probability and stochastic processes

1. experiments, models, and probabilities

theorem 1.1 demorgan's law related all three basic operations $(A \cup B)^c = (A^c \cap B^c)$

theorem 1.2 for mutually exclusive events A_1 and A_2, $P[A_1 \cup A_2] = P[A_1] + P[A_2]$

theorem 1.3
$textIfA = A_1 \cup A_2 \cup \cdots \cup A_m$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$, then

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

theorem 1.4 The probability measure $P[.]$ is a function that satisfies the following properties:

- $P[\emptyset] = 0$
- $P[A^c] = 1 - P[A]$
- For any A and B (not necessarily mutually exclusive), $P[A \cup B] = P[A] + P[B] - P[A \cap B]$
- $A \subset B, P[A] \leq P[B]$

Theorem 1.5 The probability of an event $B = s_1, s_2, \cdots, s_m$ is the sum of the probabilities of the outcomes contained in the event:

$$P[B] = \sum_{i=1}^{m} P[s_i]$$

theorem 1.6 For an experiment with sample space $S = s_1, s_2, \cdots, s_n$ in which each outcomes s_i is equally likely,

$$P[s_i] = \frac{1}{n} \quad 1 \leq i \leq n$$

theroem 1.7 A conditional probability measure $P[A|B]$ has the following properties that correspond to the axioms of probability:

Axiom 1: $P[A|B] \geq 0$

Axiom 2: $P[B|B] = 1$

Axiom 3: If $A = A_1 \cup A_2 \cup \cdots \cup A_m$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$, then

$$P[A|B] = P[A_1|B] + P[A_2|B] + \cdots + P[A_m|B]$$

Theorem 1.8 For a partition $B = B_1, B_2, \cdots, B_m$ and any event A in the sample space, let $C_i = A \cap B_i$ For $i \neq j$, the events C_i and C_j are mutually exclusive and $A = C_1 \cup C_2 \cup \cdots$

Theorem 1.9 For any event A and partition $B_1, B_2, \cdots, B_m$

$$P[A] = \sum_{i=1}^{m} P[A \cap B_i]$$

Theorem 1.10 Law of total probability

For a partition $B_1, B_2, \cdots, B_m$ with $P[B_i] > 0$ for all i,

$$P[A] = \sum_{i=1}^{m} P[A|B_i]P[B_i]$$

Theorem 1.11 Bayes' theorem

$$P[B|A] = \frac{P[A|B]P[B]}{P[A]}$$

Definition 1.1 Outcome An outcome of an experiment is a possible result of the experiment.

Definition 1.2 Sample space The sample space of an experiment is the finest-grain, mutually exclusive, collectively exhaustive set of all possible outcomes of the experiment.

Definition 1.3 Event An event is a subset of the sample space.

Definition 1.4 Axioms of Probability A probability measure $P[.]$ is a function that maps events in the sample spacce to real numbers such that

Axiom 1 For any event A, $P[A] \geq 0$

Axiom 2 $P[S] = 1$

Axiom 3 For any countable collection $A_1, A_2, \cdots$ of mutually exclusive events,

$$P[A_1 \cup A_2 \cup \cdots] = P[A_1] + P[A_2] + \cdots$$

Definition 1.5 Conditional probability The conditional probability of an event A given the occurance of the event B is

$$P[A|B] = \frac{P[AB]}{P[B]}$$

Conditional probability is defined only when $P[B] > 0$.

Definition 1.6 Two independent events Two events A and B are independent if

$$P[AB] = P[A]P[B]$$

Definition 1.7 Three Independent Events A_1, A_2, A_3 are mutually exclusive and independent if and only if

(a) A_1 and A_2 are independent

(b) A_2 and A_3 are independent

(c) A_1 and A_3 are independent

(d) $P[A_1 \cap A_2 \cap A_3] = P[A_1]P[A_2]P[A_3]$

Definition 1.8 More than Two Independent Events

If $n \geq 3$ events $A_1, A_2, \cdots, A_n$ are mutually independent if an only if

(a) all collections of $n-1$ events chosen from $A_1, A_2, \cdots, A_n$ are mutually independent,

(b) $P[A_1 \cap A_2 \cap \cdots \cap A_n] = P[A_1]P[A_2] \cdots P[A_n]$

2. Sequential Experiments

Theorem 2.1 An experiment consists of two subexperiments. If one subexperiment has k outcomes and the other has n outcomes, then the experiment has kn outcomes.

Theorem 2.2 The number of k-permutations of n distinguishable objects is

$$\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}$$

Theorem 2.4 Given m distinguishable objects, there are m^n ways to choose ith replacement an ordered sample of n objects.

Theorem 2.5 For n repitions of a subexperiment with sample space $S_sub = s_1, s_2, \cdots, s_m - 1$, the sample space S of the sequential experiment has m^n outcomes.

Theorem 2.6 The number of observation sequences for n subexperiments with sample space $S = 0, 1$ with 0 appearing n_0 times and 1 appearing $n_1 = n - n_0$ times is $\binom{n}{n_1}$.

Theorem 2.7 For n reptitions of a subexperiment with sample space $S = s_0, s_1, \cdots, s_m - 1$, the number of length $n = n_0 + n_1 + \cdots + n_{m-1}$ observation sequences with s_i appearing n_i times is

$$\binom{n}{n_0, n_1, \cdots, n_{m-1}} = \frac{n!}{n_0! n_1! \cdots n_{m-1}!}$$

Theorem 2.8 The probability of n_0 failures and n_1 successes in $n = n_0 + n_1$ independent trials is

$$P[E_{n_0, n_1}] = \binom{n}{n_1}(1-p)^{n-n_1}p_1^{n_1} = \binom{n}{n_0}(1-p)^{n_0}p^{n-n_0}$$

Theorem 2.9 A subexperiment has sample space $S = s_0, s_1, \cdots, s_m - 1$ with $P[s_i] = p_i$ for $n = n_0 + n_1 + \cdots + n_{m-1}$ independent trials, the probability of n_i occurrences of s_i, $i = 0, 1, \cdots, m-1$ is

$$P[E_{n_0, n_1, \cdots, n_{m-1}}] = \binom{n}{n_0, n_1, \cdots, n_{m-1}} p_0^{n_0} p_1^{n_1} \cdots p_{m-1}^{n_{m-1}}$$

Definition 2.1 n **choose** k For an integer $n \geq 0$, we define

$$\binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!} & k = 0, 1, \ldots, n, \\ 0 & \text{otherwise} \end{cases}$$

Definition 2.2 Multinomial coefficient For an integer $\mathrm{n} \geq 0$, we define

$$\binom{n}{n_0, n_1, \ldots, n_{m-1}} = \frac{n!}{n_0! n_1! \cdots n_{m-1}!}$$

3. Discrete Random Variables

Theorem 3.1 For a discrete random variable X with PMF $P_X(x)$ and range S_X :

(a) For any x, $P_X(x) \geq 0$

(b) $\sum_{x \in S_x} P_X(x) = 1$

(c) For any event $B \subset S_x$, The probability that X is in the set B is

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

Theorem 3.2 For any discrete random variable X with range $S_x = x_1, x_2, \ldots$ satisfying $x_1 \leq x_2 \leq \ldots$,

(a) $F_X = (-\infty) = 0$ and $F_X(\infty) = 1$

(b) For all $x' \geq x, F_X(x') \geq F_X(x)$

(c) For all $x' > x, F_X(x') > F_X(x)$

(d) $F_X(x) = F_X(x_i)$ for all x such that $x_i \leq x \leq x_{i+1}$

Theorem 3.3 For all $b > a$, $F_X(b) - F_X(a) = P[a < X \leq b]$

Theorem 3.4 The Bernoulli (p) random variable X has expected value $E[X] = p$

Theorem 3.5 The geometirc (p) random variable X has expect value $E[X] = 1/p$

Theorem 3.6

(a) For the binomial (n, p) random variable X of Definition 3.6

$$E[X] = np$$

(b) For the Pascal (k, p) random variable X of Definition 3.7

$$E[X] = k/p$$

(c) For the discrete uniform (k, l) random variable X of Definition 3.8

$$E[X] = \frac{k+l}{2}$$

Theorem 3.8 Perfom 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 $n \to \infty$, $P_{K_n}(k)$ converges to the PMF of a Poisson (α) random variable.

Theorem 3.9 For a discrete random variable X, the PMF of $Y = g(X)$ is

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

Theorem 3.10 Given a random variable X with PMF $P_X(x)$, and the derived random variable $Y = g(x)$, the expected value of Y is

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

Theorem 3.11 For any random variable X

$$E[X - \mu_X] = 0$$

Theorem 3.12 For any random variable X

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

Theorem 3.13 In the absence of observations, the minimum mean square error estimate random variable X is

$$\hat{x} = E[X]$$

Theorem 3.14

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

Theorem 3.15

$$Var[aX + b] = a^2 Var[X]$$

Theorem 3.16

(a) If X is Bernoiulli (p), then $Var[X] = p(1-p)$

(b) If X is geometric (p), then $Var[X] = (1-p)/p^2$

(c) If X is binomial (n, p), then $Var[X] = np(1-p)$

(d) If X is Pascal (k, p), then $Var[X] = k(1-p)/p^2$

(e) If X is Poisson (α) then $Var[X] = \alpha$

(f) If X is discrete uniform (k, l), then $Var[X] = (l-k)(l-k+2)/12$

Definition 3.1 Random Variable

A random variable consists of an experiment with a probability measure $P[.]$ defined on a sample space S and a function that assigns a real number to each outcome in the sample spacce of the experiment.

Definition 3.2 Discrete Random Variable X is a discrete random variable if the range of X is a countable set.

$$S_X = x_1, x_2, \ldots$$

Definition 3.3 Probability Mass Function PMF The probability mass function (PMF) of a discrete random variable X is a function that assigns a probability to each value in the range of X

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

Definition 3.4 Bernoulli (p) Random Variable X is a Bernoulli (p) random variable if the PMF of X has the form

$$P_X(x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

where the parameter p is on the range $0 < p < 1$

Definition 3.5 Geometric (p) Random Variable X is a geometric (p) random variable if the PMF of X has the form

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

where the parameter p is on the range $0 < p < 1$

Definition 3.6 Binomial (n, p) **Random Variable** X is a binomial (n, p) random variable if the PMF of X has the form

$$P_X(s) = \binom{n}{x} p^x (1-p)^{n-x}$$

where $0 < p < 1$ and n is an integer such that $n \geq 1$

Definition 3.7 Pascal (k, p) **Random Variable**

$$P_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

where $0 < p < 1$ and k is an integer such that $k \geq 1$

Definition 3.8 Discrete Uniform (k, l) **Random Variable** X is a discrete uniform (k, l) random variable if the PMF of X has the form

$$P_X(x) = \begin{cases} 1/(l-k+1) & x = k, k+1, k+2, \ldots, l \\ 0 & \text{otherwise} \end{cases}$$

where the parameters k and l are integers such that $k < l$.

Definition 3.9 Poisson (α) **Random Variable** X is a Poisson (α) random variable if the PMF of X has the form

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0, 1, 2, \ldots, \\ 0 & \text{otherwise} \end{cases}$$

where the parameter α is in the range $\alpha > 0$

Definition 3.10 Cumulative Distribution Function (CDF) The cumulative distribution function (CDF) of a discrete random variable X is a function that assigns a probability to each value in the range of X.

$$F_X(x) = P[X \leq x]$$

Definition 3.11 Mode A mode of random variable X is a number x_{mod} satisfying $P_X(x_{mod}) \geq P_X(x)$ for all x

Definition 3.12 Median A median x_{med} of random variable X is a number that satisfies

$$P[X \leq x_{med}] = 1/2, \quad P[X \geq x_{med}] = 1/2$$

Definition 3.13 Expected Value The expected value of X is

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

Definition 3.14 Derived Random Variable Each sample value y of a derived random variable Y is a mathematical function $g(x)$ of a sample value x of another random variable X. We adopt the notation $Y = g(X)$ to describe the relationship of the two random variables.

Definition 3.15 Variance The variance of random variable X is

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

Definition 3.16 Standard Deviation The standard deviation of random variable X is

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

Definition 3.17 Moments For random variable X

(a) The nth moment is $E[X^n]$

(b) The nth central moment is $E[(X - \mu_X)^n]$

4. Continuous Random Variables

Theorem 4.1 For any random variable X,

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

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

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

Theorem 4.2 For a continuous random variable X, with PMF $f_X(x)$,

(a) $f_X(x) \geq 0$ for all x,

(b) $f_X(x) = \int_{-\infty}^{x} f_X(u)\, du$,

(c) $\int_{-\infty}^{\infty} f_X(x) dx = 1$

Theorem 4.3

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

Theorem 4.4 The expected value of a function, $g(X)$, of random variable X is

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

Theorem 4.5 For any random variable X,

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

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

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

(d) $Var[aX + b] = a^2 Var[X]$

Theorem 4.6 If X is a uniform (a, b) random variable,

- The CDF of X is

$$F_X(x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \leq x \leq b \\ 1 & x > b \end{cases}$$

- The expected value of X is $E[X] = (a+b)/2$
- The varaibce of X is $Var[X] = (b-a)^2/12$

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 4.8 If X is an exponential (λ) random variable,

- The CDF of X is

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases}$$

- The expected value of X is $E[X] = 1/\lambda$
- The variance of X is $Var[X] = 1/\lambda^2$

Theorem 4.9 If X is an exponential (λ) random variable, then $K = \lceil X \rceil$ is a geometric (p) random variable with $p = 1 - e^{-\lambda}$

Theorem 4.10 If X is an Erlang (n, λ) random variable, then

(a) $E[X] = \frac{n}{(\lambda)}$

(b) $Var[X] = \frac{n}{(\lambda)^2}$

Theorem 4.11 Let K_α denote a Poisson α random variable. For any $x > 0$, the CDF of an Erlang (n, λ) random variable X satisfies,

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

Theorem 4.12 If X is a Gaussian (μ, σ) random variable, then

$$E[X] = \mu \qquad Var[X] = \sigma^2$$

Theorem 4.13 If X is a Gaussian $(\mu, \sigma), Y = aX + b$ is Gaussian $(a\mu + b, a\sigma)$

Theorem 4.14 If X is a Gaussian (μ, σ) random variable, the CDF 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 $\Phi(-z) = 1 - \Phi(z)$

Theorem 4.16 For any continuous function g(x),

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

Theorem 4.17 $\int_{-\infty}^{x} \delta(v) dv = u(x)$

Theorem 4.18 For a random variable X, we have the folloing equivalent statements:

(a) $P[X = x_0] = q$

(b) $P[x_0] = q$

(c) $F_X(x_0^+) - F_X(x_0^-) = q$

(d) $f_x(x_0) = q\delta(0)$

Definition 4.1 Cumulative Distribution Function (CDF) The cumulative distribution function (CDF) of random variable X is $F_X(x) = P[X \leq x]$

Definition 4.2 Continuous Random Variable X is a continuous random variable if the CDF F_X(x) is a continuous function.

Definition 4.3 Probability Density Function (PDF) The probability density function (PDF) of a continuous random variable X is

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Definition 4.4 Expected Value The expected value of a random variable X is

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

Definition 4.5 Uniform Random Variable X is a uniform (a, b) random variable if the PDF of X is $f_X(x)$, and where the parameter $\lambda > 0$

$$f_X(x) = \begin{cases} 1/(b-a) & a \leq x \leq b \\ 0 & otherwise \end{cases}$$

Definition 4.6 Exponential Random Variable X is an exponential (\lambda) random variable if the PDF of X is $f_X(x)$, and where the parameter $\lambda > 0$

$$f_x(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases}$$

Definition 4.7 Erlang Random Variable X is an Erlang (n, λ) random variable if the PDF of X is $f_X(x)$ where the parameter $\lambda > 0$, and the parameter $n \geq 1$ is an integer.

$$f_X(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq n \\ 0 & otherwise \end{cases}$$

Definition 4.8 Gaussian Random Variable X is a Gaussian (μ, σ) random variable if the PDF of X is $f_X(x)$ where the parameter μ can be any real number and the parameter $\sigma > 0$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

Definition 4.9 Standard Normal Random Variable The standard normal random variable Z is the Gaussian $(0, 1)$ random variable.

Definition 4.10 Standard Normal CDF The CDF of the standard normal random variable Z is

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du$$

Definition 4.11 Standard Normal Complementary CDF The standard normal complementary CDF is

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

Definition 4.12 Unit Impluse (Delta) Function Let

$$\delta(x) = \begin{cases} 1/\epsilon & -\epsilon/2 \leq x \leq \epsilon/2 \\ 0 & otherwise \end{cases}$$

The unit impulse function

$$\delta(x) = \lim_{x \to 0} d_\epsilon(x)$$

Definition 4.13 Unit Step Function The unit step function is

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

Definition 4.14 Mixed Random Variable X is a mixed random variable if and only if $F_X(x)$ contains both impluses and nonzero, finite values.