INSERT YOU ANSWERS IN THE SPACE ALLOTED FOR EACH QUESTION. DO NOT USE MORE SPACE!

MAIL THIS FORM (AND THIS FORM ONLY) TO [H.HOIJTINK@UU.NL](mailto:H.HOIJTINK@UU.NL) WHEN YOU ARE FINISHED.

YOU CAN WORK ON THE EXAM FROM 13.30 UNTIL 16.00.

**NAME:** Nina van Gerwen

**STUDENT NUMBER:** 1860852

**E-MAIL ADDRESS:** n.l.vangerwen@students.uu.nl

**Question 1 (2.5 points):**

Use the variables sex (1=boy, 2=girl) and postnumb from the datafile sesame.txt. Create using R-JAGS a Gibbs sampler for the parameters of an ANOVA with sex as factor and postnumb as the dependent variable. Create the sampler such that you can answer questions b) and c) using the output from the sampler.

1. Copy past your R and JAGS code right here.  
   RCode:  
   library(rjags)

sesame <- read.table("sesame.txt", header = TRUE)

dat <- sesame[, c(3, 17)]

dat$sex <- dat$sex - 1

n <- nrow(dat)

listdat <- list("n" = n, "sex" = dat[, 1], "postnumb" = dat[, 2])

model.def <- jags.model(file = "ModelQ1.txt",

data = listdat, n.chains = 2)

update(model.def, 1000)

parameters <- c("b0", "b1", "sigma")

set.seed(123)

results <- coda.samples(model = model.def, variable.names = parameters,

n.iter = 10000)

summary(results)

JAGS code:  
## model of the data

model{

## Likelihood

for(i in 1:n){

mu[i] <- b0 + b1 \* sex[i]

postnumb[i] ~ dnorm(mu[i], tau)

}

## Prior distributions for the coefficients

b0 ~ dnorm(0.0, 0.0001)

b1 ~ dnorm(0.0, 0.0001)

## To get the standard deviation

sigma <- 1.0/sqrt(tau)

## Prior for residual variance

tau ~ dgamma(0.001, 0.001)

}

1. Use the output from the Gibbs sampler to determine if the means of the boys and the girls are equal or not (this is not a request to compute a p-value). Give the most relevant numerical result and an interpretation of at most 20 words.  
   A: The means of boys and girls on postnumb are practically equal (Mb1 = 0.02). This means that boys and girls do not differ on postnumb.
2. Use the output from the Gibbs sampler to determine if Cohen’s d is equal to zero or not (this is not a request to compute a p-value). Give the most relevant numerical result and an interpretation of at most 20 words.   
   A: Assuming equal variance in boys and girls: Cohen’s d = 0.02264 / 12.91 = 0.001. So yes, Cohen’s d is basically equal to 0 and there is no standardized difference in postnumb between boys and girls.

**Question 2 (2.5 points):**

Whether or not two means are equal can be tested using a posterior predictive p-value and a test-statistic (do not use a discrepancy measure!). Describe the steps involved in the computation of this posterior predictive p-value (the one for testing the equality of two means).

Use short labels for each step, e.g., “Give an interpretation of the p-value”, explain each step using statistical notation and a short (max 10 words, not counting the statistical notation) elaboration, e.g., if p < .05, H0: mu1 = mu2, is rejected, otherwise it is not rejected. Please note, if you refer to, e.g., the prior distribution, you can write it as h(mu1,mu2,sigma2), you do not have to further specify it by giving the formula of the density. This holds for all the distributions and densities you will refer to in your answer.

Step 1: Specifying the null model (i.e., the means in two groups are equal)

* H0: μ1 = μ 2

Step 2: Sampling μ and sigma of both groups (i.e., gaining the posterior distributions for both groups)

* Gaining μ11, μ12, …, μ1,1000 and sigma1, sigma12, …, sigma1,1000 and  
  μ21, μ22, …, μ2,1000 and sigma2, sigma22, …, sigma2,1000

Step 3: For every sampled value, simulate a dataset with equal size as the observed dataset

* Gaining Y11, Y12, …, Y1,1000 and Y21, Y22, …, Y2,1000

Step 4: Calculate the mean difference of every simulated dataset and your observed dataset

* μY11 – μY21, μY12 – μY21, …, μY1,1000 – μY2,1000
* μ1 – μ 2

Step 5: Calculate the amount of times the mean difference in a simulated dataset is larger than in the observed dataset, given the observed dataset and null model

* *p* = P(μY1t – μY2t > μ1 – μ 2 | H0)

Step 6: If *p* < .05, H0: μ1 = μ2 is rejected and otherwise it is not.

**Question 3 (2.5 points):**

During the course you have encountered a number of statistical concepts/tools that have both Bayesian and Frequentist approaches. Consider for example the definition of probability, points estimates, certainty intervals, hypothesis testing, p-values, estimation procedures, information criteria, and so on.

Choose three statistical/concepts tools that have both Bayesian and Frequentist approaches (from the list above or other concepts/tools you may think of). For each of these concepts: 1) name a Bayesian and a frequentist approach, and 2) briefly (1-3 sentences) explain the most important difference between that Bayesian and Frequentist approach.

Statistical concept/tool 1: Probability

Frequentist approach: Relative frequency over time

Bayesian Approach: Degree of belief of an individual about an expectation

Most important difference: In frequentist, you can’t throw a coin and then say ‘It has a 50% chance to be heads.’ It either is heads or it is not. You can say, however, if I throw the coin 1000 times, it’ll be heads 50% of the time. However, in Bayesian statistics, you can state about a single event that it has a probability of being something.

Statistical concept/tool 2: Certainty intervals

Frequentist approach: Confidence Intervals

Bayesian Approach: Credible Intervals

Most important difference: A confidence interval captures uncertainty about the parameter we estimated, but we can never know whether the true parameter is in the interval or not. Whereas a credible interval captures uncertainty in the location of the parameter and can be interpreted as a probabilistic statement about the parameter.

Statistical concept/tool 3: P-values

Frequentist approach: Significance testing

Bayesian Approach: Posterior predictive p-values

Most important difference: In Frequentist statistic, the p-value gives the probability of an outcome more extreme than the outcome found under the null hypothesis and hence is used to test for significance (i.e., if smaller than .05, we can safely reject the null hypothesis). The posterior predictive p-value, however, is a method to assess model fit (i.e., whether the chosen density of a data provides a good description of the data) through sampling from the posterior and simulating data.

**Question 4 (2.5 points):**

1. Load the file “SamplesExamQ4.Rdata" into R. You can use function ‘load()’ to do this. Inside, you will find object ‘B’, which contains 4000 samples from a Bayesian MCMC sampler for parameter ‘B’.

Did these samples come from a Gibbs’ sampler or a Metropolis Hastings sampler? Motivate your answer briefly (1-3 sentences). Alongside your motivation, include the most important rcode & output you used as evidence for your answer.  
  
A: These samples came from a MH sampler. You can see this because the values stay the same for multiple iterations. The first three elements, for example, are all 0, whereas the next 6 elements are all 0.467. This is very typical for a MH sampler, because you do not always accept a new value and if you do not, you keep the old value (that’s why you see a value multiple times). In a Gibbs sampler, instead, you always get a new value, so no two values should be the same.  
Furthermore, it seems that from the 4001 sampled values, there are only 588 unique values. In other words, a lot of values were repeated multiple times.  
  
RCode:

B[1:10]  
unique(B)  
hist(B, breaks = 10000)

1. Briefly explain (1-3 sentences) when one uses a Gibbs’ sampler, and when one would use a Metropolis Hastings sampler.  
   A: One would use a Gibbs’ sampler, when they know the form of the posterior distribution and are able to get either a conjugate (same form as posterior) or semi conjugate prior (same mathematical form as the conditional posterior) for all parameters.  
   A MH sampler you would thus use when you do not know this information (i.e., when you have non-conjugate priors).