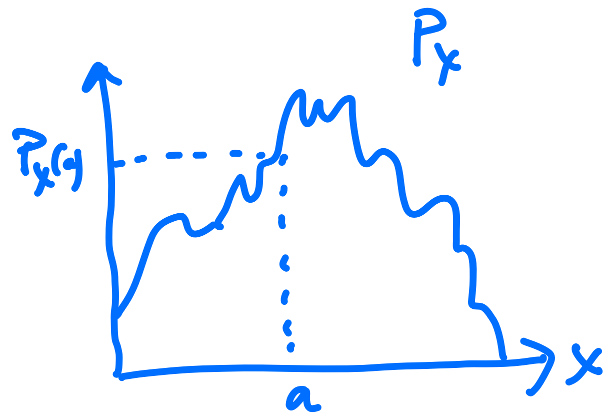
# W00 - Sampling from a probability distribution

## Summary



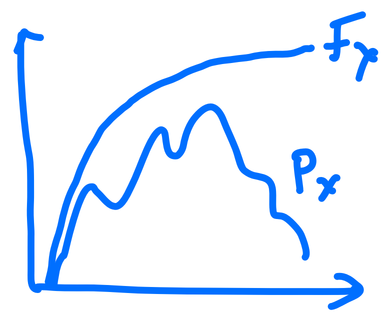

$X$ - random variable

$a \in X$

$x_{min} \leq a \leq x_{max}$

PDF: $P_X(a)$ ,, $P_X(a) \geq 0$ ,, $\displaystyle\int_{x_{min}}^{x_{max}} P_X(y)\,dy = 1$

CDF $$F_X(a) =: \int_{x_{min}}^{a} P_X(y)\,dy$$




$F_X(x_{min}) = 0$

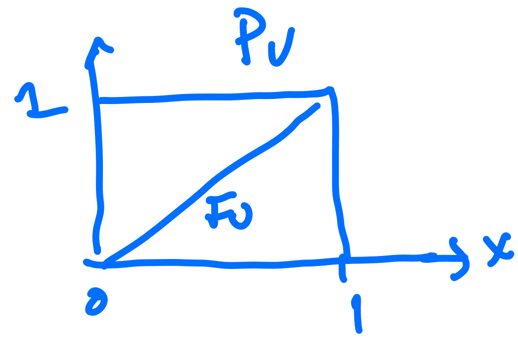
$F_X(x_{max}) = 1$

$a < b \Rightarrow F_X(a) \leq F_X(b)$

$$F_X(b) - F_X(a) = P(a \leq X \leq b) = \int_a^b P_X(z)\,dz$$

## The Uniform distribution

$U[0,1]$




$p_U(r) = 1 \ \forall r \in [0,1]$

$F_U(r) = P(U < r) = r$

For any $P_X$ on $X$:

$$F_X(x) = U[0,1]$$

Introduce $A = F_X(x)$

let's calculate

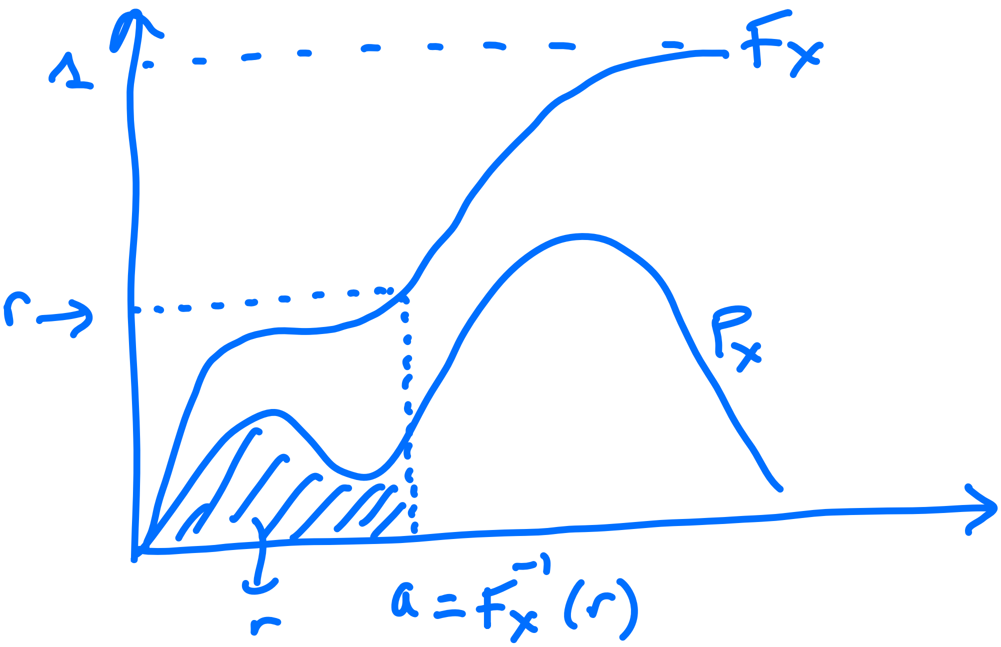
$$F_A(r) = P(A \le r) = P(F_X(X) < r)$$

$$= P(X \le F_X^{-1}(r))$$

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

$$F_A(r) = r \Rightarrow A = F_X(x) = U[0,1]\,!$$

# The inverse transformation method




$$r = F_X(a) = P_X(X \leq a)$$

$\{r_1 \ldots r_N\} \longrightarrow$ fm $U[0,1]$

$\downarrow \qquad \downarrow F_X^{-1}$

$a_1 \ldots a_N \longrightarrow$ fm $X$

① let's see it's fine

Practical example in python.

② Proof

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

Consider random variable

$$A = F_X^{-1}(U)$$

$$P(A < a) = P(F_X^{-1}(U) < a)$$

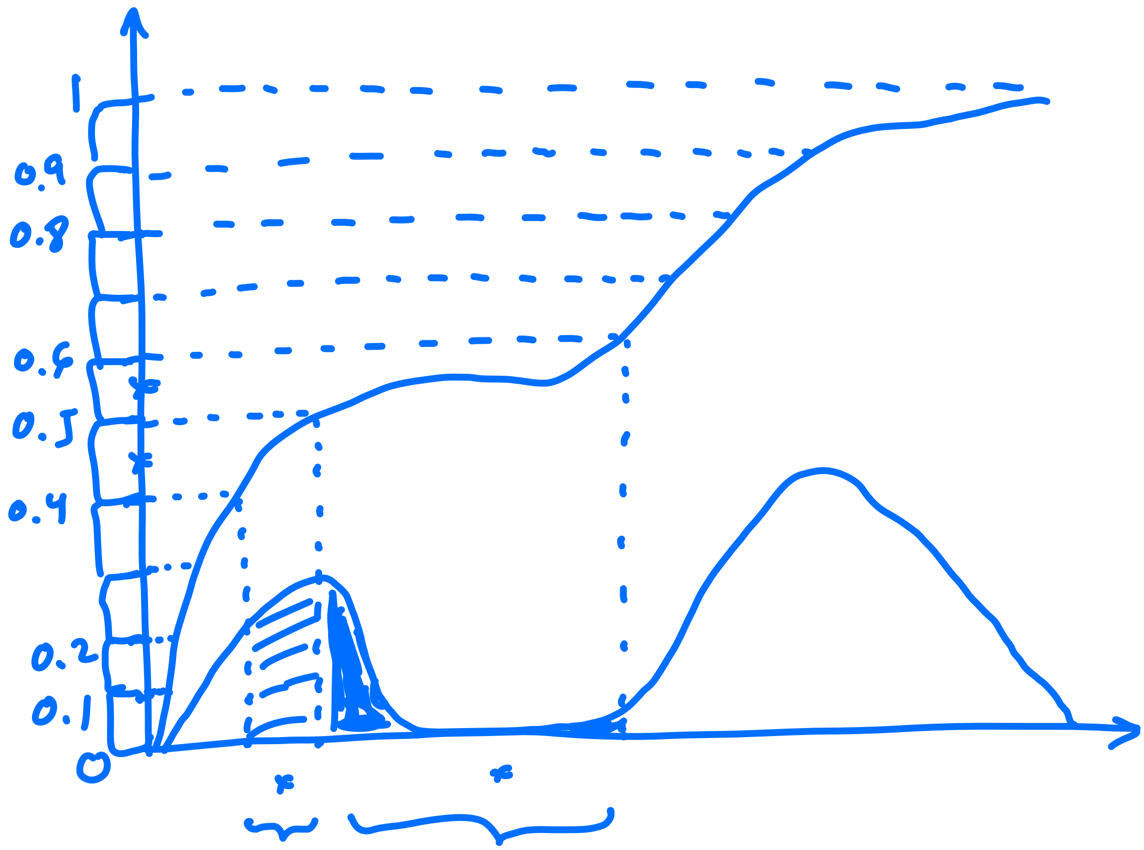
$$= P(U < F_X(a))$$

$\{P_U(r) = r\}$

$$= F_X(a)$$

then $A = X$

# More intuition



# The rejection method

- does not require $F_X$, only $P_X$
- requires another distribution $G_X$

$$P_X(x) = p(x) \quad \forall x \in X$$

$$G_X(x) = g(x) \sim \text{easy to sample from.}$$

condition $p(x) \leq M\, g(x)$


i) Sample $x_i$ from $G_X$ [inverse method for instance]

ii) sample $r_i \in U[0,1]$

iii) if $r_i < \dfrac{p(x_i)}{M g(x_i)} \rightarrow$ accept $x_i$ as sample of $P_X$

else reject $x_i$

# Importance Sampling

- doest not sample $P_x$, but allows to calculate
$$\langle f \rangle_{P_x}$$
- requires another distribution $G_x$

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

$$= \int f(x) \frac{p(x)}{g(x)}\, g(x)\, dx$$

$$= \left\langle f \cdot \frac{p}{g} \right\rangle_g$$

i) take sample for $G_x$ $\{x_1 .. x_N\}_G$

ii) calculate

$$\langle f \rangle_p = \frac{1}{N} \sum_i f(x_i) \frac{p(x_i)}{g(x_i)}$$

# Monte Carlo Sampling

To construct a random process to approximate probability distributions

$\{x_1 \dots x_N\}$ sample of $P_X$

Used for

- estimate a density
  $$P(x_b) = \frac{N_b}{N}$$
- approximate a quantity
  $$\langle f \rangle = \frac{1}{N} \sum_i f_i$$
- optimize a function

direct sampling, importance sampling, rejection sampling are examples of MC sampling.

# Markov Chain Monte Carlo (MCMC)

A MC sampling method

where the samples are correlated

"walkers"