Graphical Models

Data 102 Fall 2022

Lecture 7

Weekly Overview

- So far:
- **Today:**
  - Making decisions with feedback (online decision-making)
  - Hypothesis testing with a known alternative (Neyman-Pearson)
  - Connection to binary classification
- Next time: connecting decision-making and frequentist/Bayesian views

Recap: Statistical modeling

- Goal: find unknown parameter θ using observed data x
- How:
  - Define a probability model for the data/parameters, then use it to estimate θ from x
- Likelihood function p(x|θ): captures how likely our data are for each parameter
  - Used in both frequentist and Bayesian models
- Frequentist modeling
  - MLE (Maximum Likelihood Estimate): value of θ that makes p(x|θ) as large as possible
- Bayesian modeling
  - Define a prior p(θ): what we believe about the parameter before we see any data
  - Compute posterior p(θ|x): what we believe about the parameter after observing data
  - To get a single estimate for θ from the posterior, we can use the MAP or LMSE estimates
  - MAP: value of θ that makes p(θ|x) as large as possible
  - *MMSE:* $E_{\theta|x}[\theta]$ (*expectation of θ according to posterior p(θ|x))*

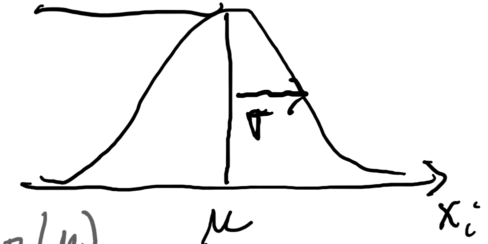
Estimating the Mean with a Gaussian Likelihood

Heights: $x_1, \dots, x_n$

Want: μ (pop. mean)

Normal likelihood $p(x_i \mid \mu) = \mathcal{N}(x_i; \mu, \sigma)$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$



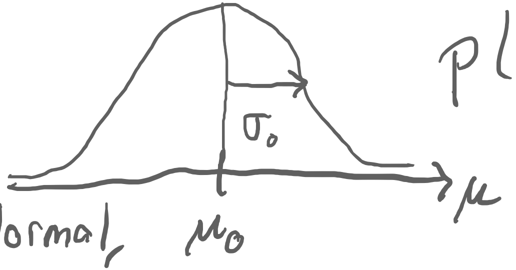
Freq: $\hat{\mu}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Bayesian: $p(\mu \mid x) = \frac{p(x \mid \mu)\, p(\mu)}{p(x)}$

$$\propto p(x \mid \mu)\, p(\mu)$$

can show (skipped) $\propto e^{-\frac{1}{\cdots}(\mu - \cdots)^2}$

Choose prior $p(\mu)$: Normal



$$p(\mu) = \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2}$$

$$= \mathcal{N}(\mu; \mu_0, \sigma_0)$$

<u>If</u> likelihood & prior are Normal, posterior is also Normal.

$$\mu \mid x_1, \dots, x_n \sim \mathcal{N}(\mu; \_\_\_\_, \_\_\_\_)$$

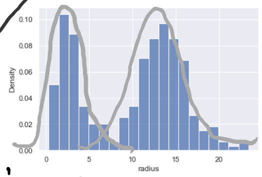
Conjugate Priors

- For some likelihood/prior pairs, the posterior will be in the same family as the prior.

$$p(\theta|x) = \frac{p(x|\theta)\cdot p(\theta)}{p(x)}$$

Likelihood	Prior	Posterior
Bernoulli	Beta	Beta
Normal	Normal	Normal
~~Geometric~~	~~Normal~~	~~???~~

Exoplanet Model



Have: $x_1, \dots, x_n$ (radii of exoplanets)

Want:

- Mean radius for each "group"
- Which planets are in each grp?

is planet i habitable?

z_i : is planet i habitable?

fixed

μ_0, σ_0 : mean/SD for "habitable"

μ_1, σ_1 : " " for "gas giant"

Posterior

$$p(\underbrace{z_1, \dots, z_n, \mu_0, \mu_1}_{\theta = (z_1, \dots, z_n, \mu_0, \mu_1)} \mid x_1, \dots, x_n)$$

Assumptions

- Planets are iid (given grp means)
- σ_0, σ_1 fixed
- Each group's radii normally dist.
- Prior means come from same dist.
  ↳ choose σ_p large

z_i is $\begin{cases} 1 & \text{if exoplanet is ish...} \\ 0 & \text{o.w.} \end{cases}$

$$\left.\begin{aligned} z_i &\overset{iid}{\sim} \text{Bernoulli}(\pi) \\ \mu_0 &\sim N(\mu_p, \sigma_p) \\ \mu_1 &\sim N(\mu_p, \sigma_p) \end{aligned}\right\} \text{priors}$$

$$\underbrace{x_i \mid z_i, \mu_0, \mu_1 \sim N(\mu_{(z_i)}, \sigma_{(z_i)})}_{\text{likelihood}}$$

Exoplanet Model: A Visual Representation

$$z_i \sim \mathrm{Bernoulli}(\pi)$$

$$\mu_k \sim \mathcal{N}(\mu_p, \sigma_p) \quad k=0, k=1$$

$$x_i | z_i, \mu_0, \mu_1 \sim \mathcal{N}(\mu_{z_i}, \sigma_{z_i})$$

Graphical Models

- Circles for RUs
- Dots for params
- Arrows indicate dependence