Example 4 (Poisson) *A Poisson random variable has infinite range* $0, 1, 2, \ldots$. *The pdf is given by the formula*

$$P(X = k) = \frac{\lambda^k}{k!}\, e^{-\lambda}, \quad k = 0, 1, \ldots$$

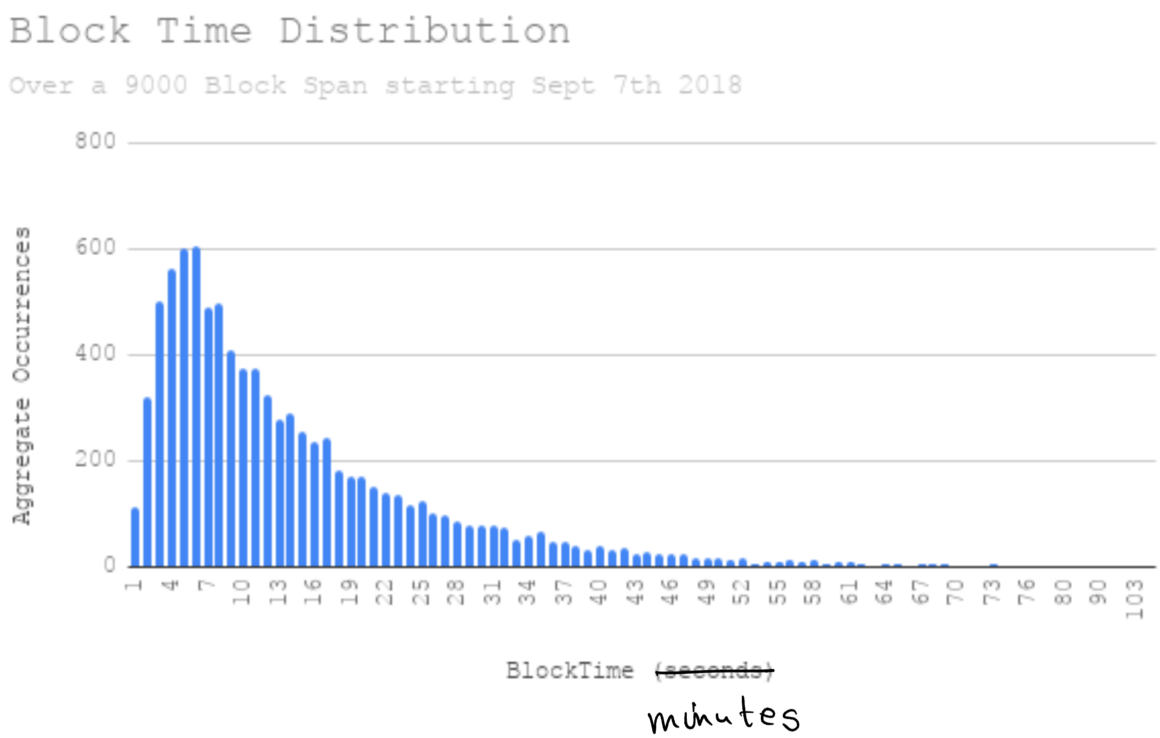
λ = lambda.

where $\lambda > 0$ *is a fixed parameter. Note that the Poisson is a useful approximation for the binomial* $\mathrm{Bin}(n, p)$ *(and is much simpler) in the case where* n *is large and* p *is small. The binomial convergence result is*

$$\lim_{\substack{n \to \infty \\ p \to 0 \\ np = \lambda}} \binom{n}{k} p^k (1-p)^{n-k} = \frac{\lambda^k}{k!}\, e^{-\lambda}$$

Compare Poisson value with exact result for previous example.

X	0	1	2	3	4	5	......
prob.	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2}{2} e^{-\lambda}$	$\frac{\lambda^3}{6} e^{-\lambda}$	......		


Block Time Distribution
Over a 9000 Block Span starting Sept 7th 2018
800
600
400
200
0
Aggregate Occurrences
1
4
7
10
13
16
19
22
25
28
31
34
37
40
43
46
49
52
55
58
61
64
67
70
73
76
80
90
103
BlockTime ~~(seconds)~~
minutes

Continuous Random Variables

If X is *continuous* then $\mathrm{Ran}(X)$ is an interval of real numbers, and the pdf is the density of probabilities around each possible value. The pdf for a continuous r.v. X is a non-negative function f_X called the probability density function. Roughly, the value $f_X(x)\,dx$ is the probability to measure X in the interval $[x, x+dx]$ where dx is a small value. More precisely, the probability to find X in the interval $[a, b]$ is

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

Note that $\mathbb{P}(X = a) = \int_a^a f_X(x)\,dx = 0$, so the probability to find any particular value of X is zero. The normalization condition is

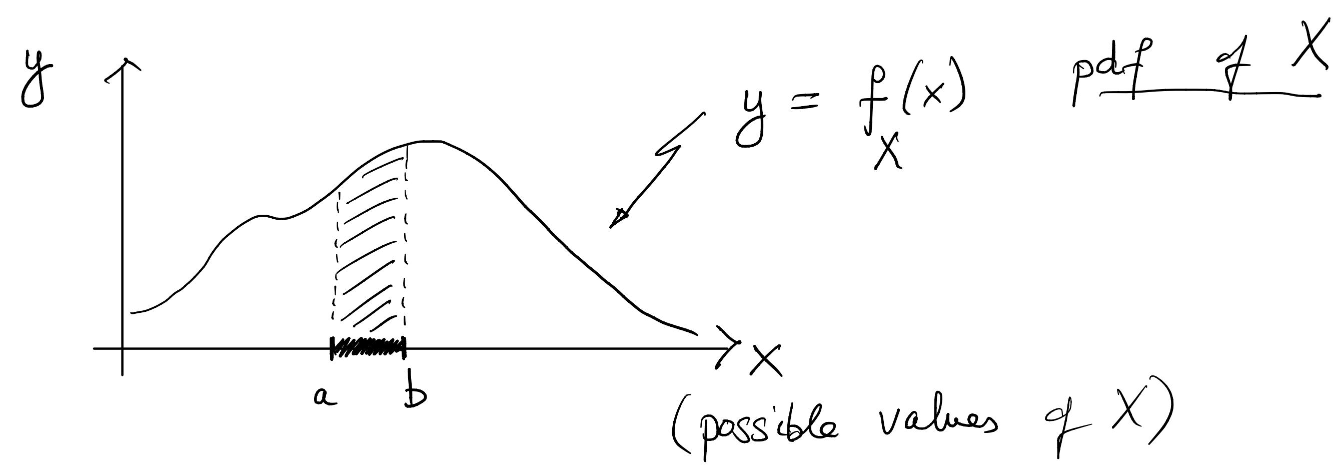
$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

Many special cases are important, we list a few here.

Example 5 (Uniform) *The r.v. X is* uniform *on the interval $[a, b]$ if X is 'equally likely' to be anywhere in the interval $[a, b]$, and has zero probability to be outside this interval. The pdf of X is constant on the interval $[a, b]$ and is zero outside this interval. That is,*

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}.$$

The special case $[a, b] = [0, 1]$ is often denoted by U, for uniform.



Prob. to measure X between a and b is equal the area above $[a, b]$ below the graph of the pdf $f_X(x)$.

$$\mathbb{P}(a \le X \le b)$$

$$= \mathbb{P}(X \in [a,b])$$

= shaded area in graph

$$= \int_a^b f_X(x)\,dx$$

Normalization property:

total area under the graph of $f_X(x)$ must equal 1.

(you haven't missed any possible values).

<u>Uniform r.v.</u>

X is uniform on $[0,1]$.

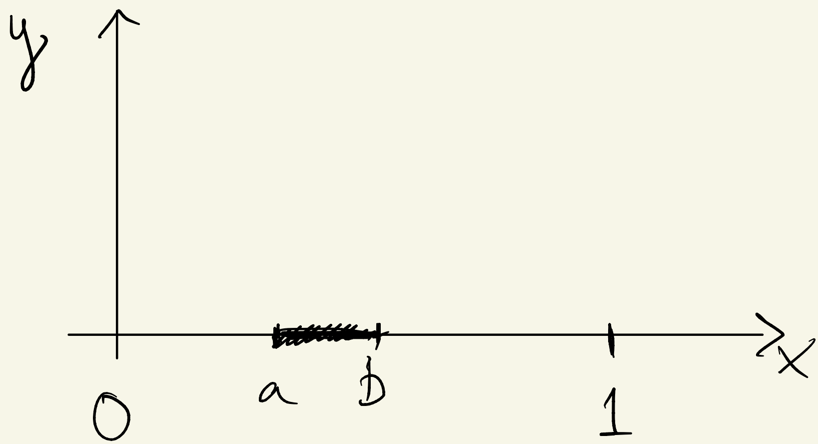
<u>Meaning</u>: X is the output of a random number generator.

$$0 \le X \le 1$$

equally likely to take any value

$\Leftrightarrow$ uniform across the interval.

<u>pdf of X:</u>



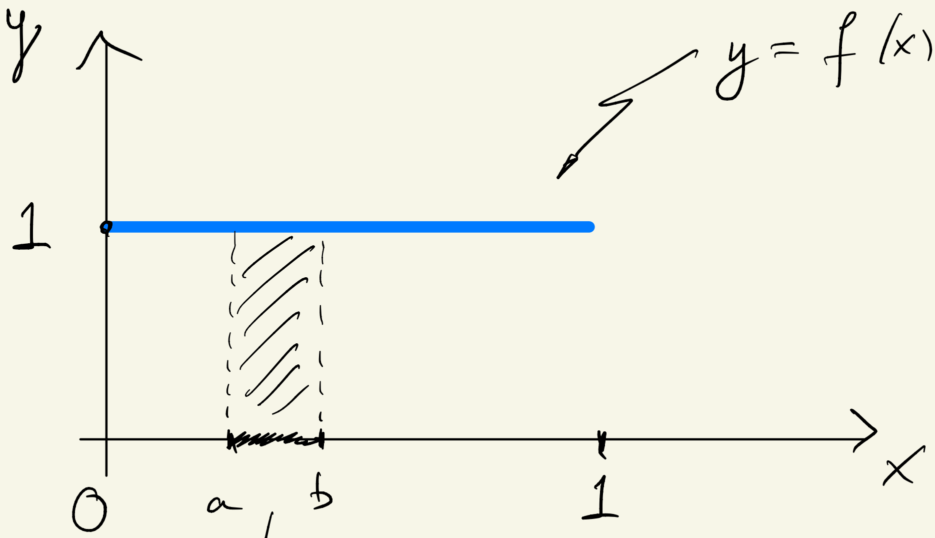
<u>Uniform:</u> $\mathbb{P}(a \le X \le b)$

$=$ prob. that a randomly thrown dart lands in $[a, b]$

$= b - a$ (length of interval $[a, b]$).

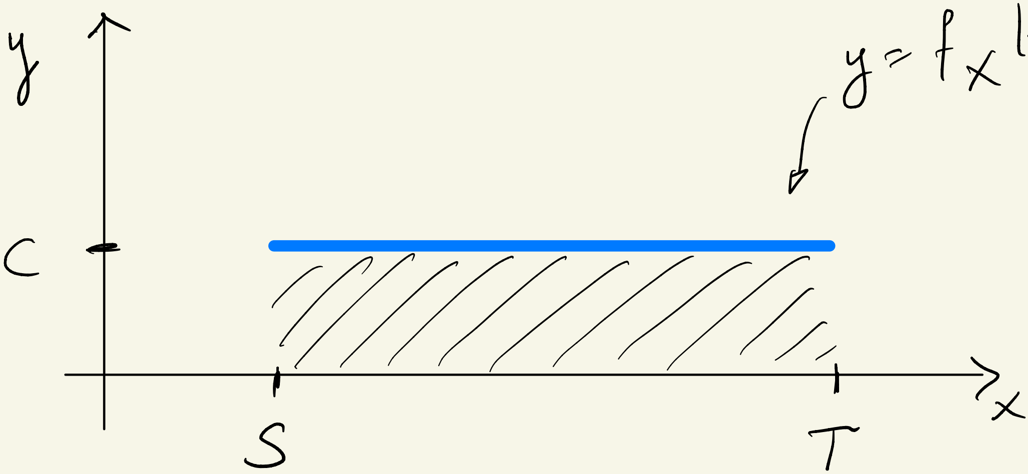
$$= \int_a^b f_X(x)\, dx$$

$$\Rightarrow \quad f_X(x) = \begin{cases} 1 & \text{for } 0 \le x \le 1 \\ 0 & \text{else.} \end{cases}$$



area $= (b-a) \cdot 1 = b-a.$

Uniform r.v. on $[S, T]$:



Normalization condition:

Total area $= 1$.

$$(T-S)\,c = 1.$$

$$c = \frac{1}{T-S}.$$

$\Rightarrow$ pdf is

$$f_X(x) = \begin{cases} \dfrac{1}{T-S} & S \leq x \leq T \\ 0 & \text{else.} \end{cases}$$

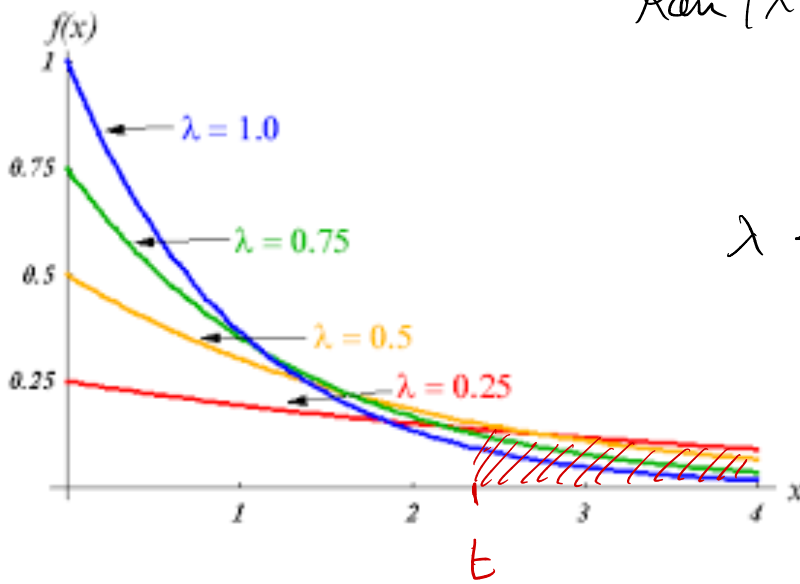
Example 6 (Exponential) *The pdf is*

$$f(x) = \begin{cases} ke^{-kx} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

$\underline{k > 0}$ positive parameter.

where $k > 0$ is called the rate. *This is often used to model the time until failure of a device.*

$\text{Ran}(X) = [0, \infty)$



$\lambda \Leftrightarrow k.$

$$\mathbb{P}(X > t) = \mathbb{P}(t < X < \infty)$$

$$= \int_t^\infty f_X(x)\,dx$$

$$= \int_t^\infty k e^{-kx}\,dx$$

$$= k\left(-\frac{1}{k}\right)e^{-kx}\Big|_t^{\infty}$$

$$= -\left(0 - e^{-kt}\right)$$

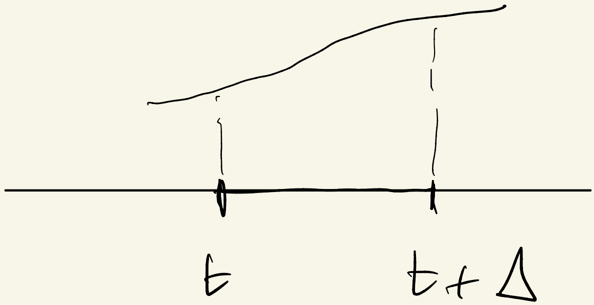
$$\mathbb{P}(X > t) = e^{-kt} \qquad (t \geq 0).$$

Important:

$\mathbb{P}(X = t) = 0$ for a continuous r.v.

What does the pdf tell you at the value t?

$f_X(t) \neq$ prob. for $X = t$



Δ is very small.

$$\mathbb{P}(t \leq X \leq t+\Delta)$$

$$= \int_t^{t+\Delta} f_X(x)\, dx$$

Because Δ is very small,

$f_X(x)$ is almost constant

in interval $[t, t+\Delta]$.

$$\Rightarrow \mathbb{P}(t \le X \le t+\Delta)$$

$$\approx \int_t^{t+\Delta} f_X(t)\, dx$$

$$= f_X(t) \cdot \Delta$$

Operational meaning of pdf:

$f_X(t)$ times length of interval

gives the prob.

So $\mathbb{P}(X>t) = \mathbb{P}(X \ge t)$

for a continuous r.v.

$$\mathbb{P}(a < X \leq b)$$

$$= \mathbb{P}(a \leq X \leq b)$$

etc.

Example 7 (Normal) *Probably the most important family of r.v.'s, widely used in statistics. The pdf is*

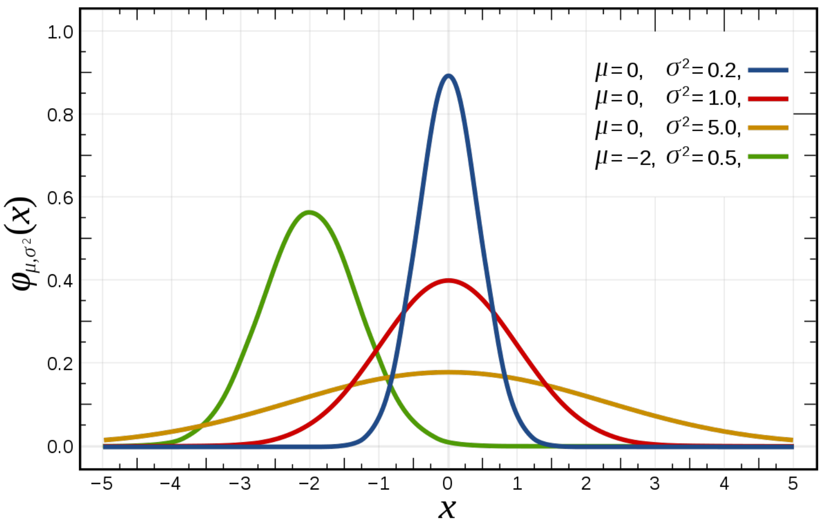
2 parameter family:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

μ = "mu"
σ^2 = sigma squared.

where μ is the mean and $\sigma^2 > 0$ is the variance. The special case $\mu = 0$, $\sigma = 1$ is called the standard normal.

$Ran(X) = (-\infty, \infty)$



bell curve

Standard normal Z.

pdf:

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

Ex:

$$\mathbb{P}(Z \geq 1) = \int_1^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

↓

no analytical formula

↓

look up in tables.

We finish up with a summary of properties of the most well-known (and useful) random variables. The first table lists the random variables along with their parameters and a useful 'story' that describes how they arise.

	Name	Story	Ran
Discrete	Bernoulli(p)	Toss a coin with probability p of turning up heads. X = Number of heads in one toss.	$\{0,1\}$
	Binomial(n,p)	Toss a coin with probability p of turning up heads. X= Number of heads in n tosses. Binomial(n,p) is the sum of n independent Bernoulli(p).	$\{0,1,2,..,n\}$
	Geometric(p)	Toss a coin with probability p of turning up heads. X = Number of tosses until the first Head.	$\{1,2,3,...\}$
	Poisson(λ)	Random calls arrive with rate λ. X = Number of calls that arrive in one time unit.	$\{0,1,2,...\}$
Continuous	Exponential(λ)	Random calls arrive with rate λ. X = time until the first arrival.	$[0,\infty)$
	Gamma(n,λ)	Random calls arrive with rate λ. X = time until the nth arrival.	$[0,\infty)$.
	Uniform(a,b)	Pick a random number X between a and b.	$[a,b]$
	Normal(μ,σ^2)	Pick an individual in a large population. X = height of the individual.	$(-\infty,\infty)$.

The second table lists the means, variances and pdf's of these random variables.

Name	Mean	Variance	pdf
Bernoulli(p)	p	$p(1-p)$	$p_X(k) = p^k\,(1-p)^{1-k}$
Binomial(n,p)	np	$np(1-p)$	$p_X(k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k(1-p)^{n-k}$
Geometric(p)	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$p_X(k) = (1-p)^{k-1}p$
Poisson(λ)	λ	λ	$p_X(k) = e^{-\lambda}\lambda^k/k!$
Exponential(λ)	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$f_X(x) = \lambda e^{-\lambda x}$
Gamma(n,λ)	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$f_X(x) = \lambda^k x^{k-1} e^{-\lambda x}/(k-1)!$
Uniform(a,b)	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$f_X(x) = 1/(b-a)$
Normal(μ,σ^2)	μ	σ^2	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-(x-\mu)^2/2\sigma^2}$

For a continuous random variable X, we can recover the pdf from the probability formula by noting that

$$f_X(x) = \frac{d}{dx} \int_{-\infty}^{x} f_X(u)\, du = \frac{d}{dx}\, \mathbb{P}(X \le x)$$

The *cumulative distribution function = cdf* is defined by

$$F_X(x) = \mathbb{P}(X \le x)$$

So we have the important relation

$$f_X(x) = \frac{d}{dx}\, F_X(x)$$

The cdf is important because it is often easier to first compute the cdf of a continuous r.v. and then take its derivative to get the pdf. Properties of the cdf:

(a) $0 \le F(x) \le 1$ for all $x \in \mathbb{R}$

(b) if $x < y$ then $F(x) \le F(y)$ (increasing function)

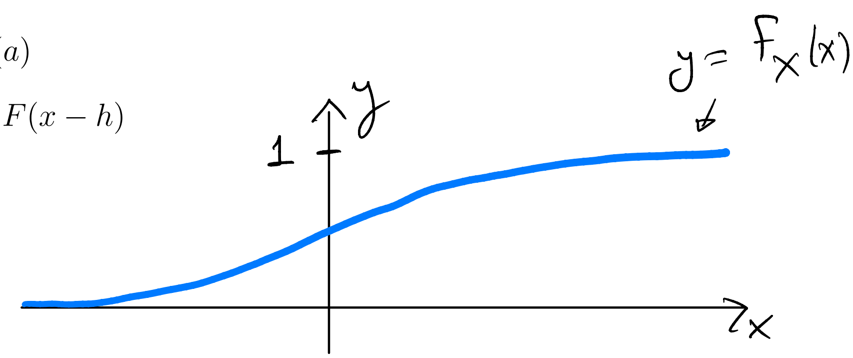
(c) $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$

(d) F is right continuous: if x_n is a decreasing sequence and $\lim x_n = x$ then $\lim F(x_n) = F(x)$

(e) $P(a < X \le b) = F(b) - F(a)$

(f) $P(X = x) = F(x) - \lim_{h\downarrow 0} F(x - h)$

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$$



Example 8 *Suppose U is uniform on $[0,1]$. Find the pdf of $X = \sqrt{U}$. Let's first find the cdf of X:* $\mathrm{Ran}(X) = [0,1]$, *and for any $x \in [0,1]$*

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(\sqrt{U} \le x) = \mathbb{P}(U \le x^2) = x^2$$

Therefore

$$f_X(x) = F_X(x)' = \begin{cases} 2x & \text{for } 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$$

U is uniform on $[0,1]$.

pdf: $f_U(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$

Another r.v. $X = \sqrt{U}$. What is pdf of X?

Note! it is not $\sqrt{f_U}$!!

First find the cdf of X.

$$F_X(x) = \mathbb{P}(X \le x)$$

$$= \mathbb{P}(\sqrt{U} \le x)$$

$$= \mathbb{P}(U \le x^2)$$

$$= \int_0^{x^2} f_U(t)\, dt$$

$$= \int_0^{x^2} 1 \cdot dt$$

$$F_X(x) = x^2. \qquad (0 \le x \le 1).$$

Great! The pdf is

$$f_X(x) = \frac{d}{dx} F_X(x)$$

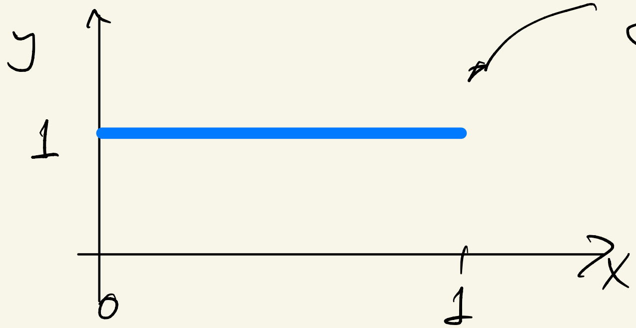
$$= \frac{d}{dx} (x^2)$$

$$= 2x \qquad (0 \le x \le 1)$$

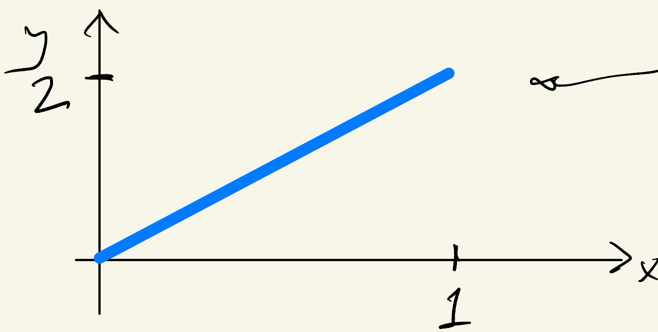
Check normalization:

$$\int_0^1 f_X(x)\,dx \stackrel{?}{=} 1$$

$$\int_0^1 2x\,dx = x^2\Big|_0^1 = 1 \checkmark$$



$y = f_u(x)$



$y = f_X(x)$

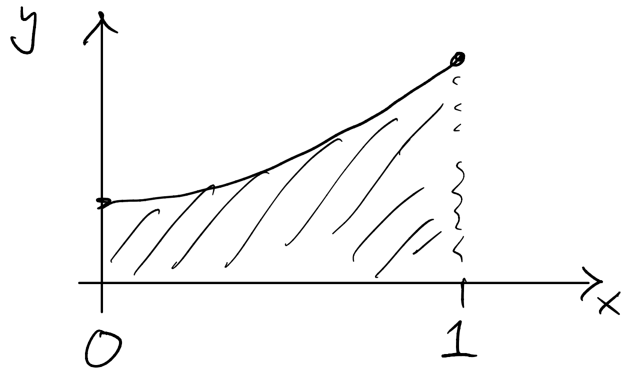
Example 9 *This example illustrates how 'randomness' can be a resource for computation. Suppose we want to find the area below the curve $y = f(x)$ over the range $0 \le x \le 1$, where*

$$f(x) = \left(3 - \sin^2 x\right)^{-1/2}$$

The integral cannot be computed exactly so we need some numerical method. Let U, V be independent uniform random numbers on $[0, 1]$. Think of these as the (x, y) coordinates of a randomly chosen point in the unit square. Then

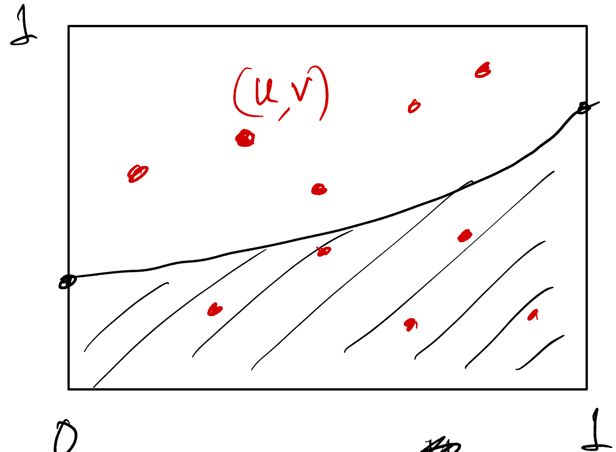
$$\mathbb{P}\left(V \le f(U)\right) = \textit{area under the curve } \{y = f(x)\} = \int_0^1 \left(3 - \sin^2 x\right)^{-1/2} dx$$

$$f(x) = \frac{1}{\sqrt{3 - \sin^2 x}}$$



Area = ?

$$= \int_0^1 \frac{1}{\sqrt{3 - \sin^2 x}}\, dx$$

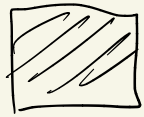


(U, V) = random point in the square.

Pick n points $(U_1, V_1), \ldots, (U_n, V_n)$ randomly distributed over the square.

$\mathbb{P}($ (U,V) is below curve $y = f(x))$

$= \mathbb{P}($ (U,V) inside shaded region$)$

= area () / aea () $(=1)$

$$\approx \int_0^1 \frac{1}{\sqrt{3-\sin^2 x}}\, dx$$

<u>Event</u>: $A = \{ (U,V)$ lies below the curve$\}$.

$\mathbb{P}(A) = \,?$

<u>Idea</u>: random points (U_1, V_1), (U_2, V_2), ..., (U_n, V_n) in square.

Let R_n = number of points which lie in shaded region.

<u>Law of Large Numbers</u>:

repeated independent trials.

$$\Rightarrow \quad \frac{R_n}{n} = \text{fraction of trials where event } A \text{ is true.}$$

$$\Rightarrow \quad \lim_{n \to \infty} \frac{R_n}{n} = \mathbb{P}(A).$$

[<u>Ex.</u> Toss a fair coin many times.

Fraction of times you get Heads:

converges to $\frac{1}{2}$].

This provides our strategy to estimate the integral: generate N independent pairs (U_i, V_i), $i = 1, \ldots, N$, and compute the fraction

$$R_N = \frac{\#\{i : V_i \leq f(U_i)\}}{N}$$

As $N \to \infty$ this will converge to $\mathbb{P}(V \leq f(U))$. Here are results for different values of N, computed using Matlab:

N	*Trial 1*	*Trial 2*	*Trial 3*	*Time per trial (sec)*
1000	*0.617*	*0.595*	*0.590*	*0.0005*
10000	*0.6153*	*0.6154*	*0.6061*	*0.003*
100000	*0.6066*	*0.6079*	*0.6071*	*0.02*
1000000	*0.6074*	*0.6067*	*0.6073*	*0.18*

The actual value of the integral is 0.6071 *so with $N = 10^6$ we are getting a good estimate.*

Joint pdf

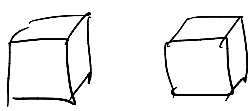
For two discrete random variables X and Y, the *joint pdf* is the list of probabilities for all possible pairs of values:

$$p_{X,Y}(x,y) = P(X = x, Y = y)$$

for all $x \in \mathrm{Ran}(X)$ and $y \in \mathrm{Ran}(Y)$.

If $\mathrm{Ran}(X)$ and $\mathrm{Ran}(Y)$ are finite, then the joint pdf is conveniently presented as a table of values.

Example 10 (Joint pdf: example) *Roll two dice, let X be the value on the first die, and let Y be the maximum of the two values.*



	$X=1$	2	3	4	5	6
$Y=1$	$\frac{1}{36}$	0	0	0	0	0
2	$\frac{1}{36}$	$\frac{2}{36}$	0	0	0	0
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{3}{36}$	0	0	0
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{4}{36}$	0	0
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	0
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$

Marginal for X:

$$\mathbb{P}(X=4) = \mathbb{P}(X=4, Y=1) + \mathbb{P}(X=4, Y=2) + \dots + \mathbb{P}(X=4, Y=6)$$
$$= 0+0+0 + \frac{4}{36} + \frac{1}{36} + \frac{1}{36} = \frac{6}{36} = \frac{1}{6}.$$

$X =$ value on first die

$Y =$ maximum of two dice.

$$\mathbb{P}(X=1, Y=1) = \mathbb{P}(\boxed{1}, \boxed{*}) = \mathbb{P}(\boxed{1}, \boxed{1}) = \frac{1}{36}.$$

$$\mathbb{P}(X=2, Y=1) = \mathbb{P}(\boxed{2}, \boxed{?}) \quad (\text{max.} = 1)$$
$$= 0.$$

$$\mathbb{P}(X=2, Y=2) = \mathbb{P}(\boxed{2}, \boxed{*}) \quad (\text{max} = 2)$$
$$= \mathbb{P}(\boxed{2}, \boxed{1}) + \mathbb{P}(\boxed{2}, \boxed{2})$$

$$= \frac{1}{36} + \frac{1}{36}$$

<u>Normalization condition:</u>

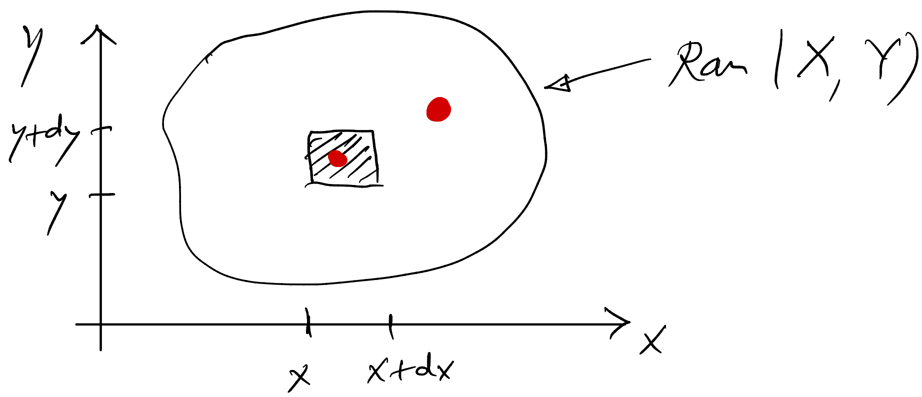
$$\text{total probability} = \sum_{x \in Ran(X)} \sum_{y \in Ran(Y)} p(x, y)$$

$$= 1.$$

For continuous random variables X and Y, the joint pdf is the function $f_{X,Y}(x, y)$ of the two real variables x, y. Again the rough meaning is that $f_{X,Y}(x, y)\, dx\, dy$ is the probability to find the pair (X, Y) in the small rectangle $[x, x + dx] \times [y, y + dy]$. The precise meaning is that the probability for the pair (X, Y) to be in some region R in the xy-plane is the double integral

$$\mathbb{P}((X, Y) \in R) = \int_R f_{X,Y}(x, y)\, dx\, dy$$

We will not be much concerned with joint pdf's for continuous random variables in this course.



$f_{X,Y}(x,y)\, dx\, dy$ = prob. to find (X, Y) in the small rectangle.

Joint pdf: marginals

We can recover the individual pdf of X from the joint pdf of (X, Y) by summing over all values of Y. Similarly we can recover the pdf of Y by summing over X.

Example 11 *Using our example above where X is the value on the first die, and Y is the maximum of the two values. Then we get p_X by holding the column fixed and summing over the rows, and we get p_Y by holding the row fixed and summing over the columns.*

	X=1	*2*	*3*	*4*	*5*	*6*
Prob	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

uniform.

	Y=1	*2*	*3*	*4*	*5*	*6*
Prob	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Expected value

The *expected value or mean* of a discrete random variable X is defined to be

$$\mathbb{E}[X] = \sum_i x_i \, P(X = x_i)$$

weighted sum of possible values of X.

where the sum runs over the possible values of X.

The operational meaning is that $\mathbb{E}[X]$ is the long-run average value of repeated measurements of the random variable X. That is, suppose that we measure the random variable X in N independent trials, and record the results as $X_1, X_2, \dots, X_N$. Then the long-run average value is

$$\overline{X}_N = N^{-1} \; (X_1 + X_2 + \cdots + X_N)$$

We will shortly see the Law of Large Numbers which implies that

$$\lim_{N \to \infty} \overline{X}_N = \mathbb{E}[X]$$

For example, if X is the outcome of rolling a die, then the numbers $\{1, 2, \dots, 6\}$ all occur with probability $1/6$. So the expected value is

$$\mathbb{E}[X] = 1(1/6) + 2(1/6) + \cdots + 6(1/6) = 7/2$$

Example 12 *Roll three fair dice. Let X be the number of different faces that appear. Find the pdf of X and compute $\mathbb{E}[X]$.*

$\boxed{1}$ $\boxed{2}$ $\boxed{5}$ $X = 3$

$\boxed{3}$ $\boxed{1}$ $\boxed{3}$ $X = 2$

$\boxed{6}$ $\boxed{6}$ $\boxed{6}$ $X = 1.$

X	1	2	3
prob.	$\frac{6}{6^3}$	$\frac{(6)(15)}{6^3}$	$\frac{(6)(5)(4)}{6^3}$

$$\mathbb{E}[X] = (1)\cdot\left(\frac{1}{36}\right) + 2\left(\frac{15}{36}\right) + 3\left(\frac{20}{36}\right)$$

$$= \frac{91}{36}$$