

APTS Statistical Modelling: Practical 1

H. E. Ogden & A. M. Overstall

April 2019

Suppose

$$y_{im} \sim \text{Poisson}(\mu(x_{im})),$$

independently, for $i = 1, \dots, n$ and $m = 1, \dots, M$, where

$$\mu(x_{im}) = 8 \exp(w(x_{im})),$$

for some function w .

Suppose $x_{im} = x_i = -10 + 20(i - 1)/(n - 1)$, $M = 3$ and

$$w(x) = 0.001(100 + x + x^2 + x^3).$$

The code below performs the following simulation study. For $b = 1, \dots, B$:

- For $i = 1, \dots, n$ and $m = 1, \dots, M$, generate

$$y_{im} \sim \text{Poisson}(\mu(x_{im})).$$

- Record the AIC for models

$$y_{im} \sim \text{Poisson}(\mu(x_{im}))$$
$$\mu(x_{im}) = \exp\left(\beta_0 + \sum_{j=1}^p \beta_j x_{im}^j\right),$$

for $p = 0, \dots, p_{\max} = 20$.

```
> B <- 1000
> n <- 25
> M <- 3
> pmax <- 20
>
> x <- rep(seq(from = -10, to = 10, length = n), each = M)
>
> mu <- function(x) {
+   8 * exp(w(x))
+ }
>
> w <- function(x) {
+   0.001 * (100 + x + x^2 + x^3)
+ }
>
> aics <- matrix(0, nrow = B, ncol = pmax)
```

```

> for (b in 1:B) {
+   y <- rpois(n = M * n, lambda = mu(x))
+   mod <- glm(y ~ 1, family = poisson)
+   aics[b, 1] <- AIC(mod)
+   formula <- "y~x"
+   mod <- glm(formula, family = poisson)
+   aics[b, 2] <- AIC(mod)
+   for (j in 3:pmax) {
+     formula <- paste(formula, " + I(x^", j - 1, ") ", sep = "")
+     mod <- glm(formula, family = poisson)
+     aics[b, j] <- AIC(mod)
+   }
+ }
>
> AICorder <- apply(aics, 1, which.min) - 1
> tAIC <- table(AICorder)
> tAIC

```

1. Investigate the performance of AIC as a model selection tool for $n = 25, 50, 100, 1000$.
2. Vary the simulation model, using

$$w(x) = \frac{1.2}{1 + \exp(-x)},$$

to see how AIC performs when the fitted models do not include the simulation model.

3. Modify the code above to compute the values of BIC and AIC_c , where

$$AIC_c = AIC + \frac{2p^2 + 2p}{n - p - 1}.$$