

# Lecture 12 Continuous Probability Distribution

BIO210 Biostatistics

---

Xi Chen

Fall, 2023

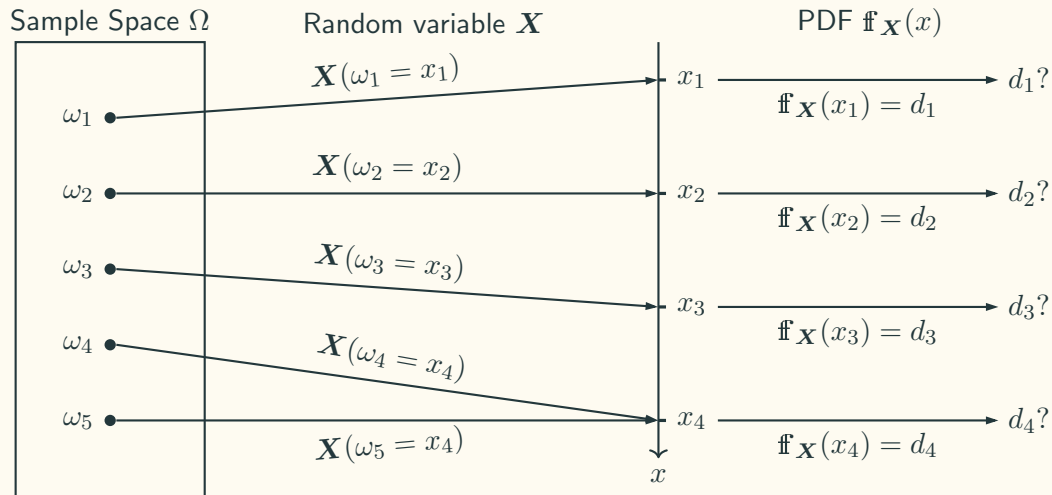
School of Life Sciences

Southern University of Science and Technology

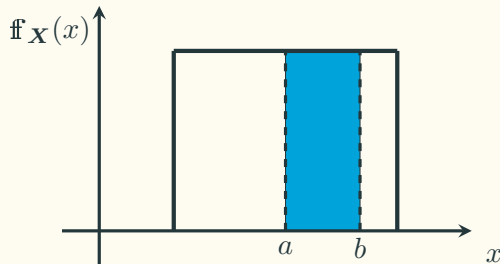
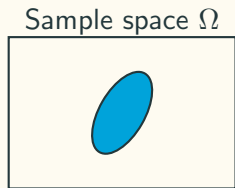


南方科技大学生命科学学院  
SUSTech · SCHOOL OF  
**LIFE SCIENCES**

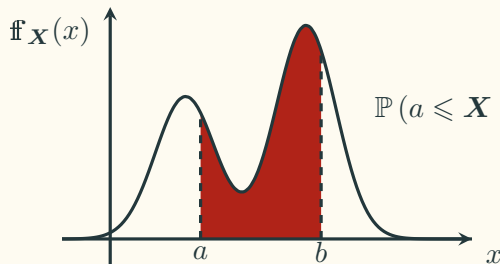
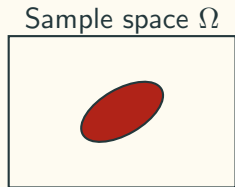
# Probability Density Function (PDF)



# Probability Density Function (PDF)

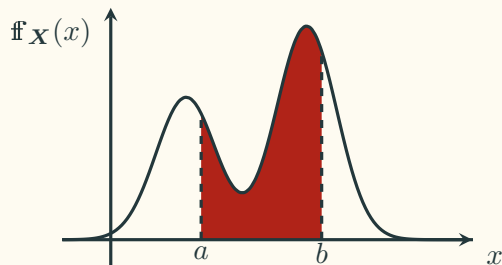


$$\mathbb{P}(a \leq X \leq b) \\ = f_X(x) \cdot (b - a)$$



$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) dx$$

# Probability Density Function (PDF)



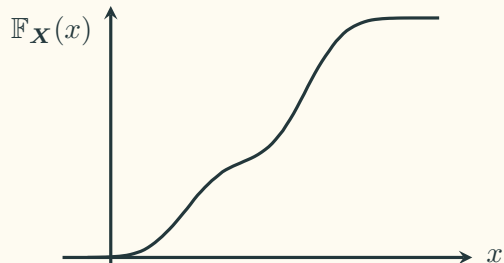
$$f_X(x) \geq 0, \quad \int_{-\infty}^{+\infty} f_X(x) dx = 1$$

$$\mathbb{P}(X = a) = ?$$

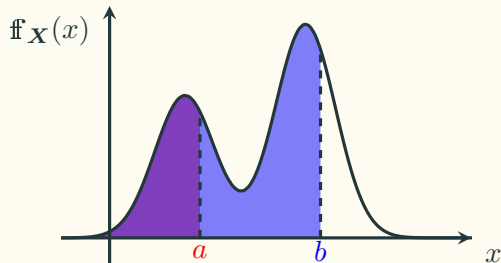
$$\begin{aligned} \mathbb{P}(x \leq X \leq x + \delta) \\ = \int_x^{x+\delta} f_X(x) dx = f_X(x) \cdot \delta \end{aligned}$$

$$f_X(x) = \frac{\mathbb{P}(x \leq X \leq X + \delta)}{\delta}$$

# Cumulative Distribution Function



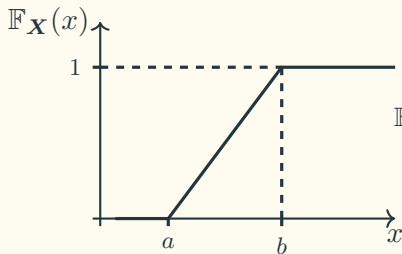
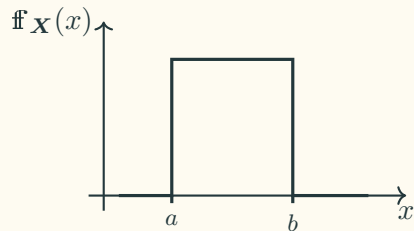
$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(t) dt$$



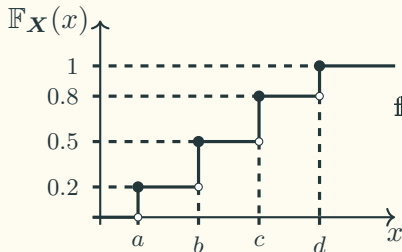
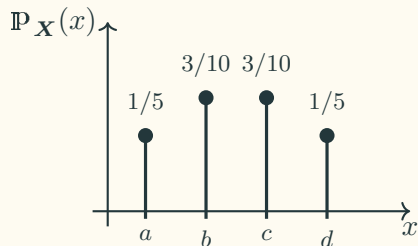
$$F_X(a) = \mathbb{P}(X \leq a) = \int_{-\infty}^a f_X(x) dx$$

$$F_X(b) = \mathbb{P}(X \leq b) = \int_{-\infty}^b f_X(x) dx$$

# Cumulative Distribution Functions (CDFs)



$$F_X(x) = \mathbb{P}(X \leq x) \\ = \int_{-\infty}^x f_X(t) dt$$



$$f_X(x) = \mathbb{P}(X \leq x) \\ = \sum_{k \leq x} P_X(k)$$

# Expectation and Variance

## The continuous case

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x \mathbf{f}_X(x) \, dx$$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) \mathbf{f}_X(x) \, dx$$

$$\begin{aligned}\text{Var}(X) &= \sigma_X^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \int_{-\infty}^{+\infty} (X - \mathbb{E}[X])^2 \mathbf{f}_X(x) \, dx \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2\end{aligned}$$

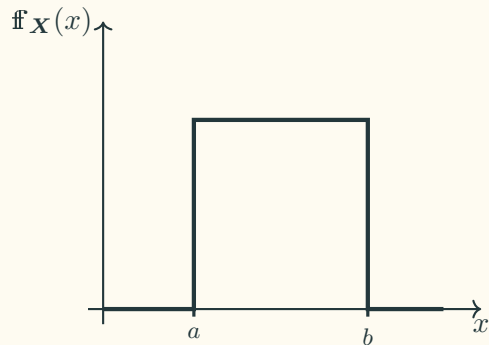
## The discrete case

$$\mathbb{E}[X] = \sum_x x \mathbf{P}_X(x)$$

$$\mathbb{E}[g(X)] = \sum_x g(x) \mathbf{P}_X(x)$$

$$\begin{aligned}\text{Var}(X) &= \sigma_X^2 = \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \sum_x (X - \mathbb{E}[X])^2 \mathbf{P}_X(x) \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2\end{aligned}$$

# Continuous Uniform Distribution



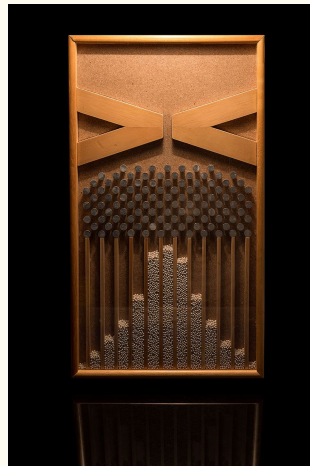
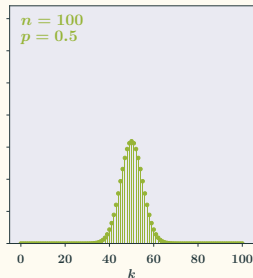
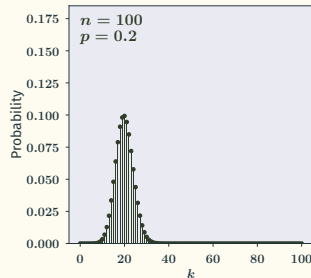
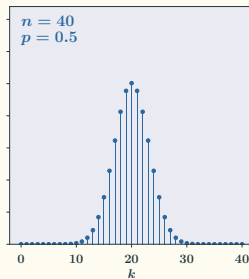
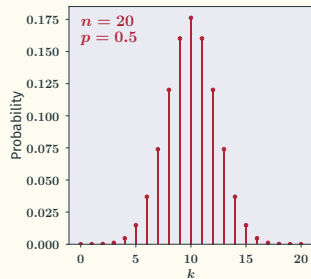
$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & x < a \text{ or } x > b \end{cases}$$

$$\mathbb{E}[X] = ?$$

$$\mathbb{V}\text{ar}(X) = ?$$



# The Idea of The Normal Distributions

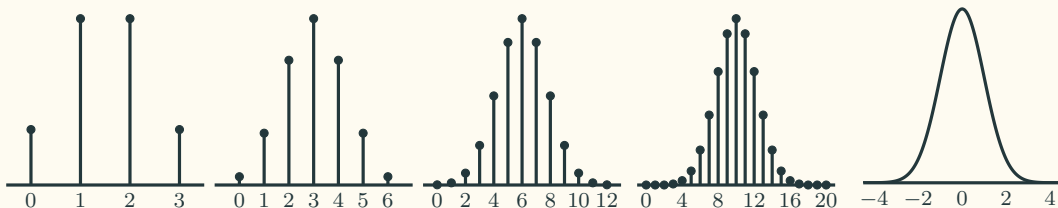


The Bean Machine  
by Francis Galton

# A Little History of The Normal Distribution - Binomial Approximation

**Abraham de Moivre:** The Doctrine of Chances, 1738

- The middle term of the binomial coefficient is approximately:  $\frac{2}{\sqrt{2\pi n}}$
- Stirling's formula:  $n! \simeq n^n e^{-n} \sqrt{2\pi n}$



# The de Moivre–Laplace Theorem

When  $n$  becomes large, and  $np, nq$  are also large:

$$\mathbb{P}_{\mathbf{X}}(k) = \binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi} \sqrt{npq}} \cdot e^{-\frac{(k - np)^2}{2npq}}, \text{ where } q = 1 - p$$

# A Little History of the Normal Distribution - The Error Curve

**Carl Friedrich Gauss:** Theoria motus corporum coelestium in sectionibus conicis solem ambientium (Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections), 1809.

- least squares
- maximum likelihood
- normal distribution

**Pierre Simon de Laplace**

- In 1782:  $\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$
- In 1810: the central limit theorem

# A Little History of the Normal Distribution - Beyond Errors

**Adolphe Quetelet:** the average man - Letters addressed to H.R.H. the grand duke of Saxe Coburg and Gotha, on the Theory of Probabilities as Applied to the Moral and Political Sciences, 1846.

TABLE 1: Chest measurement of Scottish soldiers

Girth	Frequency
33	3
34	18
35	81
36	185
37	420
38	749
39	1,073
40	1,079
41	934
42	658
43	370
44	92
45	50
46	21
47	4
48	1