① Inverse sampling (easiest)

$P(x) = $ pdf

$x_1, x_2 \dots x_n$

$P(X = x_i) = p_i$

cdf = cumul partition fn

$F(x_1) = p_1$

$F(x_2) = p_1 + p_2$

$F(x_3) = p_1 + p_2 + p_3$

$\sum p_i = 1$

contin

$F = \int_0^a f(x)\,dx$ monot. increasing

- choose $r = \text{unif}(0,1)$ y axis r
- compute inverse $F^{-1}(r)$
  $\Rightarrow$ selected item/value

Compute

1) inverse-math(cdf)

② Binary Search

Sampling: from dist $\bar{p} = (p_1, p_2, \dots p_n)$

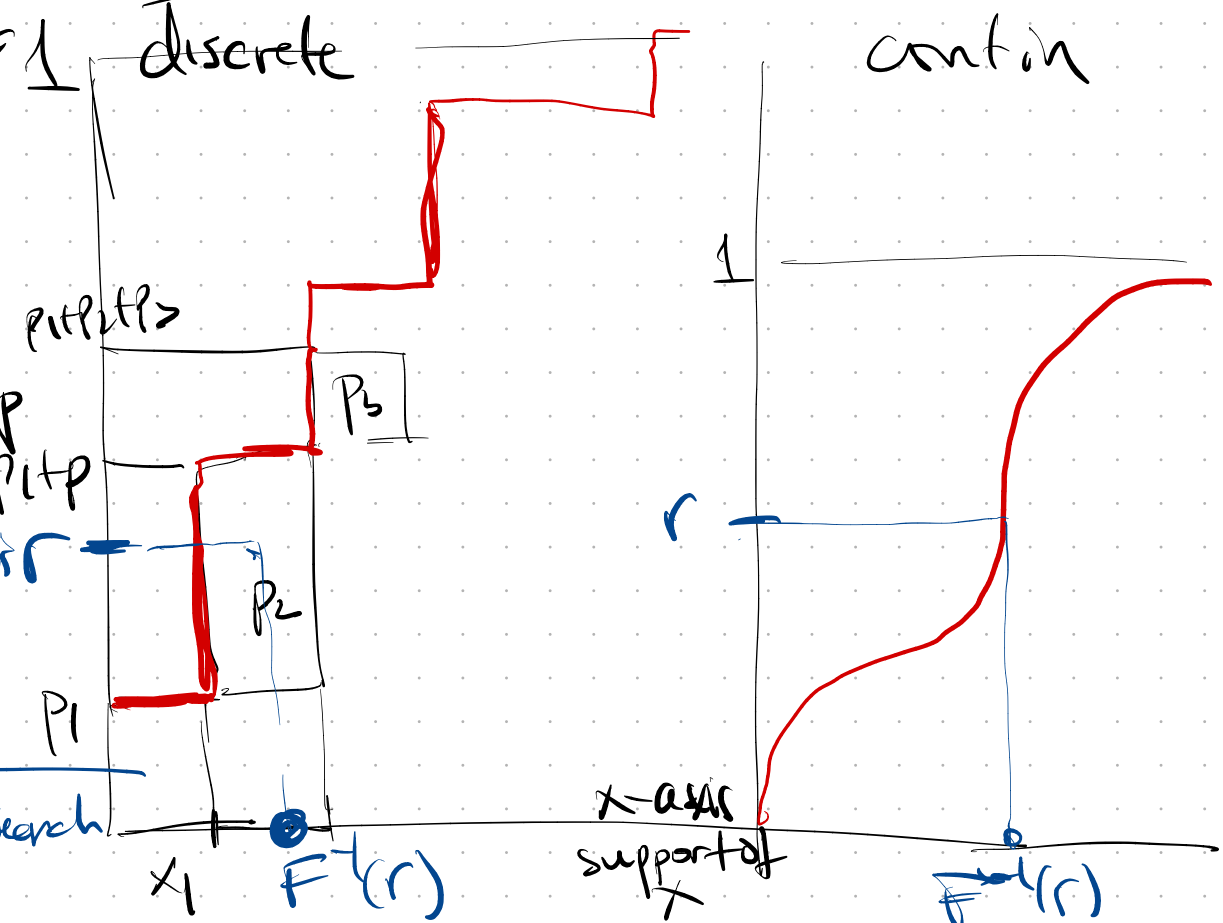
Select item k with probas p_k

"outcome $X = R.V \sim p$

p uniform $\Rightarrow$ trivial to sample

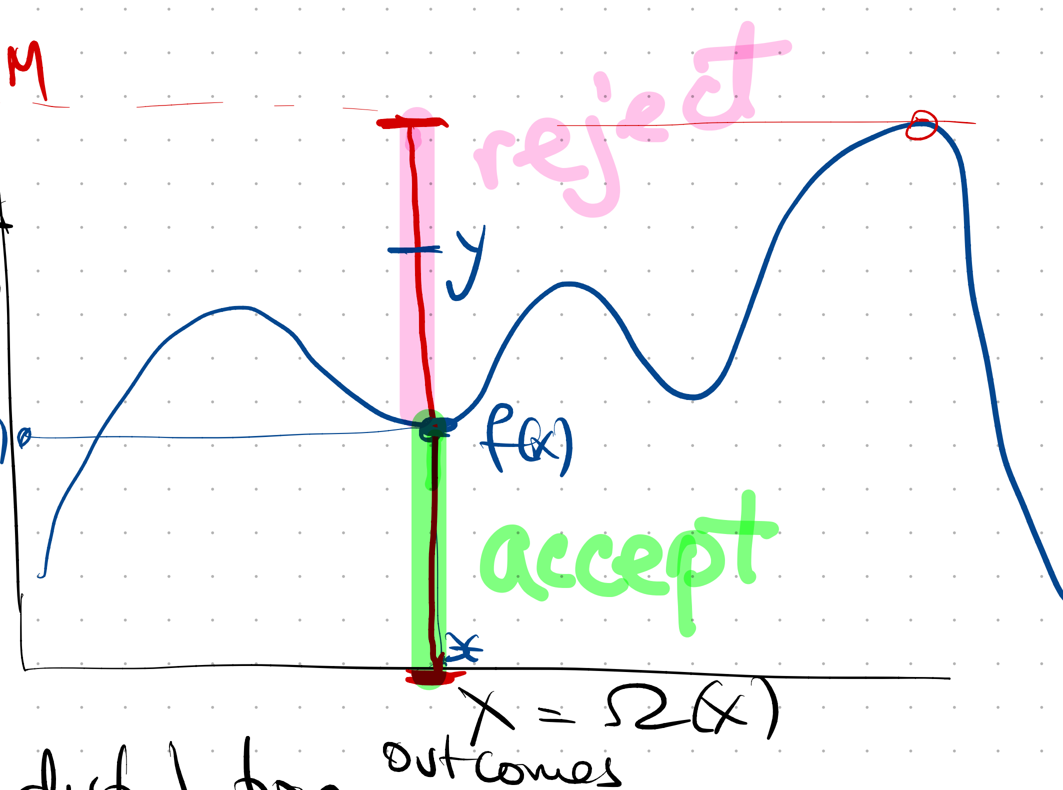
$p_i = \frac{1}{n}$ just an item unif random

practice: shuffle$(1, 2, \dots n)$, pick first



Rejection sampling (accept-reject) $f = pdf(x)$

- sample $x \in \Omega(x)$ outcome
- look at vertical line $[0\ M]$ on x
- sample uniform $y \in [0, M]$ density prop
  - if $y > f(x)$ reject x
  - if $y \le f(x)$ accept x
- repeat until accept (resample x)



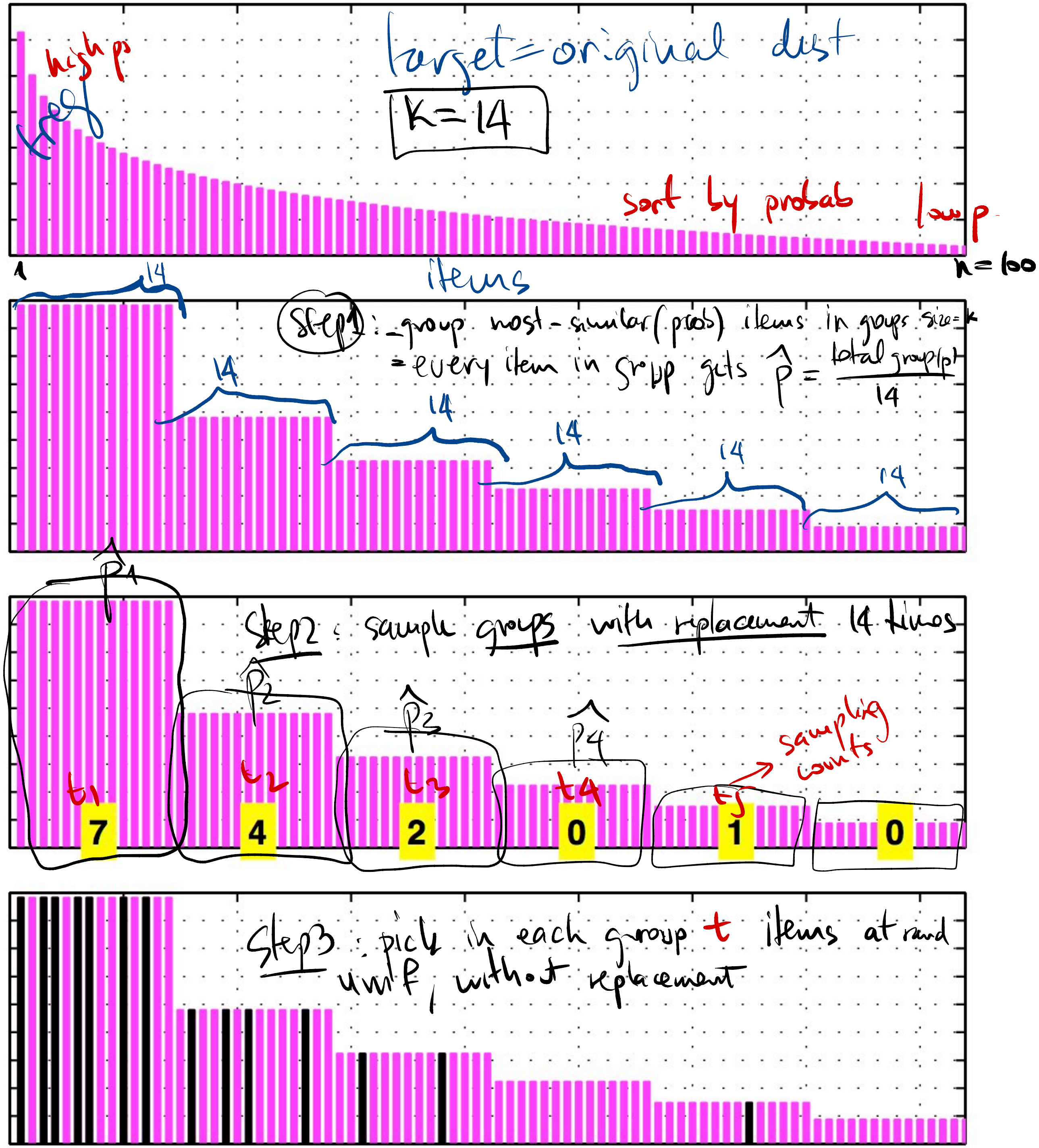
general:
- sample x from g distribution
- r = uniform$(0, 1)$
- accept if $r < \frac{f(x)}{g(x)} \cdot M$
- repeat untill accept

③ Sample K items without repeat from non-unf dist. $p_1, p_2 \cdots p_n$

- not possible exact
- K large ($K \to n$) $\Rightarrow$ all items will be selected (prob = 100%)
- $K \ll n$ we want prob $\sim p$

with repetition repl

- sample 1 item
- repeat K times


target=original dist
K=14
high p
freq
sort by probab
low p
1
n=100
items
14
Step1: group most-similar (prob) items in groups size=k
= every item in group gets $\hat{p} = \frac{\text{total group(p)}}{14}$
14
14
14
14
14
Step2: sample groups with replacement 14 times
$\hat{p}_1$
$\hat{p}_2$
$\hat{p}_3$
$\hat{p}_4$
t1
t2
t3
t4
t5
sampling counts
7
4
2
0
1
0
Step3: pick in each group t items at rand unif, without replacement