QUANTITATIVE METHODS IN ECONOMICS
Volume XXVI, No. 4, 2025, pp. 165 – 174

SKEWNESS-CORRECTED COPULA-BASED OUTLIER DETECTION FOR HIGH-FREQUENCY FINANCIAL DATA FROM THE WARSAW STOCK EXCHANGE

Marcin Dudziński https://orcid.org/0000-0003-4242-8411
Institute of Information Technology
Warsaw University of Life Sciences – SGGW, Poland
e-mail: marcin_dudzinski@sggw.edu.pl

Abstract: The objective of outlier detection is to identify rare events – the elements in data that have significantly different feature values than the rest of data. Thus, it is of high importance to conduct the outlier detection in financial data. We investigate the use of COPOD – a parameter-free anomaly detection algorithm based on application of empirical copulas and computation of the corresponding tail probabilities. Its performance is assessed on real-world data from the Warsaw Stock Exchange. We follow the theoretical framework of [Li et al. 2020] and restate its core formulas in a concise notation adapted later to the selected financial set.

Keywords: anomaly detection, Sklar's theorems, empirical copula, skewness-corrected copula-based outlier detection (COPOD) algorithm, Warsaw Stock Exchange Data

JEL classification: C51, C52

INTRODUCTION

Outliers, often known as anomalies, are data observations (instances, objects) that lie at the abnormal distances from the majority of observations in a given sample of data. Consequently, outlier points have substantially different feature values or different characteristics than the points from the remaining data. Since the existence of outliers can greatly influence the results of statistical analysis (in particular, this can distort the adequacy of statistical model), it becomes a crucial task to identify anomalous objects and later to remove them in order to secure the integrity of data, as well as to enable the proper calibration and fitting of the considered model.

https://doi.org/10.22630/MIBE.2025.26.4.15

Classical approaches of anomaly (outlier) detection range from robust covariance estimators and Mahalanobis distance-based methods to proximity-based algorithms, such as: the k-Nearest Neighbours method, the Local Outlier Factor (LOF) model, the One-class SVM algorithm, the ensemble approaches, such as the Isolation Forest setting and its generalization called the Extended Isolation Forest. In this context, the following works are worthwhile to mention: the paper by [Liu et al. 2008], where an approach based on using the Isolation Forest (IF) method is proposed, the work of [Hariri et al. 2021], where the Extended Isolation Forest (EIF) algorithm is presented, an article of [Schölkopf et al. 2001], where the method of anomaly identification using the One-Class SVM method is discussed, and the paper by [Breunig et al. 2000], where the Local Outlier Factor (LOF) – a density-based unsupervised machine learning algorithm – is considered. A comprehensive overview of anomaly detection methods is contained in [Chandola et al. 2009], while the standard monograph on outlier detection theory and its practical implications is presented in [Aggarwal 2017].

While powerful, the above mentioned approaches suffer from several limitations, namely: they often require careful hyperparameter tuning (including, e.g.: number of clusters – for clustering-based models, layer architecture – for neural network-based models, choices of individual classifiers – for ensemble models), they may not be able to cope with high-dimensional data, as well as they can be difficult to interpret in terms of marginal and joint tail probabilities.

COPOD (Copula-Based Outlier Detection) algorithm is a quite recently introduced algorithm that overcomes the mentioned limitations by providing easily interpretable method based on the definitions of empirical cumulative distribution functions and on the notion of empirical copula. The core idea behind this approach consists in approximation of the joint empirical cumulative distribution function of the data by an empirical copula obtained from marginal empirical distribution functions (edfs), and subsequently in computing the corresponding tail probabilities for sample values (observations) in order to determine their level of extremeness that is reached by calculating the corresponding outlier (anomality) scores. Since COPOD does not require any stochastic training or hyperparameter tuning, it is a deterministic, parameter-free, computationally feasible and efficient method.

A significant advantage of COPOD is that outlier scores can be decomposed into the marginal contributions, making it possible to identify which dimensions and which tails (left, right, or skewness-corrected) are responsible for indicating an observation as anomalous. For a fundamental paper on the COPOD algorithm and its skewness-corrected version, where this approach has been introduced, we refer to [Li et al. 2020].

High-frequency financial data are known to suffer from various types of irregularities, such as: recording errors, existence of extreme price jumps, bias or microstructure noise. These irregularities can seriously influence or even distort statistical analysis, more specifically the research regarding volatility estimation and risk measurement modelling. Therefore, the need to invent robust and interpretable outlier detection methods becomes vital in empirical research devoted to the fields of finance and econometrics.

The main goal of our paper is to adapt and apply the COPOD algorithm of outlier identification to real high-frequency financial data from the Warsaw Stock Exchange (GPW). We have performed an empirical study on dataset, which is a pre-processed set containing values of the following variables: price, time (in seconds), and log returns, for a single financial instrument. These values are collected in the GPW20230403000tind_for_COPOD.csv file. In order to achieve this objective, we have restated the COPOD framework in terms of empirical copulas and tail probabilities and subsequently, we have prepared a self-contained Python implementation of skewness-corrected COPOD (a version of COPOD that incorporates a skewness correction), as well as we have interpreted the detected outliers in terms of extreme returns and potential data errors. The considered dataset contains intraday transaction data for a single financial instrument from 3 April 2023.

The remainder of the paper is organized as follows. Section 2 presents theoretical framework and describes the computational implementation and design, while in Section 3 we present our numerical experiments conducted on the selected dataset and discuss the obtained results and in finalal Section 4, we conclude and summarize our research, as well as we propose the possible directions of future research. All of our numerical study has been carried out using the programming language of Python. Appropriate file with the corresponding codes can be available upon Reader's request.

SKEWNESS-CORRECTED COPOD: THEORETICAL BACKGROUND

Sklar's theorems

Construction of the COPOD algorithm and its skewness-corrected version is based on the definition of copula and conclusions from the celebrated Sklar's theorems, which may be recapitulated as follows (see, e.g.: [Sklar 1959; Schweizer and Sklar 1983; Nelsen et al. 2016]).

A copula $C\colon [0;1]^d \to [0;1]$ is defined as the joint df of a random vector $U = (U_1, U_2, \cdots, U_d)$ with the $Uniform([0;1])$ marginals. It means that:

$$C(u_1, u_2, \cdots, u_d) = P(U_1 \le u_1, U_2 \le u_2, \cdots, U_d \le u_d). \tag{1}$$

The Sklar's theorems state that, for a given joint df F with continuous marginals $F_1, F_2, \dots, F_d$, there exists a uniquely defined copula C, such that:

$$F(t_1, t_2, \dots, t_d) = C\big(F_1(t_1), F_2(t_2), \dots, F_d(t_d)\big), \tag{2}$$

and conversely that, for a given copula C and marginals $F_1, F_2, \dots, F_d$, the relation above defines a df function with marginals F_i.

A copula can be interpreted as a dependence structure, since it captures the complete dependence among random variables after eliminating the influence of their marginal distributions. The Sklar's theorems formalize this idea by establishing the correspondence between multivariate distributions and pairs consisting of marginal distributions and they allow for independent modeling of marginals and dependence structure. On the other hand, we may also choose arbitrary marginal distributions and combine them with a copula selected in order to display a desired dependence pattern. This flexibility is crucial in applications such as finance, insurance and anomaly detection.

In the present study, we approximate the joint distribution using the product copula, i.e.:

$$\hat{C}(u_1,u_2,\cdots,u_d) = \frac{1}{n}\sum_{k=1}^{n} I\big(U_{1,k} \le u_1, U_{2,k} \le u_2,\cdots,U_{d,k} \le u_d\big) = \frac{1}{n}\sum_{k=1}^{n} I\big(U_{1,k} \le u_1,\big)\cdot I\big(U_{2,k} \le u_2\big)\cdot \ldots \cdot I\big(U_{d,k} \le u_d\big) \approx u_1 \cdot u_2 \cdot \ldots u_d. \quad (3)$$

This corresponds to assuming no explicit parametric dependence structure between the marginal components. Consequently, the joint tail probability is approximated by the product of marginal tail probabilities.

For standard references on copulas and empirical copulas, applied in the Skewness-Corrected COPOD algorithm, we cite the works of [Nelsen et al. 2006] and [Deheuvels 1979].

Now, assume that we have the d-dimensional random vectors $\boldsymbol{X_i} = \big(X_{1,i}, X_{2,i},\cdots,X_{d,i}\big)$, $i = 1,2,\cdots,n$, with the joint distribution function (df) F and continuous marginal distributions: $F_1, F_2, \ldots, F_d$. We generate data from X_i and obtain n empirical d-dimensional observation vectors $\boldsymbol{x_i} = \big(x_{1,i}, x_{2,i},\cdots,x_{d,i}\big)$, $i = 1,2,\cdots,n$. Thus, our empirical data (instances) are as follows:

$$\boldsymbol{x_1} = \big(x_{1,1}, x_{2,1},\cdots,x_{d,1}\big), \boldsymbol{x_2} = \big(x_{1,2}, x_{2,2},\cdots,x_{d,2}\big),\cdots,$$
$$\boldsymbol{x_n} = \big(x_{1,n}, x_{2,n},\cdots,x_{d,n}\big).$$

Before we present the theoretical foundations of COPOD and its skewness-corrected variant, we need to introduce the following notations.

The left-tail empirical distribution functions

Having the vectors of empirical data, we define the left-tail empirical distribution function (ecdf) as follows:

$$\hat{F}_j\big(x_{j,i}\big) = \frac{1}{n}\sum_{k=1}^{n} I\big(x_{j,k} \le x_{j,i}\big), \quad (4)$$

where $j = 1,2,\cdots,d, i = 1,2,\cdots,n,$ and where – here and throughout the paper – I stands for the indicator function.

In our numerical study, the left-tail ecdf in (4) is computed using the formula:

$$\hat{F}_j\big(x_{j,i}\big) = \big\{rank_{max}(x_{j,i})\big\}/n, \quad (5)$$

where $\mathrm{rank}_{\max}(x_{j,i})$ stands for the maximum rank of $x_{j,i}$ among $\big\{x_{j,1}, x_{j,2}, \ldots, x_{j,d}\big\}$.

The right-tail empirical distribution functions

Alternatively to the left-tail case, we define the right-tail ecdfs by analogous formula, where instead of $\{x_{j,k} \leq x_{j,i}\}$, we use $\{-x_{j,k} \leq -x_{j,i}\}$ or, equivalently $\{x_{j,k} \geq x_{j,i}\}$, i.e.,

$$\hat{\hat{F}}(x_{j,i}) = \frac{1}{n}\sum_{k=1}^{n} I(-x_{j,k} \leq -x_{j,i}) = \frac{1}{n}\sum_{k=1}^{n} I(x_{j,k} \geq x_{j,i}), \quad (6)$$

where $j = 1,2,\cdots,d, i = 1,2,\cdots,n.$
In our computational study, the left-tail ecdf in (6) is calculated from formula:

$$\hat{\hat{F}}(x_{j,i}) = \{n - \mathrm{rank}_{\max}(x_{j,i})\}/n. \quad (7)$$

Skewness coefficients for each dimension

In many practical situations - including financial data - the distribution of each marginal may be skewed, and consequently – outliers may predominantly appear only in one of the tails. In order to account for this issue, a version of COPOD that incorporates a skewness correction step is employed. Namely, in the first stage of the skewness-corrected framework of the COPOD algorithm, the following skewness coefficient is considered and calculated for each dimension j:

$$b_j = \frac{\sum_{i=1}^{n}(x_{j,i}-\bar{x}_j)^2}{\left(\sqrt{\sum_{i=1}^{n}(x_{j,i}-\bar{x}_j)^2}\right)^3}, j = 1,\cdots,d. \quad (8)$$

The empirical copula observations

In the skewness-corrected COPOD variant, we define:

$$\hat{U}_{j,i} = \hat{F}_j(x_{j,i}), \hat{V}_{j,i} = \hat{\hat{F}}(x_{j,i}), \text{and:} \quad (9)$$

$$\hat{W}_{j,i} = \hat{U}_{j,i}, \text{if } b_j < 0 \text{ or } \hat{W}_{j,i} = \hat{V}_{j,i}, \text{if otherwise.} \quad (10)$$

$\hat{W}_{j,i}$ defines the skewness-corrected empirical copula observation.

Outlier scores of individual observations (for the left-tail, right-tail and skewness-corrected scenarios)

Let, for $i = 1,\cdots,n,$

$$p_{l,i} = -\sum_{j=1}^{d}\log(\hat{U}_{j,i}),\ \ p_{r,i} = -\sum_{j=1}^{d}\log(\hat{V}_{j,i}),\ \ p_{s,i} = -\sum_{j=1}^{d}\log(\hat{W}_{j,i}), \quad (11)$$

The quantities above are called as the left-tail, the right-tail and the skewness-corrected outlier scores, respectively.

The overall COPOD outlier score

The outlier scores of the individual observations $\boldsymbol{x_i} = \left(x_{1,i}, x_{2,i}, \cdots, x_{d,i}\right)$ are defined as:

$$O(\boldsymbol{x_i}) = max\left\{p_{l,i}, p_{r,i}, p_{s,i}\right\}, i = 1, \cdots, n. \tag{12}$$

Large values of $O(x_i)$ indicate strong evidence that $\boldsymbol{x_i}$ is an outlier.
We can treat the values of $O(x_i)$ as the coordinates of the vector:

$$O(\boldsymbol{X}) = [O(\boldsymbol{x_1}), O(\boldsymbol{x_2}), \cdots, O(\boldsymbol{x_n})]^T. \tag{13}$$

In order to identify anomalous observations, the empirical distribution of $\{O(\boldsymbol{x_i})\}_{i=1}^{n}$, and a quantile-based threshold is applied. In our empirical study the 95th percentile of the score distribution is used, resulting in approximately 5% of the observations being classified as outliers.

EMPIRICAL RESULTS

Dataset

The dataset considered in our numerical study consists of a pre-processed feature matrix concerning the transaction quotes of a single financial instrument traded on the Warsaw Stock Exchange (GPW).
The file GPW20230403000tind_for_COPOD.csv contains 28,624 observations and three continuous variables: *price*, *time_seconds*, and *log_return*. These features capture both the level and dynamics of the price process over a trading session. The numerical data include intraday observations for a single financial instrument from 3 April 2023.

The observations from our dataset form the high-frequency financial data. For more details on high-frequency data and financial market microstructures, we refer to [Engle 2000; O'Hara 1998; Aït-Sahalia et al. 2014]. Furthermore, in the context regarding the research study on extreme-value behavior in finance, we recommend: [Embrechts et al. 1997; Cont 2001].

We have prepared the Python code in order to implement the skewness-corrected COPOD algorithm for our datasets. This code can be available upon Reader's request. Below, we collect the results obtained through the mentioned algorithm's application.

Key Results and Conclusions Regarding an Application of the Skewness-Corrected COPOD Algorithm for the Selected Dataset

We list here the summary information on the results obtained by an application of the skewness-corrected COPOD algorithm for our dataset.

- Number of observations: **28,624**.
- Number of features: **3 (*price*, *time_seconds*, *log_return*)**,

- Outlier threshold (cutoff = 95% quantile score): $\boldsymbol{q_{0.95} \approx 7.559}$.
- Number of detected outliers: **1,432 (≈ 5% of the data)**.
- Number of observations detected as outliers: **1,432**.
- Top-20 most extreme observations detected as outliers (20 instances with the largest scores $O(x_i)$):

  $\boldsymbol{x_{10321}, x_{10317}, x_{19774}, x_{17710}, x_{10325}, x_{28133}, x_{19777}, x_{17716}, x_{19775}, x_{25678}, x_{985}, x_{17712}, x_{14266}, x_{17717}, x_{26170}, x_{989}, x_{493}, x_{19779}, x_{14763}, x_{2465}.}$
- The corresponding scores $O(x_i)$ of top 20 most extreme observations detected as outliers:

  24.993, 24.993, 24.770, 23.489, 20.176, 19.765, 19.359, 19.036, 18.674, 17.826, 17.619, 17.540, 17.294, 17.252, 17.191, 17.136.

Characterization of 20 observations with the largest COPOD scores

In this part, we present a feature-level characterization of 20 observations that achieved the largest COPOD outlier score**.** The analyzed features are: *price* – price level of the financial instrument, *time_seconds* – intraday time (in seconds), *log_return* – logarithmic return.

The observations with the largest COPOD scores are characterized by the following patterns: they have very large absolute *log-returns* (the *log-returns* empirical values are both negative and positive), they reach extremely low or extremely high *price* levels, they exhibit strong temporal concentration at specific intraday times (specifically, 22,800 seconds).

When standardized with respect to the full dataset, the following properties regarding the relation between the top 20 observations detected as outliers and the explanatory variables may be seen: these observations have moderately low ($\approx -0.9\sigma_1$) or extremely high (up to $+6\sigma_1$) values of the *price* future, values of the *time_seconds* variable are strongly concentrated around a few discrete values (($\approx -2.8\sigma_2$ for $22{,}800$ seconds), the *log_return* values are extremely heavy-tailed – approximately from $(-46)\sigma_3$ to $+102\sigma_3$, where σ_1 - σ_3 denote the corresponding standard deviations from the observations of the considered features.

Common characteristics of the top-20 skewness-corrected COPOD observations can be summarized as follows.

The most prominent property of these observations is their extreme position in the tails of the log-return distribution:

- large negative values (e.g.: −3.97, −3.59, −3.00, −2.69),
- large positive values (e.g.: +8.73, +5.21, +3.50, +2.49).

These values deeply emerge in the marginal tails and therefore exhibit very small empirical CDF values, which results in large COPOD scores.

The same top-20 skewness-corrected COPOD observations also correspond to extreme price levels:

- very low prices (e.g.: 4.60, 167.48),
- very high prices (exceeding 60,000).

The lowest detected *price* levels are small in comparison to the session *price* median of 5,458.61, while the highest *prices* represent boundary observations of the trading day.

Furthermore, with regard to the *time_seconds* feature, most of the observations occur at exactly 22,800 seconds, with a smaller number clustered around 32,000–40,500 seconds and one extreme at 61,800 seconds (the corresponding median is 47,100).

Tu sum up, we conclude that the *log_return* variable is the dominant factor of extremeness, whereas the *price* feature also influences it, but to a lesser degree. In addition, while the *time_seconds* feature is not the primary factor in the identification of rare events, but it also strengthens extremeness by placing these observations in low-density or boundary regions of the empirical distribution.

Economic and Statistical Interpretation

From an economic perspective, the observations associated with the largest values of the skewness-corrected COPOD scores $O(x_i)$ correspond to events that are highly atypical relative to the empirical distribution of intraday price dynamics. These observations are characterized by unusually large price movements, which may arise from high-impact market events, sudden liquidity shocks, or structural breaks in trading activity. In some cases, the detected extremes may also be linked to data-related issues, such as recording errors, price misprints or unpredicted corporate actions.

From a statistical viewpoint, the detected observations represent true tail events rather than local density anomalies. Their extremeness is mainly due to very small empirical tail probabilities in at least one marginal distribution, most notably the log-return component. The skewness-corrected COPOD algorithm correctly assigns large scores to such observations by aggregating marginal tail contributions through the empirical copula setting. This confirms that the method is particularly suitable for the identification of observations that lie in low-probability regions of the joint distribution, even when marginal distributions are strongly asymmetric.

Furthermore, the decomposition of the COPOD score into the left-tail, right-tail and skewness-corrected components allows for a clear interpretation of the factors that influence extremeness. In the analyzed dataset, extreme values of *the log-returns* variable have a dominant impact on the observed anomalous behavior, whereas the extreme *price* levels and the temporal concentration of *the time_seconds* quantities strengthen this impact. Such interpretability is especially desirable and valuable in financial applications, where distinguishing between economically meaningful extreme events and data recording errors or missing data is crucial.

SUMMARY AND FUTURE WORK

In our paper, we investigated the applicability of the skewness-corrected COPOD (Copula-Based Outlier Detection) algorithm to high-frequency financial data from the Warsaw Stock Exchange. Building on the theoretical framework of empirical copulas, we restated the core principles of COPOD in a concise and transparent form and provided the Python implementation used for the financial data characterized by skewness and heavy tails.

The empirical analysis, conducted on a pre-processed high-frequency dataset containing prices, intraday time and log-returns for a single financial instrument, shows that the skewness-corrected COPOD algorithm effectively identifies extreme observations corresponding to rare and economically meaningful events. The method consistently identifies approximately five percent of the observations as outliers when a quantile-based threshold is applied, while clearly separating the most extreme tail events from the majority of the data. The results confirm that log-returns constitute the primary factor of extremeness, with price levels and temporal effects providing additional impact on increase of the extremeness power.

A key advantage of the proposed approach lies in its nonparametric, deterministic, and parameter-free nature, which allows to eliminate the need for hyperparameter tuning or stochastic training. Moreover, the additive log-tail decomposition of the COPOD score secures a high degree of interpretability, which enables to investigate anomalous behavior for specific variables and distributional tails. These properties make the skewness-corrected COPOD particularly attractive for exploratory data analysis, data cleaning and risk controlling in high-frequency financial data.

Despite its strengths, the proposed approach also has some limitations. Being rank-based and static, the current formulation does not explicitly account for temporal dependence, intraday seasonality or market volatility. In addition, when applied independently to individual instruments, the method may fail to detect anomalies that arise from joint behavior across multiple assets.

Future research may therefore proceed in several directions. First, the COPOD framework could be extended to incorporate temporal structure, for example through hybrid models that combine empirical copulas with time-series features. Moreover, multivariate extensions involving multiple instruments could be explored in order to detect systemic anomalies and tail dependence. Furthermore, comparison of the skewness-corrected COPOD with alternative state-of-the-art anomaly identification methods, as well as combining it with conformal prediction techniques, could further enhance its practical relevance and robustness in large-scale financial applications. Additionally, it would be interesting in future work to apply the algorithm by incorporating parametric families of copula for dependent structures (Gaussian, Student-t copulas, etc.) instead of using an approximation by the product copula. We could also consider modeling stochastic dependencies of individual transactions by applying vine copula model to multidimensional time series (see [Nagler et al. 2022]).

REFERENCES

Aggarwal C. C. (2017) Outlier Analysis. 2nd edn. Springer, Cham. https://doi.org/10.1007/978-3-319-47578-3

Aït-Sahalia Y., Jacod J. (2014) High-Frequency Financial Econometrics. Princeton University Press, Princeton. https://doi.org/10.1515/9781400850325

Breunig M. M., Kriegel H.-P., Ng R. T, Sander, J. (2000) LOF: Identifying Density-Based Local Outliers. ACM SIGMOD Record, 29(2), 93-104.

Chandola V., Banerjee A., Kumar V. (2009) Anomaly Detection: A Survey. ACM Computing Surveys, 41(3), 15. https://doi.org/10.1145/1541880.1541882

Cont R. (2001) Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. Quant Finance 1(2), 223-236. https://doi.org/10.1080/713665670

Deheuvels P. (1979) La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. Bulletins de l'Académie Royale de Belgique, 65, 274-292.

Embrechts P., Klüppelberg C., Mikosch T. (1997) Modelling Extremal Events for Insurance and Finance. Springer-Verlag.

Engle R. F. (2000) The Econometrics of Ultra-High-Frequency Data. Econometrica, 68(1),1-22. https://doi.org/10.1111/1468-0262.00091

Hariri S., Carrasco Kind M., Brunner R. J. (2021) Extended Isolation Forest. IEEE Transactions on Knowledge and Data Engineering, 33(4), 1479-1491. https://doi.org/10.1109/TKDE.2019.2947676

Li Z., Zhao Y., Botta N., Ionescu C., Hu X. (2020) COPOD: Copula-Based Outlier Detection. [In:] Proceedings of the IEEE International Conference on Data Mining (ICDM), 1118-1123. https://doi.org/10.1109/ICDM50108.2020.00138

Liu F. T., Ting K. M., Zhou Z.-H. (2008) Isolation Forest. IEEE ICDM, 413-422.

Nagler T., Krüger D., Min A. (2022) Stationary Vine Copula Models for Multivariate Time Series. Journal of Econometrics, 227(4), 305-324.

Nelsen R. B. (2006) An Introduction to Copulas. 2nd edn. Springer, New York https://doi.org/10.1007/0-387-28678-0

O'Hara M. (1998) Market Microstructure Theory. Wiley & Sons. https://asset.quant-wiki.com/pdf/Maureen%20O%27Hara%20-%20Market%20Microstructure%20Theory%20%20-Wiley%20%281998%29.pdf?ref=fxopen.com

Schölkopf B., Platt J. C., Shawe-Taylor J., Smola A. J., Williamson R. C. (2001) Estimating the Support of a High-Dimensional Distribution. Neural Computation, 13(7), 1443-1471.

Schweizer B., Sklar A. (1983) Probabilistic Metric Spaces. North-Holland, New York.

Sklar A. (1959) Fonctions de répartition à N dimensions et leurs marges. Annales de l'ISUP, VIII (3), 229-231.