Elements of Probability

1) Experiment with measurable outcome that involves chance
↳ Examples:
  1) watch a Notre Dame v.s. Navy game and count how many yards Andrew Hendrix throws for
  2) Everyone rolls a die and we record the numbers rolled
  3) Toss a coin repeatedly until you get a head and count how many times the coin was tossed

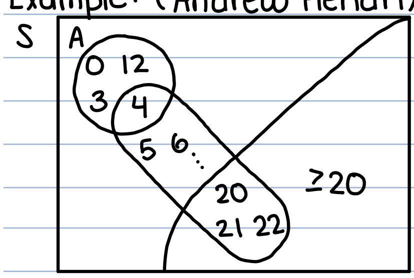
2) Sample space or set of all possible outcomes
↳ 1) For Andrew Hendrix:
  $S=\{0,1,2,3,\ldots\}=\mathbb{Z}_{\geq 0}$ or
  $S=\{0,\frac{1}{2},1,1\frac{1}{2},2,\ldots\}$ or } infinite
  $S=\{0,1,2,\ldots,500,>500\}$ or finite
  $S=\{x: x\in\mathbb{R}_{\geq 0}\}$ infinite
  2) $S=\{(x_1,x_2,x_3,\ldots,x_{52}): 1\leq x_i\leq 6$, for each $i\}$
  3) $S=\{H,TH,TTH,\ldots\}\cup\{TT\ldots$ (never get H)$\}$

3) Events i.e. sets of outcomes
↳ 1) $A=\{0,1,2,3,4\}$ (Andrew Hendrix throws <5 yards)
  $B=\{x: x\geq 20\}$ (for at least 20 yards)
  $C=\{4,5,6,\ldots,20,21\}$ (for at least 4 and at most 21)
  2) $D=\{(x_1,x_2,x_3,\ldots,x_{52})$: at least one of the x_i's is $6\}$
  3) $E=\{TTH,TTTH,\ldots\}$ (eventually get a head but it takes at least 3 tries)
↳ There are two special events:
  i) $\emptyset$ (nothing happens)
  ii) S (something happens)
↳ If event F occurs when we do the experiment, then the outcome is in F
  ↳ Example: If Andrew Hendrix throws for 4 yards, both A and C (and S) have occured but not B

Manipulating Events

Getting new events from old:
i) union ($\cup$)
ii) intersection ($\cap$)
iii) complement (^{c})

Example: (Andrew Hendrix example)



$A\cup B$ is any outcome in one of A, B, or both ($\cup$="or")
$A\cap B=\{4\}$, $A\cap B=\emptyset$ (i.e. they are disjoint) ($\cap$="and")
$A^c=\{5,6,7,\ldots\}=S\setminus A$ (i.e. everything not in A$\}$

Connection between the three: $(A\cup B\cup C)^c=A^c\cap B^c\cap C^c$

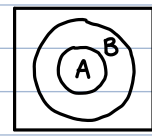
Demorgan's Laws:
1) $(A_1\cup A_2\cup\ldots\cup A_n)^c=A_1^c\cap A_2^c\cap\ldots\cap A_n^c$
2) $(A_1\cap A_2\cap\ldots\cap A_n)^c=A_1^c\cup A_2^c\cup\ldots\cup A_n^c$

A number assigned to an event can measure how likely it is

Rules of Probability

1) $0 \le P(E) \le 1$ for each E
2) $P(S) = 1$
3) For any $A_1, A_2, \ldots$ mutually exclusive, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

consequence of the rules:
1) $P(\emptyset) = 0$
↳ Proof: $S = S \cup \emptyset \cup \emptyset \cup \ldots$, $P(S) = P(S) + P(\emptyset) + P(\emptyset) + \ldots \Rightarrow P(\emptyset) = 0$
2) $P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$ if the A_i are mutually exclusive
↳ called finite additivity
3) $P(A^c) = 1 - P(A)$
↳ Proof: $S = A \cup A^c \Rightarrow P(S) = P(A \cup A^c) = P(A) + P(A^c) = 1$
4) If $A \subseteq B$ (A is a subset of B), i.e. everything in A is in B, then $P(A) \le P(B)$
↳ Proof: $B = A \cup (B \setminus A) \Rightarrow P(B) = P(A) + P(B \setminus A)$, since $0 \le P(B \setminus A) \le 1$ we have $P(A) \le P(B)$



Expirement: Pick a person at random and note their gender and if they are wearing glasses
Data:

	m	f
g	10	8
n	20	12

total: 50

$S = \{fg, fn, mg, mn\}$, 16 possible events (2^4)

Probability Examples

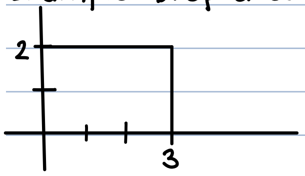
If $S = \{w_1, w_2, \ldots, w_n\}$ and we expect all outcomes to be equally likely, then
$P(S) = 1 = P(\{w_1, \ldots, w_n\}) = P(\{w_1\} \cup \{w_2\} \cup \ldots \cup \{w_n\}) \Rightarrow P(\{w_1\}) = 1/n = P(\{w_i\})$

Example: roll a 20 sided die
$S = \{1, 2, \ldots, 20\}$ all equally likely
P(each individual outcome) = $1/20$
If $E = \{15, 16, 17, 18, 19, 20\}$, $P(E) = P(15) + P(16) + \ldots + P(20) = 6/20$

In general, if all n outcomes are equally likely, then $P(E) = \frac{|E|}{|S|} = \frac{\#\text{ successful outcomes}}{\#\text{ outcomes}}$

Example (Homework): You have 2 mars bars, 1 snickers, and 1 KitKat in a bag. Draw one bar, eat it, draw a second bar, eat it, and note the two that were eaten and in which order
$S = \{(M,M), (M,S), (M,K), (S,M), (S,K), (K,M), (K,S)\}$
E = the first bar you eat is a Kit Kat $= \{(K,M), (K,S)\}$
$P(E) \neq 2/7$ because of the presence of 2 Mars bars
view the mars bars as different: M_1, M_2
There are twelve new outcomes of different probability
$E_2 = \{(K, M_1), (K, M_2), (K, S)\}$, $P(E_2) = 3/12 = 1/4$

Example: Drop a coin randomly on a 2×3 table and observe the point at which the center lands



$S=\{(x,y): 0<x<3, 0<y<2\}$ (infinite)
$U=\text{upper}=\{(x,y): 0<x<3, 1<y<2\}$, $E=\text{edge}=\{(x,y): 0<x<1 \text{ or } 2<x<3, \text{ and } 0<y<2\}$

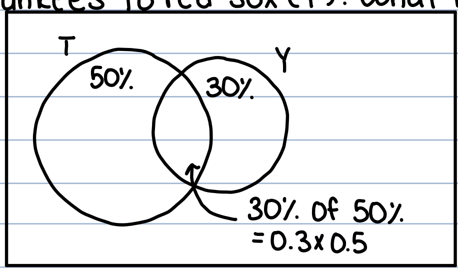
$P(E)=\frac{2}{3}=\frac{\text{Area}(E)}{\text{Area}(S)}=\frac{4}{6}=\frac{2}{3}$, $P(U)=\frac{1}{2}$

(in general $P(\text{any event})=\frac{\text{Area}(A)}{\text{Area}(S)}$)

$C=\text{center }(1\tfrac{1}{2},1)$
$P(C)=\frac{\text{Area}(\{1\tfrac{1}{2},1\})}{6}=\frac{0}{6}=0$

$P(U\cap M)=1/3$
$P(U\cup M)=P(U)+P(M)-P(U\cap M)=5/6$
If A and B are disjoint, $P(A\cup B)=P(A)+P(B)$. Otherwise, $P(A\cup B)=P(A)+P(B)-P(A\cap B)$.
$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$

Independence example: Assume 0.5 of a group is 5'8" or taller (T) and 0.3 of the group prefers yankees to red sox (Y). What is $P(Y\cap T)$? we expect it to be $0.15=0.5\times0.3$



Definition: Events A and B are independent if $P(A\cap B)=P(A)P(B)$

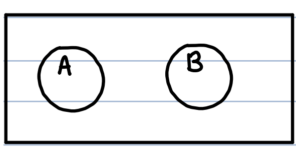
Example: roll 2 dice, red and blue, and observe the two numbers

$$S\begin{cases}(1,1),(1,2),\ldots\\(2,1),(2,2),\ldots\\ \vdots\\(6,1),(6,2),\ldots\end{cases}$$

36 outcomes, all equally likely

A= red dice is 1 or 2, B= blue dice is even
$P(A)=\frac{12}{36}=\frac{1}{3}$, $P(B)=\frac{1}{2}$, $P(A\cap B)=\frac{6}{36}=\frac{1}{6}=\frac{1}{3}\times\frac{1}{2}$
C= both dice show the same number
$P(C)=\frac{1}{6}$, $P(A)P(C)=\frac{1}{3}\times\frac{1}{6}=\frac{1}{18}$, $P(A\cap C)=\frac{2}{36}=\frac{1}{8}$
so A and B, A and C are independent
D= red dice shows 2 or 3
$P(D)=\frac{1}{3}$, $P(A)=\frac{1}{3}$, $P(A\cap D)=\frac{1}{6}\neq\frac{1}{3}\times\frac{1}{3}$ so they are dependent

Mutually Exclusive



are A and B independent? no
$P(A\cap B)=0\neq P(A)P(B)$ unless $P(A)=0$ or $P(B)=0$

Example: toss a coin until heads

$S=\{H, TH, TTH, \dots, TTT\dots\}$

$P(H)=\frac{1}{2}$, $P(TH)=\frac{1}{4}=\frac{1}{2}\frac{1}{2}$, $P(TTH)=\frac{1}{8}$, ..., $P(\overbrace{TT\dots T}^{n}H)=\left(\frac{1}{2}\right)^{n+1}$

A= eventually get a head

$P(A)=P(\{H, TH, TTH, \dots\})=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\dots=1$ (since $a+ax+ax^2+\dots=\frac{a}{1-x}$)

$\Rightarrow P(TTT\dots)=P(A^c)=1-P(A)=0$

↳ This is an example of $P(\sim)=0$ when $\sim\neq\emptyset$

Definition: 3 events A, B, and C are independent if:

i) $P(A\cap B)=P(A)P(B)$

ii) $P(A\cap C)=P(A)P(C)$

iii) $P(B\cap C)=P(B)P(C)$

iv) $P(A\cap B\cap C)=P(A)P(B)P(C)$

Example: David has to form a committee. It must have an odd number of people

$S=\{DAB, DAC, DBC, ABC\}$ all choices equally likely

A= A not on the committee, C= C not on the committee, B= B not on the committee

$P(A)=\frac{1}{2}=P(B)=P(C)$, $P(A\cap B)=\frac{1}{4}=\frac{1}{2}\frac{1}{2}=P(A\cap C)=P(B\cap C)$ so AB, BC, AC are independent

$P(A\cap B\cap C)=1/4\neq P(A)P(B)P(C)=1/8$ so A, B, C are not independent

Knowing A and B occured tells you that C does not

Suppose A and B are independent. What about A and B^c? yes

↳ Proof: $P(A\cap B^c)=P(A)-P(A\cap B)=P(A)-P(A)P(B)=P(A)(1-P(B))=P(A)P(B^c)$

More generally, if $A_1, \dots, A_n$ are independent then $P(A_1\cap A_2^c\cap A_3^c\cap\dots\cap A_{n-1}\cap A_n^c)=P(A_1)P(A_2^c)P(A_3^c)\cdots P(A_{n-1})P(A_n^c)$

Example: Fedeer and Murray play a best of three set match. Each is equally likely to win each set and sets are independent.

$S=\{MM, MFM, FMM, FF, FMF, MFF\}$

$P(\sim)$: 1/4 1/8 1/8 1/4 1/8 1/8

F = Feeder wins, $P(F)=1/2$

E= Murray wins first set, $P(E)=1/2$

Assume you know E occured. $S^{new}=\{MM, MFM, MFF\}$ $P^{new}(F)=1/4$

$1/4\to 1/2$, $1/8\to 1/4$, $1/8\to 1/4$

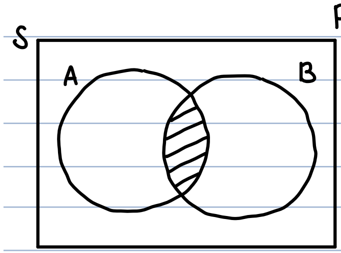
G= the match goes to three sets

$S^{new}=\{MFM, FMM, FMF, MFF\}$

$1/8\to 1/4$, $1/8\to 1/4$, $1/8\to 1/4$, $1/8\to 1/4$

Notation: $P(F|E)$ (probability of F given E), conditional probability

Definition: $P(A|B)=\frac{P(A\cap B)}{P(B)}$ if $P(B)>0$



Example: roll two dice

A = roll 10 or greater $P(A) = 6/36 = 1/6$

B = first is 6

$S^{new} = \{61, 62, 63, 64, 65, 66\}$, $P(A \backslash B) = 1/2$

$$P(A|_B) = \frac{P(A \cap B)}{P(B)} = \frac{3/36}{6/36} = \frac{3}{6} = \frac{1}{2}$$

$$P(B|_A) = \frac{P(B \cap A)}{P(A)} = \frac{3/36}{6/36} = \frac{3}{6} = \frac{1}{2}$$

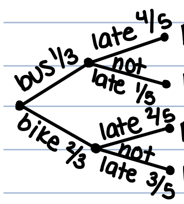
If A and B are independent and $P(A), P(B) > 0$, then $P(A|_B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$ and

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

Other direction: If $P(B|_A) = P(B)$, then $\frac{P(B \cap A)}{P(A)} = P(B) \Rightarrow$ A and B are independent

Multiplication rule: For A and B, $P(A \cap B) = P(A)P(B|_A)$

Example: Every morning, choose between bike and bus. You choose bike $2/3$ of the time and bus $1/3$ of the time. With the bus you're late $4/5$ of the time and with the bike you're late $2/5$ of the time.



$P(\text{bus} \cap \text{late}) = \frac{1}{3} \cdot \frac{4}{5} = \frac{4}{15}$

$P(\text{bus} \cap \text{not late}) = \frac{1}{15}$

$P(\text{bike} \cap \text{late}) = \frac{4}{15}$

$P(\text{bike} \cap \text{not late}) = \frac{6}{15}$

$P(\text{Bus}\,|\,\text{late}) = \frac{4/15}{4/15 + 4/15} = \frac{1}{2}$, $P(\text{late}|_{\text{bus}}) = \frac{4}{5}$

$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_2|A_1)P(A_1)P(A_3|A_1 \cap A_2)$

$$P(A \cap B \cap C) = P(A)P(B|_A)P(C|_{A \cap B}) = P(A)\frac{P(B \cap A)}{P(A)}\frac{P(C \cap A \cap B)}{P(A \cap B)}$$

Example: You have 4 envelopes, one with a prize. 4 people pick one after another. The one with the prize is the winner. Which is the best position to pick in?

A_i = person i wins

$P(A_1) = 1/4$, $P(A_2) = P(A_1^c \cap A_2) = P(A_1^c)P(A_2|_{A_1^c}) = 3/4 \cdot 1/3 = 1/4$

$P(A_3) = P(A_1^c \cap A_2^c \cap A_3) = P(A^c)P(A_2^c|A_1^c)$

$P(A_3|A_1^c \cap A_2^c) = 3/4\,(2/3)(1/2) = 1/4$, $P(A_4) = 1/4 = 1 - P(A_1) - P(A_2) - P(A_3)$

Conclusion: all positions are equally likely to win

$P(\cdot|_A)$ is a probability given any event A, get an assignment of probabilities to each possible event by conditioning on A

$P^{new} = P(E|_A)$

Fact: P^{new} is a valid assignment of probabilities i.e.:

i) $0 \le P^{new}(E) \le 1$

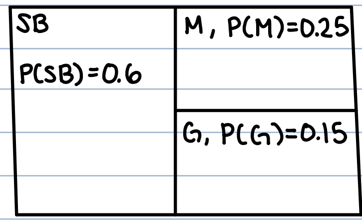
ii) $P^{new}(S) = 1$

iii) $P^{new}(\cup A_i) = \sum P^{new}(A_i)$ if A_i's disjoint

So all of our facts for P remain true for P^{new} i.e. $P(B^c|_A) = 1 - P(B|_A)$

Example: 70% of South Bend residents, 50% of Granger residents, and 60% of Mishwaka residents support Joe Donelly. The South Bend population is 120,000, Mishwaka population is 50,000, and Granger's population is 30,000.

$P(SB) = 60\% = \frac{120{,}000}{200{,}000}$



let D = supports Joe Donelly

$P(SB|D) = \frac{P(SB \cap D)}{P(D)}$

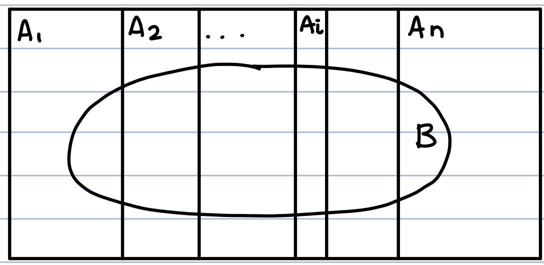
$P(D|SB) = 0.7,\ P(D|M) = 0.6,\ P(D|G) = 0.5$

$P(D) = P(D \cap SB) + P(D \cap M) + P(D \cap G) = P(SB)P(D|SB) + P(M)P(D|M) + P(G)P(D|G)$

$P(SB|D) = \frac{0.6(0.7)}{0.6(0.7) + 0.25(0.6) + 0.15(0.5)} = 0.65$

Bayes Formula

Setup:



$A_1 \cup \ldots \cup A_n = S$, A_i's partition the sample space

Data given: $P(A_1), P(A_2), \ldots, P(A_n)$ and $P(B|A_1), P(B|A_2), \ldots, P(B|A_n)$

want: $P(A_i|B) = \frac{P(B \cap A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \ldots + P(A_n)P(B|A_n)}$

(since $P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_n) = P(A_1)P(B|A_1) + \ldots + P(A_n)P(B|A_n)$ by the law of probability, Adam's law)

Bayes Theorem: If $P(B) > 0$, $A_1, \ldots, A_n$ form a partition of S, and $P(A_i) > 0$, then

$P(A_k|B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$

Example: 75% of a companys customers have "good" credit rating. 25% have "risky" rating. People with good rating are late 10% of the time and people with risky rating are late 50% of the time. Pick a random customer who is late. What is the likelyhood that they have risky credit rate?

we know: $P(R) = 0.25$, $P(G) = 0.75$, $P(L|R) = 0.5$, $P(L|G) = 0.1$

we want: $P(R|L)$

$P(R|L) = \frac{P(R \cap L)}{P(L)} = \frac{P(R)P(L|R)}{P(R)P(L|R) + P(G)P(L|G)} = \frac{(0.25)(0.5)}{(0.25)(0.5) + (0.75)(0.1)} = \frac{0.125}{0.125 + 0.75} = \frac{0.125}{0.2} = 0.625$

Example: 0.005 of the population has condition X. There is a test for X. 98% of the time X is present, the test detects it. 2% of the time X is not present, the test detects. Assume you take the test and its positive. What is the chance that you have X?

$P(X) = 0.005$, $P(X|P) = \frac{P(X \cap P)}{P(P)} = \frac{P(X)P(P|X)}{P(X)P(P|X) + P(X)P(P|X^c)} = \frac{(0.005)(0.98)}{(0.005)(0.98) + (0.995)(0.02)} \approx 0.19758$

Definition: A discrete random variable is a function X from the outcomes to the real numbers

Examples:

1) roll two dice and recall the numbers that come up. suppose we are interested in the sum of the two numbers.

random variable $X: S \to \mathbb{R}$, $X((a,b)) = a+b$, $X((3,5)) = X((4,4)) = 8$

2) note the times at which the first 3 people arrive in class. The typical outcome is $(1.46, 1.48, 1.4845)$. If I'm interested in the difference between the first and second arrival so $X(a,b,c) = b-a$

Random Variables

Examples:

1) The expirement is to go to the cinema