1 Copulas

Supervisor: Matthias Winkel

Prerequisites: BS3a Applied Probability

Type: Theoretical project with simulation component if desired

Description: Copulas describe the dependence structure of multivariate distributions (or random variables), without reference to the marginal distributions. More formally, copulas are multivariate distribution functions of random variables $(U_1, \ldots, U_n)$ on $[0,1]^n$, for which each U_i is uniformly distributed. Sklar's Theorem states that a (continuous) multivariate distribution function H can be uniquely expressed in terms of its marginal distributions and a copula. The project could start by working out a detailed proof of this theorem.

In practice, one needs specific families of copulas just as one works with specific families of distributions and there is a multitude of methods to construct copulas. The project will include the exposition of some of these. There are also several summary statistics of dependence structure such as the well-known correlation coefficient. In the context of copulas, the most important ones are Kendall's τ and Spearman's ρ.

Copulas have become very important in finance and insurance over the last few years. In insurance or risk management it is an important observation that the marginal distributions do not describe joint risks adequately; e.g., several branches of an insurance company are hit by the same events: a storm can damage cars and houses and injure people, and the insurer's risks associated with these are not independent. Similarly, in finance, prices of different stocks do not move independently; e.g. when the situation on oil markets is uncertain, a lot of stock prices react simultaneously and to a degree that depends on how important oil is for the underlying companies. The project could include a discussion of applications in insurance and/or finance.

The project will not be data-based, but there could be a simulation component.

References:

- R.B. Nelsen: *An Introduction to Copulas* . Springer Lecture Notes in Statistics 139 (1999)

- R.W.J. van der Goorbergh, C. Genest, B.J.M. Werker: Bivariate option pricing using dynamic copula models. *Insurance Math. Econom.* **37** (2005), no. 1, 101-114

- P. Embrechts, A.J. McNeil, D. Straumann: Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond (Cambridge, 1998)*, 176-223, CUP, Cambridge 2002

2 Branching processes and random trees

Supervisor: Matthias Winkel

Prerequisites: BS3a Applied Probability, B10a Martingales Through Measure Theory desirable

Type: Theoretical project with simulation component if desired

Description: Galton-Watson processes are discrete-time Markov chains that model population sizes. One should think of time as generation. The rules are such that one starts with an ancestor; each individual gives birth to an independent number of offspring, drawn from an offspring distribution on $\mathbb{N}$. Depending on the mean number of offspring, the population will die out in finite time or survive forever with positive probability. In the first case one can e.g. study the extinction time, in the second one can study the asymptotic behaviour of the population size and establish limit results.

There are variants of the model such as several types of individuals or lifetimes for the individuals. The latter leads to continuous-time Markov chains. There are even analogues with continuous population sizes (just as continuous random variables, in general, are often useful as approximations of large discrete random variables). Each of this could provide a possible direction of the project.

More explicitly, one can study the family tree of such a model (which contains more information than just the population size process), called a Galton-Watson tree. There are different codings of such random trees, and a possible direction of the project could be to study some of these.

References:

- T.E. Harris: *The theory of branching processes.* Springer 1963
- K.B. Athreya, P.E. Ney: *Branching processes.* Springer 1972
- S. Asmussen, H. Hering: *Branching processes.* Birkhäuser 1983
- J.-F. Le Gall: Random trees and spatial branching processes. *MaPhySto Lecture Notes no.9* (2000)

3 The Alpha Model

Supervisor: Matthias Winkel

Prerequisites: Some of the possible directions would benefit from B10a Martingales through measure theory. Part A Graph Theory useful but not essential.

Type: Project with possibilities for theoretical study, simulation and probably fitting models to data

Description: The Alpha Model was proposed by Daniel Ford as a model for phylogenetic trees that describe the genealogical tree relating species. Ford fitted the model to published phylogenies as available on the internet in a database called TreeBASE.

An Alpha Model tree is a random rooted binary combinatorial tree T_n with n leaves (cladogram) constructed inductively as follows. For $n = 1$ and $n = 2$ there is only one such tree, looking like an I and a Y, respectively, the vertex at the bottom being the root, vertices at the top being called leaves. Given T_n, a tree T_{n+1} is constructed as follows. For some parameter $\alpha \in (0, 1)$, associate weights with edges, weight $1-\alpha$ each with edges between a leaf and another vertex, weight α each with all other edges. Pick an edge at random with probability mass function proportional to the weights. Add two vertices. Replace the chosen edge by two edges connecting the end points to one of the new vertices and add a further edge between the new vertices.

The plan of this project is to carefully set up the model, study its properties, possibly simulate and/or fit the model to some data. There is a wide range of possibilities on the theoretical side including combinatorial and asymptotic properties, with some recent literature, to a large part at an accessible level, challenging in other parts. The data analysis side would repeat Ford's study on some more trees and compare with other models such as Aldous's beta-splitting model; further questions will arise while doing this.

References:

- D. Aldous: Probability distributions on cladograms. In *Random discrete structures (Minneapolis, MN, 1993)*, volume 76 of *IMA Vol. Math. Appli.*, pages 1017. Springer, New York, 1996.

- D. Ford: Probabilities on cladograms: introduction to the alpha model. 2005. *Preprint*, arXiv:math.PR/0511246

- J. Pitman: *Combinatorial Stochastic Processes.* Ecole d'été de Probabilités de St-Flour XXXII, *to appear in Lecture Notes in Mathematics, Springer 2006*