Student Name: ______________________________________________

# STAT 544 Mid-term Exam
# Thursday 2 March 8:00-9:20

Instructor: Jarad Niemi

**INSTRUCTIONS**

Please check to make sure you have 4 pages with writing on the front and back (some pages are marked 'intentionally left blank'). Feel free to remove the last page, i.e. the one with R code.

On the following pages you will find short answer questions related to the topics we covered in class for a total of 50 points. Please read the directions carefully.

You are allowed to use a calculator and one $8\frac{1}{2} \times 11$ sheet of paper with writing on both front and back. A non-exhaustive list of items you are not allowed to use are **cell phones, laptops, PDAs, and textbooks**. Cheating will not be tolerated. Anyone caught cheating will receive an automatic F on the exam. In addition the incident will be reported, and dealt with according to University's Academic Dishonesty regulations. Please refrain from talking to your peers, exchanging papers, writing utensils or other objects, or walking around the room. All of these activities can be considered cheating. **If you have any questions, please raise your hand.**

You will be given only the time allotted for the course; no extra time will be given.

Good Luck!

## Rescue probability

1. In Colorado, seventy percent of skiers lost during an avalanche are subsequently discovered. Of the skiers that are discovered, 60% have an emergency locator, whereas 90% of the skiers not discovered do not have such a locator. Suppose that a skier has disappeard in an avalanche. If she has an emergency locator, what is the probability that she will be discovered? (20 points)

   Answer: Let $D$ indicate the event the skier will be discovered and let $L$ indicate the event the skier has an emergency locator. We are given

$$\begin{aligned} P(L|D) &= 0.6 \\ P(L^C|D^C) &= 0.9 \\ P(D) &= 0.70 \end{aligned}$$

   We are asked to find

$$\begin{aligned} P(D|L) &= \frac{P(L|D)P(D)}{P(L|D)P(D)+P(L|D^C)P(D^C)} \\ &= \frac{P(L|D)P(D)}{P(L|D)P(D)+[1-P(L^C|D^C)][1-P(D)]} \\ &= \frac{0.6\times 0.7}{0.6\times 0.7+[1-0.9][1-0.7]} \\ &= \frac{0.42}{0.42+0.03} = 0.933 \end{aligned}$$

   So the probability the aircraft will be discovered givent that it has an emergency locator, is about 93.3%.

## Negative binomial

2. Let $Y \sim NB(r, \theta)$ where $r$ is known and $\theta$ is unknown. The probability mass function for $y$ is

$$p(y) = \binom{y+r-1}{y}(1-\theta)^r\theta^y \quad y = 0, 1, 2, \ldots$$

where $E[Y] = \frac{r\theta}{1-\theta}$.

(a) Derive the Jeffreys' prior for $\theta$. (20 points)

Answer: Jeffreys' prior is proportional to the the square root of the Fisher information.

$$\begin{aligned}
\log L(\theta) &= \log\left(\binom{y+r-1}{y}\right) + r\log(1-\theta) + y\log(\theta) \\
\frac{d}{d\theta}\log L(\theta) &= -\frac{r}{1-\theta} + \frac{y}{\theta} \\
\frac{d^2}{d\theta^2}\log L(\theta) &= -\frac{r}{(1-\theta)^2} - \frac{y}{\theta^2} \\
\mathcal{I}(\theta) &= -E\left[\frac{d^2}{d\theta^2}\log L(\theta)\right] \\
&= -E\left[-\frac{r}{(1-\theta)^2} - \frac{y}{\theta^2}\right] \\
&= \frac{r}{(1-\theta)^2} + \frac{E[y]}{\theta^2} \\
&= \frac{r}{(1-\theta)^2} + \frac{\frac{r\theta}{1-\theta}}{\theta^2} \\
&= \frac{r}{(1-\theta)^2} + \frac{r}{(1-\theta)\theta} \\
&= r\frac{\theta+(1-\theta)}{(1-\theta)^2\theta} \\
&= r\frac{1}{(1-\theta)^2\theta} \\
p_{Jeffreys}(\theta) &\propto \sqrt{\mathcal{I}(\theta)} \\
&= \sqrt{r\frac{1}{(1-\theta)^2\theta}} \\
&\propto (1-\theta)^1\theta^{1/2}
\end{aligned}$$

This is the kernel of a $Be(2, 1.5)$ distribution.

(b) Derive the posterior for $\theta$ assuming $\theta \sim Be(a, b)$. (10 points)

Answer:

$$\begin{aligned} p(\theta|y) &\propto p(y|\theta)p(\theta) \\ &\propto (1-\theta)^r \theta^y \theta^{a-1}(1-\theta)^{b-1} \\ &= \theta^{a+y-1}(1-\theta)^{b+r-1} \end{aligned}$$

Thus $\theta|y \sim Be(a + y, b + r)$.

(c) For a particular data set, you found the posterior to be $\theta|y \sim Be(150, 200)$. Provide R code to compute an equal-tail 95% credible interval for $\theta$. (4 points)

Answer:

```
qbeta(c(.025,.975), 150, 200)
```

(d) Using the Central Limit Theorem, find an approximate equal-tail 95% credible interval for $\theta$ under the posterior $\theta|y \sim Be(150, 200)$. (6 points)

Answer: We have

$$\begin{aligned} E[\theta|y] &= \tfrac{150}{150+200} = 0.429 \\ Var[\theta|y] &= \tfrac{150\times 200}{(150+200)^2(150+200+1)} \\ &= 0.0006977 \end{aligned}$$

Thus, an approximate equal-tail 95% credible interval for $\theta$ is

$$0.429 \pm 1.96 \times \sqrt{0.0006977} = (0.377, 0.481).$$

# JAGS interpretation (15 pts total)

3. Use the R/JAGS code and output, to answer the following questions

   (a) Write down the model that is being fit including priors. (6 pts)
   Answer: For a binary response $Y_{ij}$ for individual $i$ in group $j$, the model is

$$
\begin{aligned}
Y_{ij} &\overset{ind}{\sim} Ber(p_{ij}) \\
\text{logit}(p_{ij}) &= \alpha_j + b_j x_i \\
\alpha_j &\overset{ind}{\sim} N(\mu_a, \sigma_a^2) \\
\beta_j &\overset{ind}{\sim} N(\mu_b, \sigma_b^2) \\
\mu_a &\sim N(0, 100^2) \\
\mu_b &\sim N(0, 100^2) \\
\sigma_a &\sim \text{Unif}(0, 1000) \\
\sigma_b &\sim \text{Unif}(0, 1000)
\end{aligned}
$$

   (b) For **group 1**, answer the following questions to 2 decimal place

      i. Provide a 95% credible interval for $\alpha$. (1 pt)
      Answer: (-3.32,0.27)

      ii. Provide a 95% credible interval for $\beta$. (1 pt)
      Answer: (-4.63,0.25)

      iii. Provide a 95% credible interval for the success probability if `x` is zero. (2 pts)
      Answer: Because $e^{\alpha}/(1+e^{\alpha})$ is monotonic, we can evalute the endpoints of the interval from part 3(b)i. in $e^{\alpha}/(1+e^{\alpha})$ to get (0.03, 0.57).

      iv. Provide a point estimate for the success probability if `x` is one. (1 pt)
      Answer: Plug in medians for $\alpha_1$ and $\beta_1$ to get

$$\frac{e^{\hat{\alpha}_1+\hat{\beta}_1}}{1+e^{\hat{\alpha}_1+\hat{\beta}_1}} = \frac{e^{-1.33-1.51}}{1+e^{-1.33-1.51}} \approx 0.06$$

      v. Describe how you would obtain a 95% credible interval for the success probability if `x` is one. (2 pts)
      Answer: For iteration values $\alpha^{(j)}$ and $\beta^{(j)}$, calculate $e^{\alpha^{(j)}+\beta^{(j)}}/(1+e^{\alpha^{(j)}+\beta^{(j)}})$. Then take the 2.5% and 97.5% quantiles of these values.

   (c) For which groups does the estimated equal-tail 95% credible interval for the coefficient for $x$ **not** include 0? (2 pts)
   Answer: Group 4.

R/JAGS code

```
## Error:  package or namespace load failed for 'rjags':
##  .onLoad failed in loadNamespace() for 'rjags', details:
##  call:  dyn.load(file, DLLpath = DLLpath, ...)
##  error:  unable to load shared object '/usr/lib64/R/library/rjags/libs/rjags.so':
##  libjags.so.4:  cannot open shared object file:  No such file or directory
```

```
library("rjags")

## Error:  package or namespace load failed for 'rjags':
##  .onLoad failed in loadNamespace() for 'rjags', details:
##  call:  dyn.load(file, DLLpath = DLLpath, ...)
##  error:  unable to load shared object '/usr/lib64/R/library/rjags/libs/rjags.so':
##  libjags.so.4:  cannot open shared object file:  No such file or directory

model = "
model
{
  for (i in 1:N)
  {
    y[i] ~ dbern(pi[i])
    pi[i]   <- exp(eta[i]) / ( 1 + exp(eta[i]) )
    eta[i] <- alpha[g[i]] + beta[g[i]] * x[i]
  }

  for (k in 1:K)
  {
    alpha[k] ~ dnorm(mu.a, tau.a)
    beta[k]  ~ dnorm(mu.b, tau.b)
  }

  mu.a ~ dnorm(0, 0.0001)
  mu.b ~ dnorm(0, 0.0001)

  tau.a <- 1/sigma.a^2
  tau.b <- 1/sigma.b^2
  sigma.a ~ dunif(0, 1000)
  sigma.b ~ dunif(0, 1000)
}
"

dat = list(y = y, N = N, g = g, K = K, x = x)
m = jags.model(textConnection(model), dat)

## Error in jags.model(textConnection(model), dat):  could not find function "jags.model"
```

## R/JAGS output

```
res = coda.samples(m,
                   c("alpha","beta","mu.a","mu.b","sigma.a","sigma.b"),
                   n.iter=1000)

## Error in coda.samples(m, c("alpha", "beta", "mu.a", "mu.b", "sigma.a", :  could not find
function "coda.samples"

summary(res)$quantiles

## Error in summary(res):  object 'res' not found
```