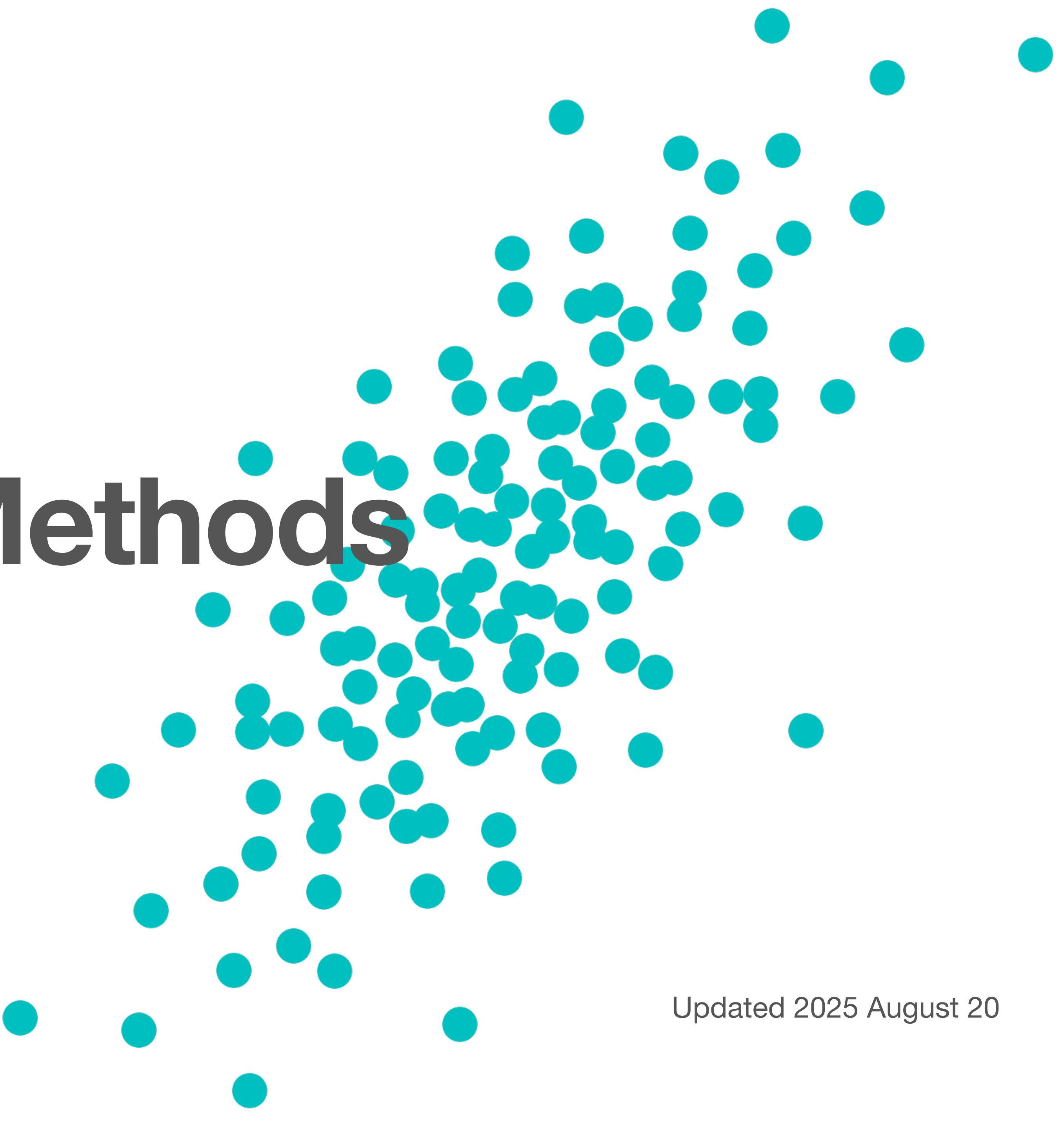


Monte Carlo Methods

MATH 565

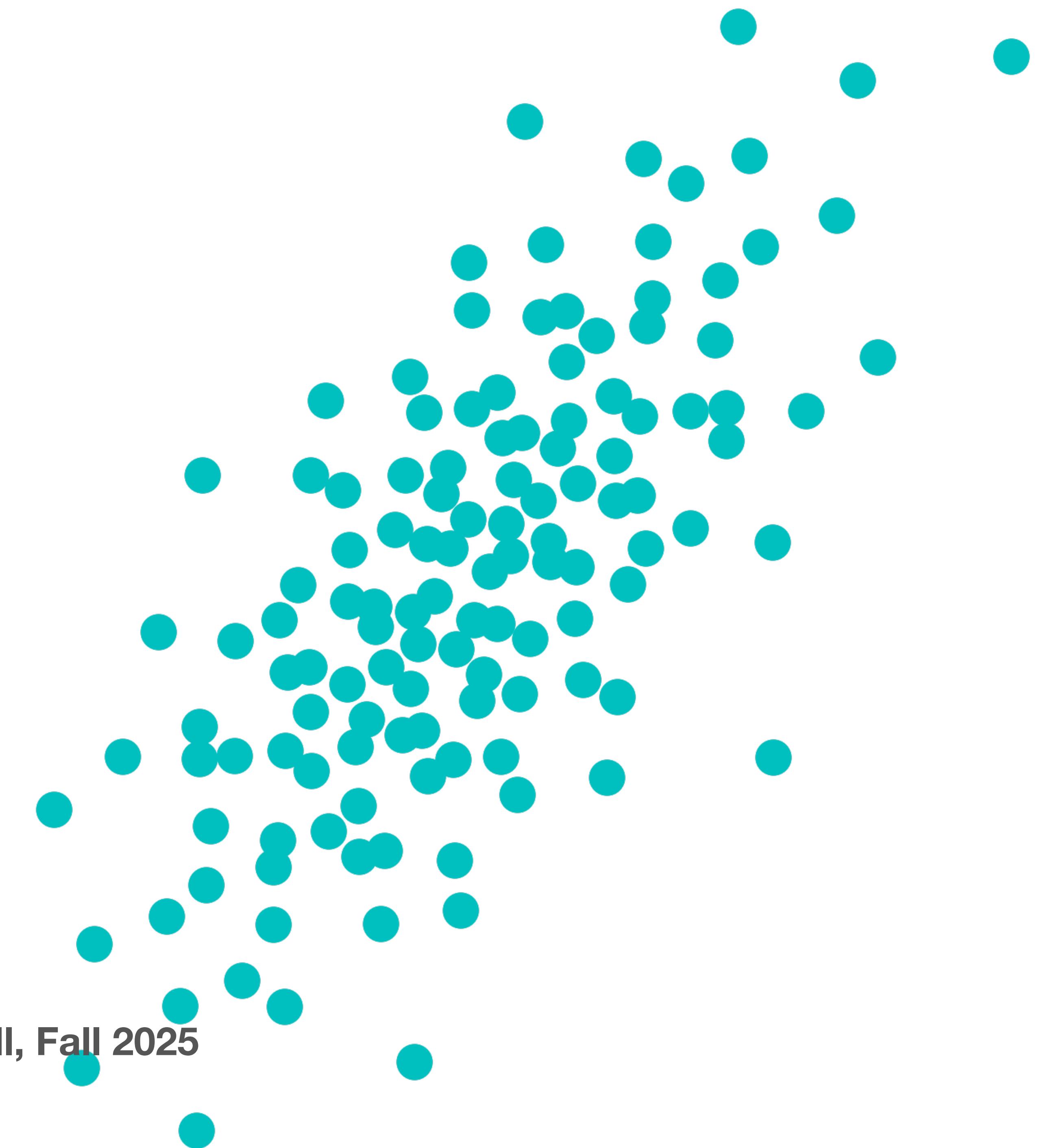
Fred Hickernell, Fall 2025

Updated 2025 August 20

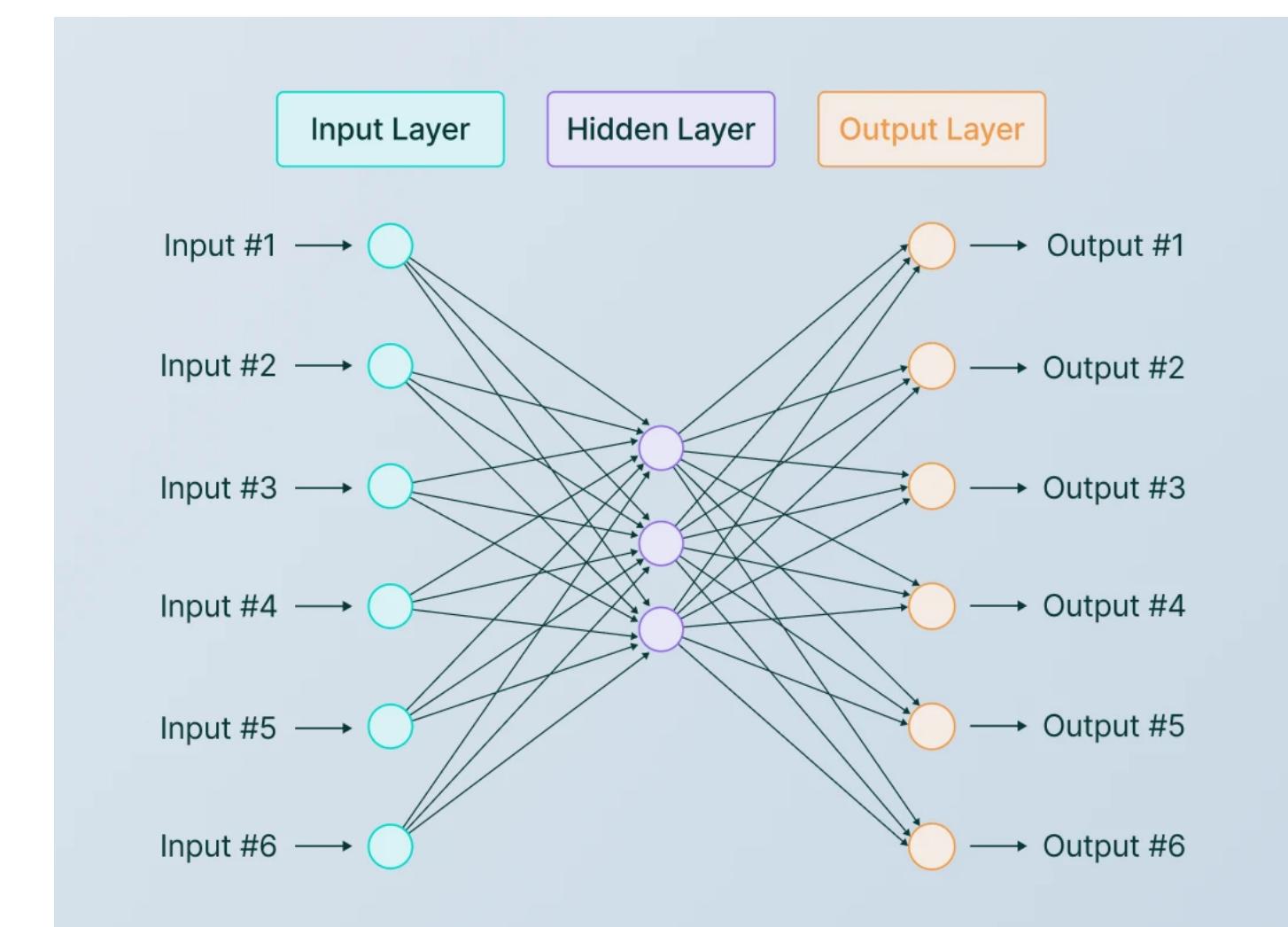
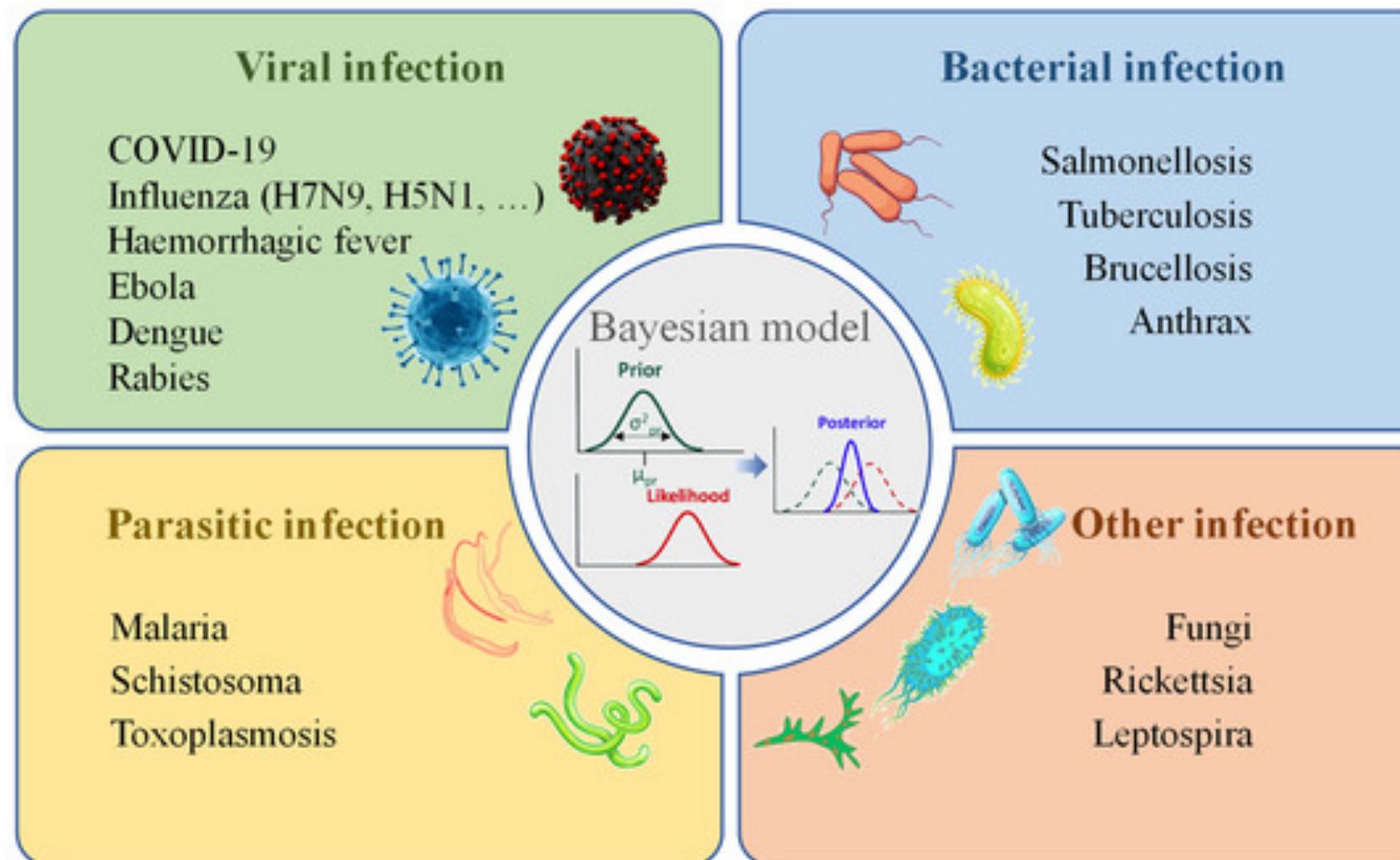
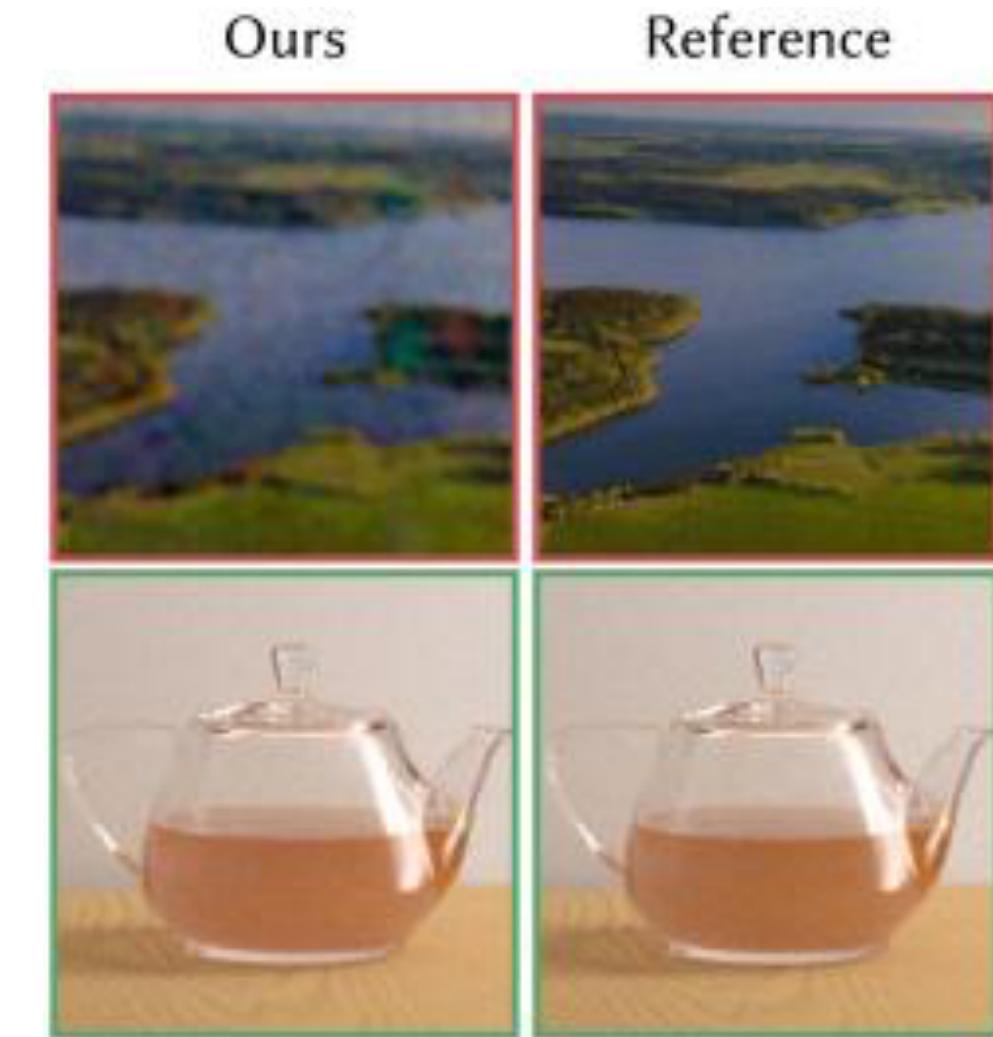
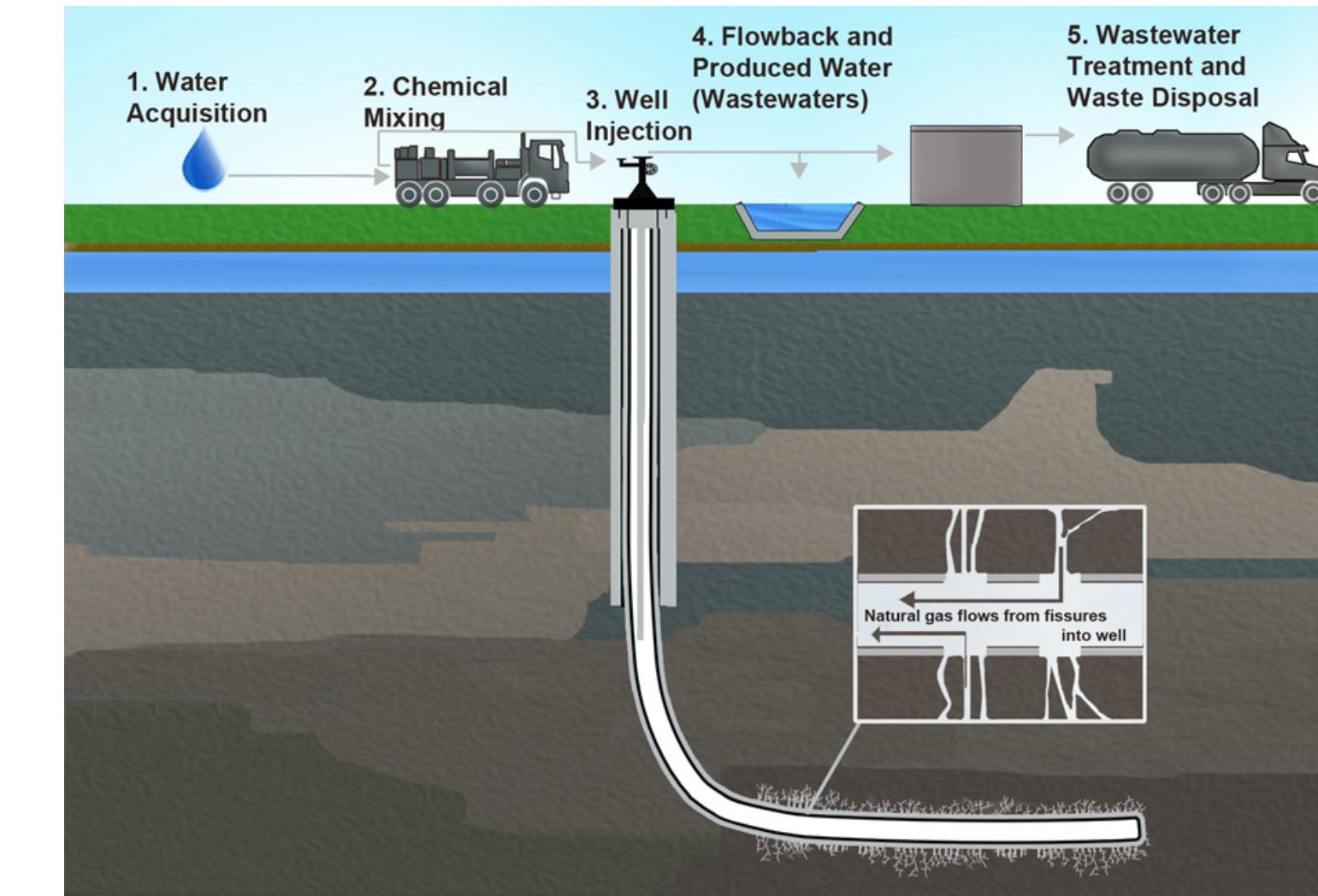


Introduction

MATH 565 Monte Carlo Methods, Fred Hickernell, Fall 2025



Monte Carlo Helps With Uncertainty [wsj 2016]





Why is there uncertainty?



Why is there uncertainty?

- Finance – **market forces**, often modeled by stochastic processes driven by Brownian motion



Why is there uncertainty?

- Finance – **market forces**, often modeled by stochastic processes driven by Brownian motion
- Engineering – **variability of system parameters**, sometimes modeled by Gaussian processes



Why is there uncertainty?

- Finance – **market forces**, often modeled by stochastic processes driven by Brownian motion
- Engineering – **variability of system parameters**, sometimes modeled by Gaussian processes
- Image rendering – **can only trace some rays**, must rely on a finite sample



Why is there uncertainty?

- Finance – **market forces**, often modeled by stochastic processes driven by Brownian motion
- Engineering – **variability of system parameters**, sometimes modeled by Gaussian processes
- Image rendering – **can only trace some rays**, must rely on a finite sample
- Bayesian inference – the **posterior probability distribution** is a combination of a prior and what is learned from data



Why is there uncertainty?

- Finance – **market forces**, often modeled by stochastic processes driven by Brownian motion
- Engineering – **variability of system parameters**, sometimes modeled by Gaussian processes
- Image rendering – **can only trace some rays**, must rely on a finite sample
- Bayesian inference – the **posterior probability distribution** is a combination of a prior and what is learned from data
- Neural networks – many parameters need to be tuned, but one cannot search in **all possible directions**



Why is there uncertainty?

- Finance – **market forces**, often modeled by stochastic processes driven by Brownian motion
- Engineering – **variability of system parameters**, sometimes modeled by Gaussian processes
- Image rendering – **can only trace some rays**, must rely on a finite sample
- Bayesian inference – the **posterior probability distribution** is a combination of a prior and what is learned from data
- Neural networks – many parameters need to be tuned, but one cannot search in **all possible directions**
- Queues – **arrival times and service times** of customers



How is this expressed quantitatively?

Y = random variable denoting **quantity of interest** = $\left\{ \begin{array}{l} \text{option payoff} \\ \text{fluid pressure} \\ \text{pixel intensity} \\ \text{statistical model parameter} \\ \text{neural network parameter} \\ \text{service time} \end{array} \right.$

= $f(\mathbf{X})$, where

\mathbf{X} = multivariate random variable with a **simpler distribution**

We want to estimate the **mean**, **variance**, **quantile**, or **probability distribution** of Y using **sample versions**



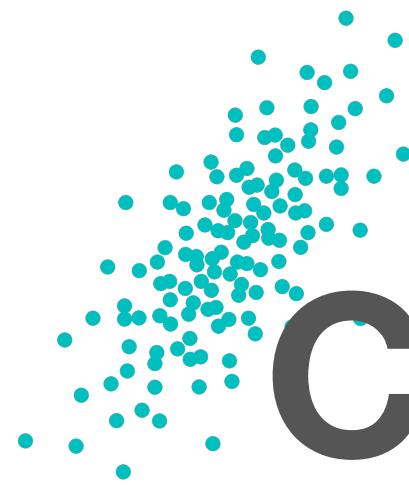
Are we there yet?

You are visiting your friend and it will require

- A **5** minute walk to the ‘L’ station
- Waiting for the train, which arrives every **20** minutes
- Traveling **35** minutes by ‘L’
- Catching a taxi at the ‘L’ destination
 - There is a **20%** chance that the car is waiting for you
 - Otherwise the average wait time is **10** minutes
- A **12** minute taxi ride

How long should you plan for the trip to take?

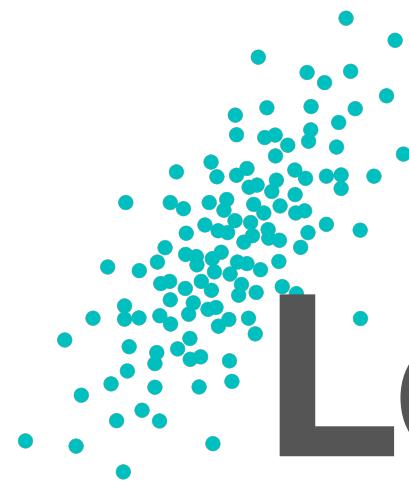
Let’s look at this Jupyter Notebook [AreWeThereYet](#) on the [class website](#)



Class website and repository

Website

Git repository



Let me know your degree program?

- Go to menti.com
- Use code 6222 6078



Why should you attend synchronously?

- You will be better keep pace?
- You will get real time answers to questions?
- You can influence the pace and direction of the course?
- You can help your peers learn and benefit from them—partake in a leaning community
- To help me know you, in case you want a reference, or to add me to your LinkedIn network



Review of probability

Let Y be a **random variable** with a sample space $\mathcal{Y} \subseteq \mathbb{R}$

$\mathbb{P}(Y \in \Omega)$ means the probability of the **event** Ω

F is the **cumulative probability distribution function** $F(y) := \mathbb{P}(Y \leq y)$, $y \in \mathcal{Y}$

Q is the **quantile function** $Q(p) := \inf\{y \in \mathcal{Y} : F(y) \geq p\}$, $0 < p < 1$

$\mathbb{E}[g(Y)] := \int_{\mathcal{Y}} g(y) dF(y)$ is the **expectation** of $g(Y)$

discrete

continuous

ρ is the **probability mass function** ρ is the **probability density function**

$$\rho(y) := \mathbb{P}(Y = y)$$

$$\rho(y) := F'(y)$$

$$\mathbb{E}[g(Y)] = \sum_{y \in \mathcal{Y}} g(y)$$

$$\mathbb{E}[g(Y)] = \int_{y \in \mathcal{Y}} g(y) dy$$



Moments

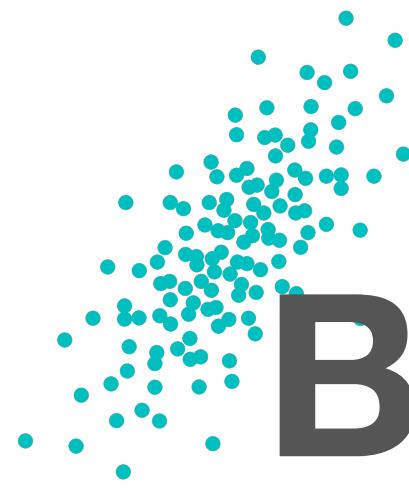
$\mu := \mathbb{E}(Y)$ is the **mean** of Y

$\sigma^2 = \text{var}(Y) := \mathbb{E}[(Y - \mu)^2]$ is the **variance** of Y

σ is the **standard deviation**

$\text{cov}(Y) := (\mathbb{E}[(Y_j - \mu_j)(Y_k - \mu_k)])_{j,k=1}^d$

is the **covariance matrix** of the random **vector** Y



Binomial Distribution

This is a coin-flipping distribution with the probability p of heads

$$Y \sim \text{Binomial}(n, p) \quad \text{scipy.stats.binom(n, p)}$$

$$\varrho(y) := \mathbb{P}(Y = y) = \binom{n}{k} p^y (1-p)^{n-y}, \quad y = 0, 1, \dots, n$$

$$F(y) = \mathbb{P}(Y \leq y) = \sum_{j=0}^{\lfloor y \rfloor} \binom{n}{j} p^j (1-p)^{n-j}$$

$$Q(p) = \inf\{y \in \{0, 1, \dots, n\} : F(y) \geq p\}$$

$$\mu = \mathbb{E}(Y) = np \quad \sigma^2 = \text{var}(Y) = np(1-p)$$



Multivariate Normal/Gaussian Distribution

Used in finance and uncertainty quantification

$Y \sim \mathcal{N}(\mu, \Sigma)$ is a d -dimensional random vector

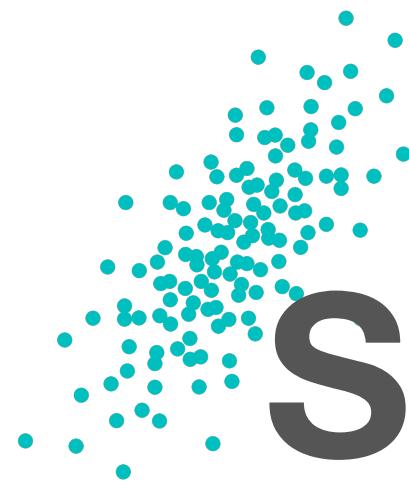
`scipy.stats.multivariate_normal(mean, cov)`

μ = mean

Σ = covariance matrix

$$\varrho(\mathbf{y}) := \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{y} - \mu)^\top \Sigma^{-1} (\mathbf{y} - \mu)\right)$$

$$F(\mathbf{y}) = \int_{(-\infty, \mathbf{y})} \varrho(\mathbf{z}) \, d\mathbf{z}$$



Sampling

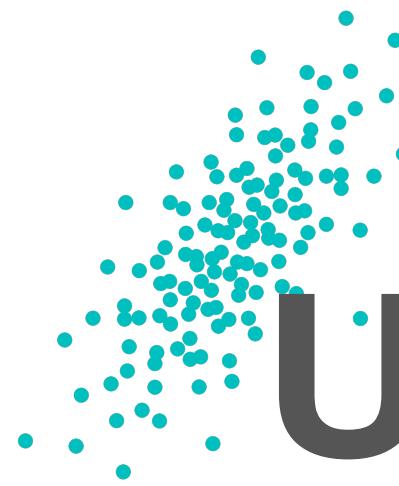
Let Y be a random variable, often $Y = f(X)$

Y_1, Y_2, \dots be a **sample**

not necessarily random or independent and identically distributed (IID)

$\hat{\mu}_n := \frac{1}{n} \sum_{i=1}^n Y_i$ is the **sample mean**

$\hat{\sigma}_n^2 := \frac{1}{n-1} \sum_{i=1}^n (Y_i - \hat{\mu}_n)^2$ is the **sample variance**



Under random sampling

Let Y be a random variable, often $Y = f(X)$

$$Y_1, Y_2, \dots \sim Y$$

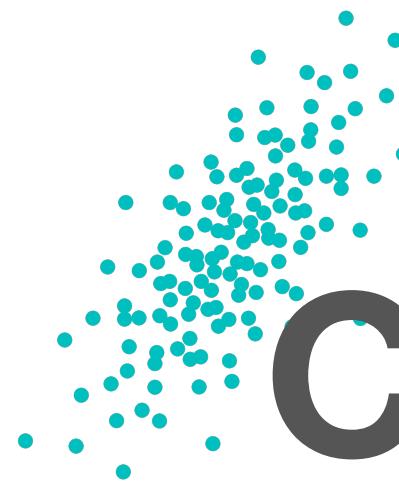
$$\mathbb{E}(\hat{\mu}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(Y_i) = \mu \text{ so } \hat{\mu}_n \text{ is unbiased}$$

If in addition, Y_1, Y_2, \dots are uncorrelated, then

$$\mathbb{E}(\hat{\sigma}_n^2) = \sigma^2 \text{ so } \hat{\sigma}_n^2 \text{ is unbiased}$$

$$\text{mse}(\hat{\mu}_n) := \mathbb{E}[(\mu - \hat{\mu}_n)^2] = \text{var}(\hat{\mu}_n) = \frac{\sigma^2}{n},$$

$$\text{rmse}(\hat{\mu}_n) = \frac{\sigma}{\sqrt{n}}$$



Central Limit Theorem

Let Y be a random variable, often $Y = f(X)$

$$Y_1, Y_2, \dots \stackrel{\text{IID}}{\sim} Y$$

The distribution of

$$\frac{\hat{\mu}_n - \mu}{\sigma/\sqrt{n}} \text{ approaches } \mathcal{N}(0, 1) \text{ as } n \rightarrow \infty$$

which means that

$$\hat{\mu}_n \dot{\sim} \mathcal{N}(\mu, \sigma^2/n) \text{ for large } n$$



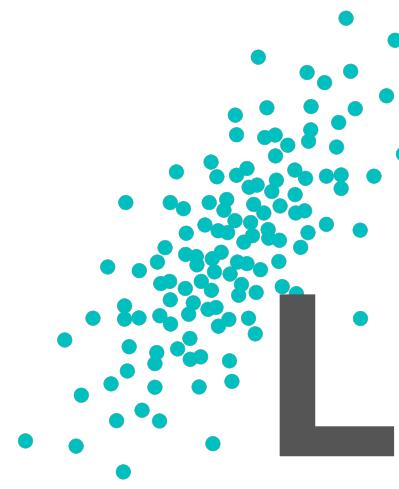
Are we there yet?

You are visiting your friend and it will require

- A **5** minute walk to the ‘L’ station
- Waiting for the train, which arrives every **20** minutes
- Traveling **35** minutes by ‘L’
- Catching a taxi at the ‘L’ destination
 - There is a **20%** chance that the car is waiting for you
 - Otherwise the average wait time is **10** minutes
- A **12** minute taxi ride

How long should you plan for the trip to take?

Let’s look at this Jupyter Notebook [AreWeThereYet](#) on the [class website](#)



Let me know your professional aspirations?

- Go to menti.com
- Use code 6222 6078