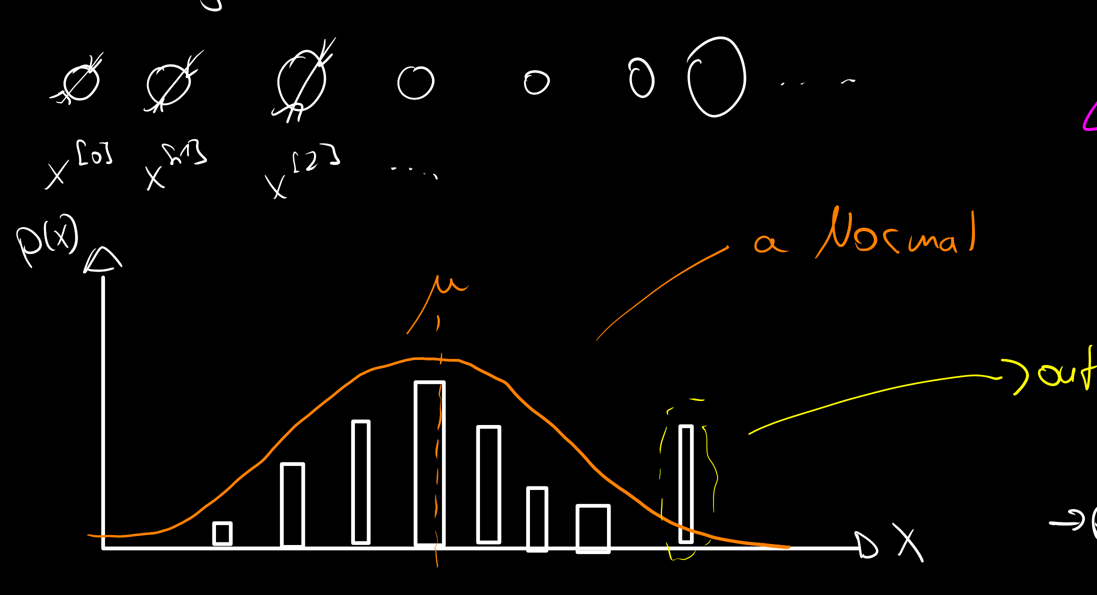
Posterior & MAP for (univariate) Normal with unknown precision

drilling holes



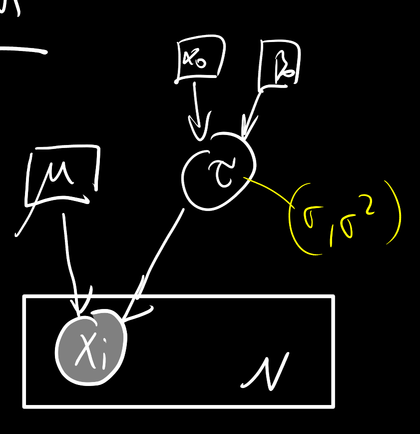
fix μ ... you know how wide the holes are

↳ σ, σ^2, τ unknown

→ MLE?

→ robuster estimate?
→ prior knowledge

DGM



$$\tau \sim \text{Gamma}(\tau; \alpha_0, \beta_0)$$

$$X_i \sim \mathcal{N}(X_i \mid \mu, \tau^{-1})$$

joint

$$p(\tau, \underline{X} = D) = p(\tau)\, p(\underline{X} = D \mid \tau)$$

$$p(\tau, D) \overset{\text{i.i.d.}}{=\!=} p(\tau) \cdot \prod_{i=0}^{N-1} p(X_i = x^{[i]} \mid \tau)$$

$$= \text{Gamma}(\tau; \alpha_0, \beta_0) \prod_{i=0}^{N-1} \mathcal{N}(X_i = x^{[i]} \mid \mu, \tau^{-1})$$

$$= \cancel{\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}}\, \tau^{\alpha_0 - 1} e^{-\beta_0 \tau} \prod_{i=0}^{N-1} \sqrt{\frac{\tau}{\cancel{2\pi}}}\, \exp\left(-\frac{\tau}{2}(x^{[i]} - \mu)^2\right) \qquad \tau^{1/2}$$

$$\sim \tau^{\alpha_0 - 1} e^{-\beta_0 \tau}\; \tau^{N/2} \exp\left(-\frac{\tau}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2\right)$$

Posterior

$$p(\tau \mid D) \overset{\text{Bayes' Rule}}{=\!=} \frac{p(D \mid \tau)\, p(\tau)}{p(D)} \sim p(D \mid \tau)\, p(\tau) = p(\tau, D)$$

Conditional is proportional to the joint

$$p(\tau \mid D) \sim \tau^{\alpha_0 - 1}\, e^{-\beta_0 \tau}\, \tau^{N/2} \exp\left(-\frac{\tau}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2\right)$$

$$= \tau^{\alpha_0 + N/2 - 1} \exp\left(-\beta_0 \tau - \frac{\tau}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2\right)$$

$$= \tau^{\underbrace{\alpha_0 + N/2}_{\alpha_N} - 1} \exp\left(-\underbrace{\left(\beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2\right)}_{\beta_N} \tau\right)$$

it's another Gamma

$$\Rightarrow p(\tau \mid D) = \text{Gamma}(\tau; \alpha_N, \beta_N)$$

$$\boxed{\alpha_N = \alpha_0 + \frac{N}{2}}$$

$$\boxed{\beta_N = \beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2}$$

$$\sigma^2_{MLE} = \frac{1}{N} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2 = \frac{1}{\tau_{MLE}}$$

$$\beta_N = \beta_0 + \frac{N}{2} \sigma^2_{MLE}$$

$$= \beta_0 + \frac{N}{2 \tau_{MLE}}$$

MAP-Estimate

→ Mode of the posterior

$$\tau_{MAP} = \underset{\tau > 0}{\text{argmax}} \left(p(\tau \mid D)\right)$$

take derivative and set to zero

$$\frac{\partial p(\tau \mid D)}{\partial \tau} = \left(\alpha_0 + \frac{N}{2} - 1\right) \cancel{\tau^{\alpha_0 + \frac{N}{2} - 2}} \cancel{\exp\left(-\left(\beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2\right) \tau\right)} \qquad \tau^{-1}$$

$$- \cancel{\tau^{\alpha_0 + \frac{N}{2} - 1}} \left(\beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2\right)$$

$$\cdot \cancel{\exp\left(-\left(\beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2\right) \tau\right)}$$

$$\overset{!}{=} 0$$

$$\Leftrightarrow \left(\alpha_0 + \frac{N}{2} - 1\right) \cdot \frac{1}{\tau} - \beta_0 - \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2 \overset{!}{=} 0$$

$$\left(\alpha_0 + \frac{N}{2} - 1\right) \cdot \frac{1}{\tau} = \beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2$$

$$\boxed{\tau_{MAP} = \frac{\alpha_0 + \frac{N}{2} - 1}{\beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2}}$$

Mode of the posterior gamma

$$\tau_{MAP} = \frac{\alpha_0 + \frac{N}{2} - 1}{\beta_0 + \frac{N}{2} \sigma^2_{MLE}} \qquad \tau := \frac{1}{\sigma^2}$$

$$\boxed{\sigma_{MAP} = \sqrt{\frac{\beta_0 + \frac{1}{2} \sum_{i=0}^{N-1} (x^{[i]} - \mu)^2}{\alpha_0 + \frac{N}{2} - 1}}}$$

"pseudo-count"

high β_0 → high variance, std
low precision

high α_0 → low variance, std
high precision