Tutorial in Bayesian modeling of perception

20160317, NYU Shanghai
Wei Ji Ma, Center for Neural Science and Department of Psychology, New York University

Bayes: the concept

Bayesian inference can explain many illusions, such as:

- Dragon illusion: https://www.youtube.com/watch?v=2K39Q9zvQoE
    - Make it yourself: see attached, from http://www.grand-illusions.com/images/articles/opticalillusions/dragon_illusion/dragon.pdf
- The Dress: http://www.cns.nyu.edu/malab/illusionstoinference/thedress/
- Bad lip reading: https://www.youtube.com/user/BadLipReading

Why do Bayesian modeling of behavior?

- Modeling behavior is useful to understand the processes underlying decisions, to characterize individual differences, and so that you have variables to correlate neural activity with
- Modeling is easiest if the stimuli have unambiguously defined features, and you parametrically vary one or more features.
- We will focus on the Bayesian observer, because: a) in many tasks people are close to Bayesian, b) because it is can serve as a starting point for constructing other models.
- Anybody can do modeling! The math might seem hard at first but after a while it is more of the same.

Bayesian modeling of a "match-to-sample" or "change detection" task

Task: Assume a one-dimensional, real-valued stimulus s. The subject's task is to determine whether two stimuli are the same ($\Delta s=0$) or different (Δs drawn from a Gaussian distribution with mean 0 and standard deviation σ_s). "Same" and "different" trials occur with equal probabilities.

Generative model: $C \rightarrow \Delta s \rightarrow \Delta x$, where: C: category (0 for same, 1 for different), Δs: physical stimulus difference, Δx: noisy measurement of the stimulus difference (drawn from Gaussian with mean Δs and standard deviation σ). On "same" trials ($C=0$), we have $\Delta s=0$, so the distribution of Δx is Gaussian with mean 0 and variance σ^2. On "different" trials ($C=1$), the external variance in Δs and the internal variance in Δx sum, so the distribution of Δx is Gaussian with mean 0 and variance $\sigma_s^2+\sigma^2$.

Inference: Posterior distribution of interest: $p(C\,|\,x)$. Reformulate as posterior ratio, $d=\frac{p(C=1\,|\,\Delta x)}{p(C=0\,|\,\Delta x)}$. Work out using the generative model:

$$d=\frac{p(\Delta x\,|\,C=1)}{p(\Delta x\,|\,C=0)}\frac{p(C=1)}{p(C=0)}=\frac{\frac{1}{\sqrt{2\pi\left(\sigma^2+\sigma_s^2\right)}}e^{-\frac{\Delta x^2}{2\left(\sigma^2+\sigma_s^2\right)}}}{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{\Delta x^2}{2\sigma^2}}}\cdot\frac{0.5}{0.5}=\sqrt{\frac{\sigma^2}{\sigma^2+\sigma_s^2}}e^{-\frac{\Delta x^2}{2}\left(\frac{1}{\sigma^2+\sigma_s^2}-\frac{1}{\sigma^2}\right)}$$

Maximum-a-posteriori decision rule:
The optimal Bayesian observer would report "different" when the probability that C=1 exceeds the probability that C=0. This is equivalent to the condition d>1, which in turn can be simplified to

$$\left|\Delta x\right|>\sigma\sqrt{\left(1+\frac{\sigma^2}{\sigma_s^2}\right)\log\left(1+\frac{\sigma_s^2}{\sigma^2}\right)}.$$

This is called a "decision rule".

Response probabilities: To obtain predictions for an experiment, we can simulate the decision-rule over many trials and obtain hit rate, false-alarm rate, and a full psychometric curve (probability of reporting "different" for different s). In this specific example, we can also write down equations for each of these quantities, but that is not possible in general.

As an example, the following Matlab code computes false-alarm and hit rates:

```
clear; close all;

Ntrials = 1e4;
sig = 2;
sigs = 5;

% C=0 trials (same)
Delta_x = randn(Ntrials,1)* sig;
p_diff = mean(abs(Delta_x) > sig * sqrt((1+sig^2/sigs^2) *
log(1+sigs^2/sig^2)));
falsealarmrate = p_diff

% C=1 trials (different)
Delta_x = randn(Ntrials,1)* sqrt(sig^2+sigs^2);
p_diff = mean(abs(Delta_x) > sig * sqrt((1+sig^2/sigs^2) * log
(1+sigs^2/sig^2)));
```

```
hitrate = p_diff
```

Once you have the code or the equations, you can play around with it to test whether the model behaves sensibly. For example, what happens to the hit and false-alarm rates when you increase σ? What when you increase σ_s?

Model fitting: The last step of Bayesian modeling is to fit the free parameters in the model to individual-subject data. In this example, there is only one free parameter, namely σ. There are many algorithms you can use for fitting, for example "fmincon" in Matlab.