To be or to have been lucky, that is the question

A. Lesage[1] and J-M. Victor[2]

Abstract: Is it possible to measure the dispersion of ex-ante chances (i.e. chances "before the event") among people, be it gambling, health, or social opportunities? We explore this question and provide some tools, including a statistical test, to evidence the actual dispersion of ex-ante chances in various areas with a focus on chronic diseases.

Introduction.

"That evening he was lucky": what do we mean by this? And even weirder when we say: "the luck turned". Does this mean that we could be visited by fortune? Or that some people are luckier than others on certain days? Of course, we cannot rule out the fact that some people may bias the chances of success simply by cheating. But is there any way to assess the dispersion of chances among gamblers (or just the fraction of cheaters)?

This kind of question is part of the field of probability calculus, which aims at determining the relative likelihoods of events. The probability calculus started during summer 1654 with the correspondence between Pascal and Fermat precisely on elementary problems of gambling. Symmetry arguments are at the heart of this calculus: for example, for an unbiased coin, the two results, heads or tails, are *a priori* equivalent and therefore have the same probability of occurrence $1/2$. This is why it is not anecdotal that Pascal wanted to give his treatise the "astonishing" title "Geometry of Chance". Another illustration of the power of symmetry arguments is the tour de force of Maxwell who managed to calculate the velocity distribution of particles in idealized gases. At the time when he derived what is called since the Maxwell–Boltzmann distribution, there was no possibility to measure this distribution. It was almost 60 years before Otto Stern could achieve the first experimental verification of this distribution [1], around the same time when he confirmed with Walther Gerlach the existence of the electron spin, for which he won the Nobel Prize in 1944. The agreement between theoretical and experimental distributions was surprisingly good.

In probability theory, events are usually associated to random variables that are measurable. For example, in the heads or tails game, heads may be associated with 1 and tails with 0. Then for a given number N of draws, one can count the number of times the heads are flipped. This number k is between 0 and N and the ratio k/N is the frequency of the heads. If the coin is unbiased, this frequency fluctuates around $1/2$ when the game (N draws for each game) is played many times. Importantly, the frequency is observed ex post, i.e. *after* the game is played, then the mean frequency is used as a *measure* of the probability of getting a head. This is the usual way of assessing probabilities in statistics. Remember that assessing probabilities for anticipating the outcome of future events is the very purpose of statistics.

Dispersion of chances is far from being limited to gamblers. Disease risk is another area where people may be and actually are unequal for genetic or environmental reasons. In this case, the result of a "draw" is whether or not you have a disease D. Gambling is then limited to one

[1] Sorbonne Université, CNRS, Physico-chimie des électrolytes et nano-systèmes interfaciaux, PHENIX, F-75005 Paris, France.
[2] Sorbonne Université, CNRS, Physique théorique de la matière condensée, LPTMC, F-75005 Paris, France.

"draw" per "gambler" and only the mean probability that an individual in a given population will become ill can be observed. But can we assess the dispersion of disease risks? And if so, how can we? As a last emblematic example, we mention social opportunities. Measuring inequality of opportunity is a crucial issue with considerable political stakes, though it is extremely difficult to assess. On this last point, we postpone the in-depth study of the measure of inequal opportunities to a further work.

In all these examples, be it gambling, disease, or social opportunity, the ex-ante chances are themselves random variables that cannot be deduced from frequency measurements nor be induced by symmetry arguments. They are *hidden variables*. We propose here some tools to assess the distribution of such hidden variables and we explore more specifically the relevance of those tools to and their consequences in the field of chronic diseases.

I. A simple draw is not enough.

Let us first assume that there is a sample of n people tossing a coin and that each of them has a probability p_i to win (hence 1-p_i to lose). In an unbiased game, all the p_i are identical and equal to $1/2$. Imagine that some gamblers are luckier, others less fortunate, hence some p_i are $> 1/2$, others are $< 1/2$. This means that the p_i are random variables that are drawn from a probability distribution f(p) that is different from $\delta\left(p - \frac{1}{2}\right)$. Let Φ and Σ^2 be the mean and variance of f(p). Let us assume now that each individual plays N times. The result of each draw j of the individual i is a random variable X_i^j, either 1 in case of success or 0 in case of failure. This is a Bernoulli process: for each i the random variables X_i^j are i.i.d. (independent, identically distributed, i.e. the probability of success p_i is the same for the N draws of i). Let us define $S_i = \sum_{j=1}^{N} X_i^j$. S_i is a random variable that follows a binomial distribution B(N, p_i). S_i is the number of times the individual i has won. The mean of S_i and its variance are

$$\langle S_i \rangle = N p_i \tag{1}$$

$$Var(S_i) = N p_i (1 - p_i) \tag{2}$$

Once every individual has played N times, we obtain an estimation of the distribution of the n random variables S_i as a histogram over the N + 1 values k = 0, 1, 2, … N. These random variables S_i are independent but *non identically* distributed as the p_i are different from one individual to another.

Just as the p_i are drawn from the distribution $f(p)$, the S_i are the realizations of a random variable S (which takes the $N + 1$ discrete values $k = 0, 1, 2, \ldots N$). The probability distribution function of S is given as follows:

$$\forall k = 0, 1, \ldots N \;\; Prob(S = k) = \int_0^1 dp f(p)\, C_N^k p^k (1 - p)^{N-k} \tag{3}$$

The mean of S is

$$\langle S \rangle = \sum_{k=0}^{N} k Prob(S = k) = \int_0^1 dp f(p) \sum_{k=0}^{N} k C_N^k p^k (1 - p)^{N-k} = \int_0^1 dp f(p)\, Np = N\langle p \rangle$$

hence

$$\langle S \rangle = N\Phi \tag{4}$$

and its variance is

$$\mathrm{Var}(S) = \langle S^2 \rangle - \langle S \rangle^2$$

where

$$\langle S^2 \rangle = \sum_{k=0}^{N} k^2 Prob(S = k) = \int_0^1 dp f(p) \sum_{k=0}^{N} k^2 C_N^k p^k (1-p)^{N-k}$$

$$= \int_0^1 dp f(p) \left[Np(1-p) + (Np)^2 \right]$$

hence

$$\langle S^2 \rangle = N(\langle p \rangle - \langle p^2 \rangle) + N^2 \langle p^2 \rangle$$

and

$$\mathrm{Var}(S) = \mathrm{N}(\langle p \rangle - \langle p^2 \rangle) + N^2 \langle p^2 \rangle - N^2 \langle p \rangle^2$$

Now

$$\langle p \rangle = \Phi$$

and

$$\langle p^2 \rangle = \Sigma^2 + \Phi^2$$

so that

$$Var(S) = N(\Phi(1-\Phi) - \Sigma^2) + N^2 \Sigma^2 \quad (5)$$

Note that within the limit $N \to \infty$, the probability distribution function of the reduced variable $\mathrm{x} = \frac{k}{N}$ (where $\mathrm{k} = 0,\ 1,\ 2,\ \ldots N$) converges to the distribution $\mathrm{f}(p)$.

Equation (5) shows that, if $\mathrm{N} = 1$, the variance $\mathrm{Var}(S) = \Phi(1-\Phi)$ *does not depend on the variance* Σ^2 of $\mathrm{f}(p)$. As a matter of fact, when $N = 1$, the gains are either 0 or 1 so that the histogram of gains has only two bins, one at 0, the other at 1. The mean of gains is Φ and the variance is $\Phi(1-\Phi)$. Neither the mean nor the variance depends on the variance Σ^2 of $f(p)$. Moreover, according to equation (3), the histogram of gains itself depends only on the mean of the distribution $\mathrm{f}(p)$:

$$Prob(S = 0) = \int_0^1 dp f(p)(1-p) = 1 - \Phi \quad (6)$$

$$Prob(S = 1) = \int_0^1 dp f(p) p = \Phi \quad (7)$$

The histogram of gains cannot therefore provide information on the dispersion of chances. For example, the two following distributions:

$$f_1(p) = \delta\left(p - \frac{1}{2}\right) \quad (8)$$

and

$$f_2(p) = \frac{1}{2}[\delta(p) + \delta(p-1)] \quad (9)$$

have the same mean $\Phi = \frac{1}{2}$, hence result in the same histograms (Figure 1 for $N = 1$). However, the variance of f_1 is null whereas the variance of f_2 is $1/4$. (Note that $1/4$ is the maximal variance that a probability distribution $f(p)$ can take). This means that *a simple draw is not enough* to extract the variance of $f(p)$ from the histogram of gains; multiple draws are necessary, though are they sufficient?

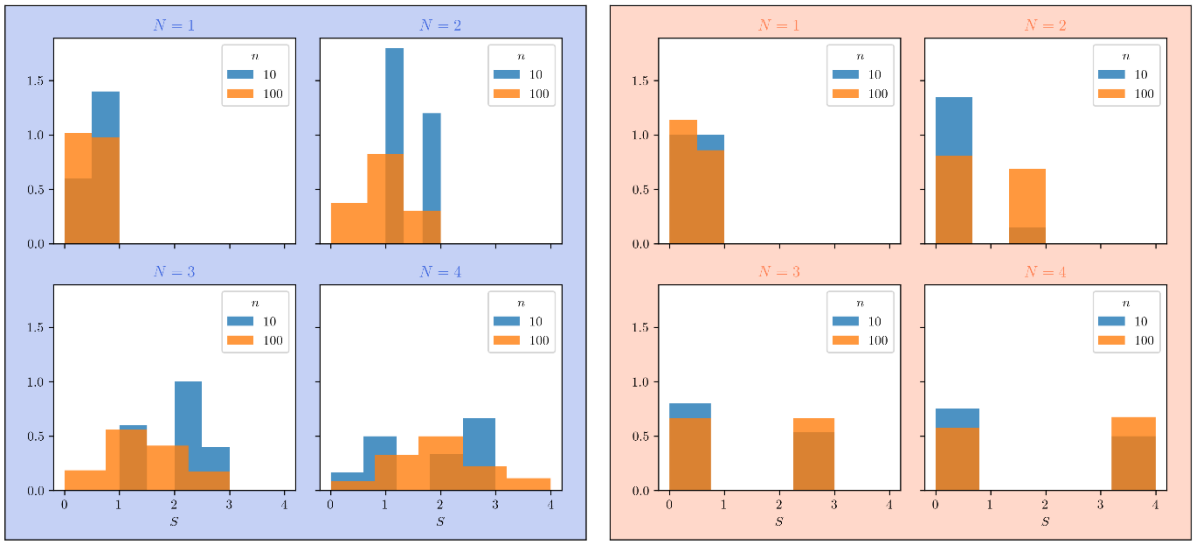

Figure 1. On the left-hand side $f_1(p) = \delta\left(p - \frac{1}{2}\right)$ *and on the right-hand side* $f_2(p) = \frac{1}{2}[\delta(p) + \delta(p-1)]$. *In each case, the histogram of success is plotted for increasing values of the number N of draws and for two numbers n of gamblers:* $n = 10$ *(in blue) and* $n = 100$ *(in orange).*

II. A statistical test of the dispersion of chances.

We first note that the histogram of gains for two draws has three bins, one at 0, the second at 1 and the third at 2, with the following values (Figure 1 for $N = 2$):

$$Prob(S = 0) = \int_0^1 dp f(p)(1-p)^2 = \langle (1-p)^2 \rangle = (1-\Phi)^2 + \Sigma^2 \qquad (10)$$

$$Prob(S = 1) = 2\int_0^1 dp f(p)p(1-p) = 2\Phi(1-\Phi) - 2\Sigma^2 \qquad (11)$$

$$Prob(S = 2) = \int_0^1 dp f(p)p^2 = \Phi^2 + \Sigma^2 \qquad (12)$$

Hence the histogram of gains now depends on (and only on) both the mean and the variance of $f(p)$. For three or more draws, we could also have access to higher order moments of $f(p)$. Nevertheless, the minimum condition for the presence of a probability dispersion is that the variance of $f(p)$is non-zero. We therefore propose to design a statistical test that will be able to discriminate between both following hypotheses:

(i) Null hypothesis H_0: everybody has the same probability Φ of gain. This means that $f(p) = \delta(p - \Phi)$ whose mean is $\langle p \rangle = \Phi$ and variance $\Sigma^2 = 0$;

(ii) Alternative hypothesis H_1: f has the same mean Φ but some people are luckier than the others so that f has a non-zero variance Σ^2.

According to H_0 the mean of N draws is Φ and the variance is $N\Phi(1-\Phi)$, whereas according to H_1 the mean of N draws is also Φ but the variance is $N(\Phi(1-\Phi)-\Sigma^2)+N^2\Sigma^2$. Hence if the variance $\mathrm{Var}(S)$ grows *linearly* with N, then all individuals have the same probability p of success. If on the contrary $\mathrm{Var}(S)$ grows *quadratically* with N then not all individuals have the same chance of success. We can therefore rephrase our hypothesis test in the following alternative based on the dependence of the variance $\mathrm{Var}(S)$ on the number N of draws:

(i) Null hypothesis H_0: the variance $\mathrm{Var}(S)$ grows *linearly* with N;
(ii) Alternative hypothesis H_1: the variance $\mathrm{Var}(S)$ grows *quadratically* with N.

Figure 2 plots the variance of the two distributions f_1 and f_2 as a function of the number N of draws for $n = 100$ gamblers. Then the Ramsey Regression Equation Specification Error Test (RESET) is a relevant tool to conclude. To be more specific, when $\mathrm{N} \geq 2$, one has to calculate the F statistic given by

$$F = \frac{\left(RSS_L - RSS_Q\right)}{\left(\frac{RSS_Q}{\mathrm{N}-1}\right)} \tag{13}$$

where RSS_L (resp. RSS_Q) is the residual sum of squares of the linear (resp. quadratic) regression. Under the null hypothesis H_0, the F statistic has an F-distribution with $(1, N-1)$ degrees of freedom. H_0 is rejected if the value of F calculated from the data is greater than the critical value of the F-distribution for some fixed false-rejection probability (usually 0.01).

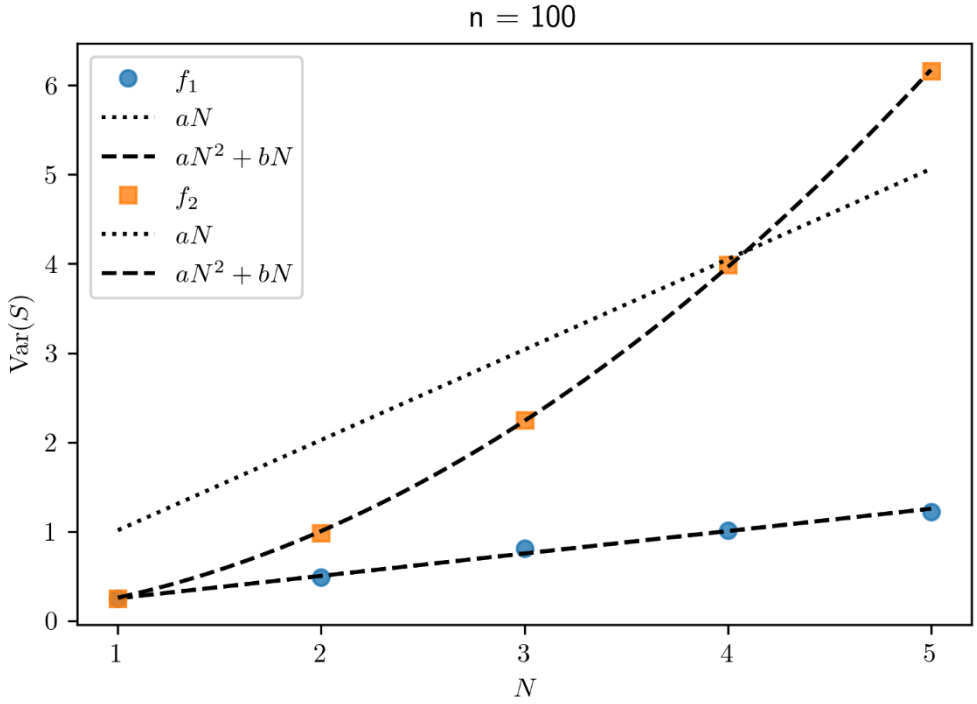

Figure 2. Linear regression fits $Var(S)$ for f_1, with $a = 0.251 \pm 0.005$ in agreement with equation (5) when $\Sigma^2 = 0$. Moreover a agrees with the expected value $\Phi(1-\Phi) = 1/4$. At odds with f_1, the linear regression does not fit $Var(S)$ for f_2 whereas

the quadratic fit is excellent, with: $a = 0.244 \pm 0.006$ *and* $b = 0.01 \pm 0.01$. *Here a agrees with the expected value* $\Sigma^2 = 1/4$ *and b with the expected value* $\Phi(1-\Phi) - \Sigma^2 = 0$.

III. Dispersion of disease risks for twins.

Inequality of risk for disease is a major public health issue. Of course, part of this inequality is known to depend on genetic and environmental factors. The mean frequency that an individual will become ill in a given population specified by genetic and environmental factors can be measured and as usual this frequency can be used as a measure of the probability to become ill. But can we assess the dispersion of disease risk, if only it exists, in this specific population? And more generally, is there any way to assess the dispersion of risk in a more objective manner, without any a priori assumption on presumed risk factors? Here comes into play a providential help from the existence of twins. Identical twins, also called monozygotic twins, have the same genome, shared the same foetal environment and generally share the same living conditions, so that they are most likely to share also the same probability to become ill, whatever the disease. Identical twins are therefore like a player betting twice. This is much related to the gambling question addressed above for $N = 2$ (two draws). Indeed, as both twins have the same probability p to have disease D, the status - ill or healthy - of each of the two twins is equivalent to the outcome – loss or gain – of each of the two draws by one and the same gambler. In this situation probability p is called a risk. Let $f(p)$ be the probability distribution function of the risk to have disease D in the population. We define the random variable S as above, i.e. $S = 0$ if both twins are healthy, $S = 1$ if only one of the two twins is ill and $S = 2$ if both twins are ill. The mean Φ and variance Σ^2 of S are given by equations (4) and (5) respectively, hence for $N = 2$

$$\langle S \rangle = 2\Phi \tag{14}$$

$$Var(S) = 2(\Phi(1-\Phi) - \Sigma^2) + 4\Sigma^2 = 2\Phi(1-\Phi) + 2\Sigma^2 \tag{15}$$

Then if $\mathrm{Var}(S)$ is significantly greater than $\langle S \rangle \left(1 - \frac{\langle S \rangle}{2}\right)$, which amounts to carry out the hypothesis test presented in the above section for $N = 2$, we can conclude that there is some dispersion of the disease risk. As we will see below the dispersion is in fact unusually large. But before that, let us calculate the concordance rate of disease D for two persons A and B. We note A_d (resp. B_d) the event " A has disease D" (resp. " B has disease D"). In genetics, concordance is the probability that a pair of individuals will both have a certain characteristic, given that one of the pair has the characteristic. Concordance rate in a population is best assessed by the probandwise rate [2], namely the conditional probability $\mathrm{P}(B_d|A_d)$ that B has disease D, knowing that A has disease D.

Of course, for two unrelated persons, this conditional probability is

$$P(B_d|A_d) = \frac{P(A_d \; and \; B_d)}{P(A_d)} = \frac{P(A_d)P(B_d)}{P(A_d)} = P(B_d) = p \tag{16}$$

and then

$$\langle P(B_d|A_d) \rangle = \langle p \rangle = \Phi \tag{17}$$

Now for two identical twins A and B, equation (16) is still true, then a naive averaging of both sides of equation (16) would necessarily lead to equation (17) as well. What is wrong? As a

matter of fact, the conditioning on A_d is not insignificant: in formula (17) it is indeed necessary to average p by considering the fact that A_d is realized, thus by using the probability density function $f_d(p)$ in the population of *affected* people and not simply $f(p)$. For identical twins, equation (17) therefore becomes

$$\langle P(B_d|A_d)\rangle = \int_0^1 p f_d(p) dp \tag{18}$$

The probability density function $f_d(p)$ can be obtained as follows:

(i) The fraction of the population that has a risk p (up to dp) to become ill is $f(p)dp$;
(ii) Among these people, a fraction p will be actually affected;
(iii) Then the fraction of affected people that has a risk p (up to dp) to become ill is $pf(p)dp$

so that

$$f_d(p)dp = \frac{pf(p)dp}{\int_0^1 pf(p)dp} \tag{19}$$

The probability density function $f_d(p)$ is therefore given by

$$f_d(p) = \frac{pf(p)}{< p >} \tag{20}$$

and the mean conditional probability $\langle P(B_d|A_d)\rangle$ is

$$\langle P(B_d|A_d)\rangle = \int_0^1 pf(p|A_d)dp = \int_0^1 p\frac{pf(p)}{\langle p\rangle}dp = \frac{1}{\langle p\rangle}\int_0^1 p^2 f(p)dp = \frac{\langle p^2\rangle}{\langle p\rangle} \tag{21}$$

We finally get the relative risk

$$RR = \frac{\langle P(B_d|A_d)\rangle}{\langle p\rangle} = \frac{\langle p^2\rangle}{\langle p\rangle^2} \tag{22}$$

which can be expressed as a function of the mean Φ and variance Σ^2 of f(p):

$$RR = \frac{\langle p^2\rangle}{\langle p\rangle^2} = \frac{\Sigma^2 + \Phi^2}{\Phi^2} = 1 + \frac{\Sigma^2}{\Phi^2} \tag{23}$$

For Crohn disease, one of the most well documented chronic disease :

$$\Phi \cong 0.0025 \tag{24}$$

and

$$RR \cong 100 \tag{25}$$

hence

$$\Sigma^2 = \Phi^2(RR - 1) \cong 0.00062 \tag{26}$$

$$\Sigma \cong 0.025 \tag{27}$$

which means that

$$\Sigma \cong 10\Phi \tag{28}$$

For Crohn disease, but this is also true for most chronic diseases, the dispersion of the risk to be affected is therefore huge indeed. The distribution $f(p)$ may then be approximated by a beta-distribution:

$$f(p;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \tag{29}$$

The parameters α and β can be expressed as a function of Φ and Σ^2:

$$\alpha = \Phi\left[\frac{\Phi(1-\Phi)}{\Sigma^2} - 1\right] \tag{30}$$

$$\beta = (1-\Phi)\left[\frac{\Phi(1-\Phi)}{\Sigma^2} - 1\right] \tag{31}$$

For Crohn disease we get:

$$\alpha \cong 0.0076 \tag{32}$$

$$\beta \cong 3 \tag{33}$$

Both probability distribution functions $f(p)$ and $f_d(p)$ for Crohn disease are plotted in Figure 3. Now the mean risk in the affected population is

$$\Phi_a = \frac{\int_0^1 p[pf(p)]dp}{\int_0^1 [pf(p)]dp} = \frac{\langle p^2 \rangle}{\langle p \rangle} = \frac{\Sigma^2 + \Phi^2}{\Phi} \tag{34}$$

As $\Sigma \cong 10\Phi$ the mean risk in the affected population is therefore

$$\Phi_a \cong 100\Phi \tag{35}$$

and the relative risk of affected people as compared to controls is

$$\frac{\Phi_a}{\Phi} \cong 100 \tag{36}$$

which means that affected people are much more predisposed to the disease than controls. Note that $\frac{\Phi_a}{\Phi}$ is equal to the relative risk RR computed above for twins.

At this stage we remark that affected people did not really have bad luck to become ill but actually had a large predisposition to become ill.

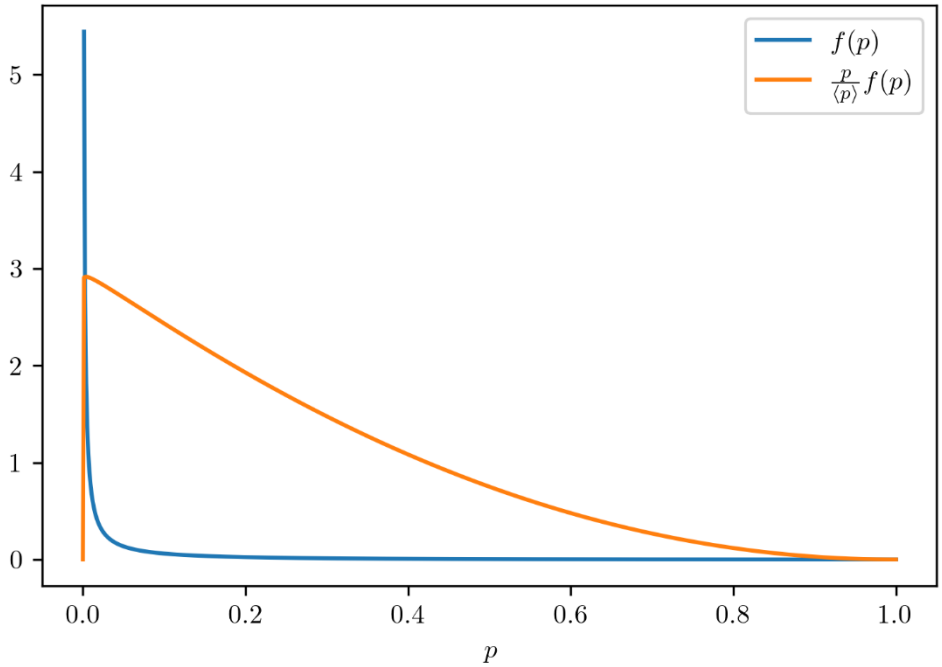

Figure 3. The risk distribution $f(p)$ among the population (in blue) is approximated by a Beta distribution of parameters $\alpha \cong 0.0076$, $\beta \cong 3$ and diverges in 0. The risk distribution $f_d = \frac{pf(p)}{\langle p \rangle}$ among affected individuals ( )⟨ ⟩ is also a Beta distribution of parameters $\alpha + 1$, β.

IV. A modular approach of chronic diseases.

Chronic diseases occur at various ages and persist throughout life. At the turn of the 2000s a new approach, called genome wide association studies (GWAS) was designed to characterize the genetic predisposition to a chronic disease. GWAS are supposed to find in particular the genes involved in a given disease, and among these genes the variants most at risk, i.e. the DNA sequences of a given gene that are more represented in the people affected by the disease. Such variants characterize the genetic predisposition to the disease. However the relative risk of most predisposition genes is hardly larger than 1, which leads to think that it is rather combinations of predisposition genes that are involved. But then a new difficulty arises, namely genetic redundancy, i.e. the fact that a given biochemical function is redundantly encoded by two or more genes. Genetic redundancy has been evidenced by knockout experiments. Knockout animals are genetically modified animals in which an existing gene has been inactivated by replacing it or disrupting it with an artificial piece of DNA. Such model animals have been created for studying the role of genes whose functions is unknown. By causing a specific gene to be inactive, and observing any differences from normal behaviour, one can infer its probable function. However knockout animals often exhibit no physiological alteration, thus evidencing that many biological functions are redundant [5]: inactivating a gene has generally no consequence on the function the gene is known to code for. This means that other genes can, and actually do implement an equivalent function. Paralagous genetic redundancy is often cited as a mechanism to account for lack of a knockout phenotype.

Suppose now that a disease D is the result of the impairment of K redundant functions, so that people become ill when and only when the whole set of K functions is impaired [3]. In other

words, as long as at least one function is performed, there is no disease. This kind of modelling has been developed recently to fit the incidence of various chronic diseases as a function of the age of onset [4]. We assume that these redundant functions are independent, i.e. the network of functions is organized in a modular way. To be more specific, each function is achieved as a module of the global physiological network (including most notably the regulatory gene network, the metabolic network and the cell signalling network) and modules operate in an independent way. The probability p to become ill is thus the product of the K probabilities p_i, $p = \prod_{i=1}^{K} p_i$, that each redundant function i is impaired. Importantly, when the function performed by module i is impaired, this module is *permissive for* disease D, but this does not mean that module i is not functioning at all. On the contrary, we suggest that module i is then switched to another functional mode which can be *protective against* another disease D'. More generally we suggest that modules may be bifunctional, i.e. that the genes that make up a module may be wired in two different ways, one being associated with a function that is protective against some disease D, the other with a function that is protective against some other disease D'.

For the sake of simplicity, we assume that all the K redundant functions have the same probability distribution function $g(p)$. Now we can relate the mean Φ and variance Σ^2 of $f(p)$ to the mean φ and variance σ^2 of $g(p)$:

$$\Phi = \varphi^K \tag{37}$$

$$\Sigma^2 = (\varphi^2 + \sigma^2)^K - \varphi^{2K} \tag{38}$$

It has been shown in [4] that the number of redundant functions is $K \cong 10$ for various "not so rare" chronic diseases. By this we mean chronic diseases that have a typical prevalence, i.e. the fraction of the population that is affected by the disease, between 0.5 per 1000 and 5 per 1000. We recall that in Europe, a disease is considered to be rare when it affects less than 1 person per 2000. Hence the mean risk for "not so rare" chronic diseases is $\Phi \cong 10^{-3}$. Equations (37) and (38) then give the mean φ and variance σ^2 of the risk distribution function $g(p)$ for one module:

$$\log(\varphi) = \frac{\log(\Phi)}{K} \cong -0.3 \tag{39}$$

hence

$$\varphi \cong \frac{1}{2} \tag{40}$$

and according to equation (28) and (38)

$$100\varphi^{2K} \cong (\varphi^2 + \sigma^2)^K - \varphi^{2K} \tag{41}$$

hence

$$\sigma^2 \cong \left(100^{\frac{1}{K}} - 1\right)\varphi^2 \tag{42}$$

As $K \cong 10$ we finally get

$$\sigma \cong 0.4 \tag{43}$$

Redundancy would thus allow an amazing variability in the combinations of possible fates of the K redundant modules involved in disease D, namely 2^K states, among which only one is permissive to disease D.

Note that σ cannot exceed 0.5, which is the maximum standard deviation that a probability distribution function $f(p)$ can take: in this extreme case, $f(p)$ is equal to $f_2(p)$ as given above in equation (9), which is the sum of two symmetric Dirac delta functions, one in 0 (protective state against disease D), the other one in 1 (permissive state for disease D). More generally, for a given module M with a mean risk φ, the maximum variance of the risk distribution $g(p)$ is $\varphi(1-\varphi)$. Whenever the variance is less than $\varphi(1-\varphi)$ there is some stochasticity in the fate of module M. The degree of stochasticity may then be defined as $\xi = 1 - x$ where x is the ratio of the actual variance over the maximal variance:

$$\xi = 1 - \frac{\sigma^2}{\varphi(1-\varphi)} \tag{44}$$

Note that the maximal dispersion $\sigma^2 = \varphi(1-\varphi)$ corresponds to the minimal degree of stochasticity $\xi = 0$. In general, the degree of stochasticity ranges between 0 and 1, from a completely deterministic behavior to a completely random one, namely:

(i) $\xi = 1$ corresponds to $g(p) = \delta(p-\varphi)$ where φ is the common probability of failure of the module M, shared by the whole population. This means that a fraction $1-\varphi$ of the population will *happen to* have the module M wired in the state that is protective against disease D while the remaining fraction φ of the population will happen to have the module M wired in the state that is permissive to disease D, hence protective against some other disease D'. This completely random behavior mimics the *bet-hedging* strategy that some plants adopt in climates that change significantly from one year to the next [6]: it is indeed advantageous for such plants to "hedge their bets", thus producing some seeds that germinate immediately in case of a rainy season and other seeds that lie dormant in case of a drought. This strategy allows these plants *and their offspring* to adapt to rapidly changing environments.

(ii) $\xi = 0$ corresponds to $g(p) = (1-\varphi)\delta(p) + \varphi\delta(p-1)$. In this case, a fraction 1-φ of the population is *determined* to be at no risk of having disease D while the remaining fraction φ is determined to be at no risk of having some other disease D'. This deterministic behavior is well adapted to a stable environment where the two states of the module M have equal fitness.

For “not so rare” chronic diseases, Equation (43) shows that the risk distribution function $g(p)$ of one module has a wide dispersion. This suggests that the degree of stochasticity ξ of one module is rather low, and even close to 0. However ξ *is not* zero, because otherwise twins would be much more alike. Why is it so? So far we have overlooked changes of the environment. However many chronic diseases have recently emerged or even exploded as a result of changes in food quality, or eating habits, or home furnishings such as a refrigerator [7]. Equation (37) shows that the mean risk φ per module must change if the prevalence Φ changes. But of course, it is more probable that only one module among the N redundant modules involved in disease D is affected by the environmental change. As the degree of stochasticity ξ of any module is low, the risk distribution $g(p)$ is close to $(1-\varphi)\delta(p) + \varphi\delta(p-1)$, hence this is practically the fraction φ of the population at risk for disease D that changes. In other words, a fraction of people with almost no risk become at very high risk, practically determined to be ill. How φ changes under environmental pressure remains to be explicited. But anyway, we speculate that evolution may have shaped the modules of the

physiological network to make them function in a rather deterministic way in stable environments but to be versatile enough so as to allow individuals and their progeny to adapt to rapidly changing environments.

Conclusion.

We have seen that hidden variables such as ex-ante chances can be actually assessed whenever multiple draws are available. Twins provide a unique means to play twice at the lottery of diseases. Of course twins are all the more relevant to assess ex-ante chances as they share the same environmental factors. In the same vein, "social twins" or more generally "social clones" would be of great help in assessing inequality of opportunities. However, controlling the environment of such social clones would be rather challenging as the issue of choice comes into play which may change people's lives with the same opportunities. Assessing the inequality of opportunities is therefore one of the most delicate, almost completely open, issues.

Since its invention in the middle of the 17th century, the probability calculus has accompanied most if not all new fields of science, especially since the beginning of the 20th century with the burst of genetics and quantum physics up to the most recent developments of quantum cognition [8], not to mention the countless applications to finance and economy.

Pascal could never complete his treatise "Geometry of Chance". This never-ending treatise is still being written, as evidenced in this special issue.

References

[1] O. Stern (1920) Zeits. f. Physik 2, 49.

[2] M. McGue (1992) When assessing twin concordance, use the probandwise not the pairwise rate. Schizophrenia Bulletin 1992, 18, 171-176.

[3] G. Debret, C. Jung, J-P. Hugot et al (2011) Genetic Susceptibility to a Complex Disease: The Key Role of Functional Redundancy. History and Philosophy of the Life Sciences Vol. 33, issue 4, 497-514.

[4] Victor J-M, Debret G, Lesne A, Pascoe L, Carrivain P, Wainrib G, et al. (2016) Network modeling of Crohn's disease incidence. PLoS ONE 11(6): e0156138.

[5] Barbaric I., Miller G., Dear T.N. (2007) Appearances can be deceiving: phenotypes of knockout mice. Briefings in Functional Genomics, Volume 6, Issue 2, 91–103.

[6] D. Cohen (1966) Optimizing reproduction in a randomly varying environment. Journal of Theoretical Biology. 12 (1): 119–129.

[7] Hugot JP, Alberti C, Berrebi D, Bingen E, Cézard JP. (2003) Crohn's disease: the cold chain hypothesis. Lancet 362(9400):2012-5.

[8] P. D. Bruza, Z. Wang, J. R. Busemeyer (2015) Quantum cognition: a new theoretical approach to psychology. Trends in Cognitive Sciences July 2015, Vol. 19, No. 7.