Math76 Summer 2020

Introduction to Bayesian Computation

Lecture 8:

Univariate Gaussian Distribution 8 July 2020

Distributions so far...

Univariate Discrete Distributions

Name	Sample Space	PMF $m(k)$	Parameters	Commonly Describes
Bernoulli	$\{0, 1\}$	$\{(1-p), p\}$	Probability p	Binary outcomes
Binomial	$\{0, 1, 2, \ldots\}$	$\binom{n}{k} p^k (1-p)^{n-k}$	Sample Size n, probability p.	Multiple binary outcomes.
Poisson	$\{01, 2, , \ldots\}$	$\frac{\lambda^k e^{-k}}{k!}$	Rate $\lambda > 0$	Count data (e.g., people / day)

Univariate Continuous Distributions

Name	Sample Space	PDF $f(x)$	Parameters	Commonly Describes
Uniform	$[a, b]$	$\frac{1}{b-a}$	Bounds a, b	Uninformative priors
Beta	$[0, 1]$	$\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$	Prior sample sizes a, b	unknown probabilities
Gamma	$[0, \infty)$	$\frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}$	Shape k and scale θ	Poisson rates, Gaussian precisions
Gaussian	$(-\infty, \infty)$	$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	Mean μ, Variance σ^2.	Averages, observation error, approximates other densities

See Table A.1 of "Bayesian Data Analysis" by Gelman et al. for a more complete list.

Gauss

How to combine multiple measurements?

1. There is some true value, we just don't know it.
2. Small errors are more likely than large errors.
3. An error of $-\epsilon$ is as likely as ϵ.
4. With several measurements of the same quality, the most likely value is their average. median for Laplace

*Image courtesy of Wikipedia

Estimate location of Mars @ particular time

x_1, x_2, x_3

$x_i - \mu = e_i$

μ ↑ truth

$$\frac{\phi'(e)}{\phi(e)} = -ke$$

$$\phi'(e) = -ke\,\phi(e)$$

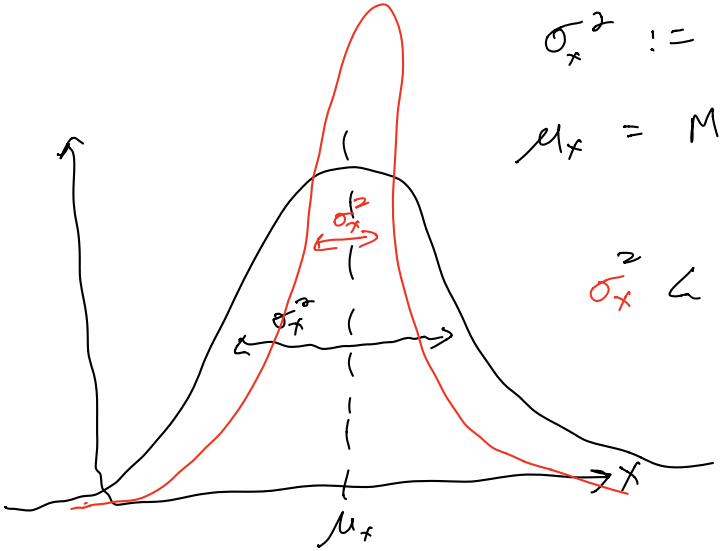
$$\phi(e) = \exp\left(-\tfrac{1}{2}ke^2\right)$$

$$\phi'(e) = -ke\exp\left(-\tfrac{1}{2}ke^2\right)$$

See "The Evolution of the Normal Distribution" by Saul Stahl for more background.

Univariate Gaussian Density

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{1}{2}\frac{(x-\mu_x)^2}{\sigma_x^2}\right]$$



σ_x^2 := Variance

μ_x = Mean

$\sigma_x^2 < \sigma_x^2$

Sums of Gaussian Random Variables

$Y = X_1 + X_2$ is a Gaussian RV, X_1, X_2 are independent

if

X_1 & X_2 are also Gaussian RVs

$$X_1 \sim N(\mu_1, \sigma_1^2)$$

$$X_2 \sim N(\mu_2, \sigma_2^2)$$

$$\mathbb{E}[Y] = \mathbb{E}[X_1 + X_2]$$

$$= \mathbb{E}[X_1] + \mathbb{E}[X_2]$$

$$\mu_Y = \mu_1 + \mu_2$$

$$\sigma_Y^2 = \sigma_1^2 + \sigma_2^2$$

Gaussian Prior and Likelihood

Prior $f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{1}{2}\frac{(x-\mu_x)^2}{\sigma_x^2}\right]$

Noise Model $y = x + \varepsilon$ $\quad \varepsilon \sim N(0, \sigma_\varepsilon^2)$

"noise term"

Posterior

$$f(x|y_{obs}) \propto \underbrace{f(y_{obs}|x)}_{\text{Likelihood Function}} \underbrace{f(x)}_{\text{Prior Density}}$$

Likelihood

$y = x + \varepsilon$

$y - x = \varepsilon$ $\quad f(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left[-\frac{1}{2}\frac{\varepsilon^2}{\sigma_\varepsilon^2}\right]$

$$f(y|x) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left[-\frac{1}{2}\frac{(y-x)^2}{\sigma_\varepsilon^2}\right]$$

$$f(x|y) \propto f(y|x)\, f(x)$$

$$= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{1}{2}\left(\underbrace{\frac{(y-x)^2}{\sigma_\varepsilon^2}}_{\text{likelihood}} + \underbrace{\frac{(x-\mu_x)^2}{\sigma_x^2}}_{\text{prior}}\right)\right]$$

$$\frac{1}{\sigma_\varepsilon^2}(y^2 - 2xy + x^2) + \frac{1}{\sigma_x^2}(x^2 - 2\mu_x x + \mu_x^2)$$

$$x^2\left(\frac{1}{\sigma_\varepsilon^2} + \frac{1}{\sigma_x^2}\right) - 2x\left(\frac{y}{\sigma_\varepsilon^2} + \frac{\mu_x}{\sigma_x^2}\right) + \frac{y^2}{\sigma_\varepsilon^2} + \frac{\mu_x^2}{\sigma_x^2}$$

arbitrary variable w

$$\log f(w) = C - \frac{1}{2}\frac{(w-\mu)^2}{\sigma_w^2} = C - \frac{1}{2}\left(\frac{1}{\sigma_w^2}(w^2 - 2\mu w + \mu^2)\right)$$

$$\frac{1}{\sigma_{post}^2} = \frac{1}{\sigma_\varepsilon^2} + \frac{1}{\sigma_x^2} \Rightarrow \sigma_{post_1}^2 = \frac{1}{\frac{1}{\sigma_\varepsilon^2} + \frac{1}{\sigma_x^2}} = \frac{\sigma_\varepsilon^2\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2}$$

$$\frac{\mu_{post}}{\sigma_{post}^2} = \frac{y}{\sigma_\varepsilon^2} + \frac{\mu_x}{\sigma_x^2}$$

$$\mu_{post} = \sigma_{post}^2\left(\frac{y}{\sigma_\varepsilon^2} + \frac{\mu_x}{\sigma_x^2}\right)$$

$$= \frac{\sigma_\varepsilon^2\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2}\left(\frac{y}{\sigma_\varepsilon^2} + \frac{\mu_x}{\sigma_x^2}\right)$$

$$\mu_{post_1} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2} y + \frac{\sigma_\varepsilon^2}{\sigma_x^2 + \sigma_\varepsilon^2}\mu$$

What happens w/ multiple obs?

$$\frac{1}{\sigma^2_{post,2}} = \frac{1}{\sigma^2_\varepsilon} + \frac{1}{\sigma^2_{post,1}}$$

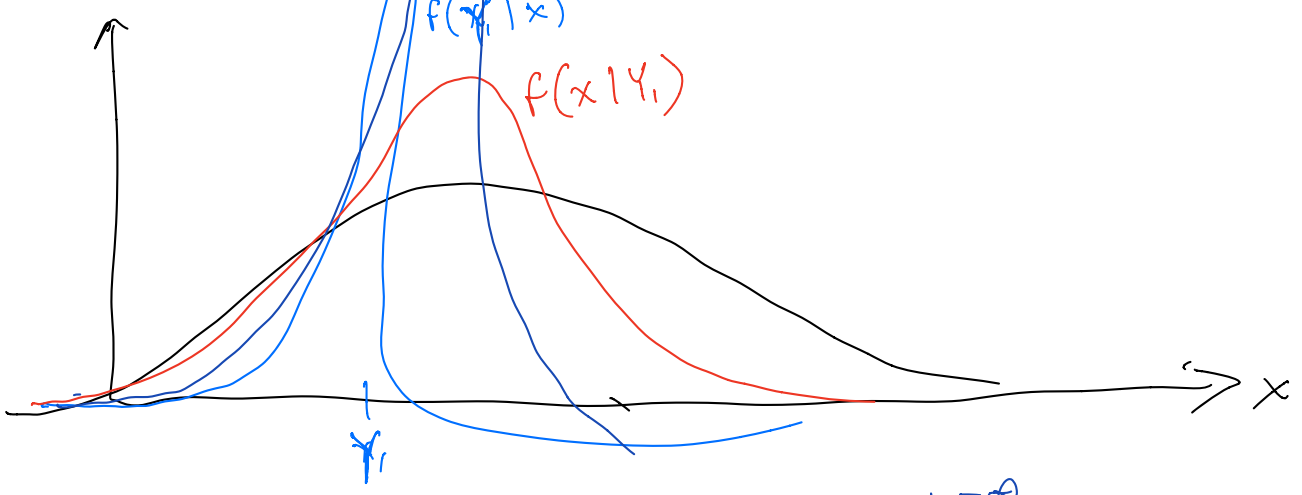
$$= \frac{1}{\sigma^2_\varepsilon} + \frac{1}{\sigma^2_\varepsilon} + \frac{1}{\sigma^2_x}$$

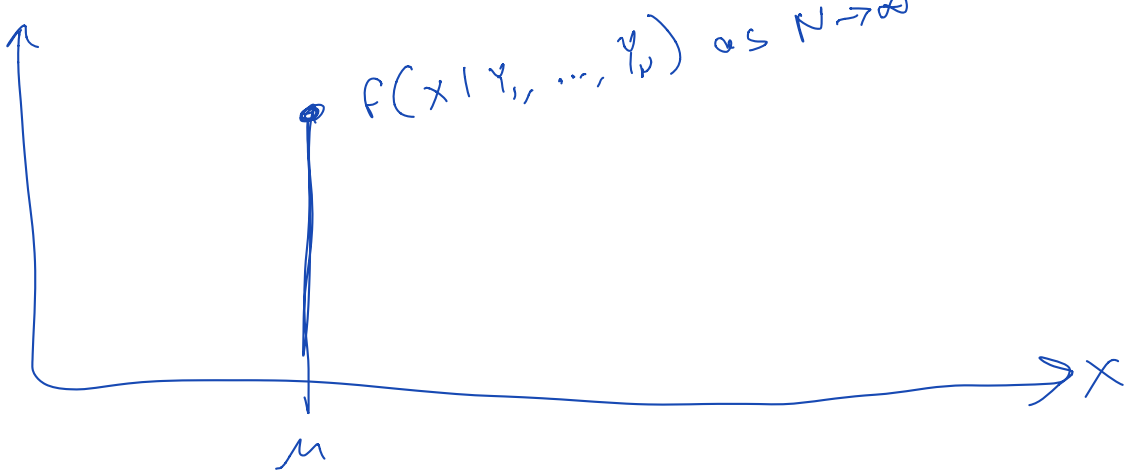
$$\sigma^2_{post,2} = \frac{1}{\frac{1}{\sigma^2_\varepsilon} + \frac{1}{\sigma^2_\varepsilon} + \frac{1}{\sigma^2_x}} = \frac{1}{\frac{2}{\sigma^2_\varepsilon} + \frac{1}{\sigma^2_x}}$$

$$= \frac{\sigma^2_\varepsilon \sigma^2_x}{2\sigma^2_x + \sigma^2_\varepsilon}$$

$$\sigma^2_{post,N} = \frac{\sigma^2_\varepsilon \sigma^2_x}{N\sigma^2_x + \sigma^2_\varepsilon}$$

$N \to \infty \quad \sigma^2_{post,N} \to 0$





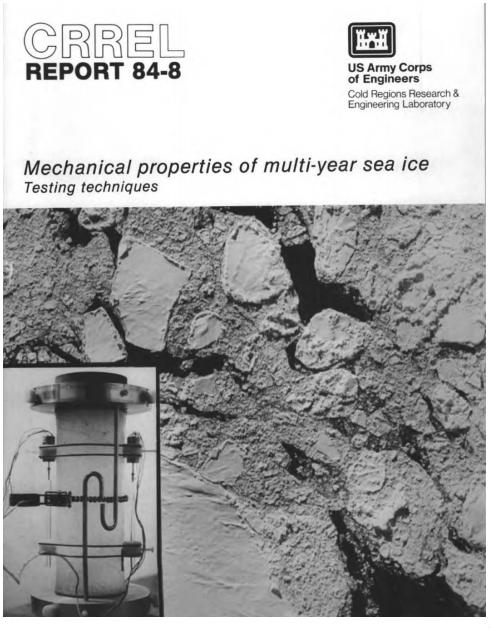
Breakout Exercise: Impact of the Prior

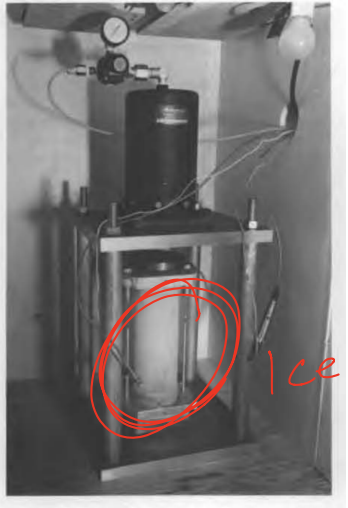
What impact does the prior have on the posterior as we get more observations?

Questions:

1. ~~How does the posterior mean change as N increases?~~
2. How does the posterior variance change as N increases?
3. What would happen if the prior was misleading?

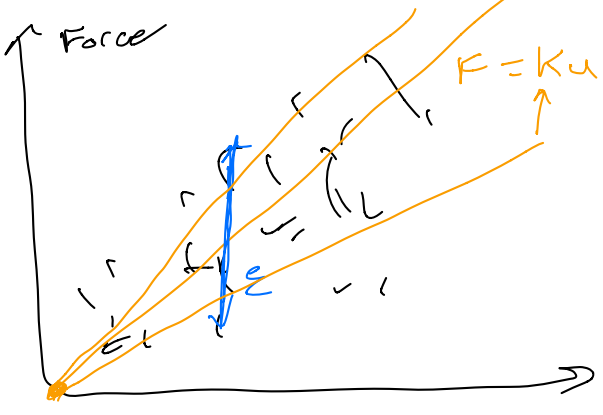
Example: Linear Regression





Ice Specimen

Figure 15. Specimen mounted for compression under constant load.



$K \sim N(\mu_K, \sigma_K^2)$

Example: Linear Regression

$F = ku + \varepsilon \Rightarrow \varepsilon = F_i - ku_i$

$\varepsilon \sim N(0, \sigma_\varepsilon^2)$

$f(k) = N(\mu_k, \sigma_k^2)$

$$\mu = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2} y + \frac{\sigma_\varepsilon^2}{\sigma_x^2 + \sigma_\varepsilon^2} \mu_x$$

$$f(F_i \mid k) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left[-\frac{1}{2} \frac{(F_i - k u_i)^2}{\sigma_\varepsilon^2}\right]$$

$$= \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left[-\frac{1}{2} \frac{u_i^2 (F_i/u_i - k)^2}{\sigma_\varepsilon^2}\right]$$

$$\rightarrow \sigma_2^2 = \frac{\sigma_\varepsilon^2}{u_i^2}$$

$$f(k \mid F_i) = N\left(\frac{\sigma_k^2}{\sigma_k^2 + \frac{\sigma_\varepsilon^2}{u_i^2}} F_i + \frac{\sigma_\varepsilon^2}{u_i^2\left(\sigma_k^2 + \frac{\sigma_\varepsilon^2}{u_i^2}\right)} \mu_k,\right.$$