Bayesian Statistics for Genetics

Lecture 3: Binomial Sampling, part 2

June, 2025

Outline

Continuing our look at Bayesian inference for binomial data:

- Prior specification
- Testing
- Logistic regression
- Predictive distributions

Prior Specification

- Particularly for small datasets, it is a good idea to examine the sensitivity of inference to the prior choice, particularly for those parameters for which there is little information in the data (e.g. variances in random effects models)
- An obvious way to find sensitivities is to compare the posterior under various priors, but experience often helps
- For subjective priors, that reflect the data analyst's belief about the unknowns, sensitivity isn't a bad thing **if** the prior can be justified.

Prior Specification

- Sometimes we can specify priors that, in some sense, allows the data to dominate the posterior. (No great name, but *weakly-informative* suggested)
- Priors can also be found (in some settings) that produce point estimates and intervals that give a good non-Bayesian properties, i.e., have good frequency properties – bias, coverage, etc. ("frequentist pursuit")
- Such priors provide a baseline to compare analyses with more substantive priors.
- Other names for such priors are *objective*, *reference* and *non-subjective*

Prior Specification

For the beta prior/binomial likelihood, recall we have to specify the Beta's parameters a and b, which are difficult to interpret.

- The posterior mean is a weighted average:

$$\mathbb{E}[\theta|y] = \frac{y+a}{N+a+b} = \frac{y}{N}\underbrace{\frac{N}{N+a+b}}_{W} + \frac{a}{a+b}\underbrace{\frac{a+b}{N+a+b}}_{1-W}.$$

- Viewing $N+a+b$ as 'total' sample size suggests a way to choose a and b
- ...where we specify the prior mean $m_{\text{prior}} = a/(a+b)$ and the "prior sample size" $N_{\text{prior}} = a+b$, then solve for a and b via

$$\begin{aligned} a &= N_{\text{prior}} \times m_{\text{prior}} \\ b &= N_{\text{prior}} \times (1 - m_{\text{prior}}). \end{aligned}$$

- Intuitively, a acts like a prior number of successes and b like a prior number of failures

A Binomial Example

- Suppose we set $N_{\text{prior}} = 5$ and $m_{\text{prior}} = \frac{2}{5}$
- It is *as if* we saw 2 successes out of 5
- Suppose we obtain data with $y = 7$, $N = 10$ and so $\frac{y}{N} = \frac{7}{10}$
- Hence $\mathsf{W} = 10/(10+5)$ and

$$\begin{aligned}\mathbb{E}[\theta|y] &= \frac{7}{10} \times \frac{10}{10+5} + \frac{2}{5} \times \frac{5}{10+5} \\ &= \frac{9}{15} = \frac{3}{5}.\end{aligned}$$

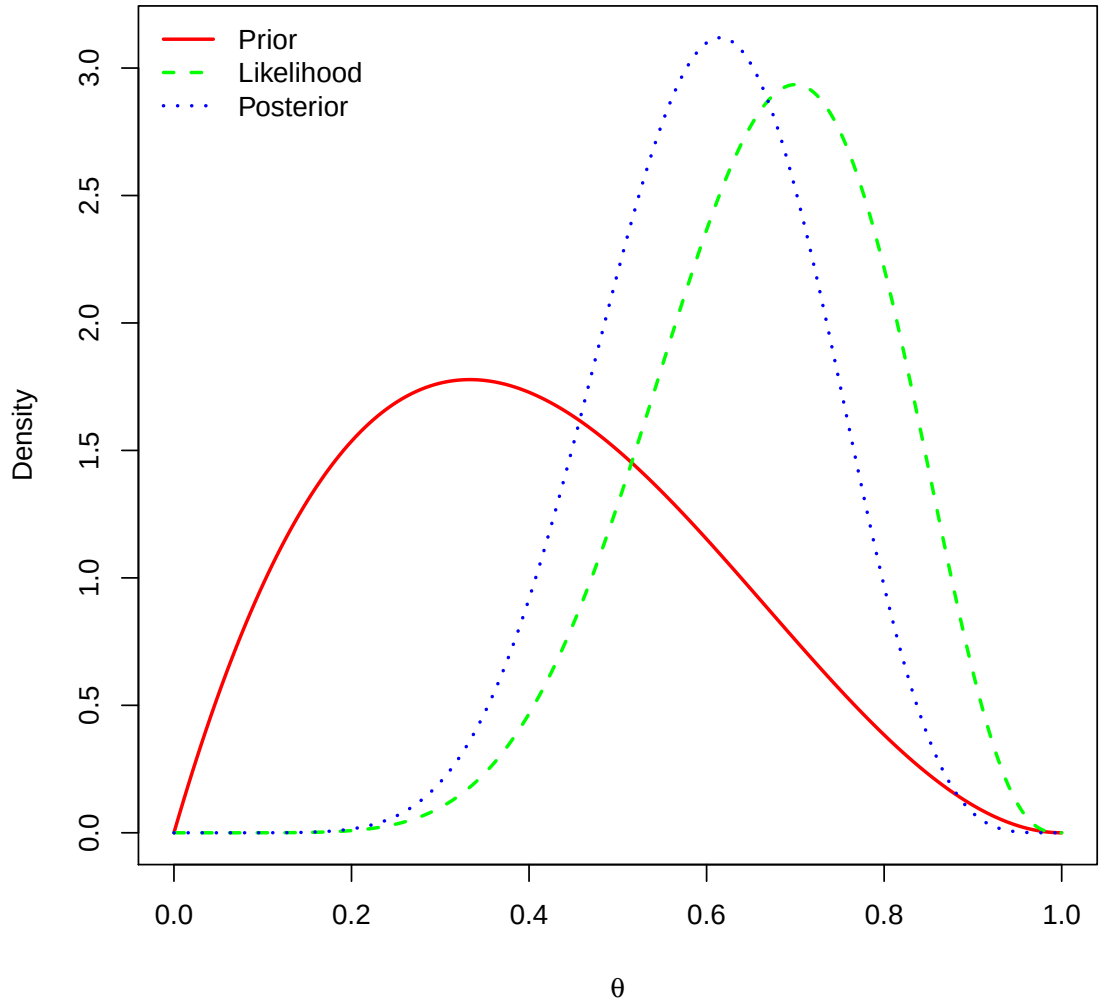
- Solving:

$$\begin{aligned}a &= N_{\text{prior}} \times m_{\text{prior}} = 5 \times \frac{2}{5} = 2 \\ b &= N_{\text{prior}} \times (1 - m_{\text{prior}}) = 5 \times \frac{3}{5} = 3\end{aligned}$$

- This gives a $\mathsf{Beta}(y + a, N - y + b) = \mathsf{Beta}(7+2, 3+3)$ posterior

A Binomial Example

Updating of a Beta(2,3) prior by likelihood proportional to a Beta(7,3) density, giving a Beta(7+2,3+3) posterior.



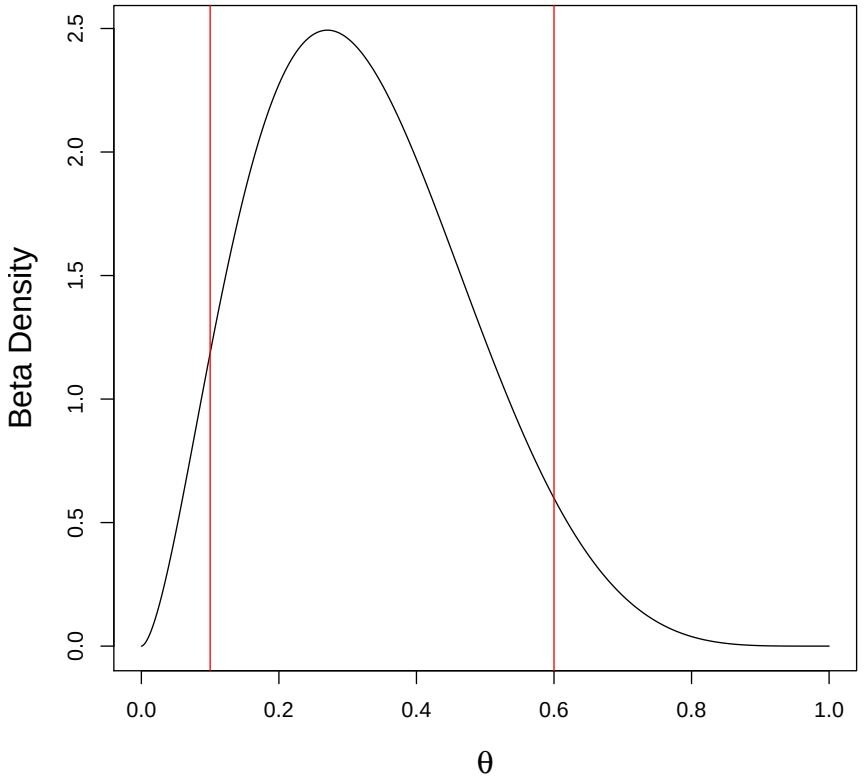
Choosing a Prior: Approach Two

A convenient & alternative way to choose a, b is specifying two prior quantiles

- For example, specify that $\Pr(\theta < 0.1) = 0.05$ and $\Pr(\theta > 0.6) = 0.05$, and find the a and b values that provide this.
- We usually find the solutions numerically; for example, solving

$$(p_1 - \Pr(\theta < q_1|a, b))^2 + (p_2 - \Pr(\theta < q_2|a, b))^2 = 0$$

  for a, b. (No prizes for elegance!)
- ...giving the Beta(2.73,5.67) prior shown



Difference in Binomial Proportions

These methods extend naturally to a pair of samples:

- Suppose we have two binomial observations:

$$\begin{aligned} Y_1|\theta_1 &\sim \text{Binomial}(N_1, \theta_1), \quad \text{for sample 1} \\ Y_2|\theta_2 &\sim \text{Binomial}(N_2, \theta_2), \quad \text{for sample 2} \end{aligned}$$

- We want inference on $\theta_1 - \theta_2$ (the *risk difference* or *absolute risk difference*)
- With independent beta priors on θ_1 and θ_2, sampling from the posterior for $p(\theta_1 - \theta_2|y_1, y_2)$ is straightforward; just sample from the beta posterior for θ_1 and θ_2 independently.

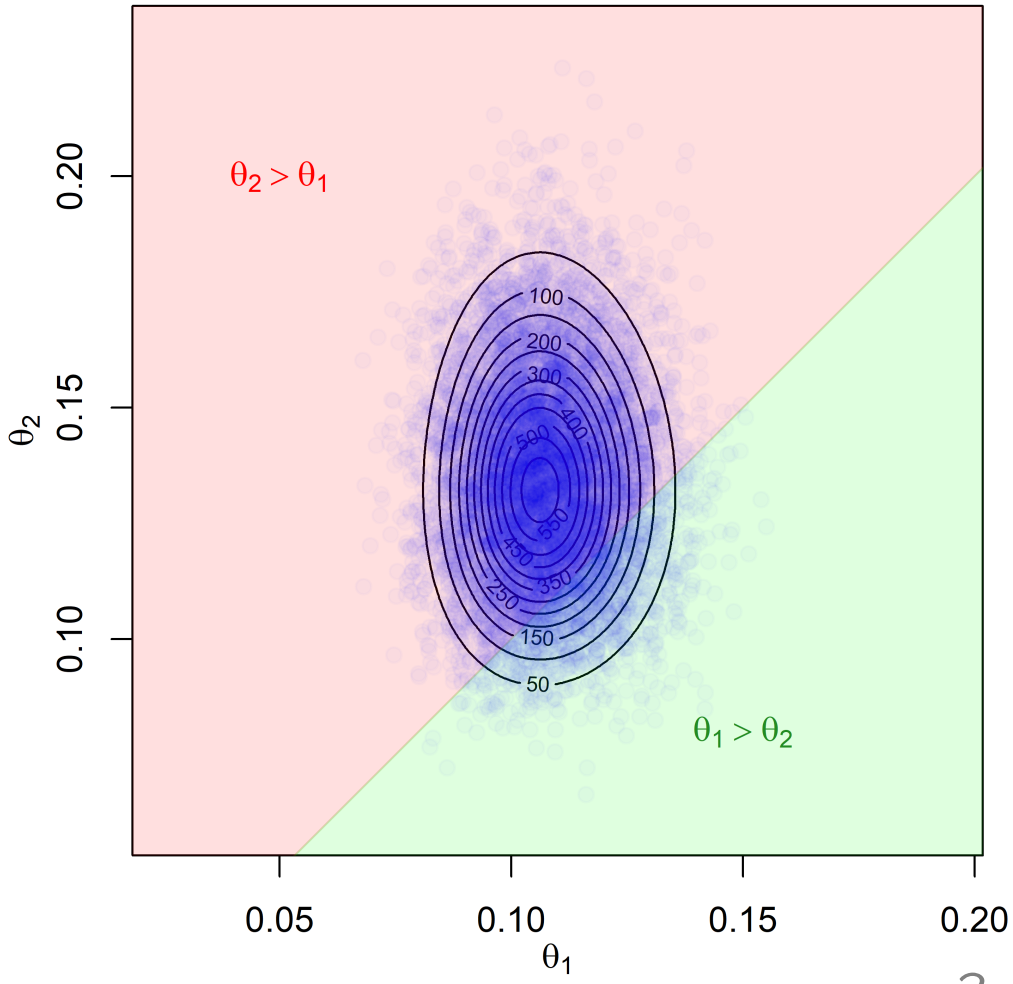
Difference in Binomial Proportions

- Savage *et al* (2008) study allele frequencies within a gene linked with skin cancer
- We want to examine differences in allele frequencies between populations.
- Here, we examine one variant in Northern European (NE) and United States (US) populations
- Let θ_1 and θ_2 be the allele frequencies in the NE and US population from which the samples were drawn, respectively.
- The allele frequencies were 10.69% and 13.21% with sample sizes of 650 and 265, in the NE and US samples, respectively. (Counts of 69/650 and 35/265)
- We assume independent Beta(1,1) priors on each of θ_1 and θ_2.

Difference in Binomial Proportions

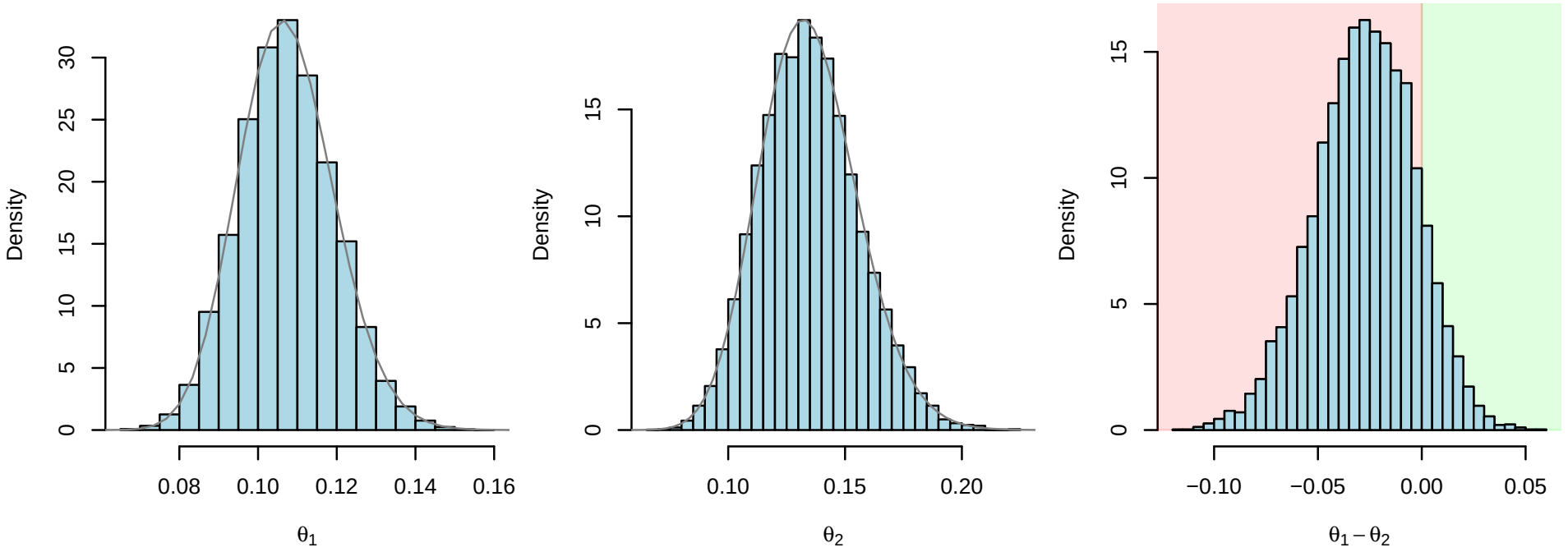
The *joint* posterior density for θ_1 and θ_2, as a contour plot with the samples superimposed:

The posterior probability that $\theta_1 > \theta_2$ is 0.12 (computed as the proportion of the samples in the green zone). The data don't strongly suggest either group has a higher/lower allele frequency.



Difference in Binomial Proportions

The exact posterior for the difference $\theta_1 - \theta_2$ is messy to work with – it's a convolution of two beta distributions. But using the samples we already took, we can get a very good approximation *without further work*;



Histograms reprsenting $p(\theta_1|y_1)$, $p(\theta_2|y_2)$ and $p(\theta_1 - \theta_2|y_1, y_2)$.

Bayes Factors for Hypothesis Testing

The *Bayes factor* (BF) provides a summary of the evidence for a particular hypothesis/model compared to another

- For null hypothesis H_0 and alternative H_1, the simplest definition is

$$\text{BF} = \frac{\Pr(y|H_0)}{\Pr(y|H_1)},$$

  i.e. the probability of the data under H_0 divided by the probability of the data under H_1. Values of BF > 1 favor H_0 while values of BF < 1 favor H_1.
- The BF is similar to the *likelihood ratio*,

$$\text{LR} = \frac{\Pr(y|H_0)}{\Pr(y|\widehat{\theta})}$$

  where $\widehat{\theta}$ is the MLE under H_1; H_0 usually specifies $\theta = 0$. BF and LR are identical if there are (unusually!) no unknown parameters in H_0 and H_1, but otherwise the BF averages over them and LR does not.

Calibration of Bayes Factors

Kass & Raftery (1995) suggest these interpretations of BFs:

1/Bayes Factor	Evidence Against H_0
1 to 3.2	Not worth more than a bare mention
3.2 to 20	Positive
20 to 150	Strong
$>$150	Very strong

These provide a guideline ("*T-shirt sizes*") but impact of right/wrong conclusions in context should also be considered.

Bayes Factor: another derivation

The *odds* of an event is the probability the event happens, divided by the probability it does not.

$$\text{Odds}[\,A\,] = \frac{\mathbb{P}[\,A\,]}{1-\mathbb{P}[\,A\,]}, \qquad \mathbb{P}[\,A\,] = \frac{\text{Odds}[\,A\,]}{1+\text{Odds}[\,A\,]}$$

If we view the 'true model' (i.e. H_0 or H_1) as an unknown, then

$$\frac{\mathbb{P}[\,H_0|y\,]}{\mathbb{P}[\,H_1|\boldsymbol{y}\,]} = \frac{\mathbb{P}[\,y|H_0\,]}{\mathbb{P}[\,y|H_1\,]} \times \frac{\mathbb{P}[\,H_0\,]}{\mathbb{P}[\,H_1\,]}$$

$$\text{i.e. Posterior Odds of } H_0 \;=\; \text{Bayes Factor} \times \text{Prior Odds}$$

- The Bayes Factor tells us how much the data *update* the prior odds for/against H_0
- Neatly, the BF does **not** depend on the prior odds of H_0, H_1 – though the prior on parameters in the null/alternative models will affect BF

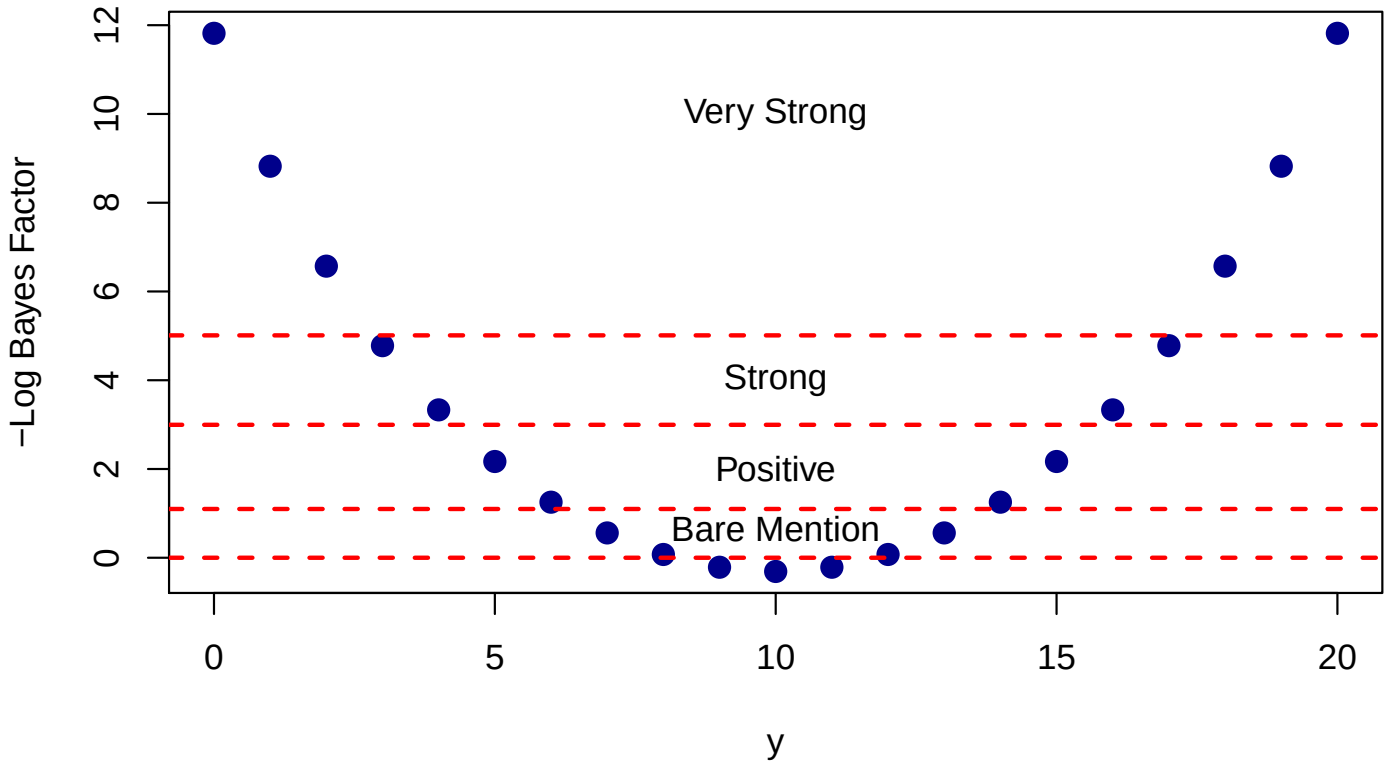
Example: Bayes Factors for Binomial Data

For the usual Binomial model with a Beta(a, b) prior, if we are interested in $H_0 : \theta = 0.5$ versus $H_1 : \theta \neq 0.5$.

The numerator and denominator of the Bayes factor are:

$$\begin{aligned}
\Pr(y|H_0) &= \binom{N}{y} 0.5^y 0.5^{N-y} \\
\Pr(y|H_1) &= \int_0^1 \binom{N}{y} \theta^y (1-\theta)^{N-y} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} d\theta \\
&= \binom{N}{y} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(y+a)\Gamma(N-y+b)}{\Gamma(N+a+b)}
\end{aligned}$$

Example: Bayes Factors for Binomial Data



Negative Log Bayes factors comparing Binomial $\theta = 0.5$ vs $\theta \neq 0.5$ for y total successes from $n = 20$ trials. The prior on θ sets $a = b = 1$, i.e. is uniform. High values indicate evidence against the null.

Analysis of ASE Data

Summarizing the approaches we've considered:

- Posterior probabilities of one region: calculate

  $$\Pr(\theta < 0.5|y)$$

  and check if it exceeds some threshold indicating further study is worthwhile, e.g.

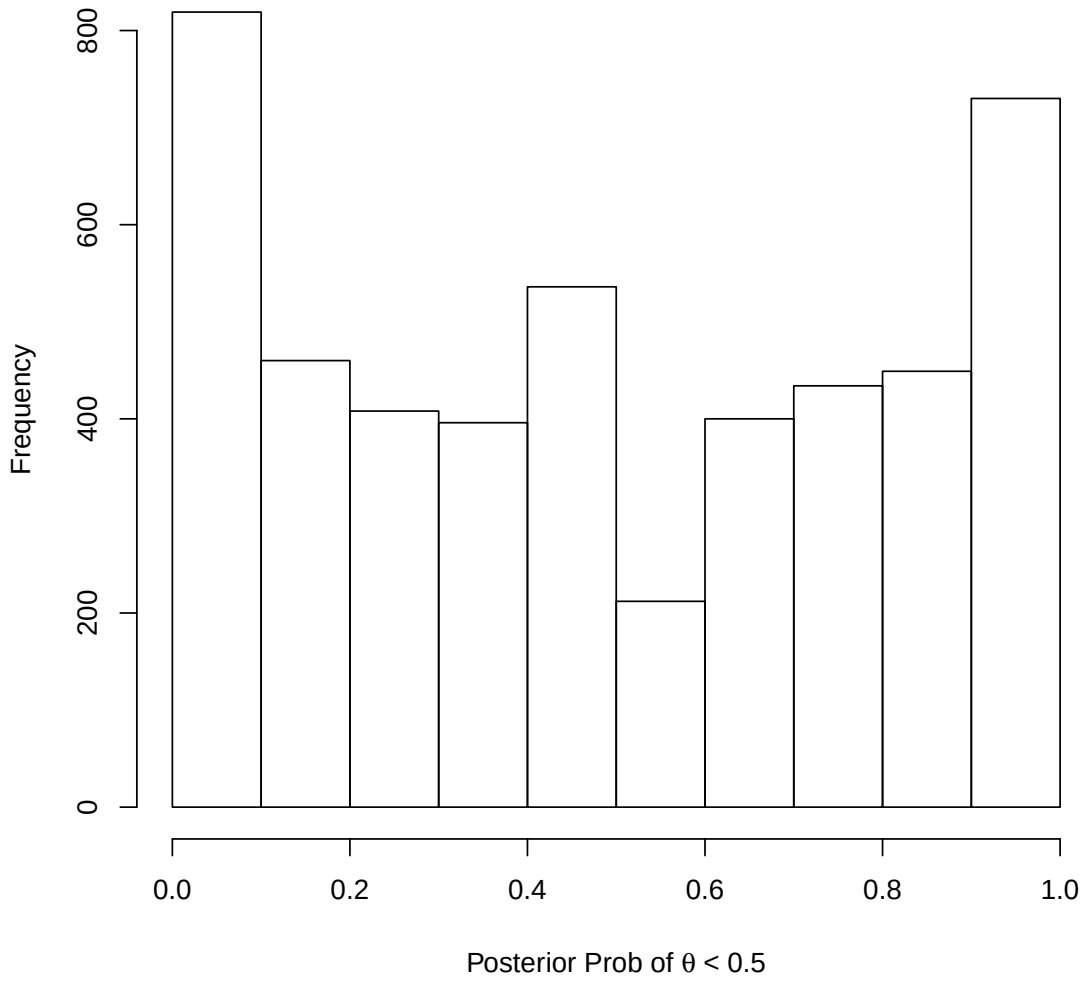
  $$\Pr(\theta < 0.5|y) < 0.01 \qquad \text{or} \qquad \Pr(\theta < 0.5|y) > 0.99$$

- Bayes Factors for $\theta = 0.5$: calculate the BF and compare it to some threshold for indicating worth of further study, e.g., if reciprocal of the Bayes factor is greater than 150
- Posterior probability of $\theta = 0.5$: like BFs, but with assumed prior support π_0 for $\theta = 0.5$, compare $\mathbb{P}[\theta = 0.5] = \frac{BF\,\pi_0/(1-\pi_0)}{1+BF\,\pi_0/(1-\pi_0)}$ to a chosen threshold

Analysis of ASE Data

Histogram of the posterior probabilities $\Pr(\theta < 0.5|y)$ for the 4,844 ASE genes.

Many have probabilities θ close to 0 or 1, indicating allele specific expression (ASE).

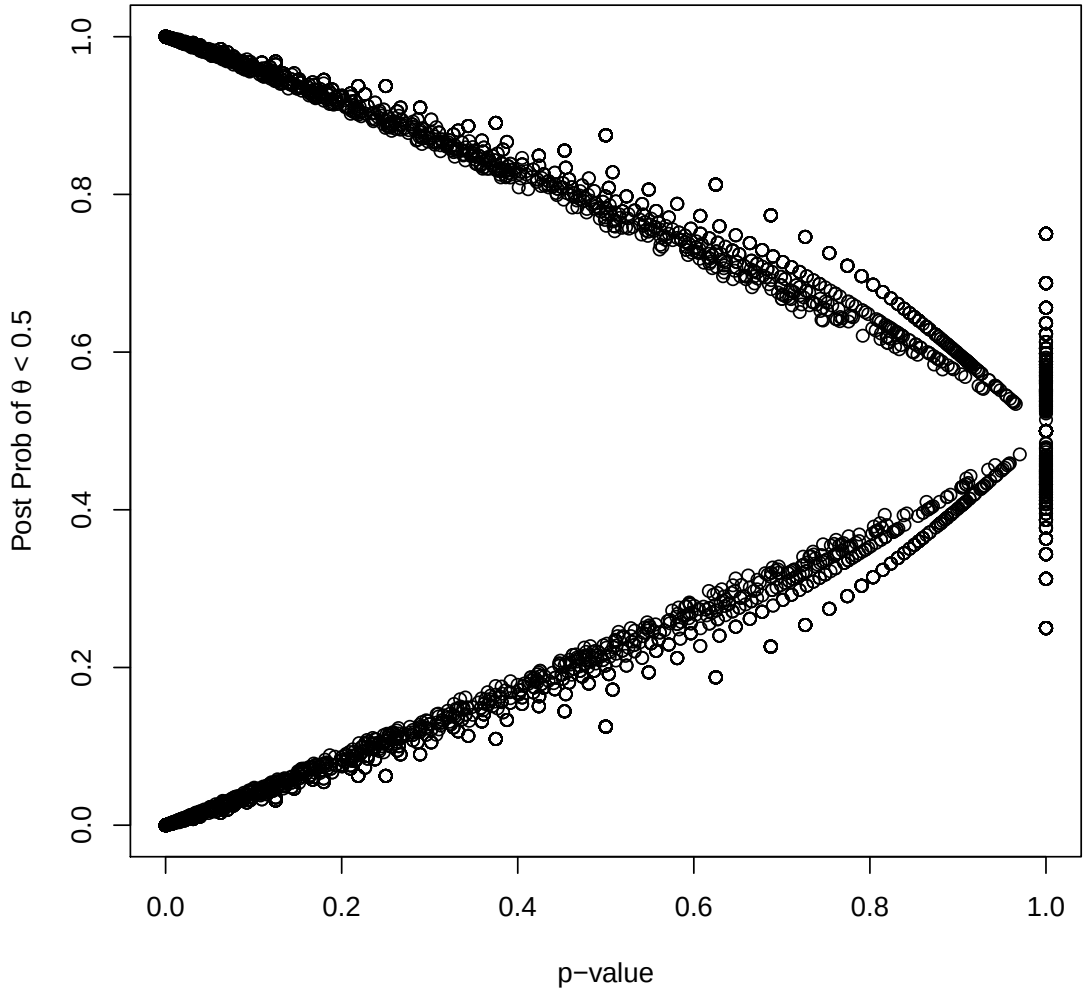


Analysis of ASE Data

Right: plotting $\Pr(\theta < 0.5|y)$ versus the p-values from an exact two-sided test.

Generally, small p-values have posterior probabilities close to 0 and 1.

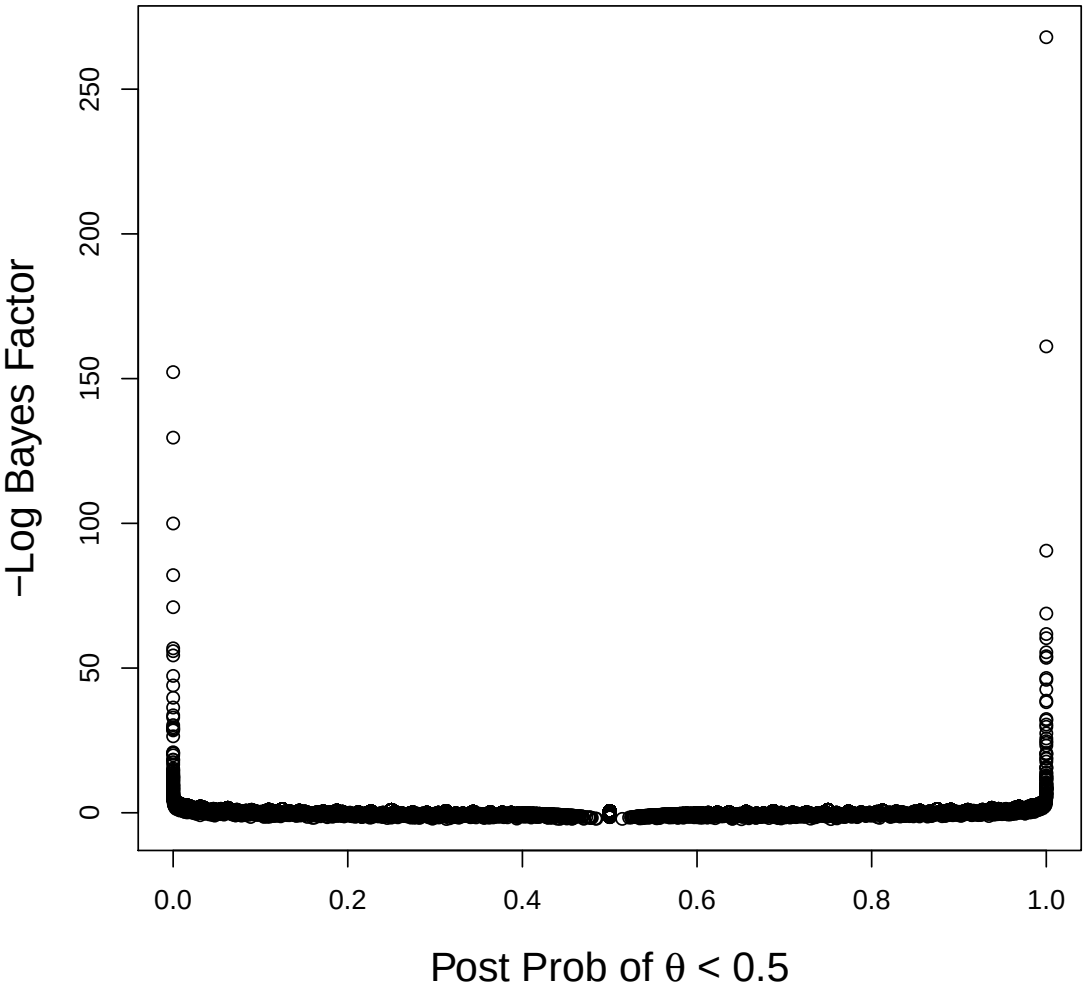
The weird lines are due to discreteness of the data.



Analysis of ASE Data

Plotting -log(BF) against posterior probability Pr($\theta < 0.5|y$) (right):

- Large BF values correspond to strong evidence of ASE
- This agrees with classical testing: large BF values corresponding to Pr($\theta < 0.5|y$) being close to 0 and 1.



Analysis of ASE Data

- Applying Bonferroni correction to control the family wise error rate at 0.05, gives a p-value threshold of $0.05/4844 = 10^{-5}$ and 111 'discoveries'. (More on this later!)
- There were 278 genes with $\Pr(\theta < 0.5|y) < 0.01$ and 242 genes with $\Pr(\theta < 0.5|y) > 0.99$.
- Following Kass & Raftery's guideline for very strong evidence that $1/BF > 150$, there would be 197 discoveries. For less stringent evidence, i.e. strong and very strong (reciprocal BF > 20) we make 359 discoveries.

For this form of hypothesis, the rankings of p-value and BF are very similar, but choosing the calibration – i.e. the threshold – remains challenging. (More on this in Session 9)

ASE output

Below are summaries of the ASE analysis – ordered according to `logBFr`, the reciprocal Bayes factor – so high numbers correspond to strong evidence against the null. `postprob` is the posterior probability of $\theta < 0.5$.

```
allvals <- data.frame(Nsum,ysum,pvals,postprob,logBFr)
oBF <- order(-logBFr)
orderallvals <- allvals[oBF,]
head(orderallvals)
     Nsum ysum         pvals     postprob   logBFr
4751  437    6 5.340324e-119 1.000000e+00 267.9572
4041  625   97  1.112231e-72 1.000000e+00 161.1355
2370  546  468  8.994944e-69 2.621622e-69 152.2517
2770  256  245  1.127211e-58 2.943484e-59 129.6198
tail(orderallvals)
     Nsum ysum     pvals  postprob    logBFr
824   761  382 0.9422103 0.4567334 -2.086604
2163  776  390 0.9142477 0.4429539 -2.091955
3153  769  384 1.0000000 0.5143722 -2.097079
2860 1076  546 0.6474878 0.3129473 -2.146555
```

Bayes Logistic Regression

To understand how binomial proportions p_i vary with covariates $\boldsymbol{x}_i$, we often turn to logistic regression models:

$$\begin{aligned} Y_i|p_i &\sim \text{Binomial}(N_i, p_i) \\ \log\left(\frac{p_i}{1-p_i}\right) &= \beta_0 + \beta_1 x_{i1} + \cdots + \beta_J x_{iJ} \end{aligned}$$

It is no longer possible to carry out a conjugate analysis by picking a convenient prior, but a common prior choice is to take

$$\pi(\beta_0, \beta_1, \ldots, \beta_J) = \prod_{j=0}^{J} N(0, \tau_j^2),$$

for fixed values τ_j^2, $j = 0, 1, \ldots, J$.

Logistic regression: LHON

We consider case-control data for Leber Hereditary Optic Neuropathy (LHON) disease with genotype data for marker rs6767450:

	CC $x = 0$	CT $x = 1$	TT $x = 2$	Total
Cases	6	8	75	89
Controls	10	66	163	239
Total	16	74	238	328

- We let $x = 0, 1, 2$ represent the number of T alleles (coded alleles, in most GWAS-type analysis – or imputed dosage)
- Let $p(x)$ represent the probability of being a case, given x copies of the T allele.

Logistic regression: LHON

For case-control studies, a popular choice is the *multiplicative odds model*:

$$\frac{p(x)}{1-p(x)} = \exp(\beta_0) \times \exp(\beta_1 x),$$

with a binomial likelihood, i.e. independent outcomes for each individual. (This is also called a *logistic regression model*.) It can also be written as stating that

$$\text{logit}\left(\Pr(Y=1|X=x)\right) = \log\left(\frac{p(x)}{1-p(x)}\right) = \beta_0 + \beta_1 x$$

or equivalently that

$$p(x) = \Pr(Y=1|X=x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}.$$

Logistic regression: LHON

Interpretation of these coefficients:

- $\exp(\beta_0)$, the odds of a sampled individual being a case, is of little interest given the case-control sampling
- $\exp(\beta_1)$ is the odds ratio describing the multiplicative change in odds for have one T allele versus zero T alleles.
- $\exp(2\beta_1)$ is the odds ratio describing the multiplicative change in odds for two T alleles versus zero T alleles.
- The odds ratio $\exp(\beta_1)$ approximates the *relative risk* a.k.a. *risk ratio*, for a rare disease
- Very usefully, under plausible conditions then $\exp(\beta_1)$'s odds ratio is the same as we'd observe in a (long!) prospective study

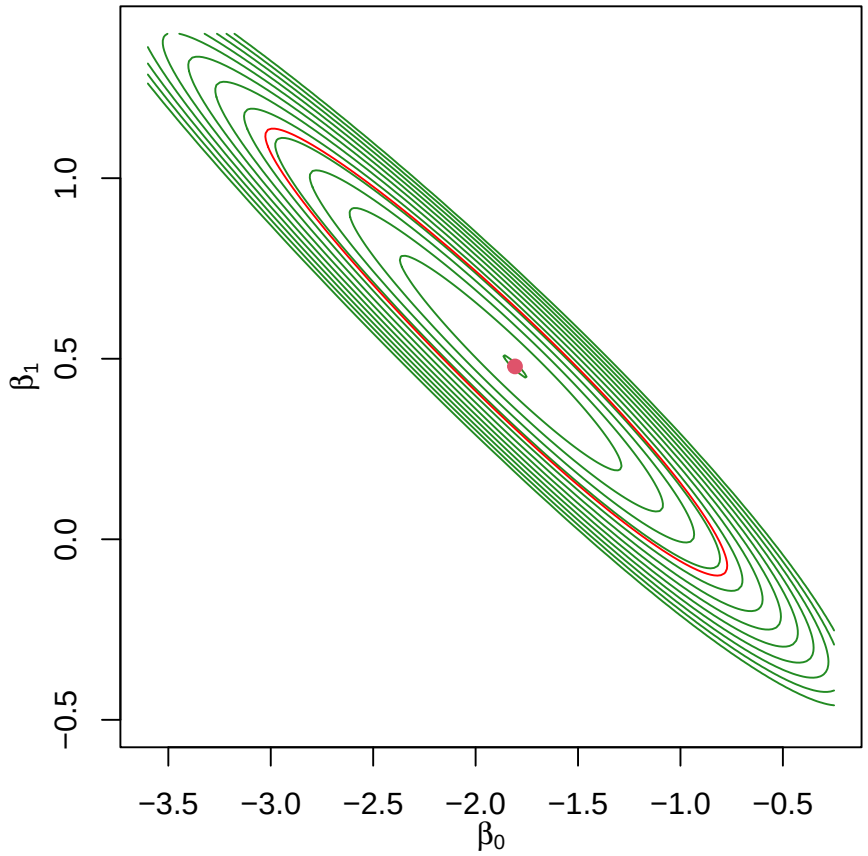
Logistic regression: LHON

A Bayesian analysis adds a prior on β_0 and β_1 – which we need to choose.

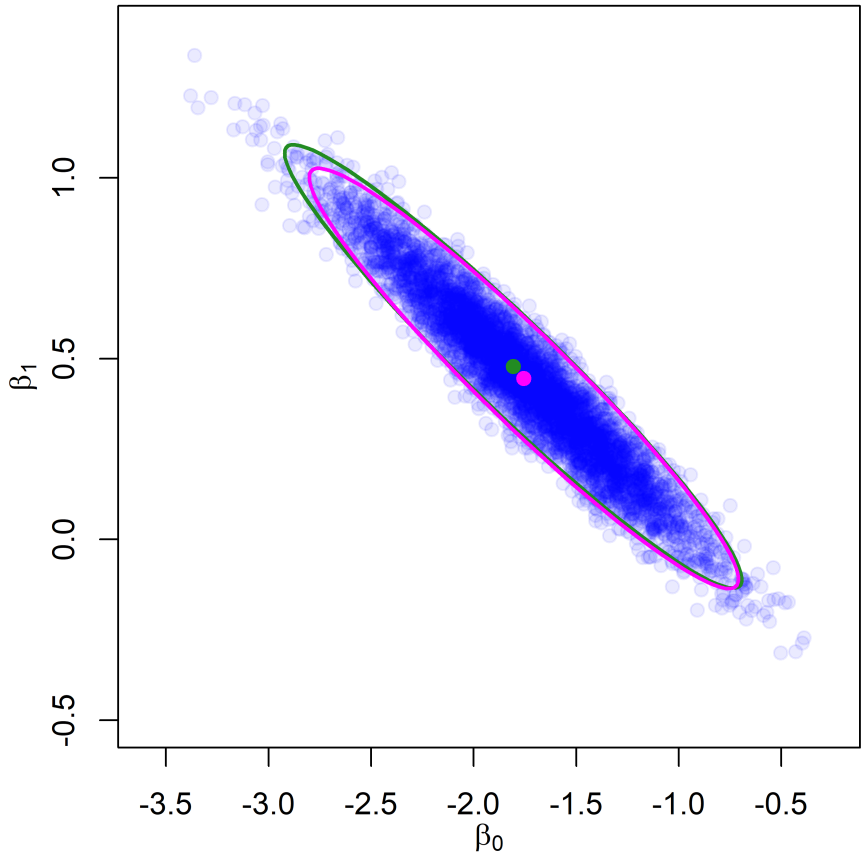
- The intercept, β_0, is typically very well-identified by the data, and so the likelihood dominates most priors; we will use a diffuse $N(0, 10)$ prior
- For the log odds ratio β_1, for a *light* constraint we use a Normal prior in which there is
  - 50% support for odds ratios above/below 1, i.e. log odds ratios above/below 0
  - 95% support for odds ratios below 5 (a large effect for a single variant)
- To center the prior at zero, we use $N(0, \sigma^2)$ – which has 95% point 1.645σ, so $\sigma = \log(5)/1.645 = 0.978$ gives the prior SD

Implementing this with rejection sampling for 5000 samples... (code on the course site)

Logistic regression: LHON



Likelihood contours ($\times 2$ increments) and approx 95% confidence region



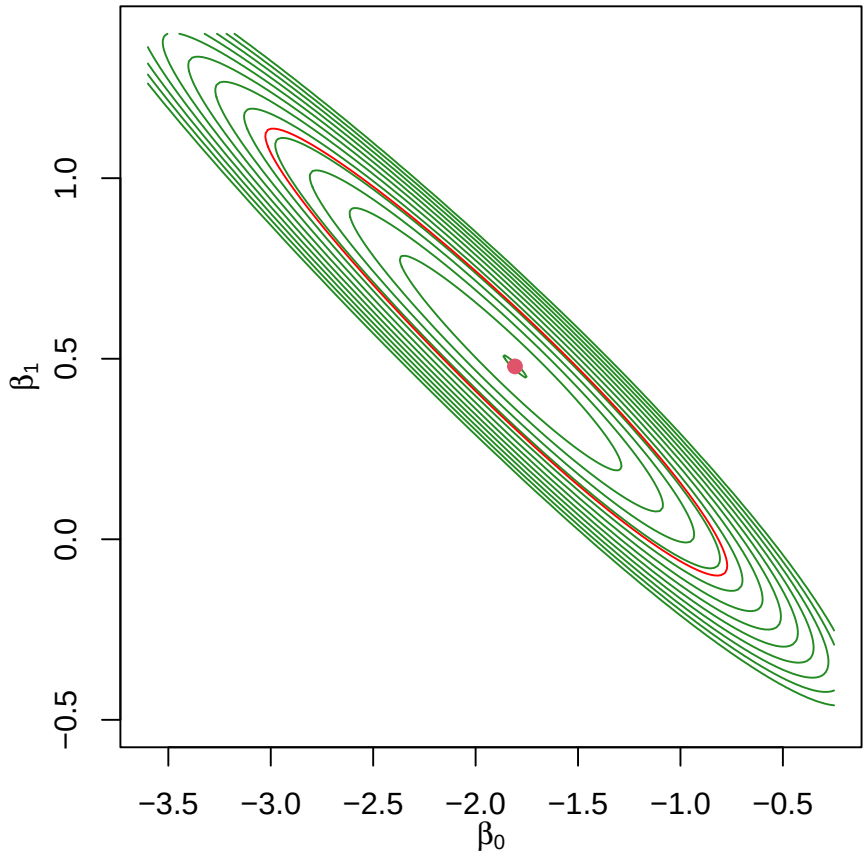
Posterior samples, approx 95% confidence region and 95% credible region

Logistic regression: LHON

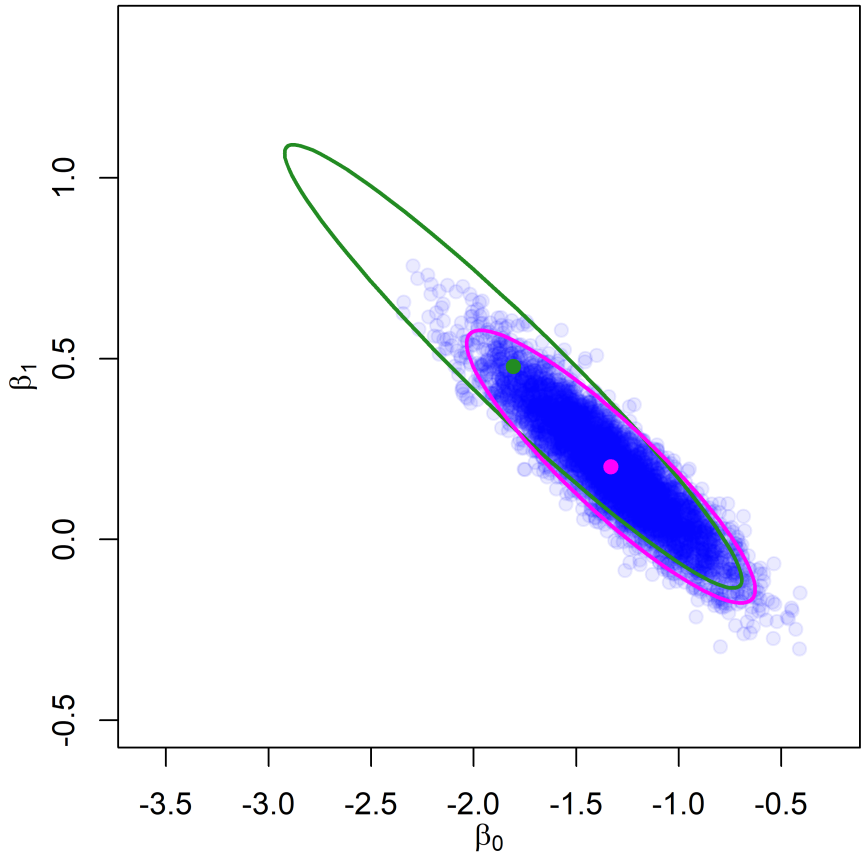
Prior belief in allele-specific odds ratios of 5 is (subjectively) optimistic: even for strongly heritable traits, most variants don't do much.

- To show the impact, we re-analyze the LHON data with the same diffues prior on β_0 but a much tighter prior on β_1
- We use a $N(0, \sigma^2)$ prior where σ^2 ensures the 97.5% point of the prior is $\log(1.5) = 0.41$
- In other words, we have prior probability 0.95 that the odds ratio lies between 2/3 and 3/2

Logistic regression: LHON



Likelihood contours (×2 increments) and approx 95% confidence region



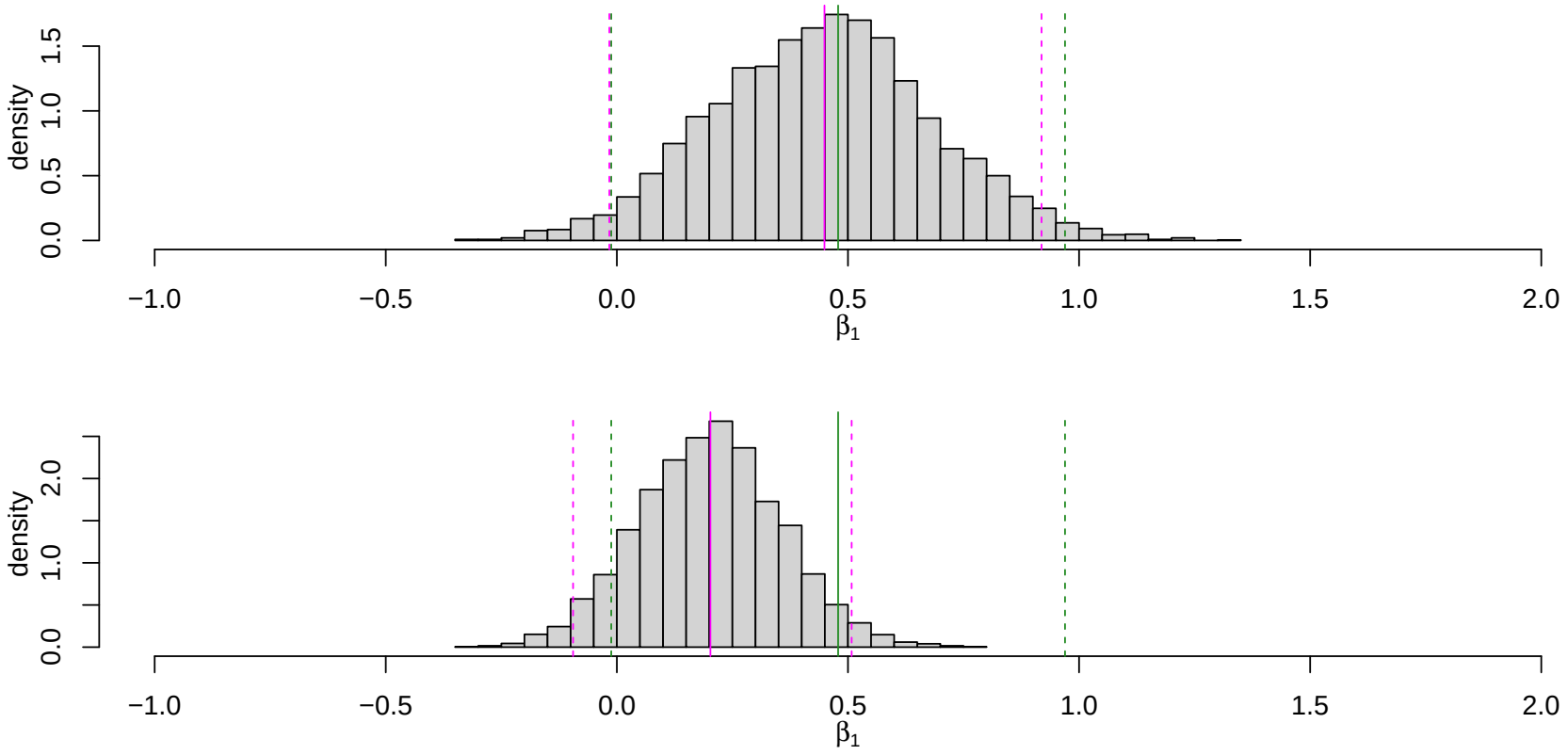
Posterior samples, approx 95% confidence region and 95% credible region

Logistic regression: LHON

- Clearly, the informative prior is adding non-trivial amounts of information
- The tighter prior on β_1 (vertical direction) ends up increasing precision on β_0 (horizontal) because the likelihood contours 'slope'
- The slope of the posterior contours is less steep – reflecting the priors impact, in addition to location and scale

Logistic regression: LHON

Histograms of the posteriors for β_1 from less (L) and more (R) informative priors, with approx 95% confidence interval and quantile-based 95% credible intervals.



Prediction

- Suppose we see y successes out of N trials, and now wish to obtain a *predictive distribution* for a future experiment with M trials.
- Let $Z \in \{0, 1, \ldots, M\}$ be that experiment's number of successes
- Predictive distribution:

$$\begin{aligned}
\Pr(z|y) &= \int_0^1 p(z, \theta|y) d\theta \\
&= \int_0^1 \Pr(z|\theta, y) p(\theta|y) d\theta \\
&= \int_0^1 \underbrace{\Pr(z|\theta)}_{\text{binomial}} \times \underbrace{p(\theta|y)}_{\text{posterior}} \, d\theta
\end{aligned}$$

  where we move between lines 2 and 3 because z is *conditionally independent* of y **given** θ, i.e.,

$$\Pr(z|\theta, y) = \Pr(z|\theta).$$

Prediction

Continuing with the calculation:

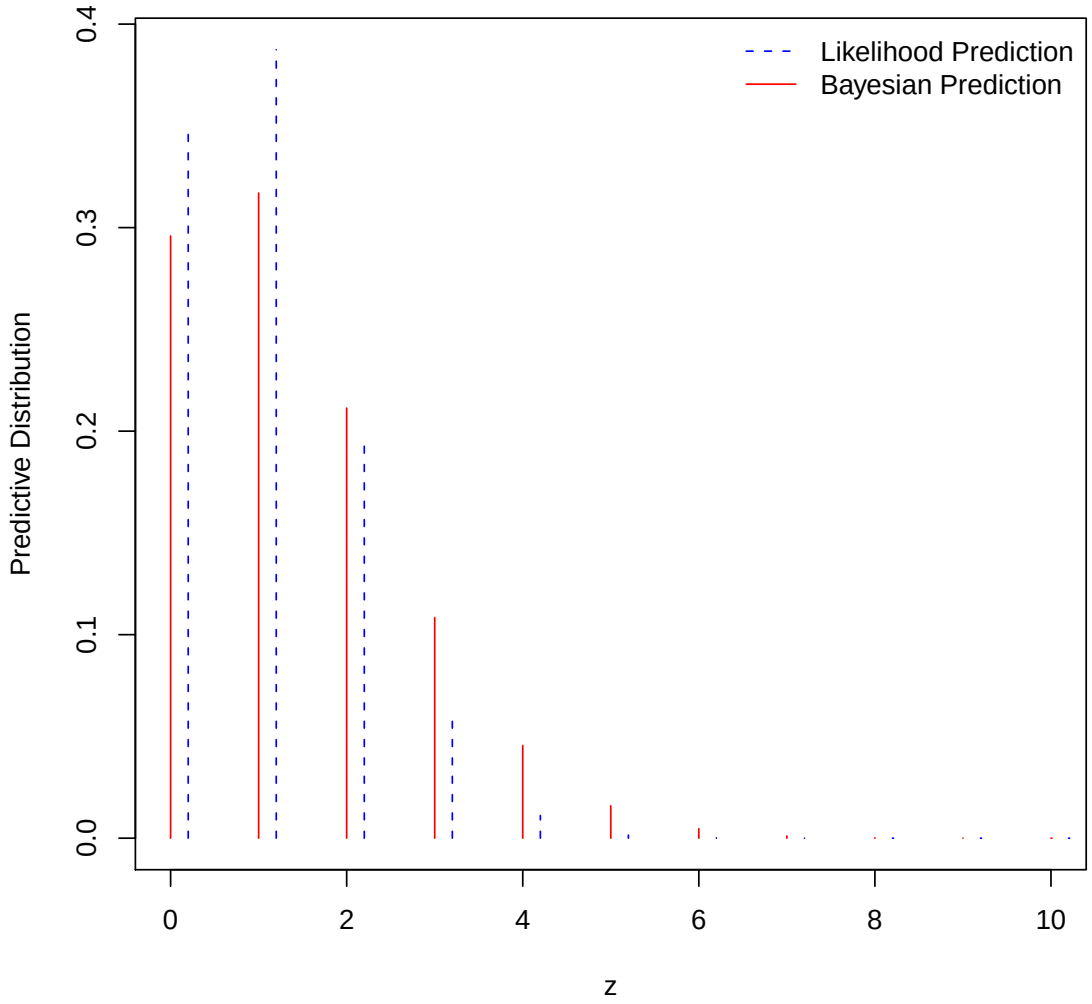
$$\begin{aligned}
\Pr(z|y) &= \int_0^1 \Pr(z|\theta) \times p(\theta|y) d\theta \\
&= \int_0^1 \left(\begin{array}{c} M \\ z \end{array} \right) \theta^z (1-\theta)^{M-z} \\
&\times \frac{\Gamma(N+a+b)}{\Gamma(y+a)\Gamma(N-y+b)} \theta^{y+a-1} (1-\theta)^{N-y+b-1} d\theta \\
&= \left(\begin{array}{c} M \\ z \end{array} \right) \frac{\Gamma(N+a+b)}{\Gamma(y+a)\Gamma(N-y+b)} \int_0^1 \theta^{y+a+z-1} (1-\theta)^{N-y+b+M-z-1} d\theta \\
&= \left(\begin{array}{c} M \\ z \end{array} \right) \frac{\Gamma(N+a+b)}{\Gamma(y+a)\Gamma(N-y+b)} \frac{\Gamma(a+y+z)\Gamma(b+N-y+M-z)}{\Gamma(a+b+N+M)}
\end{aligned}$$

for $z = 0, 1, \ldots, M$.

A likelihood approach would take the predictive distribution as Binomial$(M, \widehat{\theta})$ with $\widehat{\theta} = y/N$: this does not account for estimation uncertainty, and so tends to be anti-conservative.

Prediction

Likelihood and Bayesian predictive distribution of seeing $z = 0, 1, \ldots, M = 10$ successes, after observing $y = 2$ out of $N = 20$ successes (with $a = b = 1$).



Predictive Distribution: A General Approach

The posterior and sampling distributions won't usually combine so conveniently.

In general, we may form a *Monte Carlo* estimate of the predictive distribution:

$$\begin{aligned} p(z|y) &= \int p(z|\theta)p(\theta|y)d\theta \\ &= \mathbb{E}_{\theta|y}[p(z|\theta)] \\ &\approx \frac{1}{S}\sum_{s=1}^{S} p(z|\theta^{(s)}) \end{aligned}$$

where $\theta^{(s)} \sim p(\theta|y)$, $s = 1, \ldots, S$, is a sample from the posterior.

This provides an estimate of the predictive distribution at the point z.

Predictive Distribution: A General Approach

Alternatively, we may sample from $p(z|\theta^{(s)})$ a large number of times to reconstruct the predictive distribution.

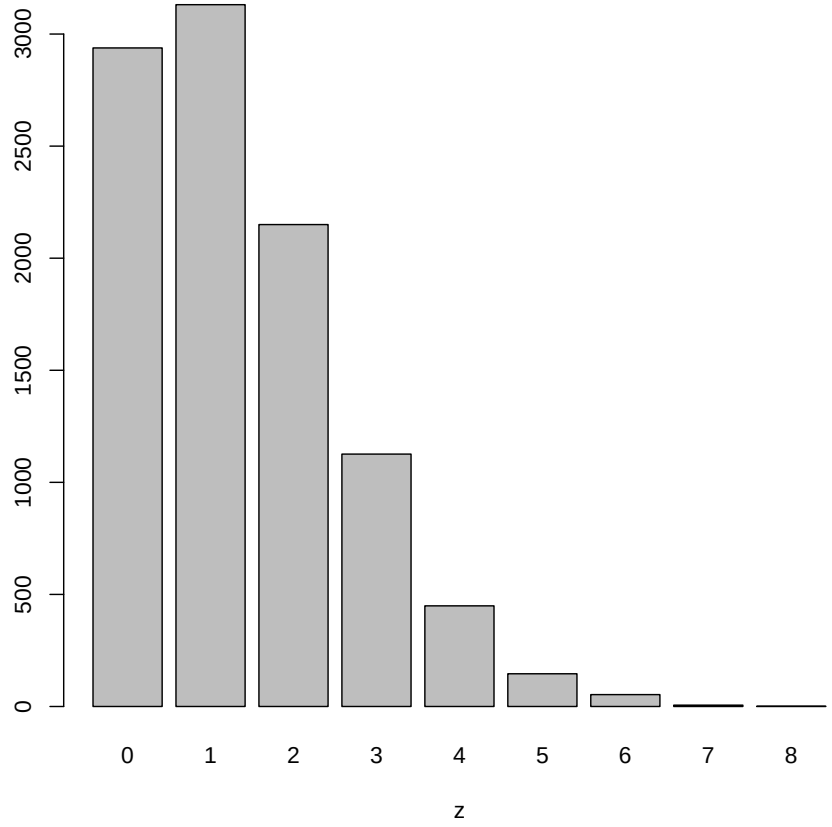
- First sample from the posterior:

$$\theta^{(s)}|y \sim p(\theta|y).$$

- Next sample from the likelihood:

$$z^{(s)}|\theta^{(s)} \sim p(z|\theta^{(s)}),$$

  for $s = 1, \ldots, S$.
- To give a sample $z^{(s)}$ from the posterior – see example, right, with $S = 10,000$ samples



Summary

- Predictions are very natural under the Bayesian approach.
- Monte Carlo sampling provides flexibility of inference
- All this lecture considered Binomial sampling, for which there is only a single parameter. For more parameters, prior specification and computing becomes more challenging...as we shall see
- For estimation and with middle to large sample sizes, conclusions from Bayesian and non-Bayesian approaches often coincide. One-sided tests are similarly ecumenical
- For two-sided testing it's more complex, as discussed in Lecture 9.

Appendix: Bayesian Sequential Updating

- We show how probabilistic beliefs are updated as we receive more data.
- Suppose the data arrives sequentially via two experiments:
  1. Experiment 1: (y_1, N_1).
  2. Experiment 2: (y_2, N_2).
- Prior 1: $\theta \sim \text{Beta}(a, b)$.
- Likelihood 1: $y_1|\theta \sim \text{Binomial}(N_1, \theta)$.
- Posterior 1: $\theta|y_1 \sim \text{Beta}(a + y_1, b + N_1 - y_1)$.
- This posterior forms the prior for experiment 2.
- Prior 2: $\theta \sim \text{Beta}(a^\star, b^\star)$ where $a^\star = a + y_1$, $b^\star = b + N_1 - y_1$.
- Likelihood 2: $y_2|\theta \sim \text{Binomial}(N_2, \theta)$.
- Posterior 2: $\theta|y_1, y_2 \sim \text{Beta}(a^\star + y_2, b^\star + N_2 - y_2)$.
- Substituting for $a^\star, b^\star$:

$$\theta|y_1, y_2 \sim \text{Beta}(a + y_1 + y_2, b + N_1 - y_1 + N_2 - y_2).$$

Appendix: Bayesian Sequential Updating

- Schematically:

$$(a, b) \to (a + y_1, b + N_1 - y_1) \to (a + y_1 + y_2, b + N_1 - y_1 + N_2 - y_2)$$

- Suppose we obtain the data in one go as $y^\star = y_1 + y_2$ successes from $N^\star = N_1 + N_2$ trials.
- The posterior is

$$\theta | y^\star \sim \text{Beta}(a + y^\star, b + N^\star - y^\star),$$

  which is the same as when we receive in two separate instances.

Appendix: Birth Weight Example

We show an example provided by Wang *et al*, with data on 189 births to women seen in a particular obstetric clinic.

The response variable `LOW` is a binary outcome indicating birth weight less than 2500 grams. We also see:

- `LOW`: Low birth weight; (0 = $\geq$ 2500g; 1 = $<$ 2500g)
- `AGE`: Mother's age
- `LWT`: Mother's weight
- `RACE`: Listed race of mother; (1 = white; 2 = black; 3 = other)
- `SMOKE`: Smoking status during pregnancy; (0 = no; 1 = yes)
- `HT`: History of hypertension; (0 = no; 1 = yes)
- `UI`: Presence of uterine irritability; (0 = no; 1 = yes)
- `FTV`: Number of physician visits during the first trimester.

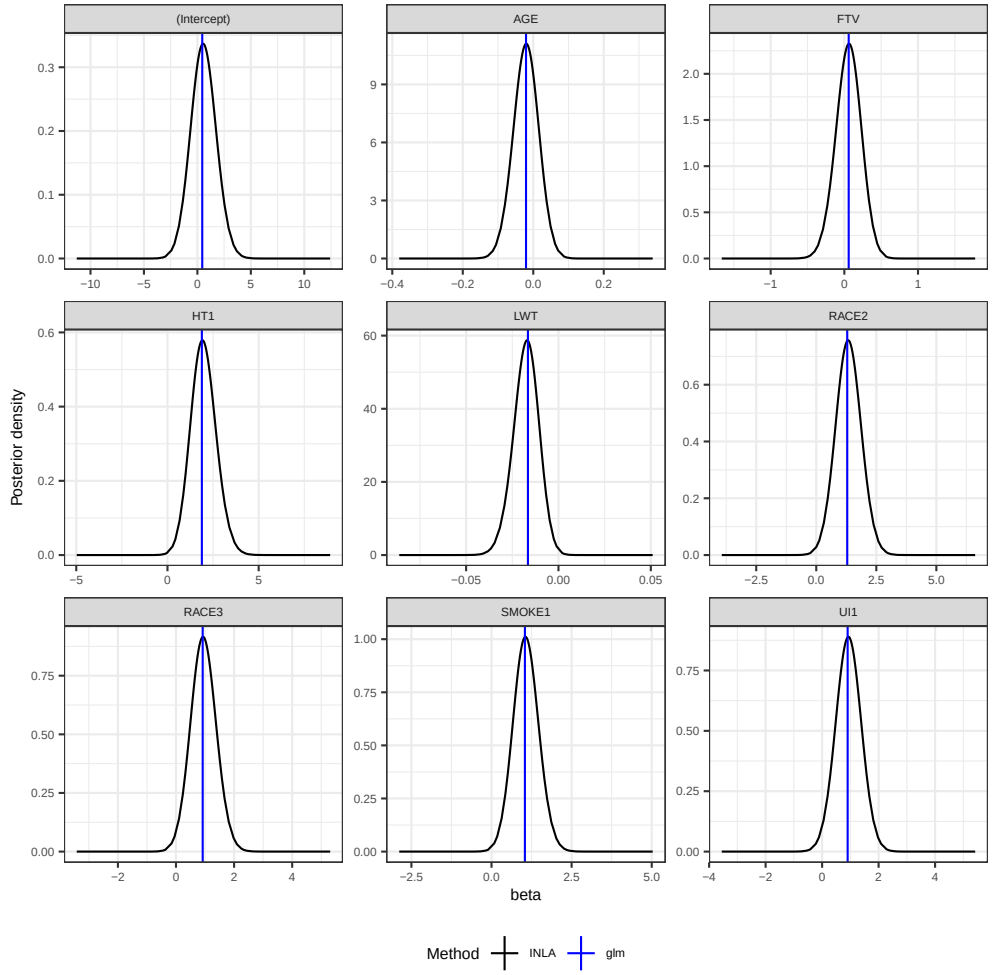
Appendix: Birth Weight Example

Under priors with large variances τ_j we obtain very similar inference under likelihood and Bayesian analyses.

	MLE	Std. Error	Posterior Mean	Posterior SD
(Intercept)	0.455	1.185	0.567	1.186
AGE	-0.021	0.036	-0.021	0.036
LWT	-0.017	0.007	-0.018	0.007
RACE2	1.290	0.528	1.340	0.528
RACE3	0.919	0.436	0.946	0.436
SMOKE1	1.042	0.395	1.075	0.395
HT1	1.885	0.695	1.974	0.694
UI1	0.904	0.449	0.933	0.449
FTV	0.059	0.172	0.056	0.172

Appendix: Birth Weight Example

Posteriors $p(\beta_j|\boldsymbol{y})$, $j = 0, \ldots, 8$.



Appendix: Birth Weight Example

Posteriors for odds $p(e^{\beta_0}|\boldsymbol{y})$ and odds ratios $p(e^{\beta_j}|\boldsymbol{y})$, $j = 1, \ldots, 8$.