Introduction to Tsallis Entropy Theory in Water Engineering

Vijay P. Singh

CRC Press
Taylor & Francis Group

Introduction to

Tsallis Entropy Theory in Water Engineering

This page intentionally left blank

Introduction to Tsallis Entropy Theory in Water Engineering

Vijay P. Singh

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

MATLAB® is a trademark of The MathWorks, Inc. and is used with permission. The MathWorks does not warrant the accuracy of the text or exercises in this book. This book's use or discussion of MATLAB® software or related products does not constitute endorsement or sponsorship by The MathWorks of a particular pedagogical approach or particular use of the MATLAB® software.

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20150813

International Standard Book Number-13: 978-1-4987-3661-9 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedicated to

my wife, Anita,
daughter, Arti,
son, Vinay,
daughter-in-law, Sonali,
and
grandsons, Ronin and Kayden

This page intentionally left blank

Contents

Preface..xv
Acknowledgments...xvii
Author xix

SECTION I Preliminaries

Chapter 1 Introduction to Tsallis Entropy Theory.....................................3

1.1 Definition of Tsallis Entropy 3
1.2 Derivation of Shannon Entropy from Tsallis Entropy..............8
1.3 Properties of Tsallis Entropy 10
1.3.1 m-Entropy...10
1.3.2 Maximum Value..11
1.3.3 Concavity...12
1.3.4 Additivity...13
1.3.5 Composibility 16
1.3.6 Interacting Subsystems..16
1.3.7 Other Features 18
1.4 Modification of Tsallis Entropy...20
1.5 Maximization 22
1.6 Partial Tsallis Entropy..24
1.7 Constrained Tsallis Entropy 26
1.8 Generalized Relative Entropies..28
References 30

Chapter 2 Probability Distributions 33

2.1 Procedure for Deriving a Probability Distribution..................33
2.1.1 Defining Tsallis Entropy 33
2.1.2 Specification of Constraints 34
2.1.3 Entropy Maximizing..38
2.1.4 Probability Distribution..39
2.1.5 Determination of the Lagrange Multipliers 39
2.1.6 Maximum Entropy 40
2.2 Maximum Entropy–Based Distributions with Regular Moments as Constraints 40
2.2.1 Mean as a Constraint..40
2.2.2 Mean and Variance as Constraints............................42
2.2.3 Variance as a Constraint..49
2.3 Use of m-Expectation 51

2.4 Choosing Expectation Value 52
2.4.1 Ordinary Expectation 53
2.4.2 Normalized *m*-Expectation 55
2.4.3 Use of *m*-Expectation 56
References 58

SECTION II Hydraulic Engineering

Chapter 3 One-Dimensional Velocity Distributions 61

3.1 Preliminaries 61
3.2 One-Dimensional Velocity Distributions 62
3.2.1 Hypothesis 62
3.2.2 Tsallis Entropy 63
3.2.3 Specification of Constraints 64
3.2.4 Maximization of Entropy 64
3.2.5 Determination of Lagrange Multipliers 66
3.2.6 Entropy of Velocity Distribution 66
3.2.7 General Velocity Distribution 66
3.3 One-Dimensional No-Constraint Velocity Distribution 66
3.4 One-Dimensional One-Constraint Velocity Distribution 69
3.4.1 Determination of the Lagrange Multipliers 70
3.4.2 Velocity Distribution 70
3.3.3 Reparameterization 74
3.4.4 Entropy 76
3.5 Relation between Mean Velocity and Maximum Velocity 81
3.6 Simplification of Velocity Distribution 83
3.7 Estimation of Mean Velocity 84
References 86

Chapter 4 Two-Dimensional Velocity Distributions 89

4.1 Preliminaries 89
4.2 2-D Velocity Distributions 90
4.3 One-Constraint Velocity Distribution Using the Chiu Coordinate System 90
4.3.1 2-D Coordinates 91
4.3.2 Formulation of a Hypothesis on Cumulative Probability Distribution 93
4.3.3 Specification of Constraints 94
4.3.4 Entropy Maximizing 94
4.3.5 Velocity Distribution 95
4.3.6 Determination of Parameters 95
4.3.7 Defining an Entropy Parameter 98

4.3.8 Maximum Entropy ... 99
4.3.9 Velocity Distribution in Terms of Parameter M ... 101
4.3.10 Construction of Isovels ... 105
4.3.11 Mean Velocity ... 105
4.3.12 Location of Mean Velocity ... 107
4.4 Velocity Distribution Using General Framework ... 108
4.4.1 Cumulative Distribution Function ... 108
4.4.2 2-D Velocity Distribution ... 109
4.4.3 Dimensionless Parameter G ... 111
4.4.4 Velocity Entropy ... 115
4.4.5 Location of Maximum Velocity ... 116
4.4.6 Construction of Isovels ... 120
4.A Appendix A ... 121
4.B Appendix B ... 122
References ... 123

Chapter 5 Suspended Sediment Concentration ... 125

5.1 Methods for Determining Sediment Concentration ... 125
5.1.1 Rouse Equation ... 126
5.1.2 Chang–Simons–Richardson Equation ... 128
5.1.3 O'Brien–Christiansen Equation ... 129
5.1.4 Lane–Kalinske Equation ... 129
5.1.5 Chiu Equation ... 129
5.1.6 Choo Equation ... 130
5.1.7 Summation ... 131
5.2 Derivation of Entropy-Based Suspended Sediment Concentration ... 131
5.2.1 Tsallis Entropy ... 132
5.2.2 Constraints ... 132
5.2.3 Maximization of Entropy ... 133
5.2.4 Cumulative Distribution Function ... 133
5.2.5 Sediment Concentration ... 135
5.2.6 Determination of the Lagrange Parameters ... 135
5.2.7 Entropy of Sediment Concentration ... 141
5.2.8 Estimation of Mean Sediment Concentration ... 143
5.2.9 Comparison with Other Concentration Distributions ... 144
5.3 Hydraulic Method ... 145
5.3.1 Probability Distribution of Travel Distance of Sediment ... 146
5.3.2 Determination of Distribution Parameters from the Random Walk Hypothesis ... 148
5.3.3 Determination of Initial Value of PDF, f_0 ... 151
5.3.4 Determination of Diffusion Coefficient ... 152
5.3.5 Sediment Concentration Distribution ... 155

5.3.6 Tsallis Entropy–Based Sediment Concentration Distribution ... 155
5.3.7 Random Walk Model-Based Distribution ... 157
References ... 161

Chapter 6 Suspended Sediment Discharge ... 165

6.1 Preliminaries ... 165
6.2 Sediment Discharge ... 166
6.2.1 Prandtl–von Karman Velocity Distribution ... 166
6.2.2 Tsallis Entropy–Based Velocity Distribution ... 167
6.2.3 Rouse Equation ... 168
6.2.4 Tsallis Entropy–Based Sediment Concentration ... 169
6.3 Suspended Sediment Discharge ... 169
6.3.1 Tsallis Entropy–Based Velocity Distribution and Sediment Concentration Distribution ... 169
6.3.2 Prandtl–von Karman Velocity Distribution and Tsallis Entropy–Based Sediment Concentration Distribution ... 170
6.3.3 Tsallis Entropy–Based Velocity Equation and Rouse Sediment Concentration Equation ... 172
6.3.4 Comparison ... 172
6.4 Modification for Sediment Discharge ... 173
References ... 174

Chapter 7 Sediment Concentration in Debris Flow ... 177

7.1 Notation and Definition ... 177
7.2 Methodology for the Determination of Debris Flow Concentration ... 178
7.2.1 Definition of Tsallis Entropy ... 178
7.2.2 Specification of Constraints ... 179
7.2.3 Maximization of Entropy ... 179
7.2.4 Determination of Lagrange Multipliers ... 180
7.2.5 Determination of Probability Density Function and Maximum Entropy ... 181
7.2.6 Cumulative Probability Distribution Hypothesis ... 186
7.2.7 Sediment Concentration Distribution ... 188
7.3 Reparameterization ... 188
7.4 Equilibrium Debris Flow Concentration ... 189
References ... 199
Additional Reading ... 199

Chapter 8 Stage–Discharge Relation 201

8.1 Methods for the Construction of Rating Curves 201
8.1.1 Graphical Method 201
8.1.2 Hydraulic Method.. 201
8.1.3 Artificial Intelligence Methods 202
8.2 Errors in Rating Curves... 202
8.3 Forms of Rating Curves 203
8.4 Randomness in Rating Curve... 204
8.5 Derivation of Rating Curves... 204
8.5.1 Definition of Tsallis Entropy 205
8.5.2 Specification of Constraints 205
8.5.3 Maximization of Entropy 206
8.5.4 Probability Distribution of Discharge 206
8.5.5 Determination of Lagrange Multipliers 207
8.5.6 Maximum Entropy 209
8.5.7 Hypothesis on Cumulative Probability Distribution of Discharge 210
8.5.8 Derivation of Rating Curve...................................... 210
8.6 Reparameterization 213
8.7 Relation between Maximum Discharge and Drainage Area 218
8.8 Relation between Mean Discharge and Drainage Area 218
8.9 Relation between Entropy Parameter and Drainage Area 219
8.10 Extension of Rating Curve 221
References 223

SECTION III Hydrology

Chapter 9 Precipitation Variability 227

9.1 Entropy as a Measure of Precipitation Uncertainty 227
9.2 Intensity Entropy 228
9.3 Apportionment Entropy of Temporal Precipitation............... 231
9.3.1 Apportionment from Empirical Data........................ 235
9.3.2 Apportionment by Random Experimentation.......... 236
9.3.3 Calculation of Precipitation Apportionment Entropy and Isoentropy Lines 236
9.4 Entropy Scaling 243
9.5 Determination of Water Resources Availability 244
9.5.1 Precipitation Data.. 244
9.5.2 Categorization of Water Resources Availability 246
9.5.3 Assessment of Water Resources Availability........... 247
References 248

Chapter 10 Infiltration 249

10.1 Preliminaries 249
10.2 Formulation of Entropy Theory 250
10.3 Methodology for the Derivation of Infiltration Equations 251
10.3.1 Principle of Maximum Entropy 251
10.3.2 Specification of Constraints 251
10.3.3 Maximization of Tsallis Entropy 252
10.3.4 Derivation of Probability Distribution and Maximum Entropy 252
10.3.5 Formulation of Continuity Equation 253
10.3.6 Cumulative Probability Distribution Hypothesis 254
10.3.7 Relation between Cumulative Infiltration and Infiltration Rate 255
10.3.8 Derivation of Infiltration Equation 255
10.4 Horton Equation 255
10.5 Kostiakov Equation 260
10.6 Philip Two-Term Equation 262
10.7 Green–Ampt Equation 265
10.8 Overton Model 268
10.9 Holtan Model 271
References 274

Chapter 11 Movement of Soil Moisture 277

11.1 Soil Moisture Zone 277
11.2 Soil Moisture Phases 278
11.3 Estimation of Soil Moisture Profile 279
11.4 Soil Moisture as a Random Variable 280
11.5 Methodology for Deriving Soil Moisture Distribution Using Tsallis Entropy 281
11.5.1 Tsallis Entropy 281
11.5.2 Principle of Maximum Entropy 281
11.5.3 Specification of Constraints 281
11.5.4 Maximization of Tsallis Entropy 282
11.5.5 Soil Moisture Profile for Wetting Phase 282
11.5.6 Moisture Profile for Drying Phase 287
11.5.7 Moisture Profiles for Mixed Phase 292
11.6 Soil Moisture Profile in Time 294
References 300

Chapter 12 Flow Duration Curve 303

12.1 Definition of Flow Duration Curve 303
12.2 Use of Flow Duration Curve 303
12.3 Construction of Flow Duration Curve 304

12.4 Derivation of Flow Duration Curve 305
12.4.1 Defining Tsallis Entropy 305
12.4.2 Specification of Constraints 306
12.4.3 Maximization of Tsallis Entropy and Probability Distribution 306
12.4.4 Derivation of Discharge PDF 307
12.4.5 Maximum Entropy 307
12.4.6 Determination of the Lagrange Multipliers 308
12.4.7 Hypothesis on Cumulative Distribution Function of Discharge 309
12.4.8 Flow Duration Curve 312
12.5 Reparameterization 315
12.6 Mean Flow and Ratio of Mean to Maximum Flow 317
12.7 Prediction of Flow Duration Curve for Ungaged Sites 321
12.8 Forecasting of Flow Duration Curve 324
12.9 Variation of Entropy with Time Scale 325
References 326

SECTION IV Water Resources Engineering

Chapter 13 Eco-Index 329

13.1 Indicators of Hydrological Alteration 330
13.2 Probability Distributions of IHA Parameters 331
13.3 Computation of Nonsatisfaction Eco-Level 331
13.4 Computation of Eco-Index 332
References 367

Chapter 14 Redundancy Measures for Water Distribution Networks 369

14.1 Optimization of Water Distribution Networks 369
14.1.1 Deterministic Optimization 369
14.1.2 Stochastic Optimization 370
14.1.3 Entropy-Based Optimization 370
14.2 Reliability Measure 370
14.3 Tsallis Entropy–Based Redundancy Measures 371
14.4 Redundancy Measures 371
14.5 Additivity Property for Independent Systems for First Approach 374
14.6 Additivity Property of Tsallis Entropy for Independent Systems for Approach Two 377
14.7 Transmission of Redundancy through Network 379
14.8 Consideration of Path Dependency 382

14.9 Modification of Redundancy Measure with Path Parameter 383
14.10 Modification of Redundancy Measure by Age Factor 384
14.11 Modification of Overall Network Redundancy 385
14.12 Relation between Redundancy and Reliability 424
References 425

Index 427

Preface

In 1988, Tsallis began to study a new form of entropy, called the Tsallis entropy, and in subsequent years, he developed the whole theory, which can be rightly referred to as the Tsallis entropy theory. This theory has since been applied to a wide spectrum of areas in physics and chemistry, and new topics applying this entropy are emerging each year. In the area of water engineering, the past few years have witnessed a range of applications of the Tsallis entropy. The literature shows the theory has enormous potential.

Currently, there seems to be no book on the Tsallis entropy for water engineering readership. Therefore, there exists a need for a book that deals with basic concepts of the Tsallis entropy theory and applications of these concepts to a range of water engineering problems. This book is an attempt to cater to this need.

The subject matter of the book is divided into 14 chapters organized in 4 sections. Section I, comprising two chapters, deals with preliminaries. Chapter 1 discusses the Tsallis entropy theory for both discrete and continuous variables. It then goes on to discuss the properties of the Tsallis entropy, partial Tsallis entropy, and constrained Tsallis entropy. The chapter is concluded with a discussion of generalized entropies. Frequency analysis constitutes the subject matter of Chapter 2. Beginning with a discussion of the procedure for deriving probability distributions, it goes on to present maximum entropy–based distributions with regular moments as constraints, the use of *m*-expectation, and choosing expectation value.

Section II consists of six chapters dealing with some aspects of hydraulics. One-dimensional velocity distributions are discussed in Chapter 3, which presents velocity distributions based on different constraints or the specification of information. It also discusses the relation between mean velocity and maximum velocity, simplification of the velocity distribution, and estimation of mean velocity. Chapter 4 presents two-dimensional velocity distributions using the Chiu coordinate system and the generalized framework. It deals with different characteristics of the velocity distribution.

Chapter 5 discusses sediment concentration. Starting with a discussion of the methods for determining sediment concentration, it presents a step-by-step procedure for the derivation of entropy-based suspended sediment concentration and the characteristics of the derived distribution. Chapter 6 treats the subject of sediment discharge in three ways. First, it considers velocity as entropy based but not sediment concentration. The second considers sediment concentration as entropy-based but not entropy-based velocity. The third considers both velocity and sediment concentration as entropy-based. The sediment concentration in debris flow is presented in Chapter 7. It presents a step-by-step methodology for determining the debris flow concentration and concludes with the treatment of reparameterization and equilibrium debris flow concentration. Chapter 8 deals with the stage–discharge rating curve. It first discusses errors and randomness in rating curves and forms thereof. It then discusses the derivation of rating curves, reparameterization, relation between maximum discharge and drainage area, relation between mean discharge and drainage area, relation between entropy parameter and drainage area, and extension of rating curves.

Hydrology is the subject of Section III, which comprises four chapters. Chapter 9 discusses precipitation variability and deals with intensity entropy, apportionment entropy, entropy scaling, hydrological zoning, and the assessment of water resources availability. Infiltration is discussed in Chapter 10, which presents the derivation of six infiltration equations, including the equations of Horton, Kostiakov, Philip, Green and Ampt, Overton, and Holtan. Chapter 11 is on soil moisture. Providing a short introduction to soil moisture profiles and their estimation, it presents the derivation of soil moisture profiles for wetting, drying, and mixed phases and the variation of soil moisture in time. Chapter 12 deals with flow duration curves. Discussing first the use and construction of flow duration curves, it presents a step-by-step procedure for deriving flow duration curves, reparameterization, mean flow and ratio of mean to maximum flow, prediction of flow duration curves for ungagged sites, forecasting of flow duration curve, and variation of entropy with time scale.

The concluding Section IV is on water resources engineering; it contains two chapters. Eco-index constitutes the subject matter of Chapter 13, containing indicators of hydrologic alteration (IHA), probability distributions of IHA parameters, and computation of nonsatisfaction eco-level and eco-index. Chapter 14 discusses measures of redundancy for water distribution networks. Presenting the optimization of water distribution networks, it deals with reliability, the Tsallis entropy, redundancy measures, the development of redundancy measures under different conditions, and the relation between redundancy and reliability.

Vijay P. Singh
College Station, Texas

Acknowledgments

The book draws from the works of tens of scientists and engineers that have been inspiring. I have tried to acknowledge these works specifically. Any omission on my part has been entirely inadvertent and I offer my apologies in advance.

I have had a number of graduate students and visiting scholars over the years who have helped me in myriad ways, and I am grateful for their help. I would particularly like to acknowledge Dr. Z. Hao from Beijing Normal University, China; Dr. Mrs. H. Cui and Dr. Clement Sohoulande from Texas A&M University; and Dr. Deepthi Rajasekhar from Stanford University. Without their support, this book would not have been completed.

Finally, I would like to take this opportunity to acknowledge the support of my brothers and sisters in India and my family here in the United States that they have given me over the years. My wife, Anita, daughter, Arti, son, Vinay, and daughter-in-law, Sonali are always there to lend me a helping hand. My grandsons, Ronin and Kayden, are my future: they make my life complete. Therefore, I dedicate this book to them, for without their support and affection, this book would not have come to fruition.

This page intentionally left blank

Author

Vijay P. Singh, PhD, DSc, PE, PH, Hon. DWRE, is a university distinguished professor and holds the Caroline and William N. Lehrer Distinguished Chair in Water Engineering at Texas A&M University, College Station, Texas. He currently serves as editor-in-chief of Springer's *Water Science and Technology Library Book Series*, the *Journal of Groundwater Research*, and *De Gruyter Open Journal of Agriculture*. He is also an associate editor of more than 15 other journals. He has won more than 72 national and international awards—including the Chow Award, Arid Lands Hydraulic Engineering Award, Torrens Award, Normal Medal, and Lifetime Achievement Award of ASCE; Linsley Award and Founders' Award of American Institute of Hydrology (AIH); and three honorary doctorates—for his technical contributions and professional service. Professor Singh has been president and senior vice-president of AIH and president of the Louisiana Section of AIH. He is a distinguished member of American Society of Civil Engineers (ASCE) and an honorary member of American Water Resources Association (AWRA), and a fellow of Environmental and Water Resources Institute (EWRI), Institution of Engineers (IE), Indian Society of Agricultural Engineers (ISAE), Indian Water Resources Society (IWRS), and Indian Association of Soil and Water Conservationists (IASWC), as well as a member or fellow of 10 international science/engineering academies. He is a member of numerous committees of ASCE and AWRA and is currently serving as chair of Watershed Council of ASCE. He has extensively published in the areas of surface water hydrology, groundwater hydrology, hydraulic engineering, irrigation engineering, environmental engineering, water resources, and stochastic and mathematical modeling.

This page intentionally left blank

Section I

Preliminaries

This page intentionally left blank

1 Introduction to Tsallis Entropy Theory

The concept of entropy originated in thermodynamics and has a history of over a century and half dating back to Clausius in 1850. In 1870, Boltzmann developed a statistical definition of entropy and hence connected it to statistical mechanics. The concept of entropy was further advanced by Gibbs in thermodynamics and by von Neumann in quantum mechanics. Outside of the world of physics, it is Shannon who developed, in the late 1940s, the mathematical foundation of entropy and connected it to information. The informational entropy is now frequently called Shannon entropy or sometimes called Boltzmann–Gibbs–Shannon entropy. Kullback and Leibler (1951) developed the principle of minimum cross entropy (POMCE) and in the late 1950s Jaynes (1957a,b) developed the principle of maximum entropy (POME). Koutsoyiannis (2013, 2014) has given an excellent historical perspective on entropy. The Shannon entropy, POME, and POMCE constitute the entropy theory that has witnessed a wide spectrum of applications in virtually every field of science and engineering and social and economic sciences, and each year new applications continue to be reported (Singh, 2013, 2014, 2015). A review of entropy applications in hydrological and earth sciences is given in Singh (1997, 2010, 2011).

In 1988, Tsallis postulated a generalization of the Boltzmann–Gibbs–Shannon entropy, now popularly called the Tsallis entropy, and discussed its mathematical properties. The definition and properties of the Tsallis entropy constitute the Tsallis entropy theory. In physics, the Tsallis entropy has received tremendous attention (Tsallis, 2001). Recently, this entropy has been applied to a number of geophysical, hydrological, and hydraulic processes. Because of its interesting properties, it is expected to receive increasing attention in water engineering in the years ahead. This chapter introduces the Tsallis entropy and presents its properties that are of particular interest in environmental and water engineering.

1.1 DEFINITION OF TSALLIS ENTROPY

First, it is useful to define the Boltzmann–Gibbs–Shannon entropy (henceforth, simply Shannon entropy). For a discrete random variable $X = \{x_i, i = 1, 2, \ldots, N\}$ that has a probability distribution $P = \{p_i, i = 1, 2, \ldots, N\}$ [p_i is the probability of $X = x_i$], the Shannon entropy H_s can be defined as

$$H_s = -k\sum_{i=1}^{N} p_i \log p_i \tag{1.1}$$

where k is a conventional positive constant and is often taken as unity and log is taken to the base of 2, e or 10, and accordingly, the unit of entropy becomes bit, nat, or docit.

Scaling p_i to p_i^m, where m is any real number, Tsallis (1988) postulated

$$H_m = k\frac{1-\sum_{i=1}^{N} p_i^m}{m-1} = \frac{k}{m-1}\sum_{i=1}^{N}\left[p_i - p_i^m\right] \tag{1.2}$$

where

H_m is the Tsallis entropy
k is often taken as unity

For $m \to 1$, the Tsallis entropy reduces to the Shannon entropy. Quantity m is often referred to as nonextensivity index or Tsallis entropy index or simply entropy index. Entropy index m characterizes the degree of nonlinearity and is related to the microscopic dynamics of the system. The value of m can be positive or negative. The Tsallis entropy is often referred to as nonextensive statistic, m-statistic, or Tsallis statistic. Tsallis (2002) noted that superextensivity, extensivity, and subextensivity occur when $m < 1$, $m = 1$, or $m > 1$, respectively. For $m \geq 0$, $m < 1$ corresponds to the rare events and $m > 1$ corresponds to frequent events (Tsallis, 1998; Niven, 2004) pointing to the stretching or compressing of the entropy curve to lower or higher maximum entropy positions.

From an informational perspective, the information gain from the occurrence of any event i is a power function and can be expressed as

$$\Delta I_i = \frac{1}{m-1}\left(1 - p_i^{m-1}\right), \quad \sum_{i=1}^{N} p_i = 1 \tag{1.3}$$

where

ΔI_i is the gain in information from an event i that occurs with probability p_i
m is any real number
N is the number of events

Equation 1.3 is a generalization of the Shannon gain function describing the information from an event expressed in logarithmic terms. For N events, the average or expected gain function is the weighted average of Equation 1.3

$$H_m = E[\Delta I_i] = \sum_{i=1}^{N} p_i\left[\frac{1}{m-1}\left(1 - p_i^{m-1}\right)\right] = \frac{1}{m-1}\sum_{i=1}^{N} p_i\left(1 - p_i^{m-1}\right) \tag{1.4}$$

where H_m is designated as the Tsallis entropy or m-entropy.

In a similar manner, the information gain for the Shannon entropy, ΔH_{si}, can be written as

$$\Delta H_{si} = -\log p_i \tag{1.5}$$

Therefore,

$$H_s = \sum_{i=1}^{N} H_{si} = -\sum_{i=1}^{N} p_i \log p_i \tag{1.6}$$

If random variable X is nonnegative continuous with a probability density function (PDF), $f(x)$, then the Shannon entropy can be written as

$$H_s(X) = H_s(f) = -k\int_0^{\infty} f(x)\log f(x)dx \tag{1.7}$$

Likewise, the Tsallis entropy can be expressed (Koutsoyiannis, 2005a,b,c) as

$$H_m(X) = H_m(f) = \frac{k}{m-1}\int_0^{\infty}\{f(x)-[f(x)]^m\}dx = \frac{k}{m-1}\left\{1-\int_0^{\infty}[f(x)]^m\right\}dx \tag{1.8}$$

Frequently, k is taken as 1. From now onward, subscript m will be deleted and H_m will be simply denoted by H.

A plot of H/k versus p for $m = -1, -0.5, 0, 0.5, 1$, and 2 is given in Figure 1.1. For $m < 0$, the Tsallis entropy is concave and for $m > 0$ it becomes convex. For $m = 0$, $H = k(N-1)$ for all p_i. For $m = 1$, it converges to the Shannon entropy. For all cases, the Tsallis entropy decreases as m increases.

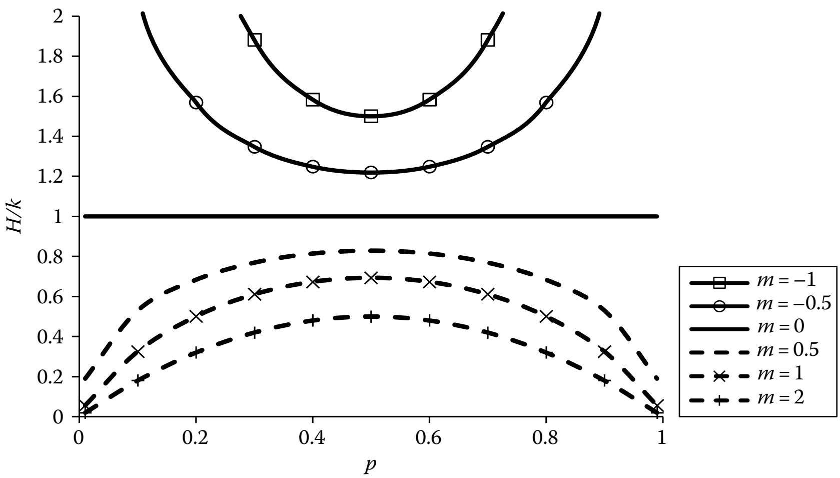

FIGURE 1.1 Plot of H/k for $N = 2$ for $m = -1, -0.5, 0, 0.5, 1$, and 2.

Example 1.1

Plot the gain function defined by the Tsallis entropy for different values of probability: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. Take k as 1, and m as −1, 0, 1, and 2. What do you conclude from this plot?

Solution

Using Equation 1.3 the gain function is computed, as shown in Table 1.1. Figure 1.2 shows the gain function for $m = -1$, 0, $m = 1$, and 2. It is seen from the figure that the gain in information decreases with the increase in the probability value regardless of the value of m. For increasing value of m, the gain diminishes for the same

TABLE 1.1
Computation of Gain Function

	ΔI			
p	$m = -1$	$m = 0$	$m = 1$	$m = 2$
0.1	49.50	9.00	2.30	0.90
0.2	12.00	4.00	1.61	0.80
0.3	5.06	2.33	1.20	0.70
0.4	2.63	1.50	0.92	0.60
0.5	1.50	1.00	0.69	0.50
0.6	0.89	0.67	0.51	0.40
0.7	0.52	0.43	0.36	0.30
0.8	0.28	0.25	0.22	0.20
0.9	0.12	0.11	0.11	0.10
1	0.00	0.00	0.00	0.00

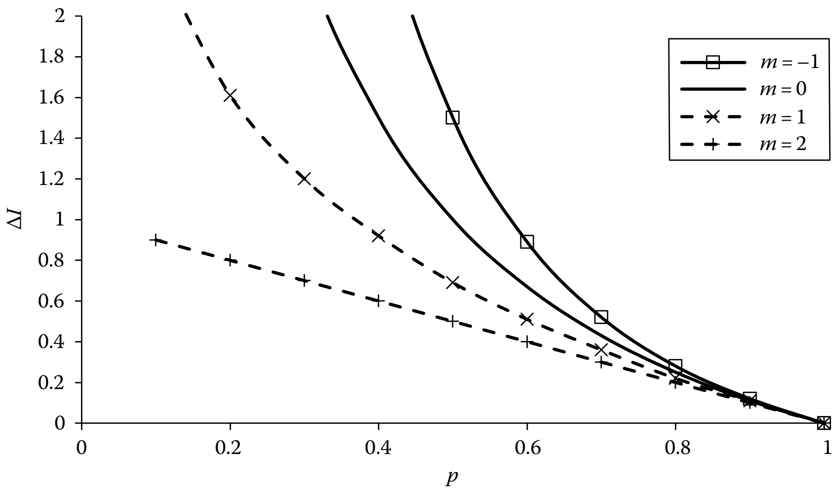

FIGURE 1.2 Gain function for $m = -1$, 0, 1, and 2.

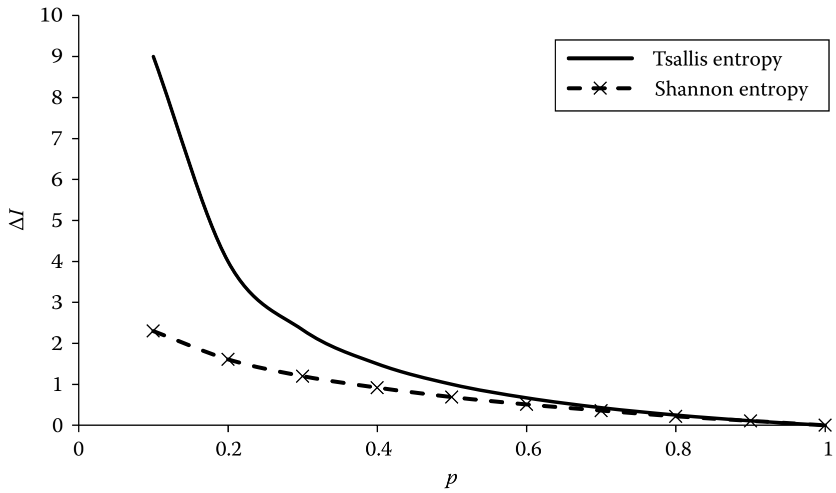

FIGURE 1.3 Comparison of the Shannon and Tsallis gain functions.

probability value. For $m = 1$, the Tsallis entropy converges to the Shannon entropy. The two gain functions are shown in Figure 1.3. The Tsallis gain function has a much longer tail showing very low values of gain as the probability increases.

Example 1.2

Consider a two-state variable taking on values x_1 and x_2. Assume that $p(x_1) = 0.0$, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. Note that $p(x_2) = 1 - p(x_1)$. Compute and plot the Tsallis entropy. Take m as 1.5 and 2.0. What do you conclude from the plot?

Solution

The Tsallis entropy is given by Equation 1.2. Let $a = p(x_1)$. For any given value of p, one can write the Tsallis entropies H_1 and H_2, respectively, for x_1 and x_2 as

$$H_1 = \frac{k}{m-1} a(1 - a^{m-1})$$

$$H_2 = \frac{k}{m-1}(1-a)[1-(1-a)^{m-1}]$$

Then, the Tsallis entropy is

$$H = H_1 + H_2$$

where each component is a weighted gain function. Thus, the Tsallis entropy is computed as shown in Table 1.2. The computed Tsallis entropy for $k = 1$ and $m = 1.5$ and 2 is shown in Figure 1.4. The Tsallis entropy plot shows a little skewness from the Shannon entropy and also predicts the maximum entropy at $p(x) = 0.5$. It also can be observed that the Tsallis entropy value decreases with an increase in the value of m.

TABLE 1.2
Computation of the Tsallis Entropy for $k = 1$, $m = 1.5$, and $m = 2$

		$m = 1.5$			$m = 2.0$		
$p(x)$	$1 - p(x)$	H_1	H_2	$H = H_1 + H_2$	H_1	H_2	$H = H_1 + H_2$
0	1	0	0	0	0	0	0
0.1	0.9	0.137	0.092	0.229	0.090	0.090	0.180
0.2	0.8	0.221	0.169	0.390	0.160	0.160	0.320
0.3	0.7	0.271	0.229	0.500	0.210	0.210	0.420
0.4	0.6	0.294	0.270	0.565	0.240	0.240	0.480
0.5	0.5	0.293	0.293	0.586	0.250	0.250	0.500
0.6	0.4	0.270	0.294	0.565	0.240	0.240	0.480
0.7	0.3	0.229	0.271	0.500	0.210	0.210	0.420
0.8	0.2	0.169	0.221	0.390	0.160	0.160	0.320
0.9	0.1	0.092	0.137	0.229	0.090	0.090	0.180
1	0	0	0	0	0	0	0

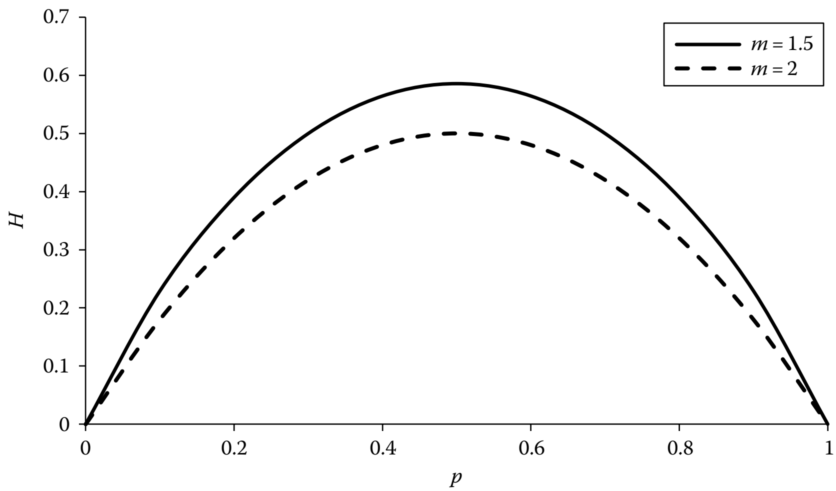

FIGURE 1.4 Tsallis entropy for $k = 1$ and $m = 1.5, 2.0$.

1.2 DERIVATION OF SHANNON ENTROPY FROM TSALLIS ENTROPY

It may be useful to show that the Tsallis entropy is a generalization of the Shannon entropy. One can express

$$p_i^m = p_i \exp[(m-1)\ln p_i] \tag{1.9}$$

The Tsallis entropy given by Equation 1.2 can be written as

$$H = \frac{k}{m-1}\left\{1-\sum_{i=1}^{N} p_i \exp[(m-1)\ln p_i]\right\} \tag{1.10}$$

It must now be shown that when m tends to unity

$$H = k\lim_{m\to 1}\frac{1-\sum_{i=1}^{N} p_i \exp[(m-1)\ln p_i]}{m-1} \tag{1.11}$$

leads to the Shannon entropy given by Equation 1.1.

Now, consider L'Hospital's rule for the division of two arbitrary functions $f(a)$ and $g(a)$:

$$\lim_{a\to b}\frac{f(a)}{g(a)}, \quad \text{if} \lim_{a\to b} g(a) = 0 \text{ or } \infty, \quad \lim_{a\to b} g(a) = 0 \text{ or } \infty \tag{1.12}$$

where b is some value and may even approach infinity. For example,

$$\lim_{m\to 1} f(m) = \lim_{m\to 1}\left(1-\sum_{i=1}^{N} p_i^m\right) = 1-\sum_{i=1}^{N} p_i = 0 \tag{1.13}$$

$$\lim_{m\to 1} g(m) = \lim_{m\to 1}(m-1) = 0 \tag{1.14}$$

Now

$$f(m) = 1-\sum_{i=1}^{N} p_i \exp[(m-1)\ln p_i] \tag{1.15}$$

or

$$f'(m) = -\sum_{i=1}^{N} p_i \ln p_i \exp[(m-1)\ln p_i] = -\sum_{i=1}^{N} p_i^m \ln p_i \tag{1.16}$$

$$g(m) = m-1 \tag{1.17}$$

$$g'(m) = 1 \tag{1.18}$$

Therefore, taking the limit on Equation 1.11,

$$\lim_{m\to 1} H = \lim_{m\to 1} \frac{f(m)}{g(m)} = \lim_{m\to 1} \frac{f'(m)}{g'(m)} = \lim_{m\to 1} \sum_{i=1}^{N} (-1) p_i^m \ln p_i$$

$$= -\lim_{m\to 1} \sum_{i=1}^{N} p_i^m \ln p_i = -\sum_{i=1}^{N} p_i \ln p_i \tag{1.19}$$

which is the Shannon entropy.

1.3 PROPERTIES OF TSALLIS ENTROPY

Following Tsallis (1988, 2004), some interesting and useful properties of the Tsallis entropy are briefly summarized here.

1.3.1 *m*-Entropy

Analogous to surprise or unexpectedness defined in the Shannon entropy, the m-surprise or m-unexpectedness is defined as $\log_m(1/p_i)$. Hence, the m-entropy can be defined as

$$H = E\left[\log_m \frac{1}{p_i}\right] \tag{1.20}$$

which coincides with the Tsallis entropy:

$$H = E\left[\frac{1 - p_i^{m-1}}{m-1}\right] \tag{1.21}$$

in which E is the expectation. Recalling the definition

$$\lim_{n\to 0}\left[\frac{w^n - 1}{n}\right] = \log w \tag{1.22}$$

where
n is any number
w is any variable

Then, Equation 1.21 is the same as Equation 1.20. For small values of n, w^n will behave as log w. A plot of function $(w^n - 1)/n$ is shown in Figure 1.5 that shows its approximation by the logarithmic function.

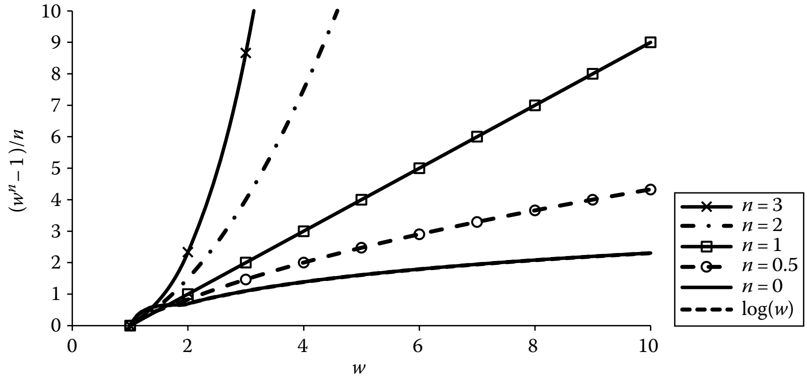

FIGURE 1.5 Plot of function $(w^n - 1)/n$ versus w for various values of n.

1.3.2 Maximum Value

Equation 1.2 attains an extreme value for all values of m when all p_i are equal, that is, $p_i = 1/N$. For $m > 0$, it attains a maximum value and for $m < 0$, it attains a minimum value. The extremum of H becomes

$$H = k\frac{N^{m-1} - 1}{1 - m} \tag{1.23}$$

If $m = 1$, applying L'Hopsital's rule to Equation 1.23 or 1.22, one gets

$$H = k \ln N \tag{1.24}$$

which is the Boltzmann entropy, H_B. Plotting H/k versus N using Equation 1.23, as shown in Figure 1.6, it is seen that H diverges for $m < 1$. The Tsallis entropy, given

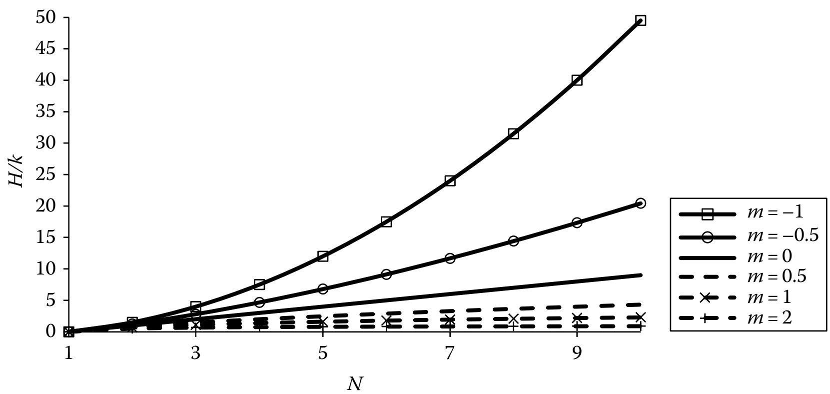

FIGURE 1.6 Plot of H/k versus N for $m = -1, -0.5, 0, 0.5, 1, 2$ when all p_i are equal (from Equation 1.18).

by Equation 1.23, diverges if $m < 1$, is maximum for $m > 1$ and is minimum for $m < 1$, and is $k(N - 1)$ for all equal p_i. Interestingly, for any value of m, the entropy extreme can be expressed in terms of the entropy for $m = 1$ as follows. For $m = 1$, Equation 1.24 can be written as $N = \exp(H/k)$. Substituting it into Equation 1.23, the result is

$$\frac{H_m}{k} = \frac{\exp[(1-m)H_B/k]-1}{1-m} \tag{1.25}$$

1.3.3 Concavity

Consider two probability distributions $P = \{p_i, i = 1, 2, \ldots, N\}$ and $Q = \{q_i, i = 1, 2, \ldots, N\}$ corresponding to a unique set of N possibilities. Then, an intermediate probability distribution $G = \{g_i, i = 1, 2, \ldots, N\}$ can be defined for a real a such that $0 < a < 1$ as

$$g_i = ap_i + (1-a)q_i \tag{1.26}$$

for all i. It can be shown that for $m > 0$,

$$H[G] \geq aH[P] + (1-a)H[Q] \tag{1.27}$$

and for $m < 0$,

$$H[G] \leq aH[P] + (1-a)H[Q] \tag{1.28}$$

Functional $H(G) \geq 0$ if $m > 0$ and is hence concave; $H(G) = 0$ if $m = 0$; and $H(G) \leq 0$ if $m < 0$ and is, therefore, convex. These inequalities, given by Equations 1.27 and 1.28, are true for $m \neq 0$ and $p_i = q_i, \forall i$.

Example 1.3

Consider $N = 3$, $m = 3$, and $P = \{0.2, 0.4, 0.4\}$ and $G = \{0.1, 0.3, 0.6\}$ and $a = 0.3$. Compute $H(P)$ and $H(G)$, and then show if Equation 1.27 holds. If $m = -0.5$, then show if Equation 1.28 holds.

Solution

$$H_m = E[\Delta I_i] = \sum_{i=1}^{N} p_i \left[\frac{1}{m-1}\left(1-p_i^{m-1}\right)\right] = \frac{1}{m-1}\sum_{i=1}^{N} p_i \left[1-p_i^{m-1}\right]$$

Given $a = 0.3$, from Equation 1.26, Q can be computed as

$$q_i = \frac{g_i - ap_i}{(1-a)}$$

$$q_1 = \frac{0.1 - 0.3 \times 0.2}{1 - 0.3} = 0.06$$

$$q_2 = \frac{0.3 - 0.3 \times 0.4}{1 - 0.3} = 0.26$$

$$q_3 = \frac{0.6 - 0.3 \times 0.4}{1 - 0.3} = 0.68$$

When m = 3,

$$H(P) = \frac{1}{3-1}[(0.2 - 0.2^3) + (0.4 - 0.4^3) + (0.4 - 0.4^3)] = 0.432$$

$$H(Q) = \frac{1}{3-1}[(0.06 - 0.06^3) + (0.26 - 0.26^3) + (0.68 - 0.68)] = 0.330$$

$$H(G) = \frac{1}{3-1}[(0.1 - 0.1^3) + (0.3 - 0.3^3) + (0.6 - 0.6^3)] = 0.378$$

$$H[G] \geq aH[P] + (1-a)H[Q] = 0.3 \times 0.432 + (1 - 0.3) \times 0.330 = 0.361$$

Equation 1.27 holds.

When m = −0.5

$$H(P) = \frac{1}{-0.5-1}[(0.2 - 0.2^{-0.5}) + (0.4 - 0.4^{-0.5}) + (0.4 - 0.4^{-0.5})] = 2.932$$

$$H(Q) = \frac{1}{-0.5-1}[(0.06 - 0.06^{-0.5}) + (0.26 - 0.26^{-0.5}) + (0.68 - 0.68^{-0.5})] = 4.242$$

$$H(G) = \frac{1}{-0.5-1}[(0.1 - 0.1^{-0.5}) + (0.3 - 0.3^{-0.5}) + (0.6 - 0.6^{-0.5})] = 3.519$$

$$H[G] \leq aH[P] + (1-a)H[Q] = 0.3 \times 2.932 + (1 - 0.3)4.242 = 3.849$$

Equation 1.28 holds.

1.3.4 Additivity

Let there be two independent systems A and B with ensembles of configurational possibilities $E^A = \{1, 2, \ldots, N\}$ with probability distribution $P^A = \{p_i^A, i = 1, 2, \ldots, N\}$ and configurational possibilities $E^B = \{1, 2, \ldots, M\}$ with probability distribution

$P^B = \{p_j^B, j = 1,2,\ldots,M\}$. Then, one needs to deal with the union of two systems $A \cup B$ and their corresponding ensembles of possibilities $E^{A \cup B} = \{(1,1), (1,2), \ldots, (i, j), \ldots, (N, M)\}$. If $p_{ij}^{A\cup B}$ represents the corresponding probabilities then by virtue of independence the joint probability will be equal to the product of individual probabilities, that is $p_{ij}^{A\cup B} = p_i^A p_j^B$ or $p_{ij}(A + B) = p_i(A)p_j(B)$ for all i and j. Hence,

$$\sum_{i,j}^{N,M}\left(p_{ij}^{A\cup B}\right)^m = \left[\sum_{i=1}^{N}\left(p_i^A\right)^m\right]\left[\sum_{j=1}^{M}\left(p_j^B\right)^m\right] \tag{1.29}$$

Taking the logarithms of Equation 1.29, one obtains

$$\log\left[\sum_{i,j}^{N,M}\left(p_{ij}^{A\cup B}\right)^m\right] = \log\left[\sum_{i=1}^{N}\left(p_i^A\right)^m\right] + \log\left[\sum_{j=1}^{M}\left(p_j^B\right)^m\right] \tag{1.30}$$

Each term of Equation 1.30 is now considered. The left side of Equation 1.30 can be written in terms of the Tsallis entropy as

$$\log\left[\sum_{i,j}^{N,M}\left(p_{ij}^{A\cup B}\right)^m\right] = \log\left\{1 - \frac{(m-1)\left[1 - \sum_{i=1,j=1}^{N,M}\left(p_{ij}^{A\cup B}\right)^m\right]}{(m-1)}\right\}$$

$$= \log[1-(m-1)H^{A\cup B}] \tag{1.31}$$

Similarly, terms on the right side of Equation 1.31 can be written as

$$\log\left[\sum_{i}^{N}\left(p_i^A\right)^m\right] = \log\left\{1 - \frac{(m-1)\left[1 - \sum_{i=1}^{N}\left(p_i^A\right)^m\right]}{(m-1)}\right\} = \log[1-(m-1)H^A] \tag{1.32}$$

$$\log\left[\sum_{j}^{M}\left(p_j^B\right)^m\right] = \log\left\{1 - \frac{(m-1)\left[1 - \sum_{i=1}^{N}\left(p_j^B\right)^m\right]}{(m-1)}\right\} = \log[1-(m-1)H^B] \tag{1.33}$$

Equation 1.31 is equal to the sum of Equations 1.32 and 1.33:

$$\log[1-(m-1)H^{A\cup B}] = \log[1-(m-1)H^A] + \log[1-(m-1)H^B] \tag{1.34}$$

Equation 1.34 can be recast as

$$1-(m-1)H^{A\cup B}=[1-(m-1)H^{A}][1-(m-1)H^{B}] \tag{1.35}$$

Equation 1.35 can be simplified as

$$1-(m-1)H^{A\cup B}=[1-(m-1)H^{A}-(m-1)H^{B}+(m-1)^{2}H^{A}H^{B}] \tag{1.36}$$

Equation 1.36 reduces to

$$H^{A\cup B}=H^{A}+H^{B}-[(m-1)H^{A}H^{B}] \tag{1.37}$$

Equation 1.37 is often expressed as

$$H(A+B)=H(A)+H(B)+(1-m)H(A)H(B) \tag{1.38}$$

Equation 1.38 can also be expressed as

$$\frac{\log[1+(1-m)H(A+B)]}{1-m}=\frac{\log[1+(1-m)H(A)]}{1-m}+\frac{\log[1+(1-m)H(B)]}{1-m} \tag{1.39}$$

In the limit as $m \to 1$, Equation 1.38 can be written as the sum of marginal entropies

$$H^{A\cup B}=H^{A}+H^{B} \quad \text{or} \quad H(A,B)=H(A)+H(B) \tag{1.40}$$

Equations 1.37 through 1.39 describe the additivity property. This property can be extended to any number of systems. In all cases, $H \geq 0$ (nonnegativity property). If systems A and B are correlated, then

$$p_{ij}^{A\cup B}\neq\left[\sum_{i=1}^{N}p_{ij}^{A\cup B}\right]\left[\sum_{j=1}^{M}p_{ij}^{A\cup B}\right] \tag{1.41}$$

for all (i, j). One may define mutual information or transinformation S as

$$T\left[\left(p_{ij}^{A\cup B}\right)\right]=H^{A\cup B}\left[\left(p_{ij}^{A\cup B}\right)\right]-H^{A}\left[\left(\sum_{i=1}^{N}p_{ij}\right)\right]-H^{B}\left[\left(\sum_{j=1}^{M}p_{ij}\right)\right] \tag{1.42}$$

Considering Equation 1.42, $T(p_{ij}) = 0$ for all m, if X and Y are independent, and Equation 1.42 will reduce to Equation 1.38. For correlated X and Y, $T(p_{ij}) < 0$ for $m = 1$, and $T(p_{ij}) = 0$ for $m = 0$. For arbitrary values of m, it will be sensitive to p_{ij}; it can take on negative or positive values for both $m < 1$ and $m > 1$ with no particular regularity and can exhibit more than one extremum.

Example 1.4

Consider a system A that has two states with probabilities $p_1^A = 0.4$ and $p_2^A = 0.6$. Consider another system designated as B with two states having probabilities $p_1^B = 0.3$ and $p_1^B = 0.7$. Both systems are independent. Compute the joint Tsallis entropy of the two systems. Take $m = 3$. Also compute the Shannon entropy.

Solution

For system A, p_1^A, p_2^A and $p_1^A + p_2^A = 1.0$. Therefore,

$$H^A = \frac{1}{3-1}[(0.4-0.4^3)+(0.6-0.6^3)] = 0.36$$

$$H^B = \frac{1}{3-1}[(0.3-0.3^3)+(0.7-0.7^3)] = 0.315$$

$$H(A+B) = 0.36+0.315-(3-1)\times 0.36\times 0.315 = 0.448$$

The joint Shannon entropy can be computed as follows:

$$H^A = -[0.4\log_2 0.4 + 0.6\log_2 0.6] = 0.971$$

$$H^B = -[0.3\log_2 0.3 + 0.7\log_2 0.7] = 0.881$$

$$H(A+B) = 0.971+0.881 = 1.852$$

In this case, the Shannon entropy is much larger than the Tsallis entropy because m is much greater than unity.

1.3.5 Composibility

The entropy $H(A + B)$ of a system comprising two subsystems A and B can be computed from the entropies of subsystems, $H(A)$ and $H(B)$, and the entropy index m.

1.3.6 Interacting Subsystems

Consider a set of N possibilities arbitrarily separated into two subsystems with N_1 and N_2 possibilities, where $N = N_1 + N_2$. Defining $P_{N_1} = \sum_{i=1}^{N_1} p_i$ and $P_{N_2} = \sum_{j=1}^{N_2} p_j$, $P_{N_1} + P_{N_2} = 1 = \sum_{k=1}^{N} p_k$. It can be shown that

$$H(P_N) = H(P_{N_1}, P_{N_2}) + P_{N_1}^m H(\{p_i | P_{N_1}\}) + P_{N_2}^m H(\{p_j | P_{N_2}\}) \tag{1.43}$$

where $\{p_i|P_{N_1}\}$ and $\{p_j|P_{N_2}\}$ are the conditional probabilities. Note that $p_i^m > p_i$ for $m < 1$ and $p_i^m < p_i$ for $m > 1$. Hence, $m < 1$ corresponds to rare events and $m > 1$ frequent events (Tsallis, 2001). This property can be extended to any number R of interacting subsystems: $N = \sum_{j=1}^{R} N_j$. Then, defining $w_j = \sum_{i=1}^{N_j} p_i, j = 1, 2, \ldots, N_j$, $\sum_{j=1}^{N} w_j = 1$, Equation 1.43 can be generalized as

$$H(\{p_i\}) = H(\{w_j\}) + \sum_{j=1}^{R} w_j^m H(\{p_i|w_j\}) \tag{1.44}$$

Here, $p_j = w_j$.

Example 1.5

Consider a set of five possibilities, p_i = {0.1, 0.15, 0.2, 0.25, 0.3}, separated into two subsets $N_1 = 3$, $p_{N_1} = \{0.1, 0.15, 0.2\}$, and $N_2 = 2$, $p_{N_2} = \{0.25, 0.3\}$. Compute the Tsallis entropy for this system. Then, use Equation 1.43 to compute the Tsallis entropy and show that both ways the entropy is the same.

Solution

First, the Tsallis entropy can be computed as

$$H(P_N) = \frac{1}{3-1}[(0.1 - 0.1^3) + (0.15 - 0.15^3) + (0.2 - 0.2^3) + (0.25 - 0.25^3) + (0.4 - 0.4^3)]$$
$$= 0.473$$

or consider as two subsystems as

$$P_{N_1} = \sum_{i=1}^{N_1} p_i = 0.1 + 0.15 + 0.2 = 0.45$$

$$P_{N_2} = \sum_{j=1}^{N_2} p_j = 0.25 + 0.3 = 0.55$$

$$H(\{p_i|P_{N_1}\}) = \frac{1}{3-1}\left\{\left[\frac{0.1}{0.45} - \left(\frac{0.1}{0.45}\right)^3\right] + \left[\frac{0.1}{0.45} - \left(\frac{0.1}{0.45}\right)^3\right] + \left[\frac{0.1}{0.45} - \left(\frac{0.1}{0.45}\right)^3\right]\right\}$$
$$= 0.432$$

$$H(\{p_i|P_{N_2}\}) = \frac{1}{3-1}\left\{\left[\frac{0.25}{0.55} - \left(\frac{0.25}{0.55}\right)^3\right] + \left[\frac{0.3}{0.55} - \left(\frac{0.3}{0.55}\right)^3\right]\right\} = 0.372$$

Thus, using Equation 1.43,

$$\begin{aligned} H(P_N) &= H(P_{N_1}, P_{N_2}) + P_{N_1}^m H(\{p_i|P_{N_1}\}) + P_{N_2}^m H(\{p_j|P_{N_2}\}) \\ &= 0.371 + 0.45^3 \times 0.432 + 0.55^3 \times 0.372 = 0.473 \end{aligned}$$

It shows that the entropy computed from Equation 1.43 is the same as given by the definition.

1.3.7 Other Features

Many complex systems exhibit a power like behavior and they may be in stationary but nonequilibrium states. This may often be the case for geomorphological systems. The Tsallis statistics (Tsallis, 2004) is particularly useful for describing such systems. This statistics exhibits three interesting features (Ferri et al., 2010). First, the PDFs, based on the Tsallis entropy, that describe metastable or stationary systems are proportional to what is called m-exponential defined as

$$\exp_m(-\alpha x) = [1-(1-m)\alpha x]^{1/(1-m)} \tag{1.45}$$

in which m and α are constants. Figure 1.7 shows a plot of Equation 1.45 for different values of α and m. In the limit $m \to 1$, m-exponential becomes the ordinary exponential, that is $\exp_1(x) = \exp(x)$. Further, if $m \to 1$ and $x = y^2$ then $\exp_m(-\alpha x)$ becomes an m-Gaussian.

The inverse of m-exponential is referred to as m-logarithm defined as

$$\ln_m(x) = \frac{x^{1-m}-1}{1-m}, \quad \ln_1(x) = \ln(x), \quad \ln_m[\exp_m(x)] = \exp[\ln_m(x)] = 1 \tag{1.46}$$

Stationary systems are characterized by nonextensivity index $m = m_{stat}$. Figure 1.8 shows a plot of Equation 1.46 for different values of m.

Second, stationary states show m-exponential sensitivity to initial conditions or weak chaos with a parameter $m = m_{sens}$. This means that small differences between adjacent states grow in an m-exponential fashion. Third, microscopic variables decrease m-exponentially with a parameter $m = m_{rel}$.

In this manner, a stationary or metastable system can be characterized by a triplet of m values, often referred to as the Tsallis m-triplet, that is $(m_{stat}, m_{sens}, m_{rel}) \neq (1, 1, 1)$, in which $m_{stat} > 1$, $m_{sens} < 1$, and $m_{rel} < 1$ (Ferri et al., 2010). Ausloos and Petroni (2007) and Petroni and Ausloos (2007) reported the values of m_{stat} for daily variation of the El Nino Southern Oscillation (ENSO) index.

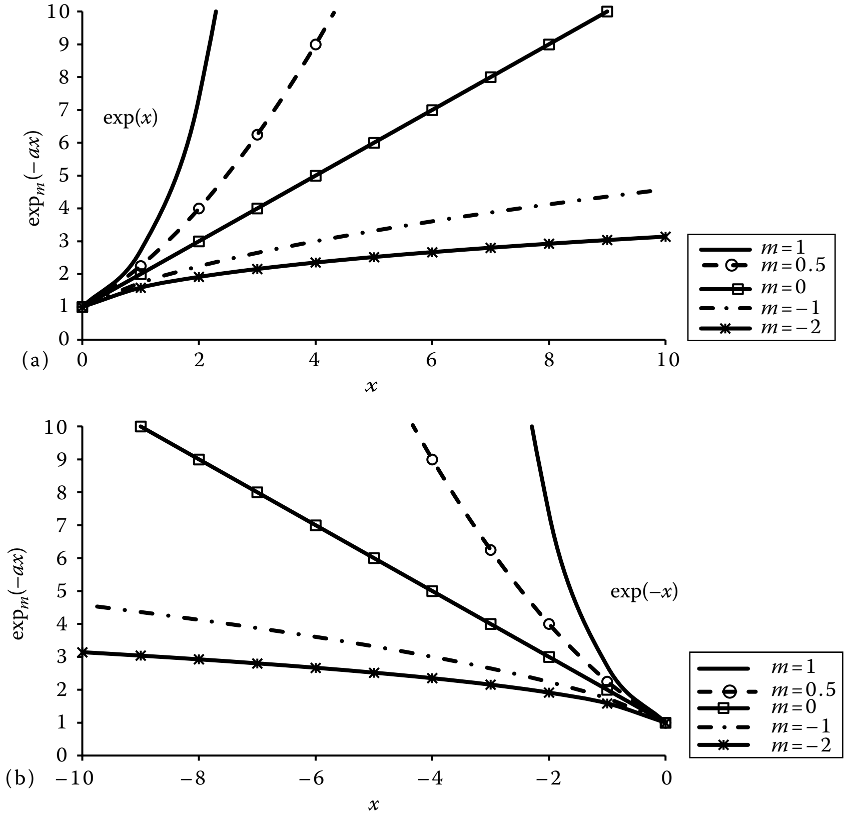

FIGURE 1.7 m-Exponential for various m values with (a) $a = -1$ and (b) $a = 1$.

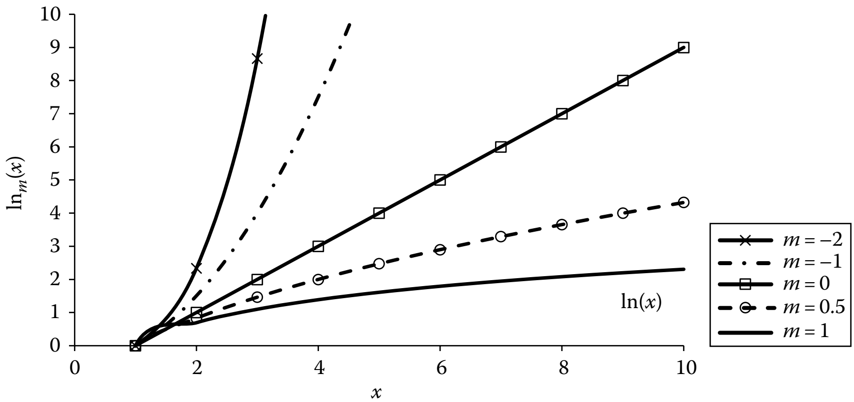

FIGURE 1.8 m-Logarithm for various m values.

1.4 MODIFICATION OF TSALLIS ENTROPY

Yamano (2001a) provided a modification of the Tsallis entropy. It may be worth recalling that the Shannon entropy function is uniquely determined not because of the definition of the mean value of information but because of the additivity of the uncertainty of information that the source contains. Considering the amount of information as the *m*-logarithmic function of probability

$$I_m(p) = -\ln_m p(x) \tag{1.47}$$

where

$$\ln_m p(x) = \frac{(p(x^{1-m})-1)}{1-m} = \frac{1-p(x^{1-m})}{m-1} \tag{1.48}$$

Function $I_m(p)$ is a monotonically decreasing function and so is –ln *p*. The unit of measurement in this case is nat, not bit. In the limit, as *m* tends to 1, the information content becomes –ln *p*.

Taking the normalized *m*-average (or escort average) of the information content or entropy, one obtains

$$\frac{-\sum_{i=1}^{N} p^m(x_i)\ln_m p(x_i)}{\sum_{i=1}^{N} p^m(x_i)} = \frac{1-\sum_{i=1}^{N} p^m(x_i)}{(m-1)\sum_{i=1}^{N} p^m(x_i)} = H_m(X) \tag{1.49}$$

which is the modified form of the Tsallis entropy and is obtained by dividing the Tsallis entropy by factor $\sum_{i=1}^{N} p^m(x_i)$.

Example 1.6

Let $N = 3$ and $p(x_i) = \{0.2, 0.3, 0.5\}$. Compute the *m*-average entropy and ordinary entropy.

Solution

Let $m = 3$, the *m*-average entropy is computed as

$$H_m(X) = \frac{1-(0.2^3+0.3^3+0.5^3)}{(3-1)(0.2^3+0.3^3+0.5^3)} = 2.625$$

and the ordinary entropy is

$$H(X) = \frac{1}{(3-1)}[(0.2-0.2^3)+(0.3-0.3^3)+(0.5-0.5^3)] = 0.42$$

where the factor

$$\sum_{i=1}^{N} p^m(x_i) = (0.2^3 + 0.3^3 + 0.5^3) = 0.16 = \frac{H(X)}{H_m(X)}$$

Yamano (2001b) discussed the properties of the modified Tsallis entropy, which are briefly presented in the following. For two random variables X and Y, their joint entropy can be expressed as

$$H_m(X,Y) = \frac{1 - \sum_{x,y} p^m(x,y)}{(m-1)\sum_{x,y} p^m(x,y)} \tag{1.50}$$

and a nonadditive conditional entropy $H_m(Y|x)$ can be written as

$$\frac{\sum_{i=1}^{N} (p^m(x_i)/[1-(m-1)H_m(Y|x)])}{\sum_{i=1}^{N} p^m(x_i)} = [1+(m-1)H_m(Y|X)]^{-1} \tag{1.51}$$

The mutual information $T_m(Y;X)$ can now be defined in the usual way as common information between X and Y, which is equal to the reduction in uncertainty in one variable due to the knowledge of another variable:

$$T_m(Y;X) = H_m(Y) - H_m(Y|X) = \frac{H_m(X) + H_m(Y) - H_m(X,Y) + (m-1)H_m(X)H_m(Y)}{1+(m-1)H_m(X)} \tag{1.52}$$

This will converge to the usual mutual information or transinformation in the additive limit m tending to 1. Following Yamano (2001b), the following relations hold for X, Y, and Z random variables:

1. $$H_m(X;Y) = H_m(X) + H_m(Y|X) + (m-1)H_m(X)H_m(Y|X) \tag{1.53}$$

2. $$H_m(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} [1+(m-1)H_m(X_{i-1},\ldots,X_1)]H_m(X_i|X_{i-1},\ldots,X_1) \tag{1.54}$$

3. $$H_m(X_1, X_2, \ldots, X_n) \le \sum_{i=1}^{n} [1+(m-1)H_m(X_{i-1},\ldots,X_1)]H_m(X_i) \tag{1.55}$$

4. $$T_m(X;Y) = H_m(X) - [1+(m-1)H_m(Y)]H_m(X|Y) \tag{1.56}$$

The mutual information becomes symmetric in X and Y:

$$T_m(X,Y) = T_m(Y,X) = H_m(X) + H_m(Y) - H_m(X,Y) \tag{1.57}$$

5.
$$[1+(m-1)H_m(X)]H_m(Y|X)$$
$$=H_m(Y,Z|X)-H_m(Z|Y,X)+(m-1)\{H_m(X)H_m(Y,Z|X)-H_m(X,Y)H_m(Z|Y,X)Z\} \tag{1.58}$$

It is seen that mutual information becomes symmetric in X and Y. In the limit m tending to 1, these relations reduce to the ones satisfied by the Shannon entropy.

6. The Kullback–Leibler (KL) cross entropy between two distributions $p(x)$ and $q(x)$ can be written in a Tsallis entropy sense as

$$D_m[p(x),q(x)] = \frac{\sum_{i=1}^{N} q^m(x_i)\ln_m q(x_i)}{\sum_{i=1}^{N} q^m(x_i)} - \frac{\sum_{i=1}^{N} p(x_i)\ln_m p(x_i)}{\sum_{i=1}^{N} p^m(x_i)} \tag{1.59}$$

The KL cross entropy satisfies

$$D_m[p(x),q(x)]\begin{cases}\geq 0\ (m>0)\\ <(m<0)\end{cases} \tag{1.60}$$

and equals 0 if $p(x) = q(x)$.

7. The generalized mutual information can be defined in terms of the generalized KL cross entropy as

$$T_m(X,Y) = D_m[P(x,y)|P(x)P(y)]$$
$$= \frac{1/(1-m)\left[1-\sum_{x,y} p(x,y)(p(x)p(y)/p(x,y))^{1-m}\right]}{\sum_{x,y} p^m(x,y)} \tag{1.61}$$

1.5 MAXIMIZATION

Consider a case where H given by Equation 1.2 is to be maximized subject to the following constraints:

$$\sum_{i=1}^{N} P_i = 1 \tag{1.62}$$

and

$$\sum_{i=1}^{N} p_i x_i = \overline{x} \tag{1.63}$$

where $\{x_i\}$ and $\bar{x}$ are real numbers. Following the method of the Lagrange multipliers, the Lagrange function can be defined as

$$L = H + \lambda_0\left[\sum_{i=1}^{N} p_i - 1\right] + \lambda_1\left[\sum_{i=1}^{N} p_i x_i - \bar{x}\right] \tag{1.64}$$

where λ_0 and λ_1 are the Lagrange multipliers. Following Tsallis (1988), Equation 1.64 can be recast as

$$L = H + \lambda_0\sum_{i=1}^{N} p_i + \lambda_0\lambda_1(m-1)\sum_{i=1}^{N} p_i x_i - [\lambda_0 + \lambda_0\lambda_1(m-1)\bar{x}] \tag{1.65}$$

It may be noted that the term within brackets on the right side of Equation 1.65 does not influence the maximization of entropy. Therefore, for entropy maximizing Equation 1.65 can simply be written as

$$L = H + \lambda_0\sum_{i=1}^{N} p_i + \lambda_0\lambda_1(m-1)\sum_{i=1}^{N} p_i x_i \tag{1.66}$$

Differentiating L in Equation 1.66 with respect to p_i and equating to zero for all i, one obtains

$$p_i = \frac{[1-\lambda_1(m-1)x_i]^{1/(m-1)}}{Z} \tag{1.67}$$

where Z is the partition function defined as

$$Z = \sum_{i=1}^{N} [1-\lambda_1(m-1)x_i]^{1/(m-1)} \tag{1.68}$$

If m tends to one, Equation 1.67 reduces

$$p_i = \frac{1}{Z}\exp(-\lambda_1 x_i) \tag{1.69}$$

in which

$$Z = \sum_{i=1}^{N} \exp(-\lambda_1 x_i) \tag{1.70}$$

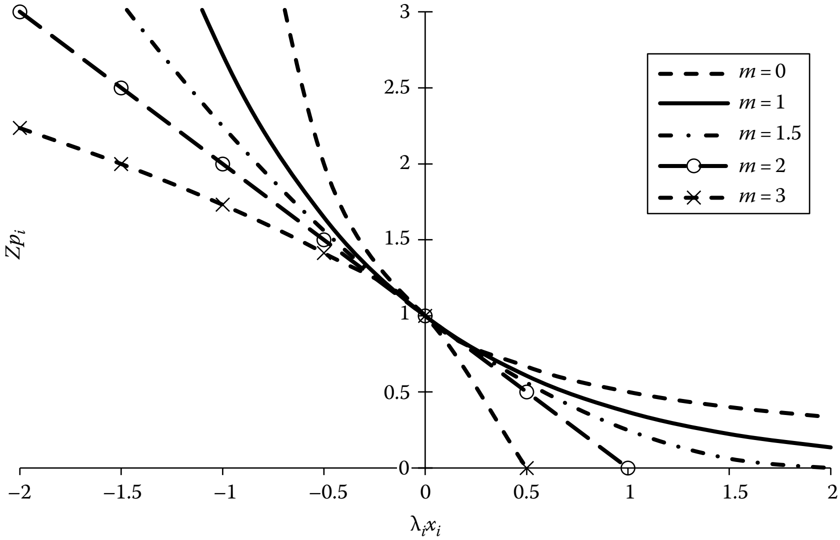

FIGURE 1.9 Plot of distribution given by Equation 1.67 parameterized by m.

Equation 1.67 expresses a power law distribution (Tsallis et al., 1998; Evans et al., 2000). This suggests that one way to obtain a power distribution is to extremize the Tsallis entropy with the constraint: $\sum_{i=1}^{N} p_i x_i^m = \overline{x^m}$, instead of $\bar{x}$. This distribution is plotted in Figure 1.9 for $m = 0, 1, 1.5, 2, 3$; the x-axis is taken as $\lambda_1 x_i$ and the y-axis is taken as Zp_i. For $m = 1$, this leads to an exponential distribution. For $m > 1$, it shows a cutoff at $\lambda_1 x_i = 1/(m - 1)$, where the slope is 0 for $m < 2$, -1 for $m = 2$, and $-\infty$ for $m > 2$ and diverges for $\lambda_1 x_i$ tending to $-\infty$. For $m < 1$, the distribution diverges at $\lambda_1 x_i = -1/(1 - m)$ and vanishes when $\lambda_1 x_i$ tends to $+\infty$.

1.6 PARTIAL TSALLIS ENTROPY

Let H_i denote the Tsallis entropy for the ith system state whose probability is p_i. Then, the Tsallis entropy for the system can be expressed as

$$H = \sum_{i=1}^{N} H_i = \sum_{i=1}^{N} H(p_i), \quad H_i = H(p_i) \tag{1.71}$$

The partial Tsallis entropy can be defined as (Niven, 2004)

$$H_i = -p_i^m \ln_m p_i = \frac{p_i - p_i^m}{m-1} \tag{1.72}$$