Discrete Probability IV

COMPSCI 230 — Discrete Math

April 13, 2017
Outline

1. Random Variables and Probability Distributions
2. Expectation
3. Repeated Bernoulli Trials
   - Distributions Related to Bernoulli Trials
From Outcomes to Numbers

• We often associate numbers to outcomes
• Example: Coin flip, bet on $H$
• Win $W = g = \$10$ if head, win $W = \ell = -\$10$ if tail
• How much do I win on average if $p = P(H) = 0.6$?
• Let the answer be $a$
• Approximate frequentist interpretation: If I play the game 1000 times, then I win close to $1000a$
• However, we play once
• We need to tie $g$ to $H$ and $\ell$ to $T$
Random Variables

- We need to tie $g$ to $H$ and $\ell$ to $T$
- $H, T$ are in the sample space $\mathcal{S}$
- $g, \ell$ are in $\mathbb{R}$
- Any function $W = f(O) : \mathcal{S} \to \mathbb{R}$ is called a random variable
- *It is neither random, nor a variable*
- It is a deterministic function of the outcome
- If you just look at the image ($W = g$ or $W = \ell$) as a variable then its value varies randomly
- Hence the (arguably confusing) name
- In the example, 

$$W : \{H, T\} \to \mathbb{R}$$

is defined as

$$W(H) = g \quad \text{and} \quad W(T) = \ell$$
Probability Distribution

- A probability function defined on a random variable is called a **probability distribution**:

\[
P(X = x) = \sum_{O \in \mathbb{S} : X(O) = x} P(O)
\]

- \(P(X = x_a) = P(O_1)\), \(P(X = x_b) = P(O_2) + P(O_3)\)

- Example 1: Roll of one fair die, \(x = X(O) = (O \mod 2)\)
  - \(X(2) = X(4) = X(6) = 0\)
  - \(X(1) = X(3) = X(5) = 1\)
  - \(P(X = 0) = P(X = 1) = 1/2\)

- Example 2: Money won when betting on \(H\) in a coin flip,
  \[
  W(O) = \begin{cases} 
  g & \text{for } O = H \\
  \ell & \text{for } O = T 
  \end{cases}
  \]
  - \(P(W = g) = 0.6\)
  - \(P(W = \ell) = 0.4\)
Clicker Test

Computer Science is ... (Not Graded)
A: boring
B: dry
C: useful
D: interesting
E: exciting
Uniform Distribution

The expected value for the outcome of the roll of a fair die is ...

Pick one (Graded)

A: 1/6
B: 3
C: 3.5
D: 4
E: 6
Binomial Distribution

The probability that there are \( k \) successes in a binomial experiment with \( n \) trials and success probability \( p \) is ...

**Pick one**

A: \( p^k \)

B: \( p^k (1 - p)^{n-k} \)

C: \( \binom{n}{k} \)

D: \( \binom{n}{k} p^k (1 - p)^{n-k} \)

E: \( \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} \)
Geometric Distribution

The expected number of trials required for the first success in a Bernoulli trial with success probability $p$ is ...

Pick one (Graded)

A: $n$
B: $1/p$
C: $(1 - p)^{n-1} p$
D: $p$
E: $p/(1 - p)$
Expectation

• I win $W = g = \$10$ with probability $P(H) = p = 0.6$
• I win $W = \ell = -\$10$ with probability $P(T) = q = (1 - p) = 0.4$
• My **expected win** is defined as
  \[ a = \mathbb{E}[W] = \mathbb{E}[f(O)] = P(H)f(H) + P(T)f(T) = pf(H) + qf(T) = pg + (1 - p)\ell = 0.6 \cdot 10 + 0.4 \cdot (-10) = \$2 \]
• If I play the game 1000 times, then I win close to 1000$\mathbb{E}[W] = \$2000$
• However, we play once: Expected win is $\$2$
• More generally the **expected value** (or **expectation**) of random variable $W = f(O)$ is
  \[ E[W] = E[f(O)] = \sum_{O \in S} f(O)P(O) = \sum_{x \in f(S)} x \sum_{O \in S \text{ : } f(O) = x} P(O) \]
• Also applies to compound experiments ($O$ is a tuple)
Computing Expectation

\[ E[W] = E[f(O)] = \sum_{O \in S} f(O) P(O) = \sum_{x \in f(S)} x \sum_{O \in S : f(O) = x} P(O) \]

Example:

<table>
<thead>
<tr>
<th>( O )</th>
<th>( f(O) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( O_1 )</td>
<td>2</td>
</tr>
<tr>
<td>( O_2 )</td>
<td>5</td>
</tr>
<tr>
<td>( O_3 )</td>
<td>1</td>
</tr>
<tr>
<td>( O_4 )</td>
<td>2</td>
</tr>
<tr>
<td>( O_5 )</td>
<td>1</td>
</tr>
<tr>
<td>( O_6 )</td>
<td>2</td>
</tr>
</tbody>
</table>

\[
2P(O_1) + 5P(O_2) + 1P(O_3) + 2P(O_4) + 1P(O_5) + 2P(O_6)
= 2[P(O_1) + P(O_4) + P(O_6)] + 1[P(O_3) + P(O_5)] + 5[P(O_2)]
\]
Expectation is Linear

- \( \mathbb{E}[aU + bV] = a\mathbb{E}[U] + b\mathbb{E}[V] \) because

  \[
  \mathbb{E}[aU + bV] = \sum_{O \in S} P(O) [aU(O) + bV(O)]
  = a \sum_{O \in S} P(O) U(O) + b \sum_{O \in S} P(O) V(O)
  = a \mathbb{E}[U] + b \mathbb{E}[V]
  \]

- Example: Upton wins or loses 8 pounds sterling. Valerie wins or loses 10 dollars
- \( \mathbb{E}[U] = 0.6 \cdot 8 - 0.4 \cdot 8 = 1.6 \) pounds
- \( \mathbb{E}[V] = 0.6 \cdot 10 - 0.4 \cdot 10 = 2 \) dollars
- Exchange rate: 1.25 dollars per pound sterling
- Combined win: \( 1.25 \mathbb{E}[U] + \mathbb{E}[V] = 1.25 \cdot 1.6 + 2 = $4 \)
- If the exchange rate changes, we do not need to recompute \( \mathbb{E}[U], \mathbb{E}[V] \)
Bernoulli Trials

- A repeated coin flip is an example of a Bernoulli trial:
  - Two outcomes per repetition
  - Fixed probability $p$ of “success”
  - Repetition outcomes are independent
- This is an unbounded repeated experiment
- Outcome $C = (C_1, C_2, \ldots)$ (Python generator?)
- Sample space $\mathcal{S}^\infty = \mathcal{S} \times \mathcal{S} \times \ldots$ where $\mathcal{S} = \{H, T\}$
- The infinite sequence $C$ is one outcome of the repeated experiment
- So we can define random variables

$$X : \mathcal{S}^\infty \rightarrow \mathbb{R}$$
Binomial Experiment

- Example 1 of random variable on Bernoulli trials: 
  \[ B_n(C) = k \text{ iff there are } k \text{ successes in the first } n \text{ trials} \]  
  [A different random variable for each \( n \)]

- What is \( P_{B_n}(k) \)?

- Choose which \( k \) are successes: \( \binom{n}{k} \)

- These are mutually exclusive choices

- Probability that the chosen \( k \) are successes: \( p^k \) 
  (because of independence)

- Probability that the remaining \( n - k \) are failures: \( (1 - p)^{n-k} \)  
  (ditto)

\[
P_{B_n}(k) = \binom{n}{k} p^k (1 - p)^{n-k}
\]

- Binomial distribution
Mean of the Binomial Distribution

\[
\mathbb{E}[B_n] = \sum_{k=0}^{n} k \ P_{B_n}(k) = \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k}
\]

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

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

\[
\mathbb{E}[B_n] = \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-1-(k-1)}
\]

\[
= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} = np \ (p + 1-p)^{n-1} = np
\]
Expected Number of Trials to First Success

• Example 2 of random variable on Bernoulli trials:

\[ N(C') = n \quad \text{iff the first } H \text{ is in trial } n \in \mathbb{N} \]

• \( N = n \) iff \( n - 1 \) Ts are followed by one \( H \)

• [and what happens after that does not matter]

• Example: \( n = 4 \). \((T, T, T, H, ...)\)

• Because of independence, \( P(N = n) = q^{n-1}p \)

• **Geometric distribution** (no upper bound on \( n \))

• \( \mathbb{E}[N] \): expected number of trials until the first success

• How many times do you need to flip a coin on average until it comes up \( H \)?
Bad Luck

If the probability of $H$ is $p = 0.6$, what is the probability that $H$ never occurs in a Bernoulli trial?

Pick one

(A: 0  B: 0.4  C: 1/0.4  D: 1/0.6  E: 0.6)

$$\lim_{n \to \infty} (1 - p)^n = 0$$
Unbounded Sample Spaces

What is the cardinality of $\mathbb{S}^\infty = \{0, 1\}^\infty$?

Pick one (Not Graded)

A: 2
B: cardinality of the naturals
C: cardinality of the reals
D: depends on the number of trials

So we are technically going beyond "discrete probability"

There will be limits...
Expected Trials to Success

The probability that a coin comes up head is $p = \frac{2}{5}$. What is the expected number of trials needed to see one head outcome?

Pick one  (Not Graded)

- **A**: 1
- **B**: 2
- **C**: 2.5
- **D**: 3
- **E**: 5