2132 Midterm

Orly Olbum

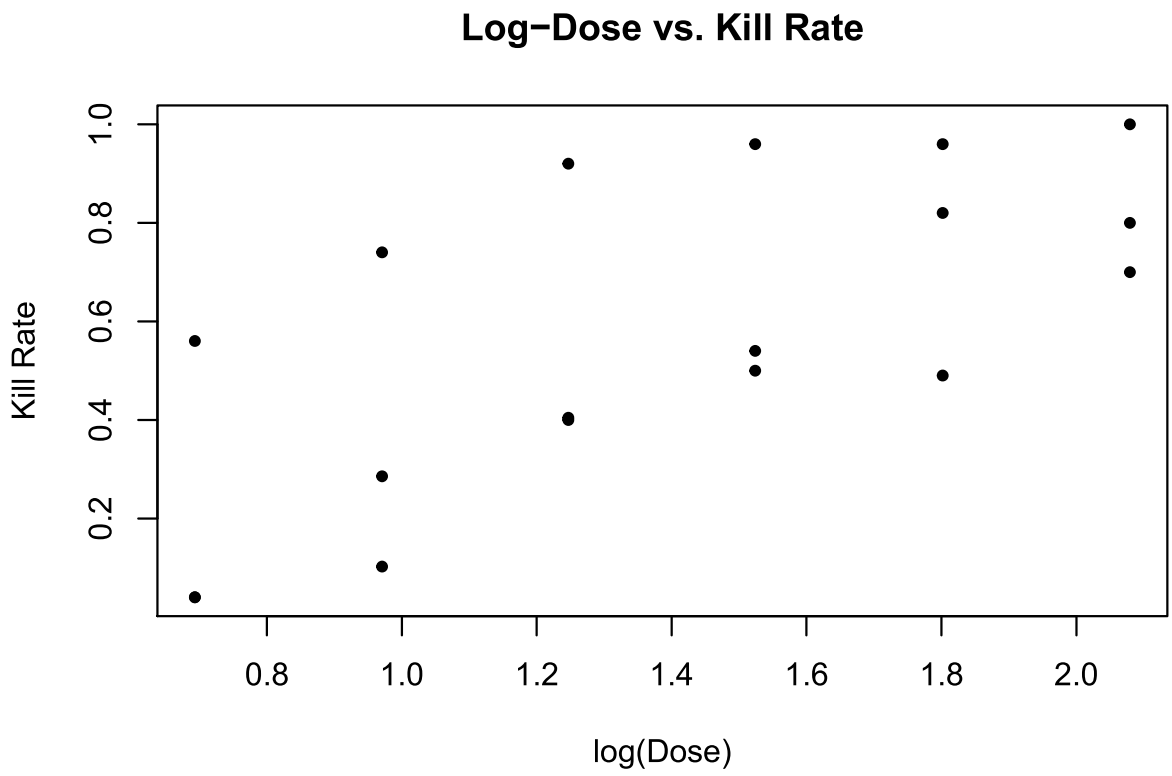
3/10/2021

Problem 1

Flour beetles Tribolium castaneum were sprayed with one of three insecticides in solution at different doses. The file "Insecticide.txt" contains data relating to the number of insects killed after a six-day period. The columns are:
- Type: A factor variable with 3 levels giving the type of insecticide. 1:=DDT, 2:=gamma-BHC, 3:=DDT+gamma-BHC
- Dose: The dose of insecticide (in mg/10cm2).
- Trials: The total number of insects.
- Successes: The number of insects killed after a six-day period.

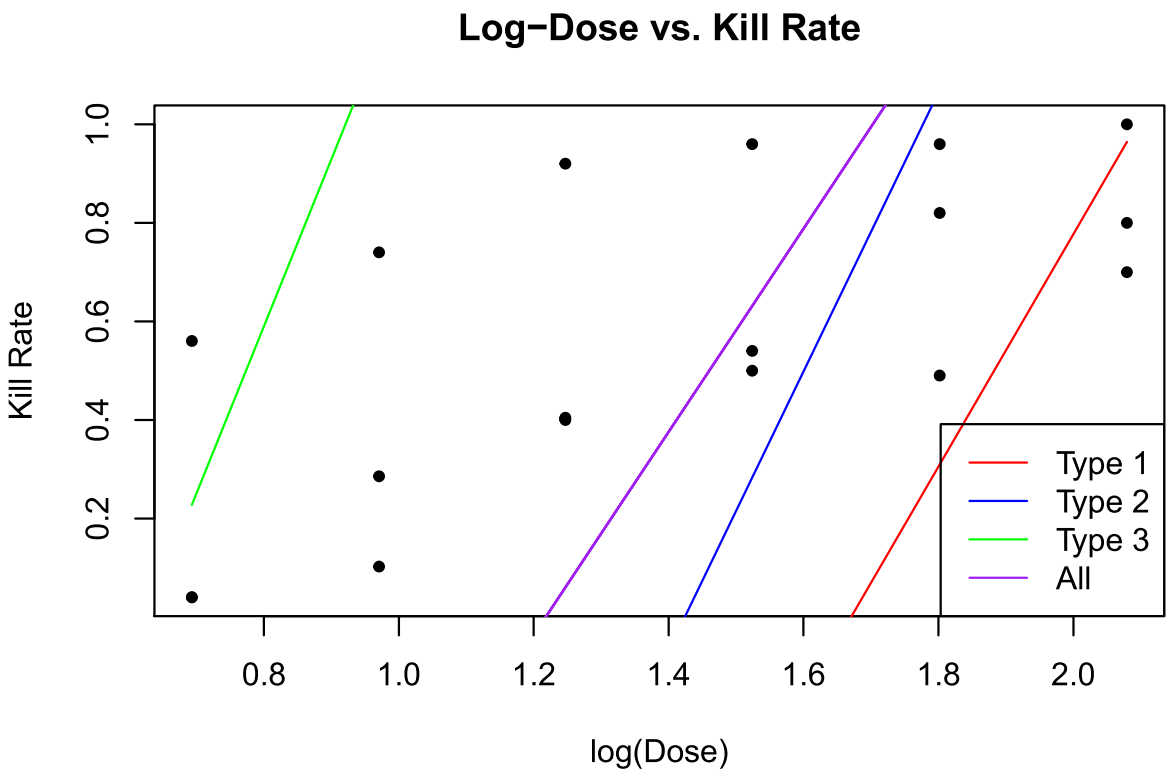
(a)

Investigate graphically the relationship between the dose, either in original units or in log units, and the kill rate. Here, kill rate is the probability an insect is killed.



(b)

On the graph for part (a), plot the linear logistic fitted curve for each of the insecticides plus the combination.



(c)

Consider the two models, one in which the relationship is described by three parallel straight lines in the log dose and and one in which the three lines are straight but not parallel. Assess the evidence against the hypothesis of parallelism.

Parallel lines would indicate a good model fit when considering a proportional odds model, whereas if the lines are straight but not parallel, the slopes are different and this contradicts the proportional odds model fit. Fit1 does not have a significant coefficient for log-dose; Fit2 does not have a significant coefficient; Fit3 does not have a significant coefficient; the fit with all types considered has a significant coefficient at an alpha of .10. Although the lines are ~sort of~ parallel, since none of the individually fitted models have significant coefficients for log-dose, we may not be content with this model fit (proportional odds).

(d)

Let Type be a 3-level factor, and let ldose be the log dose. Explain the relationship between the regression coefficients in the model formulae 'Type + ldose' and 'Type + ldose - 1'.

```
##
## Call:
## glm(formula = killrate ~ Type + ldose, family = binomial(link = "logit"),
##     data = insect)
##
```

```
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.28112  -0.11783  -0.05326   0.14433   0.33696
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -4.4940     2.5147  -1.787   0.0739 .
## Type2         0.6122     1.3879   0.441   0.6591
## Type3         3.0396     1.7643   1.723   0.0849 .
## ldose         2.7245     1.5217   1.790   0.0734 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 8.31082  on 17  degrees of freedom
## Residual deviance: 0.48343  on 14  degrees of freedom
## AIC: 18.27
##
## Number of Fisher Scoring iterations: 5


##
## Call:
## glm(formula = killrate ~ Type + ldose - 1, family = binomial(link = "logit"),
##     data = insect)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.28112  -0.11783  -0.05326   0.14433   0.33696
##
## Coefficients:
##       Estimate Std. Error z value Pr(>|z|)
## Type1   -4.494      2.515  -1.787   0.0739 .
## Type2   -3.882      2.346  -1.654   0.0980 .
## Type3   -1.454      2.011  -0.723   0.4695
## ldose    2.725      1.522   1.790   0.0734 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 8.66580  on 18  degrees of freedom
## Residual deviance: 0.48343  on 14  degrees of freedom
## AIC: 18.27
##
## Number of Fisher Scoring iterations: 5
```

In the second formula, we are omitting the intercept, which forces the model through the origin. We might do this if we are okay with (i.e., expecting to see) a zero, and in our dataset it is possible to have a 0 kill rate. When we leave in the intercept, we retain a degree of freedom because we code just two factors to indicate Type, whereas if we omit the intercept we lose a degree of freedom by adding another covariate as an indicator for the third Type (i.e., the intercept). However, we still have significan predictors in both models.

(e)

On the assumption that three parallel straight lines suffice, estimate the DDT+gamma-BHC dose required to give a 99% kill rate, and obtain a 90% confidence interval for this dose.

While the coefficients from above were not significant, the lines were ~sort of~ parallel, so we can use all of the data rather than just the Type 3 data to find a 90% CI for the designated dose.

The estimated log-dose required for 99% kill rate is 3.44. Since we are holding the assumption that the lines are parallel (sufficient for model assumptions), we use all of the data rather than just the third type of dose. The 90% CI is (1.44, 5.44).

Problem 2

MassSpect contains observed data, Pep contains expected data given P generated the spectrum in MassSpect. Let n = #mass-to-charges in Pep; let Yr = 1 if P's rth mass-to-charge exists in observed spectrum (MassSpect) and Yr = 0 if not for r = 1, . . . , n. Conditional on P having generated observed mass spectrum in MassSpect, assume Yi's are independent.

(a)

Suppose P generated the observed spectrum, and assume logit{E(Yr)} = B0 + B1log(xr), where xr is the expected relative intensity for mass-to-charge r (i.e. the rth entry of the second column in Pep.txt). Given what you know about the problem, what do you expect the sign of B1 to be?

Based on the explanation of the problem and a rough understanding of the data, I would expect B1 to be positive.

(b)

Compute the maximum likelihood estimate for B1.

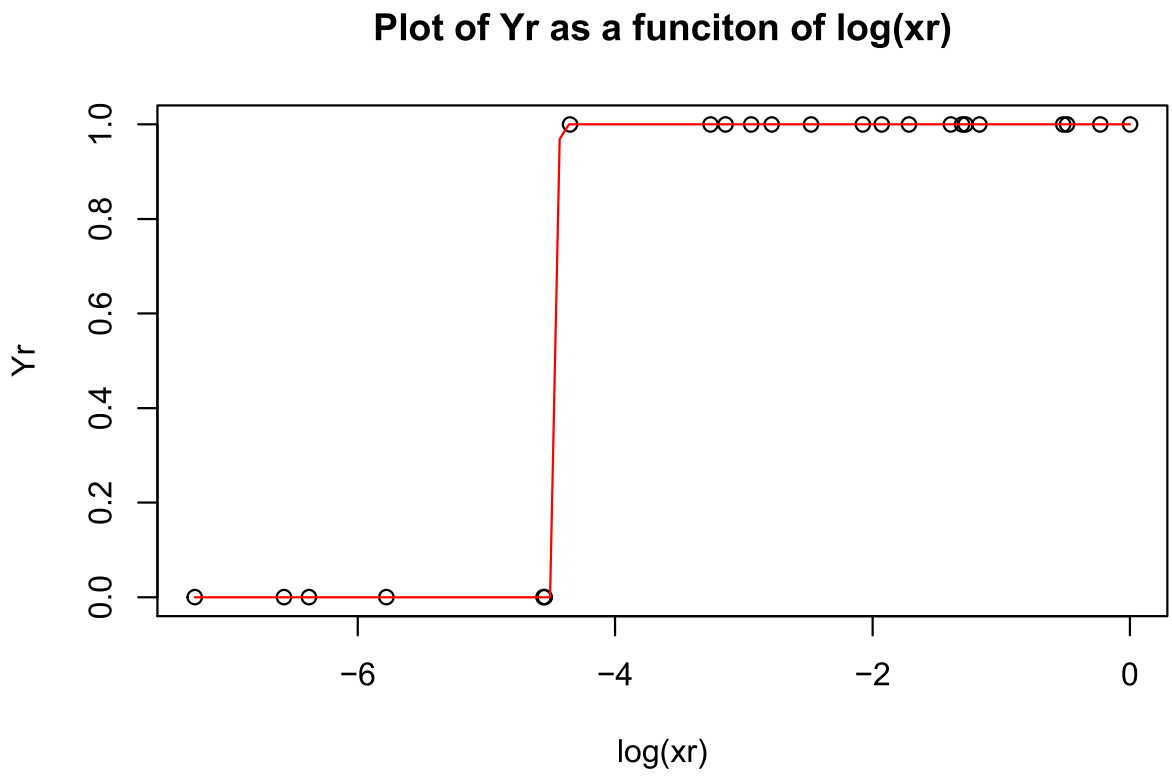
The MLE estimate for B1 is 198.2 with a standard error of 70482.6.

(i) Do you notice anything peculiar about the estimate?

The p-values associated with the coefficients are close to 1, and the coefficient values are extremely high with even more surprisingly high standard errors.

(ii) see attached

(iii) Plot Yr as a function of log(xr). In light of part (ii), are the observations you made in part (i) consistent with the plot? Explain.



Considering Yr can only take on 0 or 1, the plot does look as expected and seems like a good fit for the logit model and the observations from (i) are consistent.

Problem 2

(b) (ii) Observe data $(z_1, w_1), \dots, (z_n, w_n)$
where $z_i \in \{0,1\}$, $w_i \in \mathbb{R}^p$
$(p-1)$-dim hyperplane $\gamma \in \mathbb{R}^p$ that perfectly partitions z_i
→ for $2z_i - 1 \in \{1, -1\}$,
$(2z_i - 1)(w_i^T \gamma) > 0$, $i = 1, \dots, n$

Show that MLE for η under
$z_i \sim \text{Ber}(\pi_i)$, z_i ind.
$\text{logit}(\pi_i) = w_i^T \eta$, $i = 1, \dots, n$.
has infinite form (that $\eta \to \infty$)

$f(z_i) = \pi_i^{z_i} (1 - \pi_i)^{1 - z_i}$

$\text{logit}(\pi_i) = w_i^T \eta$

$\to \pi_i = \frac{\exp(w_i^T \eta)}{1 + \exp(w_i^T \eta)}$

$(2z_i - 1)(w_i^T \eta)$

$W = \{2(\pi_i^{z_i}(1-\pi_i)^{1-z_i}) - 1\} \cdot (w_i^T \eta)$

$= 2(\pi_i^{z_i}(1-\pi_i)^{1-z_i})(w_i^T \eta) - (w_i^T \eta)$

For $(2z_i - 1)(w_i^T \gamma) > 0$, we have
$f(z_i) \to 1$ which requires
$\pi_i = 1$ and therefore we need
$\eta \to \infty$ (infinite) because as $\eta \to \infty$,
$\pi_i \to 1$ and
$\text{logit}(\pi_i) \to 0$

(c) $\beta = (\beta_0, \beta_1)^T$

$\omega_r = (1, \log(x_r))^T \in \mathbb{R}^2$

$\underline{Y} = (Y_1, \ldots, Y_n)^T \in \mathbb{R}^n$

consider:

$Y_r | \pi_r \sim Ber(\pi_r)$, $Y_i | \pi_i$ ind

$\text{logit}(\pi_r) | \beta = \omega_r^T \beta$, $r = 1, \ldots, n$

$\beta \sim N_2(\mu, \Sigma)$

for some known constants $\mu \in \mathbb{R}^2$, $\Sigma \in \mathbb{R}^{2\times 2}$ such that Σ is pd.

(i) Let $pr(\beta | Y_n)$ be pdf of posterior distribution for $\beta | Y_n$

Show that

$$\log\{Pr(\beta | Y_n)\} = l(\beta; Y_n) + g(\beta; \mu, \Sigma) + c(Y_n)$$

$l \rightarrow$ log-likelihood

$g \rightarrow$ function depends only on β & μ, Σ

$c \rightarrow$ function depends only on Y_n

Derive expressions for l, g

(Hint: Bayes' $\rightarrow$ $pr(\beta | Y_n) = \frac{Pr(Y_n|\beta)\, pr(\beta)}{Pr(Y_n)}$

$$\log\{Pr(\beta|Y_n)\} = \log\left(\frac{Pr(Y_n|\beta)\, pr(\beta)}{Pr(Y_n)}\right)$$

$$= \underbrace{\log(Pr(Y_n|\beta))}_{\text{by definition, } = l} + \underbrace{\log(Pr(\beta))}_{= g} - \underbrace{\log(Pr(Y_n))}_{c_1 \text{ only function of } Y_n}$$

$= g$ only function of β

(ii) Show that $\log\{Pr(\beta|Y_n)\}$ is concave when treated as a function of β

say $L = \log\{Pr(\beta|Y_n)\}$

$\frac{\partial L}{\partial \beta} = 0 + g'(\beta; \mu, \Sigma) - 0$

$\partial^2 L / \partial \beta^2 = g''(\beta; \mu, \Sigma) > 0$

So, L is concave

→

(iii) Any similarities between $\log\{Pr(\beta|Y_n)\}$ and Ridge Regression from 2131?

It does seem similar → in ridge regression, we considered a penalty when modeling and subsequently estimating β. → $\hat{\beta}_{ridge} = (X^TX + \lambda I_p)^{-1} X^TY$

which is similar to

$$\log\{Pr(\beta|Y_n)\} = l + g + c$$

g: penalty on Y

l: similar to $\lambda\|\beta\|_2^2$ term

(iv) Instead of shrinking towards 0 like ridge regression, our "shrinkage estimator" for β, $g(\beta|\mu, \Sigma) = \log\{Pr(\beta)\}$, will shrink towards μ.

→

(v)

$$\hat{\beta}^{(MAP)} = \underset{\beta \in \mathbb{R}^2}{\operatorname{argmax}} \log\{pr(\beta|Y_n)\}$$

Show that $\hat{\beta}^{(MAP)}$ exists (ie, $\|\hat{\beta}^{(MAP)}\| < \infty$) and use part (ii) to argue that it is unique.

$$\hat{\beta}^{(MAP)} = \underset{\beta \in \mathbb{R}^2}{\operatorname{argmax}} \frac{pr(\beta|Y_n)}{\frac{Pr(Y_n|\beta)\,Pr(\beta)}{Pr(Y_n)}}$$ by Bayes

and since $Pr(Y_n)$ has nothing to do with β, we can drop it

$$\propto \underset{\beta \in \mathbb{R}^2}{\operatorname{argmax}} Pr(Y_n|\beta)\,Pr(\beta)$$

$$\propto \underset{\beta \in \mathbb{R}^2}{\operatorname{argmax}} \log\{Pr(Y_n|\beta)\} + \log\{pr(\beta)\}$$

$$\propto \underset{\beta \in \mathbb{R}^2}{\operatorname{argmax}} \sum_{Y_i \in Y_n} \log\{Pr(Y_i|\beta)\} + \log\{Pr(\beta)\}$$

$Y_i|\beta \sim Ber(\beta) \rightarrow Pr(Y_i|\beta) = \beta^{Y_i}(1-\beta)^{1-Y_i}$

$$\Sigma \log\{Pr(Y_i|\beta)\} = \Sigma\{Y_i \log\beta + (1-Y_i)\log(1-\beta)\}$$

$$= \Sigma Y_i \log\beta + \Sigma(1-Y_i)\log(1-\beta)$$

derive wrt $\beta \rightarrow \frac{\Sigma Y_i}{\beta} - \frac{2 - \Sigma Y_i}{1-\beta} < \infty$

$\beta \sim N_2(\mu, \Sigma) \rightarrow Pr(\beta) = \frac{1}{\sqrt{2\pi}\Sigma} \exp\{-\frac{1}{2\Sigma}(\beta-\mu)^2\}$

$$\log\{Pr(\beta)\} = -\log(\sqrt{2\pi}) - \log(\Sigma) - \frac{1}{2\Sigma}(\beta-\mu)^2$$

derive wrt $\beta \rightarrow -\frac{1}{2\Sigma} \cdot \Sigma(\beta-\mu) = -\frac{(\beta-\mu)}{\Sigma}$

so $\|\hat{\beta}^{(MAP)}\| < \infty$ and because the second derivative of g (prior) is concave from (ii), $\hat{\beta}^{(MAP)}$ is unique.

→

(vi) Find $\nabla_\beta \log\{pr(\beta|Y_n)\}$ and $\nabla_\beta^2 \log\{pr(\beta|Y_n)\}$, use to derive iterative algorithm to find $\hat{\beta}^{(MAP)}$ that uses Newton-Raphson updates. (ie, find vector function $U: \mathbb{R}^2 \to \mathbb{R}^2$ and matrix function $H: \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ such that an iteration t can be written as

$$\beta^{(t)} = \beta^{(t-1)} + H\{\beta^{(t-1)}\}\, U\{\beta^{(t-1)}\})$$

$\frac{d \log pr(\beta|Y_n)}{d\beta}$

$$\nabla_\beta \log\{pr(\beta|Y_n)\} = \nabla_\beta \log\{Pr(Y_n|\beta)\, Pr(\beta)\}$$
$$= \nabla_\beta \log\{Pr(Y_n|\beta)\} + \nabla_\beta \log\{Pr(\beta)\}$$
$$= \frac{\Sigma Y_i}{\beta} - \frac{2-\Sigma Y_i}{1-\beta} - \frac{(\beta-\mu)}{\Sigma}$$

$\frac{d^2 \log pr(\beta|Y_n)}{d\beta^2}$

$$\nabla_\beta^2 \log\{pr(\beta|Y_n)\} = \nabla_\beta^2 \log\{Pr(Y_n|\beta)\} + \nabla_\beta^2 \log\{Pr(\beta)\}$$
$$= -\frac{\Sigma Y_i}{\beta^2} - \frac{(1-\beta)(0) - (2-\Sigma Y_i)(-1)}{(1-\beta)^2} - \frac{1}{\Sigma}$$
$$= -\frac{\Sigma Y_i}{\beta^2} + \frac{2-\Sigma Y_i}{(1-\beta)^2} - \frac{1}{\Sigma}$$
$$= -\frac{1}{\beta}\cdot\frac{\Sigma Y_i}{\beta} - \frac{1}{1-\beta}\cdot\frac{2-\Sigma Y_i}{(1-\beta)} - \frac{1}{\Sigma}$$
$$= -\frac{1}{\Sigma}\cdot\frac{(\beta)(1-\beta)}{(\beta)(1-\beta)} - \frac{(1-\beta)}{\beta(1-\beta)}\cdot\frac{\Sigma Y_i}{\beta} - \frac{\beta}{\beta(1-\beta)}\cdot\frac{2-\Sigma Y_i}{(1-\beta)}$$
$$= \underbrace{\frac{1}{\beta(1-\beta)}}_{U\{\beta^{(t-1)}\}} \Big\{ \underbrace{-\frac{1}{\Sigma}(\beta)(1-\beta)}_{\beta^{(t-1)}} \underbrace{-(1-\beta)\cdot\frac{\Sigma Y_i}{\beta} - (\beta)\cdot\frac{2-\Sigma Y_i}{1-\beta}}_{H\{\beta^{(t-1)}\}} \Big\}$$

Problem 3

Consider the dataset "Fertility.txt". Your goal is to understand which factors help determine the number of children a woman from Fiji will have in her lifetime. The data matrix consists of 5 columns: - YSFM: A factor variable giving the number of years since a woman's first marriage. The levels are 1 (< 5 years), 2 (5-9 years), 3 (10-14 years), 4 (15-19 years), 5 (20-24 years), 6 (25+ years)
- Place: A factor variable with two levels (Urban and Rural), indicating if the mother lives in a urban or rural area.
- Education: A factor variable giving the mother's education level. The levels are 1 (none), 2 (lower elementary), 3 (upper elementary), 4 (secondary or higher).
- Nwomen: The total number of women sampled.
- Average: The mean number of children born to a woman in that row.

(a)

Fit an appropriate model describing the how the number of children varies with marital age, mother's abode and education. Explain the meaning of all parameters in your model, and comment on the major factors affecting fertility.

```
##
## Call:
## glm(formula = Average ~ YSFM + Place + Education, family = poisson(link = "log"),
##     data = fert)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5176  -0.2105  -0.0239   0.1725   1.1070
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.04765    0.39350   0.121 0.903625
## YSFM2        0.99525    0.43022   2.313 0.020702 *
## YSFM3        1.39972    0.41045   3.410 0.000649 ***
## YSFM4        1.57115    0.40400   3.889 0.000101 ***
## YSFM5        1.75061    0.39825   4.396 1.10e-05 ***
## YSFM6        1.75178    0.39822   4.399 1.09e-05 ***
## PlaceUrban   0.01840    0.14979   0.123 0.902251
## Education2   0.04072    0.19602   0.208 0.835441
## Education3  -0.06270    0.20119  -0.312 0.755317
## Education4  -0.66436    0.24021  -2.766 0.005679 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 68.906  on 47  degrees of freedom
## Residual deviance: 18.739  on 38  degrees of freedom
## AIC: Inf
##
## Number of Fisher Scoring iterations: 5
```

Because we are looking at count data as a dependent variable, we'll use a poisson glm model fit. YSFM is a factor variable with 6 variables. The model uses the first level as a baseline, and codes 5 others as

indicators for 1 if that level and all others are 0. If all 5 are 0, the intercept serves as the indicator for level 1. Place is a factor variable with Urban as 1 and Rural as 0. Education acts the same way as YSFM, with 4 different levels and if all are 0, the intercept acts as the indicator for level 1. The predictors with significant p-values (at 0.05) are YSFM3, YSFM4, YSFM5, YSFM6, and Education4. The intercept B0 is the log of the expected value of average children when all covariates are zero (or in this case, when a woman has level 1 YSFM, rural living, and level 1 Education). Exponential of the sum of all covariates is the expected value of average children for the covariates. Because we are using a log link, each of these coefficients represents the effect on the log of the average children a woman will have, not on the number of children itself.

(b)

Construct a 95% confidence interval for the mean number of children born to an urban woman with upper elementary education after ten years of marriage.

The 95% CI for mean number of children (4.07) with these levels is (2.58, 6.42).

(c)

Estimate the lifetime average number of children born to rural women with secondary education. Give 90% confidence limits.

Average children will be 1.13 children, with a confidence interval of (.73, 1.54). Since the highest level of YSFM has the most years since marriage, it is a good estimate for lifetime average across all women, based on the information provided.

```
## ----setup, include = FALSE, warning = FALSE, purl = TRUE----------------------------------------------------------------------
knitr::opts_chunk$set(echo = FALSE)
setwd("C:/Users/orlyo/OneDrive/Desktop/Grad School/Spring 2021/2. STAT 2132 - Applied Stat. Methods 2/Exams/Midterm")
# Problem 1
insect = read.csv("Insecticide.csv")
insect$killrate = insect$Successes/insect$Trials
insect$Type = as.factor(insect$Type)
insect$ldose = log(insect$Dose)
# Problem 2
fert = read.csv("Fertility.csv")
fert$YSFM = as.factor(fert$YSFM)
fert$Place = as.factor(fert$Place)
fert$Education = as.factor(fert$Education)
# Problem 3
spec = read.csv("MassSpec.csv")
pep = read.csv("Pep.csv")
pep$Yr = as.numeric(pep$Exp_m.to.z %in% spec$Obs_m.to.z)
pep$lexp = log(pep$Exp_Int)


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------
# plot(insect$Dose, insect$killrate, main = "Dose vs. Kill Rate", xlab = "Dose", ylab = "Kill Rate")
plot(insect$ldose, insect$killrate, main = "Log-Dose vs. Kill Rate", xlab = "log(Dose)", ylab = "Kill Rate", pch = 20)


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------
fit1 = glm(killrate ~ ldose, family = binomial(link = "logit"), data = insect[insect$Type == 1,])
fit2 = glm(killrate ~ ldose, family = binomial(link = "logit"), data = insect[insect$Type == 2,])
fit3 = glm(killrate ~ ldose, family = binomial(link = "logit"), data = insect[insect$Type == 3,])
fit_all = glm(killrate ~ ldose, family = binomial(link = "logit"), data = insect)
# summary(fit1)
# summary(fit2)
# summary(fit3)
# summary(fit_all)

# lines(insect$ldose[insect$Type == 1], fit1$fitted.values, col = "red", pch = 20)
# lines(insect$ldose[insect$Type == 2], fit2$fitted.values, col = "blue", pch = 20)
# lines(insect$ldose[insect$Type == 3], fit3$fitted.values, col = "green", pch = 20)
# lines(insect$ldose, fit_all$fitted.values, col = "purple", pch = 20)
# legend("bottomright", legend = c("Type 1", "Type 2", "Type 3", "All"), col = c("red", "blue", "green", "purple"), lty = 1)
```

```
plot(insect$ldose, insect$killrate, main = "Log-Dose vs. Kill Rate", xlab = "log(Dose)", ylab = "Kill Rate",
pch = 20)
lines(insect$ldose[insect$Type == 1], log(fit1$fitted.values/(1 - fit1$fitted.values)), col = "red", pch = 20)
lines(insect$ldose[insect$Type == 2], log(fit2$fitted.values/(1 - fit2$fitted.values)), col = "blue", pch = 20)
lines(insect$ldose[insect$Type == 3], log(fit3$fitted.values/(1 - fit3$fitted.values)), col = "green", pch =
20)
lines(insect$ldose, log(fit_all$fitted.values/(1 - fit_all$fitted.values)), col = "purple", pch = 20)
legend("bottomright", legend = c("Type 1", "Type 2", "Type 3", "All"), col = c("red", "blue", "green",
"purple"), lty = 1)


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------------
form1 = glm(killrate ~ Type + ldose, data = insect, family = binomial(link = "logit"))
summary(form1)
form2 = glm(killrate ~ Type + ldose - 1, data = insect, family = binomial(link = "logit"))
summary(form2)


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------------
# fit3 = glm(killrate ~ ldose, family = binomial(link = "logit"), data = insect[insect$Type == 3,])
fit_all = glm(killrate ~ ldose, family = binomial(link = "logit"), data = insect)
log.dose = (log(.99/(1-.99)) - summary(fit_all)$coefficients[1,1])/summary(fit_all)$coefficients[2,1]
# log.dose

lo = log.dose - 1.65*summary(fit_all)$coefficients[2, 2]
hi = log.dose + 1.65*summary(fit_all)$coefficients[2, 2]
ci = c(lo, hi)
# log.dose; ci


## ---- include = FALSE, warning = FALSE----------------------------------------------------------------------------------------------
par(mfrow = c(1, 2))
plot(spec$Obs_m.to.z, spec$Obs_Int, type = "h", main = "observed", xlab = "mass/charge", ylab =
"Relative intensity")
plot(pep$Exp_m.to.z, pep$Exp_Int, type = "h", main = "expected", xlab = "mass/charge", ylab = "relative
intensity")


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------------
n = sum(pep$Yr)
model = glm(Yr ~ log(Exp_Int), data = pep, family = binomial(link = "logit"))
# summary(model)
```

```
## ---- warning = FALSE----------------------------------------------------------------------------------------------------------------
par(mfrow = c(1, 1))
plot(pep$lexp, pep$Yr, main = "Plot of Yr as a funciton of log(xr)", xlab = "log(xr)", ylab = "Yr")
g = glm(Yr ~ lexp, family = binomial, pep)
curve(predict(g, data.frame(lexp = x), type = "resp"), add = TRUE, col = "red")


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------------
model1 = glm(Average ~ YSFM + Place + Education, data = fert, family = poisson(link = "log"))
summary(model1)


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------------
# exp(confint(model1))
log.mean = 0.04765 + 1.399525 + 0.01840 - 0.06270
new.mean = exp(log.mean)

new.data = data.frame(YSFM = "3", Place = "Urban", Education = "3")
pred = predict(model1, newdata = new.data, se.fit = TRUE, interval = "confidence")

upper = pred$fit + 1.96*pred$se.fit
lower = pred$fit - 1.96*pred$se.fit
interval = c(exp(lower), exp(upper))
# interval


## ---- warning = FALSE----------------------------------------------------------------------------------------------------------------
new = data.frame(YSFM = "6", Place = "Rural", Education = "4")
pred = predict(model1, newdata = new, se.fit = TRUE, interval = "confidence")
ci = c(pred$fit - 1.65*pred$se, pred$fit + 1.65*pred$se)
```