

# 2016 SISG Module 17: Bayesian Statistics for Genetics

## Lecture 8: Generalized Linear Modeling

Jon Wakefield

Departments of Statistics and Biostatistics  
University of Washington

## Outline

### Introduction and Motivating Examples

### Generalized Linear Models

- Definition

- Bayes Linear Model

- Bayes Logistic Regression

- Generalized Linear Mixed Models

### Approximate Bayes Inference

- The Approximation

### Conclusions

## Introduction

- In this lecture we will discuss Bayesian modeling in the context of **Generalized Linear Models (GLMs)**.
- This discussion will include the addition of random effects, i.e. the class of **Generalized Linear Mixed Models (GLMMs)**.
- Estimation via the quick **INLA** technique will be demonstrated, along with its R implementation.
- An **approximation technique** that is useful in the context of Genome Wide Association Studies (GWAS) (in which the number of tests is large) will also be introduced.
- The accompanying R code allows the analyses presented here to be replicated.

## Motivating Example: Logistic Regression

- We consider case-control data for the disease 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

- Let  $x = 0, 1, 2$  represent the number of T alleles, and  $p(x)$  the probability of being a case, given  $x$  copies of the  $T$  allele.

## Motivating Example: Logistic Regression

- For such case-control data one may fit the **multiplicative odds model**:

$$\frac{p(x)}{1 - p(x)} = \exp(\alpha) \times \exp(\theta x),$$

with a **binomial likelihood**.

- Interpretation:**

- $\exp(\alpha)$  is of little interest given the case-control sampling.
- $\exp(\theta)$  is the odds ratio describing the **multiplicative change in risk** for one T allele versus zero T alleles.
- $\exp(2\theta)$  is the odds ratio describing the **multiplicative change in risk** for two T alleles versus zero T alleles.
- Odds ratios approximate the **relative risk** for a rare disease.

A Bayesian analysis adds a prior on  $\alpha$  and  $\theta$ .

## Motivating Example: FTO Data Revisited

### Recall

- $Y = \text{weight}$
- $x_g = \text{fto heterozygote} \in \{0, 1\}$
- $x_a = \text{age in weeks} \in \{1, 2, 3, 4, 5\}$

We will examine the fit of the model

$$E[Y|x_g, x_a] = \beta_0 + \beta_g x_g + \beta_a x_a + \beta_{\text{int}} x_g x_a,$$

with independent normal errors, and compare with a Bayesian analysis.

## Motivating Example: RNA Seq with Replicates

- We report an experiment carried out in a collaboration with Caitlin Connelly and Josh Akey (UW Genome Sciences), see Connelly *et al.* (2014) for further details.
- Start with two haploid yeast strains (individuals).
- From these we obtain RNA-Seq data, where we isolate RNA from the two individuals, fragment and sequence it using next-generation sequencing, and map the sequencing reads back to the genome to generate RNA levels in the form of counts of the number of sequencing reads mapping at each gene.
- Also mate the two haploid yeast strains together to form a diploid hybrid. We again isolate RNA, fragment, and sequence it.
- Then take advantage of polymorphisms between the two strains in order to map reads to either of the two haploid individuals, giving us counts for the number of reads mapping to either one of the parental genomes in the diploid hybrid for each gene.

## Motivating Example: RNA Seq with Replicates

- We are interested in two questions from this data. First, we want to look for evidence of **trans** effects at each gene; in biological terms, this means that polymorphisms located far from the gene are responsible for differences in RNA levels.
- To detect this, look for genes where the difference between RNA levels in the haploids differs from the difference between RNA levels for the two parental strains in the diploid.
- Also interested in looking for **cis** effects, meaning polymorphisms near the gene itself are responsible for differences in RNA levels.
- We can detect **cis** effects as a difference in the count of reads mapping to each of the parental strains in the diploid at a gene.



## Motivating Example: RNA Seq Data, Statistical Model

- There are two replicates and so for each of  $N$  genes we obtain two sets of counts.
- For the diploid hybrid let  $Y_{ij}$  be the number of A alleles for gene  $i$  and replicate  $j$ , and  $N_{ij}$  is the total number of counts, so that  $N_{ij} - Y_{ij}$  is the number of T alleles  $j = 1, 2$ .
- We fit a **hierarchical logistic regression model** starting with first stage:

$$Y_{ij} | N_{ij}, p_{ij} \sim \text{binomial}(N_{ij}, p_{ij})$$

so that  $p_{ij}$  is the probability of seeing an A read for gene  $i$  and replicate  $j$ .

- At the second stage:

$$\text{logit } p_{ij} = \theta_i + \epsilon_{ij}$$

where  $\epsilon_{ij} \sim \text{normal}(0, \sigma^2)$  represent random effects that allow for excess-binomial variation.

- In the model  $\theta_i$  is a parameter of interest – if a (say) 95% posterior interval estimate contains 0 then we have evidence of **cis** effects.

## Generalized Linear Models

- **Generalized Linear Models (GLMs)** provide a very useful extension to the linear model class.
- GLMs have three elements:
  1. The responses follow an **exponential family**.
  2. The mean model is **linear** in the covariates on some scale.
  3. A **link function** relates the mean of the data to the covariates.
- In a GLM the response  $y_i$  are independently distributed and follow an **exponential family**<sup>1</sup>,  $i = 1, \dots, n$ .
- **Examples:** Normal, Poisson, binomial.

---

<sup>1</sup>so that the distribution is of the form  $p(y_i|\theta_i, \alpha) = \exp(\{y_i\theta_i - b(\theta_i)\}/\alpha + c(y_i, \alpha))$ , where  $\theta_i$  and  $\alpha$  are scalars

## Generalized Linear Models

- The **link function**  $g(\cdot)$  provides the connection between the mean  $\mu = E[Y]$  and the **linear predictor**  $\mathbf{x}\beta$ , via

$$g(\mu) = \mathbf{x}\beta,$$

where  $\mathbf{x}$  is a vector of explanatory variables and  $\beta$  is a vector of regression parameters.

- For **normal data**, the usual link is the identity

$$g(\mu) = \mu = \mathbf{x}\beta.$$

- For **binary data**, a common link is the logistic

$$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right) = \mathbf{x}\beta.$$

- For **Poisson data**, a common link is the log

$$g(\mu) = \log(\mu) = \mathbf{x}\beta.$$

## Bayesian Modeling with GLMs

- For a generic GLM, with regression parameters  $\beta$  and a scale parameter  $\alpha$ , the **posterior** is

$$p(\beta, \alpha | \mathbf{y}) \propto p(\mathbf{y} | \beta, \alpha) \times p(\beta, \alpha).$$

- An immediate question is: How to specify a **prior distribution**  $p(\beta, \alpha)$ ?
- How to perform the **computations** required to summarize the posterior distribution (including the calculation of Bayes factors)?

## Bayesian Computation

Various approaches to computation are available:

- **Conjugate analysis** — the prior combines with likelihood in such a way as to provide analytic tractability (at least for some parameters).
- **Analytical Approximations** — asymptotic arguments used (e.g. Laplace).
- **Numerical integration.**
- **Direct (Monte Carlo) sampling** from the posterior, as we have already seen.
- **Markov chain Monte Carlo** — very complex models can be implemented, for example within the free software WinBUGS.
- **Integrated nested Laplace approximation (INLA).** Cleverly combines analytical approximations and numerical integration: we illustrate the use of this method in some detail.

## Integrated Nested Laplace Approximation (INLA)

- The homepage of the INLA software is here:  
<http://www.r-inla.org/home>
- There are also lots of example links at this website.
- The fitting of many common models is described here:  
<http://www.r-inla.org/models/likelihoods>
- INLA can fit GLMs, GLMMs and many other useful model classes.

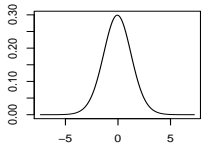
## INLA for the Linear Model

- The model is

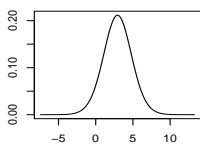
$$Y = E[Y|x_g, x_a] = \beta_0 + \beta_g x_g + \beta_a x_a + \beta_{\text{int}} x_g x_a + \epsilon$$

where  $\epsilon|\sigma^2 \sim_{iid} N(0, \sigma^2)$ .

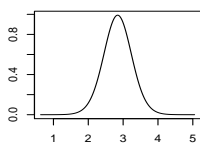
- This model has five parameters: the four fixed effects are  $\beta_0, \beta_g, \beta_a, \beta_{\text{int}}$  and the error variance is  $\sigma^2$ , which is known as a hyperparameter (note that in `inla` inference is reported for the precision  $\sigma^{-2}$ ).
- In general, posterior distributions can be summarized graphically or via numerical summaries.
- In Figures 1 and 2 give posterior marginal distributions for the fixed effects and hyperparameter  $\sigma^{-2}$ , respectively, under an analysis with relatively flat priors.

**PostDens [(Intercept)]**

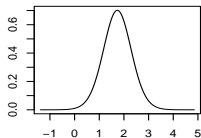
Mean = -0.061 SD = 1.371

**PostDens [linxg]**

Mean = 2.932 SD = 1.937

**PostDens [linxa]**

Mean = 2.842 SD = 0.413

**PostDens [linxint]**

Mean = 1.733 SD = 0.584

Figure 1 : Marginal distributions of the intercept and regression coefficients.



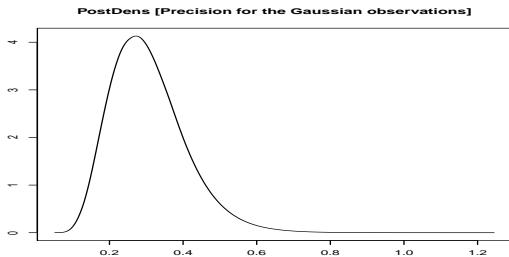


Figure 2 : Marginal distribution of the error precision.

## INLA for the Linear Model

- As with a non-Bayesian analysis, model checking is important and in Figure 3 we present a number of diagnostic plots.
- Plots:
  - (a) Normality of residuals? Sample size is quite small.
  - (b) Is the relationship with age linear?
  - (c) Mean variance relationship?
  - (d) Overall fit.
- For these data, the model assumptions look reasonable.

## FTO Diagnostic Plots

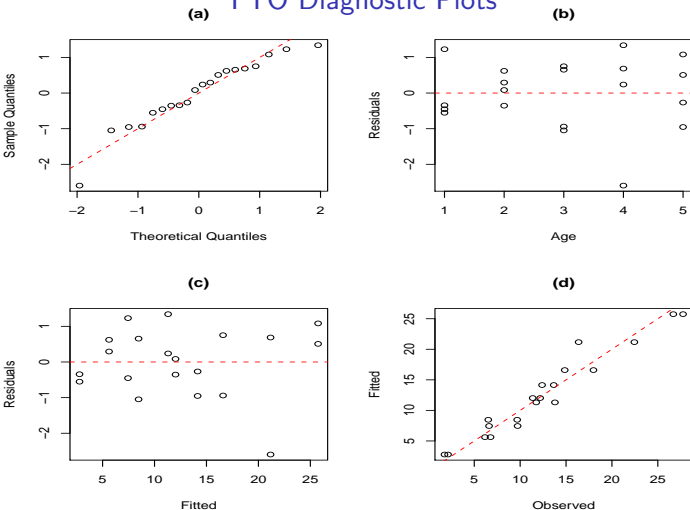


Figure 3 : Plots to assess model adequacy: (a) Normal QQ plot, (b) residuals versus age, (c) residuals versus fitted, (d) fitted versus observed.

## Bayes Logistic Regression

- The **likelihood** is

$$Y(x)|p(x) \sim \text{Binomial}(N(x), p(x)), \quad x = 0, 1, 2.$$

- Logistic link:**

$$\log\left(\frac{p(x)}{1-p(x)}\right) = \alpha + \theta x$$

- The **prior** is

$$p(\alpha, \theta) = p(\alpha) \times p(\theta)$$

with

- $\alpha \sim \text{normal}(\mu_\alpha, \sigma_\alpha)$  and
- $\theta \sim \text{normal}(\mu_\theta, \sigma_\theta)$ . where  $\mu_\alpha, \sigma_\alpha, \mu_\theta, \sigma_\theta$  are constant that are specified to reflect **prior beliefs**.

## Prior Choice for Positive Parameters

- It is convenient to specify lognormal priors for a positive parameter, for example  $\exp(\beta)$  (the odds ratio) in a logistic regression analysis.
- One may specify two quantiles of the distribution, and directly solve for the two parameters of the lognormal.
- Denote by  $\theta \sim \text{LogNormal}(\mu, \sigma)$  the lognormal distribution for a generic positive parameter  $\theta$  with  $E[\log \theta] = \mu$  and  $\text{var}(\log \theta) = \sigma^2$ , and let  $\theta_1$  and  $\theta_2$  be the  $q_1$  and  $q_2$  quantiles of this prior.
- In our example,  $\theta = \exp(\beta)$ .
- Then it is straightforward to show that

$$\mu = \log(\theta_1) \left( \frac{z_{q_2}}{z_{q_2} - z_{q_1}} \right) - \log(\theta_2) \left( \frac{z_{q_1}}{z_{q_2} - z_{q_1}} \right), \quad \sigma = \frac{\log(\theta_1) - \log(\theta_2)}{z_{q_1} - z_{q_2}}. \quad (1)$$

## Prior Choice for Positive Parameters

- As an example, suppose that for the odds ratio  $e^\beta$  we believe there is a 50% chance that the odds ratio is less than 1 and a 95% chance that it is less than 5; with  $q_1 = 0.5$ ,  $\theta_1 = 1.0$  and  $q_2 = 0.95$ ,  $\theta_2 = 5.0$ , we obtain lognormal parameters  $\mu = 0$  and  $\sigma = (\log 5)/1.645 = 0.98$ .
- The density is shown in Figure 4.

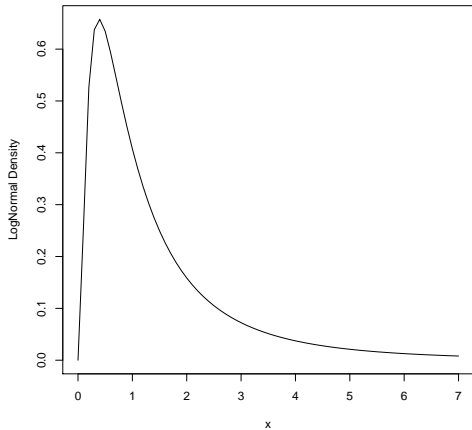


Figure 4 : Lognormal density with 50% point 1 and 95% point 5.

## Logistic Regression Example

- In the second analysis we specify

$$\alpha \sim \text{normal}(0, 1/0.1)$$

$$\theta \sim \text{normal}(0, W)$$

where  $W$  is such that the 97.5% point of the prior is  $\log(1.5)$ , i.e. we believe the odds ratio lies between  $2/3$  and  $3/2$  with probability 0.95.

- The marginal distributions are given in Figure 25



## Logistic Marginal Plots

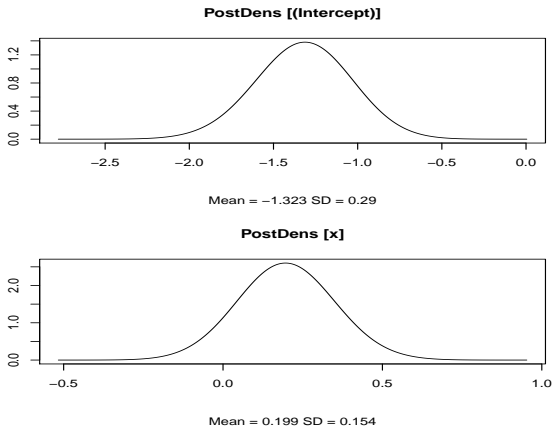


Figure 5 : Posterior marginals for the intercept  $\alpha$  and the log odds ratio  $\theta$ .

## The RNA-Seq Data: INLA Analysis

- Recall there are two replicates and so for each of  $N$  genes we obtain two sets of counts.
- For the diploid hybrid, let  $Y_{ij}$  be the number of A alleles for gene  $i$  and replicate  $j$ , and  $N_{ij}$  is the total number of counts,  $j = 1, 2$ .
- We fit a **hierarchical logistic regression model** starting with first stage:

$$Y_{ij} | N_{ij}, p_{ij} \sim \text{binomial}(N_{ij}, p_{ij})$$

so that  $p_{ij}$  is the probability of seeing an A read for gene  $i$  and replicate  $j$ .

- At the second stage:

$$\text{logit } p_{ij} = \theta_i + \epsilon_{ij}$$

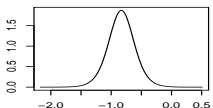
where  $\epsilon_{ij} | \sigma^2 \sim \text{normal}(0, \sigma^2)$  represent random effects that allow for excess-binomial variation; there are a pair for each gene.

- The  $\theta_i$  parameters are taken as **fixed effects** with a relatively flat prior.
- $\exp(\theta_i)$  is the odds of seeing an A read for gene  $i$ .
- Figures 6, 7 and 8 summarize inference.

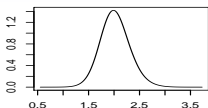
○○○○○○○○○

○○○○○  
○○○○○  
○○○○○  
●○○○○

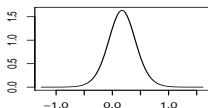
○○○○○○○○○○

**PostDens [as.factor(xvar)1]**

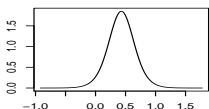
Mean = -0.832 SD = 0.23

**PostDens [as.factor(xvar)2]**

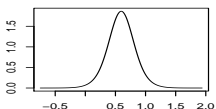
Mean = 2.022 SD = 0.293

**PostDens [as.factor(xvar)3]**

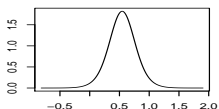
Mean = 0.173 SD = 0.256

**PostDens [as.factor(xvar)4]**

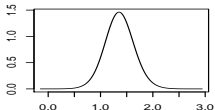
Mean = 0.43 SD = 0.232

**PostDens [as.factor(xvar)5]**

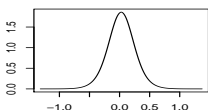
Mean = 0.596 SD = 0.23

**PostDens [as.factor(xvar)6]**

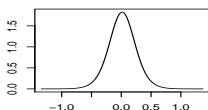
Mean = 0.546 SD = 0.235

**PostDens [as.factor(xvar)7]**

Mean = 1.363 SD = 0.283

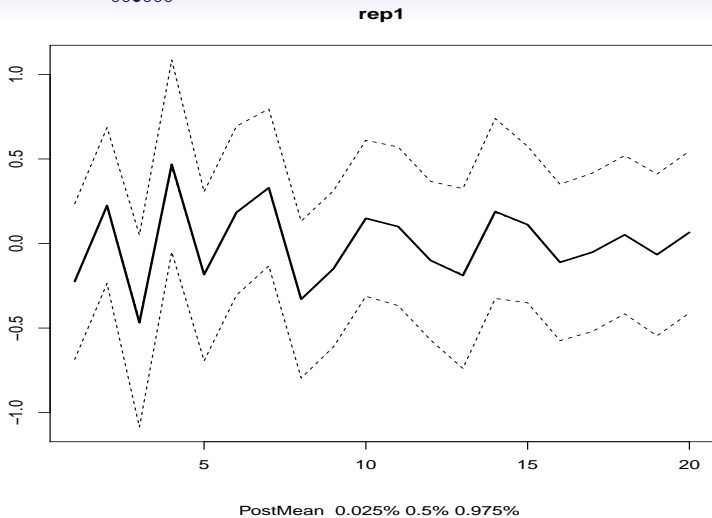
**PostDens [as.factor(xvar)8]**

Mean = 0.028 SD = 0.231

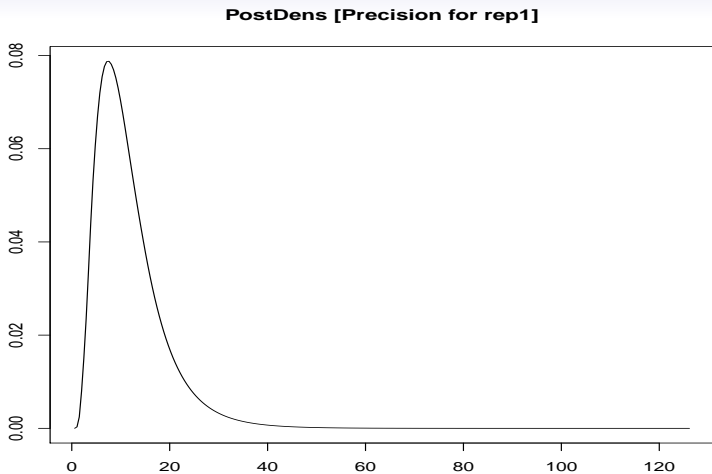
**PostDens [as.factor(xvar)9]**

Mean = 0.015 SD = 0.234

**Figure 6 :** Posterior marginals for the first 9 gene effects (compare with zero for evidence of cis effects). We plot 9 rather than all 10 for display purposes.



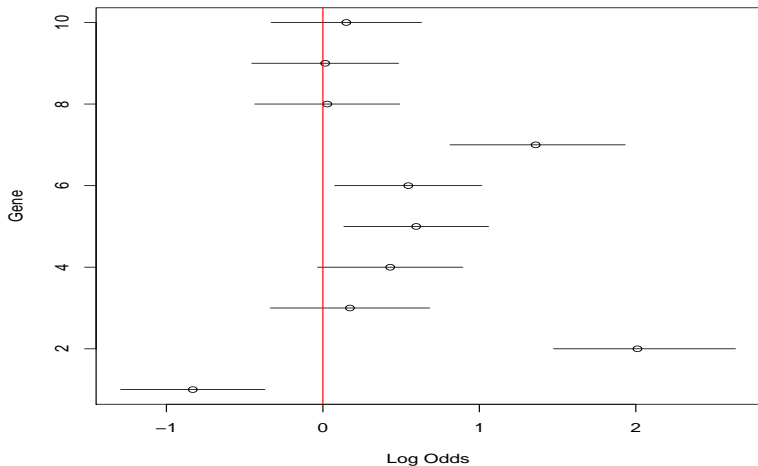
**Figure 7 :** Posterior quantiles for 20 random effects, which allow excess-binomial variation.



**Figure 8 :** Posterior marginal for precision of random effects.

## An Informative Summary for the RNA-Seq Data

- We extract the 95% intervals and posterior medians for the log odds of being an A allele.
- Comparison with 0 (in Figure 9) gives an indication of cis effects.
- Genes 1, 2, 5, 6, 7 show evidence of cis effects.



**Figure 9 :** Posterior marginal intervals for posterior of interest. Genes with posterior intervals that do not include zero, show evidence of cis effects.

## Approximate Bayes Inference

- Particularly in the context of a large number of experiments, a quick and accurate model is desirable.
- We describe such a model in the context of a [GWAS](#).
- This model is relevant when the sample size in each experiment is large.
- We first recap the [normal-normal](#) Bayes model.
- Subsequently, we describe the approximation and provide an example.



## Recall: The Normal-Normal Model

- For the model
  - **Prior:**  $\theta \sim \text{normal}(\mu_0, \tau_0^2)$  and
  - **Likelihood:**  $Y_1, \dots, Y_n | \theta \sim \text{normal}(\theta, \sigma^2)$ .
- Posterior

$$\theta | y_1, \dots, y_n \sim \text{normal}(\mu_n, \tau^2)$$

where

$$\begin{aligned} \text{var}(\theta | y_1, \dots, y_n) = \tau^2 &= [1/\tau_0^2 + n/\sigma^2]^{-1} \\ \text{Precision} = 1/\tau^2 &= 1/\tau_0^2 + n/\sigma^2 \end{aligned}$$

and

$$\begin{aligned} E[\theta | y_1, \dots, y_n] = \mu_n &= \frac{\mu_0/\tau_0^2 + \bar{y}n/\sigma^2}{1/\tau_0^2 + n/\sigma^2} \\ &= \mu_0 \left( \frac{1/\tau_0^2}{1/\tau_0^2 + n/\sigma^2} \right) + \bar{y} \left( \frac{n/\sigma^2}{1/\tau_0^2 + n/\sigma^2} \right) \end{aligned}$$

## A Normal-Normal Approximate Bayes Model

- Consider again the **logistic regression model**

$$\text{logit } p_i = \alpha + x_i\theta$$

with interest focusing on  $\theta$ .

- We require **priors** for  $\alpha, \theta$ , and some numerical/analytical technique for estimation/Bayes factor calculation.
- As discussed in Lecture 6 Wakefield (2007, 2009) considered replacing the likelihood by the approximation

$$p(\theta|\hat{\theta}) \propto p(\hat{\theta}|\theta)p(\theta)$$

where

- $\hat{\theta}|\theta \sim \text{normal}(\theta, V)$  – the asymptotic distribution of the MLE,
- $\theta \sim \text{normal}(0, W)$  – the prior on the log RR. Can choose  $W$  so that 95% of relative risks lie in some range, e.g.  $[2/3, 1.5]$ .

## Posterior Distribution

- Under the alternative, the **posterior distribution** for the log odds ratio  $\theta$  is

$$\theta|\hat{\theta} \sim \text{normal}(r\hat{\theta}, rV)$$

where

$$r = \frac{W}{V + W}.$$

- Hence, we have **shrinkage** to the prior mean of 0.
- The **posterior median for the odds ratio** is  $\exp(r\hat{\theta})$  and a 95% credible interval is

$$\exp(r\hat{\theta} \pm 1.96\sqrt{rV}).$$

- Note that as  $W \rightarrow \infty$  and/or  $V \rightarrow 0$  (which occurs as we gather more data) the non-Bayesian point and interval estimates are recovered (since  $r \rightarrow 1$ ).

## A Normal-Normal Approximate Bayes Model

- We are interested in the hypotheses:  $H_0 : \theta = 0$ ,  $H_1 : \theta \neq 0$  and evaluation of the **Bayes factor**

$$\text{BF} = \frac{p(\hat{\theta}|H_0)}{p(\hat{\theta}|H_1)}.$$

- Using the approximate likelihood and normal prior we obtain:

$$\text{Approximate Bayes Factor} = \frac{1}{\sqrt{1-r}} \exp\left(-\frac{Z^2}{2}r\right),$$

$$\text{with } Z = \frac{\hat{\theta}}{\sqrt{V}}, r = \frac{W}{V+W}.$$

## A Normal-Normal Approximate Bayes Model

- The approximation can be combined with a Prior Odds =  $\pi_0/(1 - \pi_0)$  to give

$$\text{Posterior Odds on } H_0 = \frac{\text{BFDP}}{1 - \text{BFDP}} = \text{ABF} \times \text{Prior Odds}$$

where BFDP is the **Bayesian False Discovery Probability**.

- BFDP depends on the **power**, through  $r$ .
- For **implementation**, all that we need from the data is the  $Z$ -score and the standard error  $\sqrt{V}$ , or a confidence interval.
- Hence, published results that report confidence intervals can be converted into Bayes factors for interpretation (see later lecture).
- The approximation relies on large sample sizes, so the normal distribution of the estimator provides a good summary of the information in the data.

## Combination of Data Across Studies

- Suppose we wish to combine data from **two studies** where we assume a common log odds ratio  $\theta$ .
- The estimates from the two studies are  $\hat{\theta}_1, \hat{\theta}_2$  with standard errors  $\sqrt{V_1}$  and  $\sqrt{V_2}$ .
- The Bayes factor is

$$\frac{p(\hat{\theta}_1, \hat{\theta}_2 | H_0)}{p(\hat{\theta}_1, \hat{\theta}_2 | H_1)}$$

- The approximate Bayes factor is

$$\text{ABF}(\hat{\theta}_1, \hat{\theta}_2) = \text{ABF}(\hat{\theta}_1) \times \text{ABF}(\hat{\theta}_2 | \hat{\theta}_1) \quad (2)$$

where

$$\text{ABF}(\hat{\theta}_2 | \hat{\theta}_1) = \frac{p(\hat{\theta}_2 | H_0)}{p(\hat{\theta}_2 | \hat{\theta}_1, H_1)}$$

and

$$p(\hat{\theta}_2 | \hat{\theta}_1, H_1) = E_{\theta | \hat{\theta}_1} [p(\hat{\theta}_2 | \theta)]$$

so that the density is averaged with respect to the posterior for  $\theta$ .

- **Important Point:** The Bayes factors are not independent.

## Combination of Data Across Studies

- This leads to an approximate Bayes factor (which summarizes the data from the two studies) of

$$\text{ABF}(\hat{\theta}_1, \hat{\theta}_2) = \sqrt{\frac{W}{RV_1V_2}} \exp \left\{ -\frac{1}{2} \left( Z_1^2 RV_2 + 2Z_1Z_2R\sqrt{V_1V_2} + Z_2^2 RV_1 \right) \right\}$$

where

- $R = W/(V_1W + V_2W + V_1V_2)$
- $Z_1 = \frac{\hat{\theta}_1}{\sqrt{V_1}}$  and
- $Z_2 = \frac{\hat{\theta}_2}{\sqrt{V_2}}$  are the usual  $Z$  statistics.
- The ABF will be small (evidence for  $H_1$ ) when the **absolute values** of  $Z_1$  and  $Z_2$  are **large** and they are of the **same sign**.

## Combination of Data Across Studies: The General Case

- Suppose we have  $K$  studies with estimates  $\hat{\theta}_k$  and asymptotic variances  $V_k$ ,  $k = 1, \dots, K$ .
- Assume a common underlying parameter  $\theta$ .
- The Bayes factor is given by

$$\begin{aligned}
 \text{BF}_K &= \frac{p(\hat{\theta}_1, \dots, \hat{\theta}_K | H_0)}{p(\hat{\theta}_1, \dots, \hat{\theta}_K | H_1)} \\
 &= \frac{\prod_{k=1}^K (2\pi V_k)^{-1/2} \exp\left(-\frac{\hat{\theta}_k^2}{2V_k}\right)}{\int \prod_{k=1}^K (2\pi V_k)^{-1/2} \exp\left(-\frac{(\hat{\theta}_k - \theta)^2}{2V_k}\right) (2\pi W)^{-1/2} \exp\left(-\frac{\theta^2}{2W}\right) d\theta} \\
 &= \sqrt{W \left( W^{-1} + \sum_{k=1}^K V_k^{-1} \right)} \exp \left[ -\frac{1}{2} \left( \sum_{k=1}^K \frac{\hat{\theta}_k}{V_k} \right)^2 \left( W^{-1} + \sum_{k=1}^K V_k^{-1} \right)^{-1} \right]
 \end{aligned}$$



## Combination of Studies: The General Case

- The posterior is given by

$$\theta | \hat{\theta}_1, \dots, \hat{\theta}_K \sim \text{normal}(\mu, \sigma^2)$$

where

$$\mu = \left( \sum_{k=1}^K \frac{\hat{\theta}_k}{V_k} \right) \left( W^{-1} + \sum_{k=1}^K V_k^{-1} \right)^{-1}$$

$$\sigma^2 = \left( W^{-1} + \sum_{k=1}^K V_k^{-1} \right)^{-1}$$

## Example of Combination of Studies in a GWAS

- We illustrate how reported confidence intervals can be converted to Bayesian summaries.
- Frayling *et al.* (2007) report a GWAS for Type II diabetes.
- For SNP rs9939609:

Stage	Estimate (CI)	$p$ -value	$-\log_{10}$ BF	Pr( $H_0$  data) with prior:	
				1/5,000	1/50,000
1st	1.27 (1.16–1.37)	$6.4 \times 10^{-10}$	7.28	0.00026	0.0026
2nd	1.15 (1.09–1.23)	$4.6 \times 10^{-5}$	2.72	0.905	0.990
Combined	–	–	13.8	$8 \times 10^{-11}$	$8 \times 10^{-10}$

- Combined evidence** is stronger than each **separately** since the point estimates are in agreement.
- For summarizing inference the (5%, 50%, 95%) points for the RR are:

Prior	1.00 (0.67–1.50)
First Stage	1.26 (1.17–1.36)
Combined	1.21 (1.15–1.27)

## Conclusions

- Computationally **GLMs** and **GLMMs** can now be fitted in a relatively straightforward way.
- **INLA** is very convenient and is being constantly improved.
- As with all analyses, it is crucial to check **modeling assumptions** (and there are usually more in a Bayesian analysis).
- **Markov chain Monte Carlo** provides an alternative for computation. **WinBUGS** is one popular implementation.
- Other MCMC possibilities include: **JAGS**, **BayesX**, **Stan**.

## References

- Connelly, C., Wakefield, J., and Akey, J. (2014). Evolution and architecture of chromatin accessibility and function in yeast. *PLoS Genetics*. To appear.
- Frayling, T., Timpson, N., Weedon, M., Zeggini, E., Freathy, R., and et al., C. L. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, **316**, 889–894.
- Wakefield, J. (2007). A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *American Journal of Human Genetics*, **81**, 208–227.
- Wakefield, J. (2009). Bayes factors for genome-wide association studies: comparison with  $p$ -values. *Genetic Epidemiology*, **33**, 79–86.