Class 2: Conceptual Exercise

Hey CogSci'ers :)

Here are some conceptual exercises to make sure we understand the readings (focused on chapters 2 and 3).  
This is a *group member presentation exercise*.  
That means:  
1. each member of the study group chooses a question

2. each member of the study group gets 5 minutes to prepare an answer to their question

3. each member takes turns presenting their answer to the others

4. if something was difficult, we talk about it in plenum

5. repeat

NB: some questions take longer than others. If you're done preparing before the 5 minutes are up, choose another question :)

Questions on the next pages.

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

# Exercise 2: into the readings

**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)

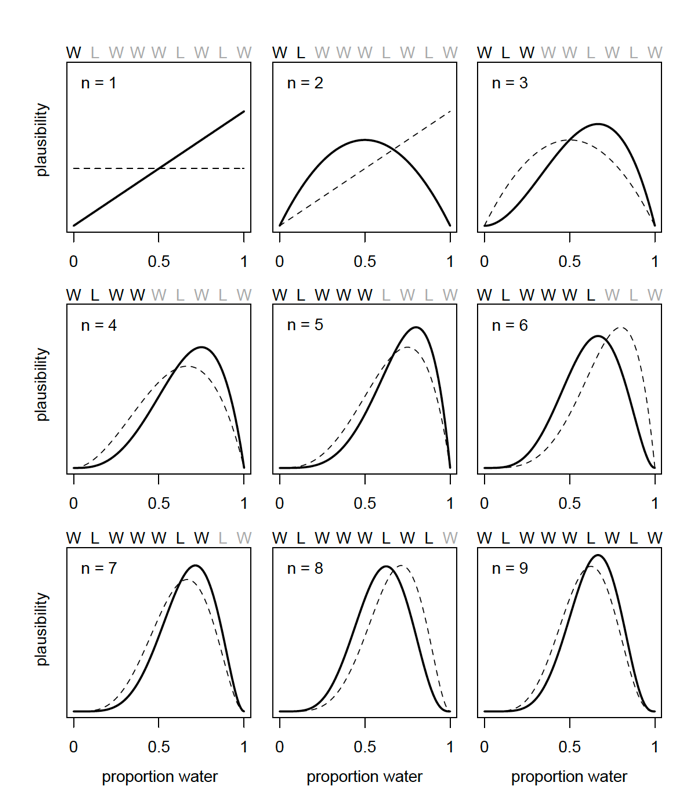
Explain both the (hypothetical) experiment, and the model used to make inferences about it

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.

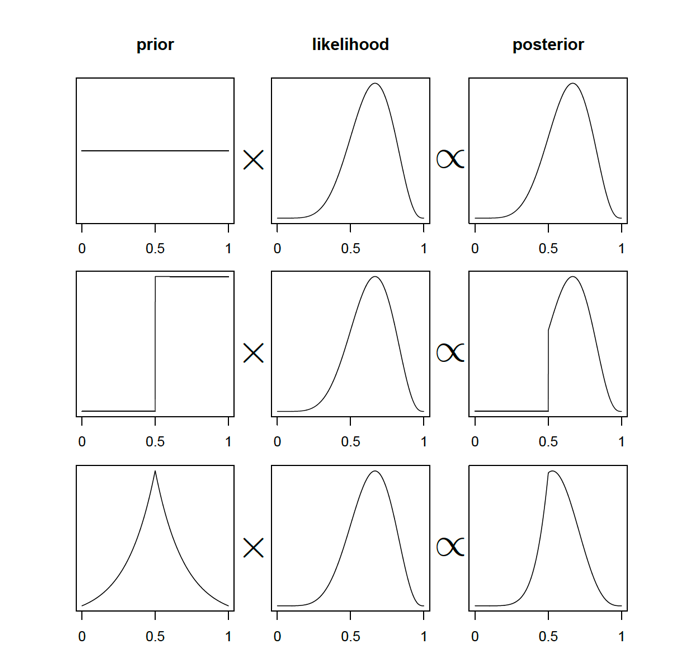
4)

Explain figure 2.5



5)

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

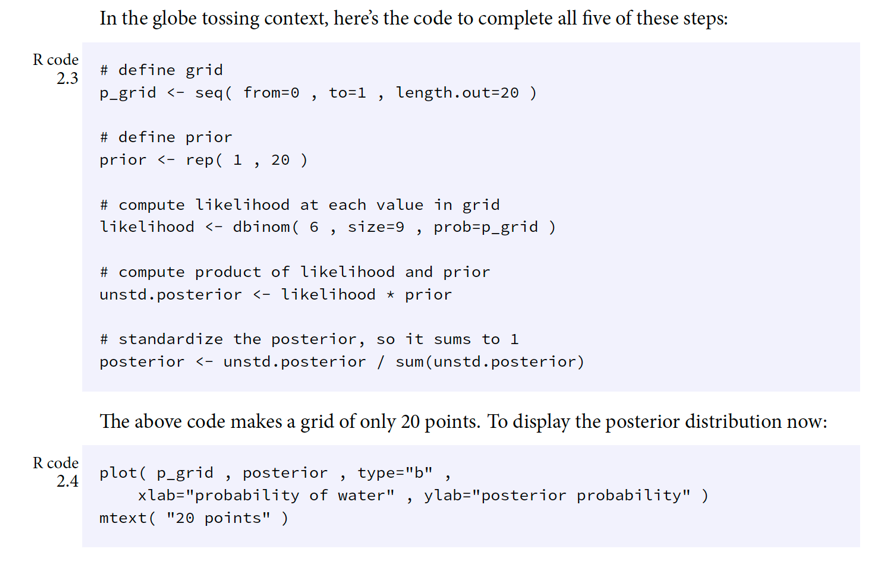


6)

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 point estimate?



7)

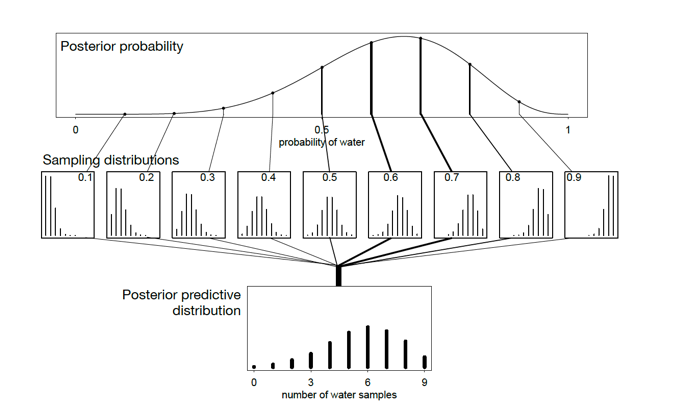
McElreath points out that using different kinds of point estimates implicitly corresponds to using different kinds of loss functions.

Explain what this means.

Give examples of different situations (from science or otherwhere) where there would be different loss functions.

8)

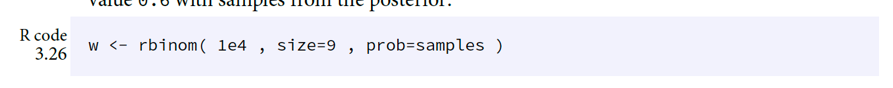
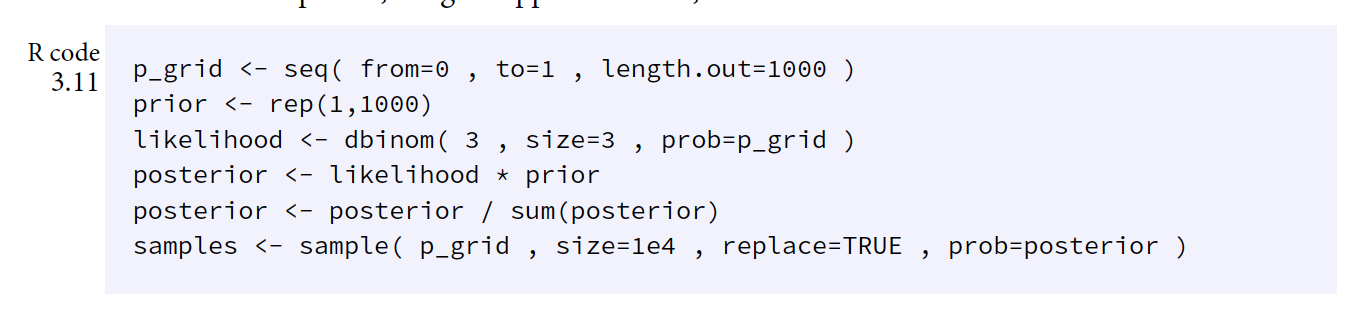
Explain Figure 3.6



Explain how the uncertainty about the parameter estimate is propagated into the posterior predictive distribution, and why it is more 'honest' than using point estimates for parameters.

9)

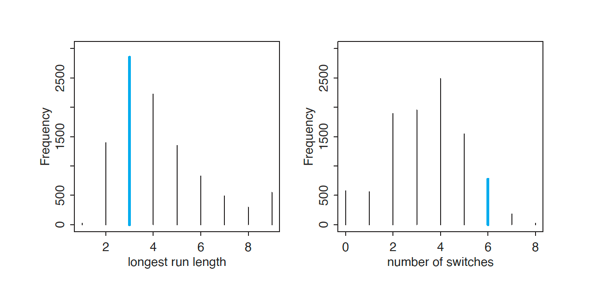
Explain what happens in R code 3.11 and 3.26.



Why does using the sampled parameter values as parameters in the likelihood function (here a binomial) amount to averaging the prediction distribution for all parameter values, weighted by those parameter values' probability?

10)

Explain figure 3.7



Why is it useful to look at the data in different ways like this?