Challenges on Imprecise Inference

for the measure of association in 2x2 tables

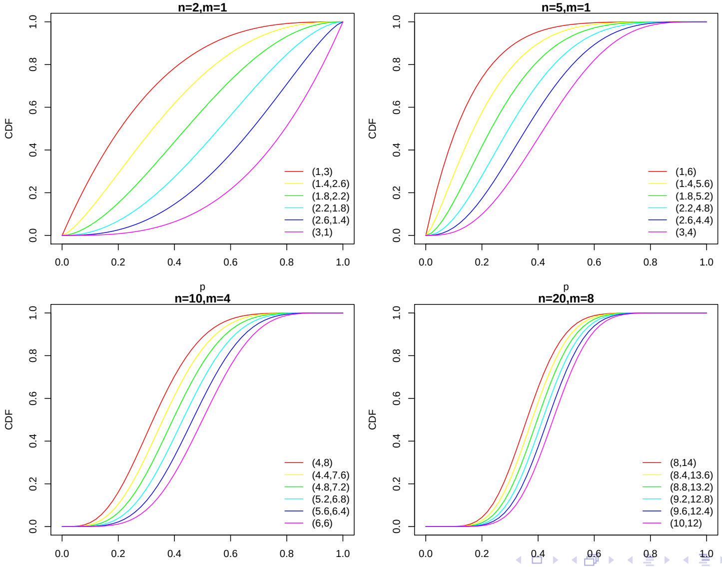
Chel Hee Lee[1] Miķelis Bickis[2]

[1]Clinical Research Support Unit
Community Health and Epidemiology
College of Medicine
University of Saskatchewan

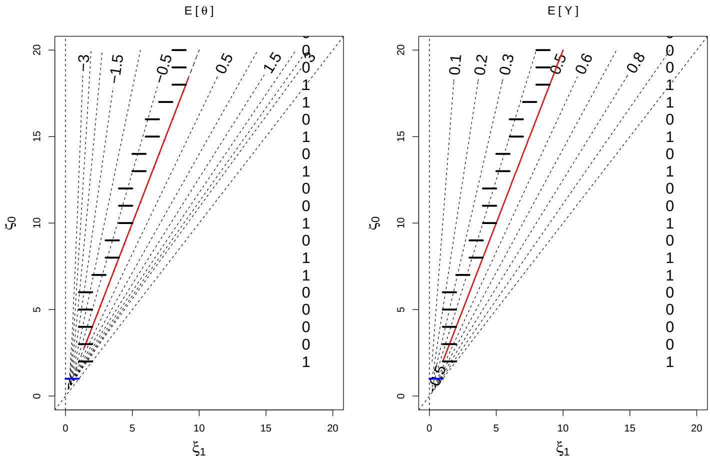
[2]Department of Mathematics and Statistics
University of Saskatchewan

9th WPMSIIP, Durham, England
6 September 2016

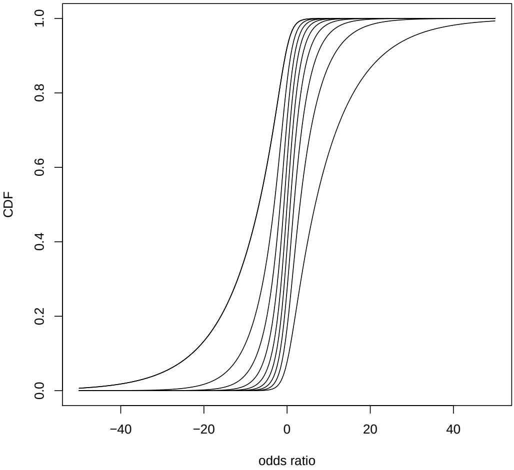
Beta-Binomial Model (Walley, 1991)



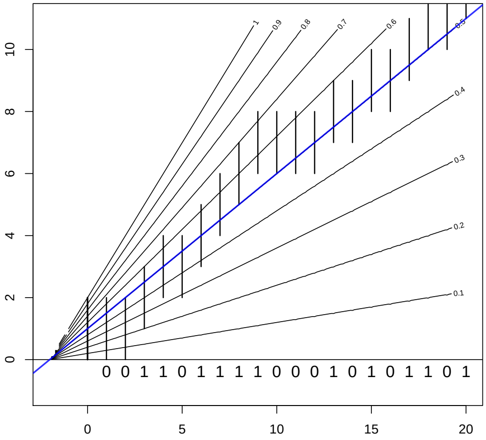
Beta-Binomial Model (Chel)



Beta-Binomial Model (Chel)



Uniform-Binomial Model



Persistent Pulmonary Hypertension (Walley, 1996, Sec. 5)

		Survived		
		Yes	No	Total
TRT	CT	6		10
	ECMO	9		9

CT means a conventional therapy and
ECMO means extracorporeal membrane oxygenation.

Assumptions

- A constant chance of survival under each treatment.
- Outcomes are independent for different babies.

Persistent Pulmonary Hypertension

Imprecise Beta Model

$$
\begin{aligned}
L(\theta_c, \theta_e | n) &\propto \theta_c^6 (1 - \theta_c)^4 \theta_e^9 \\
\pi(\theta_c, \theta_e) &\propto \theta_c^{st_c - 1} (1 - \theta_c)^{s(1-t_c)-1} \theta_e^{st_e - 1} (1 - \theta_e)^{s(1-t_e)-1} \\
\hline
\pi(\theta_c, \theta_e | n) &\propto \theta^{st_c + 5} (1 - \theta_c)^{s(1-t_c)+3} \theta_e^{st_e + 8} (1 - \theta_e)^{s(1-t_e)-1}
\end{aligned}
$$

Inferences about $\theta_e - \theta_c$

$$
\begin{aligned}
H_0 &: \theta_e \leq \theta_c \\
H_1 &: \theta_e > \theta_c
\end{aligned}
$$

This can be answered by calculating $\overline{P}(H_1 | n)$ and $\underline{P}(H_1 | n)$

Can we work with log-odds?

General form

		Y_2 (levels)		(fixed)
		Yes	No	Total
Y_1	Yes	n_{11}	n_{10}	n_1
(type)	No	n_{01}	n_{00}	n_0

$$
\begin{aligned}
& L(n_{11}, n_{01} | p_{11}, p_{01}) \\
\propto\ & p_{11}^{n_{11}}(1-p_{11})^{n_1 - n_{11}} p_{01}^{n_{01}}(1-p_{01})^{n_0 - n_{01}} \\
=\ & \exp\left\{ n_{11}\theta_1 + n_{01}\theta_2 - n_1 \log(1+e^{\theta_1}) - n_2 \log(1+e^{\theta_2}) \right\}
\end{aligned}
$$

where $\theta = \log(p_{11}/(1-p_{11}))$ and $\theta_2 = \log(p_{01}/(1-p_{01}))$. Subsuquently, log-odds ratio (LOR) is found by $\theta_1 - \theta_2$. $P(L > 0 | n_{11}, n_{01})$ can be also found numerically.

Consider All Margins Are Fixed

General form

		Y_2		(fixed)
		Yes	No	Total
Y_1	Yes	n_{11}	n_{10}	n_{1+}
	No	n_{01}	n_{00}	n_{0+}
(fixed)		n_{+1}	n_{+0}	n

$$
\begin{aligned}
& L(p_{11}, p_{1+}, p_{+1}) \\
= \; & \binom{n}{n_{ij}} p_{11}^{n_{11}} (p_{1+} - p_{11})^{n_{10}} (p_{+1} - p_{11})^{n_{01}} (1 - p_{1+} - p_{+1} + p_{11})^{n_{00}} \\
= \; & \exp\left\{ n_{11} \log(\theta) - \mathfrak{A}(\theta) - \log \binom{n}{n_{ij}} \right\}, \quad \text{where } \theta = \frac{p_{11} p_{00}}{p_{10} p_{01}}
\end{aligned}
$$

Can we work with percent agreement?

Three outcomes are considered

- $P(Y_1 = 1, Y_2 = 0) = p_1$
- $P(Y_1 = 0, Y_2 = 1) = p_2$
- $P[(Y_1 = 1, Y_2 = 1) or (Y_1 = 0, Y_2 = 0)] = p_a$

where $p_1 + p_2 + p_3 = 1$.

$$
\begin{aligned}
L(Y|p) &\propto p_1^{n_1} p_2^{n_2} p_3^{n_3} \quad \text{where } n_3 = n - n_1 - n_2 \\
&= \exp\{n_1 \left(\frac{e^{\theta_1}}{1 + e^{\theta_1} + e^{\theta_2}} \right) + n_2 \left(\frac{e^{\theta_2}}{1 + e^{\theta_1} + e^{\theta_2}} \right) \\
&\quad -n \log \left(\frac{1}{1 + e^{\theta_1} + e^{\theta_2}} \right) + \log \binom{n}{n_1, n_2, n_3} \}
\end{aligned}
$$

Subsequently, we are interested in the inference $P(p_1 - p_2 > 0) = P(\log(p_1/p_2)) = P(p_1 > p_2)$.

Cohen's Kappa

Definition

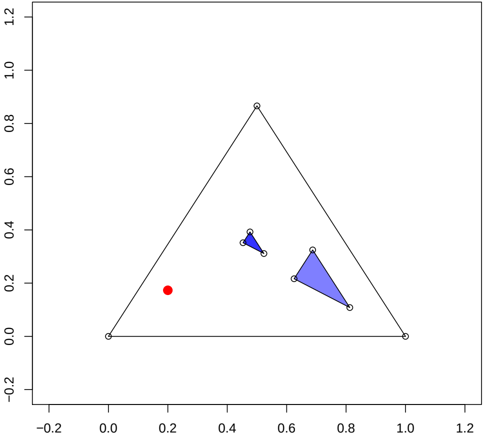
$$\begin{aligned}
\kappa(\mathbf{p}) &= \frac{(p_{11}+p_{00})-(p_{1+}p_{+1}+p_{0+}p_{+0})}{1-(p_{1+}p_{+1}+p_{0+}p_{+0})} \\
&= \frac{2(p_{11}-p_{1+}p_{+1})}{p_{1+}p_{+1}-2p_{1+}p_{+1}} \\
&= \frac{p_{00}p_{11}-p_{01}p_{10}}{p_{11}p_{00}-p_{01}p_{10}+(p_{01}+p_{10})/2}
\end{aligned}$$

Can We Work With Four Cells?

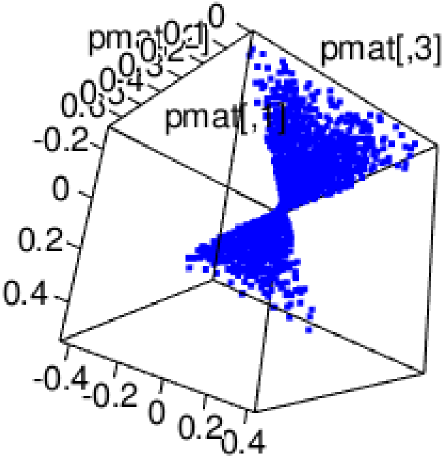
Mik's Talk

$$\log p_{ij} = \xi_1\theta_1 + \xi_2\theta_2 + \xi_3\theta_3 - \mathfrak{A}(\theta), \quad i,j = 1,2$$

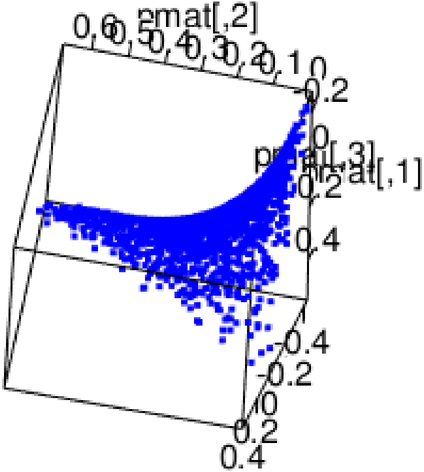
Matter of Visualization



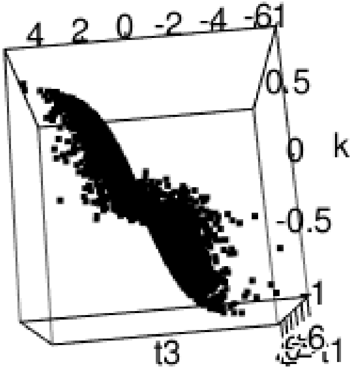
Matter of Visualization



Matter of Visualization



Matter of Visualization



References I

Diaconis, P. and Ylvisaker, D. (1979). Conjugate Priors for Exponential Families. Ann. Statist., 7(2):269–281.

Walley, P. (1991). Statistical reasoning with imprecise probabilities. Chapman and Hall, London;.

Walley, P. (1996). Inferences from Multinomial Data: Learning about a Bag of Marbles. Journal of the Royal Statistical Society. Series B (Methodological), 58(1):pp. 3–57.