STAT 423/523 Statistical Methods for Engineers and Scientists

Lecture 1: Review

Chencheng Cai

Washington State University

Sample Space

An **experiment** is any activity or process whose outcome is subject to uncertainty.

- Flip a coin (outcome: head or tail)
- Toss a die (outcome: nummber 1 to 6)
- Measure the weight of an apple (outcome: a real number)
- A patient takes a drug (outcome: recovery or not)

The **sample space** is the set of all possible outcomes of an experiment, denoted by $\mathcal{S}$.

- $\mathcal{S} = \{H, T\}$ for flipping a coin
- $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$ for tossing a die
- $\mathcal{S} = \mathbb{R}_+$ for measuring the weight of an apple
- $\mathcal{S} = \{\text{recovery}, \text{not recovery}\}$ for a patient taking a drug
- $\mathcal{S} = \{HH, HT, TH, TT\}$ for flipping a coin twice

Events

An **event** is a subset of the sample space.

- The event that observing one head when flipping a coin twice: $A = \{HT, TH\}$
- The event that observing an even number when tossing a die: $B = \{2, 4, 6\}$
- The event that the apple's weight is less than 1: $C = \{x \in \mathbb{R}_+ : x < 1\} = (0, 1)$
- A special case is the null event: $\varnothing$, which is the event that never happens.

Operations on events

- The **complement** of an event A, denoted by A', is the set of all outcomes in $\mathcal{S}$ that are not in A.
- The **union** of two events A and B, denoted by $A \cup B$, is the set of all outcomes that are in A or B.
- The **intersection** of two events A and B, denoted by $A \cap B$, is the set of all outcomes that are in both A and B.

Events

Consider the experiment that tossing a die twice.

- The sample space $\mathcal{S}$ contains $6 \times 6 = 36$ outcomes.
- The event that the the first toss is greater than the second is

$$A = \{21, 31, 32, 41, 42, 43, 51, 52, 53, 54, 61, 62, 63, 64, 65\}$$

- The event that the sum of two tosses is 7 is

$$B = \{16, 25, 34, 43, 52, 61\}$$

- The event that the sum of two tosses is 7 **and** the first toss is greater than the second is

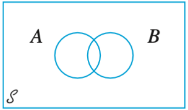
$$C = A \cap B = \{43, 52, 61\}$$

- The event that the sum of two tosses is 7 **or** the first toss is greater than the second is

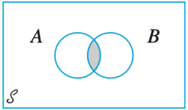
$$D = A \cup B = \{21, 31, 32, 41, 42, 43, 51, 52, 53, 54, 61, 62, 63, 64, 65, 16, 25, 34\}$$

Venn Diagram

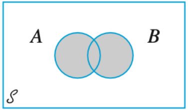
The Venn diagram is a visual representation of events.



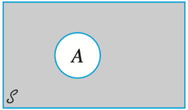
(a) Venn diagram of events A and B



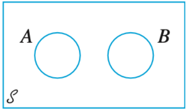
(b) Shaded region is $A \cap B$



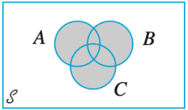
(c) Shaded region is $A \cup B$



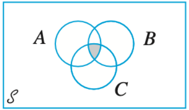
(d) Shaded region is A'



(e) Mutually exclusive events



(f) Shaded region is $A \cup B \cup C$



(g) Shaded region is $A \cap B \cap C$

For example, it is easy to see that $A \cup B = (A \cap B') \cup (A' \cap B) \cup (A \cap B)$.

Probability

The **probability** of an event A, denoted by $P(A)$, is a number between 0 and 1 that quantifies the likelihood of A occurring.

- $P(A) = 0$ means that A will never happen.
- $P(A) = 1$ means that A will always happen.

Axioms of probability

1. $P(A) \geq 0$ for any event A.
2. $P(\mathcal{S}) = 1$.
3. For any sequence of mutually disjoint events $A_1, A_2, \ldots$,

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

Remark: A and B are **mutually disjoint** if $A \cap B = \varnothing$.

Probability

Axioms of probability

1. $P(A) \geq 0$ for any event A.
2. $P(\mathcal{S}) = 1$.
3. For any sequence of mutually disjoint events $A_1, A_2, \ldots$,

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

Interpretations

1. The probability is always non-negative.
2. Because $\mathcal{S}$ contains every possible outcome, the probability of $\mathcal{S}$ is 1.
3. For disjoint events, the probability of their union is the sum of their probabilities.

Remark: The third axiom includes the finite case by setting $A_i = \varnothing$ for $i > N$.

Probability

Some properties of probability:

- $P(A') = 1 - P(A)$ for any event A. — either A happens or not.
  This can be shown by observing that $A \cap A' = \varnothing, A \cup A' = \mathcal{S}$ and by Axiom 3,
  $1 = P(\mathcal{S}) = P(A \cup A') = P(A) + P(A')$.

- $P(\varnothing) = 0$. — the null event never happens.

- $P(A) \leq 1$ for any event A. — the probability is always less than or equal to 1.

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ for any events A and B. — the **inclusion-exclusion principle**.
  This can be easily shown by the Venn diagram.

- $P(A \cup B) \leq P(A) + P(B)$ for any events A and B. — the **union bound**.

Probability

Consider tossing a die. Let the probability for each outcome as $P(\{1\}) = P(\{2\}) = \cdots = P(\{6\}) = 1/6$.

- The probability of observing an even number is $P(\{2,4,6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$. (Axiom 3)
- The probability of observing an odd number is $P(\{1,3,5\}) = 1 - P(\{2,4,6\}) = 1 - \frac{1}{2} = \frac{1}{2}$.
- The probability of observing a prime number is $P(\{2,3,5\}) = P(\{2\}) + P(\{3\}) + P(\{5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$. (Axiom 3)
- The probability of observing an even prime number is $P(\{2\}) = \frac{1}{6}$.
- Check the inclusion-exclusion principle:

$$P(\{1,3,5\} \cup \{2,3,5\}) + P(\{1,3,5\} \cap \{2,3,5\}) = P(\{1,3,5\}) + P(\{2,3,5\})$$

Conditional Probability

For any two events A and B with $P(B) > 0$, the **conditional probability** of A given B is

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

- $P(A|B)$ is the probability of A given that B has occurred.
- $P(A|B)$ is a number between 0 and 1.
- By multiplying $P(B)$ on both sides, we have the **multiplication rule**:

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

Independence

Two events A and B are **independent** if $P(A|B) = P(A)$, and are **dependent** otherwise.

- A and B are independent if and only if

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

- This definition of independence can be extended to more than two events: $A_1, A_2, \ldots, A_n$ are **mutually independent** if for any subset $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$,

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}).$$

Conditional Probability and Independence Example

Consider flip a coin twice. The sample space is $\mathcal{S} = \{HH, HT, TH, TT\}$. We can assign equal probability to each outcome (i.e. 1/4)

- The event that the first toss is head is $A = \{HH, HT\}$ with $P(A) = 1/2$.
- The event that the number of heads is 1 is $B = \{HT, TH\}$ with $P(B) = 1/2$.
- The conditional probability of A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{1/2} = \frac{1}{2} = P(A).$$

Therefore, A and B are independent.

Random Variable

For a given sample space $\mathcal{S}$ of some experiment, a **random variable** (rv) is any rule that associates a number with each outcome in $\mathcal{S}$. In mathematical language, a random variable is a function whose domain is the sample space and whose range is the set of real numbers.

- A **discrete random variable** is a random variable that can take on a countable number of values.
  - The number of heads when flipping a coin n times.
  - The number of defective items in a batch of 100.
  - The number of students in a class.
- A **continuous random variable** is a random variable that can take on an uncountable number of values.
  - The weight of an apple.
  - The time it takes to complete a task.
  - The temperature of a room.

Discrete Random Variable

The **probability distribution** or **probability mass function** (pmf) of a discrete random variable X is defined for every number x by $p(x) = P(X = x)$.

- We use the convention in the textbook that P stands for probability of events and p stands for probability distribution.

The **cumulative distribution function** (cdf) of a discrete random variable X is defined for every number x by

$$F(x) = P(X \leq x) = \sum_{y:y \leq x} p(y)$$

Discrete Random Variable

Consider toss a die twice. Let X be the sum of two tosses and let p be the pmf of X. Then

$$p(5) = P(X = 5) = P(\{14, 23, 32, 41\}) = \frac{4}{36} = \frac{1}{9}$$

The cdf of X is

$$F(5) = P(X \leq 5) = p(2) + p(3) + p(4) + p(5) = \frac{1}{36} + \frac{1}{18} + \frac{1}{12} + \frac{1}{9} = \frac{5}{18}$$

Discrete Random Variable

The **expected value** or **mean** of a discrete random variable X is defined by

$$E(X) = \sum_{x} x \cdot p(x)$$

For a function $g(X)$ of a discrete random variable X, the expected value of $g(X)$ is

$$E[g(X)] = \sum_{x} g(x) \cdot p(x)$$

The **variance** of a discrete random variable X is defined by

$$Var(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2$$

Discrete Random Variable

Let X and Y be two rvs and a and b be two constants. Then

- $E(aX + bY) = aE(X) + bE(Y)$
- $Var(aX) = a^2 Var(X)$
- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab \cdot Cov(X, Y)$

Discrete Random Variable

Common discrete distributions:

- Bernoulli
- Binomial
- Poisson
- Geometric
- Hypergeometric

Continuous Random Variables

The **probability density function** (pdf) of a continuous random variable X is a function $f(x)$ such that for any two numbers a and b with $a < b$,

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

The **cumulative distribution function** (cdf) of a continuous random variable X is defined for every number x by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Continuous Random Variables

The **expected value** or **mean** of a continuous random variable X is defined by

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

For a function $g(X)$ of a continuous random variable X, the expected value of $g(X)$ is

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f(x)dx$$

The **variance** of a continuous random variable X is defined by

$$Var(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2$$

Continuous Random Variables

Let X be a uniform random variable on the interval $[0, 1]$. Then its pdf is $f(x) = 1$ for $0 \leq x \leq 1$ and $f(x) = 0$ otherwise because for any a and b with $0 \leq a \leq b \leq 1$,

$$P(a \leq X \leq b) = \int_a^b 1 dx = b - a$$

The cdf of X is (for $x \in [0, 1]$)

$$F(x) = P(X \leq x) = \int_0^x 1 dx = x$$

and $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 1$.
The expected value of X is

$$E(X) = \int_0^1 x dx = \frac{1}{2}$$

and the variance of X is

$$Var(X) = E(X^2) - E(X)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Continuous Random Variables

Common continuous distributions:

- Uniform
- Normal
- Exponential
- Gamma
- Beta

Joint Distribution

We will take continuous random variables as an example. For discrete random variables, please replace integrals by summations.

The **joint distribution** of two continuous random variables X and Y is defined by the joint pdf $f(x, y)$ such that for any two-dimensional region A,

$$P((X, Y) \in A) = \iint_A f(x, y) dx dy$$

The **marginal distribution** of X is the pdf of X:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

The **conditional distribution** of Y given $X = x$ is the pdf of Y given $X = x$:

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

Joint Distribution

Let X and Y be two continuous random variables with joint pdf $f(x,y)$. The **expected value** of a function $g(X,Y)$ is

$$E[g(X,Y)] = \iint g(x,y)f(x,y)dxdy$$

The **covariance** of X and Y is

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

The **correlation** of X and Y is

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

Joint Distribution

X and Y are **independent** if $f(x,y) = f_X(x)f_Y(y)$ for all x and y.

X and Y are **uncorrelated** if $Cov(X,Y) = 0$.

- Independence implies uncorrelated, but uncorrelated does not imply independence.
- Example: X is a standard normal random variable and Z is a Rademarcher random variable (random ± 1). Let $Y = XZ$. Then X and Y are uncorrelated but not independent.