Joint Distribution of RVs

- In real life, we are often interested in two (or more) random variables at the same time. For example,
  - we might measure the height and weight of an object, or
  - frequency of exercise and rate of heart disease in adults,
  - level of air pollution and rate of respiratory illness in cities,
  - number of Facebook friends and age of Facebook members
- Joint distribution allows us to compute probabilities of events involving both variables and understand the relationship between the variables.

Joint Probability Mass Function (PMF)

The **joint probability mass function** of two discrete random variables X and Y is defined as

$$P_{XY}(x,y) = P(X=x, Y=y).$$

In particular, if $R_X = \{x_1, x_2, \ldots\}$ and $R_Y = \{y_1, y_2, \ldots\}$, then we can always write

$$\begin{aligned} R_{XY} &\subset R_X \times R_Y \\ &= \{(x_i, y_j) | x_i \in R_X, y_j \in R_Y\}. \end{aligned}$$

$$\sum_{(x_i, y_j) \in R_{XY}} P_{XY}(x_i, y_j) = 1$$

Joint Probability Mass Function (PMF)

We can use the joint PMF to find $P\big((X,Y) \in A\big)$ for any set $A \subset \mathbb{R}^2$. Specifically, we have

$$\boxed{P\big((X,Y) \in A\big) = \sum_{(x_i,y_j)\in(A\cap R_{XY})} P_{XY}(x_i, y_j)}$$

$$\begin{aligned} P_{XY}(x,y) &= P(X=x, Y=y) \\ &= P\big((X=x) \cap (Y=y)\big). \end{aligned}$$

Marginal PMFs

Marginal PMFs of X and Y:

$$\begin{aligned} P_X(x) &= \sum_{y_j \in R_Y} P_{XY}(x, y_j), \qquad \text{for any } x \in R_X \\ P_Y(y) &= \sum_{x_i \in R_X} P_{XY}(x_i, y), \qquad \text{for any } y \in R_Y \end{aligned} \tag{5.1}$$

Example
Consider two random variables X and Y with joint PMF given in Table

- Find P(X=0,Y≤1)
- Find the marginal PMFs of X and Y
- Find P(Y=1|X=0)
- Are X and Y independent?

	$Y=0$	$Y=1$	$Y=2$
$X=0$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{8}$
$X=1$	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{6}$

Joint Cumulative Distributive Function (CDF)

The **joint cumulative distribution function** of two random variables X and Y is defined as

$$F_{XY}(x,y) = P(X \le x, Y \le y).$$

$$\begin{aligned} F_{XY}(x,y) &= P(X \le x, Y \le y) \\ &= P\big((X \le x) \text{ and } (Y \le y)\big) = P\big((X \le x) \cap (Y \le y)\big). \end{aligned}$$

Marginal CDFs of X and Y

Marginal CDFs of X and Y:

$$F_X(x) = F_{XY}(x, \infty) = \lim_{y \to \infty} F_{XY}(x, y), \qquad \text{for any } x,$$
$$F_Y(y) = F_{XY}(\infty, y) = \lim_{x \to \infty} F_{XY}(x, y), \qquad \text{for any } y \qquad (5.2)$$

Also, note that we must have

$$F_{XY}(\infty, \infty) = 1,$$
$$F_{XY}(-\infty, y) = 0, \qquad \text{for any } y,$$
$$F_{XY}(x, -\infty) = 0, \qquad \text{for any } x.$$

Conditioning and Independence

- **Conditional PMF & CDF**

For a discrete random variable X and event A, the **conditional PMF** of X given A is defined as

$$\begin{aligned} P_{X|A}(x_i) &= P(X = x_i|A) \\ &= \frac{P(X = x_i \text{ and } A)}{P(A)}, \quad \text{for any } x_i \in R_X. \end{aligned}$$

Similarly, we define the **conditional CDF** of X given A as

$$F_{X|A}(x) = P(X \leq x|A).$$

Question

- I roll a fair die. Let X be the observed number. Find the conditional PMF of X given that we know the observed number was less than 5.

Conditional PMF of X Given Y

For discrete random variables X and Y, the **conditional PMFs** of X given Y and vice versa are defined as

$$P_{X|Y}(x_i|y_j) = \frac{P_{XY}(x_i, y_j)}{P_Y(y_j)},$$
$$P_{Y|X}(y_j|x_i) = \frac{P_{XY}(x_i, y_j)}{P_X(x_i)}$$

for any $x_i \in R_X$ and $y_j \in R_Y$.

Independent Random Variables

Two discrete random variables X and Y are independent if

$$P_{XY}(x,y) = P_X(x)P_Y(y), \quad \text{for all } x, y.$$

Equivalently, X and Y are independent if

$$F_{XY}(x,y) = F_X(x)F_Y(y), \quad \text{for all } x, y.$$

Conditional Expectation

Conditional Expectation of X:

$$E[X|A] = \sum_{x_i \in R_X} x_i P_{X|A}(x_i),$$

$$E[X|Y = y_j] = \sum_{x_i \in R_X} x_i P_{X|Y}(x_i|y_j)$$

Law of Total Probability

Law of Total Probability:

$$P(X \in A) = \sum_{y_j \in R_Y} P(X \in A | Y = y_j) P_Y(y_j), \quad \text{for any set } A.$$

Law of Total Expectation:

1. If $B_1, B_2, B_3, \ldots$ is a partition of the sample space S,

$$EX = \sum_i E[X|B_i]P(B_i) \tag{5.3}$$

2. For a random variable X and a discrete random variable Y,

$$EX = \sum_{y_j \in R_Y} E[X|Y = y_j] P_Y(y_j) \tag{5.4}$$

Functions of Two Random Variables

- Suppose that you have two discrete random variables X and Y, and suppose that Z=g(X,Y), where g: ℝ2↦ ℝ. Then, if we are interested in the PMF of Z, we can write

$$\begin{aligned} P_Z(z) &= P(g(X,Y)=z) \\ &= \sum_{(x_i,y_j)\in A_z} P_{XY}(x_i,y_j), \quad \text{where } A_z = \{(x_i,y_j)\in R_{XY} : g(x_i,y_j)=z\}. \end{aligned}$$

Note that if we are only interested in $E[g(X,Y)]$, we can directly use LOTUS, without finding $P_Z(z)$:

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

$$E[g(X,Y)] = \sum_{(x_i,y_j)\in R_{XY}} g(x_i,y_j)P_{XY}(x_i,y_j) \qquad (5.5)$$

Two Continuous Random Variables

Joint Probability Density Function (PDF)

Definition 5.1
Two random variables X and Y are **jointly continuous** if there exists a nonnegative function $f_{XY} : \mathbb{R}^2 \to \mathbb{R}$, such that, for any set $A \in \mathbb{R}^2$, we have

$$P\big((X,Y) \in A\big) = \iint_A f_{XY}(x,y)dxdy \qquad (5.15)$$

The function $f_{XY}(x,y)$ is called the **joint probability density function (PDF)** of X and Y.

In the above definition, the domain of $f_{XY}(x,y)$ is the entire $\mathbb{R}^2$. We may define the range of (X,Y) as

$$R_{XY} = \{(x,y)|f_{X,Y}(x,y) > 0\}.$$

The above double integral (Equation 5.15) exists for all sets A of practical interest. If we choose $A = \mathbb{R}^2$, then the probability of $(X,Y) \in A$ must be one, so we must have

$$\boxed{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1}$$

Marginal PDFs of X and Y from their joint PDF

Marginal PDFs

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy, \qquad \text{for all } x,$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx, \qquad \text{for all } y.$$

Example 5.15

Let X and Y be two jointly continuous random variables with joint PDF

$$f_{XY}(x,y) = \begin{cases} x + cy^2 & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ \\ 0 & \text{otherwise} \end{cases}$$

a. Find the constant c.
b. Find $P(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2})$.

Joint Cumulative Distribution Function (CDF)

The **joint cumulative function** of two random variables X and Y is defined as

$$F_{XY}(x, y) = P(X \leq x, Y \leq y).$$

The joint CDF satisfies the following properties:

1. $F_X(x) = F_{XY}(x, \infty)$, for any x (marginal CDF of X);
2. $F_Y(y) = F_{XY}(\infty, y)$, for any y (marginal CDF of Y);
3. $F_{XY}(\infty, \infty) = 1$;
4. $F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = 0$;
5. $P(x_1 < X \leq x_2,\ \ y_1 < Y \leq y_2) =$
$$F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1);$$
6. if X and Y are independent, then $F_{XY}(x, y) = F_X(x)F_Y(y)$.

Conditioning

If X is a continuous random variable, and A is the event that $a < X < b$ (where possibly $b = \infty$ or $a = -\infty$), then

$$F_{X|A}(x) = \begin{cases} 1 & x > b \\ \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} & a \le x < b \\ 0 & x < a \end{cases}$$

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{P(A)} & a \le x < b \\ 0 & \text{otherwise} \end{cases}$$

Conditioning

For two jointly continuous random variables X and Y, we can define the following conditional concepts:

1. The conditional PDF of X given $Y = y$:

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

2. The conditional probability that $X \in A$ given $Y = y$:

$$P(X \in A|Y = y) = \int_A f_{X|Y}(x|y)dx$$

3. The conditional CDF of X given $Y = y$:

$$F_{X|Y}(x|y) = P(X \leq x|Y = y) = \int_{-\infty}^{x} f_{X|Y}(x|y)dx$$

Conditional expectation and variance

$$E[X|A] = \int_{-\infty}^{\infty} x f_{X|A}(x)dx,$$
$$E[g(X)|A] = \int_{-\infty}^{\infty} g(x) f_{X|A}(x)dx,$$
$$\mathrm{Var}(X|A) = E[X^2|A] - (E[X|A])^2$$

For two jointly continuous random variables X and Y, we have:

1. Expected value of X given $Y = y$:

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

2. Conditional LOTUS:

$$E[g(X)|Y = y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y) dx$$

3. Conditional variance of X given $Y = y$:

$$Var(X|Y = y) = E[X^2|Y = y] - (E[X|Y = y])^2$$

Example

Example 5.21 Let X and Y be two jointly continuous random variables with joint PDF

$$f_{XY}(x,y) = \begin{cases} \frac{x^2}{4} + \frac{y^2}{4} + \frac{xy}{6} & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

For $0 \leq y \leq 2$, find

a. the conditional PDF of X given $Y = y$;
b. $P(X < \frac{1}{2} | Y = y)$.

- c. E[X|Y=1]
- D. Var[X|Y=1]

Independent Random Variables

Two continuous random variables X and Y are independent if

$$f_{XY}(x,y) = f_X(x) f_Y(y), \quad \text{for all } x, y.$$

Equivalently, X and Y are independent if

$$F_{XY}(x,y) = F_X(x) F_Y(y), \quad \text{for all } x, y.$$

If X and Y are independent, we have

$$E[XY] = EXEY,$$
$$E[g(X)h(Y)] = E[g(X)]E[h(Y)].$$

Example

Determine whether X and Y are independent:

a. $f_{XY}(x,y) = \begin{cases} 2e^{-x-2y} & x, y > 0 \\ 0 & \text{otherwise} \end{cases}$

b. $f_{XY}(x,y) = \begin{cases} 8xy & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$

Functions of Two Continuous Random Variables

LOTUS for two continuous random variables:

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y) \;\; dxdy \qquad (5.19)$$

Covariance and Correlation

Covariance

- The covariance gives some information about how X and Y are statistically related
  - If the covariance is **positive**, the variables tend to increase or decrease together.
  - If the covariance is **negative**, one variable tends to increase when the other decreases.
  - If the covariance is **zero**, the variables are likely independent (though not necessarily).

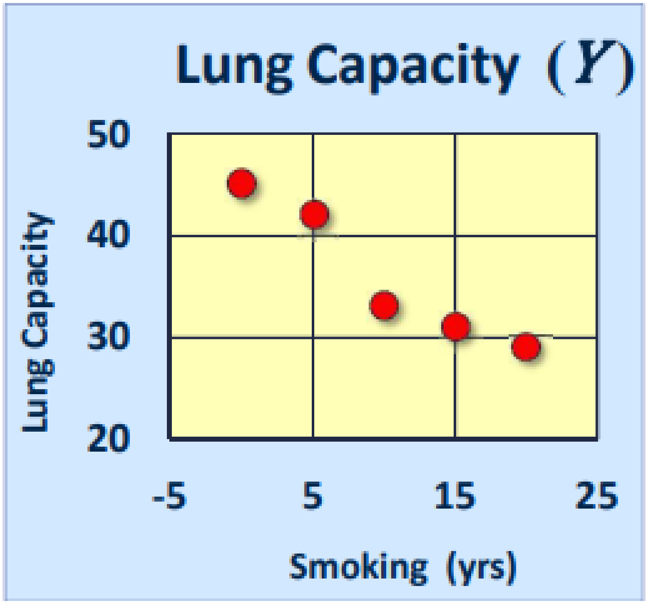
The **covariance** between X and Y is defined as

$$\mathrm{Cov}(X,Y) = E\big[(X - EX)(Y - EY)\big] = E[XY] - (EX)(EY).$$

Example

N	Cigarettes (X)	Lung Capacity (Y)
1	0	45
2	5	42
3	10	33
4	15	31
5	20	29

- Variables smoking and lung capacity *covary* inversely, like



Properties of covariance

Lemma 5.3

The covariance has the following properties:

1. $\text{Cov}(X, X) = \text{Var}(X)$;
2. if X and Y are independent then $\text{Cov}(X, Y) = 0$;
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
4. $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$;
5. $\text{Cov}(X + c, Y) = \text{Cov}(X, Y)$;
6. $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$;
7. more generally,

$$\text{Cov}\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j \text{Cov}(X_i, Y_j).$$

Example

Let X and Y be two independent N(0,1) random variables and

$Z=1+X+XY^2$

$W=1+X$.

Find Cov(Z,W)

Variance of a sum

- One of the applications of covariance is finding the variance of a sum of several random variables. In particular, if Z=X+Y, then
- Var(*Z*)=Cov(*Z*,*Z*)

  =Cov(*X*+*Y*,*X*+*Y*)

  =Cov(*X*,*X*)+Cov(*X*,*Y*)+Cov(*Y*,*X*)+Cov(*Y*,*Y*)

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

More generally, for $a, b \in \mathbb{R}$, we conclude:

$$\boxed{\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y) \qquad (5.21)}$$

Correlation Coefficient

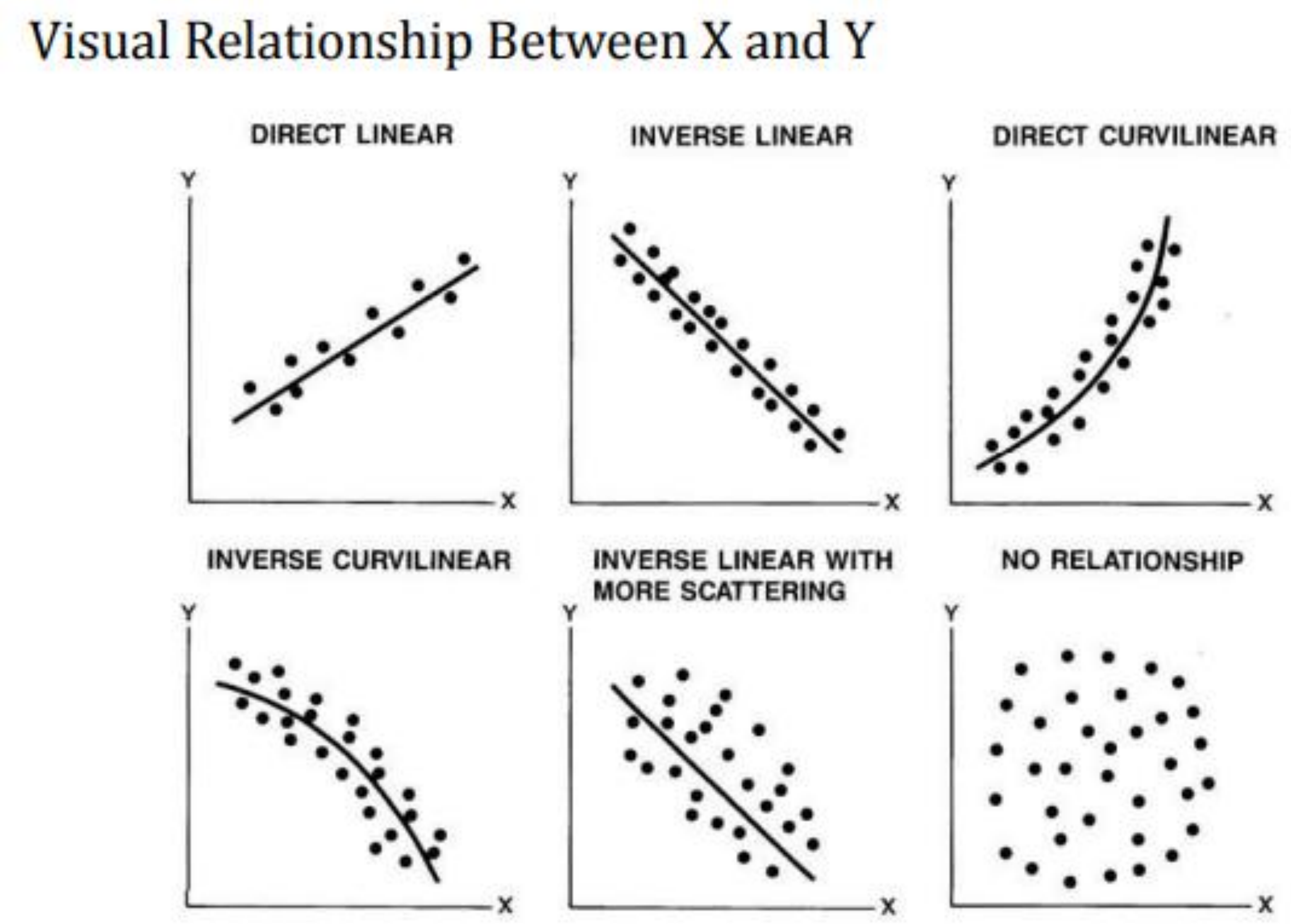
- The correlation coefficient, denoted by ρXY or ρ(X,Y), is obtained by normalizing the covariance.
- In particular, we define the correlation coefficient of two random variables X and Y as the covariance of the standardized versions of X and Y
- **Correlation is a standardized measure of the relationship between two variables, ranging from -1 to 1.**

Definition 5.2
Consider two random variables X and Y:
- If $\rho(X, Y) = 0$, we say that X and Y are **uncorrelated**.
- If $\rho(X, Y) > 0$, we say that X and Y are **positively** correlated.
- If $\rho(X, Y) < 0$, we say that X and Y are **negatively** correlated.

- Define the standardized versions of X and Y as

Correlation of RVs

- Correlation using scatter plot



$$U = \frac{X - EX}{\sigma_X}, \quad V = \frac{Y - EY}{\sigma_Y} \qquad (5.22)$$

Then,

$$\begin{aligned}
\rho_{XY} = \mathrm{Cov}(U, V) &= \mathrm{Cov}\left(\frac{X - EX}{\sigma_X}, \frac{Y - EY}{\sigma_Y}\right) \\
&= \mathrm{Cov}\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) && \text{(by Item 5 of Lemma 5.3)} \\
&= \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}.
\end{aligned}$$

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

Properties of the correlation coefficient:

Interpretation:

- $\rho = 1$: Perfect **positive** correlation (both increase together).
- $\rho = -1$: Perfect **negative** correlation (one increases, the other decreases).
- $\rho = 0$: No linear relationship.

Properties of Correlation:

1. $-1 \leq \rho(X, Y) \leq 1$.
2. **Scaling does not affect correlation:** $\rho(aX, bY) = \rho(X, Y)$ if $a, b > 0$.
3. **Correlation measures only linear relationships** (nonlinear dependencies might exist even if $\rho = 0$).

IF X and Y are uncorrelated, then

If X and Y are uncorrelated, then

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y).$$

More generally, if $X_1, X_2, \ldots, X_n$ are pairwise uncorrelated, i.e., $\rho(X_i, X_j)=0$ when $i \neq j$, then

$$\mathrm{Var}(X_1+X_2+\ldots+X_n)=\mathrm{Var}(X_1)+\mathrm{Var}(X_2)+\ldots+\mathrm{Var}(X_n).$$

Problem 1

Let X and Y be two jointly continuous random variables with joint PDF

$$f_{XY}(x,y) = \begin{cases} 2 & y+x \leq 1, x>0, y>0 \\ \\ 0 & \text{otherwise} \end{cases}$$

Find $\mathrm{Cov}(X,Y)$ and $\rho(X,Y)$.

Central Limit Theorem

- It states that, under certain conditions, the sum of a large number of random variables is approximately normal.

The Central Limit Theorem (CLT)

Let X_1,X_2,...,X_n be i.i.d. random variables with expected value $EX_i = \mu < \infty$ and variance $0 < \mathrm{Var}(X_i) = \sigma^2 < \infty$. Then, the random variable

$$Z_n = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma}$$

converges in distribution to the standard normal random variable as n goes to infinity, that is

$$\lim_{n \to \infty} P(Z_n \leq x) = \Phi(x), \qquad \text{for all } x \in \mathbb{R},$$

where $\Phi(x)$ is the standard normal CDF.

Applications of CLT

- **Hypothesis Testing –** Many statistical tests, like t-tests and z-tests, rely on CLT to assume normality in sample means.
- **Confidence Intervals –** CLT allows us to estimate population parameters using confidence intervals.
- **Quality Control in Manufacturing –** Helps monitor the variation in product quality by analyzing sample means.
- **Financial Market Analysis –** Used in risk assessment and portfolio management to approximate return distributions.
- **Election Polling & Surveys –** Polling organizations use CLT to predict election results based on sample survey data.

Common Distributions of RVs

Common Distributions of RVs

The Uniform Distribution

A random variable X is said to be *uniformly distributed* in $a \le x \le b$ if its density function is

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

and the distribution is called a *uniform distribution*.

The distribution function is given by

$$F(x) = P(X \le x) = \begin{cases} 0 & x < a \\ (x-a)/(b-a) & a \le x < b \\ 1 & x \ge b \end{cases}$$

The mean and variance are, respectively,

$$\mu = \frac{1}{2}(a+b), \qquad \sigma^2 = \frac{1}{12}(b-a)^2$$

Common Distributions of RVs

The Poisson Distribution

Let X be a discrete random variable that can take on the values 0, 1, 2, . . . such that the probability function of X is given by

$$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, 2, \ldots \tag{13}$$

where λ is a given positive constant. This distribution is called the *Poisson distribution* (after S. D. Poisson, who discovered it in the early part of the nineteenth century), and a random variable having this distribution is said to be *Poisson distributed*.

Mean	$\mu = \lambda$
Variance	$\sigma^2 = \lambda$
Standard deviation	$\sigma = \sqrt{\lambda}$

When p is small and n is fixed, Mean = λ = np, where

- **n** is the Number of Trails
- **p** is Probability of Success

Common Distributions of RVs

The Normal Distribution

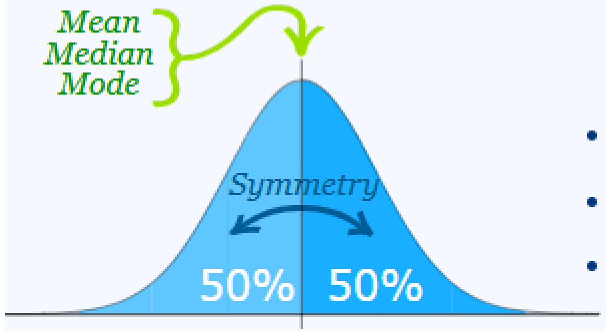
One of the most important examples of a continuous probability distribution is the *normal distribution*, sometimes called the *Gaussian distribution*. The density function for this distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty \tag{4}$$

where μ and σ are the mean and standard deviation, respectively. The corresponding distribution function is given by

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(v-\mu)^2/2\sigma^2}\, dv \tag{5}$$

If X has the distribution function given by (5), we say that the random variable X is *normally distributed* with mean μ and variance σ^2.



The **Normal Distribution** has:

- mean = median = mode
- symmetry about the center
- 50% of values less than the mean and 50% greater than the mean

Common Distributions of RVs

Standard normal distribution, also known as the z-distribution

- In this distribution, the **mean (average) is 0** and the **standard deviation (a measure of spread) is 1**.
- This creates a **bell-shaped curve** that is symmetrical around the mean ie. 0.
- The random variable of a standard normal distribution is known as the standard score or a z-score.

$$z = (X - \mu) / \sigma$$

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Where
$-\infty < z < \infty$