CS 112: Computer System Modeling Fundamentals

Prof. Jenn Wortman Vaughan

April 12, 2011

Lecture 5

Reminders & Announcements

- Quiz grades are posted on Gradebook
- Homework 2 will be posted on the course website later this week and will be due on Tuesday, April 26

Last Time

- Random variables
- Expectation
- Examples of discrete random variables
  - Bernoulli random variables
  - Binomial random variables
  - Geometric random variables
  - Poisson random variables

Today

- Computing expected values
- Variance, and how to compute it
- Using expectation to make decisions

- Zipf's law and power law distributions (not in the book)

Random Variables

A random variable is a real-valued function of the outcome of an experiment…. though we won't always write them as a function.

Example: Suppose we roll two six-sided dice.

- Sample space $\Omega = \{11, 12, 13, ..., 64, 65, 66\}$
- Random variables: the sum of the two rolls, the number of sixes rolled, the product of the two rolls, the value of the first roll, the number of even rolls, etc.

Probability Mass Functions

The probability mass function (PMF) of a random variable X is denoted p_X. For any value k that X can take on,

$$p_X(k) = P(X=k) = P(\{X=k\})$$

Example: Suppose we flip a fair coin twice. Let X be the number of heads. Then

$$p_X(0) = 1/4$$
$$p_X(1) = 1/2$$
$$p_X(2) = 1/4$$

Expectation

The expected value of a random variable X with probability mass function p_X is

$$E[X] = \sum_x x p_X(x)$$

We sometimes call this the expectation or mean of X.

This is simply the value we'd expect to get for X "on average" if we repeated the same experiment many times.

Computing Expectations

- With probability 0.6 there is low traffic and it takes me 30 minutes to get to work. The rest of the time there is high traffic and it takes me 50 minutes. How long does it take me to get to work on expectation?

Computing Expectations

- Let X be a random variables, and let Y = g(X) for an arbitrary function g. What is E[Y]?

Computing Expectations

- Let X be a random variables, and let Y = *g*(X) for an arbitrary function *g*. What is E[Y]?

Hint: By the additivity axiom, $p_Y(y) = \sum_{\{x \mid g(x) = y\}} p_X(x)$

Computing Expectations

- Let X be a random variables, and let Y = *g*(X) for an arbitrary function *g*. What is E[Y]?

Hint: By the additivity axiom, $p_Y(y) = \sum_{\{x \mid g(x)=y\}} p_X(x)$

Can compute E[Y] without computing p_Y directly!

Computing Expectations

- Let X be a random variables, and let Y = g(X) for an arbitrary function g. What is E[Y]?

  Hint: By the additivity axiom, $p_Y(y) = \sum_{\{x \mid g(x)=y\}} p_X(x)$

Can compute E[Y] without computing p_Y directly!

- Let X be a random variable, and let Y = aX + b for any scalars a and b. What is E[Y]?

Variance

The variance of a random variable measures the dispersion of X around its mean. Let $\mu = E[X]$. Then

$$\mathrm{var}(X) = E\left[(X-\mu)^2\right]$$

Variance

The variance of a random variable measures the dispersion of X around its mean. Let $\mu = E[X]$. Then

$$\text{var}(X) = E\left[(X-\mu)^2\right] = E[X^2] - \mu^2$$

Variance

The variance of a random variable measures the dispersion of X around its mean. Let $\mu = E[X]$. Then

$$\text{var}(X) = E\left[(X-\mu)^2\right] = E[X^2] - \mu^2$$

- Let's first verify that these expressions are equivalent…

Variance

The variance of a random variable measures the dispersion of X around its mean. Let $\mu = E[X]$. Then

$$\mathrm{var}(X) = E\left[(X-\mu)^2\right] = E[X^2] - \mu^2$$

- Let's first verify that these expressions are equivalent…
- Note that the variance is never negative. (When is it zero?)

Variance

The variance of a random variable measures the dispersion of X around its mean. Let $\mu = E[X]$. Then

$$\mathrm{var}(X) = E\left[(X-\mu)^2\right] = E[X^2] - \mu^2$$

- Let's first verify that these expressions are equivalent…
- Note that the variance is never negative. (When is it zero?)

The standard deviation of X is denoted σ_X, and is the square root of the variance. σ_X has the same units as X.

Computing Variance

- With probability 0.6 there is low traffic and it takes me 30 minutes to get to work. The rest of the time there is high traffic and it takes me 50 minutes. Let X be the number of minutes of my commute. E[X] = 38. What is var(X)? What is σ_X?

Computing Variance

- With probability 0.6 there is low traffic and it takes me 30 minutes to get to work. The rest of the time there is high traffic and it takes me 50 minutes. Let X be the number of minutes of my commute. E[X] = 38. What is var(X)? What is σ_X?

- Let X be a random variable, and let Y = aX + b for any scalars a and b. E[Y] = aE[X] + b. What is var(Y)? What about σ_X?

Variance of the Bernoulli

Consider a Bernoulli random variable X with parameter p.

Variance of the Bernoulli

Consider a Bernoulli random variable X with parameter *p*.

We know that the probability mass function is defined as:

$$p_X(1) = p$$

$$p_X(0) = 1-p$$

Variance of the Bernoulli

Consider a Bernoulli random variable X with parameter *p*.

We know that the probability mass function is defined as:

$$p_X(1) = p$$
$$p_X(0) = 1-p$$

What is var(X)? What is σ_X?

Discrete Uniform Random Variables

(we skipped this example in class because we were running low on time, but I'm leaving the slide here anyway… if you're curious, there is a nice description in the textbook)

Discrete Uniform Random Variables

Let a and b be arbitrary integers with $a \leq b$. Let X be a random variable that takes on every integer value in a range $[a, b]$ with equal probability.

What is p_X? What is E[X]? What is var(X)?

$$\mathrm{var}(X) = (1/12)\,(n^2-1)$$

Try to verify this on your own (or with the book).

Hint: $E[X^2] = 1/6\,(n+1)\,(2n+1)$

Expectations and Decision Making

Consider the following game: There are two questions. Question A is worth \$100 and you can answer it correctly with probability 0.8. Question B is worth \$200 and you can answer it correctly with probability 0.5, independent of your ability to answer question 1.

You may choose a question. If you answer it correctly, you get the corresponding prize and a chance to answer the other question (and receive the other prize if you're right). If your first answer is wrong, the game ends.

Which question should you attempt to answer first?

Ranking and Zipf's Law

- Suppose we take all of the terms that people search for on Google and rank them by their popularity. How often do people search for the *i*th most popular term?

Ranking and Zipf's Law

- Suppose we take all of the terms that people search for on Google and rank them by their popularity. How often do people search for the ith most popular term?

- How many friends does the ith most popular member of Facebook have?

Ranking and Zipf's Law

- Suppose we take all of the terms that people search for on Google and rank them by their popularity. How often do people search for the ith most popular term?
- How many friends does the ith most popular member of Facebook have?
- How many people live in the ith most populated city in the United States?

Ranking and Zipf's Law

- Suppose we take all of the terms that people search for on Google and rank them by their popularity. How often do people search for the ith most popular term?
- How many friends does the ith most popular member of Facebook have?
- How many people live in the ith most populated city in the United States?
- What is the net worth of the ith wealthiest person in the country?

Ranking and Zipf's Law

Let x be an item's rank, and let y be the frequency or "size" of the item.

- x = rank of a search term, y = number of searches
- x = rank of a Facebook member, y = number of friends
- x = rank of city in terms of population, y = population
- x = rank of person in terms of wealth, y = wealth

Ranking and Zipf's Law

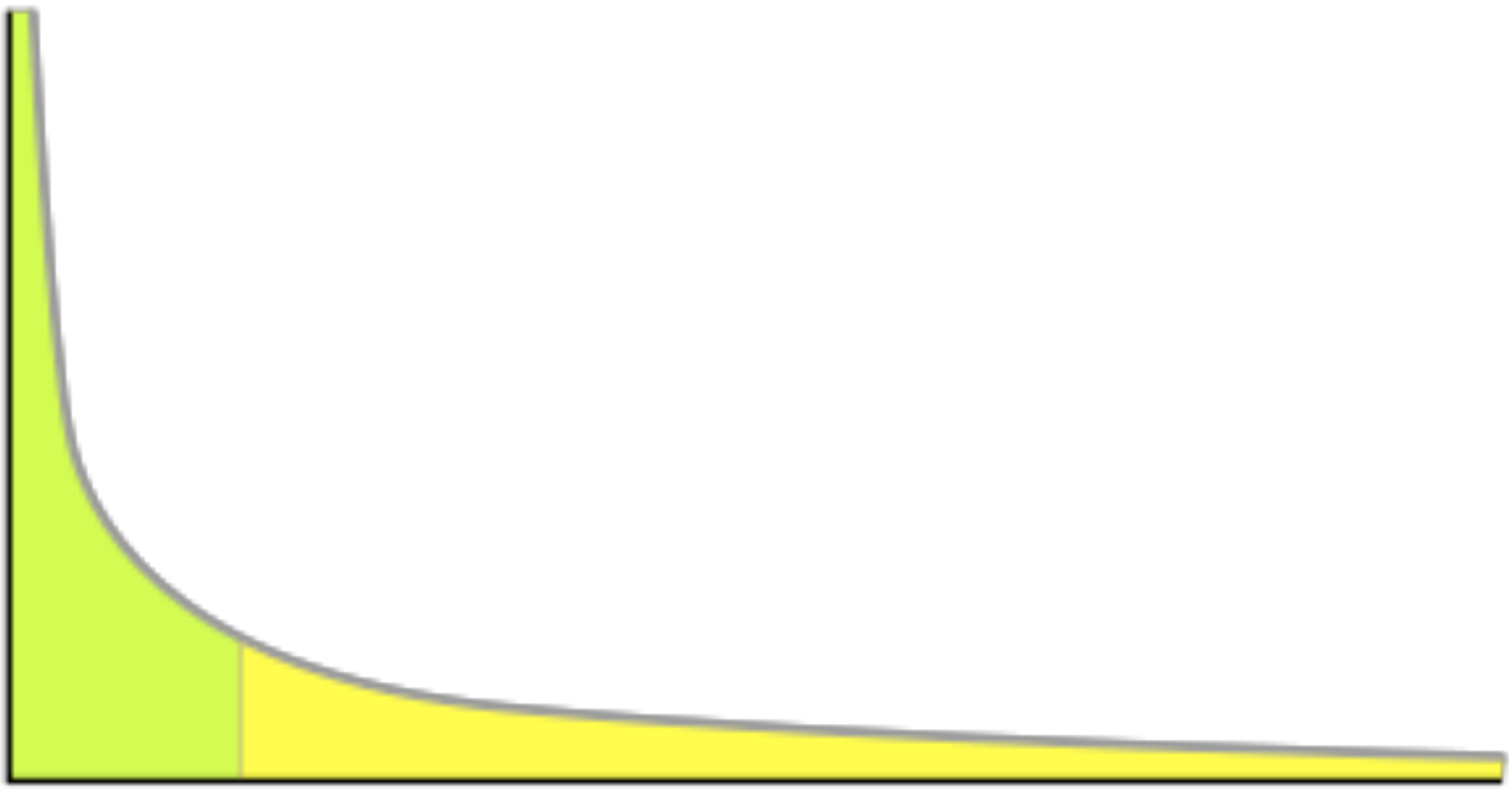
Let x be an item's rank, and let y be the frequency or "size" of the item.

- x = rank of a search term, y = number of searches
- x = rank of a Facebook member, y = number of friends
- x = rank of city in terms of population, y = population
- x = rank of person in terms of wealth, y = wealth

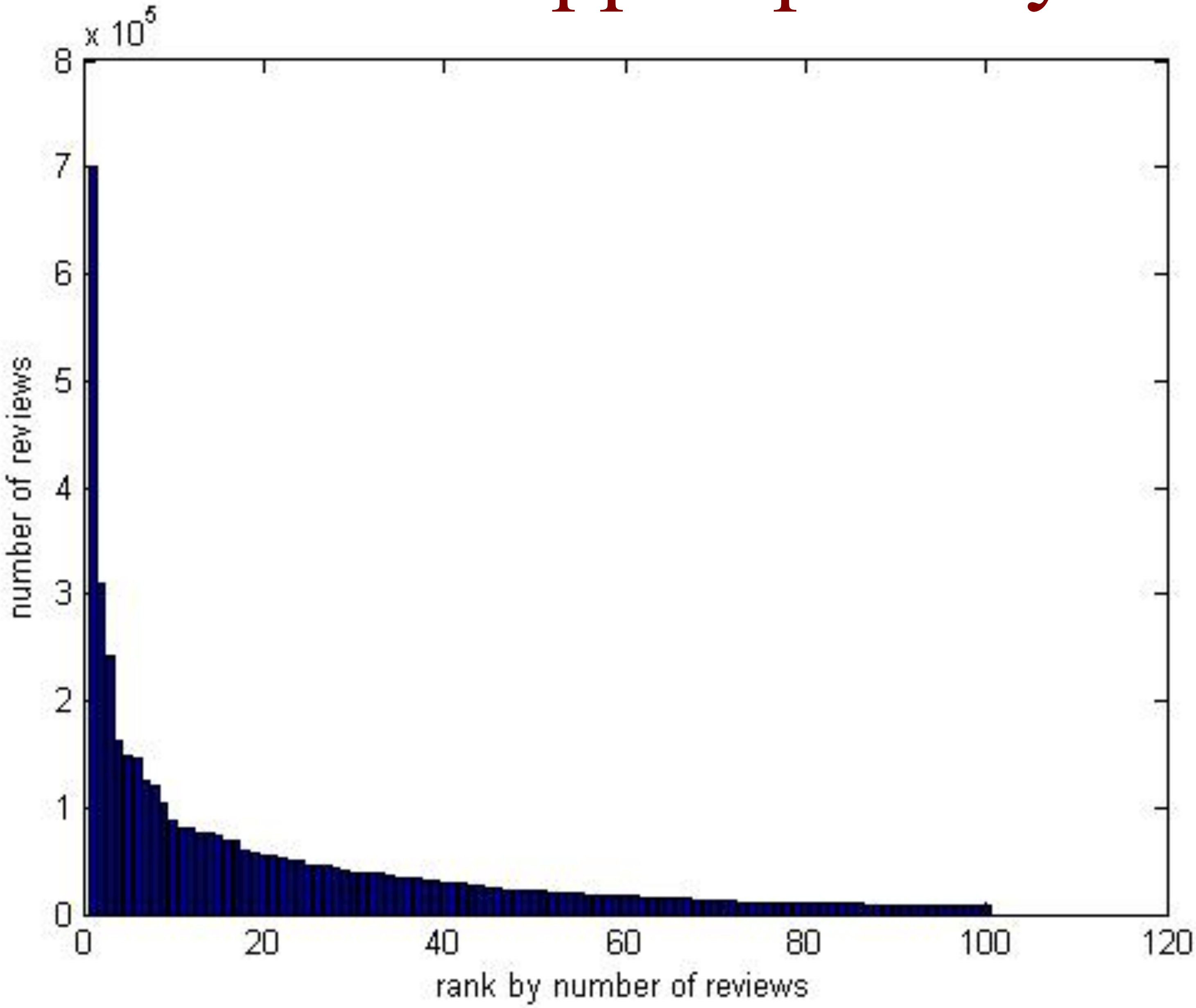
Zipf's law says that $y \sim x^{-\alpha}$ for a (usually small) constant α

Heavy Tails

Functions of the form $y \sim x^{-\alpha}$ have heavy tails



iPhone App Popularity

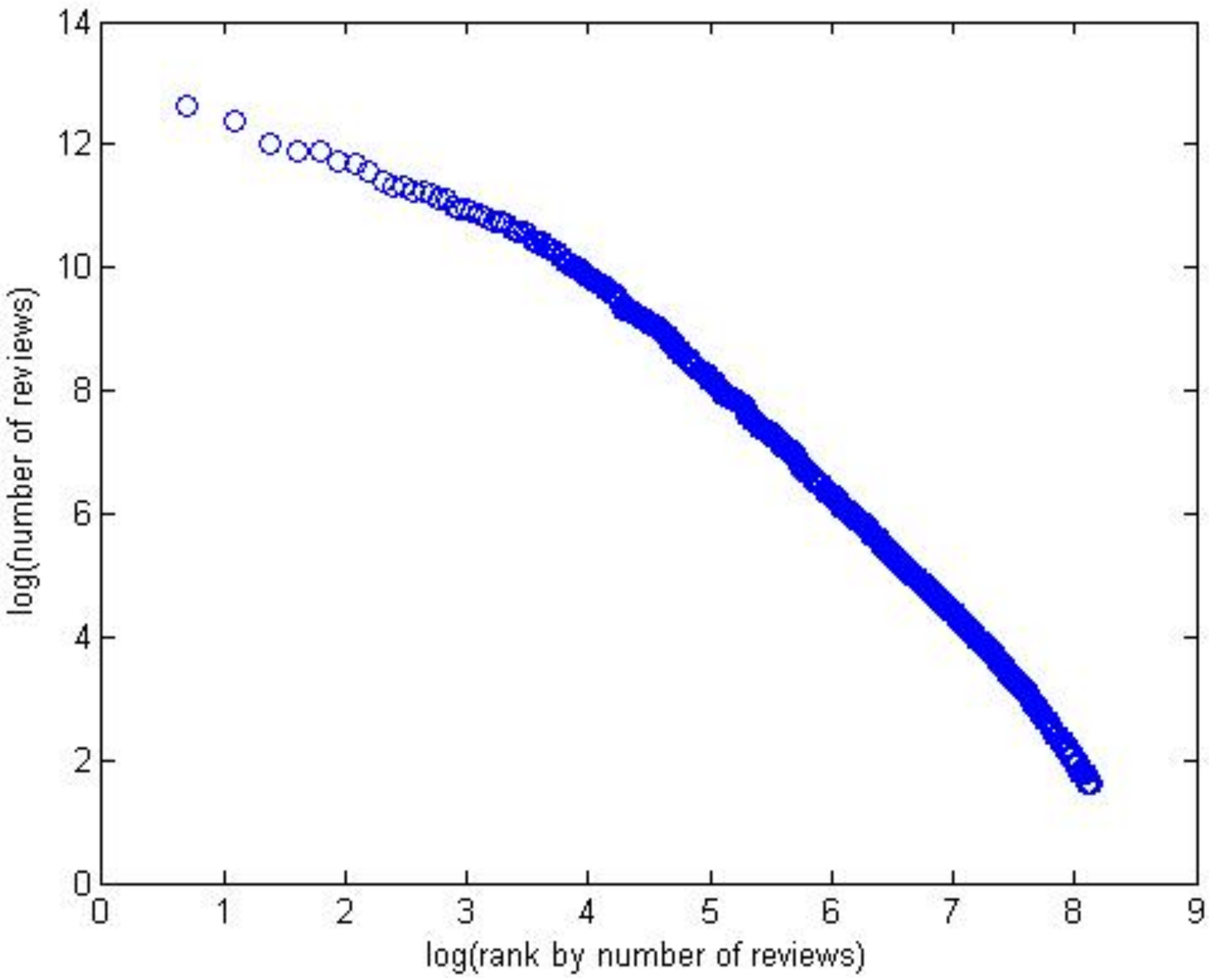


The Log-Log Plot Test

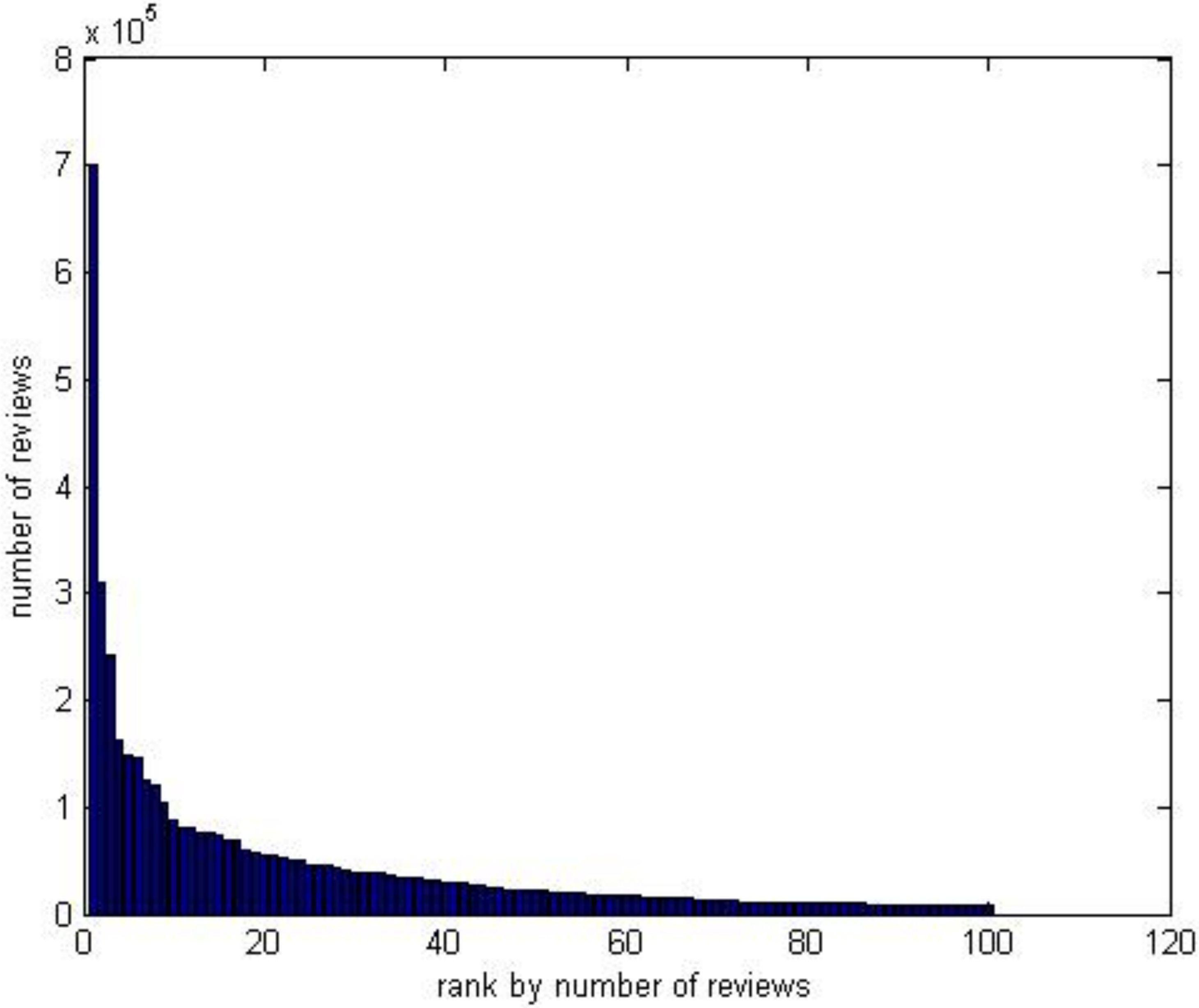
Functions of this form appear linear on a log-log plot

The Log-Log Plot Test

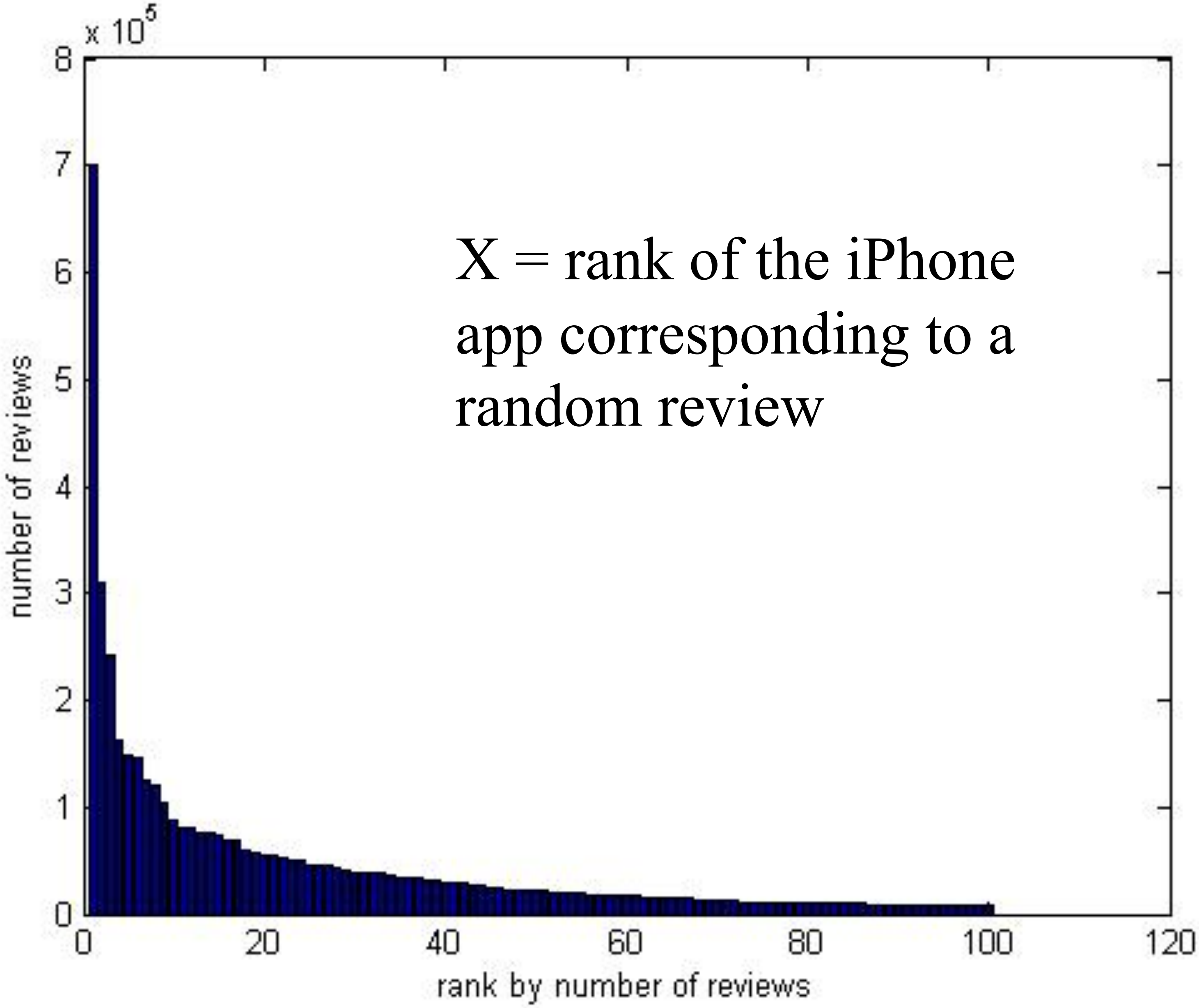
Functions of this form appear linear on a log-log plot



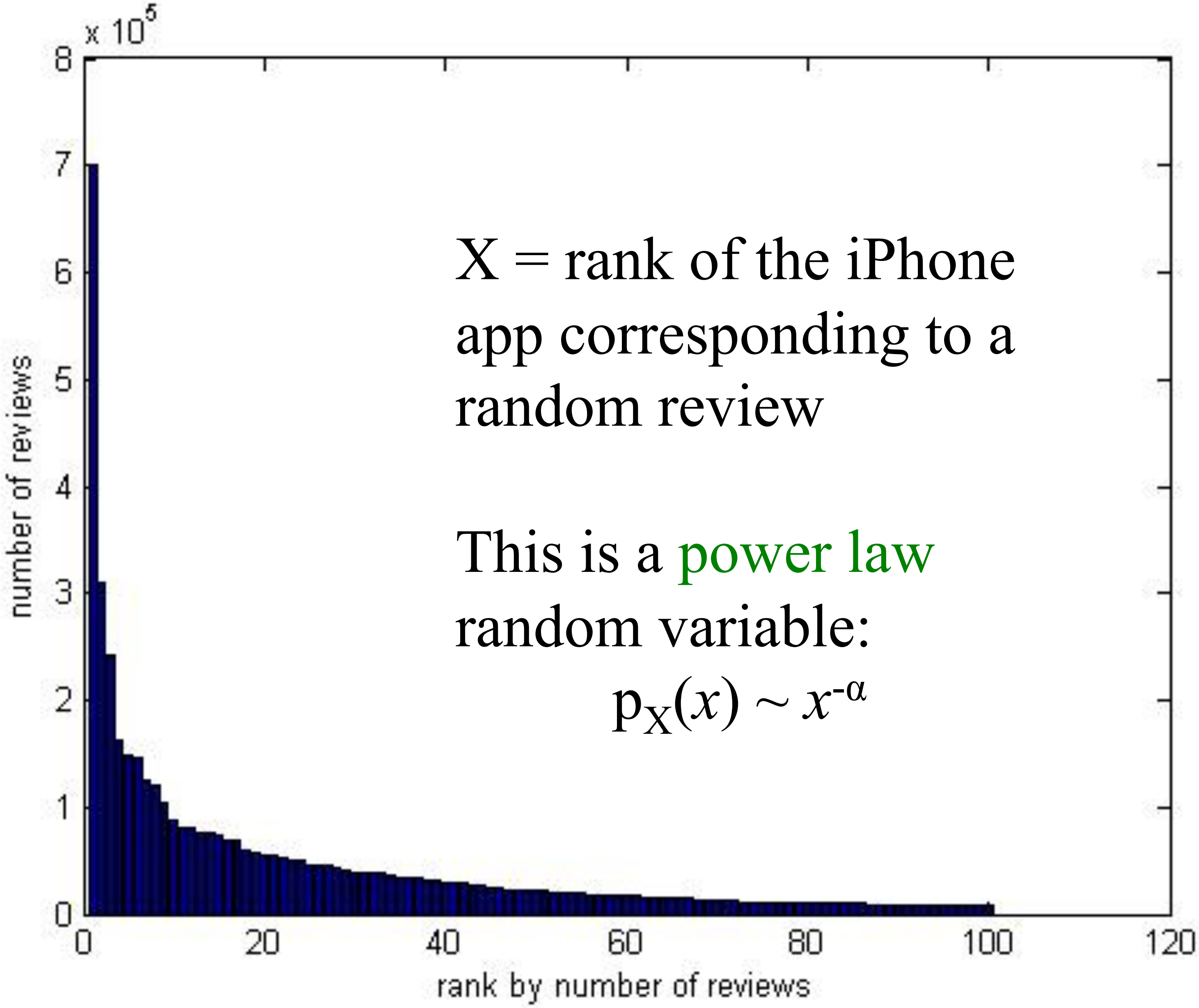
Where are the random variables?



Where are the random variables?



Where are the random variables?



X = rank of the iPhone app corresponding to a random review

This is a power law random variable:

$$p_X(x) \sim x^{-\alpha}$$

Power Law Random Variables

- Suppose we take all of the terms that people search for on Google and rank them by their popularity. Let X be the "rank" of the next term someone searches for on Google.

Power Law Random Variables

- Suppose we take all of the terms that people search for on Google and rank them by their popularity. Let X be the "rank" of the next term someone searches for on Google.

- Suppose we rank Facebook members by number of friends. Choose a random edge on the friendship graph and let X be the rank of a random node of that edge.

Power Law Random Variables

- Suppose we take all of the terms that people search for on Google and rank them by their popularity. Let X be the "rank" of the next term someone searches for on Google.

- Suppose we rank Facebook members by number of friends. Choose a random edge on the friendship graph and let X be the rank of a random node of that edge.

- Suppose we rank files by size. Let X be the rank of the file to which a random bit belongs.