

# Monte Carlo Methods

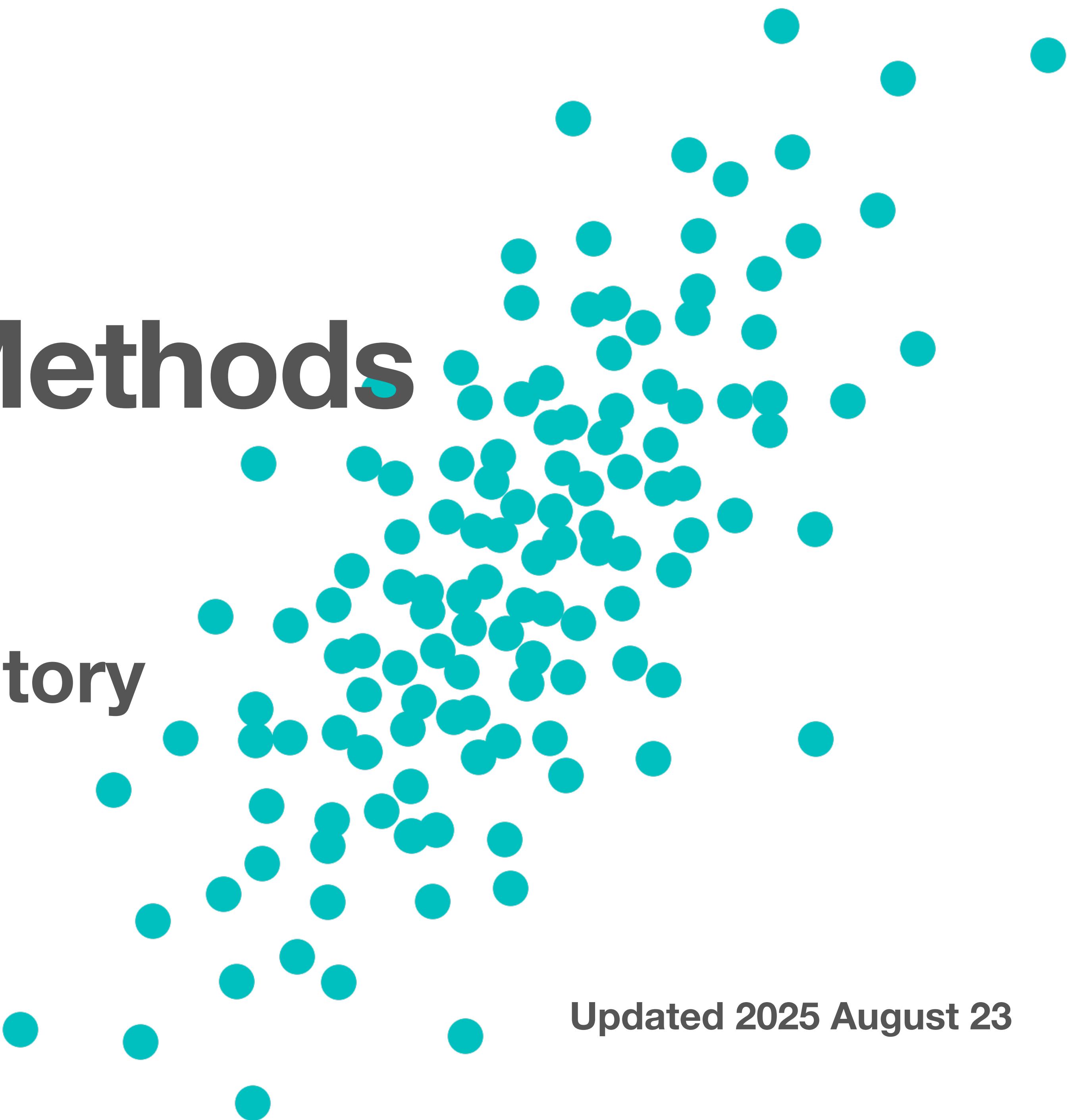
MATH 565

Git website and repository

Canvas

Fred Hickernell, Fall 2025

Updated 2025 August 23

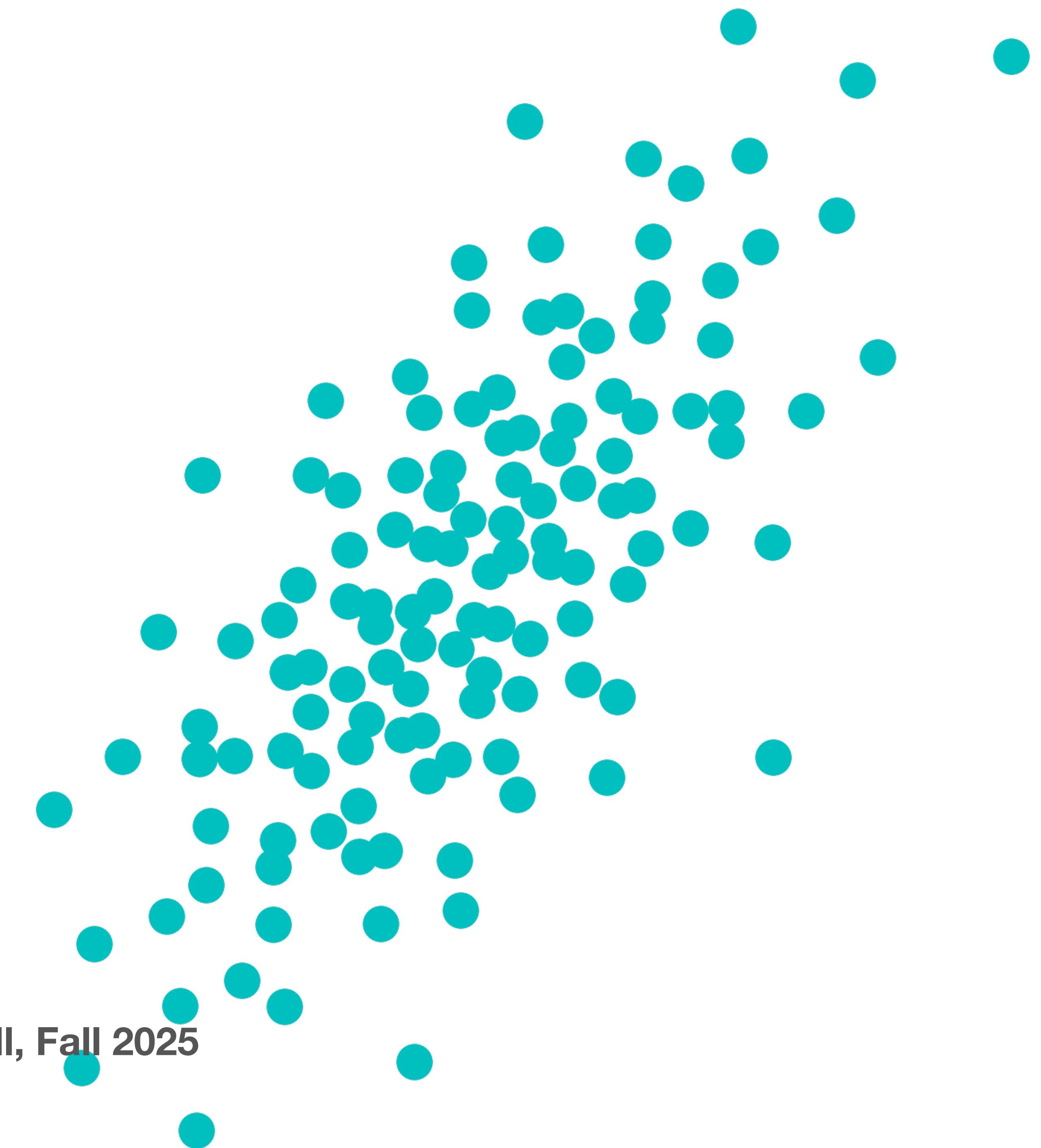


# Introduction

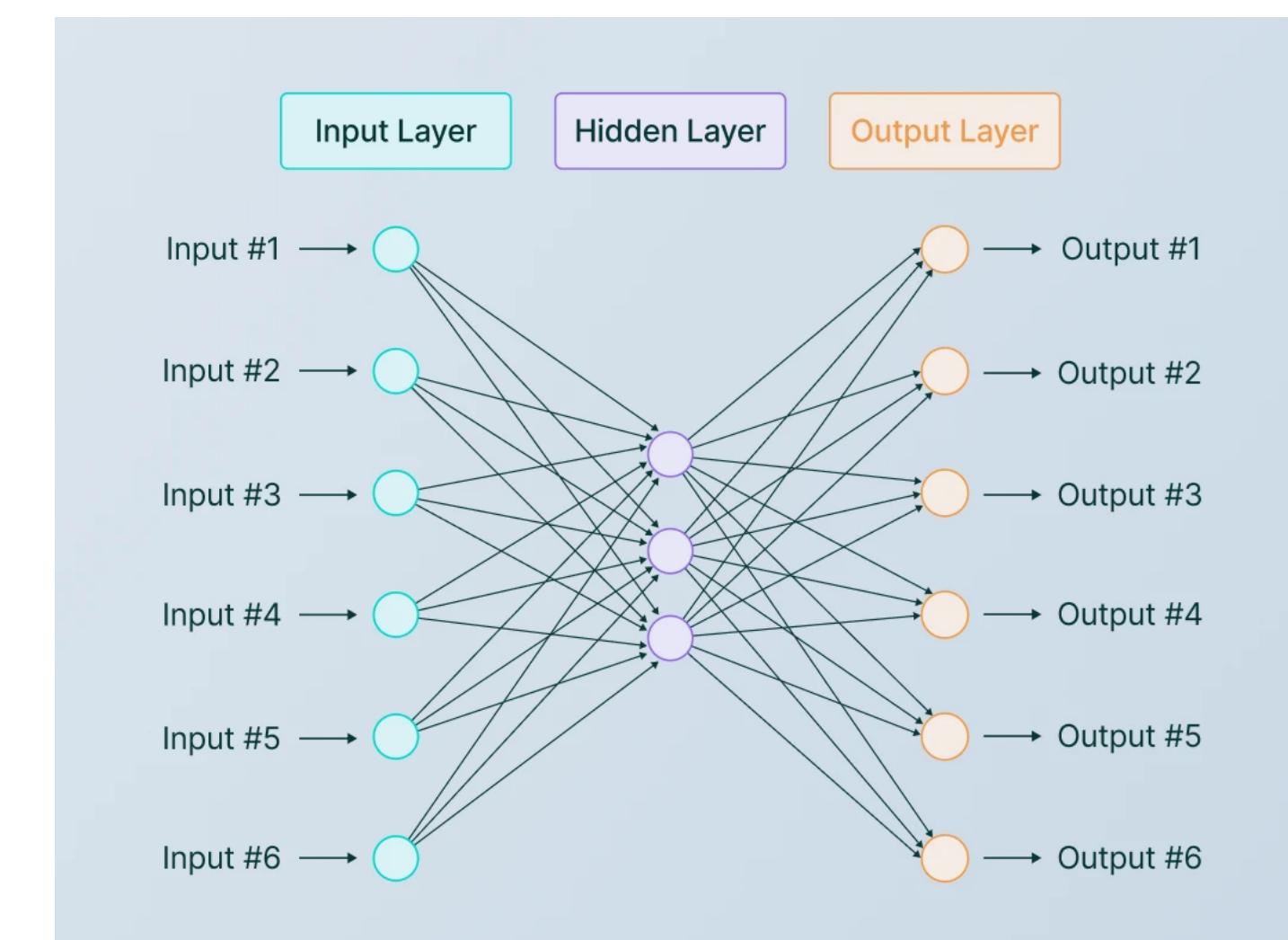
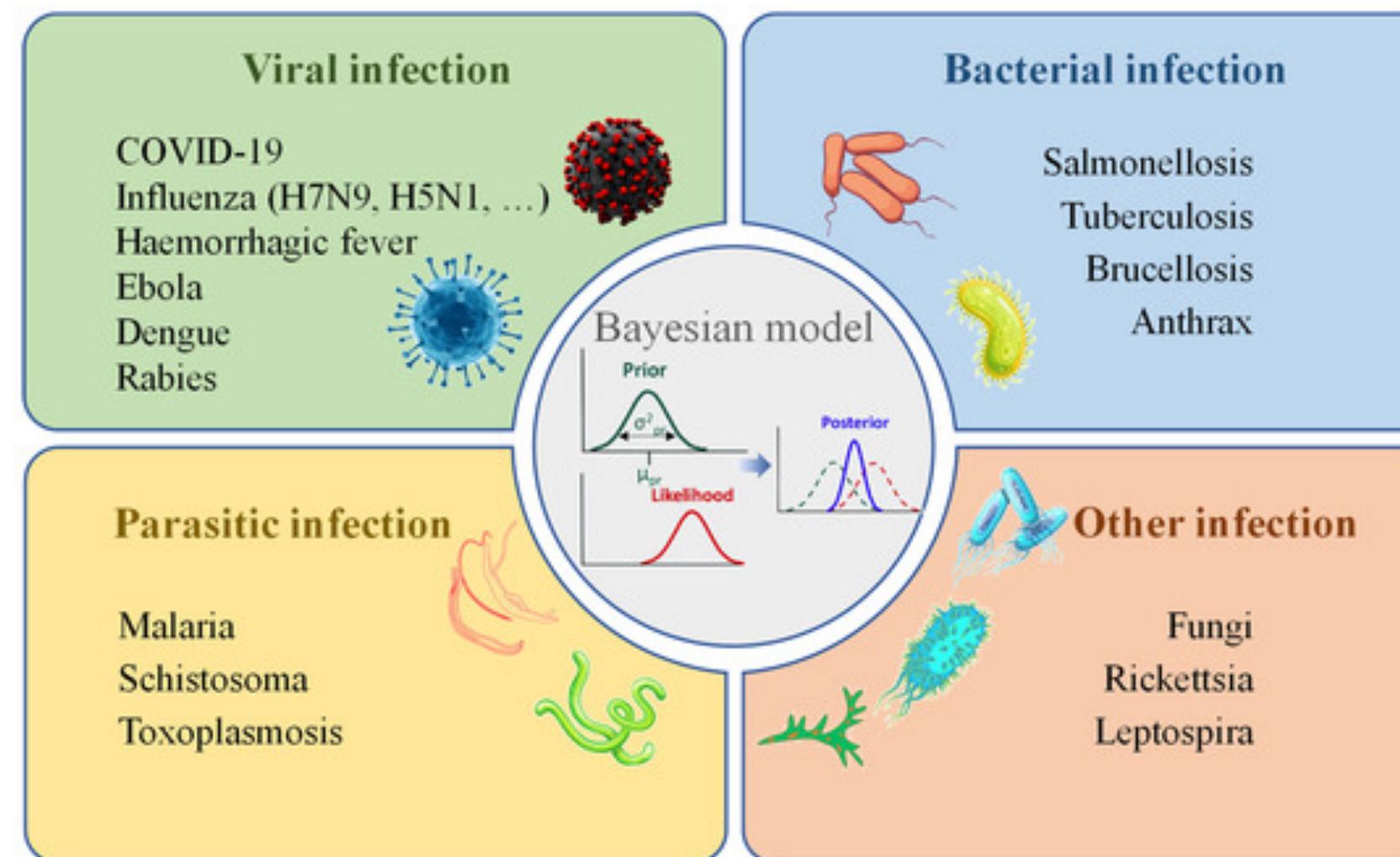
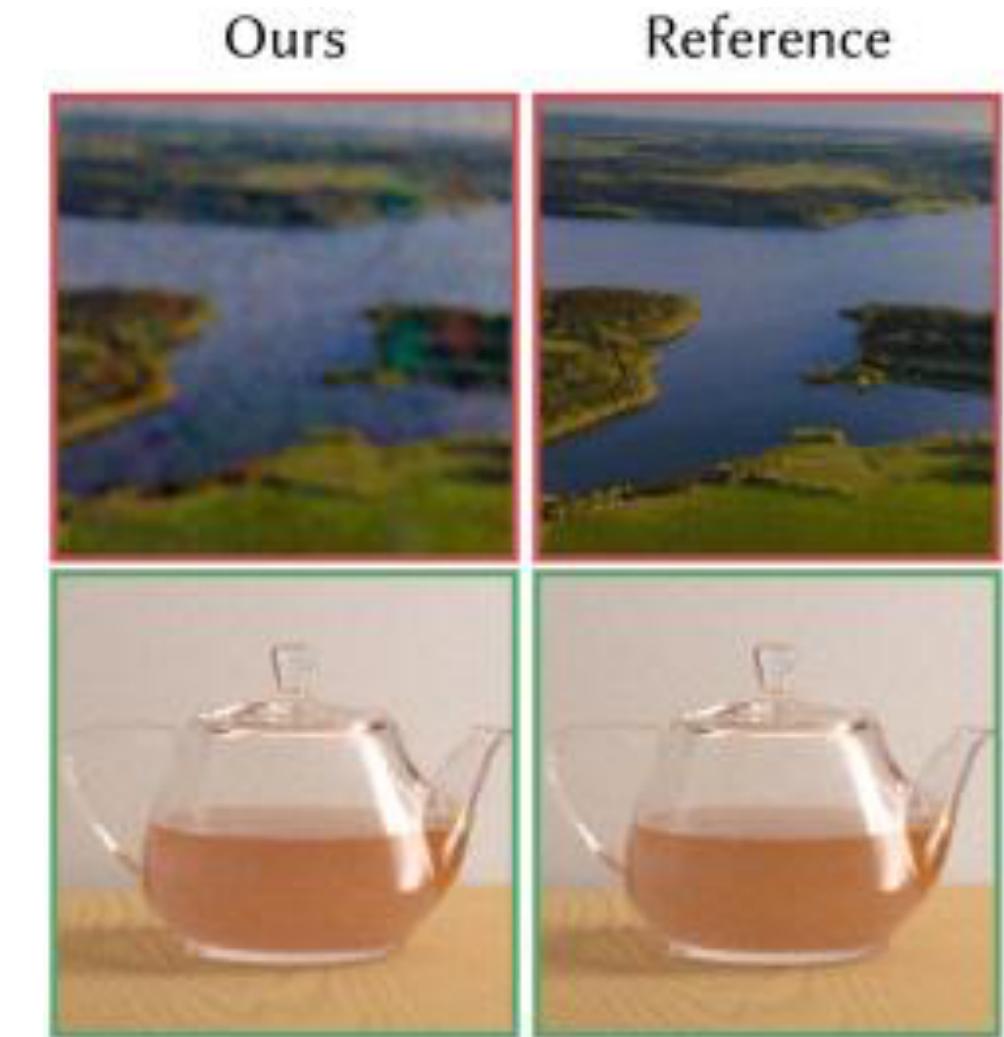
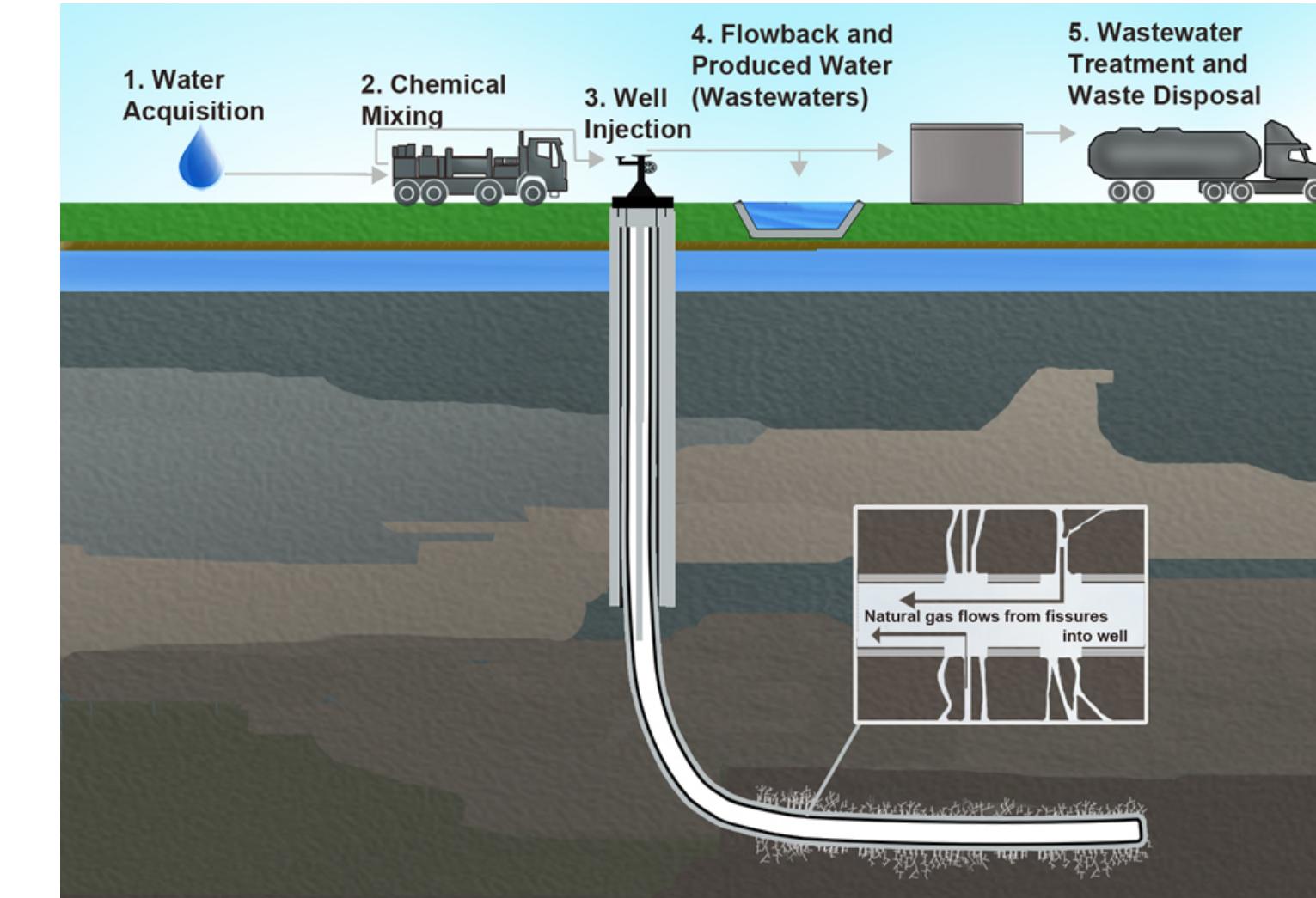
Owen, Chapters 1 & 2

Assignment 1 Due Sep 5

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**



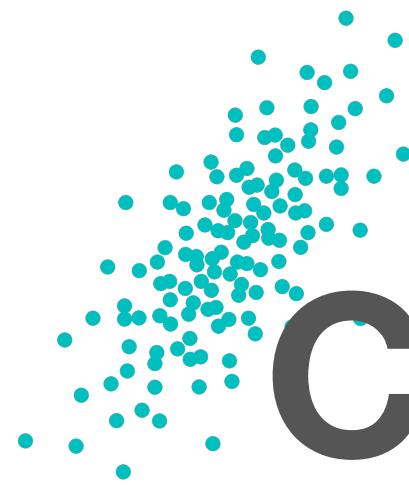
# Are we there yet?

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?

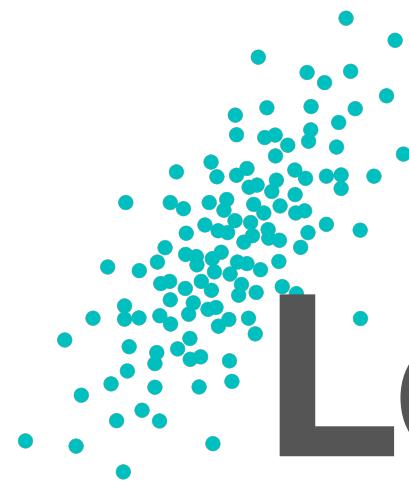
Let’s look at this Jupyter Notebook [AreWeThereYet](#) on the [class website](#)



# Class website and repository

Website

Git repository



# Let me know your degree program?

- Go to [menti.com](https://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**  $\Omega$

$F$  is the **cumulative 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
$\varrho$ is the <b>probability mass function</b> $\varrho(y) := \mathbb{P}(Y = y)$ $\mathbb{E}[g(Y)] = \sum_{y \in \mathcal{Y}} g(y) \varrho(y)$	$\varrho$ is the <b>probability density function</b> $\varrho(y) := F'(y)$ $\mathbb{E}[g(Y)] = \int_{y \in \mathcal{Y}} g(y) \varrho(y) dy$



# Moments

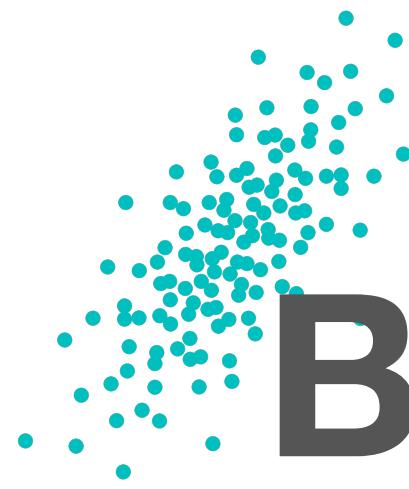
$\mu := \mathbb{E}(Y)$  is the **mean** of  $Y$

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

$\sigma$  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

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

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

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

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

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



# Multivariate Normal/Gaussian Distribution

Used in finance and uncertainty quantification

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

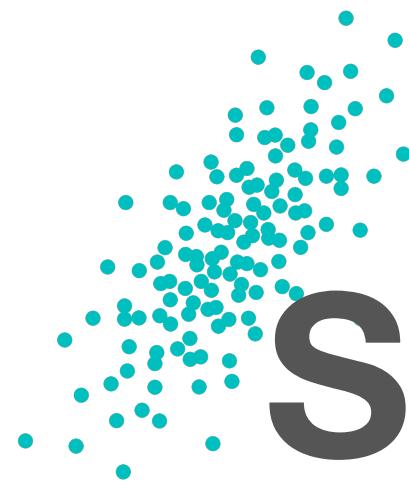
`scipy.stats.multivariate_normal(mean, cov)`

$\mu$  = mean

$\Sigma$  = 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

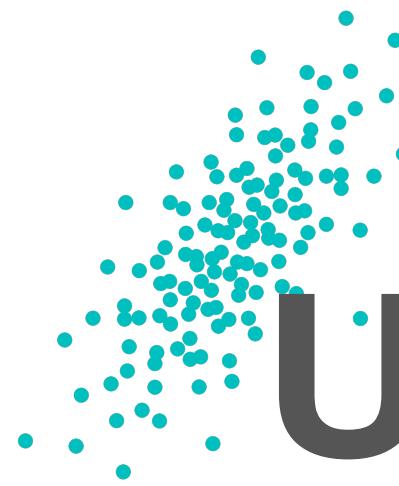
Let  $Y$  be a random variable, often  $Y = f(X)$

$Y_1, Y_2, \dots$  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  $Y = f(X)$

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

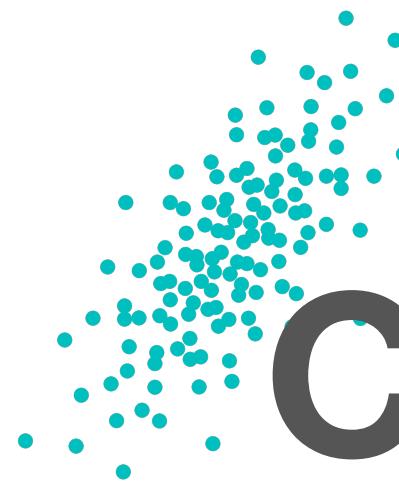
$$\mathbb{E}(\hat{\mu}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(Y_i) = \mu \text{ so } \hat{\mu}_n \text{ is unbiased}$$

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

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

which means that

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



# Monte Carlo estimates of properties of $Y$

Let  $Y$  be a random variable and let  $Y_1, Y_2, \dots$  be an IID random sample

Population Quantity	Sample Approximation
mean $\mu$	$\hat{\mu}_n \pm 2.58\hat{\sigma}_{n_0}^2$ , <code>np.mean(data)</code>
variance $\sigma^2$	$\hat{\sigma}_n^2$ , <code>np.var(data, ddof=1)</code>
cumulative distribution function $F$	empirical distribution function $\hat{F}_n$ <code>statsmodels.distributions.empirical_distribution.ECDF(data)</code>
quantile function $Q$	empirical quantile function $\hat{Q}_n$ <code>np.quantile(data, probabilities)</code>
probability density or mass function $\varrho$	histogram, <code>np.histogram(data, bins = "auto")</code> kernel density estimator, <code>scipy.stats.gaussian_kde(data)</code>

For  $\mu$  we have an interval estimator, but for the others typically point estimators



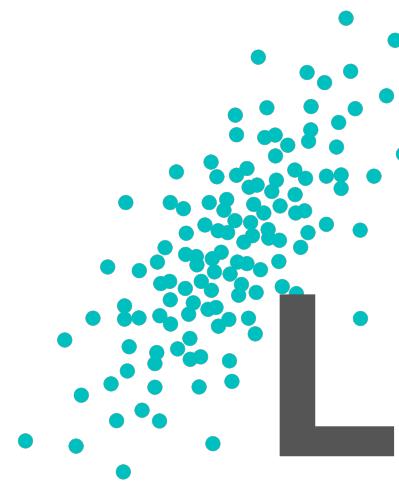
# Are we there yet?

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

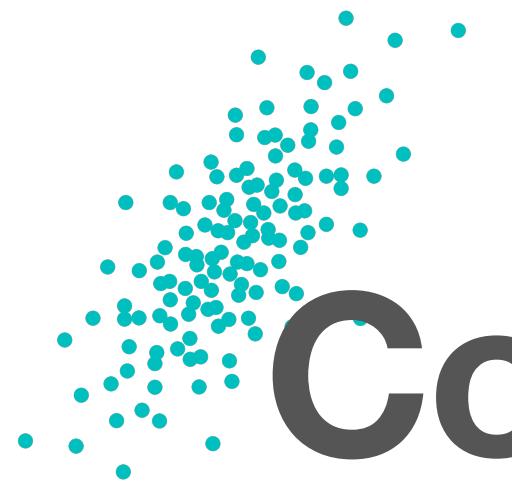
How long should you plan for the trip to take?

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



# Let me know your professional aspirations?

- Go to [menti.com](https://menti.com)
- Use code 6222 6078



# Conditional Probability

Let  $(Y, Z)$  be a vector of random variables

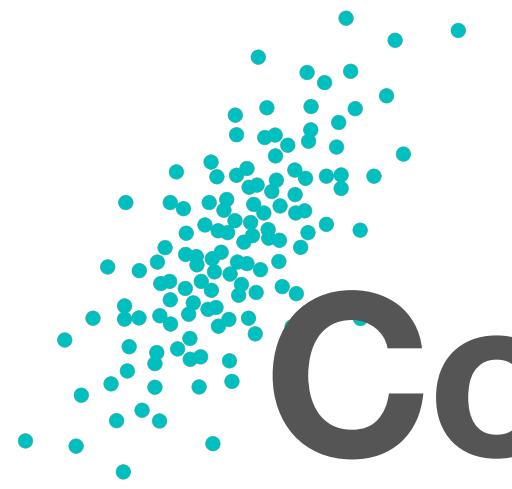
$$\mathbb{P}(Y = y \mid Z = z) = \frac{\mathbb{P}(Y = y \& Z = z)}{\mathbb{P}(Z = z)}$$

$$\mathbb{P}(Y = y \mid Z = z) = \frac{\mathbb{P}(Z = z \mid Y = y)\mathbb{P}(Y = y)}{\mathbb{P}(Z = z)} \quad \text{Bayes' Theorem}$$

$$\varrho_{Y|Z}(y \mid z) = \frac{\varrho_{Y,Z}(y, z)}{\varrho_Z(z)}$$

$$\mathbb{E}(Y) = \mathbb{E}_Z[\mathbb{E}_Y(Y \mid Z)]$$

$$\text{var}(Y) = \mathbb{E}_Z[\text{var}_Y(Y \mid Z)] + \text{var}_Z[\mathbb{E}_Y(Y \mid Z)]$$



# Conditional Monte Carlo

Let  $(Y = f(\mathbf{X}_{1:d}), \mathbf{X}_{1:d})$  be a vector of random variables

Let  $Z = \mathbf{X}_{2:d}$

If  $g(\mathbf{X}_{2:d}) = \mathbb{E}_Y(Y \mid \mathbf{X}_{2:d})$  has an analytic form, then

$\mathbb{E}(Y) = \mathbb{E}_{\mathbf{X}_{2:d}}[g(\mathbf{X}_{2:d})]$  and  $\text{var}(Y) \geq \text{var}_Z[g(\mathbf{X}_{2:d})]$

If  $h(y, \mathbf{X}_{2:d}) = \varrho_{Y|\mathbf{X}_{2:d}}(y \mid \mathbf{x}_{2:d})$  has an analytic form, then

$\varrho_Y(y) = \mathbb{E}_{\mathbf{X}_{2:d}}[h(y, \mathbf{X}_{2:d})]$



# Summary of the Introduction

- We use **sample quantities**, often using IID samples, to estimate properties of a random variable with a complex distribution, including
  - Means, variances, covariances
  - Probability density/mass functions, cumulative distribution functions, and quantile functions
- For means we have **interval** estimators, which indicate the uncertainty in our estimates, using the CLT
- For the other properties we have **point** estimators, i.e., only one approximate value for each thing to be estimated