Math 341: Probability
Second Lecture (9/15/09)

Steven J Miller
Williams College

Steven.J.Miller@williams.edu
http://www.williams.edu/go/math/sjmiller/public_html/341/

Bronfman Science Center
Williams College, September 15, 2009

Section 1.2
Events as Sets

Definitions

- Sample Space (Ω): all possible outcomes. Example: toss coin thrice: $\{HHH, \ldots, TTT\}$; toss until get head: $\{H, TH, TTH, \ldots\}$.
- Events: Subsets of sample space Ω. Example: at least 2 of 3 tosses a head: $\{HHT, HTH, THH, HHH\}$.
- Complement: $A^c = \Omega - A$.
- Field:
  - ◇ $A, B \in \mathcal{F}$ then $A \cup B$ and $A \cap B$ in $\mathcal{F}$x.
  - ◇ $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$.
  - ◇ $\varphi \in \mathcal{F}$ (so $\Omega \in \mathcal{F}$).
  - ◇ if also $A_i \in \mathcal{F}$ implies $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$ then a σ-field.

Section 1.3
Probability

Probability Measure

Finitely additive: disjoint union then $\mathbb{P}(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i)$; countably additive if the $\{A_i\}$ pairwise disjoint implies $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Probability Space

A triple $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space if Ω is a sample space with σ-field $\mathcal{F}$ and a probability measure $\mathbb{P}$ satisfying

- $\mathbb{P}(\varphi) = 0$, $\mathbb{P}(\Omega) = 1$.
- $\mathbb{P}$ is countably additive: for a disjoint union, $\mathbb{P}(\cup_{i=1}^{\infty}(A_i)) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Basic Lemmas

Lemma: For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we have

- Law of total probability: $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
- $A \subset B$ implies $\mathbb{P}(A) \leq \mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B - A)$.
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
- $\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \cdots + (-1)^{n+1}\mathbb{P}(A_1 \cap \cdots \cap A_n)$ (Inclusion - Exclusion Principle, do square-free, PNT hard).

Lemma: $A_1 \subset A_2 \subset \cdots$ and $B_1 \supset B_2 \supset \cdots$, then

- If $A = \cup_{i=1}^\infty A_i$ then $\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(\cup_{i=1}^n A_i)$.
- If $B = \cap_{i=1}^\infty B_i$ then $\mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(\cap_{i=1}^n B_i)$.

Clicker Question

What is the probability that a ‘randomly’ chosen integer is square-free?

Clicker Question

What is the probability that a 'randomly' chosen integer is square-free? Specifically, if we choose $n \in \{1, \ldots, N\}$ **uniformly, as** $N \to \infty$ **what is this probability?**

Clicker Question

What is the probability that a 'randomly' chosen integer is square-free? Specifically, if we choose $n \in \{1, \ldots, N\}$ uniformly, as $N \to \infty$ what is this probability?

Note $e \approx 2.7$, $e^2 \approx 7.4$, $e^3 \approx 20$, $\pi \approx 3$, $\pi^2 \approx 10$, $\pi^3 \approx 30$, $\sqrt{2} \approx 1.4$, $\sqrt{3} \approx 1.7$.

- (A) 0%
- (B) 10%
- (C) 45%
- (D) 60%
- (E) 75%
- (F) 90%
- (G) 100%

(H) does not exist

(I) already knew answer

Section 1.4
Conditional Probability

Definition

Conditional Probability

If $\mathbb{P}(B) > 0$ then the conditional probability of A occurring given B, denoted $\mathbb{P}(A|B)$, is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Interpretation through counting:

$$\frac{N(A \cap B)}{N(B)} = \frac{N(A \cap B)/N}{N(B)/N} \to \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Example: roll fair die twice: what is probability of a 7 or an 11 given first roll is 3? Ans: $\frac{1/36}{6/36} = 1/6$ and $\frac{0/36}{6/36} = 0$.

Definitions (cont)

Partition

A family of events $B_1, \ldots, B_n$ is a partition of Ω if the $\{B_i\}$'s are disjoint and $\cup_{i=1}^{n} B_i = \Omega$.

Always explore conditions. Countable union? Uncountable?

Partition Lemmas

Lemma

If $0 < \mathbb{P}(B) < 1$ then for any event A we have

$$\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c).$$

If the $\{B_i\}$ form a pairwise disjoint partition, then

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

Always explore conditions in a theorem!
Only useful if easier to compute $\mathbb{P}(A|B_i)$ and $\mathbb{P}(B_i)$ then $\mathbb{P}(A \cap B_i)$!

Clicker question

Question

A rare disease affects 1 in 10^5 people. A test is developed; if the person has the disease the test indicates positive 99% of the time; if the person does not have the disease then the test shows positive 1% of the time. *If your test comes back positive, what is the (approximate) probability you are infected?*

Clicker question

Question

A rare disease affects 1 in 10^5 people. A test is developed; if the person has the disease the test indicates positive 99% of the time; if the person does not have the disease then the test shows positive 1% of the time. *If your test comes back positive, what is the (approximate) probability you are infected?*

- (a) 1 / 1,000,000 (one in a million).
- (b) 1 / 100,000.
- (c) 1 / 10,000.
- (d) 1 / 1000.
- (e) 1 / 100.
- (f) 1 / 10. (g) remember answer from book

Solution to the Clicker Question

Let $A = \{\text{ill}\}$ and $B = \{+\}$.
$\mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(B)$, so

$$\mathbb{P}(\text{ill}|+)\mathbb{P}(+) = \mathbb{P}(+|\text{ill})\mathbb{P}(\text{ill})$$

.

$$\begin{aligned} \mathbb{P}(\text{ill}|+) &= \frac{\mathbb{P}(+|\text{ill})\mathbb{P}(\text{ill})}{\mathbb{P}(+)} \\ &= \frac{\mathbb{P}(+|\text{ill})\mathbb{P}(\text{ill})}{\mathbb{P}(+|\text{ill})\mathbb{P}(\text{ill}) + \mathbb{P}(+|\text{healthy})\mathbb{P}(\text{healthy})} \\ &= \frac{\frac{99}{100} \cdot 10^{-5}}{\frac{99}{100} \cdot 10^{-5} + \frac{1}{100} \cdot (1 - 10^{-5})} \end{aligned}$$

Comments

Note in this problem conditional probabilities are readily computed.

Note the probability you test positive but are healthy is 1/1011. Note if we have a population of size 10^5 then we expect one person to be sick (and there is essentially a 100% chance he'll start). There are about 10^3 healthy people who test positive. Thus the number of size to healthy should be about $1/1000$.

Question for thought: Is it better to improve the 99/100 or the 1/100?