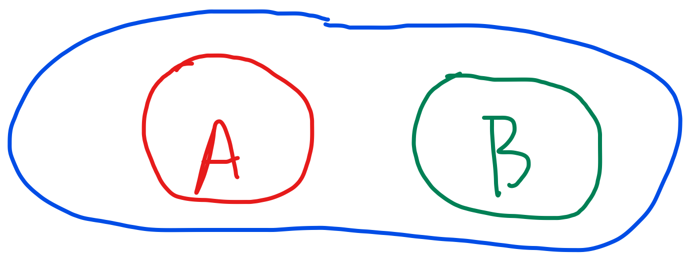
Independence vs Dependence

We had an excellent discussion on the meaning of independence, and thus of dependence. Next are some notes to recap the discussion.

Topic 1: Disjoint events A and B cannot be independent. That is, disjoint events are dependent. The reason is that belonging to event A automatically excludes the individual from belonging to event B. This means that the probability of the individual belonging to event B is equal to 0 when it is known that the individual belongs to event A.

Definition: Two events C and D are independent if knowing that one of the events has occurred tells us nothing about the probability of the other event occurring.

You now clearly see why the disjoint events A and B cannot be independent.



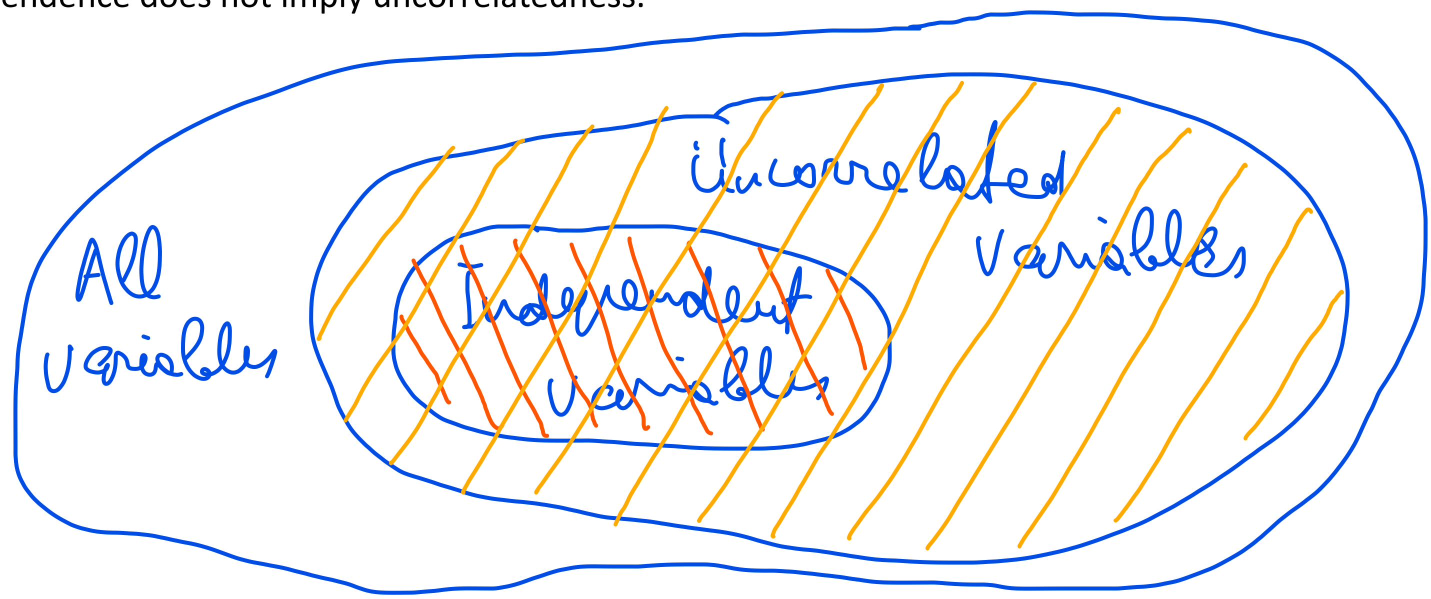
Disjoint and thus dependent

Topic 2: If two variables (e.g., some properties) X and Y are independent, then they are uncorrelated.

If, however, X and Y are uncorrelated, then they can be either independent or dependent. That is, uncorrelatedness does not imply independence.

If two variables X and Y are correlated, then they are dependent.

If, however, X and Y are dependent, then they can be either correlated or uncorrelated. That is, dependence does not imply uncorrelatedness.



Topic 3: To illustrate independent and dependence, we analyzed the following example during the class:

Let E mean “engineering student” and F “female”. My original reaction was that E and F are independent events, and the reason was this:

Suppose that a counsellor sends me a message saying that an engineering student wants to write a make-up final. From this information, I wouldn’t have a clue whether the student is female or not (if this by any chance mattered for some reason). Likewise, if I meet a female student in a cafeteria, I wouldn’t have a clue whether the student studies engineering or something else. Hence, for me, the two events E and F looked very much independent.

Once I finished this argument, a student attending the class remarked to me that there are more male engineering students than females. This bit of information changed my perspective drastically by making the events E and F dependent. Why so? Let’s replay the whole scenario again, but under the additional bit of information provided by the student.

Suppose that a counsellor sends me a message saying that an engineering student wants to write a make-up final. From this information and given that I now know that there are more males than females in the Faculty of Engineering, I conclude that the student is more likely to be male than female. Likewise, if I meet a female student in a cafeteria, I would think that the student is more likely to be from some faculty other than engineering.

Is it strange that independence of E and F has turned into dependence? No, not at all. The reason is that the bit of information about more males than females in the Faculty of Engineering has become a confounding variable, and such variables can (and often do) turn independence into dependence.

To illustrate, let B be the event that the battery in my car is charged, and T the event that the tank is full of gas. The two events B and T look independent – they have nothing to do with each other, at least as far as I can see. But suppose that I get into the car, turn the key, and the car does not start. I jump out of the car, check beneath the fuel tank for possible leakage, find nothing suspicious, and thus confidently conclude that the battery must be dead, as I can’t imagine any other reason for the car not to start. Hence, suddenly, the events B and T have become dependent: my knowledge that there is still fuel in the tank (I know I filled it in yesterday) tells me clearly that the probability that the battery is functional is 0, that is, the probability that the battery is dead is 1.