A geometric view of overdispersion

Karim Anaya-Izquierdo (The Open University)

joint work with Frank Critchley (Open University),
Paul Marriott (Waterloo) and Paul Vos (East Carolina)

WOGAS 3, Warwick 2011

Introduction

- Overdispersion is common in practice

Introduction

- Overdispersion is common in practice
- Target: Overdispersed exponential families (Binomial)

Introduction

- Overdispersion is common in practice
- Target: Overdispersed exponential families (Binomial)
- Models for overdispersed data can be expressed in terms of simple geometrical operations

Overdispersion

- Exponential family

$$f(y;\mu) = \exp(\theta(\mu)y - \psi(\theta(\mu)))\,\nu(y)$$

Overdispersion

- Exponential family

$$f(y;\mu) = \exp(\theta(\mu)y - \psi(\theta(\mu)))\,\nu(y)$$

where $E[Y] = \mu = \psi'(\theta(\mu))$

Overdispersion

- Exponential family

$$f(y;\mu) = \exp(\theta(\mu)y - \psi(\theta(\mu)))\,\nu(y)$$

where $E[Y] = \mu = \psi'(\theta(\mu))$ and $\nu(y)$ is the base measure

Overdispersion

- Exponential family

$$f(y;\mu) = \exp(\theta(\mu)y - \psi(\theta(\mu)))\,\nu(y)$$

where $E[Y] = \mu = \psi'(\theta(\mu))$ and $\nu(y)$ is the base measure

- The variance function

$$Var(Y) = \psi''(\theta(\mu)) = v(\mu)$$

so the variance depends on the mean

Overdispersion

- Exponential family

$$f(y;\mu) = \exp(\theta(\mu)y - \psi(\theta(\mu)))\,\nu(y)$$

  where $E[Y] = \mu = \psi'(\theta(\mu))$ and $\nu(y)$ is the base measure

- The variance function

$$Var(Y) = \psi''(\theta(\mu)) = v(\mu)$$

  so the variance depends on the mean

- Usually we observe in data that nominal variance is larger/smaller than one expected by exponential family

Overdispersion

- Exponential family

$$f(y;\mu) = \exp(\theta(\mu)y - \psi(\theta(\mu)))\,\nu(y)$$

where $E[Y] = \mu = \psi'(\theta(\mu))$ and $\nu(y)$ is the base measure

- The variance function

$$Var(Y) = \psi''(\theta(\mu)) = v(\mu)$$

so the variance depends on the mean

- Usually we observe in data that nominal variance is larger/smaller than one expected by exponential family
- We can proceed depending of type/cause of overdispersion by modelling a more flexible variance function eg by adding extra parameters

Overdispersion

- Need a distribution where a single parameter controls dispersion

Overdispersion

- Need a distribution where a single parameter controls dispersion
- Prototypical example:

Overdispersion

- Need a distribution where a single parameter controls dispersion
- Prototypical example: Normal distribution $N(\mu, \sigma^2)$!

Overdispersion

- Need a distribution where a single parameter controls dispersion
- Prototypical example: Normal distribution $N(\mu, \sigma^2)$!
- Generally: Exponential dispersion models

$$g(y; \theta, \psi) = \exp((y\theta - \psi(\theta))/\phi)\, \nu_\phi(y)$$

Overdispersion

- Need a distribution where a single parameter controls dispersion
- Prototypical example: Normal distribution $N(\mu, \sigma^2)$!
- Generally: Exponential dispersion models

$$g(y; \theta, \psi) = \exp((y\theta - \psi(\theta))/\phi)\, \nu_\phi(y)$$

$$\begin{aligned} E_g[Y] &= \mu = \psi'(\theta) \\ Var_g[Y] &= \phi\, \psi''(\theta(\mu)) = \phi\, v(\mu) \end{aligned}$$

Overdispersion

- Need a distribution where a single parameter controls dispersion
- Prototypical example: Normal distribution $N(\mu, \sigma^2)$!
- Generally: Exponential dispersion models

$$g(y;\theta,\psi) = \exp((y\theta - \psi(\theta))/\phi)\, \nu_\phi(y)$$

$$\begin{aligned} E_g[Y] &= \mu = \psi'(\theta) \\ Var_g[Y] &= \phi\, \psi''(\theta(\mu)) = \phi\, v(\mu) \end{aligned}$$

- Do not exist for discrete data! Jorgensen (1997)

Binomial $Y \sim Bin(n, \mu)$

- $f(y; \mu, n) = \binom{n}{x} n^{-n} \mu^y (n - \mu)^{n-y}, \quad v(y) = \binom{n}{y}$

$$
\begin{aligned}
E_f[Y] &= \mu \\
Var_f[Y] &= \frac{\mu(n-\mu)}{n} = v(\mu)
\end{aligned}
$$

Binomial $Y \sim Bin(n, \mu)$

- $f(y; \mu, n) = \binom{n}{x} n^{-n} \mu^y (n - \mu)^{n-y}, \quad \nu(y) = \binom{n}{y}$

$$E_f[Y] = \mu$$
$$Var_f[Y] = \frac{\mu(n-\mu)}{n} = v(\mu)$$

- Overdispersion by mixing:

Binomial $Y \sim Bin(n, \mu)$

- $f(y; \mu, n) = \binom{n}{x} n^{-n} \mu^{y} (n-\mu)^{n-y}, \quad \nu(y) = \binom{n}{y}$

$$
\begin{aligned}
E_f[Y] &= \mu \\
Var_f[Y] &= \frac{\mu(n-\mu)}{n} = v(\mu)
\end{aligned}
$$

- Overdispersion by mixing: $Y \mid M \sim Bin(n, M)$

Binomial $Y \sim Bin(n, \mu)$

- $f(y; \mu, n) = \binom{n}{x} n^{-n} \mu^{y} (n - \mu)^{n-y}, \quad v(y) = \binom{n}{y}$

$$
\begin{aligned}
E_f[Y] &= \mu \\
Var_f[Y] &= \frac{\mu(n-\mu)}{n} = v(\mu)
\end{aligned}
$$

- Overdispersion by mixing: $\quad Y \mid M \sim Bin(n, M)$

$$M \sim H \quad \text{such that} \quad E_H[M] = \mu \quad Var_H[M] = \tau^2 \mu (n - \mu)$$

Binomial $Y \sim Bin(n, \mu)$

- $f(y; \mu, n) = \binom{n}{x} n^{-n} \mu^{y} (n-\mu)^{n-y}, \quad \nu(y) = \binom{n}{y}$

$$E_f[Y] = \mu$$
$$Var_f[Y] = \frac{\mu(n-\mu)}{n} = v(\mu)$$

- Overdispersion by mixing: $\quad Y \mid M \sim Bin(n, M)$

$$M \sim H \quad \text{such that} \quad E_H[M] = \mu \quad Var_H[M] = \tau^2 \mu (n-\mu)$$

marginally

Binomial $Y \sim Bin(n, \mu)$

- $f(y;\mu,n) = \binom{n}{x} n^{-n} \mu^{y} (n-\mu)^{n-y}, \quad \nu(y) = \binom{n}{y}$

$$E_f[Y] = \mu$$
$$Var_f[Y] = \frac{\mu(n-\mu)}{n} = v(\mu)$$

- Overdispersion by mixing: $\quad Y \mid M \sim Bin(n, M)$

$$M \sim H \quad \text{such that} \quad E_H[M] = \mu \quad Var_H[M] = \tau^2 \mu (n-\mu)$$

marginally $g(x;\mu,\tau^2) = \int_0^n f(x;m,n) dH(m)$

Binomial $Y \sim Bin(n, \mu)$

- $f(y; \mu, n) = \binom{n}{x} n^{-n} \mu^y (n-\mu)^{n-y}, \quad \nu(y) = \binom{n}{y}$

$$\begin{aligned} E_f[Y] &= \mu \\ Var_f[Y] &= \frac{\mu(n-\mu)}{n} = v(\mu) \end{aligned}$$

- Overdispersion by mixing: $\quad Y \mid M \sim Bin(n, M)$

$M \sim H$ such that $E_H[M] = \mu \quad Var_H[M] = \tau^2 \mu (n-\mu)$

marginally $g(x; \mu, \tau^2) = \int_0^n f(x; m, n) dH(m)$

$$\begin{aligned} E_g[Y] &= \mu \\ Var_g[Y] &= v(\mu)[1 + (n-1)\tau^2] \end{aligned}$$

Binomial $Y \sim Bin(n, \mu)$

- $f(y;\mu,n) = \binom{n}{x} n^{-n} \mu^{y} (n-\mu)^{n-y}, \quad \nu(y) = \binom{n}{y}$

$$\begin{aligned} E_f[Y] &= \mu \\ Var_f[Y] &= \frac{\mu(n-\mu)}{n} = v(\mu) \end{aligned}$$

- Overdispersion by mixing: $\quad Y \mid M \sim Bin(n, M)$

$M \sim H \quad$ such that $\quad E_H[M] = \mu \quad Var_H[M] = \tau^2 \mu (n-\mu)$

marginally $g(x;\mu,\tau^2) = \int_0^n f(x;m,n) dH(m)$

$$\begin{aligned} E_g[Y] &= \mu \\ Var_g[Y] &= v(\mu)[1+(n-1)\tau^2] \end{aligned}$$

Example: Beta-Binomial

Binomial

Extension

$$f(y;\mu) \longrightarrow g(y;\mu,\phi)$$

Binomial

Extension

$$f(y;\mu) \longrightarrow g(y;\mu,\phi)$$

so that

$$g(y;\mu,0) \quad = \quad f(y;\mu)$$

Binomial

Extension

$$f(y;\mu) \longrightarrow g(y;\mu,\phi)$$

so that

$$\begin{aligned} g(y;\mu,0) &= f(y;\mu) \\ E_f[Y] &= E_g[Y] = \mu \end{aligned}$$

Binomial

Extension

$$f(y;\mu) \to g(y;\mu,\phi)$$

so that

$$\begin{aligned} g(y;\mu,0) &= f(y;\mu) \\ E_f[Y] &= E_g[Y] = \mu \\ Var_g[Y] &\geq Var_f[Y] = v(\mu) \qquad \forall \mu \end{aligned}$$

Binomial

Extension

$$f(y;\mu) \rightarrow g(y;\mu,\phi)$$

so that

$$\begin{aligned} g(y;\mu,0) &= f(y;\mu) \\ E_f[Y] &= E_g[Y] = \mu \\ Var_g[Y] &\geq Var_f[Y] = v(\mu) \qquad \forall\mu \end{aligned}$$

How do we extend in general?

Binomial

Extension

$$f(y;\mu) \to g(y;\mu,\phi)$$

so that

$$\begin{aligned} g(y;\mu,0) &= f(y;\mu) \\ E_f[Y] &= E_g[Y] = \mu \\ Var_g[Y] &\geq Var_f[Y] = v(\mu) \qquad \forall\mu \end{aligned}$$

How do we extend in general?

- Flexibility

Binomial

Extension

$$f(y;\mu) \rightarrow g(y;\mu,\phi)$$

so that

$$\begin{aligned} g(y;\mu,0) &= f(y;\mu) \\ E_f[Y] &= E_g[Y] = \mu \\ Var_g[Y] &\geq Var_f[Y] = v(\mu) \qquad \forall \mu \end{aligned}$$

How do we extend in general?

- Flexibility
- Parsimony

Binomial

Extension

$$f(y;\mu) \rightarrow g(y;\mu,\phi)$$

so that

$$\begin{aligned} g(y;\mu,0) &= f(y;\mu) \\ E_f[Y] &= E_g[Y] = \mu \\ Var_g[Y] &\geq Var_f[Y] = v(\mu) \qquad \forall \mu \end{aligned}$$

How do we extend in general?

- Flexibility
- Parsimony

Use simple dual affine geometry!!

Binomial

Extension

$$f(y;\mu) \rightarrow g(y;\mu,\phi)$$

so that

$$\begin{array}{rcl} g(y;\mu,0) & = & f(y;\mu) \\ E_f[Y] & = & E_g[Y] = \mu \\ Var_g[Y] & \geq & Var_f[Y] = v(\mu) \qquad \forall \mu \end{array}$$

How do we extend in general?

- Flexibility
- Parsimony

Use simple dual affine geometry!! -1 type and -1 type extensions

Overdispersed Binomials: $+1$ type extensions

$$
\begin{aligned}
f(y;\mu) &= \exp(y\,\theta(\mu) - \psi(\theta(\mu)))\,\nu(y) \\
&\downarrow \\
g(y;\mu,\phi) &= \exp(y\,\eta + T(y)\phi - \kappa(\eta,\phi))\,\nu(y)
\end{aligned}
$$

where $\mu = E_g[Y]$.

Overdispersed Binomials: $+1$ type extensions

$$\begin{array}{rcl} f(y;\mu) & = & \exp(y\,\theta(\mu) - \psi(\theta(\mu)))\,\nu(y) \\ & \downarrow & \\ g(y;\mu,\phi) & = & \exp(y\,\eta + T(y)\phi - \kappa(\eta,\phi))\,\nu(y) \end{array}$$

where $\mu = E_g[Y]$. Examples

Overdispersed Binomials: $+1$ type extensions

$$
\begin{aligned}
f(y;\mu) &= \exp(y\,\theta(\mu) - \psi(\theta(\mu)))\,\nu(y) \\
&\downarrow \\
g(y;\mu,\phi) &= \exp(y\,\eta + T(y)\phi - \kappa(\eta,\phi))\,\nu(y)
\end{aligned}
$$

where $\mu = E_g[Y]$. Examples

Family	$T(y)$
Double Binomial	$y\log(y/n) + (n-y)\log((n-y)/n)$
Multiplicative	$-x(n-x)$
Shrink base measure	$-\log\binom{n}{x}$

Overdispersed Binomials: $+1$ type extensions

- Locally to $\phi = 0$

$$Var(Y; \mu, \eta) = v(\mu) + \phi \, \xi(\mu)$$

where

$$\xi(\mu) = cov_f(T(Y), Y^2) - \frac{cov_f(Y, Y^2) cov_f(T(Y), Y)}{Var_f(Y)}$$

Overdispersed Binomials: $+1$ type extensions

- Locally to $\phi = 0$

$$Var(Y; \mu, \eta) = v(\mu) + \phi\, \xi(\mu)$$

where

$$\xi(\mu) = cov_f(T(Y), Y^2) - \frac{cov_f(Y, Y^2) cov_f(T(Y), Y)}{Var_f(Y)}$$

- If T(y) convex then overdispersion (Gelfand & Dalal 1990, Lindsay, 1986)

Overdispersed Binomials: $+1$ type extensions

- Locally to $\phi = 0$

$$Var(Y; \mu, \eta) = v(\mu) + \phi\, \xi(\mu)$$

where

$$\xi(\mu) = cov_f(T(Y), Y^2) - \frac{cov_f(Y, Y^2)cov_f(T(Y), Y)}{Var_f(Y)}$$

- If T(y) convex then overdispersion (Gelfand & Dalal 1990, Lindsay, 1986)
- Parameters η and $\mu = E_g(Y)$ are orthogonal (mixed parametrisation)

Double Binomial (Efron, 1986)

Consider the +1 joining of $f(y;\mu)$ and $f(y;y)$

$$g(y;\mu,\lambda) = c(\mu,\lambda)[f(y;\mu)]^{\lambda}[f(y;y)]^{1-\lambda}$$

Double Binomial (Efron, 1986)

Consider the +1 joining of $f(y;\mu)$ and $f(y;y)$

$$g(y;\mu,\lambda) = c(\mu,\lambda)[f(y;\mu)]^{\lambda}[f(y;y)]^{1-\lambda}$$

which can be written as

$$g(y;\mu,\phi) = \exp(y\eta + T(y)\phi - \kappa(\eta,\phi))$$

where

$$\phi = 1 - \lambda$$

$$\eta = \lambda\,\theta(\mu)$$

Double Binomial (Efron, 1986)

Consider the +1 joining of $f(y;\mu)$ and $f(y;y)$

$$g(y;\mu,\lambda) = c(\mu,\lambda)[f(y;\mu)]^{\lambda}[f(y;y)]^{1-\lambda}$$

which can be written as

$$g(y;\mu,\phi) = \exp(y\eta + T(y)\phi - \kappa(\eta,\phi))$$

where

$$\phi = 1-\lambda$$

$$\eta = \lambda\,\theta(\mu)$$

Interpretation

Double Binomial (Efron, 1986)

Consider the +1 joining of $f(y;\mu)$ and $f(y;y)$

$$g(y;\mu,\lambda) = c(\mu,\lambda)[f(y;\mu)]^{\lambda}[f(y;y)]^{1-\lambda}$$

which can be written as

$$g(y;\mu,\phi) = \exp(y\eta + T(y)\phi - \kappa(\eta,\phi))$$

where

$$\phi = 1 - \lambda$$

$$\eta = \lambda\,\theta(\mu)$$

Interpretation

- $\phi > 0$ overdispersion

Double Binomial (Efron, 1986)

Consider the +1 joining of $f(y;\mu)$ and $f(y;y)$

$$g(y;\mu,\lambda) = c(\mu,\lambda)[f(y;\mu)]^{\lambda}[f(y;y)]^{1-\lambda}$$

which can be written as

$$g(y;\mu,\phi) = \exp(y\eta + T(y)\phi - \kappa(\eta,\phi))$$

where

$$\phi = 1 - \lambda$$

$$\eta = \lambda\,\theta(\mu)$$

Interpretation

- $\phi > 0$ overdispersion
- $\phi < 0$ underdispersion

Shrinking base measure, (Shmueli, et al 2005)

- Consider

$$
\begin{aligned}
g(y;\mu,\lambda) &= \binom{n}{y}^{\lambda} \exp(y\eta - \kappa(\lambda,\eta)) \\
&= \binom{n}{y} \exp(y\eta + \phi T(y) - \kappa(\eta,\phi))
\end{aligned}
$$

Shrinking base measure, (Shmueli, et al 2005)

- Consider

$$
\begin{aligned}
g(y;\mu,\lambda) &= \binom{n}{y}^{\lambda} \exp(y\eta - \kappa(\lambda,\eta)) \\
&= \binom{n}{y} \exp(y\eta + \phi T(y) - \kappa(\eta,\phi))
\end{aligned}
$$

where $\phi = 1 - \lambda < 1$ and $T(y) = -\log \binom{n}{y}$

Shrinking base measure, (Shmueli, et al 2005)

- Consider

$$\begin{aligned} g(y;\mu,\lambda) &= \binom{n}{y}^{\lambda} \exp(y\eta - \kappa(\lambda,\eta)) \\ &= \binom{n}{y} \exp(y\eta + \phi T(y) - \kappa(\eta,\phi)) \end{aligned}$$

  where $\phi = 1 - \lambda < 1$ and $T(y) = -\log \binom{n}{y}$

- Interpretation

Shrinking base measure, (Shmueli, et al 2005)

- Consider

$$g(y;\mu,\lambda) = \binom{n}{y}^{\lambda} \exp(y\eta - \kappa(\lambda,\eta))$$
$$= \binom{n}{y} \exp(y\eta + \phi T(y) - \kappa(\eta,\phi))$$

where $\phi = 1 - \lambda < 1$ and $T(y) = -\log \binom{n}{y}$

- Interpretation
  - $\phi > 0$ $(\lambda < 1)$ overdispersion

Shrinking base measure, (Shmueli, et al 2005)

- Consider

$$
\begin{aligned}
g(y;\mu,\lambda) &= \binom{n}{y}^{\lambda} \exp(y\eta - \kappa(\lambda,\eta)) \\
&= \binom{n}{y} \exp(y\eta + \phi T(y) - \kappa(\eta,\phi))
\end{aligned}
$$

where $\phi = 1 - \lambda < 1$ and $T(y) = -\log \binom{n}{y}$

- Interpretation
  - $\phi > 0$ $(\lambda < 1)$ overdispersion
  - $\phi < 0$ $(\lambda > 1)$ underdispersion

Multiplicative Binomial (Altham, 1978)

- Consider

$$g(y;\mu,\phi) = \binom{n}{y} \exp(\eta y + \phi T(y) - \kappa(\eta,\phi))$$

where $T(y) = -y(n-y)$

Multiplicative Binomial (Altham, 1978)

- Consider

$$g(y; \mu, \phi) = \binom{n}{y} \exp(\eta y + \phi T(y) - \kappa(\eta, \phi))$$

where $T(y) = -y(n - y)$

- Distribution of the sum of n dependent binary variables with symmetric joint distribution and no second or higher order multiplicative interactions

Multiplicative Binomial (Altham, 1978)

- Consider

$$g(y;\mu,\phi) = \binom{n}{y} \exp(\eta y + \phi T(y) - \kappa(\eta,\phi))$$

where $T(y) = -y(n-y)$

- Distribution of the sum of n dependent binary variables with symmetric joint distribution and no second or higher order multiplicative interactions
- $Var[Y] = v(\mu) + (m-1)r(\mu,\phi)$

Multiplicative Binomial (Altham, 1978)

- Consider

$$g(y; \mu, \phi) = \binom{n}{y} \exp(\eta y + \phi T(y) - \kappa(\eta, \phi))$$

  where $T(y) = -y(n - y)$

- Distribution of the sum of n dependent binary variables with symmetric joint distribution and no second or higher order multiplicative interactions
- $Var[Y] = v(\mu) + (m-1) r(\mu, \phi)$
- Interpretation

Multiplicative Binomial (Altham, 1978)

- Consider

$$g(y;\mu,\phi) = \binom{n}{y} \exp(\eta y + \phi T(y) - \kappa(\eta,\phi))$$

where $T(y) = -y(n-y)$

- Distribution of the sum of n dependent binary variables with symmetric joint distribution and no second or higher order multiplicative interactions
- $Var[Y] = v(\mu) + (m-1)r(\mu,\phi)$
- Interpretation
  - $\phi > 0$ (positive association) overdispersion

Multiplicative Binomial (Altham, 1978)

- Consider

$$g(y;\mu,\phi) = \binom{n}{y} \exp(\eta y + \phi T(y) - \kappa(\eta,\phi))$$

  where $T(y) = -y(n-y)$
- Distribution of the sum of n dependent binary variables with symmetric joint distribution and no second or higher order multiplicative interactions
- $Var[Y] = v(\mu) + (m-1)r(\mu,\phi)$
- Interpretation
  - $\phi > 0$ (positive association) overdispersion
  - $\phi < 0$ (negative association) underdispersion

Overdispersed Binomials: $+1$ type extension

Let now $T(y)$ to depend on μ

$$
\begin{aligned}
f(y;\mu) &= \exp(y\,\theta(\mu) - \psi(\theta(\mu)))\,\nu(y) \\
&\downarrow \\
g(y;\mu,\phi) &= \exp(y\,\eta + T(y,\mu)\phi - \kappa(\eta,\phi))\,\nu(y)
\end{aligned}
$$

Overdispersed Binomials: +1 type extension

Let now $T(y)$ to depend on μ

$$
\begin{aligned}
f(y;\mu) &= \exp(y\,\theta(\mu) - \psi(\theta(\mu)))\,\nu(y) \\
&\downarrow \\
g(y;\mu,\phi) &= \exp(y\,\eta + T(y,\mu)\phi - \kappa(\eta,\phi))\,\nu(y)
\end{aligned}
$$

Examples

Family	$T(y)$
Overdispersion Test	$\dfrac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}$
Copas & Eguchi (2005)	semiparametric

Overdispersion test (Cox, 1983, Rayner & Best, 2009)

- $g(y;\mu,\phi)$ with

$$T(y;\mu) = \frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}$$

is a local mixture in the sense

Overdispersion test (Cox, 1983, Rayner & Best, 2009)

- $g(y; \mu, \phi)$ with

$$T(y; \mu) = \frac{d^2 f(x; \theta(\mu))/d\mu^2}{f(x; \theta(\mu))}$$

is a local mixture in the sense

$$\int f(x; m) dH(m) \approx g(x; \mu, \phi)$$

for small mixing distribution H s.t $E_H[M] = \mu$

Overdispersion test (Cox, 1983, Rayner & Best, 2009)

- $g(y;\mu,\phi)$ with

$$T(y;\mu) = \frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}$$

is a local mixture in the sense

$$\int f(x;m)dH(m) \approx g(x;\mu,\phi)$$

for small mixing distribution H s.t $E_H[M] = \mu$

- Neyman's smooth goodness of fit tests. Locally most powerful directed tests .

Overdispersion test (Cox, 1983, Rayner & Best, 2009)

- $g(y; \mu, \phi)$ with

$$T(y; \mu) = \frac{d^2 f(x; \theta(\mu))/d\mu^2}{f(x; \theta(\mu))}$$

is a local mixture in the sense

$$\int f(x; m) dH(m) \approx g(x; \mu, \phi)$$

for small mixing distribution H s.t $E_H[M] = \mu$

- Neyman's smooth goodness of fit tests. Locally most powerful directed tests .
- Interpretation

Overdispersion test (Cox, 1983, Rayner & Best, 2009)

- $g(y; \mu, \phi)$ with

$$T(y; \mu) = \frac{d^2 f(x; \theta(\mu)) / d\mu^2}{f(x; \theta(\mu))}$$

is a local mixture in the sense

$$\int f(x; m) dH(m) \approx g(x; \mu, \phi)$$

for small mixing distribution H s.t $E_H[M] = \mu$

- Neyman's smooth goodness of fit tests. Locally most powerful directed tests .
- Interpretation
  - $\phi > 0$ overdispersion

Overdispersion test (Cox, 1983, Rayner & Best, 2009)

- $g(y;\mu,\phi)$ with

$$T(y;\mu) = \frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}$$

  is a local mixture in the sense

$$\int f(x;m)dH(m) \approx g(x;\mu,\phi)$$

  for small mixing distribution H s.t $E_H[M] = \mu$
- Neyman's smooth goodness of fit tests. Locally most powerful directed tests .
- Interpretation
  - $\phi > 0$ overdispersion
  - $\phi < 0$ underspersion

Overdispersed Binomials: -1 type extensions

A different approach is to use local-mixture-like expressions such as

$$
\begin{aligned}
f(y;\mu) &= \exp(y\,\theta(\mu) - \psi(\theta(\mu)))\,\nu(y) \\
&\downarrow \\
g(y;\mu,\phi) &= f(y;\mu)[1+\phi\,T(y,\theta)]
\end{aligned}
$$

Overdispersed Binomials: -1 type extensions

A different approach is to use local-mixture-like expressions such as

$$\begin{array}{rcl} f(y;\mu) & = & \exp(y\,\theta(\mu) - \psi(\theta(\mu)))\,\nu(y) \\ & \downarrow & \\ g(y;\mu,\phi) & = & f(y;\mu)[1+\phi\, T(y,\theta)] \end{array}$$

Examples

Family	$T(y)$
Local Mixture order 2	$\dfrac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}$
Correlated Binomial	ditto

Binomial Local mixture: Anaya-Izquierdo & Marriott(2007)

- Modified Laplace expansion

$$\int f(y;m)\,dH(m) \sim g(y;\mu,\phi) = f(x;\mu)\left[1+\phi\frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}\right]$$

Binomial Local mixture: Anaya-Izquierdo & Marriott(2007)

- Modified Laplace expansion

$$\int f(y;m)\,dH(m) \sim g(y;\mu,\phi) = f(x;\mu)\left[1+\phi\frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}\right]$$

- $var[Y] = v(\mu) + \frac{(n-1)}{n}\phi$

Binomial Local mixture: Anaya-Izquierdo & Marriott(2007)

- Modified Laplace expansion

$$\int f(y;m)\,dH(m) \sim g(y;\mu,\phi) = f(x;\mu)\left[1+\phi\frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}\right]$$

- $var[Y] = v(\mu) + \frac{(n-1)}{n}\phi$
- μ and ϕ are orthogonal

Binomial Local mixture: Anaya-Izquierdo & Marriott(2007)

- Modified Laplace expansion

$$\int f(y;m)\,dH(m) \sim g(y;\mu,\phi) = f(x;\mu)\left[1+\phi\frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}\right]$$

- $var[Y] = v(\mu) + \frac{(n-1)}{n}\phi$
- μ and ϕ are orthogonal
- Interpretation

Binomial Local mixture: Anaya-Izquierdo & Marriott(2007)

- Modified Laplace expansion

$$\int f(y;m)\,dH(m) \sim g(y;\mu,\phi) = f(x;\mu)\left[1+\phi\frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}\right]$$

- $var[Y] = v(\mu) + \frac{(n-1)}{n}\phi$
- μ and ϕ are orthogonal
- Interpretation
    - $\phi > 0$ overdispersion

Binomial Local mixture: Anaya-Izquierdo & Marriott(2007)

- Modified Laplace expansion

$$\int f(y;m)\,dH(m) \sim g(y;\mu,\phi) = f(x;\mu)\left[1+\phi\frac{d^2 f(x;\theta(\mu))/d\mu^2}{f(x;\theta(\mu))}\right]$$

- $var[Y] = v(\mu) + \frac{(n-1)}{n}\phi$
- μ and ϕ are orthogonal
- Interpretation
  - $\phi > 0$ overdispersion
  - $\phi < 0$ underdispersion

Correlated Binomial (Kupper & Haseman, 1978)

- Distribution of the sum of n dependent binary variables where their joint distribution is symmetric with no second or higher order additive interactions

Correlated Binomial (Kupper & Haseman, 1978)

- Distribution of the sum of n dependent binary variables where their joint distribution is symmetric with no second or higher order additive interactions
- Take $\rho = \phi/(\mu(n-\mu))$ in Local mixture so ρ has the interpretation of being a pairwise correlation between the binary variables and

$$Var[Y] = v(\mu)[1 + (n-1)\rho]$$

Correlated Binomial (Kupper & Haseman, 1978)

- Distribution of the sum of n dependent binary variables where their joint distribution is symmetric with no second or higher order additive interactions
- Take $\rho = \phi/(\mu(n-\mu))$ in Local mixture so ρ has the interpretation of being a pairwise correlation between the binary variables and

$$Var[Y] = v(\mu)[1 + (n-1)\rho]$$

- In general, let $Z_1, \ldots, Z_n$ be dependent Bernoulli(p) random variables such that $Cov[Z_i, Z_j] = \phi$ for $i \neq j$. If $Y = \sum_{i=1}^{n} Z_i$

$$\begin{aligned} Var[Y] &= \sum_{i=1}^{n} Var[Z_i] + \sum_{i \neq j} Cov[Z_i, Z_j] \\ &= v(\mu)[1 + (n-1)\rho] \end{aligned}$$

Other

- Beta-Binomial, Probit/Logistic-Binomial, ...

Other

- Beta-Binomial, Probit/Logistic-Binomial, ...
- Markov-chain Binomial (Rudolfer 1990):

Other

- Beta-Binomial, Probit/Logistic-Binomial, ...
- Markov-chain Binomial (Rudolfer 1990): $Z_1, \ldots, Z_n$ are binary rv with a general transition probability

$$\begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

Other

- Beta-Binomial, Probit/Logistic-Binomial, ...
- Markov-chain Binomial (Rudolfer 1990): $Z_1, \ldots, Z_n$ are binary rv with a general transition probability

$$\begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

The sum $Y = \sum_{i=1}^{n} Z_i$ has density

$$g(y; \mu, \phi) = \frac{\alpha^y (1-\alpha)^{m-1-2y} \beta^{y+1}}{\alpha + \beta} S(y; \alpha, \beta, n)$$

$$\begin{aligned} E[Y] &= \mu \\ Var[Y] &= v(\mu)[1 + (n-1) r(\alpha, \beta)] \end{aligned}$$

Overdispersed Binomial: Mixed parametrisation

Embed the extended binomial $g(y; \mu, \phi)$ into the multinomial via

$$g(y; \mu, \phi) \mapsto (f(0; \mu, \phi), f(1; \mu, \phi), \ldots, f(n; \mu, \phi))^T$$

Overdispersed Binomial: Mixed parametrisation

Embed the extended binomial $g(y;\mu,\phi)$ into the multinomial via

$$g(y;\mu,\phi) \mapsto (f(0;\mu,\phi), f(1;\mu,\phi), \ldots, f(n;\mu,\phi))^T$$

so density of $\mathbf{U} = (U_1, \ldots, U_N)$ where $U_i = \mathbb{I}(Y = i)$ is

$$\exp(\mathbf{U}^\top \chi(\mu,\phi) - \kappa(\chi))$$

Overdispersed Binomial: Mixed parametrisation

Embed the extended binomial $g(y;\mu,\phi)$ into the multinomial via

$$g(y;\mu,\phi) \mapsto (f(0;\mu,\phi), f(1;\mu,\phi), \ldots, f(n;\mu,\phi))^T$$

so density of $\mathbf{U} = (U_1, \ldots, U_N)$ where $U_i = \mathbb{I}(Y = i)$ is

$$\exp(\mathbf{U}^\top \chi(\mu,\phi) - \kappa(\chi))$$

where $\chi(\mu,\phi) =$
$(\log(f(1,\mu,\phi)/f(0,\mu,\phi)), \ldots, \log(f(n,\mu,\phi)/f(0,\mu,\phi)))$. Define

$$\mathbf{d} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ N \end{pmatrix} \qquad \mathbf{d}_T = \begin{pmatrix} T(1) \\ T(2) \\ \vdots \\ T(N) \end{pmatrix} \qquad \mathbf{c} = \begin{bmatrix} \log \binom{N}{1} \\ \log \binom{N}{2} \\ \vdots \\ \log \binom{N}{N} \end{bmatrix}$$

Overdispersed Binomial: Mixed parametrisation

Apply linear transformation $\mathbf{Y} = A^T\mathbf{U}$ where $A = [\mathbf{d}\ \mathbf{d}_{Y^2} \parallel B]$

Overdispersed Binomial: Mixed parametrisation

Apply linear transformation $\mathbf{Y} = A^T\mathbf{U}$ where $A = [\mathbf{d}\ \mathbf{d}_{Y^2} \parallel B]$

$$\begin{aligned} E[Y_1] &= E[Y] = \mu \\ E[Y_2] &= E[Y^2] = \mu^{(2)} = Var(Y) + \mu^2 \end{aligned}$$

Overdispersed Binomial: Mixed parametrisation

Apply linear transformation $\mathbf{Y} = A^T \mathbf{U}$ where $A = [\mathbf{d}\ \mathbf{d}_{Y^2} \parallel B]$

$$\begin{aligned} E[Y_1] &= E[Y] = \mu \\ E[Y_2] &= E[Y^2] = \mu^{(2)} = Var(Y) + \mu^2 \end{aligned}$$

The density of $\mathbf{Y}$ is EF with natural parameter

$$\omega = A^{-1}\chi$$

Overdispersed Binomial: Mixed parametrisation

Apply linear transformation $\mathbf{Y} = A^T \mathbf{U}$ where $A = [\mathbf{d}\ \mathbf{d}_{Y^2} \parallel B]$

$$\begin{aligned} E[Y_1] &= E[Y] = \mu \\ E[Y_2] &= E[Y^2] = \mu^{(2)} = Var(Y) + \mu^2 \end{aligned}$$

The density of $\mathbf{Y}$ is EF with natural parameter

$$\boldsymbol{\omega} = A^{-1}\boldsymbol{\chi}$$

Now use mixed parametrisation

$$(\mu, \mu^{(2)}, \boldsymbol{\omega}^*(\mu, \phi)) \quad \text{or} \quad (\mu, Var(Y), \boldsymbol{\omega}^*) \quad \boldsymbol{\omega}^* \in \mathbb{R}^{n-2}$$

to characterise the extended Binomial $g(y; \mu, \phi)$

Overdispersed Binomial: Mixed parametrisation

Family	$Var(Y)$	ω^*
Binomial	$v(\mu)$	ω_0^*
Beta-Binomial	$\mathrm{v}(\mu)\phi$	$\omega_0^* + \mathbf{u}(\mu, \phi)$
Multiplicative	$v(\mu)\phi$	ω_0^*
Multiplicative+	$Var(Y)$	$\omega_0^* + \delta\mathbf{u}$
Markov	$v(\mu)\phi$	$\omega_0^* + \mathbf{v}(\mu, \phi)$

Summary

- Models for overdispersed data can be expressed in terms of simple geometrical operations

Summary

- Models for overdispersed data can be expressed in terms of simple geometrical operations
- Goes beyond exponential and mixture family structures

Summary

- Models for overdispersed data can be expressed in terms of simple geometrical operations
- Goes beyond exponential and mixture family structures

Work supported by EPSRC grant EP/E017878/1

References

- Jorgensen, B. (1997) *The Theory of Dispersion Models*. Chapman and Hall.
- Efron, B. (1986). Double exponential families and their use in generalized linear regression. *JASA* (81) 709–721
- Shmueli, G. et al. (2005) A useful distribution for fitting discrete data: revival of the ConwayMaxwellPoisson distribution. *JRSSC* (54) 127–142
- Altham, P.M.E (1978) Two Generalizations of the Binomial distribution. Applied Statist. (27) 162–167
- Cox, D.R. (1983) Some remarks on overdispersion. *Biometrika* (70) 269–274
- Rayner, J.C.W & Thas, O. & Best, D.J. (2009) Smooth goodness of fit tests: using R. *John Wiley and Sons*.
- Anaya-Izquierdo, K. & Marriott, P. (2007) Local mixtures of exponential families. *Bernoulli* (13) 623–640

References

- Kupper, L.L. & Haseman, J.K. (1978) The use of a correlated Binomial model for the analysis of certain toxicological experiments. *Biometrics* (34) 69–76
- Rudolfer, S.M. (1990). A markov-chain model for extra Binomial variation. *Biometrika* (77) 255–264
- Gelfand, A.E. & Dalal, S.R. (1990) A note on overdispersed exponential families. *Biometrika* (77) 55–64
- Lindsay, B. (1986) Exponential family mixture models (with Least squares estimators). *Ann. Stats.* (14) 124–137
- Copas, J.C.B & Eguchi, S. (2005) Local model uncertainty and incomplete-data bias. *JRSSB* (67) 459–513