Chapter 1

Preliminary Knowledge

Probability and Random Variables

Chap 1: Preliminary Knowledge

- Outline

Chap 1: Preliminary Knowledge

1.1 Probability Space

1.2 Random Variables

1.3 Moments of Random Variables

1.4 Special Distribution

1.5 Characteristic Functions

1.1 Probability Space

1. Probability Space (Ω , F, P)

Sample Space: Ω

- The set of all possible outcomes in any given experiments.

Borel Field(or σ Field) : F

- The collection of all possible events from the sample space.

Probability: P

- P is a probability law (i.e. probability function P(·)) that assigns a number to each event in F.

1.1 Probability Space

1. Probability Space (Ω ,F, P)

Measurable Space:

The pair (Ω ,F) is called a measurable space.

Probability Space:

The triple (Ω ,F, P) is called a probability space.

1.1 Probability Space

2. Conditional Probability

Def. The conditional event for A given B, A/B, is the event A under the stipulation that B has occurred.

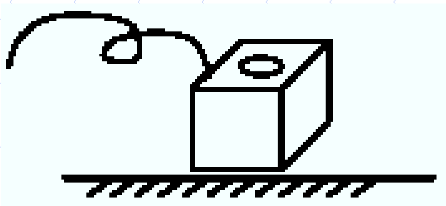
Def. If A and B are events in F with P(B) $\neq$ 0, the conditional probability of A given B is

$$P(A/B)=P(AB)/P(B).$$

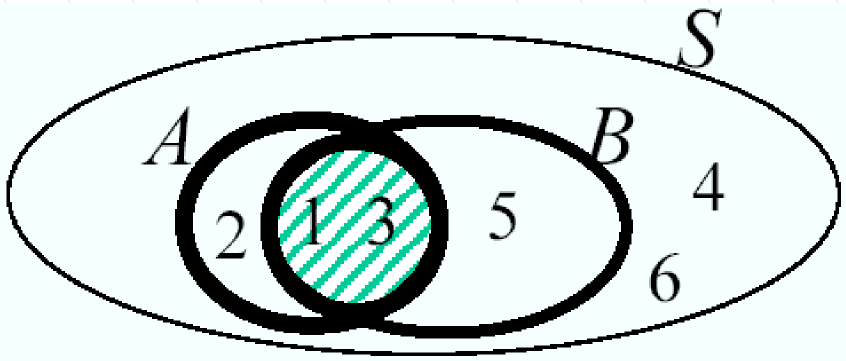
1.1 Probability Space

2. Conditional Probability

Examples:

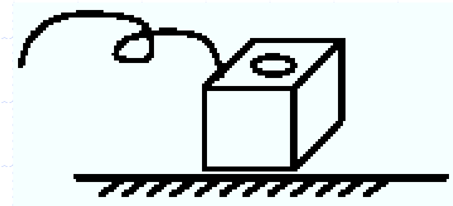


$$A = \{1,2,3\}$$
$$B = \{1,3,5\}$$



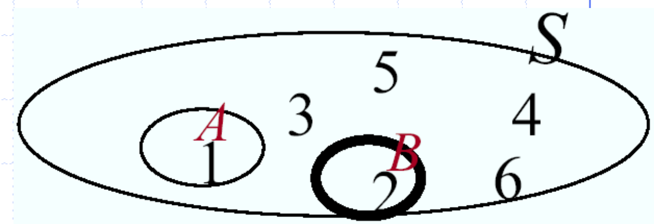
$$\rightarrow P(A/B) = \frac{P(AB)}{P(B)} = \frac{1/3}{1/2} = \frac{2}{3}$$

1.1 Probability Space

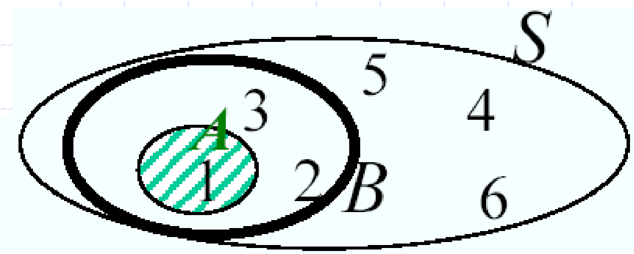


2. Conditional Probability

Properties:

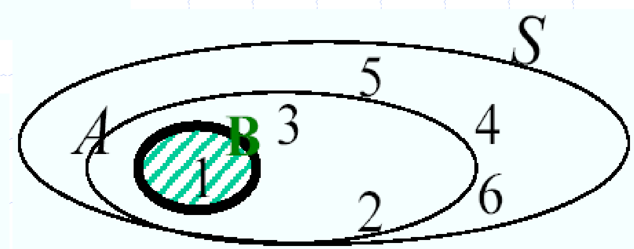


(1) $AB = 0 \rightarrow P(A/B) = 0$



(2) $A \subset B \rightarrow A \cdot B = A$

$$\rightarrow P(A/B) = \frac{P(A)}{P(B)} \geq P(A)$$



(3) $B \subset A \rightarrow AB = B$

$$\rightarrow P(A/B) = \frac{P(B)}{P(B)} = 1$$

1.1 Probability Space

3. Independent

Def. Two events A & B are independent if

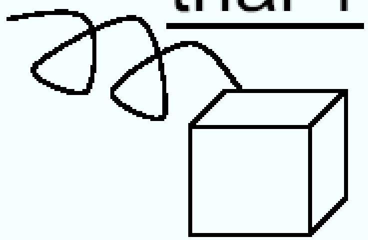
$$P(AB)=P(A)P(B)$$

$$P(A/B)=\frac{P(AB)}{P(B)}=\frac{P(A)P(B)}{P(B)}=P(A)$$

1.1 Probability Space

3. Independent

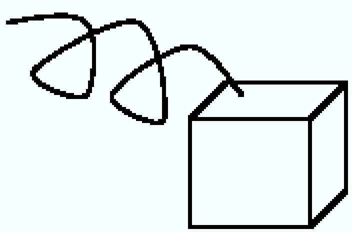
e.g. trial 1



$$S_1 = \{1,2,3,4,5,6\}$$

$$A_1 = \{1\} \qquad P(A_1) = \frac{1}{6}$$

trial 2



$$S_1 = \{1,2,3,4,5,6\}$$

$$A_2 = \{1\} \qquad P(A_2) = \frac{1}{6}$$

If A_1 & A_2 are independent, $P(A_1A_2) = P(A_1)P(A_2)$

1.1 Probability Space

3. Independent

The space of A_1A_2 is

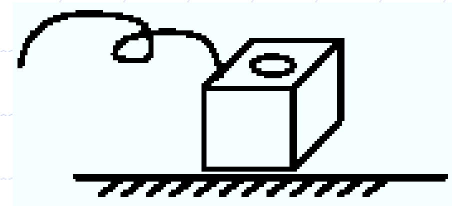
$$S = S_1 \times S_2 = \{ (1,1), (1,2), (1,3), \ldots, (6,6)\}$$

S is a new space.

Event A_1A_2 consisting of all ordered-pairs (S_{1i}, S_{2j}), $S_{1i}\ \ A_1$, $S_{2j}\ \ A_2$ is a subset of S.

1.1 Probability Space

3. Independent

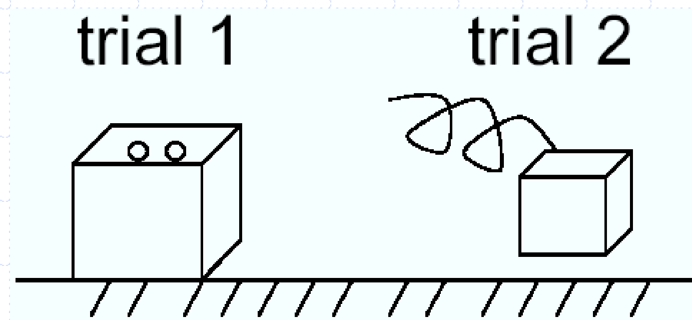


e.g.1 a single trial

$A = \{1\} \quad B = \{2\}$

A & B are Mutually Exclusive Events

$$P(A/B) = \frac{P(AB)}{P(B)} = \frac{P(\phi)}{P(B)} = 0$$



e.g. 2 two trials

$A = \{1\} \quad B = \{2\}$

A & B are Independent Events

$$P(A/B) = \frac{P(AB)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

1.1 Probability Space

3. Independent

e.g. Given

$$P(A_1) = 1/2$$
$$P(A_2) = 1/4$$
$$P(A_3) = 1/4$$
$$P(A_1A_2) = 1/8$$
$$P(A_1A_3) = 1/8$$
$$P(A_2A_3) = 1/8$$
$$P(A_1A_2A_3) = 1/32$$

Are A_1, A_2 & A_3 independent?

1.1 Probability Space

3. Independent

- For events $A_1, A_2, \ldots, A_n$ (which may or may not be independent), the probability of the simultaneous occurrence of the n events is

$$P(A_1A_2\cdots A_n)=$$

$$P(A_1)\cdot P(A_2/A_1)\cdot P(A_3/A_1A_2)\cdots P(A_n/A_1A_2\ldots A_{n-1})$$

- If events $A_1, A_2, \ldots, A_n$ are independent, then

$$P(A_1A_2\cdots A_i)=P(A_1)\cdot P(A_2)\cdots P(A_i),\ i=2,3,\ldots n$$

1.1 Probability Space

4. Theorem of total probability

$$P(B) = P(B/A_1)P(A_1) + \ldots + P(B/A_n)\cdot P(A_n)$$

5. Bayes equation

? *Help!*

Chap 1: Preliminary Knowledge

1.1 Probability Space

1.2 Random Variables

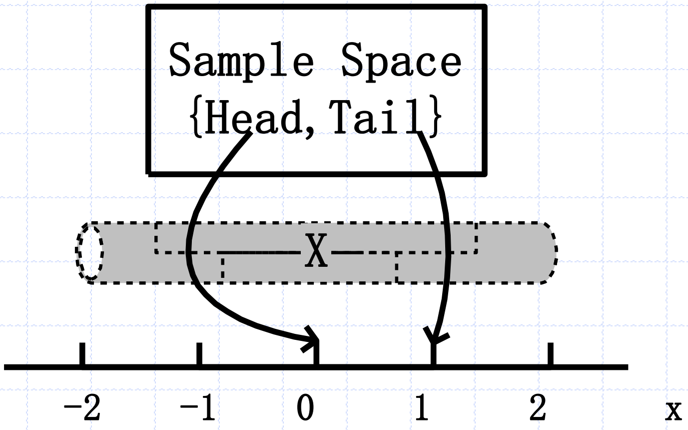
1.3 Moments of Random Variables

1.4 Special Distribution

1.5 Characteristic Functions

1.2 Random Variables

Def. A real one value function of the elements of a sample space.



Obtained by mapping a sample space to real axis.

Discrete vs. Continuous random variables

1.2 Random Variables

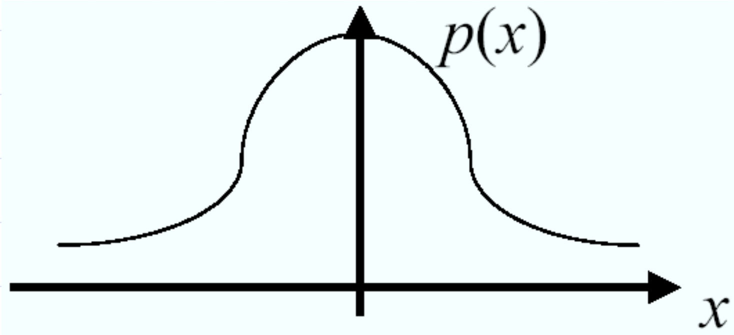
1. Probability Density Function (*p.d.f.*)

Def: Let x represent a continuous random variable in a sample space S. For each x we have a p.d.f. $p(x)$ which is a function that satisfies the following:

1) $p(x) \geq 0 \; \forall x \in S$

2) $\int_S p(x)\,dx = 1$

3) $\forall x_1 < x_2$ in S



The probability of

$$x \in [x_1, x_2] = P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} p(x)\,dx$$

1.2 Random Variables

2. Probability Distribution Function

Def: The Probability Distribution Function $F(x_1)$ of a random variable X is the probability that X is less than or equal to x_1,

$$F(x_1) = \int_{-\infty}^{x_1} p(x)dx = P(x \le x_1)$$

1.2 Random Variables

2. Probability Distribution Function

Note:

The x in $p(x)$ and $F(x)$ is not a random variable but a value of the random variable X.

The $p.d.f. p(x)$ is not a probability but a rate of change of the probability $F(x)$, i.e. $\frac{dF(x)}{dx}$

The distribution function $F(x)$ of random variable X is the probability that X has a value less than or equal to the value x.

1.2 Random Variables

3. *p.d.f.* of Linearly Combined Random Variables

Let $X_1, X_2, \ldots\ldots\ldots, X_n$ are n independent variables & $p_{x_1}(x_1), p_{x_2}(x_2), \ldots, p_{x_n}(x_n)$ are their *p.d.f s* respectively.

Let $Y1 = X_1 + X_2$, $Y2 = X_1 + X_2 + \ldots + X_n$

Then, the *p.d.f* of Y is

$$p_y(y) = p_{x_1} \otimes p_{x_2} \otimes \cdots\cdots \otimes p_{x_n}$$

1.2 Random Variables

4. Two-Dimensional Distributions

Def. The joint probability density function of two random variables *X* and *Y* is a function $p(x,y)$ that possesses the properties

$$i) \quad p(x,y) \geq 0$$

$$ii) \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\,dx\,dy = 1$$

$$iii) \quad P(x_1 \leq x \leq x_2, y_1 \leq y \leq y_2) = \int_{y_1}^{y_2}\int_{x_1}^{x_2} p(x,y)\,dx\,dy$$

1.2 Random Variables

4. Two-Dimensional Distributions

Def. The joint probability distribution function is

$$F(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} p(x,y)\,dx\,dy$$

so

$$p(x,y) = \frac{\partial^2 F(x,y)}{\partial x\,\partial y}.$$

Def. The random variables X and Y with $p.d.f.$ $p_x(x)$ and $p_y(y)$ are independent if

$$\boldsymbol{p(x,y) = p_x(x)\,p_y(y)}$$

1.2 Random Variables

4. Two-Dimensional Distributions

Def. The maginal probabiliy density functions of the variables X and Y are

$$p_X(x) = \int_{-\infty}^{\infty} p(x,y)dy \qquad p_Y(y) = \int_{-\infty}^{\infty} p(x,y)dx$$

Def. The maginal probability distribution functions are

$$F_X(x) = \int_{-\infty}^{x} p_X(x)dx = \int_{-\infty}^{x}\int_{-\infty}^{\infty} p(x,y)dydx$$

$$F_Y(y) = \int_{-\infty}^{y} p_Y(y)dy = \int_{-\infty}^{y}\int_{-\infty}^{\infty} p(x,y)dxdy$$

1.2 Random Variables

4. Two-Dimensional Distributions

E.g.

Find k for the 2D p.d.f.

$$p_{XY}(x,y)=\begin{cases} ke^{-2x-3y} & x\geq 0, y\geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Problems

Let $\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots\ldots\ldots, \boldsymbol{X_n}$ are n independent variables & $p_{x_1}(x_1), p_{x_2}(x_2), \ldots, p_{x_n}(x_n)$ are their *p.d.fs* respectively. Find the *p.d.fs of*

(1) Y=min ($\boldsymbol{X_1}, \boldsymbol{X_2}$)

(2) Y=max ($\boldsymbol{X_1}, \boldsymbol{X_2}$)

(3) Y=h(X1)

(4) Y1=h1(X1,X2)

Y2=h2(X1,X2)

Chap 1: Preliminary Knowledge

Outline

1.3 Moments of Random Variables

1. Moments

*Def.*1 The nth moment of $p(x)$ (about the origin)

$$\mathrm{E}\left[\boldsymbol{x}^{n}\right] = \int_{-\infty}^{\infty} x^{n} \cdot p(x) dx \qquad n = 1, 2, \ldots$$

*Def.*2 The nth moment of $p(x)$ about the x_0

$$\mathrm{E}\left[(\boldsymbol{x}-x_{0})^{n}\right]=\int_{-\infty}^{\infty}(x-x_{0})^{n} \cdot p(x) dx$$

$$n = 1, 2, \ldots$$

1.3 Moments of Random Variables

(1) Expectation=Mean of *X*:

The first moment (n= 1) (about the origin)

$$\overline{\boldsymbol{x}} = \mathrm{E}[\boldsymbol{x}] = \int_{-\infty}^{\infty} x \cdot p(x) dx$$

expected value = mean value = ensemble average

(2) Mean square of X:

The second moment (n=2) (about the origin)

$$\overline{\boldsymbol{x}^2} = \mathrm{E}[\boldsymbol{x}^2] = \int_{-\infty}^{\infty} x^2 \cdot p(x) dx$$

1.3 Moments of Random Variables

(3) Variance= second moment (n=2) (about the mean, i.e. central moment)

$$\sigma^2 = Var(X)$$

$$= E[(X - E(X))^2] = E(X^2) - (E(X))^2$$

$$= \overline{x^2} - \overline{x}^2$$

(4) Standard Deviation = $\sqrt{\mathrm{Var}(X)}$

1.3 Moments of Random Variables

2. Functions of Random Variable

- If $\boldsymbol{g}(X)$ is an arbitrary function of X, the expected value of $\boldsymbol{g}(X)$ is

$$\mathrm{E}[\mathbf{g}(X)] = \int_{-\infty}^{\infty} g(x) \cdot p(x) dx$$

X and $g(X)$ are continuous random variables.

1.3 Moments of Random Variables

3. Moments of 2-D p.d.f.

Def. The moments of a joint *p.d.f.* $p(x,y)$ are called joint moments

$$\mu'_{i,j} = E[x^i y^j] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i y^j \, p(x,y)\, dx\, dy$$

where $i, j = 0, 1, 2, 3, \ldots$, and the order of $\mu'_{i,j}$ is $i + j$.

Note: $$E[x] = \mu'_{1,0} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i \, p(x,y)\, dx\, dy$$

$$E[y] = \mu'_{0,1} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^j \, p(x,y)\, dx\, dy$$

1.3 Moments of Random Variables

3. Moments of 2-D p.d.f.

- If $g(X,Y)$ is an arbitrary function of X *and* Y, the expected value of $g(X,Y)$ is

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, p_{XY}(x,y)\, dx dy$$

X, Y and $g(X,Y)$ are continuous random variables.

1.3 Moments of Random Variables

3. Moments of 2-D p.d.f.

Def. The central moments (i.e. moments about the mean) are

$$\mu_{ij} = E[(\boldsymbol{x} - \overline{\boldsymbol{x}})^i (\boldsymbol{y} - \overline{\boldsymbol{y}})^j]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \overline{x})^i \, (y - \overline{y})^j \, p(x, y) \, dx \, dy$$

where $i, j = 0, 1, 2, 3, \ldots$, and the order of μ_{ij} is $i + j$.

Note:

The moment μ_{11} is called the covariance of two variables.

1.3 Moments of Random Variables

3. Moments of 2-D p.d.f.

Covariance

$$\mathrm{Cov}(X,Y)=E\left[\left(X-E(X)\right)\left(Y-E(Y)\right)\right]=E(XY)-E(X)E(Y)$$

e.g.

Find: the three 2nd order moments of the random variables *X* and *Y*.

1.3 Moments of Random Variables

4. Correlation

Correlation:

Def. The *numerical measure of the similarity between X and* Y is the normalised correlation coefficients and is defined as

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

1.3 Moments of Random Variables

4. Correlation

Properties of Correlation:

(i) ρ [-1, 1]

(ii) $\rho=0$ if *X* and *Y* are uncorrelated

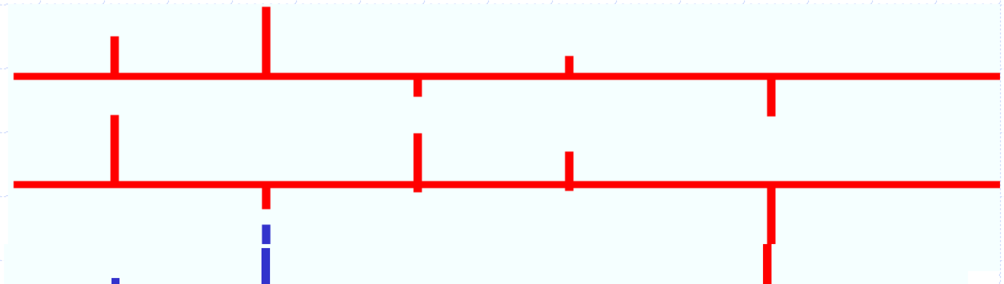
(i.e. $\mu_{11}= 0$ or Cov(X,Y)=0).

__Def:__ Random variables *X* and *Y* are uncorrelated if $\rho=0$ or $\mu_{11}= 0$ or Cov(X,Y)=0.

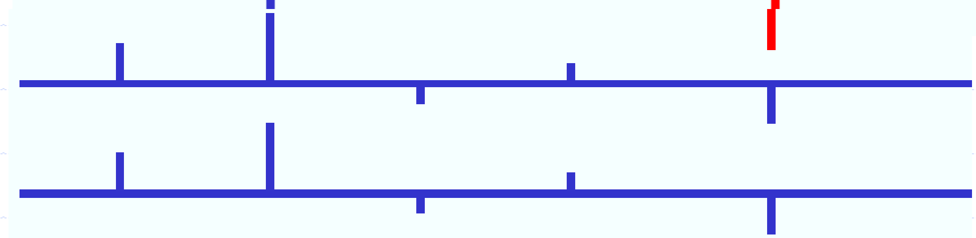
1.3 Moments of Random Variables

4. Correlation

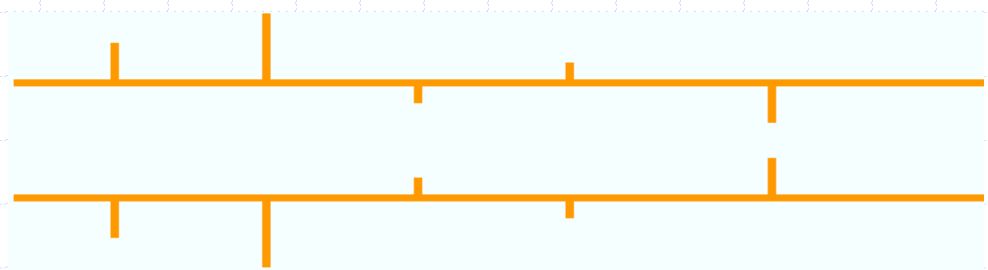
$$\rho \approx 0 \begin{cases} x \\ y \end{cases}$$



$$\rho \to 1 \begin{cases} x \\ y \end{cases}$$



$$\rho \to -1 \begin{cases} x \\ y \end{cases}$$



1.3 Moments of Random Variables

4. Correlation

Independent and Uncorrelated:

Theorem: If X and Y are statistically independent then they are uncorrelated.

e.g.

Given: X and Y are 2 independent random variables and $U=X+Y$, $V=X-Y$

Find: the condition under which U and V are uncorrelated.

1.3 Moments of Random Variables

5. If *X* and *Y* are independent then

$$\mathrm{Cov}(X,Y) = \rho_{XY} = 0$$

$$E(XY) = E(X)E(Y)$$

$$Var(X+Y) = Var(X) + Var(Y)$$

For any *X* and *Y*

$$E(X+Y) = E(X) + E(Y)$$

$$E(aX) = aE(X)$$

$$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$$

Chap 1: Preliminary Knowledge

1.1 Probability Space

1.2 Random Variables

1.3 Moments of Random Variables

1.4 Special Distribution

1.5 Characteristic Functions

1.4 Special Distributions

- Discrete
  - Bernoulli
  - Binomial
  - Geometric
  - Poisson
- Continuous
  - Uniform
  - Exponential
  - Normal

1. Discrete Distribution

1). Bernoulli Distribution

"Single coin flip", $\boldsymbol{p}$ = Pr (success)

Let N = 1 if success, 0 otherwise

$$\Pr(N=n)=\begin{cases} p, & n=1 \\ 1-p, & n=0 \end{cases}$$

$$E(N)=p$$

$$\mathrm{Var}(N)=p(1-p)$$

1. Discrete Distribution

2). Binomial Distribution

"n independent coin flips", p = Pr(success)

$N = k$ of successes

$$\Pr(N=k)=\binom{n}{k}p^k(1-p)^{n-k},\ k=0,1,\dots,n$$

$$E(N)=np$$

$$\mathrm{Var}(N)=np(1-p)$$

1. Discrete Distribution

3). Geometric Distribution

"independent coin flips" p = Pr(success)

$N = k$ of flips until (including) first success

$$\Pr(N = k) = (1-p)^{k-1} p, k = 1, 2, \ldots$$

$$E(N) = 1/p$$

$$\mathrm{Var}(N) = (1-p)/p^2$$

Memoryless property:

Have flipped k times without success.

$$\Pr(N = k + n | N > k) = (1-p)^{n-1} p \text{ (still geometric)}$$

1. Discrete Distribution

4). Poisson Distribution

"Occurrence of rare events"

λ = average rate of occurrence per period;

$N = k$ of events in an arbitrary period

$$\Pr\left(N=k\right)=\frac{\lambda^k e^{-\lambda}}{k!},\ k=0,1,2,\ldots$$

$$E\left(N\right)=\lambda$$

$$\mathrm{Var}\left(N\right)=\lambda$$

2. Continuous Distribution

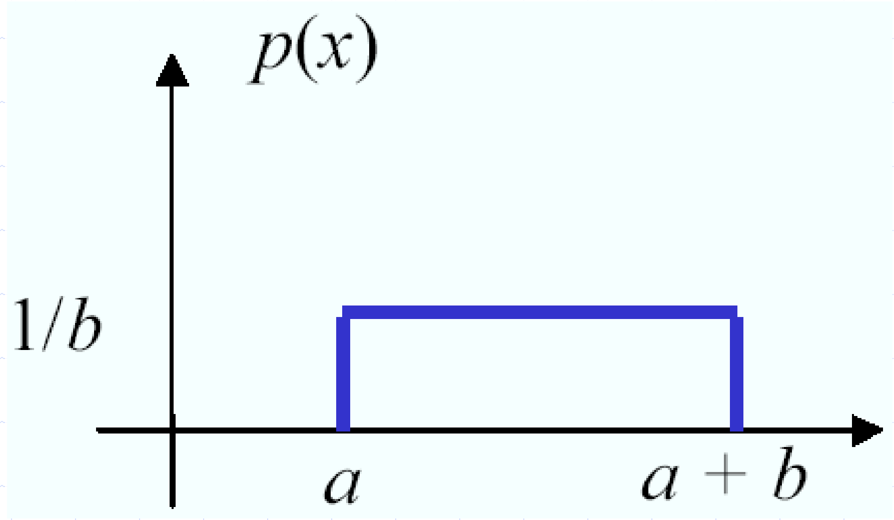
1) Uniform Distribution

Random variable X is equally likely to fall anywhere within interval (a, b).

$$f_X(x) = \frac{1}{b-a}, \quad a \le x \le b$$

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

$$\mathrm{Var}(X) = \frac{(b-a)^2}{12}$$



2. Continuous Distribution

2) Exponential Distribution

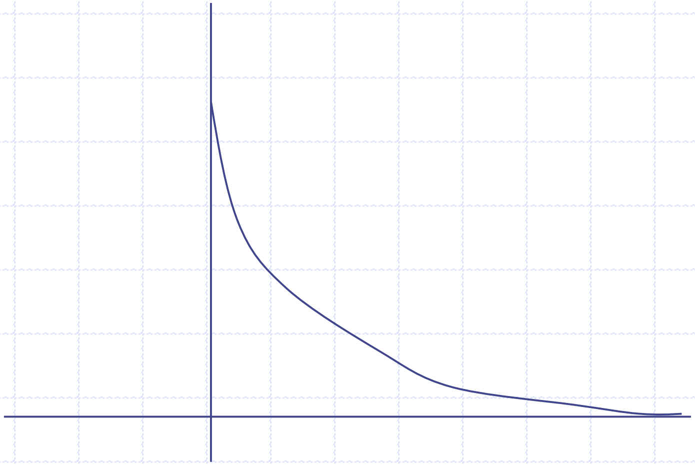
Random variable X is nonnegative and it is most likely to fall near 0.

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

$$F_X(x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

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

$$\mathrm{Var}(X) = \frac{1}{\lambda^2}$$



Also memoryless $\boxed{P\{x > s+t \mid x > s\} = P\{x > t\}}$

2. Continuous Distribution

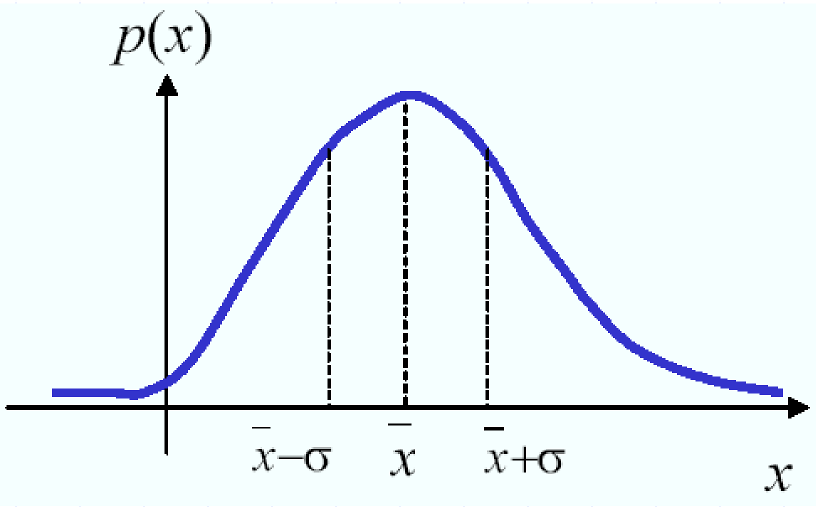
3) Normal Distribution

X follows a "bell-shaped" density function

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

$$E(X) = \mu$$

$$\mathrm{Var}(X) = \sigma^2$$



- From the central limit theorem, the distribution of the sum of independent and identically distributed random variables approaches a normal distribution as the number of summed random variables goes to infinity.

Chap 1: Preliminary Knowledge

1.5 Characteristic Function

1. Definition

The characteristic function is defined as

$$\phi_X(t) = Ee^{itX}$$

where ***t*** is a real number and ***i*** is an imaginary unit.

$$\phi_X(t) = \begin{cases} \sum_{j \in S} e^{itx_j} P(x_j) & \text{for a discrete-type random variable} \\ \int_S e^{itx} f(x) dx & \text{for a continuous-type random variable} \end{cases}$$

where S stands for the sample space.

1.5 Characteristic Functions

2. Properties of Characteristic Function

(i) $\phi(t)$ always exist since $|e^{itx}|$ is a continuous and bounded function for all finite real values of t and x.

(ii) $\phi(0)=1$

(iii) $|\phi(t)|=|E[e^{itx}]|\leq 1$

(iv) $\phi(-t)=\phi^*(t)$, where $\phi^*(t)$ denotes the complex conjugate to $\phi(t)$

1.5 Characteristic Functions

2. Properties of Characteristic Function

(v) $E[x^r] = \frac{1}{i^r}\phi^{(r)}(0)$

e.g. Obtain the mean and variance of the Bernoulli distribution whose probability mass function is given by $p(x) = p^x(1-p)^{1-x}$, where x =0 or 1.

(vi) X is a random variable. $Y=aX+b$, where a and b are constants. Then, the characteristic function of Y *is*

$$\phi_y(t) = e^{ibt}\phi_x(at)$$

1.5 Characteristic Functions

2. Properties of Characteristic Function

(vii) The random variables $X_1, X_2, \ldots, X_n$ are statistically independent. Let

$$Y = X_1 + X_2 + \ldots, + X_n$$

Then, the characteristic function of *Y is*

$$\phi_y(t) = \phi_{x_1}(t)\phi_{x_2}(t)\cdots\phi_{x_n}(t)$$

1.5 Characteristic Functions

3. Theorems 1

- If X is a discrete-type random variable, its probability mass function $\boldsymbol{p}(\boldsymbol{x})$ can be obtained by

$$p(x) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \phi(t) e^{-itx} dt$$

- If X is a continuous-type random variable, its probability density function $\boldsymbol{p}(\boldsymbol{x})$ can be obtained by

$$p(x) = \frac{1}{2\pi} \int_{-T}^{T} \phi(t) e^{-itx} dt$$

where

$$\int_{-\infty}^{\infty} |\phi(t)| dt < \infty$$

1.5 Characteristic Functions

4. Theorem 2 (Uniqueness Theorem)

Let X and Y be two random variables with characteristic functions $\phi_X(t)$ and $\phi_Y(t)$, respectively.

If $\phi_X(t) = \phi_Y(t)$ for all values of t, then X and Y have the same probability distribution.

Moment Generating Function

- Definition

The moment generating function is defined as

$$M_X(t) = E[e^{tX}]$$

where t is a real number.

$$M_X(t) = \begin{cases} \sum_{j \in S} e^{tx_j} p(x_j) & \text{for a discrete-type random variable} \\ \int_S e^{tx} f(x) dx & \text{for a continuous-type random variable} \end{cases}$$

where S stands for the sample space.

Generating Function and Probability Generating Function

- Let $\{a_n\}$ denote a sequence of numbers. The Generating Function for the sequence $\{a_n\}$ as

$$a^g(z) = \sum_{n=0}^{\infty} a_n z^n \qquad |z| < R$$

R is the convergence region.

- Let X denote a discrete random variable and

$$a_n = P\{X=n\}$$

Then $P_X(z)=a^g(z)=E[z^X]$ is called the probability generating function for the random variable X.

Homework

- 1.1
- 1.6
- What is the characteristic function of an exponential random variable X? Find $E[X^3]$.
- What is the probability generating function of an binomial random variable?
- 1.15 (2nd book)

End of chapter 1