Population models in design and analysis of dose finding experiments for binary responses

V. Fedorov, GSK

June, 2009

Outline

- Two types of mixing models
- Moment based approach
- Iterated estimators
- Design

Two mixing models

Y is the binary response and

$$\pi_{ij} = P(Y = 1 \mid i, j)$$

j stands treatment (dose), i – experimental unit (patient, center).

Two types of "mixing"

Type A $$\pi_{ij} = \eta(x_j, \gamma_i), \quad \gamma_i \sim \phi(\gamma | \theta)$$

Type B $$\pi_{ij} \sim \psi(\pi, \gamma(x_j, \theta))$$

Type A: G-probit model

$$\pi_{ij} = \pi(x_j, \gamma_i) = \int_{\infty}^{\eta(x_j, \gamma_i)} \psi(u|\alpha)du, \quad \gamma_i \sim \varphi(\gamma|\theta)$$

⊕ Probabilities π_{ij} are dependent for different j-s

Likelihood function:

$$L(\theta|\mathbf{Y}) = \prod_{i=1}^{N} \mathrm{E}[\prod_{j=1}^{K} \pi_{ij}^{Y_{ij\cdot}} (1 - \pi_{ij})^{n_{ij} - Y_{ij\cdot}} |\theta]_{\pi}$$

⊖ Easy to write difficult to compute!

⊖ Problem with building information matrix.

Type B. Beta-binomial model

$$\pi_{ij} \sim Beta(a,b)$$

$$a_{ij} = \exp\left[\theta_a^T f_a(x_{ij})\right], \qquad b_{ij} = \exp\left[\theta_b^T f_b(x_{ij})\right]$$

Marginal distribution of $Y_{ij.}$

$$\psi(y; n_{ij}, a, b) = \binom{n_{ij}}{y} \frac{B(y + a_{ij}, n_{ij} - y + b_{ij})}{B(a,b)}$$

$$L(\theta|\mathbf{Y}) = \prod_{i=1}^{N}\prod_{j=1}^{K} \binom{n_{ij}}{Y_{ij.}} \frac{B(Y_{ij.} + a_{ij}, n_{ij} - Y_{ij.} + b_{ij})}{B(a_{ij}, b_{ij})}$$

$$E[Y_{ij.}] = n_{ij}\frac{a}{a+b} = n_{ij}p$$

$$\begin{aligned}\mathrm{Var}[Y_{ij.}] &= n_{ij}\frac{ab(a+b+n_{ij})}{(a+b)^2(a+b+1)} \\ &= n_{ij}p(1-p)\{1 + \frac{\tau}{1+\tau}(n_{ij}-1)\}\end{aligned}$$

⊕ Easy to compute MLE and information* matrices

⊖ No dependence between results on different doses

*Sudhir R.,etc. Biometrical Journal 47 (2005) 2, 230–236

Quasi-linear or moment based methods I

Very modest assumptions:

$$\mathrm{E}[\pi_i] \equiv \mathbf{p} \quad \text{and} \quad \mathrm{Var}[\pi_i] = \mathbf{V}$$

Estimators for elemental parameters: $\hat{\pi}_{ij} = \overline{Y}_{ij} = \dfrac{Y_{ij\cdot}}{n_{ij}}$

Their variance*: $\mathrm{Var}[\hat{\pi}_i] = \mathbf{V}_i = \mathbf{M}_i^{-1} + \mathbf{V}$

$$\hat{\pi}_i = \{\hat{\pi}_{ij}\}_1^K \quad \text{and} \quad M_{i,jj'} = \delta_{jj'} \frac{n_{ij}}{p_j(1-p_j) - V_{jj}}$$

We are interested in estimation of **p** and **V**.

⊕ Simple computing, easy to introduce correlation between results at different doses

⊖ Potential loss of information

*Dragalin V., Fedorov V., Statist. Med. 2006; 25, 2701-2719

Quasi-linear or moment based methods II

Previous formulae are based on the formula for marginal variance:

$$\mathrm{Var}[\hat{\pi}_i] = E\left[\mathrm{Var}\left[\hat{\pi}_i \mid \pi_i\right]\right] + \mathrm{Var}\left[E\left[\hat{\pi}_i \mid \pi_i\right]\right]$$

Note.
For any continuous p.d.f. $f(p)$ with the support set $[0,1]$, mean μ_p and variance σ_p^2, we have $\sigma_p^2 < \mu_p(1-\mu_p)$, with equality only for the atomized distribution at points 0 and 1 with weights $1-\mu_p$ and μ_p, respectively. Hence, $p_j(1-p_j) - V_{jj} > 0$.

Iterated quasi-linear estimator I

Mimicking the best linear estimator leads to:

$$\hat{\mathbf{p}} = \left(\sum_{i=1}^{N} \mathbf{W}_i\right)^{-1} \sum_{i=1}^{N} \mathbf{W}_i \hat{\pi}_i$$

where $\mathbf{W}_i = \mathbf{V}_i^{-1}$, when $\mathbf{M_i}$ is regular and $\mathbf{W}_i = \mathbf{M}_i - \mathbf{M}_i(\mathbf{M}_i + \mathbf{V}^{-1})^{-1}\mathbf{M}_i$ otherwise.

.

$$\mathrm{Var}[\hat{\mathbf{p}}] = \left(\sum_{i=1}^{N} \mathbf{W}_i\right)^{-1} = \left(\sum_{i=1}^{N} \mathbf{V}_i^{-1}\right)^{-1}$$

- Problem: **p** and **V** are unknown ☹

Iterated quasi-linear estimator II

$$\hat{\mathbf{p}}^{(t+1)} = \left(\sum_{i=1}^{N} \mathbf{W}_i^{(t)} \right)^{-1} \sum_{i=1}^{N} \mathbf{W}_i^{(t)} \hat{\pi}_i$$

Where:

$$\mathbf{W}_i^{(t)} = \left[\left(\mathbf{M}_i^{(t)} \right)^{-1} + \mathbf{V}^{(t)} \right]^{-1}$$

$$\left(\mathbf{M}_i^{(t)} \right)^{-1} = \begin{pmatrix} \frac{p_1^{(t)}(1-p_1^{(t)})-V_{11}^{(t)}}{n_{i1}} & 0 & \cdots \\ 0 & \frac{p_2^{(t)}(1-p_2^{(t)})-V_{22}^{(t)}}{n_{i2}} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}$$

$$\mathbf{V}^{(t)} = \frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} \frac{Y_{i1.}}{n_{i1}} \frac{Y_{i1.}-1}{n_{i1}-1} & \frac{Y_{i1.}}{n_{i1}} \frac{Y_{i2.}}{n_{i2}} & \cdots \\ \frac{Y_{i1.}}{n_{i1}} \frac{Y_{i2.}}{n_{i2}} & \frac{Y_{i2.}}{n_{i2}} \frac{Y_{i2.}-1}{n_{i2}-1} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix} - \hat{\mathbf{p}}^{(t)} \hat{\mathbf{p}}^{(t)T}$$

Motivation for the estimator selection

The iterative estimator for $\mathbf{V}$ is motivated by the fact that

$$E\left(\frac{Y_{ij\cdot}}{n_{ij}}\frac{Y_{ij\cdot}-1}{n_{ij}-1}\mid\pi_{ij}\right)=\pi_{ij}^2,$$
$$E\left(\frac{Y_{ij\cdot}}{n_{i1}}\frac{Y_{ij'\cdot}}{n_{i2}}\mid\pi_{ij},\pi_{ij'}\right)=\pi_{ij}\pi_{ij'},$$

while

$$E(\pi_{ij}\pi_{ij'})=V_{jj'}+p_jp_{j'},$$
$$E(\pi_{ij}^2)=V_{jj}-p_j^2.$$

Locally optimal designs

Optimality criterion:

$$\mathrm{Var}[\ell^T \hat{\mathbf{p}}] = \ell^T \left(\sum_{i=1}^{N} \mathbf{V}_i^{-1} \right)^{-1} \ell$$

Two controls:

1. number of patients in each center
2. distribution of patients between treatments

$$\mathrm{Var}[\ell^\mathsf{T} \widehat{p}] = \ell^\mathsf{T} \left(\sum_{i=1}^{N} V_i^{-1} \right)^{-1} \ell \ \geq \ell^\mathsf{T} \left(\sum_{i=1}^{N} (M^{-1} + V)^{-1} \right)^{-1} \ell \ = \frac{1}{N} \ell^\mathsf{T} (M^{-1} + V) \ell$$

Thus optimal design should be balanced across all centers, i.e. $n_{ij} = n_j$ and straightforward optimization leads to

$$n_{ij} \equiv n_j \sim \sqrt{\ell_j^2 [p_j(1 - p_j) - V_{jj}]}$$

How to select population means and variances I

Response models with random parameters

Link function → $\lambda(\pi_{ij}) = \gamma_i^T \mathbf{f}(x_j) + \varepsilon_{ij}$

$\gamma_i \sim \mathcal{N}(\gamma, \Sigma), \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

Population means: $p_j \simeq \lambda^{\ominus}\left[\gamma^T \mathbf{f}(x_j) - \frac{1}{2}\ddot{\lambda}(\gamma^T \mathbf{f}(x_j))V_{jj}\right]$

Var-Cov matrix of **p** : $\mathbf{V} = \mathrm{Var}(\pi) \simeq \dot{\Lambda}^{-1}\left(\mathbf{F}^T \Sigma\, \mathbf{F} + \sigma^2 \mathbf{I}\right)\dot{\Lambda}^{-1}$

Where: $\mathbf{F} = \{\mathbf{f}(x_j)\}_1^K, \quad \dot{\Lambda} = \{diag\, \dot{\lambda}(\gamma_i^T \mathbf{f}(x_j))\}$

How to select population means and variances II

Derivations are based on Taylor's expansion:

$$\lambda(\pi_{ij}) = \gamma_i^T \mathbf{f}(x_j) + \varepsilon_{ij} \simeq \lambda(p_j) + \dot{\lambda}(p_j)(\pi_{ij} - p_j) + \frac{1}{2}\ddot{\lambda}(p_j)(\pi_{ij} - p_j)^2$$

For logit link: $\lambda(p) = \ln \frac{p}{1-p} = u, \qquad p(u) = \lambda^{\ominus}(u) = \frac{e^u}{1+e^u}$

Response described by autoregression model:

$$\lambda(\pi_{ij}) = [\gamma_i(x_j - x_{j-1})]\lambda(\pi_{i,j-1}) + \varepsilon_{ij}$$

References

- Agresti, A. and Hartzel, J. (2000). Strategies for comparing treatments on a binary response with multi-centre data. *Statistics in Medicine,* **19**: 1115–1139.
- Dragalin V., Fedorov V. (2006) Design of multi-centre trials with binary response. *Statistics in Medicine*, **25**:2701–2719.
- Sudhir P., Uditha B., and Tathagata B. (2005). Fisher Information Matrix of the Dirichlet-multinomial Distribution *Biometrical Journal,* **47**, 230–236.