Lecture 2

MS Theory - I

Fall Semester

Review

Course Overview

C&B Chapters:

1: Probability Theory
2: Transformations / Expectations
3: Common Families of Distributions
4: Multiple Random Variables
5: Properties of a Random Sample
6: Principles of Data Reduction
7: Point Estimation
8: Hypothesis Testing
9: Interval Estimation
10: Asymptotic Evaluations

Foundation — 1st Semester

Gaining Tools

Applying Tools — 2nd Semester

Chapter 1: Probability Theory

§ 1.1 Set Theory

Sample Space

Definition 1.1.1 The set, S, of all possible outcomes of a particular experiment is called the *sample space* for the experiment.

Event

Definition 1.1.2 An *event* is any collection of possible outcomes of an experiment, that is, any subset of S (including S itself).

Operations

Union of events (or sets) A & B: $A \cup B = \{x : x \in A \text{ or } x \in B\}$

Intersection of events A & B: $A \cap B = AB = \{x : x \in A \text{ and } x \in B\}$

Complementation of A: $A^c = \{x : x \notin A\}$

Properties

Theorem 1.1.4 *For any three events, A, B, and C, defined on a sample space S,*

a. Commutativity $A \cup B = B \cup A$,
$A \cap B = B \cap A$;

b. Associativity $A \cup (B \cup C) = (A \cup B) \cup C$,
$A \cap (B \cap C) = (A \cap B) \cap C$;

c. Distributive Laws $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

d. DeMorgan's Laws $(A \cup B)^c = A^c \cap B^c$,
$(A \cap B)^c = A^c \cup B^c$.

Disjoint

Definition 1.1.5 Two events A and B are *disjoint* (or *mutually exclusive*) if $A \cap B = \emptyset$. The events $A_1, A_2, \ldots$ are *pairwise disjoint* (or *mutually exclusive*) if $A_i \cap A_j = \emptyset$ for all $i \neq j$.

Review Continued

Partition

Definition 1.1.6 If $A_1, A_2, \ldots$ are pairwise disjoint and $\cup_{i=1}^{\infty} A_i = S$, then the collection $A_1, A_2, \ldots$ forms a *partition* of S.

§1.2 Basics of Probability Theory

Sigma-algebra (σ-algebra) or Borel field, $\mathcal{B}$.

Definition 1.2.1 A collection of subsets of S is called a *sigma algebra* (or *Borel field*), denoted by $\mathcal{B}$, if it satisfies the following three properties:

a. $\emptyset \in \mathcal{B}$ (the empty set is an element of $\mathcal{B}$).

b. If $A \in \mathcal{B}$, then $A^c \in \mathcal{B}$ ($\mathcal{B}$ is closed under complementation).

c. If $A_1, A_2, \ldots \in \mathcal{B}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{B}$ ($\mathcal{B}$ is closed under countable unions).

σ-algebra
- Set contains $\emptyset$ & S
- closed under complementation (If $A \in \mathcal{B}$ the $A^c \in \mathcal{B}$)
- closed under countable unions (If A and $B \in \mathcal{B}$ then $(A \cup B) \in \mathcal{B}$)
- closed under countable intersections (If A and $B \in \mathcal{B}$ then $(A \cap B) \in \mathcal{B}$)

- For this class, we will assume sets are well behaved.

Axioms of Probability or Kolmogorov's Axioms

Definition 1.2.4 Given a sample space S and an associated sigma algebra $\mathcal{B}$, a *probability function* is a function P with domain $\mathcal{B}$ that satisfies

1. $P(A) \geq 0$ for all $A \in \mathcal{B}$.

2. $P(S) = 1$.

3. If $A_1, A_2, \ldots \in \mathcal{B}$ are pairwise disjoint, then $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Axiom of Finite additivity: If $A \in \mathcal{B}$ and $B \in \mathcal{B}$ are disjoint

$$Pr(A \cup B) = Pr(A) + Pr(B)$$

Common method for "defining a legitimate probability function"

Theorem 1.2.6 *Let $S = \{s_1, \ldots, s_n\}$ be a finite set. Let $\mathcal{B}$ be any sigma algebra of subsets of S. Let $p_1, \ldots, p_n$ be nonnegative numbers that sum to 1. For any $A \in \mathcal{B}$, define $P(A)$ by*

$$P(A) = \sum_{\{i: s_i \in A\}} p_i.$$

(The sum over an empty set is defined to be 0.) Then P is a probability function on $\mathcal{B}$. This remains true if $S = \{s_1, s_2, \ldots\}$ is a countable set.

Lecture 2

MS Theory
3

Review continued

Properties of pdf:

Theorem 1.2.8 *If P is a probability function and A is any set in $\mathcal{B}$, then*

a. $P(\emptyset) = 0$, *where $\emptyset$ is the empty set;*

b. $P(A) \leq 1$;

c. $P(A^c) = 1 - P(A)$.

Theorem 1.2.9 *If P is a probability function and A and B are any sets in $\mathcal{B}$, then*

a. $P(B \cap A^c) = P(B) - P(A \cap B)$;

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

c. *If $A \subset B$, then* $P(A) \leq P(B)$.

Theorem 1.2.11 *If P is a probability function, then*

a. $P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$ *for any partition* $C_1, C_2, \ldots$;

b. $P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$ *for any sets* $A_1, A_2, \ldots$. *(Boole's Inequality)*

Counting

Table 1.2.1. *Number of possible arrangements of size r from n objects*

	Without replacement	With replacement
Ordered	$\frac{n!}{(n-r)!}$	n^r
Unordered	$\binom{n}{r}$	$\binom{n+r-1}{r}$

Sesame Street
Count von Count
(1972 - ...)
wikipedia accessed 8/7/18

Theorem 1.2.14 *If a job consists of k separate tasks, the ith of which can be done in n_i ways, $i = 1, \ldots, k$, then the entire job can be done in $n_1 \times n_2 \times \cdots \times n_k$ ways.*

Definition 1.2.16 For a positive integer n, $n!$ (read n factorial) is the product of all of the positive integers less than or equal to n. That is,

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1.$$

Furthermore, we define $0! = 1$.

Definition 1.2.17 For nonnegative integers n and r, where $n \geq r$, we define the symbol $\binom{n}{r}$, read n *choose* r, as

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}.$$

§1.2.4 Enumerating Outcomes

Assume

- S = sample space is finite
  $S = \{s_1, s_2, \dots, s_N\}$
- All outcomes equally likely
  $P(s_i) = 1/N$
- Then for any event A

$$P(A) = \sum_{s_i \in A} P(\{s_i\}) = \sum_{s_i \in A} \frac{1}{N} = \frac{\#\text{ elements in } A}{\#\text{ elements in } S}$$

sum over all s_i in A

Example 1.2.18 Poker

- 52 cards 4 suits (♦, ♥, ♣, ♠), 13 cards/suit A, 2:10, J, Q, K
- Assume Poker hand of 5 cards
- $\binom{52}{5} = 2{,}598{,}960$ possible hands
  (sample without replace, unordered)
- all hands equally likely
- Calculate probability hand has 4 Aces

- Count number hands with 4 Aces
  Ace spades, Ace clubs, Ace hearts, Ace diamonds + any other card ($n=48$)
  48 different ways

$$Pr(4 \text{ Aces}) = 48 \Big/ \binom{52}{5} < \frac{1}{50{,}000}$$

Example 1.2.19 Sample with replacement
Roll 3 sided die 2 times

unordered:	(1,1)	(2,2)	(3,3)	(1,2)	(1,3)	(2,3)
ordered:	(1,1)	(2,2)	(3,3)	(1,2) (2,1)	(1,3) (3,1)	(2,3) (3,2)
Probability:	1/9	1/9	1/9	2/9	2/9	2/9
events:	1	1	1	2	2	2

(assumes nine outcomes equally likely).

cnet.com accessed 8/7/18
3D printed 3-sided die

§1.3 Conditional Probability and Independence

Update our probabilities if have additional information.

Example 1.3.1 Draw 4 cards and get 4 Aces

- Draw 4 cards from 52 card deck.
- $\binom{52}{4} = 270{,}725$ possible combinations
- only 1 combination has 4 Aces
- $Pr(4 \text{ Aces}) = 1 \Big/ \binom{52}{4} = \frac{1}{270{,}725}$

Updated information

$P(1^{st} \text{ card} = \text{Ace}) = 4/52$
$P(2^{nd} \text{ card Ace} \mid 1^{st} \text{ card Ace}) = 3/51$
$P(3^{rd} \text{ card Ace} \mid 1^{st} \,\&\, 2^{nd} \text{ card Ace}) = \frac{2}{50}$
$P(4^{th} \text{ card Ace} \mid 1^{st}, 2^{nd}, 3^{rd} \text{ Aces}) = 1/49$

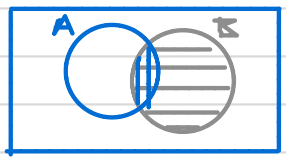
$$\frac{4}{52} \times \frac{3}{51} \times \frac{2}{50} \times \frac{1}{49} = \frac{1}{270{,}725}$$

Notation $P(A \mid B) = P(A \text{ given } B)$

Conditional Prob. continued

Definition 1.3.2 If A and B are events in S, and $P(B)>0$, then the *conditional probability of A given B*, written $P(A|B)$, is

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



$$Pr(A|B)=\frac{P(AB)}{P(B)}=\frac{P(A\cap B)}{P(B)}$$ * Note assuming $Pr(B)>0$

Example 4 Aces, 4 cards continued

Probability of 4 Aces in 4 cards $=\frac{1}{\binom{52}{4}}=\frac{\binom{4}{4} \text{ — 4 Aces choose all 4 } = \frac{4!}{4!0!}=\frac{4!}{4!\cdot 1}=1}{\binom{52}{4}}$

Probability of 3 Aces in 3 cards $=\frac{\binom{4}{3}}{\binom{52}{3}}=\frac{\frac{4!}{3!1!}}{\binom{52}{3}}=\frac{4}{\binom{52}{3}}$

Prob 2 Aces in cards $=\frac{\binom{4}{2}}{\binom{52}{2}}=\frac{\frac{4!}{2!2!}}{\binom{52}{2}}=\frac{6}{\binom{52}{2}}$

Prob 1 Ace in 1 card $=\frac{\binom{4}{1}}{\binom{52}{1}}=\frac{\frac{4!}{1!3!}}{\binom{52}{1}}=\frac{\frac{4!}{3!}}{\binom{52}{1}}=\frac{4}{\binom{52}{1}}$

Prob i Aces in i cards $=\frac{\binom{4}{i}}{\binom{52}{i}}$

$Pr(4$ aces in 4 cards $|$ (i Aces in i cards$))$

$=\frac{P(4 \text{ Aces in 4 cards} \cap (i \text{ Aces in } i \text{ cards}))}{P(i \text{ Aces in } i \text{ cards})}=\frac{P(4 \text{ Aces in 4 cards})}{P(i \text{ Aces in } i \text{ cards})}$

$=\frac{1/\binom{52}{4}}{\binom{4}{i}/\binom{52}{i}}=\frac{\binom{52}{i}}{\binom{52}{4}\binom{4}{i}}$

Example: 3 ~~prisoners~~ cows : Angus, Buttercup and Clover

- 3 cows: Angus, Buttercup and Clover are due to be 'harvested'
- Given the rise of vegetarianism in the U.S. the rancher chooses to spare one animal, whom he chooses at random.
- The rancher tells his farm boy, Wesley, who will be spared, but asks that he keep the identity of the lucky cow secret until they are released.
- Angus tries to get Wesley, who speaks bovinese, to tell him who has been spared, but Wesley refuses.
- Angus then asks which of Buttercup or Clover will be harvested.
- Wesley considers this request and answers that Buttercup will be harvested

~~Prisoner~~ Cow Problem cont.

- Wesley's reasoning: Each animal has a $1/3$ chance of being spared. Since Buttercup or Clover must be harvested, Wesley has given no info about Angus' being spared.
- Angus' reasoning: Given Buttercup will be harvested then either Angus or Clover will be spared. So Angus chance at being spared is $1/2$.

Let A = event Angus spared $\quad P(A) = 1/3$
B = event Buttercup spared $\quad P(B) = 1/3$
C = event Clover spared $\quad P(C) = 1/3$

W = event Wesley says Buttercup harvested.

$$P(A|W) = \frac{P(A \cap W)}{P(W)}$$

Summary

Cow spared	Wesley tell Angus
Angus	Buttercup harvested $\big\}$ each with equal probability
Angus	Clover harvested
Buttercup	Clover harvested
Clover	Buttercup harvested

$$\begin{aligned} P(W) &= P(\text{Wesley says Buttercup harvested}) \\ &= P(\text{Wesley says Buttercup harvested and Angus spared}) \\ &\quad + P(\text{Wesley says Buttercup harvested and Clover spared}) \\ &\quad + P(\text{Wesley says Buttercup harvested and Buttercup spared}). \\ &= (1/3)(1/2) + (1/3)(1) + (1/3)(0) \\ &= 1/6 + 2/6 = 3/6 = 1/2 \end{aligned}$$

$$P(A|W) = \frac{P(A \cap W)}{P(W)} = \frac{Pr(\text{Wesley says Buttercup. and Angus spare})}{Pr(\text{Wesley says Buttercup harvested})}$$

$$= (1/6)/1/2 = 2/6 = 1/3$$

- Angus logic flaw:
  Angus <u>falsely</u> interprets W = Buttercup harvested $= B^c$

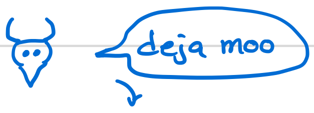
and $P(A|B^c) = \dfrac{P(A \cap B^c)}{P(B^c)} = \dfrac{P(\text{Angus spared and Buttercup harvested})}{P(\text{Buttercup harvested})}$

$$= \frac{P(\text{Angus spared})}{P(\text{Buttercup harvested})} = \frac{1/3}{2/3} = \frac{1}{2}$$

Conditional Probabilities can be 'slippery' and require <u>careful</u> interpretation.

- Follow-up Question
  What is the $P(C|W) = P(\text{Clover spared} \mid \text{Wesley says Buttercup harvested})$

RIP Buttercup....



Definition 1.3.2 If A and B are events in S, and $P(B) > 0$, then the *conditional probability of A given B*, written $P(A|B)$, is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.3.1}$$

$\xRightarrow{\text{implies}}$ $P(A \cap B) = P(A|B)\,P(B)$

by symmetry $P(A \cap B) = P(B|A)\,P(A)$

$\therefore$ (therefore) $P(A|B)\,P(B) = P(B|A)\,P(A)$

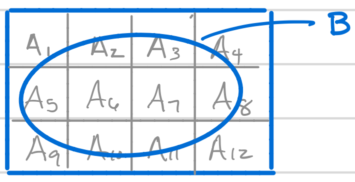
$P(A|B) = P(B|A)\dfrac{P(A)}{P(B)}$ ← Bayes Rule

General form of Bayes Rule

- Recall $P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$ for partition $C_1, C_2, \ldots$ (Thm 1.2.11)

Theorem 1.3.5 (Bayes' Rule) *Let $A_1, A_2, \ldots$ be a partition of the sample space, and let B be any set. Then, for each $i = 1, 2, \ldots$,*

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}. \quad = \frac{P(B|A_i)P(A_i)}{P(B)}$$



Example: HIV Testing (ELISA Screen)

$P(\text{test}+ \mid HIV+) = 0.98$

$P(\text{test}+ \mid \text{HIV Negative}) = .01$

$P(HIV+) = \text{prevalence} = .001$ ← depends on population

$P(\text{HIV Negative}) = 1 - .001 = .999$

Find $P(HIV+ \mid \text{test}+)$ using Baye's Rule.

$$P(HIV+ \mid \text{test}+) = \frac{P(\text{test}+ \mid HIV^+)\,P(HIV^+)}{P(\text{test}+ \mid HIV^+)P(HIV^+) + P(\text{test}+ \mid \text{HIV Neg})\,P(\text{HIV Neg})}$$

$$= (.98)(.001) \big/ [(.98)(.001) + (.01)(.999)] \cong .089$$

* Note: An ELISA test has high sensitivity $P(\text{test}+ \mid HIV^+)$.

- If test positive on ELISA, a follow test is done with high specificity $= P(\text{test}- \mid \text{HIV Neg})$.
- Bigger error in screening test is 'false negative', that is missing an HIV^+ patient.
- 2018 Tests: $P(\text{test}+ \mid HIV^+ \text{ Antibody}) = 1.0\ (.943, 1.0)$ estimate w/ 95% CI.

www.cdc.gov accessed 8/7/18.

Statistical Independence

- If B occurs, it does not effect event A, that is

$$P(A|B) = P(A)$$

If A and B are independent $(A \perp B)$; using Bayes Rule

$$P(B|A) = P(A|B)\frac{P(B)}{P(A)} = \cancel{P(A)}\frac{P(B)}{\cancel{P(A)}} = P(B)$$

so occurence of A, has no effect on B.

- Also from Bayes Rule:

since $P(A\cap B) = P(B|A)P(A)$

if $A \perp B$ $\quad P(A\cap B) = P(AB) = P(A)P(B)$.

Definition 1.3.7 Two events, A and B, are *statistically independent* if

$$P(A\cap B) = P(A)P(B). \tag{1.3.8}$$

Statistical independence also shown via:

$P(A|B) = P(A)$, $P(B>0)$

$P(B|A) = P(B)$, $P(A>0)$

Theorem 1.3.9 *If A and B are independent events, then the following pairs are also independent:*

a. *A and B^c,*

b. *A^c and B,*

c. *A^c and B^c.*

Proof of (a) in book, We'll prove (c)

Theorem 1.1.4 *For any three events, A, B, and C, defined on a sample space S,*

a. Commutativity $\quad A\cup B = B\cup A,$ $\quad A\cap B = B\cap A;$

b. Associativity $\quad A\cup(B\cup C) = (A\cup B)\cup C,$ $\quad A\cap(B\cap C) = (A\cap B)\cap C;$

c. Distributive Laws $\quad A\cap(B\cup C) = (A\cap B)\cup(A\cap C),$ $\quad A\cup(B\cap C) = (A\cup B)\cap(A\cup C);$

d. DeMorgan's Laws $\quad (A\cup B)^c = A^c\cap B^c,$ $\quad (A\cap B)^c = A^c\cup B^c.$

Theorem 1.2.9 *If P is a probability function and A and B are any sets in $\mathcal{B}$, then*

a. $P(B\cap A^c) = P(B) - P(A\cap B);$

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

c. *If $A\subset B$, then $P(A)\le P(B)$.*

Tools

c) Goal:

Show $P(A^cB^c) = P(A^c)P(B^c)$

$$P(A^cB^c) = P((A\cup B)^c) = 1 - P(A\cup B) = 1 - [P(A) + P(B) - P(A\cap B)] = 1 - P(A) - P(B) + P(A)P(B)$$
$$= (1-P(A))(1-P(B))$$
$$= P(A^c)P(B^c) \quad //$$

Quest: Extend independence def'n to ≥ 3 events

Example: Toss 2 six-sided dice; assume outcomes have equal probability $= 1/36$

$$S = \begin{bmatrix} 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \end{bmatrix} \qquad \begin{bmatrix} 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 7 & 8 & 9 & 10 & 11 \\ 7 & 8 & 9 & 10 & 11 & 12 \end{bmatrix} = \text{sums}$$

Define events

$A = \{\text{doubles}\}$ $\qquad P(A) = 6/36 = 1/6$ $\qquad P(A \cap B \cap C) = P(4,4) = 1/36$

$B = \{ 7 \leq \text{sum} \leq 10 \}$ $\qquad P(B) = 18/36 = 1/2$ $\qquad$ doubles, $7 \leq \text{sum} \leq 10$,

$C = \{\text{sum is } 2 \text{ or } 7 \text{ or } 8\}$ $\qquad P(C) = 12/36 = 1/3$ $\qquad$ $\text{sum} \in \{2, 7, 8\}$

$$P(A \cap B \cap C) = P(4,4) = 1/36$$
$$P(A)P(B)P(C) = 1/6 * 1/2 * 1/3 = 1/36$$

If $P(ABC) = P(A)P(B)P(C)$ are A, B, C independent? ?

Is $A \perp B$ $\quad P(A \text{ and } B) = P(\text{doubles and } 7 \leq \text{sum} \leq 10)$
$$= P((4,4), (5,5))$$
$$= 2/36 = 1/18$$
$$\neq P(A)P(B) = 1/6 \cdot 1/2 = 1/12$$

Is $A \perp C$ $\quad P(A \text{ and } C) = P(\text{doubles and sum} = 2, 7 \text{ or } 8)$
$$= P((1,1), (4,4)) = 1/18$$
$$= P(A)P(C) = 1/6 \cdot 1/3$$

Is $B \perp C$ $\quad P(B \text{ and } C) = P(7 \leq \text{sum} \leq 10 \text{ and sum} \in \{2, 7, 8\})$
$$= P((6,1), (5,2), (4,3), (3,4) (2,5), (1,6), (6,2), (5,3), (4,4), (3,5), (2,6)) = 11/36$$
$$\neq P(B)P(C) = (1/2)(1/3) = 1/6$$

Summary $\quad P(ABC) = P(A)P(B)P(C)$
$$P(AB) \neq P(A)P(B)$$
$$P(AC) = P(A)P(C)$$
$$P(BC) \neq P(B)P(C)$$

Conclusion: $P(ABC) = P(A)P(B)P(C)$ not sufficient to show independence.

Must Continue Quest!

Example (another counter example):

Try A, B, C with $A \perp B$, $A \perp C$, $B \perp C$

Will $P(ABC) = P(A)P(B)P(C)$

Let sample space be $3!$ permutations of $1, 2, 3$ and 3 triples of each number

Assume each event is equally likely.

$$S = \begin{bmatrix} 1,1,1 & 2,2,2 & 3,3,3 \\ 1,2,3 & 2,1,3 & 3,2,1 \\ 1,3,2 & 2,3,1 & 3,1,2 \end{bmatrix}$$

$A_i = \{ i^{\text{th}} \text{ place occupied by } 1 \}$

$A_1 = \{(1,1,1), (1,2,3), (1,3,2)\}$ $\qquad P(A_1) = 1/3$

$A_2 = \{(1,1,1), (2,1,3), (3,1,2)\}$ $\qquad P(A_2) = 1/3$

$A_3 = \{(1,1,1), (3,2,1), (2,3,1)\}$ $\qquad P(A_3) = 1/3$

$P(A_1 A_2) = P(1,1,1) = 1/9$

$P(A_1 A_3) = P(1,1,1) = 1/9$

$P(A_2 A_3) = P(1,1,1) = 1/9$

$P(A_1 A_2 A_3) = P(1,1,1) = 1/9$

$$\left.\begin{aligned} P(A_1 A_2) &= P(A_1)P(A_2) = 1/3 \cdot 1/3 = 1/9 \\ P(A_1 A_3) &= P(A_1)P(A_3) = 1/3 \cdot 1/3 = 1/9 \\ P(A_2 A_3) &= P(A_2)P(A_3) = 1/3 \cdot 1/3 = 1/9 \end{aligned}\right\} \text{ok!}$$

<u>BUT!</u> $\quad P(A_1 A_2 A_3) \neq P(A_1) P(A_2) P(A_3) = (1/3)^3 = 1/27$

Quest: Extend independence def'n to ≥ 3 events

Summary

had example where $P(ABC) = P(A)P(B)P(C)$

BUT $P(AB) \neq P(A)P(B)$
$P(BC) \neq P(B)P(C)$ } Not Pairwise $\perp$.

another example: $P(AB) = P(A)P(B)$ $A \perp B$
$P(AC) = P(A)P(C)$ $A \perp C$
$P(BC) = P(B)P(C)$ $B \perp C$

BUT! $P(ABC) \neq P(A)P(B)P(C)$

→ Showing mutual or simultaneous independence of collection of events requires a strong def'n, since

does not imply

$P(ABC) = P(A)P(B)P(C) \not\Rightarrow P(AB) = P(A)P(B)$ and $P(AC) = P(A)P(C)$ and $P(BC) = P(B)P(C)$
$\not\Leftarrow$

End of Quest:

Definition 1.3.12 A collection of events $A_1, \ldots, A_n$ are *mutually independent* if for any subcollection $A_{i_1}, \ldots, A_{i_k}$, we have

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P\left(A_{i_j}\right).$$

Summary to show A, B, C are mutually independent,

Show: $P(AB) = P(A)P(B)$
$P(AC) = P(A)P(C)$
$P(BC) = P(B)P(C)$

AND $P(ABC) = P(A)P(B)P(C)$

Example with independence:

Flip coin 3 times

$S = \boxed{\begin{matrix} HHH & HHT & HTH & THH \\ HTT & THT & TTH & TTT \end{matrix}}$

Assume each sample point is equally probable $= 1/8$

H_i = i^{th} toss is head

$H_1 = \{HHH, HHT, HTH, HTT\}$ $P(H_1) = 4/8 = 1/2$
$H_2 = \{HHH, HHT, THH, THT\}$ $P(H_2) = 4/8 = 1/2$
$H_3 = \{HHH, HTH, THH, TTH\}$ $P(H_3) = 4/8 = 1/2$

$P(H_1 \cap H_2 \cap H_3) = P(H_1H_2H_3) = P(HHH) = 1/8 \stackrel{?}{=} P(H_1)P(H_2)P(H_3) = (1/2)^3$

$P(H_1 \cap H_2) = P(H_1H_2) = P(HHH \text{ or } HHT) = \frac{2}{8} = \frac{1}{4} \stackrel{?}{=} P(H_1)P(H_2) = \frac{1}{2} \cdot \frac{1}{2} = 1/4$

$P(H_1 \cap H_3) = P(H_1H_3) = P(HHH \text{ or } HTH) = \frac{2}{8} = \frac{1}{4} \stackrel{?}{=} P(H_1)P(H_3) = \frac{1}{2} \cdot \frac{1}{2} = 1/4$

$P(H_2 \cap H_3) = P(H_2H_3) = P(HHH \text{ or } THH) = \frac{2}{8} = \frac{1}{4} \stackrel{?}{=} P(H_2)P(H_3) = \frac{1}{2} \cdot \frac{1}{2} = 1/4$

for $k=2$ or $k=3$ for subcollections $A_{i_1}, \ldots, A_{i_k}$

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

Distributions

Bernoulli and Binomial Distributions (Discrete)

Bernoulli(p)

pmf $P(X = x|p) = p^x(1-p)^{1-x}; \quad x = 0, 1; \quad 0 \le p \le 1$

sample space; parameter space

mean and variance $EX = p, \quad \text{Var}\, X = p(1-p)$

mgf $M_X(t) = (1-p) + pe^t$

Flip coin $x = 1$ if head, $x = 0$ if tail

$P(\text{head}) = p \quad P(\text{tail}) = 1-p$

$P(x=1) = p^1(1-p)^0 = p$

$P(x=0) = p^0(1-p)^1 = 1-p$

pmf:

$P(X=x|p) = p^x(1-p)^{1-x} I_{[0,1]}(x)$ indicator

Binomial(n, p)

pmf $P(X = x|n,p) = \binom{n}{x}p^x(1-p)^{n-x}; \quad x = 0, 1, 2, \ldots, n; \quad 0 \le p \le 1$

sample space; parameter space (n known)

mean and variance $EX = np, \quad \text{Var}\, X = np(1-p)$

mgf $M_X(t) = [pe^t + (1-p)]^n$

$p =$ probability success

$1-p =$ probability failure

n bernoulli trials (often n known)

$X \sim \text{bin}(n,p)$

notes Related to Binomial Theorem (Theorem 3.2.2). The *multinomial* distribution (Definition 4.6.2) is a multivariate version of the binomial distribution.

Binomial: Flip coin 3 times; $X = \#$ heads

$X \sim \text{bin}(3, p)$, $n = 3$, $p =$ prob head

pmf:

$P(X=x|n,p) = \binom{3}{x}p^x(1-p)^{3-x} \cdot I_{[0,1,2,3]}(x)$ sample space indicator

$P(X=0) = \binom{3}{0}p^0(1-p)^3 = \frac{3!}{3!\,0!}(1-p)^3 = (1-p)^3$

$P(X=1) = \binom{3}{1}p^1(1-p)^2 = \frac{3!}{2!\,1!}p(1-p)^2 = 3p(1-p)^2$

$P(X=2) = \binom{3}{2}p^2(1-p) = \frac{3!}{1!\,2!}p^2(1-p) = 3p^2(1-p)$

$P(X=3) = \binom{3}{3}p^3(1-p)^0 = \frac{3!}{0!\,3!}p^3(1-p)^0 = p^3$

§1.4 Random Variables

Definition 1.4.1 A *random variable* is a function from a sample space S into the real numbers.

Notation: random variables denoted with uppercase letters; realized value of variable (range) denoted by lower case.

Random Variable X takes on values x.

Flip coin 3 times

$S =$

HHH	HHT	HTH	THH
HTT	THT	TTH	TTT

Define random variable X = number heads in 3 tosses

s	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
X(s)	3	2	2	2	1	1	1	0

Range for $X = \mathcal{X} = \{0,1,2,3\}$

Assuming a fair coin:

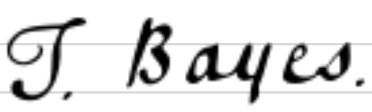
x	0	1	2	3
$P_X(X=x)$	1/8	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Random Variable realized value: 0,1,2 or 3

$P(X=3)$ Random Variable realized value

Thomas Bayes
~1701 - 1761
English statistician, philosopher and Presbyterian minister.

← Portrait of Bayes from 1936 book, but doubtful whether actually Bayes' portrait. Wikipedia, accessed 8/7/18.



← Wikipedia does think this is real signature...