DMVU2

Random Numbers and Simulation

Random Numbers

- Random numbers are numbers generated in such a way that each value has an equal chance of being chosen.
- In computers, they are produced by mathematical formulas → called **pseudo-random numbers**.
- These are widely used in experiments, games, cryptography, and simulations.

Properties of good random numbers:

1. Uniformly distributed
2. Independent of each other
3. Easy and fast to generate

Simulation

- Simulation is the process of **imitating a real-world system** on a computer to study its behavior.
- Since real systems involve uncertainty, **random numbers are used** to represent events like arrivals, service times, failures, etc.
- By running the simulation many times, we can analyze system performance without disturbing the actual system.

Applications of Simulation:

- Bank or queue systems
- Traffic flow analysis
- Weather prediction
- Inventory and supply chain
- Engineering and scientific experiments

Sampling of Continuous Distributions

What is Sampling?

- **Sampling** means generating values (data points) from a probability distribution.
- In **continuous distributions**, the variable can take any real value within a range.
- Example: Human height, waiting time, temperature → all follow continuous distributions.

Why Sampling is Needed?

- In **simulation**, we need random values that behave like the real-world data.
- Random numbers from computers are usually **Uniform(0,1)**.
- To simulate real systems, we must **transform these uniform random numbers** into values from required distributions (Normal, Exponential, etc.).

Methods of Sampling Continuous Distributions

Methods of Sampling Continuous Distributions

1. Inverse Transform Method

- Based on the cumulative distribution function (CDF).
- If $U \sim U(0,1)$, then

$$X = F^{-1}(U)$$

follows the required distribution.

- **Examples:**
  - Uniform $[a,b]$: $X = a + (b-a)U$
  - Exponential (λ): $X = -\frac{1}{\lambda}\ln U$

2. Transformation Methods

- Some distributions can be generated by transforming uniform random numbers.
- **Normal Distribution (Box-Muller method):**

If $U_1, U_2 \sim U(0,1)$,

$$Z_1 = \sqrt{-2\ln U_1}\cos(2\pi U_2), \quad Z_2 = \sqrt{-2\ln U_1}\sin(2\pi U_2)$$

both are $N(0,1)$.

3. Acceptance–Rejection Method

- Used when inverse transform is difficult.
- Idea: Propose a value from an easy distribution, accept it with a certain probability.
- Works for complex distributions.

Common Continuous Distributions

- **Uniform Distribution** – equal chance for all values.
- **Normal Distribution** – bell-shaped, common in nature.
- **Exponential Distribution** – models time between random events (e.g., arrivals in a queue).
- **Gamma / Beta Distributions** – used in reliability and Bayesian analysis.

Applications

- Queue simulation (service times ~ Exponential)
- Reliability analysis (failure times ~ Weibull)
- Financial modeling (stock returns ~ Normal/Lognormal)
- Weather, traffic, and scientific simulations

Monte Carlo Methods

- **Monte Carlo Method** is a problem-solving technique that uses **random sampling and statistical experiments** to estimate numerical results.
- Instead of solving a problem mathematically, we run **repeated simulations** using random numbers.

Steps in Monte Carlo Method

1. **Define the problem** – Identify what you want to estimate (e.g., area, probability, average).
2. **Generate random numbers** – Usually uniform random numbers from a computer.
3. **Transform random numbers** – Convert them into samples from the required distribution (Normal, Exponential, etc.).
4. **Perform simulation** – Run the process many times using these samples.
5. **Compute the estimate** – Take the average of results ⟶ this approximates the solution.

Properties

- Accuracy increases as the number of trials increases.
- Convergence rate is proportional to 1/sqrt{n} , where n = number of samples.
- Useful for **complex, high-dimensional problems** where analytical solutions are hard.

Applications

- Estimation of mathematical constants (π).
- Integration of complex functions.
- Physics and engineering simulations.
- Finance: option pricing, risk analysis.
- Operations research: queue and inventory simulations.

Small Example: Estimating π

- Imagine a square of side 2 with a circle of radius 1 inside it.
- Area of square = 4, Area of circle = π.
- Ratio = π/4.

Steps:

1. Generate random points (x, y) uniformly in the square $[-1,1] \times [-1,1]$.
2. Count how many fall inside the circle ($x^2 + y^2 \leq 1$).
3. Approximate:

$$\pi \approx 4 \times \frac{\text{Points inside circle}}{\text{Total points}}$$

Example run:

- Suppose we generate 10,000 points.
- 7,850 fall inside the circle.
- Estimate: $\pi \approx 4 \times \frac{7850}{10000} = 3.14$ (close to actual π).

Hypothesis Testing

What is Hypothesis Testing?

- **Hypothesis testing** is a statistical method used to make decisions or draw conclusions about a population **based on sample data**.
- It starts with two statements:
  - **Null Hypothesis (H_0):** Assumes no effect or no difference (the default claim).
  - **Alternative Hypothesis (H_1 or Ha):** Opposite of H_0, states that there is an effect or difference.
- Using sample data, we decide whether to **reject H_0** in favor of H_1, or **fail to reject H_0.**

Example:
H_0: "The average weight of packets is 1kg."
H_1: "The average weight is not 1kg."

Purpose of Hypothesis Testing

1. To provide an **objective rule** for decision-making under uncertainty.
2. To test **assumptions or claims** about population parameters (mean, proportion, variance).
3. To help in **scientific research** by validating theories or models.

4. To measure the **strength of evidence** against the null hypothesis.

5. To control the risk of errors (Type I and Type II).

Advantages of Hypothesis Testing

- Provides a **structured procedure** to make decisions.
- Ensures **objectivity** instead of relying on guesswork.
- Allows **quantification of risk** through significance level (α).
- Widely applicable in business, medicine, engineering, and social sciences.
- Helps in **comparing groups, treatments, or processes**.
- Can be applied to both **small and large samples** using different statistical tests (z-test, t-test, chi-square, etc.).

Rejection Region

Theory: The set of values of the test statistic where H0H0 is rejected.

Determined by: Significance level, test direction (one-tailed, two-tailed).

If test statistic falls in rejection region: Reject H0.

Example: For α=0.05α=0.05, critical z-values are ±1.96. If computed z >1.96 or <–1.96, reject H0.

Type I Error (False Positive):

- Occurs when the null hypothesis (H0*H*0) is true, but we mistakenly reject it.
- It’s like sounding an alarm for danger that isn’t there.
- The probability of making a Type I error is denoted by α*α* (significance level), often set at 0.05 (5%).
- Type I error means you believe there is an effect or difference when there isn’t one—a “false positive.”

Type II Error (False Negative):

- Happens when the null hypothesis (H0*H*0) is false, but we fail to reject it.
- It’s like missing a real danger because the alarm doesn’t sound.
- The probability of making a Type II error is denoted by β.
- **Type II error** means you miss finding an effect or difference that actually exists—a “false negative.”

Example: Medical Testing

Suppose we are testing a new drug, and our hypotheses are:

- H_0: The drug is **not effective.**
- H_1: The drug **is effective.**

Scenario	Reality	What Test Shows	Error Type
Drug not effective	Drug not effective	Conclude it's effective	Type I
Drug effective	Drug effective	Conclude not effective	Type II

Type I Error:

- We conclude the drug works (reject H_0), when it actually doesn't (false alarm).
- Consequence: Approving a useless drug.

Type II Error:

- We conclude the drug doesn't work (fail to reject H_0), when it actually does (missed detection).
- Consequence: Rejecting a beneficial drug.

Markov Process

A Markov Process is a type of stochastic process that satisfies the Markov Property:

The future state of the system depends only on the present state, not on the past sequence of states.

Mathematically:

P(Xt+1=x|Xt,Xt−1,…,X0)=P(Xt+1=x|Xt)

This property is called memoryless.

2. Key Characteristics

1. States: All possible conditions of the system (e.g., Sunny, Rainy).
2. Transitions: The movement from one state to another.
3. Transition Probabilities: Probability of moving from one state to another. Represented in a transition matrix.
4. Time:
   - Discrete-time Markov process (Markov Chain): The system evolves in fixed steps (1, 2, 3, …).

- Continuous-time Markov process: The system evolves continuously over time.

3. Example: Weather Prediction

Suppose the weather can be either Sunny (S) or Rainy (R).

- If today is Sunny, tomorrow will be Sunny with probability 0.8 and Rainy with probability 0.2.
- If today is Rainy, tomorrow will be Sunny with probability 0.4 and Rainy with probability 0.6.

☞ Transition Probability Matrix:

P=[0.8 0.2]

[0.4 0.6]

Here rows represent current state, columns represent next state.

This shows how the future (tomorrow) depends only on today's weather, not on the weather from 2 days ago.

5. Applications

- Weather Forecasting – modeling transitions between sunny, rainy, cloudy.
- Queueing Systems – future queue length depends only on current length.
- Finance – stock price movements modeled as Markov processes.
- Speech Recognition / NLP – using Hidden Markov Models.
- Biology – DNA sequence analysis.

6. Advantages

✓ Simple and powerful way to model dynamic random systems.
✓ Memoryless property makes analysis easier.
✓ Widely applicable in engineering, AI, and business systems.

7. Disadvantages

✗ Assumes only present matters – sometimes unrealistic.
✗ Transition probabilities must be known or estimated, which can be hard.
✗ Oversimplifies complex systems (e.g., weather may depend on several past days).

Hidden Markov Model (HMM)

A **Hidden Markov Model (HMM)** is a statistical model used to represent systems that are assumed to be a Markov process with unobservable (hidden) states. Unlike regular Markov models, in HMMs the actual states are not directly visible, but each state produces an observable output (emission), allowing indirect inference of the underlying state sequence.

2. Key Components

- **States:**
  The set of hidden states (e.g., "Sunny" and "Rainy" in a weather model). The sequence of states follows the Markov property (future state depends only on current state).
- **Observations (Emissions):**
  The visible outputs generated by each hidden state (e.g., "Umbrella" or "No Umbrella").
  Each state has a probability distribution over possible observations.
- **Transition Probabilities:**
  The probability of moving from one hidden state to another.
- **Emission Probabilities:**
  The probability that a particular observation is produced by a state.
- **Initial State Distribution:**
  The probability of starting in each state.

3. How HMM Works

- At every time step:
  - The model transitions from one state to another based on transition probabilities.
  - The current (hidden) state emits an observation based on the emission probabilities.
- Only observations are seen; the sequence of hidden states is inferred.

4. Diagram Explanation

- **Transition State Diagram:**
  Shows arrows between hidden states (e.g.,

“Sunny” to “Rainy”) labeled by the probability of transition.

- **Emission State Diagram:** Shows arrows from states to observations (e.g., “Sunny” emits “No Umbrella” with probability 0.8).

5. Mathematical Representation

- **States:** S={s1,s2,...,sN}
  Observations: O={o1,o2,...,oM}
- **Transition Matrix:**

  A=[aij],
  where aij=P(Next state=sj∣Current state=s)

- **Emission Matrix:**

  B=[bjk],
  where bjk=P(Observation=ok∣State=sj)

- **Initial Probabilities:** π

6. Example: Weather and Activities

Suppose you want to infer the weather (hidden state: Sunny, Rainy) based on the observation of whether people carry umbrellas.

- **States:** Sunny, Rainy
- **Observations:** Umbrella, No Umbrella
- **Transition probabilities:**
  - P(Sunny→Rainy)=0.
- **Emission probabilities:**
  - P(Umbrella∣Rainy)=0.9*P*(Umbrella∣Rainy)=0.9
  - P(Umbrella∣Sunny)=0.2*P*(Umbrella∣Sunny)=0.2

If you observe someone carrying an umbrella, you use the HMM to estimate whether it is likely raining.

7. Applications

- Speech recognition (inferring phonemes from sound)
- Part-of-speech tagging (inferring grammatical structure from words)
- Bioinformatics (gene prediction)
- Gesture recognition
- Finance and sequence prediction

8. Advantages

- Powerful for modeling sequential data where states are not directly observed.
- Efficient algorithms for training (Baum-Welch) and inference (Viterbi).

9. Disadvantages

- Requires large data to estimate probabilities accurately.
- Assumes Markov property and discrete states, which may not always fit real-world complexity.

Poisson Process

Definition

A **Poisson process** is a type of **stochastic process** that models the occurrence of random events over time (or space), where:

1. Events occur **independently** of one another.
2. The probability of a **single event** occurring in a very small interval Δt is proportional to the length of the interval (λΔt, where λ = rate of occurrence).
3. The probability of **two or more events** occurring in a very small interval is negligible.

Mathematical Detail:

- If $N(t)$ is the number of events by time t, then:

$$P(N(t)=k)=\frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

  This is the **Poisson distribution** that gives the probability of k events in interval t.

- The **inter-arrival times** between consecutive events are exponentially distributed with mean $\frac{1}{\lambda}$.

Example:

- Consider a call center receiving customer calls.
- Calls come independently.
- The average call rate λλ might be 10 calls per hour.
- The number of calls N(t)$N(t)$ in any hour follows a Poisson distribution with mean 10.
- If asked what is the probability of receiving exactly 8 calls in one hour, use the Poisson formula:

P(N(1)=8)=(10^8*e^−10)/8!

This models real-life scenarios where calls are unpredictable but average out over time.

Why Suitable for Rare Events? (Justification)

- **Rare events** are those that occur **infrequently but randomly** (e.g., earthquakes, accidents, server crashes).
- The **Poisson distribution** models the probability of a given number of events in a fixed interval, especially when events are rare and independent.
- Since the process assumes events occur **individually and unpredictably**, it aligns well with the nature of rare events.
- Example: The number of accidents on a highway per day, or number of calls arriving at a call center per minute.

Gaussian Processes (GPs)

A **Gaussian Process** is a collection of random variables, any finite number of which have a joint Gaussian (normal) distribution. It is a powerful **non-parametric Bayesian approach** used for regression, classification, and function approximation.

- **Definition:** A GP defines a distribution over functions, fully specified by a **mean function** m(x) and a **covariance function (kernel)** k(x,x').

Formal Notation

A function $f(x)$ is said to be drawn from a Gaussian Process with mean function $m(x)$ and covariance (kernel) function $k(x,x')$ as:

$$f(x) \sim \mathcal{GP}\big(m(x), k(x,x')\big)$$

- Mean Function:

$$m(x) = \mathbb{E}[f(x)]$$

This gives the expected value of the function at input x.

- Covariance (Kernel) Function:

$$k(x,x') = \mathrm{Cov}(f(x), f(x'))$$

This defines how related the function values at two points x and x' are.

Intuitive Explanation

- The **mean function** m(x)$m(x)$ can be thought of as describing the average shape of the functions sampled from the GP.
- The **kernel function** k(x,x′)$k(x,x')$ controls the smoothness and variability of functions, deciding how much f(x) and f(x′) move together.

Example

A simple example is Gaussian Process regression, where you learn a function fitting data points while controlling uncertainty using the mean and kernel functions.

- For instance, when fitting noisy observations, the GP uses the kernel to generalize and predict at new points, providing not only an estimate but also a measure of uncertainty.

Stochastic Processes

Definition

A **stochastic process** is a collection of random variables indexed by time (or another parameter), which represents the evolution of a system that changes **randomly over time**.

Formally, {X(t),t∈T}

where **X(t)** is a random variable at time *t*, and **T** is the index set (usually time).

Key Features

1. **Randomness:** Outcomes cannot be predicted with certainty.
2. **Indexing by Time:** Each time point corresponds to a random variable.
3. **State Space:** Possible values the process can take (e.g., integers, real numbers).
4. **Classification:**
   - **Discrete-time vs. Continuous-time** (based on index set).
   - **Discrete-state vs. Continuous-state** (based on state space).

Types of Stochastic Processes

- **Poisson Process:** Models random arrivals (e.g., customers arriving at a bank).
- **Markov Chain:** Future state depends only on the present state, not past history.
- **Brownian Motion:** Continuous random movement of particles.

Example: Brownian Motion

- Imagine a **pollen grain in water** under a microscope.

- It moves **randomly** due to collisions with water molecules.
- At each instant of time *t*, the position is uncertain → this sequence of random positions forms a **stochastic process**.

Advantages

- Useful for modeling **real-world uncertainty** (queues, stock prices, weather, networks).
- Provides mathematical tools for **prediction and optimization**.

Disadvantages

- Often **complex** and requires advanced probability.
- May need **large datasets** to estimate parameters.

How a Queuing System Works

A **queuing system** is a mathematical model used to study how jobs (customers, packets, or tasks) are handled when they compete for limited service resources (servers, counters, processors, etc.).

The **basic components** are:

1. **Arrival Process**
   - Jobs (customers, packets, etc.) arrive randomly over time.
   - In most cases, arrivals are modeled by a **Poisson process** (random arrivals at an average rate λ).
2. **Queue (Waiting Line)**
   - If all servers are busy, jobs must wait in a **queue**.
   - The **queue discipline** decides how jobs are picked: most commonly **FIFO (First In First Out)**, but it can also be LIFO (Last In First Out), priority-based, etc.
3. **Service Mechanism**
   - One or more **servers** provide service to jobs.
   - Service times are often modeled using an **exponential distribution** with mean service rate μ.
4. **Departure**
   - After service is completed, the job leaves the system.

Life of a Job in a Queuing System

- **Arrival** → enters system.
- **Waits** in the queue (if necessary).
- **Gets served** when a server is free.
- **Departs** after completion of service.

Steps in a Queuing System

1. **Arrival / Creation**
   A job is created (e.g., a user request or background task) and enters the queue.
2. **Queueing / Waiting**
   If all workers are occupied, the job remains in the queue until resources become available.
3. **Task Retrieval**
   When a worker is free, it fetches the next job from the queue, typically based on a scheduling policy like **FIFO (First-In-First-Out)**.
4. **Processing / Execution**
   The worker processes the job, which may involve computation, database queries, or communication with external services.
5. **Completion & Acknowledgement**
   After execution, the job is marked complete, and the system may notify the requester or update relevant records.
6. **Error Handling & Retries (Optional)**
   If processing fails, the job may be retried automatically or moved to a **Dead Letter Queue (DLQ)** for investigation.
7. **Scaling & Monitoring**
   To handle increased demand, the system can scale up workers. Performance metrics such as throughput, latency, and error rates are monitored for optimization.

Example (Bank Queue)

- Customers arrive at a bank counter.
- If the cashier is busy, they **wait in line**.
- When the cashier is free, the **next customer** is served.
- After finishing, the customer **leaves the bank**.

Auto-Regressive (AR) Process

An Auto-Regressive process of order *p* (AR(*p*)) is a time series model in which the current value depends linearly on its own previous values and a random noise term.

Mathematical Form:

$Xt=c+\phi1Xt-1+\phi2$

- *Xt*: Value at time t*t*
- *c*: Constant
- $\phi1,\phi2,...,\phi p$: Coefficients (weights for past values)
- ϵt: Random error (white noise)

Explanation:

- The AR process captures the persistence of previous values.
- For example, stock prices may depend on previous days' prices.

Example (AR(1)):

$Xt=c+\phi Xt-1+\epsilon t$

Moving Average (MA) Process

A Moving Average process of order q (MA(q*q*)) is a time series model where the current value depends on the mean of past q*q* error terms and a constant.

Mathematical Form:

$Xt=\mu+\theta1\epsilon t-1+\theta2\epsilon t-2+...+\theta q\epsilon t-q+\epsilon t$

- Xt: Value at time t
- M*:* Mean of the series (constant)
- $\theta1,...,\theta q$Coefficients (weights for past errors)
- $\epsilon t\epsilon t$: Random error (white noise)

Explanation:

- MA models the influence of random shocks/errors from previous time steps.
- For example, today's temperature could be affected by the last few days' weather disturbances.

Example (MA(1)):

$Xt=\mu+\theta\epsilon t-1+\epsilon t$

Bayesian Network

A **Bayesian Network (BN)** is a **probabilistic graphical model** that represents a set of random variables and their conditional dependencies using a **Directed Acyclic Graph (DAG)**.

- **Nodes** → random variables.
- **Edges** → conditional dependencies (cause-effect relations).
- Each node has a **Conditional Probability Distribution (CPD)** given its parent nodes.

Mathematically, a BN factorizes the joint probability distribution:

$$P(X_1, X_2, ..., X_n) = \prod_{i=1}^{n} P(X_i \mid \text{Parents}(X_i))$$

Uses / Advantages

1. **Reasoning under uncertainty** → helps in decision making with incomplete/noisy data.
2. **Efficient inference** → computes probabilities without needing the full joint distribution.
3. **Causal representation** → explicitly encodes cause-effect relationships.
4. **Applications** → medical diagnosis, spam filtering, genetics, fault detection, weather prediction.

Example

Medical Diagnosis:

- Nodes: Disease, Symptom1 (e.g., cough), Symptom2 (e.g., fever), RiskFactor (e.g., smoking).
- Edges: RiskFactor → Disease; Disease → Symptom1, Symptom2.
- Query: "Given cough and fever, what is the probability of having the disease?" The BN uses CPDs to compute this efficiently.

Hidden Markov Model (HMM) Transition State & Emission State Diagram:

An HMM has:

1. **Hidden states** – actual system conditions that we cannot observe directly.
2. **Transition probabilities** – how hidden states evolve over time.
3. **Emission probabilities** – how each hidden state produces an observable output.

So, there are two diagrams:

- **Transition State Diagram** → shows how hidden states change among themselves.
- **Emission State Diagram** → shows how hidden states generate observable symbols.

1. Transition State Diagram

- A directed graph where **nodes = hidden states** and **edges = transition probabilities**.
- Represents the **Markov property**: next state depends only on current state.

Example: Weather model
Hidden states:

- S={Sunny,Rainy}S = \{Sunny, Rainy\}S={Sunny,Rainy}

Transition probabilities:

P(Sunny→Sunny)=0.8
P(Sunny→Rainy)=0.
P(Rainy→Sunny)=0.4
P(Rainy→Rainy)=0.

```
[Sunny] --0.2--> [Rainy]
   ^ 0.8          | 0.6
   |              v
[Sunny] <--0.4-- [Rainy]
```

This means if today is Sunny, tomorrow has 80% chance Sunny, 20% chance Rainy.

2. Emission State Diagram

- Shows how each hidden state **produces observable outputs**.
- Each hidden state has arrows to observation symbols with certain probabilities.

Example (continuing Weather model):
Observations: Carry Umbrella (U) or No Umbrella (N)

Emission probabilities:

- If Sunny → P(U)=0.1,P(N)=0.
- If Rainy → P(U)=0.9,P(N)=0.

```
[Sunny] --> (Umbrella, 0.1)
   |--> (No Umbrella, 0.9)

[Rainy] --> (Umbrella, 0.9)
    |--> (No Umbrella, 0.1)
```

This means: if it is Rainy, most likely we will **see** an umbrella, even though the actual weather state is hidden.

Steps of Hypothesis Testing.

1. **State the Null and Alternative Hypotheses**
   - **Null hypothesis (H0*H*0)**: Assumes no effect or no difference (status quo).
   - **Alternative hypothesis (H1*H*1 or Ha*Ha*)**: Contradicts H0*H*0, suggests presence of an effect or difference. *Example:* H0*H*0: The new drug is not effective. Ha*Ha*: The new drug is effective.
2. **Choose the Significance Level (α*α*)**
   - Pre-decide the acceptable probability of making a Type I error (rejecting H0*H*0 when it's true).
   - Typical values: 0.05 (5%) or 0.01 (1%).
3. **Collect Data and Calculate Test Statistic**
   - Gather sample data relevant to the hypothesis.
   - Calculate an appropriate test statistic (e.g., z, t, chi-square) that measures how far the sample data diverges from H0*H*0.
4. **Determine the Critical Value or p-value**
   - **Critical value method:** Identify cutoff points for acceptance and rejection regions from statistical tables based on α*α*.
   - **p-value method:** Calculate the probability of observing a test statistic as extreme as the sample's under H0*H*0.

5. **Make a Decision**
    - **Using critical value:**
        - If test statistic is in the rejection region, reject H0$H0$.
        - Otherwise, fail to reject H0$H0$.
    - **Using p-value:**
        - If p≤α$p≤α$, reject H0$H0$.
        - If p>α$p>α$, fail to reject H0$H0$.
6. **Interpret and Report Results**
    - Present conclusions in the context of the original research question.
    - If H0$H0$ was rejected, results support the alternative hypothesis. If not rejected, there isn't sufficient evidence to support it.
    - Remember, failing to reject H0$H0$ does not prove it true.

Feature	T-Test	Z-Test	F-Test
Purpose	Compares means of two groups	Compares means/proportions	Compares variances between groups (e.g., ANOVA)
Sample Size	Small ($n < 30$)	Large ($n \geq 30$)	Any number of groups
Population Variance	Unknown	Known or large sample (by CLT)	Compares multiple variances
Distribution Used	t-distribution	Normal (z) distribution	F-distribution
Applications	Mean comparison (paired/unpaired)	Mean/proportion (one/two sample)	ANOVA, regression, equality of variance
Example	Compare average scores of two small classes	Compare average heights of two cities	Test if three fertilizers give different plant growth
Formula	$\frac{\bar{x}_1-\bar{x}_2}{s_{pooled}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$\frac{\bar{x}_1-\bar{x}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$\frac{s_1^2}{s_2^2}$ for two samples, or ANOVA F-statistics
Advantages	Good for small data, unknown variance	Simple for large data, known variance	Analyzes more than two groups simultaneously
Disadvantages	Less accurate for large samples, assumes normality	Not valid for small samples/unknown variance	Sensitive to non-normality, needs equal variances