Confidence Bound Calculation and Prediction in the Fraction-of-Time Probability Framework

Jacek Leśkow
CIMAT, Guanajuato

and

Antonio Napolitano
Dipartimento di Ingegenria Elettronica e delle Telecomunicazioni
Università di Napoli Federico II, Napoli, Italy

Abstract

The aim of this work is to introduce the concept of the confidence bound via the fraction-of-time probability approach. This approach allows to use information coming from only one realization of the phenomena under study while the observation time is increasing. The construction of the empirical and a large-sample confidence interval is presented and its implications for quantile prediction is studied.

1 Introduction

An alternative signal analysis framework that does not resort to random processes is the fraction-of-time (FOT) probability approach [4]. In such an approach, statistical parameters are defined through infinite-time averages of signals (i.e., single functions of time) rather than ensemble averages of random processes.

The adoption of the FOT probability approach is motivated by the fact that in several applications, e.g. time series in finance, signal processing, and climatology, multiple realizations of a stochastic process are not at disposal of the experimenter. Rather, there is only one realization that can be assumed to be observed for an increasing-length time interval.

The differences between the FOT and stochastic approach have a variety of implications in the study of properties of inferential techniques involved. For example, in the FOT probability framework a natural way to define estimators is through considering finite-time averages of the same quantities involved in the infinite-time averages. Thus, the kind of convergence of the estimators to be considered as the time of the observation approaches infinity is the convergence of the sequence of the finite-time averages, that is, 'pointwise,' in the 'temporal mean-square sense' [15], or in the 'sense of generalized functions (distributions)' [11]. On the contrary, in the stochastic process framework, the convergence must be demonstrated in the mean-square sense, almost everywhere sense or in the weak convergence sense.

The FOT probability approach was first introduced in [15] with reference to time-invariant statistics of ordinary functions of time. Later, it was developed in [3] and [8], and then extended to the case of distributions (generalized functions) in [11] . Moreover, in [16] an isometric isomorfism (Wold isomorfism) between a stationary stochastic process and the Hilbert space generated by a single sample path was singled out and a rigorous link between the FOT probability and the stochastic process frameworks in the stationary case was established. The case of time-variant FOT statistics of almost-cyclostationary (ACS) time-series was widely treated in [4], [5] with reference to the second-order statistics and in [7], [13] for the higher-order statistics. Finally, the Wold isomorfism was extended to the case of cyclostationary time-series in [9]. A further development in the FOT probability theory for nonstationary signals was very recently presented in [10]. In that paper a more general class of nonstationary time-series called the generalized almost-cyclostationary (GACS) time-series has been introduced and characterized.

The apper is organized as follows. In Section 2, general concepts and definitions of the FOT probability framework are reviewed. Section 3 is devoted to construction of the confidence bounds using the FOT probability. The problem of predicting quantiles (without the independence and same distribution assumption) is presented in Section 4. Finally, simulation experiment to corroborate the proposed predictors are described in Section 5. Conclusions are drawn in Section 6.

2 General definitions

Let us consider a real-valued function $x(\cdot)$ that is Lebesgue measurable on the real axis $\mathbb{R}$.

Definition 2.1 The *empirical fraction-of-time* probability distribution function $F_T(t;\xi;x)$ of $x(u)$ observed on the time interval $[t, t+T]$ is defined as

$$\begin{aligned} F_T(t;\xi;x) &\triangleq \frac{\operatorname{meas}\{u \in [t,t+T] \ : \ x(u) \leq \xi\}}{\operatorname{meas}\{u \in [t,t+T]\}} \\ &= \frac{1}{T}\int_t^{t+T} \mathcal{U}(\xi - x(u))\,\mathrm{d}u\,, \end{aligned} \tag{1}$$

where meas$\{\cdot\}$ denotes the Lebesgue measure and

$$\mathcal{U}(t) \triangleq \begin{cases} 1 & t \geq 0, \\ 0 & \text{elsewhere.} \end{cases} \tag{2}$$

It represents the proportion of time where $x(u) \leq \xi$ while $u \in [t, t+T]$.

In this section we assume that the function $x(t)$ is fixed, therefore we will drop it from the notation whenever possible.

It can be easily shown that the function $F_T(t;\xi)$ defined in (1) is a valid cumulative distribution function. In fact, it assumes values in the set $[0,1]$, is a nondecreasing function of ξ and

$$\lim_{\xi \to -\infty} F_T(t;\xi) = 0\,, \tag{3a}$$

$$\lim_{\xi \to +\infty} F_T(t;\xi) = 1\,, \tag{3b}$$

$$\lim_{\xi \to \xi_0 +} F_T(t;\xi) = F_T(t;\xi_0)\,. \tag{3c}$$

From the cumulative distribution function (1) the corresponding probability density function can be defined as

$$\begin{aligned} f_T(t;\xi) &\triangleq \frac{\mathrm{d}}{\mathrm{d}\xi} F_T(t;\xi) \\ &= \frac{1}{T}\int_t^{t+T} \delta(\xi - x(u))\,\mathrm{d}u\,, \end{aligned} \tag{4}$$

where the derivation operation and the second equality must be intended in the sense of generalized functions (distributions).

The cumulative distribution function (1) and the probability density function (4) can be used to express statistical functions of the signal $x(t)$. For example, the expected value of the distribution F_T is given by

$$\begin{aligned} \mathrm{E}\{x(u),\ u \in [t,t+T]\} &\triangleq \int_{-\infty}^{+\infty} \xi f_T(t;\xi)\,\mathrm{d}\xi \\ &= \int_{-\infty}^{+\infty} \xi \frac{1}{T}\int_t^{t+T} \delta(\xi - x(u))\,\mathrm{d}u\ \mathrm{d}\xi \\ &= \frac{1}{T}\int_t^{t+T}\int_{-\infty}^{+\infty} \xi\delta(\xi - x(u))\,\mathrm{d}\xi\ \mathrm{d}u \\ &= \frac{1}{T}\int_t^{t+T} x(u)\,\mathrm{d}u \\ &\equiv \langle x(u)\rangle_{u\in[t,t+T]}\,, \end{aligned} \tag{5}$$

that is, it is coincident with the time average of the function $x(u)$ in the interval $[t,t+T]$. In the derivation of (5) it has been assumed that the order of limit and integration operation can be interchanged and the sampling property of the Dirac delta function has been exploited.

A natural question for the function $F_T(t;\xi)$ in Definition 2.1 is to study its asymptotic behaviour when $T \to \infty$.

Definition 2.2. Assume that $\lim_{T\to\infty} F_T(t;\xi)$ exists. The (limit) *fraction-of-time* probability distribution function $F(t;\xi)$ is defined as

$$F(t;\xi) \triangleq \lim_{T\to\infty} F_T(t;\xi)$$

$$
\begin{aligned}
&= \lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} \mathcal{U}(\xi - x(u))\,\mathrm{d}u \\
&= \lim_{T\to\infty} \frac{1}{T}\int_0^{T} \mathcal{U}(\xi - x(t+t'))\,\mathrm{d}t' \,. \qquad (6)
\end{aligned}
$$

The above function can be interpreted as the proportion of time in which the values of $x(t)$, when t ranges in $\mathbb{R}$, are less or equal to ξ.

The limit in Definition 2.2 exists for a large class of functions, for example for $x(\cdot)$ nonnegative and increasing to infinity when t does so. For a complete mathematical treatment of the functions for which the limit in (6) exists the reader is referred to the paper of Urbanik [14]. The two cases that will be treated in some cases in this paper are the almost periodic functions [1] and stepwise functions.

Fact 2.3. Assume that for each $t > 0$ the limit in (6) exists. Then it does not depend on t.

Proof. We have the following obvious equality:

$$
\begin{aligned}
& \frac{1}{T}\int_t^{t+T} \mathcal{U}(\xi - x(u))\,\mathrm{d}u - \frac{1}{T}\int_0^{T} \mathcal{U}(\xi - x(u))\,\mathrm{d}u \\
=\; & \frac{1}{T}\int_T^{t+T} \mathcal{U}(\xi - x(u))\,\mathrm{d}u - \frac{1}{T}\int_0^{t} \mathcal{U}(\xi - x(u))\,\mathrm{d}u. \qquad (7)
\end{aligned}
$$

Taking $T \to \infty$ we get the desired result. The same result can be found if $t < 0$. The Fact 2.3. allows us to drop t as the argument of $F(t;\xi)$.

Note that the limit function $F(\xi)$ is in turn a valid cumulative distribution function since the same properties of $F_T(t;\xi)$ hold. Thus, the (limit) probability density function can be defined as

$$
\begin{aligned}
f(\xi) &\triangleq \frac{\mathrm{d}}{\mathrm{d}\xi}F(\xi) \\
&= \lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} \delta(\xi - x(u))\,\mathrm{d}u \,. \qquad (8)
\end{aligned}
$$

In the following , we will turn our attention to the specific form of the FOT distribution for stepwise function.

Definition 2.4. The function $x_s(t)$ is called stepwise if it is of the form

$$
x_s(t) = \sum_{k=0}^{\infty} a_k \mathcal{I}(t - kT_s) \qquad (9)
$$

where a_k are finite constants and the function $\mathcal{I}(\cdot)$ is the indicator function of the interval $[0, T_s]$.

Fact 2.5. The empirical fraction of time distribution $F_T(t;\xi)$ of the stepwise function x_s has the following form:

$$
F_T(t;\xi) = \frac{T_s}{T}\sum_{k=r+1}^{p-1} \mathcal{U}(\xi - a_k) + R(t,T;\xi)\,, \qquad (10)
$$

where, for each $t > 0$, $p = [\frac{T+t}{T_s}]$, $r = [\frac{t}{T_s}]$, $[b]$ being the integer part of b and $R(t, T; \xi)$ tends to zero when $T \to \infty$.

Proof. We have that

$$
\begin{aligned}
F_T(t;\xi) &\stackrel{\triangle}{=} \frac{1}{T}\int_t^{T+t} \mathcal{U}(\xi - x_s(u))\,\mathrm{d}u \\
&= \frac{1}{T}\int_t^{(r+1)T_s} \mathcal{U}(\xi - x_s(u))\,\mathrm{d}u + \frac{1}{T}\int_{(r+1)T_s}^{pT_s} \mathcal{U}(\xi - x_s(u))\,\mathrm{d}u \\
&\quad + \frac{1}{T}\int_{pT_s}^{t+T} \mathcal{U}(\xi - x_s(u))\,\mathrm{d}u \\
&= \frac{T_s}{T}\sum_{k=r+1}^{p-1} \mathcal{U}(\xi - a_k) + R(t, T; \xi). \qquad (11)
\end{aligned}
$$

where

$$
\begin{aligned}
R(t, T; \xi) &\stackrel{\triangle}{=} \frac{1}{T}\int_t^{(r+1)T_s} \mathcal{U}(\xi - x_s(u))\,\mathrm{d}u + \frac{1}{T}\int_{pT_s}^{t+T} \mathcal{U}(\xi - x_s(u))\,\mathrm{d}u \\
&= \frac{(r+1)T_s - t}{T}\mathcal{U}(\xi - a_r) + \frac{t + T - pT_s}{T}\mathcal{U}(\xi - a_p). \qquad (12)
\end{aligned}
$$

Now, it easy to see that for each $t > 0$ and $\xi \in \mathbb{R}$ it results $|R(t, T; \xi)| \leq \frac{T_s}{T}$ and this proves the result.

Theorem 2.6. Assume that the coefficients a_k belong to the finite alphabet $\{A_1, \ldots, A_M\}$ with $A_1 < \cdots < A_M$ and the limit FOT probabilities $p_i \stackrel{\triangle}{=} \mathrm{Prob}\{a_k = A_i\}$ exist. Then the limit function F exists.

Proof. For $\xi < A_1$ the function F is obviously zero. For $A_1 \leq \xi < A_2$, accounting for (11), it results

$$
\begin{aligned}
F_T(t, \xi) &= \frac{T_s}{T}\sum_{k=r+1}^{p-1} \mathcal{U}(\xi - a_k) + R(t, T; \xi) \\
&= \frac{1}{p - r - 2}\{\text{number of } a_k = A_1\,, k \in \{r+1, \ldots, p-1\}\} + R(t, T; \xi)\,. \qquad (13)
\end{aligned}
$$

Thus, in the limit for $T \to \infty$ one has

$$
F(\xi) = p_1\,, \quad A_1 \leq \xi < A_2\,. \qquad (14)
$$

For $A_i \leq \xi < A_{i+1}$ it results

$$
\begin{aligned}
F_T(t, \xi) &= \frac{1}{p - r - 2}\{\text{number of } a_k = A_1\,, k \in \{r+1, \ldots, p-1\}\} \\
&\quad + \frac{1}{p - r - 2}\{\text{number of } a_k = A_2\,, k \in \{r+1, \ldots, p-1\}\} \\
&\quad + \cdots \\
&\quad + \frac{1}{p - r - 2}\{\text{number of } a_k = A_i\,, k \in \{r+1, \ldots, p-1\}\} + R(t, T; \xi) \qquad (15)
\end{aligned}
$$

and, hence, in the limit for $T \to \infty$ one has

$$F(\xi) = \sum_{j=1}^{i} p_j \,, \quad A_i \le \xi < A_{i+1} \,. \tag{16}$$

Remark. The FOT distributions F_T and F can be understood in the following sense. First, we have a Lebesgue measurable function $x(t)$. This enables us to construct a probability measure P_T corresponding to the observations of $\{x(t) : 0 \le t \le T\}$. The measure P_T is defined on $\mathbb{R}$ equipped with the σ-field generated by the sets $\{t \in \mathbb{R} : x(t) \le c;\ 0 \le t \le T,\ c \in \mathbb{R}\}$. The natural link between P_T and F_T is established via the equality

$$P_T\{x(t) \le c\} \stackrel{\triangle}{=} F_T(c),$$

where $F_T(c)$ can be calculated from the Definition 2.1. Identical simple considerations can be done to illicit the link between the FOT distribution F and the measure P obtained as the limit of P_T for $T \to \infty$, provided that such limits exist.

Assume now that we have at our disposal N real valued measurable functions $x_1, \ldots, x_N$ and let us fix them for the remainder of this paragraph.

Definition 2.6. The Nth-order *FOT joint cumulative distribution function* is defined as:

$$F(\xi_1, \ldots, \xi_N, \tau_1, \ldots, \tau_N) \stackrel{\triangle}{=} \lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} \prod_{i=1}^{N} \mathcal{U}(\xi_i - x_i(u + \tau_i))\, \mathrm{d}u \,. \tag{17}$$

The function above can be interpreted as the fraction of time where jointly $x_1(u + \tau_1) \le \xi_1$, $\cdots$, $x_N(u + \tau_N) \le \xi_N$ for u in $\mathbb{R}$ and $\tau_1, \ldots, \tau_N$ fixed. It is easy to see that the function defined in (17) is a distribution function since it assumes values in $[0, 1]$, is right-continuous, non-decreasing with respect to each variable ξ_i. Moreover, the function defined in (17) verifies the consistency conditions, that is, the joint cumulative distribution function of any subset of time-series $x_n(t + \tau_n)$ with $n \in \mathcal{M} \subset \{1, \ldots, N\}$ can be obtained from (17) in the limit when the ξ_n such that n is not in $\mathcal{M}$ is set to infinity.

From (17) it follows that the Nth-order *FOT joint probability density function* for $x_1, \ldots, x_N$ is given by

$$\begin{aligned} f(\xi_1, ..., \xi_N, \tau_1, \ldots, \tau_N) &\stackrel{\triangle}{=} \frac{\partial^N}{\partial\xi_1 \cdots \partial\xi_N} F(\xi_1, ..., \xi_N, \tau_1, \ldots, \tau_N) \\ &= \lim_{T\to\infty} \frac{1}{T} \int_0^T \prod_{i=1}^{N} \delta(\xi_n - x(t + \tau_i + t'))\, \mathrm{d}t' \,. \end{aligned} \tag{18}$$

The Nth-order joint probability density functions allows to evaluate the joint statistical functions of the functions $x_1(t + \tau_1), \cdots, x_N(t + \tau_N)$. In particular, the expected value of the product $x_1(t + \tau_1)x_2(t + \tau_2)$ is given by

$$\mathrm{E}\{x(t + \tau_1)x(t + \tau_2)\}$$

$$
\begin{aligned}
&\triangleq \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \xi_1\xi_2 f(\xi_1,\xi_2,\tau_1,\tau_2)\,\mathrm{d}\xi_1\ \mathrm{d}\xi_2 \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \xi_1\xi_2 \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T} \delta(\xi_1 - x_1(u+\tau_1))\delta(\xi_2 - x_2(u+\tau_2))\,\mathrm{d}u\ \mathrm{d}\xi_1\,\mathrm{d}\xi_2 \\
&= \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T}\int_{-\infty}^{+\infty} \xi_1\delta(\xi_1 - x_1(u+\tau_1))\,\mathrm{d}\xi_1 \int_{-\infty}^{+\infty}\xi_2\delta(\xi_2 - x_2(u+\tau_2))\,\mathrm{d}\xi_2\,\mathrm{d}u \\
&= \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T} x_1(u+\tau_1)x(u+\tau_2)\,\mathrm{d}u \\
&\equiv \left\langle x_1(t'+\tau)x_2(t')\right\rangle_{t'}\,, \qquad (19)
\end{aligned}
$$

which is conincident with the temporal cross-correlation function between $x_1(t)$ and $x_2(t)$. In the derivation of (19), it has been assumed that the order of limit and integration operations can be interchanged and in the last equality $t' = u + \tau_2$ and $\tau = \tau_1 - \tau_2$ have been set.

Equations (5) and (19) show the duality existing between the FOT probability approach and the classical stochastic (stationary) process approach. In the former, the statistical functions are defined through infinite-time averages of a single function of time (or a lag product of more functions of time), while in the latter the analogous statistical functions are defined as ensemble averages of a stationary stochastic process (or a lag product of more stationary stochastic processes). When the stochastic process is also ergodic, then the functions defined in both the approaches are coincident, provided that the function of time is a nonpathological sample path of the ergodic stationary stochastic process.

It is worthwhile to underline that the functions (6), (8), (17), and (18) do not depend on the variable t and, hence, are suitable to describe the statistics of signals for which a stationary model is assumed.

The almost-cyclostationary case

Let us consider N real-valued measurable functions $x_1(t),\ldots,x_N(t)$. If the set $\Gamma_{\boldsymbol{\tau},\boldsymbol{\xi}}$ of all the frequencies of the additive sinewave components contained in the function

$$
\mathcal{U}_{\boldsymbol{x}}(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi}) \triangleq \prod_{n=1}^{N}\mathcal{U}(\xi_n - x_n(t+\tau_n)) \qquad (20)
$$

is countable for each of the column vectors $\boldsymbol{\tau} \triangleq [\tau_1,...,\tau_N]^\intercal \in \mathbb{R}^N$ and $\boldsymbol{\xi} \triangleq [\xi_1,\cdots,\xi_N]^\intercal \in \mathbb{R}^N$, then the time-series are said to be jointly generalized almost-cyclostationary in the strict sense [10]. In (20), $\mathbf{1} \triangleq [1,\ldots,1]^\intercal$ and $\boldsymbol{x} \triangleq [x_1(t+\tau_1),\ldots,x_N(t+\tau_N)]^\intercal$. Moreover, if the union over all $\boldsymbol{\tau}\in\mathbb{R}^N$ and $\boldsymbol{\xi}\in\mathbb{R}^N$ of the sets $\Gamma_{\boldsymbol{\tau},\boldsymbol{\xi}}$ is a countable set Γ, then the above functions are jointly almost-cyclostationary in the strict sense. Furthermore, it is shown in [6] that the function

$$
\begin{aligned}
F(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi}) &\triangleq \mathrm{E}^{\{\alpha\}}\left\{\mathcal{U}_{\boldsymbol{x}}(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi})\right\} \\
&= \sum_{\gamma\in\Gamma}\left\langle \mathcal{U}_{\boldsymbol{x}}(\mathbf{1}v+\boldsymbol{\tau},\boldsymbol{\xi})\,e^{-j2\pi\gamma v}\right\rangle_v\, e^{j2\pi\gamma t} \qquad (21)
\end{aligned}
$$

is a valid cumulative distribution function for each fixed value of t and $\boldsymbol{\tau}$. In (21), $\langle\cdot\rangle_v$ denotes the bilateral time average operator with respect to v and $\mathrm{E}^{\{\alpha\}}\{\cdot\}$ is the almost-periodic component extraction operator, that is, the operator that extracts all the additive sinewave components in the function specified as its argument.

The Nth-order derivative (in the sense of generalized functions) of the cumulative distribution function is defined as

$$\begin{aligned} f(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi}) &\triangleq \frac{\partial^N}{\partial\xi_1\cdots\partial\xi_N}F(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi}) \\ &= \mathrm{E}^{\{\alpha\}}\left\{\prod_{n=1}^{N}\delta(\xi_n - x_n(t+\tau_n))\right\} \end{aligned} \tag{22}$$

and turns out to be a valid probability density function for each fixed value of t and $\boldsymbol{\tau}$.

The set Γ contains at least one element $\gamma = 0$. If $\gamma = 0$ is the only element of Γ, then the N functions are said to be jointly stationary in the strict sense. If all γ's different from zero are integer multiples of a value γ_0, then the N functions are said to be jointly cyclostationary in the strict sense with the period $1/\gamma_0$. If the set Γ contains incommensurate elements, then the N functions are said to be jointly almost-cyclostationary in the strict sense. Analogous definitions can be given with reference to a single signal (i.e., $x_1(t) \equiv \cdots \equiv x_N(t) \equiv x(t)$).

Let us consider the special case of a single signal $x(t)$. If the function $x(t)$ is almost-periodic then the function (20) is also almost periodic and hence it coincides with its almost-periodic component (21). Therefore, the probability density function (22) can be expressed as

$$f(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi}) = \prod_{n=1}^{N}\delta(\xi_n - x(t+\tau_n))\,, \tag{23}$$

that is, the almost-periodic functions are the deterministic time-series in the FOT probability framework. From this property it follows that the almost-periodic extraction operator $\mathrm{E}^{\{\alpha\}}\{\cdot\}$ extracts the deterministic component of its argument, that is, it plays the same role played by the expectation operation in the stochastic process approach. Thus, it can be shown that the almost-periodic function

$$\begin{aligned} \mathcal{R}(\mathbf{1}t+\boldsymbol{\tau}) &\triangleq \mathrm{E}^{\{\alpha\}}\left\{\prod_{n=1}^{N}x_n(t+\tau_n)\right\} \\ &= \sum_{\alpha\in A}\mathcal{R}^{\alpha}(\boldsymbol{\tau})e^{j2\pi\alpha t}\,, \end{aligned} \tag{24}$$

which is called the temporal moment function, is a valid moment function, that is, it can be expressed as (see [7])

$$\mathcal{R}(\mathbf{1}t+\boldsymbol{\tau}) = \int_{\mathbb{R}^N}\prod_{n=1}^{N}\xi_n f(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\,. \tag{25}$$

In (24), $A \subseteq \Gamma$ and the functions

$$\mathcal{R}^{\alpha}(\boldsymbol{\tau}) \triangleq \left\langle \prod_{n=1}^{N} x_n(t+\tau_n) e^{-j2\pi\alpha t} \right\rangle_t \tag{26}$$

are referred to as the cyclic temporal cross-moment functions (CTCMFs).

The set A of the α such that the CTCMF is different form zero can be possibly empty. If $\alpha = 0$ is the only element of A, then the N functions are said to be jointly stationary in the wide sense. If all α's different from zero are integer multiples of a number α_0, then the N functions are said to be jointly cyclostationary in the wide sense with period $1/\alpha_0$. If the set A contains incommensurate elements, then the functions are said to be jointly almost-cyclostationary in the wide sense. Analogous definitions can be given with reference to a single function.

It can be shown that if $g[t;\boldsymbol{\xi}]$ is an almost-periodically time-varying real-valued function of $\boldsymbol{\xi} \in \mathbb{R}^N$, it results that

$$\mathrm{E}^{\{\alpha\}}\left\{g[t; x_1(t+\tau_1), \ldots, x_N(t+\tau_N)]\right\} = \int_{\mathbb{R}^N} g[t;\boldsymbol{\xi}] f(\mathbf{1}t+\boldsymbol{\tau},\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}, \tag{27}$$

which is called the fundamental theorem of the almost-periodic component extraction [4], [5], [6].

Finally, it is worth emphasizing that two ore more functions are said to be statistically independent in the FOT probability sense if their joint probability density function factorizes as the product of the marginal probability density functions [4], [5]. Consequently, any almost-periodic function (wich is a deterministic signal in the FOT probability framework) is statistically independent on any other function including itself (see [6]).

3 Confidence bounds in the FOT probability approach

The problem of finding confidence bounds in the FOT probability approach does not differ from the analogous problem in the classical stochastic approach. In this section we assume the following definition of quantile:

Definition 3.1. Let F be an arbitrary distribution function. Then the α-quantile q_α of F is defined as :

$$q_\alpha = \inf\{s \in \mathbb{R} : F(s) \geq \alpha\}\,, \tag{28}$$

where $0 < \alpha < 1$. Similarly, we define the inverse F^{-1} of a given distribution function F as $F^{-1}(u) = \inf\{s \in \mathbb{R} : F(s) \geq u\}$ for $u \in (0,1)$.

For more details see e.g. [12], Chapter 1.

From the previous section it was clear that the calculation of the empirical FOT probability distribution depended on t, the starting point of the observation. To simplify the notation in this Section we will assume that t is fixed and we will drop it from the notation. The Definition 3.1 can be applied to both empirical FOT distribution and the FOT distribution. For the empirical FOT distribution F_T an α-quantile q_α^T is such a number that the proportion of time that $x(u) \leq q_\alpha^T$ while $u \in [t, t+T]$ is equal to or greater than α. Similar interpretation can be derived for the limiting FOT distribution F. The question of finding q_α^T corresponds in finance to finding a Value-at-Risk for a given time series of returns $x(\cdot)$ on a given asset, or, in signal theory, to finding an empirical threshold corresponding to a given false alarm rate α.

Definition 3.2. The FOT confidence interval $[L, U]$ at the confidence level α is defined as the interval between $L = \frac{\alpha}{2}$-quantile and $U = (1 - \frac{\alpha}{2})$-quantile of the FOT distribution F.

Similarly, we have

Definition 3.3. The emprical FOT confidence interval $[L_T, U_T]$ at the confidence level α is defined as the interval between $L_T = \frac{\alpha}{2}$-quantile and $U_T = (1 - \frac{\alpha}{2})$-quantile of the empirical FOT distribution F_T.

The calculation of the FOT quantile follows the same lines as in the 'classical case'. For more information, see e.g. [12].

The central result of this section is the following

Theorem 3.4. Assume that the significance level α for the calculation of the quantile q_α of the limit FOT distribution F is a continuity point of F^{-1}. Then the empirical quantile q_α^T is a consistent estimate of q_α.

Proof. We use the well-known results regarding the weak convergence of the distribution functions (see [12], p. 19). Since we know (via the construction of FOT distribution F) that F_T converges to F then from the quoted result it follows that $q_\alpha^T \to q_\alpha$.

To analyze a situation when F^{-1} may have some discontinuties, (e.g., in the desired significance level α) consider the following:

Example 3.5. Let $x_s(t)$ be a stepwise function in the sense of the Definition 2.4. Assume also, that for each k, $a_k \in \{A_1, \ldots, A_M\}$ with $A_1 < \cdots < A_M$, that is coefficients a_k come from a finite set. Denote now $p_i = F(A_i) - F(A_{i-1})$, where F is the FOT distribution of $x_s(\cdot)$. Obviously, F is stepwise in this case and so is F^{-1}. By convention, assume also that $F(A_0) = 0$ and that neither of p_i is zero. The calculation of the empirical quantile can be accomplished using the following

Quantile estimation algorithm.

Step 1. For each $1 \leq i \leq M$, accounting for (11), calculate the empirical probabilities $p_i^T \triangleq F_T(A_i) - F_T(A_{i-1})$ directly using the observations of the function $x_s(\cdot)$ in the interval $[t, t+T]$, since $x_s(t) = a_k$ for $t \in [kT_s, (k+1)T_s]$.

Step 2. Find the integer k such that

$$\sum_{i=1}^{k} p_i^T < \alpha \leq \sum_{i=1}^{k+1} p_i^T.$$

Step 3. Put the empirical α-quantile $q_\alpha^T = A_{k+1}$.

It should not be overlooked that in the algortihm above the probabilities p_i^T as well as F_T depend on the starting point t.

We have the following

Corollary 3.6. The empirical quantile q_α^T calculated via the above algorithm is consistent, i.e. $q_\alpha^T \to q_\alpha$.

Proof. If the selected significance level α is a continuity point of F^{-1} then we can simply apply Theorem 3.4. If this is not the case, then q_α is a point where F has a discontinuity as well. In our case (stepwise function $x_s(t)$ as in Example 3.5) this means that there exists such k that

$$\alpha = \sum_{i=1}^{k} p_i.$$

On the other hand we know that for each i and each fixed t, $p_i^T \to p_i$ therefore we get the desired result.

Remark 3.7. It is easy to see that the quantile estimation algorithm presented above is a simple application of the Definition 3.1. Therefore, the calculation of the quantile in a general situation (arbitraty x) will follow similar routine.

4 Prediction of quantiles in the FOT probability approach

In this section we would like to present a method of prediction of the quantile using the FOT probability approach. The general assumption is that we observe a function $x(u)$ on the interval $[t, t+T]$. Then, the basic question is:

Having observed the function until time $t+T$, how to predict its empirical FOT quantile at time $t+T+\delta$?

The answer to the question above will be based entirely on the previously presented method of calculating quantiles and empirical FOT distributions. In our opinion, the FOT approach has a lot of advantages over the classical, stochastic methods. First of all, no assumption on independence of observations or stationarity of distributions is necessary.

This Section will start with the presentation of a method of predicting quantile for a case when the observed signal x_s is stepwise function in the sense of Definition 2.4 with $a_k \in \{A_1, \ldots, A_k\}$. We will call this case *a finite alphabet case*. Then, in the latter part of this Section we will put forward a quantile calculation for any continuous signal $x(t)$.

4.1 Finite alphabet case

Let us first note the following useful fact:

Fact 4.1. Assume that t and T are fixed and assume that the function $x_s(t)$ is stepwise plus $a_k \in \{A_1, \dots A_M\}$ with $A_1 < \dots < A_M$ – the finite alphabet for the coefficients a_k. Let the integers p and r have the same meaning as in Fact 2.5. Therefore

$$\begin{aligned} p_i^T &= F_T(A_i) - F_T(A_{i-1}) \\ &= \frac{T_s}{T} \sum_{k=r+1}^{p-1} [\mathcal{U}(A_i - a_k) - \mathcal{U}(A_{i-1} - a_k)] \\ &\quad + \frac{(r+1)T_s - t}{T} [\mathcal{U}(A_i - a_r) - \mathcal{U}(A_{i-1} - a_r)] \\ &\quad + \frac{t + T - pT_s}{T} [\mathcal{U}(A_i - a_p) - \mathcal{U}(A_{i-1} - a_p)] \,. \end{aligned} \tag{29}$$

Observe that the last two terms of the equality (29) are of the order T^{-1}.

Proof. The desired results follows from (11) and (12).

In the formula (29) we use the convention that $\mathcal{U}(A_0 - x) = 0$ for any x.

The method of predicting quantile is based on the study of behaviour of probabilities p_i^T when the length of observation changes from T to $T + \delta$. Obviously, the precision of the prediction depends on how large the horizon δ is. Here we will assume that $\delta = T_s$, that is, the horizon of the prediction is equal to the length of the window. In finance, this corresponds e.g., to constructing a one-day predictor of Value-at-Risk, when the data are gathered daily and possible values of the return $x(\cdot)$ come from a finite set. With that assumption in mind we have the following

Fact 4.2. Let t, T and x_s be as in Fact 4.1. Assume also that the time horizon δ of the prediction is equal to T_s. Therefore, accountig for (11) it results

$$\begin{aligned} F_{T+T_s}(\xi) &= \frac{T_s}{T + T_s} \sum_{k=r+1}^{p} \mathcal{U}(\xi - a_k) + R(t, T + T_s; \xi) \\ &= \frac{T}{T + T_s} F_T(\xi) - \frac{t + T - (p+1)T_s}{T + T_s} \mathcal{U}(\xi - a_p) + \frac{t + T - pT_s}{T + T_s} \mathcal{U}(\xi - a_{p+1}) \end{aligned} \tag{30}$$

and, hence,

$$\begin{aligned} p_i^{T+T_s} &= F_{T+T_s}(A_i) - F_{T+T_s}(A_{i-1}) \\ &= \frac{T}{T + T_s} p_i^T + R_1(t, T; i) + R_2(t, T; i) \,, \end{aligned} \tag{31}$$

where

$$R_1(t,T;i) \triangleq -\frac{t+T-(p+1)T_s}{T+T_s}\left[\mathcal{U}(A_i-a_p)-\mathcal{U}(A_{i-1}-a_p)\right] \tag{32}$$

and

$$R_2(t,T;i) \triangleq \frac{t+T-pT_s}{T+T_s}\left[\mathcal{U}(A_i-a_{p+1})-\mathcal{U}(A_{i-1}-a_{p+1})\right] \ . \tag{33}$$

Note that in (31) for $t = rT_s$ and $t+T = pT_s$ it results $R_2(t,T;i) = 0$ and the term $R_1(t,T;i)$ cannot be evaluated since it depends on the unobservable value a_p. If $pT_s < t+T < (p+1)T_s$, then $R_1(t,T;i)$ can be evaluated whereas $R_2(t,T;i)$ cannot since it depends on the unobservable value a_{p+1}.

Corollary 4.3. From Fact 4.2 it follows that for each $1 \le i \le M$ and for t, T and x_s such as in Fact 4.1 we have

$$\sum_{i=1}^{k} p_i^{T+T_s} = \frac{T}{T+T_s}\sum_{i=1}^{k} p_i^T + \sum_{i=1}^{k}\left[R_1(t,T;i)+R_2(t,T;i)\right] \ , \tag{34}$$

where, accountig for the fact that for arbitrary a_q

$$\sum_{i=1}^{k}\left[\mathcal{U}(A_i-a_q)-\mathcal{U}(A_{i-1}-a_q)\right] = \begin{cases} 1 & A_1 \le a_q \le A_k \\ 0 & \text{otherwise} \end{cases} \tag{35}$$

it results

$$\sum_{i=1}^{k} |R_n(t,T;i)| \le \frac{T_s}{T+T_s} \ , \quad n = 1,2 \ . \tag{36}$$

Moreover, (34) can be written as

$$\sum_{i=1}^{k} p_i^{T+T_s} = A(T,T_s;k) + \epsilon\, B(t,T,T_s) \ , \tag{37}$$

where

$$A(T,T_s;k) = \begin{cases} \dfrac{T}{T+T_s}\displaystyle\sum_{i=1}^{k} p_i^T & t = rT_s \ , \ t+T = pT_s \ , \\ \dfrac{T}{T+T_s}\displaystyle\sum_{i=1}^{k} p_i^T + \sum_{i=1}^{k} R_1(t,T;i) & \text{otherwise} \end{cases} \tag{38}$$

and

$$B(t,T,T_s) = \begin{cases} \dfrac{T_s}{T+T_s} & t = rT_s \ , \ t+T = pT_s \ , \\ \dfrac{t+T-pT_s}{T+T_s} & \text{otherwise} \end{cases} \tag{39}$$

in which the value of $\epsilon \in \{0,1\}$ depends on the unobservable symbol.

The problem of **finding** the quantile $q_\alpha^{T+T_s}$ consists in calculating the integer k such that

$$\sum_{i=1}^{k} p_i^{T+T_s} < \alpha \leq \sum_{i=1}^{k+1} p_i^{T+T_s} \tag{40}$$

and to put $q_\alpha^{T+T_s} = A_{k+1}$. However, on the basis of the observations in $[t, t+T]$ we cannot calculate the probabilities $p_i^{T+T_s}$ but only p_i^T. Therefore, the problem of **predicting** the quantile is calculating a predictor $\hat{q}_\alpha^{T+T_s}$ only using observations from the interval $[t, t+T]$. Consider the following

Quantile prediction algorithm for stepwise functions

Step 1 For each $i \in \{1, \ldots, M\}$ calculate the empirical probabilities p_i^T.

Step 2 Make the following procedure:

for $\epsilon' \in \{0, 1\}$

  for $\epsilon'' \in \{0, 1\}$

    find k such that

    $A(T, T_s; k) + \epsilon' B(t, T, T_s) < \alpha \leq A(T, T_s; k+1) + \epsilon'' B(t, T, T_s) \quad (*)$

    if disequality $(*)$ is true **then**

      $\hat{q}_\alpha^{T+T_s}(\epsilon', \epsilon'') = A_{k+1}$

    end

  end

end

In the above algorithm there are four cases to consider: $\epsilon' = \epsilon''$ (two cases) and $\epsilon' \neq \epsilon''$ (two cases). Therefore the proposed prediction algorithm provides at least one and at most four possible predicted values for the quantile q_α.

In that context, a natural question arises:
Given a certain precision θ, how large T is sufficient to claim that the quantile stabilizes at that precision, that is $|\widehat{q}_\alpha^{T+T_s} - q_\alpha^T| \leq \theta$?
This question has a simple answer for stepwise functions x_s as in Fact 4.1.

Remark 4.4. Assume that we would like to predict a quantile $q_\alpha^{T+T_s}$ with a given precision θ, $0 < \theta < 1$. Then it suffices to take the sample from the signal of the length T such that

$$T > \frac{4 - \theta}{\theta}$$

and put $\widehat{q}_\alpha^{T+T_s} \stackrel{\triangle}{=} q_\alpha^T$, where q_α^T is a quantile calculated according to the algorithm presented in Section 3.

Proof. It suffices to see that, according to Corollary 4.3

$$|\sum_{i=1}^{k}(p_i^{T+T_s} - p_i^{T})| \leq \frac{4T_s}{T+T_s}.$$

Therefore, at a precision θ the quantile calculation algorithm for the distribution F_{T+T_s} will not be changed if we replace $\sum_{i=1}^{k} p_i^{T+T_s}$ by $\sum_{i=1}^{k} p_i^{T}$. This argument can be seen more clearly noting that solving the inequality (*) in the quantile prediction algorithm for a significance level α is equivalent to finding a quantile from the empirical fraction-of-time probability F_T on the level $\alpha - \theta$, where θ is sufficiently small. This last observation is true do to the property of left-continuity of F^{-1}.

4.2 Continuous functions

In this part we will analyze the quantile prediction algorithm for any continuous function x. As it was seen in the previous subsection, the quantile prediction algorithm produces a (finite) set of values for the predictor $\hat{q}_\alpha^{T+T_s}$. It is easy to deduce that in the case of the continuous function x we will end up with the interval of possible predictors. This intution will be clarified in this subsection. We start with the technical result, convienient while calculating the predicted quantile.

Fact 4.5. Assume that the function x generating the FOT distributions F_T and F is continuous. Assume also that the time horizon δ is arbitrary and let $F_{T+\delta}$ be the empirical fraction of time distribution generated by x on the interval $[t, t+T+\delta]$. Then

$$F_{T+\delta}(\xi) = \frac{T}{T+\delta}F_T(\xi) + \frac{1}{T+\delta}\int_{t+T}^{t+T+\delta} \mathcal{U}(\xi - x(u))\,\mathrm{d}u. \tag{41}$$

Proof of this Fact is standard and follows from Fact 2.5.

It is easy to see that the second term of the right hand side of (41) is smaller than $\frac{\delta}{T+\delta}$. In another words,

$$F_{T+\delta}(\xi) \in [\frac{T}{T+\delta}F_T(\xi), \frac{T}{T+\delta}(F_T(\xi) + \frac{\delta}{T})].$$

This observation allows us to formulate the following

Quantile prediction algorithm for continuous functions

Step 1. Given the observation of the continuous function x calculate the F_T.
Step 2. Calculate

$$\hat{q}_\alpha^{T+\delta}(1) = \inf\{\xi : \frac{T}{T+\delta}F_T(\xi) + \frac{\delta}{T+\delta} \geq \alpha\}$$

and

$$\hat{q}_\alpha^{T+\delta}(2) = \inf\{\xi : \frac{T}{T+\delta} F_T(\xi) \geq \alpha\}.$$

It is easy to see that the above algorithm produces the interval of predictions of the form $[\hat{q}_{T+\delta}(1), \hat{q}_{T+\delta}(2)]$ and that the length of this interval, $\frac{\delta}{T+\delta}$ converges to zero while $T \to \infty$ for any fixed time horizon δ. This yields the following

Remark 4.6. Assume that we would like to predict a quantile $q_\alpha^{T+\delta}$ with a given precision θ, $0 < \theta < 1$. Then it suffices to take the sample from the signal of the length T such that

$$T > \frac{\delta \cdot (1-\theta)}{\theta}$$

and put $\widehat{q}_\alpha^{T+\delta} \stackrel{\triangle}{=} q_\alpha^T$, where q_α^T is a quantile calculated according to the algorithm presented in Section 3.

Appendix A

In this Appendix the convergence in the temporal mean-square sense of time averages and integrals is discussed. A more comprehensive treatment can be found in [15] for stationary time-series and in [2] for almost-cyclostationary time series. The more general convergence in the sense of distributions (generalized functions) of statistical functions defined starting from a single time-series is treated in [11].

In this paper, unless otherwise indicated, all the time averages are assumed to exist and to be convergent in the temporal mean-square sense (t.m.s.s.), that is, given a time-series $z(t)$ and defined

$$z_\beta(t)_T \triangleq \frac{1}{T}\int_{t-T/2}^{t+T/2} z(u)\, e^{-j2\pi\beta u}\, \mathrm{d}u \,, \tag{A.1}$$

$$z_\beta \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{+T/2} z(u)\, e^{-j2\pi\beta u}\, \mathrm{d}u \,, \tag{A.2}$$

it is assumed that

$$\lim_{T\to\infty} z_\beta(t)_T = z_\beta \qquad \text{(t.m.s.s.)}\,, \qquad \forall\beta\in\mathbb{R} \tag{A.3}$$

that is,

$$\lim_{T\to\infty} \left\langle |z_\beta(t)_T - z_\beta|^2 \right\rangle_t = 0\,, \qquad \forall\beta\in\mathbb{R}\,. \tag{A.4}$$

It can be shown that for a finite-power time-series $z(t)$ (i.e., with $\langle|z(t)|^2\rangle_t < \infty$) the set $B \triangleq \{\beta\in\mathbb{R} \,:\, z_\beta \neq 0\}$ is countable, the series $\sum_{\beta\in B}|z_\beta|^2$ is summable [2], and from (A.3) it follows that

$$\lim_{T\to\infty} \sum_{\beta\in B} z_\beta(t)_T\, e^{j2\pi\beta t} = \sum_{\beta\in B} z_\beta\, e^{j2\pi\beta t} \qquad \text{(t.m.s.s.)}\,. \tag{A.5}$$

Note that the right-hand side in (A.5) is just the almost-periodic component contained in the time-series $z(t)$.

It is worth to underline that, in the FOT probability framework, the variance of the estimators is defined in terms of the almost-periodic component extraction operation, which plays the same role played by the statistical expectation in the stochastic process framework [4], [5]. Thus, unlike the stochastic process framework, where the variance accounts for fluctuations of the estimates over the ensemble of sample paths, in the FOT probability framework the variance accounts for the fluctuations of the estimates in the time parameter, e.g., the central point of the finite-length time-series segment adopted for the estimation. Therefore, the assumption that the estimator asymptotically approaches the true value (the infinite-time average) in the mean-square sense is just equivalent to the statement that the variance of the estimator in the FOT probability sense approaches zero as the collect time approaches infinity. That is, for such time-series, estimates obtained by using different time segments are asymptotically independent of the central point of the segments.

References

[1] Besicovitch, "Almost-periodic functions," 1932.

[2] W.A. Brown, "On the theory of cyclostationary signals," Ph. D. dissertation, Dep. Elect. Eng. Comput. Sci., Univ. California, Davis, 1987.

[3] H. Furstenberg, "Stationary processes and prediction theory," Princeton University Press, Princeton, NJ, 1960.

[4] W.A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory.* Prentice Hall, Englewood Cliffs, NJ, 1987.

[5] W.A. Gardner, "An introduction to cyclostationary signals," Chapter 1 in *Cyclostationarity in Communications and Signal Processing*, W.A. Gardner, Ed., pp. 1-90, IEEE Press, New York, 1994.

[6] W.A. Gardner and W.A. Brown, "Fraction-of-time probability for time-series that exhibit cyclostationarity," *Signal Processing*, vol. 23, pp. 273-292, June 1991.

[7] W.A. Gardner and C.M. Spooner, "The cumulant theory of cyclostationary time-series, Part I: Foundation," *IEEE Trans. Signal Processing*, vol. 42, pp. 3387-3408, December 1994.

[8] E.M. Hofstetter, "Random processes," Chapter 3 in *The Mathematics of Physics and Chemistry,* vol. II, H. Margenau and G.M. Murphy eds., Princeton, NJ: D. Van Nostrand Co., 1964.

[9] H.L. Hurd and T. Koski, "On the Wold isomorfism and Wold decomposition for prediction of cyclostationary sequences," in *Proc. of Workshop on Cyclostationary Processes*, Noisy Le Grand, F, July 1996.

[10] L. Izzo and A. Napolitano, "The higher-order theory of generalized almost-cyclostationary time-series," *IEEE Trans. Signal Processing*, vol. 46, pp. 2975-2989, November 1998.

[11] E. Pfaffelhuber, "Generalized harmonic analysis for distributions," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 605-611, November 1975.

[12] R.D. Reiss, *Approximate Distributions of Order Statistics, with applications to nonparametric statistics.* Springer Series in Statistics, 1989.

[13] C.M. Spooner and W.A. Gardner, "The cumulant theory of cyclostationary time-series, Part II: Development and applications," *IEEE Trans. Signal Processing*, vol. 42, pp. 3409-3429, December 1994.

[14] K. Urbanik, "Effective processes in the sense of H. Steinhaus," *Studia Mathematica*, XVII, 1958.

[15] N. Wiener, "Generalized harmonic analysis, " *Acta Mathematica*, vol. 55, pp. 117-258, 1930.

[16] H.O.A. Wold, "On prediction in stationary time series," *Ann. Math. Statist.*, vol. 19, pp.558-567, 1948.