Introduction to Probability: Probability measure

Rosario Barone

Tor Vergata University of Rome

Statistical tools for decision making

Undergraduate Degree in Global Governance

A.Y. 2023/2024

Quantifying Uncertainty

The probability:

- Measures the chances that an event will occur.
- Forms the basis for inferential statistics.
  - Inference is used to make decisions under uncertainty: probability evaluates the uncertainty involved in those decisions.

Probability measure

Definition

Let Ω be a set $\mathcal{F}$ a collection of subsets of ω. A *probability measure*, or simply a *probability* on $(\Omega, \mathcal{F})$ is a function

$$P : \mathcal{F} \to [0, 1]$$

- To be a probability measure P must satisfy

1. $\forall A \in \mathcal{F}, \quad 0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. if A_1 e A_2 are disjoint then $P(A_1 \cup A_2) = P(A_1) + P(A_2)$

Axioms of Probability

Probability theory follows these three axioms:

1. **First Axiom**: For any event A, $P(A) \geq 0$.

2. **Secon Axiom**: $P(\Omega) = 1$, where Ω is the sample space.

3. **Third Axiom**: For a sequence of mutually exclusive events $A_1, A_2, \ldots$, $P(A_1 \cup A_2 \cup \ldots) = \sum_{i=1}^{\infty} P(A_i)$.

Properties of the probability

- **Probability of the complementary Event**: $P(A) = 1 - P(\bar{A})$
- **Probability of the empty set**: $P(\emptyset) = 0$
- **Interval of the probability of an event**: $0 \leq P(A) \leq 1$
- **Monotonicity**: If $A \subseteq B$ $P(A) \leq P(B)$
- **Union of events**: For general set A and B

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

Calassical Approach

The *probability* of an event A, denoted as $P(A)$, is a real number in the range $0 \leq P(A) \leq 1$. It represents the likelihood of event A occurring.

- Probability of an event in a finite sample space:

$$P(A) = \frac{\text{number of favorable outcomes for } A}{\text{total number of outcomes in } \Omega}$$

Example

Ex: Roll a die once, and find the probability of events A = "even number", B = "number less than 5", C = "number different from 2".

$$P(A) = \frac{3}{6} = 0.5$$

$$P(B) = \frac{4}{6} = 0.67$$

$$P(C) = \frac{5}{6} = 0.83$$

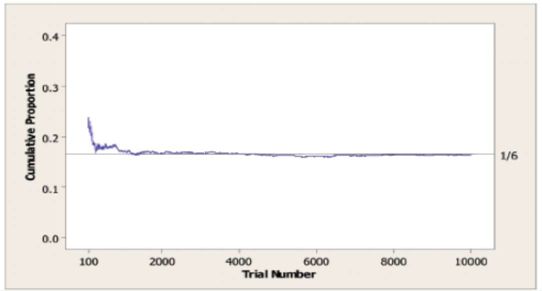
Frequentist Approach

Frequentist: Probability is defined as the proportion of times an event occurs in a long sequence of observations.

However, in the short run, the occurrence of an event may exhibit a high degree of randomness. In the long-run, this proportion becomes very predictable (Law of Large Numbers).

Example

Keeping rolling a die, the share of 6 obtained will progressively converge to 1/6:



Subjective approach

Subjective: based on individual belief, experience, information.

Example

The probability that each one of you attaches to the event "I will get 30 as final mark in Statistics".

Bayesian statistics is a branch of statistics that uses subjective probability as its foundation.

Joint Probability

Definition

Let consider the events $A, B \subset \Omega$. The Joint probability, denoted as $P(A \cap B)$, represents the probability of both event A and event B happening simultaneously.

Example

Let's consider a deck of cards. What is the probability of drawing a red card (event A) and a face card (event B) on two consecutive draws?

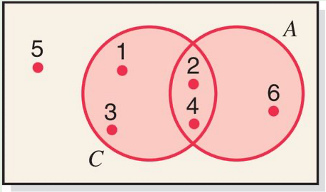
- $P(A) = \frac{26}{52} = \frac{1}{2}$ (there are 26 red cards out of 52).
- $P(B) = \frac{12}{52} = \frac{3}{13}$ (there are 12 face cards out of 52).
- $P(A \cap B) =$ Probability of drawing a red face card $= \frac{6}{52} = \frac{3}{26}$.

Joint Probability

Example

Ex: Roll a die once and consider the following events:

- A = An even number is observed $=\{2, 4, 6\}$
- B = An odd number is observed $= \{1, 3, 5\}$
- C = A number less than 5 is observed $= \{1, 2, 3, 4\}$



The joint probability of A and C is:

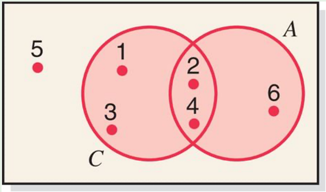
$$P(A \cap C) = \frac{2}{6}$$

Joint Probability

Example

Ex: Roll a die once and consider the following events:

- A = An even number is observed $=\{2, 4, 6\}$
- B = An odd number is observed $= \{1, 3, 5\}$
- C = A number less than 5 is observed $= \{1, 2, 3, 4\}$



The joint probability of A and C is:

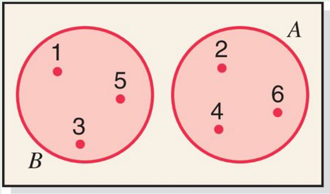
$$P(A \cap C) = \frac{2}{6}$$

Joint Probability: mutually exclusive events

Example

Ex: Roll a die once and consider the following events:

- A = An even number is observed $= \{2, 4, 6\}$
- B = An odd number is observed $= \{1, 3, 5\}$
- C = A number less than 5 is observed $= \{1, 2, 3, 4\}$



The joint probability of A and B is:

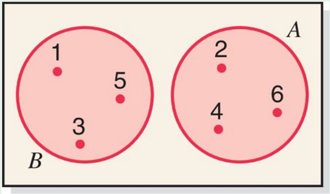
$$P(A \cap B) = 0$$

Joint Probability: mutually exclusive events

Example

Ex: Roll a die once and consider the following events:

- A = An even number is observed $=\{2, 4, 6\}$
- B = An odd number is observed $= \{1, 3, 5\}$
- C = A number less than 5 is observed $= \{1, 2, 3, 4\}$



The joint probability of A and B is:

$$P(A \cap B) = 0$$

Conditional Probability

Definition

Conditional probability is the probability of an event occurring given that another event has already occurred. It is denoted as $P(A|B)$ and read as "the probability of event A given event B."

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

- $P(A|B)$ represents the likelihood of event A occurring when we already know that event B has occurred.
- It provides a way to update probabilities based on new information.

Note: Given the definition of *conditional probability* can rewrite:

$$P(A \cap B) = P(A|B)P(B) \text{ or } P(A \cap B) = P(B|A)P(A).$$

Conditional Probability

Example

Consider a bag of colored marbles. What is the probability of drawing a red marble (A) from the bag given that we have already drawn a blue marble (B)?

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

$$P(A|B) = \frac{2/10}{3/10} = \frac{2}{3}$$

Independent Events

Definition

Two events, $A, B \subset \Omega$, are considered **independent** if the occurrence of one event does not affect the probability of the other event happening:

$$P(A|B) = P(A) \text{ or } P(B|A) = P(B).$$

- Mathematically, two events A and B are independent if $P(A \cap B) = P(A) \cdot P(B)$.
- In the case of independent events, knowing the outcome of one event provides no information about the other.

Independent Events

Example

A box contains a total of 100 DVDs that were manufactured on two machines, A and B. Of the total DVDs, 15 are defective. Of the 60 DVDs that were manufactured on Machine A, 9 are defective.
Let D be the event that a randomly selected DVD is defective, and let A be the event that a randomly selected DVD was manufactured on Machine A. Are D and A independent?

$$P(D) = \frac{15}{100} = 0.15 \quad \text{and} \quad P(D|A) = \frac{9}{60} = 0.15$$

Since $P(D|A) = 0.15 = P(D) = 0.15$, it follows that D and A are independent.

Independent Events

Example

A box contains a total of 100 DVDs that were manufactured on two machines, A and B. Of the total DVDs, 15 are defective. Of the 60 DVDs that were manufactured on Machine A, 9 are defective.
Let D be the event that a randomly selected DVD is defective, and let A be the event that a randomly selected DVD was manufactured on Machine A. Are D and A independent?

$$P(D) = \frac{15}{100} = 0.15 \quad \text{and} \quad P(D|A) = \frac{9}{60} = 0.15$$

Since $P(D|A) = 0.15 = P(D) = 0.15$, it follows that D and A are independent.

Addition Rule for Non-Disjoint Events

Definition

When two events, A and B, are not mutually exclusive, we can find the probability of either event A or event B occurring by using the addition rule for non-disjoint events:

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

where:

- $P(A \cup B)$ is the probability that either event A or event B or both occur.
- $P(A)$ is the probability of event A.
- $P(B)$ is the probability of event B.
- $P(A \cap B)$ is the probability that both events A and B occur simultaneously.

Addition Rule for Non-Disjoint Events

Example

Consider drawing a card from a standard deck of 52 playing cards. What's the probability of drawing either a red card (hearts or diamonds) or a face card (jack, queen, or king)?

$$P(\mathsf{R} \cup \mathsf{F}) = P(\mathsf{R}) + P(\mathsf{F}) - P(\mathsf{R} \cap \mathsf{F})$$

$$P(\mathsf{R} \cup \mathsf{F}) = \frac{26}{52} + \frac{12}{52} - \frac{6}{52} = \frac{32}{52}$$

Addition Rule for Non-Disjoint Events

Example

Consider drawing a card from a standard deck of 52 playing cards. What's the probability of drawing either a red card (hearts or diamonds) or a face card (jack, queen, or king)?

$$P(\mathsf{R} \cup \mathsf{F}) = P(\mathsf{R}) + P(\mathsf{F}) - P(\mathsf{R} \cap \mathsf{F})$$

$$P(\mathsf{R} \cup \mathsf{F}) = \frac{26}{52} + \frac{12}{52} - \frac{6}{52} = \frac{32}{52}$$

The Law of Total Probability

Definition

The **Law of Total Probability** states that if $\{B_i : i = 1, 2, 3, \ldots, n\}$ is a finite partition of a sample space Ω and each event B_i is measurable, then for any event $A \subset \Omega$:

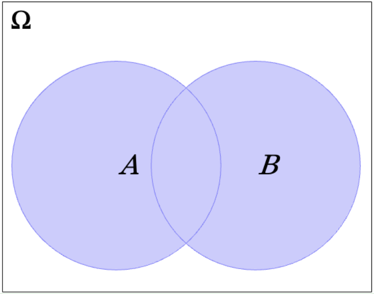
$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

- It's often used when we have multiple disjoint scenarios or partitions of the sample space.
- Since $\{B_i : i = 1, 2, 3, \ldots, n\}$ is a partition of Ω, $(B_1, B_2, \ldots, B_n)$ are mutually exclusive and exhaustive events.
- The theorem is extendable to countably infinite partitions $\{B_i : i = 1, 2, 3, \ldots\}$.

The Law of Total Probability

Example

Let consider $A, B \subset \Omega$:



by the law of total probability, we may define:

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B}).$$

Bayes Theorem

Definition

Bayes Theorem

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Where:

- $P(A|B)$ is the probability of event A occurring given that event B has occurred.
- $P(B|A)$ is the probability of event B occurring given that event A has occurred.
- $P(A)$ is the marginal probability of event A.
- $P(B)$ is the marginal probability of event B.

Bayes Theorem: Example

Example

Scenario Imagine we have a rare disease that affects 1% of the population. A medical test for this disease is not perfect. It correctly identifies the disease (true positive) 95% of the time and incorrectly diagnoses healthy individuals as having the disease (false positive) 3% of the time.

Bayes Theorem: Example

Example

Bayes' Theorem Application Let's calculate the probability that a person actually has the disease (D) given a positive test result (T^+).

- $P(D)$ is the prior probability of having the disease (1% or 0.01).
- $P(T^+|D)$ is the probability of a true positive (95% or 0.95).
- $P(T^+|\neg D)$ is the probability of a false positive (3% or 0.03).

We want to find $P(D|T^+)$.

$$P(D|T^+) = \frac{P(T^+|D) \cdot P(D)}{P(T^+|D) \cdot P(D) + P(T^+|\neg D) \cdot P(\neg D)}$$

Now, let's calculate this probability.

Contingency Tables and Joint Probabilities

Definition

Let consider $J, I \subset \Omega$. The contingency table

	I_1	I_2	**I**
J_1	f_{11}	f_{12}	f_{1+}
J_2	f_{21}	f_{22}	f_{2+}
J	f_{+1}	f_{+2}	N

showing the joint distribution of the two variables. Then:

$$P(I_i \cap J_j) = \frac{f_{ij}}{N}, \qquad \text{for } i, j = 1, 2.$$

Contingency table and marginal probabilities

Definition

Let consider $J, I \subset \Omega$. The contingency table

	I_1	I_2	**I**
J_1	f_{11}	f_{12}	f_{1+}
J_2	f_{21}	f_{22}	f_{2+}
J	f_{+1}	f_{+2}	N

showing the joint distribution of the two variables. Then:

$$P(I_i) = \frac{f_{i1} + f_{i2}}{N} = \frac{f_{+i}}{N}, \qquad \text{for } i = 1, 2.$$

$$P(J_i) = \frac{f_{j1} + f_{j2}}{N} = \frac{f_{j+}}{N}, \qquad \text{for } j = 1, 2.$$

Once we calculate joint and marginal probabilities, we can obtain the conditional probabilities.

Contingency probabilities: example

Example

	Have Shopped (S)	**Have Never Shopped** (N)	**Total**
Male (M)	500	700	1200
Female (F)	300	500	800
Total	800	1200	2000

- $P(M \cap N) = \frac{700}{2000} = \frac{7}{20}$

Contingency probabilities: example

Example

	Have Shopped (S)	**Have Never Shopped** (N)	**Total**
Male (M)	500	700	1200
Female (F)	300	500	800
Total	800	1200	2000

- $P(M|N) = \frac{P(M \cap N)}{P(N)} = \frac{7}{20} / \frac{1200}{2000} = \frac{7}{20} \cdot \frac{20}{12} = \frac{7}{12}$
- Easier way: $P(M|N) = \frac{P(M \cap N)}{P(N)} = \frac{f_{MN}}{N} / \frac{f_{+N}}{N} = \frac{f_{MN}}{N} \cdot \frac{N}{f_{+N}} = \frac{f_{MN}}{f_{+N}}$

$$P(M|N) = \frac{f_{MN}}{f_{+N}} = \frac{700}{1200} = \frac{7}{12}$$

Contingency probabilities: example

Example

	Have Shopped (S)	**Have Never Shopped** (N)	**Total**
Male (M)	500	700	1200
Female (F)	300	500	800
Total	800	1200	2000

- $P(M|N) = \frac{P(M \cap N)}{P(N)} = \frac{7}{20} / \frac{1200}{2000} = \frac{7}{20} \cdot \frac{20}{12} = \frac{7}{12}$
- Easier way: $P(M|N) = \frac{P(M \cap N)}{P(N)} = \frac{f_{MN}}{N} / \frac{f_{+N}}{N} = \frac{f_{MN}}{N} \cdot \frac{N}{f_{+N}} = \frac{f_{MN}}{f_{+N}}$

$$P(M|N) = \frac{f_{MN}}{f_{+N}} = \frac{700}{1200} = \frac{7}{12}$$

Contingency probabilities: example

Example

	Have Shopped (S)	**Have Never Shopped** (N)	**Total**
Male (M)	500	700	1200
Female (F)	300	500	800
Total	800	1200	2000

- $P(M \cup N)$?