Springer Series in Statistics

Roger B. Nelsen

An Introduction to Copulas

Second Edition

Springer

Springer Series in Statistics

Roger B. Nelsen

An Introduction to Copulas

Second Edition

Springer

Springer Series in Statistics

Advisors:
P. Bickel, P. Diggle, S. Fienberg, U. Gather,
I. Olkin, S. Zeger

Springer Series in Statistics

Alho/Spencer: Statistical Demography and Forecasting.
Andersen/Borgan/Gill/Keiding: Statistical Models Based on Counting Processes.
Atkinson/Riani: Robust Diagnostic Regression Analysis.
Atkinson/Riani/Cerioli: Exploring Multivariate Data with the Forward Search.
Berger: Statistical Decision Theory and Bayesian Analysis, 2nd edition.
Borg/Groenen: Modern Multidimensional Scaling: Theory and Applications, 2nd edition.
Brockwell/Davis: Time Series: Theory and Methods, 2nd edition.
Bucklew: Introduction to Rare Event Simulation.
Cappé/Moulines/Rydén: Inference in Hidden Markov Models.
Chan/Tong: Chaos: A Statistical Perspective.
Chen/Shao/Ibrahim: Monte Carlo Methods in Bayesian Computation.
Coles: An Introduction to Statistical Modeling of Extreme Values.
David/Edwards: Annotated Readings in the History of Statistics.
Devroye/Lugosi: Combinatorial Methods in Density Estimation.
Efromovich: Nonparametric Curve Estimation: Methods, Theory, and Applications.
Eggermont/LaRiccia: Maximum Penalized Likelihood Estimation, Volume I: Density Estimation.
Fahrmeir/Tutz: Multivariate Statistical Modelling Based on Generalized Linear Models, 2nd edition.
Fan/Yao: Nonlinear Time Series: Nonparametric and Parametric Methods.
Farebrother: Fitting Linear Relationships: A History of the Calculus of Observations 1750-1900.
Federer: Statistical Design and Analysis for Intercropping Experiments, Volume I: Two Crops.
Federer: Statistical Design and Analysis for Intercropping Experiments, Volume II: Three or More Crops.
Ghosh/Ramamoorthi: Bayesian Nonparametrics.
Glaz/Naus/Wallenstein: Scan Statistics.
Good: Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses, 2nd edition.
Good: Permutation Tests: Parametric and Bootstrap Tests of Hypotheses, 3rd edition.
Gouriéroux: ARCH Models and Financial Applications.
Gu: Smoothing Spline ANOVA Models.
Györfi/Kohler/Krzyżak/Walk: A Distribution-Free Theory of Nonparametric Regression.
Haberman: Advanced Statistics, Volume I: Description of Populations.
Hall: The Bootstrap and Edgeworth Expansion.
Härdle: Smoothing Techniques: With Implementation in S.
Harrell: Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis.
Hart: Nonparametric Smoothing and Lack-of-Fit Tests.
Hastie/Tibshirani/Friedman: The Elements of Statistical Learning: Data Mining, Inference, and Prediction.
Hedayat/Sloane/Stufken: Orthogonal Arrays: Theory and Applications.
Heyde: Quasi-Likelihood and its Application: A General Approach to Optimal Parameter Estimation.

(continued after index)

Roger B. Nelsen

An Introduction to Copulas

Second Edition

Springer

Roger B. Nelsen
Department of Mathematical Sciences
Lewis & Clark College, MSC 110
Portland, OR 97219-7899
USA
nelsen@lclark.edu

Library of Congress Control Number: 2005933254

ISBN-10: 0-387-28659-4
ISBN-13: 978-0387-28659-4

Printed in the United States of America. (SBA)

9 8 7 6 5 4 3 2 1

springeronline.com

To the memory of my parents
Ann Bain Nelsen
and
Howard Ernest Nelsen

Preface to the First Edition

In November of 1995, I was at the University of Massachusetts in Amherst for a few days to attend a symposium held, in part, to celebrate Professor Berthold Schweizer's retirement from classroom teaching. During one afternoon break, a small group of us were having coffee following several talks in which copulas were mentioned. Someone asked what one should read to learn the basics about copulas. We mentioned several references, mostly research papers and conference proceedings. I then suggested that perhaps the time was ripe for "someone" to write an introductory-level monograph on the subject. A colleague, I forget who, responded somewhat mischievously, "Good idea, Roger—why don't *you* write it?"

Although flattered by the suggestion, I let it lie until the following September, when I was in Prague to attend an international conference on distributions with fixed marginals and moment problems. In Prague, I asked Giorgio Dall'Aglio, Ingram Olkin, and Abe Sklar if they thought that there might indeed be interest in the statistical community for such a book. Encouraged by their responses and knowing that I would soon be eligible for a sabbatical, I began to give serious thought to writing an introduction to copulas.

This book is intended for students and practitioners in statistics and probability—at almost any level. The only prerequisite is a good upper-level undergraduate course in probability and mathematical statistics, although some background in nonparametric statistics would be beneficial. Knowledge of measure-theoretic probability is not required.

The book begins with the basic properties of copulas and then proceeds to present methods for constructing copulas and to discuss the role played by copulas in modeling and in the study of dependence. The focus is on bivariate copulas, although most chapters conclude with a discussion of the multivariate case. As an introduction to copulas, it is not an encyclopedic reference, and thus it is necessarily incomplete—many topics that could have been included are omitted. The reader seeking additional material on families of continuous bivariate distributions and their applications should see (Hutchinson and Lai 1990); and the reader interested in learning more about multivariate copulas and dependence should consult (Joe 1997).

There are about 150 exercises in the book. Although it is certainly not necessary to do all (or indeed any) of them, the reader is encouraged to read through the statements of the exercises before proceeding to the next section or chapter. Although some exercises do not add

anything to the exposition (e.g., "Prove Theorem 1.1.1"), many present examples, counterexamples, and supplementary topics that are often referenced in subsequent sections.

I would like to thank Lewis & Clark College for granting me a sabbatical leave in order to write this book; and my colleagues in the Department of Mathematics, Statistics, and Computer Science at Mount Holyoke College for graciously inviting me to spend the sabbatical year with them. Thanks, too, to Ingram Olkin for suggesting and encouraging that I consider publication with Springer's *Lecture Notes in Statistics*; and to John Kimmel, the executive editor for statistics at Springer, for his valuable assistance in the publication of this book.

Finally, I would like to express my gratitude and appreciation to all those with whom I have had the pleasure of working on problems related to copulas and their applications: Claudi Alsina, Jerry Frank, Greg Fredricks, Juan Quesada Molina, José Antonio Rodríguez Lallena, Carlo Sempi, Abe Sklar, and Manuel Úbeda Flores. But most of all I want to thank my good friend and mentor Berthold Schweizer, who not only introduced me to the subject but also has consistently and unselfishly aided me in the years since and who inspired me to write this book. I also want to thank Bert for his careful and critical reading of earlier drafts of the manuscript and his invaluable advice on matters mathematical and stylistic. However, it goes without saying that any and all remaining errors in the book are mine alone.

Roger B. Nelsen
Portland, Oregon
July 1998

Preface to the Second Edition

In preparing a new edition of *An Introduction to Copulas*, my goals included adding some topics omitted from the first edition while keeping the book at a level appropriate for self-study or for a graduate-level seminar. The major additions in the second edition are sections on:

- a copula transformation method;
- extreme value copulas;
- copulas with certain analytic or functional properties;
- tail dependence; and
- quasi-copulas.

There are also a number of new examples and exercises and new figures, including scatterplots of simulations from many of the families of copulas presented in the text. Typographical errors in the first edition have been corrected, and the references have been updated.

Thanks again to Lewis & Clark College for granting me a sabbatical leave in order to prepare this second edition; and to the Department of Mathematics and Statistics at Mount Holyoke College for again inviting me to spend the sabbatical year with them. Finally, I would like to thank readers of the first edition who found numerous typographical errors in the first edition and sent me suggestions for this edition.

Roger B. Nelsen
Portland, Oregon
October 2005

Contents

1 Introduction

The study of copulas and their applications in statistics is a rather modern phenomenon. Until quite recently, it was difficult to even locate the word "copula" in the statistical literature. There is no entry for "copula" in the nine volume *Encyclopedia of Statistical Sciences*, nor in the supplement volume. However, the first update volume, published in 1997, does have such an entry (Fisher 1997). The first reference in the *Current Index to Statistics* to a paper using "copula" in the title or as a keyword is in Volume 7 (1981) [the paper is (Schweizer and Wolff 1981)]—indeed, in the first eighteen volumes (1975-1992) of the *Current Index to Statistics* there are only eleven references to papers mentioning copulas. There are, however, 71 references in the next ten volumes (1993-2002).

Further evidence of the growing interest in copulas and their applications in statistics and probability in the past fifteen years is afforded by five international conferences devoted to these ideas: the "Symposium on Distributions with Given Marginals (Fréchet Classes)" in Rome in 1990; the conference on "Distributions with Fixed Marginals, Doubly Stochastic Measures, and Markov Operators" in Seattle in 1993; the conference on "Distributions with Given Marginals and Moment Problems" in Prague in 1996; the conference on "Distributions with Given Marginals and Statistical Modelling" in Barcelona in 2000; and the conference on "Dependence Modelling: Statistical Theory and Applications in Finance and Insurance" in Québec in 2004. As the titles of these conferences indicate, copulas are intimately related to study of distributions with "fixed" or "given" marginal distributions. The published proceedings of the first four conferences (Dall'Aglio et al. 1991; Rüschendorf et al. 1996; Beneš and Štěpán 1997; Cuadras et al. 2002) are among the most accessible resources for the study of copulas and their applications.

What are copulas? From one point a view, copulas are functions that join or "couple" multivariate distribution functions to their one-dimensional marginal distribution functions. Alternatively, copulas are multivariate distribution functions whose one-dimensional margins are uniform on the interval (0,1). Chapter 2 will be devoted to presenting a complete answer to this question.

Why are copulas of interest to students of probability and statistics? As Fisher (1997) answers in his article in the first update volume of the *Encyclopedia of Statistical Sciences*, "Copulas [are] of interest to statisticians for two main reasons: Firstly, as a way of studying scale-free

measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions, sometimes with a view to simulation." These topics are explored and developed in Chapters 3, 4, and 5.

The remainder of this chapter will be devoted to a brief history of the development and study of copulas. Readers interested in first-hand accounts by some of those who participated in the evolution of the subject should see the papers by Dall'Aglio (1991) and Schweizer (1991) in the proceedings of the Rome conference and the paper by Sklar (1996) in the proceedings of the Seattle conference.

The word *copula* is a Latin noun that means "a link, tie, bond" (*Cassell's Latin Dictionary*) and is used in grammar and logic to describe "that part of a proposition which connects the subject and predicate" (*Oxford English Dictionary*). The word copula was first employed in a mathematical or statistical sense by Abe Sklar (1959) in the theorem (which now bears his name) describing the functions that "join together" one-dimensional distribution functions to form multivariate distribution functions (see Theorems 2.3.3 and 2.10.9). In (Sklar 1996) we have the following account of the events leading to this use of the term copula:

> Féron (1956), in studying three-dimensional distributions had introduced auxiliary functions, defined on the unit cube, that connected such distributions with their one-dimensional margins. I saw that similar functions could be defined on the unit n-cube for all $n \geq 2$ and would similarly serve to link n-dimensional distributions to their one-dimensional margins. Having worked out the basic properties of these functions, I wrote about them to Fréchet, in English. He asked me to write a note about them in French. While writing this, I decided I needed a name for these functions. Knowing the word "copula" as a grammatical term for a word or expression that links a subject and predicate, I felt that this would make an appropriate name for a function that links a multidimensional distribution to its one-dimensional margins, and used it as such. Fréchet received my note, corrected one mathematical statement, made some minor corrections to my French, and had the note published by the Statistical Institute of the University of Paris as Sklar (1959).

But as Sklar notes, the functions themselves predate the use of the term copula. They appear in the work of Fréchet, Dall'Aglio, Féron, and many others in the study of multivariate distributions with fixed univariate marginal distributions. Indeed, many of the basic results about copulas can be traced to the early work of Wassily Hoeffding. In (Hoeffding 1940, 1941) one finds bivariate "standardized distributions" whose support is contained in the square $[-1/2,1/2]^2$ and whose margins are uniform on the interval $[-1/2,1/2]$. (As Schweizer (1991) opines, "had Hoeffding chosen the unit square $[0,1]^2$ instead of $[-1/2,1/2]^2$ for his normalization, he would have discovered copulas.")