Stat 532 Assignment 2

Kenny Flagg

September 11, 2015

1. - Gelman et al defines probabilities as numbers associated with outcomes which are "nonnegative, additive over mutually exclusive outcomes, and sum to 1 over all possible mutually exclusive outcomes" (BDA3, §1.5, p. 11).
   -

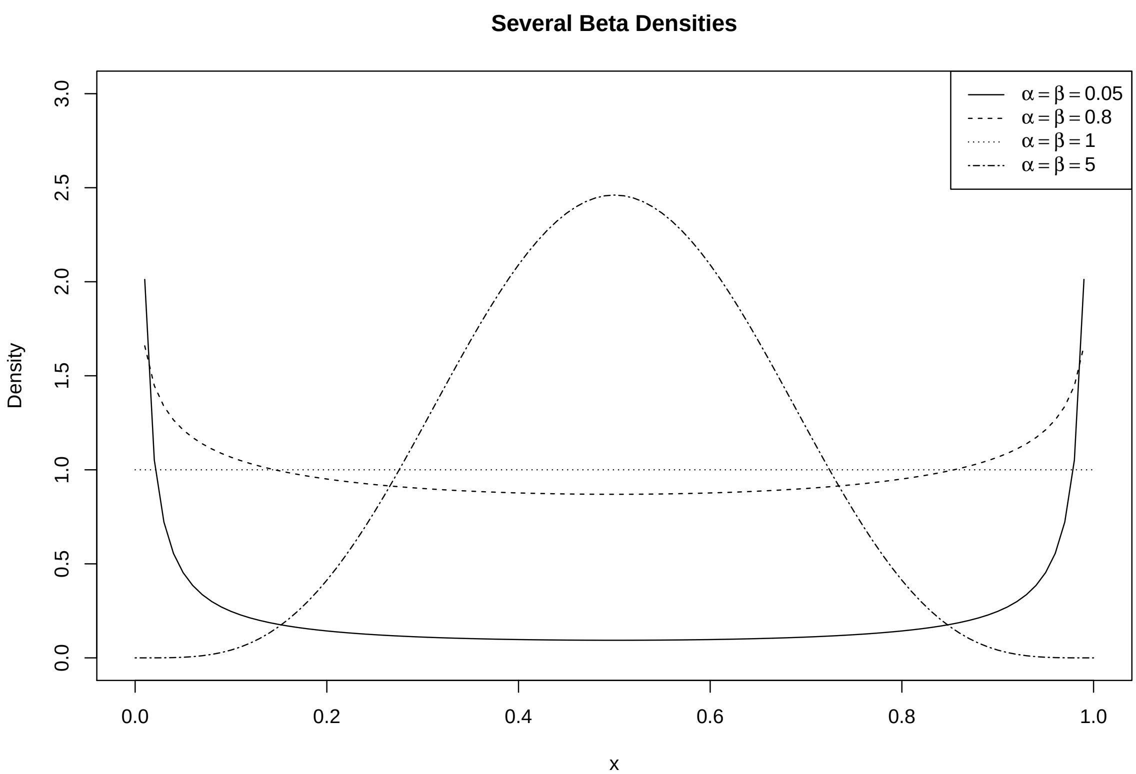
2.

3.

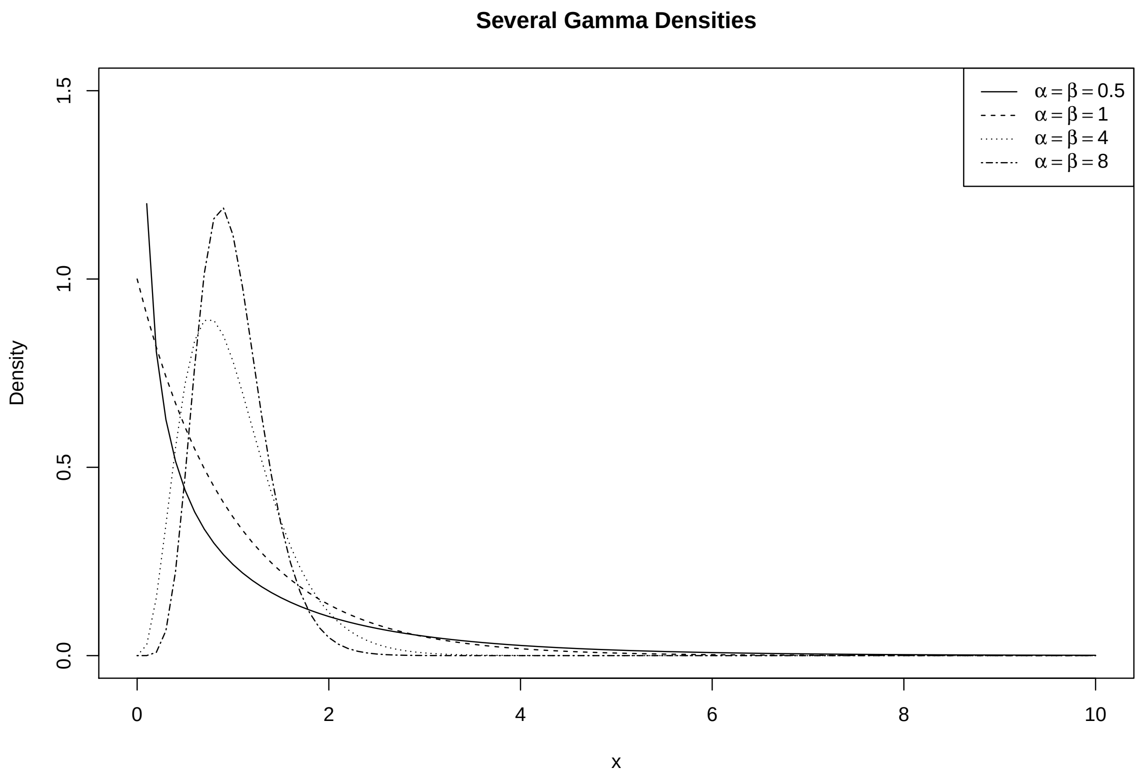
4.

5.

6. Below are several beta density curves. When $\alpha \neq \beta$, increasing α increases left-skew and increasing β increases right-skew.



7. I had not realized that R's default parameterization for the Gamma distribution uses a rate parameter instead of a scale parameter. Fortunately this agrees with the parameterization in BDA3 (where the rate parameter is called an inverse scale parameter). It makes me wonder how many times I've incorrectly used a parameterization that I did not intend to use. I've plotted several gamma density curves below.



The inverse scale parameterization of the Gamma distribution forms a conjugate prior when used to model Poisson rates and Normal precisions (inverse variances), so it is often a default choice in these situations.

8. (a) The posterior probability of infection is

$$\begin{aligned} Pr(\text{infection}|\text{test }+) &= \frac{Pr(\text{test }+|\text{infection})Pr(\text{infection})}{Pr(\text{test }+|\text{infection})Pr(\text{infection}) + Pr(\text{test }+|\text{no infection})Pr(\text{no infection})} \\ &= \frac{0.92Pr(\text{infection})}{0.92Pr(\text{infection}) + 0.08Pr(\text{no infection})} \end{aligned}$$

If the first doctor is correct, this works out to

$$Pr(\text{infection}|\text{test }+) = \frac{(0.92)(0.05)}{(0.92)(0.05) + (0.08)(0.95)} = 0.377$$

so there is a 37.7% chance that the patient is infected.
If the second doctor is right,

$$Pr(\text{infection}|\text{test }+) = \frac{(0.92)(0.1)}{(0.92)(0.1) + (0.08)(0.9)} = 0.561$$

and the patient has a 56.1% probability of being infected.

The second doctor believed the infection rate to be twice as large as what the first doctor believed. As a result, the second doctor would consider the patient as slightly more likely to be infected than not, while the first doctor would conclude that the patient is most likely not infected.

I see this as illustrating the Bayesian interpretation of probability as describing the state a knowledge about something unobserved. If we could observe the infection status without measurement error (a perfect test) then we would definitively know if the patient is infected or not. Since the patient's status is unknown, the doctors must incorporate their previous knowledge about the infection and they come to different conclusions despite both observing the same test result. If we think that objectivity means that doctors should follow the likelihood principle and reach the same conclusions, then this could be support an argument for the use of uninformative priors when experts have disagreement.

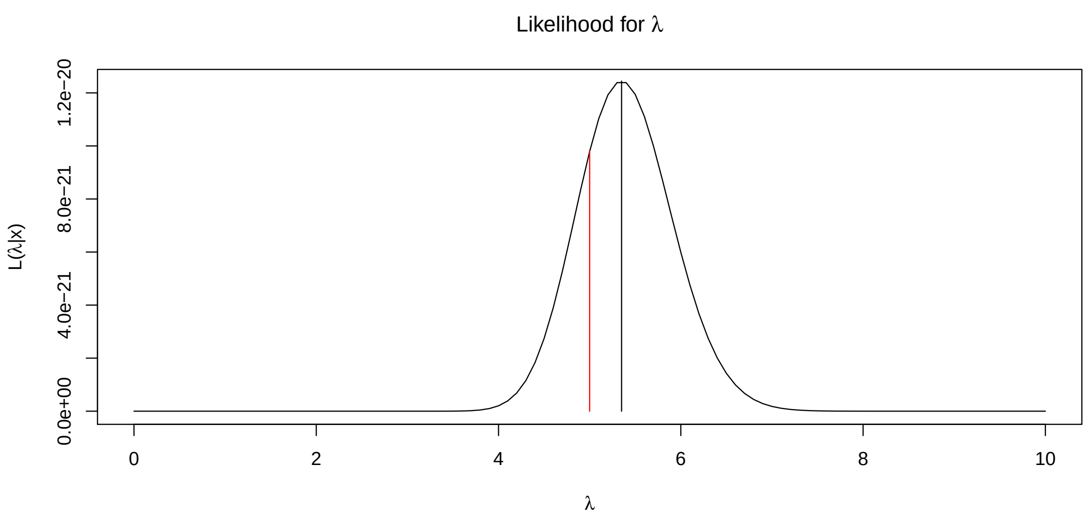
(b) $Beta(2, 20)$, $Beta(3.8, 27.2)$

9.

10. (a) 20 observations from $Poisson(\lambda = 5)$

Sample Mean: $\bar{x} = 5.35$, Sample Variance: $s^2 = 6.45$

Maximum Likelihood Estimate: $\bar{x} = 5.35$, SE: $\dfrac{s}{\sqrt{n}} = 0.56789083$



(b) 100 observations from $Poisson(\lambda = 5)$

(c) 15 observations from $N(\mu = 10, \sigma^2 = 5)$

(d) 5 observations from $N(\mu = 10, \sigma^2 = 5)$

(e) 1 observations from $Binomial(m = 100, p = 0.2)$

(f) 30 observations from $Binomial(m = 100, p = 0.2)$

(g)

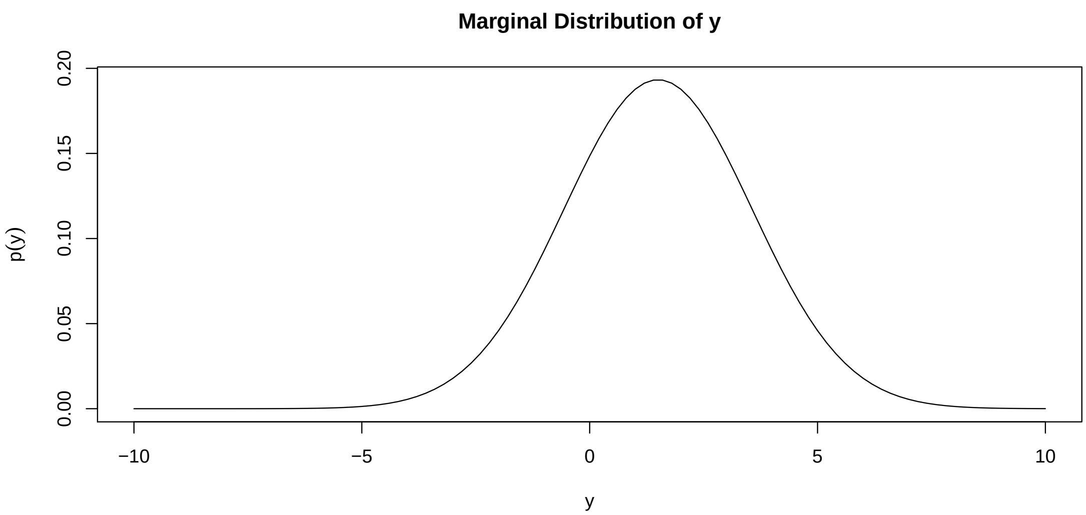
11. Problem 1.1

(a) We have $y \sim N(\theta, \sigma^2)$ with $Pr(\theta = 1) = Pr(\theta = 2) = \frac{1}{2}$. If $\sigma = 2$,

$$p(y, \theta) = \left(\frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{(y - \theta^2)}{8} \right) \right) \left(\frac{1}{2} \right) = \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{(y - \theta)^2}{8} \right); \theta = 1, 2$$

so the marginal density of y is

$$p(y) = \sum_{\theta=1}^{2} p(y,\theta) = \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{(y-1)^2}{8}\right) + \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{(y-2)^2}{8}\right)$$
$$= \frac{1}{4\sqrt{2\pi}} \left(e^{-\frac{(y-1)^2}{8}} - e^{\frac{(y-2)^2}{8}}\right).$$



(b) The posterior probability is

$$Pr(\theta = 1 | y = 1) = \frac{p(y=1, \theta=1)}{p(y=1)} = \frac{\frac{1}{4\sqrt{2\pi}} e^{-\frac{(1-1)^2}{8}}}{\frac{1}{4\sqrt{2\pi}}\left(e^{-\frac{(1-1)^2}{8}} + e^{-\frac{(1-2)^2}{8}}\right)}$$
$$= \frac{1}{1+e^{-\frac{1}{8}}} \approx 0.5312.$$

(c) For any $\sigma > 0$, the posterior distribution is

$$p(\theta|y) = \frac{p(y,\theta)}{p(y)} = \frac{\frac{1}{2\sigma\sqrt{2\pi}} e^{-\frac{(y-\theta)^2}{2\sigma^2}}}{\frac{1}{2\sigma\sqrt{2\pi}}\left(e^{-\frac{(y-1)^2}{2\sigma^2}} + e^{-\frac{(y-2)^2}{2\sigma^2}}\right)}$$
$$= \frac{e^{-\frac{(y-\theta)^2}{2\sigma^2}}}{e^{-\frac{(y-1)^2}{2\sigma^2}} + e^{-\frac{(y-2)^2}{2\sigma^2}}}; \theta = 1, 2$$

so as θ increases, each exponent approaches 0 and thus $p(\theta|y) \to \frac{1}{2}$ for $\theta = 1, 2$.

12.

13.

14.

15. (a)

(b)

16. (a)

    (b)

    (c)

17. Steps of Statistical Inference:

    (a) Ask a question