ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΝΑΥΠΗΓΩΝ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
Δ.Π.Μ.Σ: ΝΑΥΤΙΚΗ ΚΑΙ ΘΑΛΑΣΣΙΑ ΤΕΧΝΟΛΟΓΙΑ ΚΑΙ ΕΠΙΣΤΗΜΗ

Διπλωματική Εργασία:

ΣΤΑΤΙΣΤΙΚΗ ΑΚΡΑΙΩΝ ΤΙΜΩΝ. ΕΦΑΡΜΟΓΗ ΣΕ ΜΕΤΡΗΣΕΙΣ ΤΕΧΝΙΚΩΝ ΜΕΓΕΘΩΝ.

Επιμέλεια: **Χρήστος Σ. Γρηγορόπουλος**
Επιβλέπων Καθηγητής: **Ευάγγελος Π. Χίνης, Αναπλ. Καθηγητής Ε.Μ.Π.**
Σχολή Μηχανολόγων Μηχανικών
Τομέας Πυρηνικής Τεχνολογίας

Αθήνα 2017

ΠΕΡΙΛΗΨΗ

Η στατιστική ακραίων τιμών έχει σαν στόχο την εκτίμηση της πιθανότητας να συμβούν σπάνια γεγονότα τα οποία εντοπίζονται στις «ουρές» μιας κατανομής. Αυτή η πιθανότητα γίνεται εξαιρετικά δύσκολο να προσδιορισθεί σε περίπτωση που έχουμε στη διάθεση μας λίγες μετρήσεις, εξαιτίας κόστους ή δυσκολίας στη μέτρηση. Η παρούσα διπλωματική εργασία αποσκοπεί στο να περιγράψει πειραματικά μία μέθοδο στατιστικής επεξεργασίας η οποία θα δίνει αξιόπιστα αποτελέσματα εκτίμησης αστοχιών με τα λιγότερα σε πλήθος αναγκαία πειραματικά δεδομένα.

Χωρίζεται σε δύο μέρη, στο πρώτο το οποίο είναι το βιβλιογραφικό και περιλαμβάνει πολύ συνοπτικά την Θεωρία Ακραίων Τιμών και μεθόδους ανάλυσης όπως οι Block Maxima και P.O.T. και στο δεύτερο μέρος το πειραματικό, το οποίο περιλαμβάνει το πείραμα και την διαδικασία εκτίμησης της κατανομής ακραίων τιμών.

Εκτός των κλασικών μεθόδων, όπως οι Block Maxima και P.O.T υπάρχουν και άλλοι τρόποι εκτίμησης της κατανομής σε ένα πρόβλημα ακραίων τιμών, όπως με το διάγραμμα πιθανοτήτων ακραίων τιμών.

Η χρήση του ειδικού διαγράμματος πιθανότητας ως εύκολο και γρήγορο γραφικό εργαλείο στην στατιστική ανάλυση κατανομών είναι μια γενικώς αποδεκτή τεχνική. Όταν χρησιμοποιείται ένα μικρό δείγμα σε ένα πρόβλημα πιθανοτήτων, προτού η διαδικασία ολοκληρωθεί και τρέξει σε ένα πρόγραμμα στον υπολογιστή μπορεί να γίνει μία γρήγορη εκτίμηση του αποτελέσματος με αυτόν τον τρόπο.

Οι γραφικές μέθοδοι είναι σημαντικές και χρησιμοποιούνται από μηχανικούς και επιστήμονες σε καθημερινή βάση. Λόγω της απλής εφαρμογής τους και της εύκολης χρήσης τους, τα ειδικά διαγράμματα πιθανοτήτων χρησιμοποιούνται εδώ και πολλά χρόνια σε στατιστικές μελέτες. Πολύ πρακτική είναι αυτή η διαδικασία σε νέα προβλήματα, για τα οποία δεν υπάρχουν αρκετές πληροφορίες. Γενικά, στα στατιστικά

προβλήματα ακραίων τιμών ο στόχος είναι να αποκαλυφθεί η άγνωστη κατανομή, η οποία θα βοηθήσει στην αποφυγή πολλών δοκιμών και λαθών.[1] Στο πείραμά μας θα χρησιμοποιήσουμε το ειδικό διάγραμμα πιθανοτήτων ακραίων τιμών.

Περιεχόμενα

Α ΒΙΒΛΙΟΓΡΑΦΙΚΟ ΜΕΡΟΣ

1. ΘΕΩΡΙΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

1.1 ΕΙΣΑΓΩΓΗ

Το γεγονός ότι σε κάποιες εφαρμογές, προβλήματα ή μελέτες πρέπει να υπολογίζονται οι ακραίες τιμές διότι αυτές αντικατοπτρίζουν τις ακραίες συνθήκες που πρέπει να προσεγγίσουμε, απαιτεί την επικέντρωση της προσοχής πάνω στις ακραίες τιμές περισσότερο από ότι στις μέσες. Ο υπολογισμός των ακραίων τιμών διαφόρων μεγεθών μέσω της κατανομής που αυτές ακολουθούν αποτελεί τη Θεωρία Ακραίων Τιμών.

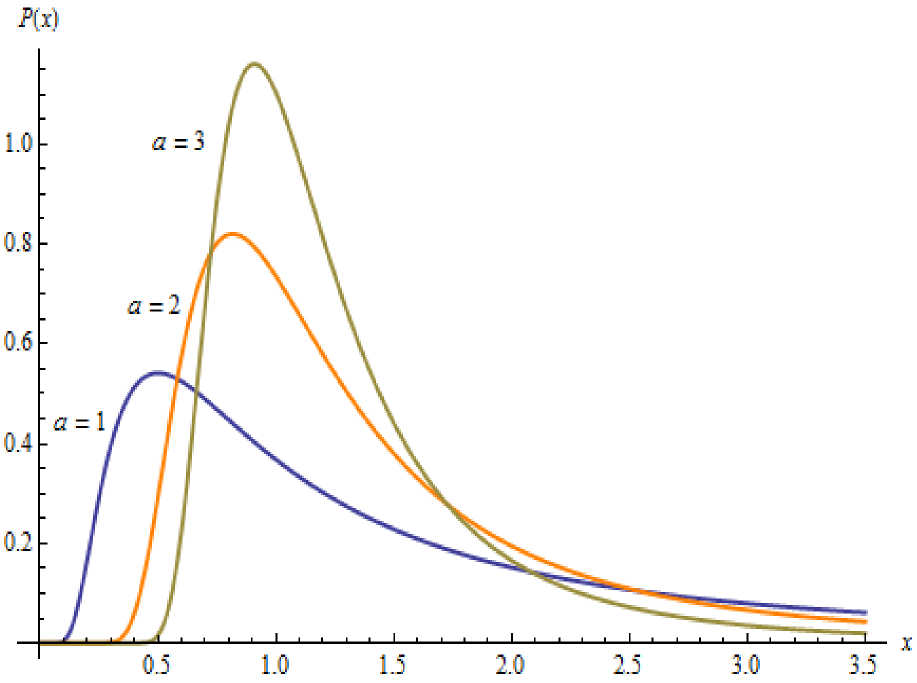
1.2 ΤΟ ΘΕΩΡΗΜΑ FISHER – TIPPETT

Έστω ότι έχουμε κάποιες τυχαίες μεταβλητές $X_1, X_2, \ldots, X_\nu$ και θέλουμε να μελετήσουμε τη μορφή της δεξιάς ουράς της F από την οποία προέρχονται οι παραπάνω τιμές. Έστω επίσης όμως είναι άγνωστη η F.
Το θεώρημα Fisher-Tippett έρχεται να λύσει αυτό το πρόβλημα, θεωρώντας ότι βασιζόμαστε στην οριακή κατανομή K της μέγιστης παρατήρησης των $X_1, X_2, \ldots, X_\nu$.

Η K θα είναι μία από τις τρείς εξής κατανομές [3]:

1. ***Κατανομή τύπου Frechet:***

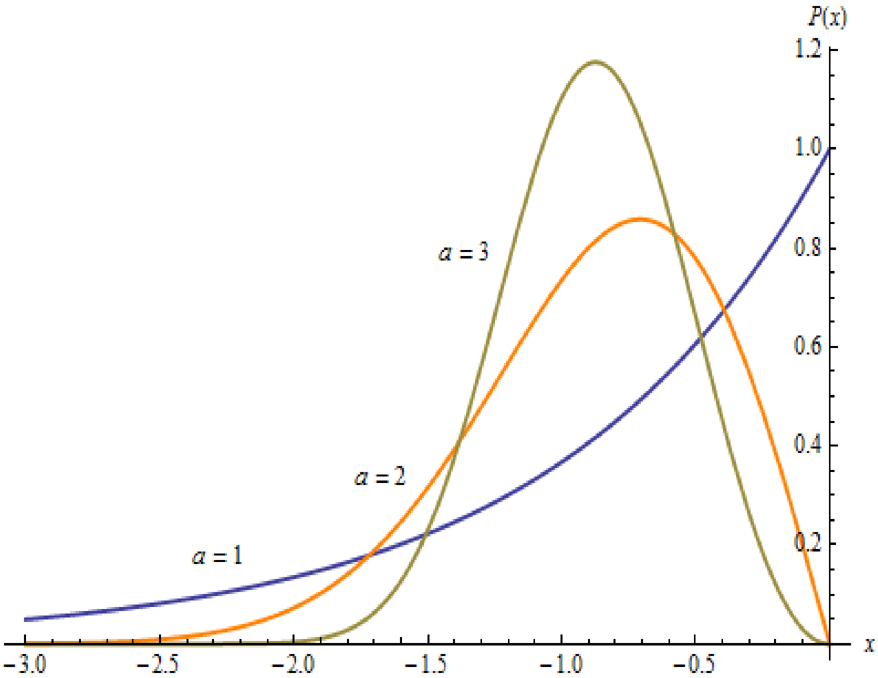
$$\boldsymbol{\Phi_\alpha(x)} = \begin{cases} \mathbf{0}, & \boldsymbol{x \le 0} \\ \boldsymbol{e^{-x^{-a}}}, & \boldsymbol{x > 0} \end{cases}, \qquad \boldsymbol{a > 0}$$



Σχήμα 1: Διάγραμμα συνάρτησης πυκνότητας πιθανότητας κατανομής τύπου Frechet για διάφορες τιμές του δείκτη α

2. ***<u>Κατανομή τύπου Weibull:</u>***

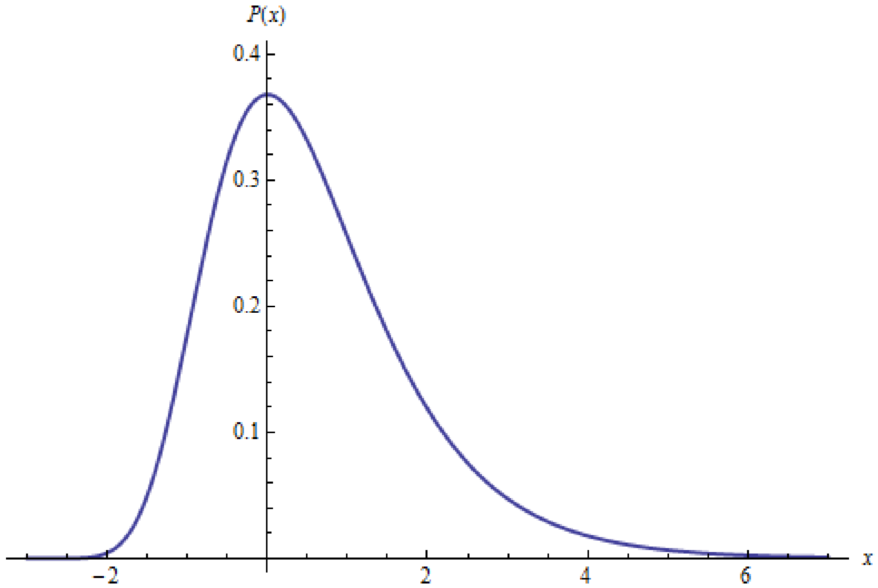
$$\boldsymbol{\Psi_{\alpha}(x)} = \begin{cases} \boldsymbol{e^{-(-x)^{a}}}, & \boldsymbol{x \leq 0} \\ \boldsymbol{1}, & \boldsymbol{x > 0} \end{cases}, \qquad \boldsymbol{a > 0}$$



Σχήμα 2: Διάγραμμα συνάρτησης πυκνότητας πιθανότητας κατανομής τύπου Weibull για διάφορες τιμές του δείκτη α

3. ***<u>Κατανομή τύπου Gumbel:</u>***

$$\boldsymbol{\Lambda(x) = e^{-e^{-x}}, x \in R}$$



Σχήμα 3: Διάγραμμα συνάρτησης πυκνότητας πιθανότητας κατανομής τύπου Gumbel

1.3 ΜΕΘΟΔΟΙ ΑΝΑΛΥΣΗΣ ΑΚΡΑΙΩΝ ΠΑΡΑΤΗΡΗΣΕΩΝ

Υπάρχουν διάφορες μέθοδοι ανάλυσης ακραίων παρατηρήσεων. Οι μέθοδοι που συναντάμε πιο συχνά είναι δύο. Η μέθοδος Block Maxima στην οποία στηριζόμαστε στις μέγιστες τιμές ανά χρονικές περιόδους και η μέθοδος Peaks Over Threshold στην οποία στηριζόμαστε στις τιμές πάνω από ένα κατώφλι.

1.3.1 ΜΕΘΟΔΟΣ BLOCK MAXIMA

Στην μέθοδο Block Maxima αρχικά χωρίζουμε τις τιμές n σε k υποσύνολα με m αριθμό, δηλαδή n=km. Από κάθε υποσύνολο επιλέγουμε τις μέγιστες τιμές οι οποίες ονομάζονται block maxima (minima αντίστοιχα οι ελάχιστες). Στη συνέχεια προσδιορίζουμε σε ποια από τις 3 κατανομές ακροτάτων ανήκουν τα block maxima και εκτιμούμε τις παραμέτρους. Συνήθως ακολουθείται μια διαφορετική αλλά ασφαλέστερη διαδικασία. Οι τρείς τύποι κατανομών ακροτάτων μετατρέπονται σε μία ενωμένη οικογένεια κατανομών η οποία ονομάζεται Γενικευμένη Κατανομή Ακραίων Τιμών ή GEV (Generalized Extreme Value distribution) . Έτσι θεωρούμε ότι τα block maxima ανήκουν στην GEV και εκτιμούμε τις παραμέτρους της. Η εκτίμηση των παραμέτρων μπορεί να γίνει με διάφορες μεθόδους όπως:

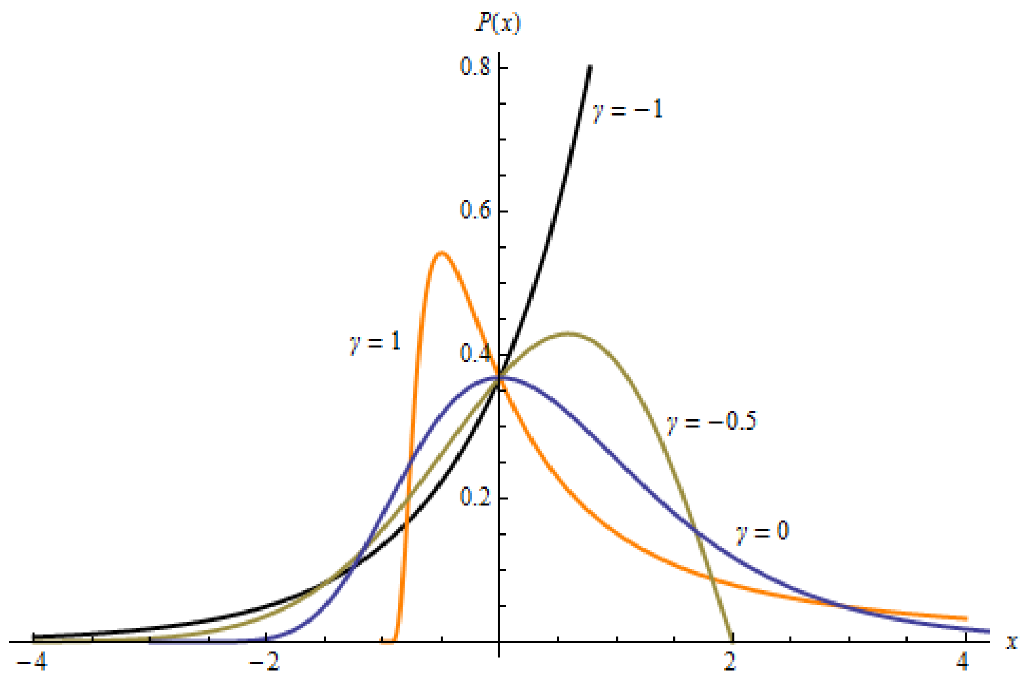
- Με τη μέθοδο των L-ροπών,
- Με τη μέθοδο της μέγιστης πιθανοφάνειας,
- Με διάφορα κατάλληλα γραφήματα [2]

1.3.2 ΓΕΝΙΚΕΥΜΕΝΗ ΚΑΤΑΝΟΜΗ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Η ΓΑΤ είναι μια οικογένεια κατανομών που αποτελείται από τις 3 ασυμπτωτικές κατανομές Gumbel, Frechet και Weibull, επίσης γνωστές σαν I, II και III ασυμπτωτικές κατανομές ακραίων τιμών. Η Θεωρία ακραίων τιμών περιλαμβάνει τον τρόπο με τον οποίο οι μέγιστες ή οι ελάχιστες τιμές από ένα πληθυσμό τιμών μπορούν να περιγραφούν από μία από τις 3 κατανομές ακραίων τιμών. Ο von Mises ένωσε τις 3 κατανομές σε μία συνδυασμένη και την ονόμασε GEV.

Ο τύπος της GEV είναι:

$$\mathbf{F(x;\mu,\sigma,\xi)} = \mathbf{exp}\left\{-\left[\mathbf{1}+\boldsymbol{\xi}\left(\frac{\boldsymbol{x}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}\right)\right]^{-\mathbf{1}/\boldsymbol{\xi}}\right\}$$



Σχήμα 4: Διάγραμμα συνάρτησης πυκνότητας πιθανότητας κατανομής τύπου Γ.Α.Τ.

1.3.3 ΜΕΘΟΔΟΣ PEAKS OVER THRESHOLD

Η μέθοδος αυτή μπορεί να χρησιμοποιηθεί για τον εντοπισμό της κατανομής των υπερβάσεων ενός συγκεκριμένου ορίου καθώς και για την εκτίμηση του σχήματος της ΄΄ουράς΄΄ της αρχικής κατανομής μιας μεταβλητής. Έστω οι τυχαίες μεταβλητές X_1, $X_{2,\ldots,}X_v$ με κοινή συνάρτηση κατανομής F και ν αρκετά μεγάλο. Η κατανομή των μέγιστων ακραίων τιμών μπορεί τότε να εκτιμηθεί από μία Γενικευμένη κατανομή Ακραίων Τιμών.[5] Η μέθοδος αυτή χρησιμοποιείται αρκετά συχνά αλλά μειονεκτεί στο γεγονός ότι εξαρτάται πολύ από την επιλογή που θα γίνει για το ΄΄κατώφλι΄΄- όριο.

1.4 ΟΙ ΜΕΘΟΔΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΤΩΝ ΜΟΝΤΕΛΩΝ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

1.4.1 Η ΜΕΘΟΔΟΣ ΤΩΝ L-ΡΟΠΩΝ

Η διαδικασία υπολογισμού των παραμέτρων με τη μέθοδο των L-ροπών είναι αντίστοιχη της μεθόδου των συνήθων ροπών. Οι *L*- Ροπές παρέχουν μέτρα της θέσης, της διασποράς, της ασυμμετρίας, της κύρτωσης και άλλων όψεων του σχήματος των κατανομών πιθανότητας ή των δειγμάτων, ωστόσο υπολογίζονται από γραμμικούς συνδυασμούς των ταξινομημένων τιμών των δεδομένων. Η μέθοδος αυτή χρησιμοποιείται αρκετά στην υδρολογία διότι παρουσιάζουν μικρότερη δειγματική μεταβλητότητα σε σύγκριση με τις κανονικές ροπές. Γενικά η μέθοδος αυτή απαιτεί μικρότερη διαδικασία υπολογισμού σε σχέση με άλλες μεθόδους όπως αυτής της Μέγιστης Πιθανοφάνειας και επίσης καταλήγει σε αξιόπιστα συμπεράσματα. [4]

1.4.2 Η ΜΕΘΟΔΟΣ ΤΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Στην μέθοδο των ελαχίστων τετραγώνων υπολογίζονται οι εκτιμήτριες με ελαχιστοποίηση ενός συγκεκριμένου αθροίσματος τετραγώνου. Η μέθοδος αυτή είναι χρήσιμη όταν πρέπει να γίνει υπολογισμός μιάς παραμέτρου με βάση ένα σύνολο παρατηρήσεων από ένα τυχαίο δείγμα, οι οποίες είναι συνδεδεμένες με την υπο εκτίμηση παράμετρο μέσω των μέσων τιμών. Γενικά, η μέθοδος των ελαχίστων τετραγώνων χρησιμοποιείται για την εκτίμηση των παραμέτρων της ανάλυσης παλινδρόμησης.

1.4.3 Η ΜΕΘΟΔΟΣ ΤΗΣ ΜΕΓΙΣΤΗΣ ΠΙΘΑΝΟΦΑΝΕΙΑΣ

Η μέθοδος της μέγιστης πιθανοφάνειας αφορά μια διαδικασία κατασκευής εκτιμητή για μία άγνωστη παράμετρο θ ή ακόμη γενικότερα για μία άγνωστη τιμή $g(\theta)$ ενός πειράματος. Με την χρησιμοποίηση του συγκεκριμένου πιθανοθεωρητικού μοντέλου όπως και κάποιων νόμων των πιθανοτήτων μπορεί να υπολογιστεί η πιθανότητα (πιθανοφάνεια) ενός ενδεχομένου που αναζητάμε. Γενικά η μέθοδος παράγει καλούς εκτιμητές ακόμα και για μεγάλο μέγεθος δείγματος.

Β. ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ

Σκοπός του πειραματικού μέρους είναι η μελέτη της ασυμπτωτικής συνάρτησης πιθανότητας (asymptotic probability function) των ελάχιστων ακραίων τιμών (smallest extreme values) που προέρχονται από ένα σχετικά μεγάλο πληθυσμό τιμών (200 τιμές) μιας τυχαίας μεταβλητής (random value). Σαν τυχαία μεταβλητή επιλέξαμε το μήκος βίδας Allen M4x30mm από χάλυβα, ποιότητας 8.8 DIN 912. Με άλλα λόγια είναι είναι η ανάλυση ενός δείγματος ακραίων τιμών για την πρόβλεψη νέων ακραίων τιμών από τον ίδιο πληθυσμό.

Όπως αναφέραμε και στο βιβλιογραφικό μέρος της διπλωματικής υπάρχουν τρείς σταθερές ασυμπτωτικές κατανομές ακραίων τιμών (συναρτήσεις πιθανότητας). Οι αρχικές κατανομές από τις οποίες προκύπτουν οι αντίστοιχες ασυμπτωτικές κατανομές είναι οι εξής τρείς: Εκθετικές, Cauchy και περιορισμένης μεταβλητής.

Η διαδικασία για τον προσδιορισμό της ασυμπτωτικής συνάρτησης πιθανότητας περιλαμβάνει τα εξής βήματα:

1) **Παίρνουμε τις μετρήσεις του (μηχανικού) χαρακτηριστικού από το δείγμα που θέλουμε να εξετάσουμε.**

2) **Από τις μετρήσεις της τυχαίας μεταβλητής-χαρακτηριστικού σχηματίζουμε πληθυσμούς τιμών ακολουθώντας την αρχική σειρά μέτρησεων.**

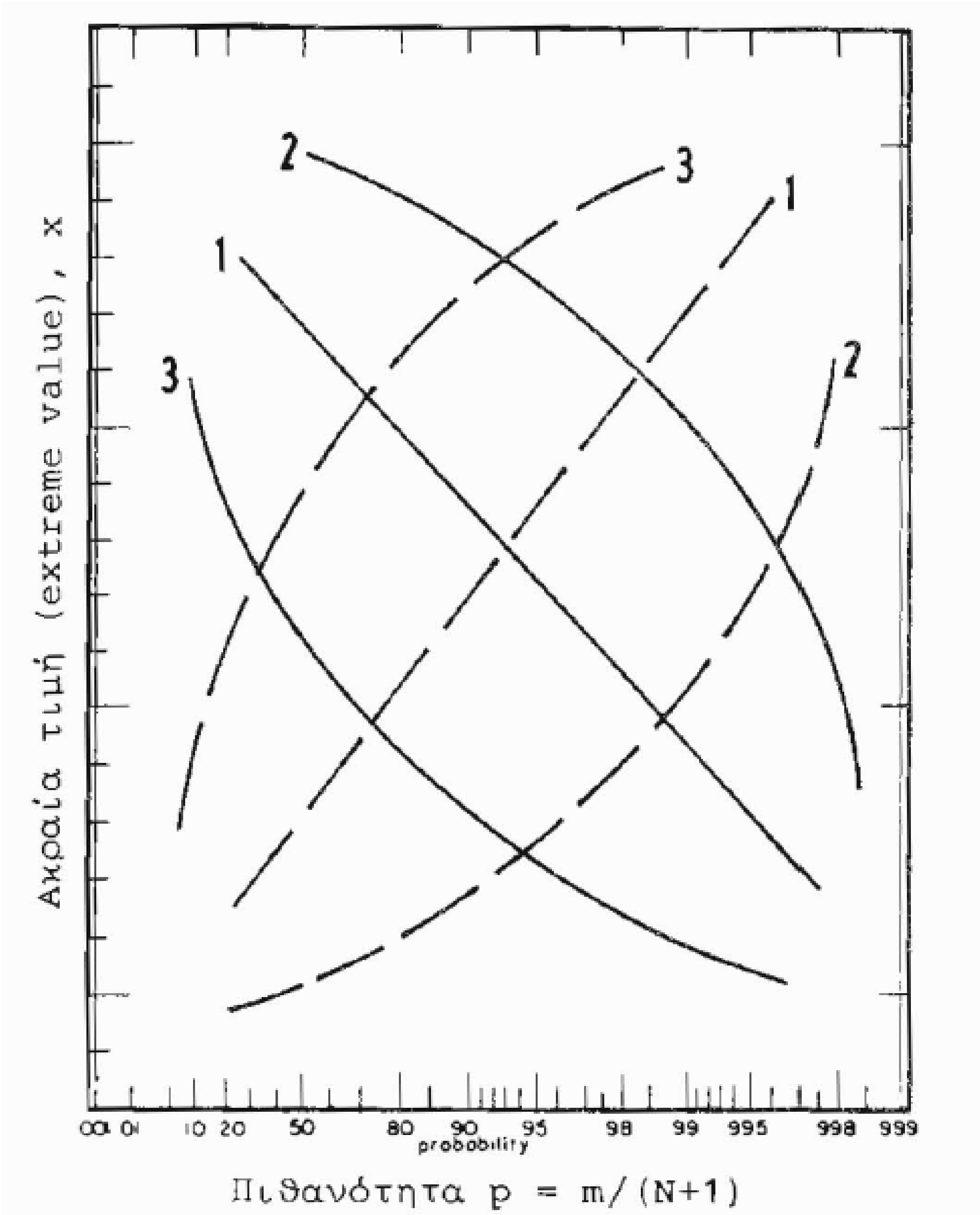
3) **Τον κάθε πληθυσμό τον διαιρούμε σε Ν υποπληθυσμούς-δείγματα με n τιμές της μεταβλητής μας για κάθε έναν.**

4) **Από κάθε υποπληθυσμό-δείγμα με n τιμές της μεταβλητής μας επιλέγεται η ελάχιστη τιμή X1 (ή αντίστοιχα η μέγιστη τιμή αν εξετάζουμε ασυμπτωτική συνάρτηση πιθανότητας μέγιστων ακραίων τιμών) επομένως από τα N δείγματα προκύπτουν N ελάχιστες τιμές (ή αντίστοιχα N μέγιστες τιμές).**

5) **Στην συνέχεια οι N ελάχιστες ακραίες τιμές X της τυχαίας μεταβλητής διατάσσονται κατά φθίνουσα τιμή και σε κάθε τιμή, δίνεται ένας αριθμός, σειράς (rank number) m, με m=1 για την μεγαλύτερη τιμή του X_1. Αν υπάρχουν περισσότερα από ένα δείγματα με την ίδια τιμή της μεταβλητής X_1, στην τιμή αυτή δίνεται η μέση τιμή των αντίστοιχων αριθμών σειράς m. Παράλληλα υπολογίζεται η πιθανότητα της ελάχιστης ακραίας τιμής για κάθε τιμή x του X_1 από την εξίσωση p=m/(N+1).**

6) **Έχοντας καταλήξει σε έναν πίνακα τιμών-σημείων (p,x) από το προηγούμενο βήμα, τοποθετούμε όλα τα σημεία σε διάγραμμα πιθανότητας ακραίων τιμών (extreme value probability paper). Στο διάγραμμα τοποθετούνται οι τιμές p στην οριζόντια κλίμακα πιθανότητας των ακραίων τιμών και x στην κατακόρυφη (γραμμική) κλίμακα. Τότε, η κατανομή των ελάχιστων ακραίων τιμών είναι η:**
   i) **Εκθετική ,**
   ii) **Cauchy,**
   iii) **Περιορισμένη**

   αν η γραμμή που διαγράφεται από τα σημεία είναι αντίστοιχα ευθεία, καμπύλη με τα κοίλα προς τα κάτω, καμπύλη με τα κοίλα προς τα πάνω. Για τις μέγιστες ακραίες τιμές ισχύουν οι διακεκομμένες γραμμές .[6]



Σχήμα 5: Καμπύλες πιθανότητας ακραίων τιμών. Με συνεχή γραμμή είναι οι καμπύλες ελαχίστων τιμών και με διακεκομμένη οι καμπύλες μεγίστων τιμών. Κατανομές: 1) Εκθετική , 2) Cauchy , 3) Περιορισμένη

Παρακάτω αναφέρονται κάποιοι ειδικοί περιορισμοί της ασυμπτωτικής θεωρίας των ακραίων τιμών [6]:

I. Ο πληθυσμός των τιμών του δείγματος από τον οποίο προέρχονται οι ακραίες τιμές, πρέπει να είναι σχετικά μεγάλος.

II. Η αρχική κατανομή, από την οποία προέρχονται οι ακραίες τιμές, πρέπει να ανήκει σε ένα από τους παραπάνω τρεις τύπους.

III. Οι τιμές-παρατηρήσεις, από τις οποίες προέρχονται οι ακραίες τιμές, θα πρέπει να είναι ανεξάρτητες μεταξύ τους.

IV. Η αρχική κατανομή από την οποία προέρχονται οι ακραίες τιμές και οι παράμετροί της πρέπει να παραμένουν σταθερές με το χρόνο (ή το χώρο) ή η επίδραση που ασκεί ο χρόνος (ή ο χώρος) στην τιμή τους πρέπει να λαμβάνεται υπόψη ή να εξαλείφεται.

V. Το δείγμα θα πρέπει να είναι αξιόπιστο και ο τρόπος με τον οποίο παίρνονται οι μετρήσεις να γίνεται κάτω από τις ίδιες ακριβώς συνθήκες, έτσι ώστε η αρχική κατανομή να παραμένει η ίδια για κάθε δείγμα. [7] , [8]

Για τη μέγιστη ακραία τιμή X_n ισχύει ο τύπος:

$$\Phi_i(x) = \Pi\{X_n \leq x\}X_n$$

και για την ελάχιστη ακραία τιμή X_1:

$$R_i(x) = 1 - F_i(x) = \Pi\{X_1 \geq x\}$$

όπου i δείκτης για τις τρεις ασυμπτωτικές κατανομές: Εκθετική (i=1), Cauchy (i=2) και Περιορισμένη (i=3).

B.1 ΔΙΕΞΑΓΩΓΗ ΠΕΙΡΑΜΑΤΟΣ

Βήμα 1: Στην συγκεκριμένη περίπτωση έχουμε πάρει 200 μετρήσεις του μήκους βίδας Allen M4x30mm από χάλυβα, ποιότητας 8.8 DIN 912

ΠΙΝΑΚΑΣ 3:Μέτρησεις μήκους βίδας Άλλεν Μ4x30 DIN 912.

A/A	L (mm)	A/A	L (mm)	A/A	L (mm)	A/A	L (mm)	A/A	L (mm)
1	29.34	41	29.37	81	29.43	121	29.49	161	29.39
2	29.31	42	29.34	82	29.44	122	29.50	162	29.54
3	29.45	43	29.33	83	29.34	123	29.45	163	29.50
4	29.38	44	29.35	84	29.48	124	29.53	164	29.44
5	29.51	45	29.44	85	29.33	125	29.45	165	29.42
6	29.50	46	29.45	86	29.51	126	29.42	166	29.34
7	29.32	47	29.50	87	29.52	127	29.39	167	29.39
8	29.53	48	29.42	88	29.43	128	29.42	168	29.58
9	29.38	49	29.45	89	29.53	129	29.52	169	29.56
10	29.54	50	29.30	90	29.45	130	29.53	170	29.37
11	29.40	51	29.50	91	29.36	131	29.29	171	29.44
12	29.54	52	29.38	92	29.46	132	29.42	172	29.37
13	29.49	53	29.43	93	29.50	133	29.40	173	29.46
14	29.50	54	29.31	94	29.60	134	29.63	174	29.39
15	29.42	55	29.29	95	29.42	135	29.40	175	29.40
16	29.24	56	29.42	96	29.36	136	29.35	176	29.52
17	29.37	57	29.28	97	29.37	137	29.60	177	29.37
18	29.52	58	29.46	98	29.18	138	29.57	178	29.56
19	29.50	59	29.47	99	29.51	139	29.42	179	29.51
20	29.31	60	29.46	100	29.57	140	29.31	180	29.40
21	29.40	61	29.50	101	29.40	141	29.40	181	29.42
22	29.42	62	29.31	102	29.41	142	29.31	182	29.50
23	29.57	63	29.33	103	29.44	143	29.42	183	29.53
24	29.43	64	29.43	104	29.53	144	29.51	184	29.36
25	29.33	65	29.39	105	29.47	145	29.38	185	29.41
26	29.36	66	29.48	106	29.53	146	29.34	186	29.48
27	29.32	67	29.50	107	29.48	147	29.37	187	29.51
28	29.53	68	29.52	108	29.38	148	29.52	188	29.48
29	29.32	69	29.54	109	29.54	149	29.49	189	29.57
30	29.39	70	29.44	110	29.35	150	29.46	190	29.61
31	29.47	71	29.47	111	29.49	151	29.31	191	29.32
32	29.51	72	29.42	112	29.40	152	29.42	192	29.38
33	29.45	73	29.35	113	29.42	153	29.55	193	29.44
34	29.42	74	29.46	114	29.53	154	29.47	194	29.37
35	29.36	75	29.43	115	29.39	155	29.50	195	29.52
36	29.41	76	29.42	116	29.49	156	29.67	196	29.28
37	29.59	77	29.37	117	29.42	157	29.43	197	29.36
38	29.32	78	29.45	118	29.51	158	29.36	198	29.54
39	29.21	79	29.41	119	29.31	159	29.38	199	29.41
40	29.55	80	29.54	120	29.17	160	29.45	200	29.45

Βήμα 2: Από τις 200 μετρήσεις που έγιναν, Πίνακας 3, σχηματίστηκαν 3 πληθυσμοί τιμών, ακολουθώντας την αρχική σειρά των μετρήσεων. Οι πληθυσμοί αυτοί αποτελούνται από Νπ: 60, 100 και 200 τιμές.

Βήμα 3: Κάθε πληθυσμός Νπ χωρίστηκε σε Ν δείγματα με μέγεθος n (n οι τιμές του X), το κάθε ένα (Νπ = Ν x n). Για παράδειγμα ο πληθυσμός Νπ=200 χωρίστηκε με 6 διαφορετικούς τρόπους. Στην $1^{η}$ περίπτωση χωρίστηκε σε 5 ομάδες των 40 δειγμάτων. Στην $2^{η}$ περίπτωση σε 8 ομάδες των 25 δειγμάτων, κ.ο.κ. Από τους παραπάνω πληθυσμούς Νπ προέκυψαν συνολικά 12 συνδυασμοί τιμών του Ν και n, Πίνακας 4.

ΠΙΝΑΚΑΣ 4: 12 συνδυασμοί τιμών του Ν και n

Νπ=Nxn=200		
Α/Α	N	n
1	5	40
2	8	25
3	10	20
4	20	10
5	25	8
6	40	5
Νπ=Nxn=100		
1	4	25
2	10	10
3	25	4
Νπ=Nxn=60		
1	5	12
2	6	10
3	12	5

Βήματα 4,5: Βρίσκουμε την ελάχιστη τιμή μήκους x_1 από κάθε δείγμα. Οι N ελάχιστες ακραίες τιμές x της τυχαίας μεταβλητής X_1 τοποθετούνται σε φθίνουσα τιμή και σε κάθε τιμή, δίνεται ένας αριθμός, σειράς (rank number) m, με m=1 για την μεγαλύτερη τιμή του X_1.

Αν υπάρχουν περισσότερα από ένα δείγματα με την ίδια τιμή της μεταβλητής X_1, στην τιμή αυτή δίνεται η μέση τιμή των αντίστοιχων αριθμών σειράς m. Η πιθανότητα της ελάχιστης ακραίας τιμής υπολογίζεται για κάθε τιμή x του X_1 από την εξίσωση $p=m/(N+1)$.

Από τα N δείγματα (από κάθε συνδυασμό τιμών N και n) προκύπτουν N ελάχιστες ακραίες τιμές (x_{11}, x_{12}, x_{13}, ...,x_{1N}) της τυχαίας μεταβλητής X_1. [6]

Ακολουθεί παρακάτω ο Πίνακας 5 ο οποίος αντιστοιχεί στον πέμπτο συνδυασμό (Α/Α=5) του Πίνακα 4 για Nπ=200. Δηλαδή από κάθε μία από τις 25 ομάδες των 8 τιμών πήραμε την ελάχιστη. Επομένως έχουμε έναν πίνακα με 25 ελάχιστες τιμές.

ΠΙΝΑΚΑΣ 5:

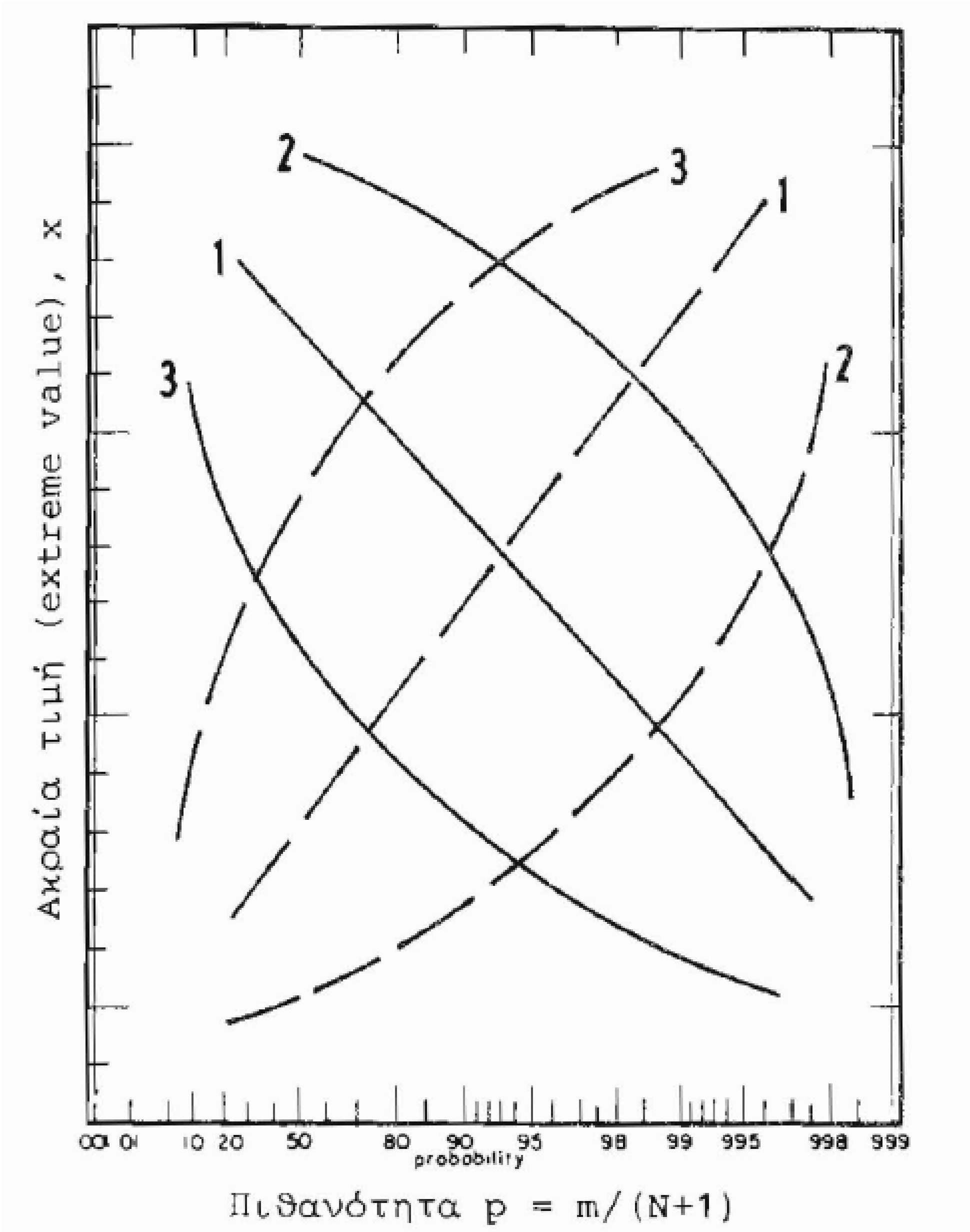
Διάταξη N=25 ελάχιστων ακραίων τιμών x του μήκους της βίδας (N_π=200, n=8) για τον υπολογισμό της πιθανότητας P= m/N +1.

Αρχικά δεδομένα		**Δεδομένα μετά τη διάταξη**		
Α/Α	**X**	**Σειρά m**	**x**	**P**
1	29.31	1	29.40	0.0384
2	29.24	2.5	29.39	0.0961
3	29.31	2.5	29.39	0.0961
4	29.32	3.5	29.37	0.1346
5	29.21	3.5	29.37	0.1346
6	29.33	7	29.36	0.2692
7	29.29	7	29.36	0.2692
8	29.28	7	29.36	0.2692
9	29.39	9.5	29.35	0.3653
10	29.35	9.5	29.35	0.3653
11	29.40	11	29.34	0.4230
12	29.36	12	29.33	0.4615
13	29.18	13.5	29.32	0.5192
14	29.35	13.5	29.32	0.5192
15	29.17	16.5	29.31	0.6346
16	29.39	16.5	29.31	0.6346
17	29.29	16.5	29.31	0.6346
18	29.31	16.5	29.31	0.6346
19	29.31	19.5	29.29	0.7307
20	29.34	19.5	29.29	0.7307
21	29.37	21	29.28	0.8076
22	29.37	22	29.24	0.8461
23	29.32	23	29.21	0.8846
24	29.36	24	29.18	0.9230
25	29.36	25	29.17	0.9615

Βήμα 6: Η πιθανότητα της ελάχιστης ακραίας τιμής υπολογίζεται από τον τύπο:

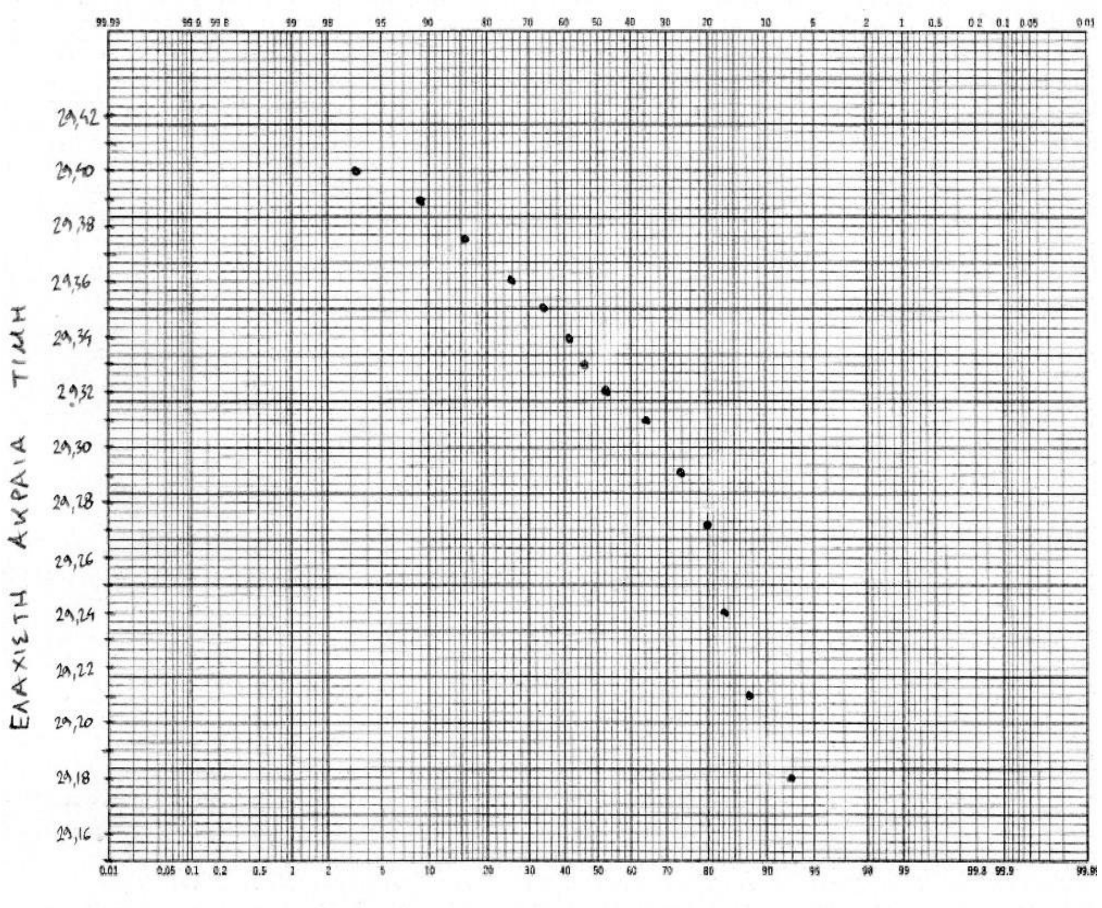
$$p = Ri(x) = \frac{m}{N+1}$$

Ο τύπος της ασυμπτωτικής κατανομής που προσεγγίζει καλύτερα τις πειραματικές μας τιμές (N ζεύγη τιμών: p, x) προσδιορίζεται γραφικώς με χρήση του ειδικού διαγράμματος πιθανότητας ακραίων τιμών extreme value probability paper, (Σχήμα 5)



Σχήμα 5: Καμπύλες πιθανότητας ακραίων τιμών (ελάχιστη η συνεχής γραμμή / μέγιστη η διακεκομμένη) στο αντίστοιχο ειδικό διάγραμμα. Κατανομές: 1) Εκθετική , 2) Cauchy , 3) Περιορισμένη.

Τα σημεία τοποθετήθηκαν στο διάγραμμα πιθανότητας ακραίων τιμών, Διάγραμμα 2. Σε όλα σχεδόν τα διαγράμματα (και των 12 συνδυασμών του Πίνακα 4), τα σημεία (p,x) φαίνεται να ακολουθούν καλύτερα την καμπύλη της κατανομής Cauchy. Επομένως, αυτή η κατανομή περιγράφει με μεγαλύτερη ακρίβεια τα πειραματικά δεδομένα.



Πιθανότητα P= m/N +1

ΔΙΆΓΡΑΜΜΑ 2: Πειραματικά σημεία (p,x) της ελάχιστης ακραίας τιμής του Πίνακα 5.

Β.2 ΜΕΛΕΤΗ ΠΕΙΡΑΜΑΤΙΚΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Διακρίνουμε από την μελέτη του Διαγράμματος 2 (p,x) ότι η $2^{η}$ ασυμπτωτική κατανομή (συνάρτηση πιθανότητας), κατανομή Cauchy, περιγράφει με μεγαλύτερη ακρίβεια τα πειραματικά δεδομένα (p,x), απ' ότι οι άλλες κατανομές.

Μπορούμε να φτάσουμε σε συμπέρασμα για την ασυμπτωτική κατανομή χωρίς να μας δίνεται ολόκληρο το μέγεθος n του δείγματος παρα μόνο έχοντας τις Ν ακραίες τιμές. Ωστόσο όταν ο αριθμός Νπ των τιμών δεν είναι μεγάλος υπάρχει πιθανότητα μεγάλου σφάλματος. Γενικά πρέπει να υπάρχει όσο το δυνατόν μεγαλύτερος αριθμός τιμών Νπ.

Η εκτίμηση της πιθανότητας ακραίων τιμών (μέγιστων ή ελάχιστων) είναι πιο ασφαλές να γίνεται με την χρήση της ασυμπτωτικής κατανομής παρά με την χρήση των τύπων της κλασικής στατιστικής (μέσης τιμής και τυπικής απόκλισης) και με την πρόβλεψη ότι η αρχική κατανομή είναι «κανονική». [6]

Η θεωρία ακραίων τιμών αποδεικνύεται πολύ χρήσιμη σε προβλήματα που έχουν να κάνουν με την αντοχή υλικών και τις μηχανικές ιδιότητες τους. Συνήθως, για να υπολογισθεί η τιμή της αντοχής ενός υλικού προσδιορίζεται μια μέση τιμή από ένα γνωστό πλήθος τιμών και στη συνέχεια αυτό διαιρείται με έναν συντελεστή ασφαλείας. Αυτό μπορεί να αποδειχθεί επισφαλές γιατί γίνεται υπόθεση ότι η το δείγμα προέρχεται από μία κανονική κατανομή και υπολογίζεται η μέση τιμή. Αυτό είναι λάθος και καλό είναι να προσδιορίζεται η αρχική κατανομή μέσω της θεωρίας ακραίων τιμών.

Άλλες εφαρμογές στις οποίες μπορεί να χρησιμοποιηθεί επιτυχημένα η Θεωρία Ακραίων Τιμών είναι:

• Ακραίων καιρικών φαινομένων όπως τυφώνες, έντονες χιονοπτώσεις, μεγάλες πλημμύρες κ.ο.κ. (Μετεωρολογία)

• Μεγάλων σεισμικών δονήσεων σε συγκεκριμένες γεωγραφικές περιοχές (Σεισμολογία)

• Πολύ υψηλών ή πολύ χαμηλών επιπέδων στάθμης των υδάτων σε συγκεκριμένες λίμνες, ποτάμια, φράγματα κ.ο.κ. (Υδρολογία)

• Καταστροφικών βλαβών ή αποτυχιών μηχανημάτων ή εξαρτημάτων, δηλαδή αποτυχιών που γίνονται σε σύντομο χρονικό διάστημα από την έναρξη λειτουργίας τους, πολύ πριν το μέσο χρόνο λειτουργίας ή ζωής τους (Αξιοπιστία Συστημάτων)

• Ηλιακών καταιγίδων που επηρεάζουν τις τηλεπικοινωνίες ή εκρήξεων υπερκαινοφανών (Αστρονομία) [2]


mm/inch
29.35
mm
OFF
ON
ZERO

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] An Introduction to Gumbel, or extreme value probability paper, Joseph F. Santner U.S. Environmental Protection Agency, 1973

[2] Extreme Value Theory (slides), 2008-10, Boutsikas Michael

[3] Πιθανοτικές κατανομές ακραίων βροχοπτώσεων. Εφαρμογή σε παγκόσμια κλίμακα. , Νεραντζάκη Σοφία , Αθήνα , Οκτώβριος 2012

[4] Gumbel, Emil Julius. 1958. Statistics of extremes. New York: Columbia University Press.

[5] Παναγιώτα Γαλιατσάτου, Στατιστικές Μέθοδοι Προσομοίωσης Ακραίων Γεγονότων. Επιμέρους εφαρμογές σε μετεωρολογική παλίρροια. , Αθήνα , 2009

[6] Γιώργος Α Σταμπολτζής, Στέφανος Ν Γιουρέλης , 1993, Στατιστική των Ακραίων Τιμών.

[7] FRECHET M. "Sur la loi de probabilite de I'ecart maximum", Ann. de la Soc. Polonaise de Math. (Cracow), 6:92, 1927.

[8] GUMBEL E.J., "Statistics of extremes", Columbia University Press, New York, 1958.