Lecture 2: Robust Cognitive Models

Jeff Rouder

Dec, 2021

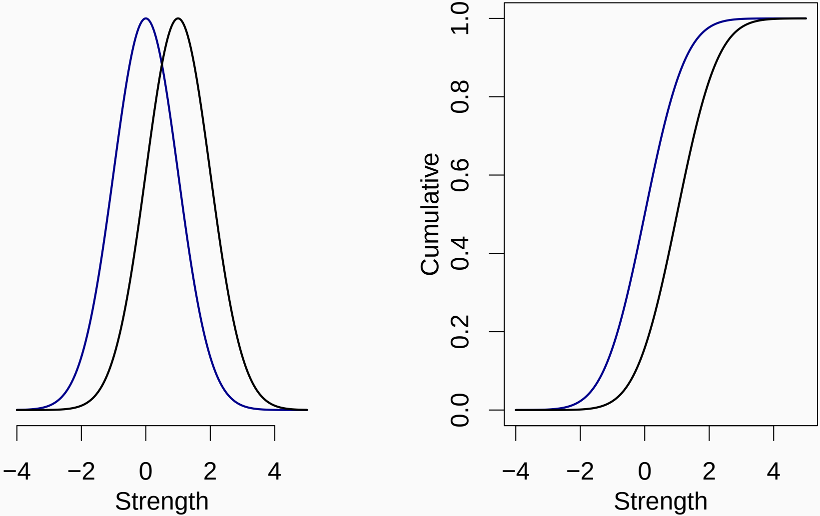
Two Signal-Dection Models

- EVSD (equal variance normal signal detection)
- UVSD (unequal variance normal signal detection)

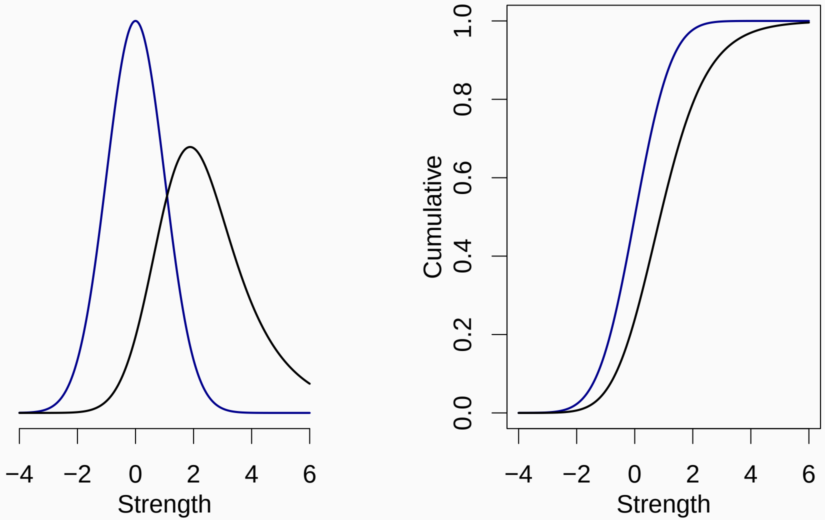
Model And Boundary Conditions

- I think it is helpful to be aware of certain cases or boundary conditions where the predictions and nature of the model become most clear.
- Boundary Condition:
  - Can studying an item reduce its mnemonic strength?
  - Does that make sense?
  - Be wary of models that specify any reduction in strength due to study.
  - Old items at least as strong as new items.

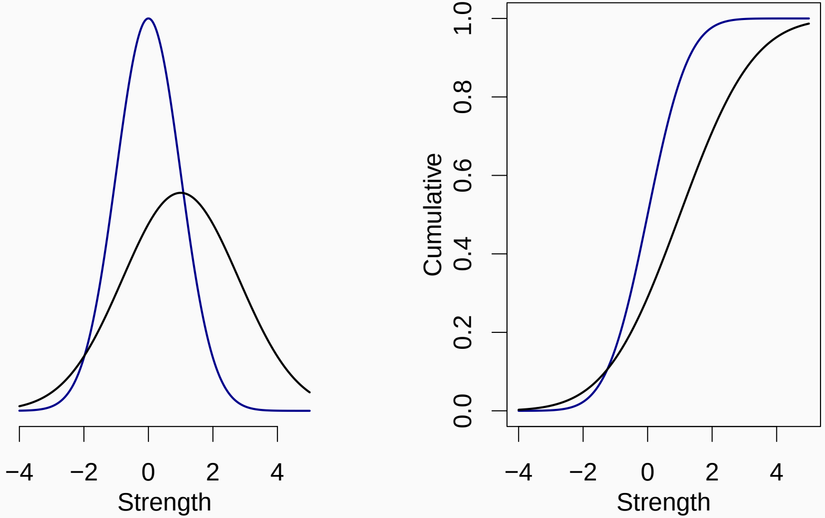
EVSD Meets Boundary so long as $\mu \geq 0$.



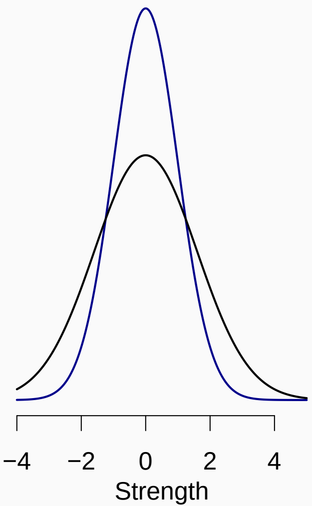
Always Stronger, Expontential Additon Study

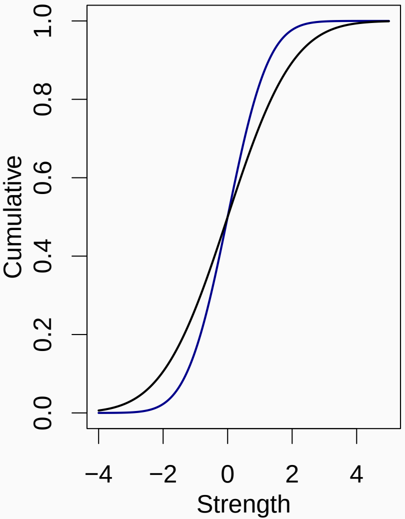


Not Always Stronger, UVSD



Worst Case





Boundary Condition

Stochastic dominance

- Relationship holds across the distribution
- Unambiguously bigger
- No crossings in cumulatives
- ROCs cant cross
- A good model should not force a violation of stochastic dominance
- UVSD does just that

Experiment With Varying Strength Conditions

- Study some words for 1 sec, 2 sec, 5 sec

Your Turn

Formally expand UVSD to include a few strength conditions.

- Certainly μ is affected by strength?
- How about σ^2?
  - constant
  - free
  - tied to μ?

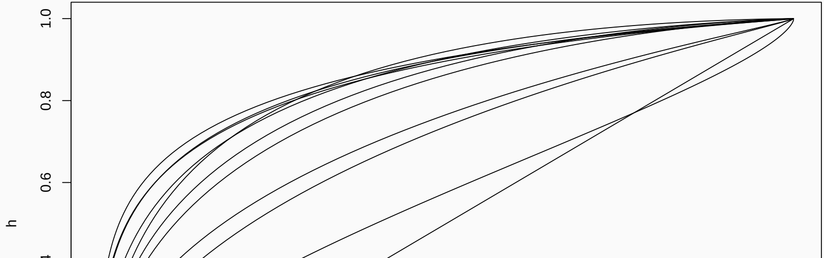
Your Turn

- Program up your multi-condition UVSD model in R
- Draw the resulting ROCs? Do they cross?
    - Always?
    - Sometimes?
    - Never?

My Turn

```r
mu=c(0,runif(10,0,2))
sigma=c(1,runif(10,1,1.5))

crit=seq(-3.5,5,.05)
h=matrix(nrow=length(crit),ncol=length(mu))
f=1-pnorm(crit,0,1)
for (i in 1:length(mu)) h[,i]=1-pnorm(crit,mu[i],sigma[i])
matplot(f,h,typ='l',col='black',lty=1,lwd=1)
```



Your Turn

- Can you envision a gamma signal detection model?
- What might it look like?
- Does it make sense?

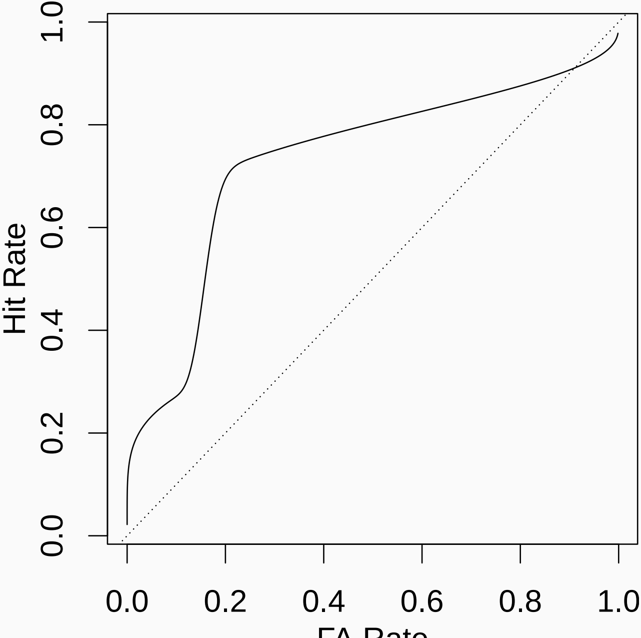
What is the core of signal detection?

- Latent strength is unidimensional
- Responses are governed by distributions relative to criteria

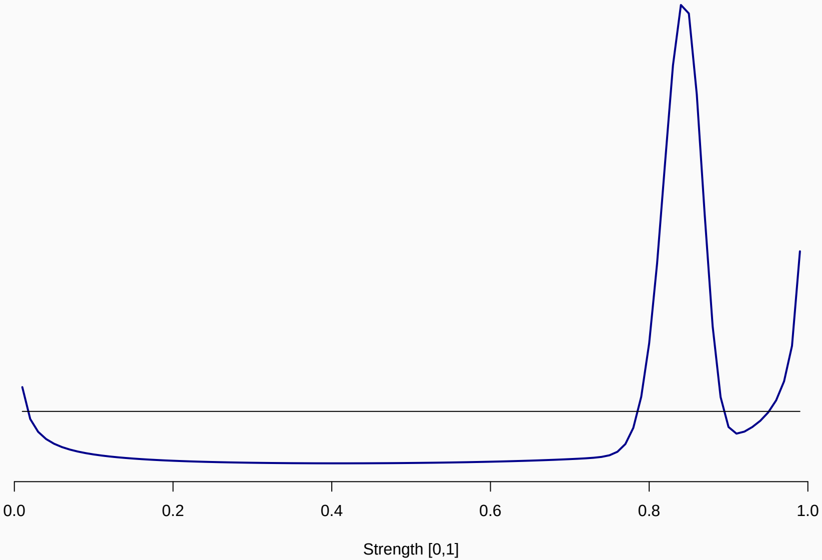
This core always hold

- It is unfalsifiable.
- Easiest proof is by construction.
  - X for new items; Y for old items
  - curve is $h(f)$.
  - set $X \ \sim \mathrm{U}(0,1)$ (for new items)
  - to calculate Y, $1 - F_Y(y) = h(y), \quad y \in [0,1]$
  - also, $f_Y(1-y) = dh(y)/dy$

Example, Is this whacky ROC signal-detection representable?



It is:



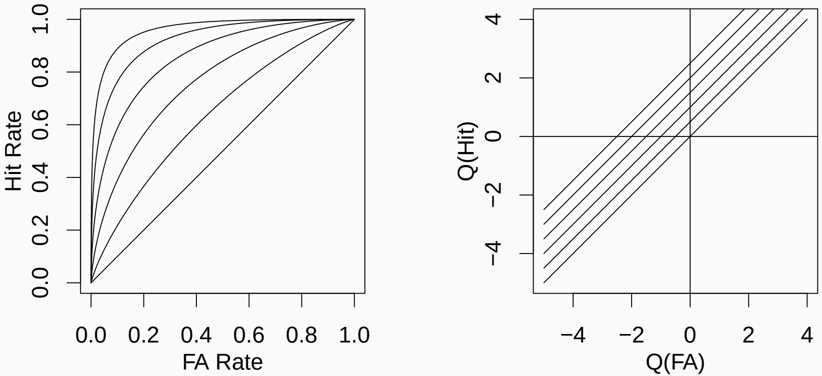
Signal Detection is Not Robust

- Either you buy the normal as part of the process (arbitrary)
- Or you lose falsifiability
- UVSD is just a hack to fit data:
  - must buy unequal variance normal
  - must live with indominance
  - example of injudicious modeling.

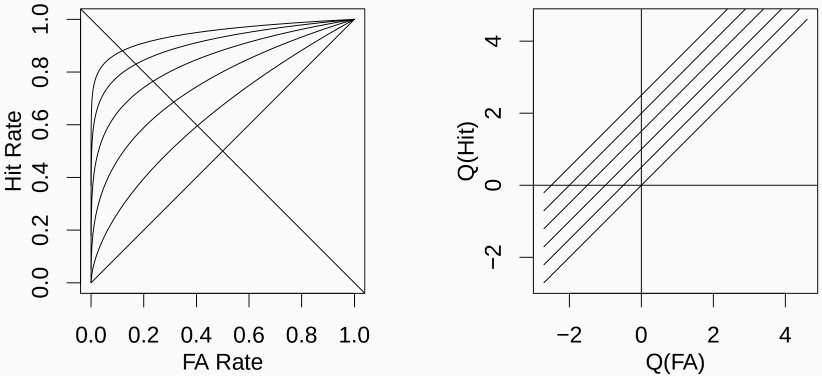
Are There Robust Signal Detection Models?

- Strength affects one part of the distribution.
- Scalar representability?

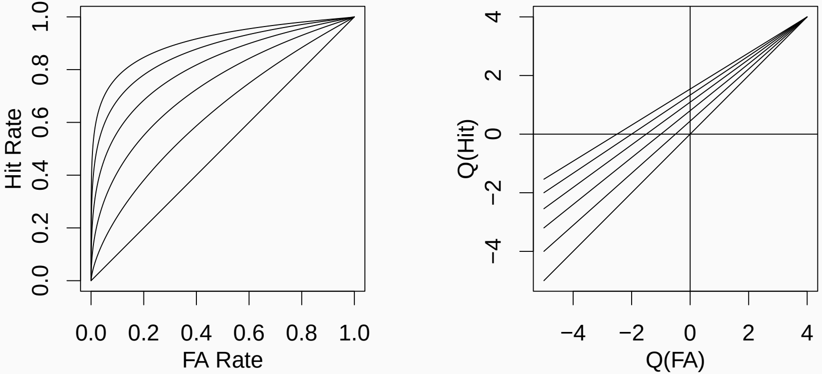
EVSD is Scalar Representable



Gamma is Scalar Representable



UVSD is Not Scalar Representable



A robust signal detection model admits a function Q that linearizes ROCs with constant slope so that there is one difference—intercept, which is strength.

Extension To Item Response Theory

I ppl each respond to J dichotomous items. Let $Y_{ij} = 0, 1$ for incorrect and correct responses, respectively.

$$Y_{ij} \sim \text{Bernoulli}(\pi_{ij})$$

Item Ordering

If item A is harder than item B, then it is harder than item B for all people.

Rasch Model

$$Y_{ij} \sim \text{Bernoulli}(\pi_{ij})$$
$$\pi_{ij} = \frac{1}{1 + \exp[-(\gamma_i - \alpha_j)]}$$

- γ_i: ability of the *i*th person
- α_j: difficulty of the *j*th item

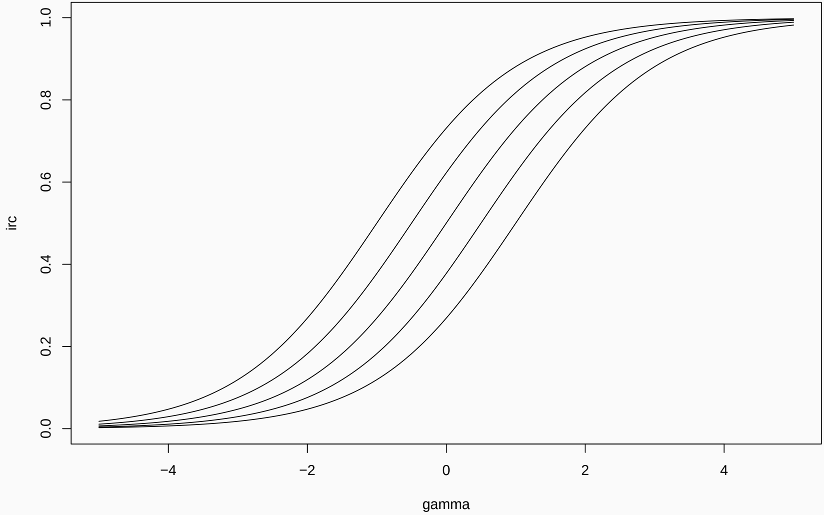
Rasch Model

- Is the Rasch model identifiable as is?
- What is needed?

Your Turn

- Does the Rasch model exhibit item ordering?
- Hint: Draw an item response curve which is a function of γ_i. It should increase as people get better.
- Now compare a few items

My Turn



2PL

Much like in signal detection, researchers thought that the Rasch model was insufficient to capture real-world response curves

$$\pi_{ij} = \frac{1}{1 + \exp[-\alpha_j(\gamma_i - \beta_j)]}$$

- Discriminability: it is a scale effect

Your Turn

Draw some response curves for 2PL

Our Turn

- 2PL is the leading IRT base model
- Changes in discriminability force violations of item ordering
- When might 2PL be appropriate?
- When might it not?

Your Turn

Can you think of general model we could use estimation to explore whether items order?

- two parameters for items (shift, scale)
- both can vary and there is the possibility of crossing
- both can vary and there is the possibility of not crossing