ST2334 Midterms
by Zong Xun

Event operations

Sample space is the set of all possible outcomes. It is an event itself, and a **sure event**. On the other hand, the **empty set**, $\emptyset$, is known as the **null event**. Event operations include union, intersection, and complement.
Union and **Intersection** can be extended to *n* events.
For union,

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cdots \cup A_n = \{x\colon x \in A_1 \text{ or } x \in A_2 \text{ or } \ldots \text{ or } x \in A_n\}$$

For intersection,

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cdots \cap A_n = \{x\colon x \in A_1 \text{ and } x \in A_2 \text{ and } \ldots \text{ and } x \in A_n\}$$

Event Relationships

- **Mutually exclusive** if $A \cap B = \emptyset$.
- If all elements in A are also elements in B, we say that B **contains** A.
- If B contains A and A contains B, then A and B are **equivalent**.

Important Properties

1. Distributive laws
2. $A \cup B = A \cup (B \cap A')$
3. $A = (A \cap B) \cup (A \cap B')$
4. De Morgan's laws
5. $A \cap B' = A - (A \cap B)$
6. If $(A \cap B') \cup (A' \cap B) = \emptyset \Rightarrow A = B$

Counting Methods

Multiplication is used used when there are n **different** events(independent) to be performed sequentially.
Addition is used when the same event can be performed by k different procedure(mutually exclusive).

Permutation

The number of ways to arrange r objects out of n.

$$P_r^n = \frac{n!}{(n-r)!} = \binom{n}{r} \times P_r^r$$

- Permutations around a circle = $(n-1)!$
- Permutations when not all objects are distinct = $\frac{n!}{n_1!n_2!\ldots n_k!}$

Combination

A selection of r objects out of n, without regard to the order.

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \binom{n}{n-r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

Probability

$$Pr(A) = lim_{n\to\infty} f_A$$

An interpretation of probability is relative frequency.

Axioms

Probability is a function on the collection of events of the sample space S, satisfying:

1. For any event A, $0 \le P(A) \le 1$
2. For the sample space, $P(S) = 1$
3. For any two mutually exclusive events A and B, i.e $A \cap B = \emptyset$, $P(A \cup B) = P(A) + P(B)$

Propositions

1. The probability of the empty set, $P(\emptyset) = 0$.
2. If $A_1, \ldots, A_n$ are mutually exclusive events ($A_i \cap A_j$ for any $i \ne j$), then
$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$$
3. $P(A') = 1 - P(A)$
4. $P(A) = P(A \cap B) \cup P(A \cap B')$
5. For any two events A and B,
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
6. If A is contained by B, then $P(A) \le P(B)$

Conditional Probability

For any two events A and B with P(A) $>$ 0, the conditional probability of B given that A has occurred is defined by

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{P(B)P(A|B)}{P(A)} \\ &= \frac{P(B)P(A|B)}{P(B)P(A|B) + P(B')P(A|B')} \end{aligned}$$

Two events are said to be **independent** if and only if $P(A \cap B) = P(A)P(B)$. We say that the probability of one event occurring does not affect the probability of another. If A and B are independent, then

- their conditional probability, $P(A|B) = P(A)$.
- $A \perp B', A' \perp B,$ and $A' \perp B'$.
- A and B cannot be mutually exclusive, supposing that $Pr(A), Pr(B) > 0$

Law of total probability

$$Pr(B) = \sum_{i=1}^{n} Pr(B \cap A_i) = \sum_{i=1}^{n} Pr(A_i)Pr(B|A_i)$$

Assuming that events $A_1, \ldots, A_n$ are mutually exclusive and exhaustive events.

Bayes' Theorem

Let $A_1, A_2, \ldots, A_n$ be a partition of the sample space S. Then

$$Pr(A_k|B) = \frac{Pr(A_k)Pr(B|A)}{\sum_{i=1}^{n} Pr(A_i)Pr(B|A_i)}$$

Bayes Theorem can be derived using conditional probability, multiplication rule and the law of total probability.

Birthday Problem

How many people in the room to guarantee that there are at least 2 people with the same birthday?
$Pr(n) = 1 - Pr(q)$, where $Pr(q)$ is the probability that every person in the room has a different birthday i.e $\frac{365(364)(363)\ldots(365-n+1)}{365^n}$.

Inverse Birthday Problem

How many people in the room to guarantee that the probability of another person with the same birthday as you is greater than 0.5? $Pr(n) = 1 - Pr(q) \ge 0.5$, where Pr(q) is the probability of someone having a different birthday i.e. $(\frac{364}{365})^n$

Monty Hall Problem

Suppose you chose to switch,
W = {Win the car}
A = {Car is behind the door of initial pick}

$$\begin{aligned} P(W) &= P(A)P(W|A) + P(A')P(W|A') \\ &= \frac{1}{3}(0) + \frac{2}{3}(1) \\ &= \frac{2}{3} \end{aligned}$$

Useful tips

- Take note of labelled objects, remove all duplicates (including duplicates caused by ordering).
- $A \cap B \ne \emptyset, A \cap C \ne \emptyset$ does not imply that $A \cap B \cap C \ne \emptyset$.
- Reduce the event into multiple smaller mutually exclusive or independent events.
- Reduce permutations and combinations problems by grouping objects. Permute objects in the group.
- Consider the negation.

Random Variables

A function X, which assigns a real number to every $s \in S$ is called a random variable. Range space, $R_x = \{x | x = X(s), s \in S\}$.

Discrete Probability Distribution

Each value of X has a certain probability, $f(x)$, and this function $f(x)$ is called the **probability mass function**. The collection of pairs $(x_i, f(x_i))$ is called the probability distribution of X. It must satisfy the following:

1. $f(x_i) \ge 0$ for all $x_i \in R_x$
2. $f(x) = 0$ for all x not in R_x
3. The sum of the probabilities must equate to 1 i.e. $\sum_{i=1}^{\infty} f(x_i) = 1$

Continuous Probability Distribution

For a continuous RV X, R_x is an interval or a collection of intervals. The **probability density function** is defined to quantify the probability that X is in a certain range. It must satisfy the following:

1. $f(x) \ge 0$ for all $x \in R_x$ and 0 if not. This means that Pr(A) = 0 does not imply that A is $\emptyset$
2. $\int_{Rx} f(x)dx = 1$
3. $P(a \le X \le b) = \int_a^b f(x)dx$. Consequently, $P(X = a) = 0$.

Cumulative Probability Distribution

For any RV X, we define its cumulative distribution function by $F(x) = P(X \le x)$. The cumulative probabilities **cannot** exceed 1.
For the **discrete** case, the c.d.f is a step function. For any two numbers $a < b$, we have

$$P(a \le X \le b) = P(X \le b) - P(X < a) = F(b) - F(a-)$$

For the **continuous** case,

$$F(x) = \int_{-\infty}^{x} f(t)dt \Rightarrow f(x) = \frac{dF(x)}{dx}$$

Some remarks:

- No matter if its continuous or discrete, F(X) is non-decreasing i.e. $x_1 < x_2, F(X_1) \le F(X_2)$.
- p.f and c.d.f have a one-to-one correspondence.
- The ranges of F(X) and f(x) should satisfy:
  - $0 \le F(X) \le 1$
  - for discrete, $0 \le f(x) \le 1$
  - for continuous, $f(x) \ge 0$, but **no need** that $f(x) \le 1$ (as long as total sum under curve is 1).
- cdf have to be **right continuous**

Expectation and Variance

For discrete,

$$E(X) = \sum_{x_i \in R_x} x_i f(x_i)$$
$$Var(X) = \sum_{x_i \in R_x} (x - \mu_x)^2 f(x)$$

For continuous,

$$E(X) = \int_{x_i \in R_x} x f(x) dx$$
$$Var(X) = \int_{x_i \in R_x} (x - \mu_x)^2 f(x) dx$$

Properties of Expectation

1. $E(aX + b) = aE(X) + b$
2. $E(X + Y) = E(X) = E(Y)$
3. Let g be an arbitrary function.
   - if X is a discrete RV, $E[g(X)] = \sum_{x \in R_x} g(x)f(x)$
   - if X is a continuous RV, $E[g(X)] = \int_{R_x} g(x)f(x)dx$
4. Let X be a positive integer-valued (excluding 0) random variable. (tut 4, qn 8)
$$E(X) = \sum_{k=1}^{\infty} P(X \ge k)$$
5. There exists probability distributions for which E(X) do not exists, i.e. $E(X) = \infty$

Properties of Variance

1. $Var(aX + b) = a^2 Var(X)$
2. Variance can also be computed using $E(X^2) - [E(X)]^2$
3. the standard deviation of X is the root of Variance.
4. variance is always greater than 0, unless Pr(X = E(X)) = 1