Supplemental Material

The Monotonicity of Second Choices

Given a signal distribution parametrized by a parameter μ for signal strength with density f_μ and cumulative distribution function F_μ, the probabilities of assigning rank 1 and rank 2 to the signal are given by $\pi_1(\mu) = \int F^{k-1} f_\mu$ and $\pi_2(\mu) = (k-1) \int F^{k-2}(1-F) f_\mu$, respectively, where F is the cumulative distribution function of the noise distribution with density f. Note the omission of subscript k. Let the signal distribution's support be an interval with lower bound l_μ that is non-decreasing in μ and upper bound u with $f_\mu(y) > 0$ for $y \in (l_\mu, u)$, and assume that the support for the noise distribution is the interval (l, u) with $l \leq \inf_\mu l_\mu$ and $f(y) > 0$ for $y \in (l, u)$, where $-\infty$ and ∞ are admissible values for the lower and upper bounds, respectively.

Assume furthermore that l_μ is differentiable for μ with $l_\mu > -\infty$, that $F_\mu(z)$ is differentiable in μ for every $z \in (l_\mu, u)$ and differentiable from the right for $z = l_\mu$. Assume that for every μ there is an open interval I_μ containing μ and a function g_μ so that $|\frac{\partial}{\partial \mu} F_\mu(z)|_{\mu=\nu}| \leq g_\mu(z)$ for all $\nu \in I_\mu$ and for all $z \in [l_\mu, u)$ and so that $\int_{l_I}^{u} g_\mu(z) f(z) \mathrm{d}z < \infty$, where $l_I = \inf I_\mu$. This is, for example, satisfied if $\frac{\partial}{\partial \mu} F_\mu(z)$ considered as a function in (μ, z) is bounded on the set of (ν, z) with $\nu \in I_\mu$ and $z \in [l_\nu, u)$. The condition states that the functions $z \to \frac{\partial}{\partial \mu} F_\mu(z) f(z)$ are dominated by an integrable function $g_\mu(z) f(z)$ locally in an interval around each μ.

Let $c_2(\mu)$ be the conditional probability of assigning rank 2 to the signal, given that rank 1 was not assigned to the signal: $c_2(\mu) = \frac{\pi_2(\mu)}{1-\pi_1(\mu)}$. We need to consider the function $H_\mu(z)$ defined by

$$\frac{\frac{\partial}{\partial \mu} F_\mu(z)}{F_\mu(z)}.$$

The following theorem will be proved:

Theorem. If $H_\mu(z)$ is monotonically increasing in z for all μ and $z \in (l_\mu, u)$, then $c_2(\mu)$ is monotonically increasing in μ.

Proof. All integrals range from lower bound l_μ to upper bound u. We will mostly suppress the bounds. Using integration by parts, we find that

$$\begin{aligned}\pi_1(\mu) &= F_\mu(z)F^{k-1}(z)|_{z=l_\mu}^{z=u} - (k-1)\int F_\mu F^{k-2} f \\ &= 1-(k-1)\int F_\mu F^{k-2} f.\end{aligned}$$

Similarly, it can be shown that

$$\pi_2(\mu) = (k-1)(1-\pi_1(\mu)) - (k-1)(k-2)\int F_\mu F^{k-3} f.$$

It follows that

$$c_2(\mu) = (k-1)-(k-2)\frac{\int F_\mu F^{k-3} f}{\int F_\mu F^{k-2} f}.$$

$c_2(\mu)$ is monotonically increasing, if the fraction on the right side, $t(\mu) = \frac{\int F_\mu F^{k-3} f}{\int F_\mu F^{k-2} f}$, is monotonically decreasing. The latter term is decreasing if its derivative with respect to μ is negative. The differentiations can be performed under the integral sign. This is ensured by the above assumption that $\frac{\partial}{\partial\mu}F_\mu(z)$ is locally dominated by an integrable function and the dominated convergence theorem (Bartle, 1966, chap. 5). Note also that l_μ is differentiable where $l_\mu > -\infty$ and that $F_\mu(l_\mu) = 0$. This yields

$$\frac{\partial}{\partial\mu}t(\mu) = \frac{(\int \frac{\partial}{\partial\mu}F_\mu F^{k-3} f)(\int F_\mu F^{k-2} f) - (\int F_\mu F^{k-3} f)(\int \frac{\partial}{\partial\mu}F_\mu F^{k-2} f)}{(\int F_\mu F^{k-2} f)^2}.$$

Dropping the denominator (it is always positive), converting the product of two integrals into a double integration, making explicit the variables y and z of integration, exchanging the roles of y and z in the second term in the numerator, and rearranging terms yields that $\frac{\partial}{\partial\mu}t(\mu)$ is negative if

$$0 > \int\int (F(z)-F(y))H_\mu(y)F_\mu(z)F_\mu(y)F^{k-3}(y)F^{k-3}(z)f(y)f(z)\mathrm{d}y\,\mathrm{d}z.$$

Let the double integral be $n(\mu)$. $H_\mu(y)$ is assumed to be increasing in y for $y \in (l_\mu, u)$. $F(y)$ is also increasing in y, because $f(y) > 0$ for $y \in (l_\mu, u)$. Hence, $0 > (F(z)-F(y))(H_\mu(y)-H_\mu(z))$ for all y and z with $y \neq z$. Let

$G(y,z) = F_\mu(z)F_\mu(y)F^{k-3}(y)F^{k-3}(z)f(y)f(z)$ and note that $G(y,z) = G(z,y)$. It follows that

$$\begin{aligned} 0 &> \int\int (F(z) - F(y))(H_\mu(y) - H_\mu(z))G(y,z)\mathrm{d}y\,\mathrm{d}z \\ &= \int\int (F(z) - F(y))H_\mu(y)G(y,z)\mathrm{d}y\mathrm{d}z - \int\int (F(z) - F(y))H_\mu(z)G(y,z)\mathrm{d}y\,\mathrm{d}z \\ &= 2n(\mu), \end{aligned}$$

by exchanging the variables of integration y and z in the last double integral.

The function $H_\mu(z)$ is related to the signal distribution's hazard function and can be shown to be increasing for many common distributions. For example, for shift distributions with $f_\mu(y) = f(y-\mu)$, the following corollary relates H_μ to the reverse hazard (Chechile, 2011), given by $r_\mu(y) = \frac{f_\mu(y)}{F_\mu(y)}$:

Corollary. For a family of shift distributions, $H_\mu(y)$ is monotonically increasing in y for all μ, if the reverse hazard is monotonically decreasing in y for all μ under the assumptions of the above theorem.

This follows from the fact that for a shift distribution $H_\mu(y) = -r_\mu(y)$ as is easy to see. Using the results reported by Chechile (2011), it immediately follows that $c_2(\mu)$ is monotonically increasing for signal and noise distributions based on normal distributions (as in the equal-variance SDT and the unequal-variance SDT), but also for the case of exponential distributions ($f_\mu(y) = k\exp(-k(y-\mu))$ with $l_\mu = \mu$), for shift distributions based on ex-Gaussian distributions, Gumbel distributions, mixtures thereof, and many other distributions. Note that the densities of these distributions are locally bounded in an interval around each μ for all y which implies that the assumptions of the theorem on local dominance are satisfied.

The distributions satisfying the conditions of the theorem are not limited to distributions with signal strength conceptualized as a shift parameter. Consider, for example, the Gamma Distribution with shape $\alpha > 0$ and scale $\mu > 0$. Here, $l_\mu = 0$ and $u = \infty$. Its distribution function is

$$F_\mu(z) = \frac{\gamma(\alpha, \frac{z}{\mu})}{\Gamma(\alpha)},$$

where γ is the lower incomplete gamma function. Its derivative with respect to μ is

$$\frac{\partial}{\partial\mu}F_\mu(z) = \frac{1}{\Gamma(\alpha)}\left(-\frac{z}{\mu^2}\right)\left(\frac{z}{\mu}\right)^{\alpha-1} e^{-\frac{z}{\mu}}.$$

Note that this function is bounded in an interval around each μ for all z. Using a series expansion of the lower incomplete gamma function (Abramowitz & Stegun, 1964, p. 262), on the other hand, leads to

$$F_\mu(z) = \left(\frac{z}{\mu}\right)^{\alpha} e^{-\frac{z}{\mu}} \sum_{i=0}^{\infty} \frac{(z/\mu)^i}{\Gamma(\alpha+i+1).}$$

Hence,

$$H_\mu(z) = -\frac{1}{\mu\Gamma(\alpha)} \frac{1}{\sum_{i=0}^{\infty} \frac{(z/\mu)^i}{\Gamma(\alpha+i+1)}},$$

which is monotonically increasing in $z > 0$.

Hierarchical-Bayesian Modeling

In this hierarchical-modelling approach (Rouder & Lu, 2005), response frequencies were directly used instead of estimated conditional probabilities. Individual rank 2 response frequencies (subscripts m and n denoting experiment and participant respectively) for weak and strong items ($d^w_{m,n}$ and $d^s_{m,n}$, respectively) among the total of non-rank 2 responses ($N^w_{m,n}$ and $N^s_{m,n}$) were modeled with a binomial model:

$$\begin{aligned} d^w_{m,n} &\sim \text{Binomial}(\theta^w_{m,n}, N^w_{m,n}), \\ d^s_{m,n} &\sim \text{Binomial}(\theta^s_{m,n}, N^s_{m,n}). \end{aligned}$$

Rate parameters $\theta^w_{m,n}$ and $\theta^s_{m,n}$ were a function of the following terms:

$$\begin{aligned} \theta^w_{m,n} &= \Phi(\phi_{m,n}), \\ \theta^s_{m,n} &= \Phi(\phi_{m,n} + \alpha_{m,n}), \end{aligned}$$

with $\phi_{m,n}$ representing the individual rate parameter (on a probit scale) and $\alpha_{m,n}$ the rate increment for strong items (also on a probit scale). These two parameters are distributed as:

$$\phi_{m,n} \sim \text{Normal}(\mu_{\phi_m}, \sigma^2_{\phi_m}),$$
$$\alpha_{m,n} \sim \text{Normal}(\mu_{\alpha_m}, \sigma^2_{\alpha_m}).$$

Parameter μ_{ϕ_m} denotes the group average rate in Experiment m while parameter $\mu_{\alpha_m} = (\delta + \beta_m) \times \sigma_{\alpha_m}$ denotes the average difference in the rates for weak and old items in Experiment m. Parameter δ corresponds to the overall effect size, β_m to a contrast-coded experiment factor ($\beta = \beta_1 = -\beta_2$), and σ_{α_m} to the variance of the difference between weak and strong items.

The following set of priors was used (Wagenmakers, Lodewyckx, Kuriyal, and Grasman, 2010):

$$\beta \sim \text{Normal}(0, 1),$$
$$\delta \sim \text{Normal}(0, 1),$$
$$\mu_{\phi_m} \sim \text{Normal}(0, 10),$$
$$\sigma_{\alpha_m} \sim \text{Uniform}(0, 10),$$
$$\sigma_{\phi_m} \sim \text{Uniform}(0, 10),$$

The estimation of posterior estimates was made with JAGS (Plummer, 2003) and R (R Core Team, 2013). Posterior samples were obtained from five Markov chain Monte Carlo chains with 250,000 iterations each. From each chain, the first 50,000 samples were removed and only every 20th subsequent sample retained for analysis. Statistic $\hat{R}$ (Gelman, Carlin, Stern, & Rubin, 2004, Chap. 11), which compares within-chain variance to between-chain variance indicated that the samples successfully converged (all $\hat{R} < 1.01$).

The Bayes Factor was calculated using the Savage-Dickey ratio method, which compares the height of the prior and posterior densities of δ at the point $\delta = 0$ (for an introduction, see Wagenmakers et al., 2010; see also Gelfand & Smith, 1990; Morey, Rouder, Pratte, & Speckman, 2011). The height of the prior was rescaled to conform to the order-restricted hypothesis $\delta > 0$.

References

Abramowitz, M. & Stegun, I. A. (1964). *Handbook of mathematical functions.* New York: Dover.

Bartle, R. G. (1966). *The elements of integration and Lebesgue measure.* New York: Wiley.

Chechile, R. A. (2011). Properties of reverse hazard functions. *Journal of Mathematical Psychology, 55*, 203-222. doi:10.1016/j.jmp.2011.03.001

Gelfand, A. & Smith, A. F. M. (1990). Sampling based approaches for calculating marginal densities.*Journal of the American statistical Association, 85,*, 398-409. doi: 10.2307/2289776

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis (2nd ed.).* London: Chapman and Hall.

Morey, R. D., Rouder, J. N., Pratte, M. S., & Speckman, P. L. (2011). Using MCMC chain outputs to efficiently estimate Bayes factors.*Journal of Mathematical Psychology, 55,*, 368-378. doi: 10.1016/j.jmp.2011.06.004

Plummer, M. (2003). *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling.*

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the Theory of Signal Detection. *Psychonomic Bulletin & Review, 12*, 573-604. doi:10.3758/BF03196750

Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: a tutorial on the Savage Dickey method.*Cognitive Psychology, 60*, 158-189. doi:10.1016/j.cogpsych.2009.12.001