Solutions to Homework 3

Shashank Sule

9/10/2019

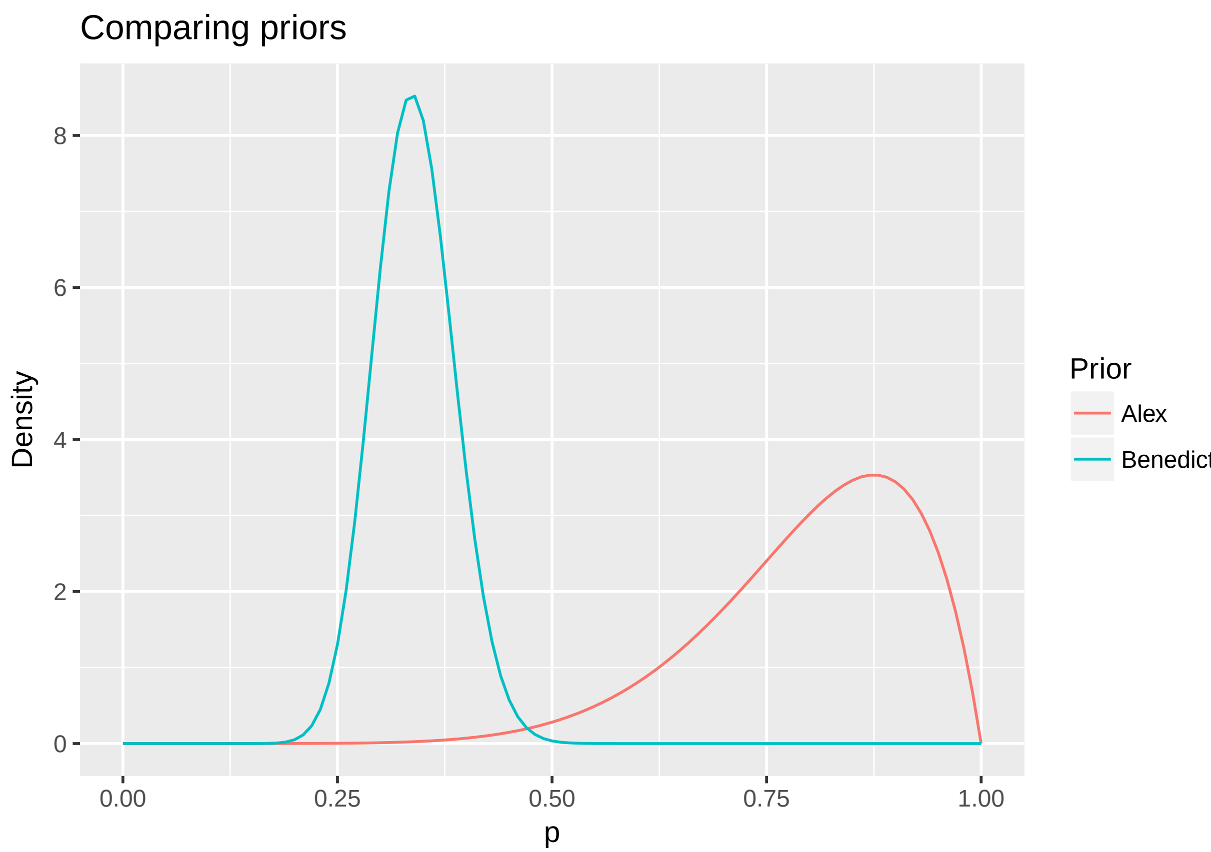
Problem 1

a) In a Beta distribution parametrized as Beta(a, b), a represents the number of successes and b the number of failures. So if there were 8 successes and 2 failures out of a data of 10 then Alex has a prior Beta(a, b).

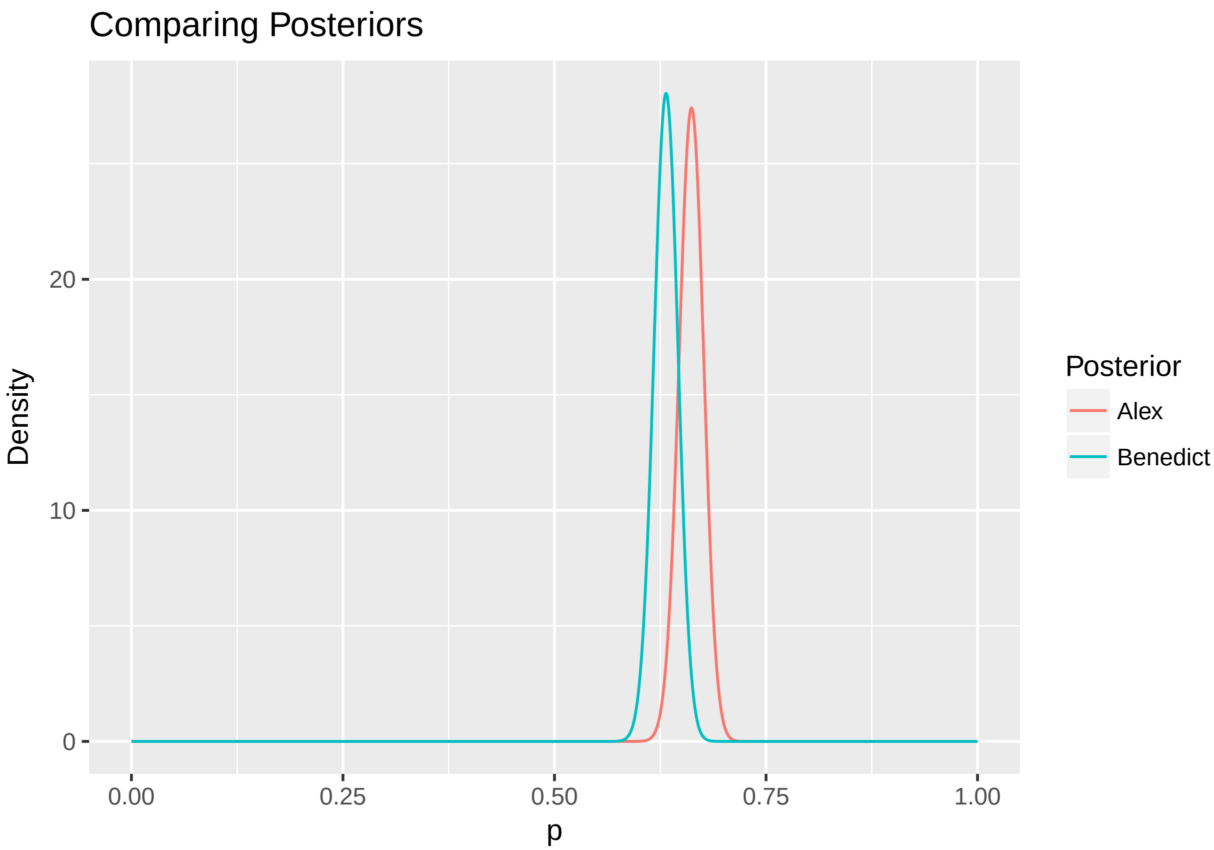
b) We compute the parameters using the `beta.select` function:

```
## [1] 34.95 67.98
```

c) Note that $y = 692$, $n = 1048$ and the posterior distribution is Beta$(a + y, b + n - y)$. So Alex has posterior Beta$(700, 358)$ and Benedict has posterior Beta$(726.95, 423.98)$.



The prior:



Credible intervals for Alex:

```
## [1] 0.6328457 0.6898269
```

Credible intervals for Benedict:

```
## [1] 0.6035532 0.6592528
```

d) To conduct a prior predictive check, we sample values from each prior and then sample a value from the likelihood function. Then we look for clustering of the sampled data around the actual data. I'll do a simulation for 1000 trials.

Alex's probability values:

```
c(sum(alex >= 692)/S, 1 - sum(alex >= 692)/S)
```

```
## [1] 0.865 0.135
```

Benedict's probability values:

```
c(sum(benedict >= 692)/S, 1 - sum(benedict >= 692)/S)
```

```
## [1] 0 1
```

Thus, neither have a good prior distribution for prediction but Benedict's is worse than Alex's because Benedict's probabiities are far more extreme (all of Benedict's data lies on one side of the actual observed value).

Problem 2

To get the odds data, I'll use Monte Carlo sampling and then transform the data. The summarized data is as follows:

```
##              Mean   Median     2.5%    97.5%
## Alex     1.958228 1.956094 1.732098 2.204190
## Benedict 1.719126 1.716058 1.528348 1.935351
```

On average, Alex's odds of kids having a TV in their room are slightly igher than those of Benedict.

Problem 3

```
##      S        5%       95%
## 1   10 0.3703484 0.7249998
## 2  100 0.4514916 0.7401867
## 3  500 0.4152143 0.7546605
## 4 1000 0.4289477 0.7400327
## 5 5000 0.4288951 0.7442373
```

I computed $S = \{10, 100, 500, 1000, 5000\}$ random samples from Beta$(15.06, 10.56)$ using the `rbeta` function and then computed the quantiles of the data using `quantile`. As the number of samples gets larger, the interval gets closer to the actual credible interval.