Supplement to "From ROC Curves to Psychological Theory"

July 13, 2013

1 Review of Density, Cumulative, and Quantile Functions.

Distributions have several different types of associated functions. The most familiar is the *density function*. Figure 1A and 1D show density functions for the normal distribution and gamma distribution, respectively. Density functions are related to histograms in that if enough data are collected then the histogram and density share the same shape. Hence, density functions serve as a first-order graphical representation of a distribution. An alternative view of a distribution is provided by the *cumulative distribution function* (CDF). CDFs describe the probability of an observation being below a specified value, and Figure 1B and 1E provides the CDFs for the normal and gamma distribution, respectively. The highlighted point in Figure 1B corresponds to a latent strength of 1.0 and a cumulative probability of .84, indicating that 84% of the mass of distribution is less than 1.0. Quantile functions are the inverse of the CDF. They describe the value (latent strength values in this case) associated with a cumulative probability, and the corresponding examples are shown in Figure 1C and 1F. The highlighted point in Figure 1C shows that the cumulative probability of .84 corresponds to a latent strength of 1.0. These three graphical representations, densities, CDFs, and quantile functions are alternative graphical representations of a distribution and may be used interchangeably.

2 Proof of Constraint on ROCs from the Discrete-State Model

The discrete state model is given by

$$h = da + (1-d)g, \tag{1}$$
$$f = db + (1-d)g. \tag{2}$$

Let A B and C denote ROC points (g, g), $(f, h)$and (b, a), respectively. The discrete-state representability condition is that point B lies on the line segment

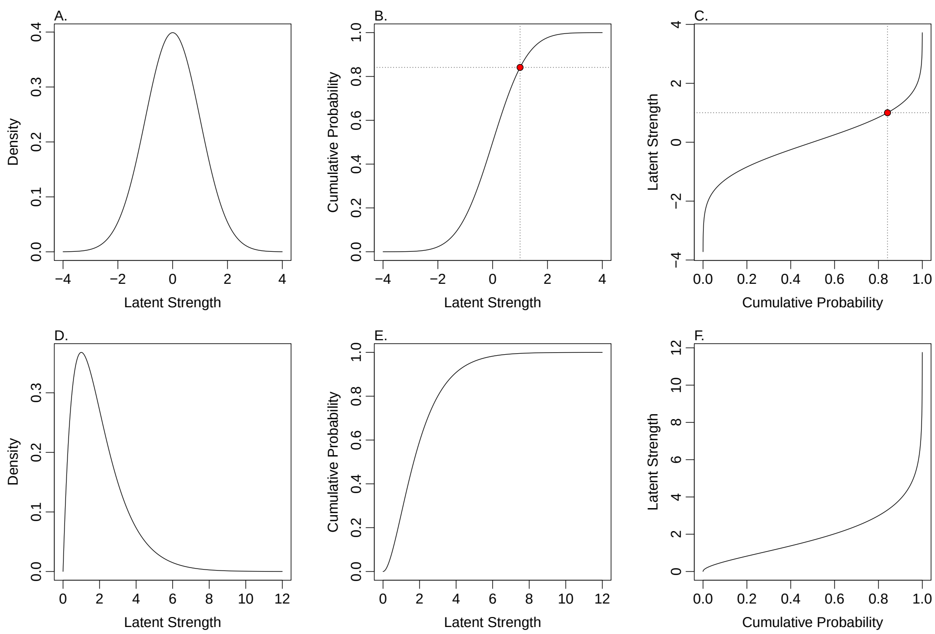

Figure 1: **A-C.** Density function, CDF, and quantile function, respectively, for a normal distribution with a mean of 0 and a variance of 1. **D-F.** Density function, CDF, and quantile function, respectively, for a gamma distribution with a shape of 2 and a scale of 1.

$\vec{AC}$, and that the distance from A to B, denoted $||AB||$ is $d||AC||$. We show that the discrete-state model implies discrete-state representability.

The first step is to show that point B lies on $\vec{AC}$. It is straightforward to show that the segment $\vec{AC}$ lies on the line

$$y = \frac{a-g}{b-g}x + \frac{g(b-a)}{b-g}$$

To show that B lies on this line, we prove that $y = h$ when $x = f$:

$$\begin{aligned} y &= \frac{a-g}{b-g}f + \frac{g(b-a)}{b-g} \\ &= \left(\frac{a-g}{b-g}\right)(db + (1-d)g) + \frac{g(b-a)}{b-g} \\ &= \frac{a(db + (1-d)g - g) + g(-db-)1-d)g + b)}{b-g} \\ &= \frac{ad(b-g) + g(1-d)(b-g)}{b-g} \\ &= da + (1-d)g, \\ &= h. \end{aligned}$$

The final step is to show that $||AB|| = d||AC||$. Note that

$$\begin{aligned} ||AB||^2 &= (h-g)^2 + (f-g)^2 \\ &= (da - (1-d)g - g)^2 + (db - (1-d)g - g)^2 \\ &= (da - dg)^2 + (db - dg)^2 \\ &= d^2((a-g)^2 + (b-g)^2). \end{aligned} \tag{3}$$

Also note that

$$||AC||^2 = (a-g)^2 + (b-g)^2.$$

Substituting this equation into (3) yields

$$||AB||^2 = d^2||AC||^2.$$

Given that $||AB||$, $||AC||$, and d are positive by definition,

$$||AB|| = d||AC||.$$

3 Proof that Mixture Constraint on Confidence Ratings is Equivalent to Discrete-State Representability

We show in this section that mixture constraint of Figure 3 in the main paper implies and is implied by discrete-state representability.

In a confidence-ratings task, the participant endorses one of I possible response options. The options are ordered such that the 1st response is the highest confidence old-item response and the Ith response is the highest confidence new-item response. Let p_{1ij} denote the probability that the participant endorses the ith option in the jth old-item condition; let p_{2ij} denote the same for the jth new-item condition; and let p_{gi}, p_{ai}, and p_{bi} denote the same when the participant is guessing, in the detect state for old items, and in the detect state for new items, respectively. The Province and Rouder (2012) mixture constraint is there exists mixture probabilities d_j for the jth condition such that

$$\begin{aligned} p_{1ij} &= d_j p_{ai} + (1-d_j)p_{gi}, \\ p_{2ii} &= d_j p_{bi} + (1-d_j)p_{go}. \end{aligned}$$

For confidence-rating tasks, hit and false alarm rates are defined as cumulative sums:

$$\begin{aligned} h_{ij} &= \sum_k^i p_{1kj}, \\ f_{ij} &= \sum_k^i p_{2kj}. \end{aligned}$$

If we define the quantities $a'_i = \sum_k^i p_{ak}$, $b'_i = \sum_k^i p_{bk}$, and $g'_i = \sum_k^i p_{gk}$, then the mixture constraint may be equivalently written as

$$\begin{aligned} h_{ij} &= d_j a'_i + (1-d_j)g'_i, \\ f_{ij} &= d_j b'_i + (1-d_j)g'_i. \end{aligned}$$

We first show this mixture constraint implies a discrete-state representation. Let A_i and C_i be the points (b'_i, a'_i) and (g'_i, g'_i), respectively, which are formed respectively by setting $d = 1$ and $d = 0$. The proof of discrete-state representability follows from the above proof with a'_i, g'_i, and b'_i serving an analogous role of a, g, and b, respectively.

The proof of the converse, that discrete-state representation implies the mixture constraint is as follows. Let $B_{ij} = (f_{ij}, h_{ij})$ and assume that B_{ij} lies on the line $\vec{AC}$ and that $||A_i B_{ij}|| = d_j ||A_i C_i||$. For the false-alarm rates (x-axis), the following holds:

$$\begin{aligned} f_{ij} &= g'_i + d_j(b_i - g'_i), \\ &= d_j b_i + (1-d_j)g'_i. \end{aligned}$$

For the hit rates (y-axis), the following holds:

$$\begin{aligned} h_{ij} &= g'_i + d_j(a_i - g'_i), \\ &= d_j a_i + (1-d_j)g'_i. \end{aligned}$$

Thus, the mixture constraint holds, and the mixture constraint and discrete-state representability are equivalent properties.

4 Proof that the Slope of ROCs is the Ratio of Signal-Detection Densities

Let h and f denote hit and false-alarm rates, respectively, and let $h(f)$ be an isosensitivity curve. Let g_s and g_n denote latent strength densities for signal and noise stimuli, respectively, in a signal detection model. The hit and false alarm rates for criteria z are given by

$$\begin{aligned} h(z) &= 1 - G_s(z), && (4) \\ f(z) &= 1 - G_n(z), && (5) \end{aligned}$$

where G_s and G_n are the CDFs for the latent strengths of signal and noise stimuli, respectively, and z is the value of the criterion. It follows immediately that

$$z(f) = G_n^{-1}(1 - f), \quad (6)$$

where G_n^{-1} is the quantile function for the latent strength of noise stimuli. Substituting (6) into (4) yields,

$$h(f) = 1 - G_s[G_n^{-1}(1 - f)]. \quad (7)$$

The slope of this function is given by the derivate dh/df:

$$\begin{aligned} \frac{dh}{df} &= -g_s[G_n^{-1}(1-f)] \times \frac{d}{df} G_n^{-1}(1-f) \\ &= -g_s[G_n^{-1}(1-f)] \frac{1}{g_n[G_n^{-1}(1-f)]} \times \frac{d}{df}(1-f) \\ &= \frac{g_s[G_n^{-1}(1-f)]}{g_n[G_n^{-1}(1-f)].} \end{aligned}$$

Substituting $z(f) = G_n^{-1}(1 - f)$ yields:

$$\frac{dh}{df} = \frac{g_s[z(f)]}{g_n[z(f)]}.$$

Note that if the latent strength of noise stimuli is uniformly distributed on $(0, 1)$, then $g_s(1 - f) = dh/df$.

5 The Relationship between Shift-Representability of ROC Curves and Signal-Detection Models

A set of signal detection models is shift representable if $g_s(z - \mu) = g_n(z)$, where g_s and g_n are density functions of strength for noise and signal stimuli, and μ is the degree of shift of the strength distribution. Equivalently, shift representability implies that $G_s(z-\mu) = G_n(z)$, where G_s and G_n are the corresponding CDFs. Hit and false alarm rates are given by

$$\begin{aligned} h &= 1 - G_n(z-\mu), \\ f &= 1 - G_n(z), \end{aligned}$$

and

$$h = 1 - G_n[G_n^{-1}(1-f) - \mu].$$

Rearranging implies that

$$1 - h = G_n[G_n^{-1}(1-f) - \mu].$$

Applying the transform $-G_n^{-1}$ to both sides yields

$$\begin{aligned} -G_n^{-1}(1-h) &= -G_n^{-1}\left(G_n\left[G_n^{-1}(1-f) - \mu\right]\right), \\ &= -G_n^{-1}(1-f) + \mu. \end{aligned}$$

Let $\theta(p) = -G_n^{-1}(1-p)$, be a transform of hit and false alarm rates. The above equation may be written as

$$\theta(h) = \theta(f) + \mu.$$

Therefore, transform $\theta(p)$ linearizes the isosensitivity curve into a space with a common slope of 1.0 and an intercept of μ.

6 Assessing Shift Representability

Shift representability provides a needed tool for assessing the complexity of processing. If a family of ROC curves are shift representable, then a simple mechanism is appropriate to describe the data; conversely, if ROCs are not shift representable then more complex mechanisms are indicated. This inferential logic raises the important question of how a researcher may know if a family of ROC curves are shift representable. The crux of the matter is establishing that there does or does not exist a linearizing transform, and this is a difficult problem, the solution to which is necessarily technical.

The current approach we use is to estimate separate linearizing functions for each ROC curve in a family. For example, if there are three repetition conditions in an experiment, we would estimate three separate linearizing functions with one for each condition. The critical question is whether these linearizing functions are the same linearizing function for a single ROC curve, and we do so by approximation. Let θ be the unknown function that linearizes a curve in a family. An estimation approach is to find a function ϕ dependent on a vector of parameters $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots)$ that reasonably approximate θ, e.g.,

$$\theta(p) \approx \phi(p, \boldsymbol{\beta}), \text{for all } p$$

where p is the relevant hit or false alarm rate. Researchers then need to pick a useful function ϕ with easy-to-estimate parameters $\boldsymbol{\beta}$. The choice of ϕ may be made by studying the domain and range of θ. Figure 8D-F shows some examples of θ, and this function has inputs between 0 and 1 and outputs that range across the entire number line, and these domains and ranges are those of quantile functions. Our approach is as follows:

$$\phi = -F_0^{-1}(1 - u(p, \boldsymbol{\beta})),$$

where F_0 is a cumulative distribution function and should be reasonably close to the CDF of the new-item distribution in the shift-representation should shift representability hold. It is not necessary that $-F_0^{-1}(1-p)$ exactly linearizes a curve, but the closer it is to θ, the more power there will be to test for shift representability. A suitable choice of F_0 for yes-no recognition is the cumulative distribution function of the log-gamma, denoted Λ, as this choice linearizes ROC curves that are about as asymmetric as observed ROC curves in the literature. A suitable choice for 2AFC paradigms is the CDF of the normal, denoted Φ,because this choice captures the symmetry of observed ROC curves for this paradigm. The critical function is $u(p, \boldsymbol{\beta})$. We use a two-parameter spline with 1/3 and 2/3 to model u:

$$u(p) = \left(1 - \frac{2}{3}\beta_1 - \frac{1}{3}\beta_2\right) p + \beta_1 \left(p - \frac{1}{3}\right) I(p - 1/3) + \beta_2 \left(p - \frac{2}{3}\right) I(p - 2/3).$$

This function is composed of three line segments that connect from $(0,0)$ to $(1/3, \beta_1)$ to $(2/3, \beta_2)$ to $(1,1)$. To preserve the needed monotonicity of g, it must be the case that $\beta_2 \geq \beta_1$. It is possible to estimate three parameters, μ_i, β_{1i}, and β_{2i} for the ith ROC curve by minimizing errors between observed and predicted hit and false-alarm rates. Shift representability holds if the β parameters do not vary across curves; i.e., $\beta_{11} = \beta_{12} = \ldots = \beta_{1I}$ and $\beta_{21} = \beta_{22} = \beta_{2I}$.