

Homework 2

Claire Rasmussen

1. Gelman states, "...probabilities are numerical quantities, defined on a set of 'outcomes,' that are nonnegative, additive over mutually exclusive outcomes, and sum to 1 over all possible mutually exclusive outcomes." We use probability to quantify or characterize the uncertainty surrounding some event. We can make mathematical statements about probability, such as, "The probability of observing a head in one toss of a fair coin is 0.5." Or we can make probability statement that are not mathematical, but largely based on our intuition or general understanding of a system, such as, "There is a good chance that if I go to the store now I'll get stuck in rush hour traffic."
2. The likelihood principle states that if \mathbf{x} and \mathbf{y} are two sample points such that $L(\theta | \mathbf{x}) = C(\mathbf{x}, \mathbf{y})L(\theta | \mathbf{y})$ for all θ , then conclusions drawn from \mathbf{x} and \mathbf{y} about θ should be the same. The frequentist criticism of the likelihood principle that Jordan made reference to is that this does not allow for information about the design of the experiment to enter into the problem. For example, let's say we have a two-sided coin and we're interested in $\theta = P(heads)$. We toss the coin 12 times and let X = the number of heads. Then

$$X \sim Binom(12, \theta)$$

Someone else tosses the coin until 3 tails are observed. Let Y = the number of heads prior to the 2nd tail. Then

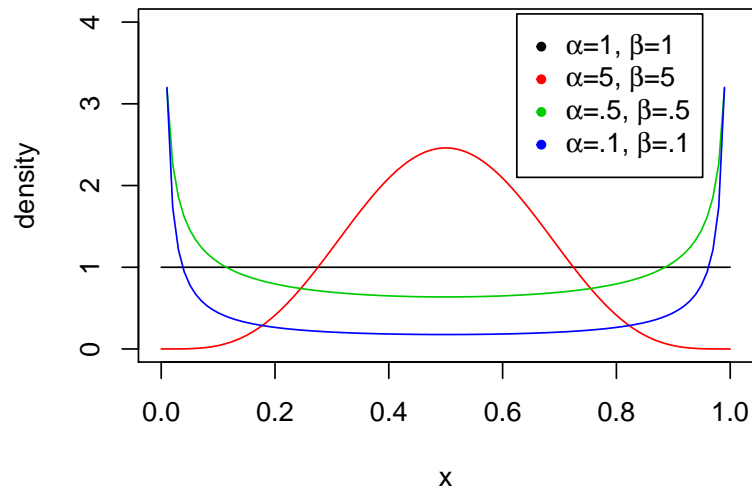
$$Y \sim NBin(3, \theta)$$

We conduct the two experiments and observe $x = 9$ and $y = 9$. If we are interested in testing

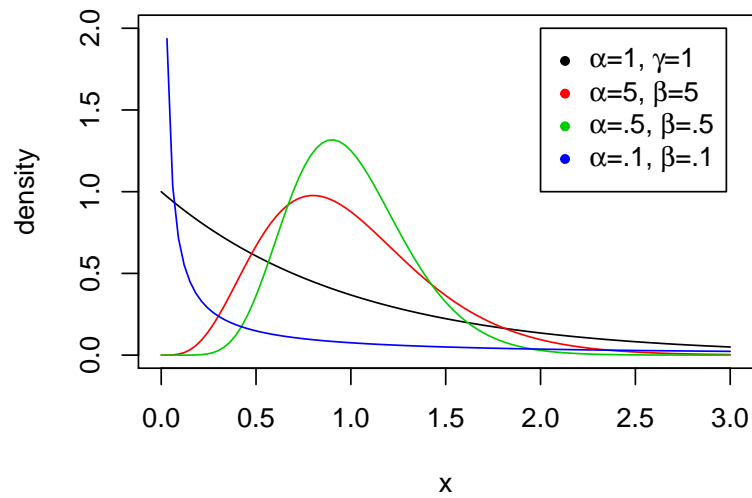
$$H_0 : \theta = 0.5 \text{ vs. } H_a : \theta > 0.5$$

we draw different conclusions if we let $\alpha = .05$ (we get p-values of 0.075 and 0.0325 in the binomial and negative binomial cases, respectively). By the likelihood principle, we should draw the same conclusion about θ , but the frequentists believe this is flawed logic since the experimental designs differ. Steve gave us this example on 1/28/15 when he first introduced the likelihood principle. One interesting point here is that he presented it as though it were an argument from the likelihoodists as to why the frequentists are wrong. I hadn't thought about from the perspective of the frequentists, that perhaps it isn't valid to expect the same data to yield the same inference if the conditions under which the data were obtained differed. However, regardless of which side one comes at this problem from, I think it really gets at why fixed level testing is a bad idea. In reality 0.075 and 0.0325 really aren't drastically different p-values.

- 3.
4. A
5. In Wilson's example, we are interested in estimating the true duckling survival rate when 100% of our sample ducklings survived. In this case, our MLE would be equal to 1, but we know intuitively the true survival rate is not one. Wilson's point was that when sample sizes are small, it is not uncommon to observe data that results in maximum likelihood estimates on the boundary of the parameter space. Bayesian methods work better here because the use of prior information (even if it is non-informative) has the effect of .
- 6.

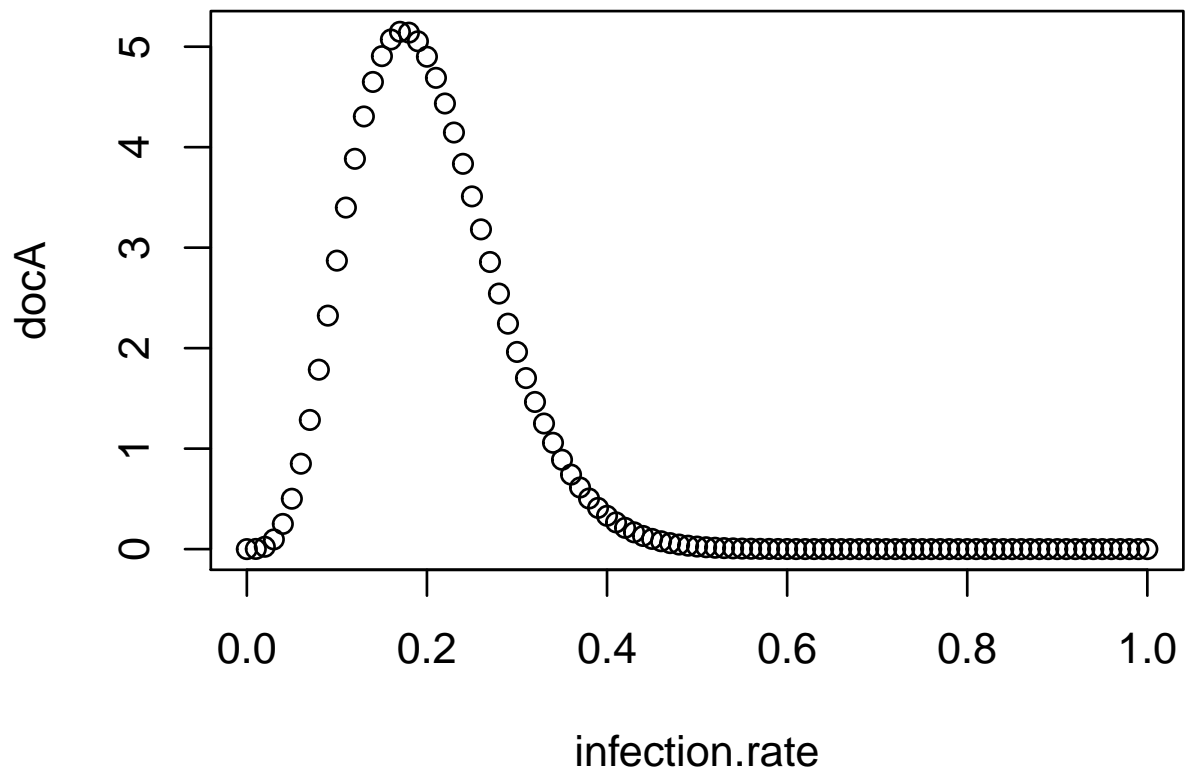


7.



$$\begin{aligned}
 8. \quad (a) \quad P(\text{Infection} \mid +\text{test}) &= \frac{P(+\text{test} \mid \text{Infection})P(\text{Infection})}{P(+\text{test})} = \frac{0.92(.05)}{P(+\text{test} \mid \text{Infection}) + P(+\text{test} \mid \text{NoInfection})} \\
 &= \frac{0.92(0.05)}{0.92 + .1} = 0.0451 \\
 P(\text{Infection} \mid +\text{test}) &= \frac{P(+\text{test} \mid \text{Infection})P(\text{Infection})}{P(+\text{test})} = \frac{0.92(.1)}{P(+\text{test} \mid \text{Infection}) + P(+\text{test} \mid \text{NoInfection})} \\
 &= \frac{0.92(0.1)}{0.92 + .1} = 0.0902
 \end{aligned}$$

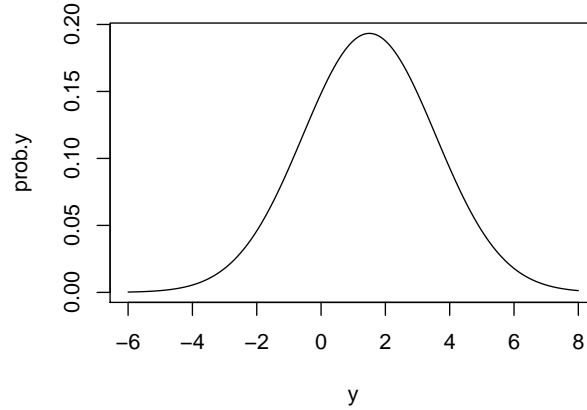
(b) ## [1] 0.99



9.

10. (a)
(b)
(c)
(d)
(e)
(f)
(g)

11. (a) For $\sigma = 2$, $f(y) = P(\theta = 1)f(y \mid \theta = 1) + P(\theta = 2)f(y \mid \theta = 2)$
 $= 0.5 * N(1, 4) + 0.5 * N(2, 4)$



(b) Note that $\pi(\theta | y) = \frac{f(y|\theta)\pi(\theta)}{f(y)} = \frac{f(y|\theta=1)P(\theta=1)+f(y|\theta=2)P(\theta=2)}{f(y)}$.

Then for $\sigma = 2$, $P(\theta = 1 | y = 1) = \frac{f(y=1|\theta=1)P(\theta=1)}{f(y=1)} = \frac{0.1995*0.5}{.1878} = 0.5312$. Intuitively this makes sense, since our observed datum was closer to $\theta = 1$ and $\theta = 2$. See appendix for the code used to compute these values.

(c)

12. $P(Identical) = 1/300$, $P(Fraternal) = 1/125$

$$P(Identical | TwinBrother) = \frac{P(Identical, TwinBrother)}{P(TwinBrother)} = \frac{P(TwinBrother|Identical)*P(Identical)}{P(TwinBrother)}$$

$$= \frac{\frac{1}{2} * \frac{1}{300}}{\frac{1}{2} * \frac{1}{300} + \frac{1}{2} * \frac{1}{125}} = \frac{5}{11}$$

13. (a) Since person A will observe the die roll, he or she will have complete information and will be able to say either the roll result in a 6 or not with probability equal to one. That is, $P_A(E) = 1$ or $P_A(E) = 0$. Person B will have only the information that the die is fair, and thus should state that the probability of observing a 6 is equal to $1/6$. That is, $P_B(E) = 1/6$.

(b) Since person A is ignorant of soccer, he or she has little or no prior information, a logic conclusion is that $P_A(E) = 0.5$. Since person B is a knowledgeable sports fan, he or she may have a more informed opinion. Thus, $P_B(E)$ would likely be either less than 0.5 or greater than 0.5.

14.

15. (a)

(b)

16. (a)

(b)

(c)

17.