LECTURE 5

• **Readings:** Sections 2.1–2.3, start 2.4

Lecture outline

• Review

• Random variables

• Probability mass function (PMF)

• Expectation
Review: Binomial probabilities

- \( n \) independent coin tosses with \( P(H) = p \)

- \( P(\text{a particular sequence}) = p \# \text{heads}(1 - p) \# \text{tails} \)

- \( P(k \text{ Hs}) = \sum_{k-H \text{ seq.}} P(\text{seq.}) \)

\[
= (\# \text{ of } k-H \text{ seqs.}) \cdot p^k(1 - p)^{n-k} \]

\[
= \binom{n}{k} p^k(1 - p)^{n-k} \]
Outcomes are not numbers

Early dice found in Williamsburg, Virginia. Photo by Joe Fudge/Daily Press.
Random variables

- Assignment of a value (number) to each possible outcome
- Mathematically: A real-valued function on $\Omega$
  - range can be discrete or continuous
- In Chapter 2:
  - range is discrete, discrete random variable
  - almost exclusively, range is a subset of the integers
- Notation:
  - random variable $X$
  - numerical value $x$
Visualization of discrete random variables

- Can have several random variables defined on the same sample space
Probability mass function (PMF)

- Also “probability law” or “probability distribution”

- Definition and notation for PMF of $X$:
  \[
p_X(x) = \Pr(X = x) = \Pr(\{\omega \in \Omega \text{ s.t. } X(\omega) = x\}) \quad [\text{more carefully}]
  \]
  - Defined for all values that $X$ can take

Basic properties of any PMF

- For any $x$ where $p_X(x)$ is defined, $p_X(x) \geq 0$

- $\sum_x p_X(x) = 1$
How to compute a PMF $p_X(x)$

- Collect all possible outcomes for which $X$ is equal to $x$
- Add their probabilities
- Repeat for all $x$

**Example:** Two independent rolls of a fair 4-sided die

\[
X = \min(F, S)
\]

<table>
<thead>
<tr>
<th></th>
<th>first roll $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>●</td>
</tr>
<tr>
<td>2</td>
<td>●</td>
</tr>
<tr>
<td>3</td>
<td>●</td>
</tr>
<tr>
<td>4</td>
<td>●</td>
</tr>
</tbody>
</table>

$p_X(x) =$
Deriving the geometric PMF

- Consider sequences of coin tosses ending with the first H
- Let outcomes be the sequences
- Let probabilities be assigned according to $P(H) = p > 0$
- Let $X$ be the length of a sequence (number of tosses until first H)
- Derive the PMF:

  $$p_X(x) = P(X = x) = P(TT \cdots TH \mid x - 1 \text{ Ts})$$
Binomial PMF

- $X$: number of heads in $n$ independent coin tosses
- $P(H) = p$
- Let $n = 4$

\[
p_X(2) = P(\text{HHTT}) + P(\text{HTHT}) + P(\text{HTTH}) + P(\text{THHT}) + P(\text{THTH}) + P(\text{TTHH})
\]

\[
= 6p^2(1 - p)^2
\]

\[
= \binom{4}{2}p^2(1 - p)^2
\]

In general:

\[
p_X(k) = \binom{n}{k}p^k(1 - p)^{n-k}, \quad k = 0, 1, \ldots, n
\]
Expectation

• Definition:

\[ E[X] = \sum_x x p_X(x) \]

• Interpretations:
  – Center of gravity of PMF
  – Average in large number of repetitions of the experiment
    (to be substantiated later in this course)

• Example: Uniform on 0, 1, \ldots, n
Properties of expectations

- Let $X$ be a r.v. and let $Y = g(X)$
  - Hard: $E[Y] = \sum y p_Y(y)$
  - Easy: $E[Y] = \sum g(x) p_X(x)$

- If $\alpha$ and $\beta$ are constants:
  - $E[\alpha] =$
  - $E[\alpha X] =$
  - $E[\alpha X + \beta] =$