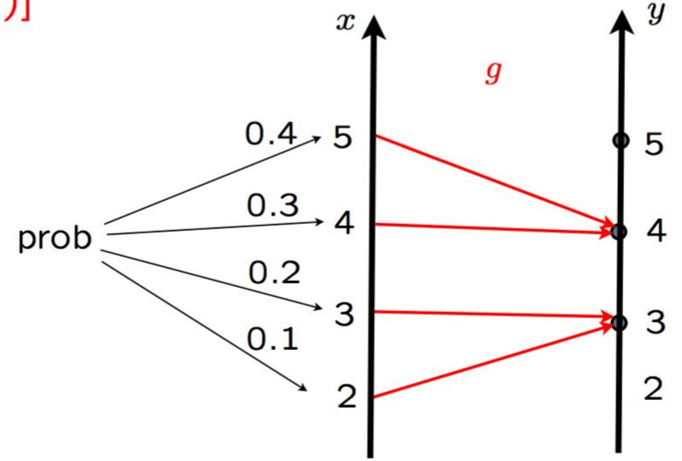
Expectation (continued), Variance

February 28 - March 2

(Clean 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: $\mathbf{E}[Y] = \sum_y y p_Y(y)$

- Averaging over x:

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

Proof:

- $\mathbf{E}[X^2] =$

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

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

- Intuitive

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

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?

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

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

$$\text{var}(X) =$$

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

Properties of the variance

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

- Notation: $\mu = \mathbf{E}[X]$
- Let $Y = X + b$
- Let $Y = aX$

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

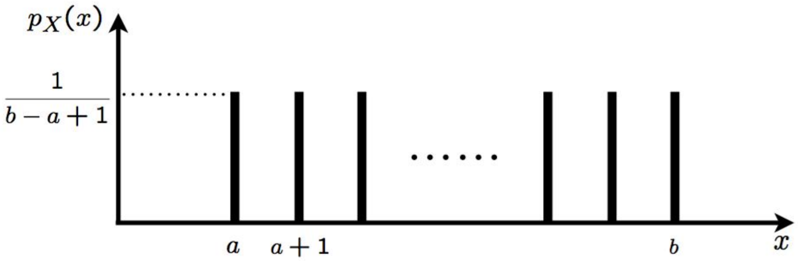
Variance of the Bernoulli

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

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

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

Variance of the uniform



Conditional PMF and expectation, given an event

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

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

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

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

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

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

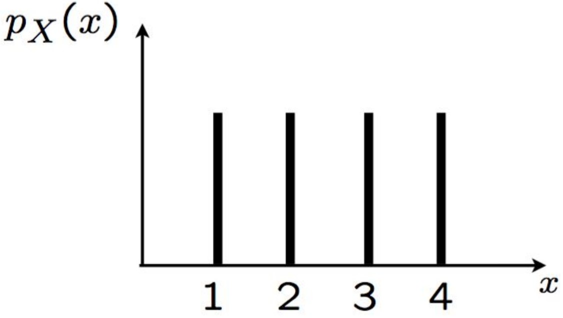
$$\mathbf{E}[X \mid A] = \sum_x x p_{X|A}(x)$$

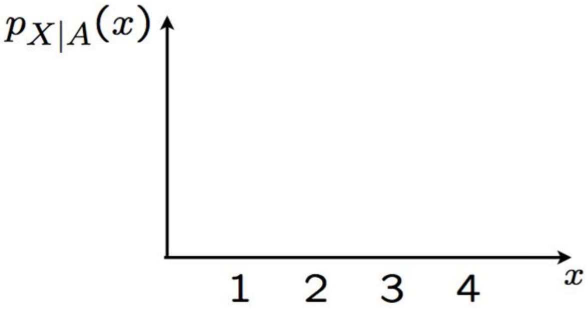
$$\mathbf{E}[g(X)] = \sum_x g(x)\, p_X(x)$$

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

Example of conditioning

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





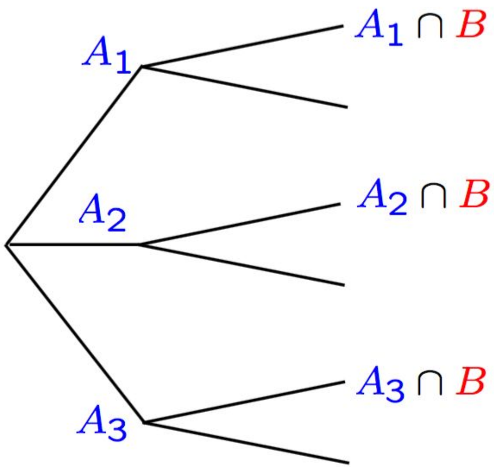
$\mathbf{E}[X] =$

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

$\text{var}(X) =$

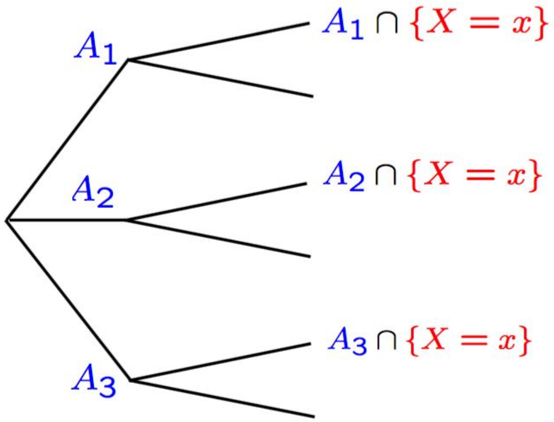
$\text{var}(X \mid A) =$

Total expectation theorem



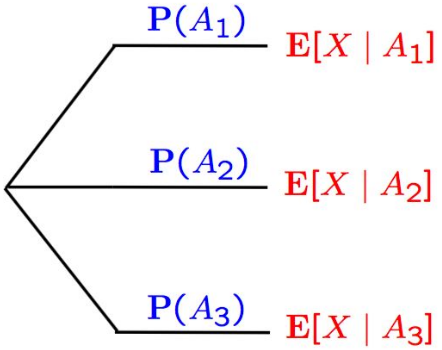
$$\mathbf{P}(B) = \mathbf{P}(A_1)\,\mathbf{P}(B \mid A_1) + \cdots + \mathbf{P}(A_n)\,\mathbf{P}(B \mid A_n)$$

Total expectation theorem



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

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

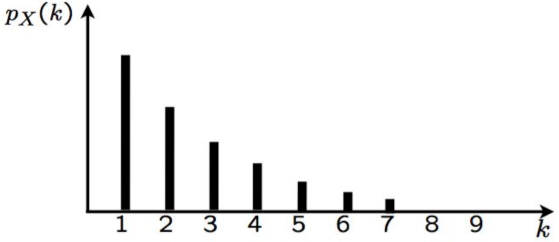


$$\boxed{\mathbf{E}[X] = \mathbf{P}(A_1)\,\mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\,\mathbf{E}[X \mid A_n]}$$

Conditioning a geometric random variable

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

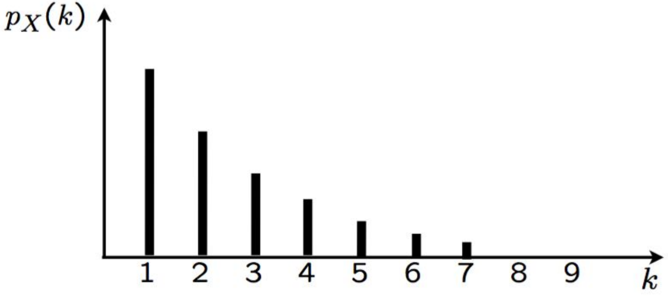
$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1, 2, \ldots$$



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

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

The mean of the geometric



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

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

Linearity of expectations

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

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

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

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

The mean of the binomial

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

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

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

$X_i = 1$ if ith trial is a success;
$X_i = 0$ otherwise

(indicator variable)

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

Independence

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

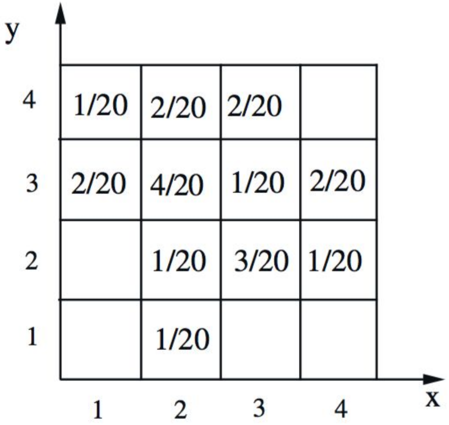
- of a r.v. and an event: $\mathbf{P}(X = x \text{ and } A) = \mathbf{P}(X = x) \cdot \mathbf{P}(A)$, for all x

- 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

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



- Independent?

- What if we condition on $X \le 2$ and $Y \ge 3$?

Independence and expectations

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

- 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: $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

Independence and variances

- Always true: $\text{var}(aX) = a^2\text{var}(X)$ $\quad$ $\text{var}(X + a) = \text{var}(X)$
- In general: $\text{var}(X + Y) \neq \text{var}(X) + \text{var}(Y)$

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

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

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)
$X_i = 0$ otherwise

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