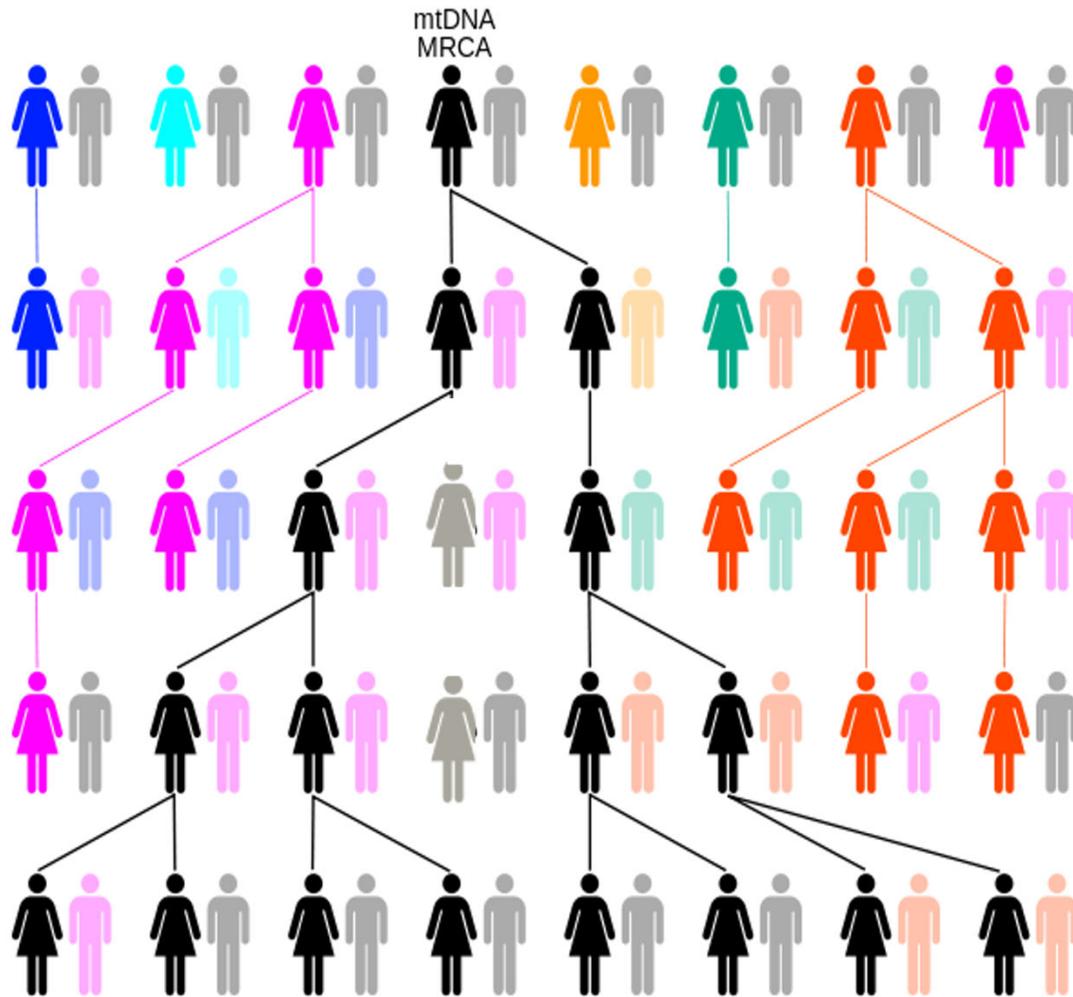
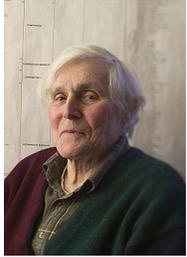


Geometric distribution in biology



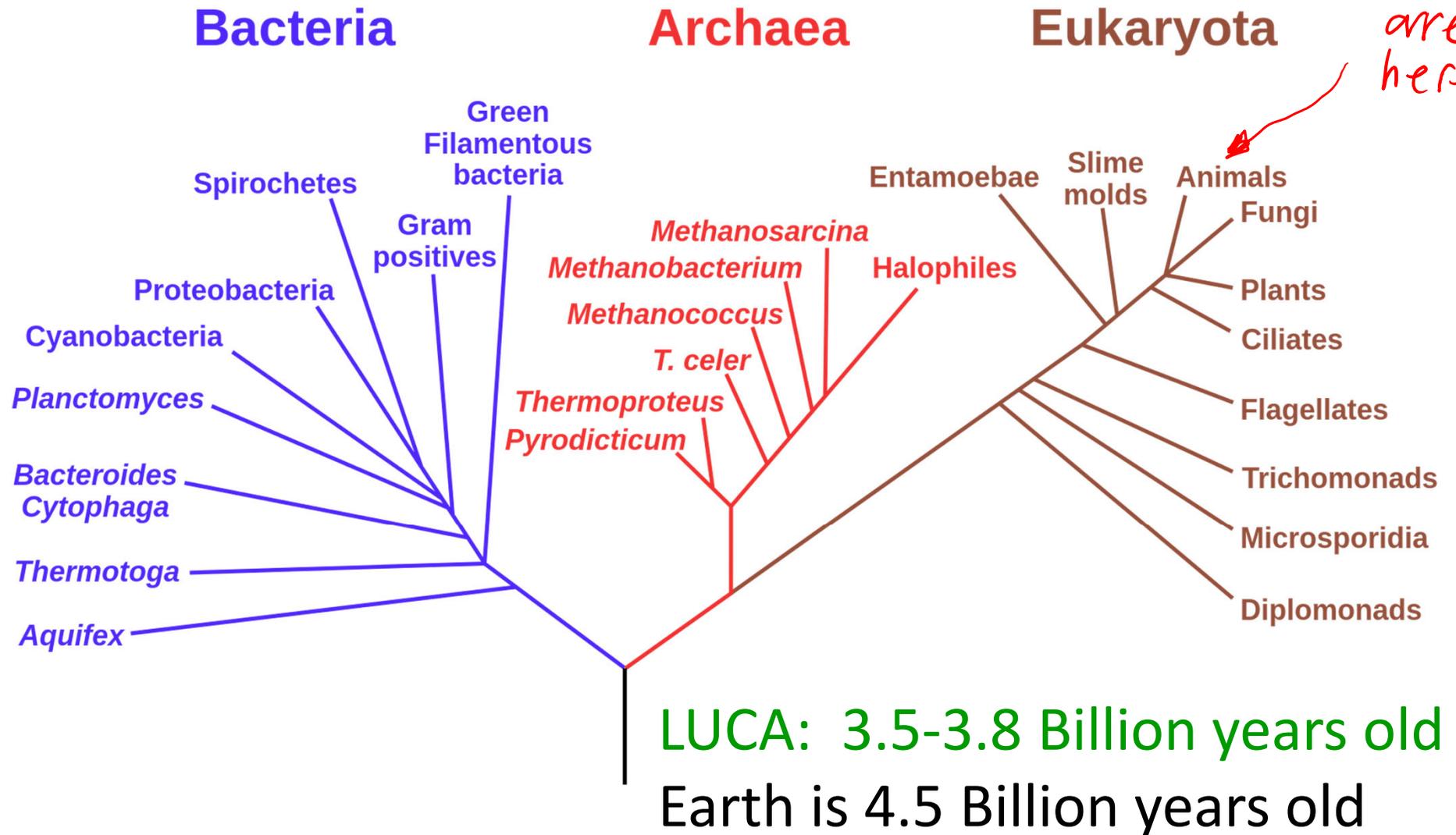
- Randomly pick two women. Random variable $T = \#$ of generations since their last maternal ancestor?
- T follows a geometric distribution
- Do these two women have the same mother? Yes: “success” in finding their last common ancestor ($p=1/N$). $P(T=1)=1/N$.
- No? “failure” ($1-p=1-1/N$). Go to their mothers and repeat the same question.
- $P(T=t)=(1-1/N)^{t-1}(1/N) \approx (1/N) \exp(-(t-1)/N)$
- t can be inferred from the density of differences on mtDNA $=2\mu t$

Last Universal Common Ancestor (LUCA)



Archaea were discovered here at UIUC in 1977 by Carl R. Woese (1928-2012) and George E. Fox

You are here



Negative Binomial Definition

- In a series of independent trials with **constant probability of success, p** , let the random variable X denote the **number of trials until r successes occur**. Then X is a **negative binomial** random variable with parameters:

$$0 < p < 1 \text{ and } r = 1, 2, 3, \dots$$

- The probability mass function is:

$$f(x) = C_{r-1}^{x-1} p^r (1-p)^{x-r} \text{ for } x = r, r+1, r+2, \dots \quad (3-11)$$

- Compare it to binomial

$$f(x) = C_x^n p^x (1-p)^{n-x} \text{ for } x = 1, 2, \dots, n$$

NOTE OF CAUTION: Matlab, Mathematica, and many other sources use X to denote the **number of failures until one gets r successes**.

We stick with **Montgomery-Runger**. For our notation use $X_{\text{matlab}} + r$

Negative Binomial Mean & Variance

- If X is a **negative binomial** random variable with parameters p and r ,

$$\mu = E(X) = \frac{r}{p} \quad \text{and} \quad \sigma^2 = V(X) = \frac{r(1-p)}{p^2} \quad (3-12)$$

- Compare to the **geometric** distribution:

$$\mu = E(X) = \frac{1}{p} \quad \text{and} \quad \sigma^2 = V(X) = \frac{(1-p)}{p^2} \quad (3-10)$$

Cancer is scary!

- Approximately 40% of men and women will be diagnosed with cancer at some point during their lifetimes (source: NCI website)

TABLE 21.2 Leading causes of death in United States in 2010. Cause of death is based on the International Classification of Diseases, Tenth Revision, 1992.

Rank	Cause of death	Number	Percent of all deaths
–	All causes	2,468,435	100.0
1	Diseases of heart	597,689	24.2
2	Malignant neoplasms	574,743	23.3
3	Chronic lower respiratory diseases	138,080	5.6
4	Cerebrovascular diseases	129,476	5.2
5	Accidents (unintentional injuries)	120,859	4.9
6	Alzheimer's disease	83,494	3.4
7	Diabetes mellitus	69,071	2.8
8	Nephritis, nephrotic syndrome, and nephrosis	50,476	2.0
9	Influenza and pneumonia	50,097	2.0
10	Intentional self-harm (suicide)	38,364	1.6

Source: National Vital Statistics Reports, 62(6) (http://www.cdc.gov/nchs/data/nvsr/nvsr62/nvsr62_06.pdf)

Table from
J. Pevsner
3rd edition

- “War on Cancer” – president Nixon 1971.
“Moonshot to Cure Cancer” – vice-president Joe Biden 2016

“War on Cancer” progress report

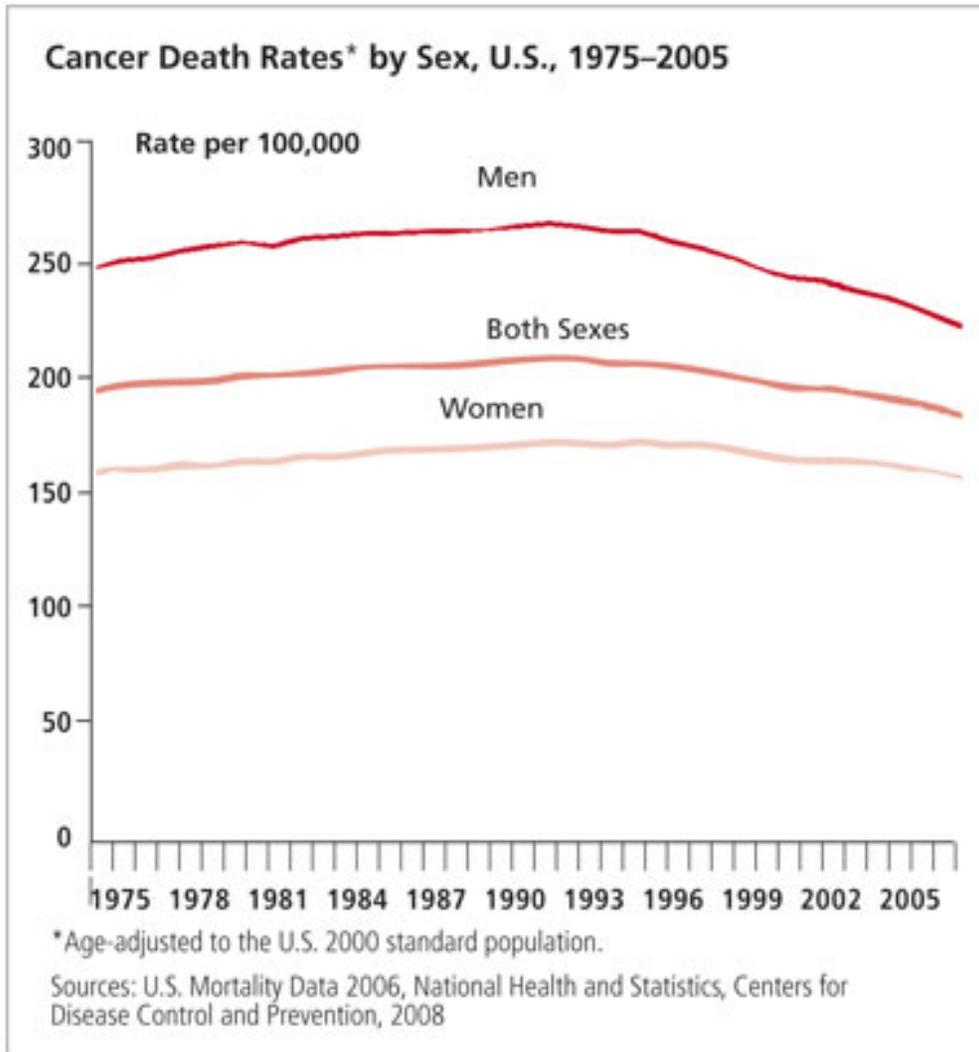


Figure 2

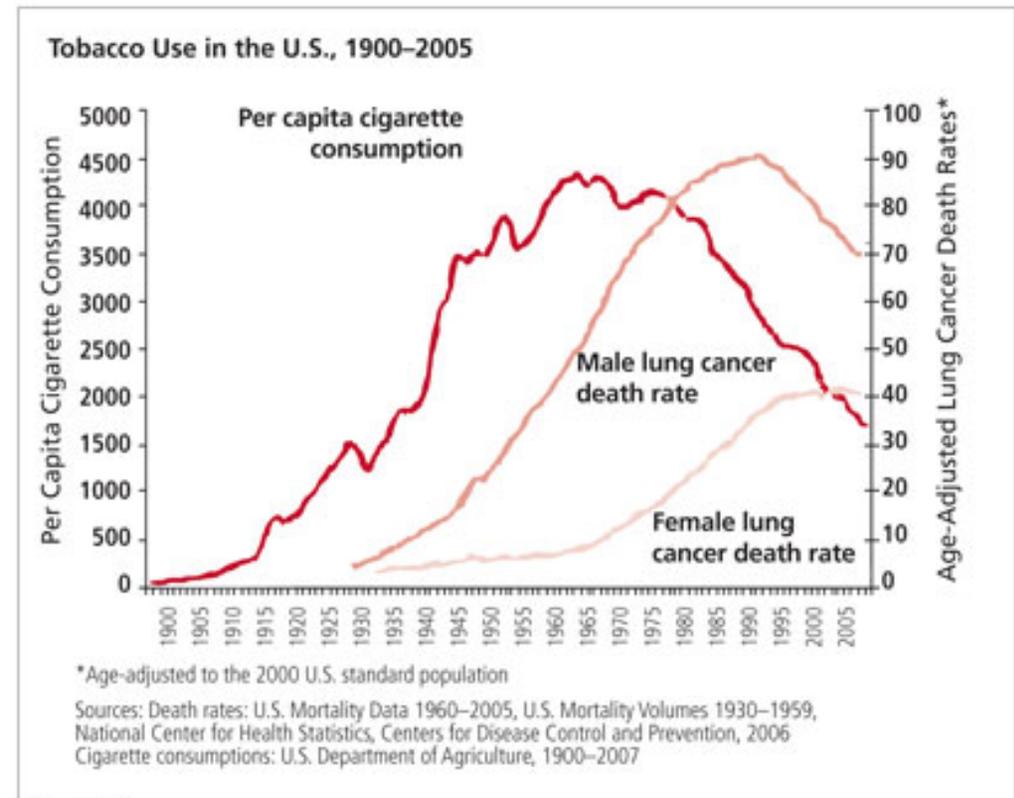


Figure 3

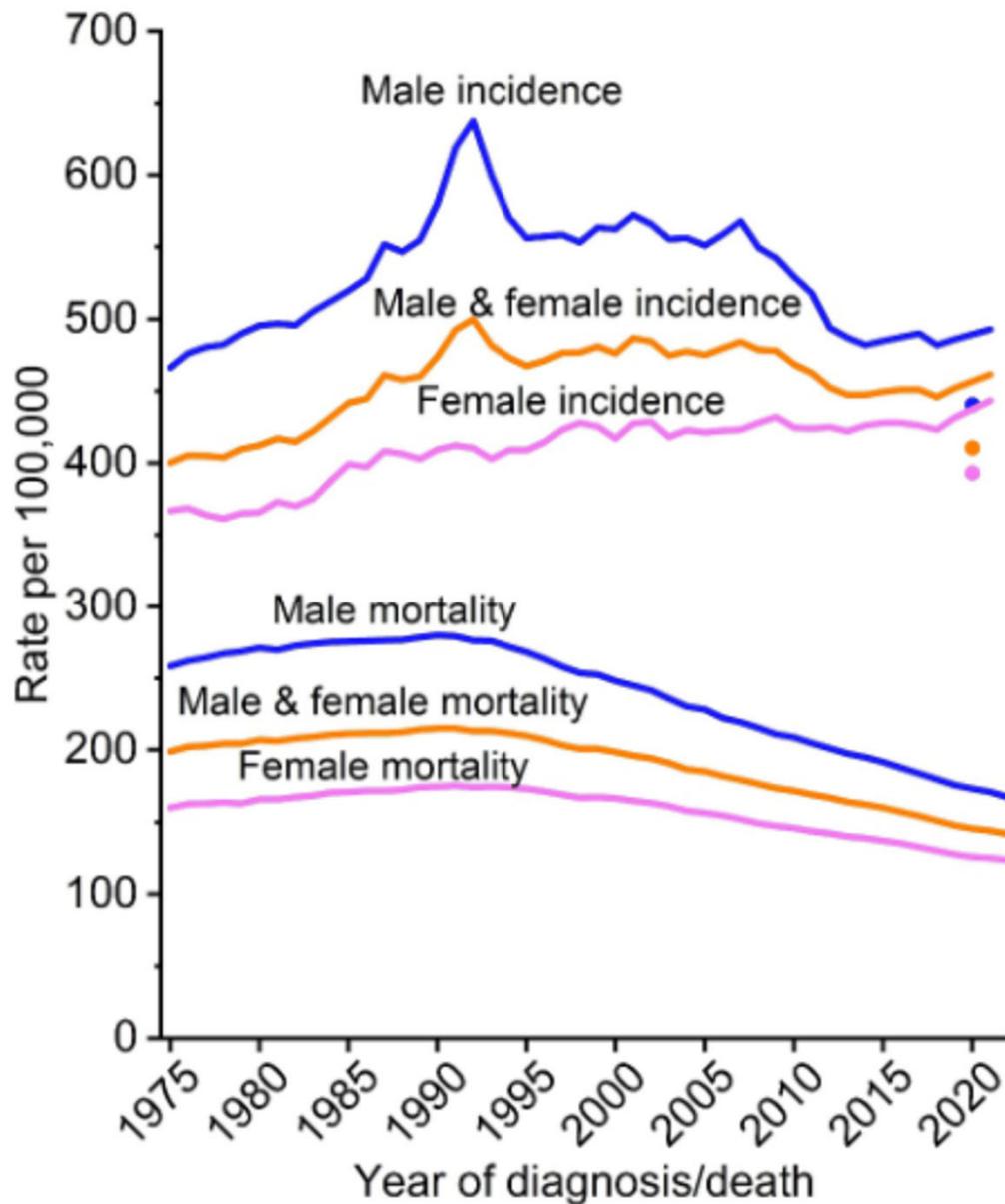


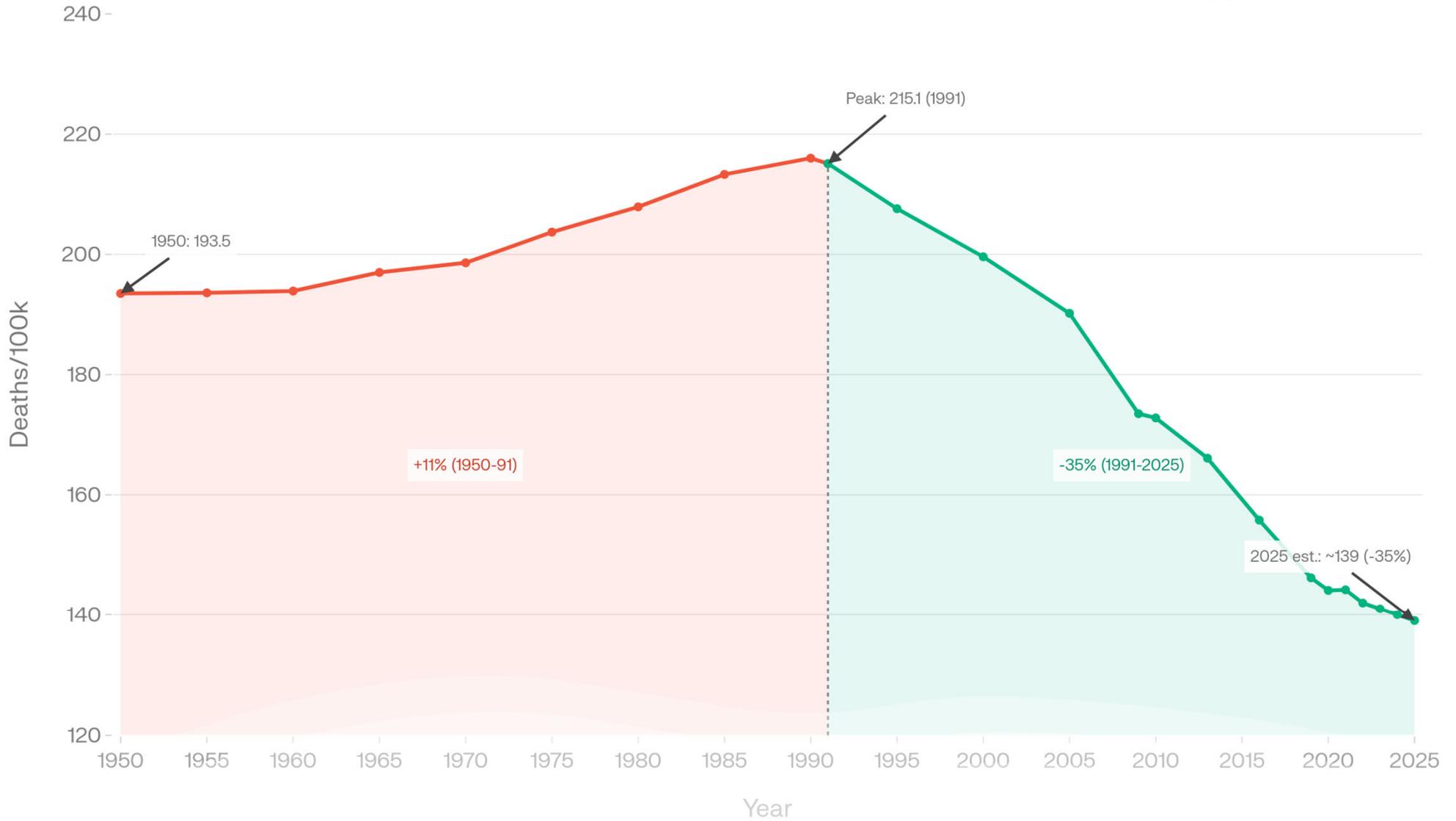
FIGURE 2 Trends in cancer incidence (1975–2021) and mortality (1975–2022) rates by sex, United States. Rates are age adjusted to the 2000 US standard population, and incidence rates are adjusted for delays in case reporting. Incidence data for 2020 are shown separate from trend lines.

Figure from:
 Siegel, R. L., Kratzer, T. B.,
 Giaquinto, A. N., Sung, H. &
 Jemal, A. Cancer statistics,
 2025. CA: A Cancer
 J. Clin. 75, 10–45 (2025).

U.S. Cancer Mortality: Rise & Fall (1950-2025)

CDC/NCHS, SEER/NCI, ACS | Age-adjusted per 100,000
—●— Rising (1950-1991) —●— Declining (1991-2025)

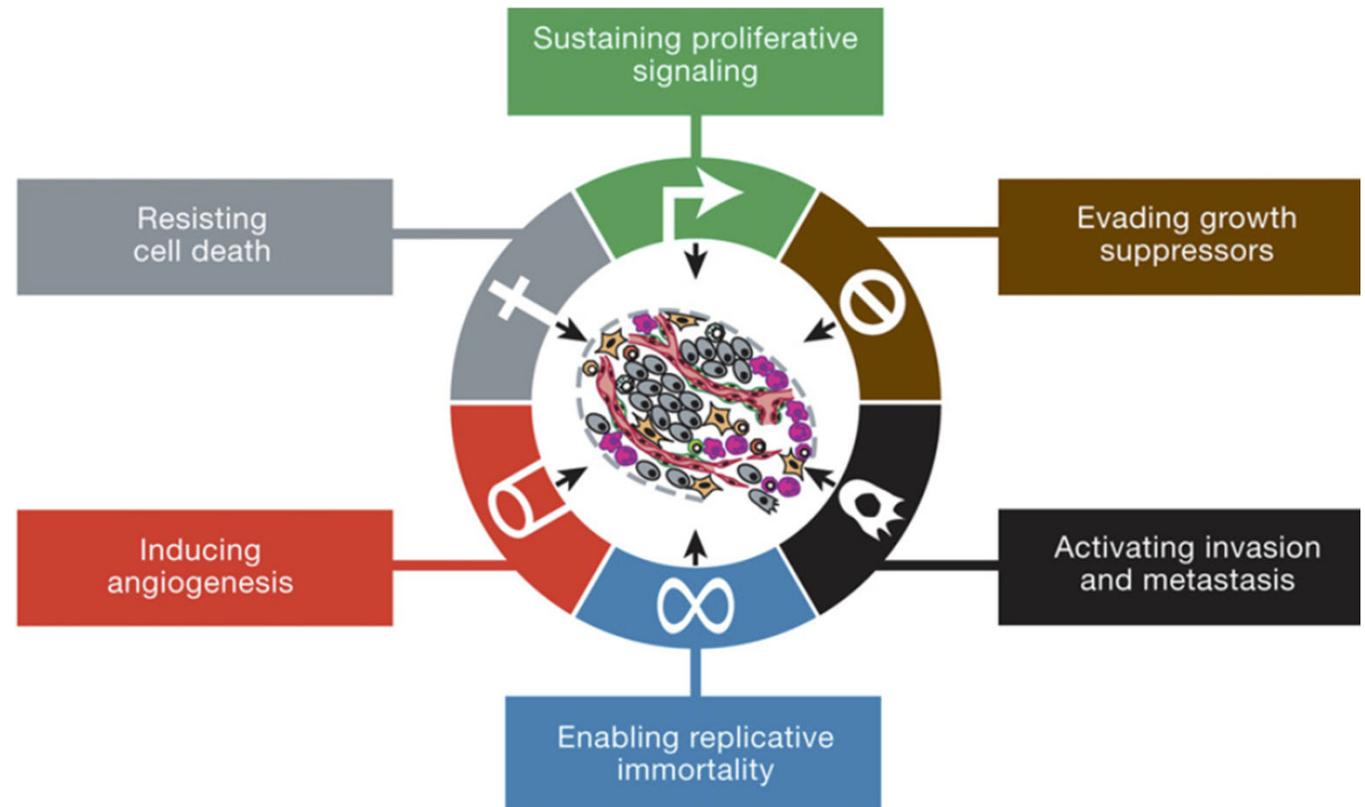
Powered by  perplexity



Probability theory and statistics
is a powerful tool to
learn new cancer biology

“Driver genes” theory

- Progression of cancer is caused by **accumulation of mutations** in a handful of **“driver” genes**
- Mutations in driver genes boost the growth of a tumor
- **Oncogenes: expression needs to be elevated** for cancer
- **Tumor suppressors (e.g. p53) need to be turned off** in cancer



Douglas Hanahan and
Robert A. Weinberg
Hallmarks of Cancer:
The Next Generation
Cell 144, 2011

Statistics of cancer incidence vs age

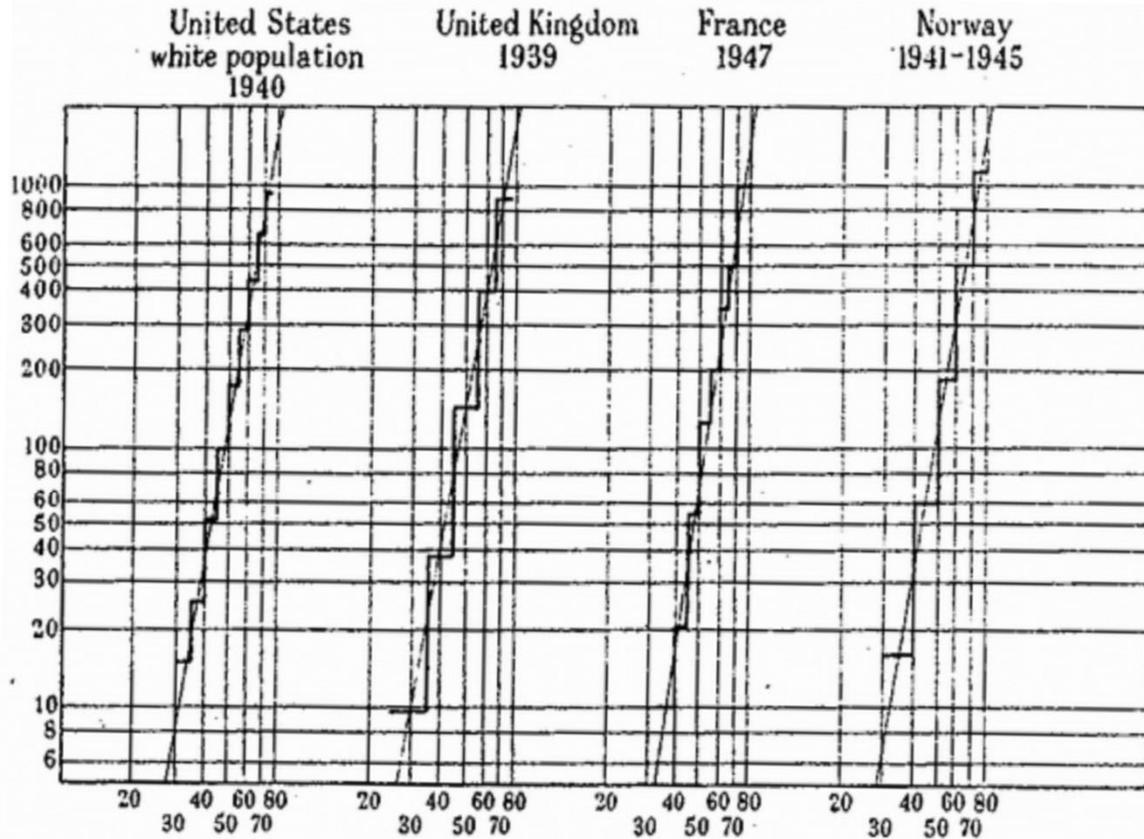


FIG. 1.—Diagram drawn to double logarithmic (log/log) scale showing the cancer death-rate (in the case of the United Kingdom, the carcinoma death-rate) in males at different ages. Deaths per 100,000 males are shown on the vertical scale, age figures on the horizontal scale.

Multi-mutation theory of cancer:
Carl O. Nordling (British J. of
Cancer, March 1953):

Cancer death rate
 $\sim (\text{patient age})^6$

It suggests the
existence of
 $k=7$ driver genes

$$P(T_{\text{cancer}} \leq t) \sim (u_1 t)(u_2 t) \dots (u_k t) \sim u_1 u_2 \dots u_k t^k$$

$$P(T_{\text{cancer}} = t) \sim \frac{d}{dt} (u_1 t)(u_2 t) \dots (u_k t) \sim k u_1 u_2 \dots u_k t^{k-1}$$

How many driver gene mutations for different types of cancer?

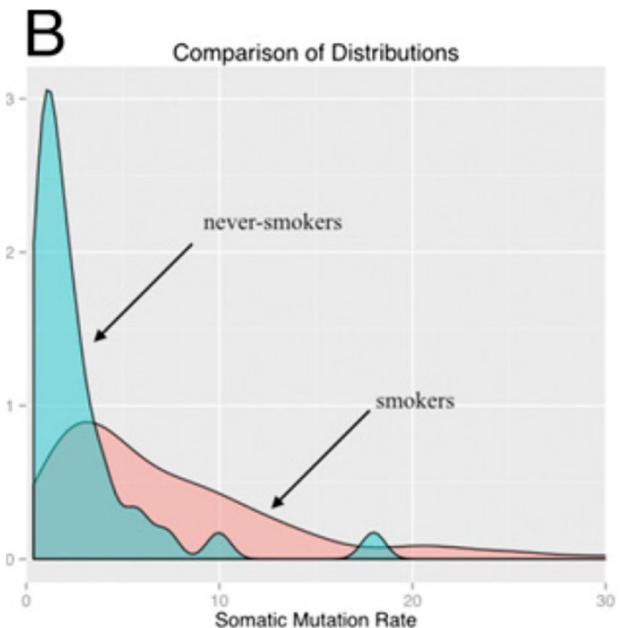
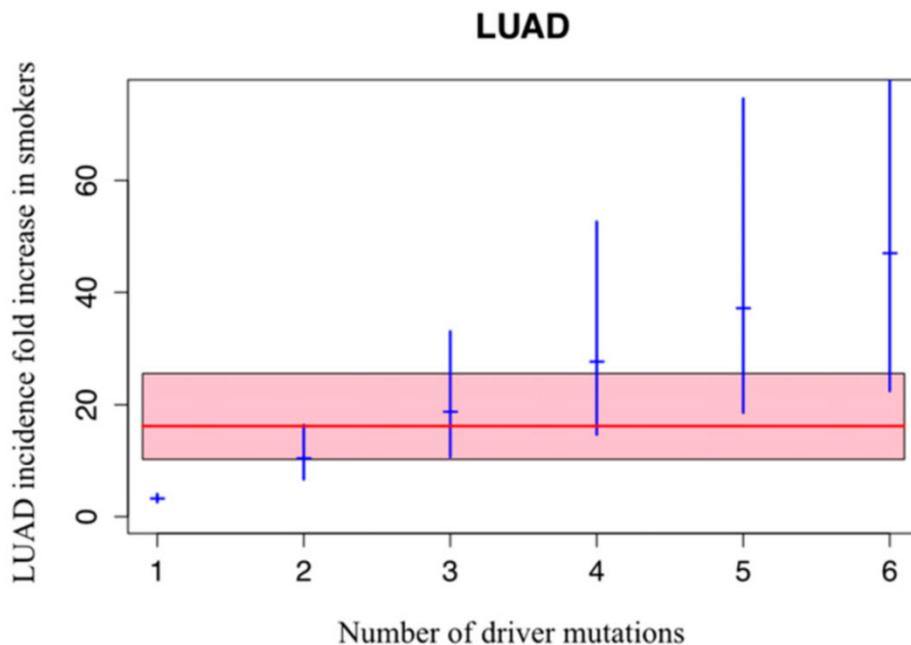
Only three driver gene mutations are required for the development of lung and colorectal cancers

Cristian Tomasetti^{a,b,1}, Luigi Marchionni^c, Martin A. Nowak^d, Giovanni Parmigiani^e, and Bert Vogelstein^{f,g,1}

^aDivision of Biostatistics and Bioinformatics, Department of Oncology, Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine, and ^bDepartment of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205; ^cCancer Biology Program, Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD 21205; ^dProgram for Evolutionary Dynamics, Department of Mathematics, Harvard University, Cambridge, MA 02138; ^eDepartment of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, MA 02215; and ^fLudwig Center for Cancer Genetics and Therapeutics and ^gHoward Hughes Medical Institute, Sidney Kimmel Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD 21205

Contributed by Bert Vogelstein, November 21, 2014 (sent for review July 31, 2014; reviewed by Zvia Agur)

Smokers have 3.23 times more mutations in lungs



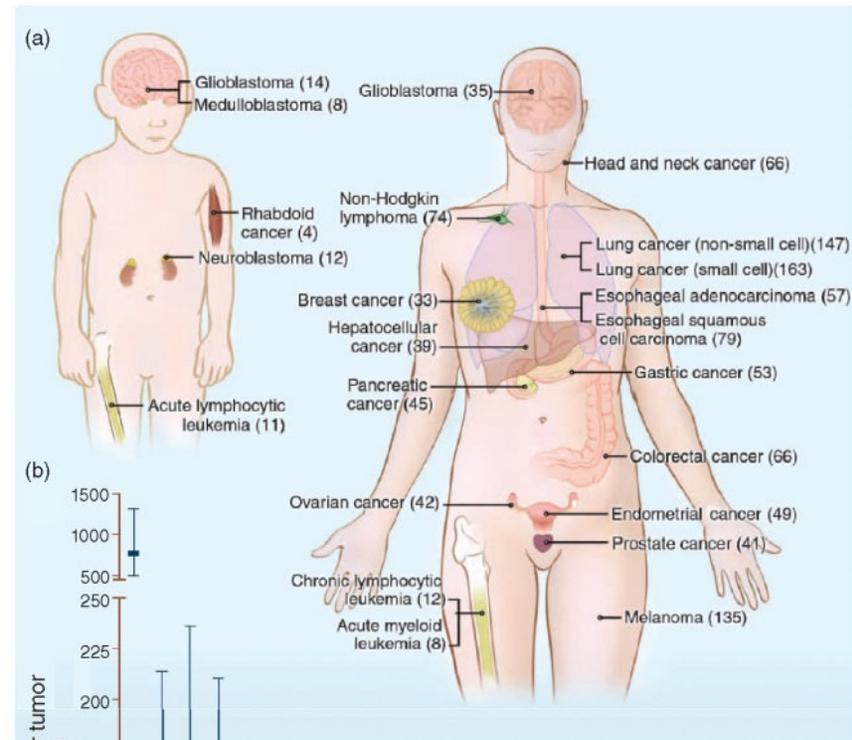
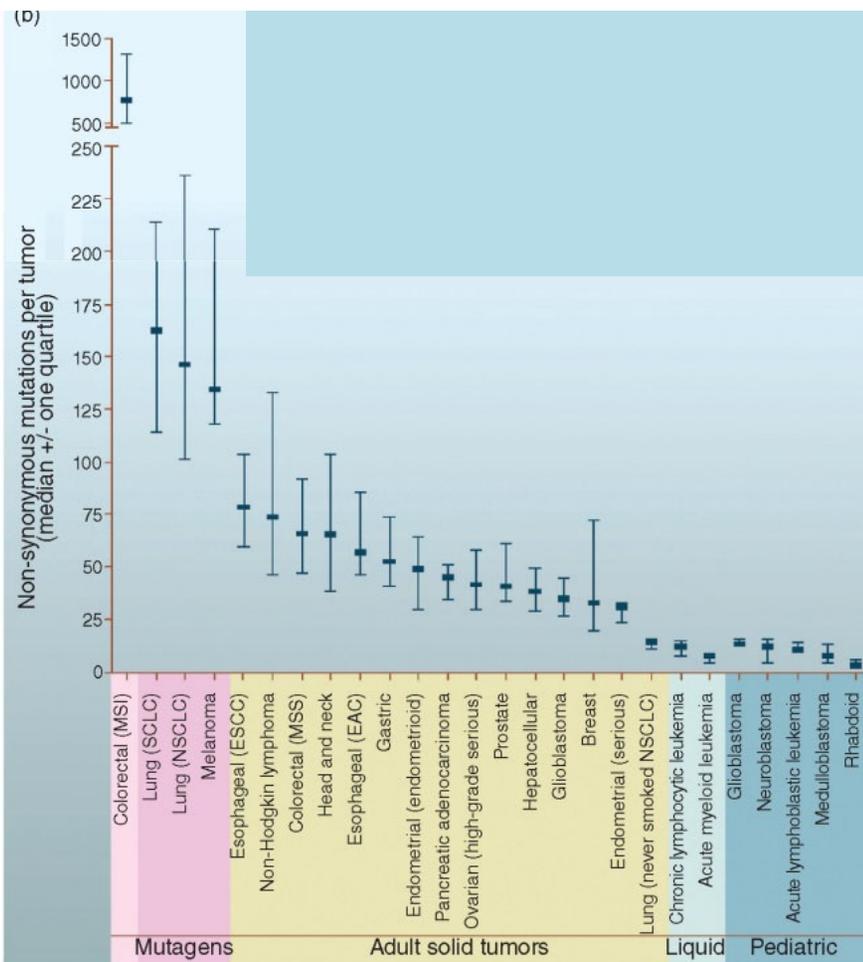


FIGURE 21.10 Somatic mutations in representative human cancers, based on genome-wide sequencing studies. (a) The genomes of adult (right) and pediatric (left) cancers are represented. Numbers in parentheses are the median number of nonsynonymous mutations per tumor. Redrawn from Vogelstein *et al.* (2013). Reproduced with permission from AAAS. (b) Median number of nonsynonymous substitutions per tumor. Horizontal bars indicate the 25% and 75% quartiles. MSI: microsatellite instability; SCLC: small cell lung cancers; NSCLC: non-small cell lung cancers; ESCC: esophageal squamous cell carcinomas; MSS: microsatellite stable; EAC: esophageal adenocarcinomas.

Bioinformatics and Functional Genomics, Third Edition, Jonathan Pevsner.
 © 2015 John Wiley & Sons, Ltd. Published 2015 by John Wiley & Sons, Ltd.
 Companion Website: www.wiley.com/go/pevsnerbioinformatics

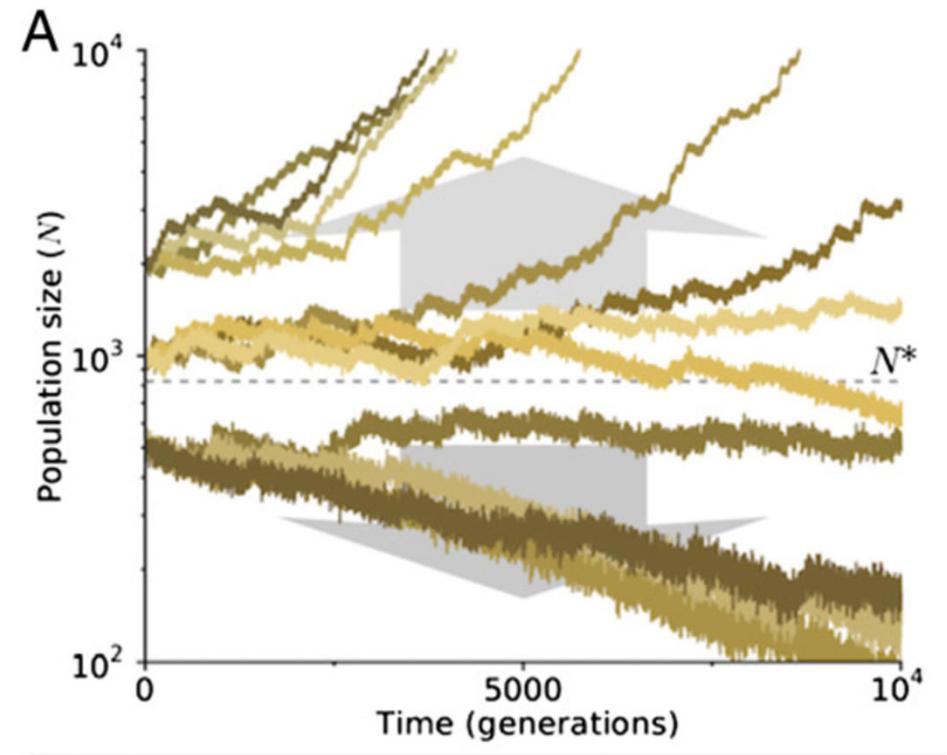
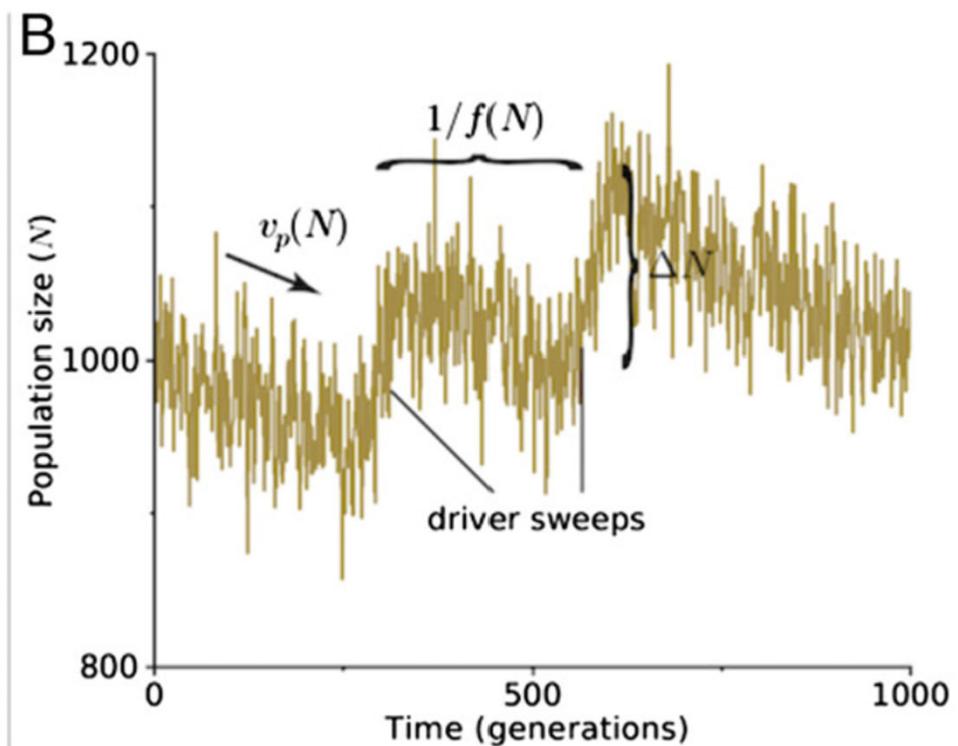
- Cancer cells carry both **“Driver”** and **“Passengers”** mutations
- **Passenger** mutations cause **little to no harm** (see later for how even little harm matters)
- Both are common as **cancers** **elevate mutation rate**

Number of passenger+driver mutations follows negative binomial distribution

- What is the **probability** to have n_p **passenger mutations** or (n_p+k) **total mutations** by the time you are diagnosed with cancer requiring k **driver mutations**?
- Let p is the probability that a mutation is a **driver** ($p = \text{Genome_target_of_driv} / (\text{Genome_target_of_driv} + \text{Genome_target_of_pass})$)
 $(1-p)$ – it is a **passenger mutation**

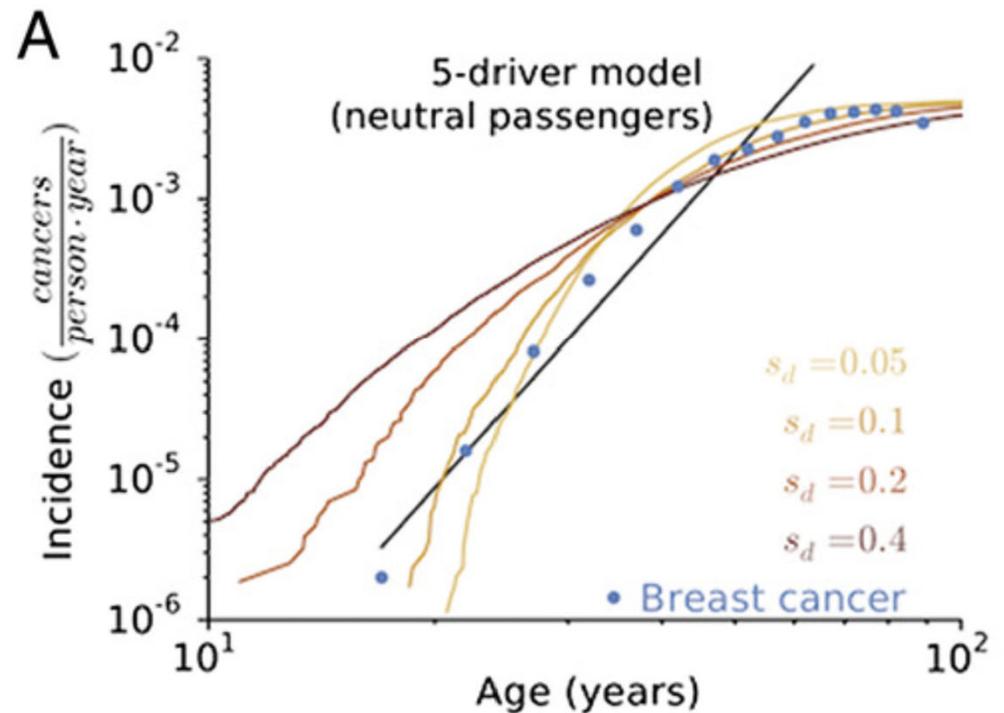
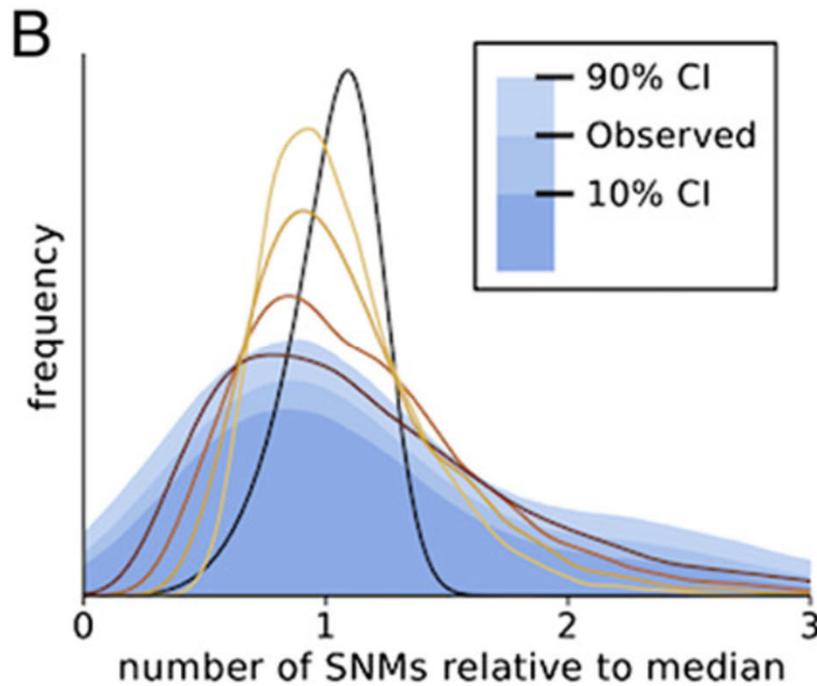
$$P(n_p + k | p, k) = \binom{n_p + k - 1}{n_p} (1-p)^{n_p} p^k$$

What if passenger mutations slow down the growth of cancer tumors?



McFarland CD, Mirny L, Korolev KS, PNAS 2014

Can we prove/quantify it using statistics?



Assume: growth rate of cancer = $(1+s_d)^{N_d} / (1+s_p)^{N_p}$

$\mu = 10^{-8}$, $\text{Target}_d = 1,400$, $\text{Target}_p = 10^7$, $s_d = 0.05$ to 0.4 , $s_p = 0.001$

s_p/s_d for breast: 0.0060 ± 0.0010 ;

melanoma: 0.016 ± 0.003 ; lung: 0.0094 ± 0.0093 ;

Blue - data on breast cancer: incidence; non-synonymous mutations

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 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 DO IGUANAS DIE
WHY AREN'T THERE DINOSAUR GHOSTS

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 SWARMS OF GNATS
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE

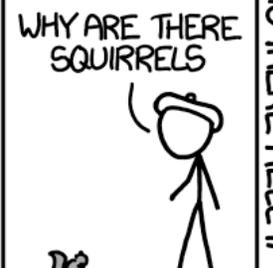
WHY ARE THERE ANTS IN MY LAPTOP
WHY IS EARTH TILTED
WHY IS SPACE BLACK
WHY IS OUTER SPACE SO COLD
WHY ARE THERE PYRAMIDS ON THE MOON
WHY IS NASA SHUTTING DOWN



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

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 ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD

WHY IS SEX SO IMPORTANT



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 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 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

Register for the midterm #1
on CBTF
and take the exam between
11am Tuesday 2/24/2026
and all day Thursday 2/26/2026

No Tuesday class next week
Thursday class is in session

Midterm Info

- Closed book exam at CBTF; no books, notes, laptops, phones...
- Calculators (not on smartphones) are provided by CBTF
- The following printout and the lecture slides will be provided

Name	Probability Distribution	Mean	Variance	Section in Book
Discrete				
Uniform	$\frac{1}{n}, a \leq b$	$\frac{(b+a)}{2}$	$\frac{(b-a+1)^2-1}{12}$	3-5
Binomial	$\binom{n}{x} p^x (1-p)^{n-x},$ $x = 0, 1, \dots, n, 0 \leq p \leq 1$	np	$np(1-p)$	3-6
Geometric	$(1-p)^{x-1} p,$ $x = 1, 2, \dots, 0 \leq p \leq 1$	$1/p$	$(1-p)/p^2$	3-7.1
Negative binomial	$\binom{x-1}{r-1} (1-p)^{x-r} p^r$ $x = r, r+1, r+2, \dots, 0 \leq p \leq 1$	r/p	$r(1-p)/p^2$	3-7.2
Poisson	$\frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots, 0 < \lambda$	λ	λ	3-9
Continuous				
Uniform	$\frac{1}{b-a}, a \leq x \leq b$	$\frac{(b+a)}{2}$	$\frac{(b-a)^2}{12}$	4-5
Normal	$\frac{1}{\sigma\sqrt{2\pi}} e^{-1/2(\frac{x-\mu}{\sigma})^2}$ $-\infty < x < \infty, -\infty < \mu < \infty, 0 < \sigma$	μ	σ^2	4-6
Exponential	$\lambda e^{-\lambda x}, 0 \leq x, 0 < \lambda$	$1/\lambda$	$1/\lambda^2$	4-8
Erlang	$\frac{\lambda^r x^{r-1} e^{-\lambda x}}{(r-1)!}, 0 < x, r = 1, 2, \dots$	r/λ	r/λ^2	4-9.1
Gamma	$\frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}, 0 < x, 0 < r, 0 < \lambda$	r/λ	r/λ^2	4-9.2

What is included in the midterm?

- Probability of events (set operations), Multiplication rules. Combinatorics
- Bayes Theorem
- Discrete Random Variables
- Other topics covered
(see HW1-HW2 for inspiration)
- No Matlab exercises (since no computers)

Probability Multiplication Rules

Combinatorics

Mr. Jones has 6 different books that he is going to put on his bookshelf. Of these, 3 are chemistry books, 2 are physics books, and 1 is a mathematics book. Jones wants to arrange his books so that two conditions are met:

(1) all the books dealing with the same subject are together on the shelf

AND

(2) all chemistry books are on the leftmost side.

How many such different arrangements are possible?

Mr. Jones has 6 different books that he is going to put on his bookshelf. Of these, 3 are chemistry books, 2 are physics books, and 1 is a mathematics book. Jones wants to arrange his books so that two conditions are met:

(1) all the books dealing with the same subject are together on the shelf

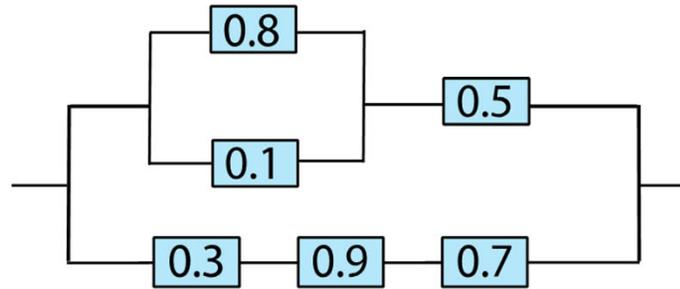
AND

(2) all chemistry books are on the leftmost side.

How many such different arrangements are possible?

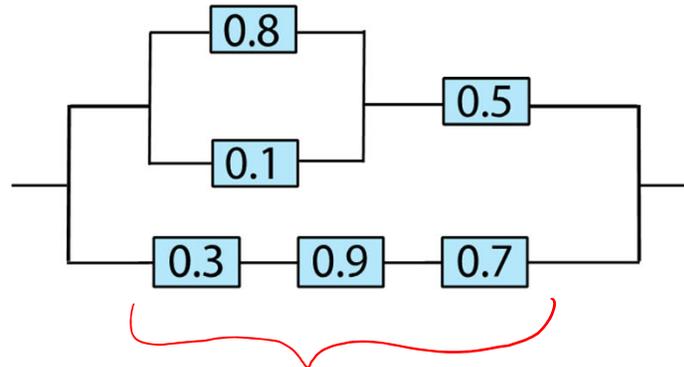
Answer: $(3! * 2! * 1!) * 2! = 24$

4. (4 points) The following circuit operates if and only if there is a path of functional devices from left to right. The probability that each device functions is as shown. Assume that the probability that a device is functional does not depend on whether or not other devices are functional. What is the probability that the circuit operates?



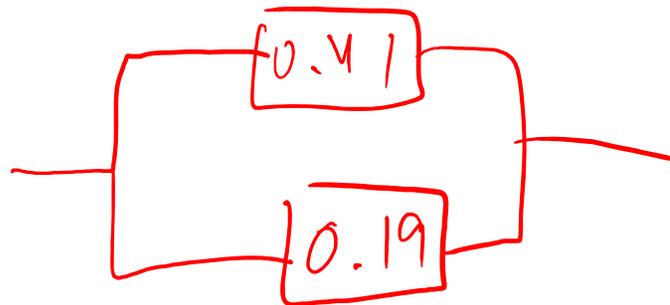
4. (4 points) The following circuit operates if and only if there is a path of functional devices from left to right. The probability that each device functions is as shown. Assume that the probability that a device is functional does not depend on whether or not other devices are functional. What is the probability that the circuit operates?

$$1 - (1 - 0.8) \cdot (1 - 0.1) = 0.82$$



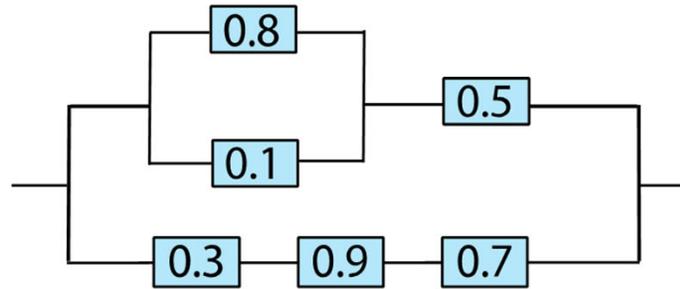
$$0.3 \cdot 0.9 \cdot 0.7 = 0.19$$

$$0.82 \times 0.5 = 0.41$$



$$1 - (1 - 0.41) \cdot (1 - 0.19) = 0.52$$

4. (4 points) The following circuit operates if and only if there is a path of functional devices from left to right. The probability that each device functions is as shown. Assume that the probability that a device is functional does not depend on whether or not other devices are functional. What is the probability that the circuit operates?



Answer: $P(\text{Operate}) = 1 - (1 - 0.3 * 0.9 * 0.7) * (1 - 0.5 * (1 - (1 - 0.8) * (1 - 0.1))) = \underline{0.52}$

Bayes theorem

In answering a question on a multiple-choice test, a student either knows the answer or he guesses. Let $\frac{1}{3}$ be the probability that he knows the answer. If he does not know the answer, he randomly guesses one out of 4 multiple choice questions. What is the conditional probability that a student knew the answer to a question given that he answered it correctly?

- A. $\frac{1}{4}$
- B. $\frac{1}{3}$
- C. $\frac{2}{3}$
- D. $\frac{1}{5}$
- E. I don't know

Get your i-clickers

In answering a question on a multiple-choice test, a student either knows the answer or he guesses. Let $\frac{1}{3}$ be the probability that he knows the answer. If he does not know the answer, he randomly guesses one out of 4 multiple choice questions. What is the conditional probability that a student knew the answer to a question given that he answered it correctly?

A. $\frac{1}{4}$

B. $\frac{1}{3}$

C. $\frac{2}{3}$

D. $\frac{1}{5}$

E. I don't know

Get your i-clickers

In answering a question on a multiple-choice test, a student either knows the answer or he guesses. Let $1/3$ be the probability that he knows the answer. If he does not know the answer, he randomly guesses one out of 4 multiple choice questions. What is the conditional probability that a student knew the answer to a question given that he answered it correctly?

Answer: $P(K)=1/3$, $P(K')=2/3$, $P(C|K)=1$, $P(C|K')=1/4$.

$P(K|C)=P(C|K)*P(K)/P(C)=1*(1/3)/(1*1/3+(1/4)*(2/3))=2/3=0.666\dots$

(10 points) Suppose that a bag contains ten coins, three of which are fair, while the remaining seven are biased: they have probability of 0.6 of heads when flipped. A coin was taken at random from the bag and flipped five times. All five flips gave heads. What's the probability that this coin is fair?

(10 points) Suppose that a bag contains ten coins, three of which are fair, the remaining seven having probability 0.6 of giving heads when flipped. A coin is taken at random from the bag and flipped five times. All five flips give heads. What's the probability that a coin is fair given the five coin flips?

Answer: Let us denote H_1 as the hypothesis that a coin is fair and H_2 that a coin is biased. The data that all five flips were heads is denoted as D . Therefore,

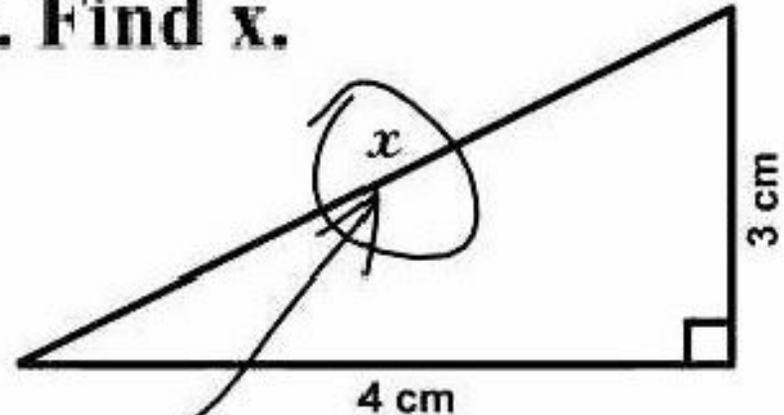
$$P(H_1|D) = P(D|H_1)P(H_1)/P(D) = 0.5^5 * 0.3 / (0.5^5 * 0.3 + 0.6^5 * 0.7) = 0.147$$

Discrete Probability Distributions

What is X in this problem?

- What is the random variable: Look for keywords:
 - Find the probability that....
 - What is the mean (or variance) of...
- What are parameters? Look for keywords:
 - Given that...
 - Assuming that...

3. Find x .



Here it is

Guide to probability distributions

- Binomial: # of samples, n , is fixed, # of successes, x , is variable

$$P(X=x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

- Geometric: # of samples, x is variable. # of successes 1 is fixed.

Success comes in the end

$$P(X=x) = (1-p)^{x-1} \cdot p$$

- Negative binomial: # of samples, x is variable. # of successes, r , is fixed
 r th success in the end

$$P(X=x) = \frac{(x-1)!}{(r-1)!(x-r)!} p^r (1-p)^{x-r}$$

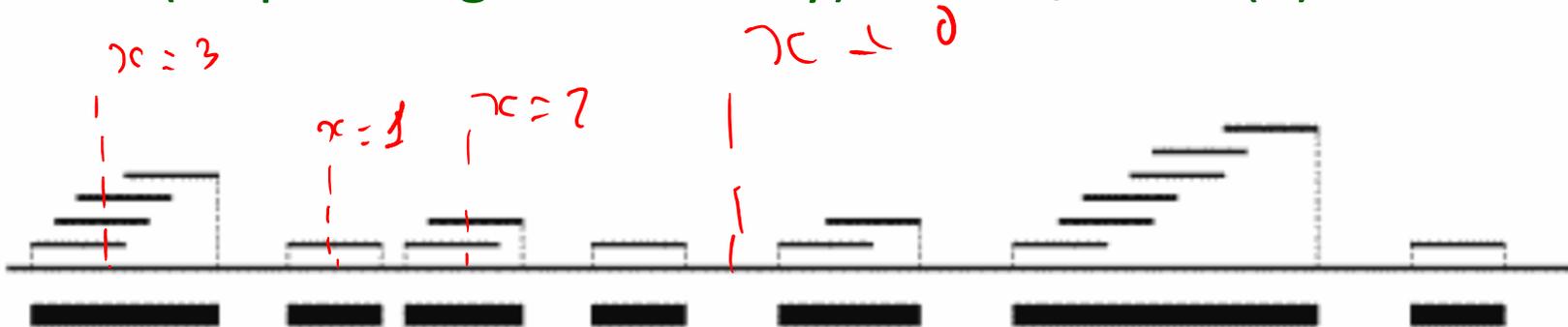
Poisson distribution in genomics

- G - genome length (in bp)
- L - short read average length
- N - number of short read sequenced
- λ - sequencing redundancy = LN/G
- x - number of short reads covering a given site on the genome

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Ewens, Grant, Chapter 5.1

Poisson as a limit of Binomial. For a given site on the genome for each short read Prob(site covered): $p=L/G$ is very small. Number of attempts (short reads): N is very large. Their product (sequencing redundancy): $\lambda = NL/G$ is $O(1)$.



Probability that a base pair in the genome is not covered by any short reads is 0.1

One randomly selects base pairs until exactly 5 uncovered base pairs are found.

Which discrete probability distribution describes the number of attempts?

- A. Poisson
- B. Binomial
- C. Geometric
- D. Negative Binomial
- E. I have no idea

Poisson	$\frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots, 0 < \lambda$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n, 0 \leq p \leq 1$
Geometric	$(1-p)^{x-1} p$ $x = 1, 2, \dots, 0 \leq p \leq 1$
Negative binomial	$\binom{x-1}{r-1} (1-p)^{x-r} p^r$ $x = r, r+1, r+2, \dots, 0 \leq p \leq 1$

Get your i-clickers

Probability that a base pair in the genome is not covered by any short reads is 0.1

One randomly selects base pairs until exactly 5 uncovered base pairs are found.

Which discrete probability distribution describes the number of attempts?

- A. Poisson
- B. Binomial
- C. Geometric
- D. Negative Binomial**
- E. I have no idea

Poisson	$\frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \dots, 0 < \lambda$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n, 0 \leq p \leq 1$
Geometric	$(1-p)^{x-1} p$ $x = 1, 2, \dots, 0 \leq p \leq 1$
Negative binomial	$\binom{x-1}{r-1} (1-p)^{x-r} p^r$ $x = r, r+1, r+2, \dots, 0 \leq p \leq 1$

Get your i-clickers

Probability that a base pair in the genome is not covered by any short reads is 0.1

One randomly selects base pairs until exactly 5 uncovered base pairs are found.

What are the values of p , r ?

- A. $p=0.5, r=5$
- B. $p=0.1, r=0.5$
- C. $p=0.1, r=5$
- D. $p=0.5, r=0.1$
- E. I have no idea

Poisson	$\frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots, 0 < \lambda$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n, 0 \leq p \leq 1$
Geometric	$(1-p)^{x-1} p$ $x = 1, 2, \dots, 0 \leq p \leq 1$
Negative binomial	$\binom{x-1}{r-1} (1-p)^{x-r} p^r$ $x = r, r+1, r+2, \dots, 0 \leq p \leq 1$

Get your i-clickers

Probability that a base pair in the genome is not covered by any short reads is 0.1

One randomly selects base pairs until exactly 5 uncovered base pairs are found.

What are the values of p , r ?

- A. $p=0.5, r=5$
- B. $p=0.1, r=0.5$
- C. $p=0.1, r=5$**
- D. $p=0.5, r=0.1$
- E. I have no idea

Poisson	$\frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots, 0 < \lambda$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n, 0 \leq p \leq 1$
Geometric	$(1-p)^{x-1} p$ $x = 1, 2, \dots, 0 \leq p \leq 1$
Negative binomial	$\binom{x-1}{r-1} (1-p)^{x-r} p^r$ $x = r, r+1, r+2, \dots, 0 \leq p \leq 1$

Get your i-clickers

Cancer happens when the gene p53 mutates.

Probability of p53 to mutate per year is 5%.

How many years before a patient gets disease?

Which discrete probability distribution would you use to answer?

- A. Poisson
- B. Binomial
- C. Geometric
- D. Negative Binomial
- E. I have no idea

Poisson	$\frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots, 0 < \lambda$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n, 0 \leq p \leq 1$
Geometric	$(1-p)^{x-1} p$ $x = 1, 2, \dots, 0 \leq p \leq 1$
Negative binomial	$\binom{x-1}{r-1} (1-p)^{x-r} p^r$ $x = r, r+1, r+2, \dots, 0 \leq p \leq 1$

Get your i-clickers

Cancer happens when the gene p53 mutates.

Probability of p53 to mutate per year is 5%.

How many years before a patient gets disease?

Which discrete probability distribution would you use to answer?

- A. Poisson
- B. Binomial
- C. Geometric**
- D. Negative Binomial
- E. I have no idea

Poisson	$\frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots, 0 < \lambda$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n, 0 \leq p \leq 1$
Geometric	$(1-p)^{x-1} p$ $x = 1, 2, \dots, 0 \leq p \leq 1$
Negative binomial	$\binom{x-1}{r-1} (1-p)^{x-r} p^r$ $x = r, r+1, r+2, \dots, 0 \leq p \leq 1$

Get your i-clickers

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 ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS THERE PHLEGM

WHY DO AMERICANS CALL IT SOCCER

WHY IS PSYCHIC WEAK TO BUG

WHY ARE MY EARS RINGING

WHY DO CHILDREN GET CANCER

WHY ARE THERE SO MANY AVENGERS

WHY IS POSEIDON ANGRY WITH ODYSSEUS

WHY ARE THE AVENGERS FIGHTING THE X MEN

WHY IS THERE ICE IN SPACE

WHY ARE THERE ANTS IN MY LAPTOP

WHY IS EARTH TILTED



WHY IS THERE AN OWL IN MY BACKYARD

WHY IS SPACE BLACK



WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS OUTER SPACE SO COLD

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY ARE THERE PYRAMIDS ON THE MOON

WHY DO OWLS ATTACK PEOPLE

WHY IS NASA SHUTTING DOWN

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE MALE AND FEMALE BIKES

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY ARE THERE GODS

WHY DO SPIDERS COME INSIDE

WHY ARE THERE TWO SPOCKS

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY IS MT VESUVIUS THERE

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY DO THEY SAY T MINUS

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE OBELISKS

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY ARE WRESTLERS ALWAYS WET

WHY DO SPIDER BITES ITCH

WHY ARE OCEANS BECOMING MORE ACIDIC

WHY IS DYING SO SCARY

WHY IS ARWEN DYING



WHY IS LIFE SO BORING
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

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 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 THERE DINOSAUR GHOSTS

WHY ARE THERE FEMALE MR NIMES

WHY IS SEX SO IMPORTANT



WHY IS THERE HELL IF GOD FORGIVES

WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD

WHY IS GPS FREE