Loss distributions

Random variable for losses

- Let L be a random variable representing losses.
- The distribution of L captures the uncertainty in potential losses.
- Commonly used distributions for modeling losses include:
  - Normal distribution
  - Log-normal distribution
  - Exponential distribution

What is random variable?

A **random variable** is a function that maps outcomes of a random experiment to real numbers. Formally:

$$X : \Omega \to \mathbb{R}$$

where:

- Ω is the **sample space** (the set of all possible outcomes),
- $X(\omega) \in \mathbb{R}$ is the value assigned to outcome $\omega \in \Omega$.

Types of random variables

- **Discrete random variables** take on a countable number of values. Examples include:
  - Number of claims in insurance
  - Number of defaults in a loan portfolio
- **Continuous random variables** take on an uncountable number of values. Examples include:
  - Loss amounts in a financial portfolio
  - Time until an event occurs (e.g., default)

Discrete random variables

Takes values in a countable set (e.g., integers).
Example: Let X be the number of heads in 3 coin tosses. Then:

$$X \in \{0, 1, 2, 3\}$$

Sample Space:

$$\Omega = \{\text{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}\}$$

Continuous random variables

Takes values in an interval or uncountable set.
Example: (Default time) Let T be the time (in seconds) until a loan defaults. Then:

$$T \in [0, \infty)$$

Probability Distribution

A random variable is associated with a **probability distribution**:

- For discrete X, we define a **probability mass function (PMF)** $p(x)$:
$$P(X = x) = p(x)$$
- For continuous X, we define a **probability density function (PDF)** $f(x)$:
$$P(a \le X \le b) = \int_a^b f(x)\, dx$$

Cumulative Distribution Function (CDF)

The **Cumulative Distribution Function (CDF)** of a random variable X is defined as:

$$F(x) = P(X \leq x)$$

- For discrete X:
$$F(x) = \sum_{t \leq x} p(t)$$
where $p(t)$ is the PMF.
- For continuous X:
$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$
where $f(t)$ is the PDF.

Properties of CDF

Properties of CDF:

- $F(x)$ is non-decreasing.
- $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.
- $P(a < X \leq b) = F(b) - F(a)$.

Discrete Example Continued

Probability Mass Function (PMF):

For the random variable X (number of heads in 3 coin tosses), the PMF is:

$$P(X = x) = \binom{3}{x} \cdot (0.5)^x \cdot (0.5)^{3-x}, \quad x = 0, 1, 2, 3$$

Table of Probabilities:

x	$P(X = x)$
0	0.125
1	0.375
2	0.375
3	0.125

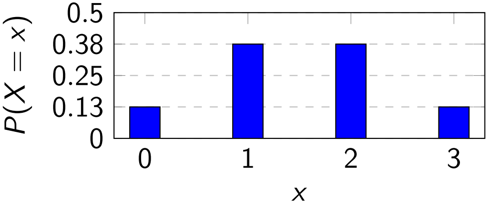
Cumulative Distribution Function (CDF)

Cumulative Distribution Function (CDF):

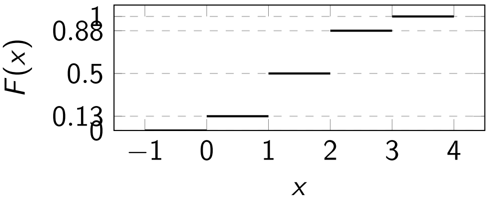
$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.125 & \text{if } 0 \leq x < 1 \\ 0.5 & \text{if } 1 \leq x < 2 \\ 0.875 & \text{if } 2 \leq x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$

Graph of PMF and CDF

PMF:



CDF:



Continuous Example: PDF and CDF

Probability Density Function (PDF):

$$f(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Cumulative Distribution Function (CDF):

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

Mean and Variance of Random Variable

Mean (Expected Value):

$$E[X] = \begin{cases} \sum_x x \cdot P(X = x) & \text{if discrete} \\ \int_{-\infty}^{\infty} x \cdot f(x)\, dx & \text{if continuous} \end{cases}$$

Variance:

$$Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

Standard Deviation:

$$\sigma(X) = \sqrt{Var(X)}$$

Discrete Example Continued: Mean and Variance

Mean (Expected Value):

$$E[X] = \sum_x x \cdot P(X = x) = 0 \cdot 0.125 + 1 \cdot 0.375 + 2 \cdot 0.375 + 3 \cdot 0.125 = 1.5$$

Variance:

$$Var(X) = E[X^2] - (E[X])^2 = \sum_x x^2 \cdot P(X = x) - (1.5)^2$$

$$= (0^2 \cdot 0.125 + 1^2 \cdot 0.375 + 2^2 \cdot 0.375 + 3^2 \cdot 0.125) - 2.25 = 0.75$$

Standard Deviation:

$$\sigma(X) = \sqrt{Var(X)} = \sqrt{0.75} \approx 0.866$$

Continuous Example Continued: Mean and Variance

Mean (Expected Value):

$$E[X] = \int_0^1 x \cdot 1 \, dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}$$

Variance:

$$Var(X) = E[X^2] - (E[X])^2 = \int_0^1 x^2 \cdot 1 \, dx - \left(\frac{1}{2}\right)^2$$

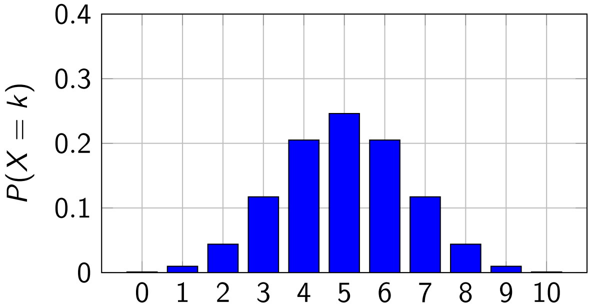
$$= \left[\frac{x^3}{3}\right]_0^1 - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Standard Deviation:

$$\sigma(X) = \sqrt{Var(X)} = \sqrt{\frac{1}{12}} = \frac{1}{2\sqrt{3}} \approx 0.289$$

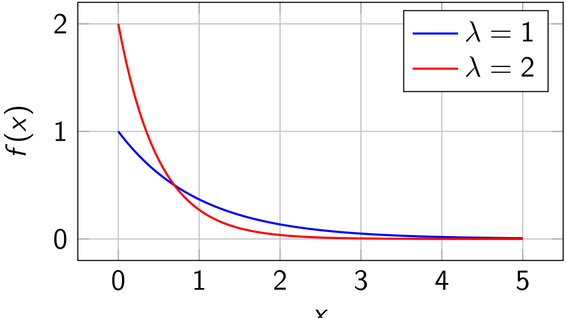
Binomial Distribution

- **Binomial Distribution:** $X \sim \text{Bin}(n, p)$
  - PMF: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k = 0, 1, \ldots, n$
  - Models the number of successes in n independent Bernoulli trials with success probability p
- Properties:
  - Mean: $E[X] = n \cdot p$
  - Variance: $\mathrm{Var}(X) = n \cdot p \cdot (1-p)$
- Example: Number of defective items in a batch of n products.



Exponential Distribution

- **Exponential Distribution:** $X \sim \text{Exp}(\lambda)$
  - PDF: $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$, $\lambda > 0$
  - Models the time between events in a Poisson process
- Properties:
  - Mean: $E[X] = \frac{1}{\lambda}$
  - Variance: $\mathrm{Var}(X) = \frac{1}{\lambda^2}$
  - Memoryless property: $P(X > s + t \mid X > s) = P(X > t)$
- Example: Time until a claim is filed in insurance.

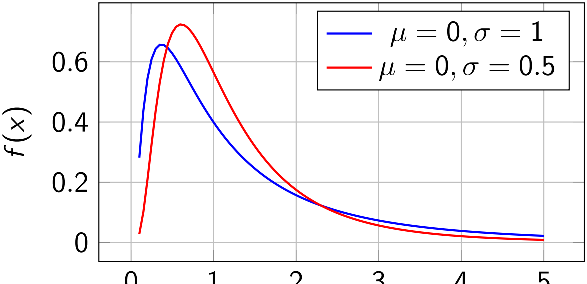


Normal Distribution

- **Normal Distribution:** $L \sim N(\mu, \sigma^2)$
  - PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
  - Symmetric around the mean μ
  - Light tails: extreme losses are rare
- Often used as a baseline model for financial returns
- Properties:
  - Mean: μ
  - Variance: σ^2
- Standard Normal: If $L \sim N(\mu, \sigma^2)$, then $Z = \frac{L-\mu}{\sigma} \sim N(0, 1)$.
- Limitation: May underestimate tail risk in financial markets

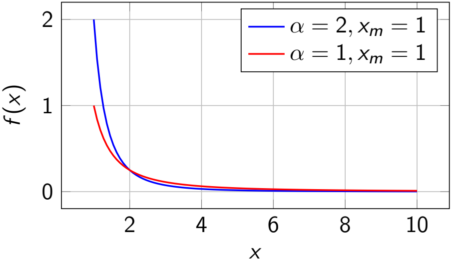
Log-Normal Distribution

- **Log-Normal Distribution:** $L \sim \text{LogNormal}(\mu, \sigma^2)$
  - If $X \sim N(\mu, \sigma^2)$, then $L = e^X$ follows a log-normal distribution.
  - PDF: $f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$ for $x > 0$.
- Properties:
  - Mean: $E[L] = e^{\mu+\sigma^2/2}$
  - Variance: $\mathrm{Var}(L) = \left(e^{\sigma^2} - 1\right) e^{2\mu+\sigma^2}$
- Applications: Modeling asset prices, insurance claims, and other non-negative data.



Heavy Tail distribution: Pareto

- **Pareto Distribution:**
  - PDF: $f(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$ for $x \geq x_m > 0$, $\alpha > 0$.
  - Used to model losses with a high probability of extreme values (catastrophic risk).
  - Heavy tail: moments may not exist if α is small.
- Example: Insurance claims, natural disasters.
- Properties:
  - Mean exists if $\alpha > 1$; $E[L] = \frac{\alpha x_m}{\alpha - 1}$.
  - Variance exists if $\alpha > 2$; $Var(L) = \frac{x_m^2 \alpha}{(\alpha-1)^2(\alpha-2)}$.



Heavy Tail distribution: Student's t

- Standard t-distribution: $t(\nu)$:

$$T = \frac{Z}{\sqrt{V/\nu}}$$

where $Z \sim N(0,1)$ and $V \sim \chi^2(\nu)$ (chi-squared with ν degrees of freedom).

- $\mathbb{E}[T] = 0$ for $\nu > 1$ and $\mathrm{Var}[T] = \frac{\nu}{\nu-2}$ for $\nu > 2$.
- t-distribution $L \sim t(\nu, \mu, \sigma^2)$:

$$T = \frac{L-\mu}{\sigma} \sim t(\nu)$$

- $\mathbb{E}[L] = \mu$ for $\nu > 1$, $\mathrm{Var}[L] = \frac{\sigma^2\nu}{\nu-2}$ for $\nu > 2$.
- Heavier tails than normal distribution; used to model financial returns.