Marcello's notes on Weight

9/1/2022

Purpose of this document

This contain some thoughts about weight, to be merged with the main chapter/paper on weight. I am keeping my notes here to avoid conflicts in updating the main file.

Different conceptions

Difference conceptions exist in the vicinity of one another:

- **Quantity:**

How much evidence you have. This satisfies monotonicity and is mostly what Kaynes had in might.

Hard to make it formally precise. Attempts to count items of evidence are not promising. They are purely syntactical. They cannot justify why, intuitively, several shaky testimonies should be *less weighty* than fewer solid testimonies.

Only sensible account is comparative. Take a body of evidence B and add an extra item of B, obtaining B+. Then, clearly B+ has more quantity of evidence than B. Comparison can only be made when one body of evidence is a subset of the other.

- **Completeness:**

Presumably there is a list of *all* items of evidence one would expect in a case, and thus completeness measures the gap between that list and the evidence actually presented.

Nance has this notion when he talks about *reasonable* completeness. A body of evidence is, for him, reasonably complete (relative to a claim H), when (a) it is not missing any relevant item of evidence that it would be cost justified (=reasonable) to obtain, and (b) it is missing a relevant item of evidence that, though currently impossible to obtain, could have been obtained with no unreasonable costs if not because of a fault (negligence, intentional destruction, etc.) of one of the parties.

How do we draw that list of *reasonably complete* evidence? This list is drawn in a number of different ways: (a) using a *script* (=this is the type of case in which one would expect to see these kinds of evidence); (b) using a case-specific shared narrative (=the following facts are well-established, and given those facts, we would expect to see these items of evidence); or (c) some combination of the two.

How do we measure the gap between the complete list of evidence and the actual evidence available? We seem to run into the same problems as with quantity.

Connection with quantity. Clearly, a body of evidence that has more quantity must also be more complete.

- **Resilience:**

The evidence currently available might support a certain claim to some degree, say the LR or the posterior probability (or some other measures of balance) tips strongly i favor of the claim. The question is, would the balance (posterior probability, LR) change in light of new evidence that might be presented? If no change, evidence is resilient. If change, the evidence is not resilience.

Need to restrict the resilience test to a *reasonable set* of possible further items of evidence, otherwise no body of evidence would count as resilient.

- **Weight (Rafal's account):**

The notion of weight (in Rafal's formal account) is related but not identical to quantity, completeness, resilience. Weight refers to (a) the weight of a distribution (=how informative the distribution is relative to the uniform distribution, which uninformative and thus has zero weight) and (b) the weight of evidence, called wDelta (=the difference between the weight of the prior distribution and the weight of the posterior distribution).

Question: How does this model of weight look in trial proceedings? How can it be applied?

Nance model: completeness plus resilience

Trial decision-making should go something like this:

- First, ask whether your evidence is reasonably complete. If yes, then assess balance (say posterior probability). If not, which is most often, then assess resilience.
- Second, relative to the reasonably missing evidence, test whether your current evidence is resilient. That is, test the resilience of your evidence (relative to a claim H) only against the evidence that is missing relative to the reasonably complete list. Since you do not know the value of that evidence (e.g. you do not know if a missing DNA test will be positive or negative), assume the worst case scenario. If current evidence is resilient, then assess balance (say posterior probability). If it is not resilient (=balance goes below the required standard), then you cannot assess balance and must reconsider (direct verdict, dismissal or more investigation might be necessary).

This two part model is followed by trial courts, as well as post trial courts and post conviction appellate judgments. See Jesse Lee Johnson v. Jeff Premo 2021 Oregon Appellate case. More on this below.

David Kaye model: gaps

Kaye agrees trial decision-making should be sensitive to gaps in the evidence. Instead of the standard,

$$P(S_p|E) > t,$$

he proposes the following revised threshold model:

$$P(S_p|E \wedge G) > t$$

The idea is to see whether the total evidence presented, E, as well as gaps G in the evidence, support the prosecutor's story S_p to the required threshold probability t.

Interestingly, G is part of the evidence together with evidence proper E. After all, absence of evidence is itself a fact and thus evidence in a broad sense.

Question: How do we come up with G? Any story, if true, induces a list of evidence we would expect. Whatver the difference between that list and the evidence actually presented E is the missing evidence G. This goes back to the assessment of reasonable completeness of the evidence.

Question: How do we assess $P(S_p|E \wedge G)$ and not just $P(S_p|E)$? What kind of evidentiary contribution does G provide?

In general, it seems that $P(S_p|E \wedge \neg G) > P(S_p|E \wedge G)$, or in other words, the presence of gaps should reduce the probability of S_p, other things being equal. So Kaye is advocating for *discounting* by the trier of facts. Nance opposes this idea.

Question: Kayes model follows a two part approach: first, assess completeness, and second, assess how gaps affect balance (say posterior probability). How does Kaye model compare to the two part model consisting of completeness plus resilience? Are they equivalent? What are the differences?

Dahlman model: information economics

Example: Salem Trial

Aggravated murder case. Victim was stabbed to death in her house.

Evidence against defendant:

- fingerprint match in victim's home
- cigarette butt in victim's home genetic match
- statement by informant that defendant said "offed the bitch to rob her"
- jewelry found with defendant matched victim's
- blood boot prints in victim's house match defedant's boot prints

Exculpating evidence:

- no genetic match between weapon and defendant
- no genetic match between blood i victim's home and defendant
- defendant's boots did not test positive for blood

Missing evidence (not presented at trial, fault of counsel and police):

- testimony by women living across the street. She saw white man enter victim's house first, then loud noises and screams followed and finally he run away. Later she saw black man enter the house and then leave. Defendant is black.

Trial:

- Defendant is convicted

Post conviction court:

- counsel was at fault for not presenting neighbor testimony
- no prejudice occurred since the testimony corroborates other evidence which was presented

Appellate courts

- counsel was at fault
- prejudice occurred since the testimony of the neighbor can be the basis of a solid defense (=white man did it, not black man who came in afterwards)

Comment: Note the two part structure, (*completeness test*) first ask whether evidence is missing and (*resilience test*) then assess whether missing evidence could have change the verdict.

Question: How would Kayes model or Rafal model of weight work in this case?

Example: Tin Box Case

Missing fingerprint evidence in a murder case due to police oversight. This from Dahlman's paper.

> In 2005 a man walks into a Swedish police station and says that he wants to turn himself in. His name is AA and he says that he has just killed an elderly woman who lives by herself in an apartment nearby. The police rush to the apartment and find the woman stabbed to death. In his confession AA explains to the police that he had heard that the woman kept a huge amount of cash in a tin box, and had knocked on her door and tricked her to let him in by pretending to work for the local church. He says that his plan was to distract the woman and quickly grab some money from the tin box, but she caught him in the act, and he panicked and stabbed her. At the police station AA pulls out a switch blade knife from his pocket and puts in on the table. The

knife is smeared in blood, and is sent to forensics, who quickly confirm that the blood belongs to the victim. The autopsy report is consistent with AAs confession. The angle of the stab wound suggests that the perpetrator is above medium height, which is somewhat odd since AA is shorter than medium, but can be explained if AA held the knife high. AA is prosecuted for murder. At the trial, AAs defense attorney says that he suspects that his client is giving a false confession to cover for someone else. AA has no criminal record, but he has two sons who both have previous convictions for burglary and assault. Both sons are above medium height, and are known to carry switch blade knives. AA insists that he did it. He claims that his sons have nothing to do with the murder, and gives the court a detailed and vivid story of how he committed it, that fits perfectly with the crime scene. During the trial the court learns that the tin box was found open at the crime scene, but was never examined for fingerprints or DNA. Apparently, the police did not consider this necessary, since AA had confessed and the victim's blood had been found on his knife. When the defense attorney tried to have the tin box examined for fingerprints or DNA, it was too late. The box had been wiped clean from the victim's blood, which had removed all potential traces from the perpetrator. The defense attorney argues that the police investigators committed a huge blunder when they missed to look for forensic traces on the tin box, since the results of this investigation could have produced evidence favorable to the defendant. If fingerprints or DNA from one of AAs sons had been found on the box, AAs confession would have been falsified. AA is acquitted. The court says in its verdict that the police should have examined the tin box for fingerprints or DNA, and explains that the evidence against AA would have been sufficient for a conviction if the tin box had been properly examined and this had not produced any evidence against the prosecution's case, but since this examination is now missing from the investigation, the evidence against AA is not sufficiently robust for the standard of proof in criminal cases.

Example: Missing Fingers

Reference class used to compute random match probability of missing finger is too generic. A difference reference class might yield a different random match probability. This is also from Dahlman's paper:

> In 2013 a beheading video is spread on the internet. The video is made with a smartphone in Syria and shows in graphic detail how a British journalist is decapitated by ISIS. The face of the executor is masked but his hands are visible and two fingers are partly missing on his right hand. A couple of months later the Swedish police receives an anonymous tip from a woman who has seen the beheading video on the internet and says that she recognizes the hand. She believes that the executor is BB, a man of Syrian origin living in Sweden. The police investigate BB and find that he made two trips from Sweden to Syria in 2013 to support ISIS in its cause. BB admits that he has participated in ISIS activities in Syria, but denies that he is the executioner in the beheading video, and claims that he has never killed anyone. A forensic image analyst compares the hand in the video with BBs hand, and report that they match. The missing fingers are severed in the same places. To assess the probability of a random match, the forensic analyst consults reference data on the prevalence of missing fingers. Searching a data base with 20 000 pictures of hands collected from the general Swedish population the forensic analyst finds 20 hands (1 in 1000) with severed fingers. At closer scrutiny, two of them (1 in 10 000) are severed in the same way as the hand in the beheading video, and match it just as well as BBs hand. The forensic analyst therefore concludes that the probability of a random match is approximately 1 in 10 000. The two matching hands in the reference data base belong to men who died before 2013 and can therefore be ruled out as suspects. BB is prosecuted for murdering the British journalist. The case for the prosecution is based on BBs affiliation with ISIS and the expert testimony of the forensic image analyst. BBs defense attorney argues that the random match probability assigned by the forensic expert is too small, since it is based on the prevalence of missing fingers in the general Swedish population and it is reasonable to assume that such injuries are more common among men that are affiliated with ISIS. In the cross-examination of the forensic expert, the defense attorney asks if it is possible that the gathering of further reference data about people affiliated with ISIS could have shown that missing fingers are considerably more frequent in this reference class, for example that 1 in 1000 rather than 1 in 10 000 are disfigured in this way. The

> forensic expert replies that this possibility cannot be ruled out. BB is acquitted. The court says in the verdict that the prosecution should have backed their case with better reference data. The court explains that a random match probability of 1 in 10 000 would have been sufficient for proof beyond reasonable doubt, given the other circumstances of the case, if this probability had been robust, but in the absence of more reference data on people affiliated with ISIS it is not sufficiently robust for the standard of proof in criminal cases.

This seems a good case to illustrate Rafal's model of weight of evidence, since the evidence here is quantitative.

Example: DNA evidence assessment

Maybe Rafal's model of weight is useful to asessment DNA evidence which has different layers of uncertainties involved – TO BE COMPLETED