

# Distribution Families

Many distribution families

Most natural and important

Theoretical and Practical significance

# Distributions

Discrete

Bernoulli

Binomial

Poisson

Geometric

Continuous

Uniform

Exponential

Normal

# Discuss

Motivation

Applications

Formulate

Visualize

Examples

Properties

$\mu$

$V$

$\sigma$

Python

Plot and experiment

# Show Distribution

Nonnegative

Sum to 1

Blendtec

Tom Dickson



# Bernoulli Distribution

Simplest non-constant distribution

Foundation of many others

$$\mu \quad V \quad \sigma$$

Repeated experiments



# Jacob Bernoulli, 1655-1705

$$\frac{1}{n}(x_1 + \dots + x_n) \rightarrow E(X)$$

Theology → mathematics

Calculus      Integrals

“Euler” number

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

e → b

Ars Conjectandi

First law of large numbers

Mentored brother Johann

Medicine → Math      Dynasty

# The simplest Distribution

Simplest

One value

5

Constant

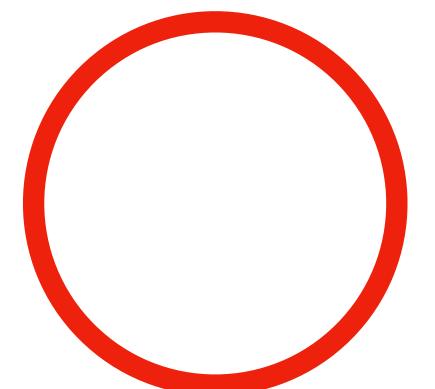
Always same

Trivial

Simplest non-trivial

Two values

Simplest values



0 and 1



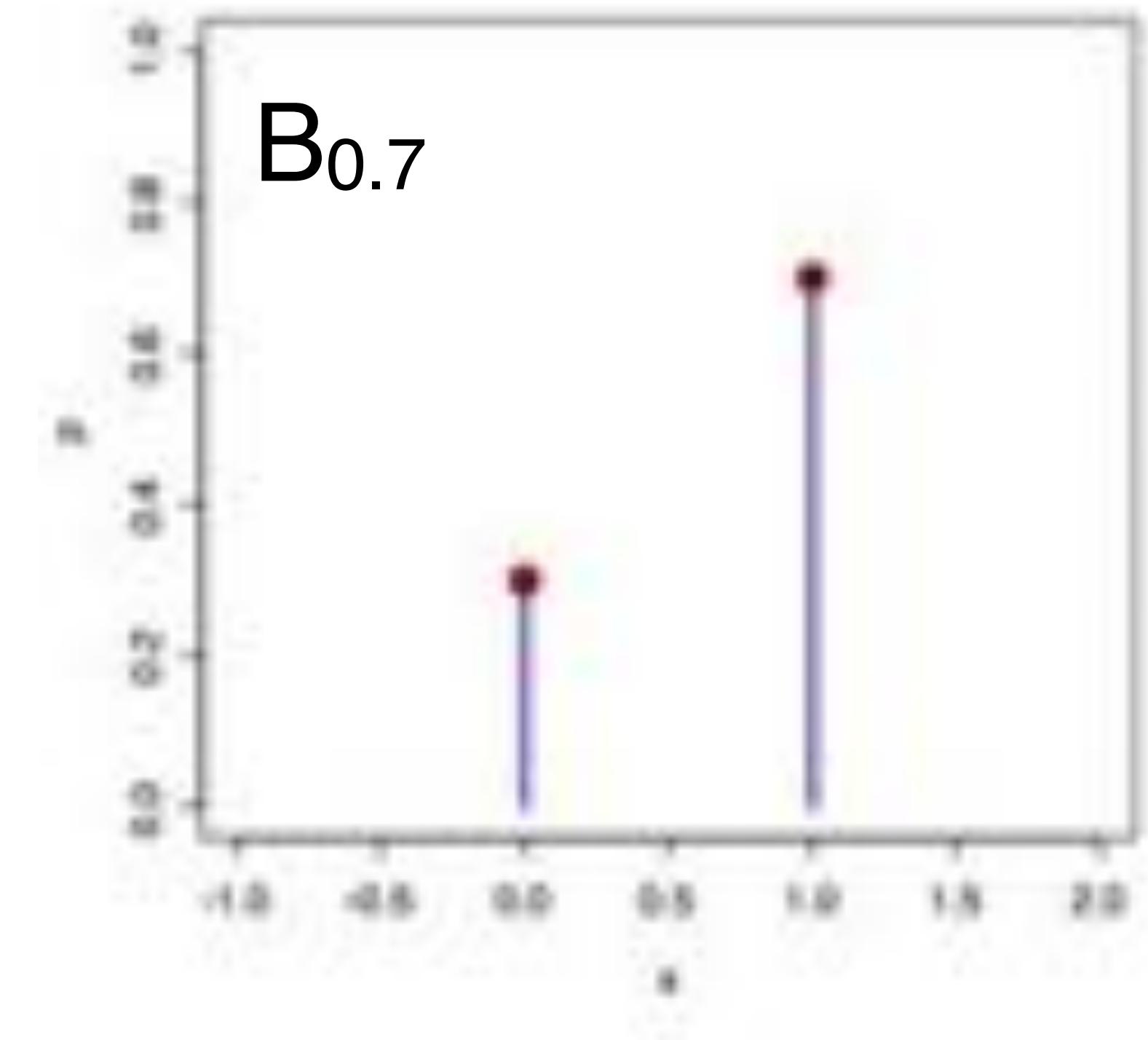
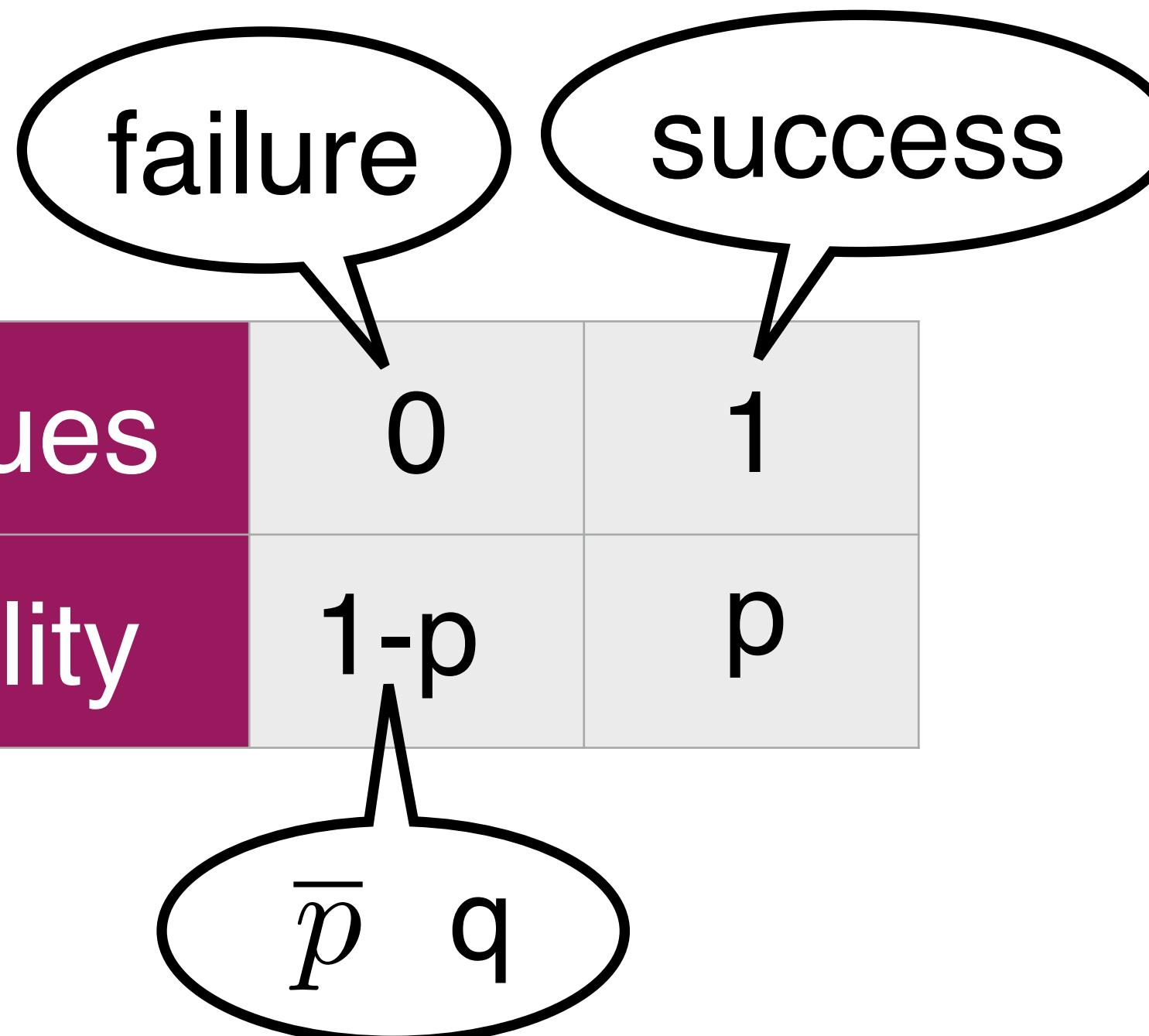
Bernoulli Coin!

# Bernoulli Distribution

$B_p$

$0 \leq p \leq 1$

Two values	0	1
Probability	$1-p$	$p$



$\Sigma$  WILL IT ADD?

$$p(0) + p(1) = (1-p)+p = 1$$

YES IT  
ADDS!

$X \sim B_p$

Bernoulli

random variable, coin, experiment, trial

# Who Cares About Two Values?

Binary version of complex events

Everyone!

Products: 80 good, 20 defective

Select one, good or not

$\sim B_{.8}$

Next child will be a boy

$\sim B_{.5}$

Generalizes to more complex variables

Patient has one of three diseases

Repeated trials yield # successes

Many important distributions

Binomial, Geometric, Poisson, Normal

# Mean

$$X \sim B_p$$

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

$$p(1) = p$$

$$EX = \sum p(x) \cdot x = (1-p) \cdot 0 + p \cdot 1 = p$$

$$X \sim B_{0.8}$$

$$EX = 0.8$$

$$EX = P(X=1)$$

Fraction of times expect to see 1

# Variance

$X \sim B_p$

$EX = p$

Variance

Easy route

$0^2 = 0$

$1^2 = 1$

$X^2 = X$

$E(X^2) = EX = p$

$$V(X) = E(X^2) - (EX)^2 = p - p^2 = p(1-p) = pq$$

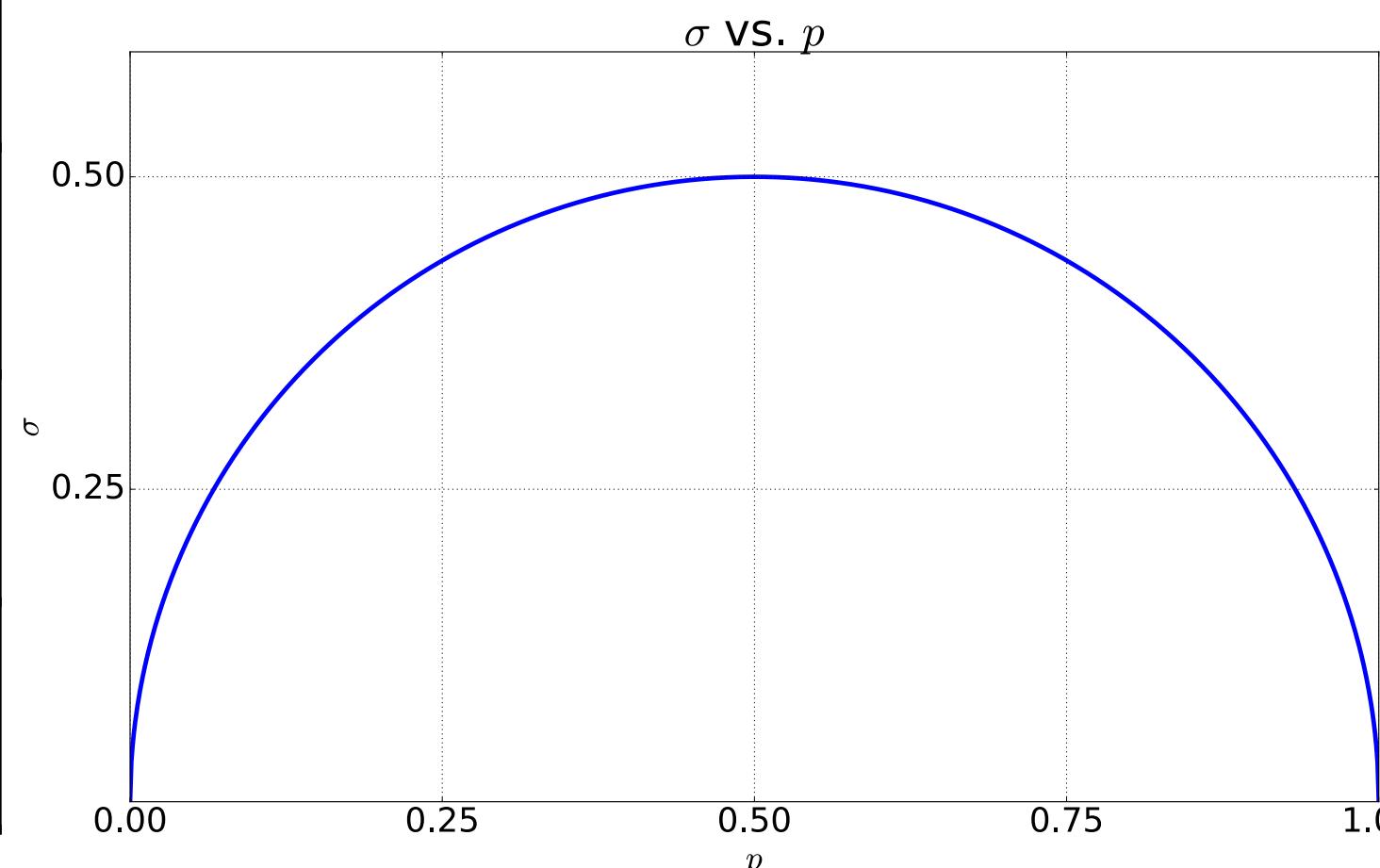
Standard Deviation

$$\sigma = \sqrt{pq}$$

$B_p$  varies most  
when  $p = \frac{1}{2}$

$E, V, \sigma$  for  
various  $p$

$p$	$EX$	$V(X)$	$\sigma$
0	0	0	0
1	1	0	0
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$



# Independent Trials

Much of  $B_p$  importance stems from multiple trials

Most common

Independent

⊤

$$0 \leq p \leq 1$$

$$X_1, X_2, X_3 \sim B_p$$

⊤

$$q \stackrel{\text{def}}{=} 1-p$$

$$P(110) = p^2 q = P(101) = P(011)$$

Generally

$$X_1, X_2, \dots, X_n \sim B_p$$

⊤

$$x^n = x_1, x_2, \dots, x_n \in \{0,1\}^n$$

$n_0$  0's and  $n_1$  1's

$$P(x_1, \dots, x_n) = p^{n_1} q^{n_0}$$

$$P(10101) = p^{n_1} q^{n_0} = p^3 q^2$$



# Typical Samples

Distribution	Typical seq.	Description	Probability
$B_0$	0000000000	constant 0	$1^{10} = 1$
$B_1$	1111111111	constant 1	$1^{10} = 1$
$B_{0.8}$	1110111011	80% 1's	$0.8^8 \cdot 0.2^2$
$B_{0.5}$	1011010010	50% 1's	$0.5^{10}$

Fair coin flip

Not most probable  
Most probable: 1...1  
Unlikely to be seen

# Bernoulli Distribution

$$\sigma = \sqrt{pq}$$

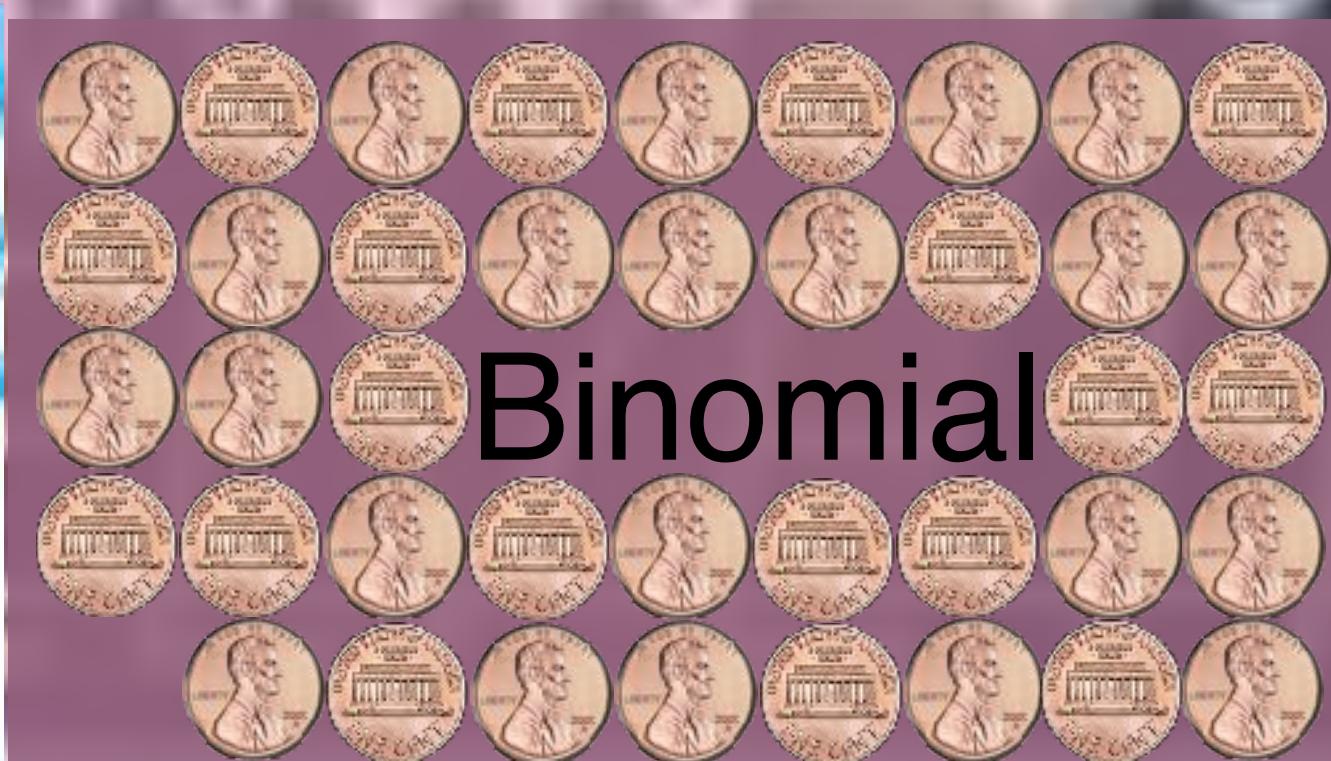
Simplest non-constant distribution

$$B_p \quad 0 \leq p \leq 1$$

$$0 \text{ and } 1 \quad p(1) = p \quad p(0) = 1-p = q$$

$$\mu = p \quad V = pq \quad \sigma = \sqrt{pq}$$

Foundation of many other distributions



A 80

Number of successes in n Bernoulli trials

Useful in many applications

$\mu, V, \sigma$

# Binomial distribution

21 heads

19 tails

# The Binomial Distribution

n independent Bernoulli experiments

$\bar{p}, q$

Each “success” with same probability  $p$ ,

“failure” with probability  $1 - p$

$B_{n,p}$  or  $B_{p,n}$  - distribution of # successes

n independent coin flips

$p(\text{heads}) = p$

$B_{n,p}$  more common,  
use  $B_{p,n}$  as generalizes  $B_p$ ,  
natural for Poisson Binomial

$B_{p,n}$  - distribution of # heads

# Applications

Positive responses to a treatment

Faulty components

Rainy days in a month

Delayed flights

# Small n

n independent experiments

Success probability p

Failure probability q = 1 - p

$b_{p,n}(k)$  - probability of k successes

n = 0

k	$b_{p,0}(k)$
0	1

n = 1

k	$b_{p,1}(k)$
0	q
1	p

$$p+q=1$$

n = 2

k	seq's	$b_{p,1}(k)$
0	00	$q^2$
1	01, 10	$2pq$
2	11	$p^2$



$$p^2 + 2pq + q^2 = (p+q)^2 = 1^2 = 1$$

# General n and k

n  $\perp\!\!\!\perp$   $B_p$  experiments

# successes  $0 \leq k \leq n$

$b_{p,n}(k) = p(k \text{ successes})$

Every k-success sequence:  $n-k$  failures, probability  $p^k \cdot q^{n-k}$

$\binom{n}{k}$  such sequences

$$= \binom{n}{k} p^k q^{n-k}$$

Distribution over  $n+1$  values

# $\Sigma$ WILL IT ADD?

$$0 \leq k \leq n$$

$$p(X = k) = b_{p,n}(k) = \binom{n}{k} p^k q^{n-k}$$

$$\begin{aligned} \sum_{k=0}^n b_{p,n}(k) &= \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \\ &= (p + q)^n \end{aligned}$$

Binomial Theorem

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$

$$= 1^n = 1$$

YES IT  
ADDS!

# Sample Distributions

n=18

$b_{p,18}(k)$



Notebook: experiment with different p and n

# Multiple Choice

Exam has 6 multiple-choice questions, each with 4 possible answers

Each question, student selects one of the 4 answers randomly

$$X = \# \text{ correct answers} \sim B_{1/4, 6}$$

$$\text{Passing: } \geq 4 \text{ correct answers} \quad P(\text{passing}) = ?$$

$$P(4) = \binom{6}{4} \cdot \left(\frac{1}{4}\right)^4 \cdot \left(\frac{3}{4}\right)^2 \approx 0.0329$$

$$P(5) = \binom{6}{5} \cdot \left(\frac{1}{4}\right)^5 \cdot \left(\frac{3}{4}\right)^1 \approx 0.00439$$

$$P(6) = \binom{6}{6} \cdot \left(\frac{1}{4}\right)^6 \cdot \left(\frac{3}{4}\right)^0 \approx 0.000244$$

$$P(\geq 4) = P(4) + P(5) + P(6) \approx 0.03759$$

# Interpretation as a Sum

$B_{p,n}$  a sum of  $n$   $B_p$

$X_1, \dots, X_n \sim B_p \perp\!\!\!\perp$

$$X \stackrel{\text{def}}{=} \sum_{i=1}^n X_i$$

$$P(X = k) = P(\text{exactly } k \text{ of } X_1, \dots, X_n \text{ are } 1) = \binom{n}{k} p^k q^{n-k} = b_{p,n}(k)$$

$X \sim B_{p,n}$

Apply to mean and variance

# Mean and Variance

$$X \sim B_{p,n}$$

$$X = \sum_{i=1}^n X_i \quad X_1, \dots, X_n \sim B_p \perp\!\!\!\perp$$

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum EX_i = \sum p = np$$



$$V(X) = V\left(\sum_{i=1}^n X_i\right) = \sum V(X_i) = \sum pq = npq$$



$$\sigma = \sqrt{npq}$$

# Multiple Choice

Exam has 6 multiple-choice questions, each with 4 possible answers

For each question, student selects one of the 4 answers randomly

$$X = \# \text{ correct answers} \sim B_{1/4, 6}$$

Mean  $EX = np = 6 \cdot 1/4 = 1.5$

Standard deviation  $\sigma = \sqrt{npq} = \sqrt{6 \cdot 1/4 \cdot 3/4} = \frac{\sqrt{18}}{4}$

# Why Vote

For simplicity odd # voters:  $2n + 1$

Each equally likely D or R

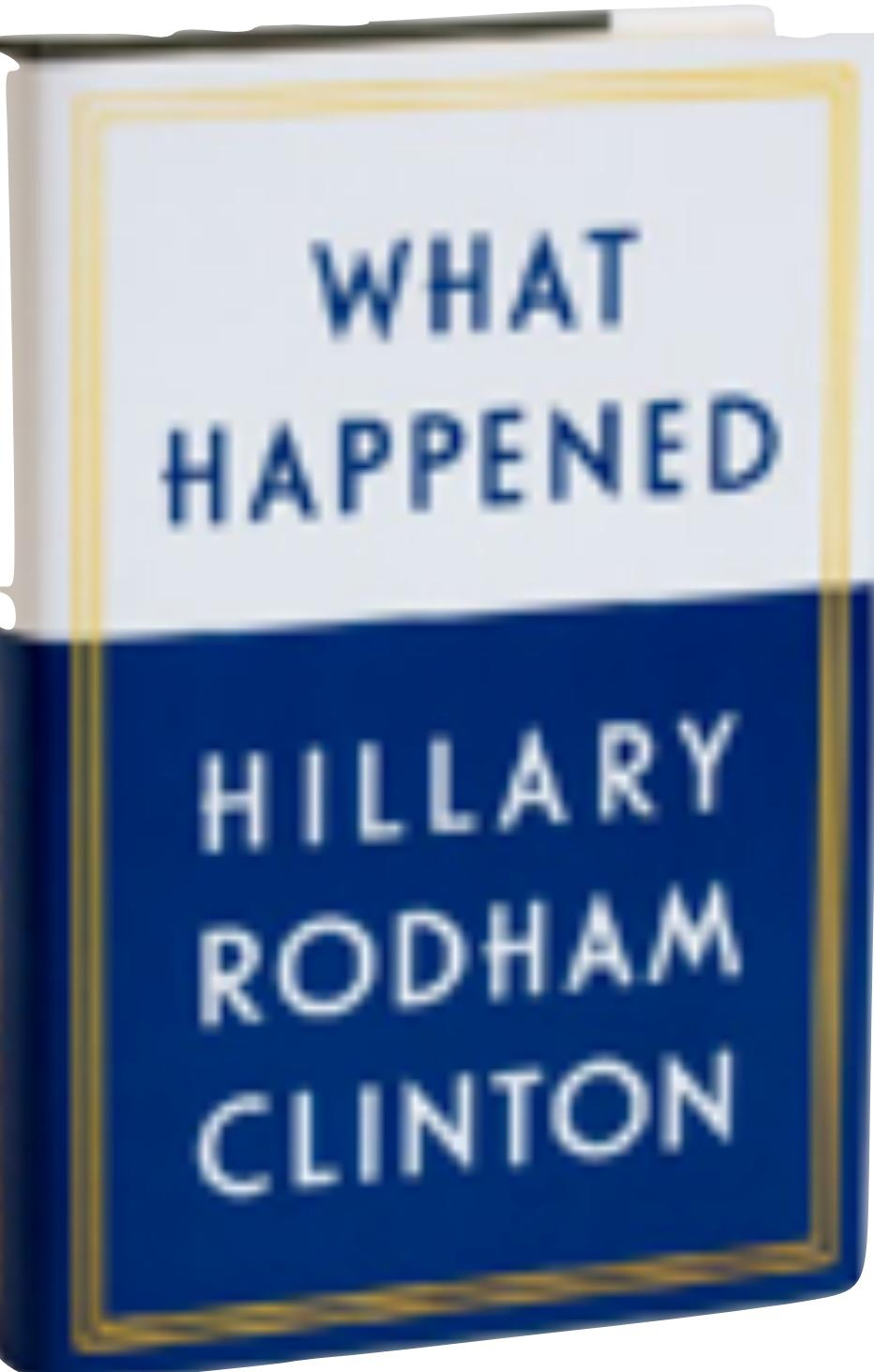
$P(\text{voter makes a difference}) = P(\text{other } 2n \text{ voters equally split})$

$$b_{p,n}(k) = \binom{n}{k} p^k q^{n-k} = \binom{2n}{n} \frac{1}{2^n} \cdot \frac{1}{2^n}$$

Stirling's Approximation

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

$$\begin{aligned} &= \frac{(2n)!}{n! \cdot n! \cdot 2^n \cdot 2^n} \\ &\approx \frac{\sqrt{2\pi \cdot 2n} \left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n} \left(\frac{n}{e}\right)^n\right)^2 2^{2n}} \\ &= \frac{1}{\sqrt{\pi n}} \end{aligned}$$



# Poisson Binomial

Generalizes the binomial distribution

$n \geq 1$	Binomial	$B_{p,n}$	For $1 \leq i \leq n$	$X_i \sim B(p)$	$\vdash X = \sum_{i=1}^n X_i$
	Poisson Binomial	$PB_{p_1, \dots, p_n}$		$X_i \sim B(p_i)$	

$PB(1/4, 2/3)$

$X_1 \sim B(1/4)$

$X_2 \sim B(2/3)$

$\vdash$

$X_1$	$X_2$	$P$	$X$
0	0	$3/4 \cdot 1/3 = 1/4$	0
0	1	$3/4 \cdot 2/3 = 1/2$	1
1	0	$1/4 \cdot 1/3 = 1/12$	1
1	1	$1/4 \cdot 2/3 = 1/6$	2

$X$	$P(x)$
0	$1/4$
1	$7/12$
2	$1/6$

# Expectation and Variance

$$X \sim PB_{p_1, p_2, \dots, p_n}$$

$$X = \sum_{i=1}^n X_i \quad X_i \sim B_{p_i}$$

$$E(X) = E\left(\sum_{i=1}^n X_i\right) \stackrel{\text{LE}}{=} \sum_{i=1}^n EX_i \stackrel{\text{B}_{pi}}{=} \sum_{i=1}^n p_i$$

$$V(X) = V\left(\sum_{i=1}^n X_i\right) \stackrel{\perp}{=} \sum_{i=1}^n V(X_i) \stackrel{\text{B}_{pi}}{=} \sum_{i=1}^n p_i(1 - p_i)$$

p(k)

No closed form

Computationally

Homework

$B_{p,n}$

Number of successes in n  $B_p$  trials

$$b_{p,n}(k) = \binom{n}{k} p^k \bar{p}^{n-k}$$

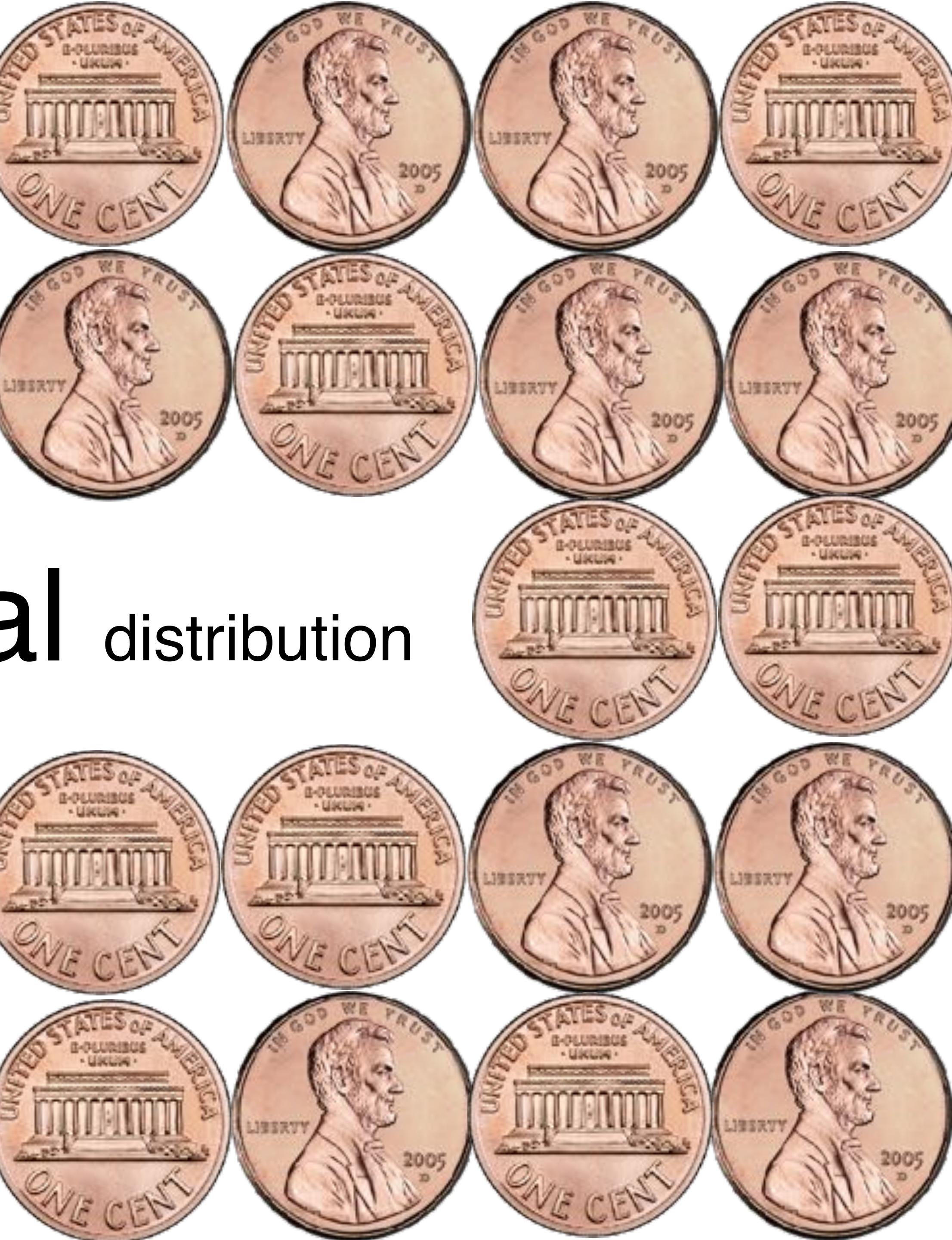
$$\mu = np$$

$$\sigma = \sqrt{np\bar{p}}$$

Why vote

# Binomial distribution

Poisson Distributions



# Coin Flips

Most basic convergence to average is  $B(p)$

Flip  $n$   $B(p)$  coins, average # 1's will approach  $np$

Probability of a sequence with  $k$  1's and  $n-k$  0's is  $p^k q^{n-k}$

Wolog assume  $p>0.5$ , then most likely is  $1^n$

Yet by WLLN with probability  $\rightarrow 1$  we see roughly  $p n$  1's and  $q n$  0's

Why do we observe these sequences and not the most likely ones?

Strength in #s. # sequences of a given composition increases near 1/2

$p n$  balances #  $\times$  probability.

Definition

Applications

Motivation

Derivation

$\mu, \sigma$

Example

# Poisson distribution

# The Poisson Distribution

Parameter

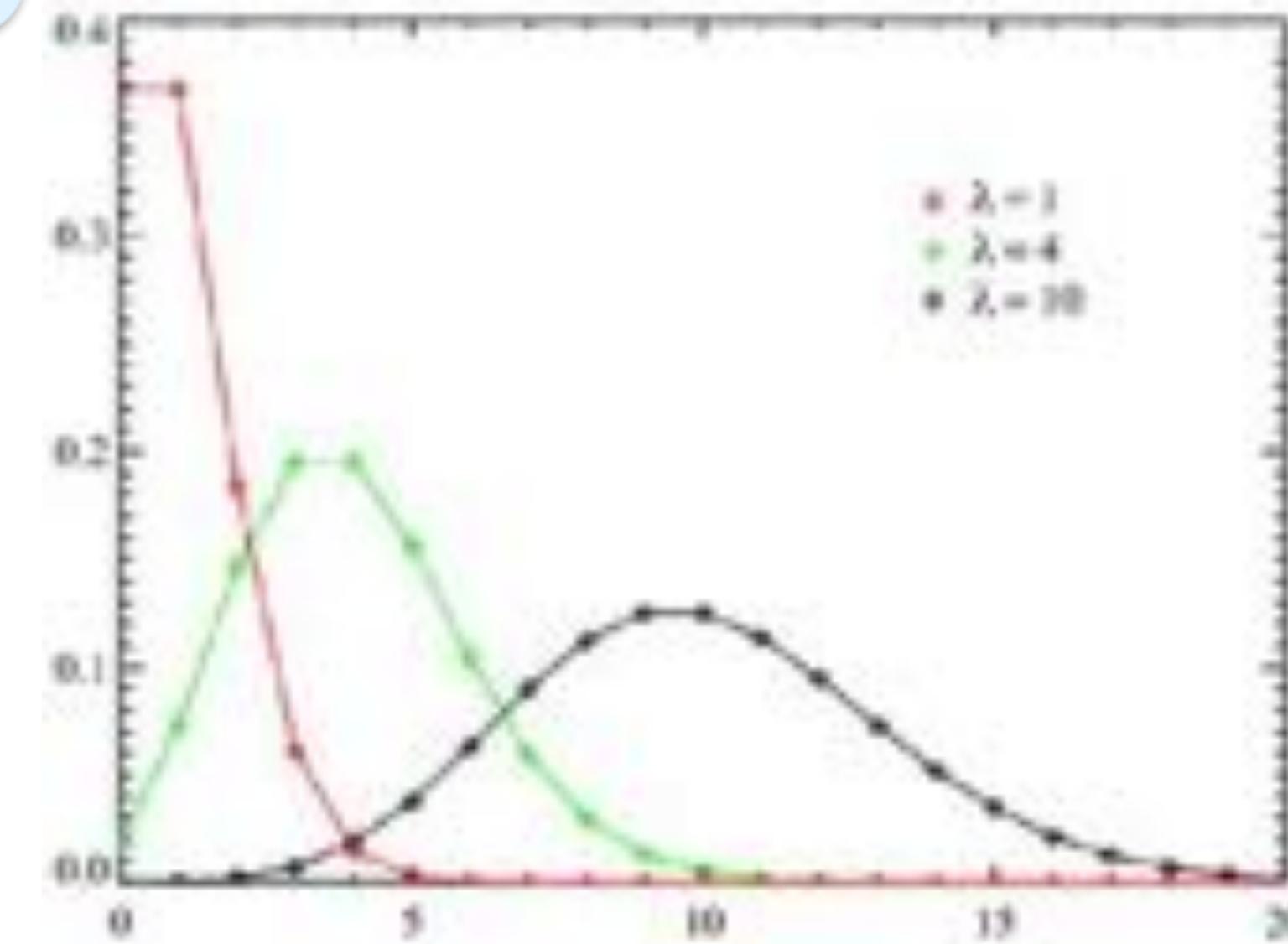
$\lambda \geq 0$  (some  $> 0$ )

Support

$N$

$$P_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

PMF



More in notebook

Significance

Approximates  $B_{p,n}$  for large  $n$  and small  $p$  so that  $np = \lambda$  is moderate

# We are Poisson

$P_\lambda$  approximates  $B_{p,n}$  for small p, large n

Numerous applications

# People clicking ad

Responses to spam

Rare-disease infections

Daily 911 calls

Daily store customers

Gallery purchasing customers

Flight no shows

Typos in a page



\$450M

# Small k

$\overbrace{\hspace{40em}}^k$

$\lambda$	$P_\lambda(k)$	0	1	2	3
General	$e^{-\lambda} \frac{\lambda^k}{k!}$	$\frac{1}{e^\lambda}$	$\frac{\lambda}{e^\lambda}$	$\frac{\lambda^2}{2e^\lambda}$	$\frac{\lambda^3}{6e^\lambda}$
1	$\frac{1}{ek!}$	$\frac{1}{e}$	$\frac{1}{e}$	$\frac{1}{2e}$	$\frac{1}{6e}$
2	$\frac{2^k}{e^{2k} k!}$	$\frac{1}{e^2}$	$\frac{2}{e^2}$	$\frac{2}{e^2}$	$\frac{4}{3e^2}$
0	$\frac{0^k}{k!}$	1	0	0	0

# Binomial Approximation

$P_\lambda$  approximates  $B_{p,n}$  for  $\lambda = pn$ , when  $n \gg 1 \gg p$

$$B_{p,n}(k) = \binom{n}{k} p^k q^{n-k} \quad q = 1 - p$$

$$p = \frac{\lambda}{n}$$

$$= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n^k}{k!} \cdot \frac{\lambda^k}{n^k} \cdot \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k}$$

Fix  $k$ , fixe  $\lambda$  while  $n \nearrow$  and  $p \searrow$

Derive Poisson

# Limit of Binomial

$$B_{p,n}(k) = \frac{\lambda^k}{k!} \cdot \frac{n^k}{n^k} \cdot \frac{(1 - \frac{\lambda}{n})^n}{(1 - \frac{\lambda}{n})^k} \rightarrow e^{-\lambda} \frac{\lambda^k}{k!} \quad \checkmark$$

$\lambda = p \cdot n$

$\lambda$  and  $k$  fixed,  $n \rightarrow \infty$

1

e<sup>-λ</sup>

1

↑

①  $\frac{n^k}{n^k} = \cancel{\frac{n}{n}} \cdot \frac{\cancel{(n-1)}}{n} \cdot \dots \cdot \frac{\cancel{(n-k+1)}}{n} \rightarrow 1 \quad \text{fixed \# (k) terms, each} \rightarrow 1$

②  $(1 - \frac{\lambda}{n})^k \rightarrow 1 \quad \text{fixed \# (k) terms, each} \rightarrow 1$

③  $(1 - \frac{\lambda}{n})^n = ((1 - \frac{\lambda}{n})^{\frac{n}{\lambda}})^\lambda \rightarrow (e^{-1})^\lambda = e^{-\lambda}$

increasing # terms, each  $\rightarrow 1$      $(1 - \frac{1}{m})^m \rightarrow e^{-1}$

# $\Sigma$ WILL IT ADD?

$$P_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad k \geq 0$$

$\geq 0$



Taylor expansion

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

$$\sum_{k=0}^{\infty} P_\lambda(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^\lambda = 1$$

YES IT  
ADDS!

# Mean and Variance

$P_\lambda$  approximates  $B_{p,n}$  for  $\lambda = np$  when  $n \gg 1 \gg p$

	$\mu$	$V$
$B_{p,n}$	$np$	$npq$
$P_\lambda$	$\lambda$	$\lambda$

← Expect

Calculate next

# Observation

$$\frac{d}{d\lambda} \lambda^k = k\lambda^{k-1} = \frac{k}{\lambda} \lambda^k$$

$$\frac{d^2}{d\lambda^2} \lambda^k = k^2 \lambda^{k-2} = \frac{k^2}{\lambda^2} \lambda^k$$

$$\frac{d^r}{d\lambda^r} \lambda^k = k^r \lambda^{k-r} = \frac{k^r}{\lambda^r} \lambda^k$$

$$k^r \lambda^k = \lambda^r \frac{d^r}{d\lambda^r} \lambda^k$$

# Falling Moments

$$X \sim P_\lambda$$

$$k^r \lambda^k = \lambda^r \frac{d^r}{d\lambda^r} \lambda^k$$

$$\begin{aligned} E(X^r) &= \sum_{k=0}^{\infty} k^r P_\lambda(k) = \sum_k k^r e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_k k^r \frac{\lambda^k}{k!} = e^{-\lambda} \sum_k \frac{\lambda^r}{k!} \frac{d^r}{d\lambda^r} \lambda^k \\ &= e^{-\lambda} \lambda^r \frac{d^r}{d\lambda^r} \sum_k \frac{\lambda^k}{k!} = e^{-\lambda} \lambda^r \frac{d^r}{d\lambda^r} e^\lambda \\ &= e^{-\lambda} \lambda^r e^\lambda = \lambda^r \end{aligned}$$

$$EX = EX^1 = \lambda$$

$$EX(X - 1) = EX^2 = \lambda^2$$

# Mean and Variance

$$EX = EX^1 = \lambda \quad \checkmark$$

$$EX(X - 1) = EX^2 = \lambda^2$$

$$E(X^2) = E(X(X - 1) + X) = E(X(X - 1)) + E(X) = \lambda^2 + \lambda$$

$$V(X) = E(X^2) - (EX)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \quad \checkmark$$

$$\sigma = \sqrt{\lambda} \quad \text{Small relative to the mean}$$

# Approximation Example

Factory produces 200 items, each defective with probability 1%

P(3 defective)?

Binomial (precise)

$$B_{0.01,200}(3) = \binom{200}{3} (0.01)^3 (0.99)^{197} \approx 0.181$$

Poisson (approximation)  $\lambda = 200 \cdot 0.01 = 2$   $P_2(3) = e^{-2} \frac{2^3}{3!} \approx 0.18$

P(some defective)?

$$B_{0.01,200}(0) = \binom{200}{0} (0.99)^{200} \approx 0.134$$

$$P_2(0) = e^{-2} \frac{2^0}{0!} = e^{-2} \approx 0.135$$

$$B_{0.01,200}(\geq 1) = 1 - 0.134 \approx 0.866$$

$$P_2(\geq 1) = 1 - 0.135 \approx 0.865$$

$$P_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

 $\lambda \geq 0$  $k \geq 0$ 

Approximates  $B_{p,n}$  for  $\lambda = np$ , when  $n \gg 1 \gg p$

# of ad clicks, rare diseases, production defects

$$\mu = \lambda$$

$$V = \lambda$$

$$\sigma = \sqrt{\lambda}$$



Geometric Distribution

Poisson  
distribution

