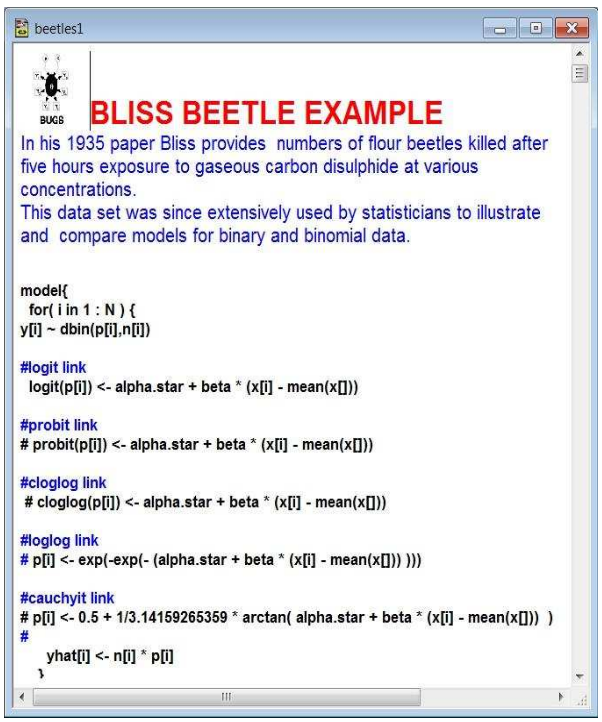
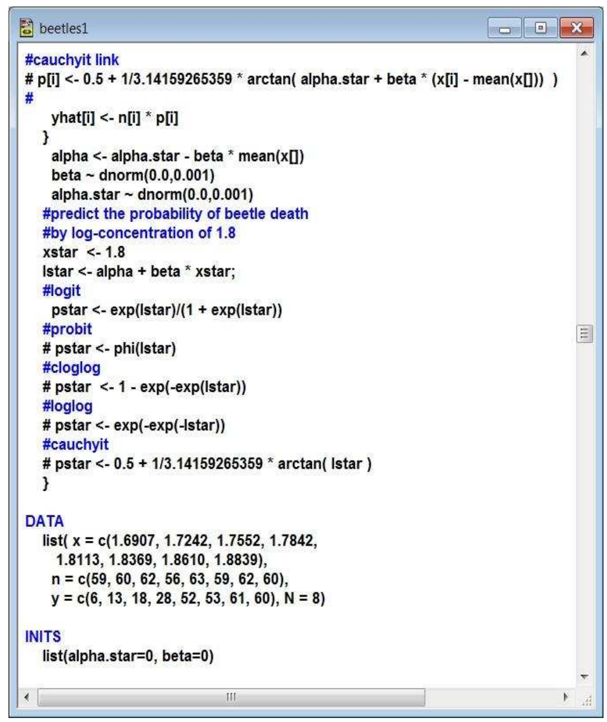
7.4 EXERCISES

BMED6420
Brani Vidakovic; Fall 2018
Consult the class slides, hints, and cited literature for the solution of exercise problems.

1. Beetles (Bliss Data). In his 1935 paper[1], Bliss provides a table showing a number of flour beetles killed after 5 hours of exposure to gaseous carbon disulfide at various concentrations. This data set has since been used extensively by statisticians to illustrate and compare models for binary and binomial data.

The program shows use of logit, probit, cloglog, loglog and cauchyit links.





2. Vasoconstriction. The data give the presence or absence ($y_i = 1$ or 0) of vasoconstriction in the skin of the fingers following inhalation of a certain volume of air (v_i) at a certain average rate (r_i). Total number of records is 39. The candidate models for analyzing the relationship are the usual `logit`, `probit`, `cloglog`, `loglog`, and `cauchyit` models.

Data are given as follows (a startup file `vasoconstriction0.odc` is also applied to save you some time):

```
y:  1,1,1,1,1,1,0,0,0,0,0,0,0,1,1,1,1,1,
```

[1] Bliss, C. I. (1935). The calculation of the dose-mortality curve. *Ann. Appl. Biol.*, **22**, 134–167.

```
0,1,0,0,0,0,1,0,1,0,1,0,1,0,0,1,1,1,0,0,1

v: 3.7, 3.5, 1.25, 0.75, 0.8, 0.7, 0.6, 1.1, 0.9, 0.9,
0.8, 0.55, 0.6, 1.4, 0.75, 2.3, 3.2, 0.85, 1.7, 1.8,
0.4, 0.95, 1.35, 1.5, 1.6, 0.6, 1.8, 0.95, 1.9, 1.6,
2.7, 2.35, 1.1, 1.1, 1.2, 0.8, 0.95, 0.75, 1.3

r: 0.825, 1.09, 2.5, 1.5, 3.2, 3.5, 0.75, 1.7, 0.75,
0.45, 0.57, 2.75, 3, 2.33, 3.75, 1.64, 1.6, 1.415,
1.06, 1.8, 2, 1.36, 1.35, 1.36, 1.78, 1.5, 1.5, 1.9,
0.95, 0.4, 0.75, 0.3, 1.83, 2.2, 2, 3.33, 1.9, 1.9, 1.625
```

(a) Transform covariates v and r as

$$x_1 = \log(10 \times v), \qquad x_2 = \log(10 \times r).$$

(b) Using WinBUGS estimate posterior means for coefficients in the logit model. Use noninformative priors on all coefficients.

(c) For a subject with $v = r = 1.5$, find the probability of vasoconstriction.

(d) Which of the five links: logit, probit, cloglog, loglog, and cauchyit, has smallest deviance. An example for use of the five links is provided: `beetles1.odc`. Uncomment and run one link at the time.

3. Caesarean Delivery: Categorical Response. Table 1 contains grouped data on infection of mothers after a C-section, collected at the Clinical Center of the University of Munich. [2] The response variable has three categories: Infection of type I, Infection of type II, and No infection. For each mother three covariates are collected:

$$\text{NOPLAN} = \begin{cases} 1 & \text{C-section was not planned,} \\ 0 & \text{Planned.} \end{cases}$$

$$\text{RISK} = \begin{cases} 1 & \text{Risk factors present,} \\ 0 & \text{No risk factors.} \end{cases}$$

$$\text{ANTIB} = \begin{cases} 1 & \text{Antibiotics given as prophylaxis,} \\ 0 & \text{No antibiotics given.} \end{cases}$$

(a) Given the covariates, establish a multinomial model, where the outcome 'No infection' serves as a baseline.

[2] Tutz, G. (2000). Die Analyse kategorialer Daten - eine anwendungsorientierte Einfhrung in Logit-Modellierung und kategoriale Regression. Oldenbourg-Verlag.

Table 1: Data on infections for 251 C-sections.

	C-section					
	Planned			Unplanned		
	Infection			Infection		
	I	II	No	I	II	No
Antibiotics						
Risk factor	0	1	17	4	7	87
No risk factor	0	0	2	0	0	0
No Antibiotics						
Risk factor	11	17	30	10	13	3
No risk factor	4	4	32	0	0	9

(b) A new C-section delivery for a mother with covariates (NOPLAN, RISK, ANTIBIO)=(1,0,0) is to be evaluated for risks of infection. What are estimated probabilities of no infection, and type I and II infections.

Hint: Consult `NHANESmulti.odc` discussed in UNIT7

4. Magnesium Ammonium Phosphate and Chrysanthemums. Walpole et al. (2007) provide data from a study on the effect of magnesium ammonium phosphate on the height of chrysanthemums, which was conducted at George Mason University in order to determine a possible optimum level of fertilization, based on the enhanced vertical growth response of the chrysanthemums. Forty chrysanthemum seedlings were assigned to 4 groups, each containing 10 plants. Each was planted in a similar pot containing a uniform growth medium. An increasing concentration of $MgNH_4PO_4$, measured in grams per bushel, was added to each plant. The 4 groups of plants were grown under uniform conditions in a greenhouse for a period of 4 weeks. The treatments and the respective changes in heights, measured in centimeters, are given in the following table:

Treatment			
50 g/bu	100 g/bu	200 g/bu	400 g/bu
13.2	16.0	7.8	21.0
12.4	12.6	14.4	14.8
12.8	14.8	20.0	19.1
17.2	13.0	15.8	15.8
13.0	14.0	17.0	18.0
14.0	23.6	27.0	26.0
14.2	14.0	19.6	21.1
21.6	17.0	18.0	22.0
15.0	22.2	20.2	25.0
20.0	24.4	23.2	18.2

Solve the problem as a Bayesian one-way ANOVA. Use STZ constraints on treatment effects.

(a) Do different concentrations of $MgNH_4PO_4$ affect the average attained height of chrysanthemums? Look at the 95% credible sets for the differences between treatment effects.

(b) Find the 95% credible set for the contrast $\mu_1 - \mu_2 - \mu_3 + \mu_4$.

5. Third-degree Burns. The data for this exercise, discussed in Fan et al. (1995), refer to $n = 435$ adults who were treated for third-degree burns by the University of Southern California General Hospital Burn Center. The patients were grouped according to the area of third-degree burns on the body. For each midpoint of the groupings "log(area +1)," the number of patients in the corresponding group who survived and the number who died from the burns was recorded:

Log(area+1)	Survived	Died
1.35	13	0
1.60	19	0
1.75	67	2
1.85	45	5
1.95	71	8
2.05	50	20
2.15	35	31
2.25	7	49
2.35	1	12

(a) Fit the logistic regression on the probability of death due to third-degree burns with the covariate `x = log(area+1)`. What is the deviance?

(b) Using your model, estimate find the posterior probability of survival for a person for which `log(area + 1)` equals 2.

(c) Repeat (a) with probit and complemenry log-log links. In terms of deviance, which morel provides the best fit.

6. Shocks! An experiment was conducted to assess the effect of small electrical currents on farm animals, with the eventual goal of understanding the effects of high-voltage powerlines on livestock. The experiment was carried out with seven cows, and six shock intensities, 0, 1, 2, 3, 4, and 5 milliamps (shocks on the order of 15 milliamps are painful for many humans.[3] Each cow was given 30 shocks, five at each intensity, in random order. The entire experiment was then repeated, so each cow received a total of 60 shocks. For each shock the response, mouth movement, was either present or absent. The data as quoted give the total number of responses, out of 70 trials, at each shock level. We ignore cow differences and differences between blocks (experiments).

[3] C. F. Dalziel, J.B. Lagen and J. L. Thurston, *Electric shocks*, *Trans IEEE* **60** (1941), 1073-1079.

Current (milliamps) x	Number of Responses y	Number of Trials n	Proportion of Responses p
0	0	70	0.000
1	9	70	0.129
2	21	70	0.300
3	47	70	0.671
4	60	70	0.857
5	63	70	0.900

As in Exercise Beetles (Bliss Data) model y as a function of x via binary regression with 5 different links and propose the link that minimizes the deviance.

7. Three WinBUGS Programs and GLM Practice. Three WinBUGS programs are supplied: iop2.odc, NBreg.odc, and terrapins.odc.

Read intro to the programs (preables in ODC files) and run them on your computer. Understand the output.