Understanding Basic Statistical Concepts
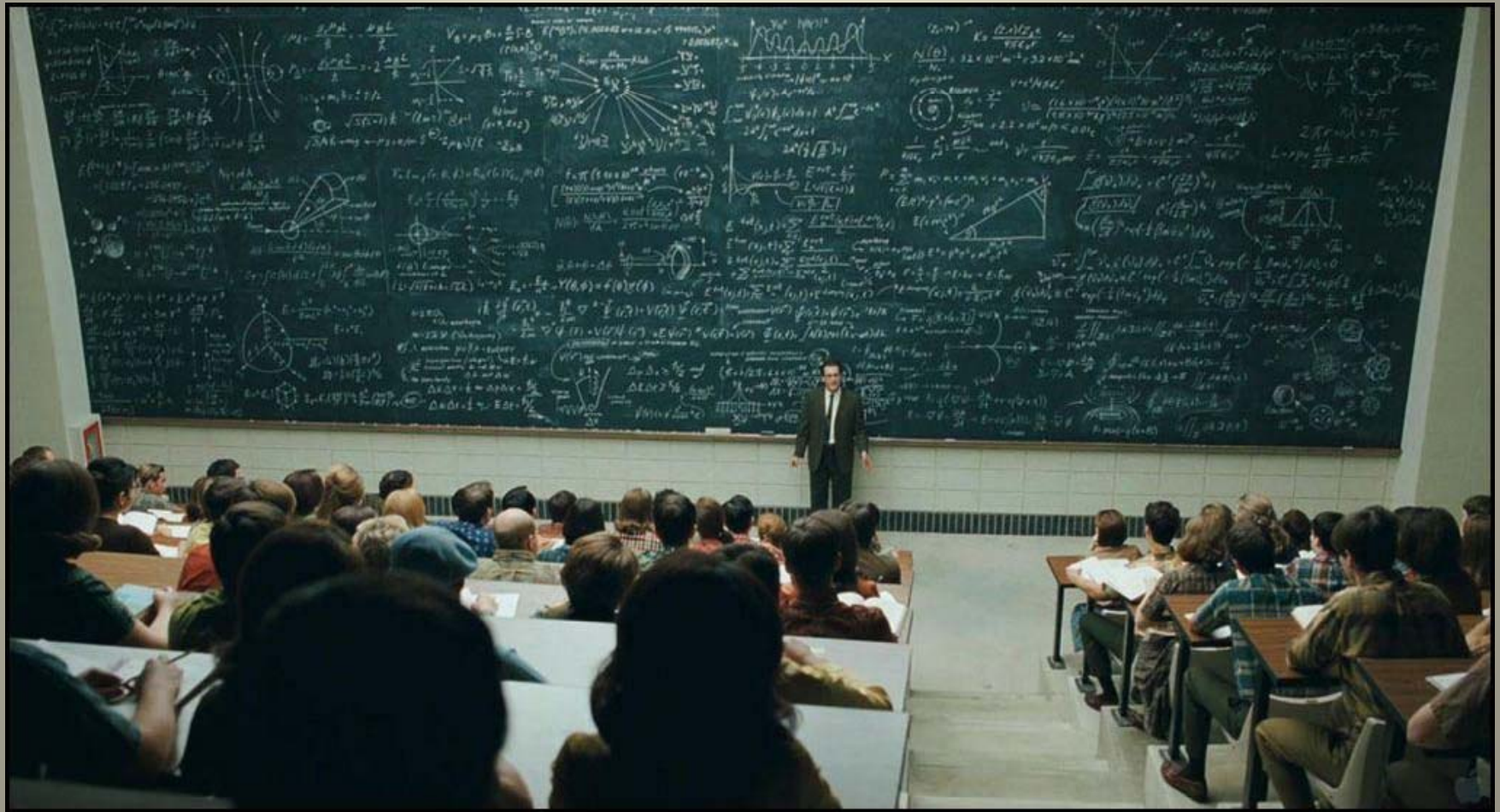# (as Applied to Fingerprints)

Presented by Michele Triplett, CLPE
at the International Association for Identification's
97th International Educational Conference
24 July 2012

Background

Why is this a hot topic?

- Recent advances in probabilistic models

- The models can be confusing

- There are different views, which makes it more confusing

- This presentation will give you BASIC information. Specifically: simple terminology and concepts.

Explaining Probabilities

Fun with Statistics

Terms

- **Statistics** – Collection and Interpretation of data.

- **Frequency** – How many times an event happens.

- **Probabilities** – The measurement of how often an event will happen when we aren't certain (rate of frequency of events).

  Probability or Rate = $\quad$ *Frequency of event*

  $\qquad\qquad\qquad\qquad\qquad$ *Total number of possible outcomes*

- **Conditional Probabilities** – Probabilities of an event happening when we know additional information.

  (Probability of someone speaking Spanish vs. Probability

  of someone speaking Spanish who is from Spain).

- **Likelihood Ratios** – Ratio of two probabilities under different circumstances, comparing the two probabilities to each other.

- **Bayesian Methods** – Conditional Probability / Bayesian Inference.

Reverend Thomas Bayes
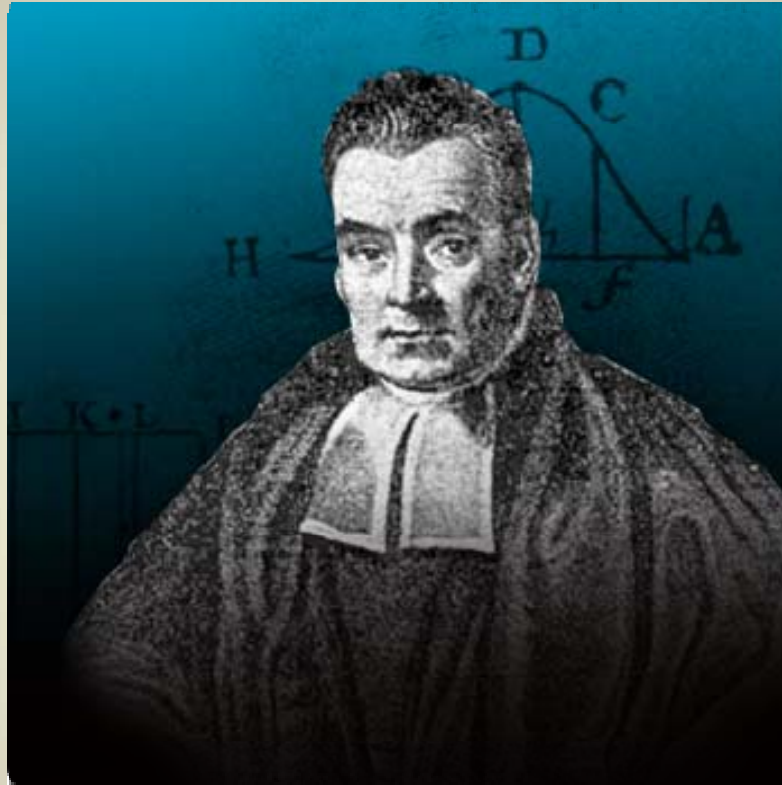
1701-1761, English Mathematician

Bayes's Theorem

- Two applications
- Both are used in situations where probabilities can't be directly calculated.
- 1) Conditional probabilities (with the ability to update the model as new _beliefs_ are established).
- 2) relate two events (known as Bayesian Inference).

Bayes Factors

scale of evidence for Bayes factors	
Bayes factor	Interpretation
B.F. < 1/10	Strong evidence for Model 2
1/10 < B.F. < 1/3	Moderate evidence for Model 2
1/3 < B.F. < 1	Weak evidence for Model 2
1 < B.F. < 3	Weak evidence for Model 1
3 < B.F < 10	Moderate evidence for Model 1
B.F. > 10	Strong evidence for Model 1

Establishing the Rate of Frequency of an Occurrence (Event)

- **Classical Probabilities** - percentage of possible outcomes (finite outcomes)

- **Empirical Probabilities -** based on a sample of observances (larger and relevant samples are better)

- **Subjective Probabilities** - user defines the probability based on <u>existing knowledge</u>. Can be dangerous if user is using a gut feeling or experience instead of <u>using current relevant knowledge.</u>

All methods are valuable for different situations. Choose the method that works best for the situation at hand.

Different Uses of Subjective Probabilities

- Yankee's winning a game based on a gut feeling.

- Yankee's winning a game based on the other team having a lot of injured players.

Accuracy of Methods

The results are as accurate as the data, the method used and the thoroughness used. This means the results of research are USELESS without knowing how they were arrived at.

Reminder

- **Classical Probabilities** – finite number of outcomes.

- **Empirical Probabilities -** based on a sample group.

- **Subjective Probabilities** - user defines the probability.

Comparing Different Methods

- Probability of it raining today
  - Classically: 1/2 outcomes = 50%
    - Assumes equally likely events
  - Empirically: Sample group = 1/100 people = 1%
    - Assumes the sample looked at accurately represents the situation
  - Subjectively: Look of the window and see it's cloudy = 85%
    - Can consider more information than supplied in the model
- Probability of me winning an election
  - Classically: 1/2 outcomes = 50%
    - Assumes equally likely events
  - Empirically: Sample group = 9/10 people = 90%
    - Assumes the sample looked at accurately represents the situation
  - Subjectively: With current knowledge I'm not running = 0%
    - Can consider more information than supplied in the model

Comparing Different Methods

- Probability of tossing a coin and getting heads
  - Classically: 1/2 outcomes = 50%
    - Assumes equally likely events
  - Empirically: Sample group of 10 tosses = could be 20%
    - Assumes the sample looked at accurately represents the situation
    - Assumes the sample group is an adequate size
  - Subjectively: = 85%
    - Can just be the gut feeling of the person
- Probability of a long commute
  - Classically: 1/2 outcomes = 50%
    - Assumes equally likely events
  - Empirically: Sample group of number of cars (low #) = 15%
    - Assumes the sample looked at accurately represents the situation
  - Subjectively: With current knowledge of fog (low #) = 90%
    - Can consider more information than supplied in the model

Comparing Different Methods

- Probability of it being busy at work on a Tuesday
  - Classically: 1/2 outcomes = 50%
    - Assumes equally likely events
  - Empirically: Sample group of Tuesdays = could be 30% (Tuesdays are usually slow)
    - Assumes the sample looked at accurately represents the situation
    - Assumes the sample group is an adequate size
  - Subjectively: = 85%
    - Knowledge of a big arrest

Benefits and Limitations

- **Classical**:
  - Benefit: Objective
  - Limitation: May not consider all relevant information
- **Empirical**:
  - Benefit: Can show that events don't have the same frequency of occurrence.
  - Limitation: Sample size may not be adequate, sample group may not be representative
- **Subjective**:
  - Benefit: Allows one to consider current relevant data
  - Limitation: Allows for not using relevant data

Example 1

Jamie gets 80% (80/100 or .8) A's in her classes. If she takes an automotive class, what is the probability of her getting an A?

Empirically: 80% (average from sample group)

Classically: 20% (1 out of 5 possible grades)

(assumes equal possibilities of each event)

Subjectively:

Could be 60% if she dislikes the topic.

Could be 99% if she likes the topic

Which Method is Best?

- Equal chance of each grade:

P(A)= 1 A out of 5 possible grades = 1/5 = 20%

- Empirically probability may show there's not an equal possibility of each event

P(A)= 80%
P(B)= 10%
P(C)= 5%
P(D)= 4%
P(F)= 1%

Classical probabilities doesn't consider all relevant information.

Empirical probabilities may show frequency of each event is different but may not be showing other current factors (desire).

Subjective probabilities allows for other relevant data.

- What if we know the lower grades were or weren't recent? Then the 'current probabilities' may be different than the 'overall probabilities'.

- Need the sample group to be relevant (current) in addition to being an adequate size.

- Need to use the best method for the situation (classical, empirical or subjective).

Example 2

P(flipping 3 tails in a row)=

Classical, fair coin = (1/2)x(1/2)x(1/2)=1/8

Classical, unfair coin = 1x1x1 = 1

Likelihood ratio of it being a fair coin = (1/8) /1

= 1/8 the likelihood of it being a fair coin or 8 times more likely that it's not a fair coin.

Empirically=?? (Try it)

Empirical probabilities are better with large sample groups (in the long run – Law of Large Numbers).

3 Tails – 1 of 8 possibilities

- HHH
- HTT
- HTH
- HHT
- TTT
- THT
- TTH
- THH

Example 3

- P(being born on January 1)=?

  Classically = 1/365

  Are birthdays equally distributed throughout the year?

  What is it dependent on (factors)?

  Which method would work best for determining the frequency of birthdates?

Frequency of an Event

Is it the same everywhere (equally distributed)?

Frequency of an Event

- The frequency of a characteristic is different in different areas (enclosure under a core v. enclosure over a core). We don't know the frequency of characteristics. We haven't found a mathematical way of establishing the frequency of ridge events.

- How many points are enough?
  - Since we haven't been able to establish the frequency of features, classical probabilities can't determine a high probability that a certain number of features are sufficiency (*one reason* classical probabilities don't work).
  - 10 common features may have less weight than 5 uncommon features. Need to consider the rarity (weight of features).
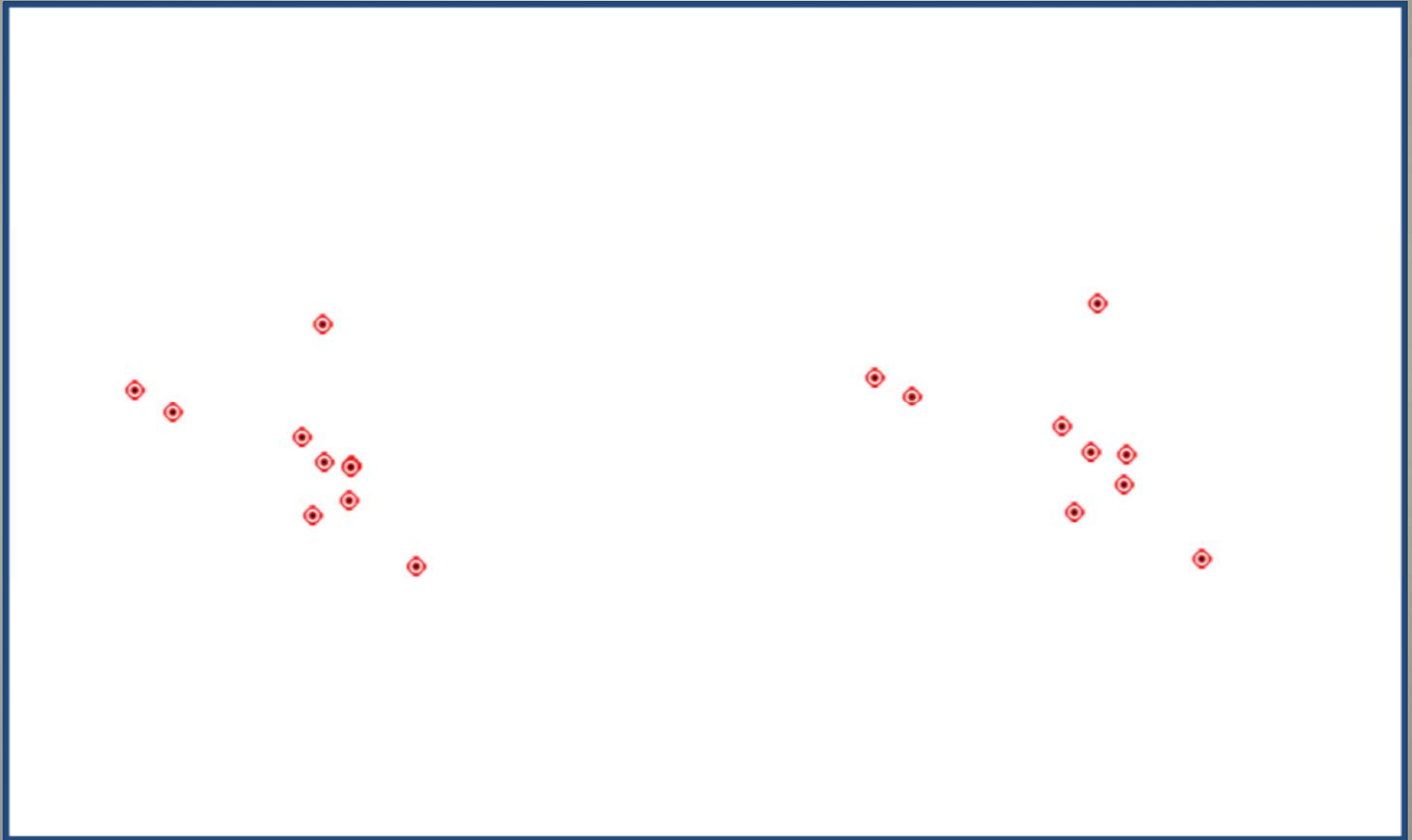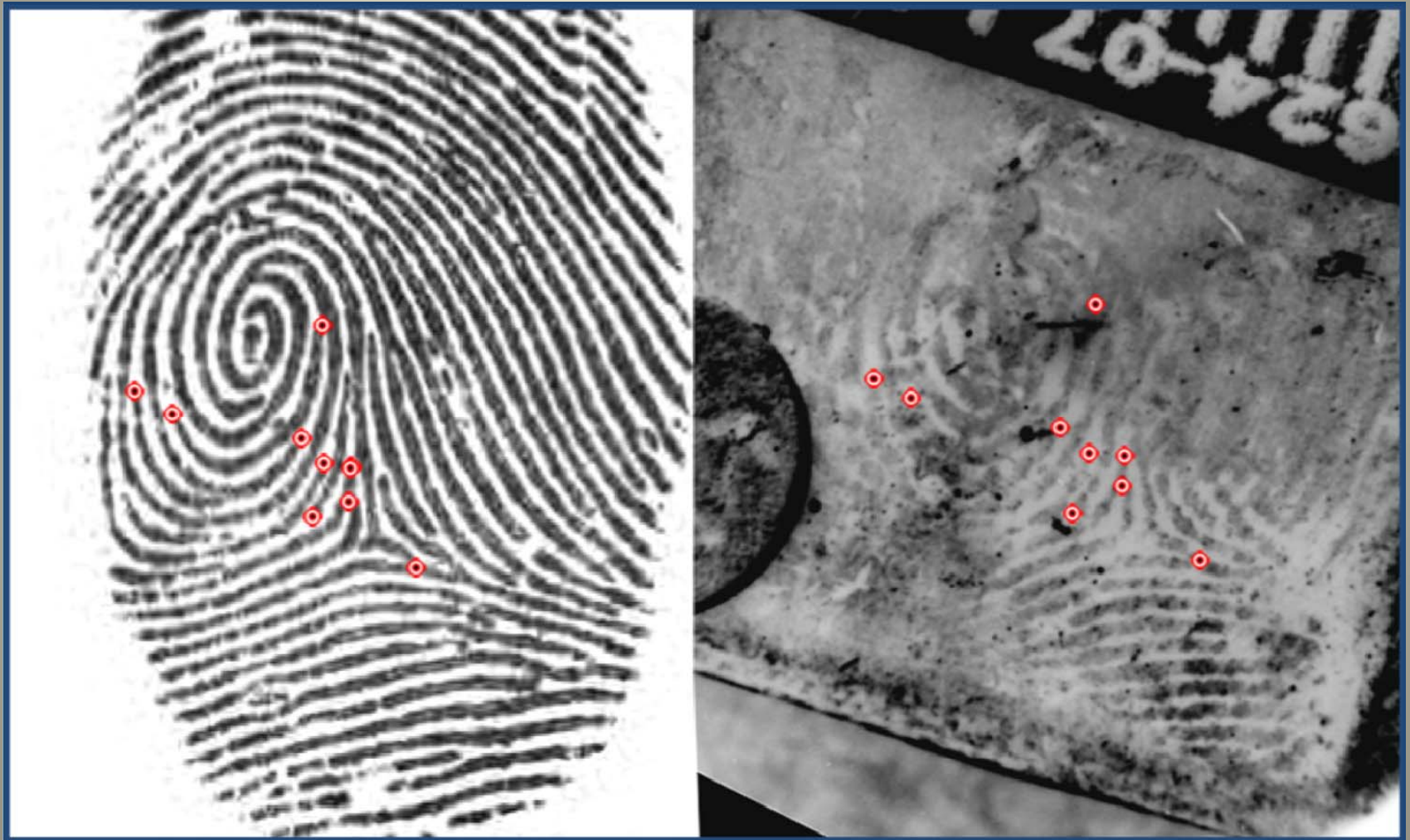
Close Non-Match
Henry Templeman
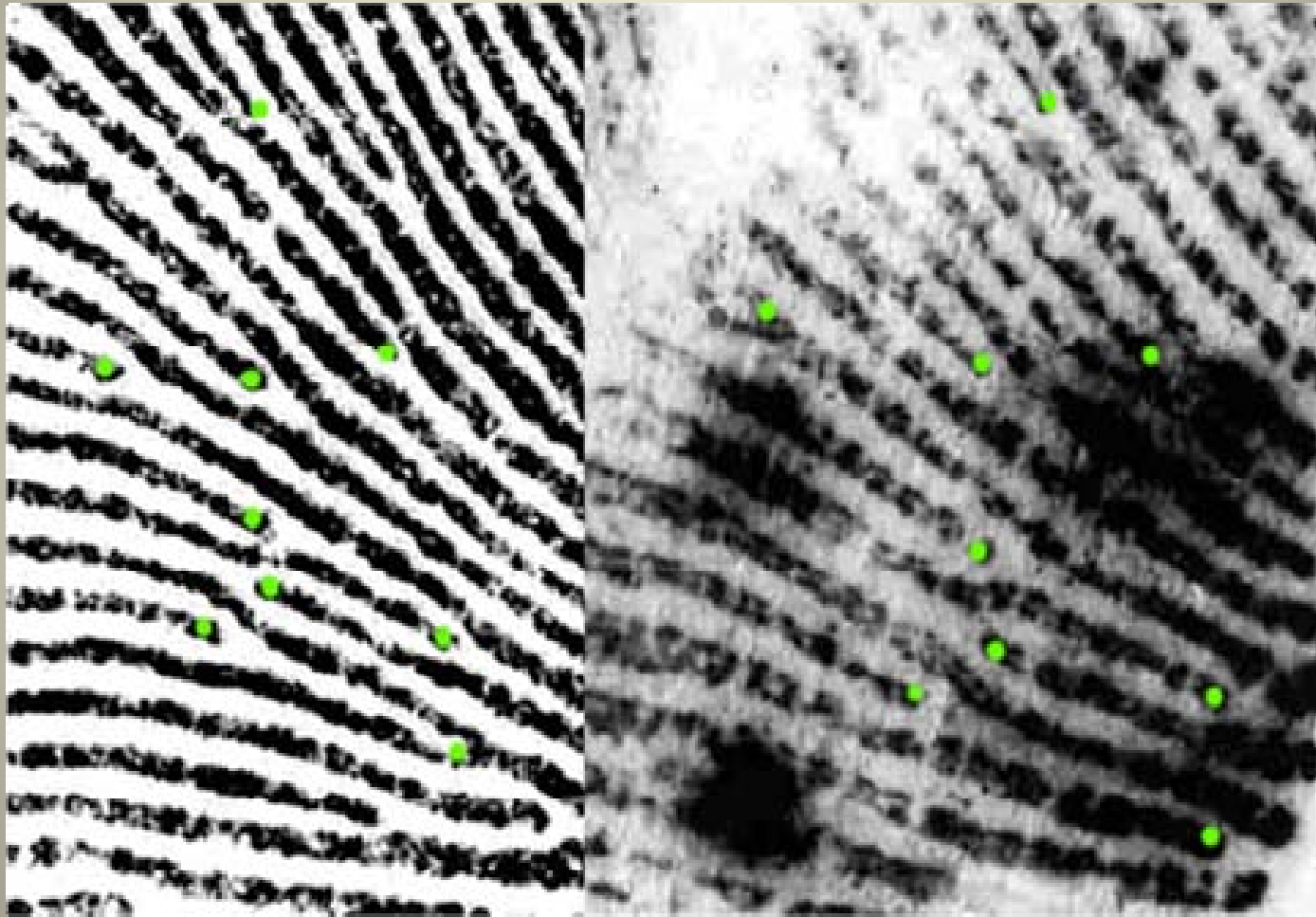
Only Considering L2…

... means you're ignoring other relevant information

Questions for the Use of Statistics

- How are probabilities and likelihood ratios going to be used?

- Use to support current conclusions (more accurate)?

- Use to arrive at additional conclusions (for current 'inconclusive decisions', lower the sufficiency needed)?

- In 2010, the IAI rescinded their resolution stating no possible, probable or likely conclusions. This was not to incorporate statistical models, it was to acknowledge that conclusions are not absolute conclusions, they are inferences (subjective probabilities).

~~The answer on 'how stats will be used' hasn't been decided yet.~~ Decided in Feb 2012.

Questions for the Use of Probabilities

- Are we looking at the probability of duplication?
- Are we looking at the probability that the conclusion is in error?

  Majority of errors are not due to the possibility of duplication, they are due to misinterpretation of data.

Misuse of Statistics

- Interpreting the data one-sidedly, ignoring relevant information (intentionally and unintentionally).

- Mathematical models may give a false sense of security in a conclusion. Need to look at:
  - Is the sample size adequate
  - Is the sample group relevance
  - Is the frequency of events a good estimate
  - Are major factors considered (perhaps: clarity, quantity, rarity (frequency), dissimilarities, intervening ridges, level 1, level 2, level 3)
  - Do the models account for misinterpretation?

Fingerprint Individuality Models

Galton (1892)

- P(fingerprint configuration)= $(1/16)$x $(1/256)$x$(1/2)^{24}$
- 24 regions
- P(correct number of ridges entering and exiting)=1/256
- P(pattern type) = 1/16
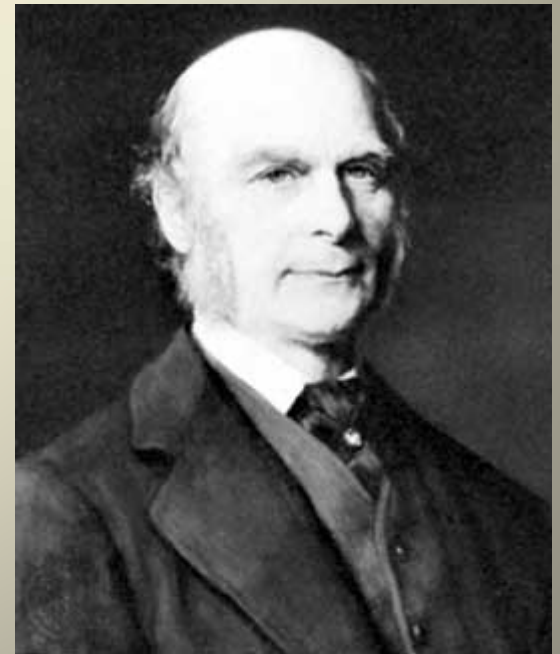
Fingerprint Individuality Models

- Henry (1900), Balthazard (1911), Bose (1917)
- 4 possible events (the events for each model were different)
- Assumed equal probability of each event
- P(event)= ¼
- P(a configuration)= $\sum$(¼) x (¼) x (¼) x (¼)… $= (¼)^n$
- For 10 minutia, one in millions
- Expectation of how many would have this configuration
  = P(a configuration)/ number of people

Fingerprint Individuality Models

- Wentworth (1918)
- P(event)= 1/50, but had no list of the 50 features
- P(a configuration)=$(1/50)^n$
- Assumed equal probability of each event

Advanced Models

- Gupta (1968)
- More possibilities of events
- Established that the frequency of events isn't equal
- 1/10 for endings and bifurcations
- 1/100 for others
- Champod (1995)
- 9 possible events
- Improved the known frequency of events
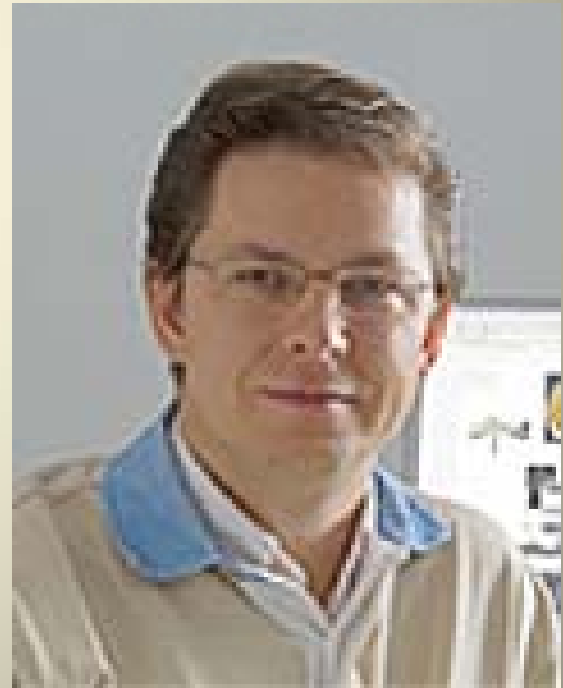
Generative Models

Generative Models incorporate ridge types (direction)

- Pankanti, Prabhakar and Jain (minutiae only) 2001
- Dass, Zhu and Jain (minutiae only) 2005
- Fang, Srihari and Srinivasan (minutiae and ridges) 2006

Differences in Models

- Use different types of features

- Use different frequencies for features

- Use different mathematical equations

- Tested their models on a different number of fingerprints (easier with computer technology)

The research has been valuable and

lead to the use in biometrics and lights-out technology!

As more information is considered, the research may be able to be used for latent prints as well.

Current Models for Partial Prints

- Christophe Champod's Model – No info
- Dactsys – Forensic Science Service (FSS) in UK
  - Original prototype developed by Cedric Neumann in 2006
- T-Model – Henry Templeman
  - Original version released in Aug 2008
- WOVI – Netherlands Forensic Institute (NFI) - No info
- Cedric Neumann's Model
  - Used for 11 points and under. Wouldn't account for the past problem cases.

Very little is published on these models.
None of these models have been
independently validated.

Current Models for Partial Prints

- Are these models an objective weight of conclusions?
  - Do they account for the major factors?
- Do they use empirical data?
  - Based on information in data bases.
- Are there subjective elements?
  - Examiners are still plotting points.

Views on Use for Partial Prints

Just starting questioning the validity of different statistical models. They haven't started evaluating models.

Estimates it'll take 2 million dollars and years to evaluate the validity of 4 models.

Individuals:

- Some think we'll be moving to probabilities or likelihood ratios quickly ; others think it is far off.
- Most people feel it's worth researching. They acknowledge we need to be open to the use of statistics but question whether models have a sound basis.

Answering Questions in Court

- Lots of models have been attempted since 1892.

- Current statistical models are useful when you have a large number of clear minutia.

- No model has been found to replicate reality accurately for low number of minutia (too much information isn't considered).

- Current models have not been externally validated.

- Our conclusions (subjective probabilities) incorporate more information than statistical models can consider at this time.

Answering in Court

- Currently, our current conclusions are based on subjective probabilities. We consider L1, L2, L3, quality, quantity, rarity, intervening ridges and dissimilarities of features.

- We make our current conclusion less subjective:

➢ By arriving at conclusions that others would arrive at (inter-subjectively testable).

➢ By being able to explain the basis behind our conclusions.

➢ By using objective data (data others would use or data others can see).

The Future

- Statistics and probabilities are not new in fingerprints but many of us have been able to avoid them.

- We're no longer able to avoid statistics since models are being recommended for use and because models are getting better (closer to being used).

Everyone will need to be able to discuss basic concepts about statistics and probabilities.

Most Important Concept

- Be able to explain the benefits and limitations of different types of probabilities.
  - Classical
  - Empirical
  - Subjective

    If you can't explain the value of the method you're using then others will exploit the limitations in an effort to discount the value of your conclusion.

- If you decide to discuss models with attorneys, you'd better be able to discuss specific factors of the models.

Terms and Concepts Not Included

- Error Rates - since errors seem so rare, a huge sample size is needed to mathematically estimate error rates... Law of Large Numbers.

- Mutually Exclusive – can't happen at the same time (boy or girl, rich or poor)

- Independent – one occurrence doesn't effect other occurrences (I: rolling dice, dependent: pulling a card from a deck of cards may be dependent on what cards were already picked).

There are different mathematical rules for different situations.

Terms and Concepts Not Included

- Sensitivity – can experiment find positive results, equation
  (or used with a different formula, incorporating false inconclusives)
- Specificity – can experiment find negative results, equation
  (or specific… not specific being non-specificity, as in L2 and L3)
  (or rarity… which is the frequency)
- Selectivity – ability to distinguish small differences
  (or used in place of the specificity equation)
  (or used with a slightly different equation, incorporating false inconclusives)
  (or rarity… which is the frequency)

These terms are used differently in chemistry and in statistics. Since fingerprints have both elements, people are using the same words for different meanings. Analysts need to consider how an author is using the word.

Quote of the Day

Statisticians are people who like math but don't have the personality to be accountants.

Questions?