

RELATIVE MEASURABILITY AND ITS APPLICATIONS

JACEK LEŚKOW

ABSTRACT. The goal of this paper is to introduce the concept of the relative measurability as a tool in building models for the analysis of time series and signals. Relative measurability allows to avoid using unverifiable assumptions such as mixing. However, with the concept of relative measurability one can get convergence of estimators and even define resampling procedures as evidenced in the research of the Author. The novelty of this article are the results corresponding to incorporating the resampling schemes into relative measurability.

1. INTRODUCTION

In recent years, there is a growing interest in providing mathematical and statistical models for large data sets. Especially in more complex data structures such as time series or signals one is left with the classical stochastic models that do not provide adequate tools to deal with the complexity of the data nor give solutions that are compatible with the needs of the data analysts. The Author of this article together with his collaborators (see e.g. [19] or [20]) since mid 2000's are proposing an alternative approach to analysis of time series and signals using the relative measurability and fraction of time approach.

The fraction of time approach was initiated as a tool of analysis of cyclostationary signals by Gardner in [11]. This approach allowed to introduce well known concepts such as expectations, moments, estimators without the burden of introducing the stochastic approach. One also has to bear in mind, that many times in the reality of the observations of time series there is only one realization available. For example, there is absolutely no chance of having second opportunity to observe the time series corresponding to the daily DIJA stock market index in any given year. The stochastic approach has in-built assumption that repeating observations one gets better precision of the estimates. For the stochastic approach for time series the absence of repetitions can be mitigated with the ergodicity assumptions, when time averages can approximate the ensemble averages. However, this is true only when the time series model has ergodic property. Verifying the property of ergodicity for real observations is next to impossible when the underlying model is not gaussian. Therefore, the approach using the fraction-of-time and relative measurability is a viable alternative.

Publication co-financed by the European Union as part of the European Social Fund within the project Center for Applications of Mathematics

J. LEŚKOW

Our paper will be organized as follows. The Section 2 contains preliminaries corresponding to the fraction of time approach, relative measurability and fundamental results from the previous research of the Author and his collaborators. Section 3 presents original approach to introduce the resampling techniques to the fraction of time context. Section 4 presents discussion on potential applications and open questions. In all considerations we assume that the considered time series or signal x(t) is a real Lebesgue measurable function.

2. PRELIMINARIES

2.0.1. Relative measurability of a function.

Definition 2.1. A function x(t) is called *relatively measurable* (RM) if and only if the set $\{t \in \mathbb{R} : x(t) \le \xi\}$ is relatively measurable for every $\xi \in \mathbb{R} - \Xi_0$, where Ξ_0 is at most a countable set of points.

For each RM function x(t) we can define a distribution function

(2.1)

$$F_{x}(\xi) \stackrel{def}{=} \mu_{R}(\{t \in \mathbb{R} : x(t) \leq \xi\})$$

$$\stackrel{def}{=} \lim_{T \to \infty} \frac{1}{T} \mu(\{t \in [-T/2, T/2] : x(t) \leq \xi\})$$

$$= \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \mathbf{1}_{(-\infty,\xi]}(x(t)) dt$$

in all points ξ where the limit exists.

In the above considerations, μ_R corresponds to the relative measure on the real line induced by the RM function $x(\cdot)$ while μ is the Lebesgue measure. For details see [19].

It is relatively easy to see (check [19]) that $F_x(\xi)$ defined above is a cdf, with one exception. It is not right-continuous, since the relative measure μ_R is not σ -additive. However, with the help of $F_x(\xi)$ one can define moments, cumulants, characteristic functions - that is all necessary elements of statistical inference for the considered signal.

We will provide now some examples of RM functions that are very important in applications.

2.0.2. Almost periodic functions. Let us recall that a function x(t) is said to be almost periodic in the sense of Bohr (B-AP) if for every $\epsilon > 0$ there exists l_{ϵ} such that for any interval I of the length l_{ϵ} there exists $\tau_{\epsilon} \in I$ such that

$$\sup_{t\in\mathbb{R}}|x(t+\tau_{\epsilon})-x(t)|<\epsilon.$$

More details on B-AP functions can be found in the classical work of Besicovitch [3].

It is very well known that any B-AP function can be uniformly approximated by trigonometric polynomial. From the research of the Author it is also clear (see [19]) that every B-AP function is RM.

Almost periodic functions are extremely important in applications, especially as models of phenomena for signals that are approximately periodic. In the engineering literature such signals are frequently called *cyclostationary*. In the recent survey of the literature (see [13]) there are at least 2500 papers published on the topic. Among those, the most prominent applications are in mechanical diagnostics (see e.g. [1]), vibroacoustics [17] and telecommunication signals processing [20]. 2.0.3. Asymptotically B-AP functions. (AAP functions). The next, more general class of functions are asymptotically almost periodic functions. The rigorous definition is as follows.

The function x(t) is AAP if it can be represented via

$$x(t) \stackrel{def}{=} x_{AP}(t) + \eta(t),$$

where $\eta(t) \in L^1_{loc}(\mathbb{R})$ and $\lim_{|t|\to\infty} \eta(t) = 0$.

In other words, asymptotically almost periodic function is a sum of an almost periodic function $x_{AP}(t)$ and a locally integrable function $\eta(t)$ that vanishes in both infinities. The need for such functions comes again from signal processing. The signal processing interpretation of the definition of AAP function is quite intuitive. The signal x(t) is AAP when it can be decomposed into a *persistent part* $x_{AP}(t)$ (note that AP functions do not vanish in infinities!) and into a *transient* part $\eta(t)$ that, for large observation interval does not contribute much to expectation, variance etc. More details on that can be found in [20].

2.0.4. *Pseudorandom functions*. We will now consider a function x(t) defined as

. .

$$x(t) \stackrel{def}{=} \cos(2\pi P([t])),$$

where P(t) is the l^{th} order polynomial with $l \ge 2$ and at least one of its coefficient is irrational. Such a function, in the signal processing context, is called a *pseudorandom function* (check [2]). It is also interesting to see that this function is not AP and is not AAP however it is RM. Another example in this class of functions is a function y(t) defined as

$$y(t) \stackrel{def}{=} \cos(\pi[P([t])]),$$

where as before $P(\cdot)$ is the l^{th} order polynomial with $l \ge 2$ and at least one of its coefficient is irrational. For more discussion on properties and applications of pseudorandom functions we refer the reader to [19].

Our goal right now is to establish the Central Limit Theorem for the sequence of independent RM functions. In order to do that, we will need concepts like: *joint relative measurability* of the sequence of RM functions, or *independence* of RM functions. We will also provide examples of such sequences of RM functions.

2.1. Joint relative measurability. Let us now assume that we have a sequence of relatively measurable functions denoted as $\varphi_1(t), \ldots, \varphi_n(t)$. We have the following

Definition 2.2. The Lebesgue measurable functions $\varphi_1(t), \ldots, \varphi_n(t), t \in R$, are said to be *jointly relatively measurable* if the limit

$$F_{\varphi_1 \cdots \varphi_n}(\xi_1, \dots, \xi_n) \stackrel{def}{=} \\ \mu_R(\{t \in \mathbb{R} \colon \varphi_1(t) \le \xi_1\} \cap \dots \cap \{t \in \mathbb{R} \colon \varphi_n(t) \le \xi_n\}) = \\ \lim_{T \to \infty} \frac{1}{T} \mu(\{t \in [-T/2, T/2] \colon \varphi_1(t) \le \xi_1, \dots, \varphi_n(t) \le \xi_n\})$$

exists for all $(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n - \Xi_0$, where Ξ_0 is at most a countable set of (n-1)-dimensional manifolds of \mathbb{R}^n .

J. LEŚKOW

This definition is required to consider a joint cdf generated by the sequence $\varphi_1(t), \ldots, \varphi_n(t)$. However, as a warning, we would like to stipulate here that there are RM functions x(t) such that even for some $\tau \in \mathbb{R}$ the original x(t) and its shifted version $x(t + \tau)$ are not jointly RM. So one needs to carefully check the sufficient conditions for the existence of the limit in the above definition. Some details on that are included in [19] and are also topic of the ongoing research of the Author and his research group.

We need to move on with our task of establishing the Central Limit Theorem for the sequence of RM functions. For that we need a concept of independence of RM functions.

2.2. Independence of RM functions. Again, as before we consider a sequence $\varphi_1(t), \ldots, \varphi_n(t)$ of relatively measurable functions. We have the following

Definition 2.3. The functions $\varphi_1(t), \ldots, \varphi_n(t)$ are said to be independent approach if and only if the sets

$$A_1 \stackrel{def}{=} \{ t \in \mathbb{R} \colon \varphi_1(t) \le \xi_1 \}, \dots, A_n \stackrel{def}{=} \{ t \in \mathbb{R} \colon \varphi_n(t) \le \xi_n \}$$

are independent $\forall (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n - \Xi_0$, where Ξ_0 is at most a countable set of (n-1)-dimensional manifolds of \mathbb{R}^n .

It is interesting to study what type of conditions enforce the independence of RM functions. The most prominent example are two sinusoids $x(t) = \sin(\pi \lambda_1 t)$ and $y(t) = \sin(\pi \lambda_2 t)$ such that the frequencies λ_1 and λ_2 are not relatively rational (in the signal processing language: *incommensurate frequencies*). Then x and y are clearly RM (since they are AP) and are also independent and jointly RM.

2.3. Central Limit Theorem of RM functions. Let us consider a sequence of real-valued functions $\{\varphi_k(t)\}_{k \in \mathbb{N}}$ such that the following assumptions are fulfilled.

(A1) For every k the function $\varphi_k(t)$ is zero mean that is

$$\lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \varphi_k(t) dt = 0.$$

(A2) The functions $\varphi_k(t)$ have temporal mean-square values

$$\lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} \varphi_k^2(t) dt = \sigma_k^2 < \infty$$

such that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n \sigma_k^2 = \sigma^2 < \infty .$$

For the Central Limit Theorem we will also need the following (A3) The functions $\varphi_k(t)$ are such that

$$a_k \stackrel{def}{=} \limsup_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} |\varphi_k(t)|^3 dt < \infty$$

with

$$\sum_{k=1}^{n} a_k = o(n^{3/2}) \quad \text{as } n \to \infty.$$

Let us now define the scaled average of the sequence $\varphi_1(t), \ldots, \varphi_n(t)$:

$$x_n(t) \stackrel{def}{=} \frac{1}{\sqrt{n}} \sum_{k=1}^n \varphi_k(t).$$

Obviously, the square root in scaling in the above formula is related to the speed of convergence of variances as stipulated in the assumption A2.

With the assumptions A1, A2, A3 we are now ready to formulate the Central Limit Theorem. **Theorem 2.4.** Let $\{\varphi_k(t)\}_{k \in \mathbb{N}}$ be a sequence of jointly RM independent functions satisfying assumptions A1, A2, and A3. We have

$$\lim_{n \to \infty} \mu_R \left(\{ t \in \mathbb{R} \colon a < x_n(t) \le b \} \right) = \mu_R \left(\{ t \in \mathbb{R} \colon a < x(t) \le b \} \right) =$$
$$= (1/2\pi\sigma^2)^{(1/2)} \int_a^b e^{-\frac{\xi^2}{2\sigma^2}} d\xi .$$

Moreover, x(t) *has a Gaussian fraction-of-time (FOT) distribution function.*

The Proof of the CLT follows from the results [19] and will not be presented here.

Remark. Note that the above CLT was obtained with only qualitative assumptions made for the underlying sequence of RM functions $\{\varphi_k(t)\}_{k\in N}$. In practice, it is relatively straightforward to check the independence of the system of functions as this can be translated to checking the relationships between the frequencies of the underlying Fourier series, if for example, the signals are B-AP. On the other hand, in the stochastic approach, in order to get a CLT for the time series or stochastic processes, one needs to introduce the mixing assumptions. To make our argument clearer, we recall now the concept of α -mixing, this concept being quite popular in recent research for nonstationary signals and stochastic processes (see e.g. [8]). The time series $\{X(t) : t \in \mathbb{Z}\}$ is called α -mixing if $\alpha_X(t) \to 0$ for $t \to \infty$, where

$$\alpha_X(t) = \sup |P(A \cap B) - P(A)P(B)|.$$

Here the supremum is taken over all $s \in \mathbb{Z}$ and all sets $A \in \mathcal{F}_X(-\infty, s)$ and $B \in \mathcal{F}_X(s+t, \infty)$, where $\mathcal{F}_X(t_1, t_2)$ corresponds to the σ -algebra generated by $\{X(t) : t_1 \leq t \leq t_2\}$.

A quick look into the above mixing condition proves the point. In order to check the validity of mixing condition in the stochastic context, one has to work with the (possibly nonstationary) distributions of the time series on hand. In reality, this can be only verified when time series is gaussian. On the other hand, the fraction-of-time (FOT) approach via relative measurability gives the CLT without any assumptions related to the stochastic distributions. It is enough to check the relationships between the frequencies of Fourier series approximating our signals.

The meaning of the result above is quite important. Provided that the technical assumptions (like e.g A1, A2 and A3) are true then the fraction of time probability related to the scaled average of the sequence of independent RM functions is becoming gaussian. Note, however, one inconvenience. In order to use this theorem in practice, one would need to estimate the asymptotic variance σ^2 . This is far from obvious but, on the other hand, is necessary to get confidence bounds on the estimates related with the scaled average. In order to solve this problem, the next section of this article is dedicated to possibilities of introducing bootstrap resampling techniques in the context of RM functions and FOT approach.

J. LEŚKOW

3. BOOTSTRAPPING THE RELATIVELY MEASURABLE FUNCTIONS

For the convenience of the readers, let us recall here the most fundamental ideas of bootstrap in the classical stochastic context. Let X_1, \ldots, X_n be an iid sequence from the distribution F. Let θ be the parameter describing the cdf F. In this context, statistical inference has three fundamental goals:

- building optimal estimators $\hat{\theta}$ of θ
- constructing the confidence intervals for θ
- testing procedures for the unknown parameter θ

Let us take for example the case of inference about the expected value $\mu = E_X F$. Maximum likelihood estimate of this parameter is the sample mean

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

We measure the error and variability of this estimate by building a confidence interval for μ . Such interval is constructed using the central limit theorem and has the form:

$$\hat{\mu}_n \pm z_{\alpha/2} \cdot SE(\hat{\mu}_n)$$

Recall that even in this simple case sometimes it is not so evident whether for small n the central limit theorem can be used. It is already well known, however, that the sampling distribution of the maximum likelihood estimate can be approximated via bootstrapping. In what follows, we give fundamentals of the bootstrap procedure.

Nonparametric bootstrap for iid sequence.

Let X_1, \ldots, X_n - iid from the distribution F. Our goal is to build a confidence interval around the mean μ . Instead of using the formula $\hat{\mu}_n \pm z_{\alpha/2} \cdot SE(\hat{\mu}_n)$ we can use nonparametric bootstrap.

- Step 1: Draw with replacement X₁^{*1},..., X_n^{*1} from X₁,..., X_n.
 Step 2: Calculate μ_n^{*1} = ¹/_n Σ_{j=1}ⁿ X_j^{*1}.
- Step 3: Repeat Step 1 and 2 B times to get $\mu_n^{*1}, \ldots, \mu_n^{*B}$.
- Step 4: Calculate the confidence interval from Step 3.

The nonparametric bootstrap defined above is *consistent*, that is the quantiles that can be derived from the procedure above are asymptotically the same as the quantiles of the normal distribution that corresponds to the sample mean for large n. However, for small n the bootstrap approximation provides better results than the normal law. For details see e.g. [8].

Now the goal of our work is to introduce the concept of bootstrap for the weak convergence of the RM functions. As before, assume that $\{\varphi_k(t)\}_{k\in N}$ be a sequence of jointly RM independent functions satisfying assumptions A1, A2, and A3. Assume also that $x_n(t)$ is the scaled average that is $x_n(t) \stackrel{def}{=} \frac{1}{\sqrt{n}} \sum_{k=1}^n \varphi_k(t)$. Our line of arguing will be now as follows. We will consider a bootstrap procedure on the

sequence $\varphi_1(t), \ldots, \varphi_n(t)$. We will then argue that the quantiles derived by bootstrap will be asymptotically the same as quantiles from the gaussian distribution obtained in the Central Limit Theorem.

Let us formulate the bootstrap algoritheorem for the sequence of independent RM functions. Recall that our objective is to replace the asymptotic gaussian relative measure with the density given by Theorem by an empirical approximation calculated on the basis of the sequence $\{\varphi_k(t)\}_{k\in\mathbb{N}}$. In order to argue that the bootstrap procedure is consistent, we will use the *m* out of *n* bootstrap presented in [22].

Bootstrapping the independent RM functions

- Step 1: Draw without replacement first bootstrap sample of the size m, that is get functions $\varphi_1^{*1}(t), \dots, \varphi_m^{*1}(t)$ from the original system $\varphi_1(t), \dots, \varphi_n(t)$. • Step 2: Calculate $x_m^{*1}(t) = \frac{1}{\sqrt{m}} \sum_{j=1}^m \varphi_j^{*1}(t)$. • Step 3: Repeat Step 1 and 2 *B* times to get the replicates $x_m^{*1}(t), \dots, x_n^{*B}(t)$.

- Step 4: Calculate confidence interval from the empirical distribution obtained in Step 3.

If one looks at the algoritheorem above, the fundamental question is the choice of the block size m. The minimal conditions presented in [22] stipulate that the consistency of this type of bootstrap procedure holds when $\frac{m}{n} \to 0$. A more detailed discussion on the choice of the block size m is the topic of the current research of the Author.

4. APPLICATIONS

As it was mentioned earlier, almost periodic functions (B-AP), asymptotically almost periodic functions (AAP), pseudorandom functions are indispensable in modern statistical inference for mechanical signals, telecommunication signals or vibroacoustic signals (see [1], [21] and [17], respectively). In order to have a clearer perspective, let us analyze x(t) - the acceleration signals of the car engine (for details see [1]). After routine filtering, such signal can be considered mean zero and cyclostationary. Each reader of this article can easily generate (and listen) to such signal by entering her or his car, starting the motor and letting it idle to listen for acceleration. One does not have to be an expert in mechanical signal processing to know that if motor works properly then we will be most likely listening to a rather smooth uniform signal of an idling motor. However, with an old clunker, it is likely that the idling signal will have occasional hick-ups or bumps. How this translates into relative measurability and central limit theorem and bootstrap? The signal generated by properly working engine, after routine filtering, will have mean zero. Now we can associate the sequence $\{\varphi_k(t)\}_{k\in\mathbb{N}}$ with several runs of the idling engine. It will be rather straightforward to check that assumptions A1, A2 and A3 hold true. If the mean of the signal is zero, then the relative measure built on $\{\varphi_k(t)\}_{k\in N}$ should converge to a gaussian mean-zero relative measure with the variance σ^2 . However, in order to run a formal test of hypotheses for the mean zero, one will have to estimate the unknown σ^2 . This can be quite problematic, given the complexity of the cyclostationary signals $\{\varphi_k(t)\}_{k\in\mathbb{N}}$ on hand. Instead, one should apply the algorithm of bootstrapping the independent RM functions to get the confidence interval from Step 4. If zero will be inside such a confidence interval then we can say that the engine works normally while idling. If not, then we better have our car inspected.

Acknowledgements. This research was supported by the grant UMO-2013/10/M/ST1/00096. The helpful discussions with my long-time research partners: Dominique Dehay and Antonio Napolitano are also acknowledged here.

REFERENCES

[1] J. Antoni, *Cyclostationarity by examples*, Mechanical Systems and Signal Processing, 2009.

- [2] J. Bass, Cours de Mathématiques, Tome III, Masson & Cie, Paris, 1971.
- [3] A.S. Besicovitch, Almost Periodic Functions, London: Cambridge University Press, 1932 and New York: Dover Publications, Inc., 1954.
- [4] P. Billingsley, Convergence of Probability Measures. 2nd Edn., John Wiley & Sons Inc., NJ, 1999.
- [5] P. Billingsley, *Probability and Measure*. Anniversary Edition, John Wiley & Sons Inc., NJ, 2012.
- [6] D. Brillinger, Times Series Data Analysis and Theory. Holden Day Inc., San Francisco, 1981.
- [7] D. Dehay, J. Leśkow, A. Napolitano, *Central limit theorem in the functional approach*, IEEE Trans. Signal Processing 61(16): 4025–4037, August 2013.
- [8] D. Dehay, A. Dudek, J. Leśkow, Subsampling for continuous time almost periodically correlated processes, Journal. Stat. Plan.Inference 150: 142–158, 2014.
- [9] R.G. Gallager, Principles of Digital Communications, Cambridge University Press, Cambridge, UK, 2008.
- [10] W.A. Gardner, Statistical Spectral Analysis: A Nonprobabilistic Theory. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [11] W.A. Gardner, W.A. Brown, *Fraction-of-time probability for time-series that exhibit cyclostationarity*, Signal Processing 23: 273–292, June 1991.
- [12] W.A. Gardner, C.M. Spooner, The cumulant theory of cyclostationary time-series, Part I: Foundation, IEEE Trans. Signal Processing 42: 3387–3408, December 1994.
- [13] W.A. Gardner, A. Napolitano, L. Paura, Cyclostationarity: Half a century of research, Signal Processing 86(4): 639–697, April 2006.
- [14] M. Kac, Statistical Independence in Probability, Analysis and Number Theory, The Mathematical Association of America, USA, 1959.
- [15] M. Kac, H. Steinhaus, Sur les fonctions indépendantes (II), Studia Mathematica 6: 59-66, 1936.
- [16] M. Kac, H. Steinhaus, Sur les foncions indépendantes IV, Studia Mathematica 7: 1–15, 1938.
- [17] J. Leśkow, Cyclostationarity and resampling for vibroacoustic signals, Acta Physica Polonica, Ser A, 2011.
- [18] J. Leśkow, A. Napolitano, 'Fraction-of-time approach in predicting Value-at Risk', with A. Napolitano, in: Lecture Notes in Mathematical Economics, editors: Leśkow, Puchet and Punzo, 183–200, Springer Verlag 2004.
- [19] J. Leśkow, A. Napolitano, *Foundations of the functional approach for signal analysis*, Signal Processing 86(12): 3796–3825, December 2006.
- [20] J. Leśkow, A. Napolitano, Non-relatively measurable functions for secure-communications signal design, Signal Processing 87(11): 2765–2780, November 2007.
- [21] A. Napolitano, Generalizations of Cyclostationary Signal Processing Spectral Analisys and Applications, John Wiley & Sons Inc. and IEEE Press, 2012.
- [22] D. Politis, J. Romano A general theory for large sample confidence regions based on subsamples under minimal assumptions, Ann. Statist. 22: 2031-2050, 1994.
- [23] D. Politis, J. Romano, M. Wolf, Subsampling, Springer Series in Statistics, 1999.
- [24] J.G. Proakis, Digital Communications, 3rd Edn., McGraw-Hill, New York, 1995.
- [25] B.G. Quinn, E.J. Hannan, *The Estimation and Tracking of Frequency*. Cambridge University Press, Cambridge, UK, 2001.
- [26] R. Salem, A. Zygumund, On lacunary trigonometric series, Proc. Nat. Acad. Sci. USA 34: 54–62, 1948.
- [27] R.J. Serfling, Approximation Theorems of Mathematical Statistics, Wiley, 1980.
- [28] C.M. Spooner and W.A. Gardner, The cumulant theory of cyclostationary time-series, Part II: Development and applications, IEEE Trans. Signal Processing 42: 3409–3429, December 1994.
- [29] A.W. van der Vaart, Asymptotic Statistics, Cambdridge University Press, Cambridge (UK), 1998.
- [30] H.L. van Trees, Detection, Estimation, and Modulation Theory, Part III. John Wiley & Sons, Inc., New York, 1971.
- [31] A. Wintner, On the asymptotic differential distribution of almost-periodic and related functions, American Journal of Mathematics 56(1): 401–406, 1934.

JACEK LEŚKOW

INSTITUTE OF MATHEMATICS, CRACOW TECHNICAL UNIVERSITY, ULICA WARSZAWSKA 24, 30-151 KRAKÓW *E-mail address*: jleskow@pk.edu.pl