Review

Wednesday, July 04, 2012
2:59 PM

Section 1.1

Life is random; probabilities rule your life.

Sources of randomness

- ignorance
- laws of nature (quantum physics, thermal noise)
- laziness

Section 1.2

Coins, dice, cards

Section 1.3

Ingredients of Probability Theory

- random experiment
- outcome ω
- sample space Ω
- event A $(\subset \Omega)$

define outcomes of interest from a random experiment

- $P(A)$ = probability of event A

Note: Consider a random experiment and a specific event A.

Ex. Toss two (fair) dice

Let A be the event that the sum is 11.

When the experiment has been performed,

the event A may $\begin{cases} \text{occur} \\ \text{not occur} \end{cases}$

The probability that it occurs is denoted by $P(A)$.

(we will see later that $P(A) = \frac{1}{18}$.)

Q: How to interpret the value of probability?

A: One way is to consider "long-run frequency interpretation."

To do this, perform the experiments n times

count the fraction of time that A occurs among these n times.

"relative frequency"

Law of large numbers (LLN)

As $n \to \infty$, the fraction will converge to $P(A)$.
(relative frequency)

Application: when n is not ∞, but large
the fraction should be close to $P(A)$.

Section 2 Set Theory

- disjoint sets ; pairwise disjoint sets; mutually exclusive sets
- partition
- countable vs uncountable sets

$\downarrow$ "discrete" $\quad$ $\downarrow$ "continuous"

- "countable" includes "finite" and "countably infinite"

Section 3 Classical Probability

- Assumptions: - finite Ω
  - equipossibility

- Formula $P(A) = \dfrac{|A|}{|\Omega|}$ ← Can't be blindly applied. Need to satisfy the assumptions above.

↑ From this formula, we see that finding the size of a set is important. This leads to Section 4.

- Properties : $P(\emptyset) = 0$, $P(\Omega) = 1$, $P(A^c) = 1 - P(A)$

- In many problems, changing how you define the sample space Ω may create equipossibility among the outcomes and therefore allow the application of classical probability theory.

  Ex. 10,000 balls (9,999 black, 1 red)

  Ex. ※ H that turn up in tossing of two coins

  3 possibilities (0, 1, 2) not equally likely

  Equipossibility can be satisfied if we consider 4 outcomes (HH, HT, TH, TT).

Section 4 Combinatorics

↑ study of counting methods

Four principles
- i) Addition ← add ※ parts in each case
- ii) Multiplication ← multiply ※ choices in each step.
- iii) Subtraction ← "at least"
- iv) Division

Ex. Scandal of Arithmetic and the Origin Inspiration of probability theory.

Four Kinds of Counting Problems

ordered (O) / unordered (U) — sampling — w/ replacement / w/o replacement

OSwR : n^r

OSw/oR : $(n)_r = \frac{n!}{\ }$ [illegible] $(n)_n = n!$ ← permutation

OSwR : n^r

OSw/oR : $(n)_r = \frac{n!}{(n-r)!}$ $(n)_n = n!$ ← permutation

USw/oR : $\binom{n}{r}$

USwR : → study/read notes if interested.

Permutation of r types of n objects :

$$\frac{n!}{n_1!\, n_2! \cdots n_r!}$$

Binomial coefficient : $\binom{n}{r} = \frac{n!}{(n-r)!\, r!}$

Event-based Probability theory

Section 5

Three axioms

- P1 : Nonnegativity $P(A) \geq 0$
- P2 : Unit normalization $P(\Omega) = 1$
- P3 : Countable additivity

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

countable disjoint union

$P(A) = 1$

- A is an almost-sure event. (a.s.)
- A occurs with probability 1. (w.p.)
- A is a support of P.

Other properties : $P(\emptyset) = 0$, $0 \leq P(A) \leq 1$

Finite additivity : $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$

When A is countable (finite or countably infinite),

$$P(A) = \sum_{a \in A} P(\{a\})$$

The sum of the probabilities of the

"The sum of the probabilities of the outcomes in the event.

$$\underbrace{P(A^c)}\ = 1 - P(A)$$

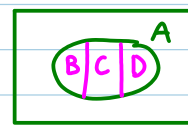
complementary probability

"Venn diagram"-based formulas

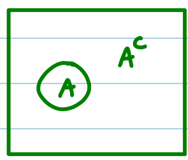
useful when you have a couple of sets

Think of the Ω as having area 1.

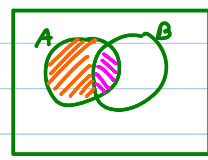
Finding probability $\Rightarrow$ Finding area.



$$P(A) = P(B) + P(C) + P(D)$$



$$P(A) + P(A^c) = 1$$



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

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

$$P(A \cup B) = P(B) + P(A \setminus B)$$

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

Important Inequalities

$$P(A \cup B) \leq P(A) + P(B)$$

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n)$$

$$P(A_1 \cup A_2 \cup \cdots\ ) \leq P(A_1) + P(A_2) + \cdots$$

Each inequality above becomes equality when the sets involved in the union are disjoint.

(We have also checked that under the assumptions of classical probability, we can evaluate the probability of event A by

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

This is nice. We started with three axioms to deal with the general settings of probability calculation. It is good to know that when the situation under consideration is the special case of classical probability, we can use the same old formula.)

Section 6.1 Conditional Probability

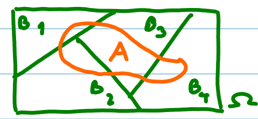
Definition $\quad P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Important properties:

$P(A^c \mid B) = 1 - P(A \mid B)$

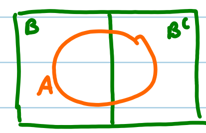
Total Probability: $P(A) = \sum_k P(A \mid B_k) P(B_k)$ when the B_k's partition Ω.

many cases (that can leads to A)



(Special Case: $P(A) = P(A \cap B) + P(A \cap B^c)$

$= P(A \mid B) P(B) + P(A \mid B^c) P(B^c)$



)

Bayes' theorem: Form 1 $\quad P(B \mid A) = P(A \mid B) \dfrac{P(B)}{P(A)}$

Form 2 $\quad P(B_k \mid A) = \dfrac{P(A \mid B_k) P(B_k)}{\sum_i P(A \mid B_i) P(B_i)}$

Remark: "Form 2" is derived from "Form 1" via the total probability theorem.

Section 6.2 Independence

Two Events

Def. $A \perp\!\!\!\perp B$ iff $P(A \cap B) = P(A) P(B)$

Three more equivalent statements

$A^c \perp\!\!\!\perp B$, $A \perp\!\!\!\perp B^c$, $A^c \perp\!\!\!\perp B^c$

$\Updownarrow$

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

$\Updownarrow$

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

When we know that $P(A) \neq 0$ and $P(B) \neq 0$, then $A \perp\!\!\!\perp B$ is also equivalent to

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

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

Three Events

Def. A, B, C are independent iff all four of the following conditions hold

1) $A \perp\!\!\!\perp B$
2) $A \perp\!\!\!\perp C$
3) $B \perp\!\!\!\perp C$

$\Leftrightarrow$ "pairwise independence"

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

See HW for examples where
- conditions (1)-(3) are not enough
- condition (4) alone is not enough

for "independence".

Section 6.3 Bernoulli Trials

Bernoulli Trial : two outcomes < success / failure

(or < A occurs / A does not occur)

(implicitly assume that the trials are independent)

→ n Bernoulli trials ← Repeat the experiment n times

→ Sequence of Bernoulli trials

For n Bernoulli trials with $\overbrace{\text{success probability}}^{\text{for each trial}} = p$,

the probability that there are exactly k successes is given by

$$\binom{n}{k} p^k (1-p)^{n-k}$$

← Note that k can be $0, 1, 2, \ldots, n$

Note that the formula above is derived under the assumption of independence among trials.

Assume this for Bernoulli trials unless stated otherwise.

For n Bernoulli trials with success probability $p = \frac{1}{n}$,

the probability that there is at least one success is given by

$$1-(1-p)^n = 1-\left(1-\frac{1}{n}\right)^n \longrightarrow 1-\frac{1}{e}$$

is given by

$$1-(1-p)^n = 1-\left(1-\frac{1}{n}\right)^n \xrightarrow[n\to\infty]{} 1-\frac{1}{e}$$

Section 7 Random Variables

Mathematically, a random variable is a function that maps the outcomes in Ω to real numbers.

To find the probability of event involving random variable X:

0) Usually, the probability will be expressed in the form

$$P[\text{ some condition(s) on } X].$$

This means finding the probability of the event defined by the set of outcomes which satisfy the given conditions when pluged into $X(\omega)$.

1) First, look at the outcomes ω inside Ω.

For each ω, check whether $X(\omega)$ satisfies the given condition(s).

2) Add the probabilities of the ω which pass the check in the previous step.