STA623 - Bayesian Data Analysis - Practical 3

Marc Henrion

2024-10-30

Practical 3

Notation

- X, Y, Z - random variables
- x, y, z - measured / observed values
- $\bar{X}$, $\bar{Y}, \bar{Z}$ - sample mean estimators for X, Y, Z
- $\bar{x}$, $\bar{y}, \bar{z}$ - sample mean estimates of X, Y, Z
- $\hat{T}$, $\hat{t}$ - given a statistic T, estimator and estimate of T
- $P(A)$ - probability of an event A occuring
- $f_X(.)$, $f_Y(.), f_Z(.)$ - probability mass / density functions of X, Y, Z; sometimes $p_X(.)$ etc. rather than $f_X(.)$
- p(.) - used as a shorthand notation for pmfs / pdfs if the use of this is unambiguous (i.e. it is clear which is the random variable)
- $X \sim F$ - X distributed according to distribution function F
- $E[X]$, $E[Y]$, $E[Z]$, $E[T]$ - the expectation of X, Y, Z, T respectively

Exercise 1

Show that the Bayes estimator $\hat{\theta}_B$ for the quadratic loss function $\mathcal{C}(\theta - \hat{\theta}) = (\theta - \hat{\theta})^2$ is given by the posterior mean. In other words, show that:

$$E[\theta|y] = \arg\min_{\hat{\theta}} \int_y \int_\Theta \mathcal{C}(\theta - \hat{\theta}) p(\theta, y) d\theta dy$$

Exercise 2

Suppose $\pi \sim \text{Beta}(2, 3)$ and $Y_1, \dots, Y_n \sim_{\text{iid}} Bernoulli(\pi)$. Further suppose we observe data $y_1, \dots, y_n$ with $n = 25, k = \sum_i y_i = 16$.

Find the following:

- posterior distribution $p(\pi|k)$ and plot it, comparing it to the prior distribution
- posterior predictive distribution $p(\tilde{y}|y_1, \dots, y_n)$
- a Bayesian point estimate $\hat{\pi}$
- the 95% quantile-based Bayesian confidence interval for π
- the 95% HPD interval

Further, compute:

- $P(\pi > 0.5|k)$
- For the following 2 hypotheses: $H_1 : \pi \in [0.3, 0.5], H_2 : \pi \in [0.5, 0.7]$, compute the prior and posterior odds and calculate the Bayes factor.

Exercise 3

Suppose $\lambda \sim \text{Gamma}(5,2)$ and $Y_1, \ldots, Y_n \sim_{\text{iid}} Poisson(\lambda)$. Further suppose we observe data $y_1, \ldots, y_n$ with $n = 18, k = \sum_i y_i = 40$.

Find the following:

- posterior distribution $p(\lambda | y_1, \ldots, y_n)$ and plot it, comparing it to the prior distribution
- posterior predictive distribution $p(\tilde{y} | y_1, \ldots, y_n)$
- the 95% quantile-based Bayesian confidence interval for λ
- the 95% HPD interval

Further, compute:

- $P(\lambda \leq 1 | y_1, \ldots, y_n)$
- For the following 2 hypotheses: $H_1 : \lambda \in [0.75, 1.25], H_2 : \lambda \in [1.75, 2.25]$, compute the prior and posterior odds and calculate the Bayes factor.