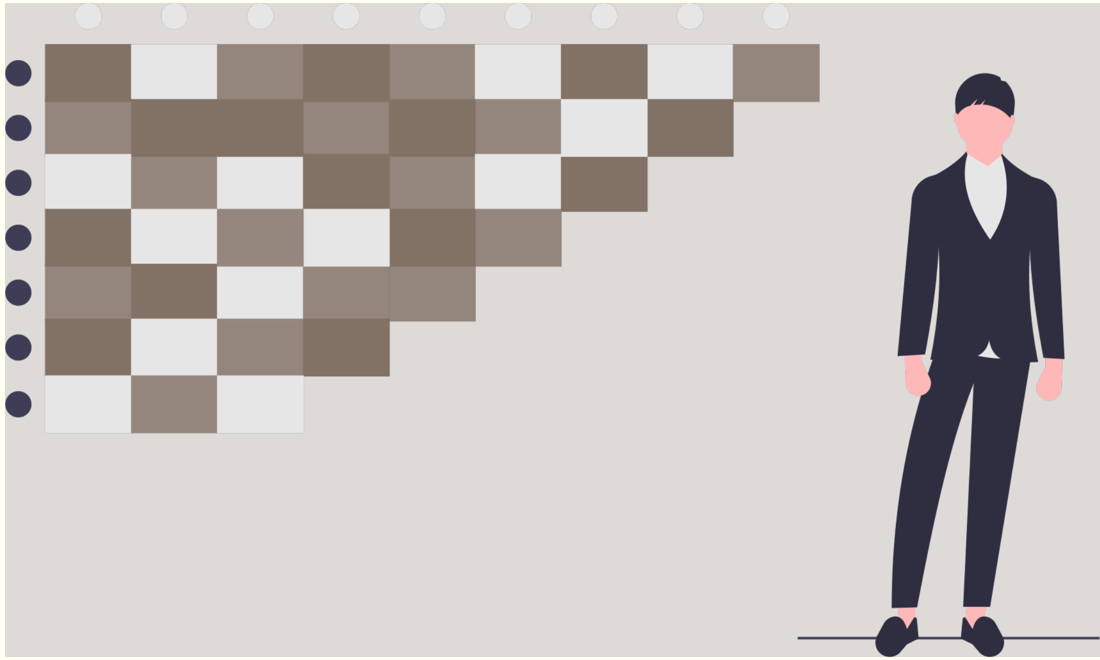
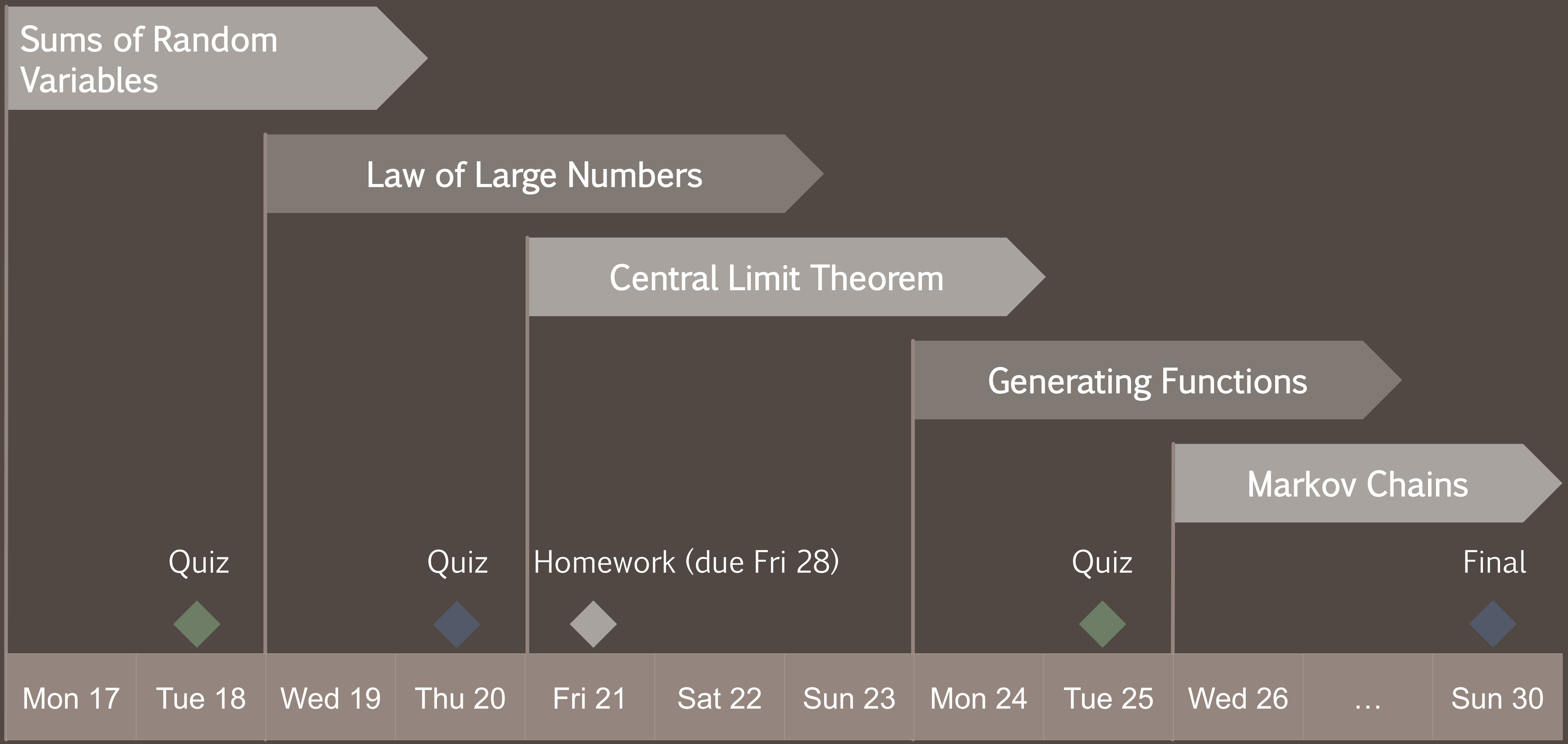
MATH 20: PROBABILITY

Sums of Independent Random Variables



Syllabus


Sums of Random Variables
Law of Large Numbers
Central Limit Theorem
Generating Functions
Markov Chains
Quiz
Quiz
Homework (due Fri 28)
Quiz
Final
Mon 17
Tue 18
Wed 19
Thu 20
Fri 21
Sat 22
Sun 23
Mon 24
Tue 25
Wed 26
…
Sun 30

Example from previous lectures

Two real numbers X and Y are chosen at random and uniformly from $[0, 1]$. Let $Z = X + Y$.

Please derive expressions for the cumulative distribution and the density function of Z.

01

$F_Z(z) = P(Z \leq z) = \cdots$

range of Z is …

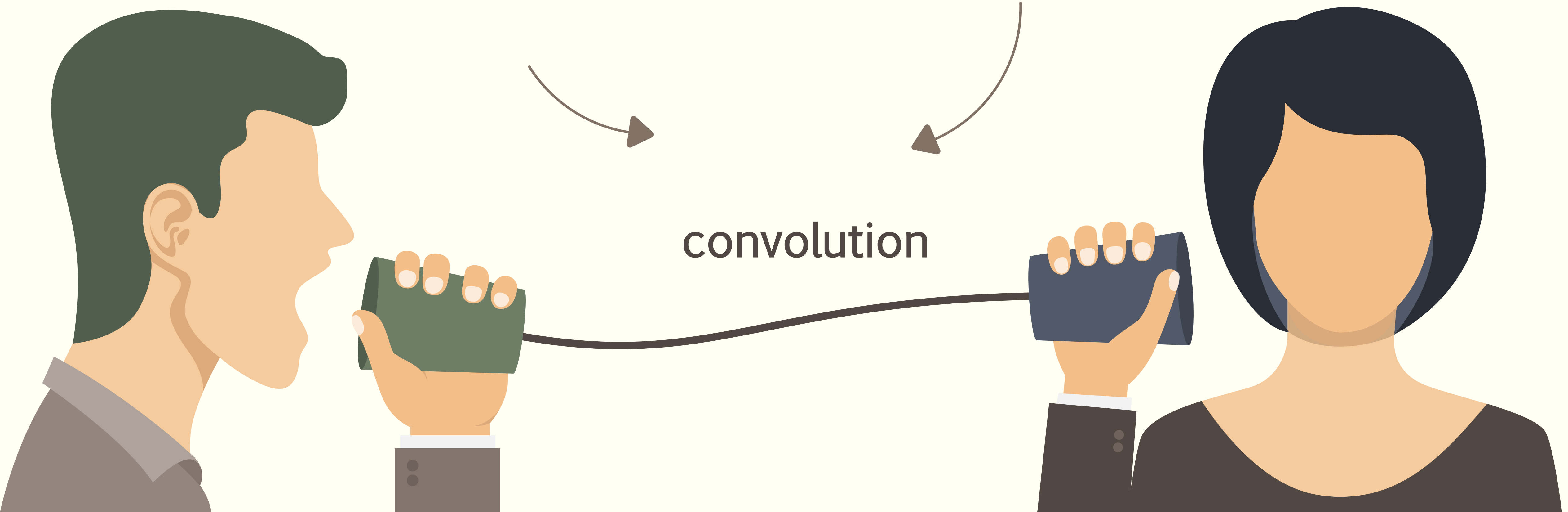
$\frac{d}{dz} F_Z(z) = f(z) = \cdots$

02

Sum of independent random variables
discrete & continuous

$$m_3(j) = \sum_k m_1(k) m_2(j-k)$$

$$(f * g)(z) = \int_{-\infty}^{+\infty} f(z-y) g(y) dy$$



Sum of discrete random variables

random variable
$Z = X + Y$
distribution function
$m_3(z)$

01

random variable X

distribution function
$m_1(x)$

independent

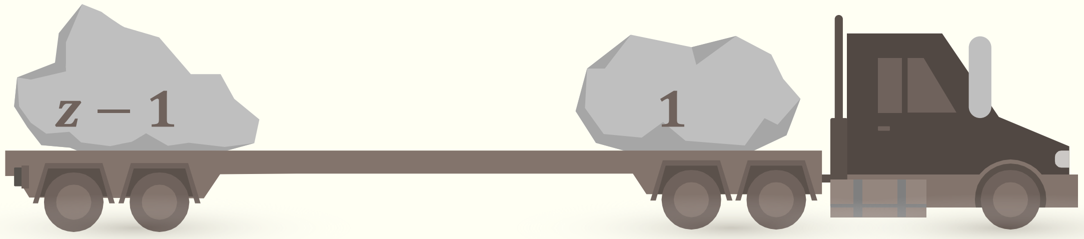
02

random variable Y

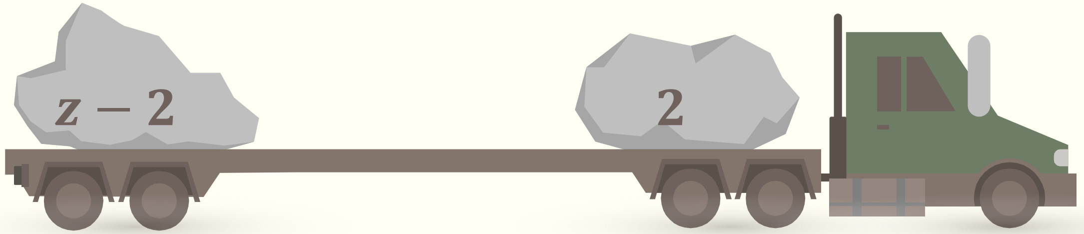
distribution function
$m_2(y)$

$$Z = X + Y$$

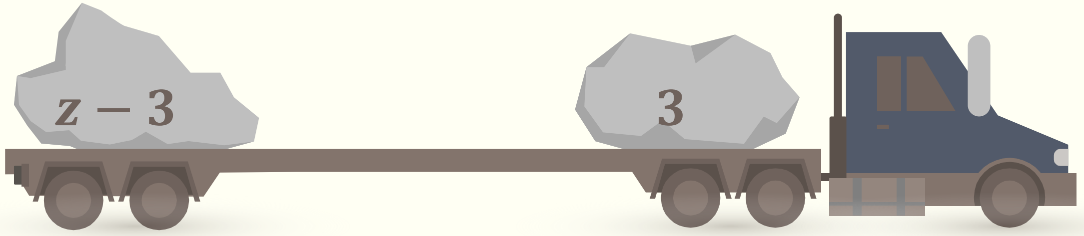
the probability that Z takes on the value z



$$z = 1 + (z - 1)$$



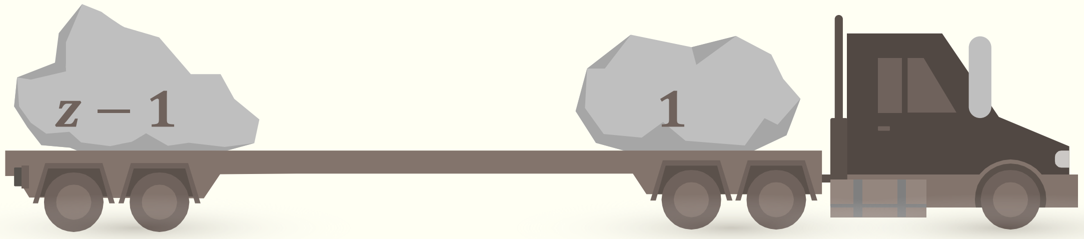
$$z = 2 + (z - 2)$$



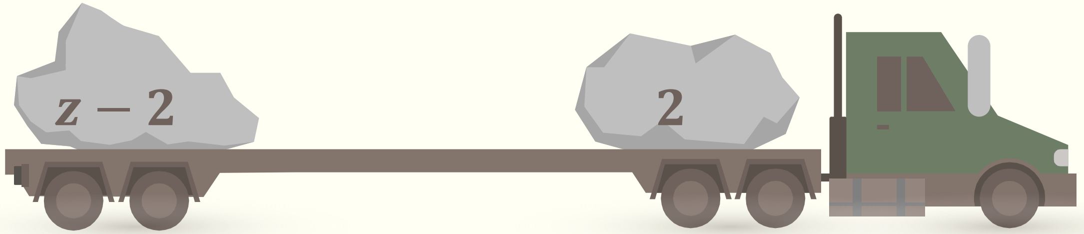
$$z = 3 + (z - 3)$$

$$Z = X + Y$$

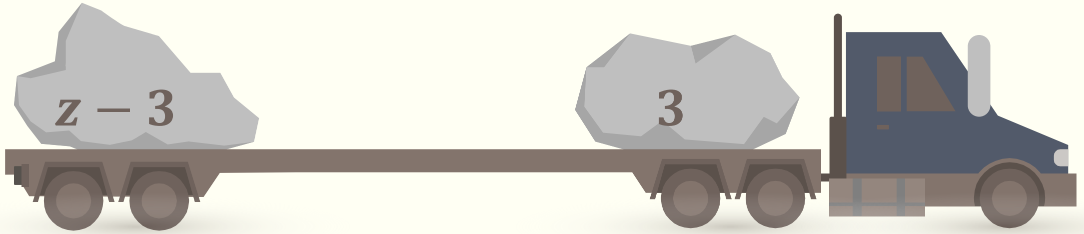
the probability that Z takes on the value z



$$z = 1 + (z - 1)$$



$$z = 2 + (z - 2)$$



$$z = 3 + (z - 3)$$

$$P(Z = z) = \sum_{k=-\infty}^{+\infty} P(X = k)P(Y = z - k)$$

Sum of discrete random variables

- Let X and Y be two **independent** integer-valued random variables, with distribution functions $m_1(x)$ and $m_2(x)$ respectively.
- Then the **convolution** of $m_1(x)$ and $m_2(x)$ is the distribution function $m_3 = m_1 * m_2$ given by

$$m_3(j) = \sum_k m_1(k)m_2(j-k),$$

  for $j = \cdots, -2, -1, 0, 1, 2, \cdots$.

- The function $m_3(x)$ is the distribution function of the random variable $Z = X + Y$.

Example

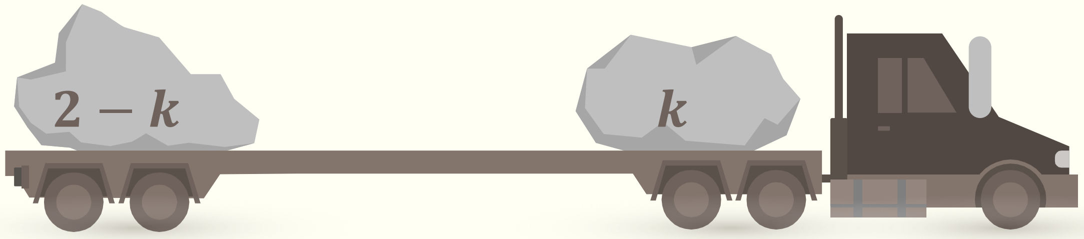
- A die is rolled twice. Let X_1 and X_2 be the outcomes, and let $S_2 = X_1 + X_2$ be the sum of these outcomes.

- Then X_1 and X_2 have the common distribution function:

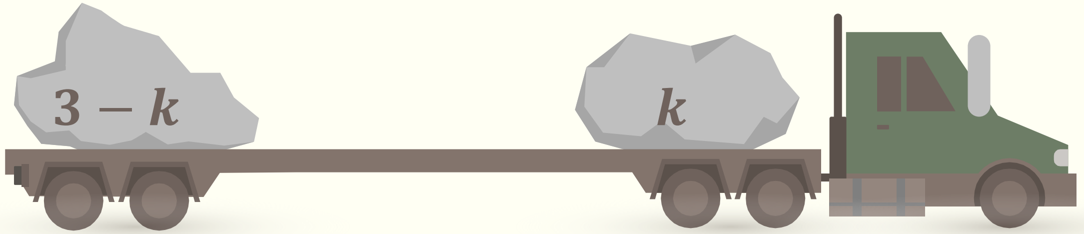
$$m = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

- The distribution function of S_2 is then the convolution of this distribution with itself.

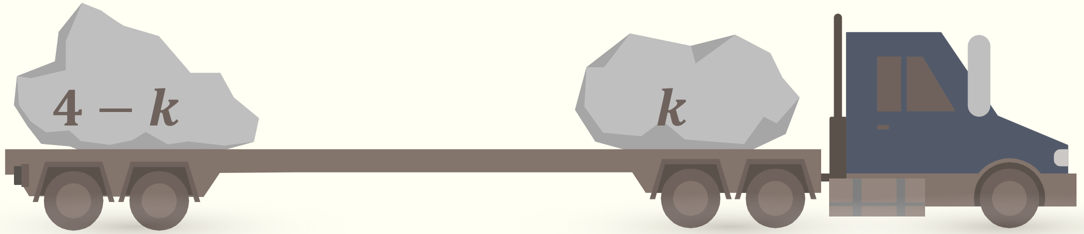
$$S_2 = X_1 + X_2$$



$$P(S_2 = 2) = m(1)m(1)$$



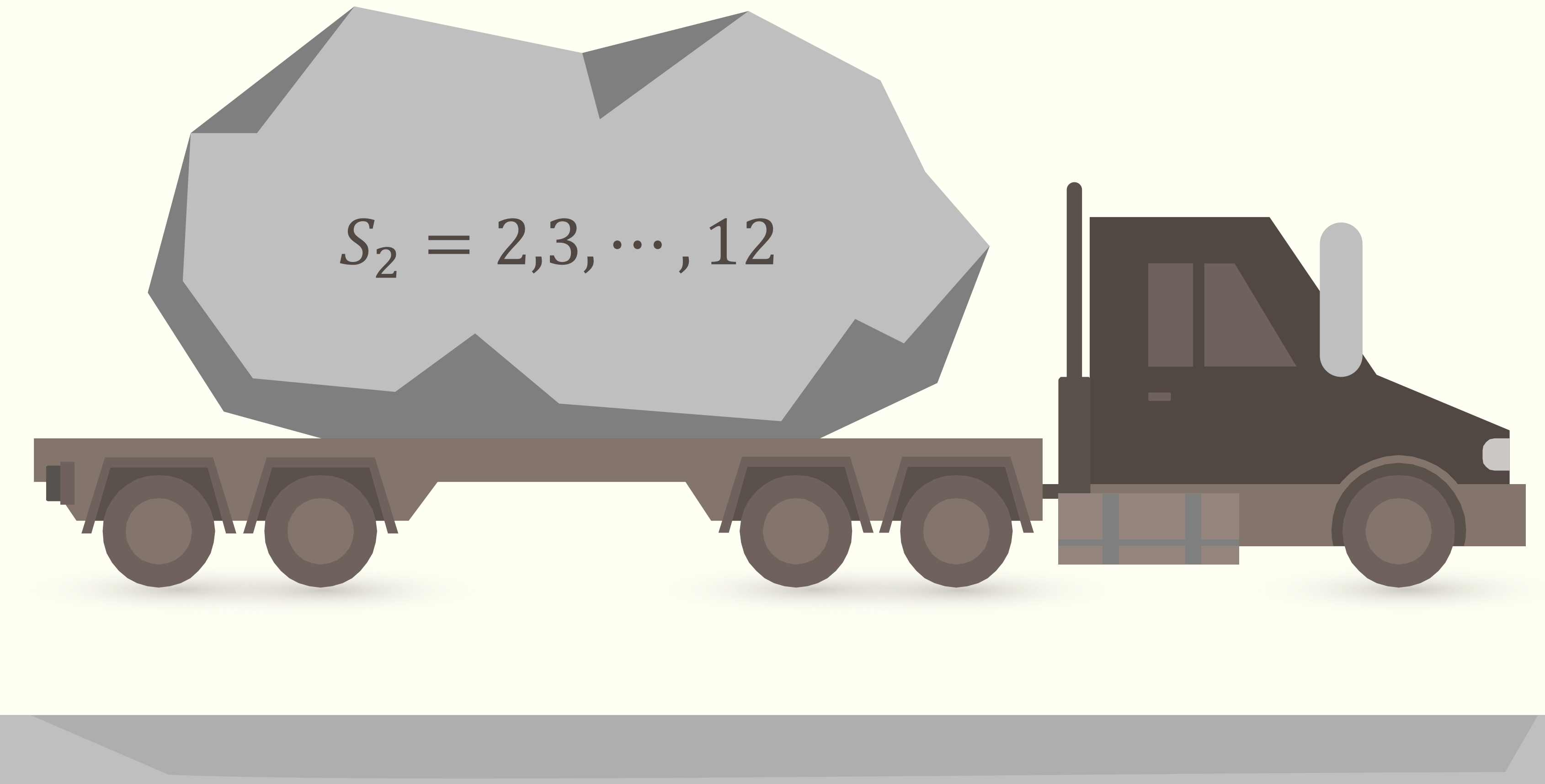
$$P(S_2 = 3) = m(1)m(2) + m(2)m(1)$$



$$P(S_2 = 4) = \cdots$$

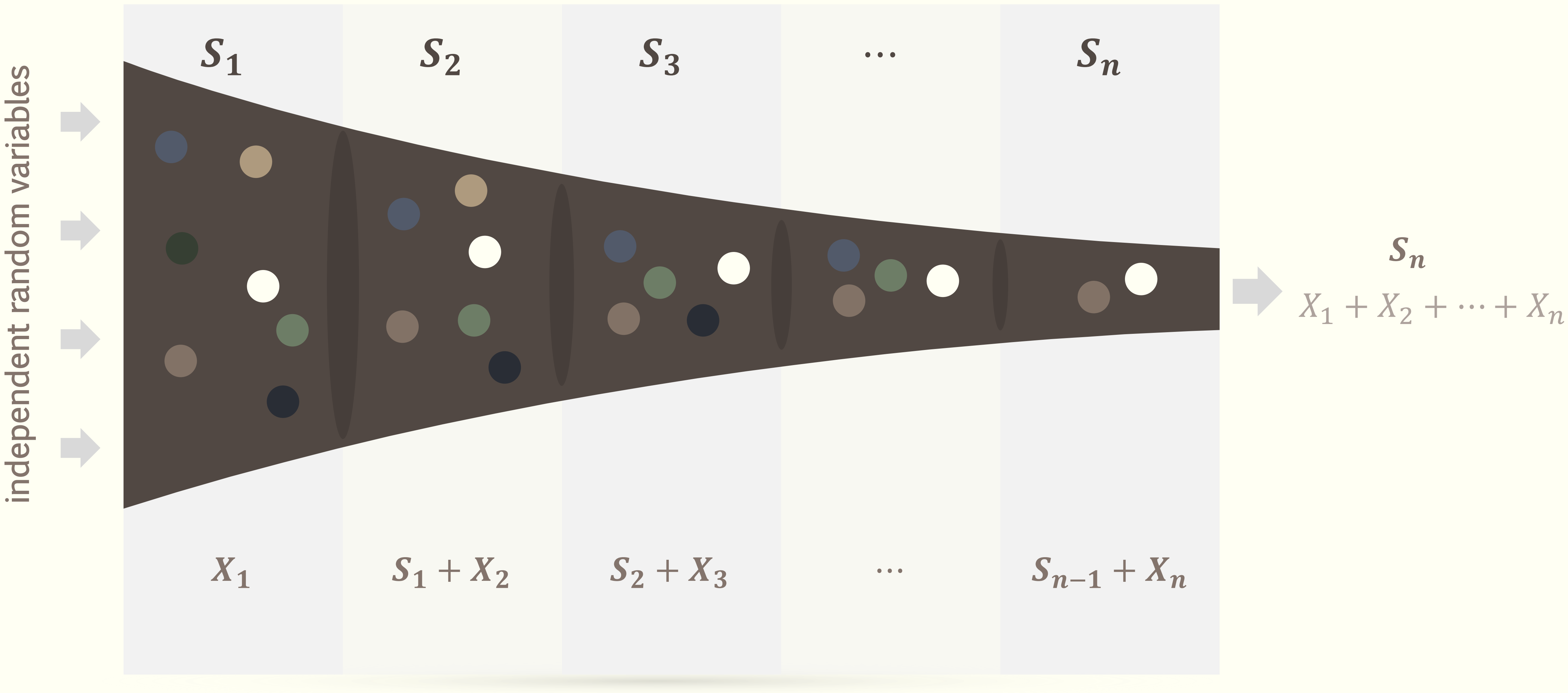
Example

$S_2 = X_1 + X_2$

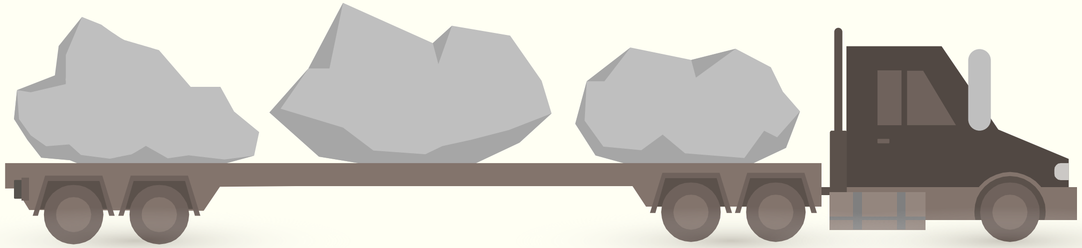


S_2	$P(S_2)$
2 and 12	$\frac{1}{36}$
3 and 11	$\frac{2}{36}$
4 and 10	$\frac{3}{36}$
5 and 9	$\frac{4}{36}$
6 and 8	$\frac{5}{36}$
7	$\frac{6}{36}$

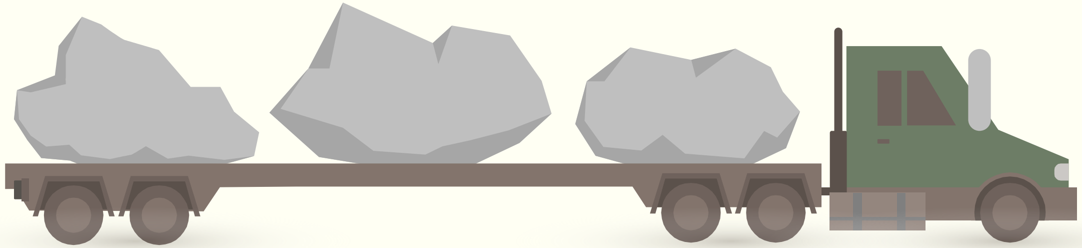
Sum of n discrete random variables



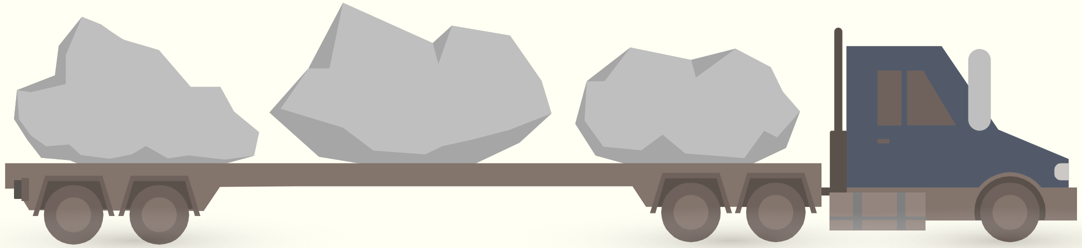
$$S_3 = X_1 + X_2 + X_3 = S_2 + X_3$$



$$P(S_3 = 3) = P(S_2 = 2)P(X_3 = 1)$$



$$P(S_3 = 4) = P(S_2 = 2)P(X_3 = 2) + P(S_2 = 3)P(X_3 = 1)$$

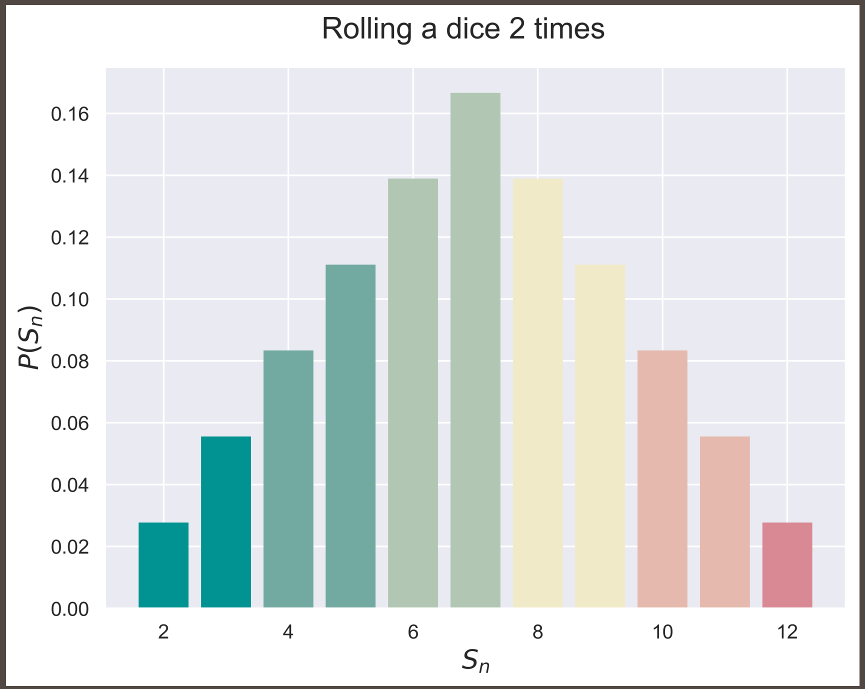


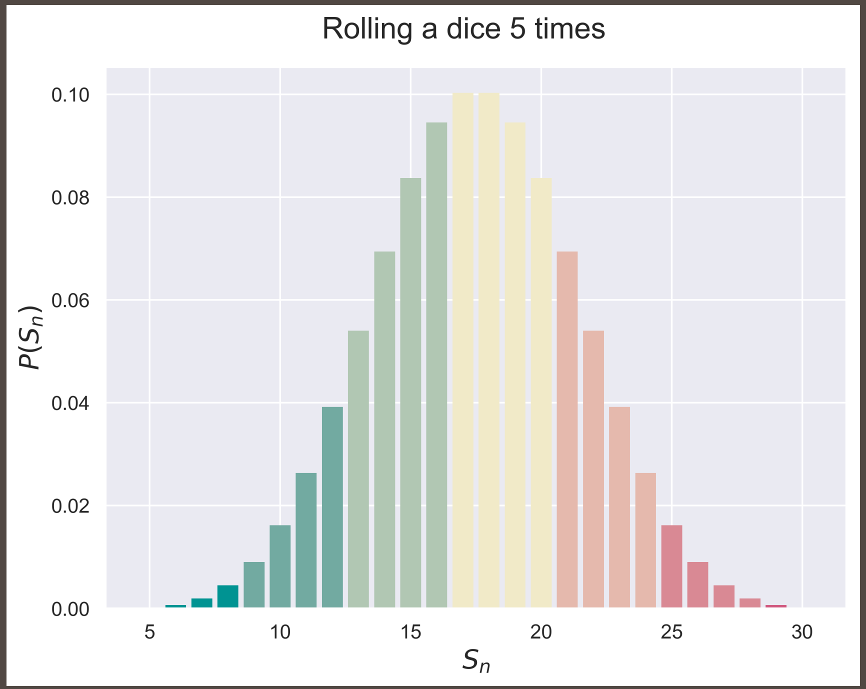
$$P(S_3 = 5) = \cdots$$

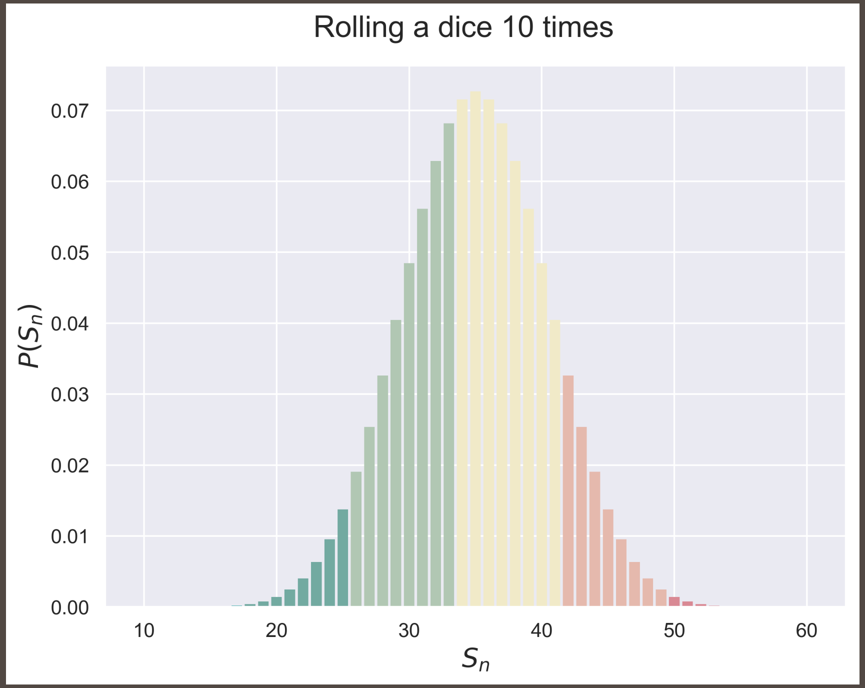
Bell-shaped curve

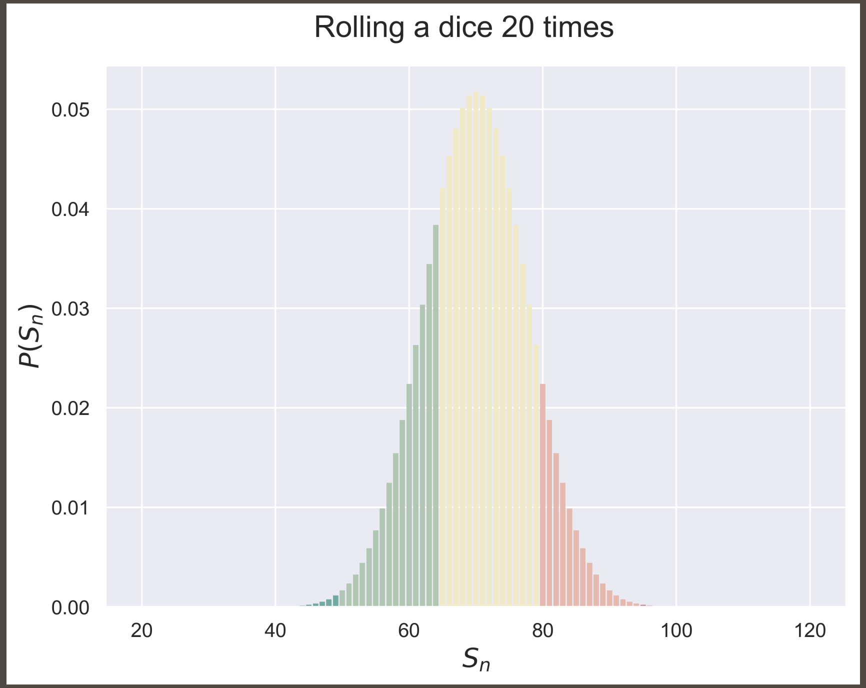
$$n \to \infty$$
$$P(S_n) \to \cdots$$

Central Limit Theorem





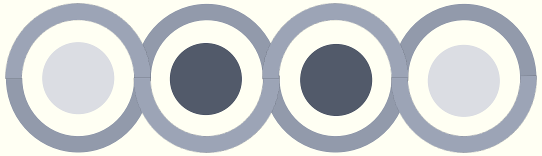




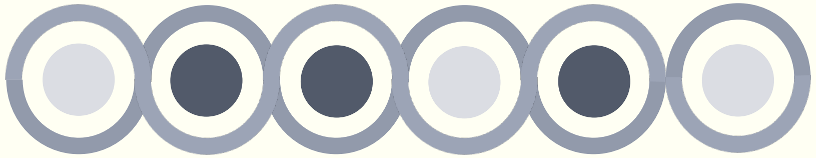
The convolution of two binomial distributions

Random variable X

binomial distribution parameters:
m and p



Random variable Z



Random variable Y

binomial distribution parameters:
n and p

The convolution of two binomial distributions

Random variable X

binomial distribution parameters: m and p

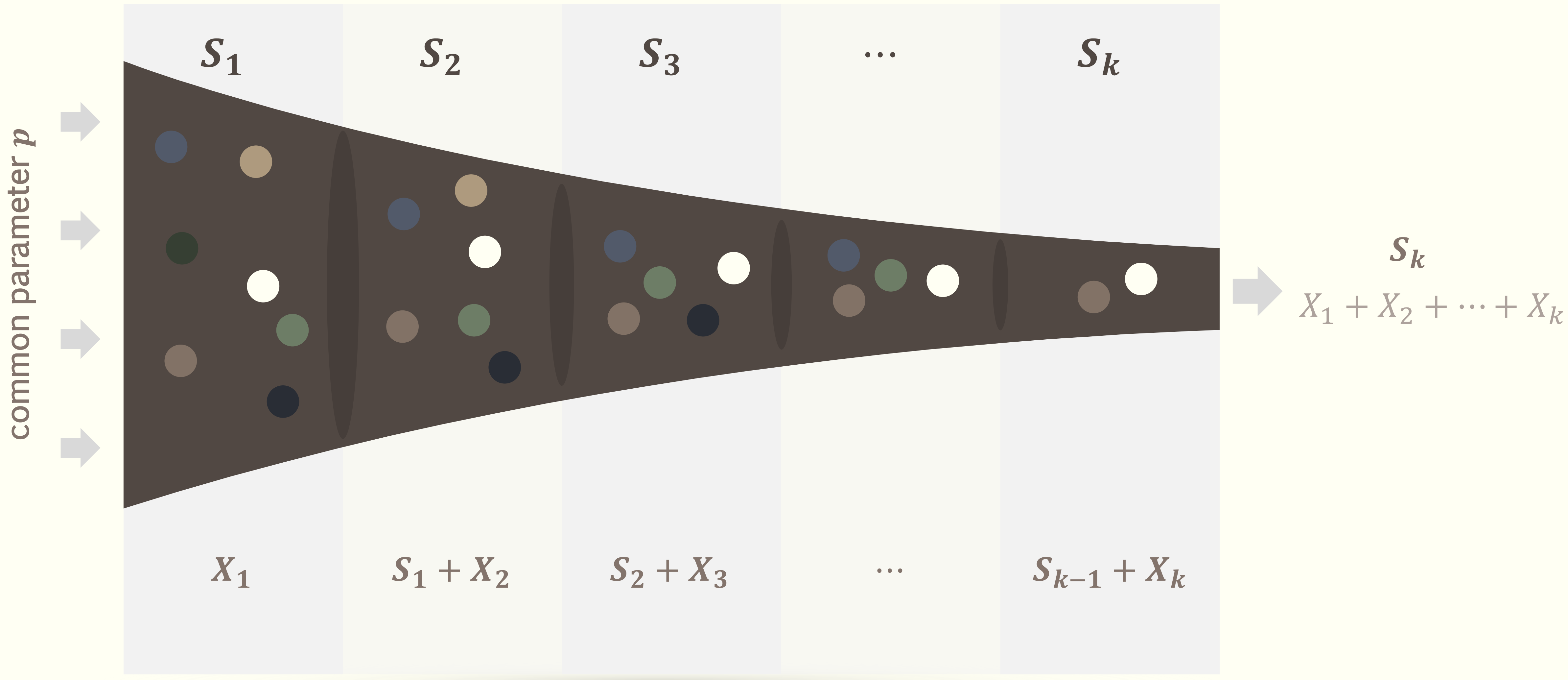
Random variable Z

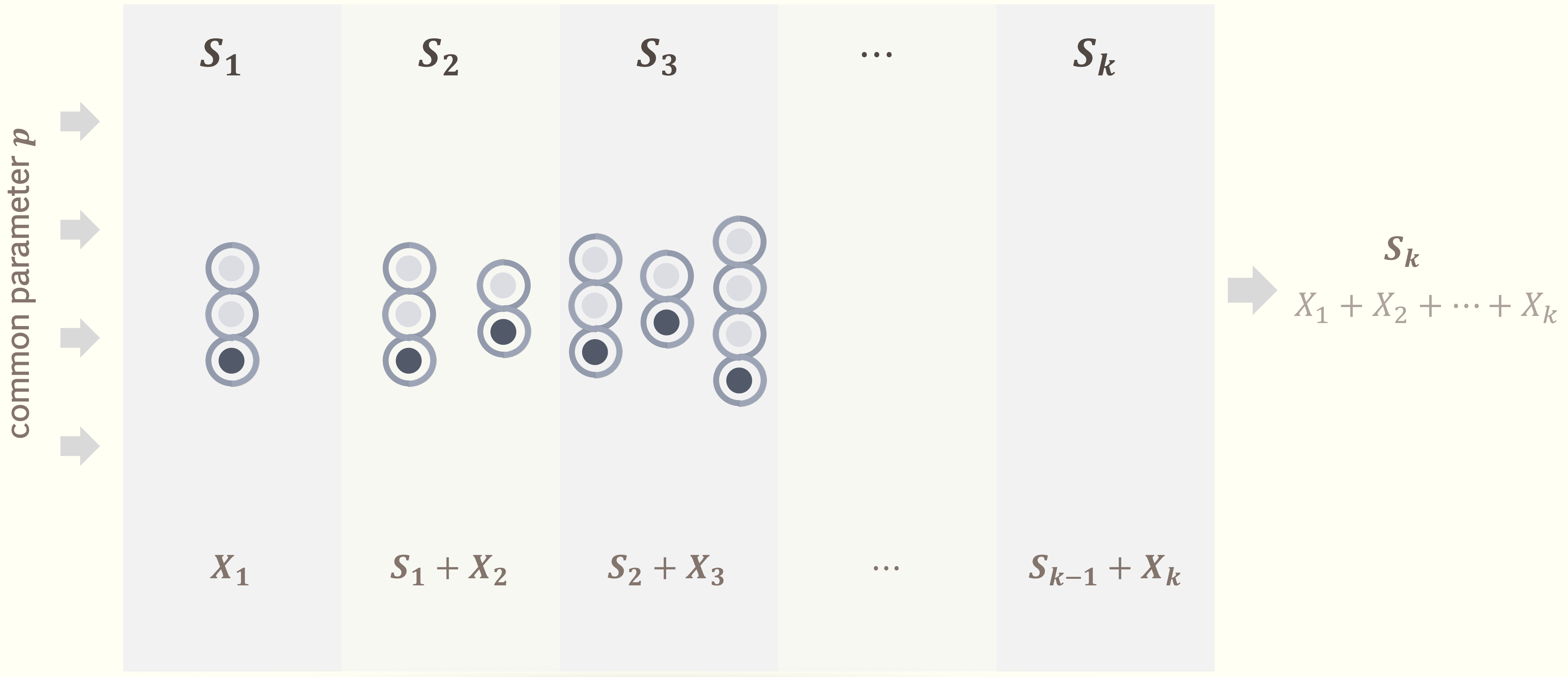
binomial distribution parameters: $m + n$ and p

Random variable Y

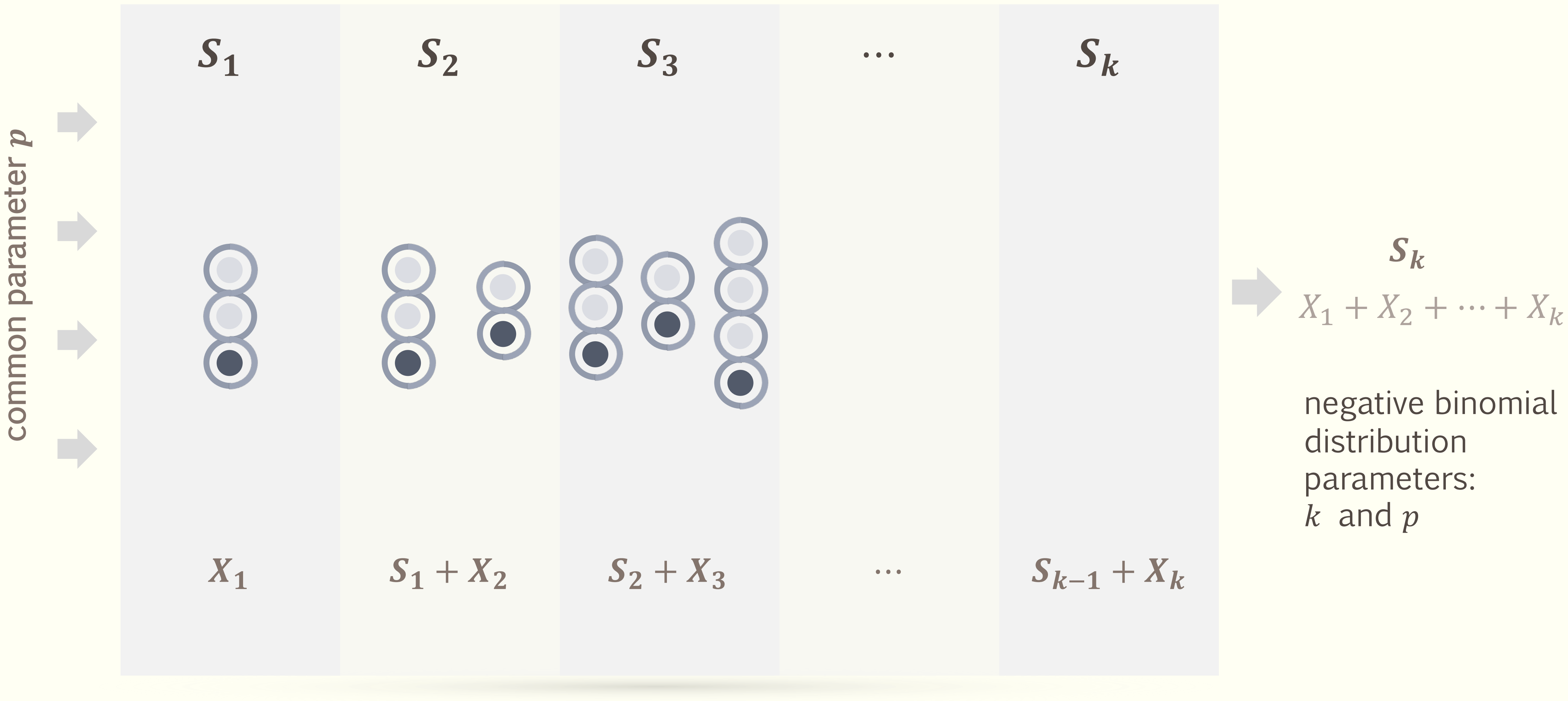
binomial distribution parameters: n and p

The convolution of k geometric distributions





X_i: the number of trials up to and including **the first success**
S_K: $\cdots$



X_i: the number of trials up to and including **the first success**
S_K: the number of trails up to and include the kth successes

Sum of continuous random variables

random variable $Z = X + Y$
density function $h(z)$

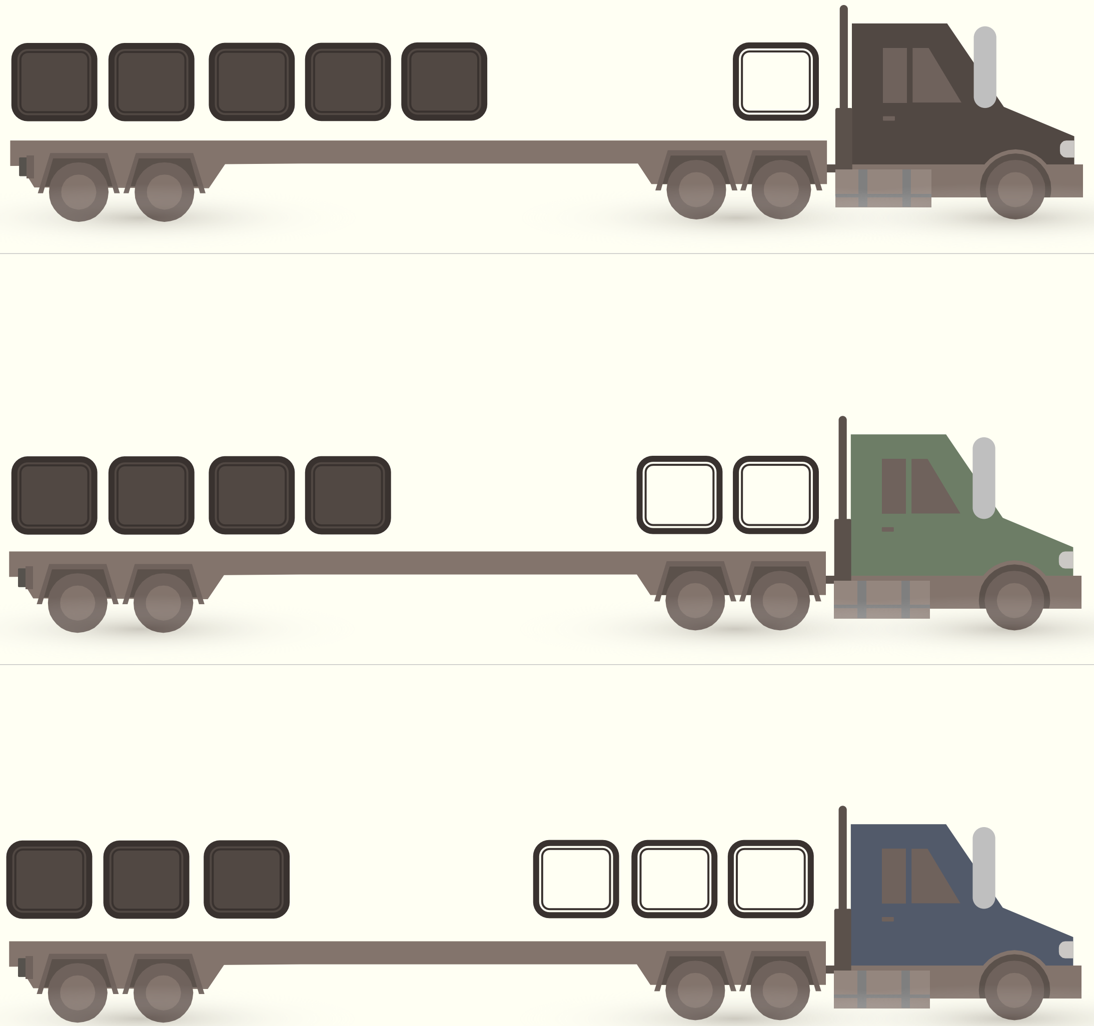
01
random variable X
density function $f(x)$

independent

02
random variable Y
density function $g(y)$

$$Z = X + Y$$

the probability that Z takes on the value z



$$z = 1 + (z - 1)$$

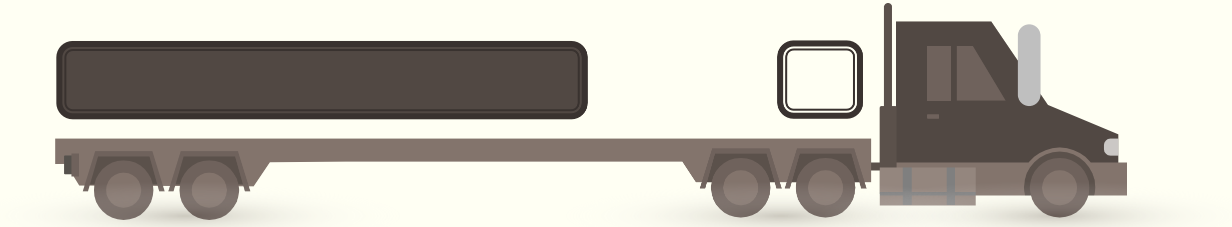
$$z = 2 + (z - 2)$$

$$z = 3 + (z - 3)$$

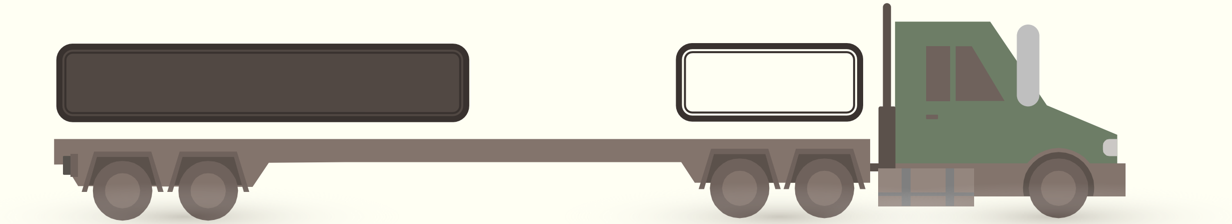
discrete

$$Z = X + Y$$

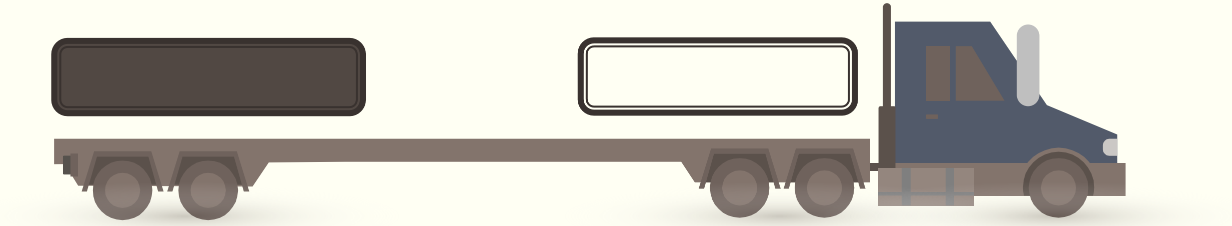
the probability that Z takes on the value z



$$z = a + (z - a)$$



$$z = b + (z - b)$$



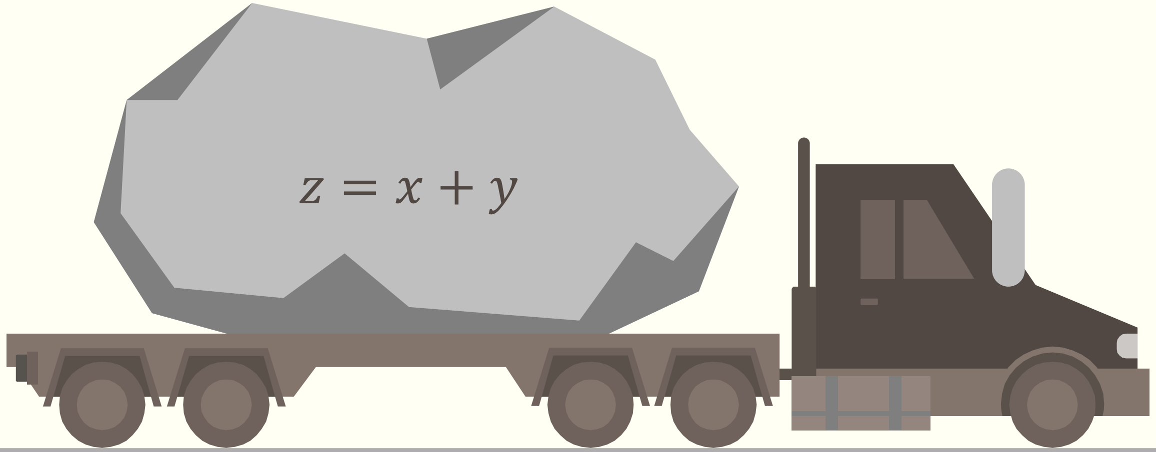
$$z = c + (z - c)$$

continuous

Convolution

- Let X and Y be two continuous random variables with density functions $f(x)$ and $g(y)$, respectively.
- Assume that both $f(x)$ and $g(y)$ are defined for all real numbers.
- Then the **convolution** $f * g$ of f and g is the function given by

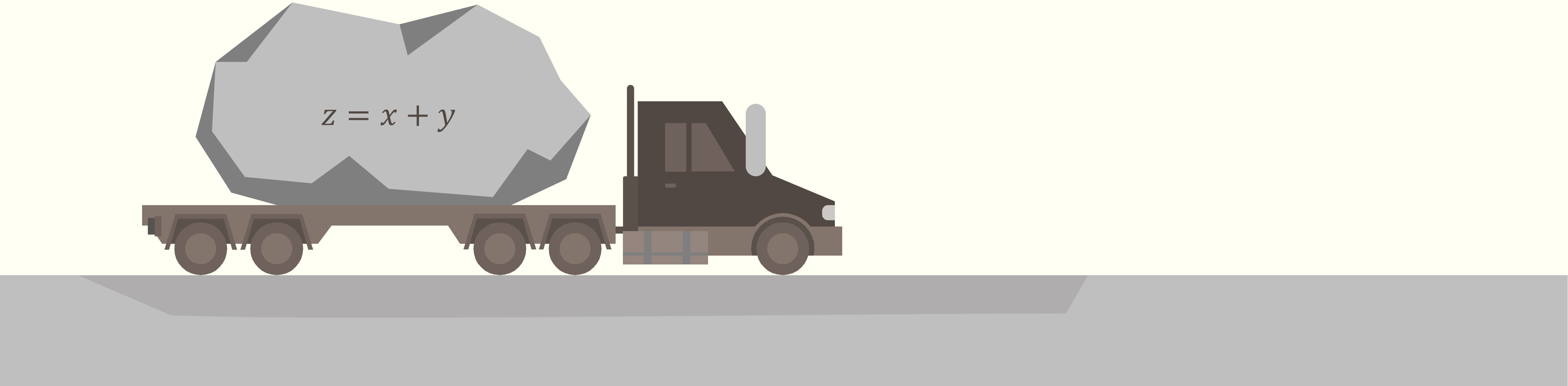
$$(f * g)(z) = \int_{-\infty}^{+\infty} f(z - y)g(y)dy.$$



Sum of continuous random variables

- Let X and Y be two **independent** random variables with density functions $f_X(x)$ and $f_Y(y)$ defined for all x.
- The the sum $Z = X + Y$ is a random variable with density function $f_Z(z)$, where f_Z is the convolution of f_X and f_Y.

$$f_Z(z) = (f_X * f_Y)(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy.$$



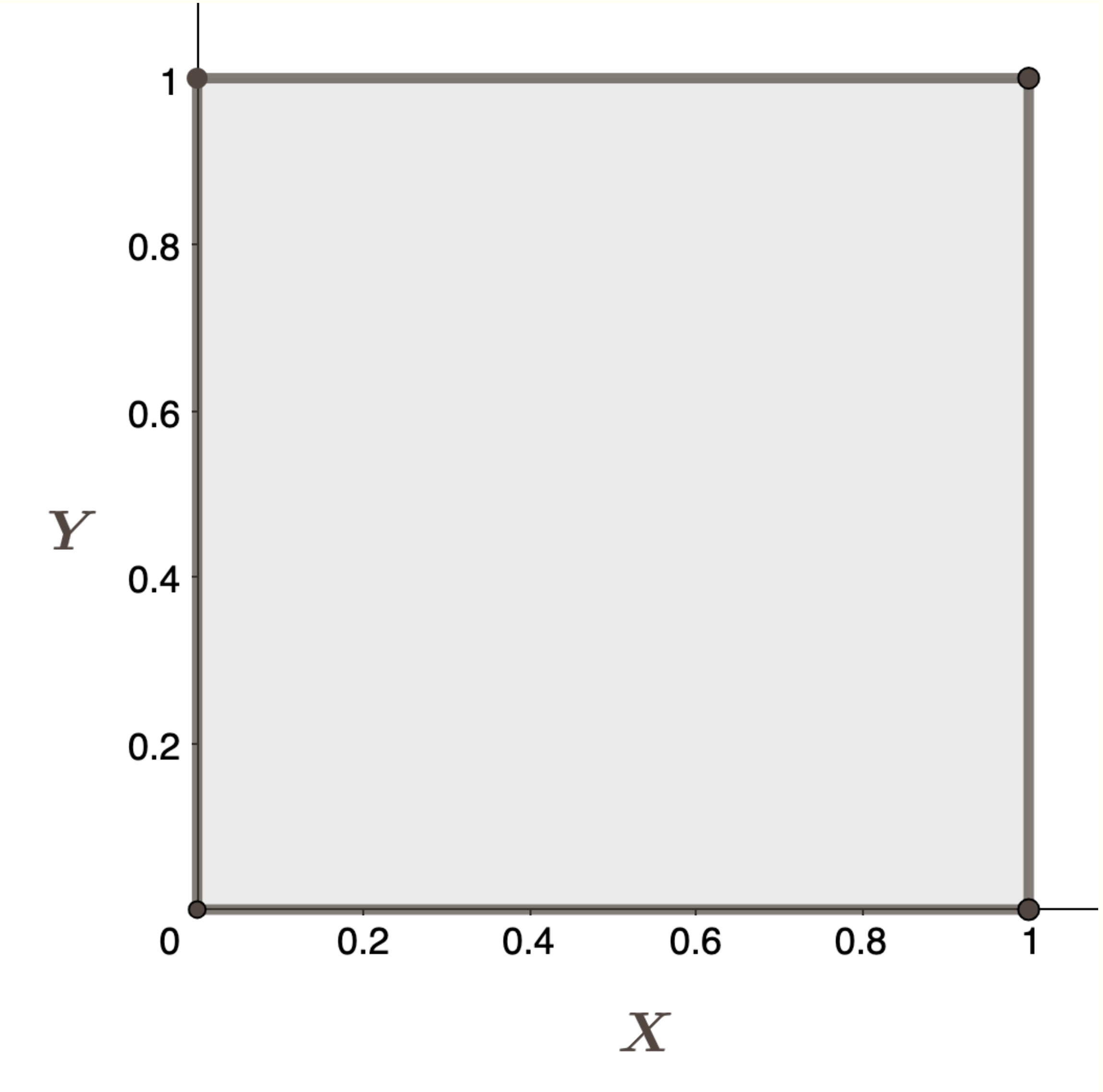
Example 1: uniform

- Suppose we choose independently two numbers at random from the interval $[0,1]$ with uniform probability density.
- What is the density of their sum?

Uniform distribution

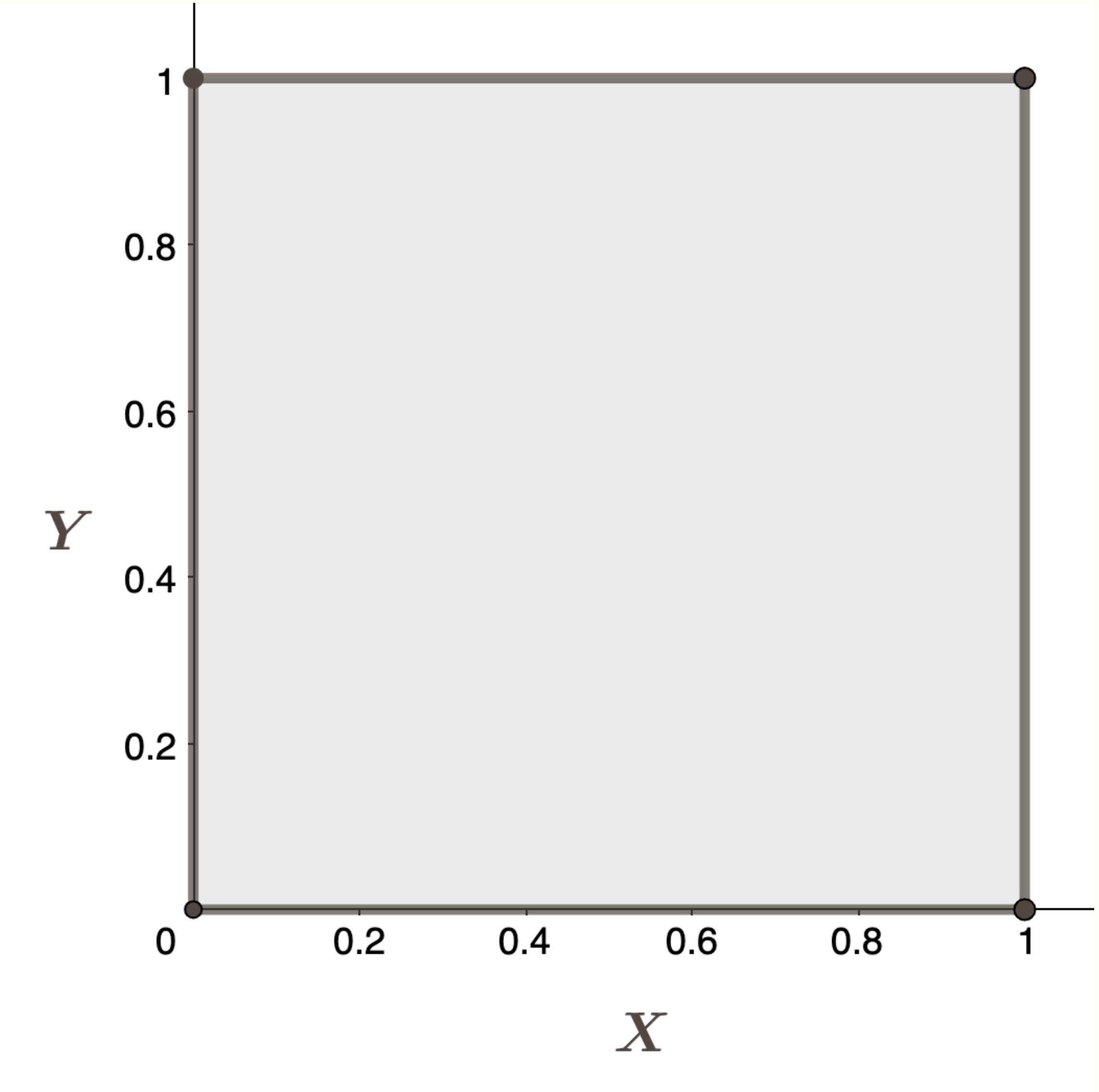
$$f_X(x) = f_Y(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0. & \text{otherwise} \end{cases}$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$$



Example 1: uniform

Uniform distribution
$$f_X(x) = f_Y(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0. & \text{otherwise} \end{cases}$$
$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$$
$$f_Z(z) = \int_0^1 f_X(z-y) dy$$



Example 1: uniform

Uniform distribution
$f_X(x) = f_Y(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0. & \text{otherwise} \end{cases}$
$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$
$f_Z(z) = \int_0^1 f_X(z-y) dy$
$0 \le z-y \le 1,\ z-1 \le y \le z$

$\mathbf{0 \le z \le 1}$

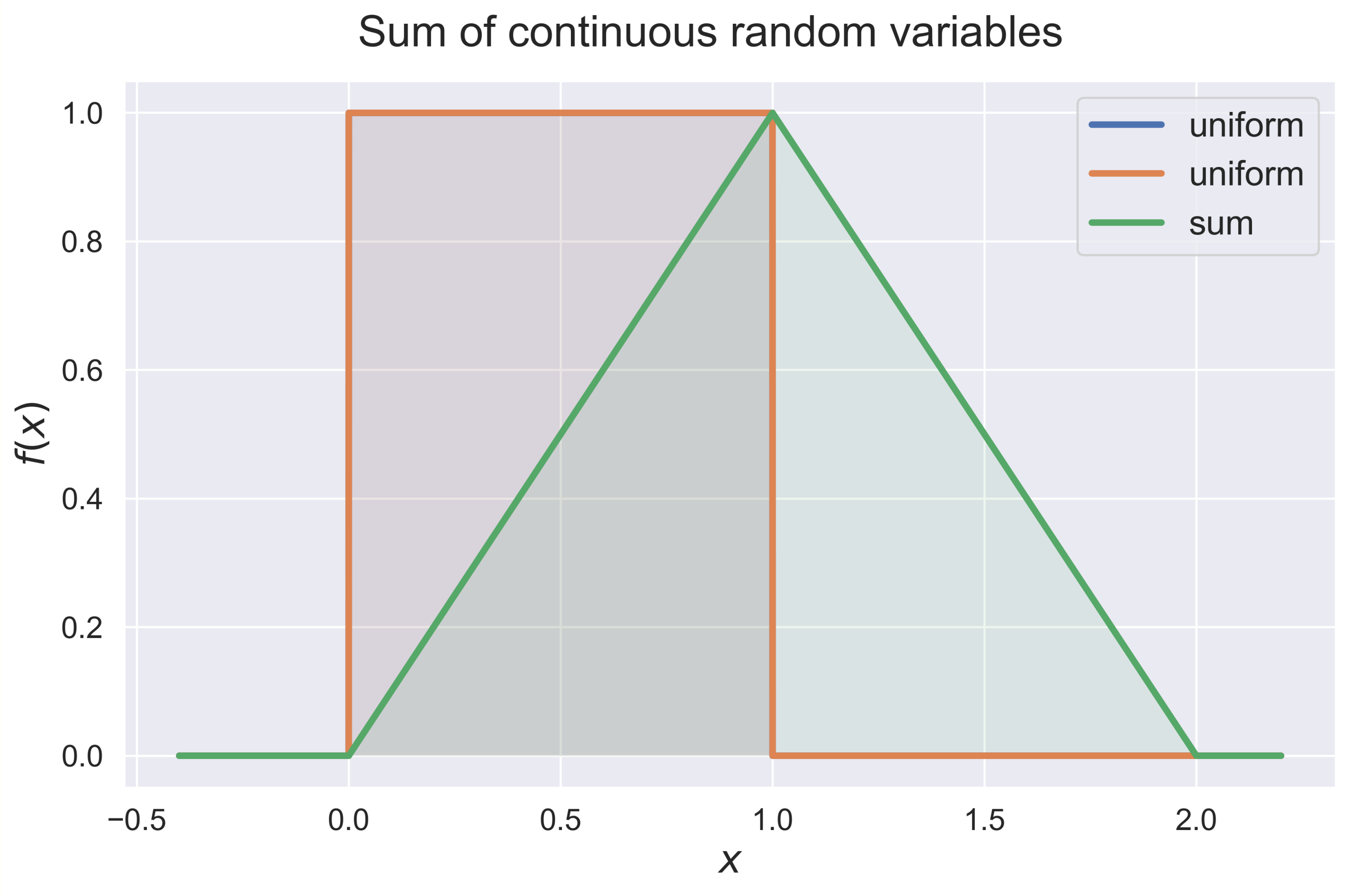
$$f_Z(z) = \int_0^z dy = z$$

$\mathbf{1 \le z \le 2}$

$$f_Z(z) = \int_{z-1}^1 dy = 2 - z$$

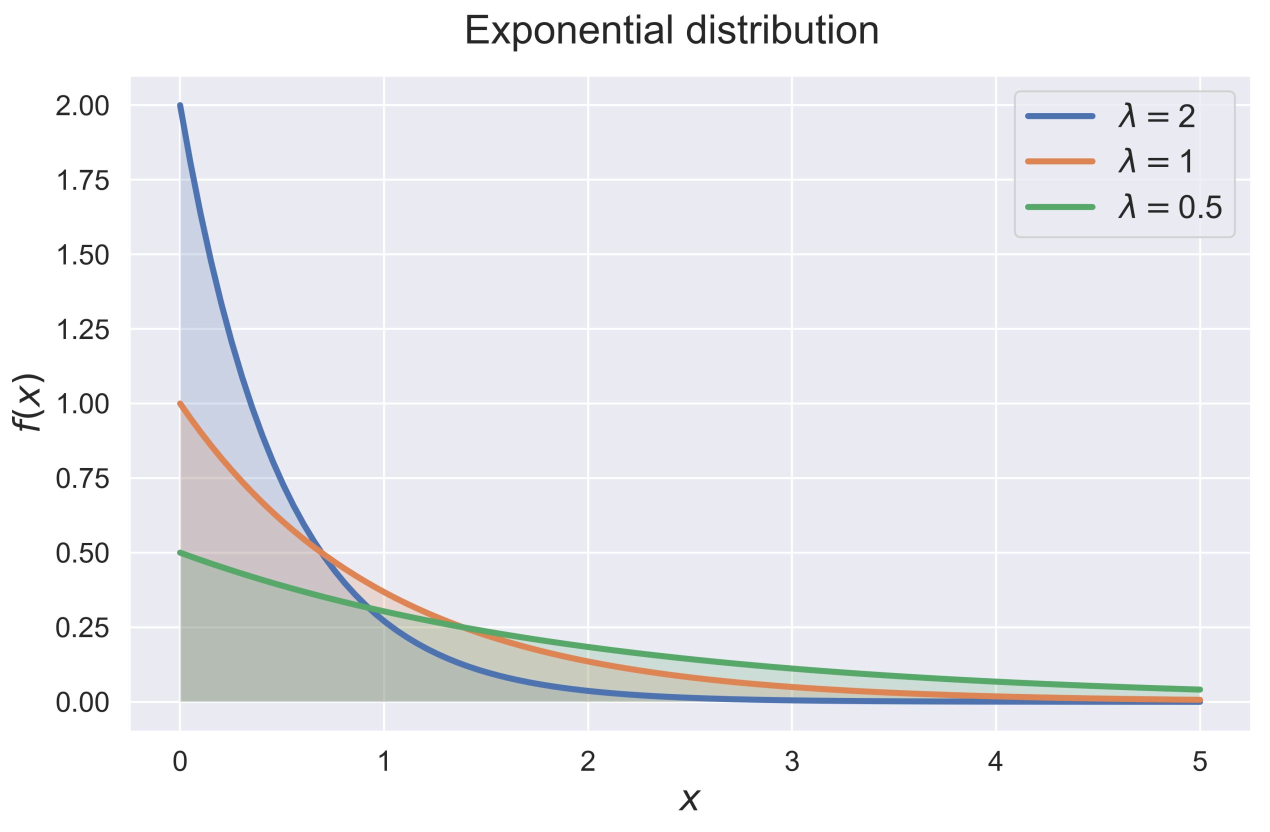
Example 1: uniform

Uniform distribution
$f_X(x) = f_Y(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0. & \text{otherwise} \end{cases}$
$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$
$f_Z(z) = \int_0^1 f_X(z-y) dy$
$f_Z(z) = \begin{cases} z, & 0 \leq z \leq 1 \\ 2-z, & 1 \leq z \leq 2 \\ 0. & \text{otherwise} \end{cases}$



Example 2: exponential

- Suppose we choose two numbers at random from the interval $[0, \infty)$ with an exponential density with parameter λ.
- What is the density of their sum?



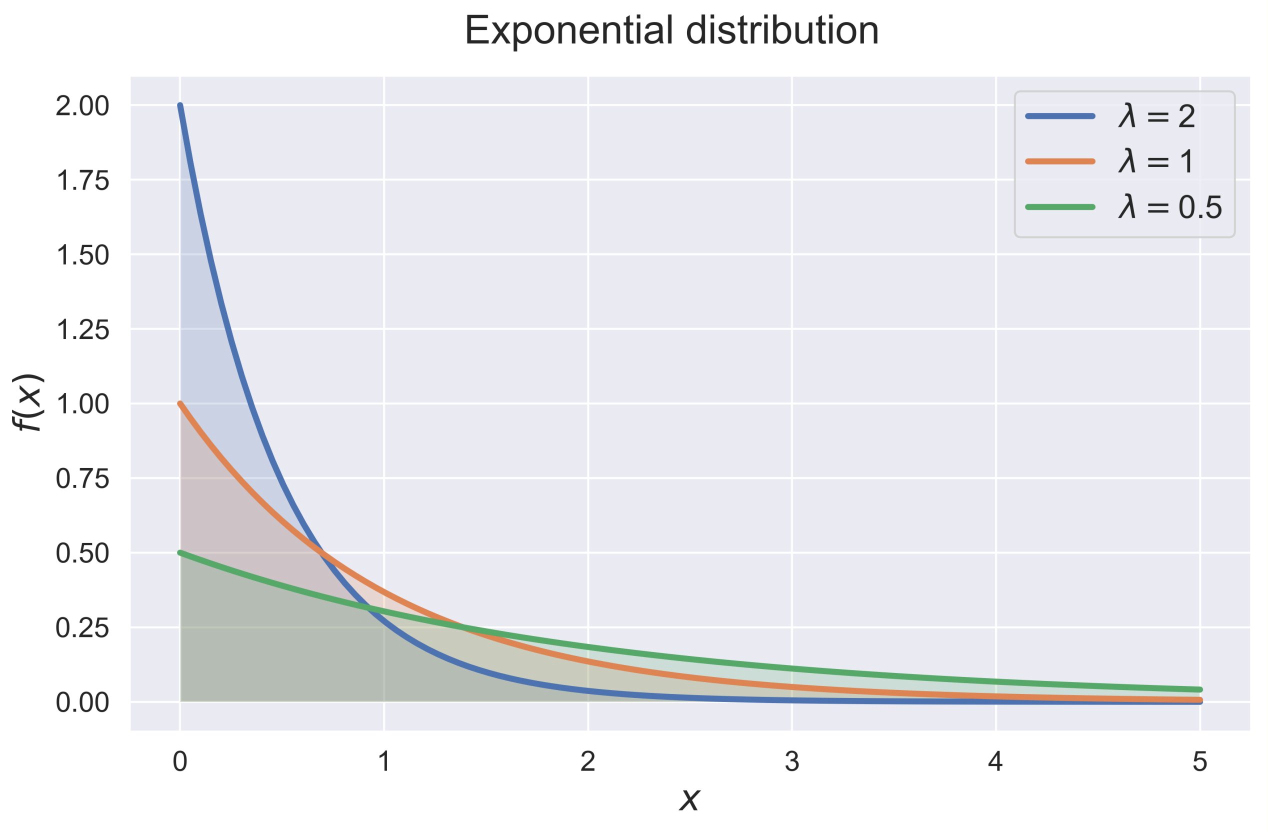
Exponential distribution

$$f_X(x) = f_Y(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0. & \text{otherwise} \end{cases}$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z - y) f_Y(y) dy$$

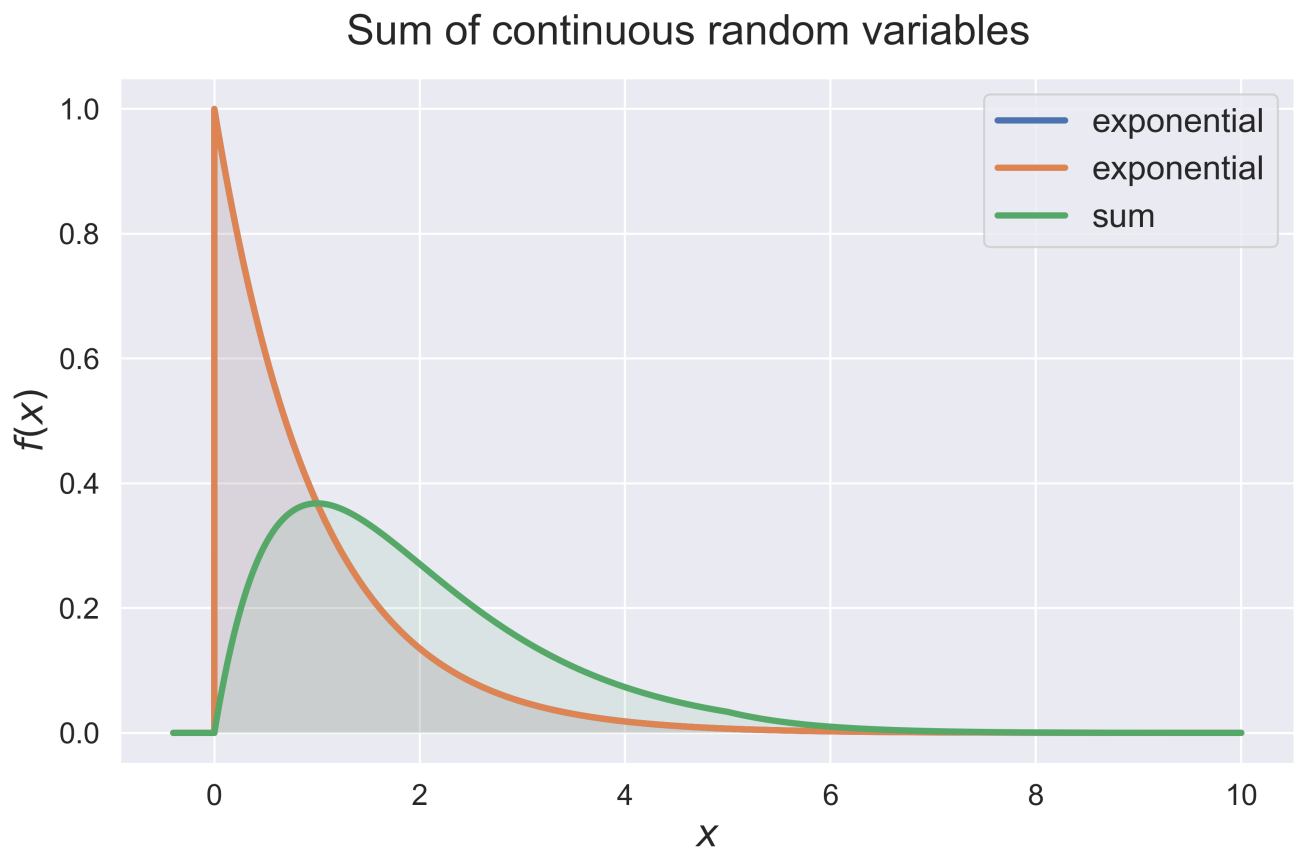
Example 2: exponential

Exponential distribution
$f_X(x) = f_Y(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0. & \text{otherwise} \end{cases}$
$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$
$f_Z(z) = \int_{0}^{z} f_X(z-y) f_Y(y) dy$



Example 2: exponential

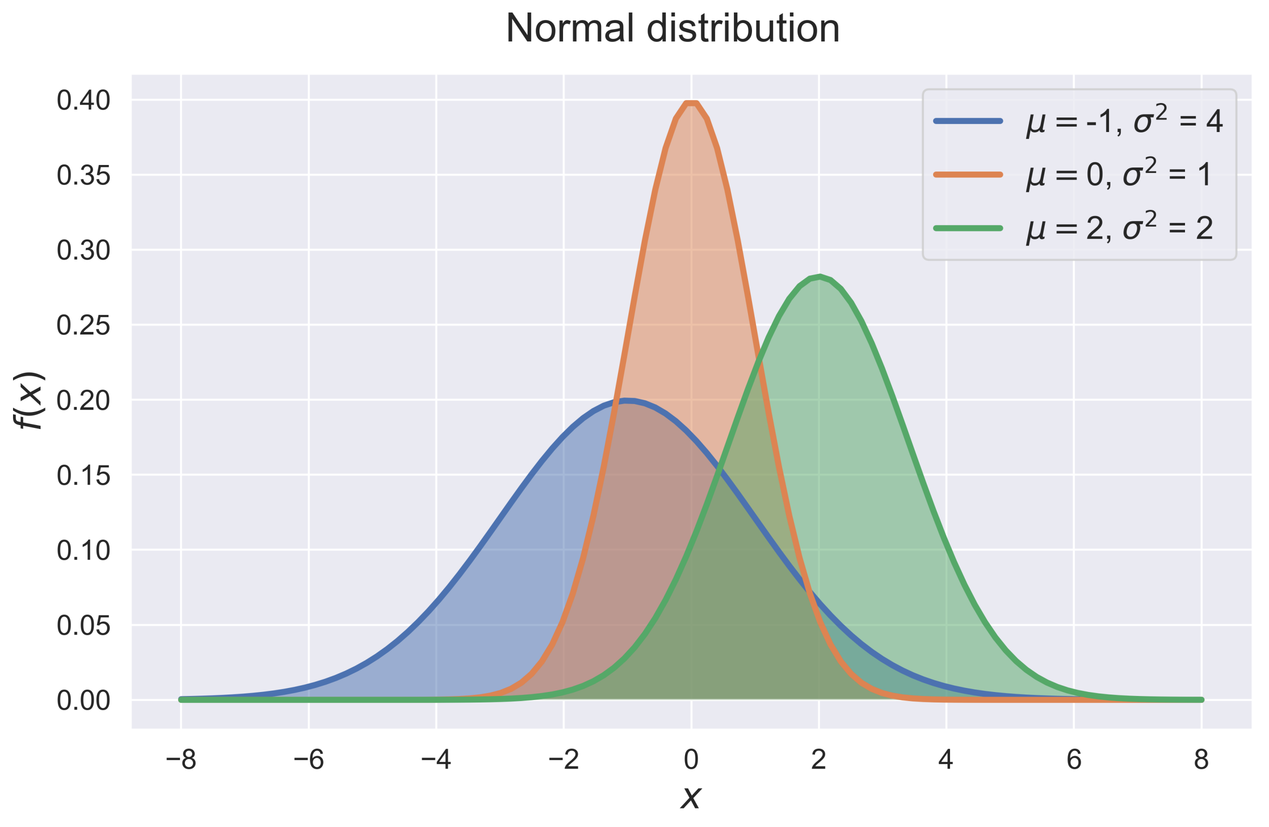
Exponential distribution
$f_X(x) = f_Y(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0. & \text{otherwise} \end{cases}$
$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$
$f_Z(z) = \int_0^z \lambda e^{-\lambda(z-y)} \lambda e^{-\lambda y} dy = \int_0^z \lambda^2 e^{-\lambda z} dy = \lambda^2 z e^{-\lambda z}$
$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z}, & z \geq 0 \\ 0. & \text{otherwise} \end{cases}$



Example 3: normal

- Suppose X and Y are two independent random variables, each with the standard normal density.
- What is the density of their sum?

Normal distribution
$$f_X(x) = f_Y(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$
$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$$

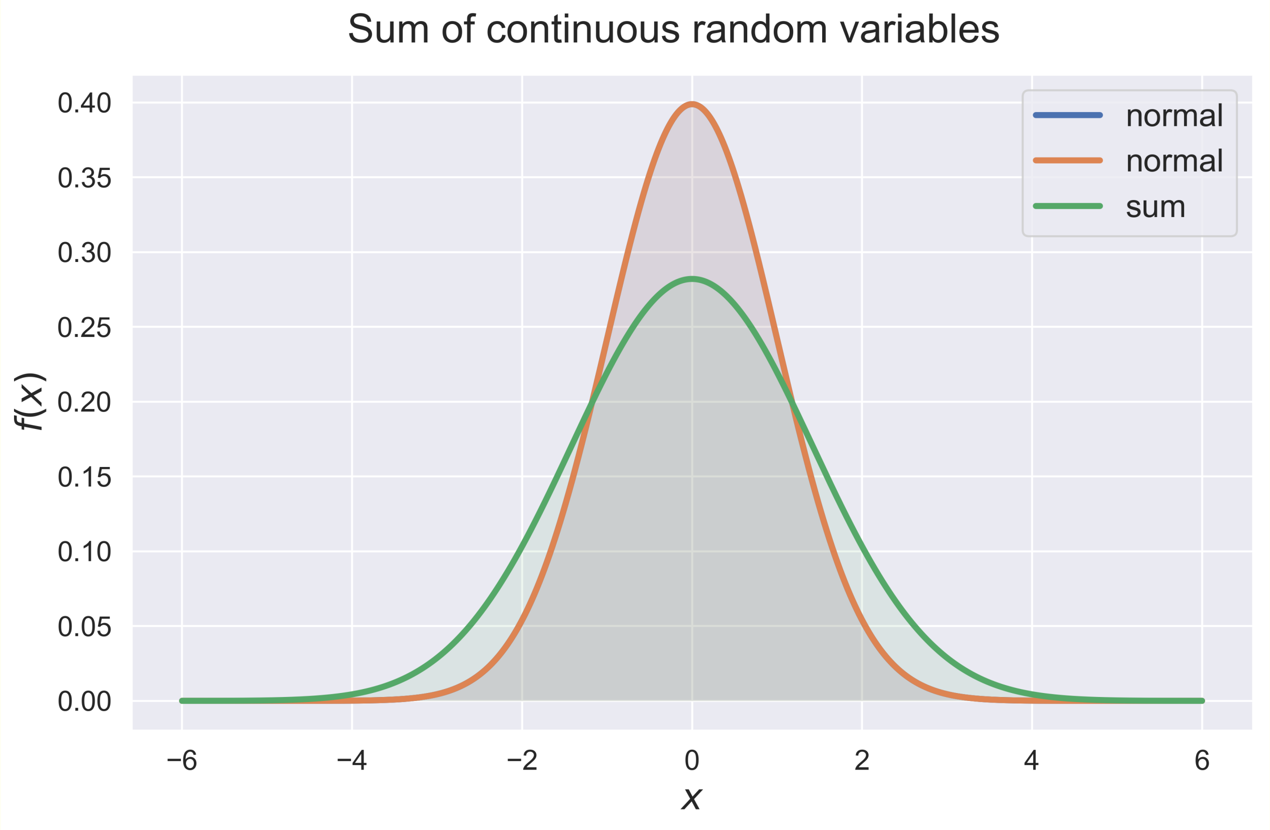


Example 3: normal

Normal distribution

$$f_X(x) = f_Y(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

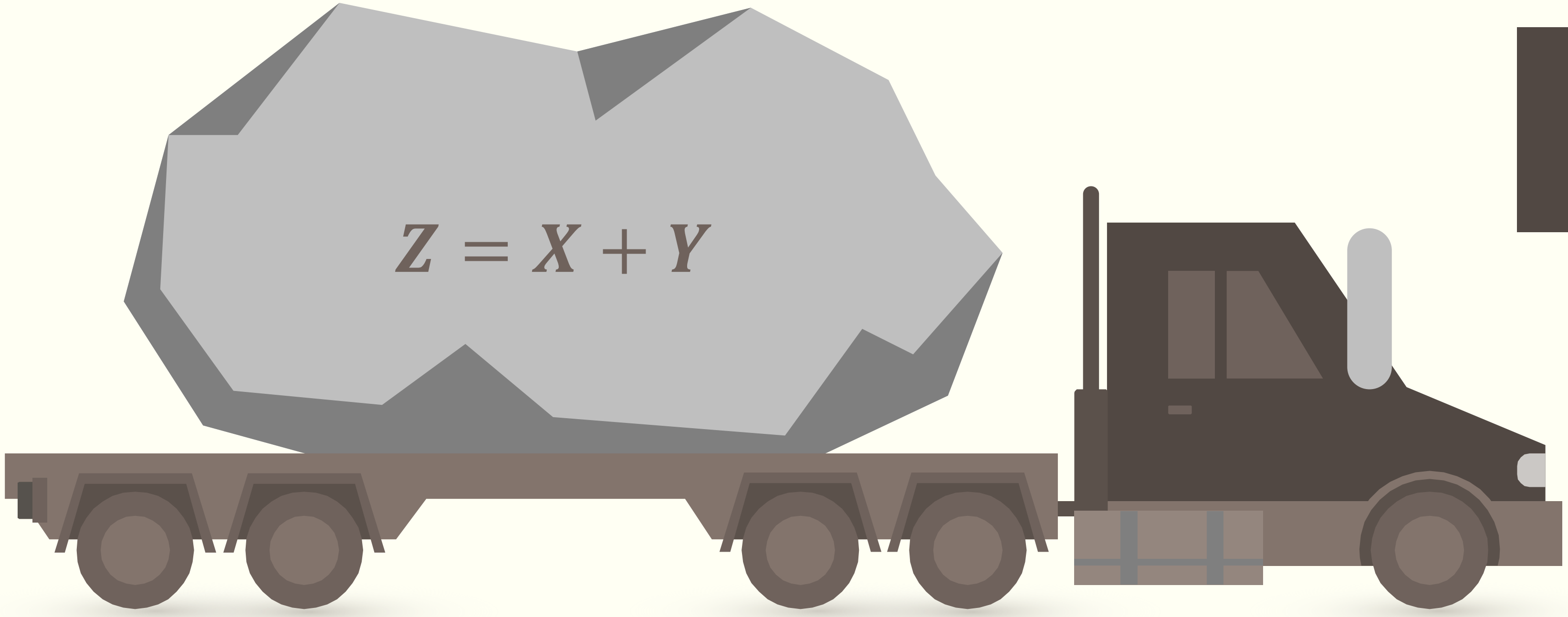
$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$$
$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-(z-y)^2/2} e^{-y^2/2} dy$$
$$= \frac{1}{\sqrt{4\pi}} e^{-z^2/4}$$



Sum of Two Independent Normal Random Variables

convolution

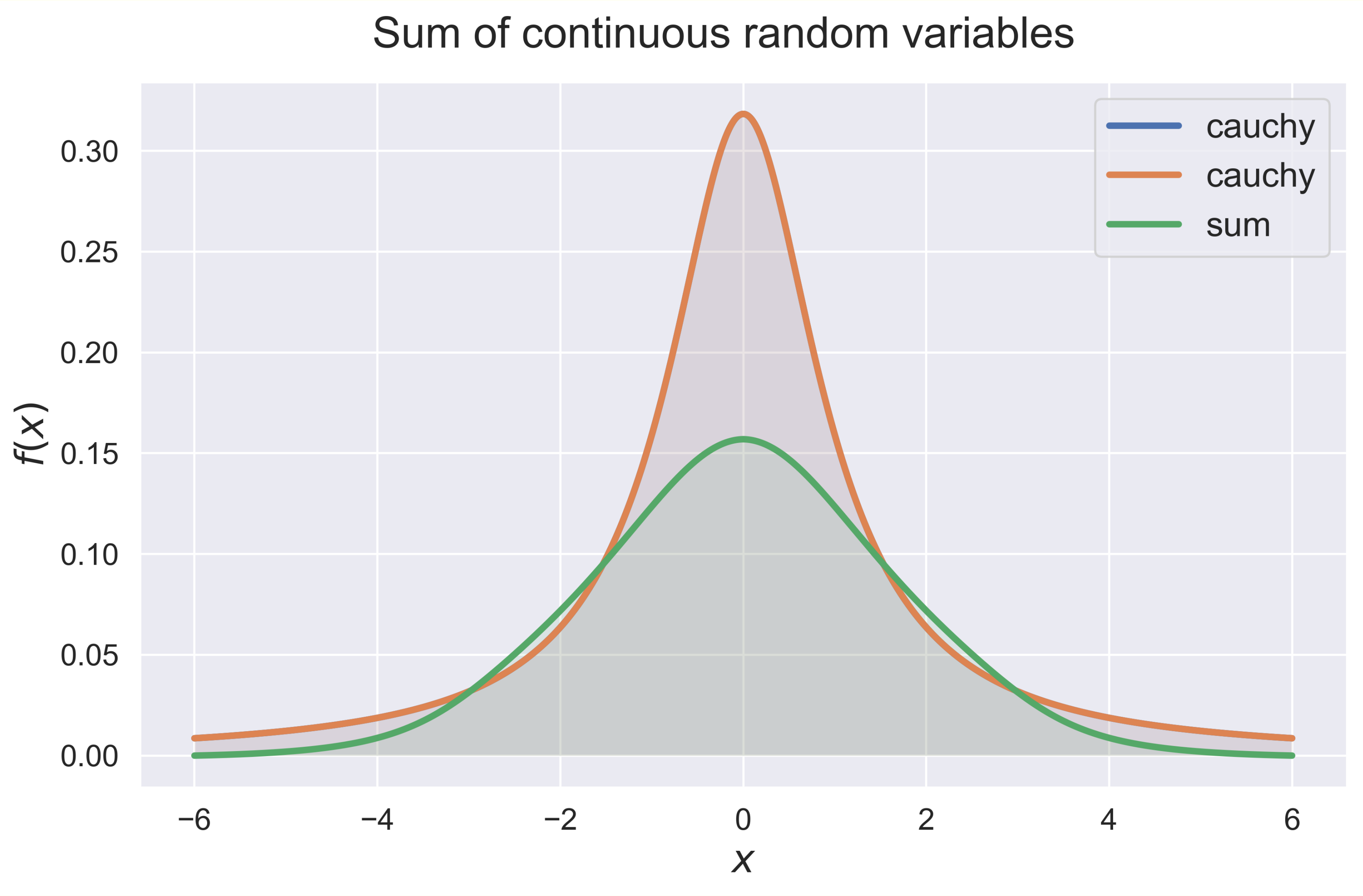
- Means: μ_1 and μ_2
- Variances: σ_1^2 and σ_2^2



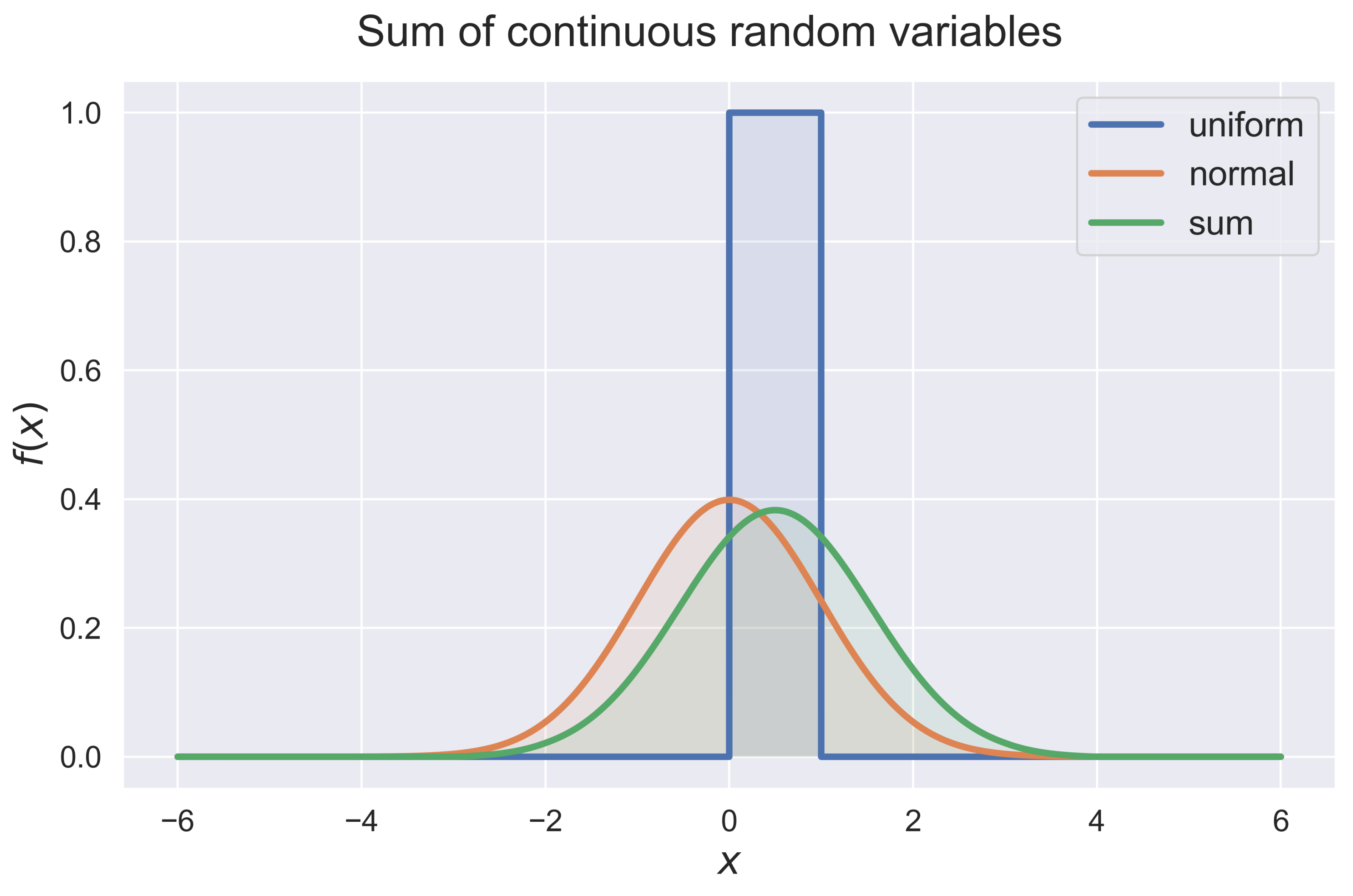
Normal + Normal = Normal

- Means: $\mu_1 + \mu_2$
- Variances: $\sigma_1^2 + \sigma_2^2$

Sum of Two Independent Random Variables (quiz)



Sum of Two Independent Random Variables



Slides 0817

- **Sum of Independent Random Variables.**

```
convolution(dist_1 = 'uniform',
            dist_2 = 'normal',
            interval = [-3, 3])
figure_convolution(dist_1 = 'uniform',
                   dist_2 = 'uniform',
                   interval = [-0.2, 1.1],
                   fsize = (10, 6), fs = 20)
```