Ch1. Probability space

1. Probability space
2. Long-run relative frequency
3. Axiom of Probability
4. Conditional Probabilities
5. Independent Events

Experiment

- (Random) Experiment ($\mathcal{E}$): an experiment whose outcome cannot be determined in advance.
  - ω : elementary outcome, simple event; possible outcome from the random experiment.
- Random experiment is the underlying (physical) dynamics generating randomness. In other words, as a result of random experiment, we observe randomness in real-life.
- Sample space (S): set of all possible outcomes of an experiment.
- Event ($A, B, C \ldots$): a subset of a sample space.
  An event A is said to occur iff the observed outcome $\omega \in A$.

Random Experiment: Examples

- Experiment: Observe a sex of a newborn baby in the hospital.

$$S = \{boy, girl\} = \{0, 1\}$$

- Experiment: Tossing two dice

$$S = \{(i, j) | i, j = 1, \ldots, 6\}$$

$$E = \{\text{the sum is 13}\}$$

- Experiment: counting the # of traffic accidents at a given intersection during a specific time interval.

$$S = \{0, 1, 2, \ldots\}$$

$$A = \{\#\ \text{accidents} \ \ \text{are} \ \leq 7\} = \{0, 1, 2, \cdots 7\}$$

Now, we want to assign chances of an event occurring. Why assign probability on an event rather than an elementary outcome? To handle continuous sample space.

Three ways to assign probability

1. Classical (uniform) model

$$P(A) = \frac{\text{\# of elements in } A}{\text{\# of elements in } S}$$

   Not realistic if $|S| = \pm\infty$.

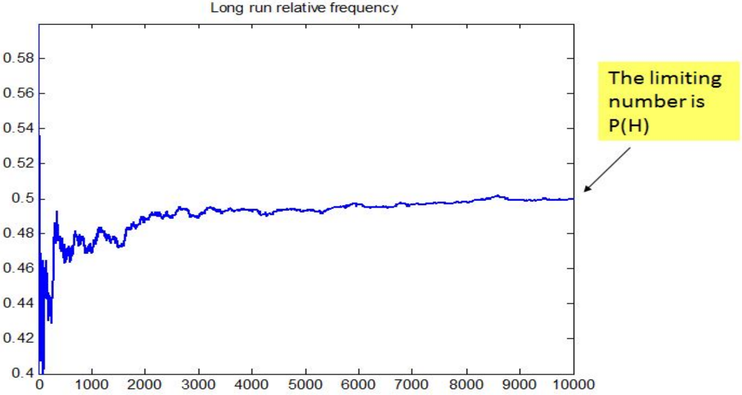
2. Long- run relative frequency
   Repeat random experiment many times under the same condition, and see the proportion of time observing event A.

$$P(A) = \lim_{n \to \infty} \frac{\#A}{n}$$

   However, mathematically rigorous treatment of long-run relative frequency is challenging. For example, existence/uniqueness of limit.

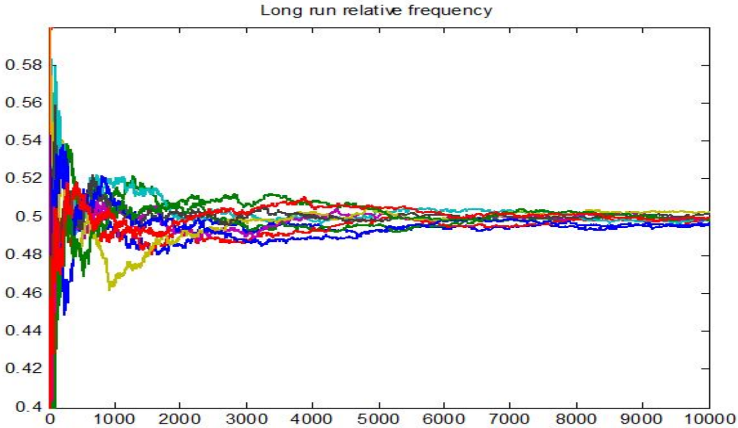
Long-run relative frequency

- Consider coin tossing and assign the probability of observing Head.



Long-run relative frequency

- (uniqueness) Will it converges to the same number on other trials also?



Axiom of Probability

- We can avoid such mathematical challenges by considering axioms.
- Axioms refer to self-evident statements. That is, accepted as true without controversy.
- Put it other way around, axioms are the fundamental (genuine) properties of probability. Observe from long-run relative frequency.

$i)$ probability never negative (non-negative).

$ii)$ $P(S) = 1$

$iii)$ Additive structure :

$$P(\{\omega_1, \omega_2\}) = P(\{\omega_1\}) + P(\{\omega_2\})$$

Axioms of Probability

Definition

A probability measure P (on a σ- field of subsets $\mathcal{F}$ of a set S) is a real-valued set function satisfying.

$i)$ $P(S) = 1$ *(add up to 1)*

$ii)$ $P(A) \geq 0$ *for all* $A \in \mathcal{F}$ *(non-negative)*

$iii)$ *If* $A_n \in \mathcal{F}, n = 1, 2 \ldots$ *are mutually disjoint sets, that is* $A_i \cap A_j = \emptyset$ *if* $i \neq j$, *then*

$$P\left(\bigcup_{i=1}^{\infty} A_n\right) = \sum_{i=1}^{\infty} P(A_n)$$

(countably additive)

A probability is an non-empty countably additive set function add up to 1.

Examples

- Experiment: Toss two coins
- $S = \{HH, HT, TH, TT\}$
- Now assign probability using axioms

- Experiment: Roll a die
- $S = \{1, 2, 3, 4, 5, 6\}$
- Now assign probability using axioms

Properties

1. $P(\emptyset) = 0$

2. $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$ for disjoint A_i's.

Properties

3. Complement law

$$P(A^c) = 1 - P(A)$$

4. $E \subset F$, then $P(E) \leq P(F)$

5. Addition law

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

Properties

5'. Inclusion- Exclusion identity

$$P(E_1 \cup \dots \cup E_n) = \sum_{i=1}^{n} P(E_i) - \sum_{i_1<i_2} P(E_{i1} \cap E_{i2}) + \dots$$

$$+(-1)^{r+1} \sum_{i_1<i_2<\dots<i_r} P(E_{i1} \cap \ldots \cap E_{ir})$$

$$\ldots + \ldots + (-1)^n P(E_1 \cap \dots \cap E_n)$$

Properties

6 Probability is a continuous set function. If $\{E_n\}$ is an increasing/decreasing sequence of events, then

$$\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n)$$

Definition (Limit of events)

Suppose E_n is increasing sequence of events $E_1 \subset E_2 \subset \ldots$, the limit of events is defined as

$$\lim_{n\to\infty} E_n = \bigcup_{i=1}^{\infty} E_i.$$

Similarly, for decreasing sequence of events $E_1 \supset E_2 \supset \ldots$,

$$\lim_{n\to\infty} E_n = \bigcap_{i=1}^{\infty} E_i$$

Properties

Indeed:

Conditional probability

Definition

The conditional probability of event E under the condition that event F happens for sure is defined as

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

whenever $P(F) > 0$.

- Experiment: Toss two dice at the same time.

$$S = \{(1,1), (1,2), \cdots, (6,6)\}$$

$F =$ First die is 4, $\quad E =$ the sum of two dice equals 6.

$$P(E|F) =$$

$$P(E|F) = \frac{n(E \cap F)}{n(F)} = \frac{n(E \cap F)/N}{n(F)/N} \approx \frac{P(E \cap F)}{P(F)}$$

Conditional probability

- Conditional probability is also probability. That is, it satisfies three axioms of probability.
- Computing probability via conditioning

$$P(E \cap F) = P(E|F)P(F)$$

In general

$$P(E_1 \cap E_2 \cap \cdots \cap E_n) = P(E_1)P(E_2|E_1)\cdots P(E_n|E_1E_2\cdots E_{n-1})$$

Computing probability via conditioning

- Example 1.8. Suppose that each of three men at a party throws his hat into the center of the room. The hats are first mixed up and then each man randomly selects a hat. What is the probability that none of the three men selects his own hat?

  Sol)

Independent events

Definition

Two events E and F are independent iff

$$P(E \cap F) = P(E)P(F)$$

- Independence implies that $P(E|F) = P(E)$ and $P(F|E) = P(F)$.
- Knowledge that $F(E)$ has occurred does not affect the probability that $E(F)$ occurs

Definition

Events $E_1, E_2, \cdots E_n$ are said to be mutually independent if for every subset $E_{1'}, E_{2'}, \cdots E_{r'}, r \leq n$

$$P(E_{1'} \cap E_{2'} \cap \cdots \cap E_{r'}) = P(E_{1'})P(E_{2'}) \cdots P(E_{r'})$$

Independent events

- Pairwise independence does not imply mutually independence. Counter example: A ball is drawn uniformly from $S = \{1, 2, 3, 4\}$. Define

  $$E = \{1, 2\} \quad F = \{1, 3\} \quad G = \{1, 4\}.$$

  Then, pairwise independence indicates that

  $$P(E \cap F) = P(E)P(F) = \frac{1}{4}$$

  $$P(E \cap G) = P(E)P(G) = \frac{1}{4}$$

  $$P(G \cap F) = P(G)P(F) = \frac{1}{4}$$

  However,

  $$\frac{1}{4} = P(EFG) \neq P(E)P(F)P(G) = \frac{1}{8}$$

Independent events

- If A and B independent, then so are A and B^c. Furthermore, A^c and B, $A^c$and B^c are all independent.

- Do not confuse independence and disjoint (mutually exclusive). They are two different concepts.

  $$P(E \cap F) = P(E)P(F) \rightarrow \text{Defined through probability}$$

  $$E \cap F = \phi \rightarrow \text{Probability is NOT required}$$

  (but we can still say $P(E \cap F) = 0$)