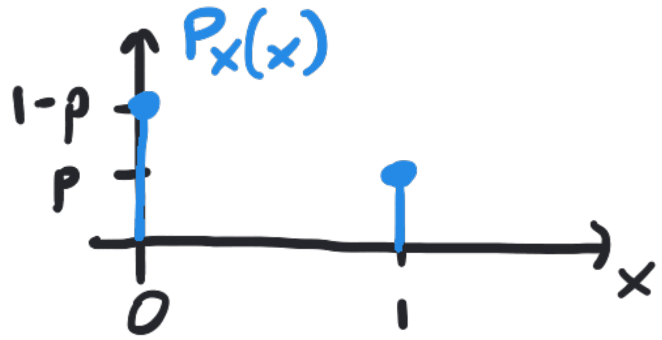
Important Families of Discrete Random Variables

- Many problems have the same underlying probability structure.
- Useful to learn a few families of random variables that show up often to avoid repetitive calculations.
- **Bernoulli**: X is Bernoulli(p) if it has PMF

$$P_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \end{cases}$$



→ Range: $R_X = \{0,1\}$

→ Mean: $\mathbb{E}[X] = p$

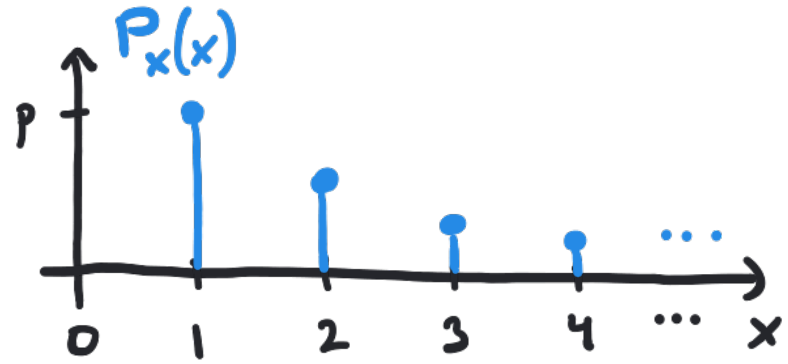
→ Variance: $\mathrm{Var}[X] = p(1-p)$

→ Interpretation: Single binary trial with success probability p.

→ Application: packet received, treatment effective, part within tolerance

- **Geometric:** X is Geometric(p) if it has PMF

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$



→ Range: $R_X = \{1,2,3,\ldots\}$

→ Mean: $\mathbb{E}[X] = \frac{1}{p}$

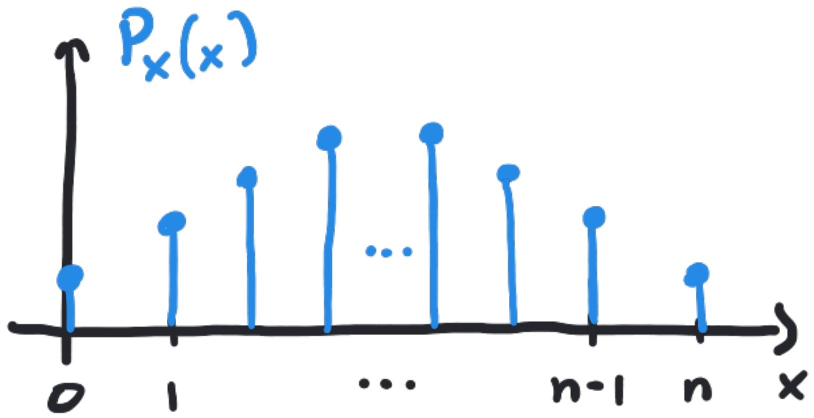
→ Variance: $\mathrm{Var}[X] = \frac{1-p}{p^2}$

→ Interpretation: Number of independent Bernoulli(p) trials until first success.

→ Application: Packet retransmissions until acknowledgment.
Patients administered drug until first cured.
Parts manufactured until first within spec.

- **Binomial:** X is a Binomial(n,p) random variable if it has PMF

$$P_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$



→ Range: $R_X = \{0, 1, 2, \ldots, n\}$

→ Mean: $\mathbb{E}[X] = np$

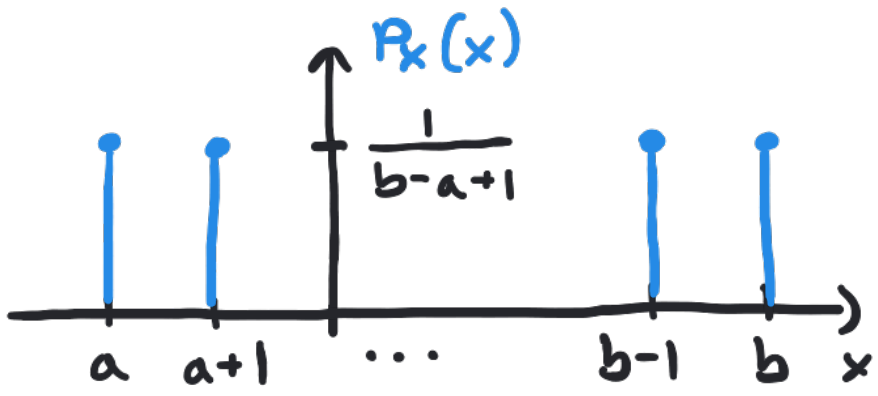
→ Variance: $\mathrm{Var}[X] = np(1-p)$

→ Interpretation: Number of successes in n independent Bernoulli(p) trials.

→ Application: Number of packets successfully received.
Number of patients cured by drug.
Number of parts within tolerance.

- **Discrete Uniform:** X is a Discrete Uniform(a,b) random variable if it has PMF

$$P_X(x) = \begin{cases} \dfrac{1}{b-a+1} & x = a, a+1, \dots, b-1, b \\ 0 & \text{otherwise} \end{cases}$$



→ Range: $R_X = \{a, a+1, \dots, b-1, b\}$

→ Mean: $\mathbb{E}[X] = \dfrac{a+b}{2}$

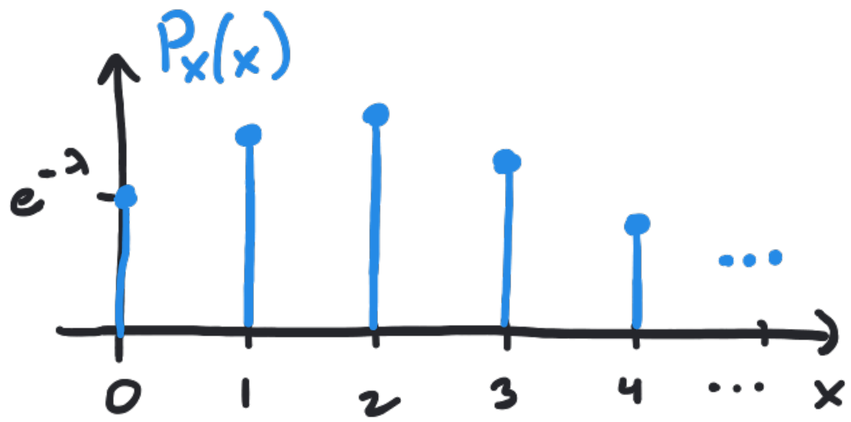
→ Variance: $\mathrm{Var}[X] = \dfrac{(b-a)(b-a+2)}{12} = \dfrac{(b-a+1)^2 - 1}{12}$

→ Interpretation: Equally likely outcomes.

→ Application: Roll of a die $(a=1, b=6)$.

- Poisson: X is Poisson(λ) if it has PMF

$$P_X(x) = \begin{cases} \dfrac{\lambda^x}{x!} e^{-\lambda} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$



→ Range: $R_X = \{0, 1, 2, \ldots\}$

→ Mean: $\mathbb{E}[X] = \lambda$

→ Variance: $\mathrm{Var}[X] = \lambda$

→ Interpretation: Number of occurrences in a fixed time window

→ Application: Number of photons hitting a CCD pixel in 1 ms
Number of neuron action potentials in 10 s
Number of turbine failures in 1 week

- **Example:** Assume that your favorite sports (or e-sports!) team wins a match with probability $\frac{3}{4}$, independently of all other matches. Let

$$X = \text{\# matches watched to see a win}$$

a) What kind of random variable is X? Geometric$\left(\frac{3}{4}\right)$ Don't forget the parameters.

b) How many matches do you need to watch, on average, to see a win? $\mathbb{E}[X] = \frac{1}{p} = \frac{1}{\left(\frac{3}{4}\right)} = \frac{4}{3}$

c) What is the probability that the first win occurs on the 3^{rd} match or later?

$$\mathbb{P}[\{X \geq 3\}] = \sum_{x=3}^{\infty} P_X(x) = \sum_{x=3}^{\infty} p(1-p)^{x-1} = \sum_{x=3}^{\infty} \frac{3}{4}\left(\frac{1}{4}\right)^{x-1}$$

Can we avoid this calculation?

Complement $\quad \|$

$$1 - \mathbb{P}[\{X < 3\}] = 1 - \mathbb{P}[\{X \leq 2\}] = 1 - \left(P_X(1) + P_X(2)\right) = 1 - \frac{3}{4}\cdot\left(\frac{1}{4}\right)^0 - \frac{3}{4}\cdot\left(\frac{1}{4}\right)^1$$

$$= 1 - \frac{3}{4} - \frac{3}{16}$$

$$= \frac{1}{16}$$

X only takes values in $R_X = \{1, 2, 3, \ldots\}$

- <u>Example</u>: Assume that your favorite sports (or e-sports!) team wins a match with probability $\frac{3}{4}$, independently of all other matches. Let

  $Y =$ # of wins in 6 matches

d) What kind of random variable is Y? Binomial$\left(6, \frac{3}{4}\right)$

e) What is $\mathrm{Var}[Y]$? $\mathbb{E}[Y^2]$?

$$\mathrm{Var}[Y] = np(1-p) = 6 \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{18}{16} = \frac{9}{8}$$

$$\mathbb{E}[Y^2] = \sum_{y \in R_Y} y^2 P_Y(y) = \sum_{y=1}^{6} y^2 \binom{6}{y} \left(\frac{3}{4}\right)^y \left(\frac{1}{4}\right)^{6-y}$$ ← Want to avoid computing this by hand.

alternate variance formula

$$= \mathrm{Var}[Y] + (\mathbb{E}[Y])^2 = \frac{9}{8} + \left(6 \cdot \frac{3}{4}\right)^2 = \frac{9}{8} + \left(\frac{9}{2}\right)^2 = \frac{9}{8} + \frac{81}{4} = \frac{171}{8}$$

f) What is the probability the team wins less than half of the matches?

$$\mathbb{P}[\{Y < 3\}] = \mathbb{P}[\{Y \leq 2\}] = \sum_{y=0}^{2} \binom{6}{y} \left(\frac{3}{4}\right)^y \left(\frac{1}{4}\right)^{6-y}$$

$$= \binom{6}{0}\left(\frac{3}{4}\right)^0\left(\frac{1}{4}\right)^6 + \binom{6}{1}\left(\frac{3}{4}\right)^1\left(\frac{1}{4}\right)^5 + \binom{6}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^4 = \frac{154}{4096}$$

- <u>Example</u>: Assume that your favorite sports (or e-sports!) team wins a match with probability $\frac{3}{4}$, independently of all other matches. Let

$$Y = \#\text{ of wins in 6 matches}$$

g) Given that your team wins less than half of the matches, what is the probability that they win exactly two matches?

This is a conditional probability question.

$A = \{Y = 2\}$ $\quad B = \{Y < 3\} = \{Y \leq 2\}$

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[\{Y=2\} \cap \{Y \leq 2\}]}{\mathbb{P}[\{Y \leq 2\}]} = \frac{\mathbb{P}[\{Y=2\}]}{\mathbb{P}[\{Y \leq 2\}]}$$

from f) →

$$= \frac{P_Y(2)}{\left(\frac{154}{4096}\right)} = \frac{\binom{6}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^4}{\left(\frac{154}{4096}\right)} = \frac{\left(\frac{135}{4096}\right)}{\left(\frac{154}{4096}\right)} = \frac{135}{154}$$