Models for Replicated Discrimination Tests: A Synthesis of Latent Class Mixture Models and Generalized Linear Mixed Models

Rune Haubo Bojesen Christensen &
Per Bruun Brockhoff

DTU Informatics
Section for Statistics
Technical University of Denmark
rhbc@imm.dtu.dk

August 13 2008

This Talk is About

- A non-standard type of mixed models
- Applicable to a range of discrimination tests
- Focus: Insight into models—not computational methods
- Motivated by examples from sensometrics and psychometrics (but also applicable in signal detection, medical decision making etc.)
- The models extend existing models by
  - Modelling the covariance structure
  - Having a close connection to psychological theory of cognitive decision making
  - Providing inference for individuals via random effect estimates

Outline

A Replicated Discrimination Test

Example:

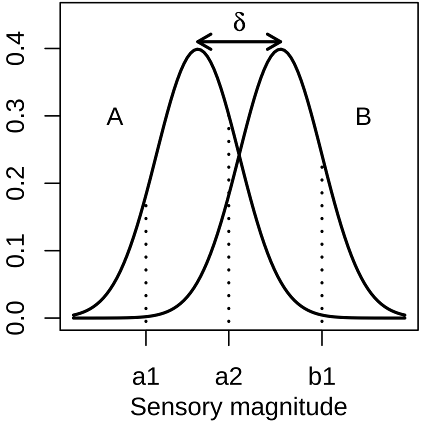
- Coke Inc. wants to substitute a sweetener in a diet coke, A with a cheaper alternative B.
- Coke Inc. employs 30 consumers in a discrimination (triangle) test
- Each consumer performs the test 10 times (replications)
- Can consumers distinguish between the two recipes?

Why Replications?

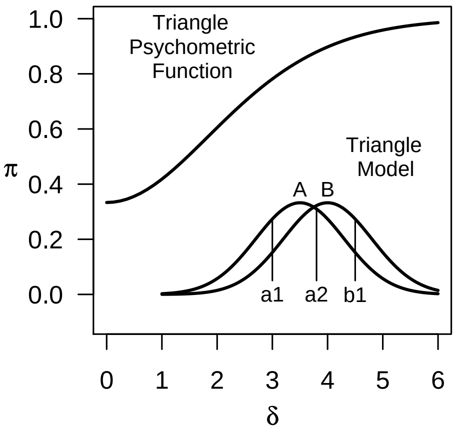
- Advantages:
  - Cheap and Easy: Substitute some assessors with replications
  - Information on difference between assessors
- Challenges:
  - Observations are often correlated and *not* independent

The Triangle Test

- Two regular products and one new product are presented to the consumer
- Two A- products $(a1, a2)$ and one B-product $(b1)$ are presented to the consumer
- Task: Identify the odd product
- $\delta = \mu_B - \mu_A$: A measure of discriminal ability and difference between products.



The Triangle Test



- Answers are binomial: A proportion of correct answers $Y_i \sim \text{Bin}(\pi_i; n_i)$ $\pi_0 \leq \pi_i < 1$
- Guessing probability: $\pi_0 = 1/3$
- Relation between π_i and δ_i:

$$\pi_i = f(\delta_i) = \int_0^\infty \left\{ \Phi\left(-z\sqrt{3} + \delta\sqrt{2/3}\right) + \Phi\left(-z\sqrt{3} - \delta\sqrt{2/3}\right) \right\} \phi(z) \text{d}z$$

- Psychometric function: Relates the probability of a correct answer to the ability to discriminate

The Basic (Naive) Model (GLM)

$$Y_i \sim \text{Bin}(\pi_i, n_i) \quad \pi_i = f_{\text{triangle}}(\delta_i) \quad \delta_i = \delta$$

- A Generalized Linear Model (GLM) with
  - Binomial distribution
  - Psychometric function as inverse link function
  - Simple linear predictor
- Assumes π and δ identical for all individuals
- Family-object; `triangle` in package sensR for use with `glm`
  - Extends discrimination tests to allow for explanatory variables
  - Prepares the way for mixed effect models for discrimination tests

Models for Replicated Discrimination Tests

- Ignore covariance structure in data
  - Basic GLM
- Marginal Models (adjust se's for overdispersion)
  - quasi-binomial GLM

- Latent Class Mixture model
- Conditional Models (model covariance structure)
  - Generalized Linear Mixed Model (GLMM)
  - Synthesis of Mixture and GLMM

Models for Replicated Discrimination Tests

- Ignore covariance structure in data
  - Basic GLM
- Marginal Models (adjust se's for overdispersion)
  - quasi-binomial GLM

- **Latent Class Mixture model**
- **Conditional Models (model covariance structure)**
  - **Generalized Linear Mixed Model (GLMM)**
  - **Synthesis of Mixture and GLMM**

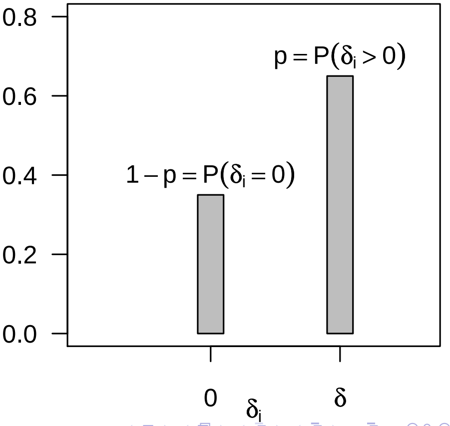
Latent Class Mixture Models

Two-class model for discriminal ability

$$P_i \sim \text{Bernoulli}(p) \quad Y_i|p_i \sim \text{Bin}(\pi_i; n_i)$$

$$\pi_i = f_{\text{triangle}}(\delta_i) \quad \delta_i = \begin{cases} 0 & \text{if} \quad p_i = 0 \\ \delta & \text{if} \quad p_i = 1 \end{cases}$$

- No dispersion among subjects with positive δ
- Often an inappropriate assumption!

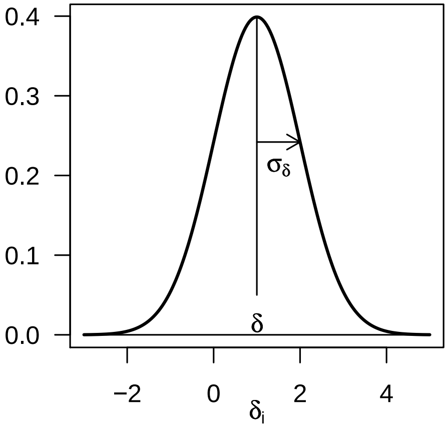


Generalized Linear Mixed Model

$$b_i \sim N(0, \sigma_\delta^2) \quad Y_i | b_i \sim \text{Bin}(\pi_i; n_i)$$

$$\pi_i = f_{\text{triangle}}(\delta_i) \quad \delta_i = \delta + b_i$$

- Assumes a continuous distribution for subjects
- Allows for dispersion among subjects
- Subjects can have $\delta < 0$!
- Impossible in the triangle test

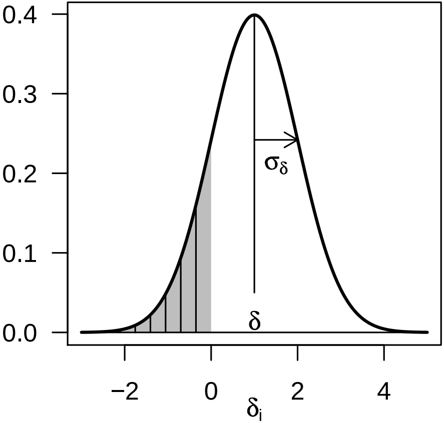


Generalized Linear Mixed Model

$$b_i \sim N(0, \sigma_\delta^2) \quad Y_i | b_i \sim \text{Bin}(\pi_i; n_i)$$

$$\pi_i = f_{\text{triangle}}(\delta_i) \quad \delta_i = \delta + b_i$$

- Assumes a continuous distribution for subjects
- Allows for dispersion among subjects
- Subjects can have $\delta < 0$!
- Impossible in the triangle test



Latent Class Mixed Model

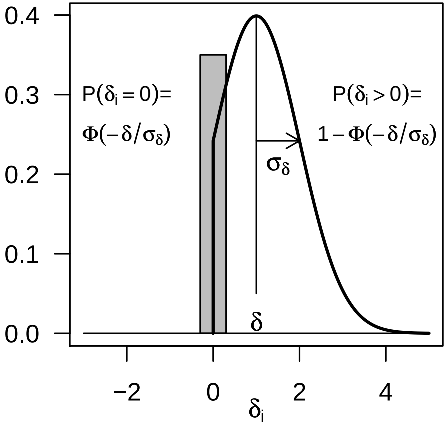
$$Y_i | b_i \sim \text{Bin}(\pi_i; n_i) \quad \delta_i \sim F(\delta, \sigma_\delta^2)$$

$$\pi_i = f_{\text{triangle}}(\delta_i) \quad \delta_i = \delta + b_i$$

$$f(\delta_i) = \begin{cases} 1 - p, & \delta_i = 0 \\ \frac{1}{\sigma_\delta}\phi\left(\frac{\delta_i - \delta}{\sigma_\delta}\right), & \delta_i > 0 \end{cases}$$

$$p = 1 - \Phi(-\delta/\sigma_\delta)$$

- One-dimensional random effect with two attributes:
  - Class probabilities $\tilde{p}_i$
  - The magnitude of discriminal ability $\tilde{\delta}_i$



Estimation in Latent Class Mixed Model

Likelihood function $\sim$ marginal density of y

$$f(y_i) = (1-p)f_1(y_i) + pf_2(y_i)$$

Likelihood at $\delta_i = 0$: $f_1(y_i) = \binom{n_i}{y_i}\pi_0^{y_i}(1-\pi_0)^{(n_i - y_i)}$

Likelihood at $\delta_i > 0$: $f_2(y_i) = \frac{1}{p}\int_0^\infty f_\pi(y_i|\delta_i)\phi((\delta_i - \delta)/\sigma_\delta)\ \sigma_\delta\ d\delta_i$

where $f_\pi(y_i|\delta_i) = \binom{n_i}{y_i}\pi_i^{y_i}(1-\pi_i)^{(n_i - y_i)}$

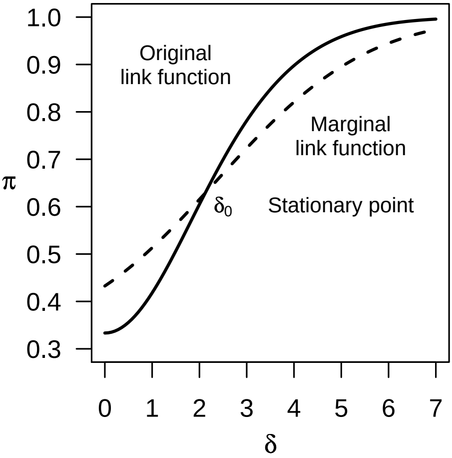
- Define likelihood function as R-function via `integrate`
- optimize with `optim`
- Structure motives an EM algorithm

Attenuation Effect

Marginal link function:

$$\begin{aligned}\pi_m &= E_{\delta_i}[E[y_i|\delta_i]] \\ &= \pi_0(1-p) \\ &+ \int_0^{\infty} f_{\text{triangle}}(\delta_i)\phi((\delta_i-\delta)/\sigma_\delta)/\sigma_\delta \; d\delta_i\end{aligned}$$

- Marginal estimates are closer to "stationary points" rather than closer to zero
- Marginal link function depends on σ_δ^2



Example Triangle Data

Model	δ	se(δ)	σ_δ	p
Basic (GLM)	1.67	0.186		
Overdisp. GLM	1.67	0.257		
Proposed Model	1.62	0.234	1.08	93.3%

- Difference in estimate of δ (attenuation effect)
- Basic model gives too small se's
- Clear variation between subjects
- Large proportion of discriminators

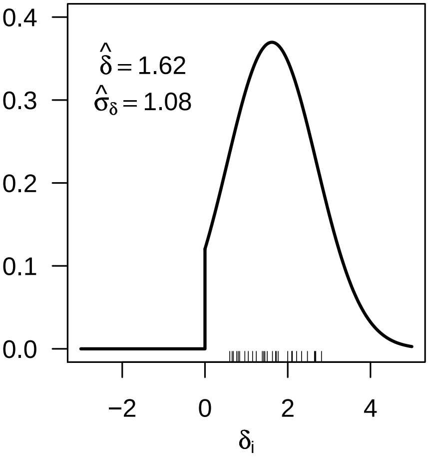
Example Triangle Data

Random effect estimates as Conditional expectations:

$$\tilde{\delta}_i = E[\delta_i | y_i] = \int_{-\infty}^{\infty} \delta_i f(\delta_i | y_i) \mathrm{d}\delta_i$$

$$= \frac{\int_{-\infty}^{\infty} \delta_i f(y_i|\delta_i) f(\delta_i)\, d\delta_i}{f(y_i)}$$

- One extra integral per individual



Example Triangle Data

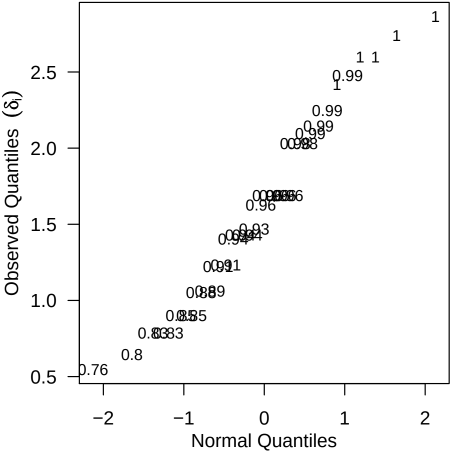
Class probabilities as Conditional expectations:

$$\tilde{p}_i = E[p_i|y_i] = \int_{-\infty}^{\infty} P_i \mathrm{d}F(P_i|y_i)$$

$$= \sum_{P_i \in (0,1)} P_i f(y_i|P_i) f(P_i)/f(y_i)$$

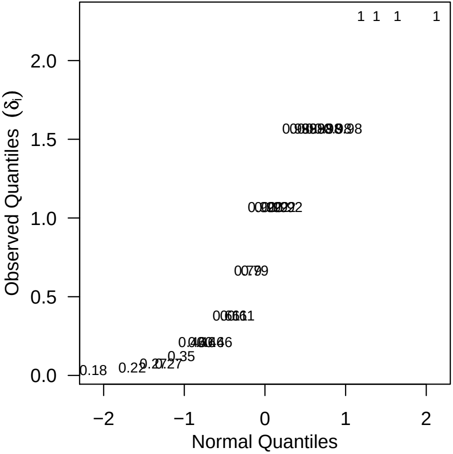
$$= \frac{p f_2(y_i)}{(1-p)f_1(y_i) + p f_2(y_i)}$$

- Random effects are almost normal
- Expect similar results from GLMM



Example: 2-AFC Test

- Random effect estimates: $\tilde{\delta}_i$, $\tilde{p}_i$
- Clear tail on the left
- Clear lower bound on δ
- Discrete nature of data more clear



Summary and Challenges

Summary

- Close connection to psychological theory
- Model-type apply to a range of discrimination test protocols (eg. triangle, duo-trio, 2-AFC and 3-AFC)
- A synthesis of Latent Class Mixture Models and GLMMs
- One-dimensional random effects with two attributes
  - Class probabilities $\tilde{p}_i$
  - The magnitude of discriminal ability $\tilde{\delta}_i$

Challenges

- Extension to additional
  - fixed effects (changes area of integration)
  - random effects (multi-dimensional integrals)
- Variance of random effect estimates
- Implementation with better computational methods

Acknowledgments

Thanks to the Program Committee

Thanks to Professor Per Bruun Brockhoff

Thank you for listening!