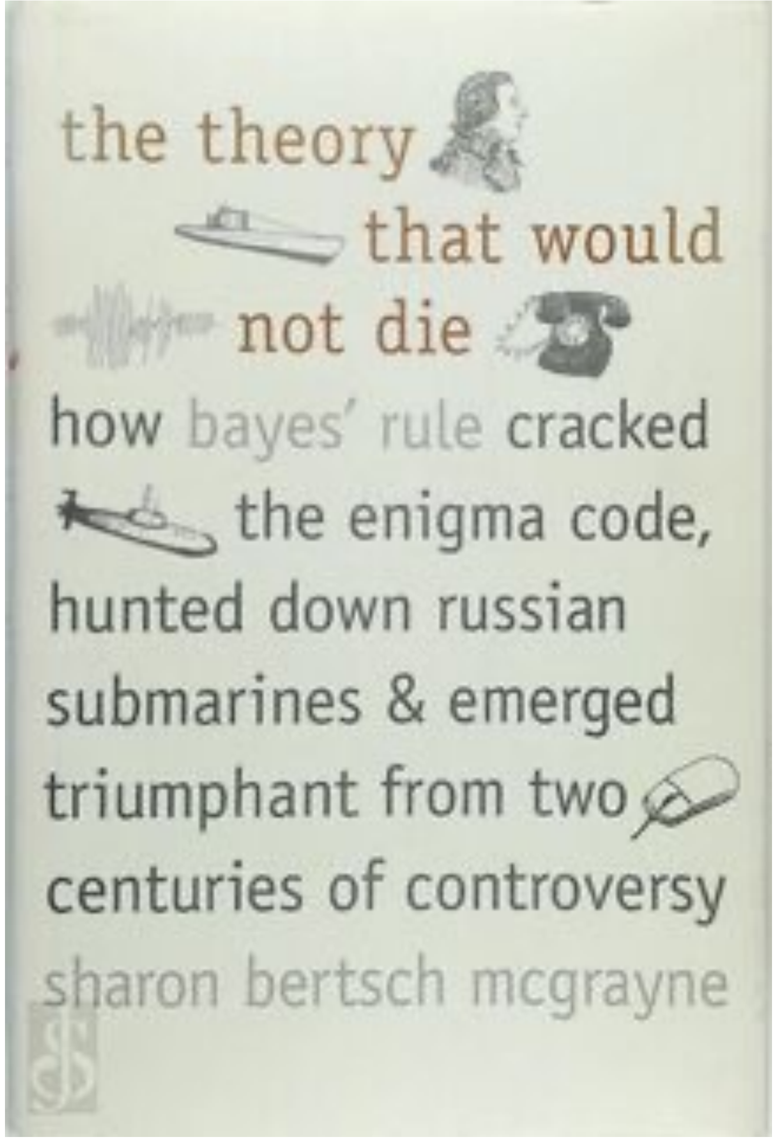
Announcements (Mon, Nov 18)

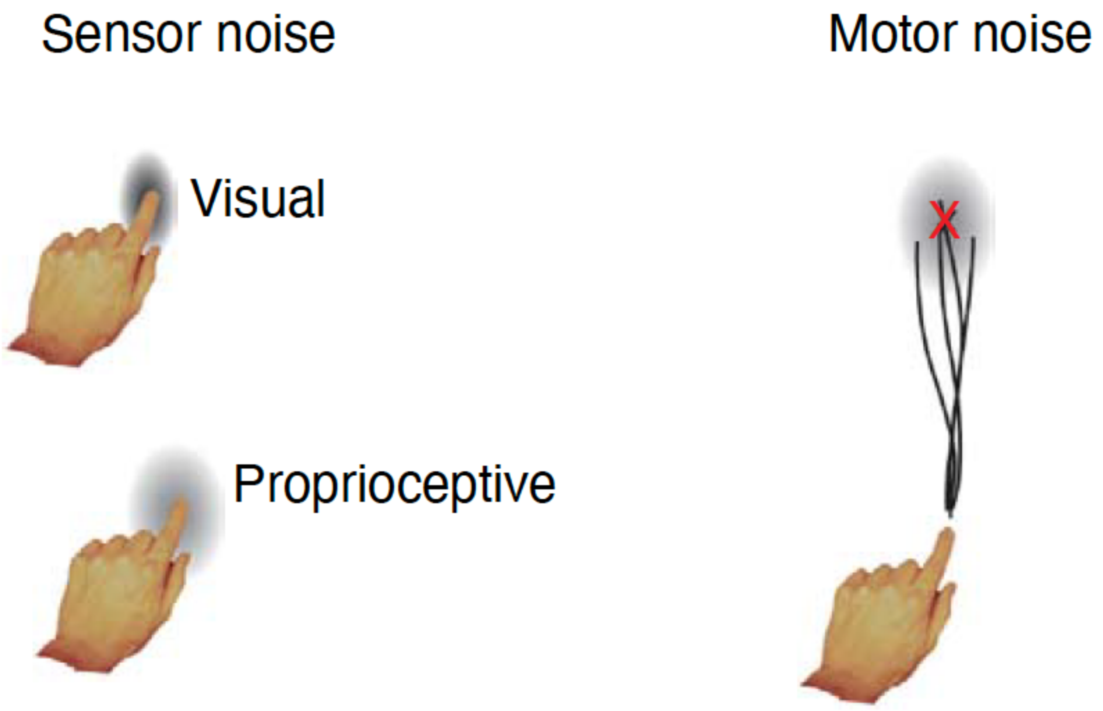
- A6:
  - About one-third of you received a 0.5/1 because you either didn't write a proper function or you made a mistake in one of the bootstrapping steps
  - Remember: a function should be flexible; if you hard code in something that is supposed to be a parameter (e.g., the # of bootstrap samples, the confidence level, etc.), then the function no longer serves its original purpose AND it's incorrect because the user thinks they can set the value of something they actually cannot
  - For each bootstrap sample, how many observations should you be (re)sampling from?

Bayesian statistics



(From KIN 482D) What sort of modelling will you learn?

- We will be modeling the human sensorimotor system, specifically with regard to perception, motor planning, and motor learning
- One of the great mysteries of sensorimotor control, and neuroscience, is how the nervous system deals with uncertainty
- **Bayesian modelling (inference)** is one of the most successful approaches to this problem – also happens to be one of the best foundations for all other types of modelling



From Kording & Wolpert (2006)

Motivating examples: Multi-sensory integration (“cue combination”)

- What do I mean by multisensory integration?

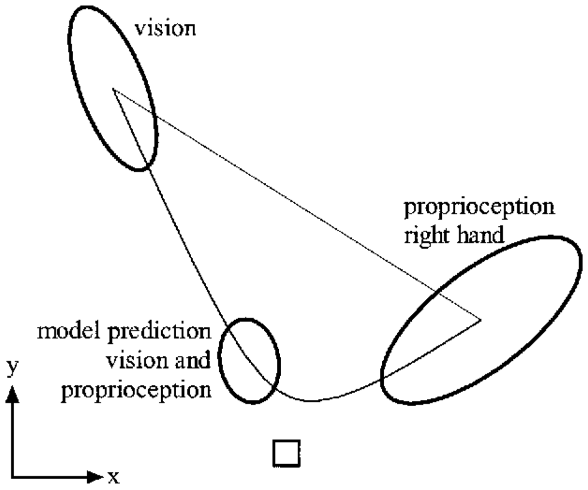
- What are some examples?

Motivating examples: Multi-sensory integration ("cue combination")

The ventriloquist effect

Motivating examples: Multi-sensory integration ("cue combination")

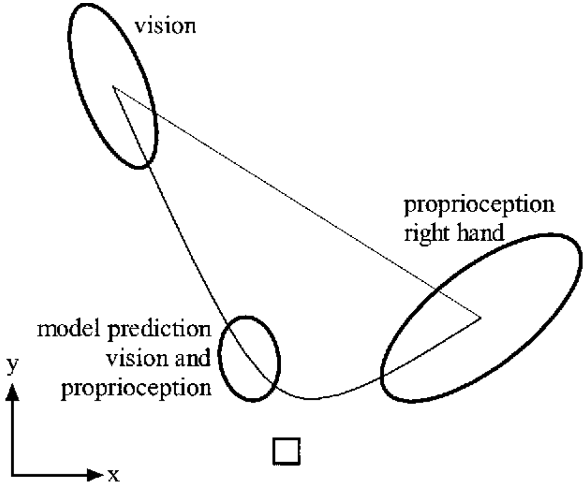
The ventriloquist effect



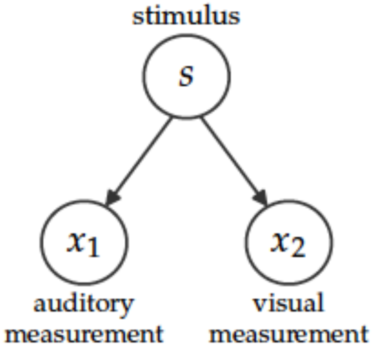
Vision and proprioception (van Beers et al 1999)

Motivating examples: Multi-sensory integration ("cue combination")

The ventriloquist effect



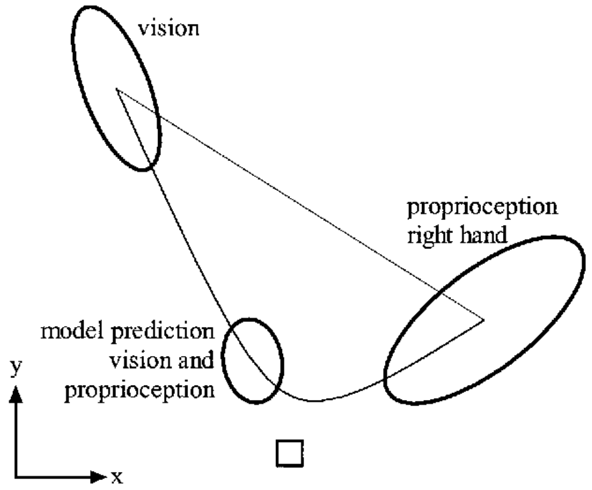
Vision and proprioception (van Beers et al 1999)



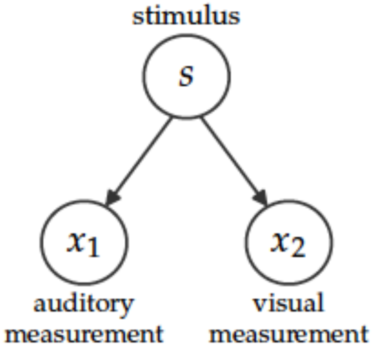
Graphical generative model

Motivating examples: Multi-sensory integration ("cue combination")

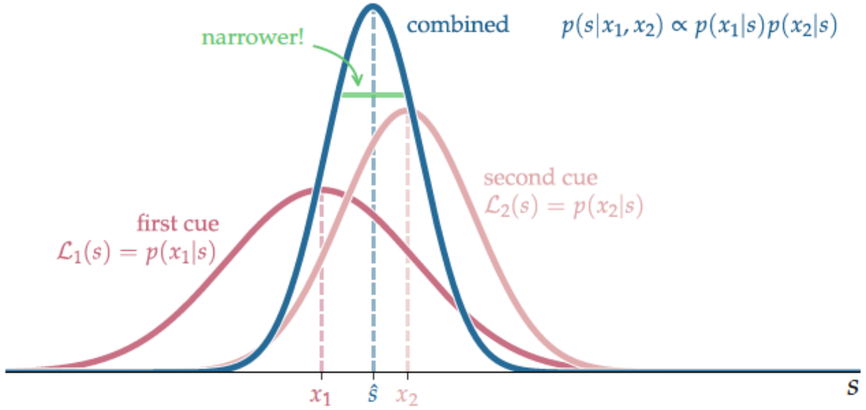
The ventriloquist effect



Vision and proprioception (van Beers et al 1999)



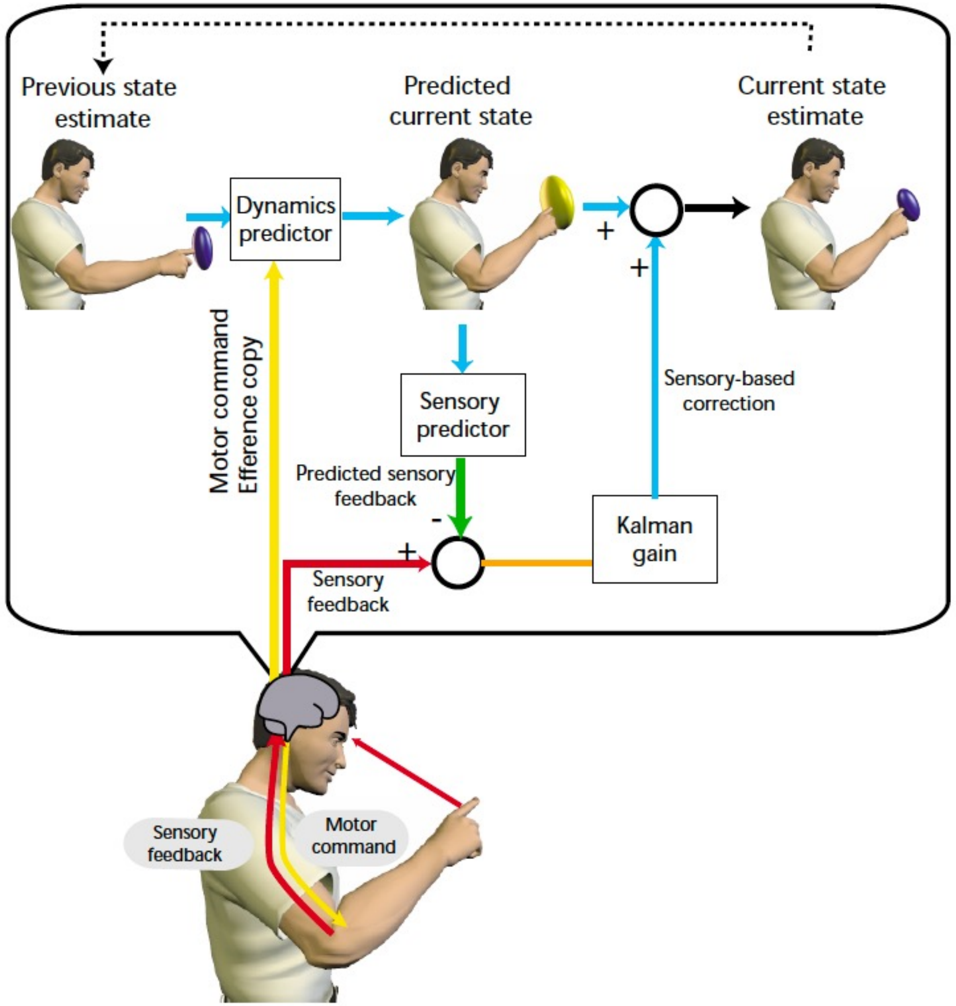
Graphical generative model



Motivating examples: sensorimotor learning

Kalman filter model

- Optimal combination of prior information with incoming measurements
- Model of sensorimotor control and learning

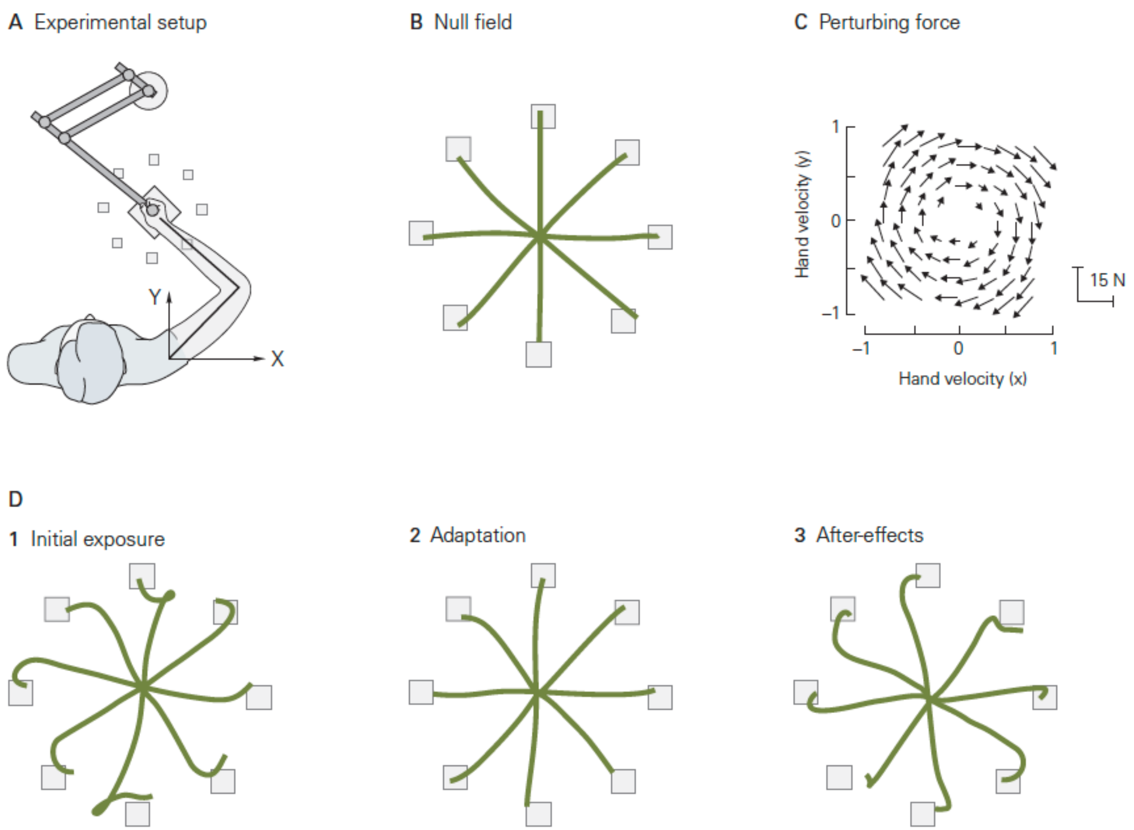


Wolpert and Ghahramani (2000)

Motivating examples: sensorimotor learning

Kalman filter model

- Optimal combination of prior information with incoming measurements
- Model of sensorimotor control and learning
  - Sensorimotor adaptation – recalibration of sensorimotor maps

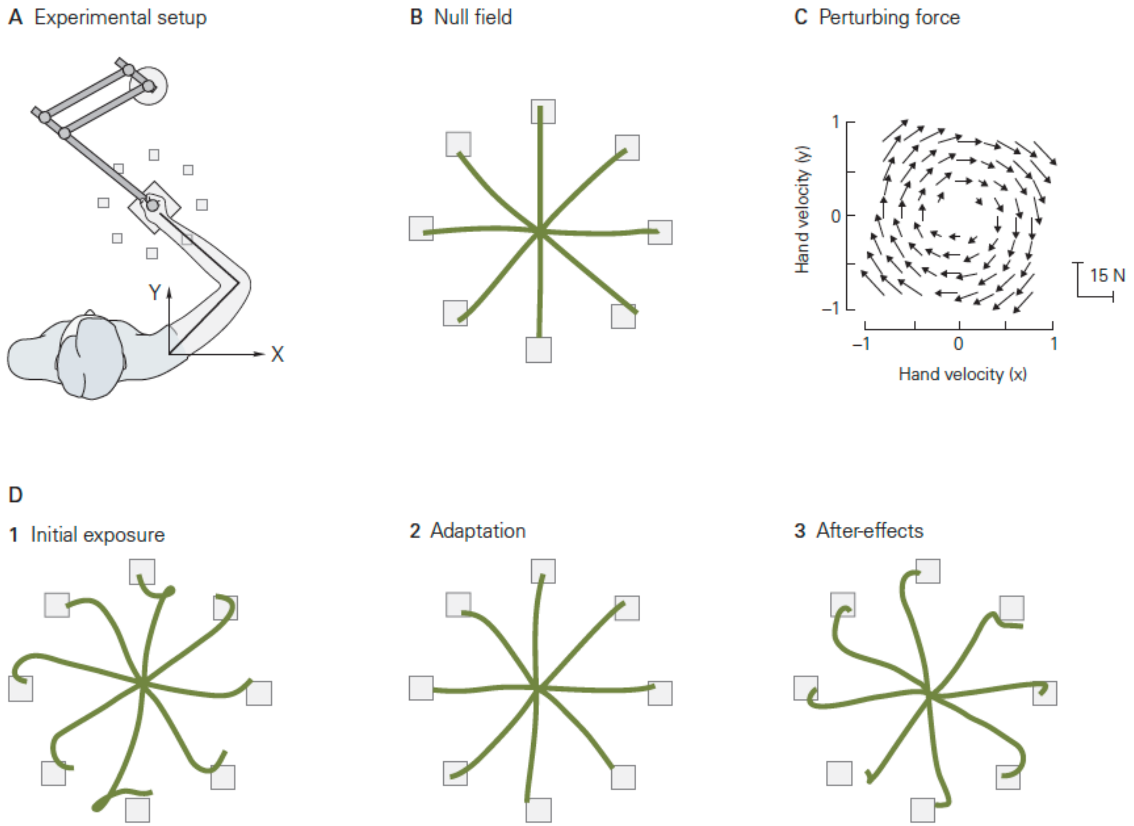


Kandel 5th ed.

Motivating examples: sensorimotor learning

Kalman filter model

- Optimal combination of prior information with incoming measurements
- Model of sensorimotor control and learning
  - Sensorimotor adaptation – recalibration of sensorimotor maps
- More generally, extremely useful algorithm used in GPS, spacecraft, brain-machine interfaces, and many other applications



Kandel 5th ed.

Why Bayesian models?

- Provides very principled framework for learning modeling, whether you're interested in perception, motor control, learning, decision-making, etc.

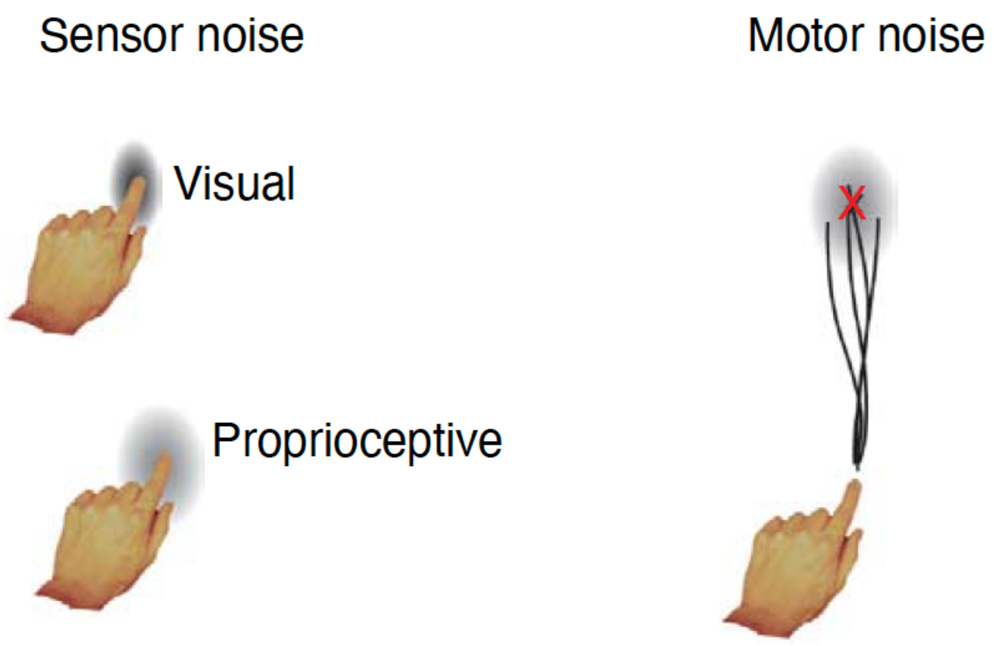
- Very successful track record (many examples throughout semester)

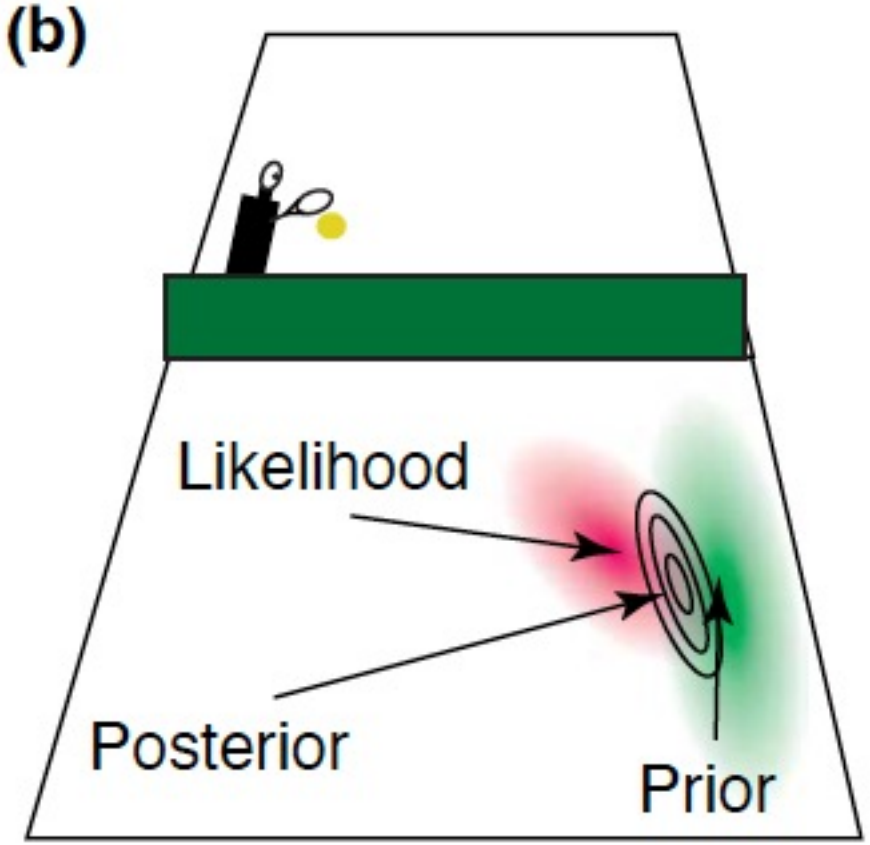
- Our brains may be Bayesian!
  - We make perceptual, motor judgments by combining prior beliefs with incoming sensory information
  - Our brains appear to do this in a statistically optimal way for many tasks

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Rev. Thomas Bayes (1702-1761)

“Bayesian decision theory in sensorimotor control”





Bayesian inference is used for sensing, guiding, planning, and predicting movement.

Kording and Wolpert (2006)

Announcements (Wed, 11/20)

- Final project rubric up on Canvas – 1 pt if you completed and turned in part 1 on time
- A7
  - (“Option” 1) For those who did not receive full marks on A6. Must resubmit **annotated** version of A6
    - Fix your code AND create a new cell explaining where you originally went wrong and how you fixed it
    - Resubmit to A6 on Canvas
  - (Option 2) Provide a one-paragraph update on where you are at with final project.
    - What have you done so far? What has gone smoothly? What has been difficult? Be specific.
    - Feel free to prep me on anything you want me to know before I see your final submission.
    - If you haven’t made progress since part 1, provide clear roadmap with dates for completing the project
    - Submit A7 notebook to A7 on Canvas

- Read version of Signal Processing chapter that is uploaded to JupyterHub – online textbook chapter is not rendering correctly

Plan for today

- Review Monday’s material
- Complete lecture 16 notebook
- Time permitting, go through some examples of Bayesian modeling

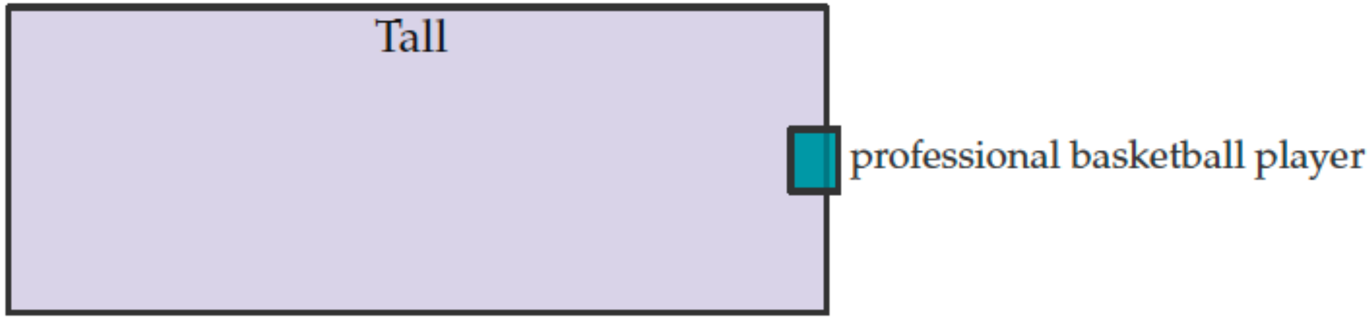
Revisiting some ideas from Monday

- I was NOT encouraging gambling ☺
  - However, it is true that understanding probability theory can help in games of chance

- Best to know how to formally (properly) deal with uncertainty
  - Clinical relevance
  - Sensorimotor control
  - Stock investments
  - Quantum phenomena

Revisiting some ideas from Monday

- The point of the diagnostic example is NOT that you shouldn't trust medical tests – you absolutely should! BUT, interpretation is not always as straight forward as we may think.
  - Prior can play a very large role (e.g., signs and symptoms of a dx plus a positive diagnostic result is very different than a positive result during a random screening)
- P(A|B) does NOT necessarily equal P(B|A)
  - The probability of testing positive given you have the disease is NOT necessarily equal to the probability of having the disease given that you test positive



$p(\text{tall}|\text{professional basketball player}) > p(\text{professional basketball player}|\text{tall})$

Figure 1.2: $p(A|B)$ does not in general equal $p(B|A)$. The area of each rectangle represents the probability of the event in the overall relevant set (here, say, all human beings). The overlap of the two rectangles represents $p(A,B)$. The sizes of the rectangles and their overlap here are conceptual and not calibrated against actual basketball participation data.

Revisiting some ideas from Monday

- Play with the prior (prevalence) in the diagnostic example to see how it changes your posterior

- Consider a prior of 0.8 (e.g., a person exhibits signs and symptoms consistent with the disease process) – now what is your posterior belief?
  - Bear in mind that the disease prevalence was serving as our prior in this example because we had no other information; however, if we were a medical professional examining a patient, our prior would incorporate signs and symptoms and not only prevalence
  - The point is that the prior incorporates our knowledge about the inferential question other than the measurement/observation/data

Inference: Is that my friend?

- The likelihood of a hypothesis is the probability of the observation *given* the hypothesized world state
- The prior is the probability of the hypothesized world state based on all knowledge you have apart from (*prior to*) the observation
- The posterior is the probability of the hypothesized world state given the observation
- Write probability statements on board

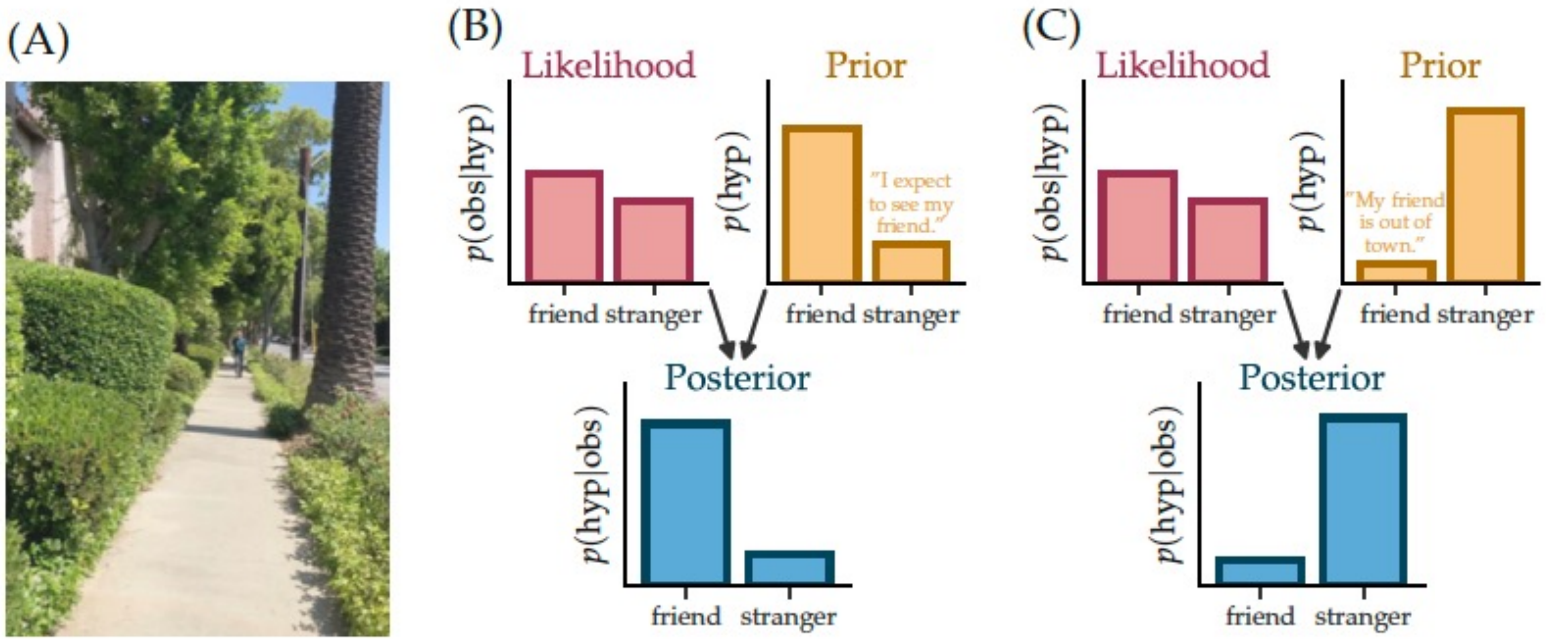


Figure 1.4: Recognizing a friend. **(A)** A visual scene offers a low-resolution view of a person in the distance who resembles your friend. **(B)** You consider the probability that the visual image would result from your friend to be greater than the probability that it would result from a stranger (likelihood function). You expected to meet your friend at this time and place (prior distribution). Therefore, you believe the person in question is probably your friend (posterior distribution). **(C)** In this alternate scenario, you thought your friend was out of town, so your prior distribution sharply favors the *stranger* hypothesis. Given the same observation (likelihood function), you conclude that the person in question is probably not your friend.

The denominator in Bayes' Rule

(joint probability)

	D	ND
+	P(+, D)	P(+, ND)
-	P(-, D)	P(-, ND)

The denominator in Bayes' Rule

(joint probability)

	D	ND
+	P(+, D)	P(+, ND)
-	P(-, D)	P(-, ND)

$P(A,B) = P(A|B)P(B)$

	D	ND
+	$P(+|D)P(D)$	$P(+|ND)P(ND)$
-	$P(-|D)P(D)$	$P(-|ND)P(ND)$

The denominator in Bayes’ Rule

(joint probability)

	D	ND
+	P(+, D)	P(+, ND)
-	P(-, D)	P(-, ND)

P(A,B) = P(A|B)P(B)

	D	ND
+	P(+|D)P(D)	P(+|ND)P(ND)
-	P(-|D)P(D)	P(-|ND)P(ND)

Marginal probabilities

	D	ND	
+	0.99 * 0.01	0.05 * 0.99	0.0594
-	0.01 * 0.01	0.95 * 0.99	0.9406
	0.01	0.99	

The denominator in Bayes' Rule

- Don't sweat too much for this class – much more important to understand prior, likelihood, and posterior
- Oftentimes, you don't need to compute the denominator (lecture notebook has examples)
- In diagnostic example, you do need to compute because we're NOT dealing with probability distributions (which can be normalized)