Recap: Random Variables

- Given a probability space $(\Omega, \mathcal{F}, P)$, a random variable is a real-valued function on Ω.
- It essentially results in an induced probability space

$$(\Omega, \mathcal{F}, P) \stackrel{X}{\rightarrow} (\Re, \mathcal{B}, P_X)$$

where $\mathcal{B}$ is the Borel σ-algebra and

$$P_X(B) = P[X \in B] = P\left(\{\omega \in \Omega \;:\; X(\omega) \in B\}\right)$$

Recap: σ-algebra

- An $\mathcal{F} \subset 2^\Omega$ is called a σ-algebra (also called σ-field) on Ω if it satisfies
  1. $\Omega \in \mathcal{F}$
  2. $A \in \mathcal{F} \;\;\Rightarrow\;\; A^c \in \mathcal{F}$
  3. $A_1, A_2, \cdots \in \mathcal{F} \;\;\Rightarrow\;\; \cup_i A_i \in \mathcal{F}$
- Thus a σ-algebra is a collection of subsets of Ω that is closed under complements and countable unions (and hence countable intersections).
- The Borel σ-algebra (on $\Re$), $\mathcal{B}$, is the smallest σ-algebra containing all intervals.
- We also have $\mathcal{B} = \sigma(\{(-\infty,\ x] : x \in \Re\}$

Recap: Distribution function of a random variable

- Let X be a random variable. It distribution function, $F_X : \Re \rightarrow \Re$, is defined by

$$F_X(x) = P[X \leq x] = P(\{\omega \in \Omega \ : \ X(\omega) \leq x\})$$

- The distribution function, F_X, completely specifies the probability measure, P_X.
- Thus, given the distribution function of X, one can (in principle) calculate probability of any event $[X \in B]$.

Recap: Properties of distribution function

- The distribution function satisfies
  1. $0 \leq F_X(x) \leq 1, \ \forall x$
  2. $F_X(-\infty) = 0; \ F_X(\infty) = 1$
  3. F_X is non-decreasing: $x_1 \leq x_2 \ \Rightarrow \ F_X(x_1) \leq F_X(x_2)$
  4. F_X is right continuous and has left-hand limits.
- Any real-valued function of a real variable satisfying the above four properties would be a distribution function of some random variable.
- We also have
  $F_X(x^+) - F_X(x^-) = F_X(x) - F_X(x^-) = P[X = x]$
  $P[a < X \leq b] = F_X(b) - F_X(a)$.

Recap

- There are two classes of random variables that we would study here.
- These are called discrete and continuous random variables.
- Note that the distribution function is defined for **all** random variables.

Discrete Random Variables

- A random variable X is said to be discrete if it takes only countably many distinct values.
- Countably many means finite or countably infinite.

Discrete Random Variable Example

- Consider three independent tosses of a fair coin.
- $\Omega = \{H,\ T\}^3$ and $X(\omega)$ is the number of H's in ω.
- This rv takes four distinct values, namely, $0, 1, 2, 3$.
- We denote this as $X \in \{0, 1, 2, 3\}$
- Let us find the distribution function of this rv
- Let us take some examples of $[X \leq x]$

$$[X \leq 0.72] \;=\; \{\omega\ :\ X(\omega) \leq 0.72\} = \{\omega\ :\ X(\omega) = 0\} = [X = 0]$$

$$\begin{aligned}[X \leq 1.57] &= \{\omega\ :\ X(\omega) \leq 1.57\} \\ &= \{\omega\ :\ X(\omega) = 0\} \cup \{\omega\ :\ X(\omega) = 1\} = [X = 0 \text{ or } 1]\end{aligned}$$

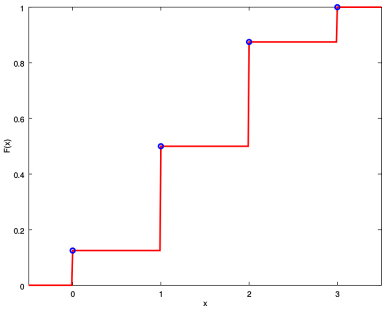
- $F_X(x) = P[X \le x]$ (Recall $X \in \{0, 1, 2, 3\}$)
- The event $[X \le x]$ for different x can be seen to be

$$[X \le x] = \begin{cases} \phi & x < 0 \\ \{TTT\} & 0 \le x < 1 \\ \{TTT, HTT, THT, TTH\} & 1 \le x < 2 \\ \Omega - \{HHH\} & 2 \le x < 3 \\ \Omega & x \ge 3 \end{cases}$$

- So, we get the distribution function as

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{8} & 0 \le x < 1 \\ \frac{4}{8} & 1 \le x < 2 \\ \frac{7}{8} & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

- The plot of this distribution function is:



- This is a stair-case function.
- It has jumps at $x = 0, 1, 2, 3$, which are the values that X takes. In between these it is constant.
- The jump at, e.g., $x = 2$ is $3/8$ which is the probability of X taking that value.

- We know that $F_X(x) - F_X(x^-) = P[X = x]$.
- For example,

$$\begin{aligned} F_X(2) - F_X(2^-) &= P[X = 2] = P(\{\omega \ : \ X(\omega) = 2\}) \\ &= P(\{THH, HTH, HHT\}) = \frac{3}{8} \end{aligned}$$

- The F_X is a stair-case function.
- It has jumps at each value assumed by X (and is constant in between)
- The height of the jump is equal to the probability of X taking that value.
- All discrete random variables would have this general form of distribution function.

- Let X be a dicrete rv and let $X \in \{a_1, a_2, \cdots, a_n\}$
- As a notation we assume: $a_1 < a_2 < \cdots < a_n$
- Let $[X = a_i] = \{\omega \ : \ X(\omega) = a_i\} = B_i$ and let $P(B_i) = q_i$.
- Since X is a function on Ω, $B_1, \cdots, B_n$ form a partition of Ω.
- Note that $q_i \geq 0$ and $\sum_{i=1}^{n} q_i = 1$.
- If $x < a_1$ then $[X \leq x] = \phi$.
- If $a_1 \leq x < a_2$ then $[X \leq x] = [X = a_1] = B_1$
- If $a_2 \leq x < a_3$ then $[X \leq x] = [X = a_1] \cup [X = a_2] = B_1 + B_2$

- Hence we can write the distribution function as

$$F_X(x) = \begin{cases} 0 & x < a_1 \\ P(B_1) & a_1 \le x < a_2 \\ P(B_1) + P(B_2) & a_2 \le x < a_3 \\ & \vdots \\ \sum_{i=1}^{k} P(B_i) & a_k \le x < a_{k+1} \\ & \vdots \\ 1 & x \ge a_n \end{cases}$$

- We can write this compactly as

$$F_X(x) = \sum_{k: a_k \le x} q_k$$

- Note that all this holds even when X takes countably infinitely many values.

- Let X be a discrete rv with $X \in \{x_1, x_2, \cdots\}$.
- Let $q_i = P[X = x_i] \quad (= P(\{\omega \ : \ X(\omega) = x_i\})\ )$
- We have $q_i \geq 0$ and $\sum_i q_i = 1$.
- If X is discrete then there is a countable set E such that $P[X \in E] = 1$.
- That is, $P[X \in \{x_1, x_2, \cdots\}] = 1$
- The distribution function of X is specified completely by these q_i

probability mass function, f_X

- Let X be a discrete rv with $X \in \{x_1, x_2, \cdots\}$.
- The probability mass function (pmf) of X is defined by

$$f_X(x_i) = P[X = x_i]; \quad f_X(x) = 0, \quad \text{for all other } x$$

- f_X is also a real-valued function of a real variable.
- We can write the definition compactly as $f_X(x) = P[X = x]$
- The distribution function (df) and the pmf are related as

$$F_X(x) = \sum_{i:x_i \leq x} f_X(x_i)$$

$$f_X(x) = F_X(x) - F_X(x^-)$$

- We can get pmf from df and df from pmf.

Properties of pmf

- The probability mass function of a discrete random variable $X \in \{x_1, x_2, \cdots\}$ satisfies
  1. $f_X(x) \geq 0, \forall x$ and $f_X(x) = 0$ if $x \neq x_i$ for some i
  2. $\sum_i f_X(x_i) = 1$
- Any function satisfying the above two would be a pmf of some discrete random variable.
- We can specify a discrete random variable by giving either F_X or f_X.
- Please remember that we have defined distribution function for any random variable. But pmf is defined only for discrete random variables

- Any discrete random variable can be specified by
  - giving the set of values of X, $\{x_1, x_2, \cdots\}$, and
  - numbers q_i such that $q_i = P[X = x_i] = f_X(x_i)$
- Note that we must have $q_i \geq 0$ and $\sum_i q_i = 1$.
- As we saw this is how we can specify a probability assignment on any countable sample space.

Computations of Probabilities for discrete rv's

- A discrete random variable is specified by giving either df or pmf. One can be obtained from the other.
- We normally specify it through the pmf.
- Given $X \in \{x_1, x_2, \cdots\}$ and f_X, we can (in principle) compute probability of any event

$$P[X \in B] = \sum_{\substack{i: \\ x_i \in B}} f_X(x_i)$$

- For example, if $X \in \{0, 1, 2, 3\}$ then

$$P[X \in [0.5,\ 1.32] \cup [2.75,\ 5.2]\ ] = f_X(1) + f_X(3)$$

- We next look at some standard discrete random variable models

Bernoulli Distribution

- Bernoulli random variable: $X \in \{0,\ 1\}$ with

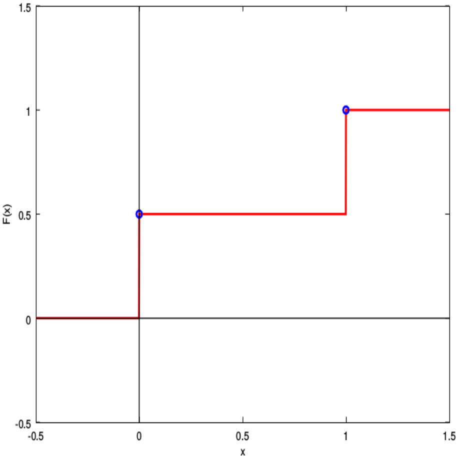
  $f_X(1) = p;\ f_X(0) = 1{-}p;$ where $0 < p < 1$ is a parameter

- This f_X is easily seen to be a pmf
- Consider $(\Omega, \mathcal{F}, P)$ with $B \in \mathcal{F}$. (The Ω here may be uncountable).
- Consider the random variable

$$I_B(\omega) = \begin{cases} 0 & \text{if } \omega \notin B \\ 1 & \text{if } \omega \in B \end{cases}$$

- It is called indicator (random variable) of B.
- $P[I_B = 1] = P(\{\omega\ :\ I_B(\omega) = 1\}) = P(B)$
- Thus, this indicator rv has Bernoulli distribution with $p = P(B)$

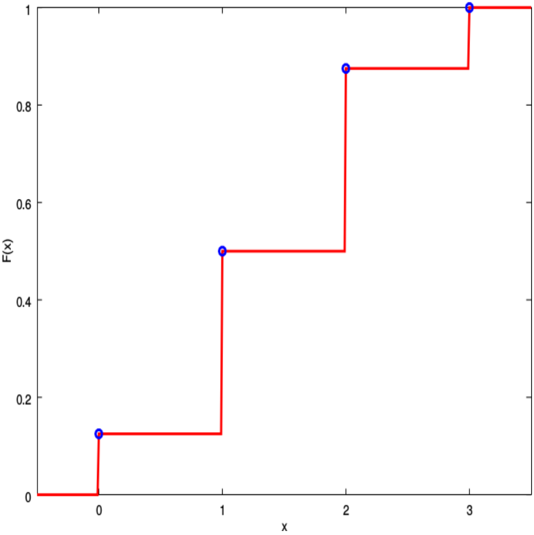
One of the df examples we saw earlier is that of Bernoulli



Binomial Distribution

- $X \in \{0,\ 1,\ \cdots,\ n\}$ with pmf
$$f_X(k) = {}^nC_k\ p^k\ (1-p)^{n-k},\ k = 0, 1, \cdots, n$$
where n, p are parameters (n is a $+$ve integer and $0 < p < 1$).
- This is easily seen to be a pmf
$$\sum_{k=0}^{n} {}^nC_k\ p^k\ (1-p)^{n-k} = (p + 1 - p)^n = 1$$
- Consider n independent tosses of coin whose probability of heads is p. If X is the number of heads then X has the above binomial distribution.
(Number of successes in n bernoulli trials)
- Any one outcome (a seq of length n) with k heads would have probability $p^k(1-p)^{n-k}$. There are ${}^nC_k$ outcomes with exactly k heads.

One of the df examples we considered was that of Binomial



Poisson Distribution

- $X \in \{0,\ 1,\ 2, \cdots\}$ with pmf

$$f_X(k) = \frac{\lambda^k\, e^{-\lambda}}{k!},\ k = 0, 1, 2, \cdots$$

  where $\lambda > 0$ is a parameter.
- We can see this to be a pmf by

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\, e^{-\lambda} = e^{\lambda}\, e^{-\lambda} = 1$$

- Poisson distribution is also useful in many applications

Geometric Distribution

- $X \in \{1,\ 2, \cdots\}$ with pmf

$$f_X(k) = (1-p)^{k-1}\, p,\ k = 1, 2, \cdots$$

  where $0 < p < 1$ is a parameter.
- Consider tossing a coin (with prob of H being p) repeatedly till we get a head. X is the toss number on which we got the first head.
- In general waiting for 'success' in independent Bernoulli trials.

Memoryless property of geometric distribution

- Suppose X is a geometric rv. Let m, n be positive integers.
- We want to calculate $P(\ [X > m+n] \mid [X > m]\ )$ (Remember that $[X > m]$ etc are events)
- Let us first calculate $P[X > n]$ for any positive integer n

$$\begin{aligned} P[X > n] &= \sum_{k=n+1}^{\infty} P[X = k] = \sum_{k=n+1}^{\infty} (1-p)^{k-1}\, p \\ &= p\, \frac{(1-p)^n}{1-(1-p)} = (1-p)^n \end{aligned}$$

(Does this also tell us what is df of geometric rv?)

- Now we can compute the required conditional probability

$$\begin{aligned} P[X > m+n \mid X > m] &= \frac{P[X > m+n, X > m]}{P[X > m]} \\ &= \frac{P[X > m+n]}{P[X > m]} \\ &= \frac{(1-p)^{m+n}}{(1-p)^{m}} = (1-p)^n \\ \Rightarrow P[X > m+n \mid X > m] &= P[X > n] \end{aligned}$$

- This is known as the memoryless property of geometric distribution
- Same as

$$P[X > m+n] = P[X > m]P[X > n]$$

- If X is a geometric random variable, it satisfies

$$P[X > m + n | X > m] = P[X > n]$$

- This is same as

$$P[X > m + n] = P[X > m] P[X > n]$$

- Does it say that $[X > m]$ is independent of $[X > n]$
- NO!
  Because $[X > m + n]$ is not equal to intersection of $[X > m]$ and $[X > n]$

Memoryless property defines geometric rv

- Suppose $X \in \{0,\ 1, \cdots\}$ is a discrete rv satisfying, for all non-negative integers, m, n

$$P[X > m + n] = P[X > m]P[X > n]$$

- We will show that X has geometric distribution
- First, note that
  $P[X > 0] = P[X > 0 + 0] = (P[X > 0])^2$
  $\Rightarrow P[X > 0]$ is either 1 or 0.
- Let us take $P[X > 0] = 1$ (and hence $P[X = 0] = 0$).

- We have, for any m,

$$\begin{aligned} P[X > m] &= P[X > (m-1)+1] \\ &= P[X > m-1]P[X > 1] \\ &= P[X > m-2]\,(P[X > 1])^2 \end{aligned}$$

- Let $q = P[X > 1]$. Iterating on the above, we get

$$P[X > m] = P[X > 0]\,(P[X > 1])^m = q^m$$

- Using this, we can get pmf of X as

$$P[X = m] = P[X > m-1] - P[X > m] = q^{m-1} - q^m = q^{m-1}(1-q)$$

- This is pmf of geometric (with $q = (1-p)$)

Continuous Random Variables

- A rv, X, is said to be continuous (or of continuous type) if its distribution function, F_X is absolutely continuous.

Absolute Continuity

- A function $g : \Re \rightarrow \Re$ is absolutely continuous on an interval, I, if given any $\epsilon > 0$ there is a $\delta > 0$ such that for any finite sequence of pair-wise disjoint subintervals, (x_k, y_k), with $x_k, y_k \in I, \ \forall k$, satisfying $\sum_k (y_k - x_k) < \delta$, we have $\sum_k |f(y_k) - f(x_k)| < \epsilon$
- A function that is absolutely continuous on a (finite) closed interval is uniformly continuous.
- If g is absolutely continuous on $[a, \ b]$ then there exists an integrable function h such that

$$g(x) = g(a) + \int_a^x h(t) \, dt, \ \ \forall x \in [a, \ b]$$

- In the above, g would be differentiable almost everywhere and h would be its derivative (wherever g is differentiable).

Continuous Random Variables

- A rv, X, is said to be continuous (or of continuous type) if its distribution function, F_X is absolutely continuous.
- That is, if there exists a function $f_X : \Re \to \Re$ such that

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt, \;\; \forall x$$

- f_X is called the probability density function (pdf) of X.
- Note that F_X here is continuous
- By the fundamental theorem of calculus, we have

$$\frac{dF_x(x)}{dx} = f_X(x), \; \forall x \text{ where } f_X \text{ is continuous}$$

Continuous Random Variables

- If X is a continuous rv then its distribution function, F_X, is continuous.
- Hence a discrete random variable is not a continuous rv!
- If a rv takes countably many values then it is discrete.
- However, if a rv takes uncoutably infinitely many distinct values, it does not necessarily imply it is of continuous type.
- As mentioned earlier, there would be many random variables that are neither discrete nor continuous.

Continuous Random Variables

- The df of a continuous rv is continuous.
- This implies
  $F_X(x) = F_X(x^+) = F_X(x^-)$
- Hence, if X is a continuous random variable then

$$P[X = x] = F_X(x) - F_X(x^-) = 0, \ \forall x$$

Continuous Random Variables

- A rv, X, is said to be continuous (or of continuous type) if its distribution function, F_X is absolutely continuous.
- The df of a continuous random variable can be written as

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt, \ \ \forall x$$

- This f_X is the probability density function (pdf) of X.

$$\frac{dF_x(x)}{dx} = f_X(x), \ \forall x \text{ where } f_X \text{ is continuous}$$

Probability Density Function

- The pdf of a continuous rv is defined to be the f_X that satisfies

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt, \quad \forall x$$

- Since $F_X(\infty) = 1$, we must have $\int_{-\infty}^{\infty} f_X(t)\, dt = 1$
- For $x_1 \leq x_2$ we need $F_X(x_1) \leq F_X(x_2)$ and hence we need

$$\begin{aligned}\int_{-\infty}^{x_1} f_X(t)\, dt \leq \int_{-\infty}^{x_2} f_X(t)\, dt \quad &\Rightarrow \quad \int_{x_1}^{x_2} f_X(t)\, dt \geq 0, \forall x_1 < x_2 \\ &\Rightarrow \quad f_X(x) \geq 0, \forall x\end{aligned}$$

Properties of pdf

- The pdf, $f_X : \Re \to \Re$, of a continuous rv satisfies
  A1. $f_X(x) \geq 0, \ \forall x$
  A2. $\int_{-\infty}^{\infty} f_X(t)\, dt = 1$
- Any f_X that satisfies the above two would be the probability density function of a continuous rv
- Given f_X satifying the above two, define

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt, \ \ \forall x$$

  This F_X satisfies
  1. $F_X(-\infty) = 0; \ F_X(\infty) = 1$
  2. F_X is non decreasing.
  3. F_X is continuous (and hence right continuous with left limits)
- This shows the the F_X is a df and hence f_X is a pdf

Continuous rv – example

- Consider a probability space with $\Omega = [0,\ 1]$ and with the 'usual' probability assignment (where probability of an interval is its length)
- Earlier we considered the rv $X(\omega) = \omega$ on this probability space.
- We found that the df for this is

$$F_X(x) = \begin{cases} 0 & \text{if } \ x < 0 \\ x & \text{if } \ 0 \leq x < 1 \\ 1 & \text{if } \ x \geq 1 \end{cases}$$

  This is absolutely continuous and we can get the pdf as

$$f_X(x) = 1 \ \text{ if } 0 < x < 1; \ (f_X(x) = 0, \ \text{ otherwise})$$

- On the same probability space, consider rv $Y(\omega) = 1 - \omega$.
- Let us find F_Y and f_Y.

- $Y(\omega) = 1 - \omega$.

$$\begin{aligned} [Y \leq y] &= \{\omega \;:\; Y(\omega) \leq y\} = \{\omega \in [0,\ 1] \;:\; 1 - \omega \leq y\} \\ &= \{\omega \in [0,\ 1] \;:\; \omega \geq 1 - y\} \\ &= \begin{cases} \phi & \text{if } \ y < 0 \\ \Omega & \text{if } \ y \geq 1 \\ [1 - y, 1] & \text{if } \ 0 \leq y < 1 \end{cases} \end{aligned}$$

- Hence the df of Y is

$$F_Y(y) = \begin{cases} 0 & \text{if } \ y < 0 \\ y & \text{if } \ 0 \leq y < 1 \\ 1 & \text{if } \ y \geq 1 \end{cases}$$

- We have $F_X = F_Y$ and thus $f_X = f_Y$. (However, note that $X(\omega) \neq Y(\omega)$ except at $\omega = 0.5$).

- Let X be a continuous rv.
- It can be specified by giving either F_X or the pdf, f_X.
- We can, in principle, compute probability of any event as

$$P[X \in B] = \int_B f_X(t)\, dt, \;\; \forall B \in \mathcal{B}$$

- In particular, we have

$$P[X \in [a,\ b]] = P[a \leq X \leq b] = \int_a^b f_X(t)\, dt = F_X(b) - F_X(a)$$

- Since the integral over the open or closed intervals is the same, we have, for continuous rv,

$$P[a \leq X \leq b] = P[a < X \leq b] = P[a \leq X < b] \text{ etc.}$$

- Recall that for a general rv

$$F_X(b) - F_X(a) = P[a < X \leq b]$$

- If X is a continuous rv, we have

$$P[a \leq X \leq b] = \int_a^b f_X(t)\, dt$$

- Thus

$$P[x \leq X \leq x + \Delta x] = \int_x^{x+\Delta x} f_X(t)\, dt \approx f_X(x)\, \Delta x$$

- That is why f_X is called probability density function.

- For any random variable, the df is defined and it is given by
$$F_X(x) = P[X \leq x] = P[X \in (-\infty,\ x]\ ]$$
- The value of $F_X(x)$ at any x is probability of some event.
- The pmf is defined only for discrete random variables as $f_X(x) = P[X = x]$
- The value of pmf is also a probability
- We use the same symbol for pdf (as for pmf), defined by
$$F_X(x) = \int_{-\infty}^{x} f_X(x)\ dx$$
- Note that the value of pdf is not a probability.
- We can say $f_X(x)\ dx \approx P[x \leq X \leq x + dx]$

A note on notation

- The df, F_X, and the pmf or pdf, f_X, are all functions defined on $\Re$.
- Hence you should not write $F_X(X \leq 5)$.
  You should write $F_X(5)$ to denote $P[X \leq 5]$.
- For a discrete rv, X, one should not write $f_X(X = 5)$.
  It is $f_X(5)$ which gives $P[X = 5]$.
- Writing $f_X(X = 5)$ when f_X is a pdf, is particularly bad.
  Note that for a continuous rv, $P[X = 5] = 0$ and $f_X(5) \neq P[X = 5]$.

- A continuous random variable is a probability model on uncountably infinite Ω.
- For this, we take $\Re$ as our sample space.
- We can specify a continuous rv either through the df or through the pdf.
- The df, F_X, of a cont rv allows you to (consistently) assign probabilities to all Borel subsets of real line.
- We next consider a few standard continuous random variables.

Uniform distribution

- X is uniform over $[a,\ b]$ when its pdf is

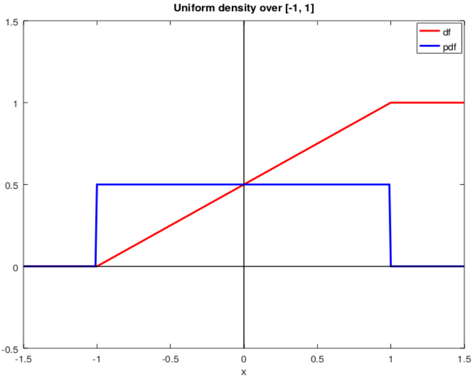
$$f_X(x) = \frac{1}{b-a}, \ \ a \le x \le b$$

($f_X(x) = 0$ for all other values of x).

- Uniform distribution over open or closed interval is essentially the same.
- When X has this distribution, we say $X \sim U[a,\ b]$
- By integrating the above, we can see the df as

$$F_X(x) = \begin{cases} \int_{-\infty}^{x} f_X(x)\ dx = \int_{-\infty}^{x} 0\ dx = 0 & \text{if } \ x < a \\ \int_{-\infty}^{a} 0\ dx + \int_{a}^{x} \frac{1}{b-a}\ dx = \frac{x-a}{b-a} & \text{if } \ a \le x < b \\ 0 + \int_{a}^{b} \frac{1}{b-a}\ dx + 0 = 1 & \text{if } \ x \ge b \end{cases}$$

- A plot of density and distribution functions of a uniform rv is given below



- Let $X \sim U[a,\ b]$. Then $f_X(x) = \frac{1}{b-a},\ a \leq x \leq b$
- Let $[c,\ d] \subset [a,\ b]$.
- Then $P[X \in [c,\ d]] = \int_c^d f_X(t)\ dt = \frac{d-c}{b-a}$
- Probability of an interval is proportional to its length.
- The earlier examples we considered are uniform over $[0,\ 1]$.

Exponential distribution

- The pdf of exponential distribution is

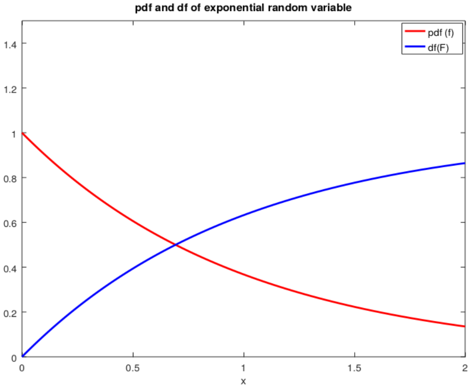
$$f_X(x) = \lambda\, e^{-\lambda x}, \;\; x > 0, \; (\lambda > 0 \text{ is a parameter})$$

(By our notation, $f_X(x) = 0$ for $x \le 0$)
- It is easy to verify $\int_0^\infty f_X(x)\, dx = 1$.
- It is easy to see that $F_x(x) = 0$ for $x \le 0$.
- For $x > 0$ we can compute F_X by integrating f_X:

$$F_X(x) = \int_0^x \lambda\, e^{-\lambda x}\, dx = \lambda \left. \frac{e^{-\lambda x}}{-\lambda} \right|_0^x = 1 - e^{-\lambda x}$$

- This also gives us: $P[X > x] = 1 - F_X(x) = e^{-\lambda x}$ for $x > 0$.

- A plot of density and distribution functions of an exponential rv is given below



exponential distribution is memoryless

- If X has exponential distribution, then, for $t, s > 0$,

$$P[X > t+s] = e^{-\lambda(t+s)} = e^{-\lambda t} \, e^{-\lambda s} = P[X > t] \, P[X > s]$$

- This gives us the memoryless property

$$P[X > t + s \mid X > t] = \frac{[P[X > t + s]}{P[X > t]} = P[X > s]$$

- Exponential distribution is a useful model for, e.g., life-time of components.
- If the distribution of a non-negative continuous random variable is memory less then it must be exponential.

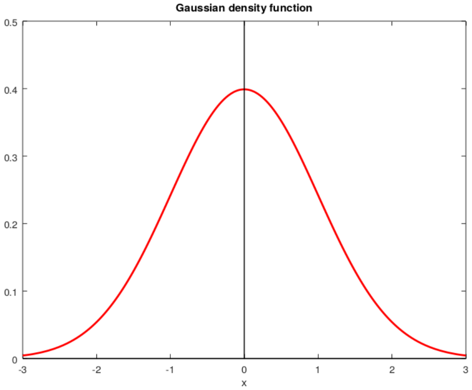
Gaussian Distribution

- The pdf of Gaussian distribution is given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \; -\infty < x < \infty$$

  where $\sigma > 0$ and $\mu \in \Re$ are parameters.
- We write $X \sim \mathcal{N}(\mu, \sigma^2)$ to denote that X has Gaussian density with parameters μ and σ.
- This is also called the Normal distribution.
- The special case where $\mu = 0$ and $\sigma^2 = 1$ is called standard Gaussian (or standard Normal) distribution.

- A plot of Gaussian density functions is given below



- $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ -\infty < x < \infty$
- Showing that the density integrates to 1 is not trivial.
- Take $\mu = 0, \sigma = 1$. Let $I = \int_{-\infty}^{\infty} f_X(x)\, dx$. Then

$$
\begin{aligned}
I^2 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-0.5x^2}\, dx \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-0.5y^2}\, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-0.5(x^2+y^2)}\, dx\, dy
\end{aligned}
$$

- Now converting the above integral into polar coordinates would allow you to show $I = 1$.
  (Left as an exercise for you!)