Introduction to Probability, Statistics and Random Processes

Chapter 4: Continuous and Mixed Random Variables

Hossein Pishro-Nik
University of Massachusetts, Amherst
Email: pishro@ecs.umass.edu

Introduction

- Discrete random variables can only take a countable number of possible values.
- Continuous random variables have a range in the form of
  - Interval on the real number line.
  - Union of non-overlapping intervals on real line.
- We also know that for any $x \in \mathbb{R}, P(X = x) = 0$.
- Analogous to the theory of discrete random variables.

Continuous Random Variables and their Distributions

- Example: Choose a real number uniformly at random in the interval $[a, b]$ and call it X.
- By uniformly at random, we mean all intervals in $[a, b]$ that have the same length have the same probability.
- Find the CDF of X.

Continuous Random Variables and their Distributions

- Uniformity implies that probability of an interval in $[a, b]$ is proportional to its length.

$$P(X \in [x_1, x_2]) \propto (x_2 - x_1)$$

- Since $P(X \in [a, b]) = 1$, we have

$$P(X \in [x_1, x_2]) = \frac{x_2 - x_1}{b - a}, \quad \text{where } a \leq x_1 \leq x_2 \leq b.$$

- From the definition of CDF, $F_X(x) = P(X \leq x)$ we get

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

Continuous Random Variables and their Distributions

- CDF for a continuous random variable uniformly distributed over $[a, b]$.

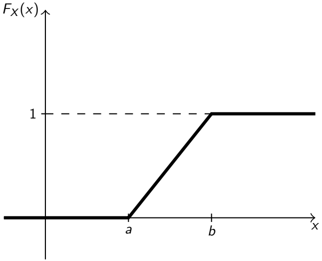

Figure: CDF for a continuous random variable uniformly distributed over $[a, b]$.

Continuous Random Variables and their Distributions

- We have the definition of a continuous random variable

> **Definition:** A random variable X with CDF $F_X(x)$ is said to be continuous if $F_X(x)$ is a continuous function for all $x \in \mathbb{R}$.

- The CDF is a continuous function with no jumps.
- No jumps is consistent with the fact that $P(X = x) = 0$ for all x.
- CDF of a continuous random variable is differentiable almost everywhere in $\mathbb{R}$.

Probability Density Function(PDF)

- For continuous random variables, CDF works but PMF does not since $P(X = x) = 0$.
- Instead we define a **Probability Density Function**.
- For a continuous random variable X, we define the function $f_X(x)$ as

$$f_X(x) = \lim_{\Delta \to 0^+} \frac{P(x < X \leq x + \Delta)}{\Delta}.$$

- $f_X(x)$ gives the probability density at point x.

Probability Density Function(PDF)

- $f_X(x)$ is limit of the probability of an interval as the length of the interval goes to 0.
- We know that

$$P(x < X \leq x + \delta) = F_X(x + \Delta) - F_X(x).$$

- Thus we get

$$f_X(x) = \lim_{\Delta \to 0} \frac{F_X(x + \Delta) - F_X(x)}{\Delta}$$
$$= \frac{dF_X(x)}{dx} = F'_X(x).$$

if $F_X(x)$ is differentiable at x.

Probability Density Function(PDF)

- We define the PDF of random variable X as

Definition: Consider a continuous random variable X with CDF $F_X(x)$. The function $f_X(x)$ is the probability density function (PDF) of X, defined by

$$f_X(x) = \frac{dF_X(x)}{dx} = F'_X(x),$$

if $F_X(x)$ is differentiable at x.

Probability Density Function(PDF)

- The PDF of a random variable with $Uniform(a, b)$ distribution is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & x < a \text{ or } x > b \end{cases}$$

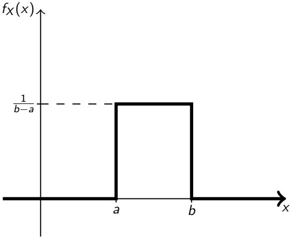

Figure: PDF for a continuous random variable uniformly distributed over $[a, b]$.

Probability Density Function(PDF)

- If $f_X(x_1) > f_X(x_2)$, we can say the value of X is more likely to be around x_1 than x_2.
- The CDF can be obtained from PDF by integration

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du.$$

- We also have

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_{a}^{b} f_X(u)du.$$

Properties of PDF

Consider a continuous random variable X with PDF $f_X(x)$. We have

1. $f_X(x) \geq 0$ for all $x \in \mathbb{R}$.
2. $\int_{-\infty}^{\infty} f_X(u)du = 1$.
3. $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(u)du$.
4. More generally, for a set A,
   $P(X \in A) = \int_A f_X(u)du$.

Range

- Range of a random variable X is the set of all possible values of the random variable.
- For a continuous random variable, we can define it as the set of all real numbers with non-zero PDF.

$$R_X = \{x | f_X(x) > 0\}$$

- R_X defined here might not show all possible values of X but the difference is unimportant.

Expected Value and Variance

- Remember the definition of expected value for a discrete random variable

$$EX = \sum_{x_k \in R_X} x_k P_X(x_k).$$

- We can write the definition of expected value of a continuous random variable as

$$\boxed{EX = \int_{-\infty}^{\infty} x f_X(x)\, dx}$$

Expected Value of a Function of a Continuous Random Variable

Law of the unconscious statistician (LOTUS) for continuous random variables:

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

- Expectation is a linear operation.
  - $E[aX + b] = aEX + b$ for all $a, b \in \mathbb{R}$
  - $E[X_1 + X_2 + ... + X_n] = EX_1 + EX_2 + ... + EX_n$ for any set of random variables $X_1, X_2, .., X_n$.

Variance

- Variance of a random variable is defined as

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

- For a continuous random variable we can write

$$\boxed{\begin{aligned} \mathrm{Var}(X) &= E\left[(X - \mu_X)^2\right] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx \\ &= EX^2 - (EX)^2 = \int_{-\infty}^{\infty} x^2 f_X(x)dx - \mu_X^2 \end{aligned}}$$

- For $a, b \in \mathbb{R}$, we have

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

Discrete vs Continuous Random Variables

Discrete	Continuous
PMF $P_X(x) = P(X = x)$	PDF $f_X(x) = \frac{dF_X(x)}{dx}$
$\sum$	$\int$
$EX = \sum_{x_k \in R_X} x_k P_X(x_k)$	$EX = \int_{-\infty}^{\infty} x f_X(x) dx$
LOTUS $E[g(x)] = \sum_{x \in R_X} g(x_k) P_X(x_k)$	LOTUS $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$

Functions of Continuous Random Variables

- If X is a continuous random variable, then $Y = g(X)$ is also a random variable.
- Note that the function of a continuous random variable might be a non-continuous random variable.
- We start by finding the CDF of Y.
- If the CDF is continuous, then we can differentiate to find the PDF of Y.

Functions of Continuous Random Variables

- Example: Let X be a $Uniform(0,1)$ random variable, and let $Y = e^X$.
  - Find the CDF and PDF of Y.
  - Find EY.
- Example: Let $X \sim Uniform(-1,1)$, and $Y = X^2$.
  - Find the CDF and PDF of Y.
  - Find EY.

The Method of Transformations

- If we want to find directly the PDF of $Y = g(X)$ and the function g satisfies some properties, we can use the *method of transformations*.

  Theorem: Suppose that X is a continuous random variable and $g : \mathbb{R} \to \mathbb{R}$ is a strictly monotonic differentiable function. Let $Y = g(X)$. Then the PDF of Y is given by

$$f_Y(y) = \begin{cases} \frac{f_X(x_1)}{|g'(x_1)|} = f_X(x_1).|\frac{dx_1}{dy}| & \text{where } g(x_1) = y \\ 0 \text{ if } g(x) = y \text{ does not have a solution} \end{cases}$$

The Method of Transformations

- We can extend the previous theorem to a more general case.
- If g is not monotonic, we can usually divide it into finite number of monotonic differentiable functions.

Theorem: Consider a continuous random variable X with domain R_X, and let $Y = g(X)$. Suppose that we can partition R_X into a finite number of intervals such that $g(x)$ is strictly monotone and differentiable on each partition. Then the PDF of Y is given by

$$f_Y(y) = \sum_{i=1}^{n} \frac{f_X(x_i)}{|g'(x_i)|} = \sum_{i=1}^{n} f_X(x_i). \left| \frac{dx_i}{dy} \right|$$

where $x_1, x_2, ..., x_n$ are real solutions to $g(x) = y$.

The Method of Transformations

- Example: Let X be a continuous random variable with PDF

$$f_X(x) = \begin{cases} 4x^3 & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and let $Y = \frac{1}{X}$. Find $f_Y(y)$.

- Example: Let X be a random variable with PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \text{for all } x \in \mathbb{R},$$

and let $Y = X^2$. Find $f_Y(y)$.

Uniform Distribution

A continuous random variable X is said to have a *Uniform* distribution over the interval $[a, b]$, shown as $X \sim Uniform(a, b)$, if its PDF is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & x < a \text{ or } x > b \end{cases}$$

Uniform Distribution

- CDF of $X \sim Uniform(a, b)$ is given by

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

- Expectation of X is given by

$$EX = \frac{a+b}{2}$$

- Variance of X is given as

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

Uniform Distribution

- PDF of $X \sim Uniform(a, b)$.

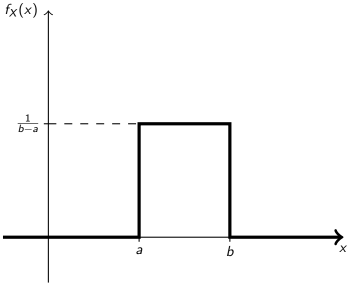

Figure: PDF for a continuous random variable uniformly distributed over (a, b).

Exponential Distribution

- Widely used in different applications to model the time elapsed between events.

A continuous random variable X is said to have an *exponential* distribution with parameter $\lambda > 0$, shown as $X \sim Exponential(\lambda)$, if its PDF is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Exponential Distribution

- CDF of $X \sim Exponential(\lambda)$ is
$$F_X(x) = (1 - e^{-\lambda x})u(x)$$
where $u(x)$ is unit step function.
- Expectation of X is given by
$$EX = \frac{1}{\lambda}$$
- Variance of X is
$$Var(X) = EX^2 - (EX)^2 = \frac{1}{\lambda^2}$$

Exponential Distribution

- PDF of $X \sim \mathit{Exponential}(\lambda)$ for $\lambda = 1, 2, 3$.

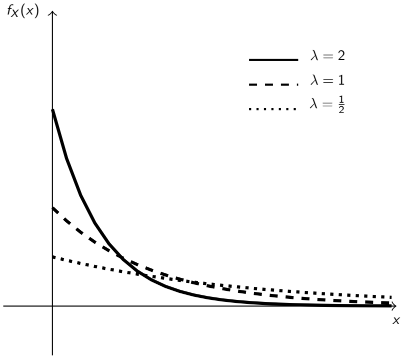

Figure: PDF of the exponential random variable

Exponential Distribution

- The exponential distribution can be viewed as the continuous analogue of the geometric distribution.
- Suppose you are tossing a coin until you observe the first heads.
- Let X be the time you observe the first heads.
- Let the Δ be the time between two tosses and probability of heads $p = \Delta.\lambda$.
- As Δ approaches 0, X converges to $Exponential(\lambda)$.

Exponential Distribution

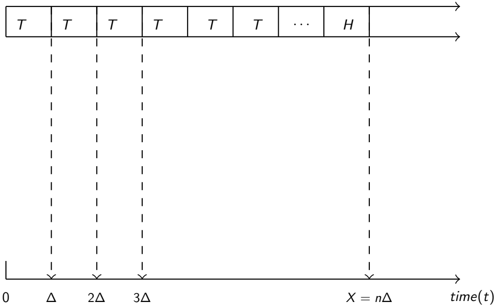

Figure: Exponential distribution can be viewed as the continuous analogue of the geometric distribution.

Exponential Distribution

If X is exponential with parameter $\lambda > 0$, then X is a *memoryless* random variable, that is

$$P(X > x + a | X > a) = P(X > x), \text{ for } a, x \geq 0.$$

- Say X is the waiting time until arrival of customer.
- From the memoryless property we have that, it does not matter how long you have waited.
- If you have not observed a customer until time a, the distribution of waiting time (from time a) is the same as when starting from at time zero.

Normal(Gaussian) Distribution

- By far the most important probability distribution.
- We will see it's importance in the Central Limit Theorem later on.
- We will first define the **standard normal random variable**.
- Other normal variables can be obtained by shifting and scaling a standard normal variable.

Normal Distribution

A continuous random variable Z is said to be a *standard normal (standard Gaussian)* random variable, shown as $Z \sim N(0,1)$, if its PDF is given by

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad \text{for all } z \in \mathbb{R}.$$

- If $Z \sim N(0,1)$, then expectation $EZ = 0$ and variance $Var(Z) = 1$.

Normal Distribution

- PDF of the standard normal random variable.

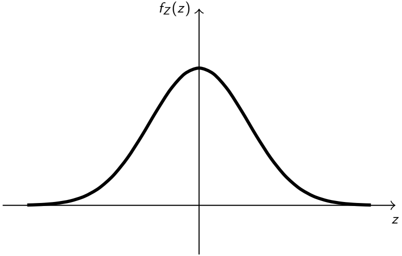

Figure: PDF of the standard normal random variable

Normal Distribution

The CDF of the standard normal distribution is denoted by the Φ function:

$$\Phi(x) = P(Z \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du.$$

- The integral in the CDF does not have a closed form solution.
- Due to it's importance, the values of $F_Z(z)$ have been calculated and readily available.
- Can use the command *normcdf* to compute $\Phi(x)$ for a given x in MATLAB.

Normal Distribution

- CDF of a standard normal variable.

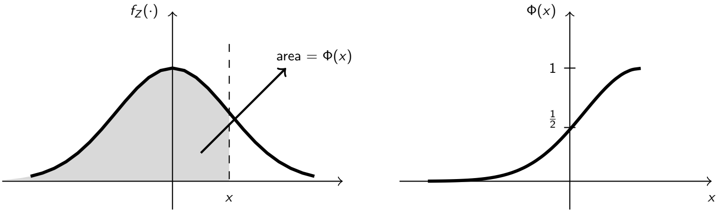

Figure: The Φ function (CDF of standard normal).

Normal Distribution

- Properties of the ϕ function include
  1. $\lim_{x \to \infty} \Phi(x) = 1, \quad \lim_{x \to -\infty} \Phi(x) = 0;$
  2. $\Phi(0) = \frac{1}{2};$
  3. $\Phi(-x) = 1 - \Phi(x)$, for all $x \in \mathbb{R}$.
- A very useful bound that we can use is

$$\frac{1}{\sqrt{2\pi}} \frac{x}{x^2+1} e^{-\frac{x^2}{2}} \leq 1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}$$

Normal Distribution

- From a standard normal random variable, we can obtain any normal variable by shifting and scaling.

If Z is a standard normal random variable and $X = \sigma Z + \mu$, then X is a normal random variable with mean μ and variance σ^2, i.e,

$$X \sim N(\mu, \sigma^2).$$

Normal Distribution

If X is a normal random variable with mean μ and variance σ^2, i.e, $X \sim N(\mu, \sigma^2)$, then

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

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

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

Normal Distribution

- PDFs for normal distributions with different mean and variance.

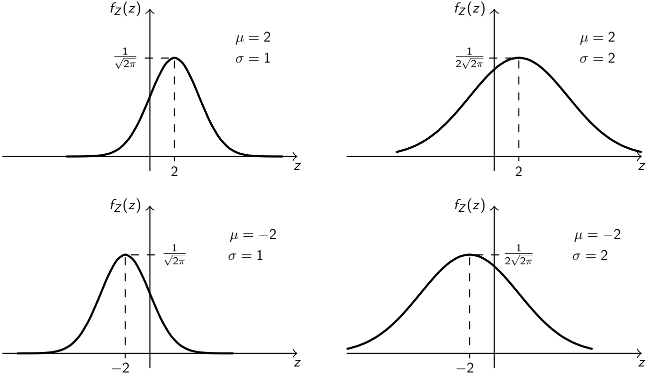

Figure: PDF for normal distribution

Normal Distribution

- An important property is that a linear transformation of a normal random variable is itself a normal random variable.
- We thus have the following theorem.

Theorem: If $X \sim N(\mu_X, \sigma_X^2)$, and $Y = aX + b$, where $a, b \in \mathbb{R}$, then $Y \sim N(\mu_Y, \sigma_Y^2)$ where

$$\mu_Y = a\mu_X + b, \;\; \sigma_Y^2 = a^2\sigma_X^2.$$

Gamma Distribution

- Widely used due it's relation to exponential and normal distributions.
- **Gamma function:**
  - Gamma function $\Gamma(x)$, is an extension of the factorial function to real (and complex) numbers.
  - If $n \in \{1, 2, 3, ...\}$, then

$$\Gamma(n) = (n-1)!$$

  - For any positive real number α, $\Gamma(\alpha)$ is defined as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \mathrm{d}x, \text{for } \alpha > 0.$$

  - For $\alpha = 1$, $\Gamma(1) = \int_0^\infty e^{-x} \mathrm{d}x = 1$.

Gamma Distribution

- Gamma function for some positive real values of α.

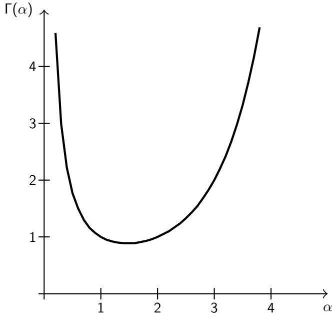

Figure: The gamma function for some positive real values of α.

Gamma Distribution

Properties of the gamma function:

For any positive real number α:

1. $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \mathrm{d}x$.
2. $\int_0^\infty x^{\alpha-1} e^{-\lambda x} \mathrm{d}x = \frac{\Gamma(\alpha)}{\lambda^\alpha}$, for $\lambda > 0$.
3. $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.
4. $\Gamma(n) = (n-1)!$, for $n = 1, 2, 3, \cdots$.
5. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Gamma Distribution

A continuous random variable X is said to have a *gamma* distribution with parameters $\alpha > 0$ and $\lambda > 0$, shown as $X \sim Gamma(\alpha, \lambda)$, if its PDF is given by

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

- $EX = \frac{\alpha}{\lambda}$ and $Var(X) = \frac{\alpha}{\lambda^2}$.
- Sum of n independent $Exponential(\lambda)$ random variables gives $Gamma(n, \lambda)$ random variable.

Gamma Distribution

- PDF of the gamma distribution for some values of α and λ.

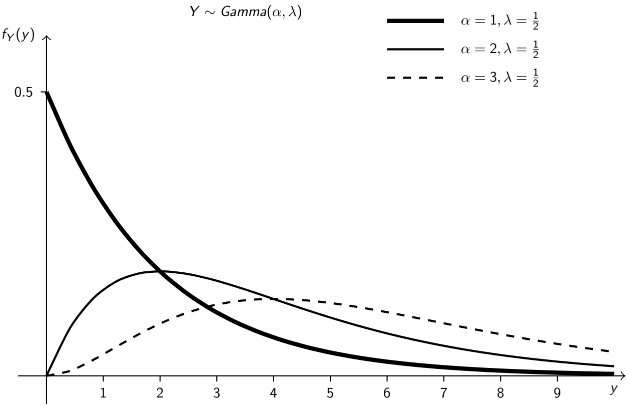

Figure: PDF of the gamma distribution for some values of α and λ.

Mixed Random Variables

- Mixed random variables have a discrete part and a continuous part.
- We can use the tools we have learned to analyze them.
- Later, we will revisit the concept of mixed random variables using the delta function.

Mixed Random Variables

- Example: Consider a mixed random variable Y with CDF $F_Y(y)$ given by

$$F_Y(y) = \begin{cases} 1 & y \geq \frac{1}{2} \\ y^2 & 0 \leq y < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

- Note that the CDF of Y is not continuous as there is a jump at $y = \frac{1}{2}$.
- It is also not in the staircase form and hence Y is not a discrete random variable.

Mixed Random Variables

- The CDF of Y

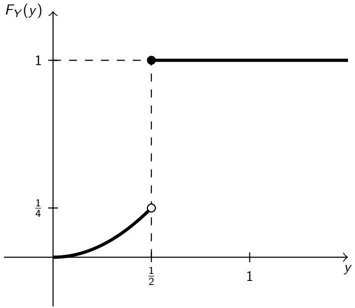

Figure: CDF of a mixed random variable Y.

Mixed Random Variables

- The CDF of Y has a continuous part and a discrete part.
- We can write

$$F_Y(y) = C(y) + D(y)$$

- $C(y)$ is the continuous part

$$C(y) = \begin{cases} \frac{1}{4} & y \geq \frac{1}{2} \\ y^2 & 0 \leq y < \frac{1}{2} \\ 0 & y < 0 \end{cases}$$

- The discrete part of $F_Y(y)$ is $D(y)$, given by

$$D(y) = \begin{cases} \frac{3}{4} & y \geq \frac{1}{2} \\ 0 & y < \frac{1}{2} \end{cases}$$

Mixed Random Variables

- We can write the CDF of any mixed random variable as a sum of a continuous function and a staircase function.

$$F_Y(y) = C(y) + D(y)$$

- Define $c(y) = \frac{dC(y)}{dy}$, wherever $C(y)$ is differentiable.
- $c(y)$ is not a valid PDF since it does not integrate to 1.
- Let $\{y_1, y_2, ...\}$ be the jump points of $D(y)$.
- We thus have

$$\int_{-\infty}^{\infty} c(y)dy + \sum_{y_k} P(Y = y_k) = 1.$$

- Expected value of Y is

$$EY = \int_{-\infty}^{\infty} yc(y)dy + \sum_{y_k} y_k P(Y = y_k).$$

Using the Delta Function

- We will use the Dirac Delta function to analyze mixed random variables.
- Remember that CDF is defined for all types of random variables.
- However PDF is only defined for continuous random variables.
- Delta functions will allow us to define PDFs for discrete and mixed random variables.

Dirac Delta Function

Properties of the delta function

Definition: We define the delta function $\delta(x)$ as an object with the following properties:

1. $\delta(x) = \begin{cases} \infty & x = 0 \\ 0 & \text{otherwise} \end{cases}$
2. $\delta(x) = \frac{d}{dx}u(x)$, where $u(x)$ is the unit step function;
3. $\int_{-\epsilon}^{\epsilon} \delta(x)dx = 1$, for any $\epsilon > 0$;
4. for any $\epsilon > 0$ and any function $g(x)$ that is continuous over $(x_0 - \epsilon, x_0 + \epsilon)$, we have

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

Using the Delta Function is PDFs of Discrete and Mixed Random Variables

- We can write the CDF of a discrete random variable X with range $R_X = \{x_1, x_2, ..\}$ and PMF $P_X(x_k)$ as
$$F_X(x) = \sum_{x_k \in R_X} P_X(x_k) u(x - x_k)$$
- We can obtain the **generalized** PDF for X by differentiating the CDF.
- Remember that we have defined the derivative of the step function as the delta function.

Using the Delta Function is PDFs of Discrete and Mixed Random Variables

For a discrete random variable X with range $R_X = \{x_1, x_2, x_3, ...\}$ and PMF $P_X(x_k)$, we define the (generalized) probability density function (PDF) as

$$f_X(x) = \sum_{x_k \in R_X} P_X(x_k)\delta(x - x_k).$$

Using the Delta Function is PDFs of Discrete and Mixed Random Variables

- The generalized PDF is defined for all random variables; allows us to use the same formula for all types of random variables.
- If the generalized PDF
  - is a sum of delta functions, it is a discrete random variable.
  - does not include delta functions, it is a continuous random variable.
  - contains both delta and non-delta functions, it is a mixed random variable.

Using the Delta Function is PDFs of Discrete and Mixed Random Variables

The (generalized) PDF of a mixed random variable can be written in the form

$$f_X(x) = \sum_k a_k \delta(x - x_k) + g(x),$$

where $a_k = P(X = x_k)$, and $g(x) \geq 0$ does not contain any delta functions. Furthermore, we have

$$\int_{-\infty}^{\infty} f_X(x)dx = \sum_k a_k + \int_{-\infty}^{\infty} g(x)dx = 1.$$