**Basic Intro to Probability**

From Jean Gallier’s Discrete Mathematics Ch 6 on Discrete Probability

---

**Why Learn About Probability?**

- Life in the real world has huge aspects of randomness or uncertainty
  - How long will it take you to get to the airport tomorrow?
  - Will you have a job tomorrow?
  - Still, we hope to understand the world and make useful predictions and inferences
  - If you leave your house 2hrs before your airplane departure, will you arrive in time?
- Randomness often follows reliable and useful laws.
- The language of these laws is the language of probability and statistics
- Play lottery for entertainment value only! (Like a movie ticket)

---

**Semi-Rigorous Definition of Probability**

- When discussing probabilities, generally in the context of some (potentially abstract) experiment
- The set of all possible outcomes $\Omega$ of an experiment must be known (could be uncountably infinite, or finite)
- Example experiment: Rolling two dice
  - Sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - $\Omega = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)\}$
  - Each member $\omega \in \Omega$ is called an elementary event
- Example (non-elementary) event: Rolling doubles
  - $B = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$

---

**Independence**

- Given a discrete probability space $(\Omega, P)$, two events $A$ and $B$ are independent if: $P(A \cap B) = P(A)P(B)$
- Two events are dependent if they are not independent.

---

**Finite Probability Space**

- Finite Probability Space: $(\Omega, P)$
  1. A nonempty, finite set $\Omega$ of outcomes or elementary events.
  2. $P: \Omega \rightarrow [0,1]$ a function (probability measure) satisfying ‘Kolomogorov’ axioms:
     - $0 \leq P(\omega) \leq 1$ for all $\omega \in \Omega$
     - $\sum_{\omega \in \Omega} P(\omega) = 1$
     - For any event $A \subseteq \Omega$, $P(A) = \sum_{\omega \in A} P(\omega)$
- Results:
  - $P(\emptyset) = 0$, $P(\Omega) = 1$
  - If $A \subseteq B$ then $P(A) \leq P(B)$
  - Since $\Omega = A \cup B$ and $A \cap B = \emptyset$, we have $P(A) + P(B) = P(\Omega)$
  - $P(A) = 1 - P(B)$
- To be more careful with countably infinite $\Omega$, need third element $P \in \mathbb{R}$ and some more axioms, but were simplifying

---

**Independence Example**

- Example: $\Omega = \{HH, HT, TH, TT\}$
  - Assume $\forall \omega \in \Omega$, $P(\omega) = \frac{1}{4}$
  - Event A: First flip is H: $A = \{HH, HT\}$
  - Event B: Second flip is H: $B = \{HH, TH\}$
  - $P(A\cap B) = P(HH) = \frac{1}{4}$
  - $P(A)P(B) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
- Independent events
  - Event A: First flip is H: $A = \{HH, HT\}$
  - Event B: Contains a T: $B = \{HT, TH, TT\}$
  - $P(A\cap B) = P(HT) = \frac{1}{4}$
  - $P(A)P(B) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
- Non-Independent events
  - Event A: First flip is H: $A = \{HH, HT\}$
  - Event B: Contains a T: $B = \{HT, TH, TT\}$
  - $P(A\cap B) = P(HT) = \frac{1}{4}$
  - $P(A)P(B) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
Random Variables

- Given some Probability Space \((\Omega, P)\), a random variable \(X: \Omega \rightarrow \mathbb{R}\) is a function defined from the sample space to the Reals
- RVs are functions! Never forget! When Max Mintz asks you what a RV is... you say “a function from the sample space to the reals”!
- RVs take on values in an experiment
- For some \(a \in \mathbb{R}\) want to know \(P(X = a)\). What does this mean? We defined probability only over \(\Omega\)
  - \(P(X = a) = P(X^{-1}(a)) = P(\{\omega \in \Omega | X(\omega) = a\})\)
  - \(P(X \leq a) = P(X^{-1}([-\infty, a])) = P(\{\omega \in \Omega | X(\omega) \leq a\})\)

PMF, PDF, CDF

- Probability Mass Function (PMF) or Probability Density Function (PDF) is
  - \(f: \mathbb{R} \rightarrow [0,1]\)
  - \(\forall a \in \mathbb{R}, f_a (\omega) = P(X = a)\)
- Cumulative Distribution Function is:
  - \(F: \mathbb{R} \rightarrow [0,1]\)
  - \(\forall a \in \mathbb{R}, F_a (\omega) = P(X \leq a)\)
  - \(F\) is monotonic nondecreasing: \(\forall x \leq y, F(x) \leq F(y)\)

Standard Distribution Example - Binomial

- Binomial Distribution
- Flip a coin that has probability \(p\) landing heads and \((1-p)\) tails \(n\) times
- Each of these flips is called a Bernoulli Trial
- Assume coin flips are independent
  - \(\Omega = \{0,1\}^n\)
  - \(X(\omega) = \# of heads in \omega\)
  - Probability of a coin flip with \(i\) heads is \(p^i (1-p)^{n-i}\)
  - There are \(\binom{n}{i}\) ways of getting an event with exactly \(i\) heads
  - \(f(\omega) = P(X = i) = \binom{n}{i}p^i (1-p)^{n-i}\)

Conditional Probability

- Given a probability space \((\Omega, P)\)
- For any two events \(A\) and \(B\), if \(P(B) \neq 0\), then we define the conditional probability \(P(A | B)\) that \(A\) occurs given that \(B\) occurs as
  - \(P(A|B) = \frac{P(A \cap B)}{P(B)}\)
- From this follows Chain Rule:
  - \(P(A \land B \land C) = P(A|B \land C)P(B \land C) = P(A|B)P(B)P(C)\)
  - Etc...
Example with Conditional Prob

- A family has two children. What is the probability that both are boys, given at least one is a boy?
- **Try it!**

Cond. Prob Example Cont’d

- \( \Omega = \{ GG, GB, BG, BB \} \)
- \( \forall \omega \in \Omega, P(\omega) = \frac{1}{4} \)
- Event \( B = (GB, BG, BB) \) of having at least one boy
- Event \( A = (BB) \) of having two boys.
- \( P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{3} \)
- Check out Monty Hall Problem in Galil

Conditional Probability Intuition

Loosely, \( \Omega \) is the various possible states of your body
H and F are binary RVs
H = Have headache
F = Have flu
HF = Have both
\[ P(H) = P(H = 1) = \frac{1}{10} \]
\[ P(F) = \frac{1}{40} \]
\[ P(H|F) = \frac{3}{7} \]
\[ P(F \land H) = ? \]
\[ P(F|H)P(H) + P(F|\bar{H})P(\bar{H}) \]

Bayes Law

- Goal: Suppose you know \( P(H|F), P(F|\bar{H}), P(H) \) can we find \( P(F|H) \)?
- Never remember! Never get confused! Always derive.
- \( P(\bar{F}|H)P(H) + P(F|\bar{H})P(\bar{H}) \)
- \( P(F|H) = \frac{P(F|H)P(H)}{P(F)} \)

Bayes Law Cont’d

- \( A = (A \cap B) \cup (A \cap \bar{B}) \)
- \( P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B}) \)
- \( P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} \)

Bayes Law Example Problem

- Doctors apply a medical test for a certain disease
- Let \( S \) be the event that the patient has the disease, and \( + \) and - the events that the test is positive or negative. We know that:
- If the patient has the disease the test is positive in 99% of the cases.
- \( P(+|S) = 0.99 \)
- However, in 2% of the cases a healthy patient tests positive.
- \( P(+|\bar{S}) = 0.02 \)
- Statistical data shows that one person out of 1000 has the disease.
- \( P(S) = 0.001 \)
- What is the probability for a patient with a positive test to be affected by the disease?
- \( P(S|+) = ? \)
- Reminder: \( P(+) = P(+|S)P(S) + P(+|\bar{S})P(\bar{S}) \)