

Quantum Econometrics: How to Explain Its Quantitative Successes and How the Resulting Formulas Are Related to Scale Invariance, Entropy, Fuzzy, and Copulas

Hung T. Nguyen^{1,2}, Kittawit Autchariyapanitkul³, Olga Kosheleva⁴,
Vladik Kreinovich⁴, and Songsak Sriboonchitta²

¹ Department of Mathematical Sciences, New Mexico State University,
Las Cruces, New Mexico 88003, USA, hunguyen@nmsu.edu

² Faculty of Economics, Chiang Mai University, Chiang Mai 50200 Thailand,
songsajecon@gmail.com

³ Faculty of Economics, Maijo University, Chiang Mai, Thailand,
kittawit_a@mju.ac.th

⁴ University of Texas at El Paso, El Paso, Texas 79968, USA,
olgak@utep.edu, vladik@utep.edu

Abstract. Many aspects of human behavior seem to be well-described by formulas of quantum physics. In this paper, we explain this phenomenon by showing that the corresponding quantum-looking formulas can be derived from the general ideas of scale invariance, fuzziness, and copulas. We also use these ideas to derive a general family of formulas that include non-quantum and quantum probabilities as particular cases – formulas that may be more adequate for describing human behavior than purely non-quantum or purely quantum ones.

1 Formulation of the Problem

Somewhat surprisingly, quantum models describe several aspects of human behavior. The main issue with which we deal in this paper is that many aspects of human behavior seem to be well-described by formulas from quantum physics; see, e.g., [1, 3, 14].

This is understandable on the qualitative level. The success of quantum models in describing several aspects of human behavior is understandable on the qualitative level (see, e.g., [8]): similar to quantum physics, every time we gain new knowledge we inevitably change the system. For example, once we learn a new dependence between the economic variables, we can make better predictions of economic phenomena and thus, change the behavior of decision makers.

But how can we explain this success on the quantitative level? The above qualitative arguments do not explain why not only *ideas* but also *formulas* from quantum physics are helpful in describing human behavior (see, e.g., [1]).

In other words, while the *qualitative* success of quantum models is reasonable, their *quantitative* success remains a mystery.

What we do in this paper. In this paper, we show that the analysis of different types of uncertainty leads to the desired quantitative explanation of the success of quantum models in describing human behavior. To be more precise, we show that this analysis leads to a general formula that includes non-quantum and quantum probabilities as particular cases. We hope that some intermediate cases of this general formula will be even more accurate in describing human behavior than the currently used formulas based on non-quantum and quantum probabilities.

The structure of the paper is as follows. In Section 2, we briefly recall the main ideas behind quantum physics, and recall how the corresponding models help in describing human behavior. In Section 3, we show how to explain the quantum formulas, and how to derive a general formula that contains non-quantum and quantum formulas as particular cases. In Sections 4 and 5, we describe how this formula is related to entropy, fuzzy, and copulas.

2 Quantum Models and How They Describe Human Behavior: A Brief Reminder

The main difference between non-quantum and quantum probabilities. According to [4], the main difference between non-quantum and quantum probabilities can be explained on the example of the following two-slot experiment.

This experiment is about particle propagation. We have a particle generator – e.g., a light source or a radio source that generates photons, or a radioactive element that generates electrons or alpha-particles. There is an array of sensors at some distance from this generator. By detecting the particles, these sensors help us estimate the probability that the original particle goes to the location x of the sensor. To be more precise, what we estimate is the probability density $\rho(x)$ corresponding to the sensor's location x .

There is a barrier between the source of the signals and the sensors. In this barrier, there are two slots that can be open or close. If both slots are closed, no particles come through, so the sensors do not detect anything. If one or both slots are open, detectors detect the particles.

We assume that the particles do not interact with each other; this is a reasonable assumption for electromagnetic waves (photos), and even for electrons – as long as their density is not too high.

Let us assume that first we open the first slot and leave the second slot closed. Let $\rho_1(x)$ denote the resulting probability density. Now, we can close the first slot, open the second slot, and measure the new probability density. We will denote this new probability density by $\rho_2(x)$. What will happen if we open both slots? What will then be the resulting probability density $\rho(x)$?

In non-quantum physics, the answer is simple. Indeed, in the new experiment, a particle reaches the sensor if it either went through the first slot or it went

through the second slot. We set up the experiment in such a way that the particle cannot go through both slots. Thus, the probability that a particle passes via one of the two slots is equal to the sum of the probability of passing through the first slot and the probability of passing through the second slot: $\rho(x) = \rho_1(x) + \rho_2(x)$.

However, this is *not* what we observe in quantum physics. In quantum physics, to properly describe uncertainty, it is not sufficient to describe the corresponding probabilities, we also need to describe a complex-valued function $\psi(x)$ – called *wave function* – for which the probability density is equal to $\rho(x) = |\psi(x)|^2$.

In this case, if we know the wave function $\psi_1(x)$ corresponding to the first-slot-open case and the wave function $\psi_2(x)$ corresponding to the second-slot-open case, then the wave function $\psi(x)$ corresponding to the case when both slots are open is equal to $\psi(x) = \psi_1(x) + \psi_2(x)$. In this case, in general,

$$\rho(x) = |\psi(x)|^2 = |\psi_1(x) + \psi_2(x)|^2 \neq |\psi_1(x)|^2 + |\psi_2(x)|^2 = \rho_1(x) + \rho_2(x).$$

For example, in situations in which the values of all wave functions are positive real numbers, we get $\psi_1(x) = \sqrt{\rho_1(x)}$, $\psi_2(x) = \sqrt{\rho_2(x)}$, and thus, $\rho(x) = \psi^2(x)$ has the form

$$\rho(x) = (\sqrt{\rho_1(x)} + \sqrt{\rho_2(x)})^2 = \rho_1(x) + \rho_2(x) + 2\sqrt{\rho_1(x) \cdot \rho_2(x)} \neq \rho_1(x) + \rho_2(x).$$

Comment. In general, the values of $\rho_i(x)$ may be complex. In this case, by using triangle inequality, the only thing that we can conclude about $|\rho(x)| = |\rho_1(x) + \rho_2(x)|$ is that

$$||\rho_1(x)| - |\rho_2(x)|| \leq |\rho(x)| \leq |\rho_1(x)| + |\rho_2(x)|.$$

By squaring all three parts of this double inequality, we can conclude that the probability density function $\rho(x) = |\psi(x)|^2$ satisfies the inequality

$$(\sqrt{\rho_1(x)} - \sqrt{\rho_2(x)})^2 \leq \rho(x) \leq (\sqrt{\rho_1(x)} + \sqrt{\rho_2(x)})^2,$$

i.e., that

$$\rho_1(x) + \rho_2(x) - 2\sqrt{\rho_1(x) \cdot \rho_2(x)} \leq \rho(x) \leq \rho_1(x) + \rho_2(x) + 2\sqrt{\rho_1(x) \cdot \rho_2(x)}.$$

How is this related to human behavior. In the early 1980s, a group of researchers from the Republic of Georgia observed the behavior of kids in a two-door room; see, e.g., [15, 16]. In some cases, both doors were open, in other cases, only one door was open. On the other side, boxes with treats were placed, and the researchers measured how frequently kids pick up treats from the box located at spatial location x .

It turns out that the older kids, after walking through the door, mostly went to the box which was the closest to this door. For these kids, the frequency $f(x)$ of selecting the box x when both doors were open was (approximately) equal to

the sum $f_1(x) + f_2(x)$ of the frequencies $f_i(x)$ corresponding to the cases when the i -th door was open and the other door was closed. In other words, older kids exhibited non-quantum behavior.

Somewhat surprising, for younger kids (3-4 years old), the frequency $f(x)$ was different from the sum $f_1(x) + f_2(x)$: not only that, it was close to the quantum formula $f(x) = (\sqrt{f_1(x)} + \sqrt{f_2(x)})^2$.

For the adults – just like for the older kids – a large part of their decision making behavior is described by the traditional (non-quantum) probabilities. However, surprisingly, some aspects of their behavior are better described by the quantum formulas; see, e.g., [1, 3].

Remaining questions. How can we explain the usability of quantum formulas? And how can we come up with formulas that take into account some similarity to quantum phenomena without requiring that all formulas are quantum ones?

In this paper, we use several uncertainty-related approaches to come up with such a more general formula.

3 How to Explain the Quantum Formulas and to Get a General Expression Containing Non-Quantum and Quantum Formulas as Particular Cases

Formulation of the problem: reminder. For each sensor location x , we know the values of the probability densities $\rho_1(x)$ and $\rho_2(x)$ corresponding to two situations when one of the slots is open. Based on these values, we need to estimate the value $\rho(x)$.

Let $f(a, b)$ denote the algorithm that transforms the known values $a = \rho_1(x)$ and $b = \rho_2(x)$ into the estimate for $\rho(x)$. In terms of this algorithm, the desired estimate has the form $\rho(x) = f(\rho_1(x), \rho_2(x))$.

The problem is: which function $f(a, b)$ is the most appropriate?

Comment. The two-slot experiment is just an example. The algorithm $f(a, b)$ can be used not only for the two-slot experiment, but for all possible situations when we need to combine the known probabilities $a = P(A)$ and $b = P(B)$ of two events A and B into an estimate $f(a, b)$ for the probability $P(A \vee B)$ of the disjunction $A \vee B$.

First natural requirement: commutativity. The estimate for $\rho(x)$ should not depend on which door we call the first one and which one we call the second one. Thus, we must have $f(\rho_1(x), \rho_2(x)) = f(\rho_2(x), \rho_1(x))$ for all possible values $\rho_1(x)$ and $\rho_2(x)$.

In other words, we must have $f(a, b) = f(b, a)$ for all possible values a and b , i.e., the function $f(a, b)$ must be commutative.

Second natural requirement: continuity. In practice, all probability values are estimated only approximately, as frequencies $\tilde{\rho}_i(x) \approx \rho_i(x)$. The more observations we have, the more accurate these approximations, i.e., the closer the estimates to the actual probabilities. It is therefore reasonable to require

that as the approximate values $\tilde{\rho}_1(x)$ and $\tilde{\rho}_2(x)$ tend to the actual values $\rho_1(x)$ and $\rho_2(x)$, the resulting estimate $f(\tilde{\rho}_1(x), \tilde{\rho}_2(x))$ should tend to the estimate $f(\rho_1(x), \rho_2(x))$ based on the actual values. In other words, it is reasonable to require that the function $f(a, b)$ be *continuous*.

Third natural requirement: monotonicity. It is also reasonable to require that if one of the probabilities $\rho_i(x)$ increases, the resulting overall probability $\rho(x)$ should increase as well. In other words, it is reasonable to require that the function $f(a, b)$ is a (non-strictly) *increasing* function of each of its variables.

Fourth natural requirement: associativity. If we have three doors instead of two, then we can estimate the probability $\rho(x)$ corresponding to the case when all 3 doors are open in two different ways:

- We can first use the algorithm $f(a, b)$ to estimate the probability density function $\rho_{12}(x) = f(\rho_1(x), \rho_2(x))$ corresponding to the case when the first two doors are open, and then again apply the same algorithm $f(a, b)$ to combine the probability density function $\rho_{12}(x)$ with the probability density function $\rho_3(x)$, resulting in the value $f(f(\rho_1(x), \rho_2(x)), \rho_3(x))$.
- Alternatively, we can first combine the probability density functions corresponding to doors 2 and 3, resulting in $\rho_{23}(x) = f(\rho_2(x), \rho_3(x))$, and then combine the resulting probability density function with $\rho_1(x)$, resulting in

$$f(\rho_1(x), \rho_{23}(x)) = f(\rho_1(x), f(\rho_2(x), \rho_3(x))).$$

It is reasonable to require that these two estimates are equal for all possible values $\rho_i(x)$, i.e., that $f(f(a, b), c) = f(a, f(b, c))$ for all a, b , and c – in other words, that the operation $f(a, b)$ is *associative*.

Fifth natural requirement: scale-invariance. By definition, the probability density is probability divided by the length or area (or volume). In principle, we can use different units for measuring length, and thus, different units for measuring area or volume. If we replace the original measuring unit with the one which is λ times smaller, the numerical values of the probability density gets multiplied by λ .

It is reasonable to require that the estimating function $f(a, b)$ should not change if we thus re-scale all the values of the probability density, that $\rho(x) = f(\rho_1(x), \rho_2(x))$ should imply $\lambda \cdot \rho(x) = f(\lambda \cdot \rho_1(x), \lambda \cdot \rho_2(x))$. Thus, we require that $f(\lambda \cdot a, \lambda \cdot b) = \lambda \cdot f(a, b)$ for all possible values a, b , and λ .

Now, we are ready to formulate our main result.

Definition 1. We say that a function $f(a, b)$ from non-negative numbers to non-negative numbers is a scale-invariant estimation function if it is commutative, associative, continuous, (non-strictly) increasing, and $f(\lambda \cdot a, \lambda \cdot b) = \lambda \cdot f(a, b)$ for all a, b , and λ .

Proposition 1. The only scale-invariant estimation functions are $f(a, b) = \min(a, b)$, $f(a, b) = \max(a, b)$, and $f(a, b) = (a^\alpha + b^\alpha)^{1/\alpha}$ for some α .

4 Proof

1°. Depending on whether the value $f(1, 1)$ is equal to 1 or not, we have two possible cases: $f(1, 1) = 1$ and when $f(1, 1) \neq 1$. Let us consider these two cases one by one.

2°. Let us first consider the case when $f(1, 1) = 1$. In this case, the value $f(0, 1)$ can be either equal to 0 or different from 0. Let us consider both subcases.

2.1°. Let us first consider the first subcase, when $f(0, 1) = 0$.

In this case, for every $b > 0$, scale invariance with $\lambda = b$ implies that

$$f(b \cdot 0, b \cdot 1) = b \cdot 0,$$

i.e., that $f(0, b) = 0$. By taking $b \rightarrow 0$ and using continuity, we also get $0 * 0 = 0$. Thus, $f(0, b) = 0$ for all b .

By commutativity, we have $f(a, 0) = 0$ for all a . So, to fully describe the operation $f(a, b)$, it is sufficient to consider the cases when $a > 0$ and $b > 0$.

2.1.1°. Let us prove, by contradiction, that in this subcase, we have $f(1, a) \leq 1$ for all a .

Indeed, let us assume that for some a , we have $b \stackrel{\text{def}}{=} f(1, a) > 1$. Then, due to associativity and $f(1, 1) = 1$, we have $f(1, b) = f(1, f(1, a)) = f(f(1, 1), a) = f(1, a) = b$.

Due to scale-invariance with $\lambda = b$, the equality $f(1, b) = b$ implies that $f(b, b^2) = b^2$. Thus, $f(1, b^2) = f(1, f(b, b^2)) = f(f(1, b), b^2) = f(b, b^2) = b^2$.

Similarly, from $f(1, b^2) = b^2$, we conclude that for $b^4 = (b^2)^2$, we have $f(1, b^4) = b^4$, and, in general, that $f(1, b^{2^n}) = b^{2^n}$ for every n .

Scale invariance with $\lambda = b^{-2^n}$ implies that $f(b^{-2^n}, 1) = 1$. In the limit $n \rightarrow \infty$, we get $f(0, 1) = 1$, which contradicts to our assumption that $f(0, 1) = 0$. This contradiction shows that indeed, $f(1, a) \leq 1$.

2.1.2°. For $a \geq 1$, monotonicity implies $1 = f(1, 1) \leq f(1, a)$, so $f(1, a) \leq 1$ implies that $f(1, a) = 1$.

Now, for any a' and b' for which $0 < a' \leq b'$, if we denote $r \stackrel{\text{def}}{=} \frac{b'}{a'} \geq 1$, then scale-invariance with $\lambda = a'$ implies that $a' \cdot f(1, r) = f(a' \cdot 1, a' \cdot r) = f(a', b')$. Here, $f(1, r) = 1$, thus $f(a', b') = a' \cdot 1 = a'$, i.e., $f(a', b') = \min(a', b')$. Due to commutativity, the same formula also holds when $a' \geq b'$. So, in this case, $f(a, b) = \min(a, b)$ for all a and b .

2.2°. Let us now consider the second subcase of the first case, when $f(0, 1) > 0$.

2.2.1°. Let us first show that in this subcase, we have $f(0, 0) = 0$.

Indeed, scale-invariance with $\lambda = 2$ implies that from $f(0, 0) = a$, we can conclude that

$$f(2 \cdot 0, 2 \cdot 0) = f(0, 0) = 2 \cdot a.$$

Thus $a = 2 \cdot a$, hence $a = 0$. The statement is proven.

2.2.2°. Let us now prove that in this subcase, $f(0, 1) = 1$.

Indeed, in this case, for $a \stackrel{\text{def}}{=} f(0, 1)$, we have, due to $f(0, 0) = 0$ and associativity, that

$$f(0, a) = f(0, f(0, 1)) = f(f(0, 0), 1) = f(0, 1) = a.$$

Here, $a > 0$, so by applying scale invariance with $\lambda = a^{-1}$, we conclude that $f(0, 1) = 1$.

2.2.3°. Let us now prove that for every $a \leq b$, we have $f(a, b) = b$. So, due to commutativity, we have $f(a, b) = \max(a, b)$ for all a and b .

Indeed, from $f(1, 1) = 1$ and $f(0, 1) = 1$, due to scale invariance with $\lambda = b$, we conclude that $f(0, b) = b$ and $f(1, b) = b$. Due to monotonicity, $0 \leq a \leq b$ implies that $b = f(0, b) \leq f(a, b) \leq f(b, b) = b$, thus $f(a, b) = b$. The statement is proven.

3°. Let us now consider the remaining case when $f(1, 1) \neq 1$.

3.1°. Let us denote $v(k) \stackrel{\text{def}}{=} f(1, f(\dots, 1) \dots)$ (k times). Then, for every m and n , the value $v(m \cdot n) = f(1, f(\dots, 1) \dots)$ ($m \cdot n$ times) can be represented as

$$f(f(1, f(\dots, 1) \dots), \dots, f(1, f(\dots, 1) \dots)),$$

where we divide the 1s into m groups with n 1s in each. For each group, we have $f(1, f(\dots, 1) \dots) = v(n)$. Thus, $v(m \cdot n) = f(v(n), f(\dots, v(n)) \dots)$ (m times).

We know that $f(1, f(\dots, 1) \dots)$ (m times) $= v(m)$. Thus, by using scale-invariance with $\lambda = v(n)$, we conclude that $v(m \cdot n) = v(m) \cdot v(n)$, i.e., that that function $v(n)$ is multiplicative. In particular, this means that for every number p and for every positive integer n , we have $v(p^n) = (v(p))^n$.

3.2°. If $v(2) = f(1, 1) > 1$, then by monotonicity, we get $v(3) = f(1, v(2)) \geq f(1, 1) = v(2)$, and, in general, $v(n + 1) \geq v(n)$. Thus, in this case, the sequence $v(n)$ is (non-strictly) increasing.

Similarly, if $v(2) = f(1, 1) < 1$, then we get $v(3) \leq v(2)$ and, in general, $v(n + 1) \leq v(n)$, i.e., in this case, the sequence $v(n)$ is strictly decreasing.

Let us consider these two cases one by one.

3.2.1°. Let us first consider the case when the sequence $v(n)$ is increasing. In this case, for every three integers m , n , and p , if $2^m \leq p^n$, then $v(2^m) \leq v(p^n)$, i.e., $(v(2))^m \leq (v(p))^n$.

For all m , n , and p , the inequality $2^m \leq p^n$ is equivalent to $m \cdot \ln(2) \leq n \cdot \ln(p)$, i.e., to $\frac{m}{n} \leq \frac{\ln(p)}{\ln(2)}$. Similarly, the inequality $(v(2))^m \geq (v(p))^n$ is equivalent to $\frac{m}{n} \leq \frac{\ln(v(p))}{\ln(v(2))}$. Thus, the above conclusion

$$\text{if } 2^m \leq p^n, \text{ then } (v(2))^m \leq (v(p))^n$$

takes the following form:

$$\text{for every rational number } \frac{m}{n}, \text{ if } \frac{m}{n} \leq \frac{\ln(p)}{\ln(2)} \text{ then } \frac{m}{n} \leq \frac{\ln(v(p))}{\ln(v(2))}.$$

Similarly, for all m' , n' , and p , if $p^{n'} \leq 2^{m'}$, then $v(p^{n'}) \leq v(2^{m'})$, i.e., $(v(p))^{n'} \leq (v(2))^{m'}$. The inequality $p^{n'} \leq 2^{m'}$ is equivalent to $n' \cdot \ln(p) \leq m' \cdot \ln(2)$, i.e., to $\frac{\ln(p)}{\ln(2)} \leq \frac{m'}{n'}$. Also, the inequality $(v(p))^{n'} \leq (v(2))^{m'}$ is equivalent to $\frac{\ln(v(p))}{\ln(v(2))} \leq \frac{m'}{n'}$. Thus, the above conclusion

$$\text{if } p^{n'} \leq 2^{m'}, \text{ then } (v(p))^{n'} \leq (v(2))^{m'}$$

takes the following form:

$$\text{for every rational number } \frac{m'}{n'}, \text{ if } \frac{\ln(p)}{\ln(2)} \leq \frac{m'}{n'} \text{ then } \frac{\ln(v(p))}{\ln(v(2))} \leq \frac{m'}{n'}.$$

Let us denote $\gamma \stackrel{\text{def}}{=} \frac{\ln(v(2))}{\ln(2)}$ and $\beta \stackrel{\text{def}}{=} \frac{\ln(v(p))}{\ln(p)}$. For every $\varepsilon > 0$, there exist rational numbers $\frac{m}{n}$ and $\frac{m'}{n'}$ for which $\gamma - \varepsilon \leq \frac{m}{n} \leq \gamma \leq \frac{m'}{n'} \leq \gamma + \varepsilon$. For these numbers, the above two properties imply that $\frac{m}{n} \leq \beta$ and $\beta \leq \frac{m'}{n'}$ and thus, that $\gamma - \varepsilon \leq \beta \leq \gamma + \varepsilon$, i.e., that $|\gamma - \beta| \leq \varepsilon$. This is true for all $\varepsilon > 0$, so we conclude that $\beta = \gamma$, i.e., that $\frac{\ln(v(p))}{\ln(v(2))} = \gamma$. Hence, $\ln(v(p)) = \gamma \cdot \ln(p)$ and thus, $v(p) = p^\gamma$ for all integers p .

3.2.2°. We can reach a similar conclusion $v(p) = p^\gamma$ when the sequence $v(n)$ is decreasing.

3.3°. By definition of $v(n)$, we have $f(v(m), v(m')) = v(m + m')$. Thus, we have

$$f(m^\gamma, (m')^\gamma) = (m + m')^\gamma.$$

By using scale-invariance with $\lambda = n^{-\gamma}$, we get

$$f\left(\frac{m^\gamma}{n^\gamma}, \frac{(m')^\gamma}{n^\gamma}\right) = \frac{(m + m')^\gamma}{n^\gamma}.$$

Thus, for $a = \frac{m^\gamma}{n^\gamma}$ and $b = \frac{(m')^\gamma}{n^\gamma}$, we get $f(a, b) = (a^\alpha + b^\alpha)^{1/\alpha}$, where $\alpha \stackrel{\text{def}}{=} 1/\gamma$.

Rational numbers $r = \frac{m}{n}$ are everywhere dense on the real line, hence the values r^γ are also everywhere dense, i.e., every real number can be approximated, with any given accuracy, by such numbers. Thus, continuity implies that $f(a, b) = (a^\alpha + b^\alpha)^{1/\alpha}$ for every two real numbers a and b .

The proposition is proven.

Discussion. For $\alpha = 1$, we get the usual formula for the probability for the event $A \vee B$ when A and B are disjoint. For $\alpha = 0.5$, we get the quantum formula. Thus, we get the desired justified general formula or which traditional probabilistic formula and the quantum formula are particular cases.

5 Relation to Entropy

What we do in this section. Let us start our analysis of the resulting general formula. In this section, we show that this formula has an interesting relation to Shannon's entropy.

Informal analysis of the problem. As we have mentioned, most aspects of human behavior and human decision making can be described by the usual probabilistic formulas. This means that while some deviations from the usual formulas are needed – to take into account some aspects of human behavior which are better described by quantum formulas – the corresponding value α should be close to the value $\alpha = 1$ corresponding to the usual probabilistic case.

Since quantum formulas seem to capture some aspects of human behavior, and quantum formulas correspond to $\alpha < 1$, this means that the actual value α is close to 1 and smaller than 1. Thus, we can conclude that $\alpha = 1 - \varepsilon$ for some small $\varepsilon > 0$.

The fact that ε is small means that we can safely ignore terms which are quadratic (or of higher order) in terms of ε , and keep only terms which are linear in ε . Let us see how the above formula can be this simplified.

Formal analysis of the problem. For every value a , we have $a^\alpha = a^{1-\varepsilon} = a \cdot a^{-\varepsilon}$. Here, since $a = \exp(\ln(a))$, we get

$$a^{-\varepsilon} = (\exp(\ln(a)))^{-\varepsilon} = \exp(-\varepsilon \cdot \ln(a)) \approx 1 - \varepsilon \cdot \ln(a).$$

Thus,

$$a^\alpha \approx a \cdot (1 - \varepsilon \cdot \ln(a)) = a - \varepsilon \cdot a \cdot \ln(a).$$

Thus,

$$(\rho_1(x))^\alpha + (\rho_2(x))^\alpha \approx \rho_1(x) + \rho_2(x) - \varepsilon \cdot (\rho_1(x) \cdot \ln(\rho_1(x)) + \rho_2(x) \cdot \ln(\rho_2(x))).$$

Similarly, since

$$\frac{1}{\alpha} = \frac{1}{1 - \varepsilon} \approx 1 + \varepsilon,$$

we get

$$a^{1/\alpha} = a + \varepsilon \cdot a \cdot \ln(a).$$

Thus,

$$\begin{aligned} \rho(x) &= ((\rho_1(x))^\alpha + (\rho_2(x))^\alpha)^{1/\alpha} \approx \\ &(\rho_1(x) + \rho_2(x) - \varepsilon \cdot (\rho_1(x) \cdot \ln(\rho_1(x)) + \\ &\rho_2(x) \cdot \ln(\rho_2(x)))) + \varepsilon \cdot (\rho_1(x) + \rho_2(x)) \cdot \ln(\rho_1(x) + \ln(\rho_2(x))). \end{aligned}$$

Resulting formula and its relation to Shannon's entropy. Thus, we arrive at the following formula:

$$\rho(x) \approx \rho_1(x) + \rho_2(x) + \varepsilon \cdot \Delta\rho(x),$$

where we denoted

$$\Delta\rho(x) \stackrel{\text{def}}{=} -\rho_1(x) \cdot \ln(\rho_1(x)) - \rho_2(x) \cdot \ln(\rho_2(x)) - \\ (-\rho_1(x) + \rho_2(x)) \cdot \ln(\rho_1(x) + \rho_2(x)).$$

The sum of expressions of the type $-\rho(x) \cdot \ln(\rho(x))$ is exactly what we see in the formula for Shannon's entropy, so we get a (unexpected but clear) connection of quantum-type effects with Shannon's entropy.

6 Fuzzy and Copula Interpretations of Our General Formula

It is not easy to interpret our formula in probabilistic terms. In general, in the traditional probability theory, the probability $P(A \vee B)$ of the disjunction $A \vee B$ cannot exceed the sum of the two corresponding probabilities: $P(A \vee B) \leq P(A) + P(B)$. Similarly, for the probability density functions $\rho_1(x)$ and $\rho_2(x)$, we should have $\rho(x) \leq \rho_1(x) + \rho_2(x)$.

However, in the quantum case,

$$\rho(x) = \rho_1(x) + \rho_2(x) + 2\sqrt{\rho_1(x) \cdot \rho_2(x)} > \rho_1(x) + \rho_2(x).$$

This is one of the results showing that quantum formulas cannot be interpreted in terms of the traditional probabilities.

A similar inequality holds for all possible values $\alpha < 1$. So what shall we do?

Comment. The fact that some aspects of human behavior cannot be adequately described in probabilistic terms is well known; see, e.g., [6]. For example, in certain situations, people estimate the probability that a person X is a professional *and* a feminist as higher than the probability that this person is a feminist – an inequality which is impossible for probabilities, since the probability of a sub-event cannot exceed the probability of the original super-event.

First option: fuzzy interpretation. A natural first option is to take into account that, since here, $P(A \vee B) > P(A) + P(B)$, the values $P(A)$, $P(B)$, and $P(A \vee B)$ are not real probabilities, they are non-probabilistic degrees of certainty. Thus, to describe these degrees, it is reasonable to consider the most well-known non-probabilistic uncertainty formalism: the formalism of fuzzy logic; see, e.g., [2, 7, 11, 13, 17].

Second option: copula interpretation. Another option is to explicitly take into account that, e.g., $\rho_1(x)$ is the probability that the particle passes through the first slot *and* did not pass through the second slot.

This is a known possible interpretation of the above feminist paradox – that when people are asked to compare the probability that X is a feminist and the probability that X is a professional and a feminist, they interpret the first option as saying that X is a feminist but *not* a professional.

If we denote by A_1 the event that the particle passed through the first slot, this means that $\rho_1(x)$ is *not* the probability of this event A_1 , but rather the probability of a composite event $A'_1 \stackrel{\text{def}}{=} A_1 \& (\neg A_2)$, i.e., the value $P(A_1) - P(A \& A_2)$. Similarly, $\rho_2(x)$ is the probability of $A'_2 \stackrel{\text{def}}{=} A_2 \& (\neg A_1)$, i.e., $P(A_2) - P(A_1 \& A_2)$.

Here, $\rho(x)$ is the probability of $A_1 \vee A_2$, i.e., the probability $P(A_1 \vee A_2) = P(A_1) + P(A_2) - P(A_1 \& A_2)$. Thus, our formula $(\rho(x))^\alpha = (\rho_1(x))^\alpha + (\rho_2(x))^\alpha$ can be interpreted as the following indirect formula for determining the corresponding *copula* $C(u, v)$ [5, 9, 10, 12], i.e., the function that described the probability $C(u, v) \stackrel{\text{def}}{=} P(A_1 \& A_2)$ in terms of the probabilities $u \stackrel{\text{def}}{=} P(A_1)$ and $v \stackrel{\text{def}}{=} P(A_2)$:

$$(u + v - C(u, v))^\alpha = (u - C(u, v))^\alpha + (v - C(u, v))^\alpha.$$

Comment. The copula interpretation is related to a fuzzy one. Indeed, in the copula interpretation, we started with the assumption that the particle cannot go through both slots, i.e., that going through the second slot is equivalent to the negation of going through the first slot. We ended up by realizing that, to make sense of the corresponding formulas for the probabilities, we need to allow a non-zero probability that the particle goes through both slots.

This is exactly what fuzzy does: instead of assuming that a person is either young or not young, it takes into account that the same person can be to some extent young and to some extent not young.

In the quantum case, we have an explicit expression for the corresponding copula. In general, from the above formula, we cannot extract an explicit expression for $C \stackrel{\text{def}}{=} C(u, v)$, but it is possible in the quantum case, when $\alpha = 1/2$. In this case, by squaring both sides of the above equation, we conclude that

$$u + v - C = u - C + v - C + 2\sqrt{(u - C) \cdot (v - C)}.$$

By cancelling equal terms on both sides and moving C to the left-hand side, we conclude that $C = 2\sqrt{(u - C) \cdot (v - C)}$. Squaring both sides of this equality, we get $C^2 = 4u \cdot v - 4(u + v) \cdot C + 4C^2$, i.e., the quadratic equation

$$3C^2 - 4(u + v) \cdot C + 4u \cdot v = 0,$$

with an explicit solution

$$C(u, v) = \frac{2(u + v) \pm \sqrt{4(u + v)^2 - 48u \cdot v}}{6} = \frac{u + v \pm \sqrt{(u + v)^2 - 12u \cdot v}}{3}.$$

Comment. It is worth mentioning that, as one can check, the resulting copula is *not* associative.

Acknowledgments

We acknowledge the partial support of the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand. This work was also supported in part by the US National Science Foundation grant HRD-1242122.

References

1. B. E. Baaquie, *Quantum Finance: Path Integrals and Hamiltonians for Options and Interest Rates*, Cambridge University Press, New York, 2004.
2. R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.
3. J. R. Busenmeyer and P. D. Bruza, *Quantum Models of Condition and Decision*, Cambridge University Press, New York, 2012.
4. R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.
5. P. Jaworski, F. Durante, W. K. Härdle, and T. Rychlik (eds.), *Copula Theory and Its Applications*, Springer Verlag, Berlin, Heidelberg, New York, 2010.
6. D. Kahneman, *Thinking Fast and Slow*, Farrar, Straus, and Giroux, New York, 2011.
7. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
8. V. Kreinovich, H. T. Nguyen, and S. Sriboonchitta, “Quantum ideas in economics beyond quantum econometrics”, *Proceedings of the 2018 International Econometrics Conference*, Hanoi, Vietnam, January 15–16, 2018, to appear.
9. J.-F. Mai and M. Scherer, *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*, World Scientific, Singapore, 2017.
10. A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton University Press, Princeton, New Jersey, 2015.
11. J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.
12. R. B. Nelsen, *An Introduction to Copulas*, Springer Verlag, Berlin, Heidelberg, New York, 2007.
13. H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
14. R. Penrose, *The Emperor’s New Mind: Concerning Computers, Minds and The Laws of Physics*, Oxford University Press, Oxford, UK, 1989.
15. *Proceedings of the Conference “Dialogue on personal computers IVERSI-85*, Tbilisi, 1985 (in Russian).
16. *Proceedings of the Conference on Semiotic Aspects of Formalizing Intelligent Activity*, Kutaisi-85, Moscow, 1985 (in Russian).
17. L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.