



Introduction aux taux d'erreur

Le but de ce document est de fournir une connaissance introductive des statistiques aux membres de l'IFIS. Il abordera des sujets importants tels que la probabilité de base, les taux d'erreur et les intervalles de confiance. Il expliquera également des problèmes actuels tels que le taux d'erreur 1 sur 18, comme la déclaré PCAST. L'objectif général est de fournir aux membres une compréhension générale de la méthodologie statistique actuellement pertinente sur le terrain. Ce n'est pas un substitut à une formation formelle en statistique. Pour faciliter la compréhension, certains sujets ont été dilués, afin que le concept puisse être compris sans entrer dans de mathématiques rigoureuses.

Si vous n'êtes pas intéressé par la lecture de l'intégralité du rapport, consultez le résumé à la fin. Il fournit quelques points brefs sur les choses que vous devez savoir.

Contents

1 Bases de Probabilité.....	4
1.1 Tableaux de Fréquences.....	5
1.2 Probabilité Conjointe	6
1.3 Probabilité Conditionnelle.....	8
2 Mesures de Performance et Taux D'erreur	9
3 Intervalles de Confiance.....	15
4 Erreur dans le rapport CCPST lié à l'étude Miami-Dade.....	17
5 Sommaire.....	19
6 Références	20

1 Bases de Probabilité

La probabilité est la représentation de la probabilité qu'un certain événement se produise. La probabilité d'un événement est représentée mathématiquement par : $P(E = \text{'un événement'})$

Par exemple, si on laisse E représenter 'la température demain', alors $P(E = \text{'pluie'})$ est 'la probabilité que la température demain soit de la pluie'. Ne vous attardez pas trop sur la lettre 'E', elle sert simplement à représenter un événement, vous pouvez utiliser n'importe quelle lettre. Un autre exemple pourrait être, si le X représente 'le résultat de la loterie', alors $P(X = \text{'gagnant'})$ est 'la probabilité que vous gagnerez la loterie'. On peut aussi dire que $CEM = \text{'le résultat d'une décision de comparaison'}$, alors $P(CEM = \text{'identification'})$ est 'la probabilité que le résultat d'une décision de comparaison soit une identification'.

Les probabilités sont représentées par des valeurs décimales ou des fractions comprises entre 0% et 100%. Les pourcentages sont équivalents aux valeurs décimales multipliés par 100, c.à-d. $1/2 = 0.5 = 50\%$, $1/4 = 0.25 = 25\%$, $0 = 0\%$, $1 = 100\%$, etc.

En résumé, on pourrait dire que T représente le résultat d'un tirage au sort, puis $P(T = \text{'têtes'}) = 1/2 = 0.5 = 50\%$.

Typiquement, la probabilité peut être interprétée comme:

$$\text{Probabilité d'un événement} = P(E = \text{'événement'}) = \frac{\text{Nombre de fois que l'événement s'est produit}}{\text{Nombre de fois où l'événement aurait pu se produire}}$$

Certains exemples sont:

$$\text{Probabilité de queues} = \frac{\text{Nombre de fois où la pièce a atteris}}{\text{Nombre de fois où la pièce a été renversé}}$$

$$\text{Probabilité qu'un male ait plus de 6'} = \frac{\text{Nombre de males plus de 6'}}{\text{Nombre de males}}$$

$$\text{La Probabilité d'un diamant étant donné que la carte est rouge} = \frac{\text{Nombre de diamants}}{\text{Nombre de cartes rouges}}$$

Nous pouvons aussi représenter les probabilités comme les probabilités. Un exemple de la probabilité que quelque chose se produise pourrait être « 1 sur 5 », ce qui signifie que nous nous attendons à ce que l'événement se produise une fois sur cinq essais. Un autre exemple pourrait être « 9 sur 10 », ce qui signifie que nous nous attendons à ce que l'événement se produise neuf fois dans dix essais. Afin de trouver la cote « 1 en X » diviser 1 par la valeur de probabilité. Voici quelques exemples,

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

De là, la côte est de en 1 (le nombre que vous venez de calculer), c.-à-d.

1 en 4, 1 en 10, 1 en 20, 1 en 24, 1 en 833

1.1 Tableaux de Fréquences

Les tableaux de fréquence sont construits en comptant le nombre de fois qu'un événement donné se produit. Supposons que nous ayons fait une expérience et compilé nos résultats. Laissez CEM représenter « le résultat d'une décision de comparaison » et GT représenter « la vérité de base de la comparaison ».

	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Dans notre pseudo-expérience, par exemple, il y a eu 10 cas où une identification a été demandée pour un essai de source différente. L'objectif est de pouvoir générer des probabilités à partir de ce tableau. Nous commençons par un exemple simple.

Exemple

Examinons les résultats possibles des décisions de comparaison. Nous laissons le CEM représenter « le résultat d'une décision de comparaison ». Remarque : Ces chiffres ne sont pas fondés sur des données réelles; ils ne servent qu'à illustrer le concept.

CEM	Identification	Non Concluant	Exclusion
	76	23	45

Quelle est la probabilité d'appeler une identification?

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

Note : Le 144 dans le dénominateur est la somme de toutes les cellules, c.-à-d. 76+23+45=144

CEM	Identification	Non Concluant	Exclusion
	76	23	45

Quelle est la probabilité d'appeler un Non concluant ?

$$P(\text{CEM} = \text{'Non Concluant'}) = \frac{23}{144} \approx 0.16 = 16\%$$

CEM	Identification	Non Concluant	Exclusion
	76	23	45

Quelle est la probabilité d'appeler une identification ou une exclusion?

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

CEM	Identification	Non Concluant	Exclusion
	76	23	45

Quelle est la probabilité de ne pas appeler un Non concluant ?

$$P(\text{CEM} \neq \text{'Non Concluant'}) = \frac{121}{144} \approx 0.84 = 84\%$$

CEM	Identification	Non Concluant	Exclusion
	76	23	45

1.2 Probabilité Conjointe

La probabilité conjointe décrit la relation entre deux événements ou plus. Nous discuterons seulement de la probabilité conjointe de deux événements, car c'est ce qui est nécessaire pour générer des taux d'erreur. Ci-dessous, nous présentons un exemple simple.

Exemple

Laissez CEM représenter « le résultat d'une décision de comparaison » et GT représente « la vérité de base de la comparaison ». Il y a un total de 1 000 essais.

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Quelle est la probabilité d'appeler une identification et la vérité de base est la même source?

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

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Quelle est la probabilité d'appeler une exclusion **et** la vérité de base est une source différente?

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

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source différent	10	120	200

Quelle est la probabilité d'appeler une identification?

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

CEM GT	Identification	Non Concluant	Exclusion
Même source	500	100	70
Source différent	10	120	200

Quelle est la probabilité que la vérité de base soit une source différente?

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

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Chaque cellule du tableau représente la fréquence conjointe de deux événements.

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

La cellule soulignée en vert dans le table ci-dessus est la fréquence articulaire qu'une identification est nommé et que le procès est de la même source. i.e. $P(\text{CEM} = \text{'Identification'} \text{ et } \text{GT} = \text{'Même Source'})$. Avons-nous la probabilité d'un seul événement, par exemple la probabilité qu'une identification ait été appelée indépendamment de la vérité sur le terrain?

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Vous faites simplement la somme de toutes les valeurs dans la colonne Identification. Cela combine tous les cas possibles où une identification peut se produire.

Qu'en est-il de la probabilité que la vérité de base soit une source différente indépendamment de la décision?

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Exactement la même idée qu'avant, maintenant nous nous concentrons sur la ligne source différente. Si nous additionnons les valeurs dans la rangée et divisons par 1000 (le nombre total d'essais dans l'expérience), nous obtenons 0,33.

1.3 Probabilité Conditionnelle

La probabilité conditionnelle est essentielle pour comprendre les taux d'erreur. Un mot qui apparaît normalement lorsqu'on examine la probabilité conditionnelle est « donné ». Par exemple, supposons que nous demandions, « Quelle est la probabilité que les Jets de Winnipeg battent le Lightning de Tampa Bay » par opposition à « quelle est la probabilité que les Jets de Winnipeg battent le Lightning de Tampa Bay étant donné que les Jets de Winnipeg ont 4 buts de plus en troisième période ». La condition fournit des renseignements supplémentaires qui changent la façon dont vous percevez le scénario. Quelques autres cas pourraient être :

« Probabilité que vous obteniez un flush » ou « Probabilité que vous obteniez un flush étant donné que vous avez reçu deux cœurs ».

« Probabilité que vous réussissiez l'examen » ou « Probabilité que vous réussissiez l'examen puisque vous avez étudié pendant 8 heures ».

« Probabilité de conclure une identification » ou « Probabilité de conclure une Identification étant donné que la vérité sur le terrain était une source différente ».

Mathématiquement, la probabilité conditionnelle s'écrit $P(A = \text{« certains événements »} \mid B = \text{« certains événements »})$. Le « \mid » peut être interprété comme « donné ». Par exemple, si nous laissons faire, CEM représente « le résultat d'une décision de comparaison » et GT représente « la vérité de base de la comparaison », puis $P(\text{CEM} = \text{« Identification »} \mid \text{GT} = \text{« Source différente »})$ est lu comme « la probabilité que vous concluiez une identification étant donné que la vérité de base était une source différente », qui est également appelé un faux positif, plus à ce sujet plus tard. Dans la section suivante, nous donnons un exemple de la façon de calculer les probabilités conditionnelles à partir d'un tableau de fréquence.

2 Mesures de Performance et Taux D'erreur

Comprendre comment les taux d'erreur sont calculés devrait être très simple si vous comprenez la probabilité conditionnelle. Vous trouverez ci-dessous une liste des mesures de rendement les plus fréquemment discutées en ce qui concerne l'identification dactyloscopique.

Taux de vrai positif (Taux d'identification vrai) = Probabilité d'appeler une identification étant donné que la vérité au sol est la même source.

Taux réel négatif (Taux réel d'exclusion) = Probabilité d'appeler une exclusion étant donné que la vérité fondamentale est une source différente.

Taux de faux positifs (Taux de fausse identification) (RPF) = Probabilité d'appeler une identification étant donné que la vérité au sol est une source différente.

Taux de faux négatifs (Taux de fausse exclusion) (FER) = Probabilité d'appeler une exclusion étant donné que la vérité fondamentale est la même source.

Nous donnons un exemple simple et calculons chacune des mesures de rendement. Pour l'exemple suivant, nous supprimons les cas non concluants pour des raisons de simplicité.

	Identification	Exclusion
Même Source	250	20
Source Différent	30	200

Exemple

$$\text{Taux Vrai Positif} = P(\text{CEM} = \text{'Identification'} \mid \text{GT} = \text{'Même Source'}) = \frac{250}{270}$$

CEM GT	Identification	Exclusion
Même Source	250	20
Source Différent	30	200

La condition change la valeur dans le dénominateur de la fraction. Dans le cas ci-dessus nous avons été donnés que la vérité de base était la même source, par conséquent nous nous concentrons seulement sur la même ligne source. Ensuite, à partir des mêmes essais sources, nous calculons le nombre de cas où l'opinion était une identification.

$$\text{Taux Négatif Réel} = P(\text{CEM} = \text{'Exclusion'} \mid \text{GT} = \text{'Source Différent'}) = \frac{200}{230}$$

CEM GT	Identification	Exclusion
Même Source	250	20
Source Différent	30	200

$$\text{Taux de Faux Positif} = P(\text{CEM} = \text{'Identification'} \mid \text{GT} = \text{'Source Différent'}) = \frac{30}{230}$$

CEM GT	Identification	Exclusion
Même Source	250	20
Source Différent	30	200

$$\text{Taux de Faux Négatifs} = P(\text{CEM} = \text{'Exclusion'} \mid \text{GT} = \text{'Même Source'}) = \frac{20}{270}$$

CEM GT	Identification	Exclusion
Même Source	250	20
Source Différent	30	200

Lorsque des non concluants sont présents, il existe deux options pour calculer certaines des mesures de rendement, qui doivent inclure les non concluants ou exclure les non concluants.

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Nous démontrons en utilisant le tableau de fréquence ci-dessus pour calculer le taux de faux positifs de deux façons différentes. Premièrement, nous démontrons le taux de faux positifs avec les Non concluants, puis nous supprimons les Non concluants.

$$\text{Taux de Faux Positif} = P(\text{CEM} = \text{'Identification'} \mid \text{GT} = \text{'Source Différent'}) = \frac{10}{330}$$

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

Taux de Faux Positif

$$= P(\text{CEM} = \text{'Identification'} \mid \text{GT} = \text{'Source Différent'} \text{ and } \text{CEM} \neq \text{Non Concluant}) = \frac{10}{210}$$

CEM GT	Identification	Non Concluant	Exclusion
Même Source	500	100	70
Source Différent	10	120	200

En ne tenant pas compte des éléments non concluants, le taux de faux positifs augmente. Habituellement, les taux de faux positifs ne comprennent pas les décisions non concluantes. L'auteur indiquera habituellement si des non concluants ont été inclus ou non dans le calcul.

Exemple

Ci-dessous nous avons le tableau des décisions de l'étude FBI/Noblis Black Box [1] annexe. Le tableau de fréquence est un peu désordonné, mais l'information est là pour obtenir les taux d'erreur.

INTRODUCTION AUX TAUX D'ERREUR

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						

Les auteurs prétendent à partir de ces données que le taux de faux positifs pour la VD (valeur pour l'identification, une étiquette utilisée pendant la phase d'analyse pour certaines organisations) est de 0,1 % et le taux de faux négatifs pour les comparaisons de VD est de 7,5 %. Il est également important de noter que les auteurs ont inclus les non concluants dans leurs calculs. Nous devrions maintenant avoir les connaissances nécessaires pour vérifier nous-mêmes ces valeurs à partir des données fournies ci-dessus.

Nous voulons d'abord calculer le taux de faux positifs pour les comparaisons VID. Nous savons donc que, pour le taux de faux positifs, nous examinons le nombre d'essais de sources différentes (non apparentées), et maintenant aussi seulement ceux qui ont été considérés comme VID. À partir des données de la boîte noire, ce nombre correspond à : 3622 + 455 + 6 = 4083 (Le FBI/Noblis a inclus les éléments non concluants dans ses calculs). Sur ces 4083 cas, 6 étaient une identification, donc,

Taux Faux Positif

$$\begin{aligned}
 &= P(\text{CEM} = \text{'Identification'} \mid \text{GT} = \text{'Source Différent'} \text{ and } \text{Val} = \text{'VID'}) \\
 &= \frac{6}{4083}
 \end{aligned}$$

INTRODUCTION AUX TAUX D'ERREUR

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						

Notre valeur finale de 6/4083 est de 0,0014695 = 0,14595 % et, avec quelques arrondis pratiques, s'avère être de 0,1 %, comme indiqué dans le rapport.

Que se passe-t-il si nous supprimons les éléments non concluants comme cela a été fait dans le rapport PCAST? Le nouveau taux de faux positifs devient,

Taux Faux Positif

$$= P(\text{CEM} = \text{'Identification'} \mid \text{GT} = \text{'Source Différent'} \text{ and Val} = \text{VID and CEM} \neq \text{'Non Concluant'}) = \frac{6}{3628}$$

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						

Notre valeur finale de 6/3628 correspond à 0.0016538 = 0.16538% ce qui signifie que les cotes sont de 1 dans 604.66. PCAST a signalé que les probabilités pour l'étude FBI/Noblis Black Box étaient de 1 sur 604.

Qu'en est-il du taux de faux négatifs? Un taux de faux négatifs signifie que nous examinons des cas de même source (appariés). De plus, nous nous limitons aux cas où la décision relative à l'analyse était VID.

$$\begin{aligned} \text{Taux Faux Négatif} &= P(\text{CEM} = \text{'Exclusion'} \mid \text{GT} = \text{'Même Source'} \text{ and } \text{Val} = \text{'VID'}) \\ &= \frac{450}{5969} \end{aligned}$$

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						

Notre valeur finale de 450/5969 est de 0,0753895 = 7,53895 %, soit 7,5 % après arrondissement. Il n'y a eu aucune erreur dans le calcul du RPF ou du RPN pour l'étude FBI/Noblis dans le rapport du PCAST. Voici un autre exemple simple de calcul du taux d'erreur.

Exemple

CEM GT	Identification	Non Concluant	Exclusion
Même Source	370	50	180
Source Différent	15	335	550

Mesure de Performance	Non Concluant?	Résultat
Taux Réel Positif	Oui	370/600
Taux Réel Négatif	Oui	550/900
Taux Faux Positif	Oui	15/900
Taux Faux Négatif	Oui	180/600
Taux Réel Positif	Non	370/550
Taux Réel Négatif	Non	550/565
Taux Faux Positif	Non	15/565
False Negative Rate	No	180/550

3 Intervalles de Confiance

Le but d'un intervalle de confiance est de fournir une bande ou une plage où le résultat réel est susceptible d'être basé sur les données recueillies à partir d'une expérience. Il est important de noter que l'intervalle de confiance ne peut parler que du paramètre d'intérêt (taux faux positif) dans les conditions d'une expérience. Par exemple, si nous devons estimer le revenu moyen d'une ville, nous ne pourrions pas généraliser ce revenu moyen à l'ensemble du Canada, ou si nous devons estimer un taux de faux positifs à l'échelle internationale, nous ne pouvons pas le préciser à un pays ou à une agence en particulier.

Par exemple, supposons que nous voulions connaître le revenu moyen à Sunnyville, au Canada. Nous menons une enquête et estimons le revenu moyen à 40 000 \$. Bien entendu, cette valeur n'est qu'une estimation fondée sur les données que nous avons recueillies. Nous allons donc également fournir l'intervalle de confiance qui donnera une limite inférieure et supérieure à l'estimation. De cette façon, nous pouvons faire des déclarations comme « il y a 95 % de chance que le revenu moyen réel à Sunnyville se situe entre 35 000 \$ et 45 000 \$ » ou « il y a 90 % de chance que le revenu moyen réel à Sunnyville se situe entre 37 500 \$ et 42 500 \$ ».

Un intervalle de confiance comporte quatre composantes :

- Le paramètre de l'intérêt (revenu moyen, taux faux positif, etc.)
- Le niveau de confiance
- La variabilité de l'expérience
- La taille de l'échantillon de l'expérience

L'intervalle de confiance le plus commun est très grossièrement construit comme ceci,

$$\text{Borne Inférieure} = \text{paramètre d'intérêt} - \frac{(\text{niveau de confiance}) \times (\text{variabilité})}{\sqrt{\text{taille d'échantillon}}}$$

$$\text{Borne Supérieure} = \text{parameter of interest} + \frac{(\text{niveau de confiance}) \times (\text{variabilité})}{\sqrt{\text{taille d'échantillon}}}$$

Revenons à notre exemple du revenu moyen et montrons comment l'intervalle de confiance pourrait être calculé. Supposons que :

- Paramètre d'intérêt = 40 000 \$
- Niveau de confiance = 95 %, ce qui se traduit par une valeur de 1,96 (non expliqué ici)
- Variabilité = 25 000 \$
- Taille de l'échantillon = 500

En utilisant ces valeurs, l'intervalle de confiance qui en résulte serait :

$$\begin{aligned} \text{Borne Inférieure} &= 40000 - \frac{1.96 \times 25000}{\sqrt{500}} = \$37,809 \\ \text{Borne Supérieure} &= 40000 + \frac{1.96 \times 25000}{\sqrt{500}} = \$42,191 \end{aligned}$$

Et nous concluons qu'il y a 95 % de chances que le revenu moyen réel à Sunnyville, au Canada, se situe entre 37 809 \$ et 42 191 \$.

Mis à part : Il est très important de noter à ce stade que les détails techniques sont sacrifiés afin de faciliter la compréhension. Par exemple, le dernier énoncé « il y a 95 % de chances que le revenu moyen réel à Sunnyville, au Canada, se situe entre 37 809 \$ et 42 191 \$, est inexact d'un point de vue statistique. Une affirmation statistiquement correcte serait « si nous devons répéter l'expérience dans les mêmes conditions, un nombre infini de fois 95 % des intervalles de confiance construits à partir des expériences contiendront la vraie valeur ». L'énoncé « correct » est habituellement un peu plus difficile à comprendre pour les gens. Pour un auditoire général, l'interprétation de « il y a 95 % de chances que le revenu moyen réel à Sunnyville, au Canada, se situe entre 37 809 \$ et 42 191 \$, est acceptable ». De plus, dans le calcul de l'intervalle de confiance, l'écart-type sert de mesure de la variabilité. Ici, nous parlons simplement de « variabilité ».

4 Erreur dans le rapport CCPST lié à l'étude Miami-Dade

Supposons que nous ayons fait une étude sur le taux d'erreur avec cinq essais pour lesquels aucune erreur n'a été commise, c.-à-d. un taux d'erreur de 0 %. Nous allons maintenant calculer l'intervalle de confiance pour le taux d'erreur de 0 % en utilisant la même méthodologie que PCAST (Binomial Exact confidence interval) :

$$\text{Borne Inférieure} = 0\%$$

$$\text{Borne Supérieure} = 52\%$$

L'intervalle de confiance a produit une limite inférieure de 0 % et une limite supérieure de 52 %. Il n'y a rien de mal avec l'intervalle de confiance, mais il pourrait y avoir quelque chose qui cloche dans la façon dont quelqu'un l'interprète. PCAST [2] aurait déclaré que le taux d'erreur pourrait être aussi élevé que 52 % ou 1 sur 2 [...] semblable à ce qu'ils ont dit dans l'étude de Miami-Dade [3], « [...] que le taux d'erreur pourrait être aussi élevé que 1 erreur dans 18 cas » [2] (le « 1 sur 18 » est basé sur une erreur de calcul comme expliqué ci-dessous). Cette déclaration est extrêmement trompeuse. Un intervalle de confiance ne vous dit pas à quel point le taux d'erreur pourrait être élevé; il fournit une zone où se trouve probablement la valeur réelle. Donc, pour notre exemple, l'énoncé le plus juste serait « il y a 95 % de chances que le taux d'erreur réel se situe entre 0 % et 52 % ».

Un intervalle de confiance peut donner une idée de la fiabilité d'une estimation. Avoir une limite inférieure de 0 % et une limite supérieure de 52 % signifie que nous ne savons pas grand-chose du taux d'erreur puisque l'intervalle de confiance est si large. Supposons plutôt que nous ayons fait une expérience où nous avons 10 erreurs sur 7000 essais, ce qui entraînerait un taux d'erreur de 0,14 %. L'intervalle de confiance qui en résulterait serait:

$$\text{Borne Inférieure} = 0.07\%$$

$$\text{Borne Supérieure} = 0.26\%$$

Il s'agit d'un intervalle de confiance beaucoup plus serré, par la suite, nous pouvons avoir une plus grande assurance dans notre taux d'erreur estimé de 0,14%, dans les conditions de l'expérience.

Nous allons maintenant passer en revue étape par étape la façon dont l'erreur « 1 sur 18 » a été commise. Nous examinons d'abord le tableau de fréquence (tableau 4) qui a été utilisé dans l'étude de Miami-Dade à la page 53.

Examens ACE [Miami-Dade Table 4]			
	Même Source	Source Différent	Totals
Identification	2,457	42	2,499
Non Concluant	446	403	849
Exclusion	235	953	1,188
Totals	3,138	1,398	4,536

Les auteurs de l'étude indiquent, à partir du tableau 4 [3], qu'il y a eu 42 identifications pour différents essais à la source. Ceci est inexact, car plus loin dans le tableau 11 [3], à la page 56, on indique qu'il n'y a que trois identifications sur les essais « source non présente » (source différente).

Décisions des participants aux essais ACE [Miami-Dade Table 11]							
Sources Présents (O/N)	# d'Empreintes Latentes	# de Décisions	Identifications Correctes	Identification Erroné	Non Concluants	Bonnes Exclusions	Exclusions Erronés
Oui	56	3,177	2,457	39	446	N/A	235
Non	24	1,359	N/A	3	403	953	N/A
Totals	80	4,536	2,457	42	849	953	235

Les auteurs calculent le taux de faux positifs, basé sur le tableau 4, avec des Non concluants comme $42/1398 = 3\%$ et le taux de faux positifs sans Non concluants comme $42/995 = 4,2\%$. Cependant, les 42 faux positifs, comme on le voit au tableau 11, se sont produits sur la même source et sur des essais de sources différentes. Dans ces circonstances, des faux positifs auraient pu se produire dans n'importe quel essai de l'expérience, par la suite, le taux de faux positifs avec des éléments non concluants devrait être de $42/4536 = 0,93\%$ et sans éléments non concluants devrait être de $42/3687 = 1,14\%$ [4].

De plus, PCAST a calculé les intervalles de confiance pour les taux de faux positifs erronés de l'étude de Miami-Dade :

Intervalle de confiance unilatéral pour le taux de faux positifs non concluants : 3,87 %

Intervalle de confiance unilatéral pour le taux de faux positifs sans conclusion : 5,43 %. La limite supérieure de l'intervalle de confiance unilatéral pour le taux de faux positifs sans éléments non concluants est de 5,43 % (0,0543), ce qui se traduit par 1 erreur dans tous les essais de 18,43 (1/0,0543), ou le fameux « 1 sur 18 ». PCAST conclut ensuite : « Le taux estimé correspond à une erreur dans 24 cas, la limite supérieure indiquant que le taux pourrait être aussi élevé que 1 erreur dans 18 cas » [2]. Comme nous l'avons déjà dit, il ne s'agit pas d'une interprétation correcte d'un intervalle de confiance et il est trompeur.

En conclusion, le taux d'erreur « 1 sur 18 » n'est pas correct, car il s'agit vraiment de la limite supérieure d'un intervalle de confiance fondé sur un taux de faux positifs incorrect. Cette erreur a été corrigée dans un article de la GRC publié dans le Journal of Forensic Identification [4] et dans une lettre de réponse du sous-comité de la crête de friction de l'Organization of Scientific Area Committees [5].

5 Sommaire

- Il n'existe pas d'étude du taux d'erreur reproduisant l'ensemble du processus de comparaison des empreintes digitales.
 - Aucun taux d'erreur ne peut être appliqué au cas en question.
 - Chaque taux d'erreur est fonction des conditions de l'expérience à partir de laquelle il a été calculé.
-
- Un intervalle de confiance fournit une mesure de l'incertitude pour un paramètre qui est estimé dans les conditions d'une expérience. Par exemple, il pourrait fournir une mesure de l'incertitude pour un taux de faux positifs estimé.
 - Un intervalle de confiance ne vous dit pas qu'un taux d'erreur « ... pourrait être aussi élevé que... » une certaine valeur.
 - Un intervalle de confiance est influencé par le paramètre que vous estimez (taux de faux positifs) ainsi que par la variabilité de vos données, la taille de l'échantillon de vos données et le niveau de confiance choisi.
-
- La valeur « 1 sur 18 » est la limite supérieure d'un intervalle de confiance unilatéral pour un taux de faux positifs mal calculé.
 - Le taux d'erreur « 1 sur 18 » n'est pas valide.
-
- Les études sur les taux d'erreur devraient être utilisées pour valider ou invalider une discipline et ne pas tenter d'estimer le taux d'erreur en pratique.

6 Références

- [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>.