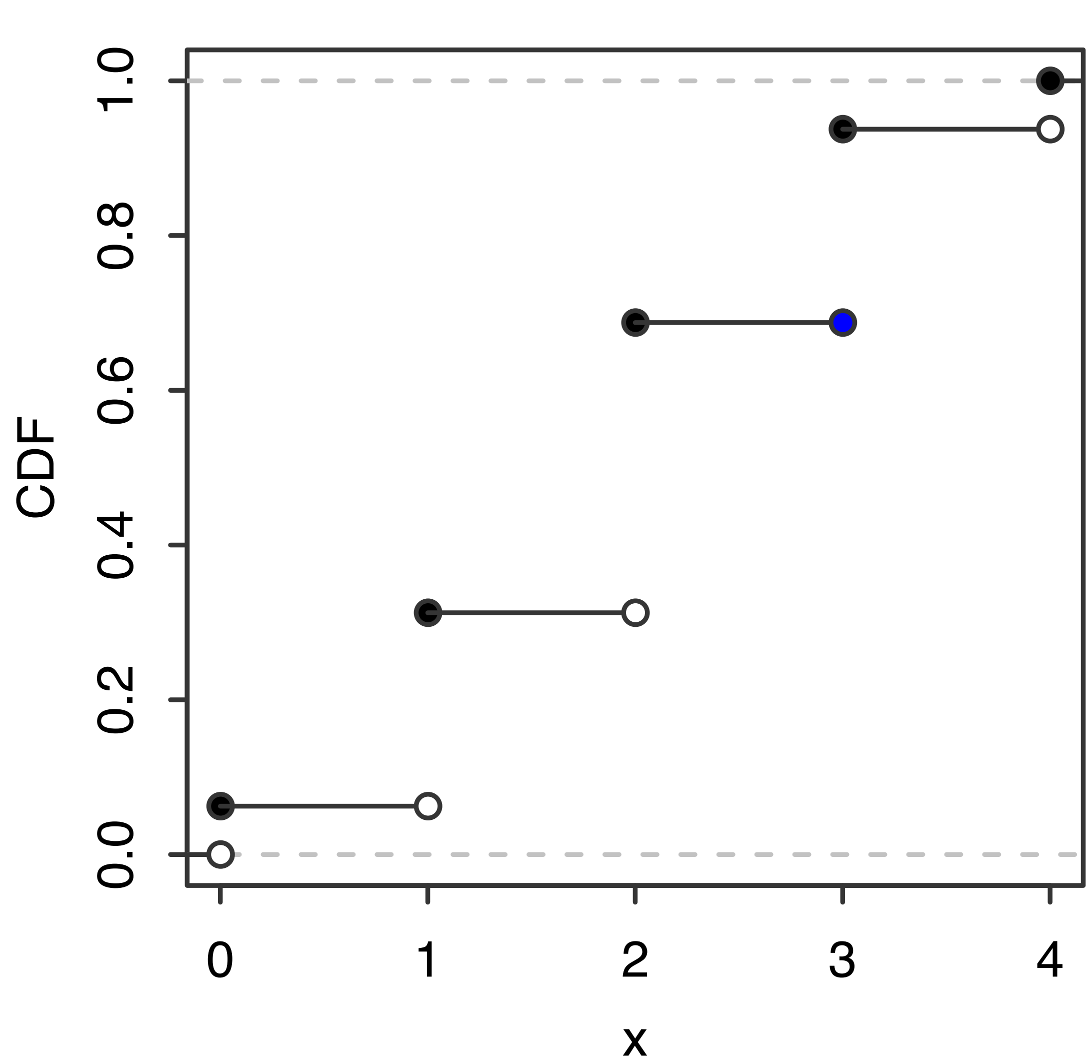
Lecture 7:

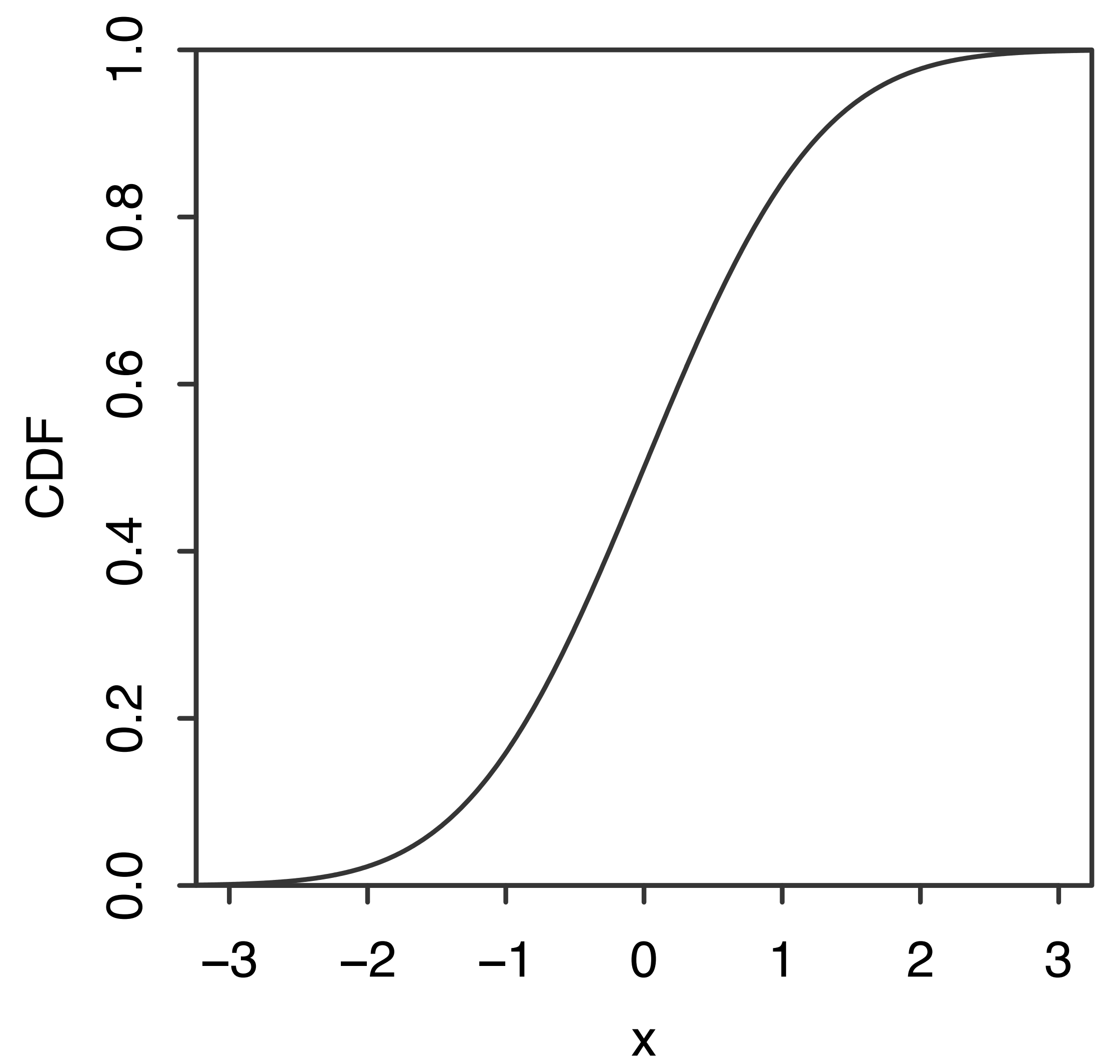
Continuous random variables

Probability density functions (PDFs)

Probability density functions (PDFs)

For a discrete r.v., the CDF "jumps" at every point in the support, and is flat elsewhere. For a **continuous** r.v. the CDF increases **smoothly**:





Probability density functions (PDFs)

Definition 5.1.1 (Continuous r.v.). An r.v. has a **continuous distribution** if its CDF is differentiable. We allow for finitely many points (including endpoints) where the CDF is continuous, but not differentiable. A **continuous r.v.** is an r.v. with such distribution.

For discrete r.v.-s, CDF is not diff.-ble in jumps. For continuous:

Definition 5.1.2 (Probability density function). For a continuous r.v. X with CDF F, the **probability density function (PDF)** of X is the derivative of the CDF, given by $f(x) = F'(x)$. The **support** of X, and its distribution, is the set of all x where $f(x) > 0$.

PDF $f(x)$ is **not** the probability that $X = x$ (in fact, for continuous r.v. $P(X = x) = 0$), but rather the **density of probability**

Probability density functions (PDFs)

Proposition 5.1.3 (PDF to CDF). Let X – continuous r.v. with PDF f.

Then the CDF of X is $F(x) = \displaystyle\int_{-\infty}^{x} f(t)\,dt$

The probability of X falling into an interval (a, b),

$$P(a \leq X \leq b) = F(b) - F(a) = \int_{a}^{b} f(x)\,dx,$$

where the endpoints can be included or excluded: $<$ / $>$ $\leftrightarrow$ $\leq$ / $\geq$

So, for an arbitrary region A,

$$P(X \in A) = \int_{A} f(x)\,dx$$

Probability density functions (PDFs)

Theorem 5.1.5 (Valid PDFs). PDF f of a continuous r.v. must satisfy:

1) Non-negativity: $f(x) \geq 0$

2) Normalisation (integrates to 1): $\displaystyle\int_{-\infty}^{+\infty} f(x)\, dx = 1$

Conversely, any such f **is** the PDF of some r.v.

Example 5.1.6 (Logistic). Logistic distribution has CDF

$F(x) = \dfrac{e^x}{1+e^x}$, $x \in \mathbb{R}$. So the PDF: $f(x) = F'(x) = \dfrac{e^x}{(1+e^x)^2}$.

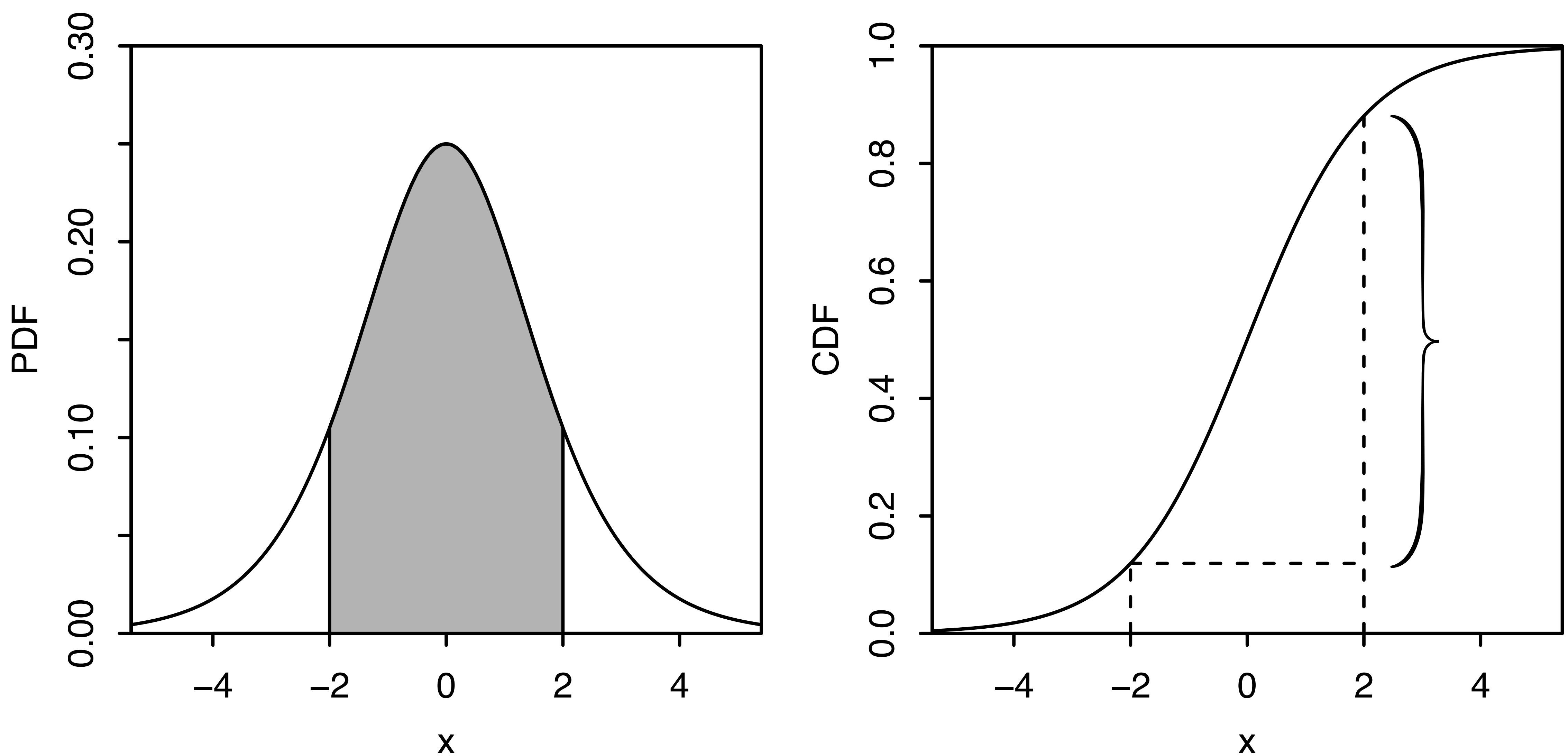
Let's find: $P(-2 < X < 2) =$

$$= \int_{-2}^{2} \frac{e^x}{(1+e^x)^2}\, dx = \int_{1+e^{-2}}^{1+e^2} \frac{1}{u^2}\, du = \left(-\frac{1}{u}\right)\Big|_{1+e^{-2}}^{1+e^2} \approx 0.76$$

Probability density functions (PDFs)

Example 5.1.6 (Logistic). Logistic distribution has CDF

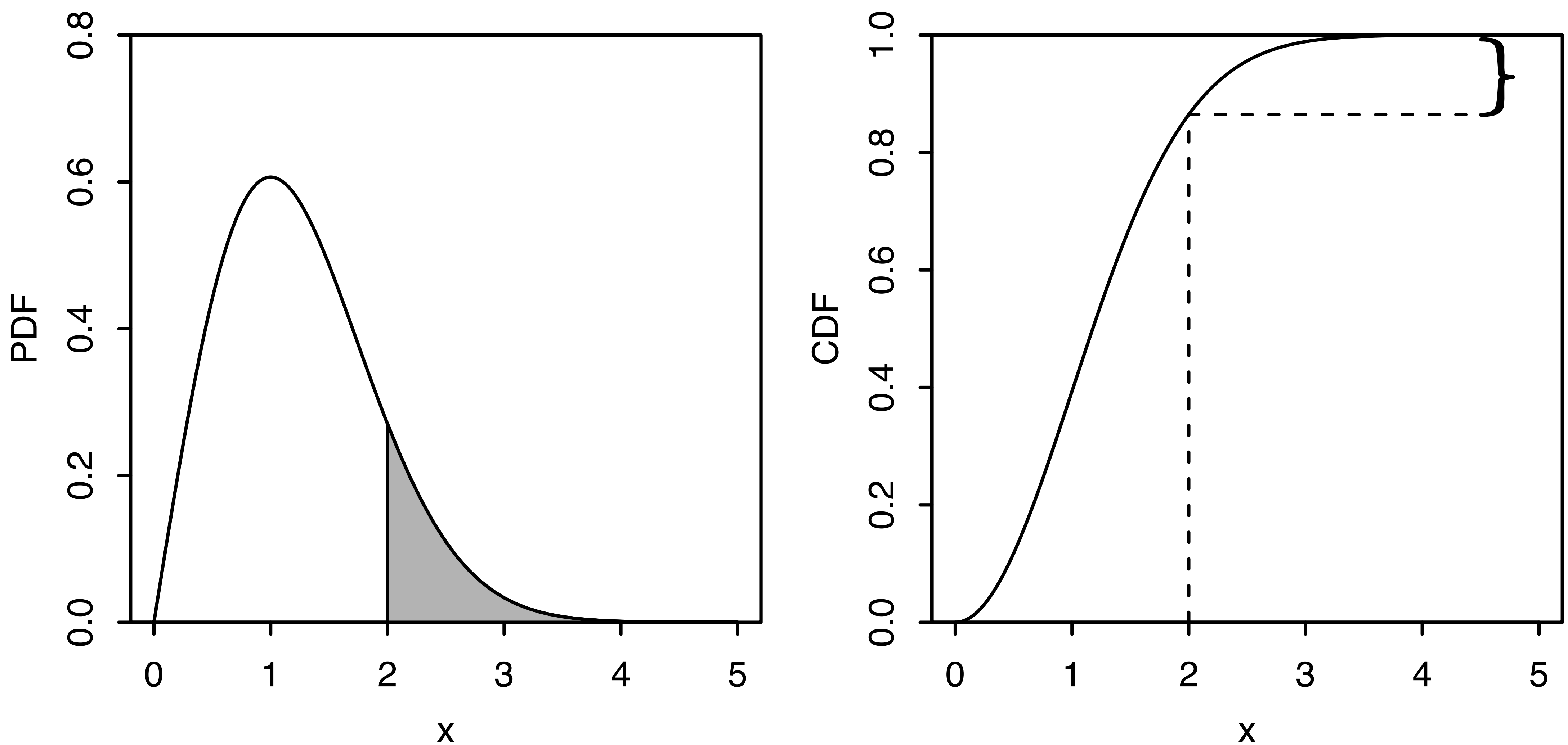
$F(x) = \dfrac{e^x}{1+e^x}$, PDF: $f(x) = \dfrac{e^x}{(1+e^x)^2}$, $P(-2 < X < 2) = 0.76$



Probability density functions (PDFs)

Example 5.1.7 (Rayleigh). The Rayleigh distribution has CDF
$F(x) = 1 - e^{-x^2/2}$, $x > 0$. PDF: $f(x) = xe^{-x^2/2}$, $x > 0$.

Let's find $P(X > 2) = \int_2^\infty xe^{-x^2/2}\,dx = 1 - F(2) \approx 0.14$



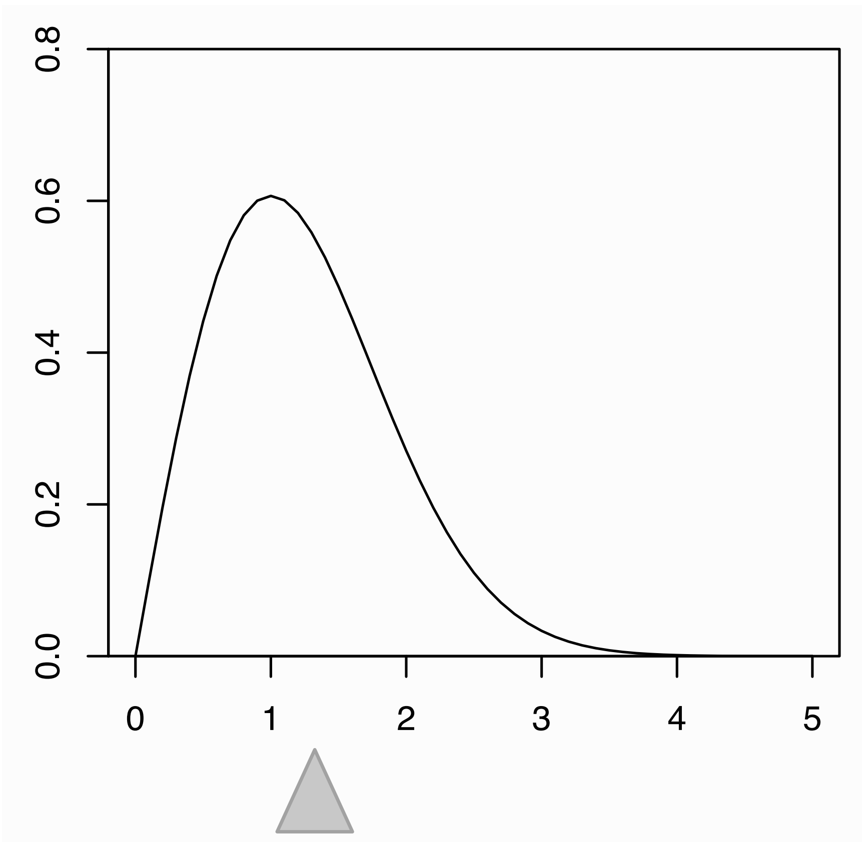
Probability density functions (PDFs)

Definition 5.1.10 (Expectation of a continuous r.v.) **Expected value** (**expectation** or **mean**) of a continuous r.v. is

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx.$$

As in the discrete case, it may not exist. If the r.v. X is dimensional – has units (centimetres, for example) – then $E(X)$ has the same units, thus $f(x)$ has inverse units (i.e. cm^{-1}, so cm $= \underset{x}{\text{cm}} \cdot \underset{f(x)}{\text{cm}^{-1}} \cdot \underset{dx}{\text{cm}}$).

Probability density functions (PDFs)



← expectation can be thought of as **centre-of-mass** of the PDF

here it is for Rayleigh distribution

Theorem 5.1.11 (LOTUS, continuous). If X is a continuous r.v. with PDF f and g is a function $g : \mathbb{R} \to \mathbb{R}$, then

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) f(x)\, dx$$

Uniform distribution

Uniform distribution

Definition 5.2.1 (Uniform distribution). A continuous r.v. U has **uniform distribution** (denote $U \sim$ Unif(a, b)) on (a, b) if its PDF is

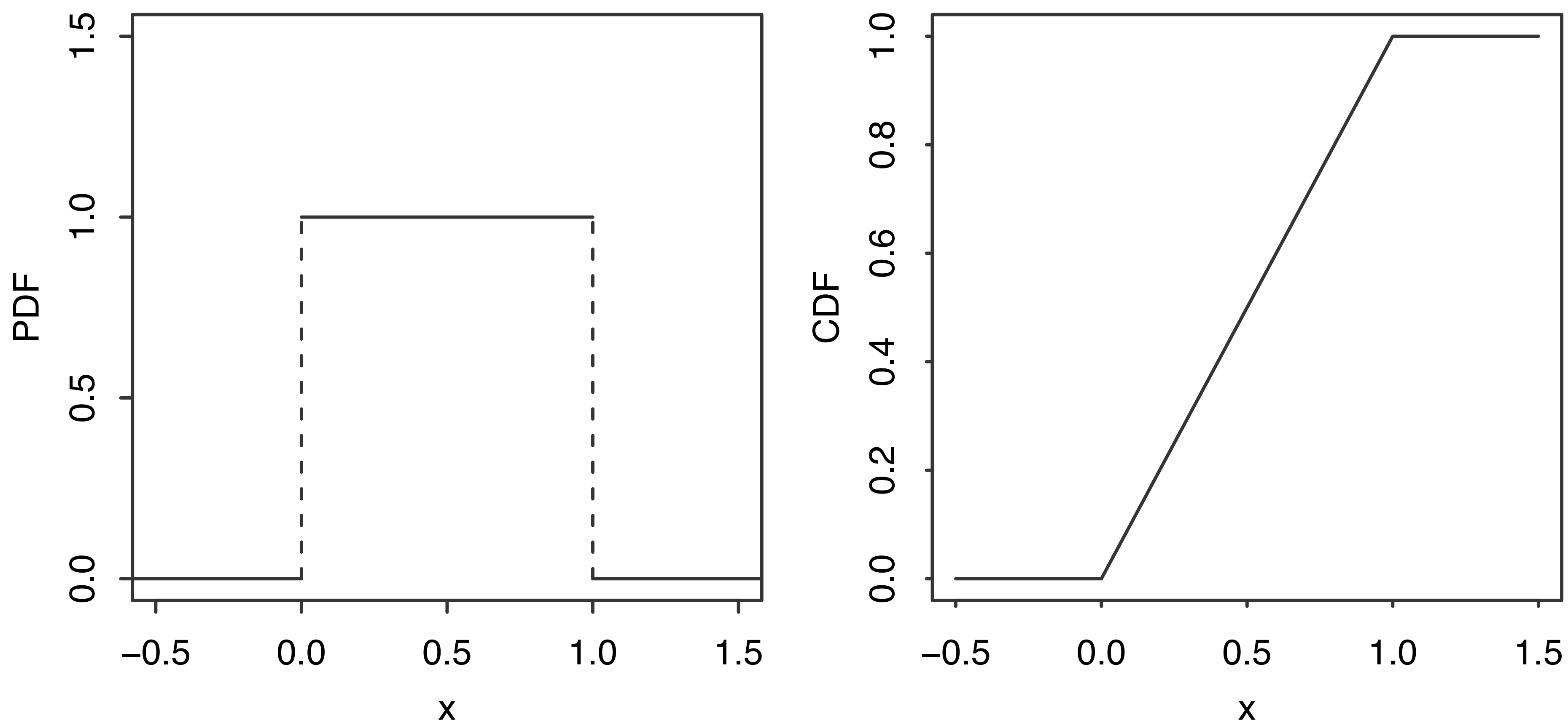
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

This is a valid PDF, since area under the rectangle of width $b - a$ and height $1/(b - a)$ is 1.

The CDF is $F(x) = \begin{cases} 0 & \text{if } x \leq a, \\ \frac{x-a}{b-a} & \text{if } a < x < b, \\ 1 & \text{if } x \geq b \end{cases}$

We will often refer to **standard uniform** – $U \sim U(0,1)$

Uniform distribution



Proposition 5.2.2 Let $U \sim \text{Unif}(a, b)$ and $(c, d) \subset (a, b)$. Then $P(x \in (c, d)) = \text{length}((c, d)) = d - c$

Proposition 5.2.3 Let $U \sim \text{Unif}(a, b)$, $(c, d) \subset (a, b)$ and $u \in (c, d)$ Then $P(U \leq u \,|\, U \in (c, d)) = \dfrac{u - c}{d - c}$

Uniform distribution

Let's find the **mean** and the **variance** of $U \sim$ Unif(a, b).

$E(U) = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a}\left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{a+b}{2}$ is quite intuitive,

For the variance, first find $E(U^2)$ with LOTUS: $E(U^2) =$

$$= \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{3} \cdot \frac{b^3 - a^3}{b-a}, \quad b^3 - a^3 = (b-a)(a^2 + ab + b^2)$$

so Var$(U) = E(U^2) - (EU)^2 = \frac{b^3 - a^3}{3(b-a)} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$.

One can do that simpler with use of:

Definition 5.2.5 (Location-scale transformation). For X – r.v.,

$Y = \sigma X + \mu$ is its loc-scale transform, with μ – loc and σ – scale

(which are some constants, $\sigma > 0$)

Uniform distribution

For $X \sim$ Unif(a, b), its loc-scale transform $Y = cX + d$ ($c > 0$) is (just a linear function of X) ***also uniform:*** $Y \sim$ Unif$(ca + d, cb + d)$!

$U \sim$ Unif$(0,1)$: $E(U) = \frac{1}{2}$, $E(U^2) = \int_0^1 x^2 dx = \frac{1}{3}$, Var$(U) = \frac{1}{12}$

So $\tilde{U} \sim$ Unif(a, b) is a loc-scale transform of U: $\tilde{U} = a + (b - a)U$

Thus:

$$E(\tilde{U}) = E(a + (b - a)U) = a + (b - a)E(U) = a + \frac{b - a}{2} = \frac{a + b}{2}$$

and $\text{Var}(\tilde{U}) = \text{Var}(a + (b - a)U) = \text{Var}((b - a)U) =$

$$= (b - a)^2 \text{Var}(U) = \frac{(b - a)^2}{12}$$

Uniform distribution

Loc-scale transform is just a linear transform of an r.v.

In fact, if one has $U \sim$ Unif$(0,1)$ – one can construct an r.v. $X = g(U)$ – some function of U – having ***any desired continuous distribution,*** $X \sim f(X)$

(If you have a uniform "coin/dice" – you can generate ***any*** r.v.!)

This fact is referred to as **Universality of the Uniform distribution**, it will be covered in detail on the seminar.

The Normal distribution

The Normal distribution

The Normal is a very famous distribution, appearing all the time in Statistics & Probability theory due to the ***central limit theorem*** – sum of a large number of i.i.d. r.v.-s has an approx. Normal distribution ***regardless*** of their individual distributions.

Definition 5.4.1 (Standard Normal distribution). A continuous r.v. Z has **standard Normal distribution** if its PDF is

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < +\infty$$

we write this as $Z \sim \mathcal{N}(0,1)$, since (we'll show that) Z has mean 0 and variance 1.

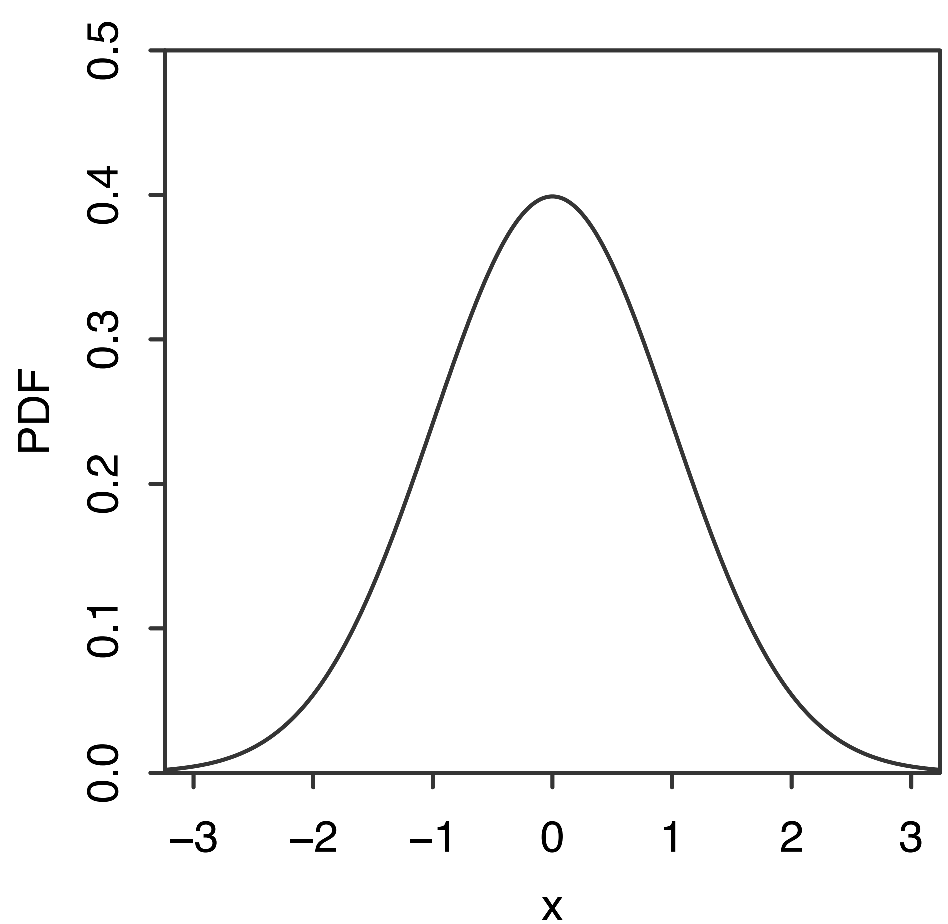
The **normalising constant** $1/\sqrt{2\pi}$ makes sure $\varphi(z)$ integrates to 1.

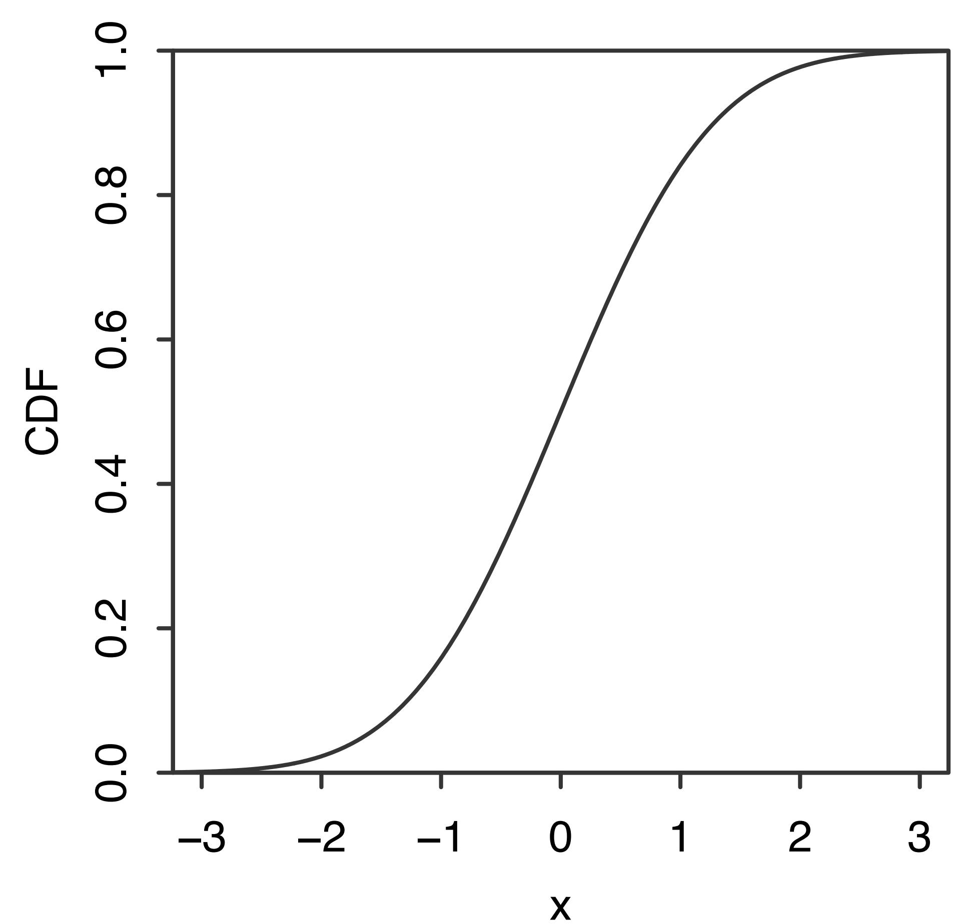
The Normal distribution

CDF of standard Normal $\Phi(z) = \int_{-\infty}^{z} \varphi(t)\,dt = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,dt$

can not be expressed in terms of elementary functions (polynomials, exponentials, sines/cosines, log-s, etc).

$\varphi(z)$, $\Phi(z)$ are standard letters for PDF and CDF of standard Normal





The Normal distribution

Several important properties of the standard Normal PDF and CDF:

1) **Symmetry of PDF**: $\varphi(z) = \varphi(-z)$ – even function

2) **Symmetry of tail areas**: $P(Z \leq -z) = P(Z \geq z)$ – that is, $\Phi(z) = 1 - \Phi(-z)$

3) **Symmetry of Z and $-Z$:** if $Z \sim \mathcal{N}(0,1)$, then $-Z \sim \mathcal{N}(0,1)$

Finding the normalising constant, $1/\sqrt{2\pi}$, can be done so: write the integral $I = \int_{-\infty}^{+\infty} e^{-z^2/2}\, dz$ ***twice:*** $I^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}}\, dxdy.$

Now in ***polar coordinates*** $x^2 + y^2 = r^2$ and $dx\, dy = r\, dr\, d\theta$, so

$$\int_0^{2\pi} \int_0^{\infty} e^{-r^2/2}\, rdr\, d\theta = \int_0^{2\pi} \left(\int_0^{\infty} e^{-u}\, du \right) d\theta = \int_0^{2\pi} 1\, d\theta = 2\pi$$

The Normal distribution

Expectation of the standard Normal = 0 by symmetry of the PDF:

$$E(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z e^{-z^2/2}\, dz = 0$$ – same for any odd n : $E(Z^n) = 0$

Variance is $\text{Var}(Z) = E(Z^2) - (EZ)^2 = E(Z^2) =$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z^2 e^{-z^2/2} dz = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z^2 e^{-z^2/2} dz$$ (for even function)

Integration by parts $u = z,\ dv = z e^{-z^2/2}\, dz,\ \ du = dz,\ v = -e^{-z^2/2}$

$$\text{Var}(Z) = \frac{2}{\sqrt{2\pi}} \left(-z e^{-z^2/2} \Big|_0^{\infty} + \int_0^{\infty} e^{-z^2/2}\, dz \right) = \frac{2}{\sqrt{2\pi}} \left(0 + \frac{\sqrt{2\pi}}{2} \right) = 1$$

With that we're ready to introduce the general Normal distribution

The Normal distribution

Definition 5.4.3 (Normal distribution). If $Z \sim \mathcal{N}(0,1)$, then $X = \mu + \sigma Z$ has **Normal distribution** with mean μ and variance σ^2 – we denote this $X \sim \mathcal{N}(\mu, \sigma^2)$.

From linearity of expectation, $E(\mu + \sigma Z) = E(\mu) + \sigma E(Z) = \mu$ and from props. of Var: $\text{Var}(\mu + \sigma Z) = \text{Var}(\sigma Z) = \sigma^2 \text{Var}(Z) = \sigma^2$

If $X = \mu + \sigma Z$, then, of course, $Z = \frac{X - \mu}{\sigma}$ – **standardised version** of X. Taking $X \to \frac{X - \mu}{\sigma}$ is thus called **standardisation** of the r.v.

The Normal distribution

Theorem 5.4.4 (Normal PDF and CDF). For $X \sim \mathcal{N}(\mu, \sigma^2)$,

the PDF is $f(x) = \varphi\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma} = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

and the CDF is $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

Here's a very useful fact to know:

Theorem 5.4.5 (68-95-99 rule). If $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$P(|X-\mu| < \sigma) \approx 0.68$$
$$P(|X-\mu| < 2\sigma) \approx 0.95$$
$$P(|X-\mu| < 3\sigma) \approx 0.997$$