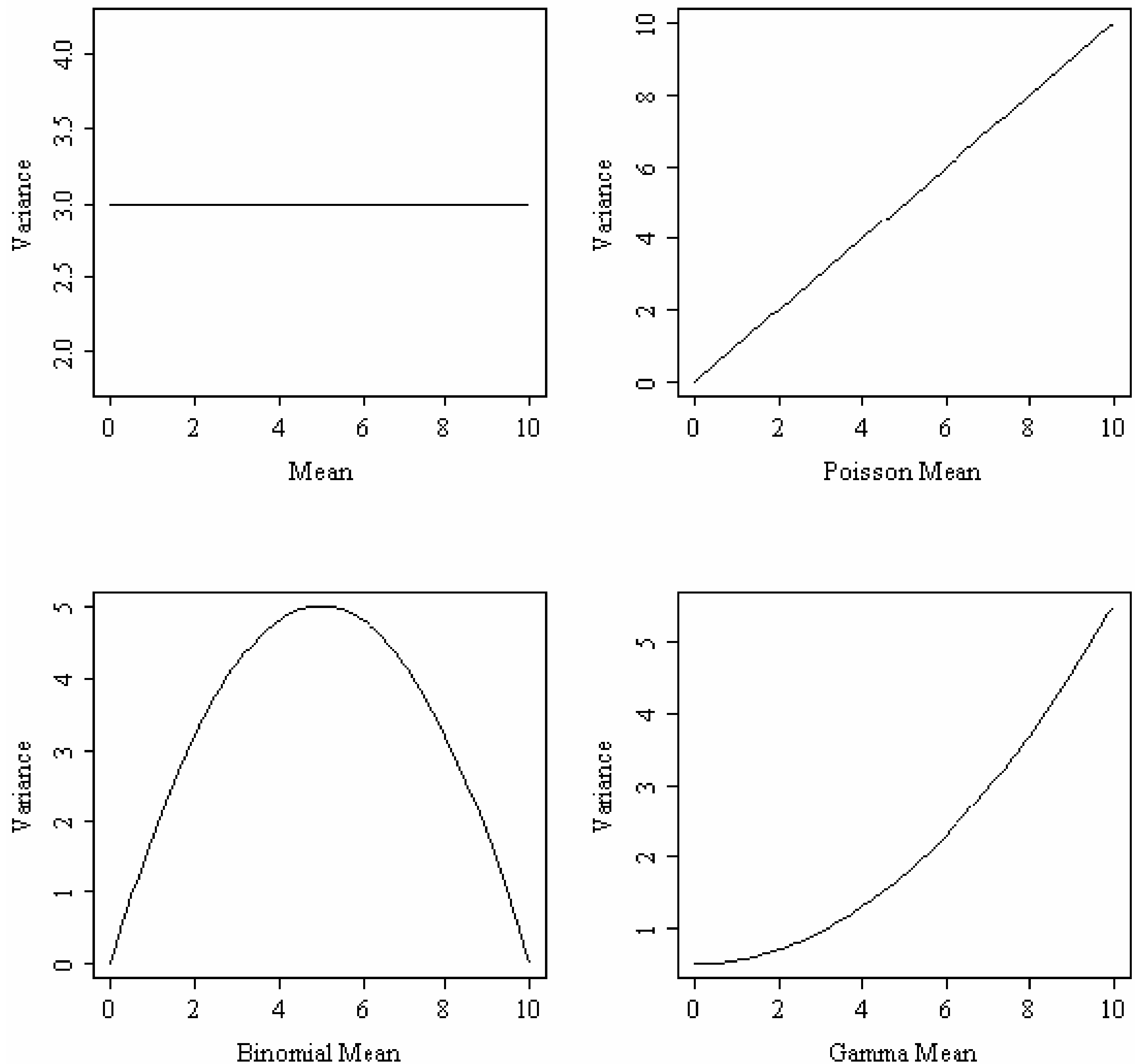
Modelos Lineares Generalizados

- Histórico

$$f(y;\theta,\phi) = exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right\} \qquad (3.1)$$



Variance
Mean
Variance
Poisson Mean
Variance
Binomial Mean
Variance
Gamma Mean

a) <u>Distribuição normal</u>

$$f(y;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(y-\mu)^2}{2\sigma^2}\right]$$

$$= \exp\left[\frac{y\mu - \dfrac{\mu^2}{2}}{\sigma^2} - \frac{1}{2}\ln\left(2\pi\sigma^2\right) - \frac{y^2}{2\sigma^2}\right]$$

Comparando com a Equação (3.1) obtemos:

$$\theta = \mu; \quad b(\theta) = \frac{\mu^2}{2}; \quad a(\phi) = \phi; \quad \phi = \sigma^2; \quad c(y,\phi) = -\frac{1}{2}\ln\left(2\pi\sigma^2\right) + \frac{y^2}{2\sigma^2}$$

Média e variância de *y*:

$$E(y) = \frac{db(\theta)}{d\theta} = \mu; \qquad \mathrm{var}(y) = \frac{d^2 b(\theta)}{d\theta^2} a(\phi) = \sigma^2$$

Algumas distribuições da família exponencial

distribuição	normal	binomial	Poisson	gama	gaussiana inversa
Notação	$N(\mu, \sigma^2)$	$B(m, \pi)/m$	$P(\lambda)$	$Ga(\nu, \frac{\nu}{\mu})$	$IG(\mu, \sigma^2)$
Suporte	$(-\infty, +\infty)$	$\{0, \frac{1}{m}, ..., 1\}$	$\{0, 1, ...\}$	$(0, +\infty)$	$(0, +\infty)$
θ	μ	$\ln\left(\frac{\pi}{1-\pi}\right)$	$\ln\lambda$	$-\frac{1}{\mu}$	$-\frac{1}{2\mu^2}$
$a(\phi)$	σ^2	$\frac{1}{m}$	1	$\frac{1}{\nu}$	σ^2
ϕ	σ^2	1	1	$\frac{1}{\nu}$	σ^2
ω	1	m	1	1	1
$c(y, \phi)$	$-\frac{1}{2}(\frac{y^2}{\phi} + \ln(2\pi\phi))$	$\ln\binom{m}{my}$	$-\ln y!$	$\nu\ln\nu - \ln\Gamma(\nu)$ $+(\nu-1)\ln y$	$-\frac{1}{2}\{\ln(2\pi\phi y^3)$ $+\frac{1}{y\phi}\}$
$b(\theta)$	$\frac{\theta^2}{2}$	$\ln(1+e^\theta)$	e^θ	$-\ln(-\theta)$	$-(-2\theta)^{1/2}$
$b'(\theta) = E(Y)$	θ	$\pi = \frac{e^\theta}{1+e^\theta}$	$\lambda = e^\theta$	$\mu = -\frac{1}{\theta}$	$\mu = (-2\theta)^{-1/2}$
$b''(\theta) = V(\mu)$	1	$\pi(1-\pi)$	λ	μ^2	μ^3
$var(Y)$	σ^2	$\frac{\pi(1-\pi)}{m}$	λ	$\frac{\mu^2}{\nu}$	$\mu^3\sigma^2$

O que é um modelo “linear”?

- Uma equação que contem
  - variáveis
  - parâmetros
  - resíduos aleatórios

.... que é linear nos parâmetros e variáveis aleatórias

O que é linear?

$Y = a + bx$

$Y = a + bx + cx^2$

$z = x^2$

$Y = a + bx + zx$

“Um” preditor linear
(linear predictor)

$$\eta_i = \sum_{j=1}^{p} x_{ib} \beta_j$$

Função de ligação
(*link function)*

$$\eta = g(\mu)$$

Função de ligação canônica para algumas distribuições da família exponencial

Distribuição	Ligação Canônica
normal	$\eta = \mu$
Poisson	$\eta = \ln \mu$
binomial	$\eta = \ln(\pi / (1 - \pi))$
gama	$\eta = 1 / \mu$
normal inversa	$\eta = 1 / \mu^2$

Desviança
(deviance)

$$\textbf{deviance} = -2\log_e(\mathcal{L}(\hat{\theta})) + 2\log_e(\mathcal{L}_s(\hat{\theta})),$$

$$= -2\left(\log_e(\mathcal{L}(\hat{\theta})) - \log_e(\mathcal{L}_s(\hat{\theta}))\right),$$

Family (Error structure)	**Deviance**
Normal	$\sum (y-\bar{y})^2$
Poisson	$2\sum y \ln(y/\mu) - (y-\mu)$
Binomial	$2\sum y \ln(y/\mu) + (n-y)\ln(n-y)/(n-\mu)$
Gamma	$2\sum (y-\mu)/y - \ln(y/\mu)$
Inverse Gaussian	$\sum (y-\mu)^2/(\mu^2 y)$