

## Computer Lab 5: Poisson Data and Binary Data

### Example: Log-linear models:

The data given in Splus build in data frame Insurance (in the library MASS) consist of the numbers of policy-holders, **Holders**, the numbers of car insurance claims made by those policyholders, **Claims**. There are three explanatory variables, **District** (four levels), **Group** (of car, four levels), and **Age** (four ordered levels). Please analyze the data up to the two way interaction with offset **log(Holders)**.

[code:]

```
library(MASS)
data=Insurance
age=factor(data$Age)
group=factor(data$Group)
dis=factor(data$District)
hol=log(data$Holders)
cla=data$Claims
data3=data.frame(hol,dis,group,age,cla)
glm1=glm(cla~offset(hol)+dis+group+age,family=poisson,data=data3)
anova(glm1,test="Chisq")
summary(glm1,cor=F)
```

### Example: Logistic regression models:

The following table refers to effects of occupations (A, B, C, D, E) and education level (P, S, C) on whether one follows politics regularly.

Follow Politics Regularly	A		B		C		D		E	
	Yes	No								
P	94	84	227	112	356	144	166	526	447	430
S	318	120	371	71	256	76	142	103	78	25
C	473	72	180	8	22	2	47	7	22	2

(a) Analyze these data using both **logit** and **probit** links. Please make conclusions.

(b) Please give the residual plots, including the one for Pearson residuals and the one for deviance residuals.

[code:]

```
### (a)
```

```
yes = c(94, 318, 473, 227, 371, 180, 356, 256, 22, 166, 142, 47, 447, 78, 22)
```

```

no = c(84, 120, 72, 112, 71, 8, 144, 76, 2, 526, 103, 7, 430, 25, 2)
data = cbind(yes, no)
occupation = factor(c(rep("A", 3), rep("B", 3), rep("C", 3), rep("D", 3),
rep("E", 3)))
education = factor(rep(c("P", "S", "C"), 5))
glm.logit = glm(data ~ occupation + education, family = binomial(link = logit))
anova(glm.logit, test = "Chisq")
summary(glm.logit, corr = F)

glm.probit = glm(data ~ occupation + education, family = binomial(link = probit))
anova(glm.probit, test = "Chisq")
summary(glm.probit, corr = F)

### (b)
par(mfrow = c(2, 2))
pResidual.logit = residuals(glm.logit, type = "pearson") ### Pearson residuals
dResidual.logit = residuals(glm.logit, type = "deviance") # ## Deviance residuals
plot(pResidual.logit, xlab = "index", ylab = "Pearson Residuals")
title("Logit Link")
plot(dResidual.logit, xlab = "index", ylab = "Deviance Residuals")
title("Logit Link")
pResidual.probit = residuals(glm.probit, type = "pearson") ### Pearson residuals
dResidual.probit = residuals(glm.probit, type = "deviance") ### Deviance residuals
plot(pResidual.probit, xlab = "index", ylab = "Pearson Residuals")
title("Probit Link")
plot(dResidual.probit, xlab = "index", ylab = "Deviance Residuals")
title("Probit Link")

```