# Fundamental Tools - Probability Theory I

MSc Financial Mathematics

The University of Warwick

September 28, 2015

# Administration

Contact info:

- Tutor: Alex Tse
- Email: S.L.Tse@warwick.ac.uk
- URL: http://www2.warwick.ac.uk/fac/sci/statistics/staff/research_students/tse

Assessment test:

- Friday 1:30 - 4:30pm
- 4 compulsory questions: 1 on linear algebra, 1 on calculus/differential equations and 2 on probability theory

# Modelling a random experiment

A random experiment can be characterised by the following 3 features:

1. What are the possible outcomes of the experiment?
2. What events can we observe? Or, what information will be revealed to us at the end of the experiment?
3. How do we assign probabilities to the events that we can observe?

# Modelling a random experiment: an example

Imagine I roll a fair die privately, and tell you if the outcome is odd or even:

1. The possible outcomes are integers from 1 to 6.
2. The information available to you is whether the roll is odd or even.
3. Probabilities are computed on basis that each outcome is equally likely, so we have 0.5 chance of obtaining odd/even.

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is essentially a collection of 3 mathematical objects representing these 3 features of a random experiment.

# Sample space

A sample space $\Omega$ is a set containing all possible outcomes of a random experiment.

- Rolling a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Flipping two coins: $\Omega = \{HH, HT, TH, TT\}$.
- Type in "=rand()" on an excel spreadsheet: $\Omega = [0, 1]$.
- Stock price path from today to time $T$:
  $\Omega =$ "a set of non-negative continuous functions on $[0, T]$".

An outcome $\omega$ is an element in $\Omega$ (i.e. $\omega \in \Omega$) to be realised at the end of the experiment, which we may or may not observe.

# Events on a sample space

An event $A$ can be represented by a subset of $\Omega$. After the realisation of a random experiment, we say "A happens" if $\omega \in A$.

- Getting an odd roll: $A = \{1, 3, 5\}$.
- Getting the same outcome in 2 coin flips: $A = \{HH, TT\}$.
- "rand()" gives a number larger than 0.5: $A = (0.5, 1]$.
- Stock price is above 2000 at time $T$: $A =$ "$S_T > 2000$".

# $\sigma$-algebra

Informally, a $\sigma$-algebra $\mathcal{F}$:

- represents the information that will be revealed to us after realisation of the random outcome;
- contains all the events that we can verify if they have happened or not after after $\omega$ is realised.

## Definition ($\sigma$-algebra)

*For $\mathcal{F}$ being a collection of subsets of $\Omega$ (i.e. events on $\Omega$), it is a $\sigma$-algebra if it satisfies the below properties:*

1. $\Omega \in \mathcal{F}$;
2. *if* $A \in \mathcal{F}$, *then* $A^C \in \mathcal{F}$;
3. *if* $A_i \in \mathcal{F}$ *for* $i = 1, 2, \ldots$, *then* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

# Motivations behind the defining properties of $\mathcal{F}$

The 3 properties of $\mathcal{F}$ are in place to ensure internal consistency of "information".

I draw a card from a poker deck of 52 cards, and only tell you the suit but not the number.

- If you can verify the event "the card drawn is a spade", you must also be able to verify the event "the card drawn is NOT a spade".
- If you can verify the event "the card drawn is a spade" and "the card drawn is a heart", you must also be able to verify the event "the card drawn is either a spade or heart".

In addition, any sensible information structure should be able to handle trivial questions like whether "the coin flip gives either a head or tail".

# Examples

1. Roll a die $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{1, 2, 3\}, \{4, 5, 6\}\}$$
$$\mathcal{F}_2 = 2^{\Omega} = \text{the set of all subsets of } \Omega \text{ (power set)}$$

   are both $\sigma$-algebras. $\mathcal{F}_1$ contains information on whether the roll is odd or not, and $\mathcal{F}_2$ contains information on the exact outcome.

2. Flip a coin twice $\Omega = \{HH, HT, TH, TT\}$.

$$\mathcal{F}_1 = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$$

   is a $\sigma$-algebra containing information on the outcome of the first flip. But

$$\mathcal{F}_2 = \{\emptyset, \Omega, \{HH, TT\}\}$$

   is not a $\sigma$-algebra.

# Generated $\sigma$-algebra

In an experiment of rolling a die, suppose we are interested in knowing whether the outcome belongs to a low-range (1-2), mid-range (3-4) or high-range (5-6). What is the minimal information required?

- The events of interested are $\{1, 2\}$, $\{3, 4\}$ and $\{5, 6\}$.
- The information of the exact outcome of the roll (represented by the power set $2^\Omega$) is sufficient, but it is an overkill.
- What we need is the smallest $\sigma$-algebra containing the three events above.

**Definition ($\sigma$-algebra generated by a collection of events)**

*Let $\mathcal{C}$ be a collection of subsets (i.e events) of $\Omega$. Then $\sigma(\mathcal{C})$, the $\sigma$-algebra generated by $\mathcal{C}$, is the smallest $\sigma$-algebra on $\Omega$ which contains $\mathcal{C}$. Alternatively, it is the intersection of all $\sigma$-algebras containing $\mathcal{C}$.*

# Generated $\sigma$-algebra: examples

In this example, the required minimal information is given by the $\sigma$-algebra generated by $\mathcal{C} = \{\{1,2\},\{3,4\},\{5,6\}\}$, then

$$\sigma(C) = \{\emptyset, \Omega, \{1,2\},\{3,4\},\{5,6\},\{3,4,5,6\},\{1,2,5,6\},\{1,2,3,4\}\}.$$

If we are interested in the exact outcome of the die, then take $\mathcal{C} = \{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}\}$, and $\sigma(\mathcal{C})$ will be the power set $2^{\Omega}$.

Except in few simple examples, it is hard to write down explicitly a generated $\sigma$-algebra. An important example of such is a Borel $\sigma$-algebra . Take $\Omega = \mathbb{R}$, it is defined as

$$\mathcal{B}(\mathbb{R}) = \sigma(\text{“collections of all open intervals in } \mathbb{R}\text{”}).$$

Conceptually it is similar to a power set generated by an infinite $\Omega$. Almost every subset of $\mathbb{R}$ that we can write down belongs to $\mathcal{B}(\mathbb{R})$.

# Probability measure

## Definition (Probability measure)

*A probability measure* $\mathbb{P}$ *defined on a* $\sigma$*-algebra* $\mathcal{F}$ *is a mapping* $\mathcal{F} \to [0, 1]$ *satisfying:*

1. $\mathbb{P}(\Omega) = 1$;
2. *For a sequence of* $A_i \in \mathcal{F}$ *where* $A_i \cap A_j = \emptyset$ *for any* $i \neq j$*, then* $\mathbb{P}(\cup_i A_i) = \sum_i \mathbb{P}(A_i)$.

From the definition, it is not hard to derive the following properties which you are likely to be familiar with already (see problem sheet):

- $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$;
- If $A \subseteq B$, then $\mathbb{P}(A) \leqslant \mathbb{P}(B)$;
- For any $A$ and $B$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- If $B_i$'s are disjoint and $\cup_i B_i = \Omega$, $\mathbb{P}(A) = \sum_i \mathbb{P}(A \cap B_i)$.

# Probability measure: examples

Typically, we assume the outcome can be directly observed at the end of the experiment and thus $\mathcal{F}$ is chosen to be the largest possible $\sigma$-algebra (i.e power set or Borel $\sigma$-algebra), and we define $\mathbb{P}$ on it. Precise definition of $\mathbb{P}$ depends on the application:

- For a countable sample space $\Omega$ where each outcome is equally likely, define $\mathbb{P}$ on $\mathcal{F} = 2^{\Omega}$ via $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$ for any $A \in \mathcal{F}$.
- To model the number of coin flip required to obtain the first head ($\Omega = \{1, 2, 3, ...\}$), define $\mathbb{P}$ on $\mathcal{F} = 2^{\Omega}$ where $\mathbb{P}$ satisfies $\mathbb{P}(\{\omega : \omega = k\}) = (1-p)^{k-1}p$. Here $p \in (0, 1)$ represents the chance of getting a head in a single flip.
- To represent a uniform random number draw from $\Omega = [0, 1]$, define $\mathbb{P}$ on $\mathcal{F} = \mathcal{B}([0, 1])$ where $\mathbb{P}$ satisfies $\mathbb{P}([a, b]) = b - a$ for $0 \leqslant a < b \leqslant 1$. Such $\mathbb{P}$ defined is called a Lebesgue measure (on $[0, 1]$).

# Independence

## Definition (Independence)

1. *Two events $A$ and $B$ are said to be independent if* $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
2. *A sequence of events $(A_i)_{i=1,2,3...}$ is said to be pairwise independent if $A_i$ and $A_j$ are independent for any $i \neq j$.*
3. *A sequence of events $A_1, A_2, ..., A_n$ is said to be independent if* $\mathbb{P}(\cap_{i=1}^n A_i) = \prod_{i=1}^n \mathbb{P}(A_i)$.

Warning: pairwise independent events are not necessarily jointly independent!

*Exercise:* Two dice are rolled. Let $A$ be the event "the sum is 7", $B$ be the event "the first die gives 3" and $C$ be the event "the second die gives 4". Are the three events pairwise independent? Are they (jointly) independent?

# Conditional probability

**Definition (Conditional probability)**

*Suppose B has positive probability of occurring, the conditional probability of A given that B has occurred is defined as*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

In case of $A$ and $B$ being independent, we have $\mathbb{P}(A|B) = \mathbb{P}(A)$. Here the knowledge of occurrence of $B$ does not change the assessment on likelihood of $A$.

Be familiar with some basic calculations involving conditional probabilities. See problem sheet.

# Principle of counting

In case where the number of outcome is finite, and each outcome has equal probability of occurrence, we determine probability via $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$. The problem reduces to finding the size of the set $A$ and $\Omega$ by counting.

*Multiplication rule:*
If there are $m$ experiments performed, and the number of outcome of the $k$-th experiment is always $n_k$ regardless of the outcomes of all other experiment, then the total number of outcomes is $n_1 \times n_2 \times \cdots \times n_m$.

The first application: For a finite set $\Omega$ of size $n$, its power set has size of $2^n$.

# $k$-permutations of $n$

We have $n$ distinct objects. $k$ of them are selected and placed along a line. What is $P_k^n$, the total number of distinguishable orderings?

Imagine each selection is an independent experiment. There are $n$ choices in filling the first spot, $n-1$ choices in filling the second spot,...,$n-k+1$ choices in filling the $k$-th spot. The number of orderings is thus

$$P_k^n = n \times (n-1) \times \cdots (n-k+1) = \frac{n!}{(n-k)!}.$$

In the special case of $k = n$, the above becomes $n!$. It is the number of permutations by shuffling $n$ objects in a line.

# $k$-combinations of $n$

We have $n$ distinct objects and $k$ of them are selected. What is $C_k^n$, the total number of possible groupings?

Consider a two-stage experiment:

1. We select $k$ objects from the $n$ objects.
2. We then place the $k$ selected objects along a line with shuffling.

This two-stage experiment is equivalent to the one in previous slide which has $P_k^n$ possible outcomes. Meanwhile:

- The number of outcomes in the first stage is $C_k^n$.
- The number of outcomes in the second stage is $k!$.
- By multiplication rule, $P_k^n = C_k^n k!$, thus

$$C_k^n = \frac{P_k^n}{k!} = \frac{n!}{(n-k)!k!}.$$

# Combinatorics: quick examples

- We need to form a team of 2 boys and 3 girls from a class with 13 boys and 17 girls. How many combinations are there? (ans: $C_2^{13} \times C_3^{17}$)
- Draw $n$ balls without replacement from an urn with $M$ red balls and $N$ black balls. What is the chance of getting $r$ red balls (and in turn $n - r$ black balls)? (ans: $\frac{C_r^M C_{n-r}^N}{C_n^{M+N}}$)
- You and the other 2 friends of yours are in a randomly shuffled queue of $n$ people. What many orderings are there such that three of you are standing next to each other? (ans: $3! \times (n - 2)!$)

See problem sheet as well for more exercises on this topic.