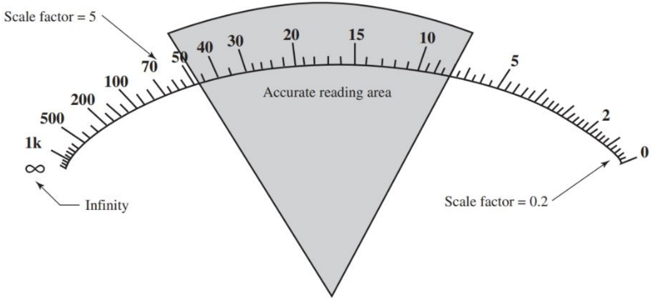
Endogenous Risk Attitudes

Nick Netzer Arthur Robson Jakub Steiner

Zurich, Simon Fraser, Zurich
Cerge-Ei

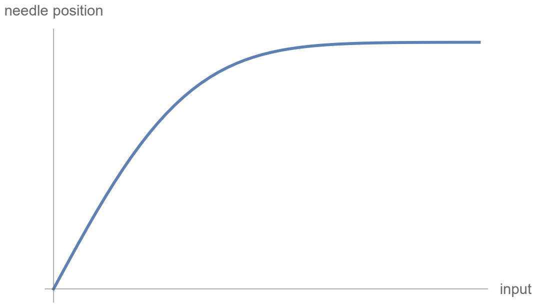
Penn State
March 2021

Nonlinear Scale in Physics



Nonlinear Scale in Physics

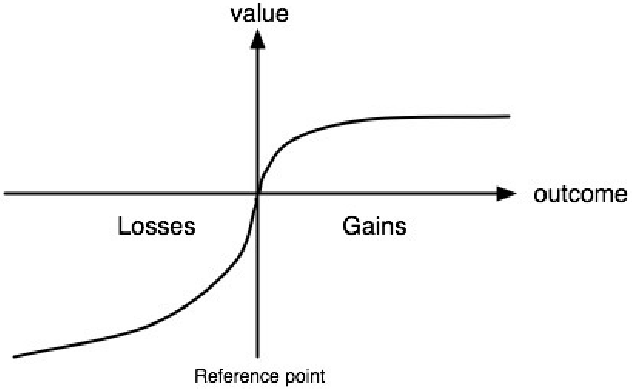
Needle position is non-linear w.r.to input



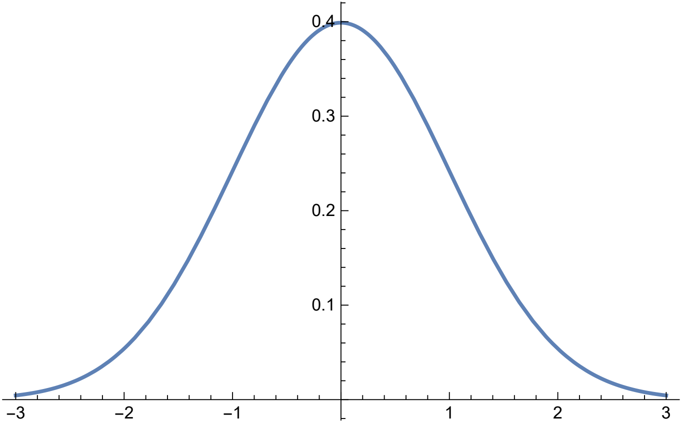
Engineers:

- invert the needle position after the measurement
- customize the non-linearity to the anticipated measurement

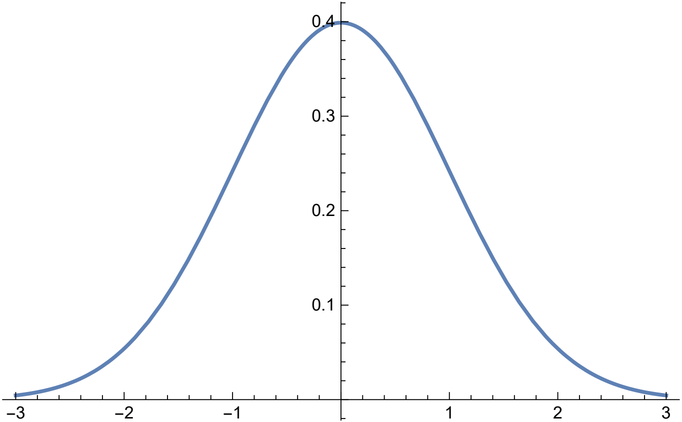
Nonlinear Scale in Psychophysics
Kahneman and Tversky



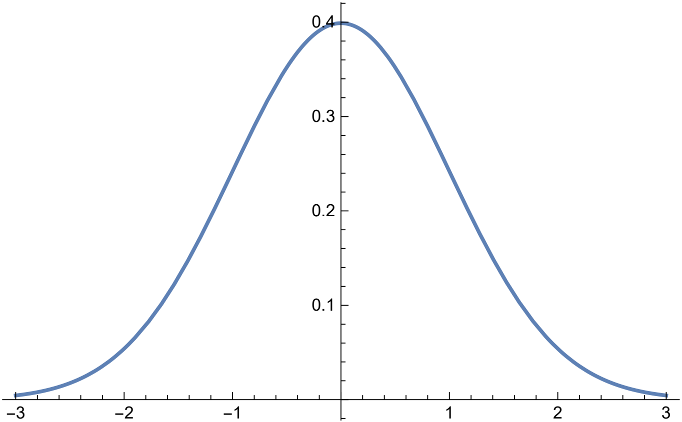
Two draws from:



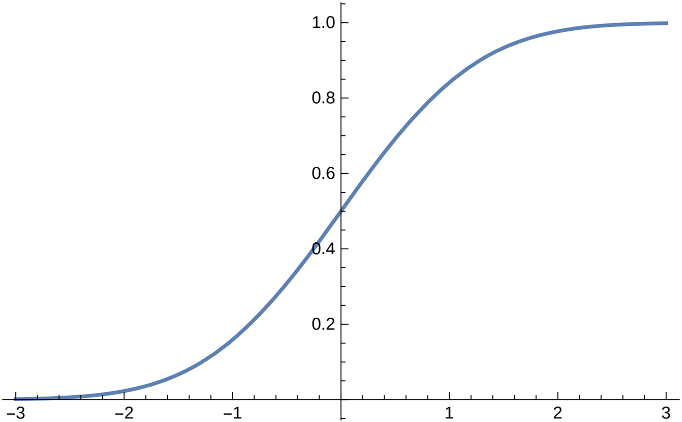
Two draws from: . . . Pick one



Choose your scale (your pointer is noisy)



Choose your scale (your pointer is noisy)



Literature

Psychophysics: Weber's law, Fechner 1860, Thurstone '27

Kahneman&Tversky '79: psychophysics rationale for s-shaped utility

Adaptive encoding of visual stimuli: Attneave '54, Barlow et al. '61, Laughlin et al. '81

Econ [riskless]: Robson '01, Netzer '09, Rayo&Becker '07 (hedonic utility)

Econ [risky]: Khaw&Li&Woodford '20, Frydman&Jin '19 (large encoding noise)

In This Paper

Optimal perception of lotteries (as opposed to simple stimuli)

- s-shaped encoding function
- over-sampling of low-prob arms

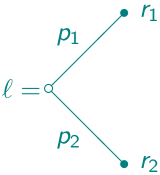
Focus on behavior (Bernoulli instead of hedonic utility)

- surprising risk ⇒ perception-driven risk attitudes
- anticipated risk ⇒ risk-neutrality

Method: asymptotic misspecified learning (White '82, Berk '66)

Table of Contents

Decision Problem



versus alternative s

Risk-neutrality: ℓ is optimal $\Leftrightarrow \sum_i p_i r_i > s$

DM observes $(p_i)_i$ and s frictionlessly

Friction in information-processing of the rewards

Rewards' Perception

Perception strategy:

- encoding function $m : \mathbb{R} \longrightarrow [\underline{m}, \overline{m}]$
- sampling frequencies $(\pi_i)_i \in \Delta(\{\text{set of arms}\})$

DM samples n signals:

- $x_k = (\hat{m}_k, i_k)$
- i_k specifies the lottery arm
- $\hat{m}_k = m(r_{i_k}) + \varepsilon_k$; iid Standard Normal noise
- sampling frequencies π_i distinct from arm probs p_i

Sophistication: DM knows conditional signal distributions

Estimation:

- MLE from a set $\mathcal{A}$ of anticipated lotteries
- or Bayesian estimator for a given prior on $\mathcal{A}$

Nearly complete information: $n \to \infty$

A posteriori optimal choice

Table of Contents

We admit redundant states because

- world'll get more complex after adaptation
- $\Rightarrow$ maladaptation

Arms i and j are payoff-equivalent if $r_i = r_j$ for all lotteries

$\mathcal{J}$ – partition of the set of all arms into payoff-equivalent classes

For now, think about $J \in \mathcal{J}$ as of a lottery arm

Ex Ante Optimization

Environment defined as distribution of the decision problems $(\mathbf{r}, s)$

- all r_J and s are iid from a Normal density

Minimize ex ante expected loss $L(n) = \mathrm{E}\left[\max\{r, s\} - \mathbb{1}_{q_n > s} r - \mathbb{1}_{q_n \leq s} s\right]$,

where r and q_n are the true and estimated lottery values

Proposition

Under a regularity condition ▸ condition

$$\lim_{n\to\infty} L(n) = \text{const.}\,\mathrm{E}\left[\sum_{J\in\mathcal{J}} \frac{p_J^2}{\pi_J m'^2(r_J)} \mid r = s\right]\frac{1}{n} + o\left(\frac{1}{n^2}\right).$$

Ex Ante Optimization

Environment defined as distribution of the decision problems $(\mathbf{r}, s)$

- all r_J and s are iid from a Normal density

Minimize ex ante expected loss $L(n) = \mathrm{E}\left[\max\{r, s\} - \mathbb{1}_{q_n > s} r - \mathbb{1}_{q_n \leq s} s\right]$,

where r and q_n are the true and estimated lottery values

Proposition

Under a regularity condition ▸ condition

$$\lim_{n \to \infty} L(n) \propto \mathrm{E}\left[\text{MSE conditional on tie}\right] + o\left(\frac{1}{n^2}\right).$$

Tie condition because small perception error distorts choice only if $r \approx s$

Ex Ante Optimization

Environment defined as distribution of the decision problems $(\mathbf{r}, s)$

- all r_J and s are iid from a Normal density

Minimize ex ante expected loss $L(n) = \mathrm{E}\left[\max\{r, s\} - \mathbb{1}_{q_n > s} r - \mathbb{1}_{q_n \leq s} s\right]$,

where r and q_n are the true and estimated lottery values

Proposition

Under a regularity condition ▸ condition

$$\lim_{n\to\infty} L(n) \propto \mathrm{E}\left[\text{MSE conditional on tie}\right] + o\left(\frac{1}{n^2}\right).$$

MSE because prob of choice distortion $\propto$ error size, and loss is too

Ex Ante Optimization

Environment defined as distribution of the decision problems $(\mathbf{r}, s)$

- all r_J and s are iid from a Normal density

Minimize ex ante expected loss $L(n) = \mathrm{E}\left[\max\{r, s\} - \mathbb{1}_{q_n > s} r - \mathbb{1}_{q_n \leq s} s\right]$, where r and q_n are the true and estimated lottery values

Proposition

Under a regularity condition ▸ condition

$$\lim_{n\to\infty} L(n) \propto \mathrm{E}\left[\sum_{J\in\mathcal{J}} p_J^2 \mathrm{MSE}(r_J) \text{ conditional on tie}\right] + o\left(\frac{1}{n^2}\right).$$

MSE is a weighted sum of MSEs for each r_J

Ex Ante Optimization

Environment defined as distribution of the decision problems $(\mathbf{r}, s)$

- all r_J and s are iid from a Normal density

Minimize ex ante expected loss $L(n) = \mathrm{E}\left[\max\{r, s\} - \mathbb{1}_{q_n > s} r - \mathbb{1}_{q_n \leq s} s\right]$, where r and q_n are the true and estimated lottery values

Proposition

Under a regularity condition ▸ condition

$$\lim_{n\to\infty} L(n) = \text{const.}\,\mathrm{E}\left[\sum_{J\in\mathcal{J}} \frac{p_J^2}{\pi_J m'^2(r_J)} \mid r = s\right]\frac{1}{n} + o\left(\frac{1}{n^2}\right).$$

MSE for r_J is mitigated by high π_J or $m'(r_J)$

Ex Ante Optimization

Environment defined as distribution of the decision problems $(\mathbf{r}, s)$

- all r_J and s are iid from a Normal density

Minimize ex ante expected loss $L(n) = \mathrm{E}\left[\max\{r, s\} - \mathbb{1}_{q_n > s} r - \mathbb{1}_{q_n \leq s} s\right]$,

where r and q_n are the true and estimated lottery values

Proposition

Under a regularity condition ▸ condition

$$\lim_{n \to \infty} L(n) = \text{const.}\, \mathrm{E}\left[\sum_{J \in \mathcal{J}} \frac{p_J^2}{\pi_J m'^2(r_J)} \mid r = s\right] \frac{1}{n} + o\left(\frac{1}{n^2}\right).$$

Tie conditioning is implied by consequentialism

Information-Processing Problem

$$\min_{m'(\cdot),(\pi_J)_J>0} \mathrm{E}\left[\sum_{J\in\mathcal{J}} \frac{p_J^2}{\pi_J m'^2(r_J)} \mid r=s\right]$$

$$\text{s.t.:} \int_{\mathbb{R}} m'(r)dr \leq \overline{m} - \underline{m}$$

$$\sum_{J\in\mathcal{J}} \pi_J = 1$$

Constraints:

- $m(\cdot)$ is bounded – your 'scale' can't be fine everywhere
- $\sum_J \pi_J = 1$ – you can't sample all the arms frequently

Optimal Perception

Proposition

1. Optimal encoding function m is s-shaped
   - convex below and concave above the reward mode
2. Over-sampling of low-prob arms
   - binary lotteries: if $p_J < 1/2$, then $\pi_J > p_i$ and vice versa
   - $I > 2$: for any two arms J, J' such that $p_J < p_{J'}$, $\frac{\pi_J}{p_J} > \frac{\pi_{J'}}{p_{J'}}$

Intuition

1. s-shape
   - $m(\cdot)$ steep at reward values that you're likely to encounter at ties
2. Over-sampling
   - diminishing return to sampling
   - over-sample the arm that you expect to be poorly informed on
   - you measure tail rewards poorly
   - low-prob arm has more spread-out rewards conditional on tie since $\sum_{J'} p_{J'} r_{J'} = s$ isn't too informative on r_J

Optimal Perception

Proposition

1. Optimal encoding function m is s-shaped
   - convex below and concave above the reward mode
2. Over-sampling of low-prob arms
   - binary lotteries: if $p_J < 1/2$, then $\pi_J > p_{J'}$ and vice versa
   - $I > 2$: for any two arms J, J' such that $p_J < p_{J'}$, $\frac{\pi_J}{p_J} > \frac{\pi_{J'}}{p_{J'}}$

Intuition

1. s-shape
   - $m(\cdot)$ steep at reward values that you're likely to encounter at ties
2. Over-sampling
   - diminishing return to sampling
   - over-sample the arm that you expect to be poorly informed on
   - you measure tail rewards poorly
   - low-prob arm has more spread-out rewards conditional on tie since $\sum_{J'} p_{J'} r_{J'} = s$ isn't too informative on r_J

Verbally

Canonical example: flying involves a small prob of accident

Accident is a tail event – hard to assess

If a nontrivial choice features a tail event, then the event has a small prob

otherwise, the choice is trivial

⇒ Small probs are often attached to tail events in nontrivial choices

Oversampling of small prob events compensates for this

Table of Contents

Example
allude to Savage '54

DM chooses whether to buy a convertible car

Reward from the convertible, r_1 or r_2, depends on weather

DM samples n signals:

- $i_k \in \{1, 2\}$ – weather for k'th sampled experience
- $\hat{m}_k = m(r_{i_k}) + \varepsilon_k$ – k'th perturbed message

Both true type probs and sampling probs are 50-50

fine DM

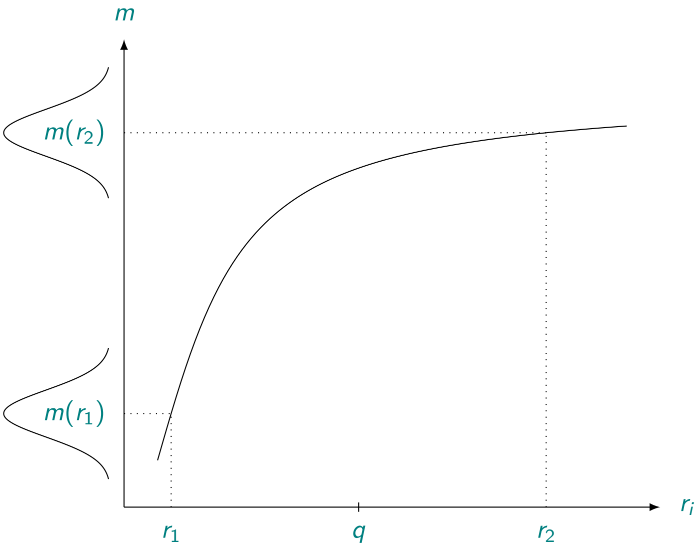
- understands role of weather
- anticipates $(r_1, r_2) \in \mathbb{R}^2$
- well-specified

coarse DM

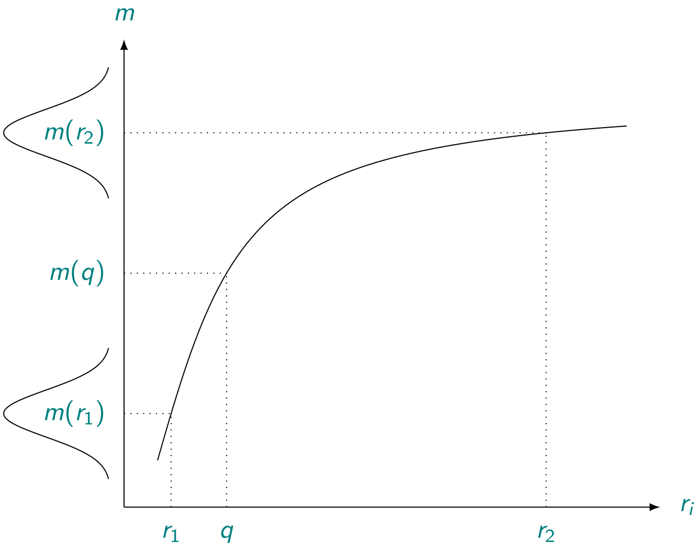
- disregards weather
- anticipates (r, r), $r \in \mathbb{R}$
- misspecified

DM's estimate of the car's value?

Fine DM $\Rightarrow$ Risk Neutrality



Coarse DM $\Rightarrow$ EUT



Paths to Misspecification

Complexity increase:

- adaptation took place in riskless world
- world got risky
- DM continues to model it as riskless

or

DM got framed:

- adaptation took place in risky world
- afterwards, DM got convinced that the next lottery is riskless

Expected-Utility Representation

DM anticipates no risk: $\mathcal{A} = \{\mathbf{r} \in \mathbb{R}^I : r_i = r_j \text{ for all arms } i, j\}$

Proposition

Prob that DM chooses the lottery in problem $(\mathbf{r}, s)$ converges to 1 (0) if

$$\sum_i \pi_i m(r_i) > (<) m(s).$$

Proof based on White '82:

- MLE $\xrightarrow{\text{a.s.}} \arg\min_{r' \in \mathcal{A}} D_{KL}(f_{\mathbf{r}}, f_{\mathbf{r}'})$
- Gaussian errors $\Rightarrow$
  - MLE of m is the convex combination of $m(r_i)$ for each arm i
  - with weights equal to the sampling frequencies

Berk '66 for the analogous result for Bayesian estimation

‘Risk Attitudes’ of Engineers

Bouncing needle caused by stochastic input

‘Risk attitudes’ emerge if

- engineer misattributes the tremble to stochasticity of measurement

Omitted Variable
illustration

Reward $\rho(\mathbf{x}, \mathbf{y})$

- $(\mathbf{x}, \mathbf{y})$ drawn from a joint density

DM omits variables $\mathbf{y}$: she thinks that the reward is $\tilde{\rho}(\mathbf{x})$

For each $\mathbf{x}$, she

- observes n signals $m(\rho(\mathbf{x}, \mathbf{y}_k)) + \varepsilon_k$
- estimates $\tilde{\rho}(\mathbf{x})$

For each $\mathbf{x}$

- the reward $\rho(\mathbf{x}, \mathbf{y})$ is a lottery since $\rho(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}$ is random
- DM conceptualizes this lottery as a riskless reward $\tilde{\rho}(\mathbf{x})$

Economist

- incorrectly thinks that DM is well-specified
- concludes that DM has Bernoulli utility $u(\cdot) = m(\cdot)$

Coarse Anticipation of Risk

$\mathcal{K}$ – a partition of the set of all arms

DM anticipates lotteries to be measurable w.r.to $\mathcal{K}$

Proposition (mixed representation)

Prob that DM chooses the lottery in problem $(\mathbf{r}, s)$ converges to 1 (0) if

$$\sum_{J \in \mathcal{K}} p_J r_J^* > (<) s,$$

where for each $J \in \mathcal{K}$:

- r_J^* is 'certainty equivalent': $m\left(r_J^*\right) = \sum_{i \in J} \frac{\pi_i}{\sum_{j \in J} \pi_j} m(r_i)$
- $p_J = \sum_{i \in J} p_i$ is the true prob of J

Corollary: risk-neutrality w.r.to anticipated lotteries

Omitted Variable (continued)

As before

- reward $\rho(\mathbf{x}, \mathbf{y})$
- DM omits $\mathbf{y}$ and estimates $\tilde{\rho}(\mathbf{x})$ using encoding m

But

- at the point of decision, observes only a signal $\mathbf{z}$ of $\mathbf{x}$

Each value of $\mathbf{z}$

- corresponds to a lottery over $\rho(\mathbf{x},\mathbf{y}) \mid \mathbf{z}$
- DM thinks the lottery is over $\tilde{\rho}(\mathbf{x}) \mid \mathbf{z}$ and computes $\mathrm{E}\left[\hat{\rho}(\mathbf{x}) \mid \mathbf{z}\right]$

Representation of DM:

- for each $\mathbf{x}$, she computes c.e. over uncertainty $\mathbf{y} \mid \mathbf{x}$ under Bernoulli utility $u = m$,
- proceeds as risk-neutral w.r.to uncertainty $\mathbf{x} \mid \mathbf{z}$

Table of Contents

Bayesian Robustness Check

Let's bridge two extreme cases:

- anticipated lotteries
- surprising lotteries

Joint limit of:

- number of signals
- precision of the prior density

We get

- robustness check
- comparative statics with respect to
    - time pressure
    - level of anticipated risk

Prior Belief

Binary lottery is drawn from prior density

$$\exp\left(-\frac{n}{\Delta}\left(r_1 - r_2\right)^2\right)$$

Prior is concentrated alongside riskless lotteries on the diagonal

Δ parametrizes the degree of the a priori anticipated risk

As $n \nearrow$, risk becomes a priori unlikely

Sampling

$a \times n$ perturbed messages

a captures decision span:

- sample size increases with a for fixed n

n has a double role. As $n \nearrow$:

- risk becomes a priori unlikely
- sample size grows

Arrow-Pratt Measure

Realized rewards $r_1 = r + \delta$, $r_2 = r - \delta$, 50-50 probs, uniform sampling

Proposition

As $n \to \infty$, DM's valuation of the lottery converges to

$$r + \frac{1}{2}\frac{m''(r)}{m'(r)}\frac{1 + a\Delta m'^2(r)}{(1 + a\Delta m'^2(r)/2)^2}\delta^2 + o(\delta^3).$$

DM:

- thinks that $r_i = r^* \pm \delta'$ for $\delta' < \delta$ (large risk is unlikely)
- then, must shift r^* relative to r to fit data (due to curvature of m)

Thinking fast/slow:

- risk attitudes decrease with time span (a)

Rabin's paradox:

- risk attitudes decrease with anticipated risk (Δ)

Conclusion

Optimal attention-allocation

- s-shaped encoding function and over-sampling of low-prob arms

Link between reward encoding and risk attitudes is subtle

- psychophysics intuition applies to surprising lotteries

Two adaptation channels

- slow: optimal encoding
- fast: anticipation of lotteries

Regularity Condition

There exists $e(\mathbf{r}, \varepsilon) \geq n \times \text{MSE}(\mathbf{r}, \varepsilon)$ such that $\text{E}\, e(\mathbf{r}, \varepsilon) < +\infty$.

$\text{MSE}(\mathbf{r}, \varepsilon)$ is of order $1/n$ because $q_i - r_i \approx \frac{\varepsilon_i}{\sqrt{\pi_i n} m'(r_i)}$

But $m(\cdot)$ gets flat at tails

$\Rightarrow$ Perception error diverges at tail rewards

RC requires reward density to vanish fast enough at tails relative to $m'(\cdot)$

It allows for application of Dominated Convergence Theorem

back