Setting a Target Test Information Function for Assembly of IRT-Based Classification Tests (2)

Kentaro Kato†

Benesse Educational Research and Development Center, Tokyo, Japan

Introduction

- Assembly of IRT-based tests requires a target test information function (TIF)
- Specification of the target TIF takes into consideration a desired level of estimation accuracy, overall characteristics of the item pool, empirical validation (simulation or trials and errors), etc.
- It would be useful if there is a systematic method to obtain a target TIF
- Focus on classification tests

◊ A refinement of the previous study presented at IMPS 2011
  - The previous study assumed a specific functional form for the risk function. It ended up with bimodal TIFs which were not very useful.
  - The current study aims at obtaining a smoother, unimodal target TIF for more practical use

Decision theoretical formulation of pass/fail classification based on IRT

- Setting (Fig. 1)
  - Use the MLE $\hat{\theta}$ to make a pass/fail decision
  - The threshold is set to 0 and the population (or prior) distribution of θ is assumed to be $N(\mu, 1)$.

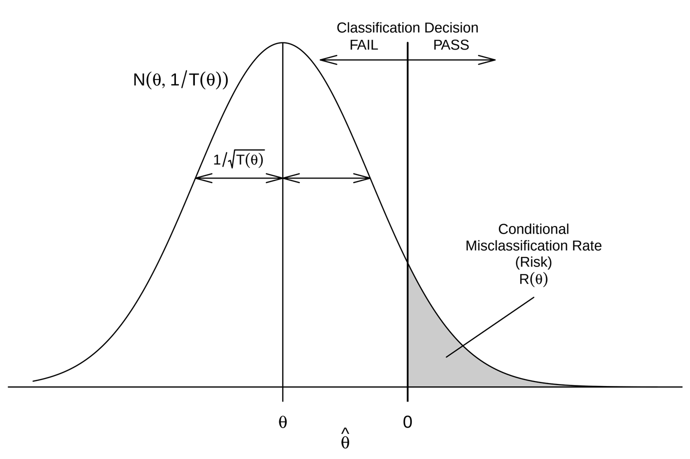

Figure 1: Pass-fail classification

- Target TIF: $T(\theta)$
- The asymptotic sampling distribution of MLE $\hat{\theta}$ is

$$\hat{\theta}|\theta \,\dot{\sim}\, N(\theta, 1/T(\theta)) \tag{1}$$

- Decision rule: $d : \Theta \rightarrow \{\text{Pass}, \text{Fail}\}$

$$d(\hat{\theta}) = \begin{cases} \text{Pass}, & \hat{\theta} > 0 \\ \text{Fail}, & \hat{\theta} \leq 0 \end{cases} \tag{2}$$

- Loss function: $L(\theta, d(\hat{\theta}))$

	Decision	
	Fail	Pass
True State	$\hat{\theta} > 0$	$\hat{\theta} \leq 0$
$\theta > 0$	1	0
$\theta \leq 0$	0	1

- Risk function: $R(\theta)$

$$\begin{aligned} R(\theta) &= E_{\hat{\theta}|\theta}[L(\theta, d(\hat{\theta}))] && (3) \\ &\approx \int_{\Theta} L(\theta, d(\hat{\theta}))\phi(\hat{\theta}|\theta, 1/T(\theta))\, d\hat{\theta} && (4) \\ &= \begin{cases} \Phi(-\theta\sqrt{T(\theta)}), & \theta > 0 \\ 1 - \Phi(-\theta\sqrt{T(\theta)}), & \theta \leq 0 \end{cases} && (5) \end{aligned}$$

where $\phi(\hat{\theta}|\theta, 1/T(\theta))$ is the pdf of $N(\theta, 1/T(\theta))$, and $\Phi(\cdot)$ is the standard normal cdf. $R(\theta)$ is the misclassification rate given θ.

- Bayes (preposterior) risk: $r(\mu)$

$$r(\mu) = E_{\theta}[R(\theta)] = \int_{\Theta} R(\theta)\phi(\theta|\mu, 1)\, d\theta \tag{6}$$

where $\phi(\theta|\mu, 1)$ is the pdf of $N(\mu, 1)$. $r(\mu)$ is the overall misclassification rate given μ.

Problem

- Specifying $T(\theta)$ which ensures the overall misclassification rate $r(\mu)$ being less than a certain value α, given the population mean μ

Method

- Assume the following functional form for the target TIF:

$$T(\theta) = \frac{z_p^2/s^2}{\theta^2/s^2 + 1} = \frac{z_p^2}{\theta^2 + s^2} \tag{7}$$

where $s > 0$ is the "scale" parameter, and $z_p = \Phi^{-1}(p)$, with $p = \lim_{\theta \to \infty} R(\theta)$, is a prespecified "height adjuster"

- The risk function (4) is then reexpressed as

$$R(\theta) = \Phi\left(-\sqrt{\frac{z_p^2\theta^2}{\theta^2 + s^2}}\right) \tag{8}$$

- Compute the Bayes risk by discrete approximation to the integral (6):

$$r(\mu) \approx f(s) = \sum_q \Phi\left(-\sqrt{\frac{z_p^2\theta_q^2}{\theta_q^2 + s^2}}\right) w_q \tag{9}$$

where θ_q, $q = 1, \ldots, Q$, represent appropriate quadrature points and w_qs are the corresponding weights which approximate $N(\mu, 1)$

- $\Phi(\cdot)$ in each summand in (9) is monotone increasing with respect to s, so given μ and $\alpha > p$ we can find a value of s such that $f(s) = \alpha$ (the overall misclassification rate)
  - Numerical computation (Newton-Raphson method) is required to find the solution of s
  - The first derivative of f necessary for computation is readily available:

$$f'(s) = \sum_q \frac{w_q s\sqrt{z_q^2\theta_q^2}}{\sqrt{2\pi}(\theta_q^2 + s^2)^{3/2}} \exp\left(-\frac{z_p^2\theta_q^2}{2(\theta_q^2 + s^2)}\right) \tag{10}$$

  - Currently p is set to $10^{-6} \iff z_p \approx -4.75$
- Values of s were computed for several combinations of μ and α (Fig. 2)

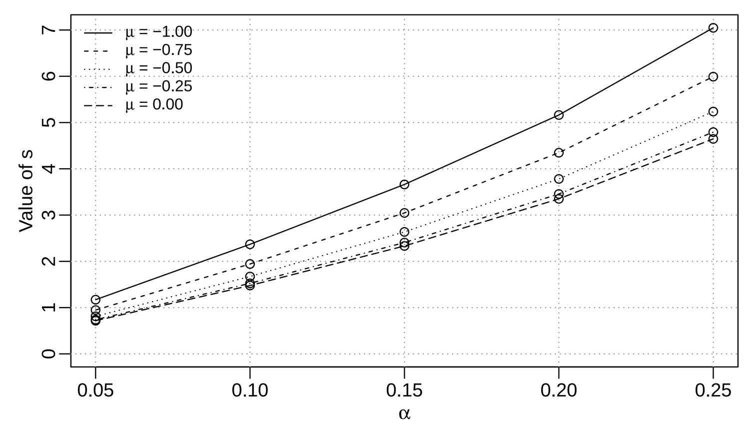

Figure 2: Optimal values of s

Results

- Target TIFs (Fig. 3)
  - The shapes of obtained TIFs are sufficiently "mild" so that they could serve as a good "reference"
  - More information is required as μ approaches to 0 (i.e., the population mean approaches to the threshold) and α becomes small (i.e., less overall misclassifications)
- Conditional misclassification rates (Fig. 4)
  - The misclassification rate distribution becomes tighter as μ approaches to 0 (i.e., the population mean approaches to the threshold)

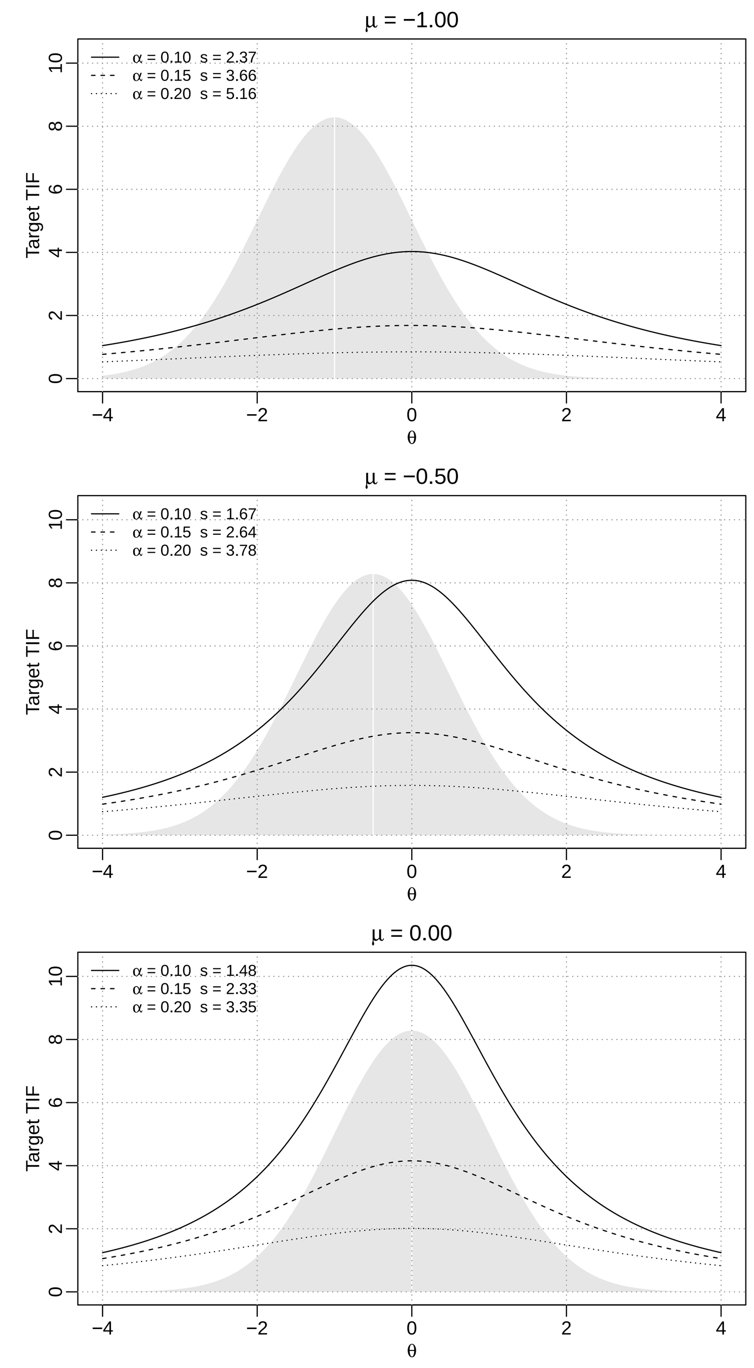

Figure 3: Target TIFs: $T(\theta)$

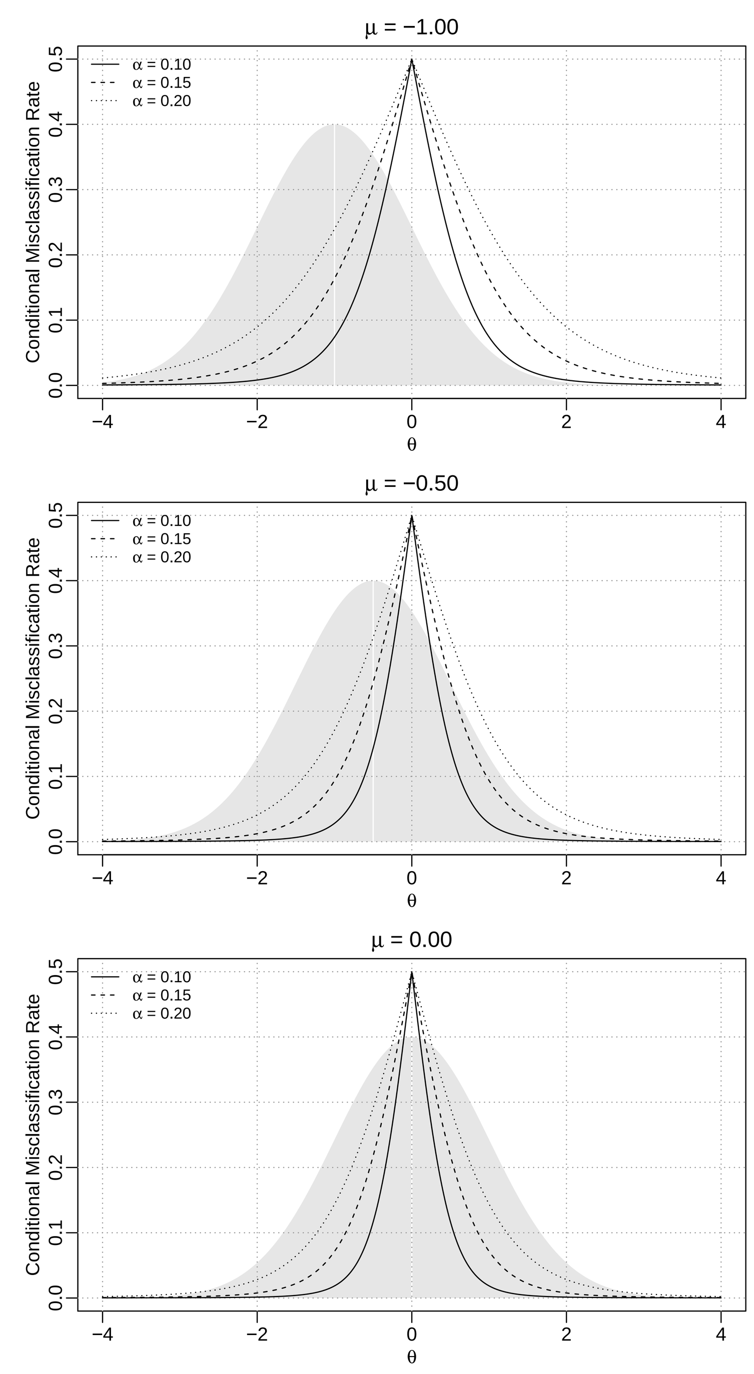

Figure 4: Conditional misclassification rates (risk functions): $R(\theta)$

Conclusions and further considerations

- Systematically obtain target TIFs which take into consideration the overall misclassification rate and the location of the population distribution of θ
  - This should not be taken as a standard, but would serve as a good starting point when one has to start from a scratch
- Limitations
  - Optimized value of s is sensitive to the value of z_p, so it should be specified with care
  - The symmetric form of $T(\theta)$ is not very efficient when μ departs from the threshold
- Possible extensions
  - Asymmetric loss functions?
  - How to approach multi-stage classifications?

† Benesse Educational Research and Development Center, 1-34 Ochiai, Tama, Tokyo, Japan 206-8686. mailto: nekotak@mail.benesse.co.jp