Package ‘GLMpack’

October 12, 2022

Type Package

Title Data and Code to Accompany Generalized Linear Models, 2nd Edition

Version 0.1.0

Description Contains all the data and functions used in Generalized Linear Models, 2nd edition, by Jeff Gill and Michelle Torres. Examples to create all models, tables, and plots are included for each data set.

License GPL (>= 3)

Depends R (>= 2.10)

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

Imports MASS, pBrackets, nnet, effects, AER, pscl, foreign, Matrix, lme4, lmtest, sandwich, censReg, plm

NeedsCompilation no

Author Jeff Gill [aut],
Michelle Torres [aut, cre],
Simon Heuberger [aut]

Maintainer Michelle Torres <smtorres@wustl.edu>

Repository CRAN

Date/Publication 2019-07-19 09:40:05 UTC

R topics documented:

---

africa *Data on inflation increase in Africa*

Description

Data for the Africa example used in chapter 7

Usage

```
data(africa)
```

Format

A data frame with 47 rows and 7 variables:

INFLATN Inflation rates

DICTATOR Number of years of personal dictatorship that occurred from independence to 1989

SIZE Area at the end of the period, in thousand square kilometers

GROWTH Average annual gross national product (GNP) rate of growth in percent from 1965 to 1989

CHURMED Number of church-operated hospitals and medical clinics as of 1973

CONSTIT the constitutional structure when not a dictatorship in ascending centrality (0 = monarchy, 1 = presidential, 2 = presidential/parliamentary mix, 3 = parliamentary)

REPRESS Violence and threats of violence by the government against opposition political activity from 1990 to 1994 ...

Examples

```
data(africa)
attach(africa)
library(lmtest)
library(plm)

## Table 7.4
y <- (INFLATN/100)[-16]
y[y > 1] <- 1
X <- cbind(DICTATOR,SIZE,GROWTH,CHURMED,CONSTIT,REPRESS)[-16,]
```

```
X[,4] <- log(X[,4]+.01)
africa.glm <- glm(y ~ X[,1:6], family=quasibinomial('logit'))
out.mat <- coeftest(africa.glm, vcov.=vcovHC(africa.glm, type="HC0"))
( out.mat <- round(cbind(out.mat[,1:2], out.mat[,1] - 1.96*out.mat[,2],
                         out.mat[,2] + 1.96*out.mat[,2]),4) )
```

---

campaign *Data on campaign contributions in California in the 2014 House elections*

Description

Data for the campaign contributions example used in chapter 6

Usage

```
data(campaign)
```

Format

A data frame with 180 rows and 16 variables:

DTRCT California district

CANDID FEC ID

CYCLE Election cycle

NAME Name of the candidate

INCUMCHALL Incumbency status

CFSCORE CFscore

CANDGENDER Gender of the candidate

PARTY Party of the candidate

TOTCONTR Contributions to the 2014 electoral campaigns in the 53 districts of California in the U.S. House of Representatives

TOTPOP Total state population

FEMALE Number of female citizens in the state

WHITE Number of white citizens in the state

HISP Number of Hispanic citizens in the state

FEMALEPCT Percentage of female citizens in the state

WHITEPCT Percentage of white citizens in the state

HISPPCT Percentage of Hispanic citizens in the state ...

Examples

```
data(campaign)
attach(campaign)
library(pBrackets)

## Gamma model
cmpgn.out <- glm(TOTCONTR ~ CANDGENDER + PARTY + INCUMCHALL + HISPPCT,
             family=Gamma(link = 'log'), data=campaign)

## Linear model
cmpgn.out_lm <- lm(TOTCONTR ~ CANDGENDER + PARTY + INCUMCHALL + HISPPCT, data=campaign)

## Table 6.8
round(glm.summary(cmpgn.out),4)
cmpgn.out$null.deviance
cmpgn.out$df.null
cmpgn.out$deviance
cmpgn.out$df.residual
logLik(cmpgn.out)
cmpgn.out$aic

## Table 6.9
summary(cmpgn.out_lm)
confint(cmpgn.out_lm)

## Figure 6.4
opar = par(mfrow=c(1,1), mar=c(5.1,4.1,4.1,2.1), oma=c(0,0,0,0))
par(mar=c(4,3,3,0),oma=c(1,1,1,1))
hist(campaign$TOTCONTR,xlab="",ylab="", yaxt="n", xaxt="n",
     xlim=c(0,9000), ylim=c(0, 130), main="",
     col = "gray40")
axis(1, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=1, lwd.ticks = 1, las=2)
title(xlab = 'Total campaign contributions (thousands of dollars)',
      ylab= "Frequency",
      line = 1.7, cex.lab=1)
title(line = 1, main="Distribution of campaign contributions", font.main=1)
par(opar)

## Figure 6.5
campaign.mu <- predict(cmpgn.out_lm)
campaign.y <- campaign$TOTCONTR
par(mfrow=c(1,3), mar=c(3,3,2,1),oma=c(1,1,1,1))
plot(campaign.mu,campaign.y,xlab="",ylab="", yaxt='n', xaxt="n", pch="+")
axis(1, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = "Fitted values", ylab="Observed values",
      line = 1.7, cex.lab=1.3)
title(main="Model Fit Plot",
      line = 1, cex.main=1.7, font.main=1)
abline(lm(campaign.y~campaign.mu)$coefficients, lwd=2)
plot(fitted(cmpgn.out_lm),resid(cmpgn.out_lm,type="pearson"),xlab="",ylab="",
```

```
     yaxt='n', xaxt="n", pch="+")
axis(1, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = "Fitted values", ylab="Pearson Residuals",
      line = 1.7, cex.lab=1.3)
title(main="Residual Dependence Plot",
      line = 1, cex.main=1.7, font.main=1)
abline(0,0, lwd=2)
plot(cmpgn.out_lm,which=2, pch="+",
     sub.caption = "", caption = "", mgp=c(1.5, 0.3, 0),
     tck=0.02, cex.axis=0.9, cex.lab=1.3, lty=1)
title(main="Normal-Quantile Plot",
      line = 1, cex.main=1.7, font.main=1)
par(opar)

## Figure 6.6
mygray = rgb(153, 153, 153, alpha = 200, maxColorValue = 255)
newdat_gender <- data.frame(CANDGENDER = c('F','M'), PARTY= rep('Democrat',2),
                          INCUMCHALL=rep("C", 2), HISPPCT=rep(mean(campaign$HISPPCT),2))
newdat_party <- data.frame(CANDGENDER = rep('M', 3), PARTY= c('Democrat','Republican',
                           'Independent'), INCUMCHALL=rep("C", 3),
                           HISPPCT=rep(mean(campaign$HISPPCT),3))
newdat_incumchall <- data.frame(CANDGENDER = rep('M', 3), PARTY= rep('Democrat',3),
                                INCUMCHALL=c('C', 'I', 'O'),
                                HISPPCT=rep(mean(campaign$HISPPCT),3))
newdat_hisiq <- data.frame(CANDGENDER = rep('M', 2), PARTY= rep('Democrat',2),
                           INCUMCHALL=rep("C", 2),
                           HISPPCT=as.numeric(summary(campaign$HISPPCT)[c(2,5)]))
newdat_hispf <- data.frame(CANDGENDER = rep('M', 200), PARTY= rep('Democrat',200),
                         INCUMCHALL=rep("C", 200), HISPPCT=seq(.1, .9, length.out = 200))
preds_gender <- predict(cmpgn.out, newdata = newdat_gender, se.fit = TRUE)
preds_party <- predict(cmpgn.out, newdata = newdat_party, se.fit = TRUE)
preds_incumchall <- predict(cmpgn.out, newdata = newdat_incumchall, se.fit = TRUE)
preds_hispiq <- predict(cmpgn.out, newdata = newdat_hisiq, se.fit = TRUE)
preds_hispf <- predict(cmpgn.out, newdata = newdat_hispf, se.fit = TRUE)
cis_gender <- round(glm.cis(preds_gender$fit, preds_gender$se.fit, 0.95,cmpgn.out$df.residual),4)
cis_party <- round(glm.cis(preds_party$fit, preds_party$se.fit, 0.95,cmpgn.out$df.residual),4)
cis_incumchall <- round(glm.cis(preds_incumchall$fit, preds_incumchall$se.fit, 0.95,
                                cmpgn.out$df.residual),4)
cis_hispiq <- round(glm.cis(preds_hispiq$fit, preds_hispiq$se.fit, 0.95,cmpgn.out$df.residual),4)
cis_hispf <- round(glm.cis(preds_hispf$fit, preds_hispf$se.fit, 0.95,cmpgn.out$df.residual),4)
iqrange = cbind(summary(campaign$HISPPCT)[c(2,5)],cis_hispiq[2,4] - cis_hispf[1,4])
par(mfrow=c(2,2), mar=c(4,3,3,0),oma=c(1,1,1,1))
plot(1:2, cis_gender[,4], type="n",xlab="",ylab="",  yaxt="n", xaxt="n",
     xlim=c(0,3), ylim=c(100, 700))
segments(1:2, cis_gender[,5], 1:2, cis_gender[,6], lwd=2, col="gray60")
points(1:2, cis_gender[,4], pch=16, cex=2.5)
text(1:2, cis_gender[,4], labels = c("F", "M"), col="white", cex=0.9)
segments(1.05, cis_gender[1,4], 1.95, cis_gender[2,4], lty=2)
brackets(1, cis_gender[1,4]+20, 2, cis_gender[1,4]+20, h = 45,  ticks = 0.5, lwd=2)
text(1.5, cis_gender[1,4]+100, bquote(hat(y)['F']-hat(y)['M'] ~ '='
     ~ .(cis_gender[2,4]-cis_gender[1,4])), cex=0.9)
axis(1, at=1:2, labels = c("Female", "Male"), tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0),
```

```
     lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = 'Gender of candidate',
      ylab="Total campaign contributions",
      line = 1.7, cex.lab=1)
title(line = 1, main="Gender of candidate", font.main=3)
plot(1:3, cis_party[,4], type="n",xlab="",ylab="",  yaxt="n", xaxt="n",
     xlim=c(0.5,3.5), ylim=c(0, 600))
segments(1:3, cis_party[,5], 1:3, cis_party[,6], lwd=2, col="gray60")
points(1:3, cis_party[,4], pch=15:17, cex=2.5)
text(1:3, cis_party[,4], labels = c("D", "R", "I"), col="white", cex=0.8)
segments(c(1.05,2.05), cis_party[1:2,4], c(1.95,2.95), cis_party[2:3,4], lty=2)
brackets(1, cis_party[2,4]+20, 2, cis_party[2,4]+20, h = 45,  ticks = 0.5, lwd=2)
brackets(3, cis_party[3,4]+20, 2, cis_party[3,4]+20, h = 45,  ticks = 0.5, lwd=2)
text(1.5, cis_party[1,4]+160, bquote(hat(y)['R']-hat(y)['D'] ~ '='
     ~ .(cis_party[2,4]-cis_party[1,4])), cex=0.9)
text(2.5, cis_party[3,4]-40, bquote(hat(y)['I']-hat(y)['R'] ~ '='
     ~ .(cis_party[3,4]-cis_party[2,4])), cex=0.9)
axis(1, at=1:3, labels = c("Democrat", "Republican", "Independent"), tck=0.03, cex.axis=0.9,
     mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = 'Party of candidate',
      ylab="Total campaign contributions",
      line = 1.7, cex.lab=1)
title(line = 1, main="Party of candidate", font.main=3)
plot(1:3, cis_incumchall[,4], type="n",xlab="",ylab="",  yaxt="n", xaxt="n",
     xlim=c(0.5,3.5), ylim=c(0, 3200))
segments(1:3, cis_incumchall[,5], 1:3, cis_incumchall[,6], lwd=2, col="gray60")
points(1:3, cis_incumchall[,4], pch=15:17, cex=1.5)
segments(c(1.05,2.05), cis_incumchall[1:2,4], c(1.95,2.95), cis_incumchall[2:3,4], lty=2)
brackets(1, cis_incumchall[2,4]+20, 2, cis_incumchall[2,4]+20, h = 105,  ticks = 0.5, lwd=2)
brackets(3, cis_incumchall[3,4]+20, 2, cis_incumchall[3,4]+20, h = 105,  ticks = 0.5, lwd=2)
text(1.5, cis_incumchall[2,4]+285, bquote(hat(y)['I']-hat(y)['C'] ~ '='
     ~ .(cis_incumchall[2,4]-cis_incumchall[1,4])), cex=0.9)
text(2.5, cis_incumchall[3,4]-200, bquote(hat(y)['O']-hat(y)['I'] ~ '='
     ~ .(cis_incumchall[3,4]-cis_incumchall[2,4])), cex=0.9)
axis(1, at=1:3, labels = c("Challenger", "Incumbent", "Open seat"), tck=0.03, cex.axis=0.9,
     mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = 'Status of candidate',
      ylab="Total campaign contributions",
      line = 1.7, cex.lab=1)
title(line = 1, main="Status of candidate", font.main=3)
plot(newdat_hispf$HISPPCT, cis_hispf[,4], type="n",xlab="",ylab="",  yaxt="n", xaxt="n",
     ylim = c(0,1100))
polygon(x = c(newdat_hispf$HISPPCT, rev(newdat_hispf$HISPPCT)),
        y = c(cis_hispf[,5], rev(cis_hispf[,6])), col = mygray, border = NA)
lines(newdat_hispf$HISPPCT, cis_hispf[,4], lwd=2)
rug(campaign$HISPPCT)
segments(iqrange[,1], cis_hispiq[,4], iqrange[,1], rep(500,2), lty=2)
segments(iqrange[1,1], 500, iqrange[2,1], 500, lty = 2)
brackets(iqrange[1,1], 510, iqrange[2,1], 510, h = 75,  ticks = 0.5, lwd=2)
text(abs((iqrange[2,1]-iqrange[1,1])/2)+iqrange[1,1], 450, 'Interquartile range', cex=0.8)
```

```
text(iqrange[,1], cis_hispiq[,4]-50, round(iqrange[,1],3), cex=0.8)
text(abs((iqrange[2,1]-iqrange[1,1])/2)+iqrange[1,1], 655,
          labels=bquote(hat(y)['Q3']-hat(y)['Q1'] ~ '=' ~ .(iqrange[1,2])), cex=0.9)
axis(1, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = '% Hispanic population in Congressional District',
      ylab="Total campaign contributions",
      line = 1.7, cex.lab=1)
title(line = 1, main="Hispanic constituency", font.main=3)
par(opar)
```

---

committee *Data on bills assigned to House committees in the 103rd and 104th Houses of Representatives*

Description

Data for the committees example used in chapters 6

Usage

```
data(committee)
```

Format

A data frame with 20 rows and 6 variables:

SIZE Number of members on the committee

SUBS Number of subcommittees

STAFF Number of staff assigned to the committee

PRESTIGE Dummy variable indicating whether or not it is a high-prestige committee

BILLS103 Number of bills in the 103rd House

BILLS104 Number of bills in the 104th House ...

Examples

```
data(committee)
attach(committee)
library(AER)
library(MASS)
library(pscl)

## Table 6.6
committee

## Table 6.7
committee.out <- glm.nb(BILLS104 ~ SIZE + SUBS * (log(STAFF)) + PRESTIGE + BILLS103)
summary.glm(committee.out)
```

```
data.frame(cbind(round(cbind(summary(committee.out)$coef[,1:2], confint(committee.out)),4)[,1],
       round(cbind(summary(committee.out)$coef[,1:2], confint(committee.out)),4)[,2],
       round(cbind(summary(committee.out)$coef[,1:2], confint(committee.out)),4)[,3],
       round(cbind(summary(committee.out)$coef[,1:2], confint(committee.out)),4)[,4]))

## Figure 6.3
z.matrix <- matrix(0,200,200)
for(i in 1:200)  {
        for(j in 1:200)  {
                if(j < 70)    z.matrix[i,j] <- 1
                if(j < 40)    z.matrix[i,j] <- 2
                if(j < 10)    z.matrix[i,j] <- 3
                if(j == 1)    z.matrix[i,j] <- 3.001
                if(j > 130)   z.matrix[i,j] <- 1
                if(j > 160)   z.matrix[i,j] <- 2
                if(j > 190)   z.matrix[i,j] <- 3
                if(j == 200)  z.matrix[i,j] <- 3.001
        }
}
pears <- resid(committee.out,type="pearson")
devs <- resid(committee.out,type="deviance")
x = seq(-2000,2000,length=200)
opar = par(mfrow=c(1,1), mar=c(5.1,4.1,4.1,2.1), oma=c(0,0,0,0))
layout(matrix(c(1,2), ncol = 1), heights = c(0.9,0.1))
par(mar=c(3,4,2,4),oma=c(2,2,1,3))
image(seq(0,51,length=200), seq(-2000,2000,length=200),z.matrix,xlim=c(0,51),ylim=c(-2000,2000),
xaxt="n",yaxt="n",xlab="",ylab="", col=rev(c("white", "gray40", "gray60", "gray80")))
points(seq(1,50,length=20),(2000/3)*pears[order(BILLS104)],pch=15)
lines(seq(1,50,length=20),(2000/3)*devs[order(BILLS104)],type="h")
abline(0,0, lwd=2)
abline(h=c((x[10]+x[9])/2,(x[40]+x[39])/2,(x[70]+x[69])/2,(x[130]+x[131])/2,
           (x[160]+x[161])/2,(x[191]+x[190])/2), lty=2)
title(xlab = "Order of Fitted Outcome Variable", ylab="Residual Effect",
      line = 1.3, cex.lab=1.3)
title(main="Model Fit Plot",
      line = 1, cex.main=1.7, font.main=1)
par(mar=c(0,1.5,1,1))
plot(0,0, type="n", axes = FALSE, xlab = "", ylab = "")
legend("center", ncol = 2,
       legend = c("Pearson", "Deviances"),
       cex=1, lty=c(0,1), pch = c(15,NA))
par(opar)
```

---

corruption *Data on censored corruption scale*

Description

Data for the corruption example used in chapter 7

Usage

```
data(corruption)
```

Format

A data frame with 83 rows and 7 variables:

ticpi85b Country-level compilation of effects into a 0 to 10 scale of increasing government corruption with an adjustment that modifies this range slightly

MSO Binary variable indicating whether the government owns a majority of key industries

LOG.PC.GDP Log of the average per capita GDP from 1975 to 1983

DEMOCRACY Polity IV democracy score from 1975 to 1983

GOVGDT Average government spending as a percentage of GDP from 1980 to 1983

ECONFREE Index of the ability of capitalists to invest and move money

FEDERAL Binary variable indicating whether the government has a federal system during this period ...

Examples

```
data(corruption)
attach(corruption)
library(censReg)

## Table 7.5
corruption.tobit.out <- censReg( ticpi85b ~ MSO + LOG.PC.GDP + DEMOCRACY + GOVGDT +
    ECONFREE*FEDERAL,
   left = 2.6, right = 10.70, data=corruption, start = NULL, nGHQ = 8, logLikOnly = FALSE)
summary(corruption.tobit.out)
summary(corruption.tobit.out)$nObs
Coefs <- summary(corruption.tobit.out)$estimate[,1]
SEs <- summary(corruption.tobit.out)$estimate[,2]
CI.low <- Coefs - SEs*1.96
CI.hi <- Coefs + SEs*1.96
( round(out.table <- cbind( Coefs, SEs, CI.low, CI.hi),4) )
```

---

dp *Data on capital punishment*

Description

Data for the capital punishment example used in chapters 4, 5, and 6

Usage

```
data(dp)
```

Format

A data frame with 17 rows and 7 variables:

EXECUTIONS The number of times that capital punishment is implemented on a state level in the United States for the year 1997

INCOME Median per capita income in dollars

PERPOVERTY Percent of the population classified as living in poverty

PERBLACK Percent of Black citizens in the population

VC100k96 Rate of violent crimes per 100,000 residents for the year before (1996)

SOUTH Dummy variable to indicate whether the state is in the South

PROPDEGREE Proportion of the population with a college degree ...

Examples

```
opar = par(mfrow=c(1,1), mar=c(5.1,4.1,4.1,2.1), oma=c(0,0,0,0))
data(dp)
attach(dp)

## Table 4.2
dp

## Table 5.1
dp.out <- glm(EXECUTIONS ~ INCOME+PERPOVERTY+PERBLACK+log(VC100k96)+
              SOUTH+PROPDEGREE, family=poisson)
dp.cis <- round(glm.summary(dp.out, alpha = 0.05),4)
round(cbind(summary(dp.out)$coef[,1:2], dp.cis),4)
round(dp.out$null.deviance,4);round(dp.out$df.null,4)
round(dp.out$deviance,4);round(dp.out$df.residual,4)
round(logLik(dp.out),4)
round(dp.out$aic,4)
round(vcov(dp.out),4) # variance covariance matrix

## Table 5.2
k <- 200
b5 <- seq(0.1, 5.4,length=k)
w <- rep(0,k)
for(i in 1:k){
  mm <- glm(EXECUTIONS ~ INCOME+PERPOVERTY+PERBLACK+log(VC100k96)+PROPDEGREE,
            offset=b5[i]*SOUTH,family=poisson)
  w[i] <- logLik(mm)
  }

f <- function(b5,x,y,maxloglik){
  mm <- glm(EXECUTIONS ~ INCOME+PERPOVERTY+PERBLACK+log(VC100k96)+PROPDEGREE,
            offset=b5*x,family=poisson)
  logLik(mm) - maxloglik + qchisq(.95,1)/2
  }
low.pll <- uniroot(f,interval=c(1.5,2), x=SOUTH, y=EXECUTIONS, maxloglik=logLik(dp.out))$root
high.pll <- uniroot(f,interval=c(3,4), x=SOUTH, y=EXECUTIONS, maxloglik=logLik(dp.out))$root
```

```
w[which.min(abs(w-low.pll))]
round(c(low.pll, high.pll),4)
cbind(round(dp.cis[,3:4],4),
      round(confint(dp.out),4))

## Table 6.2
resp <- resid(dp.out,type="response")
pears <- resid(dp.out,type="pearson")
working <- resid(dp.out,type="working")
devs <- resid(dp.out,type="deviance")
dp.mat <- cbind(rep(1,nrow(dp)), as.matrix(dp[,2:4]), as.matrix(log(dp[,5])),
                as.matrix(dp[,6]), as.matrix(dp[,7]))
dp.resid.mat <- cbind(resp,pears,working,devs)
dimnames(dp.resid.mat)[[2]] <- c("response","pearson","working","deviance")
dimnames(dp.resid.mat)[[1]] <- rownames(dp)
dp.resid.mat2 <- round(dp.resid.mat,4)
resid.df <- data.frame(cbind(dp.resid.mat2[,1], dp.resid.mat2[,2],
      dp.resid.mat2[,3], dp.resid.mat2[,4]))
colnames(resid.df) <- dimnames(dp.resid.mat)[[2]]
resid.df

## Figure 5.1
dp.mat.0 <- cbind(dp.mat[,1:5],rep(0,length=nrow(dp.mat)),dp.mat[,7])
dimnames(dp.mat.0)[[2]] <- names(dp.out$coefficients)
dp.mat.1 <- cbind(dp.mat[,1:5],rep(1,length=nrow(dp.mat)),dp.mat[,7])
dimnames(dp.mat.1)[[2]] <- names(dp.out$coefficients)
tcks = list(seq(0,140,20), seq(0,12,2), seq(0,30,5), seq(0,10,2), seq(0,30,5))
layout(matrix(c(1,1,2,2,3,3,4,4,5,6,6,7,8,8,8,8), ncol=4, byrow = TRUE),
       heights = c(0.3,0.3,0.3,0.1))
par(mar=c(3,3,2,4),oma=c(2,1,1,3))
for (i in 2:(ncol(dp.mat.0)-1))  {
  j = i-1
  if (i==6){
    i <- i+1
    plot(0,0, type = "n", axes=FALSE, xlab = "", ylab="")
  }
  ruler <- seq(min(dp.mat.0[,i]),max(dp.mat.0[,i]),length=1000)
  xbeta0 <- exp(dp.out$coefficients[-i]%*%apply(dp.mat.0[,-i],2,mean)
                + dp.out$coefficients[i]*ruler)
  xbeta1 <- exp(dp.out$coefficients[-i]%*%apply(dp.mat.1[,-i],2,mean)
                + dp.out$coefficients[i]*ruler)
  plot(ruler,xbeta0,type="l", xlab="",ylab="", yaxt="n", xaxt="n",
       ylim=c(min(xbeta0,xbeta1)-2,max(xbeta0,xbeta1)), lwd=3)
  lines(ruler,xbeta1,lty=4, lwd=2)
  axis(1, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
  axis(2, at=tcks[[j]],
       tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
 title(xlab = paste("Levels of",dimnames(dp.mat.0)[[2]][i]), ylab="Expected executions",
        line = 1.7, cex.lab=1.2)
}
plot(0,0, type = "n", axes=FALSE, xlab = "", ylab="")
par(mar=c(0,1.5,1,1))
plot(0,0, type="n", axes = FALSE, xlab = "", ylab = "")
```

```
legend("center", ncol = 2,
       legend = c("South State", "Non-South State"),
       cex=1.1, lty=c(2,1), bty="n", lwd=c(2,3))
par(opar)

## Figure 5.2
par(mar=c(3,3,1,0),oma=c(1,1,1,1))
plot(b5,w,type="l",lwd=3, xaxt="n", yaxt="n", xlab="", ylab="")
abline(h=logLik(dp.out)-qchisq(.95,1)/2,lty=3, col="gray40")
segments(dp.cis[6,3], -45, dp.cis[6,4], -45, lty=6, col="black", lwd=2)
segments(dp.cis[6,3:4], c(-45,-45), dp.cis[6,3:4], c(-55,-55), lty=3, col="gray40")
text(3.5, y=-45, "Wald Test", cex=0.9)
segments(low.pll, -42.5, high.pll, -42.5, lty=2, lwd=2, col="black")
segments(c(low.pll, high.pll), c(-55,-55), c(low.pll, high.pll),
         rep(logLik(dp.out)-qchisq(.95,1)/2,2), lty=3, col="gray40")
text(3.25, y=-42.5, "Profile log-likelihood", cex=0.9, pos=4)
axis(1, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = 'Coefficient of SOUTH', ylab="Profile log-likelihood",
      line = 1.7, cex.lab=1.2)
par(opar)

## Figure 6.1
coef.vector <- NULL
for (i in 1:length(EXECUTIONS))  {
 dp.temp <- glm(EXECUTIONS[-i] ~ INCOME[-i]+PERPOVERTY[-i]+PERBLACK[-i]+log(VC100k96)[-i]+
                   SOUTH[-i]+PROPDEGREE[-i], family=poisson)
  coef.vector <- rbind(coef.vector,dp.temp$coefficients)
}
layout(matrix(c(1,2,3,4,5,6), ncol=2, byrow = TRUE), heights = c(0.33,0.33,0.33))
par(mar=c(3,4.5,2,4),oma=c(2,1,1,3))
for(i in 2:ncol(coef.vector))  {
  x=plot(coef.vector[,i],type="b",xlab="",ylab="",  yaxt="n", xaxt="n", lwd=2,
         ylim=c(min(coef.vector[,i])-abs(min(coef.vector[,i]))*0.25,
         max(coef.vector[,i])+abs(max(coef.vector[,i]))*0.25))
  abline(h=dp.out$coefficients[i])
 axis(1, at =seq(2,16,2), tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
  if(i==2){
   axis(2, at = seq(5,35,5)/100000, labels = as.expression(paste(seq(5,35,5), "e(-5)", sep="")),
         tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
  }
  else{
   axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
  }
  title(xlab = "Index number",
        line = 1.7, cex.lab=1.2)
  title(ylab=dimnames(dp.mat.0)[[2]][i],
        line = 3.25, cex.lab=1.2)
}
par(opar)
```

glm.cis *Compute confidence intervals for predictions.*

Description

Apply an exponential transformation to the confidence intervals and predictions from binomial and Poisson models.

Usage

```
glm.cis(preds, ses, alpha, df)
```

Arguments

preds	The predictions based on the additive linear component of the model.
ses	The standard error(s) of the prediction.
alpha	The desired confidence level.
df	The desired degrees of freedom.

Value

The output is a matrix.

Examples

```
data(campaign)
attach(campaign)
cmpgn.out <- glm(TOTCONTR ~ CANDGENDER + PARTY + INCUMCHALL + HISPPCT,
                 family=Gamma(link = 'log'), data=campaign)
newdat_gender <- data.frame(CANDGENDER = c('F','M'), PARTY= rep('Democrat',2),
                           INCUMCHALL=rep("C", 2), HISPPCT=rep(mean(campaign$HISPPCT),2))
preds_gender <- predict(cmpgn.out, newdata = newdat_gender, se.fit = TRUE)
glm.cis(preds_gender$fit, preds_gender$se.fit, 0.95,cmpgn.out$df.residual)
```

glm.summary *Summarize regression output from generalized linear models.*

Description

An alternative to the summary() function.

Usage

```
glm.summary(in.object, alpha = 0.05)
```

Arguments

`in.object`	The regression output from glm().
`alpha`	A parameter defaulted to 0.05.

Value

The output is a matrix.

Examples

```
data(campaign)
attach(campaign)
cmpgn.out <- glm(TOTCONTR ~ CANDGENDER + PARTY + INCUMCHALL + HISPPCT,
              family=Gamma(link = 'log'), data=campaign)
glm.summary(cmpgn.out)
```

`glm.summary.multinom` *Summarize regression output from multinomial generalized linear models.*

Description

An alternative to the summary() function.

Usage

```
glm.summary.multinom(in.object, alpha = 0.05)
```

Arguments

`in.object`	The regression output from multinom().
`alpha`	A parameter defaulted to 0.05.

Value

The output is a list.

Examples

```
data(primary)
attach(primary)
library(nnet)
primary.out <- multinom(PRIMVOTE ~ AGE + GENDER + EDUCATION + REGION +
                         RELIGIOSITY + IDEOLOGY + RWA + TRUMPWIN, data=primary)
glm.summary.multinom(primary.out)
```

glm.vc — *Compute variance-covariance matrix.*

Description

Calculate the (unscaled) variance-covariance matrix from a generalized linear model regression output. Used in ‘GLMpack‘ within the function ‘glm.summary()‘.

Usage

```
glm.vc(obj)
```

Arguments

`obj`	The regression output from glm().

Value

The output is a matrix.

Examples

```
data(campaign)
attach(campaign)
cmpgn.out <- glm(TOTCONTR ~ CANDGENDER + PARTY + INCUMCHALL + HISPPCT,
              family=Gamma(link = 'log'), data=campaign)
glm.vc(cmpgn.out)
```

mexico — *Data on Mexico from the Mexican Family Life Survey*

Description

Data for the Mexico example used in chapter 7

Usage

```
data(mexico)
```

Format

A data frame with 644 rows and 11 variables:

SUBJECT Subject ID

HOUSEHOLD Household ID

STATE State indicator, here 5 for the state of Coahuila

MIGR Dummy variable indicating whether respondents have thought of migrating

NCRIME Number of crimes of which the female respondent has been the victim

SEV Mean seriousness of the crime

PAST Evaluations of past life conditions

FUT Evaluations of future community conditions

INC Respondents' income

AGE Respondents' age

WAVE Indicator for the wave under analysis ...

Examples

```
data(mexico)
attach(mexico)

## Table 7.3
library(lme4)
cs <- function(x) scale(x,scale=TRUE,center=TRUE) # Function to re-scale
mexico.out <- glmer(MIGR ~ cs(NCRIME) + cs(SEV) + cs(PAST) +
                cs(FUT) +cs(INC) + cs(AGE) + WAVE
              + (1|SUBJECT), data=mexico, family=binomial("logit"),
           glmerControl(optimizer="bobyqa", tolPwrss=5e-2, optCtrl = list(maxfun=10000)))
summary(mexico.out)
se <- sqrt(diag(vcov(mexico.out)))
(cis.mexico.out <- cbind(Est = fixef(mexico.out),
                        LL = fixef(mexico.out) - 1.96 * se,
                        UL = fixef(mexico.out) + 1.96 * se))
```

`peace` *Data on the characteristics of peace agreement outcomes*

Description

Data for the Peace example used in chapter 7

Usage

```
data(peace)
```

Format

A data frame with 216 rows and 9 variables:

OUTISS Ordinal variable indicating the scale of outstanding issues that were not resolved during the peace negotiations with 30 percent zero values

PP Binary variable indicating whether a rebel force is allowed to transform into a legal political party

INTCIV Binary variable indicating whether members of the rebel group are to be integrated into the civil service

AMN Binary variable indicating whether there is an amnesty provision in the agreement

PRIS Binary variable indicating whether prisoners are released

FED Binary variable indicating whether a federal state solution is included

COMIMP Binary variable indicating whether the agreement establishes a commission or committee to oversee implementation

REAFFIRM Binary variable indicating whether the agreement reaffirms earlier peace agreements

PKO Binary variable indicating whether or not the peace agreement included the deployment of peacekeeping forces ...

Examples

```
data(peace)
attach(peace)
require(pscl)

## Table 7.6
M3 <- zeroinfl(OUTISS ~ PP + INTCIV + AMN + PRIS + FED + COMIMP + REAFFIRM | PKO, data=peace)
summary(M3)
out.table.count <- cbind(summary(M3)$coef$count[,1:2],
    summary(M3)$coef$count[,1] - 1.96*summary(M3)$coef$count[,2],
    summary(M3)$coef$count[,1] + 1.96*summary(M3)$coef$count[,2])
out.table.zero <- cbind(summary(M3)$coef$zero[,1:2],
    summary(M3)$coef$zero[,1] - 1.96*summary(M3)$coef$zero[,2],
    summary(M3)$coef$zero[,1] + 1.96*summary(M3)$coef$zero[,2])
out.table.count
out.table.zero
```

`primary` *Data on the 2016 Republican Presidentical Primaries*

Description

Data for the primary example used in chapters 4 and 5

Usage

```
data(primary)
```

Format

A data frame with 706 rows and 9 variables:

PRIMVOTE Vote intention

AGE Age

GENDER Gender

EDUCATION Education

REGION Region of the country in which the respondent lives

RELIGIOSITY Religiosity

IDEOLOGY Ideology

RWA Right Wing Authoritarianism scale

TRUMPWIN Perceptions of whether Trump could win ...

Examples

```
opar = par(mfrow=c(1,1), mar=c(5.1,4.1,4.1,2.1), oma=c(0,0,0,0))
data(primary)
attach(primary)
library(nnet)
library(pBrackets)
library(effects)

## Model
primary.out <- multinom(PRIMVOTE ~ AGE + GENDER + EDUCATION + REGION +
                        RELIGIOSITY + IDEOLOGY + RWA + TRUMPWIN, data=primary)
summ.primary.out <- glm.summary.multinom(primary.out)

## Figure 4.2
par(mfrow=c(3,3), mar=c(2.5,2,2,1))
# Plot 1: Electoral preference
countsPV0 <- barplot(table(primary$PRIMVOTE), main="Electoral Preference",
        xlab="Candidates", mgp=c(1.1, 0.2, 1))
text(countsPV0[,1], rep(10,4), as.numeric(table(primary$PRIMVOTE)), cex=1.5)
# Plot 2: Age
countsAGE <- barplot(table(primary$AGE), main="Age",
                     xlab="Age categories", mgp=c(1.1, 0.2, 0))
text(countsAGE[,1], rep(10,4), as.numeric(table(primary$AGE)), cex=1.5)
# Plot 3: Gender
countsGENDER <- barplot(table(primary$GENDER), main="Gender",
                     xlab="Gender categories", mgp=c(1.1, 0.2, 0), ylim=c(0,500))
text(countsGENDER[,1], rep(25,2), as.numeric(table(primary$GENDER)), cex=1.5)
# Plot 4: Education
countsEDUC <- barplot(table(primary$EDUCATION), main="Education",
                        xlab="Schooling level", mgp=c(1.1, 0.2, 0))
text(countsEDUC[,1], rep(10,4), as.numeric(table(primary$EDUCATION)), cex=1.5)
# Plot 5: Region
countsREG <- barplot(table(primary$REGION), main="Region",
                      xlab="Region", mgp=c(1.1, 0.2, 0))
text(countsREG[,1], rep(10,4), as.numeric(table(primary$REGION)), cex=1.5)
```

```
hist(primary$RELIGIOSITY,xlab="Religiosity Score",ylab="",
     xlim=c(-1.5,2), ylim=c(0, 225), main="Religiosity",
     col = "gray70", mgp=c(1.1, 0.2, 0))
# Plot 7: Ideology
hist(primary$IDEOLOGY,xlab="Ideology Score",ylab="",
     xlim=c(-2,1.5), ylim=c(0, 150), main="Ideology",
     col = "gray70", mgp=c(1.1, 0.2, 0))
# Plot 8: Right Wing Authoritarianism
hist(primary$RWA,xlab="Right Wing Authoritarianism Score",ylab="",
     xlim=c(-2.5,2), ylim=c(0, 200), main="Authoritarianism",
     col = "gray70", mgp=c(1.1, 0.2, 0))
colnames(primary)
# Plot 9: Could Trump Win?
countsWIN <- barplot(table(primary$TRUMPWIN), main="Trump's winnability",
                     xlab="Perceptions of whether Trump could win",
                     mgp=c(1.1, 0.2, 0), ylim=c(0,550))
text(countsWIN[,1], rep(30,3), as.numeric(table(primary$TRUMPWIN)), cex=1.5)
par(opar)

## Figure 5.3
layout(matrix(1:2, ncol = 1), heights = c(0.9,0.1))
par(mar=c(3,4,0,1),oma=c(1,1,1,1))
plot(summ.primary.out[[1]][,1], type = "n", xaxt="n", yaxt="n",
     xlim=c(-10,3), ylim=c(0,60), ylab="", xlab="")
abline(v=-5, h=c(12,16,28,40,44,48,52), lwd=2)
abline(h=c(4,8,20,24,32,36,56), lty=3, col="gray60")
text(rep(-7.5,15), seq(2,58,4),
     labels = c('30-44', '45-59', '60+',
                'Male',
                'High School','Some College','Bachelor\'s degree or higher',
                'Northeast', 'South', 'West',
                'Religiosity',
                'Ideology',
                'Authoritarianism',
                'Yes', 'Don\'t know'))
segments(summ.primary.out[[1]][-1,3], seq(1,57,4), summ.primary.out[[1]][-1,4],
         seq(1,57,4), col="gray30", lwd=2)
points(summ.primary.out[[1]][-1,1], seq(1,57,4), pch=21, cex=1.4, bg="black")
text(summ.primary.out[[1]][-1,1], seq(1,57,4), labels = "C", cex = 0.7, col="white")
segments(summ.primary.out[[2]][-1,3], seq(3,59,4), summ.primary.out[[2]][-1,4],
         seq(3,59,4), col="gray30", lwd=2)
points(summ.primary.out[[2]][-1,1], seq(3,59,4), pch=21, cex=1.4, bg="white")
text(summ.primary.out[[2]][-1,1], seq(3,59,4), labels = "K", cex = 0.7, col="black")
abline(v=0, lty=2)
axis(1, tck=0.01, at = seq(-5,5,0.5),cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0,
     lwd.ticks = 1)
axis(2, tck=0.02, at = c(6,14,22,34,42,46,50,56), labels=c('AGE', 'GENDER',
     'EDUCATION', 'REGION', 'RELIGIOSITY', 'IDEOLOGY', 'RWA', 'TRUMPWIN'),
     cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 0, las=1)
title(xlab = "Coefficient",
      line = 1.7, cex.lab=1.3)
par(mar=c(0,4,0,0))
plot(0,0, type="n", axes = FALSE, xaxt="n", yaxt="n", xlab="", ylab = "")
```

```
legend("center", c("Cruz", "Kasich"), ncol=2, pch=c(21,21), pt.bg=c("black", "white"),
       pt.cex=rep(1.4,2), bty = "n")
par(opar)

## Figure 5.4
mygray = rgb(153, 153, 153, alpha = 200, maxColorValue = 255)
mygray2 = rgb(179, 179, 179, alpha = 150, maxColorValue = 255)
mygray3 = rgb(204, 204, 204, alpha = 150, maxColorValue = 255)
preds_win <- Effect("TRUMPWIN", primary.out)
preds_ideol <- Effect("IDEOLOGY", primary.out, xlevels=list(IDEOLOGY=100))
par(mfrow=c(1,2), mar=c(4,3,3,0),oma=c(1,1,1,1))
plot(1:3, preds_win$prob[,1], type="n",xlab="",ylab="",  yaxt="n", xaxt="n",
     xlim=c(0,4), ylim=c(0, 0.7))
segments(rep(1:3, 3), preds_win$lower.prob[,1:3], rep(1:3, 3), preds_win$upper.prob[,1:3],
         col=rep(c('black', 'black', 'gray60'), each=3), lty = rep(c(1,2,1), each=3))
points(rep(1:3,3), preds_win$prob[,1:3], pch=21, cex = 2,
     bg=rep(c("black", "white", "gray60"),each=3), col=rep(c("black", "black", "gray60"),each=3))
text(rep(1:3,3), preds_win$prob[,1:3], labels=rep(c("T", "C", "K"), each=3), cex = 0.8,
    bg=rep(c("black", "white", "gray60"),each=3), col=rep(c("white", "black", "black"),each=3))
axis(1, at=1:3, labels = c("No", "Yes", "DK"), tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0),
     lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = 'Perceptions of whether Trump could win the election',
      ylab="Probability of voting",
      line = 1.7, cex.lab=1)
title(line = 1, main="Winnability", font.main=3)
plot(preds_ideol$x$IDEOLOGY, preds_ideol$prob[,1], type="n",xlab="",ylab="",  yaxt="n", xaxt="n",
     xlim=c(-2,1.5), ylim=c(0, 0.7))
polygon(c(preds_ideol$x$IDEOLOGY, rev(preds_ideol$x$IDEOLOGY)),
     c(preds_ideol$lower.prob[,2], rev(preds_ideol$upper.prob[,2])), border = NA, col=mygray2)
polygon(c(preds_ideol$x$IDEOLOGY, rev(preds_ideol$x$IDEOLOGY)),
     c(preds_ideol$lower.prob[,1], rev(preds_ideol$upper.prob[,1])), border = NA, col=mygray)
polygon(c(preds_ideol$x$IDEOLOGY, rev(preds_ideol$x$IDEOLOGY)),
     c(preds_ideol$lower.prob[,3], rev(preds_ideol$upper.prob[,3])), border = NA, col=mygray3)
lines(preds_ideol$x$IDEOLOGY, preds_ideol$prob[,1], col="gray20", lwd=2)
lines(preds_ideol$x$IDEOLOGY, preds_ideol$prob[,2], col="gray40", lwd=2, lty=2)
lines(preds_ideol$x$IDEOLOGY, preds_ideol$prob[,3], col="black", lwd=2)
text(rep(1,3), c(min(preds_ideol$prob[,1]), min(preds_ideol$prob[,2]),
     max(preds_ideol$prob[,3]))+.05, labels = c('Trump', 'Cruz', 'Kasich'))
axis(1, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = 'Ideology scores',
      ylab="Probability of voting",
      line = 1.7, cex.lab=1)
title(line = 1, main="Ideology", font.main=3)
par(opar)
```

scotvote *Data on the Scottish national parliament vote*

Description

Data for the Scotland example used in chapters 4, 5, and 6

Usage

```
data(scotvote)
```

Format

A data frame with 32 rows and 7 variables:

PerYesTax Percentage of population who granting Scottish parliament taxation powers

CouncilTax Council tax collected

PerClaimantFemale Female percentage of total claims for unemployment benefits as of January 1998

StdMortalityRatio Standardized mortality rate

Active Percentage of economically active individuals relative to the population of working age

GDP GDP per council

Percentage5to15 Percentage of children aged 5 to 15 ...

Examples

```
data(scotvote)
attach(scotvote)

## Table 4.3
scotvote

## Table 5.3
scottish.vote.glm <- glm((PerYesTax) ~ CouncilTax*PerClaimantFemale+
                     StdMortalityRatio+Active+GDP+Percentage5to15,
                     family=Gamma,data=scotvote)
vote.sum <- summary(scottish.vote.glm)
round(cbind(
  vote.sum$coefficients[,1], vote.sum$coefficients[,2],
  confint(scottish.vote.glm)),4)
vote.sum

## Table 6.3
anova(scottish.vote.glm,test="F")
```

star *Data on California state data on educational policy and outcomes*

Description

Data for the STAR program example used in chapter 6

Usage

```
data(star)
```

Format

A data frame with 303 rows and 16 variables:

LOWINC Proportion of low-income students

PERASIAN Proportions of Asian students

PERBLACK Proportions of African-American students

PERHISP Proportions of Hispanic students

PERMINTE Percentage of minority teachers

AVYRSEXP Mean teacher experience in years

AVSAL Median teacher salary, including benefits, in thousands of dollars

PERSPEN Per-pupil expenditures in thousands of dollars

PTRATIO Pupil/teacher ratio in the classroom

PCTAF Percentage of students taking college credit courses

PCTCHRT Percentage of schools in the district that are charter schools

PCTYRRND Percent of schools in the district operating year-round programs

READTOT Total number of students taking the reading exam in the 9th grade

PR50RD Proportion of students scoring over the reading median in the 9th grade

MATHTOT Total number of students taking the math exam in the 9th grade

PR50M Proportion of students scoring over the math median in the 9th grade ...

Examples

```
opar = par(mfrow=c(1,1), mar=c(5.1,4.1,4.1,2.1), oma=c(0,0,0,0))
data(star)
attach(star)

## MATH MODEL
star.logit.fit <- glm(cbind(PR50M,MATHTOT-PR50M) ~ LOWINC + PERASIAN + PERBLACK + PERHISP +
                  PERMINTE * AVYRSEXP * AVSAL + PERSPEN * PTRATIO * PCTAF +
                  PCTCHRT + PCTYRRND, family=binomial(link=logit),data=star)

## READING MODEL
```

```
star.logit.fit2 <- glm(cbind(PR50RD,READTOT-PR50RD) ~ LOWINC + PERASIAN + PERBLACK + PERHISP +
                   PERMINTE * AVYRSEXP * AVSAL + PERSPEN * PTRATIO * PCTAF +
                   PCTCHRT + PCTYRRND, family=binomial(link=logit),data=star)

## Table 6.4
star.summ.mat <- round(summary(star.logit.fit)$coef, 4)
data.frame(cbind(star.summ.mat[,1], star.summ.mat[,2], "[", round(confint(star.logit.fit)[,1],4),
      " ~", round(confint(star.logit.fit)[,2],4), "]"))

## Table 6.5
mean.vector <- apply(star,2,mean)
diff.vector <- c(1,mean.vector[1:12],mean.vector[5]*mean.vector[6],mean.vector[5]*mean.vector[7],
                 mean.vector[6]*mean.vector[7],mean.vector[8]*mean.vector[9],
                 mean.vector[8]*mean.vector[10],mean.vector[9]*mean.vector[10],
                 mean.vector[5]*mean.vector[6]*mean.vector[7],
                 mean.vector[8]*mean.vector[9]*mean.vector[10])
names(diff.vector) <- names(summary(star.logit.fit2)$coef[,1])
# PERMINTE FIRST DIFFERENCE ACROSS IQR
logit <- function(vec){return(exp(vec)/(1+exp(vec)))}
logit(c(diff.vector[1:5],6.329,diff.vector[7:13],6.329*mean.vector[6],6.329*mean.vector[7],
        diff.vector[16:19],6.329*mean.vector[6]*mean.vector[7],diff.vector[21])
      %*%summary.glm(star.logit.fit)$coef[,1]) -
logit(c(diff.vector[1:5],19.180,diff.vector[7:13],19.180*mean.vector[6],19.180*mean.vector[7],
          diff.vector[16:19],19.180*mean.vector[6]*mean.vector[7],diff.vector[21])
        %*%summary.glm(star.logit.fit)$coef[,1])
# First quartile information
q1.diff.mat <- q2.diff.mat <- q3.diff.mat <- q4.diff.mat <-
  matrix(rep(diff.vector,length(diff.vector)),
                      nrow=length(diff.vector), ncol=length(diff.vector),
                      dimnames=list(names(diff.vector),names(diff.vector)))
diag(q1.diff.mat)[2:13] <- apply(star,2,summary)[2,1:12]
q1.diff.mat[14,6] <- q1.diff.mat[6,6]*q1.diff.mat[7,6]
q1.diff.mat[15,6] <- q1.diff.mat[6,6]*q1.diff.mat[8,6]
q1.diff.mat[20,6] <- q1.diff.mat[6,6]*q1.diff.mat[7,6]*q1.diff.mat[8,6]
q1.diff.mat[14,7] <- q1.diff.mat[7,7]*q1.diff.mat[6,7]
q1.diff.mat[16,7] <- q1.diff.mat[7,7]*q1.diff.mat[8,7]
q1.diff.mat[20,7] <- q1.diff.mat[6,7]*q1.diff.mat[7,7]*q1.diff.mat[8,7]
q1.diff.mat[15,8] <- q1.diff.mat[8,8]*q1.diff.mat[6,8]
q1.diff.mat[16,8] <- q1.diff.mat[8,8]*q1.diff.mat[7,8]
q1.diff.mat[20,8] <- q1.diff.mat[6,8]*q1.diff.mat[7,8]*q1.diff.mat[8,8]
q1.diff.mat[17,9] <- q1.diff.mat[9,9]*q1.diff.mat[10,9]
q1.diff.mat[18,9] <- q1.diff.mat[9,9]*q1.diff.mat[11,9]
q1.diff.mat[21,9] <- q1.diff.mat[9,9]*q1.diff.mat[10,9]*q1.diff.mat[11,9]
q1.diff.mat[17,10] <- q1.diff.mat[10,10]*q1.diff.mat[9,10]
q1.diff.mat[19,10] <- q1.diff.mat[10,10]*q1.diff.mat[11,10]
q1.diff.mat[21,10] <- q1.diff.mat[9,10]*q1.diff.mat[10,10]*q1.diff.mat[11,10]
q1.diff.mat[18,11] <- q1.diff.mat[11,11]*q1.diff.mat[9,11]
q1.diff.mat[19,11] <- q1.diff.mat[11,11]*q1.diff.mat[10,11]
q1.diff.mat[21,11] <- q1.diff.mat[9,11]*q1.diff.mat[10,11]*q1.diff.mat[11,11]
# Third quartile
diag(q2.diff.mat)[2:13] <- apply(star,2,summary)[5,1:12]
q2.diff.mat[14,6] <- q2.diff.mat[6,6]*q2.diff.mat[7,6]
q2.diff.mat[15,6] <- q2.diff.mat[6,6]*q2.diff.mat[8,6]
```

```
q2.diff.mat[20,6] <- q2.diff.mat[6,6]*q2.diff.mat[7,6]*q2.diff.mat[8,6]
q2.diff.mat[14,7] <- q2.diff.mat[7,7]*q2.diff.mat[6,7]
q2.diff.mat[16,7] <- q2.diff.mat[7,7]*q2.diff.mat[8,7]
q2.diff.mat[20,7] <- q2.diff.mat[6,7]*q2.diff.mat[7,7]*q2.diff.mat[8,7]
q2.diff.mat[15,8] <- q2.diff.mat[8,8]*q2.diff.mat[6,8]
q2.diff.mat[16,8] <- q2.diff.mat[8,8]*q2.diff.mat[7,8]
q2.diff.mat[20,8] <- q2.diff.mat[6,8]*q2.diff.mat[7,8]*q2.diff.mat[8,8]
q2.diff.mat[17,9] <- q2.diff.mat[9,9]*q2.diff.mat[10,9]
q2.diff.mat[18,9] <- q2.diff.mat[9,9]*q2.diff.mat[11,9]
q2.diff.mat[21,9] <- q2.diff.mat[9,9]*q2.diff.mat[10,9]*q2.diff.mat[11,9]
q2.diff.mat[17,10] <- q2.diff.mat[10,10]*q2.diff.mat[9,10]
q2.diff.mat[19,10] <- q2.diff.mat[10,10]*q2.diff.mat[11,10]
q2.diff.mat[21,10] <- q2.diff.mat[9,10]*q2.diff.mat[10,10]*q2.diff.mat[11,10]
q2.diff.mat[18,11] <- q2.diff.mat[11,11]*q2.diff.mat[9,11]
q2.diff.mat[19,11] <- q2.diff.mat[11,11]*q2.diff.mat[10,11]
q2.diff.mat[21,11] <- q2.diff.mat[9,11]*q2.diff.mat[10,11]*q2.diff.mat[11,11]
# Minimum
diag(q3.diff.mat)[2:13] <- apply(star,2,summary)[1,1:12]
q3.diff.mat[14,6] <- q3.diff.mat[6,6]*q3.diff.mat[7,6]
q3.diff.mat[15,6] <- q3.diff.mat[6,6]*q3.diff.mat[8,6]
q3.diff.mat[20,6] <- q3.diff.mat[6,6]*q3.diff.mat[7,6]*q3.diff.mat[8,6]
q3.diff.mat[14,7] <- q3.diff.mat[7,7]*q3.diff.mat[6,7]
q3.diff.mat[16,7] <- q3.diff.mat[7,7]*q3.diff.mat[8,7]
q3.diff.mat[20,7] <- q3.diff.mat[6,7]*q3.diff.mat[7,7]*q3.diff.mat[8,7]
q3.diff.mat[15,8] <- q3.diff.mat[8,8]*q3.diff.mat[6,8]
q3.diff.mat[16,8] <- q3.diff.mat[8,8]*q3.diff.mat[7,8]
q3.diff.mat[20,8] <- q3.diff.mat[6,8]*q3.diff.mat[7,8]*q3.diff.mat[8,8]
q3.diff.mat[17,9] <- q3.diff.mat[9,9]*q3.diff.mat[10,9]
q3.diff.mat[18,9] <- q3.diff.mat[9,9]*q3.diff.mat[11,9]
q3.diff.mat[21,9] <- q3.diff.mat[9,9]*q3.diff.mat[10,9]*q3.diff.mat[11,9]
q3.diff.mat[17,10] <- q3.diff.mat[10,10]*q3.diff.mat[9,10]
q3.diff.mat[19,10] <- q3.diff.mat[10,10]*q3.diff.mat[11,10]
q3.diff.mat[21,10] <- q3.diff.mat[9,10]*q3.diff.mat[10,10]*q3.diff.mat[11,10]
q3.diff.mat[18,11] <- q3.diff.mat[11,11]*q3.diff.mat[9,11]
q3.diff.mat[19,11] <- q3.diff.mat[11,11]*q3.diff.mat[10,11]
q3.diff.mat[21,11] <- q3.diff.mat[9,11]*q3.diff.mat[10,11]*q3.diff.mat[11,11]
diag(q4.diff.mat)[2:13] <- apply(star,2,summary)[6,1:12]
q4.diff.mat[14,6] <- q4.diff.mat[6,6]*q4.diff.mat[7,6]
q4.diff.mat[15,6] <- q4.diff.mat[6,6]*q4.diff.mat[8,6]
q4.diff.mat[20,6] <- q4.diff.mat[6,6]*q4.diff.mat[7,6]*q2.diff.mat[8,6]
q4.diff.mat[14,7] <- q4.diff.mat[7,7]*q4.diff.mat[6,7]
q4.diff.mat[16,7] <- q4.diff.mat[7,7]*q4.diff.mat[8,7]
q4.diff.mat[20,7] <- q4.diff.mat[6,7]*q4.diff.mat[7,7]*q4.diff.mat[8,7]
q4.diff.mat[15,8] <- q4.diff.mat[8,8]*q4.diff.mat[6,8]
q4.diff.mat[16,8] <- q4.diff.mat[8,8]*q4.diff.mat[7,8]
q4.diff.mat[20,8] <- q4.diff.mat[6,8]*q4.diff.mat[7,8]*q4.diff.mat[8,8]
q4.diff.mat[17,9] <- q4.diff.mat[9,9]*q4.diff.mat[10,9]
q4.diff.mat[18,9] <- q4.diff.mat[9,9]*q4.diff.mat[11,9]
q4.diff.mat[21,9] <- q4.diff.mat[9,9]*q4.diff.mat[10,9]*q4.diff.mat[11,9]
q4.diff.mat[17,10] <- q4.diff.mat[10,10]*q4.diff.mat[9,10]
q4.diff.mat[19,10] <- q4.diff.mat[10,10]*q4.diff.mat[11,10]
q4.diff.mat[21,10] <- q4.diff.mat[9,10]*q4.diff.mat[10,10]*q4.diff.mat[11,10]
q4.diff.mat[18,11] <- q4.diff.mat[11,11]*q4.diff.mat[9,11]
```

```
q4.diff.mat[19,11] <- q4.diff.mat[11,11]*q4.diff.mat[10,11]
q4.diff.mat[21,11] <- q4.diff.mat[9,11]*q4.diff.mat[10,11]*q4.diff.mat[11,11]
first_diffs <- NULL
for (i in 2:13){
        temp1 <- logit(q2.diff.mat[,i]%*%summary.glm(star.logit.fit)$coef[,1]) -
                        logit(q1.diff.mat[,i]%*%summary.glm(star.logit.fit)$coef[,1])
        temp2 <- logit(q4.diff.mat[,i]%*%summary.glm(star.logit.fit)$coef[,1]) -
          logit(q3.diff.mat[,i]%*%summary.glm(star.logit.fit)$coef[,1])
        first_diffs <- rbind(first_diffs, c(temp1,temp2))
}
first_diffs <- round(first_diffs,4)
diffs_mat <- cbind(diag(q1.diff.mat)[2:13], diag(q2.diff.mat)[2:13],
                   first_diffs[,1],
                   diag(q3.diff.mat)[2:13], diag(q4.diff.mat)[2:13],
                   first_diffs[,2])
colnames(diffs_mat) <- c("1st quartile", "3rd quartile", "Interquartile 1st diff",
                         "Min", "Max", "Full range 1st diff")
diffs_mat

star.mu <- predict.glm(star.logit.fit,type="response")
star.y <- PR50M/MATHTOT
star.n <- length(star.y)
PR50M.adj <- PR50M
for (i in 1:length(PR50M.adj))  {
  if (PR50M.adj[i] > mean(PR50M)) PR50M.adj[i] <- PR50M.adj[i] - 0.5
  if (PR50M.adj[i] < mean(PR50M)) PR50M.adj[i] <- PR50M.adj[i] + 0.5
}
par(mfrow=c(1,3), mar=c(6,3,6,2),oma=c(4,1,4,1))
plot(star.mu,star.y,xlab="",ylab="", yaxt='n', xaxt="n", pch="+")
axis(1, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = "Fitted values", ylab="Observed values",
      line = 1.7, cex.lab=1.3)
title(main="Model Fit Plot",
      line = 1, cex.main=1.7, font.main=1)
abline(lm(star.y~star.mu)$coefficients, lwd=2)
plot(fitted(star.logit.fit),resid(star.logit.fit,type="pearson"),xlab="",ylab="",
     yaxt='n', xaxt="n", pch="+")
axis(1, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = "Fitted values", ylab="Pearson Residuals",
      line = 1.7, cex.lab=1.3)
title(main="Residual Dependence Plot",
      line = 1, cex.main=1.7, font.main=1)
abline(0,0, lwd=2)
qqnorm(resid(star.logit.fit,type="deviance"),main="",xlab="",ylab="",
       yaxt='n', xaxt="n", pch="+")
axis(1, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.02, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = "Quantiles of N(0,1)", ylab="Deviance Residual Quantiles",
      line = 1.7, cex.lab=1.3)
title(main="Normal-Quantile Plot",
      line = 1, cex.main=1.7, font.main=1)
```

```
abline(-0.3,3.5, lwd=2)
par(opar)
```

---

suicide   *Data on suicides in 2009 in OECD member states*

Description

Data for the suicide example used in chapter 7

Usage

```
data(suicide)
```

Format

A data frame with 32 rows and 7 variables:

COUNTRYCODE Country code

COUNTRYNAME Name of the country

YEAR Year

DEATHS Number of suicides in the country per 100,000 individuals

GDP GDP in thousands of dollars

SUBABUSE Share of the population with alcohol or drug use disorder

TEMP Average temperature ...

Examples

```
opar = par(mfrow=c(1,1), mar=c(5.1,4.1,4.1,2.1), oma=c(0,0,0,0))
data(suicide)
attach(suicide)

## Table 7.2
# Poisson model
suic.out.p <- glm(DEATHS ~ GDP + TEMP + SUBABUSE, family = poisson)
summary(suic.out.p)
round(confint(suic.out.p),3)
coefs_poisson <- summary(suic.out.p)$coefficients[1:4,]
coefs_poisson
suic.out.qp <- glm(DEATHS ~ GDP + TEMP + SUBABUSE, family = quasipoisson)
summary(suic.out.qp)
round(confint(suic.out.qp),3)
coefs_quasipoisson <- summary(suic.out.qp)$coefficients[1:4,]
coefs_quasipoisson

## Figure 7.1
layout(matrix(c(1,2,3,4), ncol=2, byrow = TRUE))
```

```
par(mar=c(4,3,2,0),oma=c(1,1,1,1))
# Histogram #1
hist(TEMP,xlab="",ylab="",  yaxt="n", xaxt="n", main="", col="gray50", border = "gray30",
    ylim=c(0,15))
axis(1, tck=0, mgp=c(0, 0, 0), lty=1, lwd=0, lwd.ticks = 0)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=1, lwd.ticks = 1, las=2)
title(xlab = 'Mean temperature (in Celsius)', ylab="",
      line = 1.7, cex.lab=1.2)
title(line = 1, main="Temperature", font.main=3)
# Histogram #2
hist(GDP,xlab="",ylab="",  yaxt="n", xaxt="n", main="", col="gray50", border = "gray30",
     ylim=c(0,15))
axis(1, tck=0, mgp=c(0, 0, 0), lty=1, lwd=0, lwd.ticks = 0)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=1, lwd.ticks = 1, las=2)
title(xlab = 'GDP per capita (in thousands of dollars)', ylab="",
      line = 1.7, cex.lab=1.2)
title(line = 1, main="Economic Conditions", font.main=3)
# Histogram #3
hist(SUBABUSE,xlab="",ylab="",  yaxt="n", xaxt="n", main="", col="gray50", border = "gray30",
     ylim=c(0,15))
axis(1, tck=0, mgp=c(0, 0, 0), lty=1, lwd=0, lwd.ticks = 0)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=1, lwd.ticks = 1, las=2)
title(xlab = '% of population with alcohol or drug use disorders', ylab="",line = 1.7, cex.lab=1.2)
title(line = 1, main="Substance abuse", font.main=3)
# Histogram #4
hist(DEATHS,xlab="",ylab="",  yaxt="n", xaxt="n", main="", col="gray10", border = "gray20",
     ylim=c(0,15))
axis(1, tck=0, mgp=c(0, 0, 0), lty=1, lwd=0, lwd.ticks = 0)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=1, lwd.ticks = 1, las=2)
title(xlab = 'Number of suicides per 100,000 people', ylab="",
      line = 1.7, cex.lab=1.2)
title(line = 1, main="Suicide rate", font.main=3)
par(opar)

## Figure 7.2
newdat1 <- data.frame(GDP=seq(13, 70.5, 1), TEMP=rep(mean(TEMP), 58),
                      SUBABUSE=rep(mean(SUBABUSE), 58))
newdat2 <- data.frame(GDP=rep(mean(GDP), 61), TEMP=rep(mean(TEMP), 61), SUBABUSE=seq(0,6,0.1))
preds.qp.gdp <- predict(suic.out.qp, newdata = newdat1, type = "link", se.fit = TRUE)
preds.qp.subabuse <- predict(suic.out.qp, newdata = newdat2, type = "link", se.fit = TRUE)
ilink.qp <- family(suic.out.qp)$linkinv
cis.p.preds.qp.gdp <- cbind(ilink.qp(preds.qp.gdp$fit - (2 * preds.qp.gdp$se.fit)),
                            ilink.qp(preds.qp.gdp$fit + (2 * preds.qp.gdp$se.fit)))
cis.p.preds.qp.subabuse <- cbind(ilink.qp(preds.qp.subabuse$fit - (2 * preds.qp.subabuse$se.fit)),
                        ilink.qp(preds.qp.subabuse$fit + (2 * preds.qp.subabuse$se.fit)))
mygray = rgb(153, 153, 153, alpha = 200, maxColorValue = 255)
par(mar=c(4,3,1,0),oma=c(1,1,1,1), mfrow=c(1,2))
plot(newdat1$GDP, ilink.qp(preds.qp.gdp$fit), type="n",xlab="",ylab="",  yaxt="n", xaxt="n", lwd=2,
     ylim = c(0,37))
polygon(c(newdat1$GDP,rev(newdat1$GDP)), c(cis.p.preds.qp.gdp[,1],rev(cis.p.preds.qp.gdp[,2])),
        border = NA, col = mygray)
lines(newdat1$GDP, ilink.qp(preds.qp.gdp$fit))
points(GDP, DEATHS, pch="+", col="gray20", cex=0.8)
```

```
axis(1, tck=0.03, cex.axis=0.9, at=seq(20,70,10), labels = seq(20,70,10), mgp=c(0.3, 0.3, 0),
     lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = 'GDP per capita (in thousands of dollars)',
      ylab="Number of suicides per 100,000 people", line = 1.7, cex.lab=1.2)
plot(newdat2$SUBABUSE, ilink.qp(preds.qp.subabuse$fit), type="n",xlab="",ylab="",
     yaxt="n", xaxt="n", lwd=2, ylim = c(0,37))
polygon(c(newdat2$SUBABUSE,rev(newdat2$SUBABUSE)), c(cis.p.preds.qp.subabuse[,1],
        rev(cis.p.preds.qp.subabuse[,2])), border = NA, col = mygray)
lines(newdat2$SUBABUSE, ilink.qp(preds.qp.subabuse$fit))
points(SUBABUSE, DEATHS, pch="+", col="gray20", cex=0.8)
axis(1, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1)
axis(2, tck=0.03, cex.axis=0.9, mgp=c(0.3, 0.3, 0), lty=1, lwd=0, lwd.ticks = 1, las=2)
title(xlab = '% of population with alcohol or drug use disorders', ylab="",
      line = 1.7, cex.lab=1.2)
par(opar)
```

Index