

Assignment 2A: Introduction to Bayesian Statistics (Points 72)

Q1 (6 points). A city is considering building a new museum. The local paper wishes to determine the level of support for this project, and is going to conduct a poll of city residents. Out of the sample of 120 people, 74 support the city building the museum.

- What is the distribution of y , the number who support building the museum?
- Use a uniform prior for π , the proportion of the target audience that support the museum. What is the posterior distribution of π ?

Ans:

Q2 (11 points). You are going to take a random sample of voters in a city in order to estimate the proportion π who support stopping the fluoridation of the municipal water supply. Before you analyze the data, you need a prior distribution for π . You decide that your prior mean is 0.4, and your prior standard deviation is 0.1.

- Determine the beta (a , b) prior that matches her prior belief.
- What is the equivalent sample size of her prior?
- Out of the 100 city voters polled, $y = 21$ support the removal of fluoridation from the municipal water supply. Determine your posterior distribution.

Ans:

Q3 (25 points). In a research program on human health risk from recreational contact with water contaminated with pathogenic microbiological material, the National Institute of Water and Atmosphere (NIWA) instituted a study to determine the quality of New Zealand stream water at a variety of catchment types. This study is documented in McBride et al. (2002) where $n = 145$ one-liter water samples from sites identified as having a heavy environmental impact from birds (seagulls) and waterfowl. Out of these samples, $y = 12$ samples contained Giardia cysts.

- What is the distribution of y , the number of samples containing Giardia cysts?
- Let π be true probability that a one-liter water sample from this type of site contains Giardia cysts. Use a *beta* (1,4) prior for π . Find the posterior distribution of π given y .
- Summarize the posterior distribution by its first two moments.
- Find the *normal* approximation to the posterior distribution $g(\pi|y)$.
- Compute a 95% credible interval for π using the normal approximation found in part (d).

Ans:

Q4 (15 points). In the study of water quality in New Zealand streams documented in McBride et al. (2002) a high level of Campylobacter was defined as a level greater than 100 per 100 ml of stream water. $n = 145$ samples were taken from streams having a high environmental impact from birds. Out of these $y = 9$ had a high Campylobacter level. Let π be the true probability that a sample of water from this type of stream has a high Campylobacter level.

- Find the frequentist estimator for π
- Use a beta (1, 10) prior for π . Calculate the posterior distribution $g(\pi|y)$.
- Find the posterior mean and variance. What is the Bayesian estimator for π ?
- Find a 95% credible interval for π .
- Test the hypothesis

$$H_0: \pi = .10 \text{ versus } H_1: \pi \neq .10$$

at the 5% level of significance.

Ans:

Q5 (15 points). In the same study of water quality, $n = 87$ samples were taken from streams having a high environmental impact from birds. Out of these $y = 8$ had a high Campylobacter level. Let π be the true probability that a sample of water from this type of stream has a high Campylobacter level.

- Find the frequentist estimator for π
- Use a beta (1, 10) prior for π . Calculate the posterior distribution $g(\pi|y)$.
- Find the posterior mean and variance. What is the Bayesian estimator for π ?
- Test the hypothesis

$$H_0: \pi \geq .10 \text{ versus } H_1: \pi < .10$$

at the 5% level of significance.

Ans:

Updated on 30th June 2014