

数学系



Statistics of Extremes

E. J. Gumbel

0 212

23992

外文书库

E606 STATISTICS OF
EXTREMES

1389

4858

By E. J. GUMBEL
COLUMBIA UNIVERSITY

2



COLUMBIA UNIVERSITY PRESS
NEW YORK 1958

PUBLISHED 1958, COLUMBIA UNIVERSITY PRESS, NEW YORK
PUBLISHED IN GREAT BRITAIN, CANADA, INDIA, AND PAKISTAN
BY THE OXFORD UNIVERSITY PRESS
LONDON, TORONTO, BOMBAY, AND KARACHI
LIBRARY OF CONGRESS CATALOG CARD NUMBER: 57-10160

PREFACE

THE THEORY of extreme values is mentioned in most of the recent textbooks of statistics. However, this seems to be the first book devoted exclusively to these problems. It is meant for statisticians, and statistically minded scientists and engineers. To spread the application of the methods, the author has tried to keep it on an elementary level. Graphical procedures are preferred to tedious calculations. Special cases and easy generalizations are given as exercises. Therefore, the book may even meet the requirements of a textbook.

Since the line of thought pursued here is new, the reader should not expect complete solutions. There are bound to be shortcomings and gaps in a first treatment. Despite the usual prescriptions of academic habits, unsolved problems are here clearly stated. Perhaps a carping critic might maintain that more problems are raised than solved. However, this is the common fate of science.

This book is written in the hope, contrary to expectation, that humanity may profit by even a small contribution to the progress of science.

The opportunity to start this book presented itself when the necessary free time was imposed upon the author. A first draft of the manuscript was written in 1949 under a grant by the Lucius N. Littauer Foundation. This draft was revised in 1950 under a grant from the John Simon Guggenheim Memorial Foundation. The work continued while the author was engaged as a Consultant at Stanford University and later as an Adjunct Professor at Columbia University, partially under a grant from the Higgins Foundation, and under contract with the Office of Ordnance Research. Finally, a further grant from the Guggenheim Foundation made it possible to engage Professor Frank Lee (Columbia University) for the drawing of the graphs. The author wishes to state his sincere appreciation to these agencies for their help, without which he would not have had the chance to finish this book.

He is indebted to Mr. E. N. Munns, formerly Chief, Division of Forest Influence, U. S. Department of Agriculture, for valuable office help. Dr. Wolfgang Wasow of the University of California (Los Angeles) suggested changes in Chapters 1 and 3 and contributed to section 8.1. P. G. Carlson (Arthur Anderson and Co.) gave the proof for 3.2.3(7). The idea underlying 6.1.6 is due to Dr Julius Lieblein of the D. Taylor Model Basin. Paragraphs 6.2.4, 6.2.5, and 6.2.7 have been written by B. F. Kimball (N. Y.

State Public Service Commission). Most of the work in section 8.2 was done in collaboration with the late Roger D. Keeney. The Watson Scientific Computing Laboratory (Columbia University) calculated the figures given in Table 6.2.3 and put at the author's disposal an unpublished table of the probability points and the first two differences for $\Phi = .00001$ (.00001) .00600 (.00010) .08000 (.00100) .84000 (.00010) .98400 (.00001) .99995. These data facilitated much numerical work. Table 7.2.3 was calculated by Gladys R. Garabedian of Stanford University. The author takes this occasion to express his thanks for these important contributions.

Mr. Arnold Court (University of California, Berkeley) suggested applications to climatology, and Professor S. B. Littauer (Dept. of Industrial Engineering, Columbia University) proposed valuable improvements.

The author profited greatly from his collaboration with Professor A. M. Freudenthal (Columbia University) in their common work on fatigue of metals (see 7.3.5), and from the continuous interest which the Statistical Engineering Laboratory of the National Bureau of Standards took in his work.

The whole manuscript was revised and corrected by Mr. L. H. Herbach (New York University), P. G. Carlson, C. Derman (Columbia University), and A. Court. The bibliography was checked by Mr. Taro Yamane (New York University). It gives the author great pleasure to thank these collaborators, and in particular Mr. Carlson, for their friendly assistance. They certainly have contributed whatever may be good in this book. The rest should be charged against the author.

*New York,
1 October, 1957*

TABLES

1.2.9.	Normal Standard Deviation σ_n as Function of Sample Size n	39
2.1.2.	Probabilities of Modal m th Values	45
2.1.6.	Reduced Standard Errors $\sigma(\xi_m)/\sqrt{n}$ for Unit Standard Deviation	52
2.2.1.	The Probabilities $w(5, m, 5, x)$	60
2.2.7.	The Six Limiting Distributions of Exceedances	73
4.1.6.	The Three Classes of Distributions of the Exponential Type	125
4.2.2(1).	Normal Extremes	130
4.2.2(2).	Characteristic Normal Extremes u_n for Large Samples	133
4.2.2(3).	Normal Extremes and Intensity	133
5.1.1.	The Three Asymptotic Probabilities of Largest Values	157
5.1.3(1).	Asymptotic Probabilities of Initial Variates	162
5.1.3(2).	Asymptotic Probabilities of Extreme Values	164-165
5.2.4.	Semivariants of the First Asymptote	174
5.2.6.	The Three Curvatures of Largest Values	178
5.2.7.	Selected Probabilities for Normal and Largest Values	180
5.3.2.	Asymptotic Probabilities of the Three Largest Values	190
5.3.3.	Characteristics of m th Extremes	194
6.1.1.	Extreme Values for the Double Exponential Distribution	202
6.1.2(1).	Characteristic Minima	204
6.1.2(2).	Probabilities of Maxima and Minima of Largest Values	205
6.1.2(3).	The Iterated Natural Logarithm	205
6.1.5.	Standard Errors of Reduced Central Values	214
6.1.6.	The Extreme Control Band	218
6.2.2.	The Weights in Lieblein's Estimator for $N = 4, 5, 6$	215
6.2.3.	Means and Standard Deviations of Reduced Extremes	228
6.3.2.	The Standardized Flood z_m for Some Rivers	240
6.3.5.	Modal Values and Characteristic Oldest Ages, U.S.A., 1939-41	248
7.2.3.	Reduced Averages and Third Moment of the Third Asymptote	282
7.2.4.	The Four Pseudosymmetrical Cases	285
7.2.6.	Asymptotic Probabilities of Iterated Extremes	288
7.3.4.	Analysis of Droughts	301
7.3.5.	Parameters for Fatigue Failures of Nickel	304
8.1.5.	Asymptotic Distribution and Probability of ξ	316

8.1.6.	Characteristics of Reduced Extremes, Range, and Midrange	319
8.1.8.	Modal Extreme Reduced Ranges	323
8.2.3.	Transformations of Extremes to Extremal Statistics	330
8.3.1(1).	Population Standard Deviation of the Midrange	331
8.3.1(2).	Observed Midranges of Lives of Bulbs (in Ten Hours)	332
8.3.2.	Population Mean and Standard Deviation of the Range	334
8.3.4.	Characteristics for Normal Ranges	340
8.3.5(1).	Frequencies of Dew Points, Washington, D.C., 1905-1945	343
8.3.5(2).	Annual Range of Temperature in C°, Bergen, Norway, 1878-1926	343
8.3.5(3).	Annual Range of Temperature in F°, New Haven, Connec- ticut, 1907-1948	343
8.3.5(4).	Estimation of the Parameters for the Four Ranges	344

GRAPHS

1.1.9.	Characteristics for the Logarithmic Normal Distribution	19
1.2.1.	Intensity Functions $\mu(x)$	22
1.2.2(1).	Control Intervals for the Return Period	24
1.2.2(2).	Design Factor versus Design Quotient	25
1.2.5(1).	The First Laplacean and the Normal Probability	30
1.2.5(2).	The First Laplacean and the Normal Probability on Normal Paper	31
1.2.6.	Plotting Positions on Normal Paper	33
1.2.9(1).	Increase of Normal Standard Deviation as Function of the Sample Size	39
1.2.9(2).	Annual Mean Discharges, Colorado River at Bright Angel Creek, 1922-1939	40
2.1.4.	Asymptotic and Exact Standard Errors of m th Normal Values	50
2.1.6.	Control Band for Normal Central Values	54
2.2.4.	Probabilities $W(10,m,10,x)$ of Exceedances	65
2.2.5.	Sample Size n as Function of Percentage q and Probability $F_1(q)$	68
2.2.6.	Distributions and Probabilities of Rare Exceedances	71
3.1.3(1).	Return Periods of Largest Values	80
3.1.3(2).	Calculation of Quantiles from the Median Extreme	81
3.1.8.	Mean Largest Standardized Values	92
3.2.1.	Probability Paper for Normal Deviations	94
3.2.3.	The Mean Range Interpreted as Area	101
3.2.4.	The Observed Range as Tolerance Limit	105
3.2.6.	Standard Error of the Center as Function of Sample Size	109
4.1.1(1).	Averages of Largest Exponential Values	114
4.1.1(2).	Distributions of Largest Exponential Values	115
4.1.6.	The Three Classes of the Exponential Type	125
4.2.2(1).	Probabilities of Normal Extremes	129
4.2.2(2).	Quantiles of Normal Extremes	131
4.2.2(3).	Distributions of Normal Extremes for 2, 3, 5, and 10 Observations	132
4.2.2(4).	Averages of Normal Extremes in Natural Scale	134
4.2.2(5).	Occurrence Interval and Sample Size	134
4.2.2(6).	Standard Deviations for Normal Extremes, Deviates, and Ranges	135

4.2.2(7). Moment Quotients for Largest Normal Values and Deviates as Functions of Sample Size	136
4.2.2(8). Beta Diagrams for Normal Extremes, Deviates, and Ranges	137
4.2.3. Modal Largest Normal Value for $2n$ Observations	139
4.2.4. Probabilities of Studentized Normal Extreme Deviate	142
4.2.5. Characteristic Largest Value for Gamma Distribution	145
4.2.6. Asymptotic Parameters for the Logarithmic Normal Distribution	147
4.3.4. The Three Classes for the Cauchy Type	154
5.2.6. Extremal Probability Paper	177
5.2.7(1). Extreme and Normal Distributions	180
5.2.7(2). Probabilities for Extreme and Normal Standardized Values	181
5.2.7(3). Probabilities of Standardized Extremes, Range, and Midrange	182
5.2.7(4). Normal and Extremal Distribution as Function of the Probability	183
5.2.7(5). First Asymptotic Probability Traced on Logarithmic Normal Paper	184
5.3.1. Asymptotic Distributions of m th Reduced Extremes	188
5.3.2(1). Asymptotic Probabilities of m th Extremes Traced on Normal Paper	190
5.3.2(2). The Convergence of the Median to the Mode for m th Extremes	191
5.3.2(3). Asymptotic Mode and Median of m th Extremes	192
5.3.4. Cramér's Probabilities of Transformed m th Extremes	196
5.3.6. Correction for Obtaining the Composite Mode for Two Samples	199
6.1.2. Probabilities of Average Extreme Values	206
6.1.3. Modal m th Extremes $\bar{y}_m(N)$	207
6.1.4. Mean m th Extremes $\bar{y}_m(N)$	211
6.1.5. Homogeneity Test for the Most Probable Annual Flood	213
6.2.1(1). Probabilities of Extreme Exponential Values	220
6.2.1(2). Probabilities of Extreme Values for the Logistic Variate	220
6.2.1(3). Probabilities of Normal Extremes and Extreme Deviates	221
6.2.1(4). Calculated and Asymptotic Distributions of Normal Extremes for 10, 100, 1,000 Observations	222
6.2.3. Population Mean \bar{y}_N and Standard Deviation σ_N as Function of N	227
6.2.6. Correction B_N for Maximum Likelihood Method	233
6.3.1. Floods. Mississippi River, Vicksburg, Mississippi, 1898-1949	237

6.3.2.	Design Flood for Given Risk	240
6.3.3(1).	Annual Maxima of Temperatures and Their Differences from the Annual Mean. Bergen, Norway, 1857-1926	242
6.3.3(2).	Largest One-day Precipitations. Barakar Catchment Area, India	243
6.3.6.	Survivorship Function for Rubber Specimen	250
7.1.2(1).	Fréchet's Probability (II) for Different Values of k	261
7.1.2(2).	The Second Asymptotic Distribution of Largest Values	262
7.1.2(3).	The Generalized Probability $V(y)$	263
7.1.3.	Averages for the Second Asymptote	265
7.1.4(1).	Estimate of k from Coefficient of Variation	266
7.1.4(2).	Estimate of k from the Reciprocal Moments	267
7.2.1.	Probabilities III of Extreme Values, Traced on Normal Paper	275
7.2.2(1).	Probabilities III, Traced on Logarithmic Extremal Paper	277
7.2.2(2).	Influence of k on the Shape of the Distribution III of Smallest Values	278
7.2.3.	Averages of the Third Asymptote	280
7.3.1(1).	Parameter k and the Standardized Distances $A(k)$ and $B(k)$ as Functions of the Skewness,	290
7.3.1(2).	Standardized Distances from Lower Limit	291
7.3.2(1).	Estimation of k from the Coefficient of Variation	294
7.3.2(2).	Criterion for Lower Limit Zero	295
7.3.3(1).	Parameters v_n and k_n for Normal Extremes	299
7.3.3(2).	Approximations of the Probabilities of Normal Extremes by the Third Asymptote	300
7.3.4.	Droughts: Colorado and Connecticut Rivers	301
7.3.5.	Fatigue Failures of Nickel	303
8.1.4.	Asymptotic Probabilities of Extremes, Range and Midrange	312
8.1.6(1).	Asymptotic Distribution of the Range	318
8.1.6(2).	Asymptotic Probabilities of Standardized Extremes, Range and Midrange	319
8.1.8(1).	Averages of Extreme Reduced Ranges	322
8.1.8(2).	The Range of the Range	324
8.2.2.	Asymptotic Probabilities and Distributions of the Extremal Quotient for the Cauchy Type	326
8.2.3(1).	Asymptotic Probabilities $\Psi^{(2)}(\xi)$ and $\Psi^{(3)}(\xi)$ for the Geometric Range	328
8.2.3(2).	Estimation of k from the Coefficient of Variation of the Geometric Range	329
8.3.1(1).	Population Standard Deviation of Midrange	332
8.3.1(2).	Midranges of Lives of Bulbs Traced on Logistic Paper	333

8.3.2.	Population Mean Range and Standard Deviation as Function of Sample Size N	335
8.3.3(1).	Probabilities of Normal Ranges	337
8.3.3(2).	Characteristics of the Normal Range	338
8.3.5(1).	Range of Dew Points: Washington, D.C., 1905-1945	341
8.3.5(2).	Annual Range of Temperatures: Bergen, Norway, and New Haven, Connecticut	342

STATISTICS OF EXTREMES

CONTENTS

Chapter One: AIMS AND TOOLS	1
1.0. Aims	1
1.0.1. Conditions	1
1.0.2. History	2
1.0.3. The Flood Problem	4
1.0.4. Methodology	5
1.0.5. Arrangement of Contents	5
1.1. General Tools	7
1.1.1. Linear Transformations	7
1.1.2. Other Transformations	7
1.1.3. Symmetry	8
1.1.4. Measures of Dispersion	9
1.1.5. Moments	10
1.1.6. Generating Function	11
1.1.7. Convolution	13
1.1.8. The Gamma Function	15
1.1.9. The Logarithmic Normal Distribution	16
1.2. Specific Tools	20
1.2.0. Problems	20
1.2.1. The Intensity Function	20
1.2.2. The Distribution of Repeated Occurrences	21
1.2.3. Analysis of Return Periods	26
1.2.4. "Observed" Distributions	28
1.2.5. Construction of Probability Papers	28
1.2.6. The Plotting Problem	29
1.2.7. Conditions for Plotting Positions	32
1.2.8. Fitting Straight Lines on Probability Papers	34
1.2.9. Application to the Normal Distribution	38
 Chapter Two: ORDER STATISTICS AND THEIR EXCEEDANCES	 42
2.1. Order Statistics	42
2.1.0. Problems	42
2.1.1. Distributions	42

2.1.2. Averages	43
2.1.3. Distribution of Frequencies	46
2.1.4. Asymptotic Distribution of m th Central Values	47
2.1.5. The Order Statistic with Minimum Variance	51
2.1.6. Control Band	52
2.1.7. Joint Distribution of Order Statistics	53
2.1.8. Distribution of Distances	55
2.2. The Distribution of Exceedances	57
2.2.0. Introduction	57
2.2.1. Distribution of the Number of Exceedances	58
2.2.2. Moments	61
2.2.3. The Median	63
2.2.4. The Probability of Exceedances as Tolerance Limit	64
2.2.5. Extrapolation from Small Samples	67
2.2.6. Normal and Rare Exceedances	69
2.2.7. Frequent Exceedances	72
2.2.8. Summary	73
Chapter Three: EXACT DISTRIBUTION OF EXTREMES	75
3.1. Averages of Extremes	75
3.1.0. Problems	75
3.1.1. Exact Distributions	75
3.1.2. Return Periods of Largest and Large Values	78
3.1.3. Quantiles of Extremes	79
3.1.4. Characteristic Extremes	82
3.1.5. The Extremal Intensity Function	84
3.1.6. The Mode	85
3.1.7. Moments	87
3.1.8. The Maximum of the Mean Largest Value	89
3.2. Extremal Statistics	94
3.2.0. Problems	94
3.2.1. Absolute Extreme Values	94
3.2.2. Exact Distribution of Range	97
3.2.3. The Mean Range	100
3.2.4. The Range as Tolerance Limit	103
3.2.5. The Maximum of the Mean Range	106
3.2.6. Exact Distribution of the Midrange	108
3.2.7. Asymptotic Independence of Extremes	110
3.2.8. The Extremal Quotient	111

Chapter Four: ANALYTICAL STUDY OF EXTREMES	113
4.1. The Exponential Type	113
4.1.0. Problems	113
4.1.1. Largest Value for the Exponential Distribution	113
4.1.2. Order Statistics for the Exponential Distribution	116
4.1.3. L'Hôpital's Rule	118
4.1.4. Definition of the Exponential Type	120
4.1.5. The Three Classes	122
4.1.6. The Logarithmic Trend	123
4.1.7. The Characteristic Product	125
4.2. Extremes of the Exponential Type	126
4.2.0. Problems	126
4.2.1. The Logistic Distribution	126
4.2.2. Normal Extremes, Numerical Values	129
4.2.3. Analysis of Normal Extremes	136
4.2.4. Normal Extreme Deviates	140
4.2.5. Gamma Distribution	143
4.2.6. Logarithmic Normal Distribution	146
4.2.7. The Normal Distribution as a Distribution of Extremes	147
4.3. The Cauchy Type	149
4.3.0. Problems	149
4.3.1. The Exponential Type and the Existence of Moments	149
4.3.2. Pareto's Distribution	151
4.3.3. Definition of the Pareto and the Cauchy Types	152
4.3.4. Extremal Properties	153
4.3.5. Other Distributions without Moments	154
4.3.6. Summary	155
Chapter Five: THE FIRST ASYMPTOTIC DISTRIBUTION	156
5.1. The Three Asymptotes	156
5.1.0. Introduction	156
5.1.1. Preliminary Derivation	156
5.1.2. The Stability Postulate	157
5.1.3. Outline of Other Derivations	162
5.1.4. Interdependence	164
5.2. The Double Exponential Distribution	166
5.2.0. Introduction	166
5.2.1. Derivations	166
5.2.2. The Methods of Cramér and Von Mises	170

5.2.3. Mode and Median	172
5.2.4. Generating Functions	173
5.2.5. Standard and Mean Deviations	174
5.2.6. Probability Paper and Return Period	176
5.2.7. Comparison with Other Distributions	179
5.2.8. Barricelli's Generalization	184
5.3. Extreme Order Statistics	187
5.3.0. Problems	187
5.3.1. Distribution of the m th Extreme	187
5.3.2. Probabilities of the m th Extreme	189
5.3.3. Generating Functions	192
5.3.4. Cramér's Distribution of m th Extremes	194
5.3.5. Extreme Distances	197
5.3.6. The Largest Absolute Value and the Two Sample Problem	198
Chapter Six: USES OF THE FIRST ASYMPTOTE	201
6.1. Order Statistics from the Double Exponential Distribution	201
6.1.0. Problems	201
6.1.1. Maxima of Largest Values	201
6.1.2. Minima of Largest Values	203
6.1.3. Consecutive Modes	206
6.1.4. Consecutive Means and Variances	208
6.1.5. Standard Errors	212
6.1.6. Extension of the Control Band	216
6.1.7. The Control Curve of Dick and Darwin	218
6.2. Estimation of Parameters	219
6.2.0. Problems	219
6.2.1. Exponential and Normal Extremes	219
6.2.2. Use of Order Statistics	223
6.2.3. Estimates for Probability Paper	226
6.2.4. Sufficient Estimation Functions, by B. F. Kimball	229
6.2.5. Maximum Likelihood Estimations, by B. F. Kimball	231
6.2.6. Approximate Solutions	232
6.2.7. Asymptotic Variance of a Forecast, by B. F. Kimball	234
6.3. Numerical Examples	236
6.3.0. Problems	236
6.3.1. Floods	236
6.3.2. The Design Flood	238
6.3.3. Meteorological Examples	241

CONTENTS

xiii

6.3.4. Application to Aeronautics	245
6.3.5. Oldest Ages	246
6.3.6. Breaking Strength	248
6.3.7. Breakdown Voltage	249
6.3.8. Applications to Naval Engineering	251
6.3.9. An Application to Geology	254

Chapter Seven: THE SECOND AND THIRD ASYMPTOTES 255

7.1. The Second Asymptote	255
7.1.0. Problems	255
7.1.1. Fréchet's Derivation	255
7.1.2. The Cauchy Type	259
7.1.3. Averages and Moments	264
7.1.4. Estimation of the Parameters	266
7.1.5. The Increase of the Extremes	269
7.1.6. Generalization	270
7.1.7. Applications	271
7.1.8. Summary	272
7.2. The Third Asymptote	272
7.2.0. Introduction	272
7.2.1. The Von Mises Derivation	273
7.2.2. Other Derivations	276
7.2.3. Averages and Moments of Smallest Values	280
7.2.4. Special Cases	285
7.2.5. The Increase of the Extremes	287
7.2.6. The 15 Relations Among the 3 Asymptotes	288
7.3. Applications of the Third Asymptote	289
7.3.0. Problems	289
7.3.1. Estimation of the Three Parameters	289
7.3.2. Estimation of Two Parameters	293
7.3.3. Analytical Examples	298
7.3.4. Droughts	299
7.3.5. Fatigue Failures	302

Chapter Eight: THE RANGE 306

8.1. Asymptotic Distributions of Range and Midrange	306
8.1.0. Problems	306
8.1.1. The Range of Minima	307
8.1.2. Generating Function of the Range	308