

PHIL 3300 - Philosophy of Medicine

Numbers from the example in class today: (Oct 9, 2025)

Here are the charts I put up today:

	smokers	Non-smoker
men	8	32
women	6	54

	Height < 5'8"	Height at least 5'8"
smoker	6	8
Non-smoker	53	33

	Height < 5'8"	Height at least 5'8"
men	4	36
women	54	6

Is this possible? The answer is no, this is impossible. Ask yourself how many people are less than 5'8" in the group we are looking at? If you look at table 2, there are 59. If you look at table 3, 58. So my eyeballing was close, but not accurate. Lets change the last table:

	Height < 5'8"	Height at least 5'8"
men	4	36
women	55	5

Now this is actually possible. The real way to know would be to write a single table with the 8 possible combinations - how many male smokers under 5'8" in the group, etc. Any such table can be turned into three charts like this. But not all sets of three charts are consistent with any specific breakdown and if they are, they will usually be consistent with multiple. -- A good (fun!) exercise is to figure out what the full 8 row chart could look like in this case (hint: there are lots of possible tables. Assign variables  $x_1$ - $x_8$  to MSH,  $MS\sim H$ , etc. write down all the constraints, and get to work).

Okay, now lets look at some probabilities, conditional probabilities, and risks.

Probability here just refers to a frequency.

Total people = 100

$$P(M) = 40/100$$

$$P(W) = 60/100$$

$$P(S) = 14/100$$

$$P(\sim S) = 86/100$$

$$P(H) \text{ (means height is } \geq 5'8\text{'')} = 42/100$$

$$P(\sim H) = 58/100$$

$P(M\&S)$  is the probability of a conjunction - the probability that they are male and they smoke.

$$P(M\&S) = 8/100$$

$$P(W\&\sim S) = 54/100$$

$$P(S\&H) = 8/100$$

$$P(H\&W) = 5/100$$

Etc.

Conditional probability in this case is the frequency of one group within a specific restricted group. Read  $P(A|B)$  as the probability of A given B. That is, the probability of A conditional on B. By definition,  $P(A|B) = P(A\&B)/P(B)$ . So for example,  $P(S|M) = P(S\&M)/P(M) = (8/100)/(40/100) = 8/40$ . In other words, 8 of the 40 men are smokers.

Other examples:

$$P(M|S) = 8/14$$

$$P(H|M) = 36/40$$

$$P(H|S) = 8/14$$

Make sure you know how I got those numbers.

Now we can also see which variables are correlated:

A is positively correlated with B if any of these are true (these are all equivalent conditions):

$$P(A) \times P(B) < P(A\&B)$$

$P(A|B) > P(A)$   
 $P(A|B) > P(A|\sim B)$   
 $P(B|A) > P(B)$   
 $P(B|A) > P(B|\sim A)$

So for example, being male is positively correlated with being tall.

$P(M) \times P(H) < P(M\&H)$  because  $40/100 \times 42/100 < 36/100$

Here is another way to get the same answer:

$P(M|H) > P(M)$  because  $36/41 > 40/100$ .

Being tall is also positively correlated with smoking:

$P(H) \times P(S) < P(H\&S)$  because  $42/100 \times 14/100 < 8/100$

--- I should say that probabilistic independence means that all of these inequalities are equalities. So if in a population, 20% of the men were smokers and 20% of the women were smokers, then sex and smoking are independent -  $P(M) \times P(S) = P(M\&S)$ . Similarly,  $P(S) = P(S|M) = P(S|W) = 0.20$

How do we figure out which of these correlations are also causal? Being male seems causally related to being tall and being male might actually be causally related to smoking (though we want to know what we MEAN by causation here. But if it means something like ‘not an accidental relationship and no other variables screen it off and ...’.) But in class I suggested that it is plausible that height is not causally related to smoking. So height in this case is correlated with smoking, but not causally related. Why? Because this is a case of a common cause explanation. The reason that tall people are more likely to be smokers in our data set is because men are more likely to smoke and are also more likely to be tall. Hence tall people are more likely to be smokers.

In order to check that height BY ITSELF is not causally relevant, what you want is to see if height makes a difference when you hold fixed the sex. That is, among the men now check whether height is correlated with smoking. And among the women, check whether height is correlated with smoking. To do that compare:

$P(S|M)$  vs.  $P(S|M\&H)$

If it turns out that  $P(S|H) > P(S)$  but that  $P(S|H\&M) = P(S|M)$  then we say that sex “screens off” height from smoking. You can also see this by looking at conditional independence.

A and B are independent when  $P(A) \times P(B) = P(A\&B)$ . A and B are independent conditional on C if

$P(A|C) \times P(B|C) = P(A\&B|C)$ . In this case, C would screen off B from A.

If sex and height are both correlated with smoking, but sex screens off height from smoking and height does not screen off sex from smoking then sex is the cause and not height. If neither of them screen off the other, then they might both be causes (or there is some fourth variable...)

So what about this case? Well, from the information given, we can't answer that. We need to know  $P(S|M\&H)$  which is  $P(S\&M\&H)/P(M\&H)$ . But we don't know  $P(S\&M\&H)$ . So just with the information given, we can't answer the question of whether H&M are independent given M or not.