center for neural science and howard hughes medical institute, nyu

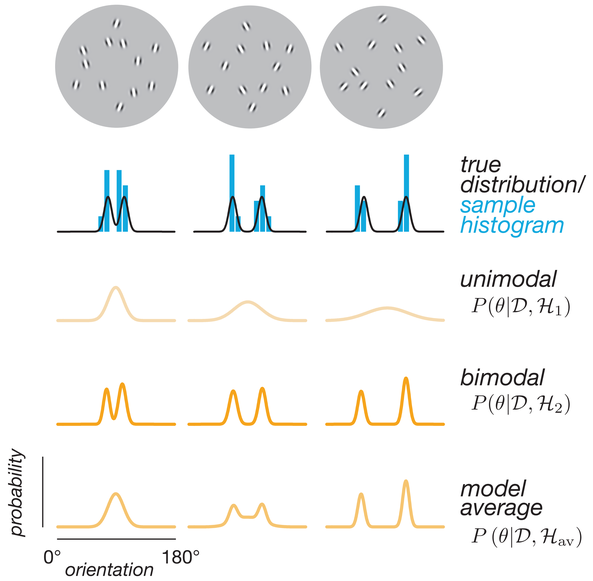
jeremy freeman, deep ganguli, and eero p. simoncelli...

Bayesian estimation of probability distributions in perception

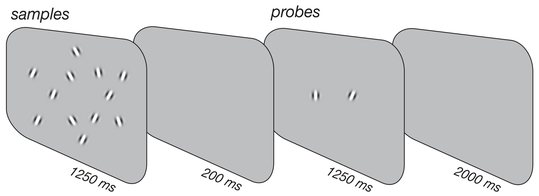
introduction

human behavior appears to reflect probabilistic inference over sensory and motor variables

what kind of probability distribution might an observer represent when shown samples from a distribution?

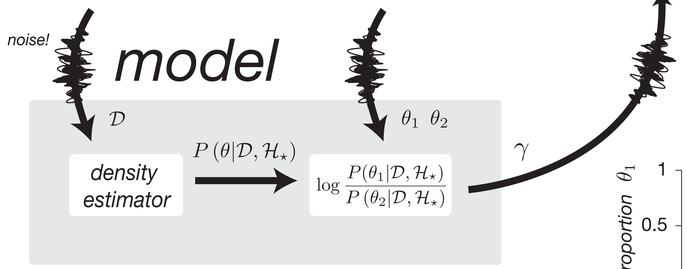


experiment



which one is more likely to have come from the set: θ_1 or θ_2 ?

model



observer model details

1 *fit a unimodal and bimodal distribution to the noisy samples*

$$P(\theta|\boldsymbol{w}_1,\mathcal{H}_1)=\mathcal{N}\left(\theta;\mu,\sigma^2\right) \quad \boldsymbol{w}_1=[\mu,\sigma]$$

$$P(\theta|\boldsymbol{w}_2,\mathcal{H}_2)=\pi\mathcal{N}\left(\theta;\mu_1,\sigma_1^2\right)+(1-\pi)\mathcal{N}\left(\theta;\mu_2,\sigma_2^2\right)$$

$$\boldsymbol{w}_2=[\pi,\mu_1,\sigma_1,\mu_2,\sigma_2]$$

$P(\boldsymbol{w}_i|\mathcal{D},\mathcal{H}_i)$ *is approximated via Bayesian variational inference [Bishop 2006]. best fitting model parameters,* $\hat{\boldsymbol{w}}_i(\mathcal{D})$, *are given by the mean of this distribution.*

2 *approximate the posterior probability of each model*

$$P(\mathcal{H}_i|\mathcal{D})\propto P(\mathcal{D}|\mathcal{H}_i)P(\mathcal{H}_i)$$

*the **prior** describes the observer's preference for bimodal vs. unimodal model*

$$P(\mathcal{D}|\mathcal{H}_i)=\int\prod_n P(\theta=\mathcal{D}_n|\boldsymbol{w}_i,\mathcal{H}_i)P(\boldsymbol{w}_i|\mathcal{H}_i)\,\mathrm{d}\boldsymbol{w}_i$$

*the **evidence** naturally penalizes the bimodal hypothesis if it "overfits" the density to the samples [MacKay 2003]*

3 *compute Bayesian model averager*

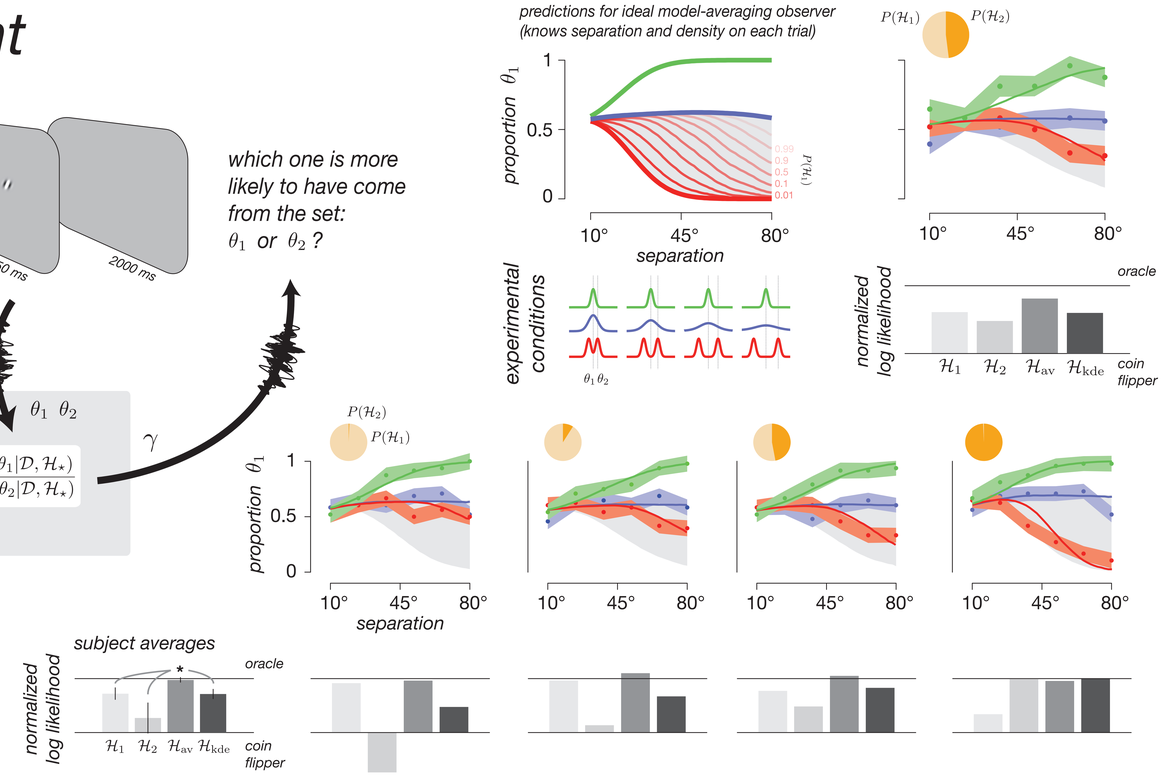
$$P(\theta|\mathcal{D},\mathcal{H}_{\mathrm{av}})=P(\theta|\hat{\boldsymbol{w}}_1(\mathcal{D}),\mathcal{H}_1)P(\mathcal{H}_1|\mathcal{D})+P(\theta|\hat{\boldsymbol{w}}_2(\mathcal{D}),\mathcal{H}_2)P(\mathcal{H}_2|\mathcal{D})$$

4 *model probability observer chooses* θ_1

$$\gamma=\log\frac{P(\theta_1|\mathcal{D},\mathcal{H}_\star)}{P(\theta_2|\mathcal{D},\mathcal{H}_\star)} \qquad P(\text{subject response}=\text{“}\theta_1\text{”}|\gamma)=\frac{1}{1+e^{-a\gamma+b}}$$

passing the log LR ratio through a logistic function models decision noise

model fitted to data



conclusion

perceptual judgements about relative probability are not consistent with exclusively unimodal or bimodal density estimates. rather, perceptual judgements are best explained by a Bayesian model averaging strategy.