Set Theory

$x \in A \cup B \iff x \in A$ or $x \in B$

$x \in A \cap B \iff x \in A$ and $x \in B$

$x \in A' \iff x \notin A$

Mutual Exclusivity

A collection of sets $A_1, \ldots, A_n$ is mutually exclusive $\iff A_i \cap A_j = \varnothing \; \forall \, i \neq j$

Collective Exhaustiveness

A collection of sets $A_1, \ldots, A_n$ is collectively exhaustive $\iff A_1 \cup \ldots \cup A_n = S$

Partitions

A partition is a set which is mutually exclusive and collectively exhaustive.

De Morgan's Law

$(A \cup B)' = A' \cap B'$ $\qquad$ $(A \cap B)' = A' \cup B'$

Axioms of Probability

1. For any event A, $P[A] \geq 0$
2. $P[S] = 1$
3. For countable collection $A = A_1, \ldots, A_n$ of mutually exclusive events,

$P[A] = \sum_{i=1}^{n} P[A_i]$

Probability Measure

$P[\varnothing] = 0$

$P[A'] = 1 - P[A]$

$P[A \cup B] = P[A] + P[B] - P[A \cap B] \; \forall \, A, B$

$A \subset B \implies P[A] \leq P[B]$

Theorem 4.18

$P[X = x_0] = P_X(x_0) = F_X(x_0^+) - F_X(x_0^-) = \frac{f_X(x_0)}{\delta(0)}$

Theorem 5.2

$P[x_1 < X \leq x_2, y_1 < Y \leq y_2] = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1)$

$P_{\mathbf{X}}(\mathbf{x}) = P[\mathbf{X} = \mathbf{x}]$, $\quad$ For vector of points $\mathbf{x}$

Theorem 5.3

For a set of discrete random variables $\mathbf{X}$ and any set B on the $\mathbf{X}$ domain, the probability of the event $\{\mathbf{X} \in B\}$ is

$P[B] = \sum_{\mathbf{x} \in B} P_{\mathbf{X}}(\mathbf{x})$

Theorem 5.7

The probability that the continuous random variables $\mathbf{X}$ are in $\mathbf{A}$

$P[A] = \int \int_A f_{\mathbf{X}}(x) \, d\mathbf{x}$

Conditional Probability

Probability of A given B:

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \qquad \text{Kolmogrov Definition}$$

$$= \frac{P[B|A]P[A]}{P[B]} \qquad \text{Bayes' Theorem}$$

$P[A|B] \geq 0 \quad P[A|A] = 1$

For countable collection $A = A_1, \ldots, A_n$ of mutually exclusive events,

$P[A|B] = \sum_{i=1}^{n} P[A_i|B]$

Law of Total Probability

For a partition $\{B_1, \ldots, B_n\}$ with $P[B_i] > 0 \; \forall \, i$, $P[A] = \sum_{i=1}^{n} P[A|B_i] \cdot P[B_i]$

Outcome

Any possible observation of an experiment.

Sample Space

Partition of all possible outcomes.

For $S = \{s_1, \ldots, s_n\}$, where each outcome s_i is equally likely,

$P[s_i] = 1/n, \quad 1 \leq i \leq n$

Event

A set of outcomes of an experiment.

Independent Events

Events A, B are independent $\iff P[A \cap B] = P[A] \cdot P[B]$

Mutual Independence

$$P[A_1 \cap \ldots \cap A_n] = \prod_{i=1}^{n} P[A_i]$$

Fundamental Theorem of Counting

An experiment with n subexperiments s has $\prod_{i=1}^{n} s_i$ outcomes.

Permutations

The number of k-permutations of n distinguishable objects is

$$(n)_k = n(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

Choose

Number of ways to choose k objects out of n distinguishable objects:

Without Replacement:	$\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}$	$n, k \in \mathbb{N} \quad 0 \leq k \leq n \in \mathbb{N}$
With Replacement:	n^k	

Number of ways to choose $n_0, \ldots, n_{m-1}$ of each subtype, out of n total objects:

$$\binom{n}{n_0, \ldots, n_{m-1}} = \frac{n!}{n_0! \cdot \ldots \cdot n_{m-1}!}$$

Random Variables

A random variable consists of an experiment with a probability measure defined on a sample space and a function that assigns a real number to each outcome in the sample space of the experiment. The set of possible values of a random variable is the range of that random variable.

Properties of Random Variables

- $F_X(x) = P[X \leq x]$.
- If $F_X(x)$ is piecewise flat with discontinuous jumps, then X is discrete.
- If $F_X(x)$ is continuous, then X is continuous.
- If $F_X(x)$ is piecewise continuous with discontinuities, then X is mixed.
- When X is discrete or mixed, $f_X(x)$ contains delta functions.

Derived Random Variable

Random variable $Y = g(X)$

Probability mass function of derived random variable $Y = g(x)$:

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x)$$

Discrete Random Variables

A random variable is discrete if its range is a countable set.

Continuous Random Variables

X is a continuous random variable if its Cumulative Distribution Function is continuous.

Mixed Random Variables

X is a mixed random variable $\iff f_X(x)$ contains both impulses and nonzero, finite values.

Random Vector

$\mathbf{X} = [X_1, \ldots, X_n]$

Unit Step Function

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$$

Unit Impulse Function

$$d_\epsilon(x) = \begin{cases} 1/\epsilon & -\epsilon/2 \leq x \leq \epsilon/2 \\ 0 & x < -\epsilon/2,\ x > \epsilon/2 \end{cases}$$

Delta Function

$$\delta(x) = \lim_{\epsilon \to 0} \begin{cases} 1/\epsilon & -\epsilon/2 \le x \le \epsilon/2 \\ 0 & x < -\epsilon/2,\ x > \epsilon/2 \end{cases}$$

Theorem 4.17

$$\int_{-\infty}^{x} \delta(v)\, dv = u(x)$$

Theorem 4.16

For any continuous function $g(x)$,

$$\int_{-\infty}^{\infty} g(x)\, \delta(x - x_0)\, dx = g(x_0)$$

Probability Density Function

$$f_X(x) = \frac{dF_X(x)}{dx} = \sum_{x_i \in S_X} P_X(x_i)\, \delta(x - x_i)$$

Theorem 4.2

For a continuous random variable X with probability density function $f_X(x)$

1. $f_X(x) \ge 0 \; \forall \; x$
2. $F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$
3. $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$

Theorem 4.3

$$P[x_1 < X \le x_2] = \int_{x_1}^{x_2} f_X(x)\, dx$$

Joint Probability Density Function

For a set of continuous random variables $\mathbf{X}$

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$$

Theorem 5.6

$f_{\mathbf{X}}(\mathbf{x}) \ge 0 \; \forall \; \mathbf{x}$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} = 1$$

Theorem 5.8

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, dx_2 \ldots dx_n$$

Probability Mass Function

Discrete subcase of the Probability Distribution Function.

Probability mass function P of the discrete random variable X:

$P_X(x) = P[X = x]$.

Theorem 5.4

For a set of discrete random variables $\mathbf{X}$ with joint probability mass function $P_{\mathbf{X}}(\mathbf{x})$,

$$P_{X_i}(x_i) = \sum_{j \ne i} \sum_{x_j \in S_{X_j}} P_{\mathbf{X}}(\mathbf{x})$$

Theorem 5.5

Joint Probability Mass Function

$$f_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}}^{(n)}(\mathbf{x})$$

Theorem 3.1:

1. For any x, $f_X(x) \ge 0$
2. $\sum_{x \in S_X} f_X(x) = 1$
3. For any event $B \subset S_X$, the probability that X is in the set B is

   $P[B] = \sum_{x \in B} f_X(x)$

Cumulative Distribution Function

Cumulative distribution function of random variable X:

$$F_X(x) = P[X \le x] = \sum_{x_i \in S_X} P_X(x_i) u(x - x_i)$$

Discrete Cumulative Distribution Function

Theorem 3.2

For any discrete random variable X with range $S_X = \{x_1, x_2, \ldots\}$ satisfying $x_n \leq x_{n+1}$,

(a) $F_X(-\infty) = 0, \quad F_X(\infty) = 1$

(b) $x_2 \geq x_1 \iff F_X(x_2) \geq F_X(x_1)$.

(c) For $x_i \in S_X$ and $\epsilon > 0$, $F_X(x_i) - F_X(x_i - \epsilon) = P_X(x_i)$

(d) $F_X(x) = F_X(x_i) \; \forall \; x : x_i \leq x < x_{i+1}$

Theorem 3.3

$F_X(b) - F_X(a) = P[a < X \leq b] \; \forall \; b \geq a$

Continuous Cumulative Distribution Function

Theorem 4.1

For any random variable X

(a) $F_X(-\infty) = 0$

(b) $F_X(\infty) = 1$

(c) $P[x_1 < X \leq x_2] = F_X(x_2) - F_X(x_1)$

Joint Cumulative Distribution Function

$F_{\mathbf{X}}(\mathbf{x}) = P[X_1 \leq x_1, \ldots, X_n \leq x_n]$

Theorem 5.1

$0 \leq F_{\mathbf{X}}(\mathbf{x}) \leq 1$

$F_{\mathbf{X}}(\infty, \ldots, \infty) = 1$

$F_{X_1}(x_1) = F_{\mathbf{X}}(x_1, \infty, \ldots, \infty)$

$F_{\mathbf{X}}(x_1, \ldots, -\infty, \ldots, x_n) = 0$

$F_{\mathbf{X}}(\mathbf{x}_1) \leq F_{\mathbf{X}}(\mathbf{x}_2), \quad \mathbf{x}_1 \leq \mathbf{x}_2$

Expected Value (Mean)

Given a random variable X with probability mass function $P_X(x)$ and the derived random variable $Y = g(X)$, the expected value of Y is:

$E[Y] = \sum_{x \in S_X} g(x) P_X(x)$

$E[X - \mu_X] = 0$

$E[aX + b] = aE[X] + b$

Mean (Expected Value)

$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x)$

Continuous Expected Value

The expected value of a continuous random variable X is

$E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx$

Theorem 4.4

The expected value of a function $g(X)$ of random variable X

$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) \, dx$

Theorem 4.5

For any random variable X,

(a) $E[X - \mu_X] = 0$

(b) $E[aX + b] = aE[X] + b$

(c) $\mathrm{Var}[X] = E[X^2] - \mu_X^2$

(d) $\mathrm{Var}[aX + b] = a^2 \, \mathrm{Var}[X]$

Multivariate Expected Value

$W = g(\mathbf{x})$

Discrete

$$E[W] = \sum_{x_1 \in S_{X_1}} \cdots \sum_{x_n \in S_{X_n}} g(\mathbf{x}) P_{\mathbf{X}}(\mathbf{x})$$

Continuous

$$E[W] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$$

Theorem 5.10

$$E\left[\sum_{i=1}^{n} a_i g_i(\mathbf{X})\right] = \sum_{i=1}^{n} a_i E[g_i(\mathbf{X})]$$

Theorem 5.11

$$E[X+Y] = E[X] + E[Y]$$

Mode

Mode x_{mod} of X satisfies $P_X(x_{\text{mod}}) \geq P_X(x)\ \forall\ x$

Median

Median x_{med} of X satisfies:

$$P[X \leq x_{\text{med}}] \geq 0.5, \qquad P[X \geq x_{\text{med}}] \geq 0.5$$

Variance

$$\text{Var}[X] = E\left[(X - \mu_X)^2\right] = E[X^2] - E[x]^2$$

Theorem 5.12

$$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2E[(X-\mu_X)(Y-\mu_Y)] = \text{Var}[X] + \text{Var}[Y] + 2\,\text{Cov}[X,Y]$$

Covariance

$$\text{Cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = r_{X,Y} - \mu_X \mu_Y$$

Correlation Coefficient

$$\rho_{X,Y} = \frac{\text{Cov}[X,Y]}{\sqrt{\text{Var}[X]\,\text{Var}[Y]}} = \frac{\text{Cov}[X,Y]}{\sigma_X \sigma_Y}$$

Theorem 5.13

If $\hat{X} = aX + b$ and $\hat{Y} = cY + d$,

$$\rho_{\hat{X},\hat{Y}} = \rho_{X,Y}$$

$$\text{Cov}[\hat{X},\hat{Y}] = ac\,\text{Cov}[X,Y]$$

Theorem 5.14

$$-1 \leq \rho_{X,Y} \leq 1$$

Theorem 5.15

If X and Y are random variables such that $Y = aX + b$,

$$\rho_{X,Y} = \begin{cases} -1 & a < 0 \\ 0 & a = 0 \\ 1 & a > 0 \end{cases}$$

Correlation

$$r_{X,Y} = E[XY]$$

Theorem 5.16

If $X = Y$, $\text{Cov}[X,Y] = \text{Var}[X] = \text{Var}[Y]$

$$r_{X,Y} = E[X^2] = E[Y^2]$$

Orthogonal Random Variables

X and Y are orthogonal $\iff r_{X,Y} = 0$

Uncorrelated Random Variables

X and Y are uncorrelated $\iff \text{Cov}[X,Y] = 0$

Theorem 5.17

For independent random variables X and Y,

$$E[g(X)h(Y)] = E[g(X)]\,E[h(Y)]$$

$$r_{X,Y} = E[X]\,E[Y]$$

$$\text{Cov}[X,Y] = \rho_{X,Y} = 0$$

$$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$$

Independent $\implies$ Uncorrelated

Standard Deviation

$$\sigma_X = \sqrt{\text{Var}[X]}$$

Families of Discrete Random Variables

Family	$f_X(x)$	$E[x]$	$\mathrm{Var}[x]$
Bernoulli Probability of success/failure	$\begin{cases} 1-p & x=0 \\ p & x=1 \end{cases}$	p	$p(1-p)$
Pascal Number of trials x until kth success.	$\binom{x-1}{k-1}p^k(1-p)^{x-k}$	$\frac{k}{p}$	$\frac{k(1-p)}{p^2}$
Geometric Number of trials x until 1st success.	$p(1-p)^{x-1}$	$\frac{1}{p}$	$\frac{(1-p)}{p^2}$
Poisson Probability of x events given an interval of time.	$\frac{\alpha^x e^{-\alpha}}{x!}$	α	α
Binomial Probability of k successes in n trials.	$\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
Discrete Uniform Probability of n equally likely events.	$1/n$	$\frac{(n+1)}{2}$	$\frac{n^2-1}{12}$

$\alpha = (\text{Rate}) \times (\text{Time Interval})$

Theorem 3.8

Perform n Bernoulli trials. In each trial, let the probability of success be α/n, where $\alpha > 0$ is a constant and $n > \alpha$. Let the random variable K_n be the number of successes in the n trials. As $n \to \infty$, $f_{K_n}(k)$ converges to the probability mass function of a Poisson (α) random variable.

Families of Continuous Random Variables

	$f_X(x)$	$F_X(x)$	$E[X]$	$\text{Var}[X]$
Uniform Probability of equally likely events.	$\begin{cases} \frac{1}{b-a} & a \leq x < b \\ 0 & x < a,\ x \geq b \end{cases}$	$\begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x \leq b \\ 1 & x > b \end{cases}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$
Exponential Probability that an event happens at time t	$\begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$	$\begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Erlang	$\begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\begin{cases} 1 - e^{-\lambda t} & x \geq 0 \\ 0 & t < 0 \end{cases}$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
Gaussian (Normal)	$\frac{\exp(-\frac{(x-\mu)^2}{2\sigma^2})}{\sqrt{2\pi\sigma^2}}$	$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$	μ	σ^2
Standard Normal Gaussian with a mean of 0 and a variance of 1	$\frac{e^{-x^2/2}}{\sqrt{2\pi}}$	$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$	0	1

$b > a$, Arrival rate $\lambda > 0$, $n \geq 1 \in \mathbb{N}$

Theorem 4.7

Let X be a uniform (a, b) random variable, where a and b are both integers. Let $K = \lceil X \rceil$. Then K is a discrete uniform $(a+1, b)$ random variable.

Theorem 4.9

If X is an exponential random variable, then $K = \lceil X \rceil$ is a geometric random variable with $p = 1 - e^{-\lambda}$.

Theorem 4.11

Let K_α denote a Poisson random variable. For any $x > 0$, the CDF $F_X(x)$ of an Erlang random variable X satisfies:

$$F_X = 1 - F_{K_{\lambda x}}(n-1) = \begin{cases} 1 - \sum_{k=0}^{n-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Theorem 4.13

If X is a Gaussian, $Y = aX + b$ is a Gaussian

Theorem 4.14

If X is a Gaussian random variable, the Cumulative Distribution Function of X is

$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

The probability that X is in the interval $(a, b]$ is:

$P[a < X \leq b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

Theorem 4.15

$\Phi(-x) = 1 - \Phi(x)$

Standard Normal Complementary Cumulative Distribution Function

$$Q(z) = P[Z > z] = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-u^2/2}\, du = 1 - \Phi(z)$$

Bivariate Gaussian Random Variables

$$f_{X,Y}(x,y) = \frac{\exp\left(-\frac{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{2\rho_{X,Y}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}{2(1-\rho_{X,Y}^2)}\right)}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}$$

Theorem 5.20

Bivariate Gaussian random variables are uncorrelated if and only if they are independent.

Theorem 5.21

If X and Y are bivariate Gaussian random variables, and

$W_1 = a_1X + b_1Y, \quad W_2 = a_2X + b_2Y$

Then for $i = 1, 2$

$E[W_i] = a_i\mu_X + b_i\mu_Y$

$\text{Var}[W_i] = a_i^2\sigma_X^2 + b_i^2\sigma_Y^2 + 2a_ib_i\rho_{X,Y}\sigma_X\sigma_Y$

$\text{Cov}[W_1, W_2] = a_1a_2\sigma_X^2 + b_1b_2\sigma_Y^2 + (a_1b_2 + a_2b_1)\rho_{X,Y}\sigma_X\sigma_Y$