Welcome back...

..to me.

Test out !!!

Don't worry. Be happy.

Look at instructions.
No collaboration.
Private message on piazza.
Note: Content can be declassified.

Turn in by Monday.
Grade by Wednesday .. night ...late ..hopefully.

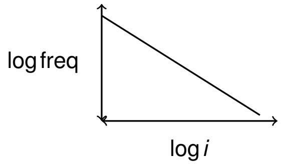
Try to get it in then or soon after!

Pareto:

20% of pods have 80% of peas.
20% of peple have 80% of land.

City populations:

ith largest city has population $\frac{p_1}{i}$.



Zipf's law. Zipf's graph.

Not a distribution.

As a distribution.

Pareto.
Income$_i \propto \frac{\text{income}_1}{i^\beta}$.
Bill Gates...then someone much less. Prelude: why? Rich get richer?
Distribution:
Pareto.
$Pr[X \geq x] \propto x^{-\alpha+1}$.
Survival function.
Note: "p.d.f." $Pr[X = x] \propto x^{-alpha}$.
See Adamic for comment on estimating for real data.
http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html

MAKE SOME DRAWINGS.

Pareto to Zipf

Zipf:
ith guy has $C\frac{1}{i^\beta}$

N people.
How many people have value more than x_i?

On expection? $NDx^{-\alpha+1}$.
ith guy has more than x_i
$\equiv i$ guys have more than x_i
$i \approx NDx_i^{-\alpha+1}$

$x_i = \frac{1}{i^{1/(1-\alpha)}}$

Relationship: $\beta = \frac{1}{1-\alpha}$

Self similarity.

Power laws.
No matter where you are there you are...

$x_{t+1} = x_t \times \gamma$.

Actually $\gamma_t \approx (1 + \beta/t)$.

Roughly constant for interval of wdith β.

Power law and philosophy.

Wow! Power laws. Cool.

Zipf: for frequency of words. For all languages!!!

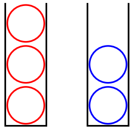
Must have something to do with the brain!

Wentian Li.

Document: "A quick brown fox jumps over the ...."

Permute the letters at random..and get a power law!!!

Polya Urns



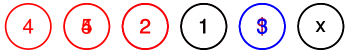
Choose bin uniformly at random.
Load on red bin?
Expectation? $n/2$
Within $n/2 \pm \sqrt{n}$ with good probability.
Approximately Gaussian with variance $\sqrt{n}/2$

Choose red bin with probability $\frac{r+1}{r+b+2}$
Expectation? $n/2$
Distribution?
Guesses?
Uniform! !!!

Permutations

Choose bin with probability $\frac{r+1}{r+b+2}$.
Claim: After n balls the $Pr[i \text{ red}] = \frac{1}{n+1}$.

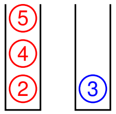
Analyse?Another process.
Start with two balls, insert n more.



Where is ball 1? Position 4.
How many red balls? 3.

Insert n balls, where oh where is ball 1?
Random permuation. Position $i \in [1, n+1]$ with prob. $\frac{1}{n+1}$
How many red balls? $j = i-1 \in [0, n]$ with prob. $\frac{1}{n+1}$.

Balls in bins? Yes!
Allocation (r, b):
choose one of $r+b$ balls or 2 bottoms.
place in corresponding bin.
$\Pr[\text{red}] = \frac{r+1}{r+b+2}$



Red balls have same distribution in two processes.

More bins.

m bins.

Uniformly at random.
Max load: $\frac{n}{m} + \sqrt{\frac{n}{m}\log n}$
Min load: $\frac{n}{m} - \sqrt{\frac{n}{m}\log n}$

Preferential Selection:
Max load: $\frac{n}{m}\log n$
Min load: n/m^2

Analysis: random permutation with m separators.
Analyse min and max size of interval.
Roughly: $(1/m)$ probability of stopping at any point.

Router Graph:
Average degree: 4
Max Degree? Uniformly random $\implies$ $Pr[\text{degree} \geq 20] \approx 10^{-4}$.
Actual high degree nodes more common:
5% of nodes have degree greater than 20.

Internet graph:
Average degree: 12.
Degree ≥ 100 with prob. $\leq 10^{-6}$.
Actual: 1% greater than 100.
Some very large.

Processes?
Preferential Attachment.
For routers?
Connect at random. Not!
For the internet graph?
Degrees too large for even that.

Internet: copy links.

Surf. Cool page. Link for mine.

Model:
Pick a random neighbor.
Copy all links.

Random Graph with average degree 4.
Plus Copy process $\rightarrow \sqrt{n}$

Routers.

Connection Game.

Process Distance:
Arrive randomly at point on unit square.
Connect to closest node.

Generate tree with average degree 1.
Max degree? $O(\log n)$.

Process Hops:
Arrive randomly at point on unit square.
Connect to first node.

Max degree? $n-1$.

Process Distance/Hops:
Arrive randomly at point on unit square.
Connect to node with $\min_{j<i} \alpha d_{ij} + h_j$.

Power law if $c \leq \alpha \leq \sqrt{n}$, $\rightarrow$ power law!

See you ...

Thursday.