The Tragedy of Conditional Probability

Thanks xkcd!  
http://xkcd.com/795/
Independence

- Two events E and F are called **independent** if:
  $$P(\text{EF}) = P(E) \times P(F)$$
  Or, equivalently: $P(E \mid F) = P(E)$

- Otherwise, they are called **dependent** events

- Three events E, F, and G independent if:
  $$P(\text{EFG}) = P(E) \times P(F) \times P(G), \text{ and}$$
  $$P(\text{EF}) = P(E) \times P(F), \text{ and}$$
  $$P(\text{EG}) = P(E) \times P(G), \text{ and}$$
  $$P(\text{FG}) = P(F) \times P(G)$$
It Always Comes Back to Dice

- Roll two 6-sided dice, yielding values $D_1$ and $D_2$
  - Let $E$ be event: $D_1 = 1$
  - Let $F$ be event: $D_2 = 1$
- What is $P(E)$, $P(F)$, and $P(EF)$?
  - $P(E) = 1/6$, $P(F) = 1/6$, $P(EF) = 1/36$
  - $P(EF) = P(E)P(F) \rightarrow E$ and $F$ independent
- Let $G$ be event: $D_1 + D_2 = 5$ \{(1, 4), (2, 3), (3, 2), (4, 1)\}
- What is $P(E)$, $P(G)$, and $P(EG)$?
  - $P(E) = 1/6$, $P(G) = 4/36 = 1/9$, $P(EG) = 1/36$
  - $P(EG) \neq P(E)P(G) \rightarrow E$ and $G$ dependent
Two Dice

• Roll two 6-sided dice, yielding values $D_1$ and $D_2$
  ▪ Let $E$ be event: $D_1 = 1$
  ▪ Let $F$ be event: $D_2 = 6$
  ▪ Are $E$ and $F$ independent? Yes!

• Let $G$ be event: $D_1 + D_2 = 7$
  ▪ Are $E$ and $G$ independent? Yes!
  ▪ $P(E) = 1/6$, $P(G) = 1/6$, $P(E \text{ and } G) = 1/36$ [roll (1, 6)]
  ▪ Are $F$ and $G$ independent? Yes!
  ▪ $P(F) = 1/6$, $P(G) = 1/6$, $P(F \text{ and } G) = 1/36$ [roll (1, 6)]
  ▪ Are $E$, $F$ and $G$ independent? No!
  ▪ $P(E \text{ and } F \text{ and } G) = 1/36 \neq 1/216 = (1/6)(1/6)(1/6)$
Sending Messages Through a Network

- Consider the following simplified network:

  - $n$ independent paths, each with probability $p$ of working
  - $E$ = working path from A to B exists. What is $P(E)$?
  - What if $p = 0.2$ and I have 5 paths?

- Solution:
  - $P(E) = 1 - P($all paths fail$) = 1 - (1 - p)^n$
  - If $p = 0.2$ and $n = 5$, $P(E) \approx 0.67$
  - If $p = 0.2$ and $n = 15$, $P(E) \approx 0.96$
Coin Flips

- Say a coin comes up heads with probability $p$
  - Each coin flip is an independent trial

- $P(n \text{ heads on } n \text{ coin flips}) = p^n$
- $P(n \text{ tails on } n \text{ coin flips}) = (1 - p)^n$

- $P(\text{first } k \text{ heads, then } n - k \text{ tails}) = p^k (1 - p)^{n-k}$

- $P(\text{exactly } k \text{ heads on } n \text{ coin flips}) = \binom{n}{k} p^k (1 - p)^{n-k}$
Random Variable

• A **Random Variable** is a real-valued function defined on a sample space

• Example:
  - 3 fair coins are flipped.
  - $Y = \text{number of “heads” on 3 coins}$
  - $Y$ is a random variable
  - $P(Y = 0) = \frac{1}{8}$  (T, T, T)
  - $P(Y = 1) = \frac{3}{8}$  (H, T, T), (T, H, T), (T, T, H)
  - $P(Y = 2) = \frac{3}{8}$  (H, H, T), (H, T, H), (T, H, H)
  - $P(Y = 3) = \frac{1}{8}$  (H, H, H)
  - $P(Y \geq 4) = 0$
Simple Game

- Urn has 11 balls (3 blue, 3 red, 5 black)
  - 3 balls drawn. +$1 for blue, -$1 for red, $0 for black
  - $Y$ = total winnings

- $P(Y = 0) = \frac{\binom{5}{3} + \binom{3}{1} \binom{3}{1} \binom{5}{1}}{\binom{11}{3}} = \frac{55}{165}$
- $P(Y = 1) = \frac{\binom{3}{1} \binom{5}{2} + \binom{3}{2} \binom{3}{1}}{\binom{11}{3}} = \frac{39}{165} = P(Y = -1)$
- $P(Y = 2) = \frac{\binom{3}{2} \binom{5}{1}}{\binom{11}{3}} = \frac{15}{165} = P(Y = -2)$
- $P(Y = 3) = \frac{\binom{3}{3}}{\binom{11}{3}} = \frac{1}{165} = P(Y = -3)$
Probability Mass Functions

• Let $X$ be a **discrete** random variable
  - if it has *countably* many values (like the integers)

• Probability Mass Function (PMF) of a discrete random variable is:
  \[ p(a) = P(X = a) \]

• Since \( \sum_{i=1}^{\infty} p(x_i) = 1 \), it follows that:
  \[ P(X = a) = \begin{cases} 
  p(x_i) \geq 0 & \text{for } i = 1, 2, \ldots \\
  p(x) = 0 & \text{otherwise}
\end{cases} \]

where $X$ can assume values $x_1, x_2, x_3, \ldots$
PMF For a Single 6-Sided Die

$X = \text{outcome of roll}$

$p(x) = \frac{1}{6}$
PMF For a Roll of Two 6-Sided Dice

\[ X = \text{total rolled} \]

\[ p(x) \]

\[ \begin{align*}
6/36 & \\
5/36 & \\
4/36 & \\
3/36 & \\
2/36 & \\
1/36 & \\
\end{align*} \]
Expected Value

• The Expected Values for a discrete random variable $X$ is defined as:

$$E[X] = \sum_{x:p(x) > 0} x \ p(x)$$

  ▪ Note: sum over all values of $x$ that have $p(x) > 0$

• $E[X]$ is often denoted as $\mu$

• Expected value also called: *Mean, Expectation, Weighted Average, Center of Mass, 1st Moment*
Expected Value Examples

\[ E[X] = \sum_{x} x \cdot p(x) \]

• Roll a 6-Sided Die. \( X \) is outcome of roll
  
  \[
  P(X=1) = P(X=2) = P(X=3) = P(X=4) = P(X=5) = P(X=6) = \frac{1}{6}
  \]
  
  \[
  E[X] = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{7}{2}
  \]

• Flip fair coin. Win $5 if “heads”, lose $2 if “tails”.
  
  Let \( Z \) = winnings, so \( E[Z] \) = “average winnings”
  
  \[
  E[Z] = 5 \cdot \frac{1}{2} + (-2) \cdot \frac{1}{2} = 2.5 - 1.0 = 1.5
  \]
Lying With Statistics

“There are three kinds of lies: lies, damned lies, and statistics”

– Mark Twain

- School has 3 classes with 5, 10 and 150 students
- Randomly choose a class with equal probability
- X = size of chosen class
- What is E[X]?
  - E[X] = 5 (1/3) + 10 (1/3) + 150 (1/3)
    = 165/3 = 55
Lying With Statistics

“There are three kinds of lies: lies, damned lies, and statistics”

– Mark Twain

• School has 3 classes with 5, 10 and 150 students
• Randomly choose a student with equal probability
• Y = size of class that student is in
• What is $E[Y]$?
  
  \[ E[Y] = 5 \left( \frac{5}{165} \right) + 10 \left( \frac{10}{165} \right) + 150 \left( \frac{150}{165} \right) \]
  
  \[ = \frac{22635}{165} \approx 137 \]

• Note: $E[Y]$ is students’ perception of class size
  
  • But $E[X]$ is what is usually reported by schools!
Linearity of Expectation

• Linearity:

\[ E[aX + b] = aE[X] + b \]

- Consider \( X = \) 6-sided die roll, \( Y = 2X - 1 \).
- \( E[X] = 3.5 \quad E[Y] = 6 \)
The St. Petersburg Paradox

• Game set-up
  - We have a fair coin (come up “heads” with $p = 0.5$)
  - Let $n =$ number of coin flips (“heads”) before first “tails”
  - You win $2^n$

• How much would you pay to play?

• Solution
  - Let $X =$ your winnings. Claim: $E[X] = \infty$
  
  Proof: $E[X] = \left(\frac{1}{2}\right)^1 2^0 + \left(\frac{1}{2}\right)^2 2^1 + \left(\frac{1}{2}\right)^3 2^2 + ... = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{i+1} 2^i$
  
  $$= \sum_{i=0}^{\infty} \frac{1}{2} = \infty$$
  
  - I’ll let you play for $1$ million... but just once! Takers?
Reminder of Geometric Series

- Geometric series: \( x^0 + x^1 + x^2 + x^3 + \ldots + x^n = \sum_{i=0}^{n} x^i \)

- A handy formula:

\[
\sum_{i=0}^{n} x^i = \frac{1 - x^{n+1}}{1 - x}
\]

- As \( n \to \infty \), and \( |x| < 1 \), then

\[
\sum_{i=0}^{n} x^i = \frac{1 - x^{n+1}}{1 - x} \to \frac{1}{1 - x}
\]
Breaking Vegas

• Consider even money bet (e.g., bet “Red” in roulette)
  ▪ $p = 18/38$ you win bet, otherwise $(1 - p)$ you lose bet
  ▪ Consider this method for how to determine our bets:
    1. $Y = $1
    2. Bet $Y$
    3. If Win then STOP
    4. If Loss then $Y = 2 * Y$, go to Step 2
  ▪ Let $Z =$ winnings upon stopping
  ▪ Claim: $E[Z] = 1$
    □ Here’s the proof in case you don’t believe me:
    \[
    E[Z] = \sum_{i=0}^{\infty} \left( \frac{20}{38} \right)^i \left( \frac{18}{38} \right)^i \left( 2^i - \sum_{j=0}^{i-1} 2^j \right) = \left( \frac{18}{38} \right) \sum_{i=0}^{\infty} \left( \frac{20}{38} \right)^i = \left( \frac{18}{38} \right) \frac{1}{1 - \frac{20}{38}} = 1
    \]
  ▪ Expected winnings $\geq 0$. Repeat process infinitely!
Vegas Breaks You

• Why doesn’t everyone do this?
  ▪ Real games have maximum bet amounts
  ▪ You have finite money
    ○ Not able to keep doubling bet beyond certain point
  ▪ Casinos can kick you out
• But, if you had:
  ▪ No betting limits, and
  ▪ Infinite money, and
  ▪ Could play as often as you want...
• Then, go for it!
  ▪ And tell me which planet you are living on
Variance

- Consider the following 3 distributions (PMFs)

- All have the same expected value, $E[X] = 3$
- But “spread” in distributions is different
- Variance = a formal quantification of “spread”
Variance

- If $X$ is a random variable with mean $\mu$ then the **variance** of $X$, denoted $\text{Var}(X)$, is:
  \[ \text{Var}(X) = E[(X - \mu)^2] \]

- Note: $\text{Var}(X) \geq 0$

- Often computed as: $\text{Var}(X) = E[X^2] - (E[X])^2$

- Also known as the **2nd Central Moment**, or **square of the Standard Deviation**
Variance of 6 Sided Die

- Let $X =$ value on roll of 6 sided die
- Recall that $E[X] = \frac{7}{2}$
- Compute $E[X^2]$

\[
E[X^2] = (1^2)\frac{1}{6} + (2^2)\frac{1}{6} + (3^2)\frac{1}{6} + (4^2)\frac{1}{6} + (5^2)\frac{1}{6} + (6^2)\frac{1}{6} = \frac{91}{6}
\]

\[
\text{Var}(X) = E[X^2] - (E[X])^2
\]

\[
= \frac{91}{6} - \left( \frac{7}{2} \right)^2 = \frac{35}{12}
\]
Jacob Bernoulli

- Jacob Bernoulli (1654-1705), also known as "James", was a Swiss mathematician
- One of many mathematicians in Bernoulli family
- The Bernoulli Random Variable is named for him
- He is my academic great-grandfather
Bernoulli Random Variable

- Experiment results in “Success” or “Failure”
  - $X$ is random indicator variable ($1 = \text{success}, \ 0 = \text{failure}$)
  - $P(X = 1) = p(1) = p \quad P(X = 0) = p(0) = 1 - p$
  - $X$ is a Bernoulli Random Variable: $X \sim \text{Ber}(p)$
  - $E[X] = p$
  - $\text{Var}(X) = p(1 - p)$

- Examples
  - coin flip
  - winning the lottery ($p$ would be very small)
Binomial Random Variable

- Consider $n$ independent trials of Ber($p$) rand. var.
  - $X$ is number of successes in $n$ trials
  - $X$ is a **Binomial** Random Variable: $X \sim \text{Bin}(n, p)$
    
    $$P(X = i) = p(i) = \binom{n}{i} p^i (1 - p)^{n-i} \quad i = 0, 1, ..., n$$

  - $E[X] = np$
  - $\text{Var}(X) = np(1 - p)$

- **Examples**
  - # of heads in $n$ coin flips
Three Coin Flips

- Three fair ("heads" with \( p = 0.5 \)) coins are flipped
  - \( X \) is number of heads
  - \( X \sim \text{Bin}(3, 0.5) \)

\[
P(X = 0) = \binom{3}{0} p^0 (1 - p)^3 = \frac{1}{8}
\]
\[
P(X = 1) = \binom{3}{1} p^1 (1 - p)^2 = \frac{3}{8}
\]
\[
P(X = 2) = \binom{3}{2} p^2 (1 - p)^1 = \frac{3}{8}
\]
\[
P(X = 3) = \binom{3}{3} p^3 (1 - p)^0 = \frac{1}{8}
\]
PMF for $X \sim \text{Bin}(10, 0.5)$
PMF for $X \sim \text{Bin}(10, 0.3)$
Genetic Inheritance

- Person has 2 genes for trait (eye color)
  - Child receives 1 gene (equally likely) from each parent
  - Child has brown eyes if either (or both) genes brown
  - Child only has blue eyes if both genes blue
  - Brown is “dominant” (d), Blue is “recessive” (r)
  - Parents each have 1 brown and 1 blue gene
- 4 children, what is $P(3 \text{ children with brown eyes})$?
  - Child has blue eyes: $p = \left(\frac{1}{2}\right) \left(\frac{1}{2}\right) = \frac{1}{4}$ (2 blue genes)
  - $P(\text{child has brown eyes}) = 1 - \left(\frac{1}{4}\right) = 0.75$
  - $X = \# \text{ of children with brown eyes}$. $X \sim \text{Bin}(4, 0.75)$

$$P(X = 3) = \binom{4}{3} (0.75)^3 (0.25)^1 \approx 0.4219$$