Continuous Random Variables

All RVs up to now have been discrete:

- Take on distinct values e.g. in set $\{1, 2, 3\}$
- Often represent binary values or counts

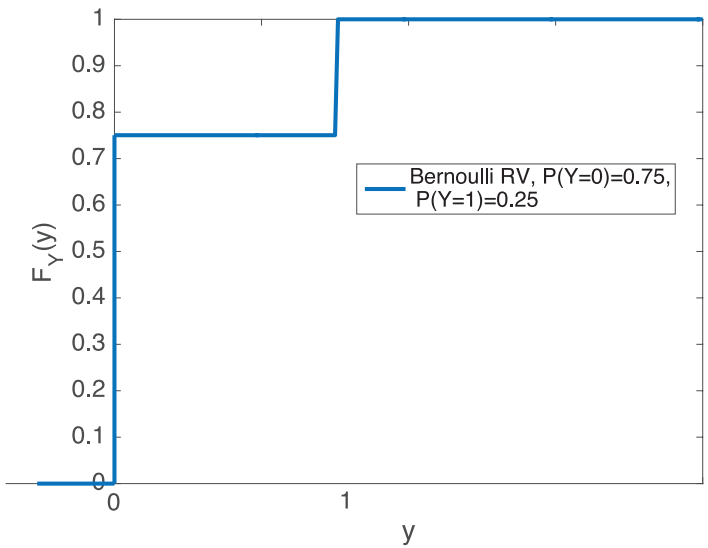
What about continuous RVs ?

- Take on real-values
- e.g. travel time to work, temperature of this room, fraction of Irish population supporting Scotland in the rugby

Cumulative Distribution Function

Suppose Y is a random variable, which may be discrete or continuous valued.

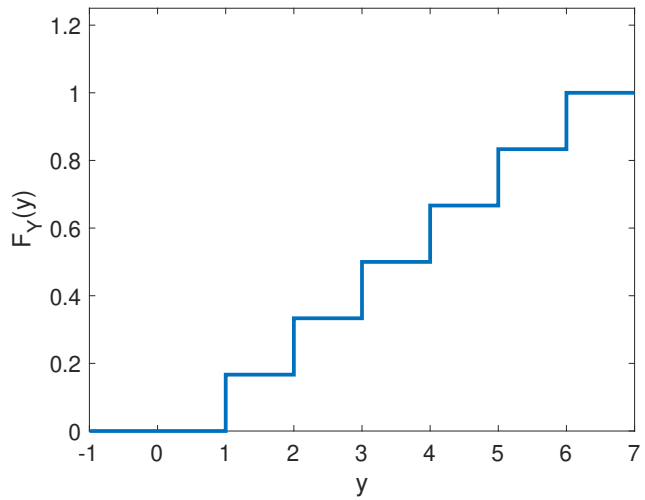
- $F_Y(y) := P(Y \leq y)$ is the **cumulative distribution function** (CDF).
- CDF exists and makes sense for both discrete and continuous valued random variables
- E.g. CDF of Bernoulli RV Y (so Y takes values 0 or 1):



- As y increases (i.e. as we move from left to right in plot) CDF either stays the same or increases, it never decreases
- $F_Y(-\infty) = 0$, $F_Y(+\infty) = 1$.

Cumulative Distribution Function

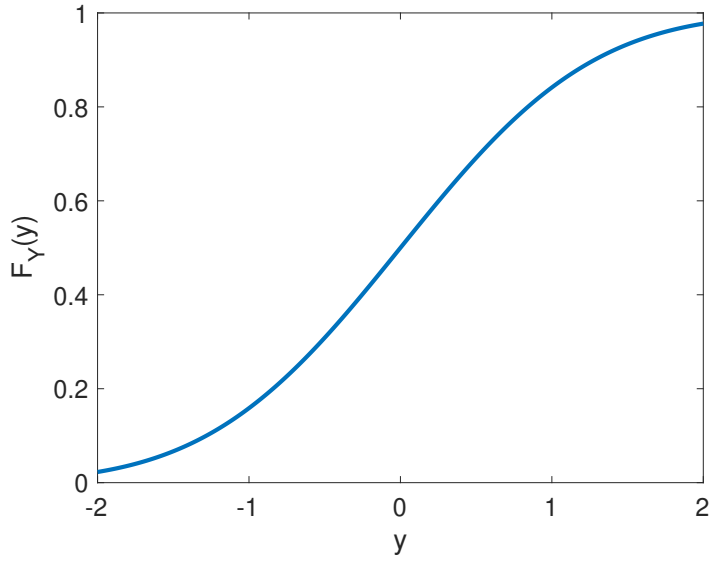
- E.g. Suppose Y is number that comes up when roll a six sided dice:



- When Y takes discrete values $\{y_1, \cdots, y_m\}$, then $F_Y(y) = \sum_{j:y_j \leq y} P(Y = y_j)$
- $F_Y(-\infty) = 0$, $F_Y(+\infty) = 1$.

Cumulative Distribution Function

CDF for continuous valued RV:



- Observe that CDF always starts at 0 and rises to 1
- CDF never decreases

Cumulative Distribution Function

Suppose Y is a random variable, which may be discrete or continuous valued.

$$P(Y \leq b) = P(Y \leq a) + P(a < Y \leq b)$$
$$i.e.\ F_Y(b) = F_Y(a) + P(a < Y \leq b)$$

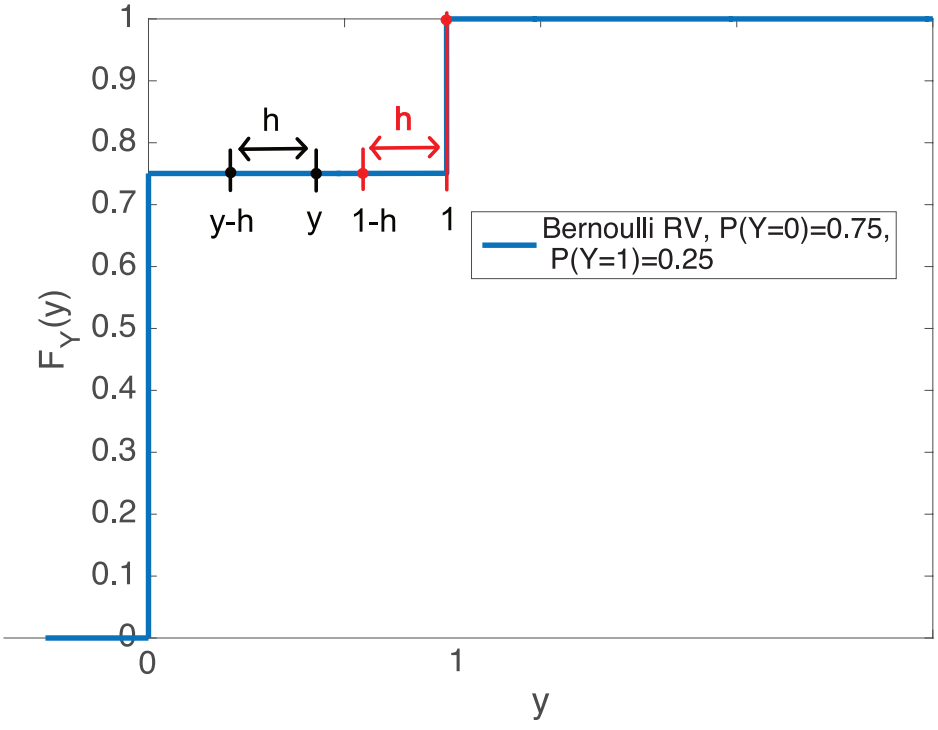
Therefore,

$$P(a < Y \leq b) = F_Y(b) - F_Y(a)$$

So given CDF of a random variable we can calculate the probability that it lies in any interval.

Discrete RVs: Recovering Probabilities From CDF

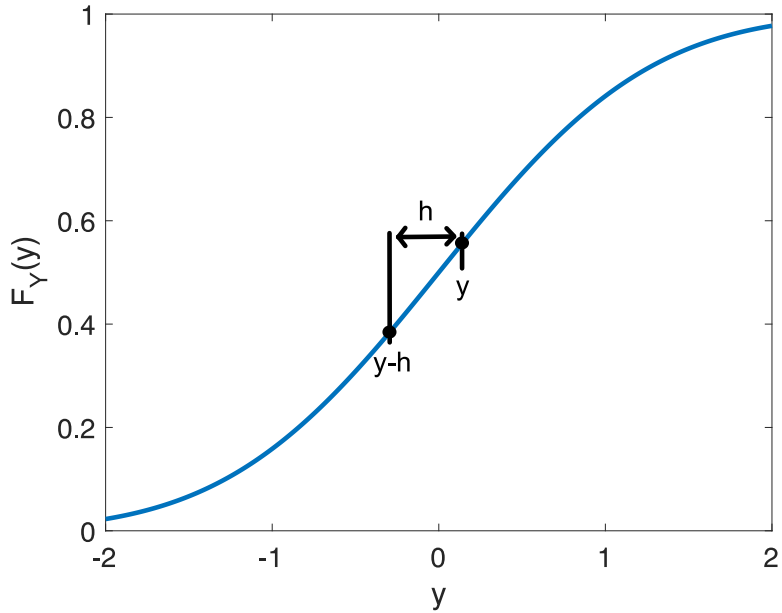
- E.g. Bernoulli RV:



- $P(y - h < Y \leq y) = F_Y(y) - F_Y(y - h)$
- What happens as $h \to 0$? Depends on whether discrete RV can take value y or not ...

Continuous Random Variables: CDF and PDF

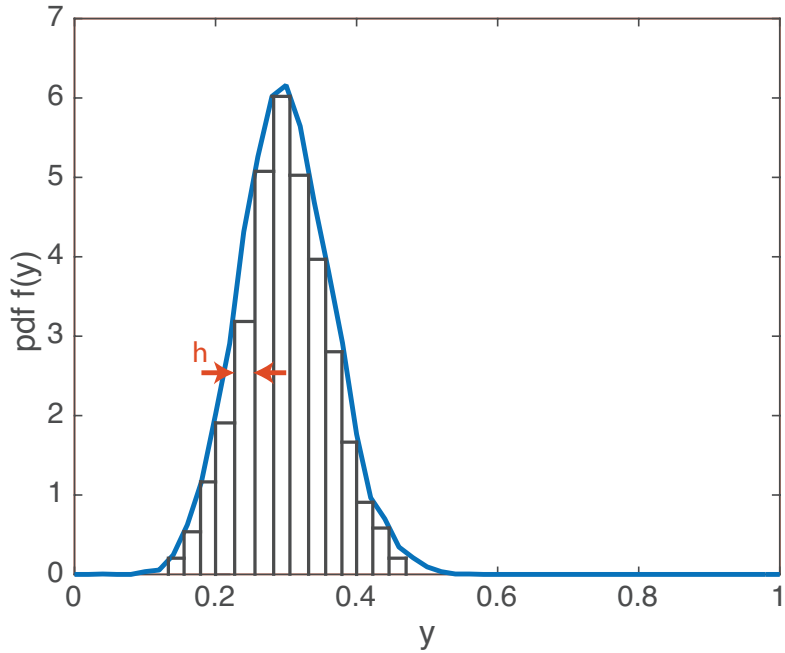
- E.g. Gaussian RV:



- $P(y - h < Y \leq y) = 0$ as $h \to 0$ i.e. $P(Y = y) = 0$
- But ... it turns out $\frac{P(y-h<Y\leq y)}{h} = \frac{F_Y(y)-F_Y(y-h)}{h}$ as $h \to 0$ is useful. Its called the **Probability Density Function** (PDF).
- PDF is **not** a probability, it can take values > 1. But it is probability-like i.e. shares many of the same properties as a probability.
- To link CDFs and PDFs properly we need to take a detour ...

How do we calculate the area under a curve ?

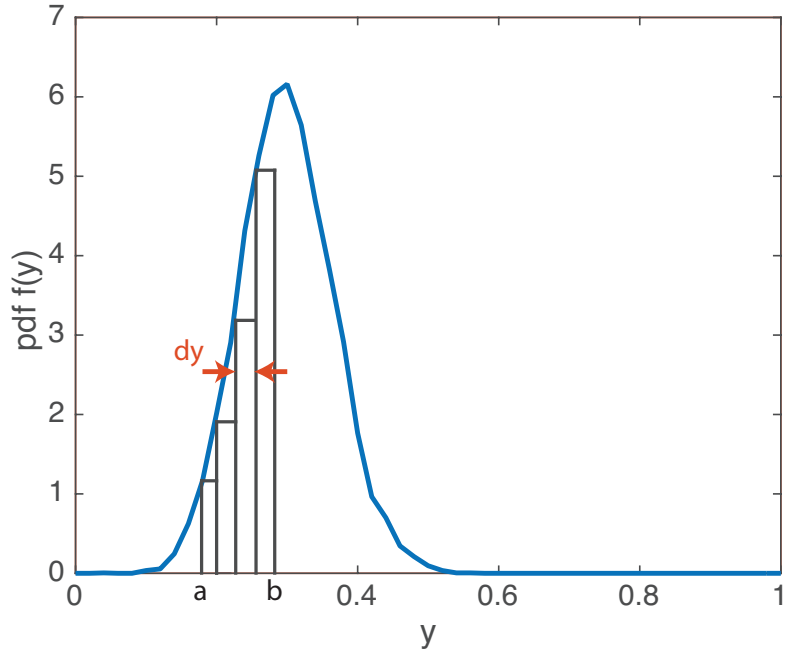
- Fit a series of rectangles under the curve, each of width h:



- We know the area under a rectangle, its the height$\times$width h
- Add up the areas of all the rectangles to get an estimate of the area under the curve
- As h gets smaller and smaller ($h \to 0$) this value becomes closer and closer to the true area[1]

[1] The maths needed to analyse this convergence is beyond this module, but if interested take a look at https://en.wikipedia.org/wiki/Riemann_integral and https://en.wikipedia.org/wiki/Lebesgue_integration.

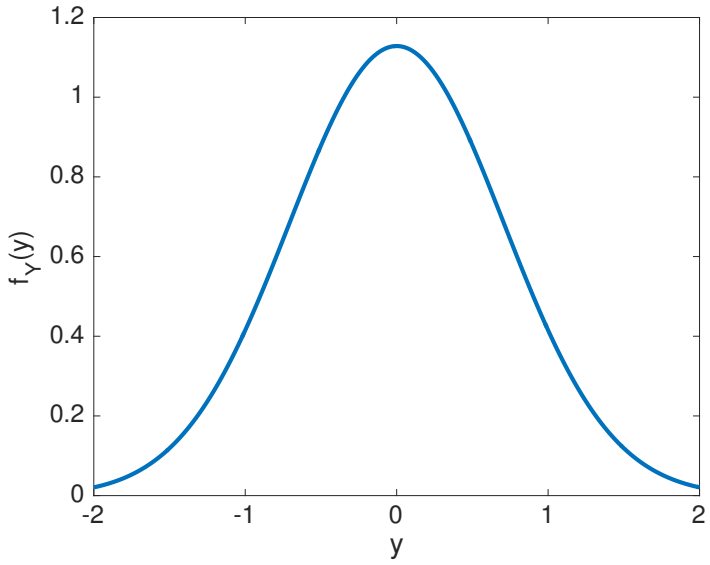
How do we calculate the area under a curve ?

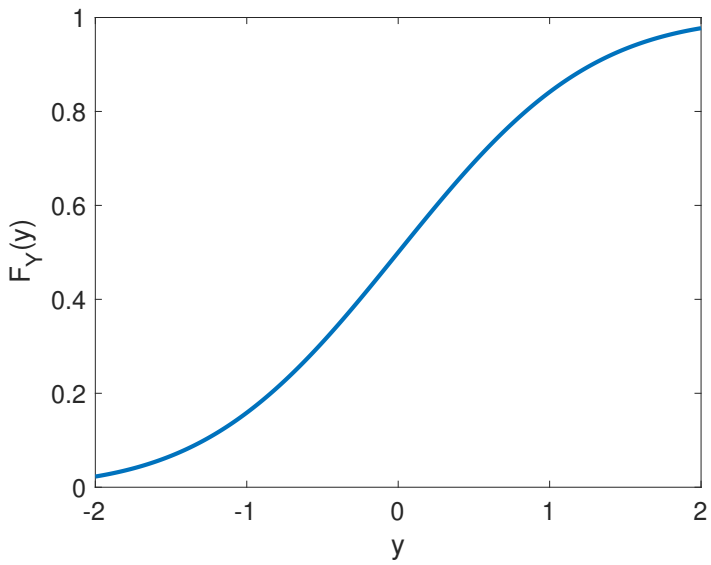


- Think of $f(y)dy$ as the area of the rectangle between y and $y + dy$ with dy infinitesimally small.
- Write the area under curve between a and b as $\int_a^b f(y)dy$
- Think of integral as the sum of areas of rectangles each of width h as $h \to 0$. Integral symbol $\int$ is supposed to be suggestive of a sum. Can think of dy as h (infinitesimally small).

How do we calculate the area under a curve ?

Example: CDF $F_Y(y)$ in right-hand plot is area under curve in left-hand plot between $-\infty$ and y i.e. $F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt$



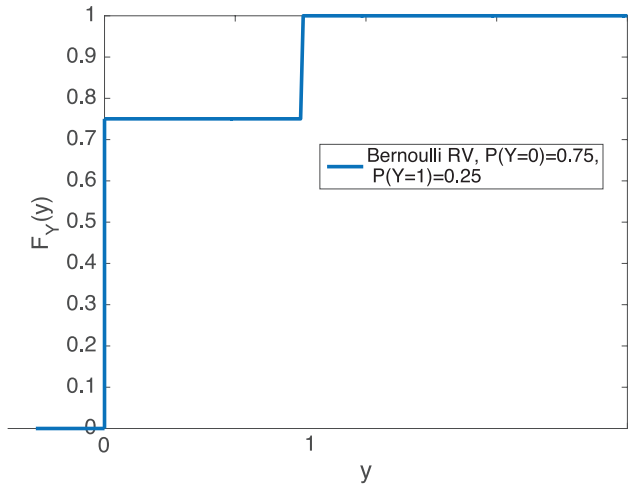


Continuous Random Variables: CDF and PDF

- For a continuous-valued random variable Y there exists a function $f_Y(y) \geq 0$ such that:

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt$$

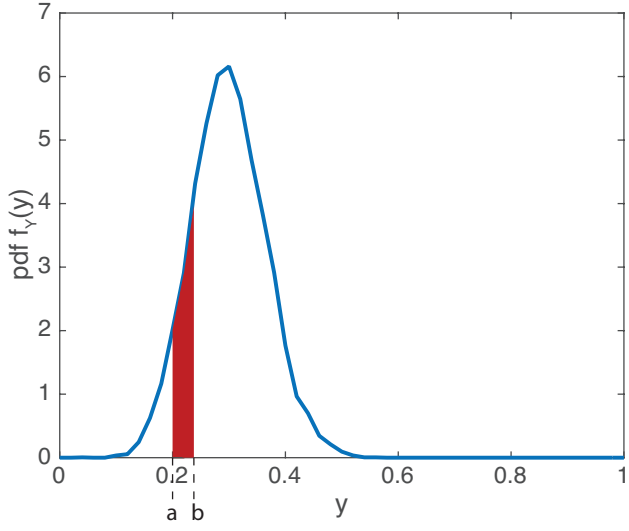
- f_Y is called the **probability density function** or **PDF** of Y.
- $\int_{-\infty}^{\infty} f(y)dy = 1$ (since $\int_{-\infty}^{\infty} f(y)dy = F_Y(\infty) = P(Y \leq \infty) = 1$)
- Note that tricky to define PDF f_Y for a discrete random variable since its CDF has "jumps" in it.



Continuous Random Variables: CDF and PDF

- It follows that

$$
\begin{aligned}
P(a < Y \leq b) \\
&= F_Y(b) - F_Y(a) \\
&= \int_{-\infty}^{b} f_Y(t)dt - \int_{-\infty}^{a} f_Y(t)dt \\
&= \int_{a}^{b} f_Y(t)dt
\end{aligned}
$$



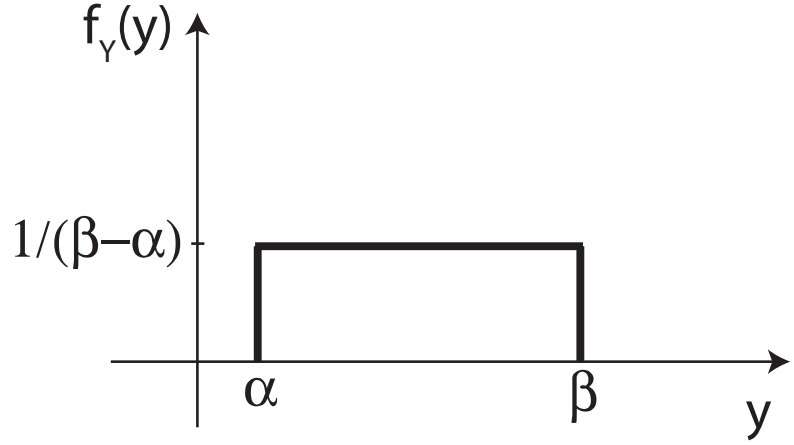
- The probability density function $f(y)$ for random variable Y is not a probability e.g. it can take values greater than 1.
- Its the area under the PDF between points a and b that is the probability $P(a < Y \leq b)$

Example: Uniform Random Variables

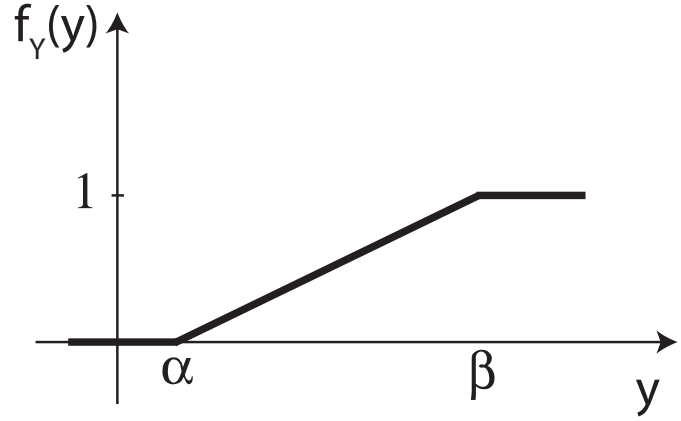
Y is a **uniform random variable** when it has PDF:

$$f_Y(y) = \begin{cases} \frac{1}{\beta-\alpha} & \textit{when } \alpha \leq y \leq \beta \\ 0 & \textit{otherwise} \end{cases}$$

PDF

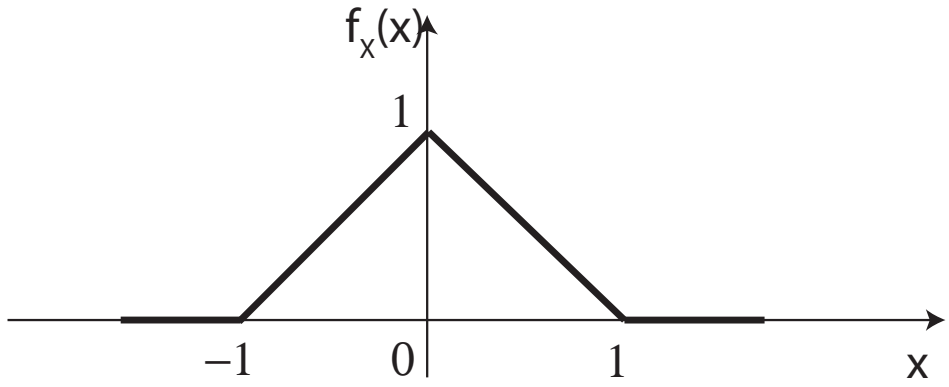


CDF



- For $\alpha \leq a \leq b \leq \beta$: $P(a \leq Y \leq b) = \frac{b-a}{\beta-\alpha}$
- rand() function in Matlab.

Example



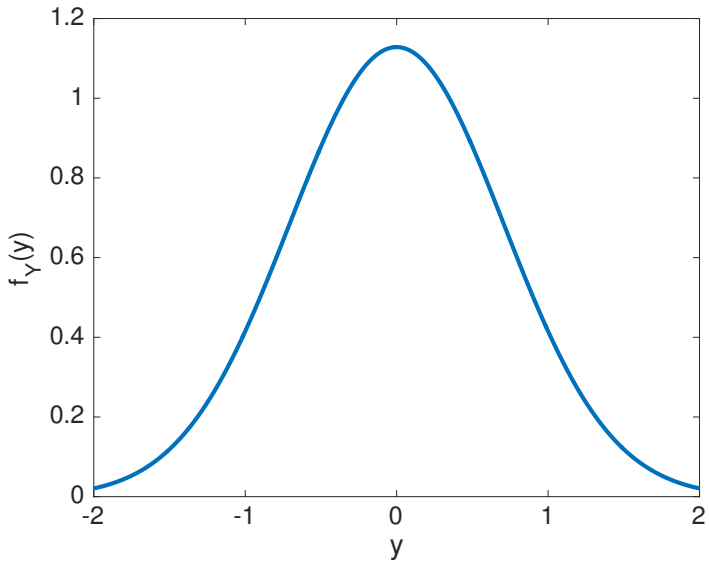
- Check the area under the PDF is 1. Area of left-hand triangle is 1/2, area of right-hand triangle same. Total is 1.
- What is $P(0 \leq X \leq 1)$? Its the area under the PDF between points 0 and 1 i.e. the area of the right-hand triangle. So $P(0 \leq X \leq 1) = 0.5$.
- What is $P(0 \leq X \leq \infty)$? $f_X(x) = 0$ for $x > 1$, so $P(0 \leq X \leq \infty) = P(0 \leq X \leq 1) = 0.5$

The Normal Distribution

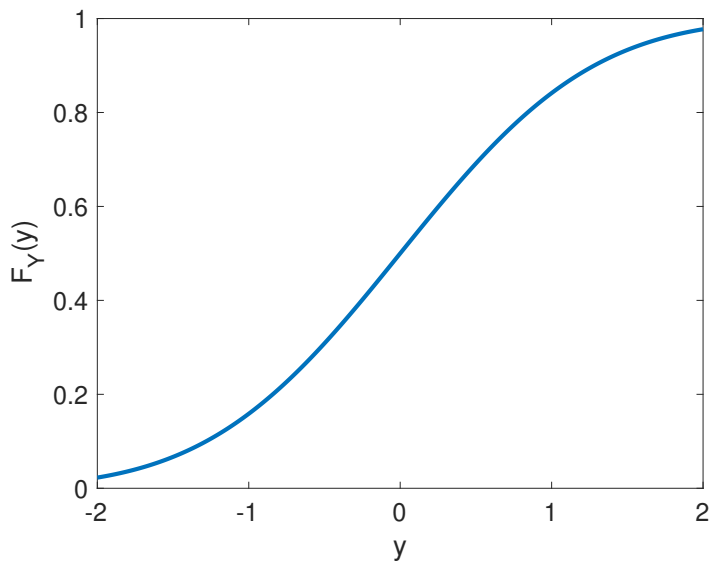
Y is a **Normal random variable** $Y \sim N(\mu, \sigma^2)$ when it has PDF:

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

PDF



CDF



$\mu = 0,\ \sigma = 1$

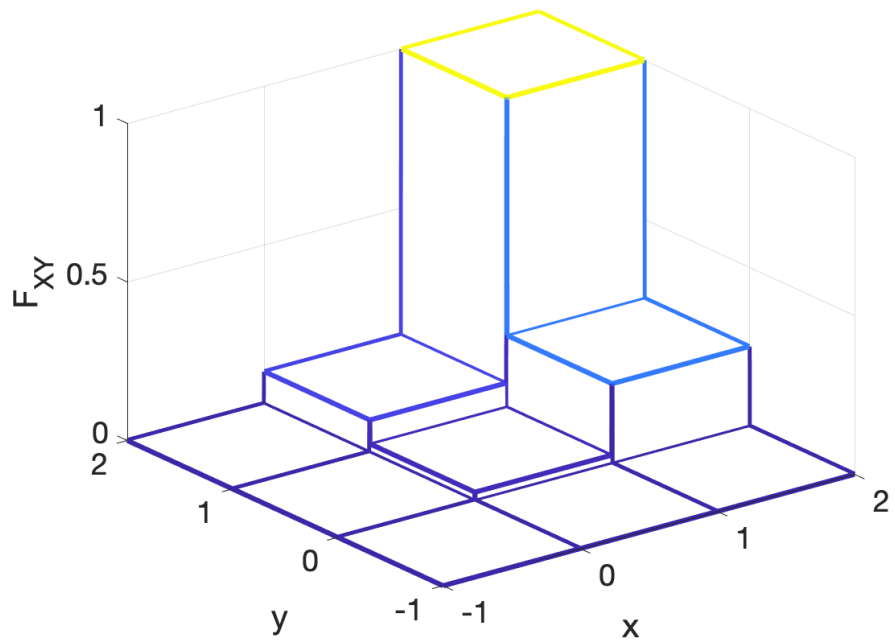
- $E[Y] = \mu,\ Var(Y) = \sigma^2$
- Symmetric about μ and defined for all real-valued x
- A Normal RV is also often called a **Gaussian random variable** and the Normal distribution referred to as the Gaussian distribution.

Joint Cumulative Distribution Function

Suppose X and Y are two random variables.

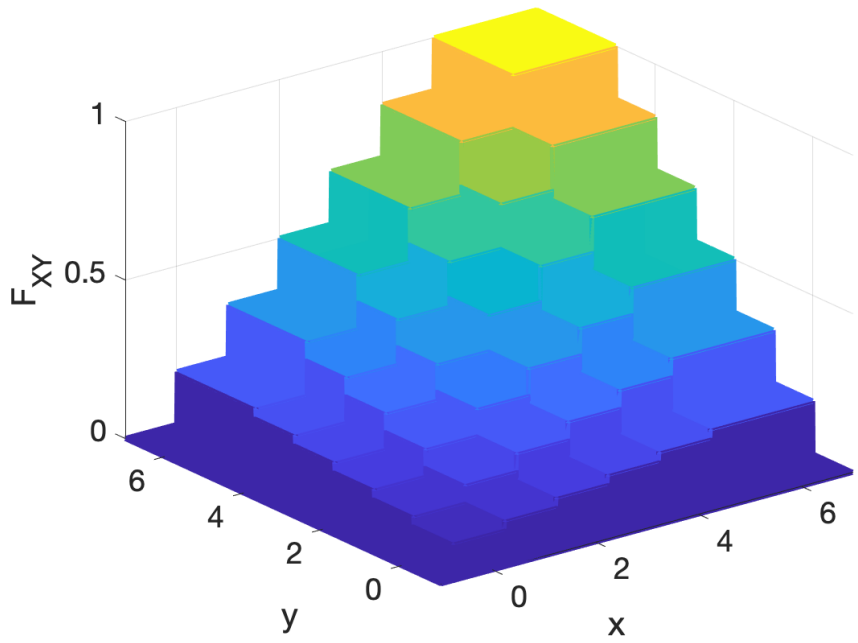
- $F_{XY}(x, y) = P(X \leq x \text{ and } Y \leq y)$ is the cumulative distribution function for X and Y

E.g. Suppose X and Y are independent Bernoulli RVs with $P(X = 0) = 0.1$, $P(X = 1) = 0.9$ and $P(Y = 0) = 0.25$, $P(Y = 1) = 0.75$. Since X and Y are independent $P(X \leq x \text{ and } Y \leq y) = P(X \leq x)P(Y \leq y)$.



Joint Cumulative Distribution Function

E.g. Suppose X and Y are independent rolls of a six-sided dice.



Joint Cumulative Distribution Function

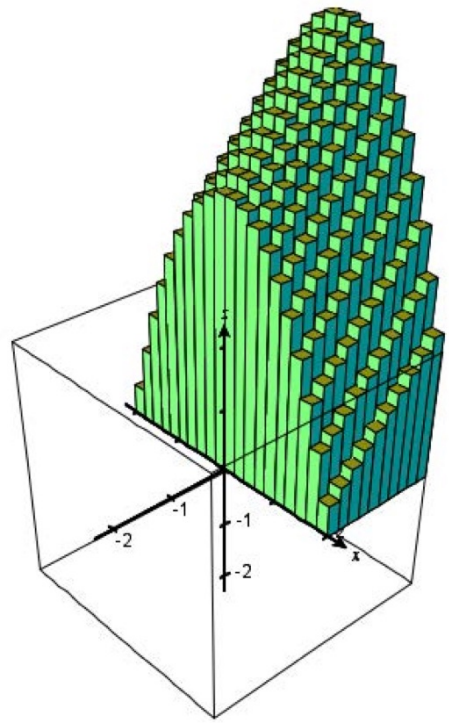
When X and Y are jointly continuous-valued random variables there exists a probability density function (PDF) $f_{XY}(x, y) \geq 0$ such that:

- $F_{XY}(x, y) = \int_{\infty}^{x} \int_{-\infty}^{y} f_{XY}(u, v) du\ dv$

Can think of

$P(u \leq X \leq u+du \text{ and } v \leq Y \leq v+dv) \approx f_{XY}$

when du, dv are infinitesimally small.



Using PDFs

PDFs are not probabilities, but we can repeat much of what we did before with discrete probabilities using PDFs and it all works provided we replace sums with integrals.

- Conditional PDFs: $f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$
- ... chain rule for PDFs:
  $f_{XY}(x,y) = f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x)$
- ... marginalising: $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$
- ... Bayes Rule for PDFs: $f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$
- Independence: X, Y are independent when $f_{XY}(x,y) = f_X(x)f_Y(y)$

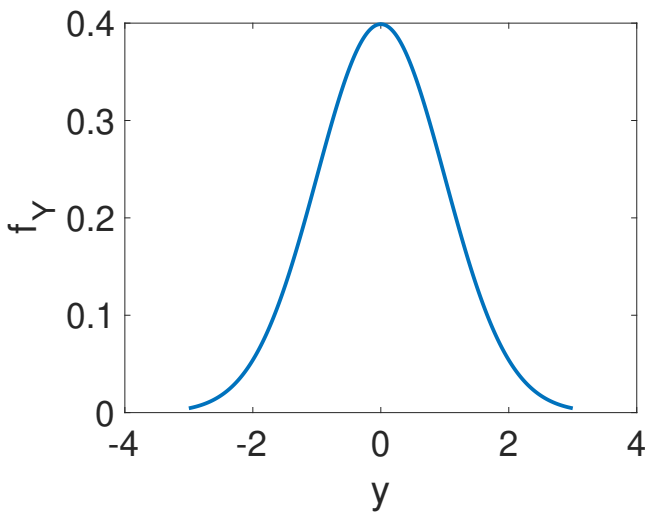
Examples

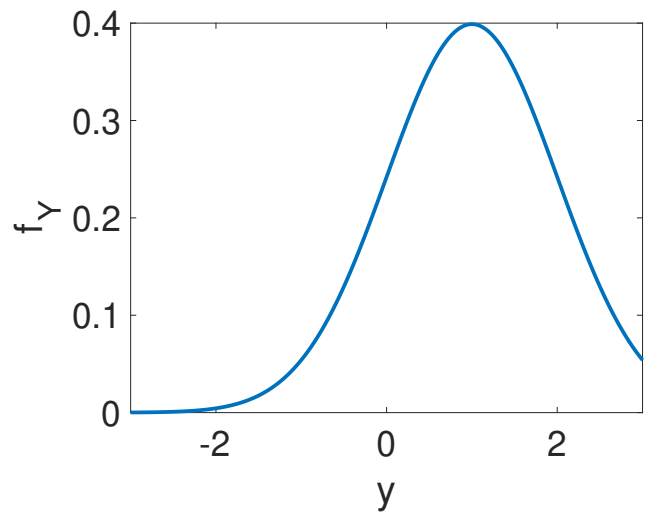
Suppose random variable $Y = M$, where $M \sim N(0, 1)$. Then

- $f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})$

Now suppose random variable $Y = x + M$, where $M \sim N(0, 1)$. Then

- $f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{(y-x)^2}{2})$
- e.g. with $x = 1$:





Examples

Suppose random variable $Y = X + M$, where $M \sim N(0, 1)$. Conditioned on $X = x$, what is the PDF of Y ?

- $f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{(y-x)^2}{2})$

Suppose that $X \sim N(0, \sigma^2)$. What is $f_{X|Y}(x|Y)$?

- Use Bayes Rule:

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}$$

$$= \frac{\frac{1}{\sqrt{2\pi}} \exp(-\frac{(y-x)^2}{2}) \times \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{x^2}{2\sigma^2})}{f_Y(y)}$$

- $f_Y(y)$ is just a normalising constant (so that the area under $f_{X|Y}(x|y)$ is 1).

Examples

Suppose random variables $X \sim N(0,1)$ and $Y \sim N(0,1)$ are independent. What is their joint PDF?

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

When independent then

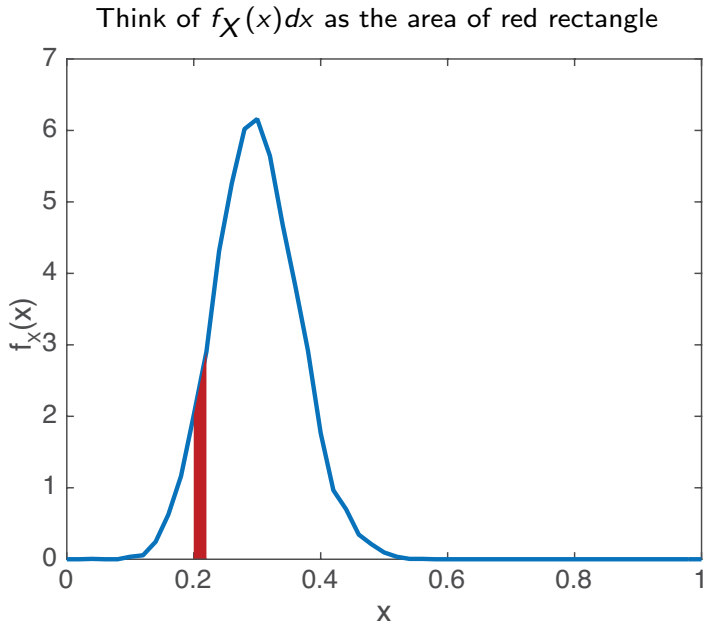
$$f_{XY}(x,y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) \times \frac{1}{\sqrt{2\pi}} \exp(-\frac{y^2}{2})$$

Expectation and Variance

For dx infinitesimally small,

$$P(x \leq X \leq x + dx) = F_X(x + dx) - F_X(x)$$
$$\approx f_X(x)dx$$

so we can think of $f_X(x)dx$ as the probability that X takes a value between x and $x + dx$.



Definitions:

For <u>discrete</u> RV X	For <u>continuous</u> RV X
$E[X] = \sum_x xP(X = x)$	$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx$
$E[X^n] = \sum_x x^n P(X = x)$	$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)dx$

As before $Var(X) = E[(X - E[X])^2) = E[X^2] - E[X]^2$.

Expectation and Variance

For both discrete and continuous random variables:

$$E[aX + b] = aE[X] + b$$
$$Var(X) = E[(X - \mu)^2] = E[X^2] - (E[X])^2$$
$$Var(aX + b) = a^2 Var(X)$$

(just replace sum with integral in previous proofs)