UNIVERSITY OF EDINBURGH
SCHOOL OF MATHEMATICS

Generalised Regression Models

GRM: Example — Poisson Regression in R/S-PLUS **Semester 1, 2022–2023**

1 Textile data set

The following discrete data are on textile faults: the cloth length (x) is the explanatory variable and the number of faults (y) is the response variable.

x:	551	651	832	375	715	868	271	630	491	372	645	441	895	458	642	492
y:	6	4	17	9	14	8	5	7	7	7	6	8	28	4	10	4
x:	543	842	905	542	522	122	657	170	738	371	735	749	495	716	952	417
y:	8	9	23	9	6	1	9	4	9	14	17	10	7	3	9	2

2 Poisson regression model

It is assumed that the responses $Y_1, \ldots, Y_n$ each have a Poisson distribution, and that the ith response Y_i has a mean

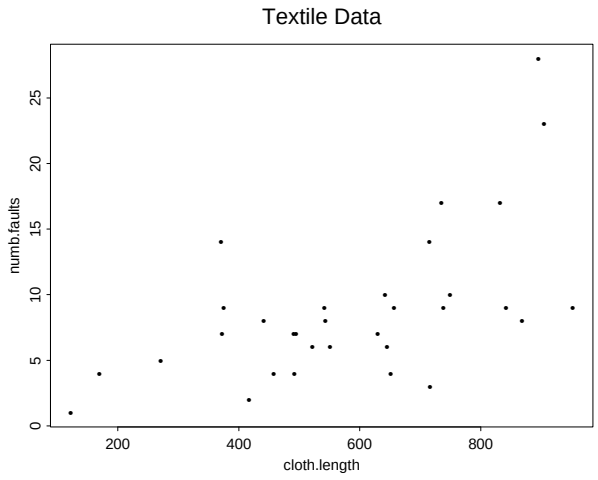
$$\mathrm{E}(Y_i) = \lambda_i = \theta_1 + \theta_2 x_i$$

which depends on the value of an explanatory variable x_i.

3 Plotting the data

```
x <- c(551,651,832,375,715,868,271,630,491,372,645,441,895,458,642,492,
    543,842,905,542,522,122,657,170,738,371,735,749,495,716,952,417)
y <- c(6,4,17,9,14,8,5,7,7,7,6,8,28,4,10,4,
    8,9,23,9,6,1,9,4,9,14,17,10,7,3,9,2)

textile.df <- data.frame(cloth.length=x,numb.faults=y)
attach(textile.df)
plot(cloth.length,numb.faults)
title('Textile Data')
```



4 Fitting the model

```
> textile.glm <- glm(formula = numb.faults ~ cloth.length,
                     family = poisson(identity))
> textile.glm
Call:
glm(formula = numb.faults ~ cloth.length, family = poisson(identity))

Coefficients:
 (Intercept) cloth.length
   0.3234857    0.0145519

Degrees of Freedom: 32 Total; 30 Residual
Residual Deviance: 64.45047
```

5 Extracting information from the `glm` object

```
> summary(textile.glm)

Call: glm(formula = numb.faults ~ cloth.length, family = poisson(identity))
Deviance Residuals:
       Min         1Q     Median       3Q      Max
 -2.798506 -1.104746 -0.2399216 0.550989 3.490582

Coefficients:
                Value  Std. Error   t value
 (Intercept) 0.3234857 1.111843792 0.2909453
cloth.length 0.0145519 0.002079579 6.9975214

(Dispersion Parameter for Poisson family taken to be 1 )

    Null Deviance: 103.7138 on 31 degrees of freedom

Residual Deviance: 64.45047 on 30 degrees of freedom

Number of Fisher Scoring Iterations: 3

Correlation of Coefficients:
             (Intercept)
cloth.length -0.9024138
```