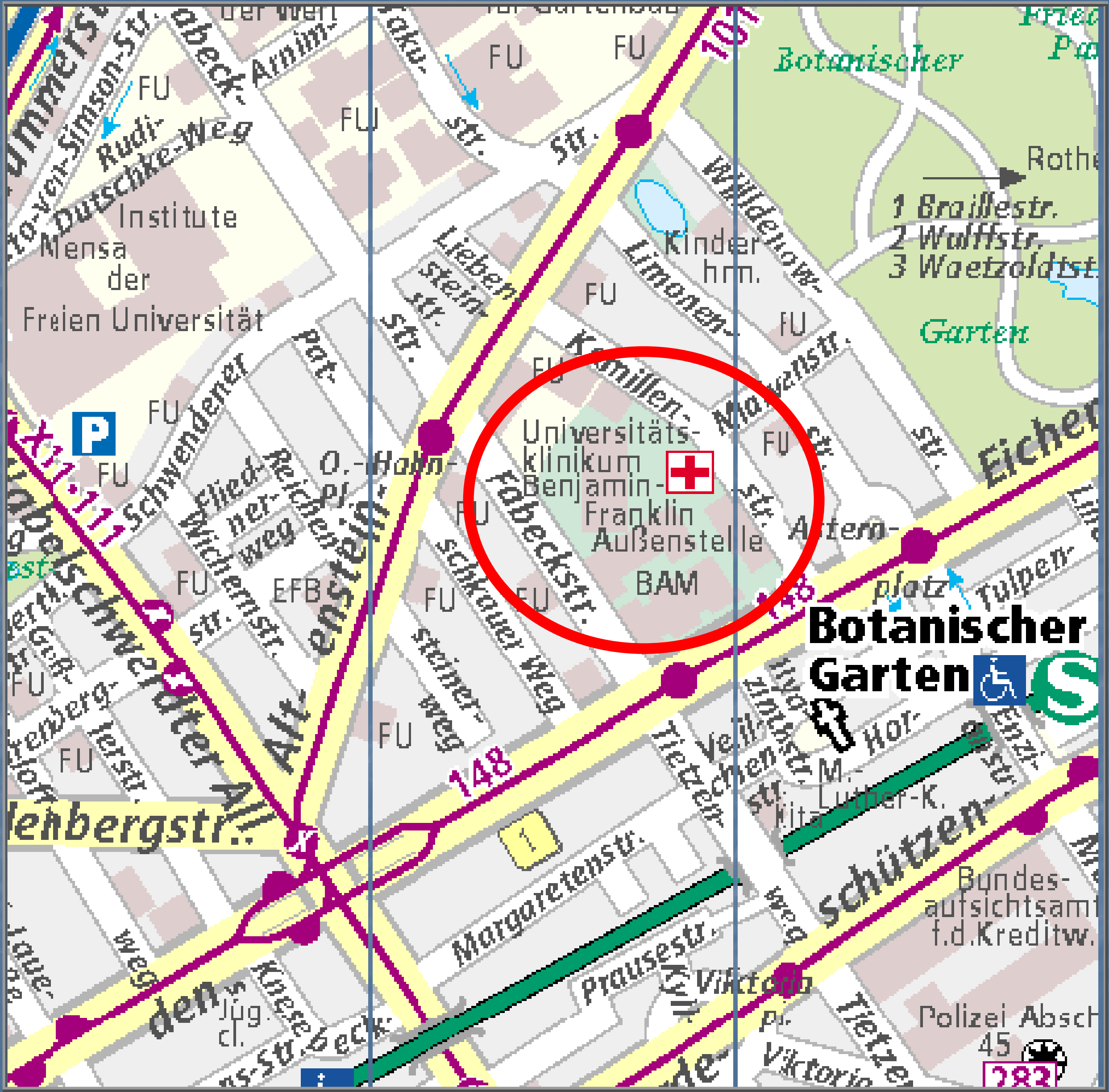
Personal Background and Areas of Interest

Prof. Dr. Dankmar Böhning

Division of International Health
Institute for Social Medicine, Epidemiology, und
Health Economics
Charité Medical School Berlin


Botanischer
Garten
Rothe
1 Braillestr.
2 Wulffstr.
3 Waetzoldtst.
Institute
Mensa
der
Freien Universität
FU
Rudi-Dutschke-Weg
Otto-von-Simson-Str.
Arnim-
Willdenow-
Limonen-
Kamillen-
Mayenstr.
Kinder
hm.
Universitäts-
klinikum
Benjamin-
Franklin
Außenstelle
BAM
Fabeckstr.
Schwendener
Pat-
O.-Hahn-
pl.
Altenstein-
Alt-
Schickauer Weg
Steiner-
weg
Lieben-
stein-
str.
Astern-
platz
Tulpen-
Botanischer
Garten
Eiche
148
101
X11, 111
Habelschwerdter Allee
EFB
Margaretenstr.
Prausestr.
Veilchen-
str.
Hyazinthstr.
M.-
Luther-K.
Kita
Schützen-
Bundes-
aufsichtsamt
f.d.Kreditw.
Polizei Absch
45
283
Viktoria
Tietzen-
weg
Jug.
cl.
Kiebitz-
Knese-
beck-
Hor-
Enzi-
an-
str.
P

Division of International Health: Staff (currently)

- Prof. Dr. Dankmar Böhning
- Dr. Ekkehart Dietz
- Ronny Kuhnert (DFG)
- Ms. Sasivimol Rattanasiri (BMZ)
- Ms. Beatrice Chew (Sekretary)
- Ms. Ina Schöttle (Research Assistant)

Overview

- History
- General Topics
- Current Areas of Interest
- Research Areas in Preperation

Overview

- History
- General Topics
- Current Areas of Interest
- Research Areas in Preperation

Personal Background

- Studies
  - Mathematics (main) and social sciences (Bielefeld and Berlin)
- Degrees
  - M.Sc. (optimal design) Dr. (algorithms)
  - Habil. (medicine: epidemiology/biometry)
- Cooperation
  - Numerous Institutions in Europe, USA, Australia, Thailand, and Philippines
- Visiting
  - 85-86 Statistics, PennState
  - 96 Psychology, Vienna
  - 98-99 Statistics, Munich
  - 04 International EpiLab, Copenhagen
  - Several Visits to Philippines and Thailand

History

- 1982: after completion of my PhD take up junior position at the Institute of Social Medicine
- 1992: v. l. in Medical Statistics and Epidemiolgy
- 2000: Award of the Title of *Professor*
- Several co-workers 1990-2004: Dietz, Kuhnert, Malzahn, Schlattmann, Stallmann, Schleinitz, ...

History

- 1981: Professor Schelp became Head of Institute for Social Medicine with emphasis on Epidemiology
- Professor Schelp active (at that time) with several intervention projects in NE Thailand

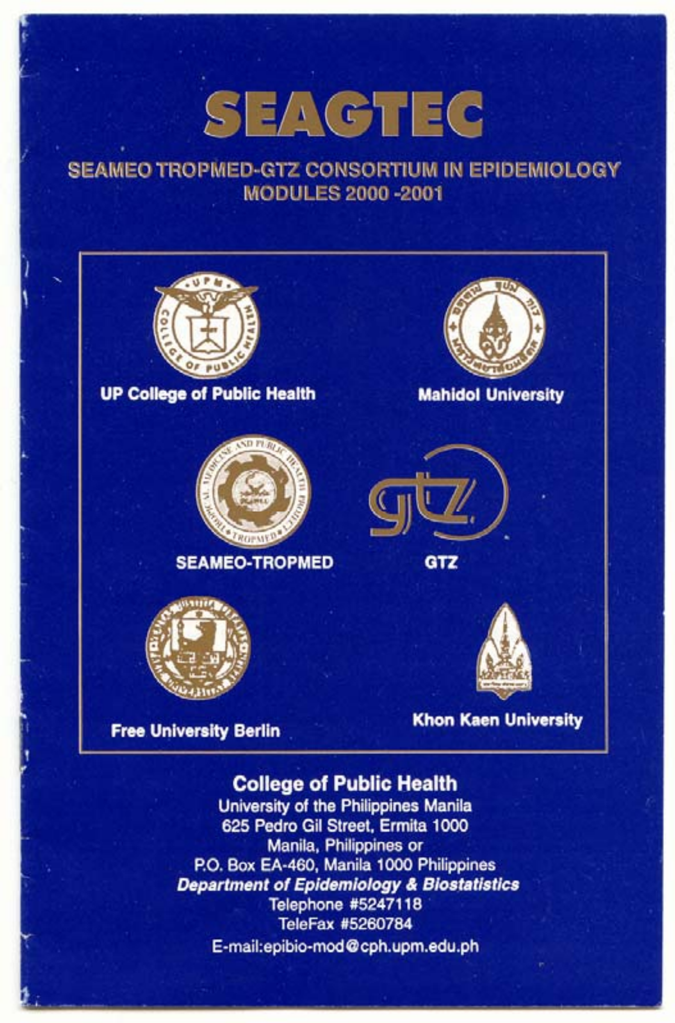
History

Besides cooperating in several projects in SE Asia

one major activity 1990-2000:

M.Sc. in Epidemiology

at UP Manila under participation of the universities of Mahidol (Bangkok, Th), Khon Kaen(Th), FU Berlin, UP Manila (Ph)



SEAGTEC

SEAMEO TROPMED-GTZ CONSORTIUM IN EPIDEMIOLOGY
MODULES 2000 -2001

UP College of Public Health

Mahidol University

SEAMEO-TROPMED

GTZ

Free University Berlin

Khon Kaen University

College of Public Health
University of the Philippines Manila
625 Pedro Gil Street, Ermita 1000
Manila, Philippines or
P.O. Box EA-460, Manila 1000 Philippines
Department of Epidemiology & Biostatistics
Telephone #5247118
TeleFax #5260784
E-mail:epibio-mod@cph.upm.edu.ph

Cooperation Projects with SE ASIA

- Partner: Faculty for Public Health, Mahidol University, Bkk, Thailand
- Prof. Chukiat Viwatwongkasem (Counterpart)
- Funding: DFG, BMZ und National Research Council of Thailand (NRCT)

Supported by German Research Foundation (DFG) in Co-operation with the National Research Coun of Thailand (NR
ภาควิชาชีวสถิติ
DEPARTMENT OF BIOSTATISTICS
ห้อง ROOM
NAME
3208 JUNYA PATTARAARCHACHAI
3208 CHAWEEWON BOONSHUYAR
3208 CHUKIAT VIWATW
3204 NUTKAMOL C
3204
3206 DECHAVUDH
3212 THAVATCHAI
PHORNPHUN C
2
ภาควิชาชีวสถิติ
DEPARTMENT OF BIOSTATISTICS
ชื่อ
ห้อง ROOM
NAME

Capture-Recapture Procedures in Public Health

Welcome all paticipants to a special lecture
"Capture-Recapture Procedures in Public Health"
Speaker : Prof. Dr. Dankmar Böhning
Organized by Department of Biostatistics faculty of Public Health,
Mahidol University. March 15 - April 16, 2004

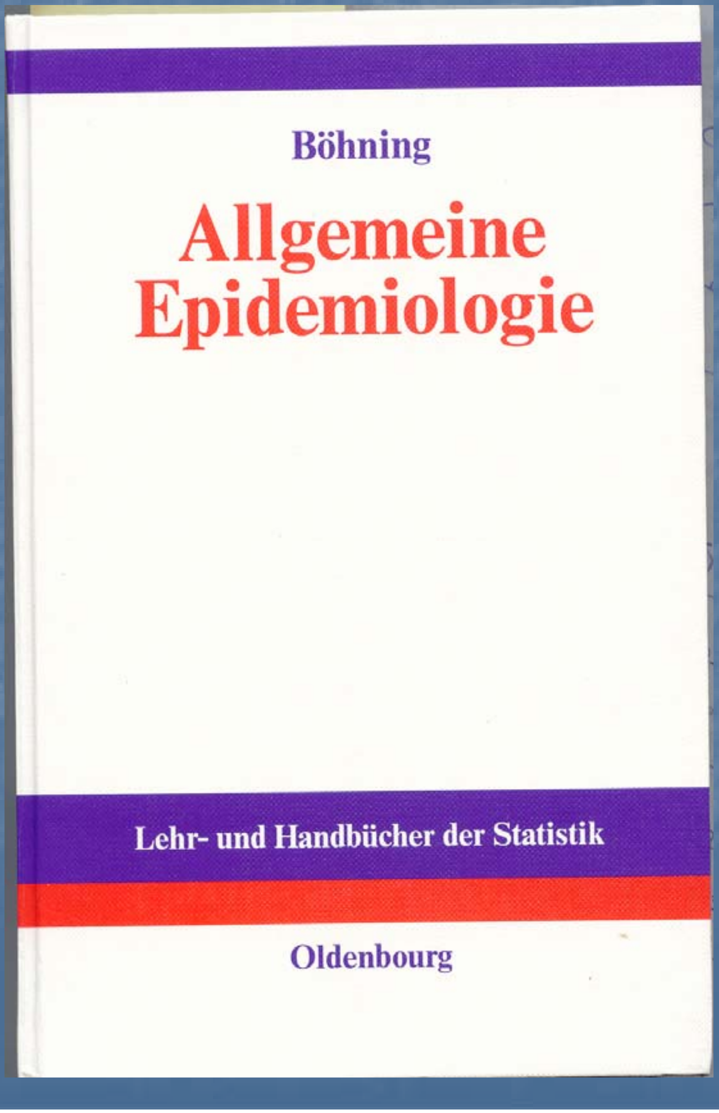
Surveillance Project on Illicit Drug Use in Thailand using Truncated Counting Distributions

Overview

- History
- General Topics
- Current Areas of Interest
- Research Areas in Preperation

General Topics

- General Epidemiology
- Problems of Inference in Epidemiology
- Epidemiologic Modelling



Böhning

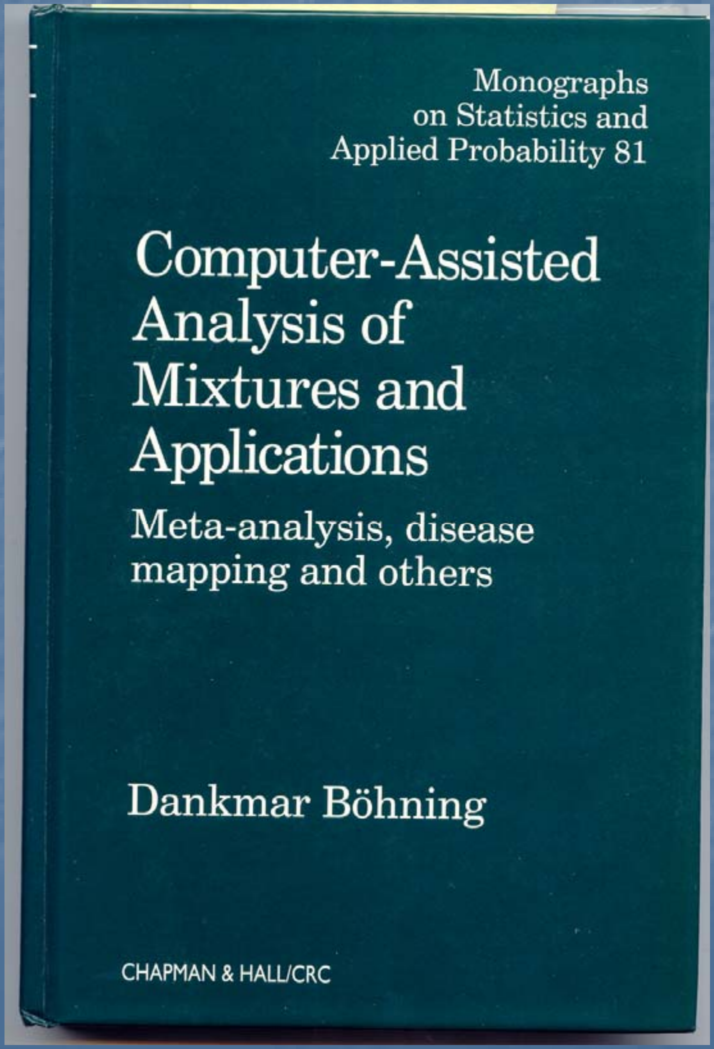
Allgemeine Epidemiologie

Lehr- und Handbücher der Statistik

Oldenbourg

General Topics

- Mixture models
- Applications in Biometry and Epidemiology



Monographs
on Statistics and
Applied Probability 81

Computer-Assisted Analysis of Mixtures and Applications

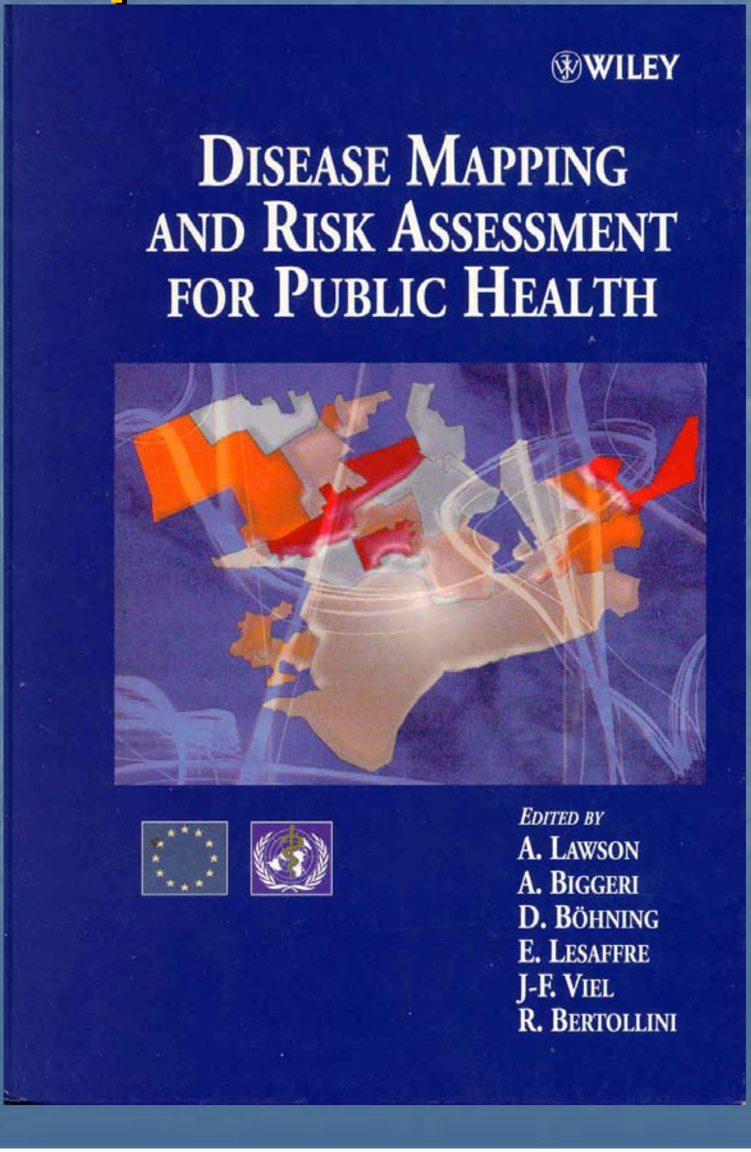
Meta-analysis, disease mapping and others

Dankmar Böhning

CHAPMAN & HALL/CRC

General Topics

- Disease Mapping and Geographical Epidemiology
- Smoothed Estimates of Geographical Risk



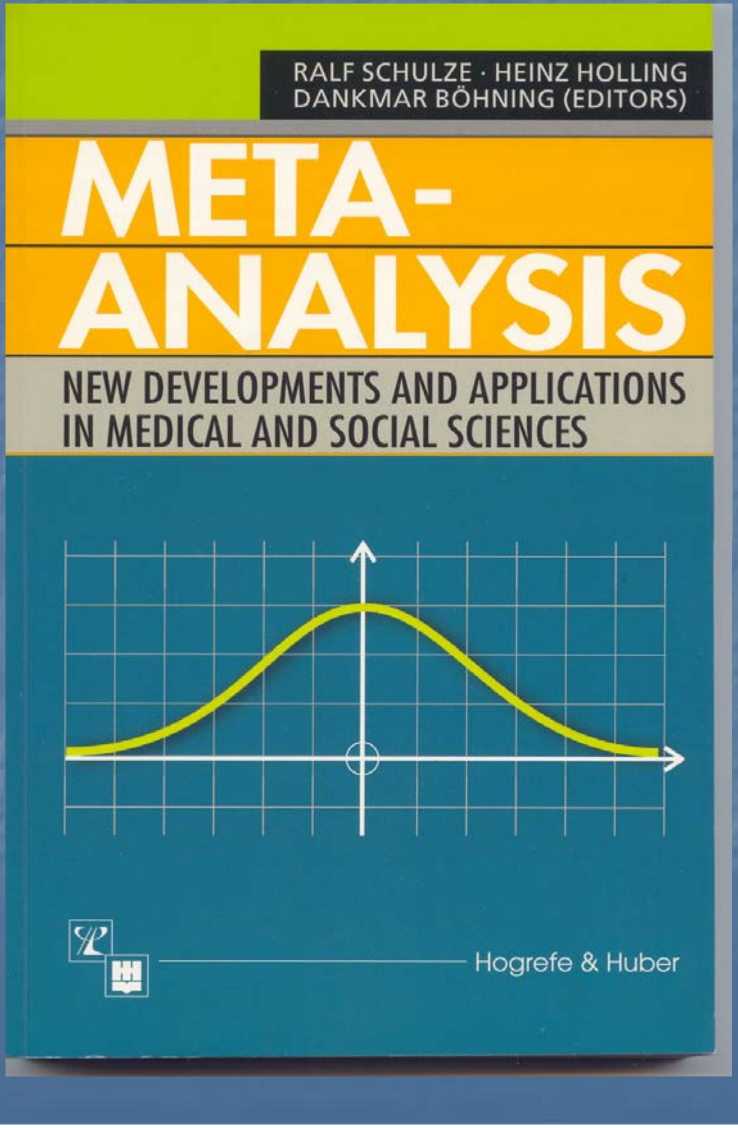
WILEY

DISEASE MAPPING AND RISK ASSESSMENT FOR PUBLIC HEALTH

EDITED BY
A. LAWSON
A. BIGGERI
D. BÖHNING
E. LESAFFRE
J-F. VIEL
R. BERTOLLINI

General Topics

- Systematic Reviews and Meta-Analysis
- Heterogeneity, Covariate and Publications Bias Modelling
- Unifying Concept



RALF SCHULZE · HEINZ HOLLING
DANKMAR BÖHNING (EDITORS)

META-ANALYSIS

NEW DEVELOPMENTS AND APPLICATIONS IN MEDICAL AND SOCIAL SCIENCES

Hogrefe & Huber

Personal Background: Editorial Board

- Biometrics (1992)
- Statistical Modelling (1999)
- Biometrical Journal (2004)

Overview

- History
- General Topics
- Current Areas of Interest
- Research Areas in Preperation

Modelling Effect- and Nuisance Parameter in Multi-Center Studies

- Typical setting: treatment- and control Arm
- For treatment arm:

x^T number of events

P^T person-time

λ^T event rate

- For control arm:

$$x^C, P^C, \lambda^C$$

Modelling Effect in Multi-Center Studies: A Typical Example

Center	Treatment		Control	
	events x_i^T	person-time P_i^T	events x_i^C	under risk P_i^C
1	29	116	21	113
2	6	73	3	121
3	30	50	23	50
4	23	180	15	172
…	…	…	…	…
59	15	60	4	60

Modelling Effect- and Nuisance Parameter in Multi-Center Studies

- parameter of interest:

  risk ratio: $\theta = \lambda^T / \lambda^C$

- nuisance parameter:

  λ^C event rate in control arm

poisson log-likelihood (for one center):

$$-\lambda^T P^T + x^T \log(\lambda^T P^T) \; -\lambda^C P^C + x^C \log(\lambda^C P^C)$$

Modelling Effect- and Nuisance Parameter in Multi-Center Studies

$$-\lambda^T P^T + x^T \log(\lambda^T P^T) \;\; -\lambda^C P^C + x^C \log(\lambda^C P^C)$$

becomes using $\theta = \lambda^T / \lambda^C$ or $\lambda^T = \theta\lambda^C$

$$-\theta\lambda^C P^T + x^T \log(\theta\lambda^C P^T) \;\; -\lambda^C P^C + x^C \log(\lambda^C P^C)$$

Keeping the parameter of interest fixed and maximizing for the nuisance parameter ...

$$\hat{\lambda}^C = \frac{x^C + x^T}{P^C + \theta P^T}$$

replacing λ^C by its estimate $\hat{\lambda}^C$

$$-\theta\hat{\lambda}^C P^T + x^T \log(\theta\hat{\lambda}^C P^T) - \hat{\lambda}^C P^C + x^C \log(\hat{\lambda}^C P^C)$$

leads to the beautiful simple Profile Log-likelihood ...

$$x^T \log(\theta) \; - (x^T + x^C)\log(P^C + \theta P^T)$$

... building the profile over all centers:

$$\sum_{i=1}^{k} x_i^T \log(\theta_i) \; - (x_i^T + x_i^C)\log(P_i^C + \theta_i P_i^T)$$

Advantages

- nuisance parameter eliminated
- Profile likelihood is simple (in this case):

$$\sum_{i=1}^{k} x_i^T \log(\theta_i) \ - (x_i^T + x_i^C) \log(P_i^C + \theta_i P_i^T)$$

- beneficial not only for effect structures but also for covariance structures (simplification of Fisher information)

Problems looked at: homogenous case

$\theta_i = \theta$ for all centers $i = 1, \ldots, k$:

$$\sum_{i=1}^{k} x_i^T \log(\theta) \ - (x_i^T + x_i^C)\log(P_i^C + \theta P_i^T)$$

Score equation for profile MLE gives:

$$\hat{\theta} = \frac{\sum_{i=1}^{k} x_i^T n_i^C w_i(\widehat{\theta})}{\sum_{i=1}^{k} x_i^C n_i^T w_i(\widehat{\theta})}, \quad w_i(\widehat{\theta}) = \frac{1}{n_i^C + \hat{\theta}\, n_i^T}$$

Problems looked at: homogenous case

$$\hat{\theta} = \frac{\sum_{i=1}^{k} x_i^T n_i^C w_i(\widehat{\theta})}{\sum_{i=1}^{k} x_i^C n_i^T w_i(\widehat{\theta})}, \quad w_i(\widehat{\theta}) = \frac{1}{n_i^C \overset{1}{\boxed{+\hat{\theta}\, n_i^T}}}$$

- Close connection to Mantel-Haenszel:
  - arms balanced then: PMLE = MH
  - Non-sparsity: PMLE and MH close
  - Sparsity: PMLE more efficient

Overview

- History
- General Topics
- Current Areas of Interest
- Research Areas in Preperation

Modelling Effect-Heterogeneity in Multi-Center Studies: Unobserved Heterogeneity

- Allowing for unobserved heterogeneity leads to mixtures of profile log-likelihoods

$$\sum_{i=1}^{k} \log \int_{\theta} [\theta^{x_i^T} / (P_i^C + \theta P_i^T)^{x_i^T + x_i^C}] \, Q(d\theta)$$

- where mixing distribution can be parametric
- or non-parametric
  - strong results on NPMLE possible using convex theory
  - estimation with EM or global ascent algorithms

Modelling Effect-Heterogeneity in Multi-Center Studies: Unobserved Heterogeneity

- Comparison with other approaches such as
  - approximating normal (problem: use empirical estimate of trial variance)

$$\sum_{i=1}^{k} \log \int_{\lambda^C} \phi((z_i - \log\theta)/\sigma_i)\, Q(d \log\theta)$$

$$\text{where } z_i \text{ obs. log-rate ratio and } \sigma_i^2 = 1/x_i^T + 1/x_i^C$$

  - multi-level approach (a la Murray Aitkin)

$$\sum_{i=1}^{k} \log \int_{\lambda^C} [\exp(-\theta\lambda^C P_i^T)(\theta\lambda^C P_i^T)^{x_i^T} \exp(-\lambda^C P_i^C)(\lambda^C P_i^C)^{x_i^C}]\, Q(d\lambda^C)$$

Modelling Effect-Heterogeneity in Multi-Center Studies: Observed Heterogeneity-Covariate Information

- Often additonal trial information is available s.a. study date, treatment modifications, patient characteristics
- Suppose information is captured in a covariate vector

z_i for center i: (GLM-type formulation)

$$\theta_i = \exp(\beta_0 + \beta' z_i)$$

Modelling Effect-Heterogeneity in Multi-Center Studies: Observed Heterogeneity-Covariate Information

Log-likelihood becomes

$$\sum_{i=1}^{k} x_i^T \log\theta_i - (x_i^T + x_i^C)\log(P_i^C + \theta_i P_i^T) \quad \text{using } \theta_i = \exp(\beta_0 + \beta' z_i)$$

$$= \sum_{i=1}^{k} x_i^T (\beta_0 + \beta' z_i) - (x_i^T + x_i^C)\log[P_i^C + \exp(\beta_0 + \beta' z_i)P_i^T]$$

- Strong results possible:
  - Hessian has simple structure
  - Hessian has lower bound (lower bound algorithm possible)
  - Guaranteed convergence to MLE

Overview

- History
- General Topics
- Current Areas of Interest
- Research Areas in Preperation

Capture-Recapture Procedures based upon Counting Distributions

- Basic objective of CR: estimate population size
- In particular of interest in areas where direct counting is difficult such as
  - a wildlife population (historic genesis)
  - how many people drive a car without license?
  - how many practicing physicians are alcohol dep.?
  - how may cases of a disease remain undetected?
- Adjustment for undercount

How many cases ***N*** in a population?

- Some mechanism identifies *n* cases
- p_0 probability of being **not** identified by the mechanism
- **Then:**

$$N = N\,p_0 + (1 - p_0)\,N$$

$= \text{unobserved} + \text{observed cases}$

$= N\,p_0 + n$

$$\widehat{N} = n/(1 - p_0)$$

(Horwitz-Thompson)

Horwitz-Thompson-Approach seems easy, but …

inclusion probability often unknown

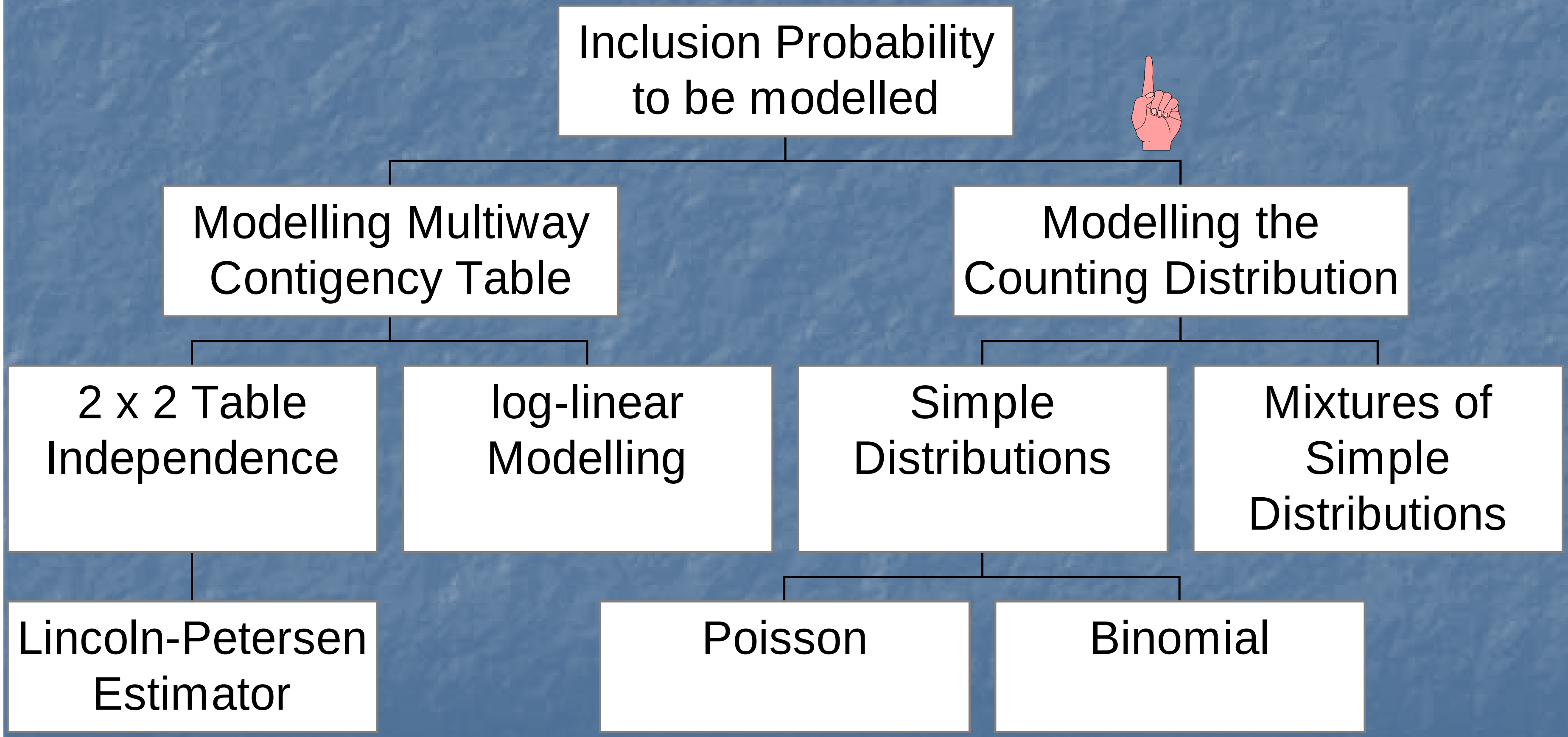
and consequently,

approaches differ in the way they estimate the inclusion probability,

or in other words, how they

model p_0

Developments



Inclusion Probability to be modelled

Modelling Multiway Contigency Table

Modelling the Counting Distribution

2 x 2 Table Independence

log-linear Modelling

Simple Distributions

Mixtures of Simple Distributions

Lincoln-Petersen Estimator

Poisson

Binomial

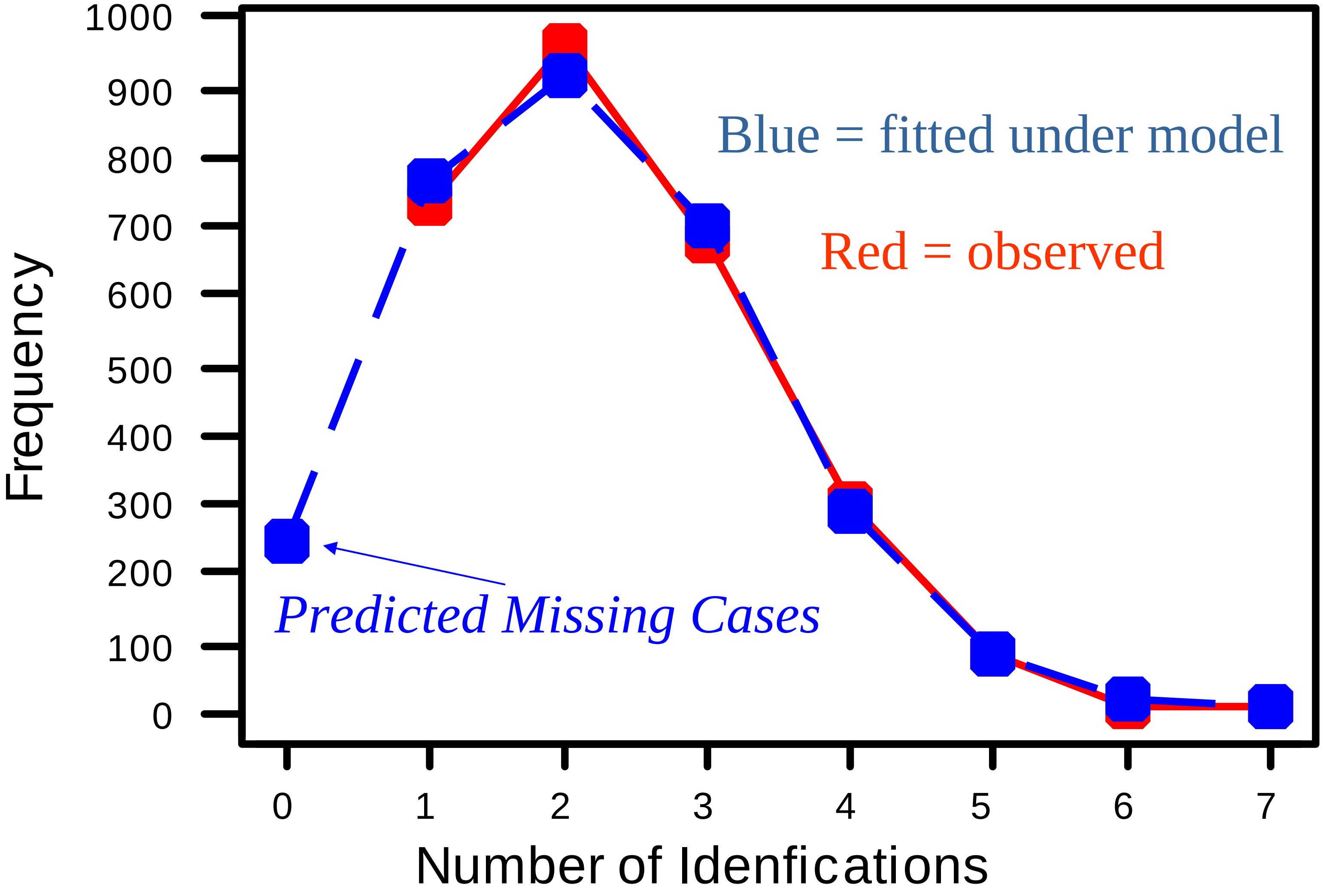
The Counting Distribution

... occurs when the mechanism can catch multiple identifications (s.a. police identifies and expells an illegal immigrant several times)

Count of identifications *i*	**Frequency of counts with *i* identifications**	**observed**
0	n_0	no
1	n_1	yes
2	n_2	yes
3	n_3	yes
4	n_4	yes
…	…	…

Distribution of Observed and Predicted Counts of Sources

for fictional data of multiple identifications



The Counting Distribution: A historic Example

- McKendrick´s cholera data
- Village in India had households with cholera cases $n_1=32$, $n_2=16$, $n_3=6$, $n_4=1$
- McKendrick ignored the houses with no cases
- Constructed an estimate (moment) based upon a Poisson assumption for the counts

Cholera Epidemic in an Indian Village (1915-1920)

House not affected, no cases

House affected, no cases

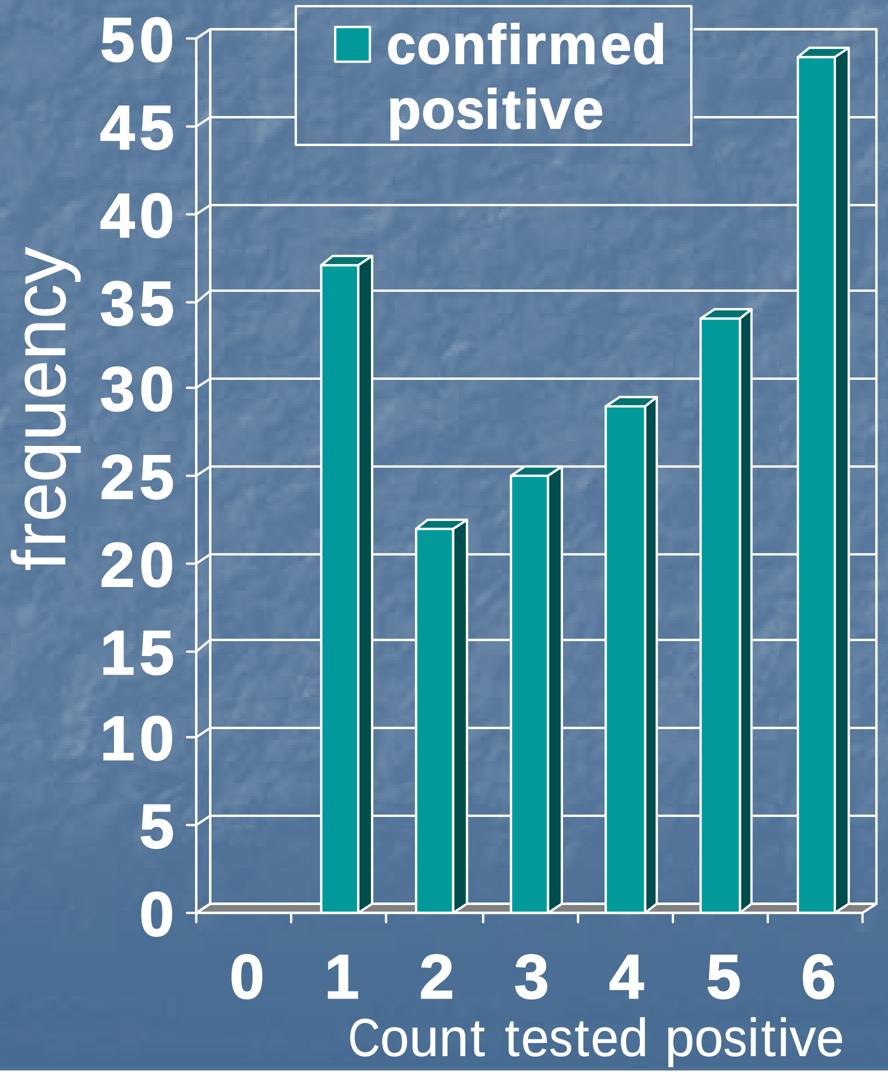
House affected, *m* cases

The counting distribution: a recent example from screening

- Lloyd & Frommer (2004, Applied Statistics) screening for bowel cancer
- 38,000 men screened in Sidney at 6 consecutive days by means of self-tesing for blood in stools
- 3,000 tested positively at least once and cancer status evaluated
- 196 were confirmed positive to have bowel cancer
- How many of 35,000 unconfirmed negative have bowel cancer?

The counting distribution: a recent example from screening

- frequency n_0 of those tested negative at all 6 times with bowel cancer is unknown
- an estimate of n_0 might be constructed from the distribution $n_{1,}$ $n_{2,}$ n_3 .... of counts



Simple Distributional Count Models

Poisson (for unobservable counts)

$$f(y,\theta) = e^{-\theta}\theta^{y} / y! \ , \ y = 0,1,2 \ldots$$

truncated Poisson (for observable counts)

$$f(y,\theta) = \frac{1}{1-e^{-\theta}} \ e^{-\theta}\theta^{y} / y! \ , \ y = 1,2 \ldots$$

Predicted Probability of a Zero:

$$p_0 = f(y,\theta) = e^{-\theta}$$

Simple Distributional Count Models

after θ is identified ...

.... probability of a zero count:

$$p_0 = f(y=0,\theta) = e^{-\theta}$$

$$\Rightarrow \widehat{N} = \frac{n}{1-p_0} = \frac{n}{1-e^{-\theta}}$$

ML-Estimation in Zero-Truncated Poisson Models

Step 1: suppose $\hat{n}_0$ would be available

$$\hat{\theta} = \frac{1}{n + \hat{n}_0} \sum_{i=1}^{m} i \, n_i$$

Step 2: suppose $\hat{\theta}$ would be available

$$\widehat{N} = \frac{n}{1 - p_0} = \frac{n}{1 - e^{-\hat{\theta}}} \Rightarrow \hat{n}_0 = \widehat{N} - n = n \frac{e^{-\hat{\theta}}}{1 - e^{-\hat{\theta}}}$$

EM-Algorithm

Step 1 (M-Step): suppose $\hat{n}_0$ would be available

$$\hat{\theta} = \frac{1}{n + \hat{n}_0} \sum_{i=1}^{m} i \; n_i$$

Step 2 (E-Step): suppose $\hat{\theta}$ would be available

$$\hat{n}_0 = E(n_0 \,|\, \hat{\theta}; n_1, n_2, \ldots) = n \frac{p_0}{1 - p_0} = n \frac{e^{-\hat{\theta}}}{1 - e^{-\hat{\theta}}}$$

ML-Estimation in Zero-Truncated Count Models

general count distribution

$$f(y,\theta)\ ,\ y = 0,1,2,...$$

assoc. zero-truncated distribution

$$\frac{1}{1-f(0,\theta)}f(y,\theta)\ ,\ y = 1,2,...$$

EM-Algorithm

Step 1 (M-Step): suppose $\hat{n}_0$ is given:

$$\hat{\theta} = MLE, \text{ based upon } \widehat{n_0}, n_1, n_2, \ldots$$

Step 2 (E-Step): suppose $\hat{\theta}$ is given:

$$\hat{n}_0 = E(n_0 \mid \hat{\theta}; n_1, n_2, \ldots) = n\frac{p_0}{1-p_0} = n\frac{f(0,\hat{\theta})}{1-f(0,\hat{\theta})}$$

More flexible and robust approach through mixtures

- Simple counting sources distributions such as Binomial and Poisson require assumptions such as homogeneity of identification probabilities that are seldom met in reality
- allowing the identification probability to vary in unobserved sub-populations will be more realistic

The mixture approach in a nutshell

mixture density:

$$f(y,\theta) \;\; = f(y,\lambda_1)q_1 + \ldots \; + f(y,\lambda_k)q_k$$

$f(y,\lambda)$ is component density

(Example: $f(y,\lambda) = e^{-\lambda}\lambda^y / y!$)

$\theta \; = \begin{pmatrix} \lambda_1 & \ldots & \lambda_k \\ q_1 & \ldots & q_k \end{pmatrix}$ is mixing distribution

Nested EM-Algorithm

Step 1 (M-Step): suppose $\hat{n}_0$ is given:

$$\hat{\theta} = MLE \text{ of mixing distribution } \theta = \begin{pmatrix} \lambda_1 & \ldots & \lambda_k \\ q_1 & \ldots & q_k \end{pmatrix}$$

provided by EM algorithm for mixtures

Step 2 (E-Step): suppose $\hat{\theta}$ is given:

$$\hat{n}_0 = E(n_0 \mid \hat{\theta}; n_1, n_2, \ldots) = n\frac{p_0}{1-p_0}$$

$$= n\frac{f(0,\hat{\theta})}{1-f(0,\hat{\theta})} = n\frac{\hat{q}_1 e^{-\hat{\lambda}_1} + \ldots + \hat{q}_k e^{-\hat{\lambda}_k}}{1-(\hat{q}_1 e^{-\hat{\lambda}_1} + \ldots + \hat{q}_k e^{-\hat{\lambda}_k})}$$

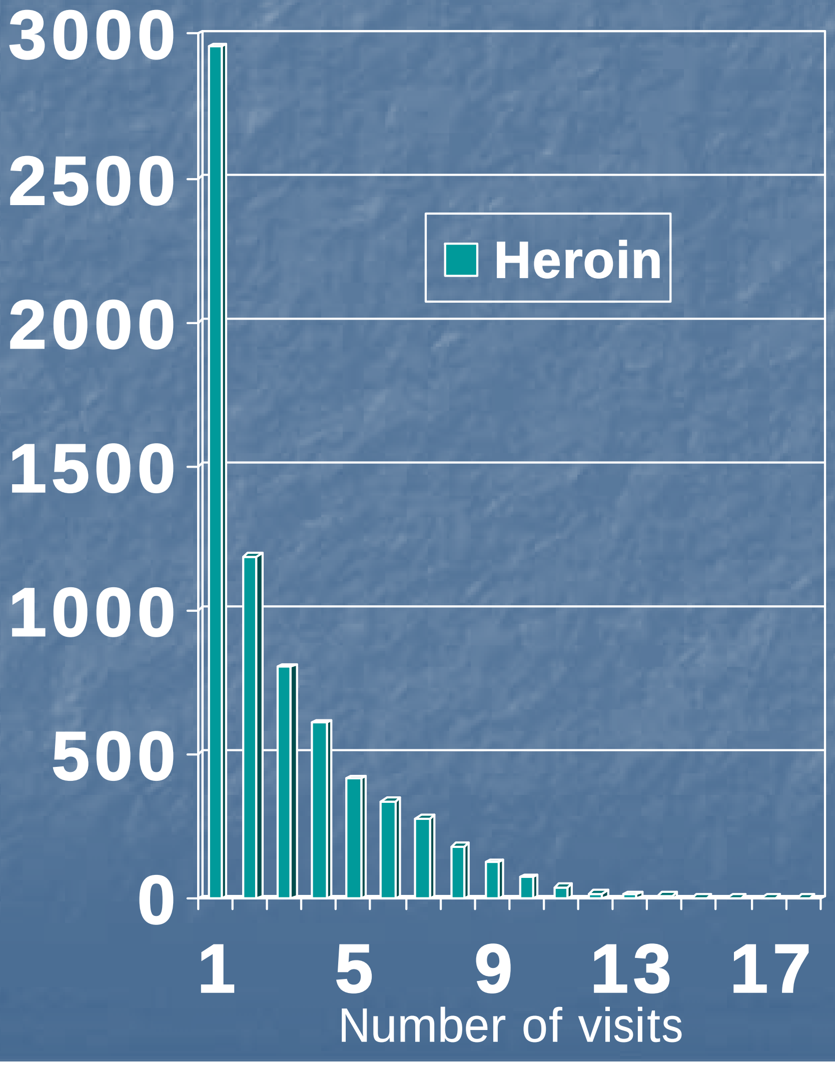
Application: surveillance study on drug use in Thailand

คณะสาธารณสุขศาสตร์ อาคาร
FACULTY OF PUBLIC HEALTH Bld.
Welcome all paticipants to a special lecture
"Capture-Recapture Procedures in Public Health"
Speaker: Prof. Dr. Dankmar Böhning
Organized by Department of Biostatistics, Faculty of Public Health
Mahidol University. March 15 - April 16, 2004

- Ministry of Public Health (Th) collects routinely data on drug use via the ONCB on drug users visiting treatment institutions
- In a pilot study (Böhning, Busaba, Chukiat et al. 2004 *EUJE)* CR-Poisson mixture model applied to data from 2002 (last quarter)
- Major emphasis on heroin and metamphetamin users

Application: surveillance study on drug use in Thailand

- Count distribution (counting number of visits) for heroin users
- n = 7,048 observed heroin users (2001, 4)

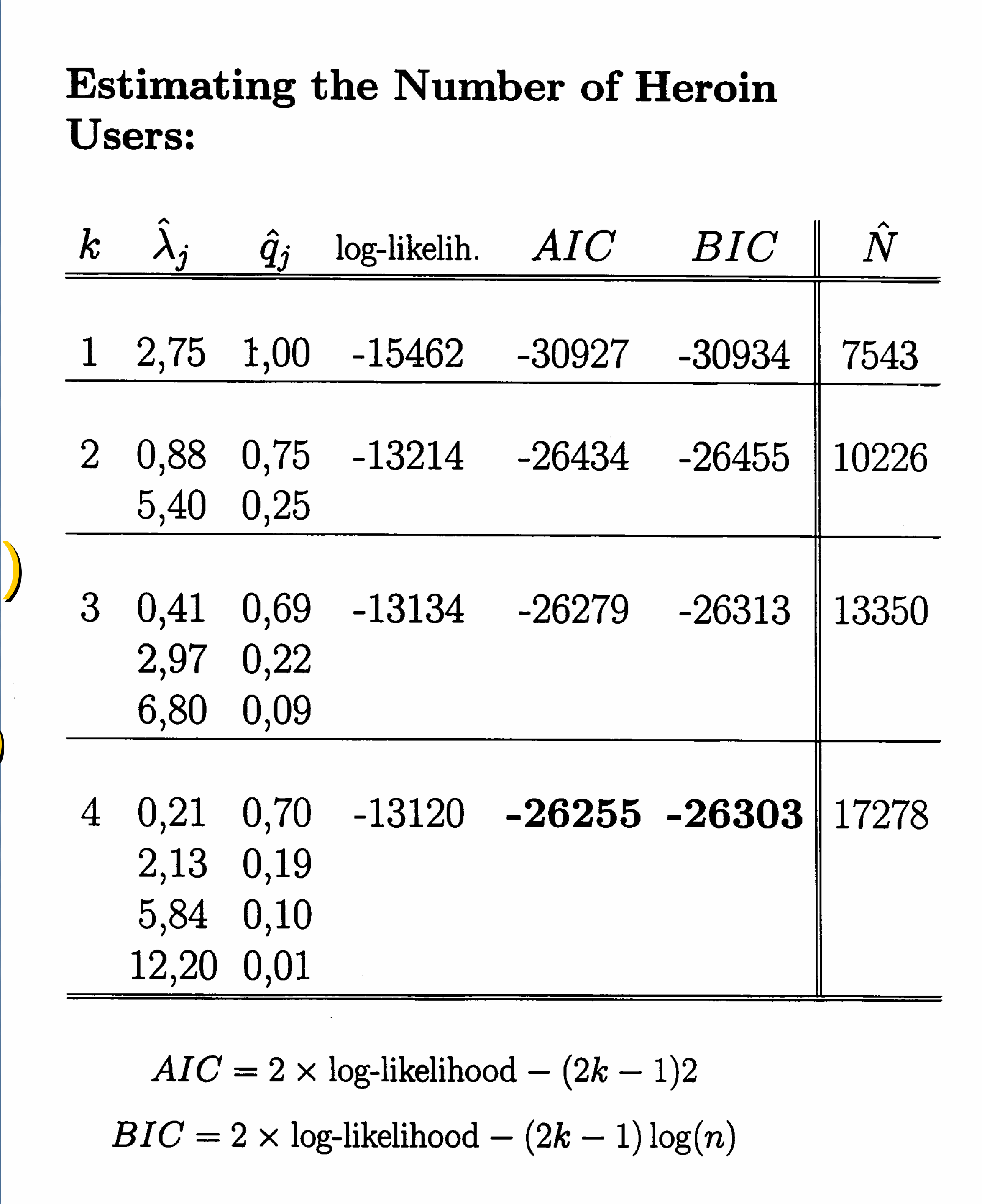


Counting contacts to treatment institutions not uncommon

- Previous modelling done primarily by practitioners with publications in
  - Addiction, Addiction Research & Theory, Journal of Drug Issues, Journal of Quantitative Cirminology
- Modelling uses primarily simple Poisson
  - simple to understand, to apply and use, and to communicate
  - however: often not appropriate
- better: semi-parametric models for counts such as Poisson mixtures

Some results

- n=7,048 (observed)
- N=17,278
- N-n=10,230 (hidden)
- Ratio: observed/hidden=0.69



Estimating the Number of Heroin Users:

k	$\hat{\lambda}_j$	$\hat{q}_j$	log-likelih.	AIC	BIC	$\hat{N}$
1	2,75	1,00	-15462	-30927	-30934	7543
2	0,88 5,40	0,75 0,25	-13214	-26434	-26455	10226
3	0,41 2,97 6,80	0,69 0,22 0,09	-13134	-26279	-26313	13350
4	0,21 2,13 5,84 12,20	0,70 0,19 0,10 0,01	-13120	**-26255**	**-26303**	17278

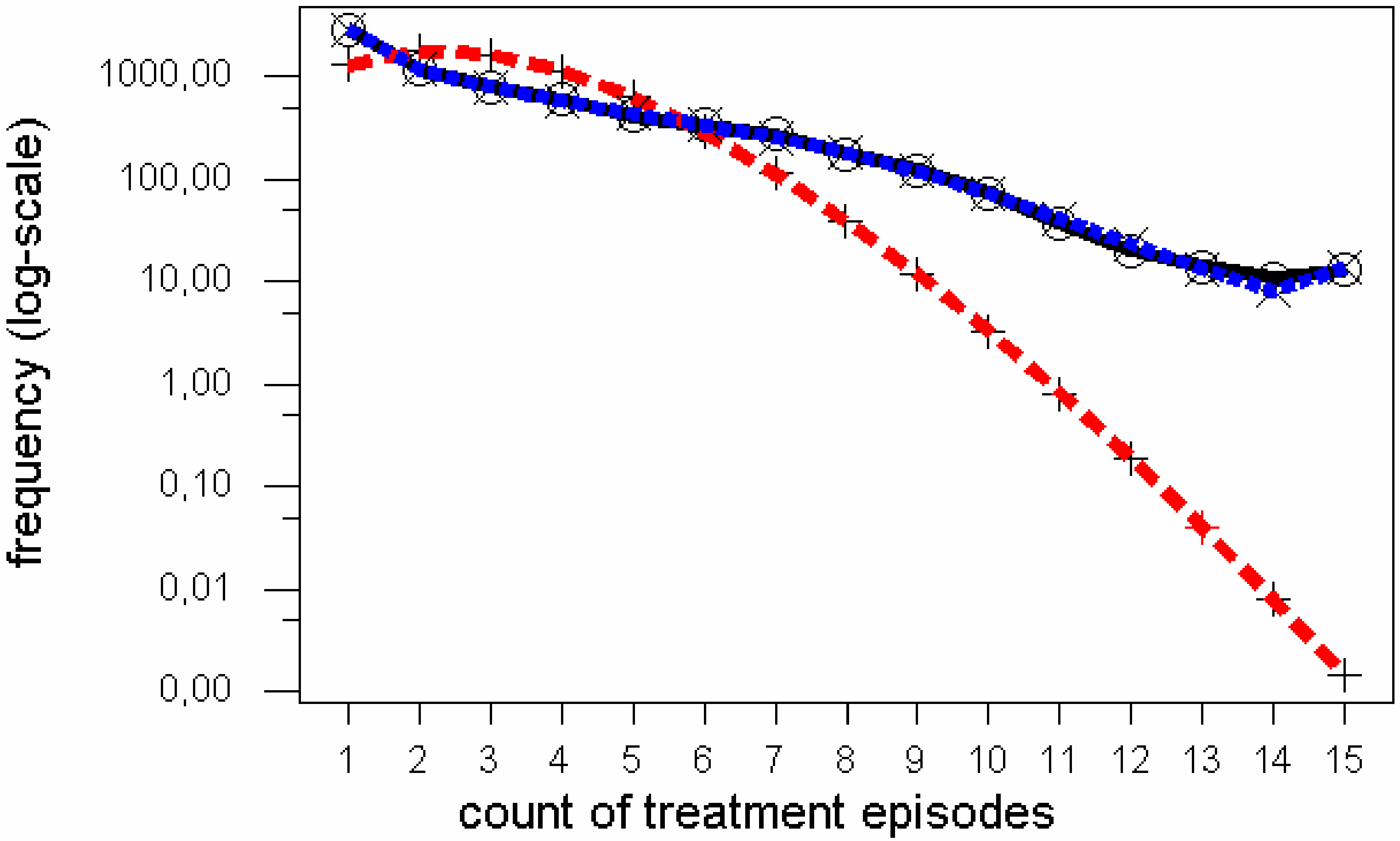
$$AIC = 2 \times \text{log-likelihood} - (2k-1)2$$

$$BIC = 2 \times \text{log-likelihood} - (2k-1)\log(n)$$

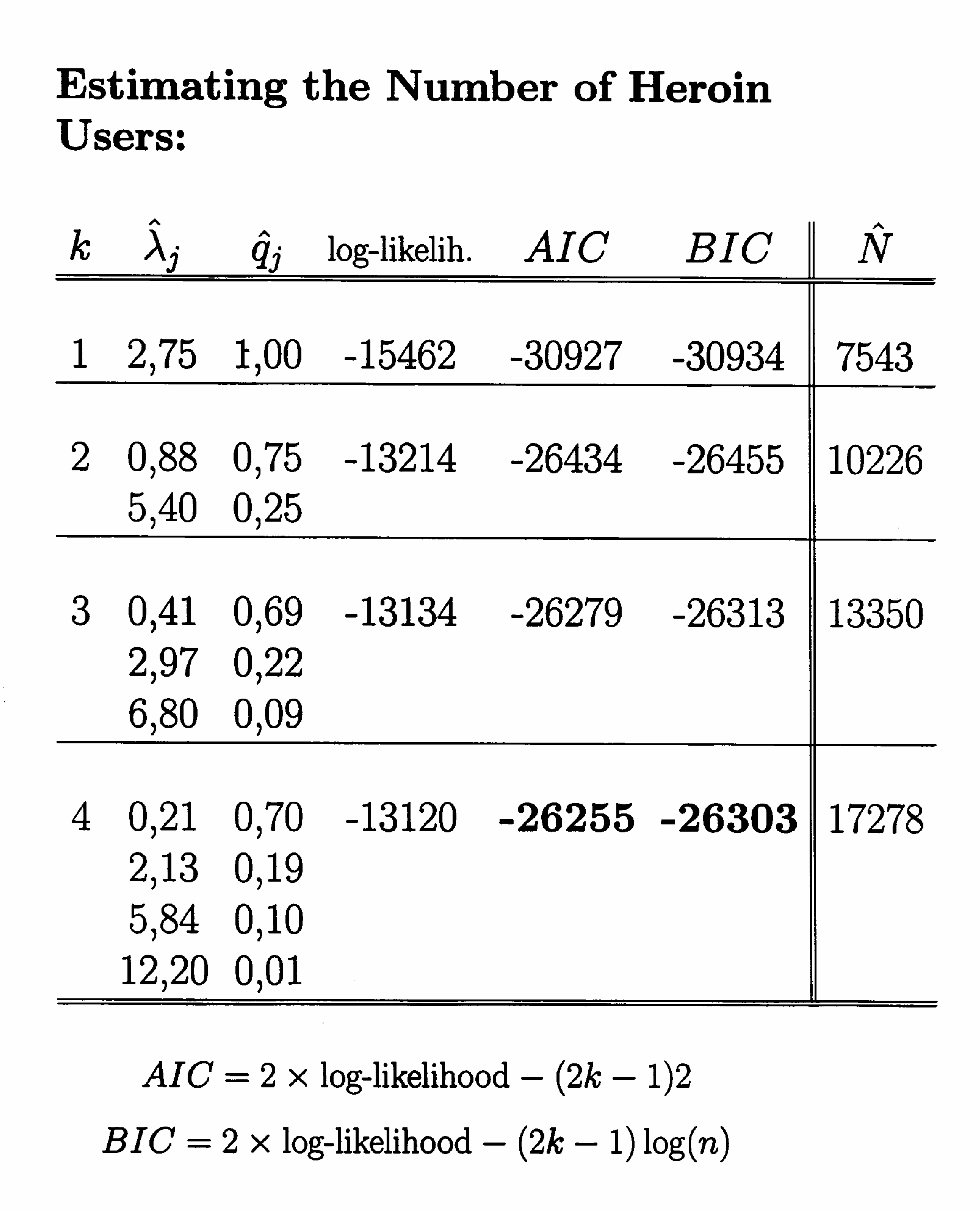
Pilot study for Bangkok, 2001 (4)

count distributions of treatment episodes for heroin users

(empirical = black; simple Poisson = red; Poisson mixture = blue)



take another look

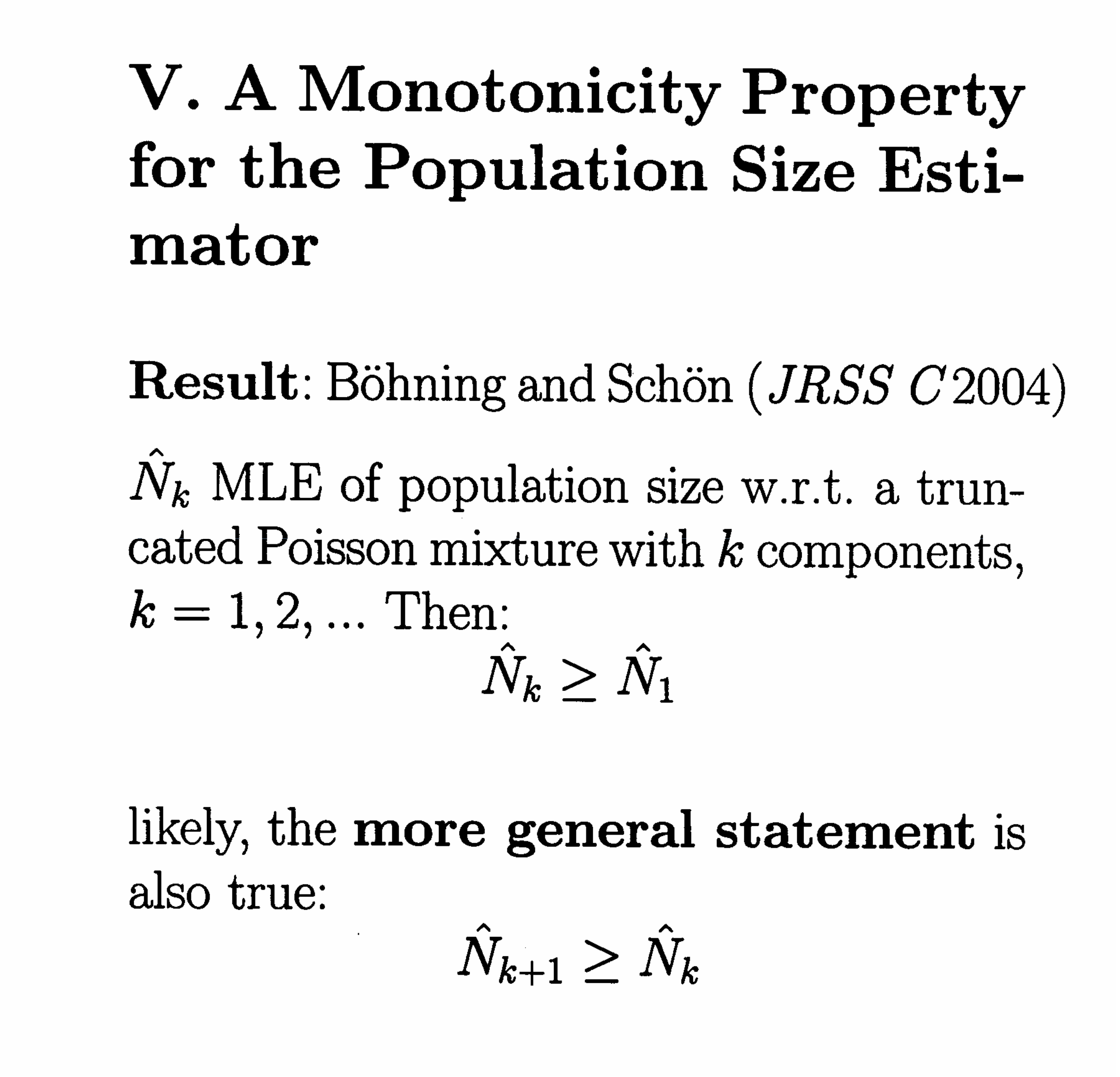


Estimating the Number of Heroin Users:

k	$\hat{\lambda}_j$	$\hat{q}_j$	log-likelih.	AIC	BIC	$\hat{N}$
1	2,75	1,00	-15462	-30927	-30934	7543
2	0,88	0,75	-13214	-26434	-26455	10226
	5,40	0,25				
3	0,41	0,69	-13134	-26279	-26313	13350
	2,97	0,22				
	6,80	0,09				
4	0,21	0,70	-13120	**-26255**	**-26303**	17278
	2,13	0,19				
	5,84	0,10				
	12,20	0,01				

$$AIC = 2 \times \text{log-likelihood} - (2k-1)2$$

$$BIC = 2 \times \text{log-likelihood} - (2k-1)\log(n)$$



V. A Monotonicity Property for the Population Size Estimator

Result: Böhning and Schön (*JRSS C* 2004)

$\hat{N}_k$ MLE of population size w.r.t. a truncated Poisson mixture with k components, $k = 1, 2, \ldots$ Then:

$$\hat{N}_k \geq \hat{N}_1$$

likely, the **more general statement** is also true:

$$\hat{N}_{k+1} \geq \hat{N}_k$$

proof uses:

- a) Jensen's inequality
- b) mean of the MLE of the mixing distribution = sample mean

Overview

- History
- General Topics
- Current Areas of Interest
- Research Areas in Preperation

Concluding Remarks

Open Problems and Research Questions

- Standard errors and confidence intervals
- Suitable modification of resampling techniques
- Validation studies
- Comparison to other approaches (Pollock-Norris or Zelterman)
- … Mixtures of binomials

very recent work in perspective

- truncated mixture of Poisson distributions
- or ...
- mixture of truncated Poisson distributions

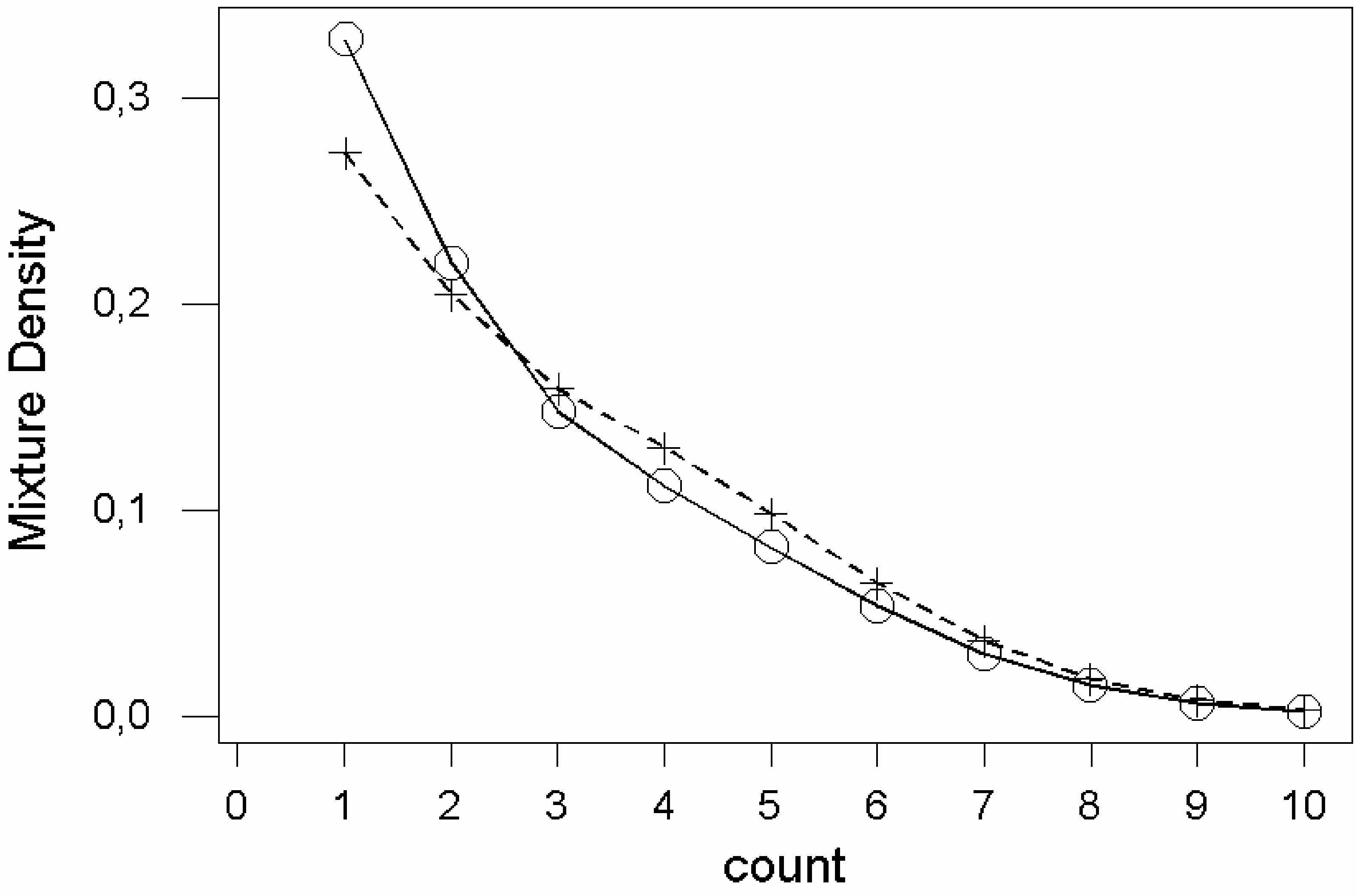
truncated Poisson mixture (dual model)

$$\frac{\sum_{j=1}^{k} q_j \; Po(y, \lambda_j)}{1 - \sum_{j=1}^{k} q_j \; Po(0, \lambda_j)}$$

mixture of truncated Poissons (primal model)

$$\sum_{j=1}^{k} q_j^{'} \frac{Po(y, \lambda_j^{'})}{1 - Po(0, \lambda_j^{'})}$$

Illustration: dual model (ring) and primal model (+) use equal weights and component means 1 and 4



truncated Poisson mixture (dual model)

$$\frac{\sum_{j=1}^{k} q_j \; Po(y,\lambda_j)}{1-\sum_{j=1}^{k} q_j \; Po(0,\lambda_j)}$$

- close to the original problem, easy to understand and to communicate
- But technical difficult, because of non-linearity

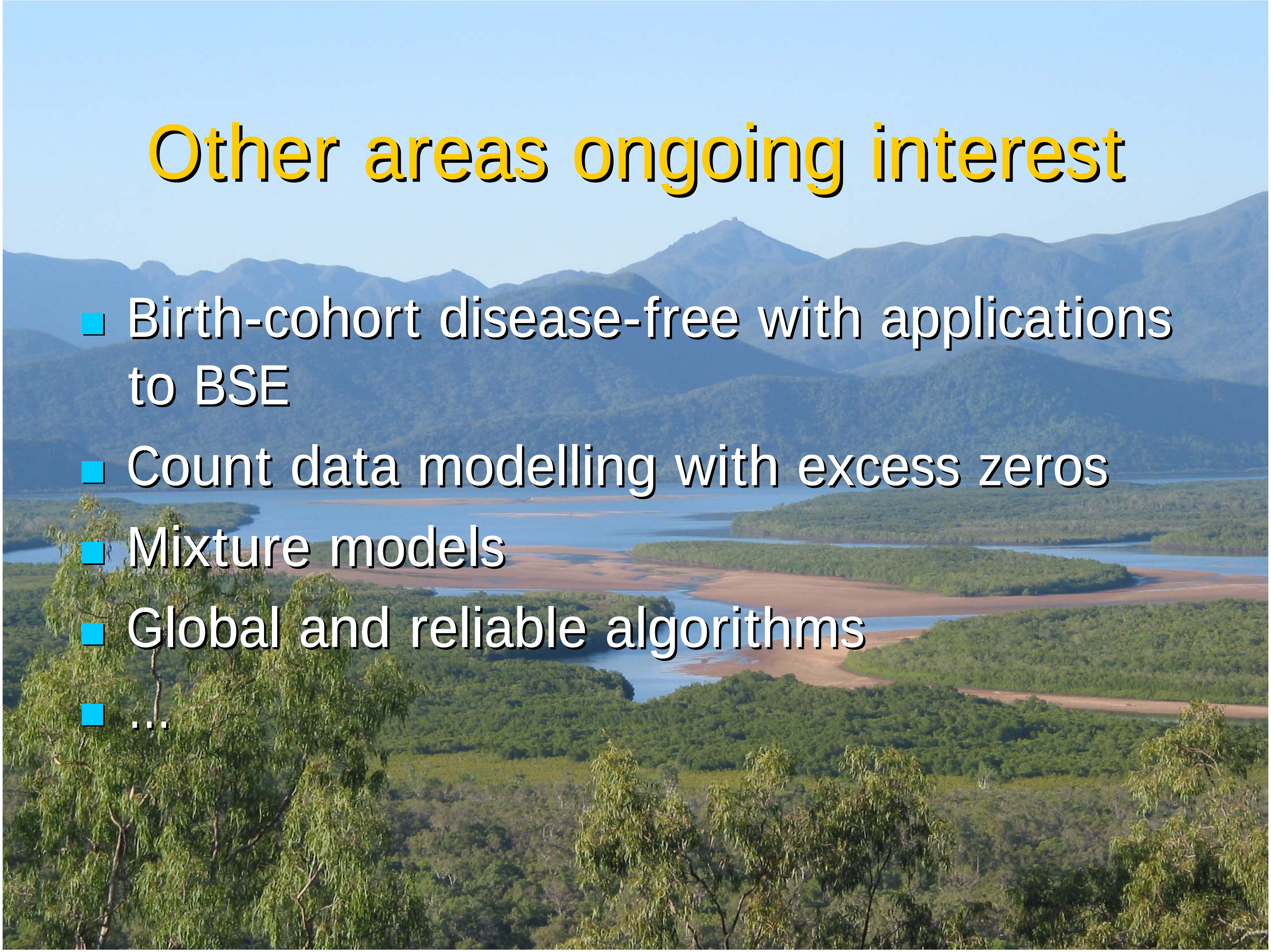
mixture of truncated Poissons (primal model)

$$\sum_{j=1}^{k} q_j^{'} \frac{Po(y, \lambda_j^{'})}{1- \ Po(0, \lambda_j^{'})}$$

- less close to the original problem
- but convex problem with strong results available on NPMLE and global ML estimation

How are dual and primal model related?

- Böhning and Kuhnert (2005, JASA)
- Both share the same likelihood surfaces
- MLEs can be explicitly transformed into each other
- $\widehat{N} = \widehat{N'}$



Other areas ongoing interest

- Birth-cohort disease-free with applications to BSE
- Count data modelling with excess zeros
- Mixture models
- Global and reliable algorithms
- ...