

Probability and Statistics for Economics Cheat Sheet

Author: Sai Zhang (email me or check my [Github](#) page)

Random experiments

The outcome in a random experiment is **unpredictable**:

- outcome is too complicated or poorly understood
- outcome is designed to be unpredictable
- coincidences, or independent chains of events

Two examples

Random card shuffle experiment: take top card from a deck and insert randomly, to complete the shuffle of n cards, we need

$$T = n + \frac{n}{2} + \dots + \frac{n}{n-1} + 1 = n \log n$$

shuffles.

Random number generator:

$$x_{n+1} = \frac{ax_n + b}{c} - \left\lfloor \frac{ax_n + b}{c} \right\rfloor$$

the remainder after dividing by c , hence $x_{n+1} \in [0, c-1]$, let $u_{n+1} = \frac{x_{n+1}}{c}$, x_0, a, b, c all be integers. For very large a and good choice of b, c , the sequence u_1, u_2, \dots is like a sequence of numbers randomly picked from $[0, 1]$

Probabilities

Probability is a number in $[0, 1]$ that measures the likelihood of an outcome or a set of outcomes.

Ways of assigning probabilities:

- **symmetry:** assume all outcomes are equally likely
- **experimental method:** relative frequency in repeated random experiment
- **subjective method:** assign probabilities using knowledge of random experiment
- **market method**

Elements of probability space

- **outcome space** Ω and outcomes $\omega \in \Omega$
- **event** $E, E \subseteq \Omega$
- **probability function/measure** $P: \mathcal{A} \rightarrow [0, 1]$: a function from a **collection** \mathcal{A} of subsets of Ω to the interval $[0, 1]$.

Classes of events

Events E_1, E_2, \dots are just sets. They also follow the algebras of sets.

Some set algebras

$$\begin{aligned} A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \\ A \cup A^C &= \Omega, A \setminus B = A \cap B^C \\ (\bigcup_{i=1}^{\infty} A_i)^C &= \bigcap_{i=1}^{\infty} A_i^C \\ (\bigcap_{i=1}^{\infty} A_i)^C &= \bigcup_{i=1}^{\infty} A_i^C \\ A \cup B &= \Omega, A \cap B = \emptyset \Leftrightarrow B = A^C \\ (A^C)^C &= A \\ A \subseteq B \Leftrightarrow A \cap B = A \Leftrightarrow A \cup B = B \Leftrightarrow A \setminus B = \emptyset \Leftrightarrow B^C \subseteq A^C \end{aligned}$$

Two special relations:

- **disjoint:** $E_1 \cap E_2 = \emptyset$
- **partition:** $\bigcup_{i=1}^{\infty} E_i = \Omega$, $\{E_i\}$ are pairwise disjoint

σ -field and Borel σ -field

Definition of σ -field

\mathcal{A} (a collection of subsets of Ω) is a σ -field if:

- 1 $\emptyset \in \mathcal{A}$
- 2 $E \in \mathcal{A} \Rightarrow E^C \in \mathcal{A}$
- 3 $E_1, E_2, \dots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{A}$

It is easy to see that: $\Omega \in \mathcal{A}$ (by 1), $\bigcap E_i^C \in \mathcal{A}$, $\bigcup E_i^C \in \mathcal{A}$, $\bigcap E_i \in \mathcal{A}$ as well

Two important σ -field:

- Trivial σ -field: $\mathcal{A} = \{\emptyset, \Omega\}$
- Largest σ -field: **powerset** of Ω , $\mathcal{P}(\Omega)$

Generating classes

When the powerset has too many events to be assigned probabilities to each event, we start from a set of events \mathcal{E} that we want to assign probabilities to. This \mathcal{E} is a **generating class**.

σ -field generated by \mathcal{E}

The **smallest** σ -field that contains \mathcal{E} is

$$\sigma(\mathcal{E}) = \{E \subseteq \Omega \mid E \in \mathcal{A}, \forall \mathcal{A} \supseteq \mathcal{E}\}$$

where \mathcal{A} can be any σ -field (including the powerset). Another way to write this definition is: for all the σ -fields that contain \mathcal{E} , $\{\mathcal{A}_i \mid \mathcal{E} \subseteq \mathcal{A}_i\}$, we have:

$$\sigma(\mathcal{E}) = \bigcap \mathcal{A}_i$$

This definition is based on the fact: **any intersection of σ -fields is a σ -field**, here is a proof:

Proof: \cap of two σ -fields is a σ -field

We can check the three definitions of σ -field: for $\mathcal{F}_1, \mathcal{F}_2$

- 1 $\emptyset \in \mathcal{F}_1, \emptyset \in \mathcal{F}_2 \Rightarrow \emptyset \in \mathcal{F}_1 \cap \mathcal{F}_2$
- 2 $E \in \mathcal{F}_1 \cap \mathcal{F}_2 \Rightarrow E^C \in \mathcal{F}_1, E^C \in \mathcal{F}_2 \Rightarrow E^C \in \mathcal{F}_1 \cap \mathcal{F}_2$
- 3 $E_1, \dots \in \mathcal{F}_1 \cap \mathcal{F}_2 \Rightarrow E_1, \dots \in \mathcal{F}_1, E_1, \dots \in \mathcal{F}_2 \Rightarrow \bigcup E_i \in \mathcal{F}_1, \bigcup E_i \in \mathcal{F}_2 \Rightarrow \bigcup E_i \in \mathcal{F}_1 \cap \mathcal{F}_2$

Borel σ -field

Borel σ -field \mathcal{B} is a σ -field on \mathbb{R} , its generating class is the set of all **open subsets** of \mathbb{R} . But \mathcal{B} 's generating class is not unique, a particularly important one is $\mathcal{E} = \{(-\infty, x] \mid x \in \mathbb{R}\}$. Here is a proof of this generating class can actually generate \mathcal{B} .

Proof: $\sigma(\mathcal{E}) = \mathcal{B}$

The proof is done in 2 steps:

- prove $\sigma(\mathcal{E}) \subseteq \mathcal{B}$: $\mathcal{E} = (-\infty, x] = \bigcap_{n=1}^{\infty} (-\infty, x + \frac{1}{n}) \in \mathcal{B}$. And by the definition of σ -field generated by \mathcal{E} , we know that $\sigma(\mathcal{E}) = \bigcap \mathcal{A}_i \subseteq \mathcal{E}$, hence $\sigma(\mathcal{E}) \subseteq \mathcal{B}$
- prove $\mathcal{B} \subseteq \sigma(\mathcal{E})$: each open set B in \mathbb{R} can be written as $B = \bigcup_{i=1}^{\infty} (a_i, b_i)$. For each (a, b) , we can rewrite it as $(a, b) = (-\infty, b) \cap (-\infty, a]^C$, where $(-\infty, b) = \bigcup_{n=1}^{\infty} (-\infty, b - \frac{1}{n}]$. By the definition of σ -field, $\bigcup_{n=1}^{\infty} (-\infty, b - \frac{1}{n}] \cap (-\infty, a]^C \in \sigma(\mathcal{E})$, hence $B \in \sigma(\mathcal{E})$.

This is actually a special case of **generating class arguments**, which is used to show that all sets in a σ -field \mathcal{A} have a certain property. It follows:

- for subsets of Ω that have the property, define the collection of them as \mathcal{E}
- show that $\mathcal{A} \subseteq \sigma(\mathcal{E})$
- show that for $\mathcal{A}_0 = \{A_{\text{property}} \in \mathcal{A}\}$ with $\mathcal{E} \subseteq \mathcal{A}_0$, \mathcal{A}_0 is a sigma field (Since $\sigma(\mathcal{E})$ is the smallest σ -field that contains \mathcal{E} , $\sigma(\mathcal{E}) \subseteq \mathcal{A}_0$)
- $\mathcal{A} \subseteq \sigma(\mathcal{E}) \subseteq \mathcal{A}_0 \subseteq \mathcal{A} \Rightarrow \mathcal{A} = \mathcal{A}_0$

Often it is difficult to show that \mathcal{A}_0 is a σ -field, hence we introduce λ -systems.

λ -systems

\mathcal{A}_0 is a λ -system if

- 1 $\Omega \in \mathcal{A}_0$
- 2 If $D_1, D_2 \in \mathcal{A}_0$ and $D_2 \subseteq D_1$, then

$$D_1 \setminus D_2 = D_1 \cap D_2^C \in \mathcal{A}_0$$

- 3 If D_n is an increasing sequence of sets in \mathcal{A}_0 , then $\bigcup_{i=1}^{\infty} D_i \in \mathcal{A}_0$

Two theorems link λ -systems and σ -fields:

Theorems of λ -systems/ σ -fields

- **Theorem 1:** If \mathcal{E} is closed under finite intersections, and if \mathcal{A}_0 is a λ -system with $\mathcal{E} \subseteq \mathcal{A}_0$, then $\sigma(\mathcal{E}) \subseteq \mathcal{A}_0$.
- **Theorem 2:** a λ -system \mathcal{A}_0 is a σ -field $\Leftrightarrow \mathcal{A}_0$ is closed in finite intersections.

Choice of σ -field

The choice of σ -field is usually determined by the nature of the outcome space:

- discrete (countable) outcome space: **powerset** of Ω shall be chosen.
- continuous outcome space (real line): **Borel σ -field** shall be chosen.

Probability measure

Definition of probability measure

A **probability measure** is a function $P: \mathcal{A} \rightarrow \mathbb{R}$ with a σ -field \mathcal{A} :

- 1 $\forall E \in \mathcal{A}, P(E) \geq 0$
- 2 $P(\Omega) = 1$
- 3 If E_1, E_2, \dots are pairwise disjoint (and countable), $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$

The usual way to specify probabilities using a probability model is to first assign probabilities to some simple collection of events \mathcal{E} , these assignments can be extended to $\sigma(\mathcal{E})$, this extension can usually be shown to be unique, giving a probability measure on $\sigma(\mathcal{E})$.

Discrete Ω : Dice roll

For the problem of rolling a single dice, the probability measure can be defined as

- σ -field \mathcal{A} : powerset of Ω
- probability measure: $P(E) = \frac{\#E}{6}$

This construction can be extended to all discrete outcome spaces: starting by assigning probabilities p_i to single outcomes and define $P(E) = \sum_{i \in E} p_i$. Here is a special case of discrete Ω : random experiments with equally likely outcomes, i.e., $\Omega = \{\omega_i\}_{i=1}^I, P(\omega_i) = \frac{1}{I}, P(E) = \frac{\#E}{I}$.

4 cases of equally likely Ω

Consider the problem of selecting k elements from n distinct elements, the number of selections N_S has 4 cases:

- Ordered without replacement: $N_S = \frac{n!}{(n-k)!}$
- Ordered with replacement: $N_S = n^k$
- Unordered without replacement: $N_S = \frac{n!}{k!(n-k)!} = \binom{n}{k}$
- Unordered with replacement: $S_n = \binom{n+k-1}{k}$

Continuous Ω : Borel σ -field \mathcal{B}

For the probability measure on \mathcal{B} can be defined as

- probabilities on $\mathcal{E} = \{(-\infty, x] \mid x \in \mathbb{R}\}$

$\mathcal{E} = \{(-\infty, x] \mid x \in \mathbb{R}\}$ is closed under finite intersection, hence, we need to show $\forall B \in \mathcal{B}, P(B) = P'(B)$, that is, \mathcal{E} has a unique probability measure.

Consider $\mathcal{B}_0 = \{B \in \mathcal{B} \mid P(B) = P'(B)\}$, which is a λ -system. For an increasing sequence B_n of events in \mathcal{B}_0 , $P(B) = P(B_1 \cup (B_2 \setminus B_1) \cup \dots \cup (B_n \setminus B_{n-1})) = P(B_1) + (P(B_2) - P(B_1)) + \dots + (P(B_n) - P(B_{n-1})) = \lim P(B_n)$ then $P(B) = \lim P(B_n) = \lim P'(B_n) = P'(B)$. Now $\mathcal{E} \subseteq \mathcal{B}_0$ is closed under finite intersections, then by Thm.1 of λ -system, $\sigma(\mathcal{E}) \subseteq \mathcal{B}_0 \subseteq \mathcal{B}$, then $\mathcal{B} = \sigma(\mathcal{E})$, hence $\forall B \in \mathcal{B}, P(B) = P'(B)$.

Properties of probability measures

- $P(E^C) = 1 - P(E), P(\emptyset) = 0, P(E) \leq 1$
- $P(E_2 \setminus E_1) = P(E_2 \cap E_1^C) = P(E_2) - P(E_2 \cap E_1)$
- $P(E_1 \cup E_2) = P(E_1 \cup (E_2 \setminus E_1)) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
- $E_1 \subseteq E_2 \Rightarrow E_2 = E_1 \cup (E_2 \setminus E_1) \Rightarrow P(E_1) \leq P(E_2)$
- $P(E_1 \cup E_2) \leq P(E_1) + P(E_2)$
- Bonferroni inequality: $P(E_2 \cap E_1) \geq P(E_1) + P(E_2) - 1$
- Law of total probability: for a partition of $\Omega, \{E_i\}, P(A) = \sum_{i=1}^{\infty} P(A \cap E_i)$
- Boole's inequality: $P(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} P(E_i)$

The proof of all these results relies on the step to transform the object set to a **union of pairwise disjoint sets**.

For all subsets of \mathbb{R} , it is impossible to assign probabilities to them. The proof takes advantage of a series of uncountable,

disjointed subsets centered at irrational numbers.

Random variable

Definition of random variables

For a probability space (Ω, \mathcal{A}, P) , a random variable X is a function $X : \Omega \rightarrow \mathbb{R}$ s.t. $\forall B \in \text{Borel } \sigma\text{-field } \mathcal{B}, E = \{\omega \mid X(\omega) \in B\} \in \mathcal{A}$.

Another way of stating it is: X takes a value in $B \Leftrightarrow \omega \in E$ i.e., event E happens. Thus, $\Pr(X \in B) = \Pr(E)$

2 things to keep in mind

- $E = X^{-1}(B)$ (X^{-1} does NOT necessarily exist)
- X is Borel measurable

Properties of X^{-1}

- $X^{-1}(B^C) = (X^{-1}(B))^C$
- $X^{-1}(\bigcup_{i=1}^\infty B_i) = \bigcup_{i=1}^\infty X^{-1}(B_i)$

Borel Measurability

Borel measurability of a random variable X (function) is established separately by the countability of outcome space Ω .

Countable Ω

For a countable $\Omega, X : \Omega \rightarrow \mathbb{R}$ is Borel measurable if $\sigma\text{-field on } \Omega, \mathcal{A}$, is the powerset of Ω

Uncountable Ω

Let $\mathcal{E} \subseteq \mathcal{B}$ be a generating class of Borel $\sigma\text{-field } \mathcal{B}$, for an uncountable Ω and its $\sigma\text{-field } \mathcal{A}, X$ is Borel measurable if $\forall E \in \mathcal{E}, X^{-1}(E) \in \mathcal{A}$.

We use a **generating class argument** to prove this:

Step 1: Define $\mathcal{C} = \{B \in \mathcal{B} \mid X^{-1}(B) \in \mathcal{A}\}$, since $\forall E \in \mathcal{E}, X^{-1}(E) \in \mathcal{A}$, we know $\mathcal{E} \subseteq \mathcal{C}$.

Step 2: Check \mathcal{C} is a $\sigma\text{-field}$:

- i $\emptyset \in \mathcal{C}$
- ii $B \in \mathcal{C} \Rightarrow X^{-1}(B) \in \mathcal{A} \Rightarrow X^{-1}(B^C) = X^{-1}(B)^C \in \mathcal{A} \Rightarrow B^C \in \mathcal{C}$
- iii $B_1, B_2, \dots \in \mathcal{C} \Rightarrow X^{-1}(B_1), \dots \in \mathcal{A} \Rightarrow X^{-1}(\bigcup_{i=1}^\infty B_i) = \bigcup_{i=1}^\infty X^{-1}(B_i) \in \mathcal{A}$

Step 3: Since \mathcal{B} is the smallest $\sigma\text{-field}$ containing \mathcal{E} (definition of generating class), \mathcal{C} is a $\sigma\text{-field}$ containing $\mathcal{E}, \mathcal{B} \subseteq \mathcal{C}$; by the definition of $\mathcal{C}, \mathcal{C} \subseteq \mathcal{B}$. Hence $\mathcal{C} = \mathcal{B}$, meaning that $\forall B \in \mathcal{B}, X^{-1}(B) \in \mathcal{A}$, and X is Borel measurable.

Borel measurability has some important applications:

Continuous $X : \mathbb{R} \rightarrow \mathbb{R}$

For the probability space $(\mathbb{R}, \mathcal{B}, P)$ and a **continuous** $X : \mathbb{R} \rightarrow \mathbb{R}$ is Borel measurable.

X_{\sup} and X_{\inf} of X_n

For a sequence of random variables, $\{X_n\}_{n=1}^\infty, X_{\sup} = \sup_n X_n$ and $X_{\inf} = \inf_n X_n$ are also Borel measurable.

A brief proof: Take (x, ∞) as a generating class for \mathcal{B}, X_n being random variables $\Rightarrow \{w \mid X_n(w) > x\} \in \mathcal{A} \Rightarrow \bigcup_n \{w \mid X_n(w) > x\} = \{w \mid X_{\sup}(w) > x\} \in \mathcal{A}$. Taking $(-\infty, x]$ as the generating class can proof for X_{\inf} .

$\lim_{n \rightarrow \infty} X_n = X$

Again, for a sequence of random variables, $\{X_n\}_{n=1}^\infty$, if $\lim_{n \rightarrow \infty} X_n = X$ exists, X is a random variable.

A brief proof: we know

$$\liminf_{n \rightarrow \infty} X_n(\omega) = \sup_n \inf_{m \geq n} X_m(\omega)$$

$$\limsup_{n \rightarrow \infty} X_n(\omega) = \inf_n \sup_{m \geq n} X_m(\omega)$$

then it is easy to see $\liminf_{n \rightarrow \infty} X_n \leq X = \lim_{n \rightarrow \infty} X_n \leq \limsup_{n \rightarrow \infty} X_n$. If $\lim_{n \rightarrow \infty} X_n = X$ exists, the three limits are all equal and all Borel measurable.

Algebras of random variables

If X, Y are random variables, $Z = X + Y$ is Borel measurable.

A brief proof: take the generating class $\mathcal{E} = \{(x, \infty)\}$, then for $A = \{\omega \mid Z(\omega) > z\}$, we have $X(\omega) + Y(\omega) > z \Leftrightarrow X(\omega) > z - Y(\omega)$.

Here is the trick: We can always find a rational number r s.t. $X(\omega) > r > z - Y(\omega)$, and $A = \bigcup_r (\{\omega \mid X(\omega) > r\} \cap \{\omega \mid Y(\omega) > z - r\})$. Since rational numbers are **countable**, X and Y are both Borel measurable, A is in \mathcal{A} , hence Z is Borel measurable.

For the collection of all Borel measurable functions $\mathcal{M} = \{X : \Omega \Rightarrow \mathbb{R}\}$ (\mathcal{M}^+ for non-negative functions), We focus on a special class:

Simple functions

$$X(\omega) = \sum_i^n \alpha_i I_{A_i}(\omega)$$

where

- I_A is the indicator function of event A
- $A_i \in \mathcal{A}, i = 1, \dots, n$ are a partition of Ω
- α_i are constants

Why are they important: each non-negative borel measurable function can be approximated by an **increasing sequence** of simple functions!

Theorem of simple functions

For $\forall X \in \mathcal{M}^+$,

$$X_n(\omega) = 2^{-n} \sum_{i=1}^{4^n} I_{X \geq \frac{i}{2^n}}(\omega)$$

satisfies $0 \leq X_1(\omega) \leq \dots \leq X_n(\omega)$ and $X_n(\omega) \rightarrow X(\omega), \forall \omega \in \Omega$.