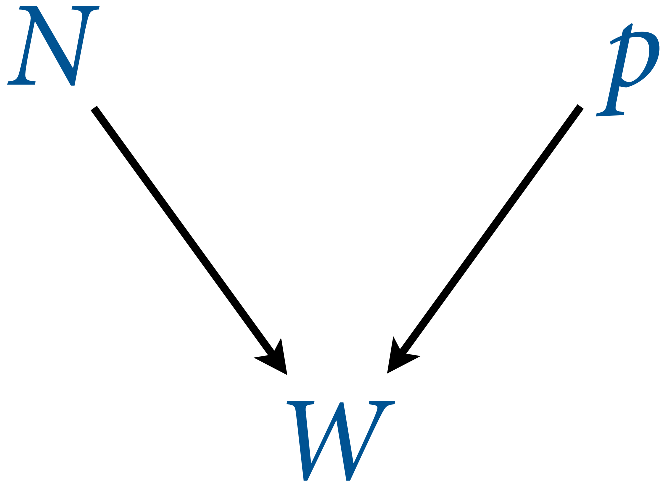
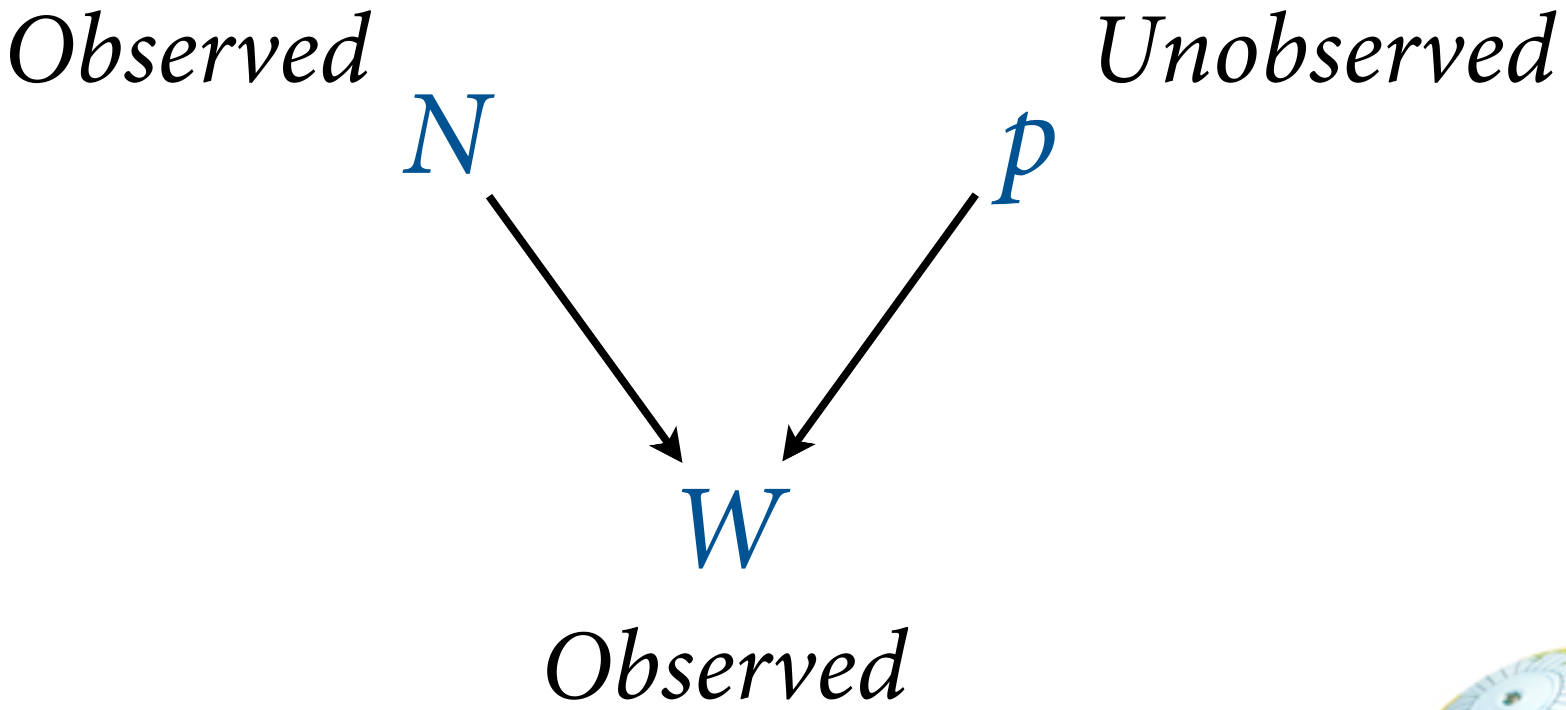

N
p
W

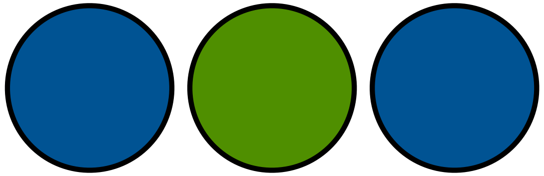

Observed
N
p
Unobserved
W
Observed

Definition of W

- Relative number of ways to see W, given N and p?
- Goal: Mathematical function to answer this question.
- The answer is a *probability distribution*.

Definition of W

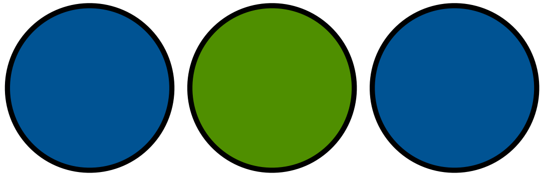
- Relative number of ways to see W, given N and p?



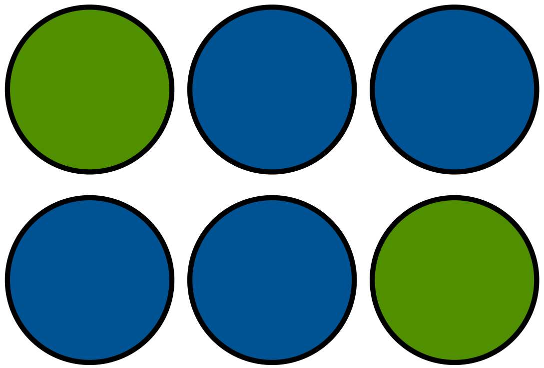
$$p \times (1-p) \times p \quad = p^2(1-p)^1$$

Definition of W

- Relative number of ways to see W, given N and p?

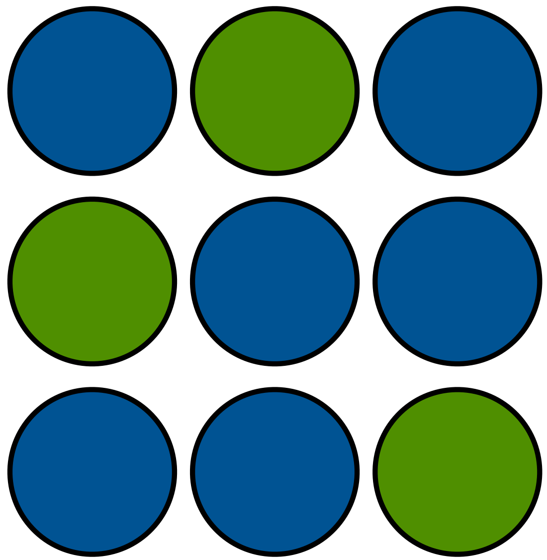


$$p \times (1-p) \times p \quad = p^2(1-p)^1$$



Definition of W

- Relative number of ways to see W, given N and p?



$$\Pr(2|3,p) = 3p^2(1-p)^1$$

W distribution (Likelihood)

number tosses

$$\Pr(W|N,p) = \frac{N!}{W!(N-W)!}p^W(1-p)^{N-W}$$

count W *probability* W

The count of W's is distributed binomially, with probability *p* of a W on each toss and *N* tosses total.

W distribution (Likelihood)

$$\Pr(W|N,p) = \frac{N!}{W!(N-W)!} p^W (1-p)^{N-W}$$

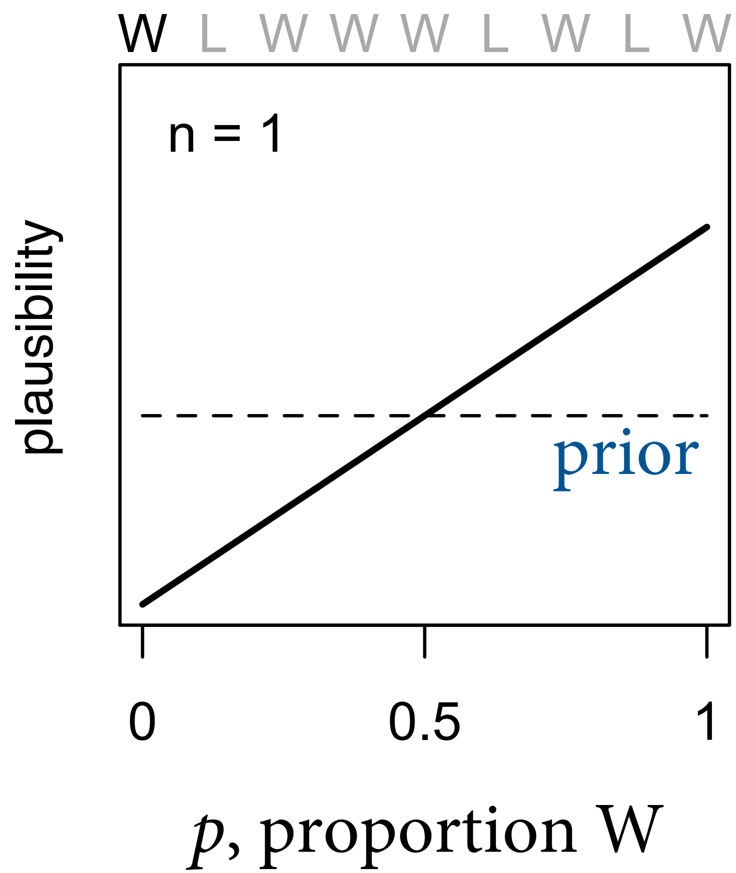
```r
dbinom( 6 , size=9 , prob=0.5 )
```

R code 2.2

```
[1] 0.1640625
```

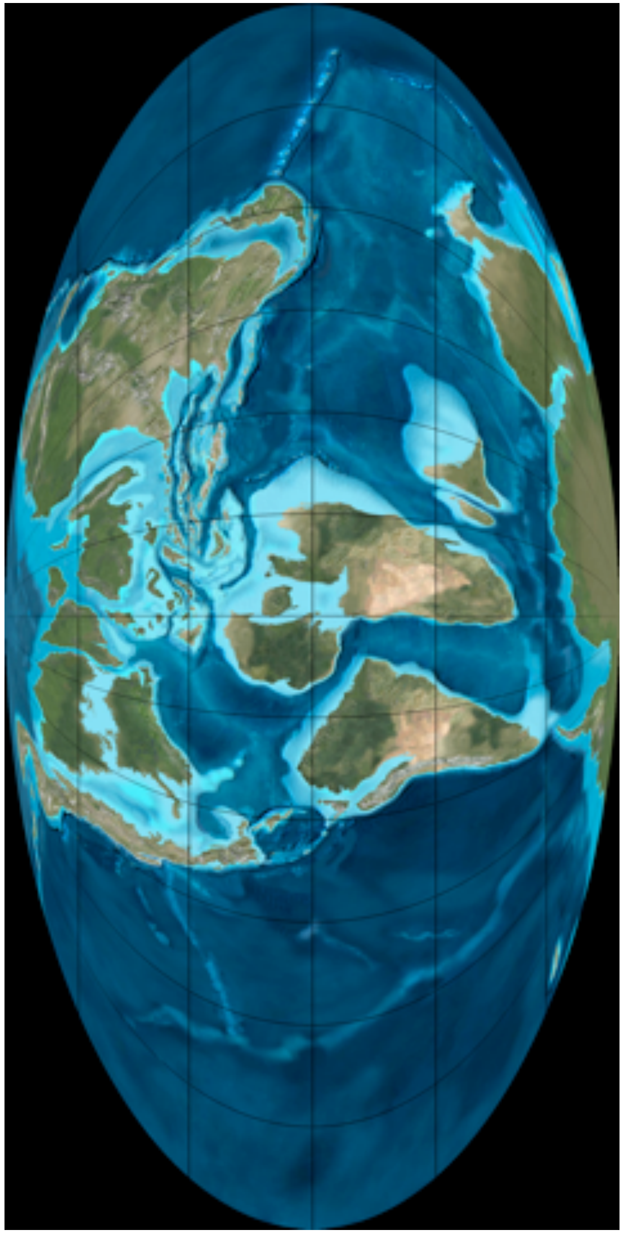
Prior probability *p*

- What the golem believes before the data arrive
- In this case, equal prior probability 0–1
- Pr(*W*) & Pr(*p*) define *prior predictive distribution*
- More on this later – it helps us build priors that make sense



Prior literature

- Huge literature on choice of prior
- Flat prior conventional & bad
  - Always know something (before data) that can improve inference
  - Are zero and one plausible values for p? Is $p < 0.5$ as plausible as $p > 0.5$?
  - There is no "true" prior
  - Just need to do better than flat
- All above equally true of likelihood



Late Cretaceous (90Mya)

The Joint Model

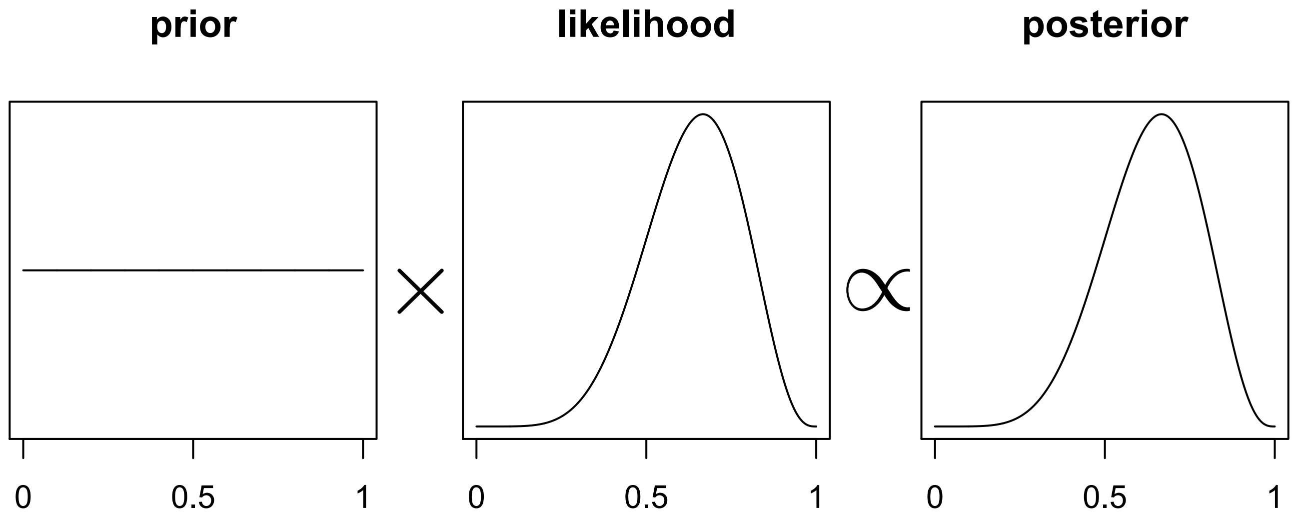
$$
\begin{aligned}
W &\sim \text{Binomial}(N, p) \\
p &\sim \text{Uniform}(0, 1)
\end{aligned}
$$

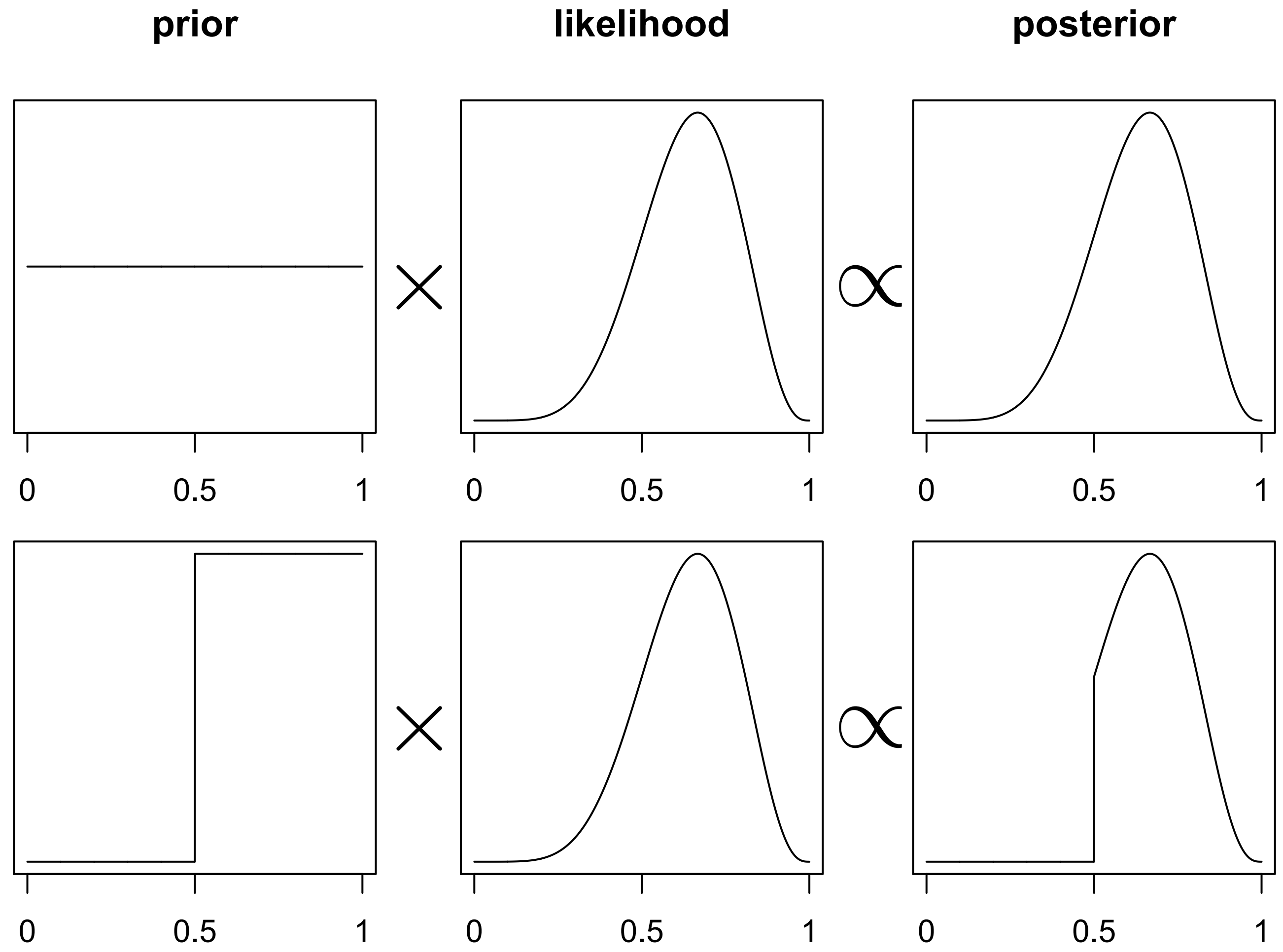
Posterior probability

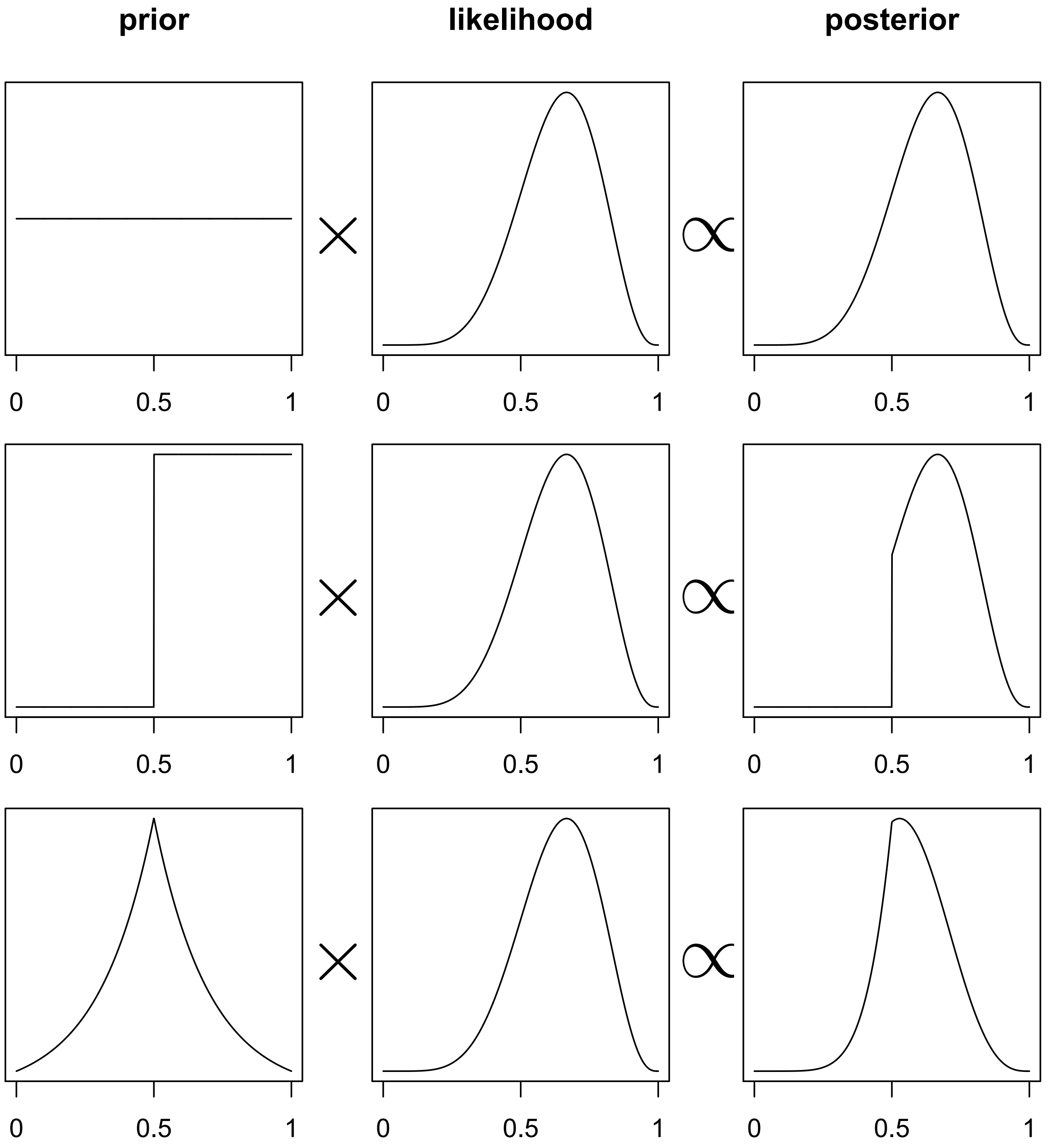
- Bayesian "estimate" is always *posterior distribution over parameters*, Pr(parameters|data)
- Here: $\Pr(p|W,N)$
- Compute using *Bayes' theorem*:

$$\Pr(p|W,N) = \frac{\Pr(W|N,p)\Pr(p)}{\sum \Pr(W|N,p)\Pr(p) \text{ for all } p}$$

$$\text{Posterior} = \frac{(\text{Prob observed variables}) \times (\text{Prior})}{\text{Normalizing constant}}$$


prior
likelihood
posterior
0
0.5
1
0
0.5
1
0
0.5
1


prior
likelihood
posterior
0
0.5
1
0
0.5
1
0
0.5
1
0
0.5
1
0
0.5
1
0
0.5
1



prior
likelihood
posterior
×
∝
0
0.5
1