Optimal inference via confidence distributions for two-by-two tables modelled as Poisson pairs: fixed and random effects

FOCUS DRIVEN STATISTICAL
INFERENCE WITH COMPLEX DATA

Céline Cunen, Nils Lid Hjort

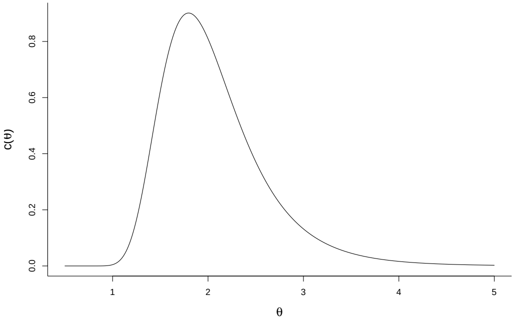
University of Oslo

30/07/2015

Plan

- Confidence distributions: a quick introduction
- Meta-analysis of 2x2 tables: the problem, different models
- Schweder and Hjort Optimality Theorem
- The optimal method for the fixed effect model
- Analysis on a real dataset
- Random effect models
- CD for the ratio between two treatment effects

Confidence distributions (CD)

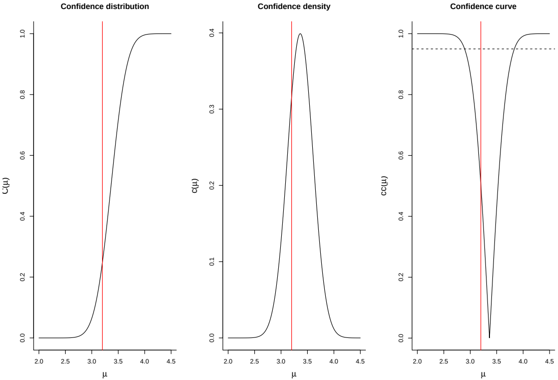


- $\approx$ a posterior without having to specify a prior
  - a sample-dependent distribution function on the parameter space
  - can be used for inference (for example for constructing CIs of all levels)

Confidence distributions (CD)

- Distribution estimator (we have point estimators, CIs and distribution estimators)

- Purely frequentist:
  - The parameter is a fixed, non-random quantity
  - 95% CI: the true parameter value will be covered by the CIs 95% of the time



Requirements for CDs

Definition

A function $C(\theta, X)$ is called a confidence distribution for a parameter θ if:

- $C(\theta, X)$ is a cumulative distribution function on the parameter space
- at the true parameter value $\theta = \theta_0$, $C(\theta_0, X)$ as a function of the random sample X follows the uniform distribution U[0,1]

- The second requirement ensures that all confidence intervals have the correct coverage.

The Problem - Meta-analysis of 2x2 tables

Study	Sample size control = $m_{0,i}$	Events control = $y_{0,i}$	Sample size treatment = $m_{1,i}$	Events treatment = $y_{1,i}$
1	39	1	43	2
2	44	4	44	4
3	107	4	110	6
4	103	5	100	7
5	110	3	106	7
6	154	4	146	11

- We are interested in finding out whether the control and treatment groups are different
- Rare events!
- Existing methods (Mantel-Haenszel, Peto) = bad for rare events
- (LLX method)

Binomial Model

$$Y_{0,i} \sim \mathrm{Bin}(m_{0,i}, p_{0,i}) \quad \text{and} \quad Y_{1,i} \sim \mathrm{Bin}(m_{1,i}, p_{1,i})$$

- with $p_{0,i} = \frac{e^{\theta_i}}{1+e^{\theta_i}}$ and $p_{1,i} = \frac{e^{\theta_i+\psi_i}}{1+e^{\theta_i+\psi_i}}$
- Odds ratio: $\mathrm{OR}_i = \frac{p_{1,i}/(1-p_{1,i})}{p_{0,i}/(1-p_{0,i})} = \frac{e^{\theta_i+\psi_i}}{e^{\theta_i}} = e^{\psi_i}$

Binomial Model

$$Y_{0,i} \sim \mathrm{Bin}(m_{0,i}, p_{0,i}) \quad \text{and} \quad Y_{1,i} \sim \mathrm{Bin}(m_{1,i}, p_{1,i})$$

- with $p_{0,i} = \frac{e^{\theta_i}}{1+e^{\theta_i}}$ and $p_{1,i} = \frac{e^{\theta_i+\psi}}{1+e^{\theta_i+\psi}}$
- Odds ratio: $\mathrm{OR} = \frac{p_{1,i}/(1-p_{1,i})}{p_{0,i}/(1-p_{0,i})} = \frac{e^{\theta_i+\psi}}{e^{\theta_i}} = e^{\psi}$
- Usual assumption: a common odds ratio (and ψ) across tables.

Binomial Model

$$Y_{0,i} \sim \mathrm{Bin}(m_{0,i}, p_{0,i}) \quad \text{and} \quad Y_{1,i} \sim \mathrm{Bin}(m_{1,i}, p_{1,i})$$

- with $p_{0,i} = \frac{e^{\theta_i}}{1+e^{\theta_i}}$ and $p_{1,i} = \frac{e^{\theta_i+\psi_i}}{1+e^{\theta_i+\psi_i}}$
- Odds ratio: $\mathrm{OR} = \frac{p_{1,i}/(1-p_{1,i})}{p_{0,i}/(1-p_{0,i})} = \frac{e^{\theta_i+\psi_i}}{e^{\theta_i}} = e^{\psi_i}$
- another possibility: a random effects model

$$\psi_i \sim \mathrm{N}(\psi_0, \tau^2).$$

Poisson Model

$$Y_{0,i} \sim \mathrm{Pois}(e_{0,i}\lambda_i) \quad \text{and} \quad Y_{1,i} \sim \mathrm{Pois}(e_{1,i}\lambda_i\gamma_i)$$

- the two models are related: $\mathrm{OR}_i = e^{\psi_i} \approx \gamma_i$ (the treatment effect)
- $\lambda_i \approx$ the risk in the control group
- $e_{0,i}$ and $e_{1,i}$ are exposure weights, chosen as $e_{0,i} = m_{0,i}/100 \quad \text{and} \quad e_{1,i} = m_{1,i}/100$
- fixed effect model: $\gamma_i = \gamma$
- or random effects model: later

Schweder and Hjort Optimality Theorem (SJS 2002)

Optimal confidence for exponential families

Let γ be the parameter of interest and λ the nuisance parameter vector in a continuous exponential model, with density

$$f(y, \gamma, \lambda) = e^{\gamma S(y) + \lambda_1 A_1(y) + \cdots + \lambda_k A_k(y) - k(\gamma, \lambda_1, \ldots, \lambda_k)}.$$

(...) Then

$$C_{S \mid A}(\gamma) = P_\gamma(S > s_{\mathrm{obs}} \mid A = a_{\mathrm{obs}})$$

is the uniformly most powerful confidence distribution.

- Uniformly most powerful confidence distribution: has lower expected confidence loss than all other CDs for a large class of loss functions and all values of (γ, λ).
- If discrete distribution use half-correction.

Optimal method for the fixed effect model

With the fixed effect model:

$$Y_{0,i} \sim \text{Pois}(e_{0,i}\lambda_i) \quad \text{and} \quad Y_{1,i} \sim \text{Pois}(e_{1,i}\lambda_i\gamma)$$

and with $z_i = y_{0,i} + y_{1,i}$ and k 2×2 tables, the log-likelihood is

$$\ell(\gamma, \lambda_1, \ldots, \lambda_k) = \sum_{i=1}^{k} \{y_{1,i} \log \gamma + z_i \log \lambda_i - (e_{0,i} + e_{1,i}\gamma)\lambda_i\}$$

and we can use the Optimality theorem.

$C_i(\gamma) = P_\gamma(Y_{1,i} > Y_{1,i,\text{obs}} \,|\, z_{i,\text{obs}}) + \frac{1}{2} P_\gamma(Y_{1,i} = Y_{1,i,\text{obs}} \,|\, z_{i,\text{obs}})$ will be the optimal CD for one table.

Optimal method for the fixed effect model

Combining information from all studies, the optimal CD for γ is

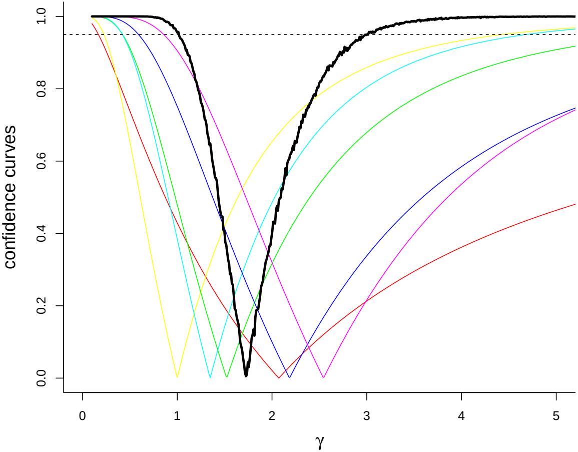
$$C^*(\gamma) = P_\gamma(B > B_{\text{obs}} \mid z_{1,\text{obs}}, \ldots, z_{k,\text{obs}}) + \frac{1}{2} P_\gamma(B = B_{\text{obs}} \mid z_{1,\text{obs}}, \ldots, z_{k,\text{obs}})$$

where $B = \sum_{i=1}^{k} Y_{1,i}$.
The conditional distribution (a binomial!)

$$Y_{1,i} \mid (Z_i = z_i) \sim \text{Bin}\left(z_i, \frac{e_{1,i}\gamma}{e_{0,i} + e_{1,i}\gamma}\right)$$

The distribution of B can be simulated.

Real dataset - Lidocaine data



- 95% confidence interval for γ: [1.01,3.01]

Random effect models

$$Y_{0,i} \sim \mathrm{Pois}(e_{0,i}\lambda_i) \quad \text{and} \quad Y_{1,i} \sim \mathrm{Pois}(e_{1,i}\lambda_i\gamma_i)$$

- Several possibilities
- For example:

$$Y_{1,i} \,|\, (Z_i = z_i, \gamma_i) \sim \mathrm{Bin}(z_i, \pi_i) \quad \text{with} \quad \pi_i = \frac{e_{1,i}\gamma_i}{e_{0,i} + e_{1,i}\gamma_i},$$

where $\pi_i \sim \mathrm{Beta}\{\tau\pi_{0,i}, \tau(1 - \pi_{0,i})\}$

and $\pi_{0,i} = e_{1,i}\gamma_0/(e_{0,i} + e_{1,i}\gamma_0)$, $\kappa = (\tau + 1)^{-1}$

- Can make CDs for γ_0 and κ, but we cannot use the Optimality theorem here

CD for the ratio between two treatment effects

$$Y_{0,i} \sim \mathrm{Pois}(e_{0,i}\lambda_i) \quad \text{and} \quad Y_{1,i} \sim \mathrm{Pois}(e_{1,i}\lambda_i\gamma_i)$$

- Interested in comparing pairs of studies:

$$\delta = \frac{\gamma_2}{\gamma_1}$$

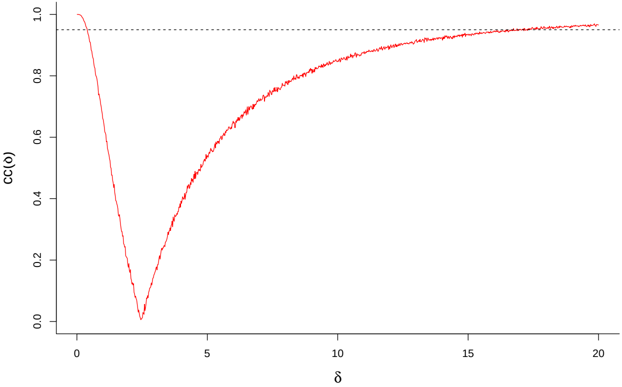
- We can use the Optimality theorem and find an optimal CD (ignoring half-correction)

$$C_{1,2}(\delta) = P_\delta(Y_{1,2} > Y_{1,2,\mathrm{obs}} \,|\, w_{\mathrm{obs}}, z_{1,\mathrm{obs}}, z_{2,\mathrm{obs}})$$

- where $W = Y_{1,1} + Y_{1,2}$, $Z_1 = Y_{0,1} + Y_{1,1}$ and $Z_2 = Y_{0,2} + Y_{1,2}$
- Based on the eccentric hypergeometric distribution

$$f(y_{1,2} \,|\, (w, z_1, z_2)) = \frac{\binom{z_1}{w-y_{1,2}}\binom{z_2}{y_{1,2}}\left(\frac{e_{0,1}e_{1,2}}{e_{1,1}e_{0,2}}\right)^{y_{1,2}}\delta^{y_{1,2}}}{\sum_{u=0}^{w}\binom{z_1}{w-u}\binom{z_2}{u}\left(\frac{e_{0,1}e_{1,2}}{e_{1,1}e_{0,2}}\right)^{u}\delta^{u}}$$

Other models - Ratio of Odds ratio



The two most different Lidocaine studies.
95% confidence interval for δ: [0.37, 17.81]

References

- Liu, D., Liu, R. Y., & Xie, M. (2014). Exact meta-analysis approach for discrete data and its application to 2x2 tables with rare events. Journal of the American Statistical Association, 109, 1450-1465.
- Schweder, T. & Hjort, N. L. (2002). Confidence and likelihood. Scandinavian Journal of Statistics, 29, 309-332.
- Schweder, T. & Hjort, N. L. (2015). Confidence, Likelihood, Probability: Inference With Confidence Distributions. Cambridge University Press.
- Xie, M. & Singh, K. (2013). Confidence Distribution, the Frequentist Distribution Estimator of a Parameter: A Review. International Statistical Review, 81, 3-39. (with discussion)