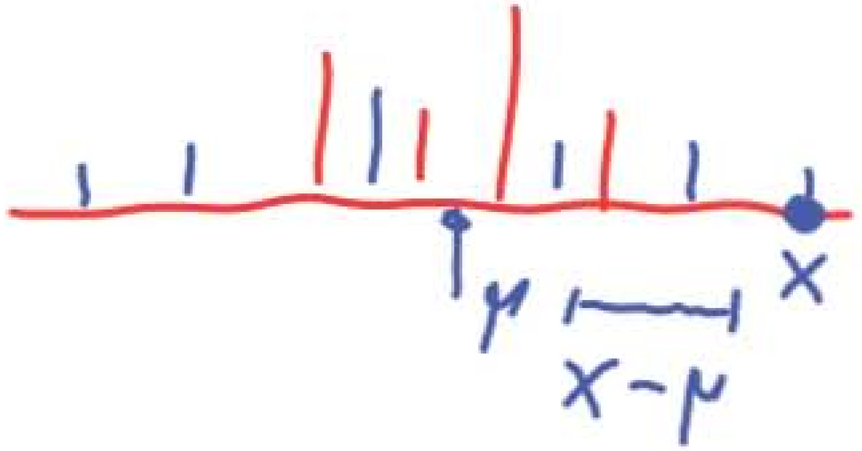
LECTURE 6: Variance; Conditioning on an event; Multiple random variables

- 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 $\mu = \mathbf{E}[X]$
- Distance from the mean: $X - \mu$
- Average distance from the mean?

$$E[X-\mu] = E[X] - \mu = \mu - \mu = 0$$

- **Definition of variance:** $\mathrm{var}(X) = \mathbf{E}[(X-\mu)^2]$

≥ 0

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

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

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

Properties of the variance

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

$\text{var}(3-4X)$
$= (-4)^2\,\text{var}(X)$
$= 16\,\text{var}(X)$

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

- Let $Y = X + b$ $\quad \nu = E[Y] = \mu + b$

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

- Let $Y = aX$ $\quad \nu = E[Y] = a\mu$

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

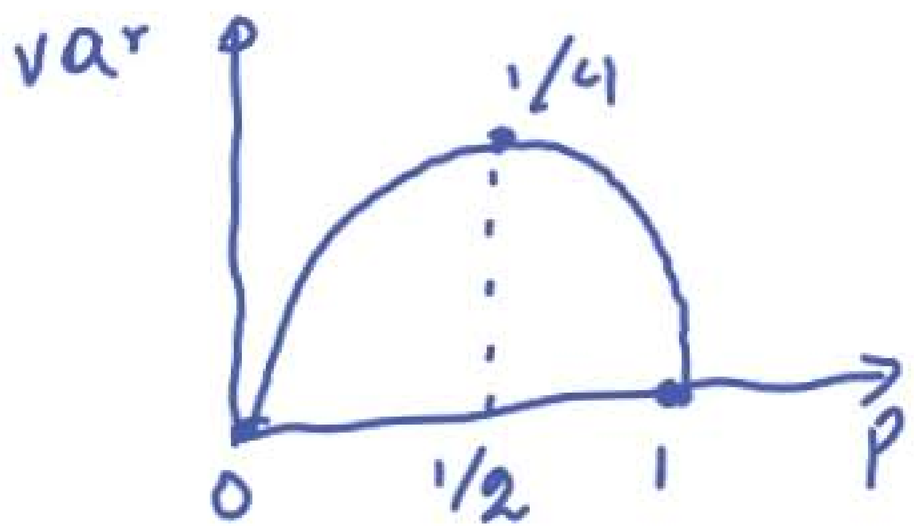
A useful formula: $\text{var}(X) = \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2$

$\text{var}(X) = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2]$

$= E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - (E[X])^2$

Variance of the Bernoulli

$E[X] = p$



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

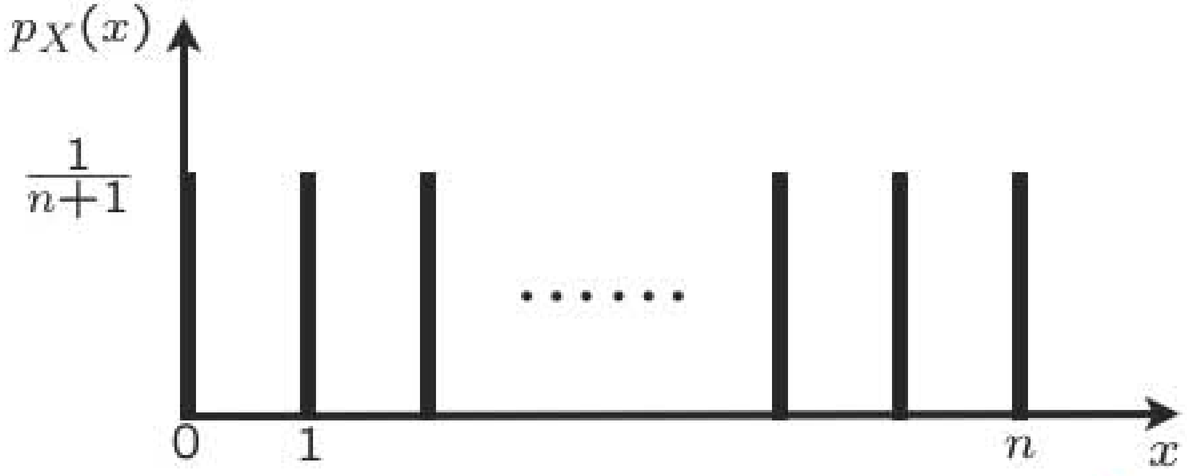
$$\text{var}(X) = \sum_x (x - \mathbf{E}[X])^2 p_X(x) = (1-p)^2 p + (0-p)^2 \cdot (1-p)$$

$$= p - 2p^2 + \cancel{p^3} + p^2 - \cancel{p^3} = p - p^2 = p(1-p)$$

$$\text{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = E[X] - (E[X])^2 = p - p^2 = \boxed{p(1-p)}$$

$X^2 = X$

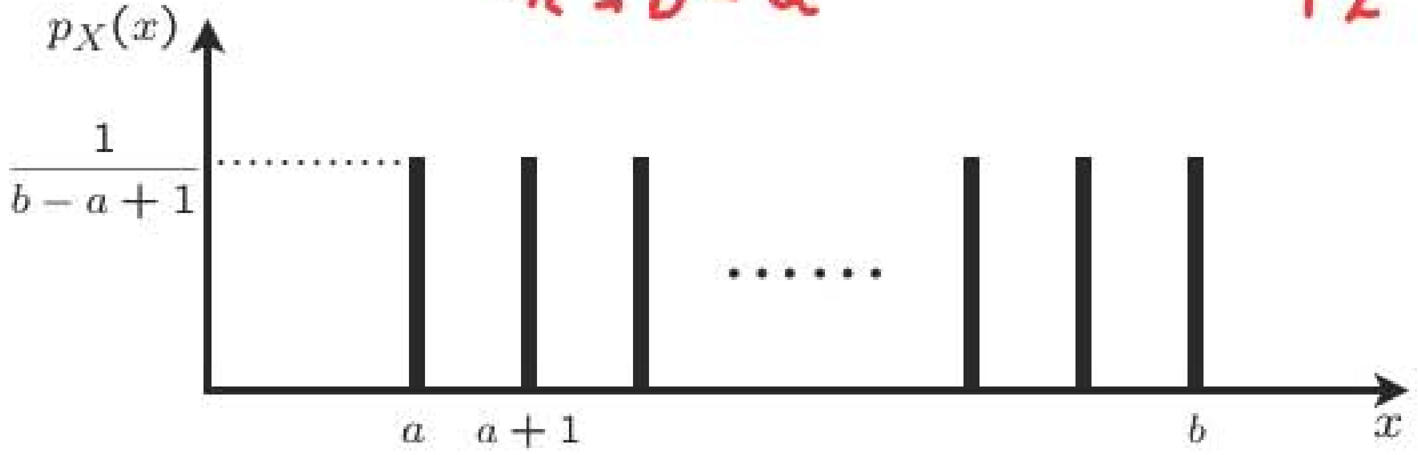
Variance of the uniform



$$\mathrm{var}(X) = E[X^2] - (E[X])^2 = \frac{1}{n+1}\underbrace{(0^2+1^2+2^2+\dots+n^2)}_{\frac{1}{6}n(n+1)(2n+1)} - \left(\frac{n}{2}\right)^2$$

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

$n = b - a$



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

Conditional PMF and expectation, given an event

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

assume $P(A) > 0$

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

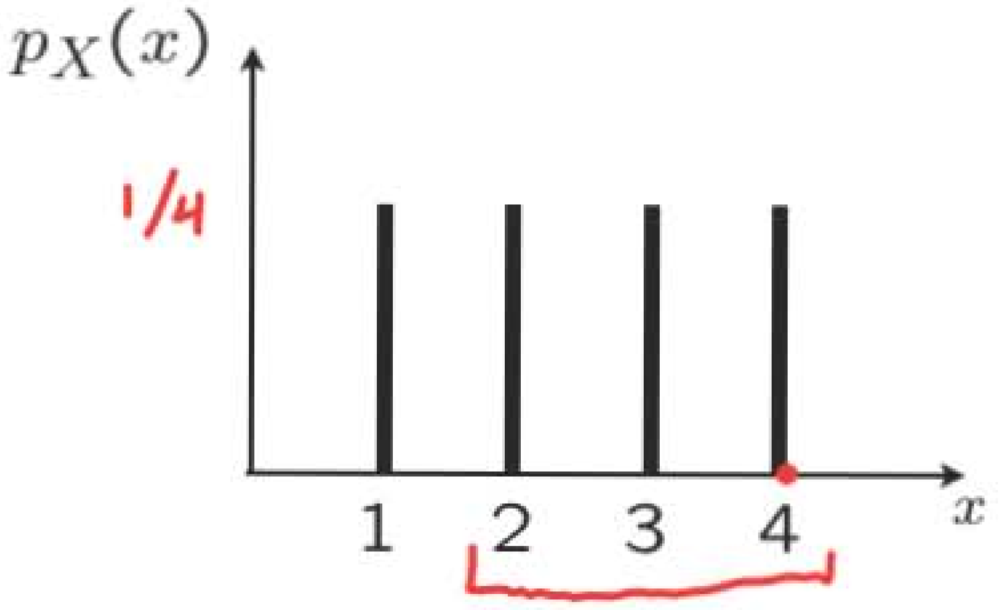
$$\sum_x p_X(x) = 1 \qquad \sum_x p_{X|A}(x) = 1$$

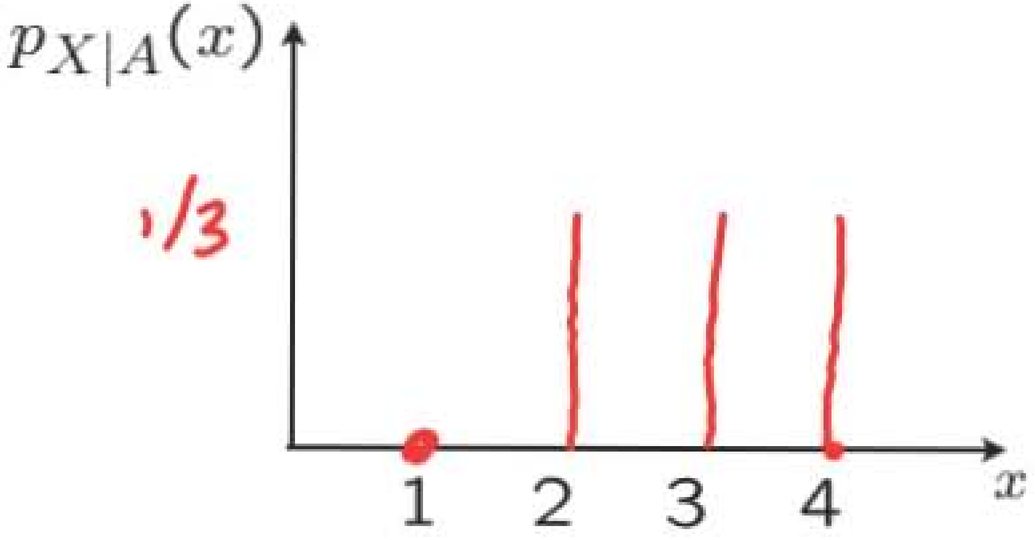
$$\mathbf{E}[X] = \sum_x x p_X(x) \qquad \mathbf{E}[X \mid A] = \sum_x x p_{X|A}(x)$$

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

Example of conditioning

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





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

$\mathrm{var}(X) = \frac{1}{12}(b-a)(b-a+2)$

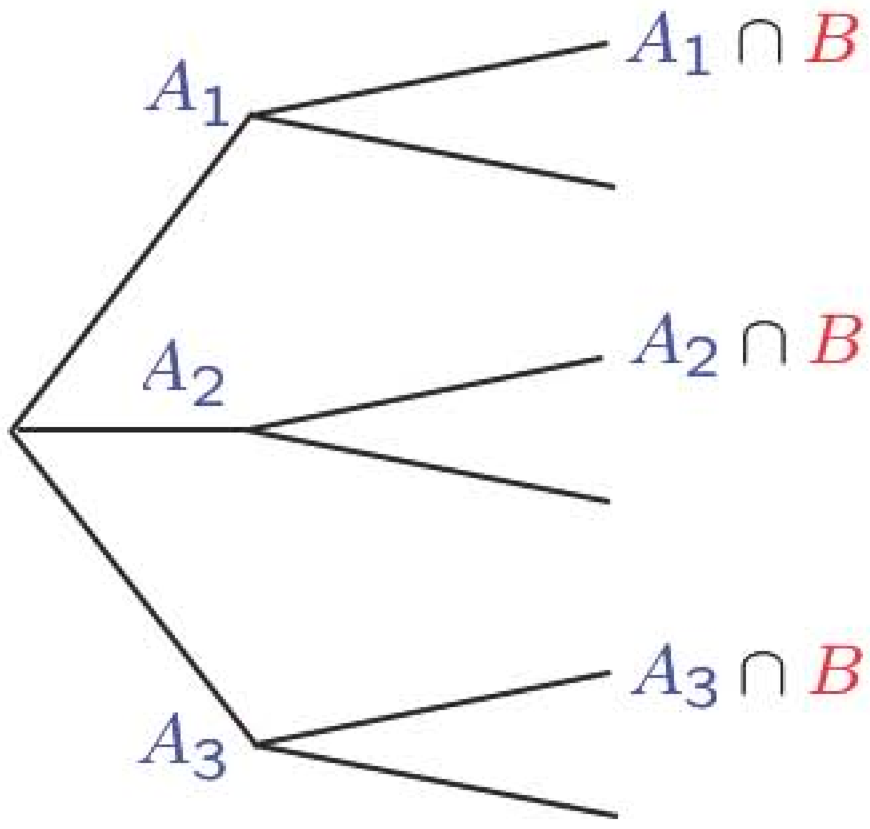
$= \frac{1}{12} 3 \cdot 5 = \frac{5}{4}$

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

$\mathrm{var}(X \mid A) = \frac{1}{3}(4-3)^2 + \frac{1}{3}(3-3)^2$

$+ \frac{1}{3}(2-3)^2 = \frac{2}{3}$

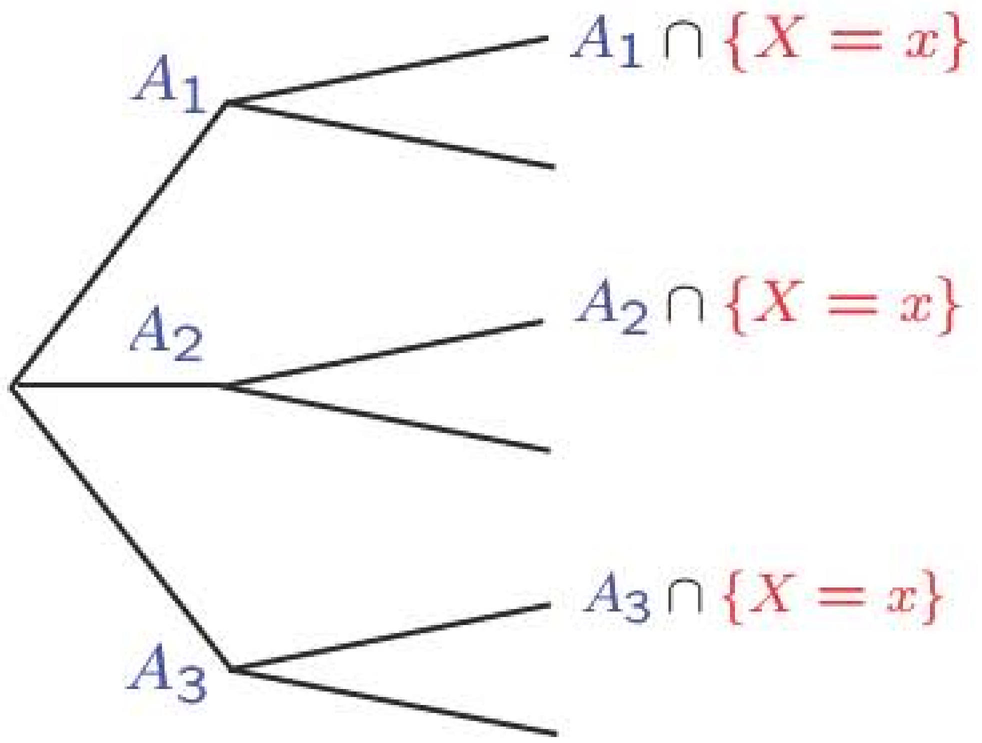
Total expectation theorem



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

$B = \{X = x\}$

Total expectation theorem



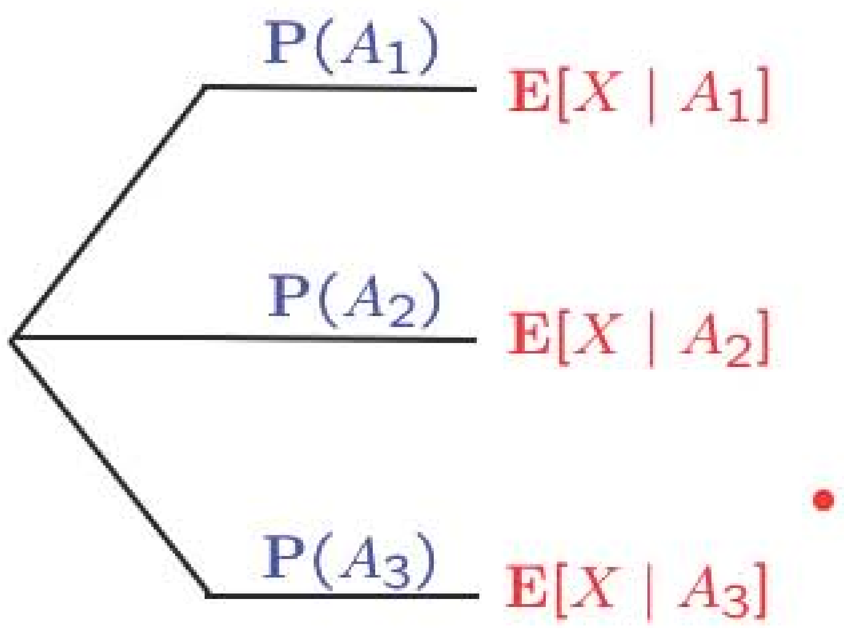
$$\mathbf{P}(B) = \mathbf{P}(A_1)\,\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\,\mathbf{P}(B \mid A_n)$$

$B = \{X = x\}$

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

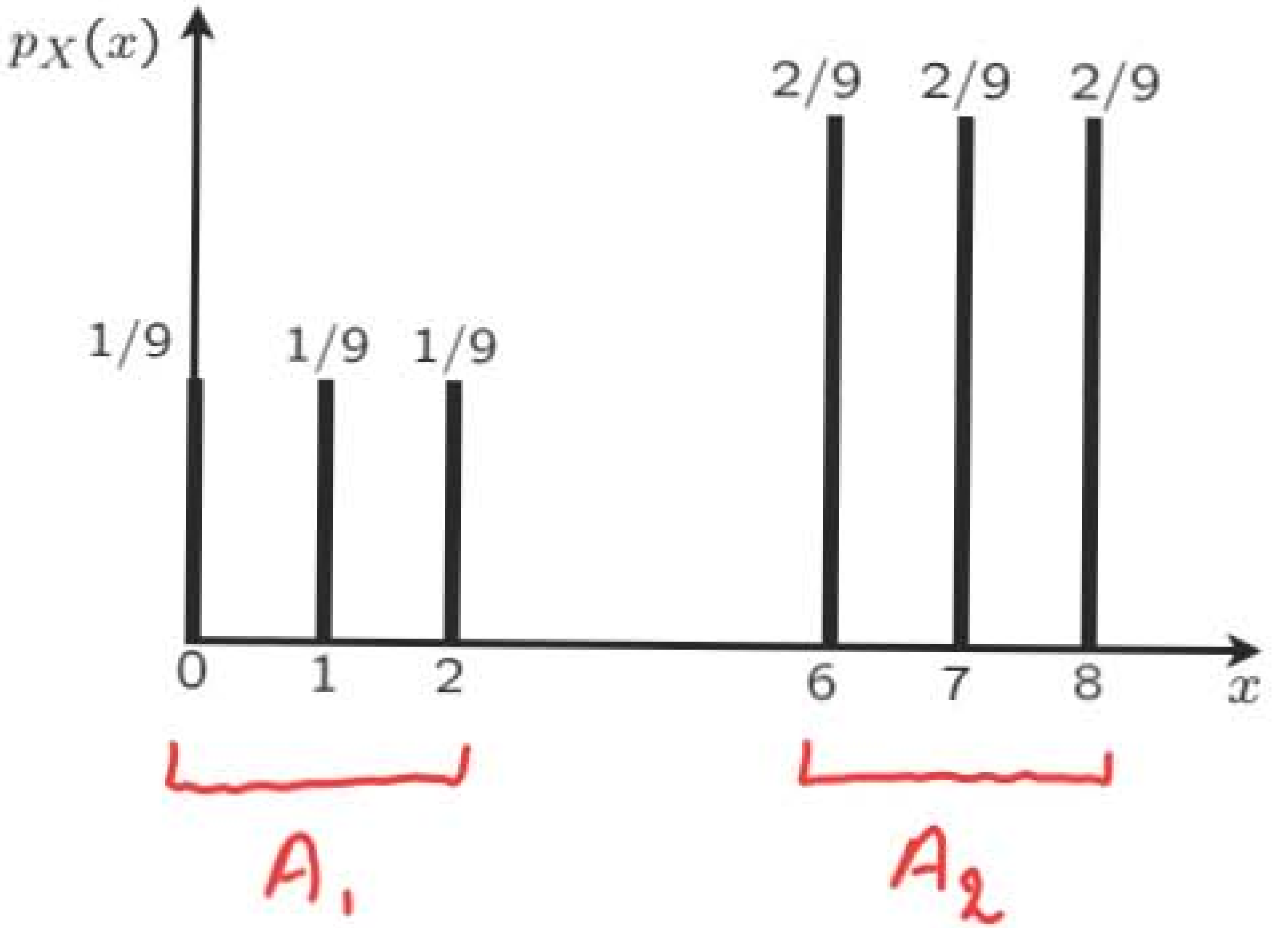
for all x

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



$$\boxed{\mathbf{E}[X] = \mathbf{P}(A_1)\,\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\,\mathbf{E}[X \mid A_n]}$$

Total expectation example



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

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

$E[X \mid A_1] = 1$

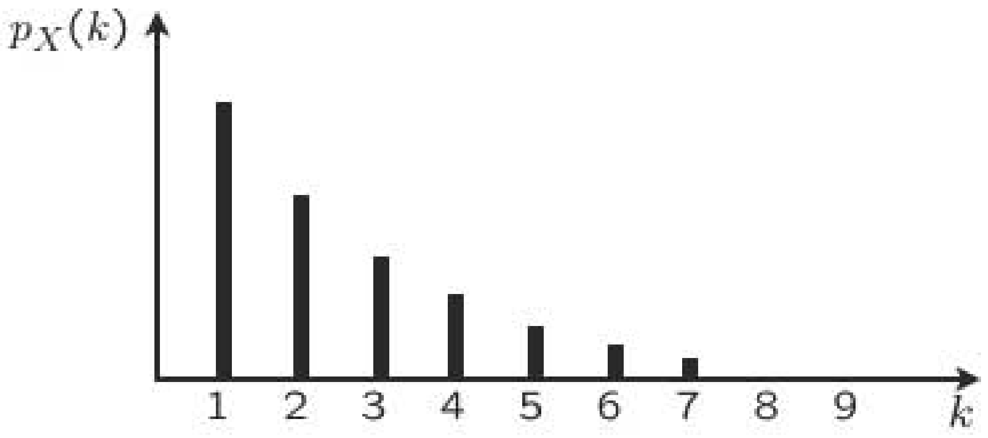
$E[X \mid A_2] = 7$

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

Conditioning a geometric random variable

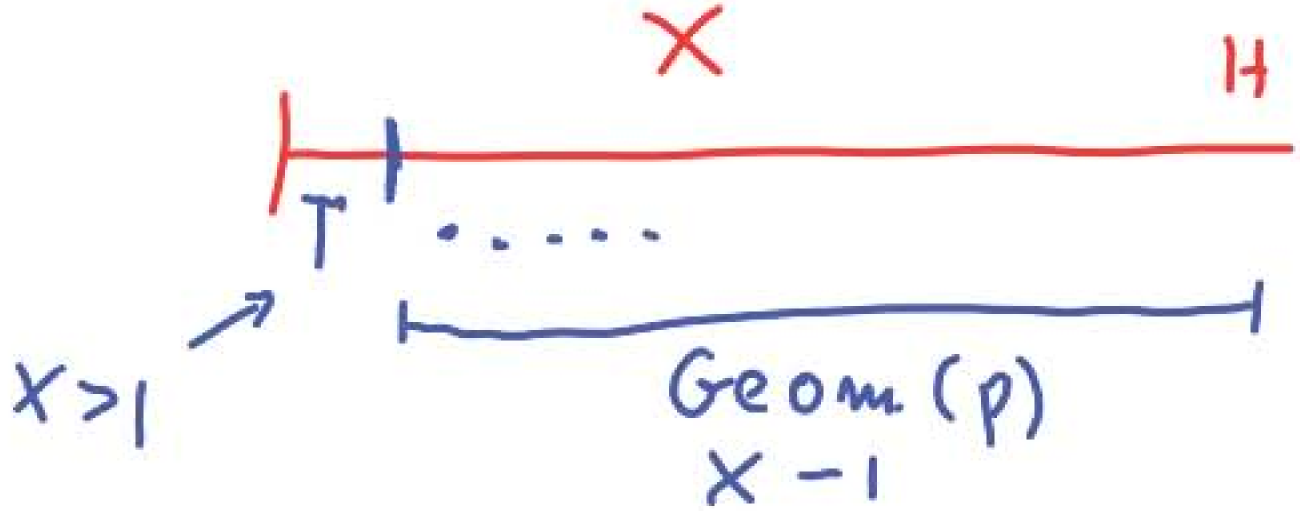
- X: number of independent coin tosses until first head; $\mathbf{P}(H) = p$

$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1, 2, \ldots$$



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

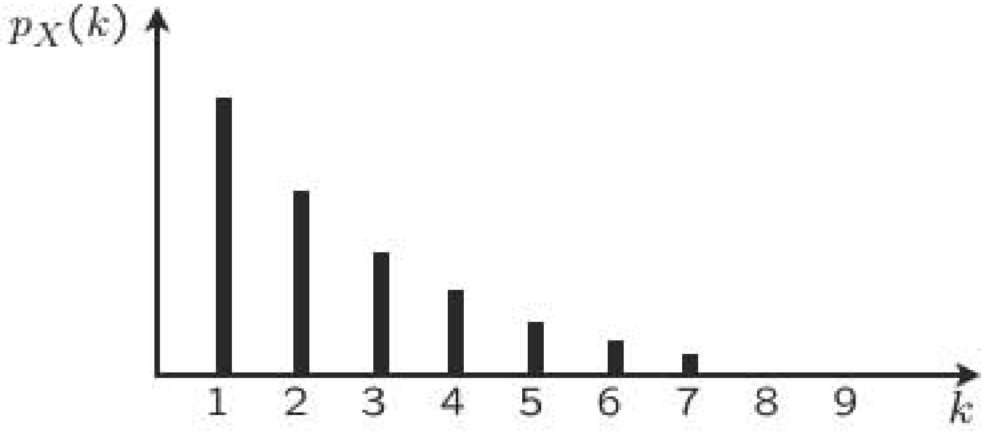
Conditioned on $X > 1$, $X - 1$ is geometric with parameter p



Conditioning a geometric random variable

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

$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1, 2, \ldots$$



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

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

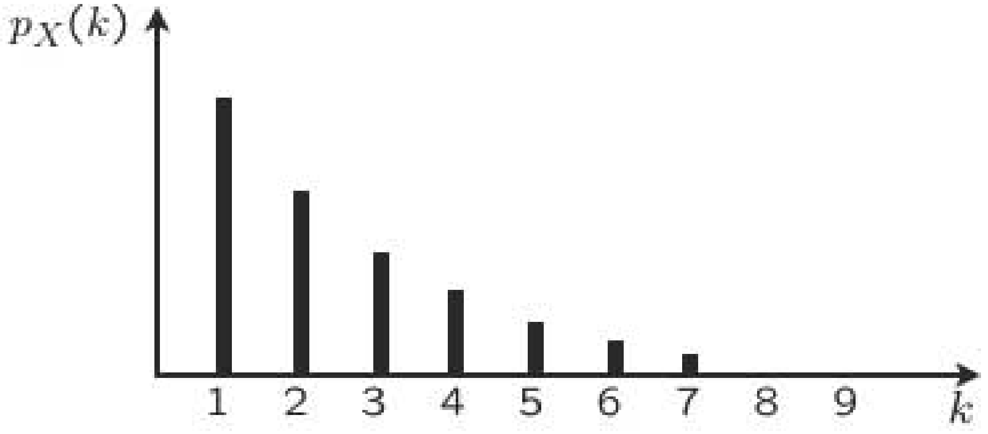
$= (1-p)^2 p = p_X(3)$

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

Conditioning a geometric random variable

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

$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1, 2, \ldots$$



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

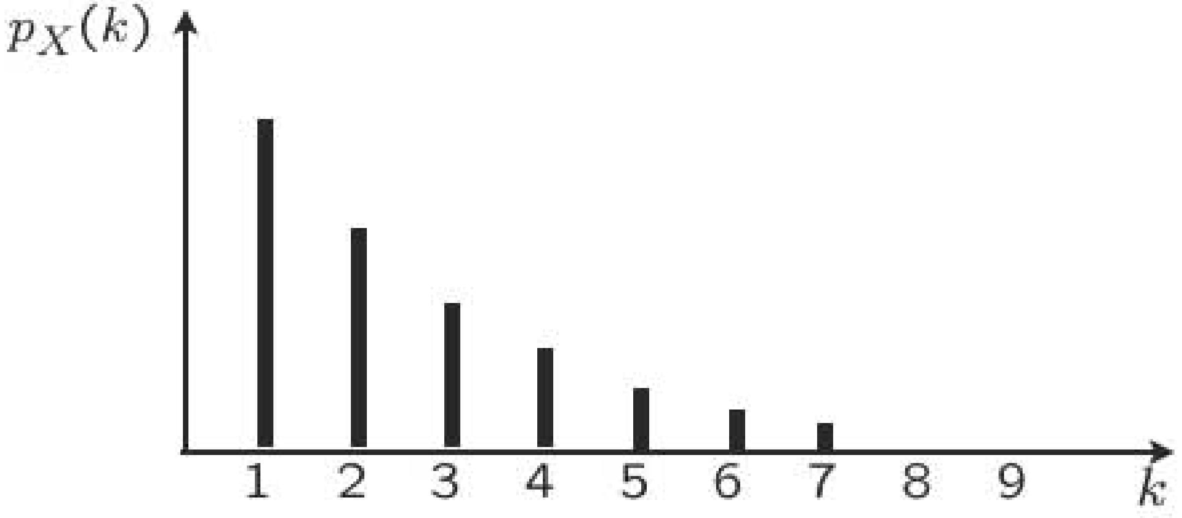
Conditioned on $X > n$, $X - n$ is geometric with parameter p

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

$$= (1-p)^2 p = p_X(3)$$

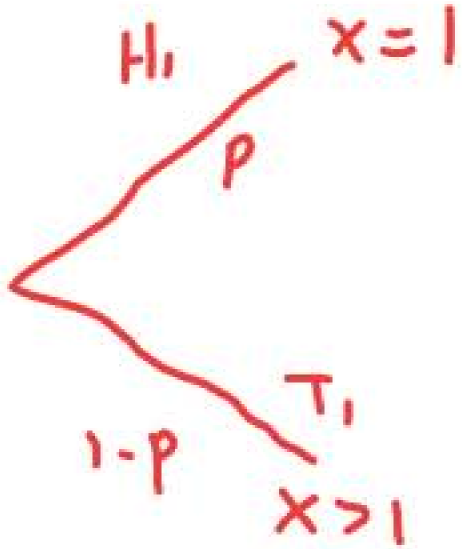
$$p_{X-1|X>1}(k) = p_X(k) = p_{X-n|X>n}(k).$$

The mean of the geometric



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

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



$$E[X] = 1 + E[X-1]$$

$$= 1 + p \cdot E[X-1 \mid X=1] + (1-p) E[X-1 \mid X>1]$$

$$= 1 + 0 + (1-p) E[X]$$

Multiple random variables and joint PMFs

$X : p_X$ ← marginal pmfs
$Y : p_Y$

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

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

y \ x	1	2	3	4
4	1/20	2/20	2/20	
3	2/20	4/20	1/20	2/20
2		1/20	3/20	1/20
1		1/20		

$p_X(3)$

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

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

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

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

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

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

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

$Z = g(X, Y)$

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

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

$E[g(X)]$

Linearity of expectations

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

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

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

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

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

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

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

$$= \sum_x x\, p_X(x) + \sum_y y\, p_Y(y) = E[X] + E[Y]$$

Linearity of expectations

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

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

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

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

The mean of the binomial

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

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

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

$X_i = 1$ if ith trial is a success; — p

$X_i = 0$ otherwise — $1-p$

(indicator variable)

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

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