BIOS 660/BIOS 672 (3 Credits):
Probability and Statistical Inference I

Jianwen Cai

https://sakai.unc.edu/portal/site/bios660-bios672-3-credits

Notes 3

Axiomatic Probability **2**
Intuitive probability . . . 3
Axiomatic probability . . . 4
Kolmogorov's Axioms of Probability . . . 5
Axioms continued . . . 6
Probability calculus . . . 7
Monotone Sequence of Events . . . 8
Boole's inequality (Union bound) . . . 9

Counting: Preliminaries **10**
Discrete Probability . . . 11
Simplest Setting . . . 12
Simple Examples . . . 13
More Preliminaries . . . 14

Counting: Ordered Samples **15**
Ordered Samples . . . 16
Sampling **with** Replacement . . . 17
Sampling **without** Replacement . . . 18
Examples . . . 19
Examples (1) . . . 20
Examples (2) . . . 21
Examples (3) . . . 22
Examples (4): Birthday Problem . . . 23

Additional Reading **24**
Additional Reading . . . 25

Intuitive probability

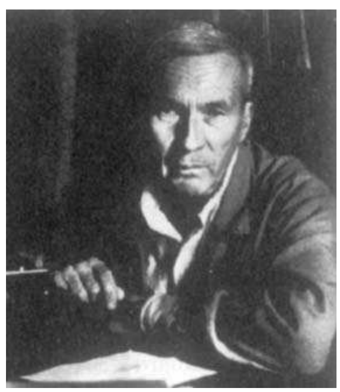
- **Random experiment**: Can be repeated indefinitely but future outcomes cannot be exactly predicted.
- **Stabilization of relative frequency**: Suppose we perform "independent" repetitions of a random experiment and count how many times an "event" E occurs. Let $f_n(E)$ be the number of occurrences in n repetitions. Then the *relative frequency* of E

  $$r_n(E) = \frac{f_n(E)}{n}$$

  "converges" to some number as $n \to \infty$. We call this number $P(E)$, the probability of E.
- Examples: rolling of a die, spin the wheel.
- But what is "event", "independent", "convergence"?

Axiomatic probability

Andrey Kolmogorov (1903–1987)



Grundbegriffe der Wahrscheinlichkeitsrechnung (1933)
Basic Concepts of Probability Calculation
Foundations of the Theory of Probability

Kolmogorov's Axioms of Probability

Suppose $E \subset \Omega$ and we want to assign a probability to E.
$P(E)$ is a real-valued function of the event E, called the probability of E, that satisfies the following:

Axioms

i. Regularity: $P(\Omega) = 1$

ii. Non-negativity: $P(E) \geq 0$

iii. Countable Additivity: If $E_1, E_2, \ldots$ are mutually exclusive $(E_i E_j = \emptyset, \;\; i \neq j)$, then $P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$ (countable additivity)

The last axiom contains a hidden assumption, namely that if $E_1, E_2, \ldots$ is an infinite sequence of events, then $\bigcup_{i=1}^{\infty} E_i$ is also an event. This is guaranteed if the events $E_1, E_2, \ldots$ belong to a σ-field.

Axioms continued

Definition: A probability measure P defined on a $\sigma-$field of subsets of Ω is a real-valued set function satisfying Kolmogorov's Axioms.

Definition - formal: A probability space is denoted by $(\Omega, \mathcal{A}, P)$ where Ω is the sample space, $\mathcal{A}$ refers to a $\sigma-$field of subsets of Ω, and P is a probability measure.

Probability calculus

If P is a probability function and A is any set in $\mathcal{A}$:

1. $P(\emptyset) = 0$
2. $P(A) \leq 1$
3. $P(A^c) = 1 - P(A)$

If P is a probability function and A, B are sets in $\mathcal{A}$:

1. $P(A - B) = P(A) - P(A \cap B)$ (Note: C&B uses $\setminus$ instead of '-')
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. $A \subset B \Rightarrow P(A) \leq P(B)$

Law of Total Probability: If P is a probability function and $\{C_1, C_2, \dots\}$ is a partition of Ω:

$$P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$$

Monotone Sequence of Events

(i) If $\{E_n\}$ is an increasing sequence of sets and $E_n \in \mathcal{A}$, then $P(\lim_n E_n) = \lim_n P(E_n)$;
(ii) If $\{E_n\}$ is a decreasing sequence of sets, then $P(\lim_n E_n) = \lim_n P(E_n)$.

Proof of (i):

Boole's inequality (Union bound)

Suppose $E_1, E_2, \ldots, E_n$ are events.

$$P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$$

Counting: Preliminaries

Discrete Probability

- **Discrete Sample Space:** A sample space is called **discrete** if it contains only finitely many points or infinitely many points which can be arranged into a simple sequence e_1, e_2, $\ldots$, i.e. a countable number of elements.
- $\mathcal{F} = 2^{\Omega}$.
- $P(\{e_1\}) + P(\{e_2\}) + \ldots = P(\Omega) = 1$, and more generally, for a set A that contains the points $a_1, a_2, \ldots, a_k$, then $P(A) = P(\{a_1\}) + P(\{a_2\}) + \ldots + P(\{a_k\})$.

Simplest Setting

Simplest setting:

1. Finite sample space with n elements, $\Omega = \{e_1, e_2, \ldots, e_n\}$
2. All sample points are equally probable.

In this case, $P(\{e_j\}) = \frac{1}{n}$, and for an event $\{a_1, \ldots, a_k\} = A \subset \Omega$,
$P(A) = kP(\{a_j\}) = \mu(A)/\mu(\Omega) = k/n$ where μ is the counting measure (counts the number of elements in a set).
$P(A)$ is simply the sum of the probabilities of each of the elements in A.

Counting and combinatorics will be concerned with identifying the number of individual elements within A.

Simple Examples

Examples: Suppose we have a standard deck of cards and we draw a single card. (probability of drawing any card is the same so they are equiprobable)

- Probability of drawing Jack of Hearts is 1/52
- Probability of drawing a Club is 13/52 = 1/4
- Probability of drawing an Ace is 4/52 = 1/13
- Probability of drawing an Ace or Heart is 16/52 = 4/13

More Preliminaries

- **Pairs:** With m elements $a_1, \ldots, a_m$, and n elements $b_1, \ldots, b_n$, it is possible to form nm pairs (a_j, b_k) containing one element from each group.
- **Multiplets:** Given r experiments that are to be performed such that the r^{th} experiment may result in any of n_r possible outcomes, then there is a total of $n_1 n_2 \ldots n_r$ possible outcomes of the r experiments.
- Examples:
  1. College committee consists of 3 freshman, 4 sophomores, 5 juniors, and 2 seniors. A subcommittee consisting of 1 person from each class is to be chosen. How many subcommittees are possible?
  2. How many 7-place license plates are possible if the first 3 places are to be occupied by letters and the final 4 by numbers?

Counting: Ordered Samples

Ordered Samples

- Many probability problems use these same ideas of counting in settings where repeated selections are taken from the same set of objects. The number of possible selection patterns (and hence the probability of a particular pattern) depends on
  1. whether or not items are replaced after selection
  2. whether or not the order of selection matters
- Suppose we have a “population” of n elements $a_1, \ldots, a_n$. Any ordered arrangement $a_{j_1}, \ldots, a_{j_r}$ of r symbols is called an *ordered sample of size* r.
  1. Sampling with replacement: each selection is made from the entire population (each of the r elements can be chosen in n ways. there are n^r possible samples)
  2. Sampling without replacement: an element, once chosen, is removed from the population. (sample cannot exceed size n)

Sampling with Replacement

Draw	Set
1	$A_1, A_2, \ldots, A_n$
2	$A_1, A_2, \ldots, A_n$
$\vdots$	
r	$A_1, A_2, \ldots, A_n$

Theorem: For a population of n elements and a prescribed sample size r, there exist n^r different samples with replacement

Example: The Braille alphabet has 6 locations that can be raised or not raised. How many combinations of raised/not raised locations are there?

Why does this differ from the actual number of Braille letters?

Sampling without Replacement

Draw	Set	
1	$A_1, A_2, \ldots, A_n$	
2	$B_1, B_2, \ldots, B_{n-1}$	$B_i's$ are $A_i's$ not drawn
3	$C_1, C_2, \ldots, C_{n-2}$	$C_i's$ are $B_i's$ not drawn
$\vdots$		
r	$X_1, X_2, \ldots, X_{n-(r-1)}$	$X_i's$ are remaining items after first $r-1$ draws

- **Theorem:** For a population of n elements and a prescribed sample size r, there exist $n(n-1)(n-2)...(n-r+1) = \frac{n!}{(n-r)!}$ different samples without replacement.
- An important special case occurs when $r = n$ (define 0!=1):

$$\frac{n!}{(n-r)!} = \frac{n!}{(n-n)!} = n(n-1)(n-2)\cdots 2 \cdot 1 = n!$$

This is the number of **permutations** of a set of size n

Examples

1. You want to take a photo of six of your friends. How many ways can you arrange them in a row?
2. A competition has 10 competitors of which 4 are Russian, 3 are from the USA, 1 is from Brazil, and 2 are from China. If the tournament result lists just the nationalities of the players in which they placed, how many outcomes are possible?
3. You want to invite those same 6 friends to dinner. How many ways can you arrange them at your round dining table?
4. 4 Martians, 3 Plutonians and 5 Jupiterians are to be seated in a row at an interplanetary conference. How many seating arrangements are possible?
5. With indistinguishable elements: How many different letter arrangements can be formed using the letters B-A-N-A-N-A?

Examples (1)

Suppose one has a population of size n. What is probability that a particular element will not be in sample of size r?

Solution *Without* Replacement:

$$\frac{(n-1)!}{(n-1-r))!} = \text{No. of samples without particular element.}$$

$$\frac{n!}{(n-r)!} = \text{No. of Samples}$$

$$\begin{aligned} q &= \text{Probability element is not in sample} \\ &= \frac{(n-1)!/(n-(r+1))!}{n!/(n-r)!} \\ &= \frac{(n-1)(n-2)\ldots(n-r)}{n(n-1)(n-2)\ldots(n-r+1)} \\ &= \frac{n-r}{n} = 1 - \frac{r}{n} \\ p &= \text{Probability element in the sample} = 1 - q = r/n \end{aligned}$$

Examples (2)

Suppose one has a population of size n. What is probability that a particular element will not be in sample of size r?

Solution *With* Replacement:

$$\begin{aligned} q &= \frac{(n-1)^r}{n^r} = \left(1 - \frac{1}{n}\right)^r \\ p &= 1 - q = 1 - \left(1 - \frac{1}{n}\right)^r \end{aligned}$$

Examples (3)

Suppose one samples r items with replacement from a total of n. What is probability of no repetitions in the sample?

$$\begin{aligned} q &= \frac{\text{No. of Samples with No Repetition}}{\text{No. of Samples}} = \frac{n!/(n-r)!}{n^r} \\ &= \frac{n \cdot (n-1) \cdot (n-2) \ldots (n-r+1)}{n \;\; n \;\; n \ldots n} \\ &= 1(1 - \frac{1}{n})(1 - \frac{2}{n}) \cdot \cdot (1 - \frac{(r-1)}{n}) \\ &\sim (1 - \frac{1 + 2 + \ldots + r - 1}{n}) \\ &= 1 - \frac{r(r-1)}{2n} \end{aligned}$$

because $1 + 2 + \ldots + r - 1 = \frac{r(r-1)}{2}$

Note: $p =$ Probability of repetitions $= 1 - q \sim \frac{(r-1)r}{2n}$

Examples (4): Birthday Problem

Application - the Birthday Problem. If there is a class of r people, what is the probability that some have birthdays on the same day? $(n = 365)$

$$p \sim \frac{r(r-1)}{2 \cdot 365}$$

Suppose $r = 25$

$$p \sim \frac{25 \cdot 24}{2 \cdot 365} = 0.82$$

Suppose $p = 1/2$. What is value of r?

$$r(r-1)/730 = 1/2$$ Solution is $r \approx \frac{39}{2} \sim 20$

Additional Reading

Additional Reading

See Chapter 1.2 in Casella and Berger.