Today: Random variables

Motivation: Roll two dice. Prob. that total is 7.

\[ S = \{(a, b) : a, b = 1, \ldots, 6 \} \]
\[ E = \{s = (a, b) : a + b = 7\} \]
\[ = \{s = (a, b) : f(s) = 7\} \]

where \( f(s) = f((a, b)) = a + b. \)

\[ \Rightarrow \] We have just written \( E \) as the level set of a function \( f : S \rightarrow \{2, \ldots, 12\} \)
\[ \Rightarrow E = \{s \in S : f(s) = 7\} \]

\[ \Rightarrow \] All events are level sets of some function.
Definition: A random variable is a function $f: S \to \mathbb{R}$.

Notation: We use letters like $X, Y, Z$... instead of $f_9, h...$

So $X: S \to \mathbb{R}$

We think of $X$ as the random number $X(s)$, where $s$ is the outcome of the experiment.
* a random variable (r.v. for short)
  * with values in a discrete/countable subset of $\mathbb{R}$ is called a discrete r.v.

  e.g. $X: S \mapsto \{1, 2, 3, \ldots\}$

    $Y: S \mapsto \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$

* If $A \subseteq \mathbb{R}$ then

  $X^{-1}(A) = \{ s \in S : X(s) \in A \}$

  is the event that $X(s)$ belongs to $A$. 
We write \( \mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A)) \)

the prob. that
\( X \) belongs to \( A \).

In other words, \( \{X \in A\} = X^{-1}(A) = \{s \in S : X(s) \in A\} \)

e.g. \( \mathbb{P}(X \geq 5) = \mathbb{P}(\{s \in S : X(s) \geq 5\}) \)

* The probability distribution of \( X \) is
the collection of probabilities \( \mathbb{P}(X \in A) \)
for all sets \( A \subset \mathbb{R} \).
Let $a \in \mathbb{R}$. We write

$$\mathbb{P}(X=a) = \mathbb{P}(X \in \{a\}) = \mathbb{P}(\{s \in S : x(s) = a\})$$

This is called the probability mass function of $X$ (p.m.f.).

$$f(a) = \mathbb{P}(X=a), \quad f: \mathbb{R} \to [0,1]$$

Remarks:

(1) If $X$ is a discrete r.v., then the p.m.f. determines the prob. distribution.

$$\forall A \subseteq \mathbb{R} \quad \mathbb{P}(X \in A) = \sum_{a \in A} \mathbb{P}(X=a)$$

$$\{x \in A\} = \bigcup_{a \in A} \{X=a\}$$
(2) The p.m.f satisfies:

\( 0 \leq \mathbb{P}(X = a) \leq 1 \) for any \( a \in \mathbb{R} \)

(3) For discrete r.v.:

\[ \sum_{a \in \mathbb{R}} \mathbb{P}(X = a) = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(S) = 1 \]

\text{normalization} \quad \rightarrow \quad \mathbb{P}(X = a)

\text{important/common random variables:}

(I) Bernoulli r.v. with parameter \( p \in [0, 1] \).

Let \( S \) be a sample space.

Let \( E \) be an event, such that \( \mathbb{P}(E) = p \).
Define $X : S \rightarrow \{0, 1\}$ by

$$X(s) = \begin{cases} 
1 & \text{if } s \in E \\
0 & \text{if } s \notin E
\end{cases}$$

This type of r.v. is called a Bernoulli r.v. with parameter $p$. (We write $\text{Ber}(p)$ for short.)

Its p.m.f. is:

$$P(X = a) = \begin{cases} 
p & \text{if } a = 1 \\
1-p & \text{if } a = 0 \\
0 & \text{otherwise}
\end{cases}$$

The p.m.f. is:

$$\{X=1\} = E, \quad \{X=0\} = E^c, \quad \{X=\emptyset\} = \emptyset$$
(II) Binomial r.v. with parameters $n \in \mathbb{N}$ and $p \in (0,1]$.

Repeat the Bernoulli experiment from (I) $n$ times, independently. This "combined" experiment has sample space:

$$\hat{S} = S^n = \left\{ (s_1, \ldots, s_n) : s_i \in S \text{ for all } i \right\}$$
Recall from last time:

- Functions $X : S \to \mathbb{R}$ are called random variables.
- Discrete r.v. are those with only countably many possible values.
- p.m.f. of $X$ is $f : \mathbb{R} \to [0,1]$ defined by $f(a) = \mathbb{P}(X = a) = \mathbb{P}\left(\{s \in S : X(s) = a\}\right)$
- Bernoulli r.v., $\text{Ber}(p)$
  - a r.v. $X$ taking values in $\{0, 1\}$
  - $p = \mathbb{P}(X = 1)$, $1 - p = \mathbb{P}(X = 0)$

Today: More common r.v. (I) Bernoulli
(II) Binomial r.v. parameters $n \in \mathbb{N}$, $p \in [0,1]$. 
Sample space, $E \subseteq S$ event with $P(E) = p$. 
Repeat experiment $n$ times (independently): 
new sample space $\hat{S} = S^n = \{(s_1, \ldots, s_n) : s_i \in E \forall i\}$ 
Define $X: \hat{S} \to \{0,1, \ldots, n\}$ by 

$$X((s_1, \ldots, s_n)) = \text{"\# of times that } E \text{ occurs"}$$

$$= \text{\# of } i \text{ such that } s_i \in E$$

Such a r.v. $X$ is called a Binomial r.v. with params $n$ and $p$, $\text{Bin}(n,p)$ for short.
p.m.f. of $X$:

$$\mathbb{P}(X=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

$k \in \{0, 1, \ldots, n\}$

- choose $k$ pos. from $\{1, \ldots, n\}$
- where $s_i \in E$

Let's check the normalization:

$$\sum_{k=0}^{n} \mathbb{P}(X=k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+(1-p))^n = 1^n = 1$$

Binomial theorem
p.m.f. of Binomial r.v.

$P(X=k)$

$\text{Bin}(20, \frac{1}{2})$

$\text{Bin}(30, \frac{1}{2})$

$\text{Bin}(40, \frac{1}{2})$
Example: A biased coin turns up Heads 60% of the time. We flip it 10 times. What is the prob. that there are at least 8 Heads?

Sol: Let $X$ denote the number of Heads. $X$ is a r.v. $X$ is a $\text{Bin}(10, 0.6)$ r.v.

$\Rightarrow \quad P(\text{at least 8 Heads}) = P(X \geq 8) = P(X \in \{8, 9, 10\}) = P(X = 8) + P(X = 9) + P(X = 10) = \binom{10}{8} \cdot 0.6^8 \cdot 0.4^2 + \binom{10}{9} \cdot 0.6^9 \cdot 0.4 + \binom{10}{10} \cdot 0.6^{10} \cdot 0.9 = \boxed{16.7\%}$
(III) Geometric r.v. with parameter \( p \in [0,1] \).

Consider an experiment with sample space \( S \).
Repeat this experiment over and over, independently and indefinitely.

New sample space \( \hat{S} = S^\infty = \{ (s_1,s_2,\ldots) : s_i \in S \} \).

Let \( E \subset S \) be an event in the original single experiment, such that \( \mathbb{P}(E) = p \).

Define \( X : \hat{S} \to \mathbb{N} \) by
\[
X((s_1,s_2,\ldots)) = \text{ "the first time that } E \text{ occurs"} = \text{ smallest } i \text{ such that } s_i \in E.
\]
Such a r.v. is called a geometric r.v. with param $p$, $\text{Geom}(p)$ for short.

P.m.f of $X$:

$$P(X=k) = \mathbb{P}(\{(s_1, s_2, \ldots): s_1, \ldots, s_{k-1} \notin E, s_k \in E\})$$

indirect dependency

$$= \mathbb{P}(\{(s_3, \ldots): s_1 > E\}) \cdot \mathbb{P}(\{(s_5, \ldots): s_2 > E\}) \cdot \mathbb{P}(\{(s_7, \ldots): s_4 > E\}) \cdot \mathbb{P}(\{(s_9, \ldots): s_6 > E\}) \cdot \mathbb{P}(\{(s_{11}, \ldots): s_7 > E\})$$

Check normalization:

$$\sum_{k=1}^{\infty} P(X=k) = p \cdot \sum_{k=1}^{\infty} (1-p)^{k-1} = p \cdot \frac{1}{1-(1-p)} = 1$$
Last time: Random variables

(I) Ber(p), Bernoulli, \[ P(X=1) = p \quad P(X=0) = 1-p \]

(II) Bin(n,p), Binomial, \[ P(X=k) = \binom{n}{k} \cdot p^k (1-p)^{n-k} \quad k=0,1,\ldots,n \]

(III) Geom(p), Geometric, \[ P(X=k) = p \cdot (1-p)^{k-1} \quad k=1,2,\ldots \]

Today: (IV) Negative Binomial r.v.
Example: Roll a die repeatedly. What is the prob. that the first even number appears on an even roll?

| Roll # | 1 | 2 | 3 | 4 | 5 | ...
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>outcome</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>....</td>
</tr>
</tbody>
</table>

Sol: Let $X$ denote the roll number on which an even number first appears.

→ We are looking for $\mathbb{P}(X \text{ is even}).$

→ $X$ is a $\text{Geom}(\frac{1}{2})$ r.v.
\[ P(X \text{ is even}) = P(X \in \{2, 4, 6, 8, 10, \ldots \}) \]
\[ = P(X = 2) + P(X = 4) + \ldots \]
\[ = \sum_{k=1}^{\infty} P(X = 2k) \]
\[ = \sum_{k=1}^{\infty} \left( \frac{1}{2} \right) \cdot (1-\frac{1}{2})^{2k-1} \]
\[ = \sum_{k=1}^{\infty} \left( \frac{1}{2} \right)^{2k} \]
\[ = \sum_{k=1}^{\infty} \left( \frac{1}{2} \right)^{k-1} + 1, \quad k-1 = l \]
\[ = \frac{1}{4} \sum_{l=0}^{\infty} \left( \frac{1}{4} \right)^{l} = \frac{1}{4} \cdot \frac{1}{1-\frac{1}{4}} = \frac{1}{3} \]
(IV) **Negative Binomial r.v.** with parameters $\mathbb{N}, p \in [0,1]$. 

Sample space $\hat{S} = S^\infty$ as for geometric r.v. 

$E \in S$ an event for the single experiment, $P(E) = p$. 

Define $X: \hat{S} \to \mathbb{N}$ by 

$$X(s, s_2, ...) = \text{"r-th time that } E \text{ occurs"}$$

$$= r \text{-th smallest } i \text{ such that } s_i \in E.$$ 

Actually, $X$ takes values in $\{r, r+1, r+2, ... \}$. 

This r.v. is called **NegBin**($r$, $p$).
p.m.f. of \( x \): \( k \in \{r, r+1, \ldots, 5\} \)

\[ P(x=k) = \binom{k-1}{r-1} \cdot p^r \cdot (1-p)^{k-r} \]

\[ \begin{array}{cccccc}
& E & E & E & E & \ldots \\
1 & 2 & 3 & 4 & 5 & 6 & 7 \\
\end{array} \]

choose \( r-1 \) places where \( E \) occurs

Example: You are playing ping-pong against a friend who is a better player, winning 60% of the serves. You were lucky and current score is: 8 - 5
What is the prob. that you will win the game?
(you reach 11 points before your friend)

So:

You need 3 more points.
Friend needs 6 more points.

Let $X$ be the # of serves when you first get to 11 points.

You win if by the time that happens, your friend has at most 20 points.

$\rightarrow X$ is a $\text{NegBin}(3,0.4)$ r.v.
After $X$ serves: you have exactly 11 points. Your friend has exactly
\[5 + (X - 3) = 12 + X\]
points.

You win $\iff 2 + X \leq 10 \iff X \leq 8$

\[P(\text{you win}) = P(X \leq 8)\]
\[= P(X \in \{3, 4, \ldots, 8\})\]
\[= P(X=3) + P(X=4) + \cdots + P(X=8)\]

p.m.f.
\[r = 3\]
\[p = 0.4\]

\[\rightarrow \quad = (\tfrac{1}{2}) \cdot 0.4^3 \cdot 0.6^0 + (\tfrac{3}{2}) \cdot 0.4^3 \cdot 0.6^1 + \ldots + (\tfrac{7}{2}) \cdot 0.4^3 \cdot 0.6^5 \approx [68\%]\]