V. Probability Distribution

Instructor: Yanlin Qi

Institute of Transportation Studies

Department of Statistics

University of California, Davis

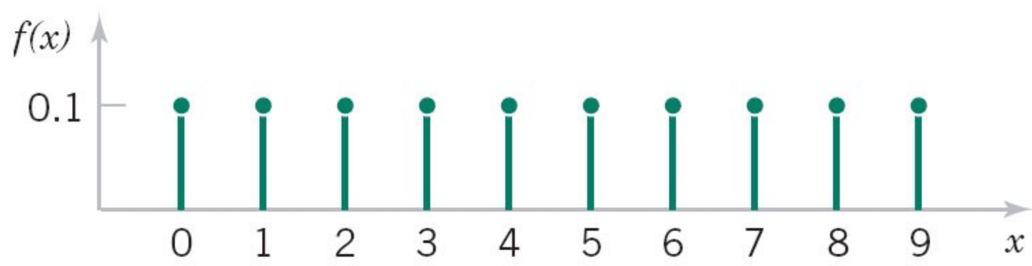
Today

- Discrete Uniform Distribution Section 3-5 and 4-5
- Binomial Distribution Section 3-6
- Related to binomial: Geometric, Negative Bin., Hypergeometric (S. 3-7, 3-8)
- Poisson Distribution Section 3-9
- Continuous Uniform Distribution Section 3-5 and 4-5
- Exponential Distribution
- Normal Distribution Section 4-6

Discrete Uniform Distribution

- The simplest discrete RV is one that assumes only a finite number of possible values, each with equal probability
- A RV X has a discrete uniform distribution if each of the n values in its range, say, $x_1, x_2, \dots x_n$, has equal probability

$$f(x_i) = 1/n$$



Mean and Variance of a DUD

- Suppose X is a discrete RV on the consecutive integers $a, a+1, a+2, \ldots, b$, for $a \le b$
- The mean of X is

$$\mu = E(X) = \frac{a+b}{2}$$

- The variance of X is

$$\sigma^2 = V(X) = \frac{(b-a+1)^2 - 1}{12}$$

Binomial Distribution

Binomial Distribution

Scenario:

- Consider an experiment like tossing a coin.
- Each toss is a **trial**.

Key Points:

- Trials are **independent** if the probability of success or failure does not change from one trial to the next.
- Example: Tossing a coin or rolling a die.

Definitions:

- p: Probability of success in a single trial
- $q=1-p$: Probability of failure in a single trial

Bernoulli Trials

- A **Bernoulli trial** (or binomial **trial**) is a random experiment with exactly two possible outcomes, "success" and "failure", in which **the probability of success is the same every time** the experiment is conducted

Binomial Distribution

- Scenario:
  - The probability of having exactly x successes in n trials:
  - PMF:

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

- $\binom{n}{x}$: Binomial coefficient, the number of ways to choose x successes (or outcomes) from n trials (or events)
- p^x: Probability of x successes
- q^{n-x} : Probability of n−x failures

Binomial Distribution - Example

- The probability of getting **exactly** 7 heads in 10 tosses of a fair coin is:

$$P(X=7)=\binom{10}{7}\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^3=\frac{10!}{7!(10-7)!}\left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^3$$

- Properties of Binomial Distribution
  - Mean: $\mu = np$

  The expected number of successes in the sample
  - Variance: $\sigma^2 = npq$

  The spread of the number of successes in the sample
  - Standard deviation: $\sigma = \sqrt{npq}$

Binomial Distribution - Example

- What is the expected value of heads if you flip a coin twelve times:

$$\mu = np \quad = (12)(0.5) = 6$$

- What is the variance:

$$\sigma^2 = npq \quad = (12)(0.5)(1.0 - 0.5) = 3.0$$

Geometric Distribution

Geometric Distribution

- **Similar to Binomial:**
  - **Bernoulli Trials:** Each trial has two possible outcomes (success or failure).
  - **Independent Trials:** The outcome of one trial does not affect the others.
- **However:**
  - **Not a Fixed Number of Trials:** The number of trials is not predetermined.
  - **Trials Continue Until Success:** We conduct the trials until the first success is obtained.
- **PMF:**

$$f(x) = (1-p)^{x-1}p$$

- Where:
- p: Probability of success on each trial.
- 1−p: Probability of failure on each trial.
- x: Number of trials until the first success

Geometric Distribution - Example

- P{a manufactured piece has a large contaminant particle}= 0.01
- Assume: Wafers are independent
- Question: What is the probability that **exactly 125 pieces** need to be analyzed before a large particle is detected?

Ans:

The 125th piece is the first one with a large particle

- X = number of samples analyzed until a large particle is detected (until a success is obtained)
  - X is geometric random variable
- P = 0.01

$$f(x) = (1-p)^{x-1}p$$

$$P(X=125) = (0.99)^{124} \times 0.01 = 0.0029$$

Geometric Distribution

- Mean

$$\mu = \frac{1}{p}$$

- Variance:

$$\sigma^2 = V(x) = {(1-p)}\big/{p^2}$$

Negative Binomial Distribution

Negative Binomial Distribution

- Generalization of Geometric Distribution
  - The random variable is the number of Bernoulli Trials needed to obtain *r* successes

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \qquad x = r, r+1, r+2, ..$$

- Mean $\mu = E(X) = {}^{r}\!/_{p}$

- Variance $\sigma^2 = V(x) = {}^{r(1-p)}\!/_{p^2}$

Negative Binomial Distribution - Example

- P{a bit received in error}=0.1
- Assume: Independent transmissions (trials)
- X = The number of bits transmitted until the **4th** error
- X has a negative binomial distribution with $r = 4$
- Q: What is P{the 4th error occur in the 10th trial}=$P(X = 10)$?
- Solution:
  - 3 errors happening anywhere in the first 9 trials
  - The final error in the last trial

$$\binom{10-1}{4-1}0.1^3 0.9^6 0.1 = \binom{9}{3}\,0.1^4 0.9^6$$

Hypergeometric Distribution

Hypergeometric Distribution

- Used to calculate probabilities when sampling without replacement
- Example:
  - Pick 3 cards without replacement. What is that probability that both cards will be aces?

Hypergeometric Distribution

❖ **Key Variables:**

- p: Probability of obtaining k successes
- k: Number of "successes" in the population
- x: Number of "successes" in the sample
- N: Size of the population
- n: Number of items sampled

❖ **Distribution Function:**

$$f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \text{ with } x = \max\{0, n+k-N\} \text{ to } \min\{k, n\}$$

Example 14

- Pick 3 cards from a deck without replacement. What is that probability that 2 cards will be aces?

p = ?
k = 4
x = 2
N = 52
n = 3

$$f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$p = \frac{({}_4C_2)\left({}_{(52\ -\ 4)}C_{(3\ -\ 2)}\right)}{{}_{52}C_3}$$

$$p = \frac{\frac{4!}{2!2!}\ \frac{48!}{47!1!}}{\frac{52!}{49!3!}} = 0.013$$

The Hypergeometric Distribution

$$f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \quad \text{with } \; x = \max\{0, n+k-N\} \text{ to } \min\{k, n\}$$

- **Mean** $\mu = E(X) = np$

$$p = \frac{k}{N}$$

- **Variance** $\sigma^2 = V(x) = np(1-p)\left(\frac{N-n}{N-1}\right)$

Summary

Binomial	The probability that the event will happen exactly x times in n trials
Geometric	The probability that a number of x trials is needed until a success is obtained
Negative Binomial	The probability that a number of x trials are needed to obtain r successes
Hypergeometric	The probability of x success is n trials/sample, from a population of size N, which has k successes (without replacement)

Poisson Distribution

Poisson Distribution

- **Definition:**
  - The Poisson distribution describes the probability of a given number of events occurring in a fixed interval of time or space.
- **Key Characteristics:**
  - Events occur randomly and independently
  - The average rate (λ) of events is known and constant
  - The time between events follows an exponential distribution
- **PMF:**

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Where:

•e≈2.718

•x: Number of events (0, 1, 2, ...)

•λ: Average rate of events

Example 3

- Given
  - Average phone calls per day = 8
  - Find the probability of 11 calls in a day

- Ans.

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

  - e = 2.718
  - λ = 8
  - x =11 (the number of successes)

$$P(X = 11) = \frac{e^{-8} \cdot 8^{11}}{11!} \approx 0.072$$

Poisson Distribution - Example

- **Example:** Earthquake Occurrences
- **Scenario:**
  - Average number of earthquakes per year: 2.3
- **Problem:**
  - Calculate the probability of having exactly 2 earthquakes in 1 year.
- **Solution:** $P(X=2) = \frac{e^{-2.3} \cdot 2.3^2}{2!} = 0.265$
- **Important Note:** $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$
  - **Use Consistent Units**
    - Average earthquakes per year: $\lambda = 2.3$
    - Average earthquakes in 2 years: $\lambda = 2.3 * 2 = 4.6$
    - Average earthquakes in 5 years: $\lambda = 2.3 * 5 = 11.5$

Mean and Variance of a Poisson RV

- The mean of a Poisson RV is $\mu = E(X) = \lambda$

- The variance of a Poisson RV is $\sigma^2 = V(X) = \lambda$

- Mean=variance for Poisson RVs
- If a phenomenon follows a Poisson distribution, information on the variance is very easily obtained

  - If the variance differs significantly from the mean, then the phenomenon cannot be modeled with a Poisson distribution

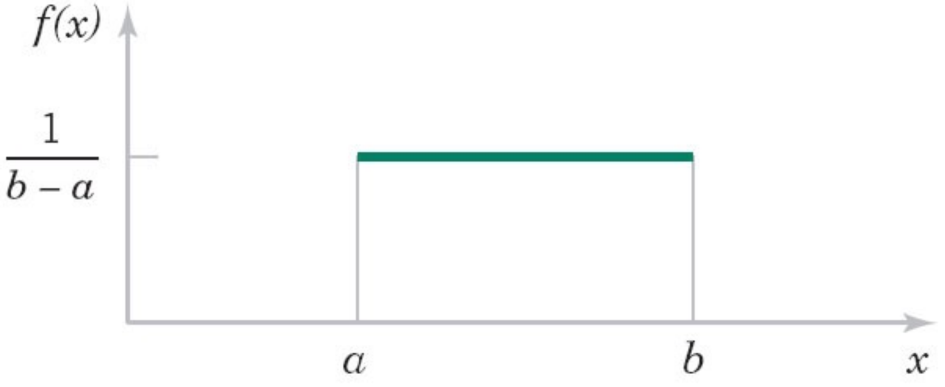
Probability Density Functions

Uniform Distribution

Continuous Uniform Distribution

- The simplest continuous RV is one that assumes values within a given interval, with equal probability
- A RV X has a continuous uniform distribution if its values have equal probability density of occurrence

$$f(x_i) = 1/(b - a)$$



Mean and Variance of a CUD

- Suppose X is a continuous RV over $a \le x \le b$
- The mean of X is

$$\mu = E(X) = \frac{a+b}{2}$$

- The variance of X is

$$\sigma^2 = V(x) = \frac{(b-a)^2}{12}$$

- The CDF is

$$F(X) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \le x < b \\ 1 & b \le x \end{cases}$$

CUD - Example

- **Scenario:**
  - Suppose we have a continuous random variable X that is uniformly distributed over the interval [2, 10].
- **Given:**
  - Lower bound a=2
  - Upper bound b = 10
- Question: To find the probability that X is less than or equal to 5:

$$F(X) = \begin{cases} 0 & \text{if } x < 2 \\ \frac{x-2}{10-2} = \frac{x-2}{8} & \text{if } 2 \leq x < 10 \\ 1 & \text{if } x \geq 10 \end{cases}$$

$$F(5) = \frac{5-2}{10-2} = \frac{3}{8} = 0.375$$

- Interpretation:
  - There is a 37.5% chance that X will be less than or equal to 5.

Exponential Distribution

Exponential Distribution

- The RV X represents the time between events in a Poisson process with a mean rate λ.
- PDF:

$$f(x) = \lambda e^{-\lambda x} \qquad 0 \le x \le \infty$$

- Mean

$$\mu = E(X) = \frac{1}{\lambda}$$

- Variance

$$\sigma^2 = V(X) = \frac{1}{\lambda^2}$$

Exponential Distribution - Example

- **Scenario:**
  - Suppose observing the time between arrivals of buses at a bus stop.
  - The average arrival rate of buses (λ) is 3 buses per hour.
- **Goal:**
  - Find the probability that the next bus will arrive within 10 minutes (which is 1/6 hour).
- **Given:**
  - λ=3 buses per hour
  - x=1/6 hour

Exponential Distribution - Example

❖ Calculation:

1. **Convert time to the same units:**
   1. Time x in hours = 10 minutes = 1/6 hour
2. **Calculate the CDF:**

$$P(X \leq x) = F(x) = 1 - e^{-\lambda x}$$

3. **Plug the values into the formula:**
   1. Plug the values into the formula:

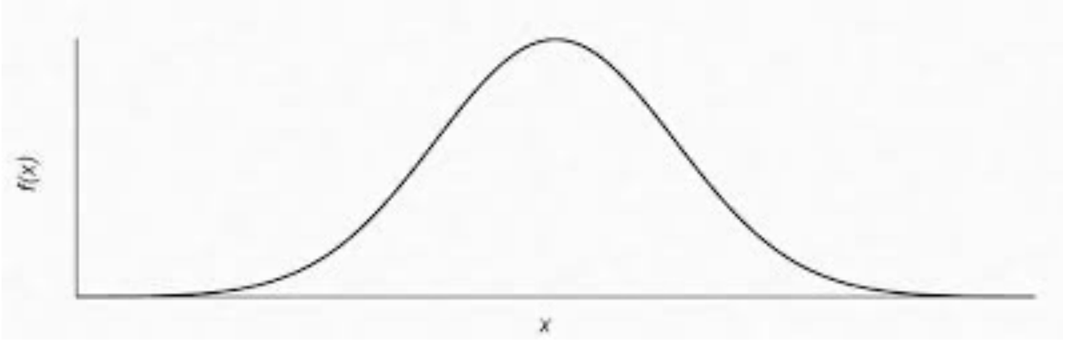
$$P(X \leq \frac{1}{6}) = 1 - e^{-3 \cdot \frac{1}{6}} = 1 - e^{-0.5}$$

Normal Distribution

Normal Distribution

Definition:

- A continuous probability distribution that is symmetrical around its mean.
- Demonstrates that data near the mean are more frequent in occurrence than data far from the mean.



Example:

- **Height of People:** Follows a normal distribution with most heights around the mean.

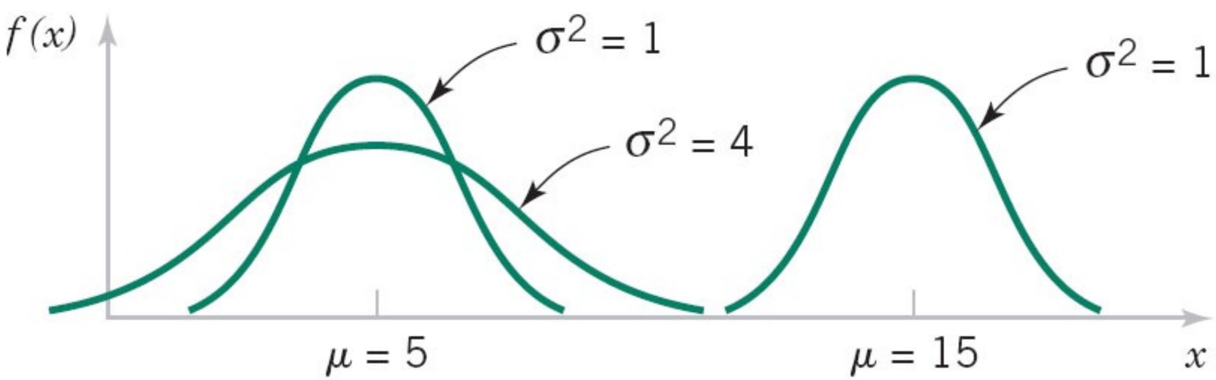
Normal (Gaussian) PDF

- The PDF of a normal RV is

$$f(X) = N(\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- with

$$E(X) = \mu \qquad V(X) = \sigma^2$$



Normal Distribution

- **Characteristics:**
  - **Symmetrical Shape:** Perfectly symmetrical around the mean.
  - **Bell Curve:** The graph is bell-shaped.
  - **Mean, Median, Mode:** All are equal in a normal distribution.
  - **Asymptotic:** Tails approach the horizontal axis but never touch it.
  - **Defined by Mean and Standard Deviation:** Characterized by mean (μ) and standard deviation (σ).

Standard Normal Distribution

- Special case with mean 0 and standard deviation 1
- PDF:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

- A standard normal RV is often denoted as Z
- The CDF of a normal RV is denoted as

$$\Phi(z) = P(Z \le z)$$

- Values for $\Phi(z)$ can be found in many Normal tables

Standard Normal Probabilities

Standard Normal Probabilities

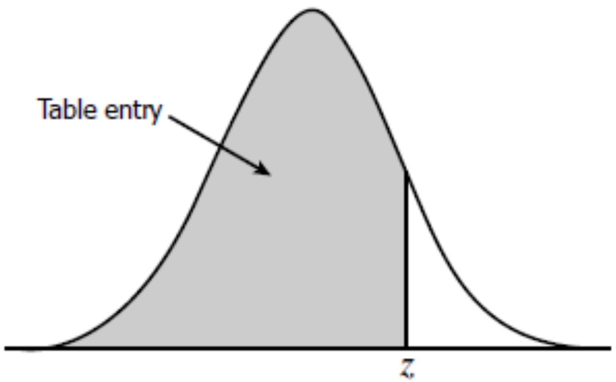


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177

Standard Normal Distribution

- Special case with mean 0 and standard deviation 1
- PDF:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

- **Properties:**
  - **68-95-99.7 Rule:**
    - 68% within one standard deviation
    - 95% within two standard deviations
    - 99.7% within three standard deviations

Standard Normal Distribution - Example

- **Example:**
- To find the cumulative probability for z=1.23:
  1. Z-score: 1.23
  2. Row: Look for 1.2
  3. Column: Look for 0.03
  4. Table Entry: The value at the intersection of the row (1.2) and the column (0.03) is approximately 0.8907.
- **Interpretation:**
  - This means that approximately 89.07% of the data lies to the left of z=1.23.

Standardizing a Normal RV

- To use the same Table for any $X = N(\mu,\sigma^2)$ we should use a simple transformation
- If X is a NRV with $E(X) = \mu$ and $V(X) = \sigma^2$, the RV

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

  is a NRV with $E(Z) = 0$ and $V(Z) = 1 \rightarrow Z = N(0,1)$
- That is, Z is a standard normal RV (SNRV)
  - It is the distance of X from its mean in terms of standard deviations
  - It is the key step to calculate a probability for an arbitrary NRV

Central Limit Theorem (CLT)

Definition:

- The sum (or average) of a large number of independent and identically distributed (i.i.d.) random variables will be approximately normally distributed.
- This holds regardless of the original distribution of the variables.

Key Points:

1. **Independence:** Variables must be independent.
2. **Identical Distribution:** Variables should be identically distributed (i.i.d.).
3. **Finite Mean and Variance:** Each variable must have a finite mean (μ) and variance (σ^2).
4. **Large Sample Size:** Distribution approaches normal as sample size increases.

Central Limit Theorem (CLT)

Mathematical Formulation:

- Let $X1, X2, \ldots, Xn$ be i.i.d. variables with mean μ and variance σ^2.
- Sum $Sn = X1 + X2 + \cdots + Xn$ is approximately normally distributed for large n.

Standardized Version:

$$Z = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

- As n approaches infinity, Z approaches $N(0,1)$.

Implications:

- CLT allows inferences about sums (or averages) of large samples, even if original data isn't normally distributed.
- Underpins many statistical methods (e.g., hypothesis testing, confidence intervals).