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

Blendtec

Tom Dickson

Will It Blend?

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!

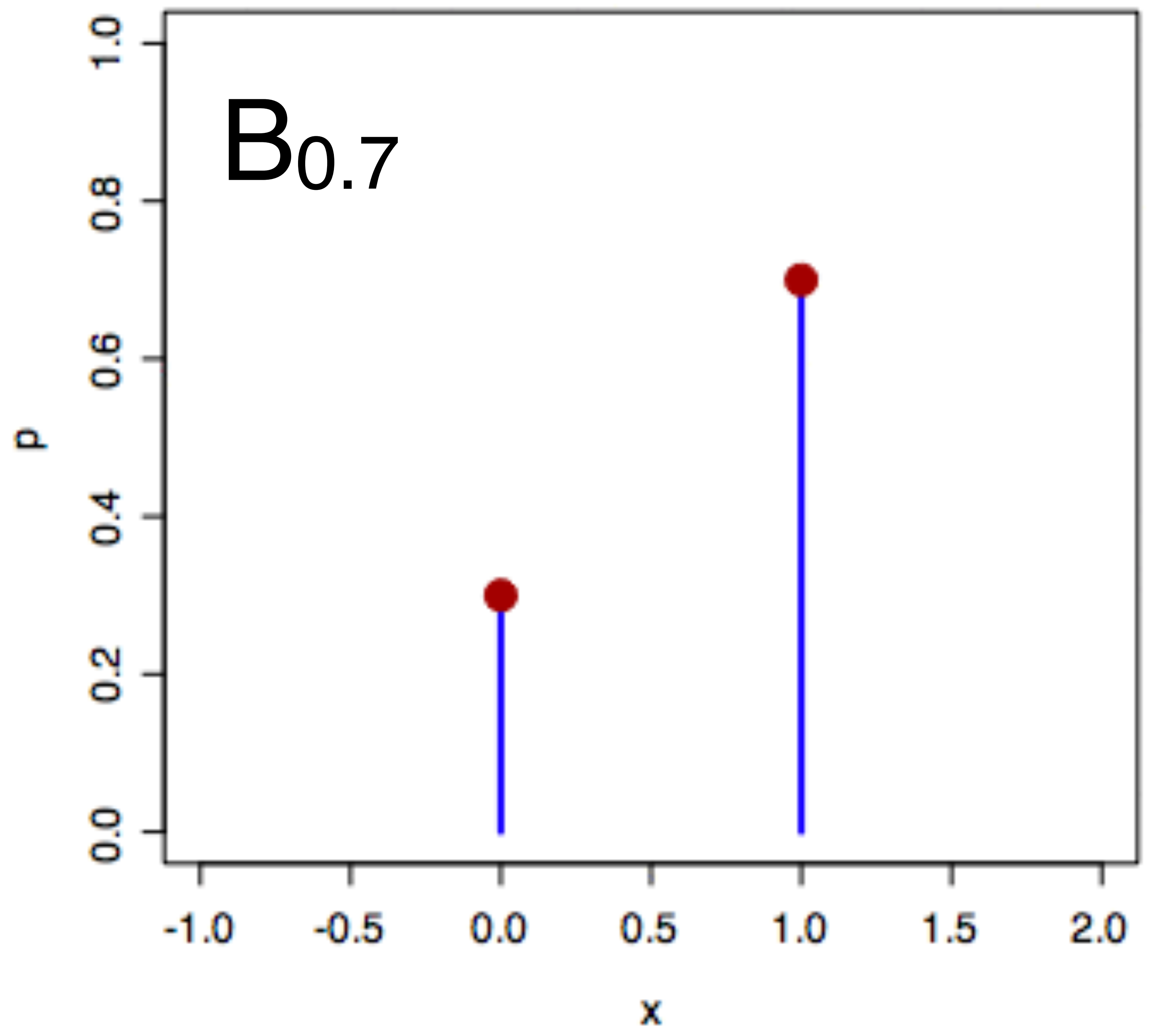
Bernoulli Distribution

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

failure success

Two values	0	1
Probability	1-p	p

$\overline{p}$ q



ΣWill 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?

Everyone!

Binary version of complex events

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

$X \sim B_p$

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

$p(1) = p$

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

$X \sim B_{0.8}$

$EX = 0.8$

$EX = P(X=1)$

Fraction of times expect to see 1

Variance

$X \sim B_p$

$EX = p$

Variance

Easy route

$0^2 = 0$

$1^2 = 1$

$X^2 = X$

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

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

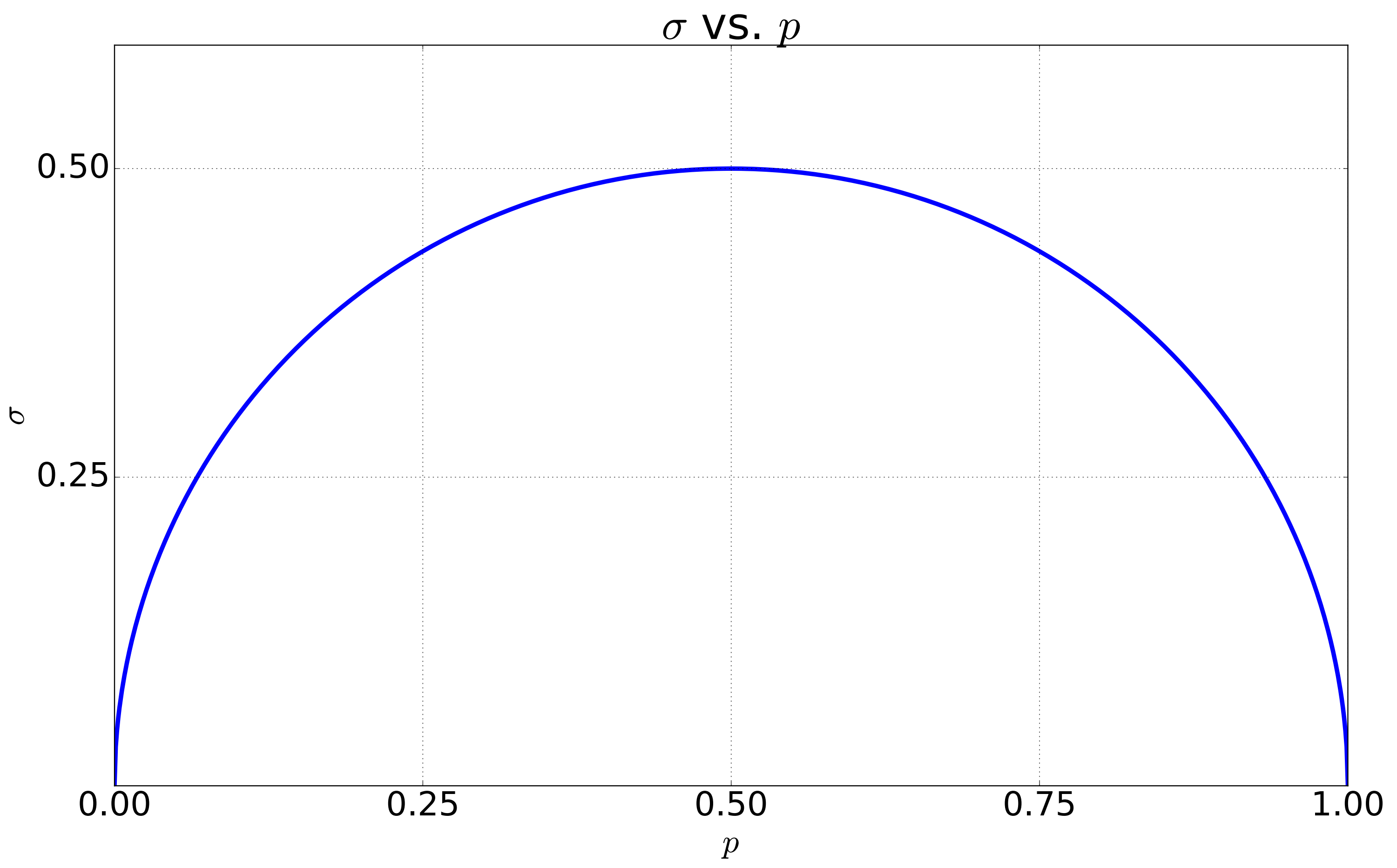
Standard Deviation

$\sigma = \sqrt{pq}$

B_p varies most when $p = ½$

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 $\perp\!\!\!\perp$

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

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

Generally $X_1, X_2, \ldots, X_n \sim B_p$ $\perp\!\!\!\perp$

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

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

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 $0 \le p \le 1$

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