Lab 4: Bayesian Statistics in R - STA 360/602

Rebecca C. Steorts

1 Agenda

Do a teacher's expectations influence student achievement? In a famous study, Rosenthal and Jacobson (1968) performed an experiment in a California elementary school to try to answer this question. At the beginning of the year, all students were given an IQ test. For each class, the researchers randomly selected around 20% of the students, and told the teacher that these students were "spurters" that could be expected to perform particularly well that year. (This was not based on the test—the spurters were randomly chosen.) At the end of the year, all students were given another IQ test. The change in IQ score for the first-grade students was:[1]

spurters (S)
$x = (18, 40, 15, 17, 20, 44, 38)$

controls (C)
$y = (-4, 0, -19, 24, 19, 10, 5, 10, 29, 13, -9, -8, 20, -1, 12, 21, -7,$
$14, 13, 20, 11, 16, 15, 27, 23, 36, -33, 34, 13, 11, -19, 21, 6, 25, 30,$
$22, -28, 15, 26, -1, -2, 43, 23, 22, 25, 16, 10, 29)$

1. Plot histograms for the change in IQ score for the two groups. Report your findings.

2. How strongly does this data support the hypothesis that the teachers? expectations caused the spurters to perform better than their classmates? IQ tests are purposefully calibrated to make the scores normally distributed, so it makes sense to use a normal model here:

$$\text{spurters: } X_1, \ldots, X_{n_S} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_S, \lambda_S^{-1})$$
$$\text{controls: } Y_1, \ldots, Y_{n_C} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_C, \lambda_C^{-1}).$$

We are interested in the difference between the means—in particular, is $\mu_S > \mu_C$? We don't know the standard deviations $\sigma_S = \lambda_S^{-1/2}$ and $\sigma_C = \lambda_C^{-1/2}$, and the sample seems too small to estimate them very well.

---

[1] The original data is not available. This data is from the `ex1321` dataset of the `R` package `Sleuth3`, which was constructed to match the summary statistics and conclusions of the original study.

it is easy using a Bayesian approach: we just need to compute the posterior probability that $\mu_S > \mu_C$:

$$\mathbb{P}(\boldsymbol{\mu}_S > \boldsymbol{\mu}_C \mid x_{1:n_S}, y_{1:n_C}).$$

Let's use independent NormalGamma priors:

$$\text{spurters: } (\boldsymbol{\mu}_S, \boldsymbol{\lambda}_S) \sim \text{NormalGamma}(m, c, a, b)$$
$$\text{controls: } (\boldsymbol{\mu}_C, \boldsymbol{\lambda}_C) \sim \text{NormalGamma}(m, c, a, b)$$

with the following hyperparameter settings, based on subjective prior knowledge:

- $m = 0$ (Don't know whether students will improve or not, on average.)
- $c = 1$ (Unsure about how big the mean change will be—prior certainty in our choice of m assessed to be equivalent to one datapoint.)
- $a = 1/2$ (Unsure about how big the standard deviation of the changes will be.)
- $b = 10^2 a$ (Standard deviation of the changes expected to be around $10 = \sqrt{b/a} = \mathbb{E}(\lambda)^{-1/2}$.)

The updated posterior parameters are

for spurters:
$$M = \frac{1 \cdot 0 + 7 \cdot 27.43}{1 + 7} = 24.0$$
$$C = 1 + 7 = 8$$
$$A = 1/2 + 7/2 = 4$$
$$B = 100/2 + \tfrac{1}{2} \cdot 7 \cdot 11.66^2 + \tfrac{1}{2}\frac{1 \cdot 7}{1 + 7}(27.43 - 0)^2 = 855.0$$

for controls:
$$M = \frac{1 \cdot 0 + 48 \cdot 12.04}{1 + 48} = 11.8$$
$$C = 1 + 48 = 49$$
$$A = 1/2 + 48/2 = 24.5$$
$$B = 100/2 + \tfrac{1}{2} \cdot 48 \cdot 16.10^2 + \tfrac{1}{2}\frac{1 \cdot 48}{1 + 48}(12.04 - 0)^2 = 6344.0$$

which implies that

$$\boldsymbol{\mu}_S, \boldsymbol{\lambda}_S \mid x_{1:n_S} \sim \text{NormalGamma}(24.0, 8, 4, 855.0)$$
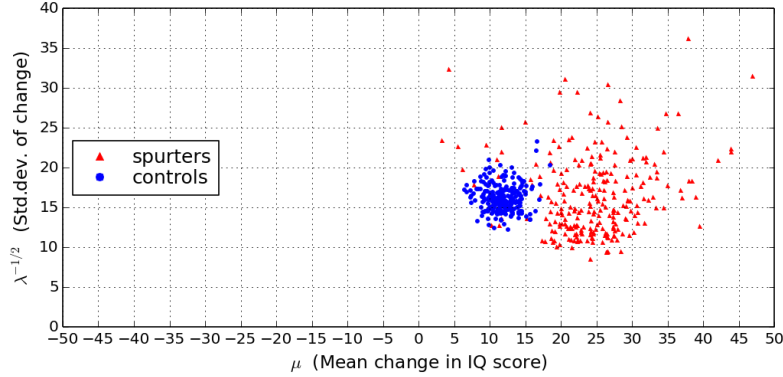$$\boldsymbol{\mu}_C, \boldsymbol{\lambda}_C \mid y_{1:n_C} \sim \text{NormalGamma}(11.8, 49, 24.5, 6344.0).$$

Figure 1: Samples of (μ, σ) from the posteriors for the two groups.

3. Based on the calculations from the previous task, provide a scatterplot of samples from the posterior distributions for the two groups. Your scatterplot should look similar to Figure 1. What are your conclusions?

4. Now, we can answer our original question: "What is the posterior probability that $\mu_S > \mu_C$?"

   The easiest way to do this is to take a bunch of samples from each of the posteriors, and see what fraction of times we have $\mu_S > \mu_C$. This is an example of a Monte Carlo approximation (much more to come on this in the future).

   To do this, we draw $N = 10^6$ samples from each posterior:

   $$(\mu_S^{(1)}, \lambda_S^{(1)}), \ldots, (\mu_S^{(N)}, \lambda_S^{(N)}) \sim \text{NormalGamma}(24.0, 8, 4, 855.0)$$
   $$(\mu_C^{(1)}, \lambda_C^{(1)}), \ldots, (\mu_C^{(N)}, \lambda_C^{(N)}) \sim \text{NormalGamma}(11.8, 49, 24.5, 6344.0)$$

   and obtain the approximation

   $$\mathbb{P}(\boldsymbol{\mu}_S > \boldsymbol{\mu}_C \mid x_{1:n_S}, y_{1:n_C}) \approx \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\left(\mu_S^{(i)} > \mu_C^{(i)}\right) = \; ??$$

   Interpret the posterior probability that you compute above.

5. Let's return back to the prior assumptions. There are a few ways that you can check that the prior conforms with our prior beliefs. Let's go back and checks these. Draw some samples from the prior and look at them—this is probably the best general strategy. See Figure 2. It's also a good idea
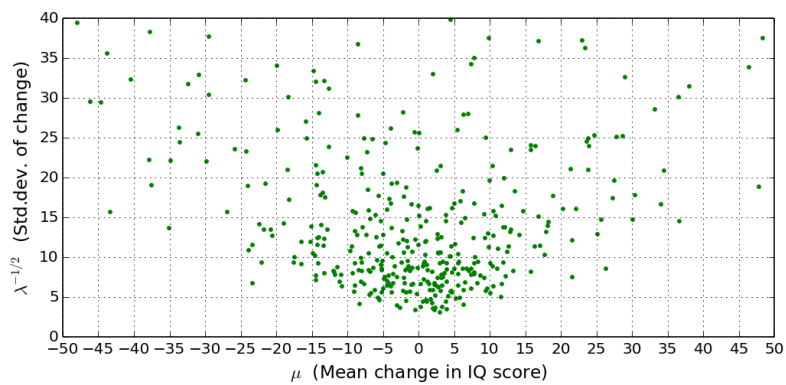
3

Figure 2: Samples of (μ, σ) from the prior.

to look at sample hypothetical datasets $X_{1:n}$ drawn using these sampled parameter values.

Please replicate a plot similar to Figure 2 and report your findings.