BACHELOR

Forensic statistics

the decision making process in court from a mathematical point of view

Pasch, E.J.M.

Award date:
2019

TU/E

Bachelor Final Project

Forensic Statistics:

The decision making process in court from a mathematical point of view.

Author:
Eline Pasch (0938501)

Supervisor:
Paulo De Andrade Serra
Second Reader:
Stella Kapodistria

July 2, 2019

Summary

Judges have a challenging job, to make sure that people are given a fair trail. In this bachelor final project, I asked myself to what extend mathematics could contribute to this. I studied literature on how Bayesian Networks could help to successfully connect the evidence to hypotheses about the guilt of a suspect. I described what shortcomes there are to this. I will go in depth on DNA data, which is quite often (mis)used in forensic statistics.

Contents

1 Introduction

In this chapter, we are going to take a look at the motivation behind this report. After this, we will establish the background needed in order to successfully set our goal .

1.1 Motivation

On the night of April 19, 1989, a female jogger was attacked and raped in New York's Central Park. Later, she was found unconscious, with a fractured skull and no memory of the assault. Initial police investigations, wrongly, focused on a group of 5 teenagers, who were in custody for a series of other attacks in the park that night.
Korey Wise was the oldest of the group, 16 years at the time, he was therefore the only one of the five tried as an adult.

There was no actual evidence against the boys, but they were convicted anyway. 11 years later, Matias Reyes, who was not one of the initial five, claimed to be guilty of this crime and evidence indeed confirmed his story. On that point Wise was released from prison after serving 11.5 years in prison for a crime he did not commit. (The Innocence Project, 2018)

Unfortunately, the story of Korey Wise is not the only one out there focusing on suspects who have been wrongly incarcerated. The Innocence Project, founded in 1992 by Peter Neufeld and Barry Scheck exonerates the wrongly convicted through DNA testing and reforms the criminal justice system to prevent future injustice. At their site, `www.innocenceproject.org`, they collect cases of wrongful convicted people and try to help them in their quest for innocence. (The Innocence Project, 2018)

When reading these stories the importance of making the right court decisions becomes very clear. Also the danger of drawing wrong conclusions because of a lack of evidence can be noted. Not just in the above discussed case of Korey Wise, but also in many other cases at their site. This is why a further investigation in this is highly important.

1.2 Background

We are going to introduce a few key concepts that will be useful for this report.

1.2.1 Forensic Science

Forensic science uses scientific principles to support or negate theories surrounding physical evidence found at a crime scene. As such, forensic scientists analyze evidence gathered or received from crime scenes and present their findings based the results of their analyses. Forensic science work generally involves one or more areas of science. Especially chemistry and biology are important to correctly handle the evidence found at the crime scene. They will, for instance, investigate chemicals and DNA samples. (EDU, n.d.)

For the interpretation of the evidence mathematics is valuable. Statistics, in particular, can be helpful for this task.

1.2.2 Forensic Statistics

Forensic Statistics, the use of statistical thinking and computations in court statements has grown, in the past 30 years. (Curran, 2009) The field of forensic statistics is very broad and dynamic, and therefore a interesting application of mathematics. Shoe prints, toxicology and handwriting tests are some examples that are often used in the court case. When evidence of this kind is found, it is important to compare the values of the suspect to the values of the evidence found at the crime scene.
In order to correctly interpret the evidence, a value corresponding to the strength or weakness of evidence towards supporting a given hypothesis. This is where mathematics comes in. The field

of Bayesian statistics could help out with contextualising the evidence in a logical way. (Sjerps, 2008)

With the use of statistics, comes the misuse of statistics, which is the case of forensic statistics might mean the conviction or acquittal of the wrong person. Therefore, it is important not only to understand the concept of forensic statistics, but also understand the possible errors that can occur. This will be done, looking at some world famous cases. Since recalculations of the numbers used in this cases might lead us to question some decisions made in the past. We can try and learn from those mistakes.

1.2.3 DNA Statistics

When talking about comparing values found at the crime scene with the values of a suspect, the most known example is the comparing of DNA fragments. This is studied in the field of DNA statistics, which is a sub field of forensic statistics.

1.3 Goal

Our goal is to find out how to correctly apply statistics in court judgements, meaning in such a way that wrong judgements are made are as infrequently as possible. Wrong judgements could either be the incarceration of an innocent person or the freedom of a guilty person.

In order to fulfill this goal, we have set some subgoals:

- I am going to review the literature that has been written on applying forensic statistics.
- I am going to look at mistakes made in the past when forensic statistics had been misused.
- I will look into DNA statistics, since the use of DNA evidence is a popular example of forensic statistics. We therefore want to understand the strength of this type of evidence.

1.4 Outline

In chapter 2, the mathematical background of forensic statistics will be discussed. This will mainly focus on the relevance of Bayesian statistics to this subject. Some tools to apply this to the forensic field will be treated. All in order to conclude a roadmap to apply mathematics to court cases.

After that, in chapter 3, a look will be taken at cases in which such a roadmap is used in famous crime cases. Both, successful and unsuccessful applications are used to give insight on the usefulness of forensic statistics. In this chapter we will also discuss some commonly made error that a judge should be aware of when applying statistics.

In chapter 4 we will dive into DNA statistics. First, we provide insight into the biological background. After that I am going to use the model to highlight how challenging it is to use DNA to identify ethnicity.

Finally, in chapter 5, I look at the conclusions we can draw from the theoretical background and model. We discuss some advantages and disadvantages of using mathematics in law. Also, we propose possible further investigations on the subject of mathematics in law.

In the appendix, chapter 6, you can find the code for the implementation of the model used in chapter 4 and the excel file belonging to it.

2 Theoretical Background

2.1 Definition of a Court Case

In the report, we often talk about court cases. There are many legal cases that could be referred to as court cases, therefore I present the definition of court case that will be used in the report. There are two main cases that are presented to court. Which are criminal cases and civil cases.

Something is called a *civil case* when one person, business, or agency sues another one because of a dispute between them, usually involving money. If someone loses a civil case, then they may be ordered to pay the other side money or to give up property, but they will not go to jail just for losing the case. It might also be possible to use forensic investigations in some of these cases, but it is not as usual. In a civil case, the plaintiff must prove his or her case by a "preponderance of the evidence" (more than 50 percent). This means that a party to a civil case can win if he or she is able to convince the judge or jury that his or her side of the case is slightly more convincing than the other side's. (The Judicial Branch of California, n.d.)

A *criminal case* happens when the government files a case in court to punish someone (the defendant) for committing a crime. If the defendant is found guilty of a crime, he or she may face jail. In criminal court, you are presumed innocent until proven guilty beyond a reasonable doubt.
A felony is the most serious kind of crime. Examples of felonies are: Robbery, Murder, Rape, Possession of illegal drugs (called "controlled substances") for sale. (The Judicial Branch of California, n.d.)

We will mainly focus on felonies in this report, since those use forensic statistics more.

It should be noted that all of these definitions are stated by the judicial branch of California and are therefore specific for courts in this area. We will use them as general definitions, since they seem like a good representation.

2.2 Setting Hypotheses

First, we look at a general way of stating hypotheses in legal cases.

In all court trails, the jury starts by assuming that the defendant is innocent. That is why, in most cases, the *null hypothesis*, H_0, is set to be the innocence of the suspect. It will then be the task of the prosecutor to show evidence that the defendant is not innocent, in other words, they will try to show that the null hypothesis is unlikely to be true. The contradicting statement they will try to prove is also called the *alternative hypothesis*, and will be noted as H_1.

This is also why a person would never found to be innocent, since this innocence was already assumed. The jury can only conclude that there was not enough evidence to reject that hypotheses, or find enough reason to reject the null hypothesis; decide the suspect is guilty. (Privitera, 2015).

Even tough these are the theoretical hypotheses belonging to the courts decision of assuming innocence, in this report we will sometimes follow examples where from literature in which they decided to turn H_0 and H_1 around, since it would make the way of reasoning easier.

2.3 Testing Hypotheses

To test if we should reject the given H_0 we have to find a value that expresses how much the evidence support this H_0. An often used way relies on Bayes' theorem. We build up to the theorem and explain how the theorem is interpreted in court cases.

2.3.1 Likelihood Ratio

We now build up to a model that relates evidence to each of the hypothesis. In order to do so, we still consider H_0 and H_1 as they are defined before, but we now question how we should evaluate the evidence (E) present. To do so, we introduce the likelihood ratio (V).

$$V = \frac{\mathbb{P}(E|H_0)}{\mathbb{P}(E|H_1)}$$

To determine the value for the likelihood ratio we find that the probability of a certain piece of evidence being found when the suspect is not guilty is divided by the probability that the same evidence is found when the suspect is not guilty. Therefore we interpret the likelihood as how much more likely it is to find a certain piece of evidence given that a person is innocent, than finding the same piece of evidence when the person is innocent. (Gastwirth, 2000)

Therefore, if we would find a likelihood of 1, we could conclude that finding a certain piece of evidence is just as likely in the case of the guilt as in the case of no guilt of the suspect (*neutral evidence*). When we would find a value bigger than 1, we find that the piece of evidence would make it more likely that the suspect is innocent (*supporting evidence*). And when the value is smaller than 1, the piece of evidence would make it seem more likely that the suspect is guilty (*attacking evidence*). (Vlek et al., 2016)

So, in conclusion, we could see the likelihood as a measure for the strength of a certain piece of evidence.

2.3.2 Prior and Posterior Probabilities

We have now seen how we could calculate a value for the evidence, but a value of a single piece of evidence does not give enough information on the guilt or innocence of a suspect. It is therefore important to compare the effect of the piece of evidence on the initial probabilities, also know as the *prior odds*. This is where Bayes's theorem could be used. Which in general could be written as

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)},$$

if $\mathbb{P}(B) > 0$.

So, the following holds, $\mathbb{P}(H_0|E) = \frac{\mathbb{P}(E|H_0)\cdot\mathbb{P}(H_0)}{\mathbb{P}(E)}$ and $\mathbb{P}(H_1|E) = \frac{\mathbb{P}(E|H_1)\cdot\mathbb{P}(H_1)}{\mathbb{P}(E)}$.

In all cases $\mathbb{P}(E)$ will be greater than 0. Since it makes no sense to make a model without any evidence present.

The H_0 represents the hypotheses that claims that the suspect is innocent and H_1 is the alternative hypothesis . Then $\mathbb{P}(H_0)$ represent the prior probability, so the probability of the suspect being innocent, based on nothing given. So, based on no evidence or other events known. Therefore $\mathbb{P}(H_0) = 0.5$ is often a logical choice. Then $\mathbb{P}(H_1) = 0.5$ follows.

Using these equalities to find a ratio between $\mathbb{P}(H_0|E)$ and $\mathbb{P}(H_0|E)$ gives us $\frac{\mathbb{P}(H_0|E)}{\mathbb{P}(H_1|E)} = \frac{\mathbb{P}(E|H_0)\cdot\mathbb{P}(H_0)\cdot\mathbb{P}(E)}{\mathbb{P}(E|H_1)\cdot\mathbb{P}(H_1)\cdot\mathbb{P}(E)}$.

So overall,we can conclude:

$$\frac{\mathbb{P}(H_0|E)}{\mathbb{P}(H_1|E)} = \frac{\mathbb{P}(E|H_0)}{\mathbb{P}(E|H_1)} \cdot \frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)}.$$

We recognize that this is equal to:

$$\frac{\mathbb{P}(H_0|E)}{\mathbb{P}(H_1|E)} = V \cdot \frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)}.$$

Which is an easy rewrite to understand the ratio between the *posterior odds*. With the posterior odds we mean the probabilities of a certain hypotheses being correct after investigating the evidence.

This is what we are interested in. We can interpreted a ratio between $\mathbb{P}(H_0|E)$ and $\mathbb{P}(H_1|E)$ as how much more likely it is that the suspect is innocent compared to the suspect being guilty, given the evidence provided. (Gastwirth, 2000)

So from the Bayes' Theorem we can conclude that the ratio of posterior odds can be found by multiplying the ratio of prior odds with the likelihood ratio.

2.3.3 Bayesian Network

Now, we have to find a way to link the evidence with the hypotheses, since in court cases the pieces of evidence should be combined to make a rightful decision. A Bayesian networks is such a model based on Bayes' Theorem that could help us providing such a framework. We explain how to compute such a network.

A Bayesian Network consist of a directed acyclic graph and conditional probability tables for each node in the graph. First, we are going to take a look at the build-up of the directed acyclic graph, after that we take a look at the determination of the values of the conditional probability table. (Vlek et al., 2016)

Definitions related to Graphs
Some definitions with respects to graphs are given.

Definition: Graph
A graph G is a pair G =(V,E) consisting of a finite set $V \neq \emptyset$ and a set E of two-element subsets of V. The elements of V are called vertices or *nodes*. An element $e = \{a, b\}$ of E is called an *edge* or link with vertices a and b. (Cohen et al., 2000)

Definition: Directed Graph
Formally, a directed graph or, for short, a digraph is a pair G =(V,E) consisting of a finite set V and a set E of ordered pairs (a, b), where $a \neq b$ are elements of V. (Cohen et al., 2000)

Definition: Walk, Path, Trail and Cycle
If there are vertices $v_0, ..., v_n$ such that $e_i = v_{i1}, v_i \in E$ for $i = 1, ..., n$, the sequence is called a *walk*; a walk for which the e_i are distinct is called a *trail*. If, in addition, the v_j are distinct, the trail is a *path*. A closed trail with $n \geq 3$, for which the v_j are distinct (except, $v_0 = v_n$), is called a *cycle* (Cohen et al., 2000)

Definition: Acyclic Graph
A graph is called acyclic if it does not contain a cycle. For a subset T of the vertex set V of a graph G we denote by G \T the induced subgraph on V \T. This graph arises from G by omitting all vertices in T and all edges incident with these vertices. (Cohen et al., 2000)

We know what a directed acyclic graph is and we will further refer to it as a DAG.

Inner Structure of the Bayesian Networks
With this knowledge of graphs, we can look into the inner structure of the Bayesian Network that we are willing to create. In order to create a Bayesian network representing the relation between hypotheses and evidence, like a network belonging to a court case, it is important to make a list of observed variables, which include the evidence belonging to the case. These represent the nodes of the graph we want to create. The evidence will represented as Boolean statements. For example "the DNA found at the crime scene matches the DNA of the suspect". This is either true or false, if it is true, the guilt of the suspect is more likely.

There are also some linking nodes needed. To link evidence to the node of the *null hypothesis*. For example the fact that a blood stain found at the crime scene belongs to the suspect could

be the variable belonging to a node. These linking nodes are sub hypothesis, which are Boolean statements whose true value is not known by the court.

Already, a clear link can be found between the two statements given above. The fact that the DNA evidence would match the DNA of the suspect (D), will make it more likely that the DNA of the suspect has indeed been found at the crime scene (E). What would make is more likely that the suspect was indeed guilty. So, if we would compute a path summarizing these nodes, we would get a path from our null hypothesis (H_0), to D, ending in E. (Fenton et al., 2010)

So the sub hypotheses and evidence nodes, will give possible explanations or scenarios. If two variables are dependent on each other, there exist an active path between them, through which an influence can flow. In other words, following a path would lead to specifications of a story line.

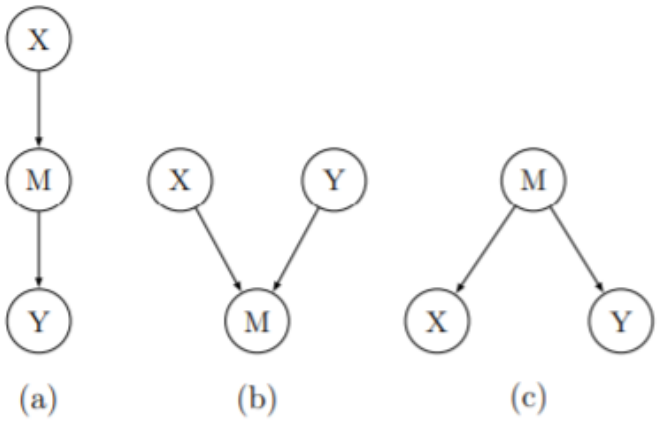


Figure 1: Structures in Trails Occurring in a Bayesian Network (Ibs, 2016)

The different substructures that can be found in a directed acyclic graph are shown in figure 1 to show dependency. An explanation of this influence and their interpretation with respect to evidence will follow. M is a node belonging to the list of observed variables, the evidence nodes, and X and Y represent nodes that do belong to the hypothesis.

a) Causal structure. In this case only 1 possible path is given. In other words, Y is directly dependent on X.

b) Common effect structure. One observation would be logical as result from 2 different events. From this structure we find that there is no path possible from X to Y (or Y to X), since M is factual and gives two independent possible explanations. Since there is no path from X to Y we can say that there is is no influence between X and Y and therefore two independent paths occur. In conclusion, following those two paths would lead to different scenarios that explain M

c) Common cause structure. One observation would have two logical explanations. Again no path between X and Y can be found, so the explanations should be independent from each other.

(Ibs, 2016)

In the example cases of O.J. Simpson that can be found in chapter 3, in figure 2, a clear example can be found of such a graph summarizing possible stories linking the evidence that was found. The above discussed structures are clearly visible, but we have to note that these structures are just the basics, expansions of b) and c) are of course possible. For example, one possible observation can also lead to more than 2 explanations.

Note, that a node never occurs twice in the network. Therefore, it should always be linked in the way just described.

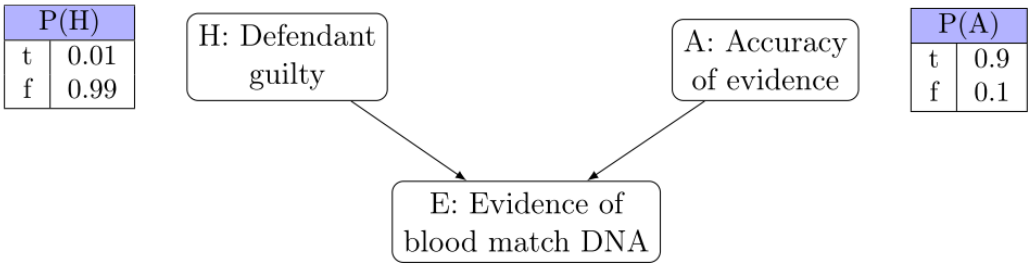
The Complete Bayesian Network
After creating patterns from different hypotheses and pieces of evidences, it is important to link them together. The first node is always the one stating the null hypothesis. From there, the evidence and underlying hypotheses are added. So there forms a DAG with paths from the hypothesis into different branches of evidence.

This direction of the structure makes senses because the the suspect's guilt increases the probability of finding incriminating evidence and also the innocence of the victim decreases the probability of finding this evidence. (Fenton et al., 2010)

Conclusion
So in conclusion, first a list of observations belonging to the investigations should be made and then possible explanations should be found to link those observations. In order to do so, a look should be taken at the dependency of these explanations and observations.

Conditional Probability Table
Conditional Probability Tables, also known as Node Probability Tables, summarize the prior beliefs for each node. For each node in the Bayesian Network a conditional probability table is added. In this table, the conditional dependencies of the variables are collected, specifying the strength of influence the nodes have on each other. At first, the prior probabilities of all nodes should be determined. This will be done by looking at the *parental nodes*, which are the nodes that have incoming information. (Ibs, 2016)

Example
We follow an example given by Ibs (2016). This example represents a simple network where a hypothesis is done and evidence is added, just in a way as we described before. Also the author added the node called "Accuracy". With this the author refers to how likely it is that the test gives the value "true" when the evidence is indeed a match. It might be strange to add a parental node that does not directly link to the hypothesis, but it could be explained easily because the accuracy of the evidence is not a consequence or deepening of the evidence node.



P(E | H, A)				
H	t		f	
A	t	f	t	f
t	1.0	0.5	10^{-6}	0.5
f	0	0.5	$1 - 10^{-6}$	0.5

Figure 2: A simple Bayesian network (Ibs, 2016)

In the tables we see that A and H do not have any parent nodes. Therefore we could give those chances by making assumptions based only on how likely it is that the given event is happening. Note that this author decided that it was way more likely that the defendant was innocent than that he/she was guilty. More often these probabilities will be set to both be 0.5, since you do not want to add this bias to the network. (Ibs, 2016)

The values of the conditional probability table belonging to the node E is a bit more interesting. Since the Boolean statement belonging to the parental nodes have 2 possible values, and the evidence node also has two possible values, there are $2^3 = 8$ values that should be determined.

We for example have to look at the combination that the defendant is guilty and that the evidence is accurate. Given this combination, how likely is it that we find a positive blood match between suspect and crime scene? This is of course 1, since we would then indeed check the DNA of the evidence with that of the suspect. Note that then automatically follows a value of $1 - 1 = 0$ for the case that the evidence would not match, since the sum of the probabilities of these values should be 1.

All other remaining values should be reasoned in a similar way.

From Bayesian Network to Likelihood Ratio
The network created now shows how likely certain events are, with respect to their parental nodes. We should now use this information to determine how likely H_0 given the entire network.
In order to do so, we have to start at the ends of the network, and work back from there. This gives new values.

Theoretical Explanation
We are interested in the probability $\mathbb{P}(H|E)$, which is the probability of our hypothesis being true, given that the evidence found is true. According to the Bayes theorem we could calculate this with the following formula:

$$\mathbb{P}(H|E) = \frac{\mathbb{P}(E|H) \cdot \mathbb{P}(H)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|H) \cdot \mathbb{P}(H)}{\mathbb{P}(E|H) \cdot \mathbb{P}(H) + \mathbb{P}(E|notH) \cdot \mathbb{P}(notH)}$$

(Fenton et al., 2010)

Example
We are interested in the probability that the hypothesis of the network in figure 2 is true, given that the evidence is true. We are not going to assume anything about the accuracy of the evidence. We consider the formula given before and the other rules from probability theory to rewrite Bayes theorem and calculate our probability.

$$\mathbb{P}(H = t|E) = \frac{\sum_A \mathbb{P}(E|H = t, A) \cdot \mathbb{P}(H = t)}{\mathbb{P}(E)}$$

Were we sum over the option of A being true and A being false.

$$= \frac{\sum_A \mathbb{P}(E|H = t, A) \cdot \mathbb{P}(H = t)}{\sum_A \sum_H \mathbb{P}(E|A, H) \cdot \mathbb{P}(H) \cdot \mathbb{P}(A)}$$

now, using the numbers from the conditional probability table, we get

$$= \frac{(1 \cdot 0.9 + 0.5 \cdot 0.1) \cdot 0.01}{1 \cdot 0.9 + 0.5 \cdot 0.1) \cdot 0.01 + (10^{-6} \cdot 0.9 + 0.5 \cdot 0.1) \cdot 0.99} \approx 0.161$$

That means, our updated, posterior probability is 0.161, while the prior probability was 0.01. So, we can conclude that it is more likely that the defendant is guilty than that it was before updating the probabilities using the evidence. (Ibs, 2016)

We know that the hypothesis is either false or true, so we can conclude

$$\mathbb{P}(H = t|E) + \mathbb{P}(H = f|E) = 1$$

From this we find $\mathbb{P}(H = f|E) \approx 0.839$. (Ibs, 2016)

Now we could use the formula found at 2.3.2 to compute the likelihood of the evidence.

$$\frac{\mathbb{P}(H_0|E)}{\mathbb{P}(H_1|E)} = V \cdot \frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)}$$

Using the values found we get $\frac{0.161}{0.839} \approx V \cdot \frac{0.01}{0.99}$. So $V \approx \frac{0.161 \cdot 0.99}{0.839 \cdot 0.01} \approx 19$.

In the next sub section, we will interpret this value.

Bayesian Network Software
For a simple structure, like the one just presented, calculations can be done by hand. Fortunately, there is also software developed specifically to do those calculations with Bayesian networks. Two possible option for Bayesian network software can be found at `https://www.agenarisk.com/` or `https://www.hugin.com/`. (Ibs, 2016)

2.4 Drawing a Conclusion

From the value of the likelihood we should draw a conclusion. In order to do so, we first look into the effect that possible mistakes will have.

2.4.1 Type I and Type II Error

When making a decision about the hypothesis, we could distinguish four cases:

1. The decision not to reject the null hypothesis could be correct.
2. The decision not to reject the null hypothesis could be incorrect. (Type II error)
3. The decision to reject the null hypothesis could be correct. (Type I error)
4. The decision to reject the null hypothesis could be incorrect.

(Privitera, 2015)

We note two different types mistakes that could be made. The incorrect decision not to reject a false null hypothesis, which is referred to as a Type II error. Or the incorrect decision to reject a true null hypothesis. This decision is an example of a Type I error. (Privitera, 2015)

If we would translate this to court cases, we distinguish cases of rightly and wrongly acquitting or convicting a victim. From this we note that there are two types of possible errors. Where wrongly incarcerating an innocent person would be considered a *type I error* and letting a guilty person go free is considered a *type II error*, when we would set our hypothesis as explained in subsection 2.2. (Privitera, 2015)

We now look at the connection between type I and type II error.

Connection Type I and Type II Error
Since we assume the null hypothesis is true, we want to control the type I errors. In other words, we want to make sure that the number of innocent people getting convicted is controlled.

At first, we want to calculate the probability of obtaining a sample outcome, given that the

value stated in the null hypothesis is true. This probability is called the *p value*. This p value is compared to the level of significance. The level, is called the *alpha level* (symbolized as α) and will be stated before an experiment. It is the largest probability we will allow to reject the null hypothesis and therefore the largest probability of committing a Type I error. (Privitera, 2015).

So, in summary, if the p-value is smaller than α, we reject the null hypothesis. So, then we could conclude that the suspect being innocent, with all the evidence given, is unlikely.

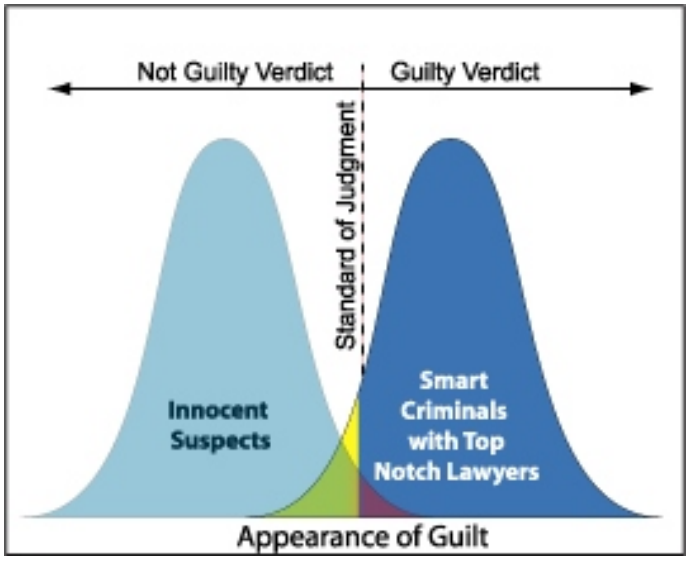


Figure 3: Error (Chl, 2013)

Figure 3 helps to understand the connection between type I and type II errors. The underlying hypothesis differ from what we mentioned before. Namely, this example uses:

H_0: The suspect is guilty.
H_1: The suspect is not guilty.

The line with the text "standard of judgements" draws a line at the point with value α on the x-axis. So for all p-values on the x-axis, before the line, H_0 will be rejected. In other words, the suspect will determined to be not guilty, since the p value is too small to accept the null hypotheses. For the p-values bigger than α, the suspect will be determined to be guilty.
In the case of the blue planes, this conclusion is correct. But there is also a possibility that the p-value given is smaller then α, while the suspect is guilty, this is represented by the yellow plane. The purple plane represents So the purple plane shows the type 1 error, while the yellow plane represents a type 2 error.

Decreasing the value of α, so shifting the "standard of judgement" to the left, would make sure that almost all criminals would get punished, but it will also increase the purple plane. So, more innocent people would get incarcerated.
A similar reasoning could be followed to conclude that increasing the value of would make sure almost no innocent suspects would get punished, but quite a lot of guilty suspect would not get free. (Chl, 2013)

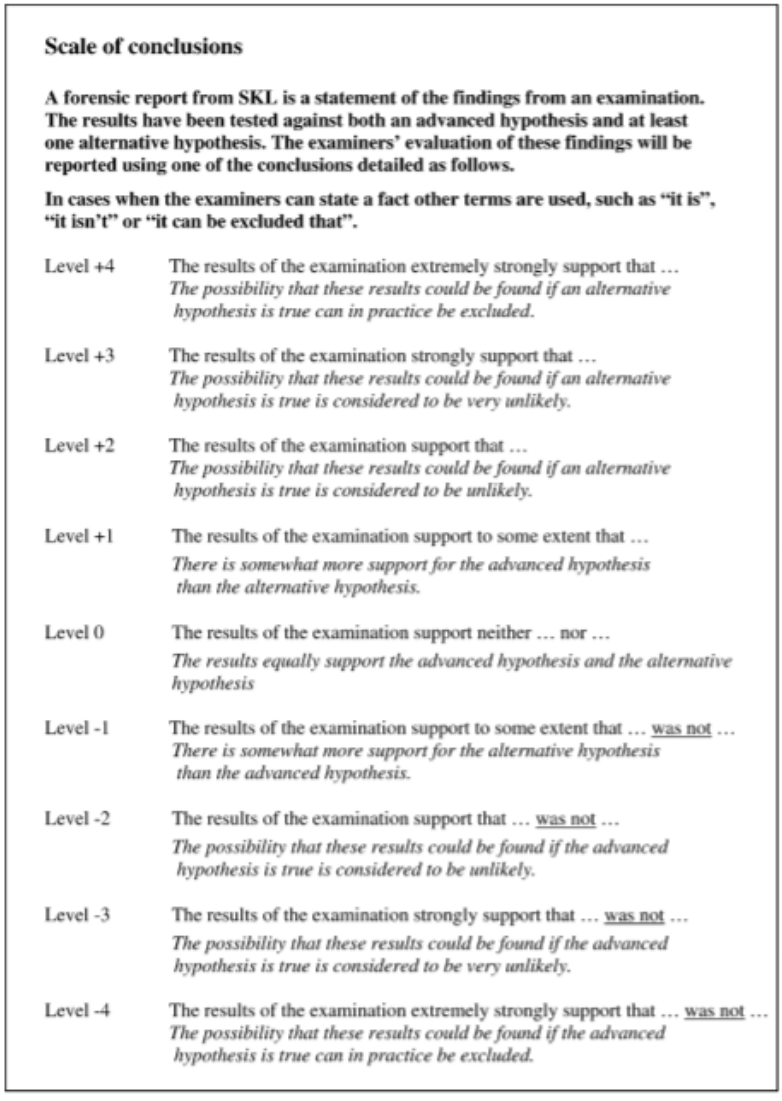
Since, we do not get such a p-value from the Bayesian network, we could not compare a p-value directly to α. But we do understand the importance of drawing a right full conclusion, in which we should also be aware of possible error and the effect those have on each other.

2.4.2 Drawing a Conclusion From the Likelihood Ratio

The Swedish National Laboratory of Forensic Science (SKL) provides a framework to translating likelihood ratios into court statements. We are going to follow their reasoning and calculations behind this framework, as provided by Nordgaard et al. (2012).

Scale of Conclusions

Forensic interpretations done at the SKL are made according to the scale of conclusions they produced. The scale has nine levels, numbered from -4 till 4. The positive numbers represent the cases where the null hypothesis is more likely than the alternative hypothesis, while the negative numbers represent situations where the alternative hypothesis is more likely. The level numbered 0 is used to represent that the null hypothesis and alternative hypothesis are both equally likely. The levels come with textual explanations as well.



Scale of conclusions

A forensic report from SKL is a statement of the findings from an examination. The results have been tested against both an advanced hypothesis and at least one alternative hypothesis. The examiners' evaluation of these findings will be reported using one of the conclusions detailed as follows.

In cases when the examiners can state a fact other terms are used, such as "it is", "it isn't" or "it can be excluded that".

Level +4	The results of the examination extremely strongly support that … *The possibility that these results could be found if an alternative hypothesis is true can in practice be excluded.*
Level +3	The results of the examination strongly support that … *The possibility that these results could be found if an alternative hypothesis is true is considered to be very unlikely.*
Level +2	The results of the examination support that … *The possibility that these results could be found if an alternative hypothesis is true is considered to be unlikely.*
Level +1	The results of the examination support to some extent that … *There is somewhat more support for the advanced hypothesis than the alternative hypothesis.*
Level 0	The results of the examination support neither … nor … *The results equally support the advanced hypothesis and the alternative hypothesis*
Level -1	The results of the examination support to some extent that … was not … *There is somewhat more support for the alternative hypothesis than the advanced hypothesis.*
Level -2	The results of the examination support that … was not … *The possibility that these results could be found if the advanced hypothesis is true is considered to be unlikely.*
Level -3	The results of the examination strongly support that … was not … *The possibility that these results could be found if the advanced hypothesis is true is considered to be very unlikely.*
Level -4	The results of the examination extremely strongly support that … was not … *The possibility that these results could be found if the advanced hypothesis is true can in practice be excluded.*

Figure 4: The english version of the scale of conclusions as used by the SKL (Nordgaard et al., 2012)

Likelihood Ratios to Scale

So, now we have to find logical values of likelihood ratios belonging to each scale. This is where Bayes theorem and uncertainty come together. In the scale represented we can note inverse statements. Therefore, we can also assume that the values of the likelihood ratios representing to the

scales are also inverse. This is why we will first try to calculate the values of the likelihood ratios belonging to the positive numbers.
In other words, we want to find values for the boundaries in the following table, and from this we want to complete the entire table.

Interval	Scale level
$1 \leq V < R_1$	0
$R_1 \leq V < R_2$	1
$R_2 \leq V < R_3$	2
$R_3 \leq V < R_4$	3
$R_4 \leq V < \infty$	4

Table 1: Outline Scale Levels Belonging to Likelihood Ratios (Nordgaard et al., 2012)

We are going to start with the formula from section 2.3.2, to produce an estimation of the value of $\mathbb{P}(H_0|E)$ in terms of V.

$$\frac{\mathbb{P}(H_0|E)}{\mathbb{P}(H_1|E)} = V \cdot \frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)}$$

Since $\mathbb{P}(H_0|E) + \mathbb{P}(H_1|E) = 1$, we can rewrite this to:

$$\frac{\mathbb{P}(H_0|E)}{1 - \mathbb{P}(H_0|E)} = V \cdot \frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)}$$

Since we mostly assume that the prior odds of the null hypothesis and alternative hypothesis are equal, we assume $\frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)} = 1$. So:

$$\frac{\mathbb{P}(H_0|E)}{1 - \mathbb{P}(H_0|E)} = V$$

From this, we deduct:

$$\mathbb{P}(H_0|E) = \frac{V}{V+1}$$

From this we can find a value of V, for which a certain probability for the posterior odd for H_0 follows. So, we have to set some values to calculate intervals.

If this posterior probability to be at least 0.99, the lowest value for V is about 100. We could say that this probability is being quite high. So it is really unlikely that the alternative hypothesis is true. Therefore we determine 100 as the lowest limit for lever 2. In other words: $R_2 = 100$.

Naturally, the upper limit of the level +4 is infinity. We fix a prosterior odd of 0.9999999 as lower limit belonging to this. Since this value not gives us almost no doubt. The V value that belong to this is 10^6. So $R_4 = 10^6$.

Since we want to have proportional intervals per scale, we want to solve:

$$log_{10}R_i - log_{10}R_{i-1} \approx k \cdot i; i = 2, 3, 4$$

Were we chose log_{10} because $1 * 10^2 = R_2$.

From this we find $k \approx 0.5$, $R_1 \approx 5.625$, $R_3 \approx 5625$ as a solution. Since these values are quite impractical from a numerical point of view, we round them up. (Nordgaard et al., 2012)

Earlier,we concluded that the scales were inverses, so $R_1 = \frac{1}{R_{-1}}$.
We now have all values to compute the following table:

Interval	Scale level
$V \leq 10^{-6}$	-4
$10^{-6} < V \leq 1/6000$	-3
$1/6000 < V \leq 1/100$	-2
$1/100 < V \leq 1/6$	-1
$1/6 \leq V < 6$	0
$1/6 \leq V < 6$	1
$6 \leq V < 6000$	2
$6000 \leq V < 10^6$	3
$10^6 \leq V$	4

Table 2: Scale Levels Belonging to Likelihood Ratios (Nordgaard et al., 2012)

So, after calculating the likelihood ratio, we are able to find the scale level in the table, in the "scale of conclusion" we can find the matching court decision.

2.4.3 Conclusion

Example
We again refer back to the example of Ibs (2016). We found that the likelihood belonging to the evidence of the case was about 19. Which would place the evidence on scale +1, under the assumption that the prior odds is 1. Which would, according to the scale of conclusions, be described as, that there is somewhat more support for the null hypothesis than for the alternative hypothesis.

2.5 Summary

We now finished a section on how Bayesian networks are produced with respect to law. Following these steps should help by providing a mathematical based conclusion from the evidence provided in a crime scene:

1. State the null hypothesis and alternative hypothesis.
2. List all evidence that could be find to support the hypotheses; write them as Boolean statements.
3. Produce a directed acyclic graph with the evidence. Add nodes with sub hypotheses, to make sure the paths in the graph are structured.
4. Make a conditional probability table per node.
5. Calculate the likelihood ratio belonging to the network.
6. Draw a conclusion; decide whether the null hypothesis or alternative hypothesis is more likely.

3 Cases

3.1 Introduction

We are interested in how to apply the theoretical approach just discussed. We will do this by discussion some cases. Also some known cases in which statistics was missused will be discussed. From this we hope to gain some knowledge about the advantages and disadvantages of the use of the Bayesian rule in forensic statistics. Hopefully, this will lead to a better understanding on how to apply the steps discussed in the legislative field.

3.2 O.J. Simpson

To study the case of O.J. Simpson the work of Olumide (2010) is used. He performed a probabilistic analysis of the evidence provided using Bayesian networks.

3.2.1 Introduction to the case

Orenthal James Simpson was suspected of the murder of his ex-wife, Nicole Brown Simpson, and another man, Ronals L. Goldman. They were found dead in Nicole's yard on the evening of June 12^{th}, 1994, with a slash of the throat and various stab wounds in other body parts.

3.2.2 Trail and Hypothesis

The case has been tried by the State of California, which defines murder as the unlawful killing of a human being or a fetus with malice aforethought. The states considers a case "First degree Murder" if the murder was premeditated. They therefore attempted to prove that:
Orenthal James Simpson was guilty of first degree murder in the deaths of Nicole Brown Simpson and Ronald L. Goldman that took place on June 12th, 1994, in Brentwood, California.
This will be considered our null hypothesis (H_0).

Simpson, on the contrary, claimed that he was not the person who committed the murders.Which is the complement of our hypothesis (H_1). (Olumide, 2010)

This contradicts the theoretical approach provided in chapter 1. The author probably made this decision, because this way connecting the facts to the hypothesis is possible. Creating story-lines assuming a persons innocence is not doable, since it would not lead to any logical explanations. This switch of H_0 and H_1 will be seen in more cases.

3.2.3 Evidence

Of all the evidence gathered for the case, the relevant evidence is divided in six different segments. Those segments will be address differently and later combined to find a aggregate likelihood.

3.2.4 Interpreting Evidence

Bayes theorem will be applied to find which of the two hypotheses is more likely given the evidence. In fact, a Bayesian network was constructed for the collection of items of evidence in all six different segments. Since the calculation of the likelihood of all segments have a similar structure, only one of them will be reflected in this report, namely the likelihood belonging to the blood strains found at the crime scene, the O.J.'s Rockingham residence and his Ford Bronco. In order to make a Bayesian network belonging to the blood stains, some insight in the DNA found at the crime scene is needed. This can be found in an article written by Thompson (1996), which focuses only on the DNA evidence present in the Simpson case and therefore was an important source in the Olumide (2010) article.
Even though the network is based on the factual DNA evidence, it is important to note that the linkage patterns and conditional probabilities could be varied within personal choices of the importance of pieces evidence. In other words, the Bayesian network created in this article is not the only possible Bayesian network and therefore not necessarily right or wrong.

3.2.5 List of Propositions

The list of propositions will represent the nodes in the Bayesian Network.

H_5	O.J. committed the two murders
A	O.J.'s blood was found at the murder scene
B	The blood of both victims and O.J.'s blood was found on his Ford Bronco
C	The blood of both victims and O.J.'s blood was found at O.J.'s Rockingham residence
D	Blood stains found on the back gate of NBS house matched O.J.'s DNA profile
D^*	Lab results supporting D
E	Bloody shoe prints found at the crime scene matched O.J.'s DNA profile
F	Bloody shoe prints at the crime scene came from Bruno Magli shoes
F^*	Expert testimony
G	O.J. owns a pair of Bruno Magli shoes
G^*	Prosecution testimony
H	Blood drops to the left of shoe prints found at crime scene matched O.J.'s DNA profile
H^*	Lab results supporting H
I	O.J. had a cut on his left finger
I^*	Detective Vannatter's testimony
J	Blood on the driver side of the exterior door handle of the Bronco matched O.J.'s DNA profile
J^*	Lab results supporting J
K	Blood on the driver's side of the interior door handle of the Bronco matched O.J.'s DNA profile
K^*	Lab results supporting K
L	Blood on the driver side carpet of the Bronco matched O.J.'s DNA profile
L^*	Lab results supporting L
M	Blood on the steering wheel of the Bronco matched both O.J.'s and Nicole's DNA profiles
M^*	Lab results supporting M
N	Blood on the console of the Bronco matched both victims' and O.J.'s DNA profiles
N^*	Lab results supporting N
O	The blood trail on the ground leading to O.J.'s Rockingham residence matched O.J.'s DNA profile
O^*	Lab results supporting O
P	Blood on the gate of O.J.'s Rockingham residence matched O.J.'s DNA profile
P^*	Lab results supporting P
Q	Blood on the floor of the master bathroom of O.J.'s Rockingham residence matched O.J.'s DNA profile
Q^*	Lab results supporting Q
R	Blood in the foyer of O.J.'s Rockingham residence matched O.J.'s DNA profile
R^*	Lab results supporting R
S	Blood found on the sock found in O.J.'s master bedroom matched both Nicole's and O.J.'s DNA profiles
S^*	Lab results supporting S
T	Blood on the right hand glove found at O.J.'s Rockingham residence matched both victims' and O.J.'s DNA profiles
T^*	Lab results supporting T
U	Someone with O.J.'s blood type bled in O.J.'s Ford Bronco
V	Blood matching both victims' blood types was found in O.J.'s Ford Bronco
W	Someone with O.J.'s blood type bled at O.J.'s Rockingham residence
X	Blood matching both victims' blood types was found at O.J.'s Rockingham residence

Table 3: Nodes used in the Bayesian Network of O.J. Simpons' case (Olumide, 2010)

3.2.6 Bayesian Network

The author constructed the network in figure 5. The 3 paths from H_5 distinguish the different locations that are important to the crime. From there, different paths are created to explain the presence of the DNA on these locations. Later, the probabilities belonging to how likely it is that these explanations are legit will be determined. To determine this, not only the evidence found is important, but also the reliability of the evidence should be taken into account. Therefore, the red nodes represent the statistics from lab results and testimonies.
All in all this network test whether this evidence would be the way it is, assuming that O.J. committed the two murders, in other words H_5 is true.

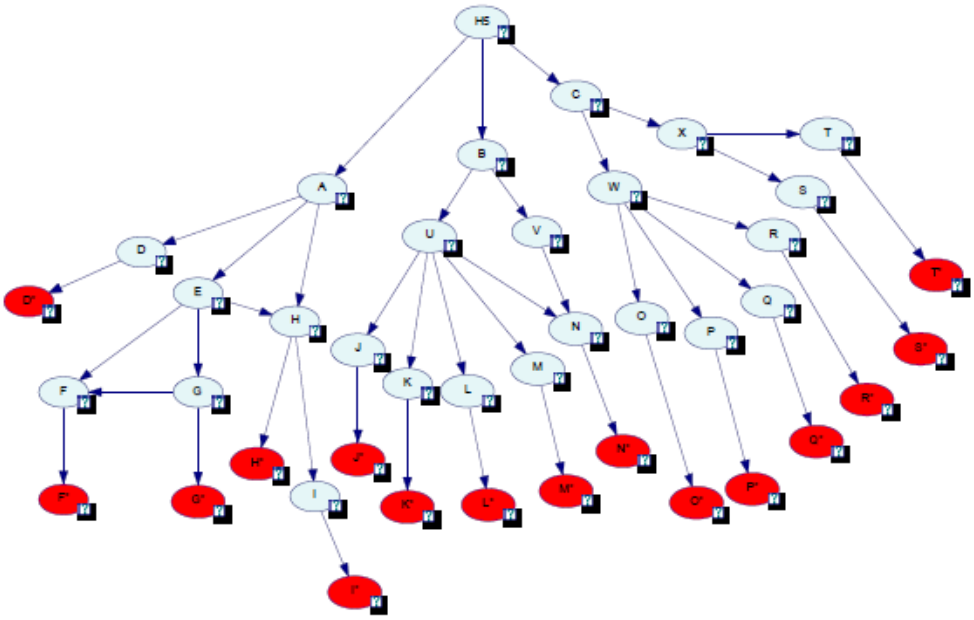


Figure 5: Bayesian Network of blood stain evidence (Olumide, 2010)

3.2.7 Setting Probabilities

As said before, the determination of the probabilities is not based only of the DNA samples found, but also on the interpretation of this evidence by the investigator. Therefore, it is important to understand the underlying choices made, in order to create a Bayesian network ourselves.

$\mathbb{P}(H_0)$ and $\mathbb{P}(H_1)$ are the only non conditional probabilities, since this is the starting point of the Bayesian network. Both H_0 and H_1 should be equally likely to occur, since there should not be assumed that the suspect is guilty, nor that he is not. Therefore the author chose $\mathbb{P}(H_0)$=$\mathbb{P}(H_1)$=0.5. Note that in the article refers to H_0 as H_5 and H_1 as H_5^c.

Since, in this case, DNA evidence is really important to check H_5 and H_5^c , the DNA was tested by three companies; the Los Angeles Police Department, the California Department of Justice DNA Laboratory and a private company called Cellmark Diagnostics.
From the DNA profiles found, the conclusion has been drawn that about 7 percent of the population has a DNA profile similar to O.J. Simpson's, 5 percent of the DNA matches with that of Ron Goldman and 3 percent has the same DNA profile as Nicole Simpson. These statistics come in handy when determining the probability that not O.J. Simpson, but another person with similar

blood profile, committed the crime.

This percentage of 0.07 % is for example used for $\mathbb{P}(U|B^c)$, which represents the probability that a person with a similar DNA profile as O.J. Simpson bled on the Ford Bronco, given that O.J. did *not* bleed there.

Another interesting thing is $\mathbb{P}(U|B)$, which represents the probability that somebody of O.J's blood profile blood, given that O.J.'s blood was found. At first, you might think that this would be 1.00, since it would be logical that when Simpson's blood was found in the car, he should have bled there. However, in this kind of conditional statements, it is important to be very specific and also think about alternative explanations. In this case, for example, it is also possible that his blood was found, because somebody has placed it there to frame O.J.. The author of the article did not really believe in that and therefore set $\mathbb{P}(U|B)$ to be 0.98. So, this probability depends on the opinion of the producer of the network, there is not specific value that should have to be picked.

Another probability that uses the DNA results directly is $\mathbb{P}(V|B^c)$, which represents the probability that blood matching both victim's found given that it is not the blood belonging to those victims. Which is precisely the chance that blood of similar DNA profiles of both victims was found. So $\mathbb{P}(V|B^c) = 0.03 \cdot 0.05 = 0.0015$. On the other hand, $\mathbb{P}(V|B)$, the probability that blood matching both victim's was found given that it is the blood belonging to those victims was found, would be 1,00. Since there is no possible way that the blood of the victims was found, but no blood of their blood profiles was found.

Also, the values belonging to the red nodes are important. Those are added to correct for false positive results given. For example the node J^* is the event that the lab results support J. Therefore $\mathbb{P}(J^*|J)$ is equal to the accuracy of the DNA test. DNA results are said to be correct 98 percent of the time. So $\mathbb{P}(J^*|J) = 0.98$. We therefore pick $\mathbb{P}(J^*|J^c) = 0.02$, even tough these chances are not complementary per definition, it seems logical that the results are either interpreted correctly or not. (Olumide, 2010)

Far from all probabilities have been discussed, but the authors way of reasoning has been clarified. This reasoning led him to the following probabilities.

No.	Prob.	Values	Prob.	Values	Prob.	Values	Prob.	Values
1	$\mathbb{P}(H_5)$	0.5	$\mathbb{P}(H\|E^cA)$	0.60	$\mathbb{P}(L\|U)$	0.85	$\mathbb{P}(P\|W)$	0.98
2	$\mathbb{P}(H_5^c)$	0.5	$\mathbb{P}(H\|E^cA^c)$	0.00	$\mathbb{P}(L\|U^c)$	0.05	$\mathbb{P}(P\|W^c)$	0.05
3	$\mathbb{P}(A\|H_5)$	0.90	$\mathbb{P}(H^*\|H)$	0.98	$\mathbb{P}(L^*\|L)$	0.95	$\mathbb{P}(P^*\|P)$	0.98
4	$\mathbb{P}(A\|H_5^c)$	0.01	$\mathbb{P}(H^*\|H^c)$	0.02	$\mathbb{P}(L^*\|L^c)$	0.07	$\mathbb{P}(P^*\|P^c)$	0.02
5	$\mathbb{P}(D\|A)$	0.75	$\mathbb{P}(I\|H)$	0.40	$\mathbb{P}(M\|U)$	0.85	$\mathbb{P}(Q\|W)$	0.98
6	$\mathbb{P}(D\|A^c)$	0.00	$\mathbb{P}(I\|H^c)$	0.05	$\mathbb{P}(M\|U^c)$	0.05	$\mathbb{P}(Q\|W^c)$	0.05
7	$\mathbb{P}(D^*\|D)$	0.98	$\mathbb{P}(I^*\|I)$	0.99	$\mathbb{P}(M^*\|M)$	0.98	$\mathbb{P}(Q^*\|Q)$	0.98
8	$\mathbb{P}(D^*\|D^c)$	0.02	$\mathbb{P}(I^*\|I^c)$	0.05	$\mathbb{P}(M^*\|M^c)$	0.02	$\mathbb{P}(Q^*\|Q^c)$	0.02
9	$\mathbb{P}(E\|A)$	0.85	$\mathbb{P}(B\|H_5)$	0.95	$\mathbb{P}(N\|VU)$	0.95	$\mathbb{P}(R\|W)$	0.98
10	$\mathbb{P}(E\|A^c)$	0.00	$\mathbb{P}(B\|H_5^c)$	0.01	$\mathbb{P}(N\|VU^c)$	0.70	$\mathbb{P}(R\|W^c)$	0.05
11	$\mathbb{P}(F\|EG)$	0.98	$\mathbb{P}(U\|B)$	0.98	$\mathbb{P}(N\|V^cU)$	0.70	$\mathbb{P}(R^*\|R)$	0.98
12	$\mathbb{P}(F\|EG^c)$	0.00	$\mathbb{P}(U\|B^c)$	0.07	$\mathbb{P}(N\|V^cU^c)$	0.05	$\mathbb{P}(R^*\|R^c)$	0.02
13	$\mathbb{P}(F\|E^cG)$	0.00	$\mathbb{P}(V\|B)$	1.00	$\mathbb{P}(N^*\|N)$	0.98	$\mathbb{P}(X\|C)$	1.00
14	$\mathbb{P}(F\|E^cG^c)$	0.00	$\mathbb{P}(V\|B^c)$	0.0015	$\mathbb{P}(N^*\|N^c)$	0.02	$\mathbb{P}(X\|C^c)$	0.0015
15	$\mathbb{P}(F^*\|F)$	0.95	$\mathbb{P}(J\|U)$	0.90	$\mathbb{P}(C\|H_5)$	0.97	$\mathbb{P}(S\|X)$	0.85
16	$\mathbb{P}(F^*\|F^c)$	0.05	$\mathbb{P}(J\|U^c)$	0.02	$\mathbb{P}(C\|H_5^c)$	0.01	$\mathbb{P}(S\|X^c)$	0.00
17	$\mathbb{P}(G\|E)$	0.20	$\mathbb{P}(J^*\|J)$	0.98	$\mathbb{P}(W\|C)$	0.98	$\mathbb{P}(S^*\|S)$	0.98
18	$\mathbb{P}(G\|E^c)$	0.05	$\mathbb{P}(J^*\|J^c)$	0.02	$\mathbb{P}(W\|C^c)$	0.07	$\mathbb{P}(S^*\|S^c)$	0.02
19	$\mathbb{P}(G^*\|G)$	0.98	$\mathbb{P}(K\|J)$	0.75	$\mathbb{P}(O\|W)$	0.98	$\mathbb{P}(T\|X)$	0.98
20	$\mathbb{P}(G^*\|G^c)$	0.02	$\mathbb{P}(K\|J^c)$	0.20	$\mathbb{P}(O\|W^c)$	0.05	$\mathbb{P}(T\|X^c)$	0.00
21	$\mathbb{P}(H\|EA)$	0.99	$\mathbb{P}(K^*\|K)$	0.98	$\mathbb{P}(O^*\|O)$	0.98	$\mathbb{P}(T^*\|T)$	0.98
22	$\mathbb{P}(H\|EA^c)$	0.00	$\mathbb{P}(K^*\|K^c)$	0.02	$\mathbb{P}(O^*\|O^c)$	0.02	$\mathbb{P}(T^*\|T^c)$	0.02

Figure 6: Prior Probabilities, O.J. Simpson Case (Olumide, 2010)

3.2.8 Calculation

After finding a Bayesian network and the likelihood ratios belonging to the propositions, a suitable probabilistic inference algorithms for Bayesian networks was applied to determine the values of the nodes in the Bayesian network. More on the application of these inference algorithms can be found in Heckerman (1995). These algorithms have been carried out by the software GeNIe and HUGIN. This led to the following values of the posterior probabilities.

Nodes	Nodes	Values
O.J. committed the two murders	H_5	1.0000
O.J.'s blood was found at the murder scene	A	1.0000
The blood of both victims and O.J.'s blood were found on his Ford Bronco	B	0.9972
The blood of both victims and O.J.'s blood were found at O.J.'s Rockingham...	C	1.0000
Blood stains found on the back gate of NBS house matched O.J.'s DNA profile	D	0.9932
Bloody shoe prints found at the crime scene matched O.J.' DNA profile	E	0.9997
Bloody shoe prints at the crime scene came from Bruno Magli shoes	F	0.9942
O.J. owns a pair of Bruno Magli shoes	G	0.9955
Blood drops to the left of shoe prints found at crime scene matched O.J.'s DNA profile	H	0.9999
O.J. had a cut on his left finger	I	0.9296
Blood on the driver side of the exterior door handle of the Bronco matched O.J.'s DNA ...	J	0.9994
Blood on the driver's side of the interior door handle of the Bronco matched O.J.'s ...	K	0.9932
Blood on the driver side carpet of the Bronco matched O.J.'s DNA profile	L	0.9964
Blood on the steering wheel of the Bronco matched both O.J.'s and Nicole's DNA profiles	M	0.9964
Blood on the console of the Bronco matched both victims and O.J.'s DNA profiles	N	0.9989
The blood trail on the ground leading to O.J.'s Rockingham residence matched O.J.'s DNA ...	O	0.9996
Blood on the gate of O.J.'s Rockingham residence matched O.J.'s DNA profile	P	0.9996
Blood on the floor of the master bathroom of O.J.'s Rockingham residence matched O.J.'s ...	Q	0.9996
Blood in the foyer of O.J.'s Rockingham residence matched O.J.'s DNA profile	R	0.9996
Blood found on the sock found in O.J.'s master bedroom matched both Nicole's and O.J.'s DNA ...	S	0.9964
Blood on the right hand glove found at O.J.'s Rockingham residence matched both victims' and ...	T	0.9996
Someone with O.J.'s blood type bled in O.J.'s Ford Bronco	U	1.0000
Blood matching both victims' blood types was found in O.J.'s Ford Bronco	V	0.9972
Someone with O.J.'s blood type bled at O.J.'s Rockingham residence	W	1.0000
Blood matching both victims' blood types was found at O.J.'s Rockingham residence	X	1.0000

Figure 7: Posterior Probabilities, O.J. Simpson Case (Olumide, 2010)

The likelihood ratio (V) belonging to the blood types is 131,120, which would mean that when only considering the blood evidence, it would be 131,120 times as likely to obtain the evidence if our H_0 is true compared to H_0 being false. (Olumide, 2010)

3.2.9 Conclusion

The aggregate likelihood ratio, the product of the likelihood ratio's found in the six different categories, is about 99,853,716,290,000. In other words, according to this article, it is 99,853,716,290,000 times as likely to obtain the evidence at hand if we suppose O.J. Simpson committed the two murders than if we suppose that he did not. Which seems quite convincing of the guilt of Simpson, but these or similar results have never been shown to the jury. After these first trails,Simpson was found not to be guilty of the murders. (Olumide, 2010)
However, in 2008, 13 years after the original trail, he got a sentence of 33 years in jail after being found guilty of 12 charges, including armed robbery and kidnapping. (Paul Vercammen, 2008)

3.2.10 The Defence Attorney's Fallacy

As said before, not only the reasoning in presumable right likelihood ratios is important, but possible mistakes should also be addressed, so we could avoid making them in the future. In the Skorupski & Wainer (2015) article a critical look has been taken at the statement done by Alan Dershowitz, an advisor to Simposon's defence attorneys. He claimed that Simpson's previous accusations of abuse of Nicole Simpson was not relevant, since only about one in 2500 men who abuse their significant other end up killing them.
We have to note that he did not ask himself the right questions, since it was already certain that

Nicole has been murdered. This fallacy is also known as the defence attorney's fallacy and will be further discussed in subsection 3.4 .
The relevant question is "how big is the chance that if somebody who has been murdered and abused, is murdered by their abuser?". Which would be about one in 3.5 according to the following calculations based on hard facts.

In this calculations, B will represent the event of a woman being abused by her significant other. M represent the event of a woman being murdered. Logically, M,B represents the event of a woman getting murdered by the man that abused her. We will use Bayes' rule to find the probability of a murdered woman being murdered by her abuser.

$$\mathbb{P}(M, B|M) = \frac{\mathbb{P}(M|M, B)\mathbb{P}(M, B)}{\mathbb{P}(M)}$$

First, determining $\mathbb{P}(M)$ is easy, since known is that in 1992 the population of women in the USA was approximately 125 million. And the number of women murdered that year was 4936. So

$$\mathbb{P}(M) = \frac{4936}{125,000,000} \approx 0.00004$$

In that same year 1432 women were murdered by their previous abusers. Which leads to

$$\mathbb{P}(M, B) = \frac{1432}{125,000,000} \approx 0.00001$$

About 3.5 million woman are abused every year. So

$$\mathbb{P}(M, B|B) = \frac{1432}{3,500,000} \approx 0.00041$$

Which leads to the conclusion of $\mathbb{P}(M, B|M)$ being 0.29. In other words, about 1 in 3.5 of murdered woman has been murdered by her abuser. This way of reasoning to the right question after a wrong question has been addressed is also sometimes revered to as the Bayesian flip. (Skorupski & Wainer, 2015)

3.3 Troy Brown

In the O.J. Simpson Case, the DNA evidence seemed quite convincing of O.J.'s guilt and performing a Bayesian flip lead to the same result. There are also other cases known where the evidence found might not have lead to the right conclusions. An example of DNA evidence used in a questionable way is the case of McDaniel v. Brown. Kaye (2009) wrote a critical report on this case, which will be studied to understand more on what could go wrong in interpretation of DNA evidence.

3.3.1 Introduction to the case

Troy Brown, was accused of sexual assault and attempted murder. He was convicted of the rape mainly based on DNA evidence. Namely, the semen discovered on the victim's bloody panties seemed to match Troy's DNA profile. More specific, Romero, the criminalist that investigated the case said that the chance that a random individual would share the same DNA profile would be 1 in 3 million. This results was used during the trial, later, more accurate recomputations would set it to be 1 in 15 million. Both numbers would indeed seem quite convincing of Troy's guilt.

3.3.2 Prosecutor's fallacy

In this case, the following hypothesis had been made:
H_0: Troy's DNA was found.
H_1: The DNA found belonged to an unrelated person.

The criminalist looking at the case told that $\mathbb{P}(Troy|Match)$, the probability that the DNA found

in the panties is the same as the DNA found in the defendant's blood, would be $1 - \frac{1}{3million} \approx 99.999967$. In fact, the number $\frac{1}{3million}$ only represents the probability that an unrelated person is the source of the DNA given that there is a match. So in fact, Romero flipped around the hypothesis. Which caused results that where convincing, but not correct. This is a common made error, called the prosecutor's fallacy.

The main problem in Romero's calculation has been that the prior probabilities are not at all used to calculate the fact that Troy was 3,000,000 more likely to have committed the crime. The statement had been made only looking at the DNA evidence. Which totally ignores other evidence in the case. (Gastwirth, 2000)

3.3.3 DNA of Brothers

Another major error in this case is that they look at a random person matching with Troy's DNA, but they did not consider the fact that Troy's four brothers could also be suspects in the case. Those brothers have a higher probability of sharing DNA with Troy.

Even when getting asked to get the brother's involved in the calculations, Romero made some mistakes. He first concluded that there was only a one in 6,500 chance of brothers have a similar DNA profile, and concluded that having more brothers would not change this probability. He then concluded that there still was a 99.98% chance of Troy not being guilty. Again making the prosecutor's fallacy.
After this, different estimations of this probability are made, even getting as high as a chance of 1 in 66. (Skorupski & Wainer, 2015)

3.3.4 Correcting for the prosecutor's fallacy

Correcting for the prosecutor's fallacy could be done following the calculations of Skorupski & Wainer (2015), which performed a Bayesian flip. We introduce "I" for innocence, "I^c" for guilt and "D" for DNA evidence matches the suspect. We will need to estimate $\mathbb{P}(I|D)$, so the probability that brown is innocent, given that a DNA match was found. We will use

$$\mathbb{P}(I|D) = \frac{\mathbb{P}(D|I)\mathbb{P}(I)}{\mathbb{P}(D)}$$

We will use $\mathbb{P}(D|I) = \frac{1}{66} \approx 0.0151$.

$\mathbb{P}(I)$ would be the probability of Brown's innocence, which is dependent on the other evidence presented in the case. Since this evidence was not that convincing yet, but definitely did not rule out the fact that Brown was guilty. We will therefore set $\mathbb{P}(I)$ to be 0.9.

To determine $\mathbb{P}(D)$ we will use

$$\mathbb{P}(D) = \mathbb{P}(D|I)\mathbb{P}(I) + \mathbb{P}(D|I^c)\mathbb{P}(I^c)$$

We know that $\mathbb{P}(I)$ is the complement of $\mathbb{P}(I^c)$. So $\mathbb{P}(I) = 1 - \mathbb{P}(I^c) = 1 - 0.9 = 0.1$.

The only probability left to estimate is $\mathbb{P}(D|I^c)$, which is the probability that the DNA matches given that Tony is guilty, so $\mathbb{P}(D|I^c) = 1$.

Using these numbers will give $\mathbb{P}(D) = 0.9 \cdot 0.0151515 + 0.1 \cdot 1 = 0.1136$.

Completing the calculation $\mathbb{P}(I|D) = \frac{0.0151 \cdot 0.1}{0.1136} \approx 0.12$. Which would not let us conclude that Troy would probably be innocent, but it does show that there is a reasonable doubt in concluding guilt. (Skorupski & Wainer, 2015)

3.3.5 Conclusion

From this case, we can conclude that DNA evidence might have some serious consequences when it is not interpreted correctly. Also, DNA evidence alone will not be enough to cover a case, since this would leave out important other evidence.

3.4 Robert Wayne Johnson

3.4.1 Introduction to the case

On the morning of July 9, 1991, a storekeeper was overpowered and raped. After this rape, the paper towels that were used by the victim to clean herself were investigated by the police. The DNA found matched that of one of the suspects; Robert Wayne Johnson. Terry Hogan, the criminalist working this case found that the odds of this match occurring randomly was only one to 312 million. Therefore the jury concluded Johnson's guilt and imprisoned him for fourteen years.
Johnson decided to petition the court for review, since he claimed innocence and therefore that the wrong conclusion was drawn based on the DNA evidence. (Feldman, 1996)

3.4.2 Probability of a Match in Two Piece of Evidence

In order to find a right interpretation of the DNA evidence we will take a look at the calculations done by Gastwirth (2000).

What went wrong in the calculations as they are done is that they calculated the probability of two unrelated individuals having the same DNA pattern to be one in 312 million. This is not the correct interpretation of this value, since this is a comparison of two unrelated pieces of evidence. Namely, E_1, the evidence of a DNA profile form a stain found at the crime scene and E_2, the evidence of the DNA profile of the suspect. Those should not be compared directly since they both contribute on a different level as evidence to find whether the suspect was present at the crime scene. In order to find a correct value belonging to the hypothesis we will do some fixing for the mistakes made.

First, we set the following hypotheses:
H_0 : Johnson was present at the crime scene
H_1 : The suspect was not present at the crime scene

Now, consider V to be the value belonging to the evidence found, so 312 million, where E is written as (E_1, E_2). Taking a look at the likelihood will give us the following calculation:

$$V = \frac{\mathbb{P}(E_1, E_2|H_0)}{\mathbb{P}(E_1, E_2|H_1)}$$

which is equal to

$$\frac{\mathbb{P}(E_1|E_2, H_0)\mathbb{P}(E_2|H_0)}{\mathbb{P}(E_1|E_2, H_1)\mathbb{P}(E_2|H_1)}$$

Now, we note that $\mathbb{P}(E_2|H_0) = \mathbb{P}(E_2|H_1)$, since the evidence matching the DNA profile of the suspect is the same, whether he was or was not present at the crime scene. This leaves us with with the following value

$$\frac{\mathbb{P}(E_1|E_2, H_0)}{\mathbb{P}(E_1|E_2, H_1)}$$

If we would then make assumptions of whether the suspect was or was not present at the crime scene, this value would get even more specific. For example, if we expect that the suspect was present, we find $\mathbb{P}(E_1|E_2, H_0)$ to be 1, since we are certain that we found a matching DNA profile at the crime scene. Also, we already specified that E_1 and E_2 should be cared for as two individual pieces of evidence. Therefore $\mathbb{P}(E_1|E_2, H_1) = \mathbb{P}(E_1|H_1)$ under the circumstance that the suspect was present at the crime scene.

This would finally lead to

$$V = \frac{1}{\mathbb{P}(E_1|H_1)}$$

Which is all that is relevant for this case.

If you would want to calculate the probability of two individuals having the same DNA profile, you could simply calculate $(\mathbb{P}(E_1|H_1))^2$, but this is definitely not relevant, since we want to measure the strength of the given evidence and only knowing this probability would not give a measure of the probability that Johnson would have been guilty either. So doing any court decisions based on this number would be a fallacy in it's own way. (Gastwirth, 2000)

3.4.3 Conclusion calculations

So, we found that 1 in 312 million is the probability that one individual selected at random from the population has the DNA profile found at the crime scene. Which is of course not the same as the probability of two unrelated individuals having the same DNA pattern, as which Hogan has set this value. So the probability given in court was misleading.

3.4.4 Probability of Finding Another Match

Again, we consider the fact that the relative frequency of the DNA pattern is 1 in 312 million. This would mean that the probability that no match is found is $1 - \frac{1}{312million}$. From this we want to deduce that finding a match with this DNA pattern of the crime scene might be easier than you expect it to be. When we consider a group of n people, the change of nobody matching the DNA found a the scene would be $(1 - \frac{1}{312million})^n$. So, the chance of finding a match would be $1 - (1 - \frac{1}{312million})^n$. If we consider n to be 1 million, we would find that there is a 0.0032 chance to find a DNA match. Which could also be interpreted as the change of an innocent person of a population of 1 million people to be wrongly matched to the DNA.

This calculation does not directly link to the Bayesian methods discussed before, but it does give an other sight on why the statement made by the court was not right. (Gastwirth, 2000)

3.4.5 Conclusion Overall Case

From this case, we can learn that it is important that pieces of evidence that might seem relevant at first should not be matched with a probability that is not relevant. In this case they made the mistake that they found a match between suspect and crime scene, but did not take into account any randomness that is possible to find this match.

3.5 Errors per Case

Per case certain errors that where made in the trail process were discussed. A short summary of the errors made follows, to have a more general idea of them.

3.5.1 O.J. Simpson

Defence Attorney's Fallacy
In the case of a defence attorney's fallacy, also known as a defense fallacy, the evidence presented is not relevant, since there was a wrong assumption that all people were equally likely to commit the crime. (Gastwirth, 2000)

3.5.2 Troy Brown

Prosecutor's Fallacy
Mathematically, this fallacy means that $\mathbb{P}(E|H_1)$ is confused with $\mathbb{P}(H_1|E)$. Mostly, it comes to

conclude that the suspect is not guilty, because of a low probability $\mathbb{P}(E|H_1)$, and finding this enough to conclude a low probability of $\mathbb{P}(H_1|E)$. (Gastwirth, 2000)

DNA of Brothers
When taking a DNA sample of DNA, not the entire DNA is investigated. Only a small fragment will be investigated. Therefore, it is possible for two people to have similar DNA fragments, without having the same complete DNA. In other words, a DNA match could be found with more than one person. It does not take a genetics expert to realise that two brothers are much more likely to have similar DNA than two randomly sampled individuals from the population. It is important to be aware that the frequency of DNA patterns is not evenly distributed within the population. (Skorupski & Wainer, 2015)

3.5.3 Robert Wayne Johnson

Probability of a Match in Two Pieces of Evidence
When adding a second piece of evidence, we should make sure that this piece of evidence does have an influence on the first evidence we had. In other words $\mathbb{P}(E_2|H) = \mathbb{P}(E_2|H \wedge E_2)$. Because when $\mathbb{P}(E_2|H) \neq \mathbb{P}(E_2|H \wedge E_2)$, instead of looking at a match between evidence and suspect, a match between two pieces of evidence is found. From this, they wrongly conclude that this match makes the suspect more suspicious. (Gastwirth, 2000)

Probability of Finding Another Match
This error has to do with concluding that a person is guilty because of a single match of DNA evidence. Therefore it is important to note that finding a match while the suspect is innocent is also possible.

When there is a population of N people and q is the frequency of the DNA pattern in this population. Then there is a $1 - (1 - q)^N$ chance of finding a innocent person in the population to match with the DNA evidence. (Gastwirth, 2000).

Example
Now we consider the entire population of the Netherlands as our value for N, which are currently about 17 million people. Since q is a frequency, we know $q \in [0, 1]$. We can even assume $q \in [0.0001, 0.9999]$, since we only want to look at reasonable values for q. (Budowle B et al., 1999) For N = 17 million, $(1-q)^N$ goes to 0 for all reasonable values of q. So, for all frequencies, there is always a chance of almost 1 of finding an innocent person in the Netherlands matching the DNA found at any crime scene. So, we are almost sure that there is always a person with a similar DNA fragment in the Netherlands, for all DNA fragments found. Therefore it is needed to not just base a decision on a DNA match only. This will be further explained in the following chapter.

4 DNA Statistics

In the cases discussed in the chapter before, we noticed the importance of DNA evidence in a forensic case. That is why we will dive a bit deeper into DNA typing and the data bases that are used presently to determine DNA matches.

4.1 DNA Typing

In short, DNA typing could be explained as the characterization of deoxyribonucleic acid (DNA) for purposes of criminal investigation. It has great potential benefits for criminal and civil justice, but there has been a lot of questions on the reliability, validity and confidentiality of the method. Therefore the Committee on DNA Technology in Forensic Science was formed in 1990. They not only developed standards for data collection and analysis but also go into depth on the legal, societal and ethical issues surrounding DNA typing. (National Research Council, 1992)

Since this report is written from a mathematical point of view, we are mostly concerned with their point of view on how to use statistics with DNA Typing. We will therefore not continue on these legal, societal and ethical issues, but it is important to be aware of their existence.

4.2 A Biological Background

In order to understand DNA typing, some basic biological background is needed.

4.2.1 Definitions

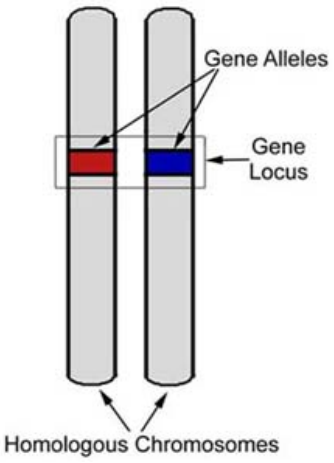


Figure 8: Simple diagram of DNA (Ashcraft, 2005)

There are two different variations of all *genes*, they are called *alleles*. A gene is located at a specific location on the DNA, this location is called the *loci*.

Because of a large number of genes, no two persons (barring identical twins) have the same DNA sequence. Unique identification with DNA typing is therefore possible, in principle, but only if enough loci are investigated. However, the DNA typing systems used today examine only a few sites of variation and have only limited resolution for measuring the variability at each site. There is a chance that two persons have DNA patterns that match at the small number of sites examined. (National Research Council, 1992)

4.2.2 CODIS

The site of the FBI, provides us with the following information.

The Combined DNA Index System, or CODIS, blends forensic science and computer technology into a tool for linking violent crimes. It enables federal, state, and local forensic laboratories to exchange and compare DNA profiles electronically, thereby linking serial violent crimes to each other and to known offenders. Using the National DNA Index System of CODIS, the National Missing Persons DNA Database also helps identify missing and unidentified individuals. In early 2015 the FBI announced 13 core loci to start the CODIS. In 2017 7 more loci where added to the CODIS. (Hares, 2012)

In other words, the CODIS is a data bank filled with information on 13 loci. This data bank is not accessible for regular people, therefore we cannot do any research into it. But we are able to simulate a data bank and do some research on that.

4.3 Modeling a DNA Data set

A simple model of a DNA data bank, based on this CODIS, will help us to look into the possibility of convicting the wrong person based on DNA evidence. In other words, we want to understand how unique DNA actually is.

4.3.1 Goal

We want to determine how easy it is to bring back a piece of DNA to the right person. In order to answer this, we will try to find how often certain DNA fragments occur in a data bank. Also, we want to investigate mistakes made within ethnic groups.

4.3.2 Assumptions

- We will not go into depth on the biological part of the research the model is, and therefore assume that the research has been done on a proper way. We will therefore also assume that the results of the project are correct.
- There is some connection know between the presence/ absence of certain pairs of loci. For simplification of the model we will assume independence instead.
- The DNA of the suspect is present at the DNA data bank.
- We will assume that the six ethnicities used are enough to make some general remarks.

4.3.3 The model

This model is based on the research done by Budowle B et al. (1999). Where they investigated the relative frequency of the alleles on the 13 loci used in FBI investigations within six different ethnicities. They investigated the relative frequency of a certain allele. Since we know that there are only 2 types of the genes, we could calculate the frequency of the other allele from this.
So we have a string of 13 relative frequencies per ethnicity.
To produces the DNA of certain person, we first produce 13 random numbers between 0 and 1 and compare this string to the numbers to the frequency of the alleles in a certain ethnicity. When the random number is smaller than the allele frequency, we will say that the allele is present, when it is bigger, we will say that the other type of gene is present. This was, a data set of 1000 string of DNA fragment per ethnicity are simulated.

After that, in a similar way, again 1000 string are produced per ethnicity. Of this string set, every string is compared to the data base. When a match is found, it is counted. We can distinguish two kind of matches; correct or incorrect. With the correct matches, we mean, matches that are matched with the right ethnicity. While the incorrect matches are matched to the wrong ethnicity. So correct matches are not necessarily matches to the right person.

The R program could be found as appendix.

4.3.4 Results and Interpretation

The program provided the following results.

Bar Plots

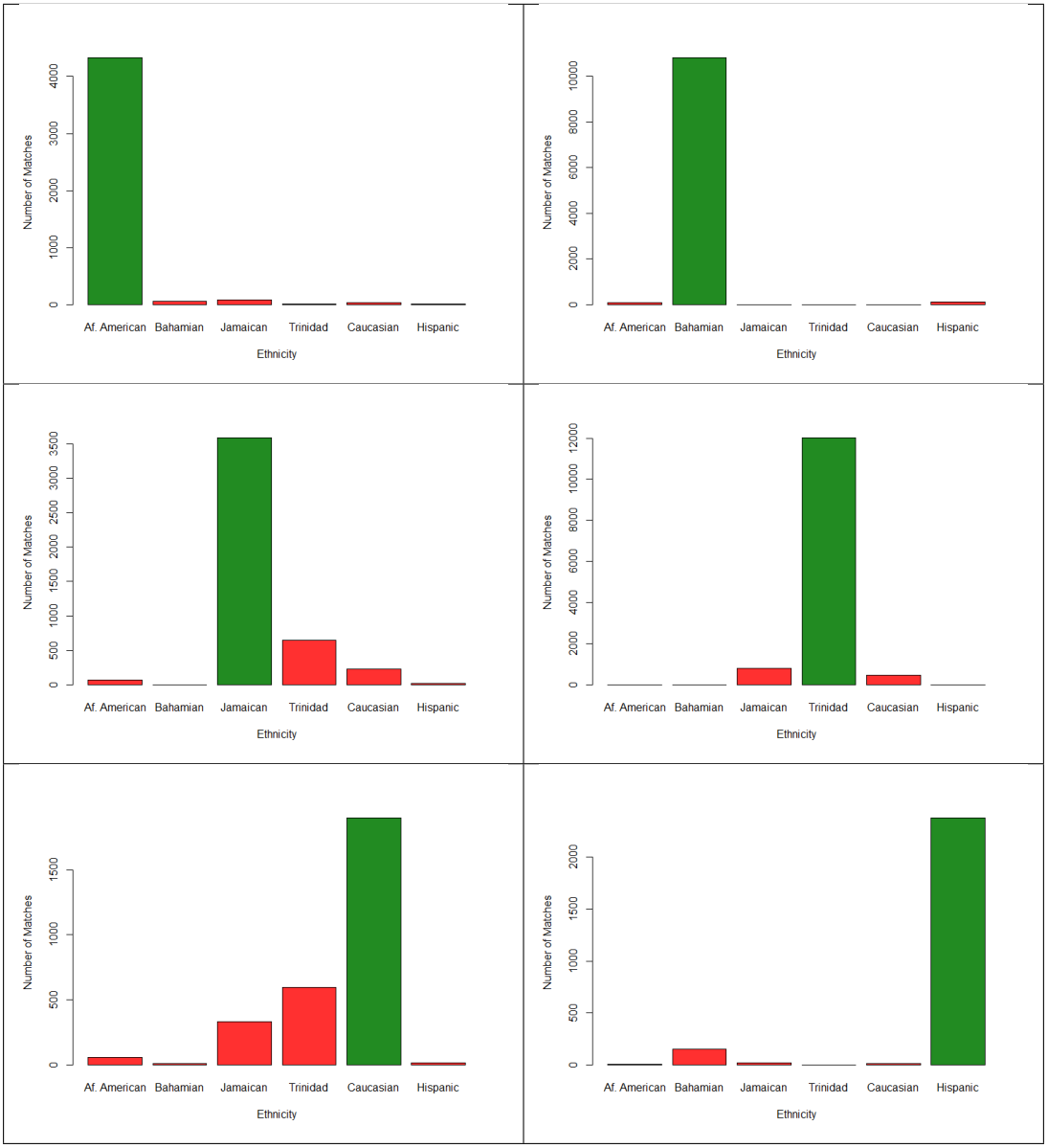


Table 4: The Number of Correct (Green) and Incorrect (Red) Matches per Ethnicity

Interpretation Bar Plots

In the bar plots above, we, find that the green bar, representing a correct match, is the biggest in all six cases. Which means that when a DNA of a certain ethnicity is found, in most cases it can be brought back to the correct ethnicity.

Important to note is the different numbers on the y-axis.A high number of correct matches might seem good at first, but this also means that there are more people appointed when entering one

piece of DNA in the data bank. Which is unfortunate.

Since we simulated the DNA of a person of a certain ethnicity, we now that the DNA that is found of a similar code within a different ethnicity could never be from the suspect. This is something that we could say, since we formed the DNA fragment ourselves. This is of course not possible to conclude from the DNA found at a crime scene.

So mistakes could be made within the ethnicity of the offender, but also between different ethnic groups. It is much more likely to make a mistake within the ethnic group.

We are curious on how many people would wrongly be matched and what ethnicity these people belong to. Therefore we are going to take a look at the following tables.

Matches within Correct Ethnicity

Ethnicity	Af. American	Bahamian	Jamaican	Trinidad	Caucasian	Hispanic
Percentage of correct matches	95.03624	97.90647	74.58403	90.54044	71.45009	93.02781

Table 5: The percentage of DNA fragments matched to the correct ethnicity

Interpretation Matches within Correct Ethnicity
We found the percentage of DNA fragments that are correctly linked to the ethnicity they came from. This has been done to check possible bias in a data bank.

It could be found that almost 29% of the Caucasian DNA would get linked to an other ethnic group. Comparing that to the fact that only some more than 2% of the Bahamian DNA gets misjudged raises some questions. We could ask ourselves if testing these specific 13 loci would be in advantage for the Caucasian population when it comes to crime scenes. This will be left for a possible further investigation.

Number of Matches

Ethnicity	Af. American	Bahamian	Jamaican	Trinidad	Caucasian	Hispanic
Number of matches	4.553	11.034	4.808	13.267	2.655	2.553
Percentage of matches	0.076	0.184	0.080	0.221	0.044	0.043

Table 6: The average number of matches per DNA fragment

Interpretation Number of Matches
We obviously conclude that it is no unique DNA fragment per person. Therefore, when a DNA match is found between the DNA at a crime scene and the DNA of a suspect, there is no certainty yet.

The Bahamian and Trinidad people have the most matches per DNA fragment, and are therefore most difficult to distinguish correctly.

Number of Wrong Matches per Ethnicity

Ethnicity	Af. American	Bahamian	Jamaican	Trinidad	Caucasian	Hispanic
Number of wrong matches	226	231	1222	1255	758	178

Table 7: The number of errors per ethnicity

Interpretation Number of Wrong Matches per Ethnicity
We find that Jamaican and Trinidad people are more often wrongly matches to evidence than African American, Bahamian or Hispanic people. Caucasian people are somewhat in between these groups. Which in real life would mean that people of Jamaican and Trinidad ethnicity are more likely to match with DNA of an other ethnic group, and therefore are more likely to get accused of a crime someone of an other ethnic group committed.

4.3.5 Conclusion

There is no unique person to be appointed when these 13 loci are investigated. It could however narrow down the suspects to a more reasonable amount, but DNA does not give a single solution.

The system is quite good at linking DNA to the right ethnic group. Which somehow concludes that the loci tested do distinguish enough.

From the test run, especially DNA of Trinidad people is hard to investigate. It is often confused with DNA from other ethnicities. Which makes that people with the Trinidad ethnicity are more likely to get linked to a crime they did not commit. We note that this is the case with this limited DNA data bank.

4.3.6 Possible improvements on the model

The model written is a simplified version of a DNA data bank. In real life, this data bank would have DNA of real people, and should for that reason not have to be simulated. It would be possible to improve the model to make it more like a real life data bank:

- In the model, independence of all loci is assumed, we know that this is not true to life. Creating a model where this dependency is used might give different results.
- Adding more ethnicities might make it more difficult to distinguish between them. So testing with a bigger set will definitely be interesting.
- Adding more loci would give us more factors to distinguish differences in DNA. So it would be useful to add them to increase the certainty of matches.

5 Conclusion and Discussion

5.1 Summary per Section

A general summary per chapter will lead us to an overall conclusion.

5.1.1 Theoretical Background

We reported a way to summarize evidence found at a crime scene by the help of Bayesian networks. By following the steps provided at this chapter, we could find a likelihood. This likelihood could then be interpreted by judges. The higher the likelihood of the system, the more likely the suspect is indeed guilty.

5.1.2 Cases

For the cases studies, we got to understand the use of Bayesian networks. We got a insight in the considerations a person should make when making a Bayesian network based on a court case.
In the case studies we also took a look at possible errors, from which we concluded that it is often quite hard to see the difference between doing a wrong or right calculation, even in big public cases.

5.1.3 DNA Statistics

The most important thing we should take away from this chapter that, even though DNA is unique per person, the DNA fragments tested are not. So, there are more people who could be linked to a certain crime scene.

The DNA fragments used at the moment do quite a good job at distinguishing ethnicities. But there is a clear difference between ethnicities. There is an advantage for Caucasians for the genes that are used nowadays.

5.2 Conclusion

Mathematics could be helpful in the field of law. It is not perfect yet, but it could be useful to make reasoning more insightful. Doing the court decision based on number will make it more fair and less based on the fact whether or not the defendant was able to convince the judge on a emotional bases.

Another advantage is that the decision is not based on the personal preferences of the judge. Different judges should come to the same conclusion without having to have a strong personal opinion about the evidence provided.

There are also some disadvantages to using mathematics. Mostly, because a judge should have a real understanding of probabilities and how to interpret them, which, we found from the errors made, is not always the case yet.
Fortunately, we are able to trace these errors and therefore make an alternative decision when we feel like a mistake has been made.

Finally, with probabilities, mistakes can still be made, since very likely situations might not be true. An example of this can be found when researching a DNA data bank.

5.3 Discussion

Various notes concerning this subject should be made in order to understand the context in which this report was written.

5.3.1 Application of Mathematics in Law

The application of mathematics in law is something that is upcoming and is quite usual for, for example, DNA investigation. The Bayesian network is not used in court yet, but it is a topic that often occurs in scientific report. In this report we established that it indeed might be a handy tool. But judges are often faced with unlikely situations and it might not be easy to apply mathematics for them. Also, just handing judges a number, without them understanding it might be dangerous. So we have to ask ourselves if regular judges should learn how to make such a network, or if we should train people on it.

5.3.2 Variation within Court Systems

Note that since the court system varies from location to location, some claims done might not be general. The statements that are discussed will be looked at from the court system that will be most relevant to the original trial.

In the cases that have been discussed in chapter 3, rape and murder have been mainly discussed. Which are not only defined differently by different courts, but also have a different punishment linked to them. This punishment is not that relevant with respect to the mathematical aspect of this report. But the definition of the alleged crime is relevant, since a judge will always use the evidence to either prove or disprove exactly the given definition. This would mean that evidence found might not be as relevant for certain courts, since the evidence found would simply not support anything that is considered illegal for that court.

This would therefore mean that certain evidence might not even be present when computing a Bayesian network of a certain situation for a certain court, while it would be present when the case would be presented to a different court.

5.3.3 Other Sciences

The field of forensic statistics is only a building block in respect to the entire field of forensic sciences. Since the report is written by somebody with a mathematical background, this has been the focus point. Noted should be that, just optimizing the mathematical aspect of making decision in court would therefore not necessarily have an enormous influence on the entire field of the decision making process.

5.4 Further investigations

Within the report, some remarks are made about possible further investigations that could be done to improve the collaboration between mathematics and law. A short summary of these remarks will follow. Some other options for further investigations are added.

5.4.1 Effect of Different Bayesian network

We could change the values in given conditional probability tables. Removing evidence or switching probabilities could change our conclusion. Playing around with it will show us the influence of the values of the nodes. Also, changing the dependencies in the graph might influence the outcomes.

5.4.2 Further investigations concerning the model

The results of the model raised the question if testing these specific 13 loci would be in advantage for the Caucasian population when it comes to crime scenes. Improving the model on a whole would also be possible as further investigation. We suggested the following improvements.

- In the model, independence of all loci is assumed, we know that this is not true to life. Creating a model where this dependency is used might give different results.
- Adding more ethnicities might make it more difficult to distinguish between them. So testing with a bigger set will definitely be interesting.

- Adding more loci would give us more factors to better distinguish DNA. So it would be useful to add them to increase the certainty of matches.

Also, looking at a real data set that is used for DNA evidence right now might me interesting. The same type of test could be done to answer if a real data set is better at giving a single person per DNA fragment.

5.4.3 Ask Specialists

For this report, we focused on the literature available on the subject. It would be interesting to talk to judges and forensic experts to ask them about their view on the use of mathematics in law. We could also ask to which extent they use mathematics right now. All this could be done in order to provide a roadmap for the use of mathematics in court, that will fit preferences of judges, forensic experts and mathematicians.

6 Appendix

6.1 Code in R

The following code was run by R to produce the results found in chapter 5

```
#load the excel sheet with the frequencies
wd <- "C:/Users/s146413/Documents/TUE/2018-2019/BEP"
setwd(wd)
library(readxl)
my_data <- read_excel("gens.xlsx",col_names = c("V1", "V2", "V3", "V4", "V5", "V6"))
my_data <- as.data.frame(sapply(my_data, as.numeric))

#a function for the production of the data set
sim <- function(my_data){
  x<- matrix(data = NA, ncol = 6, nrow = 13) #produce a matrix for 6 people with 13 genes
  for(i in 1:6){
    x[,i]<-t(runif(13)) #run 13 random numbers between 0 and 1 per person
  }
  y<-matrix(data = rep(0,13*6), ncol = 6, nrow = 13) #produce a matrix for 6
  #people with 13 genes

  for(j in 1:6){
  for (i in 1:13){
  if(x[i,j] < my_data[i,j]){ #compare the random numbers with the allele frequency
    y[i,j]<-1 #put a 1 when the allele is present
  }
  }
  }
  return(y) #return the dataset with 0's en 1's
}


DNAnr <- 1000 #determine how many DNA fragments we want to check
#produce a list of DNA for 1000 people per ethnicity
sim(my_data)
data <- list()
for(i in 1:DNAnr){
  data[[i]] <- sim(my_data) #produce the dataset
}

count <- function(ethnicity){
DNA1 <- runif(13)
for (i in 1:13){
  if(DNA1[i] < my_data[i,ethnicity]){
    DNA1[i]<-1 #put a 1 when the allele is present
  }else{
    DNA1[i]<-0 #put a 0 when the alternative allele is present
  }
}

#compare the DNA fragments to the fragments in the data bank
count <- rep(0,6)
for(i in 1:length(data)){
  for(j in 1:6){
    if(prod(DNA1==data[[i]][,j])==1){ #compare the DNA of a person with the DNA
```

```
    #from the data set
      count[j] <- count[j]+1 #count the number of matches
    }
  }
}
return(count)
}

tot <- rep(0,6) #introduce a counter to count the total number of matches for all ethnicities
match<- rep(0,6) #introduce a counter to count the total number of correct
#matches for all ethnicities
for(i in 1:6){
countA <- rep(0,6) #introduce a counter to count the number of matches per ethnicity
for(j in 1:DNAnr){
count(i)
  countA <- countA + count(i)
  match[i] <- countA[i]
}
tot <- tot + countA #count the total number of matches per ethnicity.
c <- rep(0,6)
for( k in 1:6){
  if (k==i){
 c[k] = "forestgreen" #make the correct matches green
  }
  else{
  c[k] = "firebrick1"#make the incorrect matches green
  }
}

#produce the barplots belonging to the different ethnicities
names(countA) <- c("Af. American", "Bahamian", "Jamaican", "Trinidad", "Caucasian", "Hispanic")
+
barplot(countA, col = c, xlab = "Ethnicity", ylab= "Number of Matches")
}
```

6.2 Excel table

The following table has been used in the R code

	A	B	C	D	E	F
1	0.797	0.758	0.270	0.222	0.084	0.333
2	0.328	0.790	0.655	0.229	0.063	0.928
3	0.995	0.930	0.290	0.835	0.270	0.635
4	0.696	0.168	0.284	0.896	0.786	0.061
5	0.495	0.010	0.917	0.912	0.415	0.655
6	0.920	0.263	0.573	0.523	0.622	0.512
7	0.468	0.271	0.770	0.977	0.570	0.522
8	0.298	0.173	0.051	0.068	0.395	0.993
9	0.490	0.159	0.257	0.364	0.321	0.422
10	0.983	0.439	0.354	0.014	0.620	0.205
11	0.413	0.617	0.011	0.233	0.035	0.415
12	0.137	0.911	0.116	0.392	0.423	0.910
13	0.110	0.043	0.524	0.859	0.704	0.506

References

Ashcraft, C. (2005). *Genelocus.* Retrieved 2019-06-26, from http://creationwiki.org/File:Genelocus.JPG

Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, & Keys KM. (1999). *Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians* (Tech. Rep.). Retrieved from https://pdfs.semanticscholar.org/0cae/45be08805e3b1a48051c7072a0a8af6bc401.pdf?{_}ga=2.52383277.54269664.1560857559-619004921.1559828539

Chl. (2013). *Type I and Type II Errors.* Retrieved 2019-07-01, from https://stats.stackexchange.com/posts/3568/revisions

Cohen, A. M., Cohen, H., Eisenbud, D., Singer, M. F., & Volume, B. S. (2000). *Algorithms and Computation in Mathematics* •. Retrieved from https://link-springer-com.dianus.libr.tue.nl/content/pdf/10.1007{%}2F978-3-540-72780-4.pdf

Curran, J. M. (2009). *Statistics in forensic science* (Vol. 1; Tech. Rep.). Retrieved from www.wiley.com/wires/compstats doi: 10.1002/wics.033

EDU. (n.d.). *Forensic Science Job Description.* Retrieved 2019-05-18, from https://www.crimesceneinvestigatoredu.org/forensic-scientist-job-description/

Feldman. (1996). *State v. Johnson* (Tech. Rep.). Retrieved from http://homepages.law.asu.edu/{~}kayed/class/genlaw/AZ-96-johnson.pdf

Fenton, N., Neil, M., & Lagnado, D. (2010). *A General Structure for Legal Arguments Using Bayesian Networks* (Tech. Rep.). Retrieved from https://pdfs.semanticscholar.org/d98c/f931bd44ae623e09c13600bd1b7f4c13821d.pdf?{_}ga=2.203265369.95770643.1559828539-619004921.1559828539

Gastwirth, J. L. (2000). *Statistical Science in the Courtroom* (J. L. Gastwirth, Ed.). New York, NY: Springer New York. Retrieved from http://link.springer.com/10.1007/978-1-4612-1216-4 doi: 10.1007/978-1-4612-1216-4

Hares, D. R. (2012, sep). Addendum to expanding the CODIS core loci in the United States. *Forensic Science International: Genetics*, *6*(5), e135. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S1872497312000051 doi: 10.1016/j.fsigen.2012.01.003

Heckerman, D. (1995). *A T utorial on Learning With Bayesian Networks* (Tech. Rep.). Retrieved from http://www.cis.upenn.edu/{~}mkearns/papers/barbados/heckerman.pdf

Ibs, I. C. (2016). Reasoning Applications of Bayesian Networks In Legal Reasoning by. (December). doi: 10.13140/RG.2.2.28030.18248

Kaye, D. H. (2009). *False But Highly Persuasive": How Wrong Were the Probability Estimates in McDaniel v. Brown?, 108 Mich* (Vol. 108; Tech. Rep.). Michigan: Pennsylvania State University, Dickinson School of Law. Retrieved from http://repository.law.umich.edu/mlr{_}fihttp://repository.law.umich.edu/mlr{_}fi/vol108/iss1/9

National Research Council. (1992). *DNA Technology in Forensic Science.* The National Academies Press. Retrieved from https://www.nap.edu/catalog/1866/dna-technology-in-forensic-science

Nordgaard, A., Ansell, R., Drotz, W., & Jaeger, L. (2012). Scale of conclusions for the value of evidence. *Law, Probability and Risk*, *11*. Retrieved from https://watermark.silverchair.com/mgr020.pdf?token=AQECAHi208BE49Ooan9kkhW{_}Ercy7Dm3ZL{_}9Cf3qfKAc485ysgAAAmIwggJeBgkqhkiG9w0BBwag\ggJPMIICSwIBADCCAkQGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQMEranATMpu6xzI8y-AgEQgIICFW7sMv3fvZX-imY-Ib7CBzSLv6DD5HyurIw7c1MX4-0kXjIO doi: doi:10.1093/lpr/mgr020

Olumide, K. (2010). *Probabilistic and Graphical Analysis of Evidence in O.J. Simpson's Murder Case Using Bayesian Networks* (Doctoral dissertation, The Florida State University). Retrieved from http://diginole.lib.fsu.edu/islandora/object/fsu{%}3A254040

Paul Vercammen. (2008). *O.J. Simpson guilty of armed robbery, kidnapping - CNN.com.* Retrieved 2019-03-14, from http://edition.cnn.com/2008/CRIME/10/04/oj.simpson.verdict/

Privitera, G. J. (2015). *Statistics for the Behavioral Sciences.* SAGE Publications, Inc. Retrieved from https://www.sagepub.com/sites/default/files/upm-binaries/40007{_}Chapter8.pdf

Sjerps, M. (2008). Forensic Statistics. *Netherlands Forensic Institute*, 150–155. Retrieved from http://www.nieuwarchief.nl/serie5/pdf/naw5-2008-09-2-150.pdf

Skorupski, W., & Wainer, H. (2015). The Bayesian flip Correcting the prosecutor's fallacy. *significane*, 16–20. Retrieved from https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2015.00839.x

The Innocence Project. (2018). *Korey Wise - Innocence Project.* Retrieved 2019-06-04, from https://www.innocenceproject.org/cases/korey-wise/

The Judicial Branch of California. (n.d.). *Basics of Criminal Court.* Retrieved 2019-05-27, from https://www.courts.ca.gov/10214.htm

Thompson, W. C. (1996). DNA Evidence in the O.J. Simpson Trial. *University of Colorado Law Review*, *67*, 827–857. Retrieved from https://miketgriffith.com/files/thompsonarticle.pdf

Vlek, C. S., Prakken, H., Renooij, S., & Verheij, B. (2016). A method for explaining Bayesian networks for legal evidence with scenarios. *Artif Intell Law*, *24*, 285–324. Retrieved from https://link.springer.com/content/pdf/10.1007{%}2Fs10506-016-9183-4.pdf doi: 10.1007/s10506-016-9183-4