Class 1 - Conceptual Exercises

Hey CogSci'ers :)

This Methods course is not only about showing you some new fancy analysis tricks - although there might be a bit of that too :)

Instead, the course is designed to reframe and better understand what we do when we do statistics.

Welcome to #TeamBayesian, sure :) But many things are more general as well.

Just like Richard McElreath's book, we will move between different levels:  
(that I just made these up but I think it helps frame the exercises)

**theory:** what are we trying to do with stats and modeling, how does it go well or not well, what should we be careful about etc.

**methodology:** how do different methods work concretely, what needs to happen for them to go well, what are their advantages and disadvantages

**implementation:** how do we write some code (in R) that makes nice things happen, and not something else

Today, we'll start off on the first two levels. As the course progresses, it will probably move more towards implementation.

There are exercises on the next pages.

Challenge tasks are optional and go a bit beyond the readings. They aren't necessarily difficult though.

# Exercise 1: Getting a vocabulary

Lots of terms and words are being used in statistics (and beyond). Sometimes they don't mean the same thing. Sometimes they mean something more specific than what we think. Sometimes people don't explain them, they just use them; usually it takes a bit of time (well spent!) to get comfortable with them.

Below, you find a list of terms and concepts. In Study Groups, talk about them one at a time and make sure you know what they mean, maybe make a little definition. Some will be harder than others, and some will be harder than they seem at first.

You have 15+- minutes. We will go through them quickly together afterwards. Continue with next exercise if you're done early

*What is:*

**Inference – Apply information already in your possession to a new situation. Making conclusions based on evidence.**

*What is:*

**Statistical inference – Making conclusions about population parameters based on sample statistics.**

*And how can statistics be used for inference?*

*What is:*

**Statistics – Doing sth with quantitative data.**

*What is the difference between*:

**Descriptive statistics** *and*

**Inferential statistics**

**Descriptive statistics characterizes only the data you currently have. Inferential statistics makes conclusions beyond the existing data.**

*What is:*

**A model – Simplified representation of something. / Something that stands in for or looks like something else.**

*Consider different kinds of models. Here are some examples that are not necessarily mutually exclusive. What are their differences?*

**A conceptual model – An abstract model, an idea of summarization. Purely ideas. OR A non-mathematical model.**

**A statistical model – A model that simplifies relationships between numbers/data.**

**A causal model – A model that includes causality.**

**A computational model – Software representation of a concept or process.**

**A verbal model – An abstract model in verbal form.**

**A business model**

**A LEGO model**

*Given a model (usually a statistical or computational one), what is:*

**A variable – Something that varies (can take on different values).**

*What's the difference between*

**An observed variable** *and – Data we collected.*

**An unobserved variable – We can derive it from the data.**

*What's a*

**A hidden variable *– Synonym for unobserved***

*How about a*

**Input variable***and a*

**Outcome variable**

*What is then a*

**Parameter – quantities that can’t be observed in principle**

*And how is it different from a variable? Give an example from a statistical model, and from a causal model.*

*What does it mean to*

**Estimate model parameters**

*And what is it to*

**Fit a model to data**

*What does*

**Probability**

*mean? Given a set of different types observations, why do you*

**Divide by the total number of observations**

*to get the probability for each type? What must be the*

**Sum of the probabilities of each possible outcome**

*What does probability have to do with*

**Uncertainty**

*Richard McElreath quotes Bruno Finetti:*

**PROBABILITY DOES NOT EXIST**

*Then he says probability is a*

**Device for describing uncertainty from the   
perspective of an observer with limited knowledge**

*Explain that.*

Challenge

*What is the difference between*

**Epistemic uncertainty** *and*

**Ontological uncertainty**

*and which one does McElreath seem to believe in?*

*What is a*

**Distribution**

*What is the difference between a*

**Discrete distribution** *and a*

**Continuous distribution**

*In a distribution, what is the*

**Density**

*What defines a*

**Probability distribution**

*Consider both a*

**Discrete probability distribution** *and a*

**Continuous probability distribution**

*What is*

**Probability density**

*What is on*

**The axes of common visualizations of probability distributions**

*What is a*

**Probability density function**

*How does it*

**Relate to a probability distribution**

*What are the probability density functions'*

**Input variables** *and*

**Output variables**

*What would be, in general, a*

**Density function**

*that is not a probability density function?*

*Name a few*

**Different probability density functions**

*that you know, and their corresponding*

**Probability distributions**

*Also name their*

**Parameters**

*Here are some familiar examples:*

**Uniform distribution**

**Gaussian distribution / Normal distribution**

**Binomial distribution**

*And a few more:*

**T-distribution**

**Poisson distribution**

**Beta distribution**

**Gamma distribution**

Challenge

*What would be a*

**Probability distribution in multiple dimensions**

*Why is it also known as a*

**Multivariate probability distribution**

*What is*

**Bayes' Theorem**

*What are its four components, the:*

**Prior probability P(p)**

**Likelihood P(d|p)**

**Marginal Likelihood P(d)**

**Posterior probability P(p|d)**

*(Where P means 'probability of',*

*| means 'given some',*

*p means 'parameter value' and*

*d means 'data' or 'observations')*

# Exercise 2: into the readings

Here are some conceptual discussion tasks (more or less technical) to make reading the book in the future much easier. Some of them are also just interesting (at least I think so)! In Study Groups, discuss each question one at a time. We will go through some of them together :)

**Theory Level**

1)  
Richard McElreath distinguishes between the 'large world' and the 'small world' of the model. What is the difference? Find some examples (maybe in relation to science) where knowing the difference is really important.

2)

What does it mean to say that a model is 'true'?

3)

What is meant by the ancient saying:

'All models are wrong, but some are useful'

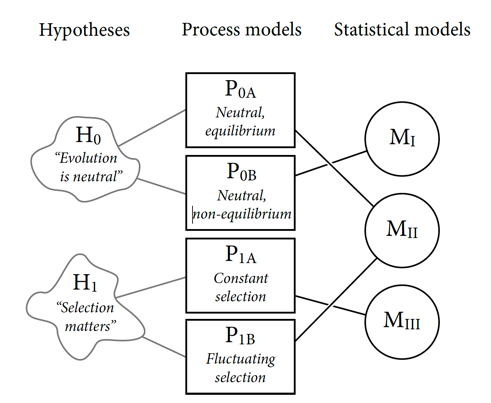
When is a model 'useful'?

Challenge

Is this point only valid for statistical and computational models, or for models in general?

4)

Explain the point of figure 1.2 below  
(the general point about hypotheses and models, not the stuff about evolution)



5)

McElreath points out three ways our inference can go wrong. We can have bad:

software

model

data

Give examples of each. 'Model' here also includes the way of estimating parameters.

**Methodology Level**

1)

Explain the marble bag example

(use figures and tables in the book)

2)

Explain the globe tossing example

(use figures and tables in the book)

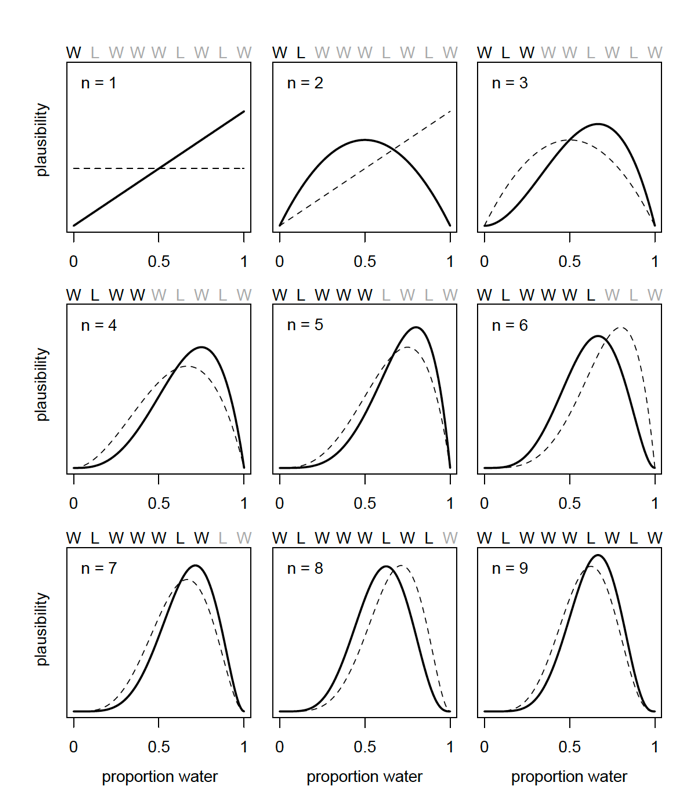
Explain both the (hypothetical) experiment, and the model used to make inferences about it

3)

Come up with a different phenomenon to make a model of,  
but where the model would have the same structure as in the globe tossing example,  
and the only difference would be what the parameters and variables represent.

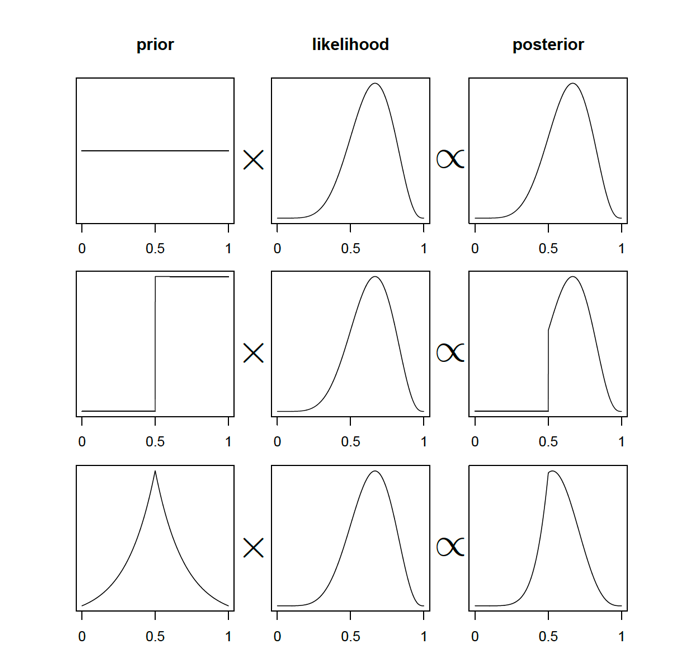
4)

Explain figure 2.5



5)

Explain figure 2.6.   
What does it mean to multiply two probability distributions together?



6)

What is the difference between having a single value and a probability distribution as a parameter value?

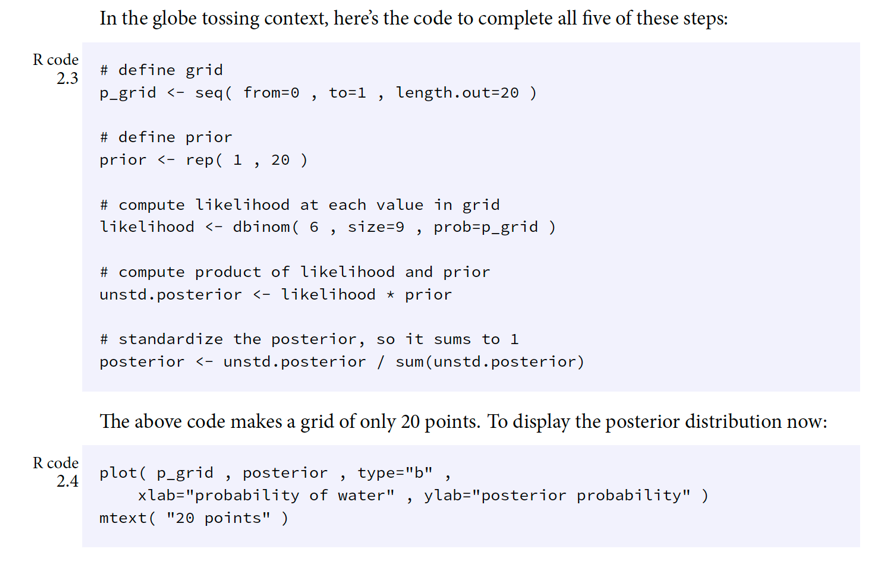
Why is it nice to have a probability distribution?

7)

Explain the grid approximation method for estimating a model's parameter values.

You can explain it based on Code 2.5.

How does it give a probability distribution, and not just a single value?



8)

Solve the exercises below on notation in probability theory.

