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

μ

V

σ

Python

Plot and experiment

Show Distribution

Nonnegative

Sum to 1

Will It Blend?

Blendtec

Tom Dickson

ΣWill It ADD?

Bernoulli Distribution

Simplest non-constant distribution

Foundation of many others

μ V σ

Repeated experiments

Jacob Bernoulli, 1655-1705

Theology → mathematics

Calculus Integrals

"Euler" number $e = \lim_{n \to \infty} (1 + \frac{1}{n})^n$ 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!

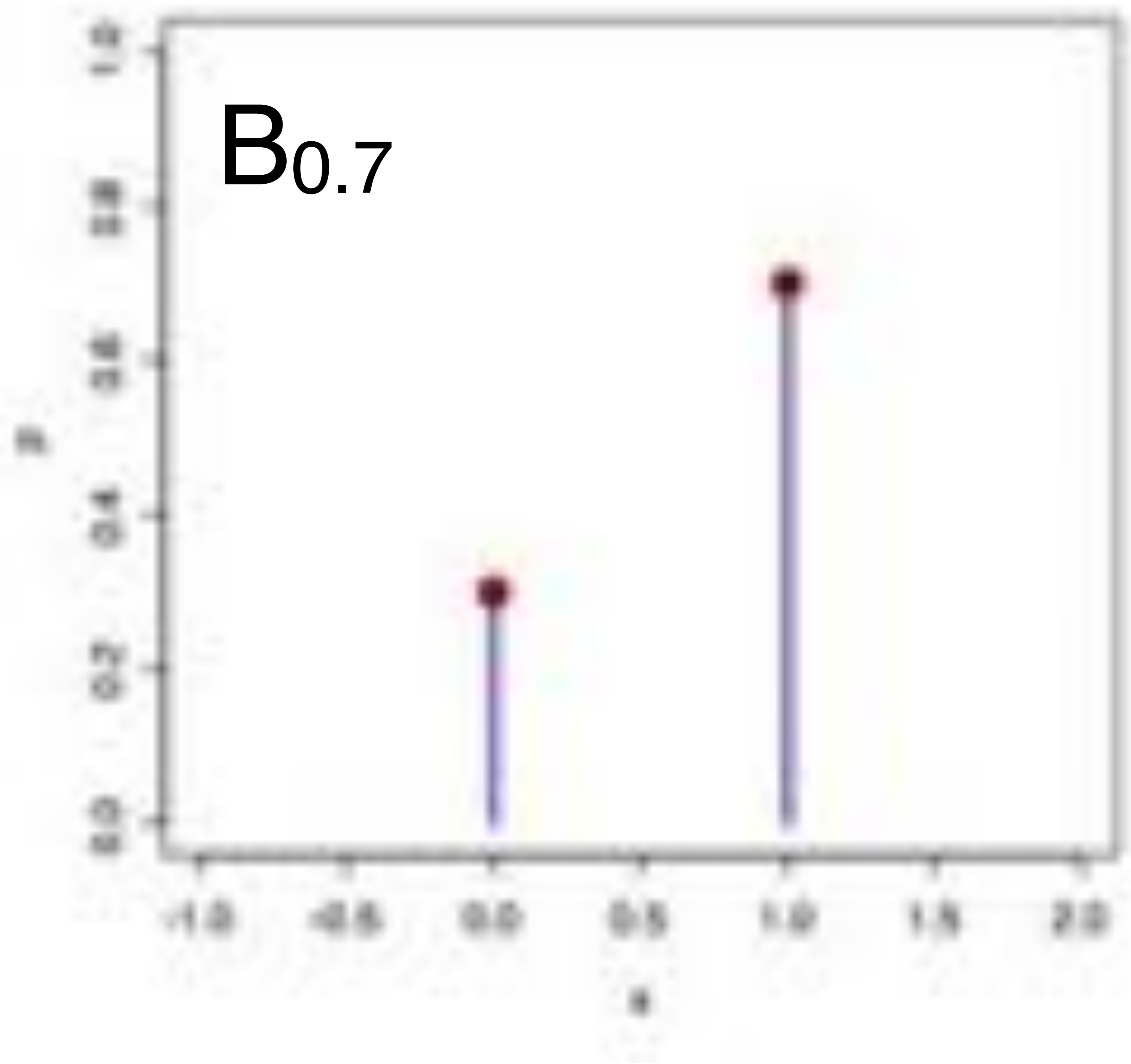
Bernoulli Distribution

B_p $0 \leq p \leq 1$

failure success

Two values	0	1
Probability	1-p	p

$\overline{p}$ q



$p(0) + p(1) = (1\text{-}p)+p = 1$

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

$X \sim B_p$ $p(0) = 1\text{-}p$ $p(1) = p$

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

$X \sim B_{0.8}$ $EX = 0.8$

$EX = P(X=1)$

Fraction of times expect to see 1

Variance

$X \sim B_p$

$EX = p$

Variance

Easy route

$0^2 = 0$ $\quad 1^2 = 1$ $\quad X^2 = X$ $\quad E(X^2) = EX = p$

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

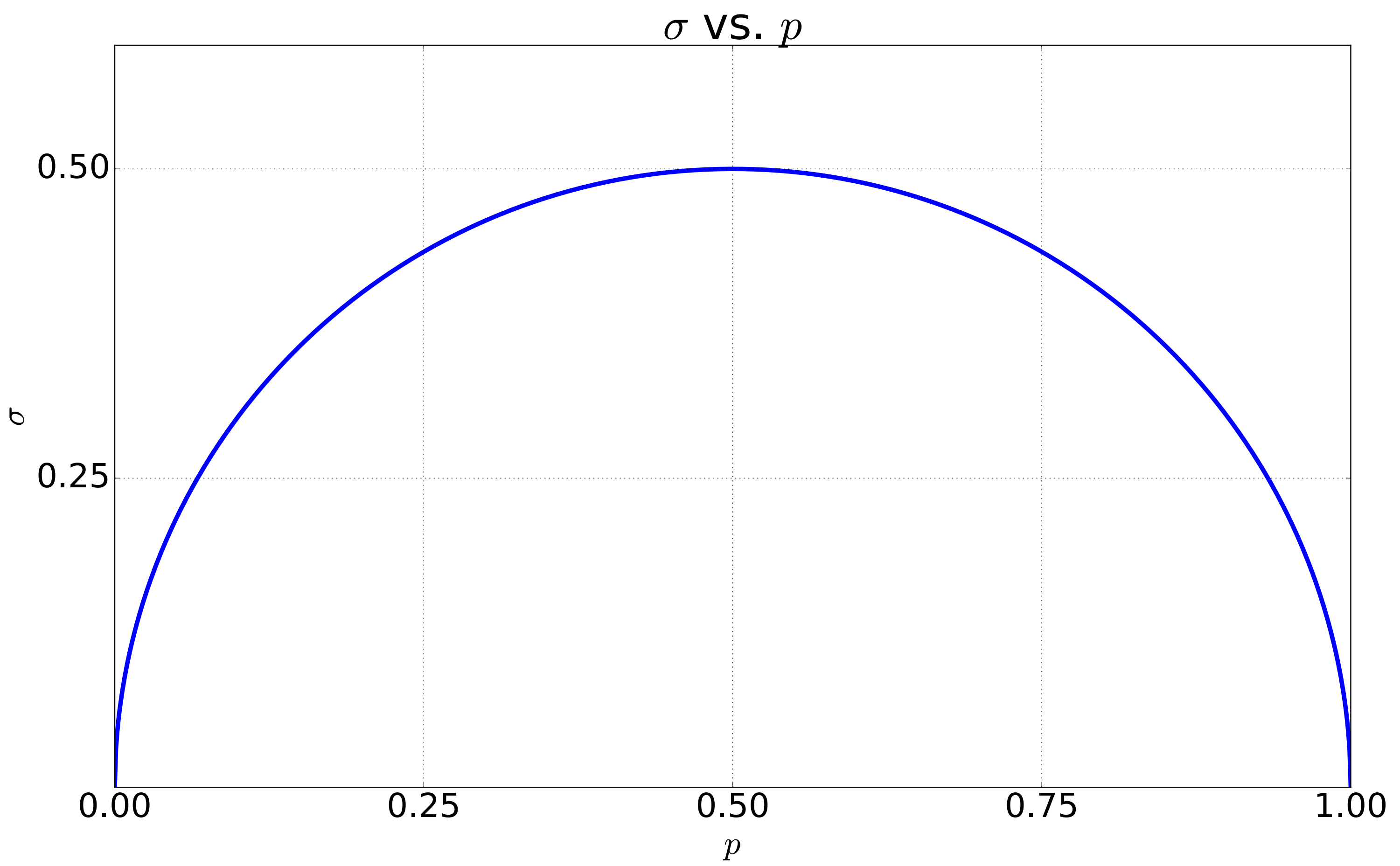
Standard Deviation

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

B_p varies most when $p = ½$

E, V, σ for various p

p	EX	V(X)	σ
0	0	0	0
1	1	0	0
½	½	¼	½



Independent Trials

Much of B_p importance stems from multiple trials

Most common | Independent | $\perp\!\!\!\perp$

$0 \le p \le 1$ | $X_1, X_2, X_3 \sim B_p$ | $\perp\!\!\!\perp$

$q \stackrel{\text{def}}{=} 1-p$ | $P(110) = p^2 q = P(101) = P(011)$

Generally | $X_1, X_2, \dots, X_n \sim B_p$ | $\perp\!\!\!\perp$

$x^n = x_1, x_2, \dots, x_n \in \{0,1\}^n$ | n_0 0's and n_1 1's

$$P(x_1, \dots, x_n) = p^{n_1} q^{n_0} \qquad P(10101) = p^{n_1} q^{n_0} = p^3 q^2$$

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}

Not most probable

Most probable: 1…1

Unlikely to be seen

Fair coin flip

Bernoulli Distribution

Simplest non-constant distribution

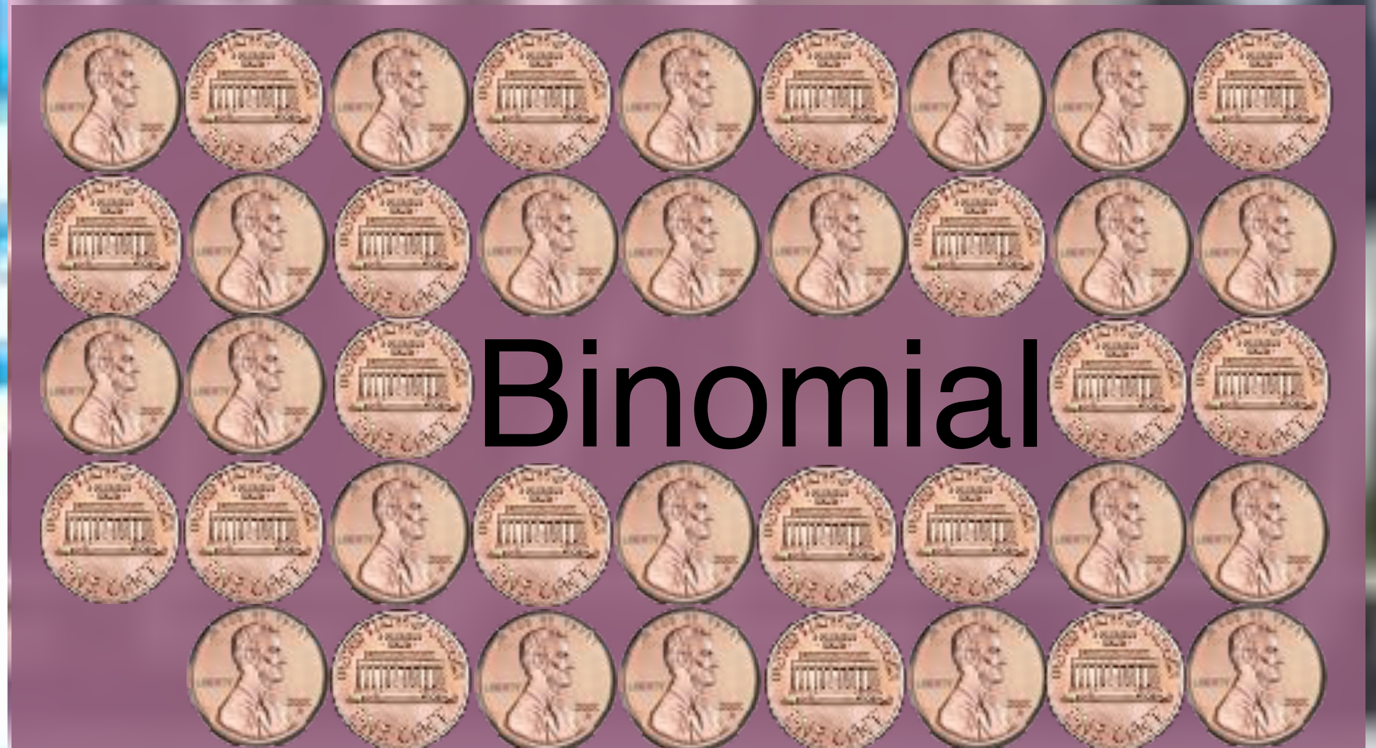
B_p $0 \leq p \leq 1$

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

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

$\sigma = \sqrt{pq}$

Foundation of many other distributions



Binomial distribution

Number of successes in n Bernoulli trials

Useful in many applications

μ, V, σ

21 heads

19 tails

The Binomial Distribution

n independent Bernoulli experiments

$\overline{p}$, q

Each “success” with same probability p, “failure" with probability 1 - p

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

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

n independent coin flips

p(heads) = p

$B_{p,n}$ - 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 $0 \le k \le n$

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

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

ΣWILL IT ADD?

$0 \le k \le n$

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

$$\sum_{k=0}^{n} b_{p,n}(k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}$$

$$= (p+q)^n$$

$$= 1^n = 1$$

YES IT ADDS!

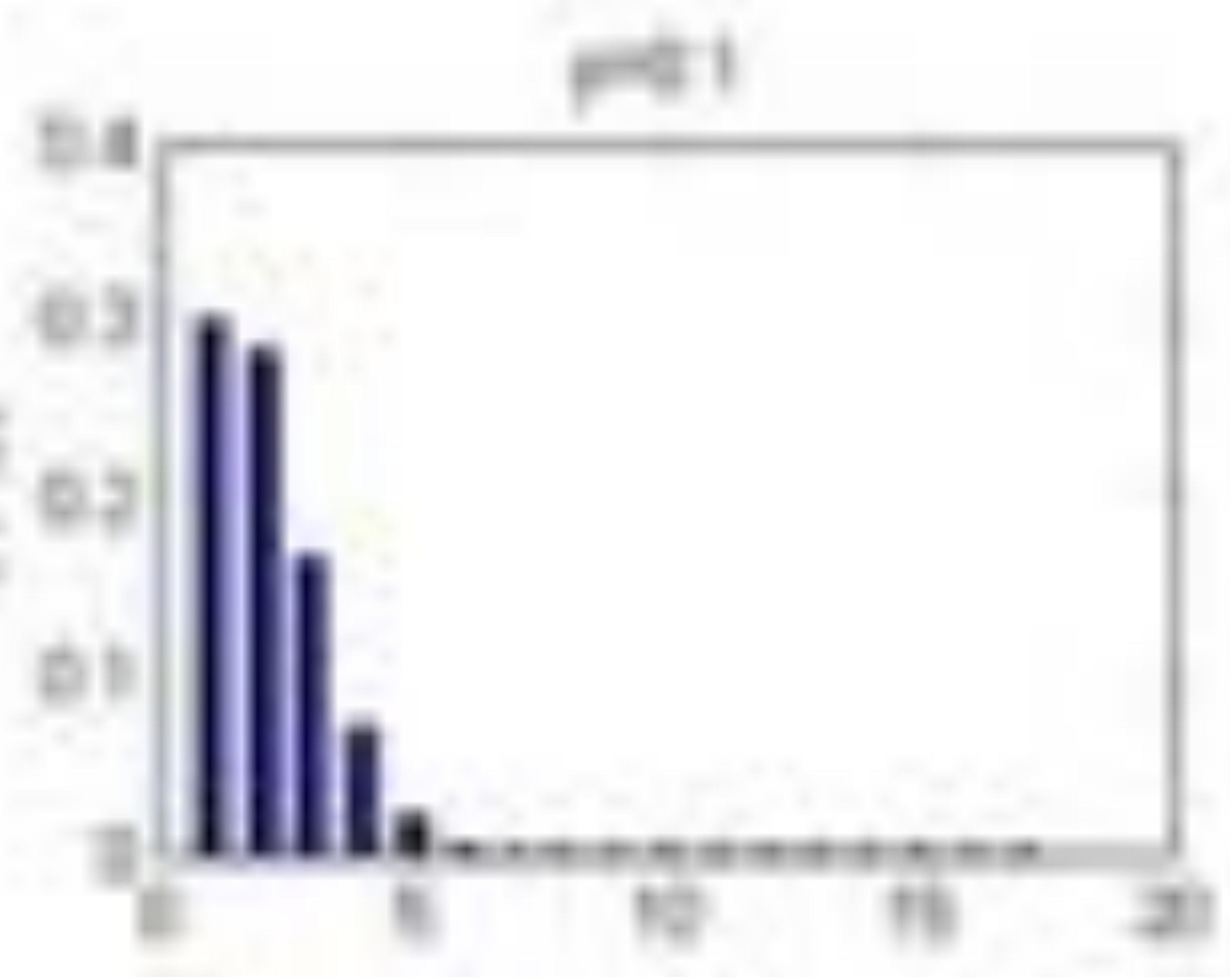
Binomial Theorem

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

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 = # correct answers

$\sim B_{¼,6}$

Passing: ≥ 4 correct answers

P(passing) = ?

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

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

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

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

Interpretation as a Sum

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

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

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

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

$X \sim B_{p,n}$

Apply to mean and variance

Mean and Variance

$X \sim B_{p,n}$

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

$$E(X) = E(\sum_{i=1}^{n} X_i) \overset{\text{LE}}{=} \sum EX_i \overset{B_p}{=} \sum p = np$$

$$V(X) = V(\sum_{i=1}^{n} X_i) \overset{\perp\!\!\!\perp}{=} \sum V(X_i) \overset{B_p}{=} \sum pq = npq$$

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

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 = # correct answers $\sim B_{¼,6}$

Mean $EX = np = 6 \cdot ¼ = 1.5$

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

Why Vote

For simplicity odd # voters: 2n + 1

Each equally likely D or R

P(voter makes a difference) = P(other 2n voters equally split)

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

$$= \binom{2n}{n} \frac{1}{2^n} \cdot \frac{1}{2^n}$$

$$= \frac{(2n)!}{n! \cdot n! \cdot 2^n \cdot 2^n}$$

$$\approx \frac{\sqrt{2\pi \cdot 2n}(\frac{2n}{e})^{2n}}{(\sqrt{2\pi n}(\frac{n}{e})^n)^2 2^{2n}}$$

$$= \frac{1}{\sqrt{\pi n}}$$

Stirling's Approximation

$$n! \sim \sqrt{2\pi n}(\tfrac{n}{e})^n$$

Poisson Binomial

Generalizes the binomial distribution

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

$PB(\frac{1}{4}, \frac{2}{3})$ $X_1 \sim B(\frac{1}{4})$ $X_2 \sim B(\frac{2}{3})$ $\perp\!\!\!\perp$

X_1	X_2	P	X
0	0	$\frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4}$	0
0	1	$\frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$	1
1	0	$\frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}$	1
1	1	$\frac{1}{4} \cdot \frac{2}{3} = \frac{1}{6}$	2

X	P(x)
0	$\frac{1}{4}$
1	$\frac{7}{12}$
2	$\frac{1}{6}$

Expectation and Variance

$$X \sim PB_{p_1, p_2, \ldots, p_n} \qquad X = \sum_{i=1}^{n} X_i \qquad X_i \sim B_{p_i}$$

$$E(X) = E(\sum_{i=1}^{n} X_i) \overset{\text{LE}}{=} \sum_{i=1}^{n} EX_i \overset{\text{Bpi}}{=} \sum_{i=1}^{n} p_i$$

$$V(X) = V(\sum_{i=1}^{n} X_i) \overset{\perp\!\!\!\perp}{=} \sum_{i=1}^{n} V(X_i) \overset{\text{Bpi}}{=} \sum_{i=1}^{n} p_i(1 - p_i)$$

p(k) | No closed form | Computationally | Homework

$B_{p,n}$ Number of successes in n B_p trials

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

$\mu = np$ $V = np\bar{p}$ $\sigma = \sqrt{np\overline{p}}$

Why vote

Binomial distribution

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 →1 we see roughly pn 1's and qn 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

pn balances # x probability.

Definition

Applications

Motivation

Derivation

μ, σ

Example

Poisson distribution

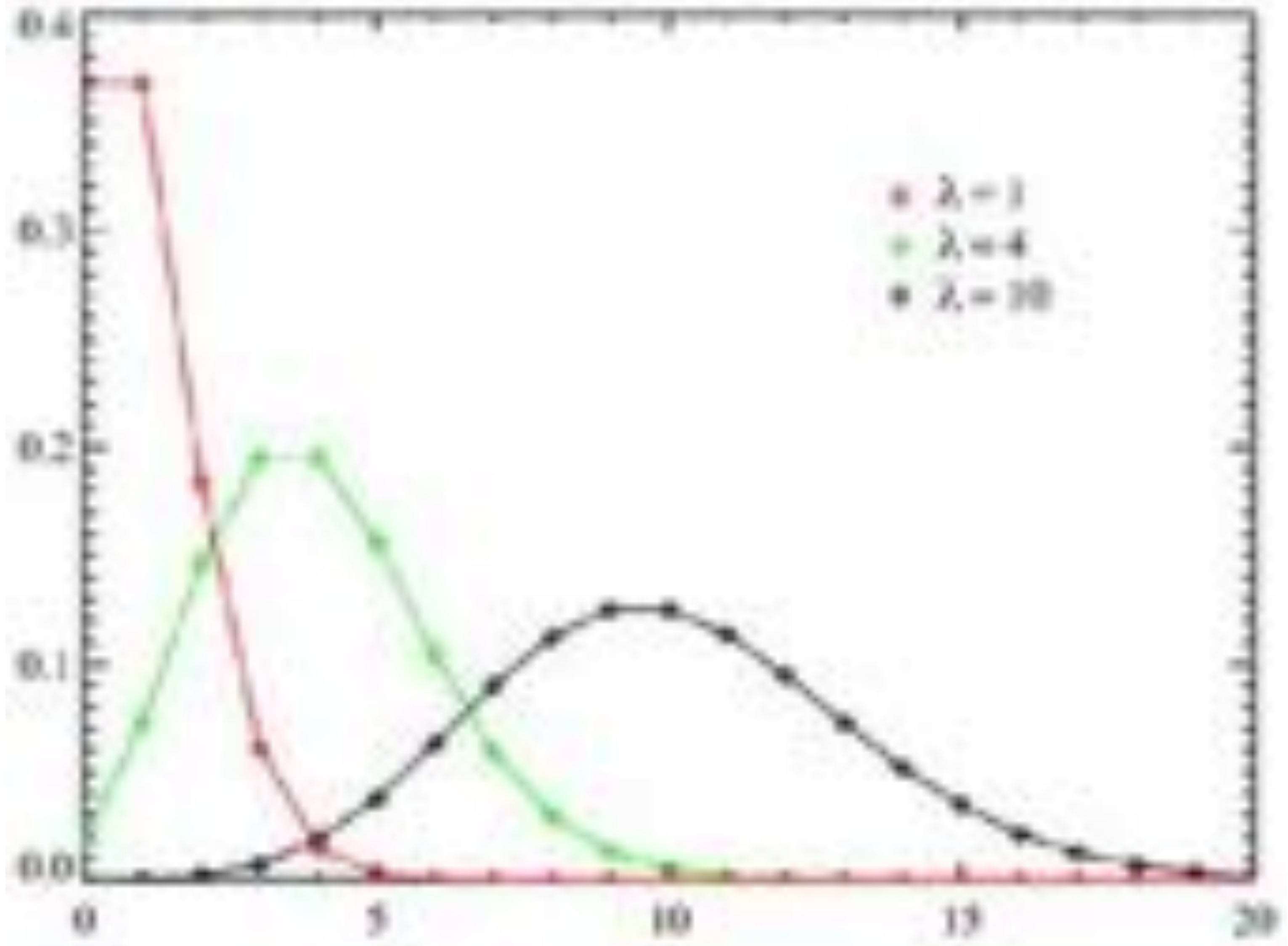
The Poisson Distribution

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

Support $\mathbb{N}$

PMF

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



More in notebook

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

We are Poisson

P_λ approximates $B_{p,n}$ for small p, large n

Numerous applications

$450M

\# 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

Small k

k

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

Binomial Approximation

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

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

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

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

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

Fix k, fixe λ while n ↗ and p ↘

Derive Poisson

Limit of Binomial

$$B_{p,n}(k) = \frac{\lambda^k}{k!}\cdot \underbrace{\cancel{\frac{n^k}{n^k}}}_{1}\cdot\frac{\cancel{(1-\frac{\lambda}{n})^n} \to e^{-\lambda}}{\cancel{(1-\frac{\lambda}{n})^k} \to 1} \to e^{-\lambda}\frac{\lambda^k}{k!}$$ ✅

$$\lambda = p\cdot n$$

λ and k fixed, n $\to \infty$

① $\frac{n^k}{n^k} = \underbrace{\cancel{\frac{n}{n}}}_{1} \cdot \underbrace{\cancel{\frac{(n-1)}{n}}}_{1} \cdot \cdots \cdot \underbrace{\cancel{\frac{(n-k+1)}{n}}}_{1} \to 1$ fixed # (k) terms, each → 1

② $(1-\frac{\lambda}{n})^k \to 1$ fixed # (k) terms, each → 1

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

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

ΣWill It ADD?

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

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

	μ	V	
$B_{p,n}$	np	npq	
P_λ	λ	λ	$\longleftarrow$ Expect

Calculate next

Observation

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

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

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

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

Falling Moments

$$X \sim P_\lambda$$

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

$$E(X^{\underline{r}}) = \sum_{k=0}^{\infty} k^{\underline{r}} P_\lambda(k) = \sum_k k^{\underline{r}} e^{-\lambda}\frac{\lambda^k}{k!}$$

$$= e^{-\lambda}\sum_k k^{\underline{r}}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_k \frac{\lambda^r}{k!}\frac{d^r}{d\lambda^r}\lambda^k$$

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

$$= e^{-\lambda}\lambda^r e^\lambda = \lambda^r$$

$$EX = EX^{\underline{1}} = \lambda \qquad EX(X-1) = EX^{\underline{2}} = \lambda^2$$

Mean and Variance

$$EX = EX^{\underline{1}} = \lambda$$ ✅

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

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

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

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