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 measuring 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 stabilizes as n gets larger and larger (assuming constant 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 $\mathrm{P}(A)$.
  - Properties:
    - It is always non-negative
    - The outcome space has probability 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?)
- Definition 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-green 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 occurring 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 multiplication 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(Getting 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 Pitman. Three boxes, Box i contains i white and one black ball. After mixing up the boxes, a ball is chosen at random. 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

3 Course outline: Random Variables through Distributions

3.1 Discrete Random Variables

- Intuition of Random Variables and Distributions/Probability density functions
- Bernoulli (p) distribution
  - The distribution on $\{0, 1\}$ defined for p between 0 and 1 by 1 has probability p and 0 has probability $1 - p$
- Uniform distribution
  - Equally likely outcomes. For a range of n outcomes, say, $\{1, 2, \ldots, n\}$, then the probability of each is $1/n$
- Binomial Probability and Distribution
  - Take event A and say that we have a success if A occurs and a failure if A does not occur. These have probability p and $q = 1-p$ respectively. Now think about the number of success in n trials, assuming the trials are independent. This is described using the **binomial probability distribution.**
  - The probability of k successes in n trials is $\binom{n}{k} p^k q^{n-k}$
  - $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)...(n-k+1)}{k(k-1)...1}$
  - Examples
    - Coin tossing
    - Throwing one or two dice
    - Birth of a child - gender
    - Disease questions
- **Mode** of Binomial Distribution
- **Mean** of Binomial Distribution (or the expected number of successes) is $\mu = np$

3.2 Introduction to Random Variables

- Definition **Random Variable RV**: Outcome of a numerical measurement of an experiment.
  - **Random:** the outcome of the measurement cannot be determined ahead of time
  - **Variable:** the outcome of the measurement is one of many possible values.

- Examples
  - coin,
  - dice,
  - birth-weight,
  - time to failure of a lightbulb
- The **range or space** of an RV is the set of values it can take
- An RV whose range is a discrete subset of the real numbers (like the whole numbers from 1 to n) is called a **discrete RV**. We'll discuss these first.
- An RV whose range is a continuous subset of the real numbers (like the interval from 1 to n) is called a **continuous RV**. We'll discuss these later.
- Probability Density Function
  - Definition **probability density function or distribution:** the function $f(k)$ gives the probability that random variable X takes on value k. This is written $\mathrm{P}(X = k)$. (Discrete vs continuous pictures.)
  - Examples
    * Number on a die
    * Number of heads in four coin tosses
  - Properties (follow from probability set function)
    * Always non-negative
    * Sum of the value of the function over all elements of outcome space is 1.
  - Note: The domain of a PDF can be extended to all real numbers by setting it equal to zero for values the random variable doesn't take.
- Expectation Value, Mean and Variance
  - Definition. The **expectation value or mean** of a random variable is the sum of its possible values weighted by their probabilities. $E(X) = \Sigma_{\text{all}\, x}\, x\mathrm{P}(X = x)$.
  - Examples
    * Number of heads when flipping four coins
    * Betting on various games
    * Examples by picture
  - Definition. The **variance** of RV X, denoted $Var(X)$, is the mean squared deviation of X from its expected value $\mu = E(X)$: $Var(X) = E[X - \mu)^2]$.
    * Note computational formula: $Var(X) = E(X)^2 - [E(X)]^2$.

- Definition. The **standard deviation** of RV X, denoted $SD(X)$, is the square root of the variance of X: $SD(X) = \sqrt{Var(X)}$.
- Intuitively, the mean is an estimate of the value around which you'd expect date on the random variable to cluster. It's usually (but not always) close to the hump in the PDF of a random variable (if it has one). Intuitively the standard deviation is an estimate of the width you'd expect of the cluster of data around the mean.

• Function of an RV

- A random variable X maybe expressed as a function of another random variable W. That is, $X = g(W)$. Here g is a function defined on the range of W and has values in the range of X. So $g(W)$ is again another random variable.
- $X = g(W)$ has its own PDF, which we can work out in terms of the distribution of W. Each $w \in W$ is assigned a unique $x \in X$ by g. (Of course an x might have many w's assigned to it.) Thus $X = x$ is an event such that W has a w such that $g(w) = x$. Thus $\mathrm{P}(X = x) = \mathrm{P}(g(W) = x) = \Sigma_{w:g(w)=x}\mathrm{P}(W = w)$.
  * Examples: X^2, coin-tossing
  * Note: Assume the distribution of X is known. Then the probability of an event determined by $g(X)$ is often found by manipulating the statement of an event. See example.
- Expectation of a Function of X
  * $E[g(X)] \neq g[E(X)]$
  * $E[g(X)] = \Sigma_{\text{all } x} g(x)\mathrm{P}(X = x)$

• Several random variables and joint PDF/distributions.

- Two RV's X and Y in the same setting. Consider the joint outcome (X, Y) as an RV
- (X, Y) has value (x, y) if $X = x$ and $Y = y$.
- Thus event $((X, Y) = (x, y))$ is the intersection of events $X = x$ and $Y = y$ which we denote $(X = x, Y = y)$.
- The distribution of (X, Y) is the joint distribution $\mathrm{P}(x, y) = \mathrm{P}(X = x, Y = y)$ such that $\mathrm{P}(x, y) \geq 0$ and $\Sigma_{\text{all } (x,y)}\mathrm{P}(x, y) = 1$.
- Example: drawing two cards without replacement.
- Probability that X and Y satisfy some condition is the sum of all $\mathrm{P}(x, y)$ over all pairs (x, y) that satisfy the condition.
- Example $X + Y$, $X - Y$, XY, $\min(XY)$, $\max(XY)$.
- Expectation of $X + Y$: $E(X + Y) = E(X) + E(Y)$

- Independent RVs
  - Random variables are independent if their joint PDF is the product of their individual PDFs. That is RV's X and Y are independent if $\mathrm{P}(X = x, Y = y) = \mathrm{P}(X = x)\mathrm{P}(Y = y)$ for all x and y.
  - This is best understood in terms of the relationship to independent events. If X and Y are independent RV's then every event determined by X is independent of every event determined by Y. That is $\mathrm{P}(X \in A, Y \in B) = \mathrm{P}(X \in A)\mathrm{P}(Y \in B)$.
  - Examples: Red die and Blue die
  - If X and Y are independent RV then expectation XY is $E(XY) = [E(X)][E(Y)]$

3.3 Continuous Random Variables

- Examples
  - Random number generator
  - Spinner on board game
  - Time to failure of a light bulb
  - Waiting for a bus
  - Birthweight
- PDF of Continuous RV
  - Definition: A continuous random variable X has probability density $f(x)$ if for each interval $[a, b]$, $\mathrm{P}(a < X < b) = \int_a^b f(x)\,dx$.
  - Note: The probability that a continuous random variable has any given value is zero. BUT the probability of its value lying in an interval can be nonzero. (This is very much like mass. The mass of a point in a solid is zero, but the mass in a tiny box in the solid is nonzero.)
  - Properties
    * Always nonnegative
    * Integral over outcome space is 1
- Expectation value $E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$
- Variance: $Var(X) = E(X^2) - [E(X)]^2$ (where $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$). Standard Deviation: $SD(X) = \sqrt{Var(X)}$.
- Independent continuous RV's
- Continuous Joint Distributions and Densities - back to double and triple integrals! See Stewart Ch 12 Section 5.

3.4 Continuous Distributions

- Uniform distribution
  - Spinner on board game
  - Expectation and variance
- Exponential distribution
  - Time to failure of a lightbulb
  - Radioactive decay
  - Graph of PDF
  - Expectation and variance
  - Graph of PDF with varying means
- Gaussian/Normal Distribution
  - Definition
  - Graph
  - Expectation and variance
  - Graphs with varying mean and variance
  - Examples