L5: Bayesian Inference

Jan 21, 2026

FoDA

Jeff M. Phillips

Recall Bayes' Rule

compute = likelihood

given = prior

$$\Pr(M \mid D) = \frac{P(D \mid M) \cdot \Pr(M)}{\Pr(D)}$$

often ignore

- Cracked windshield
  - Discrete models Ω_M

- Height

$$N_{\mu,\sigma} = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right)$$

  - Continuous model Ω_M

only likelihood, no prior

Is Normal

$$\underline{Pr(M \mid D)} \propto \underline{Pr(D \mid M)} \cdot \underline{Pr(M)}$$

which model is most likely given data observed.

\propto

$\propto$ = proportional to

$h(x) \propto C \cdot g(x)$

if $h(x) = C \cdot g(x) \quad \forall x$

for some constant C.

Continuous

$$P(m \mid D) \propto f(D \mid m) \cdot \pi(m)$$

posterior likelihood prior

unknown distribution

when $D = \{x_1, x_2, \dots, x_n\}$ $\quad x_i \sim Q$ iid

$$f(D \mid m) = \prod_{i=1}^{n} g(x_i \mid m)$$

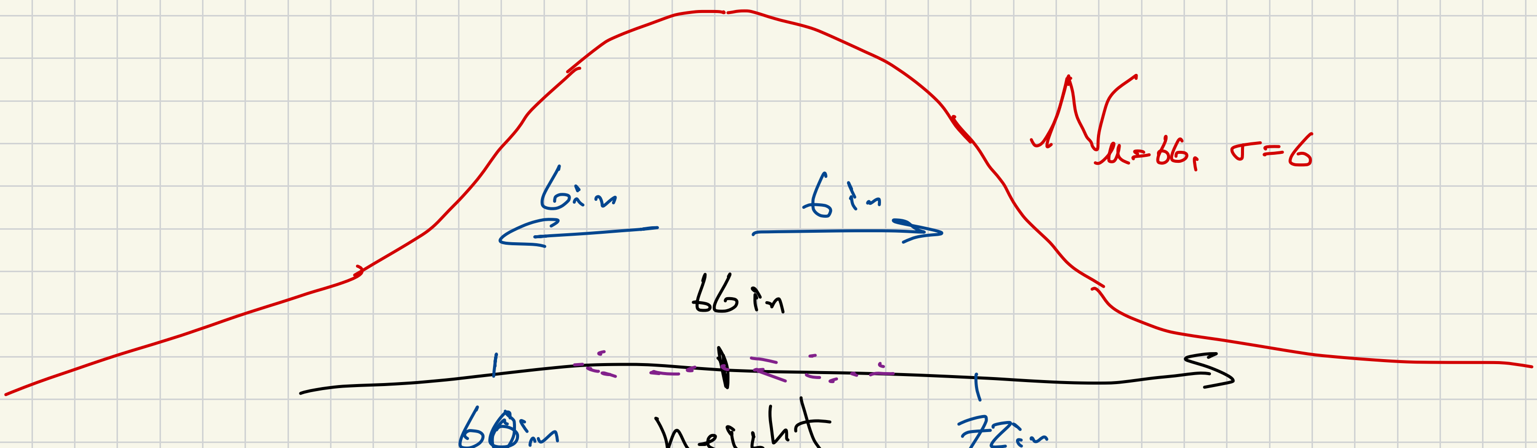
independent

Prior $\pi(m) = \mathcal{N}_{66,6}$ roughly 66 inches

~ ± 6 inches

$$\pi(m) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{(\mu - m)^2}{\sigma^2} \right)$$

guess average height of UofU students.



Data + prior $\pi(m) = \mathcal{N}_{66, 6}$

inches

likelihood $\rightarrow$ maximum likelihood estimator MLE

$D = \{x_1 \dots x_n\} \in \mathbb{R}$ (Normal) noise

$$MLE = \frac{1}{n} \sum_{i=1}^{n} x_i = 5.5 \; ?$$

freq. $\Rightarrow$ 66 inches

Prior $\pi(m) = N_{66,6}$

68 72 63

observe data $D = \{x_1, x_2, \dots x_n\}$ independent

likelihood $f(D \mid m) = \prod_{i=1}^{n} g(x_i \mid m)$

$g(x \mid m) = N_{m,\sigma=2}$

posterior $p(m \mid D) \propto f(D \mid m) \cdot \pi(m)$

$\propto \left[\prod_{i=1}^{n} g(x_i \mid m)\right] \cdot \pi(m)$

$$\ln(p(m \mid D)) \propto \ln\left(\left[\prod_{i=1}^{n} g(x_i \mid m)\right] \cdot \pi(m)\right) + C$$

log-posterior

$$\propto \sum_{i=1}^{n} \ln(g(x_i \mid m)) + \ln(\pi(m)) + C$$

$$\propto \sum_{i=1}^{n} \left[\frac{-(x_i - m)^2}{2^2} + \cancel{\ln\left(\frac{1}{\sqrt{2\pi} \cdot 2}\right)}\right] + \frac{-(66-m)^2}{6^2} + C$$

$$\propto -\left(\sum_{i=1}^{n} (x_i - m)^2 + \frac{1}{9}(66-m)^2\right) + C$$

goal MAP

$$m^* = \underset{m \in \Omega_M}{\arg\min}\ 9\sum_{i=1}^{n} (x_i - m)^2 + (66-m)^2$$

1. weighted average
2. make 9 copies each data x_i

Change prior $N_{66,\ \sigma=0.01}$

MAP

$\rightarrow m^* = \underset{m}{\arg\min} \sum_{i=1}^{n} (x_i - m)^2 + 4G(66 - m)^2$