

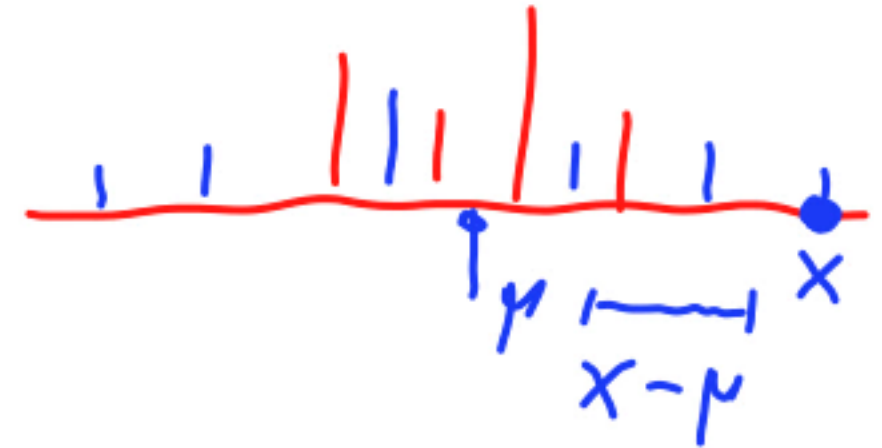
## LECTURE 6: Variance; Conditioning on an event; Multiple random variables

- Variance and its properties
  - Variance of the Bernoulli and uniform PMFs
- Conditioning a r.v. on an event
  - Conditional PMF, mean, variance
  - Total expectation theorem
- Geometric PMF
  - Memorylessness
  - Mean value
- Multiple random variables
  - Joint and marginal PMFs
  - Expected value rule
  - Linearity of expectations
- The mean of the binomial PMF

## Variance — a measure of the spread of a PMF

- Random variable  $X$ , with mean  $\mu = \mathbb{E}[X]$
- Distance from the mean:  $X - \mu$
- Average distance from the mean?

$$\mathbb{E}[X - \mu] = \mathbb{E}[X] - \mu = \mu - \mu = 0$$



- **Definition of variance:**  $\text{var}(X) = \mathbb{E}[(X - \mu)^2]$   $\geq 0$

- Calculation, using the expected value rule,  $\mathbb{E}[g(X)] = \sum_x g(x)p_X(x)$

$$g(x) = (x - \mu)^2 \quad \text{var}(X) = \mathbb{E}[g(x)] = \sum_x (x - \mu)^2 p_X(x)$$

**Standard deviation:**  $\sigma_X = \sqrt{\text{var}(X)}$

## Properties of the variance

$$\text{var}(aX + b) = a^2 \text{var}(X)$$

$$\begin{aligned}\text{var}(3 - 4x) \\ &= (-4)^2 \text{var}(x) \\ &= 16 \text{var}(x)\end{aligned}$$

- Notation:  $\mu = \mathbb{E}[X]$

- Let  $Y = X + b$        $\gamma = \mathbb{E}[Y] = \mu + b$

$$\text{var}(Y) = \mathbb{E}[(Y - \gamma)^2] = \mathbb{E}[(X + \cancel{b} - (\mu + \cancel{b}))^2] = \mathbb{E}[(X - \mu)^2] = \text{var}(X)$$

- Let  $Y = aX$        $\gamma = \mathbb{E}[Y] = a\mu$

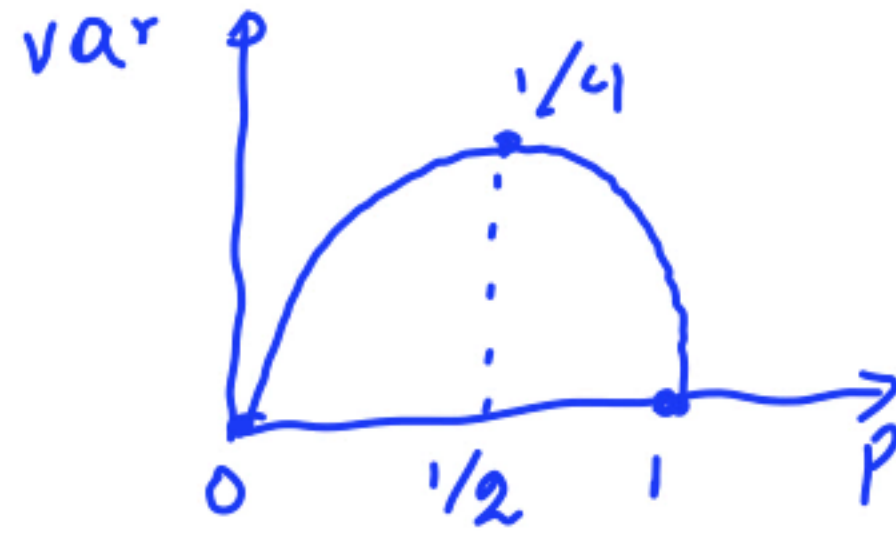
$$\text{var}(Y) = \mathbb{E}[(aX - a\mu)^2] = \mathbb{E}[a^2 (X - \mu)^2] = a^2 \mathbb{E}[(X - \mu)^2] = a^2 \text{var}(X)$$

A useful formula:  $\text{var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

$$\begin{aligned}\text{var}(X) &= \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbb{E}[X^2] - 2\mu \mathbb{E}[X] + \mu^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2\end{aligned}$$

## Variance of the Bernoulli

$$X = \begin{cases} 1, & \text{w.p. } p \\ 0, & \text{w.p. } 1 - p \end{cases} \quad E[X] = p$$



$$\begin{aligned} \text{var}(X) &= \sum_x (x - E[X])^2 p_X(x) = (1-p)^2 p + (0-p)^2 \cdot (1-p) \\ &= p - 2p^2 + \cancel{p^3} + p^2 - \cancel{p^3} = p - p^2 = p(1-p) \end{aligned}$$

$$\begin{aligned} \text{var}(X) &= E[X^2] - (E[X])^2 = E[X] - (E[X])^2 = p - p^2 = \boxed{p(1-p)} \\ X^2 &= X \end{aligned}$$

## Variance of the uniform

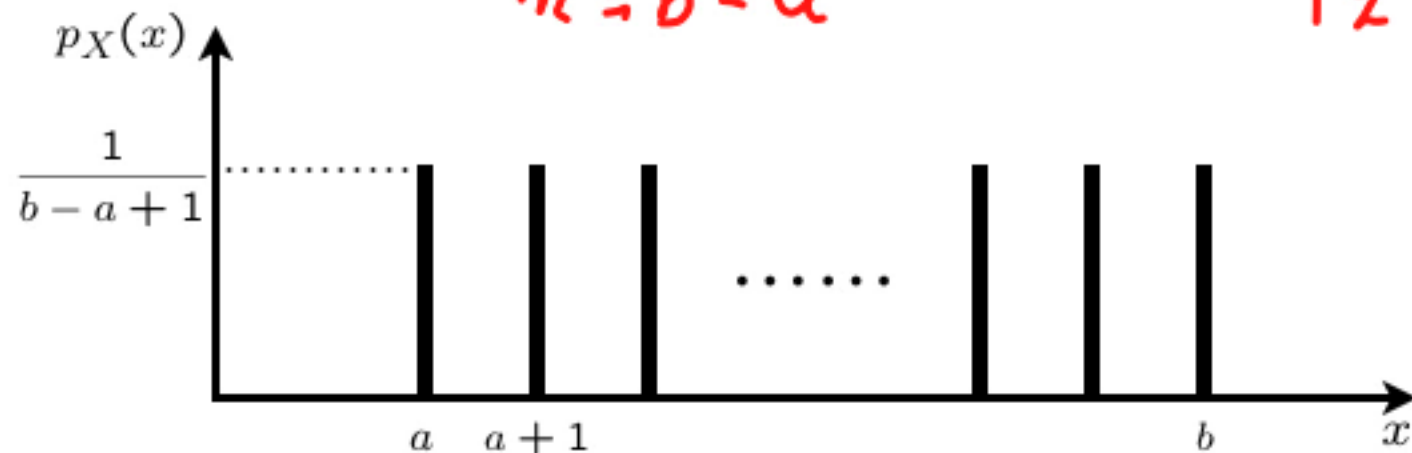


$$\frac{1}{6} n(n+1)(2n+1)$$

$$\text{var}(x) = E[x^2] - (E[x])^2 = \frac{1}{n+1} (0^2 + 1^2 + 2^2 + \dots + n^2) - \left(\frac{n}{2}\right)^2$$

$$= \frac{1}{12} n(n+2)$$

$$n = b - a$$



$$\text{Var}(x) = \frac{1}{12} (b-a)(b-a+2)$$



## Conditional PMF and expectation, given an event

- Condition on an event  $A \Rightarrow$  use conditional probabilities

$$p_X(x) = \mathbf{P}(X = x)$$

$$\underline{p_{X|A}(x)} = \underline{\mathbf{P}(X = x \mid A)}$$

assume  
 $\mathcal{P}(A) > 0$

$$\sum_x p_X(x) = 1$$

$$\sum_x p_{X|A}(x) = 1$$

$$\mathbf{E}[X] = \sum_x x p_X(x)$$

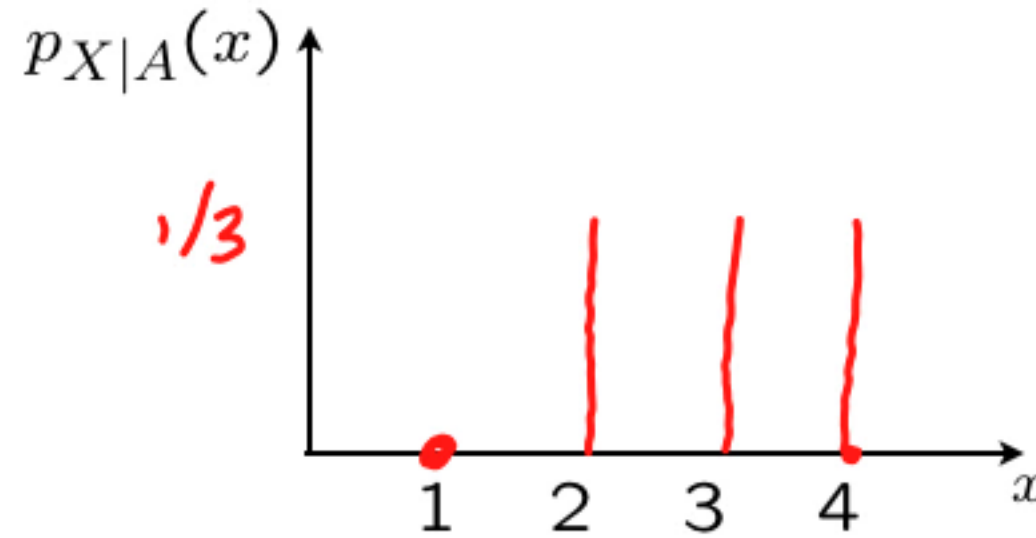
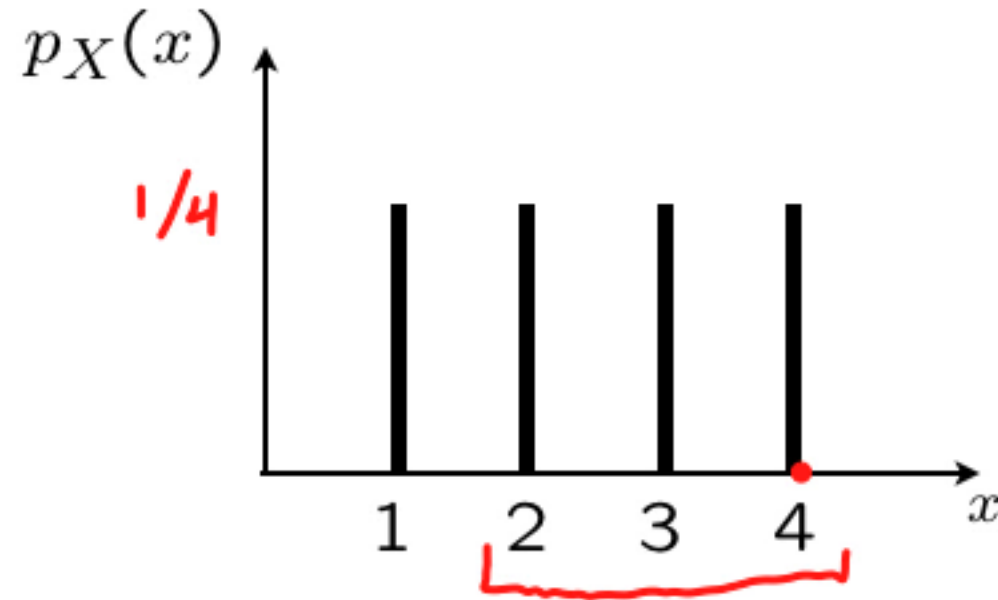
$$\mathbf{E}[X \mid A] = \sum_x x p_{X|A}(x)$$

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x)$$

$$\bullet \quad \mathbf{E}[g(X) \mid A] = \sum_x g(x) p_{X|A}(x)$$

## Example of conditioning

- Let  $A = \{X \geq 2\}$



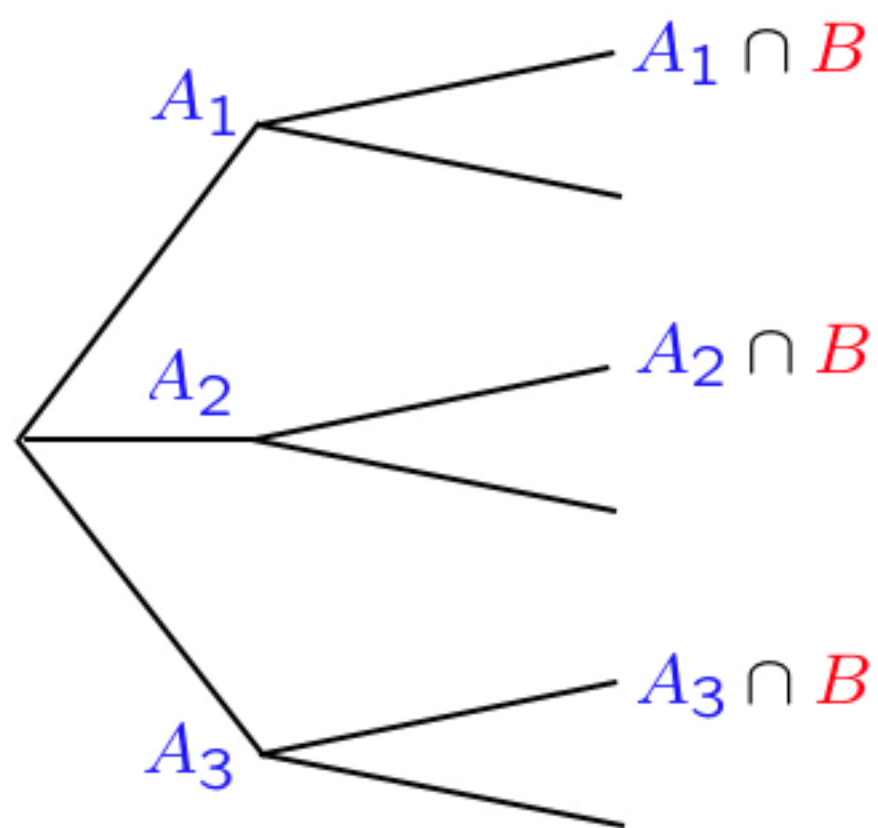
$$\mathbf{E}[X] = 2.5$$

$$\begin{aligned}\text{var}(X) &= \frac{1}{12}(b-a)(b-a+1) \\ &= \frac{1}{12} 3 \cdot 5 = \frac{5}{4}\end{aligned}$$

$$\mathbf{E}[X | A] = 3$$

$$\begin{aligned}\text{var}(X | A) &= \frac{1}{3}(4-3)^2 + \frac{1}{3}(3-3)^2 \\ &\quad + \frac{1}{3}(2-3)^2 = \frac{2}{3}\end{aligned}$$

## Total expectation theorem

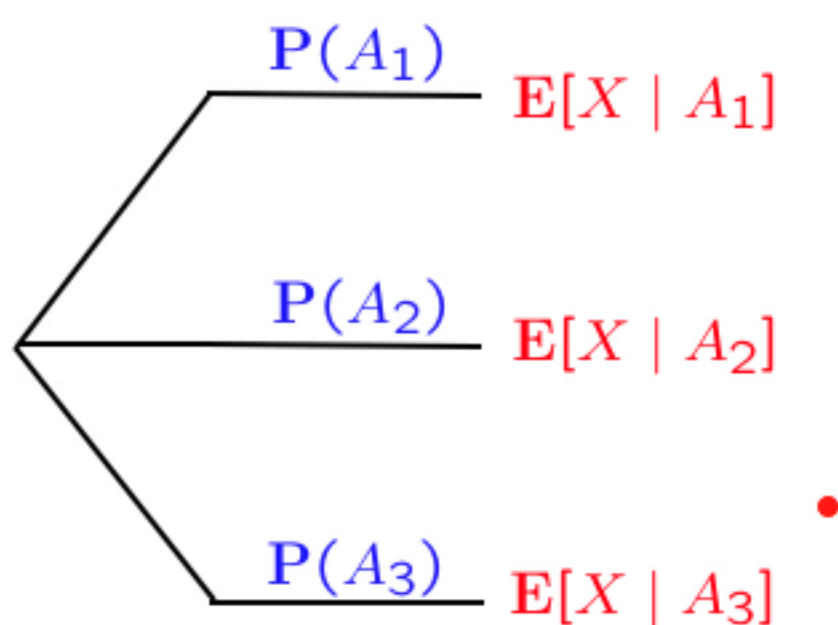
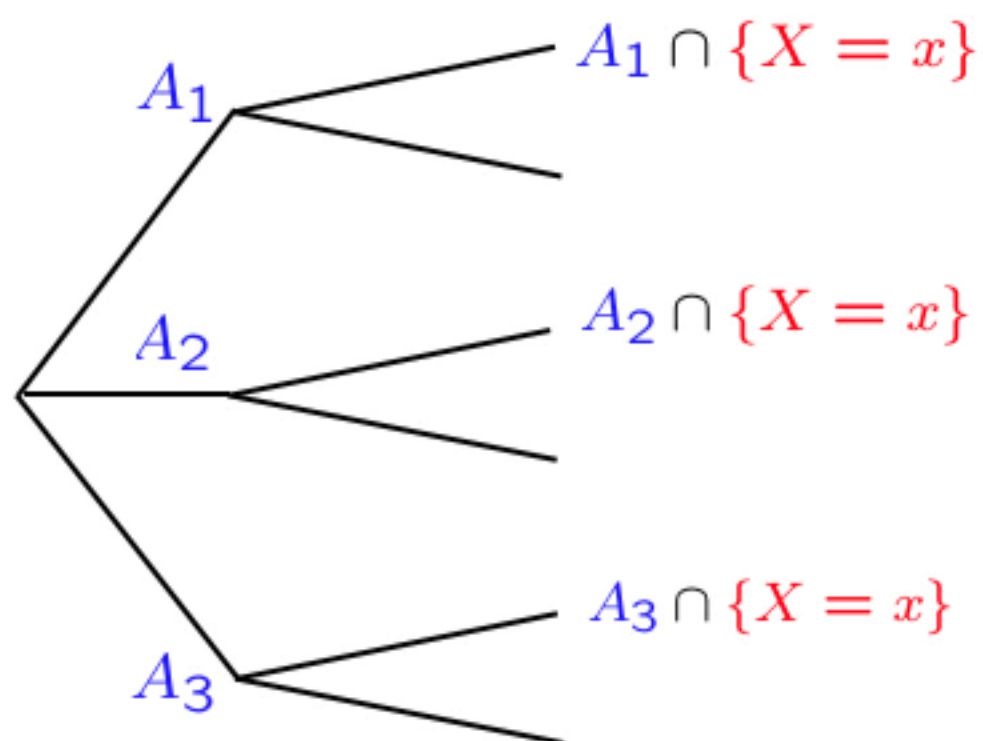


$$\mathbf{P}(B) = \mathbf{P}(A_1) \mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n) \mathbf{P}(B \mid A_n)$$

$$B = \{x = x\}$$



## Total expectation theorem



$$P(B) = P(A_1) P(B | A_1) + \cdots + P(A_n) P(B | A_n)$$

$$B = \{X = x\}$$

$$p_X(x) = P(A_1) p_{X|A_1}(x) + \cdots + P(A_n) p_{X|A_n}(x)$$

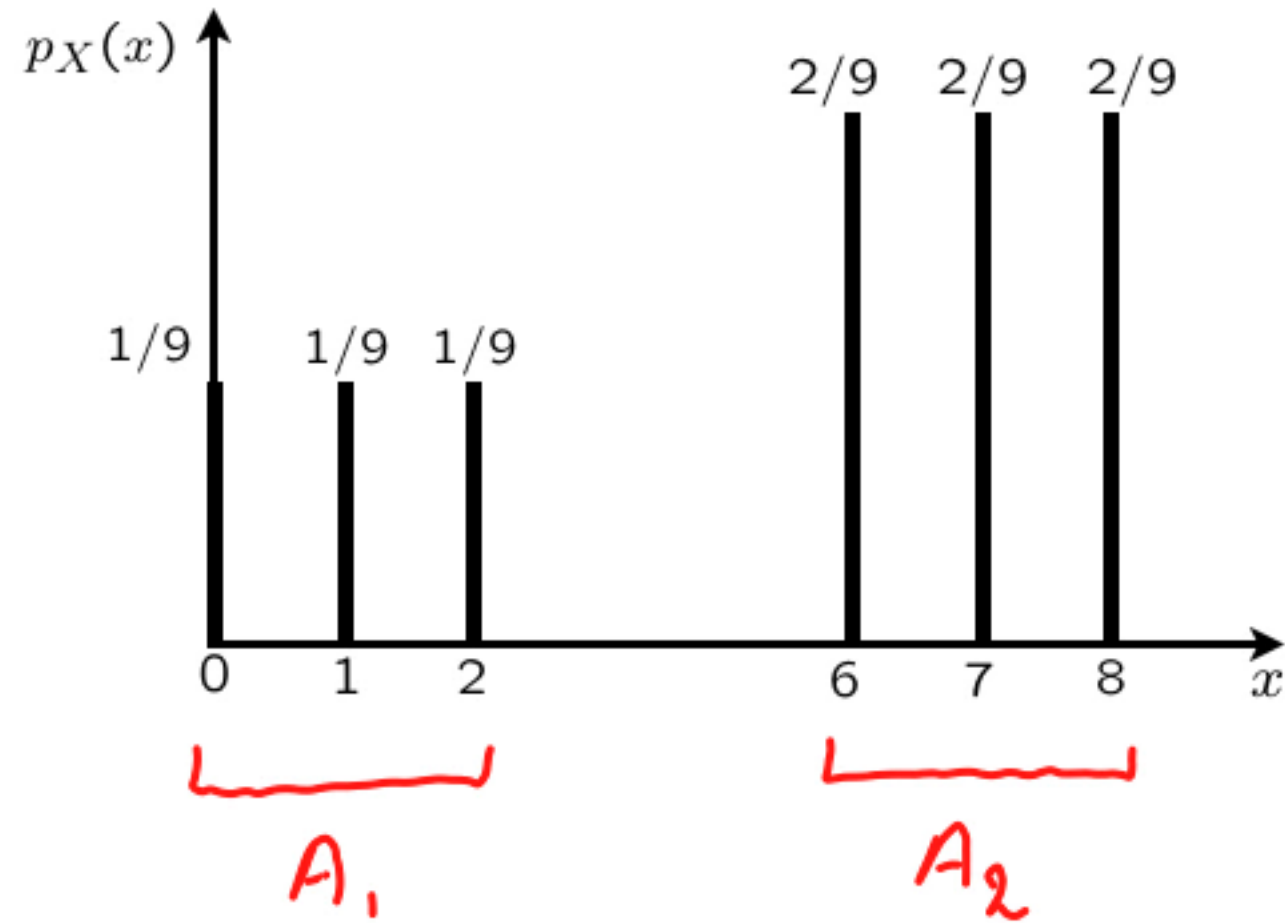
for all  $x$

$$\sum_x x p_X(x) = P(A_1) \underbrace{\sum_x x p_{X|A_1}(x)}_{E[X|A_1]} + \cdots$$

$\parallel$   
 $E[X]$

$$E[X] = P(A_1) E[X | A_1] + \cdots + P(A_n) E[X | A_n]$$

## Total expectation example



$$P(A_1) = \frac{1}{3}$$

$$P(A_2) = \frac{2}{3}$$

$$E[X|A_1] = 1$$

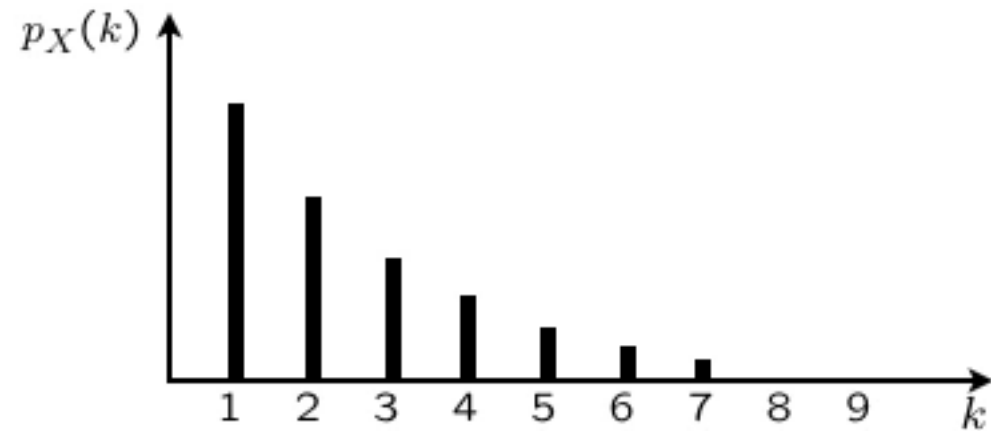
$$E[X|A_2] = 7$$

$$E[X] = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 7$$

## Conditioning a geometric random variable

- $X$ : number of independent coin tosses until first head;  $P(H) = p$

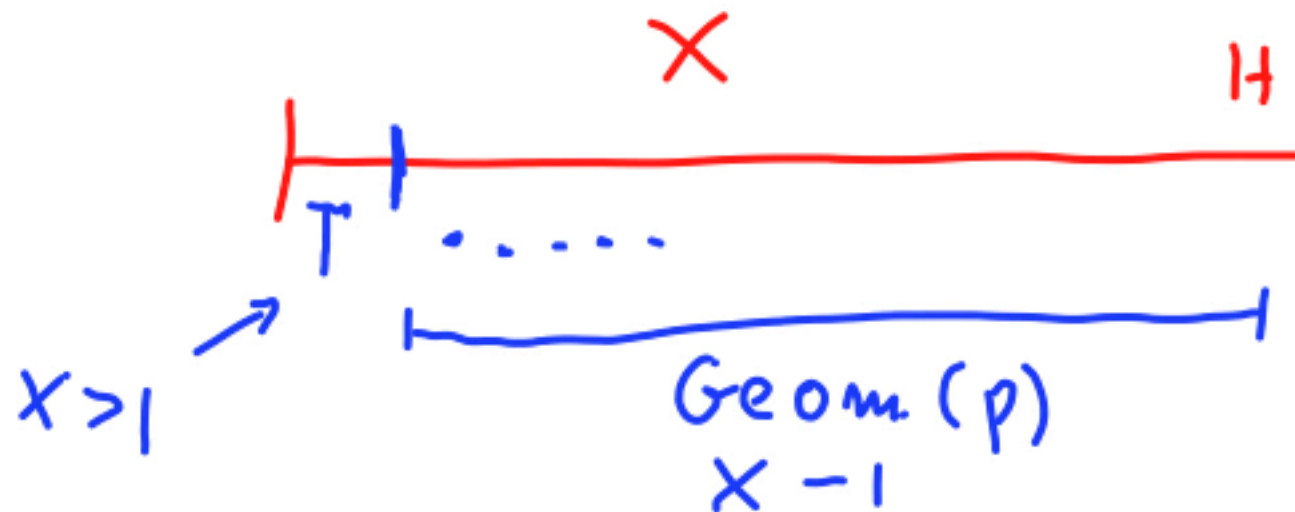
$$p_X(k) = (1 - p)^{k-1}p, \quad k = 1, 2, \dots$$



### Memorylessness:

Number of **remaining** coin tosses, conditioned on Tails in the first toss, is **Geometric**, with parameter  $p$

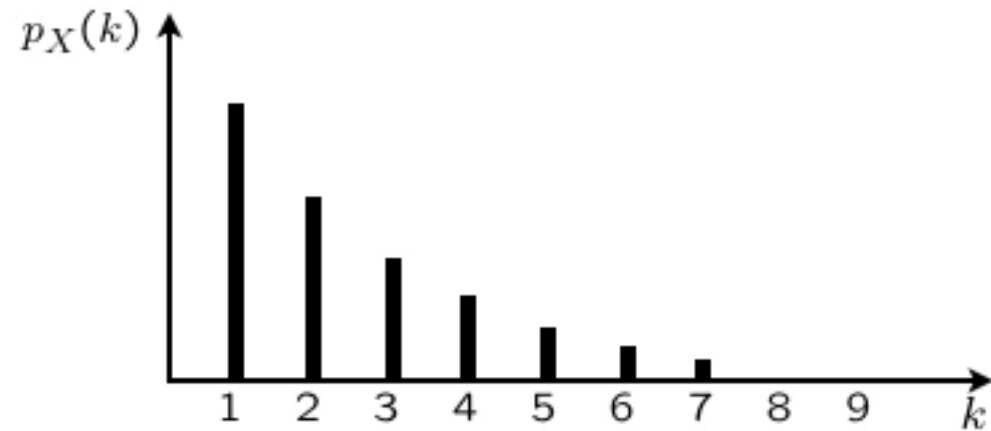
Conditioned on  $X > 1$ ,  $X - 1$  is geometric with parameter  $p$



## Conditioning a geometric random variable

- $X$ : number of independent coin tosses until first head;  $P(H) = p$

$$p_X(k) = (1-p)^{k-1}p, \quad k = 1, 2, \dots$$



### Memorylessness:

Number of **remaining** coin tosses, conditioned on Tails in the first toss, is **Geometric**, with parameter  $p$

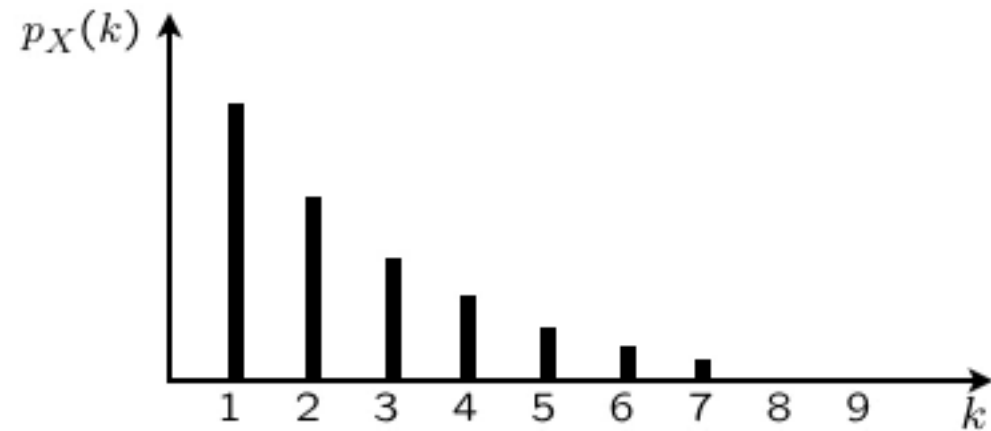
Conditioned on  $X > 1$ ,  $X - 1$  is geometric with parameter  $p$

$$\begin{aligned} p_{X-1|X>1}(3) &= P(X-1=3 | X>1) = P(T_2 T_3 H_4 | T_1) = P(T_2 T_3 H_4) \\ &= (1-p)^2 p = p_X(3) \\ p_{X-1|X>1}(k) &= p_X(k) \end{aligned}$$

## Conditioning a geometric random variable

- $X$ : number of independent coin tosses until first head;  $P(H) = p$

$$p_X(k) = (1-p)^{k-1}p, \quad k = 1, 2, \dots$$



### Memorylessness:

Number of **remaining** coin tosses, conditioned on Tails in the first toss, is **Geometric**, with parameter  $p$

Conditioned on  $X > \underline{n}$ ,  $X - \underline{n}$  is geometric with parameter  $p$

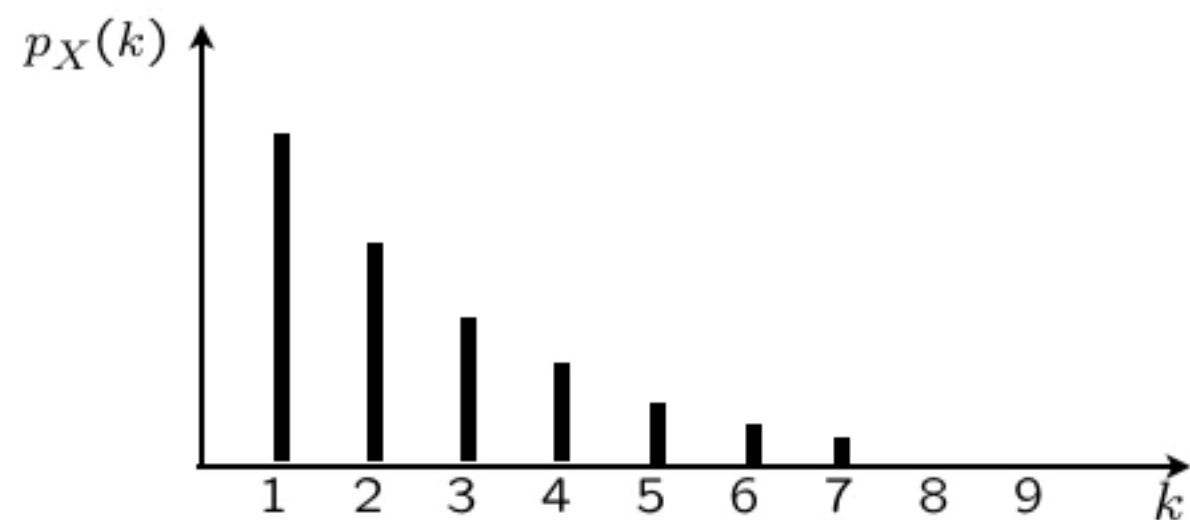
$$p_{X-1|X>1}(3) = P(X-1=3 | X>1) = P(T_2 T_3 H_4 | T_1) = P(T_2 T_3 H_4)$$

$$p_{X-1|X>1}(k) = p_X(k) = (1-p)^{k-1}p = p_X(3)$$

$$p_{X-n|X>n}(k)$$

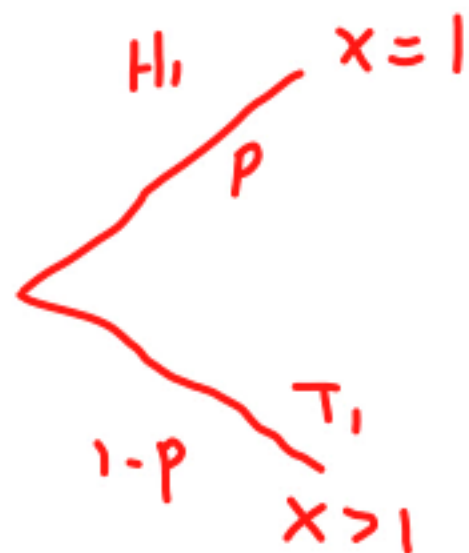


## The mean of the geometric



$$\mathbf{E}[X] = \sum_{k=1}^{\infty} k p_X(k) = \sum_{k=1}^{\infty} k (1-p)^{k-1} p$$

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



$$\begin{aligned} \mathbf{E}[x] &= 1 + \mathbf{E}[x-1] \\ &= 1 + p \cdot \mathbf{E}[x-1 | x=1] + (1-p) \mathbf{E}[x-1 | x>1] \\ &= 1 + 0 + (1-p) \mathbf{E}[x] \end{aligned}$$



## Multiple random variables and joint PMFs

$X : p_X$   
 $Y : p_Y$

*← marginal pmfs*

$P(X = Y) = \frac{2}{20}$

Joint PMF:  $p_{X,Y}(x, y) = P(X = x \text{ and } Y = y)$



$p_X(3)$

$p_{X,Y}(1, 3) = \frac{2}{20}$

$p_X(4) = \frac{1}{20} + \frac{2}{20}$

$p_Y(2) = \frac{1}{20} + \frac{3}{20} + \frac{1}{20}$

$$\sum_x \sum_y p_{X,Y}(x, y) = 1$$

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$

## More than two random variables

$$p_{X,Y,Z}(x,y,z) = \mathbf{P}(X = x \text{ and } Y = y \text{ and } Z = z)$$

$$\sum_x \sum_y \sum_z p_{X,Y,Z}(x,y,z) = 1$$

$$p_X(x) = \sum_y \sum_z p_{X,Y,Z}(x,y,z)$$

$$p_{X,Y}(x,y) = \sum_z p_{X,Y,Z}(x,y,z)$$

## Functions of multiple random variables

$$Z = g(X, Y)$$

$$\text{PMF: } p_Z(z) = \mathbf{P}(Z = z) = \mathbf{P}(g(X, Y) = \underline{z}) = \sum_{(x, y) : g(x, y) = z} p_{X, Y}(x, y)$$

$$\text{Expected value rule: } \mathbf{E}[g(X, Y)] = \sum_x \sum_y \underbrace{g(x, y)} \cdot \underline{\underline{p_{X, Y}(x, y)}}$$

$$E[g(x)]$$

## Linearity of expectations

$$\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$$

$$\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$$

$$E[X + Y] = E[g(x, y)]$$

$$(g(x, y) = x + y)$$

$$= \sum_x \sum_y (x + y) p_{x, y}(x, y)$$

$$= \sum_x \underbrace{\sum_y x p_{x, y}(x, y)} + \sum_x \sum_y y p_{x, y}(x, y)$$

$$= \sum_x x \underbrace{\sum_y p_{x, y}(x, y)} + \underbrace{\sum_x \sum_y y p_{x, y}(x, y)}$$

$$= \sum_x x p_x(x) + \sum_y y p_y(y) = E[X] + E[Y]$$

## Linearity of expectations

$$\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$$

$$\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$$

$$\mathbf{E}[X_1 + \cdots + X_n] = \mathbf{E}[X_1] + \cdots + \mathbf{E}[X_n]$$

$$\mathbf{E}[2X + 3Y - Z] = E[2x] + E[3y] - E[z] = 2E[x] + 3E[y] - E[z]$$

## The mean of the binomial

- $X$ : binomial with parameters  $n, p$ 
  - number of successes in  $n$  independent trials

$$\mathbf{E}[X] = \sum_{k=0}^n k \underbrace{\binom{n}{k} p^k (1-p)^{n-k}}_{P_X(k)}$$

$$\mathbf{E}[X] = np$$

$X_i = 1$  if  $i$ th trial is a success;  $\swarrow p$   
 $X_i = 0$  otherwise  $\searrow 1-p$  (indicator variable)

$$X = X_1 + \cdots + X_n$$

$$E[X] = \underbrace{E[X_1]}_p + \cdots + \underbrace{E[X_n]}_p = np$$