

**Αποσπάσματα από το βιβλίο
«Τεχνολογία Υδατικών Πόρων»
της Καθηγήτριας Μ.Α.Μιμίκου
Γ' έκδοση, 2006**

Κατανομές Πιθανοτήτων Υδρολογικών Μεταβλητών

Η πιθανολογική υδρολογική ανάλυση στηρίζεται στην υπόθεση χρονικής ανεξαρτησίας των υδρολογικών πραγματοποιήσεων και, αγνοώντας τη χρονική τους ακολουθία, τις κατατάσσει κατά τάξη μεγέθους. Αυτή η κατάταξη δίνει και τη βασικότερη ιστορική πληροφορία από το δείγμα των πραγματοποιήσεων της μεταβλητής που είναι η σχετική συχνότητα ή πιθανότητα υπέρβασης ή μη κάποιας συγκεκριμένης τιμής της μεταβλητής. Η κατανομή των τιμών της μεταβλητής στο πεδίο της συχνότητας ή κατανομή πιθανότητας ρυθμίζεται με βάση τα στατιστικά χαρακτηριστικά του δείγματος και αποτελεί την ταυτότητα πλέον της μεταβλητής αλλά και της διαδικασίας που κάθε υδρολογική μεταβλητή ορίζει.

Έχει λοιπόν πρωτεύουσα σημασία να γίνεται η επιλογή μέσα από τις υπάρχουσες θεωρητικές κατανομές πιθανοτήτων της καταλληλότερης για την εκάστοτε αναλύμενη μεταβλητή, αυτής δηλαδή που ταιριάζει περισσότερο στα ιστορικά στοιχεία. Στο κεφάλαιο αυτό γίνεται ακριβώς μια παρουσίαση των πιθανολογικών κατανομών που έχουν βρεθεί να προσομοιώνουν καλύτερα τις υδρολογικές μεταβλητές, των τρόπων εκτίμησης των παραμέτρων τους και των στατιστικών ελέγχων καταλληλότητας τους. Στο τέλος του κεφαλαίου αυτού γίνεται αναφορά και στους τυχαίους αριθμούς, δεδομένου ότι οι κατανομές πιθανοτήτων αποτελούν τη μοναδική τους ταυτότητα με χρήση της οποίας μπορούν να παραχθούν από τον υπολογιστή.

3.1. Γενικά για τις κατανομές πιθανοτήτων στην υδρολογία

Είναι πραγματικά πολύ μεγάλος ο αριθμός των θεωρητικών κατανομών πιθανοτήτων που η μαθηματική ανάλυση έχει προσφέρει στη διεθνή βιβλιογραφία (Chow, 1964· Benjamin και Cornell, 1970· Yevjevich, 1972· Bedient και Huber, 1992· Wanielista και άλλοι, 1997). Η πολύχρονη προσπάθεια ρύθμισης διαφόρων κατανομών πιθανοτήτων στις υδρολογικές

μεταβλητές και η εμπειρία που αποκτήθηκε οδηγούν ωστόσο σε κάποιες γενικές αρχές, που αφορούν χαρακτηριστικές ιδιότητες των κατάλληλων για διάφορες υδρολογικές μεταβλητές κατανομών, και περιορίζουν οπωσδήποτε το σύνολο των υποψηφίων πιθανολογικών κατανομών από τις οποίες πρέπει κανείς να διαλέξει. Οι ακόλουθες γενικές αρχές-ιδιότητες έχουν βρεθεί (Marcovic, 1965) ότι πρέπει να χαρακτηρίζουν τις κατανομές των ετήσιων κατακρημνίσεων και απορροών:

- Η κατανομή είναι συνεχής και ορισμένη για όλες τις θετικές τιμές της μεταβλητής.
- Ο άνω κλάδος της κατανομής είναι χωρίς όριο (unbounded).
- Ο κάτω κλάδος λαμβάνεται με όριο το μηδέν ή μια θετική τιμή.
- Η καμπύλη κατανομής πιθανοτήτων είναι ασυμπτωτική στον άξονα για μεγάλες θετικές τιμές της μεταβλητής.
- Ο μέγιστος αριθμός των παραμέτρων που περιγράφουν θεωρητικές κατανομές περιορίζεται στις τρεις.

Με βάση τα παραπάνω, η πιο γνωστή και κοινή μορφή κατανομής στην Υδρολογία είναι η μιας αιχμής κωδωνοειδής δύο κλάδων κατανομή με μεγάλο περιθώριο για τη θέση και τη μορφή της αιχμής της, χαρακτηριστικά που εκφράζονται αντίστοιχα από τους συντελεστές ασυμμετρίας και κύρτωσης.

3.2. Ορισμοί και στατιστικές παράμετροι

Οι ιστορικές πραγματοποιήσεις (μετρήσεις) μιας υδρολογικής μεταβλητής x αποτελούν το δείγμα της μεταβλητής που είναι ένα μέρος του πληθυσμού της, δηλαδή του συνόλου των πιθανών τιμών της στη συγκεκριμένη θέση. Η πιθανολογική θεώρηση του δείγματος απαιτεί την κατάταξη των τιμών της μεταβλητής κατά τάξη μεγέθους σε φθίνουσα σειρά (από τη μεγαλύτερη τιμή στη μικρότερη) ή αντίστροφα, σε αύξουσα σειρά. Οι σειρές μπορεί να είναι διακεκριμένες (discrete) ή συνεχείς (continuous). Η διαφορά της μεγαλύτερης από τη μικρότερη τιμή δίνει μια ιδέα των κλάσεων (ισοδιαστήματα τιμών της μεταβλητής) που θα πρέπει να χωριστεί το δείγμα (θα πρέπει να υπάρχουν τουλάχιστον 5 κλάσεις). Αν n είναι ο συνολικός αριθμός τιμών του δείγματος και m είναι ο αριθμός των πραγματοποιήσεων της μεταβλητής στην i^{th} κλάση ($i=1,2,...,n$) τότε η σχετική συχνότητα ή πιθανότητα γι' αυτή την κλάση είναι $p_i = m/n$. Η χάραξη των πιθανοτήτων (κατακόρυφος άξονας) για όλες τις κλάσεις τιμών της μεταβλητής (οριζόντιος άξονας) δίνει το ιστόγραμμα της κατανομής πιθανοτήτων, του οποίου το εμβαδόν θα πρέπει να ισούται με τη μονάδα. Για να γίνει αυτό θα πρέπει να διορθωθούν όλες οι τιμές p_i διαιρώντας τις με το εύρος των κλάσεων. Το ιστόγραμμα είναι η εμπειρική καμπύλη κατανομής πιθανοτήτων και πάνω σ' αυτό γίνεται η επιλογή της

θεωρητικής κατανομής $p(x)$ ανάλογα με το ποια εφαρμόζει καλύτερα.

Το πλοτάρισμα των αθροιστικών πιθανοτήτων: $\sum_{i=1}^k p_i$ για κάθε κλάση k , δίνει την εμπειρική αθροιστική καμπύλη πιθανοτήτων, στην οποία θα εφαρμοσθεί η αντίστοιχη θεωρητική: $F(x) = P(x \leq x')$ όπου x' είναι συγκεκριμένη τιμή της μεταβλητής x .

Ο υπολογισμός των παραμέτρων της κατανομής πιθανοτήτων γίνεται από τις στατιστικές παραμέτρους του ιστορικού δείγματος. Αυτές πάλι υπολογίζονται από τις στατιστικές ροπές (Feller 1957· Freeman, 1963), που για ένα δείγμα τιμών x_i ($i = 1, 2, \dots, n$) ορίζονται ως εξής:

$$r\text{-τάξεως ροπή από την αρχή: } J_r = \sum_{i=1}^n p_i x_i^r \quad (3.1)$$

$$r\text{-τάξεως κεντρική ροπή: } \mu_r = \sum_{i=1}^n p_i (x_i - \mu)^r \quad (3.2)$$

όπου p_i είναι η πιθανότητα που αντιστοιχεί στο x_i από την κατανομή πιθανοτήτων και μ είναι μια κεντρική τιμή του δείγματος. Στην Υδρολογία αρκούν συνήθως οι τέσσερις πρώτες ($r = 1, \dots, 4$) ροπές για να περιγράψουν τα χαρακτηριστικά μιας κατανομής πιθανοτήτων. Έτσι, οι στατιστικές παράμετροι του δείγματος είναι οι εξής:

Μέσος όρος (mean). Ο αριθμητικός μέσος όρος του πληθυσμού μ υπολογίζεται από την εξ. (3.1) με $r = 1$ και με την υπόθεση ότι όλες οι τιμές του έχουν την πιθανότητα: $p_i = 1/n$. Η εκτίμηση του $\bar{x}$ από το δείγμα είναι:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.3)$$

Ο γεωμετρικός μέσος όρος $\bar{x}_g$ δίνεται από την:

$$\bar{x}_g = (x_1 x_2 \dots x_n)^{\frac{1}{n}} \quad (3.4)$$

και ο λογάριθμός του ισούται με τον αριθμητικό μέσο των λογαρίθμων των τιμών του δείγματος.

Ο αρμονικός μέσος $\bar{x}_h$ δίνεται από την:

$$\bar{x}_h = \frac{n}{\sum_{i=1}^n (1/x_i)} \quad (3.5)$$

και σπάνια χρησιμοποιείται στην Υδρολογία.

Διάμεσος (median). Είναι η τιμή που αντιστοιχεί στο 50% της κατανομής πιθανοτήτων. Σε πολύ ασύμμετρες κατανομές ο μέσος όρος είναι αποπροσανατολιστικός, ενώ ο διάμεσος είναι καλύτερος αφού δίνει τη σταθερή πληροφορία ότι οι τιμές οι μεγαλύτερες ή οι μικρότερες απ' αυτόν συμβαίνουν πάντα στο μισό του χρονικού εύρους του δείγματος.

Πιθανότερος μέσος (mode). Είναι η τιμή που αντιστοιχεί στην αιχμή της κατανομής ($dp/dx = 0, d^2p/dx^2 < 0$).

Σκέδαση-Τυπική απόκλιση. Η σκέδαση του πληθυσμού σ^2 δίνεται από τη δεύτερη ($r=2$) κεντρική ροπή της εξ.(3.2) και πάλι με την παραδοχή της ίσης πιθανότητας: $p_i = 1/n$ και κεντρική τιμή το μέσο όρο. Η εκτίμηση του S^2 από το δείγμα γίνεται ως εξής:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3.6)$$

Η τετραγωνική ρίζα της σκέδασης S είναι η τυπική απόκλιση του δείγματος

$$S = \left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2} \quad (3.7)$$

Για μικρά δείγματα (μικρότερα του 30) γίνεται πολλαπλασιασμός της εξ.(3.6) επί $n/(n-1)$.

Συντελεστής ασυμμετρίας. Εκφράζει την ασυμμετρία της κατανομής (θετική για αιχμή δεξιά του μέσου όρου, αρνητική για αριστερά και μηδέν για συμμετρική κατανομή) και είναι το ηλίκο της τρίτης ($r=3$) κεντρικής ροπής προς τον κύβο της τυπικής απόκλισης. Η εκτίμηση του γ από το δείγμα γίνεται ως εξής:

$$\gamma = \frac{1}{nS^3} \sum_{i=1}^n (x_i - \bar{x})^3 \quad (3.8)$$

Για μικρά δείγματα η εξ.(3.8) πολλαπλασιάζεται επί $\frac{n^2}{(n-1)(n-2)}$.

Συντελεστής κύρτωσης. Εκφράζει την κυρτότητα της αιχμής της κατανομής και δίνεται από το πηλίκο της τέταρτης ($r = 4$) κεντρικής ροπής προς την τέταρτη δύναμη της τυπικής απόκλισης. Η εκτίμηση του k από το δείγμα γίνεται ως εξής:

$$k = \frac{1}{nS^4} \sum_{i=1}^n (x_i - \bar{x})^4 \quad (3.9)$$

Για μικρά δείγματα και πάλι γίνεται πολλαπλασιασμός της εξ.(3.9) επί $n^3/(n-1)(n-2)(n-3)$. Η κύρτωση των κατανομών συγκρίνεται συνήθως μ' αυτήν της κανονικής κατανομής που έχει $k = 3$.

Συντελεστής σκεδασιμότητας. Είναι ο λόγος της τυπικής απόκλισης προς το μέσο όρο και από το δείγμα εκτιμάται σαν:

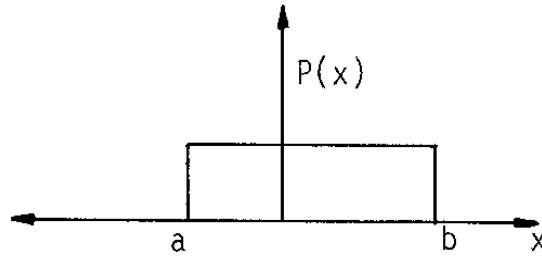
$$C_v = \frac{S}{\bar{x}} \quad (3.10)$$

3.3. Κατανομές πιθανοτήτων

Από την πληθώρα, των θεωρητικών κατανομών που διατίθενται στη διεθνή βιβλιογραφία εδώ θα αναφερθούν αυτές που έχουν υδρολογική σημασία. Οι κατανομές αναφέρονται σαν συνεχείς ή διακεκριμένες ανάλογα με το αν η μεταβλητή που προσομοιώνεται είναι συνεχής ή διακεκριμένη. Ιδιαίτερη έμφαση δίνεται στις κατανομές μεγίστων τιμών που είναι και οι πιο σημαντικές από πλευράς υδρολογικού σχεδιασμού υδατικών συστημάτων (Feller, 1957· Fisz, 1963· Hahn και Shapiro, 1967).

3.3.1. Ομοιόμορφη κατανομή

Είναι η πιο απλή κατανομή σε σχήμα ορθογωνικό, όπως φαίνεται στο Σχήμα 3.1.



Σχήμα 3.1. Ορθογωνική κατανομή.

Η (συνεχής) ομοιόμορφη κατανομή ορίζεται ως εξής:

$$\left. \begin{array}{ll} p(x)=0 & x < a \\ p(x)=1/(b-a) & a \leq x \leq b \\ p(x)=0, & b < x \end{array} \right\} \quad (3.11)$$

με τις εξής στατιστικές ιδιότητες για το μέσο όρο μ και την σκέδαση σ^2 (του πληθυσμού): $\mu = (b + a)/2$, $\sigma^2 = (b - a)^2/12$.

3.3.2. Διωνυμική κατανομή. Δοκιμές Bernoulli. Γεωμετρική κατανομή

Η διωνυμική είναι μια διακεκριμένη κατανομή που δίνει την πιθανότητα επισύμβασης ενός γεγονότος (επιτυχίας) ακριβώς x φορές σε n ανεξάρτητες προσπάθειες (δοκιμές), ενώ η πιθανότητα επιτυχίας σε κάθε δοκιμή είναι p . Η κατανομή ορίζεται ως εξής:

$$P(x) = C_x^n p^x q^{n-x}, \quad x = 0, 1, 2, \dots, n \quad (3.12)$$

όπου η πιθανότητα q είναι συμπληρωματική της p : $q=1-p$ και ο παραγοντικός δείκτης C_x^n δηλώνει τη σχέση: $C_x^n = n! / x!(n-x)!$. Η αθροιστική διωνυμική κατανομή ορίζεται ως εξής:

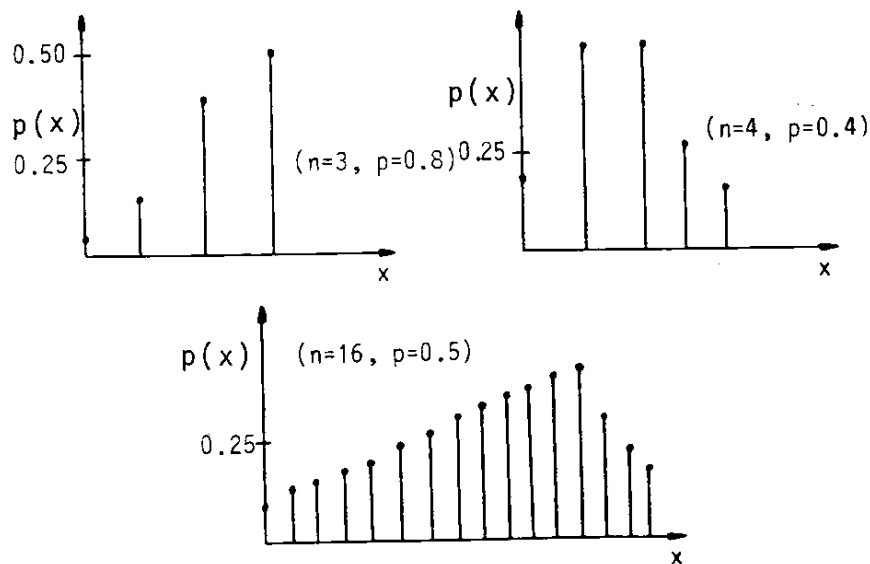
$$P(x \leq x') = C_0^n p^0 q^n + C_1^n p^1 q^{n-1} + \dots + C_{x_m}^n p^{x_m} q^{n-x_m} = \sum_{x \leq x'} C_x^n p^x q^{n-x} \quad (3.13)$$

όπου x_m είναι ο μεγαλύτερος ακέραιος που δεν ξεπερνά το x' .

Η διωνυμική ανάπτυξη της n -οστής δύναμης του μοναδιαίου αθροίσματος $(p+q)$ είναι η εξής:

$$(p + q)^n = q^n + C_1^n pq^{n-1} + \dots + C_x^n p^x q^{n-x} + \dots + p^n, \quad (3.14)$$

ενώ ισχύει: $\mu = np$ και $\sigma^2 = npq$. Για πολύ μεγάλο n ($\rightarrow \infty$) η μεταβλητή x είναι κανονική με μέσο όρο (np) και σκέδαση (npq). Έτσι, π.χ. η κατανομή για την επισύμβαση ή όχι βροχής σε μια μέρα για διάφορες τιμές του n και p δίνεται στο Σχήμα 3.2. Πίνακες της διωνυμικής κατανομής δίνονται σε εγχειρίδια στατιστικής (National Bureau of Standards, 1950).



Σχήμα 3.2. Διωνυμική κατανομή.

Όταν τα αποτελέσματα των διαφόρων προσπαθειών (δοκιμών) είναι αμοιβαία ανεξάρτητα και η πιθανότητα p επισύμβασης του γεγονότος (επιτυχίας) παραμένει αμετάβλητη έχουμε τις λεγόμενες δοκιμές Bernoulli (Bernoulli trials). Αυτές οι παραδοχές γίνονται συνήθως σε μελέτες σχεδιασμού φραγμάτων. Έτσι, αν έχουμε μια ακολουθία μέγιστων ετήσιων πλημμυρικών αιχμών για 20 χρόνια σε θέση φράγματος θεωρείται ότι τα μεγέθη είναι ανεξάρτητα μεταξύ τους και ότι η πιθανότητα επισύμβασης σε κάθε χρόνο της πλημμύρας παραμένει αμετάβλητη. Έχουμε δηλαδή 20 δοκιμές Bernoulli. Βέβαια κάποιος μπορεί να ισχυρισθεί ότι όταν συμβεί μια μεγάλη πλημμύρα, αλλάζει η διατομή στη θέση και άρα η πιθανότητα πρέπει να αλλάζει και αυτή από χρονιά σε χρονιά. Στο βαθμό που αυτό είναι αλήθεια, τότε το ομοίωμα για την πλημμυρική πιθανότητα είναι εσφαλμένο, αλλά και πάλι μπορεί να θεωρηθεί ότι ισχύει προσεγγιστικά. Ο αριθμός δοκιμών N ,

στον οποίο συμβαίνει για πρώτη φορά ένα γεγονός έχει γεωμετρική κατανομή. Η τελευταία δίνει την πιθανότητα πραγματοποίησης γεγονότος για την m -οστή φορά μετά από ακριβώς $(m+n-1)$ δοκιμές Bernoulli με πιθανότητα γεγονότος p . Είναι δε η εξής:

$$P(N=n) = P_N(n) = pq^{n-1}, n=1,2,\dots \quad (3.15)$$

με ιδιότητες $\mu = 1/p$ και $\sigma^2 = q/p^2$. Με βάση αυτή τη θεωρητική προσέγγιση, ερωτήματα όπως: Ποια είναι η πιθανότητα J να συμβεί τουλάχιστον μια φορά μια συγκεκριμένη πλημμύρα σχεδιασμού, π.χ. με ετήσια πιθανότητα επισύμβασης (υπέρβασης) $p = 1/100$ (ο παρονομαστής δηλώνει την περίοδο επαναφοράς T του γεγονότος) στα επόμενα n χρόνια ζωής του έργου, είναι απλό να απαντηθούν με τη λογική των δοκιμών Bernoulli: Η πιθανότητα q να μη συμβεί καθόλου κάθε χρόνο είναι: $q = 1 - p$, η πιθανότητα J' να μη συμβεί καθόλου στα n χρόνια είναι: $J' = (1-p)^n$ (λόγω ανεξαρτησίας των γεγονότων) άρα η πιθανότητα J να συμβεί τουλάχιστον μια φορά στα n χρόνια, που είναι η συμπληρωματική της J' , είναι η εξής:

$$J = 1 - (1 - p)^n = 1 - \left(1 - \frac{1}{T}\right)^n \quad (3.16)$$

Η πιθανότητα J λέγεται και διακινδύνευση σχεδιασμού για δεδομένη πλημμύρα ετήσιας πιθανότητας υπέρβασης p ή περιόδου επαναφοράς $T = 1/p$. Η J της εξ.(3.16) είναι και η αθροιστική κατανομή πιθανοτήτων της γεωμετρικής κατανομής.

3.3.3. Κατανομή Poisson

Σε μια στοχαστική διαδικασία Poisson τα γεγονότα συμβαίνουν τυχαία κατά τη διάρκεια του χρόνου t με μέσο χρονικό ρυθμό άφιξης λ . Αν ισχύουν οι συνθήκες (α) μονιμότητας (η πιθανότητα ενός γεγονότος μέσα σε ένα μικρό χρονικό διάστημα από t σε $t + h$ είναι περίπου λh για κάθε t), (β) μη πολλαπλασιαστικότητας (η πιθανότητα δύο ή περισσότερων γεγονότων σε μικρό χρονικό διάστημα είναι αμελητέα συγκρινόμενη με το λh) και (γ) ανεξαρτησίας (ο αριθμός των γεγονότων σε κάθε διάστημα είναι ανεξάρτητος του αριθμού κάθε άλλου χρονικού διαστήματος που δεν συμπίπτει με το προηγούμενο), τότε για κάθε διάστημα t ο αριθμός των γεγονότων x που θα συμβεί έχει μια κατανομή Poisson ως εξής:

$$P(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \quad x = 0, 1, 2, \dots \quad (3.17)$$

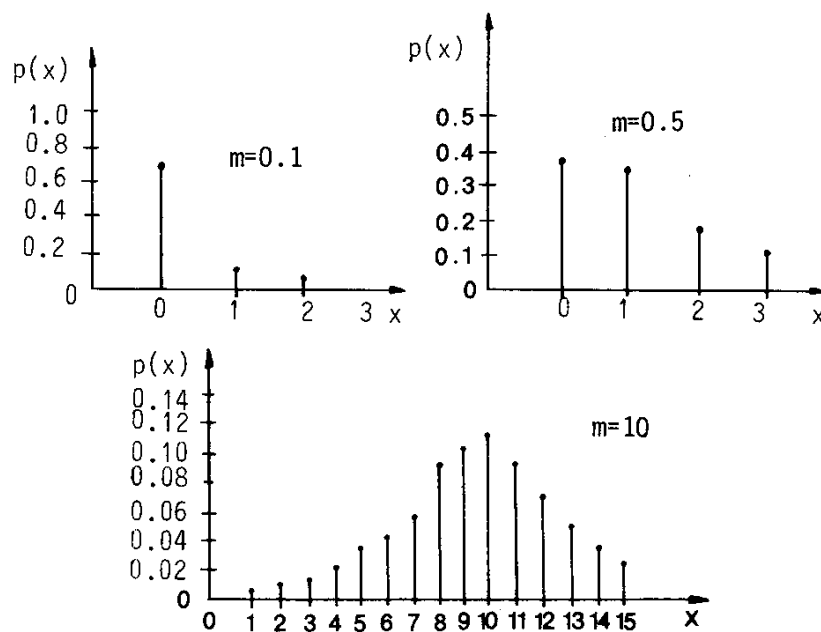
όπου $\mu = \sigma^2 = \lambda t$. Γενικά η κατανομή Poisson (διακεκριμένη) δίνεται από την:

$$P(x = x') = \frac{e^{-m} m^{x'}}{x!}, \quad m > 0, x = 0, 1, 2, \dots \quad (3.18)$$

Έστω οι αθροιστικές κατανομές ως εξής:

$$\left. \begin{aligned} P(x \geq x') &= \sum_{x=x'}^{\infty} \frac{e^{-m} m^x}{x!} \\ P(x \leq x'') &= \sum_{x=0}^{x''} \frac{e^{-m} m^x}{x!} \end{aligned} \right\} \quad (3.19)$$

Στο Σχήμα 3.3 δίνεται σχηματική η κατανομή Poisson για διάφορα m .



Σχήμα 3.3. Κατανομή Poisson.

Η κατανομή του χρόνου T μεταξύ των γεγονότων κατανομής Poisson είναι εκθετική με παράμετρο λ :

$$f_T(t) = \lambda e^{-\lambda t}, \quad t \geq 0 \quad (3.20)$$

και παραμέτρους: $\mu=1/\lambda$ και $\sigma^2=1/\lambda^2$.

Ο χρόνος t_k μεταξύ k γεγονότων κατανομής Poisson έχει κατανομή γάμμα:

$$f_{t_k}(t) = \lambda(\lambda t)^{k-1} \frac{e^{-\lambda t}}{(k-1)!}, \quad t \geq 0 \quad (3.21)$$

και παραμέτρους: $\mu = k/\lambda$ και $\sigma^2=k/\lambda^2$. Αν ο k δεν είναι ακέραιος τότε αντικαθίσταται το $(k-1)!$, με το $\Gamma(t)$ (γάμμα ολοκλήρωμα) στην εξ.(3.21). Στον Πίνακα Π-1 του παραρτήματος στο τέλος του πρώτου μέρους του βιβλίου δίνεται η αθροιστική κατανομή Poisson (μαζί με την χ^2 και τη γάμμα) ως: $F(u)$, όπου $u=(1/2)v-1$ (το v στην πρώτη στήλη του πίνακα, με μέσο όρο $\mu = m = 1/2\chi^2$). Η είσοδος στον πίνακα για την κατανομή Poisson γίνεται με την τιμή m και $v = 2(u+1)$.

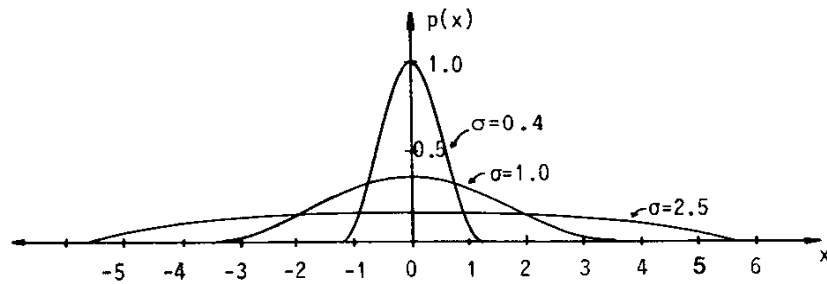
3.3.4. Κανονική κατανομή

Η κανονική κατανομή μεταβλητής x με μέσο όρο μ και σκέδαση σ είναι η εξής:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (3.22)$$

$$P(x \leq x') = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x'-\mu)^2/2\sigma^2} dx \quad (3.23)$$

και για διάφορες τιμές του σ δίνεται στο Σχήμα 3.4.



Σχήμα 3.4. Κανονική κατανομή.

Η κανονική κατανομή δίνεται συνήθως σε πίνακες (National Bureau of Standards, 1953) για αδιάστατους αριθμούς με μηδέν μέσο όρο, δηλαδή για αριθμούς που έχει αφαιρεθεί ο μέσος όρος και έχουν διαιρεθεί με την τυπική απόκλιση τους. Στον Πίνακα Π-2 του παραρτήματος δίνεται η αθροιστική τυπική κανονική κατανομή. Η χρήση πινάκων μπορεί όμως να αποφευχθεί μέσω της χρήσης διαφόρων προσεγγίσεων της κανονικής συνάρτησης κατανομής που έχουν αναπτυχθεί και δίνονται στη βιβλιογραφία (Press και άλλοι, 1987· Stedinger και άλλοι, 1993). Η χρήση των προσεγγίσεων αυτών είναι πλεονεκτικότερη όταν οι υπολογισμοί γίνονται με ηλεκτρονικό υπολογιστή.

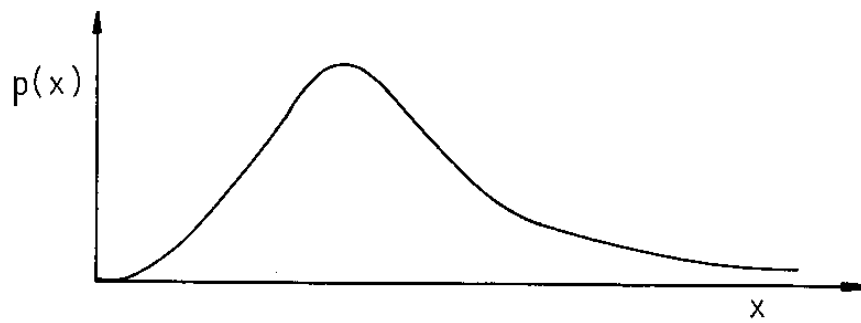
3.3.5. Λογαριθμική κανονική κατανομή (Aitchinson και Brown, 1957)

Ο λογαριθμικός μετασχηματισμός $y = \ln x$ ή $y = \ln (x-x_0)$ χρησιμοποιείται συχνά για την κανονικοποίηση μιας σειράς x . Έτσι η κατανομή με παραμέτρους $\bar{y}$ και σ_y γίνεται:

$$p(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-(y-\bar{y})^2 / 2\sigma_y^2}, \quad -\infty < y < \infty \quad (3.24)$$

ή με βάση τα x (όταν $x_0=0$ και η κατανομή διπαραμετρική)

$$p(x) = \frac{1}{\sigma_{\ln x} \sqrt{2\pi}} e^{-(\ln x - \overline{\ln x})^2 / 2\sigma_{\ln x}^2}, \quad (\text{Galton}) \quad (3.25)$$



Σχήμα 3.5. Λογαριθμική κανονική κατανομή.

Η εξ.(3.25) αρχίζει στο μηδέν και δεν έχει όριο προς τα δεξιά όπως φαίνεται στο Σχήμα 3.5.

Η λογαριθμική κανονική κατανομή χρησιμοποιείται συνήθως στην κατανομή πιθανοτήτων μεγίστων τιμών (π.χ. πλημμυρών) και θα περιγραφεί περισσότερο στο σχετικό Κεφ. 7.

3.3.6. Κατανομή γάμμα

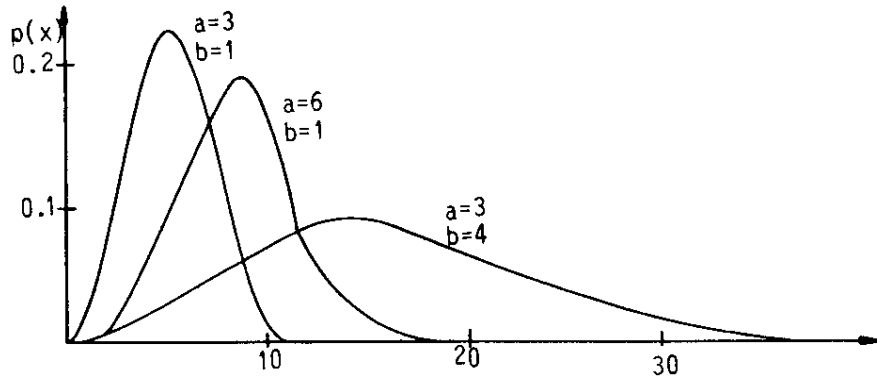
Η γάμμα κατανομή έχει την εξής μορφή:

$$p(x) = \frac{x^a e^{-x/b}}{b^{a+1} \Gamma(a+1)} \quad (3.26)$$

και η αθροιστική κατανομή,

$$P(x \leq x') = \int_{-\infty}^{x'} p(x) dx = \int_0^{x'} p(x) dx \quad (\text{ατελής συνάρτηση γάμμα}) \quad (3.27)$$

με ιδιότητες $\mu = b(a+1)$ και $\sigma^2 = b^2(a+1)$ (πίνακες της ατελούς γάμμα στον Hald, 1952). Σχηματικά η κατανομή γάμμα για διάφορες τιμές των παραμέτρων της φαίνεται στο Σχήμα 3.6. Η κατανομή γάμμα χρησιμοποιείται και για να αναπαραστήσει την κατανομή του χρόνου x_k μέχρι την k δοκιμή (ή άφιξη) μιας διαδικασίας Poisson με παραμέτρους k και λ όπως δίνεται στην εξ.(3.21). Στον Πίνακα Π-1 του παραρτήματος δίνεται η αθροιστική γάμμα ως: $1 - F(y)$, όπου $y = x^2/2\lambda$ με παραμέτρους $k = \nu/2$ και λ . Η είσοδος στον Πίνακα Π-1 γίνεται με $\nu = 2k$ και $x^2 = 2\lambda y$.



Σχήμα 3.6. Κατανομή γάμμα.

3.3.7. Κατανομές Pearson. Λογαριθμική Pearson III

Η γενική κατανομή έχει τη μορφή:

$$P(x) = e^{-\int_0^x [(a+x)/(b_0+b_1x+b_2x^2)] dx} \quad (3.28)$$

όπου a , b_0 , b_1 και b_2 είναι σταθερές. Τα κριτήρια για τον προσδιορισμό του τύπου της κατανομής είναι: $\beta_1 = \mu_3 / \mu_2^3$ και $\beta_2 = \mu_4 / \mu_2^2$, όπου μ_2 , μ_3 , μ_4 η δεύτερη, τρίτη και τέταρτη ροπή αντίστοιχα της μεταβλητής, ενώ χαρακτηριστικός είναι και ο δείκτης β που προκύπτει από τα β_1 και β_2 ως εξής:

$$\beta = \frac{\beta_1(\beta_2 + 3)^2}{4 + (4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)} \quad (3.29)$$

Για $\beta_1 = 0$, $\beta_2 = 3$, $\beta = 0$ έχουμε την κανονική κατανομή, όταν δε το $\beta_2 \neq 3$ τότε έχουμε την Pearson τύπου II. Για $\beta < 0$ έχουμε τον τύπο I, που είναι μια ασύμμετρη κατανομή με περιορισμούς και προς τις δύο κατευθύνσεις, συνήθως κωδωνοειδούς σχήματος, που θα μπορούσε να είναι και σχήματος J ή V. Η έκφραση της είναι:

Τύπος I:

$$p(x) = p_0 \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2}, \quad \text{με} \quad \frac{m_1}{a_1} = \frac{m_2}{a_2} \quad (3.30)$$

$$\text{όπου: } m_1 \text{ ή } m_2 = \frac{1}{2} \left[(r-2)^2 \pm r(r+2) \frac{\sqrt{\mu_2 \beta_1}}{2(\alpha_1 + \alpha_2)} \right],$$

$$r = \frac{6(\beta_2 - \beta_1 - 1)}{6 + 3\beta_1 - 2\beta_2}, \quad \alpha_1 + \alpha_2 = 1/2 [\mu_2 [\beta_1 (r+2)^2 + 16(r+1)]]^{1/2}$$

και

$$p_0 = \frac{N}{a_1 + a_2} \frac{m_1^{m_1} m_2^{m_2}}{(m_1 + m_2)^{m_1 + m_2}} \frac{\Gamma(m_1 + m_2 + 2)}{\Gamma(m_1 + 1)\Gamma(m_2 + 1)}$$

με N τη συνολική συχνότητα.

Η ασυμμετρία Pearson δίνεται από την: $(\beta_1^{1/2}/2)[(r+2)/(r-2)]$, ενώ οι ιδιότητες είναι: $\mu = \text{mode} - (\mu_3/2\mu_2)[(r+2)/(r-2)]$ και $\sigma^2 = \mu_2$.

Για $\beta = \infty$ ή $2\beta_2 - 3\beta_1 + 6$ προκύπτει η Pearson τύπου III, που είναι μια ασύμμετρη κατανομή με περιορισμούς προς τα αριστερά, συνήθως κωδωνοειδής, αλλά μπορεί να είναι και σχήματος J . Η έκφραση της είναι:

Τύπος III:

$$p(x) = p_0 \left(1 + \frac{x}{a} \right)^c e^{-cx/2} \quad (3.31)$$

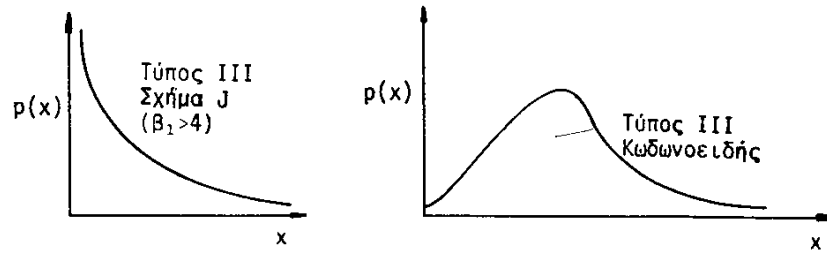
οπού:

$$c = 4/\beta_1 - 1, \alpha = c/2 (\mu_3/\sigma^2), \quad p_0 = \frac{N}{a} \frac{c^{c+1}}{e^c \Gamma(c+1)}$$

με ασυμμετρία Pearson: $\beta_1^{1/2}/2$ και ιδιότητες: $\mu = \text{mode} - \mu_3/2\mu_2$ και $\sigma^2 = \mu_2$.

Στο Σχήμα 3.7 δίνονται οι κατανομές Pearson τύπου III για διάφορα σχήματα.

Πολύ συχνά η κατανομή Pearson III εφαρμόζεται στις ασύμμετρες σειρές μεγίστων που έχουν προέλθει από λογαριθμικό μετασχηματισμό των μεγίστων ετήσιων πλημμυρικών παροχών. Η κατανομή τότε λέγεται Λογαριθμική Pearson III (Log - Pearson III). Λεπτομέρειες για την εφαρμογή αυτή θα δοθούν στο σχετικό Κεφ. 7.



Σχήμα 3.7. Κατανομές Pearson.

3.3.8. Θεωρία και ασυμπτωτικές κατανομές ακραίων τιμών (Gumbel 1954a, 1954b)

Έστω ότι Y είναι η μέγιστη τιμή των n τυχαίων μεταβλητών $x_1, x_2, \dots, x_n$. Τότε η αθροιστική πιθανότητα: $P(Y \leq y)$ ισούται με την αθροιστική πιθανότητα: P (όλες οι τιμές $x_i \leq y$). Για ανεξάρτητα μεταξύ τους x_i , η πιθανότητα γίνεται:

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(x_1 \leq y) P(x_2 \leq y) \dots P(x_n \leq y) = \\ &= F_{x_1}(y) F_{x_2}(y) \dots F_{x_n}(y) = F_{x_i}^n(y) \end{aligned} \quad (3.32)$$

Αν τώρα $f_{x_i}(y)$ είναι η κατανομή της x_i τότε:

$$F_Y(y) = dF_Y(y) / dy = n F_{x_i}^{n-1}(y) f_{x_i}(y) \quad (3.33)$$

Φαίνεται από την εξ.(3.33) ότι η επίδραση της $f_{x_i}(y)$ είναι πολύ μικρή. Γι' αυτό δεν είναι αναγκαίο να ξέρει κανείς από πρώτα την ακριβή μορφή της κατανομής της x_i για να υπολογίσει τη μορφή της κατανομής της y . Η αθροιστική $F_{x_i}(y)$ έχει σημαντική επίδραση στα άκρα της κατανομής των μεγίστων.

Οι πιο γνωστές ασυμπτωτικές κατανομές ακραίων τιμών είναι για τα μέγιστα η τύπου I (Gumbel) και η τύπου II, ενώ για τα ελάχιστα είναι η τύπου III (Weibull).

Ο τύπος I (Gumbel). Η κατανομή πιθανοτήτων είναι η εξής:

$$p(x) = a e^{-a(x-u) - e^{-a(x-u)}}, \quad -\infty < x < \infty \quad (3.34)$$

και η αθροιστική:

$$F_x(x') = P(x \leq x') = e^{-e^{-a(x'-u)}} \quad (3.35)$$

Οι παράμετροι της υπολογίζονται ως εξής:

$$u = \mu - \gamma/\alpha$$

(γ η σταθερά Euler = 0.57721). Το u είναι ο πιθανότερος μέσος όρος (mode) της μεταβλητής x , ενώ:

$$\alpha = \pi/\sigma \sqrt{6}$$

όπου μ και σ ο μέσος όρος και τυπική απόκλιση του μέγιστου x . Η ασυμμετρία της κατανομής είναι $\gamma_s = 1.139$ δηλαδή έχει σταθερή τιμή. Περισσότερες λεπτομέρειες για την κατανομή και τη χάραξη της θα δοθούν στο σχετικό Κεφ. 7.

Ο τύπος II. Έχει την έκφραση:

$$P(X \leq x') = e^{-(\theta/x')^k}, \quad 0 \leq x' < \infty \quad (3.36)$$

όπου θ είναι η αναμενόμενη μέγιστη τιμή για ένα δείγμα με n τιμές και k είναι τάξη της ροπής ανεξάρτητη από το n .

Ο τύπος III (Weibull). Έχει την εξής αθροιστική συνάρτηση:

$$P(x \leq x') = e^{-\left[\frac{x'-\varepsilon}{\theta-\varepsilon}\right]^k}, \quad -\infty < x' \leq \varepsilon \quad (3.37)$$

όπου θ είναι χαρακτηριστική ελάχιστη τιμή με πιθανότητα εμφάνισης: $P_\theta = e^{-1}$ και περίοδο επαναφοράς $T_\theta = 1.58198$ και είναι η μικρότερη τιμή των ελαχίστων τιμών ενώ k είναι η τάξη της μικρότερης παραγώγου της κατανομής που δεν εξαφανίζεται στο σημείο $x'=\varepsilon$. Η κατανομή αυτή εφαρμόζεται πολύ συχνά στην ανάλυση συχνότητας ξηρασιών (Berrettoni, 1962) (χάραξη καμπυλών διάρκειας-δριμύτητας ξηρασίας κ.λ.π.) που είναι αναγκαία στον υδρολογικό σχεδιασμό έργων διατήρησης της ποιότητας του νερού. Λεπτομέρειες γι' αυτή θα δοθούν στο σχετικό με την ανάλυση ακραίων τιμών Κεφ. 7 αυτού του μέρους του βιβλίου και για το σχεδιασμό τέτοιων έργων στο Κεφ. 3 του δεύτερου μέρους του βιβλίου.

3.4. Μέθοδοι εκτίμησης παραμέτρων κατανομών

Κατά την εκτίμηση των παραμέτρων των κατανομών πρέπει να επιτυγχάνονται οι ακόλουθες ιδιότητες, (Marković, 1965) όσον αφορά τον εκτιμητή $\hat{\theta}$ των παραμέτρων (δηλαδή το στατιστικό εκείνο χαρακτηριστικό που υπολογίζεται από το δείγμα της μεταβλητής και με βάση το οποίο γίνεται η εκτίμηση των παραμέτρων).

- Η πιθανότητα να είναι η απόλυτη τιμή της διαφοράς του εκτιμητή $\hat{\theta}$ από την πραγματική τιμή θ του πληθυσμού της μεταβλητής μικρότερη από κάθε μικρή ποσότητα ε , πρέπει να πλησιάζει τη μονάδα όσο το μέγεθος του δείγματος της μεταβλητής πλησιάζει το άπειρο: $p(|\hat{\theta} - \theta| < \varepsilon) \rightarrow 1, n \rightarrow \infty$.
- Η αναμενόμενη τιμή του εκτιμητή $\hat{\theta}$ πρέπει να είναι ίση με την τιμή θ από τον πληθυσμό: $E(\hat{\theta}) = \theta$, ενώ σαν απόκλιση ορίζεται η ποσότητα: $E(\hat{\theta}) - \theta$.
- Ανάμεσα στις πιθανές τιμές του εκτιμητή $\hat{\theta}$, εκείνη με την ελάχιστη απόκλιση θ , πρέπει να έχει την μικρότερη σκέδαση όλων.

Δύο είναι οι πιο γνωστές μέθοδοι υπολογισμού των παραμέτρων: η μέθοδος των ροπών και η μέθοδος της μέγιστης πιθανότητας (Cramer, 1946· Brunk, 1960· Beard, 1962· Lindgren και McElrath, 1966). Ωστόσο και άλλες μέθοδοι όπως η μέθοδος της μέγιστης εντροπίας ή η μέθοδος των L-ροπών (Stedinger και άλλοι, 1993) έχουν συχνά χρησιμοποιηθεί στην Τεχνική Υδρολογία.

Η μέθοδος των ροπών. Σύμφωνα μ' αυτή οι παράμετροι της κατανομής εκτιμώνται από τις στατιστικές ροπές του δείγματος της μεταβλητής. Έτσι για τις δι-παραμετρικές κατανομές ο μέσος όρος και η σκέδαση του δείγματος είναι οι δύο εκτιμητές για τον υπολογισμό των παραμέτρων.

Οι στατιστικοί αυτοί εκτιμητές από το δείγμα δεν έχουν μια μονοσήμαντη τιμή, αλλά αν κανείς σκεφθεί ότι μπορεί να πάρει πολλά δείγματα (π.χ. m τον αριθμό δείγματα $x_1, x_2, \dots, x_m$) από τον πληθυσμό μιας μεταβλητής, θα δει ότι αποτελούν οι ίδιοι οι εκτιμητές τυχαίες μεταβλητές που έχουν μια συγκεκριμένη κατανομή πιθανοτήτων, μέσο όρο, σκέδαση κ.λ.π. Ένα μέτρο του μέσου μεγέθους του τυχαίου σφάλματος που συνοδεύει τον εκτιμητή του μέσου όρου μπορεί να θεωρηθεί ότι είναι η αναμενόμενη τιμή του τετραγώνου του σφάλματος εκτίμησης ($\bar{x} - m_x$) όπου, $\bar{x}$ είναι ο τυχαίος μέσος όρος του δείγματος και m_x η μέση τιμή του (μέσος όρος των μέσων όρων $\bar{x}$ για όλες τις πιθανές δειγματοληψίες). Ο μέσος όρος m_x του μεταβλητού μέσου όρου $\bar{x}$ του δείγματος βρίσκεται ως εξής:

$$E[\bar{x}] = E\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n E[x_i] = \frac{1}{n} \sum_{i=1}^n m_x = \frac{nm_x}{n} = m_x = \mu \quad (3.38)$$

Με άλλα λόγια, η αναμενόμενη τιμή του δειγματικού μέσου όρου $\bar{x}$ είναι ίση με το μέσο όρο της μεταβλητής x .

Αποδεικνύεται ακόμα (Benjamin και Cornell, 1970) ότι η σκέδαση $\sigma_{\bar{x}}^2$ των μέσων όρων από τις πιθανές δειγματικές σειρές ενός πληθυσμού ισούται με την τυχαία σκέδαση σ_x^2 , του δείγματος (και άρα του πληθυσμού σ^2) διαιρεμένη δια του πλήθους των τιμών του δείγματος n :

$$\sigma_{\bar{x}}^2 = \frac{1}{n} \sigma_x^2 = \frac{\sigma^2}{n} \quad (3.39)$$

Οι στατιστικές τιμές του όρου m_x και της τυπικής απόκλισης $\sigma_{\bar{x}}$ είναι ανεξάρτητες από την κατανομή πιθανοτήτων του μέσου όρου $\bar{x}$ και είναι αρκετές για να απαντήσουν πολλές ερωτήσεις σχετικά με τον εκτιμητή μέσο όρο $\bar{x}$. Πάντως μπορεί να λεχθεί ότι ο εκτιμητής $\bar{x}$ θα κυμαίνεται μέσα σε συγκεκριμένα όρια, όπως: $m_x \pm k \sigma_{\bar{x}}$, όπου k θετικός αριθμός. Τα όρια αυτά μικραίνουν όσο το μέγεθος του δείγματος μεγαλώνει.

Η σκέδαση τώρα του δείγματος με μέσο όρο $\bar{x}$ ισούται με:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3.40)$$

Παίρνοντας αναμενόμενες τιμές και στα δύο μέλη της εξ.(3.40) και αναπτύσσοντας την προκύπτει:

$$E[S^2] = \frac{n-1}{n} \sigma_x^2 \quad (3.41)$$

Άρα η αναμενόμενη τιμή της δεύτερης κεντρικής ροπής του δείγματος $E[S^2]$ δεν είναι ίση με τη δεύτερη κεντρική ροπή της κατανομής πιθανότητας σ^2 αφού, για πολλά δείγματα, ο μέσος όρος της S^2 δεν εξισώνεται με την τιμή (τυχαία) της σ^2 της οποίας υποτίθεται ότι είναι εκτιμητής, αλλά είναι κατά κάτι μικρότερος. Εκτιμητές όπως το S^2 , των οποίων οι αναμενόμενες τιμές δεν είναι ίσες με την παράμετρο που εκτιμούν λέγονται εκτιμητές με απόκλιση (biased estimators). Η διόρθωση που γίνεται για εκτιμητή S^{2*} χωρίς απόκλιση

είναι η εξής:

$$S^{2*} = \frac{n}{n-1} S^2 \quad (3.42)$$

Αυτός είναι ο λόγος της διόρθωσης που γίνεται στις δεύτερες (και κατ' αναλογία στις τρίτες κ.λ.π.) ροπές για μικρά κυρίως δείγματα (ιδέ παρ. 3.2).

Η μέθοδος της μέγιστης πιθανότητας. Η συνάρτηση L (likelihood) χρησιμοποιείται για την εκτίμηση παραμέτρων. Ο κανόνας για να αποκτηθεί είναι:

Ο εκτιμητής μέγιστης πιθανότητας του θ είναι η τιμή $\hat{\theta}$ που κάνει την συνάρτηση πιθανότητας L μέγιστη. Η συνάρτηση L εξαρτάται από την παράμετρο θ και τη μεταβλητή x_i :

$$L = f(x_i, \theta) \quad (3.43)$$

Για τη συνθήκη μεγιστοποίησης τίθεται η πρώτη παράγωγος της L ως προς θ ίση με μηδέν και η εξίσωση που προκύπτει λύνεται ως προς θ :

$$\frac{\partial(\ln L)}{\partial \theta} = \frac{\partial \left[\sum_{i=1}^n \ln f(x_i, \theta) \right]}{\partial \theta} = 0 \quad (3.44)$$

Από την εξ.(3.44) για m παραμέτρους θ προκύπτουν m εξισώσεις για να λυθούν και να δώσουν m εκτιμητές παραμέτρων $\hat{\theta}$. Οι ιδιότητες του εκτιμητή $\hat{\theta}$ έχουν μελετηθεί πολύ. Έχει βρεθεί ότι έχουν περίπου κανονική κατανομή (για μεγάλα δείγματα). Οι μέσοι όροι τους είναι (ασυμπτωτικά) ίσοι με τις τιμές της αληθινής παραμέτρου θ , δηλαδή αυτοί οι εκτιμητές δεν έχουν απόκλιση (bias). Επίσης έχει δειχθεί (Freeman, 1963) ότι έχουν (ασυμπτωτικά τουλάχιστον) όλες τις ιδιότητες που αναφέρθηκαν στην αρχή. Έτσι, για μεγάλα δείγματα θα μπορούσε κανείς να πει ότι η μέθοδος αυτή είναι πιο ενδεδειγμένη. Στον Πίνακα 3.1 δίνονται οι κατανομές: Κανονική, Τριπαραμετρική Λογαριθμική - Κανονική, Τριπαραμετρική Γάμμα και Pearson IV, και ο τρόπος εκτίμησης των παραμέτρων τους με τη μέθοδο μέγιστης πιθανότητας.

Πίνακας 3.1: Κατανομές πιθανοτήτων και εκτίμηση παραμέτρων τους

Όνομα	Συνάρτηση
Διπαραμετρική Κανονική	$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
Τριπαραμετρική Λογαριθμική Κανονική	$p(x) = \frac{1}{(x-x_0)\sigma\sqrt{2\pi}} e^{-\frac{[\ln(x-x_0)-\ln\mu]^2}{2\sigma^2}}$
Τριπαραμετρική Γάμμα	$p(x) = \frac{1}{b\Gamma(a)} \left(\frac{x-x_0}{b}\right)^{a-1} e^{-\frac{x-x_0}{b}}$
Τριπαραμετρική Pearson IV	$p(x) = P_0 \left[1 + \frac{x}{a} - \frac{v^2}{r}\right]^{-m} e^{\{ \tan^{-1}(\frac{x}{a} - \frac{v}{r}) \}}$

Πίνακας 3.1: (συνέχεια)

Παράμετροι	Παρατηρήσεις	
$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$ $\hat{\sigma} = \left[\frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 \right]^{1/2}$	Χωρίς όρια, συμμετρική κωδωνοειδής	
Υπολογισμός $\hat{x}_0$ από: $\sum_{i=1}^n \frac{1}{x_i - \hat{x}_0} \left\{ \frac{1}{n} \sum_{i=1}^n \ln^2(x_i - \hat{x}_0) - \left[\frac{1}{n} \sum_{i=1}^n \ln(x_i - \hat{x}_0) \right]^2 - \frac{1}{n} \sum_{i=1}^n \ln(x_i - \hat{x}_0) \right\} + \sum_{i=1}^n \frac{\ln(x_i - \hat{x}_0)}{x_i - \hat{x}_0} = 0$ με επαναληπτικές μεθόδους $\ln \hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln(x_i - \hat{x}_0)$ $\hat{\sigma} = \frac{1}{n} \left\{ \sum_{i=1}^n [\ln(x_i - \hat{x}_0) - \ln \hat{\mu}]^2 \right\}^{1/2}$	Χωρίς όρια, σε μια κατεύθυνση, κωδωνοειδής	
Υπολογισμός $\hat{x}_0$ από: $\frac{1 + (1 + 3/4A)^{1/2}}{1 + (1 + 3/4A) - 4A} - (\bar{x} - \hat{x}_0) - \frac{1}{n} \sum_{i=1}^n \frac{1}{x_i - \hat{x}_0} = 0$ $A = \ln(\bar{x} - \hat{x}_0) - \frac{1}{n} \sum_{i=1}^n \ln(x_i - \hat{x}_0)$ με επαναληπτικές μεθόδους $\hat{\alpha} = \frac{1 + (1 + 3/4A)^{1/2}}{4A} - 0.04475(0.26)^{\hat{\alpha}}$ $\hat{b} = \frac{1}{\hat{\alpha}} (\bar{x} - \hat{x}_0)$	Χωρίς όρια, σε μια κατεύθυνση, συνήθως κωδωνοειδής ή μορφής J	
$r = \frac{6(\beta_2 - \beta_1 - 1)}{2\beta_2 - 3\beta_1 - 6}$ $m = \frac{1}{2} (r + 2)$ $v = \frac{r(r-2)\sqrt{\beta_1}}{\sqrt{16(r-1) - \beta_1(r-2)^2}}$ $P_0 = \left[\alpha \int_{-1/2\pi}^{+1/2\pi} \cos^r \theta e^{-v\theta} d\theta \right]^{-1}$ $\alpha = \frac{1}{4} \sqrt{\mu_2 [16(r-1) - \beta_1(r-2)^2]}$	$\beta_1 = \mu_3^2 / \mu_2^2$, $\beta_2 = \mu_4 / \mu_2^2$, μ_2, μ_3, μ_4 η δεύτερη, τρίτη και τέταρτη ροπή του x αντίστοιχα	Χωρίς όρια, μη συμμετρική, κωδωνοειδής

3.5. Διαστήματα εμπιστοσύνης

Η ατελής γνώση των παραμέτρων των κατανομών και η εκτίμηση τους από δειγματικές παραμέτρους που είναι οι ίδιες μεταβλητές, απαιτούν την ύπαρξη κάποιου ποσοτικού μέτρου της αβεβαιότητας αυτής που συνδέεται με κάθε εκτίμηση. Το μέτρο αυτό δίνεται από τα διαστήματα εμπιστοσύνης που περιβάλλουν και από τις δύο κατευθύνσεις (θετική και αρνητική) την κάθε εκτίμηση παραμέτρων ή κατανομών. Ο αριθμός που ποσοτικοποιεί το άνοιγμα αυτών των διαστημάτων λέγεται συντελεστής ή επίπεδο εμπιστοσύνης.

Διάστημα εμπιστοσύνης μέσου όρου. Το διάστημα εμπιστοσύνης για το μέσο όρο μ του πληθυσμού μιας κανονικής μεταβλητής με γνωστή τυπική απόκλιση σ προκύπτει από τη γνώση της κατανομής των μέσων όρων $\bar{x}$ των (m τον αριθμό) πιθανών δειγμάτων ή τιμών από τον πληθυσμό. Αυτή η κατανομή έχει βρεθεί ότι είναι κανονική με μέση τιμή ίση με μ και τυπική απόκλιση ίση με $\sigma / \sqrt{n}$. Έτσι, η πιθανότητα P να είναι το $\bar{x}$ στα όρια:

$$\mu - k \frac{\sigma}{\sqrt{n}} < \bar{x} < \mu + k \frac{\sigma}{\sqrt{n}} \quad (3.45)$$

όπου k θετικός αριθμός, προκύπτει από την εμβαδομέτρηση της κανονικής κατανομής πιθανότητας του $\bar{x}$ ως εξής:

$$\int_0^k f(t) dt = \int_0^k \frac{1}{\sqrt{2\pi}} e^{-t^2/2} = \frac{p}{2} \quad (3.46)$$

όπου t η τυποποιημένη κανονική μεταβλητή που προκύπτει από τη $\bar{x}$. Συνεπώς, το διάστημα εμπιστοσύνης με πιθανότητα p για το μέσο όρο μ είναι το εξής:

$$\bar{x} - k \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + k \frac{\sigma}{\sqrt{n}} \quad (3.47)$$

Ας υποθεθεί τώρα ότι και πάλι η μεταβλητή είναι κανονική αλλά η τυπική απόκλιση σ είναι άγνωστη. Ένα δείγμα από τον πληθυσμό έχει μέσο όρο $\bar{x}$ και τυπική απόκλιση S . Η ποσότητα: $t = \sqrt{n-1}(\bar{x} - \mu) / S$ έχει κατανομή Student t : $f_v(t)$, με $v = n - 1$ βαθμούς ελευθερίας (η κατανομή t περιγράφεται σε όλα τα στατιστικά εγχειρίδια) οπότε βρίσκεται ένα t_e τέτοιο ώστε:

$$2 \int_0^{t_e} f_v(t) dt = p = 1 - \varepsilon \quad (3.48)$$

Τότε σε ποσοστό p των δειγμάτων ισχύει:

$$-t_e \leq \sqrt{n-1}(\bar{x} - \mu) / S \leq t_e$$

$$\bar{x} - \frac{t_e S}{\sqrt{n-1}} \leq \mu \leq \bar{x} + \frac{t_e S}{\sqrt{n-1}} \quad (3.49)$$

που είναι το διάστημα εμπιστοσύνης για τον μέσο όρο μ . Για δοσμένη πιθανότητα ε προκύπτει μια τιμή του t_e από πίνακες της κατανομής t (Benjamin και Cornell, 1970). Η αθροιστική κατανομή t δίνεται στον Πίνακα Π-3 του παραρτήματος.

Διάστημα Εμπιστοσύνης σκέδασης. Διαστήματα εμπιστοσύνης πάνω στη σκέδαση σ^2 μιας κανονικά κατανομημένης μεταβλητής μπορούν να υπολογισθούν όταν η κατανομή της δειγματικής εκτίμησης S^2 είναι γνωστή. Είναι γνωστό ότι για δείγμα x_i μήκους n και μέσου όρου $\bar{x}$ ισχύει η εξ. (3.40) για τη σκέδαση S^2 . Προσθέτοντας και αφαιρώντας το m_x (μέσος όρος των $\bar{x}$) και αναπτύσσοντας τα τετράγωνα στην εξίσωση αυτή, προκύπτει:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m_x)^2 - (\bar{x} - m_x)^2 \quad (3.50)$$

Πολλαπλασιάζοντας επί n/σ^2 προκύπτει:

$$\frac{nS^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{x_i - m_x}{\sigma} \right)^2 - \left(\frac{\bar{x} - m_x}{\sigma\sqrt{1/n}} \right)^2 \quad (3.51)$$

Ο πρώτος όρος δεξιά είναι το άθροισμα τετραγώνων n ανεξάρτητων κανονικών τυχαίων μεταβλητών, κάθε μια με μηδέν μέσο όρο και μοναδιαία τυπική απόκλιση. Αυτό το άθροισμα έχει τη γνωστή κατανομή χ^2 με n βαθμούς ελευθερίας (Benjamin και Cornell, 1970) (η κατανομή χ^2 δίνεται σε όλα τα στατιστικά εγχειρίδια και εδώ στον Πίνακα Π-1 ως: $1 - F(\chi^2)$ με ν βαθμούς ελευθερίας). Ο δεύτερος όρος είναι το άθροισμα τετραγώνων τυπικών κανονικών μεταβλητών (αφού το $\bar{x}$ είναι: $N(m_x, \sigma\sqrt{1/n})$ κανονική

μεταβλητή. Άρα και αυτή έχει μια χ^2 κατανομή με 1 βαθμό ελευθερίας. Η εξ.(3.51) ξαναγράφεται ως εξής:

$$\sum_{i=1}^n \left(\frac{x_i - m_x}{\sigma} \right)^2 = \frac{nS^2}{\sigma^2} + \left(\frac{\bar{x} - m_x}{\sigma\sqrt{1/n}} \right)^2 \quad (3.52)$$

Με βάση αυτή την εξίσωση μια χ^2 κατανομημένη μεταβλητή με η βαθμούς ελευθερίας (αριστερό μέλος) είναι το άθροισμα της μεταβλητής nS^2/σ^2 και μιας χ^2 μεταβλητής με 1 βαθμό ελευθερίας. Άρα είναι μάλλον αναμενόμενο ότι και η nS^2/σ^2 είναι χ^2 -κατανομημένη με $(n-1)$ βαθμούς ελευθερίας (έτσι ώστε να εξισώνονται οι βαθμοί ελευθερίας στα δύο μέλη της εξ.(3.52)). Άρα η πιθανότητα P υπέρβασης κάποιας τιμής $\chi^2_{1-\alpha, n-1}$ στην πιθανότητα α είναι:

$$P \left[\frac{nS^2}{\sigma^2} \geq \chi^2_{1-\alpha, n-1} \right] = 1 - \alpha \quad (3.53)$$

Ανακατατάσσοντας την εξ.(3.53) προκύπτει το όριο (προς μια κατεύθυνση) για το σ^2 στο $(1-\alpha)100\%$ επίπεδο εμπιστοσύνης ως εξής:

$$P \left[\sigma^2 \leq \frac{nS^2}{\chi^2_{1-\alpha, n-1}} \right] = 1 - \alpha \quad (3.54)$$

Όπου για κάθε α αντιστοιχεί μια τιμή $\chi^2_{1-\alpha, n-1}$ που βρίσκεται από τον Πίνακα Π-1.

Διάστημα εμπιστοσύνης κατανομής. Πολλές φορές είναι επιθυμητό να αποκτηθούν τα διαστήματα εμπιστοσύνης στην κατανομή που εμπεριέχει τις παραμέτρους για τις οποίες έχουν βρεθεί τα διαστήματα εμπιστοσύνης. Αυτά βρίσκονται εύκολα με βάση τα διαστήματα εμπιστοσύνης των παραμέτρων τους (Benjamin και Cornell, 1970, σελ. 391). Έτσι π.χ., εάν για την κανονική κατανομή μιας μεταβλητής x , για το (μεταβλητό) δειγματικό μέσο όρο της $\bar{x}$ ισχύει η εξ.(3.45) για διάστημα εμπιστοσύνης σε επίπεδο p , τότε είναι προφανές ότι για την τυπική κανονική κατανομή $N(0,1)$ της: $y = (x - \mu)/\sigma$ δηλαδή την $f(y) = 1/\sqrt{2\pi} e^{-y^2/2}$, $-\infty < y < +\infty$, θα πρέπει με την ίδια πιθανότητα p να ισχύουν τα διαστήματα:

$$-k < y < +k$$

και άρα

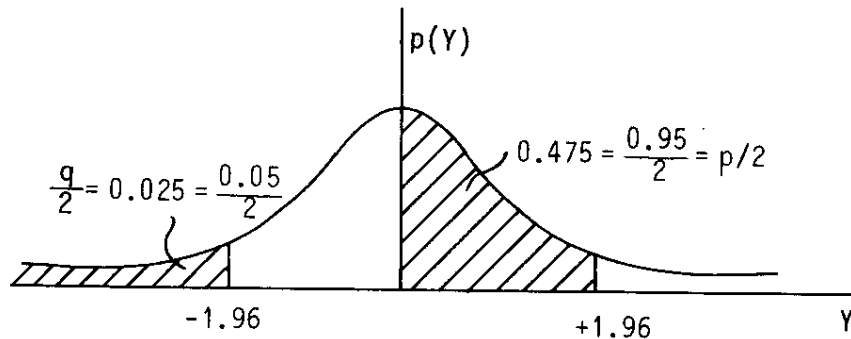
(3.55)

$$\mu - k\sigma < x < \mu + k\sigma$$

Μερικές χρήσιμες τιμές για τα διαστήματα εμπιστοσύνης της κανονικής κατανομής είναι:

k	$P(-k < y < +k)$	α (συμπληρωματική πιθανότητα)
1.00	0.683	0.317
1.96	0.950	0.050
2.00	0.954	0.046
2.58	0.990	0.010
3.00	0.997	0.003

σχηματικά δε δίνονται στο Σχήμα 3.8.



Σχήμα 3.8. Επίπεδα εμπιστοσύνης κανονικής κατανομής.

3.6. Έλεγχος καταλληλότητας κατανομών

Μετά την προσαρμογή στο εμπειρικό ιστόγραμμα κάποιου δείγματος μεταβλητής, μιας ή περισσότερων θεωρητικών κατανομών πιθανοτήτων, πρέπει να γίνει έλεγχος της καταλληλότητας της ή των κατανομών. Στην περίπτωση μάλιστα περισσότερων της μιας, πρέπει να δοθεί και απάντηση στο ερώτημα ποια κατανομή προσαρμόζει καλύτερα στα ιστορικά δεδομένα. Ο έλεγχος της καταλληλότητας μιας κατανομής βασίζεται στη θεωρία των υποθέσεων (Freund και άλλοι, 1988· Papoulis, 1990· Hirsch και άλλοι, 1993). Πιο συγκεκριμένα, γίνεται έλεγχος της υπόθεσης ότι η προσαρμοζόμενη στο δείγμα κατανομή είναι κατάλληλη για να το περιγράψει, μέσω μιας

διαδικασίας (στατιστικό τεστ) που οδηγεί είτε στην αποδοχή είτε στην απόρριψη αυτής της υπόθεσης. Από τα πιο γνωστά στατιστικά τεστ είναι το τεστ χ^2 και το τεστ Kolmogorov - Smirnov.

3.6.1. χ^2 -τεστ

Βασίζεται στον έλεγχο των αποκλίσεων μεταξύ του ιστογράμματος και των τιμών που προβλέπονται από τη θεωρητική κατανομή (Mann και Wald, 1942). Έστω μια διακεκριμένη μεταβλητή x , της οποίας γίνεται μια τυχαία δειγματοληψία n φορές. Η κατανομή του αριθμού N_i των παρατηρήσεων (της κάθε τιμής x_i του δείγματος) είναι διωνυμική με μέσο όρο και σκέδαση σύμφωνα με τις ιδιότητες της διωνυμικής κατανομής:

$$\left\{ \begin{array}{l} \mu_{N_i} = np_i, p_i = P_x(x_i) \\ \sigma_{N_i}^2 = np_i(1 - p_i) \end{array} \right\}$$

Η τυποποιημένη τυχαία απόκλιση D από το μέσο όρο της μεταβλητής N_i είναι κατανεμημένη με την τυπική κανονική κατανομή $N(0,1)$:

$$D = \frac{N_i - np_i}{\sqrt{np_i(1 - p_i)}} \quad (3.56)$$

Βρίσκεται δε (Benjamin και Cornell, 1970, παρ. 4.4.2) ότι το άθροισμα των τετραγώνων D_k :

$$D_k = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i} \quad (3.57)$$

είναι κατανεμημένο με την κατανομή χ^2 με $k - 1 - r$ βαθμούς ελευθερίας όπου r ο αριθμός των παραμέτρων της κατανομής.

Από τους σχετικούς πίνακες της κατανομής χ^2 (Πίνακας Π-1) μπορεί κανείς εύκολα να βρει την πιθανότητα: $P[D_k \geq \chi_{\alpha, k-1-r}^2] = \alpha$. Με βάση αυτά, η διαδικασία για το τεστ χ^2 είναι η εξής:

- Εκτιμούνται οι r παράμετροι της προσαρμοζόμενης κατανομής (π.χ. με τη μέθοδο της μέγιστης πιθανότητας).
- Χωρίζεται το δείγμα των στοιχείων σε k κλάσεις και προσδιορίζεται η πιθανότητα p_i μιας τυχαίας τιμής από τη θεωρητική κατανομή να πέσει μέσα σε κάθε κλάση (τουλάχιστον μια παρατήρηση σε κάθε κλάση).

- Πολλαπλασιάζεται η πιθανότητα p_i με το μέγεθος του δείγματος n . Προκύπτει έτσι ο αναμενόμενος (θεωρητικός) αριθμός E_i παρατηρήσεων για κάθε κλάση με τη συγκεκριμένη κατανομή: $E_i = np_i$.
- Μετρούνται οι πραγματικές παρατηρήσεις N_i από το δείγμα που πέφτουν μέσα σε κάθε κλάση.
- Υπολογίζεται η στατιστική παράμετρος $\chi^2 = \sum_{i=1}^k (N_i - E_i)^2 / E_i$.
- Συγκρίνεται αυτή η τιμή χ^2 με την τιμή που προκύπτει από τους πίνακες χ^2 (Πίνακας Π- 1) για $k - r - 1$ βαθμούς ελευθερίας και για συγκεκριμένη πιθανότητα. Υψηλές τιμές χ^2 δείχνουν ότι τα στοιχεία αντιτίθενται στο ομοίωμα.
- Η υπόθεση για καταλληλότητα της κατανομής απορρίπτεται αν η υπολογισμένη τιμή χ^2 πέφτει μέσα στην κρίσιμη ζώνη που ορίζεται από το επίπεδο εμπιστοσύνης 10% ή 5%.

Παράδειγμα χ^2 -τεστ

Έστω τα στοιχεία βροχής σε αριθμό σταθμών που δίνονται στον Πίνακα 3.2. Με την υπόθεση ότι οι καταιγίδες συμβαίνουν σ' αυτούς τους ανεξάρτητους σταθμούς με μέσο ρυθμό επισύμβασης μιας το χρόνο και ότι κατανέμονται κατά Poisson (Benjamin και Cornell, 1970), τότε ο αριθμός των αφίξεων βροχής X για κάθε χρόνο σε κάθε σταθμό είναι κατανεμημένος κατά Poisson με μέσο όρο ίσο με 1 ως εξής:

$$P_X(x) = \frac{(1)^x e^{-1}}{x!} = \frac{0.368}{x!}, \quad x = 0, 1, 2, \dots$$

Οι αναμενόμενες τιμές παρατηρήσεων (από την κατανομή Poisson) κάθε τιμής x σε ένα δείγμα με μέγεθος $n = 360$ φαίνονται στη στήλη (3) του Πίνακα 3.2. Αφαιρώντας κάθε αναμενόμενη τιμή από την παρατηρημένη, υψώνοντας στο τετράγωνο και διαιρώντας με τον αναμενόμενο αριθμό, αποκτώνται οι τυποποιημένες τετραγωνικές τυχαίες αποκλίσεις στη στήλη (4) του πίνακα. Οι δύο τελευταίες κατηγορίες έχουν υπολογισθεί μαζί, για να προσαρμοσθεί καλύτερα η κατανομή των στατιστικών τεστ στην κατανομή χ^2 (ο μικρότερος αριθμός παρατηρήσεων σε κάθε κλάση πρέπει να είναι μεταξύ $5 \div 10$, πάντως έχει βρεθεί ότι και μικρότεροι αριθμοί γύρω στο 2, δίνουν καλά αποτελέσματα). Η παράμετρος $D_k=19.04$ από τον πίνακα είναι κατανεμημένη κατά χ^2 με 5 βαθμούς ελευθερίας ($k - r - 1 = 7 - 1 - 1 = 5$). Η πιθανότητα της τιμής $\chi^2 = 19.04$ (από το σχετικό Πίνακα Π-1) είναι μικρότερη από 0.05. Πιο συγκεκριμένα για επίπεδο 5% ισχύει:

Πίνακας 3.2: χ^2 -τεστ

Αριθμός καταγι- δων ανά σταθμό το χρόνο (1)	Παρατη- ρημένος αριθμός επισυμ- βάσεων (2)	Παράμετρος $\lambda = 1$ (καθορισμένη)		Παράμετρος $\lambda = \bar{x} = 1.18$ (εκτιμημένη)	
		Αναμενόμενος αριθμός επισυμβάσεων (3)	Τυποποιημένες τετραγω- νικές αποκλίσεις (4)	Αναμενόμενος αριθμός επισυμβάσεων (5)	Τυποποιημένες τετραγωνικές αποκλίσεις (6)
0	105	$360(0.368) = 132.48$	$\frac{(105 - 132.48)^2}{132.48} = 5.70$	110.52	0.28
1	140	$360(0.368) = 132.48$	$\frac{(140 - 132.48)^2}{132.48} = 0.43$	130.32	0.72
2	70	$360(0.184) = 66.24$	$\frac{(70 - 66.24)^2}{66.24} = 0.21$	76.68	0.58
3	30	$360(0.061) = 21.96$	$\frac{(30 - 21.96)^2}{21.96} = 2.94$	30.24	0.002
4	12	$360(0.015) = 5.40$	$\frac{(12 - 5.40)^2}{5.40} = 8.07$	9.00	1.00
5	3	$360(0.003) = 1.08$	$\frac{(3 - 1.44)^2}{1.44} = 1.69$	2.16	0.02
6	0	$360(0.001) = 0.36$	$\frac{1.44}{0.36} = 1.44$	1.08	0.48
$\bar{x} = 1.18$	360	360.00	$\frac{D_k = 19.04}{\chi^2_{0.05,5} = 11.1}$	360.00	2.60

$$\chi^2_{0.05,5} = 11.1$$

Άρα $D_k = 19.04 \geq \chi^2_{0.05,5} = 11.1$, υπόθεση καταλληλότητας της Poisson με $\lambda=1$ θα πρέπει να απορριφθεί στο επίπεδο 5%. Αν τώρα η κατανομή διορθωθεί με $\lambda = \bar{x} = 1.18$ που είναι ο παρατηρημένος μέσος όρος αφίξεων καταιγίδων στην περιοχή των σταθμών τότε όπως φαίνεται από τις στήλες (5) και (6) του Πίνακα 3.2 η υπόθεση περνάει στο 5% αφού: $D_k = 2.60 < 9.49 = \chi^2_{0.05,4}$ (εδώ παίρνονται 4 βαθμοί ελευθερίας με την σύμπτυξη των κλάσεων σε $k = 6$).

Στατιστική Ανάλυση Ακραίων Υδρολογικών Τιμών

Οι ακραίες τιμές υδρολογικών μεταβλητών, δηλαδή οι μέγιστες και ελάχιστες παρατηρημένες τιμές μιας μεταβλητής καθ' όλη τη διάρκεια της δειγματοληψίας της, παίζουν πολύ σημαντικό ρόλο στο σχεδιασμό των υδατικών συστημάτων. Οι μέθοδοι ανάλυσης αυτών των ακραίων τιμών διακρίνονται σε προσδιοριστικές, στατιστικές - πιθανολογικές και στοχαστικές. Το βασικό πρόβλημα στο σχεδιασμό ενός έργου που σχετίζεται με τις ακραίες τιμές είναι η διακινδύνευση, δηλαδή η πιθανότητα εμφάνισης ή υπέρβασης της κρίσιμης για το σχεδιασμό ακραίας τιμής τουλάχιστον μια φορά μέσα σε συγκεκριμένο χρονικό διάστημα. Γι' αυτό και οι στατιστικές, ιδιαίτερα οι πιθανολογικές μέθοδοι, είναι εκείνες που περισσότερο χρησιμοποιούνται για την ανάλυση των ακραίων τιμών. Στο κεφάλαιο αυτό θα παρουσιαστούν τα πιο γνωστά πιθανολογικά ομοιώματα ανάλυσης ακραίων τιμών, κυρίως παροχών (πλημμυρών και ξηρασιών), όπως και η κατασκευή ομβρίων καμπυλών που προκύπτει από την ανάλυση συχνότητας της βροχής. Επίσης, θα αναφερθούν και εντοπικά ομοιώματα ανάλυσης συχνότητας πλημμυρών.

7.1. Γενικά

Τα ακραία γεγονότα μέσα σε μια υδρολογική χρονοσειρά θεωρούνται ότι αποτελούν μια ιδιαίτερη κατηγορία μέσα σε όλες τις άλλες τιμές της μεταβλητής. Έτσι από μια σειρά ημερήσιων παροχών n -ετών μπορεί κανείς να ξεχωρίσει τις μέγιστες τιμές σε κάθε χρόνο της σειράς, και έτσι να αποκτήσει μια σειρά n μεγίστων ετησίων τιμών παροχών, η οποία αναλύεται ξεχωριστά από τις υπόλοιπες. Το ίδιο μπορεί να γίνει και για τις βροχές ενός σταθμού ή μιας λεκάνης (μέσες τιμές) φυσικά χωριστά για κάθε επιθυμητή διάρκεια βροχής (π.χ. 5,10,30 λεπτών ή 1,12,24 κ.λ.π. ωρών).

Εκτός από τις ετήσιες τιμές μεγίστων, και ιδιαίτερα όταν αυτές οι τιμές δεν είναι αρκετές για ανάλυση, μπορεί κανείς να ξεχωρίσει $m_i = 1, 2, \dots, m$ τον αριθμό μέγιστες τιμές παροχών για ένα χρόνο i . Οι τιμές αυτές είναι μεγαλύτερες ή ίσες από κάποια οριακή τιμή, η οποία καθορίζεται έτσι ώστε αφ' ενός να επιτευχθεί ένα επιθυμητό m_i και αφ' ετέρου το μέγεθος αυτό να είναι συμβατό με τις ανάγκες του πλημμυρικού σχεδιασμού. Η σειρά των μεγίστων που προκύπτει έτσι λέγεται σειρά μερικής διάρκειας.

Τα πιο γνωστά ομοιώματα πιθανολογικής ανάλυσης μεγίστων ετησίων τιμών είναι η ασυμπτωτική κατανομή ακραίων τιμών τύπου I – Gumbel, η λογαριθμική κανονική κατανομή και οι κατανομές Pearson III. Γενικός κανόνας-υπόδειξη που να καθορίζει ποια από τις γνωστές κατανομές μεγίστων τιμών είναι η καλύτερη δεν έχει μέχρι στιγμής βρεθεί. Έχουν γίνει πολλές προσπάθειες και έχουν γραφεί πολλές εργασίες πάνω στο θέμα αυτό προτείνοντας τη μια ή την άλλη μέθοδο ανάλογα με το μήκος της σειράς, την ασυμμετρία κ.λ.π. (Tennessee Valley Authority, 1961· Majumbar και Sawhney, 1965· Benson, 1968· Wallis, 1976· Kottegoda, 1980). Για παράδειγμα, από την Υδρολογική Επιτροπή του Συμβουλίου Υδατικών Πόρων των Η.Π.Α. εφαρμόστηκαν έξι κατανομές σε πλημμυρικές σειρές μήκους από 40 έως 97 χρόνια σε 10 επιλεγμένους σταθμούς και έγινε η σύγκριση μεταξύ τους (Benson, 1968). Οι κατανομές ήταν: η Gumbel (τύπος I), η λογαριθμική Gumbel (διπαραμετρική τύπου II), η διπαραμετρική γάμμα, η λογαριθμική Pearson III, η λογαριθμική κανονική και η λογαριθμική κανονική τροποποιημένη με τη μέθοδο Hazen. Η επιτροπή μετά από συγκριτική ανάλυση των στοιχείων πρότεινε τη χρήση της λογαριθμικής Pearson III (της οποίας η λογαριθμική κανονική είναι μια ειδική οριακή περίπτωση με μηδενική ασυμμετρία), με την επιφύλαξη, ότι όταν τα στοιχεία δείχνουν σημαντικές αποκλίσεις από αυτήν τότε θα πρέπει να αναζητηθεί η κατανομή με τη βέλτιστη προσαρμογή. Το Συμβούλιο Υδατικών Πόρων των ΗΠΑ δημοσίευσε το 1967 (WRC, 1967) την έκθεση αρ. 15 που συστήνει στις διάφορες υπηρεσίες να χρησιμοποιούν μια ομοιόμορφη τεχνική ανάλυσης πλημμυρών, που βασίζεται ακριβώς στη χρήση της λογαριθμικής Pearson πάνω στις ετήσιες σειρές μεγίστων παροχών. Στη συνέχεια με τις εκθέσεις 17 (WRC, 1976), 17A (WRC, 1977) και 17B (WRC, 1981) βελτίωσε την τεχνική υπολογισμού, με χρήση εντοπικής ανάλυσης για τον υπολογισμό της ασυμμετρίας, για την αντιμετώπιση των ιδιαίτερα υψηλών («ασύμφωνων») τιμών του δείγματος των μεγίστων κ.λ.π., αλλά πάντα προτείνοντας για χρήση την λογαριθμική Pearson III κατανομή.

Ωστόσο υπάρχουν και άλλες εργασίες, όπως αυτή του φυσικού περιβαλλοντικού Συμβουλίου Ερευνών (N.E.R.C., 1975) που με παρόμοιας με την προηγούμενη έκτασης εργασία προτείνει την τύπου II ασυμπτωτική συνάρτηση ακραίων τιμών εξ.(3.36). Ένα ακόμη σημαντικό πρόβλημα της ανάλυσης μεγίστων (και ελαχίστων), που επηρεάζει την επιλογή της κατάλληλης κατανομής και που όμως δεν έχει ακόμα αντιμετωπισθεί σε γενική και σαφή θεωρητική βάση, είναι η ύπαρξη μέσα στο δείγμα των ακραίων τιμών μιας ή περισσότερων «ασύμφωνων» τιμών, δηλαδή ιδιαίτερα υψηλών (ή χαμηλών) τιμών σε σχέση με τις άλλες. Οπότε προκύπτουν τα ερωτήματα αν και κατά

πόσο θα πρέπει να συμπεριληφθεί η τιμή αυτή μέσα στο δείγμα (δηλαδή αν ανήκει στον ίδιο πληθυσμό με τις άλλες ή όχι) και τι επιπτώσεις θα έχει η μια ή η άλλη ενέργεια. Συνήθως η λύση δίνεται στη βάση της προσωπικής εμπειρίας του μελετητή και των στοιχείων που διατίθενται για τη φυσική εξήγηση των αιτιών του «ασύμφωνου» γεγονότος. Πάντως, είναι γεγονός ότι ιδιαίτερα στις περιπτώσεις που αφορούν την αντιπλημμυρική προστασία των μεγάλων αναπτυξιακών έργων (φράγματα κ.λ.π.), όπου τα ζητούμενα μεγέθη είναι πολύ μικρής συχνότητας και χρειάζεται ιδιαίτερα μεγάλη επέκταση του συνήθως περιορισμένου δείγματος των ιστορικών στοιχείων που υπάρχουν, θα πρέπει να δίνεται ιδιαίτερη προσοχή στην επιλογή της κατάλληλης κατανομής (να γίνονται δηλαδή τεστ καταλληλότητας κ.λ.π.) δεδομένου ότι οι διαφορές των κατανομών στους καθοδικούς κλάδους, και άρα η διαφορά τους στις επεκτάσεις, είναι σημαντικές. Όταν χρησιμοποιούνται γραφικές μέθοδοι θα πρέπει να δίνεται σημασία στη μέθοδο εκτίμησης της σχεδιαστικής θέσης κάθε σημείου σε σχέση με την περίοδο επαναφοράς του (Stripp και Young 1971). Αν παρ' όλα αυτά εξακολουθούν να υπάρχουν αμφιβολίες, τότε συνιστάται η χρήση της μέγιστης πιθανής πλημμύρας που προκύπτει από τη χρήση της μέγιστης πιθανής καταιγίδας (δες παράγραφο 2.2.10) και του μοναδιαίου υδρογραφήματος στη θέση σχεδιασμού.

Ιδιαίτερα προβλήματα ανάλυσης συχνοτήτων ακραίων τιμών παρουσιάζονται όταν στη θέση που προγραμματίζεται ένα έργο δεν υπάρχουν ιστορικά στοιχεία μετρήσεων. Σε αυτές τις περιπτώσεις συνήθως χρησιμοποιούνται εντοπικές μέθοδοι που συνδέουν την αναμενόμενη τιμή της μεταβλητής συγκεκριμένης συχνότητας με διάφορα κλιματικά και γεωμορφολογικά χαρακτηριστικά της λεκάνης απορροής. Με τα ομοιώματα αυτά γίνεται η μεταφορά της πληροφορίας από τις θέσεις μετρήσεως σε άλλες χωρίς μετρήσεις με την προϋπόθεση φυσικά της υδρολογικής ομοιογένειας. Τέλος, στις περιπτώσεις εκείνες του σχεδιασμού που χρειάζεται όχι μόνο η πληροφορία του μεγέθους της πλημμύρας σχεδιασμού, αλλά και η ακριβής συνάρτηση της στο χρόνο, χρησιμοποιούνται στοχαστικά ομοιώματα πλημμυρών (Todorovic, 1978).

7.2. Ασυμπτωτική κατανομή ακραίων τιμών τύπου I-Gumbel

Η συνάρτηση πυκνότητας πιθανότητας της ασυμπτωτικής κατανομής ακραίων τιμών τύπου I - Gumbel δίνεται από την εξ.(3.34), ενώ η αθροιστική της κατανομή δίνεται από την εξ.(3.35). Ο Gumbel (1958α) ορίζει σαν τροποποιημένη μεταβλητή (reduced variable) την:

$$y = \alpha (x-u) \quad (7.1)$$

όπου α και u οι παράμετροι της κατανομής, όπως ορίστηκαν στην παράγραφο 3.3.8. Άρα η τροποποιημένη αθροιστική κατανομή γίνεται:

$$F_Y(y) = e^{-e^{-y}} \quad (7.2)$$

Ενώ η σχέση που συνδέει τις δύο μεταβλητές είναι:

$$x = \frac{1}{a} y + u \quad (7.3)$$

Αν τώρα Φ είναι η πιθανότητα υπέρβασης ενός γεγονότος ($\Phi = P(Y > y)$) και Φ' η συμπληρωματική της ($\Phi' = P(Y \leq y) = 1 - \Phi$), από την εξ.(7.2) προκύπτει ότι:

$$y = -\ln(-\ln \Phi') \quad (7.4)$$

7.2.1. Ορισμός περιόδου επαναφοράς και σχέση της με την τροποποιημένη μεταβλητή

Έστω ότι η τροποποιημένη μεταβλητή Y_i δηλώνει μέγιστες πλημμυρικές παροχές στα χρόνια $i = 1, 2, \dots, n$ αντίστοιχα (Kotteggoda, 1980). Αν οι Y_i είναι ανεξάρτητες μεταξύ τους, η πιθανότητα ο χρόνος T μεταξύ δύο υπερβάσεων του μεγέθους y να είναι ίσος με n δίνεται από την:

$$\begin{aligned} P(\bar{T} = n) &= P(Y_1 < y)P(Y_2 < y) \dots P(Y_{n-1} < y)P(Y_n > y) = \\ &= \{P(Y < y)\}^{n-1} P(Y > y) \end{aligned} \quad (7.5)$$

Η εξ.(7.5) δηλώνει μια γεωμετρική κατανομή όπου το n παίρνει τιμές από 1 έως ∞ . Άρα η αναμενόμενη τιμή $E(\bar{T})$ βρίσκεται από την ιδιότητα της γεωμετρικής κατανομής ως εξής:

$$E(\bar{T}) = 1 / P(Y > y) = 1 / \Phi \quad (7.6)$$

Η $E(\bar{T})$ γράφεται συνήθως πιο απλά ως περίοδος επαναφοράς T και ισούται με το αντίστροφο της πιθανότητας υπέρβασης Φ . Από την εξ.(7.4) προκύπτει ότι:

$$y = -\ln \{-\ln(1 - \Phi)\} = -\ln \{-\ln(1 - 1/T)\} = -\ln \{\ln(T) - \ln(T-1)\} \quad (7.7)$$

Κάνοντας την ανάπτυξη της εξ.(7.7) σύμφωνα με το θεώρημα MacLaurin προκύπτει η εξής σχέση:

$$y = -\ln \left\{ -(-1/T) + 1/2(-1/T)^2 - 1/3(-1/T)^3 + \dots \right\}$$

Αν χρησιμοποιηθούν μόνο τρεις όροι στην ανάπτυξη αυτή τότε:

$$y \cong -\ln\left(T^{-1} + \frac{1}{2}T^{-2} + \frac{1}{3}T^{-3}\right) = \ln\{6T^3/(6T^2 + 3T + 2)\}$$

Για $T > 10$ χρόνια μπορεί κανείς με λάθος 0.5% περίπου να θεωρήσει την προσέγγιση: $y = \ln(T - 1/2)$ και για ακόμα μεγαλύτερο T ($T > 25$ χρόνια) για πρακτικούς σκοπούς την: $y = \ln T$. Η εξ.(7.7) δηλώνει μια μη γραμμική σχέση μεταξύ της y και της περιόδου επαναφοράς T . Μερικές χαρακτηριστικές τιμές των ζευγών (y , T) είναι οι εξής:

(0, 1.58) - η πιο πιθανή πλημμύρα, (0.3665, 2) - διάμεση πλημμύρα, (0.5772, 2.33) - μέση πλημμύρα.

7.2.2. Χαρτί πιθανότητας Gumbel - Σχεδίαση κατανομής

Το χαρτί πιθανότητας Gumbel χαράσσεται με βάση τη σχέση (y , T) της εξ.(7.7). Διαλέγονται δηλαδή κάποιες τιμές της T , όπως 1.01, 1.1, 1.2, 1.3, 1.5, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 100, 200 και 250 και βρίσκονται οι αντίστοιχες τιμές της y από την εξ.(7.7). Στη συνέχεια χαράσσονται κάθετες γραμμές κατά γραμμική αναλογία των τιμών y , πάνω στις οποίες γράφονται οι τιμές της T , ή της αντίστοιχης πιθανότητας υπέρβασης Φ ή της συμπληρωματικής της Φ' . Έτσι, οι τιμές y είναι σε γραμμική κλίμακα και οι περίοδοι επαναφοράς ή οι πιθανότητες είναι σε διπλή εκθετική κλίμακα. Οι οριζόντιες γραμμές του χαρτιού χαράσσονται σε απλή αριθμητική γραμμική κλίμακα. Στο χαρτί αυτό είναι φυσικό η σχέση μεταξύ x και y της εξ.(7.3) να είναι μια ευθεία γραμμή. Αντικαθιστώντας τις τιμές του y στην εξ.(7.3) με τη συνάρτηση του T από την εξ.(7.7) προκύπτει η σχέση της μέγιστης τιμής x με την περίοδο επαναφοράς T ως εξής:

$$x(T) = u - \ln\{\ln(T) - \ln(T-1)\} / \alpha \quad (7.8)$$

που είναι και πάλι ευθεία γραμμή στο χαρτί πιθανότητας Gumbel. Η χάραξη πάνω στο χαρτί Gumbel των ζευγών (x , T) από κάποιο δείγμα μεγίστων τιμών πάνω σε ευθεία γραμμή αποτελεί και μια δοκιμή καταλληλότητας της κατανομής Gumbel στην περίπτωση αυτή. Η αντιστοίχιση της μέγιστης x με την κατάλληλη περίοδο επαναφοράς T γίνεται με τη βοήθεια κάποιας σχέσης που δηλώνει τη σχεδιαστική θέση (plotting position) κάθε μέγιστης τιμής. Υπάρχουν αρκετές σχέσεις προσδιορισμού της σχεδιαστικής θέσης. Όλες υπολογίζουν την περίοδο επαναφοράς σαν συνάρτηση της σχετικής θέσης m της κάθε μέγιστης τιμής μέσα στο δείγμα, όπου οι τιμές

μεγίστων είναι καταταγμένες κατά φθίνουσα τάξη μεγέθους, π.χ. στη θέση $m=1$ είναι η μεγαλύτερη τιμή του δείγματος και στη θέση $m=n$ (n το μήκος του δείγματος) η μικρότερη τιμή. Μια απλή σχέση που εφαρμόστηκε στην Καλιφόρνια είναι η $T=n/m$, που σημαίνει ότι η μικρότερη τιμή έχει πιθανότητα $\Phi=1/T=1$, τιμή που δεν βρίσκεται σε χαρτί πιθανότητας. Ο τύπος Weibull, που συστήνεται για χρήση και από τον ίδιο τον Weibull (1958α) είναι ο εξής:

$$T = (n + 1) / m \quad (7.9)$$

Υπάρχουν και άλλοι τύποι από τους οποίους οι πιο γνωστοί είναι ο τύπος Gringorten (1963): $T = (n+0.012)/(m-0.44)$ που χρησιμοποιείται περισσότερο στις άλλες ασυμπτωτικές κατανομές ακραίων τιμών και ο τύπος Hazen (1930): $T = n / (m - 1/2)$ που συνήθως χρησιμοποιείται στην κατανομή γάμμα.

7.2.3. Διαδικασία ρύθμισης κατανομής

Υπάρχουν πολλές μέθοδοι για την εκτίμηση των παραμέτρων της κατανομής Gumbel. Η μέθοδος των ροπών, η μέθοδος Gumbel, η μέθοδος των συντελεστών συχνότητας, η μέθοδος της μέγιστης πιθανότητας, η μέθοδος της μέγιστης εντροπίας, η μέθοδος των ζυγισμένων ροπών πιθανοτήτων (Kottegoda 1980· Phien, 1987· Stedinger και άλλοι, 1993). Οι τρεις πρώτες που είναι οι πλέον χρησιμοποιούμενες, γιατί είναι πιο απλές αν και έχουν υποστεί κριτική από στατιστικής πλευράς (Phien, 1987), θα αναφερθούν εδώ.

Η μέθοδος των ροπών. Γίνεται άμεσα ο υπολογισμός των παραμέτρων α και u της κατανομής με βάση το μέσο όρο και τη σκέδαση της δειγματικής σειράς των μεγίστων x . Οι εκφράσεις των παραμέτρων με τη μέθοδο αυτή έχουν ήδη δοθεί στη σχετική παράγραφο 3.3.8, όπου αν αντικατασταθούν οι θεωρητικές τιμές του μέσου όρου μ και της τυπικής απόκλισης σ του πληθυσμού των $\bar{x}$ με τους εκτιμητές του $\bar{x}$ και S_x αντίστοιχα από το δείγμα, οι παράμετροι υπολογίζονται ως εξής:

$$\begin{aligned} \hat{\alpha} &= 1.282 / S_x \\ \hat{u} &= \bar{x} - 0.45 S_x \end{aligned} \quad (7.10)$$

Χρησιμοποιώντας αυτές τις τιμές στην εξ. (7.8) προκύπτει η εξής σχέση:

$$\hat{x}(T) = \bar{x} - S_x [0.4500 + 0.7797 \ln \{ \ln(T) - \ln(T-1) \}] \quad (7.11)$$

Χρησιμοποιώντας την εξ. (7.11) μπορεί να χαραχθεί η θεωρητική ευθεία γραμμή

της κατανομής πάνω στο χαρτί Gumbel, αφού υπολογισθούν οι τιμές $\hat{x}$ και S_x από το δείγμα των μεγίστων. Με τη μέθοδο αυτή τα σημεία (x, T) σχεδιάζονται συνήθως με τον τύπο Gringorten για το T .

Η μέθοδος Gumbel. Με τη μέθοδο αυτή η διαδικασία αρχίζει με τον υπολογισμό της μέσης τιμής $\bar{y}_n$ και της τυπικής απόκλισης σ_n της σειράς των τροποποιημένων μεταβλητών y_n . Οι τελευταίες προκύπτουν αν σε κάθε θέση m ($m = 1, 2, \dots, n$) των n τιμών του σε φθίνουσα σειρά καταταγμένου δείγματος των μεγίστων υπολογισθεί η σχεδιαστική θέση με τον τύπο Weibull: $T = (n+1)/m$ και στη συνέχεια με βάση την εξ. (7.7) υπολογισθούν τα αντίστοιχα y_n . Για τις παραμέτρους $\bar{y}_n$ και σ_n που είναι συναρτήσεις του n , υπάρχουν πίνακες υπολογισμού (Gumbel, 1954, 1958α· Kottegoda, 1980), όπως ο Πίνακας 7.1.

Πίνακας 7.1: Αναμενόμενες τιμές μέσω των όρων $\bar{y}_n$ και τυπικών αποκλίσεων σ_n της τροποποιημένης μεταβλητής κατά Gumbel

n	$\bar{y}_n$	σ_n	n	$\bar{y}_n$	σ_n	n	$\bar{y}_n$	σ_n
16	0.5154	1.0306	33	0.5388	1.1225	50	0.5485	1.1607
17	0.5177	1.0397	34	0.5396	1.1256	51	0.5489	1.1623
18	0.5198	1.0481	35	0.5403	1.1285	52	0.5493	1.1638
19	0.5217	1.0557	36	0.5411	1.1313	53	0.5497	1.1653
20	0.5236	1.0628	37	0.5417	1.1339	54	0.5501	1.1668
21	0.5252	1.0694	38	0.5424	1.1365	55	0.5504	1.1682
22	0.5268	1.0755	39	0.5430	1.1390	56	0.5508	1.1695
23	0.5282	1.0812	40	0.5436	1.1413	57	0.5511	1.1709
24	0.5296	1.0865	41	0.5442	1.1436	58	0.5515	1.1722
25	0.5309	1.0914	42	0.5448	1.1458	59	0.5518	1.1734
26	0.5321	1.0961	43	0.5453	1.1479	60	0.5521	1.1747
27	0.5332	1.1005	44	0.5458	1.1499	70	0.5548	1.1854
28	0.5343	1.1047	45	0.5463	1.1518	80	0.5569	1.1938
29	0.5353	1.1086	46	0.5468	1.1537	90	0.5586	1.2007
30	0.5362	1.1124	47	0.5472	1.1555	100	0.5600	1.2065
31	0.5371	1.1159	48	0.5477	1.1573	∞	0.5772	1.2825
32	0.5380	1.1193	49	0.5481	1.1590			

Παίρνοντας τις αναμενόμενες τιμές και στα δύο μέρη της εξ. (7.1) προκύπτει ότι:

$$E(y) = \alpha \{E(x) - u\} \quad (7.12)$$

και παίρνοντας τις αναμενόμενες τιμές πάλι αφού υψωθεί η σχέση στο τετράγωνο προκύπτει η σχέση:

$$\sigma_y^2 = \alpha^2 \sigma_x^2 \quad (7.13)$$

Αντικαθιστώντας τις $E(y) = \bar{y}_n$ και $\sigma_y^2 = \sigma_n^2$ και χρησιμοποιώντας στη θέση των $E(x)$ και σ_x^2 τους εκτιμητές τους από το δείγμα $\bar{x}$ και S_x^2 προκύπτει ότι:

$$\hat{a} = \sigma_n / S_x$$

$$\hat{u} = \bar{x} - \bar{y}_n S_x / \sigma_n \quad (7.14)$$

Άρα η προς χάραξη ευθεία γραμμή στο χαρτί Gumbel είναι η εξής:

$$\hat{x}(T) = \bar{x} - (S_x / \sigma_n) [\bar{y}_n + \ln \{\ln(T) - \ln(T-1)\}] \quad (7.15)$$

Η μέθοδος των συντελεστών συχνότητας. Σύμφωνα με τον Chow (1951, 1964) η σχέση μεταξύ x και T μπορεί να γραφεί ως εξής:

$$x(T) = \mu + K(T)\sigma \quad (7.16)$$

όπου μ , σ ο μέσος όρος και η τυπική απόκλιση αντίστοιχα του πληθυσμού των ετησίων μεγίστων και $K(T)$ συντελεστής που είναι συνάρτηση της περιόδου επαναφοράς T . Η σχέση $\{K(T), T\}$ εξαρτάται από τη μορφή της συνάρτησης πυκνότητας πιθανότητας των μεγίστων. Για τη συνάρτηση Gumbel και με τη μέθοδο των ροπών είναι φανερό ότι:

$$K(T) = -[0.4500 + 0.7797 \ln \{\ln(T) - \ln(T-1)\}] \quad (7.17)$$

Οπότε για κάθε T υπολογίζεται μια τιμή του $K(T)$. Με τη μέθοδο ρύθμισης Gumbel η σχέση γίνεται:

$$K(T) = -[\bar{y}_n + \ln \{\ln(T) - \ln(T-1)\}] / \sigma_n \quad (7.18)$$

Οπότε με βάση το μέγεθος του δείγματος n από τον Πίνακα 7.1 υπολογίζει κανείς τα $\bar{y}_n$ και σ_n και στη συνέχεια για κάθε T την τιμή του συντελεστή συχνότητας. Τέλος, χρησιμοποιώντας τις τιμές του $K(T)$ για κάθε T από την εξ.(7.17) ή την εξ.(7.18) (ανάλογα με τη μέθοδο ρύθμισης που χρησιμοποιείται) και τους εκτιμητές $\bar{x}$ και S_x από το δείγμα των x για τα μ και σ , μπορεί να υπολογισθεί η τιμή της πλημμύρας $x(T)$ από την εξ.(7.16).

Παράδειγμα ρύθμισης κατανομής Gumbel σε σειρά μεγίστων ετησίων παροχών

Σε μια θέση ποταμού μετράται η παροχή για $n = 20$ χρόνια. Οι μέγιστες ετήσιες (μέσες ημερήσιες τιμές) παροχές σε $\mu^3\delta\lambda^{-1}$ για την περίοδο αυτή δίνονται στον Πίνακα 7.2 καταταγμένες σε φθίνουσα τάξη μεγέθους και μαζί με τη σχεδιαστική θέση της κάθε μιας κατά Weibull: $T = (n + 1)/m$.

Πίνακας 7.2: Κατάταξη σειράς ημερήσιων μεγίστων ετησίων παροχών

m	$x(T)$	T	m	$x(T)$	T	m	$x(T)$	T	m	$x(T)$	T
1	330	21	6	266	3.7	11	242	1.91	16	222	1.31
2	310	10.5	7	260	3.0	12	240	1.75	17	218	1.23
3	292	7.0	8	255	2.63	13	237	1.62	18	212	1.17
4	280	5.25	9	250	2.33	14	235	1.5	19	205	1.10
5	272	4.2	10	248	2.1	15	230	1.4	20	195	1.05

Ζητείται ο υπολογισμός της κατανομής Gumbel ρυθμισμένης πάνω στη σειρά των μεγίστων και η πλημμύρα σχεδιασμού για $T = 1000$ χρόνια και $T = 100$ χρόνια.

Για τη ρύθμιση της κατανομής θα εφαρμοσθούν και οι τρεις μέθοδοι ρύθμισης.

(α) *Μέθοδος ροπών.* Η κατανομή δίνεται από την εξ.(7.11), όπου $\bar{x} = 249.95\mu^3\delta\lambda^{-1}$ και $S_x = 34.636\mu^3\delta\lambda^{-1}$. Άρα η σχέση για $T = 1000$ και 100 χρόνια γίνεται:

$$\begin{aligned}\hat{x}(1000) &= 249.95 - 34.636 [0.4500 + 0.7797 \ln \{ \ln(1000) - \ln(999) \}] = \\ &= 421\mu^3\delta\lambda^{-1}\end{aligned}$$

$$\begin{aligned}\hat{x}(100) &= 249.95 - 34.636 [0.4500 + 0.7797 \ln \{ \ln(100) - \ln(99) \}] = \\ &= 359\mu^3\delta\lambda^{-1}\end{aligned}$$

(β) *Μέθοδος Gumbel.* Οι παράμετροι $\bar{y}_n$ και σ_n για $n=20$ υπολογίζονται από τον Πίνακα 7.1 ως εξής: $\bar{y}_n = 0.5236$ και $\sigma_n = 1.0628$. Άρα η εξ.(7.15) γίνεται:

$$\begin{aligned}\hat{x}(1000) &= 249.95 - (34.636/1.0628) [0.5236 + \ln \{ \ln(1000) - \ln(999) \}] = \\ &= 458\mu^3\delta\lambda^{-1}\end{aligned}$$

$$\begin{aligned}\hat{x}(100) &= 249.95 - (34.636/1.0628) [0.5236 + \ln \{ \ln(100) - \ln(99) \}] = \\ &= 383\mu^3\delta\lambda^{-1}\end{aligned}$$

(γ) *Μέθοδος συντελεστών συχνότητας.* Χρησιμοποιώντας την εξ.(7.18) για

τον υπολογισμό των συντελεστών $K(T)$ προκύπτει:

$$\hat{K}(1000) = -[0.5236 + \ln \{ \ln(1000) - \ln(999) \}] / 1.0628 = 6.0065$$

$$\hat{K}(100) = -[0.5236 + \ln \{ \ln(100) - \ln(99) \}] / 1.0628 = 3.8357$$

Άρα οι μέγιστες παροχές είναι:

$$\hat{x}(1000) = 249.95 + 6.0065 \cdot 34.636 = 458 \mu^3 \delta\lambda^{-1}$$

$$\hat{x}(100) = 249.95 + 3.8357 \cdot 34.636 = 383 \mu^3 \delta\lambda^{-1}$$

Δηλαδή ίδιες με της μεθόδου (β) όπως εξ άλλου αναμενόταν. Στο Σχήμα 7.1 δίνεται η χάραξη των ευθειών της κατανομής Gumbel με τη μέθοδο (α) και (β). Είναι φανερό ότι η μέθοδος (β) δίνει καλύτερη προσαρμογή πάνω στα σημεία (x, T) του δείγματος.

7.2.4. Όριο ασφαλείας της κατανομής

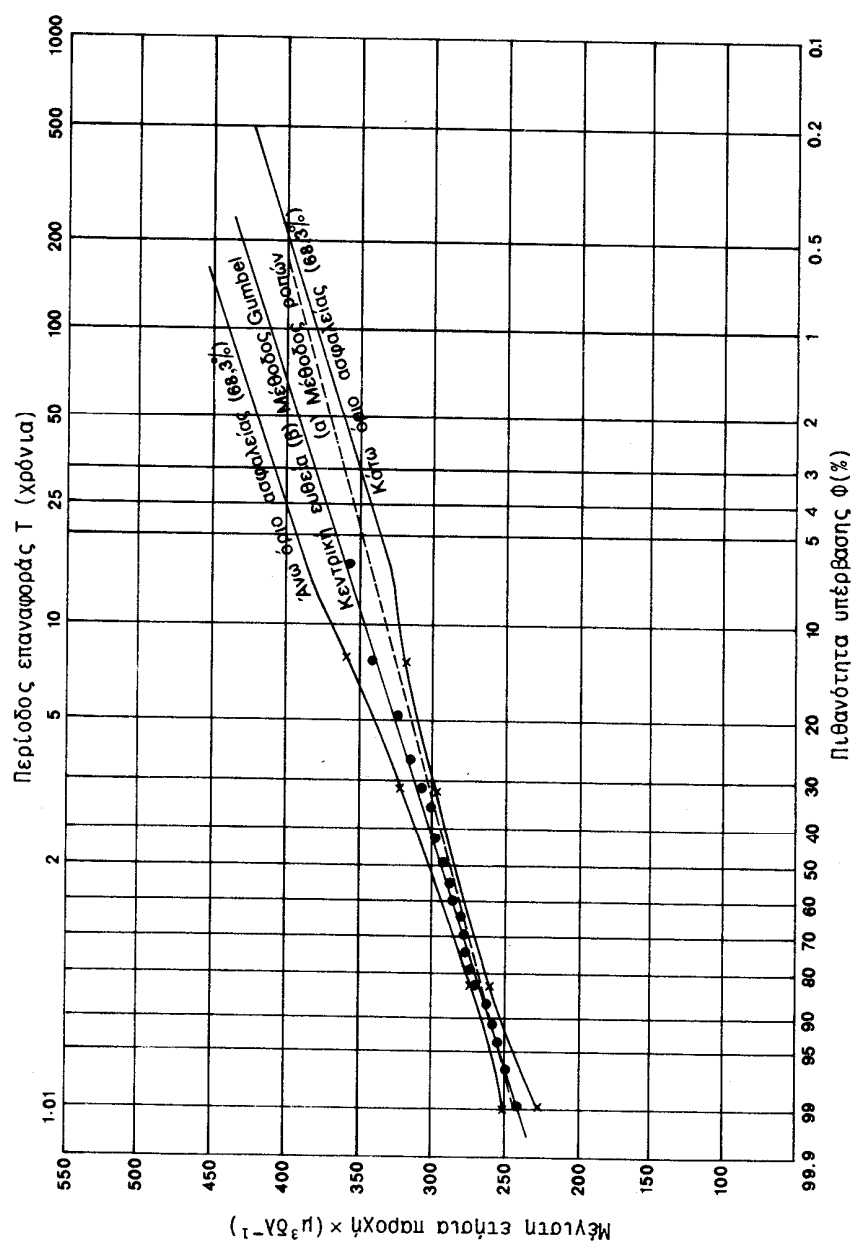
Η τιμή $x(T)$ της μεταβλητής που αντιστοιχεί σε ένα T είναι και αυτή μια μεταβλητή που συνήθως θεωρείται ότι έχει μια κανονική κατανομή (Gumbel, 1958α), όπως φαίνεται στο Σχήμα 7.2.

Άρα, η πιθανότητα P η τιμή $x(T)$ για κάθε T (στην κάθε δηλαδή θέση m), που υπολογίζεται από την κατανομή Gumbel να κυμαίνεται ανάμεσα σε k τυπικές αποκλίσεις S_m , $x(T) - kS_m < x(T) < x(T) + kS_m$ με kS_m το όριο ασφαλείας της κεντρικής τιμής $x(T)$, είναι για $k = 1$, $P = 68,3\%$, για $k = 2$, $P = 95,4\%$ για $k = 3$, $P = 99,7\%$ κ.ο.κ., όπως έχει δειχθεί και στη σχετική παράγραφο 3.5. Σε κάθε θέση m του δείγματος αντιστοιχεί μια περίοδος επαναφοράς T και μια τροποποιημένη μεταβλητή y_m . Ο Gumbel δίνει την έκφραση για την τυπική απόκλιση S_m της μεταβλητής $x(T)$ στη θέση m ως εξής:

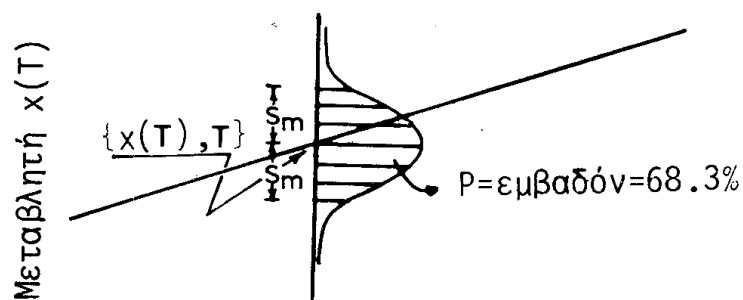
$$S_m = \frac{1}{a\sqrt{n}} f(\sigma_{y_m}) \quad (7.19)$$

Όπου $f(\sigma_{y_m})$ είναι συνάρτηση της τυπικής απόκλισης της y_m μέχρι τη θέση m υπολογιζόμενη (Gumbel, 1958α) από την:

$$f(\sigma_{y_m}) = (1/\Phi' - 1)^{1/2} / (-\ln \Phi') \quad (7.20)$$



Σχήμα 7.1. Χάραξη κατανομής Gumbel.



Σχήμα 7.2. Συνάρτηση κατανομής της μεταβλητής $x(T)$.

Στον Πίνακα 7.3 δίνονται μερικές τιμές των Φ' , y και $f(\sigma_{y_m})$ κατά Gumbel.

Πίνακας 7.3: Τιμές $f(\sigma_{y_m})$ για τα όρια εμπιστοσύνης της κατανομής Gumbel

Φ'	y	$f(\sigma_{y_m})$	Φ'	y	$f(\sigma_{y_m})$
0.01	-1.52718	(2.1607)	0.50	0.36651	1.4427
0.02	-1.36405	(1.7894)	0.55	0.51444	1.5130
0.05	-1.09719	(1.4550)	0.60	0.67173	1.5984
0.10	-0.83403	(1.3028)	0.65	0.84215	1.7034
0.15	-0.64034	1.2548	0.70	1.03093	1.8355
0.20	-0.47588	1.2427	0.75	1.24590	2.0069
0.25	-0.32663	1.2494	0.80	1.49994	2.2408
0.30	-0.18563	1.2687	0.85	1.81696	2.5849
0.35	-0.04862	1.2981	0.90	2.25037	(3.1639)
0.40	0.08742	1.3366	0.95	2.97020	(4.4721)
0.45	0.22501	1.3845	0.98	3.90194	(7.0710)
			0.99	4.60015	(10.000)

Ο Gumbel δέχεται ότι η κανονική κατανομή για την $x(T)$ στη θέση T ισχύει στο διάστημα από $\Phi' = 0.15$ έως $\Phi' = 0.85$. Στον Πίνακα 7.3 οι τιμές μέσα στις παρενθέσεις δεν ανήκουν σ' αυτό το διάστημα. Για τις εκτός του διαστήματος αυτού τιμές ο Gumbel δεν δέχεται ότι ισχύει η κανονική κατανομή. Για τις ακραίες αυτές τιμές στις θέσεις δηλαδή: $1/n + 1$ και $(n - i) / (n + 1)$ (όπου $i = 1, 2, 3$) ο Gumbel δίνει τα όρια ασφαλείας ($x_m \pm \Delta x_m$) της κεντρικής τιμής $x(T)$ για $P = 68.3\%$ και $P = 95.4\%$ όπως φαίνονται στον Πίνακα 7.4.

Πίνακας 7.4: Όρια ασφαλείας της κατανομής Gumbel για τις ακραίες τιμές του δείγματος

Θέση Φ'	Δx_m $P = 68.3\%$	Δx_m $P = 95.4\%$
$n/(n+1)$	$1.14078/\alpha$	$3.0669/\alpha$
$(n-1)/(n+1)$	$0.75409/\alpha$	$1.7820/\alpha$
$(n-2)/(n+1)$	$0.589/\alpha$	$1.35/\alpha$
$(n-3)/(n+1)$	$0.038/\alpha$	$1.17/\alpha$
$1/(n+1)$	$1.14078/(\alpha \ln n)$	$3.0699/(\alpha \ln n)$

Μετά την τελευταία τιμή τα όρια ασφαλείας βρίσκονται εκατέρωθεν της κεντρικής ευθείας παράλληλα προς αυτήν σε απόσταση $1.14078/\alpha$. Η χρήση των εξ.(7.19), (7.20) (ή του Πίνακα 7.3) και του Πίνακα 7.4 δίνει τα όρια ασφαλείας γύρω από την κεντρική τιμή $x(T)$, που υπολογίζεται από την ευθεία Gumbel για κάθε T . Έτσι για επίπεδο εμπιστοσύνης P π.χ. 68.3%, αντιστοιχεί ένα $k = 1$, άρα τα όρια ασφαλείας είναι $\pm S_m$ για κάθε m . Οπότε ακολουθεί ο υπολογισμός του S_m και η χάραξη των καμπυλών ασφαλείας γύρω από την κεντρική ευθεία της κατανομής.

Παράδειγμα εκτίμησης ορίων ασφαλείας κατανομής Gumbel

Στο προηγούμενο παράδειγμα χαράσσονται τα όρια ασφαλείας της κατανομής στην περίπτωση (β) για $P = 68.3\%$ (ένα τυπικό σφάλμα εκτίμησης). Οι παράμετροι α της κατανομής είναι ίσες με:

$$\alpha = 1.282/S_x = 1.282/34.636 = 0.037$$

Άρα ο συντελεστής $1/\alpha\sqrt{n} = 6.0434$ και για τα S_m , και Δx_m σε κάθε θέση m προκύπτει ο εξής πίνακας:

Πίνακας 7.5: Όρια ασφαλείας κατανομής Gumbel

m	Φ'	$f(\sigma_{y_m})$	$\pm S_m = \pm 6.0434 f(\sigma_{y_m})$	$\pm \Delta x_m$
$m = 1$	0.9524	-	-	30.83
$m = 2$	0.905	-	-	20.52
$m = 5$	0.762	2.0562	12.43	-
$m = 10$	0.524	1.4747	8.91	-
$m = 15$	0.286	1.2622	7.63	-
$m = 20$	0.048	-	-	10.29

Στον πίνακα αυτό στην πρώτη και στις δύο τελευταίες θέσεις τα όρια ασφαλείας $\pm 4x_m$ υπολογίσθηκαν με βάση τον Πίνακα 7.4. Στη συνέχεια τα όρια χαράχθηκαν πάνω στην κατανομή, όπως φαίνεται στο Σχήμα 7.1.

7.2.5. Παρατηρήσεις πάνω στην κατανομή Gumbel

Η κατανομή Gumbel σε σχέση με άλλες κατανομές πιθανοτήτων παρουσιάζει μερικές φορές μεγάλα λάθη στην εκτίμηση των παραμέτρων, έτσι ώστε να υπάρχουν αμφιβολίες για την επέκταση της καμπύλης. Ο Benson (1960) απέδειξε ότι για ένα υποθετικό δείγμα 100 τιμών από μια γνωστή κατανομή Gumbel οι παράμετροι που εκτιμώνται από μικρότερα δείγματα (π.χ. από 40 δείγματα των 25 τιμών) διαφέρουν σημαντικά. Βρέθηκε π.χ. ότι για να γίνει κατορθωτό η εκτίμηση για την πλημμύρα περιόδου επαναφοράς 50 και 100 ετών να κυμαίνεται μέσα σε ένα όριο $\pm 25\%$ της σωστής τιμής (με βεβαιότητα 95%), απαιτούνται δείγματα με ελάχιστο εύρος 39 και 48 χρόνια αντίστοιχα. Θεωρητικά το εύρος του δείγματος πρέπει να είναι αρκετά μεγάλο, ώστε να ισχύει η ασυμπτωτική μορφή της κατανομής Gumbel.

Ο Gumbel (1967) ανησυχούσε για τα λάθη που μπορούν να προέλθουν από περιοδικότητες και τάσεις που ενυπάρχουν στα στοιχεία, όπως επίσης και για την έλλειψη ανεξαρτησίας των τιμών, αφού διάφοροι παράγοντες όπως μετωπικές βροχές, λιώσιμο χιονιού κ.λ.π. δημιουργούν συνθήκες εξάρτησης των πλημμυρικών γεγονότων μεταξύ τους. Ένα άλλο σημείο που επίσης πρέπει να τονισθεί είναι ότι μπορούν να παραχθούν και αρνητικές τιμές της μεταβλητής, όπως π.χ. από την εξ.(7.11) ή (7.15) για πολύ μικρό T (π.χ. $T=1.01$ και για $S_x > 0.6089 \bar{x}$ στην εξ.(7.11)). Ένα άλλο χαρακτηριστικό της κατανομής που δημιουργεί δυσκολίες στην εφαρμογή της είναι η έλλειψη ευελιξίας που έχει όσον αφορά στην ασυμμετρία του δείγματος των τιμών των μεγίστων, αφού η κατανομή έχει δοσμένη σταθερή ασυμμετρία $\gamma=1.139$.

7.3. Λογαριθμική κανονική κατανομή πλημμύρων

Στις εξ.(3.24) και (3.25) δίνεται η διπαραμετρική λογαριθμική κανονική κατανομή πιθανότητας για τη μεταβλητή $y = \ln x$, δηλαδή το λογαριθμικό μετασχηματισμό της αρχικής μεταβλητής x . Στην περίπτωση που χρησιμοποιηθεί ο μετασχηματισμός: $y = \ln(x-x_0)$ προκύπτει η τριπαραμετρική λογαριθμική κατανομή. Η χρήση της λογαριθμικής κανονικής κατανομής βασίζεται στην παραδοχή (Chow, 1954) ότι τα φυσικά αίτια $\{x_1, x_2, \dots, x_r\}$ (μετεωρολογικά, γεωγραφικά κ.λ.π.) που προκαλούν τις πλημμύρες λειτουργούν πολλαπλασιαστικά ως εξής: $(x=x_1 x_2 x_3 \dots x_r)$, όπου x η προκαλούμενη πλημμύρα. Αυτό σημαίνει ότι η μεταβλητή $\ln x$ (ή $\log x$) είναι άθροισμα πολλών παραγόντων και για r αρκετά μεγάλο θα πρέπει να ακολουθεί την κανονική κατανομή σύμφωνα με το θεώρημα του κεντρικού

ορίου. Ακριβώς εξ αιτίας αυτής της παραδοχής (που φυσικά δεν είναι απόλυτα ακριβής αφού οι παράγοντες x_i έχουν διάφορες σχέσεις μεταξύ τους), η λογαριθμική κανονική κατανομή είναι μια προσέγγιση της πραγματικότητας.

Οι παράμετροι $\bar{y}$ και σ_y της τριπαμετρικής κατανομής με τη μέθοδο της μέγιστης πιθανότητας δίνονται στον Πίνακα 3.1 του Κεφ. 3. Με τη μέθοδο των ροπών προκύπτουν οι ακόλουθες σχέσεις, που συνδέουν το μέσο όρο μ , τη σκέδαση σ^2 και το συντελεστή ασυμμετρίας γ_1 του πληθυσμού x με το μέσο όρο $\bar{y}$ και σκέδαση σ_y^2 του πληθυσμού y :

$$\mu = \exp(\bar{y} + \sigma_y^2 / 2) + x_0 \quad (7.21)$$

$$\sigma^2 = \{\exp(2\bar{y} + \sigma_y^2)\} \{\exp(\sigma_y^2) - 1\} \quad (7.22)$$

$$\gamma_1 = \{\exp(3\sigma_y^2) - 3\exp(\sigma_y^2) + 2\} / \{\exp(\sigma_y^2) - 1\}^{3/2} \quad (7.23)$$

Η εξ. (7.23) χρησιμοποιείται για τον υπολογισμό της σ_y όταν ρυθμίζεται μια τριπαμετρική λογαριθμική κατανομή. Οι παράμετροι της διπαμετρικής κατανομής προκύπτουν, από τις σχέσεις των παραμέτρων της τριπαμετρικής με $x_0 = 0$. Επίσης αποδεικνύεται ότι, αν $C_v = \sigma/\mu$ είναι ο συντελεστής σκεδασιμότητας του πληθυσμού x τότε (Aitchison και Brown, 1957):

$$\gamma_1 = C_v^3 + 3C_v \quad (7.24)$$

Η λογαριθμική κανονική κατανομή σε αντίθεση με την κατανομή Gumbel είναι πιο ευέλικτη στην προσαρμογή της πάνω σε σειρές στοιχείων αφού δεν έχει συγκεκριμένη και καθορισμένη ασυμμετρία.

7.3.1. Χαρτί πιθανότητας - Σχεδίαση κατανομής

Εάν μια μεταβλητή x είναι κανονικά κατανομημένη, η παράσταση του x με ανηγμένη μεταβλητή $t = (x - \mu)/\sigma$ θα δώσει σε αριθμητική κλίμακα μια ευθεία γραμμή με σταθερά μ και κλίση σ . Έτσι κάθε x συνδέεται με ένα t που έχει μία συγκεκριμένη πιθανότητα υπέρβασης. Στην οριζόντια κλίμακα του χαρτιού μπορεί κανείς να απεικονίσει αντί για τις τιμές t τις πιθανότητες:

$$\Phi(t) = (2\pi)^{-1/2} \int_{-\infty}^t \{\exp(-t^2 / 2)\} dt$$

Οι τιμές αυτές δίνονται σε ποσοστά στις οριζόντιες κλίμακες του χαρτιού πιθανότητας σαν $\{1 - \Phi(t)\}100$ και $\Phi(t)100$. Η περίοδος επαναφοράς είναι: $T = 1 / \{1 - \Phi(t)\}$.

Το χαρτί για τη λογαριθμική κανονική κατανομή προκύπτει με τον ίδιο τρόπο, με τη διαφορά ότι η κατακόρυφη κλίμακα είναι λογαριθμική. Η

κατανομή δηλαδή προκύπτει και πάλι ευθεία γραμμή πάνω στο χαρτί με κλίση σ_y και σταθερά $\bar{y}$. Αν t_T είναι η τιμή την οποία υπερβαίνει μια τυποποιημένη

κανονική μεταβλητή με πιθανότητα: $1/T = (2\pi)^{-1/2} \int_{t_T}^{\infty} \{\exp(-t^2/2)\} dt$, τότε

η τιμή $\hat{x}(T)$ της αρχικής μεταβλητής προκύπτει (για τη διπαραμετρική κατανομή) σαν:

$$\hat{x}(T) = \exp(\bar{y} + t_T \sigma_y) \quad (7.25)$$

Αντικαθιστώντας από τις εξ.(7.21) και (7.22) (με $x_0 = 0$) προκύπτει ο συντελεστής συχνότητας: $k(T) = \{\hat{x}(T)/\mu - 1\}/C_v$, όπου $C_v = \sigma/\mu$ (σύμφωνα με τη θεώρηση της εξ.(7.16)) ως εξής (Kottagoda, 1980):

$$k(T) = C_v^{-1} \left\{ \left\{ 1/(C_v^2 + 1) \right\}^{1/2} \exp \left[t_T \left\{ \ln(C_v^2 + 1) \right\}^{1/2} \right] - 1 \right\} \quad (7.26)$$

Μια παρόμοια αλλά πιο πολύπλοκη έκφραση που συνδέει το συντελεστή συχνότητας $k(T)$ με την τυποποιημένη μεταβλητή t_T και το συντελεστή σκεδασιμότητας C_v του πληθυσμού x προκύπτει και για την τριπαραμετρική κατανομή χρησιμοποιώντας επί πλέον και την παράμετρο x_0 . Η σχεδίαση της κατανομής γίνεται πάνω στο λογαριθμικό κανονικό χαρτί πιθανότητας σαν ευθεία γραμμή σύμφωνα με την εξ.(7.25). Υπολογίζεται με βάση το C_v του δείγματος των μεγίστων x και την τιμή t_T για κάθε T , ο συντελεστής $k(T)$ από την εξ.(7.26) και στη συνέχεια η τιμή: $x(T) = \bar{x} + k(T)S_x$ όπου $\bar{x}$ και S_x οι εκτιμήσεις μέσου όρου και τυπικής απόκλισης από το δείγμα των x .

Παράδειγμα ρύθμισης λογαριθμικής κανονικής κατανομής

Πάνω στη σειρά των μεγίστων ετησίων τιμών παροχών του Πίνακα 7.2 του προηγούμενου παραδείγματος θα επιχειρηθεί η χάραξη μιας διπαραμετρικής λογαριθμικής κανονικής κατανομής. Υπολογίζεται η τυποποιημένη κανονική τιμή από τον Πίνακα Π-2 του Παραρτήματος για $T = 20$ και $T = 50$ ως εξής: $\Phi'(1.6449) = 1 - 1/20$ και $\Phi'(2.0562) = 1 - 1/50$. Άρα $t_{20} = 1.6449$ και $t_{50} = 2.0562$. Το C_v είναι ίσο με: $C_v = S_x / \bar{x} = 34.636/249.95 = 0.1386$. Από την εξ.(7.26) προκύπτει:

$$k(20) = (1/0.1386) \left(\left\{ 1/(0.1386^2 + 1) \right\}^{1/2} \exp[1.6449 \{ \ln(0.1386^2 + 1) \}^{1/2}] - 1 \right) = 1.7951$$

$$k(50) = (1/0.1386) \left(\left\{ 1/(0.1386^2 + 1) \right\}^{1/2} \exp[2.0562 \{ \ln(0.1386^2 + 1) \}^{1/2}] - 1 \right) = 2.2754.$$

Και άρα οι αντίστοιχες τιμές των μεγίστων είναι:

$$\hat{x}(20) = 249.95 + 1.7951 \cdot 34.636 = 312 \mu^3 \delta \lambda^{-1}$$

$$\hat{x}(50) = 249.95 + 2.2754 \cdot 34.636 = 329 \mu^3 \delta \lambda^{-1}$$

Η απεικόνιση των σημείων πάνω στο λογαριθμικό κανονικό χαρτί πιθανότητας γίνεται και πάλι με τον τύπο Weibull: $T = (n + 1)/m$ και η χάραξη της κατανομής πάνω στα σημεία φαίνεται στο Σχήμα 7.3(α).

7.3.2. Όρια ασφαλείας της κατανομής

Από την εξ.(7.25) για την εκτίμηση της τιμής $\hat{x}(T)$ με βάση την τυποποιημένη μεταβλητή t_T προκύπτει ότι:

$$\ln \{ \hat{x}(T) \} = \bar{y} + t_T \sigma_y \quad (7.27)$$

Ωστόσο, είναι γνωστό ότι η σκέδαση της $\bar{y}$ και της σ_y συνδέεται με τη σκέδαση σ_y^2 ως εξής: $\sigma_{\bar{y}}^2 = \sigma_y^2 / n$ και $\sigma_{\sigma_y}^2 = \sigma_y^2 / 2n$. Και οι δύο αυτές ποσότητες εκτιμούνται με βάση τη σκέδαση S_y^2 του δείγματος. Άρα:

$$\sigma_{\ln \hat{x}(T)}^2 = S_y^2 / n + t_T^2 S_y^2 / 2n \quad (7.28)$$

Η εξ.(7.28) προκύπτει αν υψωθεί στο τετράγωνο η εξ.(7.27) και ληφθούν αναμενόμενες τιμές και στα δύο μέλη της, λαμβάνοντας υπόψη ότι οι μεταβλητές y και σ_y είναι ανεξάρτητες μεταξύ τους (Shuster, 1973). Τα όρια ασφαλείας στο επίπεδο εμπιστοσύνης $100(1-\alpha) \%$ του $x(T)$ δίνονται από την εξής σχέση:

$$\exp \{ \ln \hat{x}(T) \pm t_{\alpha/2} (\sigma_{\ln \hat{x}(T)}^2)^{1/2} \} \quad (7.29)$$

όπου $t_{\alpha/2}$ είναι η κανονική τυποποιημένη μεταβλητή με πιθανότητα υπέρβασης $\alpha/2$.

Παράδειγμα εκτίμησης ορίων ασφαλείας λογαριθμικής κανονικής κατανομής

Χρησιμοποιώντας τα στοιχεία του προηγούμενου παραδείγματος υπολογίζονται τα όρια ασφαλείας στο επίπεδο 95% ($t_{\alpha/2} = 1.96$) της λογαριθμικής κανονικής κατανομής στα σημεία $T = 100, 50, 20, 5, 2, 1.01$, όπως

φαίνεται στον παρακάτω Πίνακα 7.6. Το S_v εκτιμάται από τις εξ.(7.24) και (7.23) χρησιμοποιώντας $C_v=0.1386$ και $\gamma_l=0.418$ ως εξής: $S_v=0.136$.

Πίνακας 7.6: Όρια ασφαλείας λογαριθμικής κανονικής κατανομής

T	t_T	$[\sigma_{\ln \hat{x}(T)}^2]^{1/2}$	$\ln \hat{x}(T)$ (*)	Άνω όριο ($\mu_3\delta\lambda^{-1}$)	Κάτω όριο ($\mu_3\delta\lambda^{-1}$)
100	2.3263	0.0583	5.8348	383	305
50	2.0562	0.0535	5.7961	365	296
20	1.6449	0.0465	5.7430	342	285
5	0.8416	0.0353	5.6348	300	261
2	0.000	0.0009	5.5215	250.5	249.6
1.01	-2.3263	0.0583	5.2040	204	162

(*) Οι τιμές της στήλης αυτής εκτιμήθηκαν γραφικά από το Σχήμα 7.3(α).

Η χάραξη των καμπυλών ασφαλείας για επίπεδο εμπιστοσύνης 95% φαίνεται στο Σχήμα 7.3(α).

**7.6. Στατιστική ανάλυση
μεγίστων βροχοπτώσεων.
Όμβριες καμπύλες.**

7.6. Στατιστική ανάλυση μεγίστων βροχοπτώσεων. Όμβριες καμπύλες

Έστω ότι h_i και i_i , $i=1,2,\dots,n$ είναι οι διατεταγμένες σε φθίνουσα τάξη μεγέθους μέγιστες ετήσιες τιμές ύψους βροχής h (χιλιοστ.) και έντασης i (χιλιοστ. ώρα⁻¹), σε ένα σταθμό μέτρησης βροχής για κάποια συγκεκριμένη διάρκεια t βροχής μέσα σε n χρόνια μετρήσεων. Η ανάλυση συχνότητας αυτών των σειρών ακολουθεί τη διαδικασία που περιγράφηκε στις προηγούμενες παραγράφους για τις μέγιστες ετήσιες παροχές. Επίσης, για λόγους παρόμοιους με αυτούς που αναφέρθηκαν στις παροχές, μερικές φορές χρησιμοποιούνται σειρές μερικής διάρκειας βροχοπτώσεων. Και πάλι η ανάλυση γίνεται με την ίδια διαδικασία που έχει ήδη περιγραφεί στις παροχές. Επειδή όμως στην περίπτωση των βροχών μπαίνει και μια δεύτερη μεταβλητή, η διάρκεια της, που την προσδιορίζει πέρα από το μέγεθος της, γι' αυτό η ανάλυση συνήθως συνεχίζεται, έτσι ώστε να συμπεριληφθεί και αυτή η μεταβλητή. Αποτέλεσμα αυτής της ανάλυσης είναι ο υπολογισμός των σχέσεων (i,t,T) ή (h,t,T) έντασης (ή ύψους) βροχής-διάρκειας-περιόδου επαναφοράς, που λέγονται και όμβριες καμπύλες. Οι όμβριες καμπύλες είναι απαραίτητες για τον υδρολογικό σχεδιασμό μεγάλου αριθμού υδραυλικών έργων. Με βάση αυτές μπορεί κανείς οποιαδήποτε στιγμή να υπολογίσει για κάθε διάρκεια και περίοδο επαναφοράς το ύψος και την ένταση της βροχής και γενικά να υπολογίσει ένα από τα τρία μεγέθη (h ή i , t, T) όταν ξέρει τα άλλα δύο. Από μια τέτοια εκτίμηση προκύπτει η καταιγίδα σχεδιασμού ενός έργου με συγκεκριμένη συχνότητα επανάλληψης και διάρκεια, στοιχεία που συνήθως επιλέγονται σαν κριτήρια σχεδιασμού (αποφασίζονται δηλαδή από τον μελετητή με βάση οικονομικά, γεωμορφολογικά κ.λ.π. στοιχεία του έργου). Η καταιγίδα σχεδιασμού κατάλληλα κατανομημένη στο χρόνο χρησιμοποιείται στη συνάρτηση απόκρισης της λεκάνης (π.χ. στο μοναδιαίο υδρογράφημα αν η λεκάνη είναι γραμμική) στη θέση του έργου, και έτσι αποκτάται η πλημμύρα σχεδιασμού του έργου (π.χ. υπερχειλιστή) με συγκεκριμένη συχνότητα επανάλληψης και διακινδύνευση. Περισσότερα στοιχεία γι' αυτό το σχεδιασμό θα δοθούν στο 2ο μέρος του βιβλίου στο σχετικό Κεφ. 2.

Η διαδικασία, μετά την ήδη γνωστή ανάλυση συχνότητας της διατεταγμένης σειράς μεγίστων h_i ή i_i για μια διάρκεια t_1 , συνίσταται στην επανάλληψη της ανάλυσης συχνότητας μεγίστων h_i ή i_i μερικών άλλων ακόμα διαρκειών $t_2, t_3, \dots, t_r$. Από την κατάλληλη κατανομή πιθανότητας για κάθε διάρκεια $t_1, t_2, \dots, t_r$ εκτιμάται η τιμή $h_1, h_2, \dots, h_r$ ή $i_1, i_2, \dots, i_r$ του ύψους ή της έντασης βροχής για μια συγκεκριμένη περίοδο επαναφοράς T . Οι αναλυτικές σχέσεις μιας όμβριας καμπύλης μεταξύ ύψους h ή έντασης i και διάρκειας t για συγκεκριμένο T είναι συνήθως των εξής μορφών:

(α) Απλές εκθετικές:

$$h = kt^m \text{ ή } i = kt^{m-1} \quad (7.45)$$

όπου k και m σταθερές. Οι σχέσεις της εξ.(7.45) είναι οι πιο απλές γιατί γραμμικοποιούνται εύκολα με λογαρίθμηση και προσδιορίζονται εύκολα οι τιμές των k και m με τη μέθοδο των ελαχίστων τετραγώνων.

(β) Υπερβολικές της μορφής:

$$h = \frac{kt}{t+b} \text{ ή } i = \frac{k}{t+b} \quad (7.46)$$

όπου b μια επιπλέον σταθερά,

(γ) Μικτές της μορφής:

$$h = \frac{kt}{(t+b)^m} \text{ ή } i = \frac{k}{(t+b)^m} \quad (7.47)$$

Οι σχέσεις (7.47) έχουν σχετικά μεγαλύτερη ευελιξία και γενικότητα αφού περιλαμβάνουν τη διορθωτική παράμετρο b , που διορθώνει τη χρονική κλίμακα δίνοντας τη δυνατότητα βελτιστοποίησης της σκέδασης των σημείων γύρω από την όμβρια καμπύλη. Η απόκτηση των σχέσεων (h ή i , t) και για άλλες περιόδους επαναφοράς T οδηγεί στην απόκτηση των όμβριων καμπύλων (h ή i , t , T) που έχουν, με βάση την εξ.(7.47), την εξής γενική μορφή:

$$h = \frac{ktT^a}{(t+b)^m} \text{ ή } i = \frac{kT^a}{(t+b)^m} \quad (7.48)$$

όπου a μια επιπλέον παράμετρος. Δηλαδή, οι όμβριες καμπύλες είναι ευθείες παράλληλες σε διπλά λογαριθμικό χαρτί με παράμετρο το T και άξονες h ή i και $(t+b)$ με την εξής αναλυτική έκφραση:

$$\log i = \log (kT^a) - m \log (t+b) \quad (7.49)$$

Η ρύθμιση της εξ.(7.49) αρχίζει με την εκτίμηση της διορθωτικής παραμέτρου b έτσι ώστε να βελτιστοποιείται η σκέδαση των σημείων γύρω από τις όμβριες καμπύλες (μπορεί να γίνει αναλυτικά αλλά και γραφικά, πάνω στο διπλά λογαριθμικό χαρτί). Στη συνέχεια, για κάθε T της ανάλυσης και με τη βοήθεια της μεθόδου ελαχίστων τετραγώνων εκτιμώνται τα $A_T = \log(kT^a)$ και m της εξ.(7.49). Τέλος, και με βάση τις τιμές των ζευγών (A_T , $\log T$), που ήδη είναι γνωστές, εκτιμώνται με τη μέθοδο των ελαχίστων τετραγώνων οι τιμές k και a από τη γραμμική σχέση:

$$A_T = \log k + a \log T \quad (7.50)$$

Έτσι εκτιμώνται και οι τέσσερις παράμετροι (b , m , k , a) της εξ.(7.48), που αναλυτικά εκφράζει τις όμβριες καμπύλες για τη συγκεκριμένη θέση του σταθμού. Συνήθως, είναι χρήσιμες οι όμβριες καμπύλες όταν αντιπροσωπεύουν μια ολόκληρη λεκάνη και όχι μια συγκεκριμένη θέση. Σ' αυτή την περίπτωση θα πρέπει για κάθε διάρκεια $t_1, t_2, \dots, t_r$ των αναλυόμενων βροχών να υπολογισθούν οι μέγιστες ετήσιες τιμές μέσες για όλη τη λεκάνη (π.χ. με τη μέθοδο Thiessen) και μετά να γίνει η ανάλυση με τις μέσες τιμές της λεκάνης. Ας σημειωθεί εδώ ότι για τη μεταφορά σημειακής σε επιφανειακή ένταση ισχύουν όσα έχουν αναφερθεί στο θέμα αυτό στο Κεφ. 2.

Μεγάλη σημασία στην ανάλυση συχνότητας των μεγίστων βροχοπτώσεων έχει η απόκτηση των πρωτογενών στοιχείων, δηλαδή των σειρών h_i ή i_i για κάθε διάρκεια βροχής. Ιδιαίτερα για διάρκειες βροχής διαφορετικές από την 24ωρη και τα πολλαπλάσια της απαραίτητο είναι στη θέση ανάλυσης να υπάρχει βροχογράφος, όργανο δηλαδή συνεχούς καταγραφής της βροχής. Από τις ταινίες του βροχογράφου (συνήθως εβδομαδιαίες ή ημερήσιες) και για κάθε χρόνο επιλέγεται το διάστημα $t_1, t_2, \dots, t_r$ ωρών με το μεγαλύτερο ετήσιο ύψος. Η βροχόπτωση που επιλέγεται πρέπει να είναι συνεχής ή με ελάχιστες διακοπές, έτσι ώστε να μη δικαιολογείται η διακοπή του γεγονότος που αναλύεται και η έναρξη ενός νέου. Με άλλα λόγια κάθε μέγιστο θα αναφέρεται σε ένα ενιαίο στατιστικά μετεωρολογικό συμβάν και όχι σε δύο διαδοχικά. Επίσης, προσοχή θα πρέπει να δίνεται στην περίπτωση που υπολογίζεται η μέγιστη h_i ή i_i μιας διάρκειας t μέση όμως για μια λεκάνη. Βρίσκεται για ένα σταθμό η μέγιστη αυτή τιμή (συνήθως σ' αυτόν που παρουσιάζει την υψηλότερη βροχόπτωση από όλους τους άλλους) και στη συνέχεια βρίσκονται οι τιμές στους άλλους σταθμούς της λεκάνης για την ίδια καταιγίδα στην ίδια διάρκεια χρόνου και τέλος η μέση τιμή αυτών, σύμφωνα π.χ. με τη μέθοδο Thiessen. Για το αν αυτή η τιμή είναι πράγματι η μέγιστη μέση της λεκάνης για τη συγκεκριμένη διάρκεια, θα πρέπει να γίνει μια επαλήθευση, εκτιμώντας τη μέγιστη μέση τιμή της λεκάνης για τη διάρκεια αυτή αρχίζοντας τώρα από τη μέγιστη ετήσια σημειακή ίδιας διάρκειας σε ένα άλλο σταθμό της λεκάνης, είτε εξαντλώντας πρακτικά την πιθανότητα να υπάρχει άλλη μέγιστη τιμή μέσα στο έτος αυτό.

Τέλος, σε πολλές χώρες το θέμα της κατάρτισης όμβριων καμπυλών έχει αντιμετωπιστεί μαζικά για εκτεταμένες γεωγραφικές περιοχές. Κατασκευάστηκαν έτσι χάρτες που δίνουν έτοιμα στοιχεία για κάθε περιοχή, βάσει των οποίων μπορεί να καταρτιστούν εύκολα όμβριες καμπύλες σε οποιοδήποτε σημείο, χωρίς να απαιτούνται τα πρωτογενή ιστορικά δεδομένα. Στις ΗΠΑ έχουν κατασκευαστεί τέτοιοι χάρτες από το U.S. Weather Bureau που χρονολογούνται από το 1961 (Viessman και άλλοι, 1989· Wanielista και άλλοι, 1997). Οι χάρτες αυτοί απεικονίζουν ισοϋέτιες καμπύλες μεγίστων υψών βροχής για διάφορες διάρκειες βροχόπτωσης και διάφορες περιόδους επαναφοράς. Στην Ελλάδα έχει γίνει μια πρώτη πιλοτική διερεύνηση προς την

κατεύθυνση κατασκευής ανάλογων χαρτών (Κοζώνης, 1995) με εφαρμογή στη Στερεά Ελλάδα. Οι χάρτες που διατίθενται δεν μπορούν να θεωρηθούν ως τελικοί και επιχειρησιακά αξιοποιήσιμοι, αλλά αποτελούν ένα πρώτο βήμα προς την κατεύθυνση της κατασκευής ενός άτλαντα της χώρας.

Παράδειγμα υπολογισμού ομβρίων καμπυλών

Στον Πίνακα 7.10 δίνονται οι μέγιστες μέσες τιμές βροχοπτώσεων h (χιλιοστά) για διάρκειες: $t= 12,24,36,48$ και 72 ωρών σε μια λεκάνη στη ΒΔ Ελλάδα έκτασης 820 τετρ. χιλιομ. και για μια σειρά $n = 18$ ετών: 1965-1983.

Οι παύλες στον Πίνακα δηλώνουν έλλειψη του αντίστοιχου ύψους βροχής. Για την απόκτηση των στοιχείων χρησιμοποιήθηκαν τρεις βροχογράφοι και 11 βροχόμετρα στη λεκάνη. Τα στοιχεία από τα βροχόμετρα (ημερήσιας ανάγνωσης) εκτιμήθηκαν με βάση την αναλογική σχέση τους ως προς κάποιο βροχογράφο αναφοράς της λεκάνης. Κάθε μέγιστο h διάρκειας t από το βροχογράφο διαιρεμένο με το ημερήσιο ύψος H_{BF} του βροχογράφου τη συγκεκριμένη ημέρα που ανήκει το γεγονός (αν ανήκει σε δύο διαδοχικές ημέρες η αναλογία γίνεται και για τις δύο ημέρες, π.χ. αν μια 12-ωρη βροχή διαρκεί στον βροχογράφο από τις 8 μ.μ. μιας ημέρας μέχρι τις 8 π.μ. της επομένης οι τέσσερις πρώτες ώρες (ύψος h_1) θα χρησιμοποιηθούν για την αναλογία που θα χρησιμοποιηθεί στο βροχόμετρο την προηγούμενη ημέρα και οι άλλες 8 (ύψους $h_2 = h - h_1$) για την αναλογία που θα χρησιμοποιηθεί την επόμενη ημέρα) πολλαπλασιάζεται με το αντίστοιχο ημερήσιο H_{BP} του βροχομέτρου: $H_{BP}h/H_{BF}$, για να δώσει το αντίστοιχο ύψος βροχής για την ίδια διάρκεια t στο βροχόμετρο. Οι μέσες τιμές στη λεκάνη για όλες τις διάρκειες υπολογίστηκαν με τη μέθοδο Thiessen.

Πίνακας 7.10: Μέγιστες ετήσιες μέσες βροχοπτώσεις h (χιλιοστ.)

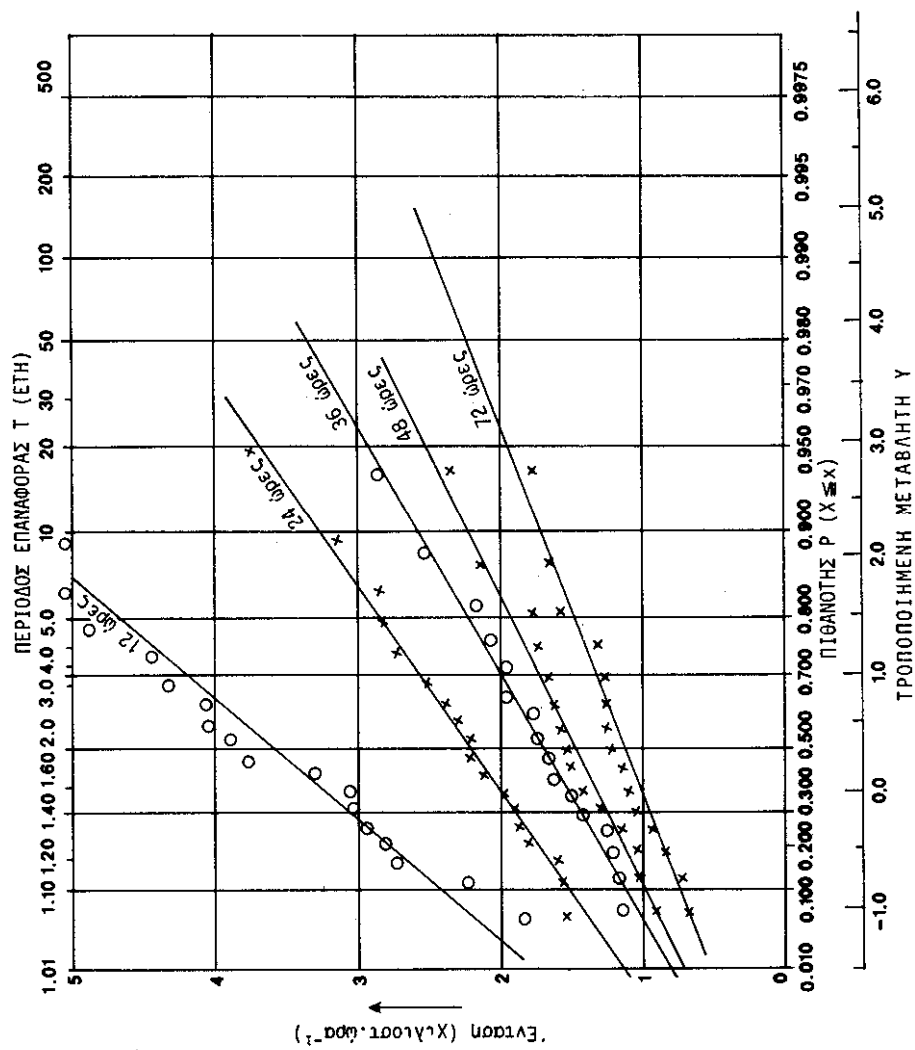
ΕΤΟΣ	12-ωρο	24-ωρο	36-ωρο	48-ωρο	72-ωρο
1965-66	58.7	89.8	—	—	—
1966-67	60.4	67.1	—	77.9	82.9
1967-68	48.5	51.2	64.4	—	—
1968-69	46.8	52.9	54.7	62.7	79.8
1969-70	35.4	52.9	71.0	76.4	91.8
1970-71	61.3	65.7	101.6	112.1	113.5
1971-72	22.2	43.1	45.1	56.1	63.7
1972-73	27.3	45.9	58.1	73.0	73.8
1973-74	33.8	38.2	44.5	50.4	53.6
1974-75	36.7	38.6	43.1	50.9	68.2
1975-76	33.0	38.0	42.5	44.1	50.6

1976-77	51.8	56.1	75.2	84.7	126.3
1977-78	39.7	47.7	63.8	70.8	92.6
1978-79	60.2	67.4	69.7	80.6	89.2
1979-80	45.5	57.7	60.2	—	93.2
1980-81	36.5	45.2	51.9	74.9	90.3
1981-82	52.9	75.0	78.7	82.9	—
1982-83	48.7	60.1	90.4	102.6	118.6

Τα στοιχεία στη συνέχεια μετατράπηκαν σε εντάσεις: $i = h/t$, και κατετάγησαν σε σειρά με φθίνουσα τάξη μεγέθους i_i ($i = 1, 2, \dots, n$ με i_n τη μεγαλύτερη) για όλες τις διάρκειες. Η εφαρμογή της κατανομής Gumbel εξ.(7.15), βρέθηκε να αντιπροσωπεύει πολύ καλά τη συχνότητα εμφάνισης των μεγίστων ετησίων βροχών για όλες τις διάρκειες. Στον Πίνακα 7.11 δίνονται οι παράμετροι $\bar{y}_n$, σ_n της εξ.(7.15) και α και u της εξ.(7.8) για τη ρύθμιση των κατανομών Gumbel για όλες τις διάρκειες. Η χάραξη των καμπυλών σαν ευθειών σε χαρτί πιθανότητας φαίνεται στο Σχήμα 7.6.

Πίνακας 7.11: Παράμετροι κατανομής Gumbel στην ανάλυση των μεγίστων βροχοπτώσεων

Παράμετρος	Διάρκεια (ώρες)				
	12	24	36	48	72
$\bar{y}_n$	0.5202	0.5202	0.5157	0.5128	0.5128
σ_n	1.0493	1.0493	1.0316	1.02057	1.02057
α	1.0564	1.8060	2.1577	2.5982	3.2943
u	3.2093	2.0098	1.5229	1.3313	1.0376



Σχήμα 7.6. Κατανομές Gumbel για όμβριες καμπύλες.

Στη συνέχεια από τις κατανομές Gumbel εκτιμήθηκαν για κάθε διάρκεια και για περιόδους επαναφοράς: $T = 10, 20, 50, 100, 1000$ και 10000 χρόνια, οι τιμές της έντασης i (χιλιοστ. ώρα⁻¹) όπως φαίνονται στον Πίνακα 7.12.

Πίνακας 7.12: Εντάσεις βροχής (από Gumbel) για διάφορες περιόδους επαναφοράς και διάρκειες

Περίοδος επαναφοράς T (χρόνια)	Διάρκεια (ώρες)						
	12	24	36	48	72	A_T	m
10	5.34	3.26	2.57	2.20	1.72	1.39	0.63
20	6.02	3.65	2.90	2.47	1.94	1.44	0.63
50	6.90	4.17	3.33	2.83	2.22	1.50	0.63
100	7.56	4.56	3.65	3.10	2.43	1.54	0.63
1000	9.75	5.83	4.72	3.99	3.13	1.65	0.63
10000	11.93	7.11	5.79	4.88	3.83	1.74	0.63

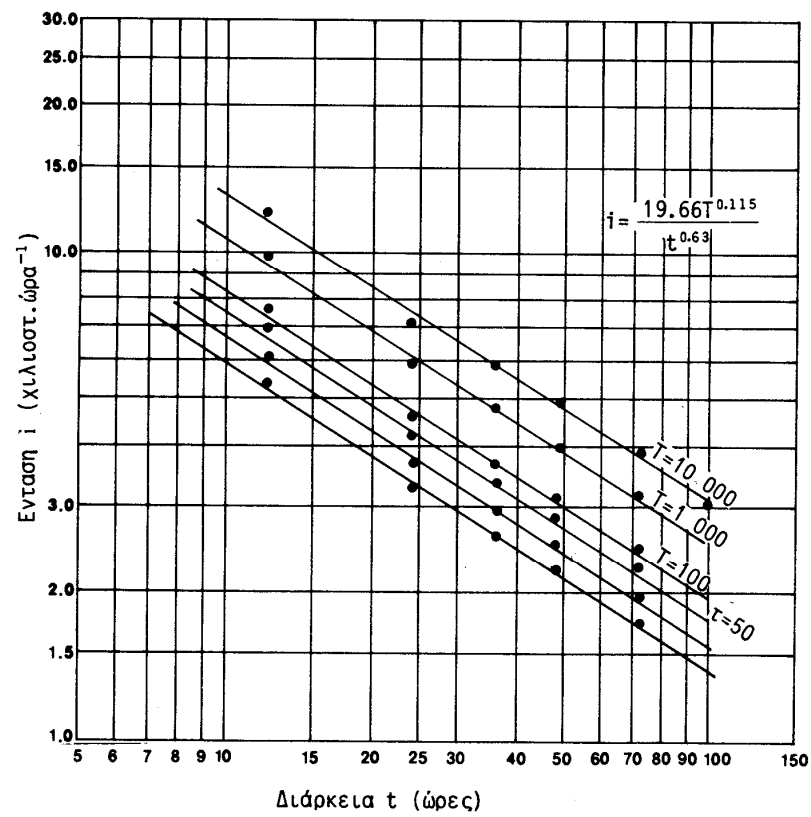
Για κάθε T συσχετίστηκαν γραμμικά τα ζεύγη ($\log i$, $\log t$) και σχεδιάστηκαν στο Σχήμα 7.7. Η παράμετρος b εκτιμήθηκε ίση με μηδέν (σχετικά μικρή διασπορά των σημείων γύρω από τις όμβριες καμπύλες) και με τη βοήθεια των ελαχίστων τετραγώνων εκτιμήθηκαν τα m και $A_T = \log(kT^a)$ της εξ. (7.49), που φαίνονται στις δύο τελευταίες στήλες του Πίνακα 7.12.

Για κάθε T χαράχθηκαν οι εξ. (7.49) σαν ευθείες γραμμές σε διπλά λογαριθμικό χαρτί στο Σχήμα 7.7. Το σταθερό m δηλώνει ακριβώς ότι οι ευθείες για κάθε T στο διπλά λογαριθμικό χαρτί του Σχήματος 7.7 είναι παράλληλες. Τέλος, συσχετίζοντας με γραμμική σχέση τις αντίστοιχες τιμές A_T και $\log T$ από τον Πίνακα 7.12, υπολογίστηκαν τα k και a ίσα με:

$k = 19.66$ και $a = 0.115$. Άρα οι όμβριες καμπύλες για την υπ' όψη λεκάνη έχουν την εξής αναλυτική γενική σχέση:

$$i = \frac{19.66T^{0.115}}{t^{0.63}}$$

Είναι αξιοσημείωτο ότι οι ευθείες γραμμές που προκύπτουν από αυτή την τελευταία γενική για όλα τα T σχέση διαφέρουν λίγο από τις ήδη χαραγμένες ευθείες στο Σχήμα 7.7, που είναι οι ευθείες ελαχίστων τετραγώνων ($\log i$, $\log t$) για κάθε T . Αυτό, οφείλεται στο βαθμό συσχέτισης της ευθείας (A_T , $\log T^a$) που είναι διάφορος της μονάδας ($r = 0.99$).



Σχήμα 7.7. Όμβριες καμπύλες.