Lecture 11: Psychology as a Science

Introduction to Probability

Dr Lincoln Colling
9 Dec 2021

UNIVERSITY
OF SUSSEX

Probability

What do we mean by "probability"?

It might seem like there's an easy answer to this question, but there's *at least* three senses of **probability**.

These different senses after often employed in different contexts, because they make more sense in some contexts and not others

The three I'll cover are:

- The **classical view** of probability
- The **frequency view** of probability
- The **subjective view** of probability

The classical view

The *classical view* is often used in the context of games of chance like roulette and lotteries

We can sum it up as follows:

> If we have an (exhaustive) list of **events** that can be produce by some (exhaustive) list of equiprobable outcomes (the number of events and outcomes need not be the same), the the **probability** of a particular event occurring is just **the proportion of outcomes that produce that event**.

To make it concrete we'll think about flipping coins. If we flip two coins the possible outcomes that can occur are:

> HH, HT, TH, TT

The classical view

If we're interested in a particular event—for example, the event of "obtaining at least one head from two flips"—then we just count the number of outcomes that produce that event.

> **HH**, **HT**, **TH**, TT

Three out of four outcomes would produce the event of "at least one head", so the probability is $\frac{3}{4}$ or 0.75

If you're viewing probability like this, it's very important to be clear about what counts as a possible outcome.

E.g., When playing the lottery, how many outcomes are there?

- Two outcomes? You pick the correct numbers or you don't? So the the probability of winning is $\frac{1}{2}$?
- Of course not! There's 45,057,474 possible outcomes, and 1 leads to you winning with 45,057,473 leading to you not winning!

The frequency view

When you take a frequency view of probability you're making a claim about **how often, over some long period of time** some event occurs.

- The frequency view is often the view that we take in science. If we wanted to assign a probability to the claim "drug X lowers depression", we can't just think of each possible outcomes that **could** occur when people take Drug X and then count up how many lead to lower depression and how many do not.

- No way to make an exhaustive list of every possible outcome!

- But we can run an experiment where we give Drug X and see whether it lowers depression. And we can repeat this many many times. Then we count up the proportion of experiments in which depression was lowered.

- That is then the probability that Drug X lowers depression.

The subjective view (credences)

Consider the following statement:

> The England cricket team will lose the upcoming test series against South Africa

There is a sense in which you can assign a probability to this

- But it isn't the classical kind—we can't just enumerate all the possible outcomes that lead to this event
- Nor is it the frequency kind—we can't repeat the 2020/2021 cricket tour over and over and see how often England lose.

When we talk about probability in this context mean something like *degree of belief*, *credence*, or *subjective probability*.

Probability in this context is the answer to the question "how sure are you that the England cricket team will lose the upcoming test series against Australia?"

Calculating with probability

The different views of probability have got to do with what the numbers **mean**, but once we have the have the numbers there's no real disagreements about how we do calculations with those numbers[1]

Some properties of probabilities

When we attach numbers to probabilities those numbers must range from 0 to 1

- If an event has probability 0 then it is impossible
- If an event has probability 1 then it is guaranteed

These two simple rules can help us to check our calculations with probabilities. If we get a value more than 1 or a value less than 0, then something has gone wrong!

[1] Probabilities don't always have to have **numbers** attached. There is a sense in which something can be **more probably** than something else with numbers being attached.

The addition law

Whenever two events are *mutually exclusive*:

> The probability that at least one them occurs is the **sum** of the their individual probabilities

If we flip a coin, one of two things can happen. It can land Heads, or it can land Tails. It's can't land heads **and** tails (*mutually exclusive*), and one of those things must happen (it's a list of all possible events)

- What's the probability that at least one of the those events happens? Since one of those events must happen the probability must be 1

- But we can work it out from the individual probabilities

1. $\frac{1}{2}$ possible outcomes produces Heads—P(Heads) = 0.50

2. $\frac{1}{2}$ possible outcomes produces Tails—P(Tails) = 0.50

The probabilities of at least one of **Heads** or **Tails** occurring is 0.5 + 0.5 = 1

Mutually exclusive and non mutually exclusive events

Consider a deck of cards:

1. What is the probability pulling out a Spade or a Club?

2. What is the probability of pulling out a Spade or a Ace

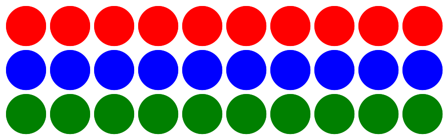
In situation (1) the events are *mutually exclusively* or *disjoint*. A card can't be a Spade AND a Club. It will either be a Space, a Club, or something else. The addition rule applies:

- P(Spade) + P(Club) = Probability of selecting a spade or a club.

In situation (2) the events are **not** mutually exclusive. A card out be both a Spade and an Ace.

- So we need a different rules

To make this clear, we'll take a look at an example



The display shows 10 red circles, 10 circles, and 10 circles.

- The probability of selecting a red circle is $\frac{10}{30}$ or 0.333
- The probability of selecting a blue circle is $\frac{10}{30}$ or 0.333
- The probability of selecting a green circle is $\frac{10}{30}$ or 0.333

Selecting **blue** and selecting **red** are **mutually exclusive**. This means that if you select **only one circle**, that circle can't be both blue **and** red. Although we can't work out of the probability of selecting blue **and** red, we could ask about the probability of selecting, for example, a circle that is blue **or** red.

To work out this probability we apply the **addition rule**.

Mathematically, we can say: $P(A \cup B) = P(A) + P(B)$

Let's put some numbers to it:

Number of red

10

Number of green

10

Number of blue

10

Colours

red blue green

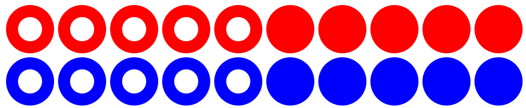
The probability of selecting red **or** blue is:

$P(\text{red} \cup \text{blue})$ = 0.333 + 0.333 = 0.667

We can work this out just by counting:

10 (red) + 10 (blue) = 20

20 ÷ 30 (total number of circles) = 0.667



Now we'll change the display so that some circles have white dots and some have no dots.

Now we can ask, what is the probability of selecting a **red** circle **OR** a circle with a white dot.

- The probability of selecting a circle—P(Red)—is $\frac{10}{20}$ = 0.5

Next, we'll calculate the probability of selecting a dot with a white dot.

- The probability of selecting a circle with a dot—P(Dot)—is $\frac{10}{20}$ = 0.5

Number of blue

10

Number of blue with dots

5

Number of red

10

Number of red with dots

5

Show all

Only show blue

Only show red

Only show with dots

Only show without dots

If we want to know the probability of selecting a circle that is **red** or **has a white dot** can we just apply the **addition rule**?

If we did we'd get $\frac{10}{20}$ + $\frac{10}{20}$ = $\frac{20}{20}$ or 1!.

This **isn't correct**. It **can't be correct**. We can see this but adjusting the sliders above. Set the first three sliders to 10, and the last slider (Number of red with dots) to 5. Clicking the button below with set the sliders for you.

Set sliders

Notice that when we apply the addition rule you end up **double counting** some circles.

Show doubled counted dots

The circles that are both red **and** have dots are counted twice. So instead of just adding the two probabilities, we add the two probabilities and then **subtract** the ones that were double counted. This gives us $\frac{10}{20}$ + $\frac{10}{20}$ - $\frac{5}{20}$ = $\frac{15}{20}$ or 0.75. This obeys our rules of being 1 or less.

Two or more events

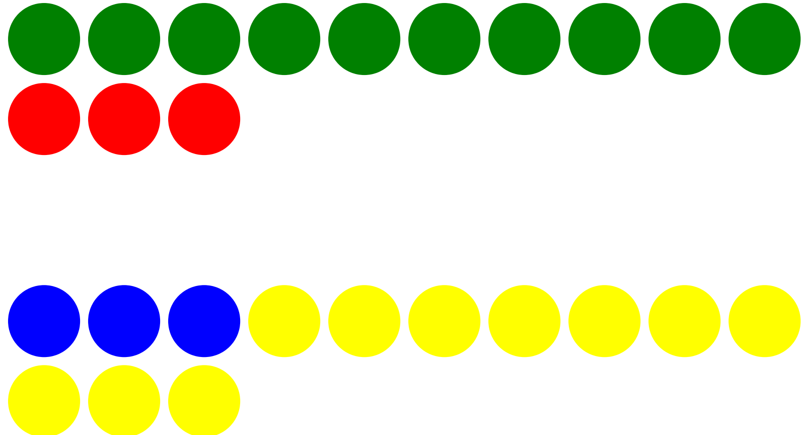
In the last example we asked about the probability of selecting a **red circle** or a **circle with a white dot**

In this example we're dealing with an event that only produces **one** outcome.

- Selecting a circle that is red
- Selecting a circle with a white dot
- Selecting a circle that is blue

But we can make things more complex so that we're dealing with events that produce **multiple** outcomes.

There's a few different scenarios that can happen when we're dealing with multiple events, so we'll start simple and then get more complex...



In this scenario we have **two sets** of coloured circles. You will pick one circle from each set.

- The probability of selecting a green circle is $\frac{10}{13}$ and the probability of selecting a red circle is $\frac{3}{13}$
- The probability of selecting a blue circle is $\frac{3}{13}$ and the probability of selecting a yellow circle is $\frac{10}{13}$

If you have two picks, one from the first set and one from the second, then what is the probability of selecting a **green** circle and a **yellow** circle? We can't just add $\frac{10}{13} + \frac{10}{13}$ because that would equal 1.538!

Instead, we multiply: $\frac{10}{13} \times \frac{10}{13} = \frac{100}{169}$

Here our event can produce outcomes such as:

- Selecting a blue circle and a red circle
- Selecting a green circle and yellow circle etc

Two or more events

coins 0

When there are **0** coins there are **1** possible sequences.

- We can also just count when we're dealing with multiple events
- But this is often easier to do when we draw probability trees

Independence

In the previous example, the two choices were **independent**

This just means that the outcome of either choice doesn't influence the probability of the other choice

That is, we can calculate the probability of each event without considering **anything about the other event**

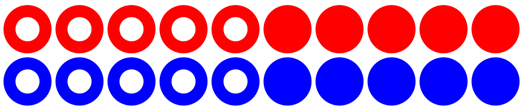
When this is the case, we can calculate the probabilities of both events occurring just by multiplying the two probabilities

But sometimes this isn't the case... sometimes the probability of a second event is **dependent** on the first event

Let us look at a simple example...

In these examples, rather than selecting **two** circles we'll ask about the probability of a single circle having **two** features

Conditional probability



Change values

We could ask a question like:

> What is the probability of selecting a blue circle and a circle with a white dot in **two** selections

To answer this, we'd just work out $\frac{10}{20} \times \frac{10}{20} = \frac{100}{400}$, or 0.25 because the result of the first draw doesn't influence the result of the second draw (they're independent).

But what if we instead want to answer the question:

> What is the probablity of selecting a blue circle and a circle with a white dot in **one** selection.

Now the two things are no longer independent. First we work out the probability of selecting Blue ($\frac{10}{20}$), and then we select Blue...

Select blue

Then we work out the probability selecting a circle with a white dot ($\frac{5}{10}$)

If multiply $\frac{10}{20} \times \frac{5}{10} = \frac{5}{20}$ or we could just count!

Show all

Conditional probabilities

In the last example I introduced the idea of a **conditional probability**

- We knew P(Blue): The probability of a circle being blue *independently of whether it had a dot or not*
- And P(Dot): The probability of a circle having a dot *independently of whether it was red or not*
- But to answer our question we needed to know a conditional probability

P(Blue) × P(Blue|Dot): $\frac{20}{35} \times \frac{15}{20} = \frac{15}{35}$

We could of also done it the other

P(Dot) × P(Dot|Blue): $\frac{20}{35} \times \frac{15}{20} = \frac{15}{35}$

Or we could just count the number of circles (out of all the circles) that are both Blue **and** have a white dot

This example makes it seem like conditional probability is pretty easy (it's all about counting the correct circles!), but it can be tricky to understand!

Conditional probabilities and independence

Although it didn't seem like it, the first example with the **two sets** of circles also involved conditional probabilities

However, P(Green) was equal to P(Green|Yellow) and P(Yellow) was equal to P(Yellow|Green)

- The probability of picking Green didn't change given that we'd already picked Yellow
- And the probability of picking Yellow didn't change given that we'd already picked Green

This is the **mathematical definition of independence**

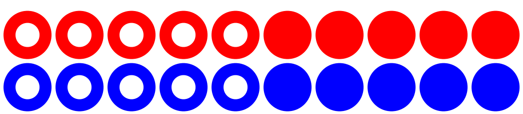
Working with conditional probabilities

In our blue circle white dot example we saw that P(Blue|Dot) and P(Dot|Blue) were equal

- P(Blue|Dot) = $\frac{15}{20}$
- P(Dot|Blue) = $\frac{15}{20}$

But conditional probabilities and their inverse are not always equal

- P(Red|Dot) = $\frac{5}{20}$
- P(Dot|Red) = $\frac{5}{15}$



Show all
Only show blue
Only show red
Only show with dots
Only show without dots

Change values

Bayes theorem

There's a mathematical formula that related P(A|B) to P(B|A). This formula is know as **Bayes theorem**.

Bayes theorem is very useful for thinking about conditional probabilities, because conditional probabilities can sometimes be incredibly *unintuitive*

Consider the following example:

Does a positive test mean somebody is actually sick?

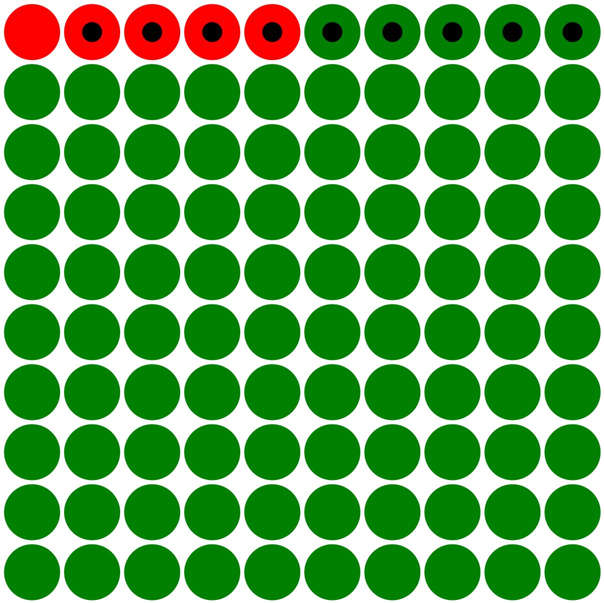
There is a test for an illness. The test has the following properties:

1. About 80% of people that **actually have the illness** will test positive
2. Only ~5% of people **that don't have the illness** will test positive

Somebody, who may be sick or healthy, takes the test and tests positive...

Is that person actually sick?

Does a positive test mean somebody is actually sick?



Incidence

Common ◯ Rare ◉

Positive tests only ☐

80% or $\frac{4}{5}$ sick people test **positive** test and 5.263% or $\frac{5}{95}$ healthy people test **postive**. Given a **positive test**, there's a $\frac{4}{9}$ chance or 44.444% probability that the person is actually sick.

The probability of the person actually being sick **depends** on the incidence of the disease.

If the disease is rare then there's a low probability that the person is actually sick
If the disease is common then there's a high probability that the person is actually sick

Bayes theorem

We can work out the answer to the previous question just by counting the dots, but we can also use Bayes theorem.

Bayes theorem is given as:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

or

$$P(\text{🤮} \mid \text{✅}) = \frac{P(\text{✅} \mid \text{🤮}) \times P(\text{🤮})}{P(\text{✅})}$$

and when we put numbers to it...

$$\frac{4}{9} = \frac{\frac{4}{5} \times \frac{5}{100}}{\frac{9}{100}}$$

Note that the crucial values here are P(🤮) and P(✅). These are sometimes referred to as the prior probabilities or unconditional probabilities.

If you change the value of P(🤮) then you're changing how rare or common the disease

Reasoning with Bayes theorem and conditional probabilities

- Reasoning about conditional probabilities like the testing example can be difficult between people often forget about the P(🤮) and P(✅) bits.
- But we ignore P(🤮) and P(✅) we can see it's easy to make mistakes!
- Another common error is to confuse P(🤮|✅) and P(✅|🤮) or to think P(🤮|✅) = P(✅|🤮)
- But we saw from our earlier example that this isn't the case
  - You can think of the following example to help remind you of this. P(Lives in London | Is Boris Johnson) = 1, but P(Is Boris Johnson | Lives in London) = 1 in 9 Million

The media and (the scientific literature) is unfortunately littered with examples of people getting muddled with conditional probabilities

And some of these confusions can actually be **dangerous**!

I'll just pick out two more examples to finish on...

Does the Covid vaccine work?

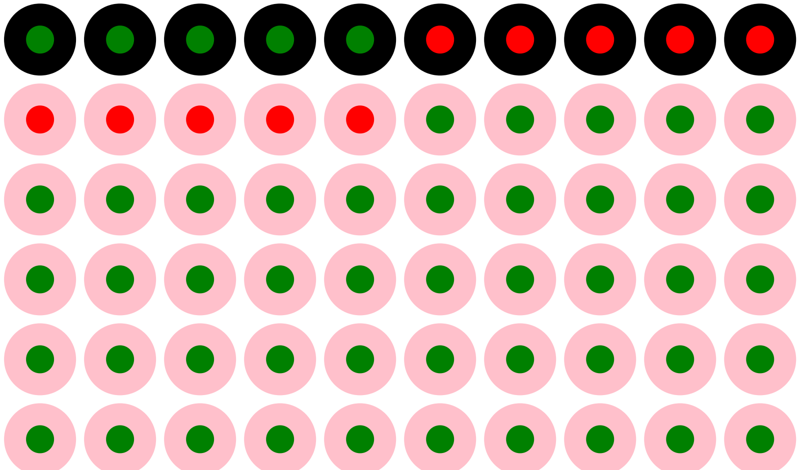
You might have heard the following statistic in the media/online

> 50% of people die from Covid have been vaccinated

I've seen this stat on social media along with the claim that it *shows that the Covid vaccine doesn't work*

Let's assume the state is accurate. Does what does this mean the vaccine doesn't work?

Does the Covid vaccine work?



Vaccination rate

Low High

Show deaths only

Vaccination rate is 80%.

Given a **person has died**, there's a $\frac{5}{10}$ chance or 50% probability that the person was vaccinated.

The vaccine **works** because:

- 50% or $\frac{5}{10}$ unvaccinated people die.
- 10% or $\frac{5}{50}$ vaccinated people.

In this example we're keeping the vaccine efficacy constant (50% chance of dying in not vaccinated and 10% chance of dying if vaccinated). The vaccinate doesn't prevent all deaths, but it does offer protection!

Racial bias in police shootings

A few years ago Johnson *et al* published a study (in a very prestigious journal) about **racial bias** in **police** shooting.

Their finding can be summed up as follows:

> There is no racial bias in police shootings because people shot by police are more likely to be White than Black

This was picked up by the conservative media (e.g., Fox news) to show that organisations like BLM were fighting against a problem that didn't exist!

But is the reasoning correct, and do the data show what Johnson *et al* claim?

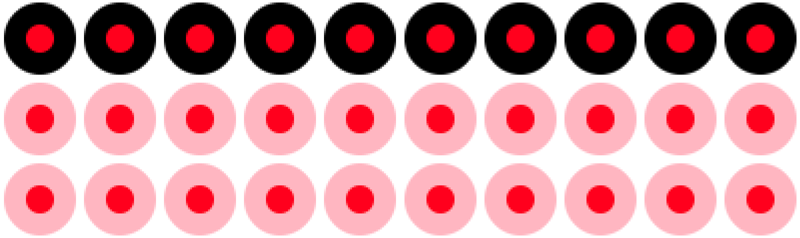
NO!

Johnson *et al*, the journal reviewers, the journal editors[1], and the media were looking at whether P(Black|Shot) was larger than P(White|Shot) when they should've been looking at whether P(Shot|Black) was larger than P(Shot|White)

[1] This paper has now been retracted from the journal after a campaign that started on twitter, but the damage is maybe already done

Racial bias in police shootings

Let's firt look at the data[1] Johnson *et al* present



In graphic shows the people shot by the police.

The probablity that a person is White (P(White|Shot)) is $\frac{20}{30}$ or 66.67%

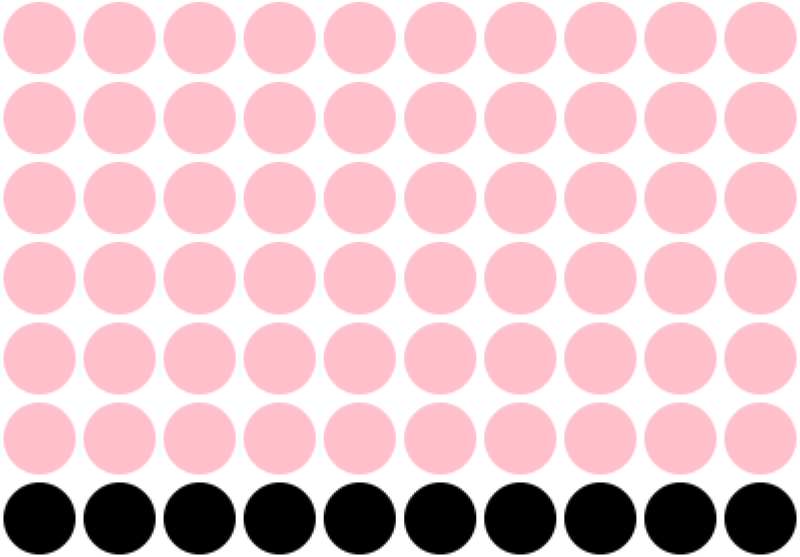
The probablity that a person is Black (P(Black|Shot)) is $\frac{10}{30}$ or 33.33%

These are the two probabilities that Johnson *et al* look at.

[1]These aren't the actual data, but I'm simplified it to make things easier

Racial bais in police shootings

But let's add some additional data. These are the people that have had *encounters* will police that **didn't** end in a shooting

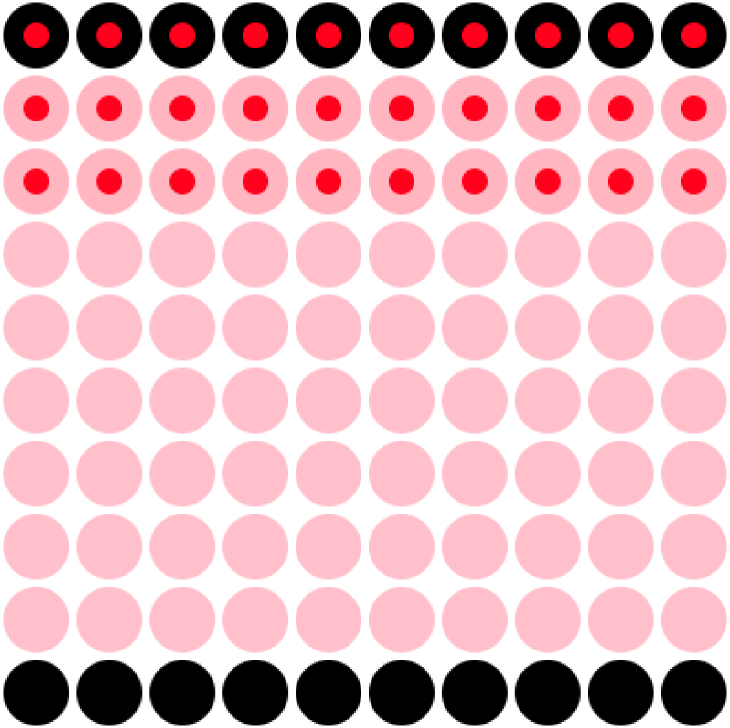


Jonson *et al* didn't report this data, so I'm made this up for illustration

We need this data, because instead of looking at P(Black|Shot)/P(White|Shot) we need to look at P(Shot|Black) and P(Shot|White)

Racial bais in police shootings

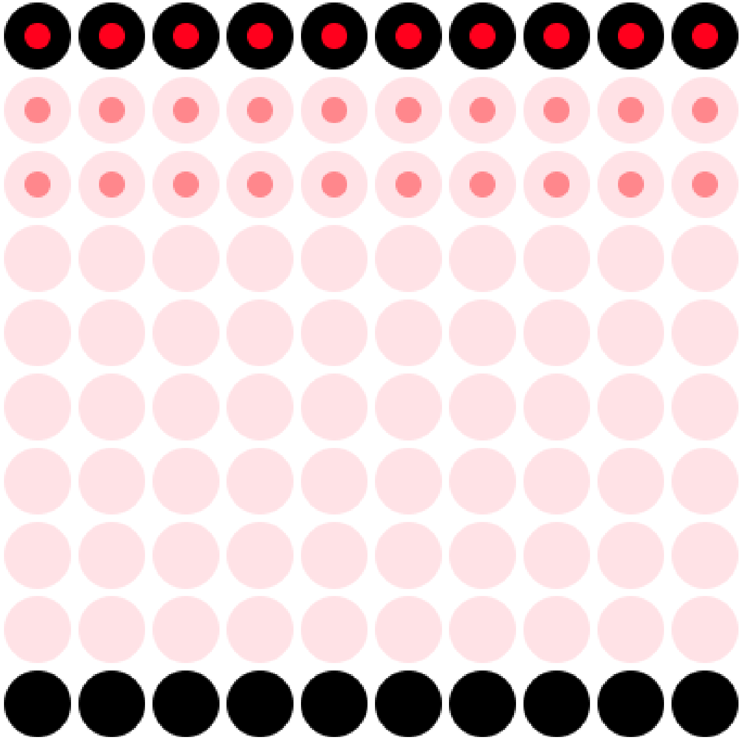
Putting the it all together we see this:



These are all the *encounters* that occurred between the police and civilans including those that ended in the police shooting a civilans and those that did not.

Racial bais in police shootings

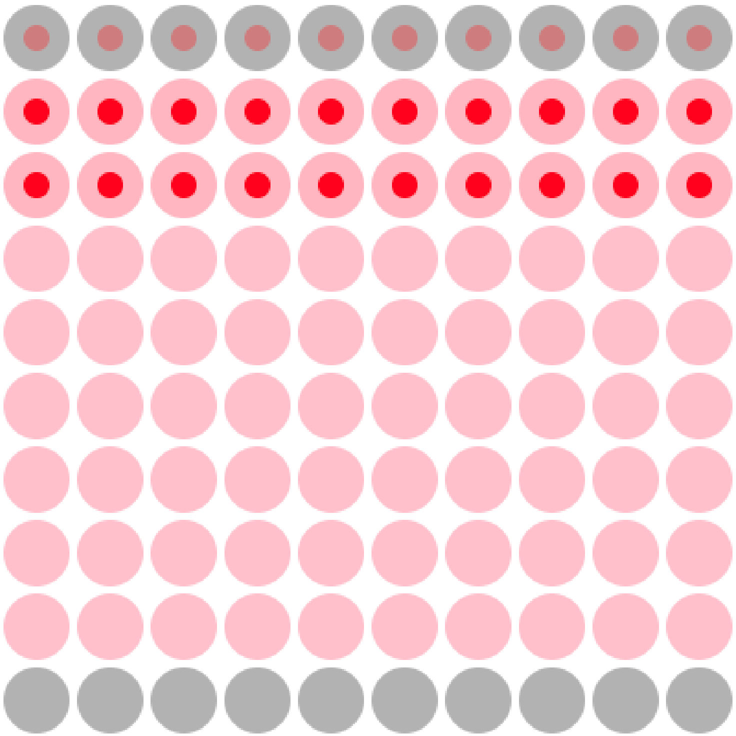
Now we **condition on being Black**



This allows us to see P(Shot|Black), which gives $\frac{10}{20}$ or 50%

Racial bais in police shootings

And **condition on being White**

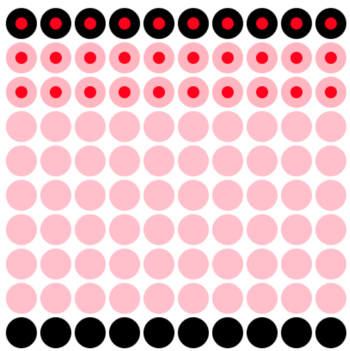


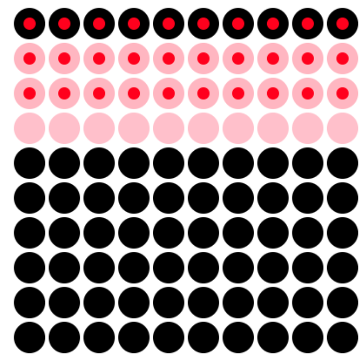
Which allows us to see P(Shot|White), which gives $\frac{20}{80}$ or 25%

These data, at least, suggest there is a bias in Police shooting

Racial bais in police shootings

Unfortunately, Johsnon *et al* didn't collect the data they need to draw their collusions





- Both figures above as consistent with P(Black|Shot) = 0.33 and P(White|Shot) = 0.67
- But one give P(Shot|Black) = 0.5 and P(Shot|White) = 0.25
- And the other gives P(Shot|Black) = 0.14 and P(Shot|White) = 0.67

Either one of these *could*[1] be the case, but Johnson *et al*'s data can't tell us this and therefore, they have absolutely no basis to support their claim

[1]Give the racial makeup of the US the first seems more plausible than the second