ΥΔΡΟΣΚΟΠΙΟ

ΠΡΟΓΡΑΜΜΑ STRIDE ΕΛΛΑΣ

ΔΗΜΙΟΥΡΓΙΑ ΕΘΝΙΚΗΣ ΤΡΑΠΕΖΑΣ
ΥΔΡΟΛΟΓΙΚΗΣ ΚΑΙ ΜΕΤΕΩΡΟΛΟΓΙΚΗΣ
ΠΛΗΡΟΦΟΡΙΑΣ

ΕΘΝΙΚΗ ΜΕΤΕΩΡΟΛΟΓΙΚΗ ΥΠΗΡΕΣΙΑ

HELLENIC NATIONAL METEOROLOGICAL SERVICE

ΣΧΕΔΙΑΣΜΟΣ ΣΤΑΤΙΣΤΙΚΗΣ ΕΠΕΞΕΡΓΑΣΙΑΣ ΔΕΔΟΜΕΝΩΝ

DEVELOPMENT OF TECHNIQUES FOR STATISTICAL DATA ANALYSIS

Θ. Χαραντώνης, Γ. Σακελλαρίδης και Ι. Παπαγεωργίου

Th. Charantonis, G. Sakellarides and J. Papageorgiou

Αριθμός τεύχους 5/10
Report number

ΑΘΗΝΑ - ΙΟΥΛΙΟΣ 1993
ATHENS - JULY 1993

HYDROSCOPE

STRIDE HELLAS PROGRAMME

DEVELOPMENT OF A NATIONAL DATA
BANK FOR HYDROLOGICAL AND
METEOROLOGICAL INFORMATION

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ

Στην αρχή δίνονται οι ορισμοί των στατιστικών παραμέτρων και παρουσιάζεται η βασική θεωρία της στατιστικής επεξεργασίας μετεωρολογικών δεδομένων. Στη συνέχεια περιγράφεται ο τρόπος στατιστικής επεξεργασίας κάθε μιας μετεωρολογικής παραμέτρου ανάλογα με τη χρονική διάρκεια (απλές, μηνιαίες και ετήσιες τιμές). Η στατιστική επεξεργασία των δεδομένων του ανέμου γίνεται ξεχωριστά λόγω της ιδιαιτερότητάς τους.

ABSTRACT

Firstly, the appropriate definitions of statistical parameters are given and the basic theory of statistical manipulation of meteorological data is presented. Continuously, the methods of statistical calculations for each meteorological parameter are described according to the period of analysis (single, monthly and annual values). The statistical calculations for the wind are presented separately because of the different type of data.

1. ΕΙΣΑΓΩΓΗ

Καθημερινά κάθε μετεωρολογικός σταθμός καταγράφει τιμές διάφορων παραμέτρων που επηρεάζουν τον καιρό και το κλίμα. Η καταγραφή μπορεί να είναι συνεχής (αυτογραφικά όργανα), ή/και να λαμβάνει χώρα ορισμένες χρονικές στιγμές (με όργανα αμέσης ανάγνωσης ή και χωρίς όργανα με εκτίμηση της τιμής ορισμένων παραμέτρων από τον παρατηρητή). Η συχνότητα των παρατηρήσεων και οι παράμετροι που μετρούνται εξαρτάται από το είδος του μετεωρολογικού σταθμού (συνοπτικός ή κλιματολογικός ή αγρομετεωρολογικός κλπ.).

Ετσι, ένας τεράστιος αριθμός δεδομένων προστίθενται κάθε μέρα στα, ήδη μεγάλα, αρχεία των μετεωρολογικών Υπηρεσιών όλου του κόσμου. Οι παρατηρήσεις αυτές χρησιμοποιούνται:

α. 'Αμεσα.

Από τους προγνώστες μετεωρολόγους, τους αγρομετεωρολόγους κλπ, για να έχουν μια σαφή εικόνα του καιρού που επικρατεί στην χώρα τους και να παίρνουν τα κατάλληλα μέτρα όπου και όταν χρειαστεί, ενώ ένα μέρος απ'αυτές αποτελούν τις αρχικές συνθήκες που απαιτού-απαιτούνται στα προγνωστικά μοντέλα καιρού.

β. 'Εμεσα.

Από τους κλιματολόγους υδρολόγους κλπ. για να βγάλουν διάφορες στατιστικές παραμέτρους (όπως κλιματολογικές, υδρολογικές, αγρομετεωρολογικές κλπ.) και απ'αυτές χρήσιμα συμπεράσματα για το κλίμα της μελετούμενης περιοχής, σχέσεις παραμέτρων στον ίδιο ή διαφορετικούς σταθμούς κλπ.

Τις μετεωρολογικές παραμέτρους, ανάλογα με τον χρόνο και τον τρόπο που γίνεται η παρατήρηση τους, μπορούμε να τις κατατάξουμε στις παρακάτω κατηγορίες:

<u>1. Ανάλογα με τον χρόνο.</u>

1.α. Σε συνεχή χρόνο.

Στην κατηγορία αυτή ανήκουν οι παράμετροι που μετρούνται με καταγραφικά αυτογραφικά όργανα (θερμογράφοι, βαρογράφοι, βροχογράφοι, κλπ.).

1.β. Σε διακριτό ή ασυνεχή χρόνο.

Η κατηγορία αυτή μπορεί να χωριστεί σε:

1.β.1. Σταθερής "περιόδου δείγματος".

Εδώ ανήκουν οι παράμετροι που μετρούνται σε σταθερά χρονικά διαστήματα, π.χ. μισής ώρας ή τριών ωρών κλπ., με άμεση ανάγνωση των καταλλήλων οργάνων (θερμόμετρα, βαρόμετρα, βροχόμετρα κλπ.) ή με εκτίμηση (νέφωση, ορατότητα κλπ.).

1.β.2. Μη σταθερής περιόδου.

Εδώ ανήκουν οι παράμετροι που η μέτρηση ή εκτίμησή τους γίνεται μόνο όταν παραστεί ανάγκη ή όταν οι παράμετροι εκφράζουν κάποιο καιρικό ή φαινολογικό φαινόμενο (όπως αστραπή, βροντή κλπ.).

2. Ανάλογα με τον τρόπο.

2.α. Συνεχής.

Είναι αυτές που μετρούνται με καταγραφικά όργανα και οι τιμές τους είναι πραγματικοί αριθμοί, δηλαδή το σύνολο των τιμών τους είναι ένα διάστημα του συνόλου R των πραγματικών αριθμών.

2.β. Ασυνεχείς ή διακριτές.

2.β.1. Μετρούμενες.

Είναι οι παράμετροι των οποίων η τιμή λαμβάνεται με όργανα άμεσης ανάγνωσης και το σύνολο των τιμών τους είναι ένα πεπερασμένο υποσύνολο του συνόλου Q των ρητών αριθμών, που εξαρτάται από την απαιτούμενη ακρίβεια και την ακρίβεια των χρησιμοποιούμενων οργάνων. Στην μετεωρολογία για την θερμοκρασία η απαιτούμενη ακρίβεια είναι της τάξης του 0.1°C πίεσης 0.1hPa κλπ. Ετσι ακόμα και όταν στο θερμόμετρο μπορούσαμε να διαβάσουμε π.χ. την τιμή 9.25°C, αυτή θα αναγραφόταν σαν 9.2°C.

2.β.2. Εκτιμώμενες.

Στην κατηγορία αυτή ανήκουν οι παράμετροι των οποίων οι τιμές εκτιμούνται από τον παρατηρητή, όπως η ποσότητα και το είδος των νεφών, η ορατότητα, το ύψος κύματος κλπ. Το σύνολο των τιμών στην κατηγορία αυτή είναι κυρίως ένα πεπερασμένο υποσύνολο του συνόλου N των φυσικών αριθμών. Στην κατηγορία αυτή μπορούν να προστεθούν και παράμετροι των οποίων η τιμή είναι ένας αριθμός που δείχνει αν το φαινόμενο έλαβε χώρα ή όχι (π.χ. αστραπή). Επειδή ακόμα και στην περίπτωση των παραμέτρων που οι τιμές τους καταγράφονται με αυτογραφικά όργανα (συνεχής και σε συνεχή χρόνο) η ανάγνωση των τιμών γίνεται είτε χειρονακτικά (πολύ μικρή "διακριτική ικανότητα") είτε ηλεκτρονικά με Scanner (μεγάλη αλλά πεπερασμένη "διακριτική ικανότητα" τόσο ως προς τον χρόνο της ανάγνωσης όσο και ως προς την τιμή), θεωρούμε ότι οι μετεωρολογικές παράμετροι είναι διακριτές σε ασυνεχή χρόνο, γεγονός που διευκολύνει σημαντικά την στατιστική επεξεργασία τους. Ακόμα, οι πλήθους M τιμές (X,Y) μιας παραμέτρου Y του σταθμού X μπορούν να θεωρηθούν σαν τα αποτελέσματα ενός πειράματος που επαναλαμβάνεται M φορές στο σταθμό X, για την μέτρηση των τιμών της παραμέτρου Y. Το μεγάλο πλήθος των τιμών καθιστά την στατιστική επεξεργασία τους απαραίτητη.

Με την στατιστική επεξεργασία απ'όλο τον όγκο των πληροφοριών γίνεται προσπάθεια υπολογισμού των βασικών στατιστικών παραμέτρων και των χαρακτηριστικών τιμών που η παρουσίαση τους υπό μορφή πινάκων ή διαγραμμάτων, (πολύ μικρότερου όγκου, από τον όγκο που έχουν όλες οι τιμές) επιτρέπουν την εξαγωγή συμπερασμάτων για την παράμετρο (ή τις παραμέτρους) που εξετάζεται (ονται) με τρόπο που να εξυπηρετεί τόσο τις ανάγκες τις μετεωρολογίας, όσο και ένα μεγάλο πλήθος χρηστών και ερευνητών.

Η εργασία αυτή γράφεται με γνώμονα να περιγράψει την στατιστική επεξεργασία που πρέπει να γίνει στις τιμές των μετεωρολογικών παραμέτρων και για να βοηθήσει την ομάδα σύνταξης των προγραμμάτων της στατιστικής επεξεργασίας, ώστε τα αποτελέσματα να είναι εύκολο να χρησιμοποιηθούν από τους διάφορους χρήστες και να συμβαδίζουν με τα διεθνή πρότυπα. Ανάλογα λογισμικά έχουν αναπτυχθεί στην ΕΜΥ από τον μετεωρολόγο κ. Γ. Κορνάρο για τις "κλιματολογικές" παραμέτρους όλων των Μετεωρολογικών παραμέτρων, από τον μετεωρολόγο κ. Α. Νιάνιο για τις "υδρομετεωρολογικές" παραμέτρους και από τους μετεωρολόγους κ.κ. Ι. Βερόπουλο και Θ.Χαραντώνη για τις "αγρομετεωρολογικές" παραμέτρους.

2. ΔΕΙΓΜΑΤΙΚΟΣ ΧΩΡΟΣ - ΓΕΓΟΝΟΤΑ - ΔΙΑΜΕΛΙΣΜΟΣ - ΣΥΧΝΟΤΗΤΕΣ - ΙΣΤΟΓΡΑΜΜΑ - ΠΙΘΑΝΟΤΗΤΕΣ

2.1 Δειγματικός χώρος.

Δειγματικός χώρος μιας παραμέτρου Χ είναι το σύνολο S_x όλων των δυνατών τιμών που μπορεί να πάρει η παράμετρος.

Δειγματικός χώρος μιας παραμέτρου Χ, σ'ένα πείραμα ποσοτικών ή ποιοτικών μετρήσεών της, καλείται το σύνολο S_x όλων των τιμών που πήρε η παράμετρος κατά την διάρκεια του πειράματος.

Αν το πλήθος των στοιχείων του S_x είναι πεπερασμένο τότε η παράμετρος ονομάζεται πεπερασμένη διακριτή. Αν το S_x είναι αριθμήσιμο (ισοδύναμο με το σύνολο N των φυσικών αριθμών) ονομάζεται άπειρη διακριτή και αν το S_x είναι ισοδύναμο με το σύνολο των πραγματικών αριθμών R λέγεται συνεχής. Οι πεπερασμένες και οι άπειρες διακριτές τυχαίες μεταβλητές θα ονομάζονται απλά <u>διακριτές</u>.

Το σύνολο όλων των υποσυνόλων ενός συνόλου Α (δυναμοσύνολο του Α) παριστάνεται με 2^A. Αν το πλήθος των στοιχείων του Α είναι cardA=N τότε το πλήθος των στοιχείων του 2^A είναι ίσο με $card 2^A=2^N$.

<u>Παράδειγμα 2.1.</u>

Αν η παράμετρος Χ είναι η θερμοκρασίαΤ, τότε σαν S_x μπορούμε να θεωρήσουμε το σύνολο-διάστημα

$S_x=(-273.16°C\ ,\ +\infty)$

αφού η θερμοκρασία μπορεί να πάρει τιμές μεγαλύτερες από -273.16°C (απόλυτο μηδέν).

<u>Παράδειγμα 2.2.</u>

Αν μετρείται η μέγιστη ημερήσια θερμοκρασία σ'ένα σταθμό και κατά την διάρκεια των μετρήσεων βρέθηκε ελάχιστη τιμή της η -10°C και μεγίστη τιμή της η +50°C τότε ο δειγματικός χώρος της παραμέτρου αυτής, στο σταθμό αυτό και για την συγκεκριμένη χρονική περίοδο διάρκειας των μετρήσεων, θεωρούμε το σύνολο:

$S=[-10,50]$

<u>Παράδειγμα 2.3.</u>

Είναι γνωστό ότι η νέφωση Ν εκτιμάται από τον παρατηρητή σε όγδοα (μερικές φορές σε δέκατα) καλυμμένου ουρανού και οι τιμές που παίρνει είναι οι 0,1,2,3,4,5,6,7,8,9 (9 όταν ο ουρανός είναι αόρατος π.χ. σε περίπτωση ομίχλης). 'Ετσι ο δειγματικός χώρος που θα πάρουμε για την νέφωση θα είναι ο:

$S=\{0,1,2,3,4,5,6,7,8,9\}$

2.2. Γεγονότα ή συμβάντα

Κάθε υποσύνολο A του δειγματικού χώρου S ($A\in 2^S$) ονομάζεται γεγονός ή/και συμβάν.

Οταν: A=S το A λέγεται βέβαιο γεγονός,

αν $A=\emptyset$ ($\emptyset$ το κενό σύνολο) τότε το A ονομάζεται αδύνατο γεγονός, και τέλος,

αν $A=\{\alpha\}$ (μονοσύνολο) το A λέγεται απλό γεγονός.

Παράδειγμα 2.4.

Αν S=[10,20] τότε τα σύνολα: $A_1=\emptyset$, $A_2=\{15\}$, $A_3=\{18\}$, $A_4=(12,15)$ είναι γεγονότα του S (τα A_2 και A_3 απλά γεγονότα).

Αν A,B είναι δύο γεγονότα του S και $A\cap B=A.B=\emptyset$ τότε τα γεγονότα αυτά ονομάζονται ασυμβίβαστα ή ξένα μεταξύ τους γεγονότα.

2.3. Τυχαία μεταβλητή.

Πολλές φορές τα αποτελέσματα ενός πειράματος μπορεί να είναι ποιοτικά ή πολύ μεγάλοι αριθμοί που είναι δύσκολοι στο χειρισμό τους. Ετσι για παράδειγμα το είδος των νεφών (Stratus,Cumulus κλπ.) είναι ένα ποιοτικό αποτέλεσμα, η αντιστοίχηση δε κάποιου αριθμού στο καθ'ένα απ'αυτά, διευκολύνει την επεξεργασία τους. Εστω λοιπόν μια παράμετρος A, S ο δειγματικός της χώρος και 2^S το σύνολο όλων των υποσυνόλων του S (δυναμοσύνολο του S).

Κάθε συνάρτηση $X:2^S \longrightarrow R_X$, όπου $R_X \subset R$ ονομάζεται τυχαία μεταβλητή. Είναι λοιπόν η τυχαία μεταβλητή μια συνάρτηση με πεδίο ορισμού το σύνολο των γεγονότων και πεδίο τιμών κάποιο υποσύνολο R_X του συνόλο R των πραγματικών αριθμών. Κάθε υποσύνολο $B\subset R_X$ είναι εικόνα κάποιου (ή κάποιων) συνόλου $Y\in 2^S$, δηλαδή όταν δοθεί ένα υποσύνολο B του R_X μπορούμε να βρούμε ένα υποσύνολο Y του 2^S έτσι ώστε X(Y)=B.

Αν οι τιμές της παραμέτρου A είναι πραγματικοί αριθμοί και θέσουμε $\chi=\{\chi\}$ τότε η συνάρτηση $X(\{x\})=X(\chi)=x$ είναι μία τυχαία μεταβλητή (η ταυτοτική συνάρτηση).

Παράδειγμα 2.5.

Οι τιμές της πίεσης P στην επιφάνεια της θάλασσας στα πλάτη μας κυμαίνεται μεταξύ των τιμών 950hPa και 1060hPa. Η συνάρτηση X(P)=P-950 είναι μια τυχαία μεταβλητή με πεδίο τιμών το σύνολο $R_X=[0,110]$, που περιέχει αριθμούς που είναι πιο εύκολοι στον χειρισμό τους από τους αντίστοιχους αριθμούς του S=[950,1060].

Παράδειγμα 2.6.

Μία μέρα χαρακτηρίζεται σαν βροχερή όταν κατά την διάρκειά της έβρεξε, σαν ημέρα χιονιού αν κατά την διάρκειά της σημειώθηκε χιονόπτωση. Αν θέλαμε να σχηματίσουμε τον δειγματικό χώρο ενός πειράματος στο οποίο ζητάμε τον αριθμό των βροχερών ημερών ή των ημερών χιονιού αυτός θα έπρεπε να είναι:

S={ημέρα χωρίς βροχή ή χιόνι, βροχερή ημέρα, ημέρα χιονιού, ημέρα που σημειώθηκε και βροχή και χιόνι}.

Η παράσταση των στοιχείων του S είναι αρκετά εκτεταμένη. Αν θεωρήσουμε την συνάρτηση:

X:S—>R με:

X(ημέρα χωρίς βροχή ή χιόνι)=0

X(βροχερή ημέρα)=1

X(ημέρα χιονιού)=2

X(ημέρα που σημειώθηκε και βροχή και χιόνι)=3

τότε ο δειγματικός χώρος X(S) που αντιστοιχεί στον S είναι:

X(S)={0,1,2,3}

που είναι πιο εύχρηστος από τον ίδιο τον S.

2.4. Διαμελισμοί συνόλων.

2.4.1. Διαμελισμός συνόλου.

'Εστω A ένα διάφορο του κενού σύνολο ($A \neq \emptyset$) και

$Q=\{A_1, A_2, \ldots A_\kappa\}$

ένα σύνολο κ υποσυνόλων του A.

Αν η ένωση των συνόλων-στοιχείων του Q είναι ίση με το A τότε λέμε ότι το Q καλύπτει το A ή ότι το Q είναι μια κάλυψη του A.

Επομένως:

(Το $Q=\{A_1, A_2, \ldots A_\kappa\}$ με $A_i \subset A$, $i \in I_\kappa$, καλύπτει το A) $\Longleftrightarrow$ ($\cup^\kappa_{i=1} A_i = A$).

Αν ακόμα τα σύνολα του Q είναι ξένα μεταξύ τους, δηλαδή αν:

$A_i \cap A_j = A_i . A_j = \emptyset$ όταν $i \neq j$

τότε λέμε ότι το Q αποτελεί ένα διαμελισμό του A.

Αν υποθέσουμε τώρα ότι το A είναι ο δειγματικός χώρος S μιας παραμέτρου X και $Q=\{A_1, A_2, \ldots A_\kappa\}$ ένας διαμελισμός του S τότε τα στοιχεία του Q είναι γεγονότα του S.

<u>Παράδειγμα 2.7.</u>

Αν:

S=[-10,35]

ο δειγματικός χώρος της ημερήσιας ελάχιστης θερμοκρασίας ενός σταθμού τότε τα σύνολα:

$Q_1=\{A_1=[-10,10], A_2=[10,35]\}$

$Q_2=\{A_1=[-10,-5), A_2=[-5,0], A_3=[0,5], A_4=[5,15], A_5=[15,25], A_6=[25,35]\}$

αποτελούν διαμελισμούς του S και το σύνολο:

$Q_3=\{A1=[-10,20], A2=[10,35]\}$

αποτελεί κάλυψη του S όχι όμως διαμελισμό του.

<u>Παράδειγμα 2.8.</u>

Αν σταθμός λειτούργησε για n χρόνια και X η τυχαία μεταβλητή με τιμές τις τιμές μιας παραμέτρου που σημειώθηκαν στις διάφορες ημερομηνίες και ώρες των παρατηρήσεων τότε το σύνολο:

$Q=\{A_1,A_2,...A_n\}$

με:

$A_\kappa=\{x/x$ τιμές της παραμέτρου κατά τις ημερομηνίες και ώρες παρατήρησης του έτους κ, $\kappa \in I_n\}$

αποτελεί ένα διαμελισμό των τιμών της X.

<u>Παράδειγμα 2.9.</u>

Αν το σύνολο των τιμών μιας παραμέτρου X είναι πεπερασμένο, τότε και ο δειγματικός της χώρος S θα είναι πεπερασμένος. Εστω ότι:

$S=\{\alpha_1,\alpha_2,...\alpha_m\}$

τότε το σύνολο

$Q=\{\{\alpha_1\},\{\alpha_2\},...,\{\alpha_m\}\}$

είναι ένας διαμελισμός του S.

Για τις μετεωρολογικές παραμέτρους ο παραπάνω διαμελισμός παίζει σπουδαίο ρόλο αφού όπως αναφέρθηκε στην εισαγωγή, οι μετεωρολογικές παράμετροι είτε μετρούνται με όργανα άμεσης ανάγνωσης είτε με αυτογραφικά όργανα, μπορούν να θεωρηθούν σαν διακριτές.

Το πλήθος κ των συνόλων (που λέγονται και κλάσεις) από τα οποία πρέπει να αποτελείται ένας διαμελισμός, για την στατιστική επεξεργασία μιας τυχαίας μεταβλητής, είναι συνάρτηση του πλήθους N των διατιθεμένων τιμών της παραμέτρου και δίνεται από τον τύπο:

$\kappa=5\lambda o\gamma N$

ενώ για ισομήκεις κλάσεις μπορεί να χρησιμοποιείται (Ψώινος 1978) και ο τύπος:

$\kappa=1+3.3\lambda o\gamma N$

2.4.2. Διαδοχικός διαμελισμός.

Εστω A ένα σύνολο και

$Q=\{A_1,,A_2,...A_\kappa\}$

ένας διαμελισμός του A σε κ υποσύνολά του.

Θεωρούμε ότι το σύνολο A_m, $m\in I_\kappa$ διαμελίζεται με τον διαμελισμό

$Q_m=\{A_m(1),A_m(2),...A_m(\lambda_m)\}$

με $\lambda_m>1$ για τουλάχιστον ένα από τα $m\in I_\kappa$.

Το σύνολο:

$$
\begin{aligned}
QQ=\{\ & A_1(1),\ A_1(2),...A_1(\lambda 1),\\
& A_2(1),\ A_2(2),...A_2(\lambda 2),\\
& ....,\\
& A_\kappa(1),\ A_\kappa(2),...A_\kappa(\lambda_\kappa)\ \}
\end{aligned}
$$

αποτελεί επίσης ένα διαμελισμό του A που ονομάζεται διαδοχικός διαμελισμός του συνόλου A ως προς τον διαμελισμό Q ή απλά διαδοχικός διαμελισμός του A.

Παράδειγμα 2.10.

Στο παράδειγμα 2.8 για κάθε σύνολο A_κ, $\kappa \in I_n$ θεωρούμε το σύνολο:

$Q_\kappa=\{A_\kappa(1),A_\kappa(2),\ldots A_\kappa(12)\}$

όπου $A_\kappa(i)=\{$τιμές της παραμέτρου κατά τις ημερομηνίες του ίδιου έτους κ και του ίδιου μήνα $i\in I_{12}\}$.

Το σύνολο:

$QQ=\{A_\kappa(i)/\kappa\in I_n,\ i\in I_{12}\}$

αποτελεί ένα διαμελισμό του συνόλου των τιμών της X κατά τις ημερομηνίες και ώρες παρατήρησης, που είναι ένας διαδοχικός διαμελισμός του διαμελισμού Q.

2.4.3. Συνδιαμελισμός.

Εστω A και B δύο σύνολα και Q_A, Q_B διαμελισμοί των συνόλων A και B αντίστοιχα. Το καρτεσιανό γινόμενο:

$Q_A * Q_B = Q$

είναι ένας διαμελισμός του καρτεσιανού γινομένου A*B που ονομάζεται **συνδιαμελισμός των συνόλων A και B, αντίστοιχος στους Q_A και Q_B** ή απλά συνδιαμελισμός των A και B.

Παράδειγμα 2.11.

Είναι γνωστό ότι πολλά άτομα υποφέρουν στον "καύσωνα". Αν σαν μέρα καύσωνα σ'ένα τόπο ορίσουμε την ημέρα κατά την οποία η μέγιστη θερμοκρασία ξεπερνά τους 36.5°C και η ελάχιστη θερμοκρασία είναι μεγαλύτερη από τους 28°C, τότε το σύνολο:

Γ={ημέρα καύσωνα στον σταθμό X}

είναι ίσο με:

Γ={ημέρες με $Tmax \geq 36.5$°C και $Tmin \geq 28$°C}.

Θεωρούμε τα σύνολα:

S={'Ολες οι ημέρομηνίες των μετρήσεων της θερμοκρασίας}

A={ημέρα με $Tmax \geq 36.5$°C},

C_SA={ημέρα με $Tmax < 36.5$°C}

B={ημέρα με $Tmin \geq 28$°C}

C_SB={ημέρα με $Tmin < 28$°C}

Τα σύνολα:

$Q_1=\{A, C_SA\}$

και

$Q_2=\{B, C_SB\}$

αποτελούν διαμελισμό του συνόλου S αφού προφανώς:

$S=A\cup C_SA$, $S=B\cup C_SB$, $A\cap C_SA=A.C_SA=\emptyset$ και $B\cap C_SB=B.C_SB=\emptyset$

Το σύνολο $Q=Q_1*Q_2$ αποτελεί συνδιαμελισμό του συνόλου $S*S=S^2$.

Παράδειγμα 2.12.

Οι ώρες πραγματοποίησης των "συνοπτικών παρατηρήσεων" σ'ένα συνοπτικό σταθμό είναι τα στοιχεία του συνόλου:

HOSO={00UTC, 03UTC, 06UTC, 09UTC,12UTC, 15UTC, 18UTC, 21UTC}

που τα αντιστοιχούμε στους 8 αριθμούς $i\in I_8$ που δίνονται από τον τύπο:

$i=1+IW/3$

όπου $i \in HOSO$ η ώρα συνοπτικής παρατήρησης.

Βτους συνοπτικούς σταθμούς εκτός των άλων μετεωρολογικών παραμέτρων μετρείται και η διεύθυνση DD και η ένταση FFF του ανέμου. Για την DD δεχόμαστε την κωδικοποίηση:

DD=1 όταν ο άνεμος είναι βόρειος (διεύθυνση N)
DD=2 " " " βορειοανατολικός (διεύθυνση NE)
DD=3 " " " ανατολικός (διεύθυνση E)
DD=4 " " " νοτιοανατολικός (διεύθυνση SE)
DD=5 " " " νότιος (διεύθυνση S)
DD=6 " " " νοτιοδυτικός (διεύθυνση SW)
DD=7 " " " δυτικός (διεύθυνση W)
DD=8 " " " βορειοδυτικός (διεύθυνση NW)
DD=9 όταν ο άνεμος είναι μεταβλητός ή επικρατεί άπνοια

και για την ένταση FFF (σε Beaufort B) την παρακάτω κωδικοποίηση:

FFF=1 όταν επικρατεί άπνοια
FFF=2 όταν $0<FFF<5$
FFF=i όταν $4<i<7$
FFF=7 όταν $6<FFF<9$
FFF=8 όταν $FFF>8$.

Θεωρούμε τα σύνολα:

$Q_{DD}=\{A_1,A_2..,...A_9\}$ με $A_i=\{DD/DD=i\}$

$Q_{FFF}=\{B_1,B_2,...,B_8\}$ με $B_i=\{FFF/FFF=i\}$

και

$Q_T=\{T_1,T_2,...,T_8\}$

όπου:

$T_i=\{$παρατηρήσεις της i ώρας$\}$.

Το σύνολο:

$Q=Q_T*Q_{DD}*Q_{FFF}$

είναι ένας συνδιαμελισμός του συνόλου των παρατηρήσεων του σταθμού, που βασίζεται στην ώρα παρατήρησης, την διεύθυνση και την ένταση του ανέμου.

2.5. <u>Τάξη-διατεταγμένη τάξη, των στοιχείων ενός διατάξιμου συνόλου.</u>

'Ενα σύνολο λέγεται <u>διατάξιμο</u> αν, με κάποιο τρόπο, μπορούμε να κατατάξουμε τα στοιχεία του ανάλογα με το μέγεθός τους σε αύξουσα ή φθίνουσα σειρά.

Εστω το σύνολο $A=\{a_1,a_2,...a_n\}$. Το στοιχείο a_i, $i\in I_n$ λέμε ότι καταλαμβάνει την i θέση ή την i τάξη. Αν οι a_i, $i\in I_n$ είναι πραγματικοί αριθμοί, έστω

$\omega_1=\min\{a_i,\ i\in I_n\}$, $\omega_2=\min\{a_i$ με $a_i\neq\omega_1\}$,...,

$w_k=\min\{a_i$ με $a_i\neq\omega_\lambda,\ \lambda\in I_{k-1}\}$, $w_n=\max\{a_i,\ i\in I_n\}$

Προφανώς $\omega_1<\omega_2<...<\omega_n$ και για κάθε δείκτη $i\in I_n$ υπάρχει μοναδικός δείκτης $j\in I_n$ έτσι ώστε: $\omega_i=a_j$, ενώ ακόμα ισχύει:

$A=\{a_1, a_2, ... a_n\}=\{\omega_1, \omega_2, ... \omega_n\}$.

Ο αριθμός i για τον οποίο ισχύει:

$\omega_i = a_j$

λέγεται διατεταγμένη τάξη του a_j όταν τα στοιχεία του συνόλου A διατάσσονται κατά αύξουσα σειρά. Επειδή τα στοιχεία του A, όταν αυτό γραφεί με την μορφή:

$A=\{\omega_n, \omega_{n-1}, ..., \omega_1\}$

είναι διατεταγμένα με αντίθετη σειρά διάταξης από την προηγούμενη (δηλαδή φθίνουσα), ο αριθμός $\kappa=n+1-i$ είναι διατεταγμένη τάξη του a_j όταν τα στοιχεία του A διατάσσονται κατά φθίνουσα σειρά.

Παράδειγμα 2.13.

Εστω $A=\{15,12,17,0\}$. Η τάξη του 15 είναι 1 αυτή του 12 η 2, του 17 η 3 και του 0 η 4. Αν τα στοιχεία του A τα κατατάξουμε κατά σειρά αύξοντος μεγέθους τότε το 0 θα κατέχει την τάξη 1 το 12 την 2, τον 15 την 3 και το 17 την 4, που εί ναι οι διατεταγμένη τάξη των στοιχείων του A σε αύξουσα διάταξη.

Κατά την διαδικασία των μετρήσεων των μετεωρολογικών παραμέτρων, μια τιμή μιας παραμέτρου μπορεί να εμφανιστεί περισσότερες από μία φορές σε διαφορετικές όμως χρονικές στιγμές.

Εστω $S=\{a_1,, a_2, ... a_n\}$ ο δειγματικός χώρος μιας διακριτής παραμέτρου X και $\kappa_1, \kappa_2 .., .. \kappa_n$ το πλήθος των μετρήσεων κατά τις οποίες η τιμή της παραμέτρου X ήταν a_i, $i \in I_m$. Αν μετά την διάταξη των $a_1, a_2, ... a_n$, κατά αύξουσα (ή φθίνουσα) τάξη μεγέθους προκύψει ότι:

$S=\{\omega_1, \omega_2, ... \omega_n\}$ με $\omega_i < \omega_j$, $i<\zeta$

($\omega_i > \omega_j$, $i<s$ για φθίνουσα διάταξη)

και αν $a_\kappa = \omega_\lambda$ (λ η τάξη του a_κ), τότε αν γράψουμε τις πλήθους:

$M = \Sigma^{n}_{i=1} \kappa_i$

τιμές που μετρήθηκαν κατά την διάρκεια λειτουργίας του σταθμού, κατά τάξη αύξοντος (ή φθίνοντος) μεγέθους, αυτές θα έχουν την μορφή:

$$\begin{array}{ccccccc} \lambda_1 & & \lambda_2 & & \cdots & & \lambda_n \\ \omega_1, \omega_1, \ldots \omega_1, & & \omega_2, \omega_2, \ldots \omega_2, & & \ldots, & & \omega_n, \omega_n, \ldots \omega_n \\ 1 \quad 2 \quad\quad \lambda_1 & & \lambda_1+1 \;\; \lambda_1+2 \ldots \lambda_1+\lambda_2 & & \ldots & & \lambda_1+\lambda_2+\ldots\lambda_{n-1}+1, \ldots, M \end{array}$$

όπου $\lambda_i = \kappa_j$ το πλήθος των τιμών του στοιχείου $\omega_\lambda = a_\mu$ Τον μέσο όρο των λ_i αριθμών, από $\lambda_1+\lambda_2+...+\lambda_{i-1}+1=i_1$ μέχρι τον αριθμό $i_2=i_1+\lambda_i$ τον ονομάζουμε μέση τάξη του στοιχείου ω_i.

Παράδειγμα 2.14.

Εστω οι δέκα (M=10) τιμές της ελάχιστης θερμοκρασίας το πρώτο δεκαήμερο του έτους ήταν: 2,5,2,4,5,5,6,7,6,3. Ο δειγματικός χώρος για το δεκαήμερο αυτό είναι:

$S=\{2,5,4,6,7,3\}$

με $\kappa_1=2, \kappa_2=3, \kappa_3=1, \kappa_4=2, \kappa_5=1, \kappa_6=1$

Το S γράφεται:

$S=\{2,3,4,5,6,7\}$,

οπότε:

$\omega_1=\alpha_1=2$ με $\lambda_1=\kappa_1=2$,

$\omega_2=\alpha_6=3$ με $\lambda_2=\kappa_6=1$,

$\omega_3=\alpha_3=4$ με $\lambda_3=\kappa_3=1$,

$\omega_4=\alpha_2=5$ με $\lambda_4=\alpha_2=3$,

$\omega_5=\alpha_4=6$ με $\lambda_5=\kappa_4=2$

και

$\omega_6=\alpha_5=7$ με $\lambda_6=\alpha_5=1$

Οι δέκα παρατηρήσεις γράφονται κατά τάξη αύξοντος μεγέθους: 2,2,3,4,5,5,5,6,6,7, με αντίστοιχες τάξεις τις: 1 2 3 4 5 6 7 8 9 10 οπότε οι διατεταγμένες τάξεις τους είναι η 1.5 του 2 (1+2)/2, η 3 του 3, η 4 του 4, η 6 του 5 (5+6+7/3), η 8.5 του 6 (8+9/2) και τέλος η 10 του 7.

2.6. Χρονοσειρά.

Κατά τους Box and Jenkins 1970, ένα σύνολο τιμών μιας παραμέτρου X που παρατηρήθηκαν κατά τις χρονικές στιγμές $\tau_1<\tau_2<\ldots<\tau_\nu$ αποτελεί μια χρονοσειρά της παραμέτρου. Αν τ_1 είναι η χρονική στιγμή που ένας σταθμός πραγματοποίησε την πρώτη παρατήρηση για την μετεωρολογική παράμετρο X τότε αν με τ_2 παραστήσουμε την δεύτερη, τ_3 την τρίτη κ.ο.κ τ_ν την $\nu^{\iota o\sigma\tau\eta}$ παρατήρηση, θα ισχύει:

$\tau_1<\tau_2<\ldots<\tau_\nu$

και επομένως το σύνολο των τιμών $X(\tau_1), X(\tau_2), \ldots X(\tau_\nu)$ αποτελεί μια χρονοσειρά της X. Αν με ΔT_i παραστήσουμε το χρονικό διάστημα που μεσολάβησε μεταξύ της i και της i+1 παρατήρησης, $i \in I_{\nu-1}$ θα είναι:

$$\Delta T_i=\tau_{i+1}-\tau_i \tag{2.1}$$

Οταν συμβεί:

$$\Delta T_1=\Delta T_2=\ldots\Delta T_{\nu-1}=\Delta T \tag{2.2}$$

τότε το σταθερό αυτό χρονικό διάστημα μεταξύ δύο διαδοχικών παρατηρήσεων ονομάζεται <u>περίοδος δείγματος</u>. Στην περίπτωση αυτή οι αριθμοί $\tau_1, \tau_2, \ldots\tau_\nu$ αποτελούν αριθμητική πρόοδο με λόγο ΔΤ και επομένως θα ισχύει:

$$\tau_\kappa=\tau_1+(\kappa-1)*\Delta T \quad \kappa \in I_\nu \tag{2.3}$$

και οι τιμές $X(\tau_i)$ της παραμέτρου X μπορούν να γραφούν:

$$x_1, x_2, \ldots x_\nu \text{ με } x_\kappa=X(\tau_\kappa) \tag{2.4}$$

Αν οι τιμές της X λαμβάνονται σε διακριτούς χρόνους (όργανα άμεσης ανάγνωσης ή εκτίμησης) τότε η χρονοσειρά λέγεται διακριτή. Αν η σειρά προέρχεται από συνεχής μετρήσεις στο χρόνο, τότε ονομάζεται συνεχής.

<u>Παράδειγμα 2.15.</u>

Οι τιμές της ημερήσιας ελάχιστης ή μέγιστης θερμοκρασίας ή του αρθοιστικού ημερήσιου ύψους βροχής, μπορύν να θεωρηθούν σαν χρονοσειρές με περίοδο δείγματος ΔT=24 ώρες.

Παράδειγμα 2.16.

Οι τιμές μιας παραμέτρου X που μετράται κατά τις κύριες συνοπτικές ώρες {00,06,12,18} (UTC) αποτελούν μια χρονοσειρά με περίοδο δείγματος ΔT=6h, και οι οκτώ ημερήσιες τιμές μιας παραμέτρου, που μετράται σ'όλες τις συνοπτικές ώρες, αποτελούν μια χρονοσειρά, με περίοδο δείγματος ΔT=3h.

Παράδειγμα 2.17.

Από τις τιμές των παραμέτρων που μετρούνται με αυτογραφικά όργανα, μπορεί να οριστεί ένα μεγάλος αριθμός από χρονοσειρές, που προκύπτουν από την καταγραφείσα καμπύλη με λήψη σημείων διαφορετικών περιόδων δείγματος ΔT. Το πλήθος των τιμών από τις οποίες αποτελείται κάθε τέτοια χρονοσειρά είναι αντιστρόφως ανάλογο του ΔT.

Παράδειγμα 2.18.

Αν θεωρήσουμε τους μέσους όρους των τιμών μιας παραμέτρου για κάθε ένα δεκαήμερο του έτους και ο σταθμός διαθέτει αρχείο η ετών με μετρήσεις της παραμέτρου, τότε οι 36*η αυτές τιμές, μπορούν να θεωρηθούν σαν τα στοιχεία μιας χρονοσε ιράς. Από τα παραπάνω παραδείγματα προκύπτει ότι για μια παράμετρο μπορούν να οριστούν πολλές χρονοσειρές. Τρόποι κατασκευής χρονοσειρών δίνονται και στο τεύχος 5/9 του 'ΥΔΡΟΣΚΟΠΙΟ'. Ενδιαφέρουσες χρονοσειρές για την μετεωρολογία είναι οι N χρονοσειρές των ημερήσιων τιμών των παραμέτρων που μετρούνται N φορές την ημέρα. Αν N=8 δηλαδή αν η παράμετρος μετράται οκτώ φορές την ημέρα (κατά τις συνοπτικές ώρες) οι οκτώ χρονοσειρές των ημερήσιων, (μία για κάθε ώρα παρατήρησης), τιμών, παρουσιάζουν ιδιαίτερο ενδιαφέρον. Από τον τρόπο γραφής (εξίσωση (2.4)) των τιμών μιας χρονοσειράς, προκύτπει ότι η χρονοσειρά είναι μια πεπερασμένη ή άπειρη ακολουθία τιμών της παραμέτρου X. Η υπακολουθία ή μερική ακολουθία που προκύπτει από την χρονοσειρά όταν οι δεικτες κ, με τους οποίους συνοδεύονται οι τιμές της χρονοσειράς, πάρουν ειδική μορφή (π.χ. άρτιοι δείκτες, δηλαδή κ=2m) ονομάζεται **μερική χρονοσειρά** της υπόψη χρονοσειράς και αποτελεί επίσης μια χρονοσειρά. Οταν από τις τιμές της χρονοσειράς x_{κ} λάβει κανείς μόνο τις τιμές που οι δείκτες της επαληθεύουν μια συνθήκη της μορφής $\kappa_1<\lambda<\kappa_2$ τότε η δημιουργούμενη πεπερασμένη χρονοσειρά ονομάζεται **τμήμα της χρονοσειράς** με άκρα κ_1 και κ_2.

Ενα σύνολο K τιμών, με K≤M από τις M τιμές μιας χρονοσειράς ονομάζεται **δείγμα μεγέθους K** και αν οι K τιμές εκλεγούν με τυχαίο τρόπο τότε το σύνολο αυτό ονομάζεται **τυχαίο δείγμα μεγέθους K.**

Εστω ότι οι τιμές μιας παραμέτρου X άρχισαν να καταγράφονται από την πρώτη Ιανουαρίου κάποιου έτους, που παριστάνουμε με YEAR-

START, και ότι οι τιμές λαμβανόμενες με περίοδο δείγματος ΔΤ, είναι αρχειοθετημένες για NUMYEAR χρόνια μέχρι την τριακοστή πρώτη Δεκεμβρίου του έτους YEAREND. Θα είναι λοιπόν:

$$NUMYEAR=YEAREND-YEARSTART+1 \tag{2.5}$$

Οι ημέρες που μεσολάβησαν από την πρώτη μέχρι την τελευταία καταγραφή είναι:

$$NumDays=365*K+K_1 \tag{2.6}$$

όπου K_1 το πλήθος των δίσεκτων ετών της περιόδου των NUMYEAR ετών. Αν θεωρήσουμε ότι οι τιμές της X λαμβάνονται όλες τις συνοπτικές ώρες παρατήρησης (ΔΤ=3h), τις 0,3,6,9,12,15,18,21UTC, τότε θα υπάρχουν οκτώ τιμές της X ανά ημέρα, και οι αρχειοθετημένες τιμές της X θα είναι σε πλήθος:

$$M=8*NumDays \tag{2.7}$$

Την κάθε ώρα παρατήρησης IW (όπου IW∈HOSO=Hour Of Synoptique Observation) αντιστοιχούμε στον αριθμό:

$$i=1+IW/3 \tag{2.8}$$

Από την χρονοσειρά x(1),x(2),...x(M) όλων των M τιμών της χρονοσειράς της παραμέτρου X, μπορούν να προκύψουν οκτώ μερικές χρονοσειρές, μια για κάθε ώρα παρατήρησης, με NumDays στοιχεία η κάθε μία, τα:

$$x(1,i),\ x(2,i),\ \ldots,\ x(NumDays,i) \text{ με } (i\in I_8) \tag{2.9}$$

που συνδέονται με τα στοιχεία x(j), $j\in I_M$, με την σχέση:

$$x(m,i)=x(j)=x[m+8*(i-1)] \tag{2.10}$$

και μπορούν να θεωρηθούν σαν τιμές της παραμέτρου X(i), $i\in I_8$, που πάρθηκαν με περίοδο δείγματος ΔΤ=24h (ημερήσιες τιμές της παραμέτορυ X, ορισμένης συνοπτικής ώρας IW=(i-1)*3). Οι οκτώ παραπάνω μερικές χρονοσειρές προέκυψαν από τον διαμελισμό Q του συνόλου I_M, σε οκτώ υποσύνολά του A_i, $i\in I_8$ με στοιχεία:

$$A_i=\{j\in I_M \text{ και } i=1+j\bmod 8\} \tag{2.11}$$

όπου jmod8 το υπόλοιπο της διαίρεσης του φυσικού j με το 8. Από κάθε μία από τις παραπάνω οκτώ χρονοσειρές μπορούν να δημιουργηθούν NUMYEAR τμήματα της, με διαμελισμό του συνόλου A_i (διαδοχικό του διαμελισμού Q), στα σύνολα $A_{\kappa,i}$, $i\in I_8$, $\kappa\in I_{NUMYEAR}$ με:

$$A_{\kappa,i}=\{j\in I_{NumDays},\ i\in I_8,\ 1+365*(\kappa-1)+\kappa_1\leq j\leq 365*\kappa+\kappa_1+\kappa_2\} \tag{2.12}$$

όπου $\kappa_2\in\{0,1\}$, ανάλογα με το αν το έτους YEAR(κ) είναι δίσεκτο ή όχι αντίστοιχα, και κ_1 ο αριθμός των δίσεκτων ετών κατά τα έτη YEAR(1)=YEARSTART έως YEAR(κ) ελλατωμένος κατά 1. Το έτος YEAR(κ) συνδέεται με το κ με την σχέση:

$$YEAR(\kappa)=YEARSTART-1+\kappa \tag{2.13}$$

Το κάθε σύνολο $A_{i,\kappa}$ περιέχει λοιπόν 366 ή 365 ημερήσιες τιμές (ανάλογα με το αν το έτος YEAR(κ) είναι δίσεκτο ή όχι, αντίστοιχα. Τα στοιχεία του κάθε ενός από τα κ τμήματα της χρονοσειράς, που δημιουργούνται από τον διαδοχικό διαμελισμό του συνόλου I_M μπορούν να γραφούν με την μορφή:

$$x(\kappa,1,i),\ x(\kappa,2,i),\ \ldots,x[\kappa,(365+a_\kappa),i] \tag{2.14}$$

με $a_\kappa\in\{0,1\}$ ανάλογα με το αν το έτος YEAR(κ), που δίνεται από την (2.13) είναι ή όχι δίσεκτο. Το καθένα από τα 8*NUMYEAR τμήματα της χρονοσειράς $x(j)$, $J \in I_M$ όλων των τιμών της παραμέτρου X, περιέχει τι ς 365 ή 366 τιμές της παραμέτρου, τους έτους YEAR (κ), $\kappa \in I_{NUMYEAR}$, κατά την IW=3*(i-1) συνοπτική ώρα παρατήρησης ($i \in I_8$). Αν διαμελίσουμε το σύνολο των $365+\alpha_\kappa$ ημερών του έτους YEAR(κ), στις ημέρες του κάθε μήνα $m \in I_{12}$ ή του κάθε δεκαήμερου $n \in I_{36}$, προκύπτουν δώδεκα ή τριάντα έξι αντίστοιχα τμήματα της χρονοσειράς (δηλαδή συνολικά 8*NUMYEAR*V, $V \in \{12,36\}$) που το καθένα περιέχει j τιμές, όπου το j δίνεται από:

α. Για διαμελισμό του έτους σε δώδεκα μήνες:

J=30 αν $m \in \{4,6,9,11\}$

J=31 αν $m \in \{1,3,5,7,8,10,12\}$

J=28 αν m=2 και YEAR(κ) όχι δίσεκτο

J=29 αν m=2 και YEAR(κ) δίσεκτο

β. Για διαμελισμό του έτους σε δεκαήμερα:

J=10 αν $n \neq 3*m$ ή αν $n \in \{12,18,27,33\}$

J=11 αν $n \in \{3,9,15,21,24,30,36\}$

J=8 αν n=6 και YEAR(κ) όχι δέσεκτο

j=9 αν n=6 και YEAR(κ) δίσεκτο

Τα J στοιχεία ($J \in A=\{28,29,30,31\}$ ή $J \in B=\{8,9,10.11\}$) ανάλογα με τον διαμελισμό σε μήνες ($v=mI_{12}$), ή σε δεκαήμερα ($v=n \in I_{36}$), αντίστοιχα), μπορούν να γραφούν:

$$x(\kappa,v,1,i),\ x(\kappa,v,2,i),\ldots,\ x(\kappa,v,J,i) \qquad (2.15)$$

δημιουργώντας τα τμήματα-σύνολα X(κ,v,i), των ημερήσιων τιμών της παραμέτρου X του έτους YEAR(κ), κατά τον αντίστοιχο κωδικό αριθμό i της ώρας παρατήρησης IW, του μήνα ή του δεκαήμερου v.

Για την μετεωρολογία και γενικότερα τις εφαρμοσμένες επιστήμες, μεγάλο ενδιαφέρον παρουσιάζει η χρονοσειρά που δημιουργείται από τις διαφορές των διαδοχικών όρων μιας χρονοσειράς που μετρήθηκαν με σταθερή διαφορά χρόνου k περιόδων:

$$\Delta^k \chi_i = x_{i+k} - x_i \qquad (2.16)$$

Για k=1 συμφωνούμε να γράφουμε: $\Delta^1 \chi_i = \Delta \chi_i$.

3. ΣΥΧΝΟΤΗΤΕΣ - ΣΤΟΙΧΕΙΑ ΠΙΘΑΝΟΤΗΤΩΝ

3.1. Συχνότητες απλών γεγονότων.

Εστω S ο δειγματικός χώρος μιας παραμέτρου X και $\alpha \in S$ μία τιμή της παραμέτρου. Το σύνολο {α} το ονομάσαμε στα προηγούμενα απλό γεγονός. Υποθέτουμε ότι η χρονοσειρά της παραμέτρου, που έδωσε για δειγματικό χώρο τον S, αποτελείται από M στοιχεία (M διαδοχικές παρατηρήσεις). Αν κατά τις M παρατηρήσεις της παραμέτρου X, οι τιμές της παραμέτρου ήταν ίσες με α για N(α) φορές, τότε ο αριθμός:

$N(\alpha)=$ πλήθος περιπτώσεων κατά τις οποίες $x_i=\alpha,\ i\in I_M$ (3.1)

ονομάζεται απόλυτη συχνότητα του γεγονότος {α} στις M επαναλήψεις του πειράματος μέτρησης των τιμών της X, ή απλά απόλυτη συχνότητα του {α} (ή της τιμής α) στις M επαναλήψεις-φορές. Ο αριθμός:

$f(\alpha)=N(\alpha)/M$ (3.2)

ονομάζεται σχετική συχνότητα του γεγονότος {α} (ή της τιμής α) στις M επαναλήψεις του πειράματος μέτρησης των τιμών X ή πιο απλά σχετική συχνότητα του {α}. Τέλος ο αριθμός:

$f(\%\alpha)=100*f(\alpha)$ (3.3)

ονομάζεται επί τοις εκατό συχνότητα του γεγονότος {α} κατά τις M μετρήσεις της X, ή απλούστερα επί τα εκατό συχνότητα του {α} (ή της τιμής α). Αν ο δειγματικός χώρος είναι πεπερασμένος (δηλαδή έχει πεπερασμένου πλήθους στοιχεία) και n είναι το πλήθος των στοιχείων του, δηλαδή $S=\{\alpha_1,\alpha_2,...\alpha_n\}$ τότε:

$$\sum_{i=1}^{n} N(\alpha_i)=M, \qquad \sum_{i=1}^{n} f(\alpha_i)=1, \qquad \sum_{i=1}^{n} f(\%\alpha_i)=100 \qquad (3.4)$$

3.2. Συχνότητα γεγονότος.

Η έννοια της συχνότητας ενός απλού γεγονότος, μπορεί να επεκταθεί εύκολα και για τα σύνθετα γεγονότα. Εστω $A\in 2^{S}$ ένα γεγονός, του διακριτού δειγματικού χώρου $S=\{\alpha_1,\alpha_2,...\alpha_n\}$ μιας παραμέτρου X, που προέκυψε από ένα πείραμα που επαναλήφθηκε M φορές για την μέτρηση των τιμών της X. Αν το γεγονός A είναι το κενό σύνολο ($A=\emptyset$) τότε ορίζονται σαν απόλυτη συχνότητα, σχετική συχνότητα και επί τοις εκατό συχνότητά του, ο αριθμός 0 δηλαδή:

$N(\emptyset)=f(\emptyset)=f(\%\emptyset)=0$ (3.5)

Αν το γεγονός A αποτελείται από περισσότερα του ενός στοιχεία, έστω ν, τότε το $A=\{\beta_1,\beta_2,...\beta_\nu\}$ γράφεται σαν ένωση ν ξένων μεταξύ τους συνόλων:

$$A=\bigcup_{i=1}^{\nu}\{\beta_i\}$$

και η απόλυτη, ή σχετική, ή σχετική επί τα εκατό, συχνότητες ορίζονται αντίστοιχα από τους αριθμούς:

$$N(A)=\sum_{i=1}^{\nu} N(\beta_i), \quad f(A)=\sum_{i=1}^{\nu} N(\beta_i)/M \quad \text{και} \quad f(\%A)=100*f(A) \qquad (3.6)$$

Μετά απ'αυτά προκύπτει εύκολα ότι:

$$N(S)=M, \ f(S)=1, \ f(\%S)=100 \qquad (3.7)$$

Στους ορισμούς των συχνοτήτων (απόλυτη, σχετική, σχετική επί τοις εκατό) του γεγονότος A πρέπει να προστίθεται και η έκφραση "στις M τιμές της παραμέτρου X". Ειδικά μάλιστα για τον ορισμό της απόλυτης συχνότητας η προσθήκη της παραπάνω έκφρασης ε ίναι απαραίτητη, ενώ όταν το M είναι αρκετά μεγάλο, στους ορισμούς των σχετικών συχνοτήτων, η έκφραση, δεν είναι απαραίτητη. Η σχετική συχνότητα $f(A)$ ενός γεγονότος A μπορεί να θεωρηθεί σαν μία συνάρτηση με πεδίο ορισμού το δυναμοσύνολο 2^S του δειγματικού χώρου S μιας τυχαίας μεταβλητής και πεδίο τιμών το σύνολο-διάστη-στημα $[0,1]$. Η τιμή-εικόνα $f(A)$ του γεγονότος A είναι το πηλίκο του αριθμού $N(A)$, των περιπτώσεων κατά τις οποίες οι τιμές της παραμέτρου X ανήκαν στο σύνολο A κατά τις M μετρήσεις της X, δια του αριθμού M.

Η συνάρτηση αυτή ονομάζεται συνάρτηση πυκνότητας πιθανότητας (σ.π.π.). Η συνάρτηση πυκνότητας πιθανότητας $f(x)$ ικανοποιεί τις:

$$\text{f1:} \quad f(x)\geq 0 \quad \forall x\in R \qquad (3.8)$$

$$\text{f2:} \quad \sum^{M} f(x)=1 \ \text{για διακριτή τυχαία μεταβλητή} \qquad (3.9)$$

$$\text{f3:} \quad \int_{-\infty}^{+\infty} f(x)dx=1 \quad \text{για συνεχή τυχαία μεταβλητή} \qquad (3.10)$$

<u>Παράδειγμα 3.1.</u>

Εστω X η παράμετρος με τιμές X=0 όταν κατά την διάρκεια μιας μέρας δεν σημειώθηκε βροχή, X=1 όταν το ύψος της βροχής κατά την διάρκεια της ημέρας ήταν μικρότερο από 5mm και X=10 όταν στην διάρκεια της ημέρας το ύψος βροχής ήταν μεγαλύτερο ή ίσο από 5mm. Ο δειγματικός χώρος της X είναι ο: $S=\{0,1,10\}$ και στις M=10000 μέρες που μετρήθηκε η X σ'ένα σταθμό βρέθηκαν 8000 περιπτώσεις με X=0, 1500 περιπτώσεις με X=1 και 500 περιπτώσεις με X=10. Είναι λοιπόν:

$N(0)=8000, \ f(0)=0.8, \ f(\%0)=80,$

$N(1)=1500 \ f(1)=0.15, \ f(\%1)=15$ και

$N(10)=500, \ f(10)=0.05, \ f(\%10)=5.$

Αν ζητούσαμε τις συχνότητες του γεγονότος A={ημέρα βροχής} τότε:

$A=\{1,10\}$

και:

$N(A)=2000, \ f(A)=0.2, \ f(\%A)=20$

3.3. <u>Συχνότητα γεγονότος A υπό συνθήκη B.</u>

Πολλές φορές στην πράξη, αντί για τις συχνότητες (απόλυτη, ή σχετικές) ενός γεγονότος A, ενδιαφερόμαστε να βρούμε τις συχνότητες του A όταν συμβαίνει κάποιο γεγονός-συνθήκη B.

Παράδειγμα 3.2.

Εστω το γεγονός:

A={ Αριθμός παρατήρησης κατά την οποία η ένταση του ανέμου ήταν μεγαλύτερη από 35knots }

του δειγματικού χώρου:

S={αύξοντας αριθμός παρατήρησης της έντασης του ανέμου}

και B_1 και B_2 τα γεγονότα:

B_1={αύξοντας αριθμός παρατήρησης κατά την οποία η διεύθυνση του ανέμου κυμαινόταν μεταξύ 150° και 220° }.

B_2={αύξοντας αριθμός παρατήρησης κατά την οποία η ένταση του ανέμου στην παρατήρηση που έγινε πριν από 6 ώρες ήταν μικρότερη από 20knots}.

Μπορεί να ενδιαφερόμαστε για τα γεγονότα (που παριστάνουμε με A/B_1, A/B_2 αντίστοιχα):

A/B_1 = {ένταση FF>35knots όταν ξέρουμε ότι 150°≤DD≤220°}

A/B2 = {ένταση FF>35knots όταν πριν από 6 ώρες FF<20 knots}

Υποθέτουμε ότι:

$N(S)=M$, $N(A)\leq M$ $0<N(B_1)\leq M$, $0<N(B_2)\leq M$, $N(A\cap B_1)$ και $N(A\cap B_2)$ είναι οι απόλυτες συχνότητες των γεγονότων $S, A, B_1, B_2, A\cap B_1$ και $A\cap B_2$ αντίστοιχα. Η συχνότητα εμφάνισης του A όταν επαληθεύεται η συνθήκη B πρέπει να οριστεί με το σκεπτικό ότι το ρόλο του δειγματικού χώρου παίζει τώρα το σύνολο B_1 ή το B_2 και το ρόλου του γεγονότος, το γεγονός $A\cap B_1$ ή το $A\cap B_2$, ανάλογα με την περίπτωση.

Ετσι:

$$f(A/B_1)=\frac{N(A\cap B_1)}{N(B_1)}=\frac{N(A\cap B_1)/M}{N(B_1)/M}=\frac{f(A\cap B_1)}{f(B_1)}, \quad f(\%\ A/B_1)=100*f(A/B_1)$$

και ομοίως:

$$f(A/B_2)=\frac{N(A\cap B_2)}{N(B_2)}=\frac{N(A\cap B_2)/M}{N(B_2)/M}=\frac{f(A\cap B_2)}{f(B_2)}, \quad f(\%\ A/B_2)=100*f(A/B_2)$$

Μετά το παράδειγμα αυτό είμαστε πλέον σε θέση να δώσουμε τον ορισμό των και σχετικών συχνοτήτων ενός γεγονότος A υπό την συνθήκη B.

Ορίζεται σαν σχετική συχνότητα του γεγονότος A ως προς την συνθήκη B ο αριθμός:

$$f(A/B)=f(A\cap B)/f(B) \tag{3.11}$$

και σχετική επί τα εκατό συχνότητα του γεγονότος A ως προς την συνθήκη B ο αριθμός:

$$f(\%A/B)=100*f(A/B) \tag{3.12}$$

Στον ορισμό παραλήφθηκε σαν προφανής ο περιορισμός $N(B)\neq 0$, ενώ δεν αναφέρεται το πλήθος M των παρατηρήσεων από το οποίο ο δειγματικός χώρος προήλθε, για να μην γίνει πολύ μακρυσκελής ο ορισμός.

Παράδειγμα 3.3.

Υποθέτουμε ότι κατά τις ημέρες που η ημερήσια μέγιστη θερμοκρασία Tmax είναι μεγαλύτερη από 36°C παρατηρείται δυσμενής επίπτωση στην υγεία ορισμένων ανθρώπων. Η επίπτωση είναι μεγαλύτερη όταν η ελάχιστη θερμοκρασία Tmin την ίδια μέρα είναι μεγαλύτερη από 28°C.

Ορίζουμε τα σύνολα:

S={ημερομηνίες μέτρησης της θερμοκρασίας}

A={ημερομηνίες με Tmax>36°C},

B={ημερομηνίες με Tmin>28°C}

$A \cap B$={ημερομηνίες κατά τις οποίες Tmax>36°C και Tmin>28°C},

και έστω:

$M=N(S)=10000$, $NA=2000$, $N(B)=1800$, $NA\cap B=1000$.

Θα είναι: $f(A)= 0.2$, $f(B)= 0.18$, $f(A\cap B)=0.1$ και

$f(A/B)=f(A\cap B)/f(B)=0.1/0.18=0.555$

$f(B/A)=f(A\cap B)/f(A) = 0.1/0.2 = 0.5$

Από το παράδειγμα αυτό προκύπτει ότι η πληροφορία που μας δόθηκε (συνθήκη B) για την ελάχιστη θερμοκρασία αύξησε την (στενά συνδεδεμένη με την έννοια της σχετικής συχνότητας, όπως θα δούμε στην επόμενη παράγραφο), πιθανότητα να αισθανθούν άσχημα ορισμένα άτομα.

Παράδειγμα 3.4.

Ενα πρόβλημα που παρουσιάζεται και έχει ενδιαφέρον για ενεργειακούς σκοπούς (Χαραντώνης 1990) είναι το παρακάτω.

Θα πρέπει να εγκατασταθεί σέ ένα τόπο ανεμογεννήτρια ή συλλέκτης ηλιακής ενέργειας ή ταυτόχρονα τα δύο συστήματα;

Εστω ότι μια ανεμογεννήτρια παράγει αρκετή ενέργεια για να καλύψει τις ενεργειακές ανάγκες όταν η μέση ημερήσια ένταση του ανέμου ξεπερνά τα 10m/sec και για τον ίδιο λόγο ένα ηλιακό σύστημα χρειάζεται ηλιοφάνεια μεγαλύετερη από 3h/ημέρα.

Χρησιμοποιούμε για δειγματικό χώρο τον δειγματικό χώρο του προηγούμενου παραδείγματος και για γεγονότα A και B τα παρακάτω:

A={ημερομηνίες με μέση ένταση ανέμου FF≥10m/sec},

B={ημερομηνίες με διάρκεια ηλιοφάνειας η≥3h},

οπότε:

$A\cap B$={ημερομηνίες με FF≥10ms^{-1} και η≥3h}

Εστω $f(A)=0.5$, $f(B)=0.75$ και $f(A\cap B)=0.3$. Θα είναι λοιπόν: $f(A/B)=0.3/0.75=0.4$ και $f(B/A)=0.3/0.5=$ 0.6 Εδώ φαίνεται ότι η εγκατάσταση ηλιακού συστήματος σε τοποθεσία που είναι εγκατεστημένη μόνο ανεμογεννήτρια ανεβάζει την αξιοπιστία του σύνθετου συστήματος (ανεμογεννήτρια-ηλιακός συλλέκτης) κατά 0.167 (0.667-0.5) και η εγκατάσταση ανεμογεννήτριας σε τοποθεσία που είναι εγκατεστημένος ηλιακός συλλέκτης ανεβάζει την αξιοπιστία του συστήματος κατά 0.2 (0.8-0.6). Τα αντίστοιχα ποσοστά βελτίωσης είναι:

$0.167/0.5 \approx 33\%$ και $0.2/0.6=33\%$ επίσης.

3.4 <u>Συνάρτηση αθροιστικής συχνότητας πιθανότητας.</u>

Στην παράγραφο 2.5 υποθέσαμε ότι από τις M τιμές μιας παραμέτρου X που μετρήθηκαν, υπήρχαν μόνο $\kappa \leq M$ τιμές που ήταν διαφορετικές μεταξύ τους, που, για τον ορισμό της διατεταγμένης τάξης των τιμών αυτών, τις διατάξαμε κατά σειρά αύξοντος μεγέθους.

Εστω λοιπόν $S=\{x_1,x_2,\ldots x_\kappa\}$ με $x_1<x_2<\ldots x_\kappa$, και $N_1,N_2,\ldots N_\kappa$, $f(x_1),f(x_2),\ldots f(x_\kappa)$, οι απόλυτες και οι σχετικές συχνότητες των απλών γεγονότων $\{x_i\}, i\in I_\kappa$.

Ορίζουμε σαν συνάρτηση αθροιστικής πιθανότητας (σ.α.π.), την συνάρτηση $F:2^S \longrightarrow [0,1]$, που επαληθεύει τις παρακάτω συνθήκες:

F1. $F(x)\geq 0 \quad \forall x\in R$ (3.13)

F2. $F(-\infty)=\lim\limits_{x\to-\infty} F(x)=0 \quad F(+\infty)=\lim\limits_{x\to+\infty} F(x)=1$ (3.14)

F3. $F(x)\leq F(y) \iff x\leq y$ (3.15)

F4a. $F(x)=\int_{-\infty}^{x} f(t)dt$ (3.16)

για τυχαία μεταβλητή με συνεχή δειγματικό χώρο και

F4b. $F(x_k)=\sum_{i=1}^{k} f(x_i)$ με $x_i\leq x_k$ (3.17)

για διακριτή τυχαία μεταβλητή.

Οι πρώτες τρείς από τις παραπάνω ιδιότητες στην περίπτωση του πεπερασμένου διακριτού χώρου S, που περιγράφεται παραπάνω γίνονται:

F1.a $F(x)\geq 0 \ \forall x\in R$

F2.a $F(x)=0 \ \forall x<x_1 \quad F(x)=1 \quad \forall x\geq x_k$

F3.a $F(x_\kappa)<F(x_\lambda) \iff \kappa<\lambda$ (ή $x_\kappa<x_\lambda$ όταν ο S δεν είναι διατετεγμένος).

Από την (3.16) προκύπτει ότι:

$f(x)=dF(x)/dx$ (3.18)

Ανάλογα με την επί τα εκατό σχετική συχνότητα μπορεί να οριστεί και η επί τα εκατό αθροιστική σχετική συχνότητα.

Για την F(x) ισχύει ακόμη:

$F(x_1\leq x\leq x_2)=F(x\leq x_2)-F(x<x_1)$

τόσο για διακριτή όσο και για συνεχή τυχαία μεταβλητή.

<u>Παράδειγμα 3.4.</u>

Για τις τιμές της ελάχιστης θερμοκρασίας που δόθηκαν στο παράδειγμα 2.14 είναι:

$f(2)=0.2$, $f(3)=0.1$, $f(4)=0.1$, $f(5)=0.3$, $f(6)=0.2$ και $f(7)=0.1$ οπότε: $F(x)=0$ όταν $x<2$, $F(x)=0.2$ όταν $2\leq x<3$, $F(x)=0.3$ όταν $3\leq x<4$, $F(x)=0.4$ όταν $4\leq x<5$, $F(x)=0.7$ όταν $5\leq x<6$, $F(x)=0.9$ όταν $6\leq x<7$, και $F(x)=1$ όταν $x\geq 7$.

3.5. Ιστογράμματα-καμπύλες συχνοτήτων.

Εστω $X: S \longrightarrow R$ μια τυχαία μεταβλητή με πεδίο ορισμού τον δειγματικό χώρο μιας παραμέτρου και πεδίο τιμών ένα υποσύνολο του R. Εστω ακόμη $f(x)$ και $F(x)$ η (σ.π.π.) και η (σ.α.π.) της X αντίστοιχα. Αν στο επίπεδο χoy φέρουμε τα ευθύγραμμα τμήματα που ενώνουν τα σημεία (x_i,o) και $(x_i,f(x_i))$, το σχήμα που θα προκύψει ονομάζεται **ιστόγραμμα** (ή ραβδόγραμμα ή χάρτης πιθανότητας (Goldberg 1960).

Αν ενώσουμε μεταξύ τους τα σημεία με συντεταγμένες:

$$(x,y)=\left((x_i+x_{i+1})/2 \ , \ [f(x_i)+f(x_{i+1})]/2 \right)$$

η προκύπτουσα καμπύλη ονομάζεται **πολύγωνο συχνοτήτων** ή **καμπύλη συχνοτήτων**. Τέλος αν κατασκευάσουμε το διάγραμμα της $F(x)$ τότε το προκύπτον σχήμα ονομάζεται **διάγραμμα αθροιστικής συχνότητας**. Σε περιπτώσεις συνεχών τυχαίων μεταβλητών τόσο η $f(x)$ όσο και η $F(x)$ είναι συνεχείς συναρτήσεις και τα διαγράμματα τους είναι συνεχής καμπύλες.

3.6. Στοιχεία πιθανοτήτων.

3.6.1. Πιθανότητα ή Συνάρτηση πιθανότητας.

Μια συνάρτηση P με πεδίο ορισμού $D_P=2^S$ όπου S ο δειγματικός χώρος μιας τυχαίας μεταβλητής ονομάζεται **συνάρτηση πιθανότητας** όταν ικανοποιεί τις παρακάτω τρείς συνθήκες:

P1: $0\leq P(A)\leq 1 \quad \forall A\in 2^S$ (3.19)

P2: $P(S)=1$ (3.20)

P3: $P(A\cup B)=P(A)+P(B)$ όταν $A\cap B=\emptyset$ (3.21)

και η τιμή της P(A) ονομάζεται **πιθανότητα του γεγονότος A**. Η ιδιότητα P3 μπορεί να επεκταθεί για την ένωση **n** ξένων (ή ασυμβίβαστων) μεταξύ τους γεγονότων-συνόλων δηλαδή:

P3: $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$ όταν $\forall i\neq j \quad A_i\cap A_j=\emptyset$ (3.22)

Για δύο και τρία τυχαία γεγονότα A,B,Γ, αποδεικνύεται ότι ισχύουν:

$P(A\cup B)=P(A)+P(B)-P(A\cap B)$ (3.23)

$P(A\cup B\cup \Gamma)=P(A)=P(B)+P(\Gamma)-P(A\cap B)-P(A\cap \Gamma)-P(B\cap \Gamma)+P(A\cap B\cap \Gamma)$ (3.24)

Είναι προφανές ότι άν $B=C_SA$ τότε σύμφωνα με την (3.21) θα ισχύει:

$P(B)=P(C_SA)=1-P(A)$, διότι $A\cup B=S$ και $A\cap B=\emptyset$,

όπου ο συμβολισμός C_XY σημαίνει το συμπλήρωμα του συνόλου Y ως προς το σύνολο X.

Η συνάρτηση πιθανότητας P και η σχετική συχνότητα f ενός γεγονότος A συνδέονται μεταξύ τους από την σχέση:

$$P(x)= \lim_{M \to +\infty} f(x) \qquad (3.25)$$

Ακόμα για μεγάλες τιμές του M αν με A συμβολίσουμε το γεγονός $A=\{X\leq x\}$ δηλαδή το σύνολο των τιμών της τυχαίας μεταβλητής X, που είναι μικρότερες από την τιμή x, τότε η P(A) και η F(A) (σ.α.π.) συνδέονται από την σχέση:

$$F(x)=F(A)=P(A)=P(X\leq x) \qquad (3.26)$$

ή ακόμη για M αρκετά μεγάλο:

$$F(x)=\begin{cases} P(-\infty<X\leq x)=\Sigma f(t) & \text{όταν η X είναι διακριτή και } t\leq x \\ \int_{-\infty}^{x} f(t)dt & \text{όταν η X είναι συνεχής} \end{cases} \qquad (3.27)$$

<u>Παράδειγμα 3.5.</u>

Διαισθητικά, η πιθανότητα εμφάνισης των αριθμών 1,2,3,4,5,6 κατά την ρήψη ενός "τίμιου" κύβου είναι 1/6. Ρίπτοντας ένα κύβο 100 φορές βρέθηκαν:

f(1)=0.2,f(2)=0.2,(f(3)=0.15,F(4)=0.15,f(5)=0.20 και f(6)=0.1.

Εδώ το M=100 είναι αρκετά μικρό για να θεωρήσουμε ότι f(i)=P(i) για κάθε $i\in I_6$.

Οπως θα δούμε παρακάτω, ο μέσος όρος μιας διακριτής μεταβλητής X που μετρήθηκε M φορές, δίνεται από την:

$$E(X)= \sum_{i=1}^{M} x_i/M \qquad (3.28)$$

ή αν $y_1,y_2,...y_k$ είναι οι κ διάφορες μεταξύ της X κατά τις M επαναλήψεις του πειράματος μέτρησης των τιμών της και $f(y_i)$, $i\in I_k$ οι αντίστοιχες σχετικές συχνότητες τους τότε:

$$E(X)= \sum_{i=1}^{k} y_i f(y_i). \qquad (3.29)$$

Στο παραπάνω παράδειγμα 3.5 με f(i)=P(i)=1/6 προκύπτει: E(X)=3.5, ενώ για τις 100 μόνο ρήψεις του παραδείγματος αυτού παίρνουμε:

E(X)=1*0.2+2*0.2+3*0.15+4*0.15+5*0.2+6*0.1=3.25.

3.6.2. Υπό συνθήκη πιθανότητα - Στοχαστικά ανεξάρτητα γεγονότα.

Αν B είναι ένα γεγονός για το οποίο $P(B)\neq 0$ και A ένα τυχαίο γεγονός τότε ο αριθμός:

$$P(A/B)= P(A\cap B)/P(B) \qquad (3.30)$$

ονομάζεται υπο συνθήκη πιθανότητα του γεγονότος A δοθέντος ότι συμβαίνει το γεγονός B, σε αντιστοιχία με την σχετική συχνότητα του A υπό την συνθήκη B.

Αν: $P(A\cap B)=P(A).P(B) \qquad (3.31)$

τότε τα γεγονότα A και B ονομάζονται στοχαστικά ανεξάρτητα ή απλά ανεξάρτητα.

Για τα ανεξάρτητα γεγονότα A και B όταν ισχύει $P(A).P(B)\neq 0$ τότε:

$$P(A/B)=P(A) \text{ και } P(B/A)=P(B) \quad (3.32)$$

3.7. Συνάρτηση πυκνότητας πιθανότητας -Συνάρτησης αθροιστικής πιθανότητας, πολυδιάστατης τυχαίας μεταβλητής. Πίνακες συνάφειας.

Εστω ότι κατά την διεξαγωγή ενός πειράματος μετρούνται ή και καταγράφονται ή και εκτιμούνται ταυτόχρονα πολλές παράμετροι που χαρακτηρίζουν την κατάσταση του μελετούμενου συστήματος. Για την μετεωρολογία αυτός είναι ο κανόνας, αφού το σύστημα "καιρός" χαρακτηρίζεται από πολλές παραμέτρους, οι οποίες μετρούνται στους σταθμούς μέτρησης και τον χαρακτηρίζουν. Ετσι σε κάθε παρατήρηση λαμβάνονται οι τιμές της θερμοκρασίας, της πίεσης, του ανέμου κλπ. Εστω λοιπόν $X_1,X_2,...X_\kappa$ $\kappa\geq 2$ τυχαίες μεταβλητές που μετρούνται ταυτόχρονα και:

$x_{1,1},x_{1,2},...x_{1,M},$

$x_{2,1},x_{2,2},...x_{2,M},$

. . . .

$x_{\kappa,1},x_{\kappa,2},...x_{\kappa,M}$

M τιμές της κάθε μιας απ'αυτές.

Υποθέτουμε ότι υπάρχουν $\lambda_1,\lambda_2,...\lambda_\kappa$ διάφορες μεταξύ τους τιμές της κάθε παραμέτρου λ_i, $i\in I_\kappa$, που υποθέτουμε ότι ταξινομήθηκαν κατά τάξη αύξοντος μεγέθους (η υπόθεση αυτή δεν είναι απαραίτητη), οι:

$$\omega_i(1)<\omega_i(2)<...<\omega_i(\lambda_i) \quad (3.33)$$

με αντίστοιχες σχετικές και αθροιστικές συχνότητες τις:

$$f(\omega_i(1)),\ f(\omega_i(2)),\ ...f(\omega_i(\lambda_i)) \quad (3.34)$$

$$F(x_{i,j}\leq w_i(1)),\ F(x_{i,j}\leq w_i(2)),...F(x_{i,j}\leq w_i(\lambda_i)) \quad j\in I_M \quad (3.35)$$

Εισάγεται η κ-διάστατη τυχαία μετεβλητή:

$$X=(X_1,X_2,...X_\kappa) \quad (3.36)$$

της οποίας η τιμή κατά την χρονική στιγμή t δίνεται από την:

$$X(t)=(X_1(t),\ X_2(t),...X_\kappa(t)) \quad (3.37)$$

Το σύνολο:

$$Q_i=\{\omega_i(1),\omega_i(2),...\omega_i(\lambda_i) \quad i\in I_\kappa \quad (3.38)$$

αποτελεί ένα διαμετων τιμών της Xi και επομένως το σύνολο:

$$Q=Q_1*Q_2*...Q_\kappa \quad (3.39)$$

αποτελεί ένα συνδυαμελισμό του συνόλου:

$$D_\kappa=[\omega_1(1),\omega_1(\lambda i)]*[\omega_2(1),\ \omega_2(\lambda_2)]*...[\omega_\kappa(1),\omega_\kappa(\lambda_\kappa)] \quad (3.40)$$

θ έ τ ο υ μ ε :

$$f(\omega_1(i_1),\omega_2(i_2),...\omega_\kappa(i_\kappa))=$$
$$=f(X_1=\omega_1(i_1),X_2=\omega_2(i_2),...,X_\kappa=\omega_\kappa(i_\kappa)) \quad (3.41)$$

και: $$F(\omega_1(i_1),\omega_2(i_2),...,\omega_\kappa(i_\lambda))=$$
$$=F(X_1\leq\omega_1(i_1),\ X_2\leq\omega_2(i_2),...X_\kappa\leq\omega_\kappa(i_\kappa)) \quad (3.42)$$

Οι συναρτήσεις που ορίζονται από τις (3.41) και (3.42) ονομάζονται συνάρτηση πυκνότητας πιθανότητας και συνάρτηση αθροιστικής πιθανότητας της κ-διάστατης τυχαίας μεταβλητής X που ορίζεται από τις (3.36 και (3.37) αντίστοιχα.

Αν $M \longrightarrow +\infty$ τότε:

$$f(\omega_1(i_1),\omega_2(i_2),\ldots\omega_\kappa(i_1))=P(\omega_1(i_1),\omega_2(i_2),\ldots\omega_\kappa(i\kappa) \quad (3.43)$$

και

$$F(\omega_1(i_1),\omega_2(i_2),\ldots\omega_\kappa(i_\kappa))=P(-\infty<X_1\leq\omega_1(i_1),\ -\infty<X_2\leq\omega_2(i_2),\ \ldots,\ -\infty<X_\kappa\leq\omega_\kappa(i_\kappa)) \quad (3.44)$$

Για $\kappa=2$ η τυχαία μεταβλητή $X=(X_1,X_2)$ λέγεται δυδιάστατη. Ο δυδιάστατος πίνακας με στοιχεία:

$$f(i,j)=f(X_1=\omega_1(i_1),X_2=\omega_2(i_2)) \quad (3.45)$$

ονομάζεται πίνακας συνάφειας των X_1 και X_2 (για $\kappa>2$ ο πίνακας συνάφειας των $X_1,X_2,\ldots X_\kappa$ είναι κ-διάστατος). Σε περίπτωση συνεχών τυχαίων μεταβλητών, ή συνδυασμού συνεχών και διακριτών τυχαίων μεταβλητών οι παραπάνω σχέσεις τροποποιούνται ανάλογα όπως γίνεται και για τις μονοδιάστατες τυχαίες μεταβλητές.

Για την δυδιάστατη διακριτή τυχαία μεταβλητή H(X,Y), αν $x_1<x_2<\ldots x_\nu$ είναι οι ν διάφορες μεταξύ τους τιμές που έλαβε η X και $y_1<y_2<\ldots y_\lambda$ οι λ διάφορες μεταξύ τους τιμές της Y σ'ένα πείραμα M ταυτόχρονων μετρήσεων των τιμών της X και της Y, οι πίνακες συνάφειας, με τιμές τις απόλυτες συχνότητες N(i,j) και τις σχετικές συχνότητες $f(i,j)=N(i,j)/M$, όπου N(i,j) το πλήθος των ζευγών που περιέχονται στο σύνολο $A=\{(x_i,y_j)/X=x_i$ και $Y=y_j\}$ έχουν την μορφή των πινάκων 3.1α και 3.1β αντίστοιχα. Η δημιουργία των κ-διαστάσεων πινάκων συνάφειας μιας κ-διαστάεων τυχαίας μεταβλητής γίνεται με τον ίδιο τρόπο που έγιναν οι πίνακες συνάφειας για την δυδιάστατη τυχαία μεταβλητή. Επειδή το σύνολο $Q_Y=\{y_1,y_2,\ldots y_\lambda\}$ αποτελεί διαμελισμό του δειγματικού χώρου των M παρατηρήσεων θα είναι:

$$N(x=x_i)=\sum_{j=1}^{\lambda} N(i,j)=N(i,*) \quad (3.46)$$

$$f(x=x_i)=\sum_{j=1}^{\lambda} f(i,j)=f(i,*) \quad (3.47)$$

Ομοίως επειδή το $Q_X=\{x_1,x_2,,,x_\kappa\}$ αποτελεί διαμελισμό του δειγματικού χώρου θα είναι:

$$N(y=y_j)=\sum^{\kappa} N(i,j)=N(*,j) \quad (3.48)$$

$$f(y=y_j)=\sum_{i=1} f(i,j)=f(*,j) \quad (3.49)$$

Αν για δύο τυχαίες μεταβλητές ισχύει:

$$f(i,j)=f(i,*)*f(*,j) \quad (3.50)$$

τότε οι τυχαίες μεταβλητές θα λέγονται στοχαστικά ανεξάρτητες ή απλούστερα ανεξάρτητες.

Γενικότερα, αν $A_1,A_2,\ldots A_\kappa$ είναι τυχαία γεγονότα των τυχαίων μεταβλητών $X_1,X_2,\ldots X_\kappa$, μιας κ-διαστάσεων τυχαίας μεταβλητής

$X=(X_1,X_2,...X_\kappa)$ που ορίζεται απ'αυτές και $f(A_1),f(A_2)...f(A_\kappa)$ οι σχετικές συχνότητες των γεγονότων $A_1,A_2,...A_\kappa$ όπως ορίστηκαν από την (3.2) τότε οι τυχαίες μεταβλητές $X_1,X_2,...X_\kappa$ θα λέγονται ανεξάρτητες μεταξύ τους αν:

$$f(A_1,A_2,...A_\kappa) = f(A_1)*f(A_2)*...f(A_\kappa) \qquad (3.51)$$

Οι αριθμοί $N(i,*)$, $N(*,j)$, $f(i,*)$ και $f(*,j)$ ονομάζονται απόλυτες και σχετικές περιθωριακές συχνότητες της δυδιάστατης διακριτής τυχαίας μεταβλητής $H(X,Y)$, αντίστοιχα.

Πίνακας 3.1.α Πίνακας συνάφειας απόλυτων συχνοτήτων $N(i,j)$ της διακριτής δυδιάστατης τυχαίας μεταβλητής $H=(X,Y)$.

X \ Y	y_1	y_2	y_3	....	y_λ	$\sum_{j=1}^{\lambda} N(i,j)$
x_1	N(1,1)	N(1,2)	N(1,3)	....	N(1,λ)	N(1,*)
x_2	N(2,1)	N(2,2)	N(2,3)	....	N(2,λ)	N(2,*)
x_3	N(3,1)	N(3,2)	N(3,3)	....	N(3,λ)	N(3,*)
.	.	.	.	....	.	.
.	.	.	.	....	.	.
.	.	.	.	....	.	.
x_κ	N(κ,1)	N(κ,2)	N(κ,3)	....	N(κ,λ)	N(κ,*)
$\sum_{i=1}^{\kappa} N(i,j)$	N(*,1)	N(*,2)	N(*,3)	...	N(*,λ)	M

Πίνακας 3.1.β Πίνακας συνάφειας σχετικών συχνοτήτων $f(i,j)$ της διακριτής δυδιάστατης τυχαίας μεταβλητής $H=(X,Y)$.

X \ Y	y_1	y_2	y_3	....	y_λ	$\sum_{j=1}^{\lambda} f(i,j)$
x_1	f(1,1)	f(1,2)	f(1,3)	....	f(1,λ)	f(1,*)
x_2	f(2,1)	f(2,2)	f(2,3)	....	f(2,λ)	f(2,*)
x_3	f(3,1)	f(3,2)	f(3,3)	....	f(3,λ)	f(3,*)
.	.	.	.	....	.	.
.	.	.	.	....	.	.
.	.	.	.	....	.	.
x_κ	f(κ,1)	f(κ,2)	f(κ,3)	....	f(κ,λ)	f(κ,*)
$\sum_{i=1}^{\kappa} f(i,j)$	f(*,1)	f(*,2)	f(*,3)	...	f(*,λ)	1

4. ΠΑΡΑΜΕΤΡΟΙ ΘΕΣΗΣ

4.1. Ακρες τιμές, εύρος τιμών.

Εστω $x_1, x_2, \ldots x_M$ μια χρονοσειρά M τιμών της τυχαίας μεταβλητής X και $S=\{\omega_1, \omega_2, \ldots \omega_\kappa\}$ με $\omega_i < \omega_j$ όταν $i<j$, ο διατεταγμένος δειγματικός χώρος της χρονοσειράς αυτής. Ο αριθμός ω_1 είναι (και ονομάζεται) η ελάχιστη τιμή της X και ο αριθμός ω_κ μέγιστη τιμή, κατά την διάρκεια του πειράματος των M μετρήσεων των τιμών της, δηλαδή:

$$\omega_1=\min\{x_i/i\in I_M\} \text{ και } \omega_\kappa=\max\{x_i/i\in I_M\} \tag{4.1}$$

Εάν η τυχαία μεταβλητή X είναι συνεχής τότε αν $D=[\alpha,\beta]$ είναι το πεδίο τιμών της, η ελάχιστη τιμή της X είναι ο αριθμός α και η μέγιστη τιμή ο αριθμός β, και ο αριθμός $EYP=\beta-\alpha$ ονομάζεται εύρος των τιμών της X. Εάν οι τιμές της X μετρούνται για n χρόνια αλλά μόνο ορισμένη περίοδο κάθε χρόνο (π.χ. μόνο τον Ιανουάριο) και α_i, β_i, $i\in I_n$ είναι η ελάχιστη και η μέγιστη τιμή κάθε έτους τότε ο αριθμός $\alpha=\min\{\alpha_i,\ i\in I_n\}$ ονομάζεται απόλυτα ελάχιστη τιμή της περιόδου και ο αριθμός $\beta=\max\{\beta i, i\in I\eta\}$ απόλυτα μέγιστη τιμή της περιόδου. Οι αριθμοί $EYP_i=\beta_i-\alpha_i$ $i\in I_n$ και $EYP=\beta-\alpha$ είναι τα ετήσια εύρη και το απόλυτο εύρος, αντίστοιχα, των τιμών της παραμέτρου.

4.2. Απόσταση εκατοστών.

Αν $x_1, x_2, \ldots x_M$ είναι οι M τιμές της παραμέτρου X που μετρήθηκαν κατά την διάρκεια ενός πειράματος και $y_1, y_2, \ldots y_\kappa$ οι κ διάφορες μεταξύ τους τιμές της, με συχνότητες $f(y_i)$, $i\in I_\kappa$ τότε:

$$F(y_1\le x_m\le y_\kappa)=1 \quad \forall m\in I_M.$$

Εστω y_λ, y_μ είναι δύο τιμές του διαστήματος $[y_1, y_\kappa]$ και έστω $\alpha\in[0,0.5]$ ένας θετικός μικρότερος του 0.5.

Αν:

$$F(y_1\le x_i\le y_\lambda)=\alpha,\ F(y_\mu< x_i\le y_\kappa)=\alpha \tag{4.2}$$

ή:

$$F(x_i>y_\lambda)=1-\alpha \text{ και } F(x_i\le y_\mu)=1-\alpha \tag{4.3}$$

προκύπτει:

$$F(y_\lambda\le x_i\le y_\mu)=1-2\alpha \tag{4.4}$$

Η διαφορά $\Delta y=y_\mu-y_\lambda$ ονομάζεται απόσταση σημείων ορισμένης πιθανότητας (2α). Αν ειδικά $\alpha=0.25$ τότε η Δy ονομάζεται ενδοτεταρτημοριακό εύρος. Αν η X ακολουθεί την κανονική κατανομή τότε ισχύει ο εμπειρικός τύπος: $\Delta y=\sigma/4$ όπου σ η τυπική απόκλιση της X.

4.3. Μέση τιμή.

Εστω X μια τυχαία μεταβλητή. Αν η τυχαία μεταβλητή είναι διακριτή και στις M μετρήσεις των τιμών της, αυτές ήταν οι $x_1, x_2, \ldots x_M$, εκ των οποίων οι n: $\omega_1 < \omega_2 < \ldots \omega_n$ ήταν διαφορετικές μεταξύ τους και είχαν σχετικές συχνότητες $f(\omega_i), i \in I_n$ τότε ορίζεται σαν μέση τιμή ή και μαθηματική ελπίδα ή δειγματικός μέσος όρος των M τιμών της X ο αριθμός:

$$E(X) = \sum_{i=1}^{M} x_i/M = \sum_{j=1}^{n} \omega_j f(\omega_j) \qquad (4.5)$$

Εάν ο M είναι πολύ μεγάλος ($M \longrightarrow +\infty$) τότε $f(\omega_i) = P(\omega_i)$ και η (4.5) γίνεται:

$$\mu_x = \Sigma\, \omega_i P(\omega_i) \quad i \in I_M \qquad (4.6)$$

Ο συμβολισμός μ_x αντί του E(x) χρησιμοποιείται εδώ για να δείξει ότι ο υπολογισμός έγινε για το σύνολο των δυνατών τιμών της X και στην περίπτωση αυτή για τον μ_x χρησιμοποιείται μόνο ο όρος μέσος όρος. Αν η τυχαία μεταβλητή είναι συνεχής, f(x) η συνάρτηση πυκνότητας πιθανότητας της και α και β η ελάχιστη και η μέγιστη τιμή της κατά την διάρκεια καταγραφής των τιμών της (το α μπορεί να είναι το $-\infty$ και το β το $+\infty$ τότε:

$$E(x) = \int_{\alpha}^{\beta} x f(x) dx \qquad (4.7)$$

και αν ο χρόνος καταγραφής των τιμών της είναι πολύ μεγάλος (τείνει στο $+\infty$) τότε αντί του E(x) χρησιμοποιείται το μx δηλαδή:

$$\mu_x = \int_{\alpha}^{\beta} x f(x) dx \qquad (4.8)$$

Εστω g(x) μία συνάρτηση με πεδίο ορισμό το πεδίο τιμών της X. Για την εύρεση της μέσης τιμής E[g(x)] της g(x) οι παραπάνω τύποι γίνονται:

$$E[g(x)] = \sum_{i=1}^{M} g(x_i)/M = \sum_{j=1}^{n} g(\omega_j).f(\omega_j) \qquad (4.9)$$

$$E[g(x)] = \int_{\alpha}^{\beta} g[x(t)].f(t) dt \qquad (4.10)$$

Από τις ιδιότητες του αθροίσματος (Σ) και του ολοκληρώματος προκύπτουν εύκολα και οι σχέσεις-ιδιότητες της μέσης τιμής:

E1: $E(\alpha x + \beta) = \alpha E(x) + \beta$ (4.11)

E2: $E[\varphi(X) + g(Y)] = E[\varphi(X)] + E[g(Y)]$ (4.12)

όπου X και Y δύο τυχαίες μεταβλητές με το ίδιο πεδίο τιμών D και φ,g δύο συναρτήσεις που έχουν κοινό πεδίο ορισμού το D.

Από την E1 για α=0 προκύπτει:

$$E(\beta) = \beta \qquad (4.13)$$

Από την ίδια σχέση για β=-E(x) και α=1 προκύπτει:

$$E[x - E(x)] = 0 \qquad (4.14)$$

Αν $Q=\{A_1,A_2,...A_\kappa\}$ είναι ένας διαμελισμός του συνόλου $[\omega_i,\omega_n]$ σε κ ισομήκεις κλάσεις σχετικής συχνότητας $f(A_i)$ και a_i το μέσο της κλάσης A_i, τότε ο $E(X)$ δίνεται προσεγκιστικά από την:

$$E(X)=\sum_{i=1}^{\kappa} a_i f(A_i) \tag{4.15}$$

Οταν η μέση τιμή υπολογίζεται από την παραπάνω σχέση, πολύ συγγραφείς (Kendall and Stuart 1976) προτείνουν την διόρθωση κατά Separd. Η διόρθωση αυτή είναι πολύ μικρή.

4.4. Επικρατούσα τιμή και διάμεσος. (Mode και median).

4.4.1. Επικρατούσα τιμή.

Εστω $\omega_1<\omega_2,...\omega_n$ είναι οι n διάφορες μεταξύ τους τιμές της τυχαίας μεταβλητής X κατά τις $M\geq n$ μετρήσεις των τιμών της, και $f(\omega_i)$ είναι η σχετική συχνότητα εμφάνισης του αριθμού ω_i. Αν υπάρχει ακριβώς ένας δείκτης λ ($\lambda\in I_n$) έτσι ώστε: $f(\omega_\lambda)>f(\omega_i)$ για κάθε $i\in I_n$ με $i\neq\lambda$, τότε ο αριθμός ω_i ονομάζεται επικρατούσα τιμή της X κατά τις M μετρήσεις των τιμών της και παριστάνεται με Mode(X). Οι Panofski και Brief (1963) προτείνουν σαν επικρατούσα τιμή, την τιμή x_λ στην οποία η $f(x)$ παρουσιάζει τοπικό μέγιστο, δηλαδή:

$$[df(x)/dx]_{x=x_\lambda}=0 \text{ και } [d^2f(x)/dx^2]_{x=x_\lambda}<0. \tag{4.16}$$

Τέτοια σημεία όμως μπορεί η $f(x)$ να έχει περισσότερα από ένα.

Αν διαμελιστεί το σύνολο $S=[\omega_1,\omega_n]$ σε $\kappa\geq 2$ ισομήκη διαστήματα-γεγονότα τα $A_1,A_2,...A_\kappa$, μήκους $(\omega_n-\omega_1)/\kappa$, και αν $f(A_1)$, $f(A_2)...,f(A_\kappa)$ είναι οι σχετικές συχνότητες των γεγονότων αυτών, τότε αν υπάρχει μοναδικό λ ($\lambda\in I_\kappa$) για το οποίο:

$f(A_\lambda)>f(A_i)$ $i\in I_\kappa$ με $i\neq\lambda$

τότε το Αλ ορίζεται σαν διάστημα επικρατούσας τιμής.

<u>Παράδειγμα 4.1</u>

Αν 1,2,3,2,5,4,3,3,4,7 είναι οι δέκα τιμές της τυχαίας μεταβλητής X στις M=10 μετρήσεις των τιμών της τότε: $f(1)=0.1, f(2)=0.2, f(3)=.0.3, f(4)=0.2, f(5)=0.1$ και $f(7)=0.1$. Προφανώς λοιπόν Mode(X)=3. Αν το σύνολο [1,7] διαμελιστεί σε 4 ισομήκη διαστήματα τα $A_1=[1,2.5)$, $A_2=[2.5,4)$, $A_3=[4,5.5)$ και $A_4=[5.5,7]$ τότε: $f(A_1)=0.3$, $f(A_2)=0.3$, $f(A_3)=0.3$, και $f(A_4)=0.1$, οπότε δεν υπάρχει με τον διαμελισμό αυτό διάστημα επικρατούσας τιμής.

4.4.2. Διάμεσος ή διάμεσος τιμή (Median).

Εάν και πάλι $\omega_1<\omega_2<...\omega_n$ είναι οι n διάφορες μεταξύ τους τιμές της τυχαίας μεταβλητής X κατά τις M μετρήσεις των τιμών της, $f(\omega_i)$ είναι οι αντίστοιχες συχνότητες των ω_i, και $F(\omega_i)$ είναι οι αντίστοιχες αθροιστικές συχνότητες των ω_i, δηλαδή:

$$F(\omega_i)=\sum_{m=1}^{i} f(\omega_m)$$

τότε αν υπάρχει δείκτης $\lambda \in I_n$ έτσι ώστε:

$$F(\omega_\lambda)=0.5 \qquad (4.17)$$

τότε η τιμή ω_λ ονομάζεται διάμεσος τιμή της X (στις M μετρήσεις των τιμών της) ή/και διάμεσος και παριστάνεται με **median(X)**.

Αν δεν υπάρχει δείκτης λ για τον οποίο $F(\omega_\lambda)=0.5$ τότε θα υπάρχει δείκτης ν έτσι ώστε:

$$F(\omega_\nu)<0.5 \text{ και } F(\omega_{\nu+1})>0.5 \qquad (4.18)$$

Στην περίπτωση αυτή το διάστημα $(\omega_\nu,\omega_{\nu+1})$ ονομάζεται διάστημα της διαμέσου της X, και η διάμεσος τιμή υπολογίζεται προσεγγιστικά με γραμμική παρεμβολή. Αν η τυχαία μεταβλητή είναι συνεχής και f(x) είναι η συνάρτηση πυκνότητας πιθανότητας της τότε:

$$F(x)=\int_\alpha^x f(t)dt=\int_x^\beta f(t)dt=0.5 \qquad (4.19)$$

όπου α και β η ελάχιστη και η μέγιστη τιμή της X αντίστοιχα κατά την διάρκεια καταγραφής των τιμών της. Αν x_o είναι η τιμή της x για την οποία ισχύει η (4.19) τότε η xo ονομάζεται διάμεσος τιμή της X κατά την διάρκεια καταγραφής των τιμών της.

4.4.3. Σχέση μέσης-επικρατούσας και διαμέσου τιμής της X.

Η σχέση που συνδέει τις τιμές E(X), Mode(X) και median (X) κατά τους Panofski και Brief (1963) είναι:

$$E(X)-mode(X)=3[E(X)-median(X)] \qquad (4.20)$$

Μια παράμετρος X για την οποία **mode(X)>E(X)** ονομάζεται θετικά ασύμμετρη, ενώ αν **mode(X)<E(X)** ονομάζεται αρνητικά ασύμμετρη. Τέλος άν **mode(X)=median(X)=E(X)** η παράμετρος ονομάζεται συμμετρική.

5. ΠΑΡΑΜΕΤΡΟΙ ΔΙΑΣΠΟΡΑΣ

5.1. Διακύμανση-Τυπική απόκλιση-Ροπές

5.1.1. Διακύμανση-Τυπική Απόκλιση.

Θεωρούμε τις τυχαίες μεταβλητές X και Y με τιμές που δίνονται στον πίνακα (5.1). Προφανώς οι δύο αυτές μεταβλητές έχουν ίδιους μέσους όρους $E(X)=E(Y)=1$ και ενώ οι τιμές της X δεν μεταβάλλονται καθόλου αυτές της Y κυμαίνονται από το -1 μέχρι το 3. Ενα μέτρο του πόσο μεταβάλλονται οι τιμές μιας τυχαίας μεταβλητής είναι η διακύμανση ή μεταβλητότητα της κατά την διάρκεια των M μετρήσεων των τιμών της, που ορίζεται από την:

$$Var(X)=S_x^2=E[X-E(X)]^2=\sum_{i=1}^{M}[X^2-E(X)]^2/M=E(X^2)-[E(X)]^2 \qquad (5.1)$$

όπου $x_1,x_2,\ldots x_M$ οι M τιμές της X κατά τις M μετρήσεις των τιμών της. Επειδή η S_x^2 έχει διαφορετικές μονάδες από την X (οι μονάδες της X στο τετράγωνο) εισάγεται η έννοια της τυπικής απόκλισης (standard deviation) που ορίζεται από την:

$$S_x=SQRT\tau S_x^2)=SQRT(Var(X)) \qquad (5.2)$$

Οπως στην περίπτωση της μέσης τιμής, όπου αντί του συμβολισμού E(X) χρησιμοποιείται ο συμβολισμός μx, όταν M—>+∞, έτσι και για την διακύμανση χρησιμοποιείται ο συμβολισμός σ_x^2 (και ο σ_x για την τυπική απόκλιση) και είναι:

$$\sigma_x^2 = E(X-\mu_x)^2 \qquad (5.3)$$

Υπό την προυπόθεση ότι M—>+∞ και $E(X-\mu)^2<+\infty$ η (5.3) μπορεί να γραφεί:

$$\sigma\chi_2=E(X-\mu_x)^2=E(X^2)-\mu_x^2 \qquad (5.4)$$

Από τις ιδιότητες των συμβόλων Σ (άθροισης) και ∫ (ολοκλη-ρώματος) προκύπτουν εύκολα και οι σχέσεις-ιδιότητες της διακύ-μανσης:

V1: $Var(\alpha*X+\beta)=\alpha^2*Var(X)$ (5.5)

V2: $S_{\alpha*x+\beta}=|\alpha|*S_x$ (5.6)

Πρέπει να αναφερθεί ότι για μικρές τιμές του πλήθους M των διατιθέμενων τιμών της παραμέτρου X, η διακύμανση πρέπει να υπολογίζεται από την:

$$Var(X)=M*S_x^2/(M-1) \qquad (5.7)$$

που λέγεται δειγματική διακύμανση, όπου Sx η διακύμανση που υπολογίζεται από την (5.1). Αν $X_1,X2,\ldots X\eta$ είναι η ανεξάρτητες μεταβλητές με ίδιες κατανομές συχνοτήτων δηλαδή $E(X)=E(X_i)$, $\sigma_x=\sigma_{xi}$, $i\in I_n$ και ίδιο δειγματικό χώρο τότε, η τυχαία μεταβλητή Y που ορίζεται από την:

$$Y=\sum_{i=1}^{n} x_i/n \qquad (5.8)$$

Πίνακας 5.1. Τιμές X και Y για την κατανόηση του ρόλου της διακύμανσης και της τυπικής απόκλισης.

X	1	1	1	1	1	1	E(X)=1
Y	1	-1	2	0	3	1	E(Y)=1

έχει:

$E(Y)=E(X)$ και $\sigma_Y^2=\sigma_X^2/n$ ι (5.9)

Η ισότητα $E(Y)=E(X)$ ισχύει ανεξάρτητα από το αν οι $X_1,\ldots X_n$ είναι ανεξάρτητες ή όχι.

Παράδειγμα 5.1

Αν σε σταθμό που λειτουργεί για K χρόνια παραστήσουμε με $x_{i,j}$, $i\in I_d$, $J\in I_n$, τις τιμές της θερμοκρασίας της 06UTC (ή κάποιας άλλλης ώρας) που μετρήθηκαν κατά την διάρκεια ενός μήνα με d ημέρες $d\in A=\{28,29,30,31\}$ (π.χ. για τον Ιανουάριο d=31, για τον Φεβρουάριο d=29 ή d=28 ανάλογα με το αν ο χρόνος είναι δίσεκτος ή όχι), ή κάθε μέρα του χρόνου (d=366 ή d=365 ανάλογα με το αν ο χρόνος είναι δίσεκτος ή όχι) και θεωρήσουμε ότι ο αριθμός M=d*K είναι αρκετά μεγάλος θα έχουμε:

$$\mu(T_{06})=\sum_{i=1}^{d}\sum_{j=1}^{\kappa} x_{i,j}/(d^*\kappa) \quad \sigma^2(T_{06})=\sum_{i=1}^{d}\sum_{j=1}^{\kappa}[x_{ij}-\mu(T_{06})]^2/(d^*\kappa)$$

$$E[T_{06}(j)]=\sum_{i=1}^{d} x_{ij}/d \quad E(T_{06})=\sum_{j=1}^{\kappa} E[T_{06}(j)/\kappa$$

και τέλος:

$Var(T_{06})=\Sigma[E(T_{06}(j)-E(T_{06})]^2/\kappa$ με $j\in I_\kappa$.

Αν υποθέσουμε ότι οι d τιμές κάθε έτους αποτελούν ένα τυχαίο δείγμα και ότι οι τιμές του κάθε έτους είναι ανεξάρτητες των τι-μών των άλλων ετών, τότε έχουμε κ ανεξάρτητα δείγματα και σύμφωνα με την παραπάνω πρόταση θα είναι:

$\mu(T_{06})=E(T_{06})=\Sigma\{E[T_{06}(j)]\}/\kappa \quad j\in I_\kappa$

$Var(T_{06})=\sigma^2(T_{06})/\kappa$.

Από τις τιμές του πίνακα (5.1) προκύπτει εύκολα ότι:

$Var(X)=S_X^2=0$ και $Var(Y)=S_Y^2=16/6-1=5/3$.

5.2. Ροπές κ τάξης.

Κατά τον ορισμό του μέσου όρου της συνάρτησης g(x), μιας τυχαίας μεταβλητής X, δεν έγινε καμμιά υπόθεση για την μορφή της g(x). Στην ειδική περίπτωση που η g(x) έχει την μορφή:

$g(x)=(X-\alpha)^\kappa$ (5.10)

τότε ο μέσος όρος της, E[g(x)] δίνεται από:

E[g(X)]=E[(x-α)K]=i=1ΣM(xi-α)κ/M=i=1Ση(yi-α)κf(yi) (5.11)

για την περίπτωση που η X είναι διακριτή και την:

$$E[g(x)] = E[(x-\alpha)^{\kappa}] = \int_{y_1}^{y_2} yf(y)dy \qquad (5.12)$$

όταν η X είναι συνεχής.

Ο αριθμός $E[(x-\alpha)^{\kappa}]$ ονομάζεται ροπή κ-τάξης της X ως προς α.

Αν α=0, ο $E(X^{\kappa})$ ονομάζεται ροπή κ-τάξης της X ως προς την αρχή ή απλά ροπή κ-τάξης της X και παριστάνεται με μ'_{κ} και αν α=E(X) τότε ο αριθμός $E\{[X-E(x)]^{\kappa}\}$ ονομάζεται ροπή κ-τάξης ως προς τον μέσο όρο ή κεντρική ροπή κ τάξης και παριστάνεται με $\mu_{\kappa}E(X)$ ή με μ_{κ}. Στην ειδική περίπτωση που α=E(X) και κ=2 βρίσκουμε την διακύμανση της X. Ο λόγος:

$$Kur(X)=\mu_4 E(X)/(\mu_2 E(X)^2=\mu_4 E(X)/Var^2(X) \qquad (5.13)$$

ονομάζεται κυρτότητα της καμπύλης της συνάρτησης πυκνότητας πιθανότητας f(x) της X ή απλούστερα κυρτότητα της X. Ο αριθμός αυτός αποτελεί ένα μέτρο της συγκέντρωσης των τιμών της X γύρω από τον μέσο όρο της. Πολλές τιμές κοντά στο μέσο όρο δίνουν kur(x)>3 ενώ λίγες τιμές κοντά σ'αυτόν δίνουν kur( )<3. Ο λόγος:

$$Skew(X)=\mu_3 E(X)/Var^{1.5}(X) \qquad (5.14)$$

ονομάζεται ασυμμετρία της (σ.π.π.) της X (ή απλά της X)

5.3. <u>Ανισότητα του Chebyshef.</u>

5.3.1. Νόμος των μεγάλων αριθμών.

Αν μ_x(η E(X)) ο μέσος όρος μιας τυχαίας μεταβλητής, $\sigma_x>0$ η τυπική της απόκλιση, α>0 ένας τυχόν θετικός πραγματικός αριθμός τότε η πιθανότητα να απέχει η τιμή της X από την μ_x, περισσότερο από το α είναι μικρότερη από τον λόγο $\sigma^2{}_x/\alpha^2$, δηλαδή:

$$P(\ |X-\mu_x|>\alpha\ ) < \sigma^2{}_x/\alpha^2 \qquad (5.15)$$

Η παραπάνω πρόταση είναι γνωστή σαν θεώρημα του Chebyshef και η (5.15) σαν ανισότητα του Chebyshef.

5.3.2. Νόμος των μεγάλων αριθμών.

Αν μ_x και σ_x είναι ο μέσος όρος και η τυπική απόκλιση των τιμών μιας τυχαίας μεταβλητής X, α ένας θετικός πραγματικός αριθμός και E(X) ο μέσος όρος η τιμών της X που εκλέχθηκαν με τυχαίο τρόπο (τυχαίο δείγμα μεγέθους n) τότε αυξάνοντας το n απεριόριστα η πιθανότητα:

$$P(\mu_x-\alpha\leq E(X)\leq\mu_x+\alpha)$$

πλησιάζει στην μονάδα. Η πρόταση αυτή είναι γνωστή σαν νόμος των μεγάλων αριθμών. Επειδή ο α είναι τυχόν αριθμός μπορεί κανείς να τον θεωρήσει όσο μικρό θέλει. Ετσι αυξάνοντας το μέγεθος n του δείγματος, ο μέσος όρος του E(X), μπορεί να πλησιάσει όσο θέλουμε την μέση τιμή μ_x της τυχαίας μεταβλητής.

Ο νόμος αυτός προκύπτει από την ανισότητα του Chebyshef (5.15) σε συνδυασμό με την εξίσωση (5.9) ως εξής: Αν αντί της τυχαίας μεταβλητής X θεωρήσουμε την τυχαία μεταβλητή E(X) (μέσος όρος τυχαίου δείγματος μεγέθους n από τις τιμές της X) και παραστήσουμε με Var(X) την διακύμανση των η αυτών τιμών, τότε η 5.15 γράφεται:

$$P(E(X)-\mu_x)\leq\alpha) > 1-Var(X)/(\kappa^*\alpha^2) \qquad (5.16)$$

Λαμβάνοτας υπόψη την (5.9) η (6.16) γίνεται:

$$P(|E(X)-\mu_x|\leq\alpha) > 1-\sigma^2/(n*\alpha^2) \quad (5.17)$$

από την οποία προκύπτει ότι:

$$\lim_{n \to +\infty} P(|E(X)-\mu_x|\leq\alpha)=1 \quad (5.18)$$

Μια εφαρμογή του νόμου αυτού δίνεται στο παρακάτω παράδειγμα.

Παράδειγμα 5.2.

Εστω $x_{i,j}$ οι M=d*κ τιμές της θερμοκρασίας της 06UTC του παραδείγματος 5.1 και ότι μας είναι άγνωστη η $\mu(T_{06})$. Αν E(To6) και Var(To6) είναι η μέση τιμή και η διακύμανση των M τιμών τότε σύμφωνα με την (5.17) θα είναι:

$$P(\ |E(T_{06})-\mu(T_{06})|>\alpha\ ) > 1-Var(T_{06})/(M*\alpha^2)$$

Αν α=0.1°C, κ=10 χρόνια, d=365 και $Var(T_{06})=36.5$ τότε:

$$P(\ |E(T_{06})-\mu(T_{06})|\leq 0.1\ ) > 1-36.5/(10*365*0.01)=0.99$$

δηλαδή η άγνωστη μέση τιμή της παραμέτρου T_{06}, με πιθανότητα μεγαλύτερη από 0.99 βρίσκεται στο διάστημα:

$$D=[E(T_{06})-0.1,\ E(T_{06})+0.1].$$

Από το παράδειγμα αυτό προκύπτει ότι με πολύ μικρό σφάλμα να δέχεται κανείς σαν την αληθή μέση τιμή μιας τυχαίας μεταβλητής, την μέση τιμή που προκύπτει από το σύνολο των διατιθέμενων τιμών, όταν τα χρόνια παρατήρησης είναι πάνω από 20 ή ακόμα και πάνω από 10. Οταν ο μέσος όρος μιας παραμέτρου υπολογίζεται από το σύνολο των διατιθεμένων τιμών θα αναφέρεται απλά σαν μέσος όρος της παραμέτρου, ενώ αν υπολογί-ζεται με βάση ορισμένο πλήθος τιμών θα αναφέρεται σαν δειγματικός μέσος όρος της παραμέτρου. Το ίδιο ισχύει και για τις υπόλοιπες στατιστικές παραμέτρους (π.χ. την τυπική απόκλιση, τις άκρες τιμές, την συνάρτηση αθροιστικής πιθανότητας κλπ.).

5.4. Συνδιακύμανση.

Εστω δύο τυχαίες μεταβλητές X και Y που έχουν τον ίδιο δειγματικό χώρο S, E(X) και E(Y) οι μέσες τιμές τους και Var(X), Var(Y) οι διακυμάνσεις τους. Ορίζουμε την τυχαία μεταβλητή H από την σχέση:

$$H(x)=[X(x)-E(X)]*[Y(x)-E(Y)] \quad x\in S \quad (5.19)$$

Η μέση τιμή E(H) της H ονομάζεται συνδιακύμανση των X και Y και παριστάνεται με Cov(X,Y), δηλαδή:

$$Cov(X,Y)=E\{[(X-E(X)]*[Y-E(Y)]\}=E(X*Y)-E(X)*E(Y) \quad (5.20)$$

Για την Cov(X,Y) ισχύει προφανώς και η:

$$Cov(X,Y)= Cov(Y,X) \quad (5.21)$$

Σε περίπτωση που η συνδιακύμανση υπολογίζεται από όλες τις δυνατές τιμές των τυχαίων μεταβλητών X και Y θα την παριστάνουμε με COV(X,Y), αντί του Cov(X,Y) που υπολογίζεται από ένα δείγμα μεγέθους Μ ζεύγων (x_i, y_i). Θα γράφουμε γενικά Cov(X,Y) όταν πρόκειται για δειγματική συνδιακύμανση ή την συνδιακύμανση και θα χρησιμοποιείται ο Cov(X,Y) μόνο για την συνδιακύμανση. Οταν Y=X τότε:

$$\mathrm{Cov}(X,X)=\mathrm{Var}(X) \tag{5.22}$$

Παράδειγμα 5.3.

Εστω το πρώτο δεκαήμερο του Ιανουαρίου οι τιμές της θερμοκρασίας της 12UTC και οι τιμές της αντίστοιχης πίεσης (διορθωμένη ως προς την θερμοκρασία και το υψόμετρο του σταθμού είναι οι εμφανιζόμενες στον πίνακα 5.2. Αν λάβουμε σαν δειγματικό χώρο τον $S=\{i/i\in I_{10}\}$ με στοιχεία τον χρόνο παρατήρησης και θέσουμε $x_i=T_{12}(i), y_i=P_{12}(i)$ τότε όπως προκύπτει από τον πίνακα 5.2 θα είναι:

$$\mathrm{Cov}(X,Y)=\mathrm{Cov}(T_{12},P_{12})=-4.432$$

Πίνακας 5.2. Τιμές της θερμοκρασίας και της πίεσης της 12UTC κατά το πρώτο δεκαήμερο του Ιανουαρίου. Με X' παριστάνεται η διαφορά X-E(X) με $X\in\{T_{12},P_{12}\}$.

α/α	T_{12}	P_{12}	T'_{12}	P'_{12}	T'*P'
1	12.3	1019.2	1.6	-7.1	-11.36
2	11.7	1023.5	1.0	-2.8	-2.80
3	14.5	1028.2	2.8	1.9	5.32
4	9.8	1030.1	-0.9	3.8	-3.42
5	7.6	1034.1	-3.1	7.8	-24.18
6	9.0	1030.2	-1.7	3.9	-6.63
7	10.0	1028.0	-0.7	1.7	-1.19
8	10.1	1025.3	-0.6	-1.0	0.60
9	12.5	1022.8	1.8	-3.5	-6.30
10	9.5	1021.6	-1.2	-4.7	5.64
Σ	107.0	10263.0	0	0	-44.32
E	10.7	1026.3	0	0	-4.432

5.5. Τυποποιημένη ή κανονικοποιημένη τυχαία μεταβλητή.

Αν μια τυχαία μεταβλητή X έχει μέση τιμή μ_x (ή δειγματική μέση τιμή την E(X)) και διακύμανση $\sigma^2_x=\mathrm{Var}(X)$, τότε η τυχαία μεταβλητή:

$$Z=\frac{X-E(X)}{s_x} \quad \left[\text{ή } = \frac{X-\mu_x}{\sigma_x} \text{ όταν εργαζόμαστε με όλο τον πληθυσμό}\right] \tag{5.23}$$

με τιμές:

$z= (\chi-\mu_x)/\sigma_x = [\chi-E(X)]/\sigma_x$ (5.24)

όπως προκύπτει από τις ιδιότητες E1 και V1 (σχέσεις 4.8 και 5.5 αντίστοιχα), θέτοντας $\alpha=1/\sigma_x$ και $\beta=-\mu_x/\sigma_x$ έχει $\mu_z=0$, $\sigma_z=1$.

Η τυχαία μεταβλητή Z που ορίζεται από την 5.23 ονομάζεται τυποποιημένη ή κανονικοποιημένη τυχαία μεταβλητή της X και παριστάνεται με Z ή με X^{*} και ο μετασχηματισμός (5.23) λέγεται κανονικός μετασχηματισμός. Από την (5.20) προκύπτει:

$Cov(\alpha*X+b,\ c*Y+d)=E[(\alpha*X+b)*(c*Y+d)]-E(\alpha*X+b)*E(c*Y+d)=$
$=E(\alpha*X*c*Y+\alpha*d*X+b*c*Y+b*d)-[\alpha*E(X)+b]*[c*E(Y)+d]=$
$=\alpha*b*E(X*Y)+\alpha*d*E(X)+b*c*E(Y)+b*d-$
$-\alpha*c*E(X)*(E(Y)-\alpha*d*E(X)-b*c*E(Y)-b*d=$
$=\alpha*b*[E(X*Y)-E(X)*(E(Y)]=\alpha*b*Cov(X,Y)$ (5.25)

Αν στην 5.25 θέσουμε $\alpha=1/\sigma_x$, $b=-\mu_x/\sigma_x$, $c=\sigma_y$, $d=-\mu_y/\sigma_y$ προκύπτει η συνδιακύμανση των κανονικιποιημένων τυχαίων μεταβλητών X^{*} και Y^{*}:

$X^{*}= (X-\mu_x)/\sigma_x$ και $Y^{*}= [Y-E(Y)]/\sigma_y$ (5.26)

που είναι ίση με:

$Cov(X^{*},Y^{*})=Cov(X,Y)/(\sigma_x*\sigma_y)$ (5.27)

5.6. <u>Συντελεστής συσχέτισης.</u>

Ο αριθμός $r(X,Y)=Cov(X^{*},Y^{*})$ που ορίζεται από την (5.27) ονομάζεται συντελεστής γραμμικής συσχέτισης ή απλά συντελεστής συσχέτισης των τυχαίων μεταβλητών X και Y. Οταν αντί ενός δείγματος μεγέθους M χρησιμοποιηθεί ολόκληρος ο πληθυσμός των τιμών των X και Y τότε αντί του r(X,Y) (δειγματική τιμή) χρησιμοποιείται ο συμβολισμός ρ(X,Y) δηλαδή $\rho(X,Y)=COV(X^{*},Y^{*})$.

Αν $r(X,Y)=0$ τότε οι X και Y λέγονται ασυσχέτιστες (γραμμικά), ενώ αν $r(X,Y)\neq 0$ θα λέγονται συσχετισμένες (γραμμικά).

Αν $Var(X)*Var(Y)=0$ ο r(X,Y) δεν έχει νόημα και τότε δεχόμαστε εξ ορισμού $r(X,Y)=0$ (Goldberg 1960).

Αν λοιπόν $Var(X)*Var(Y)\neq 0$ τότε $r(X,Y)=0 \Longrightarrow Cov(X,Y)=0$

Αποδεικνύεται εύκολα ότι ο r(X,Y) επαληθεύει την διπλή ανισότητα:

$-1\leq r(X,Y)\leq 1$ (5.28)

Οταν συμβεί $r(X,Y)>0$ τότε οι X,Y λέμε ότι παρουσιάζουν θετική συσχέτιση ή είναι θετικά συσχετισμένες και αν $r(X,Y)<0$ τότε οι X, Y λέγονται αρνητικά συσχετισμένες ή ότι έχουν αρνητική συσχέτιση (στο υπό μελέτη δείγμα). Στην ειδική περίπτωση που $r(X,Y)=1$ (ή $r(X,Y)=-1$) τότε οι X,Y λέμε ότι είναι τέλεια συσχετισμένες θετικά (ή αρνητικά). Στην τελευταία αυτή περίπτωση υπάρχουν δύο πραγματικοί αριθμοί α και β με $\alpha\neq 0$ έτσι ώστε:

$Y=\alpha*X+\beta$ (5.29)

Με βάση τον ορισμό της Cov(X,Y), η μεταβλητότητα της τυχαίας μεταβλητής H=X+Y, με τιμές το άθροισμα των τιμών των τυχαίων μεταβλητών X και Y (που πρέπει να έχουν ίδιο δειγματικό χώρο) θα είναι ίση με:

$$Var(H)=Var(X+Y)=E[X+Y-E(X+Y)]^2=$$
$$=E[X-E(X)]^2+E[Y-E(Y)]^2+2*E[X-E(X)]*[Y-E(Y)]=$$
$$=Var(X)+Var(Y)+2*Cov(X,Y) \quad (5.30)$$

και αυτή της τυχαίας μεταβλητής Θ=X-Y θα είναι ίση με:

$$Var(\Theta)=Var(X-Y)=Var(X)+Var(Y)-2*Cov(X,Y) \quad (5.31)$$

όπως εύκολα μπορεί να αποδειχθεί.

Στην παράγραφο 3.7.2 ορίστηκε η έννοια των ανεξάρτητων γεγονότων από την: $P(A \cap B)=P(A)*P(B)$. Οι τυχαίες μεταβλητές X και Y με ίδιο δειγματικό χώρο και συναρτήσεις πυκνότητας πιθανότητας $f_X(x)$ και $f_Y(\chi)$ αντίστοιχα και $f_{XY}(x,y)$ από κοινού συνάρτηση πυκνότητας πιθανότητας που ορίζεται από την:

$$f_{XY}(x,y)=f(X=x \text{ και } Y=y) \quad (5.32)$$

θα λέμε ότι είναι ανεξάρτητες εάν:

$$f_{XY}(x,y)=f_X(x)*f_Y(y) \quad (5.33)$$

Για ανεξάρτητες τυχαίες μεταβλητές ισχύει:

$$E(X*Y)=E(X)*E(Y) \quad (5.34)$$

οπότε βάση της (5.20) για ανεξάρτητες τυχαίες μεταβλητές θα ισχύει:

$$(X,Y \text{ ανεξάρτητες}) \Longrightarrow Cov\ (X,Y)=0 \quad (5.35)$$

και βάση των (5.30) και (5.31):

$$(X,Y \text{ ανεξάρτητες}) \Longrightarrow Var(X \pm Y)=Var(X)+Var(Y) \quad (5.36)$$

Η ιδιότητα αυτή των ανεξάρτητων τυχαίων μεταβλητών χρησιμοποιήθηκε στην 5.9. Η σχέση (5.30) μπορεί να γενικευθεί για η τυχαίες μεταβλητές με τον ίδιο δειγματικό χώρο ως εξής:

Αν $X_1,X_2,...Xn$ είναι η τυχαίες μεταβλητές και $\alpha_1,\alpha_2,...\alpha_n$ n πραγματικοί αριθμοί τότε:

$$Var(\alpha_1 X_1+\alpha_2 X2+...\alpha_n X_n)=Var(\sum_{i=1}^{n} \alpha_i{}^2 Var(X_i)+2*\sum_{j>i}\sum_{i=1}^{n} \alpha_i\alpha_j*Cov(X_iX_j) \quad (5.37)$$

Παράδειγμα 5.3.

Εστω ότι κατά δέκα ταυτόχρονες μετρήσεις των παραμέτρων X και Y σ'ένα σταθμό βρέθηκαν οι τιμές που δίνονται στον πίνακα 5.3.

Παρατηρούμε ότι E(X)=0, δηλαδή X-E(X)=X, οπότε:

$Cov(X,Y)=E(X-E(X))*(Y-E(Y))=E[X*(Y-E(Y)]$.

Από τον πίνακα βλέπει κανείς ότι: Cov(X,Y)=0 και επομένως και C(X,Y)=0. Σύμφωνα με τα παραπάνω οι X και Y είναι ασυσχέτιστες. Στην πραγματικότητα όμως οι τιμές της Y προκύπτουν απ'αυτές της X με βάση τον τύπο: $Y=X^2-10$. Βλέπουμε λοιπόν ότι δύο συναρτήσεις μπορεί να είναι γραμμικά ασυσχέτιστες χωρίς να είναι ανεξάρτητες, αλλά να συνδέονται με σχέση άλλης μορφής από την γραμμική.

Παράδειγμα 5.4.

Το σύνολο S των ημερών που έγιναν παρατηρήσεις του κλάσματος ηλιοφάνειας και της νέφωσης σ'ένα σταθμό διαμελίσαμε σε τρεις

Πίνακας 5.3. Τιμές των παραμέτρων X και Y σε δέκα ταυτόχρονες μετρήσεις των τιμών τους.

α/α	X	Y	Y-E(Y)	X*(Y-E(Y))
1	-1	-9	-10	10
2	2	-6	-7	-14
3	5	15	14	70
4	3	-1	-2	-6
5	1	-9	-10	-10
6	-2	-6	-7	14
7	-3	-1	-2	6
8	-5	15	14	-70
9	-4	6	5	-20
10	4	6	5	20
Σ	0	10	0	0
E	0	1	0	0

κλάσεις ως προς κάθε μια από τις δύο αυτές παραμέτρους, τις:

A_1={ημέρες με κλάσμα ηλιοφάνειας μεγαλύτερο του 0.9}

A_2={ημέρες με κλάσμα ηλιοφάνειας ίσο με 0}

$A_3=S-A_1-A_2$ ως προς την ηλιοφάνεια και

B_1={ημέρες με νέφωση μικρότερη από 1/8}

B_2={ημέρες με νέφωση ίση με 8/8}

$B_3=S-B_1-B_2$

Οι παρατηρηθείσες συχνότητες του συνδιαμελισμού δίνονται στον πίνακα 5.4. Υποθέτουμε ότι $X(A_i)=i$ και $Y(B_i)=i$ Από τις τιμές του πίνακα προκύπτει:

E(X)=1*0.5+2*0.2+3*0.3=1.8

E(Y)=1*0.5+2*0.2+3*0.3=1.8

Var(X)=12*0.5+22*0.2+32*0.3-1.82=0.76=Var(X)

Cov(X,Y)=E(X-E(X)E(Y-E(Y))=

=(-0.8)*(-0.8)*0.5+0.2*0.2*0.2+1.2*1.2*0.3=

=0.32+0.08+0.432=0.76

Επομένως r(X,Y)=1.

Στο παράδειγμα αυτό βλέπει κανείς ότι $f(i,j)=0$ όταν $i \neq j$ και $f(i,i)>0$, δηλαδή οι διάφορες του μηδενός συχνότητες κείνται μόνο στην κύρια διαγώνιο του πίνακα συνάφειας και ότι στην περίπτωση αυτή r(X,Y)=1. Το συμπέρασμα αυτό μπορεί να γενικευτεί ως εξής:

Αν οι σχετικές συχνότητες στον τετραγωνικό πίνακα συνάφειας δύο τυχαίων μεταβλητών X,Y, που κείνται εκτός της κύριας διαγωνίου είναι μηδέν, τότε r(X,Y)=1.

Μάλιστα ισχύει και η παρακάτω πρόταση:

Αν οι σχετικές συχνότητες, στοιχεία του τετραγωνικού πίνακα συνάφειας δύο τυχαίων μεταβλητών X και Y, που κείνται εκτός της δευτερεύουσας διαγωνίου του πίνακα είναι μηδέν τότε r(X,Y)=-1.

Στις παραπάνω δύο περιπτώσεις μπορούν να βρεθούν δύο σταθερές α και β έτσι ώστε: $y=\alpha*X+\beta$.

Πίνακας 5.4. Σχετικές συχνότητες $f(A_i)$ $f(B_i)$ και $f(A_i \cap B_i)$ του παραδείγματος 5.4.

\X \ Y \	A_1	A_2	A_3	P(Y)
B_1	0.5	0	0	0.5
B_2	0	0.2	0	0.2
B_3	0	0	0.3	0.3
P(X)	0.5	0.2	0.3	1

5.7. Πίνακας συνδιακύμανσης-Πίνακας συσχέτισης.

Εστω οι τυχαίες μεταβλητές $X_1,X_2,\ldots X_n$ με ίδιο δειγματικό χώρο S και $Cov(X_i,X_j)$ με $(i,j)\in I_n$. οι $n*n$ συνδιακυμάνσεις τους, όπου για i=j $Cov(X_i,X_i)=Var(X_i)$. Σχηματίζουμε τον τετραγωνικό πίνακα A με στοιχεία τα:

$$A = \begin{bmatrix} Cov(X_1,X_1) & Cov(X_1,X_2) & \ldots\ldots & Cov(X_1,X_n) \\ Cov(X_2,X_1) & Cov(X_2,X_2) & \ldots\ldots & Cov(X_2,X_n) \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ Cov(X_n,X_1) & Cov(X_n,X_2) & \ldots\ldots & Cov(X_n,X_n) \end{bmatrix} \quad (5.37)$$

Ο πίνακας αυτός ονομάζεται πίνακας συνδιακύμανσης των X_1, $X_2,\ldots X_n$ και είναι συμμετρικός ως προς την κύρια διαγώνιο αφού $Cov(X,Y)=Cov(Y,X)$. Αν $\rho(X_i,X_j)$ είναι οι συντελεστές συσχέτισης των X_i και X_j ($\rho(X_i,X_i)=1$), ο τετραγωνικός πίνακας:

$$B = \begin{bmatrix} \rho(X_1,X_1) & \rho(X_1,X_2) & \ldots\ldots & \rho(X_1,X_n) \\ \rho(X_2,X_1) & \rho(X_2,X_2) & \ldots\ldots & \rho(X_2,X_n) \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \rho(X_n,X_1) & \rho(X_n,X_2) & \ldots\ldots & \rho(X_n,X_n) \end{bmatrix} \quad (5.38)$$

ονομάζεται πίνακας συσχέτισης των $X_1,X_2,\ldots X_n$ και είναι επίσης συμμετρικός διότι: $\rho(X,Y)=\rho(Y,X)$.

Στην περίπτωση που μελετάται ένα δείγμα των τυχαίων μεταβλητών αντί του $\rho(X.Y)$ θέτουμε $r(X.Y)$.

Αν $x_1,x_2,\ldots x_M$ οι M τιμές μιας χρονοσειράς της τυχαίας μεταβλητής X, ορίζουμε την τυχαία μεταβλητή X_κ, $\kappa\in I_n$ $\kappa<<<M$ (ο

συμβολισμός <<< σημαίνει πολύ μικρότερο), από την:

$$x(t)=x_{t+\kappa-1} \quad t\in I_{M-\kappa} \qquad (5.39)$$

δηλαδή:

$X_1(1)=x_1$, $X_1(2)=\chi_2$ ..... $X_1(M)=x_M$ M τιμές.
$X_2(1)=\chi_2$, $X_2(2)=x_3$ ..... $X_2(M-1)=x_M$ M-1 τιμές.
$X_n(1)=x_n$, $X_n(2)=x_n+1$ .... $X_n(M)$ M-n+1 τιμές.

Από κάθε μία από τις παραπάνω η τυχαίες μεταβλητές θεωρούμε τις πρώτες M-n τιμές $X_\kappa(1)$ .... $X_\kappa(M-\kappa+1)$, $\kappa\in I_n$.

Από τα παραπάνω προκύπτει το οι τιμές της κάθε τυχαίας μεταβλητής X_κ, προκύπτουν από τις τιμές της X, με χρονική μετατόπιση των τιμών κατά κ-1.

Αν με Cov(i,j) παραστήσουμε την συνδιακύμανση των X_i,X_j, δηλαδή:

$$Cov(i,j)=Cov(X_i,Xj) \qquad (5.40)$$

και με ρ(i,j) τον συντελεστή συσχέτισης των X_i,X_j, δηλαδή:

$$\rho(i,j)=\rho(X_i,X_j) \qquad (5.41)$$

οι αντίστοιχοι των Α και Β πίνακες ονομάζονται αντίστοιχα πίνακας αυτοσυνδιακύμανσης και αυτοσυσχέτισης. Στην περίπτωση που $M \longrightarrow +\infty$ αποδεικνύεται ότι:

$$(i\in I_n \text{ και } j\in I_n \text{ με } |i-j|=\kappa) \Longrightarrow \rho(i,j)=\rho_\kappa=ct \qquad (5.42)$$

όπου ct=σταθερά. Ο αριθμός ρ_κ ονομάζεται συντελεστής αυτοσυσχέτισης κ τάξεως, και ο πίνακας Β γράφεται:

$$B_n = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \rho_4 \cdots & \rho_{n-3} & \rho_{n-2} & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \rho_2 & \rho_3 \cdots & \rho_{n-4} & \rho_{n-3} & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \rho_1 & \rho_2 \cdots & \rho_{n-5} & \rho_{n-4} & \rho_{n-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \rho_{n-4} & \rho_{n-5} \cdots & \rho_2 & \rho_1 & 1 \end{bmatrix} \qquad (5.43)$$

Οι Box and Jenkins 1969, αποδεικνύουν ότι η ορίζουσα του πίνακα B_n είναι θετικά ορισμένη. Στην περίπτωση που το μέγεθος του δείγματος είναι μικρό αντί του ρ_κ χρησιμοποιείται ο συμβολισμός r_κ.

5.8 <u>Παλινδρόμηση.</u>

Από τον νόμο-εξίσωση:

$$P=\rho*T*R \qquad (5.44)$$

των τελείων αερίων, που συνδέει την πίεση P την απόλυτη θερμοκρασία T(°K) και την πυκνότητα ρ, ενός τέλειου αερίου (R παγκόσμια σταθερά), προκύπτει ότι αν διατηρηθεί σταθερή η τιμή μιας των P,T,ρ, έστω της ρ, τότε σε κ άθε τιμή της T θα αντιστοιχεί μια και μόνο τιμή της P.

Πιθανώς αν διατηρείται η ρ σταθερή και μετρούνται οι τιμές των P και T, τα προκύπτοντα ζεύγη να μην επαληθεύουν ακριβώς την

(5.44) εξαιτίας της ατέλειας των οργάνων ή εξαιτίας σφαλμάτων κατά την λήψη των τιμών απ'αυτά. Οταν όμως τα όργανα είναι "τέλεια" και οι μετρήσεις γίνονται χωρίς σφάλματα τότε τα ζεύγη (P,T) που θα μετρούνται κάθε φορά θα βρίσκονται επί της ευθείας: $P=\alpha$ όπου $\alpha=\rho*R$. Στην περίπτωση αυτή υπάρχει μια συνάρτηση $\varphi(P,\rho,T)=0$, εδώ η $\varphi(P,T,\rho)=P-\rho TR$ πουυ συνδέει τις διατεταγμένες τριάδες (P,T,ρ).

Η εξίσωση:

$$F=m*\gamma \qquad (5.45)$$

που συνδέει την μάζα m, την δύναμη F και την επιτάχυνση μπορεί να γραφεί:

$\varphi(F,m,\gamma)=F-m*\gamma=0$.

Ομοίως η αξία Y ενός εμπορεύματος του οποίου η τιμή μονάδας είναι α, δίνεται από την:

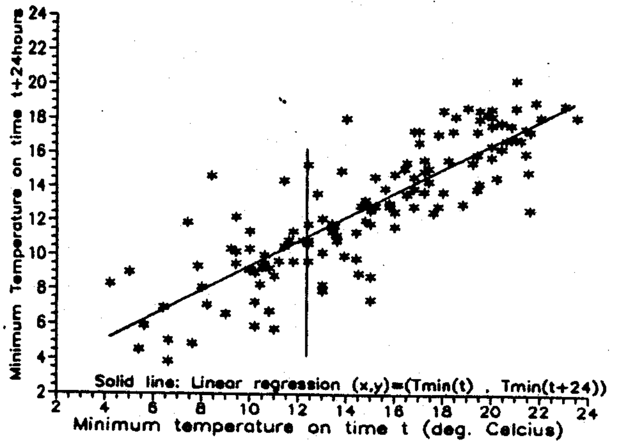
$$Y=\alpha*x \qquad (5.46)$$

όπου x οι μονάδες του εμπορεύματος και μπορεί να γραφεί: $\varphi(x,y)=y-\alpha*x$. Στις παραπάνω περιπτώσεις η τιμή μιας από τις μεταβλητές, όταν οι υπόλοιπες μεταβλητές που εμφανίζονται στην συνάρτηση φ πάρουν συγκεκριμένη τιμή, είναι συγκεκριμένη. Οταν οι $n\geq 2$ τυχαίες μεταβλητές $X_1,X2,\ldots X_n$ συνδέονται μεταξύ τους με μία συναρτησιακή σχέση της μορφής:

$$\varphi(x_1,x_2,\ldots x_n)=0 \qquad (5.47)$$

λέμε ότι είναι **συναρτησιακά εξαρτημένες**.

Εστω X και Y οι τυχαίες μεταβλητές με τιμές τις αντίστοιχες τιμές δύο παραμέτρων που μετρούνται κατά την διάρκεια ενός πειράματος. Αν για παράδειγμα η X δίνει τις μετρούμενες τιμές της θερμοκρασίας η Y θα μπορούσε να είναι η αντίστοιχη πίεση, ή η θερμοκρασία μετά χρόνο Δt κλπ. Αν σε σύστημα συντεταγμένων με άξονες X και Y τοποθετήσουμε τα ζεύγη (x,y) που μετρήθηκαν κατά την διάρκεια ενός πειράματος και τις αντίστοιχες τιμές μετά χρόνο Δt, θα εμφανιστεί ένα σχήμα της μορφής του σχήματος 5.1.



Σχήμα 5.1 Διάγραμμα Διασποράς των X και Y όπου X οι τιμές της θερμοκρασίας την D ημέρα και Y οι αντίστοιχες τιμές την D+1 ημέρα ($Y(t)=X(t+\Delta t)$ με $\Delta t=24h$).

Το διάγραμμα αυτό ονομάζεται **διάγραμμα διασποράς**, και η θέση των σημείων στο συγκεκριμένο διάγραμμα εξαρτάται από την διαφορά Δt του χρόνου κατά τον οποίο πάρθηκαν οι παρατηρήσεις της Y. Στο διάγραμμα αυτό παρατηρεί κανείς ότι στην τιμή x_i της X αντιστοιχούν περισσότερες από μία τιμές της τυχαίας μεταβλητής Y και επομένως δεν είναι δυνατόν να οριστεί μία συνάρτηση της μορφής $\varphi(x,y)=0$ που να συνδέει τις τιμές των X και Y. Για κάθε μία τιμή x από τις τιμές της X θεωρούμε τον μέσο όρο:

$$y_*=E(Y/X=x) \qquad (5.48)$$

Η y_* όντας μοναδική για κάθε συγκεκριμένη τιμή του x, μπορεί να οριστεί μια συνάρτηση της μορφής $\varphi(x,y_*)=0$ που να συνδέει τα ζεύγη (x,y_*). Η συνάρτηση $\varphi(x,y_*)=0$ ονομάζεται **συνάρτηση παλινδρόμησης** και η καμπύλη που δίνει η γραφική της παράσταση, καμπύλη ή γραμμή παλινδομήσεως. Εάν η επίλυση της $\varphi(x,y_*)=0$ ως προς y_* δώσει εξίσωση της μορφής $y_*=g(x)$ και η $g(x)$ είναι γραμμική τότε η παλινδρόμηση ονομάζεται **γραμμική παλινδρόμηση**, αλλοιώς ονομάζεται **μη γραμμική ή καμπυλόγραμμη**. Εάν για κάθε τιμή της x είναι $g(x)=ct$ (σταθερά) δεν υπάρχει καμμιά εξάρτηση των τιμών y_* από τις τιμές της x. Στην περίπτωση αυτή λέμε ότι δεν υπάρχει **στατιστική εξάρτηση** μεταξύ των τιμών των μεταβλητών X και Y ενώ αν η $g(x)$ δεν είναι σταθερά λέμε ότι οι X και Y είναι στατιστικά εξαρτημένες. Η αναζήτηση της κατάλληλης συνάρτησης της μορφής $y_*=g(x)$ που συνδέει τις τιμές της X (ανεξάρτητη μεταβλητή) με τις αντίστοιχες τιμές Y (εξαρτημένη μεταβλητή) έχει σε πολλές περιπτώσεις μεγάλη σημασία. Για παράδειγμα η μέτρηση της απόλυτης υγρασίας του αέρα θα απαιτούσε την "απόλυτη" ξήρανση ενός δείγματος του, με σημαντικό κόστος, ενώ αυτή μπορεί να βρεθεί έμμεσα από άλλες μετρούμενες παραμέτρους. Στο παράδειγμα 5.3 είδαμε ότι ενώ ο συντελεστής συσχέτισης ήταν μηδέν, οι τυχαίες μεταβλητές ήταν συναρτησιακά εξαρτημένες και στο παράδειγμα 5.4 είδαμε ότι όταν $\rho=1$ τότε υπάρχει γραμμική εξάρτηση της μορφής $Y=\alpha*x+\beta$. Επομένως ο συντελεστής $\rho(x,y)$ δεν μπορεί από μόνος του να δώσει απάντηση στο αν υπάρχει ή όχι στατιστική ή συναρτησιακή εξάρτηση μεταξύ δύο τυχαίων μεταβλητών και κυρίως (εκτός της περίπτωσης $\rho(x,y)=\pm1$) δεν μπορεί να βοηθήσει στην εύρεση της κατάλληλης συνάρτησης. Η διορατικότητα και η εμπειρία του ερευνητή στην διάγνωση της κατάλληλης συνάρτησης από το διάγραμμα διασποράς καθώς και η εμπειρία του σε σχέση με τις μετεωρολογικές παραμέτρους, παραμένουν ακόμα τα βασικότερα βοηθήματα του, για την εύρεση της κατάλληλης συνάρτησης, από τις διαθέσιμες τιμές, με όσο το δυνατό λιγότερες παραμέτρους και με την απλούστερη μορφή. Η εύρεση της κατάλληλης συνάρτησης που συνδέει δύο ή περισσότερες μεταβλητές, βασίζεται στην μέθοδο των ελάχιστων τετραγώνων, ανεξάρτητα από την συνάρτηση που κατά περίπτωση αναζητείται.

Η ανάπτυξη της μεθόδου αυτής είναι έξω από έξω από τους σκοπούς της εργασίας αυτής και ο αναγνώστης μπορεί να ανατρέξει γι'αυτήν σ'οποιοδήποτε εγχειρίδιο στατιστικής.

5.8.1 Γραμική παλινδρόμηση. Ευθεία παλινδρόμησης.

Οταν η στατιστική εξάρτηση δύο τυχαίων μεταβλητών X και Y είναι γραμμική η συνάρτηση που συνδέει τις μέσες τιμές, των τιμών της εξαρτημένης μεταβλητής (που θεωρούμε χωρίς περιορισμό της γενικότητας ότι είναι η Y), για δοθείσα τιμή x της ανεξάρτητης μεταβλητής X, (δηλαδή τις τιμές y_x που ορίζονται από την 5.48) και τις αντίστοιχες τιμές x της X, έχει την μορφή:

$$y_x=\alpha_{yx}*\ x+\beta_{yx} \qquad (5.49)$$

που ονομάζεται ευθεία παλινδρόμησης ή ευθεία των ελαχίστων τετραγώνων. Στην περίπτωση αυτή οι σταθερές α_{yx} και β_{yx} που παριστάνουν τον συντελεστή διεύθυνσης της ευθείας (5.48) και την μετατόπιση της y ως προς την αρχή των αξόνων αντίστοιχα, δίνονατι από τις σχέσεις:

$$\beta_{yx}=E(Y)-\alpha_{yx}*E(X) \qquad (5.50)$$

$$\alpha_{yx}=Cov(X,Y)/Var(X) \qquad (5.51)$$

Οι δείκτες y,x στα α και β δείχνουν ότι εξαρτημένη μεταβλητή είναι η Y και ανεξάρτητη η X. Ο συντελεστής α_{yx} ονομάζεται συντελεστής της γραμμικής παλινδρόμησης παλινδρόμησης της Y ως προς την X. Οι Cov(X,Y), Var(X), E(X) και E(Y) E(Y) βρίσκονται σύμφωνα με τους τύπους (5.20), (5.1), (4.5), αντίστοιχα, από τα M ζεύγη τιμών (x_i,y_i) που διατίθενται. Μια άλλη εύχρηστη μορφή της (5.51) είναι και η:

$$\alpha_{yx} = [\sum_{i=1}^{M} x_i*y_i - M*E(X)*E(Y)]/[\sum_{i=1}^{M} x_i^2 - M*E^2(X)] \qquad (5.52)$$

Αν ζητούσαμε την σχέση $x_y=\alpha_{xy}*y+\beta_{xy}$, δηλαδή αν θεωρούσαμε την Y για ανεξάρτητη μεταβλητή η και την X για εξαρτημένη τότε οι α_{xy} και β_{xy} θα δίνονταν από τις αντίστοιχες των (5.50) και (5.51) σχέσεις:

$$\beta_{xy}=E(X)-\alpha_{xy}*E(Y) \qquad (5.53)$$

$$\alpha_{xy}=Cov(X,Y)/Var(Y) \qquad (5.54)$$

Πολλαπλασιάζοντας τις (5.51) και (5.54) βρίσκεται:

$$r^2(X,Y)=\alpha_{y,x}*\alpha_{x,y} \qquad (5.55)$$

Ο αριθμός $r^2(X,Y)$ (ο $\rho^2(X,Y)$ σε περίπτωση άπειρου δείγματος) ονομάζεται συντελεστής προσδιορισμού (στο εξεταζόμενο δείγμα). Αν με y(x) παραστήσουμε την τιμή της Y που προκύπτει από την (5.49) για δοθείσα τιμή x της X, για τιμές των α_{yx} και β_{yx} που δίνονται από τις (5.51) και (5.50), τότε, αν $y_1,y_2,...y_k$ είναι οι κ τιμές της Y για το δεδομένο x και y_x η μέση τιμής τους, οι διαφορές:

$$\Delta y_i(x)=y_i-y(x) \qquad (5.56)$$

δεν είναι όλες μηδέν και το μοντέλο που περιγράφεται από την (5.49) είναι τόσο καλύτερο όσο, οι διαφορές (5.56) είναι μικρότερες. Ο αριθμός:

$$S_e=[M/(M-2)]^{0.5}*\{E[Y-Y(x)]^2\}^{0.5} \qquad (5.57)$$

ονομάζεται **τυπικό σφάλμα της εκτίμησης** και όσο μικρότερο είναι αυτό, τόσο λιγότερο θα διαφέρουν οι εκτιμούμενες από τις παρατηρούμενες τιμές. Αποδεικνύεται ότι:

$$S_\varepsilon^2=(M-1)*(S_Y^2-a_{yx}^2*S_X^2)/(M-2) \quad (5.58)$$

όπου M το μέγεθος του δείγματος των ζευγών (x_i,y_i) που διατίθενται και S_Y^2, S_X^2 οι δειγματικές τυπικές αποκλίσεις, που δίνονται από την (5.7). Ο S_ε^2 δίνει μια καλή εκτίμηση της διασποράς των σφαλμάτων. Αποδεικνύεται ακόμα ότι:

$$\sum_{i=1}^{M}[y_i-E(Y)]^2 = \sum_{i=1}^{M}[y_i-y(x)]^2 + \sum_{i=1}^{M}[y(x)-E(Y)]^2 \quad (5.59)$$

Το πρώτο μέρος της εξίσωσης αυτής ονομάζεται ολική μεταβολή ή ολικό σφάλμα της Y και ο πρώτος προσθετέος του δεύτερου μέρους υπόλοιπο μεταβολής. Τέλος ο δεύτερος προσθετέος στο δεύτερο μέρος της παραπάνω ισότητας είναι η μεταβλητότητα ή το σφάλμα που οφείλεται στην παλινδρόμιση. Η (5.59) λέγει ουσιαστικά ότι:

Το ολικό σφάλμα = σφάλμα οφειλόμενο στην παλινδρόμηση
(εξηγούμενο σφάλμα)
+ανεξήγητο σφάλμα (υπόλοιπο σφάλμα).

Αν το πλήθος M των ζευγών (x_i,y_i) που χρησιμοποιούνται για την εύρεση των a_{yx} και β_{yx} της (5.49) ήταν πολύ μεγάλο $(M\longrightarrow+\infty)$, παριστάνουμε το a_{yx} με α και το β_{yx} με β. Εστω ότι οι a_{yx} και β_{yx} υπολογίζονται από n, μεγέθους M δείγματα, και $a_{yx}(i)$, $\beta_{yx}(i)$ οι αντίστοιχες τιμές που βρίσκονται από το $i\in I_n$ δείγμα, τότε:

$E(a_{yx})=a=\Sigma a_{yx}(i)/n$ και (5.60)

$E(\beta_{yx})=\beta=\Sigma\beta_{yx}(i)/n$ με $i\in I_n$

$Var(a_{yx})=S_\varepsilon^2/[S_x^2*(M-1)]$ και (5.61)

$Var(\beta_{yx})=S_\varepsilon^2*\{(1/M)+E^2(X)/[M-1)*S_x^2]\}$

όπου $\varepsilon_j=y_j-[a_{yx}*x_j+\beta_{yx}]=\Delta y(x)$ το σφάλμα που γίνεται στην $j\in I_M$ παρατήρηση όταν δεχτούμε σαν τιμή της Y την $y(x_i)$ αντί της y_i. Οταν $\varepsilon\approx N(\varepsilon,0,S_\varepsilon^2)$ με S_ε να υπολογίζεται από την (5.57) τότε:

$a_{yx}\approx N(a_{yx},a,Var(a_{yx}))$ και $\beta_{yx}\approx N(\beta_{yx},\beta,Var(\beta_{yx}))$ (5.62)

Οι τυχαίες μεταβλητές:

$t_a=(a_{yx}-a)/S_a$ και $t_\beta=(\beta_{yx}-\beta)/S\beta$ (5.63)

ακολουθούν την κατανομή t με $v=M-2$ βαθμούς ελευθερίας. Επί πλέον όταν $\varepsilon\approx N(\varepsilon,0,S_\varepsilon^2)$ για δοθείσα τιμή x της X για την μέση τιμή $E(y_x)=E[Y/(X=x$ και $y=a_{yx}*x+\beta_{yx})]$ της Y ισχύει:

$E(y_x)=a_{yx}*x+\beta_{yx}$ και $Var(y_x)=S_\varepsilon^2*[(1/M)+UU]$ (5.64)

όπου

$UU=[x-E(x)]^2/[(M-1)*S_x^2]$

και για την y_x ισχύει:

$y_x \approx N(y_x,a_{yx}*x+\beta_{yx},Var(y_x))$.

Στην περίπτωση αυτή όταν δοθεί μια στάθμη σημαντικότητας (σ.σ):α, το διάστημα εμπιστοσύνης 100*(1-α) για την τιμή $E(y_x)$ είναι το:

$$D=[a_{yx}*x+\beta_{yx}-Var(y_x)*t_{M-2,a/2}\ ,\ a_{yx}*x+\beta_{yx}+Var(y_x)*t_{M-2,a/2}]= =[A_x,B_x]. \quad (5.65)$$

Αν, για τις διάφορες τιμές του X, ενώσουμε μεταξύ τους τα σημεία (x,A_x) καθώς και τα σημεία (x,B_x), προκύπτει η ζώνη εμπιστοσύνης των τιμών $E(y_x)$ για τη δοθείσα (σ.σ):α (σχήμα 5.2).

Οταν ζητάτε το (Δ.Ε.):D' της τιμής του y_x για δοθέν x αυτό (Κούνια,κ.α. 1984) δίνεται από την:

$$D' = [\omega_1,\omega_2] \quad (5.66)$$

όπου:

$$\omega_1=a_{yx}*x+\beta_{yx}-t_{M-2,a/2}*SS \text{ και } \omega_2=a_{yx}*x+\beta_{yx}+t_{M-2,a/2}*SS \quad (5.67)$$

και

$$SS^2=S_e{}^2*[1+(1/M)+UU] \quad (5.68)$$

Τα σημεία (x,ω_1) αν ενωθούν μεταξύ τους δίνουν μια γραμμή που ονομάζεται **κατώτερο όριο πρόβλεψης** της Y και τα σημεία (x,ω_2) δίνουν το **ανώτερο όριο πρόβλεψης** (σχήμα (5.2)).

Η (5.59) μπορεί να πάρει την μορφή:

$$1-r^2=\Sigma[y(\chi)-E(y)]^2/\Sigma[y_i-E(y)]^2 \quad (5.69)$$

Από την (5.69) προκύπτει ότι ο αριθμός $1-r^2$ δίνει το ποσοστό της ολικής μεταβλητότητας, που οφείλεται στο σφάλμα που γίνεται όταν σαν τιμή της μεταβλητής Y λάβουμε την τιμή y_x που δίνεται από την (5.49) αντί της παρατηρηθείσας τιμής y_i. Στις εφαρμογές ζητάτε η εύρεση της τιμής του ρ (που προκύπτει από όλον τον πληθυσμό, που θεωρείται άπειρος) και η τιμή r που βρίκεται από τα M ζεύγη που διατίθενται σε κάθε περίπτωση, αποτελεί μια εκτίμηση του ρ. Αναφέρθηκε προηγουμένως ότι όταν ρ=0 (ρ(x,y)=0) τότε οι X,Y είναι γραμμικά ασυσχέτιστες.

Αποδεικνύεται ότι όταν:

$\rho\neq0$ τότε η τυχαία μεταβλητή Ξ με τιμές:

$$\xi=0.5*\ln[(1+r)/(1-r)] \quad (5.70)$$

ακολουθεί την κανονική κατανομή (δες § 6.1), $\Xi\approx N(\Xi,\mu_\xi,\sigma_\xi{}^2)$ με:

$$\mu_\xi=0.5*\ln[(1+\rho)/(1-\rho)] \text{ και } \sigma_\xi{}^2=1/(M-3) \quad (5.71)$$

και όταν:

ρ=0 η τυχαία μεταβλητή:

$$t_r=r*[(M-2)/(1-r^2)]^{0.5} \quad (5.72)$$

ακολυθεί την t (δες § 6.8) με v=M-2 βαθμούς ελευθερίας.

Ετσι όταν βρεθεί η r(X,Y), από ένα δείγμα μεγέθους M ζευγών (x_i,y_i), τιμών των τυχαίων μεταβλητών X και Y και ζητείτε να ελεγχθεί αν οι X και Y είναι γραμμικά εξαρτημένες ή όχι (δηλαδή αν $\rho(x,y)\neq0$ ή ρ(x,y)=0) τότε το (Δ.Ε.) της ρ θα δίνεται από την:

$$D=[\ \{\exp(r_1)-1\}/AA\ ,\ \{\exp(r_2)-1\}/AA\ ] \quad (5.73)$$

με: $AA=\exp(r_1)+\exp(r_2)+1$, $BB=Z_{a/2}/(M-3)^{0.5}$ και

$$r_1=\xi-BB \qquad r_2=\xi+BB \quad (5.74)$$

όταν η ελεγχόμενη υπόθεση Ho για (σ.σ):α (Δες ΥΔΡΟΣΚΟΠΙΟ τεύχος 11/4) είναι η:

$Ho:\rho\neq0$

και διάστημα εμπιστοσύνης του το:

$$D=R-[d_1,d_2] \quad (5.75)$$

με $d_1=\{(M-2)/(1-r^2)\}^{0.5}*t_{M-3,a/2}*|r|$

και $d_2=\{(M-2)/(1-r^2)\}^{0.5}*t_{M-2,a/2}*|r|$

όταν η ελεγχόμενη υπόθεση Ho σε (σ.σ):α είναι η: Ho:ρ=0, όπου οι τιμές των Z_a και $t_{v,a}$ δίνονται στους πίνακες 6.2 και 6.3.

5.8.2 Πολλαπλή γραμμική συσχέτιση.

Εστω οι n+1 τυχαίες μεταβλητές X_1, X_2, ... X_n, Y και $(x_{1,i}, x_{2,i}, \ldots x_{n,i}, y_i)$ $i \in I_M$ ένα, n+1 διαστάσεων, μεγέθους M δείγμα των τιμών τους. Υποθέτουμε ότι ζητάμε να ελέγξουμε κατά πόσο οι τιμές y_i της μεταβλητής Y (εξαρτημένΟ μεταβλητή) συνδέεται με γραμμική σχέση της μορφής:

$$y_i = \beta o + \sum_{i=1}^{n} \beta_i * x_i + \varepsilon_i \qquad (5.76)$$

όπου, β_o, β_i $i \in I_n$ σταθερές και $e_i = y_i - y_x$, με:

$$y_x = \beta_o + \Sigma \beta_i * x_i \quad i \in I_n \qquad (5.77)$$

η τιμή του y που υπολογίζεται από την (5.76) για δοθέν $x = (x_1, x_2, \ldots x_n)$.

Υπό μορφήν πινάκων η (5.76) γράφεται:

$$Y^t = X * B^t + \epsilon^t \qquad (5.78)$$

όπου:

$$Y^t = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_M \end{bmatrix} \qquad \beta^t = \begin{bmatrix} \beta_o \\ \beta_1 \\ \beta_2 \\ \cdot \\ \beta_n \end{bmatrix} \qquad \epsilon^t = \begin{bmatrix} \varepsilon_o \\ \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \varepsilon_n \end{bmatrix}$$

Παριστάνουμε με X^t τον ανάστροφο του πίνακα X και υποθέτουμε ότι ο πίνακας X^t*X είναι μη ιδιάζων (η ορίζουσα $|x^t*x| \neq 0$), οπότε υπάρχει ο αντίστροφος του $(X^t*X)^{-1}$. Με τις παραπάνω προυποθέσεις αποδεικνύεται (με εφαρμοή της μεθόδου των ελαχίστων τετραγώνων), ότι οι τιμές των β_o και β_i $i \in I_n$ που ελαχιστοποιούν το $\Sigma\, \varepsilon_i^2 = \varepsilon^2$ $i \in I_M$, δίνονται από την:

$$\beta^t = (X^t * X)^{-1} * X^t * Y^t \qquad (5.79)$$

Αν υποθέσουμε ότι έχουμε η μεγέθους M δείγματα και ότι:

$E(\epsilon^t) = 0^t$, $Var(\epsilon^t) = \sigma^2 * I_M$ όπου $E(\beta^t) = \beta^t$ και $Var(\beta^t_i) = \omega_{ii}\sigma^2$

όπου:

$$0^t = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \qquad I^M = \begin{bmatrix} 1 & 0 & 0 \ldots 0 \\ 0 & 1 & 0 \ldots 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot \\ \cdot & \cdot & \cdot\cdot\cdot\cdot \\ 0 & 0 & 0 \ldots 1 \end{bmatrix}$$

τότε:

$E(\beta^t) = \beta^t$ και $Var(\beta_i) = w_{ii} * \sigma^2$, με ω_{ii} το στοιχείο του πίνακα $(X^t*X)^{-1}$ που καταλαμβάνει την i,i θέση (κύρια διαγώνιος).

Μία εκτίμηση για το σ^2 από ένα δείγμα μεγέθους M δίνεται από την σχέση:

$$\sigma^2 = (Y*Y^t - \beta^t * X^t * Y^t)/(M-\kappa-1) \quad \text{με } i \in I_M \qquad (5.80)$$

Η ανάλογη προς την (5.66) σχέση στην περίπτωση πολλαπλής συσχέτισης γράφεται:

$$R^2 = 1 - S_{xy}/S_y$$

με:

$$S_{xy} = Y*Y^t - \beta^t * X^t . Y^t \quad \text{και} \quad S_y{}^2 = Var(Y).$$

Ο αριθμός R ονομάζεται συντελεστής πολλαπλής γραμμικής συσχέτισης και δίνει ένα μέτρο του πόσο καλά εκφράζονται οι τιμές της Y συναρτήσει των $X_1, X_2, \ldots X_\kappa$.

Στην περίπτωση που:

$Y=X_i(t+\kappa*\Delta t)$, $X_i=X_1[t+(i-1)*\Delta t]$ $i \in I_\kappa$ (5.81)

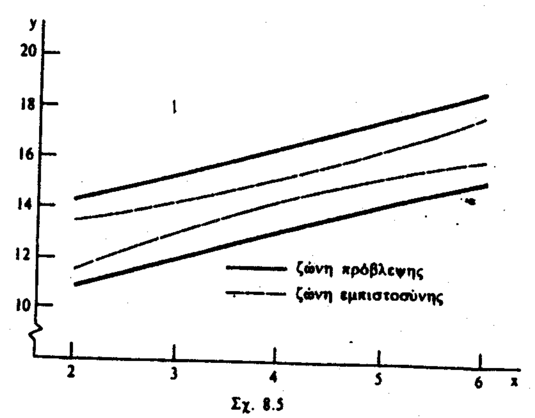
το μοντέλο που περιγράφεται από την (5.75) ονομάζεται αυτοσυσχετισμένο κ-τάξης.

5.8.3 Μη γραμμική (καμπυλόγραμμη) παλινδρόμηση.

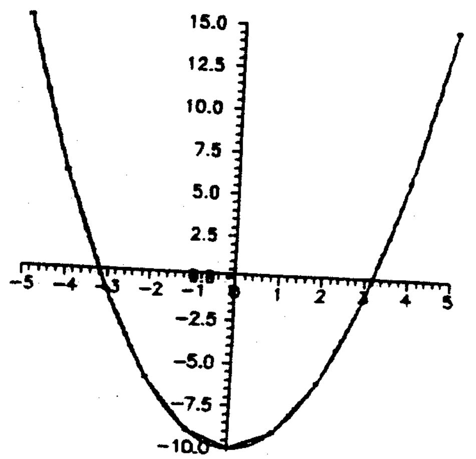
Αν το διάγραμμα διασποράς δύο τυχαίων μεταβλητών δείχνει ότι η στατιστική εξάρτηση μεταξύ των (X,Y) δεν είναι γραμμική, τότε θα έχει την μορφή $yx=g(x)$ και ανάλογα με το είδος της συνάρτησης $g(x)$ (πχ. πολυωνυμική, εκθετική, υπερβολική κλπ.) μιλάμε για στατιστική εξάρτηση μορφής $g(x)$. Η εύρεση της συνάρτησης $g(x)$ από ένα δείγμα μεγέθους M, ζευγών (xi,yi) βρίσκεται με την μέθοδο των ελαχίστων τετραγώνων όπως φαίνεται στο παρακάτω παράδειγμα.

<u>Παράδειγμα 5.5</u>

Το διάγραμμα διασποράς, ενός μεγέθους M ζευγών (x_i,y_i) των τυχαίων μεταβλητών X,Y δίνεται στο σχήμα 5.3. Ενας έμπειρος ερευνητής μπορεί να δει εύκολα ότι η συνάρτηση που συνδέει τα ζεύγη (x,y_x) είναι της μορφής: $y_x=\alpha\chi_2+\beta x+\gamma$. Θέτοντας $X_1=X$ και $X_2=X^2$ η παραπάνω σχέση γράφεται: $Yx=\beta o+\beta_1 x_1+\beta_2 x^2$ όπου $\beta o=\gamma$, $\alpha_1=\beta$ και $\alpha_2=\alpha$, παίρνει δηλαδή την μορφή της εξίσωσης (5.75) και οι συντελεστές $(\gamma,\beta,\alpha)=(\beta o,\beta_1,\beta_2)$ βρίσκονται με την μέθοδο που περιγράφηκε στην προηγούμενη παράγραφο. Οι τιμές των χ και y δίνονται στον πίνακα 5.3 και η προκύπτουσα σχέση $Y=X^2-10$ έχει συζητηθεί στο παράδειγμα 5.4.



Σχήμα 5.2. Ζώνη εμπιστοσύνης- ανώτερο και κατώτερο όριο πρόβλεψης.



Σχήμα 5.3. Καμπυλόγραμμη συσχέτιση. (Πολυώνυμο $2^{ου}$ βαθμού).

6. ΘΕΩΡΗΤΙΚΕΣ ΚΑΤΑΝΟΜΕΣ ΣΥΧΝΟΤΗΤΩΝ. ΕΛΕΓΧΟΣ ΠΡΟΣΑΡΜΟΓΗΣ. ΣΥΣΧΕΤΙΣΕΙΣ.

Στην εισαγωγή αναφέρθηκε ότι οι μετεωρολγικές παράμετροι θεωρούνται σαν διακριτές τυχαίες μεταβλητές (συναρτήσεις) σε διακριδιακριτό (μη συνεχή) χρόνο και δόθηκαν οι λόγοι για τους οποίους ευσταθεί η υπόθεση αυτή. Στο κεφάλαιο 2 δόθηκαν οι ορισμοί της συνάρτησης πυκνότητας πιθανότητας (σ.π.π.), f(x) και της συνάρτησης αθρο στικής πιθανότητας (σ.α.π.), F(x). Τις συναρτήσεις αυτές τις ονομάζουμε συναρτήσεις κατανομών (κατανομή πυκνότητας και αρθοιστική κατανομή αντίστοιχα) και σε ορισμένες περιπτώσεις, οι τιμές των συναρτήσεων αυτών είναι περίπου ταυτότσημες με τις τιμές ορισμένων θεωρητικών συναρτήσεων. Οι συναρτήσεις αυτές (που θα ονομάζουμε θεωρητικές συναρτήσεις κατανομών) μπορεί να έχουν πεδίο τιμών ένα πεπερασμένο ή αριθμήσιμα άπειρο (ισοδύναμο του N) σύνολο οπότε μιλάμε για διακριτές κατανομές ή να ορίζονται για κάθε τιμή μεταξύ δύο τιμών τους οπότε μιλάμε για συνεχείς κατανομές. Οταν οι τιμές της συνάρτησης κατανομής μιας τυχαίας μεταβλητής X ταυτίζονται (ή πλησιάζουν πολύ τις αντίστοιχες τιμές μιας θεωρητικής συνάρτησης κατανομής g, θα λέμε ότι η X ακολουθεί την κατανομή g, ή ότι η X **προσαρμόζεται** στην g. Οι πιο εύχρηστες θεωρητικές συναρτήσεις κατανομών στην μετεωρολογία είναι:

η κανονική κατανομή,
η κατανομή του Poisson,
η Διωνυμμική κατανομή,
η κατανομή του Gumbel,
η κατανομή του Weibull,
η αρνητική εκθετική
και η ομοιόμορφη κατανομή.

Η κατανομή του Student, η X^2 και η κατανομή F, μαζί με την κανονική κατανομή, αποτελούν τις σημαντικώτερες κατανομές που ακολουθούν οι στατιστικές παράμετροι (μέσες τιμές, τυπικές αποκλίσεις, συντελεστές συσχέτισης κλπ.) των τυχαίων μεταβλητών και χρησιμοποιούνται για τους στατιστικούς ελέγχους.

6.1. <u>Η κανονική κατανομή.</u>

6.1.1. Κανονική ή φυσική κατανομή ή κατανομή του Gauss.

Μια τυχαία μεταβλητή X (και κατ'επέκταση η παράμετρος που αυτή αντιπροσωπεύει) με πεδίο ορισμού το R (συνεχής) λέμε ότι ακολουθεί την κανονική κατανομή όταν έχει συνάρτηση πυκνότητας πιθανότητας f(x), που δίνεται από την:

f(χ)=exp[-(x-μ)²/(2σ²)]/[σ*SQRT(2π)] μ∈R, σ>0 (6.1)

Οι μ και σ λέγονται παράμετροι της κατανομής και αντί της έκφρασης "η παράμετρος X ακολουθεί την κανονική κατανομή" γράφουμε $X \approx N(X,\mu,\sigma^2)$. Αποδεικνύεται ότι:

$$E(X)=\mu \text{ και } \sigma^2=Var(X) \quad (6.2)$$

Κατά τους Κούνια κ.α. (1984), όταν οι τιμές μιας τυχαίας μεταβλητής επηρεάζονται από πολυάριθμους παράγοντες, μπορούμε να ισχυριστούμε ότι η παράμετρος αυτή ακολουθεί την κανονική κατανομή.

Αν η $X \approx N(X,\mu,\sigma^2)$ τότε για την κανονικοποιημένη ή/και τυποποιημένη τυχαία μεταβλητή $Z=X^{*}=(X-\mu)/\sigma$ που προκύπτει από την X με τον κανονικό μετασχηματικό (παρ.5.5) θα είναι:

$$Z \approx N(Z,0,1) \quad (6.3)$$

Η συνάρτηση πυκνότητας πιθανότητας της f(z) θα δίνεται από την:

$$f(z)=exp(-z^2/2)/SQRT(2\pi) \quad z \in R \quad (6.4)$$

και ισχύει:

$E(Z)=0$, $Var(Z)=1$, $f(-z)=f(z)$

Ακόμα αν $X \approx N(X,\mu,62)$ τότε:

$$Mode\ (X)=Median(X)=E(X)=\mu \quad (6.5)$$

$$Kur(X)=3, Asym(X)=0 \quad (6.6)$$

Η μορφή της καμπύλης f(z) δίνεται στο σχήμα 1.

Στο σχήμα 2 δίνεται το εμβαδόν (γραμμοσκιασμένο) που περιέχεται από τις:

$Z=z$ και $0 \le y \le f(z)$

Το εμβαδόν αυτό σαν συνάρτηση του z δίνεται από την:

$$\Phi(z)=P(-\infty<Z\le z)=[1/SQRT(2\pi)]*\int_{-\infty}^{z} e^{-x^2/2}dx \quad (6.7)$$

Στον πίνακα 6.1 δίνονται οι τιμές Φ(z) συναρτήσει του z για τιμές του z από -3 έως 3 με βήμα 0.01.

Ο πίνακας 6.2 περιέχει τις τιμές του εμβαδού (πιθανότητας) του γραμμοσκιασμένου τμήματος του σχήματος 3 δηλαδή τις τιμές:

$$P(0\le Z\le z)=[1/SQRT(2\pi)]*\int_{0}^{z} e^{-x^2/2}\,dx \quad (6.8)$$

συναρτήσει του z για $z \in [0,3]$, με βήμα 0.01. Το παραπάνω εμβαδό μπορεί να υπολογιστεί αρκετά απλά με ολοκλήρωση, εφαρμόζοντας τον κανόνα του τραπεζίου. Για $z>0$ ισχύουν:

$$\Phi(z)=0.5+P(0\le Z\le z) \quad (6.9)$$

$$P(-z\le Z\le 0)=P(0\le Z\le z) \quad (6.10)$$

$$P(-z\le Z\le z)=2*P(0\le Z\le z) \quad (6.11)$$

Στην μετεωρολογία πολλές μεταβλητές (θερμοκρασία στην επιφάνεια και καθ'ύψος, πίεση κλπ.) ακολουθούν κανονική κατανομή τόσο στο σύνολο των τιμών όσο και κατά μήνα ή ακόμη και κατά δεκαήμερο.

6.1.2. Ελεγχος προσαρμογής της ατανομής της τυχαίας μεταβλητής X στην κανονική κατανομή.

Η διαπίστωση για το αν μια τυχαία μεταβλητή ακολουθεί την κανονική κατανομή γίνεται με εφαρμογή του X^2 test ως εξής:

Εστω $x_1, x_2, \ldots x_M$ οι Μ τιμές μιας τυχαίας μεταβλητής Χ κατά την διάρκεια ενός πειράματος μετρήσεων των τιμών της και $\omega_1 < \omega_2 < \ldots \omega_n$ $n \leq M$ οι n διάφορες μεταξύ τους τιμές της Χ διατεταγμένες κατ'αύξουσα τάξη μεγέθους. Διαμελίζεται το διάστημα $D=[\omega_1, \omega_n]$ σε:

$$\kappa=1+3.3*\lambda o\gamma M \quad ή \quad \kappa=5*\lambda o\gamma M \tag{6.12}$$

ισομήκη διαστήματα-καταστάσεις τις:

$$A_1=[\xi_o, \xi_1],\ A_2=(\xi_1, \xi_2],\ A_\kappa=(\xi_{\kappa-1}, \xi\kappa]$$

με:

$$\xi_o=\omega_1,\ \xi_\kappa=\omega_n \text{ και } \xi_i=\xi_{i-1}+(\omega_n-\omega_i)/\kappa \quad i \in I_\kappa \tag{6.13}$$

και βρίσκουμε τις απόλυτες συχνότητες $N(A_i)$. Αν $E(X)$ και σ_x ο μέσος όρος και η τυπική απόκλιση των Μ τιμών της Χ, βρίσκουμε τις τιμές z_i των άκρων των διαστημάτων A_i που προκύπτουν από τις τιμές ξ_i με βάση το κανονικό μετασχηματισμό:

$$z_i=[\xi_i-E(X)]/\sigma_x \tag{6.14}$$

και κατόπιν βρίσκουμε τις αθροιστικές πιθανότητες $\Phi(z_i)$ από τον πίνακα 6.1 και τις απόλυτες θεωρητικές αθροιστικές συσχνότητες $g(z_i)=M*\Phi(z_i)$ με βάση τον πίνακα 6.1 και τέλος τις απόλυτες θεωρητικές συχνότητες $\Theta(A_i)$ των συνόλων A_i από την:

$$\Theta(A_i)=g(z_{i+1})-g(z_i) \quad i \in I_{\kappa \cup (o)} \tag{6.15}$$

Μετά απ'αυτά βρίσκεται ο αριθμός $X_v{}^2$ που δίνεται από την:

$$X_v{}^2=\Sigma[\Theta(A_i)-N(A_i)]^2/\Theta(A_i) \tag{6.16}$$

Αν υπάρχουν διαστήματα-καταστάσεις με συχνότητα $\Theta(Ai)<5$ τότε ή τις ενοποιούμε με γειτονικές τους ή κάνουμε νέο διαμελισμό με μικρότερο αριθμό διαστημάτων-κλάσεων. Τον $X_v{}^2$ που βρίσκεται με την (6.15) το συγκρίνουμε με θεωρητικό $X^2{}_{v,a}$ (η κατανομή $X_v{}^2$ θα αναλυθεί στην παρ.6.9) όπου v οι βαθμοί ελευθερίας (πλήθος καταστάσεων μείον αριθμός χρησιμοποιηθέντων περιορισμών και για την περίπτωση μας κ-3 διότι χρησιμοποιείται το Ε(Χ), το σ_x και το πλήθος Μ των παρατηρήσεων) και στάθμη σημαντικότητας α (κυρίως για α λαμβάνεται η τιμή α=0.05). Αν συμβεί:

$$X_v{}^2 > X^2{}_{v,a} \tag{6.17}$$

τότε απορρίπτουμε, σε στάθμη σημαντικότητας (σ.σ):α, την υπόθεση ότι οι τιμές της Χ ακολουθούν την εκλεγείσα κατανομή (εδώ την κανονική), ενώ αν:

$$X_v{}^2 \leq X_{v,a} \tag{6.18}$$

τότε η προσαρμογή θεωρείται καλή. Ο πίνακας (6.4) δίνει τις τιμές του $X^2{}_{v,a}$, για διάφορες τιμές των v και α. Η μέθοδος ελέγχου προσαρμογής της κατανομής της παραμέτρου στην φυσική κατανομή, όπως αναπτύχθηκε παραπάνω, μπορεί να εφαρμοστεί και για τον έλεγχο της προσαρμογής της κατανομής της Χ σ'οποιαδήποτε άλλη θεωρητική κατανομή, χρησιμοποιώντας την (6.16) για την εύρεση του $X_v{}^2$. Παράδειγμα εφαρμογής της παραπάνω διαδικασίας (αριθμητικό) δίνεται και στο τεύχος 5/9 του ΥΔΡΟΣΚΟΠΙΟ.

6.2. Η κατανομή του Poisson.

Εστω X μια τυχαία μεταβλητή, S ο δειγματικός της χώρος και $A \in 2^S$ ένα γεγονός. Το απλό γεγονός X(t), δηλαδή την τιμή της παραμέτρου X την χρονική στιγμή t την ονομάζουμε συμβάν. Λέμε ότι το γεγονός A συμβαίνει την χρονική στιγμή t όταν $X(t) \in A$.

Αν οι τιμές της X ικανοποιούν τις παρακάτω συνθήκες:

I. Τα πλήθη των "συμβάντων" σε ξένα μεταξύ τους χρονικά διαστήματα είναι ανεξάρτητα μεταξύ τους.

II. Η πιθανότητα να έχουμε ένα συμβάν σε χρόνο Δt ($\Delta t \longrightarrow 0$) είναι ίση με:

$$P(\text{ένα συμβάν σε χρόνο } \Delta t)=\lambda*\Delta t+0(\Delta t)*\Delta t \qquad (6.19)$$

III. Η πιθανότητα να συμβούν δύο ή περισσότερα συμβάντα σε χρόνο Δt ($\Delta t \longrightarrow 0$) είναι ίση με 0, δηλαδή:

$$P(\text{τουλάχιστον δύο συμβάντα σε χρόνο } \Delta t)=0(\Delta t)*\Delta t \qquad (6.20)$$

Η συνάρτηση 0(Δt) είναι άπειροστο μεγαλύτερης τάξης του Δt, δηλαδή

$$\lim 0(\Delta t)/\Delta t = 0 \quad (\Delta t \longrightarrow 0) \qquad (6.21)$$

Η τυχαία μεταβλητή που πληρεί τις παραπάνω προϋποθέσεις λέμε ότι ακολουθεί την κατανομή του Poisson με παράμετρο λ. Αν E(X) και Var(X) είναι ο μέσος όρος και η διακύμανση των τιμών μιας τυχαίας μεταβλητής X που ακολουθεί την κατανομή του Poisson τότε:

$$\lambda=E(X)=Var(X) \qquad (6.22)$$

και η συνάρτηση πυκνότητας πιθανότητας της δίνεται από την:

$$f(x)=P(x)=e^{-\lambda}*\lambda^{x}/x! \qquad (6.23)$$

Το x στην (6.23) παίρνει τις τιμές 0,1,2,... δηλαδή $x \in N$ και επομένως η κατανομή του Poisson είναι διακριτή. Η P(x) δίνει την πιθανότητα να συμβεί ένα γεγονός A x φορές σε άπειρο αριθμό επαναλήψεων του πειράματος. Η κατανομή του Poisson έχει πολλές ομοιότητες με την διωνυμική κατανομή που περιγράφεται στην παράγραφο 6.6. Χρησιμοποιείται για σπάνια φαινόμενα. π.χ. αν ο αριθμός των ημερών με χιονόπτωση στην Αθήνα τα τελευταία 30 χρόνια ήταν 90 τότε το φαινόμενο "χιονίζει μια μέρα στην Αθήνα" μπορεί να θεωρηθεί σπάνιο και να εξεταστεί με την κατανομή του Poisson με λ=90/30=3(ημερες/χρόνο).

Επειδή $\Sigma f(x)=\Sigma P(x)=1$ $x \in N$ όταν ζητείται η εύρεση της πιθανότητας πραγματοποίησης ενός γεγονότος το πολύ Xo φορές, αυτή θα δίνεται από:

$$P(x \leq x_o)=1-e^{-\lambda}*(\Sigma\lambda^{i}/i!) \quad i=0,1,2..,x_o \qquad (6.24)$$

Παράδειγμα 6.1.

Από το αρχείο των παρατηρήσεων ενός σταθμού εγκατεστημένου στο αεροδρόμιο κάποιας πόλης, βρέθηκε ότι κατά μέσο όρο το αεροδρόμιο παραμένει κλειστό λόγω άσχημεν καιρικών συνθηκών (ομίχλη, πάγος, ισχυροί πλάγιοι άνεμοι) πέντε φορές το χρόνο. Θεωρόντας ότι το γεγονός να κλείσει το αεροδρόμιο είναι "σπάνιο", και ότι η τυχαία μεταβλητή, που δίνει τον αριθμό των περιπτώσεων κατά τις οποίες το αεροδρόμιο είναι κλειστό, πληρεί τις συνθήκες I,II και III, η πιθανότητα να είναι κλειστό το αεροδρόμιο x φορές ανά έτος δίνεται από την (6.23) με λ=5.

6.3. Η κατανομή Weibull.

Η συνεχής τυχαία μεταβλητή X λέμε ότι ακολουθεί την κατανομή του Weibull όταν η συνάρτηση πυκνότητας πιθανότητας (σ.π.π.)της f(x) δίνεται από την:

$$f(x)=\alpha*\beta*x^{\beta-1}\exp(-\alpha*x^{\beta}) \qquad x>0 \quad \alpha>0 \quad \beta>0 \tag{6.25}$$

και συνάρτηση αθροιστικής πιθανότητας $F(x)=P(-\infty<X\leq x)$ την:

$$F(x)=P(-\infty<X\leq x)=1-\ln\alpha x^{\beta} \tag{6.26}$$

Η κατανομή αυτή χρησιμοποιείται για θετικές τυχαίες μεταβλητές (X>0) και κυρίως για την ένταση του ανέμου (Χαραντώνης και Παπασταμόπουλος 1982). Οι α και β λέγονται παράμετροι της κατανομής αυτής και όταν η τυχαία μεταβλητή X ακολουθεί την κατανομή αυτή γράφουμε:

$X\approx W(x,\alpha,\beta)$.

Η διαδικασία για τον έλεγχο της προσαρμογής της τυχαίας μεταβλητής X στην κατανομή του Weibull είναι ίδια μ'αυτή που ακολουθήθηκε στην παρ.6.1.2. Οι βαθμοί ελευθερίας και στην περίπτωση αυτή όπως και στην κανονική κατανομή είναι:

$v=\kappa-3$

αφού από τις τιμές του δείγματος υπολογίζονται οι παράμετροι α και β. Η εύρεση των παραμέτρων α και β της κατανομής του Weibull γίνεται ως εξής:

Εστω $x_1,x_2,...x_M$ οι τιμές της τυχαίας μεταβλητής X με $x_i>0$ $i\in I_M$, $\omega_1<\omega_2<...<\omega_n$ οι $n\leq M$ διάφορες μεταξύ τους τιμές της και $F(\omega_i)$ οι σχετικές αθροιστικές συχνότητες των ω_i $i\in I_n$. Αν η X ακολουθεί την κατανομή του Weibull τότε σύμφωνα με την (6.26) θα ισχύει:

$$f(\omega_i)=1-\exp(-\alpha*\omega_i^{\beta}) \tag{6.27}$$

θέτοντας $\alpha=1/c^{\beta}$ η (6.27) γίνεται:

$$F(w_i)=1-\exp[-(w_i/c)^{\beta}] \tag{6.28}$$

ή

$$\exp[-(w_i/c)^{\beta}]=1-F(w_i) \tag{6.29}$$

που με λογαρίθμηση (δύο φορές) δίνει:

$$\beta\ln\omega_i-\beta\ln c=\ln\{-\ln[1-F(\omega_i)]\} \tag{6.30}$$

θέτοντας $\beta=-\beta\ln c$, $A=\beta$, $\kappa=\ln\omega$, $Y=\ln[\{-\ln[1-F(\omega)]\}$ η (6.30) παίρνει την μορφή:

$$Y_i=A*u_i+B \tag{6.31}$$

από την οποία με εφαρμογή της μεθόδου των ελαχίστων τετραγώνων (παρ. 5.8.1.) προκύπτουν οι τιμές των A και B, ίσες με:

$$A=Cov(U,Y)/Var(Y)^{0.5} \qquad B=E(Y)-A*E(U) \tag{6.32}$$

οπότε:

$$\beta=Cov(\ln w, \ln\{-\ln[1-F(w)]\})/[Var(\ln w)]^{0.5} \tag{6.33}$$

και

$\beta*\ln c=E(\ln\{-\ln[1-F(\omega)]\}$ (6.34)

από τις οποίες προκύπτει η τιμή του c και κατόπιν η τιμή της παραμέτρου α. Η μέση τιμή E(X) και η διακύμανση της τυχαίας μεταβλητής X≈W(X,α,β) είναι:

$E(X)=\alpha^{-(1/\beta)}*\Gamma(1+1/\beta)$ (6.35)

$Var(X)=\alpha^{-(2/\beta)}*\{\Gamma(1+2/\beta)-[\Gamma(1+1/\beta)]^2\}$ (6.36)

(όπου Γ(x) η συνάρτηση Γ που αναπτύσσεται στην επόμενη παράγραφο).

6.4. Η Γάμα κατανομή -Εκθετική κατανομή.

Η συνάρτηση πυκνότητας πιθανότητας f(x) της θετικής συνεχούς τυχαίας μεταβλητής X που ακολουθεί την Γάμα κατανομή με παραμέτρους β και γ (δηλαδή X≈G(X,β,γ)), είναι η:

$f(x)=x^{\gamma-1}*e^{-x/\beta}/[\beta^{\gamma}*\Gamma(\gamma)]$ με $\chi>0,\ \beta>0,\ \gamma>0$ (6.37)

με

$E(x)=\beta*\gamma$, και $\sigma_x^2=Var(X)=\beta^2*\gamma$ (6.38)

Θέτοντας A=ln[E(x)]-E(lnx) μια "καλή" (optimum) εκτίμηση των β και γ δίνεται (Panofsky and Briet 1963), από τις:

$\gamma'=[1+SQRT(1+0.75*A)]/(4*A)$ $\beta'=E(X)/\gamma'$ (6.39)

Η Γ(x) είναι συνάρτηση "γάμα" με τιμές:

$$\Gamma(x)=\int_0^{+\infty} t^{x-1}*e^{-t}\,dt \qquad (6.40)$$

από την οποία προκύπτουν οι:

Γ.1 $\Gamma(x)=(x-1)!$ αν $\chi\in N^*$

Γ.2 $\Gamma(x)=(x-1)*\Gamma(x-1)$ για $x>1$

Γ.3 $\Gamma(x)=(x-[x])*\Gamma\{x-[x]\}*[\chi-1]!$

Γ.4 $\Gamma(1/2)=SQRT(\pi)$

Η κατανομή Γάμα χρησιμοποιείται κυρίως σε υδρολογικές παραμέτρους, όπως η βροχή, η τάση των υδρατμών, η εξάτμηση κλπ.

Ειδική περίπτωση της κατανομής G(x,β,γ) αποτελεί η περίπτωση κατά την οποία β=1/λ και γ=1, οπότε η (6.37) δίνει:

$f(x)=\lambda e^{-\lambda x}$ $x>0,\ \lambda>0$ $(f(x)=0$ για $x\leq 0)$ (6.41)

με συνάρτηση αθροιστικής πιθανότητας:

$$F(x)=\int_0^{x} f(t)dt=1-e^{-\lambda x} \quad (F(x)=0 \text{ για } x\leq 0) \qquad (6.42)$$

και μέση τιμή και διακύμανση, τις:

$E(x)=1/\lambda$ και $Var(X)=1/\lambda^2$ (6.43)

αντίστοιχα.

Η τυχαία μεταβλητή X που έχει (σ.π.π.) την f(x) που δίνεται από την (6.40) λέμε ότι ακολουθεί την εκθετική κατανομή με παράμετρο λ>0 και γράφεται X≈Exp(X,λ).

6.5. Η κατανομή του Gumbel. Περίοδος επιστροφής.

Η συνάρτηση αθροιστικής συχνότητας F(x) της τυχαίας μεταβλητής X που ακολουθεί την κατανομή του Gumbel (Remenieras 1959) δίνεται από την:

$$F(x)=\exp\{-\exp[-\alpha*(\chi-\chi_o)/\sigma]\} \qquad (6.44)$$

με $\alpha=1/(0.78*\sigma)$ και $\chi_o=E(X)-0.577/\alpha$

όπου E(X) η μέση τιμή των τιμών της X και $\sigma=[Var(x)]^{0.5}$ η τυπική της απόκλιση.

Από την (6.44) με εφαρμογή της (3.18) προκύπτει εύκολα η συνάρτηση πυκνότητας πιθανότητας της X.

Την κατανομή αυτή χρησιμοποιούμε για τις ελάχιστες και μέγιστες τιμές της παραμέτρου X σε ορισμένη περίοδο, π.χ. δεκαήμερο, μήνα, έτος. Ο Μονόπωλης 1978 ορίζει το x_o ίσο με το Mode των τιμών της X.

Η $F(x)=P(-\infty<X\leq x)$ εκφράζει την πιθανότητα του γεγονότος:

Α={η τιμή της X παραμένει μικρότερη από x}.

Σε περίπτωση που μελετούνται οι μέγιστες τιμές σε διάφορα, ξένα μεταξύ τους χρονικά διαστήματα (έτη, μήνες, δεκαήμερα), ενδιαφέρεται κανείς για την πιθανότητα του γεγονότος:

Β={η τιμή της X είναι μεγαλύτερη από x}

που ισούται με 1-P. Ο αριθμός αυτός συνδέεται με την περίοδο επιστροφής (δηλαδή μέσο χρονικό διάστημα που μεσολαβεί μεταξύ δύο διαδοχικών εμφανίσεων του γεγονότος B, ή κάποιου άλλου γεγονότος) με την σχέση:

$$T=1/(1-p) \qquad (6.45)$$

Η περίοδος επιστροφής T (Ruturn period) της τιμής x_i με $x_1<x_2<...<x_n$ μιας τυχαίας μεταβλητής X με τιμές $X(t)=x_i=\max\{\omega_i/t\in t_i\}$, τις μέγιστες τιμές (ή ελάχιστες τιμές οπότε $X(t)=\min\{w_i/t\in t_i\}$) μιας παραμέτρου Ω στις n ίσης διάρκειας χρονικές περιόδους $t_i\in I_n$, με $\Omega(t)=\omega_t$ δίνεται με καλή προσέγγιση από την:

$$T(x_i)=i/(n+1) \qquad (6.46)$$

όπου i είναι η τάξη του x_i. Από τις (6.44) και (6.45), προκύπτει ότι:

$$1-p=1-F(x)=\exp[-\exp(-y)]=1/T \quad \text{με } y=\alpha*(\chi-\chi_o) \qquad (6.47)$$

και ακόμα

$$y=-\{0.834+2.3*\ln[\ln(T/(T-1))]\} \qquad (6.48)$$

Τιμές του Y συναρτήσει του T με βάση την (6.48) δίνονται στον Πίνακα 6.5. Για την εύρεση της τιμής X_T της X για περίοδο επανάληψης T ο Chow 1951 δίνει τον τύπο:

$$X_T=E(X)+K*S_x \qquad (6.49)$$

όπου E(X) και S_x η μέση τιμή και η τυπική απόκλιση των διτιθε-μένων τιμών (μεγίστων ή ελαχίστων ορισμένης περιόδου) και K παράγοντες που παίρνει τιμές συνάρτήσει των T και n

$$K=(y_T-Y_n)/S_n \qquad (6.50)$$

Πίνακας 6.5. Τιμές του $y=\alpha*(x-x_o)$ συναρτήσει της περιόδου επιστροφής Τ.

T	2	5	10	25	50	100
y	0.3665	1.4994	2.2502	3.1985	3.9019	4.6001

όπου Y_n, S_n η μέση τιμή και ητυπική απόκλιση των n τιμών y_i που προκύπτουν από την (6.47) για Τ από 2 μέχρι n+1, και y_T η αντίστοιχη τιμή του y για την περίοδο επιστροφής Τ, για την οποία ζητάτε η τιμή X_T.

Για την ανάλυση των άκρων τιμών απαιτείται μεγάλος αριθμός ομοιογενών παρατηρήσεων και η χρονοσειρά των τιμών να μην παρουσιάζει θετικές ή αρνητικές τάσεις. Ακόμη η εμφάνιση κάποιας άκρας τιμής πρέπει να είναι ανεξάρτητη της προηγούμενης της. Ετσι π.χ. για τον έλεγχο των ετήσιων άκρων θερμοκρασιών η εμφάνιση της ελάχιστης τιμής του έτους YEAR την 31 Δεκεμβρίου και αυτής του YEAR+1 την 1 Ιανουαρίου δεν μπορεί να θεωρηθεί σαν ανεξάρτητα γεγονότα, αφού μπορεί να είναι αποτέλεσμα του ίδιου συστήματος. Ακόμη το πλήθος των διατιθεμένων τιμών πρέπει να είναι αρκετά μεγάλο, και οπωσδήποτε όχι μικρότερο από 15.

6.6. Η Διωνυμική κατανομή.

Εστω S ο δειγματικός χώρος ενός πειράματος μετρήσεων των τιμών μιας παραμέτρου και $Q=\{A,B\}$ ένας διαμελισμός του S σε δύο σύνολα. Ορίζουμε την τυχαία μεταβλητή Χ από την:

$$X(t)=\begin{cases} 1 & \text{όταν} \quad \chi \in A \\ 0 & \text{όταν} \quad \chi \in B=C_s A \end{cases} \tag{6.51}$$

όπου xt η τιμή της παραμέτρου κατά την μέτρηση την χρονική στιγμή t (η Χ λέγεται **διχοτόμος ή χαρακτηριστική συνάρτηση** του συνόλου Α). Υποθέτουμε ακόμη ότι οι τιμές x_t, x_τ είναι ανεξάρτητες μεταξύ τους και ότι η πιθανότητα των συνόλων Α και Β είναι σταθερές:

$$P(x_t \in A)=P(A)=p \text{ και } P(x_t \in B)=P(x_t \in C_s A)=P(B)=q=1-p \tag{6.52}$$

Μια παράμετρος που πληρεί την (6.52) και της οποίας οι τιμές σε διάφορους χρόνους είναι ανεξάρτητες μεταξύ τους ονομάζεται **διαδικασία Bernoulli**. Θεωρούμε τώρα την τυχαία μεταβλητή:

$$Y(M)=\Sigma\ X(x_t) \quad t \in I_M \tag{6.53}$$

που δίνει το άθροισμα των τιμών της Χ σε Μ μετρήσεις.

Αν Κ είναι το πλήθος των $x_t \in A$, το πλήθος των $x_t \in B$ θα είναι ίσο με Μ-Κ και εξαιτίας της ανεξαρτησίας των τιμών της παραμέτρου, η πιθανότητα του γεγονότος:

Γ={από τις Μ τιμές της διατεταγμένης Μιαδας $(x_1, x_2, \ldots x_M)$, ακριβώς Κ τιμές ανήκουν στο σύνολο Α}

είναι ίση με:

$$P(\Gamma)=P(x_1)*P(x_2)*\ldots*P(x_M)=p^K*q^{M-K} \tag{6.54}$$

Επειδή το πλήθος των Μιάδων με Κ τιμές $x_i \in A$ είναι ίσο με

(M:K) προκύπτει ότι:

$$f(K)=P[Y(M)=K]=(M:K)*p^{K}*q^{M-K} \quad (6.55)$$

όπου:

$$(M:K)=\binom{M}{K}= M!/[K!*(M-K)!]$$

Η μέση τιμή των περιπτώσεων στις οποίες $x_t \in A$ σε M μετρήσεις των τιμών της παραμέτρου, και η διακύμανσή τους είναι αντίστοιχα:

$$E(y)=M*P \text{ και } Var(Y)=M*p*q \quad (6.56)$$

Η διακριτή τυχαία μεταβλητή Y, που έχει συνάρτηση πιθανότητας που δίνεται από την (6.55) και δίνει τον αριθμό K των "επιτυχιών" (σαν επιτυχία ορίζουμε την περίπτωση $X(x_t)=1$) σε M μετρήσεις των τιμών μιας παραμέτρου λέμε ότι ακολουθεί την διωνυμική κατανομή και γράφουμε:

$$Y \approx B(Y,M,K).$$

<u>Παράδειγμα 6.2.</u>

Η πιθανότητα να "κλείσει" η Μαλακάσα εξ εξαιτίας χιονιού κατά το χειμώνα είναι P=0.05. Ποιά η πιθανότητα να είναι κλειστή η Μαλακάσα το χειμώνα τουλάχιστον 2 φορές σε διάστημα 10 ημερών; Υπο έτουμε ότι τα γεγονότα {κλειστή η Μαλακάσα} είναι ανεξάρτητα μεταξύ τους. Με την υπόθεση αυτή και με M=10 K≥2 προκύπτει:

$$P(Y(10)\geq 2)=1-P[Y(10)=0]-P[Y(10)=1]=0.086$$

6.7. <u>Γεωμετρική κατανομή -Αρνητική διωνυμική κατανομή.</u>

Εστω ότι σε μια διαδικασία Bernoulli ζητάτε ο αριθμός κ των απαιτούμενων επαναλήψεων για να έχουμε την πρώτη επιτυχία. Τότε οι κ τιμές της παραμέτρου θα επαληθεύουν:

$x_1 \in B$, $x_2 \in B$, ...$x_{\kappa-1} \in B$ και $x_\kappa \in A$ και το γεγονός:

$$\Gamma=\{(x_1,x_2,x_3...x_\kappa)/x_i \in B \text{ με } i \in I_{\kappa-1} \text{ και } x_\kappa \in A\} \quad (6.57)$$

εξαιτίας της ανεξαρτησίας των τιμών της μετρούμενης παραμέτρου θα έχει πιθανότητα:

$$f(\kappa)=P(\kappa)=P(X(x_\kappa)=1)*\prod_{i=1}^{\kappa-1} P(X(x_i))=p*q^{1-\kappa}=p*(1-p)^{1-\kappa} \quad (6.58)$$

Η μέση τιμή και η διακύμανση του αριθμού κ των απαιτούμενων μετρήσεων για την εμφάνιση της πρώτης επιτυχίας είναι:

$$E(\Gamma)=E[X(\chi_\kappa)=1 \text{ όταν } X(\chi_i)=0 \text{ για } i \in I_{\kappa-1}]=1/p \quad (6.59)$$

$$Var(\Gamma)=(1-p)/p^2 \quad (6.60)$$

Η τυχαία μεταβλητή Γ με (σ.π.π.) την f(κ) που δίνεται από την (6.58) λέμε ότι ακολουθεί την Γεωμετρική κατανομή με παράμετρο p. Εάν αντί για τον αριθμό κ των απαιτούμενων μετρήσεων των τιμών της παραμέτρου, μέχρι την εμφάνιση της πρώτης επιτυχίας ζητάμε τον αριθμό κ των μετρήσεων μέχρι την ν επιτυχία, τότε ή συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής της οποίας η τιμή είναι ο αριθμός κ≥ν δίνεται από την:

$$f(\kappa,\nu)=(\kappa-1:\nu-1)*p^{\nu}*q^{\kappa-\nu} \quad \kappa\geq\nu \quad (6.61)$$

και η τυχαία μεταβλητή Y(κ,ν) που έχει (σ.π.π.) την f(x) που δίνεται από την (6.61) λέμε ότι ακολουθεί την αρνητική διωνυμική κατανομή (ή κατανομή του Polya και οι παράμετροι της είναι οι p και v. Για v=1 ή (6.61) δίνει την (6.58), δηλαδή η γεωμετρική κατανομή είναι μια ειδική περίπτωση της κατανομής του Polya. Η μέση τιμή και η διακύμανση της Y είναι:

$$E(Y)=v/p \quad \text{και} \quad Var(Y)=v*(1-p)/p^2 \tag{6.62}$$

6.8. Ομοιόμορφη κατανομή.

Αν η συνάρτηση πυκνότητας πιθανότητας μιας συνεχούς τυχαίας μεταβλητής X είναι θετική και σταθερή σ'ένα διάστημα [α,β] β>α και έχει μηδενική τιμή εκτός αυτού τότε λέμε ότι ακολουθεί την ομοιόμορφη κατανομή στο διάστημα [α,β] και γράφουμε: X≈U(α,β). Σύμφωνα με τα παραπάνω, η (σ.π.π.) της τυχαίας μεταβλητής που ακολουθεί ομοιόμορφη κατανομή δίνεται από την:

$$f(x)=\begin{cases} 1/(\beta-\alpha) & \text{όταν } \chi\in[\alpha,\beta] \\ 0 & \text{όταν } \chi\in R-[\alpha,\beta] \end{cases} \tag{6.63}$$

Η μέση τιμή και η διακύμανση της τυχαίας μεταβλητής που ορίζεται από την (6.63) δίνονται από τις:

$$E(X)=(\alpha+\beta)/2 \tag{6.64}$$

$$Var(X)=(\beta-\alpha)^2/2 \tag{6.65}$$

και η συνάρτηση αθροιστικής πιθανότητας της θα δίνεται από την:

$$f(x)=\begin{cases} 0 & \text{όταν } \chi<\alpha \\ (\chi-\alpha)/(\beta-\alpha) & \text{όταν } \chi\in(\alpha,\beta) \\ 1 & \text{όταν } \chi\geq\beta \end{cases} \tag{6.65}$$

6.9. Η t_v κατανομή ή κατανομή του Student.

Η συνεχής τυχαία μεταβλητή X_v που έχει (σ.π.π.) την:

$$f(x,v)=\frac{\Gamma[(v+1)/2]*(1+\chi^2/v)^{-[(v+1)/2]}}{\Gamma(v/2)*SQRT(v*\pi)} \tag{6.66}$$

λέμε ότι ακολουθεί την t_v-κατανομή με v βαθμούς ελευθερίας.

Η μέση τιμή της και η διακύμανση της (που ορίζεται για v≥3) δίνονται από τις:

$$E(X_v)=0 \quad Var(X)=v/(v-2) \tag{6.67}$$

Για v—>+∞ (και ουσιαστικά για v≥30) οι τιμές της f(x,v) είναι ανεξάρτητες του v και με πολύ καλή προσέγγιση ισχύει:

$$f(x,v)=f(x)=e^{-x^2/2}/SQRT(2*\pi) \quad \text{όταν } v\geq30 \tag{6.68}$$

δηλαδή ίση με την (σ.π.π.) της κανονικοποιημένης τυχαίας μεταβλητής (εξίσωση 6.4).

Η t_v-κατανομή αποτελεί ένα από τα σημαντικότερα βοηθήματα της στατιστικής ανάλυσης, αφού σ'αυτή στηρίζεται, το ονομαζόμενο, t-test (ή Student test), που εφαρμόζεται για τον έλεγχο στατιστικών υποθέσεων. Δίνουμε χωρίς απόδειξη τα παρακάτω δύο θεωρήματα.

Θεώρημα 6.1. Αν η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή ($X \approx N(x,\mu,\sigma^2)$ και $x_1,x_2,\ldots x_M$ είναι ένα τυχαίο δείγμα, μεγέθους M, των τιμών της και E(X), S_x είναι η μέση τιμή και η τυπική απόκλιση των M τιμών του δείγματος τότε η τυχαία μεταβλητή t_{M-1} που ορίζεται από την:

$$t_{M-1}=[E(X)-\mu]/[S_x*SQRT(M)] \tag{6.69}$$

ακολουθεί την t κατανομή με v=M-1 βαθμούς ελευθερίας.

Θεώρημα 6.2. Εστω X και Y δύο ανεξάρτητες τυχαίες μεταβλητές με:

$X \approx N(X,\mu_x,\sigma_x^2)$ και $Y \approx N(Y,\mu_y,\sigma_y^2)$

Αν $x_1,x_2,\ldots x_\kappa$ ένα τυχαίο, μεγέθους κ, δείγμα από τις τιμές της X, $y_1,y_2,\ldots y_\lambda$ ένα, μεγέθους λ, τυχαίο δείγμα από τις τιμές της Y και E(X), E(Y), S_x και S_y είναι αντίστοιχες μέσες τιμές και οι αντίστοιχες τυπικές αποκλίσεις που υπολογίζονται με βάση τις τιμές των δειγμάτων, τότε η τυχαία μεταβλητή:

$$t_{\kappa+\lambda-2}=\frac{[E(X)-E(Y)]-(\mu_x-\mu_y)}{[(\kappa-1)S_x^2+(\lambda-1)S_y^2]^{0.5}} * \frac{[(\kappa+\lambda-2)*\kappa*\lambda]^{0.5}}{(\kappa+\lambda)^{0.5}} \tag{6.70}$$

ακολουθεί την t-κατανομή με v=κ+λ-2 βαθμούς ελευθερίας. Ο πίνακας 6.3 περιέχει, συναρτήσει του v, για διάφορες τιμές του αριθμού α τις τιμές $t_{v,a}$ για τις οποίες:

$$P(|t_v|>t_{v,a})=a \text{ ή } P(t_v<-t_{v,a})+P(t_v>t_{v,a})=2a \text{ ή}$$
$$P(-t_{v,a}<t_v<t_{v,a})=1-2*a \tag{6.71}$$

6.9 Η X^2 κατανομή (διαβάζεται Xi-τετράγωνο).

Η (6.37) για β=2 και γ=v/2 γίνεται

$$f(x,v)=x^{v/2-1}e^{-x/2}/2^{v/2}*\Gamma(v/2) \tag{6.72}$$

Η τυχαία μεταβλητή X της οποίας η (σ.π.π.) δίνεται από την (6.72) λέμε ότι ακολουθεί την X^2 κατανομή με v βαθμούς ελευθερίας. Η μέση τιμή και η διακύμανσή της, όπως προκύπτει από την (6.38) για β=2 και γ=v/2 είναι:

$E(X_v^2)=v$ και $Var(X_v^2)=2*v$ αντίστοιχα.

Στον πίνακα 6.4 δίνονται οι τιμές $X^2_{v,a}$, συναρτήσει των v και α, για τις οποίες:

$$P(X_v^2>X^2_{v,a})=a \tag{6.73}$$

Ο πίνακας 6.4 περιέχει τιμές του v≤100, διότι για v>100, είναι $X_v^2(X^2,v,2v)$, ενώ στο α δίνονται μόνο οι τιμές 0,1, 0.05, 0.025, 0.01 και 0.005 και οι συμπληρωματικές τους ως προς την μονάδα. Οι τιμές αυτές είναι οι πλέον χρησιμοποιούμενες στην στατιστική ανάλυση. Για v>30 ο Fisher (Gramer,1946) απέδειξε ότι η τυχαία μεταβλητή $X=SQRT(2*X_v^2)$, με πολύ καλή προσέγγιση ακολουθεί την κανονική κατανομή $X \approx N(X,SQRT(2v-1),1)$ και η τυχαία μεταβλητή $Z=X-SQRT(2*v-1)$ είναι: $Z \approx N(Z,0,1)$.

Η X_v^2 κατανομή (μαζί με την κανονική και την t_v κατανομή), είναι από τις πλέον χρησιμοποιούμενες στην στατιστική ανάλυση, και ο τρόπος που χρησιμοποιείται στον έλεγχο της προσαρμογής των δεδομένων στην κανονική κατανομή δόθηκε στην παρ.6.1.2. Παραδείγματα της εφαρμογής της X_v^2 κατανομής τόσο στον έλεγχο προσαρμογής όσο και στον έλεγχο της ανεξαρτησίας και της ομοιογένειας δίνονται στο τεύχος 5/9 του ΥΔΡΟΣΚΟΠΙΟ, όπου δίνονται και σχετικά με την κατανομή X^2 θεωρήματα.

6.10. <u>Η F-κατανομή. Ανάλυση της διασποράς.</u>

6.10.1 Η F-κατανομή.

Εστω η τυχαία μεταβλητή X με (σ.π.π.) την:

$$f(x)=\begin{cases} \Gamma(\kappa,\lambda)*\chi^{\kappa/2-1}*(1-\chi)^{\lambda/2-1} & \text{όταν } 0<\chi<1 \\ 0 & \text{όταν } \chi\le 0 \text{ ή } \chi\ge 1 \end{cases} \tag{6.74}$$

όπου:

$$\Gamma(\kappa,\lambda)=\Gamma[(\kappa+\lambda)/2]/[\Gamma(\kappa/2)*\Gamma(\lambda/2)] \tag{6.75}$$

θέτοντας:

$$Y=\kappa*X/[\lambda*(1-X)] \tag{6.76}$$

η (σ.π.π.) της Y δίνεται από την:

$$f(y)=\begin{cases} \Gamma(\kappa,\lambda)*y^{\lambda/2-1}*(\kappa/\lambda)^{\kappa/2}*(1+\kappa*y/2) & \text{όταν } y>0 \\ 0 & \text{όταν } y\le 0 \end{cases} \tag{6.77}$$

Η τυχαία μεταβλητή Y της οποίας η (σ.π.π.) είναι η f(y) που δίνεται από την (6.77) λέμε ότι ακολουθεί την F κατανομή με κ,λ βαθμούς ελευθερίας και γράφουμε **$Y \approx F(Y,\kappa,\lambda)$**.

Ο πίνακας 6.6 δίνει τις τιμές του $F_{\kappa,\lambda,\alpha}$ για διάφορες τιμές των κ και λ και για α=0.1 και α=0.05, για τις οποίες:

$$P(F>F_{\kappa,\lambda,\alpha})=\alpha \tag{6.78}$$

Η κατανομή F χρησιμοποιείται για τον έλεγχο της τυπικής απόκλισης, όπως περιγράφεται στο τεύχος 5/9 του ΥΔΡΟΣΚΟΠΙΟ, όπου δίνεται και παράδειγμα εφαρμογής του F-test, όπως ονομάζεται ο έλεγχος που στηρίζεται στην F-κατανομή, καθώς και τα σχετικά θεωρήματα στα οποία βασίζεται το F-test. Τέλος την κατανομή αυτή χρησιμοποιούμε για τον έλεγχο της ομοιγένειας μέσω της ανάλυσης της διασποράς, μέθοδος που συνοπτικά περιγράφεται αμέσως παρακάτω, επειδή συμπληρώνει τις μεθόδους ελέγχου ομοιογένειας που αναπτύχθηκαν στο τεύχος 5/9 του ΥΔΡΟΣΚΟΠΙΟ.

6.10.2. Ανάλυση της διασποράς

Εστω η χρονοσειρά $x_1,x_2,\ldots x_M$, M τιμών μιας παραμέτρου X που "μετρήθηκαν" κατά την διάρκεια ενός πειράματος. Υποθέτουμε ότι οι συνθήκες του λήψης των τιμών δεν ήταν οι ίδιες καθ'όλη την διάρκεια του πειράματος και έστω:

$$\left.\begin{array}{l} a_1(1), a_1(2), \dots a_1(i_1) \\ a_2(1), a_2(2), \dots a_2(i_2) \\ \dots\dots\dots\dots\dots\dots \\ \dots\dots\dots\dots\dots\dots \\ a_k(1), a_k(2), \dots a_k(i_k) \end{array}\right\} \qquad (6.79)$$

οι $i_1, i_2, \dots, i_k$ τιμές της X που μετρήθηκαν στις k συνθήκες $A_1, A_2, \dots, A_k$ με:

$$i_1+i_2+i_3+\dots i_k=\Sigma i_j=M \text{ με } j\in I_k \qquad (6.80)$$

<u>Παράδειγμα 6.3.</u>

Εστω ότι οι x_i, $i\in I_M$ είναι οι τιμές της ελάχιστης θερμοκρασίας ενός σταθμού ο οποίος μετά από ορισμένο χρόνο λειτουργίας του στη θέση Σ_1 (i_1 τιμές με συνθήκες μετρήσεων A_1), μεταφέρθηκε στην θέση Σ_2 (συνθήκες A_2) όπου άλλαξε το ελαχιστοβάθμιο θερμόμετρο με το οποίο πραγματοποιούσε τις μετρήσεις του μετά από i_2 μετρήσεις και αφού πραγματοποίησε i_3 μετρήσεις με το νέο θερμόμετρο (συνθήκες A_3), άλλαξε τον μετεωρολογικό κλωβό, με νέου τύπου, όπου κάνει ακόμα τις μετρήσεις του (i_4 τιμές σε συνθήκες A_4). Στην περίπτωση αυτή υπάρχουν k=4 διαφορετικές συνθήκες μετρήσεων και θέτοντας:

$I=i_1$, $J=i_1+i_2$, $K=J+i_3$

θα έχουμε:

$a_1(1)=x_1$, $a_1(2)=x_2, \dots,$ $a_1(i_1)=x_I$

$a_2(1)=x_{I+1}$, $a_2(2)=x_{I+2}, \dots,$ $a_2(i_2)=x_J$

$a_3(1)=x_{J+1}$, $a_3(2)=x_{J+2}, \dots,$ $a_3(i_3)=x_K$

$a_4(1)=x_{K+1}$, $a_4(2)=x_{K+2}, \dots,$ $a_4(i_4)=x_M$

<u>Παράδειγμα 6.4.</u>

Εστω T_1, $T_2, \dots,$ T_M και P_1, $P_2, \dots,$ P_M οι M τιμές της θερμοκρασίας και της πίεσης της 12UTC κατά τον Ιούλιο (M=31*NETH όπου NETH τα έτη λειτουργίας του σταθμού), και P_{min} και P_{max} η ελάχιστη και η μέγιστη πίεση την 12UTC στον σταθμό κατά τον Ιούλιο στη διάρκεια των NETH ετών. Θεωρούμε τον διαμελισμό Q_P του συνόλου $[P_{min}, P_{max}]$ σε τέσσερα ισομήκη διαστήματα-καταστάσεις, τα $B_1=[P_{min}, P_1]$, $B_2=(P_1, P_2]$, $B_3=(P_2, P_3]$ και $B_4=(P_3, P_{max}]$, όπου:

$P_i=(P_{max}-P_{min})*i/4$, $i\in I_3$.

Το σύνολο I_M των M ημερών διαμελίζεται στα τέσσερα υποσύνολά του A_i, $i\in I_4$ όπου:

$A_i=\{j\in I_M \text{ και } P_j\in B_i\}$

Με βάση τα σύνολα-συνθήκες A_i δημιουργούνται οι i_1, i_2, i_3 και i_4 τιμές της μορφής $a_i(l)$, $i\in I_4$, $l\in I_s$ όπου $s\in\{i_1, i_2, i_3, i_4\}$.

Από τα πλέον συνηθισμένα προβλήματα που ανακύπτουν κατά την στατιστική ανάλυση των χρονοσειρών της μετεωρολογικής παραμέτρου (και οποιασδήποτε άλλης παραμέτρου) X, είναι και τα παρακάτω:

1. Μια ή περισσότερες από τις συνθήκες μέτρησης των τιμών της παραμέτρου, επιδρά κατά τρόπο που οι προκύπτουσες τιμές να

διαφέρουν σημαντικά από τις αντίστοιχες τιμές της παραμέτρου που θα είχαν ληφθεί με άλλες συνθήκες;

2. Εχουν όλες οι τιμές της χρονοσειράς ληφθεί κάτα από τις ίδιες συνθήκες μετρήσεων ή μ'άλλα λόγια είναι ομοιογενής η χρονοσειρά των τιμών της X;

3. Οι αλλαγές που γίνονται στο περιβάλλον του σταθμού εξαιτίας ανθρώπινων δραστηριοτήτων (πχ άνοιγμα δρόμου, επέκταση οικιστικής περιοχής κλπ) επέδρασαν σημαντικά στις τιμές ορισμένων μετεωρολογικών παραμέτρων, οπότε οι αντίστοιχες χρονοσειρές των τιμών τους να μην μπορούν να θεωρηθούν ομοιγενείς;

Απάντηση σε προβλήματα της μορφής αυτής μπορεί να δώσει η εφαρμογή του t-test, όταν οι συνθήκες είναι δύο μόνο και γνωστές, ενώ ανεξάρτητα από το πλήθος των συνθηκών την απάντηση μπορεί κανείς να την αναζητήση με εφαρμογή των ελέγχων ομοιογένειας (που δόθηκαν στο τεύχος 5/9 του ΥΔΡΟΣΚΟΠΙΟ), καθώς και η αναλυόμενη παρακάτω μέθοδος της ανάλυσης της διασποράς (ή της διακύμανσης), γνωστής στην διεθνή βιβλιογραφία σαν ANOVA. Επειδή κατά την ανάλυση της διασποράς για τον έλεγχο της ομοιογένειας των χρονοσειρών ενδειαφερόμαστε κυρίως για την επίδραση της παρόδου των ετών (ένας παράγοντας οχρόνος), δίνεται παρακάτω η μέθοδος της ανάλυσης της διασποράς με ένα παράγοντα. Για την ταυτόχρονη επίδραση περισσότερων παραγόντων μπορεί κανείς να ανατρέξει στους Κούνια, κ.α. 1984, Kendal and Stuart 1969, Αδαμόπουλος 1963, καθώς και σε άλλα εγχειρίδια της στατιστικής.

Για την ανάλυση της διασποράς οι M τιμές χ_i, $i \in I_M$ της χρονοσειράς γράφονται με την μορφή (6.79) και υπολογίζονται οι K μέσοι όροι των i_m tim;vn $a_i(i_m)$ της X στις συνθήκες A_m:

$$E(A_K = \Sigma\, a_m(j)/i_m \quad j \in I_i \tag{6.81}$$

και ο γενικός μέσος όρος a της χρονοσειράς:

$$E(X) = \Sigma \chi_i/M \quad \text{με } i \in I_M \tag{6.82}$$

Κατόπιν υπολογίζεται το άθροισμα των τετραγώνων των διαφορών:

$$S_o^2 = \Sigma\, [x_j - E(X)]^2 \quad j \in I_M \tag{6.83}$$

$$S_\upsilon^2 = \{\Sigma\, x_j^2 \text{ με } j \in I_M\} - \{\Sigma\, E^2(A_i) \text{ με } i_K\} \tag{6.84}$$

και τέλος το:

$$S_\alpha^2 = S_o^2 - S_\upsilon^2 = \{\Sigma\, E^2(A_{i)\ \text{με}\ i \epsilon I K}\} - E^2(X) \tag{6.85}$$

και τέλος ο αριθμός:

$$F_{\nu,\lambda} = (S_\alpha^2/\nu)/(S_\upsilon^2/\lambda) \tag{6.86}$$

όπου $\nu = K-1$ και $\lambda = M-K$ οι βαθμοί ελευθερίας των S_α και S_o αντίστοιχα.

Η S_o ονομάζεται ολική μεταβολή, η S_α μεταβολή μεταξύ των δειγμάτων (και είναι αυτή που οφείλεται στην επίδραση των συνθηκών-παραγόντων του πειράματος) και η S_υ ονομάζεται υπόλοιπη μεταβολή ή/και σφάλμα (και οφείλεται σε τυχαίους παράγοντες που δεν είναι δυνατόν να ελεγχθούν).

Για κάθε μια από τις S_o, S_α, S_υ ονομάζεται μέση μεταβολή το ηλίκο της με τους αντίστοιχους βαθμούς ελευθερίας της, που είναι $M-1$ για την S_o, $K-1$ για την S_a και $M-k$ για την S_u.

Αποδεικνύεται ότι ο $F_{v,\lambda}$ ακολουθεί την κατανομή F με τους αντίστοιχους βαθμούς ελευθερίας. Αν για δοθείσα στάθμη σημαντικότητας (σ.σ): α (κυρίως α=0.05) συμβεί:

$$F_{v,\lambda} > F_{v,\lambda,a} \qquad (6.87)$$

η υπόθεση H_o (H_o: Οι K μέσες τιμές $E(A_i)$ με $i \in I_K$ των τιμών της παραμέτρου X που υπολογίστηκαν από τις $i_1, i_2, \ldots, i_K$ τιμές, που προέκυψαν από μετρήσεις στις συνθήκες $A_1, A_2, \ldots, A_K$ αντίστοιχα, δεν διαφέρουν σημαντικά σε στάθμη σημαντικότητας α) απορρίπτεται και γίνεται δεκτή η εναλλακτική της υπόθεση H_1 (H_1: κάποια από τις k μέσες τιμές διαφέρει σημαντικά από τις άλλες).

Οι τιμές $F_{v,\lambda,a}$ δίνονται στον πίνακα της F κατανομής στο παράρτημα των στατιστικών πινάκων.

Παράδειγμα. 6.5

Στον πίνακα 6.9 δίνονται οι μέσες μηνιαίες τιμές της ελάχιστης θερμοκρασίας κατά τα έτη 1961 μέχρι 1990, στον σταθμό της Λάρισσας. Ζητάτε να απαντήσουμε στο ερώτημα;

Είναι η μέση ελάχιστη θερμοκρασία κάποιου έτους σημαντικά διαφορετική από την κανονική μέση ελάχιστη θερμοκρασία του σταθμού;

Η υπόθεση H_o στην περίπτωση αυτή είναι:

H_o: $T_{61}=T_{62}=T_{63}=\ldots=T_{90}=\mu(Tmin)$

σε στάθμη σημαντικότητας (σ.σ): α, όπου $\mu(T_{min})=8.7°C$, η κανονική τιμή της ελάχιστης θερμοκρασίας στη Λάρισσα.

Εργαζόμενοι όπως περιγράφεται παραπάνω βρίσκουμε:

$i_1=i_2=\ldots=i_{30}=12$, $k=30$, $M=360$ και

$(S_a^2/v)=2.24$, $(S_u^2/\lambda)=4.12$ και $v=29$, $\lambda=330$, $F_{v,\lambda}=0.055$

Επειδή $0.055=F_{v,\lambda} \leq F_{29,330,0.05} \approx 1.46$

η H_o γίνεται δεκτή στην ζητούμενη στάθμη σημαντικότητας και επομένως μπορούμε να πούμε ότι η σειρά των τιμών της ελάχιστης θερμοκρασίας στον παραπάνω σταθμό είναι ομοιογενής.

6.11. Αλλες κατανομές.

6.11.1. Πολυωνυμική κατανομή.

Αν η τυχαία μεταβλητή X έχει πεδίο τιμών το διάστημα [α,β] και η συνάρτηση πυκνότητας πιθανότητας της δίνεται από την:

$$f(x)=a_\kappa x^\kappa+a_{\kappa-1}x^{\kappa-1}+\ldots+a_1x+a_o \quad \text{όταν } x\in[a,\beta] \qquad (6.88)$$

και $f(x)=0$ όταν $x<a$ ή $x>\beta$

λέμε ότι ακολουθεί πολυωνυμική κατανομή, με $\kappa+1$ παραμέτρους τις a_o, a_i $i \in I_\kappa$.

Η εύρεση των παραμέτρων από τις $M>\kappa+1$ τιμές ενός δείγματος γίνεται με την μέθοδο των ελαχίστων τετραγώνων. (παρ.5.8.1), όπου θέτουμε $X_1=X$, $X_2=X^2$, $X_3=X^3 \ldots X_\kappa=X^\kappa$

6.11.2 Λογαριθμοκανονική κατανομή (Log-Normal)

Αν X είναι μια θετική τυχαία μεταβλητή και η τυχαία μεταβλητή Y=lnX ακολουθεί την κανονική κατανομή ($Y \approx N(Y,\mu_Y,\sigma_Y^2)$) τότε η X λέμε ότι ακολουθεί την λογαριθμοκανονική κατανομή. Η (σ.π.π.) της X στην περίπτωση αυτή δίνεται (Κούνια 1984) από την:

$$f(x)=\frac{\exp[-(\ln x-\mu_Y)^2/(2*\sigma_Y^2)]}{x*\sigma*SQRT(2*\pi)} \qquad (6.89)$$

Στον πίνακα 6.7 δίνονται, για τις κατανομές που περιγράφονται στις προηγούμενες παραγράφους, οι συναρτήσεις πυκνότητας, οι παράμετροι που τις προσδιορίζουν, η μέση τιμή τους και η διακύμανσή τους. Στον πίνακα 6.8 δίνονται οι πλέον πιθανές κατανομές των σημαντικότερων μετεωρολογικών παραμέτρων.

7 ΣΤΑΤΙΣΤΙΚΑ ΜΕΤΕΩΡΟΛΟΓΙΚΩΝ ΠΑΡΑΜΕΤΡΩΝ.

7.1. Παράσταση απλών τιμών και στατιστικών παραμέτρων.

7.1.1. Παράσταση απλών τιμών.

Ονομάζουμε απλή τιμή μιας μετεωρολογικής παραμέτρου X, την τιμή που είχε η X μια ορισμένη χρονική στιγμή, σε κάποιο συγκεκριμένο σταθμό μέτρησης ή εκτίμησης ή καταγραφής ή υπολογισμού των τιμών της. Το σύνολο όλων των απλών τιμών της X, θεωρούμε ότι είναι αρχειοθετημένο σ'ένα πίνακα X πέντε διαστάσεων:

$$X=X\ (K,L,M^{\alpha},J,N) \qquad (7.1)$$

με στοιχεία:

$$x=X(\kappa,l,m^{\alpha},j,n) \qquad (7.2)$$

τις απλές τιμές της παραμέτρου, όπου:

K: Το πλήθος των σταθμών που διαθέτουν τιμές της παραμέτρου X

κ: Ο κωδικός αριθμός του $\kappa \in I_{\kappa}$ σταθμού (π.χ. κ=1 για τον σταθμό της Ε.Μ.Υ. στο Ελληνικό, κ=2 για τον σταθμό της ΔΕΗ στην Καβάλα κλπ.)

L: Το πλήθος των ετών κατά τα οποία ο σταθμός κ διαθέτει τιμές της παραμέτρου X. Αν YEARSTART(κ) είναι το έτος κατά το οποίο άρχισε η παρατήρηση (η μέτρηση ή η εκτίμηση ή η καταγραφή ή ο υπολογισμός) των τιμών της X και YEAREND(κ) το τελευταίο έτος, μέχρι το οποίο το αρχείο των τιμών της παραμέτρου είναι ενημερωμένο τότε:

$$L=L(\kappa)=YEAREND(\kappa)\ -\ YEARSTART(\kappa)+1 \qquad (7.3)$$

l: Ο αύξων αριθμός του έτους παρατήρησης. Το έτος YEAR και ο $l \in I_i$ συνδέονται μεταξύ τους με την:

$$l=l(\kappa)=YEAR+1-YEARSTART(\kappa) \qquad (7.4)$$

M: Το πλήθος των υποσυνόλων στα οποία διαμελίζεται το έτος. Οι βασικές τιμές του M είναι:
M=12 (μήνες) και M=36 (δεκαήμερα).

m: Ο αύξων αριθμός του μήνα ($m \in I_{12}$) ή του δεκαήμερου ($m \in I_{36}$) μέσα στο έτος.

α: Ο εκθέτης του M παίρνει τις τιμές α=μ (μήνας) για την περίπτωση M=12 και α=δ (δεκαήμερο) για την περίπτωση που M=36.

J: Ο αριθμός των ημερών του μήνα ή του δεκαήμερου. Οι τιμές που παίρνει το J, ανάλογα με την τιμή του a και του m που δόθηκαν και στο τέλος του δευτέρου κεφαλαίου είναι:

1. Αν $a=\mu$ τότε $j\in\{28,29,30,31\}$ και ειδικά:
   $J=31$ για $m\in\{1,3,5,7,8,10,12\}$
   $J=30$ για $m\in\{4,6,9,11\}$
   $J=29$ για $m=2$ και δίσεκτο έτος.
   $J=28$ για $m=2$ και μη δίσεκτο έτος.
2. Αν $a=\delta$ τότε $J\in\{8,9,10,11\}$ και ειδικά:
   $J=11$ για $m\in\{3,9,15,21,24,30,36\}$
   $J=9$ για $m=6$ και δίσεκτο έτος
   $J=8$ για $m=6$ και μη δίσεκτο έτος
   $J=10$ για $m\in\{I_{36}-\{3,6,9,15,21,24,30,36\}\}$

j: Ο αύξων αριθμός της ημέρας μέσα στο μήνα ή το δεκαήμερο.

N: Ο αριθμός των διατιθέμενων τιμών μέσα στην ημέρα.

n: Ο αύξων αριθμός της παρατήρησης μέσα στην ημέρα ($n\in I_N$). Ο αριθμός αυτός δείχνει ταυτόχρονα και την ώρα παρατήρησης, σε συνάρτηση με το N.

Στην εισαγωγή έγινε φανερό ότι ο αριθμός N, των ημερήσιων τιμών, εξαρτάται από την παράμετρο X και τον σταθμό κ.

Οι επικρατέστρες τιμές του N είναι:

N=1 Μία τιμή ανά ημέρα.

Στην κατηγορία αυτή ανήκουν παράμετροι όπως: Μέγιστη, ελάχιστη τιμή και εύρος οποιασδήποτε συνεχούς παραμέτρου, η ηλιοφάνεια, η ηλιακή ακτινοβολία, η εξάτμιση κλπ. Στην ίδια κατηγορία μπορεί να κατατάξει και τις απλές τιμές μιας συγκεκριμένης παρατήρησης κάθε ημέρας π.χ. η θερμοκρασία της 6UTC ή η πίεση της 12UTC κλπ., έστω και αν η παράμετρος μετράται περισσότερες από μια φορές την ημέρα. Ακόμα στην ίδια κατηγορία μπορεί να συμπεριληφθεί και κάθε ημερήσια τιμή που προκύπτει από την μαθηματική επεξεργασία των N>1 διαθέσιμων ημερήσιων τιμών, όπως για παράδειγμα η μέση ημερήσια τιμή (θερμοκρασία, πίεση, υγρασία κλπ.) η αθροιστική τιμή της ημέρας (ύψος βροχής, χιονιού κλπ.). Τέλος στην κατηγορία αυτή πρέπει να ενταχθούν και οι παρατηρήσεις ανώτερης ατμόσφαιρας στο Ηράκλειο και την Θεσσαλονίκη, αφού στους σταθμούς αυτούς πραγματοποιείται μόνο μια παρατήρηση ανά ημέρα.

N=2 Παράμετροι των οποίων οι τιμές λαμβάνονται δύο φορές την ημέρα.

Στην κατηγορία αυτή εμπίπτουν οι παρατηρήσεις ανώτερης ατμόσφαιρας στο Ελληνικό (0UTC και 12UTC), καθώς και ορισμένες άλλες παράμετροι των οποίων οι τιμές λαμβάνονται δύο φορές την ημέρα, όπως το ύψος βροχής των μετεωρολογικών σταθμών της ΕΜΥ (παρατηρήσεις στις 6UTC και 18UTC και τιμές τα ύψη βροχής από 18UTC της προηγούμενης μέχρι την 6UTC της ημέρας η μία και από 6UTC μέχρι 18UTC η άλλη), η εξάτμιση κλπ.

N=3

Αρκετοί κλιματολογικοί και γεωργικοί σταθμοί πραγματοτοποιούν τρείς παρατηρήσεις ανά ημέρα, στις 6UTC, 12UTC και 18UTC. Παράμετροι που ανήκουν στην κατηγορία αυτή είναι:

Η θερμοκρασία εδάφους, η θερμοκραία σε διάφορα βάθη στο έδαφος (κυρίως 5,10,15,20,25,50,75,100cm) κλπ.

N=8

Οι συνοπτικοί σταθμοί πραγματοποιούν τις παρατηρήσεις τους κατά τις λεγόμενες συνοπτικές ώρες παρατήρησης (0.3.6.9,12,15,18 και 21UTC). Από τις τιμές-ώρες αυτές οι 0,6,12 και 18UTC ονομάζονται κύριες συνοπτικές ώρες παρατήρησης.

N=24

Οι μετεωρολογικοί σταθμοί που βρίσκονται στα αεροδρόμια εκτελούν παρατηρήσεις ανά μία ώρα.

N=48

Οι σταθμοί που εδρεύουν στα μεγάλα αεροδρόμια μετρούν ή εκτιμούν τις μετεωρολογικές παραμέτρους (θερμοκρασία, πίεση, νέφωση, ορατότητα, άνεμο κλπ.) ανά ημίωρο δίνοντας σαρανταοκτώ τιμές ανά ημέρα.

N

Για τις παραμέτρους που οι τιμές τους καταγράφονται με αυτογραφικά όργανα μπορεί κανείς να διαλέξει την περίοδο δείγματος και να λάβει ακόμα και περισσότερες από σαρανταοκτώ τιμές. Η μικρότερη τιμή της περιόδου δείγματος που κυρίως λαμβάνεται είναι ΔT=5min δίνοντας N=288 τιμές/ημέρα. Εχοντας κανείς τις N=288 τιμές μπορεί να διαλέξει N=24 λαμβάνοντας μια τιμή για κάθε δώδεκα τιμές (ΔT=1h) N=48 (ΔT=30min), N=12 (DT=2h), N=96 (ΔT=15min), N=144 (ΔT=10min). Τιμές N>48 απαιτούνται (για την ώρα) μόνο για την βροχή, που για κατασκευαστι-κές εφαρμογές χρειάζονται συχνά και τιμές έντασης βροχής πενταλέπτου διάρκειας.

Για την παράσταση ενός υποσυνόλου, του συνόλο $X(K,L,M^{\alpha},J,N)$, το οποίο προκύπτει για ορισμένη τιμή ενός συνδυασμού μέχρι και τεσσάρων (διότι στην περίπτωση που και οι πέντε δείκτες πάρουν συγκεκριμένη τιμή, το προκύπτον υποσύνολο αποτελείται από ένα μόνο στοιχείο, την απλή τιμή $x(\kappa,l,m,j,n)$) από τους πέντε δείκτες που προσδιορίζουν το σύνολο $X(K,L,M^{\alpha},J,N)$, οι αντίστοιχοι δείκτες θα αντικαθίστανται από την συγκεκριμένη τιμή τους.

Για την περίπτωση που μελετάται το σύνολο των απλών τιμών ορισμένων ετών, ο δείκτης L αντικαθίσταται από το αρχικό έτος της περιόδου που μελετάται με δείκτη το πλήθος των ετών που μελετούνται. Οταν N>1 και ζητάτε να μελετηθεί το σύνολο των τιμών μιας συγκεκριμένης ώρας με κωδικό $n \in I_N$, τον N αντικαθιστούμε με το n με δείκτη τον N. Συμφωνείται εδώ η αντιστοιχεία:

n=(ώρα παρατήρησης+1) όταν N=24 (ώρες παρατήρησης σε UTC)
n(ώρα παρατήρησης+3)/3 όταν N=8 (συνοπτικές ώρες σε UTC)
n(ώρα παρατήρησης+6)/6 όταν N=4
(κύριες συνοπτικές ώρες σε UTC)
n(ώρα παρατήρησης)/6 όταν N=3
(για τις ώρες παρατήρησης 06,12 και 18UTC)

Παράδειγμα 7.1

Για την παράμετρο X που μετράται σε K σταθμούς οκτώ φορές την ημέρα:

α. Το σύνολο $X(3,5,12^{\mu},J,8)$ περιλαμβάνει τις $v=8*31=248$ τιμές του σταθμού με κωδικό $\kappa=3 \leq K$, κατά το πέμπτο έτος λειτουργίας του σταθμού ($l=5 \leq L$) κατά τον Δεκέμβριο ($M=12$, $\alpha=\mu$).

β. Το σύνολο $X(\kappa,61_{10},M^{\delta},J,3_8)$ περιέχει σαν στοιχεία τις τιμές της παραμέτρου X σ'όλους τους σταθμούς που "μετρούν" την X, κατά την δεκαετία 1961-1970 ($L=61_{10}$), όλα τα δεκαήμερα (M^{δ}),κατά την τρίτη κατά σειράν από την οκτώ ημερήσιες παρατηρήσεις (δηλαδή τις παρατηρήσεις της 6UTC).

γ. Το σύνολο $X(3,61_{30},30^{\delta},J,3_8)$ περιέχει τις τιμές του σταθμού με κωδικό $\kappa=3 \leq K$ κατά την τριακονταετία 1961-1990 (κανονική τιμή), το τριακοστό δεκαήμερο (21-30 Οκτωβρίου) κατά την 6UTC.

δ. Τα σύνολα $X(3,5,6^{\mu},6,8)$ και $X(3,5,6^{\delta},6,8)$ περιέχουν τις οκτώ ημερήσιες τιμές του σταθμού $\kappa=3 \leq K$ κατά το $l=5 \leq L$ έτος της λειτουργίας του, την έκτη Ιουνίου (6^{μ}) και την 26 Φεβρουαρίου (6^{δ}) αντίστοιχα, για την παράμετρο X.

Στον πίνακα 7.1 δίνονται για τις κυριώτερες μετεωρολογικές παραμέτρους X, ο προτεινόμενος συμβολισμός τους, οι σύνηθες τιμές των διαθέσιμων τιμών ανά ημέρα, η απαιτούμενη ακρίβεια των τιμών, ο αύξων αριθμός τους, με τον οποίο θα αναφέρονται στα επόμενα, και οι κωδικοί αριθμοί που δείχνουν τον τρόπο λήψης των τιμών ως εξής:

1: Οργανα άμεσης ανάγνωσης
2: Εκτίμηση από τον παρατηρητή
3: Ανάγνωση από ταινία αυτογραφικού οργάνου
4: Υπολογισμός ή εκτίμηση από άλλες παραμέτρους
5: Από τηλεμετρικά όργανα (RADAR ή Μετεωρολογικοί Δορυφόροι)
6: Υπολογισμός που βασίζεται σε τιμές αυτογραφικών οργάνων.

Επισημαίνεται ότι το γράμμα Ν στη στήλη "Συνήθης τιμές του Ν" δείχνει ότι μπορεί να έχει κανείς όσες τιμές θέλει αρκεί να διαλέξει κατάλληλη περίοδο δείγματος. Ακόμα την στιγμή αυτή εκτός από τον ΕΛΓΑ που διαθέτει ορισμένο πλήθος τιμών από ψηφιακά RADAR, δεν υπάρχει αρχείο τηλεμετρικών παρατηρήσεων.

7.1.2. Παράσταση των ημερήσιων τιμών των μετεωρολογικών παραμέτρων.

Οι ημερήσιες τιμές μιας συγκεκριμένης ώρας της ημέρας αποτελούν ταυτόχρονα απλές και ημερήσιες τιμές και η παράστασή τους σαν απλές τιμές συζητήθηκε στην προηγούμενη παράγραφο. Στην προηγούμενη παράγραφο δόθηκε και η παράσταση των παραμέτρων των οποίων διατίθεται μόνο μία τιμή ανά ημέρα (N=1 π.χ. ημερήσια μεγίστη ή ελαχίστη τιμή). Στην παράγραφο αυτή δίνεται ο συμβολισμός των ημερήσιων τιμών που προκύπτουν από την μαθηματική επεξεργασία ν ($1 \leq \nu \leq N$) ημερήσιων τιμών, της μορφής $x(\kappa,l,m^{\alpha},j,n)$ με $n \in I_N$ και κ,l,m,j σταθερά. Για την παράσταση του συνόλου των ημερήσιων τιμών που προκύπτουν, στην θέση του δείκτη Ν του συνόλου $X(K,L,M^{\alpha},J,N)$, τοποθετείται κατάλληλος δείκτης ή έκφραση που να παριστάνει όσο το δυνατό καλύτερα την τιμή που προκύπτει από την μαθηματική αυτή επεξεργασία. Ο ίδιος συμβολισμός χρησιμοποιείται και για τον δείκτη n των απλών τιμών-στοιχείων των αντίστοιχων συνόλων που προκύπτουν. Οι σημαντικότερες μορφές της μαθηματικής επεξεργασίας, από τις οποίες θα προκύψουν οι ημερήσιες τιμές της μετεωρολογικής παραμέτρου X και οι προτεινόμενοι δείκτες (σε παρένθεση) που θα αντικατθιστούν, στην αντίστοιχη περίπτωση, τον δείκτη Ν είναι:

1. Η μέση ημερήσια τιμή από όλες τις διαθέσιμες τιμές δηλαδή ν=Ν, με αντίστοιχο δείκτη τον Ε (N=E):

$$x(\kappa,l,m^{\alpha},j,E)=\Sigma\ x(\kappa,l,m^{\alpha},j,n)/N \quad n \in I_N \tag{7.5}$$

2. Η μέση τιμή των οκτώ συνοπτικών ωρών, δηλαδή ν=8 ($N=E_8$)

$$x(\kappa,l,m^{\alpha},j,E_8)=\Sigma\ x(\kappa,l,m^{\alpha},j,n)/8 \quad n \in I_8 \tag{7.6}$$

3. Η μέση τιμή που προκύπτει από το ημιάθροισμα της μέγιστης και της ελάχιστης τιμής της ημέρας, που προκύπτουν από όργανα άμεσες ανάγνωσης (κυρίως για την θερμοκρασία και την σχετική υγρασία) ($N=E_{2/2}$):

$$x(\kappa,l,m^{\alpha},j,E_{2/2})=0.5*[x(\kappa,l,m^{\alpha},j,MAX)+x(\kappa,l,m^{\alpha},j,MIN)] \tag{7.7}$$

4. Η μέση ημερήσια τιμή από τις τρεις τιμές που μετρήθηκαν κατά τις ώρες 6,12 και 18UTC ($N=E_3$):

$$x(\kappa,l,m^{\alpha},j,E_3)=\Sigma\ x(\kappa,l,m^{\alpha},j,n_3)/3 \tag{7.8}$$

5. Η μέση ημερήσια τιμή, από τις τρεις παραπάνω ώρες, στην οποία όμως η παρατήρηση της 18UTC λαμβάνεται δύο φορές ($N=E_{B3}$):

$$x(\kappa,l,m^{\alpha},j,E_{B3})=\Sigma\ \alpha_n*x(\kappa,l,m^{\alpha},j,n_3)/4 \tag{7.9}$$

όπου $\alpha_1=\alpha_2=1$ και $\alpha_3=2$

6. Το ημερήσιο εύρος των τιμών που προκύπτει από όργανα άμεσης ανάγνωσης (N=ΕΥΡ):

$$x(\kappa,l,m^{\alpha},j,ΕΥΡ)=x(\kappa,l,m^{\alpha},j,MAX)-x(\kappa,l,m^{\alpha},j,MIN) \tag{7.10}$$

7. Η ελάχιστη μέγιστη τιμή και το εύρος που προκύπτουν από τις N τιμές της ταινίας του αυτογραφικού οργάνου που καταγράφει τις τιμές της X (N=max, N=min και N=ευρ.αντίστοιχα):

$$x(\kappa,l,m^{\alpha},j,max)=max\{x(\kappa,l,m^{\alpha},j,n)\ /\ n\in I^{N}\} \quad (7.11)$$

$$x(\kappa,l,m^{\alpha},j,min)=min\{x(\kappa,l,m^{\alpha},j,n)\ /\ n\in I_{N}\} \quad (7.12)$$

$$x(\kappa,l,m^{\alpha},j,\varepsilon\upsilon\rho)=x(\kappa,l,m^{\alpha},j,max)-x(\kappa,l,m,j,min) \quad (7.13)$$

Εκτός της θερμοκρασίας, για την οποία η ελέχιστη και η μέγιστη τιμή προκύπτουν από όργανα άμεσης ανάγνωσης, οι παραπάνω τιμές των υπόλοιπων μετεωρολογικών παραμέτρων πρέπει να υπολογίζονται μετά από διόρθωση των τιμών των αυτογραφικών οργάνων, σύμφωνα με τις οδηγίες διόρθωσης που περιγράφονται στην παρ.4.1.1 του τεύχους 10/2 του ΥΔΡΟΣΚΟΠΙΟ.

8. Το ημερήσιο άρθοισμα ν (1<ν≤N) ημερήσιων τιμών ($N=\Sigma_{\nu/N}$), κυρίως για την βροχή και το χιόνι ή τις βαθμοημέρες πάνω ($N=T_{\alpha T o}$) ή κάτω ($N=T_{\kappa T o}$) από μια βασική θερμοκρασία To, που ορίζονται από τις:

$$x(\kappa,l,m^{\alpha},j,\Sigma_{\nu/N})=\Sigma\ x(\kappa,l,m^{\alpha},j,n)\quad n\in I_{\nu} \quad (7.14)$$

για το ημερήσιο ύψος της βροχής και

$$x(\kappa,l,m^{\alpha},j,T_{\alpha T o})=\Sigma\ y(\kappa,l,m^{\alpha},j,n)\quad n\in I_{\nu},\ w\in\{y,\varphi\} \quad (7.15)$$

όπου:

$$y(\kappa,l,m^{\alpha},j,n)=\begin{cases} x(\kappa,l,m^{\alpha},j,n)-To & \text{όταν } \chi(\kappa,l,m^{\alpha},j,n)\geq To \\ 0 & \text{όταν } \chi(\kappa,l,m^{\alpha},j,n)<To \end{cases} \quad (7.16)$$

και:

$$x(\kappa,l,m^{\alpha},j,T_{\kappa T o})=\Sigma\ \varphi(\kappa,l,m^{\alpha},j,n)\quad n\in I_{\nu},\ w\in\{y,\varphi\} \quad (7.17)$$

όπου:

$$\varphi(\kappa,l,m^{\alpha},j,n)=\begin{cases} To-x(\kappa,l,m^{\alpha},j,n) & \text{όταν } \chi(\kappa,l,m^{\alpha},j,n)\leq To \\ 0 & \text{όταν } \chi(\kappa,l,m^{\alpha},j,n)>To \end{cases} \quad (7.18)$$

9. Οι βαθμοημέρες για την ελάχιστη και την μέγιστη θερμοκρασία ($N=MM_{\alpha T o}$ ή $N=MM_{\beta T o}$ για βαθμοημέρες πάνω ή κάτω από την βασική θερμοκρασία To αντίστοιχα) που ορίζονται από τις:

$$x(\kappa,l,m^{\alpha},j,MM_{\alpha T o})=\begin{cases} x(\kappa,l,m^{\alpha},j,MM)-To & \text{όταν } \chi(\kappa,l,m^{\alpha},j,MM)\geq To \\ 0 & \text{όταν } \chi(\kappa,l,m^{\alpha},j,MM)<To \end{cases} \quad (7.19)$$

και:

$$\chi(\kappa,l,m^{\alpha},j,MM_{\beta T o})=\begin{cases} To-x(\kappa,l,m^{\alpha},j,MM) & \text{όταν } \chi(\kappa,l,m^{\alpha},j,MM)\leq To \\ 0 & \text{όταν } \chi(\kappa,l,m^{\alpha},j,MM)>To \end{cases} \quad (7.20)$$

όπου MM∈{MIN,MAX}. Οι βασικές τιμές για τις οποίες πρέπει να υπολογίζονται οι ημερήσιες τιμές των παραπάνω παραμέτρων είναι:

To∈{-4,-2,0,2,4,5,10,15,18,20,25,30,35} (τιμές σε °C).

και πρέπει μέσω του προγράμματος που θα συνταχθεί να δίνεται η δυνατότητα επιλογής της To από τον χρήστη.

10. Εμφάνιση ή όχι φαινομένου ή γεγονότος κατά την διάρκεια της ημέρας (N=Γ). Για τις μετεωρολογικές παραμέτρους των οποίων οι απλές τιμές είναι 1 ή 0 ανάλογα με το αν κατά την ώρα n της παρατήρησης συνέβαινε ή όχι κάποιο γεγονός, η ημερήσια τιμή ορίζεται επίσης ίση με 1 ή 0 ανάλογα με τον αν κατά την διάρκεια της ημέρας το γεγονός ή το φαινόμενο συνέβη τουλάχιστον μια φορά και δίνεται από την συνάρτηση:

$$\chi(\kappa,l,m^{\alpha},j,\Gamma)=\begin{cases} 0 \text{ όταν } \chi(\kappa,l,m^{\alpha},j,n_{\Gamma})=0 & \text{για όλα τα } n\in I_N \\ 1 \text{ όταν } \chi(\kappa,l,m^{\alpha},j,n_{\Gamma})=1 & \text{για ένα τουλάχιστο } n \end{cases} \tag{7.21}$$

όπου:

$$\chi(\kappa,l,m^{\alpha},j,n_{\Gamma})=\begin{cases} 0 \text{ όταν το γεγονός δεν συνέβει στην n μέτρηση} \\ 1 \text{ όταν το γεγονός Γ συνέβει στην n μέτρηση.} \end{cases}$$

Παράμετροι της κατηγορίας αυτής είναι οι παράμετροι που στον πίνακα 7.1 έχουν α/α=12, πρέπει όμως να είναι δυνατό στο χρήστη να ορίσει και για άλλες παραμέτρους, ανάλογα με τις ανάγκες του, ορίζοντας μέσω menu τα γεγονότα που τον ενδιαφέρουν π.χ.

$$\chi(\kappa,l,m^{\alpha},j,\Gamma)=\begin{cases} 0 & \text{όταν } P_{\bullet\alpha}(\kappa,i,m^{\alpha},j,n)\leq 1000\text{hPa} \\ 1 & \text{όταν } P_{\bullet\alpha}(\kappa,i,m^{\alpha},j,n)>1000\text{hPa} \end{cases}$$

ή:

$$\chi(\kappa,l,m^{\alpha},j,\Gamma)=\begin{cases} 0 & \text{όταν } T_{max}>30^{\circ}\text{C και έβρεξε.} \\ 1 & \text{αλλοιώς} \end{cases}$$

Αν με A παραστήσουμε τον δείκτη που αντικαθιστά τον N τότε στα στοιχεία $x(\kappa,l,m^{\alpha},j,A)$ του συνόλου $X(K,L,M^{\alpha},J,A)$ των ημερήσιων τιμών της παραμέτρου X μόνο οι τέσσερις δείκτες (οι κ,l,m^{α},j) είναι ελεύθεροι. Για τα υποσύνολα του $X(K,L,M^{\alpha},J,A)$ που προκύπτουν για ορισμένη τιμή ενός συνδυασμού μέχρι και τριών από τους κ,l,m,j (αν και οι τέσσερις λάβουν ορισμένη τιμή το υποσύνολο περιέχει σαν μοναδικό στοιχείο την ημερήσια τιμή), χρησιμοποιείται ται η παράσταση που συμφωνήθηκε στην παράγραφο 7.1.1. Στον πίνακα 7.2 δίνονται για τις αντίστοιχες παραμέτρους του πίνακα 7.1 οι ημερήσιες τιμές που υπολογίζονται κατά περίπτωση, και ο αντίστοιχος τύπος υπολογισμού που χρησιμοποιείται. Οι μονάδες μέτρησης και η ζητούμενη ακρίβεια στις ημερήσιες τιμές είναι η ίδια μ'αυτή την απλών τιμών.

<u>Παράδειγμα 7.2.</u>

Για την παράσταση του συνόλου που περιέχει τις μέσες ημερήσιες τιμές, του σταθμού με κωδικό $\kappa=15\leq K$ του έτους $l=5\leq L$ κατά τον Ιούλιο ($M^{\alpha}=7^{\mu}$), ο συμβολισμός:

$X(15,5,7^{\mu},J,E)$

δείχνει ότι ο υπολογισμός γίνεται από το σύνολο των διαθέσιμων τιμών και ο συμβολισμός:

$X(15,5,7^{\mu},J,E_{2/2})$

δείχνει ότι οι τιμές του συνόλου υπολογίστηκαν με τον τύπο (7.7).

Τέλος το:

$BP(10,71_{20},M^{\mu},J,\Sigma_{2/2})$

είναι το σύνολο των τιμών των ημερήσιων αθροισμάτων της βροχής (X=BP) που προκύπτει από τις τιμές του ύψους βροχής την 6UTC και την 18UTC ($N=\Sigma_{2/2}$), στο σταθμό με κωδικό $\kappa=10\leq K$, κατά την περίοδο 1971-1990 ($L=71_{20}$) όλους τους μήνες ($M^{\alpha}=M^{\mu}$) και όλες τις ημέρες.

7.1.3. Παράσταση Στατιστικών τιμών των μετεωρολογικών παραμέτρων.

7.1.3.1.α. Μηνιαίες-δεκαήμερες τιμές κάθε έτους ξεχωριστά.

Για την παράσταση του συνόλου των μηνιαίων ή δεκαήμερων τιμών που προκύπτουν από την μαθηματική επεξεργασία των j διαθέσιμων ημερήσιων τιμών του μήνα ή του δεκαήμερου που μελετάται, ο δείκτης J αντικαθίσταται από κατάλληλη δείκτη ή παράσταση, που εκφράζει όσο το δυνατό καλύτερα την προκύπτουσα από την κατά περίπτωση μαθηματική επεξεργασία (όπως έγινε στην περίπτωση του υπολογισμού την ημερήσιων τιμών, από τις απλές τιμές της ημέρας).

Στην περίπτωση αυτή τον δείκτη N αντικαθιστούμε με δείκτη ή παράσταση που να δείχνει την προέλευση των ημερήσιων τιμών, όπως συζητήθηκε στην προηγούμενη παράγραφο. Για την παράσταση μιας οποιασδήποτε ημερήσιας τιμής χρησιμοποιείται ο δείκτης A π ου μπορεί να είναι ένα οποιοδήποτε, (κατάλληλο για την μελετούμενη παράμετρο), στοιχείο του συνόλου:

$\{E,E_{2/2},E_3,E_8,E_{83},\Sigma,\Sigma_{\nu/N},\Gamma,\alpha To,\kappa To,N,\eta_N,\ldots\}$

Οι μηνιαίες τιμές κάθε έτους ξεχωριστά που πρέπει να καταχωρηθούν σε ξεχωριστό αρχείο, μαζί με το πλήθος ν των διαθέσιμων ημερήσιων για τον υπολογισμό τους τιμών δίνονται στον πίνακα 7.3.

Οι βασικότερες μηνιαίες ή δεκαήμερες τιμές που υπολογίζονται για κάθε έτους ξεχωριστά είναι οι παρακάτω:

1. **Μέση μηνιαία τιμή (j=E)**
2. **Αθροιστική τιμή του μήνα (j=Σ)**
3. **Ελάχιστη και μέγιστη τιμή (j=MIN και j=MAX αντίστοιχα)**
4. **Εύρος τιμών (συνεχείς παράμετροι) (j=EYP)**
5. **Συχνότητα εμφάνισης γεγονότος-φαινομένου (j=ΓF)**

<u>Παράδειγμα 7.3.</u>

α. Το σύνολο των μέσων μηνιαίων τιμών της πρώτης από τις οκτώ ημερήσιες παρατηρήσεις κάθε μέρας του μήνα συμβολίζεται με:

$X(K,L,M^{\mu},E,1_8)$

β. Ο συμβολισμός:

$X(K,L,3^{\delta},\Sigma,\Sigma_{2/2})$

παριστάνει το σύνολο των τιμών των αρθοιστικών τιμών του τρίτου δεκαήμερου (21-31 Ιανουαρίου), που προκύπτουν από τις αθροιστικές ημερήσιες τιμές, που υπολογίστηκαν από τις δύο ημερήσιες τιμές (6UTC και 18UTC) για όλους τους σταθμούς που διαθέτουν τι μές της παραμέτρου X και για όλα τα χρόνια.

γ. Το σύνολο των μέσων μηνιαίων τιμών της μέγιστης θερμοκρασίας (X=T,N=MAX) του σταθμού με κωδικό κ=1ЄΚ, είναι το:

$X(1,K,M\mu,E,MAX)$

δ. Με:

$X(1,51_{30},1^{\delta},\Sigma,T_{o5^{\circ}})$

παριστάνεται το σύνολο των τριάντα αθροιστικών τιμών, του πρώτου δεκαήμερου, κατά τα έτη 1951-1980, στο σταθμό με κωδικό $\kappa=1\leq K$, των ημερήσιων βαθμοημερών πάνω από την βασική θερμοκρασία To=5°C.

ε. Το:

$X(1,61_{30},1^{\mu},\Gamma F,E_{2/2})$

είναι το σύνολο των συχνοτήτων εμφάνισης γεγονότος (j=ΓF) που συνδέεται με τις μέσες ημερήσιες τιμές που προκύπτουν από την (7.7) για τον σταθμό με κωδικό κ=1, την τριακονταετία 1961-1990 ($L=61_{30}$) κατά τον Ιανουάριο ($M^{\alpha}=1^{\mu}$). Ενα τέτοιο γεγονός θα μπορούσε να ήταν το:

$\Gamma=\{5\leq x(1,l,1^{\mu},j,E_{2/2})\leq 15 \text{ με } l\in\{61,62,...90\},\ j\in I_{31}\}$

7.1.3.1.β. Μηνιαίες-δεκαήμερες τιμές πολλών ετών.

Οι μηνιαία ή δεκαήμερη τιμή ενός συγκεκριμένου μήνα $m^{\mu}\in I_{12}$ ή δεκαήμερου $m^{\delta}\in I_{36}$ που αναφέρεται σε $1<\lambda\leq L$ έτη, μπορεί να προκύψει με τους εξής δύο τρόπους:

α. Από την μαθηματική επεξεργασία των μηνιαίων (δεκαήμερων) τιμών που υπολογίστηκαν για κάθε έτους ξεχωριστά (λ τιμές). Οι σπουδαιότερες τιμές που πρέπει να υπολογιστούν στην περίπτωση αυτή είναι η μέση τιμή (των μέσων τιμών ή των αθροισμάτων, ή των περιπτώσεων της επικράτησης διαφόρων φαινομένων-γεγονότων), η τυπική απόκλιση των παραπάνω (εντός παρένθεσης) τιμών, καθώς και η ελάχιστη και η μέγιστη τιμή τους.

β. Από την μαθηματική επεξεργασία του συνόλου των ημερήσιων τιμών του μήνα ή του δεκαήμερου (J*λ τιμές). Οι βασικότερες τιμές που πρέπει να προκύψουν από την επεξεργασία των ημερήσιων τιμών, για λ=L, λ=10 και L=30 είναι:

β.1. Η μέση τιμή όλων των ημερήσιων τιμών του μήνα ή του δεκαήμερου.

β.2. Η διακύμανση και η τυπική τους απόκλιση.

β.3. Η μέγιστη και ελάχιστη τιμή.

β.4. Η συνχότητα f(x) εμφάνισης κάθε τιμής μεταξύ της ελάχιστης και μέγιστης τιμής.

β.5. Η εύρεση των τιμών $x_{m,p}$, της παραμέτρου για τις οποίες η αθροιστική συχνότητα

$F(x_{m,p})=F(x\leq x_{m,p})=p/100$.

Οι πλέον ενδιαφέρουσες τιμές του p είναι τα στοιχεία του συνόλου:

$p \in \{0.5, 1, 2.5, 5, 10, 25, 50, 75, 90, 95, 97.5, 99, 99.5\}$

β.6. Για συνεχείς μόνο παραμέτρους η εύρεση της συχνότητας $f(A_i)$ των τ συνόλων A_i, $i \in I_\tau$, με $\tau=[5*\log(j*\lambda)]$ (σύμφωνα με τον τύπο (5.5), τιμές που θα χρειαστούν για τον έλεγχο της προσαρμογής των τιμών της X σε κατάλληλη κατανομή συχνοτήτων.

β.7. Η Κυρτώτητα και η ασσυμετρία (συνεχείς παράμετροι).

Κατά την περίπτωση που ο υπολογισμός των τιμών γίνεται με βάση τις μηνιαίες ή δεκαήμερες τιμές κάθε έτους ο δείκτης J αντικαθίσταται από δείκτη ή παράσταση που δείχνει όσο καλύτερα γίνεται την μαθηματική επεξεργασία από την οποία προήλθαν οι λ μηνιαίες ή δεκαήμερες τιμές, όπως περιγράφηκε στην παράγραφο 7.1.3.1.α. και ο δείκτης L των ετών με δείκτη ή παράσταση που να δείχνει όσο καλύτερα γίνεται το είδος της μαθηματικής επεξεργασίας των τιμών και σε παρένθεση τέσσερα ψηφία, τα δύο πρώτα το έτους αρχής (τα δύο πρώτα ψηφία) και τα δύο τελευταία το τελικό έτος της περιόδου των λ ετών (δύο τελευταία ψηφία). Οταν λ=L (δηλαδή οι υπολογισμοί γίνονται με τις τιμές όλων των ετών) αντί των τεσσάρων ψηφίων (αρχής και τέλους της περιόδου) τίθεται το L. Στην περίπτωση που οι υπολογισμοί βασίζονται στις ημερήσιες τιμές, ο δείκτης L των ετών αντικαθίσταται και πάλι από δείκτη ή παράσταση που δείχνει όσο καλύτερα γίνεται την μαθηματική επεξεργασία από την οποία προήλθαν οι τιμές, και σε παρένθεση τα τέσσερα ψηφία, όπως παραπάνω, και ο δείκτης J αντικαθίσταται από το γράμμα D που δείχνει ότι οι υπολογισμοί βασίζονται στις διαθέσιμες ημερήσιες τιμές.

Παράδειγμα 7.4.

α. Το σύνολο:

$X(3, E(6171), M^{\mu}, E, E_{2/2})$

περιέχει σαν στοιχεία τις δώδεκα μέσες μηνιαίες τιμές ($M^{\alpha}=M^{\mu}$) της περιόδου 1961-1971 (L=E(6171)), που υπολογίστηκαν με βάση τις μέσες μηνιαίες τιμές (J=E), ο υπολογισμός των οποίων στηρίχτηκε στις μέσες ημερήσιες τιμές (N=E2/2) που υπολογίστηκαν με βάση τον τύπο (7.7), στον σταθμό με κωδικό $\kappa=3\leq K$.

β. Το σύνολο:

$T(3, SD(6171), M^{s}, D, MAX)$

περιέχει τις τριάνταέξη τυπικές αποκλίσεις ($M^{\alpha}=M^{s}$, L=SD(6171)) των τιμών της μέγιστης θερμοκρασίας (X=T, N=MAX), που υπολογίστηκαν με βάση τις ημερήσιες τιμές (J=D), της περιόδου 1961-1971, στον σταθμό με κωδικό $\kappa=3\leq K$.

γ. Με:

$BP(K, E(L), 1^{\mu}, \Sigma, \Sigma_{2/2})$

παριστάνεται το σύνολο των μέσων τιμών του Ιανουαρίου ($M^{\alpha}=1^{\mu}$), όλων των ετών (L=E(L)), για την βροχή (X=BP), που υπολογίστηκαν με βάση τα μηνιαία αθροίσματα (J=Σ) που προέκυψαν από τα ημερήσια αθροίσματα βάση του τύπου (7.14), όλων των Κ σταθμών.

δ. Η:

$T_{850}(1,Var(L),1^{δ},D,1_2)$

είναι η διακύμανση της θερμοκρασίας στα 850hPa ($X=T_{850}$) κατά την πρώτη από τις δύο διαθέσιμες ημερήσιες τιμές ($N=1_2$ δηλαδή παρατήρηση της OUTC) που υπολογίστηκε με βάση τις ημερήσιες τιμές (J=D) όλων των ετών (L=Var(L)) κατά το πρώτο δεκαήμερο του έτους ($M^{α}=1^{δ}$).

ε. Το σύνολο:

$X(K,E(L),12^{μ},T_{α10}(D),E_8)$

περιέχει τις Κ μέσες τιμές των "βαθμοημερών" της μέσης ημερήσιας τιμής της Χ, που προκύπτει από την (7.6), πάνω από την βασική τιμή To=10°C ($J=T_{α10}(D)$,$N=E_8$) του Δεκεμβρίου ($M^{α}=12^{μ}$) που υπολογίζονται με βάση όλες τις ημερήσιες τιμές όλων των ετών (L=E(L)) κάθε σταθμού.

7.1.3.2. Ετήσιες τιμές - Τιμές λ ετών (1<λ≤L).

7.1.3.2.α. Ετήσια τιμή συγκεκριμένου έτους.

Η ετήσια τιμή του έτους με κωδικό αριθμό $1 \in I_1$, μιας παραμέτρου Χ που μετράται σ'ένα σταθμό κ≤Κ, μπορεί να προκύψει από την μαθηματική επεξεργασία:

α. Των δώδεκα μηνιαίων ή τριάντα έξη δεκαήμερων τιμών και

β. Από τις 365 (ή 366) ημερήσιες τιμές.

Οι βασικότερες ετήσιες τιμές που υπολογίζονται, είναι συνάρτιση της παραμέτρου Χ που εξετάζεται κατά περίπτωση, είναι:

1. Μέση τιμή.
2. Αθροιστική τιμή.
3. Μέγιστη, ελάχιστη τιμή και εύρος.
4. Συχνότητα εμφάνισης γεγονότων-φαινομένων.
5. Ημερομηνία πρώτης και τελευταίας εμφάνισης γεγονότος-φαινομένου.

Για την παράσταση των συνόλων των ετήσιων τιμών, για τους δείκτες j και Ν εφαρμόζονται τα όσα αναφέρθηκαν στις προηγούμενες παραγράφους. Ο δείκτης $M^{α}$ των μηνών αντικαθίσταται στην περίπτωση αυτή με την παράσταση ή τον δείκτη που παριστάνει με τον καλύτερο τρόπο την μαθηματική επεξεργασία των τιμών και σε παρένθεση οι αριθμοί 12, 36, 365 ανάλογα με το αν η μαθηματική επεξεργασία γίνεται στις μηνιαίες, τις δεκαήμερες ή τις ημερήσιες τιμές αντίστοιχα.

<u>Παράδειγμα 7.5.</u>

α. Το:

$BP(1,1,\Sigma(12),\Sigma,\Sigma_{2/2})$

είναι το ετήσιο άθροισμα των δώδεκα μηνιαίων αρθοισμάτων, των ημερήσιων αθροισμάτων της βροχής κατά το πρώτο έτος λειτουργίας (L=1) του σταθμού με κωδικό κ=1.

β. Το:

DP(K,L,E(365),D,E_8)

είναι το σύνολο (με K*L στοιχεία) των μέσων ετήσιων τιμών του σημείου δρόσου (X=DP, πίνακας 7.1) που υπολογίζεται με βάση τις μέσες ημερήσιες (M^a=E(365), N=E_8), όταν η ημερήσια τιμή υπολογίζεται με βάση τις τιμές των οκτώ συνοπτικών ωρών παρατήρησης.

7.1.3.2.β. Ετήσιες τιμές λ ετών ($1<\lambda\leq L$).

Για την παράσταση των ετήσιων τιμών λ ετών ($1<\lambda\leq L$), ο δείκτης M^a αντικαθίσταται από δείκτη ή παράσταση που να εκφράζει όσο το δυνατό καλύτερα την μαθηματική επεξεργασία των στοιχείων του συνόλου που χρησιμοποιείται (όπως και στην περίπτωση των ετήσιων τιμών συγκεκριμένου έτους) και ο δείκτης L των ετών αντικαθίσταται με δείκτη όπως και στην περίπτωση της παράστασης των μηνιαίων ή δεκαήμερων τιμών λ ετών, που αναπτύχθηκε στην παράγραφο 7.1.3.1.β. Οι στατιστικές-κλιματολογικές παράμετροι που πρέπει να υπολογιστούν και να καταχωρηθούν στην περίπτωση αυτή είναι συνάρτηση της μετεωρολογικής παραμέτρου X που μελετάται και οι βασικότερες απ'αυτές είναι:

1. Η μέση τιμή, η ελάχιστη και η μέγιστη τιμή, το εύρος, η διακύμανση και η τυπική απόκλιση των λ ετήσιων τιμών.

2. Χαρακτηριστικές τιμές της X, ως εξής:

Αν D_x=[MIN,MAX] είναι το διάστημα εντός του οποίου ανήκουν οι ημερήσιες τιμές της X, ACCU η ακρίβεια τους και EYP=MAX-MIN τότε το πλήθος των δυνατών τιμών της X στο χρονικό διάστημα των λ ετών είναι ίσο με:

$$v=1+EYP/ACCU \tag{7.22}$$

Αν παραστήσουμε με ω_i, $i\in I_v$ τις τιμές αυτές, βρίσκουμε τις συχνότητες $f(\omega_i)$, τιμές που μπορούν να χρησιμοποιηθούν για τον έλεγχο της προσαρμογής της παραμέτρου σε θεωρητική συνάρτηση κατανομής (παράγραφος 6.1.2) καθώς και τις αθροιστικές συχνότητες $F(\omega_i)$.

Ακόμη βρίσκουμε τις τιμές $\omega_{0.005}$, $\omega_{0.995}$ και ω_p με:

$p=0.01*i$, $i\in I_{100}$,

για τις οποίες:

$$F(x\leq w_p)=p \tag{7.23}$$

Οι τιμές του p οι οποίες ενδιαφέρον όταν λ=10 ή λ=30 είναι στοιχεία του συνόλου:

$p\in\{0.005,0.025,0.05,0.1,0.2,0.25,0.5,0.75,0.8,0.9,0.95$
$0.95, 0.975,0.995\}$

Για p=0.5 η τιμή ω_p είναι η διάμεσος των ημερήσιων τιμών της X κατά την εξεταζόμενη περίοδο.

7.2. Τύποι υπολογισμού.

Οι χρησιμοποιούμενοι τύποι για τον υπολογισμό των σημαντικότερων στατιστικών παραμέτρων από Μ διαθέσιμες τιμές δόθηκαν στα προηγούμενα κεφάλαια, ενώ στην παράγραφο 7.1.2. δόθηκαν οι τύποι υπολογισμού των βασικών ημερήσιων τιμών από το σύνολο των διαθέσιμων ημερήσιων τιμών ή μόνο από ένα μέρος αυτών.

7.2.1. Μηνιαίες-δεκαήμερες τιμές συγκεκριμένου μήνα (m^{μ}) ή δεκαήμερου (m^{δ}) για συγκεκριμένο έτος l.

1. <u>Μέση τιμή κάθε ώρας παρατήρησης.</u>

Αυτή υπολογίζεται από την:

$$x(\kappa,l,m^{\alpha},E,n)= \Sigma\ (x(\kappa,l,m^{\alpha},j,n)/J \quad j\in I_J \tag{7.24}$$

όπου J ο αριθμός των ημερών του μήνα m^{μ} ή δεκαήμερου m^{α}.

Η παραπάνω τιμή υπολογίζεται για όλες τις παραμέτρους του πίνακα 7.1 των οποίων οι διαθέσιμες τιμές N είναι περισσότερες από 1 και στον πίνακα 7.2 δεν υπολογίζεται το Σ ή οι βαθμοημέρες. Οταν ο δείκτης η αντιστοιχεί σε κάποιο συνοπτική ώρα, τότε η υπολογιζόμενη μέση τιμή πρέπει να αρχειοθετείται.

2. <u>Αθροιστική τιμή κάθε ώρας.</u>

$$x(\kappa,l,m^{\alpha},\Sigma,n)= \Sigma\ x(\kappa,l,m^{\alpha},j,n) \quad j\in I_J \tag{7.25}$$

Η αθροιστική αυτή τιμή υπολογίζεται κυρίως για την βροχή (με N=2) καθώς και για τις βαθμοημέρες ως προς κάποια βασική θερμοκρασία και για τις οκτώ συνοπτικές ώρες παρατήρησης.

3. <u>Η μέση τιμή που προκύπτει από τις ημερήσιες τιμές:</u>

$$x(\kappa,l,m^{\alpha},E,A)= \Sigma\ x(\kappa,l,m^{\alpha},j,A)/J \quad j\in I_J \tag{7.26}$$

όπου $A\neq\Sigma$ η παράμετρος-δείκτης, που δείχνει τον τρόπο υπολογισμού των ημερήσιων τιμών (κυρίως Α∈{Ε,Ε8,Ε3,Ε83,MIN,MAX,ΕΥΡ}).

4. <u>Αθροιστική τιμή που προκύπτει από ημερήσιες τιμές.</u>

$$x(\kappa,l,m^{\alpha},\Sigma,A)= \Sigma\ x(\kappa,l,m^{\alpha},j,A) \quad j\in I_J \tag{7.27}$$

όπου το Α δεν είναι στοιχείο του συνόλου {Ε,MIN,MAX,ΕΥΡ}, η παράμετρος που δείχνει την προέλευση των ημερήσιων τιμών και κυρίως Α∈{$\Sigma_{N/N}$, $\Sigma_{2/2}$, αΤο, κΤο, ΜΜΜαΤο, ΜΜΜκΤο}.

5. <u>Συχνότητα εμφάνισης ορισμένου γεγονότος-φαινομένου.</u>

$$x(\kappa,l,m^{\alpha},\Gamma,\Gamma)= \Sigma\ x(\kappa,l,m^{\alpha},j,\Gamma) \tag{7.28}$$

όπου η τιμή x(κ,l,mα,j,Γ) δίνεται από την (7.21) Τέτοιες συχνότητες υπολογίζονται (κυρίως) για τις τιμές των παραμέτρων με α/α 12 του πίνακα 7.1.

6. <u>Ελάχιστη, μέγιστη τιμή, εύρος τιμών ορισμένης ώρας.</u>

$$x(\kappa,l,m^{\alpha},MMM,n)=mmm\{x(\kappa,l,m^{\alpha},j,n) \text{ με } j\in J\} \tag{7.29}$$

όπου j το πλήθος των ημερών του μήνα και ΜΜΜ∈{MIN,MAX} ανάλογα με το αν mmm=max ή mmm=min.

$$x(\kappa,l,m^{\alpha},EYP,n)=x(\kappa,l,m^{\alpha},MAX,n)-x(\kappa,l,m^{\alpha},MIN,n) \tag{7.30}$$

Οι τιμές που δίνονται από τις (7.28) και (7.29) υπολογίζονται κυρίως για την θερμοκρασία του αέρα (T_{α}), την θερμοκρασία του εδάφους (T_{ε} ή T_{aa}), την θερμοκρασία του υγρού θερμομέτρου T_w, το σημείο δρόσου DP, την πίεση και τις παραμέτρους υγρασίας (3.1 έως

3.4 του πίνακα 7.1).

7. Ελάχιστη, μέγιστη τιμή και εύρος από ημερήσιες τιμές.

$x(\kappa,1,m^{\alpha},MMM,A)=mmm\{x(\kappa,1,m^{\alpha},j,A)$ με $j\in J\}$ (7.31)

όπου οι MMM,mmm και J όπως περιγράφονται παραπάνω και A παράσταση-δείκτης που δείχνει την προέλευση των ημερήσιων τιμών (κυρίως MIN,MAX,E). Το αντίστοιχο εύρος των τιμών δίνεται από την:

$x(\kappa,1,m^{\alpha},EYP,A)=x(\kappa,1,m^{\alpha},MAX,A)-x(\kappa,1,m^{\alpha},MIN,A)$ (7.32)

7.2.2. Μηνιαίες-δεκαήμερες τιμές συγκεκριμένου μήνα (m^{μ}) ή δεκαήμερου (m^{δ}), λ ετών ($1<\lambda\leq L$).

Αν λ_1 και λ_2 είναι το πρώτο και τελευταίο έτος της περιόδου των λ ετών βρίσκονται:

1. Μέση τιμή

$$x(\kappa,E(\lambda_1\lambda_2),m^{\alpha},B,A)=\sum_{1=s}^{s+\lambda} x(\kappa,1,m^{\alpha},B,A)/\lambda \qquad (7.32)$$

όπου $s=\lambda_1+1-YEARSTART$ (τύπος 7.4) και A,B δείκτες που αντικαθιστούν τους N και J αντίστοιχα και δείχνουν την προέλευση των μηνιαίων ή δεκαήμερων τιμών που χρησιμοποιούνται.

Ο A μπορεί να πάρει κατάλληλη για την X τιμή από το σύνολο: $\{n\in I_N,\ E(E,E_8,E_3,$κλπ.$),\ \Sigma(\Sigma_{N/N},\Sigma_{2/2}$ κλπ.$),\ \Gamma,\ T_{\alpha\tau o},$ κλπ.$\}$ και το B είναι στοιχείο του συνόλου $\{E,\Sigma,\Gamma\}$ ανάλογα και πάλι από την εξεταζόμενη παράμετρο X.

2. Διακύμανση και τυπική απόκλιση.

Η διακύμανση (Var) και η τυπική απόκλιση (SD) υπολογίζονται τόσο με βάση τις μέσες μηνιαίες τιμές των Λ=λ ετών (λ τιμές) όσο και με βάση τις ημερήσιες τιμές, βάσει των τύπων (5.1), (5.7) και (5.2) αντίστοιχα.

Η μέση τιμή των Λ τιμών στις δύο περιπτώσεις συμπίπτουν (ιδιότητα της μέσης τιμής) με την μέση τιμή που δίνεται από την (7.32). Ειδικά για παραμέτρους στις οποίες το A ή το B παίρνει την τιμή Σ οι υπολογισμοί γίνονται μόνο με βάση τα μηνιαία ή δεκαήμερα αθροίσματα. (Βροχή, χιόνι, συχνότητα εμφάνισης φαινομένων, βαθμοημέρες).

3. Ελάχιστη, μέγιστη τιμή, εύρος τιμών.

Η ελάχιστη και η μέγιστη τιμή μιας παραμέτρου για μια περίοδο λ ετών υπολογίζεται από ανάλογες προς τις (7.28) και (7.30) σχέσεις, όπου τώρα η ελάχιστη και η μέγιστη τιμή υπολογίζονται από τις αντίστοιχες λ τιμές. Το εύρος των τιμών δίνεται από την:

$$x(\kappa,EYP(\lambda_1\lambda_2),m^{\alpha},D,A)=x(\kappa,MAX(\lambda_1\lambda_2),m^{\alpha},D,A)-$$
$$-x(\kappa,MIN(\lambda_1\lambda_2),m^{\alpha},D,A) \qquad (7.33)$$

όπου ο δείκτης D δείχνει ότι οι ελάχιστη και η μέγιστη τιμή υπολογίστηκαν από το σύνολο των ημερήσιων τιμών).

4. Σχετικές και αθροιστικές συχνότητες απλών και σύνθετων γεγονότων

Το μέγιστο πλήθος των τιμών που έλαβε η X κατά τον μήνα m^{μ} ή το δεκαήμερο m^{a} είναι:

$$v=x(\kappa,EYP(\lambda_1\lambda_2),m^{a},D,A)/ACCU \quad (7.34)$$

όπου ACCU η ακρίβεια με την οποία καταχωρούνται οι τμές της X. Αν $x_i(m^{a})$, $i\in I_v$ είναι οι v αυτές τιμές, βρίσκονται οι σχετικές συχνότητες τους $F[x_i(m^{a})]$ καθώς και οι συχνότητες των συνόλων $F(A_i)$, $i\in I_s$ όπου το s δίνεται από την:

s=[5*λογ(Λ)] (τύπος ανάλογος του 6.11).

και τα A_i είναι τα σύνολα στα οποία διαμελίζεται το σύνολο των τιμών της X. Ο υπολογισμός των $f[x_i(m^{a})]$ και των $f(A_i)$ απαιτείται κατά τον έλεγχο της προσαρμογής της καμπύλης συχνοτήτων της X ως προς κατάλληλη με την μελετούμενη παράμετρο θεωρητική κατανομή της (παράγραφος 6.1.2.).

5. Οι τιμές $\omega_{m,p}$ για τις οποίες ισχύουν οι σχέσεις:

$$F(x\le\omega_{m,p})=p \quad (7.35)$$

με p=0.01*i και $i\in I100$ καθώς και για p=0.005 και p=0.995 Από τις παραπάνω τιμές πρέπει να αρχειοθετούνται οι τιμές ωm,p για p=0.005, α=0.995 που απαιτούνται κατά τον έλεγχο 1/100 (τεύχος 5/9 ΥΔΡΟΣΚΟΠΙΟ) καθώς και για p=0.025, p=0.05, p=0.1, p=0.2, p=0.25, p=0.5, p=0.75, p=0.8, p=0.9, p=0.95 και p=0.975, που απαιτούνται για την κατασκευή καμπύλων συχνοτήτων. Ειδικά για p=0.5 βρίσκεται η διάμεσος των τιμών της X στο συγκεκριμένο μήνα ή το συγκεκριμένο δεκαήμερο, στην περίοδο των λ ετών.

7.2.3. Κανόνες υπολογισμού-αρχειοθέτισης των μηνιαίων τιμών.

Κατά την ανάπτυξη των τύπων υπολογισμού που έγινε στις δύο προηγούμενες παραγράφους είχε υποτεθεί (σιωπηλά), ότι το αρχείο τιμών της X περιείχε για κάθε μέρα όλες τις N απλές τιμές της. Αυτό όμως αποτελεί ιδιανική περίπτωση και είναι πολύ μικρός ο αριθμός των σταθμών που διαθέτουν "πλήρη" αρχεία.

Στην πραγματικότητα σχεδόν όλοι οι σταθμοί, άλλος περισσότερο και άλλος λιγότερο, έχουν ελλείψεις στις μετρούμενες τιμές τους (οφειλόμενες σε διάφορους λόγους), που πολλές φορές συμπληρώνονται είτε από τα αναγνώσματα των ταινιών των αυτογραφικών οργάνων (για τις παραμέτρους που οι τιμές τους καταγράφονται από αυτογραφικά όργανα) είτε με άλλες μεθόδους (υπάρχει σχετικό τεύχος στο ΥΔΡΟΣΚΟΠΙΟ). Στην περίπτωση που δεν είναι δυνατό να συμπληρωθούν με κάποιο τρόπο οι ελλείψεις πρέπει να λαμβάνεται σοβαρά το πλήθος των ελλειπουσών τιμών κατά τον υπολογισμό των μηνιαίων τιμών.

Ο WMO 1989, προτείνει τον λεγόμενο "**κανόνα 3/5**" σύμφωνα με τον οποίο, μια στατιστική παράμετρος αναφερόμενη σε μηνιαία βάση, υπολογίζεται όταν το πλήθος των ελλειπουσών ημερήσιων τιμών δεν ξεπερνά τις πέντε ελλείψεις ή όταν οι συνεχόμενες ημέρες των ελλείψεων δεν ξεπερά τις τρείς.

Ετσι αν οι διατιθέμενες ημερήσιες τιμές του Ιανουαρίου κάποιου έτους είναι 25, ή αν κατά τον Απρίλιο διατίθενται 26 τιμές και οι οι τέσσερις ελλείπουσες τιμές είναι τεσσάρων διαδοχικών ημερών, δεν πρέπει να υπολογίζονται οι μέσες τιμές των μηνών αυτών. Ο WMO είναι αυστηρότερος στην περίπτωση που οι παράμετροι που υπολογίζονται είναι αθροιστικές τιμές, για τις οποίες απαιτείται να υπολογίζονται από όλες τις ημερήσιες τιμές δηλαδή δεν επιτρέπεται να υπάρχει έλλειψη.

Πολλές φορές, για πρακτικούς λόγους, οι κανόνες αυτοί δεν εφαρμόζονται και υπολογίζονται μηνιαίες τιμές ακόμα και σε περιπτώσεις που το πλήθος των ελλείψεων είναι μεγαλύτερο του επιτρεπόμενου.

Στις περιπτώσεις αυτές τις υπολογιζόμενες τιμές πρέπει να τις συνοδεύει ένας κωδικός "αξιοπιστίας" και να μην λαμβάνονται οι τιμές αυτές υπόψη κατά τον υπολογισμό των ετήσιων τιμών βάσει των μηνιαίων τιμών, οπότε ουσιαστικά δεν πρέπει να υπολογίζονται οι ετήσιες τιμές.

7.2.4. Ετήσιες τιμές.

7.2.4.α. Ενός συγκεκριμένου έτους $l\in I_\Sigma$.

Η ετήσια τιμή μιας παραμέτρου X ενός συγκεκριμένου έτους $l\in I_\Sigma$, μπορούν να υπολογιστούν με δύο τρόπους:

1. Από τις Λ ($\Lambda\in\{365,366\}$) ημερήσιες τιμές της παραμέτρου.
2. Από τις Λ=12 μηνιαίες τιμές ή τις Λ=36 δεκαήμερες τιμές.

Για απλούστευση τις γραφής την μηνιαία ή την δεκαήμερη τιμή του μήνα ή του δεκαήμερου m^α την παριστάνουμε με:

$$x(\kappa,l,m^\alpha,A)=x(\kappa,l,m^\alpha,B,A) \quad \alpha\in\{\mu=\text{μήνας},\ \delta=\text{δεκαήμαρο}\} \qquad (7.36)$$

όπου Α κατάλληλος δείκτης-παράσταση που δείχνει την προέλευση των ημερήσιων τιμών (όπως συζητήθηκε στην παράγραφο 7.1.2) και Β ο ανάλογος δείκτης-παράσταση που δείχνει τι είναι η $x(\kappa,i,m^\alpha,B,A)$ και με:

$$x(\kappa,l,i,A)=x(\kappa,l,m^\alpha,j,A) \quad \text{με } i\in I_\Lambda \text{ και } \Lambda\in\{365,366\} \qquad (7.37)$$

τις ημερήσιες τιμές της παραμέτρου, ανάλογα με το αν το μελετούμενο έτος YEAR (εξίσωση (7.4)) είναι δίσεκτο ή όχι. Ο αριθμός i είναι ο αύξοντας αριθμός της ημέρας μέσα στο έτος (γνωστός σαν Julian Day).

Μετά απ'αυτά, ανάλογα με την παράμετρο, υπολογίζονται και αρχειοθετούνται οι παρακάτω ετήσιες τιμές.

1. Ετήσιος μέσος όρος:

$$x(\kappa,l,E_\Lambda,A)= \Sigma\ x(\kappa,l,i,A)/\Lambda \quad \text{με } i\in I_\Lambda \qquad (7.38)$$

με $\Lambda\in\{12,36,365,366\}$, ανάλογα με το αν ο μέσος όρος υπολογίστηκε από τις μηνιαίες ή τις δεκαήμερες ή τέλος τις ημερήσιες τιμές.

Οι τιμές που προκύπτουν για την μέση ετήσια τιμή μιας παραμέτρου όταν Λ=12, λίγο μόνο διαφέρουν (στο δεύτερο δεκαδικό ψηφίο) από τις αντίστοιχες τιμές που βρίσκονται όταν $\Lambda\in\{365,366\}$ όπως εύκολα μπορεί να το διαπιστώσει κανείς και από τις τιμές του πίνακα 6.9.

Την ίδια σχεδόν τιμή βρίσκει κανείς υπολογίζοντας την ετήσια μέση τιμή από τις τριάντα έξη δεκαήμερες τιμές.

Κατά τον υπολογισμό των ετήσιων μέσων τιμών πρέπει να λαμβάνεται σοβαρά υπόψη ο κανόνας 3/5 και όταν κάποιος μήνας του έτους δεν πληρεί τον κανόνα αυτό πρέπει να τίθεται στην προ κύπτουσα ετήσια τιμή (όταν υπολογιστεί) ένας δείκτης αξιοπιστίας της.

2. Η διακύμανση και η τυπική απόκλιση των Λ $\Lambda \in \{365,366\}$ ημερήσιων τιμών καθώς και των $\Lambda \in \{12,36\}$ μηνιαίων ή δεκαήμερων τιμών, σύμφωνα με τον τύπο (5.1).

3. Η ελάχιστη, η μέγιστη τιμή και το εύρος των ημερήσιων τιμών, σύμφωνα με την (4.1), και οι αντίστοιχες τιμές που προκύπτουν με βάση τις δώδεκα μηνιαίες μέσες τιμές. Ειδικά το εύρος που υπολογίζεται με βάση τον τύπο:

$$T(\kappa,l,ΕΥΡ)=\max[T(\kappa,l,m^{\mu},E,MAX)]-\min[T(\kappa,l,m^{\mu},E,MIN)] \qquad (7.39)$$

με $m^{\mu} \in I_{12}$ ονομάζεται ετήσιο θερμομετρικό (ή/και θερμοκρασιακό) εύρος.

4. Ετήσιο άθροισμα (που συμπίπτει υπολογιζόμενο με τις $\Lambda \in \{365,366\}$ ημερήσιες τιμές ή τις δώδεκα μηνιαίες ή τις 36 δεκαήμερες αθροιστικές τιμές) και δίνεται από τον τύπο:

$$x(\kappa,l,\Sigma_{\Lambda},A)=\sum_{i=1}^{\Lambda} x(\kappa,l,i,A) \qquad (7.40)$$

με $A=\Sigma$ όταν ο υπολογισμός γίνεται με βάση τα ημερήσια αθροίσματα, η $A=n$ όταν η ετήσια τιμή αναφέρεται σε ορισμένη ώρα και για την βροχή ή γενικότερα για τον υετό (βροχή ή χιόνι) και $A=T_{\alpha\tau o}$, ή $A=T_{\kappa\tau o}$ για βαθμοημέρες άνω ή κάτω ορισμένων τιμής Το της θερμοκρασίας.

5. Αριθμός ημερών επικράτησης ορισμένου γεγονότος ή φαινόμενου.

Ο αριθμός αυτός προκύπτει από το άθροισμα των δώδεκα μηνιαίων τιμών επικράτησης του φαινομένου-γεγονότος κατά το συγκεκριμένο έτος, τιμές που υπολογίστηκαν στο προηγούμενο βήμα (μηνιαίες τιμες για το συγκεκριμένο έτος), ή από τις $\Lambda \in \{365,366\}$ ημερήσιες τιμές. Στο υπό εκπόνιση πρόγραμμα της στατιστικής επεξεργασίας των χρονοσειρών των μετεωρολογικών παραμέτρων, πρέπει να προβλεφθεί ρουτίνα υπολογισμού του παραπάνω αριθμού για οποιοδήποτε φαινόμενο-γεγονός επιθυμεί να ορίσει ο μελετητής της χρονοσειράς. Για τις τρείς μετεωρολογικούς παραμέτρους θερμοκρασία του αέρα, βροχή, άνεμο, πρέπει να υπολογίζονται και να αρχειοθετούνται οι παρακάτω αριθμοί ημερών, επικράτησης των παρακάτω φαινομένων - γεγονότων.

<u>Για την θερμοκρασία του αέρα.</u>

5.1 Μερικού παγετού ($Tmin<0°C$)

5.2 Ολικού παγετού ($Tmax<0°C$)

5.3 Εντονου και πολύ έντονου παγετού ($Tmin<5°C$ και $Tmin<10°C$ αντίστοιχα)

5.4 Αριθμός ημερών καύσωνα ($Tmax>35°C$), (ΥΔΡΟΣΚΟΠΙΟ τεύχος 10/2)

5.5 Αριθμός ημερών με $E(T)>30°C$, $T_{max}>25°C$, $T_{min}<5°C$

5.6 Η ελεύθερη παγετού περίοδος του έτους, δηλαδή η περίοδος (τελευταία ημερομηνία του χειμώνα ή της άνοιξης και η πρώτη ημερομηνία του φθινόπωρου και μέχρι το τέλος τους χρόνου) κατά την οποία η θερμοκρασία παρέμεινε πάνω από 0°C.

Για την βροχή.

5.7 Βροχής (ημέρες που σημειώθηκε βροχόπτωση, ανεξάρτητα της ποσότητας).

5.8 Βροχής πάνω από 0.1mm, πάνω από 1mm.

5.9 Εντονης βροχής (πάνω από 5mm).

Για τον άνεμο.

5.10 Απνοιας (ταχύτητα 0 m/sec).

5.11 Ισχυρών ανέμων (ταχύτητα μεγαλύτερη ή ίση από 6 Beaufort σε κάποια παρατήρηση).

5.12 Θυελλωδών ανέμων (ταχύτητα μεγαλύτερη από 7 Beaufort).

7.2.4.β. Ολων των ετών, δεκαετιών, τριάντα ετών (κανονικές τιμές).

7.2.4.β.1. Τιμές υπολογιζόμενες από τις ετήσιες τιμές.

Αν $X(\kappa,L,G_\lambda,A)$ είναι το σύνολο των L ετήσιων τιμών της παραμέτρου X, στο σταθμό με κωδικό $\kappa \in I_\kappa$, όπου οι A, G_λ δείχνουν την προέλευση (τρόπο υπολογισμού) των ετήσιων τιμών, όπως συζητήθηκε παραπάνω και $x(\kappa,l,G_\lambda,A)$ η ετήσια τιμή του έτους με κωδικό έτους $l \in I_L$ τότε με τους τύπους:

$$x(\kappa,E(L),G_\lambda,A)= \Sigma\ x(\kappa,l,G_\lambda,A)/L \qquad (7.41)$$

$$x(\kappa,E(L_1,L_2),G_\lambda,A)= \Sigma\ x(\kappa,l,G_\lambda,A)/LL \quad l \in I_{LL} \qquad (7.42)$$

βρίσκονται αντίστοιχα η μέση τιμή των L ετών (όλα τα έτη) και η μέση τιμή από το έτος L_1 μέχρι το έτους L_2, $(LL=L_2-L_1+1)$.

$$x(\kappa,MM(L),G_\lambda,A)=mm\{x(\kappa,l,G_\lambda,A) \text{ με } l \in I_L\} \qquad (7.43)$$

$$x(\kappa,MM(L_1,L_2),G_\lambda,A)=mm\{x(\kappa,l,G_\lambda,A) \text{ με } l \in I_{LL}\} \qquad (7.44)$$

όπου: $LL=L_2-L_1+1\}$

βρίσκονται οι μέγιστες MM=MAX, mm=max και οι ελάχιστες MM=MIN, mm=min ετήσιες τιμές, και με:

$$x(\kappa,EYP(R),G_\lambda,A)=x(\kappa,MAX(R),G_\lambda,A)-x(\kappa,MIN(R),G_\lambda,A) \qquad (7.45)$$

το εύρος των ετήσιων τιμών των L (R=L) ετών ή της περιόδου L_1,L_2 $(R=L_1L_2)$.

Από τις L ετήσιες τιμές $y_1=x(\kappa,l,G_\lambda,A)$ του συνόλου $X(\kappa,L,G_\lambda,A)$ βρίσκονται και οι συντελεστές α_{yx}, β_{yx},R_{xy} των ευθειών παλινδρόμησης των ξευγών $(x,y)=(l,y_1)$ από τους οποίους προκύπτει αν υπάρχει "μακραίωνα τάση", ή παλινδρόμιση των παραπάνω τιμών. Οι τιμές αυτές είναι απαραίτητες για τον έλεγχο της ομοιογένειας της χρονοσειράς των τιμών της παραμέτρου και συμπληρώνει τους ανάλογους ελέγχους που προτείνονται στο τεύχος 5/9 του ΥΔΡΟΣΚΟΠΙΟ. Για παράδειγμα αναφέρεται ότι με βάση της ετήσιες τιμές της ελάχιστης θερμοκρασίας της περιόδου 1961-1990 που δίνονται στον πίνακα 6.9, η εξίσωση $y_1=\alpha_{y1}*l+\beta_{y1}$ έδωσε:

$$y_1 = -0.0114*l+8.8$$

και έχουν:

r_{y1}=-0.2189.

7.2.4.β.2. Τιμές που υπολογίζονται με βάση τις διαθέσιμες ημερήσιες τιμές.

Τα στοιχεία x(κ,l,m^{α},j,A) του συνόλου X(κ,L,M^{α},J,A), των ημερήσιων τιμών της μετεωρολογικής παραμέτρου X, στον σταθμό $\kappa \in I_{\kappa}$, όπου A κατάλληλος δείκτης ή παράσταση που δείχνει την προέλευση των ημερήσιων τιμών, πορούν να γραφούν και με μορφή:

x_1, x_2, ... x_{Λ}, ή x_i, $i \in I_{\Lambda}$

όπου Λ το πλήθος των ημερών κατά τα L έτη. Γραμμένες με την μορφή αυτή οι ημερήσιες τιμές, μπορούν να θεωρηθούν σαν τιμές μιας χρονοσειράς, και με εφαρμογή των τύπων που δόθηκαν στα κεφάλαια 2 μέχρι 6, ανάλογα και με την παράμετρο X, να βρεθούν:

1. Η ελάχιστη, η μέγιστη τιμή, το διάστημα (πεδίο) των τιμών της και το εύρος (παράγραφος 4.1).

2. Η απόσταση εκατοστών (θερμοκρασία-πίεση) και το ενδοτεταρτημοριακό εύρος (παράγραφος 4.2).

3. Η μέση τιμή (τύποι (4.5) και (4.6).

4. Η επικρατούσα τιμή ή το διάστημα επικρατούσας τιμής (παράγραφος 4.4.1) και η διάμεσος τιμή (παράγραφος 4.4.2).

5. Η διακύμανση και η τυπική απόκλιση (παράγραφος 5.1.1.).

6. Η κυρτώτητα (ή/και κύρτωση) και η ασυμμετρία (τύποι (5.13) και (5.14) αντίστοιχα.

7. Ο πίνακας αυτοδιακύμανσης B_n (ειδικά για την θερμοκρασία και την πίεση μπορεί κανείς να υπολογίσει τον πίνακα αυτοδιακύμακύμανσης και με τις Λ*N όπου N ο αριθμός των ανά ημέρα παρατηρήσεων) για διάφορες τιμές του n. Προτείνεται ο υπολογισμός του B_n για n=5 για την περίπτωση υπολογισμού με βάση τις ημερήσιες τιμές και για n=24 όταν χρησιμοποιούνται οι τιμές των οκτώ συνοπτικών παρατηρήσεων (τρείς ημέρες).

8. Οι συχνότητες f(α), που δίνονται από την (3.2) και οι αθροιστικές συχνότητες F(α), που δίνονται από την (3.17) ή την (3.32).

8. ΣΤΑΤΙΣΤΙΚΑ ΣΤΟΙΧΕΙΑ ΤΟΥ ΑΝΕΜΟΥ.

8.1 Χρήσιμοι ορισμοί και μετασχηματισμοί.

Ο άνεμος, σε αντιδιαστολή με τις άλλες μετεωρολογικές παραμέτρους, είναι μέγεθος διανυσματικό, οριζόμενο από το μέτρο του και την διεύθυνση του ως προς ένα ορισμένο σύστημα πολικών συντεταγμένων. Οι απλές τιμές του ανέμου, σύμφωνα με τους συμβολισμούς που εισήχθησαν στο προηγούμενο κεφάλαιο, και ειδικά στον πίνακα 7.1, θα είναι της μορφής

$$(\rho,\theta)=(\Phi(\kappa,l,m,j,n),\Omega(\kappa,l,m,j,n))=W(\kappa,l,m,j,n) \qquad (8.1)$$

όπου:

$\Omega=WD_s$ όταν πρόκειται για την διεύθυνση του ανέμου επιφανείας και

$\Omega=WD_{dd}$ όταν πρόκειται για την διεύθυνση του ανέμου στην ισοβαρική επιφάνεια dd (σε hPa),

και η ένταση (ταχύτητα του) Φ σύμφωνα με τους συμβολισμούς του πίνακα 7.1 και ανάλογα με τις μονάδες που χρησιμοποιούνται για την μέτρησή της, έχει τις τιμές:

$\Phi=VFS_{MS}$ για την ταχύτητα σε m/s,
$\Phi=VFS_{B}$ για την ταχύτητα σε Beaufort
$\Phi=VFS_{K}$ για την ταχύτητα σε κόμβους (Knots)
$\Phi=VFdd_{K}$ για άνεμο στην ισοβαρική επιφάνεια dd (Knots)
$\Phi=VFdd_{MS}$ για την ισοβαρική επιφάνεια dd σε m/s.

Παρατήρηση.

Η διεύθυνση του ανέμου στην μετεωρολογία μετράται κατά την ανάδρομη φορά και με αρχή μέτρησης των γωνιών τον βορά του τόπου μετρήσεων. Ετσι για την εύρεση της διεύθυνσης $\Omega(\kappa,l,m,j,n)$ του ανέμου σε κανονικό σύστημα πολικών συντεταγμένων και να είναι δυνατή η μετάβαση σε ορθοκανονικό σύστημα καρτεσιανών συντεταγμένων, πρέπει η διεύθυνση $\Omega(\kappa,l,m,j,n)$ να αντικαθίσταται από την τιμή:

$$\Omega(\kappa,l,m,j,n)=[270-\Omega(\kappa,l,m,j,n)]*\pi/180 \qquad (8.2)$$

όπου η ισότητα έχει την έννοια της αντικατάστασης, που χρησιμοποιείται στον προγραμματισμό των ηλεκτρονικών υπολογιστών.

Η διεύθυνση του διανυσματικού ανέμου, που θα προκύψει από τις ορθοκανονικές συντεταγμένες του ανέμου χρησιμοποιώντας την (8.2), στο αρχικό σύστημα πολικών συντεταγμένων (της μετεωρολογίας), γίγίνεται με τον αντίστροφο του (8.2) μετασχηματισμό:

$$\Omega(\kappa,l,m,j,n)=270-[180*\Omega(\kappa,l,m,j,n)]/\pi \qquad (8.3)$$

Αν N είναι ο αριθμός των μετρήσεων του ανέμου ανά ημέρα, (για κλιματολογικούς σκοπούς ενδειαφέρον παρουσιάζει η περίπτωση N=3 και οι τρεις αυτές τιμές αναφέρονται στις παρατηρήσεις των συνοπτικών ωρών 6, 12 και 18UTC), βρίσκουμετις παρακάτω τιμές:

1. <u>Την μέση ένταση ή μέσο μέτρο</u> (Resultant Wind), θεωρόντας τον άνεμο σαν συνάρτηση μόνο της έντασής του, δηλαδή:

$$WF(\kappa,l,m,j,n)=\Phi(\kappa,l,m,j,n) \quad (8.4)$$

χρησιμοποιώντας τους αντίστοιχους του κεφαλαίου 7 τύπους:

$$WF(\kappa,l,m,j,E_N)=\Sigma_{n\in N}\Phi(\kappa,l,m,j,n)/N \quad (8.5)$$

για την ημερήσια τιμή μιας συγκεκριμένης μέρας j,

$$WF(\kappa,l,m,E_J,E_N)=\Sigma_{j\in J}WF(\kappa,l,m,j,E_N)/J \quad (8.6)$$

όπου J είναι ο αριθμός των ημερών του μήνα ή του δεκαήμερου m ($m\in I_{12}$ για μηνιαίες και $m\in I_{36}$ για τιμές δεκαημέρου), για την μηνιαία ή την δεκαήμερη τιμή του συγκεκριμένου έτους:

$$WF(\kappa,E_\Lambda,m,E_J,E_N)=\Sigma\ [WF(\kappa,l,m,E_J,E_N)]/\Lambda \quad \text{με } l\in I_\Lambda \quad (8.7)$$

για την μέση τιμή ενός μήνα ή δεκαήμερου m όλων των διατιθέμενων τιμών (Λ=L) ή συγκεκριμένου αριθμού ετών, (αλλάξοντας το σύνολο I_Λ με το σύνολο των ετών για τα οποία επιθυμούμε την εύρεση της μέσης τιμής):

$$WF(\kappa,l,E,E_J,E_N)=\Sigma\ WF(\kappa,l,m,E_J,E_N)/12 \quad m\in I12 \quad (8.8)$$

για την μέση τιμή ενός συγκεκριμένου έτους, $l\in I_L$,τιμή που μπορεί να υπολογιστεί και με βάση τις τριάντα έξη μέσες δεκαήμερες τιμές του έτους και στην οποία δεν λαμβάνεται υπόψη το διαφορετικό πλήθος των ημερών των διαφόρων μηνών και τέλος η:

$$WF(\kappa,E_\Lambda,E,E_{JJ},E_N)=\Sigma[WF(\kappa,l,E,E_{JJ},E_N)/\Lambda \quad \text{με } l\in I_\Lambda \quad (8.9)$$

για την μέση τιμή Λ ετών και ειδικά, όλων των διατιθέμενων ετών (Λ=L), δεκαετιών και τριάντα ετών (κανονικές τιμές).

2. <u>Τον μέσο διανυσματικό άνεμο</u>, που προκύπτει ως εξής:
Θεωρούμε τις προβολές WX(κ,l,m,j,n) και WY(κ,l,m,j,n) του διανύσματος W(κ,l,m,j,n) στους άξονες X και Y ενός ορθοκανονικού συστήματος καρτεσιανών συντεταγμένων με:

$$WX(\kappa,l,m,j,n)=\Phi(\kappa,l,m,j,n)*\cos[\Omega(\kappa,l,m,j,n)] \quad (8.10)$$

$$WY(\kappa,l,m,j,n)=\Phi(\kappa,l,m,j,n)*\sin[\Omega(\kappa,l,m,j,n)] \quad (8.11)$$

όπου οι γωνίες Ω έχουν μετασχηματιστεί σύμφωνα με την (8.2). Με βάση τις τιμές αυτές υπολογίζονται οι πολικές συντεταγμένες (μέτρο ρ και διεύθυνση θ) του μέσου διανυσματικού ανέμου στις παρακάτω περιπτώσεις:

α. <u>Μία συγκεκριμένη ημέρα j στον συγκεκριμένο σταθμό κ, ενός συγκεκριμένου έτους l</u> από τις εξισώσεις:

$$\rho=V(\kappa,l,m,j,E_N)=SQRT\{[WX(\kappa,l,m,j,E_N)]^2+[WY(\kappa,l,m,j,E_N)]^2\} \quad (8.12)$$

$$\theta=D(\kappa,l,m,j,E_N)=\arctan\left[\frac{WY(\ ,l,m,j,E_N)}{WX(\ ,l,m,j,E_N)}\right] \quad (8.13)$$

όπου:

$$WZ(\kappa,l,m,j,E_N)=\Sigma_{n\in IN}WZ(\kappa,l,m,j,n)/N \quad \text{με } Z\in\{X,Y\} \quad (8.14)$$

β. <u>Τον μήνα ή το δεκαήμερο m, για συγκεκριμένο έτος l, και συγκεκριμένο σταθμό κ</u>, με βάση τις ημερήσιες τιμές των WX και WY από τις εξισώσεις:

$$\rho=V(\kappa,l,m,E_j,E_N)=SQRT\{[WX(\kappa,l,m,E_j,E_N)]^2+[WY(\kappa,l,m,E_j,E_N)]^2\} \quad (8.15)$$

$$\theta=D(\kappa,l,m,E_j,E_N)=arctann\left[\frac{WY(,l,m,E_j,E_N)}{WX(,l,m,E_j,E_N)}\right] \quad (8.16)$$

$$WZ(\kappa,l,m,E_j,E_N)=\Sigma WZ(\kappa,l,m,j,E_N)/J \quad \text{με } j\in I_J \quad (8.17)$$

όπου $Z\in\{X,Y\}$ και J ο αριθμός των ημερών του μελετούμενου μήνα ή δεκαήμερου.

γ. <u>Τον μήνα ή το δεκαήμερο m για Λ χρόνια</u> και κυρίως για όλα τα χρόνια (Λ=L) που διαθέτει ανεμολογικές παρατηρήσεις ο σταθμός κ, τις δεκαετίες της μορφής $\Lambda 0_{10}$ όπου Λ=0,1,..,9, και τις τριακονταετίες της μορφής $\Lambda 0_{30}$, που υπολογίζονται με εφαρμογή ανάλογων προς τους παραπάνω τύπους:

$$\rho=V(\kappa,E,m,E_j,E_N)=SQRT\{[WX(\kappa,E,m,E_j,E_N)]^2+[WY(\kappa,E,m,E_j,E_N)]^2\} \quad (8.18)$$

$$\theta=D(\kappa,E,m,E_j,E_N)=arctann\left[\frac{WY(\kappa,E,m,E_j,E_N)}{WX(\kappa,E,m,E_j,E_N)}\right] \quad (8.19)$$

$$WZ(\kappa,E,m,E_j,E_N)=\Sigma WZ(\kappa,E,m,E_j,E_N)/\Lambda \quad j\in I_\Lambda \quad (8.20)$$

όπου $Z\in\{X,Y\}$, J ο αριθμός των ημερών του μελετούμενου μήνα ή δεκαήμερου και Λ ο αριθμός των μελετούμενων ετών.

δ. <u>Ενα συγκεκριμένο έτος $l\in I_L$</u> που υπολογίζεται με βάση τις δώδεκα μηνιαίες τιμές (ή τις τριάντα έξη δεκαήμερες), με τρόπο ανάλογο με τον περιγραφέντα παραπάνω και ανάλογο με τον τύπο (8.8).

ε. <u>Λ ετών</u>, που υπολογίζεται με βάση τις ετήσιες τιμές των Λ ετών και ειδικά για Λ=L (όλα τα διατιθέμενα χρόνια), τις δεκαετίες και τις τριακονταετίες, όπως περιγράφονται στην περίπτωση γ.

8.2. <u>Παράμετροι προς υπολογισμό και καταχώρηση.</u>

Από το σύνολο όλων των καταχωρημένων στο αρχείο ανεμολογικών παρατηρήσεων του σταθμού, βρίσκονται:

α.1 Οι μέγιστες εντάσεις κάθε μήνα ή δεκαήμερου του κάθε έτους,

α.2 Οι αντίστοιχες τιμές κάθε μήνα ή δεκαήμερου όλων των ετών ή Λ ετών,

α.3 Οι μέγιστες εντάσεις κάθε έτους, και Λ ετών (κυρίως όλων των ετών ή δεκαετιών ή τριακονταετιών).

Οι μέγιστες αυτές εντάσεις βρίσκονται τόσο συνολικά όσο και ξεχωριστά για κάθε διεύθυνση (οκτώ ή τριάντα έξη διευθύνσεις, ανάλογα με τον σταθμό).

β.1 Περιπτώσεις (συχνότητες) άπνοιας ανά μήνα-δεκαήμερο και έτος τόσο ξεχωριστά για κάθε έτος όσο και για Λ (Λ≤L) έτη.

β.2 Οι ανόλογες περιπτώσεις σε ετήσια βάση, τόσο ξεχωριστά για κάθε έτος όσο και για Λ≤L έτη (δεκαετίες-τριακονταετίες).

γ. Περιπτώσεις (και σχετικές ή σχετικές επί τα εκατό συχνότητες που προκύπτουν απ'αυτές και τις αντίστοιχες αθροιστικές τιμές) ανά μήνα ή δεκαήμερο και έτος (ξεχωριστά για κάθε έτος και για Λ έτη), κατά τις οποίες το μέτρο (η ένταση) του ανέμου ξεπερνούσε τα 4,5,6,7,8 και 9 Beaufort.

Οι περιπτώσεις αυτές υπολογίζονται για κάθε διεύθυνση ξεχωριστά όσο και στο σύνολο τους. Η διεύθυνση με την μεγαλύτερη συχνότητα εμφάνισης ονομάζεται επικρατούσα διεύθυνση και παίζει σπουδαίο ρόλο στους κατασκευαστικούς σχεδιασμούς κτηρίων γεφυρών κλπ.

δ. Περιπτώσεις, όπως ακριβώς στην προηγούμενη γ περίπτωση, αλλά για τιμές έντασης 5,6,7,8,9,10,12,15,20 και 25m/s, τιμές που έχουν μεγάλη σημασία στους κατασκευαστικούς σχεδιασμούς και κυρίως στον σχεδιασμό της εγκατάστασης ανεμογεννητριών.

Η καταχώρηση των συχνοτήτων ανά διεύθυνση και ένταση επιτρέπει την κατασκευή του ροδοδιαγράμματος του ανέμου ή, όπως απλούστερα λέγεται, του ρόδου του ανέμου, σε δεκαήμερη ή μηνιαία ή ετήσια βάση.

ε. Τιμές της έντασης V_p του ανέμου για τις οποίες η αθροιστική συχνότητα (ανά μήνα ή δεκαήμερο ή έτος των Λ, Λ≤L, ετών) ήταν μικρότερη ή ίση του p/100 και για τιμές του p από το σύνολο:

$p\in\{0,5,10,20,25,50,75,80,90,95,100\}$

δηλαδή:

$$F(\rho\leq V_p)=p/100 \tag{8.21}$$

Με τις τιμές αυτές μπορούμε να κατασκευάσουμε διαγράμματα με άξονες τον χρόνο (δεκαήμερα ή μήνες) και την ένταση του ανέμου, στα οποία να χαραχθούν οι καμπύλες ίσων πιθανοτήτων p (σχήμα 8.1).

στ. Αθροιστικές (%) συχνότητες $F(V_o,\lambda)$ των περιπτώσεων κατά τις οποίες (ανεξάρτητα της διεύθυνσης) η ένταση V του ανέμου ήταν μεγαλύτερη από την τιμή V_o ($V_o>0$) για λ (λ≥1) συνεχόμενες παρατηρήσεις του ανέμου και για διάφορες τιμές του V_o.

Οι αθροιστικές αυτές συχνότητες $F(V_o,\lambda)$, υπολογίζονται ως εξής:

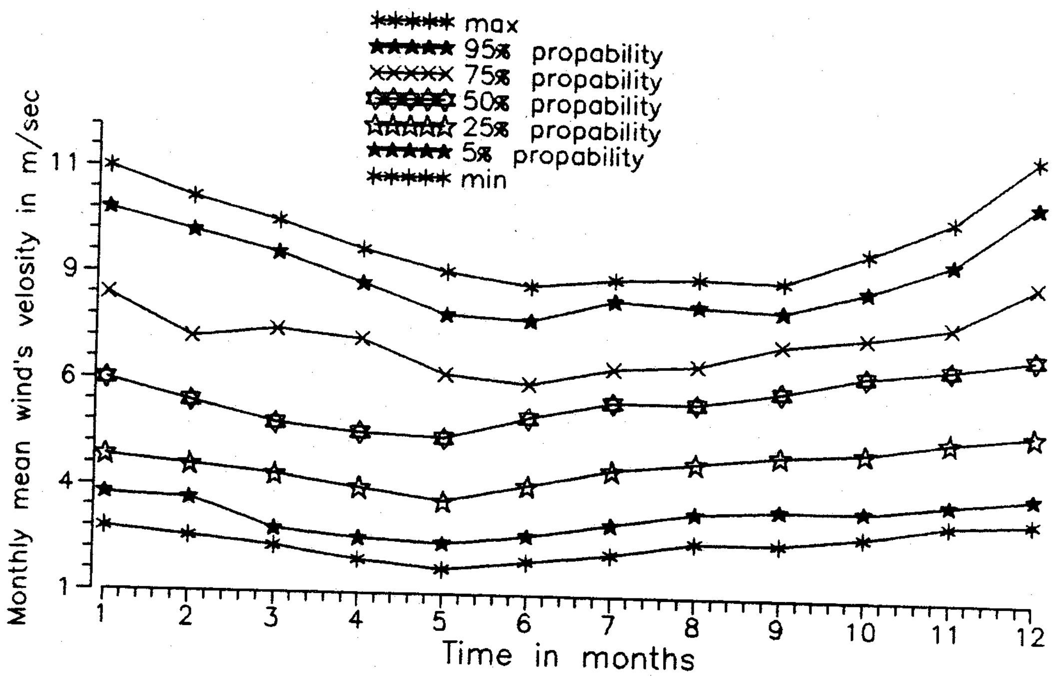
Θεωρούμε ότι το σύνολο των εντάσεων του ανέμου αποτελείται από M τιμές, που ελήφθησαν με περίοδο δείγματος ΔT, τις:

$y_1=V(t_o)$, $y_2=V(t_o+\Delta T)$, $y_3=V(t_o+2*\Delta T)$,..., $y_M=V[t_o+(M-1)*\Delta T]$

όπου t_o η χρονική στιγμή της πρώτης παρατήρησης. Με την μορφή αυτή προκύπτει η χρονοσειρά $\{y_i\}$ με $i\in I_M$.

Για μια τιμή V_o της έντασης του ανέμου θέτουμε:

$$x(i)=\begin{cases}1 & \text{αν } y_i\geq V_o\\ 0 & \text{αν } y_i<V_o\end{cases} \qquad i\in I_M \tag{8.22}$$



Σχήμα 8.1 Καμπύλες μέσων μηνιαίων τιμών της έντασης του ανέμου σε υποθετικό σταθμό, με ορισμένη αθροιστική συχνότητα.

$$x_\lambda(i)=\begin{cases} 1 & \text{αν } \chi(i-\lambda+1)=1 \quad i\in M_\lambda=I_{M-\lambda+1},\quad \lambda\in N^* \\ 0 & \text{αν } \chi(i-\lambda+1)=0 \text{ για ένα τουλάχιστο } \lambda \end{cases} \tag{8.23}$$

Για $\lambda=1$ η (8.23) δίνει την (8.22)

Μετά απ'αυτά:

$$F(V_o,\lambda)= 100*[\Sigma x_\lambda(i)]/(M-\lambda+1) \quad i\in I_M \tag{8.24}$$

Οι τιμές του λ μέχρι το οποίο βρίσκονται οι $F(V_o,\lambda)$ είναι συνάρτηση της μέγιστης έντασης V_o του ανέμου και του ΔT και σε διάγραμμα με άξονες τον χρόνο t και την ταχύτητα V η καμπύλη με $F(V_o,\lambda)=0$ έχει συνήθως την μορφή του σχήματος 8.2 (Desire Le Gourieres 1982). Το διάγραμμα αυτό είναι ένα από τα σημαντικώτερα βοηθήματα κατά τον υπολογισμό της διαθέσιμης αιολικής ενέργειας στην περιοχή του σταθμού και στον σχεδιασμό της βέλτιστης διαμέτρου και τον αριθμό των πτερυγίων της προς εγκατάσταση ανεμογεννήτριας που πρέπει να τοποθετηθεί.

στ. Ο λόγος:

$$a=\rho/WF \tag{8.25}$$

του μέτρου ρ του διανυσματικού ανέμου προς τον μέσο άνεμο (σε μηνιαία ή δεκαήμερη ή ετήσια βάση για Λ έτη). Ο λόγος αυτός α παίρνει τιμές στο διάστημα [0,1] και μεγάλες τιμές του δείχνουν εμμονή του ανέμου στην επικρατούσα διεύθυνση.

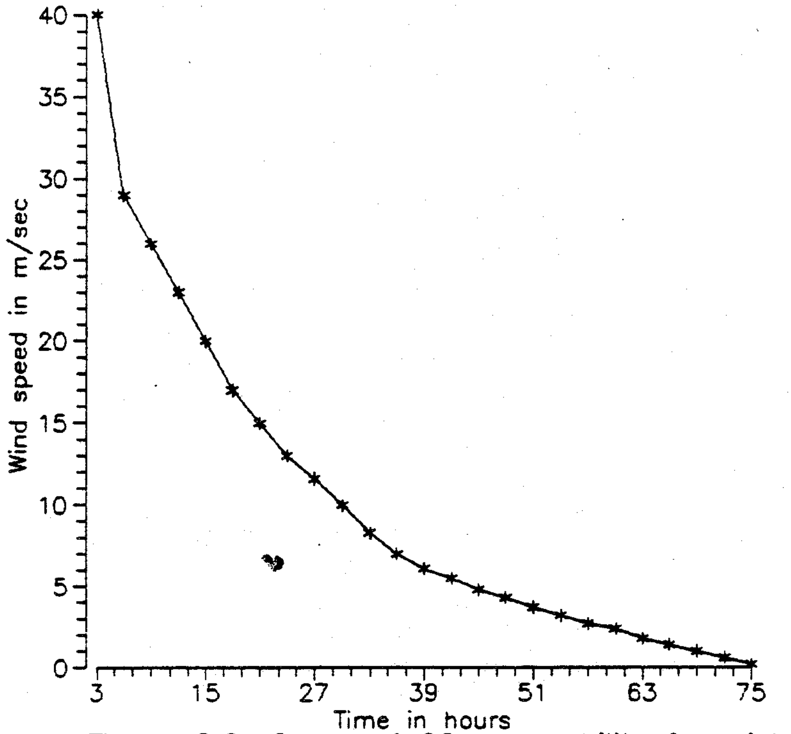


Figure 8.2. Curve of 99% propability for wich the wind's speed remain less than less than V_0 for t consecutives hours

ΑΝΑΦΟΡΕΣ

1. Αδαμόπουλος,Λ.,Α.: Στατιστική. Πανεπιστημιακαί παραδόσεις. Τρεις τόμοι. Θεσσαλονίκη 1963.
2. Βασιλείου,Χ.,Π.: Στοχαστικές μέθοδοι στις επιχειρησιακές έρευνες. Εκδόσεις Χριστοδουλίδη. Θεσσαλονίκη 1985.
3. Κούνια,Σ.- Κολυβά-Μαχαίρα,Φ.- Μπαγιάτη,Κ.- Μπόρα-Σέντα,Ε.: Εισαγωγή στην Στατιστική.
Εκδόσεις Χριστοδουλίδη. Θεσσαλονίκη 1990.
4. Μονόπωλης,Γ.: Στοιχεία Υδρολογίας.
ΓΕΑ/ΕΜΥ, Αθήνα. Δεύτερη Εκδοση 1978.
5. Χαραντώνης,Θ. και Παπασταμόπουλος,Ι.: Θεωρητικό μοντέλο προ-προσαρμογής της εντάσεως του ανέμου στην επιφάνεια και τα 900MBS στην Αθήνα. Διατριβή για την απόκτιση του Μεταπτυχιακού ενδεικτικού Μετεωρολογίας. Πανεπιστήμιο Αθηνών, 1982.
6. Χαραντώνης Θ.: Μοντέλο Prog για πρόγνωση της ελάχιστης θερ-μορκρασίας. Εφαρμογή στο Ελληνικό και την Νέα Φιλαδέλφεια. ΓΕΑ/ΕΜΥ. Αριθμός μελέτης ΕΜΥ:17. 1992.
7. Χαραντώνης Θ.: Συνοπτική θεωρία των αλύσεων Markov και εφαρ-μογή της στη μελέτη της μέγιστης ημερήσιας θερμοκρασίας στην Αθήνα. ΓΕΑ/ΕΜΥ, 1988.
8. Ψώινος,Δ.,Π.: Στατιστική. Θεσσαλονίκη 1978

Εκτός των παραπάνω χρησιμοποιήθηκαν και αρκετά από τα τεύχη του ΥΔΡΟΣΚΟΠΙΟ κυρίως τα με αριθμό τεύχους: 5/9 και 10/2 και 3/1.

---

9. Box,P.E.G. and Jenkins,M.,G.1969: Time Series Analysis fore-casting and Control. HOLDEN-DAY, San Francisco, Dusseldorf, Johannesburg,London,Panama,Singapure,Sydney,Toronto. Library of Congress Catalog. Card Number 77-79534. Printed in the USA, 1970.
10. Charantonis,Th.1989: Solor and Wind Energy in the isle of Naxos. 2ο Ευρωπαικό Συμπόσιο "Soft energy Sources an Systems at the local level", Chania,Crete,Greece.
11. Chow,T.,V.1951: A general formula for hydrologic frequency analysis. Trans.Am.Geoph.Un.Vol32,p.p.231-237,April 1951
12. Gourieres (Le), D.: Energie Eolienne. Theorie, conception et calcul pratique des Installations. EUROLLES Paris 1982.
13. Goldberg,S: Probability, an introduction. Μετάφραση στην Ελληνική με τίτλο: Πιθανότης. Εισαγωγή από τον Κ.Γ.Παναγάκη (Εκδόσεις Πεχλιβανίδη και ΣΙΑ) 1962.
14. Gramer,H.:Mathematical Methods of Statistics. Princeston Univ.Press, Princeston,N.J.1946 p.p.251. 1946.

15. Justus,C.,G and Hargraves,R. and Yalcin,A.: Nation Wide Assessment of Potential Output from Wind-Powered Generators. Journal of Applied Meteorology, Vol.15, No 7, July 1976 p.p.673-678.

16. Kendall,M.,G. and Stuart,A.: The Advanced Theory of Statistics. Char. Griffin & Company Limited. Lovdon & High Wycombe, 1976. Τρεις τόμοι.

17. Panofsky,A.,H. and Brier,-.,G.: Some Applications of Statistics to Meteorology. Mineral Industries Continuing Education College of Mineral Industries. The Pennsylvania State University. University Park, Pennsylvania. Firs edition 1963.

18. Remenieras,G.1959: Elements d' hydrologie Appliquee. Collection Armand Colin No 343 Section Electricite, Librairie Armand Colin.

19. World Meteorological Organisation (WMO): Calculation of Monthly and Annual 30-Years Standard Normals. WMO-TD/No 341, WCDP-No.10. March 1989.

20. WMO.: Guide to Climatological Practices. WMO No.100

ΠΑΡΑΡΤΗΜΑ
ΣΤΑΤΙΣΤΙΚΩΝ ΠΙΝΑΚΩΝ

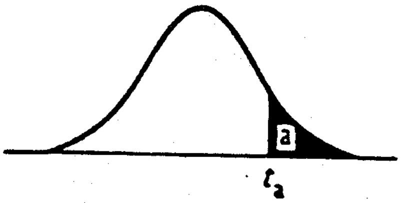
Πίνακας 6.1 Τιμές της πιθανότητας $P(0 \leq Z \leq Z_{\alpha})$ συναρτήσει του Z_{α} για την τυχαία μεταβλητή Z που ακολουθεί την κανονική κατανομή με μέση τιμή μ=0 και τυπική απόκλιση σ=1, $Z \approx N(Z,0,1)$

Z_{α}	.00	.001	.002	. 003	.004	.005	.006	.007	.008	.009
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1737	.1773	.1808	.1844	.1880
.5	.1915	.1950	.1985	.2020	.2054	.2089	.2123	.2157	.2191	.2224
.6	.2258	.2291	.2324	.2357	.2390	.2422	.2454	.2486	.2518	.2549
.7	.2581	.2612	.2643	.2674	.2704	.2734	.2764	.2794	.2824	.2853
.8	.2882	.2911	.2939	.2968	.2996	.3024	.3052	.3079	.3106	.3133
.9	.3160	.3187	.3213	.3239	.3265	.3290	.3315	.3341	.3365	.3390
1.0	.3414	.3438	.3462	.3486	.3509	.3532	.3555	.3578	.3600	.3622
1.1	.3644	.3666	.3687	.3709	.3730	.3750	.3771	.3791	.3811	.3831
1.2	.3850	.3870	.3889	.3908	.3926	.3945	.3963	.3981	.3998	.4016
1.3	.4032	.4050	.4067	.4084	.4100	.4116	.4132	.4148	.4163	.4179
1.4	.4194	.4209	.4223	.4238	.4252	.4266	.4280	.4294	.4307	.4320
1.5	.4333	.4346	.4359	.4371	.4384	.4396	.4408	.4419	.4431	.4442
1.6	.4453	.4464	.4475	.4486	.4496	.4507	.4517	.4527	.4537	.4546
1.7	.4556	.4565	.4574	.4583	.4592	.4601	.4610	.4618	.4626	.4634
1.8	.4642	.4650	.4658	.4665	.4673	.4680	.4687	.4694	.4701	.4708
1.9	.4715	.4721	.4727	.4734	.4740	.4746	.4752	.4758	.4763	.4769
2.0	.4774	.4780	.4785	.4790	.4795	.4800	.4805	.4810	.4814	.4819
2.1	.4823	.4827	.4832	.4836	.4840	.4844	.4848	.4852	.4856	.4859
2.2	.4863	.4866	.4870	.4873	.4876	.4880	.4883	.4886	.4889	.4892
2.3	.4895	.4897	.4900	.4903	.4905	.4908	.4910	.4913	.4915	.4918
2.4	.4920	.4922	.4924	.4926	.4928	.4930	.4932	.4934	.4936	.4938
2.5	.4940	.4942	.4943	.4945	.4946	.4948	.4950	.4951	.4953	.4954
2.6	.4955	.4957	.4958	.4959	.4960	.4962	.4963	.4964	.4965	.4966
2.7	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974	.4975	.4976
2.8	.4976	.4977	.4978	.4979	.4979	.4980	.4981	.4981	.4982	.4983
2.9	.4983	.4984	.4984	.4985	.4986	.4986	.4987	.4987	.4988	.4988
3.0	.4988	.4989	.4989	.4990	.4990	.4991	.4991	.4991	.4992	.4992
3.1	.4992	.4993	.4993	.4993	.4994	.4994	.4994	.4994	.4995	.4995
3.2	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4997	.4997	.4997
3.3	.4997	.4997	.4997	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.4	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.5000
3.5	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

Πίνακας 6.2 Τιμές της πιθανότητας $P(-Z_\alpha \leq Z \leq Z_\alpha)$ συναρτήσει του Z_α, για την τυχαία μεταβλητή Z που ακολουθεί την κανονική κατανομή, με μέση τιμή μ=0 και τυπική απόκλιση σ=1.

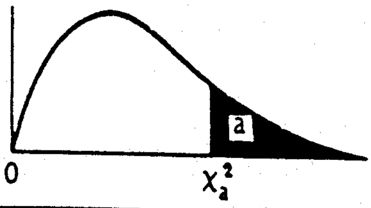
Z_α	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.0	.0000	.0080	.0160	.0239	.0319	.0399	.0478	.0558	.0638	.0717
.1	.0797	.0876	.0955	.1034	.1113	.1192	.1271	.1350	.1429	.1507
.2	.1585	.1663	.1741	.1819	.1897	.1974	.2051	.2129	.2205	.2282
.3	.2358	.2435	.2511	.2586	.2662	.2737	.2812	.2886	.2961	.3035
.4	.3109	.3182	.3255	.3328	.3401	.3473	.3545	.3617	.3688	.3759
.5	.3830	.3900	.3970	.4039	.4109	.4177	.4246	.4314	.4381	.4449
.6	.4516	.4582	.4648	.4714	.4779	.4844	.4908	.4972	.5036	.5099
.7	.5162	.5224	.5286	.5347	.5408	.5468	.5528	.5588	.5647	.5706
.8	.5764	.5822	.5879	.5936	.5992	.6048	.6103	.6158	.6213	.6267
.9	.6320	.6373	.6426	.6478	.6529	.6580	.6631	.6681	.6731	.6780
1.0	.6828	.6877	.6924	.6972	.7018	.7065	.7110	.7156	.7200	.7245
1.1	.7288	.7332	.7375	.7417	.7459	.7500	.7541	.7582	.7622	.7662
1.2	.7701	.7739	.7777	.7815	.7852	.7889	.7925	.7961	.7997	.8032
1.3	.8066	.8100	.8134	.8167	.8200	.8232	.8264	.8296	.8327	.8357
1.4	.8387	.8417	.8446	.8475	.8504	.8532	.8560	.8587	.8614	.8640
1.5	.8667	.8692	.8718	.8743	.8767	.8791	.8815	.8839	.8862	.8885
1.6	.8907	.8929	.8951	.8972	.8993	.9014	.9034	.9054	.9073	.9093
1.7	.9112	.9130	.9149	.9167	.9185	.9202	.9219	.9236	.9252	.9269
1.8	.9285	.9300	.9316	.9331	.9346	.9360	.9374	.9388	.9402	.9416
1.9	.9429	.9442	.9455	.9467	.9480	.9492	.9503	.9515	.9526	.9538
2.0	.9548	.9559	.9570	.9580	.9590	.9600	.9610	.9619	.9628	.9637
2.1	.9646	.9655	.9664	.9672	.9680	.9688	.9696	.9704	.9711	.9718
2.2	.9726	.9733	.9739	.9746	.9753	.9759	.9765	.9772	.9778	.9783
2.3	.9789	.9795	.9800	.9806	.9811	.9816	.9821	.9826	.9831	.9835
2.4	.9840	.9844	.9849	.9853	.9857	.9861	.9865	.9869	.9872	.9876
2.5	.9880	.9883	.9886	.9890	.9893	.9896	.9899	.9902	.9905	.9908
2.6	.9911	.9913	.9916	.9918	.9921	.9923	.9926	.9928	.9930	.9932
2.7	.9935	.9937	.9939	.9941	.9942	.9944	.9946	.9948	.9950	.9951
2.8	.9953	.9954	.9956	.9957	.9959	.9960	.9962	.9963	.9964	.9965
2.9	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974	.9975	.9976
3.0	.9977	.9978	.9979	.9980	.9980	.9981	.9982	.9983	.9983	.9984
3.1	.9985	.9985	.9986	.9986	.9987	.9988	.9988	.9989	.9989	.9990
3.2	.9990	.9991	.9991	.9992	.9992	.9992	.9993	.9993	.9994	.9994
3.3	.9994	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9997	.9997
3.4	.9997	.9997	.9998	.9998	.9998	.9998	.9999	.9999	.9999	.9999
3.5	.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Πινακας 6.3. Των τιμών της t- κατανομής ώστε $P(t_v > t_{v,\alpha}) = \alpha$.



Β.ε	a = .10	a = .05	a = .025	a = .010	a = .005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
inf.	1.282	1.645	1.960	2.326	2.576

Πινακας 6.4. Των τιμών $X^2_{\nu,\alpha}$ της X^2 κατανομής για τις οποιες ισχύει $P(X^2 > X^2_{\nu,\alpha}) = \alpha$.

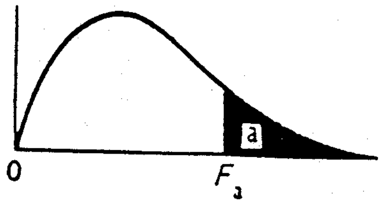


Β.ε	a = .995	a = .990	a = .975	a = .950	a = .900
1	0.0000393	0.0001571	0.0009821	0.0039321	0.0157908
2	0.0100251	0.0201007	0.0506356	0.102587	0.210720
3	0.0717212	0.114832	0.215795	0.351846	0.584375
4	0.206990	0.297110	0.484419	0.710721	1.063623
5	0.411740	0.554300	0.831211	1.145476	1.61031
6	0.675727	0.872085	1.237347	1.63539	2.20413
7	0.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518
11	2.60321	3.05347	3.81575	4.57481	5.57779
12	3.07382	3.57056	4.40379	5.22603	6.30380
13	3.56503	4.10691	5.00874	5.89186	7.04150
14	4.07468	4.66043	5.62872	6.57063	7.78953
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509
20	7.43386	8.26040	9.59083	10.8508	12.4426
21	8.03366	8.89720	10.28293	11.5913	13.2396
22	8.64272	9.54249	10.9823	12.3380	14.0415
23	9.26042	10.19567	11.6885	13.0905	14.8479
24	9.88623	10.8564	12.4011	13.8484	15.6587
25	10.5197	11.5240	13.1197	14.6114	16.4734
26	11.1603	12.1981	13.8439	15.3791	17.2919
27	11.8076	12.8786	14.5733	16.1513	18.1138
28	12.4613	13.5648	15.3079	16.9279	18.9392
29	13.1211	14.2565	16.0471	17.7083	19.7677
30	13.7867	14.9535	16.7908	18.4926	20.5992
40	20.7065	22.1643	24.4331	26.5093	29.0505
50	27.9907	29.7067	32.3574	34.7642	37.6886
60	35.5346	37.4848	40.4817	43.1879	46.4589
70	43.2752	45.4418	48.7576	51.7393	55.3290
80	51.1720	53.5400	57.1532	60.3915	64.2778
90	59.1963	61.7541	65.6466	69.1260	73.2912
100	67.3276	70.0648	74.2219	77.9295	82.3581

Πινακας 6.4 . Των τιμών $X^2_{\nu,\alpha}$ της X^2 κατανομής για τις οποίες ισχύει $P(X^2 > X^2_{\nu,\alpha}) = \alpha$.

a = .10	a = .05	a = .025	a = .010	a = .005	B.E.
2.70554	3.84146	5.02389	6.63490	7.87944	1
4.60517	5.99147	7.37776	9.21034	10.5966	2
6.25139	7.81473	9.34840	11.3449	12.8381	3
7.77944	9.48773	11.1433	13.2767	14.8602	4
9.23635	11.0705	12.8325	15.0863	16.7496	5
10.6446	12.5916	14.4494	16.8119	18.5476	6
12.0170	14.0671	16.0128	18.4753	20.2777	7
13.3616	15.5073	17.5346	20.0902	21.9550	8
14.6837	16.9190	19.0228	21.6660	23.5893	9
15.9871	18.3070	20.4831	23.2093	25.1882	10
17.2750	19.6751	21.9200	24.7250	26.7569	11
18.5494	21.0261	23.3367	26.2170	28.2995	12
19.8119	22.3621	24.7356	27.6883	29.8194	13
21.0642	23.6848	26.1190	29.1413	31.3193	14
22.3072	24.9958	27.4884	30.5779	32.8013	15
23.5418	26.2962	28.8454	31.9999	34.2672	16
24.7690	27.5871	30.1910	33.4087	35.7185	17
25.9894	28.8693	31.5264	34.8053	37.1564	18
27.2036	30.1435	32.8523	36.1908	38.5822	19
28.4120	31.4104	34.1696	37.5662	39.9968	20
29.6151	32.6705	35.4789	38.9321	41.4010	21
30.8133	33.9244	36.7807	40.2894	42.7956	22
32.0069	35.1725	38.0757	41.6384	44.1813	23
33.1963	36.4151	39.3641	42.9798	45.5585	24
34.3816	37.6525	40.6465	44.3141	46.9278	25
35.5631	38.8852	41.9232	45.6417	48.2899	26
36.7412	40.1133	43.1944	46.9630	49.6449	27
37.9159	41.3372	44.4607	48.2782	50.9933	28
39.0875	42.5569	45.7222	49.5879	52.3356	29
40.2560	43.7729	46.9792	50.8922	53.6720	30
51.8050	55.7585	59.3417	63.6907	66.7659	40
63.1671	67.5048	71.4202	76.1539	79.4900	50
74.3970	79.0819	83.2976	88.3794	91.9517	60
85.5271	90.5312	95.0231	100.425	104.215	70
96.5782	101.879	106.629	112.329	116.321	80
107.565	113.145	118.136	124.116	128.299	90
118.498	124.342	129.561	135.807	140.169	100

Πινακας 6.5.1. Των τιμών $F_{v1,v2,\alpha}$ της F-κατανομής για τις οποίες ισχύει $P(F>F_{v1,v2,\alpha})=\alpha$.



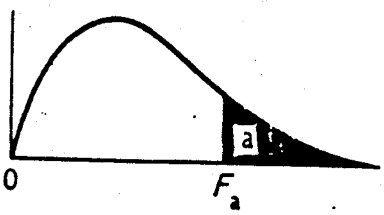
Βαθμοι Ελευθεριας (a = .05)

v_1 / v_2	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Πίνακας 6.5.1.Των τιμών $F_{v1,v2,\alpha}$ της F-κατανομής για τις οποίες ισχύει $P(F>F_{v1,v2,\alpha})=\alpha$.

10	12	15	20	24	30	40	60	120	∞	v_1 / v_2
241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3	1
19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50	2
8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53	3
5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	4
4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36	5
4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	6
3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23	7
3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93	8
3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71	9
2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54	10
2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40	11
2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30	12
2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21	13
2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13	14
2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	15
2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01	16
2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	17
2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	18
2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	19
2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	20
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81	21
2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78	22
2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76	23
2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73	24
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71	25
2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69	26
2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67	27
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65	28
2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64	29
2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62	30
2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51	40
1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39	60
1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25	120
1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00	∞

Πίνακας 6.5.2 Των τιμών $F_{v_1,v_2,\alpha}$ της F-κατανομής για τις οποίες ισχύει $P(F>F_{v_1,v_2,\alpha})=\alpha$.



Βαθμοί ελευθερίας (a = .01)

v_2 \ v_1	1	2	3	4	5	6	7	8	9
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Πινακας 6.5.2 Των τιμών $F_{\nu_1,\nu_2,\alpha}$ της F-κατανομής για τις οποίες ισχύει $P(F>F_{\nu_1,\nu_2,\alpha})=\alpha$.

10	12	15	20	24	30	40	60	120	∞	ν_1 / ν_2
6056	6106	6157	6209	6235	6261	6287	6313	6339	6366	1
99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50	2
27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13	3
14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46	4
10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02	5
7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88	6
6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65	7
5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86	8
5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31	9
4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91	10
4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60	11
4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36	12
4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17	13
3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00	14
3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87	15
3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75	16
3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65	17
3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57	18
3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49	19
3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42	20
3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36	21
3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31	22
3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26	23
3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21	24
3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17	25
3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13	26
3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10	27
3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06	28
3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03	29
2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01	30
2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80	40
2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60	60
2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38	120
2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00	∞

Πίνακας 6.7. Βασικές κατανομές συχνοτήτων της τυχαίας μεταβλητής X με X(t)=x.

α/α	Κατανομή της X	σ.π.π ή σ.α.π f(x) ή F(x)	Τιμές Παραμέτρων	E(X) μέσος όρος	Var(x) διακύμανση	Τιμές χ
		Διακριτές κατανομές.				
1	Bernoulli	$f(x)=p^{x}*q^{1-x}$	$q=1-p \quad 0<p<1$	p	$p*q$	$\chi\in\{0,1\}$
2	Διωνυμική	$f(x)=\begin{bmatrix} M \\ x \end{bmatrix}*p^{x}*q^{M-x}$	$q=1-p \quad 0<p<1$	Mp	Mpq	$x<M$, $M\in N$
3	Poisson	$f(x)=\lambda^{x}*e^{-\lambda}/\chi!$	$\lambda>0$	λ	λ	$x\in N$
4	Γεωμετρική	$f(x)=p*q^{x-1}$	$q=1-p \quad 0<p<1$	$1/p$	q/p^2	$x\in N^*$
5	Αρνητική διωνυμική	$f(x)=\begin{bmatrix} M-1 \\ x-1 \end{bmatrix}*p^{x}*q^{M-x}$	$q=1-p \quad 0<p<1$	x/p	xq/p^2	$x<M$, $M\in N^*$
		Συνεχείς κατανομές.				
1	Ομοιόμορφη	$f(x)=1/(\beta-\alpha)$	$\beta>\alpha$	$(\alpha+\beta)/2$	$(\beta-\alpha)^2/2$	$x\in[\alpha,\beta]$
2	Κανονική	$f(x)=\dfrac{\exp[-(x-\mu)^2/(2*\sigma^2)]}{\sigma*(2*\pi)^{0.5}}$	$\mu\in R,\ \sigma>0$	μ	σ^2	$x\in R$
3	Λογαριθμο-κανονική	$f(x)=\dfrac{\exp[-(\ln x-\mu)^2/(2*\sigma^2)]}{\chi*\sigma*(2*\pi)^{0.5}}$	$\mu\in R,\ \sigma>0$	$E(\ln X)$	$Var(\ln X)$	$x>0$
4	Weibull	$f(x)=\alpha\beta\chi^{\beta-1}\exp(-\alpha\chi^{\beta})$	$\alpha>0 \quad \beta>0$	$\alpha^{1/\beta}\Gamma(1+1/\beta)$	$\alpha^{-2/\beta}\Gamma(1+2/\beta)-\Gamma^2(1+1/\beta)$	$x>0$
5	Γεωμετρική	$f(x)=\lambda*e^{-\lambda x}$	$\lambda>0$	$1/\lambda$	$1/\lambda^2$	$x>0$
6	Γάμα	$f(x)=\dfrac{x^{\gamma-1}*e^{-x/\beta}}{\beta^{\gamma}*\Gamma(\gamma)}$	$\beta>0 \quad \gamma>0$	$\beta*\gamma$	$\beta^2*\gamma$	$x>0$
7	Gumbell	$f(x)=\exp\{-\exp[-a*(x-x_0)]\}$	$a>0$	$x_0+0.577/a$	$(0.78*a)^{-2}$	$x\in R$
8	t_v	$f(x)=\dfrac{\Gamma(A)*(1+x/v)^{-(v+1)/2}}{\Gamma(v/2)}$	$A=(v+1)/2$, $v\geq3,\ v\in N$	0	$v/(v-2)$	$x\in R$
9	X^2	$f(x)=\dfrac{x^{0.5v-1}*\exp(-x/2)}{\Gamma(v/2)*2^{v/2}}$	$v\in N^*$	v	$2*v$	$x>0$
10	$F_{\kappa,\lambda}$	$f(x)=\dfrac{A*B*\chi^{0.5\lambda-1}(1+\kappa\chi/\lambda)}{\Gamma(\kappa/2+\lambda/2)/C}$	$A=(\kappa/\lambda)^{\kappa/2}$, $B=\Gamma(\kappa/2+\lambda/2)$, $C=\Gamma(\kappa/2)*\Gamma(\lambda/2)$	$\lambda/(\lambda-2)$, $\lambda>2$	$2(\kappa+\lambda-2)/[\kappa*(\lambda-4)]$, $\lambda>4$	$x>0$, $\kappa\in N^*$, $\lambda\in N^*$
11	Πολυωνιμική	$f(x)=\Sigma\alpha_i*x^{i}$	$i\in I_k \quad \alpha_i\in R$	$\int_{-\infty}^{\infty} x*f(x)dx$	$\int_{-\infty}^{\infty} x^2*f(x)dx$	$x\in R$

Πίνακας 6.8. Κατανομές των βασικών μετεωρολογικών παραμέτρων.

α/α	Παράμετρος	Κατανομή συχνότητας (πλέον πιθανή)
1	Θερμοκρασία	
α	Απλές τιμές (πολλών ετών) στο σύνολό τους ή ξεχωριστά κάθε ώρας παρατήρησης.	
1	Ολες οι διαθέσιμες τιμές	Κανονική
2	Ολες οι διαθέσιμες τιμές για ορισμένο μήνα ή δεκαήμερου.	Σχεδόν κανονική με μικρή ασσυμετρία και κυρτώτητα εξαρτώμενη από τον μήνα ή το δεκαήμερο που εξετάζεται.
3	Ελάχιστη ή μέγιστη ημέρας	Κανονική
4	Μέση ημερήσια	Κανονική
5	Βαθμοημέρες άνω ή κάτω της T_o	Κανονική
6	Ημερομηνίες αρχής ή τέλους περιόδου του έτους με θερμο- πάνω ή κάτω της τιμής T_o	Κανονική
7	Συχνότητες εμφάνισης τιμών κάτω της T_o ή άνω της T_1 ή μεταξύ T_o και T_1.	Poisson ή Διωνιμική ή σχεδόν κανονική για περίοδο πολλών ετών.
β	Μέσοι όροι (έτους ή μήνα ή δεκαήμερου) πολλών ετών	
1	Ορισμένης ώρας παρατήρησης	Κανονική
2	Ελάχιστη ή μέγιστη ημέρας	Κανονική
3	Μέση ημερήσια.	Κανονική
γ	Απόλυτες άκρες τιμές έτους ή μήνα ή δεκαήμερου, ορισμένης ώρας για όλη την ημέρα.	Gumbell ή άλλη κατανομή άκρων τιμών
2	Πίεση στο σταθμό ή στην επιφάνεια της θάλασσας, ή ύψος ισοβαρικών επιφανειών.	Περίπου ή ίδια με την κατανομή της θερμοκρασίας, ανάλογα με τις εξετα ζόμενες τιμές (ημερήσιες, μηνιαίες άκρες τιμές κλπ.
3	Αριθμός ημερών φαινομένων, όπως καταιγίδα, χιόνι, ομίχλη όμβρο, πάχνη, παγετό κλπ, ανά έτος ή μήνα ή και δεκαήμερο.	Διωνυμική ή Poisson ή αρνητική διωνυμική
4	Νέφωση σύνολο τιμών ή για ξεχωριστά για κάθε ώρα ή για κάθε μήνα ή για κάθε δεκαήμερο.	Πολυωνιμική ή διωνυμική για την περίπτωση εξέτασης νέφωσης πάνω ή κάτω από ορισμένη τιμή.
5	Βροχή ή χιόνι	
1	Ημέρες βροχής ή χιονιού	Αρνητική διωνυμική.
2	Συχνότητα ημερών ισχυρών βροχοπτώσεων	Poisson ή αρνητική διωνυμική.
3	Μέγιστες εντάσεις ορισμένης διάρκειας	Gumbell ή διπλή αρνητική εκθετική
4	Αθροισμα (ετήσιο ή μηνιαίο ή δεκαήμερο)	Γάμα κατανομή, τείνουσα στην κανονική για πολλά έτη ετήσιων αθροισματων.

Πίνακας 6.8. Κατανομές των βασικών μετεωρολογικών παραμέτρων. (Συνέχεια του πίνακα).

α/α		Παράμετρος	Κατανομή συχνότητας (πλέον πιθανή)
6		Ηλιοφάνεια και ακτινοβολία	
	1	Ετήσια αθροίσματα πολλών ετών	Γάμα, τείνουσα στην κανονική όταν διατίθενται αρκετά χρόνια.
	2	Μηνιαία-δεκαήμερα αθροίσματα	Γάμα.
	3	Ημερήσιες τιμές έτους ή μήνα ή δεκαήμερου	Γάμα ή πολυωνιμική πολυωνιμική για τις ημερήσιες τιμές του μήνα ή δεκαήμερου
7		Ανεμος	
	1	Μέση ταχύτητα	Weibull ή Γάμα
	2	Μέγιστος άνεμος (ταχύτητα)	Κατανομή άκρων τιμών (πχ Gumbell)
8		Υγρασία (μέση ετήσια ή μηνιαία ή δεκαήμερη)	Σχεδόν κανονική (μικρή θετική ασσυμετρία)

Πίνακας 6.9 Μέσες τιμές της ελάχιστης θερμοκρασίας στη Λάρισσα την τριακονταετία 1961 έως 1990, ανά μήνα και συνολικά για το έτος. Στην τελευταία σειρά του πίνακα δίνονται οι μέσες τιμές της περιόδου 1961-1990 (normal).

Ετ.	Ιαν.	Φεβ.	Μαρ.	Απρ.	Μαϊ.	Ιουν	Ιουλ	Αυγ.	Σεπ.	Οκτ.	Νοε.	Δεκ.	E_{365}	E_{12}
61	1.5	.2	3.5	7.8	10.9	15.6	17.3	16.9	13.2	10.4	6.8	2.1	8.88	8.83
62	2.0	-.2	5.2	5.0	10.5	15.2	18.3	18.2	15.7	12.2	9.2	1.6	9.47	9.41
63	.2	2.8	2.2	7.2	11.6	14.9	18.5	17.6	14.7	11.7	6.6	3.1	9.31	9.27
64	-2.2	1.0	5.4	5.4	10.3	15.1	16.6	15.7	13.3	11.0	6.9	3.0	8.48	8.45
65	2.1	-.8	2.3	6.2	9.3	14.1	17.9	16.0	12.9	7.6	3.8	1.3	7.78	7.72
66	-1.1	2.9	3.4	6.9	9.3	15.3	17.3	19.1	14.3	13.9	7.1	2.9	9.31	9.27
67	-1.1	.0	2.2	6.2	11.0	13.7	17.6	18.4	15.0	11.1	6.0	2.0	8.55	8.49
68	-4.4	3.0	1.3	6.8	13.2	15.6	18.4	16.8	14.7	9.5	8.6	3.1	8.87	8.87
69	1.2	3.3	5.5	5.0	12.1	16.4	17.1	17.4	15.9	8.3	5.0	4.0	9.32	9.28
70	4.1	2.4	3.9	6.7	10.5	14.7	17.7	17.6	14.1	8.6	3.7	.1	8.71	8.67
71	4.2	1.2	1.8	6.2	10.6	15.0	15.8	17.3	13.1	7.5	4.2	-.1	8.11	8.07
72	3.8	3.6	2.4	8.1	10.5	15.3	18.0	16.8	14.3	7.6	3.8	2.3	8.89	8.88
73	1.3	1.9	2.0	6.2	11.4	16.3	18.7	17.2	15.7	10.4	3.0	1.6	8.87	8.83
74	2.1	3.2	4.5	6.0	9.8	15.2	16.8	17.6	13.7	10.3	5.9	-.6	8.73	8.70
75	-1.2	.0	3.7	6.9	12.1	15.4	17.9	16.6	14.8	9.2	6.5	.9	8.60	8.55
76	-.9	1.3	4.4	6.7	11.4	14.7	16.9	15.2	12.4	11.2	6.2	1.6	8.44	8.42
77	.6	2.6	2.2	4.8	11.3	14.5	18.5	17.2	13.6	6.8	5.9	.6	8.24	8.21
78	.6	3.0	3.7	5.8	9.9	13.9	17.2	15.3	12.7	9.2	1.7	3.8	8.11	8.08
79	-.5	3.7	4.6	6.1	11.1	15.6	17.8	17.5	13.1	10.2	7.1	2.1	9.06	9.03
80	.5	1.6	4.0	5.5	9.7	14.2	16.7	17.8	14.0	10.8	7.4	1.0	8.62	8.59
81	-1.5	.8	4.7	5.3	8.7	15.7	17.2	16.1	13.9	11.2	1.6	2.7	8.08	8.03
82	.7	-.1	2.5	6.0	10.0	15.0	16.3	17.0	14.2	11.7	3.8	3.0	8.39	8.33
83	-1.6	-1.5	1.4	6.4	11.1	14.6	18.0	16.4	13.0	8.2	6.1	2.4	7.94	7.89
84	.8	4.4	4.3	6.1	10.2	13.5	15.4	16.3	13.5	11.3	7.0	2.6	8.79	8.77
85	2.5	-2.0	5.1	7.6	13.1	15.2	17.4	17.3	13.2	7.9	7.3	3.6	9.08	9.01
86	1.4	3.1	5.6	5.4	11.5	15.3	17.4	17.9	13.9	9.6	4.4	-1.7	8.70	8.66
87	2.1	3.7	.2	6.4	9.8	14.9	18.7	17.4	15.4	10.9	5.7	2.9	9.04	9.01
88	4.3	1.1	2.4	6.0	11.0	15.6	20.6	17.7	15.1	9.2	2.0	-1.2	8.68	8.64
89	-2.3	.4	4.9	6.9	10.2	14.8	18.6	17.1	14.9	7.8	5.3	.9	8.32	8.28
90	-2.3	1.0	2.1	7.6	10.9	16.1	18.9	16.5	13.8	10.0	7.1	4.1	8.85	8.81
E	.6	1.6	3.4	6.3	10.8	15.0	17.7	17.1	14.1	9.8	5.5	1.9	8.67	8.64

Πίνακας 7.1 Μετεωρολογικοί παράμετροι. Κωδικοί αριθμοί (α/α), κωδικοί του τρόπου προέλευσης των τιμών, πλήθος N των διαθέσιμων τιμών ανά ημέρα, μονάδες μέτρησης, ακρίβεια και προτεινόμενος συμβολισμός τους.

α/α	Μετεωρολογική παράμετρος X	Συμβολι-λισμός	Συνηθεις τιμές του N	Προέ-λευση	Μονάδες μετρ.	Ακρί-βεια
1.1	Θερμοκαρασία αέρα (T)	T_a	3,6,8,24,48,N	1,3	°C	0.1
1.2	T υγρού θερμομέτρου	T_w	3,6,8,24,48	1	>>	>>
1.3	Σημείο δρόσου	DP	3,6,8,24,48	4	>>	>>
1.4	T γυμνού εδάφους	T_e	3,4,8,N	1,3	>>	>>
1.5	T σε διάφορα βάθη dd(cm)	T_{dd}	3,4,8,N	1,3	>>	>>
1.6	T επιφάνειας θαλάσσης	$T_\bullet$	1,2,3,8	1	>>	>>
1.7	T >> λιμνών	T_λ	1,2,3,8	1	>>	>>
1.8	Μέγιστη T αέρα	T_a	1,2 (κυρίως 1)	1,3	>>	>>
1.9	Ελάχιστη T αέρα	T_a	1,2 (κυρίως 1)	1,3	>>	>>
1.10	T αέρα σταθερών ισοβαρι-κών επιφανειών ddd(hPa)	T_{ddd}	1,2	1	>>	>> >>
1.11	T υγρού θερμομέτρου των ισοβαρικών επιφανειών.	$T_{w.ddd}$	1,2	1	>>	>> >>
1.12	Αλλες T (Δυναμική κλπ)			2	>>	>>
2.1	Πίεση (Ατμοσ.) σταθμού(P)	$P_{\sigma\tau}$	3,6,8,24,48,N	1,3	hPa	>>
2.2	P στην μέση σταθμη θαλασ.	$P_{\theta\alpha}$	3,6,8,24,48,N	4	>>	>>
2.3	Ελάχιστη P (θάλασσας)	$P_{\theta\alpha}$	1	3,4	>>	>>
2.4	Μέγιστη P (θάλασσας)	$P_{\theta\alpha}$	1	3,4	>>	>>
2.5	Υψος ισοβ. επιφ. ddd(hPa)	Y_{ddd}	1,2	4	gm	10
3.1	Σχετική υγρασία	RH	3,6,8,24,48,N	1,3,4	%	1
3.2	Ελάχιστη σχετική υγρασία	RH	1	3,4	>>	>>
3.3	Μέγιστη σχετική υγρασία	RH	1	3,4	>>	>>
3.4	Απόλυτη υγρασία	ABH	3,6,8,24	4	gr/m^3	0.01
3.5	Μερική τάση των υδρατμών	e_w	3,6,8,24	4	hPa	0.1
4.1	Εξάτμιση Pich	PICHE	2	1	mm	0.1
4.2	Εξάτμιση Pan	PAN	2	1	>>	>>
4.3	Δυναμική εξάτμιση	EVD	1,2,4,8	4	>>	>>
4.4	Δυναμική εξατμισοδιαπνοή	EVTD	1,2,4,8	4	>>	>>
4.5	Πραγματική >>	REEV	1,2,4,8	4	>>	>>
5.1	Ολική Νέφωση	CL	3,6,8,24,48	2,5	όγδοα	1
5.2	Χαμηλή >>	CLX	3,6,8,24,48	2,5	>>	>>
5.3	Υψηλή >>	CLY	3,6,8,24,48	2,5		>>
5.4	Cumulonimbus	CLCB	3,6,8,24,48	2,5		>>
5.5	Υπόλοιπες νεφώσεις	CLAL	3,6,8,24,48	2,5		>>
6.1	Ηλιοφάνεια	HL ή n	1,2,24	3	ώρες	0.1
6.2	Κλάσμα ηλιοφάνειας	HLR	1	4		0.1
7.1	Ολική Ηλιακή ακτινοβολία	GL ή RA	1,2,24	3	$WM^{-2}h^{-1}$	0.1
7.2	Σχετική ολική GL	GLR	1,2,24	4		>>
7.3	Αμεση ηλιακή ακτινοβολία	R_a	1,2,24	3	>>	>>
7.4	R_a σχετική	RR_a	1,2,24	4		>>

Πίνακας 7.1 Συνέχεια του πίνακα από την προηγούμενη σελίδα.

α/α	Μετεωρολογική παράμετρος X	Συμβολισμός	Σύνηθεις τιμές του N	Προέλευση	Μονάδες μετρ.	Ακρίβεια
7.5	Διάχυτη ηλιακή ακτινοβ.	R_d	1,2,24	4	>>	>>
8	Ορατότητα (οριζόντια)	VV	3,6,8,24,48	2	m (Km)	
9.1 9.2 9.3	Υψος βροχής Διάρκεια βροχής Ενταση βροχής	BP BPΔ BPE	2,4,8,N 2,N 2,N	1,3 6 4	mm	0.1
10	Υψος χιονιού	SN	1	1	mm	0.1
11.1 11.2 11.3 11.4 11.5 11.6 11.7	Ταχύτητα του ανέμου (FF) FF FF FF ισοβαρ. επιφ. ddd(hPa) Διεύθυνση (DD) του FF DD του FF_{ddd} Διαδρομή ανέμου	FFMS FFB FFK FF_{ddd} DD DD_{ddd} DIADR	8,24,48,N 8,24,N 8,24,48,N 2,1,4 8,24,48,N 2,1,4 2	1,2,3 2 1,3 1 1,2,3 1 1	ms^{-1} B Knots Knots ή μοίρες >> 100m	1 1 1 ms^{-1} 10° >> 1
12	Παράμετροι με τιμή 1 ή 0 ανάλογα με το αν την ώρα παρατήρησης συνέβη ή όχι κάποιο φαινόμενο ή γεγονός όπως: ομίχλη, πάχνη, δρόσος, αστραπή, βροντή, χιόνι, χαλάζη, όμβρος κλπ	Τέσσερα πρώτα γράμματα της λέξης πχ. DROS	3,6,8,24	2		
13	Παράμετροι με τιμή 1 ή 0 ανάλογα με το αν ολόκληρη την ημέρα της παρατήρησης συνέβη ή όχι κάποιο γεγονός ή φαινόμενο (από τα παραπάνω ή άλλο οριζόμενο από τον μελετητή, όπως άνεμος πάνω από 5Beaufort ή θερμοκρασία μικρότερη από 0°C κλπ).	Οπως παραπάνω και στην θέση του N των τιμών τίθεται 01	1	4		
14	Βαθμοί κάτω από την τιμή To ή πάνω από την τιμή To για την θερμοκρασία αέρα και την ελάχιστη-μέγιστη	T_{ato} ή T_{kto}	Οσες οι τιμές της To ή για την ελάχιστη ή την μέγιστη.	4	°C	0.1

Κωδικοί προέλευσης τιμών:

1. Οργανα άμεσης ανάγνωσης. 2. Εκτίμηση από τον παρατηρητή.
3. Ανάγνωση της τιμής από ταινία αυτογραφικού οργάνου.
4. Υπολογισμός ή εκτίμηση της τιμής από τις τιμές άλλων παραμέτρων.
5. Από τηλεμετρικά όργανα (RADAR, μετεωρολογικοί δορυφόροι).
6. Υπολογισμός που βασίζεται σε τιμές αυτογραφικών οργάνων.

Πίνακας 7.2 Ημερήσιες τιμές των μετεωρολογικών παραμέτρων που υπολογίζονται με βάση τις απλές τιμές των αντίστοιχων παραμέτρων του πίνακα 7.1. (Οι μονάδες και η ακρίβεια των τιμών είναι αυτή του παραπάνω πίνακα 7.1.) Δείκτης ή παράσταση που αντικαθιστά τον δείκτη N ή τον n (συμβολισμός) και αντίστοιχοι τύποι με βάση τους οποίους υπολογίζονται.

Μετεωρολογική παράμετρος X α/α πίνακα 7.1	Συμβολισμός και αντίστοιχος τύπος υπολογισμού.
1.1	E(7.5),E_8(7.6),E_3(7.8),$E_{2/2}$(7.7),E_{B3}(7.9),Γ(7.21) $T_{\alpha\tau o}$(7.17),$T_{\kappa\tau o}$(7.15),ΕΥΡ(7.10)
1.2 και 1.3	E(7.5),E_8(7.6),E_3(7.8),E_{B3}(7.9),MIN,MAX,ΕΥΡ(7.10)
1.4 και 1.5	E_3(7.8),E(7.5),MIN,MAX,ΕΥΡ(7.10)
1.6 και 1.7	E(7.5),E_3(7.8),E_8(7.6),MAX,MIN,ΕΥΡ(7.10)
1.8 και 1.9	$MM_{\alpha\tau o}$(7.20),$MM_{\kappa\tau o}$(7.19),Γ(7.21)
1.10 και 1.11	E(7.5) (υπάρχουν το πολύ δύο τιμές ανά ημέρα),ΕΥΡ(7.10)
1.12	E(7.5),E_3(7.8),E_8(7.6),MAX,MIN,ΕΥΡ(7.10)
2.1 και 2.2 2.5	Οπως 1.2 και 1.3. Οπως 1.10 και 1.11.
3.1, 3.4 και 3.5	Οπως 1.2 και 1.3
4.1 μέχρι 4.5	$\Sigma_{N/N}$(7.14)
5.1 μέχρι 5.5	E(7.5),E_8(7.6),Γ(7.21)
6.1	Διατίθεται κυρίως μία τιμή ανά ημέρα η $\Sigma_{N/N}$,Γ(7.21)
7.1, 7.3 και 7.5	Οπως η 6.1 και 6.2
8	Γ(7.21) (κυρίως ομίχλη ή ορατότητα κάτω από VV_o)
9.1 9.2 9.3	$\Sigma_{2/2}$(7.14),$\Sigma_{N/N}$(7.14),Γ(7.21) $\Sigma_{2/2}$(7.14),$\Sigma_{N/N}$ MAX
10	$\Sigma_{2/2}$(7.14)
11.1 μέχρι 11.3	E(8.5),MAX,Γ(7.21), και μέσο διανυσματικό άνεμο από τις E(8.12),E(8.14)
12	Γ(7.21)
14	$\Sigma_{N/N}$

Πίνακας 7.3 Μηνιαίες ή δεκαήμερες τιμές των μετεωρολογικών παραμέτρων που υπολογίζονται με βάση τις ημερήσιες τιμές των αντίστοιχων παραμέτρων των πινάκων 7.1 και 7.2. Οι μονάδες και η ακρίβεια των τιμών είναι αυτή που δίνεται στον πίνακα 7.1. Συμβολισμός (δείκτης ή παράσταση) που αντικαθιστά τον δείκτη J ή τον j, των ημερών του μήνα ή του δεκαήμερου και αντίστοιχος συμβολισμός του τρόπου υπολογισμού των χρησιμοποιούμενων ημερήσιων τιμών (σε παράνθεση). Οι μηνιαίες-δεκαήμερες τιμές υπολογίζονται για κάθε έτος ξεχωριστά, δηλαδή με βάση τις J∈{28,29,30,31} ή τις J∈{8,9,10,11} ημερήσιες τιμές του μήνα ή του δεκαήμερου αντίστοιχα.

Μετεωρολογική παράμετρος X α/α πίνακα 7.1	Συμβολισμός του τρόπου υπολογισμού που αντικαθιστά τον δείκτη J και σε παρένθεση παράσταση που δείχει το είδος των ημερήσιων τιμών (πίνακες 7.1 και 7.2).
1.1	$E(n,E,E_8,E_3,E_{2/2},E_{B3})$, ΓF(Γ), $\Sigma(T_{\alpha T o}, T_{\kappa T o})$, MIN(ΕΥΡ), MAX(ΕΥΡ) και E,MIN,MAX,ΕΥΡ,Γ κάθε παρατήρησης ξεχωριστά.
1.2 και 1.3	$E(n,E,E_8,E_3,E_{B3})$, MIN, MAX, ΕΥΡ
1.4 και 1.5	$E(n,E_3,E)$,MIN,MAX,ΕΥΡ,ΓF(Γ)
1.6 και 1.7	$E(n,E,E_3)$,MAX,MIN,ΕΥΡ
1.8 και 1.9	$E,\Sigma(MMT_{\alpha T o}),\Sigma(MMT_{\kappa T o})$,ΓF(Γ),MIN,MAX
1.10 και 1.11	E(n,E) MAX,MIN,ΕΥΡ
1.12	$E(n,E,E_3,E_8)$,MAX,MIN,ΕΥΡ
2.1 και 2.2	Οπως 1.2 και 1.3.
2.3 και 2.4	E,MIN,MAX
2.5	Οπως 1.10 και 1.11.
3.1, 3.4 και 3.5	Οπως 1.2 και 1.3
4.1 μέχρι 4.5	$\Sigma(\Sigma_{N/N})$,MAX
5.1 μέχρι 5.5	E(E),ΓF (Γ={ημέρες με νέφωση 0/8 ή με νέφωση 8/8})
6.1	$\Sigma(\Sigma_{N/N})$,ΓF (Γ={ημέρα με HΛ=η=0 ή με HLR≥α και α∈[0,1]
7.1, 7.3 και 7.5	Οπως η 6.1 και 6.2
8	ΓF (κυρίως ομίχλη ή ορατότητα κάτω από VV_o=1000m)
9.1	$\Sigma(\Sigma_{2/2},\Sigma_{N/N})$,ΓF
9.2	$\Sigma(\Sigma_{2/2},\Sigma_{N/N})$
9.3	MAX (διάρκειας 5,10,15,20,30,60,120,720min)
10	$\Sigma(\Sigma_{2/2})$,ΓF (Γ={χιόνισε την μελετούμενη ημέρα)
11.1 μέχρι 11.3	$E(E,E_3,E_8)$,MAX,ΓF
12	ΓF (Ορισμένης ώρας και στο σύνολο της ημέρας)
14	$\Sigma(\Sigma_{N/N})$

Πίνακας 7.4 Ετήσιες τιμές των μετεωρολογικών παραμέτρων που υπολογίζονται με βάση τις ημερήσιες τιμές των αντίστοιχων παραμέτρων των πινάκων 7.1 και 7.2. Οι μονάδες και η ακρίβεια των τιμών είναι αυτή που δίνεται στον πίνακα 7.1. Οι ετήσιες τιμές υπολογίζονται για κάθε έτος ξεχωριστά, δηλαδή με βάση τις Λ∈{365,366} ημερήσιες τιμές του έτους υπολογισμών.

Μετεωρολογική παράμετρος X α/α πίνακα 7.1	Συμβολισμός του τρόπου υπολογισμού που αντικαθιστά τον δείκτη M και σε παρένθεση παράσταση που δείχει το είδος των ημερήσιων τιμών (πίνακες 7.1 και 7.2). (J=D)
1.1	$E(n,E,E_8,E_3,E_{2/2},E_{B3})$, ΓF, $\Sigma(T_{\alpha\tau o}, T_{\kappa\tau o})$, MIN(ΕΥΡ), MAX(ΕΥΡ) και E,MIN,MAX,ΕΥΡ,Γ κάθε παρατήρησης ξεχωριστά.
1.2 και 1.3	$E(n,E,E_8,E_3,E_{B3})$, MIN, MAX, ΕΥΡ
1.4 και 1.5	$E(n,E_3,E)$,MIN,MAX,ΕΥΡ,ΓF(Γ)
1.6 και 1.7	$E(n,E,E_3)$,MAX,MIN,ΕΥΡ
1.8 και 1.9	E,$\Sigma(MMT_{\alpha\tau o})$,$\Sigma(MMT_{\kappa\tau o})$,ΓF, MIN,MAX
1.10 και 1.11	E(n,E) MAX,MIN,ΕΥΡ
1.12	$E(n,E,E_3,E_8)$,MAX,MIN,ΕΥΡ
2.1 και 2.2	Οπως 1.2 και 1.3.
2.3 και 2.4	E,MIN,MAX
2.5	Οπως 1.10 και 1.11.
3.1, 3.4 και 3.5	Οπως 1.2 και 1.3
4.1 μέχρι 4.5	$\Sigma(\Sigma_{N/N})$,MAX
5.1 μέχρι 5.5	E(E),ΓF (Γ={ημέρες με νέφωση 0/8 ή με νέφωση 8/8})
6.1	$\Sigma(\Sigma_{N/N})$,ΓF (Γ={ημέρα με ΗΛ=η=0 ή με HLR≥α και α∈[0,1]
7.1, 7.3 και 7.5	Οπως η 6.1 και 6.2
8	ΓF (κυρίως ομίχλη ή ορατότητα κάτω από VV_o=1000m)
9.1	$\Sigma(\Sigma_{2/2},\Sigma_{N/N})$,ΓF
9.2	$\Sigma(\Sigma_{2/2},\Sigma_{N/N})$
9.3	MAX (διάρκειας 5,10,15,20,30,60,120,720min)
10	$\Sigma(\Sigma_{2/2})$,ΓF (Γ={χιόνισε την μελετούμενη ημέρα)
11.1 μέχρι 11.3	$E(E,E_3,E_8)$,MAX,ΓF
12	ΓF (ορισμένης ώρας και στο σύνολο της ημέρας)
14	$\Sigma(\Sigma_{N/N})$