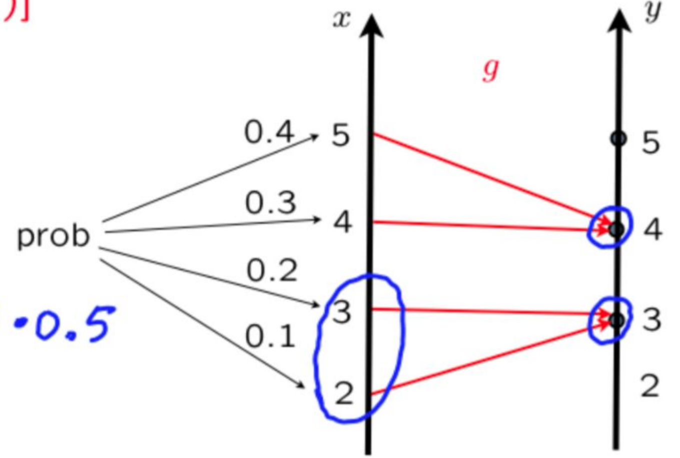
Expectation (continued), Variance

February 28 - March 2

(Annotated Slides)

The expected value rule, for calculating $\mathbf{E}[g(X)]$



- Let X be a r.v. and let $Y = g(X)$
- Averaging over y: $\mathbf{E}[Y] = \sum_y y\, p_Y(y)$

  $3\cdot(0.1+0.2)+4\cdot(0.3+0.4)$

- Averaging over x: $3\cdot 0.1 + 3\cdot 0.2 + 4\cdot 0.3 + 4\cdot 0.5$

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

Proof: $\sum_y \sum_{x:g(x)=y} g(x)\, p_X(x)$

$$= \sum_y \sum_{x:g(x)=y} y\, p_X(x) = \sum_y y \underbrace{\sum_{x:g(x)=y} p_X(x)}$$

$$= \sum_y y\, p_Y(y) = E[Y]$$

- $\mathbf{E}[X^2] = \sum_x x^2 p_X(x)$

  $g(x) = x^2$

- Caution: In general, $\mathbf{E}[g(X)] \neq g(\mathbf{E}[X])$

  $E[X^2] \neq (E[X])^2$

Linearity of expectation: $\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$

X = salary $\quad E[X]$ = average salary

Y = new salary $= 2X + 100 \qquad E[Y] = E[2X+100] = 2E[X] + 100$

- Intuitive

- **Derivation**, based on the expected value rule:

$g(x) = ax + b$

$Y = g(X)$

$$E[Y] = \sum_x g(x) p_X(x)$$

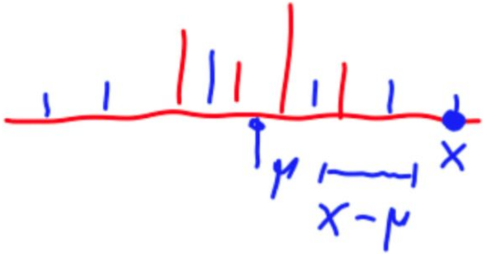
$$= \sum_x (ax+b) p_X(x) = a \sum_x x p_X(x) + b \underbrace{\sum_x p_X(x)}_{1}$$

$$= aE[X] + b$$

$E[g(X)] = g(E[X])$.

exceptional g

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:** $\text{var}(X) = \mathbf{E}[(X-\mu)^2]$ ≥ 0

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

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

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

Properties of the variance

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

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

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

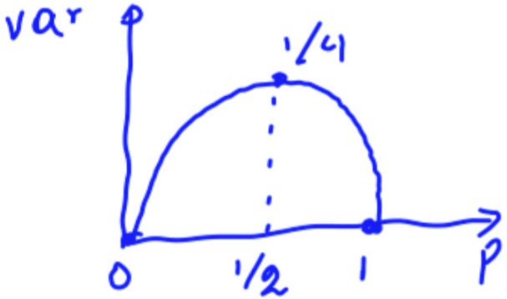
$\text{var}(X) = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2]$

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

Variance of the Bernoulli

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

$E[x] = p$



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

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

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

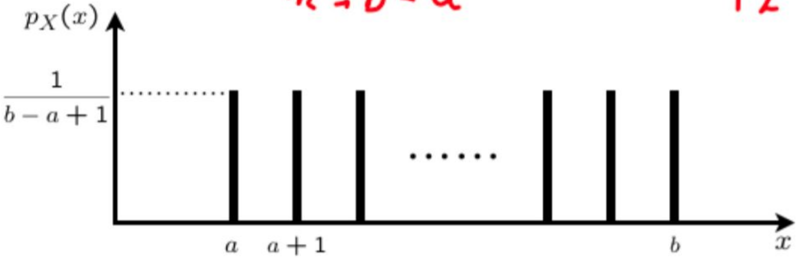
$X^2 = X$

Variance of the uniform

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

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

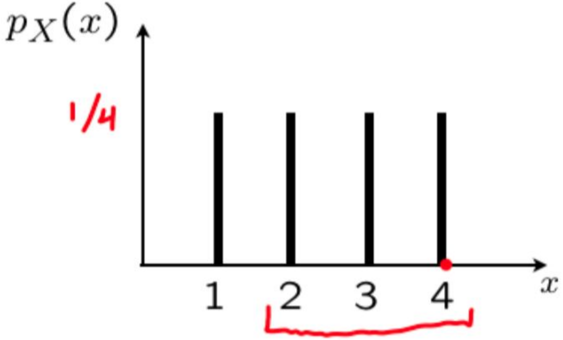
$n = b - a$

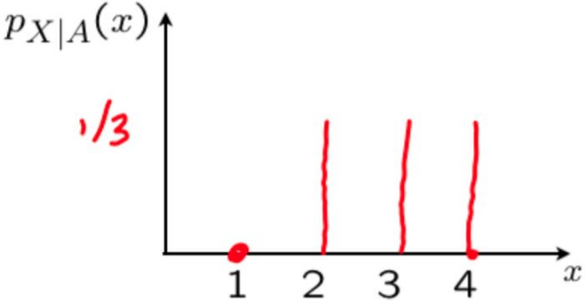


$$\mathrm{Var}(X) = \frac{1}{12}(b-a)(b-a+2)$$

Example of conditioning

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





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

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

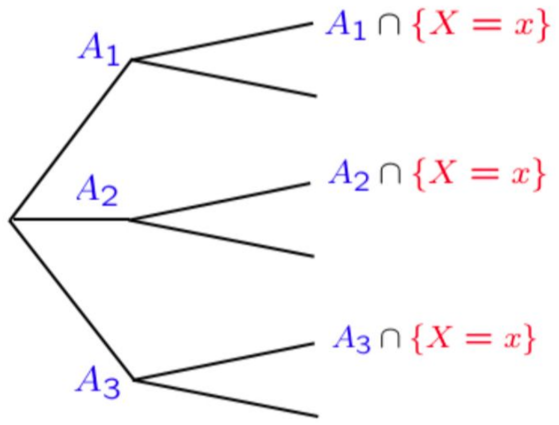
$\text{var}(X) = \frac{1}{12}(b-a)(b-a+2)$

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

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

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

Total expectation theorem



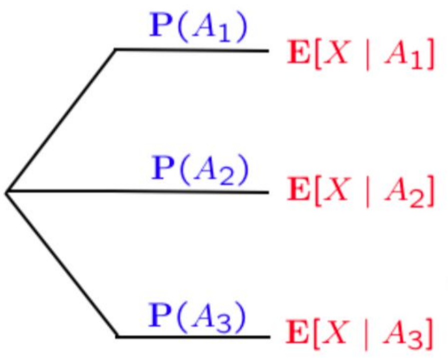
$$\mathbf{P}(B) = \mathbf{P}(A_1)\,\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\,\mathbf{P}(B \mid A_n)$$

$B = \{X = x\}$

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

for all x

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

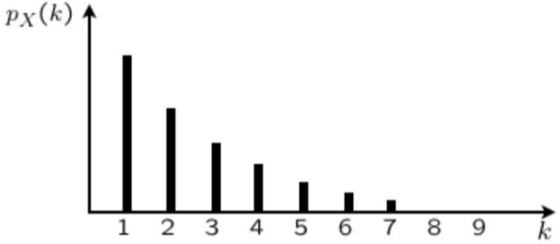


$$\boxed{\mathbf{E}[X] = \mathbf{P}(A_1)\,\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\,\mathbf{E}[X \mid A_n]}$$

Conditioning a geometric random variable

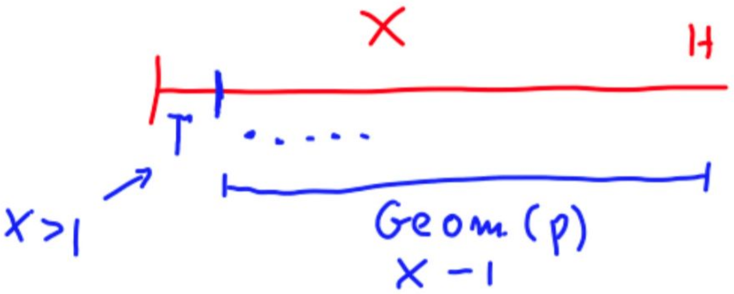
- X: number of independent coin tosses until first head; $\mathbf{P}(H) = p$

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



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

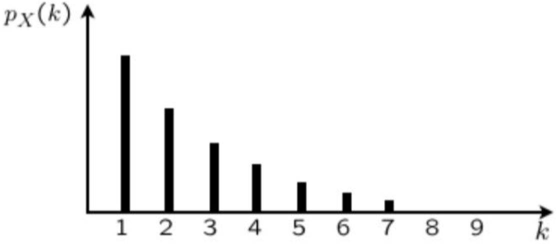
Conditioned on $X > 1$, $X - 1$ is geometric with parameter p



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

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

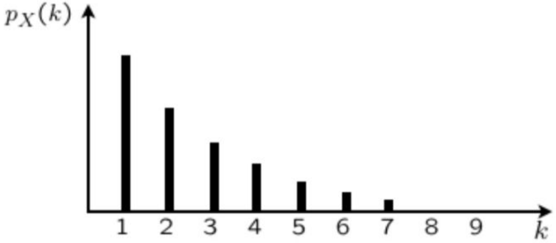
$$= (1-p)^2 p = p_X(3)$$

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

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

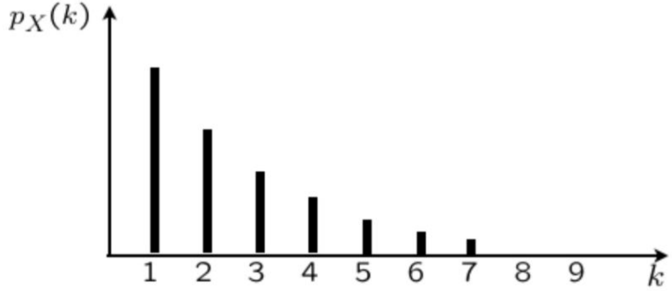
Conditioned on $X > n$, $X - n$ is geometric with parameter p

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

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

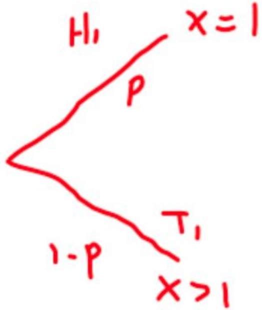
$$p_{X-1|X>1}(k) = p_X(k) = p_{X-n|X>n}(k).$$

The mean of the geometric



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

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



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

Multiple random variables and joint PMFs

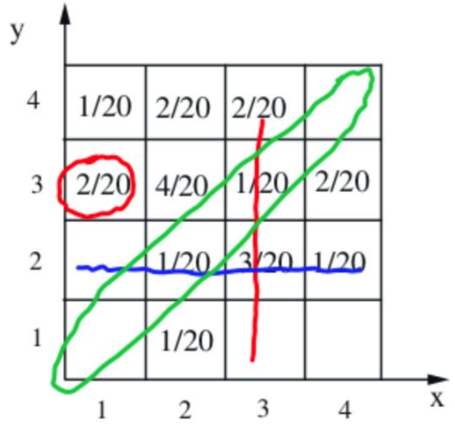
marginal pmfs

$X : p_X$

$Y : p_Y$

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

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



$P_X(3)$

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

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

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

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

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

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

Linearity of expectations

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

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

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

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

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

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

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

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

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

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

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

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

$X_i = 0$ otherwise — $1-p$

(indicator variable)

$X = X_1 + \cdots + X_n$

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

Independence

- of two events: $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$ $\qquad$ $\mathbf{P}(A \mid B) = \mathbf{P}(A)$

- of a r.v. and an event: $\mathbf{P}(X = x \text{ and } A) = \mathbf{P}(X = x) \cdot \mathbf{P}(A)$, for all x

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

- of two r.v.'s: $\mathbf{P}(X = x \text{ and } Y = y) = \mathbf{P}(X = x) \cdot \mathbf{P}(Y = y)$, for all x, y

$p_{X|Y}(x|y) = p_X(x)$ $\qquad$ $p_{X,Y}(x, y) = p_X(x)\, p_Y(y)$, for all x, y

$p_{Y|X}(y|x) = p_Y(y)$

X, Y, Z are **independent** if:
$p_{X,Y,Z}(x, y, z) = p_X(x)\, p_Y(y)\, p_Z(z)$, for all x, y, z

Example: independence and conditional independence

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		

1/3	1/9	2/9
2/3	2/9	4/9
	1/3	2/3

- Independent? No

$p_X(1) = 3/20$

$p_{X|Y}(1|1) = 0$

- What if we condition on $X \leq 2$ and $Y \geq 3$?

Yes

Independence and expectations

- In general: $\mathbf{E}[g(X,Y)] \neq g\big(\mathbf{E}[X], \mathbf{E}[Y]\big)$

always true

- Exceptions: $\mathbf{E}[aX+b] = a\mathbf{E}[X]+b$ $\quad\quad$ $\mathbf{E}[X+Y+Z] = \mathbf{E}[X]+\mathbf{E}[Y]+\mathbf{E}[Z]$

If X, Y are **independent:** $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$

$g(X)$ and $h(Y)$ are also independent: $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

$E[g(X,Y)] \quad g(x,y) = xy$

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

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

Independence and variances

- Always true: $\text{var}(aX) = a^2\text{var}(X)$ $\qquad$ $\text{var}(X + a) = \text{var}(X)$

- In general: $\text{var}(X + Y) \neq \text{var}(X) + \text{var}(Y)$

> If X, Y are independent: $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$

assume
$E[X] = E[Y] = 0$
$E[XY] = E[X]E[Y] = 0$

$$\text{var}(X+Y) = E[(X+Y)^2] = E[X^2 + 2XY + Y^2]$$
$$= E[X^2] + 2\underline{E[XY]} + E[Y^2] = \text{var}(X) + \text{var}(Y)$$

- Examples:
  - If $X = Y$: $\text{var}(X + Y) =$ $\text{var}(2X) = 4\text{var}(X)$
  - If $X = -Y$: $\text{var}(X + Y) =$ $\text{var}(0) = 0$
  - If X, Y independent: $\text{var}(X - 3Y) =$ $\text{var}(X) + \text{var}(-3Y) = \text{var}(X) + 9\text{var}(Y)$

Variance of the binomial

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

$X_i = 1$ if ith trial is a success; (indicator variable) independent

$X_i = 0$ otherwise

$X = X_1 + \cdots + X_n$

$\boxed{\mathrm{var}(X)} = \mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n)$

$= n \cdot \mathrm{var}(X_1) = \boxed{np(1-p)}$