Analysis of Posterior Distribution and Conjugate Priors

Ben Russell

March 27, 2024

1)

Given a likelihood function $p_X(x|\theta) = h(x)\exp(\theta t(x) - a(\theta))$, and a prior distribution $p_\theta(\theta) \propto \exp(\tau\theta - \gamma \cdot a(\theta))$, our goal is to find the posterior distribution of θ and to show that p_θ is a conjugate prior for the parameter θ.

The likelihood of observing $X_1 = x_1, \ldots, X_N = x_N$ given θ is

$$p_X(X_1 = x_1, \ldots, X_N = x_N|\theta) = \prod_{i=1}^{N} h(x_i)\exp(\theta t(x_i) - Na(\theta))$$

This leads us to the posterior distribution of θ,

$$p_\theta(\theta|X_1 = x_1, \ldots, X_N = x_N) \propto \exp\left(\theta\left(\sum_{i=1}^{N} t(x_i) + \tau\right) - (N+\gamma)a(\theta)\right)$$

The form of this posterior distribution shows that $p_\theta(\theta)$ is indeed a conjugate prior for $p_X(x|\theta)$ with respect to θ, meaning that the posterior distribution is in the same family as the prior distribution.

2)

Given the information:

- The probability that an email is spam, $P(\text{Spam}) = 0.565$.
- The probability that an email is not spam, $P(\neg\text{Spam}) = 1 - P(\text{Spam}) = 0.435$.
- The probability that an email is flagged as spam given that it is actually spam, $P(\text{Flagged as Spam}|\text{Spam}) = 0.98$.
- The probability that an email is flagged as spam given that it is not spam (false positive), $P(\text{Flagged as Spam}|\neg\text{Spam}) = 0.05$.

We want to find the probability that an email is indeed spam given that it has been flagged as spam, $P(\text{Spam}|\text{Flagged as Spam})$.

Applying Bayes' theorem:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

In our context:

$$P(\text{Spam}|\text{Flagged as Spam}) = \frac{P(\text{Flagged as Spam}|\text{Spam}) \cdot P(\text{Spam})}{P(\text{Flagged as Spam})}$$

Where $P(\text{Flagged as Spam})$ is calculated as:

$$P(\text{Flagged as Spam}) = P(\text{Flagged as Spam}|\text{Spam})\cdot P(\text{Spam}) + P(\text{Flagged as Spam}|\neg\text{Spam})\cdot P(\neg\text{Spam})$$

Substituting the given values:

$$P(\text{Spam}|\text{Flagged as Spam}) = \frac{0.98 \cdot 0.565}{(0.98 \cdot 0.565) + (0.05 \cdot 0.435)}$$

This calculation leads to $P(\text{Spam}|\text{Flagged as Spam}) \approx 0.9622$, meaning there is approximately a 96.22% chance that an email is indeed spam given that it has been flagged as spam by the software.

3) Given the mean and variance of the prior distribution of p, a Beta distribution is a suitable choice due to its conjugacy with the binomial likelihood function. The prior parameters α and β are derived based on the provided mean and variance. After observing 26 heads in 50 tosses, the posterior distribution parameters, α' and β', are updated accordingly.

0.1 Posterior Distribution

The posterior distribution, being a Beta distribution, is updated as follows:

$$\alpha' = \alpha + k = 40.00,$$
$$\beta' = \beta + n - k = 37.00.$$

The posterior distribution is $\text{Beta}(\alpha', \beta')$.

0.2 Maximum A Posteriori (MAP) Estimate

The MAP estimate of p is given by:

$$p_{\text{MAP}} = \frac{\alpha' - 1}{\alpha' + \beta' - 2} \approx 0.520.$$

The posterior distribution indicates our updated belief about the probability p of landing heads, incorporating both our prior belief and the evidence from the coin tosses. The MAP estimate provides the most likely value of p based on the observed data.

4)

Given a set of independent and identically distributed observations $X_1, \ldots, X_N$ with a probability density function (pdf) proportional to $x^{b-1}1_{[0,1]}(x)$, where $b > 0$ is an unknown parameter and $1_{[0,1]}(x)$ is the indicator function defined as:

$$1_{[0,1]}(x) = \begin{cases} 1, & \text{for } x \in [0,1] \\ 0, & \text{otherwise} \end{cases}$$

For the parameter b, we consider as a prior the Gamma distribution, i.e.,

$$p_b(b) \propto b^{\alpha-1}e^{-\beta b},$$

defined for $b > 0$, where $\alpha, \beta > 0$.

To determine the maximum a posteriori (MAP) estimator of b given observations $X_1 = x_1, \ldots, X_N = x_N$, we use Bayes' theorem. The posterior distribution is proportional to the product of the likelihood function and the prior distribution of b.

The likelihood function is:

$$L(b|x_1, \ldots, x_N) \propto \left(\prod_{i=1}^{N} x_i \right)^{b-1}$$

The log of the posterior distribution is:

$$\log p(b|x_1, \ldots, x_N) = (b-1) \sum_{i=1}^{N} \log x_i + (\alpha - 1) \log b - \beta b + \text{const}$$

Taking the derivative of the log posterior with respect to b, setting it to zero, and solving for b gives the MAP estimator:

$$b_{\text{MAP}} = \frac{\alpha - 1}{\beta - \sum_{i=1}^{N} \log x_i}$$

This expression provides the MAP estimate of b based on the prior Gamma distribution parameters α and β, and the observed data through $\sum_{i=1}^{N} \log x_i$.