STAT 457 Homework 05

Martha Eichlersmith

2019-11-16

Problem 1

Consider two urns each containing an unknown mixture of blue and white marbles. A random sample of size 18 (with replacement) is drawn from urn #1 and a random sample of size 6 (with replacement) is drawn from urn #2. Of the 18 selected marbles from urn #1, 14 are blue. The corresponding number of blue marbles from urn #2 is 2.

$$
\begin{aligned}
L(\pi|Y) &\propto \pi^y(1-\pi)^{n-y} = \pi^{14}(1-\pi)^4 \sim \text{Binomial}, n_\pi = 18,\ y_\pi = 14 \\
&\implies p(\pi \mid Y) \sim \text{Beta}(y+\alpha_0, n-y+\beta_0) \quad \text{where} \quad p(\pi) \sim \text{Beta}(\alpha_0, \beta_0)
\end{aligned}
$$

$$
\begin{aligned}
L(\psi|Y) &\propto \psi^y(1-\psi)^{n-y} = \psi^{2}(1-\psi)^4 \sim \text{Binomial}, n_\pi = 6,\ y_\pi = 2 \\
&\implies p(\psi \mid Y) \sim \text{Beta}(y+\alpha_0, n-y+\beta_0) \quad \text{where} \quad p(\psi) \sim \text{Beta}(\alpha_0, \beta_0)
\end{aligned}
$$

Problem 1a

Let π denote the proportion of blue marbles in urn #1 and let ψ denote the corresponding proportion in urn #2. Under the (i) Haldane, (ii) flat and (iii) non-informative priors, compute $p\left(\ln\left[\frac{\pi}{1-\pi}\right] > \ln\left[\frac{\psi}{1-\psi}\right] \mid \text{data}\right)$ using the normal approximation.

$$
p\left(\ln\left[\frac{\pi}{1-\pi}\right] > \ln\left[\frac{\psi}{1-\psi}\right] \mid \text{data}\right) = p\left(\ln\left[\frac{\pi}{1-\pi}\right] - \ln\left[\frac{\psi}{1-\psi}\right] > 0 \mid \text{data}\right)
$$

Normal Approx: Probability difference in logodds is greater than 0

Prior	Haldane - Beta(0, 0)	Flat - Beta(1, 1)	Non-informative - Beta(.5, .5)
p-value	0.96533	0.97788	0.97285

Problem 1b

Repeat (1a) by drawing deviates from the appropriate beta distributions. Quantify the Monte Carlo error in your value.

Haldane Prior - Beta(0, 0)

Iterations	10000	1e+05	1e+06
p-value	0.90144	0.90114	0.90140
Standard Error	0.00901	0.00284	0.00090

Flat Prior - Beta(1, 1)

Iterations	10000	1e+05	1e+06
p-value	0.90339	0.90072	0.90116
Standard Error	0.00883	0.00284	0.00090

Non-informative Prior - Beta(.5, .5)

Iterations	10000	1e+05	1e+06
p-value	0.89969	0.90234	0.90110

Non-informative Prior - Beta(.5, .5)

Standard Error	0.00887	0.00284	0.00090

‘

Problem 1c

Compare your results in (1a) and (1b) to the p-value obtained via Fisher's exact test.

Problem 1d

Add delinquency problem

Problem 2

Suppose a sample of size n is drawn at random and with replacement from some population. For large n the sample proportion ($\hat{p}$) is normally distributed with mean p and variance $\frac{p(1-p)}{n}$. Find the asymptotic distribution of $2\sin^{-1}\sqrt{\hat{p}}$ using the delta method.

Problem 3

Let $x_1, \cdots, x_n$ be an iid sample from $\mathcal{N}(\theta, 1)$ and let $y_1, \cdots, y_n$ be an independent iid sample from $\mathcal{N}(\phi, 1)$. Derive the distribution of $\overline{x}/\overline{y}$ (where $\overline{y} \neq 0$) via the delta method.

Problem 4

197 animals are distributed into four categories: $Y = (y_1, y_2, y_3, y_4)$ according to the genetic linkage model $\left(\frac{2+\theta}{4}, \frac{1-\theta}{4}, \frac{1-\theta}{4}, \frac{\theta}{4}\right)$. In HW#4 you derived the likelihood for the data $Y = (125, 18, 20, 34)$ and you derived the likelihood for the data $Y = (14, 0, 1, 15)$. In that homework, you also used Newton-Raphson algorithm to obtain the MLE ($\hat{\theta}$) of θ and the standard error of $\hat{\theta}$.

Problem 4a

Plot the normalized likelihood and the associated normal approximation in the same figure for the data $Y = (125, 18, 20, 34)$. Discuss the adequacy of the normal approximation.

Problem 4b

Repeat (4a) for $Y = (14, 0, 1, 5)$

Problem 5

Use Laplace's method (second order) to compute the posterior mean (under a flat prior) for the genetic linkage model for both data sets.