

Emory University
MATH 361 Mathematical Statistics I
Learning Notes

Jiuru Lyu

September 1, 2023

Contents

1	Prerequisites	2
2	Probability	5
2.1	Sample Space and Probability	5
2.2	Conditional Probability and Independence	8

1 Prerequisites

Definition 1.0.1 (Geometric Series). A geometric series has the form

$$\sum_{n=1}^{\infty} ar^{n-1} = a + ar + ar^2 + \dots$$

If $|r| < 1$, then the series converges to $\frac{a}{1-r}$. Otherwise, it diverges.

Example 1.0.2 Does the series $\sum_{n=1}^{\infty} 2^{2n} 3^{1-n}$ converge or diverge?

Solution 1.

Note that

$$2^{2n} 3^{1-n} = (2^2)^n 3^{1-n} = 4^n \left(\frac{1}{3}\right)^{n-1} = 4 \cdot 4^{n-1} \left(\frac{1}{3}\right)^{n-1} = 4 \left(\frac{4}{3}\right)^{n-1}.$$

So,

$$\sum_{n=1}^{\infty} 2^{2n} 3^{1-n} = \sum_{n=1}^{\infty} 4 \left(\frac{4}{3}\right)^{n-1}$$

is a geometric series, with $a = 4$ and $r = \frac{4}{3}$.

Since $|r| = \left|\frac{4}{3}\right| = \frac{4}{3} > 1$, the series diverges. □

Definition 1.0.3 (Taylor Series). The Taylor series expanded about a of a differentiable function f is

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!} (x-a)^2 + \dots$$

Definition 1.0.4 (Maclaurin Series). The Taylor series expanded about $a = 0$.

Remark. The Maclaurin Series of e^x is given by $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

Theorem 1.0.5 Binomial Expansion

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k},$$

where $\binom{n}{k}$ is read as “ n choose k ” and can also be written as nCk .

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1) \cdots (n-k+1)}{k!}.$$

Theorem 1.0.6 Integration by Parts

$$\int u \, dv = uv - \int v \, du.$$

Example 1.0.7 Evaluate $\int x e^{-x} \, dx$.

Solution 2.

Let $u = x$, $dv = e^{-x} \, dx$. So, $du = dx$ and $v = \int e^{-x} \, dx = -e^{-x}$. Then,

$$\int x e^{-x} \, dx = -x e^{-x} - \int -e^{-x} \, dx = -x e^{-x} - e^{-x} + C.$$

□

Definition 1.0.8 (Type I Improper Integral). If $\int_a^t f(x) \, dx$ exists for all $t > 0$, then

$$\int_a^\infty f(x) \, dx = \lim_{t \rightarrow \infty} \int_a^t f(x) \, dx.$$

Example 1.0.9 Evaluate $\int_0^\infty x e^{-x} \, dx$.

Solution 3.

$$\begin{aligned} \int_0^\infty x e^{-x} \, dx &= \lim_{t \rightarrow \infty} \int_0^t x e^{-x} \, dx = \lim_{t \rightarrow \infty} \left[-x e^{-x} - e^{-x} \right]_0^t \\ &= \lim_{t \rightarrow \infty} \left(-t e^{-t} - e^{-t} + 1 \right) \\ &= -\lim_{t \rightarrow \infty} \left(\frac{t}{e^t} \right) - \lim_{t \rightarrow \infty} e^{-t} + 1 \\ &= -\lim_{t \rightarrow \infty} \left(\frac{1}{e^t} \right) - 0 + 1 = -0 - 0 + 1 = 1. \end{aligned}$$

□

Example 1.0.10 Double Integrals over Irregular Domains.

Consider

$$\iint_D 4xy - y^4 \, dA,$$

where D is the region bounded between $y = \sqrt{x}$ and $y = x^3$.

Evaluate this double integral over D .

Solution 4.

Firstly, we draw the diagram representing D as follows:



$$\begin{aligned}\iint_D 4xy - y^3 \, dA &= \int_0^1 \int_{x^3}^{\sqrt{x}} 4xy - y^3 \, dy \, dx = \int_0^1 \left[2xy^2 - \frac{1}{4}y^4 \right]_{x^3}^{\sqrt{x}} dx \\ &= \int_0^1 2x(x - x^6) - \frac{1}{4}(x^2 - x^{12}) \, dx \\ &= \int_0^1 2x^2 - 2x^7 - \frac{1}{4}x^2 + \frac{1}{4}x^{12} \, dx \\ &= \left[\frac{2}{3}x^3 - \frac{1}{4}x^8 - \frac{1}{12}x^3 + \frac{1}{52}x^{13} \right]_0^1 \\ &= \frac{2}{3} - \frac{1}{4} - \frac{1}{12} + \frac{1}{52} = \frac{55}{156}.\end{aligned}$$

□

2 Probability

2.1 Sample Space and Probability

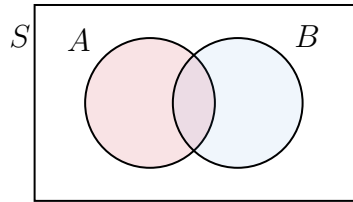
Definition 2.1.1 (Experiment). An *experiment* is a procedure with well-defined outcome.

Definition 2.1.2 (Sample Space/ S). The *sample space*, denoted as S is the set of all possible outcomes of an experiment.

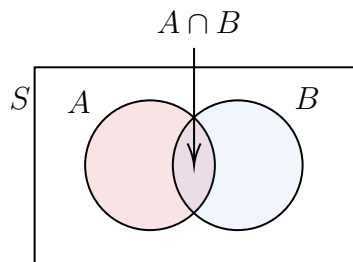
Definition 2.1.3 (Event). An *event* is a collection of outcomes.

Example 2.1.4 Consider flipping two coins. Use H to represent heads and T to represent tails. Then, $S = \{HH, HT, TH, TT\}$. Event “one heads” = $\{HT, TH\}$, and the event “at least one heads” = $\{HT, TH, HH\}$.

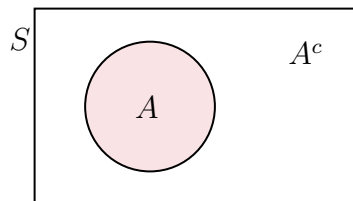
Definition 2.1.5 (Union/ \cup). $A \cup B$ is the *union* of A and B , meaning everything in A and everything in B .



Definition 2.1.6 (Intersection/ \cap). $A \cap B$ is the *intersection* of A and B , everything in both A and B .



Definition 2.1.7 (Complement/ A^c). A^c denotes the *complement* of A , meaning everything in S that is not in A .



Corollary 2.1.8 $A \cap A^c = \{\} = \emptyset$.

Definition 2.1.9 (Mutually Exclusive). Two sets A and B over the same sample space are *mutually exclusive* if they have no outcomes in common. i.e., $A \cap B = \emptyset$.

Remark. A and A^c are mutually exclusive, but not all sets mutually exclusive are complements of each other.

Definition 2.1.10 (Probability Function). Let A be an event over a sample space S . Then, $P(A)$ denotes the *probability* of A and P is the *probability function*. The probability function P assigns a number $P(A)$ for each event $A \subseteq S$.

Axiom 2.1.11 Kolmogorov Axioms

1. Let A be an event in S , then $P(A) \geq 0$.
2. $P(S) = 1$.
3. If A and B are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
4. If A_1, \dots, A_n, \dots are mutually exclusive sets, then

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

Proposition 2.1.12 $P(A^c) = 1 - P(A)$.

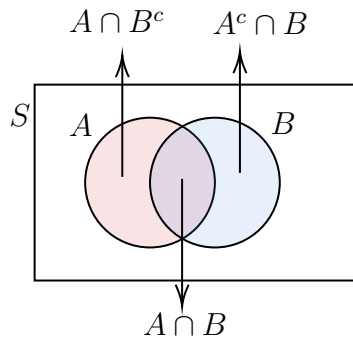
Proof 1. Note that $P(S) = 1$. Since $A^c \cup A = S$, we have $P(A \cup A^c) = 1$. Since A and A^c are mutually exclusive, $P(A \cup A^c) = P(A) + P(A^c) = 1$. So, $P(A^c) = 1 - P(A)$. ■

Proposition 2.1.13 $P(\emptyset) = 0$.

Proof 2. Note that $P(S) = 1$. Then, $P(S^c) = 1 - P(S)$. By definition, we know $S^c = \emptyset$. So, $P(\emptyset) = 1 - P(S) = 1 - 1 = 0$. ■

Proposition 2.1.14 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof 3. Consider the following Venn diagram:



Note that $P(A) = P(A \cap B) + P(A \cap B^c)$ and $P(B) = P(A \cap B) + P(A^c \cap B)$. So, we have

$$P(A) + P(B) = \boxed{P(A \cap B^c) + P(A^c \cap B) + P(A \cap B)} + P(A \cap B). \quad (1)$$

From the Venn diagram, we notice that $P(A \cap B^c) + P(A^c \cap B) + P(A \cap B)$ is exactly $P(A \cup B)$. So, Eq. (1) becomes $P(A) + P(B) = P(A \cup B) + P(A \cap B)$. That is exactly what is required: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. ■

Definition 2.1.15 (Classical Probability). In a discrete and finite case, S is finite and all outcomes are equally likely, and the probability function is defined as

$$P(A) = \frac{|A|}{|S|},$$

where $|A|$ is the cardinality of A and $|S|$ is the cardinality of S .

Example 2.1.16 Despite the definition of classical probability (probability function defined for a discrete and finite case), there are other definitions of probability functions:

1. Discrete and Countably Infinite:

Let $S = \mathbb{N}$ be the set of natural numbers. Then,

$$P(k) = \frac{1}{2^k}.$$

It can also be verified that

$$P(S) = \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

2. Continuous and Uncountably Infinite:

Let $S = [0, 1]$. Suppose E is a subset of $[0, 1]$ such that $\int_E dx$ is defined. Then,

$$P(E) = \int_E dx,$$

and it can also be verified that $P(S) = 1$.

2.2 Conditional Probability and Independence

Definition 2.2.1 (Conditional Probability). We read $P(A|B)$ as the probability of A given B . Knowing B occurs, we create a new sample space, in which the probability of A occurs changes:

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

Corollary 2.2.2 $P(A \cap B) = P(A|B)P(B)$

Example 2.2.3 Find the probability of dealing A first, 2 second, and 3 third.

Solution 1.

$$\begin{aligned} P(\text{dealing } A, 2, 3) &= P(A \text{ first})P(2 \text{ second}|A \text{ first})P(3 \text{ third}|A \text{ first} \cap 2 \text{ second}) \\ &= \frac{4}{52} \cdot \frac{4}{51} \cdot \frac{4}{50} \end{aligned}$$

□

Corollary 2.2.4 $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_2 \cap A_1)$.

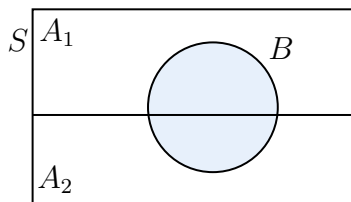
Theorem 2.2.5 The Law of Total Probability

Suppose the sample space $S = A_1 \cup A_2 \cup \cdots \cup A_n$, with $A_i \cap A_j = \emptyset \quad \forall i \neq j$. Then,

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_n).$$

Remark. This theorem gives us a nice way to partition the sample space.

Example 2.2.6



As represented in the diagram above, $P(B) = P(B \cap A_1) + P(B \cap A_2)$.

Theorem 2.2.7 Bayes Theorem

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

Definition 2.2.8 (Independence). Events A and B are *independent* if $P(A|B) = P(A)$, meaning the occurrence of B does not affect the occurrence of A .

Corollary 2.2.9 If A and B are independent, then $P(A \cap B) = P(A|B)P(B) = P(A)P(B)$.

Example 2.2.10 Coronary Artery Disease (CAD)

The probability of someone having CAD is 60%. In a study of 101 patients, 37 of them are known to NOT have CAD and 64 are known to have CAD. Of the 37 patients without CAD, 34 had negative tests while 3 had positive tests. Of the 64 with CAD, 54 had positive tests and 10 had negative tests. Find the probability of a patient has CAD given positive test.

Solution 2.

Let $T+$ be positive test, $T-$ be negative test, $D+$ be presence of CAD, and $D-$ be absence of CAD. Then, from the problem, we have

$$P(D+) = 0.6; \quad P(D-) = 1 - P(D+) = 0.4$$

and

$$P(T+|D+) = \frac{54}{64} \approx 0.84; \quad P(T-|D-) = \frac{34}{37} \approx 0.92; \quad P(T+|D-) = \frac{3}{37} \approx 0.08.$$

Then, by Bayes Theorem,

$$\begin{aligned} P(D+|T+) &= \frac{P(T+|D+)P(D+)}{P(T+|D+)P(D+) + P(T+|D-)P(D-)} \\ &= \frac{0.84 \times 0.6}{0.84 \times 0.6 + 0.08 \times 0.4} \approx \boxed{0.94}. \end{aligned}$$

□