Πιθανοθεωρητικές και στατιστικές μέθοδοι στην τεχνική υδρολογία

Δημήτρης Κουτσογιάννης
Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων, Εθνικό Μετσόβιο Πολυτεχνείο,
Ηρώων Πολυτεχνείου 5, 157 00 Ζωγράφου

1 Σύνδεση υδρολογίας και πιθανοθεωρίας

Υπάρχουν τρεις τουλάχιστον λόγοι για τους οποίους η πιθανοθεωρία και η στατιστική αποτελούν το βασικό επιστημονικό εργαλείο της τεχνικής υδρολογίας, σε βαθμό που να θεωρείται η τεχνική υδρολογία ως ιδεώδες πεδίο εφαρμογής αυτών των μαθηματικών κλάδων:

1. Η *τύχη* ενυπάρχει σε όλα τα φαινόμενα της υδρολογίας, ξεκινώντας από την καταιγίδα ή την ανομβρία και καταλήγοντας στην πλημμύρα και την ξηρασία. Έτσι οι νόμοι των πιθανοτήτων κυβερνούν σε μεγαλύτερο ή μικρότερο βαθμό αυτά τα φαινόμενα. Ο όρος *τύχη* χρησιμοποιείται εδώ με την έννοια της μη (αιτιοκρατικής) προβλεψιμότητας. Το θέμα αυτό αναλύεται περισσότερο αμέσως πιο κάτω, στην ένθετη ενότητα “Μη προβλεψιμότητα και τύχη”
2. Η υδρολογική πληροφορία απαρτίζεται από μετρήσεις υδρολογικών διεργασιών, η επεξεργασία των οποίων προϋποθέτει τη χρήση στατιστικών μεθόδων. Ο έλεγχος των σφαλμάτων των μετρήσεων και η συμπλήρωση των ελλείψεων ιστορικών δειγμάτων είναι δύο τυπικές περιπτώσεις που αντιμετωπίζονται με χρήση στατιστικών μεθόδων.
3. Στην Τεχνική Υδρολογία η λήψη τεχνικο-οικονομικών αποφάσεων, που αναφέρονται είτε στο σχεδιασμό έργων είτε στη λειτουργία τους, γίνεται πάντα υπό καθεστώς αβεβαιότητας. Η ορθολογικότερη προσέγγιση-ποσοτικοποίηση της αβεβαιότητας επιτυγχάνεται με τη θεωρία πιθανοτήτων.

Επισημαίνεται, ωστόσο, με έμφαση ότι και οι τελειότερες στατιστικές τεχνικές δεν μπορούν να υποκαταστήσουν:

- την έλλειψη ή ανεπάρκεια της ιστορικής πληροφορίας (μετρήσεων της φυσικής διεργασίας), άρα και την ανεκτίμητη αξία της,
- την έλλειψη αξιοπιστίας της ιστορικής πληροφορίας, η οποία εξαρτάται από τη σοβαρότητα με την οποία οι αρμόδιες υπηρεσίες αντιμετωπίζουν τη συλλογή, καταγραφή, ταξινόμηση και επεξεργασία της, και
- την *έξυπνη* προσέγγιση των προβλημάτων, δηλαδή το συνδυασμό κοινής λογικής με μια ευρύτητα πνεύματος που κατευθύνει τη μαθηματική ανάλυση και σύνθεση, αποφεύγοντας την αδέξια εφαρμογή τυποποιημένων λύσεων.

Μη προβλεψιμότητα και τύχη

Ένα φυσικό φαινόμενο, που διέπεται από προσδιοριστικούς (ντετερμινιστικούς) νόμους, είναι πάντα προβλέψιμο; Η φυσική μας διαίσθηση απαντά θετικά στο ερώτημα αυτό, με την προϋπόθεση ότι είναι γνωστοί οι νόμοι που διέπουν το φαινόμενο. Ίσως γιατί η διαίσθηση μας βασίζεται σε συστήματα που διέπονται από *γραμμικούς* νόμους. Ωστόσο η πραγματικότητα για την πλειονότητα των φυσικών συστημάτων είναι διαφορετική. Τις τελευταίες δεκαετίες έχει διερευνηθεί μια μεγάλη ποικιλία *μη γραμμικών συστημάτων*, φυσικών ή μαθηματικών, τα οποία εμφανίζουν μη περιοδική, ακανόνιστη, απρόβλεπτη, τυχαία συμπεριφορά, αν και διέπονται από προσδιοριστικούς νόμους. Η συμπεριφορά αυτή έχει γίνει γνωστή με τον όρο *προσδιοριστικό χάος*. Τα υδρολογικά και μετεωρολογικά συστήματα είναι τυπικοί αντιπρόσωποι αυτής της συμπεριφοράς, από το μικροσκοπική κλίμακα (π.χ. τύρβη) μέχρι τη μακροσκοπική (πλημμύρες, ξηρασίες κτλ.).

Ο πρώτος που αναγνώρισε αυτή τη συμπεριφορά των μη γραμμικών συστημάτων ήταν ο διάσημος φυσικομαθηματικός Henri Poincaré που πριν από 100 περίπου χρόνια έγραφε στο "Επιστήμη και μέθοδος" (από τον Stewart, 1990):

> *"Μια πολύ ανεπαίσθητη αιτία, η οποία μας διαφεύγει, καθορίζει ένα αξιοσημείωτο φαινόμενο που δεν μπορούμε να μη δούμε, και τότε λέμε ότι το φαινόμενο αυτό οφείλεται στην τύχη. Αν μπορούσαμε να ξέρουμε επακριβώς τους νόμους της φύσης και την κατάσταση του σύμπαντος στην αρχική του στιγμή, θα μπορούσαμε να προβλέψουμε επακριβώς την κατάσταση αυτού του ίδιου του σύμπαντος σε μια μεταγενέστερη χρονική στιγμή. Αλλά, ακόμα και αν οι φυσικοί νόμοι δεν είχαν άλλα μυστικά από εμάς, θα μπορούσαμε να ξέρουμε την αρχική κατάσταση μόνο κατά προσέγγιση. Αν αυτό μας επιτρέπει να προβλέψουμε τη μεταγενέστερη κατάσταση με τον ίδιο βαθμό προσέγγισης, αυτό αρκεί να πούμε ότι το φαινόμενο είχε προβλεφθεί, ότι υπόκειται σε νόμους. Όμως, το ζήτημα δεν είναι πάντοτε έτσι: υπάρχει περίπτωση οι πολύ λεπτές διαφορές στις αρχικές συνθήκες να παράγουν πολύ μεγαλύτερες διαφορές στα τελικά φαινόμενα, ένα ελάχιστο σφάλμα στην αρχή να προκαλεί ένα τεράστιο σφάλμα στο τέλος. Η πρόβλεψη τότε γίνεται αδύνατη κι έτσι έχουμε το φαινόμενο της τύχης.*
>
> *Γιατί άραγε οι μετεωρολόγοι συναντούν τόσο μεγάλες δυσκολίες στην πρόβλεψη του καιρού; Γιατί οι βροχές και οι καταιγίδες μας φαίνονται ότι έρχονται στην τύχη, με αποτέλεσμα πολλοί άνθρωποι να θεωρούν εντελώς φυσικό να προσεύχονται για να μην έρθει βροχή ή για να σταματήσει, ενώ τους φαίνεται γελοίο να προσευχηθούν για μια έκλειψη; Παρατηρούμε ότι οι μεγάλες διαταραχές συμβαίνουν γενικά σε περιοχές όπου η ατμόσφαιρα βρίσκεται σε ασταθή ισορροπία. Οι μετεωρολόγοι γνωρίζουν ότι αυτή η ισορροπία είναι ασταθής, ότι εμφανίζεται κάπου ένας κυκλώνας – όμως δεν μπορούν να ξέρουν πού. Ένα δέκατο του βαθμού περισσότερο ή λιγότερο σε κάποιο σημείο και ο κυκλώνας ξεσπά εδώ και όχι εκεί, σπέρνοντας τον όλεθρο σε χώρες που θα τον είχαν αποφύγει. Αυτό θα μπορούσαμε να το προβλέψουμε αν ξέραμε αυτό το δέκατο του βαθμού, οι παρατηρήσεις όμως δεν είναι αρκετά πυκνές ούτε αρκετά ακριβείς – και για το λόγο αυτό όλα φαίνονται να οφείλονται στον παράγοντα τύχη."*

Ας σημειωθεί ότι ο Poincarι ασχολήθηκε κυρίως με κινήσεις αστρικών σωμάτων, όπου και κει, αντίθετα με την κρατούσα αντίληψη που θέλει να επικρατεί πλήρης τάξη και περιοδικότητα, ανακάλυψε χαοτικά φαινόμενα (βλ. π.χ. Schroeder, 1990). Οι ανακαλύψεις του Poincarι, ωστόσο, έμειναν ανεξερεύνητες μέχρι τη δεκαετία του 1960, οπότε η εισαγωγή των ηλεκτρονικών υπολογιστών επέτρεψε την πληρέστερη μελέτη των μη γραμμικών συστημάτων. Στις αρχές της δεκαετίας του 1960 ο μετεωρολόγος Edward Lorenz τελείως τυχαία, κάνοντας ένα πείραμα σε ηλεκτρονικό υπολογιστή, επιβεβαίωσε τη χαοτική συμπεριφορά της εξέλιξης του καιρού. Ο Lorenz κατάρτισε ένα απλοποιημένο μοντέλο που προσέγγιζε τα φαινόμενα μεταφοράς στην ατμόσφαιρα, το οποίο αποτελείται από τρεις συνήθεις μη γραμμικές διαφορικές εξισώσεις (βλ. Cooper, 1989). Τις εξισώσεις αυτές τις επέλυε αριθμητικά με έναν από τους πρώτους ηλεκτρονικούς υπολογιστές. Ο ίδιος περιγράφει την ανακάλυψή του με τον ακόλουθο τρόπο:

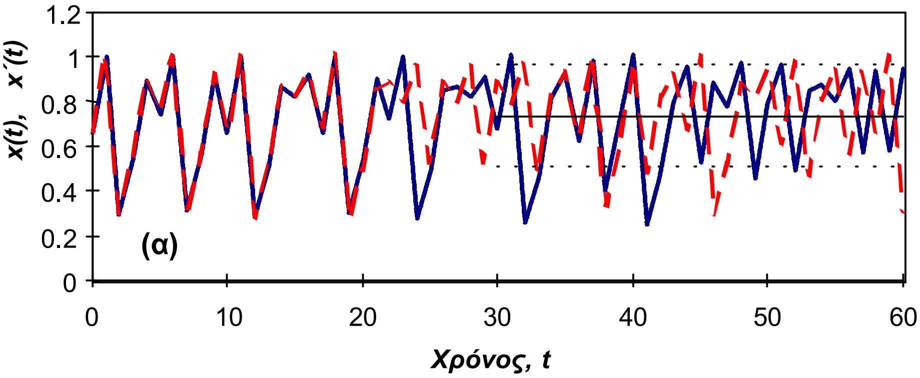
> *"Κατά τη διάρκεια των υπολογισμών μας αποφασίσαμε να εξετάσουμε μια από τις λύσεις σε μεγαλύτερη λεπτομέρεια, και διαλέξαμε κάποιες ενδιάμεσες συνθήκες που είχαν εκτυπωθεί από τον υπολογιστή, τις οποίες εισαγάγαμε ως νέες αρχικές συνθήκες. Όταν γυρίσαμε στον υπολογιστή, μια ώρα αργότερα, και αυτός είχε προσομοιώσει περίπου δύο μήνες "καιρού", ανακαλύψαμε ότι υπήρχε πλήρης διαφωνία με την προηγούμενη λύση. Στην αρχή αυτό το αποδώσαμε σε πρόβλημα της μηχανής, πράγμα που δεν ήταν ασυνήθιστο, αλλά γρήγορα αντιληφθήκαμε ότι οι δύο λύσεις προέρχονταν από τις ίδιες συνθήκες. Οι υπολογισμοί έγιναν με υπολογιστική ακρίβεια περίπου έξι δεκαδικών ψηφίων, αλλά οι εκτύπωση περιείχε μόνο τρία, έτσι ώστε οι νέες αρχικές συνθήκες αποτελούνταν από τις παλιές συν ορισμένες μικρές αποκλίσεις (διαταραχές). Αυτές οι αποκλίσεις μεγεθύνονταν σχεδόν εκθετικά, και διπλασιάζονταν σε περίπου τέσσερις μέρες, έτσι που σε δύο μήνες οι δύο λύσεις έπαιρναν ξεχωριστούς δρόμους."*

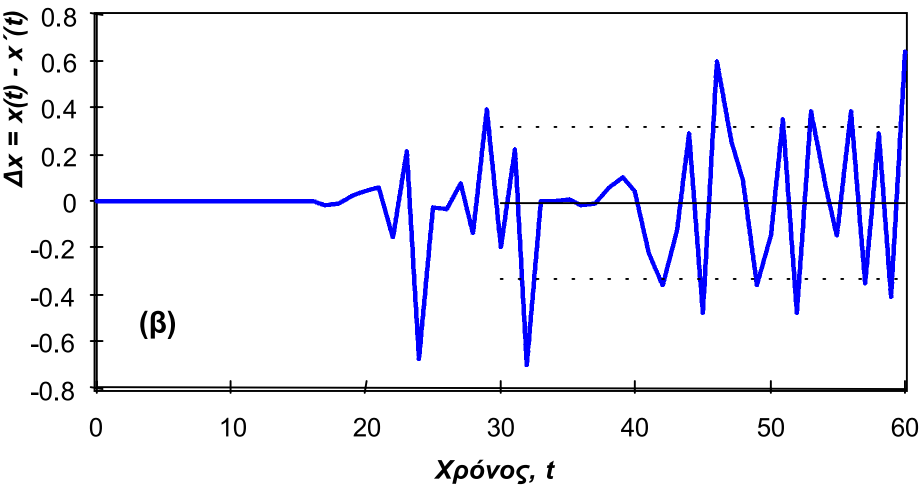
Ένα πολύ απλό παράδειγμα για να δούμε τη μεγέθυνση των διαταραχών στην εξέλιξη ενός μη γραμμικού φαινομένου δίνεται στο Σχήμα 1. Ας θεωρήσουμε ένα σύστημα το οποίο περιγράφεται από μια μοναδική μεταβλητή $x(t)$, όπου t είναι ο χρόνος που στο παράδειγμα θεωρείται διακριτός. Θεωρούμε ότι η κατάσταση του συστήματος στην τρέχουσα χρονική στιγμή t εξαρτάται από την κατάστασή του στην αμέσως προηγούμενη χρονική στιγμή t -1, και ότι η μετάβαση από τη μια στιγμή στην άλλη περιγράφεται από την εξίσωση.

$$x(t) = k\,x(t-1)[1 - x(t-1)] \tag{1}$$

όπου k είναι αριθμητική σταθερά. Η εξίσωση αυτή είναι γνωστή στη βιβλιογραφία (βλ. π.χ. Cooper, 1989, Schroeder, 1990, Stewart, 1990, όπου υπάρχει αναλυτικότερη μελέτη της εξίσωσης) ως *λογιστική απεικόνιση* και βέβαια είναι μια από τις απλούστερες μη γραμμικές εξισώσεις.

Τονίζεται ότι το σύστημα που περιγράφει η εξίσωση είναι πλήρως κλειστό και προσδιοριστικό δηλαδή δεν μεσολαβεί στην εξέλιξή του κανένας εξωτερικός παράγοντας. ούτε προστίθεται καμιά τυχαία συνιστώσα. Παρατηρούμε ότι μετά από κάποιο αρχικό χρόνο (30 χρονικές μονάδες στο παράδειγμα μας) δύο "λύσεις" που αρχικά έχουν ασήμαντη απόκλιση, μόνο $1*10^{-6}$ για t = 0, αποκλίνουν τελείως και κάθε μια ακολουθεί το δικό της ανεξάρτητο δρόμο. Εύκολα συμπεραίνουμε ότι η προσδιοριστική πρόγνωση είναι πολύ αποτελεσματική για μικρά χρονικά διαστήματα (οι αποκλίσεις των δύο "λύσεων" στο Σχήμα 1 είναι δυσδιάκριτες) αλλά οδηγεί σε πολύ σημαντικά σφάλματα για μεγαλύτερα χρονικά διαστήματα. Για τα μεγάλα διαστήματα η στατιστική πρόγνωση είναι πιο αποτελεσματική, αφού και μόνο η στατιστική μέση τιμή δίνει μικρότερο σφάλμα στην εκτίμηση της πραγματικής εξέλιξης της $x(t)$ από το σφάλμα που δίνει η $x΄(t)$. Επιπλέον, η στατιστική προσέγγιση έχει τη δυνατότητα να προσδιορίζει όρια μέσα στα οποία κυμαίνεται μια μεταβλητή, για καθορισμένη πιθανότητα.





Σχήμα 1 Εξέλιξη της λογιστικής απεικόνισης (εξίσωση 1) για σταθερά k = 3.7. Στο διάγραμμα (α) εμφανίζεται με χοντρή συνεχή γραμμή η πραγματική εξέλιξη της $x(t)$ για αρχική τιμή $x(0)$ = 0.660001. και με διακεκομμένη η εξέλιξη $x΄(t)$ αν η αρχική τιμή ληφθεί (π.χ. για λόγους στρογγύλευσης) ίση με 0.66 (η διαφορά από την πραγματική τιμή είναι μόνο $1*10^{-6}$). Στο διάγραμμα (β) δίνεται η διαφορά των $x(t)$ και $x΄(t)$ συναρτήσει του t. Και στα δύο διαγράμματα έχει απεικονιστεί επίσης η μέση τιμή της περιόδου $30 \le t \le 60$ (συνεχείς ευθείες) και το μέσο σφάλμα γύρω από τη μέση τιμή (διακεκομμένες ευθείες). Το μέσο σφάλμα είναι 0.19 για το διάγραμμα (α) και 0.31 για το διάγραμμα (β)

Βεβαίως τα φυσικά υδρομετεωρολογικά συστήματα είναι απείρως πολυπλοκότερα από το σύστημα του παραπάνω παραδείγματος. Εύκολα λοιπόν μπορεί να συναγάγει κανείς την αδυναμία πλήρους προσδιοριστικής προσέγγισής τους και την αναγκαιότητα της πιθανοθεωρητικής προσέγγισης.

2 Γενικές έννοιες θεωρίας πιθανοτήτων

2.1 Θεμελίωση θεωρίας πιθανοτήτων

Η σύγχρονη προσέγγιση της θεωρίας πιθανοτήτων βασίζεται στη θεωρία συνόλων και είναι πολύ απλή. Δομείται από ορισμένες θεμελιώδεις έννοιες και τρία αξιώματα., όπως αναπτύσσονται παρακάτω.

Δειγματικός χώρος
Ορίζεται ως το σύνολο του οποίου τα στοιχεία ω αντιστοιχούν στις δυνατές εκβάσεις ενός πειράματος, ή μιας παρατήρησης ή μέτρησης (συμβολικά Ω).

Γεγονός
Είναι κάθε υποσύνολο του δειγματικού χώρου. *Στοιχειώδες* είναι ένα γεγονός με ένα μόνο στοιχείο. *Βέβαιο γεγονός* είναι ο δειγματικός χώρος (Ω). *Αδύνατο γεγονός* είναι το κενό σύνολο ($\varnothing$).

Οικογένεια γεγονότων
Είναι μια συλλογή $\mathcal{F}$ υποσυνόλων (γεγονότων) A του Ω.

Πιθανότητα
Ορίζεται αξιωματικά. ως μια συνάρτηση P επί της οικογένειας γεγονότων $\mathcal{F}$. Μέσω αυτής, σε κάθε γεγονός A αντιστοιχίζουμε έναν αριθμό $P(A)$ που λέγεται *πιθανότητα του γεγονότος A*.

Αξιώματα της θεωρίας πιθανοτήτων
Αποτελούν συνθήκες τις οποίες οφείλει να ικανοποιεί η συνάρτηση P, και είναι τα ακόλουθα:

I. $$P(A) \geq 0$$

II. $$P(\Omega) = 1$$

(2)

IIIα. $P(A \cup B) = P(A) + P(B)$, με την προϋπόθεση ότι $A \cap B = \varnothing$

IIIβ. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, με την προϋπόθεση ότι $A_i \cap A_j = \varnothing, i \neq j$

Σημειώνεται ότι το αξίωμα IIIβ αποτελεί επέκταση του IIIα για άπειρο αριθμό γεγονότων, αλλά δεν είναι συνέπεια του IIIα, γι' αυτό και εισάγεται ως ανεξάρτητο αξίωμα.

Υπενθυμίζεται ότι η ένωση δύο γεγονότων (συμβολικά $A \cup B$, ή $A+B$, διαβάζεται "A ή Β") είναι επίσης ένα γεγονός που πραγματοποιείται αν πραγματοποιηθεί ένα από τα δύο γεγονότα. Αντίστοιχα, η τομή δύο γεγονότων (συμβολικά $A \cap B$, ή AB, διαβάζεται "A και Β") είναι ένα γεγονός που πραγματοποιείται αν πραγματοποιηθούν ταυτόχρονα και τα δύο γεγονότα. Τέλος δύο γεγονότα που η τομή τους είναι το αδύνατο γεγονός (όπως τα A και B στο αξίωμα IIIα) λέμε ότι είναι *αμοιβαία αποκλειόμενα*.

Τυχαία μεταβλητή
Είναι μια συνάρτηση ορισμένη επί του δειγματικού χώρου. Μέσω αυτής, σε κάθε δυνατή έκβαση ω αντιστοιχίζουμε, βάσει ενός προκαθορισμένου κανόνα, έναν αριθμό $X(\omega)$. Η τυχαία μεταβλητή μπορεί να είναι διακριτή ή συνεχής, ανάλογα αν ο δειγματικός χώρος είναι διακριτός (π.χ. πεπερασμένο σύνολο ή άπειρο αλλά αριθμήσιμο) ή συνεχής (άπειρο μη αριθμήσιμο σύνολο). Για λόγους απλοποίησης στο συμβολισμό μιας τυχαίας μεταβλητής παραλείπουμε το στοιχείο ω στο οποίο αντιστοιχεί και γράφουμε απλώς X.

Συνάρτηση κατανομής
Είναι η συνάρτηση της πραγματικής μεταβλητής x, που δίνεται από την εξίσωση

$$F_X(x) = P(X \le x) \tag{3}$$

όπου X είναι οποιαδήποτε τυχαία μεταβλητή. Η συνάρτηση κατανομής ορίζεται για κάθε πραγματικό αριθμό x, από -∞ μέχρι +∞. Είναι πάντα αύξουσα συνάρτηση, συνεχής από δεξιά, και υπακούει στη σχέση

$$0 = F_X(-\infty) \le F_X(x) \le F_X(+\infty) = 1 \tag{4}$$

Συχνά αποκαλείται και αθροιστική συνάρτηση κατανομής ή και πιθανότητα μη υπέρβασης.

Πιθανότητα υπέρβασης
Είναι η παράσταση

$$F_{1_X}(x) = P(X > x) = 1 - F_X(x) \tag{5}$$

που είναι φθίνουσα συνάρτηση και υπακούει στη σχέση

$$1 = F_{1X}(-\infty) \ge F_{1X}(x) \ge F_{1X}(+\infty) = 0 \tag{6}$$

Συνάρτηση πυκνότητας πιθανότητας
Είναι η παράγωγος της συνάρτησης κατανομής, ήτοι η

$$f_X(x) = \frac{dF(x)}{dx} \tag{7}$$

Για συνεχείς τυχαίες μεταβλητές η συνάρτηση αυτή ορίζεται παντού, δε συμβαίνει όμως το ίδιο στην περίπτωση των διακριτών μεταβλητών. Οι βασικές της ιδιότητες που προκύπτουν άμεσα από τον ορισμό της είναι

$$f_X(x) \ge 0, \int_{-\infty}^{\infty} f_X(x)dx = 1 \tag{8}$$

Σημειώνεται ότι η συνάρτηση πυκνότητας πιθανότητας δεν παριστάνει πιθανότητα και γι' αυτό μπορεί να πάρει τιμές μεγαλύτερες της μονάδας. Η σχέση της με την πιθανότητα προσδιορίζεται από την εξίσωση

$$f_X(x) = \lim_{\Delta x \to 0} \frac{P(x \le X \le x + \Delta x)}{\Delta x} \tag{9}$$

Εφαρμογή 2.1

Για καλύτερη κατανόηση των βασικών εννοιών της θεωρίας πιθανοτήτων δίνουμε το ακόλουθο παράδειγμα από το χώρο της υδρολογίας. Σε ένα συγκεκριμένο τόπο και μια συγκεκριμένη περίοδο του έτους μας ενδιαφέρει η μαθηματική περιγραφή του ενδεχομένου αν βρέχει κατά τη διάρκεια μιας μέρας ή όχι. Θεωρούμε το φυσικό αυτό φαινόμενο ως ισοδύναμο με ένα πείραμα τύχης με δύο δυνατές εκβάσεις: βροχερή μέρα (συμβολικά Β) και στεγνή μέρα (συμβολικά Σ).

Ο δειγματικός χώρος είναι βέβαια πεπερασμένος και διακριτός. Συγκεκριμένα

$$\Omega = \{\mathrm{B}, \Sigma\}$$

Αντίστοιχα, η οικογένεια γεγονότων περιλαμβάνει όλα τα υποσύνολα του δειγματικού χώρου, άρα

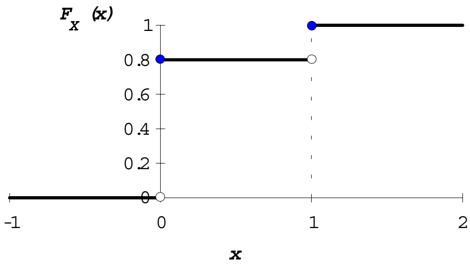
$$\mathcal{F} = \{\varnothing, \{\mathrm{B}\}, \{\Sigma\}, \Omega\}$$

Για να ορίσουμε την πιθανότητα επί του $\mathcal{F}$ φτάνει να ορίσουμε την πιθανότητα ενός από τα δύο ενδεχόμενα, έστω την P(Β). Για να γίνει αυτό με αξιόπιστο τρόπο χρειάζεται να έχουμε στη διάθεσή μας ιστορικά δεδομένα και να χρησιμοποιήσουμε τις αρχές της στατιστικής. Για την ώρα, ας θεωρήσουμε αυθαίρετα ότι $P(\mathrm{B}) = 0.2$. Οι υπόλοιπες πιθανότητες προκύπτουν με βάση τα αξιώματα. Έτσι έχουμε $P(\varnothing) = 0$ και $P(\Omega) = 1$. Δεδομένου ότι τα Β και Σ είναι αμοιβαία αποκλειόμενα ισχύει $P(\mathrm{B}) + P(\Sigma) = P(\mathrm{B}+\Sigma) = P(\Omega) = 1$, άρα $P(\Sigma) = 0.8$.

Ας ορίσουμε την τυχαία μεταβλητή X με τον ακόλουθο τρόπο

$$X(\Sigma) = 0, X(\mathrm{B}) = 1$$

Μπορούμε τώρα να υπολογίσουμε εύκολα τη συνάρτηση κατανομής της X. Για $x < 0$ έχουμε $F_X(x) = P(X \le x) = 0$ (δεδομένου ότι η X, όπως την ορίσαμε, δεν μπορεί να πάρει αρνητικές τιμές). Για $0 \le x < 1$ έχουμε $F_X(x) = P(X \le x) = P(X = 0) = 0.8$. Τέλος για $1 \le x$ έχουμε $F_X(x) = P(X \le x) = P(X = 0) + P(X = 1) = 1$. Η γραφική παράσταση της συνάρτησης κατανομής φαίνεται στο Σχήμα 2. Το κλιμακωτό σχήμα είναι χαρακτηριστικό για διακριτές τυχαίες μεταβλητές. Αντίθετα, για συνεχείς τυχαίες μεταβλητές η συνάρτηση κατανομής είναι παντού συνεχής.



Σχήμα 2 Συνάρτηση κατανομής της εφαρμογής 1.2.

Κλείνουμε την εφαρμογή αυτή με δύο παρατηρήσεις. Πρώτον, η πιθανότητα βροχερής ημέρας και κατά συνέπεια και η συνάρτηση κατανομής εξαρτάται από τη συγκεκριμένη περίοδο του έτους (π.χ. μήνας), καθώς επίσης και από το συγκεκριμένο τόπο, στον οποίο αναφερόμαστε. Δεύτερον, στο συγκεκριμένο πιθανοθεωρητικό μοντέλο που κατασκευάσαμε δε γίνεται καθόλου αναφορά στην αλληλουχία βροχερών-στεγνών ημερών. Αυτό βέβαια δεν είναι λάθος, απλώς περιορίζει την προγνωστική αξία του μοντέλου. Ένα πληρέστερο μοντέλο θα περιέγραφε π.χ. ξεχωριστά την πιθανότητα μιας στεγνής μέρας που ακολουθεί αμέσως μετά από επίσης στεγνή μέρα, η οποία, όπως εμπειρικά γνωρίζουμε, είναι αυξημένη σε σχέση με την πιθανότητα στεγνής μέρας ύστερα από βροχερή μέρα. Τέτοιου είδους μοντέλα, που λαμβάνουν υπόψη τη χρονική αλληλουχία γεγονότων, στηρίζονται στη θεωρία των *στοχαστικών ανελίξεων* και ανήκουν στον ιδιαίτερο κλάδο της *στοχαστικής υδρολογίας*, η ανάπτυξη του οποίου ξεφεύγει από το σκοπό αυτού του κειμένου.

Ανεξάρτητα γεγονότα

Δύο γεγονότα A και B λέγονται *ανεξάρτητα* (ή *στοχαστικώς ανεξάρτητα*), αν

$$P(A \cap B) = P(A)P(B) \tag{10}$$

Σε αντίθετη περίπτωση τα γεγονότα λέγονται στοχαστικώς εξαρτημένα. Ο ορισμός της στοχαστικής ανεξαρτησίας επεκτείνεται και για μια πολλαπλότητα γεγονότων A_1, A_2, …, τα οποία λέγονται ανεξάρτητα αν ισχύει

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_n}) \tag{11}$$

για κάθε πεπερασμένο σύνολο διακεκριμένων δεικτών i_1, i_2, …, i_n.

2.2 Αναμενόμενες τιμές, ροπές, άλλες παράμετροι κατανομών

Αν X είναι συνεχής τυχαία μεταβλητή και $g(X)$ είναι συνάρτηση της X, τότε ορίζεται ως *αναμενόμενη τιμή* ή *προσδοκία της* $g(X)$ το μέγεθος

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \tag{12}$$

Η αντίστοιχη σχέση για διακριτή τυχαία μεταβλητή *X*, που παίρνει τις τιμές $x_1, x_2, \ldots$ είναι

$$E[g(X)] = \sum_{i=1}^{\infty} g(x_i) P(X = x_i) \tag{13}$$

Ειδικότερα:

1. Για $g(X) = X^r$, όπου $r = 0, 1, 2, \ldots$, το μέγεθος

$$m_X^{(r)} := E[X^r] \tag{14}$$

ονομάζεται *ροπή περί την αρχή* (ή απλώς *ροπή*)*τάξης r* της *X*.

2. Για $g(X) = (X - m_X)^r$, το μέγεθος

$$\mu_X^{(r)} := E[(X - m_X)^r] \tag{15}$$

ονομάζεται *κεντρική ροπή τάξης r* της *X*.

Μέση τιμή

Για $g(X) = X$, το μέγεθος

$$m_X := E[X] \tag{16}$$

(δηλαδή η ροπή τάξης 1) ονομάζεται *αναμενόμενη τιμή* ή *μέση τιμή* της *X*. Εναλλακτικά για τη μέση τιμή χρησιμοποιείται και το σύμβολο μ_X.

Διασπορά

Για $g(X) = (X - m_X)^2$, το μέγεθος

$$\sigma_X^2 := \mu_X^{(2)} = E[(X - m_X)^2] \tag{17}$$

(δηλαδή η κεντρική ροπή τάξης 2) ονομάζεται *διασπορά* της *X*. Η διασπορά συμβολίζεται ακόμη και ως $\mathrm{Var}[X]$.

Παράμετροι θέσης

Η μέση τιμή ουσιαστικά περιγράφει τη θέση του κέντρου βάρους του σχήματος που ορίζει η συνάρτηση πυκνότητας πιθανότητας με τον οριζόντιο άξονα (βλ. Σχ. 3α). Είναι επίσης ισοδύναμη με τη στατική ροπή περί τον κατακόρυφο άξονα του εν λόγω γεωμετρικού σχήματος (δεδομένου ότι το εμβαδό του σχήματος είναι ίσο με 1). Άλλες παράμετροι θέσης είναι

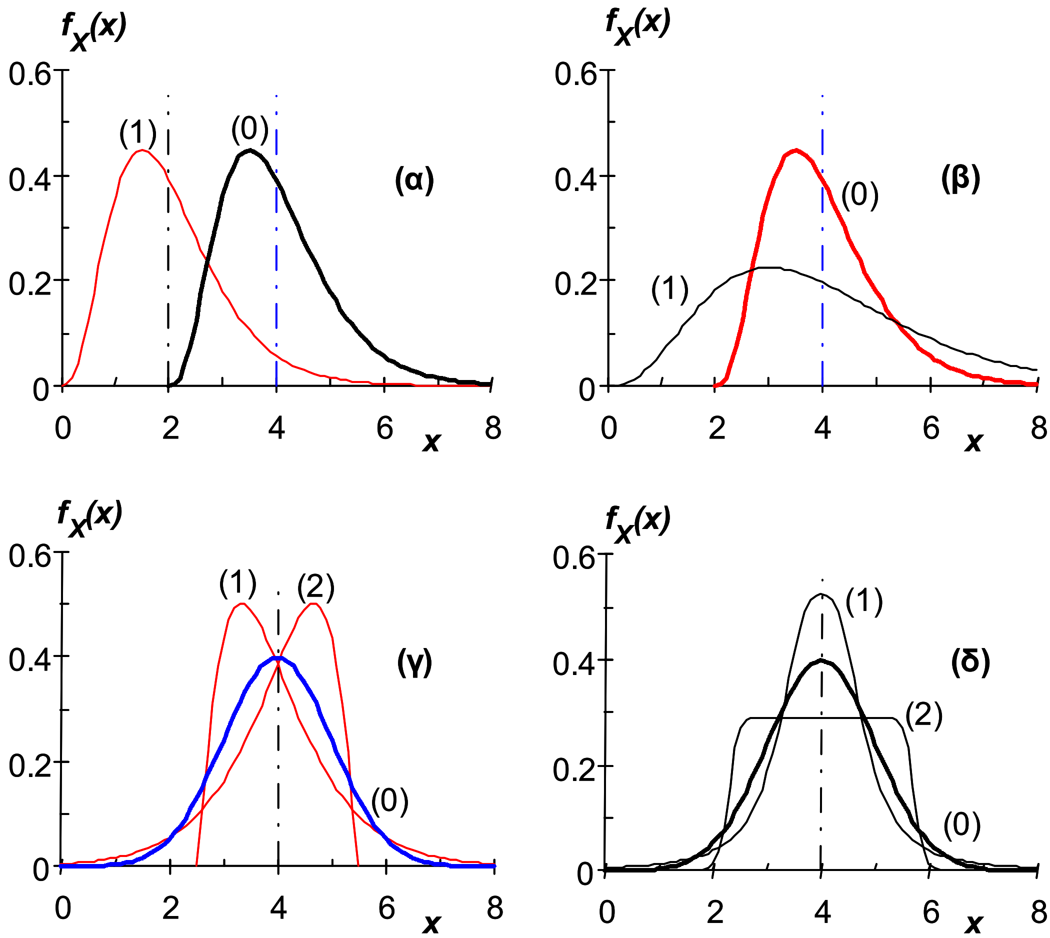
1. Η *πιθανότερη τιμή* ή *κορυφή*, συμβολικά x_p, είναι η τιμή της μεταβλητής *x* για την οποία η $f_X(x)$ γίνεται μέγιστη, προκειμένου για συνεχείς μεταβλητές, ή που έχει τη μεγαλύτερη πιθανότητα να συμβεί, αν η μεταβλητή είναι διακριτή. .

2. Η *διάμεσος*, συμβολικά x_m, είναι εκείνη η τιμή της μεταβλητής για την οποία ισχύει $P(X \le x_m) = P(X \ge x_m) = 1/2$. [*] Έτσι, η διάμεσος αντιστοιχεί στο σημείο που χωρίζει την καμπύλη πυκνότητας πιθανότητας σε δύο ισοδύναμα τμήματα με εμβαδό 1/2.

Γενικά η μέση τιμή, η κορυφή και η διάμεσος δεν ταυτίζονται, εκτός αν η κατανομή είναι συμμετρική και μονοκόρυφη.

---

[*] Αυστηρά η σχέση αυτή ισχύει μόνο για συνεχείς μεταβλητές. Κατ' αναλογία προσδιορίζεται η διάμεσος σε διακριτές μεταβλητές.



Σχήμα 3 Σχέση των χαρακτηριστικών παραμέτρων κατανομής με το σχήμα της συνάρτησης πυκνότητας πιθανότητας: (α) Επίδραση της μέσης τιμής. Οι καμπύλες (0) και (1) έχουν μέσες τιμές 4 και 2, αντίστοιχα, ενώ και οι δύο έχουν τυπική απόκλιση 1, συντελεστή ασυμμετρίας 1 και συντελεστή κύρτωσης 4.5. (β) Επίδραση της τυπικής απόκλισης. Οι καμπύλες (0) και (1) έχουν τυπική απόκλιση 1 και 2 αντίστοιχα, ενώ και οι δύο έχουν μέση τιμή 4, συντελεστή ασυμμετρίας 1 και συντελεστή κύρτωσης 4.5. (γ) Επίδραση του συντελεστή ασυμμετρίας. Οι καμπύλες (0), (1) και (2) έχουν συντελεστή ασυμμετρίας 0, +1.33 και -1.33 αντίστοιχα, ενώ και οι τρεις έχουν μέση τιμή 4 και τυπική απόκλιση 1 (ο συντελεστής κύρτωσης είναι 3 για την καμπύλη (0) και 5.67 για τις καμπύλες (1) και (2)). (δ) Επίδραση του συντελεστή κύρτωσης. Οι καμπύλες (0), (1) και (2) έχουν συντελεστή κύρτωσης 3, 5 και 2, αντίστοιχα, ενώ και οι τρεις έχουν μέση τιμή 4, τυπική απόκλιση 1 και συντελεστή ασυμμετρίας 0.

Παράμετροι διασποράς

Η διασπορά μιας τυχαίας μεταβλητής δείχνει το μέγεθος της συγκέντρωσης της πυκνότητας πιθανότητας γύρω από τη μέση τιμή. Έτσι, μικρή διασπορά δείχνει συγκεντρωμένη κατανομή (βλ. Σχ. 3β). Η οριακή τιμή μηδενικής διασποράς αντιστοιχεί σε μεταβλητή που παίρνει μία μόνο τιμή με πλήρη βεβαιότητα (αρνητική τιμή της διασποράς είναι αδύνατη). Το γεωμετρικό αντίστοιχο της διασποράς είναι η ροπή αδρανείας περί τον κατακόρυφο κεντροβαρικό άξονα του σχήματος που ορίζει η συνάρτηση πυκνότητας πιθανότητας με τον οριζόντιο άξονα.

Η τετραγωνική ρίζα της διασποράς,

$$\sigma_X = \sqrt{\mathrm{Var}[X]} \tag{18}$$

που έχει διαστάσεις ίδιες με την τυχαία μεταβλητή, λέγεται *τυπική απόκλιση*. Η αδιάστατη παράμετρος

$$C_{v_X} = \frac{\sigma_X}{m_X} \tag{19}$$

λέγεται *συντελεστής μεταβλητότητας*.

Εναλλακτική παράμετρος διασποράς, που πάντως δεν χρησιμοποιείται πολύ συχνά, είναι το κεντρικό πλάτος που ορίζεται ως η διαφορά $x_{0.75} - x_{0.25}$, δηλαδή η διαφορά των τιμών της μεταβλητής που αντιστοιχούν σε τιμές της συνάρτησης κατανομής 0.75 και 0.25. Το εμβαδό της συνάρτησης πυκνότητας πιθανότητας ανάμεσα σε αυτές τις δύο θέσεις είναι ίσο με 0.5.

Παράμετροι ασυμμετρίας

Η τρίτη κεντρική ροπή χρησιμοποιείται για την περιγραφή της ασυμμετρίας της κατανομής. Μηδενική τιμή της τρίτης ροπής δείχνει συμμετρική κατανομή. Αν η τρίτη ροπή είναι μεγαλύτερη ή μικρότερη από το μηδέν λέμε ότι η κατανομή είναι θετικά ασύμμετρη ή αρνητικά ασύμμετρη, αντίστοιχα (βλ. Σχ. 3γ). Στην περίπτωση θετικά ασύμμετρης μονοκόρυφης κατανομής ισχύει $x_p \le x_m \le m_X$ ενώ η ανάστροφη ανισότητα ισχύει για αρνητικά ασύμμετρη κατανομή. Αποτελεσματικότερο μέτρο της ασυμμετρίας της κατανομής είναι ο αδιάστατος *συντελεστής ασυμμετρίας* που ορίζεται από τη σχέση

$$C_{s_X} = \frac{\mu_X^{(3)}}{\sigma_X^3} \tag{20}$$

Παράμετροι κύρτωσης

Ο όρος κύρτωση περιγράφει το πόσο αιχμηρή ή όχι είναι η συνάρτηση πυκνότητας πιθανότητας γύρω από την αιχμή της. Μέτρο αυτής της ιδιότητας είναι η τέταρτη κεντρική ροπή, ή καλύτερα ο αδιάστατος συντελεστής κύρτωσης που ορίζεται μέσω της τελευταίας από τη σχέση

$$C_{k_X} = \frac{\mu_X^{(4)}}{\sigma_X^4} \tag{21}$$

Ο συντελεστής αυτός συνήθως συγκρίνεται με την τιμή 3, η οποία αντιστοιχεί στην κανονική κατανομή (βλ. ενότητα 5). Τιμή του συντελεστή μεγαλύτερη του 3 αντιστοιχεί σε λεπτόκυρτη (αιχμηρή) κατανομή ενώ τιμή μικρότερη του 3 αντιστοιχεί σε πλατύκυρτη (επίπεδη) κατανομή (βλ. Σχ. 3γ).

Εφαρμογή 2.2

Κατά τις μέρες που βρέχει το ημερήσιο ύψος βροχής, X, μετρημένο σε συγκεκριμένο σταθμό και εκφρασμένο σε mm βρέθηκε ότι ακολουθεί την εκθετική κατανομή με συνάρτηση κατανομής

$$F_X(x) = 1 - e^{-\lambda x}, x \ge 0$$

όπου λ= 0.05 mm^{-1}.* *Ζητείται ο υπολογισμός των παραμέτρων θέσης, διασποράς, ασυμμετρίας και κύρτωσης της κατανομής.*

Παραγωγίζοντας τη συνάρτηση κατανομής υπολογίζουμε τη συνάρτηση πυκνότητας πιθανότητας

$$f_X(x) = \lambda e^{-\lambda x}, x \ge 0$$

Προκειμένου να υπολογίσουμε τη μέση τιμή, εφαρμόζουμε την εξίσωση (12) για $g(X) = X$ και έχουμε

$$m_X = E[X] = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^{\infty} \lambda x e^{-\lambda x} dx$$

και μετά τις πράξεις

* Επειδή η τυχαία μεταβλητή στο παράδειγμα ορίζεται μόνο για τις μέρες που βρέχει, η κατανομή της είναι συνεχής παντού. Αν ορίζονταν για όλες τις μέρες, και τις μη βροχερές, τότε η τιμή $X = 0$ θα είχε συγκεκριμένη πιθανότητα διαφορετική από το μηδέν, και έτσι η συνάρτηση κατανομής θα είχε πήδημα στη θέση $x = 0$.

$$m_X = 1/\lambda = 20 \text{ mm}$$

Με παρόμοιο τρόπο βρίσκουμε για τυχόν $r \geq 0$

$$m_X^{(r)} = E\left[X^r\right] = r!/\lambda^r$$

και τελικά, εφαρμόζοντας τις εξισώσεις (17) και (15)

$$\sigma_X^2 = 1/\lambda^2 = 400 \text{ mm}^2, \; \mu_X^3 = 2/\lambda^3 = 16000 \text{ mm}^3$$

$$\mu_X^4 = 9/\lambda^4 = 1440000 \text{ mm}^4$$

Η πιθανότερη τιμή είναι προφανώς μηδέν (βλ. Σχ, 4). Η διάμεσος τιμή προσδιορίζεται από τη σχέση

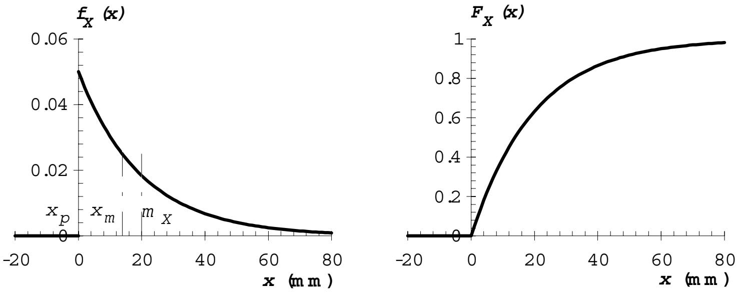
$$F_X(x_m) = 1/2 \Leftrightarrow 1 - e^{-\lambda x_m} = 1/2 \Leftrightarrow x_m = \ln 2/\lambda = 13.9 \text{ mm}$$

Παρατηρούμε ότι ισχύει η ανισοτική σχέση $x_p \leq x_m \leq m_X$ που χαρακτηρίζει τις θετικά ασύμμετρες κατανομές.

Η τυπική απόκλιση είναι σ_X = 20 mm και ο συντελεστής μεταβλητότητας $C_{v_X} = 1$. Αυτή η τιμή του συντελεστή μεταβλητότητας δείχνει πολύ μεγάλη διασπορά της κατανομής.

Ο συντελεστής ασυμμετρίας υπολογίζεται από τη σχέση (20) και είναι $C_{s_X} = \left(2/\lambda^3\right)/(1/\lambda^3) = 2$. Αυτό επιβεβαιώνει τη θετική ασυμμετρία της κατανομής, πράγμα που επιβεβαιώνεται επίσης και από το Σχήμα 4. Ειδικότερα, παρατηρούμε ότι η συνάρτηση πυκνότητας πιθανότητας του παραδείγματος έχει *σχήμα ανεστραμμένου J*, σε αντίθεση με τις συνηθέστερες κατανομές (π.χ. στο Σχ. 3) που έχουν *κωδωνοειδές σχήμα*.

Ο συντελεστής κύρτωσης υπολογίζεται από τη σχέση (21) και είναι $C_{k_X} = \left(9/\lambda^4\right)/(1/\lambda^4) = 9$. Αυτό δείχνει ότι η κατανομή είναι λεπτόκυρτη, όπως επιβεβαιώνεται και από το Σχήμα 4.



Σχήμα 4 Συνάρτηση πυκνότητας πιθανότητας και συνάρτηση κατανομής του ημερήσιου ύψους βροχής της εφαρμογής 2.4

Χρονική κλίμακα και σχήμα κατανομής

Στο παραπάνω παράδειγμα είδαμε ότι η κατανομή ενός φυσικού μεγέθους έντονα τυχαίου, όπως είναι η βροχόπτωση, σε μια μικρή χρονική κλίμακα, όπως είναι η ημερήσια, εμφανίζει μεγάλη διασπορά, έντονη θετική ασυμμετρία και σχήμα πυκνότητας πιθανότητας τύπου ανεστραμμένου J. Αυτό συμβαίνει επειδή το κύριο σώμα των τιμών του μεγέθους εμφανίζεται κοντά στο μηδέν, ενώ παράλληλα εμφανίζονται και εξαιρετικά μεγάλες (θετικές) τιμές του μεγέθους, με μικρή βέβαια πιθανότητα. Το γεγονός ότι οι αρνητικές τιμές του φυσικού μεγέθους αποκλείονται οδηγεί σε θετικά ασύμμετρες κατανομές, με μεγάλη "ουρά" προς τα δεξιά, που είναι και οι συχνότερες στην υδρολογία. Ωστόσο, καθώς προχωρούμε σε μεγαλύτερες χρονικές κλίμακες, π.χ. από την ημερήσια στην μηνιαία, η μέση τιμή του μεγέθους αυξάνει, χωρίς αντίστοιχου βαθμού αύξηση των μεγαλύτερων ροπών, πράγμα που οδηγεί σε μικρότερους συντελεστές διασποράς, ασυμμετρίας και κύρτωσης. Έτσι οι κατανομές έχουν πια κωδωνοειδές σχήμα και τείνουν να είναι συμμετρικές. Υπάρχουν σοβαροί θεωρητικοί λόγοι που οδηγούν σε αυτή τη συμπεριφορά για τις μεγάλες χρονικές κλίμακες, οι οποίοι μαθηματικά περιγράφονται από το κεντρικό οριακό θεώρημα.

2.3 Από κοινού και περιθώριες συναρτήσεις κατανομής

Στα παραπάνω εδάφια αναπτύχθηκαν οι έννοιες της θεωρίας πιθανοτήτων που αναφέρονται σε μια μεμονωμένη τυχαία μεταβλητή. Συχνά όμως ενδιαφέρει η ταυτόχρονη μελέτη δύο μεταβλητών. Έστω το ζεύγος τυχαίων μεταβλητών (X, Y) που, σύμφωνα με τα παραπάνω, είναι συναρτήσεις των δειγματικών χώρων (Ω_X, Ω_Y) αντίστοιχα. Η τομή γεγονότων $\{X \le x\} \cap \{Y \le y\} = \{X \le x, Y \le y\}$ είναι επίσης γεγονός του δειγματικού χώρου $\Omega_{XY} = \Omega_X \times \Omega_Y$. Βάσει αυτού του γεγονότος ορίζεται η *από κοινού συνάρτηση κατανομής* του ζεύγους μεταβλητών (X, Y) ως η συνάρτηση των πραγματικών μεταβλητών (x, y):

$$F_{XY}(x,y) = P(X \le x, Y \le y) \tag{22}$$

Οι δείκτες X, Y συχνά παραλείπονται, εφόσον δεν υπάρχει κίνδυνος σύγχυσης. Με την προϋπόθεση ότι η F_{XY} είναι παραγωγίσιμη, η συνάρτηση

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} \tag{23}$$

είναι η από κοινού συνάρτηση πυκνότητας πιθανότητας των μεταβλητών. Προφανώς ισχύει

$$F_{XY}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(\xi,\omega) d\omega d\xi \tag{24}$$

Αναμενόμενες τιμές - ροπές

Η αναμενόμενη τιμή ή προσδοκία της συνάρτησης $g(X, Y)$ ορίζεται από την

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y) dy dx \tag{25}$$

Το μέγεθος $E[X^p Y^q]$ ονομάζεται *από κοινού ροπή τάξης* $p + q$ των X και Y. Αντίστοιχα, το μέγεθος $E[(X - m_X)^p (Y - m_Y)^q]$ ονομάζεται *από κοινού κεντρική ροπή τάξης* $p + q$ των X και Y.

Συνδιασπορά

Από τις κεντρικές ροπές συχνότερα χρησιμοποιείται η

$$\sigma_{XY} := E[(X - m_X)(Y - m_Y)] = E[XY] - m_X m_Y \tag{26}$$

που ονομάζεται και *συνδιασπορά* των X και Y. Εναλλακτικά η συνδιασπορά συμβολίζεται ως $\mathrm{Cov}[X, Y]$.

Συντελεστής συσχέτισης

Διαιρώντας τη συνδιασπορά με τις τυπικές αποκλίσεις σ_X και σ_Y παίρνουμε το *συντελεστή συσχέτισης*

$$\rho_{XY} := \frac{\mathrm{Cov}[X, Y]}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}} \equiv \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{27}$$

ο οποίος είναι αδιάστατος και με τιμές $-1 \le \rho_{XY} \le 1$. Ο συντελεστής αυτός αποτελεί σημαντική παράμετρο για τη μελέτη της συσχέτισης μεταξύ δύο μεταβλητών.

Ανεξαρτησία μεταβλητών

Οι μεταβλητές (X, Y) λέγονται *ανεξάρτητες* αν για κάθε ζεύγος τιμών (x, y) ισχύει

$$F_{XY}(x,y) = F_X(x) F_Y(y) \tag{28}$$

Σε αυτή την περίπτωση ισχύει επίσης η

$$f_{XY}(x, y) = f_X(x) f_Y(y) \tag{29}$$

η οποία είναι ισοδύναμη με την (28), καθώς και οι ακόλουθες

$$\sigma_{XY} = 0 \Leftrightarrow \rho_{XY} = 0 \Leftrightarrow E[XY] = E[X]E[Y] \tag{30}$$

οι οποίες αποτελούν απλές συνέπειες της (28) και όχι ικανές συνθήκες για να είναι ανεξάρτητες οι μεταβλητές (X, Y). Δύο μεταβλητές (X, Y) για τις οποίες ισχύει η (30) λέγονται *ασυσχέτιστες*.

2.4 Η έννοια της στοχαστικής ανέλιξης

Η θεωρία των στοχαστικών ανελίξεων αποτελεί ιδιαίτερο κλάδο της θεωρίας πιθανοτήτων, ο οποίος μάλιστα είναι από τους πιο προχωρημένους. Ο σκοπός μας εδώ δεν είναι να δώσουμε πλήρη εικόνα αυτού του κλάδου, αλλά μόνο ορισμένες εισαγωγικές έννοιες που είναι χρήσιμες για την ορθή κατανόηση της μαθηματικής περιγραφής των υδρολογικών διεργασιών και των προϋποθέσεων στις οποίες στηρίζεται αυτή η περιγραφή. Οι έννοιες αυτές και η σύνδεση τους με την τεχνική υδρολογία θα διασαφηνιστούν στο εδάφιο 4.1.

Οι στοχαστικές ανελίξεις αποτελούν οικογένειες τυχαίων μεταβλητών, όπως αυτές του προηγούμενου εδαφίου, που, όμως, μπορεί να είναι και απειροπληθείς. Έτσι, *στοχαστική ανέλιξη* ονομάζεται μια οικογένεια μεταβλητών X_t όπου t είναι παράμετρος που παίρνει τιμές από ένα κατάλληλο σύνολο T (Papoulis, 1965, Taylor and Karlin, 1984). Εναλλακτικά χρησιμοποιείται και ο συμβολισμός $X(t)$ αντί του X_t. Αν και γενικά το *δεικτοσύνολο* T μπορεί να είναι οποιοδήποτε σύνολο, συχνότατα παριστάνει χρόνο. Σε περίπτωση που το δεικτοσύνολο αντιστοιχεί σε διακριτές μονάδες χρόνου, $T = \{0, 1, 2, \ldots\}$, έχουμε μια *ανέλιξη σε διακριτό χρόνο*. Αντίστοιχα, αν το δεικτοσύνολο αντιστοιχεί σε συνεχή χρόνο, δηλαδή $T = [0, \infty)$, μιλούμε για *ανέλιξη σε συνεχή χρόνο*.

Χρονοσειρά

Μια υλοποίηση της στοχαστικής ανέλιξης, δηλαδή ένα σύνολο παρατηρήσεων $x(t)$ της $X(t)$, για μεταβαλλόμενο χρόνο t, ονομάζεται *δειγματοσυνάρτηση* ή *χρονοσειρά* της ανέλιξης.

Ροπές - αυτοσυνδιασπορά - αυτοσυσχέτιση

Οι ροπές ορίζονται και εδώ με τον ίδιο τρόπο όπως στις προηγούμενες ενότητες. Ιδιαίτερο ενδιαφέρον παρουσιάζουν:

1. Η *μέση τιμή της ανέλιξης*, δηλαδή η αναμενόμενη τιμή της μεταβλητής $X(t)$:

$$\mu(t) = E[X(t)] = \int_{-\infty}^{\infty} x f(x; t) dt \tag{31}$$

2. Η *αυτοσυνδιασπορά* της ανέλιξης, δηλαδή η συνδιασπορά των τυχαίων μεταβλητών $X(t_1)$ και $X(t_2)$:

$$C(t_1, t_2) = \mathrm{Cov}[X(t_1), X(t_2)] = E[(X(t_1) - \mu(t_1))(X(t_2) - \mu(t_2))] \tag{32}$$

Σημειώνεται ότι η διασπορά της ανέλιξης, δηλαδή η διασπορά της μεταβλητής $X(t)$, είναι $\mathrm{Var}[X(t)] = C(t, t)$. Κατά συνέπεια ο *συντελεστής αυτοσυσχέτισης*, δηλαδή ο συντελεστής συσχέτισης των τυχαίων μεταβλητών $X(t_1)$ και $X(t_2)$ ορίζεται από τη σχέση

$$\rho(t_1, t_2) = \frac{\mathrm{Cov}[X(t_1), X(t_2)]}{\sqrt{Var[X(t_1)] Var[X(t_2)]}} = \frac{C(t_1, t_2)}{\sqrt{C(t_1, t_1) C(t_2, t_2)}} \tag{33}$$

Στασιμότητα

Στη γενική περίπτωση οι στατιστικές παράμετροι μιας στοχαστικής ανέλιξης, π.χ. η μέση τιμή και η διασπορά της, είναι συναρτήσεις του χρόνου και κατά συνέπεια μεταβάλλονται με τη μεταβολή του χρόνου. Μια ειδική κατηγορία ανελίξεων, σαφώς απλούστερων στη μελέτη, είναι οι στάσιμες ανελίξεις στις οποίες δεν υπάρχει μεταβολή των στατιστικών χαρακτηριστικών με την πάροδο του χρόνο. Έτσι, μια στοχαστική ανέλιξη λέγεται *στάσιμη με την αυστηρή έννοια*, ή απλώς *στάσιμη*, όταν η συνάρτηση κατανομής της δεν επηρεάζεται από τη μετατόπιση του χρόνου, δηλαδή αν, για τυχούσα χρονική μετατόπιση τ, η συνάρτηση κατανομής οποιασδήποτε τάξης της $X(t + \tau)$ ταυτίζεται με τη συνάρτηση κατανομής της ίδιας τάξης της $X(\tau)$. Λέγεται δε *στάσιμη με την ευρεία* (ή *ελαστική*) *έννοια* αν η μέση τιμή της είναι σταθερή και η αυτοσυνδιασπορά της εξαρτάται μόνο από τη διαφορά του χρόνου, δηλαδή αν

$$E[X(t)] = \mu = \text{σταθερό} \qquad [(X(t)-\mu)(X(t+\tau)-\mu)] = C(\tau) \tag{34}$$

Εργοδικότητα

Η έννοια της εργοδικότητας μιας στοχαστικής ανέλιξης έχει σχέση με το πρόβλημα του προσδιορισμού της κατανομής της από μια απλή σειρά παρατηρήσεών της. Έτσι μια στάσιμη στοχαστική ανέλιξη είναι *εργοδική* αν κάθε παράμετρος της κατανομής μπορεί να προσδιοριστεί από μια απλή δειγματοσυνάρτηση της ανέλιξης. Δεδομένου ότι οι παράμετροι υπολογίζονται ως χρονικές μέσες τιμές, ο παραπάνω ορισμός εκφράζεται και με τον εξής τρόπο: Μια ανέλιξη είναι εργοδική αν οι χρονικοί μέσοι είναι ίσοι με τους συνολικούς μέσους (δηλαδή τις αναμενόμενες τιμές). Για παράδειγμα, μια ανέλιξη είναι εργοδική ως προς τη μέση τιμή αν

$$E[X(t)] = \lim_{N\to\infty} \frac{1}{N}\sum_{t=0}^{N} X(t) \qquad \text{(για ανέλιξη διακριτού χρόνου)}$$

$$E[X(t)] = \lim_{T\to\infty} \frac{1}{T}\int_{0}^{T} X(t)dt \qquad \text{(για ανέλιξη συνεχούς χρόνου)} \tag{35}$$

Το αριστερό μέλος των παραπάνω παριστάνει το συνολικό μέσο, δηλαδή την αναμενόμενη τιμή, ενώ το δεξιό μέλος παριστάνει το χρονικό μέσο, στην οριακή περίπτωση άπειρου χρόνου. Ας σημειωθεί ότι ενώ το αριστερό μέλος είναι παράμετρος και όχι τυχαία μεταβλητή, το δεξιό μέλος είναι τυχαία μεταβλητή (ως άθροισμα ή ολοκλήρωμα τυχαίων μεταβλητών). Η εξίσωση λοιπόν μιας παραμέτρου με μια τυχαία μεταβλητή υπονοεί ότι η τυχαία μεταβλητή έχει μηδενική διασπορά. Αυτή ακριβώς είναι και η προϋπόθεση που κάνει την ανέλιξη εργοδική, προϋπόθεση που, ας σημειωθεί, δεν ισχύει απαραίτητα για κάθε στοχαστική ανέλιξη.

3 Γενικές έννοιες στατιστικής (επιγραμματικά)

Στατιστικός πληθυσμός

Στη στατιστική ο όρος *πληθυσμός* αναφέρεται σε κάθε συλλογή αντικειμένων, των οποίων ενδιαφέρουν οι μετρικές ιδιότητες. Οι πληθυσμοί μπορεί να είναι πεπερασμένοι (και συγκεκριμένοι), όπως για παράδειγμα ο πληθυσμός όλων των ποταμών της Ελλάδας, ή άπειροι (και, κατά κανόνα, αφαιρετικά ορισμένοι). Στην στατιστική υδρολογία, συνήθως οι πληθυσμοί που ενδιαφέρουν ανήκουν στη δεύτερη κατηγορία, πχ. ο πληθυσμός των δυνατών παρατηρήσεων της παροχής ενός ποταμού.

Στατιστικό δείγμα

Είναι ένα πεπερασμένο υποσύνολο του πληθυσμού, για παράδειγμα οι παροχές που μετρήθηκαν σε ένα ποταμό μια δεδομένη περίοδο. Το δείγμα χαρακτηρίζεται τυχαίο όταν κάθε στοιχείο του πληθυσμού έχει την ίδια πιθανότητα να συμπεριληφθεί σε αυτό.

Τα βασικά προβλήματα της στατιστικής

Τα βασικά προβλήματα της στατιστικής είναι τα ακόλουθα δύο:

1. η ποσοτική περιγραφή του πληθυσμού, δηλαδή η εξαγωγή ποσοτικών συμπερασμάτων για τον πληθυσμό από το διαθέσιμο δείγμα, και
2. η ποσοτικοποίηση της αβεβαιότητας, δηλαδή η ποσοτικοποίηση της πληροφορίας την οποία παρέχει το δείγμα για τον πληθυσμό, σε όρους βαθμού αβεβαιότητας.

Εκτιμήτριες και εκτιμήσεις

Ένας μαθηματικός τύπος ή, γενικότερα, μια διαδικασία για την περιγραφή των ιδιοτήτων ενός πληθυσμού ονομάζεται *εκτιμήτρια*. Το αριθμητικό αποτέλεσμα της διαδικασίας ονομάζεται *εκτίμηση*.

Εκτιμήτριες τυπικών στατιστικών παραμέτρων

Με δεδομένο ένα δείγμα που περιλαμβάνει n δεδομένα, καθένα από τα οποία περιγράφεται ποσοτικά από την τυχαία μεταβλητή X_i (i =1, ..., n), οι ακόλουθες εκτιμήτριες εφαρμόζονται προκειμένου να εκτιμηθούν οι βασικές στατιστικές παράμετροι του πληθυσμού

Μέση τιμή

$$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n} \tag{36}$$

Διασπορά

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \frac{1}{n-1} \sum_{i=1}^{n} X_i^2 - \frac{n}{n-1} \overline{X}^2 \tag{37}$$

Κεντρική ροπή τάξης k

$$\hat{M}^{(k)} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^k \tag{38}$$

4 Ειδικές έννοιες θεωρίας πιθανοτήτων στην υδρολογία

4.1 Πιθανοθεωρητική περιγραφή υδρολογικών διεργασιών

Από την οπτική γωνία της θεωρίας πιθανοτήτων οι υδρολογικές διεργασίες είναι στοχαστικές ανελίξεις. Ας πάρουμε για παράδειγμα την παροχή σε μια συγκεκριμένη θέση ενός ποταμού. Σε κάθε χρονική στιγμή t μέσα στο συνεχή χρόνο η παροχή $X(t)$ είναι μια τυχαία μεταβλητή, με την έννοια ότι δεν υπάρχει προσδιοριστική μέθοδος καθορισμού της τιμής της παροχής με πλήρη βεβαιότητα. Έτσι λοιπόν η $X(t)$, σύμφωνα με τον ορισμό που δόθηκε παραπάνω, είναι στοχαστική ανέλιξη σε συνεχή χρόνο, ενώ μια σειρά μετρήσεών της σε τακτά χρονικά διαστήματα αποτελεί μια χρονοσειρά.

Στο σημείο αυτό θα πρέπει να δώσουμε ορισμένες διευκρινήσεις προκειμένου να αποφευχθούν παρανοήσεις στις οποίες είναι εύκολο να οδηγηθεί κανείς από την εισαγωγή της έννοιας της στοχαστικής ανέλιξης για την περιγραφή μιας φυσικής διεργασίας.

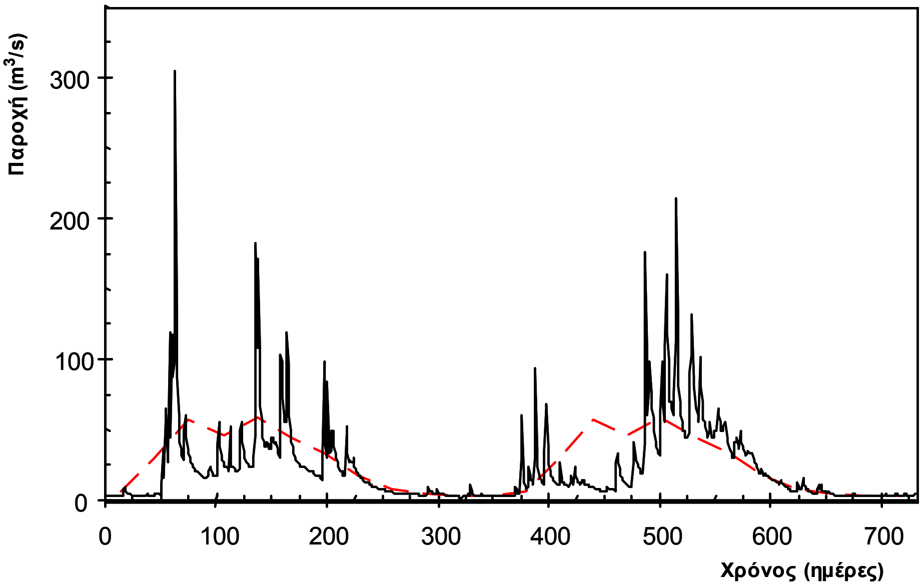
Κατ΄ αρχήν, το γεγονός ότι μια φυσική φυσική διεργασία περιγράφεται από μια στοχαστική ανέλιξη δεν σημαίνει ότι η πρώτη δεν υπακούει σε κανενός είδους αιτιοκρατία. Αντίθετα, είναι γνωστό ότι τα υδρολογικά μεγέθη εμφανίζουν περιοδικές διακυμάνσεις μέσα στη διάρκεια ενός έτους, οι οποίες προφανώς οφείλονται στην ετήσια περιστροφή του άξονα της γης και στα μετεωρολογικά φαινόμενα που αυτή επισύρει. Αυτές οι περιοδικές διακυμάνσεις αποτελούν την ονομαζόμενη *προσδιοριστική συνιστώσα* των διεργασιών. Η στοχαστική ανέλιξη $X(t)$ δεν έχει δυσκολία να συμπεριλάβει και να περιγράψει μαθηματικά και αυτή τη συνιστώσα. Συχνά μάλιστα γίνεται διάκριση των δύο συνιστωσών θεωρώντας ότι

$$X(t) = D(t) + \Xi(t) \tag{39}$$

όπου $D(t)$ είναι η προσδιοριστική (εν προκειμένω περιοδική) συνιστώσα και $\Xi(t)$ είναι η στοχαστική συνιστώσα. Στο Σχήμα 5 παρουσιάζεται η μεταβολή της παροχής ενός ποταμού στη διάρκεια δύο ετών, όπου είναι εμφανής η ετήσια περιοδικότητα.

Σε πολλά κείμενα υδρολογίας (π.χ. Haan, 1977, Kottegoda, 1980) στην προσδιοριστική συνιστώσα συγκαταλέγονται ακόμη οι *τάσεις*, δηλαδή οι μακροχρόνιες, πολυετείς, βαθμιαίες μεταβολές των μέσων χαρακτηριστικών της $X(t)$, καθώς και τα *άλματα*, δηλαδή οι απότομες αλλαγές στα μέσα χαρακτηριστικά της $X(t)$, οι οποίες εν συνεχεία διατηρούνται επίσης για μεγάλα χρονικά διαστήματα. Ωστόσο, εκτός από σπάνιες περιπτώσεις, οι μηχανισμοί που προκαλούν αυτές τις τάσεις και τα άλματα δεν είναι γνωστοί. Κατά συνέπεια, δεν είναι προβλέψιμος ο τόπος με τον οποίο θα εξελιχθούν αυτά τα φαινόμενα, αντίθετα με τα φαινόμενα που οφείλονται στην περιοδικότητα της γης, τα οποία είναι βέβαια προβλέψιμα. Με αυτή την έννοια, στο κείμενο αυτό οι τάσεις και τα άλματα θεωρούνται κατά κανόνα ως τυχαίες διακυμάνσεις, παρά ως προσδιοριστικές, οι οποίες όμως συντελούνται σε πολύ μεγαλύτερη χρονική κλίμακα (π.χ. δεκαετιών, αιώνων κοκ) από την κλίμακα των συνήθων τυχαίων διακυμάνσεων

Ειδικότερα, το στοχαστικό μέρος της ανέλιξης δεν είναι πλήρως τυχαίο, αλλά, όπως λέμε, έχει *στοχαστική δομή* ή *μνήμη*. Αυτό σημαίνει ότι υπάρχει στοχαστική εξάρτηση των τιμών της ανέλιξης σε γειτονικές χρονικές στιγμές, ή διαφορετικά, ισχυρός συντελεστής αυτοσυσχέτισης της ανέλιξης. Για παράδειγμα, αν η τιμή της παροχής ενός ποταμού στο χρόνο t_0 είναι $X(t_0)$ = 500 m^3/s, είναι απίθανο να είναι μετά από ένα μικρό χρονικό διάστημα Δt π.χ. 1 ώρα $X(t_0 + \Delta t)$ = 0.5 m^3/s. Το πιθανότερο είναι ότι θα κυμανθεί γύρω από μια τιμή κοντά στα 500 m^3/s.



Σχήμα 5 Διακύμανση της ημερήσιας παροχής του ποταμού Ευήνου στη θέση Πόρος Ρηγανίου (υδρολογικά έτη 1971-72 και 1972-73 - έναρξη χρόνου 1.10.1971). Με διακεκομμένη γραμμή φαίνονται οι υπερετήσιες μέσες τιμές για κάθε μήνα υπολογισμένες από το ιστορικό δείγμα 1970-71 μέχρι 1989-90.

Ανελίξεις σε διακριτό χρόνο

Η μελέτη της πλήρους ανέλιξης σε συνεχή χρόνο μιας υδρολογικής μεταβλητής $X(t)$ είναι αρκετά δύσκολο πρόβλημα, αλλά και δεν είναι απαραίτητη για τα περισσότερα από τα πρακτικά προβλήματα του μηχανικού. Για το λόγο αυτό ξεφεύγει από τους στόχους αυτού του κειμένου. Θα εντοπίσουμε την προσοχή μας σε ορισμένες παράγωγες ανελίξεις, οι οποίες αναφέρονται σε διακριτό χρόνο και αναλύονται ευκολότερα. Πριν προχωρήσουμε στον ορισμό αυτών των ανελίξεων, ας ορίσουμε δύο χρονικά μεγέθη. Το πρώτο είναι το *υδρολογικό έτος*, μέσα στο οποίο πραγματοποιείται ένας πλήρης κύκλος περιοδικών υδρολογικών διακυμάνσεων. Η διάρκειά του D ταυτίζεται με τη διάρκεια του ημερολογιακού έτους, συνήθως όμως ξεκινάει όχι στις 1 Ιανουαρίου, αλλά στην αρχή της βροχερής περιόδου του έτους. Έτσι, στην Ελλάδα, κατά σύμβαση, το υδρολογικό έτος ξεκινάει στις 1 Οκτωβρίου. Το δεύτερο είναι το *χρονικό βήμα* Δ, δηλαδή το χρονικό παράθυρο μέσα από το οποίο βλέπουμε την ανέλιξη. Σε αντίθεση με το υδρολογικό έτος, το χρονικό βήμα δεν είναι σταθερό, αλλά εξαρτάται από τη συγκεκριμένο πρόβλημα που αντιμετωπίζουμε. Συνήθως κυμαίνεται από μερικά λεπτά της ώρας μέχρι ένα υδρολογικό έτος. Βοηθητικό για τους ορισμούς που θα ακολουθήσουν είναι το Σχήμα 6 στο οποίο ο συνεχής χρόνος t (στον οριζόντιο άξονα) μετριέται σε υδρολογικά έτη, ενώ για σχεδιαστικούς λόγους έχει θεωρηθεί $\Delta = D/4$.

Η πρώτη απλοποίηση της πλήρους ανέλιξης $X(t)$ (Σχ. 6(1)) αφορά στη διακριτοποίηση του χρόνου (Σχ. 6(2)). Το διακριτό χρόνο k τον ορίζουμε χωρίζοντας το συνεχή χρόνο σε διαστήματα μεγέθους Δ. Έτσι οι τιμές $k = 1, 2, \dots$, αντιστοιχούν στα χρονικά διαστήματα $[0, \Delta)$, $[\Delta, 2\Delta)$, κοκ. Ως τιμή της νέας ανέλιξης $X_\Delta(k)$ στη χρονική θέση k ορίζουμε το χρονικό μέσο της $X(t)$ στο αντίστοιχο διάστημα, ήτοι

$$X_\Delta(k) = \frac{1}{\Delta} \int_{(k-1)\Delta}^{k\Delta} X(t)\,dt \tag{40}$$

Αν για παράδειγμα η $X(t)$ παριστάνει τη στιγμιαία παροχή ποταμού, και το Δ ληφθεί μία ημέρα ή ένας μήνας, τότε η $X_\Delta(k)$ παριστάνει την ημερήσια (ακριβέστερα: τη μέση ημερήσια) ή τη μηνιαία (ακριβέστερα: τη μέση μηνιαία) παροχή, αντίστοιχα. Μερικές φορές ως παράγωγο μέγεθος παίρνουμε όχι το χρονικό μέσο αλλά το χρονικό άθροισμα στο αντίστοιχο διάστημα Δ δηλαδή το μέγεθος

$$X^*_\Delta(k) = \int_{(k-1)\Delta}^{k\Delta} X(t)\,dt \tag{41}$$

Στο παραπάνω παράδειγμα το μέγεθος $X^*_\Delta(k)$ παριστάνει τον ημερήσιο ή το μηνιαίο όγκο απορροής. Αντίστοιχα αν η $X(t)$ παριστάνει τη στιγμιαία ένταση βροχής σε δεδομένο σημείο μιας λεκάνης απορροής, και το Δ ληφθεί μία ημέρα ή ένας μήνας, τότε η $X^*_\Delta(k)$ παριστάνει το ημερήσιο ή το μηνιαίο ύψος βροχής, αντίστοιχα.

Αν και η διακριτοποίηση του χρόνου είναι ένα βήμα προς την απλοποίηση της μελέτης μιας υδρολογικής διεργασίας, η μαθηματική περιγραφή της $X_\Delta(k)$ ή της $X^*_\Delta(k)$ είναι ακόμη αρκετά πολύπλοκη, γιατί προϋποθέτει την ανάλυση της περιοδικότητας και της ισχυρής αυτοσυσχέτισης της ανέλιξης. Έτσι, η μελέτη και αυτού του τύπου των ανελίξεων δεν καλύπτεται από αυτό το εισαγωγικό εγχειρίδιο τεχνικής υδρολογίας και ο ενδιαφερόμενος αναγνώστης παραπέμπεται στους Haan (1977), Kottegoda (1980) Bras and Rodriguez-Iturbe (1985) και Salas et al. (1988). Οι επόμενες απλοποιήσεις είναι αρκετά δραστικότερες και οι ανελίξεις που προκύπτουν είναι πολύ ευκολότερο να μελετηθούν αλλά και είναι πολύ χρησιμότερες στα πρακτικά προβλήματα του μηχανικού.

Αν στην ανέλιξη $X_\Delta(k)$ (ή την $X^*_\Delta(k)$) θεωρήσουμε χρονικό βήμα ίσο με ένα υδρολογικό έτος ($\Delta = D$) τότε παίρνουμε την *ετήσια ανέλιξη* συμβολικά $X_D(\tau)$, όπου χρησιμοποιήσαμε ως σύμβολο του διακριτού χρόνου, ο οποίος πλέον εκφράζεται σε ακέραιες μονάδες υδρολογικού έτους, το τ αντί του k (Σχ. 6(3)). Έτσι έχουμε

$$X_D(\tau) = \frac{1}{D} \int_{(\tau-1)\Delta}^{\tau\Delta} X(t)\,dt, X^*_D(\tau) = \int_{(\tau-1)\Delta}^{\tau\Delta} X(t)\,dt \tag{42}$$

Σε αυτή την ανέλιξη πλέον έχει εξαλειφθεί τελείως η περιοδικότητα, δεδομένου ότι δεν είναι "ορατά" χρονικά διαστήματα μικρότερα του έτους, και έχει περιοριστεί η αυτοσυσχέτιση σε μεγάλο βαθμό, λόγω του μεγάλου χρονικού βήματος ολοκλήρωσης. Εξ άλλου η ανέλιξη αυτή, η οποία που εκφράζει τη διαδοχή στα διάφορα υδρολογικά έτη μιας συνολικής ετήσιας υδρολογικής ποσότητας, είναι επαρκής για προβλήματα εκτίμησης υδατικού δυναμικού.

Ένας τρόπος για να πάμε σε μικρότερο από το ετήσιο χρονικό βήμα, εξαλείφοντας παράλληλα την περιοδικότητα και την αυτοσυσχέτιση, φαίνεται στο Σχήμα 6(4). Σε κάθε υδρολογικό έτος $\tau = 1, 2, \ldots,$ παίρνουμε ένα μόνο διάστημα μεγέθους Δ, και συγκεκριμένα το $[(\tau - 1)D + (k - 1)\Delta, (\tau - 1)D + k\Delta)$. Το k εδώ είναι ένας δεδομένος ακέραιος με δυνατές τιμές k = 1, 2, …, D/Δ (στο Σχ. 6(4) έχει ληφθεί k = 1). Η προκύπτουσα ανέλιξη είναι:

$$Y_\Delta(\tau) = \frac{1}{\Delta} \int_{(\tau-1)D+(k-1)\Delta}^{(\tau-1)D+k\Delta} X(t)dt, Y^*_\Delta(\tau) = \int_{(\tau-1)D+(k-1)\Delta}^{(\tau-1)D+k\Delta} X(t)dt \quad (43)$$

Για παράδειγμα αν η η $X(t)$ παριστάνει τη στιγμιαία παροχή ποταμού, το Δ ληφθεί ένας μήνας και το k = 1, τότε η $Y_\Delta(\tau)$ είναι η σειρά των (μέσων) μηνιαίων παροχών του Οκτωβρίου κάθε έτους και η $Y^*_\Delta(\tau)$ είναι η σειρά των αντίστοιχων όγκων απορροής.

Ανελίξεις ακροτάτων

Σε πολλά προβλήματα αντί των μέσων ή των αθροιστικών μεγεθών δεδομένης διάρκειας ενδιαφέρουν τα ακρότατα μεγέθη, δηλαδή τα μέγιστα (για τη μελέτη πλημμυρών) ή τα ελάχιστα (για τη μελέτη ξηρασιών). Για τη μελέτη αυτών των μεγεθών σχηματίζουμε αντίστοιχες ανελίξεις. Έτσι, στο Σχήμα 6(5) έχει σχηματιστεί η ανέλιξη των *στιγμιαίων ετήσιων μεγίστων* $Z_0(\tau)$. Σε κάθε υδρολογικό έτος πήραμε μόνο μια τιμή, τη στιγμιαία μέγιστη τιμή που εμφανίζεται κατά τη διάρκεια όλου του υδρολογικού έτους, δηλαδή

$$Z_0(\tau) = \max_{\tau-1 \le t < \tau} \{X(t)\} \quad (44)$$

Με αντίστοιχο τρόπο σχηματίζεται και η ανέλιξη των *στιγμιαίων ετήσιων ελαχίστων*. Αυτές οι ανελίξεις δεν εμφανίζουν περιοδικότητα (μια τιμή μόνο ανά έτος) ούτε αυτοσυσχέτιση (απομακρυσμένες χρονικά τιμές που προκύπτουν από τελείως διαφορετικά υδρομετεωρολογικά φαινόμενα) και γι' αυτό είναι εύκολο να μελετηθούν.

Αν ενδιαφέρει η μέγιστη τιμή του μεγέθους μέσα σε μια δεδομένη διάρκεια Δ, τότε ορίζουμε και μελετούμε αντίστοιχα τις *ανελίξεις ετήσιων μεγίστων δεδομένης διάρκειας*, ήτοι (Σχ. 6(6))

$$Z_D(\tau) = \max_{\tau-1 \le s < \tau-\Delta} \left\{ \frac{1}{\Delta} \int_s^{s+\Delta} X(t)dt \right\}, Z^*_D(\tau) = \max_{\tau-1 \le s < \tau-\Delta} \left\{ \int_s^{s+\Delta} X(t)dt \right\} \quad (45)$$

Σημειώνεται ότι ο ορισμός των ανελίξεων αυτών βασίστηκε στη ανέλιξη συνεχούς χρόνου $X(t)$. Εναλλακτικά, αλλά με μικρότερη ακρίβεια, μπορούν να οριστούν από τις ανελίξεις διακριτού χρόνου (Σχ. 6(7)) με τις σχέσεις

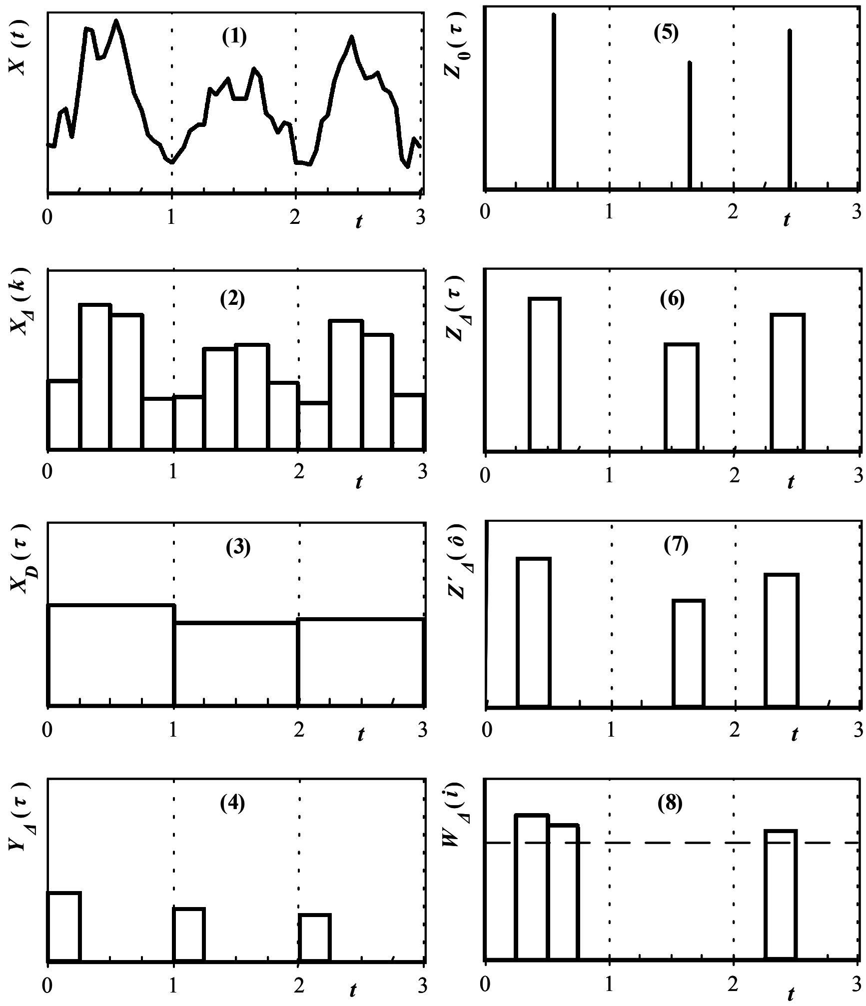
$$\alpha'_\Delta(\tau) = \max_{k_1 \le k \le k_2} \{X_\Delta(k)\}, \alpha'^*_\Delta(\tau) = \max_{k_1 \le k \le k_2} \{X^*_\Delta(k)\} \quad (46)$$

όπου $k_1 := (\tau - 1)D / \Delta + 1$, $k_2 := \tau D / \Delta$. Στα Σχήματα 6(6) και 6(7) είναι εμφανές ότι οι οι μεταβλητές $Z_\Delta(\tau)$ και $Z'_\Delta(\tau)$ δεν ταυτίζονται ούτε ως προς το μέγεθος ούτε ως προς τη χρονική τοποθέτηση, χωρίς πάντως να διαφέρουν πολύ. Με ανάλογο τρόπο σχηματίζονται και οι ανελίξεις των *ετήσιων ελαχίστων δεδομένης διάρκειας*. Στα προβλήματα πλημμυρών και ξηρασιών οι τυπικές τιμές του χρονικού βήματος Δ κυμαίνονται από μερικά λεπτά (π.χ. στις καταιγίδες σχεδιασμού δικτύων αποχέτευσης ομβρίων) μέχρι μερικές μέρες ή ακόμη και ένα μήνα (π.χ. στη μελέτη ποιότητας υδάτων ποταμού σε συνθήκες ξηρασίας).

Μια τελευταία σειρά μεγίστων που ονομάζεται *σειρά μερικής διάρκειας* φαίνεται στο Σχήμα 6(8). Η σειρά αυτή, όπως και αυτή του Σχήματος 6(7), προκύπτει από την ανέλιξη διακριτού χρόνου $X_\Delta(k)$. Αντί όμως να πάρουμε τη μέγιστη τιμή κάθε υδρολογικού έτους, σχηματίζουμε τη σειρά όλων των τιμών που υπερβαίνουν ένα κατώφλι c, ανεξάρτητα από τη θέση των τιμών αυτών στα διάφορα υδρολογικά έτη, δηλαδή

$$\{W_\Delta(i), i = 1,2,\ldots\} = \{X_\Delta(k) : X_\Delta(k) \geq c, k = 1,2,\ldots\} \tag{47}$$

(Αντίστοιχα ορίζεται και η ανέλιξη $W^*_\Delta(i)$ από την $X^*_\Delta(k)$). Εδώ η μεταβλητή i που επέχει θέση χρόνου στην πραγματικότητα αντιπροσωπεύει απλώς τον αύξοντα αριθμό που έχει η κάθε τιμή στη σειρά των χρονικά διαδοχικών τιμών. Το κατώφλι c συνήθως επιλέγεται έτσι ώστε *κατά μέσο όρο* να αντιστοιχεί μια τιμή μεγαλύτερη απο το κατώφλι ανά έτος. Έτσι, στο Σχήμα 6(8) το κατώφλι που αντιπροσωπεύεται από την οριζόντια διακεκομμένη γραμμή έχει επιλεγεί έτσι ώστε να υπάρχουν τρεις τιμές σε σύνολο τριών υδρολογικών ετών. Στο συγκεκριμένο παράδειγμα βλέπουμε ότι στο πρώτο υδρολογικό έτος προκύπτουν δύο τιμές πάνω από το κατώφλι, ενώ στο δεύτερο καμιά. Το γεγονός ότι στις σειρές μερικής διάρκειας μπορεί να εμφανίζονται τιμές που αντιστοιχούν σε γειτονικές θέσεις του πραγματικού χρόνου (όπως οι δύο τιμές του πρώτου υδρολογικού έτους του παραδείγματος) ενδέχεται να εισάγει μη αμελητέα στοχαστική εξάρτηση στις διαδοχικές τιμές της ανέλιξης. Αν είναι επιθυμητή η κατασκευή σειράς ανεξάρτητων τιμών θα πρέπει είτε να τεθεί και ένα κατώφλι ελάχιστης χρονικής απόστασης διαδοχικών τιμών, είτε να χρησιμοποιηθούν άλλες εμπειρικές μέθοδοι (βλ. Kottegoda, 1980, σ. 247).



Σχήμα 6 Βοηθητικό σκαρίφημα για τον ορισμό των διάφορων τύπων ανελίξεων (βλ. περιγραφή στο κείμενο).

Βασικές απλουστευτικές παραδοχές

Πιο πάνω διευκρινήσαμε ότι στο εισαγωγικό αυτό κείμενο δεν καλύπτεται ούτε η μελέτη των ανελίξεων συνεχούς χρόνου, ούτε των πλήρων ανελίξεων διακριτού χρόνου. Ορίσαμε, ωστόσο, άλλους έξι τύπους ανελίξεων στις οποίες ο “χρόνος” είναι διακριτός και δεν ταυτίζεται κατ' ανάγκη με τον πραγματικό χρόνο. Επίσης, καθεμιά από τις ανελίξεις περιλαμβάνει ακριβώς μία τιμή ανά έτος, εκτός από την ανέλιξη μερικής διάρκειας που περιλαμβάνει κατά μέσο μία τιμή ανά έτος. Για τη μελέτη μας θα κάνουμε από τώρα τις ακόλουθες παραδοχές:

1. Οι ανελίξεις είναι στάσιμες. Θεωρούμε ότι η κατανομή κάθε μεταβλητής παραμένει ίδια από έτος σε έτος.
2. Οι ανελίξεις είναι εργοδικές: Θεωρούμε ότι οι αναμενόμενες τιμές είναι ίσες με τους χρονικούς μέσους.
3. Οι μεταβλητές που αντιστοιχούν σε διαφορετικούς χρόνους είναι στοχαστικά ανεξάρτητες.

Για να διαλευκάνουμε το νόημα αυτών των παραδοχών δίνουμε ένα παράδειγμα. Έστω ότι η $X(t)$ αντιπροσωπεύει τη στιγμιαία παροχή στο χρόνο t, οπότε σύμφωνα με τα παραπάνω η $X_D(\tau)$ είναι η μέση ετήσια παροχή του υδρολογικού έτους τ. Έστω ότι διαθέτουμε μετρήσεις 30 ετών, ξέρουμε δηλαδή τις τιμές $x_D(1)$, …, $x_D(30)$, των τυχαίων μεταβλητών $X_D(1)$, …, $X_D(30)$. Είναι βέβαια προφανές, αλλά ωστόσο τονίζουμε ότι για κάθε μεταβλητή μπορούμε να έχουμε μόνο μια τιμή (δεν μπορούμε να κάνουμε πολλαπλά πειράματα τύχης με διαφορετικές εκβάσεις ώστε να σχηματίσουμε ένα δείγμα πολλαπλών τιμών για την ίδια μεταβλητή, π.χ. την ετήσια παροχή του υδρολογικού έτους με αριθμό 25). Τίθεται τώρα το ερώτημα αν με αυτά τα δεδομένα μπορούμε να εκτιμήσουμε τη μέση τιμή της παροχής του επόμενου υδρολογικού έτους, του υπ' αριθμόν 31. Η απάντηση είναι αρνητική αν δεν ισχύει η πρώτη από τις παραπάνω παραδοχές. Αν η κατανομή της ετήσιας παροχής είναι διαφορετική κάθε έτος, τότε δεν μπορούμε να συνδέσουμε τα προηγούμενα χρόνια με το τρέχον.* Αλλά εξακολουθεί να είναι αρνητική και αν ακόμη ισχύει η πρώτη παραδοχή αλλά όχι η δεύτερη. Στην πραγματικότητα, αυτό που μπορούμε να υπολογίσουμε από τα διαθέσιμα δεδομένα είναι ο χρονικός μέσος όρος, δηλαδή η μέση υπερετήσια παροχή των προηγούμενων τριάντα ετών. Η εργοδικότητα μας εξασφαλίζει ότι αυτός ο χρονικός μέσος τείνει προς τη θεωρητική μέση τιμή της κάθε μεταβλητής. Έτσι οι δύο πρώτες παραδοχές είναι θεμελιώδεις. Η τρίτη παραδοχή, της ανεξαρτησίας, είναι λιγότερο σημαντική, αλλά βοηθά πολύ σε περιπτώσεις που μας ενδιαφέρει η πιθανότητα διαδοχικών γεγονότων, όπως στο παράδειγμα που ακολουθεί στο τέλος αυτού του εδαφίου.

Τονίζουμε ότι η πρώτη από τις πιο πάνω παραδοχές αποτελεί απλώς μια προσέγγιση, δεδομένου ότι τα υδρομετεωρολογικά φαινόμενα δεν περιγράφονται με αυστηρώς στάσιμες ανελίξεις. Η ορθότητα της δεύτερης παραδοχής συναρτάται σε μεγάλο βαθμό από την ορθότητα της πρώτης. Για την τρίτη παραδοχή έχει ήδη γίνει συζήτηση: οι ανελίξεις ετήσιων τιμών και μέγιστων τιμών μερικής διάρκειας δεν είναι πάντα απαλλαγμένες από στοχαστική εξάρτηση, η οποία πάντως είναι αρκετά μικρή και μπορεί να αγνοηθεί, ιδίως όταν μας ενδιαφέρει η περιθώρια συνάρτηση κατανομής ενός μεμονωμένου μεγέθους. Είναι απαραίτητο πάντως να τονίσουμε ότι στην υδρολογία δεν υπάρχουν γενικά αποδεκτές *a priori* παραδοχές και γι' αυτό πάντα θα πρέπει να γίνεται έλεγχος των παραδοχών *a posteriori*.

Τελικό συμπέρασμα

Οι παραδοχές της στασιμότητας, της εργοδικότητας και της ανεξαρτησίας που έγιναν για τις ετήσιες ανελίξεις ετησίων τιμών, μηνιαίων τιμών συγκεκριμένου μήνα, ετήσιων μεγίστων ή ελαχίστων (στιγμιαίων ή δεδομένης διάρκειας) και μεγίστων μερικής διάρκειας, οδηγούν σε μια τελική απλοποίηση που διευκολύνει πολύ τη μελέτη: Μπορούμε καθένα απ' αυτά τα μεγέθη να το θεωρήσουμε ως μια απλή τυχαία μεταβλητή, και τις τιμές του στα διάφορα υδρολογικά έτη να τις θεωρήσουμε ως διαφορετικές εμφανίσεις της ίδιας τυχαίας μεταβλητής. Έτσι μπορούμε να αντικαταστήσουμε την έννοια της ανέλιξης με την έννοια της απλής τυχαίας μεταβλητής. Για παράδειγμα, δεν μιλούμε για την ανέλιξη των ετήσιων παροχών αλλά για την τυχαία μεταβλητή της

* Υπάρχουν πάντως τρόποι να αντιμετωπιστεί η μη στασιμότητα, οι οποίοι ξεφεύγουν από το σκοπό του παρόντος κειμένου.

ετήσιας παροχής. Αντίστοιχα, αντί να μιλούμε για τη χρονοσειρά των ετήσιων παροχών μιλούμε για το δείγμα της ετήσιας παροχής, αγνοώντας τη χρονική διάταξη των ιστορικών δεδομένων των διαδοχικών υδρολογικών ετών.

Εφαρμογή 4.1

Με βάση τα ιστορικά δεδομένα παροχών ενός ποταμού που διατίθενται για μια μακρά περίοδο, εκτιμήσαμε ότι το ενδεχόμενο να εμφανιστεί ετήσιος όγκος απορροής μικρότερος από 500 hm^3 έχει πολύ μικρή πιθανότητα, ίση με 10^{-2}. (α) Να υπολογιστεί η πιθανότητα να εμφανιστεί επί πέντε διαδοχικά χρόνια όγκος απορροής μικρότερος από 500 hm^3.

Με την προϋπόθεση ότι ισχύουν οι παραδοχές της στασιμότητας, της εργοδικότητας και της ανεξαρτησίας η ζητούμενη πιθανότητα είναι απλώς $\left(10^{-2}\right)^5 = 10^{-10}$. Πρόκειται βέβαια για εξαιρετικά μικρή πιθανότητα. Ωστόσο έχει παρατηρηθεί ότι στη φύση τέτοια φαινόμενα διαδοχής εξαιρετικά ξηρών ετών δεν είναι τόσο απίθανα..

Αν δεν ισχύει η παραδοχή της ανεξαρτησίας, τότε κατά κανόνα η συσχέτιση των τιμών των διαφορετικών ετών (που περιγράφεται από το συντελεστή αυτοσυσχέτισης της ανέλιξης για χρονικό βήμα ένα) είναι θετική. Σε αυτή την περίπτωση η πιθανότητα του υπόψη ενδεχομένου θα είναι μεγαλύτερη από 10^{-10}, αλλά δεν μπορούμε να υπολογίσουμε την τιμή της χωρίς να χρησιμοποιήσουμε προχωρημένες τεχνικές στοχαστικών ανελίξεων.

(β) Αν τα προηγούμενα τέσσερα χρόνια έχει εμφανιστεί όγκος απορροής μικρότερος από 500 hm^3, ποια είναι η πιθανότητα να εμφανιστεί και εφέτος το ίδιο ενδεχόμενο;

Προφανώς, αν ισχύει η παραδοχή της ανεξαρτησίας, η πιθανότητα που ζητείται είναι 10^{-2}. Το τι συνέβη τα προηγούμενα χρόνια δεν επηρεάζει το παρόν αν υπάρχει ανεξαρτησία των γεγονότων. Αν δεχτούμε ότι υπάρχει θετική συσχέτιση μεταξύ των διαδοχικών ετών, τότε η πιθανότητα θα είναι μεγαλύτερη, αφού είναι πιθανότερο ένα ξηρό υδρολογικό έτος να ακολουθείται από επίσης ξηρό.

Ενδεχομένως, η διαισθητική απάντηση σε αυτό το ερώτημα να είναι αντίθετη: η διαίσθηση, παρασυρμένη από την εξαιρετικά μικρότερη πιθανότητα της διαδοχής πέντε διαδοχικών ξηρών ετών, που όπως υπολογίσαμε είναι 10^{-10}, έχει την τάση να αποδώσει πιθανότητα μικρότερη από 10^{-2}, ίσως και 10^{-10}. Πρόκειται όμως για καθαρή πλάνη.*

4.2 Η έννοια της περιόδου επαναφοράς

Η έννοια της *περιόδου επαναφοράς* είναι βασική για τη στατιστική υδρολογία, αν και δεν χρησιμοποιείται στην καθεαυτό θεωρία πιθανοτήτων. Σημειώνουμε εξαρχής ότι η έννοια αυτή είναι υπεύθυνη για πολλές παρανοήσεις που θα εκτεθούν παρακάτω προκειμένου να διασαφηνιστεί όσο γίνεται καλύτερα.

Αρχικά, η περίοδος επαναφοράς, T, μιας δεδομένης τιμής x της τυχαίας μεταβλητής X (η οποία στην πραγματικότητα αντιπροσωπεύει μια στοχαστική ανέλιξη) ορίζεται ως ο *μέσος* αριθμός χρονικών διαστημάτων (εν προκειμένω υδρολογικών ετών) που μεσολαβεί μεταξύ δύο διαδοχικών εμφανίσεων της τυχαίας μεταβλητής με μέγεθος μεγαλύτερο ή ίσο της δεδομένης τιμής x. Για παράδειγμα, αν η τιμή 500 m^3/s της μέγιστης ετήσιας παροχής έχει περίοδο επαναφοράς 50, αυτό σημαίνει ότι κατά μέσο όρο μεσολαβεί ένα διάστημα 50 ετών ανάμεσα σε δύο εμφανίσεις πλημμυρικής παροχής μεγαλύτερης ή ίσης των 500 m^3/s. Αποδεικνύεται (Kottegoda, 1980) ότι η περίοδος επαναφοράς της τιμής x είναι

$$T = \frac{1}{P(X > x)} = \frac{1}{F_{1_X}(x)} = \frac{1}{1 - F_X(x)} \tag{48}$$

δηλαδή η περίοδος επαναφοράς είναι το αντίστροφο της πιθανότητας υπέρβασης. Οι προϋποθέσεις για να ισχύει η παραπάνω σχέση είναι (α) να είναι συνεχής η τυχαία μεταβλητή και (β) να ισχύει η παραδοχή ανεξαρτησίας του προηγούμενου εδαφίου, δηλαδή κάθε εμφάνιση να είναι στοχαστικά

* Βέβαια δεν είναι η μοναδική πλάνη του κοινού νου γύρω από τις πιθανότητες. Σε τέτοιου είδους φαινομενικά “αθώες ” πλάνες έχουν στηριχθεί πολλές απάτες, ιδίως σε τυχερά παιχνίδια, κληρώσεις κ.ο.κ.

ανεξάρτητη από τις προηγούμενες και επόμενές της. Δεδομένου ότι οι προϋποθέσεις αυτές ισχύουν κατά κανόνα για τα μεγέθη που εξετάζουμε σε αυτό το κείμενο, μπορούμε να θεωρούμε την εξίσωση (48) ως ισοδύναμο ορισμό της περιόδου επαναφοράς.

Τα παραπάνω ισχύουν κατά βάση για τις περιπτώσεις που ενδιαφέρουν μεγάλες τιμές υδρολογικών μεγεθών, π.χ. πλημμύρες. Κατ' αναλογία ορίζεται η περίοδος επαναφοράς για μικρά μεγέθη, π.χ. για παροχές ξηρασίας. Στην περίπτωση αυτή, στον αρχικό ορισμό θα πρέπει να αντικαταστήσουμε τη φράση “μεγαλύτερο ή ίσο” με τη φράση “μικρότερο ή ίσο” ή ισοδύναμα να θεωρήσουμε

$$T = \frac{1}{P(X \le x)} = \frac{1}{F_X(x)} \tag{49}$$

Συμπερασματικά λοιπόν, στα πλαίσια αυτού του κειμένου θα θεωρούμε την περίοδο επαναφοράς είτε ως το αντίστροφο της πιθανότητας υπέρβασης, όταν ενδιαφέρει μέγιστο μέγεθος είτε ως το αντίστροφο της πιθανότητας μη υπέρβασης (συνάρτησης κατανομής) όταν ενδιαφέρει ελάχιστο μέγεθος.

Ας έρθουμε τώρα στην κριτική της έννοιας της περιόδου επαναφοράς. Κατ' αρχήν η διττή σημασία της τείνει να δημιουργήσει σύγχυση. Επίσης ο όρος “περίοδος” δημιουργεί τον κίνδυνο να εκληφθεί το μέγεθος στο οποίο αναφέρει ως κυμαινόμενο περιοδικά, πράγμα που είναι τελείως εσφαλμένο αφού αναφέρεται σε τυχαίες μεταβλητές. Για το λόγο αυτό έχει συχνά χρησιμοποιηθεί, χωρίς όμως να επικρατήσει, ο όρος *διάστημα επανόδου*. Ωστόσο, και εδώ υπάρχει σύγχυση, γιατί ο τελευταίος όρος έχει χρησιμοποιηθεί συχνά και με άλλο νόημα, ήτοι του πραγματικού (και όχι του μέσου) χρόνου ανάμεσα σε διαδοχικές υπερβάσεις μιας τιμής, οπότε έχει την έννοια τυχαίας μεταβλητής (π.χ. στους Chow et al., 1988). Ένα τελευταίο πρόβλημα αφορά τη διάσταση του μεγέθους *T*. Όπως έχει οριστεί παραπάνω, το μέγεθος είναι αδιάστατο και ως αδιάστατο χρησιμοποιείται σε όλες τις μαθηματικές σχέσεις. Ωστόσο, στην πράξη αποδίδονται στο μέγεθος διαστάσεις χρόνου με μονάδα ίση με το υδρολογικό έτος. Έτσι λέμε π.χ. ότι η (μέγιστη ετήσια) πλημμυρική παροχή των 500 m^3/s έχει περίοδο επαναφοράς 50 ετών, ή αλλιώς ότι η πλημμύρα 50ετίας είναι 500 m^3/s.

Παρ' όλα τα παραπάνω προβλήματα προβλήματα η περίοδος επαναφοράς έχει χρησιμοποιηθεί ευρέως στην υδρολογία, γιατί γίνεται αμεσότερα αντιληπτή από την έννοια της πιθανότητας και γιατί η αριθμητική της έκφραση είναι ανετότερη από αυτή της πιθανότητας. Για παράδειγμα είναι ανετότερη η έκφραση “πλημμύρα 50ετίας” από την έκφραση “πλημμύρα πιθανότητας υπέρβασης 0.02”. Ωστόσο, συνδυάζοντας τους δύο εκφράσεις, θα μπορούσαμε να πούμε ακριβέστερα “πλημμύρα πιθανότητας υπέρβασης 1 : 50”.

Εφαρμογή 4.2

α. Σε ποια πιθανότητα υπέρβασης και ποια πιθανότητα μη υπέρβασης αντιστοιχούν οι πλημμύρες 2, 5, 10, 20, 50, 100, 500, 1000, 5 000 και 10 000 ετών;

Με βάση την εξίσωση (48) εύκολα βρίσκουμε τις τιμές του Πίνακα 1.

β. Σε ποια πιθανότητα υπέρβασης και ποια πιθανότητα μη υπέρβασης αντιστοιχούν οι ξηρασίες των ίδιων περιόδων επαναφοράς, όπως στο ερώτημα α;

Η απάντηση δίνεται από τον Πίνακα 1, αν αντιμετατεθούν αμοιβαία οι στήλες της πιθανότητας υπέρβασης και πιθανότητας μη υπέρβασης.

Πίνακας 1 Αντιστοιχία περιόδου επαναφοράς, πιθανότητας υπέρβασης και πιθανότητας μη υπέρβασης της Εφαρμογής 4.2.

Περίοδος επαναφοράς T	**Πιθανότητα υπέρβασης** F1 (%)	**Πιθανότητα μη υπέρβασης** F (%)
2	50	50
5	20	80
10	10	90
20	5	95
50	2	98
100	1	99
500	0.2	99.8
1000	0.1	99.9
5 000	0.02	99.98
10 000	0.001	99.99

4.3 Η έννοια της διακινδύνευσης

Κατά το σχεδιασμό ενός έργου είναι συνηθέστατη πρακτική να υπολογίζεται με βάση τα διαθέσιμα δεδομένα η *τιμή σχεδιασμού* ή *φόρτιση σχεδιασμού* L, αλλά η *ικανότητα του έργου* C να υιοθετείται μεγαλύτερη, σε τρόπο ώστε να υπάρχει δεδομένο περιθώριο ασφάλειας

$$\mathrm{SM} := C - L \tag{50}$$

ή δεδομένος, μεγαλύτερος από 1, *συντελεστής ασφάλειας*

$$\mathrm{SF} := \frac{C}{L} \tag{51}$$

Στο χώρο των κατασκευών πολιτικού μηχανικού το μέγεθος L μπορεί, για παράδειγμα, να είναι το δυσμενέστερο φορτίο λειτουργίας ενός υποστυλώματος οπότε το C είναι το φορτίο που προκαλεί θραύση του υποστυλώματος. Στο χώρο της τεχνικής υδρολογίας, το L μπορεί να είναι, για παράδειγμα, η παροχή σχεδιασμού ενός υπερχειλιστή φράγματος οπότε το μέγεθος M είναι η παροχετευτικότητα του υπερχειλιστή, δηλαδή η παροχή που μπορεί να διοδευτεί με ασφάλεια από τον υπερχειλιστή χωρίς κίνδυνο υπερπήδησης του φράγματος.

Η υιοθέτηση δεδομένης συγκεκριμένης αριθμητικής τιμής για το συντελεστή ασφάλειας (π.χ. SF = 3) ή για το περιθώριο ασφάλειας, αποτελεί εμπειρική προσέγγιση στην αντιμετώπιση της αβεβαιότητας που υπάρχει στην εκτίμηση της φόρτισης ενός έργου ή της ικανότητάς του. Ωστόσο, αυτή η πρακτική δεν δίνει πολύ σαφές μέτρο της ασφάλειας και μάλιστα δίνει την εσφαλμένη εντύπωση της πλήρους εξάλειψης του κινδύνου αστοχίας του συγκεκριμένου έργου. Αντίθετα, η προσέγγιση της αβεβαιότητας μέσω της θεωρίας πιθανοτήτων αποτελεί τη μόνη ορθολογική απάντηση στην ποσοτικοποίηση του κινδύνου αστοχίας. Σύμφωνα με αυτή την προσέγγιση, τα μεγέθη SM και SF θεωρούνται τυχαίες μεταβλητές, βάσει των οποίων ορίζεται, ως μέτρο του κινδύνου αστοχίας, η *διακινδύνευση* (ή *επικινδυνότητα - risk*):

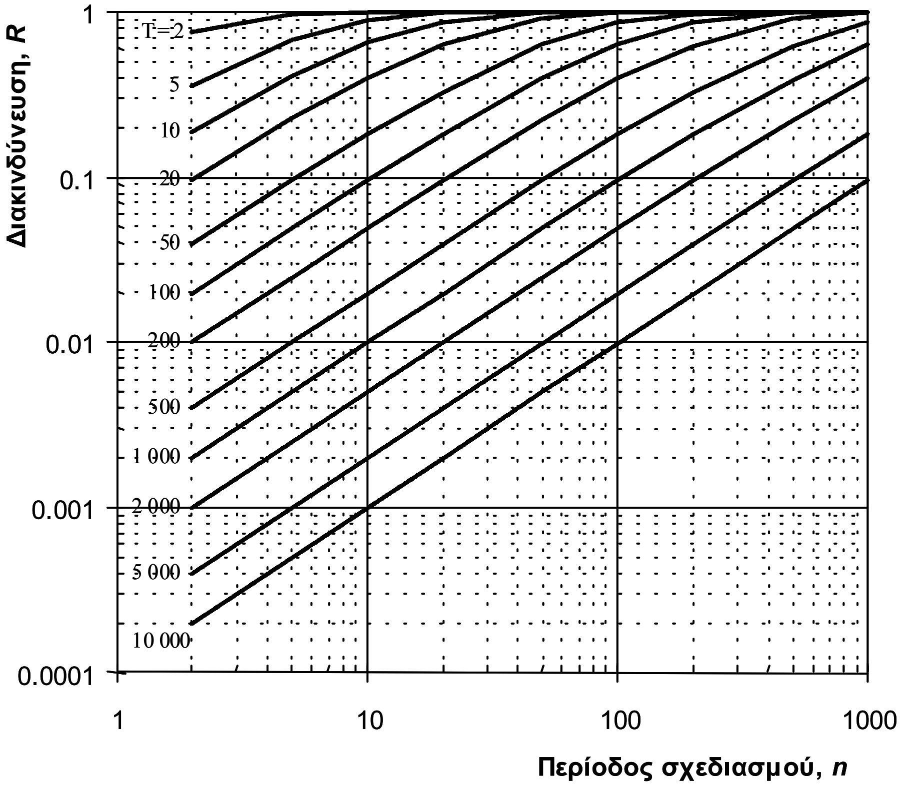
$$R := P(\mathrm{SF} < 1) = P(\mathrm{SM} < 0) \tag{52}$$

Στην τυπικότερη κατηγορία προβλημάτων της τεχνικής υδρολογίας, τα προβλήματα σχεδιασμού, η ικανότητα του έργου (π.χ. η παροχετευτικότητα ή η αποθηκευτική ικανότητα) μπορεί να θεωρηθεί ότι έχει δεδομένη τιμή με πλήρη βεβαιότητα ($C = c$), οπότε δεν θεωρείται ως τυχαία μεταβλητή.* Σε αυτή την περίπτωση η διακινδύνευση είναι

* Ο αναγνώστης που ενδιαφέρεται για συνθετότερα προβλήματα, όπου και η ικανότητα είναι τυχαία μεταβλητή, παραπέμπεται στους Chow et al., 1988.

$$R = P(L > c) = 1 - P(L \le c) \tag{53}$$

και οφείλεται μόνο στην ενυπάρχουσα αβεβαιότητα των υδρολογικών διεργασιών. Για το λόγο αυτό μιλούμε για *φυσική* ή *ενυπάρχουσα διακινδύνευση*.



Σχήμα 7 Γραφική απεικόνιση της σχέσης των χαρακτηριστικών μεγεθών υδρολογικού σχεδιασμού (εξίσωση 55).

Το μέγεθος L εξαρτάται από τη μεταβλητότητα της φυσικής διεργασίας που ενδιαφέρει (πχ. της πλημμυρικής παροχής) καθώς και από τη διάρκεια που θα εκτεθεί το έργο στο φυσικό κίνδυνο, δηλαδή τη *διάρκεια ζωής* του έργου. Εάν X παριστάνει το μέγιστο μέγεθος της φυσικής διεργασίας σε ετήσια βάση (π.χ. τη μέγιστη ετήσια πλημμύρα) και n είναι η διάρκεια ζωής του έργου, τότε το γεγονός $\{L \le c\}$ ισοδυναμεί με n διαδοχικές εμφανίσεις του γεγονότος $\{X \le c\}$. Για να μην έχουμε, δηλαδή, υπέρβαση του μεγέθους c σε όλη τη διάρκεια ζωής του έργου θα πρέπει να μην έχουμε υπέρβαση σε όλα τα n χρόνια αυτής της διάρκειας. Θεωρώντας ότι οι πλημμύρες των διαδοχικών ετών είναι στοχαστικά ανεξάρτητες παίρνουμε

$$R = 1 - [P(X \le c)]^n = 1 - [F_X(c)]^n \tag{54}$$

όπου F_X είναι η συνάρτηση κατανομής του φυσικού μεγέθους που μας ενδιαφέρει, σε ετήσια βάση. Αντικαθιστώντας τη με την αντίστοιχη περίοδο επαναφοράς παίρνουμε την ακόλουθη βασική σχέση που συνδέει τα τρία βασικά μεγέθη υδρολογικού σχεδιασμού ενός έργου, δηλαδή την περίοδο επαναφοράς, τη διάρκεια ζωής και τη διακινδύνευση:

$$R = 1 - \left(1 - \frac{1}{T}\right)^n \tag{55}$$

Γραφική απεικόνιση της (55) δίνεται στο Σχήμα 7 για χαρακτηριστικές περιόδους επαναφοράς.

Ισοδύναμη έκφραση της (55) είναι η ακόλουθη, η οποία έχει επιλυθεί ως προς την περίοδο επαναφοράς:

$$T = \frac{1}{1-(1-R)^{1/n}} \tag{56}$$

Τέλος, δεδομένου ότι $\ln(1-x)^n = n\ln(1-x) = n(-x - x^2/2 - \cdots) \approx -nx$, παίρνουμε και την ακόλουθη προσεγγιστική έκφραση του R:

$$R \approx 1 - e^{n/T} \tag{57}$$

η οποία ισχύει με σφάλμα < 1% για $T \geq 50$.

Εφαρμογή 4.3

α. Διώρυγα εκτροπής σχεδιάζεται να λειτουργήσει κατά την περίοδο κατασκευής φράγματος που εκτιμάται σε 5 χρόνια. Ποια πρέπει να είναι η περίοδος επαναφοράς της πλημμύρας σχεδιασμού, ώστε η διακινδύνευση να μην υπερβαίνει το 10%;

Από την (56) παίρνουμε

$$T = \frac{1}{1-(1-R)^{1/n}} = \frac{1}{1-(1-0.1)^{1/5}} = 47.9$$

Στρογγυλεύουμε σε $T = 50$ χρόνια.

β. Πόση είναι η διακινδύνευση αν ένα έργο σχεδιαστεί με περίοδο επαναφοράς ίση με τη διάρκεια ζωής του;

Αν δεχτούμε ότι η διάρκεια ζωής του έργου είναι αρκετά μεγάλη (≥50 χρόνια), τότε από την (57) παίρνουμε $R = 1 - e^{-1} = 0.632 = 63.2\%$. Διαφορετικά η διακινδύνευση υπολογίζεται από την (55).

Αναφορές

Bras, R.L. and Rodriguez-Iturbe, I., *Random functions in hydrology*, Addison-Wesley, 1985

Chow, V. T., Maidment, D. R., and Mays, L. W., *Applied hydrology*, McGraw-Hill, 1988

Cooper, N. G., *From cardinals to chaos, Reflections on the life and legacy of Stanislaw Ulam*, Cambridge Univ, Press, 1989.

Haan, C. T., Statistical methods in hydrology, The Iowa State University Press, USA, 1977

Kottegoda, N. T., *Stochastic water resources technology*, Macmillan Press, London, 1980.

Papoulis, A., *Probability, random variables and stochastic processes*, McGraw-Hill, 1965.

Salas, J. D., Delleur, J. W., Yevjevich, V. and Lane, W. L., *Applied modelling of hydrologic time series*, Water Resources Publications, Littleton, Co., USA, 1988.

Schroeder, M., *Fractals, chaos, power laws, Minutes from an infinite paradise*, W. H. Freeman and Co., 1990.

Stewart, I., *Does God play dice?*, Penguin, 1990 (Ελληνική μετάφραση Κ. Σαμαρά, *Παίζει ο Θεός ζάρια;*, Εκδόσεις Κωσταράκη, 1991).

Taylor, H. M. and Karlin, S., *An introduction to stochastic modelling*, Academic Press, 1984.