Summaries Dissertation Bayesian Circular Statistics

Kees Mulder

5/18/2019

Contents

English 350 words

Researchers often analyze data that is either numerical, such as height in centimeters, or is divided into categories, such as level of education. However, you can also encounter observed angles, such as wind directions in degrees from north. In that case, the data are best visualized on a circle, and are therefore called *circular data*. Not just wind directions provide circular data. We run into them in almost all fields, from psychology to astronomy.

Why is circular data different? Moving one way around the circle means that at some point, we end up back where we started, because $0° = 360°$. As a result, a lot of the bread-and-butter statistician's toolkit, even something as simple as the mean, can not be used on circular data. Instead, we have to use specialized circular statistics methods.

We take a look at several applications, and provide new ways to analyze circular data for practical problems, usually using solutions from the field of *Bayesian statistics*.

- In cognitive psychology there are experiments with a circular outcome. To relate both numerical and categorical predictors to the circular outcome, we made a new *circular regression* model, which uses the predictors in a better way than earlier models.
- Testing whether a group of directions is spread evenly around the circle can be done better by using the a hypothesis test using the *Bayes Factor*.
- Different musical genres are listened to at different times of the day. We have created a model that can capture precisely in what way these differ.
- Eye movements also have a direction. By creating a method that can model these, we were able to test differences between infants and adults in the way they look around the world.
- In criminology, crime times over the 24-hour clock are circular observations. We show that the way they are treated can be improved, and provide new ways to do that.

Finally, we've created an `R` package, `circbayes`, that can perform these analyses in a user-friendly way. As a result, the field of Bayesian circular statistics has both been expanded in the scope of its analyses, as well as the accessibility of its methods.

English 250 words

Researchers often analyze data that is either numerical, such as height in centimeters, or is divided into categories, such as level of education. However, you can also encounter data like wind directions in degrees. Such data are best visualized on a circle, and are therefore called *circular data*. We run into this type of data in almost all fields, from psychology to astronomy.

Why is circular data different? Moving one way around the circle means that at some point, we end up back where we started, because $0° = 360°$. As a result, a lot of the statistician's toolkit, even something as simple as the mean, can not be used on circular data.

We take a look at several applications, and provide new ways to analyze circular data for practical problems, usually using solutions from *Bayesian statistics*.

For example, in cognitive psychology of haptic behavior there are experiments with a circular outcome. To relate both numerical and categorical predictors to the circular outcome, we made a new *circular regression* model, which uses the predictors in a better way than earlier models. Other problems we worked on are testing whether directions are spread evenly on the circle, analysis of the times at which people listened to certain music genres, models for eye movement directions obtained in eye tracking research, and modeling crime times in criminology.

Finally, we've created an `R` package, `circbayes`, that can perform these analyses in a user-friendly way. As a result, the field of Bayesian circular statistics has both been expanded in the scope of its analyses, as well as the accessibility of its methods.

Dutch 250 words

We analyseren vaak data die of numeriek is, zoals lengte in centimeters, of categorisch, zoals opleidingsniveau. Je kunt echter ook data tegenkomen zoals windrichtingen in graden. Zulk soort data kun je het best visualiseren op een cirkel, zoals een kompas. Daarom noemen we dit *circulaire data.* We lopen dit soort data overal tegen het lijf, van psychologie tot astronomie.

Maar wat is er dan anders aan circulaire data? Als je rond zou lopen op de cirkel, kom je op een gegeven moment weer terug waar je begonnen bent, omdat je een heel rondje hebt gelopen. Dat kun je ook zien omdat $0° = 360°$. Daardoor kunnen we zelfs simpele statistieken, zoals het gemiddelde, niet meer gebruiken op circulaire data.

Wij hebben nieuwe analysemethoden gemaakt voor de praktische problemen in de analyse van circulaire data. Vaak gebruiken we daar kennis uit de *Bayesiaanse statistiek*.

In cognitieve psychologie zijn er bijvoorbeeld experimenten met een circulaire uitkomst. Om te kunnen zien hoe die samenhangt met numerieke en categorische predictors, hebben we een *circulair regressiemodel* ontwikkeld, dat de predictors op een betere manier gebruikt dan eerdere modellen. Andere problemen waar we aan hebben gewerkt zijn het testen of een groep richtingen gelijk verdeeld is over de cirkel, analyse van de tijden waarop mensen naar bepaalde muziekgenres luisteren, modellen voor oogbewegingsrichtingen die we in *eye tracking* onderzoek tegenkomen, en het modelleren van misdaadtijden in criminologie.

Tot slot hebben we een `R` package gemaakt, `circbayes`, waarmee deze analyses op een gebruiksvriendelijke manier kunnen worden uitgevoerd. Het eindresultaat is dat we het veld van Bayesiaanse circulaire statistiek hebben uitgebreid in de mogelijke analyses, als ook de toegankelijkheid van de methoden.