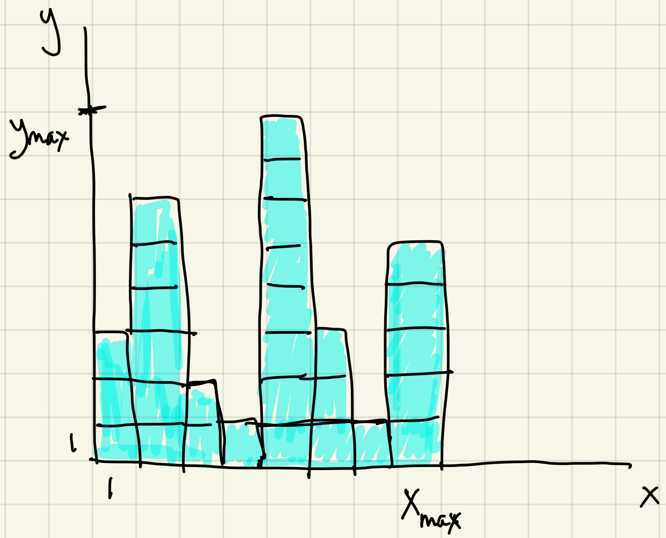
The "accept-reject" method

(AKA rejection sampling)

Given 2-dimensional domain



(pick from uniform on much bigger domain)

Goal output random (uniformly distributed) blue block

Naive rejection sampling algorithm:

Repeat forever:
 Pick random $x \in [x_{max}]$
 $y \in [y_{max}]$

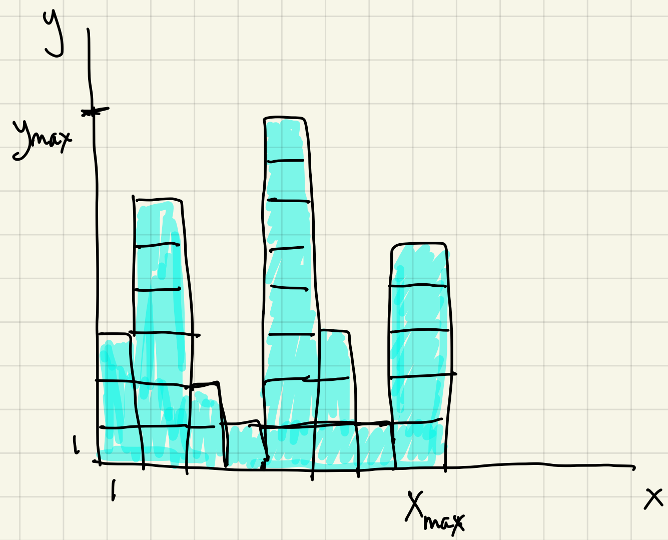
 If (x,y) is a blue block, output (x,y) + halt

Analysis: output (x,y) with prob $\frac{1}{\#\text{blue blocks}}$

Expected runtime per output: $\frac{1}{\text{fraction covered by blue blocks}} = \frac{x_{max} \cdot y_{max}}{\#\text{blue blocks}}$

Better algorithm?

assume we know # blue blocks at every x: b_x



Repeat forever

- pick random $x \in [X_{max}]$
- pick random $y \in [b_x]$
- toss coin with prob $\frac{b_x}{\max_y b_y}$
  - if heads, output (x,y) + halt

Analysis: pick (x,y) with prob $\frac{1}{X_{max}} \cdot \frac{1}{b_x}$

output (x,y) with prob $\frac{1}{X_{max}} \cdot \frac{1}{b_x} \cdot \frac{b_x}{\max_y b_y} = \frac{1}{X_{max} \cdot \max_y b_y}$ $\Big\}$ Same for each blue block so uniform distribution

Expected runtime per output: $\frac{X_{max} \cdot \max_y b_y}{\underbrace{\#\text{blue blocks}}_{X_{max} \cdot \text{ave } b_y}}$ $\Big\}$ when is this better?

e.g. if average $b_y \geq \frac{1}{2} \max_y b_y$

then this is 2.

Example in class:

after bucketing nodes by degree,

within a bucket all nodes had degree

$$(1+\beta)^i \le \text{degree} \le (1+\beta)^{i+1}$$

so

$$\text{max degree} \le (1+\beta)^{i+1}$$
$$\text{ave degree} \ge (1+\beta)^i$$