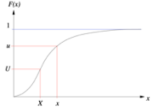
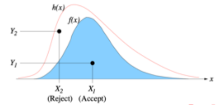
Distribution sampling

For r.v. X with support $supp(X)$, the objective is to define a sampling function: $U(0,1) \rightarrow supp(X)$ in terms of X's density/cdf (or pmf/cdf) function.

1. **The Inverse Transform method:** Suppose X is a continuous r.v. with cdf $F(x) = P(X \leq x)$ and that we are trying to sample X, Let U ~ U(0,1). Because F(x) increases monotonically, we have: $P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$, so **set $\boldsymbol{U = F(X)}$ and invert to give $\boldsymbol{X = F^{-1}(U)}$.**





(For Method 1) (For Method 2)

We can use the inverse transform method to sample a discrete r.v., X by inverting its cumulative distribution function, $F_X(x)$ (a"step function"). If $U \sim U(0,1)$, then the inverse transform methods returns $X = min\{x : F(x) \geq U\}$.

2. **The Acceptance-Rejection (AR) Method:** If F(x) cannot be explicitly inverted (e.g., normal cdf) we can sometimes work with the corresponding density function f(x). We choose a density function g(x) easy to sample from. Now we try to find a constant, c, so that **c g(x) = h(x) dominates f (x) for all x (i.e. c g(x)>f(x))**:

By construction, c is the area under h(x): $c = c\int_x g(x)dx = \int_x h(x)dx$.

To find c: We need to maximize $\frac{f(x)}{g(x)}$: Differentiate $\frac{f(x)}{g(x)}$, let $\frac{d}{dx}\frac{f(x)}{g(x)} = 0$ to find the maximum value of x, calculate c using the value of x ($c = max\frac{f(x)}{g(x)}$).