c. \( X \sim \text{Bin}(\pi, m) \), \( Y \sim \text{Bin}(\rho, n) \).

   i. “noninformative” “reference” prior on both.

   ii. Likelihood \( \pi^X (1 - \pi)^{m-X} \rho^Y (1 - \rho)^{n-Y} \)

   iii. Prior \( \pi^{-1} (1 - \pi)^{-1} \rho^{-1} (1 - \rho)^{-1} \)

   iv. More interesting parameterization \( \delta = \pi - \rho \in (-1, 1) \),
   \( \tau = \pi + \rho \in (|\delta|, 2 - |\delta|) \)

   v. \( \pi = (\delta + \tau)/2 \), \( \rho = (\tau - \delta)/2 \)

   vi. Posterior \( (\delta + \tau)^{X-1} (1 - \delta - \tau)^{m-X-1} (\tau - \delta)^{Y-1} (1 - \tau + \delta)^{n-Y-1} \)

   • The jacobian of the \((\pi, \rho) \rightarrow (\delta, \tau)\) transformation is constant, and will wash out of calculation.

WMS: 16.4-16.5

L. Bayesian hypothesis testing.

1. As before, decide between \( H_0 : \theta \in \Omega_0 \) vs. \( H_A : \theta \in \Omega_a \).
   a. Here I used notation similar to that of frequentist analysis.
   b. At present, no “null” and “alternate” subtext.

2. Choose hypothesis with highest posterior probability.

3. Often report posterior odds \( \frac{P[\Omega_0|\text{data}]}{P[\Omega_a|\text{data}]} \)

4. Factor \( B \) by which prior odds \( \frac{P[\Omega_0]}{P[\Omega_a]} \) was changed is
Lecture 13

called *Bayes factor*.

a. \[ B = \frac{P[\Omega_0|\text{data}] P[\Omega_a]}{P[\Omega_a|\text{data}] P[\Omega_0]} \]

b. When hypothesis \( \Omega_0 \) and \( \Omega_a \) are both simple, Bayes factor is the likelihood ratio.

c. Point hypotheses are only workable if there’s positive prior probability on them.

\[ \text{B: 4.6} \]

M. Bayesian Hierarchical Models

1. Bayesian alternative to frequentist random effects modeling.

2. Setup:

\[
X_{11}, \ldots, X_{1n_1} \sim i.i.d. \mathcal{N}(\theta_1, \sigma^2) \\
X_{21}, \ldots, X_{2n_2} \sim i.i.d. \mathcal{N}(\theta_2, \sigma^2) \\
\vdots \\
X_{k1}, \ldots, X_{kn_k} \sim i.i.d. \mathcal{N}(\theta_k, \sigma^2)
\]

3. \( \theta_k \sim i.i.d. \mathcal{N}(\mu, \tau^2) \)

4. \( \mu, \sigma \) and \( \tau \) given non-informative prior.

5. Since these are all conjugate priors, one can produce a normal posterior for \( \mu \).

6. Cf. frequentist approach \( X_{ji} = \mu + \eta_j + \epsilon_{ij}, \eta_j \sim \mathcal{N}(\sigma^2) \),
\[ \varepsilon_{ij} \sim \mathcal{N}(\tau^2). \]

7. Results similar.