Sample Spaces

- **Sample space**, S, is set of all possible outcomes of an experiment
  - Coin flip: S = {Head, Tails}
  - Flipping two coins: S = {(H, H), (H, T), (T, H), (T, T)}
  - Roll of 6-sided die: S = {1, 2, 3, 4, 5, 6}
  - # emails in a day: S = $\{x \mid x \in \mathbf{Z}, x \geq 0\}$ (non-neg. ints)
  - YouTube hrs. in day: S = $\{x \mid x \in \mathbf{R}, 0 \leq x \leq 24\}$

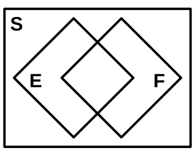
Events

- **Event**, E, is some subset of S ($E \subseteq S$)
  - Coin flip is heads: E = {Head}
  - ≥ 1 head on 2 coin flips: E = {(H, H), (H, T), (T, H)}
  - Roll of die is 3 or less: E = {1, 2, 3}
  - # emails in a day ≤ 20: E = $\{x \mid x \in \mathbf{Z}, 0 \leq x \leq 20\}$
  - Wasted day (>5 YT hrs.): E = $\{x \mid x \in \mathbf{R}, x > 5\}$

Note: When Ross uses: $\subset$, he really means: $\subseteq$

Set operations on Events

- Say E and F are events in S

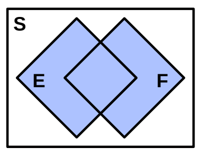


Set operations on Events

- Say E and F are events in S

Event that is in E <u>or</u> F

$E \cup F$



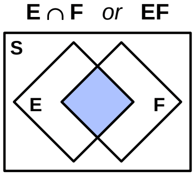
- S = {1, 2, 3, 4, 5, 6} die roll outcome
- E = {1, 2} F = {2, 3} $E \cup F$ = {1, 2, 3}

Set operations on Events

- Say E and F are events in S

Event that is in E <u>and</u> F

$E \cap F$ *or* **EF**



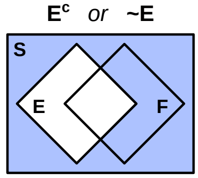
- S = {1, 2, 3, 4, 5, 6} die roll outcome
- E = {1, 2} F = {2, 3} E F = {2}
- **Note:** *mutually exclusive* events means E F = ∅

Set operations on Events

- Say E and F are events in S

Event that is not in E (called complement of E)

E^c *or* **~E**



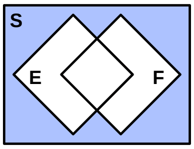
- S = {1, 2, 3, 4, 5, 6} die roll outcome
- E = {1, 2} E^c = {3, 4, 5, 6}

Set operations on Events

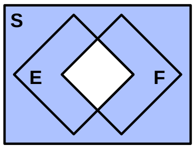
- Say E and F are events in S

DeMorgan's Laws

$(E \cup F)^c = E^c \cap F^c$



$(E \cap F)^c = E^c \cup F^c$



$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c \qquad \left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c$$

Axioms of Probability

- Probability as relative frequency of event:

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

- Axiom 1: $0 \le P(E) \le 1$
- Axiom 2: $P(S) = 1$
- Axiom 3: If E and F mutually exclusive ($E \cap F = \varnothing$), then $P(E) + P(F) = P(E \cup F)$

For any sequence of mutually exclusive events $E_1, E_2, \ldots$

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Implications of Axioms

- $P(E^c) = 1 - P(E)$ $\quad (= P(S) - P(E))$
- If $E \subseteq F$, then $P(E) \le P(F)$
- $P(E \cup F) = P(E) + P(F) - P(EF)$
  - This is just Inclusion-Exclusion Identity for Probability

General form of Inclusion-Exclusion Identity:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{r=1}^{n} (-1)^{(r+1)} \sum_{i_1 < \ldots < i_r} P(E_{i_1} E_{i_2} \ldots E_{i_r})$$

Equally Likely Outcomes

- Some sample spaces have equally likely outcomes
  - Coin flip: S = {Head, Tails}
  - Flipping two coins: S = {(H, H), (H, T), (T, H), (T, T)}
  - Roll of 6-sided die: S = {1, 2, 3, 4, 5, 6}
- P(Each outcome) = $\frac{1}{|S|}$
- In that case, P(E) = $\frac{\text{number of outcomes in E}}{\text{number of outcomes in S}} = \frac{|E|}{|S|}$

Rolling Two Dice

- Roll two 6-sided dice.
  - What is P(sum = 7)?
- S = {(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)}
- E = {(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)}
- P(sum = 7) = |E|/|S| = 6/36 = 1/6

Twinkies and Ding Dongs

- 4 Twinkies and 3 Ding Dongs in a Bag. 3 drawn.
  - What is P(1 Twinkie and 2 Ding Dongs drawn)?
- Ordered:
  - Pick 3 ordered items: |S| = 7 * 6 * 5 = 210
  - Pick Twinkie as either 1st, 2nd, or 3rd item:
    |E| = (4 * 3 * 2) + (3 * 4 * 2) + (3 * 2 * 4) = 72
  - P(1 Twinkie, 2 Ding Dongs) = 72/210 = 12/35
- Unordered:
  - |S| = $\binom{7}{3}$ = 35
  - |E| = $\binom{4}{1}\binom{3}{2}$ = 12
  - P(1 Twinkie, 2 Ding Dongs drawn) = 12/35

Chip Defect Detection

- n chips manufactured, 1 of which is defective.
- k chips randomly selected from n for testing.
  - What is P(defective chip is in k selected chips)?
- $|S| = \binom{n}{k}$
- $|E| = \binom{1}{1}\binom{n-1}{k-1}$
- P(defective chip is in k selected chips)

$$= \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\frac{(n-1)!}{(k-1)!(n-k)!}}{\frac{n!}{k!(n-k)!}} = \frac{k}{n}$$

Any Straight in Poker

- Consider 5 card poker hands.
  - "straight" is 5 consecutive rank cards of any suit
  - What is P(straight)?
  - Note: this is a little different than the textbook
- $|S| = \binom{52}{5}$
- $|E| = 10\binom{4}{1}^5$
- P(straight) = $\frac{10\binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394$

"Official" Straight in Poker

- Consider 5 card poker hands.
  - "straight" is 5 consecutive rank cards of any suit
  - "straight flush" is 5 consecutive rank cards of same suit
  - What is P(straight, but not straight flush)?
- $|S| = \binom{52}{5}$
- $|E| = 10\binom{4}{1}^5 - 10\binom{4}{1}$
- P(straight) = $\frac{10\binom{4}{1}^5 - 10\binom{4}{1}}{\binom{52}{5}} \approx 0.00392$

Card Flipping

- 52 card deck. Cards flipped one at a time.
  - After first ace (of any suit) appears, consider next card
  - Is P(next card = Ace Spades) < P(next card = 2 Clubs)?
  - Initially, might think so, but consider the two cases:
- First note: |S| = 52! (all cards shuffled)
- Case 1: Take Ace Spades out of deck
  - Shuffle left over 51 cards, add Ace Spades after first ace
  - |E| = 51! * 1 (only 1 place Ace Spades can be added)
- Case 2: Do same as case 1, but...
  - Replace "Ace Spades" with "2 Clubs" in description
  - But |E| and |S| are the same as case 1
  - So P(next card = Ace Spade) **=** P(next card = 2 Clubs)

Selecting Programmers

- Say 28% of all students program in Java
  - 7% program in C++
  - 5% program in Java and C++
- What percentage of students do not program in Java or C++
  - Let A = event that a random student programs in Java
  - Let B = event that a random student programs in C++
  - 1 – P(A U B) = 1 – [P(A) + P(B) – P(AB)]
    = 1 – (0.28 + 0.07 – 0.05) = 0.7 $\rightarrow$ 70%
- What percentage programs in C++, but not Java?
  - $P(A^c B) = P(B) - P(AB) = 0.07 - 0.05 = 0.02 \rightarrow 2\%$

Birthdays

- What is the probability that of n people, none share the same birthday (regardless of year)?
  - $|S| = (365)^n$
  - $|E| = (365)(364)...(365 - n + 1)$
  - P(no matching birthdays)
    $= (365)(364)...(365 - n + 1)/(365)^n$
- Interesting values of n
  - n = 23: P(no matching birthdays) < ½ (least such n)
  - n = 75: P(no matching birthdays) < 1/3,000
  - n = 100: P(no matching birthdays) < 1/3,000,000
  - n = 150:
    P(no matching birthdays) < 1/3,000,000,000,000,000

Birthdays

- What is the probability that of *n* other people, none of them share the same birthday as **<u>you</u>**?
  - $|S| = (365)^n$
  - $|E| = (364)^n$
  - P(no birthdays matching yours) = $(364)^n/(365)^n$
- Interesting values of *n*
  - *n* = 23: P(no matching birthdays) ≈ 0.9388
  - *n* = 190: P(no matching birthdays) ≈ 0.5938
    - Anyone born on May 10th?
    - Is today anyone's birthday?
  - *n* = 253: P(no matching birthdays) ≈ 0.4995
    - Least such *n* for which P(no matching birthdays) < ½
  - Why are these probabilities much higher then before?