Lesson 9

Тест, 8 вопроса

1
Баллы

1.
What is the advantage of using a link function such as the logit transform for logistic regression?

- (selected) It ensures that the success probability ($\mathrm{E}(y)$ if y is Bernoulli) is between 0 and 1 without requiring any constraints on the x variables or the β coefficients.
- It ensures that the β coefficients lie between 0 and 1 for all values of predictors x.
- It ensures that $\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$ is between 0 and 1 using log transformations of the β coefficients.
- It makes the β coefficients interpretable directly as probabilities.

1
Баллы

2.

Logistic regression works with binomial likelihoods in addition to Bernoulli likelihoods. If the response y_i is a number of successes in n_i independent trials each with ϕ_i success probability, we can still model ϕ_i with a linear model using the logit transformation.

As an example, consider the OME data in the `MASS` package in R. The data consist of experimental results from tests of auditory perception in children. Under varying conditions and for multiple trials under each condition, children either correctly or incorrectly identified the source of changing signals.

Although the independence of the trails and results are questionable, we'll try fitting a logistic regression to these data. First, we'll explore the relationships briefly with the following code:

```
library("MASS")
data("OME")
?OME # background on the data
head(OME)

any(is.na(OME)) # check for missing values
dat = subset(OME, OME != "N/A") # manually remove OME missing values identified with "N/A"
dat$OME = factor(dat$OME)
str(dat)

plot(dat$Age, dat$Correct / dat$Trials )
plot(dat$OME, dat$Correct / dat$Trials )
plot(dat$Loud, dat$Correct / dat$Trials )
plot(dat$Noise, dat$Correct / dat$Trials )

```

We are interested how these variables relate to the probability of successfully identifying the source of changes in sound. Of these four variables, which appears to have the weakest association with the probability of success?

- ○ Age in months
- ◉ OME: degree of otitis media with effusion (low or high)
- ○ Loudness of stimulus in decibels
- ○ Noise: stimulus type (coherent or incoherent)

1
Баллы

3.

Next, we'll fit a reference logistic regression model with noninformative prior in R. We can do this with the **glm** function, providing the model formula as with the usual **lm**, except now the response is the observed proportion of correct responses. We must also indicate how many trials were run for each experiment using the **weights** argument.

```
mod_glm = glm(Correct/Trials ~ Age + OME + Loud + Noise, data=dat, weights
    =Trials, family="binomial")
summary(mod_glm)
```

To get an idea of how the model fits, we can create residual (using a special type of residual for non-normal likelihoods) and in-sample prediction plots.

```
plot(residuals(mod_glm, type="deviance"))
plot(fitted(mod_glm), dat$Correct/dat$Trials)
```

It appears from the second plot that the model is not very precise (some model predictions were far from the observed proportion of correct responses). Nevertheless, it can be informative about the relationships among the variables.

Report the posterior mode estimate of the coefficient for low OME. Round your answer to two decimal places.

-0.24

1
Баллы

4.

Next, we will fit a similar model in JAGS. To make the results comparable to those of the reference model, we will use the same configuration of covariates. We can extract this information from the reference model using **model.matrix**.

```
X = model.matrix(mod_glm)[,-1] # -1 removes the column of 1s for the intercept
head(X)
```

The data include categorical covariates which R codes as dummy variables (as with ANOVA). Hence we have an indicator variable for whether OME is at its low level and another indicating whether the Noise is incoherent. The intercept is then associated with this baseline group. Ignoring the continuous variables **Age** and **Loud**, what are the characteristics of this baseline group?

- ○ Children with low OME exposed to coherent sound.
- ○ Children with low OME exposed to incoherent sound.
- ● Children with high OME exposed to coherent sound.
- ○ Children with high OME exposed to incoherent sound.

Lesson 9

Тест, 8 вопроса

1
Баллы

5.

Now complete the following code (as well as the code from previous questions) to fit the JAGS model with the fairly noninformative priors given. Use three chains with at least 5,000 iterations in each.

```
mod_string = " model {
  for (i in 1:length(y)) {
    y[i] ~ dbin(phi[i], n[i])
    logit(phi[i]) = b0 + b[1]*Age[i] + b[2]*OMElow[i] + b[3]*Loud[i] + b[4]*Noiseincoherent[i]
  }

  b0 ~ dnorm(0.0, 1.0/5.0^2)
  for (j in 1:4) {
    b[j] ~ dnorm(0.0, 1.0/4.0^2)
  }

} "

data_jags = as.list(as.data.frame(X))
data_jags$y = dat$Correct # this will not work if there are missing values in dat (because they would be ignored by model.matrix). Always make sure that the data are accurately pre-processed for JAGS.
data_jags$n = dat$Trials
str(data_jags) # make sure that all variables have the same number of observations (712).

```

Because there are many data points, the MCMC will take some time to run.

Before analyzing the results, perform some MCMC diagnostic checks. What does the Raftery and Lewis diagnostic (`raftery.diag()`) suggest about these chains?

- ( ) The dependence factor for many of the variables is large (>5.0), indicating weak autocorrelation in the chains. We would not require a large number of iterations to reliably produce 95% probability intervals for the parameters.
- (•) The dependence factor for many of the variables is large (>5.0), indicating strong autocorrelation in the chains. We would require a large number of iterations to reliably produce 95% probability intervals for the parameters.
- ( ) The scale reduction factor for many variables is large (>5.0), indicating that the different chains are not exploring the same space yet. We need to run a longer burn-in period.
- ( ) The scale reduction factor for many variables is large (>5.0), indicating that the different chains are exploring the same space. We have used a sufficient burn-in time.

1
Баллы

Lesson 9

Тест, 8 вопроса

6.

Although `OMElow` is the predictor with weakest statistical association to probability of correct responses, the posterior probability that its coefficient β_2 is negative is still greater than 0.9. How do we interpret this (most likely) negative coefficient in the context of our models?

- (selected) While holding all other predictors constant, low OME is associated with a lower probability of correct responses than high OME.
- While holding all other predictors constant, low OME is associated with a decrease of magnitude $|\beta_2|$ in the probability of correct responses while high OME is associated with an increase $|\beta_2|$.
- While holding all other predictors constant, low OME is associated with an increase of magnitude $|\beta_2|$ in the probability of correct responses while high OME is associated with a decrease of $|\beta_2|$.
- While holding all other predictors constant, low OME is associated with a higher probability of correct responses than high OME.

1
Баллы

7.

Using the posterior mean estimates of the model coefficients, create a point estimate of the probability of correct responses for a child of age 60 months, with high OME, using a coherent stimulus of 50 decibels. Round your answer to two decimal places.

Hint: First calculate the linear part by multiplying the variables by the coefficients and adding them up (call this xb). Once you have that, apply the inverse of the link function to transform it into a probability estimate.

0.92

1
Баллы

8.

Use the posterior mean estimates of the model coefficients to create point estimates of the probability of correct responses for each observation in the original data. To do this, follow the steps outlined in the lesson to create a vector of these probabilities called **`phat`** (using our notation from this quiz, it would be $\hat{\phi}$).

Once you have **`phat`**, calculate the proportion of in-sample observations that are correctly classified according to the following criterion: the model prediction and observed correct response rate are either both higher than 0.7 or both lower than 0.7. Round your answer to two decimal places.

Hint: Use the following code:

```
(tab0.7 = table(phat > 0.7, (dat$Correct / dat$Trials) > 0.7))
sum(diag(tab0.7)) / sum(tab0.7)
```

0.84

Оплатить курс