Continuous Probability Distribution

- uniform Distribution
- exponential Distribution
- Normal Distribution
- Standard Normal Process

Example: Continuous Random Variables

Ex.: modeling the lifetime of a device

- Time is a **continuous random variable**
- Random Time is typically modeled as **exponential distribution**
- We assume that with **average** lifetime of a device is 2 years

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2}, & \text{x} \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

- Probability that the device's life is between 2 and 3 years is:

$$P(2 \leq x \leq 3) = \frac{1}{2} \int_2^3 e^{-x/2} dx = 0.14$$

The life time Ex.

- **Cumulative Distribution Function**: A device has the CDF:

$$F(x) = \frac{1}{2}\int_0^x e^{-t/2}dt = 1 - e^{-x/2}$$

  - The probability that the device lasts for less than 2 years:

$$P(0 \le X \le 2) = F(2) - F(0) = F(2) = 1 - e^{-1} = 0.632$$

  - The probability that it lasts between 2 and 3 years:

$$P(2 \le X \le 3) = F(3) - F(2) = (1 - e^{-(3/2)}) - (1 - e^{-1}) = 0.145$$

The life time Ex.

Expected Value and Variance

- The mean of life of the previous device is:

$$E(X) = \frac{1}{2}\int_0^\infty xe^{-x/2}dx = -xe^{-x/2}\Big|_0^\infty + \int_0^\infty e^{-x/2}dx = 2$$

- To compute variance of *X*, we first compute *E(X²)*:

$$E(X^2) = \frac{1}{2}\int_0^\infty x^2e^{-x/2}dx = -x^2e^{-x/2}\Big|_0^\infty + \int_0^\infty e^{-x/2}dx = 8$$

- Hence, the variance and standard deviation of the device's life are:

$$V(X) = 8 - 2^2 = 4$$
$$\sigma = \sqrt{V(X)} = 2$$

Exponential Distribution

- **The memoryless property:** In probability theory, **memoryless** is a property of certain probability distributions: the **exponential distributions** and the **geometric distributions**, wherein any derived probability from a set of random samples is distinct and has no information (i.e. "memory") of earlier samples.
- Formally, the memoryless property is:

  For all s and t greater or equal to 0:

$$p\left(X > s + t \mid X > s\right) = p\left(X > t\right)$$

- This means that the future event do not depend on the past event, but only on the present event

Normal Distribution

- The **Normal distribution**, also called the **Gaussian distribution**, is an important family of continuous probability distributions, applicable in many fields.
- Each member of the family may be defined by two parameters, **location** and **scale**: **the mean** ("average", μ) and **variance** (standard deviation squared, σ^2) respectively.
- The importance of the normal distribution as a model of quantitative phenomena in the **natural** and **behavioral** sciences is due in part to the **Central Limit Theorem**.
- It is usually used to model system error (e.g. channel error), the distribution of natural phenomena, height, weight, etc.

Normal or Gaussian Distribution

- A **continuous random variable** X, taking all real values in the range (-∞,+∞) is said to follow a **Normal distribution** with parameters **μ** and **σ** if it has the following PDF and CDF:

$$\text{PDF: } f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

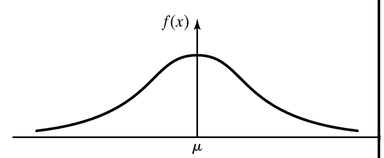
$$\text{CDF: } F(x) = \frac{1}{2}\cdot\left(1+erf\left(\frac{x-\mu}{\sigma\cdot\sqrt{2}}\right)\right)$$ where

$$\text{Error Function: } erf(x) = \frac{2}{\sqrt{\pi}} \cdot \int_0^x \exp\left(-t^2\right)$$

- The Normal distribution is denoted as

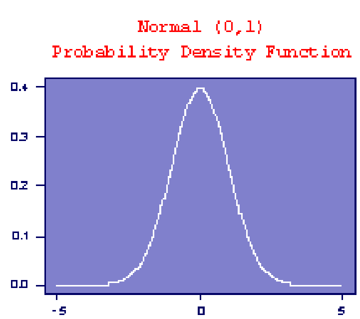
$$X \sim N\left(\mu, \sigma^2\right)$$

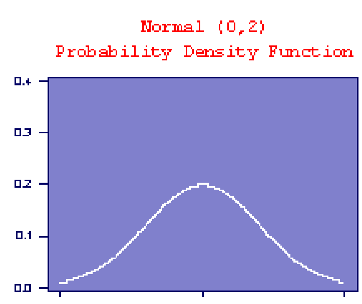
- This probability density function (PDF) is
  - a symmetrical, bell-shaped curve,
  - centered at its expected value μ.
  - The variance is σ^2.



Normal distribution

- Example
- The simplest case of the normal distribution, known as the **Standard Normal Distribution**, has expected value zero and variance one. This is written as $N(0,1)$.





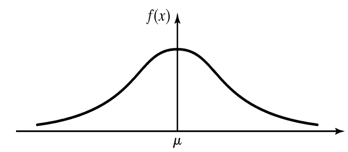
Normal or Gaussian Distribution

- A **continuous random variable** X, taking all real values in the range (-∞,+∞) is said to follow a **Normal distribution** with parameters **μ** and **σ** if it has the following PDF and CDF:

$$\text{PDF: } f(x) = \frac{1}{\sigma\cdot\sqrt{2\pi}}\cdot\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

$$\text{CDF: } F(x) = \frac{1}{2}\cdot\left(1+erf\left(\frac{x-\mu}{\sigma\cdot\sqrt{2}}\right)\right)$$ where $$\text{Error Function: } erf(x) = \frac{2}{\sqrt{\pi}}\cdot\int_0^x \exp\left(-t^2\right)$$

- The Normal distribution is denoted as $X \sim N\left(\mu,\sigma^2\right)$
- This probability density function (PDF) is
  - a symmetrical, bell-shaped curve,
  - centered at its expected value μ.
  - The variance is σ^2.



Standard Normal Distribution

Independent of μ and σ, using the **standard normal distribution**:

- Transformation of variables: let

$$Z \sim N(0,1)$$

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

$$F(x) = P(X \le x) = P\left(Z \le \frac{x-\mu}{\sigma}\right)$$

$$= \int_{-\infty}^{(x-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$= \int_{-\infty}^{(x-\mu)/\sigma} \phi(z)dz = \Phi\left(\tfrac{x-\mu}{\sigma}\right)$$, where $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$

Standard Normal Distribution

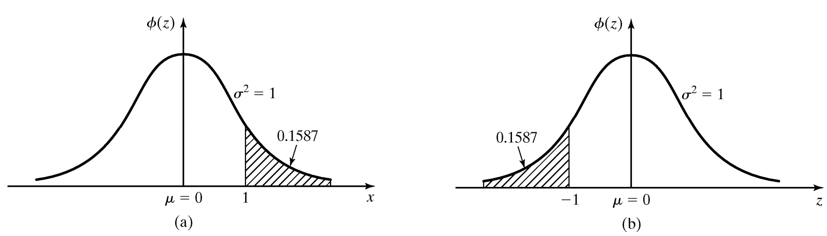
- Note that $f_Z(x)$ *is positive for all* $-\infty < x < \infty$ *hence Z takes on all real* values, its range is the entire real line. Also note that *is an even function* $f_Z(x)$
- The graph of $f_Z(x)$ *is a bell-shaped curve, symmetric about the y-axis.*
- This curve is called a gaussian curve. Its maximum is attained at *x = 0,* then it decreases on both sides of its top point. Actually, it decreases very fast.

Normal Distribution

- Example: The time required to load a transporting truck, *X*, is distributed as *N(12,4)*
  - The probability that the truck is loaded in less than 10 hours:

$$F(10) = \Phi\left(\frac{10-12}{2}\right) = \Phi(-1) = 0.1587$$

  - Using the symmetry property, $\Phi(1)$ is the complement of Φ (-1)



Empirical Distributions

- An Empirical Distribution is a distribution whose parameters are the observed values in a sample of data.
  - May be used when it is impossible or unnecessary to establish that a random variable has any particular parametric distribution.
  - **Advantage**: no assumption beyond the observed values in the sample.
  - **Disadvantage**: sample might not cover the entire range of possible values.

Empirical Distributions

- In statistics, an **empirical distribution function** is a **cumulative probability distribution function** that concentrates probability 1/n at each of the n numbers in a sample.
- Let $x_1, \ldots, x_n$ be iid random variables in with the CDF equal to *F(x)*.
- The empirical distribution function $F_n(x)$ based on sample $x_1, \ldots, x_n$ is a step function defined by

$$F_n(x) = \frac{\text{number of element in the sample } \le x}{n} = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x)$$

where *I*(*A*) is the **indicator of event** *A*. $I(X_i \le x) = \begin{cases} 1 & \text{if } (X_i \le x) \\ 0 & \text{otherwise} \end{cases}$

- For a fixed value *x*, $I(X_i \le x)$ is a **Bernoulli (Trial)** random variable with parameter $\boldsymbol{p=F(x)}$, hence $nF_n(x)$ is a **binomial** random variable with mean $nF(x)$ and variance $nF(x)(1-F(x))$.

Functions of Random Variables

- Example: let $X = U(0,1)$ and $Y = \frac{-1}{\lambda}\ln(1-X),\ \lambda > 0$

Find $F_Y(y)$

Functions of Random Variables

- Example: let $X = U(0,1)$ and $Y = \frac{-1}{\lambda}\ln(1-X),\ \lambda > 0$

Find $F_Y(y)$

$$0 < X < 1$$
$$-1 < -X < 0$$
$$0 < 1 - X < 1$$
$$-\infty < \ln(1-X) < 0$$
$$0 < Y < \infty$$

$$F_Y(y) = P(Y \le y) = P(\frac{-1}{\lambda}\ln(1-X) \le y)$$
$$= P(1 - X \ge e^{-\lambda y}) = P(X \le 1 - e^{-\lambda y})$$

$$F_Y(y) = 1 - e^{-\lambda y}$$
$$f_Y(y) = \lambda e^{-\lambda y}$$

Mean-Variance

Properties of the mean:

$$1) E[cX] = cE[X]$$
$$2) E[c_1 X_1 + c_2 X_2] = c_1 E[X_1] + c_2 E[X_2]$$

Properties of the variance:

$$1) Var(X) \geq 0$$
$$2) Var(cX) = c^2 Var(X)$$
$$3) Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$

Rules

- Exr.
$$1) X = c \Rightarrow Var(X) = 0$$
$$2) Y = aX \Rightarrow Var(Y) = a^2 Var(X)$$
$$3) Y = X + b \Rightarrow Var(Y) = Var(X)$$
$$4) Y = X_1 + X_2 \Rightarrow Var(Y) = Var(X_1) + Var(X_2)$$
- Question Is
$$Var(X - Y) = Var(X) - Var(Y)$$

Example

Let X and Y be independent and

E(X)= 5, E(Y)=-3, and $\sigma_X = 2, \sigma_Y = 3$

Find the mean and he Std Deviation of

Z=3X-2Y-2

E[Z]= 3E[X]-2E[Y]-2 = 15+6-2=19

Var(Z)= 9Var(X)+4Var(Y) = 9.4+4.9 = 72

$\sigma_Z = \sqrt{72}$

Covariance

- The covariance is a measure of dependency between two variables
- Def. Cov(X,Y)= E[(X-E[X]) * (Y-E[Y])

= E[X,Y]-E[X]*E[Y]

For dependent Variables X, Y

Var(X+Y)=Var(X)+Var(Y) + 2E[X,Y] – 2E[X].E[Y]

= Var(X)+Var(Y)+2Cov(X,Y)

Remark: Large Covariance → Large dependency

But If the covariance is 0, it does not mean they are independent.

Rules

(Prove)

1. Cov(X,X)=Var(X)
2. Cov(X,Y)=Cov(Y,X)
3. Cov. Is linear i.e.

Cov(aX+bY,Z)=aCov(X,Z)+bCov(Y,Z)

Exr. For X,Y independent U(0,1)

Find $$Cov(X+2Y, X^2-Y)$$

Correlation Analysis

- In probability theory and statistics, **correlation** (often referred to as a correlation coefficient) indicates **the strength and direction of a linear relationship between two random variables**.
- **Correlation** refers to the departure of two variables from independence. In this broad sense there are several coefficients, measuring the degree of correlation, adapted to the nature of the data.

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y},$$

Correlation Analysis

- Remark: $-1 \le \rho_{(X,Y)} \le 1$
- The sign of ρ indicates the direction of the relationship;
  - ρ near 0 indicates no linear relationship,
  - ρ near 1 or -1 indicates a strong linear relationship.

Correlation Analysis

- Remark: $-1 \le \rho_{(X,Y)} \le 1$
- The sign of ρ indicates the direction of the relationship;
  - ρ near 0 indicates no linear relationship,
  - ρ near 1 or -1 indicates a strong linear relationship.

- Example: For X= U(0,1) find $\rho_{(X,X^2)}$

$Var(X) = \frac{1}{12}$ and $Var(X^2) = E[X^4] - (E[X^2])^2 = \frac{1}{5} - (\frac{1}{3})^2 = \frac{4}{45}$

$$\rho(X, X^2) = \frac{Cov(X, X^2)}{\sigma_X \sigma_{X^2}} = \frac{1/12}{\sqrt{1/12}\sqrt{4/45}} \approx 0.996$$

i.e. very strong dependency