοιχα ο μέσος και η κανονική απόκλιση για την περίοδο j στην οποίαν ανήκει η παρατήρηση Xi

(iv) Ταυτοποίηση, εκτίμηση και επαλήθευση των μοντέλων ARIMA

Η ταυτοποίηση των μοντέλων ARIMA, δηλαδή η επιλογή της κατηγορίας (ARIMA, ARIMA) που αντιπροσωπεύει καλύτερα την συμπεριφορά των παρατηρήσεων, γίνεται με την εξέταση των συναρτήσεων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης (Ξενάκης 1991, Makridakis κ.ά. 1983).

Η συνάρτηση μερικής αυτοσυσχέτισης δίνεται από την εξίσωση:

$$\Phi_{\copyright\copyright} = \frac{\left|R_K^*\right|}{\left|R_k\right|}$$

Όπου:

$$R_k = \begin{bmatrix} 1 & r_1 & \cdots & r_{k-1} \\ r_1 & 1 & \cdots & r_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & & \cdots & 1 \end{bmatrix} \qquad (2.84)$$

και $\left|R_K^*\right|$ είναι η ορίζουσα του μητρώου R_K^* που προκύπτει από το μητρώο R_k όταν αντικατασταθεί η τελευταία στήλη με τη στήλη:

$$\left[r_1 \bullet r_2 \bullet\bullet\bullet r_k\right]'$$

Η εκτίμηση των παραμέτρων *φ^ θ*, και *σ„* μπορεί, για την περίπτωση των μοντέλων AR να γίνει με επίλυση συστήματος γραμμικών εξισώσεων αλλά η μέθοδος αυτή στηρίζεται στην υπόθεση ότι τα υπόλοιπα *at* είναι ασυσχέτιστα με τις παρατηρήσεις, που συνήθως δεν είναι ρεαλιστική. Ο συνήθης τρόπος εκτίμησης των παραμέτρων είναι η χρήση μη γραμμικών εξισώσεων που λύνονται με διαδοχικές προσεγγίσεις (Box και Jenkins 1970). Ο διαγνωστικός έλεγχος της καταλληλότητας του μοντέλου γίνεται με την εξέταση της συμπεριφοράς των υπολοίπων. Χρησιμοποιούμε την συνάρτηση:

$$Q = n(n+2)\sum_{i=1}^{k}(n-1)^{-1} r_i^2$$

όπου:

k : αριθμός συντελεστών αυτοσυσχέτισης

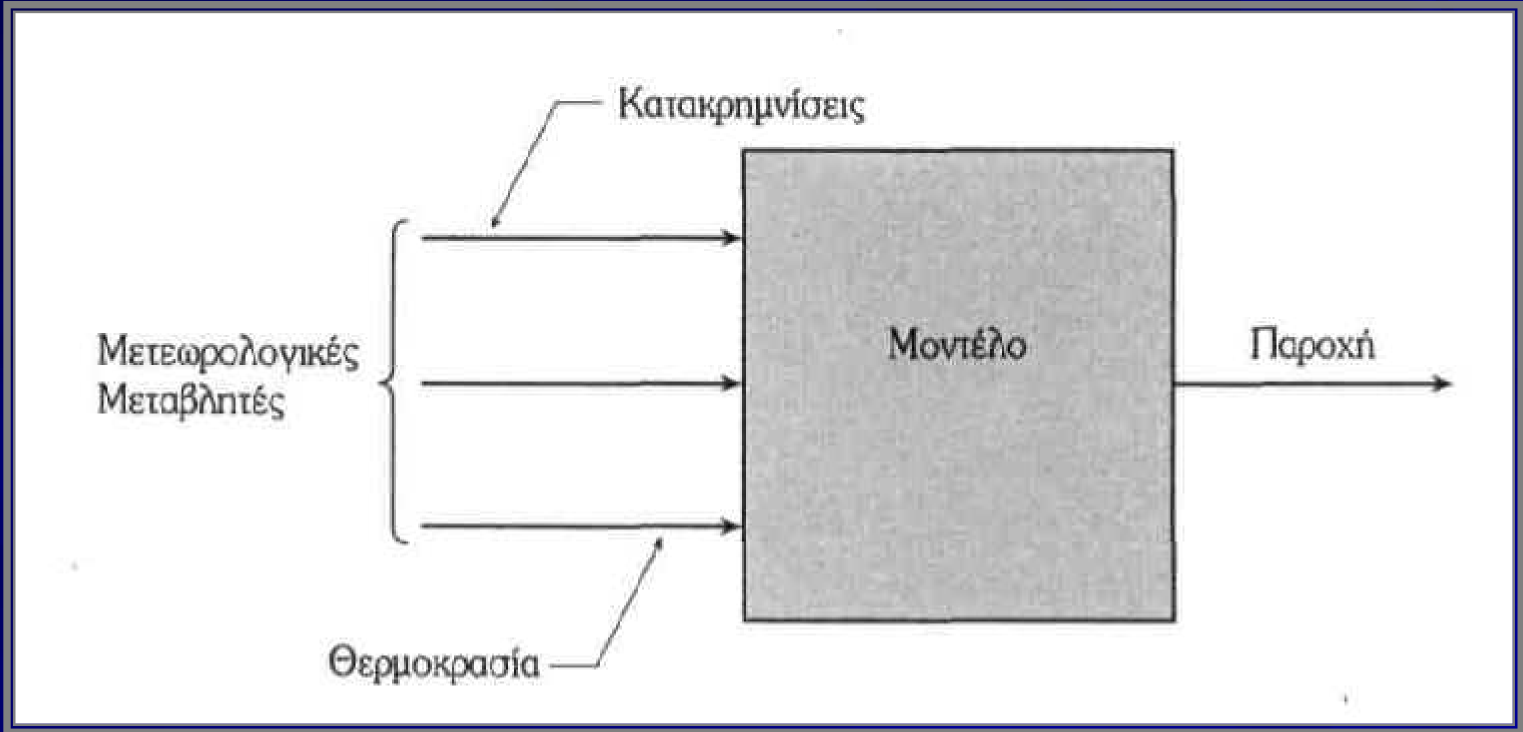
r_i : συντελεστής αυτοσυσχέτισης υπολοίπων τάξης i

Στην περίπτωση που το μοντέλο είναι ικανοποιητικό, τα υπόλοιπα είναι ασυσχέτιστα και η συνάρτηση Q ακολουθεί την κατανομή χ^2 με (k - (ρ + q +1)) βαθμούς ελευθερίας όπου (ρ + q +1) είναι ο αριθμός των παραμέτρων του μοντέλου.

2.13 ΣΤΟΧΑΣΤΙΚΗ ΥΔΡΟΛΟΓΙΑ

2.13.1 Σύντομη Ανασκόπηση

Η προσομοίωση υδρολογικών συστημάτων βασίζεται σε προσδιοριστικά ή στοχαστικά μοντέλα. Τα προσδιοριστικά μοντέλα που περιγράφονται στο Κεφάλαιο 14 έχουν συνήθως τη μορφή του Σχ. 2.12 και οι εξισώσεις τους απεικονίζουν κατά προσέγγιση τα φυσικά φαινόμενα που λαμβάνουν χώρα στη λεκάνη (π.χ. εξατμισοδιαπνοή, διήθηση, εδαφική υγρασία, αποθήκευση, κτλ.). Στα μοντέλα αυτά η ανάλυση των στατιστικών ιδιοτήτων των υπολοίπων δηλαδή των διαφορών μεταξύ των παρατηρήσεων και των υπολογιζόμενων παροχών γίνεται εκ των υστέρων για την επαλήθευση του μοντέλου. Αντίθετα, στα στοχαστικά μοντέλα δεν χρησιμοποιούνται εξισώσεις που περιγράφουν φυσικά φαινόμενα αλλά μαθηματικές εξισώσεις που απεικονίζουν τις στατιστικές ιδιότητες των μεταβλητών της εισόδου και της εξόδου του υδρολογικού συστήματος. Η προσαρμογή του μοντέλου θεωρείται ικανοποιητική αν τα υπόλοιπα είναι τυχαίοι αριθμοί χωρίς αυτοσυσχέτιση με τυπική απόκλιση μικρότερη από την τυπική απόκλιση των παρατηρήσεων.



Σχ. 2.12: *Σχηματική παράσταση προσδιοριστικού μοντέλου.*

Τα στοχαστικά υδρολογικά μοντέλα χρησιμοποιούνται για την εξερεύνηση της δομής των χρονοσειρών για δύο σκοπούς:

1. Την πρόβλεψη των μελλοντικών τιμών της μεταβλητής.
2. Την δημιουργία προσομοιωμένων τιμών της μεταβλητής που χρησιμοποιούνται για τη διαστασιολόγηση υδραυλικών έργων (Bras και Rodriguez -Iturbe, 1985, Μιμίκου, 1990).

<u>Πρόβλεψη</u>

Αν μια σειρά υδρολογικών παρατηρήσεων $\chi_1,\ldots,\chi_2,\ldots,\chi_\eta$ αποτελείται από ανεξάρτητες παρατηρήσεις, η παρατήρηση $\chi_{\eta+1}$ δεν μπορεί να προβλεφθεί από τις προηγούμενες και η καλύτερη πρόβλεψη της είναι ο μέσος της κατανομής των παρατηρήσεων $\chi_\iota,\ldots\chi_\eta$. Αντίθετα, αν οι παρατηρήσεις δεν είναι ανεξάρτητες μπορούμε να χρησιμοποιήσουμε τις προηγούμενες παρατηρήσεις για την πρόβλεψη, με βάση ένα στατιστικό μοντέλο (π.χ. τύπου ARIMA).

Προσομοίωση

Η προσομοίωση ετησίων παρατηρήσεων γίνεται ως εξής:

1. Προσαρμόζεται ένα στοχαστικό μοντέλο (π.χ. ARIMA) στις ετήσιες παρατηρήσεις.

2. Παράγεται μια σειρά τυχαίων αριθμών που ακολουθούν την κατανομή των υπολοίπων του μοντέλου.

3. Χρησιμοποιείται η εξίσωση του μοντέλου για την παραγωγή των προσομοιωμένων παροχών που, αν το μοντέλο έχει επιλεγεί σωστά, έχουν κατά προσέγγιση τα ίδια στατιστικά χαρακτηριστικά με τις παρατηρήσεις.

Στην περίπτωση μηνιαίων τιμών οι συντελεστές συσχέτισης διαδοχικών , μηνών μεταβάλλονται εποχιακά και πρέπει να χρησιμοποιηθεί μοντέλο που λαβαίνει υπ' όψη αυτή την ιδιότητα.

Το πρώτο μοντέλο αυτού του τύπου προτάθηκε από τους Thomas και Fiering1962, (Τσακίρης 1985) και δίνεται από την εξίσωση

$$x_{i+1} = \bar{x}_{j+1} + \quad {}_j (x_i - \bar{x}_j) + £_i s_{j+1} (1 - r_j^2)^{1/2} \qquad (2.86)$$

Όπου:

Xi και x_{i+1} :παροχές των μηνών I και i+1 αντίστοιχα(i=1,2,…,n)

$\bar{x}_{j+1}$ και $\bar{x}_j$:μέσες παροχές των μηνών j και j+1 αντίστοιχα(j=1,2,…,12)

βj :συντελεστής παλιδρόμησης

S_{j+1} :τυπική απόκλιση της παροχής του μηνός j+1

r_j :συντελεστής συσχέτισης μεταξύ x_j και x_{j+1}

$£_i S_{j+1} \sqrt{1 - r_j^2}$:τυχαία μεταβλητή που ακολουθεί την κατανομή των υπολοίπων της παλιδρόμησης

$$\bar{x}_{i+1} = \bar{x}_{j+1} + \quad {}_j (x_i - \bar{x}_j)$$

Το μοντέλο αυτό αποτελείται βασικά από τις εξισώσεις παλινδρόμησης των δώδεκα ζευγών διαδοχικών μηνών και υπολογίζεται με τον εξής τρόπο:

1. Εκτιμούμε τους μέσους xj και τις κανονικές αποκλίσεις sj για κάθε μήνα, και τους συντελεστές συσχέτισης *rj* και γραμμικής παλινδρόμησης *βj* για όλα τα ζεύγη μηνών (j, *j* + 1).
2. Παράγουμε μια σειρά ανεξαρτήτων τυχαίων αριθμών $£_i S_{j+1}\sqrt{1-r_j^2}$

που ακολουθούν την ίδια κατανομή με τους αριθμούς $(x_{i+1} - \hat{x}_{i+1})$ όπου

$\hat{x}_{i+1}$ =εκτιμημένη τιμή από παλινδρόμηση

$$\hat{x}_{i+1} = \hat{x}_{j+1} + \quad_j (x_i - \overline{x}_j) \qquad (2.87)$$

3. Χρησιμοποιώντας την παραπάνω εξίσωση παράγουμε προσομοιωμένες τιμές χ_{j+1} για κάθε μήνα

Λόγω της χρησιμότητας των προσομοιωμένων παροχών, το θέμα αυτό ερευνήθηκε εντατικά τις τελευταίες δεκαετίες. Τα μοντέλα που προτάθηκαν καλύπτουν τις ακόλουθες περιπτώσεις:

1. Μοντέλα προσομοίωσης εποχιακών παροχών αλληλοσυσχετισμένων σταθμών (Matalas, 1967, Τσακίρης, 1985).

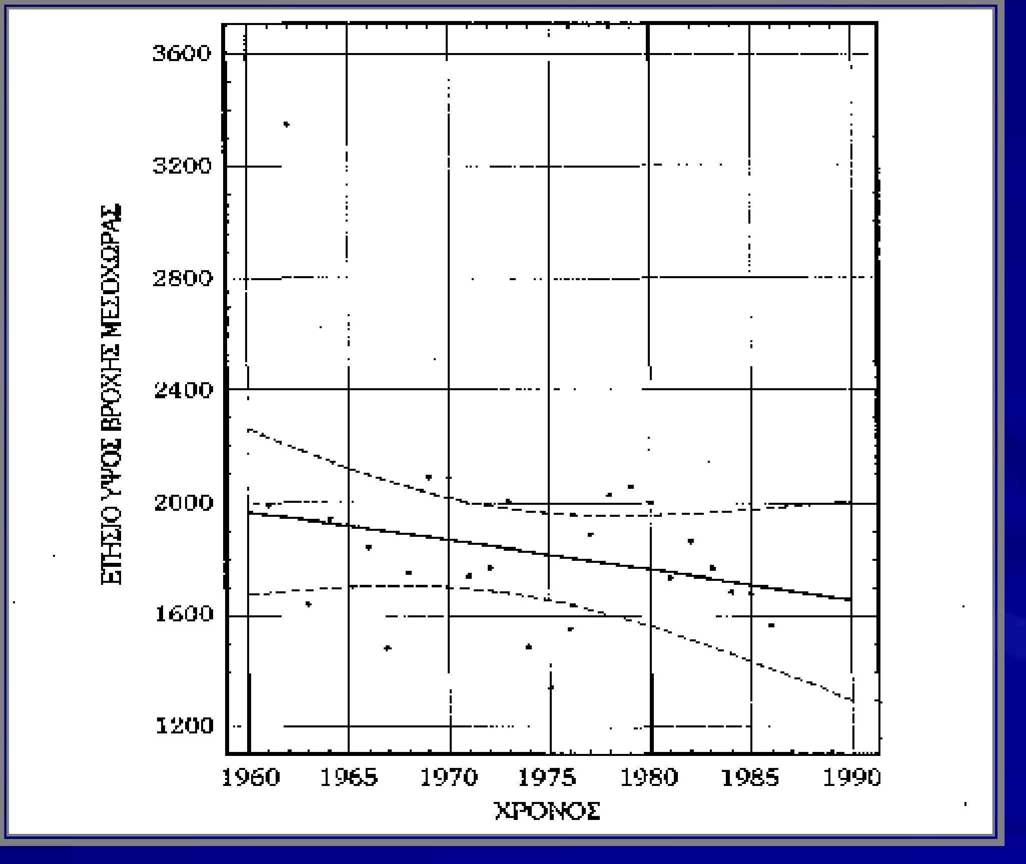
2. Μοντέλα επιμερισμού προσομοίωσης μηνιαίων παροχών που διατηρούν και τις ιδιότητες των ετησίων παροχών (Valencia και Schaake 1973).

3. Μοντέλα που διατηρούν τις ιδιότητες μακροχρόνιας εμμονής των χρονοσειρών (Συντελεστής Hurst).

Λεπτομέρειες των μοντέλων αυτών δίνονται στα βιβλία Salas κ.ά., (1980), Bras και Iturbe (1985), και Μιμίκου (1990).

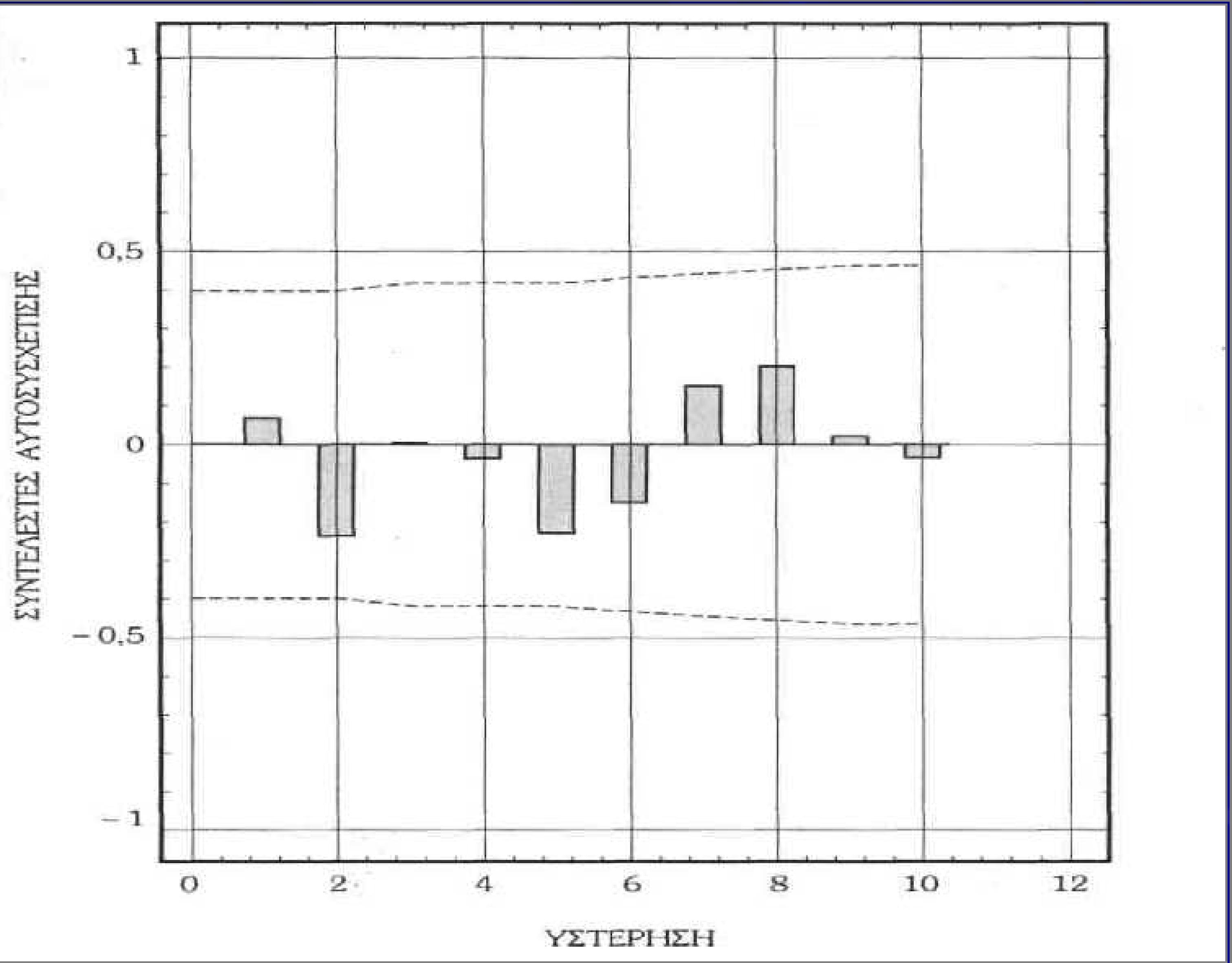
2.13.2 Παράδειγμα

Στο Σχήμα 2.8 δίνεται η σειρά ετήσιων βροχοπτώσεων στο σταθμό Μεσοχώρας Αχελώου. Στο Σχήμα 2.13 δίνεται η ίδια σειρά μαζί με την ευθεία παλινδρόμησης συναρτήσει του χρόνου.



Σχ. 2.13: *Σειρά ετήσιων βροχομετρικών υψών στο σταθμό Μεσοχώρας της λεκάνης του Αχελώου μαζί με την ευθεία παλινδρόμησης.*

Η κλίση της ευθείας αυτής (αρνητική τάση) δεν είναι στατιστικά σημαντική και συνεπώς οι ετήσιες τιμές της βροχόπτωσης μπορούν να θεωρηθούν τυχαίες (r = -0.218, P =0.27). Στο ίδιο συμπέρασμα καταλήγουμε αν χρησιμοποιήσουμε τους συντελεστές συσχέτισης βαθμού Spearman και Kendall μεταξύ της σειράς βροχοπτώσεων και του χρόνου. Οι συντελεστές αυτοί είναι Αντίστοιχα -1206 και -0.1198 και δεν είναι στατιστικά σημαντικοί (P (Spearman) = 0.54, P (Kendall) = 0.38). Το συμπέρασμα αυτό επιβεβαιώνεται από τις δοκιμές ροών. Στο Σχήμα 2.14 φαίνεται η συνάρτηση αυτοσυσχέτισης της χρονοσειράς. Όλοι οι συντελεστές αυτοσυσχέτισης είναι μέσα στο διάστημα εμπιστοσύνης και συνεπώς η δοκιμή αυτή επιβεβαιώνει το συμπέρασμα των προηγουμένων δοκιμών ότι η σειρά αποτελείται από τυχαίες παρατηρήσεις.



Σχ. 2.14: *Η συνάρτηση αυτοσυσχέτισης της χρονοσεφάς*