



Introduction to Error Rates

The purpose of this document is to provide an introductory knowledge of statistics for IFIS members. It will discuss important topics such as basic probability, error rates and confidence intervals. It will also explain current issues such as the '1 in 18' error rate as stated by PCAST. The overall goal is to provide members with a general understanding of the statistical methodology currently relevant in the field. It is not a substitute for formal training in statistics. For easier understanding some topics have been 'watered down' so that the concept can be understood without getting into rigorous mathematics.

If you are not interested in reading the entire report look at the summary section at the end. It provides some brief points of the things you need to know.

Contents

1 Probability Basics	4
1.1 Frequency Tables	5
1.2 Joint Probability	6
1.3 Conditional Probability	8
2 Performance Measures and Error Rates.....	9
3 Confidence Intervals	14
4 Error in PCAST Report related to the Miami-Dade Study	16
5 Summary	18
6 References	19

1 Probability Basics

Probability is the representation of the likelihood that a certain event will occur. The probability of an event is mathematically represented as, $P(E = \text{'some event'})$

For example, if we let E represent 'the weather tomorrow', then $P(E = \text{'rain'})$ is 'the probability that the weather tomorrow is rain'. Don't get too fixated on the letter 'E', this is just used to represent an event, any letter can be used. Another example could be, let X represent 'the outcome of the lottery', then $P(X = \text{'win'})$ is 'the probability that you win the lottery'. Or we could say let DEC = 'the outcome of a comparison decision', then $P(\text{DEC} = \text{'Identification'})$ is 'the probability that the outcome of a comparison decision is an Identification'.

Probabilities are represented by decimal values or fractions between 0 and 1 or as percentages between 0% and 100%. The percentages are equivalent to the decimal values multiplied by 100, i.e. $1/2 = 0.5 = 50\%$, $1/4 = 0.25 = 25\%$, $0 = 0\%$, $1 = 100\%$, etc.

Putting it all together, we could say, let F represent 'the outcome of a coin flip', then $P(F = \text{'heads'}) = 1/2 = 0.5 = 50\%$.

Typically, probability can be interpreted as:

$$\text{Probability of an event} = P(E = \text{'event'}) = \frac{\text{Number of times the event occurred}}{\text{Number of times the event could have occurred}}$$

Some examples are:

$$\text{Probability of tails} = \frac{\text{Number of times the coin landed tails up}}{\text{Number of times the coin was flipped}}$$

$$\text{Probability of a male being over 6'} = \frac{\text{Number of males over 6'}}{\text{Number of males}}$$

$$\text{Probability of a diamond given the card is red} = \frac{\text{Number of diamonds}}{\text{Number of red cards}}$$

We can also represent probability as the odds. An example of the odds of something occurring could be '1 in 5' which means we expect the event to occur once in five trials. Another example could be '9 in 10' which means we expect the event to occur nine times in ten trials. In order to find the '1 in X' odds divide 1 by the probability value. Here are a few examples,

$$\frac{1}{0.25} = 4, \quad \frac{1}{0.1} = 10, \quad \frac{1}{0.05} = 20, \quad \frac{1}{0.0423} \approx 23.6 \approx 24, \quad \frac{1}{0.0012} \approx 833.3 \approx 833$$

From there the odds are 1 in (the number you just calculated), i.e.

$$1 \text{ in } 4, \quad 1 \text{ in } 10, \quad 1 \text{ in } 20, \quad 1 \text{ in } 24, \quad 1 \text{ in } 833$$

1.1 Frequency Tables

Frequency tables are constructed by counting the number of times a certain event occurs. Suppose we did an experiment and tabulated our results. Let DEC represent ‘the outcome of a comparison decision’ and GT represent ‘the ground truth of the comparison’.

	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

In our pseudo-experiment, for example, there were 10 cases where an Identification was called on a different source trial. The goal is to be able to generate probabilities from this table. We begin with a simple example.

Example

Let’s look at the possible outcomes of comparison decisions. We let DEC represent ‘the outcome of a comparison decision’. Note: These numbers are not based on real data; they are just to help illustrate the concept.

DEC	Identification	Inconclusive	Exclusion
	76	23	45

What is the probability of calling an Identification?

$$P(\text{DEC} = \text{'Identification'}) = \frac{76}{144} \approx 0.53 = 53\%$$

Note: The 144 in the denominator is the sum of all the cells, i.e. 76+23+45=144

DEC	Identification	Inconclusive	Exclusion
	76	23	45

What is the probability of calling an Inconclusive?

$$P(\text{DEC} = \text{'Inconclusive'}) = \frac{23}{144} \approx 0.16 = 16\%$$

DEC	Identification	Inconclusive	Exclusion
	76	23	45

What is the probability of calling an Identification or an Exclusion?

$$P(DEC = \text{'Identification'} \text{ or } DEC = \text{'Exclusion'}) = \frac{121}{144} \approx 0.84 = 84\%$$

DEC	Identification	Inconclusive	Exclusion
	76	23	45

What is the probability of **not** calling an Inconclusive?

$$P(DEC \neq \text{'Inconclusive'}) = \frac{121}{144} \approx 0.84 = 84\%$$

DEC	Identification	Inconclusive	Exclusion
	76	23	45

1.2 Joint Probability

Joint probability describes the relationship between two or more events. We will only discuss joint probability of two events as this is what is necessary in order to generate error rates. Below we show a simple example.

Example

Let DEC represent 'the outcome of a comparison decision' and GT represent 'the ground truth of the comparison'. There are a total of 1000 trials.

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

What is the probability of calling an Identification **and** the ground truth is same source?

$$P(DEC = \text{'Identification'} \text{ and } GT = \text{'Same Source'}) = \frac{500}{1000} = 0.5 = 50\%$$

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

What is the probability of calling an Exclusion **and** the ground truth is different source?

$$P(DEC = \text{'Exclusion'} \text{ and } GT = \text{'Different Source'}) = \frac{200}{1000} = 0.2 = 20\%$$

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

What is the probability of calling an Identification?

$$P(DEC = \text{'Identification'}) = \frac{500 + 10}{1000} = \frac{510}{1000} = 0.51 = 51\%$$

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

What is the probability of the ground truth being different source?

$$P(GT = \text{'Different Source'}) = \frac{10 + 120 + 200}{1000} = \frac{330}{1000} = 0.33 = 33\%$$

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

Each individual cell in the table represents the joint frequency of two events.

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

The cell highlighted in green in the above image is the joint frequency that an Identification is called and the trial is same source, i.e. $P(DEC = \text{'Identification'} \text{ and } GT = \text{'Same Source'})$. But how do we get the probability of only one event, for example the probability that an Identification was called regardless of the ground truth?

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

You simply sum all of the values in the Identification column. This combines all the possible cases when an Identification can occur.

What about the probability that the ground truth is different source regardless of the decision?

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

Exact same idea as before, now we just focus on the different source row. If we sum the values in the row and divide by 1000 (the total number of trials in the experiment) we get 0.33.

1.3 Conditional Probability

Conditional probability is key for understanding error rates. A word which normally appears when looking at conditional probability is, 'given'. For example, suppose we asked, 'what is the probability the Winnipeg Jets beat the Tampa Bay Lightning' as opposed to 'what is the probability the Winnipeg Jets beat the Tampa Bay Lightning **given** that the Winnipeg Jets are up 4 goals in the third period'. The condition provides additional information which changes how you perceive the scenario. A couple more cases could be:

'Probability you get a flush' vs 'Probability you get a flush **given** that you were dealt two hearts'.

'Probability you pass the exam' vs 'Probability you pass the exam **given** you studied for 8 hours'.

'Probability you conclude an Identification' vs 'Probability you conclude an Identification **given** that the ground truth was different source'.

Mathematically, conditional probability is written as $P(A = \text{'some event'} \mid B = \text{'some event'})$. The '|' can be read as 'given'.

For example if we let, DEC represent 'the outcome of a comparison decision' and GT represent 'the ground truth of the comparison', then, $P(\text{DEC} = \text{'Identification'} \mid \text{GT} = \text{'Different Source'})$ is read as 'the probability you conclude an Identification **given** that the ground truth was different source', which is also called a False Positive, more on that later. We provide an example in the next section of how to calculate conditional probabilities from a frequency table.

2 Performance Measures and Error Rates

Understanding how error rates are calculated should be very straightforward if you understand conditional probability. Below we provide a list of the most frequently discussed performance measures with respect to fingerprint identification.

True Positive Rate (True Identification Rate) = Probability of calling an Identification given the ground truth is same source.

True Negative Rate (True Exclusion Rate) = Probability of calling an Exclusion given the ground truth is different source.

False Positive Rate (False Identification Rate) (FPR) = Probability of calling an Identification given the ground truth is different source.

False Negative Rate (False Exclusion Rate) (FER) = Probability of calling an Exclusion given the ground truth is same source.

We provide a simple example and calculate each of the performance measures. For the following example we remove the Inconclusive cases for simplicity.

	Identification	Exclusion
Same Source	250	20
Different Source	30	200

Example

$$\text{True Positive Rate} = P(\text{DEC} = \text{'Identification'} \mid \text{GT} = \text{'Same Source'}) = \frac{250}{270}$$

DEC GT	Identification	Exclusion
Same Source	250	20
Different Source	30	200

The condition changes the value in the denominator of the fraction. In the case above we were given that the ground truth was same source, as a result we only focus on the same source row. Then, out of the same source trials, we calculate the number of cases where the opinion was an identification.

$$\text{True Negative Rate} = P(\text{DEC} = \text{'Exclusion'} \mid \text{GT} = \text{'Different Source'}) = \frac{200}{230}$$

DEC GT	Identification	Exclusion
Same Source	250	20
Different Source	30	200

$$\text{False Positive Rate} = P(\text{DEC} = \text{'Identification'} \mid \text{GT} = \text{'Different Source'}) = \frac{30}{230}$$

DEC GT	Identification	Exclusion
Same Source	250	20
Different Source	30	200

$$\text{False Negative Rate} = P(\text{DEC} = \text{'Exclusion'} \mid \text{GT} = \text{'Same Source'}) = \frac{20}{270}$$

DEC GT	Identification	Exclusion
Same Source	250	20
Different Source	30	200

When Inconclusives are present there are two options for calculating some of the performance measures, which are to include the Inconclusives or exclude the Inconclusives.

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

We demonstrate using the frequency table above calculating the False Positive Rate two different ways. First we demonstrate the False Positive Rate with the Inconclusives followed by removing the Inconclusives.

$$\text{False Positive Rate} = P(\text{DEC} = \text{'Identification'} \mid \text{GT} = \text{'Different Source'}) = \frac{10}{330}$$

DEC GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

False Positive Rate

$$= P(DEC = 'Identification' \mid GT = 'Different Source' \text{ and } DEC \neq \text{Inconclusive}) = \frac{10}{210}$$

DEC \ GT	Identification	Inconclusive	Exclusion
Same Source	500	100	70
Different Source	10	120	200

By not including the Inconclusives the False Positive Rate increases. Typically, False Positive Rates will not include Inconclusive decisions. The author will usually state whether or not Inconclusives were included in the calculation.

Example

Below we have the table of decisions from the FBI/Noblis Black Box study [1] appendix. The frequency table is a little on the messy side, but the information is there to get error rates.

Comparison Decision	Latent Value	Total	Mates	Non-mates	% of mated pairs			% of non-mated pairs		
					PRES	CMP	VID	PRES	CMP	VID
(not compared)	NV	3,947	3,389	558	29.3%			10.1%		
Exclusion	VEO	486	161	325	1.4%	2.0%		5.9%	6.5%	
Exclusion	VID	4,072	450	3,622	3.9%	5.5%	7.5%	65.3%	72.7%	88.7%
Inconclusive	VEO	2,596	2,019	577	17.4%	24.7%		10.4%	11.6%	
Inconclusive	VID	2,311	1,856	455	16.0%	22.7%	31.1%	8.2%	9.1%	11.1%
Individualization	VEO	40	40	0	0.3%	0.5%		0.0%	0.0%	
Individualization	VID	3,669	3,663	6	31.6%	44.7%	61.4%	0.1%	0.1%	0.1%
Totals		17,121	11,578	5,543	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Total comparisons	Value (either)	13,174	8,189	4,985						
Total comparisons	VID	10,052	5,969	4,083						

The authors claim from this data that the False Positive Rate for VID (Value for Identification, a label used during the analysis phase for some organizations) comparisons is 0.1% and False Negative Rate for VID comparisons is 7.5%. It's also important to note that the authors included the Inconclusives in their calculations. We should now have the knowledge to check those values ourselves from the data provided above.

First we want to calculate the False Positive Rate for VID comparisons. So, we know that for the False Positive Rate we are looking at the number of different source (non-mates) trials, and now also only those which were considered VID. From the Black Box data this number works out to be: 3622 + 455 + 6 = 4083 (The FBI/Noblis included Inconclusives in their calculations). Out of those 4083 cases 6 were an Identification, therefore,

False Positive Rate

$$= P(DEC = 'Identification' | GT = 'Different Source' and Val = 'VID')$$

$$= \frac{6}{4083}$$

Comparison Decision	Latent Value	Total	Mates	Non-mates	% of mated pairs			% of non-mated pairs		
					PRES	CMP	VID	PRES	CMP	VID
(not compared)	NV	3,947	3,389	558	29.3%			10.1%		
Exclusion	VEO	486	161	325	1.4%	2.0%		5.9%	6.5%	
Exclusion	VID	4,072	450	3,622	3.9%	5.5%	7.5%	65.3%	72.7%	88.7%
Inconclusive	VEO	2,596	2,019	577	17.4%	24.7%		10.4%	11.6%	
Inconclusive	VID	2,311	1,856	455	16.0%	22.7%	31.1%	8.2%	9.1%	11.1%
Individualization	VEO	40	40	0	0.3%	0.5%		0.0%	0.0%	
Individualization	VID	3,669	3,663	6	31.6%	44.7%	61.4%	0.1%	0.1%	0.1%
Totals		17,121	11,578	5,543	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Total comparisons	Value (either)	13,174	8,189	4,985						
Total comparisons	VID	10,052	5,969	4,083						

Our final value of 6/4083 works out to be 0.0014695 = 0.14595% and with some convenient rounding turns out to be 0.1%, as stated in the report.

What if we remove the Inconclusives as was done in the PCAST report? The new False Positive Rate becomes,

False Positive Rate

$$= P(DEC = 'Identification' | GT = 'Different Source' and Val = 'VID and DEC \neq 'Inconclusive')$$

$$= \frac{6}{3628}$$

Comparison Decision	Latent Value	Total	Mates	Non-mates	% of mated pairs			% of non-mated pairs		
					PRES	CMP	VID	PRES	CMP	VID
(not compared)	NV	3,947	3,389	558	29.3%			10.1%		
Exclusion	VEO	486	161	325	1.4%	2.0%		5.9%	6.5%	
Exclusion	VID	4,072	450	3,622	3.9%	5.5%	7.5%	65.3%	72.7%	88.7%
Inconclusive	VEO	2,596	2,019	577	17.4%	24.7%		10.4%	11.6%	
Inconclusive	VID	2,311	1,856	455	16.0%	22.7%	31.1%	8.2%	9.1%	11.1%
Individualization	VEO	40	40	0	0.3%	0.5%		0.0%	0.0%	
Individualization	VID	3,669	3,663	6	31.6%	44.7%	61.4%	0.1%	0.1%	0.1%
Totals		17,121	11,578	5,543	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Total comparisons	Value (either)	13,174	8,189	4,985						
Total comparisons	VID	10,052	5,969	4,083						

INTRODUCTION TO ERROR RATES

Our final value of 6/3628 works out to be 0.0016538 = 0.16538% which means the odds are 1 in 604.66. PCAST reported the odds for the FBI/Noblis Black Box study as 1 in 604.

What about the False Negative Rate? A False Negative means we are looking at same source (mated) cases. As well, we restrict to the cases where the Analysis decision was VID.

$$\text{False Negative Rate} = P(\text{DEC} = \text{'Exclusion'} \mid \text{GT} = \text{'Same Source'} \text{ and } \text{Val} = \text{'VID'})$$

$$= \frac{450}{5969}$$

Comparison Decision	Latent Value	Total	Mates	Non-mates	% of mated pairs			% of non-mated pairs		
					PRES	CMP	VID	PRES	CMP	VID
(not compared)	NV	3,947	3,389	558	29.3%			10.1%		
Exclusion	VEO	486	161	325	1.4%	2.0%		5.9%	6.5%	
Exclusion	VID	4,072	450	3,622	3.9%	5.5%	7.5%	65.3%	72.7%	88.7%
Inconclusive	VEO	2,596	2,019	577	17.4%	24.7%		10.4%	11.6%	
Inconclusive	VID	2,311	1,856	455	16.0%	22.7%	31.1%	8.2%	9.1%	11.1%
Individualization	VEO	40	40	0	0.3%	0.5%		0.0%	0.0%	
Individualization	VID	3,669	3,663	6	31.6%	44.7%	61.4%	0.1%	0.1%	0.1%
Totals		17,121	11,578	5,543	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Total comparisons	Value (either)	13,174	8,189	4,985						
Total comparisons	VID	10,052	5,969	4,083						

Our final value of 450/5969 works out to be 0.0753895 = 7.53895% which after rounding is 7.5%. There were no mistakes in the calculation of the FPR or FNR for the FBI/Noblis study in the PCAST report. Below we provide another simple example of error rate calculations.

Example

DEC \ GT	Identification	Inconclusive	Exclusion
Same Source	370	50	180
Different Source	15	335	550

Performance Measure	Inconclusives?	Result
True Positive Rate	Yes	370/600
True Negative Rate	Yes	550/900
False Positive Rate	Yes	15/900
False Negative Rate	Yes	180/600
True Positive Rate	No	370/550
True Negative Rate	No	550/565
False Positive Rate	No	15/565
False Negative Rate	No	180/550

3 Confidence Intervals

The purpose of a confidence interval is to provide a band or a range where the true result is likely to be based on the data collected from an experiment. It is important to note that the confidence interval can only speak to the parameter of interest (false positive rate) under the conditions of an experiment. For example, if we were to estimate the average income for a town, we cannot generalize that average income to all of Canada, or if we were to estimate a False Positive Rate at an international level, we cannot specify it to a particular country or agency.

For example, suppose we want to know the average income in Sunnyville, Canada. We conduct a survey and estimate the average income to be \$40,000. But of course that value is only an estimate based off of the data we collected. So we will also provide the confidence interval which will provide a lower and upper bound on the estimate. That way we can make statements like: 'there is a 95% chance that the true average income in Sunnyville is between \$35,000 and \$45,000' or 'there is a 90% chance that the true average income in Sunnyville is between \$37,500 and \$42,500'.

A confidence interval has four components:

- The parameter of interest (average income, false positive rate, etc.)
- The level of confidence
- The variability of the experiment
- The sample size of the experiment

The most common confidence interval is very roughly constructed like this,

$$\text{Lower Bound} = \text{parameter of interest} - \frac{(\text{level of confidence}) \times (\text{variability})}{\sqrt{\text{sample size}}}$$

$$\text{Upper Bound} = \text{parameter of interest} + \frac{(\text{level of confidence}) \times (\text{variability})}{\sqrt{\text{sample size}}}$$

Let's go back to our average income example, and show how the confidence interval could be calculated. Assume that,

- Parameter of interest = \$40,000
- Level of confidence = 95%, which translates to a value of 1.96 (not explained here)
- Variability = \$25,000
- Sample size = 500

Using these values, the resulting confidence interval would be,

$$\text{Lower Bound} = 40000 - \frac{1.96 \times 25000}{\sqrt{500}} = \$37,809$$

$$\text{Upper Bound} = 40000 + \frac{1.96 \times 25000}{\sqrt{500}} = \$42,191$$

And what we would conclude is that there is a 95% chance that the true average income in Sunnyville, Canada is between \$37,809 and \$42,191.

Aside: It is very important to note at this point that technicalities are being sacrificed in order for easier understanding. For example, the last statement 'there is a 95% chance that the true average income in Sunnyville, Canada is between \$37,809 and \$42,191', is incorrect from a statistical perspective. A statistically correct statement would be 'if we were to repeat the experiment under the same conditions an infinite number of times 95% of the confidence intervals constructed from the experiments will contain the true value'. The 'correct' statement is usually a little more confusing for people to understand. For a general audience the interpretation of 'there is a 95% chance that the true average income in Sunnyville, Canada is between \$37,809 and \$42,191' is fine. As well, in the confidence interval calculation, the standard deviation is used as the measure of variability. Here we simply refer to it as 'variability'.

4 Error in PCAST Report related to the Miami-Dade Study

Suppose we did an error rate study with 5 trials on which no errors occurred, i.e. an error rate of 0%. Now we will calculate the confidence interval for the 0% error rate using the same methodology as PCAST (Binomial Exact confidence interval)¹:

$$\text{Lower Bound} = 0\%$$

$$\text{Upper Bound} = 52\%$$

The confidence interval returned a lower bound of 0% and an upper bound of 52%. There is nothing wrong with the confidence interval, but there could be something wrong with how someone interprets it. PCAST [2] would have stated that the error rate could be as high as 52% or 1 in 2...similar to what they said with the Miami-Dade study [3], "...that the error rate could be as high as 1 error in 18 cases" [2] (the '1 in 18' is based on a calculation error as explain below)². This statement is extremely misleading. A confidence interval does not tell you how high the error rate could be; it provides an area where the true value probably lies. So a more correct statement for our example would be, 'there is a 95% chance that the true error rate is between 0% and 52%'.

A confidence interval can provide some idea for the reliability of an estimate. Having a lower bound of 0% and an upper bound of 52% means that we really don't know that much about the error rate since the confidence interval is so wide. Instead suppose we did an experiment where we had 10 errors on 7000 trials which would result in an error rate of 0.14%. The resulting confidence interval would be³:

$$\text{Lower Bound} = 0.07\%$$

$$\text{Upper Bound} = 0.26\%$$

This is a much tighter confidence interval, subsequently, we can have greater assurance in our estimated error rate of 0.14%, under the conditions of the experiment.

We will now go through step-by-step on how the '1 in 18' mistake was made. First we look at the frequency table (Table 4) that was used from the Miami-Dade study on page 53.

ACE Examinations [Miami-Dade Table 4]			
	Same Source	Difference Source	Totals
Identification	2,457	42	2,499
Inconclusive	446	403	849
Exclusion	235	953	1,188
Totals	3,138	1,398	4,536

¹ To calculate this confidence interval go to <http://www.sample-size.net/confidence-interval-proportion/> and enter 5 for N, 0 for x, and 95 for CL.

² Refer to Wilkinson, D.; Richard, D.; Hockey, D. Expert Fingerprint Testimony Post-PCAST—A Canadian Case Study. J. For. Ident. 2018, 68 (3), 299-331.

³ To calculate this confidence interval go to <http://www.sample-size.net/confidence-interval-proportion/> and enter 7000 for N, 10 for x, and 95 for CL. It rounds the values from 0.0007 to 0.001 and 0.0026 to 0.003

The authors of the study indicate from Table 4 [3] that there were 42 Identifications on different source trials. This is incorrect as later in Table 11 [3] on page 56 indicates that there are only 3 Identifications on 'source not present' (different source) trials.

ACE Trials Participant Decisions [Miami-Dade Table 11]							
Source Present (Y/N)	# of Latent Prints	# of Decisions	Correct Identifications	Erroneous Identifications	Inconclusives	Correct Exclusions	Erroneous Exclusions
Yes	56	3,177	2,457	39	446	N/A	235
No	24	1,359	N/A	3	403	953	N/A
Totals	80	4,536	2,457	42	849	953	235

The authors calculate the False Positive Rate, based on Table 4, with Inconclusives as $42/1398 = 3\%$ and the False Positive Rate without Inconclusives as $42/995 = 4.2\%$, however, the 42 False Positives as seen in Table 11 occurred on both same source and different source trials⁴. Under these circumstances, false positives could have occurred on any trial in the experiment, subsequently, the False Positive Rate with Inconclusives should be $42/4536 = 0.93\%$ and without Inconclusives should be $42/3687 = 1.14\%$ [4].

Furthermore, PCAST calculated confidence intervals for the incorrect False Positive Rates from the Miami-Dade study:

One Sided Confidence Interval for False Positive Rate with Inconclusives: 3.87%

One Sided Confidence Interval for False Positive Rate without Inconclusives: 5.43%

The upper bound of the one sided confidence interval for the False Positive Rate without Inconclusives is 5.43% (0.0543) which translates to 1 error in every 18.43 ($1/0.0543$) trials, or the now infamous '1 in 18'.

PCAST then concludes, "The estimated rate corresponds to 1 error in 24 cases, with the upper bound indicating that the rate could be as high as 1 error in 18 cases" [2]. As stated earlier this is not a correct interpretation of a confidence interval and is misleading.

In conclusion, the '1 in 18' error rate is not correct, as it really is the upper bound of a confidence interval based on an incorrect False Positive Rate. This error was corrected in both an RCMP article published in the Journal of Forensic Identification [4] and a response letter from the Organization of Scientific Area Committees Friction Ridge Subcommittee [5].

⁴ Refer to R. v. Borynk Article as published in Wilkinson, D.; Richard, D.; Hockey, D. Expert Fingerprint Testimony Post-PCAST—A Canadian Case Study. J. For. Ident. 2018, 68 (3), 299-331 for more information. In the study the examiners were comparing a latent to sets of ten exemplars. If the same source exemplar was present, the trial was deemed as 'source present'. The examiner could still incorrectly identify the latent to the wrong card or one of the nine other exemplars from the correct source.

5 Summary

- There is no error rate study which replicates the entire fingerprint comparison process.
- There is no error rate which can be applied to the case at hand.
- Each error rate is subject to the conditions of the experiment from which it was derived.

- A confidence interval provides a measure of uncertainty for a parameter which is being estimated under the conditions of an experiment. For example, it could provide a measure of uncertainty for an estimated false positive rate.
- A confidence interval does not tell you that an error rate “...could be as high as...” some value.
- A confidence interval is affected by the parameter which you are estimating (false positive rate) as well as the variability in your data, the sample size of your data, and the level of confidence selected.

- The ‘1 in 18’ is the upper bound of a one-sided confidence interval for an incorrectly calculated false positive rate.
- The ‘1 in 18’ is not a valid error rate.

- Error rate studies should be used to validate or invalidate a discipline and not attempt to estimate the error rate in practise.

6 References

[1] Ulery, B. T.; Hicklin, R. A.; Buscaglia, J.; Roberts, M. A. Accuracy and Reliability of Forensic Latent Fingerprint Decisions. *PNAS* **2011**, *108* (19), 7733–7738.

[2] President’s Council of Advisors on Science and Technology (PCAST). Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods; Executive Office of the President’s Council of Advisors on Science and Technology: Washington, D.C., 2016.

[3] Pacheco, I.; Cerchiai, B.; Stoiloff, S. Miami-Dade Research Study for the Reliability of the ACE-V Process: Accuracy and Precision in Latent Fingerprint Examinations; NCJRS Doc. No. 248534, Dec, 2014.

[4] Wilkinson, D.; Richard, D.; Hockey, D. Expert Fingerprint Testimony Post-PCAST—A Canadian Case Study. *J. For. Ident.* 2018, *68* (3), 299-331

[5] OSAC Friction Ridge Subcommittee’s Response to the President’s Council of Advisors on Science & Technology’s (PCAST) Request for Additional References. Submitted December 14, 2016. <https://www.nist.gov/topics/forensic-science/friction-ridge-subcommittee>.