Weak priors and model checking

Stat 340: Bayesian Statistics

1. Weakly informative prior distributions

2. Model checking

(Problem topics 1 & 5)

Example

Setting Assessing proportion of U.S. transportation industry workers who use drugs on the job.

Data RS of size $n = 10$ taken; $Y = 2$ positive tests

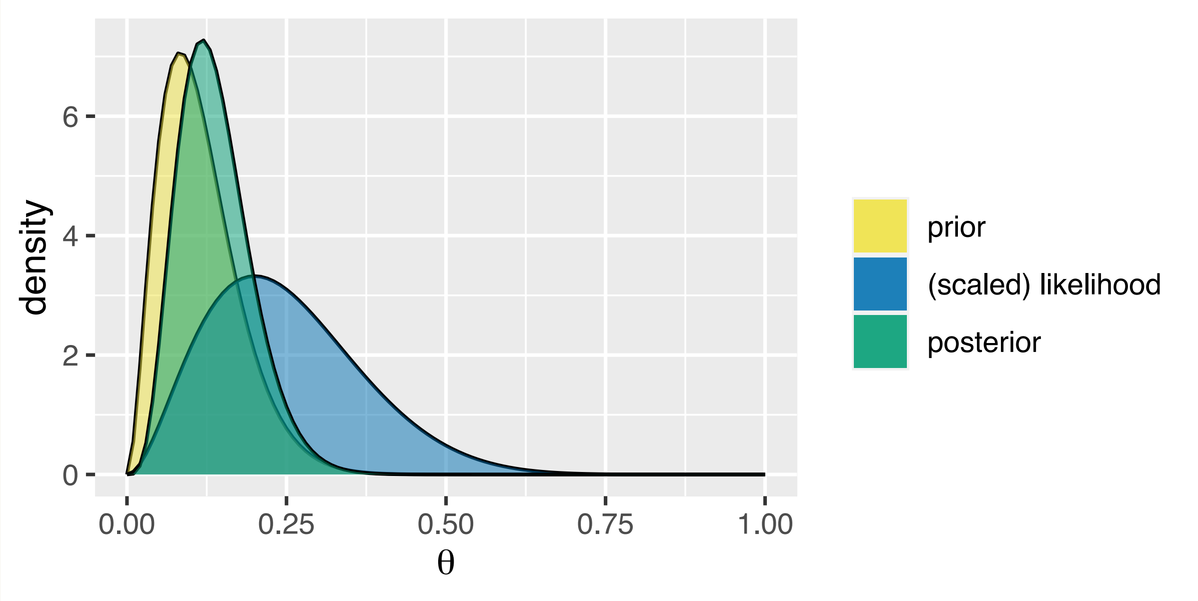
Likelihood $Y|\theta \sim \mathrm{Binomial}(n = 10, \theta)$

Prior Based on prior testing, $\mathrm{Beta}(a = 3, b = 23)$

Posterior $\theta|Y \sim \mathrm{Beta}(5, 31)$

Example

- Posterior $\propto$ Prior $\times$ Likelihood
- Posterior can't have density anywhere prior density is 0
- If prior "dominates" the likelihood, then the data have very little to contribute to the analysis!



1. Vauge/weakly informative priors

> *[A] prior which is dominated by the likelihood is one which does not change very much over the region in which the likelihood is appreciable and does not assume large values outside that range."*
>
> *For such a prior distribution we can approximate the result of Bayes' formula (theorem) by substituting a constant for the prior distribution.*
>
> *-Box and Tiao (1973)*

Box, G. E., & Tiao, G. C. (1973). *Bayesian inference in statistical analysis.*

A weak prior

Setting Assessing proportion of U.S. transportation industry workers who use drugs on the job.

Data RS of size $n = 10$ taken; $Y = 2$ positive tests

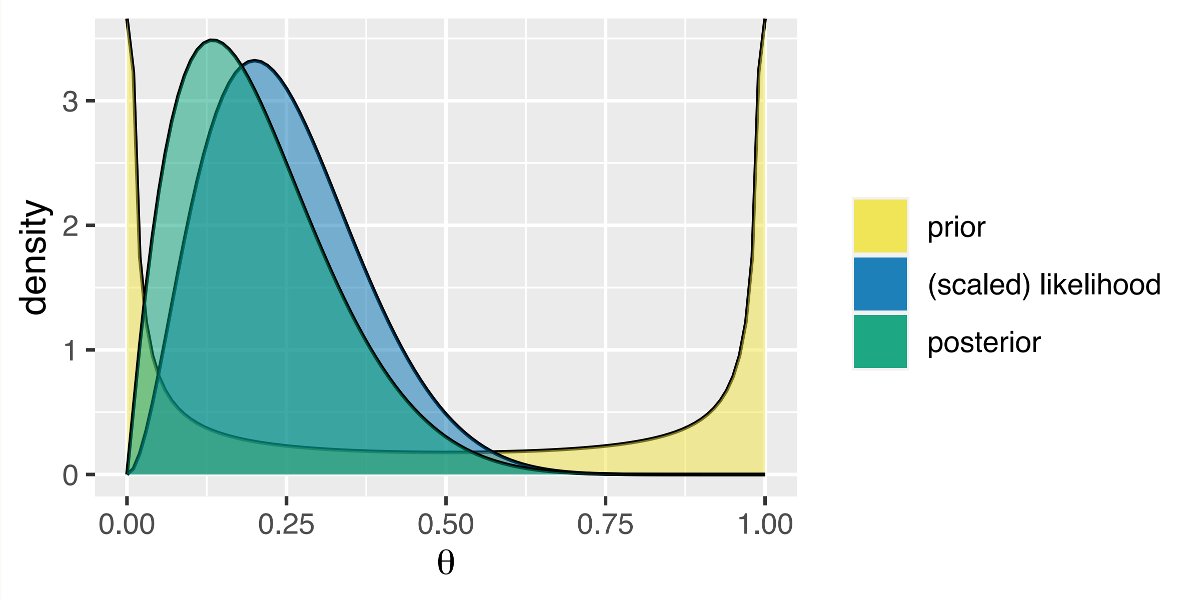
Likelihood $Y|\theta \sim \text{Binomial}(n = 10, \theta)$

Prior No prior info so analyst sets $\text{Beta}(a = 1/10, b = 1/10)$

Posterior $\theta|Y \sim \text{Beta}(2.1, 10.1)$

A weak prior

- Likelihood dominates the prior
- The data almost entirely determine the posterior

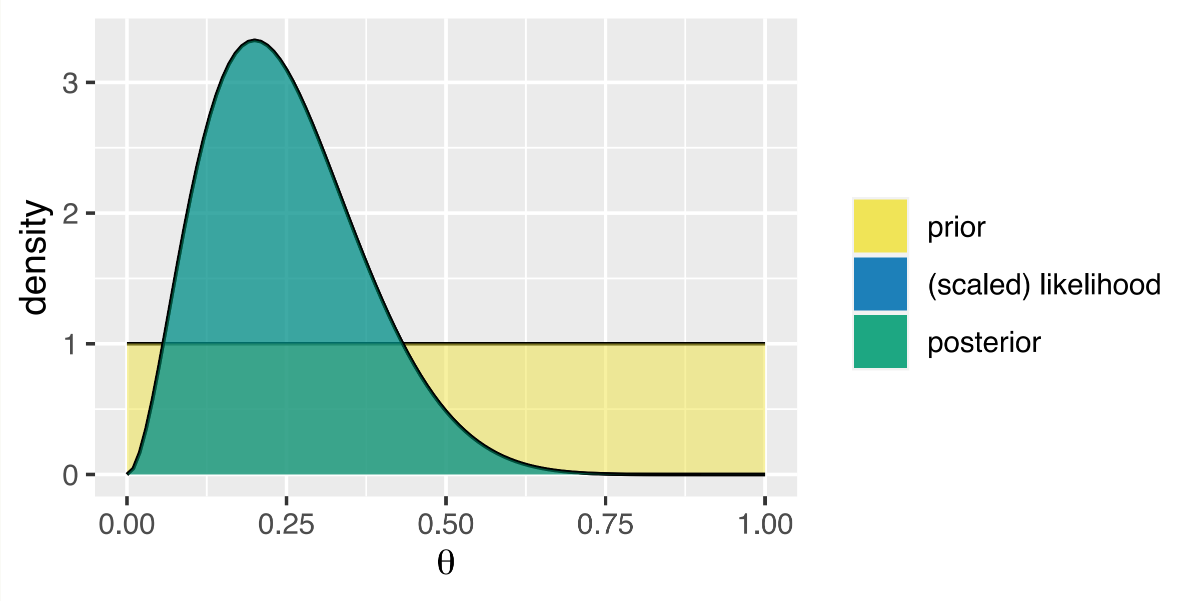


A flat prior

Setting	Assessing proportion of U.S. transportation industry workers who use drugs on the job.
Data	RS of size $n = 10$ taken; $Y = 2$ positive tests
Likelihood	$Y\|\theta \sim \mathrm{Binomial}(n = 10, \theta)$
Prior	No prior info so analyst sets $\mathrm{Beta}(a = 1, b = 1)$
Posterior	$\theta\|Y \sim \mathrm{Beta}(3, 11)$

A flat prior

- Likelihood dominates the prior
- The data almost entirely determine the posterior



2. Reference priors

Many definitions:

- Conventional or default choice (Kass and Wasserman, 1996)
- Representing ignorance in some formal sense (Bernardo, 1979)

Check out the catalog of default priors

Kass, R. E., & Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435), 1343-1370.

Bernardo, J. (1979). Reference Posterior Distributions for Bayesian Inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2), 113-14.

Binomial reference prior

Data RS of size $n = 10$ taken; $Y = 2$ positive tests

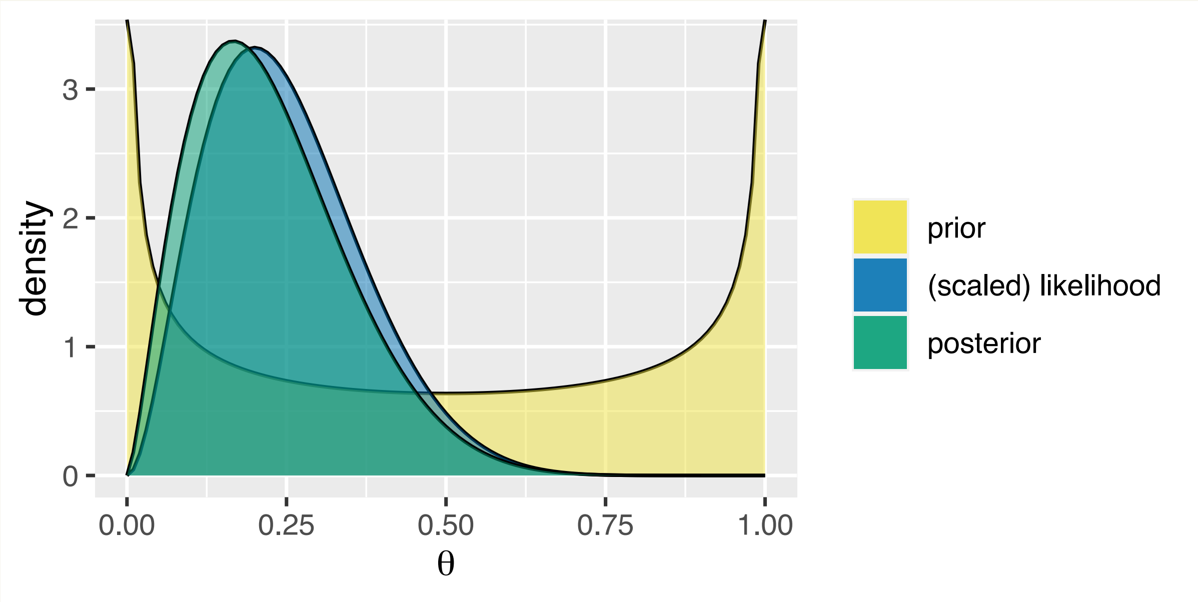
Likelihood $Y|\theta \sim \text{Binomial}(n = 10, \theta)$

Prior Analyst want to be "objective", so uses reference prior
$\theta \sim \text{Beta}(1/2, 1/2)$

Posterior $\theta|Y \sim \text{Beta}(2.5, 10.5)$

Binomial reference prior

- Likelihood dominates the prior
- The data almost entirely determine the posterior



Model checking

Posterior predictive checking

Key idea

If a model fits, then if should be able to generate data that look similar to the observed data

- Draw new samples of the same size from the model
- Are there systematic differences?

Example

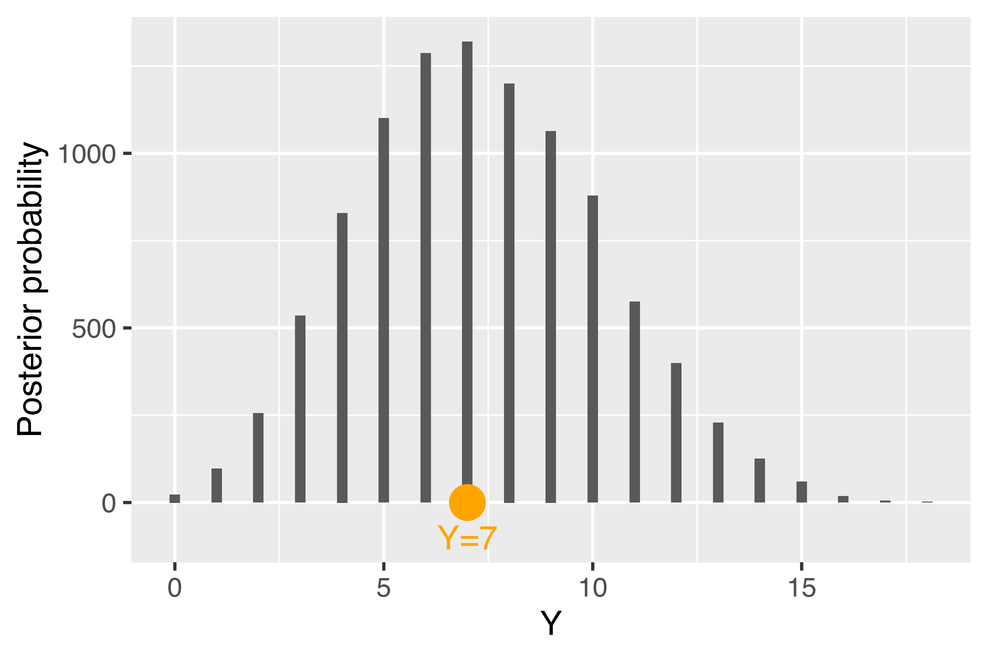
OBSERVED DATA

$Y = 7$

MODEL

$$
\begin{aligned}
Y|p &\sim \text{Binomial}(20, p) \\
p &\sim \text{Beta}(1, 1) \\
\implies p|Y &\sim \text{Beta}(8, 14)
\end{aligned}
$$

Are there systematic differences?



Example (continued)

OBSERVED DATA

Looking at the data, you see a temporal pattern

1 1 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0

MODEL

$$\begin{aligned} Y|p &\sim \text{Binomial}(20, p) \\ p &\sim \text{Beta}(1, 1) \\ \implies p|Y &\sim \text{Beta}(8, 14) \end{aligned}$$

Are there systematic differences?

0 1 0 1 1 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0

0 0 0 0 1 0 0 1 1 1 0 1 1 1 0 0 1 1 0 0

1 0 0 0 0 1 0 0 1 1 1 1 1 0 1 1 1 0 1 0

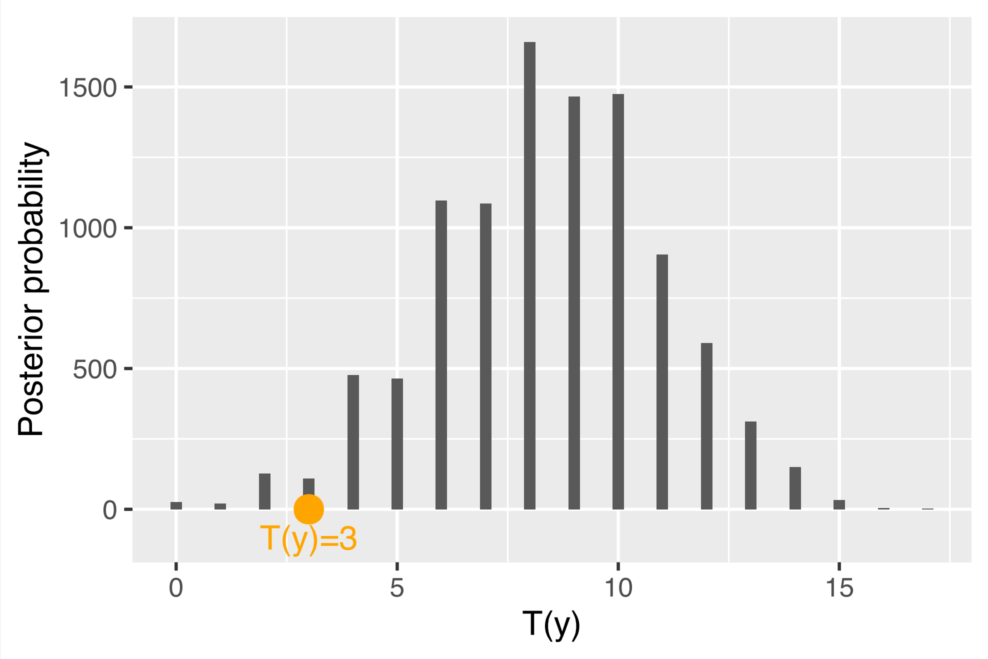
0 0 0 1 0 1 0 1 0 1 1 0 1 0 1 1 0 0 0 1

0 0 0 1 1 1 1 0 1 0 1 1 0 1 0 0 0 1 1 0

Example (continued)

OBSERVED DATA: 1 1 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0

- Consider $T(y) =$ number of switches between 0 and 1 in the sequence.
- Calculate for $T(y, p)$ for replicate data sets, $\mathbf{y}_{\text{rep}}$
- Calculate for $T(y, p)$ for observed data set
- Approximate $P(T(\mathbf{y}_{\text{rep}}) \geq T(\mathbf{y}))$



$$P(T(\mathbf{y}_{\text{rep}}) \geq T(\mathbf{y})) \approx 0.9828$$