Queen Mary
University of London

Maths & Stats Pre-Sessional

Probability Distributions

Lecturer: Claudio Vallar
School of Economics and Finance

Probability Distributions

In this session:

- We will review some basic statistical concepts in probability theory
- Probability rules
- Probability distribution functions for discrete and continuous variables
- Normal Distribution

For more extensive reading, refer to Chapter 3 and 5 of Newbold, P., Carlson, W., and Thorne, B. (2010). Statistics for Business and Economics, Pearson, 7th Edition

Definitions

Definitions. A **random experiment** is a process leading to two or more possible outcomes, without knowing exactly which outcome will occur.

Basic outcomes are the possible outcomes of a random experiment.

A **sample space** (S) is the set of all basic outcomes.

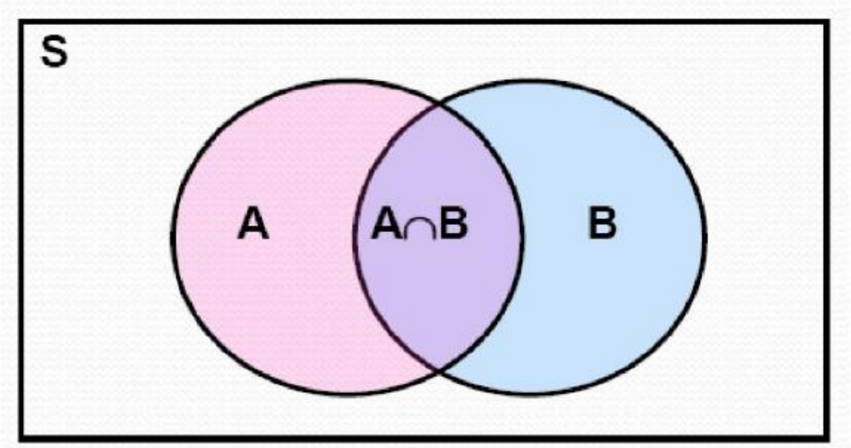
An **event** (E) is any subset of the basic outcomes from the sample space. An event occurs if the random experiment results in one of its constituent basic outcomes. The null event represents the absence of a basic outcome and is denoted by $\emptyset$.

Example:

- Random experiment: Throwing a dice.
- The basic outcomes are 1, 2, 3, 4, 5, 6.
- The sample space S = {1, 2, 3, 4, 5, 6}
- An event of scoring even: {2, 4, 6}. An event of scoring < 4: {1, 2, 3}

Key Definition

Definition. The **intersection** of events A and B, $A \cap B$, is a set of outcomes that belong to both events A and B.



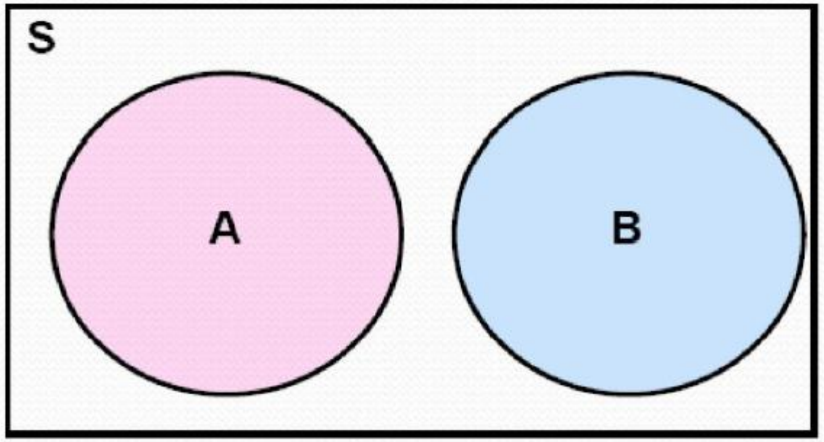
Example (dice).

If A= {1,2,3} and B = { 2,4,6}, then the intersection of A and B, denoted, A ∩ B, is

$$A \cap B = 2 .$$

Key Definition

Definition. If the events A and B have no common basic outcomes, they are called **mutually exclusive**, and their intersection, $A \cap B$, is said to be the empty set, $\emptyset$.

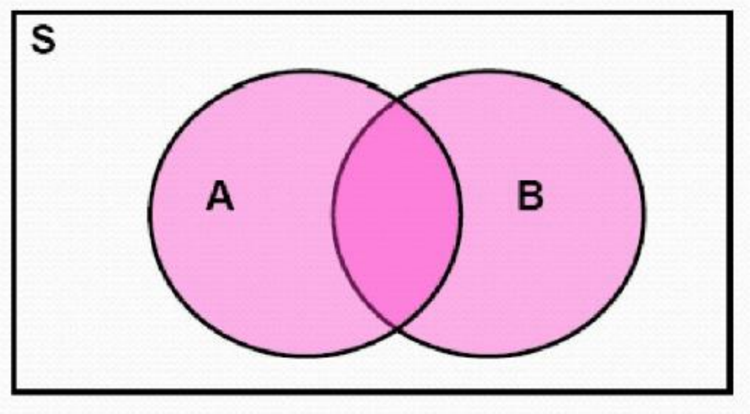


Example (dice).

If A= {1,3,5} and B = { 2,4,6}, then A and Bare mutually exclusive

Key Definition

Definition. Let A and B be two events in the sample space, S. Their **union**, $A \cup B$, is the set of all basic outcomes in S that belong to at least one of these two events.



Example (dice).

If A= {1,2,3} and B = { 2,4,6}, then the union of A and B, denoted, A ∪ B, is

$$A \cup B = \{1,2,3,4,6\}\ .$$

Key Definition

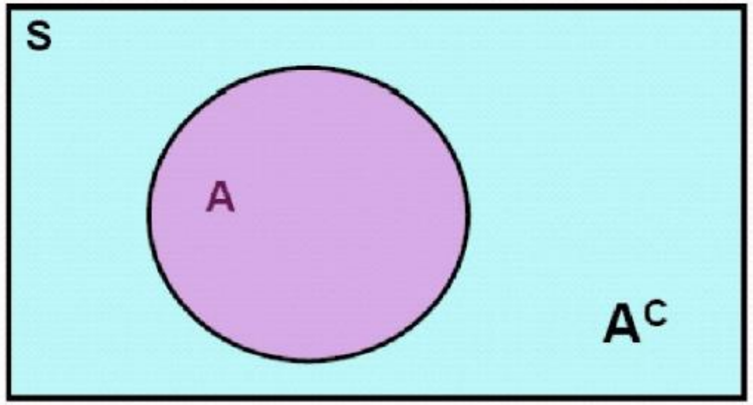
Definition. Given K events $E_1, E_2, \dots, E_K$ in the sample space, S, if $E_1 \cup E_2 \cup \cdots \cup E_K = S$, these K events are said to be **collectively exhaustive**.

Example (dice).

If A= {1,2} and B = { 2,3,4,5,6}, then A ∪ B = *S*. Hence, A and B are collectively exhaustive.

Key Definition

Definition. Let A be an event in the sample space, S. The set of basic outcomes of a random experiment belonging to S but not to A is called the **complement** of A and is denoted by $\bar{A}$ or A^c.



<u>Example (dice)</u>.

If A= {1,2,3} and S = { 1,2,3,4,5,6}, then the complement of A is

$$\bar{A} = \{4,5,6\} .$$

Key Definition

Example.

Let S = {1,2,3,4,5,6,7,8}, A = {3,7,8} and B = {2,3,6,8}

Find:

1. $A \cup B$
2. $A \cap B$
3. $\bar{A}$ and $\bar{B}$

Key Definition

- Probability measures how often (likely) an event will occur. The probability of event A occurring is denoted by P(A).

- For the measure of likelihood to make sense, we require probabilities to satisfy three conditions:

  1. If A is any event in the sample space, S, then $0 \leq P(A) \leq 1$.

  2. Let A be an event in S and let O_i denote the basic outcomes. Then, P(A) = $\sum_A P(O_i)$, where the notation implies that the summation extends over all the basic outcomes in A.

  3. P(S)= 1.

Probability Rules: Complement Rule and Conditional Probability

Complement Rule:

$$P(\bar{A}) = 1 - P(A).$$

Conditional Probability:

What is the probability that event A occurs, given that event B has already happened?

Definition. Let A and B be two events. The **conditional probability** of event A, given that event B has occurred, is denoted by $P(A|B)$ and is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ provided that } P(B) > 0.$$

Similarly,

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \text{ provided that } P(A) > 0.$$

Probability Rules: Addition and Multiplication Rules

Addition Rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

If A and B are mutually exclusive, then the addition rule is reduced to

$$P(A \cup B) = P(A) + P(B).$$

Multiplication Rule:

What is the probability that event A occurs, given that event B has already happened?

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Statistical Independence

Definition. Let A and B be two events. These events are **statistically independent** if and only if

$$P(A \cap B) = P(A)P(B).$$

This implies that if A and B are statistically independent, then

$$P(A|B) = P(A) \text{ if } P(B) > 0.$$

$$P(B|A) = P(B) \text{ if } P(A) > 0.$$

Example in Finance

Type of Fund	AB Fund	SM Fund	GW Fund	FC Fund	DD Fund	MX Fund	ZK Fund
Equity Fund	X	X	X		X		
Hybrid Fund		X	X	X	X	X	
Money Market Fund	X	X	X				X

- The table above compares seven investment funds in terms of the types of funds that each fund operates. There are different types of funds under consideration, notably equity fund, money market fund, hybrid fund (i.e., a mixture of equity and money market instruments). The seven investment funds operate one or more of these fund types.

Example in Finance

Identify:

1. Equity Fund ∪ Hybrid Fund ∪ Money Market Fund:
2. Equity Fund ∩ Money Market Fund ∩ Hybrid Fund
3. Equity Fund ∩ Money Market Fund
4. Equity Fund ∩ Hybrid Fund
5. Funds without other fund types save for hybrid fund.
6. Which funds have money market and hybrid funds exclusively?.
7. Which funds have equity funds exclusively?

Example in Finance

1 . Equity Fund ∪ Hybrid Fund ∪ Money Market Fund:

The universe of all investment funds {AB Fund, SM Fund, GW Fund, DD Fund, FC Fund, MX Fund}

2. Equity Fund ∩ Money Market Fund ∩ Hybrid Fund

Only two funds {SM Fund, GW Fund} have all three fund types save only a hedge fund.

3 . Equity Fund ∩ Money Market Fund

The AB Fund is the only fund that has both equity and money market funds only.

4. Equity Fund ∩ Hybrid Fund

The DD Fund is the only fund that has both equity and hybrid funds only.

5. The FC Fund and MX Fund are the only funds without other fund types save for hybrid fund.

6. There are no funds that have money market and hybrid funds exclusively.

7. There are no funds that have equity funds exclusively.

Example in Finance

Benefits of Venn Diagrams:

- **Visual organization**: they assist in the visual depiction of information, which helps to understand the logic behind the relationships of particular elements.
- **Assist in making decisions**: Venn diagrams assist in making decisions between two or more choices. It makes it easier to compare and contrast.
- **Reason through logic**: they help to reason complex issues through logic.
- **Detect data patterns**: It is easier to detect data patterns that may not have been evident. Patterns such as probabilities and correlations are easily deduced.

Maths & Stats Pre-Sessional

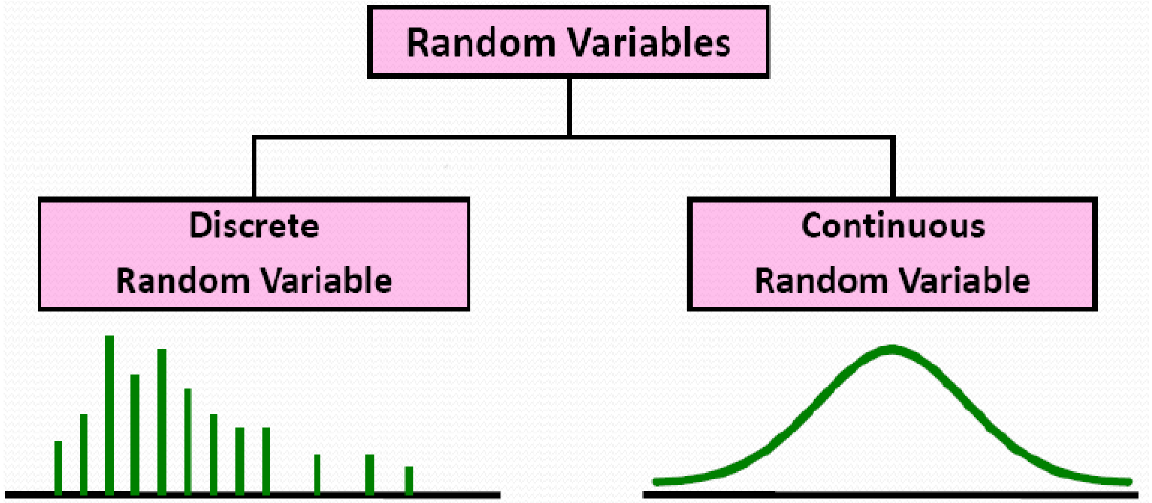
Random Variables: PDF and CDF

Lecturer: Claudio Vallar
School of Economics and Finance

Random Variable

Definition: A **random variable** is a variable that takes on numerical values realized by the outcomes in the sample space generated by a random experiment.

Random variables can be discrete or continuous.



Discrete Random Variable: takes on a finite number of values (e.g coin toss, die roll)

Continuous Random Variable: takes on any value in a real interval (e.g. temperature)

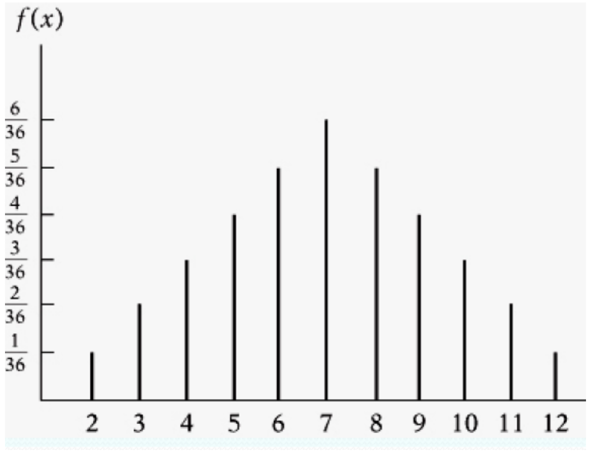
Probability Density Function (PDF)

The **probability density function (PDF)** of X summarizes the information concerning the possible outcomes of X and the corresponding probabilities. The notation used is $f(x)$.

For a <u>*discrete random variable*</u>: $f(x) = prob(X = x)$

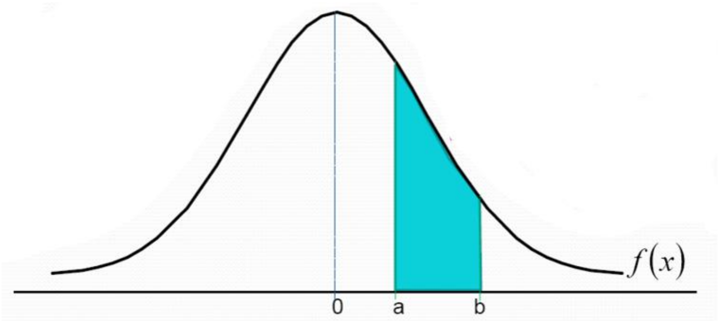
The axioms of probability require that:

- $0 \leq Prob(X = x) \leq 1$
- $\sum_x f(x) = 1$



For a <u>*continuous random variable*</u>: the probability associated with any particular point is zero and:

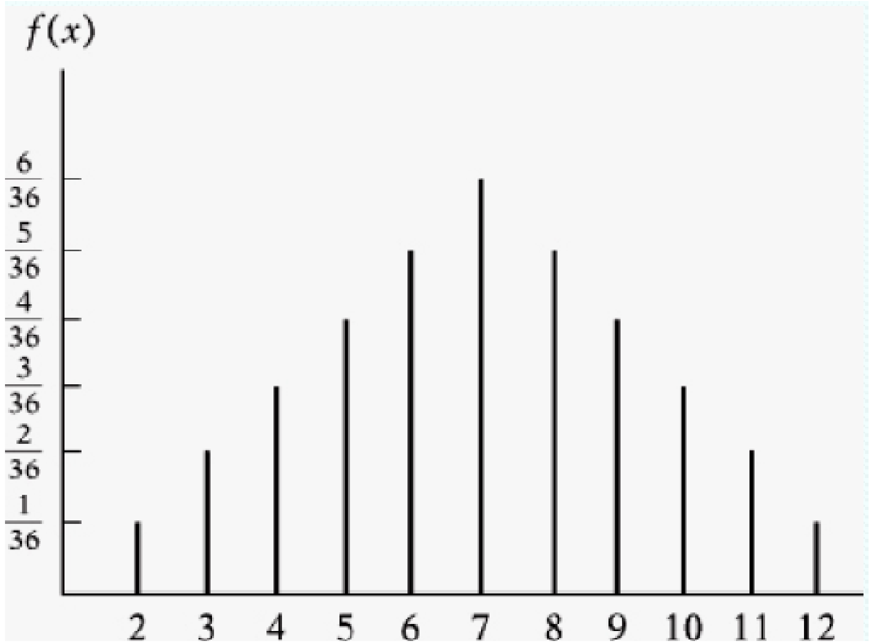
- $\text{Prob}(a \leq x \leq b) = \int_a^b f(x)dx \geq 0$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$



Probability Distribution Functions: Discrete RVs

Example. In a throw of two dice, the random variable X can take on the values

x	2	3	4	5	6	7	8	9	10	11	12
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

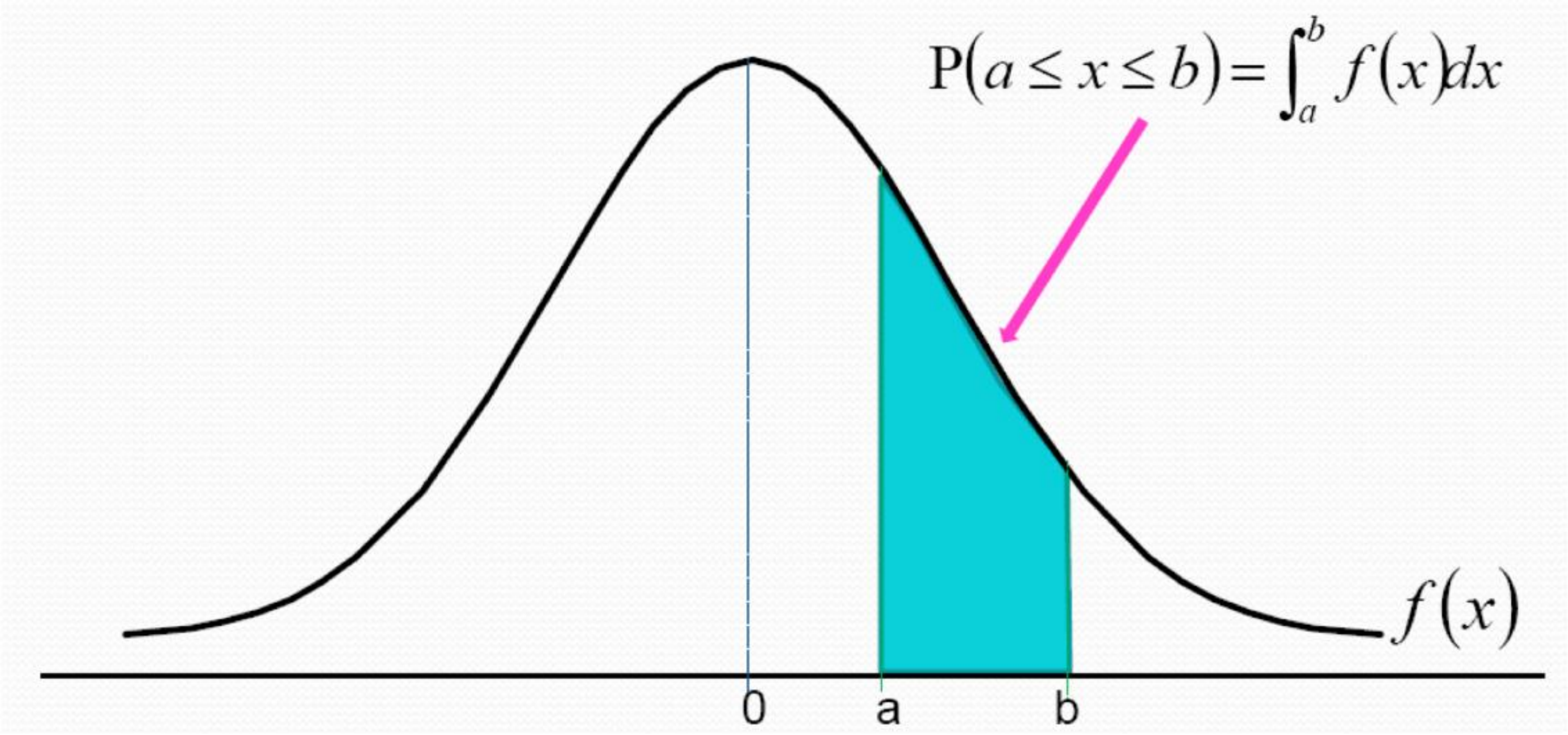


Probability Density Functions: Continuous RVs

The probability density function, *f(x)*, of the random variable is a function with the following properties:

1. *f(x)*> 0 for all *x*.
2. The area under *f(x)* over all values of *X* within its range is equal to 1.
3. Let *a* < *b* be two positive values of random variable *X*. Then the probability that *X* lies between *a* and *b* is the area under the probability density function between these two points.
4. The cumulative distribution function, $F(x_0)$, is the area under the probability density function, *f(x)*, up to x_0.

Probability Density Functions: Continuous RVs



Cumulative Density Function (PDF)

The **cumulative density function (CDF)** describes the probability that the random variable is less than or equal to a particular value. The notation used is $F(x)$.

In both the continuous and discrete cases, the CDF of a random variable must satisfy the following properties:

- $0 \leq F(x_i) \leq 1, \quad \forall x_i$
- $if\ x_1 < x_2, then\ F(x_1) < F(x_2).$ CDF is a non-decreasing function of x
- $\lim_{x \to +\infty} F(x) = 1$
- $\lim_{x \to -\infty} F(x) = 0$

Cumulative Density Function (PDF)

Two important properties of CDFs that are useful for computing probabilities are the following:

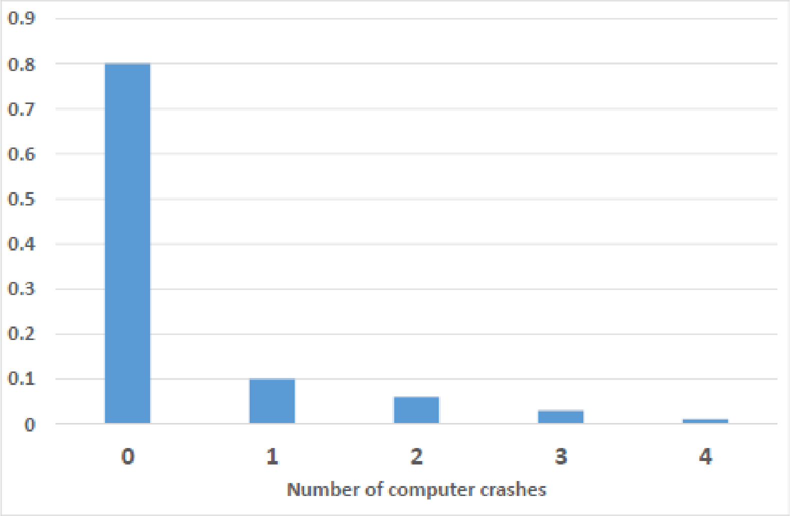
- For any number c, $\mathrm{Prob}(X > c) = 1 - F(c)$
- For any numbers a < b, $\mathrm{Prob}(a < X \leq b) = F(b) - F(a)$

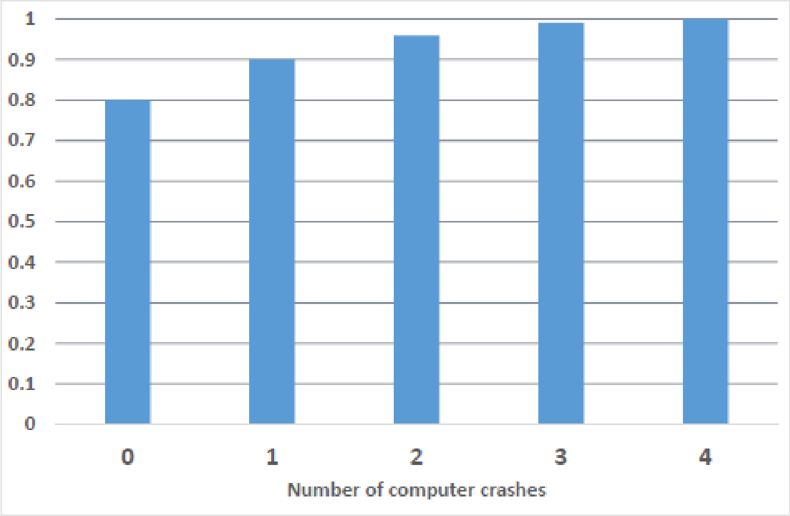
Example: Discrete Random Variable

Example: computer crashing N times in a day

	outcome				
	0	1	2	3	4
pdf	0.8	0.1	0.06	0.03	0.01
cdf	0.8	0.9	0.96	0.99	1

Figure: PDF (left) and CDF (right)

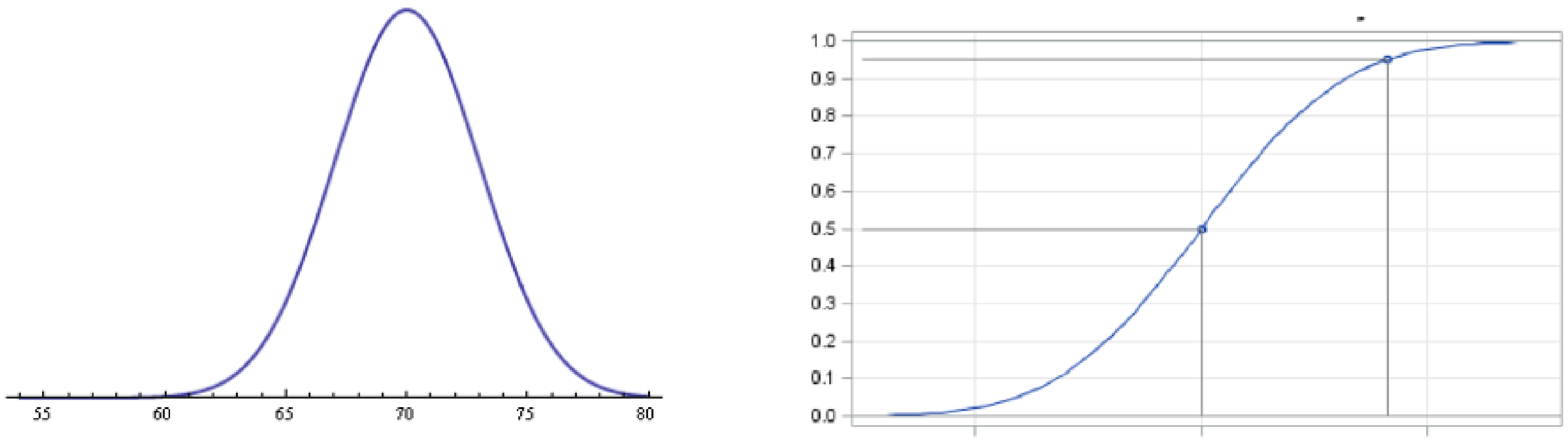




Example: Continuous Random Variable

Example:

Figure: PDF (left) and CDF (right)



Joint Distributions

Let X and Y be two random variables. The **Joint Probability Density Function** expresses the probability that simultaneously X takes the value x and Y takes the value y, as a function of x and y. That is:

$$P_{X,Y}(x, y) = P(X = x \cap Y = y) \text{ for all pairs of } (x, y)$$

Marginal Distributions

A **marginal probability density** or marginal probability distribution is defined with respect to an individual variable. To obtain the marginal distributions from the joint density, it is necessary to sum or integrate out the other variable.

$$P_X(x) = \begin{cases} \sum_y P(x,y) & in\ the\ discrete\ case \\ \int_y P(x,s)dxds & in\ the\ continuous\ case \end{cases}$$

And similar for $f_Y(y)$.

Example of two discrete random variables

Example: There are two random variables: X denotes bond rating and Y denotes bond yield.

The joint probability distribution of X and Y is given in the table on the right.

The marginal distribution of X is given in the last row.

The marginal distribution of Y is given in the last column.

Bond Rating (X)

Bond Yield (Y)	BBB	BB	B	P(y)
8.50%	0.26	0.1	0	0.36
11.50%	0.04	0.28	0.04	0.36
17.50%	0	0.02	0.26	0.28
P(x)	0.3	0.4	0.3	1

Maths & Stats Pre-Sessional

Moments of a distribution

Lecturer: Claudio Vallar
School of Economics and Finance

Expected Value

Moments: summary statistics of the probability distribution

The **mean**, or **expected value**, of a random variable is:

$$\mu_x = E(X) = \begin{cases} \sum xP(x) & if\ X\ is\ discrete \\ \int xf(x)dx & if\ X\ is\ continuous \end{cases}$$

- Measure of central tendency
- Weighted average of the possible values of X with the probability f (x) serving as weights

Variance

The **variance** of a random variable is:

$$\sigma_x^2 = var(X) = E[(X - \mu_x)^2] = \begin{cases} \sum (x - \mu_x)^2 P(x) & if\ X\ is\ discrete \\ \int (x - \mu_x)^2 f(x) dx & if\ X\ is\ continuous \end{cases}$$

- Measure of dispersion
- How the values of X are spread around its mean
- Standard Deviation: $\sigma_x = \sqrt{var(X)}$

Mean and Variance of a Random Variable

Example. In a throw of two dice, the random variable X can take on the values

$\boldsymbol{x}$	2	3	4	5	6	7	8	9	10	11	12
$\boldsymbol{P}(\boldsymbol{x})$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Mean and variance?

$$\mu_X = \sum_x xP(x) = 2 \times \frac{1}{36} + 3\times \frac{2}{36} + \cdots + 11\times \frac{2}{36} + 12\,\frac{1}{36} = 7.$$

$$\sigma_X^2 = \sum_x (x-\mu_X)^2 P(x) = E[X^2] - \mu_X^2 = 2^2 \times \frac{1}{36} + 3^2\times \frac{2}{36} + \cdots + 11^2\times \frac{2}{36} + 12^2\,\frac{1}{36} - 7^2$$

$$= 5.83.$$

Properties

Properties of Expected Value:

- $E(a) = a$
- $E(aX + b) = aE(X) + b$
- $E(aX + bY) = aE(X) + bE(Y)$

Properties of Variance:

- $Var(a) = 0$
- $Var(aX + b) = a^2Var(X)$
- $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$
- $Var(aX - bY) = a^2Var(X) + b^2Var(Y) - 2abCov(X, Y)$
- $Var(aX + bY) = a^2Var(X) + b^2Var(Y)$ if X and Y are independent

Covariance of two Random Variables

The **covariance** of a random variable is:

$$covar(X,Y) = E[(X-\mu_x)(Y-\mu_Y)] = E(XY) - \mu_X\,\mu_Y =$$

$$= \begin{cases} \sum (x-\mu_x)^2 P(x) \quad if\ X\ is\ discrete \\ \int (x-\mu_x)^2 f(x)dx \quad if\ X\ is\ continuous \end{cases}$$

Properties of covariances:

- $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X,Y)$
- $Cov(X,Y) = 0$, if X and Y are independent
- $Covar(a + b, c + dY) = bd\ Cov(X,Y)$

Maths & Stats Pre-Sessional

Distributions of Random Variables

Lecturer: Claudio Vallar
School of Economics and Finance

Probability Distributions

Certain experimental situations naturally give rise to specific probability distributions.

Economists have used a wide variety of other distributions

Continuous Random Variables:

- Normal distribution
- distributions derived from the Normal (Student t, Chi Square and F distribution)

Discrete Random Variables

- Bernoulli distribution
- Poisson distribution

Example of Discrete Probability Distributions: Binomial

- Example:

Imagine running an experiment with a success rate of *p* independently for *n* times. What is the probability of having *x* successes from this series of experiments?

This probability distribution follows a **binomial distribution**, *B(n, p).*

It must satisfy four conditions:

1. Fixed number of trials (").
2. Each trial can result in one of only two possible outcomes: success or failure.
3. The probability of success in a single trial (!) is constant.
4. Trials are independent.

Example of Discrete Probability Distributions: Binomial

The probability density function of a binomial random variable X, is

$$P(x) = \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x}.$$

Note that n and p are given. n is the number of trials and p is the probability of success in each trial. The outcome x is the number of successes in the sample, $x = 0,1,2,\cdots,n$.

Example. Flip a coin 4 times. Let X be the number of heads.

$$P(x) = \frac{4!}{x!\,(4-x)!} 0.5^x (1-0.5)^{4-x} = \frac{1.5}{x!\,(4-x)!}.$$

So $P(0) = 0.0625, P(1) = 0.25, P(2) = 0.375, P(3) = 0.25, P(4) = 0.0625.$

Example of Discrete Probability Distributions: Binomial

- Mean and variance of the binomial distribution:

$$E(X) = np;$$

$$Var(X) = np\,(1-p).$$

- Example.:

If there are 10 trials and the probability of success in each trial is 0.5, what are *E(X)* and *Var(X)*?

$$E(X) = 10 * 0.5 = 5;$$

$$Var(X) = 10 * 0.5 * (1 – 0.5) = 2.5.$$

Continuous Probability Distributions: Normal

- A continuous random variable X is said to be **normally distributed** if its probability density function has the form:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{(x-\mu)^2}{\sigma^2}\right)}$$

Where e and π are mathematical constants

μ is the mean

σ is the standard deviation

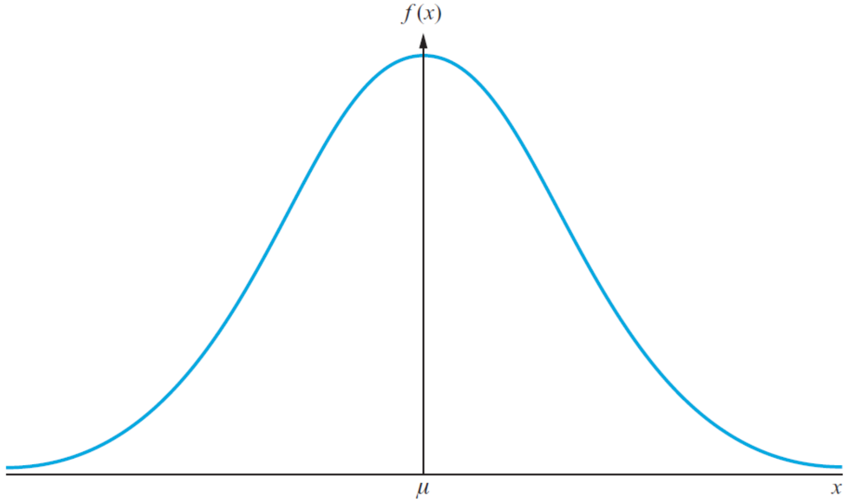
x is any value of the continuous variable

- The normal distribution is determined by its first two moments (**mean μ and variance σ^2**)

Normal Distribution

Given the mean μ and variance σ^2 the normal distribution is defined using the notation

$\boldsymbol{x \sim N(\mu, \sigma^2)}$

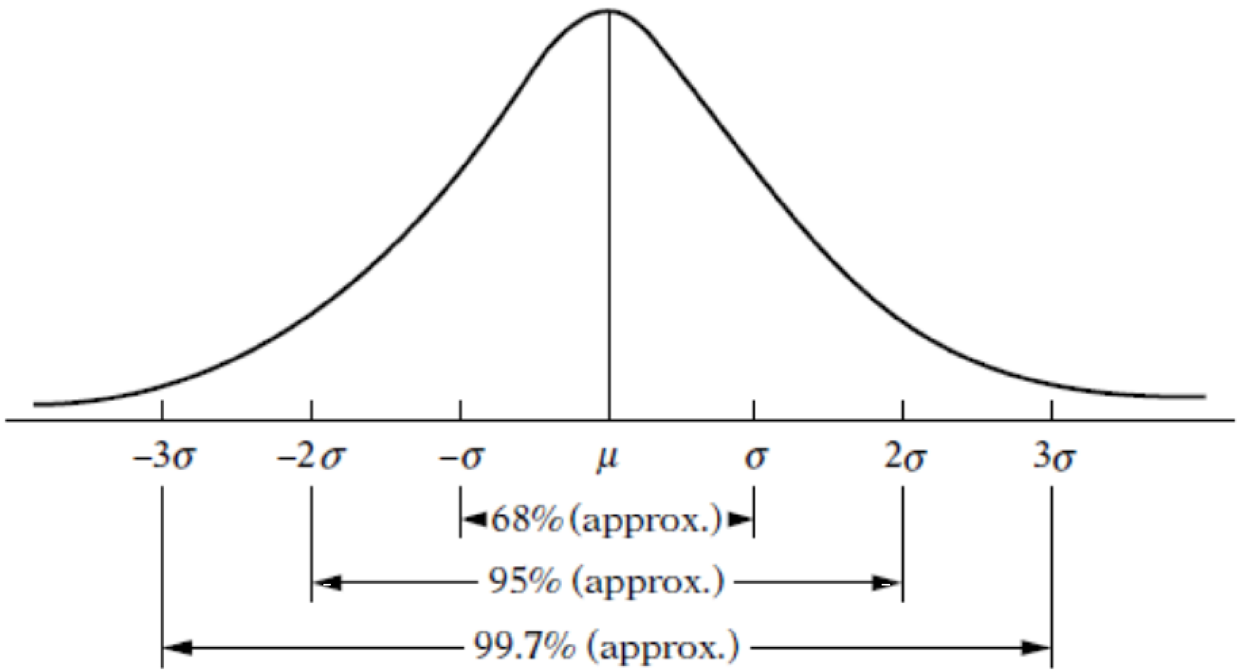


This gives the symmetric bell-shaped curve that is centred on the expectation and has a dispersion that is determined by the standard deviation.

The Normal distribution is symmetric (i.e. mean=median=mode), its skewness = 0 and its kurtosis = 3

Normal Distribution

The empirical rule, or three-sigma rule, states that about 68% of the area under the normal curve lies between $\mu \pm \sigma$, 95% of the area between $\mu \pm 2\sigma$, and about 99.7% of the area between $\mu \pm 3\sigma$



Standard Normal Distribution

The **standard normal distribution** is a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$:

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

- A value from any normal distribution can be transformed into its corresponding value on a standard normal distribution using the following formula:

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

  Standard normal data values are called z-scores.

- A table of standardized normal values (Z table) can then be used to find the probability for a specific z-score: $Prob(Z < z)$
- The z-score allows to compare two scores that are from different normal distributions.

Chi-Square distribution

Let $Z_1, Z_2, \ldots., Z_n$ be independent random variables, each distributed as standard normal, then:

$$W = \sum_{i=1}^{m} Z_i^2 = \chi_m^2$$

Where m is the degrees of freedom.

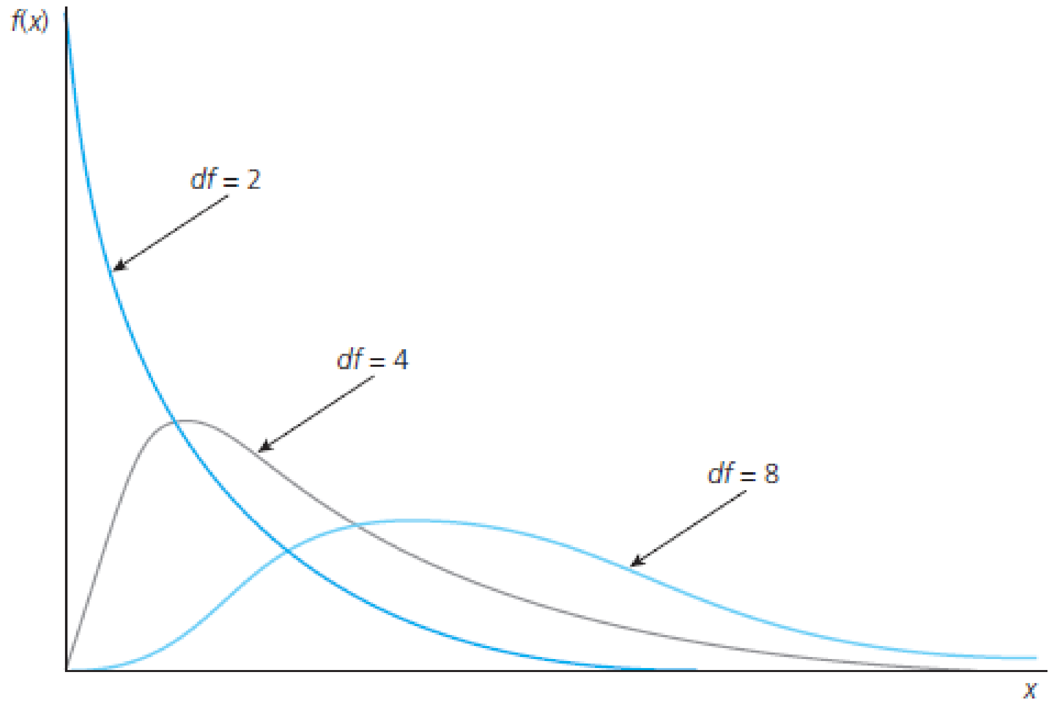
W has a **chi-square distribution with *m* degrees of freedom**.

<u>Properties</u>:

- The mean of a χ_m^2 variable is m and its variance is $2m$.
- If W_{1,m_1} and W_{2,m_2} are two independent chi-square distributions, then $W_{1,m_1} + W_{2,m_2}$ is also a chi-square distribution with df = $m_1 + m_2$

Chi-Square distribution

The pdf for chi-square distributions with varying degrees of freedom is as below:



χ_m^2 is a skewed distribution to the right but as df increases, it becomes symmetric

Student's t distribution

Let $Z_1 \sim N(0,1)$ and $W \sim \chi^2_m$ where Z_1 and W are independently distributed. Then:

$$t_m = \frac{Z_1}{\sqrt{W/m}} = \frac{Z_1\sqrt{m}}{\sqrt{W}}$$

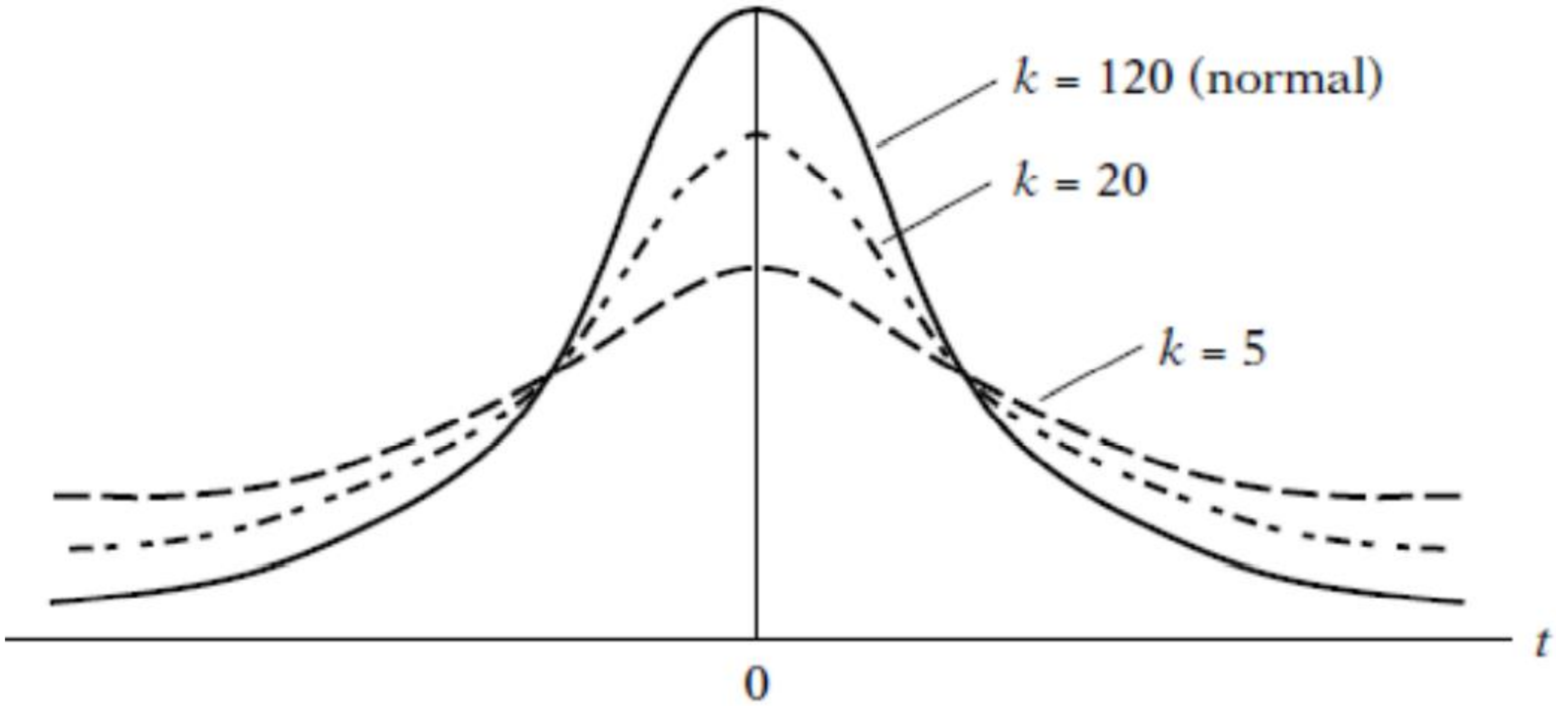
follows a **Student's t distribution** with m degrees of freedom.

<u>Properties</u>:

- It has the same shape as the normal distribution but it is flatter
- It is symmetric
- It has zero mean and variance $m/(m-2)$

Student's t distribution

As the number of degrees of freedom increases, t_m approximates a normal distribution



F distribution

Let W_1 and W_2 be independent chi-squared variables with m_1 and m_2 degrees of freedom, respectively, then:

$$\frac{W_1/m_1}{W_2/m_2} \sim F_{m_1,m_2,}$$

This is knows as **F distribution** with m_1 and m_2 degrees of freedom.