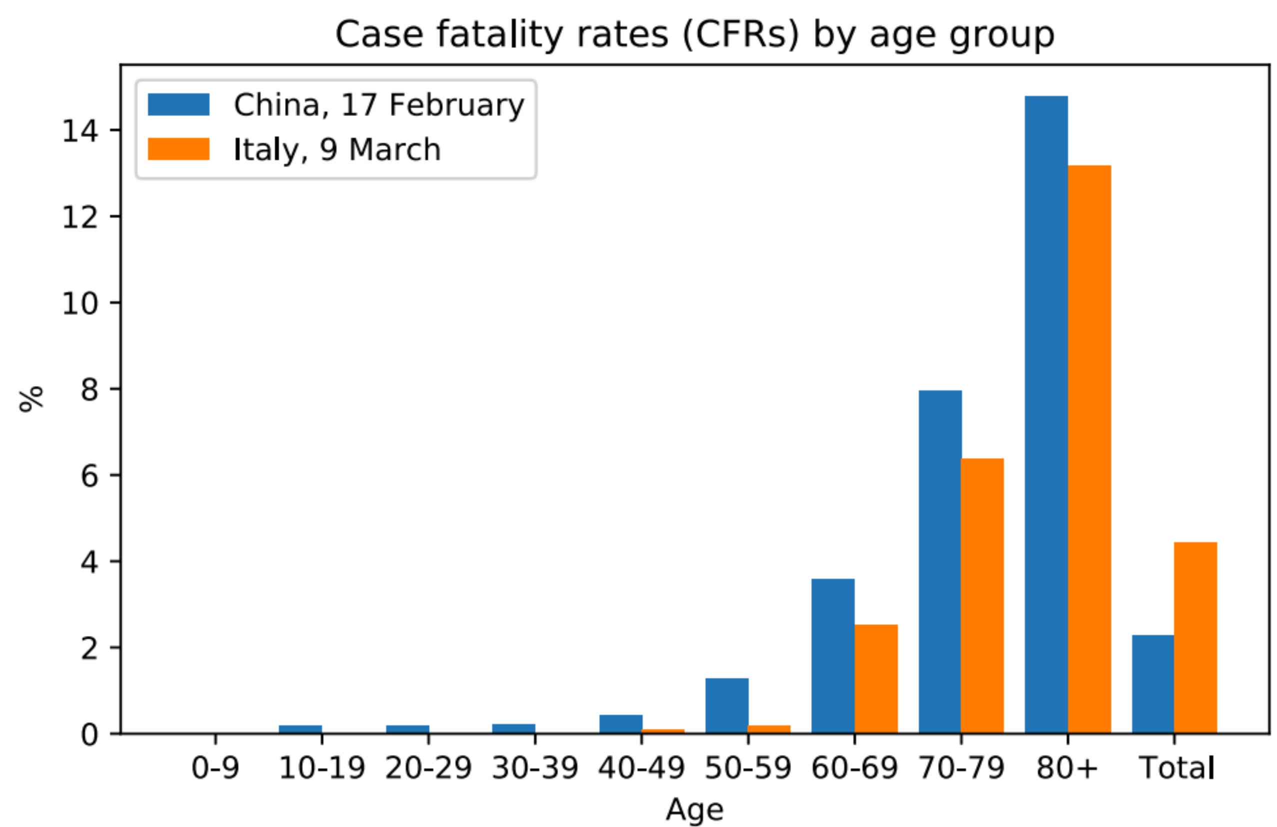
Epidemiology - Part 1

Feb. 11, 2025

Recap question:

Feb. 11, 2025

1. Critique the following incorrect statement: in the Israeli ICU, 515 people are hospitalized due to COVID. Of these, 214 were not vaccinated and 301 were vaccinated. Therefore, the vaccines do not work (data from Sept. 2021)
2. Explain why the last bar of this graph does not contradict the rest of the bars

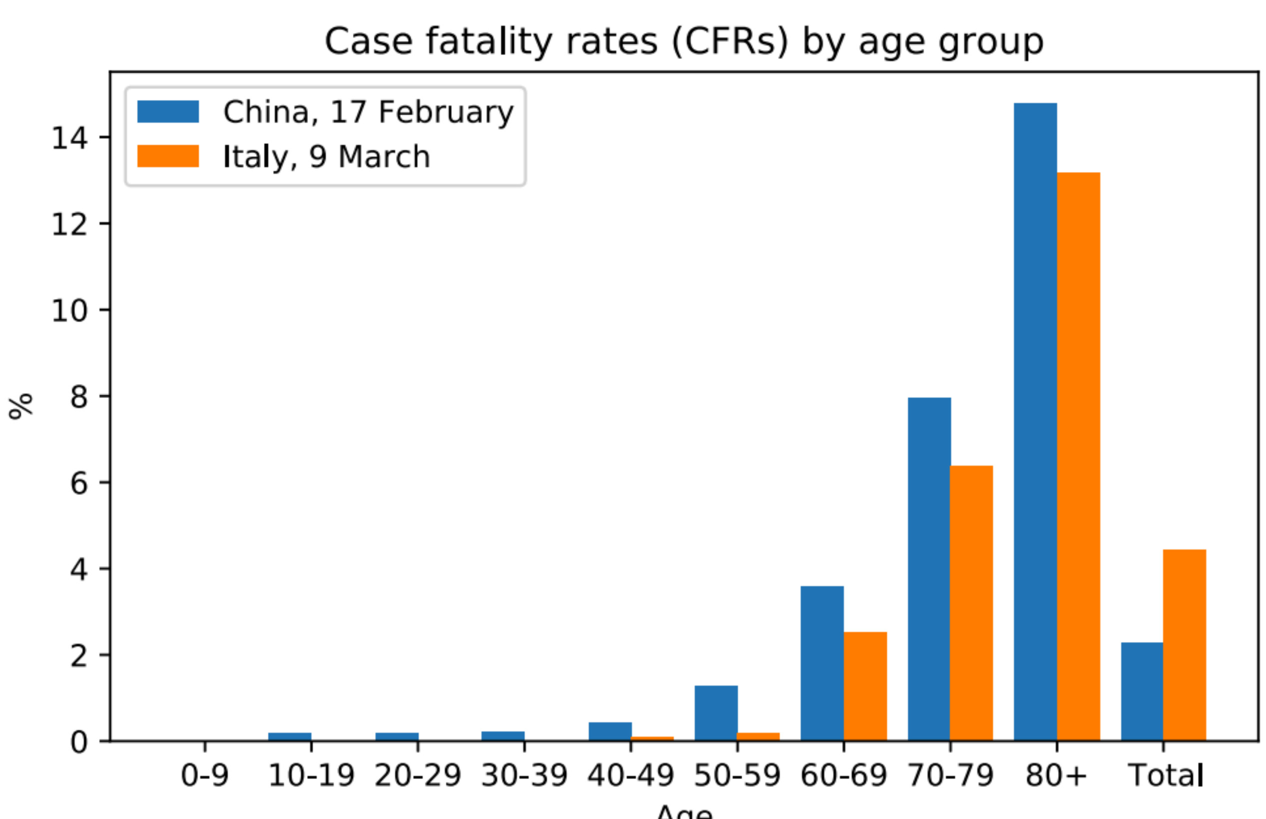


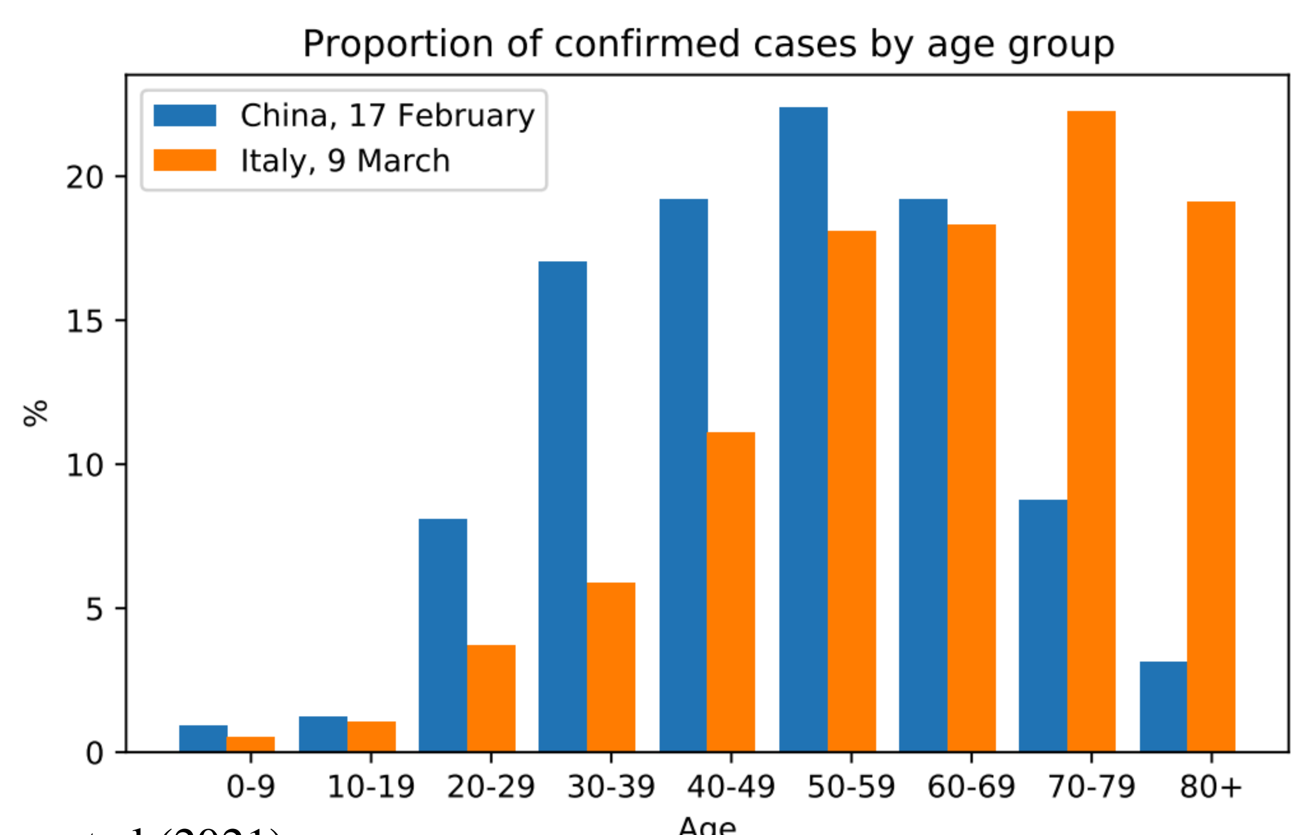
von Kügelgen et al (2021)

Recap question:

Feb. 11, 2025

1. There could be a Simpson's paradox lurking behind the numbers, …
2. This is Simpson's paradox: there were more infected in the oldest age group in Italy, which raised overall mortality levels





von Kügelgen et al (2021)

Epidemiology - Part 1

Feb. 11, 2025

By the end of this lecture, you will be able to:

1. Recap p-values
2. Explain Dorfman's testing method
3. Calculate sensitivity and specificity of repeated COVID tests

Last bits of hypothesis testing

Hypothesis testing

Consider the following example:
48 bank supervisors (all male) were given the same personnel file and asked whether the person should be promoted or not. The files are identical, except that 24 of the files were assigned to belong to male employees and 24 to females. Of the 48 files, 35 were promoted, 21 of which belonged to males and the rest belonged to females.

The percentage of men promoted $= \frac{21}{24} = 88\%$

The percentage of women promoted $= \frac{14}{24} = 58\%$

So there is a 30% difference between men and women promoted. Could this have been due to chance?

Terminology:

H_0 : null hypothesis that there is no discrimination, i.e. each person is equally likely to get promoted.

H_A: alternative hypothesis, i.e., there is some sort of discrimination. Ofter this is just the complement of H_0.

p-value: probability of obtaining results that are at least as extreme as the observed results, if the null hypothesis were correct, i.e.,

$$p = P(\text{difference at least } 30\% | H_0) \approx 0.01$$

What value of the p - value is significant?

This is a value judgement and differs between fields, depending on how serious it would be to draw the wrong conclusion. Given a level α of significance, we reject the null hypothesis if we witness the p - value is smaller than α .

$\alpha = 0.05$ or $\alpha = 0.01$ is common in biomedical research

$\alpha = 0.0000003$ was used when discovering the Higgs boson

	H_0 is true	H_0 is false
Do not reject H_0	Correct inference	Type II error
Reject H_0	Type I error	Correct inference

Example 2:

Tom has two roommates: Ryan and Hugo. Every week, Tom draws a name out of a bucket to randomly select the roommates to take the trash out that week. Hugo suspects that Tom is cheating, so he starts keeping track of the draws, and he finds that out of 12 draws, Tom didn't get picked even once!

H_0 : Tom is not cheating so each roommate gets picked 1/3 of the time

H_A : Tom is cheating

P(Tom not picked in a given draw $| H_0$) $=$

P(Tom not picked in 12 consecutive draws $| H_0$)

$=$

Hint: the probabilities of **independent** events multiply $P(A \text{ and } B) = P(A) \cdot P(B)$

$p =$

P(Tom not picked in a given draw $| H_0) = \frac{2}{3}$

P(Tom not picked in 12 consecutive draws $| H_0$)

$=$

Hint: the probabilities of **independent** events multiply $P(A \text{ and } B) = P(A) \cdot P(B)$

$p =$

P(Tom not picked in a given draw $| H_0$) $= \frac{2}{3}$

P(Tom not picked in 12 consecutive draws $| H_0$)

$$= \left(\frac{2}{3}\right)^{12} \approx 0.8\%$$

$p =$

P(Tom not picked in a given draw $| H_0) = \frac{2}{3}$

P(Tom not picked in 12 consecutive draws $| H_0$)

$$= \left(\frac{2}{3}\right)^{12} \approx 0.8\%$$

$$p = P(\text{Tom not picked in at least 12 draws}|H_0)$$

P(Tom not picked in a given draw $| H_0) = \dfrac{2}{3}$

P(Tom not picked in 12 consecutive draws $| H_0$)

$$= \left(\frac{2}{3}\right)^{12} \approx 0.8\%$$

$$p = P(\text{Tom not picked in at least 12 draws}|H_0)$$
$$= P(\text{Tom not picked in 12 draws}|H_0) \approx 0.008$$

↑ because there were 12 draws in total

What value of the p - value is significant?

This is a value judgement and differs between fields, depending on how serious it would be to draw the wrong conclusion

$p = 0.05$ or $p = 0.01$ is common in biomedical research

$p = 0.0000003$ was used when discovering the Higgs boson

	H_0 is true	H_0 is false
Do not reject H_0	Correct inference	Type II error
Reject H_0	Type I error	Correct inference

Do not reject $\neq$ Accept!!!

Drawbacks of the Framework of Hypothesis Testing

- More inclined towards not rejecting the null hypothesis
- The result depends on the choice of null hypothesis

Epidemiology is the study and analysis of the incidence, distribution, and possible control of diseases and other factors relating to health.

Types of COVID-19 tests

Diagnostic tests: show if you have an active coronavirus infection.

- molecular tests
- antigen tests

Antibody tests: show if you have been infected by coronavirus in the past

Used at Princeton

Molecular tests (Ex: RT-PCR)	Antigen tests
saliva, nasal or throat swab	nasal or throat swab
detect genetic materials of the virus	detect specific proteins of the virus
highly sensitive and specific	highly specific but less sensitive
a few hours to days	less than an hour

Used at Princeton

Molecular tests (Ex: RT-PCR)	Antigen tests
saliva, nasal or throat swab	nasal or throat swab
detect genetic materials of the virus	detect specific proteins of the virus
highly sensitive and specific	highly specific but less sensitive
a few hours to days	less than an hour

Challenges in testing

Resources: having an adequate number of trained health professionals to collect and process samples along with an adequate supply of reagents and testing machines

Time: having the test results available promptly to avoid infected people from unknowingly spreading the virus to others.

At Princeton

Undergraduate Student Data

Week Ending ▾	Tests	Positive Cases	Positivity Rate	Change in Positivity Rate
Feb 25, 2022	5,870	326	5.55%	0.66%
Feb 18, 2022	6,131	300	4.89%	4.51%
Feb 11, 2022	5,537	21	0.38%	0.09%
Feb 4, 2022	8,015	23	0.29%	-1.03%
Jan 28, 2022	9,344	123	1.32%	-0.57%

51 - 55 / 129 ‹ ›

Graduate Student Data

Week Ending ▾	Tests	Positive Cases	Positivity Rate	Change in Positivity rate
Feb 18, 2022	2,141	5	0.23%	-0.07%
Feb 11, 2022	2,319	7	0.3%	0.07%
Feb 4, 2022	3,408	8	0.23%	-0.25%
Jan 28, 2022	3,491	17	0.49%	-0.38%
Jan 21, 2022	2,641	23	0.87%	-0.17%

51 - 55 / 128 ‹ ›

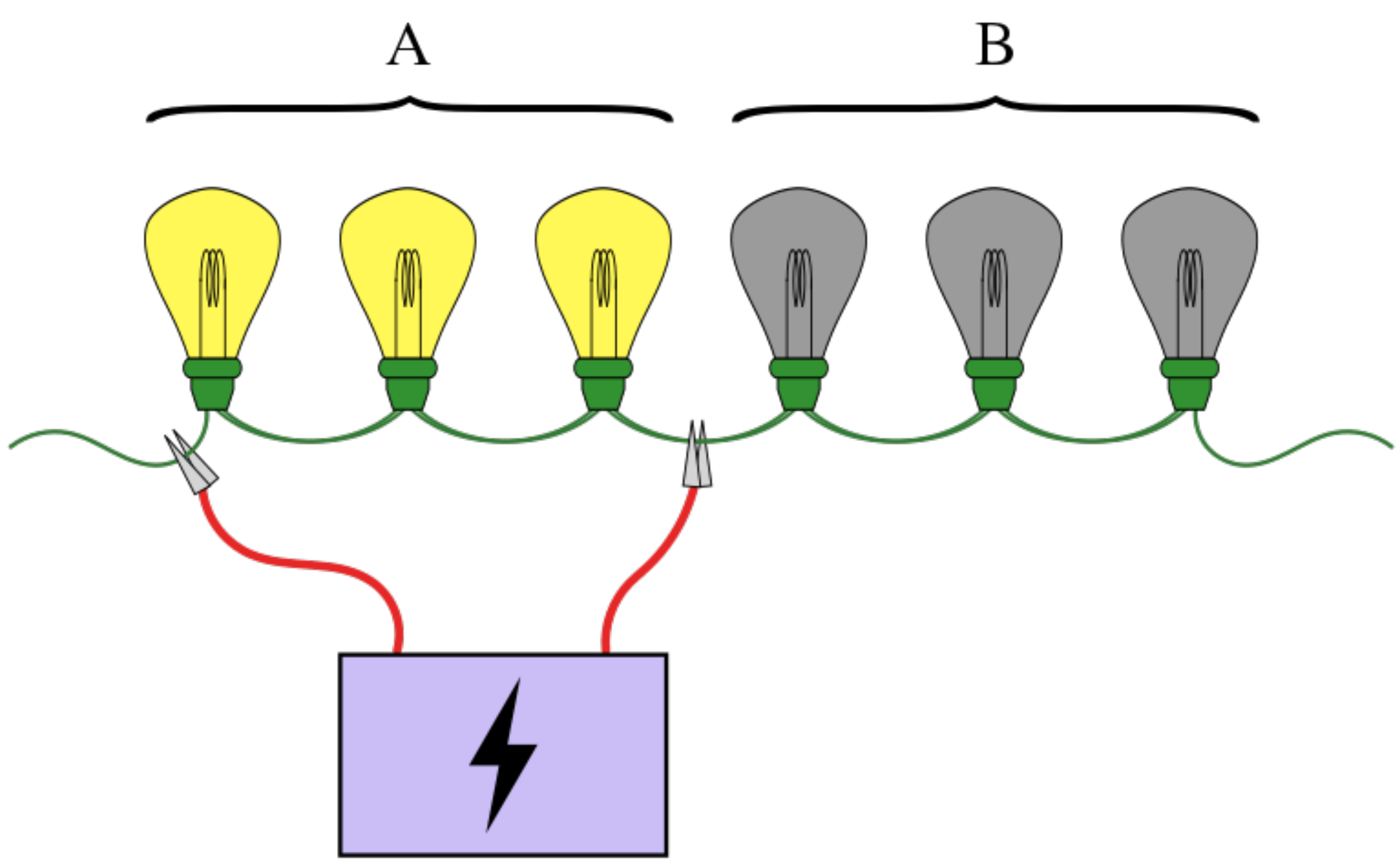
Faculty, Staff & Other Data

Week Ending ▾	Tests	Positive Cases	Positivity Rate	Change in Positivity Rate
Feb 25, 2022	5,360	15	0.28%	0.02%
Feb 18, 2022	5,477	14	0.26%	-0.01%
Feb 11, 2022	5,593	15	0.27%	-0.28%
Feb 4, 2022	5,445	30	0.55%	-0.37%
Jan 28, 2022	4,025	37	0.92%	-0.85%

51 - 55 / 129 ‹ ›

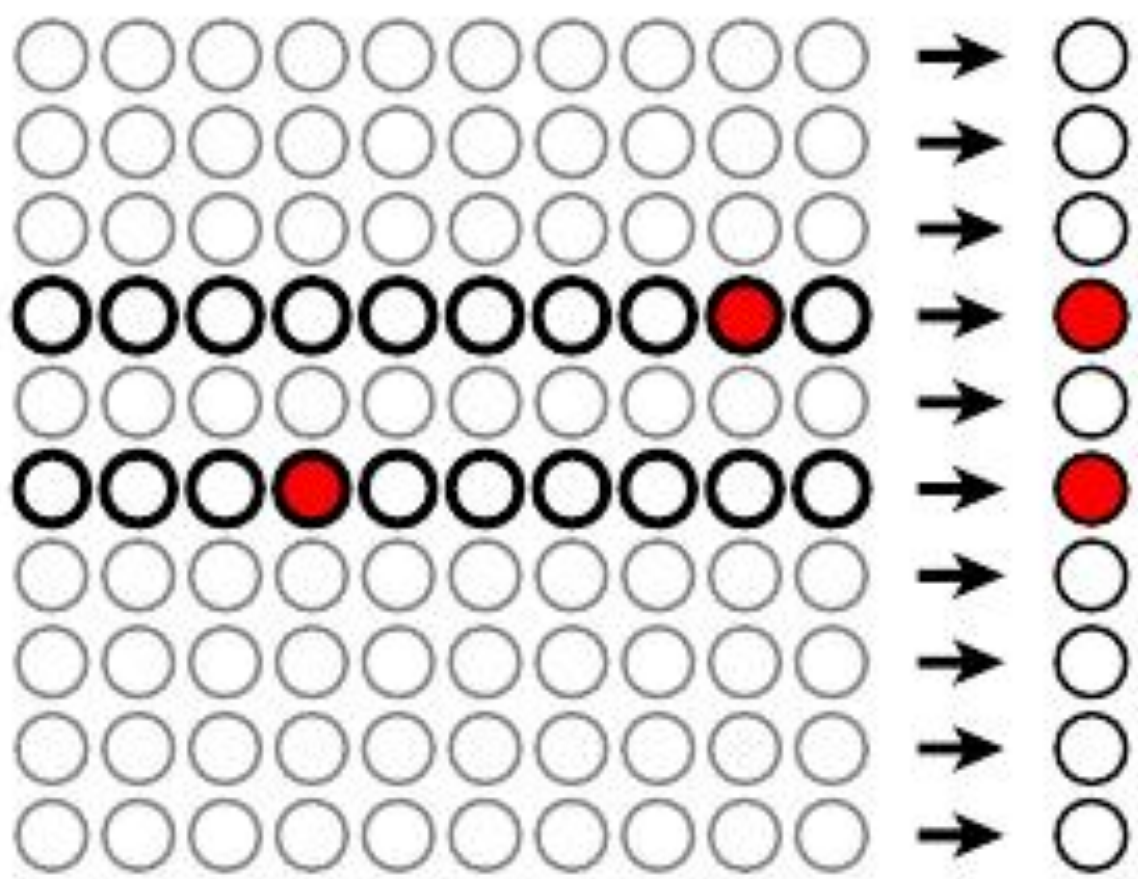
Pool testing

combine samples from many patients into testing pools strategically rather than testing samples from each individual patient separately



Dorfman's method

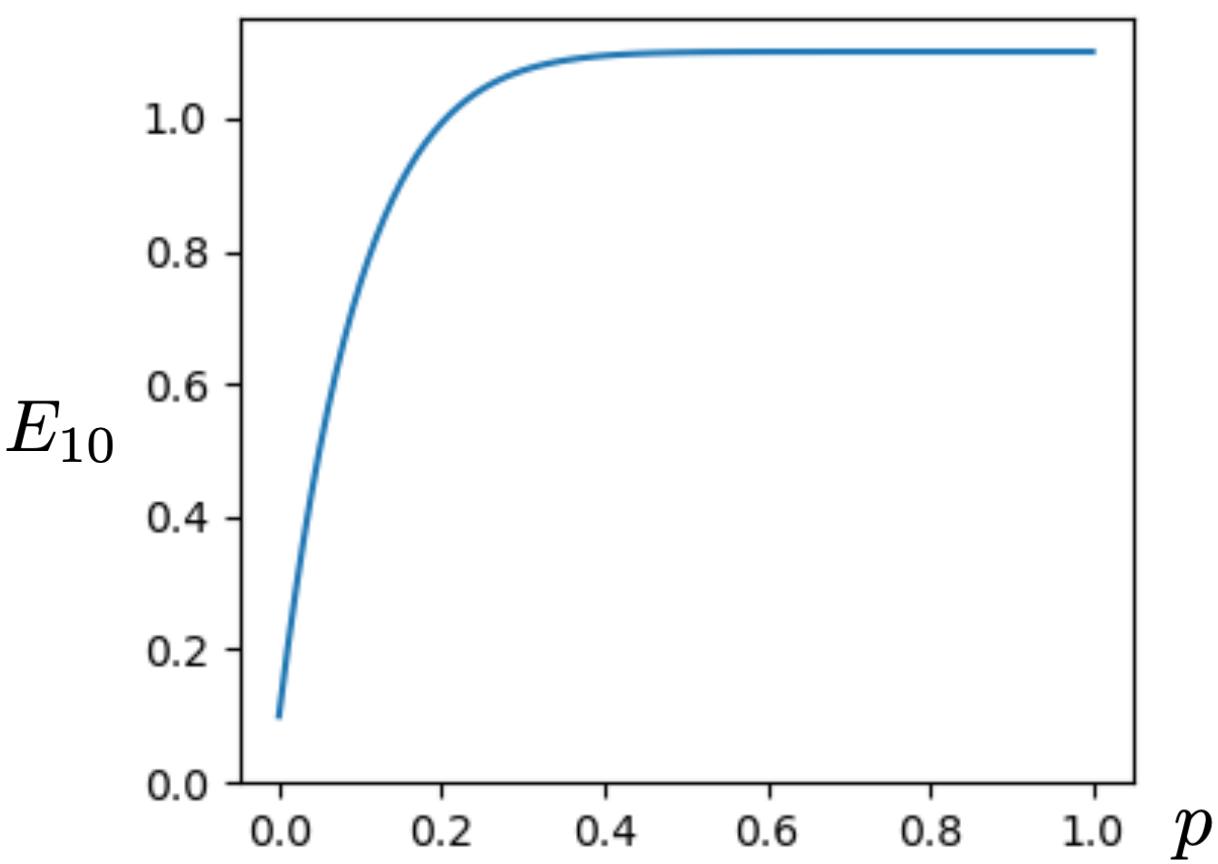
If the test result of a pool is negative, everyone in that pool is free of infection. If a pool tests positively, each individual in the pool is then tested.



Dorfman's method is useful when the population infection rate, or **prevalence**, is low. Consider the scenario that there is an infected person in every pool instead. How many tests are needed?

If k^2 samples are grouped into k pools with k samples each, with prevalence p, the expected number of tests per person is

$$E_k(p) = \frac{1}{k} + (1 - q^k), \quad q = 1 - p$$



Asymptomatic weekly positivity rate at PU less than 10% at the peak

At Princeton, pooling was used starting Feb 2021

THE DAILY PRINCETONIAN

MASTHEAD | ADVERTISE | CONTACT US NEWS OPINION SPORTS FEATURES THE PROSPECT MULTIMEDIA PODCAST CARTOON DATA PUZZLES HUMOR PROJE(

Try our latest crossword

Follow us on Instagram

Try our

NEWS

Princeton’s COVID-19 lab approved to begin “pooling” samples

Evelyn Doskoch

February 16, 2021 | 7:24pm EST

At Princeton, pooling was used starting Feb 2021

THE DAILY PRINCETONIAN

MASTHEAD | ADVERTISE | CONTACT US NEWS OPINION SPORTS FEATURES THE PROSPECT MULTIMEDIA PODCAST CARTOON DATA PUZZLES HUMOR PROJE

Try our latest crossword Follow us on Instagram Try our

NEWS

Princeton's COVID-19 lab approved to begin "pooling" samples

Evelyn Doskoch

February 16, 2021 | 7:24pm EST

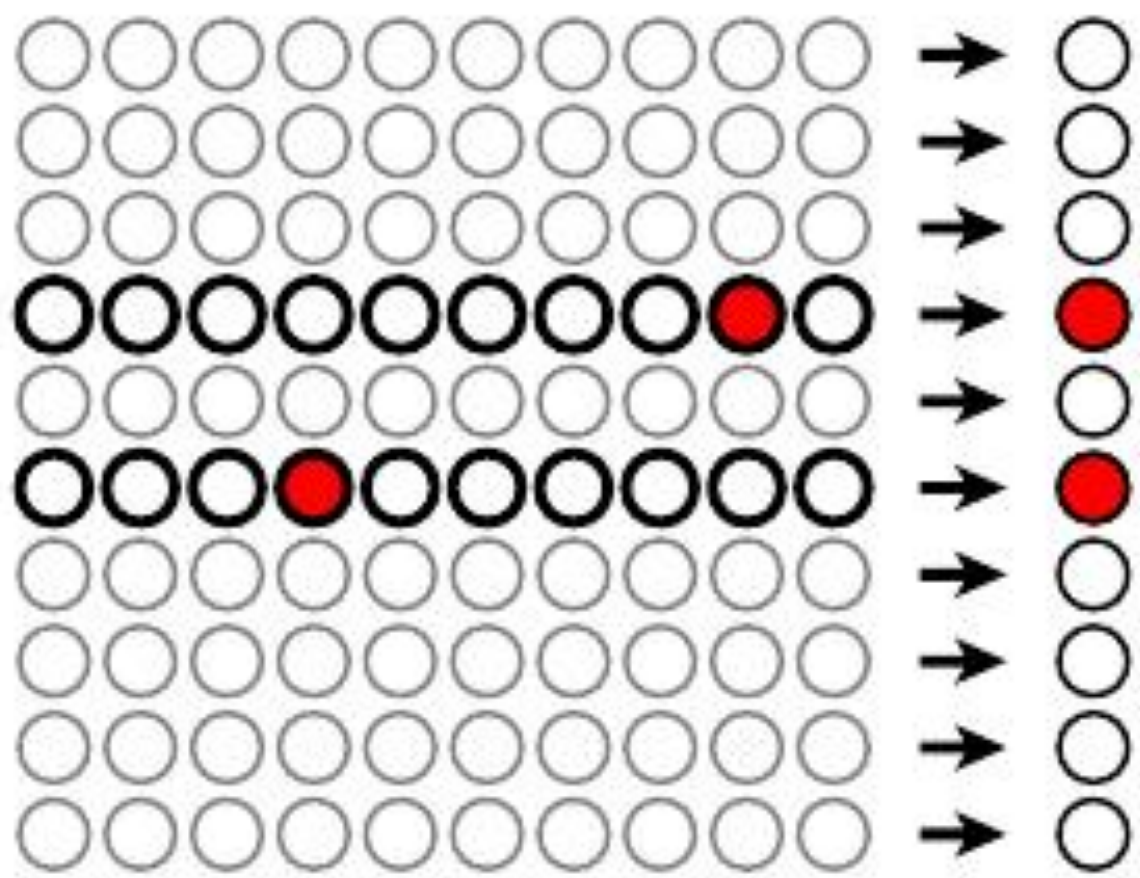
After some experiments, settled on 4 individual samples combined into one pool, and a positive result of the pool triggered individual tests of the 4 samples.

Saves work when there are less than ~10% positive cases

Non-perfect tests

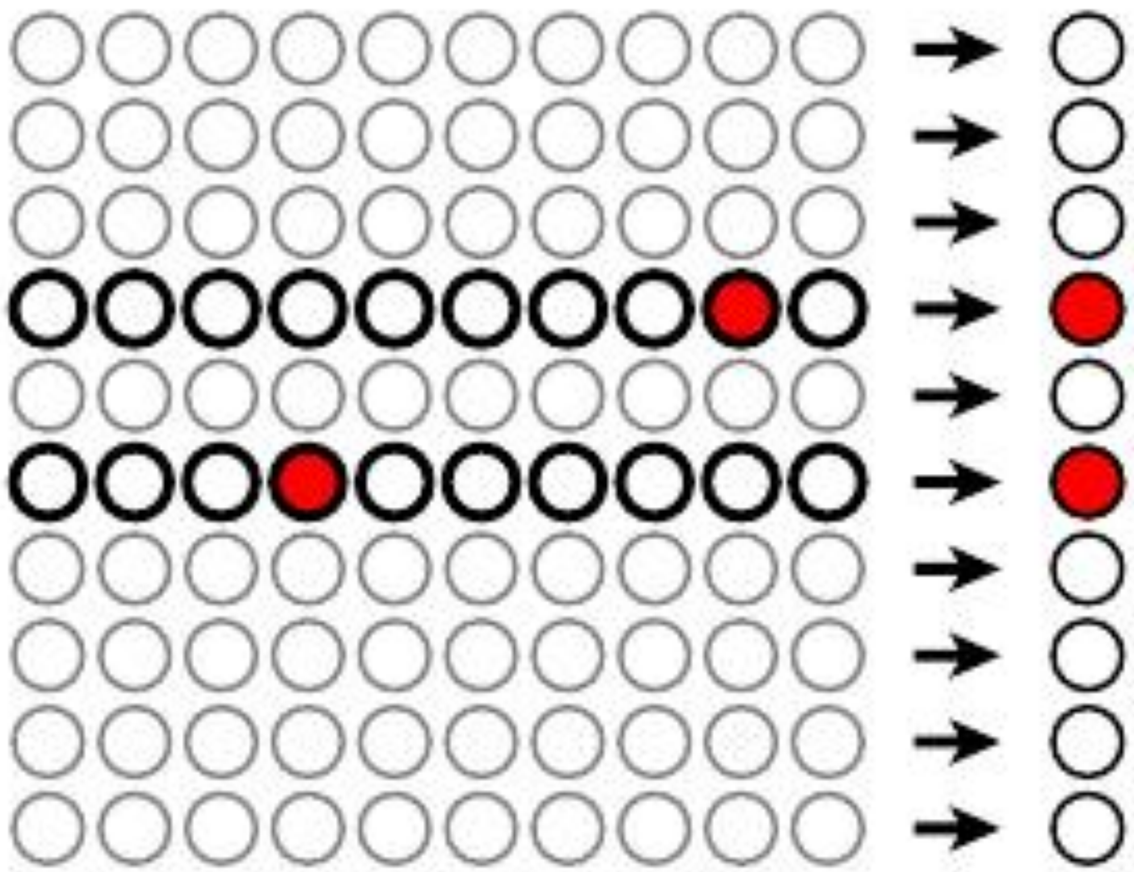
When the test has sensitivity S_e and specificity S_p, the expected number of tests per person is

$$E_k(p) = \frac{1}{k} + S_e(1 - q^k) + (1 - S_p)q^k$$

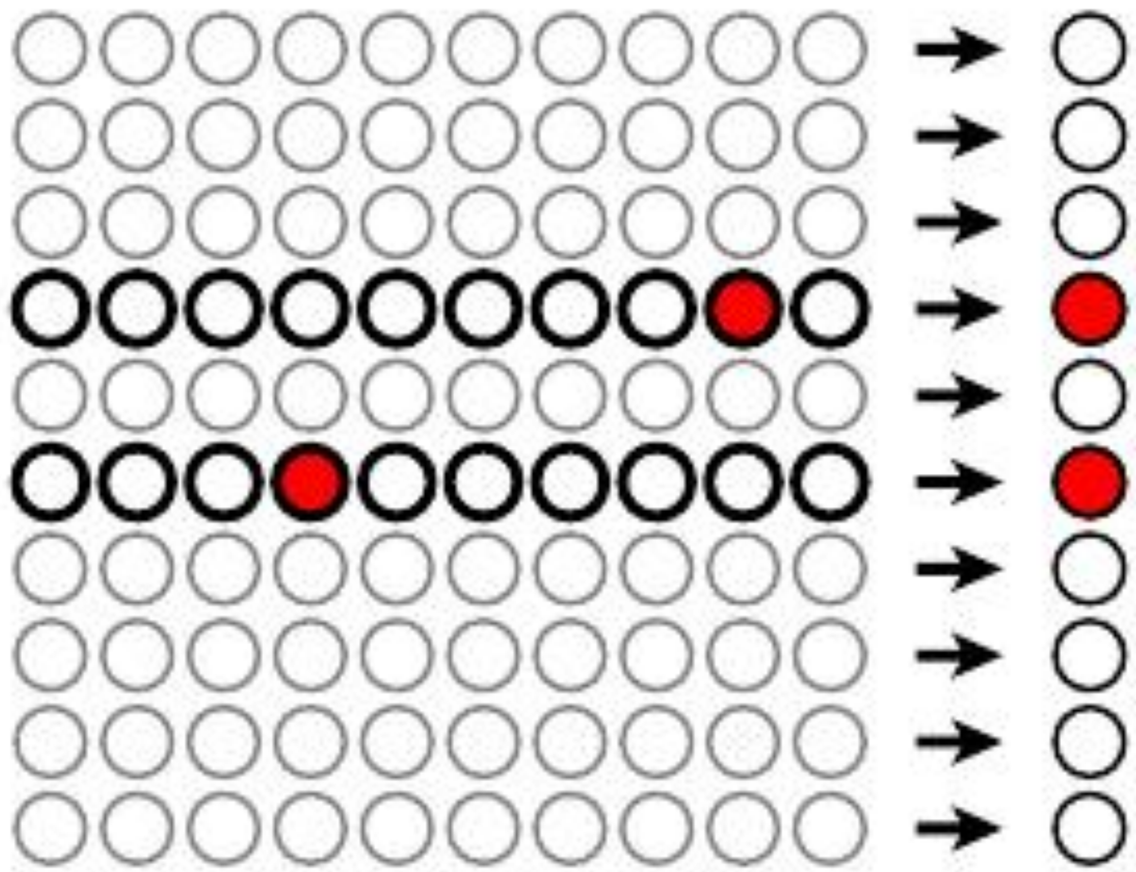


Sensitivity = probability that test shows positive given that you are infected
Specificity = probability that test shows negative given that you are not infected

What happens to sensitivity and specificity for each individual?

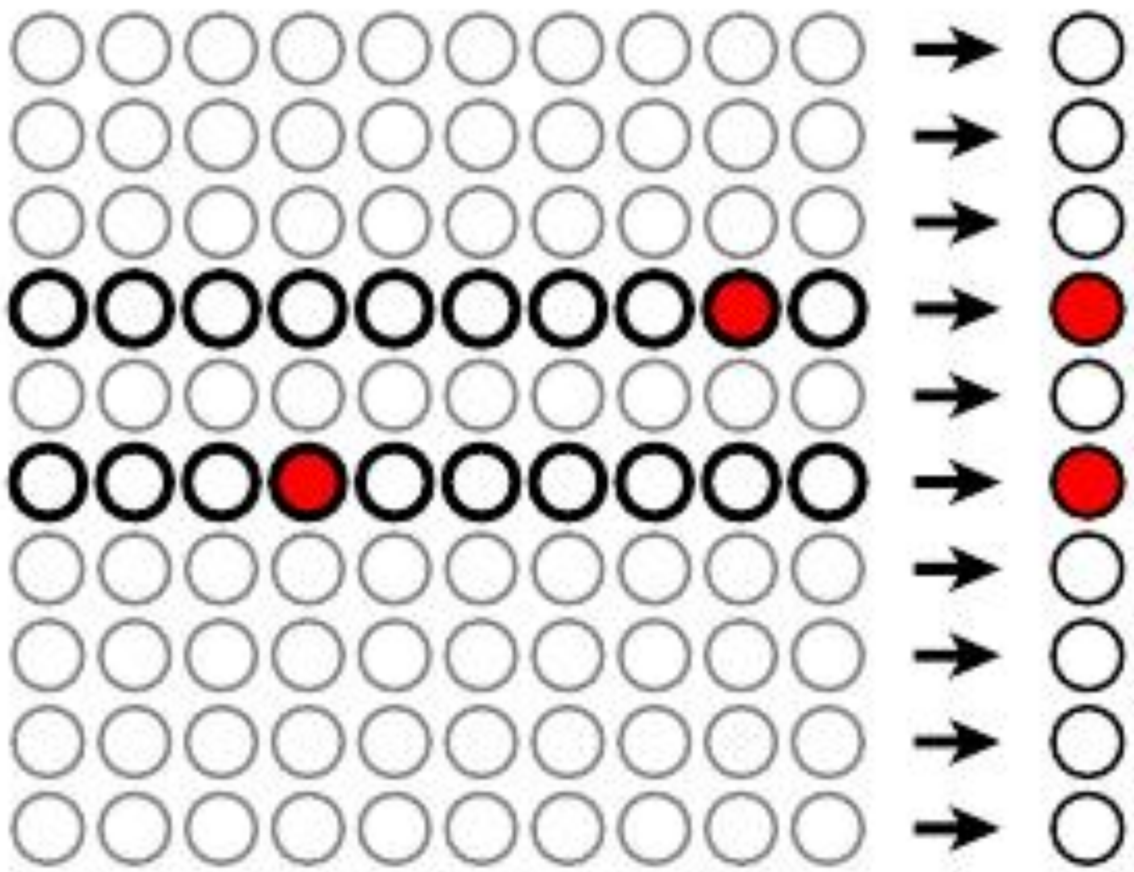


What happens to sensitivity and specificity for each individual?



An infected person needs two tests (pool and subsequent individual test) to be positive to be correctly shown as positive.

What happens to sensitivity and specificity for each individual?



An infected person needs two tests (pool and subsequent individual test) to be positive to be correctly shown as positive.
Expect sensitivity to decrease, but specificity to increase!

The sensitivity and specificity of Dorfman's method is

$$S_e' = S_e^2$$

and

$$S_p' = 1 - (1 - S_p)P_{k-1},\ \text{where}$$
$$P_{k-1} = S_e(1 - q^{k-1}) + (1 - S_p)q^{k-1}$$

For $S_e = S_p = 0.95,\ p = 0.01$, we have

$S_e' = 0.90,\ S_p' = 0.99$

Repeated testing

Repeated testing

Recall what we did for a single test: Consider one COVID test with 84% sensitivity and 99% specificity. (I = infected), and assume 10% of the population is infected

P(positive | I) = 0.84	True positive
P(negative | I) = 0.16	False negative
P(negative | nI) = 0.99	True negative
P(positive | nI) = 0.01	False positive

We used Bayes' rule to get the formula:

$$P(I|\text{positive}) = \frac{P(\text{positive}|I)P(I)}{P(\text{positive}|I)P(I) + P(\text{positive}|nI)P(nI)}$$

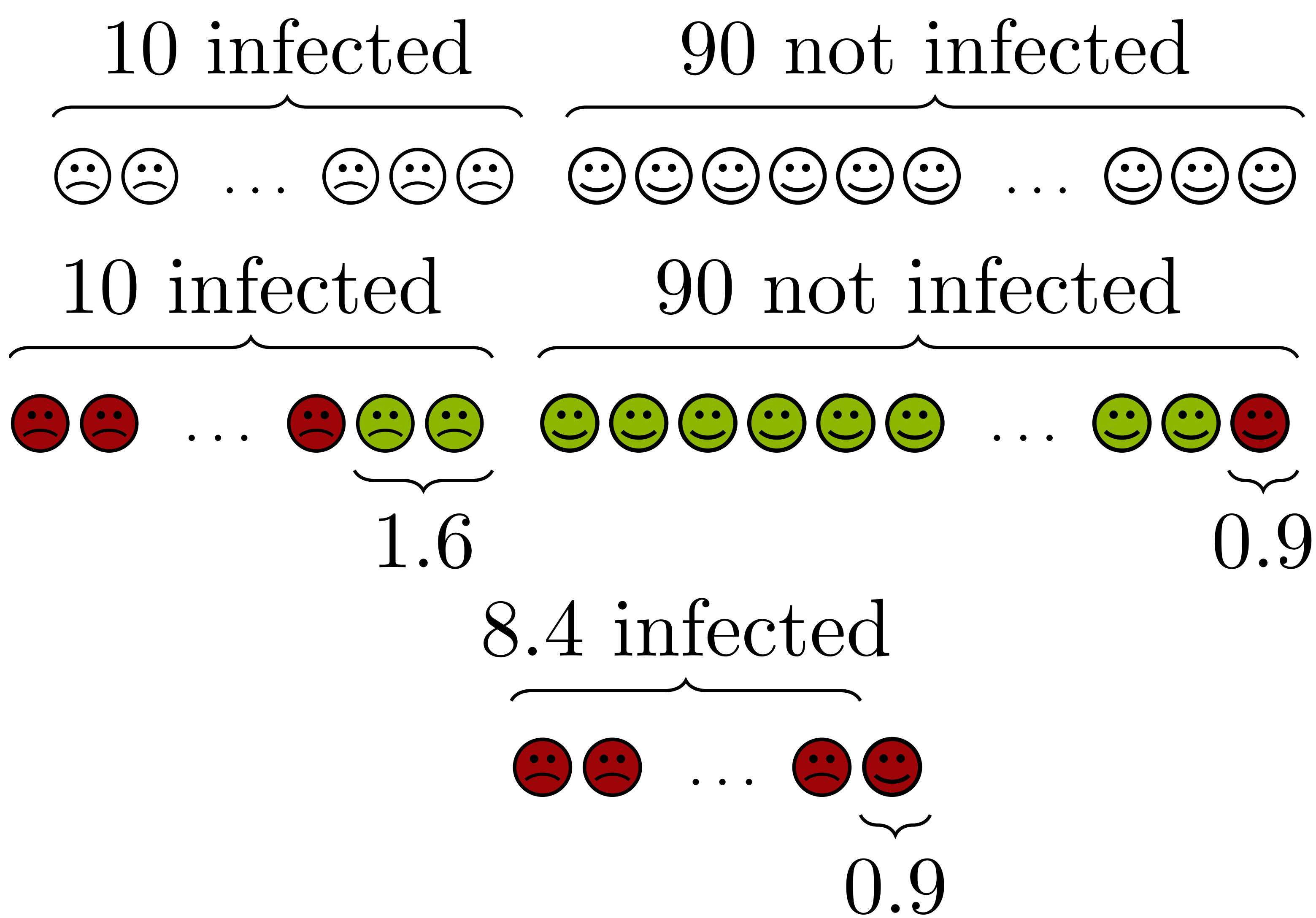
Example 1:

P(test positive | I) = 0.84 — True positive

P(test negative | I) = 0.16 — False negative

P(test negative | nI) = 0.99 — True negative

P(test positive | nI) = 0.01 — False positive



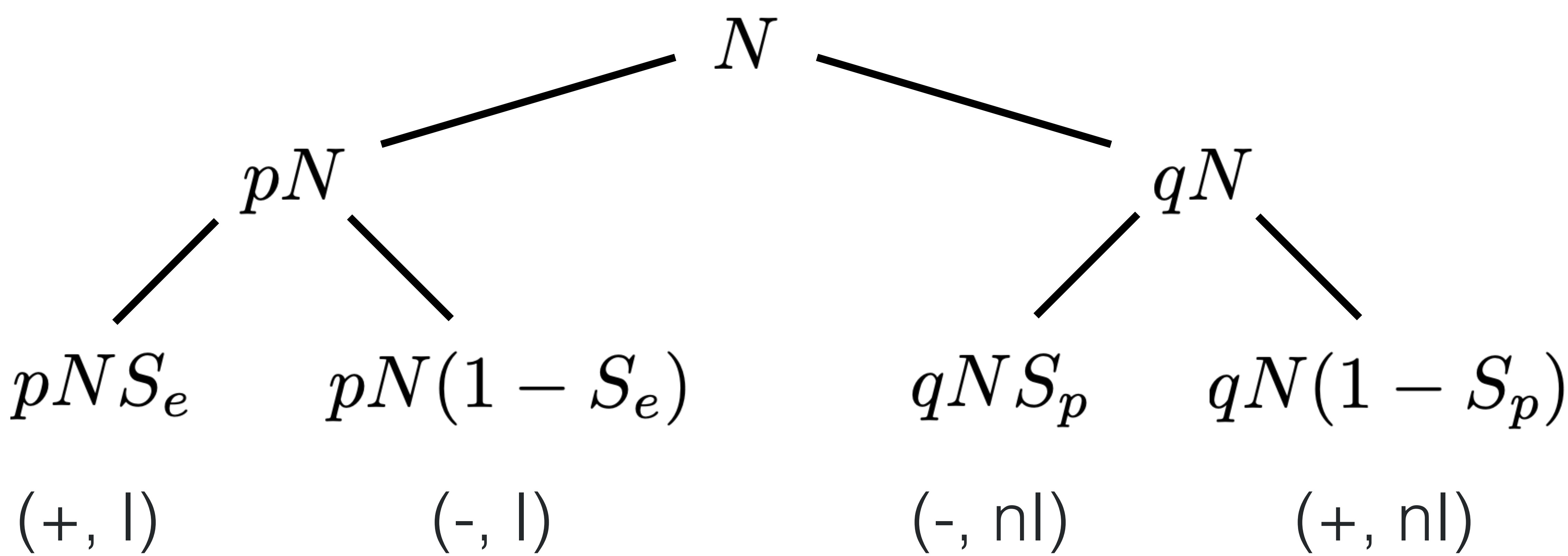
Repeated testing

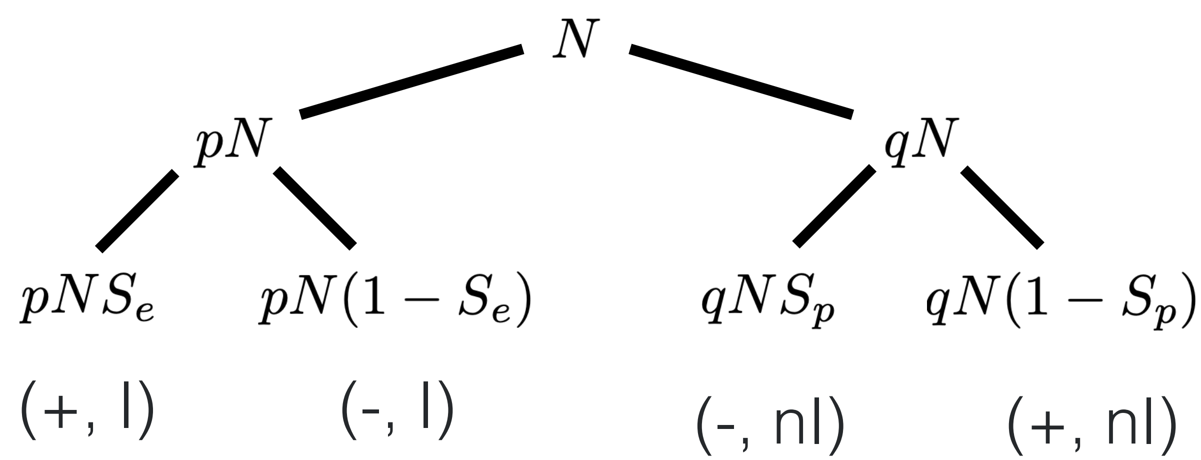
Graphical illustration of the math:

S_e : sensitivity, S_p : specificity, N : population size

p : prevalence, $q = 1 - p$,

After the first test,

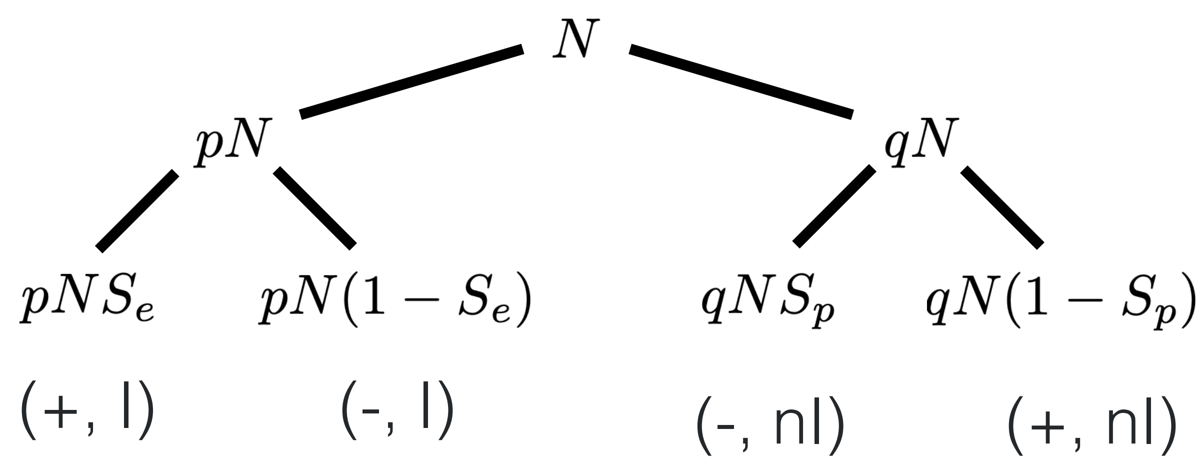




$$N_+ = pNS_e + qN(1-S_p)$$

$$P(I|+) = \frac{pNS_e}{pNS_e + qN(1-S_p)} = \frac{pS_e}{pS_e + q(1-S_p)}$$

$$P(nI|+) = \frac{q(1-S_p)}{pS_e + q(1-S_p)}$$



$$N_+ = pNS_e + qN(1-S_p)$$

$$P(I|+) = \frac{pNS_e}{pNS_e + qN(1-S_p)} = \frac{pS_e}{pS_e + q(1-S_p)}$$

$$P(nI|+) = \frac{q(1-S_p)}{pS_e + q(1-S_p)}$$

For $S_e = 0.9,\ S_p = 0.95,\ p = 0.05,$

$$P(I|+) \approx 0.49$$

When testing again we do the same math, but only look at the individuals who tested positive in the first test, so the probability to be infected is now $p(I|+)$

In other words, we replace:

$$N \to N' = N_+ \text{ and } p \to p' = p(I|+)$$

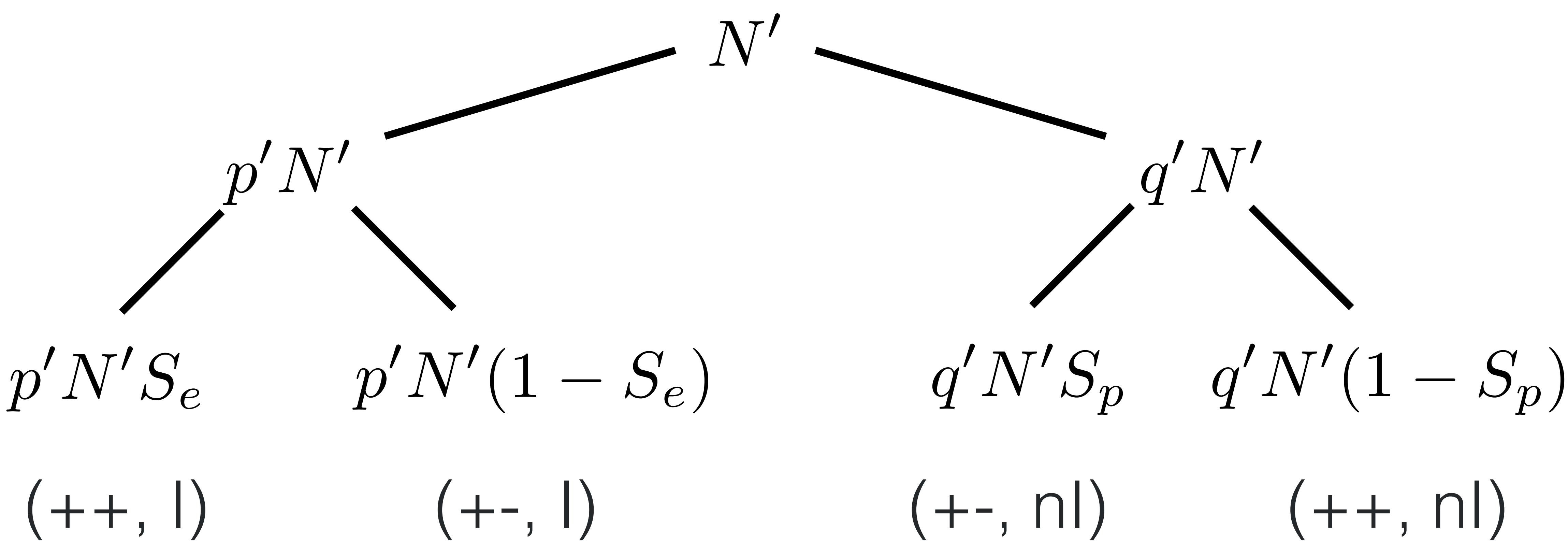
$$q \to q' = 1 - p'$$

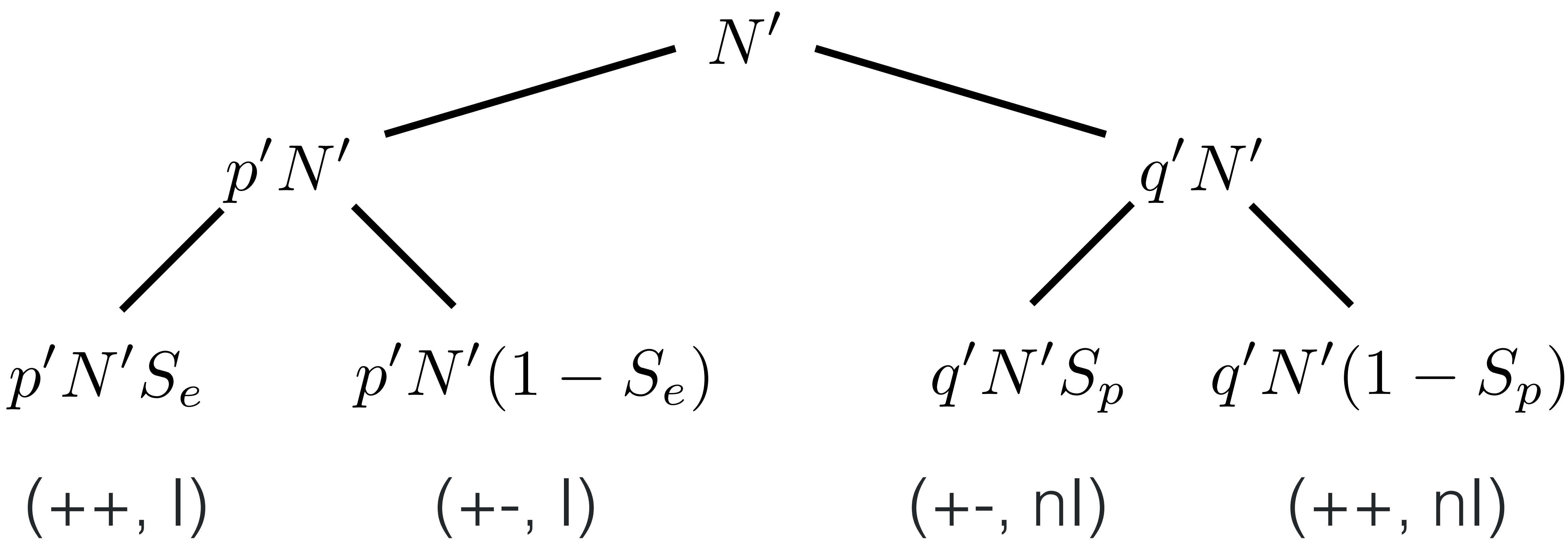
When testing again we do the same math, but only look at the individuals who tested positive in the first test, so the probability to be infected is now $p(I|+)$

In other words, we replace:

$$N \rightarrow N' = N_+ \text{ and } p \rightarrow p' = p(I|+)$$

$$q \rightarrow q' = 1 - p'$$





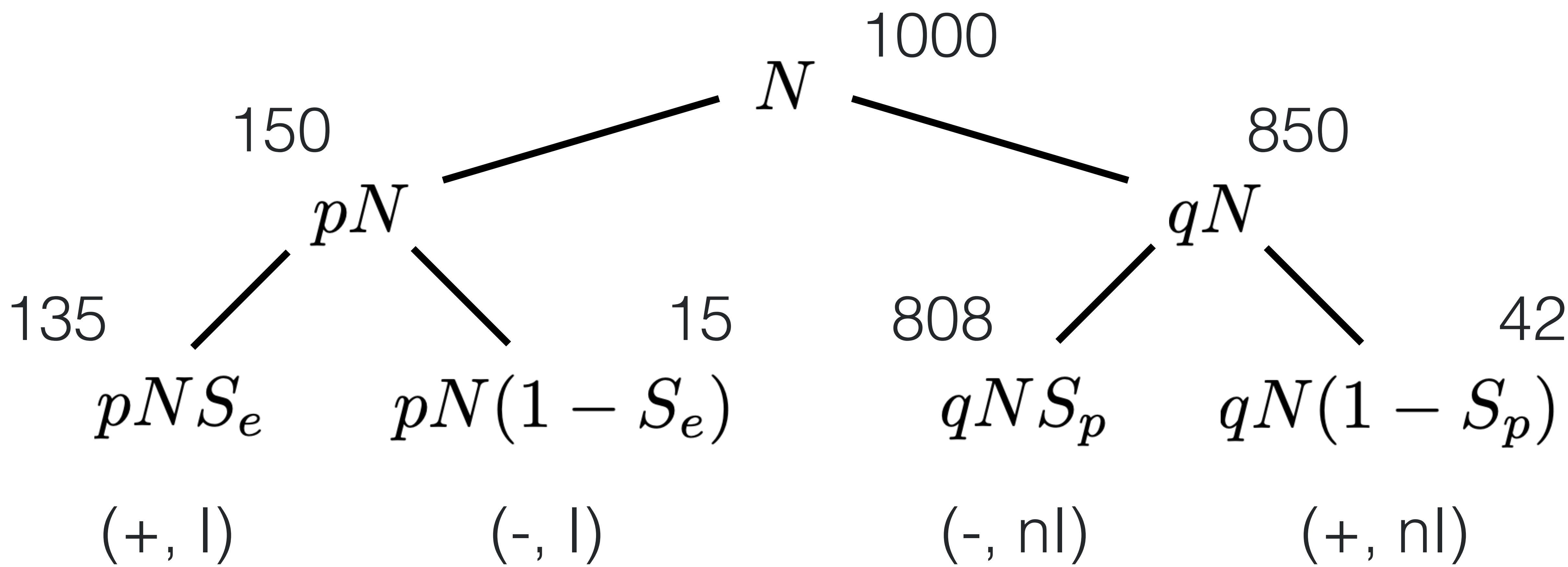
$$
\begin{aligned}
P(I|++) &= \frac{p'N'S_e}{p'N'S_e + q'N'(1-S_p)} = \frac{p'S_e}{p'S_e + q'(1-S_p)} \\
&= \frac{pS_e^2}{pS_e^2 + q(1-S_p)^2} \approx 0.94
\end{aligned}
$$

If we perform a third test, we test the samples whose first two test results are positive, and the same calculations give

$$p(I|+++) = \frac{pS_e^3}{pS_e^3 + q(1 - S_p)^3} \approx 0.997$$

Estimating prevalence

$$S_e = 0.9,\ \ S_p = 0.95,\ \ p = 0.15$$



test number	number of positives N_+	ratio r	N_+/r
1	177	0.9	197
2	124	0.81	152
3	109	0.729	150

If we neglect the number of false positives, the number of positives after k tests should be approximately S_e^k times the number of positive cases in the population, which can be used to estimate the total number of positives.

References

1. http://www.ams.org/publicoutreach/feature-column/fc-2020-10
2. http://www.ams.org/publicoutreach/feature-column/fc-2020-09