PROBABILITY REVIEW

BAYESIAN STATISTICS FOR ECOLOGISTS

IGB 12. TO 19. NOVEMBER 2018

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon (“model”)

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon ("model")
- I have some general knowledge about the question

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon ("model")
- I have some general ("prior") knowledge about the question

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon ("model")
- I have some general ("prior") knowledge about the question
- I gather additional knowledge

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon (“model”)
- I have some general (“prior”) knowledge about the question
- I gather additional knowledge (“data”)

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon (“model”)
- I have some general (“prior”) knowledge about the question
- I gather additional knowledge (“data”)

What is the probability that my **model is correct** given what I **already know about it** and **what I’ve learned**?

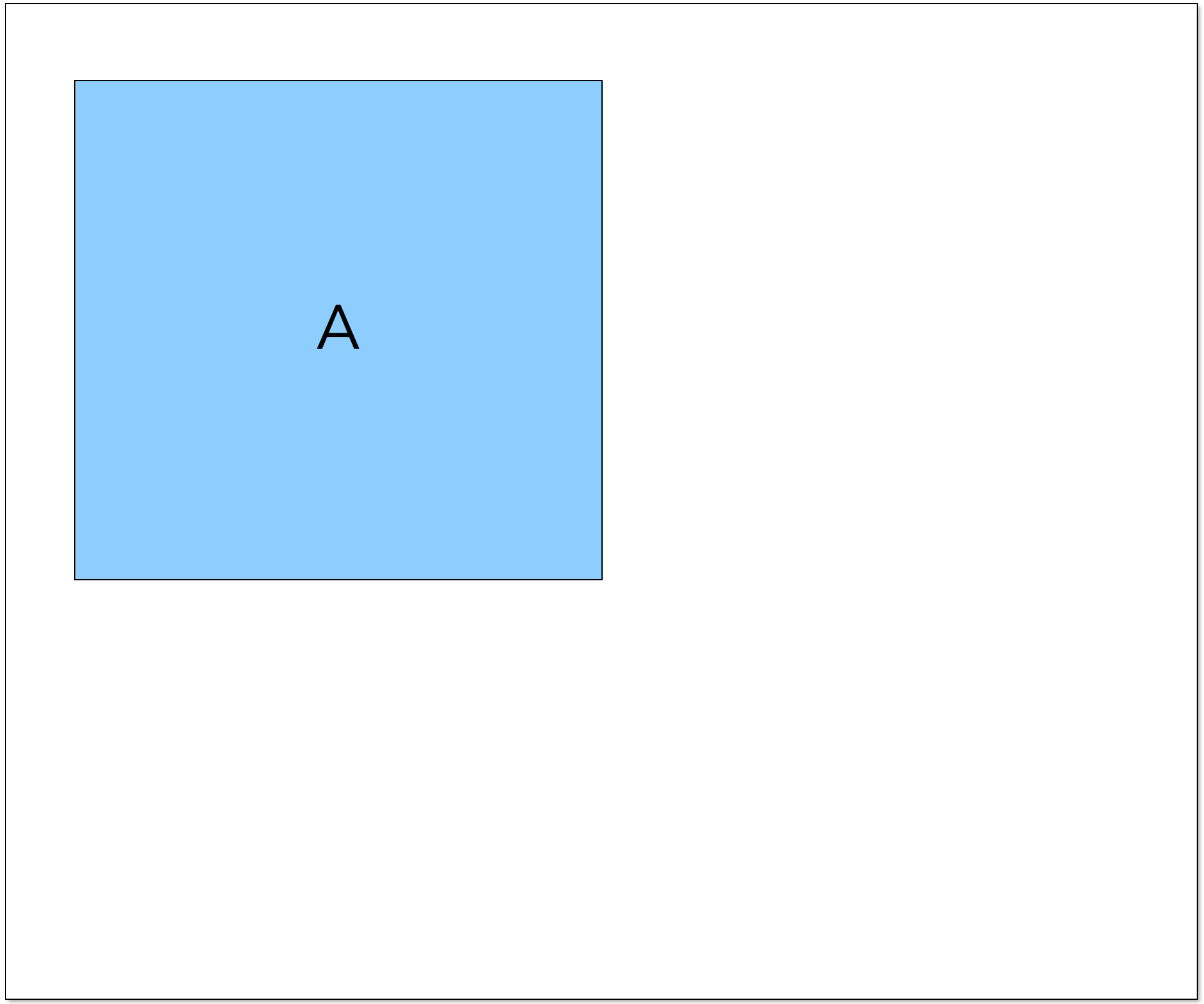
STATISTICAL INDEPENDENCE

- We wish to make a statistical conclusion from some vector of observations $\mathbf{y}_{1..i}$
- We assume that the y_i's are **i.i.d.**
- What does this mean?

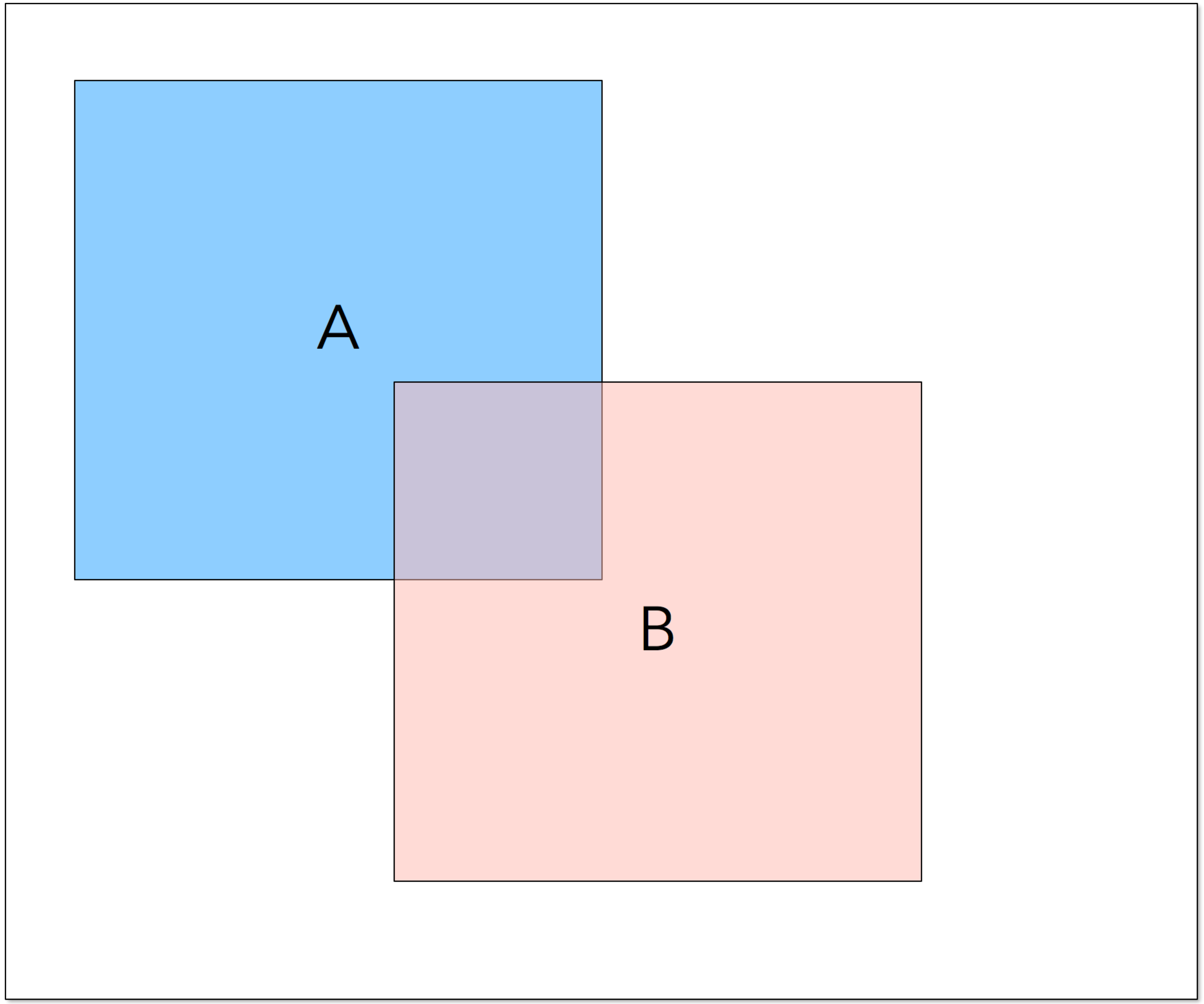
STATISTICAL INDEPENDENCE

- Two events **A** and **B** are independent if the probability of **A** occurring does not depend on whether **B** has occurred, and vice-versa

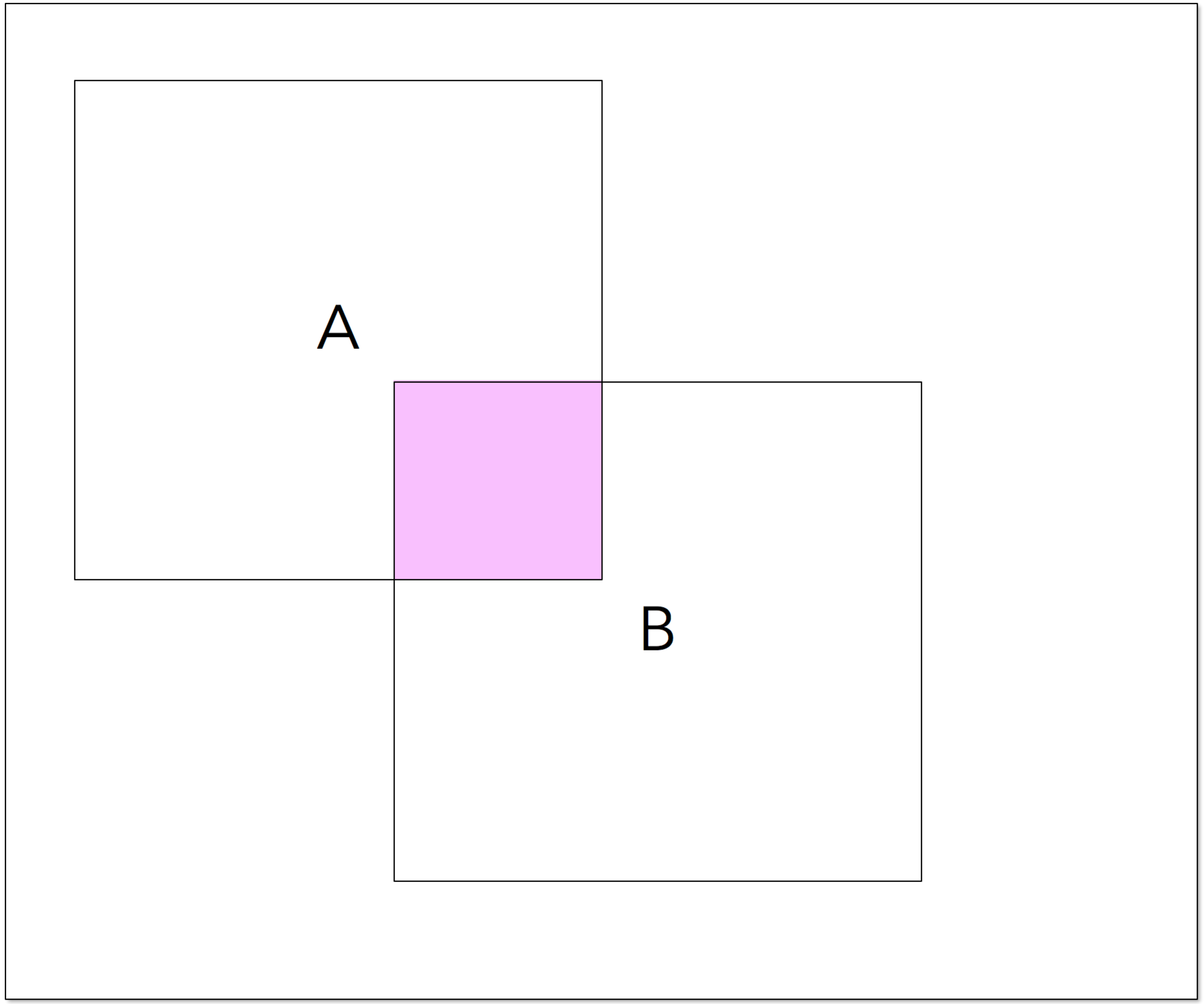
Total sample space



pr(A) – marginal probability of A



pr(A) + pr(B) - the probability of either one OR both



A ∩ B (A and B) – also written pr(A,B)

STATISTICAL INDEPENDENCE

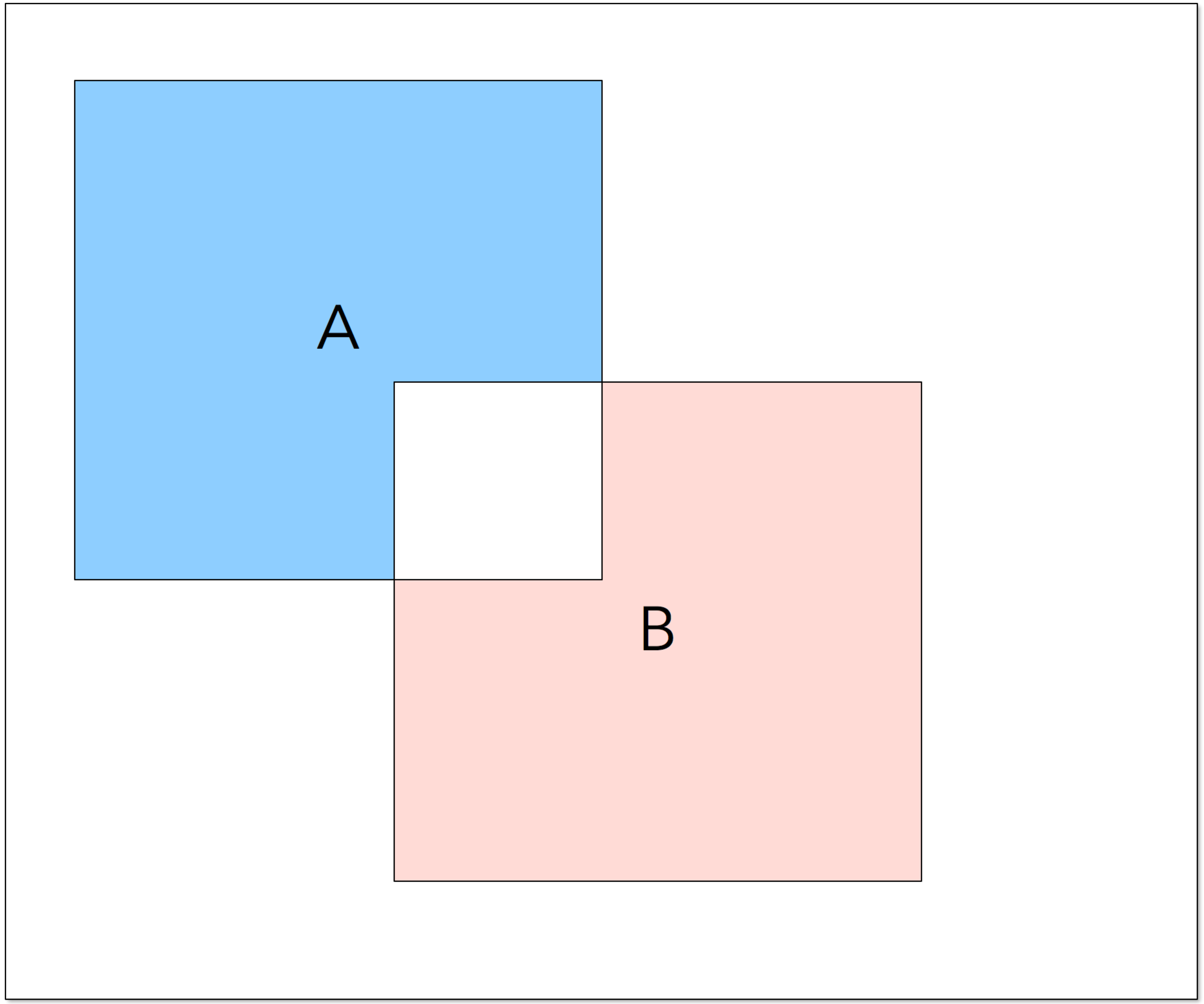
- Two events **A** and **B** are independent if the probability of **A** occurring does not depend on whether **B** has occurred, and vice-versa
- If **A** and **B** are independent:
  - pr(A,B) = pr(A)pr(B)

STATISTICAL INDEPENDENCE

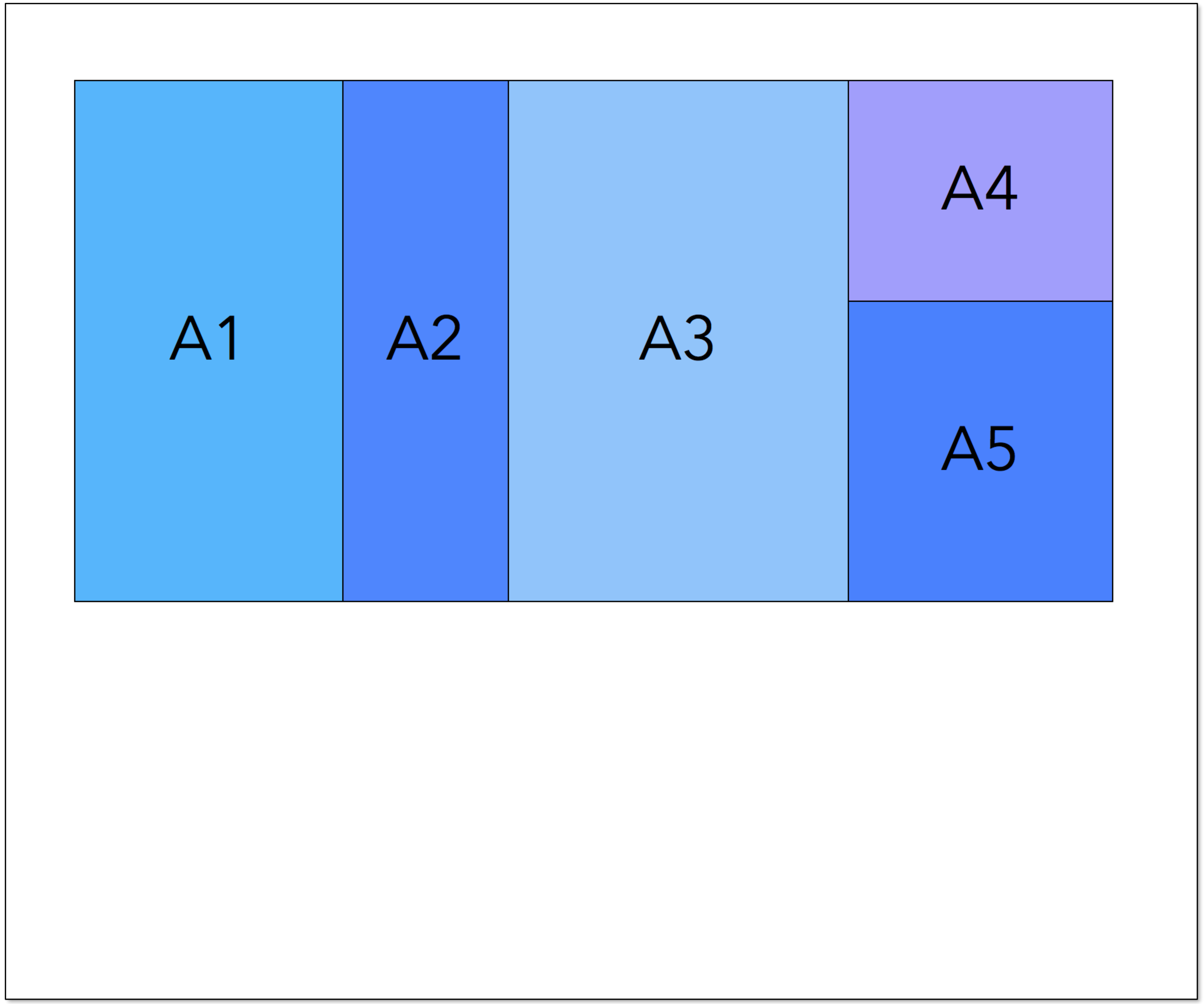
- Two events **A** and **B** are independent if the probability of **A** occurring does not depend on whether **B** has occurred, and vice-versa
- If **A** and **B** are independent:
    - pr(A,B) = pr(A)pr(B)
- If a vector **y** is i.id., then all y's are independent, and

$$y_i \sim \mathbb{D}(.)$$

- y is a random variable drawn from a statistical distribution (more later)

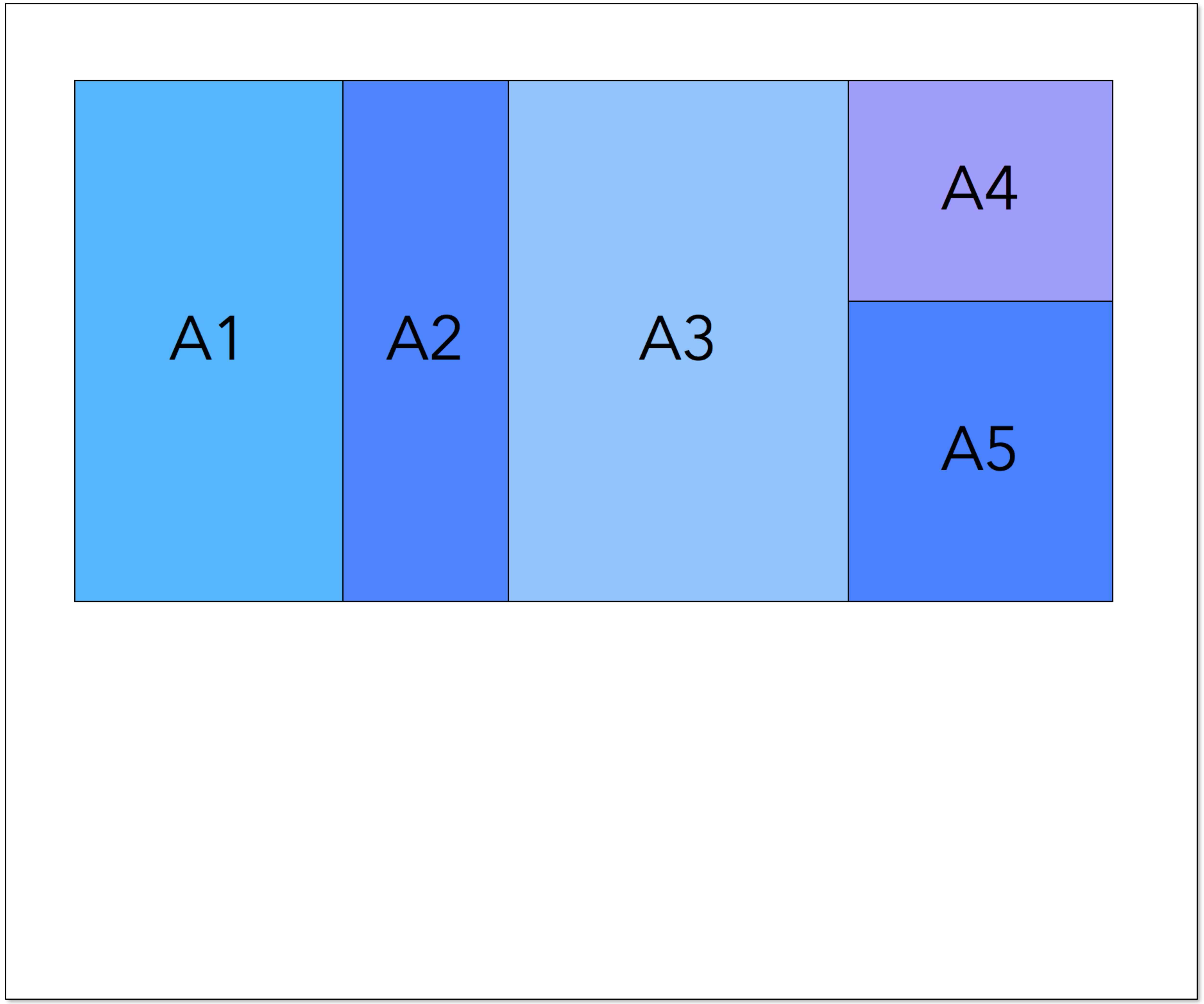


$A \cup B$ (A or B) = pr(A) + pr(B) - pr(A,B)



Partitioning:

$$pr(A) = pr(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5) = pr(A_1) + pr(A_2) + pr(A_3) + pr(A_4) + pr(A_5)$$



If the **A** encompasses all possibilities, then pr(A) = 1

CONDITIONAL PROBABILITY

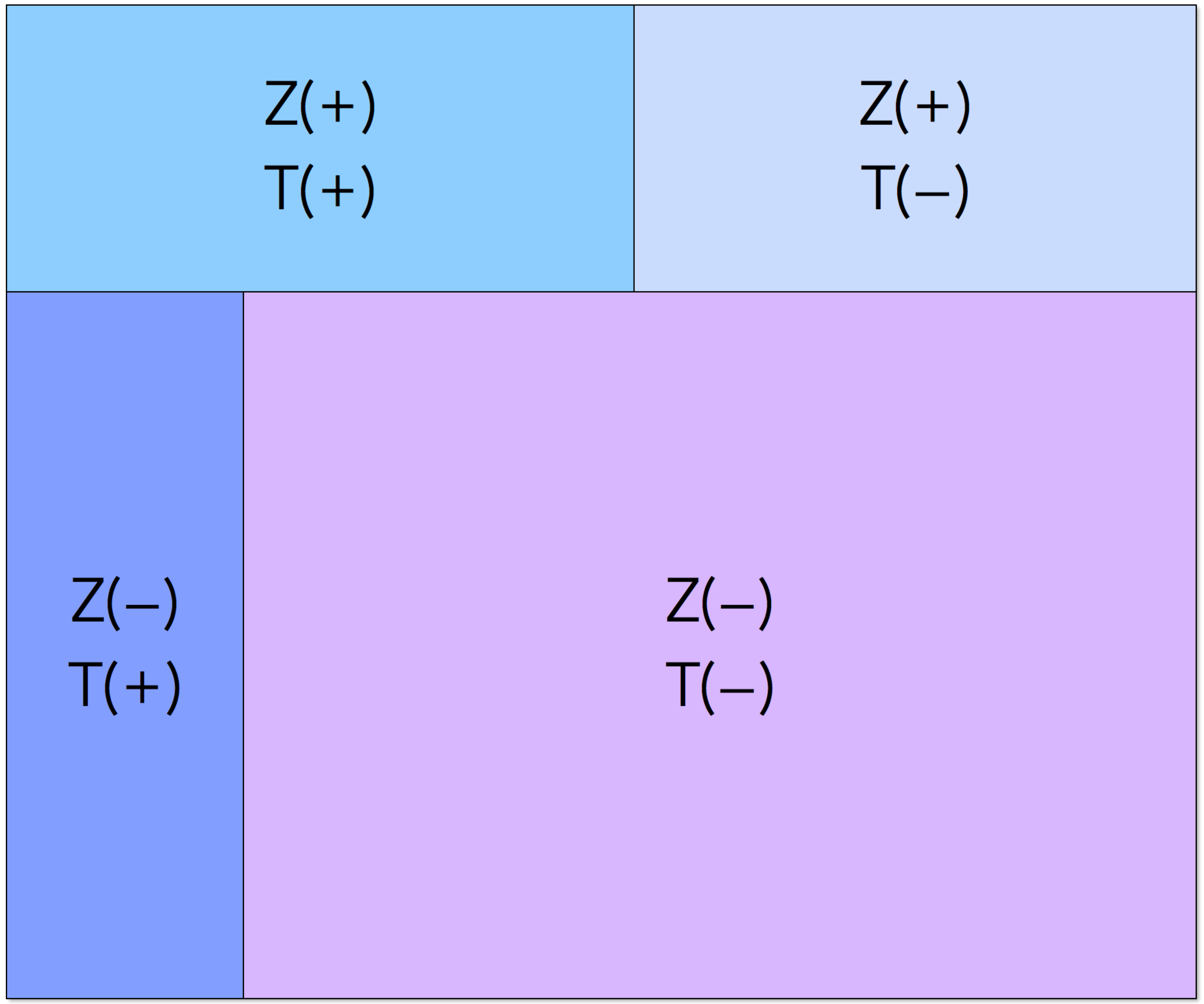
- pr(**A** | **B**) is the probability that A occurs if B has already occurred.
- The **joint** probability of two events is equal to the product of their **conditional** probability and the **marginal** probability of the conditioning event (product rule)
    - pr(**A**,**B**) = pr(**A** | **B**)pr(**B**)
- If **A** and **B** are independent, then pr(**A** | **B**) = pr(**A**)

CONDITIONAL PROBABILITY PRACTICE

- Will you become a zombie?
  - Assume 0.1% of people are infected but don't know it yet
  - We have a test that has a 0.5% false negative rate (test is negative when you are infected) and a 1% false positive rate (test is positive when you are not a zombie)
- You take the test, and the result is positive. What is the probability that you are actually a zombie?

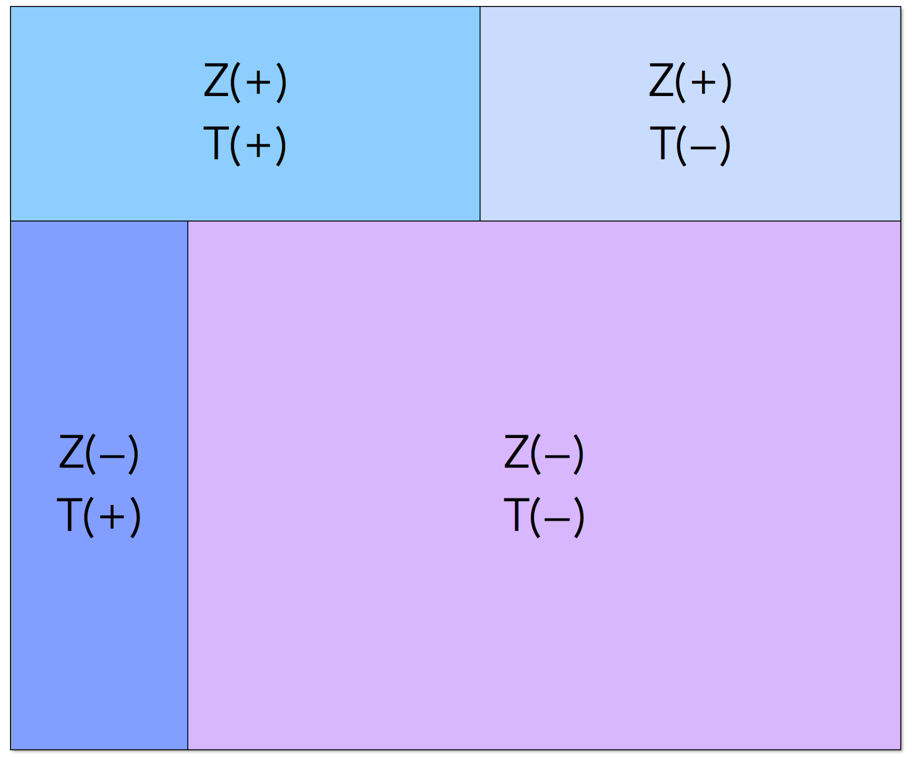
ZOMBIE?

- 0.1% infected
- 1% false positive: Z(-)T(+)
- 0.5% false negative: Z(+)T(-)



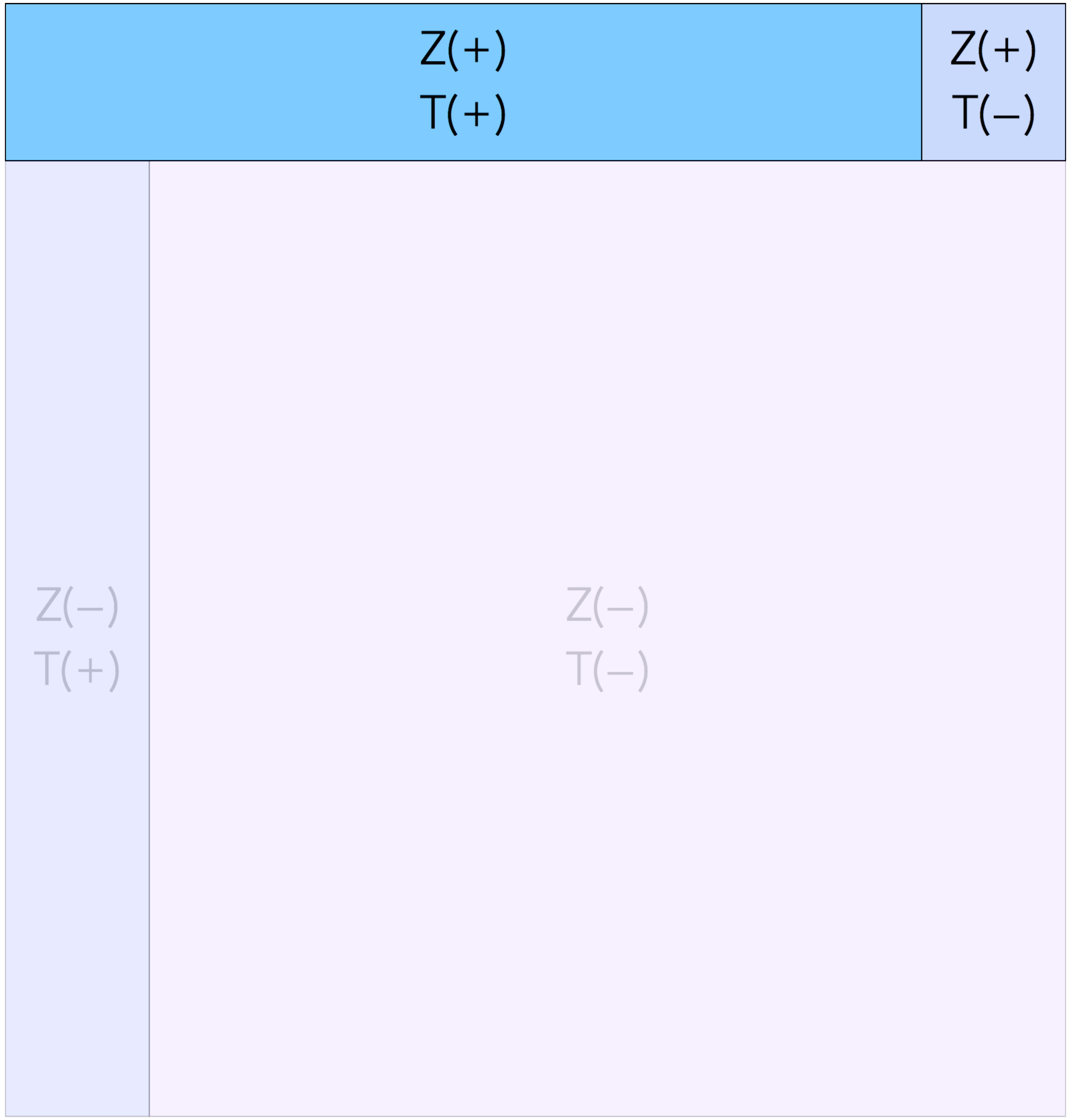
ZOMBIE?

- We want the probability of zombiness given a positive test:
    - pr(**Z** | **T**)
- pr(**Z**) = 0.001
- pr(**Z**,!**T**) = pr(!**T** | **Z**)pr(**Z**) = 0.005 * 0.001 = 0.000005
- pr(**Z**,**T**) = pr(**Z**) - pr(**Z**,!**T**) = 0.001 - 0.000005 = 0.000995
- pr(!**Z**,**T**) = [1 - pr(**Z**)][pr(**T** | !**Z**)] = 0.999 * 0.01 = 0.00999
- pr(!**Z**,!**T**) = 1 - pr(!**Z**,**T**) = 0.99001



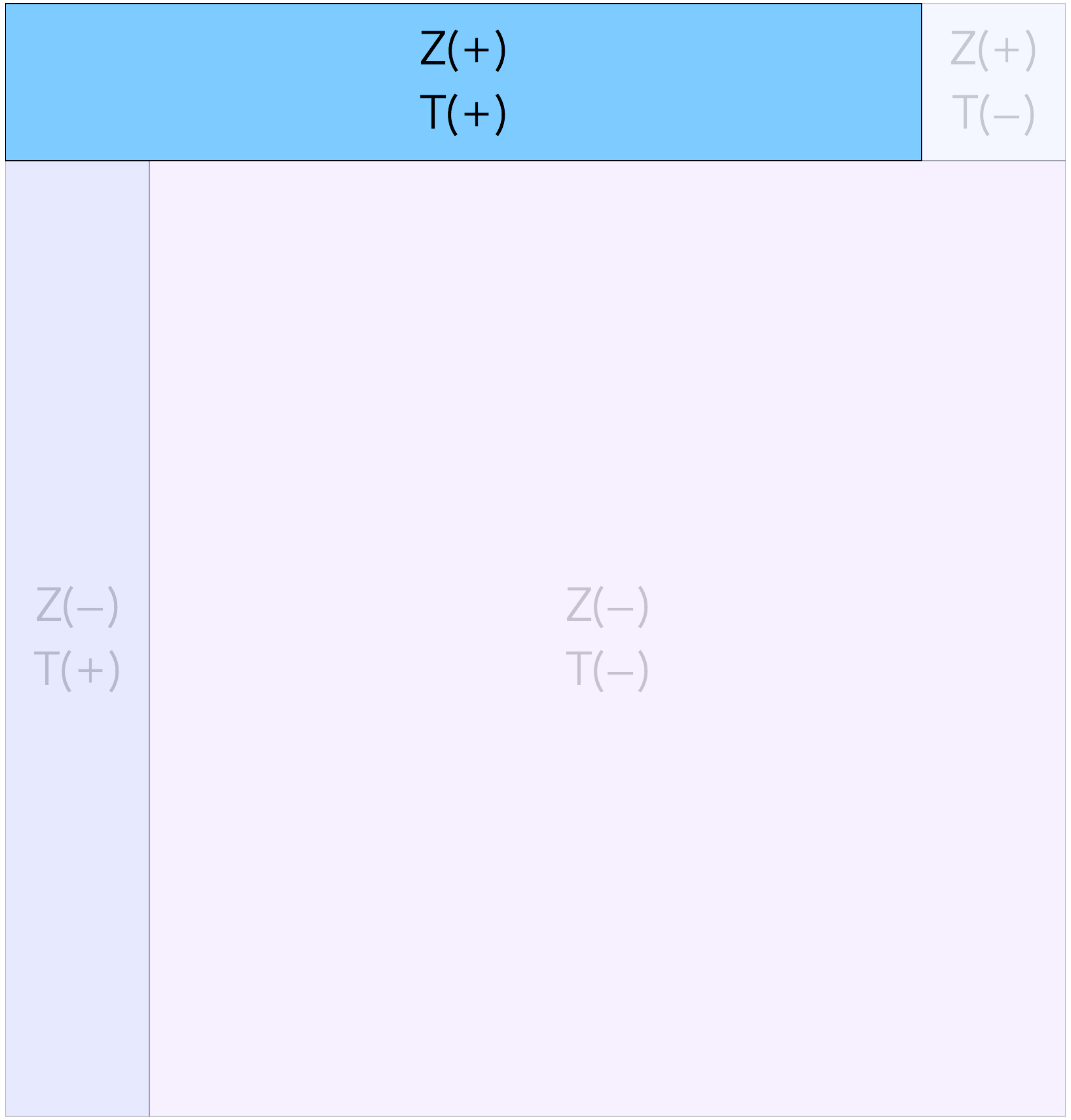
YOU ARE PROBABLY NOT A ZOMBIE

- In a population of 1,000,000 people:
  - 1000 are zombies…



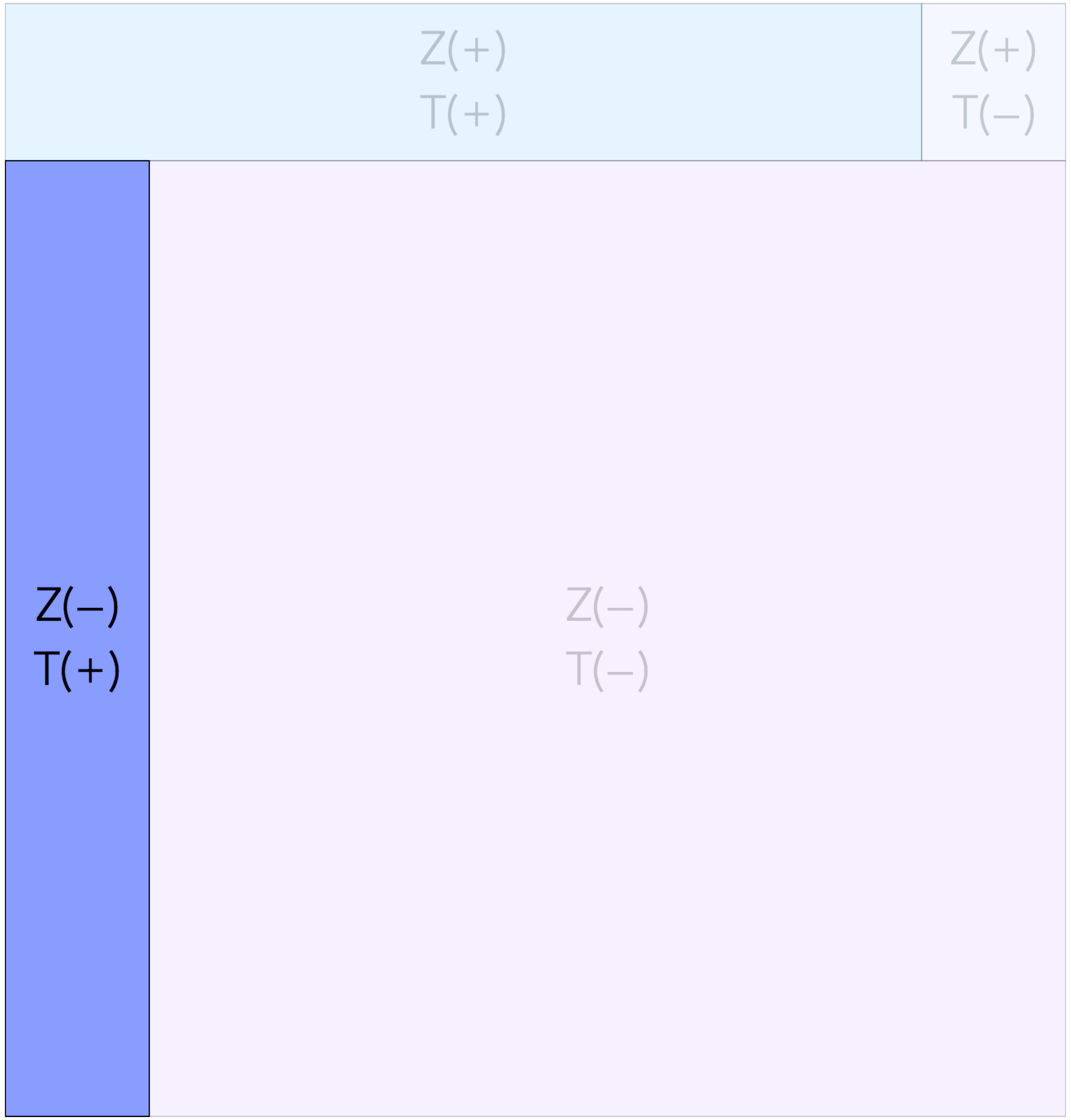
YOU ARE PROBABLY NOT A ZOMBIE

- In a population of 1,000,000 people:
    - 1000 are zombies…
    - of which 995 will test positive



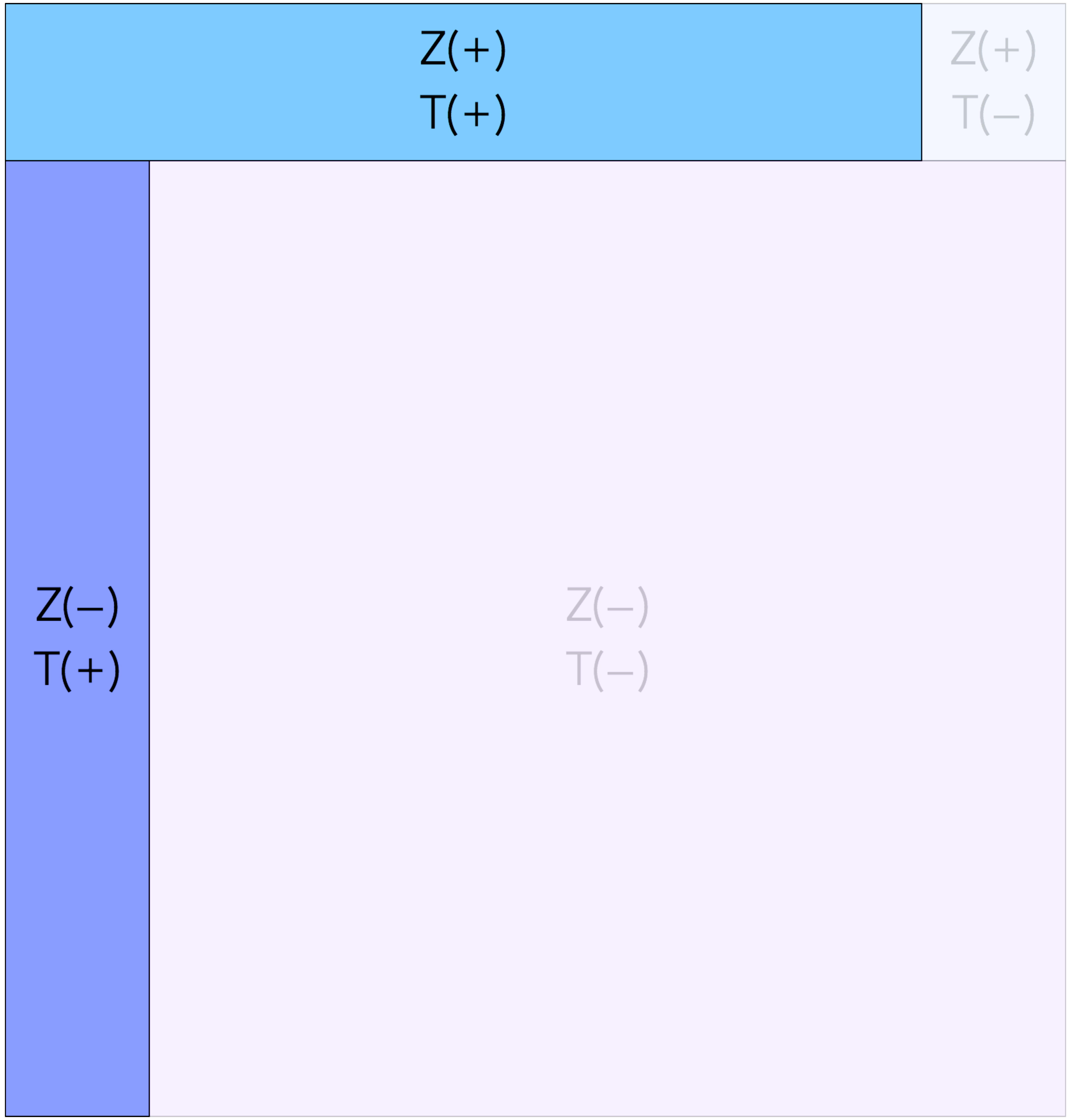
YOU ARE PROBABLY NOT A ZOMBIE

- In a population of 1,000,000 people:
  - 1,000 are zombies…
  - of which 995 will test positive
  - 999,000 are not zombies, of which 9,990 will test positive



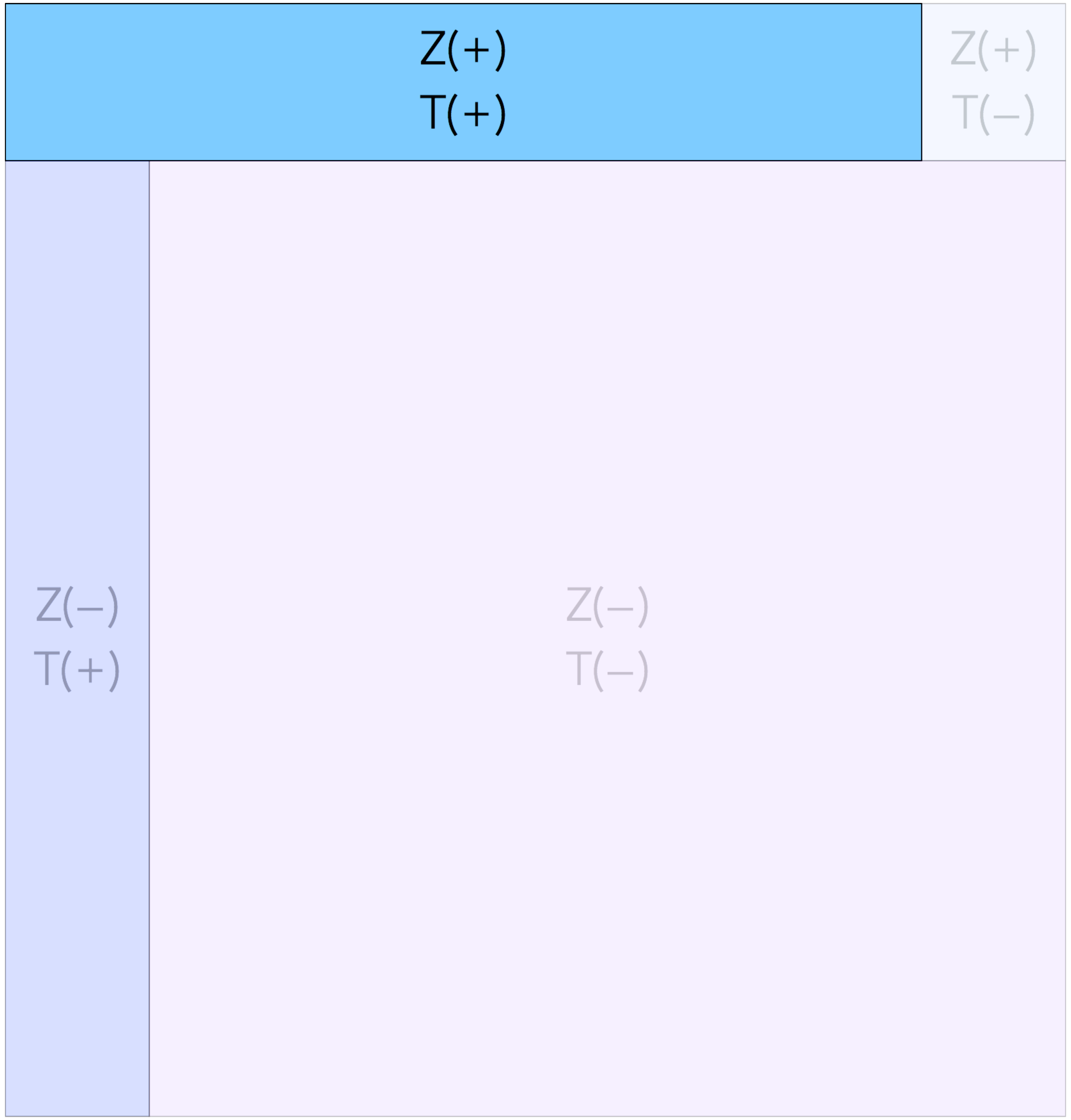
YOU ARE PROBABLY NOT A ZOMBIE

- In a population of 1,000,000 people:
    - 1,000 are zombies…
    - of which 995 will test positive
    - 999,000 are not zombies, of which 9,990 will test positive
    - 10,985 will test positive overall



YOU ARE PROBABLY NOT A ZOMBIE

- In a population of 1,000,000 people:
  - 1,000 are zombies…
  - of which 995 will test positive
  - 999,000 are not zombies, of which 9,990 will test positive
  - 10,985 will test positive overall
  - of which 995 are zombies: ~9.1% chance of being a zombie



ANOTHER WAY TO THINK ABOUT THIS

$$pr(Z|T) = ?$$

$$pr(Z,T) = pr(Z|T)pr(T) = pr(T|Z)pr(Z)$$

$$pr(Z|T) = \frac{pr(Z,T)}{pr(T)}$$

$$pr(Z|T) = \frac{pr(T|Z)pr(Z)}{pr(T,Z) + pr(T,!Z)}$$

ANOTHER WAY TO THINK ABOUT THIS

$$pr(Z|T) = \frac{pr(T|Z)pr(Z)}{pr(T,Z) + pr(T,!Z)}$$

$$pr(Z|T) = \frac{pr(T|Z)pr(Z)}{pr(T|Z)pr(Z) + pr(T,!Z)(1 - pr(Z))}$$

$$pr(Z|T) = \frac{(1-0.005)\times 0.001}{0.995\times 0.001+0.01\times 0.999} = 0.091$$

CHAIN RULE

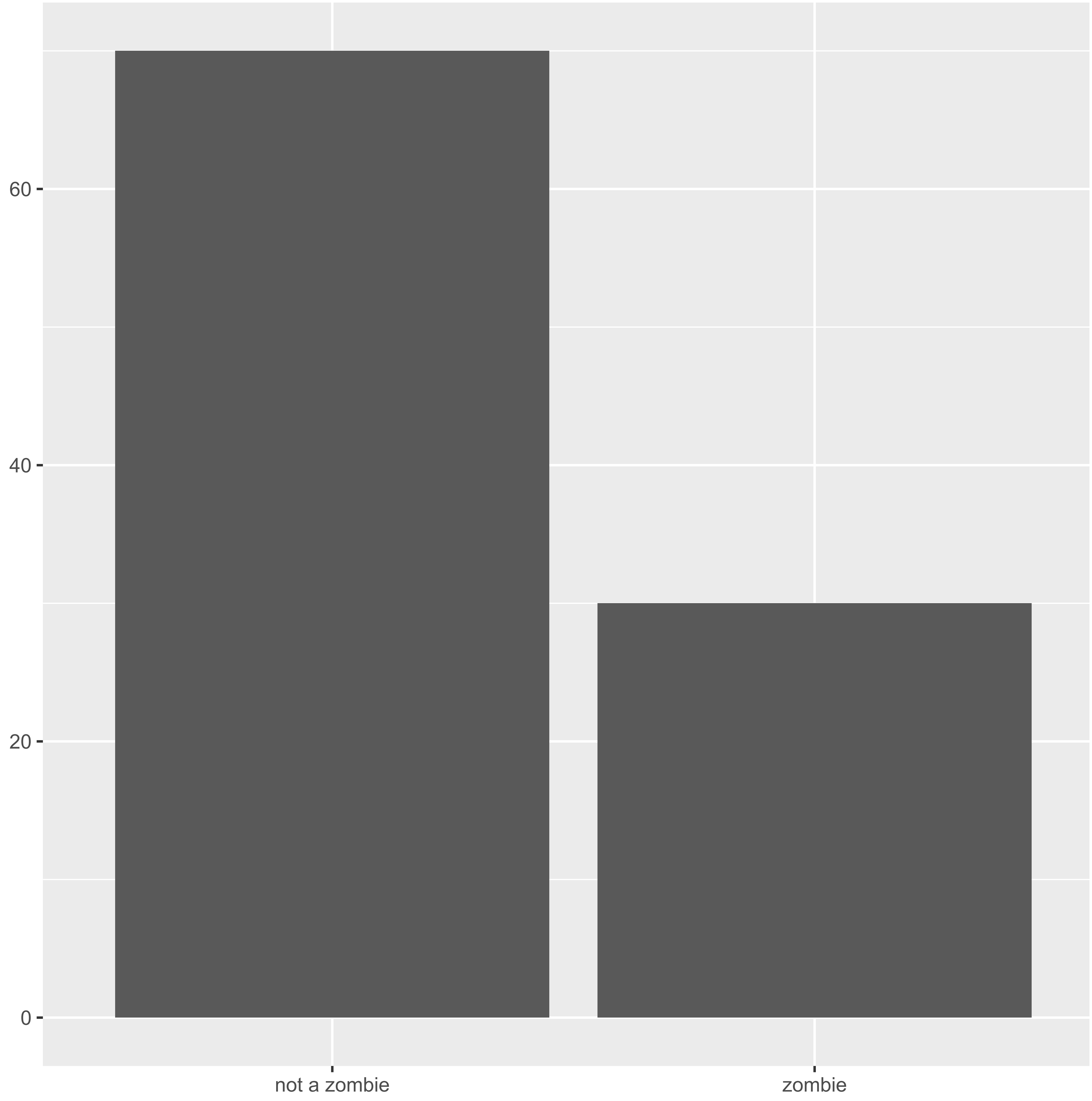
$$pr(A, B) = pr(A|B)pr(B)$$

This generalises to:

$$pr\left(\bigcap_{i=1..n} E_i\right) = pr\left(E_n \bigcap_{i=1..n-1} E_i\right) * pr\left(\bigcap_{i=1..n-1} E_i\right)$$

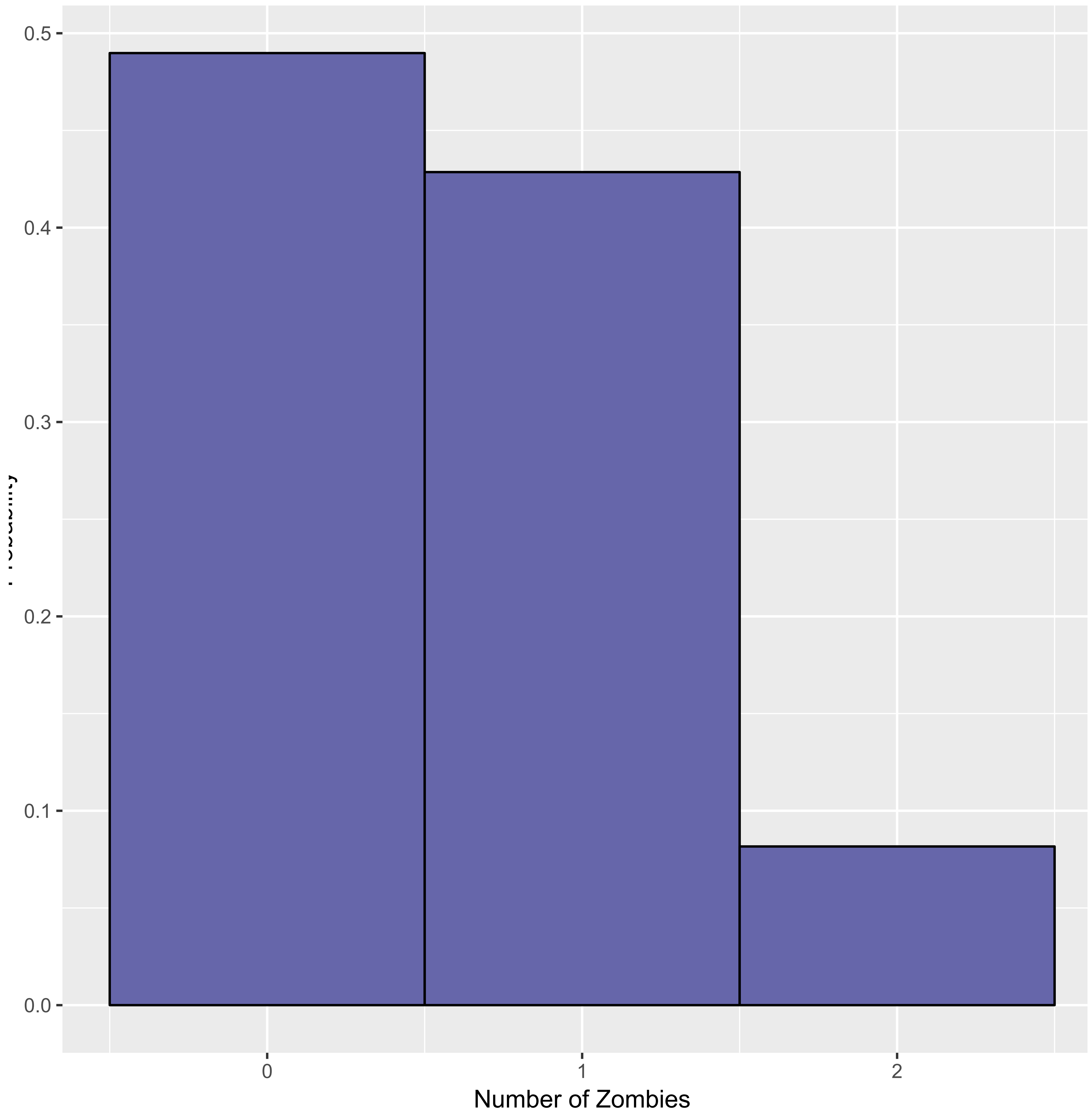
SO OUR TEST ISN'T GOOD ENOUGH

- Solution: Make more zombies
- If pr(**Z**) = 0.3, and we draw one person at random, we have two outcomes:
  - **zombie** (30% chance), or **not a zombie** (70% chance)
- Trivially, if we do this 100 times, we expect the results to look like this:



DRAW 2 PEOPLE

- We now have 3 outcomes: 0, 1, or 2 zombies
- If the draws are independent
  - [0] = (1-[**Z**])(1 - [**Z**]) = 0.49
  - [1] = 2[**Z**](1 - [**Z**]) = 0.42
  - [2] = [**Z**][**Z**] = 0.09
- This is a **probability distribution** - literally the distributions of total probability (=1) over the possible outcomes



GENERALISING

- If we take **n** draws with probability of the event (e.g., being a zombie) **p**, can we compute the probability of **k** events?

$$pr(k|n,p) = f(n,p) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- This is the **probability mass function (PMF)** of the **binomial distribution**

PMF, PDF, CDF

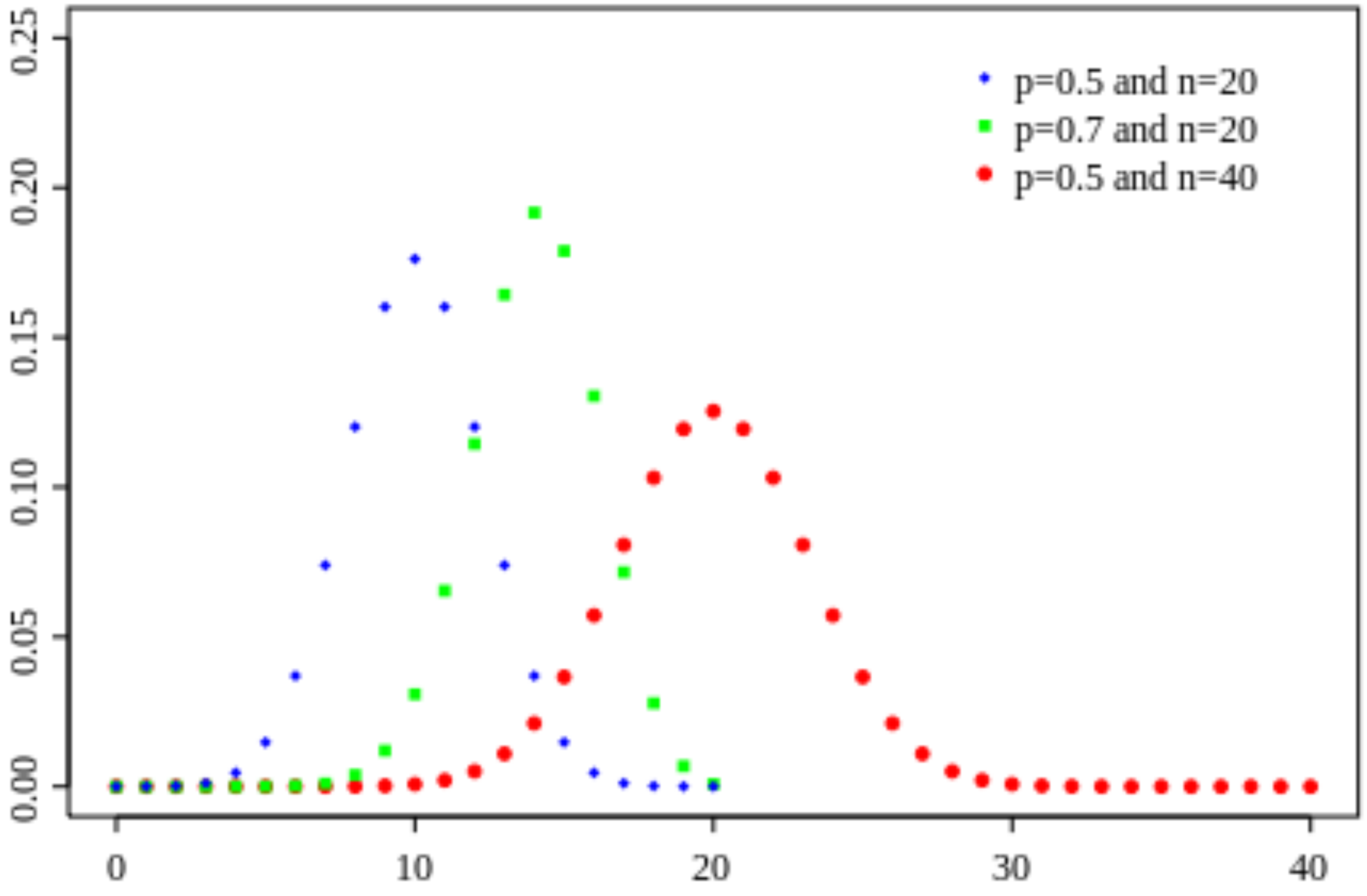
- A **probability mass function** (discrete distributions) or **probability density function** (continuous distributions) **f(x)** with parameters **x**:
  - is defined on an interval [a, b] (may be infinite)
  - is positive
  - is regular (one value for f for each value of x, finite derivative), and:

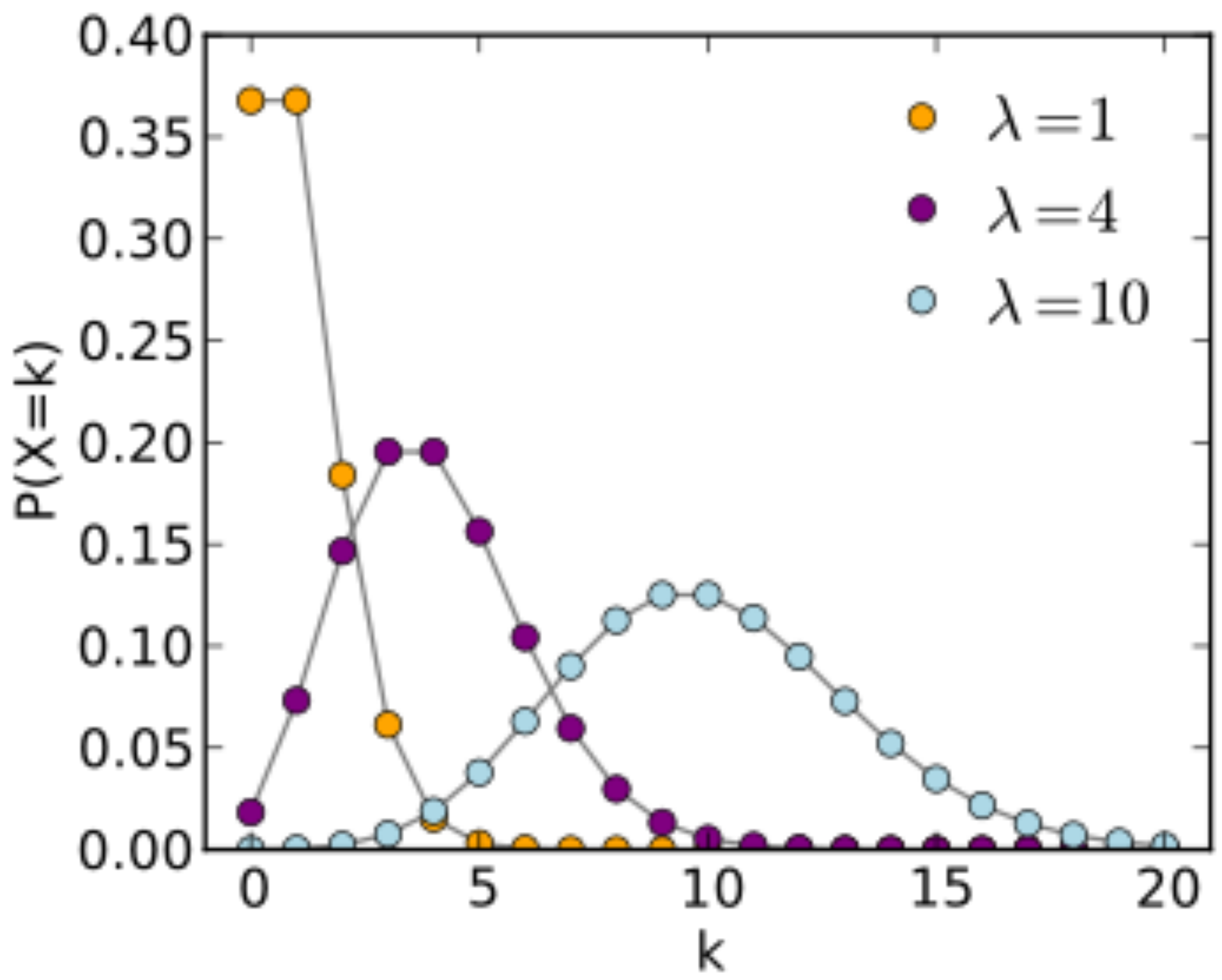
$$\int_a^b f(x)dx = 1$$

- For every PDF there is a corresponding **cumulative density function** g(x,i) describing the probability of a value between a and i

DISCRETE DISTRIBUTIONS

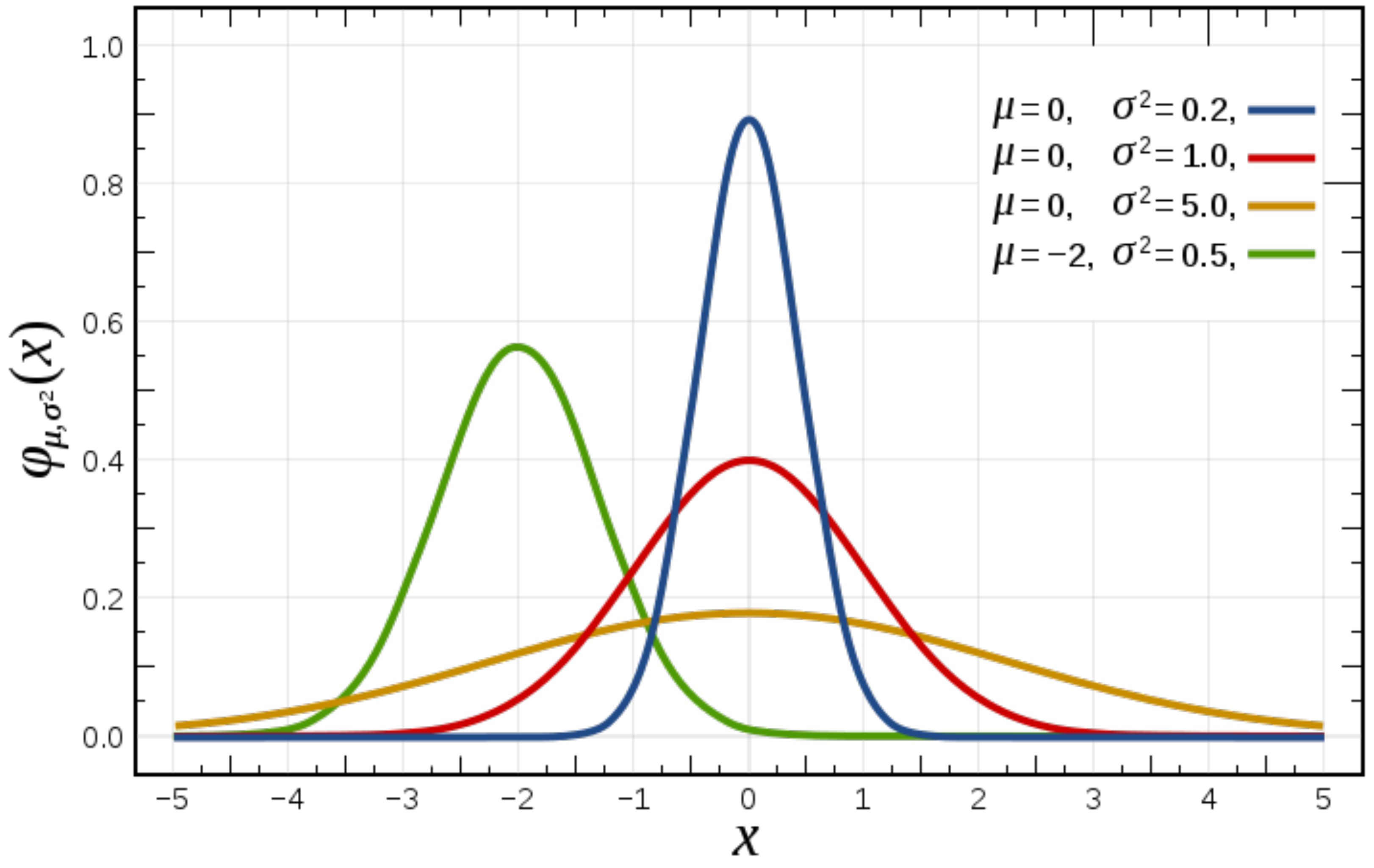
- Binomial: f(n,p)
- Poisson: f(λ)

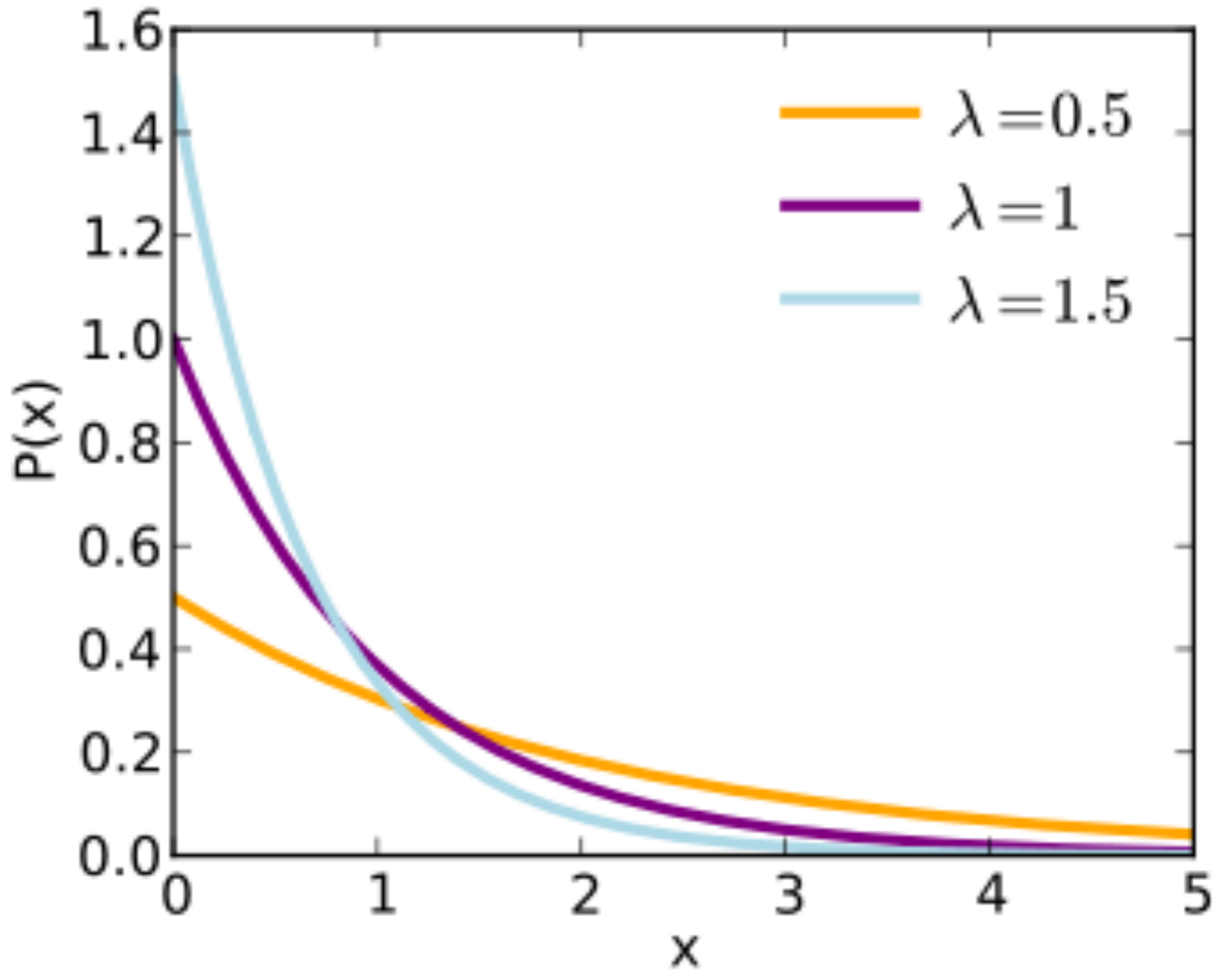


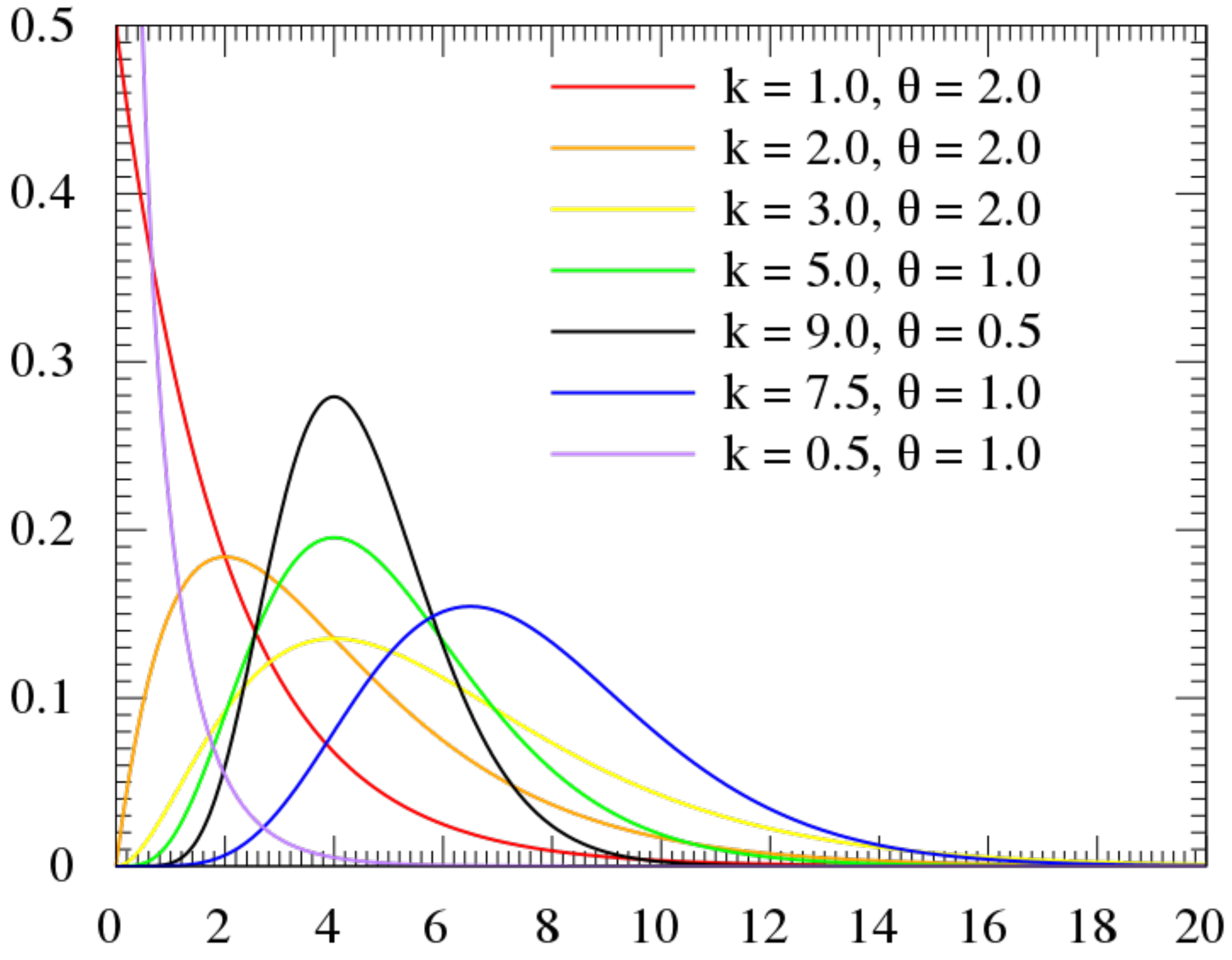


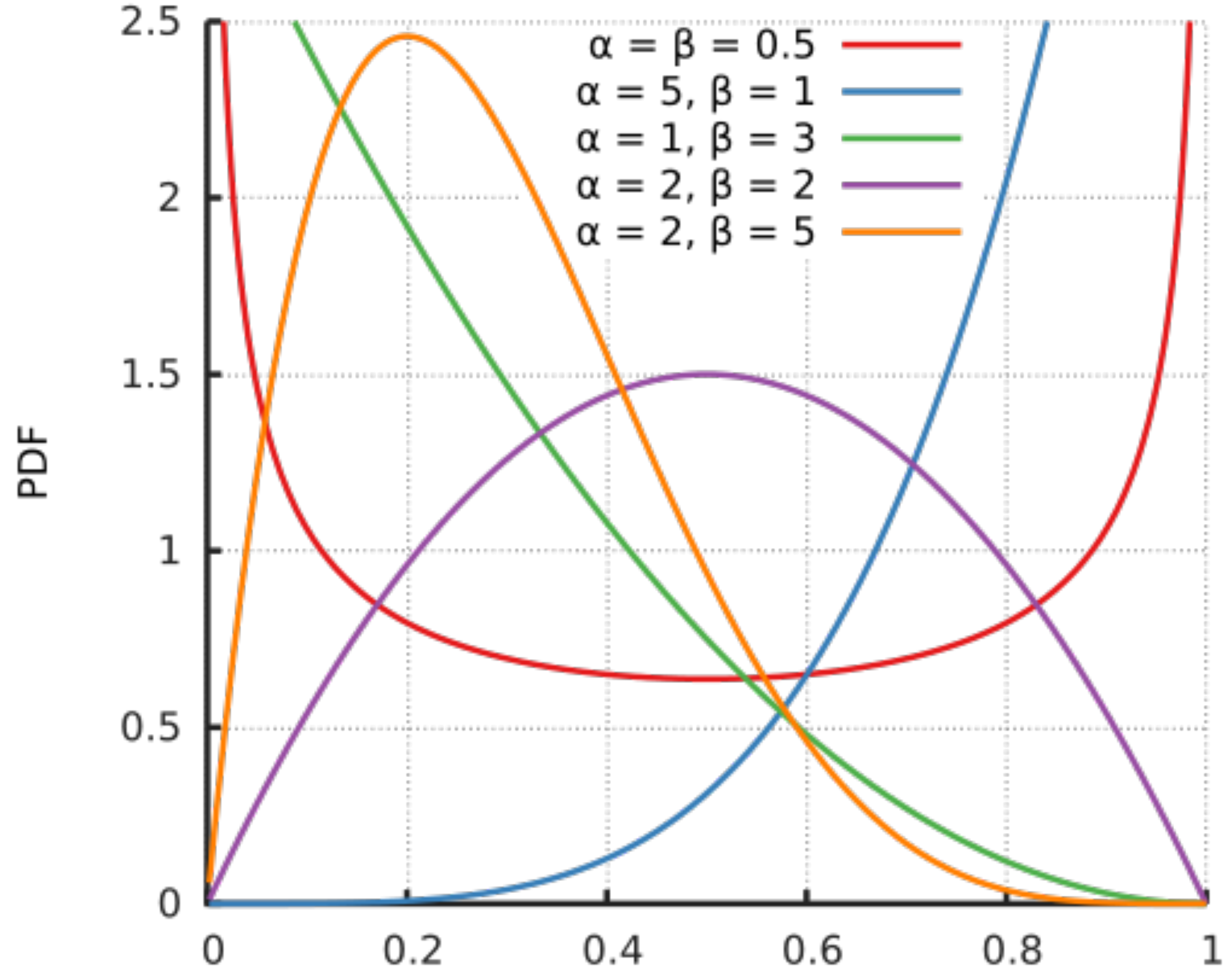
CONTINUOUS DISTRIBUTIONS

- Normal: f(μ,σ)
- Exponential: f(λ)
- Gamma: f(k,θ)
- Beta: f(α, β)









DISTRIBUTIONS IN R

- R has families of distributions:
    - binom, pois, norm, exp, gamma, beta
- And functions for each:
    - probability **d**ensity, cumulative **p**robability, **q**uantiles, and **r**andom draws
- `dbinom(3, 10, 0.1)` will return the probability of getting exactly 3 events ("successes") in 10 trials with a probability of 0.1

R PRACTICE

- What is the probability of observing exactly 6 events in a minute for a poisson process with lambda = 3.6 events/minute
- What about for observing **6 or more** events?
- Is a **probability density** the same as a **probability**?
    - What is the maximum probability density of a normal distribution with mean=0 and sd=0.1, and at what value (**x**) does this occur?
    - If it's not the same, what is the probability of observing **x**?
    - Do these answers make sense? How?
    - What is the probability of observing a value of **x ± 0.02**?
- For the same normal distribution, find the value **x** such that the probability of observing **x or less** is 0.4. What is **x** if the probability of observing **greater than x** is 0.4?

A MORE NATURAL WAY TO THINK ABOUT STATISTICS

- I want to describe some phenomenon (“model”)
- I have some general (“prior”) knowledge about the question
- I gather additional knowledge (“data”)

What is the probability that my **model is correct** given what I **already know about it** and **what I’ve learned**?

BAYES' THEOREM

- We really want pr(model | data); call model = θ and data = X
- From the product rule, we know that:

$$pr(\theta|X) = \frac{pr(\theta, X)}{pr(X)}$$

$$pr(\theta|X) = \frac{pr(X|\theta)pr(\theta)}{pr(X)}$$

$$pr(Z|T) = \frac{pr(T|Z)pr(Z)}{pr(T, Z) + pr(T, !Z)}$$

BAYES' THEOREM

posterior probability

likelihood

prior probability

$$pr(\theta|X) = \frac{pr(X|\theta)pr(\theta)}{pr(X)}$$

normalising constant

For Bayesian inference, each of these terms is actually a **probability distribution**

BAYES' THEOREM

$$pr(Z|T) = \frac{pr(T|Z)pr(Z)}{pr(T,Z) + pr(T,!Z)}$$

- How to evaluate pr(X)?
- For the zombie example, we added up all the different ways one could test positive
- How to do this for a continuous PDF?

$$pr(\theta|X) = \frac{pr(X|\theta)pr(\theta)}{\int pr(X|\theta)pr(\theta)d\theta}$$

- This is a hard problem, we will come back to it