Discrete versus Continuous Random Variables

Discrete Random Variable	**Continuous Random Variable**
$p(x_i) = P(X = x_i)$ 1. $p(x_i) \geq 0$, for all i 2. $\sum_{i=1}^{\infty} p(x_i) = 1$	$f(x)$ 1. $f(x) \geq 0$, for all x in R_X 2. $\int_{R_X} f(x)dx = 1$ 3. $f(x) = 0$, if x is not in R_X
$p(X \leq x) = \sum_{x_i \leq x} p(x_i)$	$p(X \leq x) = \int_{-\infty}^{x} f(t)dt = 0$ $p(a \leq X \leq b) = \int_{a}^{b} f(x)dx$

Mean and Variance

E(X) the expected value

Discrete: $E(x) = \sum_i x_i p(x_i)$

Continuous: $E(x) = \int_0^{\infty} xf(x)$

Var(x) the variance

Discrete $$Var(X) = \sum_{i=0}^{n} (x_i - \mu_i)^2 . p(x_i) = \sum_{i=0}^{n} x_i^2 . p(x_i) - \left(\sum_{i=0}^{n} x_i . p(x_i) \right)^2$$

Continuous $$Var(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 . f(x)dx = \left(\int_{-\infty}^{\infty} x^2 . f(x)dx \right)^2 - \mu_x^2$$

Example: Mean and variance

- When a die is thrown, each of the possible faces 1, 2, 3, 4, 5, 6 (the xi's) has a probability of 1/6 (the p(xi)'s) of showing. The expected value of the face showing is therefore:

  $\mu = E(X) = (1 \times 1/6) + (2 \times 1/6) + (3 \times 1/6) + (4 \times 1/6) + (5 \times 1/6) + (6 \times 1/6) = 3.5$

- Notice that, in this case, E(X) is 3.5, which is not a possible value of X.

Variance

- The **variance** is a measure of the 'spread' of a distribution about its average value.
- Variance is symbolized by or Var(X) or σ^2
  - The **mean** is a way to describe the location of a distribution,
  - the **variance** is a way to capture its scale or degree of being spread out. The unit of variance is the square of the unit of the original variable.

Variance

- The Variance of the random variable X is defined as:

$$V(X) = \sigma_X^2 = E\left(X - E(X)\right)^2 = E\left(X^2\right) - E(X)^2$$

  - where E(X) is the expected value of the random variable X.

- The **standard deviation** is defined as the square root of the variance, i.e.:

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{V(X)} = s$$

Coefficient of Variation

- The Coefficient of Variance of the random variable X is defined as:

$$CV(X) = \frac{V(X)}{E(X)} = \frac{\sigma_X}{\mu_X}$$

Discrete Probability Distribution

- Bernoulli Trials
- Binomial Distribution
- Geometric Distribution
- Poisson Distribution
- Poisson Process

Bernoulli Trials

Any simple trial with two possible outcomes. p and q

EX: Tossing a coin, repeat, with counting # of success *p* *“ the number of heads”*

Then # of failure *q*=(1-*p)* , *“ the number of tails”*

P(HHT) =p.p.q

P(TTT)=q.q.q

If we have k as the number of successes and n-k failures

Then the probability is $p^k q^{n-k}$

Binomial Random Variable

If we have $X : S \to \{0,1,2,3\}$

Where X is the number of successes

$X(sss) = 3$

$X(sfs) = X(ssf) = 2$ ….etc

X now is a random variable.

X is named Binomial random variable resulted from n Bernoulli trials denoted: $b(n, p)$

Modeling of Random Events with Two-States

Binomial Random Variable

Now the probability that $X = k$ $0 \le k \le n$

that is all strings with k success and $n - k$ fails, there are $\binom{n}{k}$ different ways

$$P(X = k) = \binom{n}{k} p^k q^{n-k}$$

Remark: $\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$

Geometric Random Variable

- Consider independent Bernoulli trials are performed until success $s, fs, ffs, fffs, \ldots$

$$P(X = n) = P(fff \ldots fs) = q^{n-1} p = pq^{n-1}$$

- Remark: $\sum_{n=1}^{\infty} pq^{n-1} = p(1 + q + q^2 + \ldots) = p \frac{1}{1-q} = 1$
- Exr: For a geometric variable x compute $P(X > k)$

Geometric Random Variable

PMF: $p(X = k) = \begin{cases} q^{k-1} p, & k = 0,1,2,\ldots,n \\ 0, & \text{otherwise} \end{cases}$

CDF: $\mathrm{F}(X) = p(X \le k) = 1 - (1-p)^k$

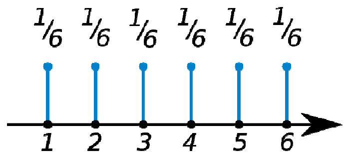
Expected Value : $E[X] = \frac{1}{p}$

Variance : $V[X] = \sigma^2 = \frac{q}{p^2} = \frac{1-p}{p^2}$

Uniform Random Variable

- An R.V. Takes values $1,2,3,\ldots n$ with equal probabilities

$$P(X=k)=\frac{1}{n}$$



Poisson

- For $X=b(n,p)$ with large *n* and small *p*, it is useless to compute the exact $P(X=k)$, as it involves huge calculations of factorial *n*

$$P(X=k)=\binom{n}{k}p^k q^{n-k}=\frac{n(n-1)....(n-k+1)}{fact(k)}p^k q^{n-k}$$

- For large n n-k is approximated to n

$$P(X=k)\approx\frac{n^k}{k!}p^k(1-p)^n=\frac{(np)^k}{k!}\left[(1-p)^{\frac{1}{p}}\right]^{np}$$

$$\lim(1-x)^{\frac{1}{x}}=e^{-1}$$

$$P(X=k)=\frac{(np)^k}{k!}e^{-np}$$

$$P(X=k)=\frac{(\lambda)^k}{k!}e^{-\lambda},\lambda=np$$

Poisson

- Ex. A production line with .4 percent of its items are defective , n=500 items are taken for a quality control. What is the probability that 0, 1, 3 items of them are defective
- That is X=b(500,.004) aprox. To Poisson

$$P(X=k)=\frac{(\lambda)^k}{k!}e^{-\lambda}$$
$$\lambda=500*.004=2$$
$$P(x=0)=e^{-2}$$
$$P(x=1)=2e^{-2}$$
$$P(x=3)=\frac{4}{3}e^{-2}$$

Example: Poisson Distribution

- The number of cars that enter the parking follows a Poisson distribution with a mean rate equal to λ = 20 **cars/hour**
  - The probability of having **exactly** **15 cars** entering the parking in **one hour**:

$$p(15)=P(X=15)=\frac{20^{15}}{15!}\cdot\exp(-20)=0.051649$$

Applications of Poisson

- **Context:** number of events occurring in a fixed period of time
  - Events occur with a known average rate and are independent
- Possion distribution is characterized by the **average rate** λ
  - The average number of arrival in the fixed time period.

- **Examples**
  - The number of cars passing a fixed point in a 5 minute interval. Average rate: λ = 3 cars/5 minutes
  - The number of calls received by a switchboard during a given period of time. Average rate: λ =3 call/minutes
  - The number of message coming to a router per second
  - The number of travelers arriving to the airport for flight registration

Poisson Distribution

- The Poisson distribution with the average rate parameter λ

PMF: $$p(k) = P(X = k) = \begin{cases} \frac{\lambda^k}{k!}\exp(-\lambda) & \text{for } k = 0,1,2,\ \ldots. \\ 0, & \text{otherwise} \end{cases}$$

CDF: $$F(k) = p(X \le k) = \sum_{i=0}^{k} \frac{\lambda^i}{i!} \cdot \exp(-\lambda)$$

Expected value: $E[X] = \lambda$

Variance: $V[X] = \lambda$

Continuous Probability Distribution

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

Continuous Uniform Distribution

- The **continuous uniform distribution** is a family of probability distributions such that for each member of the family, all intervals of the same length on the distribution's support are **equally probable**
- A random variable X is **uniformly distributed** on the interval $[a,b]$, $U(a,b)$, if its PDF and CDF are:

PDF: $$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

CDF: $$F(x) = \begin{cases} 0, & x \prec a \\ \frac{x-a}{b-a}, & a \le x \prec b \\ 1, & x \ge b \end{cases}$$

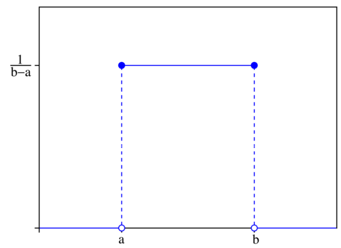
Expected value: $$E[X] = \frac{a+b}{2}$$

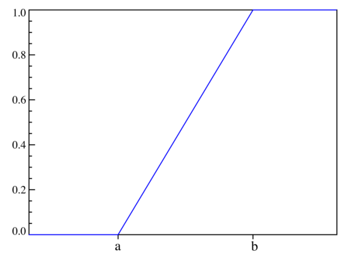
Variance: $$V[X] = \frac{(a+b)^2}{12}$$

Uniform Distribution U(a,b)

- The PDF is $f(x) = const = \frac{1}{b-a}$
- **Properties**
  - $p(x1 \le X \le x2)$ is proportional to the length of the interval

$$F(X_2) - F(X_1) = \frac{X_2 - X_1}{b-a}$$





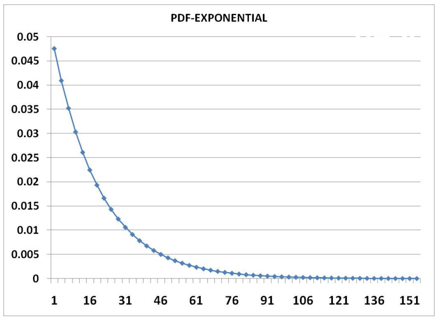
- **Special case:** a **standard uniform** distribution U(0,1).
  - Very useful for random number generators in simulators

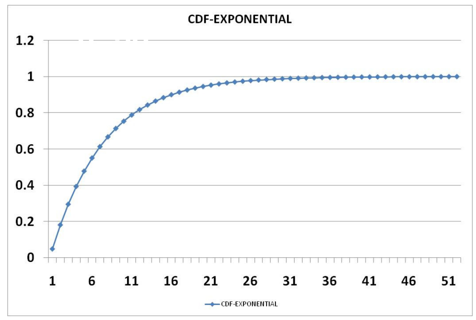
Exponential Distribution

$$\text{PDF:} \quad f(x) = \begin{cases} \lambda \cdot \exp(-\lambda \cdot x), & x \ge 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\text{CDF:} \quad F(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, & x \ge 0 \end{cases}$$

Exponential Distribution





$$f(x) = \begin{cases} \frac{1}{20} \cdot \exp\left(-\frac{x}{20}\right), & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \exp\left(-\frac{x}{20}\right), & x \geq 0 \end{cases}$$

Exponential Distribution

- (Special interest)The **exponential distribution** describes the times between events in a Poisson process, in which events occur continuously and independently at a constant average rate.
- A random variable X is **exponentially distributed** with parameter $\mu = 1/\lambda > 0$ if its PDF and CDF are:

PDF: $f(x) = \begin{cases} \lambda \cdot \exp(-\lambda \cdot x), & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \Rightarrow f(x) = \begin{cases} \frac{1}{\mu} \cdot \exp\left(-\frac{x}{\mu}\right), & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$

CDF: $F(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, & x \geq 0 \end{cases} \Rightarrow F(x) = \begin{cases} 0, & x < 0 \\ 1 - \exp\left(-\frac{x}{\mu}\right), & x \geq 0 \end{cases}$

Expected value: $E[X] = \frac{1}{\lambda} = \mu$

Variance: $V[X] = \frac{1}{\lambda^2} = \mu^2$

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, σ2) 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:

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

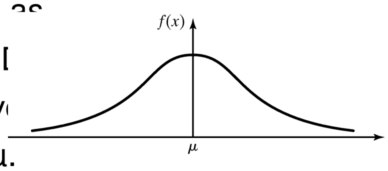
$$\text{CDF: } F(x) = \frac{1}{2} \cdot \left(1 + erf\left(\frac{x-\mu}{\sigma \cdot \sqrt{2}}\right)\right)$$

where

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

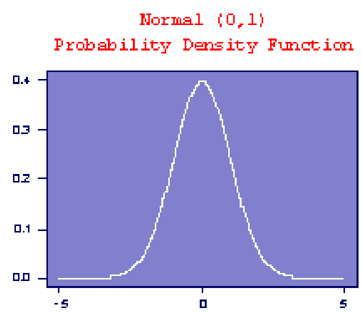
$$X \sim N\left(\mu, \sigma^2\right)$$

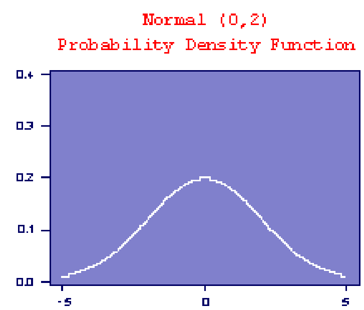
- The Normal distribution is denoted as
- This probability density function (PD
  - a symmetrical, bell-shaped curv
  - centered at its expected value μ.



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 $N(0,1)$.





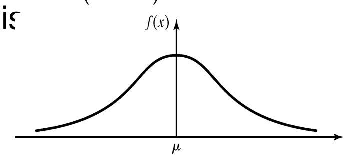
Normal or Gaussian Distribution

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

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

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

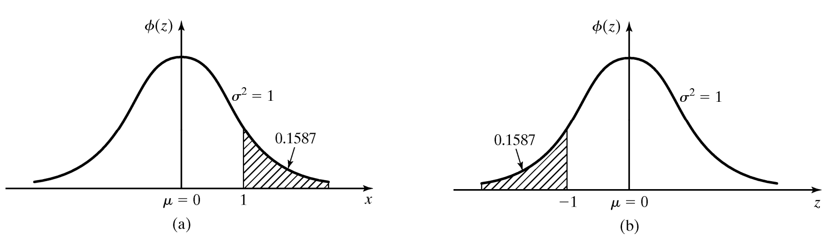
- Note that $f_Z(x)$ *is positive for all* $-\infty < x < \infty$ *, hence Z takes on all real* values, its range is the entire real line. Also note that $f_Z(x)$ *is an even function*
- The graph of $f_Z(x)$ *is a bell-shaped curve, symmetric about the y-axis.*
- This curve is called a gaussian curve. 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

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

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

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