Copula Modeling for Clinical Trials

Nathan T. James

October 12, 2018

1 Introduction

Background

Clinical trials are considered the gold standard for evaluated of medical intervention [cite].

Principles of beneficence, etc.. (check ICH guidelines) mean that all interventions must in principle be evaluated for both efficacy and safety and further these benefits must outweigh the risks [cite]. Traditionally the efficacy and safety effects are evaluated using a series of trials with four sequential phases. Phase I are first-in-human trials designs usually performed on a small number of healthy participants to determine dosing (and other PK/PD?). Phase II trials are often first to test intervention in patients with the disease of interest and are used to determine efficacy and make go/no-go decision about whether to continue drug development. Phase III trials confirm the efficacy found in early studies and gather more information about adverse events. These studies are usually the pivotal studies required by regulatory agencies to approve a drug for marketing. Finally, Phase IV studies are ongoing surveillance ... [cite]

Usually the effect of treatment on efficacy and safety outcomes are explored with separate models for each type of outcome

Further, many studies include several outcomes (endpoints) of interest represent the effects of the intervention on a multiple (pathologies/disease processes?) and interest may

Same with multiple types of adverse event

For all these situations, modeling outcomes jointly uses data more efficiently [CITE] and provides a better characterization of the intervention across multiple domains[word choice?] simultaneously

(Can occur at multiple phases - phase I/II dosing toxicity/efficacy, phase III benefit-risk , phase IV? adverse risk assessment)

Several strategies have been proposed for such joint models including blah, blah [cite]. In this report I will focus on use of copulas for joint models[2].

Copulas

In the usual case, each outcome is modeled by a univariate function of treatment and and other covariates

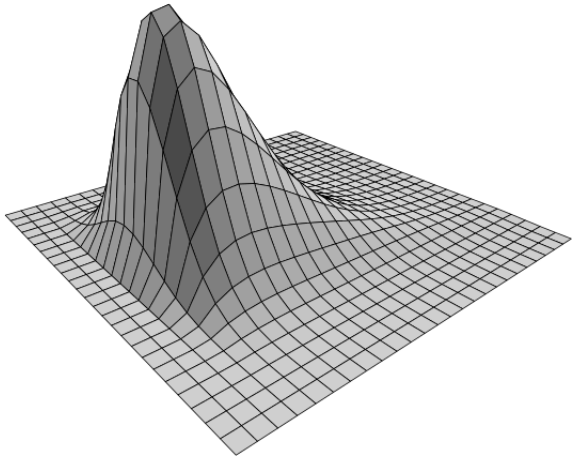

Figure 1: Gumbel copula

A copula is a mathematical function that combines univariate distributions to create a multivariate distribution. Importantly, Sklar's Theorem[cite] guarantees that a proper multivariate distribution function can be constructed by combining any univariate distribution functions (any?? check discrete) and a copula

this is an example of a citation [6].

A distinct advantage of the copula approach is the ability to "couple" arbitrary univariate models, thereby allowing modeling of the marginal univariate outcomes to be separated from modeling of the copula, which contains information about the correlation (more general term) between the outcomes

Further details on copulas and copula regression are presented below

this is another citation [5]

We can also include a picture as seen in figure 1. it's a placeholder for now

Inference

Inference on the resulting multivariate distribution function can be used to determine how the outcomes are correlated and can be performed under several paradigms

-Frequentist

-Bayesian

Principled method of combining prior beliefs concerning parameters with additional data

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)\,d\theta} \tag{1}$$

In context of clinical trials - include elicited clinical expertise, previous (historical) studies

Inference based on posterior probabilities more interpretable

Table 1: Popular Copulas

Independence	Normal
t-Copula	other elliptical??
Archimedean	Clayton
Comprehensive	

Adaptive designs??

-Likelihood??

2 Joint Copula models

Copula families

Much more background on copulas

Elliptical

Archimedean

Combinations/Comprehensive

Here is some additional text and we can also reference a table (see table 1). The incivility was the sum of features present within a block [3].

Copula regression

This is a whole section on copula regression

$$E(Z(s)) = \mu(s) = \mu, \text{for all } s \in D \tag{2}$$

$$Var(Z(s)) = \sigma^2(s) = \sigma^2, \text{for all } s \in D \tag{3}$$

$$Cov(Z(s_1), Z(s_2)) = C(s_i - s_j), \text{for all } s_i \neq s_j \in D \tag{4}$$

Equations 2 and 3 ensure that the mean and variance are constant and independent of location throughout the region D. Equation 4 ensures that the covariance depends only on the difference between two locations, rather than the locations themselves. $C(\cdot)$ in equation 4 is called the covariance function.

([7], [1]).

We use the R package "copula" [4].

3 Applications

References

[1] Swati Biswas, Diane D Liu, J Jack Lee, and Donald A Berry. Bayesian clinical trials at the university of texas m. d. anderson cancer center. 6(3):205–216.

[2] Maria J. Costa and Thomas Drury. Bayesian joint modelling of benefit and risk in drug development.

[3] Maria J. Costa, Weili He, Yannis Jemiai, Yueqin Zhao, and Carl Di Casoli. The case for a bayesian approach to benefit-risk assessment:: Overview and future directions. 51(5):568–574.

[4] Marius Hofert, Ivan Kojadinovic, Martin Maechler, and Jun Yan. copula: Multivariate dependence with copulas.

[5] Harry Joe. *Dependence modeling with copulas*. Number 134 in Monographs on statistics and applied probability. CRC Press, Taylor & Francis Group.

[6] Roger B. Nelsen. *An introduction to copulas*. Springer series in statistics. Springer, 2nd ed edition.

[7] Michael Stanley Smith. Bayesian approaches to copula modelling. page 33.

Appendix

Appendix Subsection 1

Next, a streetmap was developed based on the 2007 TIGER/Line shapefiles. The TIGER shapefiles containing the location of streets, rail lines, and other passages in Baltimore city were merged with an accompanying relationship table containing information about street name, direction, zip code, and address ranges for each block face as well as indicator fields for type of linear feature (road, rail, other). Utilizing these indicator fields, the dataset was narrowed to include only road segments.

Appendix Subsection 2

something more