COMPUTATIONAL ASTROPHYSICS

Observatorio
Astronómico
Nacional

Computational Astrophysics

03A. Introduction to Statistics

Eduard Larrañaga
Observatorio Astronómico Nacional
Universidad Nacional de Colombia

Kolmogorov Axioms

Kolmogorov Axioms

Let Ω be a collection of possible elementary events and let A and B be two events such that $A, B \in \Omega$.

The *probability* of occurrence of the event A is a real number denoted as $P(A)$, satisfying the following axioms (Kolmogorov),

1. $P(A) \geq 0$ for each A
2. $P(\Omega) = 1$
3. For all countable disjoint sets $A_1, A_2, \ldots \in \Omega$,

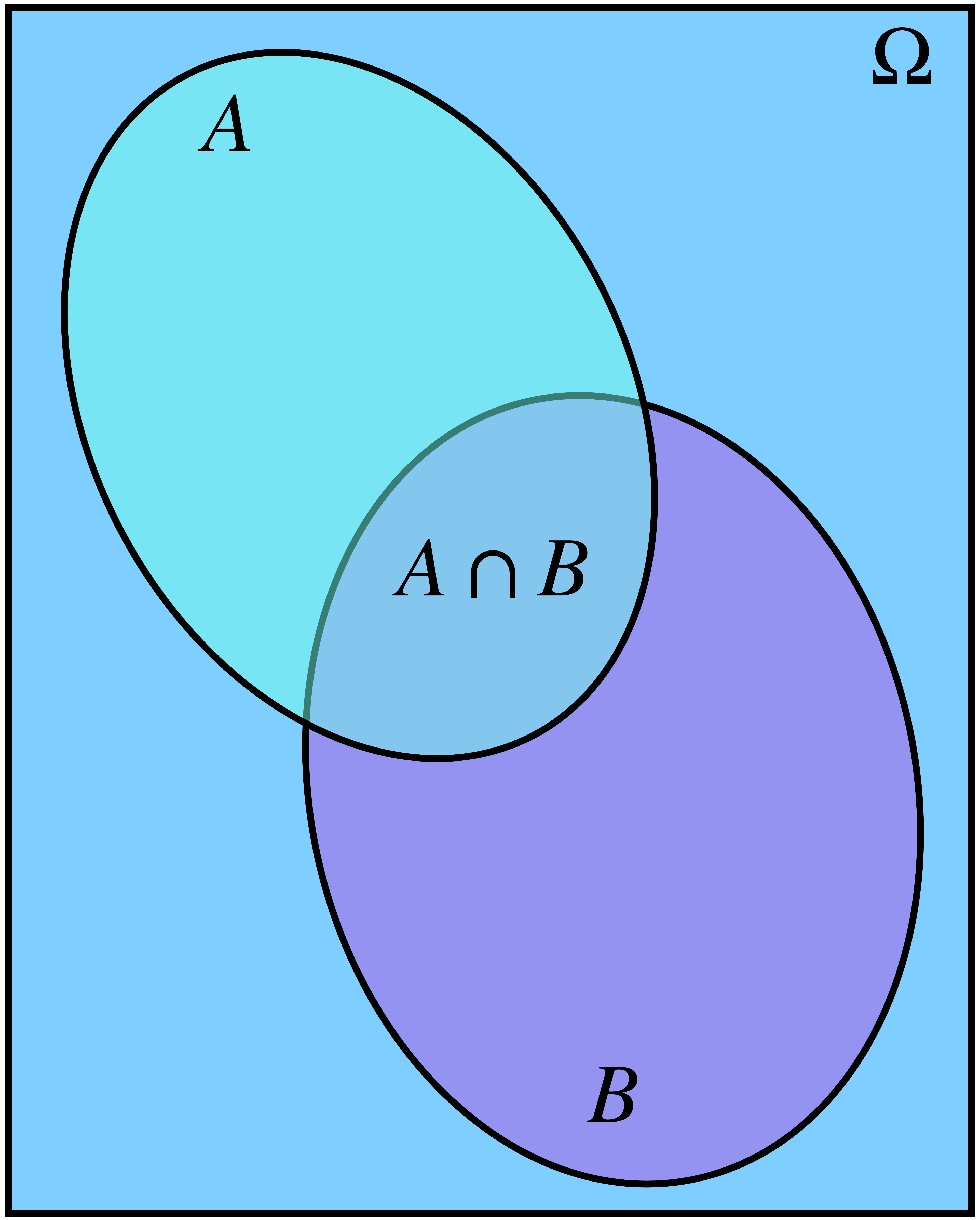
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$$

Kolmogorov Axioms

As a result of the above axioms, we have the following results,

$$P(A) + P(A^c) = 1$$

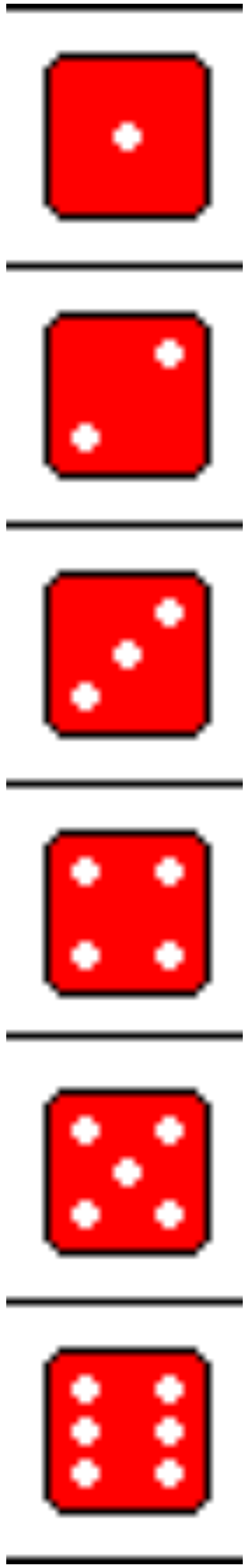
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Random Variables

A **random**, or **stochastic**, variable is that whose value results from the measurement of a quantity that is subject to random variations. It can take on a set of possible different values, each with an associated probability.

There are **discrete** and **continuos** random variables.



$$P(⚀) = \frac{1}{6}$$

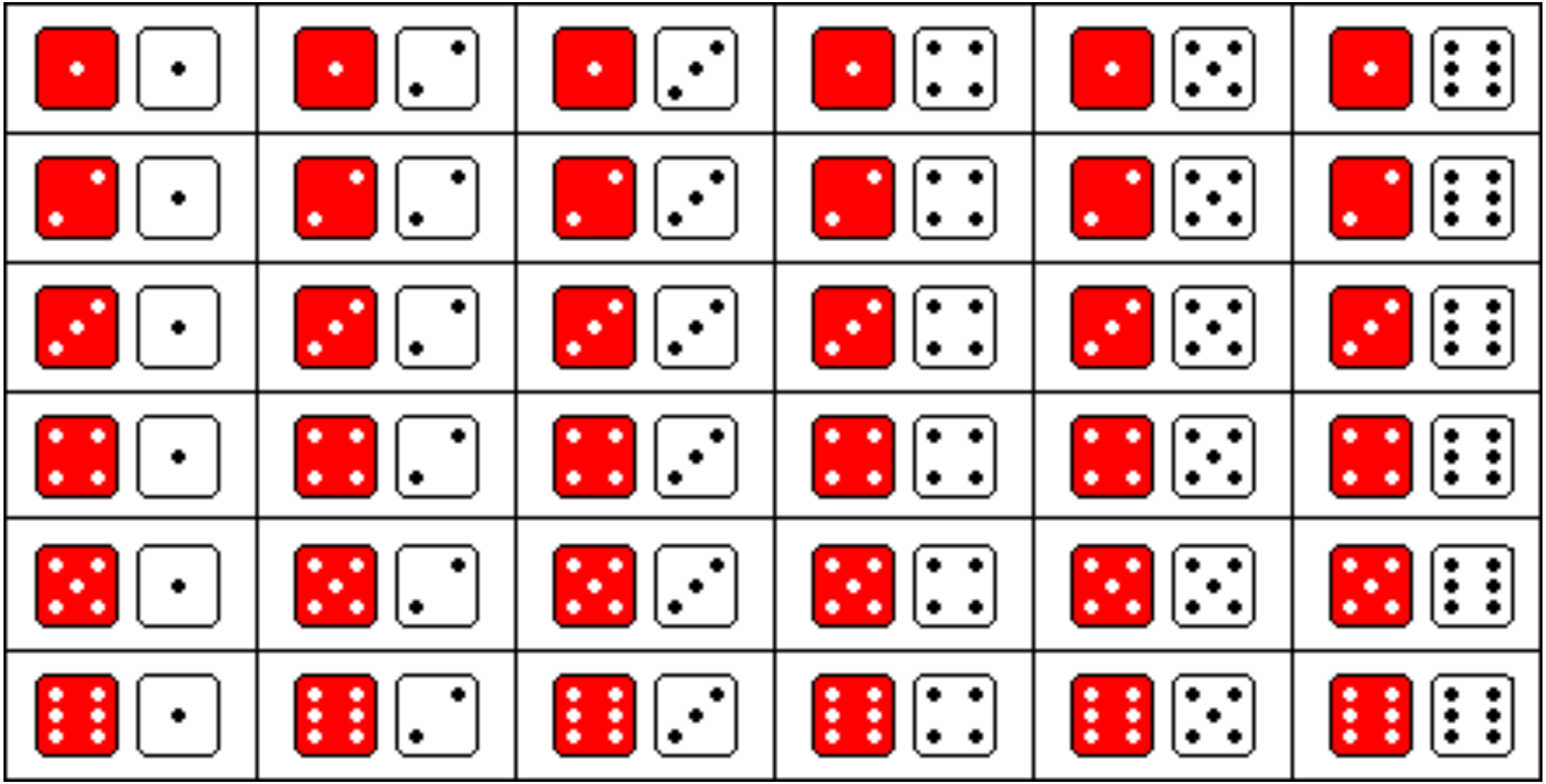
$$P(⚁) = \frac{1}{6}$$

$$P(⚂) = \frac{1}{6}$$

$$P(⚃) = \frac{1}{6}$$

$$P(⚄) = \frac{1}{6}$$

$$P(⚅) = \frac{1}{6}$$

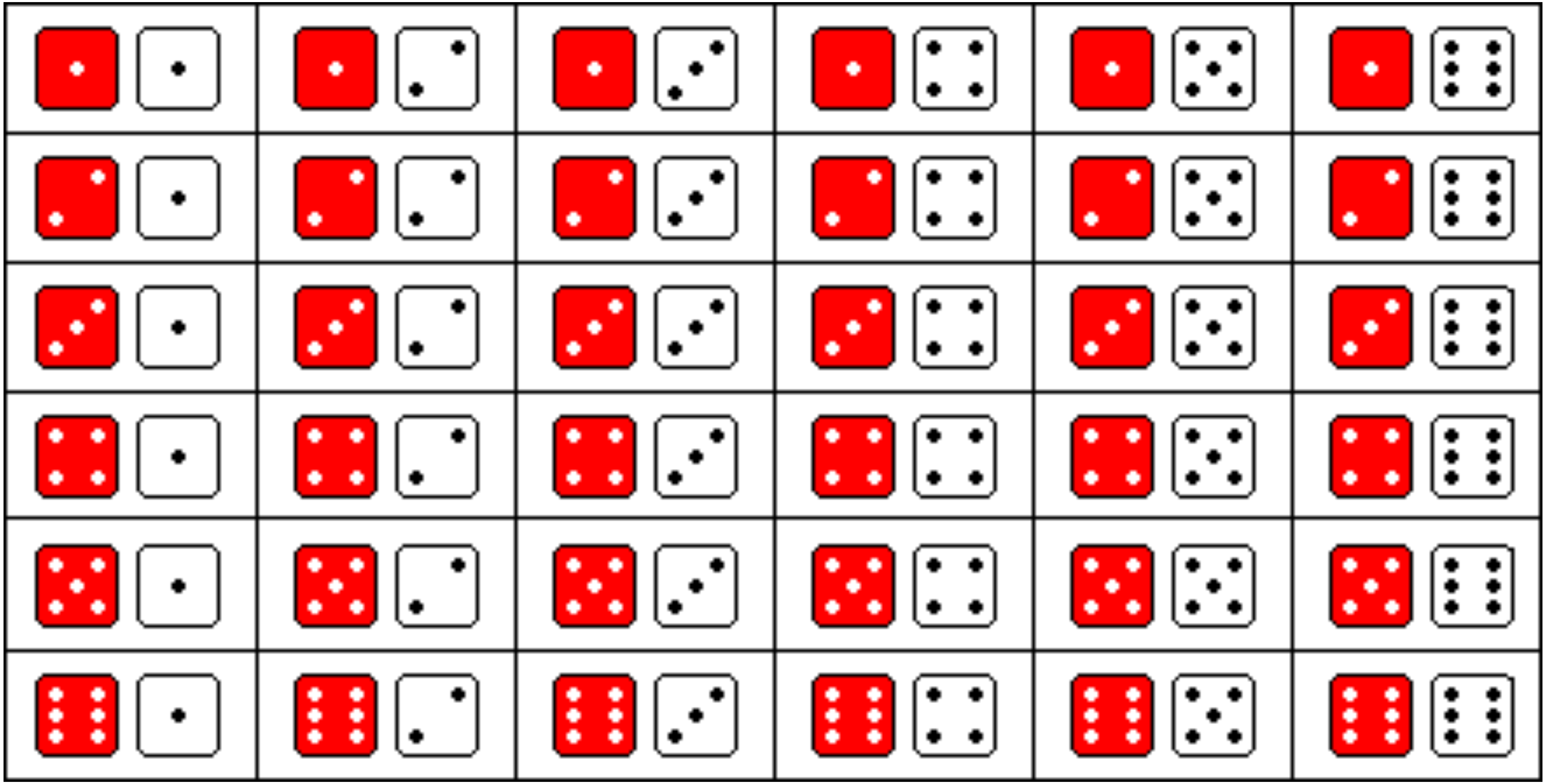


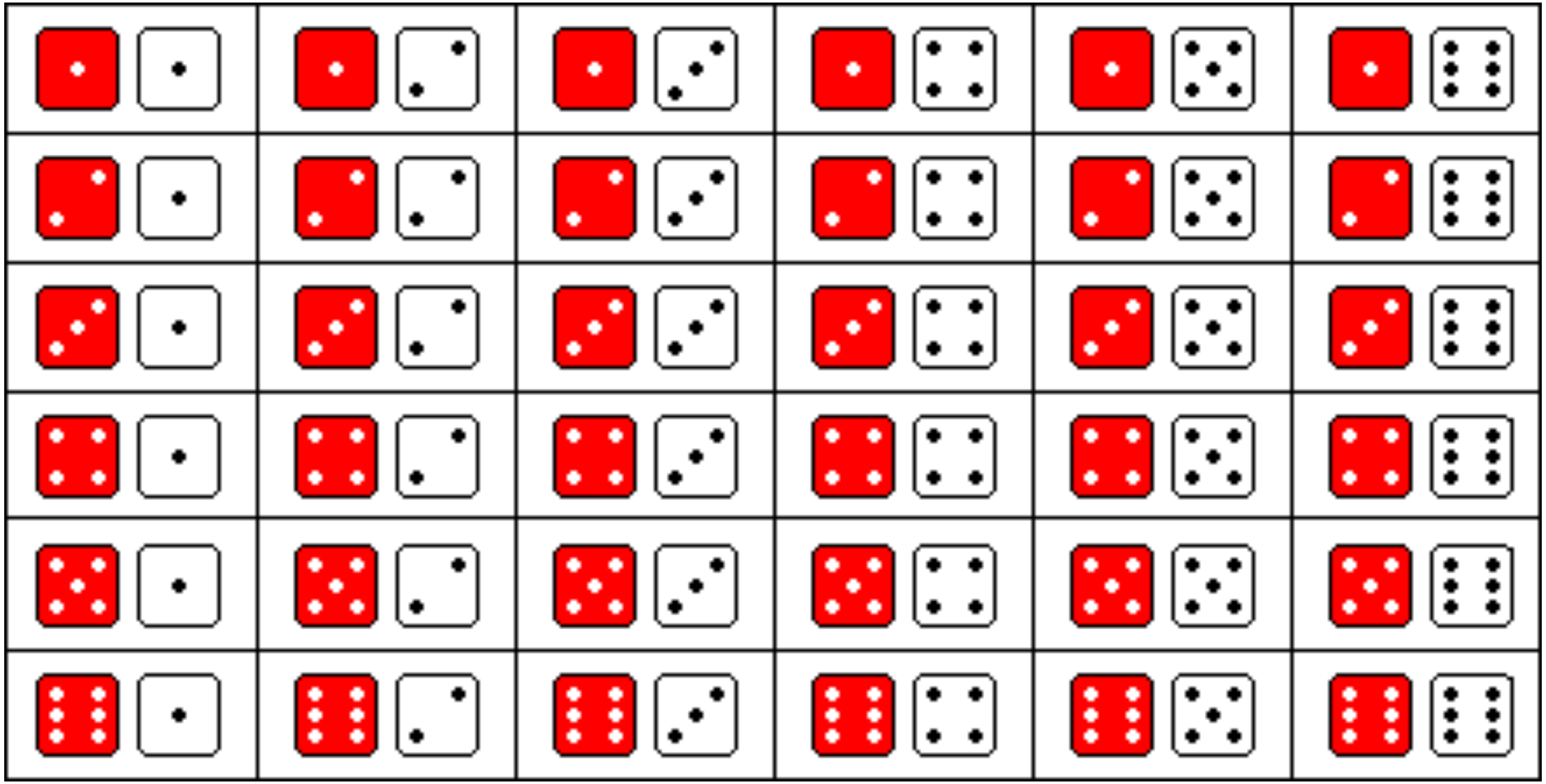
Independent Random Variables
Unconditional Probability

Two random variables, x and y are *independent* if

$$p(x, y) = p(x \cap y) = p(x)p(y)$$

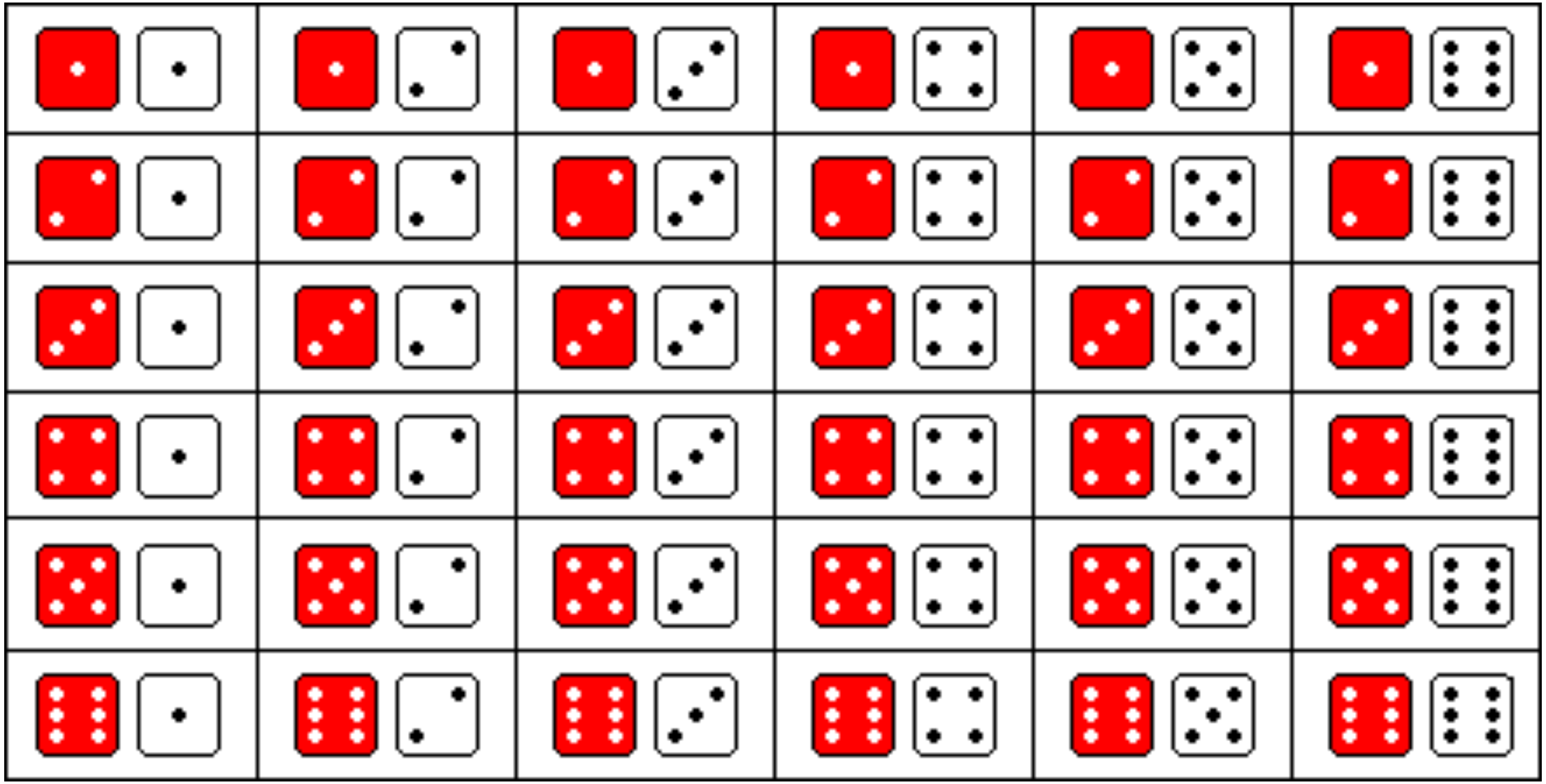
i.e., the knowledge of the value of x tells nothing about the value of y.





$P($$,$$) = P($$\cap$$)$

$\frac{1}{6}$ $\frac{1}{6}$



$P($  ,  $) = P($  $\cap$  $) = \dfrac{1}{6} \times \dfrac{1}{6} = \dfrac{1}{36}$

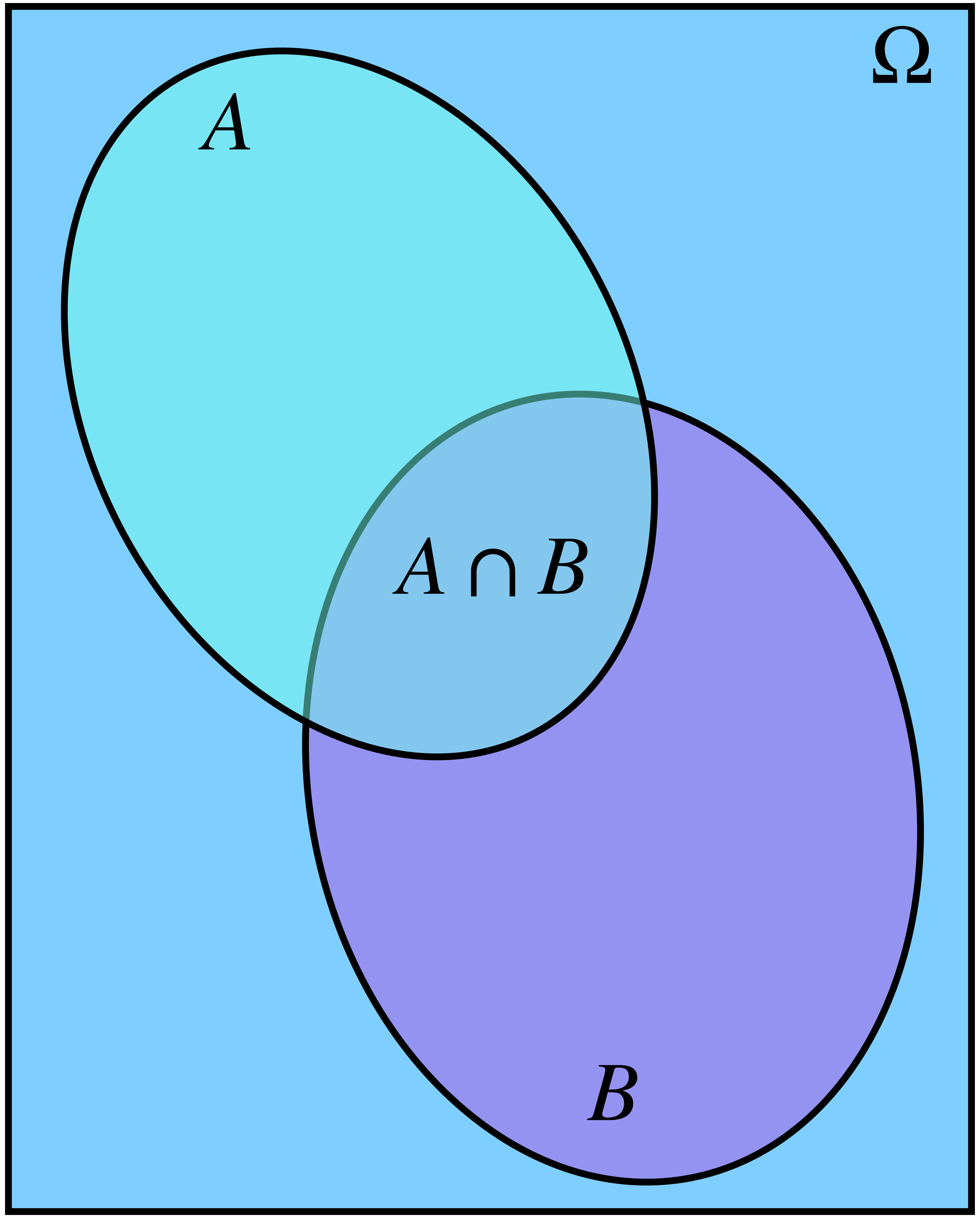
Conditional Probabilities

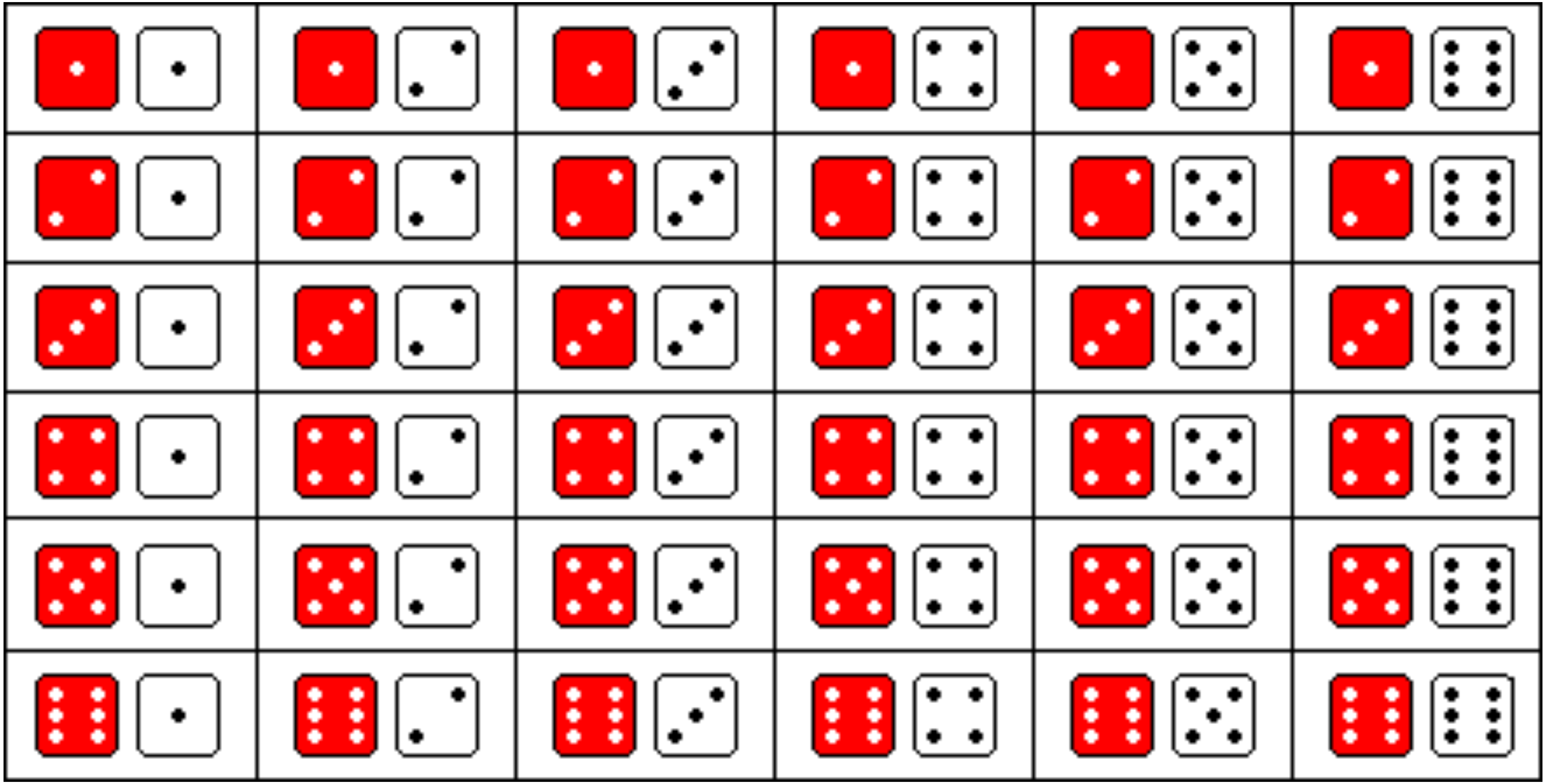
When the variables are *not independent*, we have a *conditional probability, defined by*

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

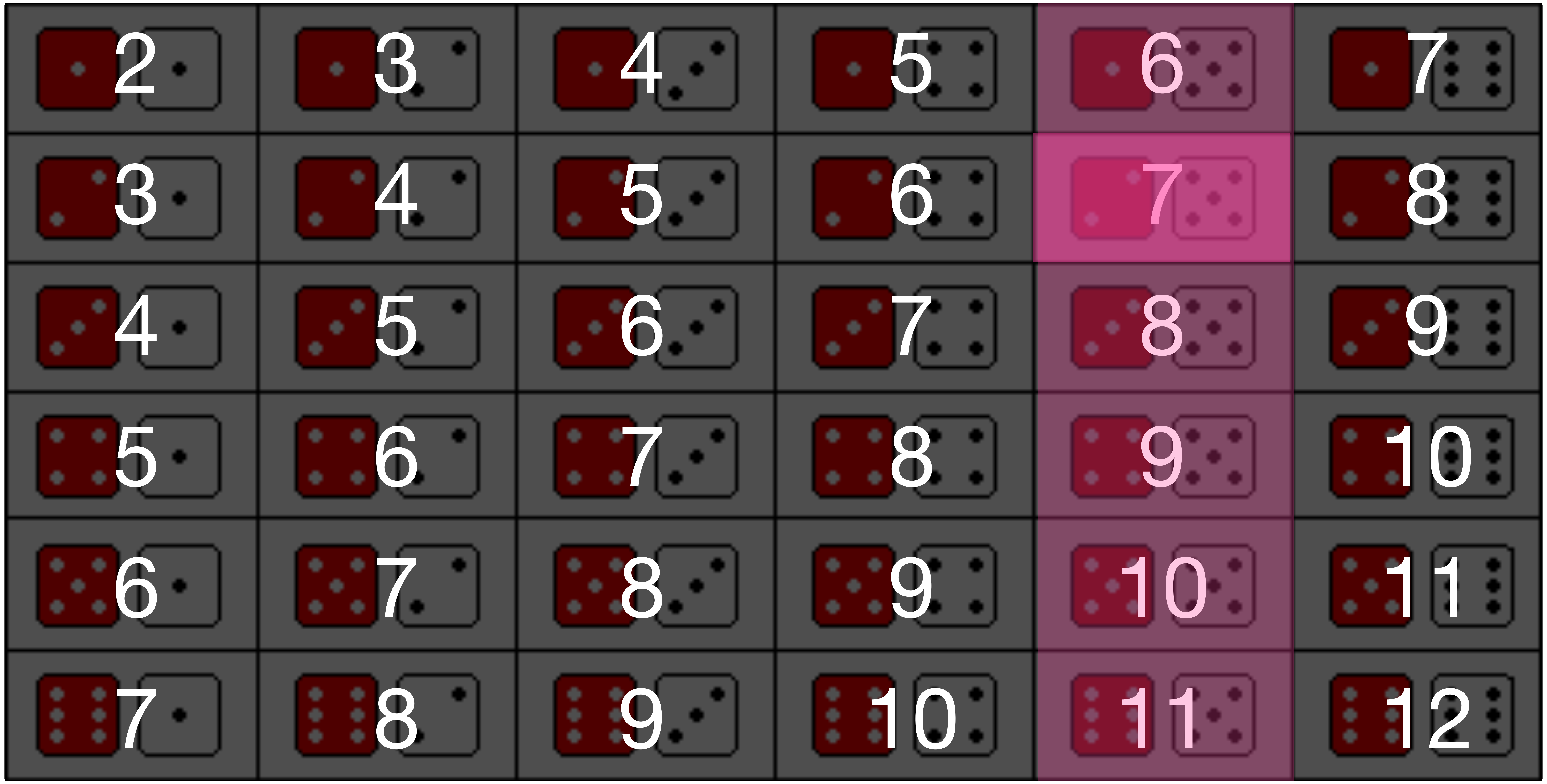
$P(A \mid B)$:

Probability of event A given that B occurs.



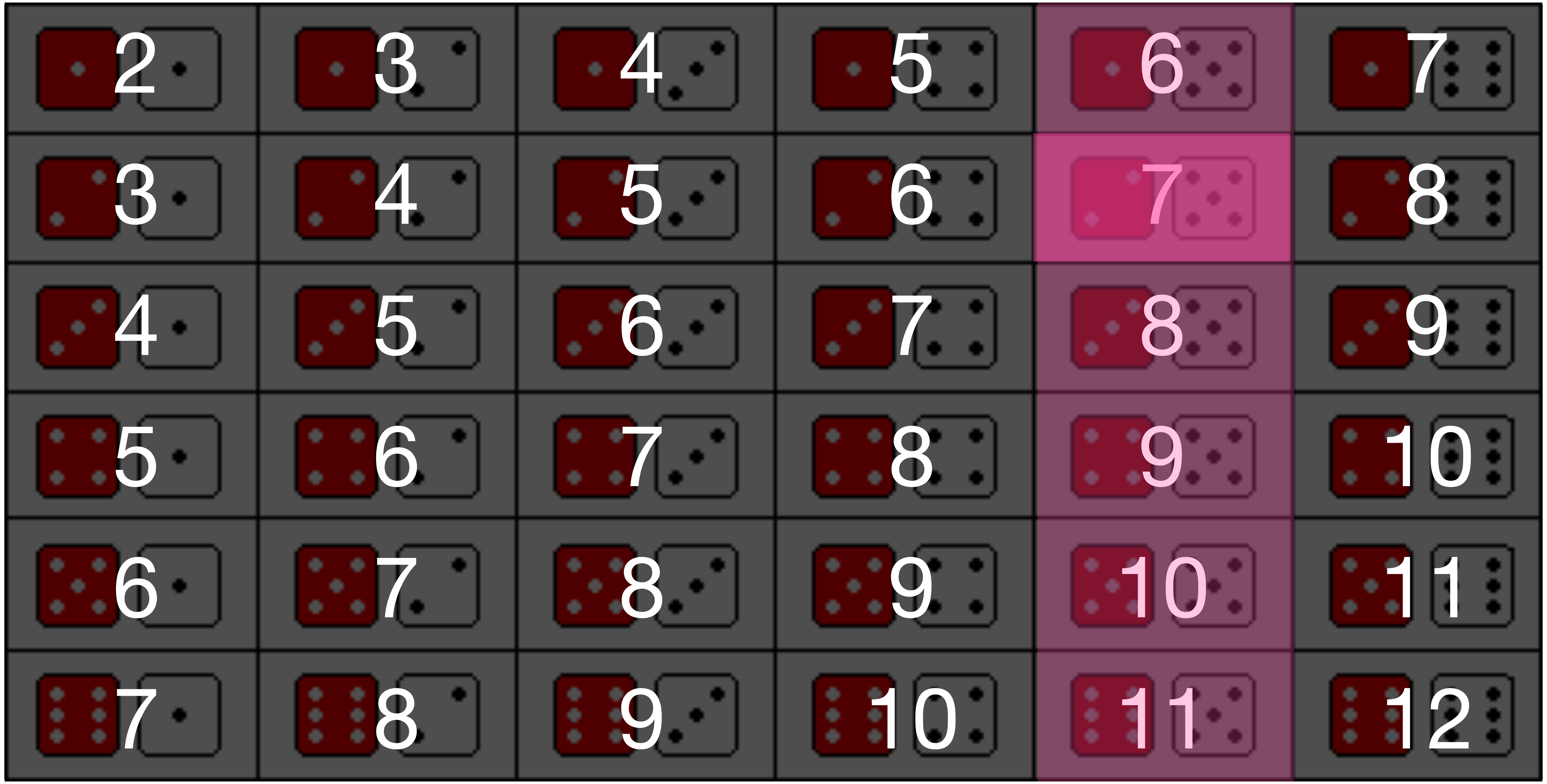



2 3 4 5 6 7
3 4 5 6 7 8
4 5 6 7 8 9
5 6 7 8 9 10
6 7 8 9 10 11
7 8 9 10 11 12



$P($ 7 , ⚄ $) = P($ 7 $\cap$ ⚄ $)$

$\frac{1}{6}$ $\frac{1}{6}$



$$P(7, ⚄) = P(7 \cap ⚄) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

Conditional Probabilities

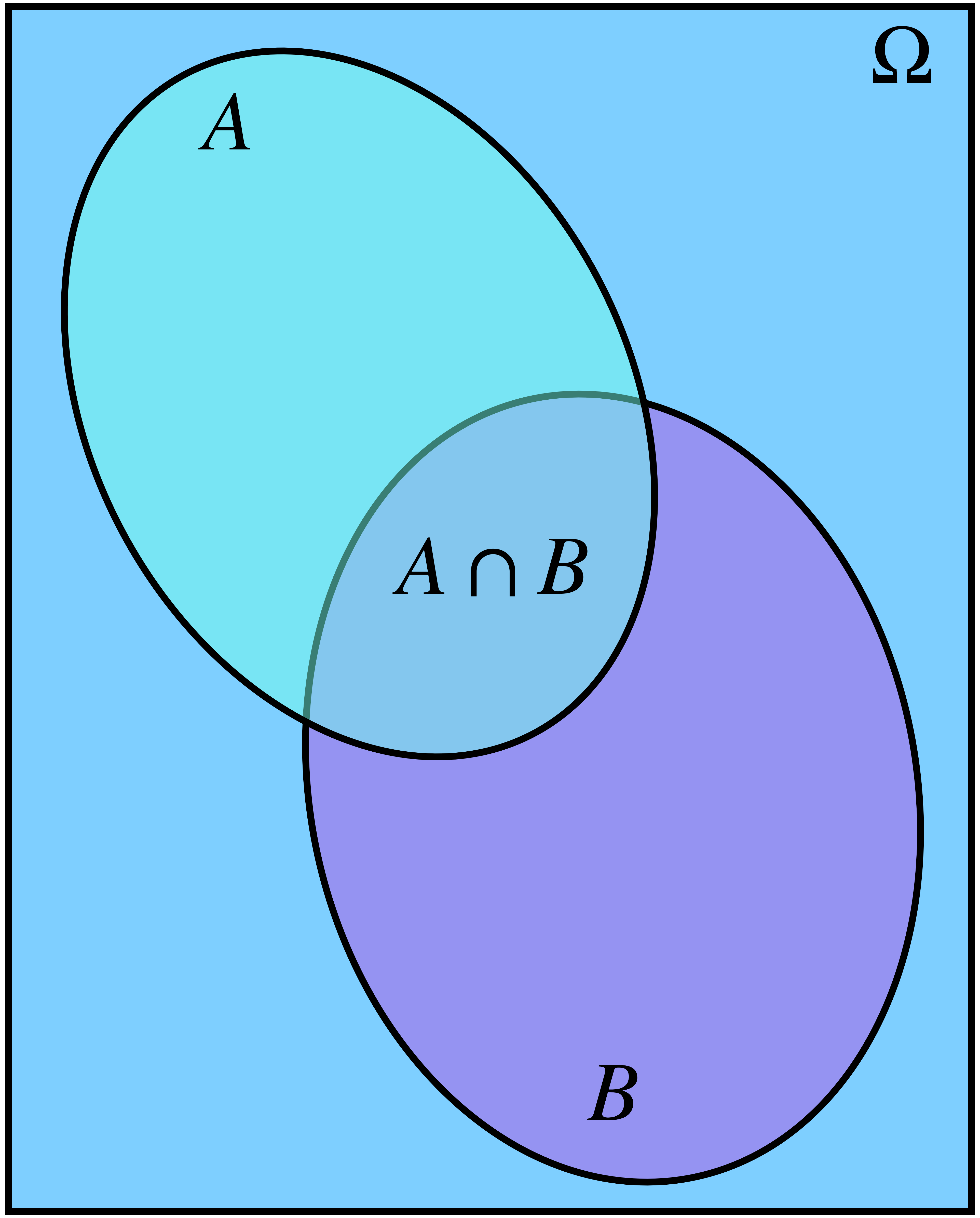
When the variables are *not independent, we have a conditional probability,*

.

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

$P(A \mid B)$:

Probability of event A given that B occurs.



Law of Total Probability

Given a set $B_1, B_2, \ldots, B_N \in \Omega$ of disjoints events and such that

$$\bigcup_{i=1}^{\infty} B_i = \Omega,$$

then

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A \mid B_i) P(B_i)$$

$P(\ 8\ )$

$$P(8) = 0 + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}$$

$$P(8) = 0 + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}$$

$$P(8) = \frac{5}{36}$$

Law of Total Probability

Given a set $B_1, B_2, \ldots, B_N \in \Omega$ of disjoints events, such that

$$\bigcup_{i=1}^{\infty} B_i = \Omega$$

and assuming that an event C is not mutually exclusive with A or any of the events B_i, then

$$P(A \mid C) = \sum_i P(A \mid C \cap B_i) P(B_i \mid C)$$

Random Variables and Bayes’ Rule

Random Variables and Probability Density Functions

A **random**, or **stochastic**, variable is that whose value results from the measurement of a quantity that is subject to random variations. It can take on a set of possible different values, each with an associated probability.

There are **discrete** and **continuos** random variables.

The function which ascribes a probability value to each outcome of the random variable is called **probability density function** or **pdf**.

Random Variables and Probability Density Functions

Independent Identically Distributed or **iid** random variables are drawn from the same distribution and are independent. Two random variables, x and y are *independent* if

$$p(x, y) = p(x)p(y)$$

i.e., the knowledge of the value of x tells nothing about the value of y.

When the variables are *not independent, we have*

$$p(x, y) = p(x \mid y)p(y) = p(y \mid x)p(x).$$

Marginal Probability Functions

The **Marginal Probability Function** is defined as

$$p(x) = \int p(x, y)dy$$

Or equivalently

$$p(x) = \int p(x \mid y)p(y)dy$$

Bayes' Rule

From the above results, we obtain the **Bayes' rule,**

$$p(y \mid x) = \frac{p(x \mid y)p(y)}{p(x)} = \frac{p(x \mid y)p(y)}{\int p(x \mid y)p(y)dy}.$$

For discrete random variables, y_i, this expression is writen as

$$p(y_j \mid x) = \frac{p(x \mid y_j)p(y_j)}{p(x)} = \frac{p(x \mid y_j)p(y_j)}{\sum_i p(x \mid y_i)p(y_i)}.$$

Descriptive Statistics. Continuous Variable

Some Definitions in Descriptive Statistics. Continuous Variable

Consider an arbitrary distribution function $h(x)$ of a continuous variable. In the following, we present some important definitions of descriptive statistics.

Arithmetic Mean (Expectation value)

$$\bar{x} = E(x) = \mu = \int_{-\infty}^{\infty} x h(x) dx$$

Some Definitions in Descriptive Statistics. Continuous Variable

Variance

$$\textbf{var} = \sigma^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 h(x) dx$$

Standard Deviation

$$\sigma = \sqrt{\sigma^2} = \sqrt{\int_{-\infty}^{\infty} (x - \bar{x})^2 h(x) dx}$$

*Variance is also known as the second central moment of the distribution.

Some Definitions in Descriptive Statistics. Continuous Variable

Skewness

$$\Sigma = \int_{-\infty}^{\infty} \left(\frac{x - \bar{x}}{\sigma} \right)^3 h(x)dx$$

Kurtosis

$$K = \int_{-\infty}^{\infty} \left(\frac{x - \bar{x}}{\sigma} \right)^4 h(x)dx - 3$$

* Skewness and Kurtosis are the central moments of order 3 and 4, respectively, for the distribution.

Some Definitions in Descriptive Statistics. Continuous Variable

Absolute deviation about x_0

$$\delta = \int_{-\infty}^{\infty} \left| x - x_0 \right| h(x) dx$$

Mode (most probable value for unimodal functions) : x_m

$$\left. \frac{dh(x)}{dx} \right|_{x=x_m} = 0$$

Some Definitions in Descriptive Statistics. Continuous Variable

$p\,\%$ **quantiles:** q_p

$$\frac{p}{100} = \int_{-\infty}^{q_p} h(x)dx$$

p is called a percentile.

The Uniform Distribution

The simplest example of a probability distribution function is the **Uniform (top-hat or box) Distribution**, given by the relation

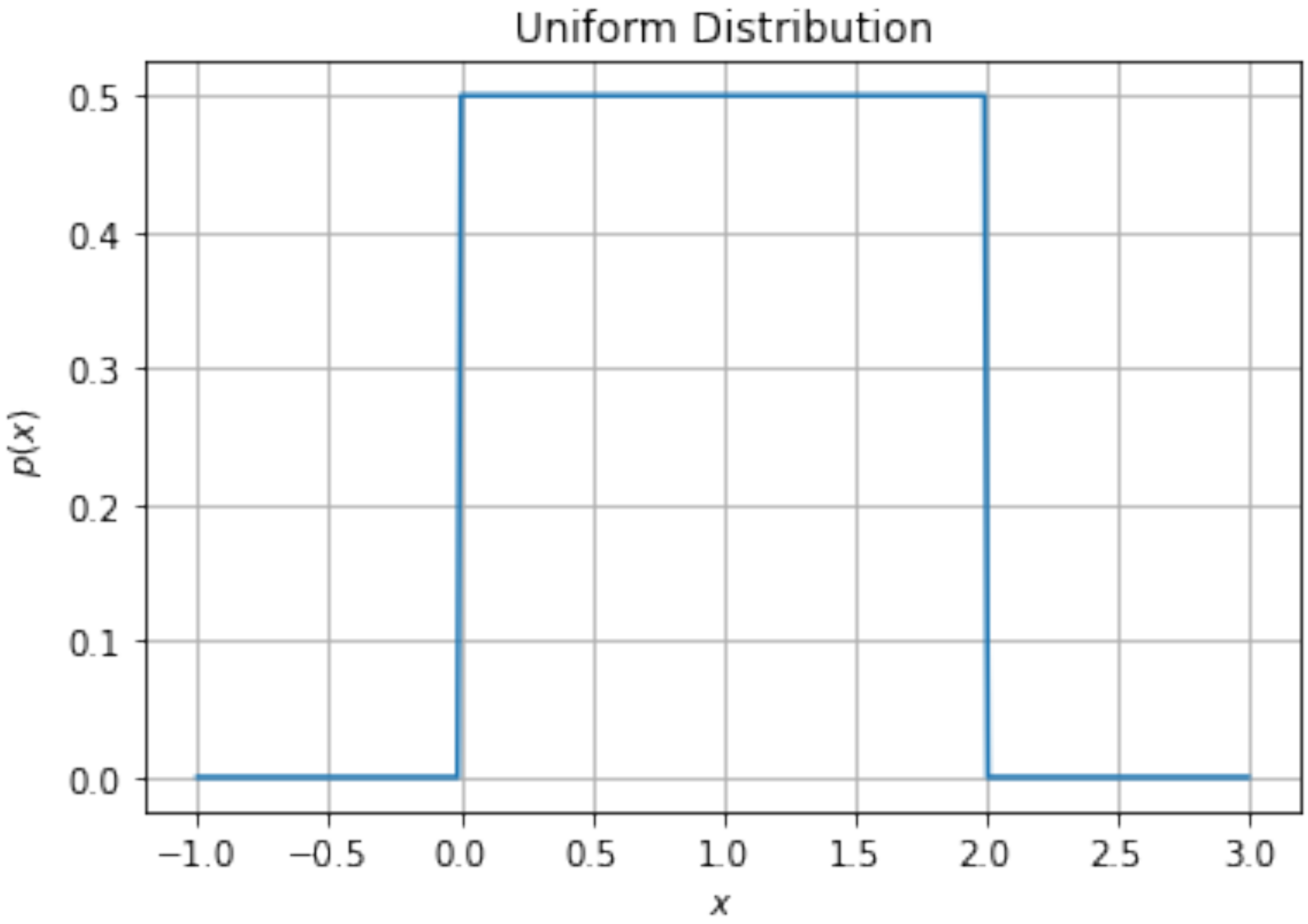
$$p(x;\mu,W) = \begin{cases} \frac{1}{W} & \text{for } \left|x-\mu\right| \leq \frac{W}{2} \\ 0 & \text{otherwise .} \end{cases}$$

The constant W is the width of the box. Some well known results are:

* $\sigma = \dfrac{W}{\sqrt{12}} \sim 0.3W$

* Skewness: $\Sigma = 0$

* Kurtosis: $K = -1.2$ (i.e. the distribution is platykurtic)

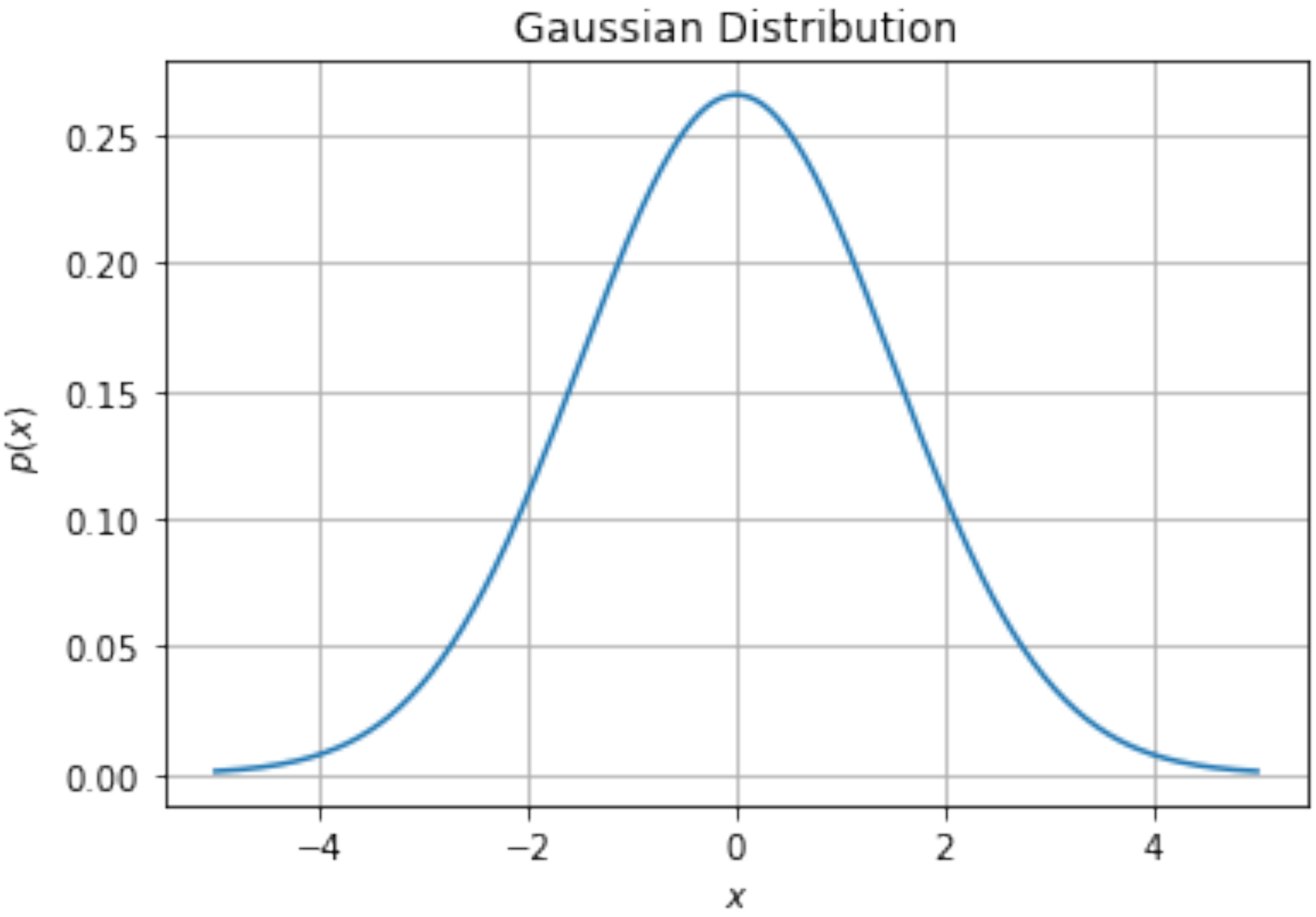

Uniform Distribution
0.5
0.4
0.3
0.2
0.1
0.0
p(x)
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
x

The Gaussian Distribution

Other important example is the **Gaussian (or Normal) Distribution**, given by

$$p(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Note that this distribution incorporates explicitly the average μ and the standard deviation σ.


Gaussian Distribution
0.25
0.20
0.15
0.10
0.05
0.00
p(x)
−4
−2
0
2
4
x

Descriptive Statistics. Discrete Variable

Some Definitions in Descriptive Statistics. Discrete Variable

Now consider an arbitrary distribution function $w(k)$ of a discrete variable. The definitions of descriptive statistics are given below, defining $w_i = w(k_i)$,

Arithmetic Mean (Expectation value)

$$\bar{k} = E(k) = \mu = \frac{\sum_{i=1}^{N} k_i w_i}{\sum_{i=1}^{N} w_i}$$

Some Definitions in Descriptive Statistics. Discrete Variable

Variance

$$\textbf{var} = \sigma^2 = \frac{\sum_{i=1}^{N} (k_i - \bar{k})^2 w_i}{\sum_{i=1}^{N} w_i}$$

Standard Deviation

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{N} (k_i - \bar{k})^2 w_i}{\sum_{i=1}^{N} w_i}}$$

*Variance is also known as the second central moment of the distribution.

Some Definitions in Descriptive Statistics. Discrete Variable

Skewness

$$\Sigma = \frac{\sum_{i=1}^{N} \left(\frac{k_i - \bar{k}}{\sigma} \right)^3 w_i}{\sum_{i=1}^{N} w_i}$$

Kurtosis

$$K = \frac{\sum_{i=1}^{N} \left(\frac{k_i - \bar{k}}{\sigma} \right)^4 w_i}{\sum_{i=1}^{N} w_i} - 3$$

* Skewness and Kurtosis are the central moments of order 3 and 4, respectively, for the distribution.

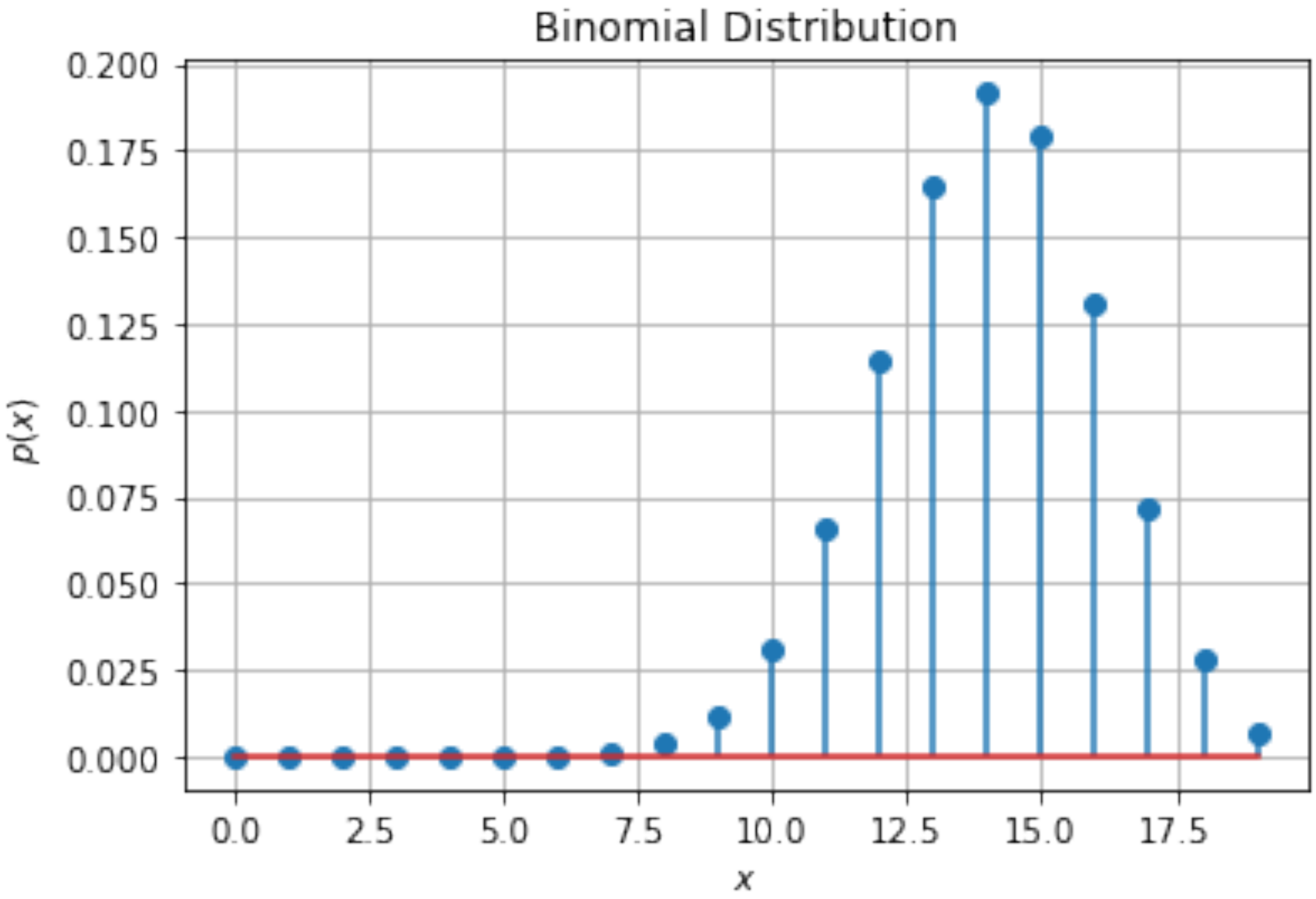
The Binomial Distribution

The Binomial distribution describes a variable that can take only two discrete values, say 0 or1. If the probability to obtain the number 1 is b, the distribution of the discrete variable k that measures how many times the result 1 occurred in N trials is given by the expression

$$p(k; b, N) = \frac{N!}{k!(N-k)!} b^k (1-b)^{N-k}$$

The particular case $N = 1$ is known as the *Bernoulli distribution*.

For the binomial distribution, the expected value of successes is $\bar{k} = bN$ and its standard deviation is $\sigma = \sqrt{Nb(1-b)}$.



Binomial Distribution
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
p(x)
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
x

@astronomiaOAN

AstronomiaOAN

@astronomiaOAN