Course Description: Bayesian data analysis and multilevel modeling
7.5hp

| Course leader     | Marco Tullio Liuzza | Phone: 08-164601  
Email: marco.tullio.liuzza@psychology.su.se |
|-------------------|----------------------|-------------------|
| Teachers          | Marco Tullio Liuzza  | Phone: 08-164601  
Email: marco.tullio.liuzza@psychology.su.se |
|                   | Mats Nilsson         | Phone: 08-164605  
Email: mnn@psychology.su.se |
|                   | Stefan Wiens         | Phone: 08-163933  
Email: sws@psychology.su.se |
| Course administrator | Monika Karlsson   | Phone: 08-163845  
Email: monika.karlsson@psychology.su.se |
| Examiner          | Marco Tullio Liuzza  | Phone: 08-164601  
Email: marco.tullio.liuzza@psychology.su.se |

Language: English
Course level: Doctoral level

Eligibility criteria: Accepted for studies at doctoral level within social sciences including public health

Main field of study: Psychology

Host department: Department of Psychology, Stockholm University

Sign up: http://www.psychology.su.se/utbildning/alla-
utbildningar/forskniv%C3%A5/studiehandbok-
forskniv%C3%A5/anm%C3%A4lan-till-kurser-inom-forskarutbildningen-
1.30731

Note that the number of participants will be limited and registration is on a first-come, first-served basis.
Bayesian data analysis and multilevel modeling

The course will start with a general seminar to discuss the benefits of Bayesian approaches to data analysis for Psychology scholars. The first week of the course covers Bayesian analysis: its theory, implementation in software (R and JASP), and applied use on the students’ own data. Even though other alternative approaches will be presented, we will focus on the Bayes Factor approach to hypothesis testing, because this approach is applicable to most of the research designs in Psychology, and is implemented in popular R packages such as BayesFactor (Morey and Rouder, 2015) and online calculators.

Besides the theoretical aspects, the course will offer actual programming in R to show how to generate priors, likelihood and posteriors on the basis of practical examples. Students will be introduced to the user-friendly, free software JASP (JASP Team, 2016).

The final goal of the first week is to allow students to perform Bayesian data analyses on their own data and to plan a design using a Sequential Bayes Factor Design analysis (Schönbrodt and Wagemakers, 2016).

The second week of the course covers multilevel models (MLMs) because these are more robust to unequal sample sizes, missing data, and extreme cases, as estimation of individual level and group level parameters inform each other in an updating process (e.g. German & Hill, 2007). Also, MLM can be nicely combined with Bayesian inference (McElreath, 2016). Thus, Bayesian approaches are useful to compare different multilevel models. Since “All models are wrong, but some are useful” (Box, 1976), model selection requires a good balance between correctly describing the data and predicting new ones.

Statistical inference is not a universal procedure that provides a surrogate of objective truth (Gigerenzer, & Marewski 2014). Abandoning this misbelief forces scholars to think carefully about how they are testing their hypotheses and what they can reasonably infer from their data. To illustrate the advantage of MLM, categorical data are often analyzed as accuracy rates (in percent), but Jaeger (2008) has shown that this approach can lead to spurious results, as linear models such ANOVAs rest on assumptions that are not tenable when dealing with categorical data. These issues are avoided in Generalized Linear Mixed Models (AKA Generalized Multilevel Models, GMLMs) for binomially distributed outcome. GMLMs combine the advantages of ordinary logit models with the ability to account for random subject and item effects at once. Also, the use of GMLMs provides a much more flexible approach that allows to model continuous variables on a trial-by-trial basis.
Therefore, the second week of the course introduces multilevel models (MLMs) more broadly and teaches how to implement them in the popular R package lme4 (Bates, Maechler, Bolker, Walker, 2015). Particular emphasis will be placed on (re)introducing linear models from scratch and showing how all the most popular analytic approaches used in Psychology (from t-tests to ANOVAs and ANCOVAs) are nothing but specific linear models. We will then focus on the random effects (intercepts, slopes and their covariance) in linear models and show how to model them in lme4. Last, we will reconcile the topics treated during the first week and multilevel modeling by setting and updating priors on MLMs parameters and will show how to adopt a Bayesian approach in model comparison.

**Expected learning outcomes**
After the course, students will be able to compute Bayes Factors for their own datasets and to perform a Bayesian multilevel model analysis.

**Examination**
At the end of each week participants will have to pass a quiz to assess their knowledge of the basic concepts of the course. One week after the end of the course participants will have to submit a final, written assignment that will require the analysis of real or simulated data which will be then reported as a results paragraph. Participants need to submit commented code in R and JAGS/Stan.

**Grade and grade criteria**
The course is graded on a pass/fail basis: Pass: For a passing grade, the doctoral student has completed the examination requirements and thereby shown that the expected learning outcomes are achieved. Fail: The examination task has been solved insufficiently, in such a way that the expected learning outcomes are not met.

**Dates**
The course will occur on weeks 11 and 12 of the year 2017. See preliminary schedule below.
Reading lists

Books


Articles


### Preliminary schedule

<table>
<thead>
<tr>
<th>Datum</th>
<th>Dag</th>
<th>Plats</th>
<th>Tid</th>
<th>Tema</th>
<th>Lärare</th>
<th>Undervisningstyp</th>
</tr>
</thead>
<tbody>
<tr>
<td>170502</td>
<td>Tis</td>
<td>GEL</td>
<td>13-16</td>
<td>Introduction and Probability</td>
<td>MTL</td>
<td>Förel.</td>
</tr>
<tr>
<td>170503</td>
<td>Ons</td>
<td>GEL</td>
<td>13-16</td>
<td>Simulations and Likelihood intervals</td>
<td>MN</td>
<td>Förel.</td>
</tr>
<tr>
<td>170504</td>
<td>Tor</td>
<td>GEL</td>
<td>13-16</td>
<td>Bayesian approaches</td>
<td>MTL</td>
<td>Förel.</td>
</tr>
<tr>
<td>170505</td>
<td>Fre</td>
<td>GEL</td>
<td>13-16</td>
<td>Bayes Factor</td>
<td>SW</td>
<td>Förel.</td>
</tr>
<tr>
<td>170508</td>
<td>Mån</td>
<td>GEL</td>
<td>13-16</td>
<td>Bayes Factor seminar</td>
<td>SW</td>
<td>Handl.möte</td>
</tr>
<tr>
<td>170509</td>
<td>Tis</td>
<td>GEL</td>
<td>13-16</td>
<td>Linear regression</td>
<td>MN</td>
<td>Förel.</td>
</tr>
<tr>
<td>170510</td>
<td>Ons</td>
<td>GEL</td>
<td>13-16</td>
<td>Multilevel models</td>
<td>MTL</td>
<td>Förel.</td>
</tr>
<tr>
<td>170511</td>
<td>Tor</td>
<td>GEL</td>
<td>13-16</td>
<td>Generalized Multilevel models</td>
<td>MTL</td>
<td>Förel.</td>
</tr>
<tr>
<td>170512</td>
<td>Fre</td>
<td>GEL</td>
<td>13-16</td>
<td>Bayesian multilevel models</td>
<td>MTL</td>
<td>Förel.</td>
</tr>
<tr>
<td>170515</td>
<td>Mån</td>
<td>GEL</td>
<td>13-16</td>
<td>Bayesian multilevel models</td>
<td>MTL</td>
<td>Handl.möte</td>
</tr>
</tbody>
</table>