Emory University

MATH 361 Mathematical Statistics I

Learning Notes

Jiuru Lyu

August 29, 2023

Contents

1 Prerequisites

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

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

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

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

Solution 1.

Note that

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

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 $r = \dfrac{4}{3}$.

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

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

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

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

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

Theorem 1.0.5 Binomial Expansion

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

where $\dbinom{n}{k}$ 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

$$\int u \, \mathrm{d}v = uv - \int v \, \mathrm{d}u.$$

Example 1.0.7 Evaluate $\displaystyle\int xe^{-x} \, \mathrm{d}x$.

Solution 2.

Let $u = x, \ \mathrm{d}v = e^{-x} \, \mathrm{d}x$. So, $\mathrm{d}u = \mathrm{d}x$ and $v = \displaystyle\int e^{-x} \, \mathrm{d}x = -e^{-x}$. Then,

$$\int xe^{-x} \, \mathrm{d}x = -xe^{-x} - \int -e^{-x} \, \mathrm{d}x = -xe^{-x} - e^{-x} + C.$$

□

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

$$\int_a^\infty f(x) \, \mathrm{d}x = \lim_{t\to\infty} \int_a^t f(x) \, \mathrm{d}x.$$

Example 1.0.9 Evaluate $\displaystyle\int_0^\infty xe^{-x} \, \mathrm{d}x$.

Solution 3.

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

□

Example 1.0.10 Double Integrals over Irregular Domains.

Consider

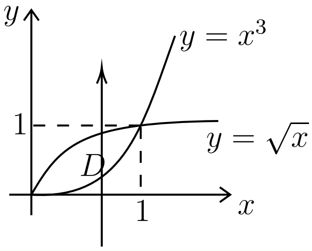
$$\iint_D 4xy - y^4 \, \mathrm{d}A,$$

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:



$$\iint_D 4xy - y^3 \,\mathrm{d}A = \int_0^1 \int_{x^3}^{\sqrt{x}} 4xy - y^3 \,\mathrm{d}y\mathrm{d}x = \int_0^1 \left[2xy^2 - \frac{1}{4}y^4\right]_{x^3}^{\sqrt{x}} \mathrm{d}x$$
$$= \int_0^1 2x\big(x - x^6\big) - \frac{1}{4}\big(x^2 - x^{12}\big)\,\mathrm{d}x$$
$$= \int_0^1 2x^2 - 2x^7 - \frac{1}{4}x^2 + \frac{1}{4}x^{12}\,\mathrm{d}x$$
$$= \left[\frac{2}{3}x^3 - \frac{1}{4}x^8 - \frac{1}{12}x^3 + \frac{1}{52}x^{13}\right]_0^1$$
$$= \frac{2}{3} - \frac{1}{4} - \frac{1}{12} + \frac{1}{52} = \frac{55}{156}.$$

□

2 Probability

2.1 Sample Space and Probability

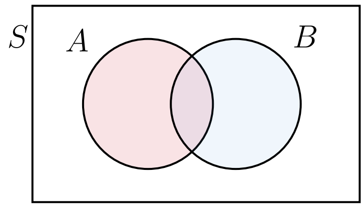
Definition 2.1.1 (Experiment). An *experiment* is a procedure with well-defined outcome.

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

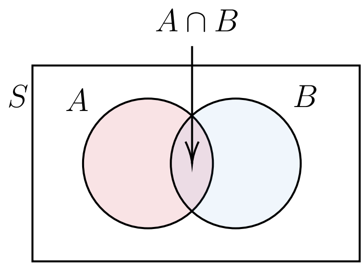
Definition 2.1.3 (Event). An *event* is a collection of outcomes.

> **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"$= \{HT, TH, HH\}$.

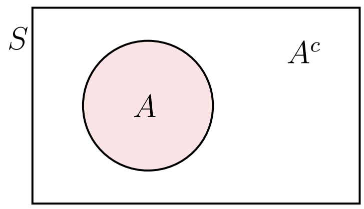
Definition 2.1.5 (Union/$\cup$). $A \cup B$ is the *union* of A and B, meaning everything in A and everything in B.



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 $A \cap A^c = \{\} = \varnothing$.

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., $A \cap B = \varnothing$.

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 $A_1, \ldots, A_n, \ldots$ 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 $P(A^c) = 1 - P(A)$.

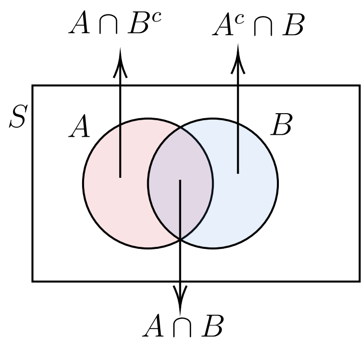
Proof 1. Note that $P(S) = 1$. Since $A^c \cup A = S$, we have $P(A \cap A^c) = 1$. Since A and A^c are mutually exclusive, $P(A \cup A^c) = P(A) + P(A^c) = 1$. So, $P(A^c) = 1 - P(A)$. ■

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

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

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

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

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 $\int_E \mathrm{d}x$ is defined. Then,

   $$P(E) = \int_E \mathrm{d}x,$$

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