STAT 207, Spring 2026
Homework # 2
Released April 14. Due April 27 on Canvas

Instructions: You may discuss the homework problems in small groups, but you must write up the final solutions and codes yourself. You can either type your work in LaTeX or write down on papers and scan. Please make sure that all handwriting are visible and please combine all pages into a single PDF file (in the correct order).

For any problems that involve coding, you must provide written answers and also include your codes. You can either include your codes at the end of the homework and label which questions they correspond to, or include as part of the answer (e.g., in the R Markdown style). You will receive no credit if you submit only codes or only written answers.

1. BDA3 Problem 5.3. Instead of printing the table, visualize the pair-wise probabilities of one being better than the other as an 8×8 matrix.

2. BDA3 Problem 5.13

3. In this question, we will derive a Normal-Inverse-Wishart model for analyzing movie ratings. Consider the dataset being $X \in \mathbb{R}^{n \times p}$ for n raters and p movies. Assume every rater assigns a score to each of the p movies and the score has been pre-processed so that they can be modeled as Normally distributed random variables. We consider the model

$$X_{ij} \sim_{ind} N(u_i^T v_j, \sigma^2), \quad i = 1, ..., n, j = 1, ..., p$$

where $u_i \in \mathbb{R}^K$ is a K dimensional latent user feature vector, and $v_j \in \mathbb{R}^K$ is a K dimensional latent movie feature vector. We put independent Gaussian priors on u_i and v_j so that

$$u_i | \mu_u, \Sigma_u \sim_{ind} N(\mu_u, \Sigma_u), \quad i = 1, ..., n$$

$$v_j | \mu_v, \Sigma_v \sim_{ind} N(\mu_v, \Sigma_v), \quad j = 1, ..., p$$

For the hyperpriors, we let

$$\mu_u | \Sigma_u \sim N(\mu_0, \Sigma_u / \kappa_0), \quad \Sigma_u \sim InvWishart(\nu_0, \Lambda_0)$$

$$\mu_v | \Sigma_v \sim N(\mu_0, \Sigma_v / \kappa_0), \quad \Sigma_v \sim InvWishart(\nu_0, \Lambda_0)$$

   (a) Write out the steps of a Gibbs sampler to analyze the data.

   (b) This is not a question for the homework, but just a thought exercise. What if there are missing cells in this matrix. That is, not everyone has watched all movies. How would you still fit this model and use it for predicting the missing cell (what would user X rate movie Y if the user had watched it). You do not need to submit any answer to this question.

4. In this question, we will analyze data on 10 power plant pumps using a Poisson gamma model. The number of failures Y_i is assumed to follow a Poisson distribution

$$Y_i|\theta_i \sim_{ind} Poisson(\theta_i t_i), \quad i = 1, ..., 10$$

where θ_i is the failure rate for pump i and t_i is the length of operation time of the pump (in 1000s of hours). The data is as follows:

pump	1	2	3	4	5	6	7	8	9	10
t_i	94.3	15.7	62.9	126	5.24	31.4	1.05	1.05	2.1	10.5
y_i	5	1	5	14	3	19	1	1	4	22

We consider the conjugate gamma prior distribution for the failure rate

$$\theta_i|\alpha, \beta \sim_{ind} Gamma(\alpha, \beta), \quad i = 1, ..., 10$$

with hyperpriors $\alpha \sim Exp(1)$ and $\beta \sim Gamma(0.1, 1.0)$. All Gamma distribution uses the shape and rate parameters.

(a) Write out the steps of a Metropolis-Hastings within Gibbs sampling algorithm to analyze these data.

(b) Apply the algorithm to the data and show histograms of the posterior marginal distributions for α and β, and a scatter plot of the bivariate posterior distribution.

(c) Analytically integrate θ_i from the posterior and derive (up to proportionality) the posterior $p(\alpha, \beta|y)$.

(d) Construct a Metropolis-Hastings algorithm to sample from the posterior $p(\alpha, \beta|y)$ without sampling θ_i.

(e) Repeat part (b) using this reduced sampler.

(f) Describe how you can draw samples from $p(\theta_i|y)$ from the reduced sampler

A Codes for Q4(e)

```
logpost <- function(t,y,phi1,phi2){
    n <- length(t)
    term1 <- (n*exp(phi1)-0.9)*phi2 - n*lgamma(exp(phi1))
    term2 <- sum(lgamma(y+exp(phi1)))
    term3 <- -sum((y+exp(phi1))*log(t+exp(phi2))) - exp(phi1) - exp(phi2)
    logpost <- term1 + term2 + term3 + phi1 + phi2
    logpost
}
t <- c(94.3,15.7,62.9,126,5.24,31.4,1.05,1.05,2.1,10.5)
y <- c(5,1,5,13,3,19,1,1,4,22)
nsim <- 5000
phi1 <- phi2 <- NULL
phi1[1] <- phi2[1] <- 0
for (j in 2:nsim){
    oldpost1 <- logpost(t,y,phi1[j-1],phi2[j-1])
```

```
    phi1new <- rnorm(1,phi1[j-1],.2)
    newpost1 <- logpost(t,y,phi1new,phi2[j-1])
    phi1[j] <- ifelse(log(runif(1)) < newpost1 - oldpost1,phi1new,phi1[j-1])
    oldpost2 <- logpost(t,y,phi1[j],phi2[j-1])
    phi2new <- rnorm(1,phi2[j-1],.2)
    newpost2 <- logpost(t,y,phi1[j],phi2new)
    phi2[j] <- ifelse(log(runif(1)) < newpost2 - oldpost2,phi2new,phi2[j-1])
}
burnin <- 1000


index <- seq(burnin+1,nsim)
pdf("../HW/figures/HW2-poisson-gamma-MH-2.pdf", width = 10, height = 7)
par(mfrow=c(3,2))
plot(exp(phi1[index]),ylab=expression(alpha),xlab="Iteration", type = "l")
plot(exp(phi2[index]),ylab=expression(beta),xlab="Iteration", type = "l")
hist(exp(phi1[index]),xlab=expression(alpha),main="")
hist(exp(phi2[index]),xlab=expression(beta),main="")
plot(exp(phi1[index]),exp(phi2[index]),xlab=expression(alpha),
     ylab=expression(beta))
 dev.off()
```