Poisson Regression

Tate Jacobson

Department of Statistics
Oregon State University, Corvallis, Oregon

Spring 2026

Reading: Ch 22 of *The Sleuth*

Outline

1. Motivating Example: Case Study 22.1.1

2. Poisson Distribution

3. Poisson Regression
   - Interpreting Coefficients
   - Residuals

Motivating Example: Case Study 22.1.1

Case Study 22.1.1

These data report the ages and number of successful matings among 41 African elephants over an 8 year study period. We expect older elephants to have more successful matings, since

1. Male elephants continue to grow throughout their lives.
2. Larger males are likely to be more successful at mating.

Question: What is the relationship between age and the number of successful matings?

Elephant data

Display 22.1 **p. 645**

Age at beginning of study and number of successful matings for 41 African elephants

Age	Matings	Age	Matings	Age	Matings
27	0	33	3	39	1
28	1	33	3	41	3
28	1	33	3	42	4
28	1	33	2	43	0
28	3	34	1	43	2
29	0	34	1	43	3
29	0	34	2	43	4
29	0	34	3	43	9
29	2	36	5	44	3
29	2	36	6	45	5
29	2	37	1	47	7
30	1	37	1	48	2
32	2	37	6	52	9
33	4	38	2		

Why not binomial?

The matings are counted over a fixed period of time (8 years)

- As such, the number of matings *does not have a total number of trials or "size" associated with it* and it does not make sense to model it with a binomial distribution.
- Instead, we use the **Poisson distribution** to model this kind of count data.

Poisson Distribution

Poisson Responses

We can think of other examples of count data that do not have a total or “size” associated with them:

1. Number of crabs caught in traps over a 10 hour period
2. Number of times an internet site is visited over 24 hours
3. Number of hot dogs Joey Chestnut eats in 10 minutes

All of these are counts obtained over some interval of time and/or space and would be better modeled with a Poisson distribution than a binomial distribution.

Poisson Responses

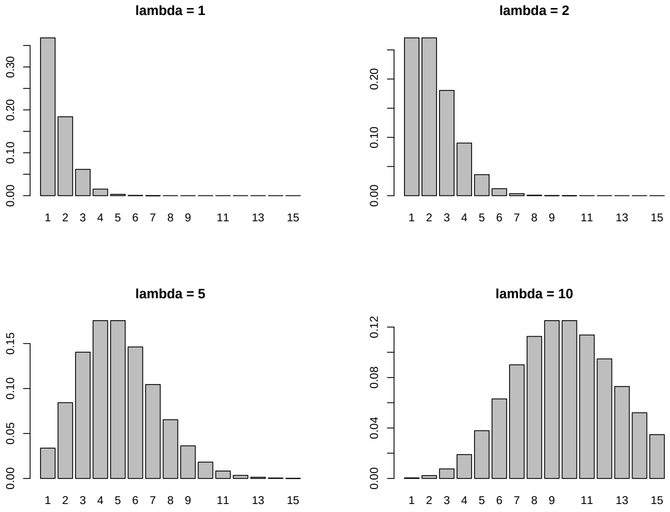
The Poisson probability distribution has distribution function

$$Pr(X = x) = \frac{\lambda^x e^{-\lambda}}{x!},$$

for $x = 0, 1, 2, \ldots$.

- The parameter λ is often called the rate parameter.
- If $X \sim Poisson(\lambda)$, then
  - mean$(X) = \lambda$
  - variance$(X) = \lambda$
- Like the Binomial distribution, the Poisson distribution has only one parameter to describe both the mean and the variance.

Poisson Probability Functions



Poisson Regression

Poisson Regression

The Poisson regression model has two key assumptions

$$\begin{aligned} Y &\sim Poisson(\lambda), \text{ where } \lambda = \mu\{Y|X_1, \ldots, X_p\} \\ \log(\lambda) &= \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p \end{aligned}$$

We see that this is a GLM with link function $g(\lambda) = \log(\lambda)$.

GLM Tools

Because the Poisson regression model is a GLM, tools for other GLMs carry over:

- We use maximum likelihood estimation to estimate $\beta_0, \ldots, \beta_p$
- We can use Wald's test and CIs for individual coefficients
- We can compare models with the drop-in-deviance test

(More on all this later.)

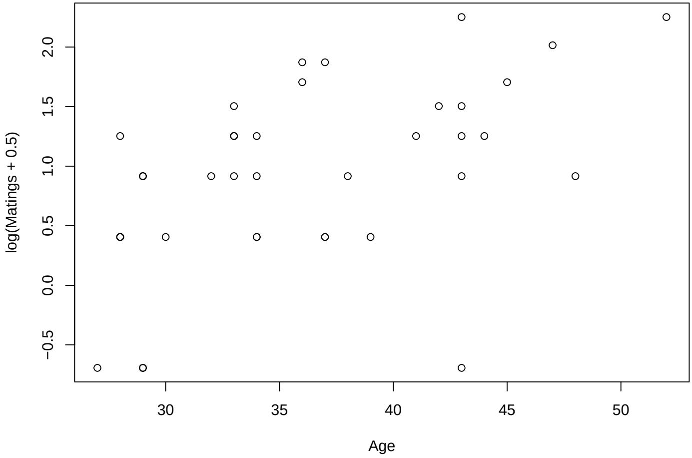
Exploratory Plots

As in binomial logistic regression, we can **plot an empirical version of the link against predictors**:

- in binomial regression:
  - logit link: $\log\left(\frac{\pi}{1-\pi}\right)$
  - empirical logits: $\log\left(\frac{Y_i}{m_i - Y_i}\right)$
- in Poisson regression:
  - log link: $\log(\lambda_i)$
  - empirical version: $\log(Y_i)^*$

*Note: We need to use a transformation like $\log(Y_i + 0.5)$ if any counts are 0.

Empirical Logs



Why not just log-linear regression?

We might wonder, why not just use multiple linear regression (MLR) with a transformed response?

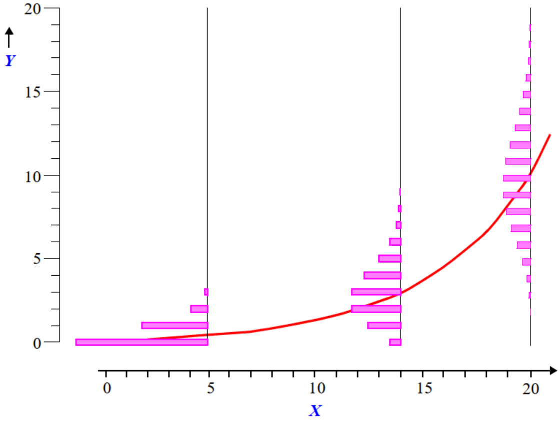
- The MLR model assumes the residuals have constant variance
- The Poisson regression model, on the other hand, assumes that variance$(Y|X_1, \ldots, X_p) = \lambda$

In practice, we can fit a multiple linear regression model with a log-transformed response and check whether the constant variance assumption is met.

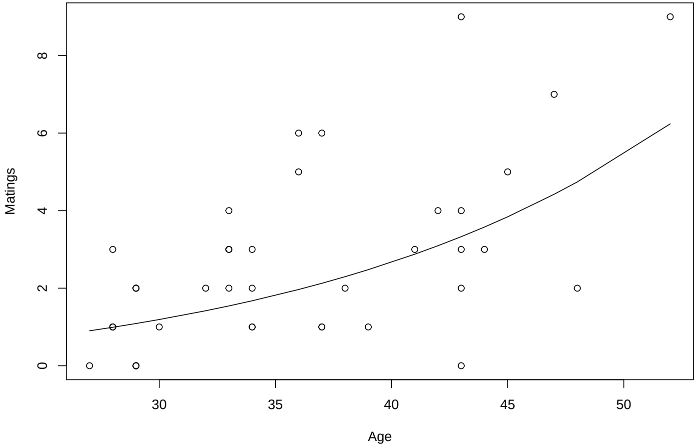
Poisson Variance

Display 22.5 **p. 650**

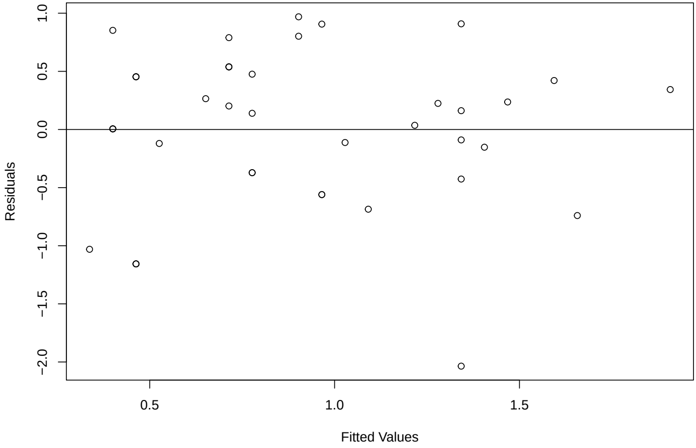
A representation of a log-linear model in which the distribution of *Y*, as a function of *X*, is Poisson with mean μ, and $\log(\mu) = -1.7 + .20X$; the histograms are the Poisson distributions of *Y* at three values of *X*



Log-Linear Fit



Log-Linear Residuals



There are two reasons to use Poisson regression here

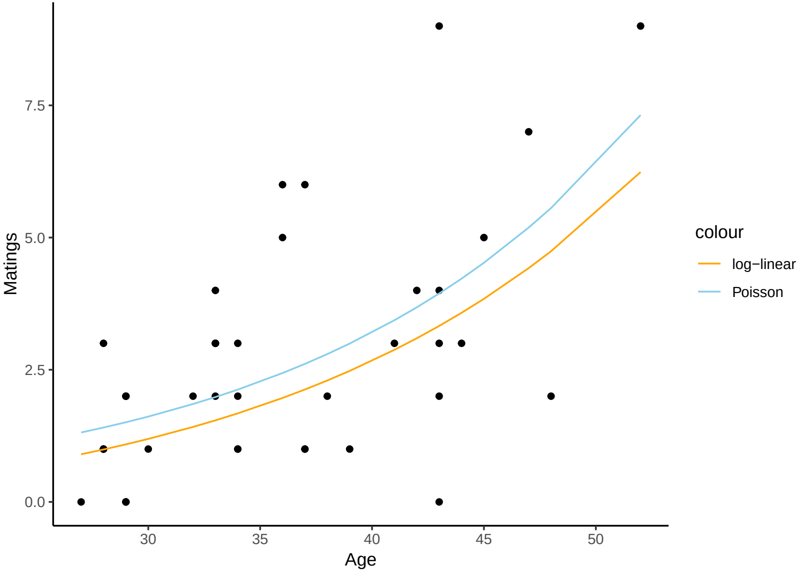
- The response is a count, so clearly not normally distributed
- Based on the residual plot, there's some evidence of non-constant variance, so *another* multiple linear regression assumption (homoscedasticity) is violated.

Poisson Regression in R

We fit the Poisson regression model in R:

```
> glm1 <- glm( Matings ~ Age, family = poisson )
> summary(glm1)

Call:
glm(formula = Matings ~ Age, family = poisson)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.58201    0.54462  -2.905  0.00368 **
Age          0.06869    0.01375   4.997 5.81e-07 ***
---
Signif. codes:  0 ?***? 0.001 ?**? 0.01 ?*? 0.05 ?.? 0.1 ? ? 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 75.372  on 40  degrees of freedom
Residual deviance: 51.012  on 39  degrees of freedom
AIC: 156.46

Number of Fisher Scoring iterations: 5
```

Poisson vs Log-Linear Fit



Interpreting Poisson Regression Coefficients

The estimated model for the number of matings is

$$\log(\hat{\lambda}) = -1.582 + 0.069 \times Age$$

- For each additional year of Age, $\log(\hat{\lambda})$ increases by 0.069.
- On the unlogged scale, this corresponds to a multiplicative change in $\hat{\lambda}$ of . . . .
- For interpretation remember λ is the mean of the response.

Question: What's the multiplicative change in $\hat{\lambda}$ for an increase in Age of 10 years?

Residuals from Poisson Regression

As in binomial logistic regression, we can compute the deviance and Pearson residuals:

- The deviance residuals take the form:

$$D_i = \text{sign}(Y_i - \hat{\lambda}_i)\sqrt{2\left[Y_i \log\left(\frac{Y_i}{\hat{\lambda}_i}\right) - Y_i + \hat{\lambda}_i\right]},$$

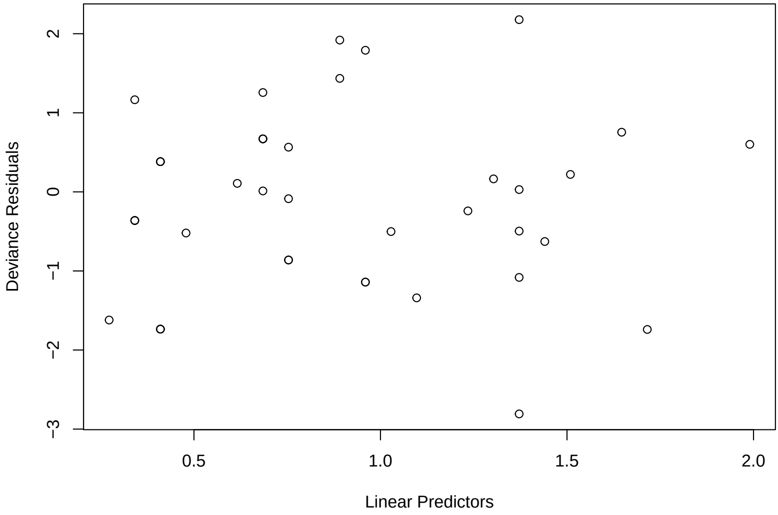
  where $\hat{\lambda}_i$ is estimated mean value for observation i.
- And the Pearson residuals are:

$$P_i = \frac{Y_i - \hat{\lambda}_i}{\sqrt{\hat{\lambda}_i}}.$$

As before,

- the **Residual Deviance** that R reports is the sum of the squared deviance residuals
- the Pearson residuals are a little more intuitively appealing, but the two sets of residuals are often quite similar

Deviance Residuals for Elephant Model



Recap

Material covered: Ch 22.1 - 22.3 of *The Sleuth*

- Poisson regression
- Poisson regression model checking