SM-1402 Basic Statistics

Chapter 4 (Continuous Distributions)

Dr. Haziq Jamil

Semester 2, 2019/20

Continuous Distribution

- This chapter concerns **continuous random variables**.
- When a variable is continuous, it can take on infinitely many, uncountable values, for example
  - The mass (grams) of a bag of sugar packaged by a particular machine.
  - The time taken (minutes) to perform a task.
  - The height (cm) of a three-year-old girl.
  - The lifetime (hours) of a 100-watt light bulb.

Continuous Distribution

- A continuous random variable X is described by its **probability density function** (pdf).
- It is specified for the range of values for which x is valid (support).
- The pdf can be illustrated by a curve $y = f(x)$. Note that this function cannot be negative.
- Probabilities are given by the **area under the (pdf) curve**. It is sometimes possible to find an area by geometry (e.g. area of triangles, rectangles, etc.). However, often areas need to be calculated using integration.

Continuous Distribution

Example 5.1: X *is the delay in hours of a flight from Kuching, where*

$$f(x) = 0.2 - 0.02x, \qquad 0 \leq x \leq 10$$

Find

(a) the probability that the delay will be less than four hours.

(b) the probability that the delay will be between two and six hours.

Continuous Distribution

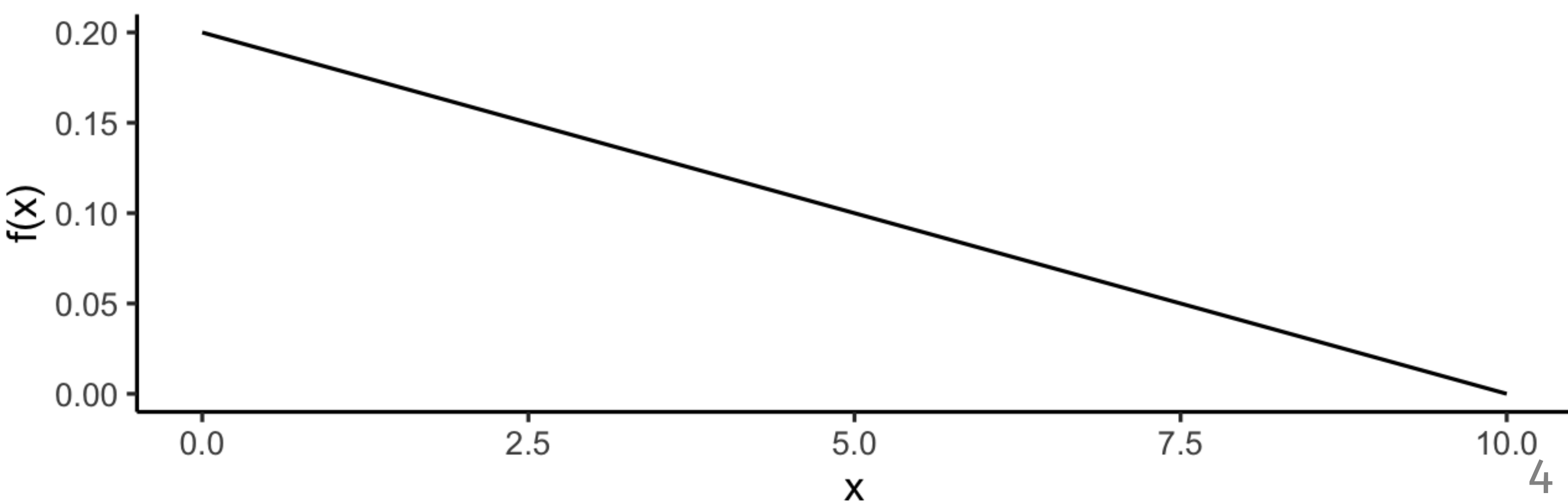
Example 5.1: X *is the delay in hours of a flight from Kuching, where*

$$f(x) = 0.2 - 0.02x, \qquad 0 \le x \le 10$$

Find

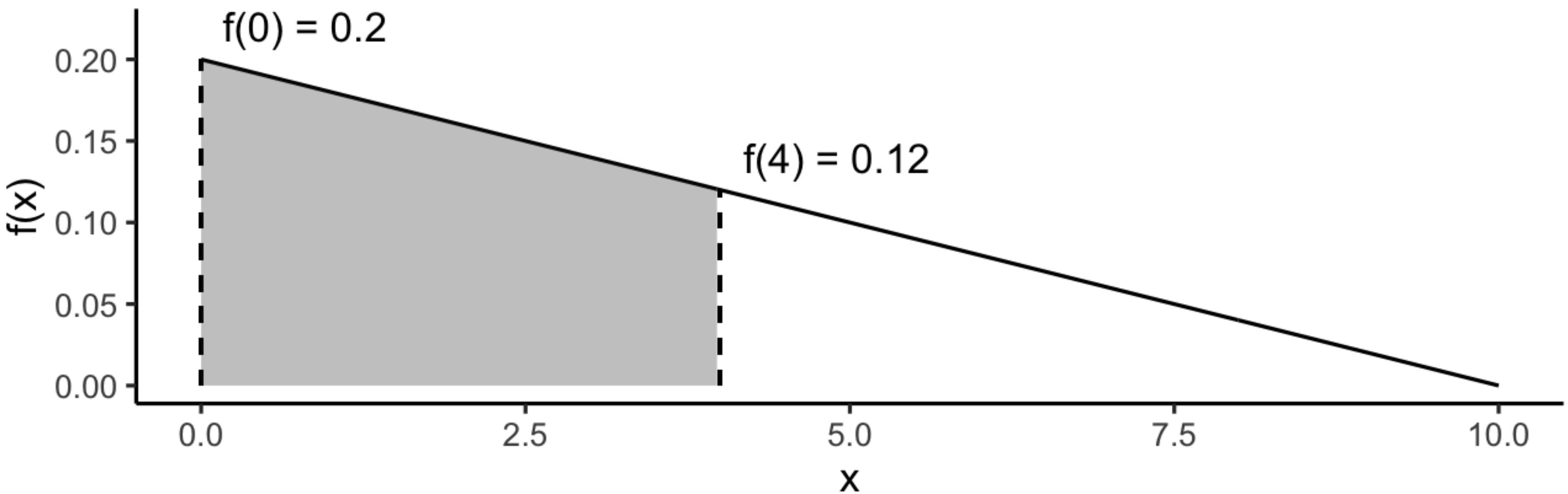
(a) the probability that the delay will be less than four hours.

(b) the probability that the delay will be between two and six hours.



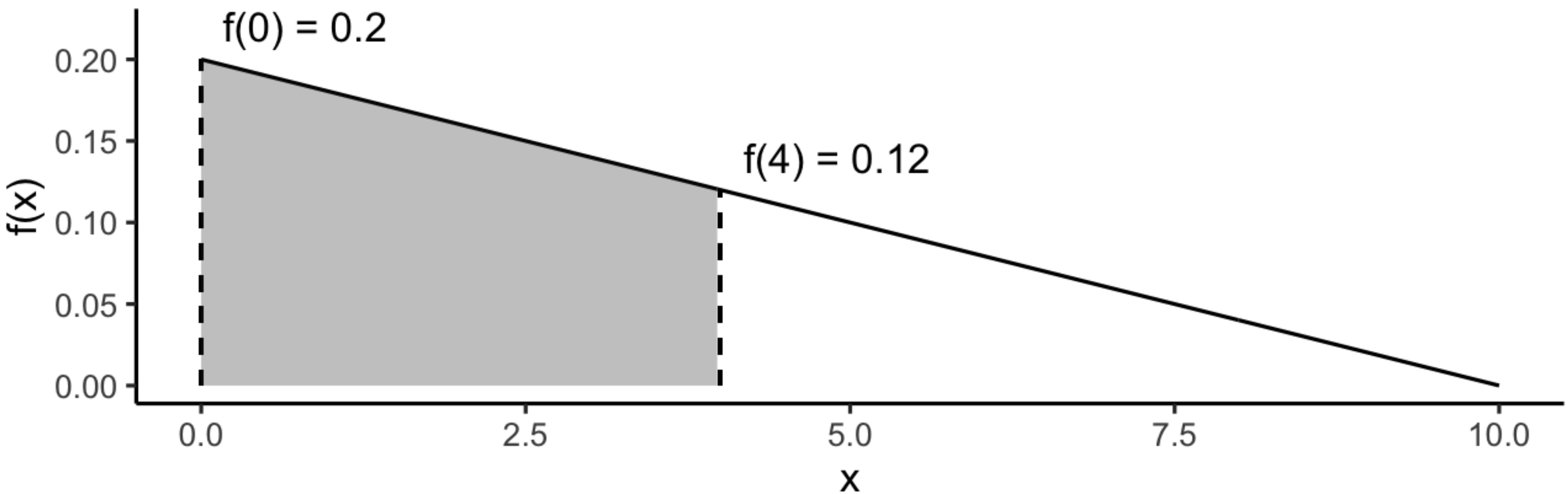
Continuous Distribution

(a) the probability that the delay will be less than four hours.



Continuous Distribution

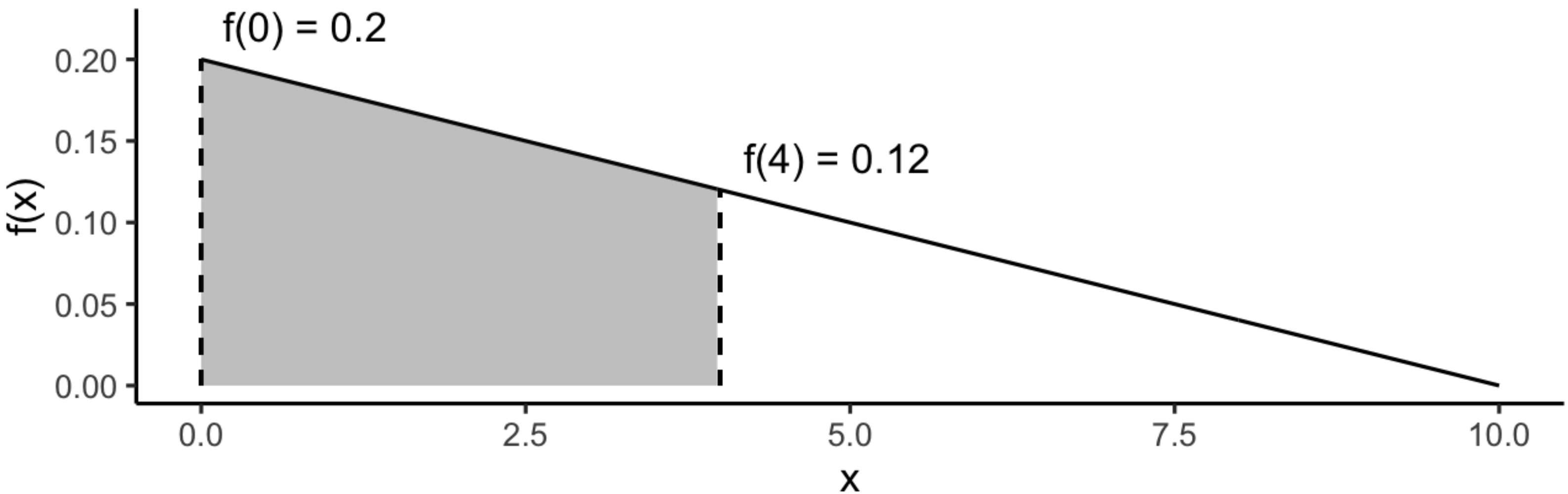
(a) the probability that the delay will be less than four hours.



Area = $\frac{1}{2}(a+b)h = \frac{1}{2}(0.2+0.12)(4) = 0.64$

Continuous Distribution

(a) the probability that the delay will be less than four hours.

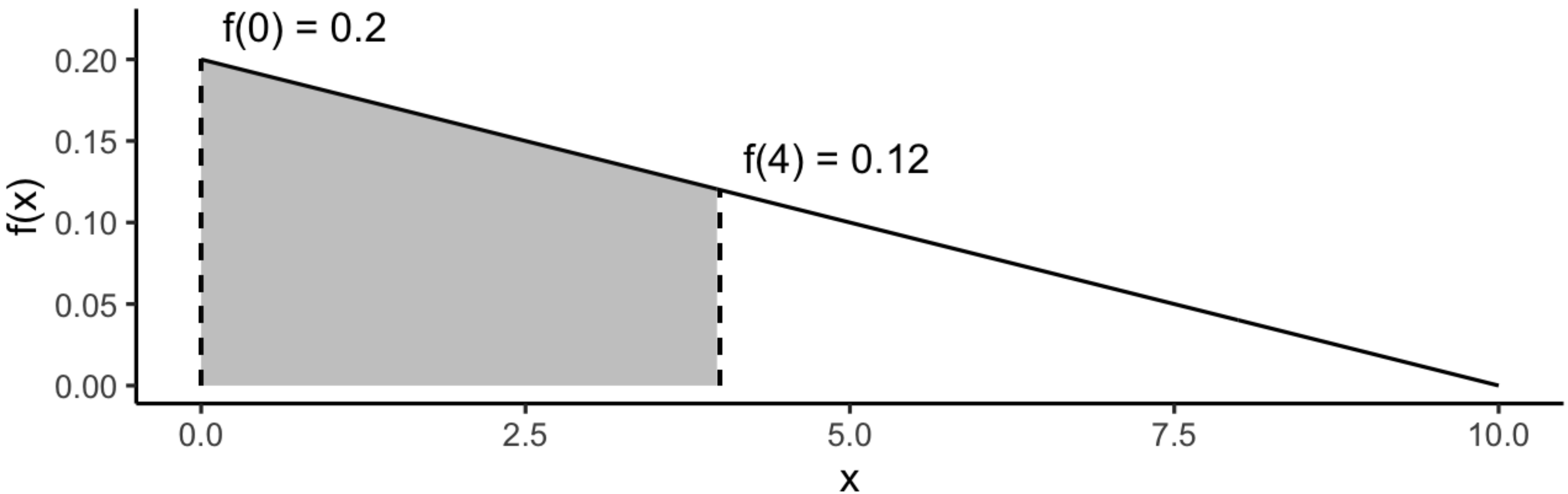


Area = $\frac{1}{2}(a+b)h = \frac{1}{2}(0.2+0.12)(4) = 0.64$

Alternatively, area = $\int_0^4 0.2 - 0.02x\,dx = [0.2x - 0.01x^2]_0^4 = 0.64$

Continuous Distribution

(a) the probability that the delay will be less than four hours.



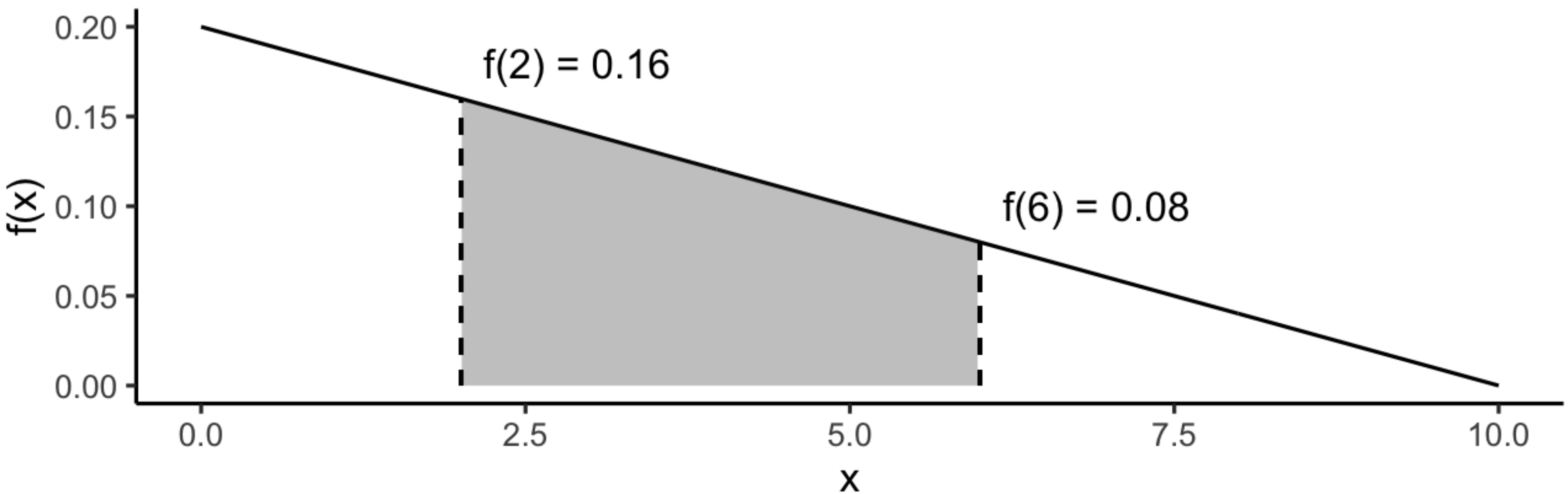
Area = $\frac{1}{2}(a+b)h = \frac{1}{2}(0.2+0.12)(4) = 0.64$

Alternatively, area = $\int_0^4 0.2 - 0.02x\,dx = [0.2x - 0.01x^2]_0^4 = 0.64$

$\therefore P(X \leq 4) = 0.64$

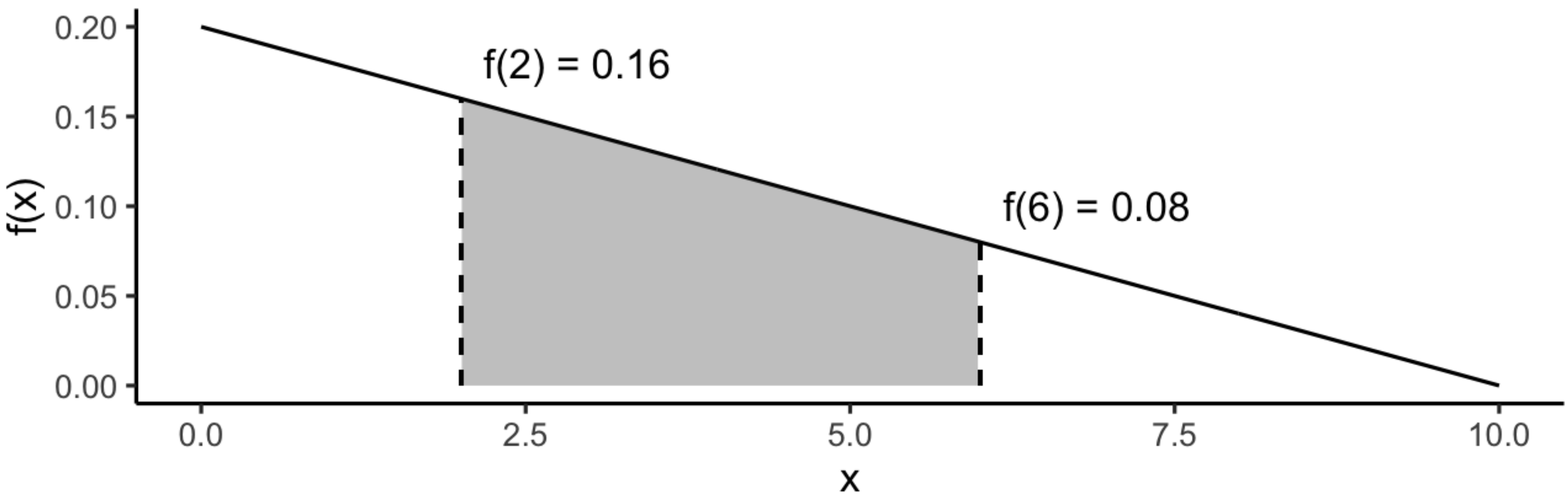
Continuous Distribution

(b) the probability that the delay will be between two and six hours.



Continuous Distribution

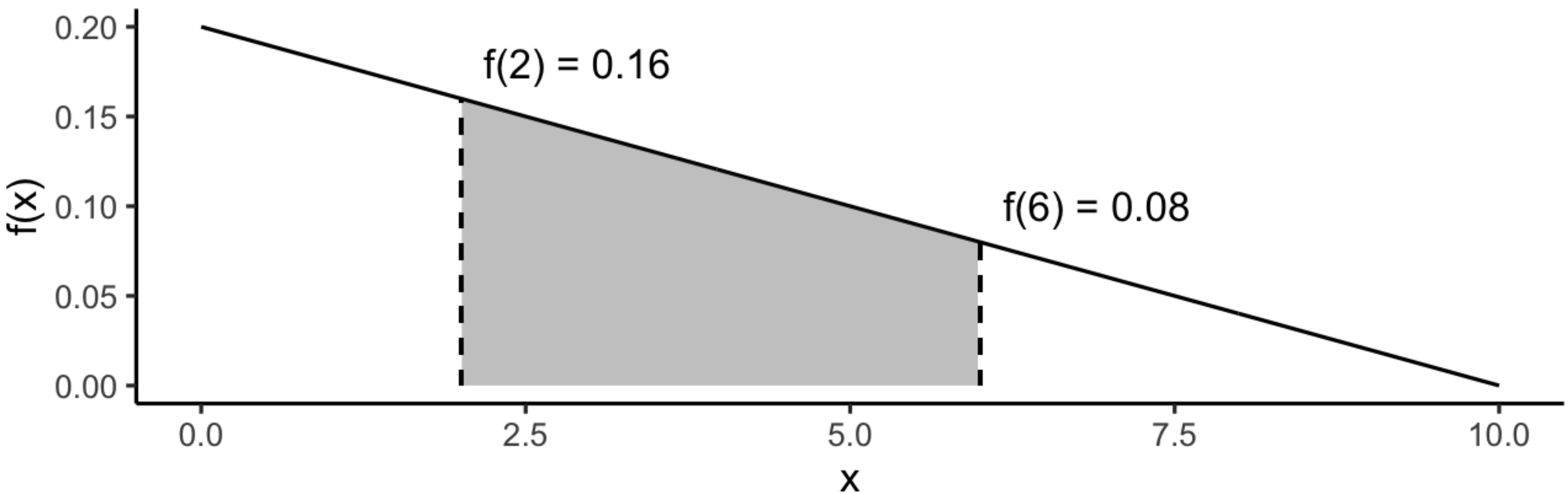
(b) the probability that the delay will be between two and six hours.



Area = $\frac{1}{2}(a+b)h = \frac{1}{2}(0.16+0.08)(4) = 0.48$

Continuous Distribution

(b) the probability that the delay will be between two and six hours.

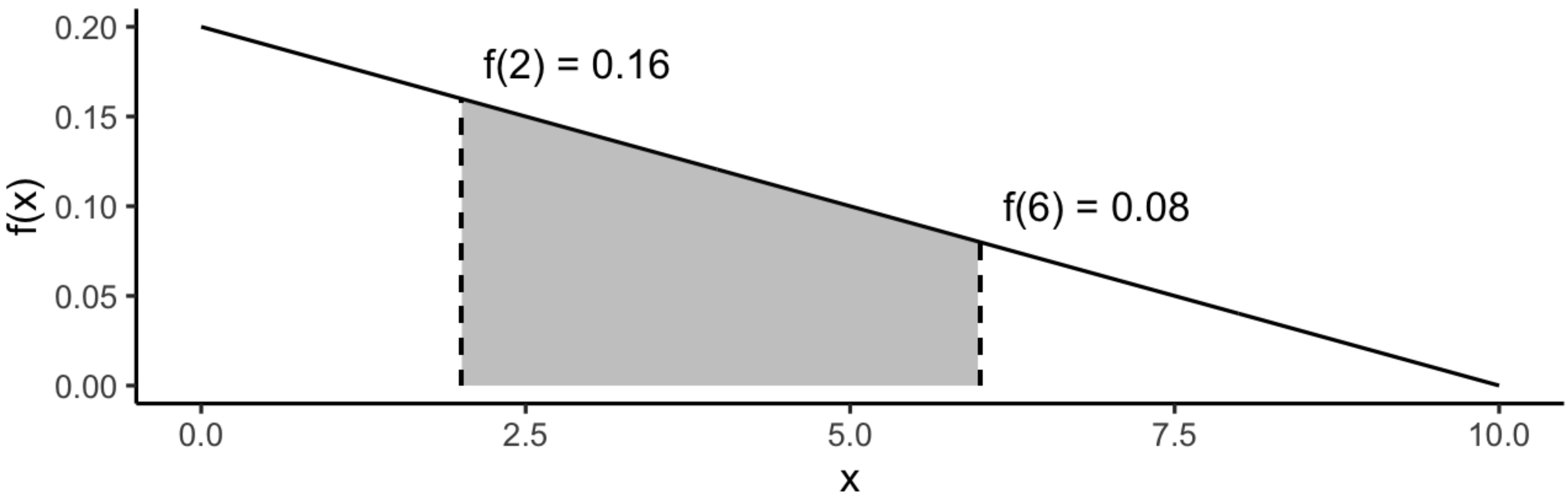


Area = $\frac{1}{2}(a+b)h = \frac{1}{2}(0.16+0.08)(4) = 0.48$

Alternatively, area = $\int_0^4 0.2 - 0.02x\,dx = [0.2x - 0.01x^2]_2^6 = 0.48$

Continuous Distribution

(b) the probability that the delay will be between two and six hours.



Area = $\frac{1}{2}(a+b)h = \frac{1}{2}(0.16+0.08)(4) = 0.48$

Alternatively, area = $\int_0^4 0.2 - 0.02x\,dx = [0.2x - 0.01x^2]_2^6 = 0.48$

$\therefore P(X \leq 4) = 0.48$

Continuous Distribution

- It is not possible to find the probability that the delay is, say, **exactly** three hours.
- If you try to integrate, then

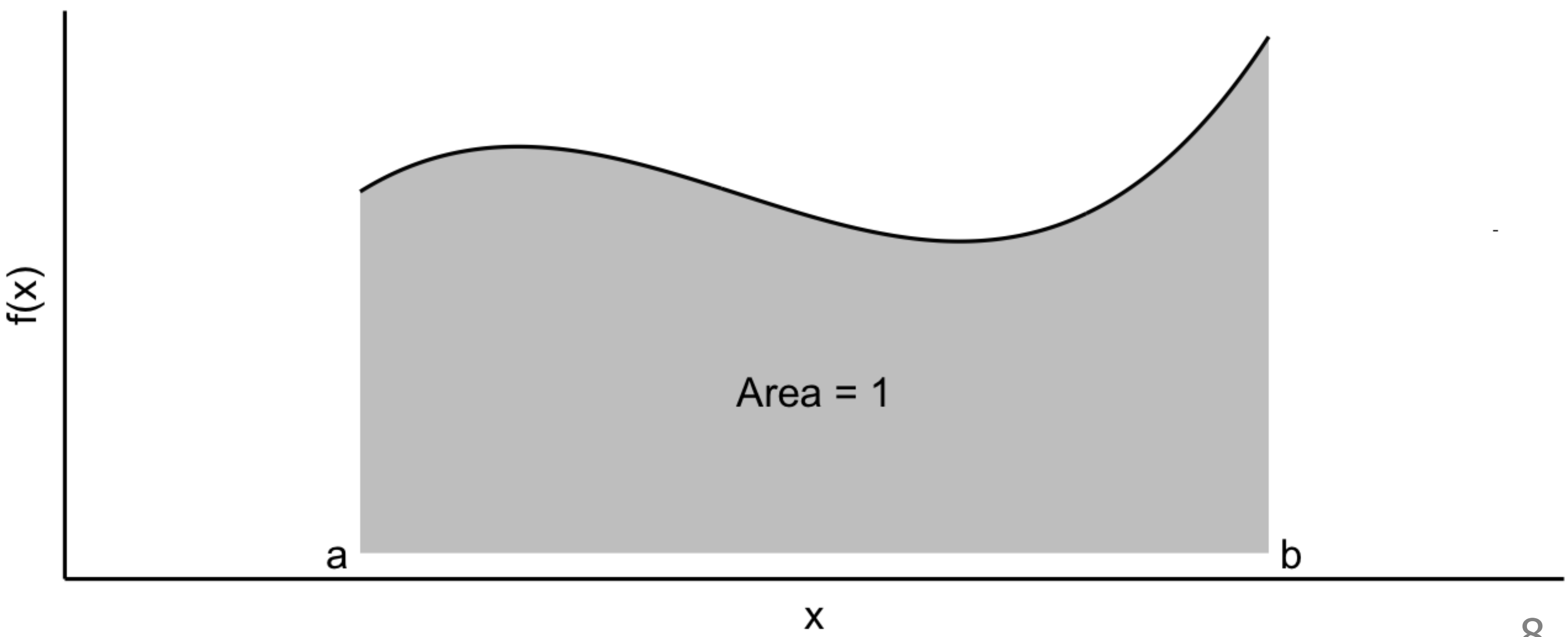
$$\int_3^3 0.2 - 0.02x\,dx = [0.2x - 0.01x^2]_3^3 = 0$$

- You can only find the probability that X lies within a range.
- Also, for the same reason, it does not matter whether $\leq$ or $<$ was used. Therefore, these are all the same
    - $P(2 \leq X \leq 6)$
    - $P(2 < X \leq 6)$
    - $P(2 \leq X < 6)$
    - $P(2 < X < 6)$

Continuous Distribution

For a continuous random variable X with pdf $f(x)$ valid over the range $[a, b]$,

$$\int_a^b f(x)\,dx = 1$$



Continuous Distribution

Example 5.2: A continuous random variable has pdf $f(x) = kx^2$ *for* $0 < x < 4$*. Find* $P(1 \leq X \leq 3)$*.*

Continuous Distribution

Example 5.2: A continuous random variable has pdf $f(x) = kx^2$ *for* $0 < x < 4$. *Find* $P(1 \le X \le 3)$.

First, solve for k. Since the total area must equal 1,

$$P(0 \le X \le 4) = \int_0^4 kx^2\,dx = \left[\frac{kx^3}{3}\right]_0^4 = 1$$

$$\Rightarrow k4^3/3 - 0 = 1$$

$$\Rightarrow k = 3/64$$

Continuous Distribution

Example 5.2: A continuous random variable has pdf $f(x) = kx^2$ *for* $0 < x < 4$. *Find* $P(1 \leq X \leq 3)$.

Then,

$$P(1 \leq X \leq 3) = \int_1^3 kx^2 \, dx = \left[\frac{kx^3}{3}\right]_1^3 = \frac{k}{3}(3^3 - 1) = 26k/3$$

So $P(1 \leq X \leq 3) = 3/64 \times 26/3 = 26/64 = 0.406$

Expectation

For a continuous random variable X with pdf $f(x)$, the expected value of X is given by the formula

$$E(X) = \int_x x f(x)\, dx$$

$E(X)$ is referred to as the *mean* or *expectation* of X and is often denoted by μ.

Expectation

Example 5.3: X *has pdf* $f(x) = x^2/9$ *for* $0 < x < 3$*. Find (a)* μ*; and (b)* $P(X < \mu)$.

Expectation

Example 5.3: X *has pdf* $f(x) = x^2/9$ *for* $0 < x < 3$*. Find (a)* μ*; and (b)* $P(X < \mu)$.

(a)

$$\begin{aligned} \mu = \int_0^3 x f(x)\,dx &= \int_0^3 x \cdot x^2/9\,dx \\ &= \frac{1}{9}\left[x^4/4\right]_0^3 \\ &= 81/36 = 2.25 \end{aligned}$$

Expectation

Example 5.3: X *has pdf* $f(x) = x^2/9$ *for* $0 < x < 3$. *Find (a)* μ*; and (b)* $P(X < \mu)$.

(a)

$$\begin{aligned}\mu = \int_0^3 x f(x)\,dx &= \int_0^3 x \cdot x^2/9\,dx \\ &= \frac{1}{9}\left[x^4/4\right]_0^3 \\ &= 81/36 = 2.25\end{aligned}$$

(b)

$P(X < \mu) = P(X < 2.25) = \int_0^{2.25} x^2/9\,dx = [x^3/27]_0^{2.25} = 0.42$

Expectation

The expectation of **any function** of X can also be derived[1] as

$$E\big(g(X)\big) = \int_x g(x) f(x)\, dx$$

For example,

- $E(X^2) = \int_x x^2 f(x)\, dx$
- $E(1/X) = \int_x \frac{1}{x} f(x)\, dx$
- $E(e^{tx}) = \int_x e^{tx} f(x)\, dx$

[1] Assuming existence!

Expectation

The properties of expectations also apply to continuous random variables: In general, for constants a, b and c, and two random variables X and Y,

- $E(a) = a$
- $E(aX) = aE(X)$
- $E(aX + bY + c) = aE(X) + bE(Y) + c$

Also, for two functions of X, $g(X)$ and $h(X)$,

- $E\big(g(X) + h(X)\big) = E\big(g(X)\big) + E\big(h(X)\big)$

Variance

As in the discrete case, the variance is given by

$$\mathrm{Var}(X) = E\big((X-\mu)^2\big)$$

and for continuous r.v. X with pdf $f(x)$,

$$\begin{aligned}
\mathrm{Var}(X) &= \int_x (x-\mu)^2 f(x)\,dx \\
&= \int_x x^2 f(x)\,dx - \mu^2 \\
&= \int_x x^2 f(x)\,dx - E^2(X) \\
&= E(X^2) - E^2(X)
\end{aligned}$$

Variance

An as before, the following results also hold when X is continuous (a and b are constants)

- $\mathrm{Var}(X) \geq 0$
- $\mathrm{Var}(a) = 0$
- $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$

Given another r.v. Y such that X and Y are **independent**, and a constant c, then

- $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$
- $\mathrm{Var}(aX \pm bY + c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$

Variance

Example 5.6: The continuous random variable X *has pdf* $f(x) = x/8$, *for* $0 < x < 4$. *Find the variance of* X.

Variance

Example 5.6: The continuous random variable X has pdf $f(x) = x/8$, for $0 < x < 4$. Find the variance of X.

First find $E(X)$ and $E(X^2)$

$$E(X) = \int_0^4 x \cdot \frac{x}{8}\, dx = \left[\frac{x^3}{24}\right]_0^4 = 64/24 = 8/3$$

Variance

Example 5.6: The continuous random variable X has pdf $f(x) = x/8$, for $0 < x < 4$. Find the variance of X.

First find $E(X)$ and $E(X^2)$

$$E(X) = \int_0^4 x \cdot \frac{x}{8}\,dx = \left[\frac{x^3}{24}\right]_0^4 = 64/24 = 8/3$$

$$E(X^2) = \int_0^4 x^2 \cdot \frac{x}{8}\,dx = \left[\frac{x^4}{32}\right]_0^4 = 256/32 = 8$$

Variance

Example 5.6: The continuous random variable X has pdf $f(x) = x/8$, for $0 < x < 4$. Find the variance of X.

First find $E(X)$ and $E(X^2)$

$$E(X) = \int_0^4 x \cdot \frac{x}{8}\,dx = \left[\frac{x^3}{24}\right]_0^4 = 64/24 = 8/3$$

$$E(X^2) = \int_0^4 x^2 \cdot \frac{x}{8}\,dx = \left[\frac{x^4}{32}\right]_0^4 = 256/32 = 8$$

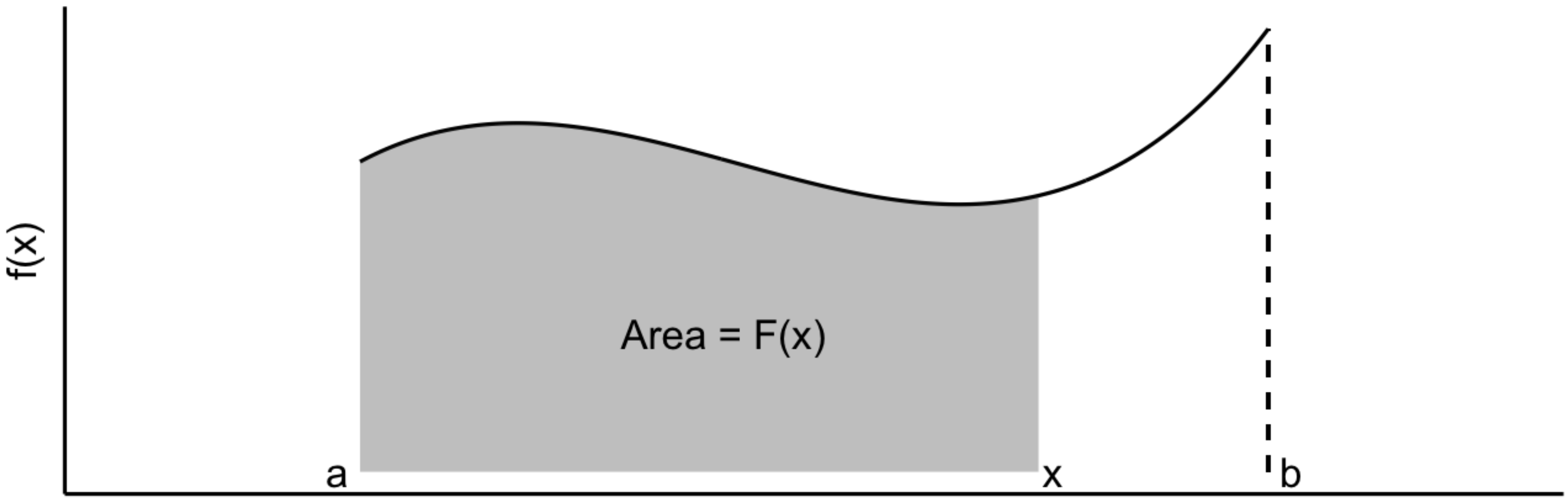
Thus, $\mathrm{Var}(X) = E(X^2) - E^2(X) = 8 - (8/3)^2 = 8/9$

Cumulative distribution function

Recall that the cdf is defined as $F(x) = P(X \leq x)$.

For a continuous r.v. X with pdf $f(x)$, the cdf is

$$F(x) = \int_{-\infty}^{x} f(\tilde{x})\, d\tilde{x}$$



Cumulative distribution function

Example: X *has pdf* $f(x) = x/8$, *for* $0 < x < 4$. *Find* $P(1 < X < 3)$ *using the cdf.*

Cumulative distribution function

Example: X *has pdf* $f(x) = x/8$*, for* $0 < x < 4$*. Find* $P(1 < X < 3)$ *using the cdf.*

First find the cdf

$$\begin{aligned} F(x) = \int_{-\infty}^{x} \tilde{x}/8 \, d\tilde{x} &= \int_{0}^{x} \tilde{x}/8 \, d\tilde{x} \\ &= [\tilde{x}^2/16]_0^x \\ &= x^2/16 \end{aligned}$$

Cumulative distribution function

Example: X has pdf $f(x) = x/8$, for $0 < x < 4$. Find $P(1 < X < 3)$ using the cdf.

First find the cdf

$$
\begin{aligned}
F(x) = \int_{-\infty}^{x} \tilde{x}/8 \, d\tilde{x} &= \int_{0}^{x} \tilde{x}/8 \, d\tilde{x} \\
&= [\tilde{x}^2/16]_0^x \\
&= x^2/16
\end{aligned}
$$

$$
\begin{aligned}
P(1 < X < 3) &= F(3) - F(1) \\
&= 3^2/16 - 1/16 \\
&= 8/16 = 0.5
\end{aligned}
$$

Special continuous distributions: The normal distributions

The normal distribution

The normal distribution is the most important distribution in statistics.

- Many naturally occurring phenomena can be *modelled* as following a normal distribution.
- The central limit theorem (CLT): The disribution of the mean of a sample tends to converge to a normal distribution, as more and more samples are collected.
- Often, the normal distribution is used for the error term in standard statistical models (e.g. linear regression).

The normal distribution

Let X be disributed according to a normal distribution with mean μ and variance σ^2. We write $X \sim \mathrm{N}(\mu, \sigma^2)$. Then its pdf is

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

The normal distribution

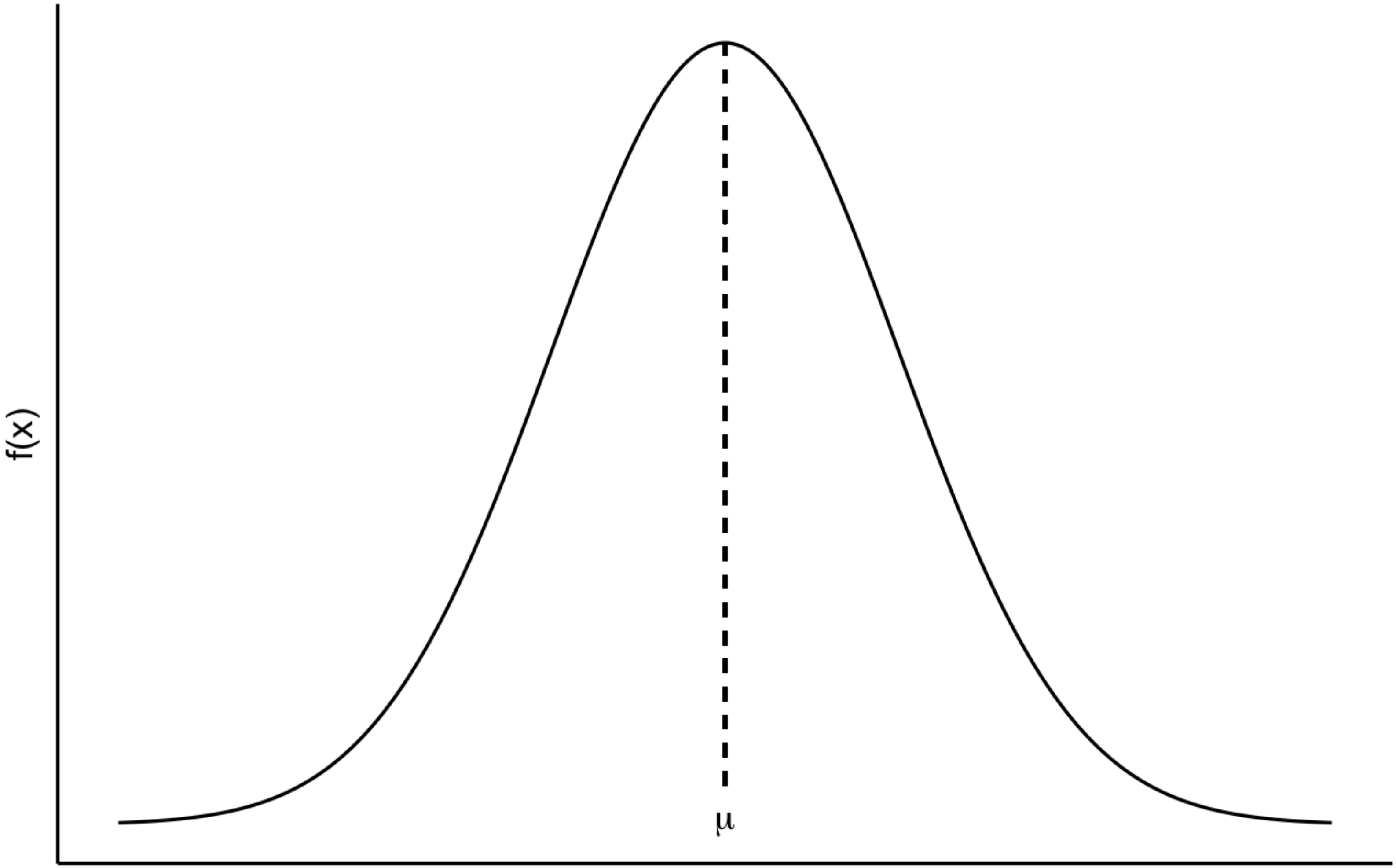
Let X be disributed according to a normal distribution with mean μ and variance σ^2. We write $X \sim \mathrm{N}(\mu, \sigma^2)$. Then its pdf is

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

<u>Properties</u>

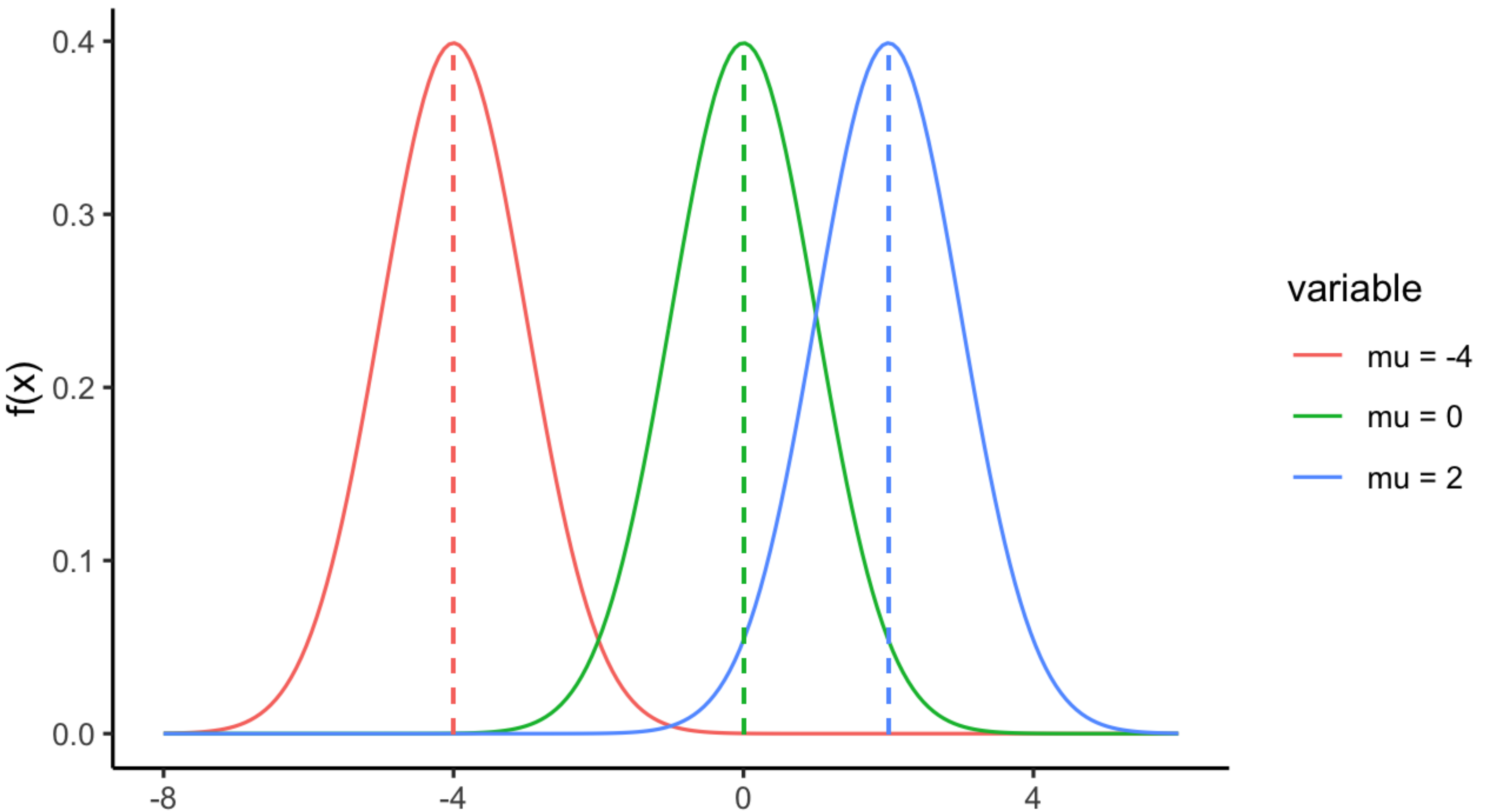
- The parameters of the normal distribution are the mean $E(X) = \mu$ and the variance $\mathrm{Var}(X) = \sigma^2$.
- Support: $x \in \mathbb{R}$.
- The normal distribution is **symmetric** about μ, and **centred** at μ.
- Due to symmetry, median = mean = μ.
- The mode is also μ (peaks at the mean).

The normal distribution



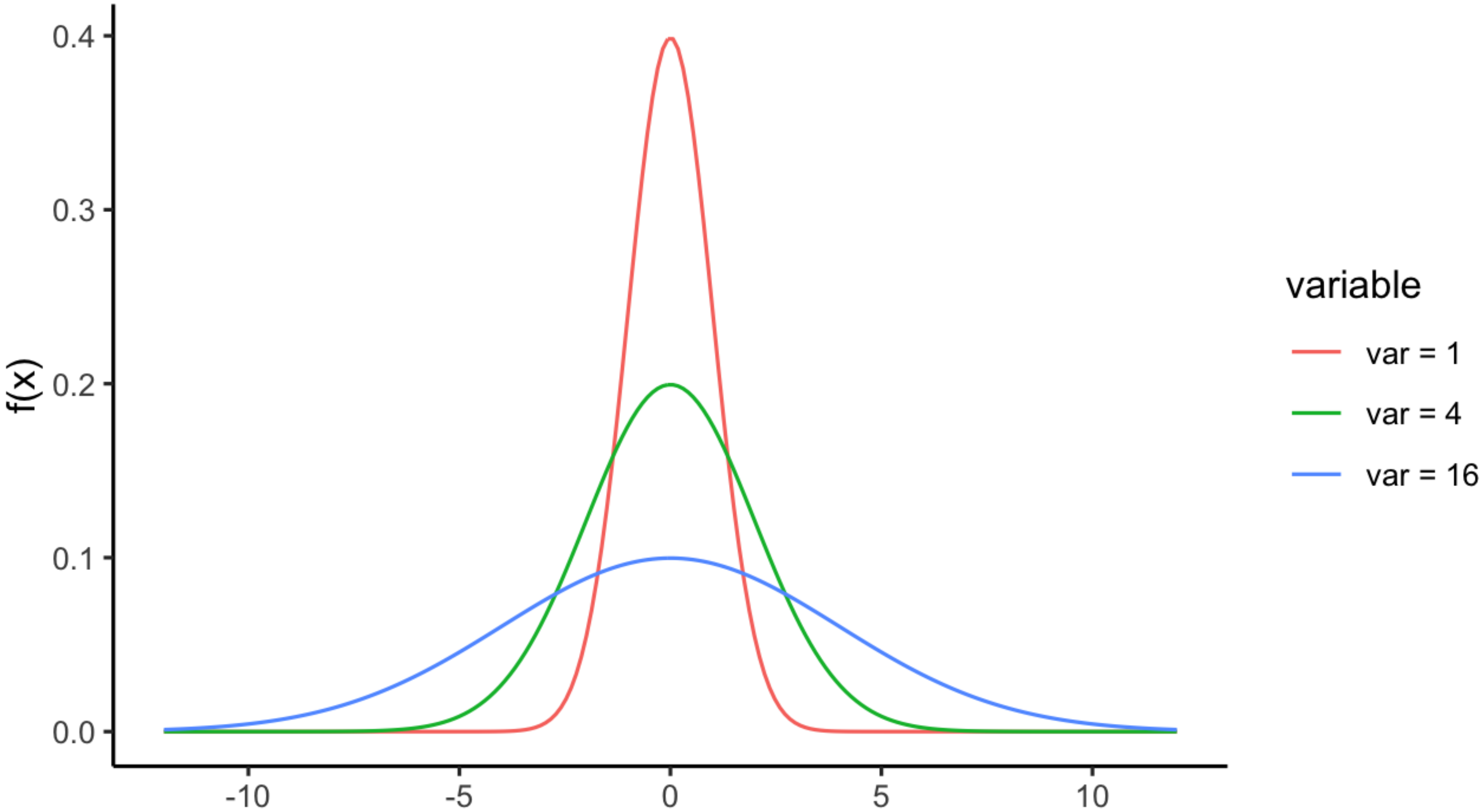
The normal distribution

Effect of changing the mean (location). Variance $\sigma^2 = 1$.



The normal distribution

Effect of changing the variance (scale/spread). Mean $\mu = 0$.



The standard normal

The most important normal distribution is the special case when $\mu = 0$ and $\sigma^2 = 1$. This is called the **standard normal** distribution, denoted

$$Z \sim \mathrm{N}(0, 1)$$

with pdf

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

The pdf of Z is denoted specially by the greek letter ϕ, pronounced 'phi'.

The standard normal

Any normal r.v. can be **standardized** by

- subtracting the mean
- dividing by the standard deviation

Let $X \sim \mathrm{N}(\mu, \sigma^2)$. Then

$$Z := \frac{X - \mu}{\sigma} \sim \mathrm{N}(0, 1)$$

The standard normal

Calculating probabilities of the (standard) normal distribution is not possible to do explicitly using integration. That is,

$$P(a \leq Z \leq b) = \int_a^b \phi(z)\, dz$$

is **not tractable** (no closed form exists).

Instead, we rely on **computer approximations**, or by manually searching for the values from a statistical table. Statistical tables will show values of the standard normal cdf

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \phi(\tilde{z})\, d\tilde{z}$$

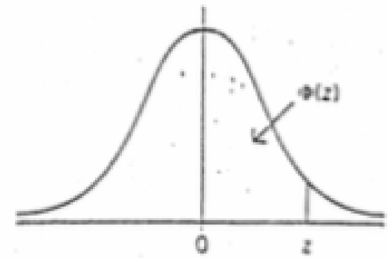
The standard normal

THE STANDARD NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$ where

$$\Phi(z) = P(Z < z).$$

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	ADD								
											1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23

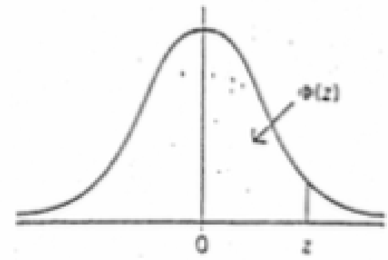
The standard normal

THE STANDARD NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$ where

$\Phi(z) = P(Z < z)$.

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	ADD 1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23

Find (a) $P(Z < 0.85)$, (b) $P(Z > 0.85)$

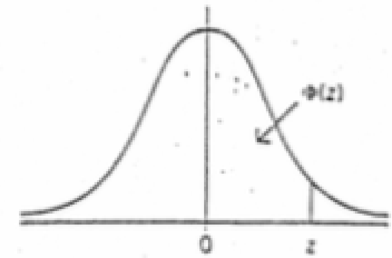
The standard normal

THE STANDARD NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$ where

$\Phi(z) = P(Z < z)$.

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	ADD 1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23

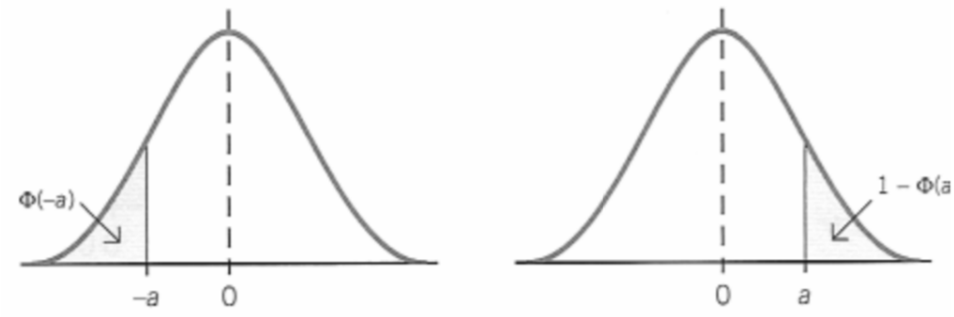
Find (a) $P(Z < 0.85)$, (b) $P(Z > 0.85)$

(a) 0.8023, (b) $1 - 0.8023 = 0.1977$

The standard normal

Finding probabilities involving negative values of z.

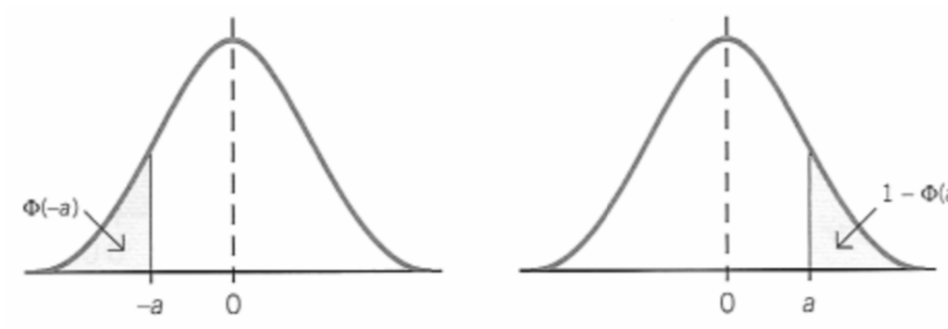
$$P(Z < -z) = \Phi(-z) = 1 - \Phi(z)$$



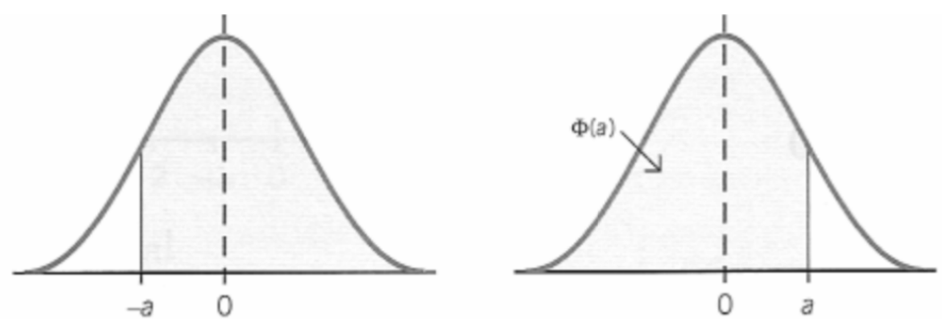
The standard normal

Finding probabilities involving negative values of z.

$$P(Z < -z) = \Phi(-z) = 1 - \Phi(z)$$



$$P(Z > -z) = 1 - P(Z < -z) = 1 - (1 - \Phi(z)) = \Phi(z)$$



The standard normal

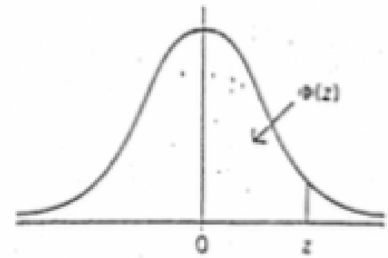
Example: Find $P(Z < -0.913)$ *and* $P(Z > -0.913)$.

THE STANDARD NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$ where

$\Phi(z) = P(Z < z)$.

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



											1	2	3	4	5	6	7	8	9
z	0	1	2	3	4	5	6	7	8	9	ADD								
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23

The standard normal

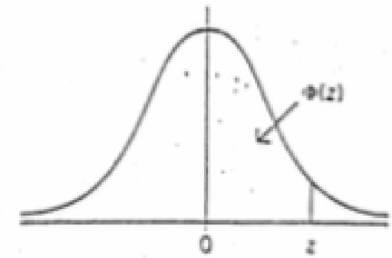
Example: Find $P(Z < -0.913)$ *and* $P(Z > -0.913)$.

THE STANDARD NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$ where

$\Phi(z) = P(Z < z)$.

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
											ADD								
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23

$$P(Z < -0.913) = 1 - \Phi(0.913) = 1 - 0.8194 = 0.1806.$$

The standard normal

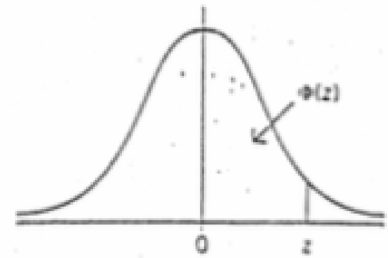
Example: Find $P(Z < -0.913)$ *and* $P(Z > -0.913)$.

THE STANDARD NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$ where

$\Phi(z) = P(Z < z)$.

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
											ADD								
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23

$$P(Z > -0.913) = \Phi(0.913) = 0.8194.$$

The standard normal

Important results worth learning

- $P(Z \le -a) = 1 - \Phi(a)$
- $P(Z \le -a) = \Phi(a)$
- $P(a \le Z \le b) = \Phi(b) - \Phi(a)$
- $P(-a \le Z \le b) = \Phi(a) + \Phi(b) - 1$
- $P(-a \le Z \le -b) = \Phi(b) - \Phi(a)$
- $P(|Z| \le a) = P(-a \le Z \le a) = 2\Phi(a) - 1$
- $P(|Z| \ge a) = P(\{Z < -a\} \cup \{Z > a\}) = 2\big(1 - \Phi(a)\big)$

The standard normal

Example 6.3: Lengths of metal strips produced by a machine are normally distributed with mean length of 150cm and a standard deviation of 10cm. Find the probability that the length of a randomly selected strip is

(a) shorter than 165cm.

(b) within 5cm of the mean.

The standard normal

Example 6.3: Lengths of metal strips produced by a machine are normally distributed with mean length of 150cm and a standard deviation of 10cm. Find the probability that the length of a randomly selected strip is

(a) shorter than 165cm.

(b) within 5cm of the mean.

Let X represent the lenghts of metal strips produced. Then $X \sim \mathrm{N}(150, 10^2)$.

(a)

$$\begin{aligned} P(X < 165) = P(Z < (165 - 150)/10) = P(Z < 1.5) \\ = \Phi(1.5) = 0.9332 \end{aligned}$$

The standard normal

(b) To be within 5cm of the mean implies that the value X should not be more than 155cm and should not be less than 145cm. That is, $145 < X < 155$.

The standard normal

(b) To be within 5cm of the mean implies that the value X should not be more than 155cm and should not be less than 145cm. That is, $145 < X < 155$.

Thus,

$$\begin{aligned} P(145 < X < 155) &= P\left(\frac{145-150}{10} < Z < \frac{155-150}{10}\right) \\ &= P(-0.5 < Z < 0.5) \\ &= 2\Phi(0.5) - 1 \\ &= 2 \times 0.6915 - 1 \\ &= 0.383 \end{aligned}$$

Linear functions of normal r.v.

If $X \sim \mathrm{N}(\mu, \sigma^2)$ and a and b are constants (with $b \neq 0$), then:

$$a + bX \sim \mathrm{N}(a + b\mu, b^2\sigma^2)$$

Linear functions of normal r.v.

If $X \sim \mathrm{N}(\mu, \sigma^2)$ and a and b are constants (with $b \neq 0$), then:

$$a + bX \sim \mathrm{N}(a + b\mu, b^2\sigma^2)$$

If X_1 and X_2 are *independent* normal random variables, such that $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$, then:

$$X_1 \pm X_2 \sim \mathrm{N}(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$$

Note that variances are *added* even when dealing with the difference between independent random variables.

Linear functions of normal r.v.

Example: An invester has the choice of two out of four investments: X_1*,* X_2*,* X_3 *and* X_4*. The profits (in BND 1,000 per annum) from these may be assumed to be indpendently distributed and:*

- Profit from $X_1 \sim \mathrm{N}(2, 1)$
- Profit from $X_2 \sim \mathrm{N}(3, 3)$
- Profit from $X_3 \sim \mathrm{N}(1, 0.25)$
- Profit from $X_4 \sim \mathrm{N}(2.5, 4)$

Which pair of investments should the investor choose in order to maximise the probability of making a total profit of at least BND 2,000? What is the maximum probability?

Linear functions of normal r.v.

Example: An invester has the choice of two out of four investments: X_1, X_2, X_3 and X_4. The profits (in BND 1,000 per annum) from these may be assumed to be indpendently distributed and:

- Profit from $X_1 \sim \mathrm{N}(2, 1)$
- Profit from $X_2 \sim \mathrm{N}(3, 3)$
- Profit from $X_3 \sim \mathrm{N}(1, 0.25)$
- Profit from $X_4 \sim \mathrm{N}(2.5, 4)$

Which pair of investments should the investor choose in order to maximise the probability of making a total profit of at least BND 2,000? What is the maximum probability?

Let's sum the ${}^4C_2 = 6$ possible pairwise combinations of normal r.v. Then calculate the Z value. We want to maximise $P(Z > 2)$ (the distribution which gives us the smallest Z value).

Linear functions of normal r.v.

- $X_1 + X_2 \sim \mathrm{N}(5, 4)$, $Z = -1.50$
- $X_1 + X_3 \sim \mathrm{N}(3, 1.25)$, $Z = -0.89$
- $X_1 + X_4 \sim \mathrm{N}(4.5, 5)$, $Z = -1.12$
- $X_2 + X_3 \sim \mathrm{N}(4, 3.25)$, $Z = -1.11$
- $X_2 + X_4 \sim \mathrm{N}(5.5, 7)$, $Z = -1.32$
- $X_3 + X_4 \sim \mathrm{N}(3.5, 4.25)$, $Z = -0.73$

Linear functions of normal r.v.

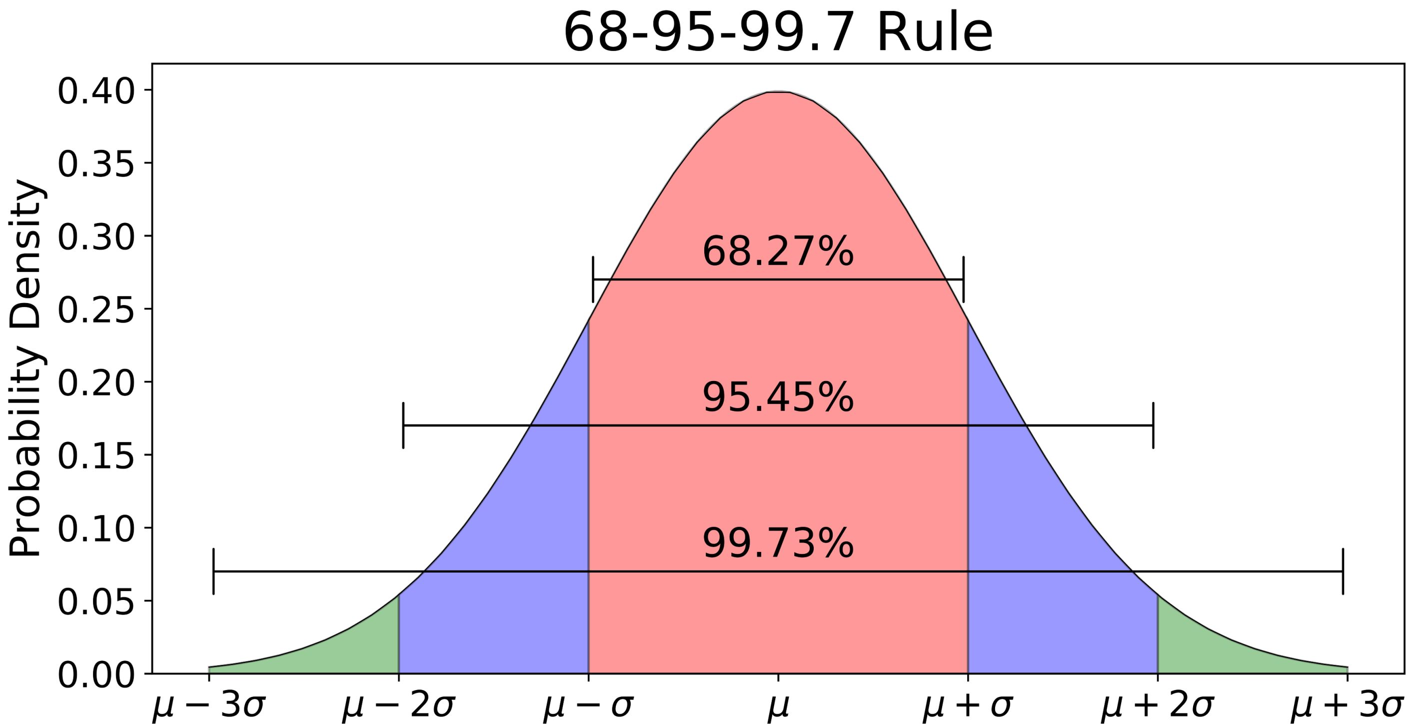
- $X_1 + X_2 \sim \mathrm{N}(5, 4)$, $Z = -1.50$
- $X_1 + X_3 \sim \mathrm{N}(3, 1.25)$, $Z = -0.89$
- $X_1 + X_4 \sim \mathrm{N}(4.5, 5)$, $Z = -1.12$
- $X_2 + X_3 \sim \mathrm{N}(4, 3.25)$, $Z = -1.11$
- $X_2 + X_4 \sim \mathrm{N}(5.5, 7)$, $Z = -1.32$
- $X_3 + X_4 \sim \mathrm{N}(3.5, 4.25)$, $Z = -0.73$

Evidently, the combination of X_1 and X_2 gives us a mean profit of BND 5,000 with a standard deviation of 2. This gives the smalles Z values and thus the highes probability of out all possible pairs of investments.

$$P(Z > -1.5) = \Phi(1.5) = 0.9332$$

68-95-99.7 rule

Also known as the empirical rule, is a shorthand to remember the percentage of values that lie within a band around the mean in a normal distribution.



Normal approximations to discrete distributions

Normal approximation to binomial

Recall that $X \sim \mathrm{Bin}(n, p)$ has these properties

- $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$
- $E(X) = np$
- $Var(X) = np(1-p)$

Normal approximation to binomial

Recall that $X \sim \text{Bin}(n, p)$ has these properties

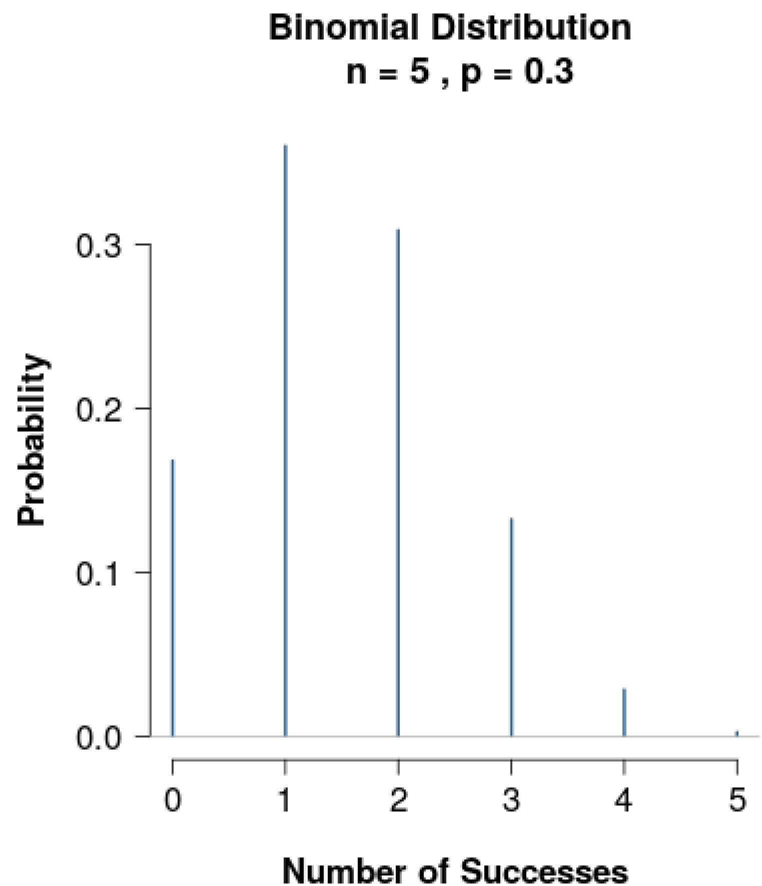
- $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$
- $E(X) = np$
- $Var(X) = np(1-p)$

Consider computing these probabilities below, when $n = 100$:

- $P(X = 50)$ (try calculating $\binom{100}{50}$ and p^{50} for various values of p—what do you observe?)
- $P(X > 50)$
- $P(25 < X < 75)$ (how many computations must you do?)

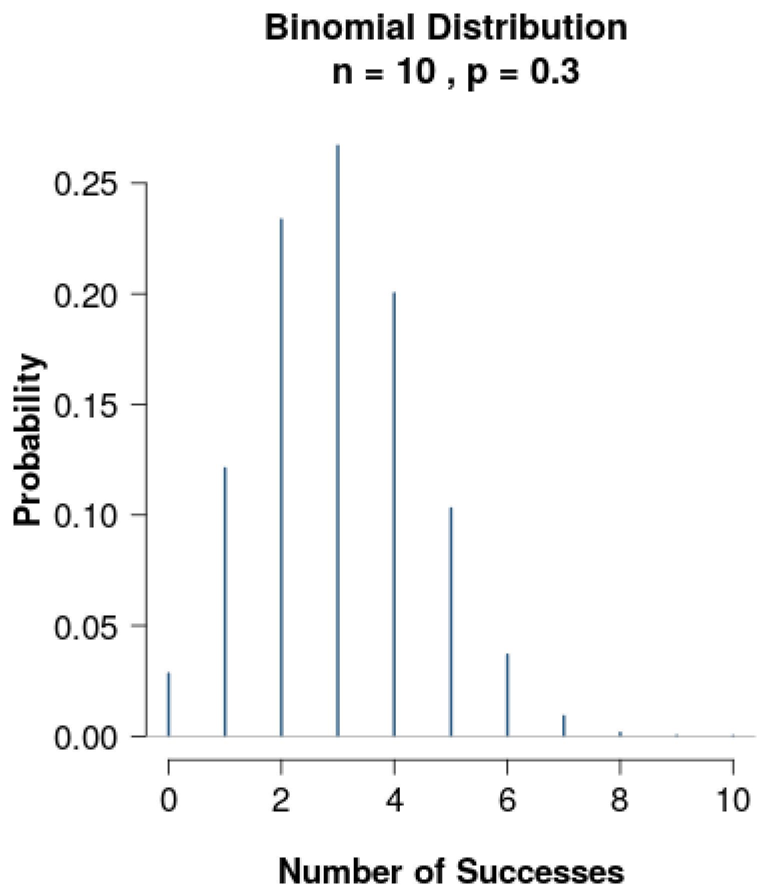
Normal approximation to binomial

What does the distribution actually look like? Fix $p = 0.3$ and let $n \to \infty$:



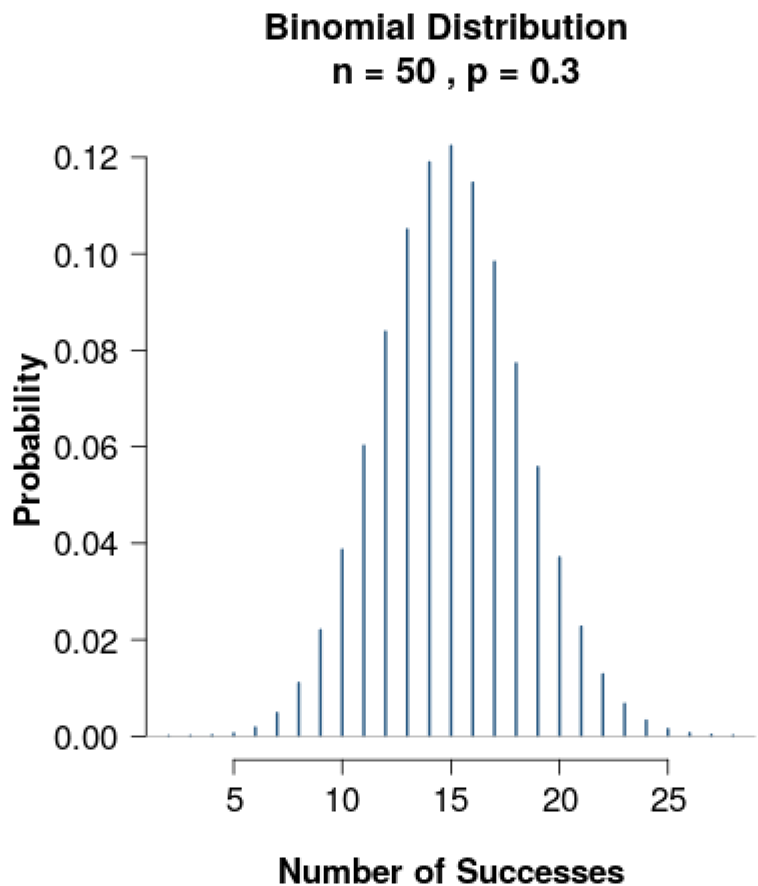
Normal approximation to binomial

What does the distribution actually look like? Fix $p = 0.3$ and let $n \to \infty$:



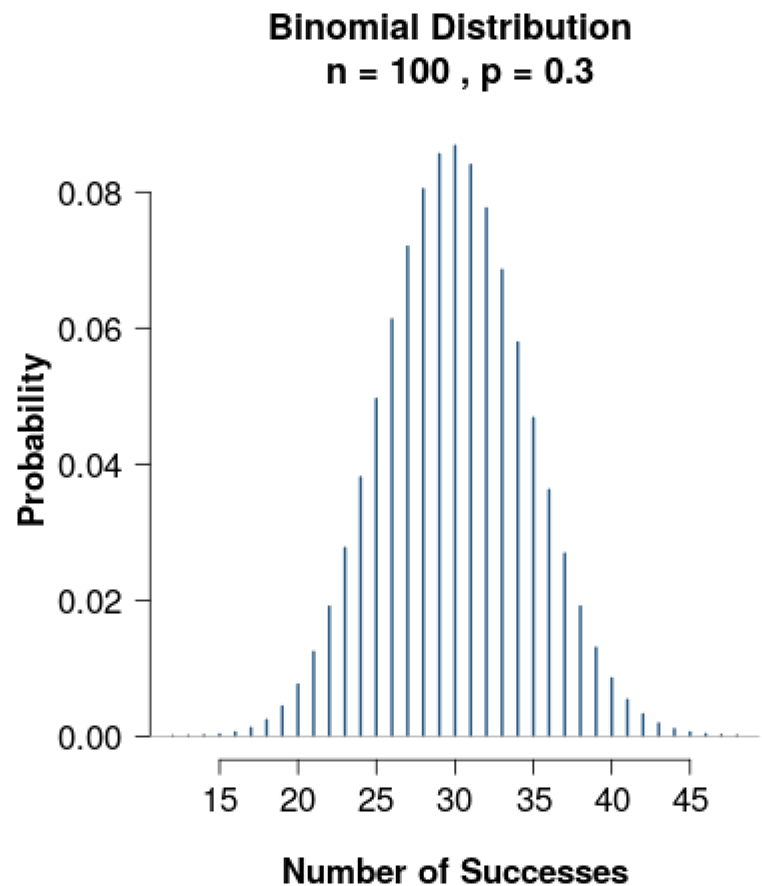
Normal approximation to binomial

What does the distribution actually look like? Fix $p = 0.3$ and let $n \to \infty$:



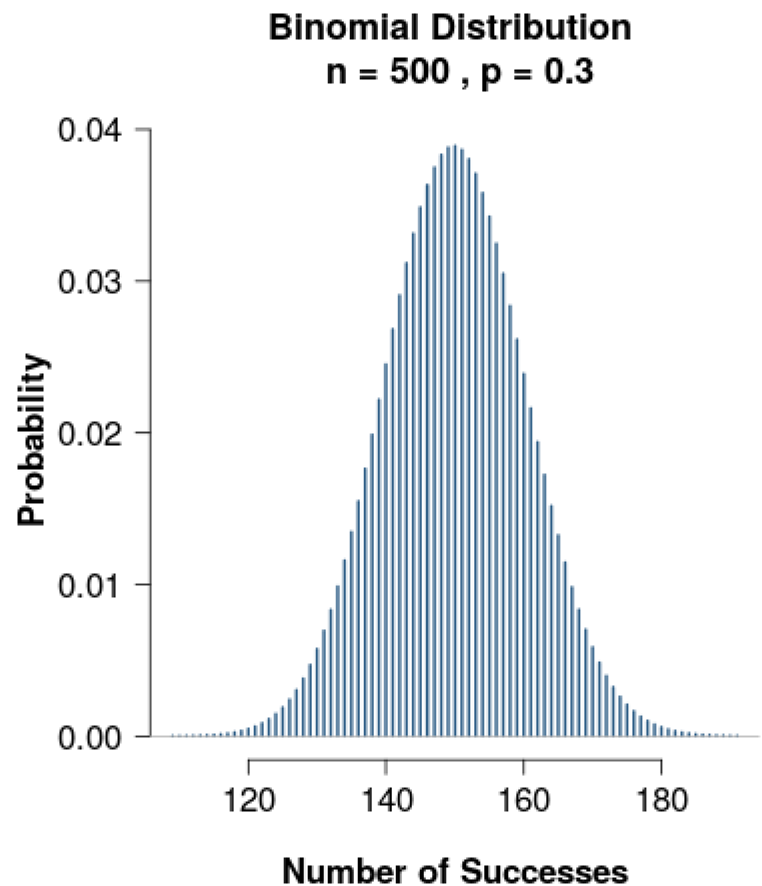
Normal approximation to binomial

What does the distribution actually look like? Fix $p = 0.3$ and let $n \to \infty$:



Normal approximation to binomial

What does the distribution actually look like? Fix $p = 0.3$ and let $n \to \infty$:



Normal approximation to binomial

When `(np >5)` and `(n(1-p)>5)`, it is appropriate to use the normal approximation to the binomial

As $n \to \infty, \mathrm{Bin}(n, p) \to N\big(np, np(1-p)\big)$.

Normal approximation to binomial

When `(np >5)` and `(n(1-p)>5)`, it is appropriate to use the normal approximation to the binomial

As $n \to \infty$, $\text{Bin}(n,p) \to N\big(np, np(1-p)\big)$. This means that for $X \sim \text{Bin}(n,p)$,

$$
\begin{aligned}
P(X < x) &\approx P\left(\frac{X - np}{\sqrt{np(1-p)}} < \frac{x - np}{\sqrt{np(1-p)}}\right) \\
&= \Phi\left(\frac{x - np}{\sqrt{np(1-p)}}\right)
\end{aligned}
$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution.

Continuity correction

For discrete distributions,

$$P(X \leq x) = P(X < x + 1)$$

But we know for continuous distributions, $P(X \leq x) = P(X < x)$.
So when a discrete X is approximated by a continuous distribution, we must apply **continuity correction**.

Continuity correction

For discrete distributions,

$$P(X \leq x) = P(X < x + 1)$$

But we know for continuous distributions, $P(X \leq x) = P(X < x)$. So when a discrete X is approximated by a continuous distribution, we must apply **continuity correction**.

Roughly speaking,

$$P(X \leq x) \approx P(X_{cts} < x + 1/2)$$

Continuity correction

Apply these continuity corrections in your calculations!

Discrete	Continous
$X = c$	$c - 0.5 < X < c + 0.5$
$X < c$	$X < c + 0.5$
$X \leq c$	$X < c + 0.5$
$X > c$	$X > c - 0.5$
$X \geq c$	$X > c - 0.5$

Continuity correction

Apply these continuity corrections in your calculations!

Discrete	Continous
$X = c$	$c - 0.5 < X < c + 0.5$
$X < c$	$X < c + 0.5$
$X \leq c$	$X < c + 0.5$
$X > c$	$X > c - 0.5$
$X \geq c$	$X > c - 0.5$

Questions

- What is continuity correction for $a < X < b$?
- What is $P(X > 0)$?
- What is $P(X = -1)$?

Continuity correction

Example 6.8: Find the probability of obtaining 4, 5, 6 or 7 heads when a fair coin is tossed 12 times.

Continuity correction

Example 6.8: Find the probability of obtaining 4, 5, 6 or 7 heads when a fair coin is tossed 12 times.

Let X be the number of heads when coin is tossed 12 times. $X \sim \mathrm{Bin}(12, 0.5)$. We are interested in $P(X = 4, 5, 6, 7)$. Since np and $n(1-p)$ are both greater than 5, we can use the normal approximation $X \sim N(6, 3)$.

Continuity correction

Example 6.8: Find the probability of obtaining 4, 5, 6 or 7 heads when a fair coin is tossed 12 times.

Let X be the number of heads when coin is tossed 12 times. $X \sim \mathrm{Bin}(12, 0.5)$. We are interested in $P(X = 4, 5, 6, 7)$. Since np and $n(1-p)$ are both greater than 5, we can use the normal approximation $X \sim N(6, 3)$.

$$\begin{aligned}
P(4 \leq X \leq 7) &\approx P(3.5 < X < 7.5) \\
&= P\left(\frac{3.5 - 6}{\sqrt{3}} < Z < \frac{7.5 - 6}{\sqrt{3}}\right) \\
&= P(-1.44 < Z < 0.866) \\
&= \Phi(0.866) - \Phi(-1.44) \\
&= 0.8067549 - 0.0749337 = 0.732
\end{aligned}$$

Normal approximation to Poisson

Recall that $X \sim \text{Pois}(\lambda)$ has these properties

- $P(X = x) = e^{-\lambda}\lambda^x/x!$
- $E(X) = \lambda$
- $Var(X) = \lambda$

Normal approximation to Poisson

Recall that $X \sim \mathrm{Pois}(\lambda)$ has these properties

- $P(X = x) = e^{-\lambda}\lambda^x / x!$
- $E(X) = \lambda$
- $Var(X) = \lambda$

For large λ (say $\lambda = 100$), try computing

- $e^{-\lambda}$
- 100^{100}

Normal approximation to Poisson

Recall that $X \sim \text{Pois}(\lambda)$ has these properties

- $P(X = x) = e^{-\lambda}\lambda^x/x!$
- $E(X) = \lambda$
- $Var(X) = \lambda$

For large λ (say $\lambda = 100$), try computing

- $e^{-\lambda}$
- 100^{100}

This can be problematic if computations not handled carefully (just like the binomial case).

Normal approximation to Poisson

When `(\lambda > 15)`, it is appropriate to use the normal approximation to the Poisson

As $n \to \infty, \mathrm{Pois}(\lambda) \to N\left(\lambda, \lambda\right)$.

Normal approximation to Poisson

When `(\lambda > 15)`, it is appropriate to use the normal approximation to the Poisson

As $n \to \infty$, $\mathrm{Pois}(\lambda) \to N\big(\lambda, \lambda\big)$. This means that for $X \sim \mathrm{Pois}(\lambda)$,

$$
\begin{aligned}
P(X < x) &\approx P\left(\frac{X-\lambda}{\sqrt{\lambda}} < \frac{x-\lambda}{\sqrt{\lambda}}\right) \\
&= \Phi\left(\frac{x-\lambda}{\sqrt{\lambda}}\right)
\end{aligned}
$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. But also don't forget *continuity correction*!

Normal approximation to Poisson

Example 6.10: A radioactive disintegration gives counts that follow a Poisson distribution with a mean count of 25 per second. Find the probability that in a one-second interval the count is between 23 and 27 inclusive.

Normal approximation to Poisson

Example 6.10: A radioactive disintegration gives counts that follow a Poisson distribution with a mean count of 25 per second. Find the probability that in a one-second interval the count is between 23 and 27 inclusive.

Denote the radioactive disintegration counts by X. Then $X \sim \mathrm{Pois}(25)$. We are interested in $P(23 \leq X \leq 27)$. Since $\lambda > 15$, we can apply a normal approximation.

Normal approximation to Poisson

Example 6.10: A radioactive disintegration gives counts that follow a Poisson distribution with a mean count of 25 per second. Find the probability that in a one-second interval the count is between 23 and 27 inclusive.

Denote the radioactive disintegration counts by X. Then $X \sim \mathrm{Pois}(25)$. We are interested in $P(23 \leq X \leq 27)$. Since $\lambda > 15$, we can apply a normal approximation.

$$\begin{aligned} P(23 \leq X \leq 27) &\approx P(22.5 < X < 27.5) \\ &= P\left(\frac{22.5 - 25}{\sqrt{25}} < Z < \frac{27.5 - 25}{\sqrt{25}}\right) \\ &= P(-0.5 < Z < 0.5) \\ &= 2\Phi(0.5) - 1 \\ &= 2(0.6914625) - 1 = 0.383 \end{aligned}$$

Special continuous distributions: The exponential distributions

Exponential distribution

Suppose you are interested in

- The amount of time until an earthquake occurs.
- The time between two lightbulbs failing.
- The length (in minutes) of faculty staff meetings at UBD.
- The average waiting time at a hospital's A&E.

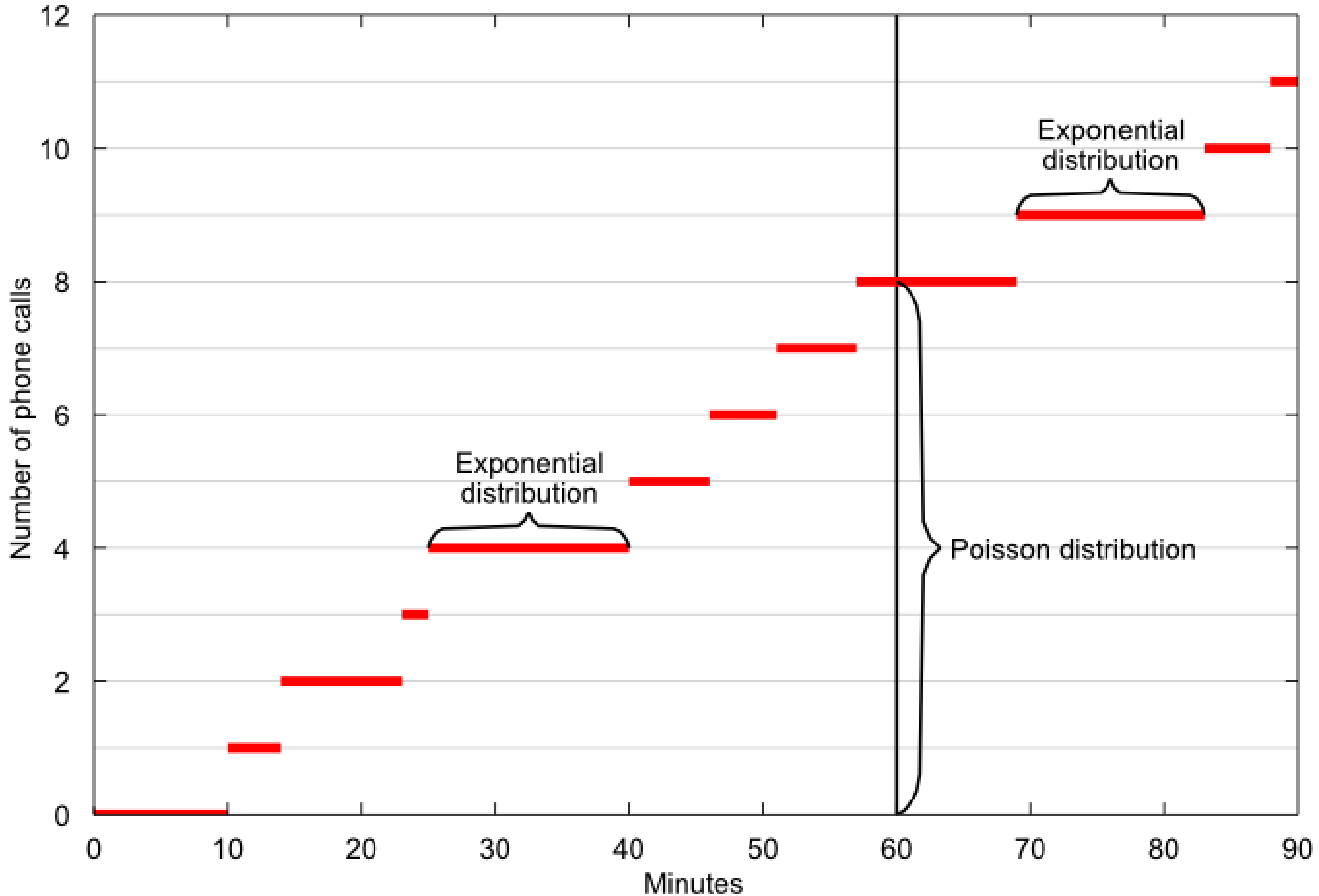
Exponential distribution

Suppose you are interested in

- The amount of time until an earthquake occurs.
- The time between two lightbulbs failing.
- The length (in minutes) of faculty staff meetings at UBD.
- The average waiting time at a hospital's A&E.

The exponential distribution is the probability distribution of the time between events in a Poisson point process, i.e., a process in which events occur continuously and independently at a constant average rate.

Exponential distribution



Exponential distribution

Let X be disributed according to an exponential distribution with rate λ. We write $X \sim \mathrm{Exp}(\lambda)$. Then its pdf is

$$f(x) = \lambda e^{-\lambda x}$$

Exponential distribution

Let X be disributed according to an exponential distribution with rate λ. We write $X \sim \mathrm{Exp}(\lambda)$. Then its pdf is

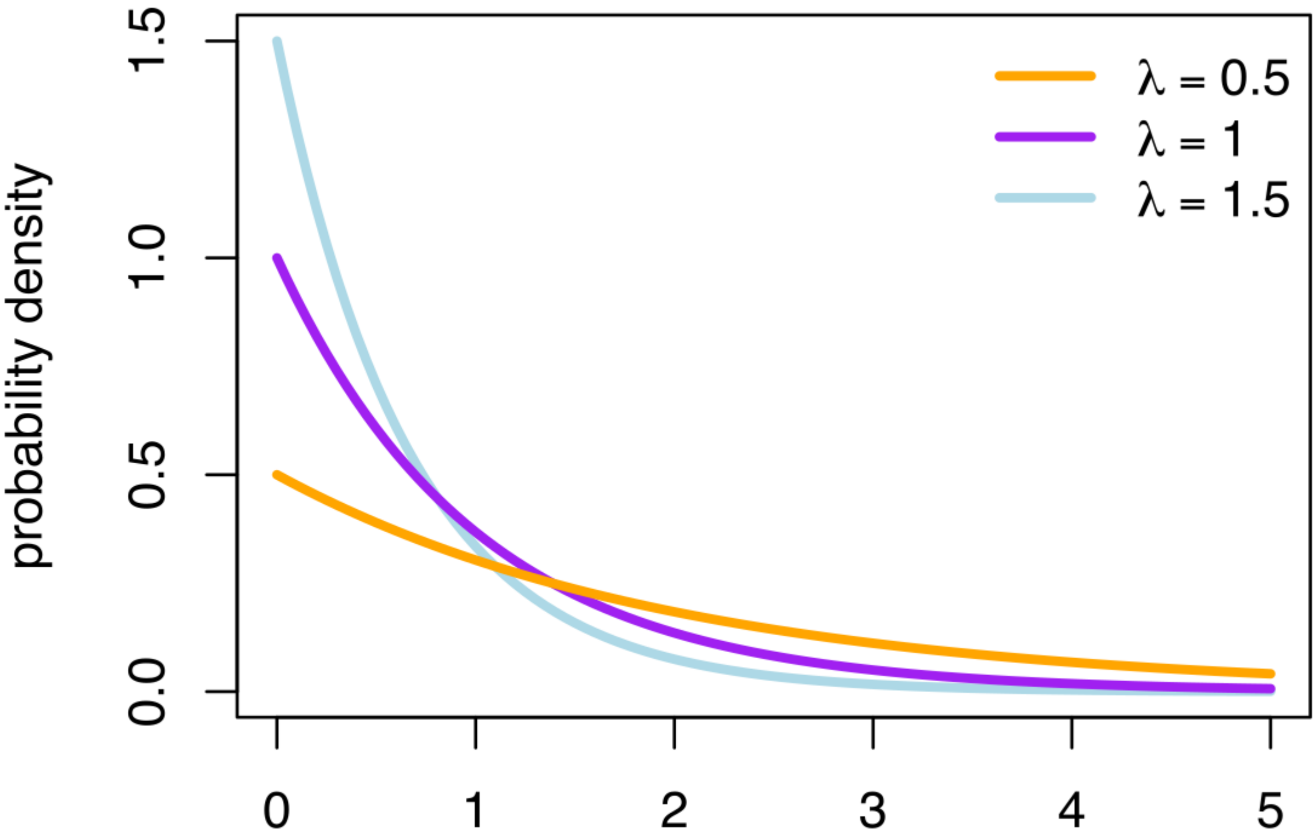
$$f(x) = \lambda e^{-\lambda x}$$

<u>Properties</u>

- The parameter of the exponential distribution is the rate $\lambda > 0$.
- Support: $x > 0$.
- $E(X) = 1/\lambda$.
- $Var(X) = 1/\lambda^2$.

Exponential distribution

The pdf experiences "exponential decay"— long wait times between two events occuring becomes more and more unlikely.



Exponential distribution

Example: Suppose $\lambda = 3$*, find* $P(0.5 < X < 1)$ *and* $P(X < 6)$*.*

Exponential distribution

Example: Suppose $\lambda = 3$, *find* $P(0.5 < X < 1)$ *and* $P(X < 6)$.

$$\begin{aligned} P(0.5 < X < 1) &= \int_{0.5}^{1} 3e^{-3x} dx \\ &= \left[-e^{-3x}\right]_{0.5}^{1} \\ &= e^{-1.5} - e^{-3} = 0.173 \end{aligned}$$

Exponential distribution

Example: Suppose $\lambda = 3$, find $P(0.5 < X < 1)$ and $P(X < 6)$.

$$\begin{aligned} P(0.5 < X < 1) &= \int_{0.5}^{1} 3e^{-3x} dx \\ &= [-e^{-3x}]_{0.5}^{1} \\ &= e^{-1.5} - e^{-3} = 0.173 \end{aligned}$$

$$\begin{aligned} P(X < 6) &= \int_{0}^{6} 3e^{-3x} dx \\ &= [-e^{-3x}]_{0}^{6} \\ &= 1 - e^{-18} \approx 1 \end{aligned}$$

Memoryless property

The exponential distribution has what is called the "memoryless" property. Given that you know the last event occured in x amount of time, the next occurence is independent of this.

Memoryless property

The exponential distribution has what is called the "memoryless" property. Given that you know the last event occured in x amount of time, the next occurence is independent of this.

$$P(X > s + t | X > s) = P(X > t)$$

Memoryless property

The exponential distribution has what is called the "memoryless" property. Given that you know the last event occured in x amount of time, the next occurence is independent of this.

$$P(X > s + t | X > s) = P(X > t)$$

For example, assume that bus waiting times are exponentially distributed. A memoryless wait for a bus would mean that the probability that a bus arrived in the next minute is the same whether you just got to the station or if you've been sitting there for twenty minutes already.

https://perplex.city/memorylessness-at-the-bus-stop-f2c97c59e420

END