Πιθανότητες και Στατιστική --- εφαρμογές σε ακραία υδρομετεωρολογικά φαινόμενα

Δρ Μ.Σπηλιώτη

Επίκουρος Καθηγητής ΔΠΘ

Βασικές πηγές

- Τσακίρης Γ., 2013. Υδατικοί Πόροι Ι, Τεχνική Υδρολογία. Εκδόσεις Συμμετρία.
- Μπέλλος Β., 2010. Τεχνική Υδρολογία.
- Σακκάς, 2003. Τεχνική υδρολογία.
- Μιμίκου Μ., 1994. Τεχνολογία Υδατικών Πόρων.
- Χρυσάνθου Βλ., 2018. Σημειώσεις Τεχνικής Υδρολογίας
- Μυλωνάς και Παπαδόπουλος 2017. Πιθανότητες και Στατιστική για Μηχανικούς, Εκδόσεις Τζιόλα.
- Ζιούτας Γ, 2012. Πιθανότητες και Στατιστική για Μηχανικούς. Εκδόσεις Σοφία
- H. ALFREDO, S. ANG, W. H. TANG (Μετάφραση Δ. Παναγιωτακόπουλος) (2003). Εφαρμογές Πιθανοτήτων και Στατιστικής στη Μελετη και Προγραμματισμό Τεχνικών Έργων. Εκδόσεις ΑΦΟΙ ΚΥΡΙΑΚΙΔΗ.
- Mimikou M, Baltas E., Tsihrintzis, V., 2016. Hydrology and water resources systems analysis. CRC Press

Βασικές έννοιες

Πιθανότητες

- Ιδεατό
- Τυχαία μεταβλητή
- Συνεχείς και διακριτές κατανομές
- Ταίριασμα μεταξύ στατιστικού δείγματος και θεωρητικής κατανομής πιθανότητας

Πιθανότητες

- Η τύχη είναι άμεσα συνδεδεμένη με τα υδρολογικά φαινόμενα (πλημμύρες, ξηρασίες) με αποτέλεσμα να περιγράφονται σε μικρό ή μεγάλο βαθμό από τη θεωρία των πιθανοτήτων
- Η τεχνική υδρολογία στηρίζεται σε μετρήσεις φυσικών μεταβλητών που η επεξεργασία τους προϋποθέτει τη χρήση στατιστικών μεθόδων (έλεγχος των σφαλμάτων των μετρήσεων, συμπλήρωση ελλείψεων ιστορικών δειγμάτων και κυρίως επέκταση χρονοσειρών) και το συνταίριασμα τους με τις πιθανότητες (κόσμος του ιδεατού)
- Η λήψη αποφάσεων για το σχεδιασμό και τη βέλτιστη λειτουργία των υδραυλικών έργων και των υδατικών συστημάτων γενικότερα, γίνεται πάντοτε υπό καθεστώς αβεβαιότητας, η οποία μπορεί να ποσοτικοποιηθεί με την θεωρία των πιθανοτήτων

Εμπειρική κατανομή πιθανότητας

Εμπειρική Πιθανότητα.

Ας υποθέσουμε ότι ένα στοχαστικό πείραμα, με δειγματικό χώρο Ω μπορεί να εκτελεσθεί με τις ίδιες συνθήκες απεριόριστες φορές. Έστω ότι σε ν εκτελέσεις του πειράματος, το ενδεχόμενο A έλαβε χώρα $n_\nu(A)$ φορές. Τότε η πιθανότητα του ενδεχομένου A θα είναι:

$$P(A) = \lim_{\nu \to \infty} \frac{n_\nu(A)}{\nu}.$$

Η εμπειρική αυτή πιθανότητα, είναι:

(1) μη αρνητική: $P(A) \geq 0$

(2) καν/νη: $P(\Omega) = 1$

(3) προσθετική: $P(A \cup B) = P(A) + P(B)$,

για αμοιβαίως αποκλειόμενα ενδεχόμενα A και B.

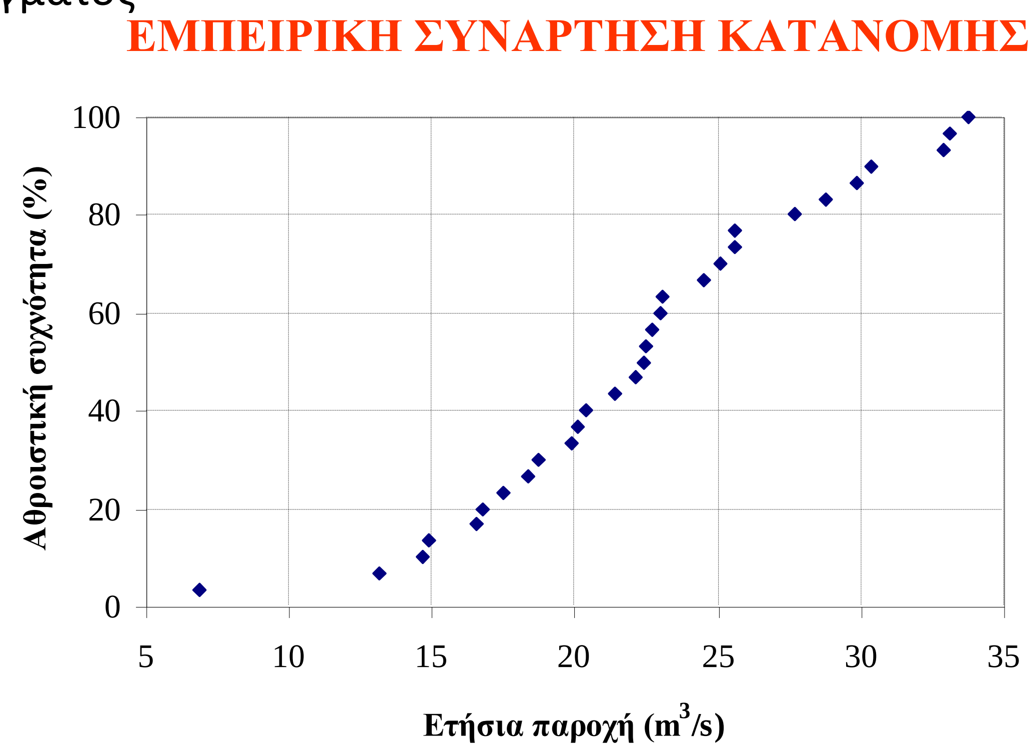
ΕΜΠΕΙΡΙΚΗ ΠΙΘΑΝΟΤΗΤΑ

Η πιθανότητα να λάβει μια τιμή η μεταβλητή βάσει των πραγματικών μετρήσεων (των τιμών του δείγματος) . Τα δεδομένα τοποθετούνται με φθίνουσα σειρά και η εμπειρική πιθανότητα υπολογίζεται από τη σχέση:

$$F_X(x) = 1 - \frac{m}{N+1} \qquad Weibull(1939)$$

Υδρολογία, χρήση ανοικτών σειρών, παρανομαστής N+1

όπου m=αύξων αριθμός (δεδομένα κατά φθίνουσα σειρά)
N= μέγεθος δείγματος



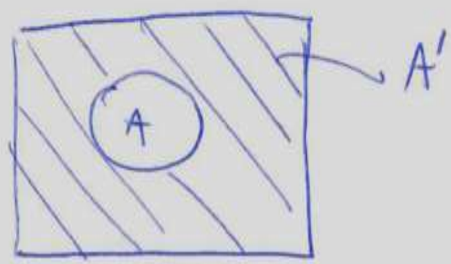
Βασικές έννοιες πιθανοτήτων

Συμπλήρωμα

Συμπλήρωμα:

Αν A' συμπλήρωμα ενός ενδεχομένου A, ως προς το χώρο Ω, τότε

$$P(A') = 1 - P(A)$$



Ξένα ενδεχόμενα και η πιθανότητα της **ένωσης** και της **τομή τους**

Αν δύο οποιαδήποτε γεγονότα είναι **ξένα** μεταξύ τους ($A_1 \cap A_2 = \emptyset$) π.χ. (η ομορφιά να είναι φτωχή $\cap$ η ομορφιά να είναι πλούσια $= \emptyset$)

ισχύει

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \quad (\text{ένωση "ή"})$$

προφανώς $P(A_1 \cap A_2) = 0$.

Τομή ανεξάρτητων ενδεχομένων

Τομή (και)

Για A και B στοχαστικά ανεξάρτητα ενδεχόμενα.

(η πραγματοποίηση του A ενδεχομένου, δεν επηρεάζει την πραγματοποίηση του B ενδεχομένου) τότε!

$$P(A \cap B) = P(A) \cdot P(B)$$

Πιθανότητα της ένωσης δύο γεγονότων

Για την πιθανότητα της ένωσης δύο γεγονότων ισχύει

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ένωση

τομή

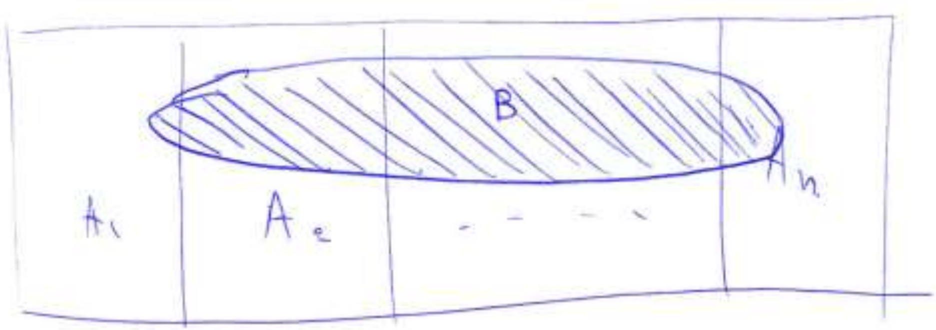
Δεσμευμένη πιθανότητα:

$P(A/B) =$ πιθανότητα να συμβεί το ενδεχόμενο A με δεδομένο ότι έχει συμβεί το B $=$

$$= \frac{P(A\cap B)}{P(B)} = \text{δεσμευμένη πιθανότητα}$$

$$\boxed{P(B) = P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + \dots + P(B/A_n)P(A_n)}$$

$A_1 \cup A_2 \cdots \cup A_n = \Omega\,,\quad A_i \cap A_j = \varnothing$



Θεώρημα Bayes

$$P(E|B) = \frac{P(B|E)P(E)}{\sum_i P(B|A_i)P(A_i)}$$

Ανεξάρτητα Ενδεχόμενα

Δύο ενδεχόμενα A και B με $P(A) > 0$ και $P(B) > 0$ λέγονται ανεξάρτητα, αν και μόνον αν $P(A|B) = P(A)$ και $P(B|A) = P(B)$.

$\downarrow$

Για δύο ανεξάρτητα ενδεχόμενα (όχι ξένα)

$$P(A \cap B) = P(A) \cdot P(B)$$

$\downarrow$ A,B ανεξάρτητα ενδεχόμενα.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) =$$

$$\underbrace{P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)}_{\text{για } A,B \text{ ανεξάρτητα ενδεχόμενα.}}$$

Πιθανότητα τομής σε μη ανεξάρτητα ενδεχόμενα

$$P(A \cap B) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B).$$

Έννοια δεσμευμένης πιθανότητας

ΠΕΡΙΟΔΟΣ ΕΠΑΝΑΦΟΡΑΣ - ΔΙΑΚΙΝΔΥΝΕΥΣΗ

- Περίοδος επαναφοράς, Τ μιας δεδομένης τιμής x της τυχαίας μεταβλητής Χ είναι ο μέσος αριθμός χρονικών διαστημάτων (εν προκειμένω υδρολογικών ετών) που μεσολαβεί μεταξύ 2 διαδοχικών εμφανίσεων της τυχαίας μεταβλητής με μέγεθος μεγαλύτερο ή ίσο της δεδομένης τιμής x.

$$\boxed{P(X \geq x) = \frac{1}{T}}$$

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

- Διακινδύνευση είναι η πιθανότητα *R* να πραγματοποιηθεί μέσα σε *n* έτη τιμή που αντιστοιχεί σε περίοδο επαναφοράς *T.*

Πιθανότητα υπέρβασης σε *n* έτη

(Διακινδύνευση): να συμβεί τουλάχιστον μία φορά σε n χρόνια η πλημμύρα με περίοδο επαναφοράς Τ)

$$R=1-(1-1/T)^n$$

Risk

Έστω η 50 χρόνων πλημμύρα. $\left(p = \frac{1}{50} = 0.02\right)$

Ποια η πιθανότητα ότι μία τουλάχιστον πλ θα συμβεί στα επόμενα 30 χρόνια μία τουλάχιστον πλημμύρα με T = 50 χρόνια?

$$\text{Risk} = 1 - \underbrace{(1 - 0.02)}_{\text{καμία πλημμύρα}}{}^{30} = 0.455$$

στα 30 χρόνια.

$1 - 0.02 =$ καμία πλημμύρα.

$\underbrace{(1-0.02) \cdots\cdots (1-0.02)} =$ καμία πλ. στα 30 χρόνια

$1 - (1-0.02)^{30} =$ τουλάχιστον μία πλ στα 30 χρόνια $=$

$= \left(\overline{\text{καμία πλημμύρα στα 30 χρόνια}}\right)$

Αν η κατασκευή έχει μια περίοδο των 100 χρόνια,
τότε

$$Risk = 1 - \left(1 - \frac{1}{100}\right)^{30} = 0.26$$

Παράδειγμα: Πιθανότητα εμφάνισης μιας αιχμής πλημμύρας με παροχή $Q \geqslant x$, όπου x η αιχμή πλημμύρας με περίοδο επαναφοράς $T = 50$ έτη, κατά τη διάρκεια μιας περιόδου τριών ετών ($n = 3$):

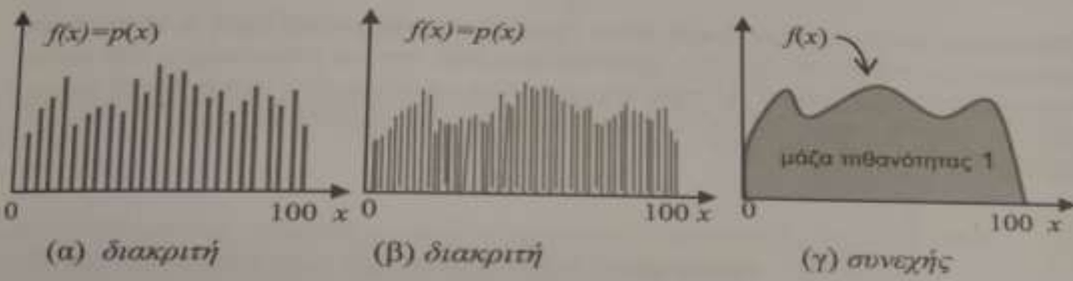
$$p_3 = 1 - \left(1 - \frac{1}{50}\right)^3 = 0.059 \simeq 6\%$$

Συνεχής τ.μ., συνάρτηση πυκνότητας πιθανότητας, η κανονική κατανομή

Ζιούτας Γ, 2012. Πιθανότητες και Στατιστική για Μηχανικούς. Εκδόσεις Σοφία

... σαν ένα συνεχές.

Έστω ότι έχουμε μία διακριτή τυχαία μεταβλητή X η οποία μπορεί να πάρει 25 διακριτές τιμές από το 0 μέχρι το 100 με συνάρτηση μάζας πιθανότητας όπως φαίνεται στο Σχήμα 4.5 (α). Ας πούμε τώρα ότι η τυχαία μεταβλητή X μπορεί να πάρει περισσότερες τιμές στο διάστημα $0 \le x \le 100$, όπου ο αριθμός τιμών είναι μεγάλος αλλά πεπερασμένος, όπως 1, 2, 3, … , 98, 99, 100. Η μάζα πιθανότητας των 25 μπάρων απλώνεται τώρα σε περισσότερα σημεία της X, όπως παριστάνεται στην περίπτωση (β) του Σχήματος 4.5.



Σχήμα 4.5 (α), (β), συναρτήσεις μάζας πιθανότητας, (γ) συνάρτηση πυκνότητας πιθανότητας

Σε κάθε μία από τις δυνατές τιμές x_i αντιστοιχεί μία πιθανότητα $p_i = p(x_i) = P(X=x_i)$, όπου i=1,2,…100, των οποίων το άθροισμα ισούται με την μονάδα. Αναφέραμε παραπάνω ότι η τυχαία μεταβλητή X ενδέχεται να πάρει οιαδήποτε τιμή μέσα στο διάστημα, $0 \le x \le 100$. Αφού οι τιμές που μπορεί να πάρει η X είναι άπειρες μη μετρήσιμες, τι θα συμβεί με τις πιθανότητες $p_i = p(x_i)$; Οι πιθανότητες p_i είναι άπειρες τον αριθμό και το άθροισμά τους θα οδηγούσε σε τιμή μεγαλύτερη της μονάδας. Οι πιθανότητες αυτές έχουν μηδενική τιμή και δεν έχουν νόημα. Εκείνο που πρέπει να κάνουμε είναι να σκορπίσουμε την μάζα των πιθανοτήτων $p(x_i)$ που ορίζονται μόνο για τις τιμές $x_1, x_2, x_3, \ldots, x_n$ σε όλο το διάστημα, $0 \le x \le 100$, όπως φαίνεται στο Σχήμα 4.5 (γ). Προκύπτει έτσι μία συνάρτηση f ορισμένη (για το παρόν παράδειγμα) για όλες τις τιμές x, $0 \le x \le 100$, η οποία εκφράζει την πυκνότητα της μάζας πιθανότητας σε όλο το πεδίο τιμών της X.

$$P(a<X\leq b) = \int_a^b f_X(x)\,dx \tag{3.3}$$

Με βάση την (3.3), η συνάρτηση κατανομής είναι

$$F_X(x) = P(X\leq x) = \int_{-\infty}^{x} f_X(\xi)\,d\xi \tag{3.4}$$

και αν η πρώτη παράγωγος της $F_X(x)$ υπάρχει, τότε

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{3.5}$$

Πρέπει ίσως να επαναληφθεί ότι η συνάρτηση $f_X(x)$ δεν εκφράζει πιθανότητες. Όμως το $f_X(x)dx = P(x < X \leq x + dx)$ είναι η πιθανότητα ότι οι τιμές της X θα είναι στο διάστημα $[x, x + dx]$.

... την κατανομή πιθανό-

Συνάρτηση πυκνότητα πιθανότητας (για συνεχής τ.μ.)

Ορισμός: Ονομάζουμε συνάρτηση πυκνότητας πιθανότητας, μιας συνεχούς τυχαίας μεταβλητής X τη συνάρτηση $f_X(x)$, για την οποία ισχύουν:

(1) $f_X(x) \geq 0, \ \forall x \in R.$ (θετικές τιμές)

(2) $\displaystyle\int_{-\infty}^{+\infty} f_X(x)\,dx = 1.$

(3) η πιθανότητα του γεγονότος $X \in (a,b)$:

$$P(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

Η συνάρτηση $f_X(x)$ <u>δεν</u> περιγράφει πιθανότητα.

Ορισμός: $f_X(x)dx = P(x \le X \le x+dx)$ ιδ.

$f_X(x)dx$ = η πιθανότητα ότι οι τιμές της X θα είναι στο διάστημα $[x, x+dx)$

ΘΕΩΡΗΤΙΚΗ ΠΙΘΑΝΟΤΗΤΑ
(Συνάρτηση Κατανομής και αθροιστική πιθανότητα)

Συνάρτηση πυκνότητας πιθανότητας

Χ τυχαία μεταβλητή

Συνάρτηση πυκνότητας πιθανότητας
(συνάρτηση κατανομής)

$$f_X(X) = \frac{dF_X(x)}{dx}$$

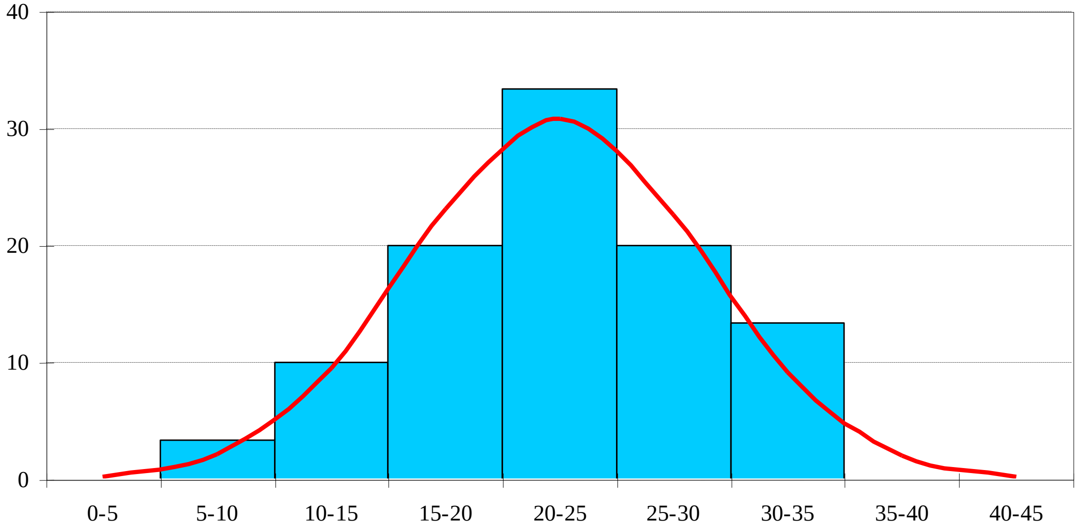
Πιθανότητα μη υπέρβασης
Η πιθανότητα η τυχαία μεταβλητή να είναι μικρότερη ή ίση της δεδομένης τιμής x

$$F_X(x) = P(X \leq x) \; (\alpha\theta\rho o\iota\sigma\tau\iota\kappa\acute{\eta}\, \Pi\iota\theta\alpha\nu\acute{o}\tau\eta\tau\alpha)$$
$$0 = F_X(-\infty) \leq F_X(x) \leq F_X(+\infty) = 1$$

Πιθανότητα υπέρβασης
Η πιθανότητα η τυχαία μεταβλητή να είναι μεγαλύτερη της δεδομένης τιμής x

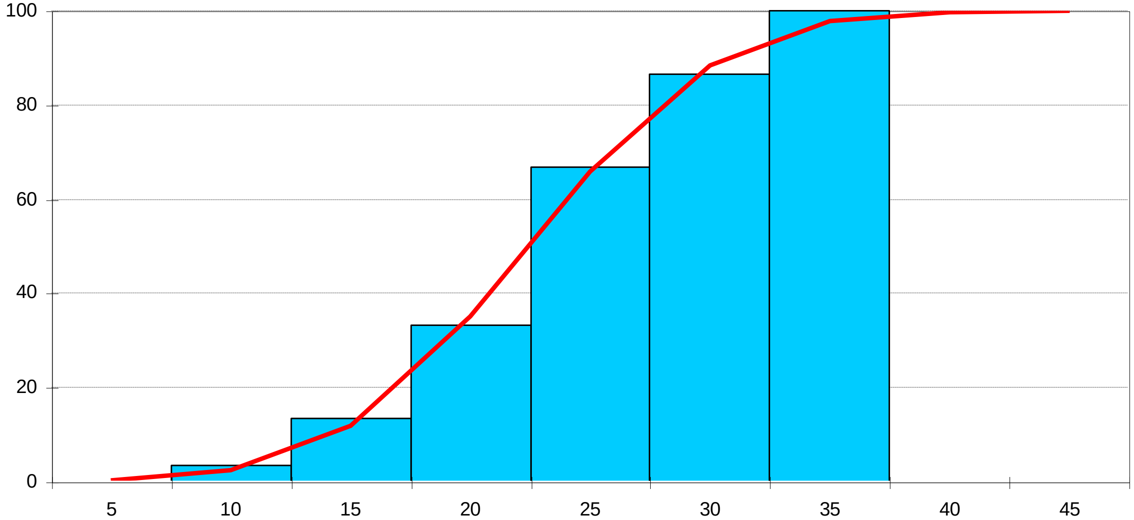
$$F_{1X} = P(X \geq x) = 1 - F_X(x)$$

Διάγραμμα συνάρτησης πυκνότητας πιθανότητας (συνάρτηση κατανομής)



Η συνάρτηση πυκνότητας πιθανότητας ισούται με την πρώτη παράγωγο της αθροιστικής πιθανότητας

Διάγραμμα αθροιστικής πιθανότητας



Η αθροιστική πιθανότητα δηλώνει την πιθανότητα όπου μια τυχαία μεταβλητή είναι ίση ή μικρότερη από κάποια δοθείσα τιμή

Βασικές έννοιες
(συνεχείς και διακριτές κατανομές)

Οι τιμές των Υδρολογικών δεδομένων είναι δυνατόν να θεωρηθούν τυχαίες οπότε μελετώνται σαν Στατιστικές μεταβλητές.

Πληθυσμός στη Στατιστική είναι ένα σύνολο στοιχείων με κοινές ιδιότητες π.χ. τα ετήσια ύψη βροχής ενός βροχομετρικού σταθμού.

Δείγμα είναι ένα τμήμα του πληθυσμού.

Οι μεταβλητές διακρίνονται σε

διακεκριμένες όπως π.χ. ο αριθμός των εμφανίσεων μιας τιμής της παροχής ενός ποταμού σε μία θέση μεγαλύτερης από μια συγκεκριμένη τιμή κατά τη διάρκεια μιας χρονικής περιόδου και

συνεχείς όπως π.χ. οι τιμές της μέσης ημερήσιας παροχής κατά τη διάρκεια μιας χρονικής περιόδου.

- Σε πολλά προβλήματα (υδρολογίας, υδρογεωλογίας, μετεωρολογίας κ.α) **οι μεταβλητές θεωρούνται τυχαίες** (π.χ. απορροή, βροχόπτωση κτλ.).
- Κάθε τυχαία μεταβλητή μπορεί να περιγραφεί από μια συνάρτηση πυκνότητας πιθανότητας και την αντίστοιχη αθροιστική πιθανότητα κάποια κατανομής.
- Η εύρεση της κατανομής που ακολουθεί η τυχαία μεταβλητή είναι κρίσιμης σημασίας διότι γίνεται εφικτή η πρόβλεψη εμφάνισης ή μη μιας τιμής της μεταβλητής κάτω από μια πιθανότητα.

- **Η εύρεση της κατανομής γίνεται με το ταίριασμα της θεωρητικής πιθανότητας (που υποθέτουμε ότι ακολουθεί το δείγμα) με την εμπειρική (που υπολογίζεται βάσει Weibull).**
- Για το ταίριασμα αυτό αξιοποιούνται εργαλεία της **Στατιστικής.** Υπολογίζονται οι στατιστικές παράμετροι του μέσου όρου, της τυπικής απόκλισης, του συντελεστή μεταβλητότητας και του συντελεστή ασυμμετρίας από το δείγμα.

Κανονική κατανομή (Normal distribution)

Συνάρτηση πυκνότητας πιθανότητας:

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{2.32}$$

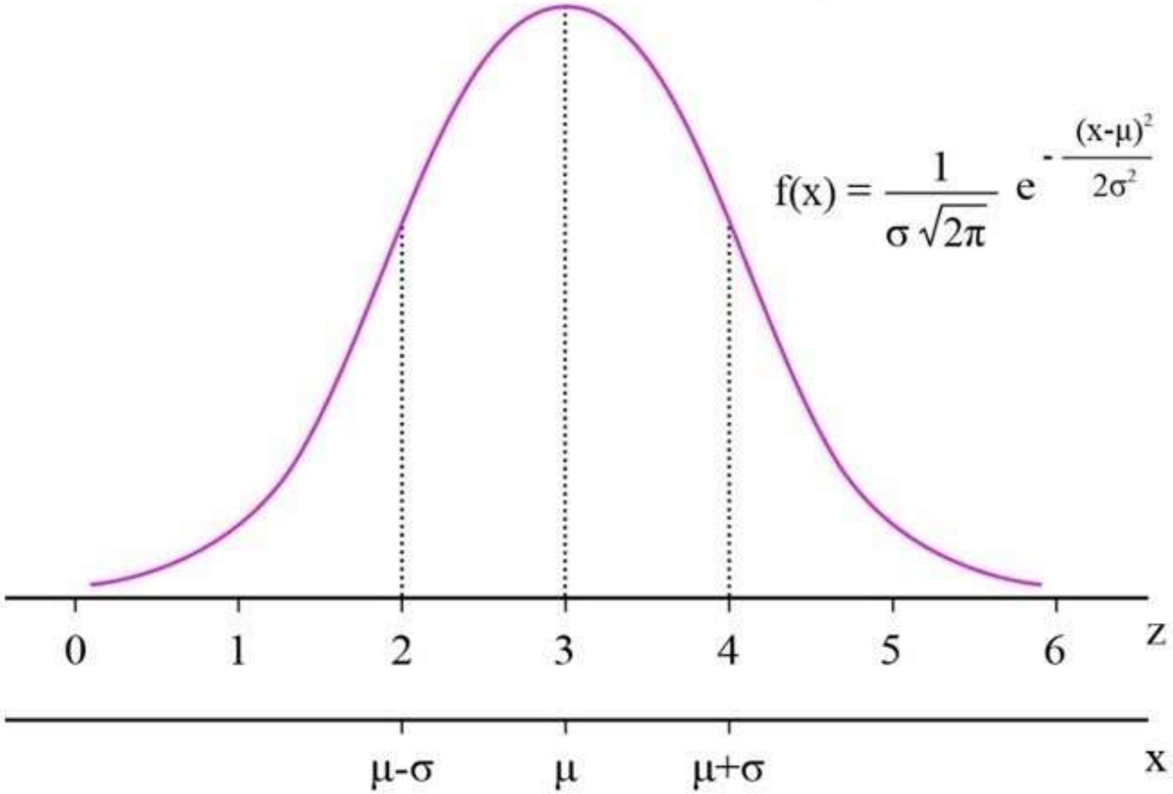
Αθροιστική συνάρτηση πιθανότητας:

$$p(X\leq x)=F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}p(x)dx \tag{2.33}$$

αν $z=\dfrac{x-\mu}{\sigma}$ $\quad p(z)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}=\sigma p(x)$ $\quad$ (2.34)

$$F(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}p(z)dz \tag{2.35}$$

Κανονική κατανομή



Πίνακας με τις τιμές της μεταβλητής Ζ. Στην κανονική κατανομή ο παράγοντας συχνότητας Κ συμπίπτει με τη΄μεταβλητή Ζ.

Ο παράγοντας συχνότητας για την κανονική κατανομή είναι ίσος με την ανηγμένη μεταβλητή, $k_T = z$.

Πίν. 2.1: *Τιμές του z για δοσμένες τιμές της αθροιστικής πιθανότητας P(z) της Τυπικής Κανονικής Κατανομής*

$P(z)$ %	z	$P(z)$ %	z	$P(z)$ %	z	$P(z)$ %	z	$P(z)$ %	z	$P(z)$ %	z
1	− 2.326	21	− 0.806	41	− 0.228	61	0.279	81	0.878	99.1	2.366
2	− 2.054	22	− 0.772	42	− 0.202	62	0.305	82	0.915	99.2	2.409
3	− 1.881	23	− 0.739	43	− 0.176	63	0.332	83	0.954	99.3	2.457
4	− 1.751	24	− 0.706	44	− 0.151	64	0.358	84	0.994	99.4	2.512
5	− 1.645	25	− 0.674	45	− 0.126	65	0.385	85	1.036	99.5	2.576
6	− 1.555	26	− 0.643	46	− 0.100	66	0.412	86	1.080	99.6	2.652
7	− 1.476	27	− 0.613	47	− 0.075	67	0.440	87	1.126	99.7	2.748
8	− 1.405	28	− 0.583	48	− 0.050	68	0.467	88	1.175	99.8	2.878
9	− 1.341	29	− 0.553	49	− 0.025	69	0.496	89	1.227	99.9	3.090
10	− 1.382	30	− 0.524	50	0.000	70	0.524	90	1.282		
11	− 1.227	31	− 0.496	51	0.025	71	0.553	91	1.341	99.91	3.121
12	− 1.175	32	− 0.468	52	0.050	72	0.583	92	1.405	99.92	3.156
13	− 1.126	33	− 0.440	53	0.075	73	0.613	93	1.476	99.93	3.195
14	− 1.080	34	− 0.412	54	0.100	74	0.643	94	1.555	99.94	3.239
15	− 1.036	35	− 0.385	55	0.126	75	0.674	95	1.645	99.95	3.291
16	− 0.994	36	− 0.358	56	0.151	76	0.706	96	1.751	99.96	3.353
17	− 0.954	37	− 0.332	57	0.176	77	0.739	97	1.881	99.97	3.432
18	− 0.915	38	− 0.305	58	0.202	78	0.772	98	2.054	99.98	3.540
19	− 0.878	39	− 0.279	59	0.228	79	0.806	99	2.326	99.99	3.719
20	− 0.842	40	− 0.253	60	0.253	80	0.842				

Παράδειγμα:

Αν σε ένα πληθυσμό $\mu = 20$ και $\sigma = 5$ να ευρεθεί η πιθανότητα $X < 2$.
Υπολογίζεται η ανοιγμένη μεταβλητή z:

$$z = (x - \mu)/\sigma = (2 - 20)/5 = -3.6$$

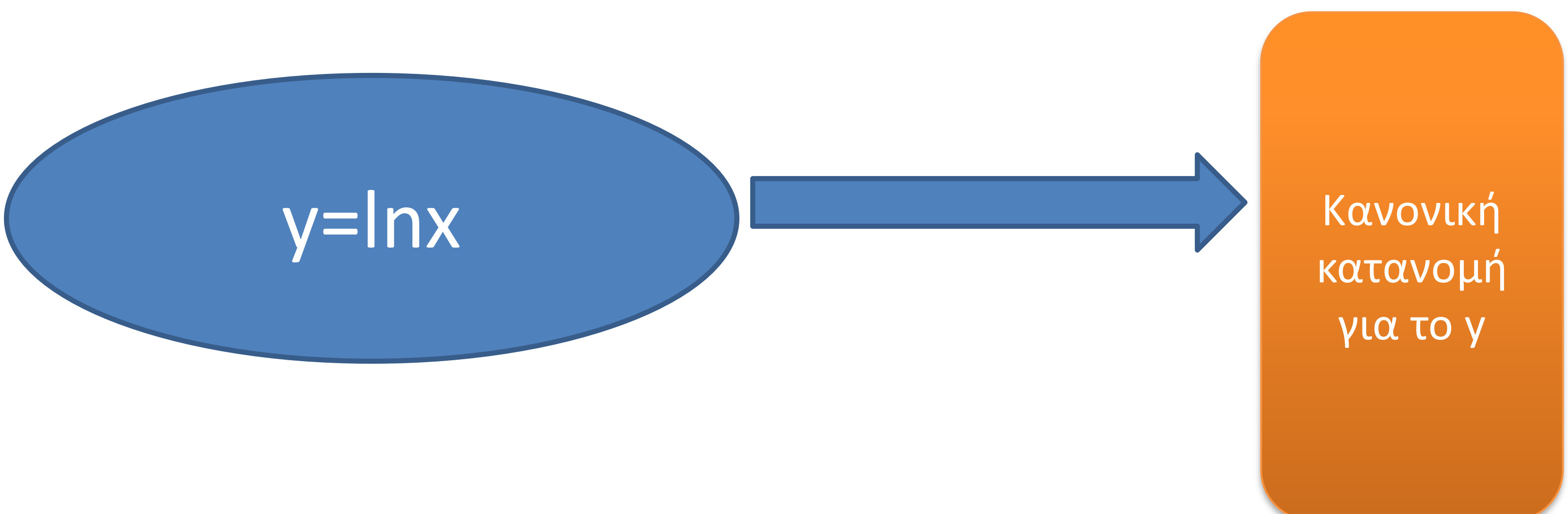
Από τους πίνακες της Κανονικής κατανομής (Παραρτ. Πιν. 1) για z = -3.6

Λογαριθμικές Κατανομές
(περίπτωση Λογαριθμοκανονικής κατανομής)

Λογαριθμοκοκανονική κατανομή (Log-normal ή κατανομή Galton)

Στην κατανομή αυτή η κανονική κατανομή προσαρμόζεται στους Νεπέριους λογαρίθμους της στατιστικής μεταβλητής οπότε αν x_i είναι η στατιστική μεταβλητή τότε στην κανονική κατανομή προσαρμόζεται η μεταβλητή $y_i = ln\, x_i$.

Οι αντίστοιχοι υπολογισμοί γίνονται όπως στην κατανομή Log Pearson III.



Οι στατιστικές παράμετροι της αρχικής μεταβλητής x συνδέονται με τις αντίστοιχες παραμέτρους της μεταβλητής y με τις ακόλουθες σχέσεις

$$\mu = e^{\mu_y + \frac{1}{2}\sigma_y^2} \tag{4.37}$$

$$\sigma = \mu\left[e^{\sigma_y^2} - 1\right]^{1/2} \tag{4.38}$$

$$C_v = \left[e^{\sigma_y^2} - 1\right]^{1/2} \tag{4.39}$$

$$C_s = 3C_v + C_v^3 \tag{4.40}$$

Στατιστικό Δείγμα

Αμερόληπτοι εκτιμητές

Στατιστικό δείγμα - αμερόληπτοι εκτιμητές.

- $\boxed{\text{Μέσος όρος}}$

$$\bar{X} = \frac{\sum_{i=1}^{N} x_i}{N}$$

, x_i: μέτρηση

N: πλήθος δείγματος

- $\boxed{\text{Τυπική απόκλιση}}$

$$\hat{s} = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N-1}}$$

- άλλοι αμερόληπτοι εκτιμητές

Συντελεστής ασυμμετρίας

$$C_s \simeq \frac{a}{\hat{s}^3} =$$

$a \rightarrow$ ασυμμετρία.

$$= \frac{N}{(N-1)(N-2)} \sum (x_i-\bar{x})^3$$

- Ο **μέσος όρος** (average) στον πληθυσμό παρίσταται συνήθως με μ ενώ σε δείγμα με $\bar{x}$

$$\bar{x} = \frac{1}{N}\sum_{1}^{k} N_i x_i = \frac{1}{k}\sum x_i \qquad (2.22)$$

- Η **διάμεσος** (median) είναι η τιμή που αντιστοιχεί στο 50% της κατανομής των πιθανοτήτων.

- Η **πιθανότερη τιμή** (mode) είναι η τιμή που εμφανίζεται πιο συχνά. Στις συνεχείς συναρτήσεις είναι η μέγιστη τιμή της συνάρτησης πυκνότητας πιθανότητας.

- **Η μεταβλητότητα** (variance) είναι:

$$\sigma^2 = \frac{1}{N}\sum (x_i - \bar{x})^2 \qquad (2.23)$$

για μικρά δείγματα πολλαπλασιάζεται με την ποσότητα $\frac{N}{N-1}$

- Η **τυπική απόκλιση** (standard devistion) η οποία παρίσταται με σ ενώ σε δείγμα με s

$$\sigma = \sqrt{\mu_2} \qquad s = \sqrt{\frac{1}{N-1}\sum N_i (x_i - \bar{x})^2} \qquad (2.24)$$

- **Τυπικό σφάλμα** (standard error) ονομάζεται η ποσότητα: $\frac{s}{\sqrt{N}}$

- **Ο συντελεστής μεταβλητότητας** (coefficient of variation)

$$c_v = \frac{\sigma}{\mu} \approx \frac{s}{\bar{x}} \qquad (2.25)$$

- **Η ασυμμετρία** $\alpha - \mu_3$ για μικρά δείγματα είναι:

$$a = \frac{N}{(N-1)(N-2)}\sum (x_i - \bar{x})^3 \qquad (2.26)$$

- **Ο συντελεστής ασυμμετρίας**

$$c_s = \frac{\alpha}{\sigma^3} \approx \frac{a}{s^3} \qquad (2.27)$$

– Κύρτωση:

Δύο κατανομές μπορούν να έχουν την ίδια μέση τιμή και τυπική απόκλιση αλλά να διαφέρουν μεταξύ τους ως προς την "κύρτωση" (αιχμηρότητα της καμπύλης συχνότητας πιθανότητας)

Υπάρχουν διάφορα μέτρα κύρτωσης.

Συντελεστής Κύρτωσης: κ

$$K = \frac{1}{N s^4} \sum_{i=1}^{n} (x_i - \bar{x})^4 .$$

για μικρά δείγματα γίνεται πολλαπλασιασμός

επί $\frac{N^3}{(N-1)(N-2)(N-3)}$.

Για την κανονική κατανομή (ιδεατός κόσμος),

$\kappa = 3$. (Μιμίκου, 1994)

• Η εντολή KURT του Excel υπολογίζει το παρακάτω μέτρο: (Μηλιάδη και Ασημακόπουλος)

$$\left\{ \frac{N(N+1)}{(N-1)(N-2)(N-3)} \sum_{i=1}^{N} \left(\frac{x_i - \bar{x}}{s} \right)^4 \right\} - \frac{3(N-1)^2}{(N-2)(N-3)}$$

Συνταίριασμα μεταξύ στατιστικού δείγματος και θεωρητικών κατανομών

Γιατί χρησιμοποιώ θεωρητικές κατανομές πιθανοτήτων?

- Γιατί χρησιμοποιώ θεωρητικές κατανομές πιθανοτήτων και όχι την εμπειρική κατανομή?

Συνήθως, η περίοδος επαναφοράς (πχ στο χιλιασμό) είναι μεγαλύτερη (ή και πολλαπλάσια) από το μέγεθος του δείγματος. Επομένως καταφεύγουμε σε θεωρητικές συναρτήσεις συχνότητας πιθανότητας και τις προσαρμόζουμε στο στατιστικό δείγμα

Πχ την "προσαρμογή" (ρύθμιση) των θεωρητικών κατανομών πιθανότητας υπάρχουν διάφορες μέθοδοι πχ

1. η μέθοδος των ροπών
2. η μέθοδος της μέγιστης πιθανοφάνειας
3. Ελαχίστων τετραγώνων

κλπ

Ειδικά στην υδρολογία χρη/ται και οι

4. γραφική μέθοδος
5. παράγοντα συχνότητας

Εκτίμηση παραμέτρων

Μέθοδος των ροπών

Μέθοδος των ροπών

Ⓐ Θεωρητικής κατανομής πιθανότητας

$$\mu_x = E[X] = \begin{cases} \sum_j x_j f_x(x_j) & \text{διακριτή} \\ \int_{-\infty}^{+\infty} x f_x(x)\, dx & \text{συνεχής} \end{cases}$$

Επομένως, η μ_x είναι το κέντρο βάρους της θεωρητικής συνάρτησης πυκνότητας πιθανότητας

(B) Διασπορά

$$Var[X] = E\left[(X-\mu_X)^2\right] = \begin{cases} \sum_j (x_j-\mu_X)^2 f_X(x_j) & \text{διακριτή} \\ \int_{-\infty}^{+\infty} (x-\mu_X)^2 f_X(x)\,dx & \text{συνεχής} \end{cases}$$

$Var[X] < Var[Y]$

(μέτρο του ανοίγματος της θαμπ)

↳ τυπική απόκλιση $= \sqrt{Var[X]}$ = ίδιες μονάδες με τυχαία μεταβλητή

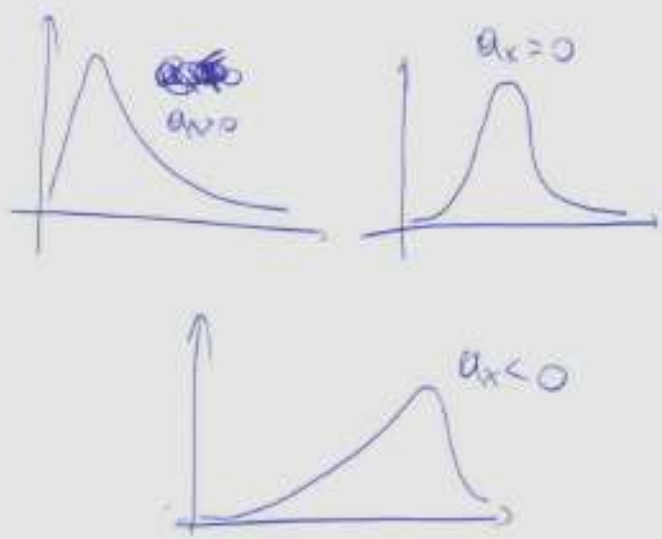
(Γ) Ασυμμετρία.

$$a_x = E\left[(X-\mu_x)^3\right]$$

$$\downarrow$$

συντελεστής ασυμμετρίας $\gamma_x = \dfrac{a_x}{\sigma_x^3}$

a_x = μέτρο της συμμετρίας.



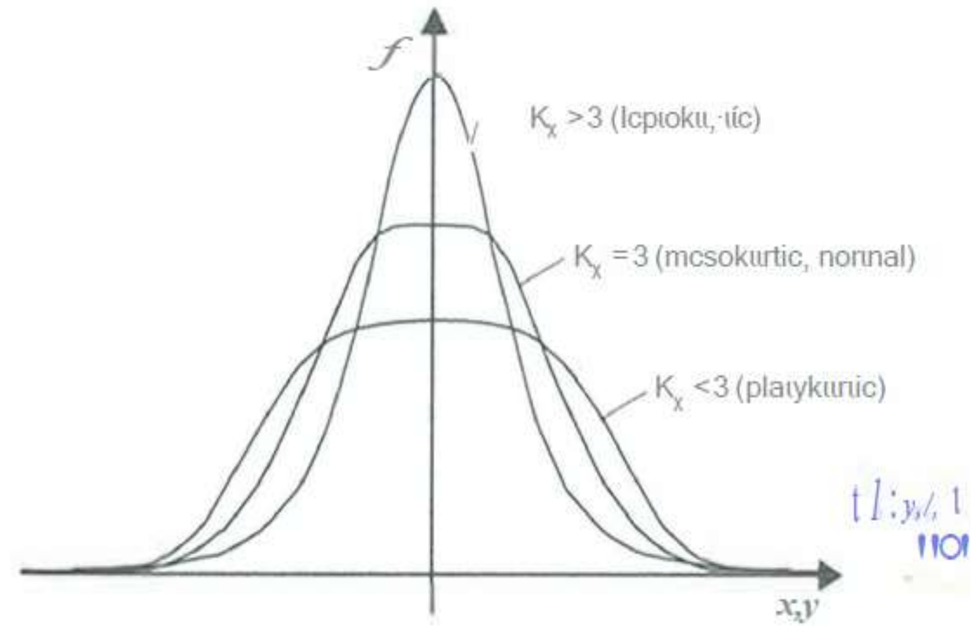
Κύρτωση

$$a_x = E\left[(X - \mu_x)^4\right]$$

$$\downarrow$$

$$K_x = \frac{a_x}{\sigma_x^4}$$ = μέτρο της αιχμηρότητας της θ.σ.π.π.

Κύρτωση



Πίνακας 5.1. Συνήθεις Κατανομές και οι Χαρακτηρίζουσες Παράμετροι

Κατανομή	Συνάρτηση Πυκνότητας Πιθανότητας (ΣΠΠ) ή Μάζας Πιθανότητας (ΣΜΠ)	Παράμετροι	Σχέση με Μέση τιμή και Διασπορά
Διωνυμική	$p_X(x)=\binom{n}{x}p^x(1-p)^{n-x},\ x=0,1,2,\ldots,n$	p	$E(X)=np$ $Var(X)=np(1-p)$
Γεωμετρική	$p_X(x)=p(1-p)^{x-1},\ x=0,1,2,\ldots$	p	$E(X)=1/p$ $Var(X)=(1-p)/p^2$
Poisson	$p_X(x)=\frac{(\nu t)^x}{x!}e^{-\nu t},\ x=0,1,2,\ldots$	ν	$E(X)=\nu t$ $Var(X)=\nu t$
Εκθετική	$f_X(x)=\frac{1}{\lambda}e^{-x/\lambda},\ x\geq 0$	λ	$E(X)=\lambda$ $Var(X)=\lambda^2$
Γάμμα	$f_X(x)=\frac{\nu(\nu x)^{k-1}e^{-\nu x}}{\Gamma(k)},\ x\geq 0$	ν, k	$E(X)=k/\nu$ $Var(X)=k/\nu^2$
Κανονική	$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$ $-\infty<x<\infty$	μ, σ	$E(X)=\mu$ $Var(X)=\sigma^2$
Λογαριθμική Κανονική	$f_X(x)=\frac{1}{\sqrt{2\pi}\zeta x}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\lambda}{\zeta}\right)^2\right]$ $x\geq 0$	λ, ζ	$E(X)=\exp(\lambda+\frac{1}{2}\zeta^2)$ $Var(X)=E^2(X)[e^{\zeta^2}-1]$
Rayleigh	$f_X(x)=\frac{x}{\alpha^2}\exp\left[-\frac{1}{2}\left(\frac{x}{\alpha}\right)^2\right],\ x\geq 0$	α	$E(X)=\sqrt{\frac{\pi}{2}}\alpha$ $Var(X)=\left(2-\frac{\pi}{2}\right)\alpha^2$
Ομοιόμορφη	$f_X(x)=\frac{1}{b-a},\ a<x<b$	a, b	$E(X)=(a+b)/2$ $Var(X)=\frac{1}{12}(b-a)^2$
Τριγωνική	$f_X(x)=\frac{2}{b-a}\left(\frac{x-a}{u-a}\right),\ a\leq x\leq u$ $=\frac{2}{b-a}\left(\frac{b-x}{b-u}\right),\ u\leq x\leq b$	a, b, u	$E(X)=\frac{1}{3}(a+b+u)$ $Var(X)=\frac{1}{18}(a^2+b^2+c^2-ab-au-bu)$
Βήτα	$f_x(x)=\frac{1}{B(q,r)}\frac{(x-a)^{q-1}(b-x)^{r-1}}{(b-a)^{q+r-1}},$ $a\leq x\leq b$	a, b, q, r	$E(X)=a+\frac{q}{q+r}(b-a)$ $Var(X)=\frac{qr}{(q+r)^2(q+r+1)}(b-a)^2$

Οι παραπάνω ροπές προσδιορίζονται από τη θεωρητική κατανομή πιθανότητας ως συνάρτηση των παραμέτρων της κατανομής και…

Σύγκριση αμερόληπτων εκτιμητών και ροπών

- **Εξίσωση**
- των αμερόληπτων εκτιμητών μέσου όρου, τυπικής απόκλισης, συντελεστών κύρτωσης και ασυμμετρίας **του δείγματος,**
- με αυτές όπως προκύπτων **με τη μέθοδο των ροπών** **από τις θεωρητικές κατανομές πιθανότητας** (βλπ. αμέσως προηγούμενες διαφάνειες).

Ανάλυση παράγοντα συχνότητας

Παράγοντας Συχνότητας

- Αρχικά επιλέγεται κατανομή
- O Chow πρότεινε την ***«αντίστροφη πορεία»*** από την πιθανότητα να προσδιοριστεί το κατώφλι της τυχαίας μεταβλητής x_T

$$x_T = \bar{x} + K_T s$$

- όπου

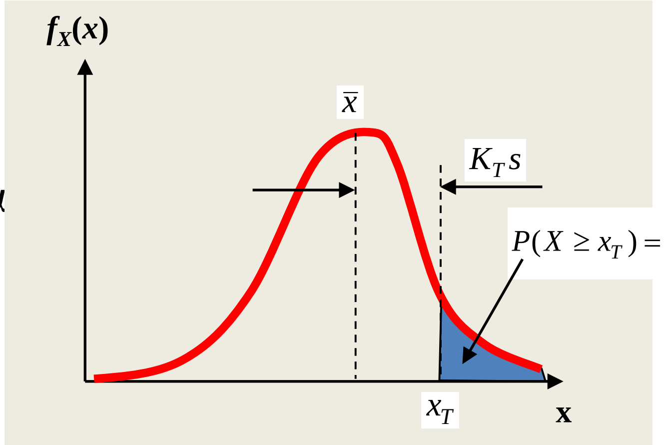
$x_T =$ Εκτιμώμενο κατώφλι

$K_T =$ π.σ με βάση την πιθανότητα και την θεωρητική κα

$T =$ περιόδος επαναφοράς

$\bar{x} =$ μέσος όρος δείγματος

$s = τυπική\ απόκλιση\ δείγματος$



Κανονική κατανομή

- Πυκνότητα πιθανότητας

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$K_T = \frac{x_T - \overline{x}}{s} = z_T$$

- Z=K

$$x_T = \overline{x} + K_T s = \overline{x} + z_T s$$

- Π.Χ.: Περίοδος επαναφοράς μεγίστων 50 χρόνων

$$T = 50;\ p = \frac{1}{50} = 0.02 \rightarrow P\left(Q \geq q\right) = \frac{1}{50} \rightarrow P\left(Q \leq q\right) = 1 - \frac{1}{50} = 0.98$$

$$\rightarrow K_{50} = z_{50} = 2.054$$

Από την Εξ. 2.57 προκύπτει ότι

$$K = \frac{x - \overline{x}}{s} \quad \text{άρα} \quad K = z$$

2.4.2 Κατανομή μεγίστων ακραίων τιμών

$$y=\frac{a+x}{c} \qquad a=\gamma c-\mu \qquad c=\frac{\sqrt{6}}{\pi}\sigma \tag{2.61}$$

$$\chi=\mu+\sigma\left[-\frac{\sqrt{6}}{\pi}(\gamma-y)\right] \tag{2.62}$$

$$y=-\ln\ln(\frac{T}{T-1}) \tag{2.63}$$

$$K=-\frac{\sqrt{6}}{\pi}\left[\gamma+\ln\ln(\frac{T}{T-1})\right] \quad N\rightarrow\infty \tag{2.64}$$

για N πεπερασμένο από την εκτίμηση των παραμέτρων a ; c προκύπτει :

$$a_n=c_n y_n-\bar{x} \qquad c_n=\frac{s}{\sigma_n} \tag{2.65}$$

$$x=\bar{x}+s\left[\frac{y-\bar{y}_n}{\sigma_n}\right] \quad K=-\frac{\ln\ln(\frac{T}{T-1})+y_n}{\sigma_n} \tag{2.66}$$

Έλεγχος αν η θεωρητική κατανομή πιθανότητας που επιλέχθηκε είναι κατάλληλη

Τεστ καταλληλότητας

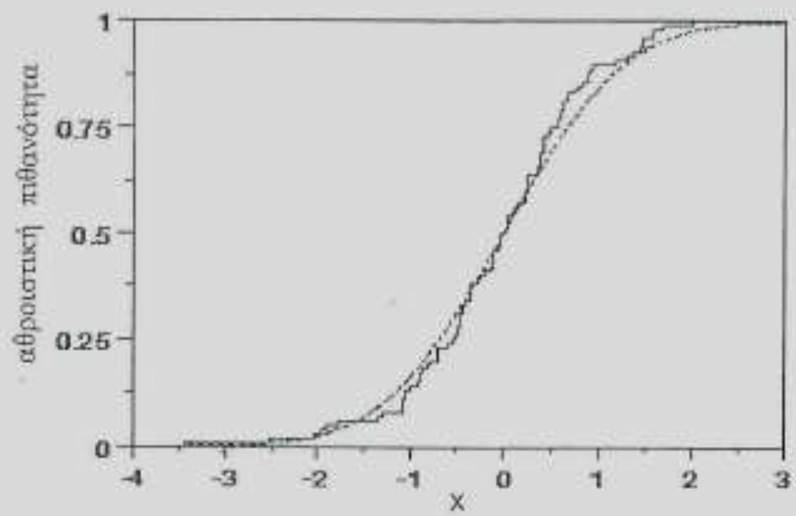
2.6.2 Κριτήριο Kolmogorov – Smirnov

Το κριτήριο αυτό στηρίζεται στην εμπειρική συνάρτηση πιθανότητας E_N που ορίζεται ως εξής

$$E_N = \sum N_i / N \tag{2.78}$$

όπου N ο αριθμός των παρατηρήσεων $x_1, x_2, \ldots\ldots x_N$ και N_i είναι ο αριθμός παρατηρήσεων μικρότερος του x_i όταν τα x_i είναι διαταγμένα κατά αύξουσα σειρά

Στο Σχήμα 2.19 φαίνεται η γραφική παράσταση της σύγκρισης μιας εμπειρικής κατανομής και της κανονικής αθροιστικής κατανομής για 100 τυχαίους αριθμούς. Το κριτήριο Kolmogorov – Smirnov στηρίζεται στη μέγιστη απόκλιση των δύο κατανομών.



Σχ. 2.16 Εμπειρική κατανομή και Κανονική αθροιστική κατανομή

Το κριτήριο αυτό μαθηματικά εκφράζεται ως εξής :

$$D = \max_{i \leq i \leq N} \left| F(x_i) - \sum N_i / N \right| \tag{2.79}$$

όπου η $F(x_i)$ η θεωρητική αθροιστική κατανομή της κατανομής που

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

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

Θεωρητικές κατανομές πιθανοτήτων

Διακριτές Τυχαίες Μεταβλητές

Σε κάθε πείραμα τύχης με μόνο δύο αποτελέσματα, τα οποία λέμε επιτυχία και αποτυχία, με πιθανότητες p και 1-p αντίστοιχα, η συνάρτηση πιθανότητας της τυχαίας μεταβλητής X: "ο αριθμός επιτυχιών σε n επαναλήψεις"

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n.$$

Στην περίπτωση αυτή λέμε ότι η X ακολουθεί διωνυμική κατανομή με παραμέτρους n, p.

Πηγ (Μυλωνάς και Παπαδόπουλος, 2017)

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

Επιτυχία ή αποτυχία

Κατανομή Poison.

1) Διακριτή κατανομή πιθανοτήτων.

2) Διακριτή μεταβλητή X η οποία περιστάνει τον αριθμό εμφάνισης του A στο χρονικό διάστημα t ονομάζεται μεταβλητή Poisson:

$$F(X_t = x) = e^{-\lambda t} \frac{(\lambda t)^x}{x!},$$

$x = 0, 1, 2, \ldots$

λ: μέσος αριθμός εμφάνισης γεγονότος A στη μονάδα του χρόνου.

Ισχύει: $E(X_t) = \lambda t = Var(X_t)$

ροπή πρώτης τάξης

δεύτερης τάξης

Πιθανότητα εμφάνισης του γεγονότος x φορές σε διάστημα t

Συνεχείς Τυχαίες μεταβλητές

Αναφέρθηκαν

- Κανονική
- Λογαριθμοκανονική

Κατανομές Erlang και Γάμμα.

Προκύπτουν από την κατανομή Poisson.

Κατανομή Erlang

Μία τυχαία μεταβλητή X_r, η οποία περιγράφει το χρονικό διάστημα μέχρι την εμφάνιση r γεγονότων σε μια διαδικασία Poisson είναι μια Erlang μεταβλητή με συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \begin{cases} \dfrac{\lambda(\lambda x)^{r-1} e^{-\lambda x}}{(r-1)!}, & x>0 \text{ και } r=1,2,\dots, \\ 0 & \text{αλλού} \end{cases}$$

Παρατηρώ εκεί για $(r-1)!$ μπορεί να χρησιμοποιηθεί η $\Gamma(r)$ συνάρτηση γάμμα.

Επίσης; συμβολισμοί $\quad \theta = r$

$$\lambda = \frac{1}{a}.$$

Γάμμα κατανομή για μη φυσικό αριθμό το r. →

Πιθανότητα με Erlang κατανομή

- Πιθανότητα ο χρόνος μεταξύ r ακριβώς συμβάντων να είναι μικρότερος του x

$$P\left(X_r < x\right) = 1 - \sum_{k=0}^{r-1} \frac{\left(\lambda x\right)^{\kappa}}{\kappa!} e^{-\lambda x} = F_{X_r}\left(x\right)$$

Δι-παραμετρική Γάμμα Κατανομή

Συνάρτηση πυκνότητας πιθανότητας

$$\boxed{f(x) = \frac{\left(\frac{x}{a}\right)^{b-1} \exp\left(-\frac{x}{a}\right)}{a\,\Gamma(b)}}$$

γάμμα συνάρτηση πυκνότητα πιθανότητας.

$a, b > 0.$

Ορισμός: $\Gamma(b) = \int_0^\infty t^{b-1} e^{-t}\, dt$ ← γάμμα συνάρτηση

αν b φυσικός αριθμός $\Gamma(b) = (b-1)!$

Ιδιότητες: $\Gamma(1) = \Gamma(2) = 1$, $\Gamma(b+1) = b\,\Gamma(b)$, $b > 0$.

Εκτίμηση παραμέτρων με τη μέθοδο των ροπών

$$b = \frac{\bar{X}^2}{S^2}, \quad a = \frac{S^2}{\bar{X}} \quad (b > 1.)$$

διαφορετικά →

Excel: GAMMADIST

Στον τύπο του ιδιατού, $C_s = \frac{2}{\sqrt{b}}$

Ρύθμιση Γάμμα Κατανομής:

Προσεγγιστικά: (Thom, 1958)

$$\beta = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right), \quad \alpha = \frac{\bar{x}}{\beta}$$

$$A = \ln(\bar{x}) - \frac{\sum \ln x}{n}$$

↑ σύνολο δεδομένων.

(σε πολλά συγγράμματα ο συμβολισμός α, β, είναι εκ διαμέτρου αντίθετος).

Κατανομή Weibull

→ εφαρμόζεται σε κατανομή που λαμβάνει μόνο θετικές τιμές
→ έχει θετική ασυμμετρία

· Η συνάρτηση πυκνότητας πιθανότητας είναι:

$$f(X=x) = \left(\frac{a}{\beta}\right)\left(\frac{x}{\beta}\right)^{a-1} \cdot \exp\left[-\left(\frac{x}{\beta}\right)^{a}\right]$$

· Η αθροιστική πιθανότητα:

$$P(X \leq x) = \int_0^x f(x)\,dx = 1 - \exp\left[-\left(\frac{x}{\beta}\right)^{a}\right]$$

a: παράμετρος μορφής

β: ίδιες μονάδες με τη μεταβλητή X.

Οι εκτιμητές $\hat{\mu}, \hat{\sigma}$ και a, β είναι περίπλοκες.

Εκτίμηση a,b, Weibull με

γραμμική παλινδρόμηση:

$$\ln\{-\ln(1-P(X\le x))\} = -a\ln b + a\ln x$$

$$\boxed{\Psi = B + Ax}$$

$$\left\{ \text{όπου} \quad \Psi = \ln\{-\ln(1-P(X\le x))\}, \quad x = \ln x \right\}$$

Οπότε τελικά

$$A = a$$

$$b = \exp\left(-\frac{B}{a}\right)$$

Τριπαραμετρική Pearson (α, β, γ)

Συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \frac{1}{\alpha \Gamma(\beta)} \left[\frac{x-\gamma}{\alpha}\right]^{\beta-1} e^{-\left[\frac{x-\gamma}{\alpha}\right]}$$

συνάρτηση γάμμα

ασύμμετρη:

$$\left.\begin{array}{ll} \text{κατώτατο όριο} & x=\gamma \\ \text{άνω όριο} & x=\infty \\ \text{μέγιστη τιμή} & x=\mu-\alpha \end{array}\right\} \text{(Σακκάς, 2007)}$$

για $y = \frac{x-\gamma}{\alpha} \longrightarrow$ Σπη γάμμα 1 παραμέτρου.

όπου: $\Gamma(\beta) = \int_0^\infty t^{\beta-1} e^{-t}\, dt$, συνάρτηση γάμμα.

Με τη μέθοδο των ροπών:

$$\alpha = \frac{1}{2} s C_s, \quad \beta = \left[\frac{2}{C_s}\right]^2, \quad \gamma = \bar{x} - 2\frac{s}{C_s}$$

(Στον κόσμο του ιδεατού $C_s = \frac{2}{\sqrt{\beta}}$)

ξ σε άλλα συγγράμματα

Μέθοδος Παράγοντα Συχνότητας στη Pearson III κατανομή.

- Μέθοδος παράγοντα συχνότητας:

$$X = \mu + K_T \sigma$$

$$\downarrow$$

$$\boxed{X \simeq \bar{X} + K_T \cdot S}$$

K_T: εξαρτάται από την περίοδο επαναφοράς και το είδος της πιθαν/κής κατανομής

("αντίστροφο πρόβλημα")

- Pearson III κατανομή:

$$\lambda = \frac{1}{6} C_s$$

$$\downarrow$$

$$\boxed{K_T = \frac{1}{3\lambda}\left[(1+\lambda z-\lambda^2)^3 - 1\right]}$$

z: τιμή της ανηγμένης μεταβλητής z για την κανονική κατανομή ως συνάρτηση της περιόδου επαναφοράς.

Εκθετική κατανομή:

Προκύπτει από τη διαδικασία Poisson

- Poisson: αριθμός επιτυχιών σε συγκεκριμένο διάστημα t.
- Εκθετική κατανομή: Χρονικό διάστημα μεταξύ δύο επιτυχιών αποτελούν την εκθετική τυχαία μεταβλητή X.
- Συνεχής κατανομή
- ΣΠΠ:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{για } 0 \leq x < \infty \\ 0 & \text{αλλού} \end{cases}$$

πχ Πιθανότητα διάστημα μεταξύ δύο σεισμών να είναι μεγαλύτερο ή ίσο από X έτη

X μέρα.

Εκθετική κατανομή, χρόνος μεταξύ συμβάντων.
γνωστός ο μέσος χρόνος μεταξύ 2 συμβάντων.

Μέσα στο έτος, 110 ανεξάρτητα συμβάντα βροχόπτωσης λαμβάνουν χώρα με μέση διάρκεια 5hr.

• Επομένως, αγνοώντας την εποχικότητα ο μέσος χρόνος μεταξύ καταιγίδων θα είναι:

$$\bar{t} = \frac{\overset{\text{ανά έτος}}{8760hr} - 110 \cdot 5.3}{\underset{\uparrow \text{καταιγίδες}}{110}} = 74.3 \text{ ώρες} \rightarrow$$

$$\rightarrow \lambda = \frac{1}{\bar{t}} = 0.0135\, hr^{-1}.$$

Η πιθανότητα ότι θα μεσολαβήσουν 4 ημέρες $= 96hr$, μεταξύ δύο γεγονότων θα είναι:

$$Prob(t \geq 96) = 1 - F(96) =$$

$$= e^{-0.0135 \cdot 96} = 0.27.$$

Ποια είναι η πιθανότητα ότι μεταξύ 2 καταιγίδων θα μεσολαβήσουν ακριβώς 12?

$$Prob(t = 12) = 0. \ (\overset{\text{πιθανότητα}}{\text{συνεχούς μεταβλητής}})$$

$$Prob(t \leq 12) = F(12) = 1 - e^{-0.0135 \cdot 12} = 1 - 0.85 = 0.15$$