Masco: follow-along with R

2025-10-17

Poisson Likelihood with a Gamma Prior

A city hospital tracks the number of patients visiting the emergency room (ER) each day. A senior hospital administrator estimates that the ER typically sees around 50 patients per day.

Let Y_i be the number of visits on day i.

We model the daily number of ER visits as a Poisson random variable:

$$Y_i \mid \lambda \sim \text{Poisson}(\lambda), \quad i = 1, \dots, n.$$

where λ is the expected number of daily ER visits.

We assume that λ follows a Gamma distribution:

$$\lambda \sim \text{Gamma}(10, \theta).$$

```r
# Daily ER visits data of last two weeks
er_visits <- c(48,52,50,55,47,49,68,51,46,54,48,45,50,70)
```

Questions:

1. How can we use the administrator's information into the prior?
2. Write the code using `Rjags` to sample from the posterior distribution.
3. Generate a sample of size 2000 from the posterior distribution and plot its summary.
4. Compare the posterior distribution with the distribution $\text{Gamma}\Big(10 + \texttt{sum(er\_visits)}, 0.2 + 14\Big)$.
5. Do you think that the information provided by the administrator is relevant, compared to the posterior distribution?

Change point detection in time series

In this exercise, we explore change point detection, a problem that has been extensively studied in statistics.

1. Read Chapter 5.4 of Lee's Bayesian Cognitive Modeling.
2. Implement the corresponding Bayesian change point detection model.

```r
library(rjags)
true_mu1 <- 1 # value of the first mean
true_mu2 <- 2 # value of the second mean
sigma <- 0.5

set.seed(12)
# A time series of length 50, with a change point at 30
X <- c(rnorm(30, true_mu1, sigma), rnorm(20, true_mu2, sigma))
n <- length(X)
```

```
model_string <- "
model {
    # Prior

    # Likelihood
}
"

data_list <- list(X = X, n = n)

jags <- jags.model(textConnection(model_string), data = data_list, n.chains = 3)
update(jags, 1000)

# Parameters of interest
params <- c("mu1", "mu2", "sigma", "tau")

# Sampling from the posterior distribution
samp <- coda.samples(jags, params, n.iter = 5000)
```

- Run the following code to visualize your results.

```
library(coda)

# Extract posterior samples
tau_samples <- as.matrix(samp[, "tau"])
mu1_samples <- as.matrix(samp[, "mu1"])
mu2_samples <- as.matrix(samp[, "mu2"])

# Posterior estimates
tau_est <- mean(tau_samples)
tau_est_int <- quantile(tau_samples, probs = c(0.025, 0.975))
mu1_est <- mean(mu1_samples)
mu2_est <- mean(mu2_samples)

# Plot the time series
plot(X, type = "l", col = "blue", lwd = 2,
     xlab = "Time / Index", ylab = "Value",
     main = "Time Series with Estimated Change Point and Means")

# Change point
abline(v = tau_est, col = "red", lwd = 2, lty = 2)         # posterior mean
abline(v = tau_est_int, col = "red", lwd = 1, lty = 3)     # 95% CI

# Estimated means as horizontal lines
lines(1:floor(tau_est), rep(mu1_est, floor(tau_est)), col = "green", lwd = 2)
lines(ceiling(tau_est):length(X), rep(mu2_est, length(X) - ceiling(tau_est) + 1), col = "green", lwd = 

# Legend
legend("topleft",
       legend = c("X", "Estimated tau", "95% CI tau", "Estimated means"),
       col = c("blue", "red", "red", "green"),
       lty = c(1,2,3,1), lwd = c(2,2,1,2))
```

- In what ways might this model be relevant or useful for your Project 1 analysis?

Comparaison betweenn the means of two populations

1. Read Chapter 8 of Lee's Bayesian Cognitive Modeling.
2. Apply this model to the fructose dataset.