Motivation: “Named” Random Variables

Random Variables that show up all over the place.
- Easily solve a problem by recognizing it’s a special case of one of these random variables.

Each RV introduced today will show:
- A general situation it models
- Its name and parameters
- Its PMF, Expectation, and Variance
- Example scenarios you can use it
Welcome to the Zoo! (Preview)

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Probability Mass Function</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X \sim \text{Unif}(a, b)$</td>
<td>$P(X = k) = \frac{1}{b - a + 1}$</td>
<td>$E[X] = \frac{a + b}{2}$</td>
<td>$Var(X) = \frac{(b - a)(b - a + 2)}{12}$</td>
</tr>
<tr>
<td>$X \sim \text{Ber}(p)$</td>
<td>$P(X = 1) = p, P(X = 0) = 1 - p$</td>
<td>$E[X] = p$</td>
<td>$Var(X) = p(1 - p)$</td>
</tr>
<tr>
<td>$X \sim \text{Bin}(n, p)$</td>
<td>$P(X = k) = \binom{n}{k}p^k(1 - p)^{n-k}$</td>
<td>$E[X] = np$</td>
<td>$Var(X) = np(1 - p)$</td>
</tr>
<tr>
<td>$X \sim \text{Geo}(p)$</td>
<td>$P(X = k) = (1 - p)^{k-1}p$</td>
<td>$E[X] = \frac{1}{p}$</td>
<td>$Var(X) = \frac{1-p}{p^2}$</td>
</tr>
<tr>
<td>$X \sim \text{NegBin}(r, p)$</td>
<td>$P(X = k) = \binom{k-1}{r-1}p^r(1-p)^{k-r}$</td>
<td>$E[X] = \frac{r}{p}$</td>
<td>$Var(X) = \frac{r(1-p)}{p^2}$</td>
</tr>
<tr>
<td>$X \sim \text{HypGeo}(N, K, n)$</td>
<td>$P(X = k) = \binom{K}{k} \frac{(N-K)^{n-k}}{N^n}$</td>
<td>$E[X] = \frac{nK}{N}$</td>
<td>$Var(X) = n^2 \frac{K(N-K)(N-n)}{N^2(N-1)}$</td>
</tr>
</tbody>
</table>
Agenda

- Discrete Uniform Random Variables
- Bernoulli Random Variables
- Binomial Random Variables
- Geometric Random Variables
- Applications
Discrete Uniform Random Variables

A discrete random variable $X$ **equally likely** to take any (int.) value between integers $a$ and $b$ (inclusive), is **uniform**.

**Notation:** $X \sim \text{Unif}(a, b)$

**PMF:**

$$\Pr(X=k) = \frac{1}{b-a+1}$$

**Expectation:**

$$E(X) = \sum_{k=a}^{b} k \cdot \frac{1}{b-a+1} = \frac{a+b}{2}$$

**Variance:**

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

**Example:** value shown on one roll of a fair die

$$X \sim \text{Unif}(1, 6)$$
Discrete Uniform Random Variables

A discrete random variable $X$ equally likely to take any (int.) value between integers $a$ and $b$ (inclusive), is uniform.

**Notation:** $X \sim \text{Unif}(a, b)$

**PMF:** $\Pr(X = i) = \frac{1}{b - a + 1}$

**Expectation:** $E[X] = \frac{a + b}{2}$

**Variance:** $\text{Var}(X) = \frac{(b - a)(b - a + 2)}{12}$

**Example:** value shown on one roll of a fair die is $\text{Unif}(1, 6)$:
- $\Pr(X = i) = 1/6$
- $E[X] = 7/2$
- $\text{Var}(X) = 35/12$
Agenda

• Discrete Uniform Random Variables
• Bernoulli Random Variables
• Binomial Random Variables
• Geometric Random Variables
• Applications
Bernoulli Random Variables

A random variable $X$ that takes value 1 ("Success") with probability $p$, and 0 ("Failure") otherwise. $X$ is called a Bernoulli random variable.

**Notation:** $X \sim \text{Ber}(p)$

**PMF:** $\Pr(X = 1) = p$, $\Pr(X = 0) = 1 - p$

**Expectation:**

$E(X) = p$

**Variance:**

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

\[ E(X^2) = 1^2 p + 0^2 (1-p) = p \]

\[ \text{Var}(X) = p - p^2 = p(1-p) \]

Poll:

<table>
<thead>
<tr>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$p$</td>
</tr>
<tr>
<td>$p$</td>
<td>$1 - p$</td>
</tr>
<tr>
<td>$p$</td>
<td>$p(1-p)$</td>
</tr>
<tr>
<td>$p$</td>
<td>$p(1-p)$</td>
</tr>
</tbody>
</table>

https://pollev.com/annakarlin185
Bernoulli Random Variables

A random variable $X$ that takes value 1 (“Success”) with probability $p$, and 0 (“Failure”) otherwise. $X$ is called a Bernoulli random variable.

Notation: $X \sim \text{Ber}(p)$

PMF: $\Pr(X = 1) = p$, $\Pr(X = 0) = 1 - p$

Expectation: $E[X] = p$     Note: $E[X^2] = p$

Variance: $\text{Var}(X) = E[X^2] - E[X]^2 = p - p^2 = p(1 - p)$

Examples:
- Coin flip
- Randomly guessing on a MC test question
- A server in a cluster fails
Agenda

- Discrete Uniform Random Variables
- Bernoulli Random Variables
- Binomial Random Variables
- Geometric Random Variables
- Applications
Binomial Random Variables

A discrete random variable $X$ that is the number of successes in $n$ independent random variables $Y_i \sim \text{Ber}(p)$. $X$ is a Binomial random variable where $X = \sum_{i=1}^{n} Y_i$

Examples:

- # of heads in $n$ coin flips
- # of 1s in a randomly generated $n$ bit string
- # of servers that fail in a cluster of $n$ computers
- # of bit errors in file written to disk
- # of elements in a bucket of a large hash table

Poll:

https://pollev.com/annakarlin185

$\Pr(X = k) =$

a. $p^k (1 - p)^{n-k}$
b. $np$
c. $\binom{n}{k} p^k (1 - p)^{n-k}$
d. $\binom{n}{n-k} p^k (1 - p)^{n-k}$
Binomial Random Variables

A discrete random variable $X$ that is the number of successes in $n$ independent random variables $Y_i \sim \text{Ber}(p)$. $X$ is a Binomial random variable where $X = \sum_{i=1}^{n} Y_i$

**Notation:** $X \sim \text{Bin}(n, p)$

**PMF:** $\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$  

**Expectation:**

**Variance:**

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

**Poll:**

<table>
<thead>
<tr>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a. \quad p$</td>
<td>$p$</td>
</tr>
<tr>
<td>$b. \quad np$</td>
<td>$np(1-p)$</td>
</tr>
<tr>
<td>$c. \quad np$</td>
<td>$np^2$</td>
</tr>
<tr>
<td>$d. \quad np$</td>
<td>$n^2p$</td>
</tr>
</tbody>
</table>

https://pollev.com/annakarlin185

$$\text{Var}(X_1 + X_2 + \ldots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \ldots + \text{Var}(X_n)$$

if $X_i$'s mutually indep.
Binomial Random Variables

A discrete random variable $X$ that is the number of successes in $n$ independent random variables $Y_i \sim \text{Ber}(p)$. $X$ is a Binomial random variable where $X = \sum_{i=1}^{n} Y_i$

Notation: $X \sim \text{Bin}(n, p)$

PMF: $\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$

Expectation: $E[X] = np$

Variance: $\text{Var}(X) = np(1 - p)$
Mean, Variance of the Binomial

If $Y_1, Y_2, ..., Y_n \sim \text{Ber}(p)$ and independent (i.i.d), then
$X = \sum_{i=1}^{n} Y_i$, $X \sim \text{Bin}(n, p)$

Claim $E[X] = np$

$$E[X] = E\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} E[Y_i] = nE[Y_1] = np$$

Claim $Var(X) = np(1 - p)$

$$Var(X) = Var\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} Var(Y_i) = nVar(Y_1) = np(1 - p)$$

by independence
Binomial PMFs

PMF for $X \sim \text{Bin}(10, 0.5)$

PMF for $X \sim \text{Bin}(10, 0.25)$

$p = 0.25$

\[\text{mean} = \text{expectation}\]
Example

Sending a binary message of length 1024 bits over a network with probability 0.999 of correctly sending each bit in the message without corruption (independent of other bits). Let $X$ be the number of corrupted bits. What is $E[X]$?

Poll:

https://pollev.com/ annakarlin185

a. 1022.99
b. 1.024
b. 1.02298
d. 1
e. Not enough information to compute
Brain Break
Agenda

- Discrete Uniform Random Variables
- Bernoulli Random Variables
- Binomial Random Variables
- Geometric and other Random Variables
Geometric Random Variables

A discrete random variable $X$ that models the number of independent trials $Y_i \sim \text{Ber}(p)$ before seeing the first success. $X$ is called a Geometric random variable with parameter $p$.

**Notation:** $X \sim \text{Geo}(p)$

**PMF:** $\Pr(X = k) = (1-p)^{k-1} p$

**Expectation:** $E(X) = \frac{1}{p}$

**Variance:**

**Examples:**
- # of coin flips until first head
- # of random guesses on MC questions until you get one right
- # of random guesses at a password until you hit it
Geometric Random Variables

A discrete random variable $X$ that models the number of independent trials $Y_i \sim \text{Ber}(p)$ before seeing the first success. $X$ is called a geometric random variable with parameter $p$.

Notation: $X \sim \text{Geo}(p)$

PMF: $\Pr(X = k) = (1 - p)^{k-1} p$

Expectation: $E[X] = \frac{1}{p}$

Variance: $\text{Var}(X) = \frac{1-p}{p^2}$

Examples:

- # of coin flips until first head
- # of random guesses on MC questions until you get one right
- # of random guesses at a password until you hit it
Example: Music Lessons

Your music teacher requires you to play a 1000 note song without mistake. You have been practicing, so you have a probability of 0.999 of getting each note correct (independent of the others). If you mess up a single note in the song, you must start over and play from the beginning. Let $X$ be the number of times you have to play the song from the start. What is $E[X]$?

$X \sim \text{Geo}(0.001)$

$\Pr(\text{successfully play all the way thru}) = (0.999)^{1000}$

$Y$: # notes I play until I mess up a note.

$\Pr(Y = 10) = 0.999 \times 0.001$
Negative Binomial Random Variables

A discrete random variable $X$ that models the number of independent trials $Y_i \sim Ber(p)$ before seeing the $r^{th}$ success. Equivalently, $X = \sum_{i=1}^{r} Z_i$ where $Z_i \sim Geo(p)$. $X$ is called a Negative Binomial random variable with parameters $r, p$.

Notation: $X \sim \text{NegBin}(r, p)$

PMF:

Expectation:

Variance:
Negative Binomial Random Variables

A discrete random variable $X$ that models the number of independent trials $Y_i \sim Ber(p)$ before seeing the $r^{th}$ success. Equivalently, $X = \sum_{i=1}^{r} Z_i$ where $Z_i \sim Geo(p)$. $X$ is called a **Negative Binomial** random variable with parameters $r, p$.

**Notation:** $X \sim \text{NegBin}(r, p)$

**PMF:** $\Pr(X = k) = \binom{k-1}{r-1} p^r (1 - p)^{k-r}$

**Expectation:** $E[X] = \frac{r}{p}$

**Variance:** $\text{Var}(X) = \frac{r(1-p)}{p^2}$
Hypergeometric Random Variables

A discrete random variable $X$ that measures the number of white balls you draw when you draw $n$ balls uniformly at random from a total of $N$ of which $K$ are white and the rest are black. $X$ is called a Hypergeometric RV with parameters $N, K, n$.

Notation: $X \sim \text{HypGeo}(N, K, n)$

PMF:

$$\Pr(X = k) = \binom{K}{k} \binom{N-K}{n-k} \binom{N}{n}$$

Expectation:
Hypergeometric Random Variables

A discrete random variable $X$ that measures the number of white balls you draw when you draw $n$ balls uniformly at random from a total of $N$ of which $K$ are white and the rest are black. $X$ is called a Hypergeometric RV with parameters $N, K, n$.

**Notation:** $X \sim \text{HypGeo}(N, K, n)$

**PMF:** $\Pr(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$

**Expectation:** $\mathbb{E}[X] = n \frac{K}{N}$

**Variance:** $\text{Var}(X) = n \frac{K(N-K)(N-n)}{N^2(N-1)}$
Hope you enjoyed the zoo!

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Density</th>
<th>Expectation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X \sim \text{Unif}(a, b)$</td>
<td>$P(X = k) = \frac{1}{b - a + 1}$</td>
<td>$E[X] = \frac{a + b}{2}$</td>
<td>$Var(X) = \frac{(b - a)(b - a + 2)}{12}$</td>
</tr>
<tr>
<td>$X \sim \text{Ber}(p)$</td>
<td>$P(X = 1) = p, P(X = 0) = 1 - p$</td>
<td>$E[X] = p$</td>
<td>$Var(X) = p(1 - p)$</td>
</tr>
<tr>
<td>$X \sim \text{Bin}(n, p)$</td>
<td>$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$</td>
<td>$E[X] = np$</td>
<td>$Var(X) = np(1 - p)$</td>
</tr>
<tr>
<td>$X \sim \text{Geo}(p)$</td>
<td>$P(X = k) = (1 - p)^{k-1}p$</td>
<td>$E[X] = \frac{1}{p}$</td>
<td>$Var(X) = \frac{1 - p}{p^2}$</td>
</tr>
<tr>
<td>$X \sim \text{NegBin}(r, p)$</td>
<td>$P(X = k) = \binom{k - 1}{r - 1} p^r (1 - p)^{k-r}$</td>
<td>$E[X] = \frac{r}{p}$</td>
<td>$Var(X) = \frac{r(1 - p)}{p^2}$</td>
</tr>
<tr>
<td>$X \sim \text{HypGeo}(N, K, n)$</td>
<td>$P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$</td>
<td>$E[X] = \frac{nK}{N}$</td>
<td>$Var(X) = \frac{nK(N - K)(N - n)}{N^2(N - 1)}$</td>
</tr>
</tbody>
</table>