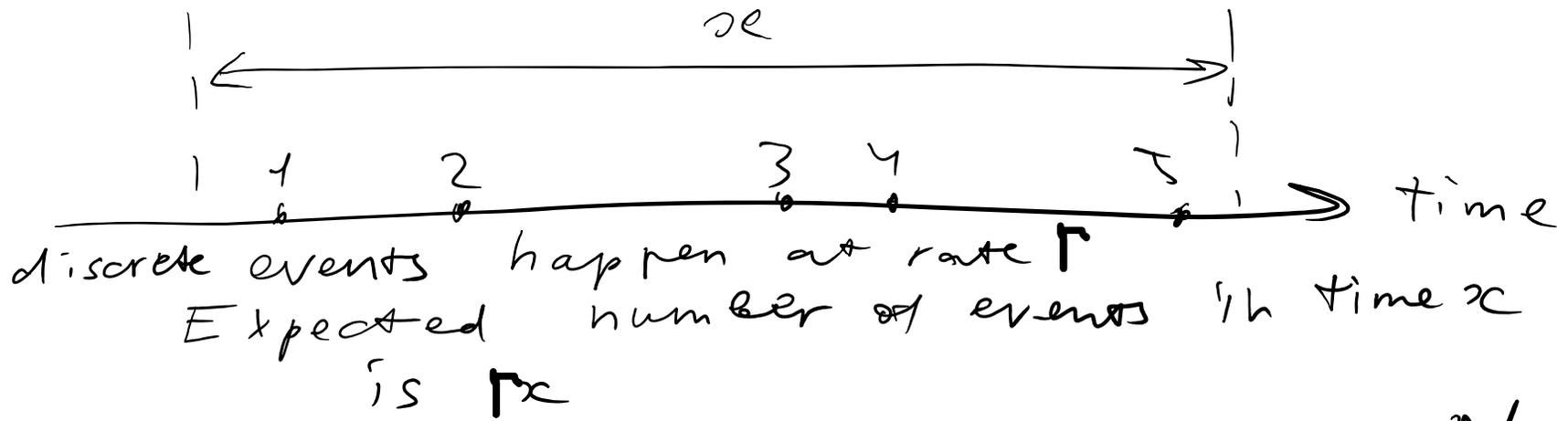


Constant rate (Poisson) process



# Constant rate (Poisson) process



The actual number of events  $N_x$  is a Poisson distributed discrete random variable

$$P(N_x = n) = \frac{(\Gamma x)^n}{n!} e^{-\Gamma x}$$

Why Poisson?

Divide  $x$  into many tiny intervals of length  $\Delta x$

$$p = \Gamma \Delta x$$

$$L = x / \Delta x$$

$$\text{Prob}(N=n) = \binom{L}{n} p^n (1-p)^{L-n}$$

↓  $p \sim \Delta x \rightarrow 0, L \sim \frac{1}{\Delta x} \rightarrow \infty$

$$E(N_x) = pL = \Gamma x$$

Poisson

# Constant rate (AKA Poisson) processes

- Let's assume that proteins are produced by ribosomes in the cell at a **rate  $r$  per second**.
- **The expected number of proteins** produced in  **$x$  seconds** is  **$r \cdot x$** .
- The actual number of proteins  $N_x$  is a **discrete random variable** following a **Poisson distribution** with mean  $r \cdot x$ :

$$P_N(N_x=n) = \exp(-r \cdot x) (r \cdot x)^n / n! \quad E(N_x) = rx$$

- Why Discrete Poisson Distribution?
  - Divide time into many tiny intervals of length  $\Delta x \ll 1/r$
  - The probability of success (protein production) per interval is small:  $p_{\text{success}} = r\Delta x \ll 1$ ,
  - The number of intervals is large:  $n = x/\Delta x \gg 1$
  - Mean is constant:  $r = E(N_x) = p_{\text{success}} \cdot n = r\Delta x \cdot x/\Delta x = r \cdot x$
  - In the limit  $\Delta x \ll x$ ,  $p_{\text{success}}$  is small and  $n$  is large, thus Binomial distribution  $\rightarrow$  Poisson distribution

# Exponential Distribution Definition

**Exponential random variable**  $X$  describes interval between two successes of a constant rate (Poisson) random process with success rate  $r$  per unit interval.

The probability density function of  $X$  is:

$$f(x) = re^{-rx} \quad \text{for } 0 \leq x < \infty$$

Closely related to the discrete **geometric distribution**

$$f(x) = p(1-p)^{x-1} = p e^{(x-1) \ln(1-p)} \approx pe^{-px} \quad \text{for small } p$$



To summarize constant rate processes:

$r$  - rate per unit of length time length TD

$N(x)$  - discrete number of events

in time  $x$

Poisson: 
$$P(N(x)=n) = \frac{(r \cdot x)^n}{n!} e^{-r \cdot x}$$

Time interval  $X$  between successive events is a continuously distributed random variable

Its PDF is  $f(x) = e^{-rx}$

# What is the interval $X$ between two successes of a constant rate process?

- $X$  is a **continuous random variable**
- **CCDF:  $P_X(X > x) = P_N(N_X = 0) = \exp(-r \cdot x)$ .**
  - Remember:  $P_N(N_X = n) = \exp(-r \cdot x) (r \cdot x)^n / n!$
- **PDF:  $f_X(x) = -dCCDF_X(x)/dx = r \cdot \exp(-r \cdot x)$**
- We started with a discrete Poisson distribution where time  $x$  was a parameter
- We ended up with a **continuous exponential distribution**

# Exponential Mean & Variance

If the random variable  $X$  has an exponential distribution with rate  $r$ ,

$$\mu = E(X) = \frac{1}{r} \quad \text{and} \quad \sigma^2 = V(X) = \frac{1}{r^2} \quad (4-15)$$

Note that, for the:

- Poisson distribution: **mean** = **variance**
- Exponential distribution: **mean** = **standard deviation** = **variance**<sup>0.5</sup>



# Biochemical Reaction Time

- The time  $x$  (in minutes) until an enzyme catalyzes a biochemical reaction and generates a product is approximated by this CCDF:

$$F_{>}(x) = e^{-2x} \text{ for } 0 \leq x$$

Here the rate of this process is  $r=2 \text{ min}^{-1}$  and  $1/r=0.5 \text{ min}$  is the average time between successive products of this enzyme

- What is the PDF?

$$f(x) = -\frac{dF_{>}(x)}{dx} = -\frac{d}{dx} e^{-2x} = 2e^{-2x} \text{ for } 0 \leq x$$

- What **proportion of reactions will not generate another product within 0.5 minutes of the previous product?**

$$P(X > 0.5) = F_{>}(0.5) = e^{-2 * 0.5} = 0.37$$

We observed our enzyme for 1 minute and no product has been generated:

The product is “overdue”

What is the probability that a product will not appear during the next 0.5 minutes?

$$F_{>}(x) = e^{-2x}$$

$$F_{>}(0.5) \approx 0.37$$

$$F_{>}(1.5) \approx 0.05$$

$$F_{>}(1.0) \approx 0.13$$

- A. 0.32
- B. 0.37
- C. 0.08
- D. 0.24
- E. I have no idea

Get your i-clickers

We observed our enzyme for 1 minute and no product has been generated:

The product is “overdue”

What is the probability that a product will not appear during the next 0.5 minutes?

$$F_{>}(x) = e^{-2x}$$

$$F_{>}(0.5) \approx 0.37$$

$$F_{>}(1.5) \approx 0.05$$

$$F_{>}(1.0) \approx 0.13$$

A. 0.32

B. 0.37

C. 0.08

D. 0.24

E. I have no idea

Get your i-clickers



Memoryless property of the exponential distribution

$$P(X > t+s | X > s) = P(X > t)$$

$$\begin{aligned} P(X > t+s | X > s) &= \frac{P(X > t+s, X > s)}{P(X > s)} = \\ &= \frac{\exp(-r(t+s))}{\exp(-rs)} = \exp(-rt) = \\ &= P(X > t) \end{aligned}$$

Exponential is the only memoryless distribution

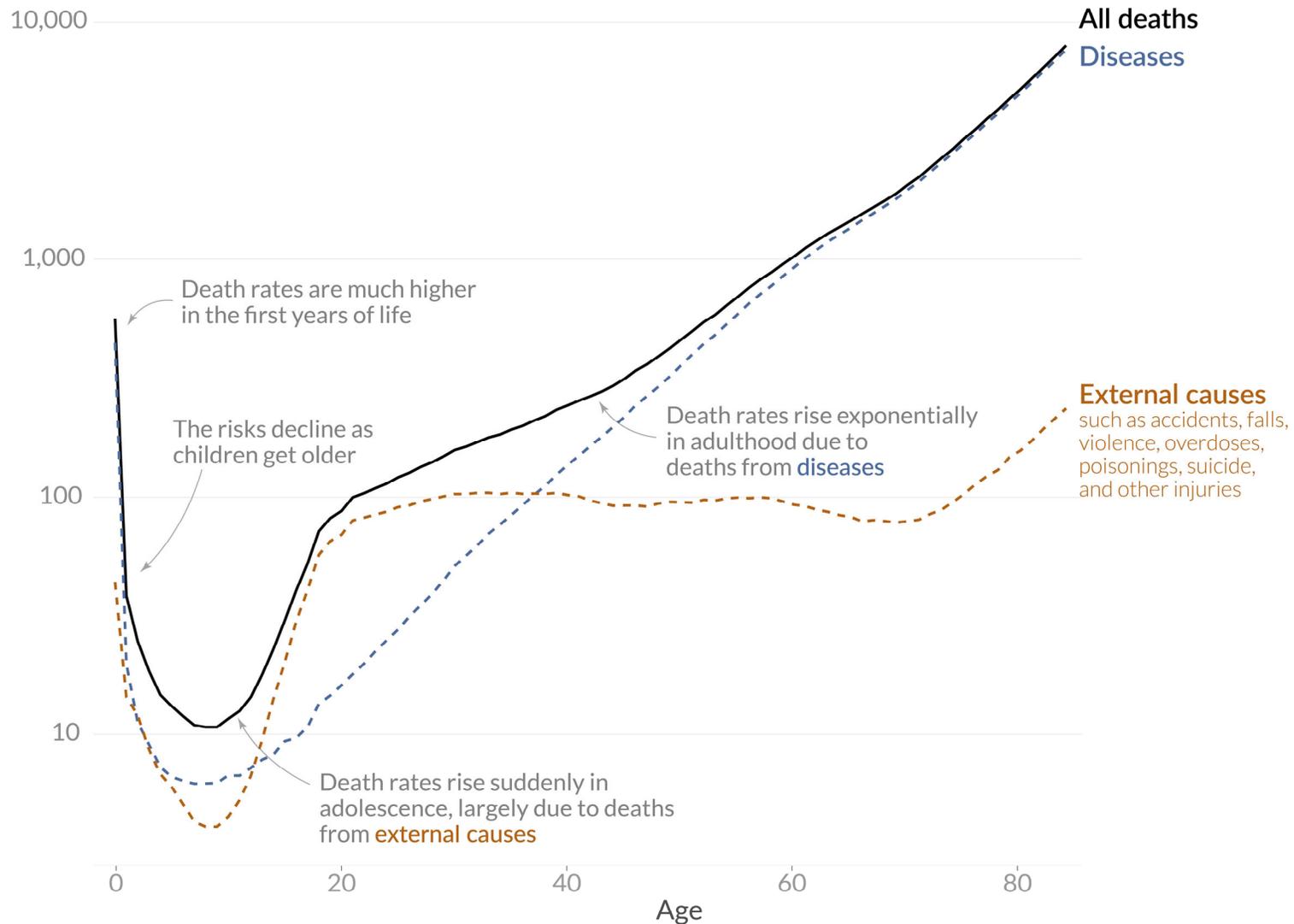
# Survivorship curve: surviving fraction vs age



# Death rates across ages

National data from the United States between 2018 and 2021.

Annual death rate, per 100,000 people (log scale)



Note: Period death rates using ICD-10 categories. 'Diseases' includes all categories except 'external causes' and 'signs, symptoms and abnormal findings'.

Source: United States Centers for Disease Control and Prevention, via CDC Wonder database

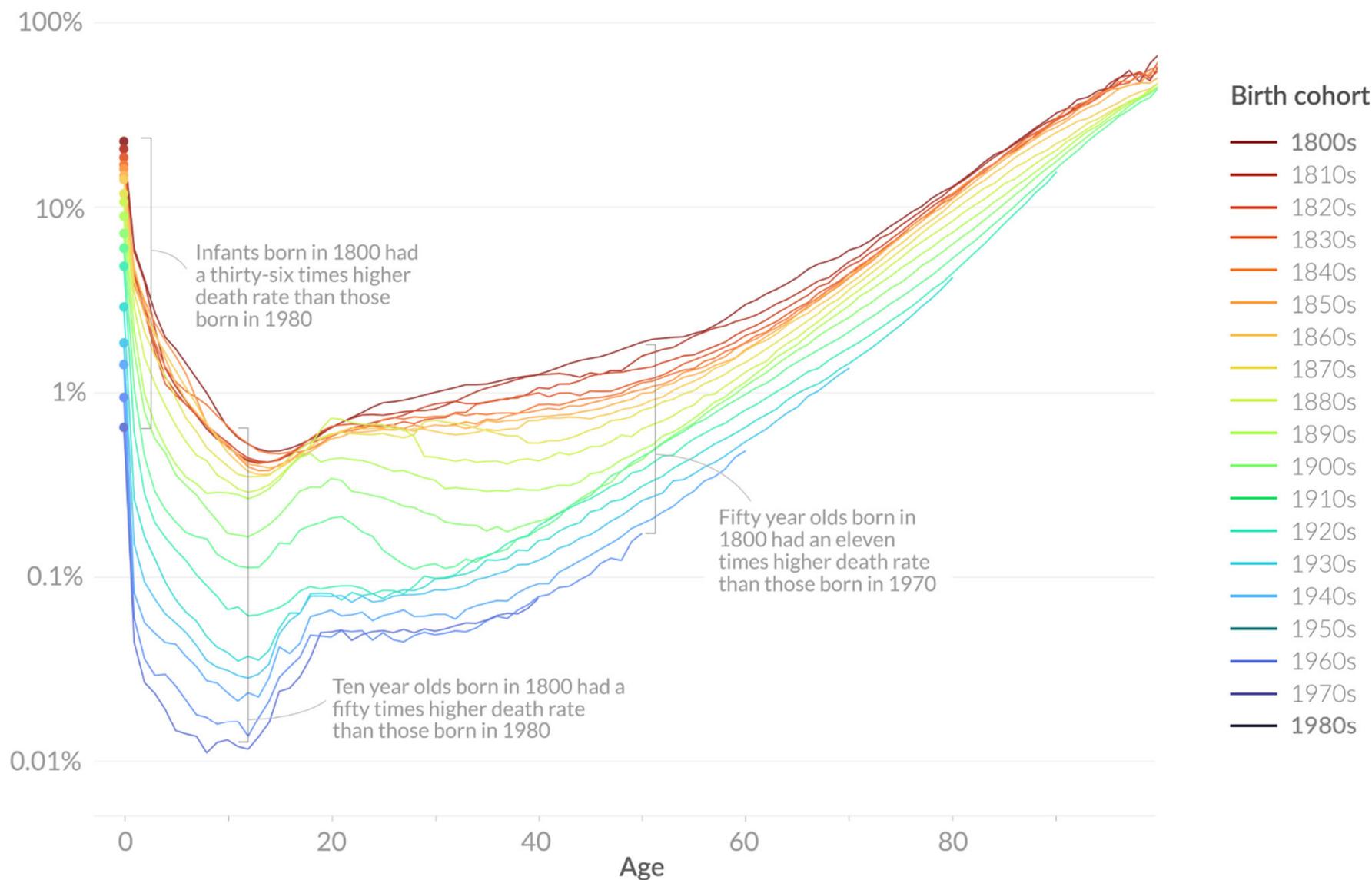
OurWorldinData.org – Research and data to make progress against the world's largest problems.

Licensed under CC-BY by the author Saloni Dattani

# Death rates have declined across the lifespan

Cohort data from Sweden where long-term data is available. Annual death rates at age 0 are shown as dots.

Death rate (log scale)



**Note:** Lines begin for age groups who were included in the dataset, once data collection began. Lines end for those who have not yet reached a given age. Death rates above age 95 are not shown due to uncertainties.

**Source:** Human Mortality Database. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France).

# It is different for species/genera

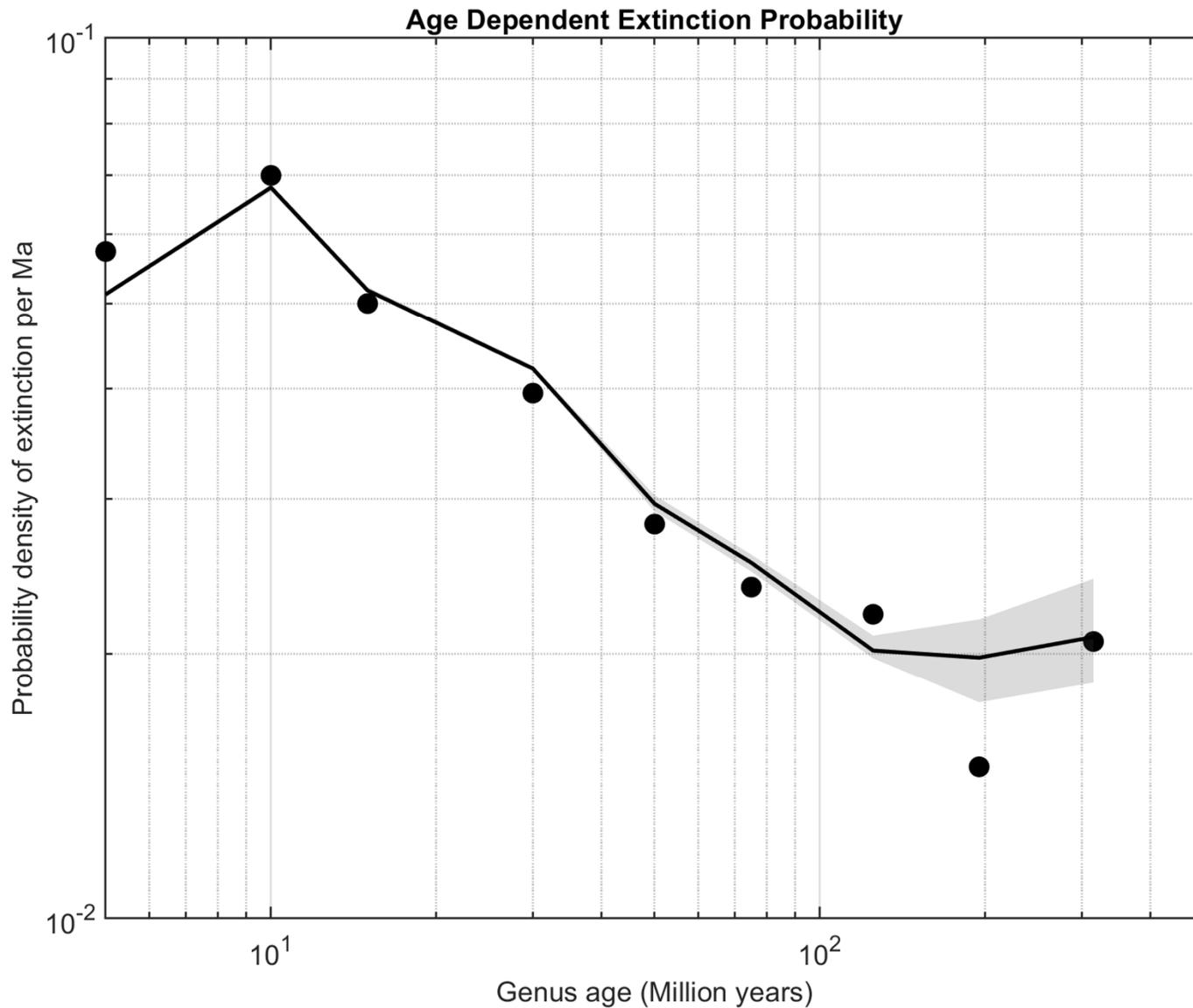


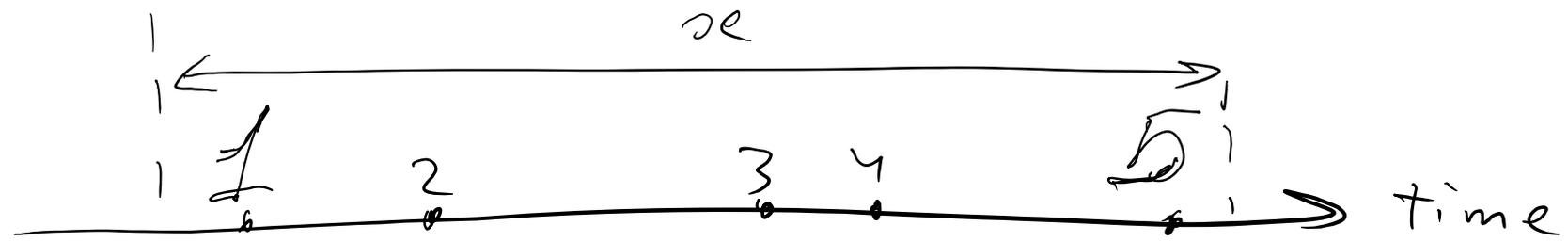
Figure from “Good genes or good luck? Both, if you want to survive mass extinctions”  
Mikhail Tikhonov, Kim Sneppen, Stefan Bornholdt, and Sergei Maslov

# Erlang Distribution

- The Erlang distribution is a generalization of the exponential distribution.
- The **exponential distribution** models the time interval to the **1<sup>st</sup> event**, while the
- **Erlang distribution** models the time interval to the  **$k^{\text{th}}$  event**, i.e., a sum of  $k$  exponentially distributed variables.
- The exponential, as well as Erlang distributions, is based on the constant rate (or Poisson) process.



Constant rate (Poisson) process



Events happen independently  
from each other at  
constant rate =  $r$  ;  $E[N_x] = rx$

$X$  follows Erlang distribution

$$P(X > x) = \sum_{n=0}^{r-1} P(N_x = n) =$$
$$= \sum_{n=0}^{r-1} \frac{(rx)^n}{n!} e^{-rx}$$

# Erlang Distribution

Generalizes the Exponential Distribution:

waiting time between event 0 and **event k**  
(constant rate process with rate=**r**)

$$P(X > x) = \sum_{m=0}^{k-1} \frac{e^{-rx} (rx)^m}{m!} = 1 - F(x)$$

Differentiating  $F(x)$  we find that all terms in the sum except the last one cancel each other:

$$f(x) = \frac{r^k x^{k-1} e^{-rx}}{(k-1)!} \quad \text{for } x > 0 \quad \text{and } k = 1, 2, 3, \dots$$

# Gamma Distribution

The random variable  $X$  with a probability density function:

$$f(x) = \frac{r^k x^{k-1} e^{-rx}}{\Gamma(k)}, \text{ for } x > 0 \quad (4-18)$$

has a gamma random distribution with parameters  $r > 0$  and  $k > 0$ . If  $k$  is a positive integer, then  $X$  has an Erlang distribution.



$$f(x) = \frac{r^k x^{k-1} e^{-rx}}{\Gamma(k)}, \text{ for } x > 0$$

$$\int_0^{+\infty} f(x) dx = 1, \text{ Hence}$$

$$\Gamma(k) = \int_0^{+\infty} r^k x^{k-1} e^{-rx} dx = \int_0^{+\infty} y^{k-1} e^{-y} dy$$

Comparing with Erlang distribution  
for integer k one gets

$$\Gamma(k) = (k-1)!$$

# Gamma Function

The gamma function is the generalization of the factorial function for  $r > 0$ , not just non-negative integers.

$$\Gamma(k) = \int_0^{\infty} y^{k-1} e^{-y} dy, \quad \text{for } r > 0 \quad (4-17)$$

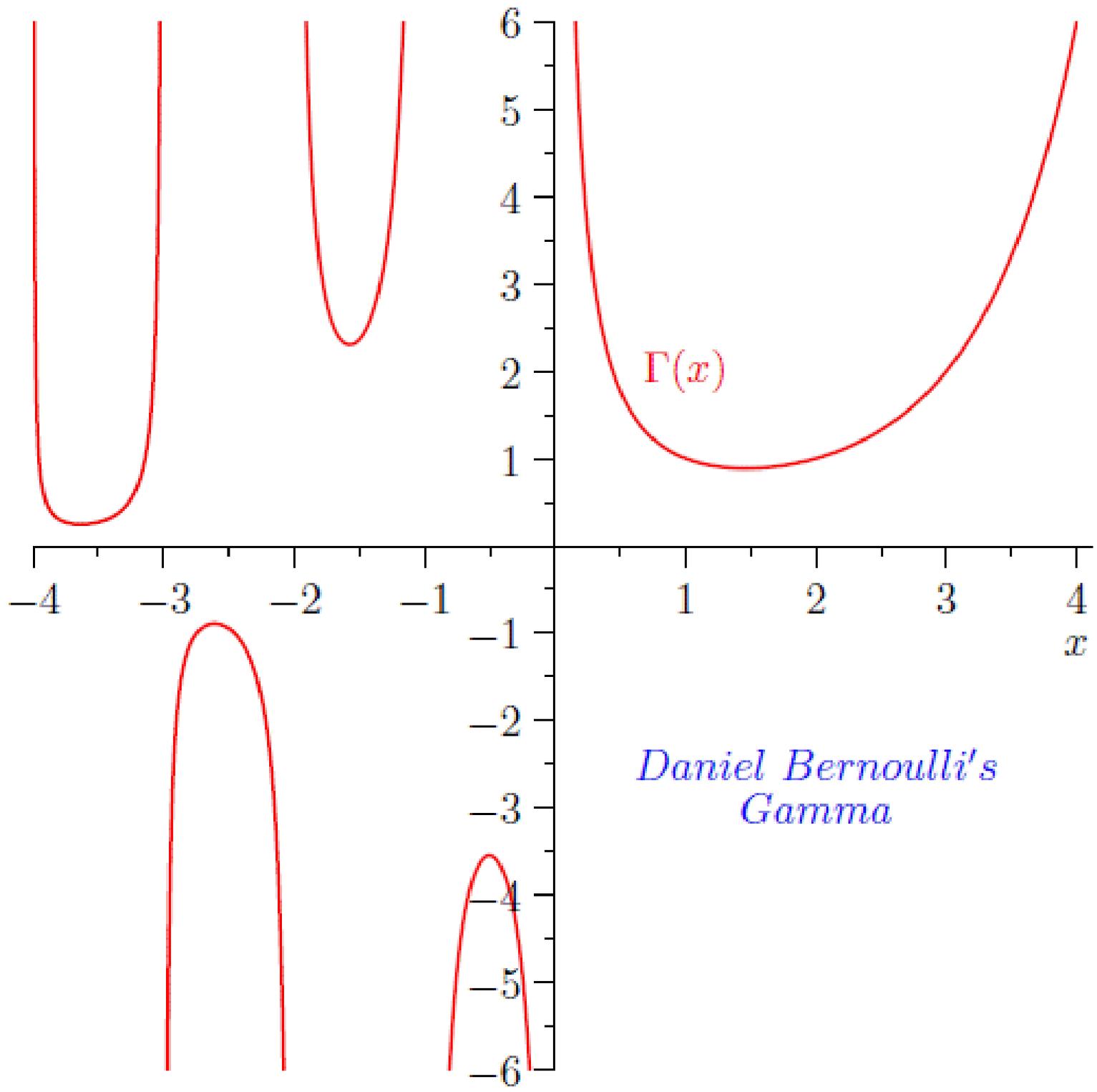
Properties of the gamma function

$$\Gamma(1) = 1$$

$$\Gamma(k) = (k - 1)\Gamma(k - 1) \quad \text{recursive property}$$

$$\Gamma(k) = (k - 1)! \quad \text{factorial function}$$

$$\Gamma\left(\frac{1}{2}\right) = \pi^{\frac{1}{2}} = 1.77 = \left(-\frac{1}{2}\right)! \quad \text{interesting fact}$$



*Daniel Bernoulli's  
Gamma*

SOLO

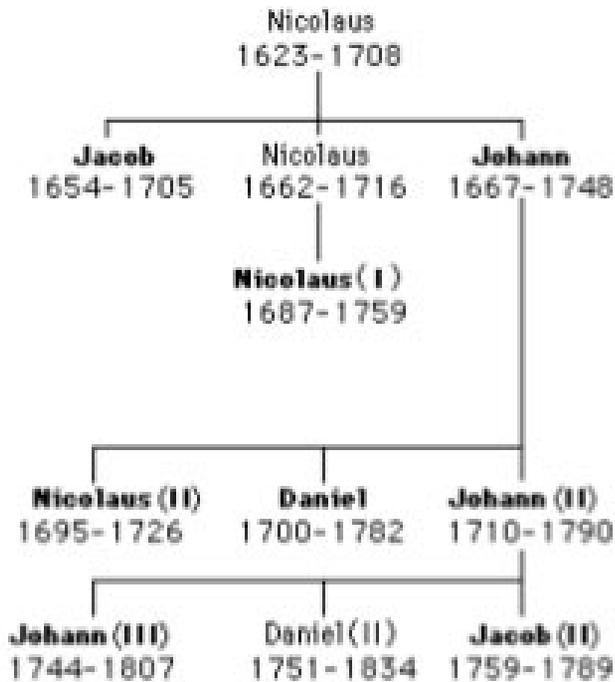
# BERNOULLI FAMILY

Bernoulli trials

## SOLO HERMELIN

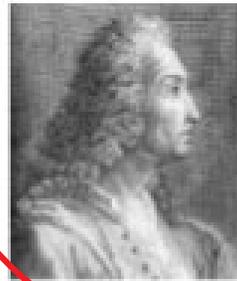
<http://www.solohermelin.com>

### The Bernoulli family



Those shown in **bold** above are in our archive

See This



Jacob  
1654-1705



Johann  
1667-1748



Nicolaus II  
1695-1720



Daniel  
1700-1782



Johann II  
1710-1790



Johann III  
1744-1807



Jacob II  
1759-1789

Gamma function

# Mean & Variance of the Erlang and Gamma

- If  $X$  is an **Erlang** (or more generally **Gamma**) distributed **random variable** with parameters  $r$  and  $k$ ,

$$\mu = E(X) = k/r \quad \text{and} \quad \sigma^2 = V(X) = k/r^2 \quad (4-19)$$

- Generalization of exponential results:  
 $\mu = E(X) = 1/r$  and  $\sigma^2 = V(X) = 1/r^2$  or

Negative binomial results:

$$\mu = E(X) = k/p \quad \text{and} \quad \sigma^2 = V(X) = k(1-p) / p^2$$

# Matlab exercise:

- Generate a sample of 100,000 variables with exponential distribution with  $r = 0.1$  using `random('Exponential'...)` command
- Generate a sample of 100,000 variables with “Harry Potter” Gamma distribution with  $r = 0.1$  and  $k = 9 \frac{3}{4}$  (9.75)
- Calculate mean and compare it to  $1/r$  (Exponential) and  $k/r$  (Gamma)
- Calculate standard deviation and compare it to  $1/r$  (Exponential) and  $\sqrt{k}/r$  (Gamma)
- Plot semilog-y plots of PDFs and CCDFs.

# Matlab exercise: Exponential

- `Stats=100000; r=0.1;`
- `r2=random('Exponential', 1./r, Stats,1);`  
**%Matlab uses mean=1/r as a parameter**
- `disp([mean(r2),1./r]); disp([std(r2),1./r]);`
- `step=0.1; [a,b]=hist(r2,0:step:max(r2));`
- `pdf_e=a./sum(a)./step;`
- `figure; subplot(1,2,1); semilogy(b,pdf_e,'ko-');`
- `X=0:0.01:100;`
- `for m=1:length(X);`  
    `cdf_e(m)=sum(r2>X(m))./Stats;`  
`end;`
- `subplot(1,2,2); semilogy(X,cdf_e,'ko-');`

# Matlab exercise: Gamma

- `Stats=100000; r=0.1; k=9.75;`
- `r2=random('Gamma', k,1./r, Stats,1);`
- `disp([mean(r2),k./r]);`
- `disp([std(r2),sqrt(k)./r]);`
- `step=0.1; [a,b]=hist(r2,0:step:max(r2));`
- `pdf_g=a./sum(a)./step;`
- `figure;`
- `subplot(1,2,1); semilogy(b,pdf_g,'ko-'); hold on;`
- `x=0:0.01:max(r2); clear cdf_g;`
- `for m=1:length(x);`
- `cdf_g(m)=sum(r2>x(m))./Stats;`
- `end;`
- `subplot(1,2,2); semilogy(x,cdf_g,'rd-');`

Credit: XKCD  
comics

WHY ARE THERE SLAVES IN THE BIBLE

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS  
WHY ARE AMERICANS AFRAID OF DRAGONS

WHY IS HTTPS CROSSED OUT IN RED  
WHY IS THERE A LINE THROUGH HTTPS  
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK  
WHY IS HTTPS IMPORTANT

# QUESTIONS

FOUND IN GOOGLE AUTOCOMPLETE



WHY ARE THERE WEEKS  
WHY DO I FEEL DIZZY

WHY AREN'T ECONOMISTS RICH

WHY DO AMERICANS CALL IT SOCCER

WHY ARE MY EARS RINGING

WHY ARE THERE SO MANY AVENGERS

WHY ARE THE AVENGERS FIGHTING THE X MEN  
WHY IS WOLVERINE NOT IN THE AVENGERS

## WHY ARE THERE ANTS IN MY LAPTOP

WHY IS EARTH TILTED

WHY IS SPACE BLACK  
WHY IS OUTER SPACE SO COLD  
WHY ARE THERE PYRAMIDS ON THE MOON  
WHY IS NASA SHUTTING DOWN



WHY IS THERE AN OWL IN MY BACKYARD

WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY DO OWLS ATTACK PEOPLE

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE GODS

WHY ARE THERE TWO SPOCKS

WHY IS MT VESUVIUS THERE

WHY DO THEY SAY T MINUS

WHY ARE THERE OBELISKS

WHY ARE WRESTLERS ALWAYS WET

WHY ARE OCEANS BECOMING MORE ACIDIC

WHY IS ARWEN DYING

WHY AREN'T MY QUAIL LAYING EGGS  
WHY AREN'T MY QUAIL EGGS HATCHING

WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

WHY IS LIFE SO BORING

WHY ARE CIGARETTES LEGAL

WHY ARE THERE DUCKS IN MY POOL

WHY IS JESUS WHITE

WHY IS THERE LIQUID IN MY EAR

WHY DO Q TIPS FEEL GOOD

WHY DO GOOD PEOPLE DIE



WHY ARE ULTRASOUNDS IMPORTANT  
WHY ARE ULTRASOUND MACHINES EXPENSIVE  
WHY IS STEALING WRONG

WHY ARE DOGS AFRAID OF FIREWORKS  
WHY IS THERE NO KING IN ENGLAND

WHY DO WHALES JUMP  
WHY ARE WITCHES GREEN

WHY ARE THERE MIRRORS ABOVE BEDS

WHY DO I SAY UH

WHY IS SEA SALT BETTER

WHY ARE THERE TREES IN THE MIDDLE OF FIELDS

WHY IS THERE NOT A POKEMON MMO

WHY IS THERE LAUGHING IN TV SHOWS

WHY ARE THERE DOORS ON THE FREEWAY

WHY ARE THERE SO MANY SVCHOST.EXE RUNNING

WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA

WHY ARE THERE SCARY SOUNDS IN MINECRAFT

WHY IS THERE KICKING IN MY STOMACH

WHY ARE THERE TWO SLASHES AFTER HTTP

WHY ARE THERE CELEBRITIES

WHY DO SNAKES EXIST

WHY DO OYSTERS HAVE PEARLS

WHY ARE DUCKS CALLED DUCKS

WHY DO THEY CALL IT THE CLAP

WHY ARE KYLE AND CARTMAN FRIENDS

WHY IS THERE AN ARROW ON AANG'S HEAD

WHY ARE TEXT MESSAGES BLUE

WHY ARE THERE MUSTACHES ON CLOTHES

WHY ARE THERE MUSTACHES ON CARS

WHY ARE THERE MUSTACHES EVERYWHERE

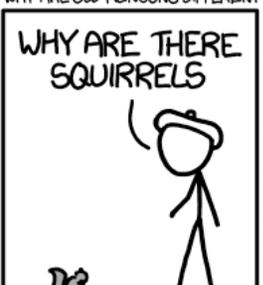
WHY ARE THERE SO MANY BIRDS IN OHIO

WHY IS THERE SO MUCH RAIN IN OHIO

WHY IS OHIO WEATHER SO WEIRD

WHY ARE THERE MALE AND FEMALE BIKES

WHY ARE THERE BRIDESMAIDS  
WHY DO DYING PEOPLE REACH UP  
WHY AREN'T THERE VARICOSE ARTERIES  
WHY ARE OLD KUNGONS DIFFERENT



WHY IS PROGRAMMING SO HARD  
WHY IS THERE A 0 OHM RESISTOR  
WHY DO AMERICANS HATE SOCCER  
WHY DO RHYMES SOUND GOOD

WHY DO TREES DIE

WHY IS THERE NO SOUND ON CNN

WHY AREN'T POKEMON REAL

WHY AREN'T BULLETS SHARP  
WHY DO DREAMS SEEM SO REAL

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY DO SPIDERS COME INSIDE

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY DO SPIDER BITES ITCH

WHY IS DYING SO SCARY

WHY IS THERE NO GPS IN LAPTOPS

WHY DO KNEES CLICK

WHY AREN'T THERE E GRADES

WHY IS SEX SO IMPORTANT



WHY AREN'T THERE DINOSAUR GHOSTS

WHY ARE THERE FEMALE MR NIMES

WHY IS GPS FREE