1 Poisson distribution

\[ Y \sim \text{Pois}(\lambda) \Rightarrow P(Y = k) = \exp(-\lambda) \frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots, \quad (\lambda > 0) \]

Facts:

- \( E(Y) = \text{var}(Y) = \lambda \)
- \( X \sim \text{Bin}(n, p) \) with \( n \) big and \( p \) small then \( X \sim \text{Pois}(np) \)
- Memoryless renewal process
  \(\Leftrightarrow\) time to event is exponentially distributed
  \(\Leftrightarrow\) number of events in time \( T \) is Poisson
- \( Y_i \sim \text{Pois}(\lambda_i) \) independent \(\Rightarrow\) \( \sum Y_i \sim \text{Pois}(\sum \lambda_i) \)

Very useful in “real life”. Examples (under proper assumptions):

1. Number of customers arriving at bank teller in an hour
2. Number of grades a Netflix movie gets in a year
3. Number of mutations in a genomic locus in \( X \) generations (or sum over genome)

2 Poisson regression

Like Gaussian and Binomial (logistic) regression, we assume:

\[ Y | X = x \sim \text{Pois}(\lambda(x)) \]

Like in Binomial we can try to assume \( \lambda(x) = x^T \beta \) but does not guarantee \( \lambda > 0 \).
Solution is to use link functions and assume \( \log(\lambda(x)) = x^T \beta \).

3 Variance stabilizing transformations

Assume we insist on using square error loss for modeling or (especially) evaluation. Probabilistically squared error corresponds to Gaussian log likelihood, of homoskedastic i.e.,

- Symmetric.
- \( \text{Var}(Y | X) \) is fixed throughout covariate space.
Second assumption especially is problematic. Say $Y | X \sim Pois(\lambda(X))$ and $\lambda$ is large, then:

- Normal appx. is not a problem: $Pois(\lambda) \approx N(\lambda, \lambda)$
- But variance linearly depends on mean $\Rightarrow$ homoskedasticity is a problem

Can we transform $Y$ in such a way as to solve this problem?

General setup:

\[
E(Y) = \mu, \quad Var(Y) = \sigma^2 = \Omega(\mu)
\]

Examples:

1. Gaussian: $\sigma^2$ is an independent parameter, can assume $\Omega(\mu) =$constant.
2. Bernoulli: $\Omega(\mu) = \mu(1 - \mu)$
3. Poisson: $\Omega(\mu) = \mu$

We are looking for $f(Y)$ such that $Var(f(Y))$ will be roughly constant (independent of $\mu$), like in the normal case. Taylor expansion:

\[
f(Y) = f(\mu) + (Y - \mu)f'(\mu) + O((Y - \mu)^2)
\]

\[
(f(Y) - f(\mu))^2 \approx (Y - \mu)^2(f'(\mu))^2
\]

Abusing additional probability law (which one?) we write:

\[
Var(f(Y)) \approx Var(Y)(f'(\mu))^2
\]

So to (approximately) stabilize variance we want:

\[
(f'(\mu))^2 = 1/\Omega(\mu)
\]

Poisson: \[(f'(\mu))^2 = 1/\mu \Rightarrow f(\mu) = 2\sqrt{\mu} \Rightarrow Var(\sqrt{Y}) \approx 1/4
\]

Binomial variance stabilizing transformation:

\[
X \sim Bin(n, p) \Rightarrow var(\arcsin(\sqrt{X/n})) \approx 1/2
\]