Math 21a Biochem

PROBABILITY

Fall 2004

1 Introduction

Text: *Probability* by Jim Pitman, New York: Springer-Verlag c1993. Found on reserve in the Cabot Library.

This is a brief outline of the last part of the course. It is intended as a list of topics and examples rather than a complete set of lecture notes. It should provide a framework that you can fill in with your lecture notes, problem sets, and if needed the text on reserve in Cabot library. The outline is broken up into the topics we have covered. Note that different topics take different amount of lecture time to cover; some less than, some more than one class. Please don't hesitate to contact us if you have questions about the material in the course.

2 Course outline: Introduction through Conditional Probability

2.1 Motivation and Introduction to notation

- The virus game
- Equally likely events
  - Coin tossing $\{HH, HT, TH, TT\}$
  - Throwing fair dice
  - Winning games of chance: example the lottery or rock/paper/scissors
- Law of averages
  - Relative frequency of an event: the ratio measureing how often something occurs in a sequence of observations. (Example number of heads in n coin tosses.)
  - General rule: relative frequencies based on larger numbers of observations fluctuate less than those of small numbers of observations.

  - This is the empirical **law of averages:** the relative frequency of an event based on n trials stabilises as n gets larger and larger (assuming contant conditions).

- Other events
  - Gender of children B/G. (Total probability one.) A family has two children. You know that one of the children is a boy. What is the probability the other child is a boy? $\{BB, BG, GB, GG\}$
  - Lethal Doses (LD50). This is the amount of poison you need to consume to have a 50% chance of dying. (Measured as a percentage of body weight.)
  - Dartboards: you are equally likely to hit each point on a dart board. So, P(hitting region)=Area(region)/Area(dartboard). (See figure.)

- The language of events in probability
  - Experiment
  - Outcome, outcome space
  - Event

- Dictionary between the language of events and sets
  - See figure from
  - Examples above used to illustrate the following
  - Outcome space Ω
  - Event, impossible event
  - Not A
  - Either A or B
  - Both A and B
  - A and B are mutually exclusive

- The probability set function
  - The probability set function is a function that assigns to an event, the real number that reflects the likelihood of that event: denoted $\mathrm{P}(A)$.
  - Properites:
    - It is always non-negative
    - The outcome space has probablility one: $\mathrm{P}(\Omega) = 1$
    - For any event A, $0 \leq \mathrm{P}(A) \leq 1$

- Example: What is the probability you will end up in a particular Harvard House? This is a randomized process. (Think of Quincy vs Dunster house - how many of the available slots are there in each?)
- Defintion of **partitions:** Event B is partitioned into n events $B_1, \ldots, B_n$ if $B = B_1 \cup \ldots \cup B_n$ and the events $B_1, \ldots, B_n$ are mutually exclusive (that is every outcome is in one and only one of the B_i).
  - Non-negative $\mathrm{P}(B) \geq 0$
  - Addition: if $B_1, \ldots, B_n$ is a partition of B, then $\mathrm{P}(B) = \mathrm{P}(B_1) + \ldots + \mathrm{P}(B_n)$.
  - Total one: $\mathrm{P}(\Omega) = 1$
- Use partitions, Venn diagrams and the language of logic to prove
  - Complement rule: $\mathrm{P}(\text{not}A) = \mathrm{P}(A^c) = 1 - \mathrm{P}(A)$
  - Subset/Difference rule: If A is a subset of B, then $\mathrm{P}(A) \leq \mathrm{P}(B)$ and the difference in these probabilities is the probability that B occurs and A does not, that is $\mathrm{P}(B \cap A^c) = \mathrm{P}(B) - \mathrm{P}(A)$.
  - Inclusion-Exclusion rule: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.

2.2 Dependent vs Independent Events

- How to define independent and dependent events?
- Examples independent events
  - Drawing a card from a deck of cards, replacing it, then drawing another
  - Rolling a Dice and tossing a coin or "lucky" dice outcomes verses lottery outcomes
  - Gender and Harvard House selection
  - Midterm grades and Harvard House selection (time/causality)
  - Tides in Sydney Harbour vs traffic crossing the Sydney Harbour Bridge
- Examples dependent events
  - Drawing a card from deck of cards, then drawing another without replacing the first
  - Dying given LD50 for alcohol and dying given LD50 cocaine. What if you've done both?!?
  - Getting TB and getting HIV

- Gender and red-greeen color blindness
- Salary and GPA? How would we know?

- Definition **independent events:** events have no influence on each other (also see conditional probability later)

- Definition **dependent events:** not independent

- Calculating $\mathrm{P}(A \cap B)$
  - For independent events $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$
  - For dependent events it is harder and turns into the multiplication rule below.

- Definition **multiplication rule:** $\mathrm{P}(A \cap B) = \mathrm{P}(A|B)\mathrm{P}(B)$
  - $\mathrm{P}(A|B)$ is the probability of A given B. Note $A|B$ notation: $A|B$ not a set!
  - The intuition for the multiplication rule is an idea of time or causality
  - Think of A as an event determined by some overall outcome which can be thought of as occuring by stages, and B is some event depending just on the first stage. If you think of B as happening before A, then we can think of the muliplication rule as $\mathrm{P}(A \cap B) = \mathrm{P}(B)\mathrm{P}(A|B)$. In words: the chance of B followed by A (or $\mathrm{P}(B \cap A)$) is the chance of B times the chance of A given B.

- Multiplication rule in action: example of two electrical components with failure rates. Note use of tree diagram in solution.

2.3 Conditional Probability

- Motivation: the influence of additional information on the assignment of probabilities

- Examples
  - Given 2 urns, both with different numbers of red and blue balls. Find the probability of drawing a blue ball, given it came from urn 1.
  - P(Contracting malaria from infections bite | sickle cell trait)
  - P(Gettng infected from infectious contact | not immune)
  - Coins again: Find the probability of getting two heads in three tosses of a coin, given the first toss is a head.

- Definition **Conditional Probability**
  - $\mathrm{P}(A|B)$ is the probability that event A has occurred given that event B has occurred $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$

- Note $A|B$ notation - $A|B$ not a set!
- This is not symmetric, that is $\mathrm{P}(A|B) \neq \mathrm{P}(B|A)$

- Examples
  - Picking cards from a deck. Suppose two cards are dealt from a deck of 52 cards. What is the probability that the second card is black?
  - Electric component example illustrating the multiplication rule: Probability of second component working and also probability that exactly one works.

- Rule of Averaged Conditional Probability
  - Recall multiplication rule and partitions
  - Definition **Averaged Conditional Probability** For a partition $B_1 \dots B_n$ of Ω,

    $$\mathrm{P}(A) = \mathrm{P}(A|B_1)\mathrm{P}(B) + \dots \mathrm{P}(A|B_n)\mathrm{P}(B_n)$$

    That is, the overall probability $\mathrm{P}(A)$ is the weighted average of the conditional probabilities $\mathrm{P}(A|B_i)$ with weights $\mathrm{P}(B_i)$.
  - Proof follows from a) definition of partition b) additivity of probability set function c) use of multiplication rule for each term above
  - Venn diagram

- Independent events again.
  - Apply averaged conditional probability to independent events. This just gives the multiplication rule again $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$. Thus $\mathrm{P}(A|B) = \mathrm{P}(A)$.
  - Example: the probability of a girl given heads in a coin toss is just the probability of a girl.

- Bayes' Rule
  - Example in Pittman. Three boxes, Box i contains i white and one black ball. After mixing up the boxes, a ball is chosen at raondom. Which box would you guess if the ball is white and what is your chance of guessing right?
  - Disease testing example. This one worked in detail to illustrate Bayes' rule stated below.
    * P(positive test | having disease)
    * P(negative test | having disease)
    * P(positive test | not having disease)
    * P(negative test | not having disease)

    * If you get a positive result, what is the probability that you actually have the disease?
  - Risk groups
    * Having a relative with breast cancer increases likelihood of getting breast cancer
    * Being intravenous drug user increases risk of HIV
    * Being poor?

- Definition **Bayes' Rule** For a partition $B_1, \ldots, B_n$ of all possible outcomes,

$$\mathrm{P}(B_i|A) = \frac{\mathrm{P}(A|B_i)\mathrm{P}(B_i)}{\mathrm{P}(A|B_1)\mathrm{P}(B_1) + \ldots + \mathrm{P}(A|B_n)\mathrm{P}(B_n)} \quad (i = 1, \ldots, n)$$

  - Follows from definition of conditional probability and the rule of averaged conditional probability
  - Bayes' rule is simply a generalization of this result for a partition of the outcome space into more than two parts

- Example: Monty Hall problem