

Certitudes collective et incertitudes individuelles, les données massives changent-elles la donne ?

Arthur Charpentier

LES COLLOQUES
CERISY 

SCOR
FOUNDATION FOR SCIENCE

Un peu de contexte

Détour historique

Durkheim et la notion de faits sociaux

Weber et l'action sociale

Prédictibilité et rationalité (encore un détour)

Prédire le suicide

Émile Durkheim et le suicide

Prédire le risque de suicide à l'ère de l'IA

Probabilités ?

Interpréter les probabilités

Calibration

Le groupe ou l'individu ?

La sagesse des foules (vox populi)

Les marchés prédictifs

De l'aspect temporel



Contexte (très rapidement, Charpentier (2022, 2024))

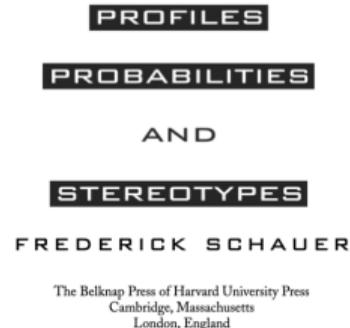
“Technology is neither good nor bad; nor is it neutral,” Kranzberg (1986).

“Ce qui est unique dans le domaine de l’assurance, c’est que même la discrimination statistique qui, par définition, est dépourvue d’intentions malveillantes, pose des problèmes moraux et juridiques importants. Pourquoi ? Parce que, d’une part, les décideurs politiques voudraient que les assureurs traitent leurs assurés de manière égale, sans discrimination fondée sur la race, le sexe, l’âge ou d’autres caractéristiques, même s’il est statistiquement logique d’opérer une discrimination. [...] D’autre part, au coeur de l’activité l’activité d’assurance réside dans la discrimination entre les assurés à risque et ceux qui ne le sont pas. Mais le risque est souvent statistiquement corrélé avec les mêmes caractéristiques que celles que les décideurs politiques voudraient interdire aux assureurs de prendre en compte,” Avraham (2017).



Contexte (très rapidement)

“Être **actuaire**, c’est être un **spécialiste de la généralisation**, et les actuaires s’engagent dans une forme de prise de décision que l’on appelle parfois actuarielle. Les actuaires guident les compagnies d’assurance dans la prise de décisions concernant de **grandes catégories** qui ont pour effet d’attribuer à l’ensemble de la catégorie certaines caractéristiques qui sont indiquées de manière probabiliste par l’appartenance à la catégorie, mais qui peuvent ne pas être possédées par un membre particulier de la catégorie,” Schauer (2006).



generalization is the stock in trade of the insurance industry. Indeed, the insurance industry has its own name for this kind of decisionmaking. To be an *actuary* is to be a specialist in generalization, and actuaries engage in a form of decisionmaking that is sometimes called *actuarial*. Actuaries guide insurance companies in making decisions about large categories (teenage males living in northern New Jersey) that have the effect of attributing to the entire category certain characteristics (carelessness in driving) that are probabilistically indicated by membership in the category, but that still may not be possessed by a particular member of the category (this *particular* teenage male living in northern New Jersey).

Occasionally the actuarial generalizations of the insurance industry become controversial. One example is the use of generalizations about the comparative safety of different neighborhoods as a basis for setting the rates for homeowners' insurance or determining the willing-

Détour historique, Préambule

“Au cours des deux premières décennies du siècle, de nombreuses compagnies ont transformé ces connaissances qualitatives en catégories écrites formelles, bien que de manière incomplète et idiosyncratique. Bien que ces systèmes de classification des risques soient restés relativement primitifs, ils ont néanmoins aidé les entreprises à prendre des décisions concernant les risques, en particulier à fixer les tarifs, qui étaient l’expression quantitative de la compréhension qualitative du danger par les assureurs. Les tarifs et les catégories de risques étaient en constante évolution, les assureurs entamant un dialogue avec les clients potentiels et le paysage,”

Eating Smoke

Fire in Urban America, 1800–1950

Mark Tebeau

CHAPTER TWO

The Business of Safety

The American Fire Insurance Industry, 1800–1850

As they cultivated new business early in the nineteenth century, insurance firms began to focus on developing a better understanding of the problem of fire and on setting guidelines for everyday business practices. During the first three decades of the century, several activities became central to the fire insurance business: surveying a risk, corresponding with field representatives and customers about hazards and rates, and compiling records of surveys and transactions in ledgers, and classifying danger. By the 1810s, companies transformed such informal procedures into formal written guidelines and organizational structures. In particular, the industry diversified its risks, and underwriters established rudimentary distinctions between different sorts of property, manufacturing activities, and construction methods. Initially such divisions resided in the minds of company secretaries—in an expanding qualitative knowledge base about fire danger that they developed from their own experience. During the first two decades of the century, many companies transformed these qualitative understandings into formal written categories, albeit incompletely and in an idiosyncratic fashion. Although such systems of classifying risk remained relatively primitive, they nonetheless helped to guide firms in making decisions about risks—especially in setting rates, which were quantitative expressions of insurers’ qualitative understandings of danger. Both rates and categories of risk were in a constant state of flux, as underwriters entered into a dialogue with prospective customers and the landscape.

The Johns Hopkins University Press
Baltimore

Détour historique, Préambule

“En 1852, Horace Binney, éminent avocat de Philadelphie, (...) oppose les méthodes de tarification incertaines des assureurs incendie aux pratiques plus saines des assureurs vie, qui reposent sur une méthode actuarielle. Il ridiculise le secteur pour ne pas avoir créé une ‘table de mortalité’ des pertes dues aux incendies et décrit les activités du secteur comme non scientifiques. (...) Dès 1850, de nombreux assureurs incendie ont pris conscience des limites de leur approche.

Même des sociétés de premier plan reconnaissent qu’elles ne peuvent prédire ni la fréquence et l’étendue des incendie ni leurs coûts financiers avec un degré de certitude acceptable. (...) Plus important encore, ils ont commencé à développer et à appliquer le raisonnement statistique à leurs activités, en particulier à leurs efforts de catégorisation,

Eating Smoke

Fire in Urban America, 1800–1950

Mark Tebeau

CHAPTER THREE

Statistics, Maps, and Morals

Making Fire Risk Objective, 1850–1875

In April 1852, prominent Philadelphia lawyer Horace Binney excoriated the fire insurance industry at an odd moment. In the keynote address of a gala event celebrating the one hundredth anniversary of the Philadelphia Contributionship for Loss from Fire, he censured fire underwriters for what he characterized as a haphazard approach to their business. Binney contrasted fire underwriters' uncertain rating methods with the sounder practices of life underwriters, which were based in actuarial method. He ridiculed the industry for not having created a “mortality table” of fire loss, and described the industry's business as unscientific. Binney urged insurers to intensify their program of observing the landscape and to form organizations that advanced the industry's common interests, arguing that such activities and associations would help fire insurers to organize their industry rationally and to understand the problem of fire. He also suggested a wholesale reconsideration of public fire defenses. Amazingly, it is unlikely that his remarks offended the audience. By 1850 many fire insurers realized the limitations of their approach. Even leading firms, such as Aetna, recognized that they could predict neither the frequency and extent of fires nor their financial costs with an acceptable degree of certainty.¹

The Johns Hopkins University Press
Baltimore

Détour historique, Préambule

“Comme Binney l’avait prévu, l’utilisation des statistiques a transformé la manière dont les assureurs incendie produisaient des connaissances sur le risque d’incendie. L’élaboration de tables actuarielles a marqué un tournant qualitatif et quantitatif important pour le secteur, d’autant plus qu’elle s’est déroulée parallèlement à l’expansion d’autres technologies de gestion de l’information. Les assureurs ont représenté le risque d’incendie urbain sous forme de ratios statistiques, ont objectivé le danger d’incendie dans les cartes des villes et ont forgé de nouvelles relations institutionnelles. Dans les vingt années qui ont suivi la critique de Binney, les assureurs incendie ont considérablement modifié leurs habitudes,” Tebeau (2003)

Eating Smoke

Fire in Urban America, 1800–1950

Mark Tebeau

CHAPTER THREE

Statistics, Maps, and Morals

Making Fire Risk Objective, 1850–1875

In April 1852, prominent Philadelphia lawyer Horace Binney excoriated the fire insurance industry at an odd moment. In the keynote address of a gala event celebrating the one hundredth anniversary of the Philadelphia Contributionship for Loss from Fire, he censured fire underwriters for what he characterized as a haphazard approach to their business. Binney contrasted fire underwriters' uncertain rating methods with the sounder practices of life underwriters, which were based in actuarial method. He ridiculed the industry for not having created a “mortality table” of fire loss, and described the industry's business as unscientific. Binney urged insurers to intensify their program of observing the landscape and to form organizations that advanced the industry's common interests, arguing that such activities and associations would help fire insurers to organize their industry rationally and to understand the problem of fire. He also suggested a wholesale reconsideration of public fire defenses. Amazingly, it is unlikely that his remarks offended the audience. By 1850 many fire insurers realized the limitations of their approach. Even leading firms, such as Aetna, recognized that they could predict neither the frequency and extent of fires nor their financial costs with an acceptable degree of certainty.¹

The Johns Hopkins University Press
Baltimore

Détour historique, Préambule

Pour terminer,

“En outre, les villes du pays connaissent de plus en plus souvent des incendies dramatiques et des dizaines, voire des centaines, de compagnies d’assurance font faillite à la suite de ces incendies. Par exemple, après le désastreux incendie de New York en 1835, vingt-trois des vingt-six compagnies locales new-yorkaises ont fait faillite. D’autres grands incendies du XIXe siècle ont donné lieu à des statistiques similaires,” Tebeau (2003)

i.e. les données permettent mieux comprendre les risques,

Ici, on regardera les modèles dans le but de segmenter.

Eating Smoke

Fire in Urban America, 1800–1950

Mark Tebeau

CHAPTER THREE

Statistics, Maps, and Morals

Making Fire Risk Objective, 1850–1875

In April 1852, prominent Philadelphia lawyer Horace Binney excoriated the fire insurance industry at an odd moment. In the keynote address of a gala event celebrating the one hundredth anniversary of the Philadelphia Contributionship for Loss from Fire, he censured fire underwriters for what he characterized as a haphazard approach to their business. Binney contrasted fire underwriters' uncertain rating methods with the sounder practices of life underwriters, which were based in actuarial method. He ridiculed the industry for not having created a “mortality table” of fire loss, and described the industry's business as unscientific. Binney urged insurers to intensify their program of observing the landscape and to form organizations that advanced the industry's common interests, arguing that such activities and associations would help fire insurers to organize their industry rationally and to understand the problem of fire. He also suggested a wholesale reconsideration of public fire defenses. Amazingly, it is unlikely that his remarks offended the audience. By 1850 many fire insurers realized the limitations of their approach. Even leading firms, such as Aetna, recognized that they could predict neither the frequency and extent of fires nor their financial costs with an acceptable degree of certainty.¹

The Johns Hopkins University Press
Baltimore

Détour historique, de l'aggrégation

“la certitude du principe sous-jacent du nouveau régime, qui était fondé sur l’espérance de la continuité de la vie qui, bien que les vies des hommes prises séparément soient incertaines, est cependant réductible à une certitude dans un ensemble de vies”,” Daston (2021), citant Society for Equitable Assurances (1764).

LORRAINE DASTON

Classical Probability in the Enlightenment

PRINCETON UNIVERSITY PRESS

generality of these persons of less importance than a provision for themselves in sickness, or old age, or at a time, when they may be disabled from labour.”²⁴⁴ The prospectus was at pains to distinguish its premiums from the flat rates other companies charged, “be the life ever so young and healthy,” and included sample premiums so that the reader might make the comparison himself. Above all, the prospectus stressed the certainty of the underlying principle of the new scheme, which was “grounded upon the expectancy of the continuance of life; which, although the lives of men separately taken, are uncertain, yet in an aggregate of lives is reducible to a certainty.”

²⁴⁴ *A Short Account of the Society for Equitable Assurances on Lives and Survivorships* (London, 2 August 1764).

Détour historique, Durkheim et la notion de faits sociaux

“Quant à la question qui a été l’origine de ce travail, c’est celle des rapports de la personnalité individuelle et de la solidarité sociale. [...] Comment peut-il être à la fois plus personnel et plus solidaire ?” Durkheim (1893), préface.

“c’est ainsi que chaque individu a son histoire, quoique les bases de l’organisation physique et morale soient les mêmes chez tous. En fait, quand on est entré quelque peu en contact avec les phénomènes sociaux, on est, au contraire, surpris de l’étonnante régularité avec laquelle ils se reproduisent dans les mêmes circonstances. Même les pratiques les plus minutieuses et, en apparence, les plus puériles, se répètent avec la plus étonnante uniformité.”

Durkheim (1895), chapitre V - Règles relatives à l’explication des faits sociaux.



Détour historique, Durkheim et la notion de faits sociaux

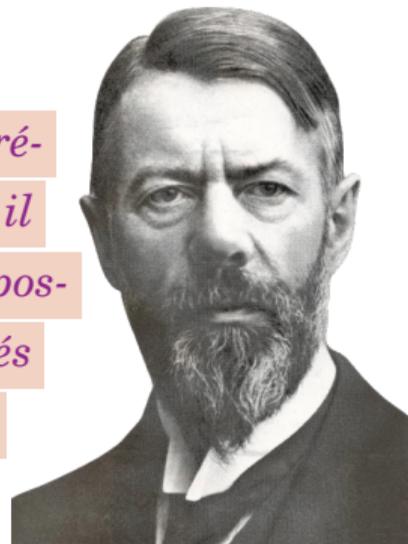
Dans *Les règles de la méthode sociologique* (Durkheim (1895)), il définit les **faits sociaux** comme des manières d'agir, de penser et de sentir qui existent en dehors des consciences individuelles et qui exercent une contrainte sur les individus.

Ces faits sociaux sont caractéristiques des groupes et des sociétés, et ils sont régis par des lois propres qui peuvent être étudiées de manière scientifique.

Dans son étude sur le **suicide** (Durkheim (1897)), il montre que malgré la nature hautement personnelle et apparemment imprévisible de l'acte suicidaire, les taux de suicide présentent des régularités frappantes selon des variables telles que la religion, l'état civil, et les conditions économiques.

Détour historique, Weber et l'action sociale

“En fin de compte, il en est de ces faits comme des autres données incompréhensibles: de même que l'individu qui agit pratiquement, l'étude compréhensive les considère comme de simples 'data' avec lesquels il faut compter. Si peu que ce soit le cas jusqu'à présent, il est possible qu'à l'avenir la recherche découvre aussi des régularités incompréhensibles dans un comportement significativement différencié [besonderes Verhalten].” Weber (1922).



Dans *Wirtschaft und Gesellschaft* (Weber (1922)), il propose une approche centrée sur le concept d'**action sociale**. Pour Weber, l'action sociale est une action à laquelle les individus attribuent un sens subjectif et qui prend en compte les actions des autres.

Détour historique, Weber et l'action sociale

“Nous appelons sociologie (au sens où nous entendons ici ce terme utilisé avec beaucoup d'équivoques) une science qui se propose de comprendre par interprétation [deutend verstehen] l'activité sociale et par là d'expliquer causalement [ursiichlich erkliiren] son déroulement et ses effets.” Weber (1922).

Weber reconnaît que bien que les actions individuelles soient motivées par des significations subjectives variées et donc puissent sembler imprévisibles, lorsqu'elles sont considérées collectivement, elles tendent également à suivre des schémas compréhensibles.

Dans son étude de l'éthique protestante et l'esprit du capitalisme (Weber (1905)), il reconnaît que des régularités émergent des actions individuelles lorsque celles-ci sont observées dans leur ensemble.

Prédictibilité et rationalité

Herbert Simon et sa théorie de la **rationalité limitée**, suggère que, malgré les limites cognitives des individus, les rendant imprévisibles, des modèles prévisibles émergent dans la prise de décision collective.

“Dans les expériences réelles avec le jeu, il s’avère que le comportement coopératif se produit assez fréquemment et qu’une stratégie ‘tit-for-tat’ (‘coopération-réciprocité-pardon’ se comporter de manière coopérative jusqu’à ce que l’autre joueur agresse ; puis agresser une fois mais revenir à la coopération si l’autre joueur le fait également) donne presque toujours des récompenses plus élevées que d’autres stratégies. Roy Radner a montré (communication personnelle) que si les joueurs s’efforcent d’obtenir un gain satisfaisant plutôt qu’un gain optimal, la solution coopérative peut être stable. La rationalité limitée semble produire de meilleurs résultats que la rationalité non limitée dans ce type de situation concurrentielle,” Simon (1969).



Détour historique, retour sur le suicide

“Dès lors, nous pouvons constituer les types sociaux du suicide, non en les classant directement d’après leurs caractères préalablement décrits, mais en classant les causes qui les produisent. Sans nous préoccuper de savoir pourquoi ils se différencient les uns des autres, nous chercherons tout de suite quelles sont les conditions sociales dont ils dépendent ; puis nous grouperons ces conditions suivant leurs ressemblances et leurs différences en un certain nombre de classes séparées, et nous pourrons être certains qu’à chacune de ces classes correspondra un type déterminé de suicide .

En un mot, notre classification , au lieu d’être morphologique, sera, d’emblée, étiologique (...) Si l’on veut savoir de quels confluent divers résulte le suicide considéré comme phénomène collectif, c’est sous sa forme collective, c’est-à-dire à travers les données statistiques , qu’il faut, l’envisager. C’est le taux social qu’il faut directement prendre pour objet d’analyse.” Durkheim (1897).



Détour historique, retour sur le suicide

Durkheim montre que le suicide est largement influencé par le degré d'intégration et de régulation sociales.

Son analyse sociologique cherche à démontrer que le suicide est un phénomène social et que les variations dans les taux de suicide peuvent être expliquées par des facteurs sociaux plutôt que simplement par des causes individuelles ou psychologiques.

Par exemple, il observe que les taux de suicide sont systématiquement plus élevés dans les sociétés protestantes que dans les sociétés catholiques, une différence qu'il attribue à des niveaux plus élevés d'intégration sociale dans ces dernières.

Cette régularité montre que les comportements individuels, lorsqu'ils sont agrégés, révèlent des tendances prévisibles à l'échelle collective.

Prédire le risque de suicide à l'ère de l'IA

125 ans après, peut-on utiliser les **données massives** et l'**apprentissage machine** (a.k.a. "**Intelligence Artificielle**"), peut-on mieux prédire les risques individuels ?

“Cette analyse a porté sur 365 études (3,428 tailles d’effet totales des facteurs de risque) réalisées au cours des 50 dernières années. La présente méta-analyse à effets aléatoires a produit plusieurs résultats inattendus : dans les analyses du rapport des cotes, du rapport des risques et de la précision diagnostique, la prédiction n’était que légèrement supérieure au hasard pour tous les résultats ,”
Franklin et al. (2017).

Voir aussi Rakesh (2017) ou Gomes de Andrade et al. (2018).

Prédire le risque de suicide à l'ère de l'IA

Approche 1, par les données (individuelles) de santé.

Étude au Québec, [Gholi Zadeh Kharrat et al. \(2024\)](#) et [Wang et al. \(2024\)](#).

Des individus ont été considérés comme des cas d'âge de 15 ans et plus et étaient décédés d'un suicide entre le 1er janvier 2002 et le 31 décembre 2019 ($n = 18,339$).

Les contrôles étaient un échantillon aléatoire de 1% de la population québécoise âgée de 15 ans et plus de chaque année, qui était en vie au 31 décembre de chaque année, de 2002 à 2019 ($n = 1,307,370$). Nous avons inclus 103 caractéristiques, y compris des facteurs individuels, programmatiques, systémiques et communautaires, mesurées jusqu'à cinq ans avant les événements suicides.

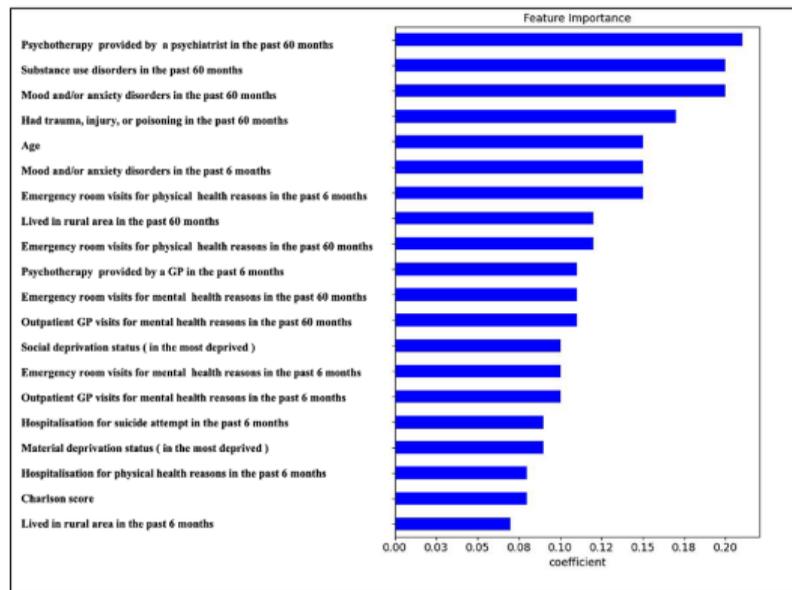
La question de savoir si les facteurs de risque de suicide diffèrent entre la population à haut risque et la population en général est une zone qui reste pour la plupart inexplorée.

Prédire le risque de suicide à l'ère de l'IA

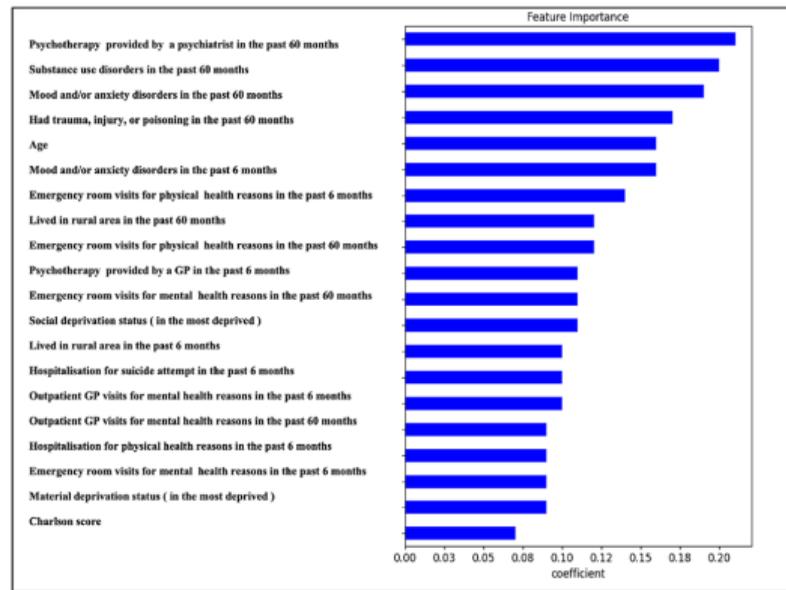
- Substance use disorder (alcohol, drug), 3, 6, 12, 24, 36, 48, 60
- Mood, anxiety, depressive, bipolar disorder, schizophrenia, autism,
- Non-Pharmacological treatments for mental health disorders (hospitalisations for mental health reasons, general practitioner visits for mental health reasons, psychotherapy visits)
- Physical diagnosis (Dementia, Neurological disease, Endocrine system disorder, Trauma, Respiratory disorder, Infectious disease, Digestive disorder, Cardiovascular disorder, Cancer)
- Non-pharmacological treatments for physical health disorders (hospitalisations for physical health reasons)
- Age, Sex, Location, Region
- Environmental Variables (Deprivation Index, Health System Environment, Quality of Care Indicators)

Prédire le risque de suicide à l'ère de l'IA

(A)

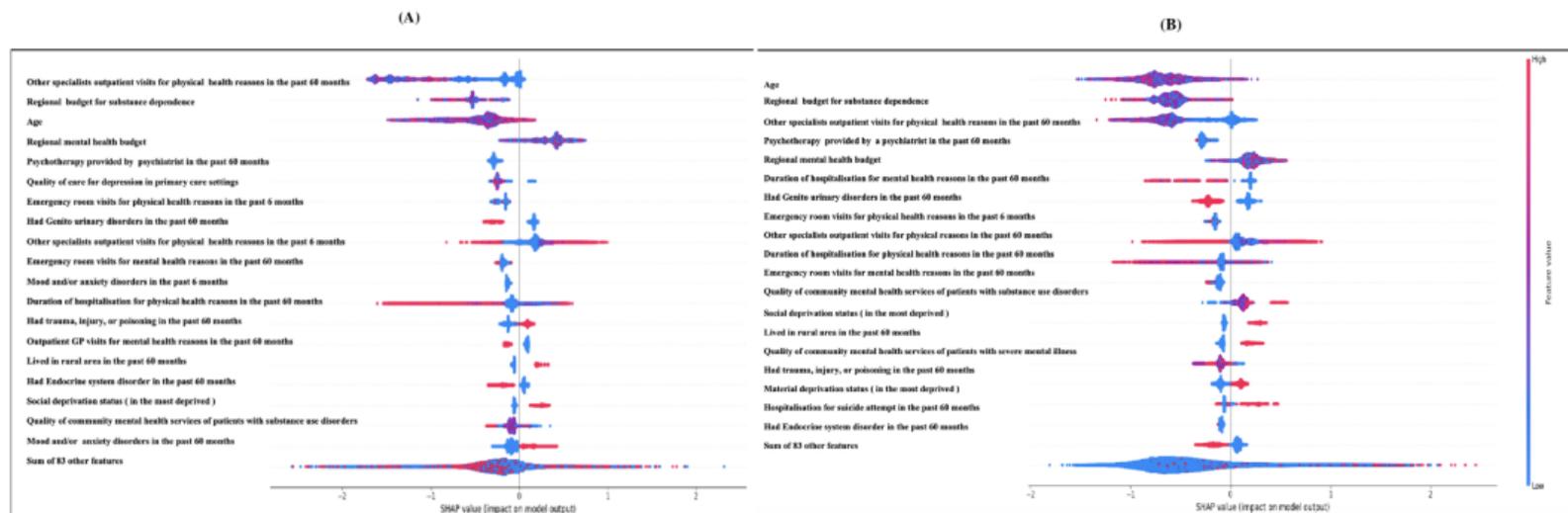


(B)



(Source Gholi Zadeh Kharrat et al. (2024), Wang et al. (2024))

Prédire le risque de suicide à l'ère de l'IA



(Source [Gholi Zadeh Kharrat et al. \(2024\)](#), [Wang et al. \(2024\)](#))

L'ensemble de l'étude a été réalisée rétrospectivement à l'aide de données administratives de santé de l'INSPQ pour la prévision du risque de suicide. Les modèles ont correctement identifié environ 31 à 38% des hommes morts par suicide / 40 à 47% des femmes décédées par suicide.

Prédire le risque de suicide à l'ère de l'IA

Approche 2, par les publications en ligne (réseaux sociaux)

1) Zirikly et al. (2019), Predicting the degree of suicide risk in Reddit posts

Le processus de collecte de données et de jugement a été mené sur des affectations provenant de forums désignés sur le suicide, tels que la "surveillance du suicide" sur Reddit

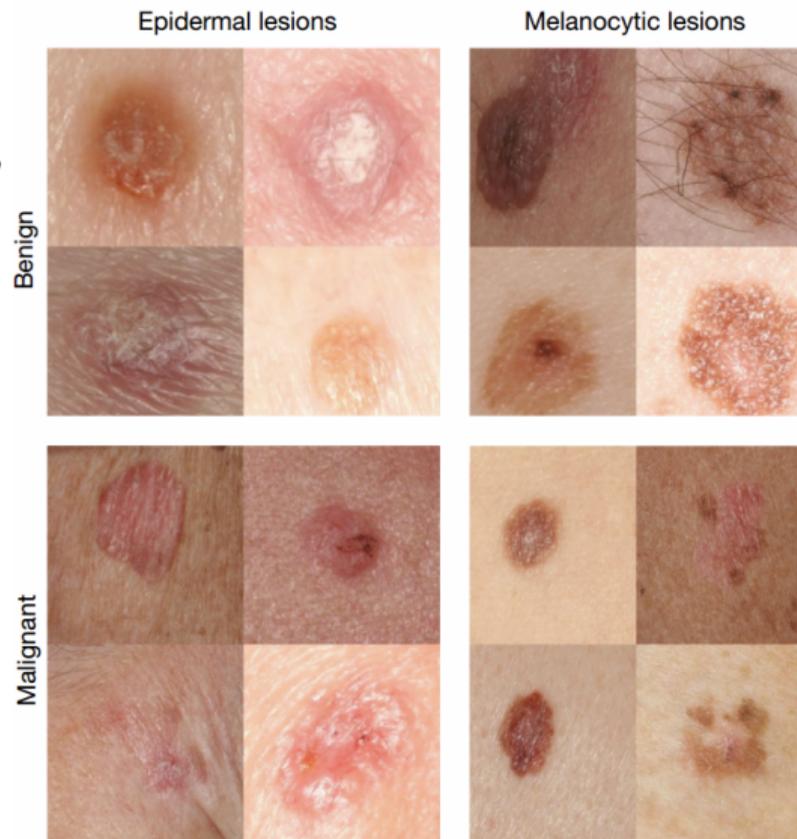
2) Ophir et al. (2020), , Deep neural networks detect suicide risk from textual facebook posts

83,292 messages rédigés par 1,002 utilisateurs authentifiés de Facebook, accompagnés d'informations psychosociales valables sur les utilisateurs.

Après avoir lu et signé le formulaire de consentement, les participants ont rempli huit mesures de psycho-diagnostic et nous ont donné une autorisation unique de télécharger leurs messages Facebook jusqu'à 12 mois avant la date de la recherche.

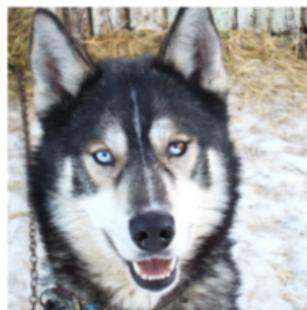
Prévision d'évènements binaires et de scores

À partir d'une image x , prédire y (noté \hat{y}),
où $y \in \{\text{Husky, Loup}\}$
où $y \in \{\text{Maligne, Benigne}\}$

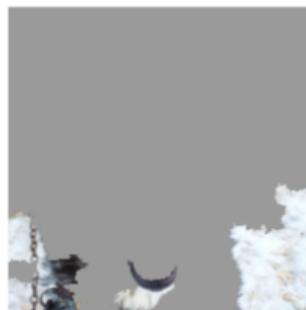


Prévision d'évènements binaires et de scores

À partir d'une image x , prédire y (noté \hat{y}),
où $y \in \{\text{Husky, Loup}\}$



(a) Husky classified as wolf



(b) Explanation

Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.

“Sur une collection de 60 images supplémentaires, le classificateur prédit ‘Loup’ s’il y a de la neige (ou un fond clair en bas), et ‘Husky’ sinon, indépendamment de la couleur de l’animal, de sa position, de sa pose, etc.”, Ribeiro et al. (2016)

	Before	After
Trusted the bad model	10 out of 27	3 out of 27
Snow as a potential feature	12 out of 27	25 out of 27

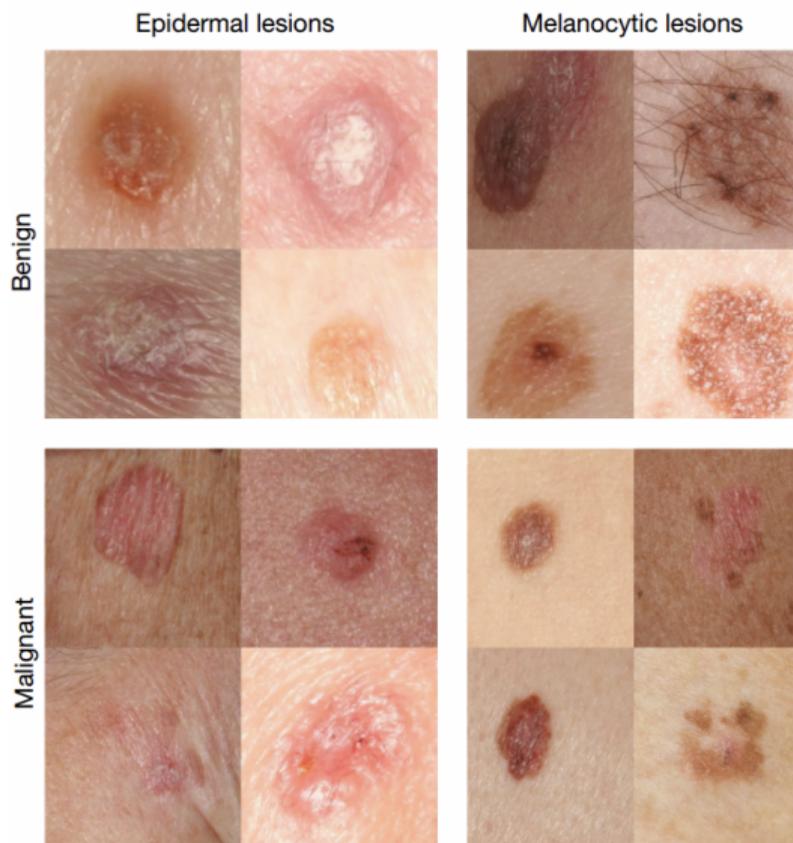
Table 2: “Husky vs Wolf” experiment results.

Prévision d'évènements binaires et de scores

À partir d'une image x , prédire y (noté \hat{y}),
où $y \in \{\text{Maligne, Benigne}\}$

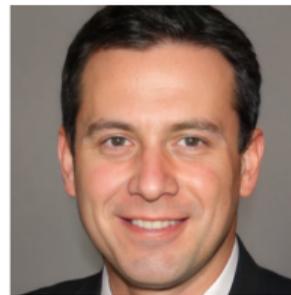
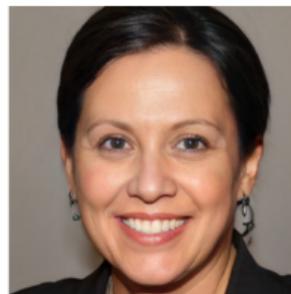
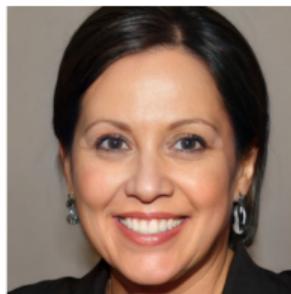
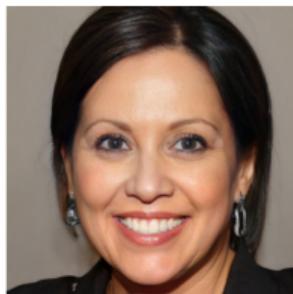
Selon [Esteva et al. \(2017\)](#) et [Winkler et al. \(2019\)](#) on peut utiliser des réseaux de neurones profonds pour détecter des cancers de la peau.

“Ainsi, dans l'ensemble des images de biopsie, si une image contenait une règle, l'algorithme était plus susceptible de qualifier une tumeur de maligne, car la présence d'une règle était corrélée à une plus grande probabilité qu'une lésion soit cancéreuse.”, [Daily Beast](#)

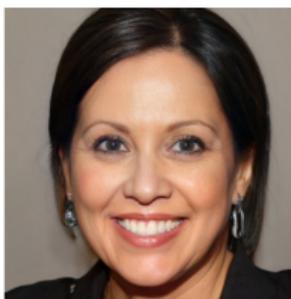


Prévision d'évènements binaires et de scores

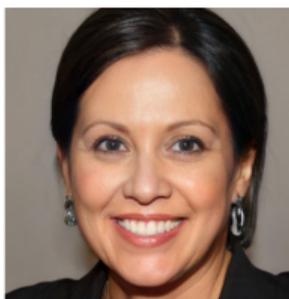
À partir d'une image x , prédire \hat{y} , où $y \in \{\text{Femme, Homme}\}$



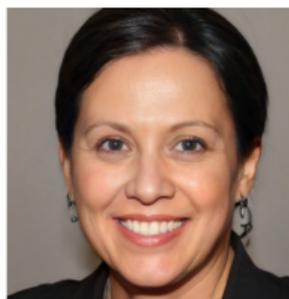
Prévision d'évènements binaires et de scores



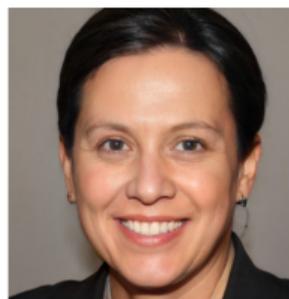
female (0.984)
male (0.016)



female (0.983)
male (0.017)



female (0.982)
male (0.018)



female (0.960)
male (0.040)



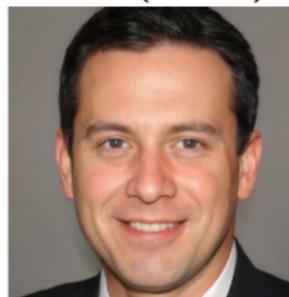
female (0.009)
male (0.991)



female (0.013)
male (0.987)



female (0.014)
male (0.986)



female (0.015)
male (0.985)

Prévision d'évènements binaires et de scores

Dans de nombreux modèles actuariels, on va chercher à prédire une variable $y \in \{0, 1\}$, à l'aide de variables prédictives, $\mathbf{x} = (x_1, \dots, x_k)$,

- tentative de suicide (ou pas) dans l'année
- décès (ou pas) dans l'année, d'au moins une personne dans un couple
- survenance d'un incendie (ou pas) dans l'année
- survenance d'un accident responsable (ou pas) dans l'année
- déclaration frauduleuse d'un sinistre (ou pas)

(dans ce cas attention du biais de sélection lors de la modélisation)

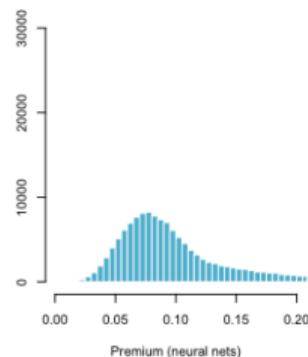
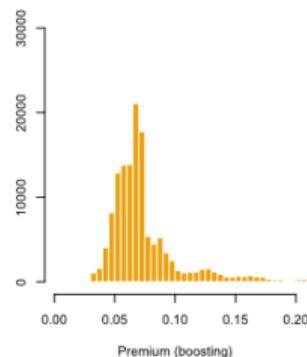
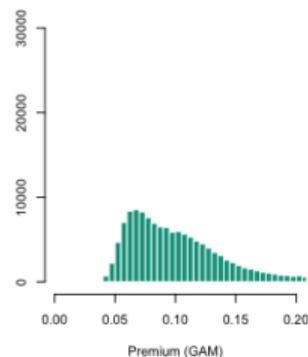
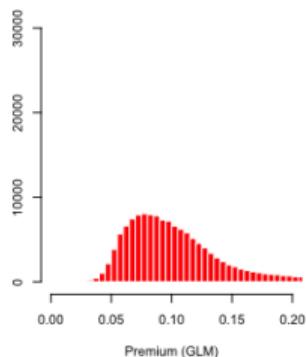
Dans la littérature en apprentissage machine, on cherche à prédire la classe, $\hat{y} \in \{0, 1\}$

En tarification, **on ne cherche pas à prédire qui aura un accident**, mais la probabilité qu'une personne ayant les caractéristique \mathbf{x} ait un accident...

Prévision d'évènements binaires et de scores

Considérons la **fréquence annuelle de sinistres** en assurance automobile (`freMTPL2freq` du package `CASDataset`, Denuit et al. (2021)).

	$\hat{\varsigma}^{\text{glm}}$	$\hat{\varsigma}^{\text{gam}}$	$\hat{\varsigma}^{\text{bst}}$	$\hat{\varsigma}^{\text{nnet}}$
moyenne des $\hat{\varsigma}(\mathbf{x}_i)$'s	0.1087	0.1092	0.0820	0.1230
10% quantile	0.0605	0.0598	0.0498	0.0529
90% quantile	0.1682	0.1713	0.1244	0.2051



Interpréter les probabilités

“Lorsqu’on parle de ‘probabilité de décès’, le sens exact de cette expression ne peut être défini que de la manière suivante. Nous ne devons pas penser à un individu, mais à

une certaine classe dans son ensemble, par exemple ‘tous les hommes assurés âgés de quarante et un ans vivant dans un pays donné et n’exerçant pas certaines professions dangereuses’. Une probabilité de décès est attachée à la classe d’hommes ou à une autre classe qui peut être définie de manière similaire. Nous ne pouvons rien dire sur la probabilité de décès d’un individu, même si nous connaissons en détail son état de vie et de santé. L’expression ‘probabilité de

décès’, lorsqu’elle se réfère à une seule personne, n’a aucune signification ,” von Mises (1928, 1939).



Interpréter les probabilités

“La probabilité est utilisée comme un substitut de la vérité tant que la vérité est inconnue [...] La première interprétation de la probabilité d'un événement unique est le degré d'attente avec lequel un événement est anticipé [...]”



“si la probabilité a quelque chose à voir avec la fiabilité des prédictions, l'énoncé de la probabilité doit être vérifiable en termes d'occurrence de l'événement prédit,”

Reichenbach (1971).

THE THEORY OF PROBABILITY

An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability

By HANS REICHENBACH

PROFESSOR OF PHILOSOPHY IN THE UNIVERSITY OF CALIFORNIA AT LOS ANGELES

UNIVERSITY OF CALIFORNIA PRESS

BERKELEY AND LOS ANGELES · 1949

§ 71. Attempts at a Single-Case Interpretation of Probability

After the discussion of the frequency meaning of probability, the investigation must turn to linguistic forms in which the concept of probability refers to an individual event. It is on this ground that the frequency interpretation has been questioned. Some logicians have argued that such usage is based on a different concept of probability, which is not reducible to frequencies. Is the existence of two disparate concepts of probability an inescapable consequence of the usage of language?

The first interpretation of the probability of single events is the *degree of expectation* with which an event is anticipated. The feeling of expectation certainly represents a psychological factor the existence of which is indisputable; it even shows degrees of intensity corresponding to the degrees of probability; it, however, arises from the fact that the degree of expectation varies from person to person and depends on more factors than the degree of the probability of the event to which the expectation refers. Apart from the probability of an event, emotional associations will influence the feeling of expectation. If it is a desirable event, as, for instance, the passing of an examination, optimistic persons will anticipate it with too-certain expectations, whereas pessimistic persons will think of it in terms of too-uncertain expectations.

“Si l’on nous demande de trouver la probabilité d’un événement futur particulier, nous devons d’abord intégrer ce cas dans une classe de référence appropriée. Un événement ou une chose particulière peut être incorporé dans de nombreuses classes de référence, d’où résulteront des probabilités différentes,” Reichenbach (1971).

THE THEORY OF PROBABILITY

An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability

By HANS REICHENBACH

PROFESSOR OF PHILOSOPHY IN THE UNIVERSITY OF CALIFORNIA AT LOS ANGELES

UNIVERSITY OF CALIFORNIA PRESS

BERKELEY AND LOS ANGELES · 1949

§ 72. The Frequency Interpretation of the Probability of the Single Case

The analysis of meaning has suffered from too close an attachment to psychological considerations. The meaning of a sentence has been identified with the mental images associated with the utterance of the sentence. Such conception leads to meanings varying from person to person; and it will not help to find the meaning that a man would adopt if he had a clear insight into the implications of his words. Logic is interested not in what a man means but in what he should mean, that is, in the meaning that, if assumed for his words, would make his words compatible with his actions.

If we are asked to find the probability holding for an individual future event, we must first incorporate the case in a suitable reference class. An individual thing or event may be incorporated in many reference classes, from which different probabilities will result. This ambiguity has been called the *problem of the reference class*. Assume that a case of illness can be characterized by its inclusion in the class of cases of tuberculosis. If additional information is obtained from an X-ray, the same case may be incorporated in the class of serious cases of tuberculosis. Depending on the classification, different probabilities will result for the prospective issue of the illness.

We then proceed by considering the *narrowest class for which reliable statistics can be compiled*. If we are confronted by two overlapping classes, we shall choose their common class. Thus, if a man is 21 years old and has tuberculosis, we shall regard the class of persons of 21 who have tuberculosis. Classes that are known to be irrelevant for the statistical result may be disregarded. A class C is irrelevant with respect to the reference class A and the attribute class B if the transition to the common class $A \cdot C$ does not change the probability, that is, if $P(A \cdot C, B) = P(A, B)$. For instance, the class of persons having the same initials is irrelevant for the life expectation of a person.

Interpréter les probabilités

“Il existe plusieurs écoles de pensée concernant l’interprétation des probabilités, dont aucune n’est exempte de défauts, de contradictions internes ou de paradoxes (...) Il n’existe pas de classification standard des interprétations des probabilités, et même les plus populaires peuvent subir des variations subtiles d’un texte à l’autre.” de Elía and Laprise (2005).

Les principales interprétations des probabilités sont, Hájek (2024)

- Probabilité classique
- Probabilité logique
- Probabilité subjective
- Interprétations en fréquence
- Interprétations de la propension
- Interprétations du meilleur système

Interpréter les probabilités

Il existe de nombreux paradoxes philosophiques quand on parle de probabilité (et de hasard), e.g. *je lance une pièce de monnaie, qui retombe, hors de ma vue*

$$\mathbb{P}(X = \text{face}) = \mathbb{P}(X = \text{pile}) = 1/2 ?$$

$$\mathbb{P}(X = \text{face}) = 1 \text{ ou } \mathbb{P}(X = \text{pile}) = 1 ?$$

ou dans un contexte juridique *“Ecoutez, soit il l’a fait, soit il ne l’a pas fait. S’il l’a fait, il est coupable à 100%. coupable à 100% et s’il ne l’a pas fait, il est coupable à 0%. de culpabilité comme une probabilité entre les deux n’a pas de sens et n’a pas sa place dans la loi,”* cité dans [Fenton and Neil \(2018\)](#).

“Nous souhaitons que les probabilités de classe estimées reflètent la véritable probabilité sous-jacente de l’échantillon,” [Kuhn and Johnson \(2013\)](#)

Calibration ou Interpretation en Fréquence

“Supposons que le Met Office indique que la probabilité de pluie demain dans votre région est de 80%. Il ne dit pas qu’il pleuvra sur 80% de la superficie de votre région et qu’il ne pleuvra pas sur les 20% restants. Ils ne disent pas non plus qu’il pleuvra pendant 80% du temps. Ce qu’ils disent, c’est qu’il y a 80% de chances qu’il pleuve à n’importe quel endroit de la région, par exemple dans votre jardin. [...] Une prévision de 80% de chances de pluie dans votre région devrait signifier en gros que, pendant environ 80% des jours où les conditions météorologiques sont les mêmes que celles de demain, il pleuvra là où vous êtes. [...] S’il ne pleut pas dans votre jardin demain, la prévision à 80% n’était pas erronée, car elle ne disait pas que la pluie était certaine. Mais si l’on considère une longue série de jours pour lesquels le Met Office a indiqué que la probabilité de pluie était de 80%, on peut s’attendre à ce qu’il ait plu pour environ 80% d’entre eux.” McConway (2021)

Calibration ou Interpretation en Fréquence

“Supposons qu’un prévisionniste attribue séquentiellement des probabilités aux événements. Il est bien calibré si, par exemple, parmi les événements auxquels il attribue une probabilité de 30%, la **proportion à long terme** qui se produit effectivement s’avère être de 30%,” Dawid (1982)

“Parmi toutes les fois où vous avez dit qu’il y avait 40% de chances qu’il pleuve, combien de fois la pluie s’est-elle réellement produite ? Si, **sur le long terme**, il a vraiment plu environ 40% du temps, cela signifie que vos prévisions étaient bien calibrées,” Silver (2012)

Un modèle \hat{s} est dit **balancé** (Charpentier (2024)) si $\mathbb{E}[Y] = \mathbb{E}[\hat{s}(\mathbf{X})]$.

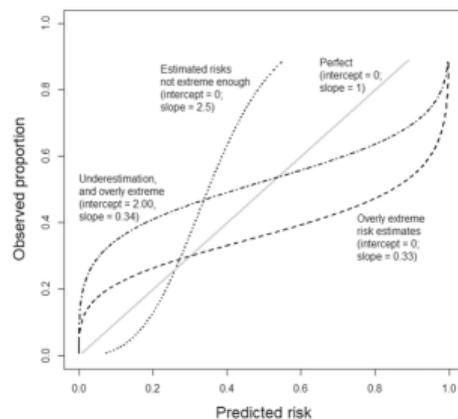
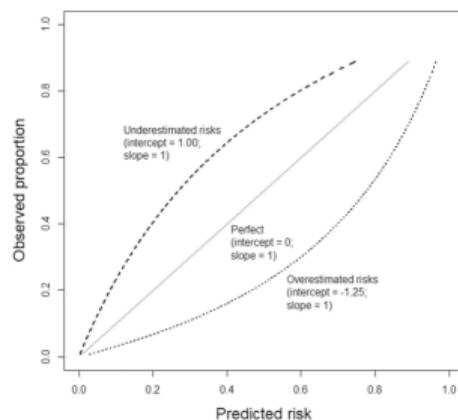
Un modèle \hat{s} est dit **bien calibré** (Schervish (1989)) ou **localement balancé** (Charpentier (2024)) si

$$\mathbb{E}[Y \mid \hat{s}(\mathbf{X}) = p] = p, \quad \forall p \in [0, 1].$$

Calibration ou Interpretation en Fréquence

Comme l'expliquent Van Calster et al. (2019) *“parmi les patients présentant un risque estimé de 20%, nous nous attendons à ce que 20 patients sur 100 soient atteints ou développent l'événement,”*

- Si 40 personnes sur 100 dans ce groupe sont atteintes de la maladie, le risque est **sous-estimé**.
- Si nous observons que dans ce groupe, 10 personnes sur 100 sont atteintes de la maladie, nous avons **surestimé** le risque.



La sagesse des foules (vox populi)

17°0 at Moyet, Basotland, on August 23. The mean yearly value of the absolute maxima was 86°9, and of the corresponding minima 41°6. The mean temperature for the year was 57°9, below the average. The stormiest month was October, and the calmest was April.

We have also received the official meteorological year-books for South Australia (1904) and Mysore (1905). Both of these works contain valuable means for previous years.

Forty Years of Southern New Mexico Climate. Bulletin No. 50 of the New Mexico College of Agriculture contains the meteorological data recorded at the experimental station from 1862 to 1902 inclusive, together with results of temperature and rainfall observations at other stations in the Mesilla Valley for most of the years between 1851 and 1860, published some years ago by General Greely in a "Report on the Climate of New Mexico." The station is situated in lat. 31° 15' N., long. 106° 45' W., and is 3568 feet above sea-level. The data have a general application to those portions of southern New Mexico with an altitude less than 4000 feet. The mean annual temperature for the whole period was 61°6, mean maximum (fourteen years) 70°3, mean minimum 41°4, absolute maximum 106° (which occurred several times), absolute minimum 1° (December, 1865). The mean annual rainfall was 8.8 inches; the smallest yearly amount was 3.5 inches, in 1873, the largest 17.1 inches, in 1905. Most of the rain falls during July, August, and September. The relative humidity is low, the mean annual amount being about 51 per cent. The bulletin was prepared by J. D. Tinsley, vice-director of the station.

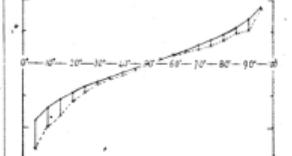
Distribution of the estimates of the dressed weight of a particular living ox, made by 287 different persons.

Degrees of the length of Array = 100	Estimates in lbs.	Cases		Errors of Observed over Normal
		Observed dressed from 1007 lbs.	Normal $\mu = 127$	
5	1074	-133	-90	+43
10	1109	-98	-70	+28
15	1126	-81	-57	+24
20	1148	-59	-40	+19
25	1162	-45	-37	+8
30	1174	-33	-30	+3
35	1184	-26	-21	+5
40	1188	-19	-14	+5
45	1197	-10	-7	+3
50	1207	0	0	0
55	1214	+7	+7	0
60	1219	+12	+14	-2
65	1225	+18	+21	-3
70	1230	+23	+29	-6
75	1239	+29	+37	-8
80	1243	+36	+46	-10
85	1244	+47	+57	-10
90	1207	+52	+70	-18
95	1193	+66	+90	-24

Fig. 16. The first and third quartiles, stand at 27 and 77 respectively. The median or middlemost value, stands at 50. The dressed weight proved to be 1207 lbs.

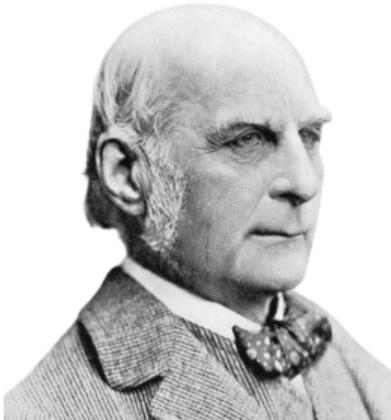
According to the democratic principle of "one vote one value," the middlemost estimate expresses the *vox populi*, every other estimate being condemned as too low or too high by a majority of the voters (for fuller explanation see "One Vote, One Value," NATURE, February 28, p. 414). Now the middlemost estimate is 1207 lbs., and the weight of the dressed ox proved to be 1198 lb.; so the *vox populi* was in this case 9 lbs., or 0.8 per cent of the whole weight, too high. The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood, and became rapidly more sparse as the distance from it increased.

Diagram, from the tabular values.



The positions here is the normal curve with $\mu = 127$. The broken line is drawn from the observations. The line connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lb. above the middlemost (37 per cent.), and another quarter deviated more than 29 lb. below it (24 per cent.), therefore the range of the two middle quarters, that is, of the middlemost half, lay within those limits. It would be an equal chance that the estimate written on any card picked at random out of the collection lay within or without those limits. In other words, the "probable error" of a single observation may be reckoned as $1/2(45+29)$, or 37 lb. (34 per cent.). Taking this for the μ of the normal curve that is best adapted for comparison with the observed values, the results are obtained which appear in above table, and graphically in the diagram.



“Ce résultat est, je pense, plus crédible que ce que l'on aurait pu attendre d'un jugement démocratique,”
Galton (1907).

(This result is, I think, more creditable to the trustworthiness of a democratic judgment than might have been expected)

Meteorological Observations in Germany.—The results of the observations made under the system of the Deutsche Seewarte, Hamburg, for 1905, at ten stations of the second order, and at fifty-six storm-warning stations, have been received. This is the twenty-eighth yearly volume published by the Seewarte, and forms part of the series of German meteorological year-books. We have frequently referred to this excellent series, and the volume in question is similar in all respects to its predecessors; it contains most valuable data relating to the North Sea and Baltic coasts. We note that the sunshine at Hamburg was only 29 per cent. of the possible amount, and that there were 103 sunless days; the rainfall was 55.9 inches, the rainy days being 172 in number.

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the Wood of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for *id.* each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly sent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiased by passion and uninfluenced by strategy and the like. The spectators few detested practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge fairly was probably much the same in either case.

After wearing thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arranged these in order of the magnitudes of the estimates, and converted the cent. quarters, and lbs. in which they were made, into lbs., under which form they will be treated.

Les marchés prédictifs (ou les paris comme outils de prévision)

En **finance (mathématique)**, loi du prix unique (Harrison and Pliska (1981), Delbaen and Schachermayer (1994) voir aussi Varian (1987)).

En **assurance**, les prix dépendent du portefeuille (**données**) de l'assureur et du **modèle utilisé**

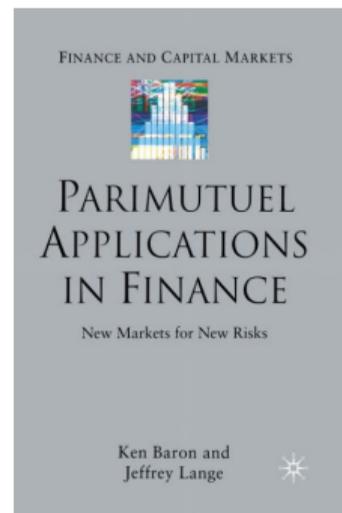
Mais il existe un entre-deux...

Très proche du concept de "*pari néerlandais*," dutch book, présenté par De Finetti (1931)

"*marché des paris*," betting market, Rhode and Strumpf (2004)

"*marché des prédictions*," prediction market, Polgreen et al. (2007)

"*parimutuel*," Baron and Lange (2006)



De l'aspect temporel dans les modèles prédictifs

Que peut-on prédire un an en avance ?

“Si nous adoptons une attitude rationnelle à l'égard du problème de la valeur, par exemple en nous plaçant dans la situation de quelqu'un qui doit déterminer la répartition de ses ressources pour une année ou cinq ans à l'avance, nous obtiendrons une vision différente de la valeur de l'argent,” Knight (1921).

L'incertitude augmente à mesure que l'horizon temporel s'allonge, ce qui rend les prévisions à long terme intrinsèquement plus difficiles.

RISK, UNCERTAINTY AND PROFIT

BY

FRANK H. KNIGHT, PH.D.

ASSOCIATE PROFESSOR OF ECONOMICS IN THE STATE UNIVERSITY
OF IOWA

BOSTON AND NEW YORK

HOUGHTON MIFFLIN COMPANY

The Riverside Press Cambridge

1921

conomic system, on the basis of that sort of calculation.¹ If we take a rational attitude toward the problem of value — as, for example, by the device, previously suggested, of placing ourselves in the position of one who had to determine the apportionment of his resources for a year or five years in advance — we shall get a different view of it. Then the earlier units are no different from the later ones, on either side of the balance; up to a certain point the balance is positive, then it suddenly becomes negative, and when the balance is struck the debits and credits are equal. There is a sort of Emersonian principle of Compensation applicable to every item; each is worth what it costs, but also costs what it is worth.

De l'aspect temporel dans les modèles prédictifs

Prévision à court terme (1 minute) : Les facteurs de risque pertinents sont immédiats et directement observables. Par exemple, si un conducteur est distrait, en excès de vitesse ou sur le point de rencontrer un danger soudain comme une autre voiture qui fait une embardée, ces facteurs sont directement liés à l'instant présent. Ils peuvent être observés, mesurés et utilisés pour prédire le risque immédiat d'accident avec un degré de confiance relativement élevé.

Prévision à long terme (1 an) : Il devient beaucoup plus complexe de prédire si une personne aura un accident de voiture. Les facteurs de risque sont diffus, imprévisibles et évoluent avec le temps. Par exemple, si les habitudes générales d'un conducteur, comme le fait de conduire fréquemment ou d'avoir des antécédents d'accidents, peuvent suggérer un risque plus élevé, il s'agit d'indicateurs généraux qui ne tiennent pas compte des conditions ou des comportements spécifiques qui se produiront sur la route à un moment donné.

Miller and Schuster (1983), Michon (1985) ou Angell and von Buseck (1985)

References

- Angell, L. S. and von Buseck, C. S. (1985). An exploratory study of vehicle type in alcohol-related crashes. *Human behavior and traffic safety*, pages 285–335.
- Avraham, R. (2017). Discrimination and insurance. In Lippert-Rasmussen, K., editor, *Handbook of the Ethics of Discrimination*, pages 335–347. Routledge.
- Baron, K. and Lange, J. (2006). *Parimutuel applications in finance: new markets for new risks*. Springer.
- Charpentier, A. (2022). *Assurance, biais, discrimination et équité*. Institut Louis Bachelier.
- Charpentier, A. (2024). *Insurance: biases, discrimination and fairness*. Springer Verlag.
- Daston, L. (2021). *Classical probability in the Enlightenment*. Princeton University Press.
- Dawid, A. P. (1982). The well-calibrated bayesian. *Journal of the American Statistical Association*, 77(379):605–610.
- de Elía, R. and Laprise, R. (2005). Diversity in interpretations of probability: Implications for weather forecasting. *Monthly Weather Review*, 133(5):1129–1143.
- De Finetti, B. (1931). Sul significato soggettivo della probabilità. *Fundamenta mathematicae*, 17.
- Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, 300(1):463–520.

References

- Denuit, M., Charpentier, A., and Trufin, J. (2021). Autocalibration and tweedie-dominance for insurance pricing with machine learning. *Insurance: Mathematics and Economics*, 101:485–497.
- Durkheim, É. (1893). *De la division du travail social*. Félix Alcan.
- Durkheim, É. (1895). *Les règles de la méthode sociologique*. Félix Alcan.
- Durkheim, É. (1897). *Le Suicide : Étude de sociologie*. Félix Alcan.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639):115–118.
- Fenton, N. and Neil, M. (2018). *Risk assessment and decision analysis with Bayesian networks*. Crc Press.
- Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., Musacchio, K. M., Jaroszewski, A. C., Chang, B. P., and Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological bulletin*, 143(2):187.
- Galton, F. (1907). Vox populi. *Nature*, 75:450–451.

References

- Gholi Zadeh Kharrat, F., Gagne, C., Lesage, A., Gariépy, G., Pelletier, J.-F., Brousseau-Paradis, C., Rochette, L., Pelletier, E., Lévesque, P., Mohammed, M., et al. (2024). Explainable artificial intelligence models for predicting risk of suicide using health administrative data in quebec. *PLoS one*, 19(4):e0301117.
- Gomes de Andrade, N. N., Pawson, D., Muriello, D., Donahue, L., and Guadagno, J. (2018). Ethics and artificial intelligence: suicide prevention on facebook. *Philosophy & Technology*, 31:669–684.
- Hájek, A. (2024). Interpretations of probability. *Stanford Encyclopedia of Philosophy*.
- Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications*, 11(3):215–260.
- Knight, F. H. (1921). *Risk, uncertainty and profit*. Houghton Mifflin Company.
- Kranzberg, M. (1986). Technology and history: "kranzberg's laws". *Technology and culture*, 27(3):544–560.
- Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modeling*. Springer.
- Michon, J. A. (1985). A critical view of driver behavior models: what do we know, what should we do? In *Human behavior and traffic safety*, pages 485–524. Springer.
- Miller, T. and Schuster, D. H. (1983). Long-term predictability of driver behavior. *Accident Analysis & Prevention*, 15(1):11–22.

References

- Ophir, Y., Tikochinski, R., Asterhan, C. S., Sisso, I., and Reichart, R. (2020). Deep neural networks detect suicide risk from textual facebook posts. *Scientific reports*, 10(1):16685.
- Polgreen, P. M., Nelson, F. D., Neumann, G. R., and Weinstein, R. A. (2007). Use of prediction markets to forecast infectious disease activity. *Clinical Infectious Diseases*, 44(2):272–279.
- Rakesh, G. (2017). Suicide prediction with machine learning. *American Journal of Psychiatry Residents' Journal*.
- Reichenbach, H. (1971). *The theory of probability*. University of California Press.
- Rhode, P. W. and Strumpf, K. S. (2004). Historical presidential betting markets. *Journal of Economic Perspectives*, 18(2):127–142.
- Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). "why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144.
- Schauer, F. (2006). *Profiles, probabilities, and stereotypes*. Harvard University Press.
- Schervish, M. J. (1989). A General Method for Comparing Probability Assessors. *The Annals of Statistics*, 17(4):1856–1879.
- Silver, N. (2012). *The signal and the noise: Why so many predictions fail-but some don't*. Penguin.

References

- Simon, H. A. (1969). *The Sciences of the Artificial*. MIT Press.
- Society for Equitable Assurances (1764). *Short Account of the Society for Equitable Assurances on Lives and Survivorships*. London.
- Tebeau, M. (2003). *Eating smoke: Fire in urban America, 1800–1950*. Johns Hopkins University Press.
- Van Calster, B., McLernon, D. J., Van Smeden, M., Wynants, L., and Steyerberg, E. W. (2019). Calibration: the achilles heel of predictive analytics. *BMC medicine*, 17(1):1–7.
- Varian, H. R. (1987). The arbitrage principle in financial economics. *Journal of Economic Perspectives*, 1(2):55–72.
- von Mises, R. (1928). *Wahrscheinlichkeit Statistik und Wahrheit*. Springer.
- von Mises, R. (1939). *Probability, statistics and truth*. Macmillan.
- Wang, J., Kharrat, F. G. Z., Gariépy, G., Gagné, C., Pelletier, J.-F., Massamba, V. K., Lévesque, P., Mohammed, M., Lesage, A., et al. (2024). Predicting the population risk of suicide using routinely collected health administrative data in quebec, canada: Model-based synthetic estimation study. *JMIR public health and surveillance*, 10(1):e52773.
- Weber, M. (1905). *Die protestantische Ethik und der Geist des Kapitalismus*. Tübingen: Mohr.
- Weber, M. (1922). *Wirtschaft und Gesellschaft*. Tübingen: Mohr.

References

- Winkler, J. K., Fink, C., Toberer, F., Enk, A., Deinlein, T., Hofmann-Wellenhof, R., Thomas, L., Lallas, A., Blum, A., Stolz, W., et al. (2019). Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition. *JAMA dermatology*, 155(10):1135–1141.
- Zirikly, A., Resnik, P., Uzuner, O., and Hollingshead, K. (2019). Clpsych 2019 shared task: Predicting the degree of suicide risk in reddit posts. In *Proceedings of the sixth workshop on computational linguistics and clinical psychology*, pages 24–33.