Review

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}, \qquad \text{assuming } \mathrm{P}(B) > 0$$

- Multiplication rule:

$$\mathrm{P}(A \cap B) = \mathrm{P}(B) \cdot \mathrm{P}(A \mid B) = \mathrm{P}(A) \cdot \mathrm{P}(B \mid A)$$

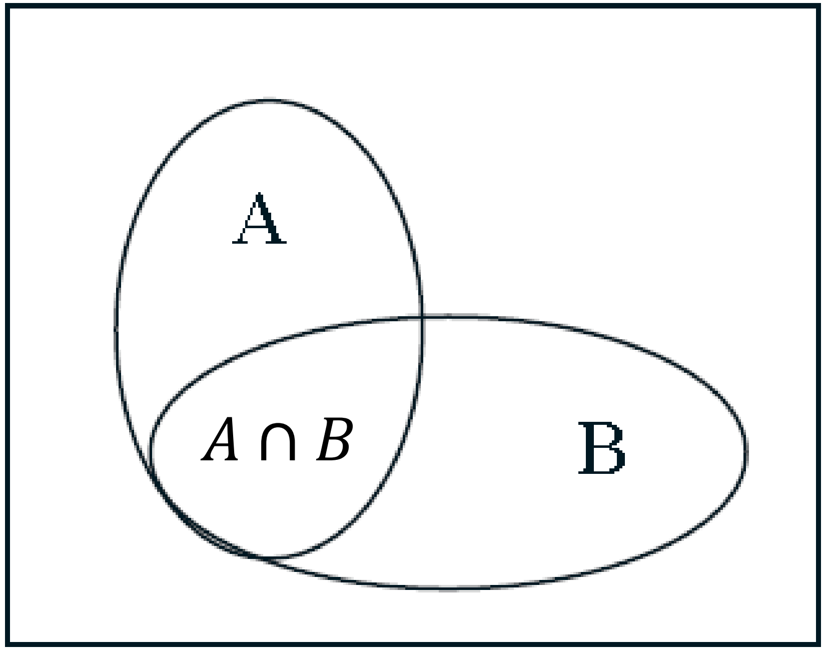
- Total probability theorem:

$$\mathrm{P}(B) = \mathrm{P}(A)\mathrm{P}(B \mid A) + \mathrm{P}(A^c)\mathrm{P}(B \mid A^c)$$

- Bayes rule:

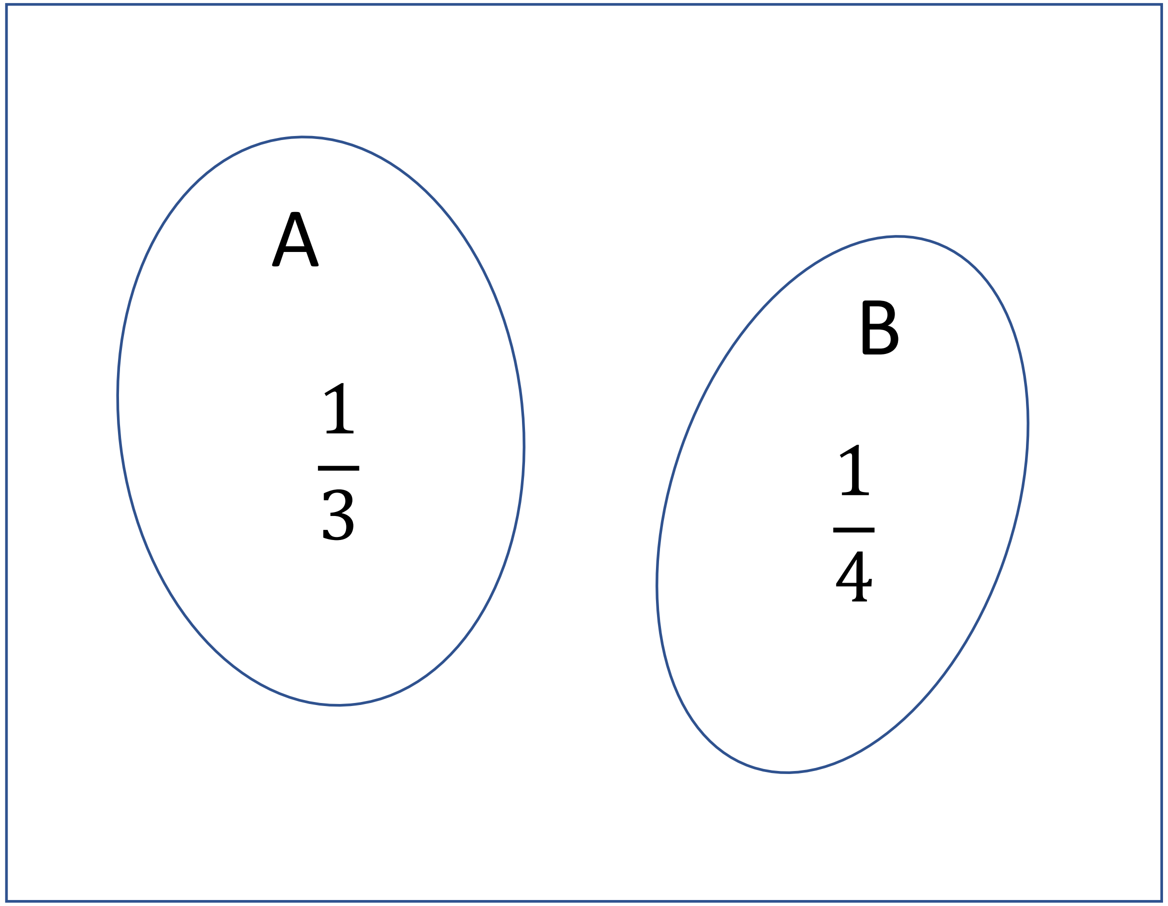
$$\mathrm{P}(A_i \mid B) = \frac{\mathrm{P}(A_i)\mathrm{P}(B \mid A_i)}{\mathrm{P}(B)}$$

Independence of two events



- **"Defn:"** $\mathrm{P}(B \mid A) = \mathrm{P}(B)$
  - "occurrence of A provides no information about B's occurrence"
- Recall that $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B \mid A)$
- **Defn:** $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$
- Symmetric with respect to A and B
  - applies even if $\mathrm{P}(A) = 0$
  - implies $\mathrm{P}(A \mid B) = \mathrm{P}(A)$

INDEPENDENCE is not to be confused with DISJOINTNESS



$$P(A \cap B) = 0$$

NOT EQUAL

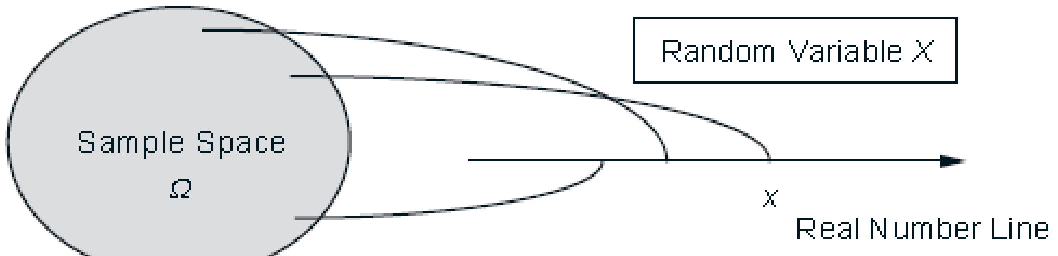
$$P(A)P(B) = \frac{1}{12}$$

$$P(A) = \frac{1}{3}$$

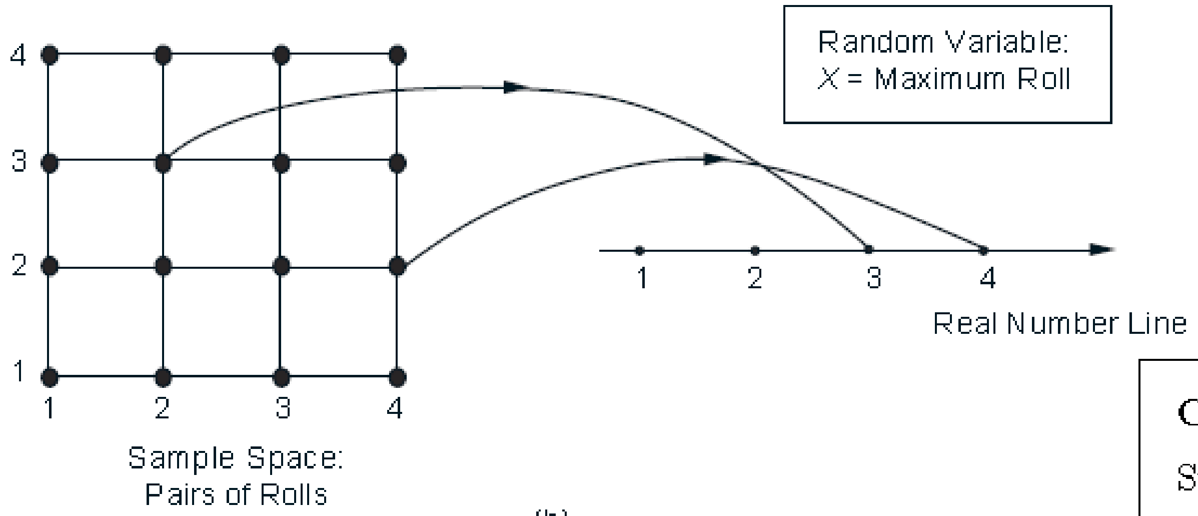
NOT EQUAL

$$P(A|B) = 0 \neq P(A)$$

Discrete Random Variables



(a)

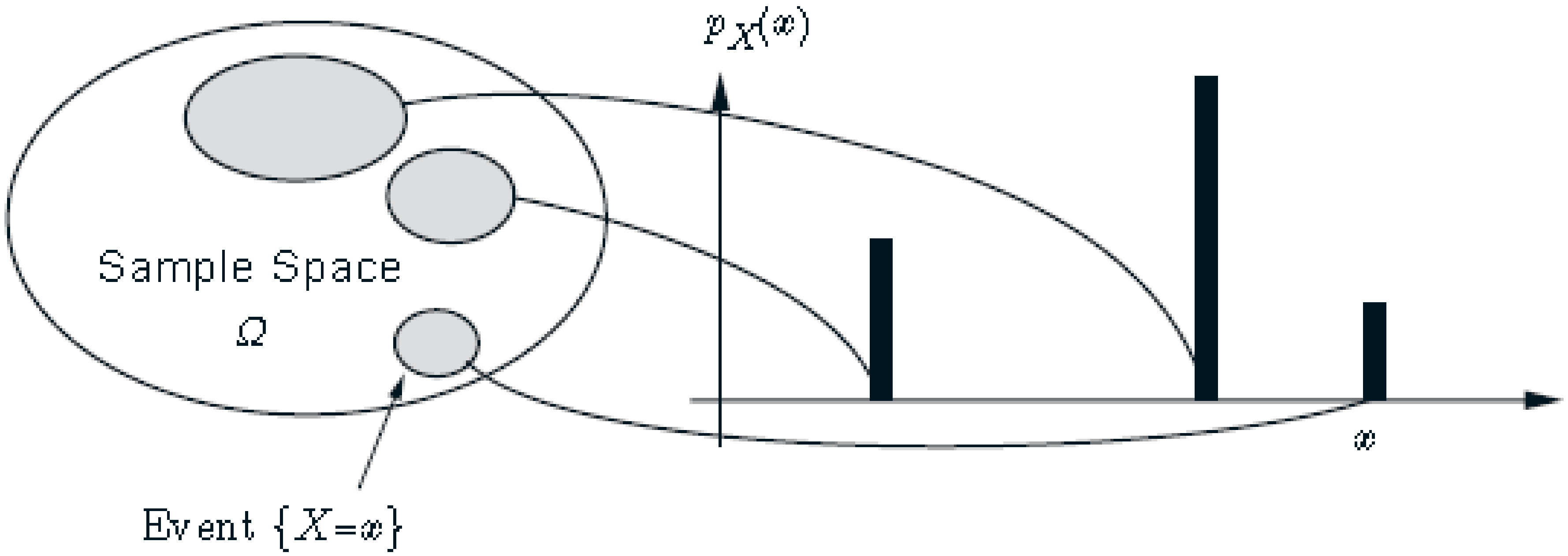


(b)

Concepts Related to Discrete Random Variables

Starting with a probabilistic model of an experiment:

- A **discrete random variable** is a real-valued function of the outcome of the experiment that can take a finite or countably infinite number of values.
- A (discrete) random variable has an associated **probability mass function** (PMF), which gives the probability of each numerical value that the random variable can take.
- A **function of a random variable** defines another random variable, whose PMF can be obtained from the PMF of the original random variable.



- Notation:

$$
\begin{aligned}
p_X(x) &= \mathrm{P}(X = x) \\
&= \mathrm{P}(\{\omega \in \Omega \text{ s.t. } X(\omega) = x\})
\end{aligned}
$$

- $p_X(x) \geq 0 \qquad \sum_x p_X(x) = 1$

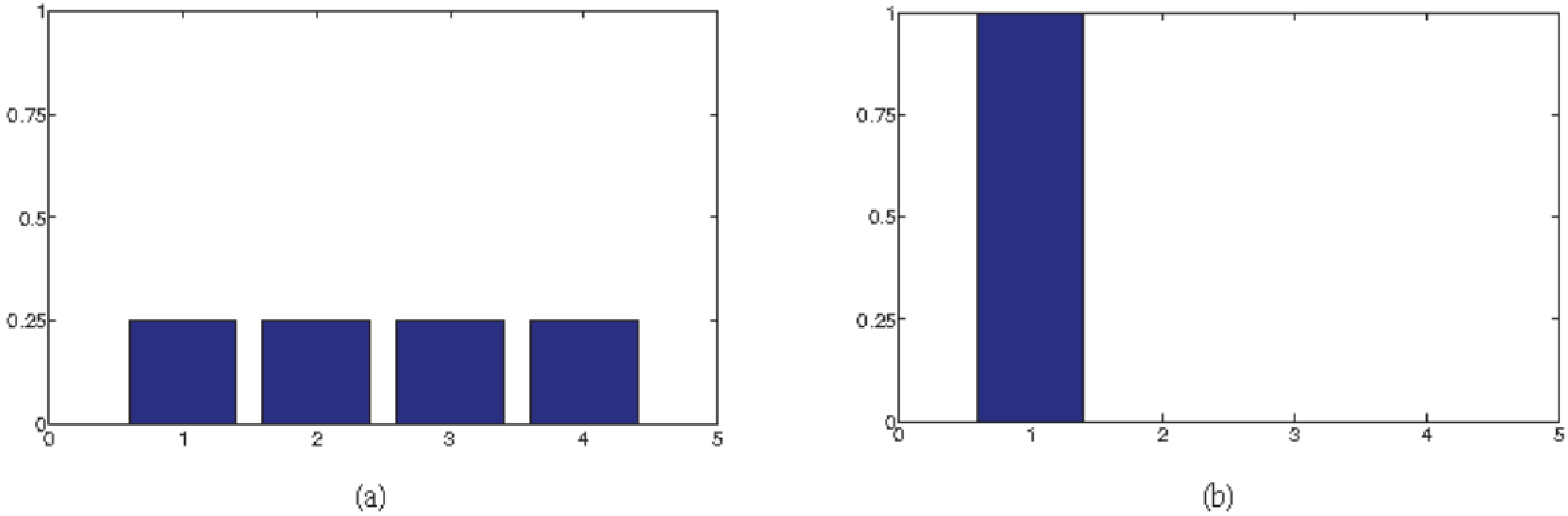
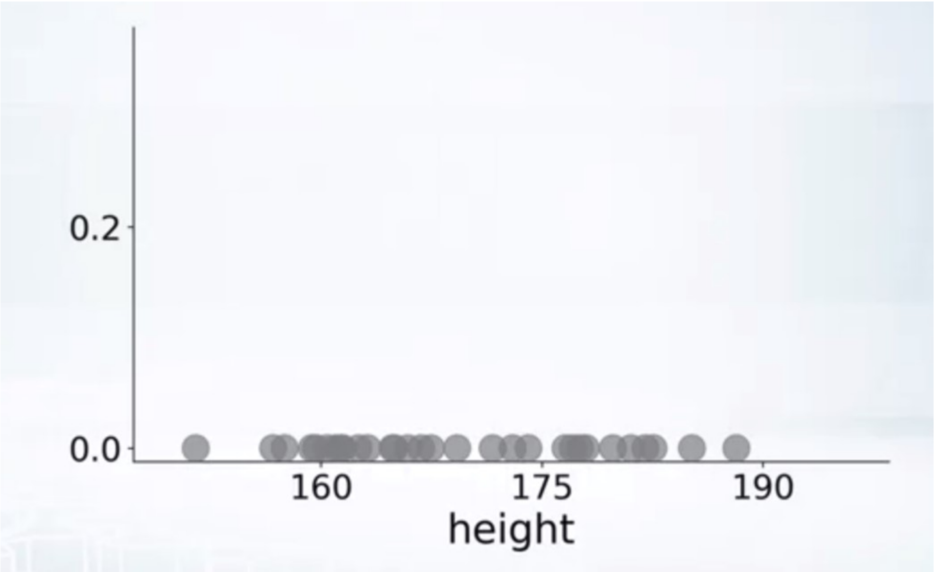
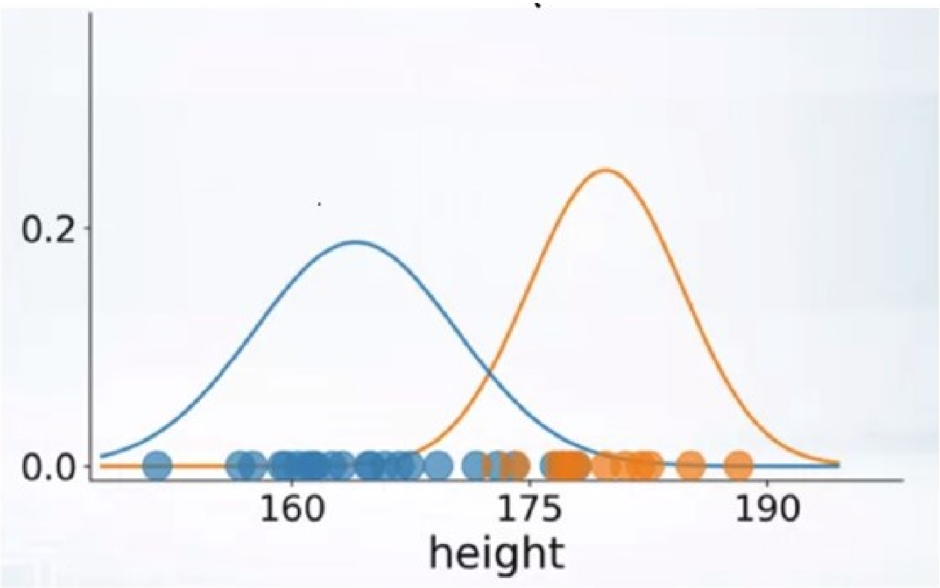
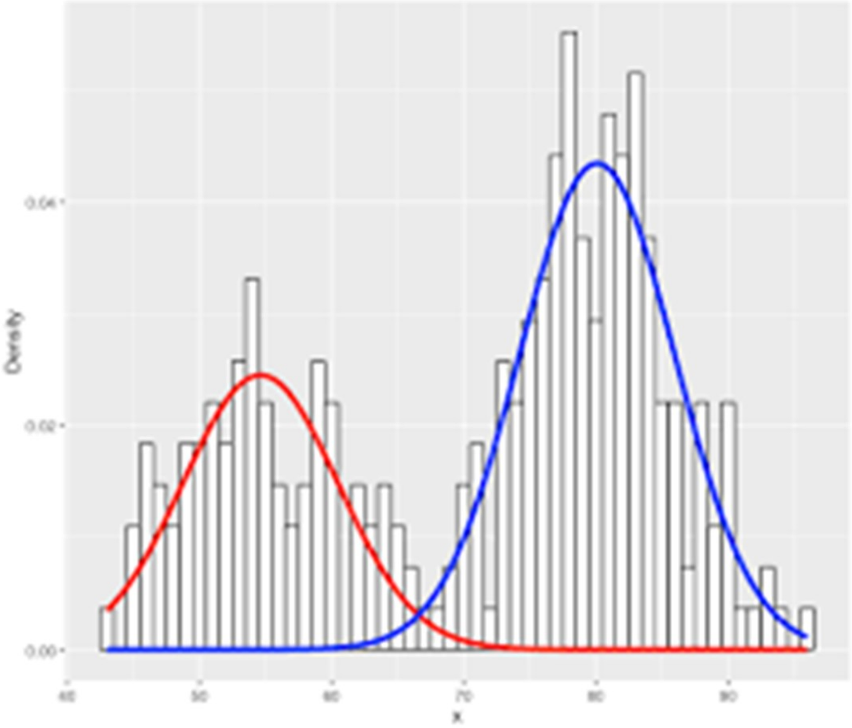
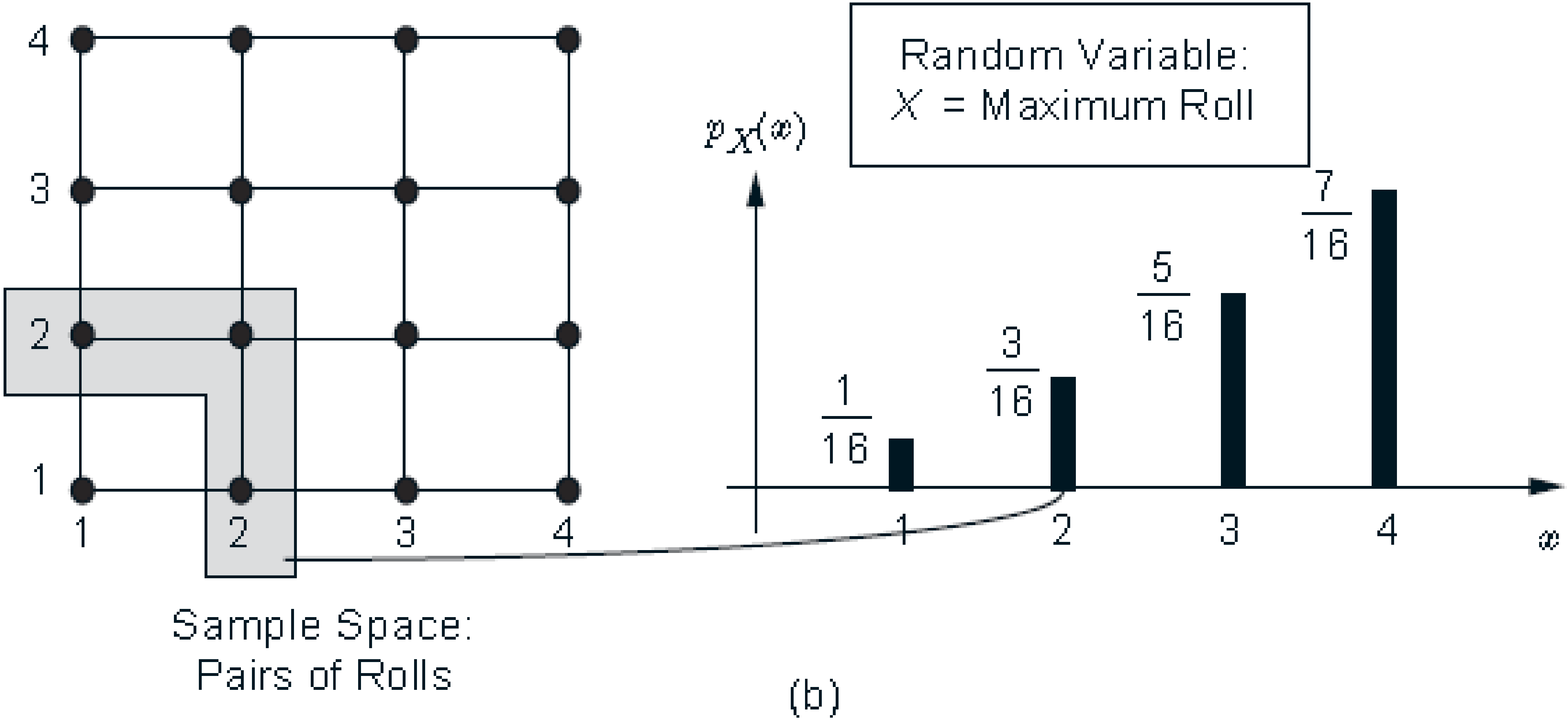


Figure 2.1 (A) a uniform distribution on $\{1, 2, 3, 4\}$, with $p(x = k) = 1/4$. (b) a degenerate distribution $p(x) = 1$ if $x = 1$ and $p(x) = 0$ if $x \in \{2, 3, 4\}$. Figure generated by `discreteProbDistFig`.


0.2
0.0
160
175
190
height


0.2
0.0
160
175
190
height


0.04
0.02
0.00
Density
x


Random Variable:
X = Maximum Roll
$p_X(x)$
1/16
3/16
5/16
7/16
1 2 3 4
x
Sample Space:
Pairs of Rolls
(b)

The Bernoulli Random Variable

$$X = \begin{cases} 1 & \text{if a head,} \\ 0 & \text{if a tail.} \end{cases}$$

Its PMF is

$$p_X(x) = \begin{cases} p & \text{if } x = 1, \\ 1 - p & \text{if } x = 0. \end{cases}$$

INDEPENDENT AND IDENTICALLY DISTRIBUTED

IID

Each attempt is independent and has the same probability of success: the attempts are INDEPENDENT AND IDENTICALLY DISTRIBUTED or IID.

GEOMETRIC DISTRIBUTION

Example 5.2.1. Suppose we flip a fair coin several times in a row until we get a head. Let X be the total number of flips. What is the distribution over X?

$$P(X = n) = (1/2)^n \text{ for positive integer } n.$$

Example 5.2.2. Suppose we perform the same type of experiment as in the previous example, but our coin isn't fair: the probability of getting a head is some value p. Now what is the distribution over X, the number of flips required to get a head?

$$\mathrm{P}(TT\cdots TH)$$

$$P(X = n) = (1 - p)^{n-1} p \qquad \text{for positive integer } n.$$

GEOMETRIC DISTRIBUTION

$$P(X = n) = (1 - p)^{n-1} p \qquad \text{for positive integer } n.$$

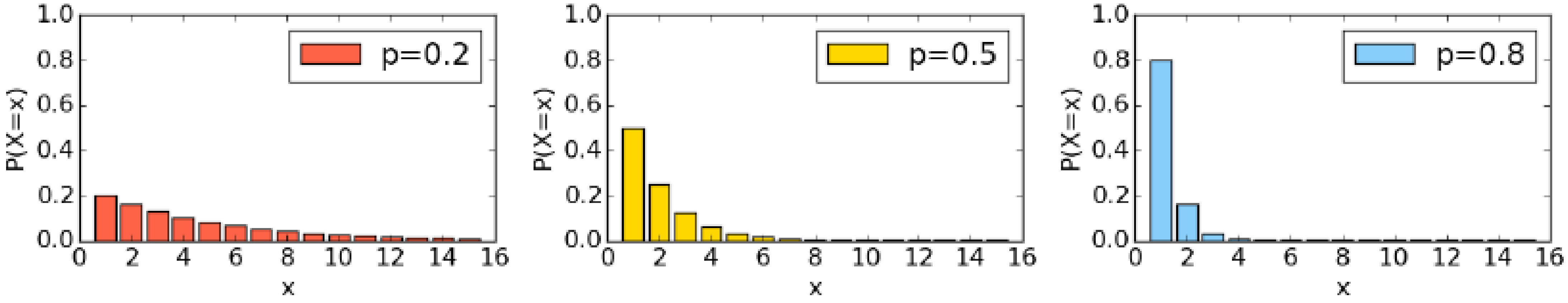


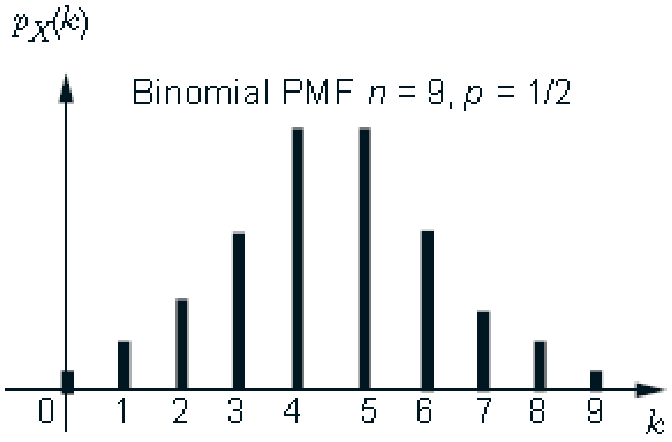
Figure 4: Three examples of geometric distributions, with parameter values of 0.2, 0.5, and 0.8. Possible values of X extend to infinity, but we only show values up to 15.

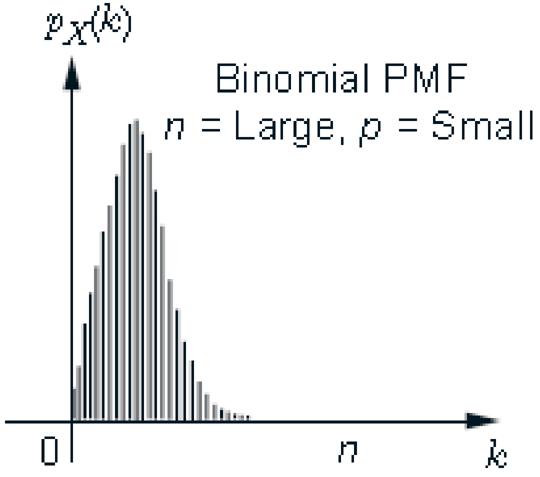
The Binomial Random Variable

A biased coin is tossed n times. At each toss, the coin comes up a head with probability p, and a tail with probability $1-p$, independently of prior tosses. Let X be the number of heads in the n-toss sequence. We refer to X as a **binomial random variable with parameters n and p**. The PMF of X consists of the binomial probabilities that were calculated in Section 1.4:

$$p_X(k) = \mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n.$$

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





Marginal Distribution Calculation ➔ Integrating "out" the unwanted variable

verbose	concise		
$P(X=x)=1-P(X\neq x)$	$P(x)=1-P(\neg x)$	complement	(18)
$P(X=x)=\sum_y P(X=x,Y=y)$	$P(x)=\sum_y P(x,y)$	law of total prob	(19)
$P(X=x\mid Y=y)=\dfrac{P(X=x,Y=y)}{P(Y=y)}$	$P(x\mid y)=\dfrac{P(x,y)}{P(y)}$	defn of cond prob	(20)
$P(X=x,Y=y)=P(X=x\mid Y=y)P(y)$	$P(x,y)=P(x\mid y)P(y)$	product rule	(21)
$P(X=x\mid Y=y)=\dfrac{P(Y=y\mid X=x)P(X=x)}{P(Y=y)}$	$P(x\mid y)=\dfrac{P(y\mid x)P(x)}{P(y)}$	Bayes' Rule	(22)

JOINT PMFS OF MULTIPLE RANDOM VARIABLES

$$p_{X,Y}(x,y) = \mathbf{P}(X = x, Y = y)$$

$$\mathbf{P}\big((X,Y) \in A\big) = \sum_{(x,y)\in A} p_{X,Y}(x,y).$$

marginal PMFs:

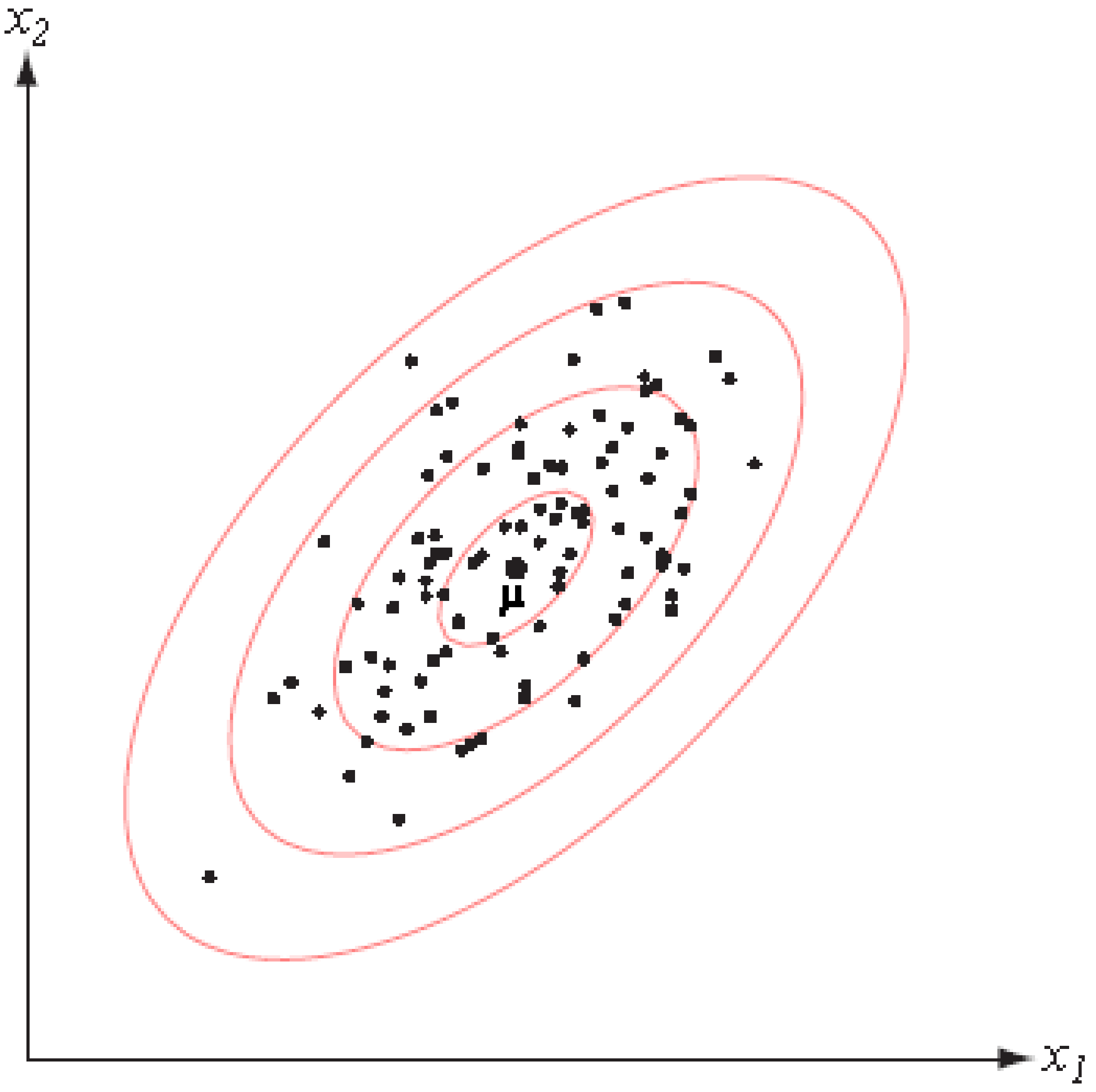
$$p_X(x) = \sum_y p_{X,Y}(x,y),$$

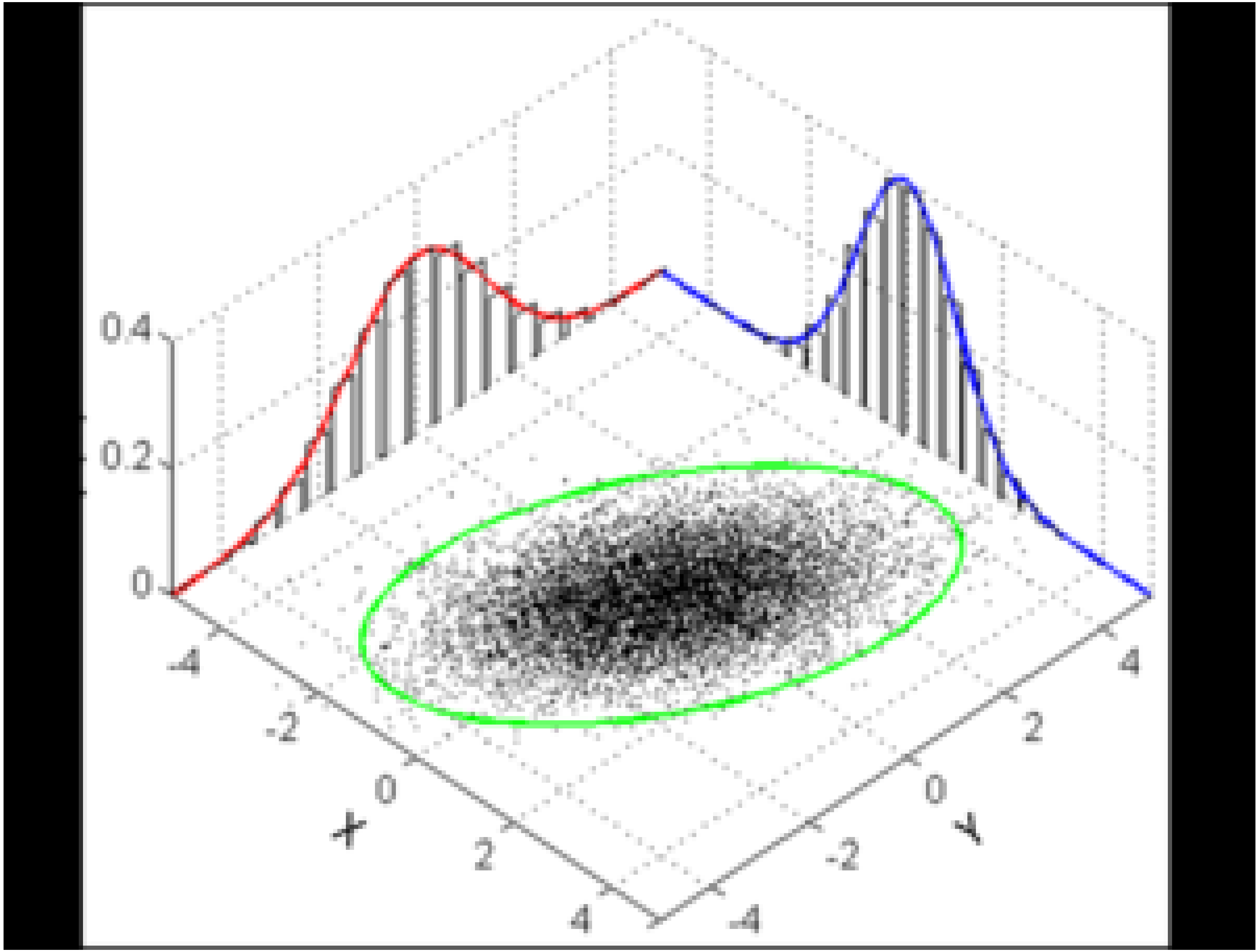
$$p_Y(y) = \sum_x p_{X,Y}(x,y).$$

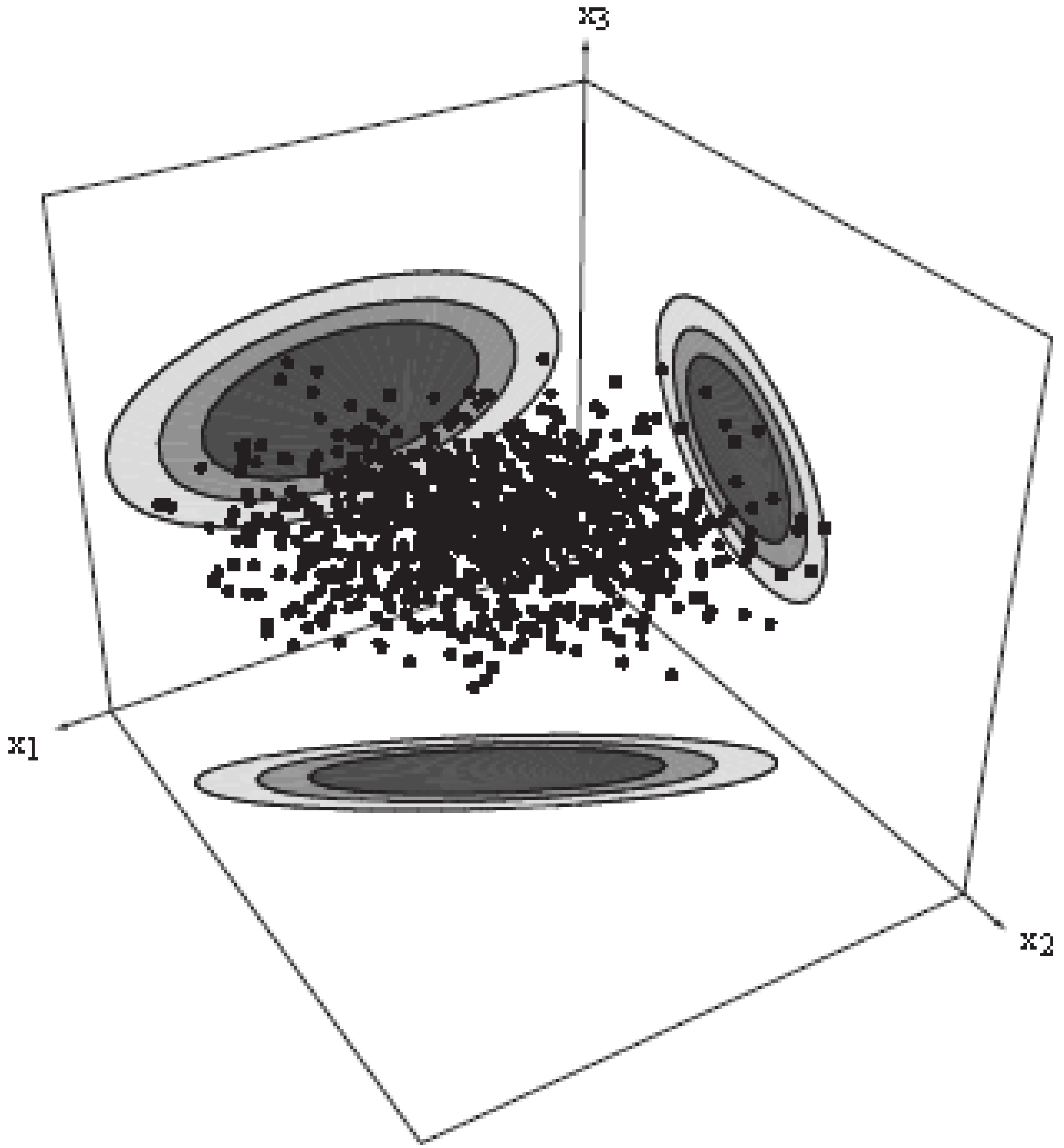
Joint PMF $P_{X,Y}(x,y)$
in tabular form

y \ x	1	2	3	4	Row Sums: Marginal PMF $P_Y(y)$
4	0	1/20	1/20	1/20	3/20
3	1/20	2/20	3/20	1/20	7/20
2	1/20	2/20	3/20	1/20	7/20
1	1/20	1/20	1/20	0	3/20
Column Sums: Marginal PMF $P_X(x)$	3/20	6/20	8/20	3/20	

$$\mathbf{P}\big((X,Y)\in A\big) = \sum_{(x,y)\in A} p_{X,Y}(x,y).$$



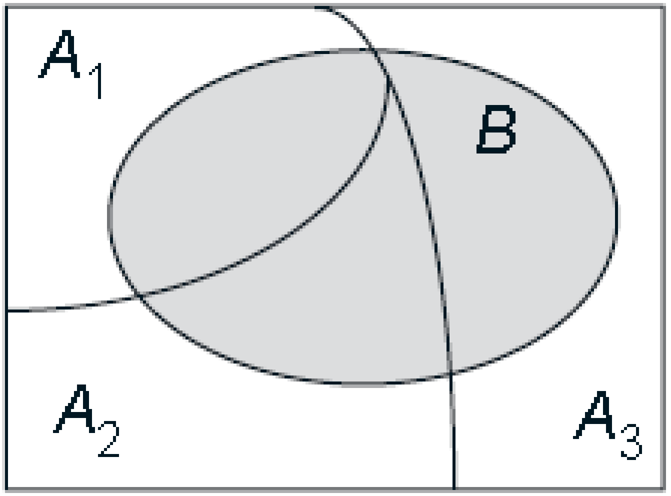

0.4
0.2
0
-4
-2
0
2
4
X
4
2
0
-2
-4
Y



x3
x1
x2

$$p(x) = \int_{-\infty}^{\infty} p(x, y)\, dy$$

law of tot prob, continuous version



$$B = (A_1 \cap B) \cup \cdots \cup (A_n \cap B).$$

$$\mathbf{P}(B) = \mathbf{P}(A_1 \cap B) + \cdots + \mathbf{P}(A_n \cap B).$$

$$P(B) = \sum_{i=1}^{n} P(A_i, B)$$

if X_1,X_2,...Xn are **continuous random variables**, then the marginal probability density function should be

$$f_{X_i}(x_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_{i-1} dx_{i+1} \ldots dx_n.$$

EXPECTED VALUE
EXPECTATION

MEAN

μ

Expectation

We define the **expected value** (also called the **expectation** or the **mean**) of a random variable X, with PMF $p_X(x)$, by†

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

Expected Value Rule for Functions of Random Variables

Let X be a random variable with PMF $p_X(x)$, and let $g(X)$ be a real-valued function of X. Then, the expected value of the random variable $g(X)$ is given by

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

$$M = \frac{m_1k_1 + m_2k_2 + \cdots + m_nk_n}{k}.$$

$$p_i \approx \frac{k_i}{k}, \qquad i = 1, \ldots, n.$$

$$M = \frac{m_1k_1 + m_2k_2 + \cdots + m_nk_n}{k} \approx m_1p_1 + m_2p_2 + \cdots + m_np_n.$$

$$M = \sum_{i=1}^{n} m_ip_i \qquad \Longrightarrow \qquad \mathbf{E}[X] = \sum_x xp_X(x).$$

Expected Value Rule for Functions of Random Variables

Let X be a random variable with PMF $p_X(x)$, and let $g(X)$ be a real-valued function of X. Then, the expected value of the random variable $g(X)$ is given by

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

$$\begin{aligned}
\mathbf{E}[g(X)] &= \mathbf{E}[Y] \\
&= \sum_y y p_Y(y) \\
&= \sum_y y \sum_{\{x \mid g(x)=y\}} p_X(x) \\
&= \sum_y \sum_{\{x \mid g(x)=y\}} y p_X(x) \\
&= \sum_y \sum_{\{x \mid g(x)=y\}} g(x) p_X(x) \\
&= \sum_x g(x) p_X(x).
\end{aligned}$$

VARIANCE

$$\mathrm{var}(X) = \mathbf{E}\left[\left(X - \mathbf{E}[X]\right)^2\right].$$

STANDARD DEVIATION σ

$$\sigma_X = \sqrt{\mathrm{var}(X)}.$$

$$\mathrm{var}(X) = \mathbf{E}\left[\left(X - \mathbf{E}[X]\right)^2\right] = \sum_x \left(x - \mathbf{E}[X]\right)^2 p_X(x).$$

Variance in Terms of Moments Expression

$$\mathrm{var}(X) = \mathbf{E}[X^2] - \left(\mathbf{E}[X]\right)^2.$$

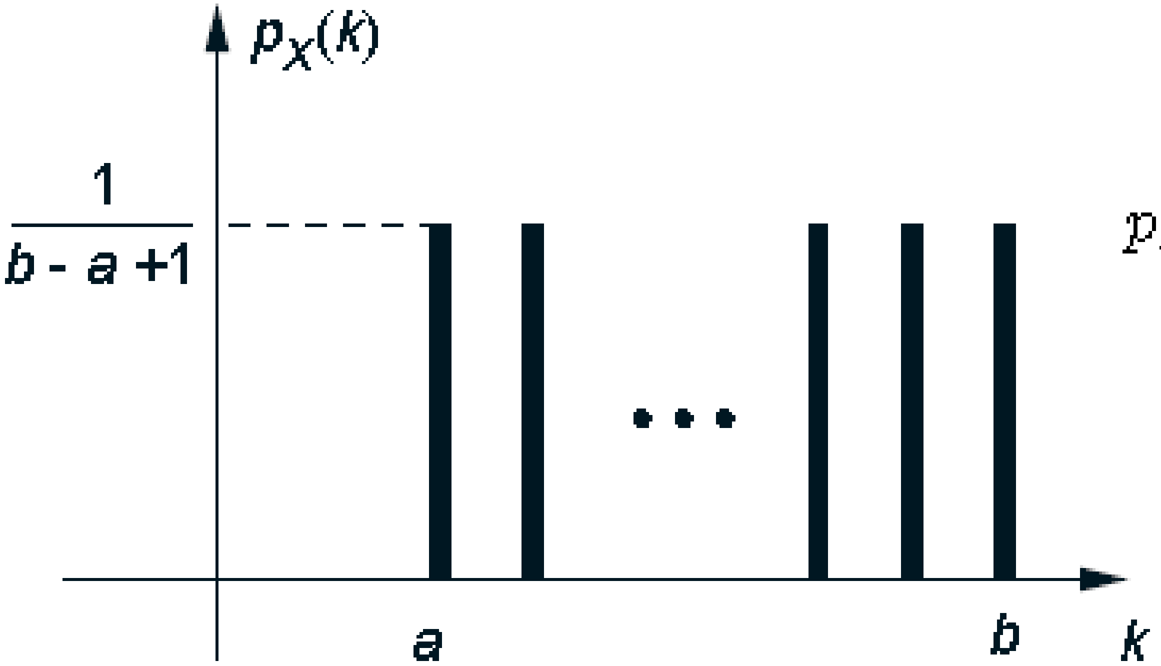
nth moment

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

Discrete Uniform Random Variable.

roll of a fair six-sided die

$$p_X(k) = \begin{cases} 1/6 & \text{if } k = 1, 2, 3, 4, 5, 6, \\ 0 & \text{otherwise.} \end{cases}$$



$$p_X(k) = \begin{cases} \dfrac{1}{b - a + 1} & \text{if } k = a, a+1, \ldots, b, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathbf{E}[X] = \frac{a + b}{2},$$

$$\text{var}(X) = \frac{(b - a)(b - a + 2)}{12}.$$