Bernoulli and Binomial

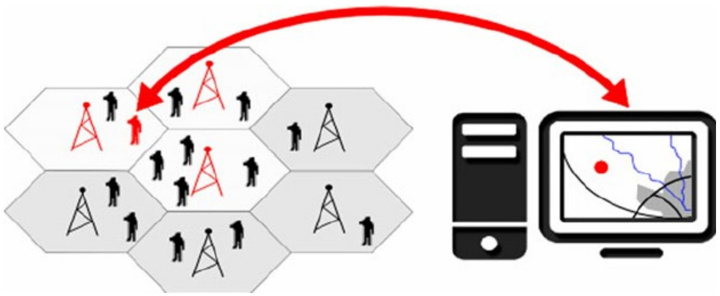
Will Monroe
July 10, 2017

with materials by
Mehran Sahami
and Chris Piech

image:
Antoine Taveneaux

Announcements: Problem Set 2

Due **this Wednesday**, 7/12, at 12:30pm (before class).



(Cell phone location sensing)

Announcements: Midterm

Two weeks from tomorrow:

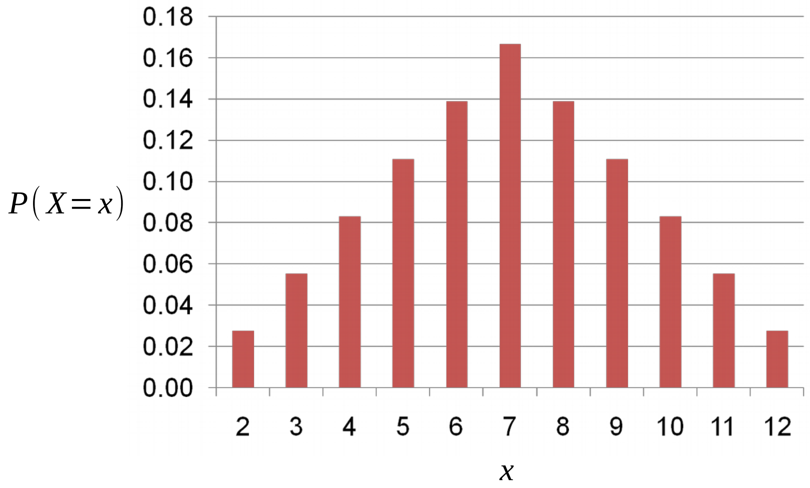
Tuesday, July 25, 7:00-9:00pm

Tell me by the end of this week if you have a conflict!

Review: Random variables

A **random variable** takes on values probabilistically.

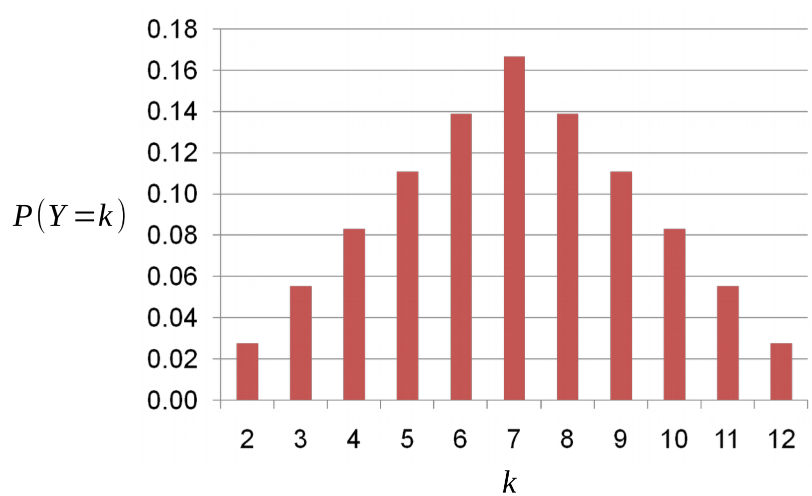
$$P(X=2)=\frac{1}{36}$$



Review: Probability mass function

The **probability mass function** (PMF) of a random variable is a function from values of the variable to probabilities.

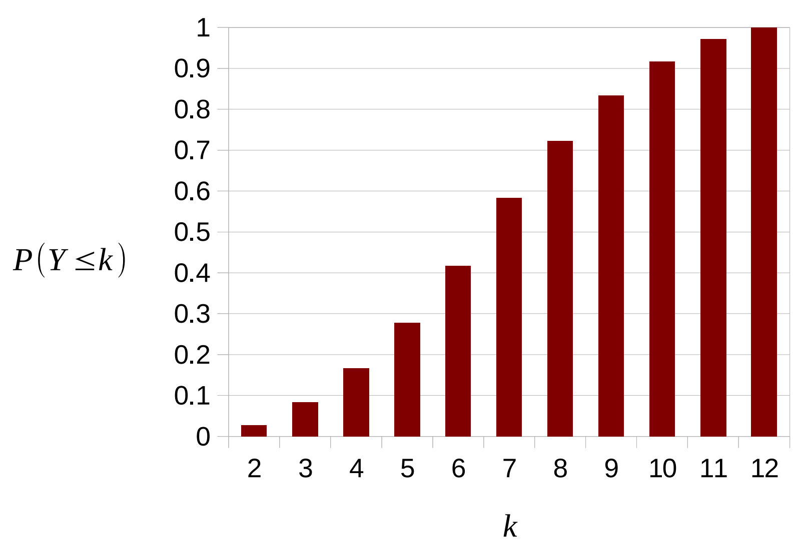
$$p_Y(k) = P(Y = k)$$



Review: Cumulative distribution function

The **cumulative distribution function** (CDF) of a random variable is a function giving the probability that the random variable is **less than or equal to** a value.

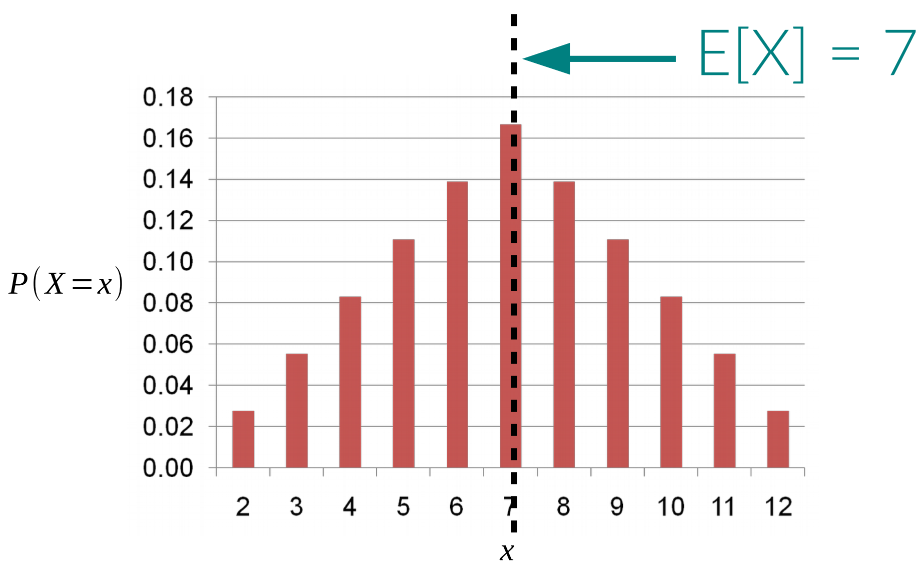
$$F_Y(k) = P(Y \leq k)$$



Review: Expectation

The **expectation** of a random variable is the “**average**” value of the variable (weighted by probability).

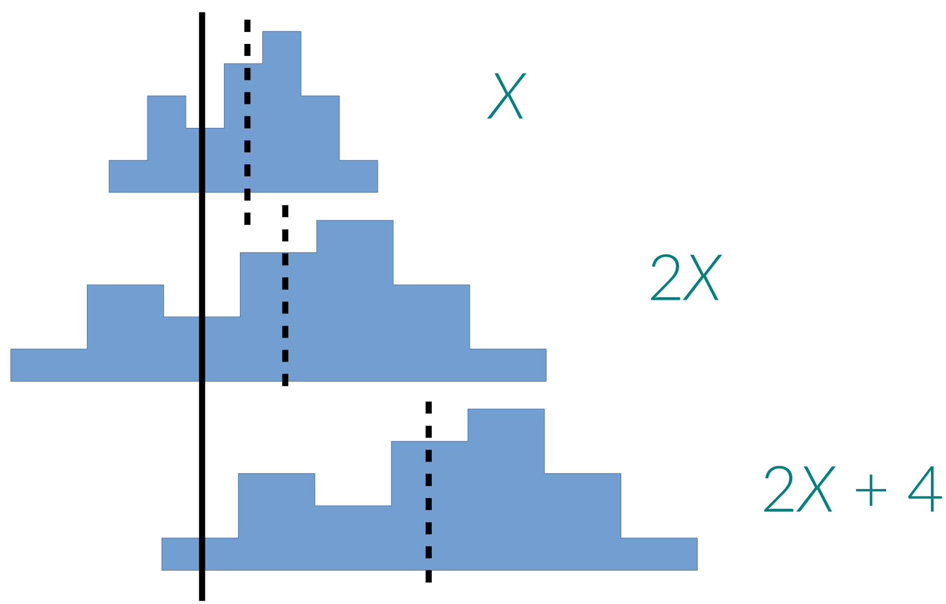
$$E[X] = \sum_{x:p(x)>0} p(x) \cdot x$$



Review: Linearity of expectation

Adding random variables or constants? **Add** the expectations. Multiplying by a <u>constant</u>? **Multiply** the expectation by the constant.

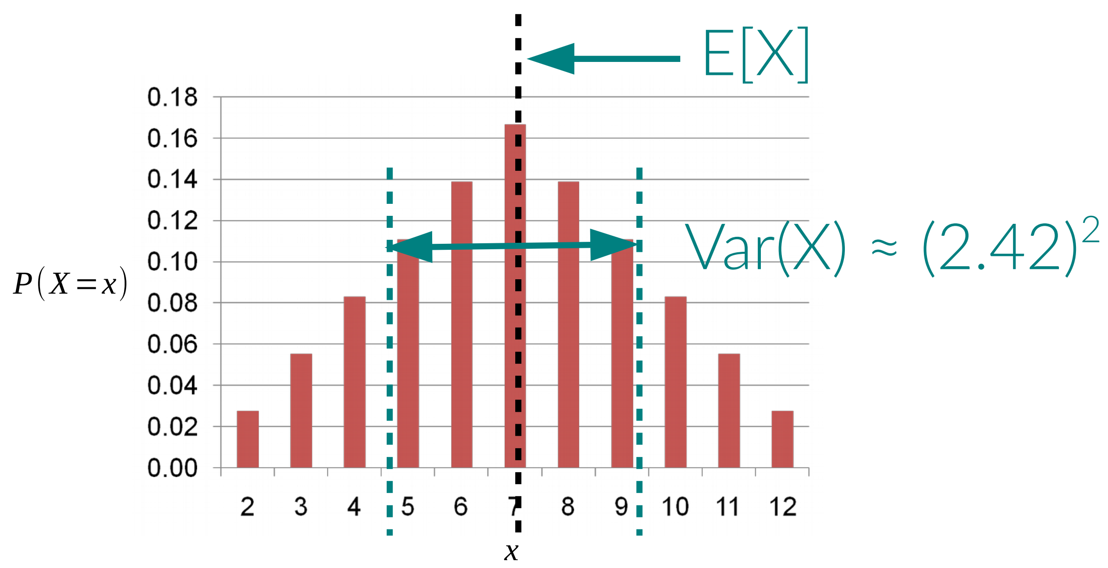
$$E[aX+bY+c]=aE[X]+bE[Y]+c$$



Review: Variance

Variance is the average **square** of the **distance** of a variable from the expectation. Variance measures the “**spread**” of the variable.

$$\mathrm{Var}(X) = E[(X - E[X])^2]$$
$$= E[X^2] - (E[X])^2$$



Today: Basic distributions

Many types of random variables come up repeatedly. Known frequently-occurring **distributions** lets you do computations without deriving formulas from scratch.

variable → X, family → Bin, parameters → n, p

$$X \sim \mathrm{Bin}(n, p)$$

We have ________ independent __________,
INTEGER / PLURAL NOUN

each of which _________ with probability
VERB ENDING IN -S

_________. How many of the _________
REAL NUMBER / REPEAT PLURAL NOUN

_________?
REPEAT VERB -S

Bernoulli random variable

An indicator variable (a possibly biased coin flip) obeys a **Bernoulli distribution**. Bernoulli random variables can be 0 or 1.

$$X \sim \mathrm{Ber}(p)$$

$$p_X(1) = p$$
$$p_X(0) = 1 - p \qquad \text{(0 elsewhere)}$$

Review: Indicator variable

An **indicator variable** is a “Boolean” variable, which takes values 0 or 1 corresponding to whether an event takes place.

$$I = \mathbb{1}[A] = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

Bernoulli: Fact sheet

$$X \sim \text{Ber}(p)$$

probability of "success" (e.g., heads)

Program crashes

Run a program, crashes with prob. p, works with prob. $(1 - p)$

X: 1 if program crashes

$P(X = 1) = p$
$P(X = 0) = 1 - p$

$X \sim \text{Ber}(p)$

Ad revenue

Serve an ad, clicked with prob. p, ignored with prob. $(1 - p)$

C: 1 if ad is clicked

$P(C = 1) = p$

$P(C = 0) = 1 - p$

$C \sim \text{Ber}(p)$

Bernoulli: Fact sheet

$$X \sim \mathrm{Ber}(p)$$

?

probability of "success" (heads, ad click, ...)

PMF:

$$p_X(1) = p$$
$$p_X(0) = 1 - p \qquad \text{(0 elsewhere)}$$

expectation:

image (right): Gabriela Serrano

Expectation of an indicator variable

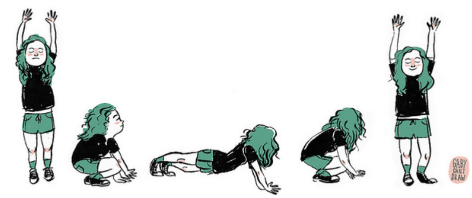
$$I = \mathbb{1}[A] = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

$$E[I] = P(A) \cdot 1 + [1 - P(A)] \cdot 0$$

$$= P(A)$$

Bernoulli: Fact sheet

$$X \sim \text{Ber}(p)$$



?

probability of "success" (heads, ad click, ...)

PMF:

$$p_X(1) = p$$
$$p_X(0) = 1 - p \qquad \text{(0 elsewhere)}$$

expectation: $E[X] = p$

variance:

image (right): Gabriela Serrano

Variance of a Bernoulli RV

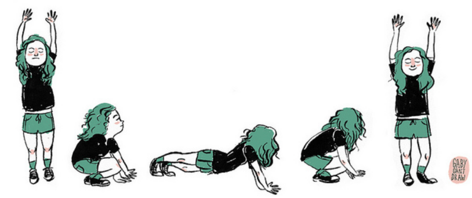
$$p_X(1) = p$$
$$p_X(0) = 1 - p \qquad \text{(0 elsewhere)}$$

$$E[X^2] = p \cdot 1^2 + (1-p) \cdot 0^2$$
$$= p$$

$$\text{Var}(X) = E[X^2] - (E[X])^2$$
$$= p - (p)^2$$
$$= p \cdot (1-p)$$

Bernoulli: Fact sheet

$$X \sim \text{Ber}(p)$$



?

probability of “success” (heads, ad click, ...)

PMF:

$$p_X(1) = p$$

$$p_X(0) = 1 - p \qquad \text{(0 elsewhere)}$$

expectation: $E[X] = p$

variance: $\text{Var}(X) = p(1-p)$

image (right): Gabriela Serrano

Jacob Bernoulli

Swiss mathematician (1654-1705)

Came from a family of competitive mathematicians (!)

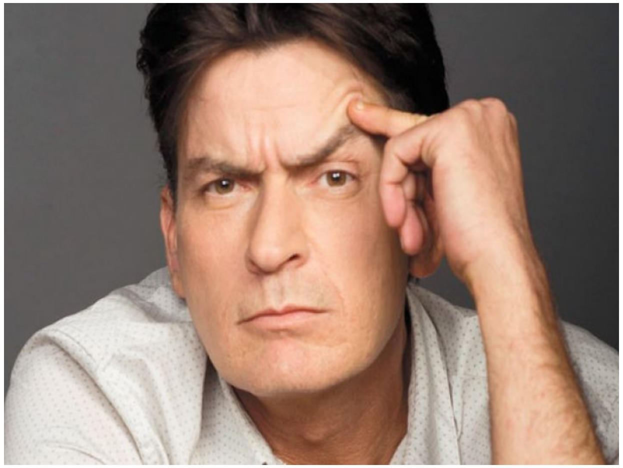


image (right): The Gazette Review

Jacob Bernoulli

Swiss mathematician (1654-1705)

Came from a family of competitive mathematicians (!)

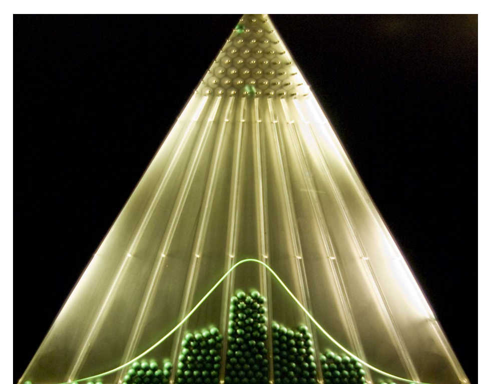
image (right): The Gazette Review

Binomial random variable

The **number of heads** on *n* (possibly biased) coin flips obeys a **binomial distribution**.

$$X \sim \text{Bin}(n, p)$$

$$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{if } k \in \mathbb{N}, 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}$$



Binomial: Fact sheet

number of trials (flips, program runs, ...)

$$X \sim \text{Bin}(n, p)$$

probability of “success” (heads, crash, ...)

Program crashes

n runs of program, each crashes with prob. p, works with prob. $(1 - p)$

H: number of crashes

$$P(H = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

H ~ Bin(n, p)

Ad revenue

n ads served, each clicked with prob. *p*, ignored with prob. (1 – *p*)

H: number of clicks

$$P(H = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

<u>H ~ Bin(*n*, *p*)</u>

Binomial: Fact sheet

number of trials (flips, program runs, ...)

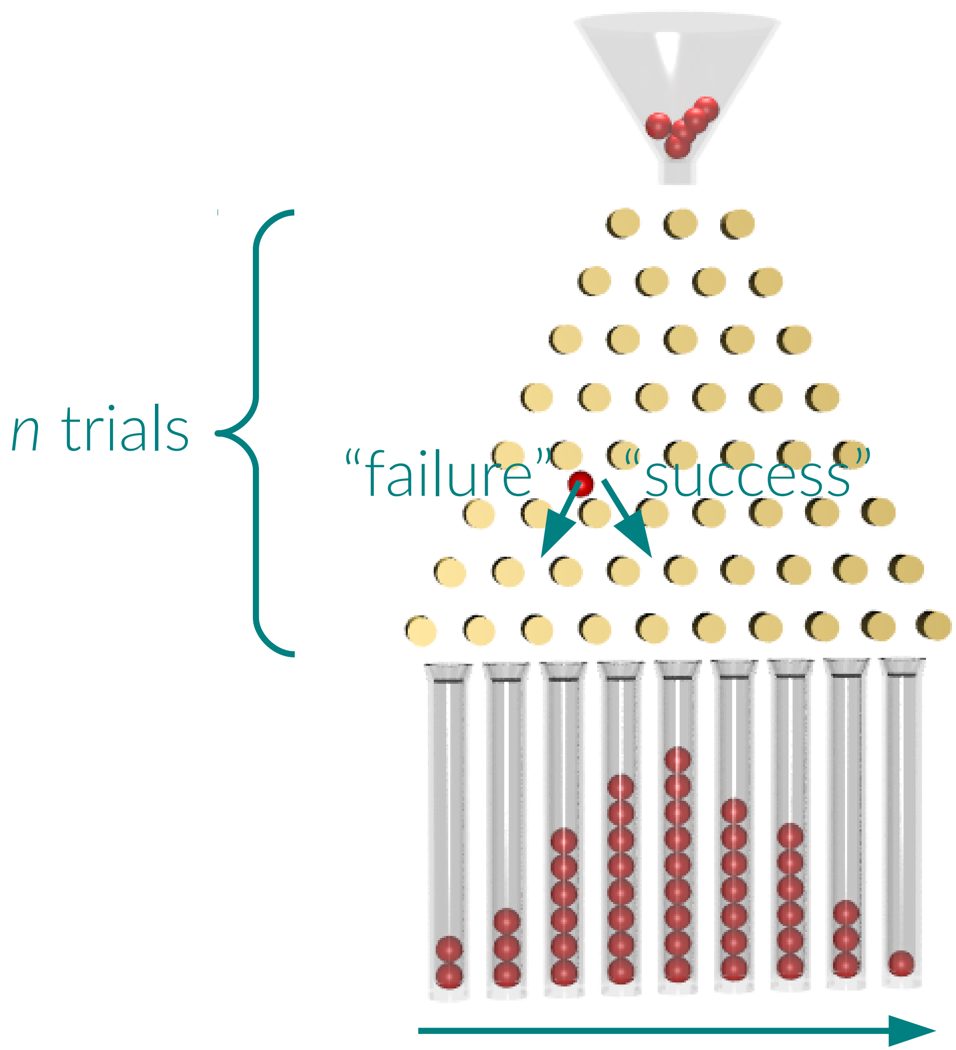
$$X \sim \text{Bin}(n, p)$$

probability of "success" (heads, crash, ...)

PMF:

$$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{if } k \in \mathbb{N}, 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}$$

The Galton board

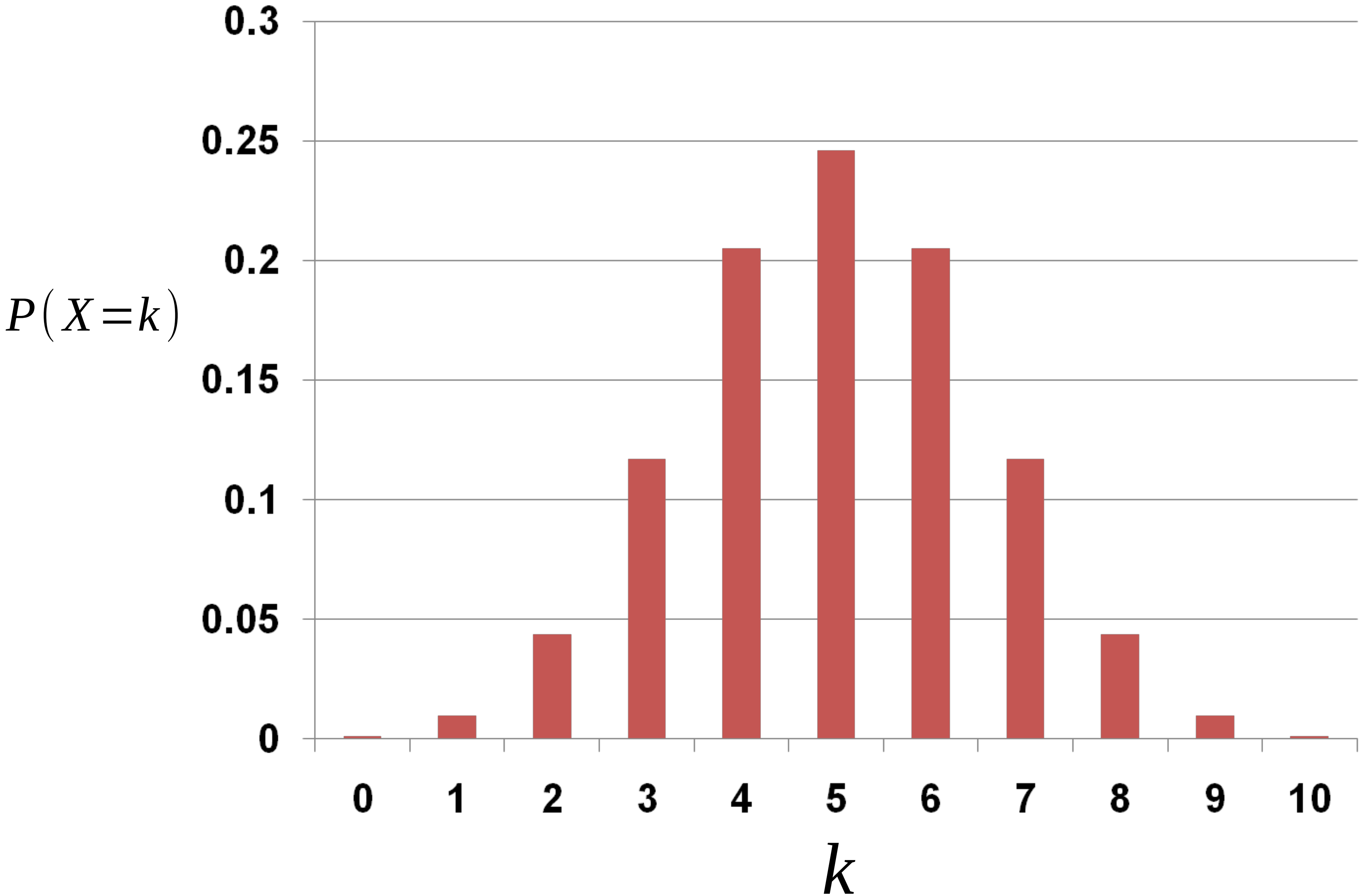


$$X \sim \mathrm{Bin}(n, 0.5)$$

X = final position = number of "successes"

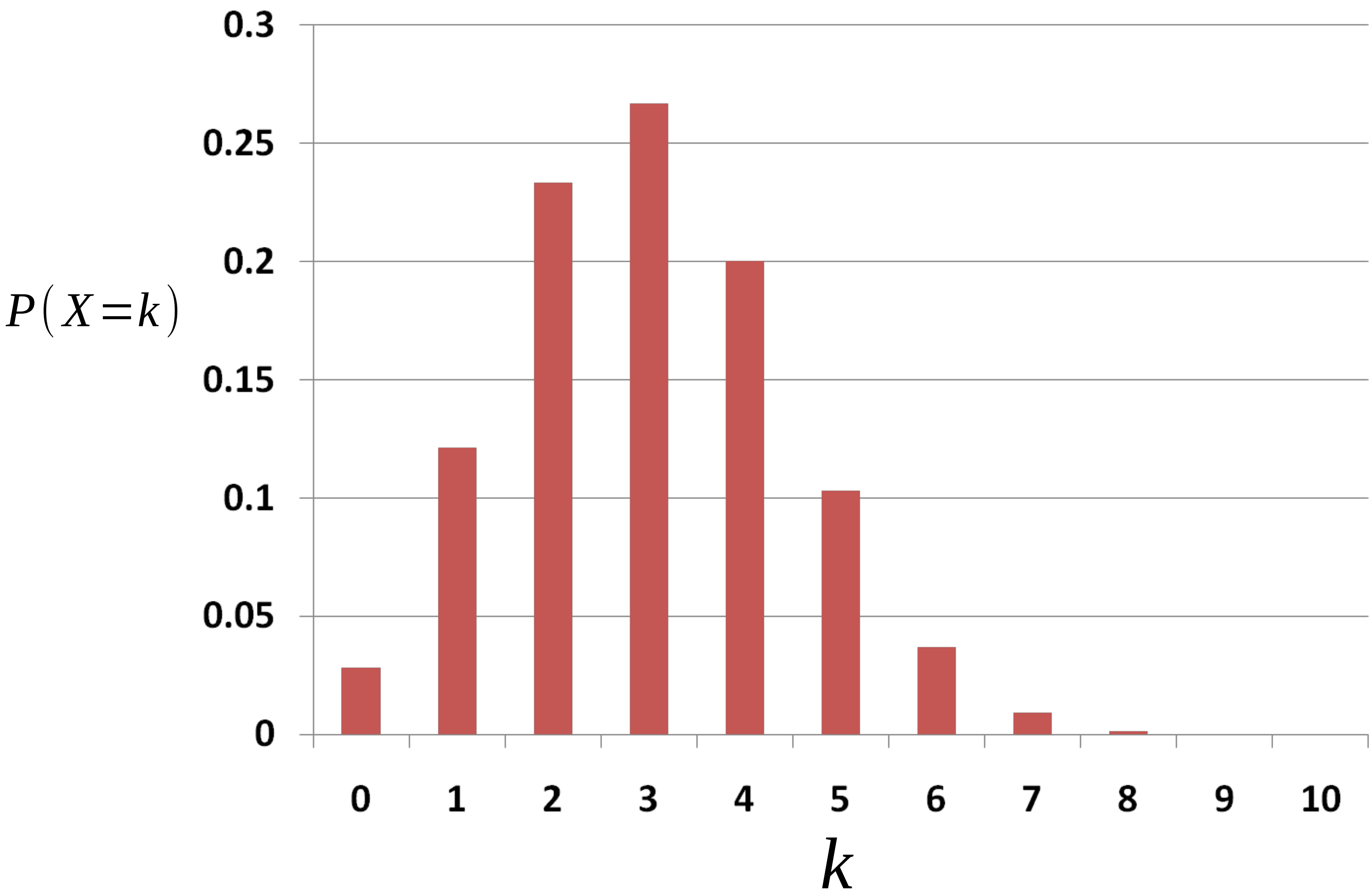
PMF of Binomial

$$X \sim \text{Bin}(10, 0.5)$$



PMF of Binomial

$$X \sim \text{Bin}(10, 0.3)$$



Break time!

Binomial: Fact sheet

number of trials (flips, program runs, ...)

$$X \sim \text{Bin}(n, p)$$

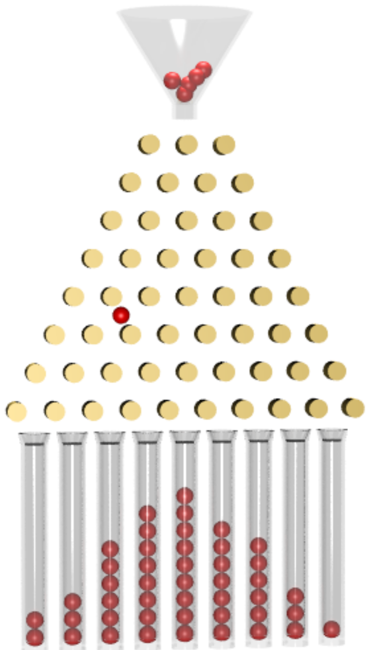
probability of "success" (heads, crash, ...)

PMF:

$$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{if } k \in \mathbb{N}, 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}$$

expectation:

Expectation of a binomial



$$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{if } k \in \mathbb{N}, 0 \le k \le n \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \sum_{k=0}^{n} P(X=k) \cdot k$$

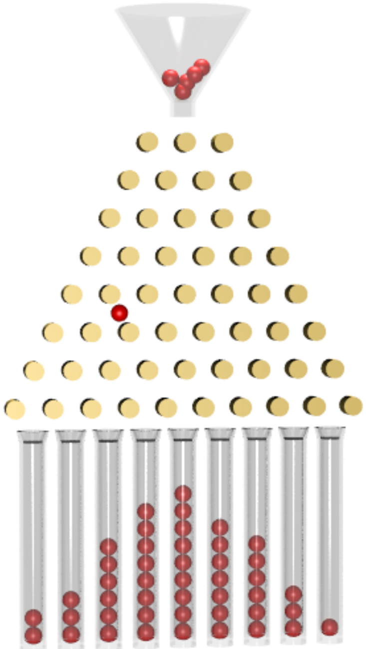
$$= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} k$$

$$= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} n \cdot \binom{n-1}{k-1} p \cdot p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = (p+1-p)^{n-1} = np$$

Expectation of a binomial



X = number of "successes"
X_i = indicator variable for success on i-th trial

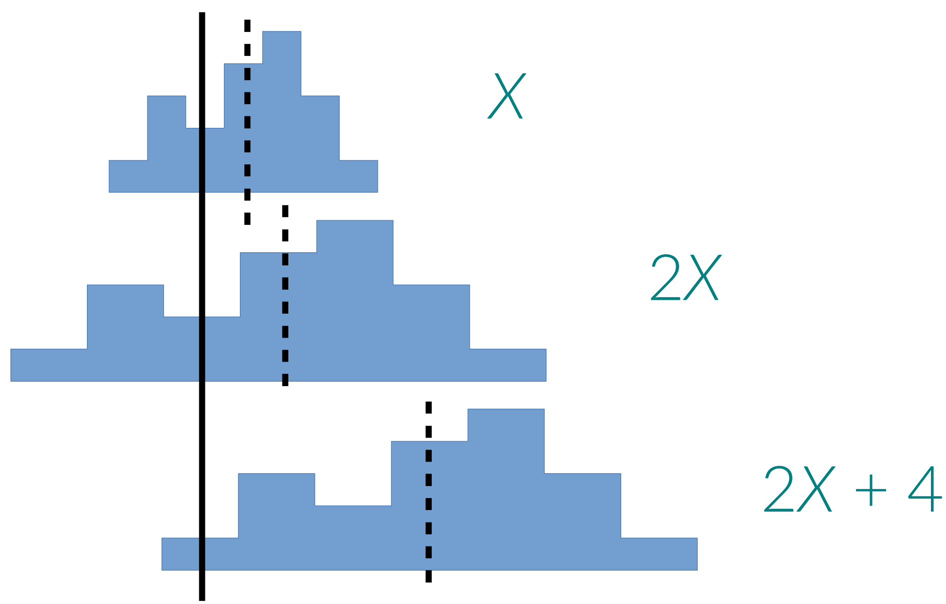
$$X = \sum_{i=1}^{n} X_i \qquad X_i \sim \mathrm{Ber}(p)$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right]$$

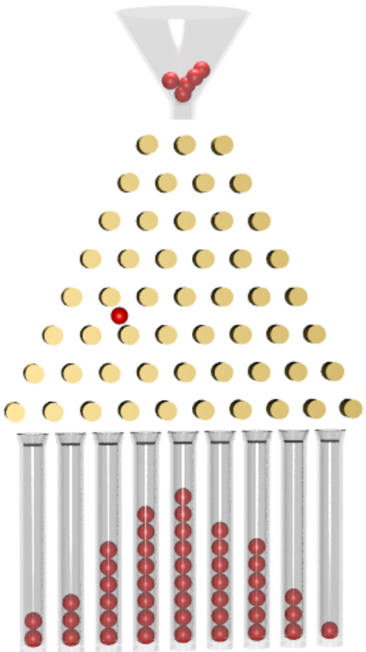
Review: Linearity of expectation

Adding random variables or constants? **Add** the expectations. Multiplying by a <u>constant</u>? **Multiply** the expectation by the constant.

$$E[aX+bY+c]=aE[X]+bE[Y]+c$$



Expectation of a binomial



X = number of "successes"
X_i = indicator variable for success on i-th trial

$$X = \sum_{i=1}^{n} X_i \qquad X_i \sim \mathrm{Ber}(p)$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

$$= \sum_{i=1}^{n} p$$

$$= np$$

Binomial: Fact sheet

number of trials (flips, program runs, ...)

$$X \sim \text{Bin}(n, p)$$

probability of "success" (heads, crash, ...)

PMF:

$$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{if } k \in \mathbb{N}, 0 \leq k \leq n \\ 0 & \text{otherwise} \end{cases}$$

expectation: $E[X] = np$

variance: $\text{Var}(X) = np(1-p)$

note: $\text{Ber}(p) = \text{Bin}(1, p)$

Eye color

Parents each have one brown (B) and one blue (b) gene.*
Brown is dominant: Bb → brown eyes.

Parents have 4 children.
X: number of children with brown eyes

$$E[X]=np=4\cdot 0.75=3$$

$$X\sim \mathrm{Bin}(4,0.75)$$

*Don't get your genetics information from CS 109!
Eye color is influenced by more than one gene.

Eye color

Parents each have one brown (B) and one blue (b) gene.*
Brown is dominant: Bb → brown eyes.

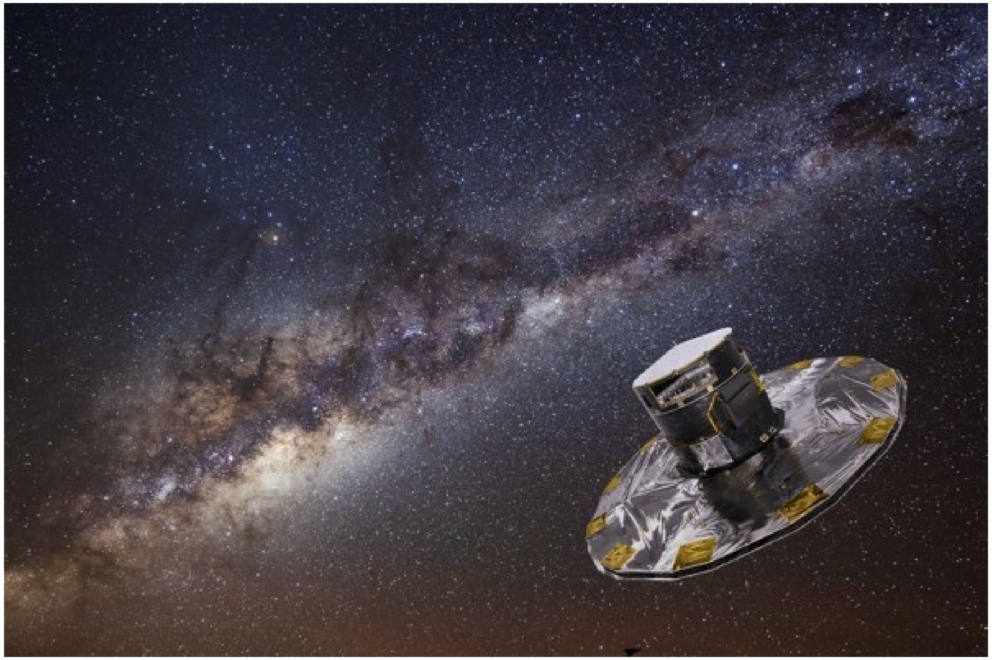
Parents have 4 children.
X: number of children with brown eyes

$$P(X=3)=\binom{4}{3}(0.75)^3(0.25)^1=4\cdot\frac{3^3}{4^4}\approx 0.422$$

$$X\sim\text{Bin}(4,0.75)$$

*Don't get your genetics information from CS 109!
Eye color is influenced by more than one gene.

Sending satellite messages



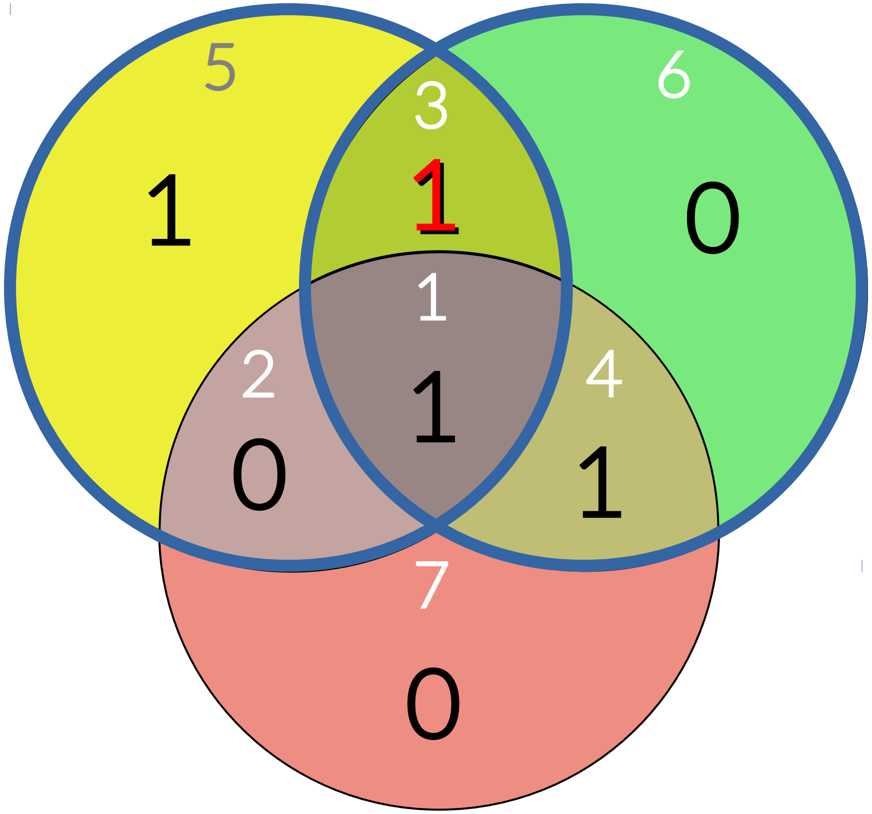
Sending a 4-bit message through space. Each bit corrupted (flipped) with probability p = 0.1.

X: number of bits flipped
X ~ Bin(4, 0.1)

(bit flip = "success".
not much of a success!)

$$P(X=0)=\binom{4}{0}(0.1)^0(0.9)^{4-0}$$
$$=(0.9)^4$$
$$\approx 0.656$$

Hamming codes

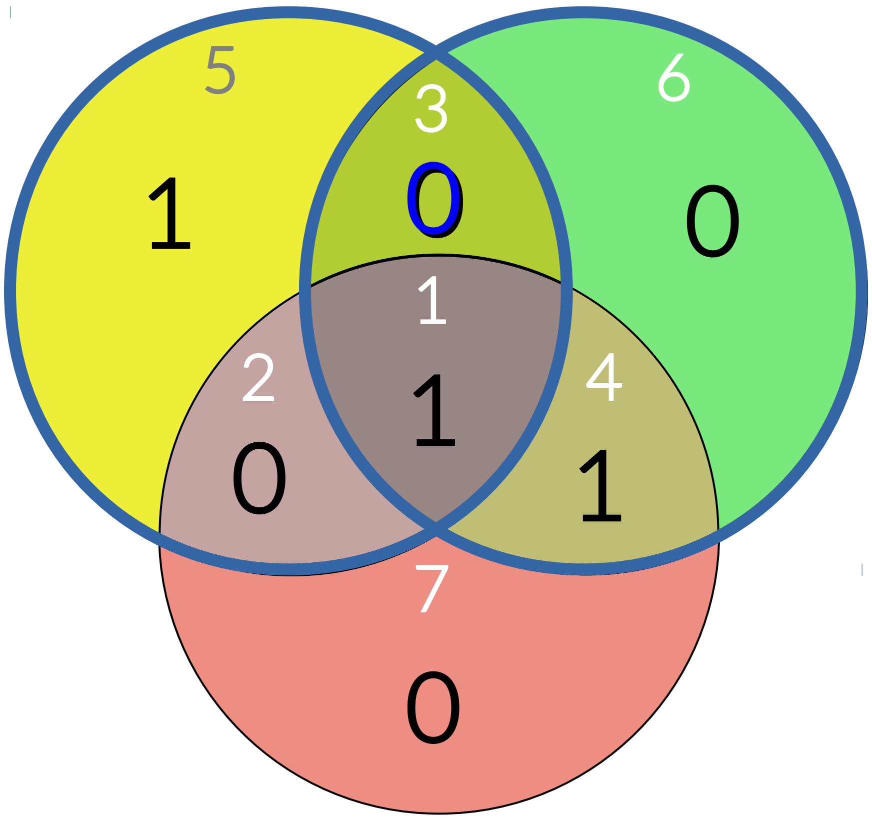


Message: 1 0 0 1

Send as: 1 0 0 1 1 0 0

Receive: 1 0 1 1 1 0 0

Hamming codes



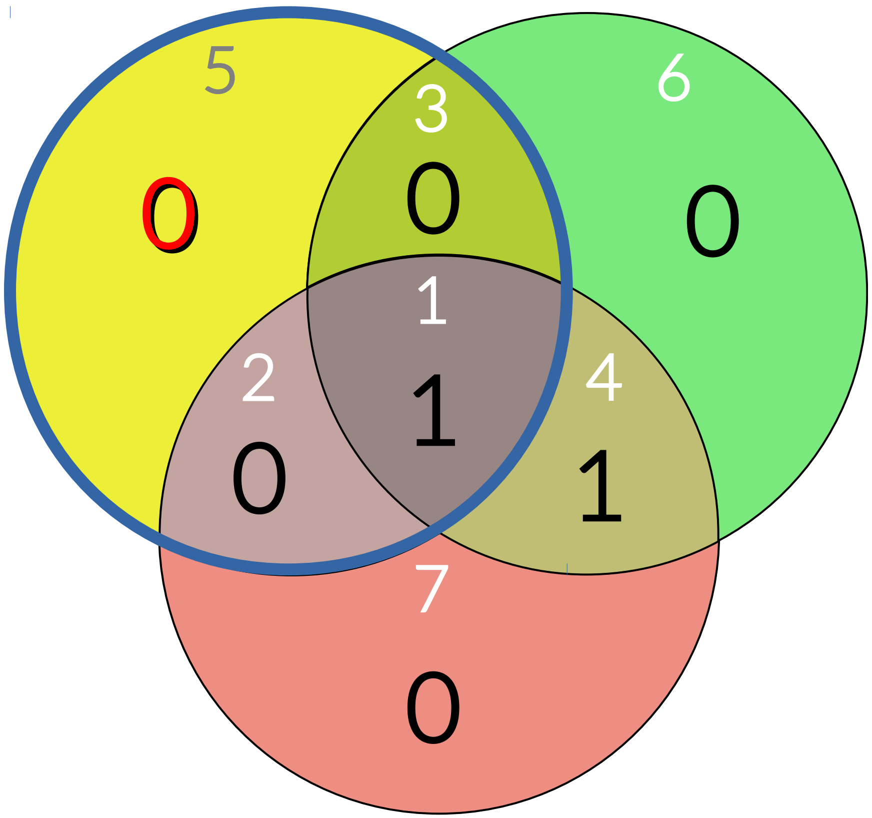
Message: 1 0 0 1

Send as: 1 0 0 1 1 0 0

Receive: 1 0 1 1 1 0 0

Correct to: 1 0 0 1 1 0 0

Hamming codes

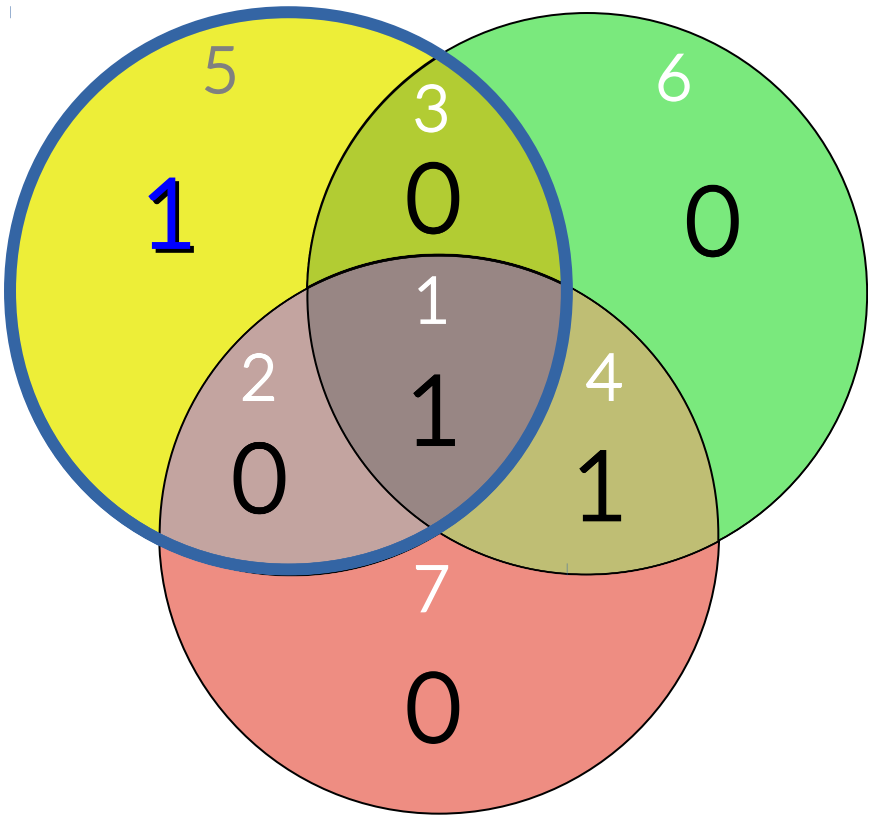


Message: 1 0 0 1

Send as: 1 0 0 1 1 0 0

Receive: 1 0 0 1 0 0 0

Hamming codes



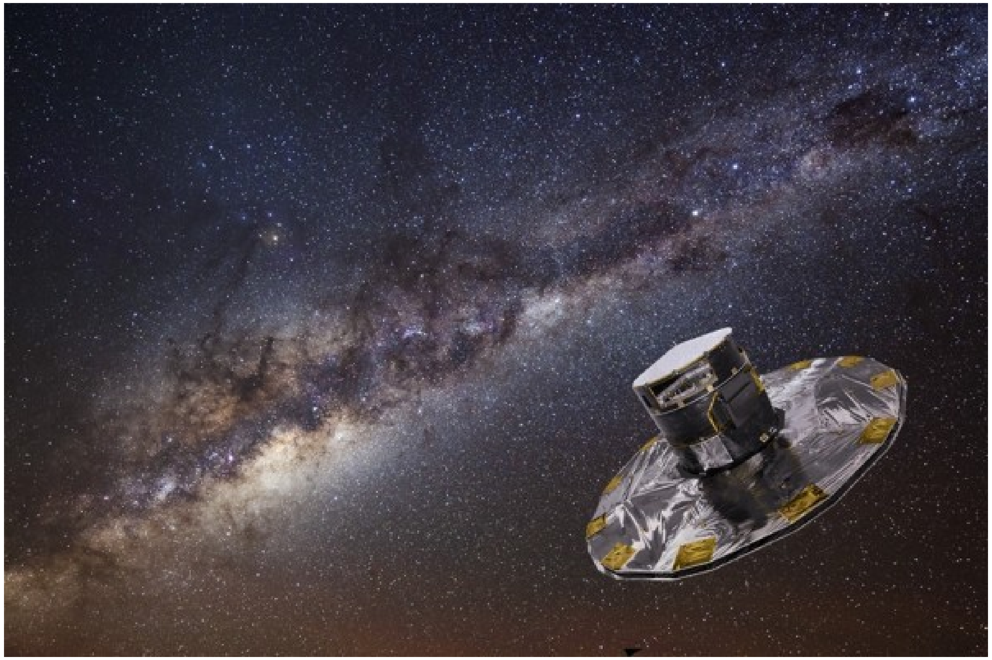
Message: 1 0 0 1

Send as: 1 0 0 1 1 0 0

Receive: 1 0 0 1 0 0 0

Correct to: 1 0 0 1 1 0 0

Sending satellite messages



Sending a 4-bit message through space. Each bit corrupted (flipped) with probability p = 0.1.

X: number of bits flipped
X ~ Bin(4, 0.1)

$$P(X \leq 1) = P(X=0) + P(X=1)$$
$$= \binom{7}{0}(0.1)^0(0.9)^{7-0} + \binom{7}{1}(0.1)^1(0.9)^{7-1}$$
$$= (0.9)^7 + 7 \cdot (0.1) \cdot (0.9)^6$$
$$\approx 0.478 + 0.372 = 0.850$$