Why Clayton and Gumbel Copulas: A Symmetry-Based Explanation

Vladik Kreinovich[1], Hung T. Nguyen[2,3], and Songsak Sriboonchitta[3]

[1] Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA, vladik@utep.edu
[2] Department of Mathematical Sciences, New Mexico State University
Las Cruces, New Mexico 88003, USA, hunguyen@nmsu.edu
[3] Department of Economics, Chiang Mai University
Chiang Mai, Thailand, songsak@econ.chiangmai.ac.th

Abstract. In econometrics, many distributions are non-Gaussian. To describe dependence between non-Gaussian variables, it is usually not sufficient to provide their correlation: it is desirable to also know the corresponding copula. There are many different families of copulas; which family shall we use? In many econometric applications, two families of copulas have been most efficient: the Clayton and the Gumbel copulas. In this paper, we provide a theoretical explanation for this empirical efficiency, by showing that these copulas naturally follow from reasonable symmetry assumptions. This symmetry justification also allows us to provide recommendations about which families of copulas we should use when we need a more accurate description of dependence.

Keywords: Archimedean copulas, econometrics, symmetries

1 Formulation of the Problem

Copulas are needed. Traditionally, in statistics the dependence between random variables η, ν, ..., is described by their correlation. This description is well justified in the frequent cases when the joint distribution is Gaussian: in this case, to describe the joint distribution, i.e., to describe the corresponding cumulative distribution function $P(\eta \leq x \,\&\, \nu \leq y \,\&\, \ldots)$, it is sufficient to describe the marginal distribution $F_\eta(x) \stackrel{\text{def}}{=} P(\eta \leq x)$, $F_\nu(y) \stackrel{\text{def}}{=} P(\nu \leq y)$, ..., of each of the variables, and the correlations between each pairs of variables.

In many practical situations, e.g., in economics, the distributions are often non-Gaussian; see, e.g., [4]. For non-Gaussian pair of variables (η, ν), in general, it is not enough to know the distribution of each variables and the correlations between them, we need more information about the dependence. Such information is provided, e.g., by a *copula*, i.e., by a function $C(u, v)$ for which $P(\eta \leq x \,\&\, \nu \leq y) = C(F_\eta(x), F_\nu(y))$; see, e.g., [3–5].

Usually, Archimedean copulas are used, i.e., copulas of the form $C(u, v) = \psi(\psi^{-1}(u) + \psi^{-1}(v))$ for some decreasing *generator function* $\psi(x)$ that maps $[0, \infty)$ into $(0, 1]$.

Most efficient Archimedean copulas. In econometric applications, the following two classes of Archimedean copulas turned out to be most efficient [4]: the Frank copulas $C(u,v)=-\dfrac{1}{\theta}\cdot\ln\left(1-\dfrac{(1-\exp(-\theta\cdot u))\cdot(1-\exp(-\theta\cdot v))}{1-\exp(-\theta)}\right)$, the Clayton copulas $C(u,v)=\left(u^{-\theta}+v^{-\theta}-1\right)^{-1/\theta}$ and the Gumbel copulas $C(u,v)=\exp\left(\left((-\ln(u))^{-\theta}+(-\ln(v))^{-\theta}-1\right)^{-1/\theta}\right)$. The efficiency of Frank's copulas is clear: Frank copulas are the only Archimedean copulas which satisfy the natural condition $C(u,v)+C(u,1-v)+C(1-u,v)+C(1-u,1-v)=1$ that describes the intuitive idea that for every two events U and V, we should have $P(U\,\&\,V)+P(U\,\&\,\neg V)+P(\neg U\,\&\,V)+P(\neg U\,\&\,\neg V)=1$ (see, e.g., [3–5]). But why Clayton and Gumbel copulas?

Main question: why Clayton and Gumbel copulas? In principle, there are many different copulas. So why did the above two classes turned out to be the most efficient in econometrics?

Auxiliary question: what if these copulas are not sufficient? While at present, the above two classes of copulas provide a good description of all observed dependencies, in the future, we will need to describe this dependence in more detail, so we will larger classes of copulas. Which classes should we use?

What we do in this paper. In this paper, we provide an answers to both questions. Specifically, we show that natural symmetry-based ideas indeed explain the efficiency of the above classes of copulas, and that these same ideas can lead us, if necessary, to more general classes.

2 Why Symmetries

Symmetries as a fundamental description of knowledge: brief reminder. Symmetries are one of the fundamental concepts of modern physics. The reason for their ubiquity is that most of our knowledge is based on symmetry; see, e.g., [1].

Indeed, how do we gain any knowledge about the physical world? Let us start with a simple example: we observe many times that the Sun rises every morning, and we conclude that it will rise again. This conclusion is based on the implicit assumption that the dynamics of the Solar system does not change when we move from one day to another.

Similarly, we drop a rock, and it falls down with an acceleration of 9.81 m/sec^2. We repeat this experiment at different locations on the Earth, we repeat it at the same location turning to different places, and we always get the same acceleration. We therefore conclude that at all locations on the Earth surface, no matter what our orientation is, the rock will drop with the same acceleration. This means that no matter how we shift or rotate, the fundamental laws of physics do not change.

In general, when we formulate a physical law based on observations, we thus implicitly assume that new situations are similar to the already observed ones, so the regularities that we observed earlier will happen in future situations as well. This idea has been formalized in modern physics, to the extend that may physical theories (starting with the quarks theory) are formulated not in terms of differential equations as before, but explicitly in terms of appropriate symmetries [1]. Moreover, it was discovered that many fundamental physical equations – e.g., Maxwell equations of electrodynamics, Schrödinder's equations of quantum mechanics, Einstein's equations of General Relativity etc. – can be uniquely derived from the corresponding symmetries; see, e.g., [2].

In view of efficiency of symmetries in physics, it is reasonable to use them in other disciplines as well; for example, in [6], we have shown that symmetries can be efficiently applied to computing.

Basic symmetries. In this paper, we will use the *basic* symmetries that come from the fact that the numerical value of a physical quantity depends on the choice of the measuring unit and on the choice of a starting point.

Let us start with the choice of the measuring unit. For example, when we measure lengths and instead of using meters, start using centimeters – a unit which is $\lambda = 100$ times smaller than the meter – instead of the original numerical values x, we get new values $x' = \lambda \cdot x$ which are λ times larger. Many fundamental physical processes do not have any preferred unit of length; for such processes, it is reasonable to require that the corresponding equations do not change if we simply change the units. The corresponding transformations are called *scalings*, and invariance under such transformation is known as *scale-invariance.* It is worth mentioning that scale-invariance is an important part of symmetry-based derivation of the fundamental physical equations presented in [2].

Another basic symmetry is the possibility to select different starting points for measurements. For example, when we measure time, we can arbitrarily select the starting point: instead of the usual calendar that starts at Year 0, we can start, as the French Revolution proposed, so start with the date of the Revolution. In this case, instead of the original numerical value x, we get a new value $x' = x + s$, where s is the difference between the starting points (e.g., $s = -1789$ for the French Revolution). Many fundamental physical quantities like time do not have any preferred starting point; for such processes, it is reasonable to require that the corresponding equations do not change if we simply change the starting point. The corresponding transformations are called *shifts*, and invariance under such transformation is known as *shift-invariance.*

3 Invariant Functions Corresponding to Basic Symmetries

Example of invariance. A power law $y = x^a$ has the following invariance property: if we change a unit in which we measure x, then in the new units, we get the exact same formula – provided that we also appropriately changing a measuring unit for y. Let us explain this property in detail.

If we replace the original unit for measuring x by a new measuring unit which is λ times larger, then all the numerical values get decreased by a factor of λ. In other words, instead of the original values x, we have new values $x' = \frac{x}{\lambda}$. How will the dependence of y on x look in the new units?

From the above dependence of x' on x, we conclude that $x = \lambda \cdot x'$. Substituting this expression into the formula $y = x^a$, we conclude that $y = (\lambda \cdot x')^a = \lambda^a \cdot (x')^a$. This new formula is different from the original formula – it has not only x' raised to the power a, it also has a multiplicative constant $\mu \stackrel{\text{def}}{=} \lambda^a$. However, we can make this formula exactly the same if we also select a new unit for y: namely, a unit which is μ times larger than the original one. Now, instead of the original values y, we get new values $y' = \frac{y}{\mu}$. In these new units, due to $y = \mu \cdot (x')^a$, the dependence of y on x takes the form $y' = (x')^a$ – i.e., exactly the same form as in the previous units.

Towards a general description of invariant functions corresponding to basic symmetries. Let us now provide a general description of invariant functions corresponding to basic symmetries. As we have mentioned, there are two types of basic symmetries: scaling (corresponding to a change in measuring unit) and shift (corresponding to the change in the starting point). When we are looking for invariant functions $y = f(x)$, we have 2 possible symmetries for x and 2 possible symmetries for y, so we need to consider all $2 \times 2 = 4$ possible combinations of these symmetries. Let us describe what happens in all these 4 cases: scale $\rightarrow$ scale, scale $\rightarrow$ shift, shift $\rightarrow$ scale, and shift $\rightarrow$ shift.

Definition 1. *A differentiable function $f(x)$ is called* scale-to-scale invariant *if for every λ, there exists a μ for which $f(\lambda \cdot x) = \mu \cdot f(x)$.*

Comment. In this case, if we replace x with $x' = \lambda \cdot x$, we can get the same dependence $y' = f(x')$ if we replace y with $y' = \frac{y}{\mu}$.

Proposition 1. *A function $f(x)$ is scale-to-scale invariant if and only if it has the form $f(x) = A \cdot x^a$ for some real numbers A and a.*

F or readers' convenience, all the proofs are placed in a special (last) section.

Definition 2. *A differentiable function $f(x)$ is called* scale-to-shift invariant *if for every λ, there exists an s for which $f(\lambda \cdot x) = f(x) + s$.*

Comment. In this case, if we replace x with $x' = \lambda \cdot x$, we can get the same dependence $y' = f(x')$ if we replace y with $y' = y - s$.

Proposition 2. *A function $f(x)$ is scale-to-shift invariant if and only if it has the form $f(x) = A \cdot \ln(x) + b$ for some real numbers A and b.*

Definition 3. *A differentiable function $f(x)$ is called* shift-to-scale invariant *if for every s, there exists a λ for which $f(x+s) = \lambda \cdot f(x)$.*

Comment. In this case, if we replace x with $x' = x + s$, we can get the same dependence $y' = f(x')$ if we replace y with $y' = \frac{y}{\lambda}$.

Proposition 3. *A function $f(x)$ is shift-to-scale invariant if and only if it has the form $f(x) = A \cdot \exp(k \cdot x)$ for some real numbers A and k.*

Definition 4. *A differentiable function $f(x)$ is called* shift-to-shift invariant *if for every s, there exists a b for which $f(x+s) = f(x) + b$.*

Comment. In this case, if we replace x with $x' = x + s$, we can get the same dependence $y' = f(x')$ if we replace y with $y' = y - b$.

Proposition 4. *A function $f(x)$ is shift-to-shift invariant if and only if it has the form $f(x) = A \cdot x + c$ for some real numbers A and c.*

Definition 5. *A function is called* invariant *if it is either scale-to-scale invariant, or scale-to-shift invariant, or shift-to-scale invariant, or shift-to-shift invariant.*

Discission. Not all physical dependencies are invariant. Specifically, when the mappings $y = f(x)$ from x to y and $z = g(y)$ are both invariant with respect to the same symmetries, then their composition $z = g(f(x))$ is also invariant with respect to the same symmetries. In general, however, the mappings $y = f(x)$ and $z = g(y)$ correspond to different symmetries; in this case, their composition is not necessarily invariant.

In view of this observation, if we want to use symmetries but cannot find an invariant function, we should be look for functions which are compositions of two invariant functions (if necessary, compositions of three, etc.) Let us apply this approach to our problem of finding appropriate copulas.

4 Why Scalings and Shifts Can Be Applied to Probabilities

Our objective is to derive the formulas for copulas, i.e., formulas that transform probabilities into probabilities. Due to the fundamental nature of symmetries, we plan to use symmetries in this derivation. The basic symmetries are scalings and shifts, symmetries which are justified by the possibility to select different measuring units and different starting points. The challenge here is that this justification cannot be directly applied to probabilities: since probabilities are limited by the interval $[0, 1]$, for probabilities, 0 is a natural starting point, and 1 is a natural measuring unit.

We will show, however, that, by using arguments which are slightly more complex than in the general case, we can still justify the use of scalings and shifts in the probabilistic context. Indeed, the possibility of scaling naturally comes from the fact that most of the probabilities that we analyze are, in effect, *conditional* probabilities, and the numerical values of these probabilities change if we change the context. Let us give a simplified example.

Let us assume that we want to invest in stable stocks, and we have selected several such stocks, i.e., stocks that only experience a drastic change in price when the market as a whole starts changing. We want to gauge stability of several such stocks. One way to estimate such a stability is divide the number of days c when this stock drastically changed by the total number of days N during which we kept the records. Alternatively, since the stock only changes when the market itself changes, we can divide c by the total number of days $n < N$ when the market drastically changed. The two resulting probabilities $p = \frac{c}{N}$ and $p' = \frac{c}{n}$ differ by a multiplicative constant $p' = \lambda \cdot p$, where $\lambda \stackrel{\text{def}}{=} \frac{n}{N}$. Thus, in econometric applications, scaling makes sense for probabilities.

Let us give another simplified example. Suppose that we have a stock which always fluctuates when the market changes *and* also sometimes experiences drastic changes of its own. How can we estimate the stability of this stock? One way is to divide the number of days c when this stock drastically changed by the total number of days N during which we kept the record of this stock, and get an estimate $p = \frac{c}{N}$. However, since we know that this stock always changes when the market changes, it makes sense to only consider days when the market itself was stable, i.e., to use the estimate $p' = \frac{c-n}{N-n}$. Since $n \ll N$, we have $p' \approx \frac{c-n}{N} = p + s$, where $s \stackrel{\text{def}}{=} -\frac{n}{N}$. Thus, in econometric applications, shifts also make sense for probabilities.

5 Archimedean Copulas with Whose Generators Are Either Invariant or Compositions of Two Invariant Functions

Now that we have given arguments that symmetries – including basic symmetries such as scalings and shifts – can be applied to econometric copulas, let us describe the corresponding results. Let us start by describing all the Archimedean copulas in which the generator is invariant.

Proposition 5. *The only Archimedean copula with an invariant generator is the copula $C(u, v) = u \cdot v$ corresponding to independence.*

Discussion. This result shows that to describe dependence, it is not sufficient to use invariant generators, we need to consider *compositions* of invariant generators.

Proposition 6. *The only Archimedean copulas in which a generator is a composition of two invariant functions are the following:*

- *Clayton copulas* $C(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}$*;*
- *the the Gumbel copulas* $C(u,v) = \exp\left(\left((-\ln(u))^{-\theta} + (-\ln(v))^{-\theta} - 1\right)^{-1/\theta}\right)$*;*
- *the copulas* $C(u,v) = \dfrac{1}{L}\exp\left(\dfrac{1}{\ell}\cdot\ln(u\cdot L)\cdot\ln(v\cdot L)\right)$*, with* $\ell = \ln(L)$*.*

Comment. Please notice that we have an additional family of copulas.

Discussion. This approach leads to a natural answer to a question of which copulas we should use when we the approximation provided by the existing Archimedean copulas is no longer sufficient:

- we should first try Archimedean copulas whose generator function is a composition of *three* invariant functions,
- if needed, we should move to Archimedean copulas whose generator function is a composition of *four* invariant functions,
- etc.

Example. The generator $\psi(x) = -\dfrac{1}{\theta}\cdot\ln(1-(1-\exp(-\theta))\cdot\exp(-x))$ of Frank copulas can be obtained as a composition of three invariant transfromations: first, we apply an invariant function $y = f(x) = (1-\exp(-\theta))\cdot\exp(-x)$, then an invariant function $z = g(y) = 1 - y$, and finally, an invariant function $t = h(z) = -\dfrac{1}{\theta}\cdot\ln(z)$.

6 Proofs

Proof of Proposition 1. It is easy to check that every function $f(x) = A\cdot x^a$ is scale-to-scale invariant. Vice versa, let $f(x)$ be a scale-to-scale invariant function. By definition, this means that for every λ, there exists a μ (depending on this λ) for which $f(\lambda\cdot x) = \mu(\lambda)\cdot f(x)$.

This property is trivially true when $f(x) = 0$ for all x. It is therefore sufficient to consider the cases when the function $f(x)$ is not identically 0. Let us prove, by contradiction, that in such cases, the function $f(x)$ cannot attain zero values for $x \neq 0$. Indeed, if $f(x_0) = 0$ for some $x_0 \neq 0$, then, for every other x, we will get $f(x) = \mu\left(\dfrac{x}{x_0}\right)\cdot f(x_0) = 0$. So, $f(x) \neq 0$ for $x \neq 0$.

Here, the function $f(x)$ is differentiable, the function $f(\lambda\cdot x)$ is also differentiable, and thus, their ratio $\mu(\lambda) = \dfrac{f(\lambda\cdot x)}{f(x)}$ is also differentiable. Differentiating both sides of the equation $f(\lambda\cdot x) = \mu(\lambda)\cdot f(x)$ by λ, we conclude that $x\cdot f'(\lambda\cdot x) = \mu'(\lambda)\cdot f(x)$. In particular, for $\lambda = 1$, we get $x\cdot f'(x) = \mu_0\cdot f(x)$, where we denoted $\mu_0 \stackrel{\text{def}}{=} \mu'(1)$. This equation can be rewritten as $x\cdot\dfrac{df}{dx} = \mu_0\cdot f$. In this equation, we can separate variables by moving all the terms containing df and f to the left side and all the terms containing dx and x to the right side. As a result, we get $\frac{df}{f} = \mu_0 \cdot \frac{dx}{x}$. Integrating both sides, we get $\ln(f) = \mu_0 \cdot \ln(x) + C$ for some constant C. Thus, we conclude that $f(x) = \exp(\ln(f(x))) = \exp(\mu_0 \cdot \ln(x) + C) = \exp(C) \cdot x^{\mu_0}$, which is exactly the desired form for the transformation $f(x)$, with $A = \exp(C)$ and $a = \mu_0$. The proposition is proven.

Proof of Proposition 2. It is easy to check that every function $f(x) = A \cdot \ln(x) + b$ is scale-to-shift invariant. Vice versa, let $f(x)$ be a scale-to-shift invariant function. By definition, this means that for every λ, there exists an s (depending on this λ) for which $f(\lambda \cdot x) = f(x) + s(\lambda)$.

Here, the function $f(x)$ is differentiable, the function $f(\lambda \cdot x)$ is also differentiable, and thus, their difference $s(\lambda) = f(\lambda \cdot x) - f(x)$ is also differentiable. Differentiating both sides of the equation $f(\lambda \cdot x) = f(x) + s(\lambda)$ by λ, we conclude that $x \cdot f'(\lambda \cdot x) = s'(\lambda)$. In particular, for $\lambda = 1$, we get $x \cdot f'(x) = s_0$, where we denoted $s_0 \stackrel{\text{def}}{=} s'(1)$. This equation can be rewritten as $x \cdot \frac{df}{dx} = s_0$. In this equation, we can separate variables by moving all the terms containing df and f to the left side and all the terms containing dx and x to the right side. As a result, we get $df = s_0 \cdot \frac{dx}{x}$. Integrating both sides, we get $f = s_0 \cdot \ln(x) + C$ for some constant C. This is exactly the desired form for the transformation $f(x)$. The proposition is proven.

Proof of Proposition 3. It is easy to check that every function

$$f(x) = A \cdot \exp(k \cdot x)$$

is shift-to-scale invariant. Vice versa, let $f(x)$ be a shift-to-scale invariant function. By definition, this means that for every shift s, there exists a λ (depending on this s) for which $f(x + s) = \lambda(s) \cdot f(x)$.

This property is trivially true when $f(x) = 0$ for all x. It is therefore sufficient to consider the cases when the function $f(x)$ is not identically 0. Let us prove, by contradiction, that in such cases, the function $f(x)$ cannot attain zero values at any x. Indeed, if $f(x_0) = 0$ for some x_0, then, for every other x, we will get $f(x) = \lambda(x - x_0) \cdot f(x_0) = 0$. So, $f(x) \neq 0$ for all x.

Here, the function $f(x)$ is differentiable, the function $f(x + s)$ is also differentiable, and thus, their ratio $\lambda(s) = \frac{f(x + s)}{f(x)}$ is also differentiable. Differentiating both sides of the equation $f(x + s) = \lambda(s) \cdot f(x)$ by s, we conclude that $f'(x + s) = \lambda'(s) \cdot f(x)$. In particular, for $s = 0$, we get $f'(x) = \lambda_0 \cdot f(x)$, where we denoted $\lambda_0 \stackrel{\text{def}}{=} \lambda'(0)$. This equation can be rewritten as $\frac{df}{dx} = \lambda_0 \cdot f$. In this equation, we can separate variables by moving all the terms containing df and f to the left side and all the terms containing dx and x to the right side. As a result, we get $\frac{df}{f} = \lambda_0 \cdot dx$. Integrating both sides, we get $\ln(f) = \lambda_0 \cdot x + C$ for some constant C. Thus, we conclude that $f(x) = \exp(\ln(f(x))) = \exp(\lambda_0 \cdot x + C) = \exp(C) \cdot \exp(\lambda_0 \cdot x)$, which is exactly the desired form for the transformation $f(x)$, with $A = \exp(C)$ and $k = \lambda_0$. The proposition is proven.

Proof of Proposition 4. It is easy to check that every function $f(x) = A \cdot x + b$ is shift-to-shift invariant. Vice versa, let $f(x)$ be a shift-to-shift invariant function. By definition, this means that for every shift s, there exists a b (depending on this s) for which $f(x + s) = f(x) + b(s)$.

Here, the function $f(x)$ is differentiable, the function $f(x + s)$ is also differentiable, and thus, their difference $b(s) = f(x + s) - f(x)$ is also differentiable. Differentiating both sides of the equation $f(x + s) = f(x) + b(s)$ by s, we conclude that $f'(x + s) = b'(s)$. In particular, for $a = 0$, we get $f'(x) = b_0$, where we denoted $b_0 \stackrel{\text{def}}{=} b'(0)$. Integrating this equation, we get $f = b_0 \cdot x + C$ for some constant C. This is exactly the desired form for the transformation $f(x)$. The proposition is proven.

Proof of Proposition 5. A generator $\psi(x)$ of an Archimedean copula should map 0 into 1 and ∞ into 0: $\psi(0) = 1$ and $\psi(\infty) \stackrel{\text{def}}{=} \lim\limits_{x \to \infty} = 0$. One can easily check that most invariant functions do not satisfy this property:

- the function $f(x) = A \cdot x^a$ does not satisfy the property $f(0) = 1$;
- the function $f(x) = A \cdot \ln(x) + b$ does not satisfy the property $f(0) = 1$, and
- the function $f(x) = A \cdot x + b$ does not satisfy the property $f(\infty) = 0$.

The only remaining invariant function is $f(x) = A \cdot \exp(k \cdot x)$. For this function, from $f(0) = 1$, we conclude that $A = 1$, and from $f(\infty) = 0$, that $k < 0$. One can check that for this generator function $\psi(x) = \exp(-|k| \cdot x)$, the inverse is equal to $\psi^{-1}(u) = -\dfrac{1}{|k|} \cdot \ln(u)$, and thus, the corresponding copula has the form $C(u, v) = \psi(\psi^{-1}(u) + \psi^{-1}(v)) = u \cdot v$. The proposition is proven.

Proof of Proposition 6. We have 4 types of invariant functions $f(x)$ and 4 types of invariant functions $g(y)$, so we have $4 \times 4 = 16$ possible compositions $\psi(x) = g(f(x))$. Let us consider them one by one.

1°. Let us first consider the case when $f(x) = A \cdot x^a$.

1.1°. If $g(y)$ is of the same type $g(y) = B \cdot y^b$, then the composition is also of the same type, and we already know, from the proof of Proposition 5, that a function of this type cannot be a generator.

1.2°. If $g(y) = B \cdot \ln(y) + b$, then the composition has the form

$$\psi(x) = g(f(x)) = B \cdot \ln(A \cdot x^a) + b = (B \cdot a) \cdot \ln(x) + (B \cdot \ln(A) + b).$$

In this case, we cannot have $\psi(0) = 1$.

1.3°. If $g(y) = B \cdot \exp(k \cdot y)$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot \exp(k \cdot A \cdot x^a)$. The condition $\psi(0) = 1$ leads to $B = 1$ and $a > 0$, the condition $\psi(\infty) = 0$ leads to $k \cdot A < 0$. One can easily check that in this case, we get the Gumbel copula.

1.4°. If $g(y) = B \cdot y + b$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot A \cdot x^a + b$. The condition $\psi(\infty) = 0$ leads to 0, so $\psi(x) = (B \cdot A) \cdot x^a$, and the equality $\psi(0) = 1$ is not possible.

2°. Let us first consider the case when $f(x) = A \cdot \ln(x) + b$.

2.1°. If $g(y) = B \cdot y^a$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot (A \cdot \ln(x) + b)^a$. This function cannot satisfy the property $\psi(0) = 1$.

2.2°. If $g(y) = B \cdot \ln(y) + a$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot \ln(A \cdot \ln(x) + b) + a$, then we also cannot have the property $\psi(0) = 1$.

2.3°. If $g(y) = B \cdot \exp(k \cdot y)$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot \exp(k \cdot A \cdot \ln(x) + k \cdot b) = (B \cdot \exp(k \cdot b)) \cdot x^{k \cdot A}$, so we cannot have $\psi(0) = 1$.

2.4°. If $g(y) = B \cdot y + a$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot A \cdot \ln(x) + (B \cdot A + b)$, we also cannot have $\psi(0) = 1$.

3°. Let us consider the case when $f(x) = A \cdot \exp(k \cdot x)$.

3.1°. If $g(y) = B \cdot y^a$, then the composition has the form $\psi(x) = g(f(x)) = (B \cdot A^a) \cdot \exp((k \cdot a) \cdot x)$. We already know, from the proof of Proposition 5, that such generator functions lead to the independence copula.

3.2°. If $g(y) = B \cdot \ln(y) + a$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot \ln(A \cdot \exp(k \cdot x)) + a = (B \cdot k) \cdot x + (B \cdot \ln(A) + a)$, so we cannot have $\psi(\infty) = 0$.

3.3°. If $g(y) = B \cdot \exp(a \cdot y)$, then the composition takes the form

$$\psi(x) = B \cdot \exp(\ell \cdot \exp(k \cdot x)),$$

for $\ell = a \cdot A$. The condition $\psi(0) = 1$ leads to $B \cdot \exp(\ell) = 1$, so $B = \exp(-\ell)$, and $\psi(x) = \exp(\ell \cdot (\exp(k \cdot x) - 1))$.

Let us describe the copula corresponding to this generator. For that, let us first find an explicit expression for the inverse function $\psi^{-1}(u)$. From the condition that $\psi(x) = u$, we conclude that $\ell \cdot (\exp(k \cdot x) - 1) = \ln(u)$, hence $\exp(k \cdot x) = 1 + \dfrac{\ln(u)}{\ell} = \dfrac{\ln(u \cdot L)}{\ell}$, where we denoted $L \stackrel{\text{def}}{=} \exp(\ell)$. Thus, $k \cdot x = \ln\left(\dfrac{\ln(u \cdot L)}{\ell}\right)$, and $x = \dfrac{1}{k} \cdot \ln\left(\dfrac{\ln(u \cdot L)}{\ell}\right)$.

To find $C(u, v)$, we compute $x + y$, where $x = \psi^{-1}(u)$ and $y = \psi^{-1}(v)$, then compute $z = x + y$ and $C(u, v) = \psi(z)$. Here,

$$z = x + y = \frac{1}{k} \cdot \left[\ln\left(\frac{\ln(u \cdot L)}{\ell}\right) + \ln\left(\frac{\ln(v \cdot L)}{\ell}\right)\right],$$

hence $k \cdot z = \ln\left(\frac{\ln(u \cdot L)}{\ell}\right) + \ln\left(\frac{\ln(v \cdot L)}{\ell}\right)$ and $\exp(k \cdot z) = \frac{\ln(u \cdot L)}{\ell} \cdot \frac{\ln(v \cdot L)}{\ell}$.
Thus, $\ell \cdot (\exp(k \cdot z) - 1) = \frac{\ln(u \cdot L) \cdot \ln(v \cdot L)}{\ell} - \ell$, so for

$$C(u, v) = \exp(\ell \cdot (\exp(k \cdot z) - 1)),$$

we get the desired expression.

3.4°. If $g(y) = B \cdot y + a$, then the composition takes the form $\psi(x) = g(f(x)) = (B \cdot A) \cdot \exp(k \cdot x) + a$. The condition $\psi(\infty) = 0$ leads to $a = 0$, so we get an exponential generator function which, as we have mentioned, leads to the independence copula.

4°. Finally, let us consider the case when $f(x) = A \cdot x + b$.

4.1°. If $g(y) = B \cdot y^a$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot (A \cdot x + b)^a$. This generator function leads to the Clayton copulas.

4.2°. If $g(y) = B \cdot \ln(y) + a$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot \ln(A \cdot x + b) + a$. For this function, the condition $\psi(\infty) = 0$ cannot be satisfied.

4.3°. If $g(y) = B \cdot \exp(k \cdot y)$, then the composition takes the form $\psi(x) = g(f(x)) = B \cdot \exp(k \cdot A \cdot \ln(x) + k \cdot b) = (B \cdot \exp(k \cdot b) \cdot x^{k \cdot A}$. This function cannot satisfy the condition $\psi(0) = 1$.

4.4°. If $g(y) = B \cdot y + a$, then the composition $\psi(x) = g(f(x))$ is also a linear function, so we cannot have $\psi(0) = 1$.

The proposition is proven.

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 (Cyber-ShARE Center of Excellence) and DUE-0926721, by Grant 1 T36 GM078000-01 from the National Institutes of Health, and by a grant on F-transforms from the Office of Naval Research.

References

1. Feynman, R., Leighton, R., Sands, M.: The Feynman Lectures on Physics, Addison Wesley, Boston, Massachusetts (2005)
2. Finkelstein, A.M., Kreinovich, V., Zapatrin, R.R.: Fundamental physical equations uniquely determined by their symmetry groups. Springer Lecture Notes in Mathematics, Springer- Verlag, Berlin, Heidelberg, New York, Vol. 1214, 159–170 (1986)
3. Jaworski, P., Durante, F., Härdle, W.K., Ruchlik, T. (eds.), Copula Theory and Its Applications, Springer Verlag, Berlin, Heidelberg, New York (2010)
4. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management: Concepts, Techniques, Tools. Princeton University Press, Princeton, New Jersey (2005)
5. Nelsen, R.B.: An Introduction to Copulas. Springer Verlag, Berlin, Heidelberg, New York (1999)
6. Nguyen, H.T., Kreinovich, K.: Applications of Continuous Mathematics to Computer Science. Kluwer, Dordrecht (1997)