comparing two proportions using Bayes factors: assumptions

Dr. Merlise Clyde

Duke University

parent's perceptions on bullying

	male	female
yes	34	61
no	52	61
not sure	4	0
total	90	122

Note: in the following analyses, "no" and "not sure" are combined

$$H_1:\ p_{\text{male}} = p_{\text{female}}$$

$$H_2:\ p_{\text{male}} \neq p_{\text{female}}$$

Source: https://commons.wikimedia.org/wiki/File:Bully_Free_Zone.jpg

assumptions for inference about probabilities

- independence within groups
- independence between groups
- common response rate between groups under H_2

  $$P(\text{Response}_{i,\text{male}} = \text{Yes} | p_{\text{male}}, H_2) = p_{\text{male}}$$

  $$P(\text{Response}_{i,\text{female}} = \text{Yes} | p_{\text{female}}, H_2) = p_{\text{female}}$$

- common response rate within groups under H_1

  $$P(\text{Response}_{i,\text{male}} = \text{Yes} | p,\ H_1) = p$$

  $$P(\text{Response}_{i,\text{female}} = \text{Yes} | p,\ H_1) = p$$

- do not need to check sample size!

conjugate prior and posterior distributions for beta-binomial

$$p_{\text{male}} \sim \text{Beta}(a_m,\ b_m)$$

a_m prior number of males reporting bullying

b_m prior number of males not reporting bullying

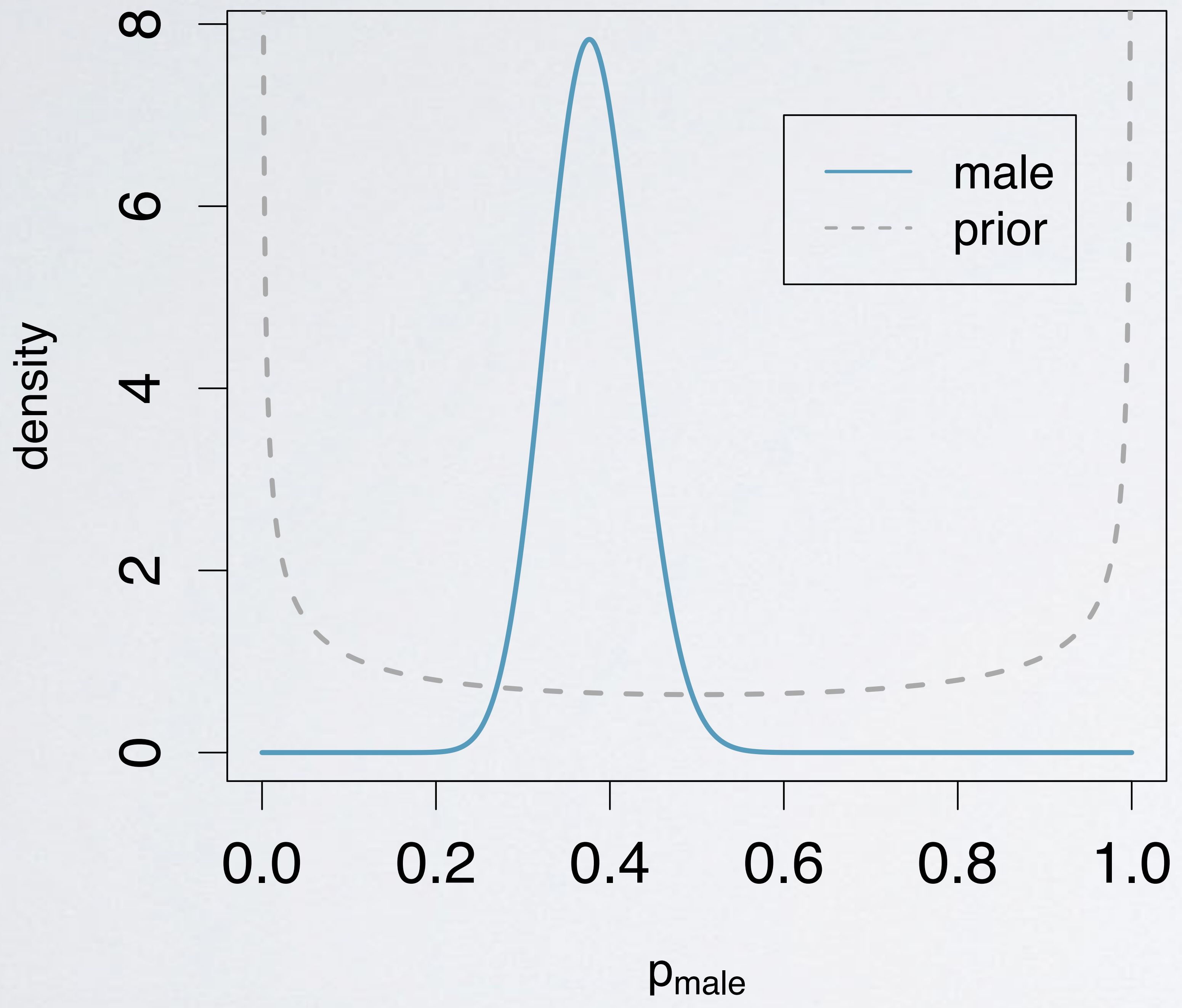
$a_m + b_m$ prior sample size

$$\Rightarrow p_{\text{male}} \mid \text{data} \sim \text{Beta}(R_m + a_m,\ n_m - R_m + b_m)$$

prior and posterior distributions for males under H_2

prior $p_{\text{male}} \sim \text{Beta}(1/2,\ 1/2)$

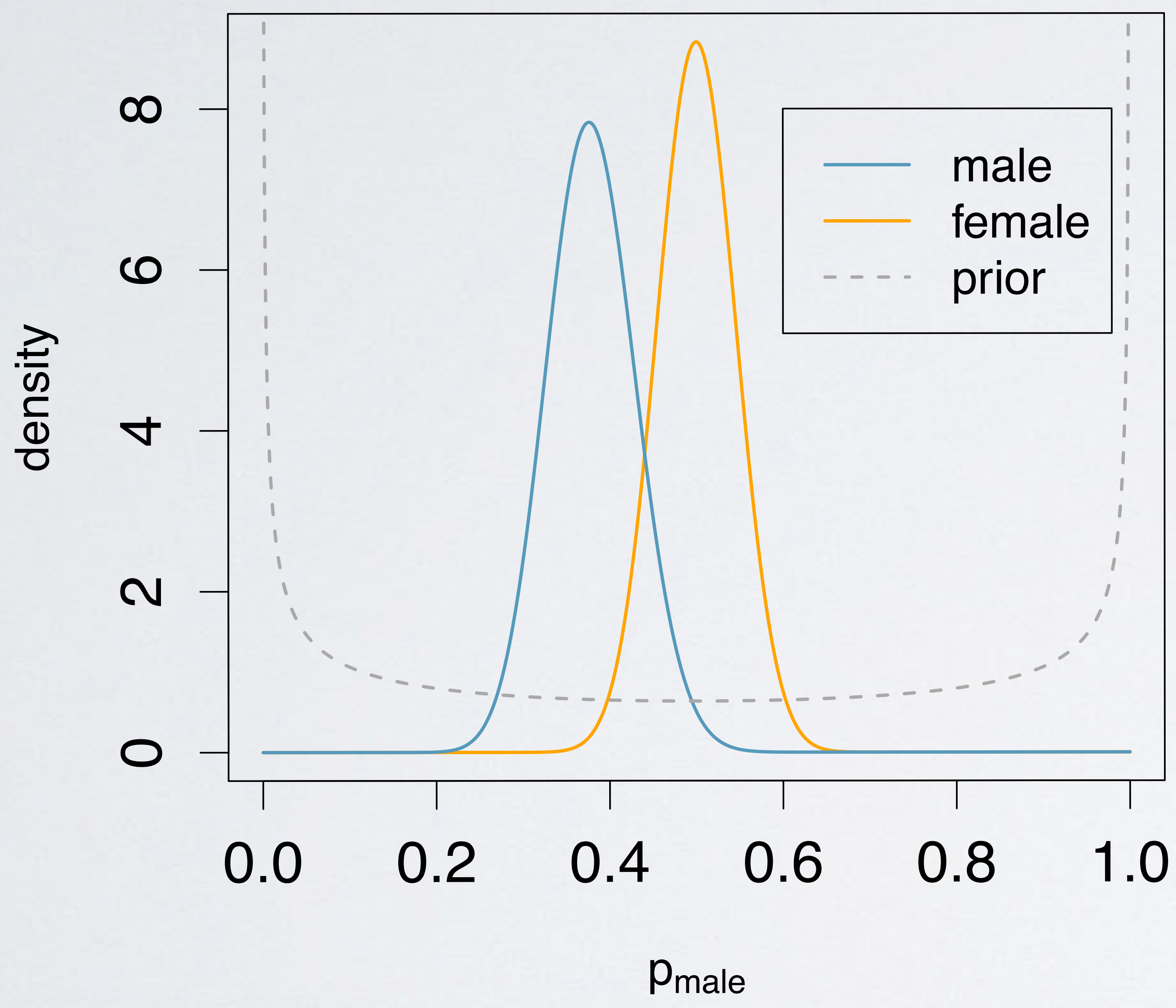
posterior $p_{\text{male}} \mid \text{data} \sim \text{Beta}(34 + 1/2,\ 56 + 1/2)$



prior and posterior distributions for females under H_2

prior $p_{\text{female}} \sim \text{Beta}(1/2,\, 1/2)$

posterior $p_{\text{female}} \mid \text{data} \sim \text{Beta}(61 + 1/2,\, 61 + 1/2)$



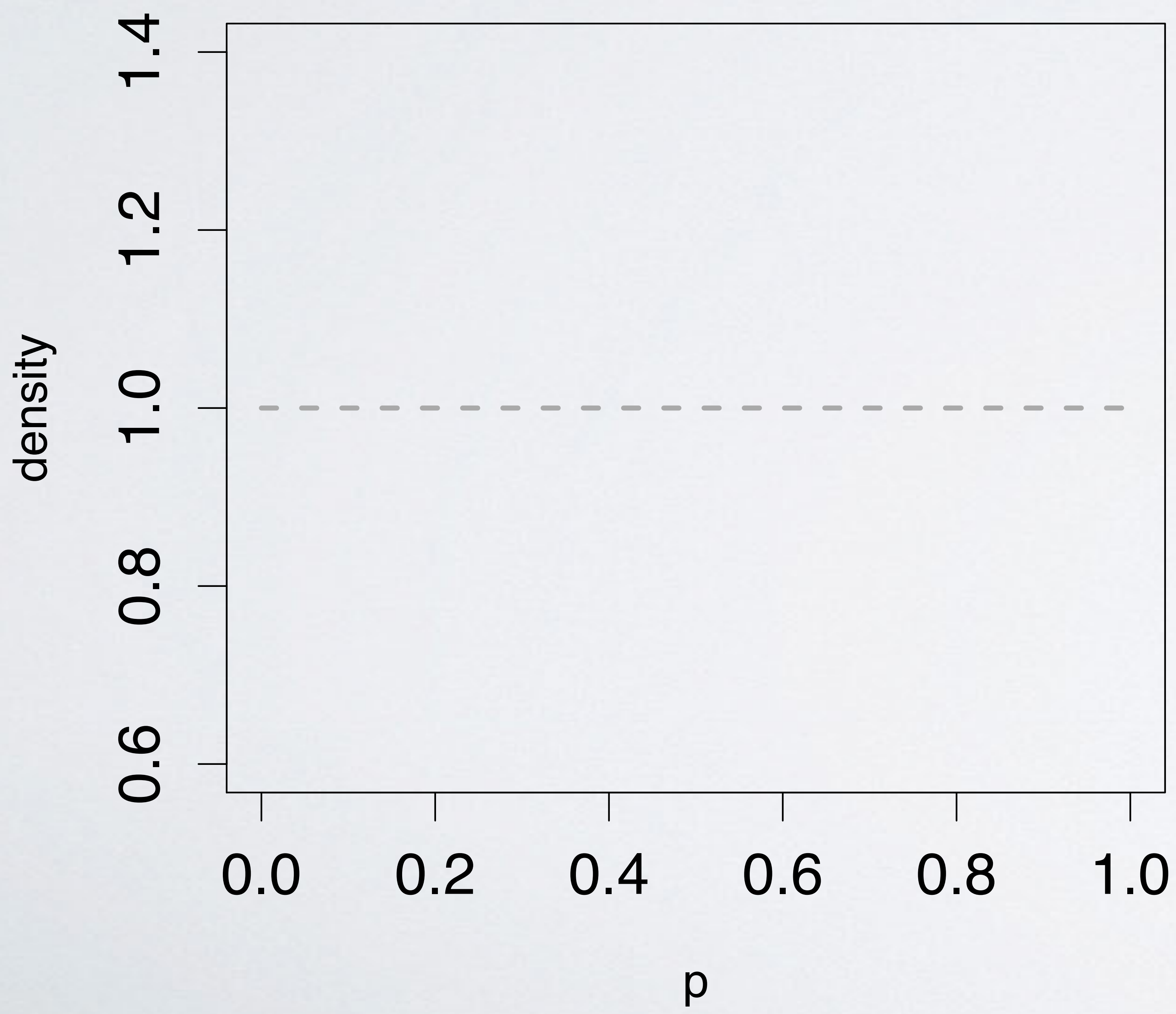
pooled prior distributions under H_1

$$p(p) \propto p^{a_m - 1}(1-p)^{b_m - 1} p^{a_f - 1}(1-p)^{b_f - 1}$$

$$= p^{a_m + a_f - 1}(1-p)^{b_m + b_f - 1}$$

$$\Rightarrow p \sim \text{Beta}(a_m + a_f, b_m + b_f)$$

$$p \sim \text{Beta}(1/2 + 1/2, 1/2 + 1/2)$$



pooled prior and posterior distributions under H_1

	male	female
yes	34	61
no+not sure	56	61
total	90	122

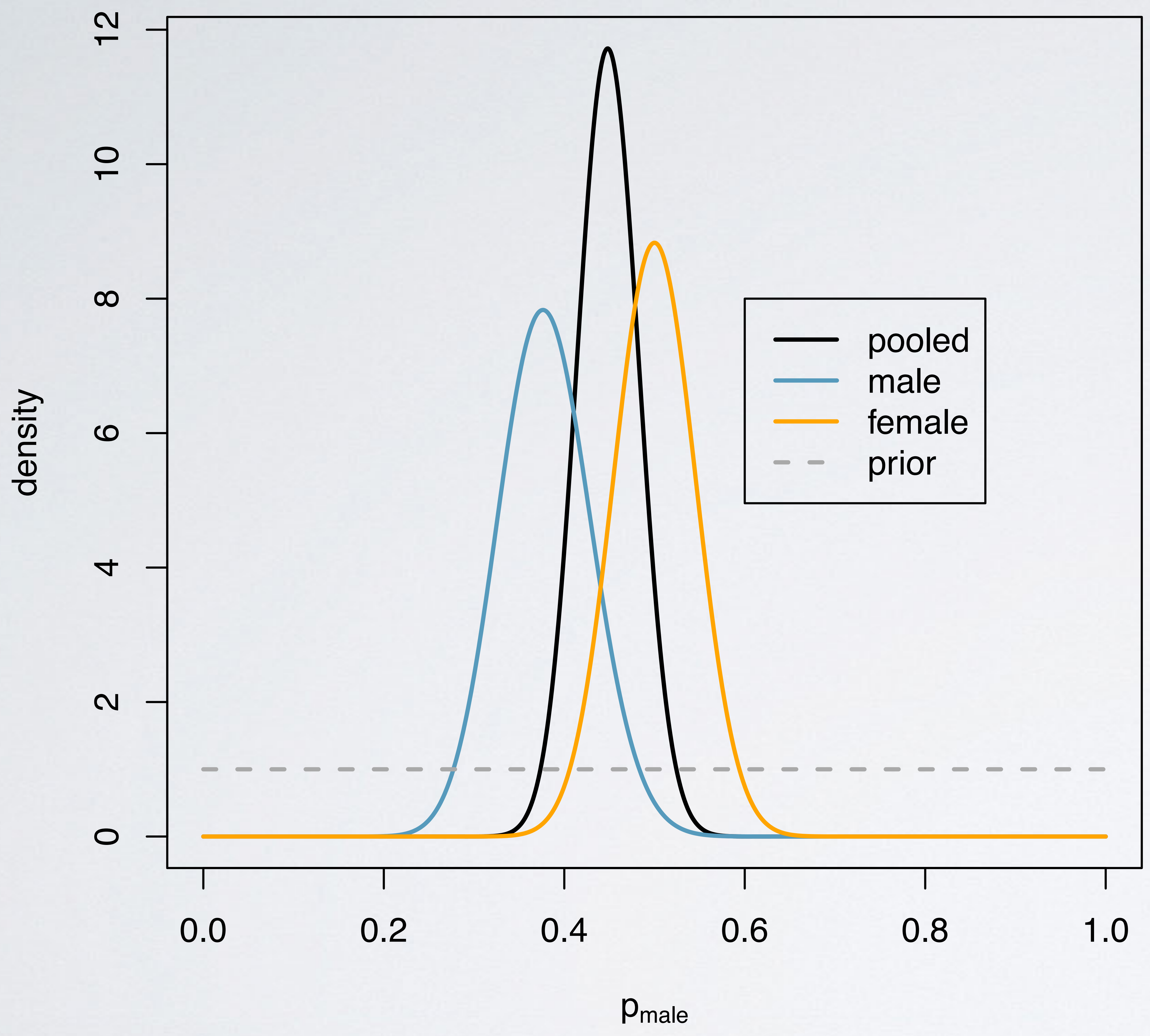
prior $p \sim \text{Beta}(1, 1)$

likelihood $R_{\text{male}} \sim \text{Bin}(90, p)$ $R_{\text{female}} \sim \text{Bin}(122, p)$

posterior $\propto p^{34}(1 - p)^{56} p^{61}(1 - p)^{61} p^{1-1}(1 - p)^{1-1}$

$p \mid \text{data} \sim \text{Beta}(95 + 1, 117 + 1)$

pooled prior and posterior distributions under H_1



summary

- inference with paired or matched normal samples
- default prior distributions under the two hypotheses
- posterior distributions for proportions under H_2 and H_1

next:

- Bayes factors and posterior probability