SISG 2017 Module 13
Bayesian Statistics for Genetics

Jon Wakefield & Ken Rice

UW Dept of Biostatistics

July, 2017

Welcome!

- This course provides an introduction to Bayesian statistics, as used in analysis of genetic data
- Applications will include: Hardy-Weinberg testing and estimation, detection of allele-specific expression, QTL mapping, testing in genome-wide association studies, mixture models, multiple testing in high throughput genomics
- The module provides a foundation for Module 17 (MCMC for Genetics) – which discusses the required calculations in much more depth

Logistics

- Course Text: P.D. Hoff (2009), A First Course in Bayesian Statistical Methods, Springer. [UW access]
- Supplementary Text: J.C. Wakefield (2013), Bayesian and Frequentist Regression Methods, Springer. [UW access]
- We will demonstrate analyses in R – you are encouraged to follow along
- Code and other materials (course notes, papers) are available at the course website:
  http://faculty.washington.edu/kenrice/sisgbayes/
- If you have no/limited R experience please pair up with someone who does. Experienced people – do this anyway!

Overview: Day 1

1. (Rice): Why Bayes? Introduction
2. (Wakefield): Review of Probability – start Binomial sampling
3. (Wakefield) Binomial Sampling
4. (Rice) Linear Regression. Continuous sampling

Overview: Day 2

5. (Wakefield) Multinomial Sampling

6. (Rice) Model Selection and Averaging

7. (Wakefield) Generalized linear modeling

8. (Rice) Meta analysis

Overview: Day 2.5

9. (Wakefield) Bayesian and frequentist testing: Single tests and multiple tests. Model comparison

10. (Rice) Decision theory. Software (WinBUGS/JAGS/INLA/Stan)