Generalized Linear Models and Transformed Residuals

Henrik Bengtsson[1]

[1]Mathematical Statistics
Centre for Mathematical Sciences
Lund University

March 5, 2004

Outline

Introduction

Cumulants

Transforms

Transforms for the exponential family

Transformed residuals

The idea behind *transformed residuals*

$$r_T(Y,\theta) = \frac{h(Y) - E_\theta[h(Y)]}{D_\theta[h(Y)]}$$

is to find a transform $h(\cdot)$ such that the distribution of $h(Y)$ is as "Normal as possible".

Two cases; choose $h(\cdot)$ such that

- **(Variance-stabilizing residuals)**
  the μ-asymptotic variance of $h(Y)$ is constant in θ,
- **(Anscombe residuals)**
  the μ-asymptotic skewness of $h(Y)$ is zero.

Cumulants and the exponential family

Consider the exponential family, which has density function

$$f_Y(y) \propto \exp\{y\theta - b(\theta) + c(y)\}.$$

The moment generating function of Y is

$$M_Y(t) = E[exp(tY)] = \exp\{b(\theta + t) - b(\theta)\}.$$

The *cumulant generating function* of Y is

$$K_Y(t) = \ln\{M_Y(t)\} = b(\theta + t) - b(\theta).$$

Power series expansion of the cumulant function

The *cumulant generating function* of Y is

$$K_Y(t) = \ln\{M_Y(t)\}$$

It has a power series expansion

$$K_Y(t) = \sum_{r=1}^{\infty} \kappa_r[Y] \frac{t^r}{r!}$$

where $\{\kappa_r[Y]\}_{r=1}^{\infty}$ are *the cumulants of* Y. The four first are

$$\kappa_1[Y] = E[Y] = \mu, \ \kappa_2[Y] = V[Y] = \sigma^2,$$
$$\kappa_3[Y] = E[(Y - E[Y])^3], \ \kappa_4[Y] = E[(Y - E[Y])^4] - 3\sigma^4$$

One definition of *skewness* is $S_\theta[Y] = \kappa_3[Y]/\sigma^3$ and is a *scale-free measure of asymmetry*.

Normal distribution cumulants

If $Y \in N(\mu, \sigma^2)$, then $\kappa_1[Y] = \mu$, $\kappa_2[Y] = \sigma^2$ and $\kappa_r[Y] = 0$ for $r > 2$. Thus,

$$K_Y(t) = \mu t + \sigma^2 \frac{t^2}{2} + 0 \cdot \frac{t^3}{3!} + 0 \cdot \frac{t^4}{4!} + \dots$$

In other words, the cumulants $\kappa_r[Y]; r > 2$ can be seen as a *measures of non-Normality*.

μ-asymptotic normality

Example: For $Y \in Po(\mu)$, the cumulant g.f. is $K_Y(t) = \mu(\exp(t) - 1)$, which gives that $\kappa_r[Y] = \mu; \forall r$, that is

$$K_Y(t) = \mu t + \mu \frac{t^2}{2} + \mu \frac{t^3}{3!} + \dots$$

Hence, the skewness is $S_\theta[Y] = \kappa_3[Y]/\sigma^3 = \mu/\mu^3 = \mu^{-2}$. When $\mu \to \infty$ we have that $S_\theta[Y] \to 0$, that is, “more Normal”, which we recognize from CLT.

N-asymptotic normality (of the mean)

Let $Y_1, Y_2, \ldots, Y_N$ be iid ($\sim Y$). One can show that the cumulants of the mean $\bar{Y} = \sum_{n=1}^{N} Y_n$ are

$$\kappa_r[\bar{Y}] = \frac{1}{N^{r-1}} \kappa_r[Y].$$

Standardize $Z = \frac{\bar{Y}-\mu}{\sigma/\sqrt{N}}$. Its cumulants are

$$\kappa_1[Z] = 0, \;\; \kappa_2[Z] = 1, \;\; \kappa_r[Z] = \frac{1}{\sigma^r N^{r/2-1}} \kappa_r[Y]; (r > 2).$$

Thus, for $r > 2$ we have that $\kappa_r[Z] \to 0$ as $N \to \infty$, that is, $Z \sim N(0,1)$ as $N \to \infty$ (CLT).

Cumulants and transforms

Let $\mu = E_\theta[Y]$ and $\sigma^2 = V_\theta[Y]$ so that $E_\theta[\bar{Y}] = \mu$ and $V_\theta[\bar{Y}] = \sigma^2/N$. Moreover, let $h(\cdot)$ be a (nice) transform. Second order Gaussian approximation (aka the Delta method) gives that

$$E_\theta[h(\bar{Y})] = h(\mu) + \frac{h''(\mu)}{2} \cdot \frac{\sigma^2}{N} + O(N^{-2})$$

$$V_\theta[h(\bar{Y})] = [h'(\mu)]^2 \cdot \frac{\sigma^2}{N} + O(N^{-2})$$

$$\kappa_3[h(\bar{Y})] = [h'(\mu)]^3 \cdot \frac{\kappa_3[Y]}{N^2} + 3[h'(\mu)]^2 \cdot h''(\mu) \cdot \frac{\sigma^4}{N^2} + O(N^{-3})$$

Variance stabilizing transforms

Now, if $([h'(\mu)]^2 \cdot \sigma^2) = a$ (a constant), then

$$V_\theta[h(\bar{Y})] = [h'(\mu)]^2 \cdot \frac{\sigma^2}{N} + O(N^{-2}) = \frac{a}{N} + O(N^{-2}).$$

In other words, the variance of the transformed variable is independent of the mean μ.

Example: Let $Y \sim Po(\mu)$. Then the transform $h(y) = y^{1/2}$ is stabalizing the variance because

$$V_\theta[\sqrt{\bar{Y}}] = V_\theta[h(\bar{Y})] = [\frac{1}{2}\mu^{-1/2}]^2 \cdot \mu + O(N^{-2}) = \frac{1}{4} + O(N^{-2}).$$

Symmetrizing transforms

The skewness of $h(\bar{Y})$ is approximately zero, if

$$[h'(\mu)]^2 \cdot \kappa_3[Y] + 3 \cdot h''(\mu) \cdot \sigma^4 = 0$$

such that the third cumulant becomes

$$\begin{aligned} \kappa_3[h(\bar{Y})] = [h'(\mu)]^3 \cdot \frac{\kappa_3[Y]}{N^2} + 3[h'(\mu)]^2 \cdot h''(\mu) \cdot \frac{\sigma^4}{N^2} + O(N^{-3}) \\ = O(N^{-3}). \end{aligned}$$

Symmetrizing transforms for the Poisson distribution

Example: Let $Y \sim Po(\mu)$ so that $\kappa_r[Y] = \mu; \forall r$. Then

$$[h'(\mu)]^2 \cdot \kappa_3[Y] + 3 \cdot h''(\mu) \cdot \sigma^4 = 0 \Leftrightarrow [h'(\mu)]^2 + 3 \cdot h''(\mu) \cdot \mu = 0$$

One solution is that $h(\mu) = a\mu^{2/3} + b$ where a and b are constants. For $a = 1$ and $b = 0$ the variance of $h(\bar{Y}) = \bar{Y}^{2/3}$ becomes

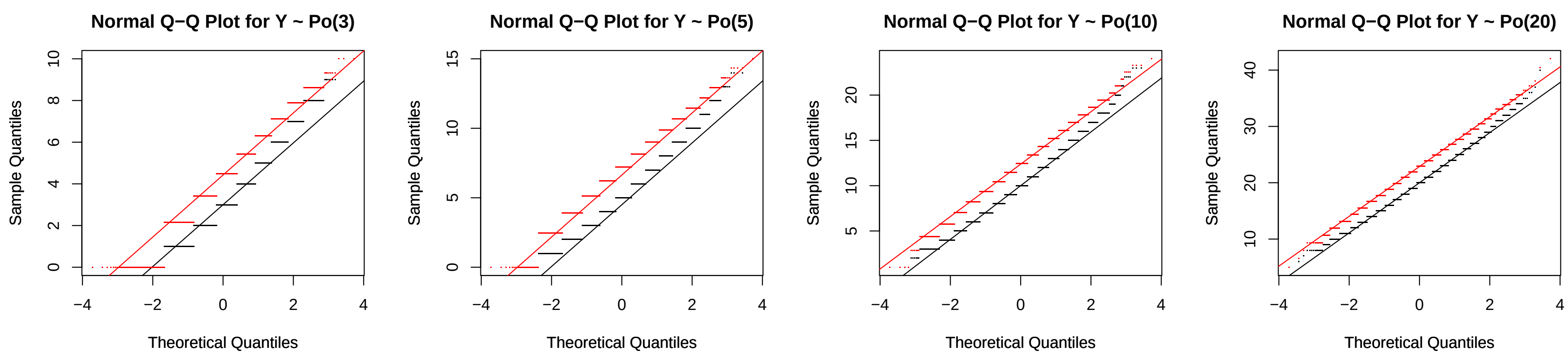
$$V_\theta[h(\bar{Y})] = [h'(\mu)]^2 \cdot \frac{\mu}{N} + O(N^{-2}) = [(2/3)\mu^{-1/3}]^2 \cdot \frac{\mu}{N} + O(N^{-2})$$
$$= (2/3)^2\mu^{1/3} \cdot \frac{1}{N} + O(N^{-2}),$$

or equivalent, the standard deviation becomes

$$D_\theta[h(\bar{Y})] = (2/3)\mu^{1/6}/N + O(N^{-2}).$$

Simulation: Symmetrization of Poisson samples.

Simulation of $N = 5000$ samples from $Y \in Po(\mu)$ with $\mu = 3, 5, 10, 20$ and the symmetrization transform $h(Y) = Y^{2/3}$:



Symmetrizing and variance stabilization Poisson transform

So, for $Y \sim Po(\mu)$, the transform $h(Y) = Y^{2/3}$ "makes" the distribution symmetric ($\kappa_3[h(Y)] \approx 0$) and the the standard deviation "becomes" $D_\theta[h(Y)] \approx (2/3)\mu^{1/6}$. The mean is $E[h(Y)] \approx \mu^{2/3}$.

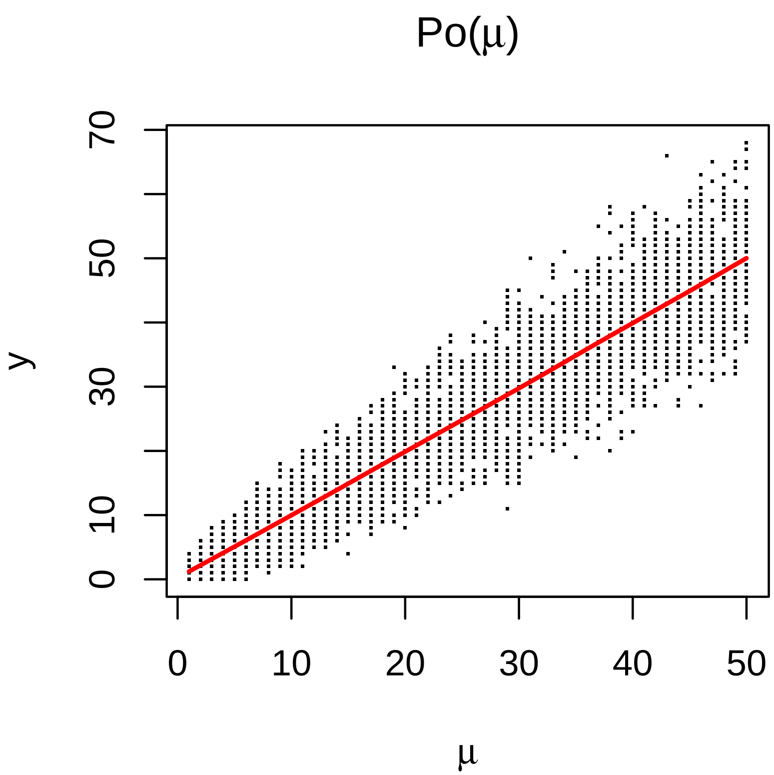
However, the transform

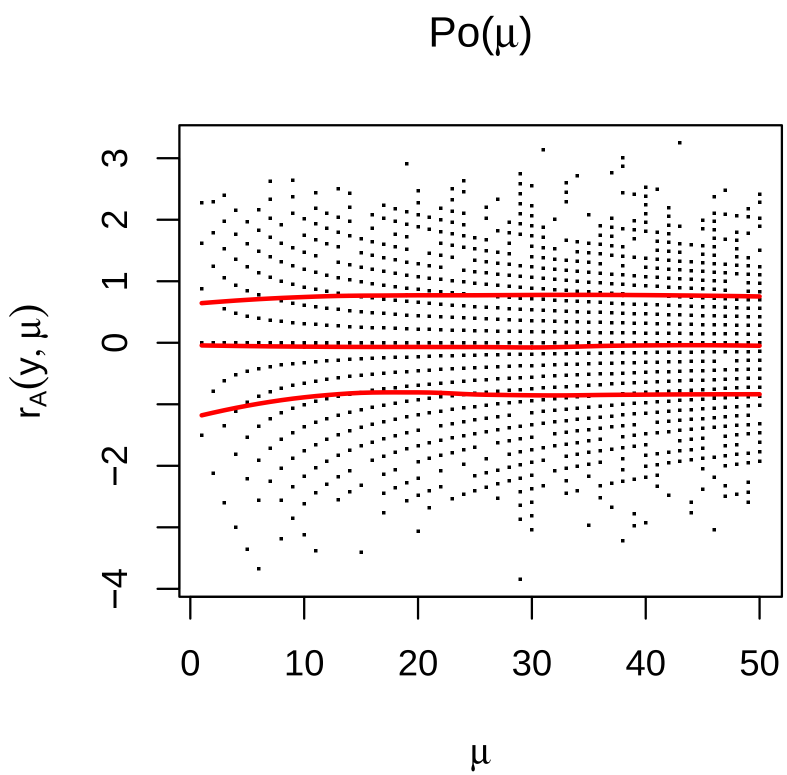
$$Z = Z(Y) = \frac{h(Y) - E[h(Y)]}{D[h(Y)]} \approx \frac{Y^{2/3} - \mu^{2/3}}{\frac{2}{3}\mu^{1/6}}$$

is approximately $Z \sim N(0, 1)$, which is exactly the *Anscombe residuals for the Poisson distribution*;

$$r_A(Y, \mu) = \frac{3}{2} \cdot \frac{Y^{2/3} - \mu^{2/3}}{\mu^{1/6}}.$$

Simulation: Poission distribution and Anscombe residuals





$$E_\mu[Y] = \mu$$
$$V_\mu[Y] = \mu$$
$$S_\mu[Y] = \mu^{-2}$$

$$E_\mu[r_A(Y)] = -0.0312 \approx 0$$
$$V_\mu[r_A(Y)] = 1.019 \approx 1$$
$$S_\mu[r_A(Y)] = -0.00738 \approx 0$$

Exponential family revisited

For an exponential family given by

$$f_Y(y) \propto \exp\{y\theta - b(\theta) + c(y)\}$$

the cumulant generating function of Y is

$$K_Y(t) = \ln\{M_Y(t)\} = b(\theta + t) - b(\theta)$$

so that $\kappa_r[Y] = b^{(k)}(\theta)$. For this reason, the variance can be written as a function of the mean:

$$V_\theta[Y] = \sigma^2 = \sigma^2(\mu) = b''(\theta)((b')^{-1}(\mu))$$

so that, since $b'(\theta) = \kappa_1[Y] = \mu$,

$$\frac{d}{d\mu}\sigma^2(\mu) = \frac{b^{(3)}(\theta(\mu))}{b^{(2)}(\theta(\mu))} = \frac{\kappa_3(\mu)}{\kappa^2(\mu)} = \frac{\kappa_3(\mu)}{\sigma^2(\mu)}.$$

Exponential family revisited...

If we let $V_\theta[Y] = V(\mu) = \sigma^2(\mu)$, then $\kappa_3(\mu) = V(\mu)/V'(\mu)$ and the symmetrizing transformation solves

$$[h'(\mu)]^2 \cdot \kappa_3[Y] + 3 \cdot h''(\mu) \cdot \sigma^4 = 0$$
$$\Leftrightarrow h'(\mu)V'(\mu)V(\mu) + 3h''(\mu)V^2(\mu) = 0$$

or equivalent

$$\frac{d}{d\mu}\left[(h'(\mu))^3 V(\mu)\right] = 0.$$

A solution to this equation is

$$h(u) = \int_{-\infty}^{u} \frac{1}{[V(\theta)]^{1/3}} d\theta.$$

This is what Wedderburn (19??) showed hold for all distributions in GLM, i.e. exponential family distributions.

Poisson distribution revisited

For $Y \in Po(\mu)$ we have that $E_\mu[Y] = \mu$ and $V_\mu[Y] = \mu$. Thus according to Wedderburn's result the transform that makes $h(Y)$ most Normal is

$$h(u) = \int_0^u \frac{1}{\theta^{1/3}} d\theta = \left[\frac{3}{2}\theta^{2/3}\right]_0^u = \frac{2}{3}u^{2/3}.$$

We have showed that this transform is not stabilizing the variance; recall $D_\theta[h(Y)] \approx (2/3)\mu^{1/6}$, but dividing by the standard deviation the Anscombe residuals obtain this

$$r_A(Y, \mu) = \frac{3}{2} \cdot \frac{Y^{2/3} - \mu^{2/3}}{\mu^{1/6}}.$$

For Further Reading

Donald A. Pierce and Daniel W. Schafer.
Residuals in Generalized Linear Models.
Journal of American Statistical Association, Vol. 81, No. 396, Theory and Methods, December 1986