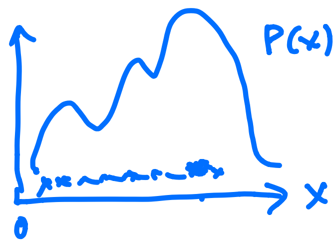
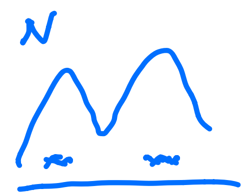
W00 Sampling from a probability distribution

X – variable

$P(x)$ – probability density function.

$\{x_i\}_{i=1}^N$ $x_1, \dots, x_N \sim$ according to the $P(x)$



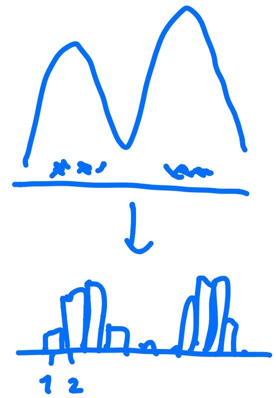


N large

$$\langle f(x) \rangle_p = \int f(x) \cdot p(x) \cdot dx$$

$\{x_i\}_{i=1}^{N}$

$$\simeq \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$



$$\simeq \frac{1}{N} \sum_b N_b f(x_b)$$

$$= \sum_b \frac{N_b}{N} f(x_b)$$

$p(x_b)$

$f(x) = (x-\mu)^2$

$\langle f(x) \rangle = \frac{1}{N} \sum_i (x_i - \mu)^2$

$= \sigma^2$

$f(x) = x$

$\mu = \langle x \rangle = \frac{1}{N} \sum_i x_i$ "

Why sampling?

i) to generate alternative hypothesis.

$\{data\}$ — model $P(data|model)$

data fit H_0 $\{samples\ H_0\}$?

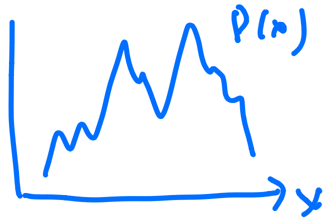
ii) $P(x)$ is complicated

$\{x_i\}$ allows do averages of any function.

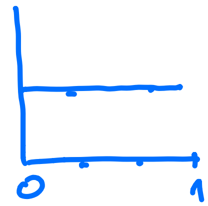
$$\langle f(x) \rangle \simeq \frac{1}{N} \sum_i f(x_i)$$

iii) $P(y)$ $P(m)$ → sample parameter values m.

Main point. an arbitrary $P(x)$



$r = \text{random}() \in [0,1]$



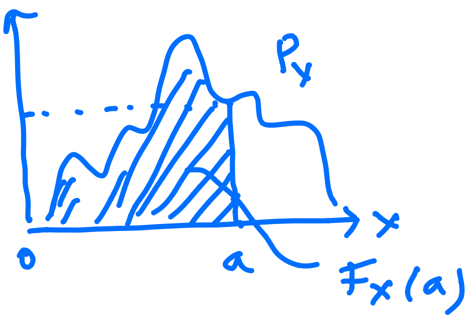
$\{x_i\}$ fm P $\longrightarrow$ Uniform distribn $[0,1]$

PDF, CDF

PDF = probability density function $P_X(x) = P(X = x)$

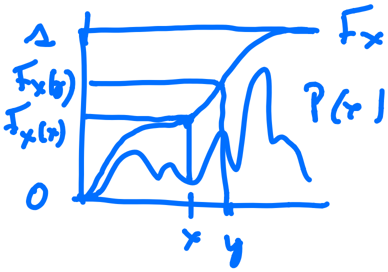
CDF = cumulative " " $F_X(x) = P(X \le x)$

$P_X(a) = P(X = a)$



$$F_X(a) = \int_0^a P_X(x)\,dx$$

$$= \sum_{x=0}^{a} P(x)$$

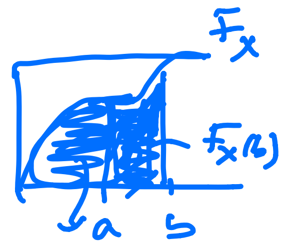


$$\lim_{x \to -\infty} F_X(x) = 0$$

$$\lim_{x \to +\infty} F_X(x) = 1$$

- always increasing.

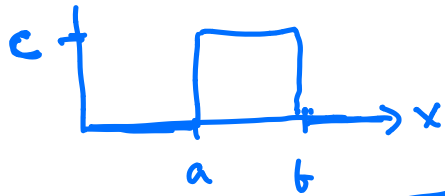
$x \leq y \quad F_X(x) \leq F_X(y)$



$F_X(a) = P(X \leq a)$

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

The Uniform Distribution U in $x \in [a,b]$



$$P_U(x) = \begin{cases} 0 & x<a \\ c & a \le x \le b \\ 0 & x>b \end{cases}$$

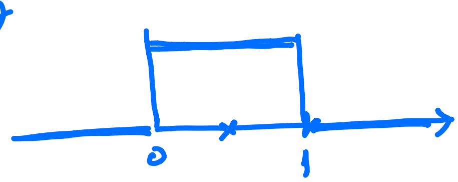
$a \le x \le b$ $\quad P_U(x) = P_U(y) = c$

$\quad\ \ y$

$P_U($

$a=0$

$b=1$



① c $(a, b$?

② $F_U(x)$

U $$P_U(x) = \begin{cases} 0 & x < a \\ c & a \le x \le b \\ 0 & x > b. \end{cases}$$

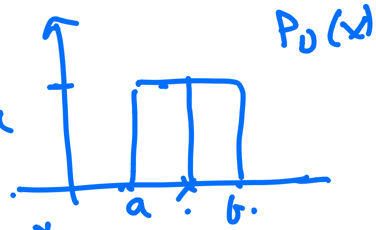
ⓒ? $F_U(x)$

① $\underline{c} = 1$

$$P_U(x) = \begin{cases} 0 & x < a \\ c & a \le x \le b \\ 0 & x > b. \end{cases}$$

② $\underline{F_U(x)}$

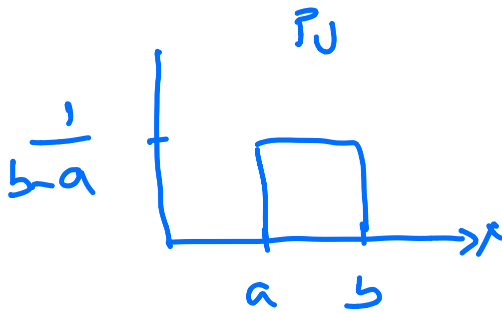
$c = \frac{1}{b-a}$

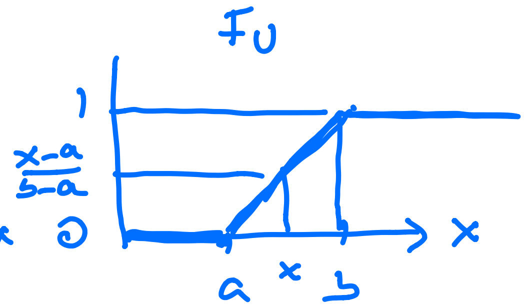


$$F_U(x) = \int_a^x P_U(x)\,dx = \int_a^x c\,dx = c(x-a)$$

$$x=b \quad \underline{F_U(x=b) = 1} = \int_a^b c\,dx = c(b-a) = 1$$

$$c = \frac{1}{b-a}$$





$$F_U(x) = c(x-a) = \frac{x-a}{b-a}$$

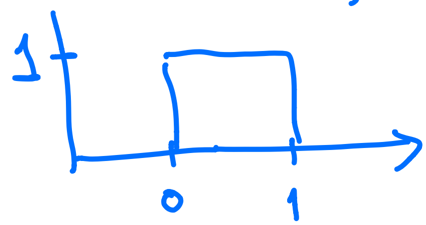
$x=a \quad F_U=0$

$x=b \quad F_U=1$

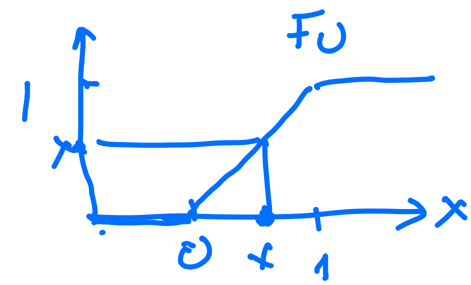
$\left.\begin{array}{l} a=0 \\ b=1 \end{array}\right]$ random

$a=0$ Standar Uniform dist $[0,1]$

$b=1$ $P_U(x) = 1 \quad (0 \le x \le 1)$



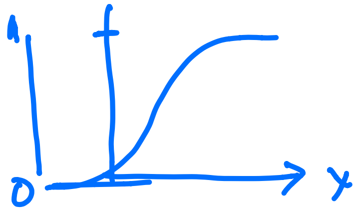
$$F_U(x) = \frac{x-a}{b-a} = x$$



$$F_U(x) = x$$

$0 \le x \le 1$

$X \quad F_X \quad 0 \le \underbrace{F_X(x)}_{\text{CDF - random \#}} \le 1$



$$P(U \le F_X(x)) = F_X(x)$$

$P_U(x) = x$

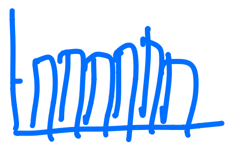
$\boxed{F_X(x) \text{ is } U[0,1]}$

$\boxed{F_X(x)}$ random variable.

$\{X_1, \dots X_N\}$ $P(x)$.

$\downarrow$

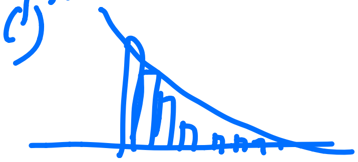
$u_1 = F_X(x_1)$, ... $u_i = F_X(x_i)$, ..., $u_N = F_X(x_N)$



1000

xs = scipy.stats.expon.rvs( )

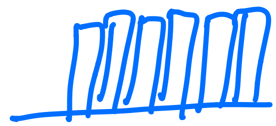
hist(xs)



$F_E(x)$

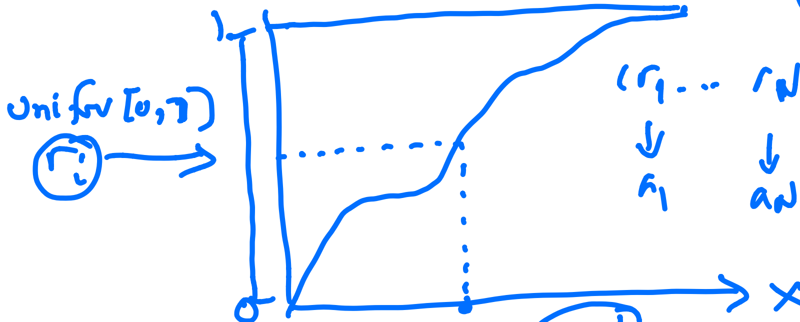
ys = scipy.stats.expon.cdf(xi)

hist(ys)



The Inverse Tranformation Method

F_X



$(r_1 \dots r_N)$ $U[0,1]$

$\downarrow \qquad \downarrow$

$a_1 \quad a_N$ sampled F_X

$a_i = F_X^{-1}(r)$

$F_X(a) = r$

$F_X(x)$ uniform.