Continuous Random Variables

Lesson 06

Digital Finance

Outline

From Discrete to Continuous

Continuous random variable: Takes any value in an interval (like height, time, temperature).
Why is $P(X = \text{exactly } 1.7000...) = 0$?

- Infinite precision = infinitely many possible values
- Any single point has probability $1/\infty = 0$
- Instead, ask: $P(1.69 < X < 1.71)$ – that's non-zero!

Key shift in thinking:

- Discrete: count probability at each point (bars)
- Continuous: measure probability over intervals (areas)

Think of continuous probability as "area under a curve," not "height at a point."

Probability Density Function (PDF)

Visual idea: Draw a curve. The area under the curve between two points = probability.
PDF rules:

- Curve never goes below zero
- Total area under curve = 1 (100%)

Finding probability: $P(a \leq X \leq b)$ = shaded area from a to b
Common confusion: Height $f(x)$ is NOT probability!

- Height can exceed 1 (imagine a narrow, tall curve)
- Only the *area* is probability

If you know calculus: area = $\int_a^b f(x)dx$. If not: think "shaded region."

CDF for Continuous Variables

CDF $F(x) = P(X \leq x) =$ "probability of being at most x"
Visual interpretation:

- CDF at any point = total area under PDF up to that point
- CDF always goes from 0 (far left) to 1 (far right)
- CDF is a smooth, increasing curve for continuous variables

Practical calculations:

- $P(X > a) = 1 - F(a)$ (complement rule)
- $P(a < X < b) = F(b) - F(a)$ (subtract CDFs)

CDF accumulates probability. Use tables/software to find F(x) values.

Continuous Uniform Distribution

Equal probability over interval $[a, b]$:

$$X \sim \text{Uniform}(a, b)$$

PDF:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Properties:

- $E[X] = \frac{a+b}{2}$
- $\text{Var}(X) = \frac{(b-a)^2}{12}$

Used for "complete ignorance" – all values equally likely.

Normal (Gaussian) Distribution

The most important distribution in statistics!
What it looks like: Bell-shaped, perfectly symmetric

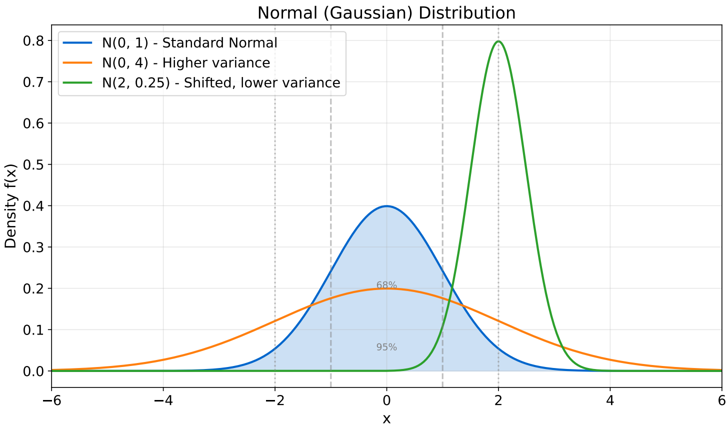
- Peak at the center (mean μ)
- Tails extend infinitely but probability gets tiny
- Wider bell = more spread (larger σ)

Notation: $X \sim N(\mu, \sigma^2)$ means "X follows a normal distribution"

- μ (mu) = mean = where the bell is centered
- σ (sigma) = std dev = how wide the bell is

Formula exists but memorizing it isn't necessary – focus on the shape and properties.

Normal Distribution: Visualization



Different parameters change location and spread.

Standard Normal Distribution

Special case: $Z \sim N(0,1)$ **Standardization:** Any normal can become standard

$$Z = \frac{X - \mu}{\sigma}$$

Empirical rule (68-95-99.7):

- 68% of data within 1 std dev of mean
- 95% within 2 std dev
- 99.7% within 3 std dev

Z-scores measure "how many standard deviations from the mean."

Exponential Distribution

Time until event occurs:

$$X \sim \text{Exponential}(\lambda)$$

PDF:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

Properties:

- $E[X] = 1/\lambda$
- $\text{Var}(X) = 1/\lambda^2$
- Memoryless: $P(X > s + t | X > s) = P(X > t)$

Continuous analog of geometric distribution.

Log-Normal Distribution

If $\ln(X)$ **is normal, then** X **is log-normal:**

$$X \sim \text{LogNormal}(\mu, \sigma^2) \Leftrightarrow \ln(X) \sim N(\mu, \sigma^2)$$

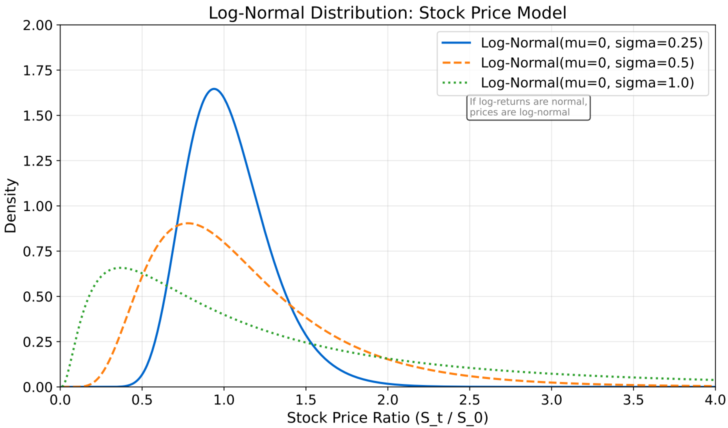
Properties:

- Always positive: $X > 0$
- Right-skewed
- Multiplicative processes

Finance: Stock prices modeled as log-normal

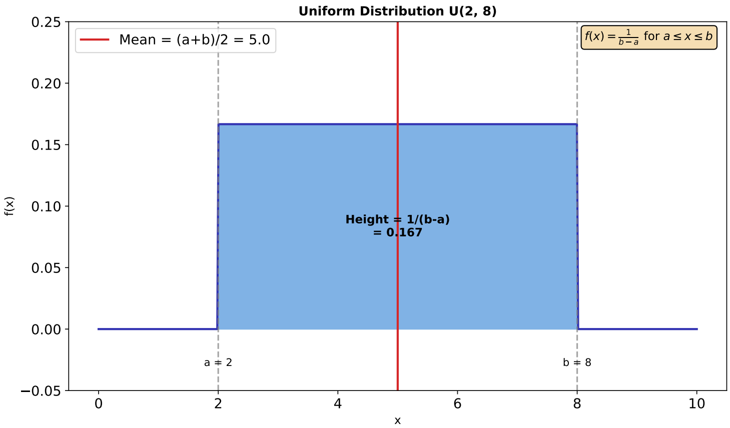
Log-returns are normal; price levels are log-normal.

Log-Normal in Finance



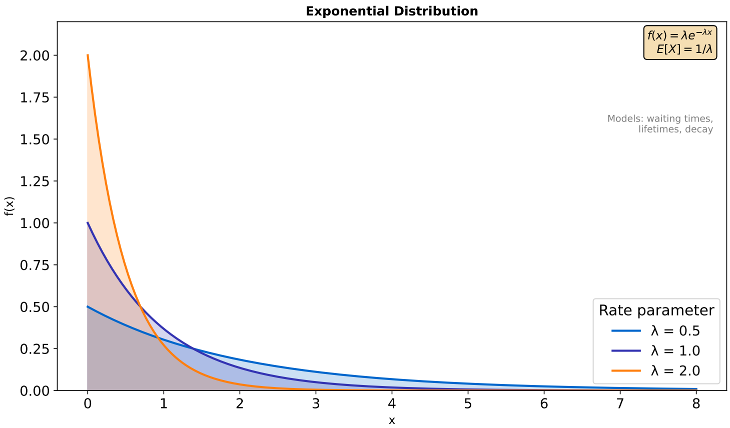
Stock prices can't go negative, log-normal enforces this.

Uniform Distribution



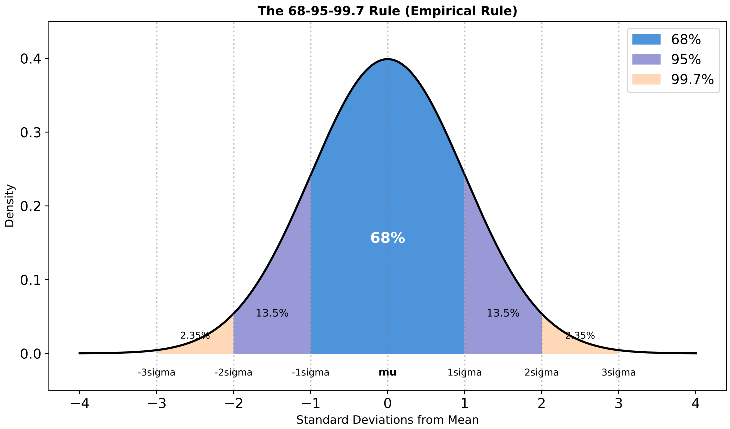
All values equally likely in the interval.

Exponential Distribution



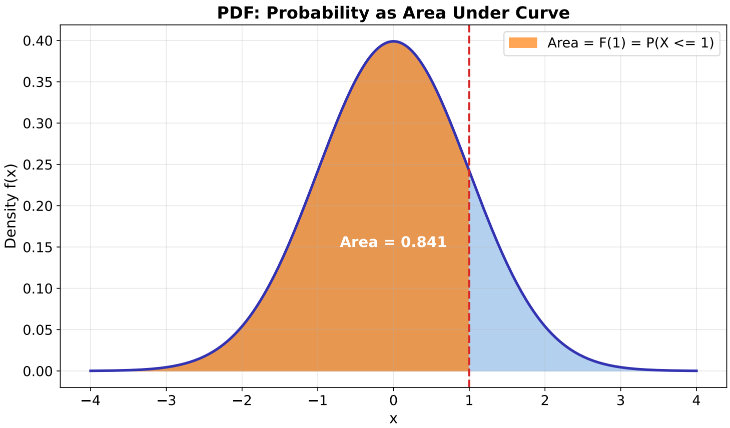
Models waiting times; memoryless property.

Normal Distribution: 68-95-99.7 Rule



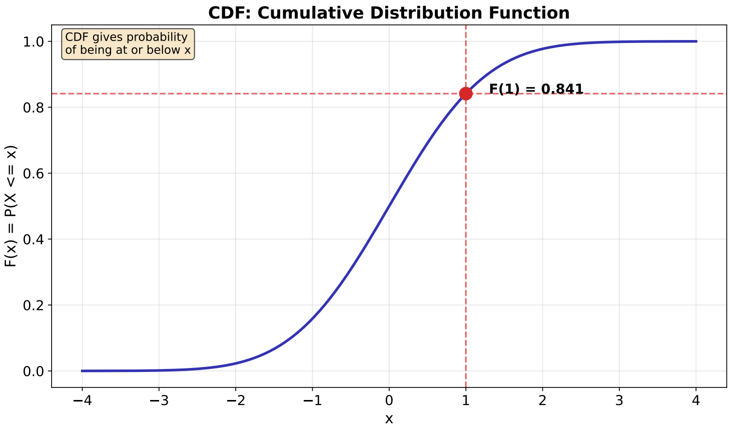
Most data falls within 3 standard deviations.

PDF: Probability as Area



Shaded area represents probability P(X ¡= 1).

CDF: Cumulative Distribution



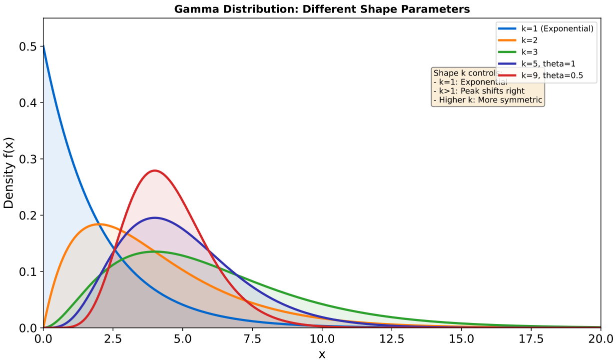
CDF is smooth for continuous distributions.

Distribution Summary

Distribution	Support	Mean	Variance
Uniform(a, b)	$[a, b]$	$(a+b)/2$	$(b-a)^2/12$
Normal(μ, σ^2)	$\mathbb{R}$	μ	σ^2
Exponential(λ)	$[0, \infty)$	$1/\lambda$	$1/\lambda^2$
Log-Normal(μ, σ^2)	$(0, \infty)$	$e^{\mu+\sigma^2/2}$	$(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$

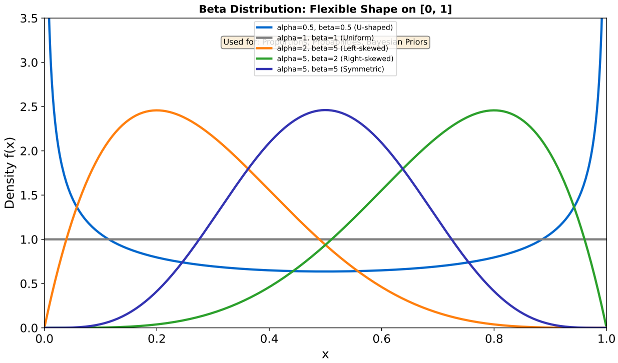
Next lesson: Joint Distributions

Gamma Distribution



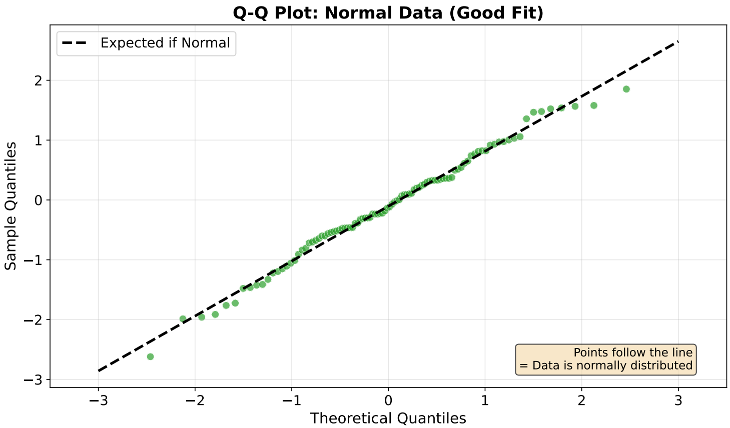
PDF: $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$**; E[X]** $= \alpha/\beta$**, Var(X)** $= \alpha/\beta^2$**.**

Beta Distribution



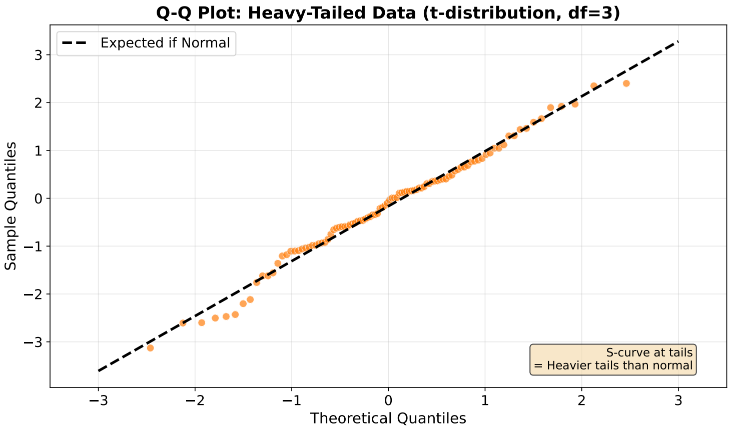
PDF: $f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ **on [0,1]; E[X] =** $\alpha/(\alpha+\beta)$**.**

Q-Q Plot: Normal Data (Good Fit)



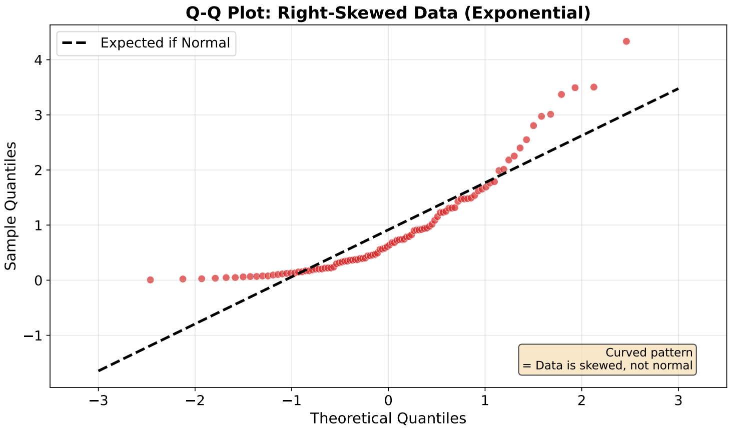
Points on line = data follows normal distribution.

Q-Q Plot: Heavy-Tailed Data



S-curve at tails = more extreme values than normal predicts.

Q-Q Plot: Skewed Data



Curved pattern = data is skewed, not normally distributed.