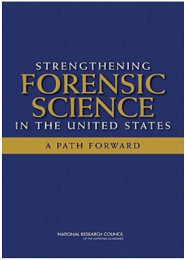
Statistical Thinking for Forensic Practitioners

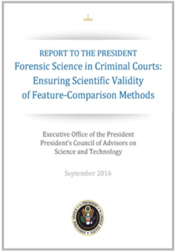
Hal Stern
University of California, Irvine

csafe
Center for Statistics and Applications in Forensic Evidence

October / November 2022

Why are we here? - Interesting times in forensic science





Organization of OSAC Scientific Area Committees for Forensic Science

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

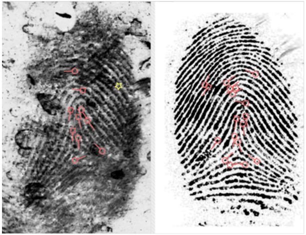
csafe
Center for Statistics and
Applications in Forensic Evidence

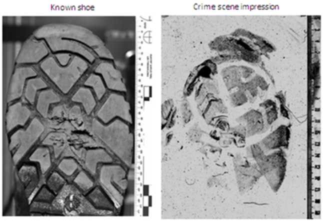
Context for the Short Course

- Daubert standard governs admission of scientific expert testimony in federal courts
  - Judge as gatekeeper
  - Relevant factors for judge to consider include peer review, known error rate, standards, etc.
  - Some states still use Frye standard of "general acceptance" in the relevant scientific community
- FRE 702 requires testimony be "based on sufficient facts or data" and use "reliable principles and methods" "reliably applied ... to the facts of the case"
- National Academies of Science (2009) and PCAST (2016) reports raise questions about the scientific foundation of pattern matching (and other types) of evidence
- Increased attention on the role of probability and statistics

Context for the Short Course

- Many different forensic questions
- Focus of this discussion is questions of source determination
- Do evidence samples (e.g., from a crime scene and a suspect) come from the same source? Examples to be discussed include
  - DNA
  - Trace evidence (e.g., glass)
  - Pattern evidence (e.g., fingerpints, shoe prints)





Outline

- **Part 1 - Probability Concepts and Their Relevance to Forensic Science**
  - **review of probability concepts**
  - **conditional probability and independence**
  - **Bayes' Theorem and likelihood ratio**
- Part 2 - Data, Measurement, Reliability and Expert Opinion
  - collecting data
  - measurement, variability, reliability and accuracy
  - forensic evidence evaluation as expert opinion / black box studies
- Part 3 - Statistical Inference and the Two-Stage Approach to Evidence
  - estimatiion, confidence intervals, significance tests
  - two-stage approach (significance test/coincidence probability)
- Part 4 - The Likelihood Ratio Approach - Strengths and Weaknesses
  - introducing the likelihood ratio
  - examples – the good, the bad, and the ugly

Learning Objectives for Part 1

- Understand the difference between population and sample, and the role of probability
- Understand the definition of probability and how it is used to characterize uncertainty
- Understand the meaning of conditional probability and independence
- Understand how Bayes Theorem works and its relationship to the likelihood ratio

The Role of Probability

People (CA) v. Collins (1968)

- An elderly woman walking in an alley was attacked from behind and robbed
- She saw a young woman with blonde hair running away
- Other witnesses saw a woman with blonde hair in a ponytail get into a yellow car driven by a black man with a mustache and beard
- Police were eventually led to an interracial couple living in the area with a yellow Lincoln

The Role of Probability

People (CA) v. Collins (1968)

- Prosecution gave estimates of the frequency of the characteristics identified by the witnesses
  - Black man with a beard: 1 out of 10
  - Man with a mustache: 1 out of 4
  - White woman with blonde hair: 1 out of 3
  - Woman with a ponytail: 1 out of 10
  - Interracical couple in a car: 1 out of 1,000
  - Yellow car: 1 out of 10
- Prosecution did not indicate basis for the numbers (more on this later)

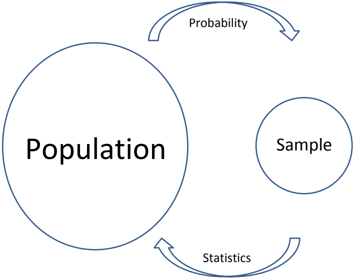
The Role of Probability

People (CA) v. Collins (1968)

- Expert witness for the prosecution: mathematics professor
  - Given all of the probabilities by the prosecutor
  - Asked to combine them all to result in the probability of finding all the characteristics in one couple
  - Multiplied them all together (citing the product rule for independent events):
    "1 out of 10" x "1 out of 4" x ...
  - Resulting probability is 1 out of 12 million
- Prosecutor's conclusion:
  - A couple which matches all of the witness observations is so rare that the couple on trial must be the couple that committed the robbery

Probability and Statistics Preliminaries

"The Big Picture"



- Population = universe of objects of interest
  Sample = objects available for study
- Probability: population $\rightarrow$ sample (deductive)
- Statistics: sample $\rightarrow$ population (inductive)

Probability and Statistics Preliminaries

The Big Picture in Practice

- Applications
  - Drug seizure (population = 100 bags; sample chosen for analysis)
  - Glass fragments (two populations = glass from crime scene and glass from suspect; take samples from each)
  - Forensic accounting
    (population = all transactions; sample chosen for analysis)
- Relevance to pattern evidence
  - Interested in variation among samples from a population of "same source" impressions (e.g., distortion in latent prints)
  - Interested in variation among samples from a relevant population of alternative sources

What is probability?

- Probability is the mathematical language of uncertainty
- The probability of an event is a number (between 0 and 1) describing the likelihood that the event occurs
- Applications are very broad. Example of events include:
  - measurement of glass refractive index is between 1.52 and 1.53
  - randomly chosen finger has a loop pattern
  - the proposition that the crime scene evidence and the suspect evidence have a common source is true
- Notation:
  $\Pr(E) =$ probability of the event E
  $\Pr(\bar{E}) = \Pr(E^c) =$ probability E does not occur $= 1 - \Pr(E)$
  - $\Pr(E) = 1$ - event is sure to happen
  - $\Pr(E) = 0$ - event never happens

What is probability?

- Interpretations of probability
  - long run frequency of occurrence of event
    (must be a repeatable experiment such as toss of a coin or roll of a die)
  - subjective belief of likelihood of an event
    (probability Angels win the baseball World Series)
- Where do probabilities come from?
  - empirical evidence / data
  - mathematical models
  - subjective opinion

Probability
Probability and Odds

- Probabilities are related to odds
  - odds are ratios of probabilities
  - odds in favor of event Y are defined as
    $O_f = P(Y)/P(\bar{Y}) = P(Y)/(1 - P(Y))$
  - odds against event Y are defined as
    $O_a = P(\bar{Y})/P(Y) = (1 - P(Y))/P(Y)$
  - if we are given the odds against event Y, then $P(Y) = 1/(O_a + 1)$
  - e.g., if $O_a = 4$ ("4 to 1 against") then $P(Y) = .2$
    - if you bet \$1 that Y will happen then
      20% of the time you win \$4
      80% of the time you lose \$1
      (note: you will break even in the long run)

Probability

- Probability questions can be confusing
  - results are not always intuitive
  - subtle differences in wording can lead to major differences in the answer
- Examples include:
  - The Monty Hall problem
  - The birthday problem (see next slide)
- But probability concepts are critical to reasoning about uncertainty

Test yourself

The birthday problem

- Suppose there are 50 people in a room. The probability that at least two share a birthday is closest to ...
  - 0.15
  - 0.33
  - 0.50
  - 0.95

Test yourself

The birthday problem - answer

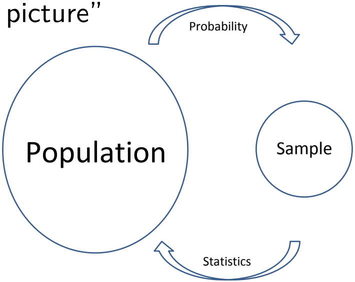
- Suppose there are 50 people in a room. The probability that at least two share a birthday is closest to ...
    - The correct answer is 0.95. This is not very intuitive!
    - Our intuition can fail because we often think about the probability that one of those people would match my birthday (about .15)
    - In fact with 50 people there are many comparisions (A with B, A with C, ...) so lots of opportunities
    - The proof works by trying to find the probability of no matching birthdays
    - The first person can be on any day (365/365), the 2nd person on 364/365 (to avoid the first), the 3rd person on 363/365, etc.
    - If you multiple 365/365 x 364/365 x 363/365 ....x 317/365 x 316/365 it turns out to be quite small

Conditional Probability

- Consider the example of an individual flying from LAX to JFK and worried about a delay
- Based on historical data we might believe that $\Pr(\text{delay}) = 0.27$
- Now suppose we learn that the weather forecast calls for thunderstorms in New York
  - May now believe the probability of a delay is higher
  - This leads to the notion of **conditional probability**, what is the probability of a delay given that thunderstorms are forecast?
  - Notation: We write $\Pr(\text{delay} \mid \text{thunderstorms})$ ...with the verical bar serving as shorthand for "given" or "conditional on" or "given the condition"
  - Perhaps we conclude $\Pr(\text{delay} \mid \text{thunderstorms}) = 0.50$

Understanding Conditional Probability

- Recall the "big picture"



- For the first statement, $\Pr(\text{delay}) = 0.27$, the population consists of all LAX-JFK flights
- For the second statement, $\Pr(\text{delay} \mid \text{thunderstorms}) = 0.50$, the population consists of LAX-JFK flights on days with forecasts of thunderstorms
- Conditional probability changes the information we have and changes the population we are talking about

Understanding Conditional Probability

- Study of sentencing of 362 black convicted murderers in Georgia in the 1980s found that 59 were sentenced to death
- Murderers categorized by race of victim and sentence received

	Death Penalty	No DP	Total
White victim	45	85	130
Black victim	14	218	232
Total	59	303	362

- $\text{P(Death Penalty)} = 59/362 = .16$

Test yourself

Conditional probability

- Murderers categorized by race of victim and sentence received

	Death Penalty	No DP	Total
White victim	45	85	130
Black victim	14	218	232
Total	59	303	362

- Suppose we focus only on crimes in which the victim was white. What is the probability that the convicted murderer received the death penalty in that case? We could write this as $P(\text{Death Penalty} \mid \text{White victim})$.
    - $45 \ / \ 130 = .35$
    - $14 \ / \ 232 = .06$
    - $45 \ / \ 59 = .76$
    - $85 \ / \ 303 = .28$
    - can't tell from the information in the table

Test yourself

Conditional probability - answer

- Murderers categorized by race of victim and sentence received

	Death Penalty	No DP	Total
White victim	45	85	130
Black victim	14	218	232
Total	59	303	362

- Suppose we focus only on crimes in which the victim was white. What is the probability that the convicted murderer received the death penalty in that case? We could write this as $P(\text{Death Penalty} \mid \text{White victim})$.
  - The correct answer is $45/130 = .35$
  - Conditional probability asks us to think about a subset of the population (in this case the 130 cases with white victims)
  - Among that reduced population the proportion receiving the death penalty is higher than in the overall population

Understanding Conditional Probability

- Study of sentencing of 362 black convicted murderers in Georgia in the 1980s found that 59 were sentenced to death
- Murderers categorized by race of victim and sentence received

	Death Penalty	No DP	Total
White victim	45	85	130
Black victim	14	218	232
Total	59	303	362

- $P(\text{Death Penalty}) = 59/362 = .16$
- $P(\text{Death Penalty} \mid \text{White Victim}) = 45/130 = .35$
- $P(\text{Death Penalty} \mid \text{Black Victim}) = 14/232 = .06$
- Note: A number of important factors are not included (e.g., context of murder)

Understanding Conditional Probability

- Consider the following data regarding the use of consecutive matching striae (CMS) as a criterion for deciding whether a pair of bullets have the same source
- Li, 2012 thesis, U Cent. Okla. - max CMS in comparing groove impressions (9mm bullets) from known matches and known non-matches

	2	3	4	5	6	7	8	Total
Known Matches	55	54	23	11	2	0	1	146
Known Non-Matches	48	11	1	0	0	0	0	60

- $\Pr(CMS \geq 5 \mid \text{known match}) = 14/146 = .10$
- $\Pr(CMS \geq 5 \mid \text{known nonmatch}) = 0/60 = .00$
- $\Pr(\text{ known match} \mid CMS = 4) = 23/24 = .96$ (note that this conditional probability depends on the mix of match/nonmatch in the sample)

Conditional Probability and Independence

- Sometimes the additional information doesn't change the probability of an event
- Famous classroom examples include coin flips, dice rolls
- Suppose I have pasta for dinner the day before my flight. Presumably ....
  $\text{Pr(flight delay} \mid \text{pasta for dinner)} = \text{Pr(flight delay)}$
- We would then say that having a flight delay is **independent** of what I had for dinner
- A well-known example of independent events in forensic science is the independence of DNA markers found on different chromosomes

Independence and the Product Rule

- We may want to know the probability that two different events both happen
  - What is the probability that my flight is delayed and my luggage is lost?
  - What is the probability that I get a head on my first coin toss and my second coin toss?
- This can be complicated to compute because the likelihood of the second event may depend on whether the first has happened. An example:
  - Pr(Yankees win World Series) = 0.15 (as of Oct 11)
  - Pr(Yankees win World Series | Judge (star) gets hurt) = 0.06
- In general the probability that two events both occur is found by $\Pr(A \text{ and } B) = \Pr(A)\Pr(B|A) = \Pr(B)\Pr(A|B)$
- If two events are independent, then there is a simple product rule. We can just multiply probabilities
  $\Pr(A \text{ and } B) = \Pr(A)\Pr(B)$

People (CA) vs Collins - a cautionary tale

- Recall that expert witness testified that if observed characteristics have specified probabilities and are independent, then probability of observing a couple matching on all characteristics is $1/12$ million
- Collins couple were found guilty
- Malcolm Collins appealed claiming the probability evidence was prejudicial
- What should the court do?

Test yourself

The Collins case

- What should the court do in the Collins appeal? Pick all of the answers that seem right.
  - Let the conviction stand. The expert is right.
  - Let the conviction stand. The expert is irrelevant, but the eyewitness testimony is compelling.
  - Overturn the conviction. There is no basis for those probabilities.
  - Overturn the conviction. The assumption of independence is not justified here.
  - Overturn the conviction. Statistics has no place in the courtroom.

People (CA) vs Collins - a cautionary tale

- Collins couple were found guilty
- Malcolm Collins appealed claiming the probability evidence was prejudicial
- California Supreme Court reversed the conviction
  - Court indicated testimony lacked an adequate foundation
    - Inadequate evidentiary foundation for probabilities
    - Inadequate proof of statistical independence
  - Court found testimony and prosecutor's use distracted the jury from its proper role

People (CA) vs Collins - a cautionary tale

- The Court's concerns
  - Prosecution did not provide any sources for the probabilities supplied
  - Need to have some empirical basis for the probabilities (e.g., $\Pr(\text{man with mustache}) = 0.47$ in France 2016)
  - Suspect that some of the characteristics are not independent (e.g., $\Pr(\text{beard} \mid \text{mustache}) = 0.91$ in France 2016)
  - Dependencies of this type will lower the probability (and make the evidence less convincing)
  - Mathematics as a distraction: the prosecution's argument provides no guidance to the jury on the critical issue of whether the Collins committed the crime
    - possibility of eyewitness error / disguise
    - possibility of more than one couple matching the description
- Interesting to note that the Court's last concern about mathematics as a distraction has been overcome; calculations like those used in the Collins case are regularly used in analyses of DNA evidence

State (CT) vs Skipper

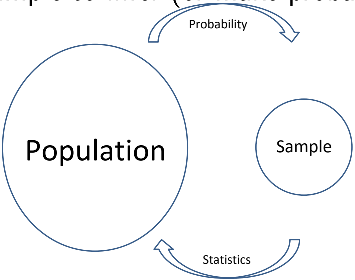
- Collins case introduces us to
  - probabilities of simple events (e.g., probability of blond woman)
  - conditional probability
  - product rule for independent events
- Much current discussion is focused on more sophisticated uses of probability
- Introduce these ideas through a second case, State (CT) vs Skipper (1994)

State (CT) vs Skipper

- Defendant charged with sexual assault
- State's expert witness reported on results of a genetic paternity test
- Expert reported a paternity index (likelihood ratio) of 3496 (probability that defendant would produce a child with the given genotype is 3496 times as large as the probability that a random male would produce such a child)
- Expert indicated the paternity index could be converted into a statistic that gave the defendant's probability of paternity
- He did so and reported the probability of paternity $= 3496/3497 = 0.9997$
- To disentagle the statistical issues in this case we need
    - Bayes' theorem (more advanced probability)
    - A framework for assessing forensic evidence

Bayes' Theorem (or Rule)

- Thomas Bayes was an English mathematician, philosopher and minister
- Famous among statisticians for his mathematical work on "inverse probability"
- Recall that in our "big picture" (below) probability tells us how to go from knowledge about the population to what we can expect to see in a sample
- Inverse probability (now known as Bayesian statistics) refers to using our observed sample to infer (or make probability statements) about the population



Bayes' Theorem - Gunshot residue example

- Consider a diagnostic test for gunshot residue on an individual
  - Let G denote the event that gunshot residue is present (we will say "not G" to denote the opposite event)
  - Let T denote the event that our diagnostic test is positive (indicates gunshot residue is present) and "not T" to indicate a negative test

True Status	Test Result T	Test Result Not T
G	True Positive	False Negative
Not G	False Positive	True Negative

  - We frequently have information about the performance of the test
  - $P(T \mid G)$ = true positive rate, sensitivity
  - $P(\text{not } T \mid \text{ not } G)$ = true negative rate, specificity
  - $p(T|\text{ not } G)$ = false positive rate
  - $p(\text{not } T|G)$ = false negative rate
  - People sometimes refer to false positives as Type I errors and false negatives as Type II errors; we will argue today that that is not appropriate language for diagnostic tests

Bayes' Theorem - Gunshot residue example

- Bayes' Theorem provides a means of taking information we have about the test (how the test performs on people whose status is known) to infer the status of an individual who has been tested
  - We know $\Pr(T \mid G)$ and $\Pr(T \mid \text{ not } G)$
  - Bayes' Theorem allows us to calculate $\Pr(G \mid T)$
- To do this **Bayes' Theorem also requires some "prior" information** about the prevalence of gunshot residue in the population of interest (i.e., $\Pr(G)$)
- A couple of key points
  - This "prior" information is important and its not clear where it should come from
  - In general, $P(T \mid G) \neq P(G \mid T)$ (sensitivity of the test is not the same as our certainty given the test result)

Bayes' Theorem - Gunshot residue example

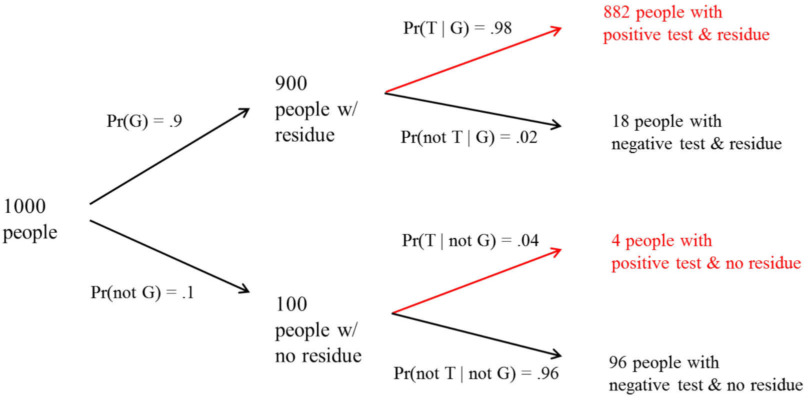
- How does it work?
  - Assume $\Pr(T \mid G) = 0.98$ (the test is very sensitive - many true positives)
  - Assume $\Pr(\text{not } T \mid \text{not } G) = 0.96$ (the test is pretty specific - low rate of false positives)
  - For now assume $\Pr(G) = 0.90$ in the population of interest
  - We test an individual and get a positive test result
  - What can we say about the probability that the individual actually has gunshot residue on them
- There is a mathematical formula for this ....

$$P(G \mid T) = \frac{P(G \text{ and } T)}{P(T)} = \frac{P(T \mid G)P(G)}{P(T|G)P(G) + P(T|\text{not } G)P(\text{not } G)} = \frac{.98 * .9}{(.98 * .9 + .04 * .1)} = .995$$

but it is easier to think about this with a picture

Bayes' Theorem - Gunshot residue example

- Suppose we have a population of 1000 individuals



- Conclusion: Note that there are 886 positive tests and 882 are "true" $\Pr(G|T) = \Pr(\text{residue} \mid \text{positive}) = 882/886 = .995$

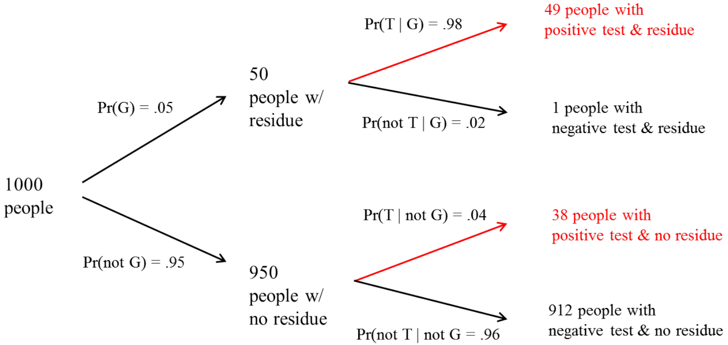
Bayes' Theorem - Gunshot residue example

- Can sometimes get surprising results from Bayes' Rule
- Return to the diagnostic test for gunshot residue example
  - assume $P(T \mid G) = .98$ (sensitivity)
  - assume $\Pr(\text{not } T \mid \text{not } G) = 0.96$ (specificity)
  - now assume $P(G) = .05$ **(low prevalence)**
    (i.e., testing in a population where gun usage is rare)

Test yourself
Bayes' Theorem

- In the setting of the previous slide (sensitivity = .98, specificity = .96, prevalence = .05), suppose the gunshot residue test returns a positive result. The probability that the individual actually has gunshot residue on their hand is closest to ....
  - 1.00
  - 0.75
  - 0.50
  - 0.25
  - 0.00

Bayes' Theorem - Gunshot residue example



- Conclusion: Note that there are 87 positive tests and 49 are "true" $\Pr(G|T) = \Pr(\text{residue} \mid \text{positive}) = 49/87 = .56$
- The prior information matters a great deal when interpreting the test result
- Same phenomenon can happen with drug testing, medical diagnostics

Conditional Probability / Bayes' Theorem in the Courtroom

- E = evidence
  - DNA markers from the crime scene sample and suspect sample
  - Measurements on glass fragments from crime scene / suspect's clothing
  - Image of bullet cartridges found at crime scene / test fire from suspect's weapon
- H_s = "same source" proposition (two samples have same source)
  H_d = "different source" proposition (two samples w/ different sources)
- Then
  $\Pr(E \mid H_s)$ = probability of seeing evidence if suspect is the source
  $\Pr(E \mid H_d)$ = probability of seeing evidence if suspect is not the source
- And
  $\Pr(H_s \mid E)$ = probability suspect is the source given the evidence
  $\Pr(H_d \mid E)$ = probability suspect is not the source given the evidence

Bayes' Theorem and the Likelihood Ratio - concept

- E = evidence
- H_s = "same source" proposition (two samples come from the same source)
  H_d = "different source" proposition (two samples come from different sources)
- Bayes' Theorem can be used to move from statements about the evidence to statements about the propositions/hypotheses
- But as with the gunshot residue it requires prior information about the likelihood of the propositions
- Bayes' Theorem can be written in several different forms. Here it is helpful to write in terms of odds (ratios of probabilities)

$$\frac{P(H_s \mid E)}{P(H_d \mid E)} = \frac{P(E \mid H_s)}{P(E \mid H_d)} \frac{P(H_s)}{P(H_d)}$$

Bayes' Theorem and the Likelihood Ratio - concept

- Bayes' Theorem

$$\frac{P(H_s \mid E)}{P(H_d \mid E)} = \frac{P(E \mid H_s)}{P(E \mid H_d)} \frac{P(H_s)}{P(H_d)}$$

- Term on far right is "a priori" odds in favor of the same source proposition (the prior information)
- Term in the middle is known as the likelihood ratio (LR) or Bayes factor (BF)
- Left hand side is "a posteriori" odds in favor of the same source proposition
- We will discuss in much more detail later
- For now the key point is the distinction between the likelihood ratio and the posterior odds

Bayes' Theorem and the Likelihood Ratio - concept

- Recall our gunshot residue example
  E = evidence = positive test
  H_s = suspect has gunshot residue
  H_d = suspect doesn't have gunshot residue
  - $LR = P(E \mid H_s)/P(E \mid H_d) = .98/.04 = 24.5$
  - In high prevalance case
    prior odds are $9:1$ and posterior odds are $220.5:1$
    (posterior probability $= .995$)
  - In low prevalance case
    prior odds are $1:19$ and posterior odds are $24.5:19$
    (posterior probability $= .56$)
  - Note: Can also compute likelihood ratio if evidence were a negative test (turns out to be $.02/.96 = 1/48$ which is not the reciprocal of the LR for the positive test)
- Much more on the likelihood ratio later in the short course

State (CT) vs Skipper - the role of prior information

- Expert witness testified that paternity index was 3496 (something like a likelihood ratio)
- Expert also offered to turn this index into a probability that Skipper was the father (.9997)
- Skipper was convicted
- He filed appeal claiming the statistical evidence was improperly admitted

Test yourself
The Skipper case

- What should the court do with the Skipper appeal?
    - Let the conviction stand because the likelihood ratio is very high
    - Let the conviction stand because the probability that Skipper is the father is .9997
    - Overturn the conviction because there is no role for statistics in this case
    - Overturn the conviction because the expert overstepped in creating a probability out of the likelihood ratio

State (CT) vs Skipper - the role of prior information

- Skipper was convicted
- He filed appeal claiming the statistical evidence was improperly admitted
- State Supreme Court found the expert's application of Bayes' Theorem was inconsistent with the presumption of innocence and remanded for new trial
    - Court determined that the conversion done by the expert to go from LR to posterior odds assumed prior probability of paternity was 0.50
    - Found this to violate presumption of innocence

State (CT) vs Skipper

- E = genetic evidence
- H_d = "defendant is the father" proposition
  H_r = "random man is the father" proposition
- Bayes' Theorem

$$\frac{P(H_d \mid E)}{P(H_r \mid E)} = \frac{P(E \mid H_d)}{P(E \mid H_r)} \frac{P(H_d)}{P(H_r)}$$

- Expert testified that $\Pr(E \mid H_d) / \Pr(E \mid H_r) = LR = 3496$ (evidence is much more likely if defendant is father than if a random man is the father)
- Expert assumed prior odds of 1-to-1 (50% probability for H_d)
- Expert computed posterior odds are 3496-to-1 which gives $\Pr(H_d \mid E) = 3496/3497 = .9997$

State (CT) vs Skipper

- The difference between $\Pr(E \mid H_d)$ and $\Pr(H_d \mid E)$ is critical!
- $Pr(E \mid H_d)$ is a statement about the probability of seeing the evidence if suspect is the father
- $\Pr(H_d \mid E)$ is a statement about the probability the defendant is the father based on the observed evidence
  - It seems like we want this quantity
  - But getting it depends on specifying the pre-evidence probability of the defendant being the father
  - The Court here found it is not appropriate for forensic expert to form pre-evidence opinions about the hypothesis of guilt/paternity
- Statements about the evidence (i.e., the components of the LR) are where the forensic expertise lies

Probability

A short recap

- Probability is the mathematical language of uncertainty
- Provides a common scale (0 to 1) for describing the chance that an event will occur
- Conditional probability is a key concept ...
  the probability of an event depends on what information is considered
- Independent events can be powerful (allows us to multiply probabilities as is common in DNA analysis)
- Bayes' Rule is a mathematical result showing how we should update our probabilities
  - leads naturally to thinking about the likelihood ratio as a summary of the evidence (more later)

Probability

Some key takeaways about probability

- What is the basis for probabilities that are provided? Is there supporting evidence?
- The assumption of independence is powerful ... but needs to be confirmed
- Probability statements need to be carefully interpreted:
  - what events are we assigning probabilities to
  - what information are we assuming to be true
- Distinction between Pr(evidence | hypothesis) vs Pr(hypothesis | evidence)