University of Macerata (Italy)

Department of Economics and Law

Quantitative Methods for Economics Policy

cycle: XXXII

University of Paris 1

Panthéon-Sorbonne (France)

Laboratoire du Centre d'Économie

de la Sorbonne (CES)

Economic Sciences

Reliability stochastic systems and rational expectations

Ph.D. Candidate

Dr. Jessica Riccioni

University of Macerata

Ph.D. Advisor

Prof. Roy Cerqueti

Coordinator PhD Programme

Prof. Luca De Benedictis

University of Paris 1

Panthéon-Sorbonne

Ph.D. Advisor

Prof. Jorgen-Vitting Andersen

Coordinator PhD Programme

Prof. Mouez Fodh

Year: 2020

Dedication

Francesco, Family and Friends

Acknowledgements

I would like to express my sincere gratitude to Prof. Roy Cerqueti for his unwavering support and belief in me before, during and after my PhD.

I am deeply grateful to Prof. Jorgen-Vitting Andersen for his invaluable advice and patience during my PhD study.

Their experience, competence and generous knowledge have inspired me in all the time of my academic research and daily life.

Contents

Introduction and literature review

Assessing the failure time of systems with interconnected components is a significant problem in reliability theory and leads to important questions in mathematical statistics and probability modelling.

This PhD thesis aims to develop research in this field by proposing a stochastic model for evaluating the expected time of failure of a system from a rational expectations perspective.

In particular, our aim is to address a reliability problem by exploiting the characteristics of k-out-of-n systems (with homogeneous and heterogeneous components) through rational expectations. Accordingly, the proposed approach consists of two indispensable frameworks: the framework for reliability theory with special focus on k-out-of-n systems and the framework for rational expectations.

We investigate this line of research from two perspectives which have similar characteristics, but will be explored in two different contexts.

Our main interest is the analysis of reliability systems, an area in which many investigations have been carried out, and in particular, studies have focused on the prediction of failure times under various assumptions. There are two different approaches to dealing with the study of reliability systems: a probabilistic approach by analyzing the probability distribution of a system's failure time and a Bayesian computational approach by estimating the average failure time of a system, conditioned by the description of a scenario in which the evolution of the given reliability system is observed.

Regarding the first approach, many contributions have appeared in the literature over the years, applying different methods in various fields. These research papers

have focused on the reliability function. This function is used in lifetime data models. It indicates the probability over the duration of a system and it is a function of time.

Navarro et al. [1] further developed coherent systems with dependent components. Their aim was to compare systems by considering properties of their components. They developed theorems and proofs showing that properties of coherent systems could be selected on the basis of properties of k-out-of-n systems and mixtures. Khaledi and Shaked [2] considered the residual life of systems with equal or different types of component using a stochastic comparison. They considered systems of n components such that, if r components (with $r < n$) have failed, the systems are still active. They found upper and lower bounds on the expected residual life of the systems described above by considering some applications. Navarro et al. [3] extended the classical signature-based mixture representation to the systems with exchangeable components.

Navarro and Spizzichino [4] analysed a stochastic comparison of systems with component lifetimes sharing the same copula. They investigated different vectors of marginal distributions of series systems without redundancy whose reliability function is given by the product of the reliability functions of its component lifetime. They also obtained theoretical results concerning parallel systems and for general coherent systems with independent or dependent components. The same type of research was carried out by Di Crescenzo and Pellerey [5], although they considered systems with components linked via suitable mixtures. These authors provided results for systems of two components in series and for systems of two components in parallel, showing that their lifetimes can be improved by replacing these systems with certain mixtures of duplicated (dependent) components.

Navarro et al. [6] determined ordering properties for coherent systems by taking into account the system reliability function as a distorted function of the common component reliability function. They provided many examples to validate the procedure. This type of study can be extended to the rank-dependent expected utility model.

Navarro et al., continued their research on coherent systems with dependent or

independent components in [7]. They demonstrated sufficient conditions on the components and lifetimes and on the number of components, to improve the reliability of the whole system according to different stochastic orders. The examples included support their study applied to the optimal random allocation of components in series and parallel systems.

Still using the stochastic order-based approach, Gupta et al. [8], compared the residual lifetime and the inactivity time of failed components of coherent systems with the lifetime of a system that had the same structure and the same dependence. Among the results they found those for the likelihood ratio order, reversed failure rate order, failure rate order, and the usual stochastic order.

Azaron et al. [9] introduced a new approach to determine the reliability function of time-dependent systems with standby redundancy, assuming that not all elements of the system are set to function from time zero. At the beginning the reliability graph works because the input and output elements are connected to the first path (main elements) whose components are functioning. There is a switch from one path to the next when components fail. When all connections between input and output are broken, the system fails.

Recently, in the context of modern digital systems, Borgonovo et al. [10] exploited the importance of system components in a computationally efficient way in system design, proposing a measure for time-independent reliability analysis. Through examples of coherent and non-coherent systems and a realistic digital system application, they introduced a new measure, that has probabilistic and geometric interpretations, and the concept of time consistency.

Note also Parsa et al. [11], which described a new stochastic order based on the Gini-type index (GT index), useful as a tool to gain information on the ageing properties of reliability systems, and thereby demonstrating the different characteristics of active or already failed components. The GT index has been applied to series systems, parallel systems and parallel-series systems with shared components.

A contribution about the mean time to failure and availability of semi-Markov missions, has been published by Çekyay and Özekici [12]. They studied generaliza-

tions and extensions of this research, building a model with two characteristics: (i) durations and sequence of the phases are random; (ii) the lifetimes of the components have general distributions.

Other recent studies, such as Zarezadeh et al. [13], investigated the joint reliability of two coherent systems with shared components, obtaining a pseudo-mixture representation for the joint distribution of the systems failure time. A coherent system is a system without irrelevant components and with a structure function that increases in each of the n components. They provided many cases of coherent systems with shared components and an application in which two groups of these systems were compared stochastically.

Oe et al. [14] used autoregressive models to predict the failure of a stochastic system through four types of performance index of the variations. They detected the failure of a cutting tool for a lathe and predicted the width of flank wear by using time series analysis. For this purpose, they considered four types of performance index used in the prediction methods: quadratic distance of AR parameter differences, variance of the residuals, Kullback information and distance of the Kullback information (divergence measure).

Regarding Bayesian approaches to the reliability problem, many researchers investigated system reliability in the operational research field and developed a variety of methods.

Some of them have focused on estimating the mean failure time and reliability function of a system through asymmetric loss functions. Bhattacharya [15] dealt with a priori information on the values that the parameters can assume in the context of life testing techniques. Basu and Ebrahimi [20] on the research line of life testing problems, calculated some reliability Bayesian estimators of the mean lifetime using asymmetric loss functions. El-Sayyad [16] discussed a comparison between new unbiased estimators for the parameter in the exponential distribution and the corresponding Bayesian estimators derived. In his paper Canfield [17] approached a reliability estimation problem by providing a solution that contains the prior information in the Bayesian theory and the loss function through the exponential model

of reliability. Considering that in some estimation and prediction problems the use of symmetric loss functions may be inappropriate, Varian [18] considered the properties of optimal forecasts under asymmetric loss from a theoretical perspective introducing a Linex loss function in the context of real estate assessment. Zellner [19] discussed properties of estimation and prediction procedures based on Varian's asymmetric Linex loss function. He delved into examples on normal n-mean problem, on least squares estimator of a multiple regression coefficient, on linear combination of regression coefficients.

Among the Bayesian approaches, we find papers in several fields.

Van Noortwijk et al. [21] determined a Bayesian failure model for hydraulic structures based on observable deterioration characteristics. They exploited an independent probability density function that is a mixture of exponentials through the non-negative increments of deterioration. The knowledge available in this model is about the average amount of deterioration per unit time.

Gunawan and Papalambros [22] used Bayesian statistics to develop a reliability-based optimization method for problems with incomplete information in engineering design. Considering a Bayesian binomial inference technique, they estimated a Beta distribution through a uniform prior with a precision proportional to the number of samples.

Mastran [23], Mastran and Singpurwalla [24], Barlow [25], Martz et al. [26], Martz and Waller [27] and Peng et al. [28] proposed a Bayesian approach in reliability analysis with multilevel heterogeneous data sets (binomial data, lifetime data and degradation data). Specifically, Mastran presented a procedure that adopt a squared-error loss function for two cases: (i) procedure for attribute (pass-fail) test data whose results are valid for any systems; (ii) procedure for variables (time-to-fail) data whose results are valid for non-exponential data. Mastran and Singpurwalla illustrated a method for coherent structure of independent components about prior data. They proposed examples on series system of independent components and on paralle system with component interdependence. Barlow studied a procedure about lifetime data reliability for evaluating moments of the system distribution. He

proposed a combination of information between components and systems. Martz et al. assumed a Beta prior distribution for the development of a Bayesian reliability method for series system of independent binomial subsystems and components.They presented a very detailed procedure and explanatory examples about either test or prior data at three or more configuration levels in the system. Martz and Waller extended this procedure to arbitrary system configurations of series/parallel subsystems of other subsystems or components. Peng et al. integrated multilevel pass-fail, lifetime, and degradation data in the context of reliability analysis.

Kim et al. [29] developed a method for predicting failures of a partially observable failing system that can be applied to a wide range of deteriorating stochastic systems with multivariate condition monitoring data. They assumed system with 3-state continuous time homogeneous Markov process.

Aktekin and Caglar [30] studied a software reliability model, with modeling of a multiplicative failure rate whose components evolve stochastically over testing stages. They discussed its Bayesian estimation and modeled parameters via Markov chain Monte Carlo methods.

Jiang and Skorupski [90] adopted a Bayesian approach to estimate variance components in a multivariate generalizability theory. They described a BUGS code for the implementation of several examples.

In our study we also dealt with systems with dependent components as in the papers of Navarro et al. [1, 6, 7], Khaledi and Shaked [2], Navarro and Spizzichino [4], Gupta et al. [8], Azaron et al. [9] and Oe et al. [14].

However, our research belongs to the second group of publications. We are not concerned with studying the reliability function, but we use a Bayesian approach to estimate the average failure time of stochastic systems, where the failures depend on the number and importance of the components.

Since the reliability of a system depends on the reliability of its components, within this very extensive research field we focus our attention on k-out-of-n systems. The k-out-of-n systems are those in which a system of n components requires at least k components to operate. The stochastic systems we simulate in our study fit exactly

into this type of approach. In the literature many researchers have studied k-out-of-n systems in which the failure process of each component depends on the conditions in its operating environment.

Da Costa et al. [32] used in reliability theory a martingale approach to apply active redundancy or minimal standby redundancy depending on the type of systems treated.

Eryilmaz [44] carried out an in-depth analysis of the concept of mean residual life as a fundamental characteristic that has been widely used in dynamic reliability analysis.

Wang et al. [34] considered the reliability estimation problem of weighted k-out-of-n multi-state systems. They derived the unbiased system reliability estimator and the corresponding unbiased covariance estimator applying the universal generating function method and giving three practical examples.

Zhang et al. [46] proposed a model to incorporate the information from environmental observation in the evaluation of system performances. They used Monte-Carlo simulations in numerical examples to confirm the accuracy and the efficiency of the model.

Some researchers, such as Barlow and Heidtmann [36], Jain and Gopal [37] and Rai et al. [38], evaluated the reliability of k-out-of-n systems by creating and applying different methods. Others, such as Pham and Upadhyaya [39] and Hecht and Hecht [40], measured the efficiency of k-out-of-n systems for more accurately predicting failure rates of very high reliability systems.

The type of component distribution is fundamental and is typically assumed to be negative-exponential, with consequent very consistent prediction errors when different distributions are used. In this direction, Li and Zuo [41], Sarhan [42], Asadi and Bayramoglu [43] and Eryilmaz [44] used Markov models to analyze of failure rates in the case of components with negative-exponential distributions. On the other hand, Van Gemund and Reijns [45] calculated the average failure time of k-out-of-n systems with a single cold standby for arbitrary distributions, trying to minimize the prediction errors in cases where the distribution shapes differ from the negative-

exponential one. The latter distribution seems to be the best for predicting the average failure times of a system with interconnected components, because the greater the weight of the component within the system, the faster its reliability decreases.

There are also reliability studies of k-out-of-n systems with heterogeneous components and random weights, in which each component may make a different contribution to the system reliability. Those by Xie and Pham [47], Li and Zuo [48], Eryilmaz and Bozbulut [49], Eryilmaz [50, 51] and Zhang [52, 53] are particularly interesting.

Applications of reliability analysis to k-out-of-n systems with homogeneous components (see Milczek [54] and Ge and Wang [55]) are less frequent.

It is also worth mentioning some papers, such as Boland and El-Neweihi [56], Taghipour and Kassaei [57] and Sheu et al. [58], in which the components fail randomly somesubsequent to shocks. In fact, the distribution of these failure times at time t follows a Poisson process. In the first case homogeneous components are studied, and in the other two cases, non-homogeneous components.

Here we deal with both kinds of system, composed of homogeneous or heterogeneous components with the intention of understanding, through an extensive simulation approach, which type of initial distribution is better in the process of predicting the mean failure time of stochastic systems with interconnected components. In our case there is an aspect of randomness that concerns the selection of the components and the process of determining their failure. However, this is not a Poisson process, as the failure time of the components is deterministic and depends on the increasingly assumed importance of the components, as will be explained in more details in the following.

Throughout the literature we see how the reliability of a system is linked to its ability not to fail within a certain time. There is therefore a relationship with the time, evaluated on the basis of the information the system provides us with. This allows us to estimate the behavior and the impact on system reliability.

Depending on the information stored by the system, we have the possibility to evaluate its reliability by following and predicting its failure over time. The investigation of ways to exploit available information and the discipline of forecasting direct

us towards the second framework that concerns us here: rational expectations.

In this thesis we address the problem posed above by developing it in the context of rational expectations.

In the modern economic theory, agents are assumed to adjust their expectations to anticipate future decisions. Over the years in the study of economic behavior, two main hypotheses have been put forward: adaptive expectations and rational expectations.

Adaptive expectations assumes that people form their expectations by giving importance to past events when predicting future outcomes. They played a key role in the 1960s and 1970s (see Friedman [59] and Chow [60]).

Instead, with rational expectations, people make their decisions (which on average will be correct) based on all the best information available (see for example Muth [61], Lucas [62], Sargent et al. [63], Sargent and Wallace [64] and Barro [65]). Rational expectations are a fundamental assumption in many theoretical models with implications for economic analysis, and thanks to the increasing accessibility of big data in recent years, studies have been carried out on the use of rational expectations to identify prediction errors in large samples. So, rational expectations are important in any situation in which behavior is influenced by expectations (see Maddock and Carter [66]).

They can be applied in macroeconomic models, in microeconomic situations and in a wide selection of other models.

Starting from the general equilibrium, the insights of rational expectations can be deduced analytically. From the equation of supply and demand and the assumptions of rational expectation, we obtain:

$$p_t^\star = \mathbb{E}[p_t|\mathbb{I}_{t-1}] \tag{1}$$

where $p_t^\star$ is the price expectation, $\mathbb{E}$ is the expected value operator, p_t is the price at time t, and $\mathbb{I}_{t-1}$ is all the information available at time $t-1$.

Many authors have studied rational expectations over the years in several fields.

For example, Hansen and Singleton [67] estimated and tested parameters of eco-

nomic agents dynamic objective functions within nonlinear rational expectations models.Their model did not require a priori assumptions but only the specification of the economic context in which to apply it.

In the economic and financial field, Blanchard and Watson [68] examined rational expectations in the financial markets through the study of speculative bubbles, while Delcey and Serge [69] studied the relationship between the efficient market hypothesis and rational expectations.

In the field of operations research, there is a paper by Atıcı et al. [71] on discrete time domains. They generalized the Cagan's model of hyperinflation. They changed the idea of the time horizon of the original model without considering only the time intervals represented by the integers. The time domain of their generalized model therefore includes all the real numbers of the considered time horizon.

Research by Becker et al. [70] campared the rational expectations hypotheses with the bounds and likelihood heuristic to explain the average forecasting behavior.

In our study, we apply rational expectations to reliability systems. Through rational expectations, we can predict failure times on the basis of the knowledge about the past experience of our stochastic systems and condition the results on the information provided by the system – with interacting elements with unknown random lifetimes – thus obtaining a continuous gain on prediction performance.

To the best of our knowledge, there are no similar contributions in the existing literature that address reliability problems with rational expectations through stochastic systems.

We aim to demonstrate how the prediction of failure times improves with the enhancement of the information collected, causing the failure of stochastic systems through the failure of the interconnected components. The intention is to use all available information efficiently in order to process it and make forecasts.

Generally, in the literature (see for example Sanyal et al. [72], Krishnamurthy and Mathur [73], Gokhale et al. [74], and Yacoub et al. [75]), reliability problems are addressed using scenario analysis on real systems that we will call *"in-vivo"* systems. After the failure of a system, scenario analysis is performed to understand what

happened to that specific system. Here, we propose a new methodology for forecasting using rational expectations in the field of reliability theory. In fact, for the predictions we use all the information contained in an *"information set"*, and specifically, all the different realizations of the weights of the nodes. In our approach we record the failure data of a very large number of systems and provide the results of different scenarios on our information set.

We assess the reliability of our stochastic systems on the basis of several statistical indicators calculated on the realizations of the weights, and we also take into consideration some initial weight distributions to obtain comparisons both between the various cases studied and with respect to what is contained in the existing literature. In our research we compare different types of initial weight distribution to investigate whether the results are consistent with the literature and identify the distribution with the highest predictive power among the cases treated. Rational expectations are given, in this context, by the expected value of the failure time under the constraint of the realization of a given configuration.

The method is inspired by the one proposed by Andersen and Sornette [76, 77] for the prediction of failure time of the overall system, conditioned on the information revealed by the damage occurring up until the time at which the system is being evaluated (the configurations of the theoretical setting). Their idea was influenced by the method known as "reverse tracing of precursors" (RTP) (see Keilis-Borol et al. [78], Shebalin et al. [79]) for earthquake prediction based on seismicity patterns.

However, unlike them, we insert an interaction into our systems in such a way that the failure of a component affects the rest of the still active system. The reallocation rule works according to preference relations among the components. Once one of the components fails, its relevance is reallocated to the remaining active components. We seek to provide the most reliable estimate of the failure time of a stochastic system, whose components are interconnected through rational expectations.

The question we want to answer is: which kind of information available can be used to understand what is the optimal prediction?

To do this we have developed two different approaches that will be explained in

detail in the following and which will be explored from two perspectives, characterized by peculiar aspects that will allow us to study two separate contexts. In both cases the failure time is the variable of interest to be predicted.

In the first context, our intention is to analyze how much the functioning of the system and its probability of failure depend on the level of the statistical value under study. We link the information set (the levels of the statistical indicators of the weight components of stochastic systems) with the residual failure times. An average residual failure time is assigned to each observed value. We then calculate the errors with respect to the residual failure times of the statistical indicator values of the *in-vivo* systems. A comparative analysis of the error distributions will allow us to understand how to obtain the best prediction. Time remains implicit in this approach. We use rational expectations by conditioning them to all available information and we represent errors by taking away the time from the final visualization of the results and focusing on the values of the statistical indicators and the errors associated with each value. It will be useful to understand how the statistical indicators behave individually and how the levels of errors vary according to the size class they belong to.

In the second context we exploit rational expectations to find the average prediction error conditioned on the percentiles of the information set (the levels of the statistical indicators of the weight components of stochastic systems). As time evolves, the likelihood of the failure of the systems increases and we can see from the results of this model how we gain predictive power. In this model the time dependence is direct. We issue an optimal prediction of the failure time at a specific time t and we see how the model performs as time advances. In this application, after studying the trends of the individual indicators, a comparative analysis will be presented for the different forms of conditioning proposed. An important comparative study between the different methods concludes this second line of research.

We closely follow the method od prediction studied by Andersen and Sornette [76, 77].

The evolution dynamics leading to the failure of the system is explored from a

theoretical point of view, as well as through these two extensive simulation contexts. A large number of analyses are then built up to simulate the failure of the system and we are able to identify the connections among the configurations and the lifetime of the system. In so doing, we derive a probability distribution for the random time of failure of the system conditioned on the configurations.

Set against this background, we seek improvement along five important statistical dimensions, valid for both models: variance, kurtosis, skewness, Gini coefficient, and Shannon entropy.

In Chapter 1 we propose a theoretical framework for the evolution of the system lifetime which is valid for both contexts of research presented above. The theoretical setting connects the two macro frameworks underlying this thesis: rational expectations and reliability of stochastic systems.

From here, we continue towards the two computational models, both depending on time and on the statistical indicators being analysed, to obtain a complete view of what is proposed in the theoretical model.

In Chapter 2 the statistical measure predictor role model with aggregated time is implemented with an extensive simulation analysis to validate the effectiveness of the theoretical setting. The results will be presented by emphasizing the role of statistical measures through a cross-sectional analysis over time. We aggregate over time in the display of the final graphs.

Chapter 3 is devoted to the time predictor role model with conditioning on statistical measure percentiles. We present a new algorithm with high computational complexity as we focus on time by proposing three levels of prediction linked to three percentiles of the distributions of statistical indicators. This allows us to present both individual and aggregated results with a final comparison between the various statistical methods proposed.

For both contexts we started from the difficulty of the methods for predicting the failure of reliability systems based on proper monitoring of the system that fails over time. The general goal is the implementation of a procedure that can be useful for any system with interactive components.

These models are two complementary studies that show two different approaches, focusing first on statistical indicators and then on time.

This gives us an overview of this combination of the two research areas we wanted to bring together to contribute to the proposed literature in an innovative way.

This could have real practical relevance in economics and finance, for example for banking networks, but also for assessing the systemic risk of a country, the Eurozone, or sovereign credit, among other things.

We draw our conclusions at the end of the thesis.

Chapter 1

The theoretical framework

1.1 The model

We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ containing all the random quantities used throughout the paper. We denote the expected value operator related to the probability measure $\mathbb{P}$ by $\mathbb{E}$.

We denote the *reliability system* – or, simply, *system* – by $\mathbf{S}$, and assume that it is composed by n *components* denoted by $C_1, \ldots, C_n$ and collected in a set $\mathcal{C}$.

As we will see, in our framework the system can be considered of k-out-of-n type in that it fails when some of its components fail.

The *state* of $\mathbf{S}$ is a binary quantity. If the system is *active* and works, then its state is 1. Otherwise, the state of $\mathbf{S}$ is 0, and the system is said to be *failed*. The state of $\mathbf{S}$ evolves in time, and we denote by $Y(t)$ the state of the system at time $t \geq 0$. At the beginning of the analysis (time $t = 0$) the system is naturally assumed to be in state 1.

Analogously, the state of the j-th component C_j at time t is denoted by $Y_j(t)$, and it takes value 1 when C_j is active and 0 when C_j is failed. At time $t = 0$ we have $Y_j(0) = 1$, for each $j = 1, \ldots, n$.

1.2 Main assumptions on the system

We now point out three natural assumptions of our model, which are tailored on standard reliability theory: first, the different components of the system are assumed to be not homogeneous in terms of their relevance; second, the components of the system are interconnected and exhibit different levels of interconnection; third, relevance and interconnection levels change over time, according to the change of the status of the components of the system.

We enter the details.

For each $j = 1, \dots, n$ and $t \geq 0$, the *relative importance of the component* C_j over the entire system at time t is measured through $\alpha_j(t)$, where $\alpha_j : [0, +\infty) \to [0, 1]$ and $\sum_{j=1}^{n} \alpha_j(t) = 1$, for each t.

For each $t \geq 0$, we collect the $\alpha(t)$'s in a time-varying vector $\mathbf{a}(t) = (\alpha_j(t))_j$, where

$$\mathbf{a} : [0, +\infty) \to [0, 1]^n \qquad \text{such that} \qquad \mathrm{t} \mapsto \mathbf{a}(\mathrm{t}). \tag{1.1}$$

If a component is not active at time t, then its relevance for the system is null. Moreover, each active component has positive relative relevance, i.e. the system does not contain irrelevant active components. Formally,

$$\alpha_j(t) = 0 \Leftrightarrow Y_j(t) = 0. \tag{1.2}$$

Condition (1.2) is useful, in that it allows to describe the status of the components of the system directly through the α's.

For each $j = 1, \dots, n$, the relative relevance of C_j changes in correspondence of the variation of the state of one the components of the system. Once a component fails, then it disappears from the reliability system – i.e., its relative relevance becomes null – and the relative relevances of the components of the remaining active ones are modified on the basis of a suitably defined *reallocation rule*.

Below we illustrate a blob graph containing the operation of the proportional reallocation rule:

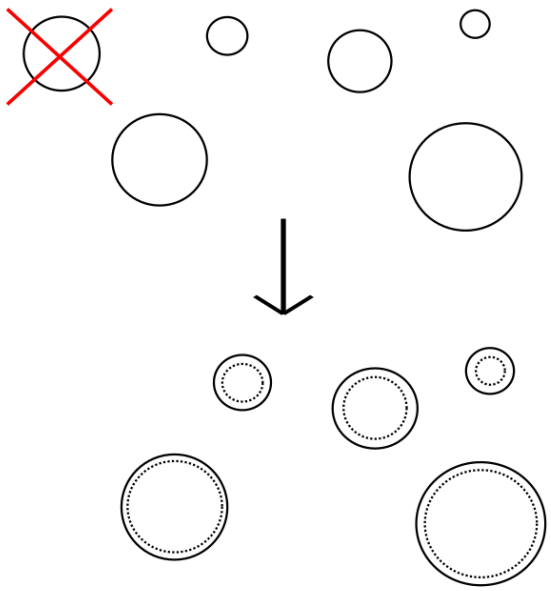

FIGURE 1.1: A blob graph representing the proportional reallocation rule: [Top] A component has failed and is deleted from the system. [Bottom] The relevance of the failed component is reallocated over the remaining active components proportionally to their α's (bubble size) before the failure.

Next example proposes a way to build a reallocation rule.

Example 1. *Consider a system* $\mathbf{S}$ *whose components set is* $\mathcal{C} = \{C_1, C_2, C_3, C_4, C_5\}$. *Assume that, at time* $t = 0$, *we have* $\alpha_1(0) = 0.1$, $\alpha_2(0) = 0.15$, $\alpha_3(0) = 0.3$, $\alpha_4(0) = 0.2$, $\alpha_5(0) = 0.25$.

Now, suppose that the first failure of one of the components of the system occurs at time $t = 7$, *when* C_3 *fails. Of course,* $\alpha_j(t) = \alpha_j(0)$, *for each* $t \in [0, 7)$ *and* $j = 1, 2, 3, 4, 5$. *Moreover,* $\alpha_3(7) = 0$.

We consider a specific reallocation rule, which states that the relevance is reallocated over the remaining active components proportionally to their α*'s before the failure. This means that*

$$\alpha_1(7) = \frac{0.1}{0.1 + 0.15 + 0.2 + 0.25},\ \alpha_2(7) = \frac{0.15}{0.1 + 0.15 + 0.2 + 0.25},$$
$$\alpha_3(7) = 0,\ \alpha_4(7) = \frac{0.2}{0.1 + 0.15 + 0.2 + 0.25},\ \alpha_5(7) = \frac{0.25}{0.1 + 0.15 + 0.2 + 0.25}.$$

In general, if τ_1, τ_2 *are the dates of two consecutive failures, with* $\tau_1 < \tau_2$, *we have*

$$\alpha_j(\tau_2) = \frac{\alpha_j(\tau_1)\mathbf{1}_{\{Y_j(\tau_2)=1\}}}{\sum_{i=1}^{5} \alpha_i(\tau_1)\mathbf{1}_{\{Y_i(\tau_2)=1\}}}, \qquad j = 1, 2, 3, 4, 5.$$

The α*'s are step functions, whose jumps occur in correspondence to the failure of one of the components.*

For what concerns the *interconnections among the components*, we define their time varying relative levels through functions of type $w_{ij} : [0, +\infty) \to [0, 1]$, for each $i, j = 1, \ldots, n$, so that $w_{ij}(t)$ is the relative level of the interconnection between C_i and C_j at time $t \geq 0$. We assume that arcs are oriented, so that in general $w_{ij}(t) \neq w_{ji}(t)$, for each t. Moreover, by construction, $\sum_{i,j=1}^{n} w_{ij}(t) = 1$, for each t. We also assume that self-connections do not exist in our framework, i.e. $w_{ii}(t) = 0$, for each i and t.

For each $t \geq 0$, the $w(t)$'s are collected in a time-varying vector $\mathbf{w}(t) = (w_{ij}(t))_{i,j}$, with

$$\mathbf{w} : [0, +\infty) \to [0, 1]^{n \times n} \qquad \text{such that} \qquad \mathrm{t} \mapsto \mathbf{w}(\mathrm{t}). \tag{1.3}$$

If C_i is a not active component at time t, then $w_{ij}(t) = w_{ji}(t) = 0$, for each $j = 1, \ldots, n$. This condition simply formalizes that a failed component is disconnected

from the system. Such a statement suggests that the failure of a component might generate disconnections among the components of the system.

The behavior of the w's is analogous to that of the α's. Also in this case, the relative levels of interconnections change when one of the components of $\mathbf{S}$ change its state, and there is a reallocation rule for the remaining levels of interconnections.

We synthesize the reallocation rules of the weights on nodes and arcs broadly by $\mathcal{R}$.

Therefore, a natural rewriting of the system $\mathbf{S}$ with components in $\mathcal{C}$ and reallocation rule $\mathcal{R}$ at time t is then

$$\mathbf{S}(t) = \{\mathbf{a}(t), \mathbf{w}(t)\}. \tag{1.4}$$

Notice that (1.4) highlights the observable features of the system with a given set of components and a specific reallocation rule, i.e. the weights on the nodes and on the arcs. Thus, according to (1.4), we can say that the $\{\bar{\mathbf{a}}, \bar{\mathbf{w}}\}$ is an *observation* of the system at a given time, where $\bar{\mathbf{a}} \in [0,1]^n$ and $\bar{\mathbf{w}} \in [0,1]^{n \times n}$.

When needed, we will conveniently remove the dependence on t from the quantities in formula (1.4).

1.3 The structure of the system

To capture the dependence of the state of $\mathbf{S}$ on the ones of its components, we simply introduce a function $\phi : \{0,1\}^n \to \{0,1\}$

$$Y(t) = \phi(Y_1(t), \ldots, Y_n(t)). \tag{1.5}$$

In reliability theory, ϕ is usually denoted as the *structure function* of the system.

We denote the elements of $\{0,1\}^n$ as configurations of the states of the components of the system or, briefly, *configurations*.

Function ϕ in (1.5) has the role of clustering the set of configurations in two subsets: the ones leading to the failure (F) of the system and those associated to the not failed (NF) system. Thus, we say that $K_F \subseteq \{0,1\}^n$ is the collection of

configurations such that $\phi(x_F) = 0$, for each $x_F \in K_F$ while $K_{NF} \subseteq \{0,1\}^n$ is the collection of configurations such that $\phi(x_{NF}) = 1$, for each $x_{NF} \in K_{NF}$. By definition, $\{K_F, K_{NF}\}$ is a partition of $\{0,1\}^n$.

In order to describe a systemic risk problem, some requirements on ϕ are needed.

First, $(0,\dots,0) \in K_F$ and $(1,\dots,1) \in K_{NF}$. This condition means that when all the components of the system are active (not active), then the system is active (not active) as well.

Second, ϕ is non-decreasing with respect to its components. This has an intuitive explanation: the failure of one of the components of the system might worsen the state of the system and cannot improve it.

Third, each component is able to determine the failure of the system. Formally, this condition states that for each $j = 1,\dots,n$ there exists $\big(y_1,\dots,y_{j-1},y_{j+1},\dots,y_n\big) \in \{0,1\}^{n-1}$ such that $(y_1,\dots,y_{j-1},1,y_{j+1},\dots,y_n) \in K_{NF}$ and $(y_1,\dots,y_{j-1},0,y_{j+1},\dots,y_n) \in K_F$.

1.4 Failure of the system and rational expectations

As said above, time $t = 0$ represents *today* – the starting point of the observation of the evolution of the system –. At time $t = 0$ all the components are active and the system works.

The failure of the system is then a random event, which occurs when the system achieves one of the configurations belonging to K_F.

We define the *system lifetime* as:

$$\mathcal{T} := \inf\{t \geq 0 | \phi(Y_1(t),\dots,Y_n(t)) = 0\}. \tag{1.6}$$

Analogously, the n-dimensional vector of *components lifetimes* is $\mathbf{X} = (X_1,\dots,X_n)$, where

$$X_j = \inf\{t > 0 \,|\, Y_j(t) = 0\}. \tag{1.7}$$

To be as general as possible, we assume that the failure lifetimes of the components of the system $\{X_1,\dots,X_n\}$ are not independent random variable and do not share the

same distribution. In fact, at each components' failure, the α's and the w's modify in accord to the reallocation rule $\mathcal{R}$; this modify also the probability of subsequent failures of the system components in the very natural case of failures dependent on the weights.

Moreover, we can reasonably assume that the failure of the system coincides with the failure of one its components.

To fix ideas, we provide an example.

Example 2. *Assume that* $\mathcal{C} = \{C_1, C_2, C_3, C_4, C_5\}$ *and*

$$\mathbf{a}(0) = (0.1, 0.5, 0.2, 0.1, 0.1), \qquad \mathbf{w}(0) = \begin{pmatrix} 0 & 0 & 0.1 & 0.1 & 0 \\ 0 & 0 & 0.1 & 0 & 0 \\ 0.1 & 0.1 & 0 & 0.1 & 0.05 \\ 0.1 & 0 & 0.1 & 0 & 0.05 \\ 0 & 0 & 0.05 & 0.05 & 0 \end{pmatrix}$$

Suppose that the reallocation rules $\mathcal{R}$ *for relative relevance and interconnection levels are of proportional type, as in Example 1. Such reallocations are implemented if the system is not failed.*

Furthermore, assume that the failure of a component has a twofold nature: by one side, it can be due by an idiosyncratic shock; by the other side, it can be driven by the failure of the other components. Specifically, we hypothesize that if a given component fails, then the components connected only to it fail as well, independently from their level of interconnections. Differently, the idiosyncratic shocks are assumed to be captured by a Poisson Process with parameter λ *– giving the timing of the failures – jointly with a uniform process over* $\mathcal{C}$*, independent on the Poisson Process – which identifies the failed component.*

Moreover, suppose that the system fails at the first time in which components with aggregated relative relevance greater than 0.4 fail.

Now, suppose that the first failure is observed at time $t = 8$*, when* C_2 *fails. Then, automatically,* C_3 *fails as well, since it is connected only to* C_2*. The aggregate relative relevance before the failures is* $\alpha_2(8^-) + \alpha_3(8^-) = 0.5 + 0.2 > 0.4$*, and the system fails.*

A rational expectations approach is used for the computation of the expectation of the random time in which the system fails. Specifically, we will compute the expected value of $\mathcal{T}$ conditioned to the specific values of the weights $\mathbf{a}(t)$ and $\mathbf{w}(t)$ measured at time t.

We denote by RE the rational expectations of the time $\mathcal{T}$ given all the possible observations of the system. Specifically, for $t \geq 0$, we set

$$RE_t = \left\{ \mathbb{E}\left[\mathcal{T} \,|\, \{\bar{\mathbf{a}}(t), \bar{\mathbf{w}}(t)\} \right] \,:\, \bar{\mathbf{a}}(t) \in [0,1]^n,\, \bar{\mathbf{w}}(t) \in [0,1]^{n \times n} \right\}. \tag{1.8}$$

Formula (1.8) provides the expected value of the lifetime of $\mathbf{S}(t)$ in correspondence of any observation of the system.

1.5 Preliminaries for the numerical validation of the theoretical setting

In this section we provide the specifications and the main definitions for implementing extensive simulations, with the aim of reproducing synthetically a large number of systems. In doing so, we offer the basis of a concrete validation of the theoretical model, which will be carried out in the next chapters.

The synthetic systems form a set, called hereafter *"information set"* (and the systems will be denoted by $\mathbf{S}^I$'s). The systems in the *information set* will be observed till their failure, and the failure times will be properly store.

As we will see, we will cluster the systems on the basis of the observations $\{\bar{\mathbf{a}}, \bar{\mathbf{w}}\}$, and any observation will be associated to the residual failure time of the corresponding system. Implicitly, time plays a key role in this context. Indeed, systems evolve and a generic $\{\bar{\mathbf{a}}, \bar{\mathbf{w}}\}$ is linked to the specific time t in which it is observed.

According to formula (1.8), rational expectations at time t are therefore computed as the expected values – statistical means, in this case – of the residual lifetimes of the *information set* systems at time t conditioned by all the possible observations available at time $t-1$.

We set $\mathbb{E}\left[\mathcal{T} \mid \{\bar{\mathbf{a}}, \bar{\mathbf{w}}\}\right] = +\infty$ when the observation $\{\bar{\mathbf{a}}, \bar{\mathbf{w}}\}$ does not appear in the set of the simulated observations.

The efficiency of the rational expectations is tested by means of a new set of systems – the *"in vivo"* systems (and such systems will be denoted by $\mathbf{S}^V$'s) – whose observations and residual failure times are compared with the ones obtained in the *information set*.

1.6 Specifications of the reliability system

We here describe the reliability system framework on the ground of the experiments we deal with. Some requirements go in the precise direction of reducing the computational complexity of the problem.

First of all, we assume that $w_{ij}(t) = \frac{1}{n(t)(n(t)-1)}$ when $i \neq j$ and $w_{ij}(t) = 0$ for $i = j$, for each i, j, $t \geq 0$ and where $n(t) = 1, \ldots, n$ is the number of components not failed at time t. This condition means that all the active components of the system are mutually connected, and connections are equivalent in terms of their entities. As already stated in the theoretical modeling, loops are not allowed.

The assumption on $\mathbf{w}$ allows to think at the observations of the system by considering only $\mathbf{a}$. Thus, we will refer hereafter only to the weights α's in order to deal with the rational expectations. In this respect, we rewrite formula (1.8) for time t as follows

$$RE_t = \left\{\mathbb{E}\left[\mathcal{T} \mid \bar{\mathbf{a}}(t)\right] \,:\, \bar{\mathbf{a}}(t) \in [0,1]^n\right\}. \tag{1.9}$$

The procedure for the failure of the components works in a stepwise form. We present it at a generic time $t = 1, 2, \ldots$. At the beginning, a random extraction of one active component is implemented, for identifying the candidate failed component. Say that C_j is the extracted component. Then, we sample a random number r. If $\alpha_j(t) > r$, then the component C_j fails at time t, and $\alpha_j(s) = 0$, for each $s > t$. Conversely, when $\alpha_j(t) \leq r$, then the component C_j does not fail. Then, the procedure is reiterated at time $t + 1$ and so on, till the failure of the system.

The reallocation rule is of proportional type, as in Example 1, so that the relevance of the failed component is reallocated over the remaining active components proportionally to their α's before the failure. Such a reallocation rule, together with the condition for the failure described above, give a clear intuition on the presence of stochastic dependence among the lifetimes of the system components.

The system is assumed to fail the first time in which the number of failed components becomes higher than the one of the active components.

Chapter 2

The role of statistical measure in predicting the failure with aggregated time

In the context of reliability systems, our goal is to demonstrate an improvement in prediction performance using rational expectations by conditioning the residual failure times of simulated stochastic systems to the memory of past events. The key to our approach is the information stored over time through rational expectations with a particul focus on the effect of different levels of statistical indicators on prediction results.

We use two sets of systems for our analysis: one (*information set*) to record the information and calculate rational expectations, and another (*in-vivo*) to compare the failure times of real systems with those predicted. The prediction errors are obtained by comparing the *information set* systems and the *in-vivo* systems.

In this first context, our attention is focused on all the values of the statistical indicators that represent the synthetic measures of our systems.

Time remains implicit and, in the final representations of the errors committed in the prediction of failure times, we will directly relate the errors to the values of the statistical indicators. In order to deeply investigate our predictions we take into account three different quantiles of the error distribution: 10%, 50% and 90% of the

distribution. We thus have a comparison between small, medium and large errors.

It will be interesting to understand how results change depending on the the characteristic considered for the analysis and on the initial distribution of the weights of the configurations.

2.1 Proposal and presentation of the simulation procedure

2.1.1 Overview of the simulation procedure

To pursue our target, the observations of the systems will be opportunely synthesized by means of statistical indicators, pointing to central tendency, variability, shape and concentration information content of the distribution of the weights of the components.

Specifically we will deal with: (i) the variance of the α's, to evaluate the variability of the observations and therefore the dispersion of the weights; (ii) the kurtosis of the α's, to study if the weights are heavy-tailed or light-tailed with respect to a normal distribution; (iii) the skewness of the α's, to provide a measure if the symmetry of the observations; (iv) the Gini coefficient of the weights α's, to evaluate the inequality within the observations; (v) the Shannon entropy of the α's, to estimate the degree of complexity of information and the mutability of the observations.

The simulation procedure will be divided for convenience into five parts.

1. We build the *information set*. We denote by $K\_simulations$ the number of simulated systems that will be recorded in the *information set*, so that the generic *information set* system will be denoted by $\mathbf{S}_k^I$, with $k = 1, \ldots, K\_simulations$.

   Each system is created from the beginning – where all the components are active – to its failure. Components fail according to the procedure described in Subsection 1.6. For each system $\mathbf{S}_k^I$ we collect the information by one of the five statistical indicators (i)-(v) presented above related to the specific observation

$\bar{\mathbf{a}}(t) = (\alpha_1(t), \ldots, \alpha_n(t))$ – where the reference to t indicates that the observation is considered at time t. Here, we introduce the failure time $\mathcal{FT}_k$ of the related system – where the index k points to the k-th system $\mathbf{S}_k^I$ of the *information set*; the residual failure time $\mathcal{T}_k(t)$ of each observation $\bar{\mathbf{a}}(t)$ and depends on k and t; the number of active components $n_k(t)$. We will follow the systems from the initial time until their failure with the time $t = 0, 1, 2, \ldots, \mathcal{FT}_k$.

By introducing opportune notation, we will collect for each k-th system:

- All the times until the failure $\mathcal{FT}_k$ with $t = 0, 1, 2, \ldots, \mathcal{FT}_k$.
- For each time t:
  - We record a set of α's of the observations $\bar{\mathbf{a}}$ that depends on k, t and $n_k(t)$ (all the weights of the $\mathbf{S}^I$'s);
  - Each set of α's is associated with one value $\bar{\star}_I$ of the statistical indicator represented by $\star$, with $\star$ =variance, kurtosis, skewness, Gini coefficient and Shannon entropy. Any observed value $\bar{\star}_I$ depends on k and t, while $n_k(t)$ intervenes directly in the computation of the considered statistical indicator at time t and for system k. The values are rounded to 3 decimal places.
- All the failure times $\mathcal{FT}$'s of systems $\mathbf{S}^I$'s in a vector that depends on the k-th system $\mathbf{S}_k^I$ and indicated with $\mathcal{FT}_k$.
- From the $\mathcal{FT}$'s we derive all the residual failure times $\mathcal{T}$'s in a matrix $\mathcal{T}_I$, that depends on the k-th system $\mathbf{S}_k^I$ and the observed value $\bar{\star}_I$ ($\mathcal{T}_k(\bar{\star}_I)$). The residual failure time measures the time until the failure of each $\bar{\star}_I$. Specifically:

$$\mathcal{T}_k(\bar{\star}_I) = \mathcal{FT}_k - t, \forall t = 0, \ldots, \mathcal{FT}, \tag{2.1}$$

2. Now, we are in the position of constructing the rational expectations in (1.9) on the basis of the $\mathbf{S}^I$'s.

   At this aim, we build an application that links each observed value $\bar{\star}_I$ of the statistical indicator represented by $\star$, with $\star$ =variance, kurtosis, skewness, Gini

coefficient and Shannon entropy, with a set of residual failure times contained in $\mathcal{T}_I$. We do it for all the *information set* systems $\mathbf{S}^I$'s.

We compute the arithmetic mean of eah set of residual failure times in correspondence of the $\bar{\star}_I$'s, namely $\mathbb{E}[\mathcal{T}_{RE}(\bar{\star}_I)]$. In so doing, we obtain the rational expectations of formula (1.9) associated to each $\bar{\star}_I$.

3. The *in-vivo* systems are built according to the same procedure of the *information set* systems. We denote by $X\_simulations$ the number of simulated *in-vivo* systems, and the generic *in-vivo* system is $\mathbf{S}^V_x$, with $x = 1, \ldots, X\_simulations$.

   In this case, by introducing opportune notation, we will record for each x-th system:

   - All the times until the failure $\mathcal{FT}_x$ with $t = 0, 1, 2, \ldots, \mathcal{FT}_x$, where $\mathcal{FT}_x$ is the failure time of the related system – the index x points to the x-th system $\mathbf{S}^V_k$ of the *in-vivo*.
   - For each time t:
     - We record a set of α's of the observations $\bar{\mathbf{a}}$ that depends on x, t and $n_x(t)$ (all the weights of the $\mathbf{S}^V$'s), where $n_x(t)$ is the number of active components.
     - Each set of α's is associated with one value $\bar{\star}_V$ of the statistical indicator represented by $\star$, with $\star$ =variance, kurtosis, skewness, Gini coefficient and Shannon entropy. Any observed value $\bar{\star}_V$ depends on x and t. The values are rounded to 3 decimal places.
   - All the failure times $\mathcal{FT}$'s of systems $\mathbf{S}^V$'s, for each x, in a vector that depends on the x-th system $\mathbf{S}^V_x$ (indicated with $\mathcal{FT}_x$).
   - From the $\mathcal{FT}$'s, all the residual failure times $\mathcal{T}$'s in a matrix $\mathcal{T}_V$, that depends on the x-th system $\mathbf{S}^V_x$ and the observed value $\bar{\star}_V$ ($\mathcal{T}_x(\bar{\star}_V)$). Every value $\bar{\star}_V$ is related to a residual failure time $\mathcal{T}_x(\bar{\star}_V)$.

   Here, we consider all the values $\bar{\star}_V$'s. These values are different a priori from the $\bar{\star}_C$'s found in the first part of the simulation procedure.

4. Now we compare the observed values of the *in-vivo* systems $\bar{\star}_V$'s with the observed values of the *information set* systems $\bar{\star}_I$'s when *in-vivo* and *information set* systems show the "same observation".

   We therefore need to replace every $\bar{\star}_V$ that has never been observed in the *information set* with the value $\bar{\star}_I$ which represents the "same observation". The "same observation" in the *information set* is the observed value $\bar{\star}_I$ which is closest to the value $\bar{\star}_V$ we are replacing.

   Specifically, we compute the minimum distance between each $\bar{\star}_V$ and the values $\bar{\star}_I$'s.

   Once we find the closest observation of $\bar{\star_V}$ among the $\bar{\star}_I$'s, we build a function that assigns to each $\bar{\star}_I$ – considered the "same observation" – the residual failure times contained in $\mathcal{T}_V$, namely $\mathcal{T}_x(\bar{\star}_I)$, linked to the $\bar{\star}_V$'s that have been replaced.

   Then, we consider, for each $\bar{\star}_I$, the error between the rational expectations and all the residual failure times of the in vivo systems assigned to the specific $\bar{\star}_I$. Specifically,

$$E_{\bar{\star}_{RE}}(\bar{\star}_I) = |\mathbb{E}[\mathcal{T}_{RE}(\bar{\star}_I)] - \mathcal{T}_x(\bar{\star}_I)|, \tag{2.2}$$

   where $\mathcal{T}_x(\bar{\star}_I)$ is the residual failure time of the in vivo systems assigned to the $\bar{\star}_I$'s to replace rational expectations.

   We create, for each $\bar{\star}_I$, a distribution of errors between the expected residual times calculated with the rational expectations and the residual failure times of the *in-vivo* systems.

5. For comparison purposes, we use the distribution of the errors in rational expectations prediction for the computation of the quantiles $q = 10\%$, $q = 50\%$ and $q = 90\%$ of the error distribution.

   In this way, we quantify the impact of the error prediction comparing $E_{\bar{\star}_{RE\_10}}$, $E_{\bar{\star}_{RE\_50}}$ and $E_{\bar{\star}_{RE\_90}}$ with the benchmark E_B for the different analysis carried

out in this work (variance, kurtosis, skewness, Gini coefficient and Shannon entropy).

The fifth part consists of the following steps:

1. Create the distribution of the quantiles $q = 10\%$, $q = 50\%$ and $q = 90\%$ of the errors distribution in rational expectations prediction ($E_{\bar{\star}_{RE}}$) for the $\bar{\star}_I$'s. We need to construct statistics over predictions in order to assess their quality. We decide to focus our attention on fixed quantiles at level q of the distributions of the errors: $q = 10\%$, $q = 50\%$ and $q = 90\%$, i.e. the 10^{th}, 50^{th} and the 90^{th} quantile. Let us consider errors in ascending order, so that the 10^{th} quantile represents the smallest errors in our forecast of residual failure times. Instead, the 90^{th} quantile identifies the biggest errors made in predicting the residual failure times of stochastic systems. The 50^{th} quantile gives us the median of errors, so the central errors of the distribution. E_{RE_10}, E_{RE_50} and E_{RE_90} are, respectively, the 10^{th} (Error 1), 50^{th} (Error 2) and the 90^{th} (Error 3) of the distribution of the errors.

   These quantiles take into account all the information recorded in the *information set* over time.

2. To obtain greater comprehensibility in the interpretation of the results, proceed with the clustering of the $\bar{\star}_I$'s (in ascending order) in equal class intervals. The class width is 50, except for the last class that includes the values excluded from previous classes.

   Compute for each class the average of the errors distribution created in the previous steps (errors distribution, average of the errors distribution and percentiles).

3. Create the benchmark errors given by the errors made without the use of rational expectations: E_B.

4. To understand how rational expectations allow us to obtain the optimal prediction of residual failure times using the information stored in the past, it is

necessary to examine the comparison between $E_{\bar{\star}_{RE\_10}}$, $E_{\bar{\star}_{RE\_50}}$, $E_{\bar{\star}_{RE\_90}}$ and E_B of the several indicators considered in the analysis (variance, kurtosis, skewness, Gini coefficient and Shannon entropy).

2.1.2 Parameter set

The parameters are set as follow:

TABLE 2.1: Parameter set

Parameter	**Value**
number of components	10
K_simulations	10.000
X_simulations	10.000

The α's and r are generated from different types of distributions which provide five possible cases for each analysed characteristic of the configurations (see Subsection 2.1.1).

Consider:

- Uniform distribution in (0,1) type;
- Beta distribution type, which has support in (0,1), composed of two positive shape parameters, denoted by α and β that represent the exponents of the random variable and control the shape of the distribution. Depending on the value assigned to the two parameters, we will get several distributions from which we can randomly extract our weights. We consider four combinations of shape parameters (see Fig. 2.1):
  - $\alpha = 1$ and $\beta = 3$ that is an asymmetric distribution more concentrated over the values close to zero;
  - $\alpha = \beta = 0.5$ that is a symmetric distribution bimodal over the extremes 0 and 1;

- $\alpha = \beta = 2$ that corresponds to a platykurtic symmetric distribution centered in 0.5;
- $\alpha = 1$ and $\beta = 0.5$ that is an asymmetric distribution on the right with a high concentration of values close to 1.
- $\alpha = \beta = 1$ is a specific distribution that corresponds to the uniform distribution defined in (0,1).

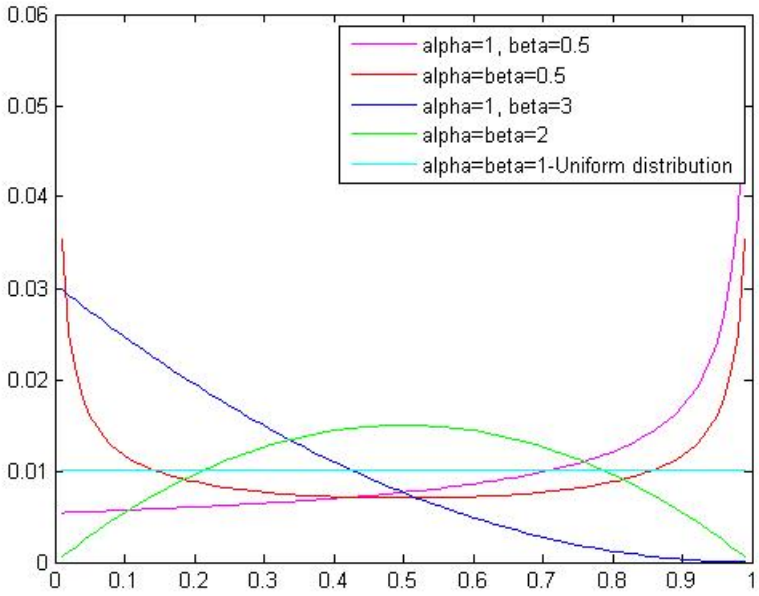

FIGURE 2.1: Beta distribution with four different combinations of shape parameters and uniform distribution.

We therefore carried out five different analysis according to the five characteristics of the α's distribution: variance, kurtosis, skewness, Gini coefficient and Shannon entropy.

The following paragraph will show the graphs related to the study of the several indicators in combination with the initial weights distributions.

2.2 Results and discussion

In this section we present the results of the study and the related discussion. The paragraph will be divided into two blocks: in the first part of the discussion we will investigate the relationship between the values of the statistical indicators and levels of error; in the second part we will go into the composition of errors by studying their frequency.

We now focus our attention on the various statistical measures under analysis.

We start from the variance. Fig. 2.2 shows the prediction errors (in absolute value) for the analysis of the variance considering all the different five types of initial weight distribution. Magenta fixed line is the E_B. Red line, blue line and light blue line, blue line and red line are respectively $E_{\bar{\star}_{RE\_10}}$ (Errors 1), $E_{\bar{\star}_{RE\_50}}$ (Errors 2) and $E_{\bar{\star}_{RE\_90}}$ (Errors 3).

Variance is a good estimator of the dispersion and the volatility of the weights of each system around the mean.

As the subgraphs of Fig. 2.2 show, we can note a decreasing trend of errors towards zero in all the different cases taken into consideration. High volatility of the system components is linked to the reduction of errors made in the prediction and it is due to the decrease of the uncertainty about the systems failure time. Data are very spread out from the mean and we are able to capture a wider range of values.

Variance is very performing and the errors stabilize around zero for all initial weights distributions except for the 50^{th} and the 90^{th} percentile of the Beta distribution with $\alpha = 1$ and $\beta = 3$. In this case from a certain threshold of variance onwards we observe a small increase in errors which emphasizes the fact that excessive fluctuations affect the predictability of the variance for medium or very large errors. When we detect a high variance we pay close attention to data without generalizing the values never seen before. So, for the negative exponential distribution, too high values and an excessive turbulence within the observations, however, are overfitting and that can highlight distortions that lead the model to become less good at predicting new data.

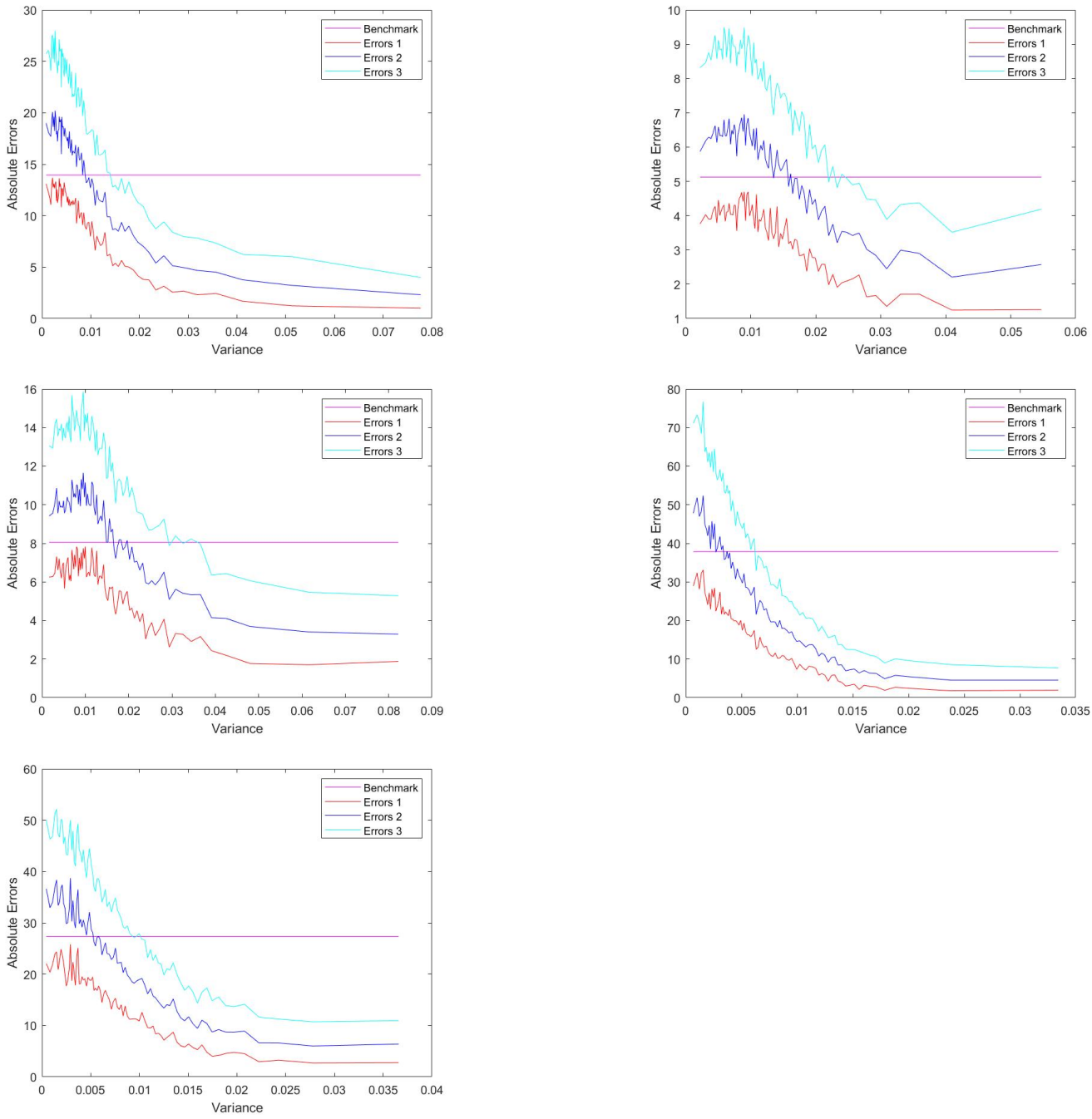

FIGURE 2.2: Prediction errors (in absolute value) for the $E_{\bar{\star}_{RE\_10}}$, the $E_{\bar{\star}_{RE\_50}}$ and the $E_{\bar{\star}_{RE\_90}}$ of the variance considering all the initial weight distributions. [Top] (Left) Uniform distribution; (Right) Beta distribution with $\alpha = 1$ and $\beta = 3$. [Center] (Left) Beta distribution with $\alpha = \beta = 0.5$; (Right) Beta distribution with $\alpha = \beta = 2$. [Bottom] (Left) Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

With low variance values, the model is underfitting the residual failure times and it is unable to identify the real moment of system failure.

From Fig. 2.2 in the initial parts of the various curves we observe a very scattered trend characterized by many spikes with errors that stabilize at a certain level of variance. From these levels of variance, the prediction of the systems failure times is almost perfect. In these windows of the paths the components are very dependent on each other; this allows us to predict how long it will take for the system to fail.

Moreover, we can note a different trend of the E_{RE_90} compared to the E_{RE_10}. The line of the biggest errors decreases faster toward zero, while in the line of the smallest errors there is a slight flat behavior. This flat behavior depends on the fact that when we have a very low variance (close to zero), in a big error the gain in the prediction is more evident than what we will see in a small error.

The different subgraphs also display the change in the initial distribution of components.

According to the existing literature the ideal shape of the weights distribution that minimizes prediction errors is the negative exponential one (see Li and Zuo [41], Sarhan [42], Asadi and Bayramoglu [43] and Eryilmaz [44]). In our study it corresponds to the Beta distribution with $\alpha = 1$ and $\beta = 3$.

In fact, the best predictive efficacy belongs to the asymmetric case with $\alpha = 1$ and $\beta = 3$ and to the symmetric one with $\alpha = \beta = 0.5$. In the first case the predictive gain is very high, although there is a worsening in the final trend which corresponds to very high levels of variance.

However, in other cases, we always obtain excellent results through the use of rational expectations with increasing levels of variance. In fact, we appreciate a constant decrease to near zero. The simmetric Beta with $\alpha = \beta = 2$ together with the asymmetric Beta with $\alpha = 1$ and $\beta = 0.5$ are the cases that perform worse as regards to the errors in absolute value, but in correspondence with appropriate variance values they reach errors close to zero with trends flattening out due to ever increasing values of the indicator.

We can therefore deduce from this analysis the levels of variance that provide an

optimal balance without overfitting or underfitting the residual failure times. This specific application suggests that the study of the variance can be a powerful tool to predict the residual failure times of a stochastic system in case of high values of variance, which are crucial to understanding the behavior of prediction models.

Jiang et al. [80] dealt with software fault prediction models by establishing that the lower the variance, the more reliable the system is. Also Twomey and Smith [81], who dealt with error estimator methods for evaluating prediction models, confirmed with their studies that good performance is linked to low variance.

On the contrary, however, in our study, high variance values are more informative and allow a better prediction, performing better than lower values and fewer fluctuations.

We now analyse the kurtosis. Kurtosis is a measure of the "tailedness" of the distribution through the description of the tails and the identification of the outliers.

In Fig. 2.3 we see the prediction errors concerning kurtosis. Error 1 $E_{\bar{\star}_{RE\_10}}$, Errors 2 represents $E_{\bar{\star}_{RE\_50}}$ and Error 3 represents $E_{\bar{\star}_{RE\_90}}$.

The situation is scattered.

As in the study of variance, the two distributions that perform better and provide lower and decreasing errors are Beta distribution with $\alpha = \beta = 0.5$ and Beta distribution with $\alpha = 1$ and $\beta = 3$. At a glance, the latter analysis is the best performing.

In the first three sub-charts we can observe a different error trend between E_{RE_90} line and E_{RE_10} line. The trend of the smallest errors is almost linear with constant errors for all the kurtosis values recorded in the simulations. Instead, the biggest errors follow a path that reaches maximum errors in correspondence of low kurtosis values and therefore in presence of a platykurtic distribution of the weights. Kurtosis loses predictive power when distribution has a lower likelihood of extreme weights as compared to a normal distribution. When the weights are distributed around their average and we observe thinner tails, the use of rational expectations proves to be ineffective.

Our model, instead, becomes successful when there is a positive excess of kurtosis

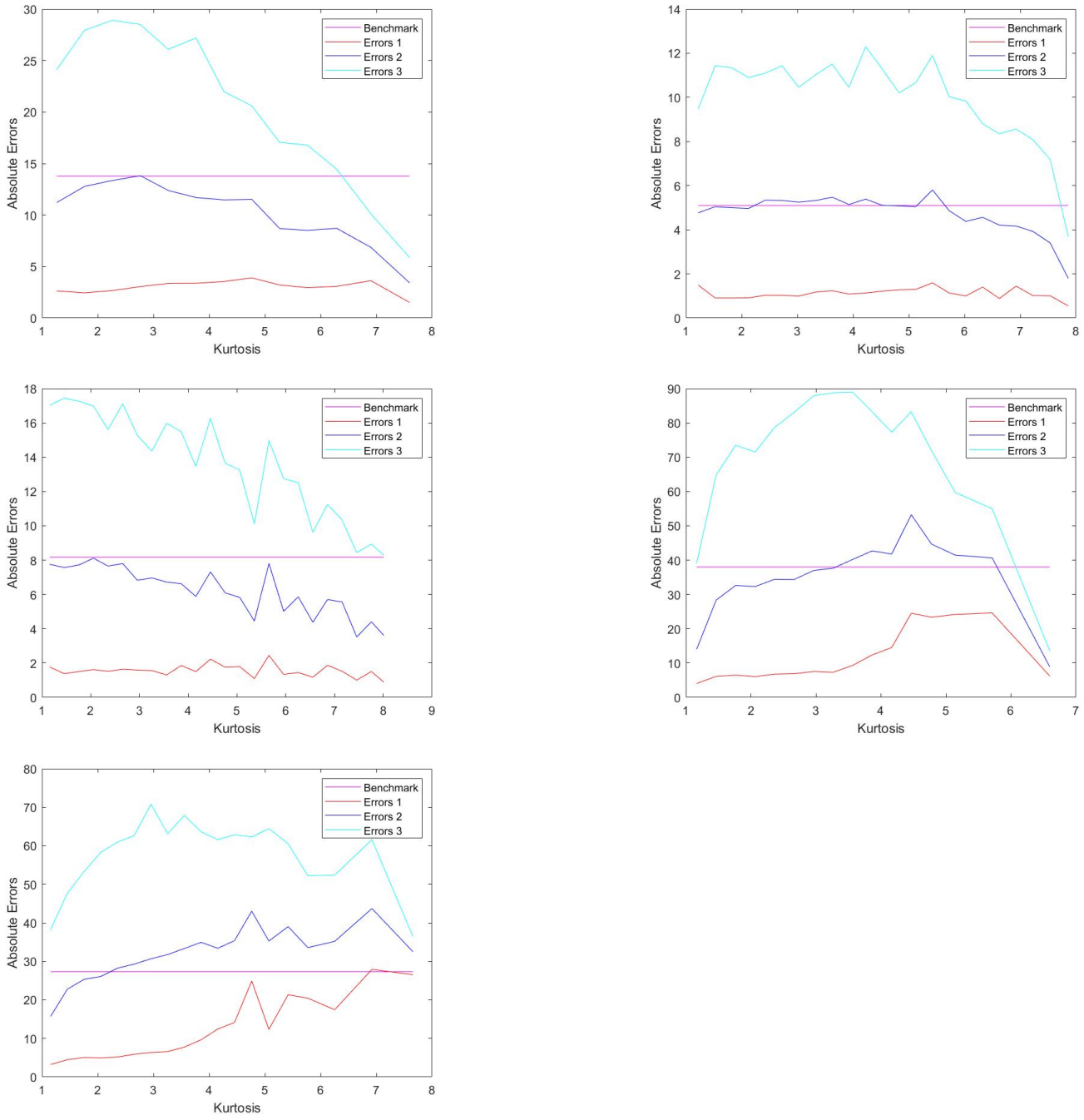

FIGURE 2.3: Here we present the study of kurtosis. See for the details Fig 2.2.

and the data are fat-tailed distributed. If we take into account even the most extreme values, we obtain an increasingly inclusive range of cataloged wights. The application of rational expectations attests to be an excellent method of forecasting the residual failure times considering even the values with greater distance from the mean.

The situation is completely reversed if we start from a symmetric weight distribution centered in 0.5 (Beta distribution with $\alpha = \beta = 2$). In correspondence with distributions with non-fat tails, kurtosis does not work well. This case represents an exception showing anomalous behavior. It provides very high error levels. It does not work because the initial values assigned to the weights extracted from this distribution are concentrated in 0.5 and it underestimates the extreme values close to 0 and 1, which are considered in the other distributions. The minimum errors committed in predicting failure times through rational expectations are those at kurtosis levels around 8. It can only be used if we have distributions with extreme values.

If, on the other hand, the weights are distributed with an asymmetry to the right (Beta distribution with $\alpha = 1$ and $\beta = 0.5$), kurtosis is a good predictor of thin tail distributions with continuously rising curves as the values increase.

Switch now to the skewness.

In Fig. 2.4 we can see the different trends of skewness errors.

Through the skewness of the components weights, we investigate the nature and dynamics of the shape of the distribution. Skewness measures how close the data are to the symmetry or the lack of symmetry in data distribution.

The skewness has been used in the prediction of future market returns with excellent results (see Jondeau et al. [82]). Also in our study it is a good predictor depending on the values assumed by the distribution.

Compared to the previous indicators, skewness fluctuates more in the central part of the curves and exhibits large spikes, especially at values close to zero when the data are fairly symmetric. The highest errors are achieved in conditions of symmetry, negative asymmetry or slight positive asymmetry of the distribution. Regardless of the initial distribution considered, a forecast bias is evident around a skewness range that goes from -1.5 to +2.5.

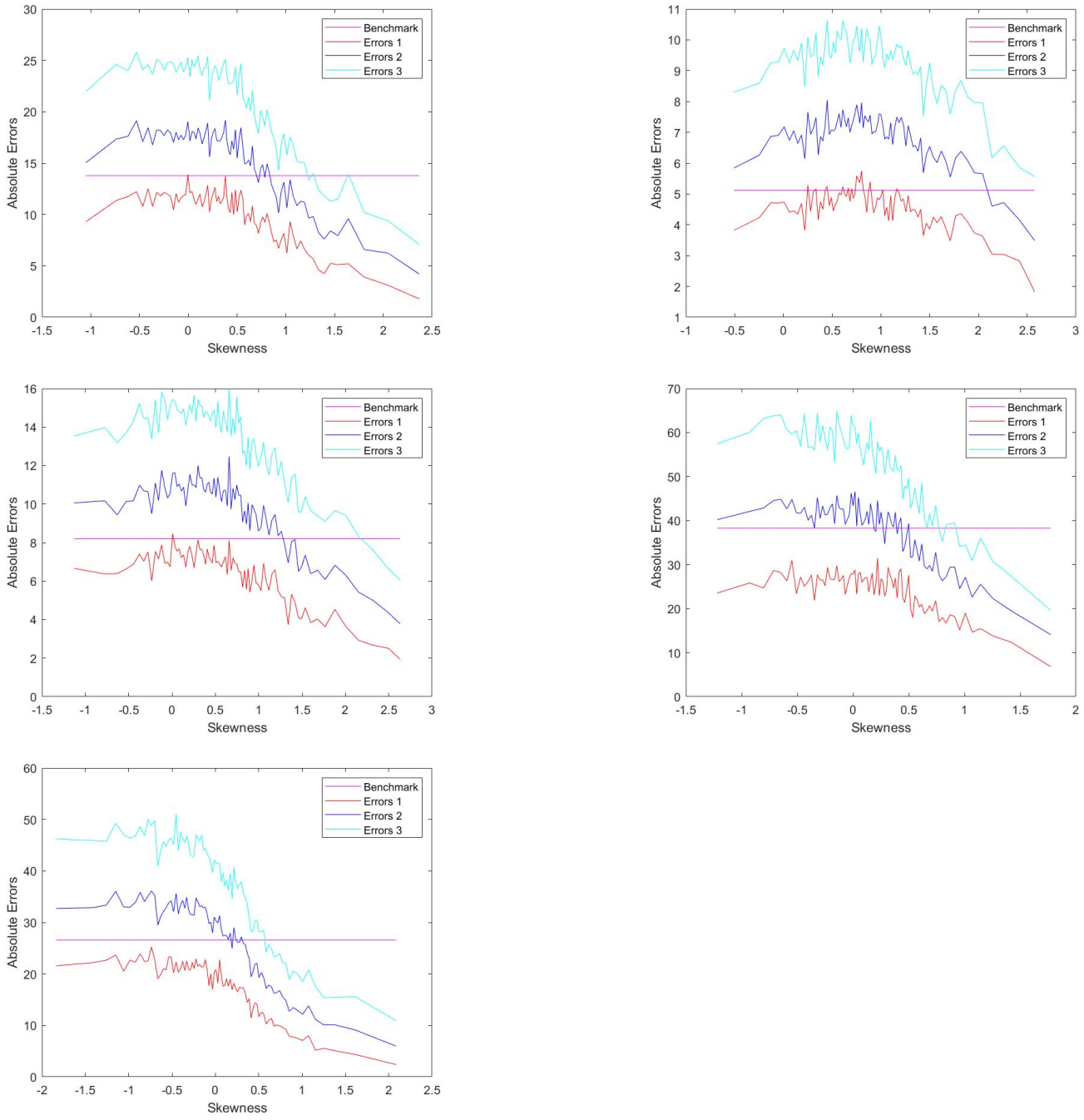

FIGURE 2.4: Prediction errors for the skewness. The five sub-graphs contain the five initial weight distributions. Refer to the caption of Fig. 2.2.

Overall, the trend of average errors and percentiles is very scattered and we do not notice big differences between the various curves. The best accuracy is obtained by applying our predictive model in cases of Beta distribution with $\alpha = \beta = 0.5$ and Beta distribution with $\alpha = 1$ and $\beta = 3$. The least predictive gain and very high maximum error levels are reached in the case of Beta distribution with $\alpha = \beta = 2$ (as in the study of the two previous indicators).

While generally highly skewed data affect the accuracy of the predictive model (see Larasati et al. [83]), the results shown in Fig. 2.4 provide evidence that extremely positively skewed weights pick up the best information with good predictive gain.

As part of the use of statistical moments in forecasting models, Reijns and Gemund [84], Amari et al. [85], Ramberg et al. [86] and Kinateder and Papavassiliou [87], state that the lowest moments of the distributions (second moment in our case, i.e. the variance) are more efficient than the highest moments which are attested to be less stable and reliable (third, skewness and fourth, kurtosis). With our analysis we confirm what these researchers concluded.

In fact, the variance is the one that comes closest to zero errors, with some features of uncertainty in the case of very low variability (components very similar to each other). The situation worsens for skewness and kurtosis. The third statistical moment is characterized by less regularity and mixed trends, but with the achievement of very low error values. The fourth statistical moment has a more flattered shape but it is less efficient overall.

Fig. 2.5 displays the prediction errors (in absolute value) for the Gini coefficient considering all the different five types of initial weights distribution.

We can consider the Gini coefficient as a measure of accuracy to assess the performance of our prediction model of residual failure times. In fact, the increase in the coefficient corresponds to an increase in inequality among the values taken into consideration in the analysis. Values close to 1 indicate that the range of possible weights cataloged and on which the study is based is very wide.

Not many researchers have used the Gini coefficient as a prediction tool. Ooghe and Spaenjers [88] have argued that it is very useful to use the Gini when dealing

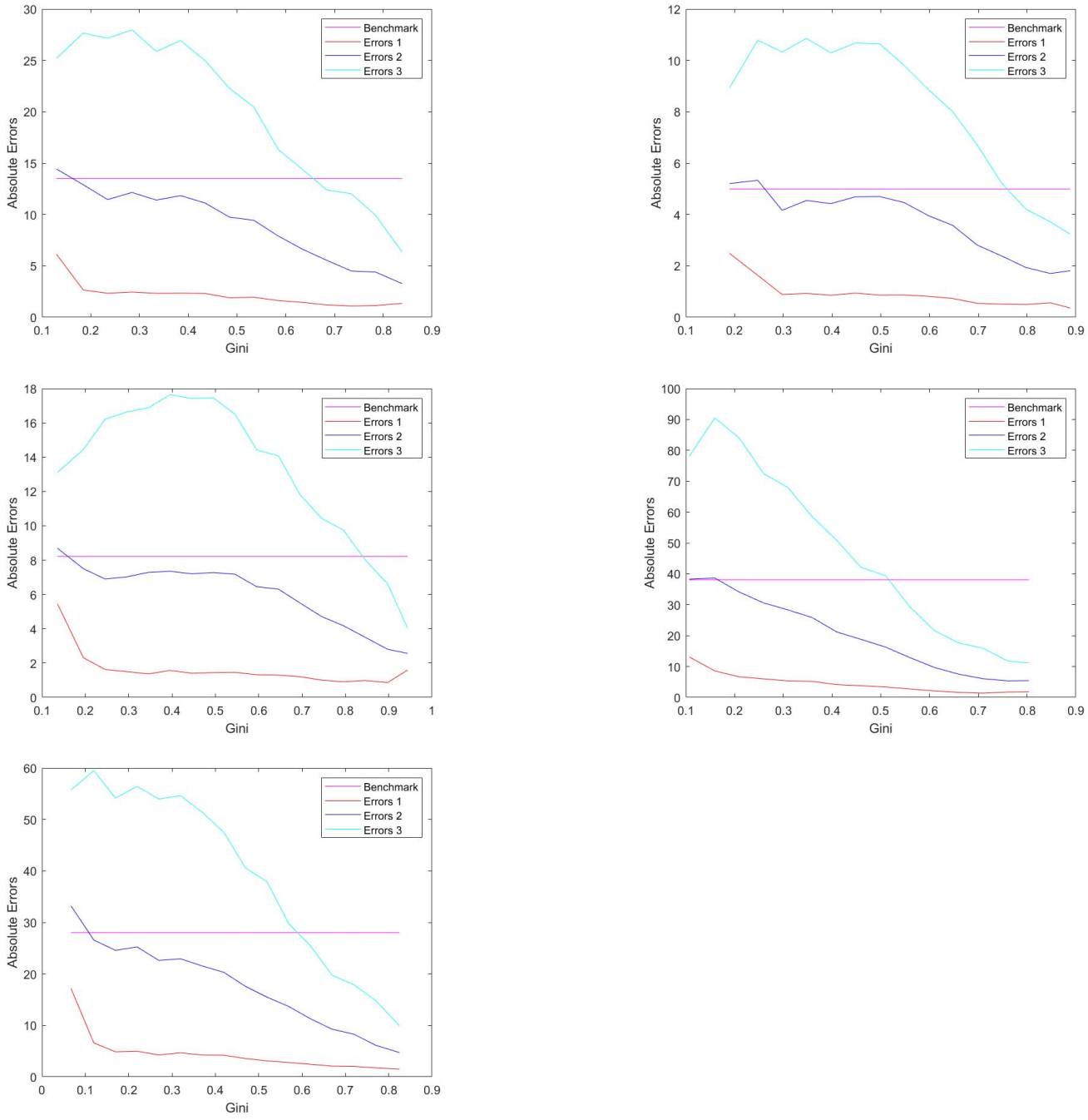

FIGURE 2.5: Case of the Gini coefficient. Magenta line of errors is the E_B. Red line, blue line and light blue line are $E_{\bar{\star}_{RE\_10}}$ (Errors 1), $E_{\bar{\star}_{RE\_50}}$ (Errors 2) and $E_{\bar{\star}_{RE\_90}}$ (Errors 3), respectively (see 2.2).

with failure prediction models.

The behavior of errors is similar to that of kurtosis in some scenarios, but there are differences to note. The comments previously left for kurtosis about the different trends of the three curves can however also be extended for the study of this indicator.

First of all, the 10^{th} percentile has a linear trend but for values of the Gini coefficient near zero (the initial trend) we observe a peak that decreases rapidly settling at error levels close to zero.

Secondly, the symmetric Beta distribution with $\alpha = \beta = 2$ and the asymmetric Beta with $\alpha = 1$ and $\beta = 0.5$ behave in the opposite way to that of kurtosis. In fact, although the level of initial errors is very high, we can visualize a predictive gain that grows exponentially as the values of Gini coefficient increase.

We can appreciate visually through error analysis the improvement in the accuracy of the predictions for increasing values of the coefficient up to the attainment of prediction errors tending to zero. Near the maximum polarization of the components we reach excellent levels of prediction.

For very small Gini values, we can deduce that we are in the initial time of the simulations, when no node or few nodes have failed. In fact, the values of the weights will be more equidistributed and the equidistribution will decrease as the components fail and the weights of the failed nodes are proportionally redistributed among the components still active. Low values of the coefficient correspond to less cataloged information and greater uncertainty associated with forecasts through rational expectations.

We end this first part of analysis by focusing on the Shannon entropy illustrated in Fig. 2.6.

There are no contributions in literature on the use of entropy as a prediction tool. Thus our analysis fills this gap.

The results obtained show a very irregular error trend characterized by many spikes that alternate touching a very wide range of values. Any improvement is very random and the use of rational expectations is not effective as a prediction tool.

For each type of initial distribution – by comparing the values with the previous

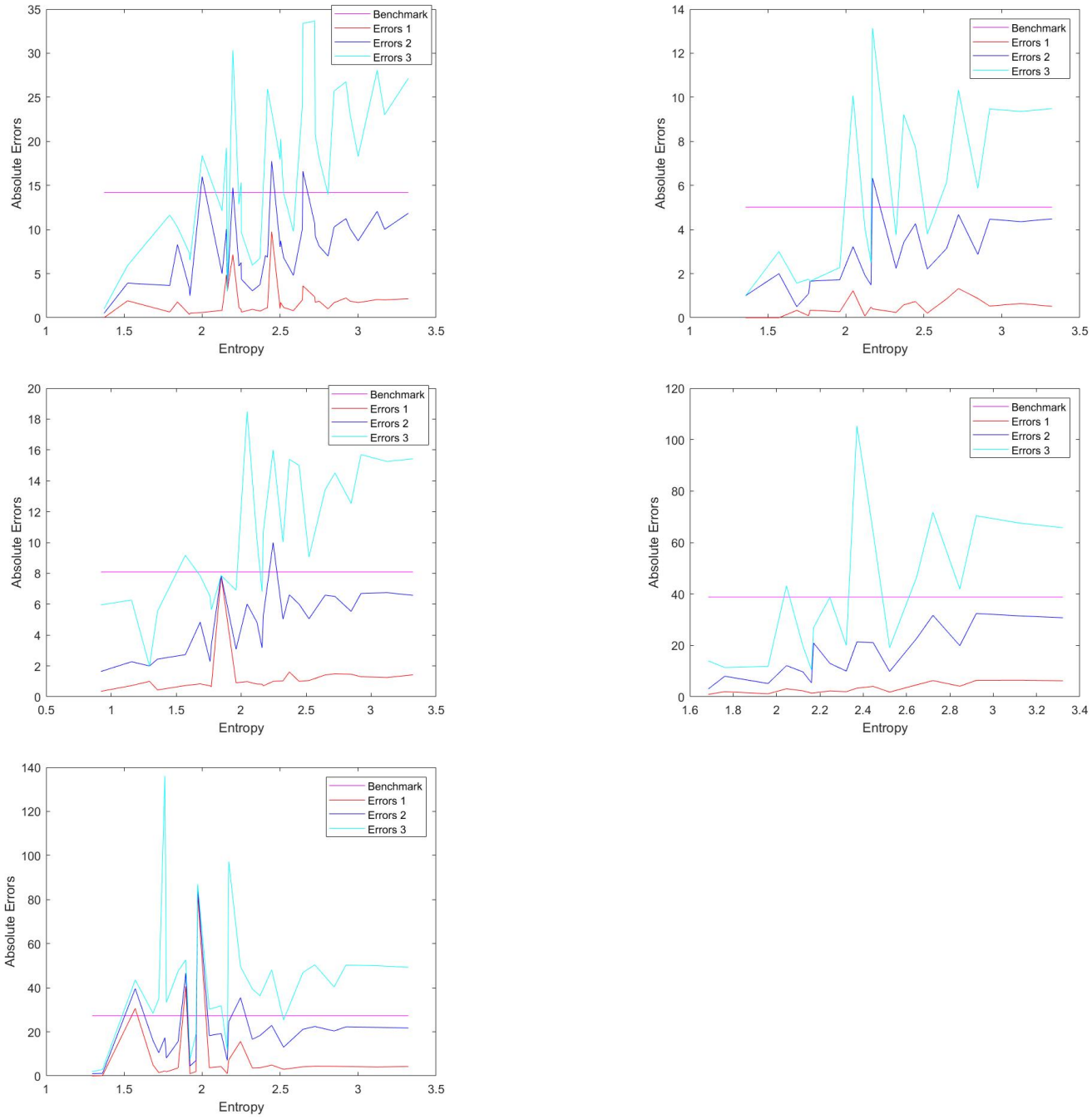

FIGURE 2.6: Shannon entropy analysis. As in the previous charts: [Top] (Left) Uniform distribution; (Right) Beta distribution with $\alpha = 1$ and $\beta = 3$. [Center] (Left) Beta distribution with $\alpha = \beta = 0.5$; (Right) Beta distribution with $\alpha = \beta = 2$. [Bottom] (Left) Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

analysis – the errors are the worst ever obtained and the possibility of distortions occurring by applying our model on entropy is evident.

In this forecasting model the Shannon entropy is the least informative and least useful statistical indicator. There is no regularity in the trends and we fail to capture information on which entropy values are best suited for better predicting failure times.

Comments relating to previous charts need to be corroborated by further analysis.

We proceed with a second part in which we can add investigations performed by the illustration of a series of histograms containing the absolute frequencies of the prediction errors.

The aim is to understand the extent of the errors committed during our study according to the statistical indicator and the initial distribution of the weights entered as input.

In each graph there are two sub-graphs, one in which the error values are divided into classes of width 10, with the residual values included in the last class whose size is variable; and another where we eliminate the zero value (no errors made in the prediction) from the figure, so that we can provide a representation of the frequencies for each error value. The analysis of the variance is reported in Fig. 2.7.

In Figs. 2.8 and 2.9 we show the error frequencies for kurtosis and skewness, respectively.

We can observe the analysis of the distribution concentration in Fig. 2.10 where we find the histograms for the five different initial distributions of the Gini coefficient.

Ultimately, we find Shannon entropy (see Fig. 2.11).

The comments for this second analysis can be quite generalized for all statistical indicators. In fact, the trend of the error frequencies is very similar in the various cases, with variance, kurtosis and Gini coefficient which are the best prediction tools, with more linear and regular trends, while skewness and Shannon entropy show deviations that invalidate the study.

And this confirms the first part of the analysis.

In general, observing the frequency of the prediction errors, these reach a very high level for errors located in the lower classes, and then decrease in a strictly monotonous

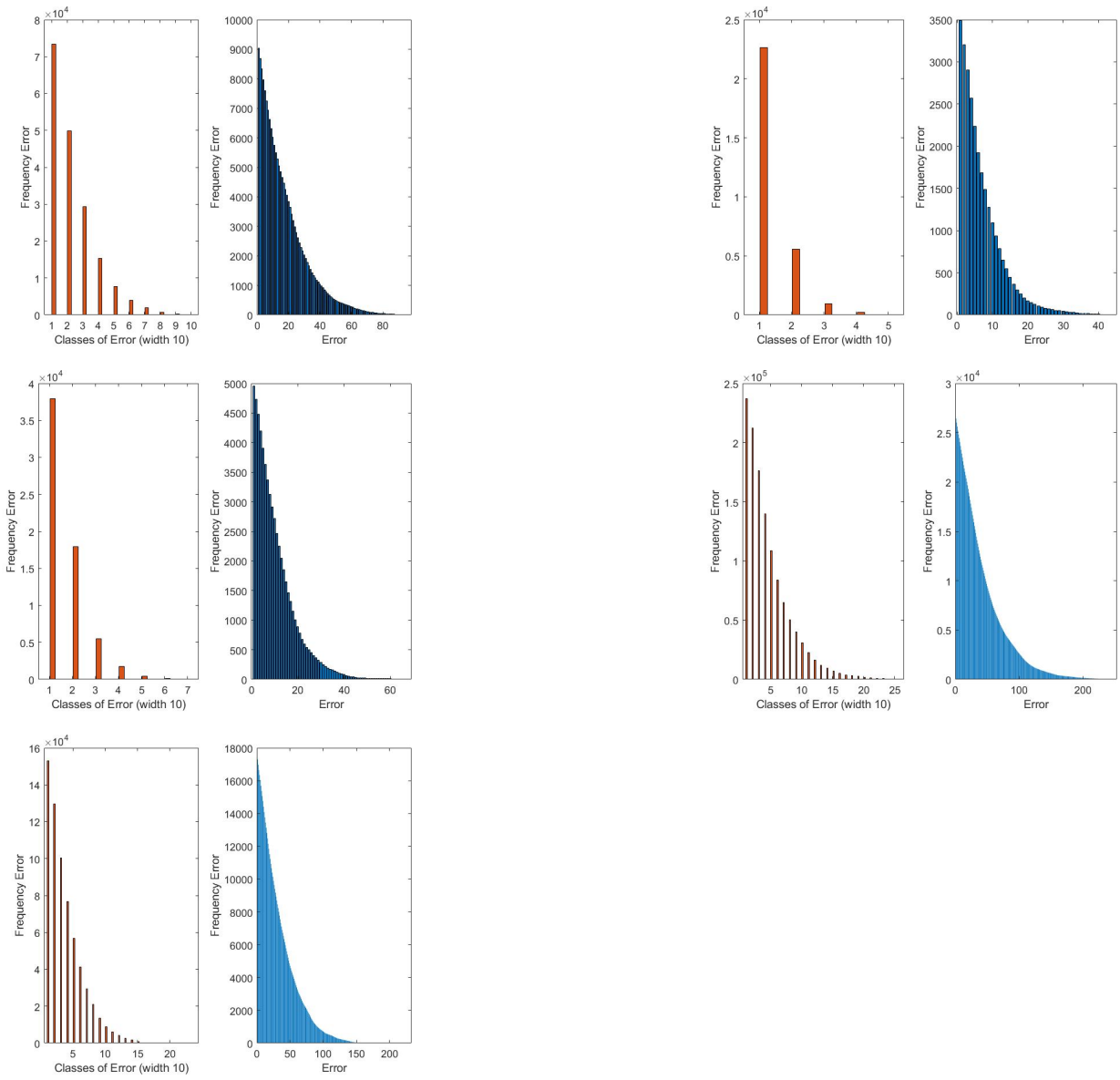

FIGURE 2.7: Histograms of the frequency of the variance prediction errors. In the first sub-graph we show the frequency for error classes of width 10 (each bar contains 10 error values in ascending order, except the last class where the final error values are inserted). In the second sub-graph, instead, we have the error frequency of each single value excluding zero from the representation. These are the errors related to the analysis of variance considering all the initial weight distributions. [Top] (Left) Uniform distribution; (Right) Beta distribution with $\alpha = 1$ and $\beta = 3$. [Center] (Left) Beta distribution with $\alpha = \beta = 0.5$; (Right) Beta distribution with $\alpha = \beta = 2$. [Bottom] (Left) Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

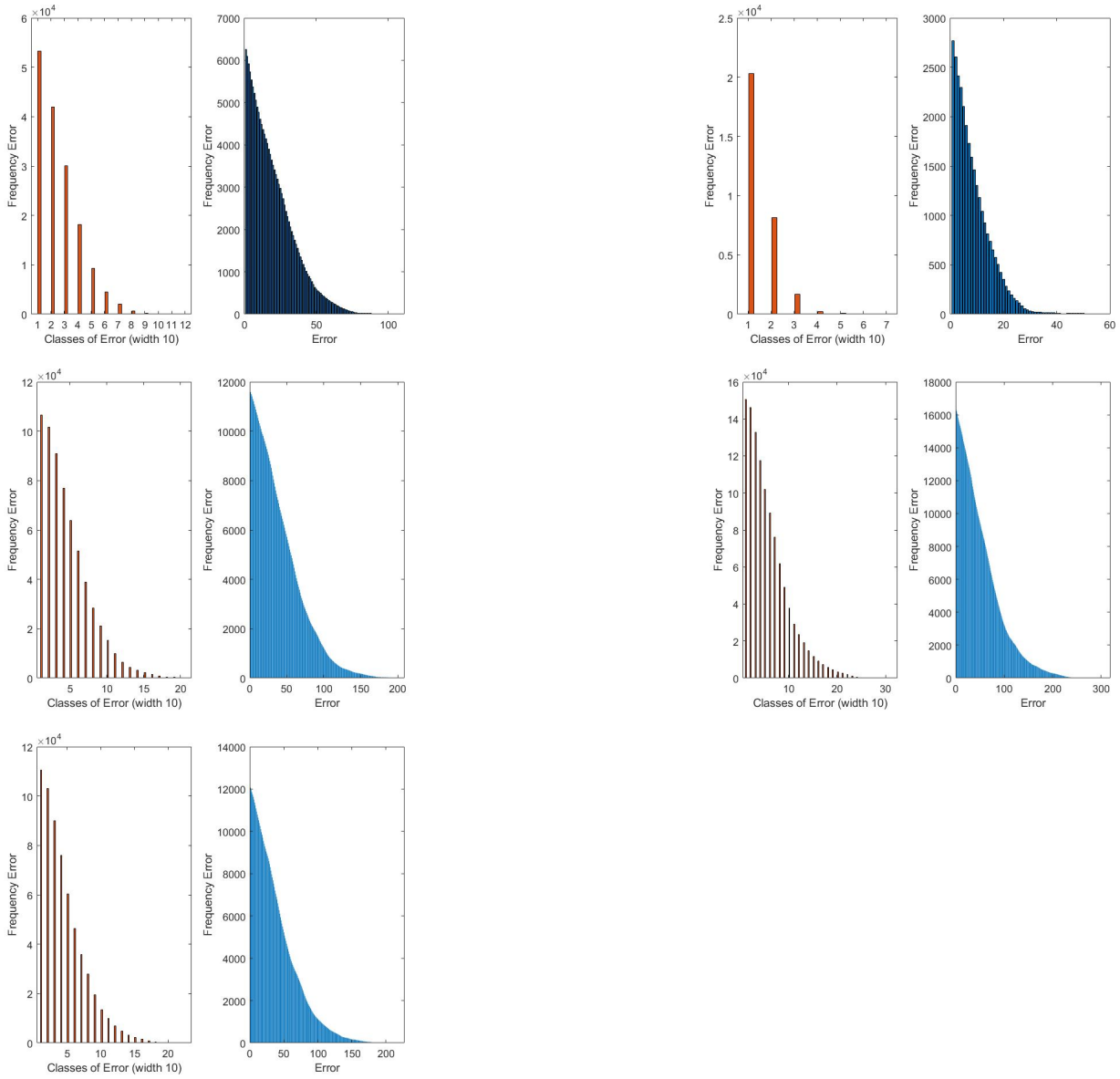

FIGURE 2.8: Histograms related to the frequency of the kurtosis prediction errors. See caption of Fig. 2.7 for a complete description of the sub-graphs.

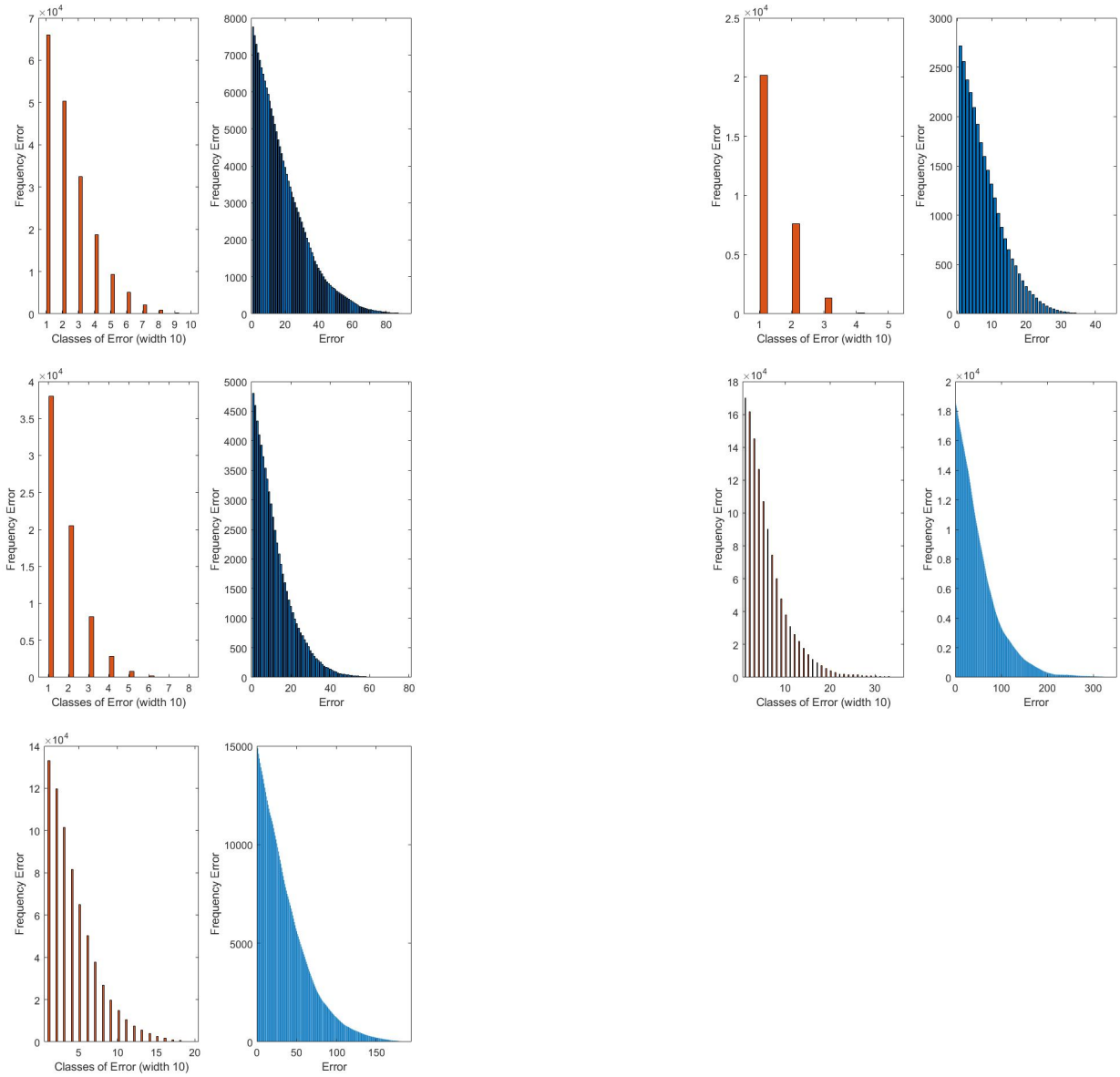

FIGURE 2.9: Here the analysis of the frequency error of skewness. More details in Fig. 2.7.

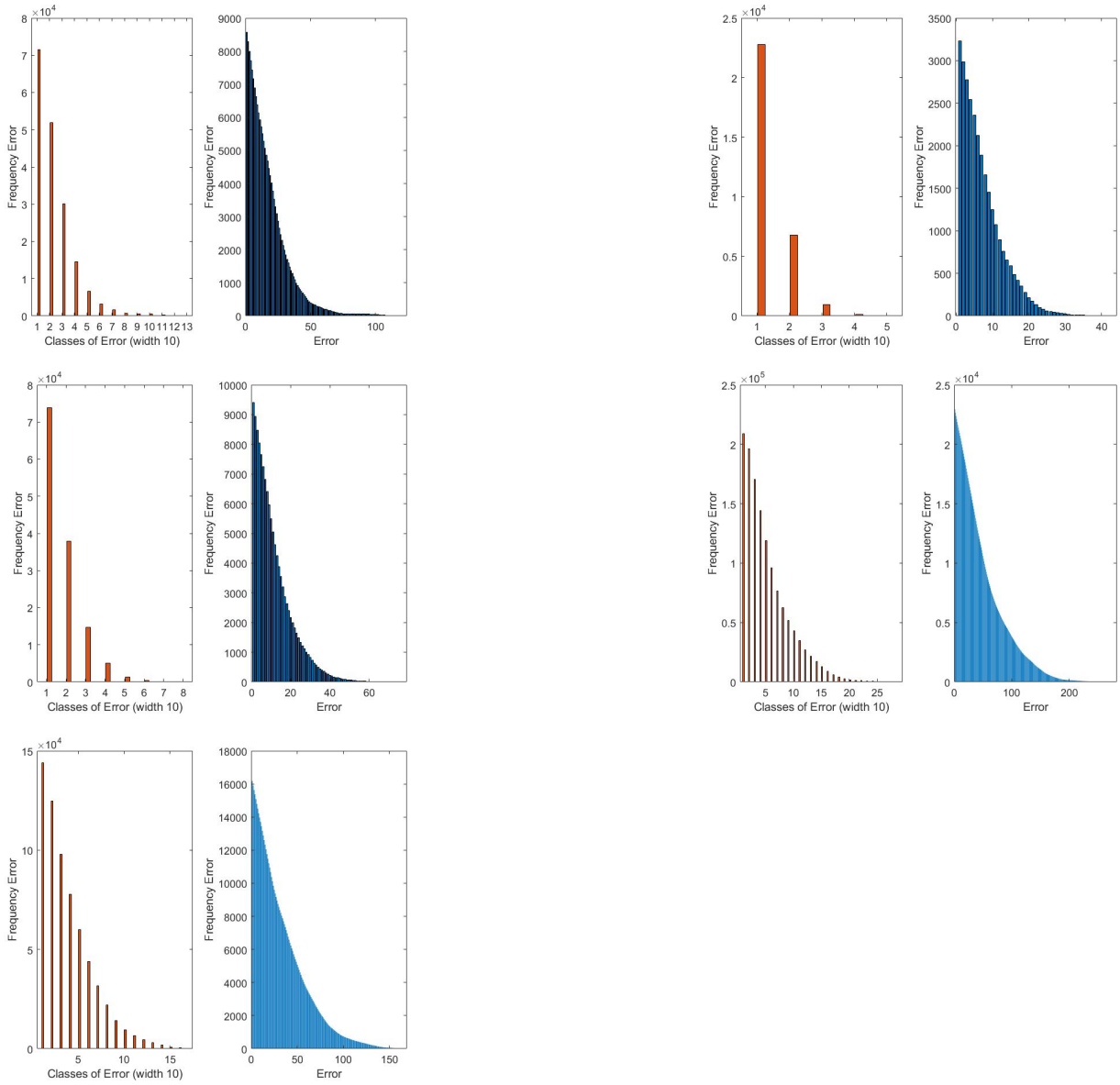

FIGURE 2.10: Frequency errors related to the study of Gini coefficient considering five initial weight distributions. [Top] (Left) Uniform distribution; (Right) Beta distribution with $\alpha = 1$ and $\beta = 3$. [Center] (Left) Beta distribution with $\alpha = \beta = 0.5$; (Right) Beta distribution with $\alpha = \beta = 2$. [Bottom] (Left) Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

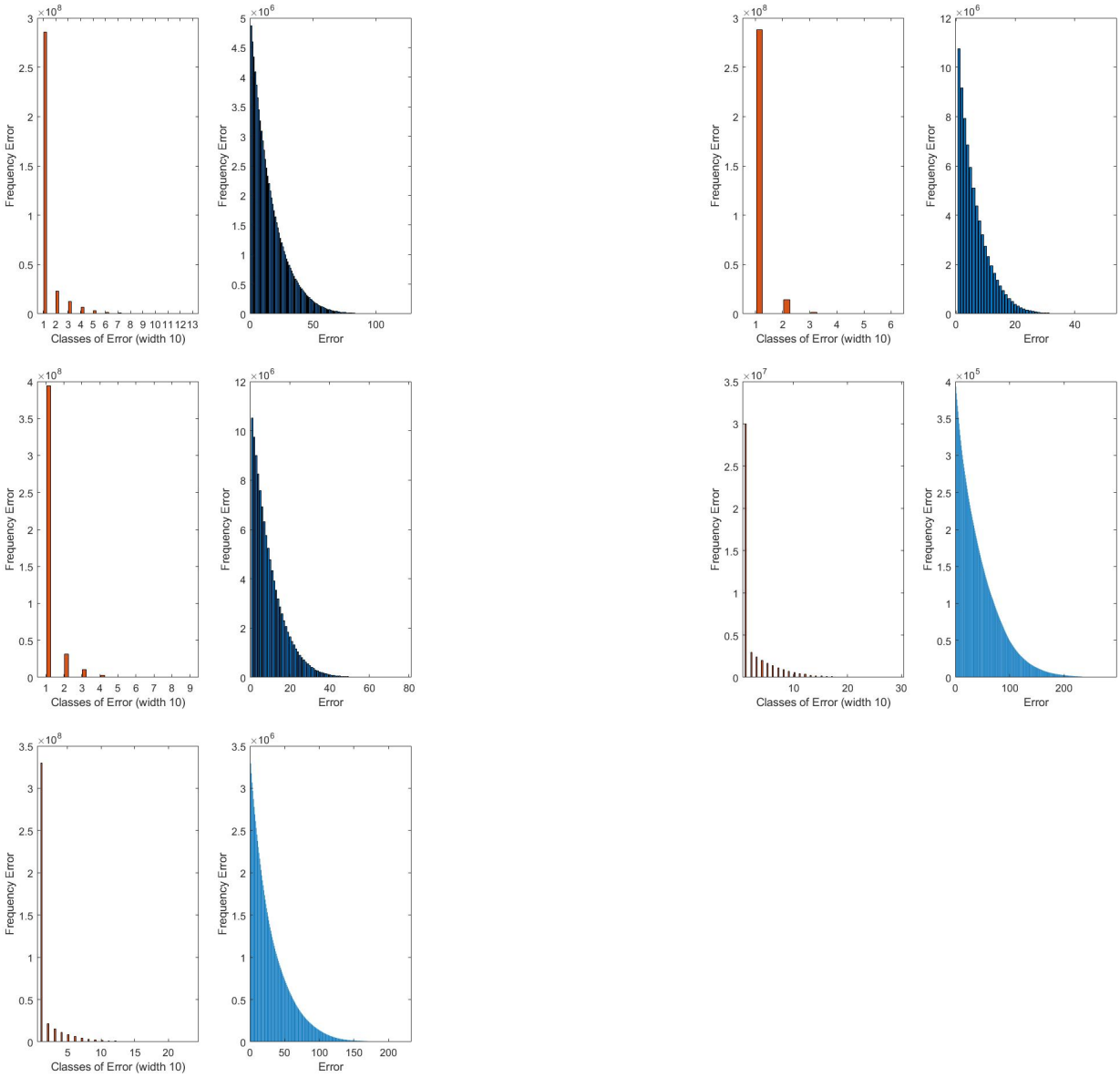

FIGURE 2.11: Histograms of the frequency of the prediction errors. Refer to the caption of Fig. 2.7 for a comprehensive explanation of Shannon entropy graphs.

manner as the error values increase (higher classes).

The predictive ability is not trivial because the error values with higher frequency are found when the error is close to zero and the trend of the curve falls very quickly.

It is evident that the errors corresponding to 0 and 1 (zero or almost zero errors) are the most frequent. This confirms the effectiveness of our analysis which provides a very consistent reliability rate with many correct predictions. The frequencies decrease dramatically in correspondence with high prediction errors.

From these graphs we can get an overview of the order of magnitude of these curves. In fact, there is an inversely proportional relationship between the frequency of errors (ordinate axis) and the error classes (abscissa axis).

The most dense and excessive errors are those concerning the symmetric Beta distributions with $\alpha = \beta = 2$ and the asymmetric one with $\alpha = 1$ and $\beta = 0.5$. In these cases the frequency of the lowest errors is much higher than other scenarios. Very often our predictions are correct. The high number of error classes, however, give us the perception of the high probability of making very large errors when the predictions are wrong.

Best predictive gain ever for Beta distribution with $\alpha = 1$ and $\beta = 3$. This confirms that the exponential negative distribution is the one that is most suitable in a context of forecasting failure times. And also with this aspect of the analysis we are consistent with the study of the first part which deals with each individual indicator in relation to errors in absolute value.

Chapter 3

The role of time in predicting the failure with conditioning on statistical measure percentiles

Also in this second model the aim is to verify the reliability of systems. We exploit rational expectations to carry out a study for the failure time prediction of stochastic systems that is directly dependent on time.

We consider two different sets of systems, *information set* systems and *in-vivo* systems, in order to exploit the information available to understand which real systems are very close to the cataloged one.

Here, we focus both on time and on the levels of the statistical tools that synthesize our systems.

The reliability of systems involves the passing of time and the increase in information provided up to the time we carry out the analysis.

In order to obtain the optimal prediction of the failure time of a system, we implement a comparative study between statistical measures and how they minimize the errors made in the predictions as a function of time.

We are witnessing an increase in the computational complexity of the procedure by turning the analysis as a function of time. This direct time dependence affects the forecasts of failure times of stochastic systems. We take into account the levels of

statistical indicators through three studies each conditioned to different percentiles of our data set.

The question we want to answer is: is the information we have about specific *in-vivo* systems useful for predicting when the system will fail?

And again: which error prediction conditioned on specific quantiles of the distribution of statistical indicators are closest to zero?

To provide an answer, we compare the errors made in the predictions through rational expectations, conditioning them to three percentiles of the distribution of the statistical indicator considered: 10^{th} percentile (low value), 50^{th} percentile (average value) and 90^{th} percentile (high value).

The 10^{th} percentile selects those realizations so that their largest cluster is smaller than 90% of the distribution values which have a bigger largest cluster. On the contrary, the 90^{th} percentile represents those values which have a big cluster to the point that only 10% of the realizations have a larger cluster. The 50^{th} percentile is instead placed at the center of this distribution.

These curves are compared with the benchmark (errors between the failure times of the *information set* and the failure times of the *in-vivo* systems).

The aim is to investigate which is the best combination (indicator-initial weights distribution) to achieve the optimal prediction. In this model we are therefore able to carry out an overall comparative analysis.

3.1 Highlights of the simulation algorithm steps

Also in this second application we want to examine the predictive power of the five summary measures used in the previous context, and specifically: (i) the variance of the α's; (ii) the kurtosis of the α's; (iii) the skewness of the α's; (iv) the Gini coefficient of the weights α's; (v) the Shannon entropy of the α's.

The simulation procedure will consist of seven phases:

1. We build the *information set* of systems. We denote by $K\_simulations$ the number of simulated systems that will be recorded in the *information set*,

so that the generic *information set* system will be denoted by $\mathbf{S}_k^I$, with $k = 1, \ldots, K\_simulations$.

Each system is created from the beginning – where all the components are active – to its failure. Components fail according to the procedure described in Subsection 1.6. For each system $\mathbf{S}_k^I$ we collect the information by one of the five statistical indicators (i)-(v) presented above related to the specific observation $\bar{\mathbf{a}}(t) = (\alpha_1(t), \ldots, \alpha_n(t))$ – where the reference to t indicates that the observation is considered at time t. Here, we introduce the failure time $\mathcal{FT}_k$ of the related system – where the index k points to the k-th system $\mathbf{S}_k^I$ of the *information set*; the number of active components $n_k(t)$. We will follow the systems from the initial time until their failure with the time $t = 0, 1, 2, \ldots, \mathcal{FT}_k$.

Keeping the notation introduced in the previous application, we will collect for each k-th system:

- All the times until the failure $\mathcal{FT}_k$ with $t = 0, 1, 2, \ldots, \mathcal{FT}_k$.
- For each time t:
  - We record a set of α's of the observations $\bar{\mathbf{a}}$ that depends on k, t and $n_k(t)$ (all the weights of the $\mathbf{S}^I$'s);
  - Each set of α's is associated with one value $\bar{\star}_I$ of the statistical indicator represented by $\star$, with $\star$ =variance, kurtosis, skewness, Gini coefficient and Shannon entropy. Any observed value $\bar{\star}_I$ depends on k and t, while $n_k(t)$ intervenes directly in the computation of the considered statistical indicator at time t and for system k. We collect all values $\bar{\star}_I$'s in a matrix ($\mathbf{I}^\star$).
- All the failure times $\mathcal{FT}$'s of systems $\mathbf{S}^I$'s in a vector $\mathcal{FT}^I$ that depends on the k-th system $\mathbf{S}_k^I$ and indicated with $\mathcal{FT}_k$.

2. At this point we follow the same procedure and we simulate a new series of systems which we will call *in-vivo* systems (namely $\mathbf{S}^V$'s). These new systems will be necessary for the calculation of rational expectations. We denote by

X_simulations the number of simulated *in-vivo* systems, and the generic *in-vivo* system is $\mathbf{S}_x^V$, with $x = 1, \ldots, X\_simulations$.

In this case, by using opportune notation, introduced in the previous context, we will record for each x-th system:

- All the times until the failure $\mathcal{FT}_x$ with $t = 0, 1, 2, \ldots, \mathcal{FT}_x$, where $\mathcal{FT}_x$ is the failure time of the related system – the index x points to the x-th system $\mathbf{S}_k^V$ of the *in-vivo*.
- For each time t:
  - We record a set of α's of the observations $\bar{\mathbf{a}}$ that depends on x, t and $n_x(t)$ (all the weights of the $\mathbf{S}^V$'s), where $n_x(t)$ is the number of active components.
  - Each set of α's is associated with one value $\bar{\star}_V$ of the statistical indicator represented by $\star$, with $\star$ =variance, kurtosis, skewness, Gini coefficient and Shannon entropy. Any observed value $\bar{\star}_V$ depends on x and t. We collect all values $\bar{\star}_V$'s in a matrix ($\mathbf{V}^\star$).
- All the failure times $\mathcal{FT}$'s of systems $\mathbf{S}^V$'s, for each x, in a vector $\mathcal{FT}^V$ that depends on the x-th system $\mathbf{S}_x^V$ (indicated with $\mathcal{FT}_x$).

Here, we consider all the values $\bar{\star}_V$'s. These values are different a priori from the $\bar{\star}_I$'s found in the first part of the simulation procedure.

3. Now, we compute the rational expectations in (1.9) on the basis of the $\mathbf{S}^I$'s.

   - Check the tolerance threshold condition (*Condition*).

     We use the $\mathbf{S}^I$'s for the prediction of failure times of the $\mathbf{S}^V$'s, applying a condition considering a tolerance threshold for rational expectations. Specifically, we look at the generic statistical indicator level $\bar{*}_I \in \mathbf{I}^\star$ at time $\bar{t}$ such that the following *Condition* holds

$$|\mathbf{I}^\star(\bar{t}) - \bar{*}_V(\bar{t})| < T \tag{3.1}$$

where T is the tolerance level entered in the initial setup.

We check the *Condition* for each observed value $\bar{*}_V$ with respect to all the $\bar{*}_I$'s $\in \mathbf{I}^\star$ at the specific time $\bar{t}$.

- We store all the $\mathcal{FT}^I$'s of the systems associated to $\mathbf{I}^\star$, i.e. satisfying *Condition* depending on time $\bar{t}$, disregarding the other systems.
- We compute the arithmetic mean of each set of $\mathcal{FT}^I$'s in correspondence of any time $\bar{t}$, namely $\mathbb{E}[\mathcal{FT}_{RE}(\bar{\star}_V)]$. In so doing, we obtain the rational expectations of formula (1.9) associated to each $\bar{\star}_V$.

4. In order to assess the quality of such predictions, we focus our attention on fixed percentiles $p^\star$ of the distribution of $\mathbf{I}^\star$ maintaining the dependence on time with $t = 0, 1, 2, \ldots, \mathcal{FT}_x$.

   Specifically, the percentiles are: $p^\star = 10\%$, $p^\star = 50\%$ and $p^\star = 90\%$.

   For the prediction, we identify all the $\mathbf{S}^V$'s that, at each time $\bar{t}$, belong to the observed $p^\star$.

   To do this, we look for all the $\bar{\star}_V(\bar{t})$ which are considered the "same observation" with respect to the specific $p^\star$ (time dependent).

   The "same observation" is each observed value $\bar{\star}_V(\bar{t})$ which is closest to the value $p^\star(\bar{t})$ considering a tolerance distance which is linked to the statistical measure analysed. In particular, 80% of the standard deviation of the $\mathbf{V}^\star$.

   Specifically, we compute the distance between each $\bar{*}_V(\bar{t})$ and the fixed $p^\star(\bar{t})$:

$$|\bar{*}_V(\bar{t}) - p^\star(\bar{t})| < 0.8 * std(\mathbf{V}^\star) \tag{3.2}$$

   Once we find the closest observations of $p^\star(\bar{t})$ among the $\bar{\star}_V$'s, we assign them in a matrix namely $\bar{\star}_V|p^\star$ that depends on x and t which contains the set of $\mathbf{S}^V$'s belonging to the reference percentile $p^\star$.

5. We compute the average prediction error conditioned on the $p^\star$ between the rational expectations $\mathbb{E}[\mathcal{FT}_{RE}(\bar{\star}_V)]$ and the $\mathcal{FT}^V$'s connected to the $\bar{\star}_V|p^\star$.

Specifically,

$$E_{RE}|p^{\star} = \mathbb{E}[|\mathbb{E}[\mathcal{FT}_{RE}(\bar{\star}_V)] - \mathcal{FT}_x|\bar{\star}_V|p^{\star}|] \tag{3.3}$$

Each $E_{RE}|p^{\star}$ depends on time t.

Finally, we obtain three different distributions of errors conditioned to the percentiles:

- $E_{RE}|10\%$
- $E_{RE}|50\%$
- $E_{RE}|90\%$

We create, for each $\bar{\star}_I$, a distribution of errors between the expected failure times calculated with the rational expectations and the failure times of the *in-vivo* systems.

6. To have the possibility of quantifying the impact of the rational expectation predictions, we now assess a naive benchmark error given by the errors made without the use of rational expectations, as follows

$$E_B = \mathbb{E}[|\mathbb{E}[\mathcal{FT}^I] - \mathcal{FT}_x|] \tag{3.4}$$

For comparison purposes, we normalize all the times t with $t = 0, 1, 2, \ldots, \mathcal{FT}_x$.

In this way, we quantify the impact of the error prediction comparing E_B, $E_{RE}|10\%$, $E_{RE}|50\%$ and $E_{RE}|90\%$ for the several indicators considered in the analysis (variance, kurtosis, skewness, Gini coefficient and Shannon entropy).

7. To obtain greater comprehensibility in the interpretation of the results, proceed with the graphic representation to check which conditioning on the three percentiles is the optimal prediction of failure times using the information stored in the past.

It is necessary to examine the comparison between $E_{RE}|10\%$, $E_{RE}|50\%$, $E_{RE}|90\%$ and E_B.

3.2 Setting of the simulation algorithm

Given the increasing complexity of the simulation algorithm compared to the previous model, the number of simulations is reduced for both the *information set* and *in-vivo* systems.

The parameters are those indicated in the following table:

TABLE 3.1: Parameter set

Parameter	**Value**
number of components	10
$K\_simulations$	5.000
$X\_simulations$	5.000

As already stated above in Subsection 2.1.2 , we extract the α's and r from different types of distributions which provide five possible cases for each analysed characteristic of the configurations (see Fig. 2.1).

Specifically:

- Uniform distribution in (0,1) type;
- Four Beta distribution type:
  - $\alpha = 1$ and $\beta = 3$;
  - $\alpha = \beta = 0.5$;
  - $\alpha = \beta = 2$
  - $\alpha = 1$ and $\beta = 0.5$;

The value assigned to the *tolerance level* will depend on the specific case we want to simulate. We evaluated three different *tolerance level*:

1. $tol_1 = 0.005$;

2. $tol_2 = 0.05$;

3. $tol_3 = 0.5$.

In the following results not all *tolerance level* will be shown but only the case $tol_2 = 0.05$ because it is the case in which the statistics seem more stable.

However, they are available upon request.

We therefore carried out five types of analysis according to five different characteristics of the α's distribution: variance, kurtosis, skewness, Gini coefficient and Shannon entropy.

Results and discussion are presented in the next section.

3.3 Results and discussion

The graphs for each single statistical predictor will be initially discussed according to the different conditioning, to conclude with a comparative analysis on which one represents the best prediction tool for stochastic systems in a context of rational expectations.

For each statistical indicator we design five different scenarios based on five initial distributions of component weights in order to understand which distribution is best suited to this analysis.

In this model we also evaluate which of the initial weight distributions proposed is the best performing one.

The exponential negative distribution is the best when dealing with problems related to the evaluation of the failure time of systems with interconnected components. As we will see in the following results, also for this second framework the consistency with literature is confirmed (see Li and Zuo [41], Sarhan [42], Asadi and Bayramoglu [43] and Eryilmaz [44]).

Now we move on to the individual analysis of variance, kurtosis, skewness, Gini coefficient and Shannon entropy.

A general comment regarding almost all the results linked to this model, concerns time. In fact, in the time horizon that goes from the starting time to the normalized failure time of the systems, we see an evident predictive gain depending on the case examined. We appreciate an improvement in percentile-conditioned error curves as available information increases.

The first analysis is through the study of the behavior of variance. Variance is a measure of statistical dispersion of the realizations depending on time.

We then go on to capture the variability of the component weights, and how much they fluctuate with respect to the average values of the time-to-time distribution.

Figs. 3.1, 3.2, 3.3, 3.4 and 3.5 show the prediction errors (in absolute value) for $p = 10\%$, $p = 50\%$ and $p = 90\%$ about the analysis of the variance, considering all different five types of initial weight distribution (Uniform distribution, Beta distribution with $\alpha = 1$ and $\beta = 3$, Beta distribution with $\alpha = \beta = 0.5$, Beta distribution with $\alpha = \beta = 0.5$ and Beta distribution with $\alpha = 1$ and $\beta = 0.5$, respectively) and $tol_2 = 0.05$. Green triangles correspond to to the 90^{th} percentile of the *information set*. Blue crosses represent the errors conditioned to the 10^{th} percentile of the *information set* and red dots are the errors conditioned to the 50^{th} percentile of the *information set*. The fixed magenta line is the benchmark.

Through the different tolerances we want to discover the predictive gain or loss depending on the cluster on which we calculate the rational expectations: small tolerance means that we compare very similar values, while the increase in tolerance leads us to an analysis between an ever wider range of values.

Looking at the graphs as a whole (and this remark can be extended to subsequent results with some exceptions included), the first thing that stands out is that the four curves coincide for small t. Therefore, both the benchmark and the various conditional error curves have an equal or very similar starting level.

The reason is that in the initial trends no system has yet failed and we still have no information. So, there are no substantial differences in the absence of information in the prediction of failure time of stochastic systems and the errors converge to the same point.

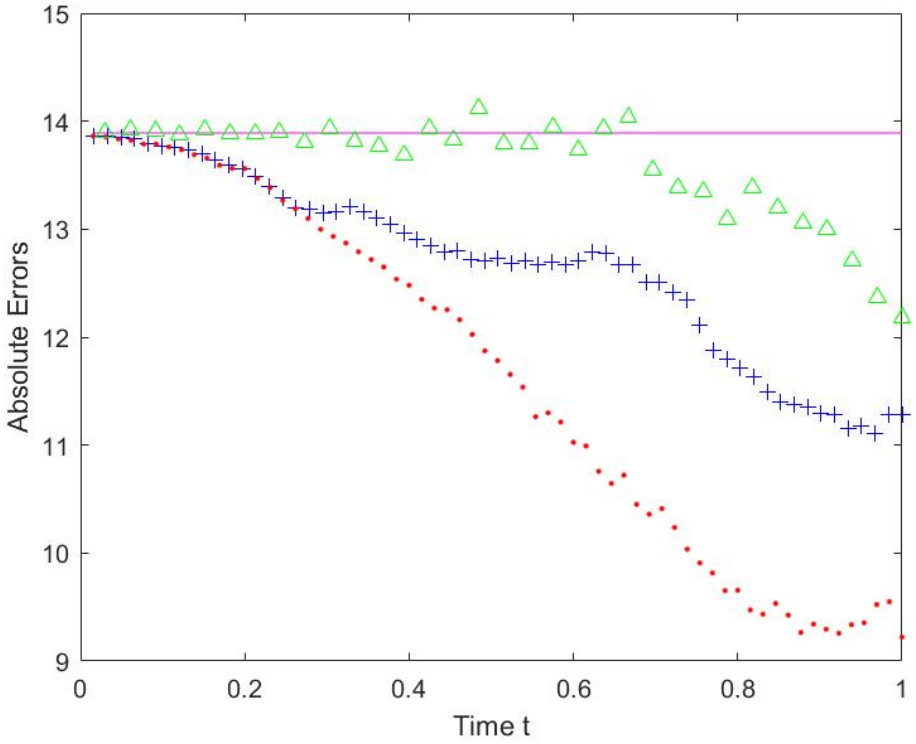

FIGURE 3.1: Time evolution of prediction errors: case of variance and uniform distribution.

For larger t the information stored increased and we can appreciate a deviation between the benchmark and errors conditioned by the percentiles which showcase predictive gain linked to the use of information.

In general, these graphs confirm the predictive power of variance and the gain in prediction accuracy, both when we condition the forecast on the $p = 10\%$, $p = 50\%$ and $p = 90\%$ of the variance distribution.

The value assumed by the variance has a great influence on the errors made in our failure time forecast and is in turn related to the tolerance level used in the initial setup.

By focusing on the initial distribution of weights (on which the failure of the system depends) we can assert that the results are in line with what has been stated by various researchers over time (see Li and Zuo [41], Sarhan [42], Asadi and Bayramoglu [43] and Eryilmaz [44]).

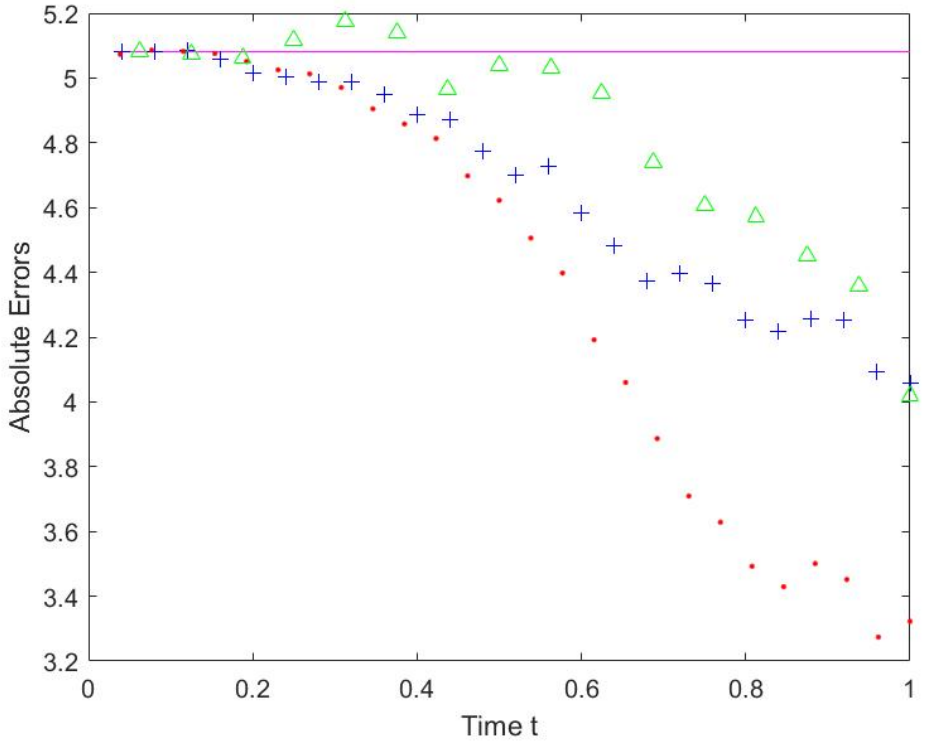

FIGURE 3.2: Time evolution of prediction errors: case of variance and Beta distribution with $\alpha = 1$ and $\beta = 3$.

The asymmetric Beta distribution with $\alpha = 1$ and $\beta = 3$ (trend similar to the negative exponential one) is the ideal distribution that allows us to minimize errors in terms of absolute value and the errors are much more delayed in time. We also get good results with the symmetric distribution with $\alpha = \beta = 0.5$ and the Uniform one.

The situation gets worse in the asymmetric case with $\alpha = 1$ and $\beta = 0.5$ and is even more drastic in the symmetric one with $\alpha = \beta = 0.5$. The errors in absolute value increase exponentially compared to the previous cases and the curves are much more irregular and scattered. Although in these outcomes there seems to be a predictive improvement linked to time, we must nevertheless prefer the previous distributions both in terms of levels of error and of regularity in the conditioned paths.

So, after the initial part of the trends, where the information we have is very random so there is no gain in the use of rational expectations, the variance proves to be a good prediction tool. Accuracy in performance depends on various factors that

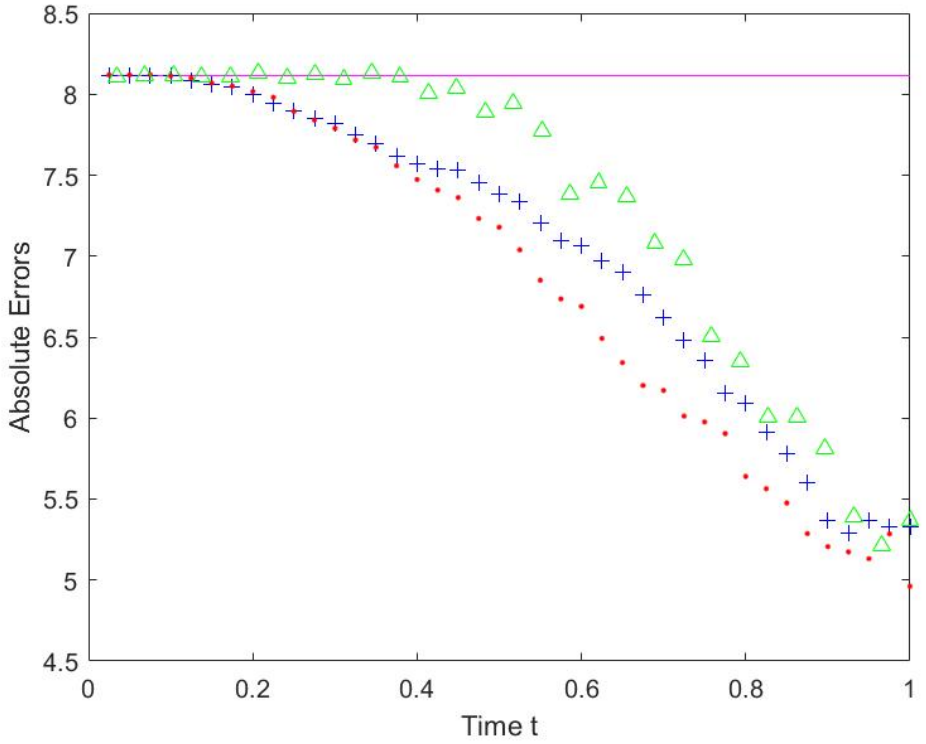

FIGURE 3.3: Time evolution of prediction errors: case of variance and Beta distribution with $\alpha = \beta = 0.5$.

we include as inputs into the analysis, such as tolerance level or initial distribution. In the context of prediction models, especially when the data sets being analyzed are not very large (see Parzen ([89]) and Jiang et al. ([90])), a low variance is considered more stable and reliable while a high variance is a sign of instability.

In our analysis, however, the results are mixed and there is no uniqueness between the various cases. And what information to take into account when predicting failure times depends on the tolerance level applied.

Here, we show the results of the second indicator used as a predictive tool: kurtosis.

With kurtosis we want to investigate the existence of extreme values between the weights of the components of the systems, determining the heaviness of the distribution tails.

As already seen in the case of variance, Figs. 3.6, 3.7, 3.8, 3.9 and 3.10 show the prediction errors of kurtosis conditioned to the same percentiles for $p = 10\%$, $p = 50\%$

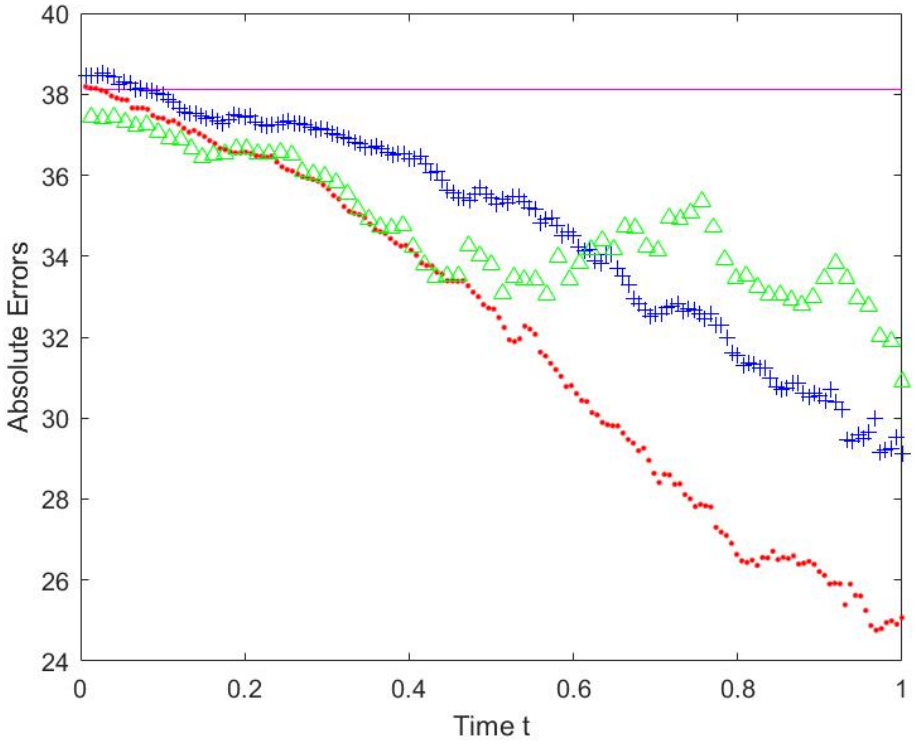

FIGURE 3.4: Time evolution of prediction errors: case of variance and Beta distribution with $\alpha = \beta = 2$.

and $p = 90\%$. Refer to the same details as the variance figures.

Generally, high values of kurtosis perform worse than low values.

It is intuitive that, since kurtosis is a risk index, if we obtain very high values (as in the case of conditioning with respect to $p = 90\%$ of the distribution), they are joint to the increasing probability of obtaining components that are very different from each other and with an importance (in terms of probable system failure) polarized towards extreme values (few components with very high weights and with a high probability of failure). We therefore suffer in terms of predictive gain.

Observing the evolution of our systems from time $t = 0$ to their failure in the case ok kurtosis, there is an initial flat behavior that depends on the fact that when we have little information available, the initial trend is dominated by the case.

As in the analysis of variance, the initial distribution that performs best in terms of forecasting is the Beta distribution with $\alpha = 1$ and $\beta = 3$, just as stated by other

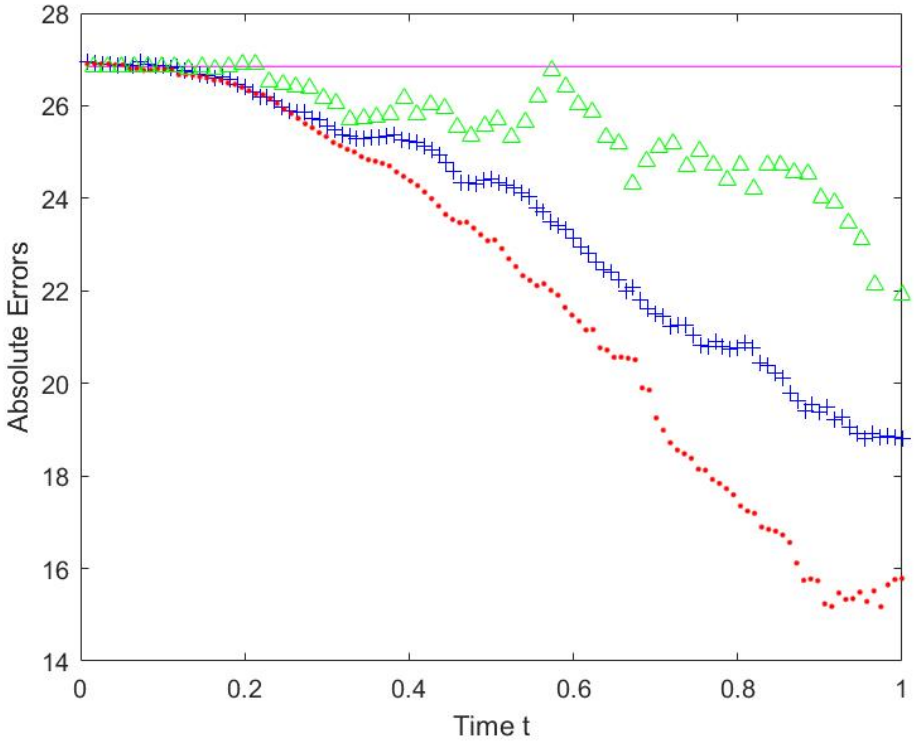

FIGURE 3.5: Time evolution of prediction errors: case of variance and Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

studies. Even the uniform distribution and the Beta one with $\alpha = \beta = 0.5$ are very robust in this context, while the two remaining distributions are not reliable, reaching very high error values and fluctuating and irregular trends throughout the path (especially for the line of errors conditioned on the 90^{th} percentile of the *information set*).

In view of the statement above, the kurtosis results are noticeable for medium and low values therefore in correspondence with mesokurtic and platykurtic weight distributions, while our model loses efficacy for leptokurtic weight distributions.

We now move on to the third moment of our weight distributions: skewness. How does skewness affect prediction modeling of stochastic systems?

The graphs relating to skewness analysis are illustrated below. Fig. 3.11 shows the Uniform distribution. Fig 3.12 instead contains the results about Beta distribution with $\alpha = 1$ and $\beta = 3$. In Figs. 3.13, 3.14 we can see the two symmetric Beta

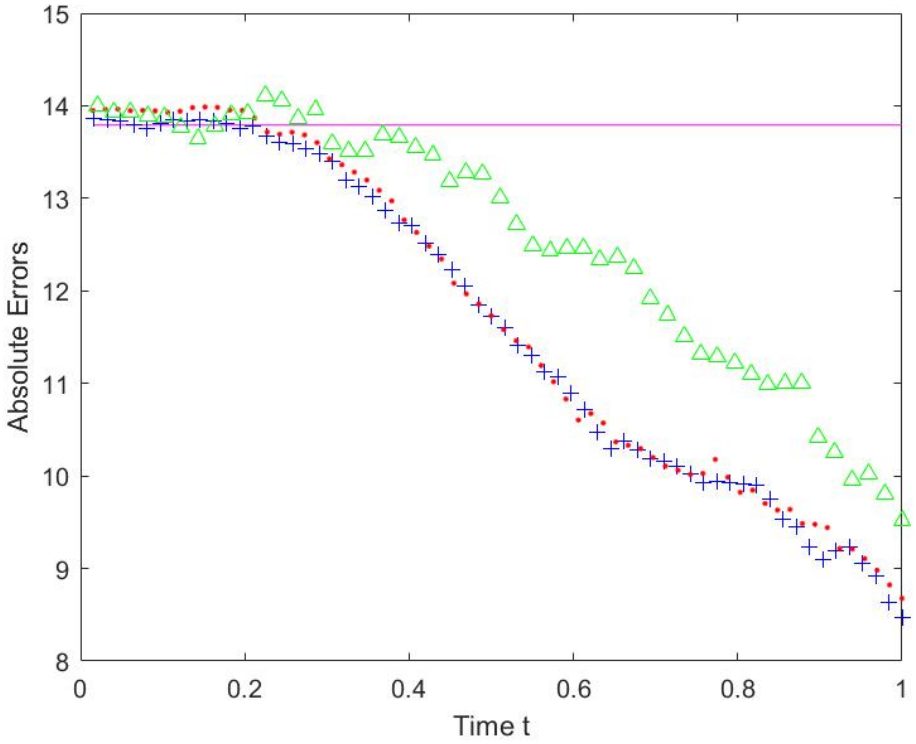

FIGURE 3.6: Time evolution of prediction errors: case of kurtosis and uniform distribution.

distribution with $\alpha = \beta = 0.5$ and $\alpha = \beta = 2$, respectively. Fig. 3.15 displays Beta distribution with $\alpha = 1$ and $\beta = 0.5$. We condition the prediction errors to to three percentiles of the statistical indicators $p = 10\%$, $p = 50\%$ and $p = 90\%$. The *tolerance level* correponds to $tol_2 = 0.05$. The fixed magenta line represents the benchmark. Green triangles, Blue crosses and red dots ere the 90^{th}, 10^{th} and 50^{th} percentiles of the *information set*, respectively.

In literature, several studies have been carried out in the field of reliability theory – specifically, k-out-of-n systems – which exploit the various statistical moments to validate their models. We refer for example to Reijns and Gemund ([84]) or Amari et al. ([85]).

According to these authors, lower moments are more robust than higher moments.

We have already seen how kurtosis (fourth statistical moment) is however a good predictor indicator. We will find out later which of all indicators is the best.

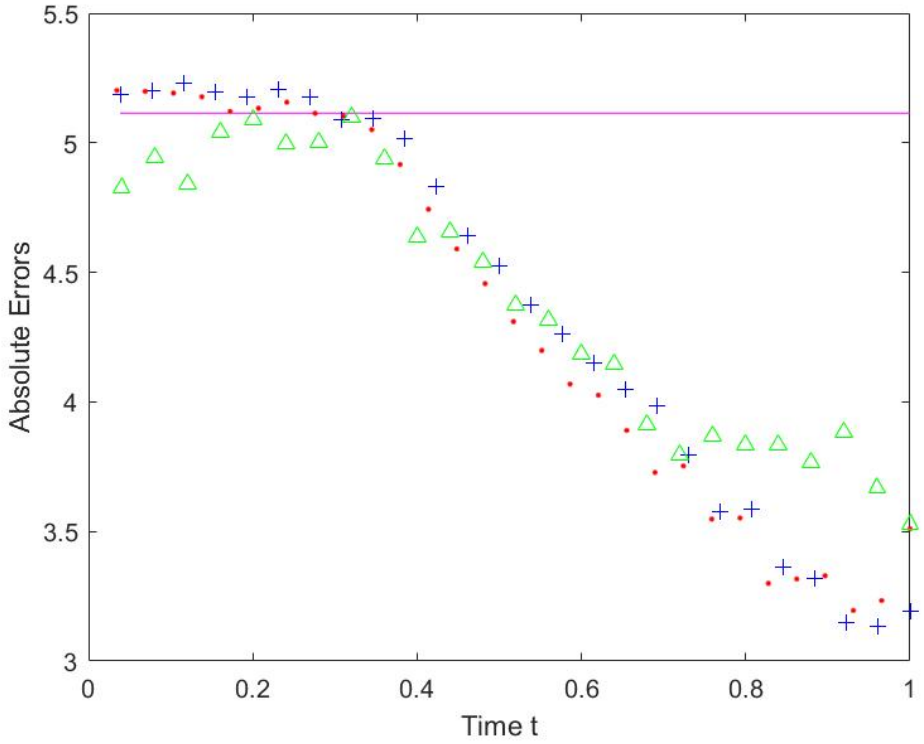

FIGURE 3.7: Time evolution of prediction errors: case of kurtosis and Beta distribution with $\alpha = 1$ and $\beta = 3$.

The values of skewness are sensitive to fluctuations (we should appreciate a loss of precision that affects our predictive model).

For some distributions (particularly for Beta with $\alpha = 1$ and $\beta = 3$ and the symmetric one with $\alpha = \beta = 0.5$ and Uniform distributions), skewness enables us to improve our predictions over time. There is no obvious difference between the various conditionings, so having large (positive) skewness, low (negative) or medium values does not affect the accuracy of the analysis.

Therefore, for the purposes of this application and in the three scenarios mentioned above, we are not interested in knowing whether the weight distribution of stochastic systems is symmetric or asymmetric towards the right or left tail. The behaviors of the three curves are in fact quite regular and constantly decreasing.

If instead we move to the Figs. 3.14 and 3.15 (Beta distribution with $\alpha = \beta = 2$ and with $\alpha = 1$ and $\beta = 0.5$), the analysis immediately appears much more imprecise

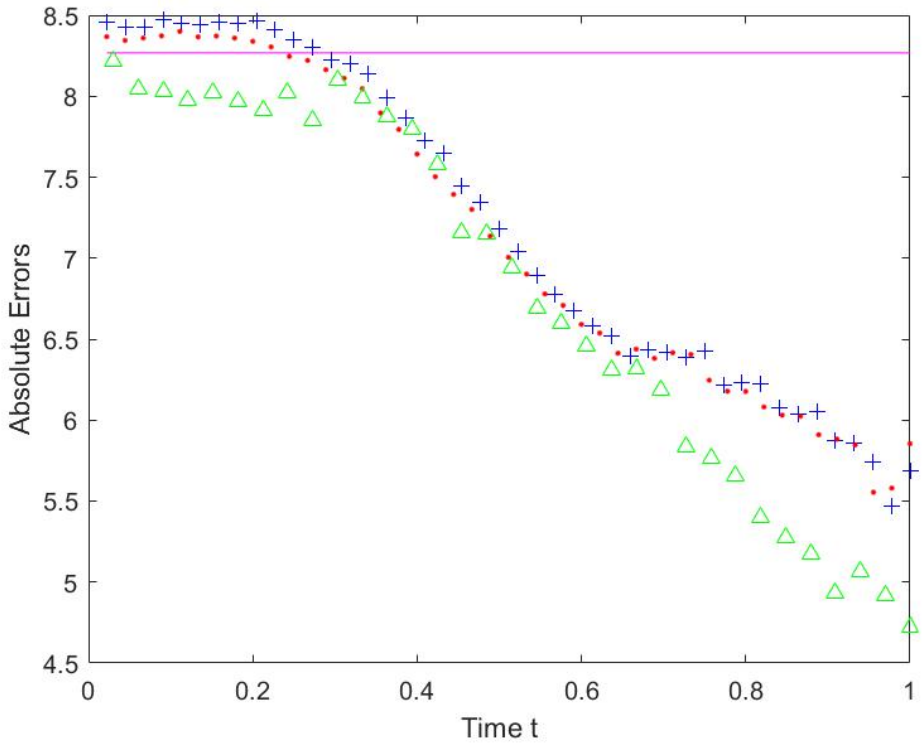

FIGURE 3.8: Time evolution of prediction errors: case of kurtosis and Beta distribution with $\alpha = \beta = 0.5$.

and not very useful.

For symmetric or low asymmetric datasets the trend remains good, while when the distributions move towards the tails and the values become abnormal upwards or downwards, the predictive capacity of skewness is lost and these values are not within the valid range for the reliability values. The results are penalized in term of tendency and errors.

We now pass to a concentration coefficient which is the Gini coefficient.

The use of the Gini coefficient as a performance measure is not usually used in failure prediction models. However, according to the research of Ooghe and Spaenjers ([88]) it is a very powerful and attractive tool in our field of study.

In fact, it is able to capture the total power of the model considering all weights without overestimating or underestimating any of them.

Refer to the details of the graphs stated above for the previous statistical indicators

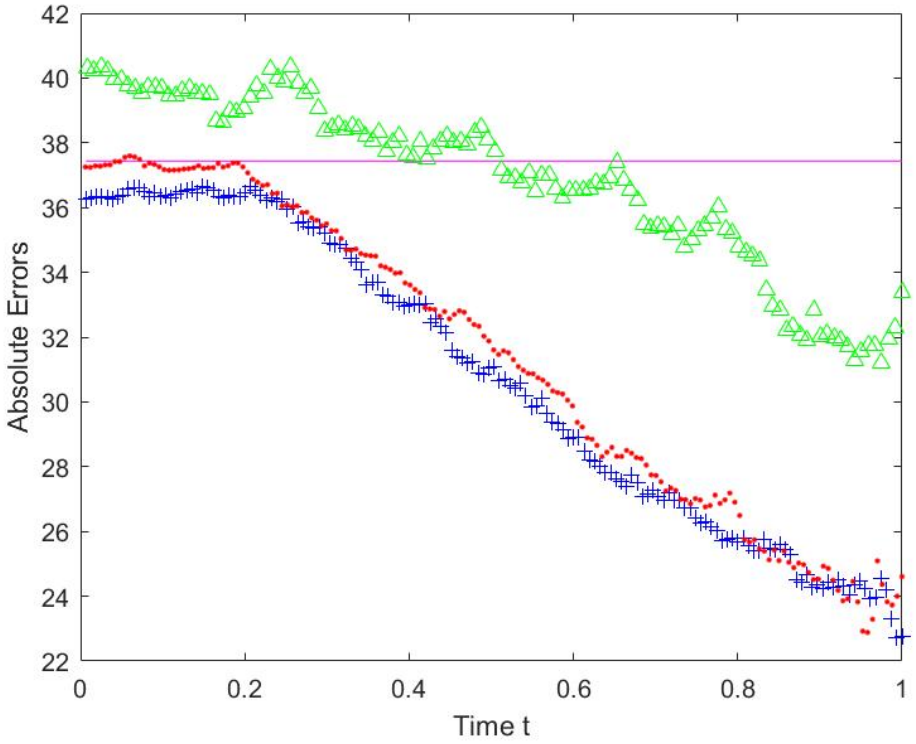

FIGURE 3.9: Time evolution of prediction errors: case of kurtosis and Beta distribution with $\alpha = \beta = 2$.

for Figs. 3.16, 3.17, 3.18, 3.19 and 3.20. In these five figures are exhibited the prediction errors of Gini coefficient.

The result obtained is truly consistent.

This indicator provides the predictive ability of rational expectations that impoves with the reaching of the time of failure.

Regarding the several perspectives linked to the initial weight distributions, the situation is very similar to the previous analysis. We are always in line with the existing literature, which confirms the predominance of the exponential negative distribution (Beta distribution with $\alpha = 1$ and $\beta = 3$).

It is surprising how the best result is linked to very polarized distributions with high weights more concentrated on a few components. And this makes sense because, since the failure of the system is linked to the number of components that stop working, the probability of extracting a component with a very high weight (therefore

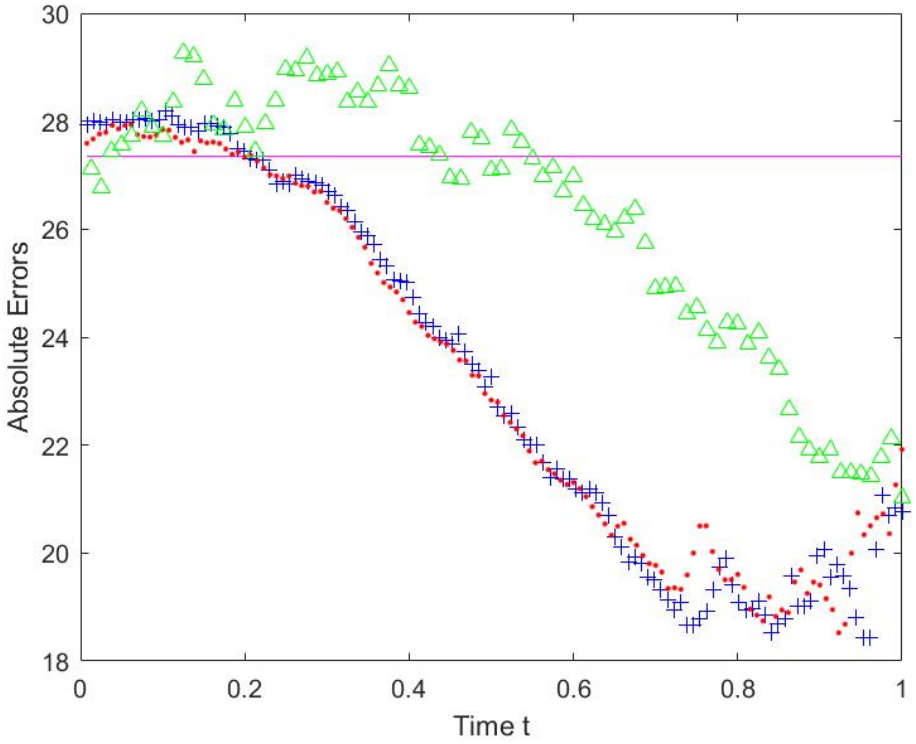

FIGURE 3.10: Time evolution of prediction errors: case of kurtosis and Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

more sensitive to possible failure) is much lower. This supports the reliability of the system. It must also be said, however, that components with very high weights are more vulnerable and contribute to our uncertainty prediction process that does not seem to affect the results of errors conditioning on the 90^{th} percentile.

While the error curve relating to the high values (and medium values) of the Gini coefficient decreases monotonically, when the weights are equally distributed (low values of the coefficient), the trends are in any case decreasing but more scattered and characterized by less linear regularity.

Gini coefficient stands out as an excellent predictor.

The last indicator to be analyzed is Shannon entropy.

Also in this case the prediction errors are presented in the Figs. 3.21, 3.22, 3.23, 3.24 and 3.25. We always consider all the different initial weight distribution type (Uniform distribution, Beta distribution with $\alpha = 1$ and $\beta = 3$, Beta distribution with

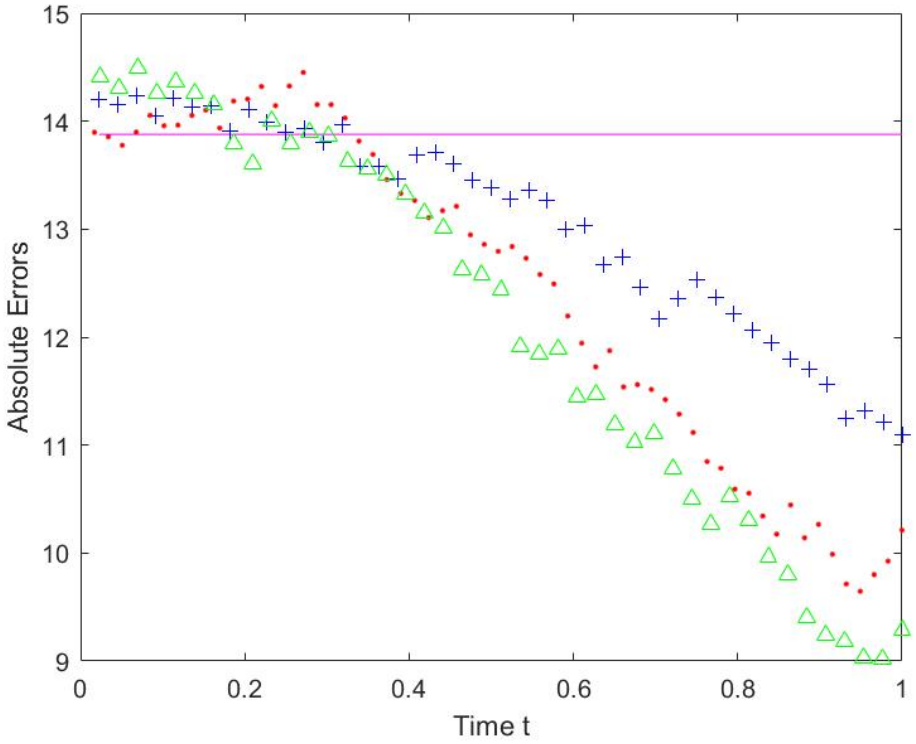

FIGURE 3.11: Time evolution of prediction errors: case of skewness and uniform distribution.

$\alpha = \beta = 0.5$, Beta distribution with $\alpha = \beta = 0.5$ and Beta distribution with $\alpha = 1$ and $\beta = 0.5$, respectively). As already seen for the previous statistical measures we represent three conditionings depicted in the graphs by the three curves: blue crosses indicate the 10^{th} percentile, red dots indicate the 50^{th} percentile and green triangles indicate the 90^{th} percentile. The fixed magenta line is the benchmark.

Shannon entropy is a quantitative measurement of information which captures the amount of information related to the reference random processes.

There is a lack of papers in literature using Shannon entropy in predictive models.

It is exploited in computational mechanics, materials science, and physics to investigate the fatigue lifetime problem (see Zhang et al. [91]) and in the field of business failure prediction (see Bal et al. [92]).

We try to place it in this context and subsequently to compare it with the other statistical indicators. The amount of information gained taking into account the

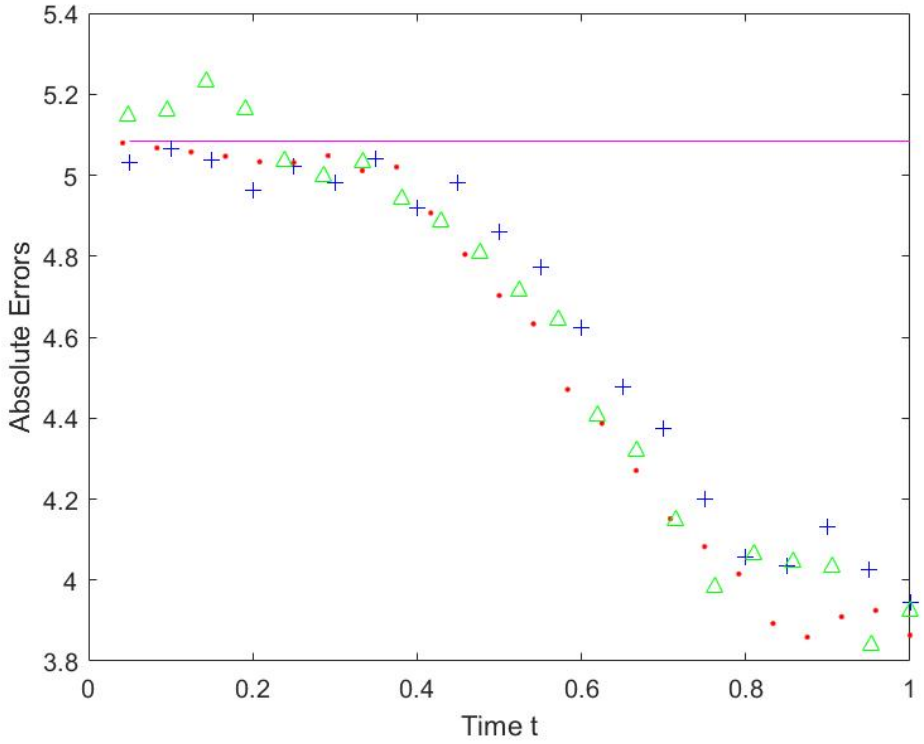

FIGURE 3.12: Time evolution of prediction errors: case of skewness and Beta distribution with $\alpha = 1$ and $\beta = 3$.

Shannon entropy is very useful.

According to the passage of time, the results shown in the five scenarios are the outcomes of a good prediction. In fact, we notice an initial phase of the curves that is flat and very close to the benchmark. Indeed, if we are very far from the reasonable level of information, then the path is quite random (the use of rational expectations still does not come out).

The predictive gain grows exponentially as well as the knowledge gained through the information stored with the use of rational expectations increases. The increase in information is linked to the passage of time.

We can assert that the results are solid for high values of Shannon entropy and the probability of predicting more efficiently increases. In fact, the curve of the errors conditioning on the $p = 90\%$ of the *information set* distribution, is the one that performs better and decreases faster than the low values trend that show a less steep

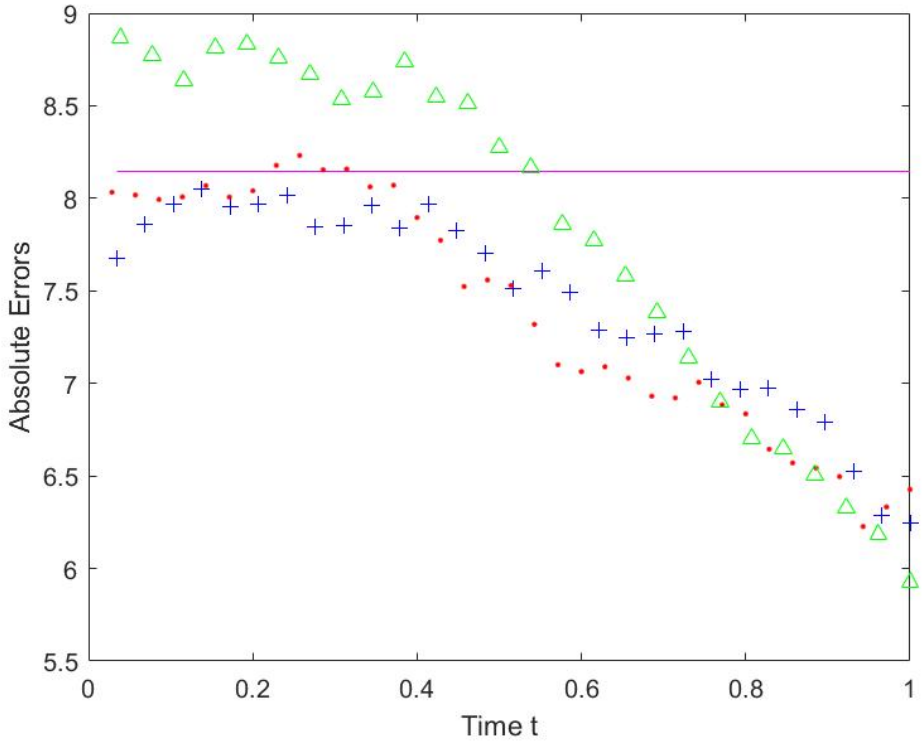

FIGURE 3.13: Time evolution of prediction errors: case of skewness and Beta distribution with $\alpha = \beta = 0.5$.

and less predictive path.

Overall, entropy is a very important factor in understanding stochastic systems prediction model.

To better understand which statistical indicator provides the best prediction and which the trend that minimizes errors depending on the percentiles is, a comparison between the statistical indicator and the percentiles $p = 10\%$, $p = 50\%$ and $p = 90\%$ is now necessary. In this analysis we take into account the different distributions of the initial weights.

Over time, many studies have been collected on the large number of different prediction accuracy statistics. To review what the literature says, in measurements and in predictive models, lower moments are more robust and consistent than higher moments (see Reijns and Gemund [84], Amari et al. [85], Ramberg et al. [86], Kinateder and Papavassiliou [87]). These researchers validate the use of moments of

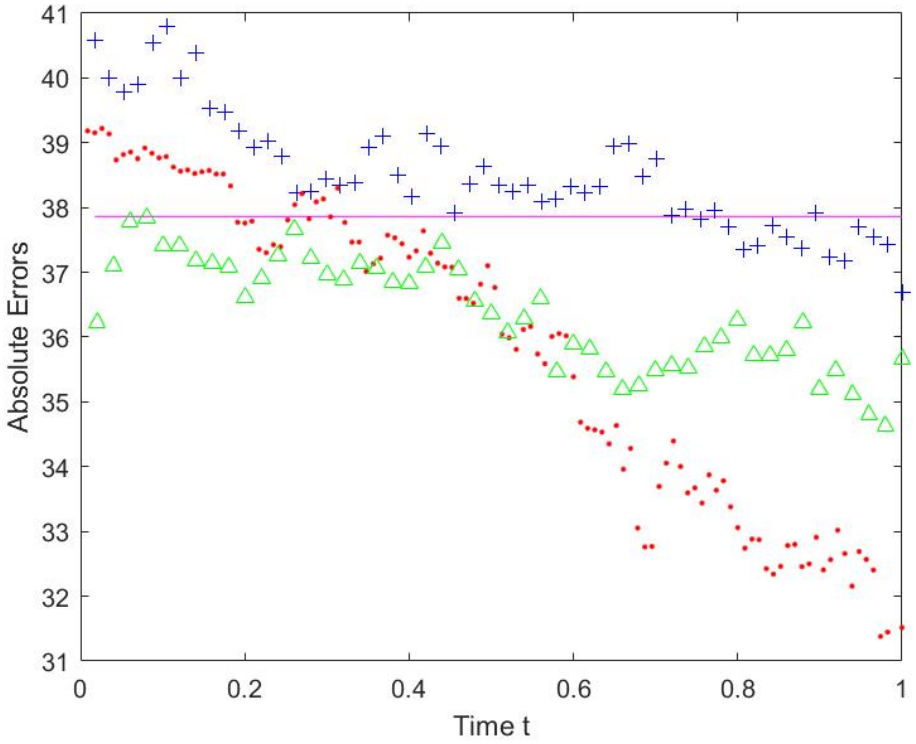

FIGURE 3.14: Time evolution of prediction errors: case of skewness and Beta distribution with $\alpha = \beta = 2$.

distributions for quantitative analysis.

Here we want to investigate the relationship between our measures of interest in order to strengthen the results obtained.

Therefore, in addition to the second, third and fourth statistical moments (variance, skewness and kurtosis), we have also added the Gini coefficient (little used in these research fields) and Shannon entropy (never used in this type of applications).

Studied individually, each indicator is imperfect and incomplete because we do not have access to all the information available, while with this comparison we can get an idea of which the best prediction tool is, depending on the distribution assumed by the weights of the components.

The curves show, as a function of time t, the behavior of the statistical measures considered according to the conditioning.

In principle, the results of the several scenarios in terms of the initial errors match

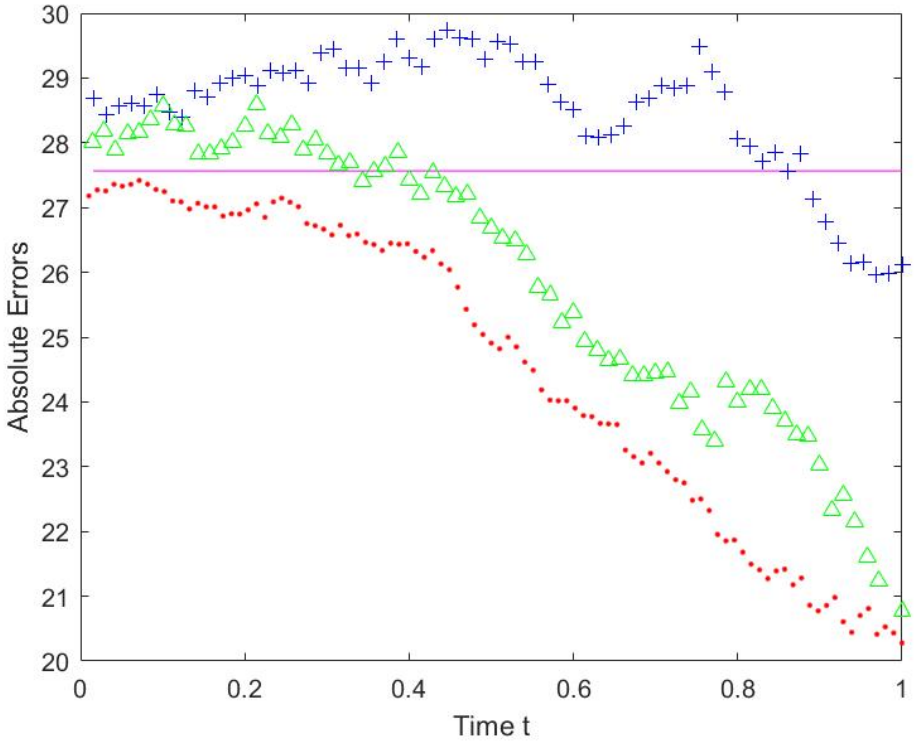

FIGURE 3.15: Time evolution of prediction errors: case of skewness and Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

with each other. We do not detect significant differences in the analysis from this point of view.

For what concerns the different initial distributions, we are consistent with the literature (as we have already seen in the single previous analysis) and the best is the one corresponding to the negative-exponential: the Beta distribution with $\alpha = 1$ and $\beta = 3$. It is the one with the most predictive power regardless of the statistical indicator.

Good results obtained with the symmetric Beta with $\alpha = \beta = 0.5$ distribution but also the Uniform one. Both minimize errors and allow us to predict failure times with an acceptable predictive gain.

Errors in absolute value grow dramatically, and the damage in the predictive process for all statistical tools is truly evident both for the asymmetric Beta distribution with $\alpha = 1$ and $\beta = 0.5$, but even more so for the symmetric one with $\alpha = \beta = 2$

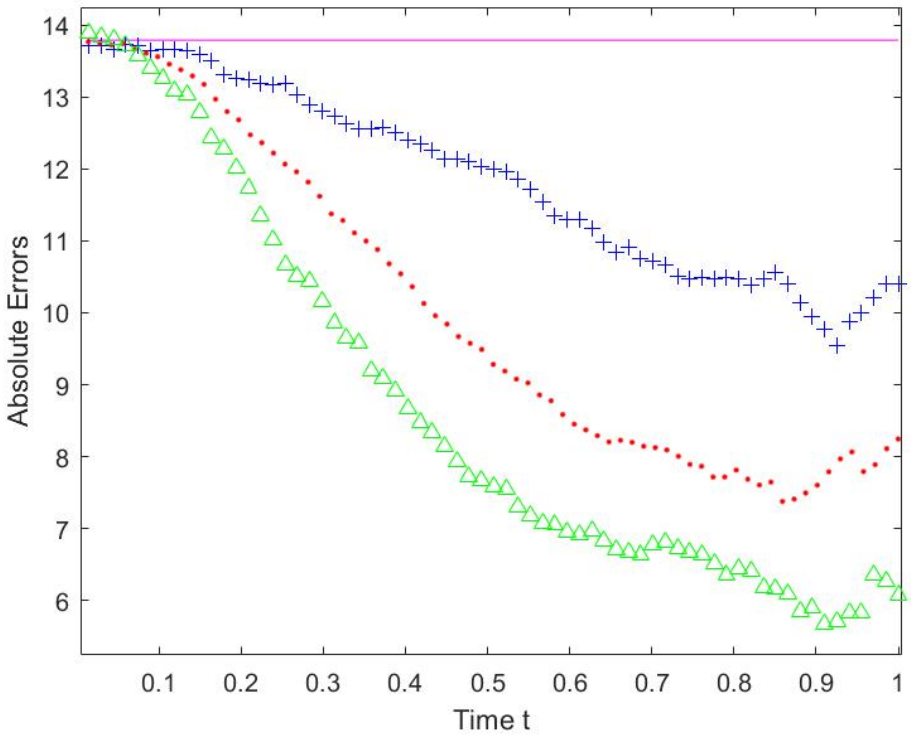

FIGURE 3.16: Time evolution of prediction errors: case of Gini coefficient and uniform distribution.

which reaches very high error levels. The latter is the absolute worst and the one that produces the poorest performances.

Figs. 3.26, 3.27 and 3.28 show the results conditioned on the uniform distribution.

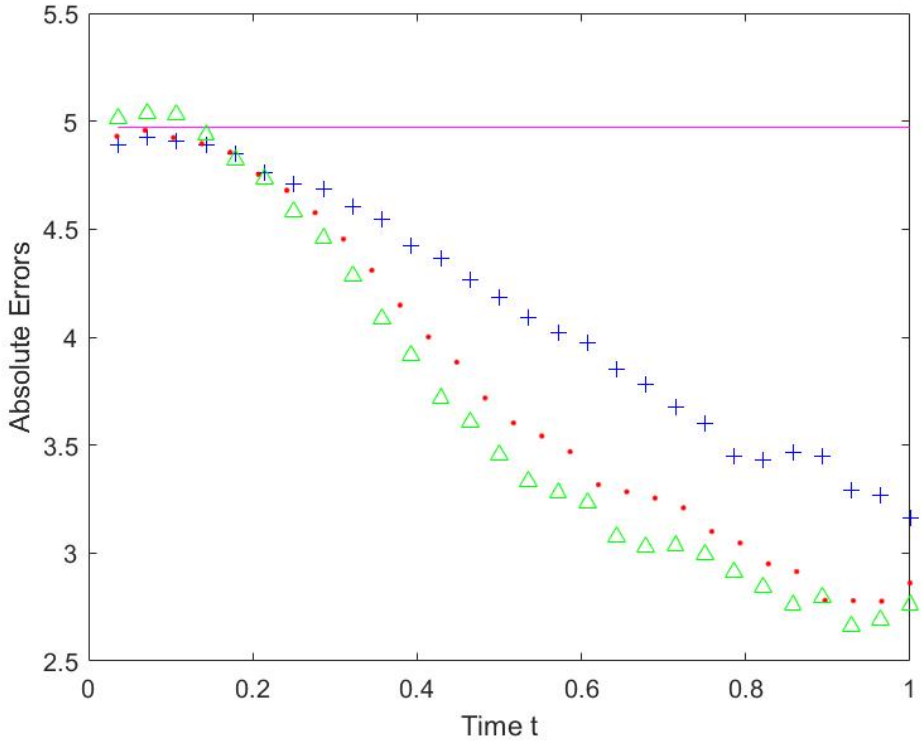

FIGURE 3.17: Time evolution of prediction errors: case of Gini coefficient and Beta distribution with $\alpha = 1$ and $\beta = 3$.

We proceed with the results of the Beta distribution with $\alpha = 1$ and $\beta = 3$ shown in Figs. 3.29, 3.30 and 3.31.

From an asymmetric Beta distribution, we continue with a symmetric Beta distribution with $\alpha = \beta = 0.5$ (see Figs. 3.32, 3.33 and 3.34).

In Figs. 3.35, 3.36 and 3.37 you can see the results of another symmetric Beta distribution (with $\alpha = \beta = 2$).

The last graphs relate to the asymmetric Beta distribution with with $\alpha = 1$ and $\beta = 0.5$ (see Figs. 3.38, 3.39 and 3.40).

In cases where the indicators show their ability to reach high values, usually abnormal (all conditioning at the 90^{th} percentile), Gini coefficient is the indicator that picks up the best information and allows an excellent improvement in the accuracy of the predictions.

Shannon entropy also proves to be a good predictive tool when we compare very

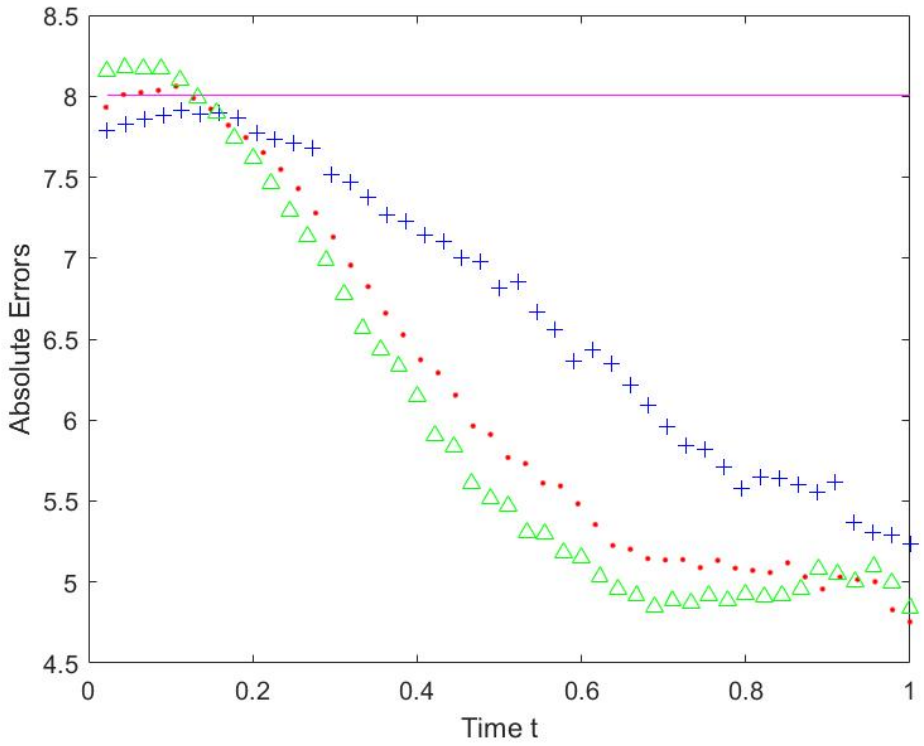

FIGURE 3.18: Time evolution of prediction errors: case of Gini coefficient and Beta distribution with $\alpha = \beta = 0.5$.

high levels of indicators.

On the other hand, when the analysis shifts towards low or average values of the distributions (10^{th} and 50^{th} percentiles conditioning), the Shannon entropy is the indicator that allows us to obtain the best result with a much faster and clear drop in the errors and earlier over time than the others. It is an excellent forecast tool with a very high predictive gain.

The Gini coefficient also gives us a satisfactory prediction capacity, in order to evaluate the correct use of rational expectations in the prediction of failure times of our stochastic systems.

For variance and kurtosis, the results are mixed.

Compared to what is stated in the literature, kurtosis performs better than variance (some exceptions excluded) despite the greater sensitivity of the fourth statistical moment to the fluctuations of our variables. Therefore, these two indicators do not

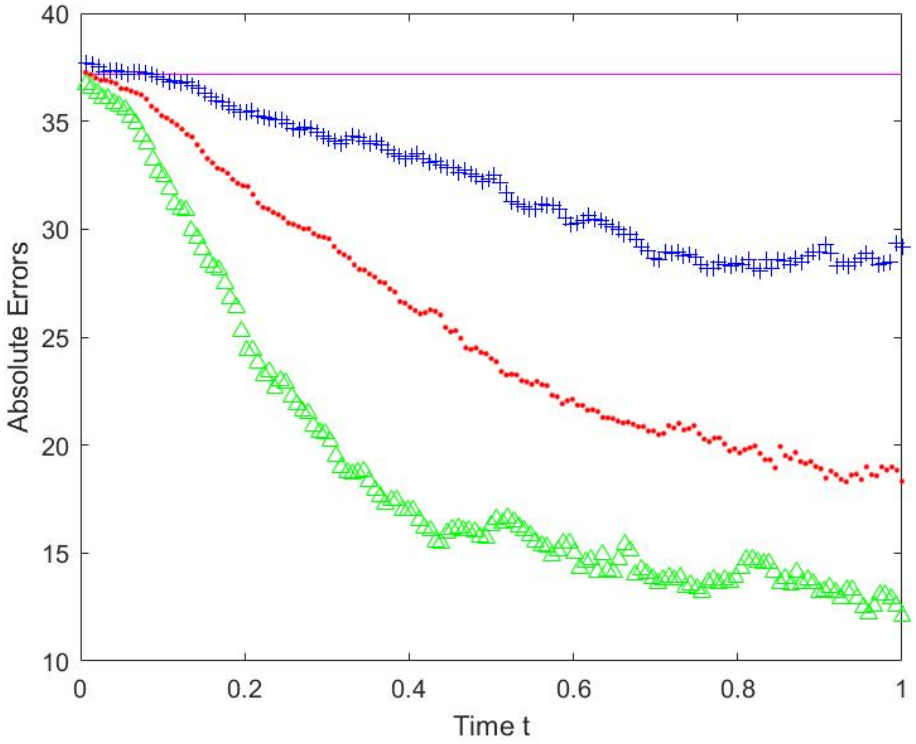

FIGURE 3.19: Time evolution of prediction errors: case of Gini coefficient and Beta distribution with $\alpha = \beta = 2$.

optimize the information available with a slightly decreasing error trend, any improvement is very random and the use of rational expectations is not as effective as a prediction tool.

The less efficient and less suitable indicator in the use of rational expectations is the skewness. It is the predictor which exploits the recorded information worse causing us to lose predictive accuracy.

In conclusion, we can say that our application consistently demonstrates a very high accuracy in the field of failure time prediction through rational expectations.

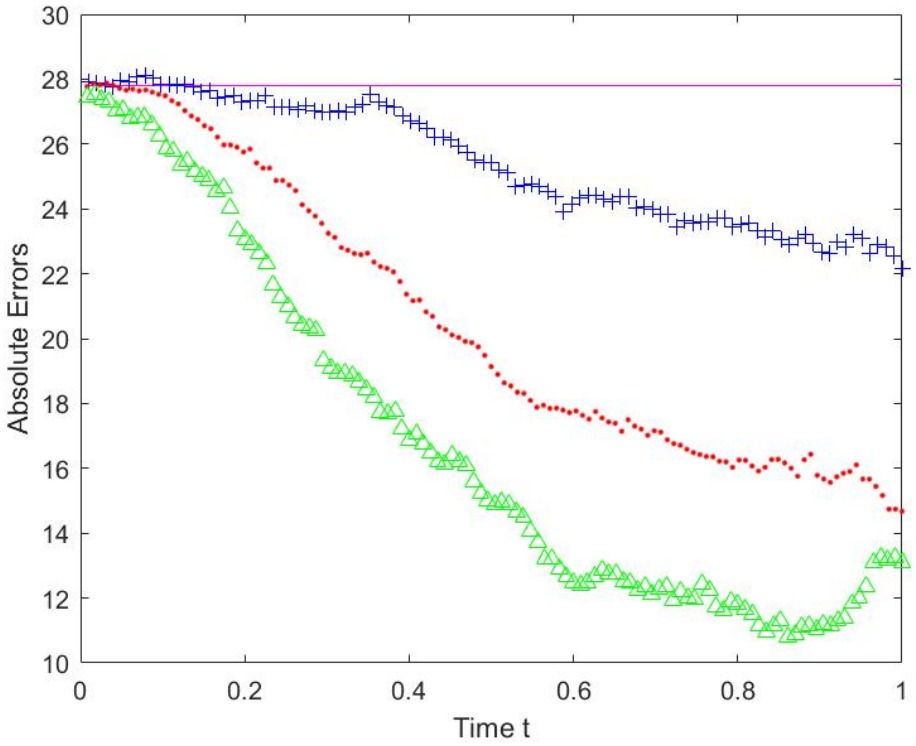

FIGURE 3.20: Time evolution of prediction errors: case of Gini coefficient and Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

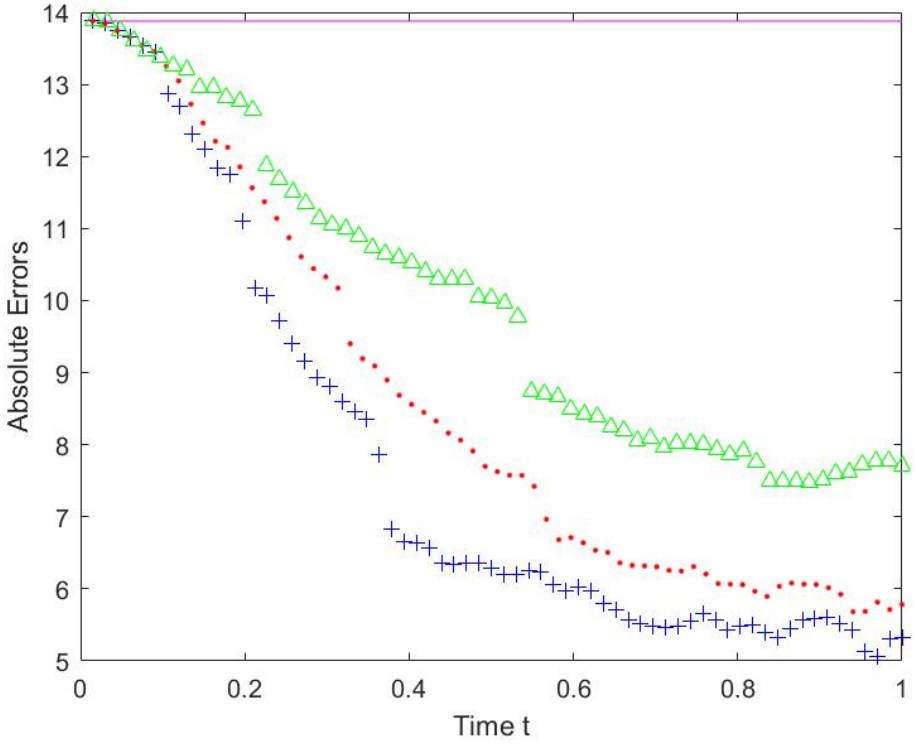

FIGURE 3.21: Time evolution of prediction errors: case of Shannon entropy and uniform distribution.

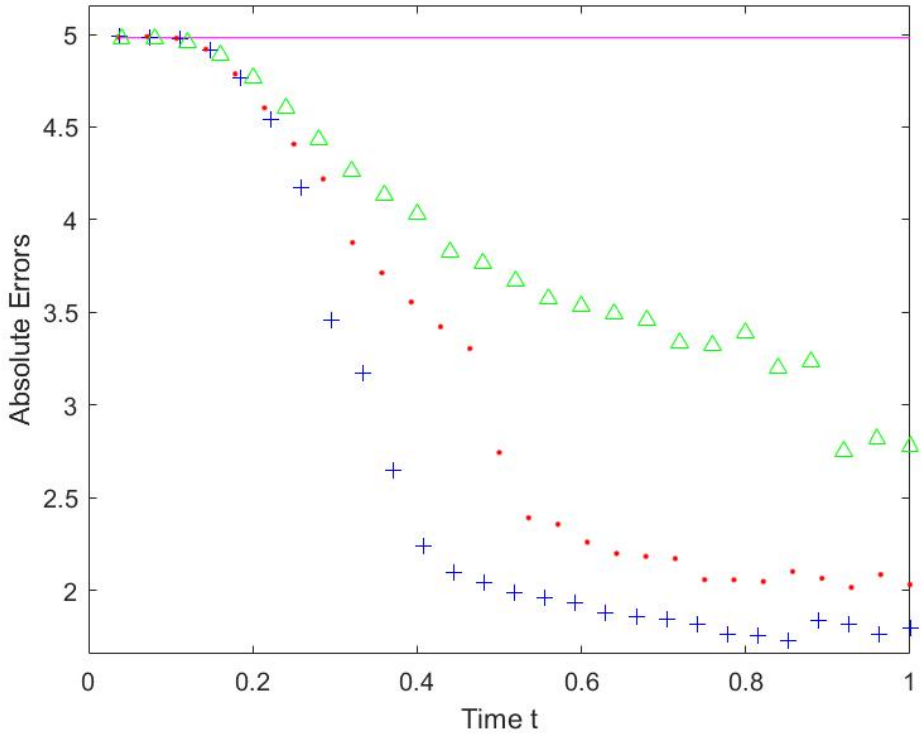

FIGURE 3.22: Time evolution of prediction errors: case of Shannon entropy and Beta distribution with $\alpha = 1$ and $\beta = 3$.

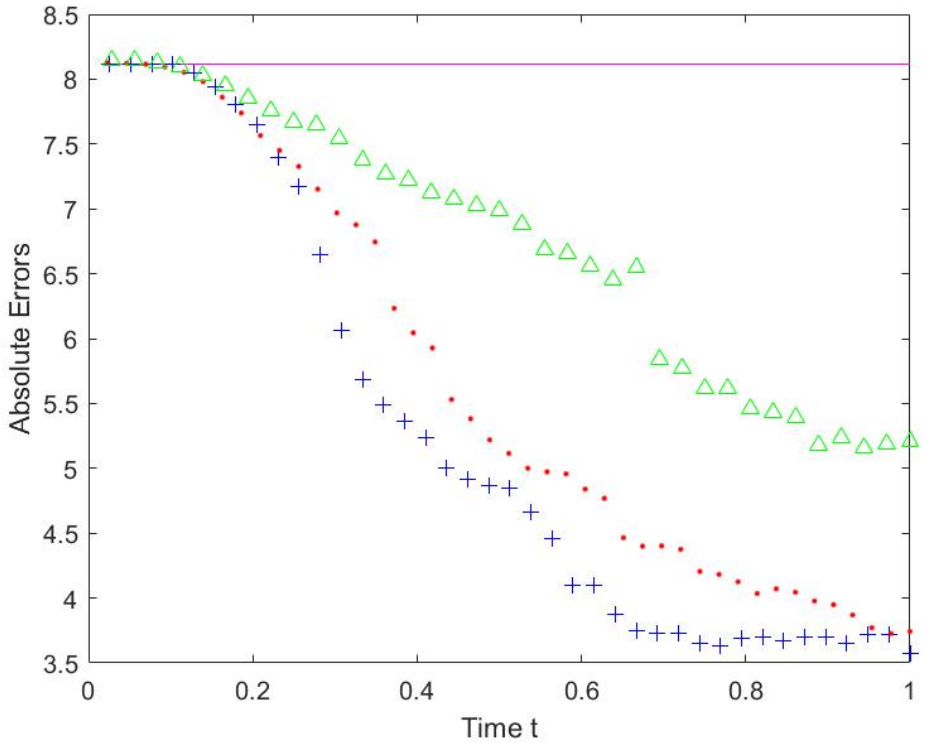

FIGURE 3.23: Time evolution of prediction errors: case of Shannon entropy and Beta distribution with $\alpha = \beta = 0.5$.

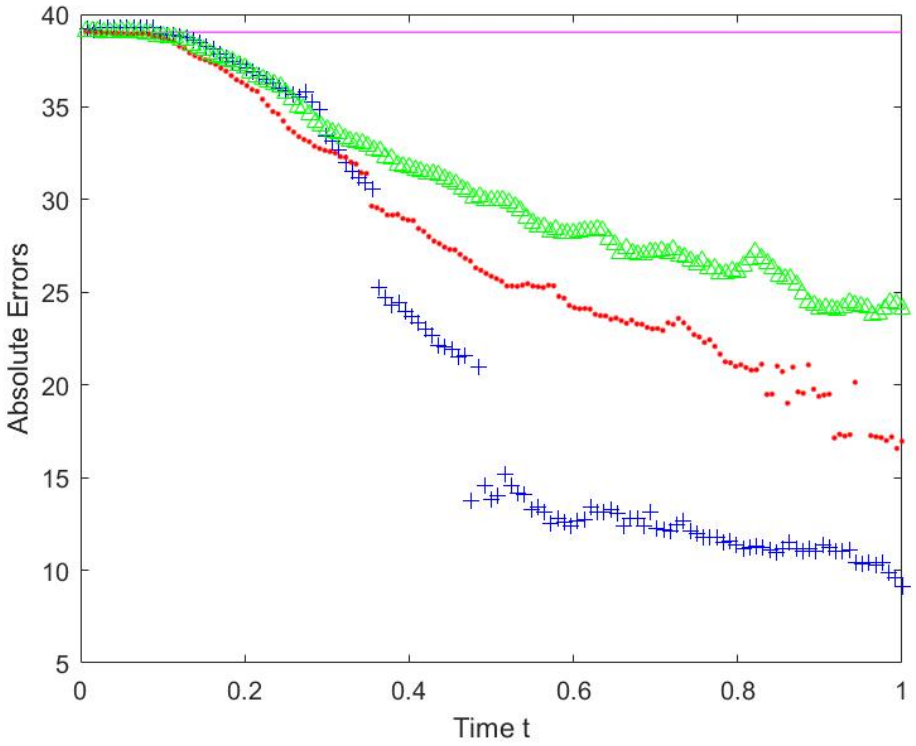

FIGURE 3.24: Time evolution of prediction errors: case of Shannon entropy and Beta distribution with $\alpha = \beta = 2$.

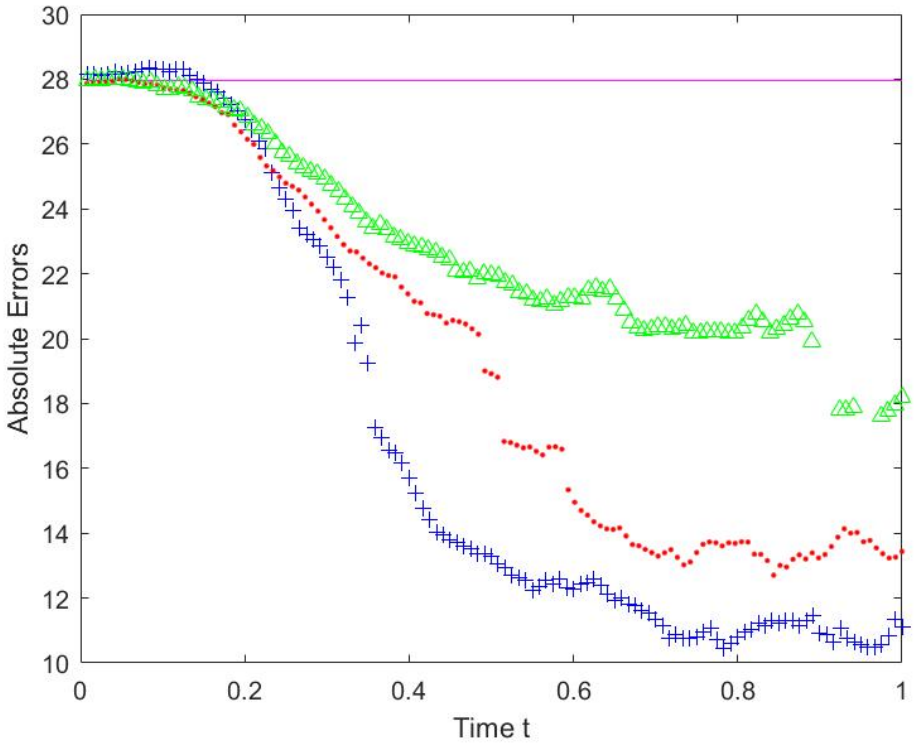

FIGURE 3.25: Time evolution of prediction errors: case of Shannon entropy and Beta distribution with $\alpha = 1$ and $\beta = 0.5$.

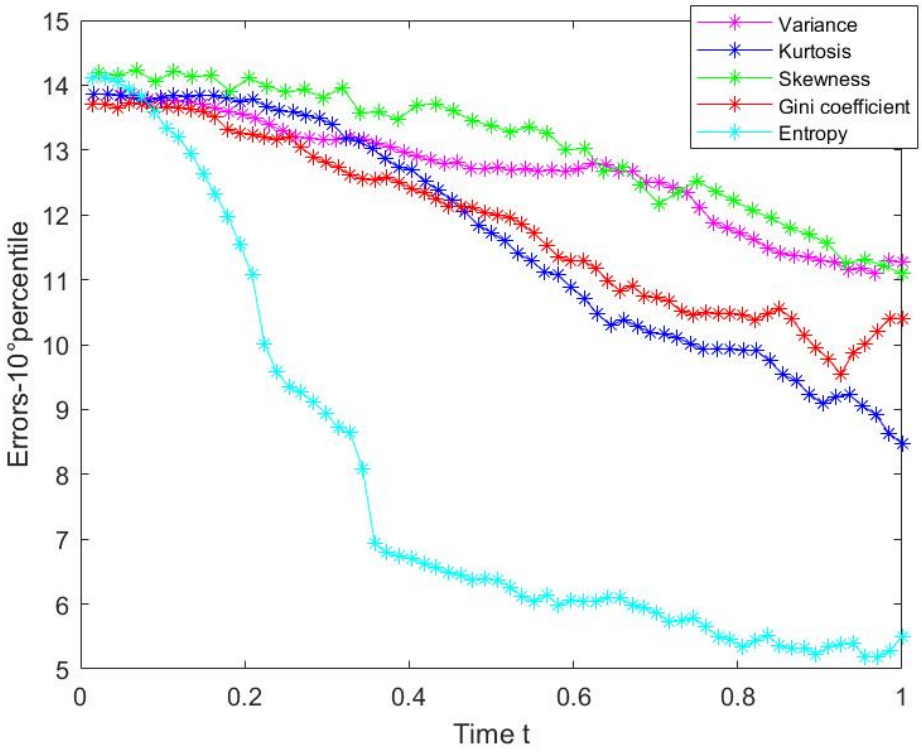

FIGURE 3.26: Comparison between variance (purple line), kurtosis (blue line), skewness (green line), Gini coefficient (red line) and Shannon entropy (light blue line) considering the conditioning of the errors on the 10^{th} percentile of the distribution. Case of uniform distribution.

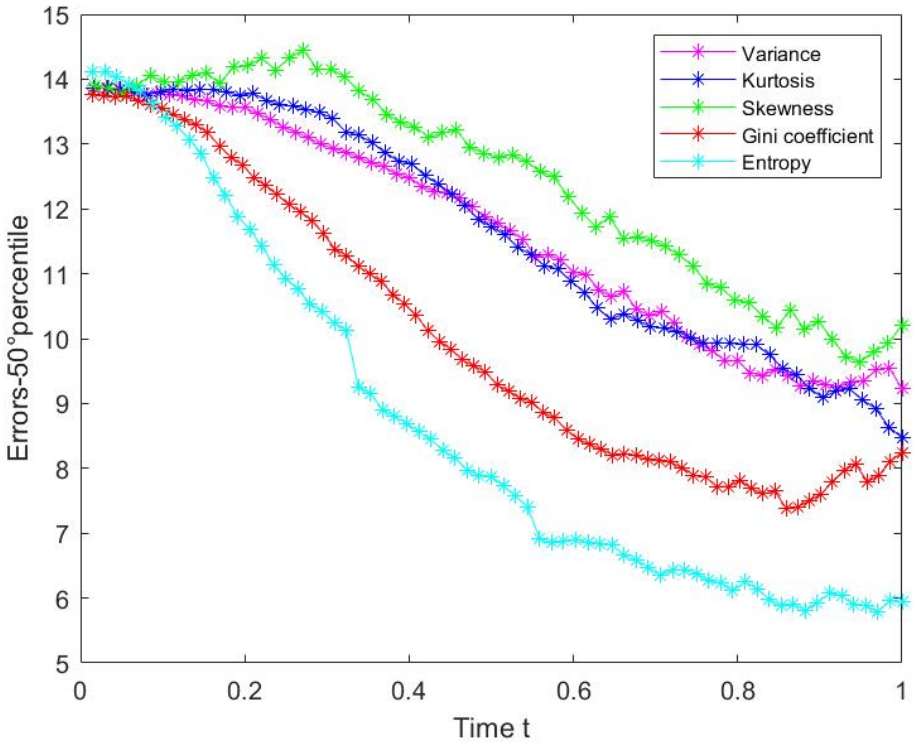

FIGURE 3.27: Case of the 50^{th} percentile with the uniform distribution.

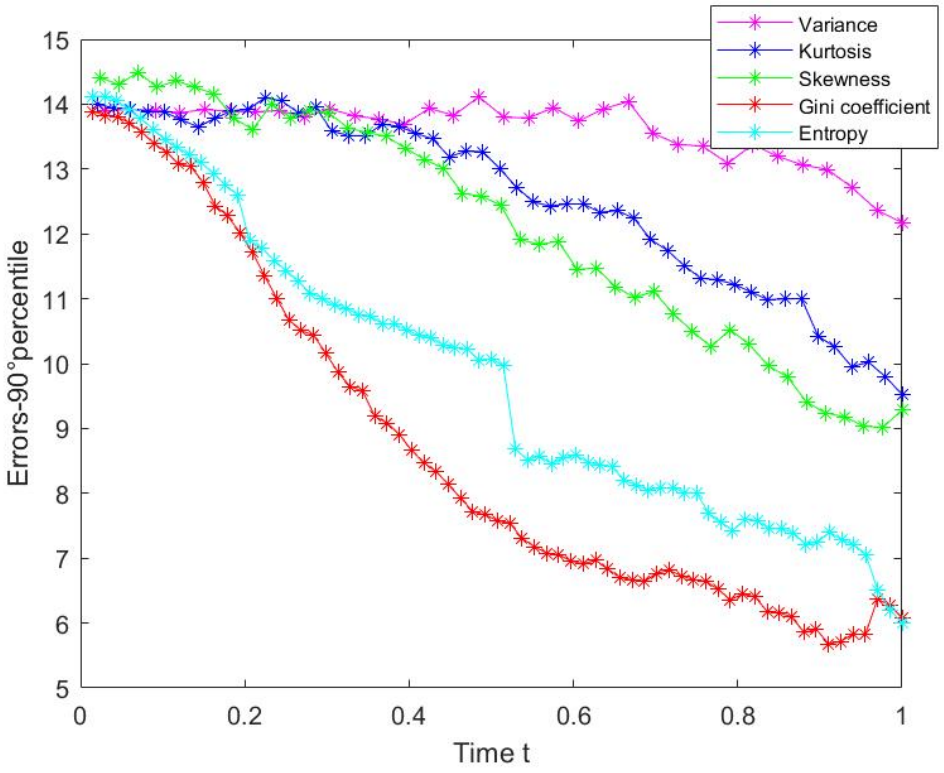

FIGURE 3.28: Prediction errors in the case of uniform distribution conditioned on the 90^{th} percentile.

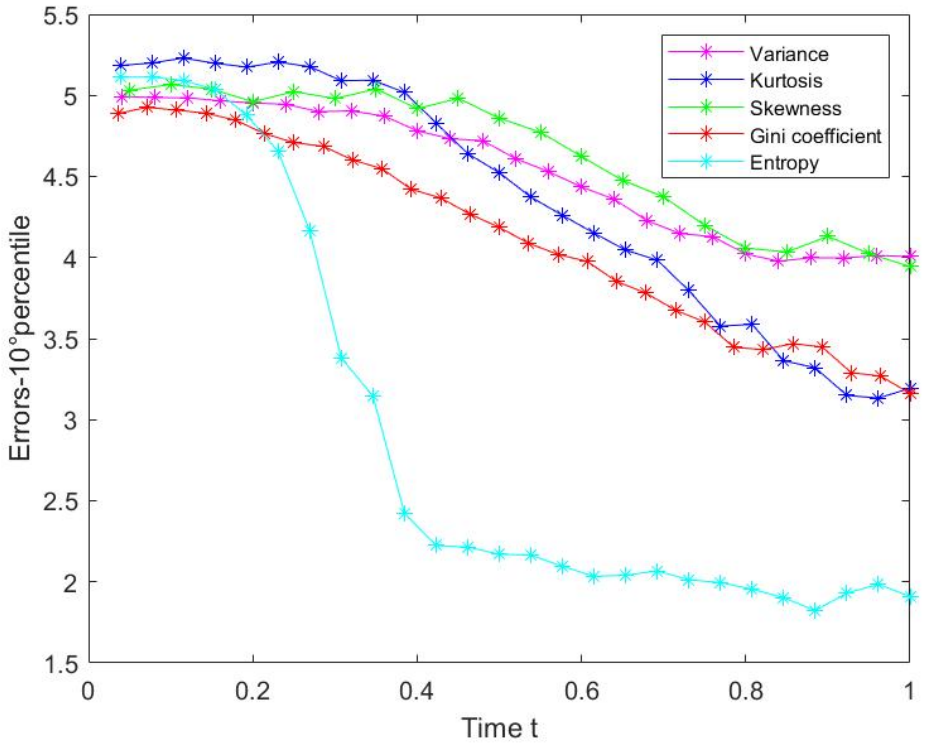

FIGURE 3.29: Prediction errors (in absolute value) for Beta distribution with $\alpha = 1$ and $\beta = 3$ conditioned on the 10^{th} percentile of the distribution.

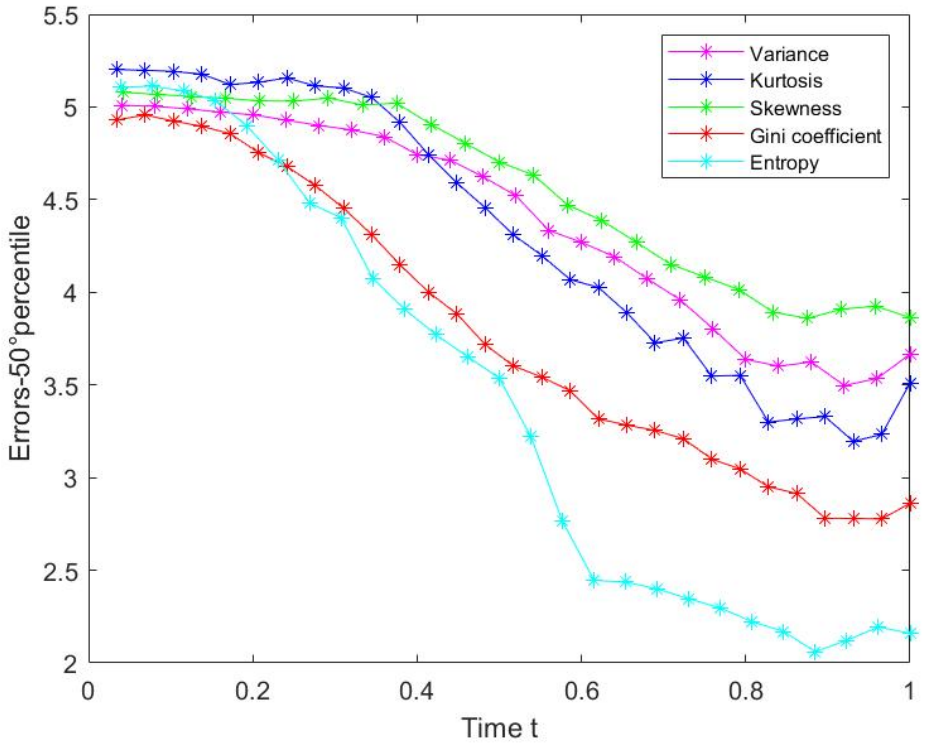

FIGURE 3.30: Beta distribution with $\alpha = 1$ and $\beta = 3$ with errors conditioned on the 50^{th} percentile of the statistical measure distributions.

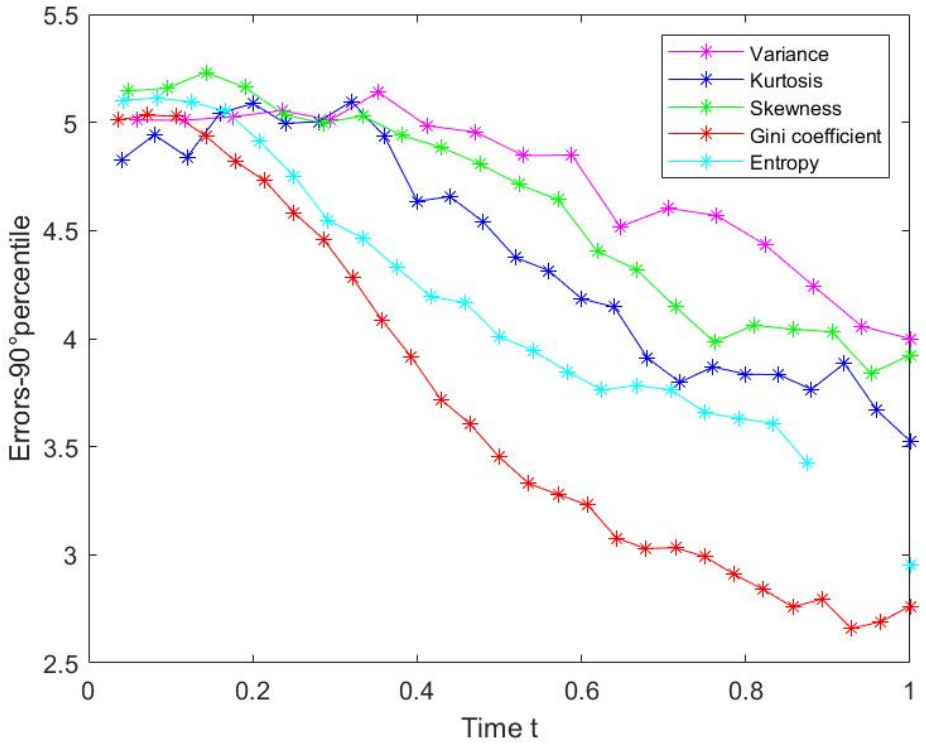

FIGURE 3.31: Beta distribution with $\alpha = 1$ and $\beta = 3$ in the case of the errors conditioned on the 90^{th} percentile of the distributions.

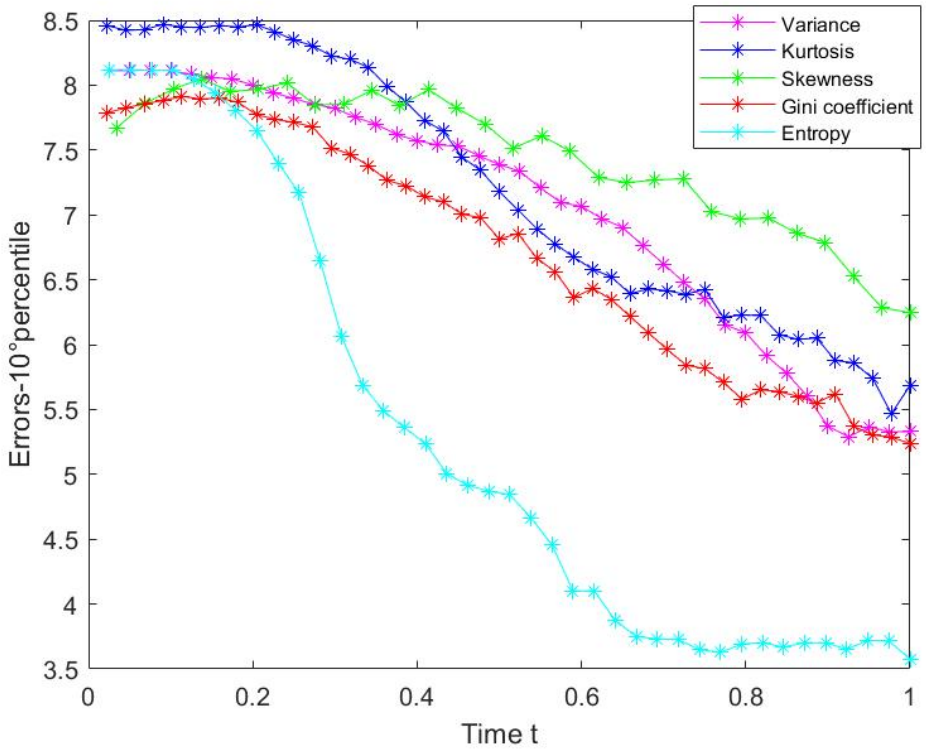

FIGURE 3.32: Errors (in absolute value) for Beta distribution with $\alpha = \beta = 0.5$ conditioned on the 10^{th} percentile.

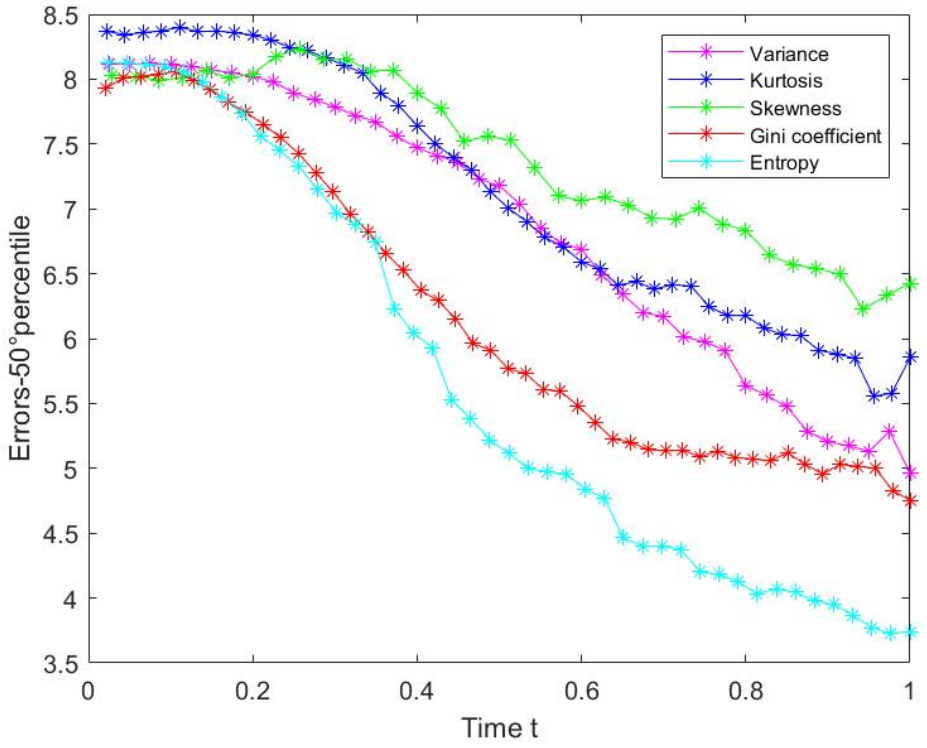

FIGURE 3.33: Prediction errors as in the previous graph but with conditioning on 50^{th} percentile.

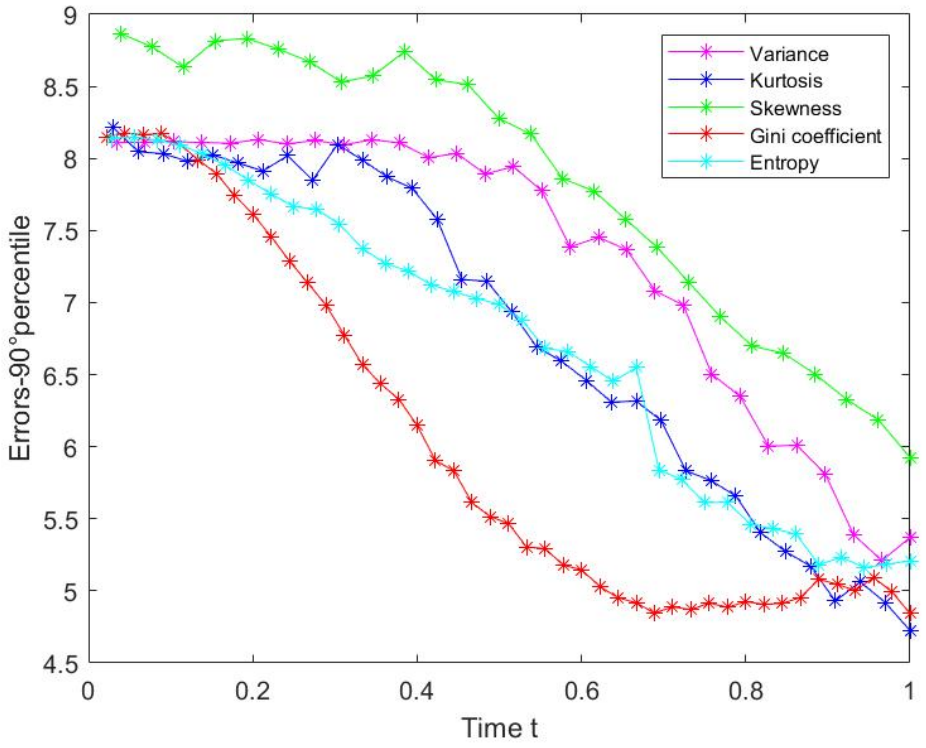

FIGURE 3.34: Beta with $\alpha = \beta = 0.5$ and conditional distribution on the 90^{th} percentile.

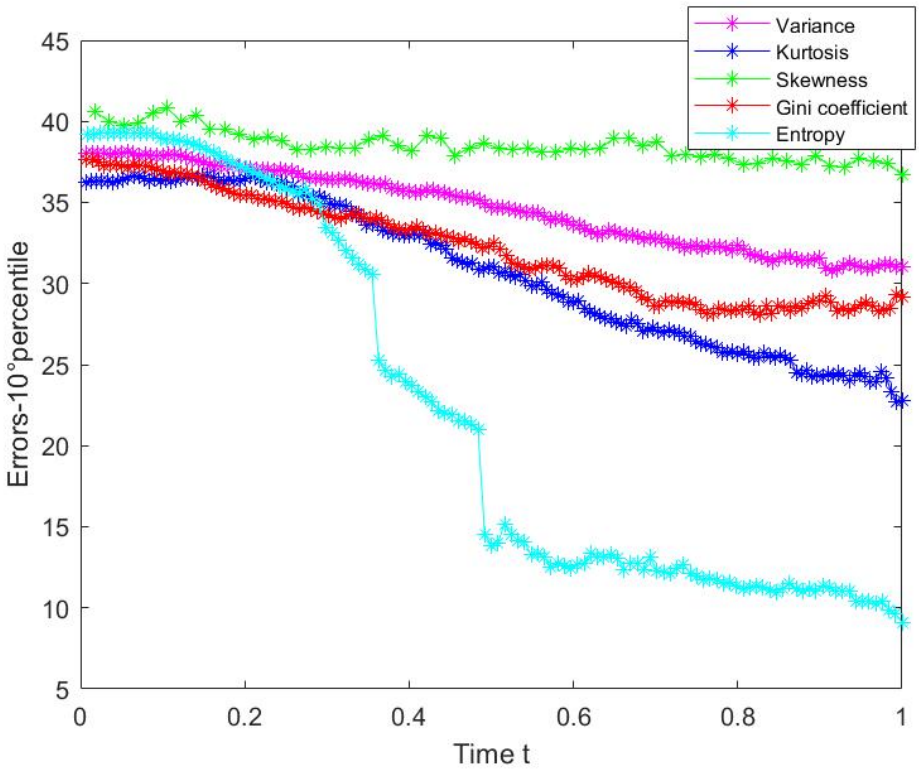

FIGURE 3.35: Overall graph with the analysis of variance, kurtosis, skewness, Gini coefficient and Shannon entropy. Beta distribution with $\alpha = \beta = 2$ conditioned on the 10^{th} percentile.

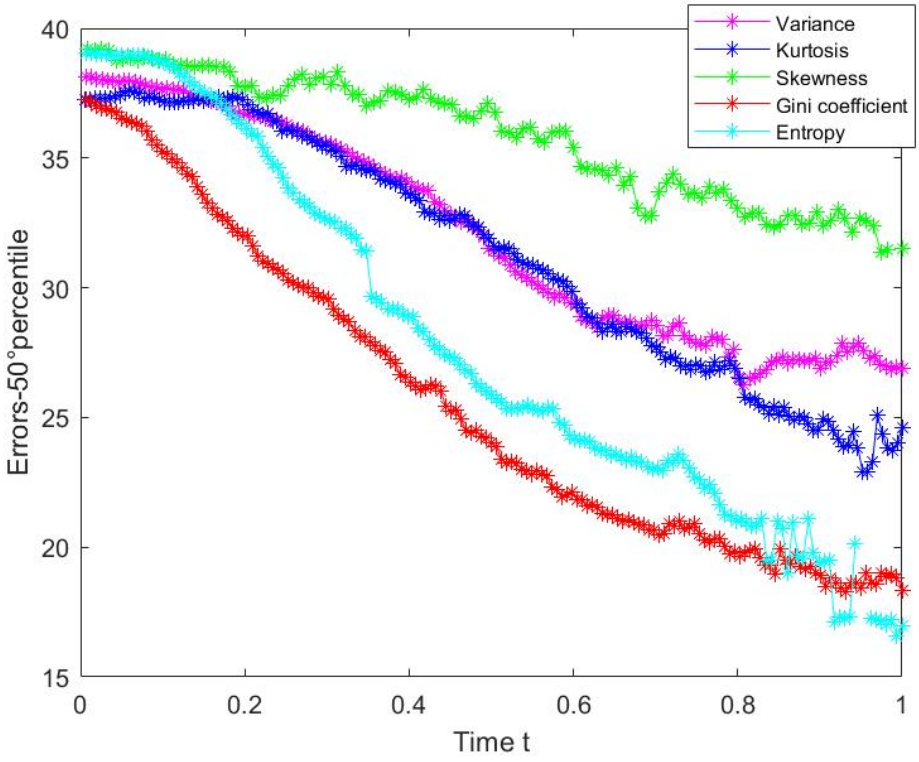

FIGURE 3.36: Here the case of Beta distribution with $\alpha = \beta = 2$ and the 50^{th} percentile.

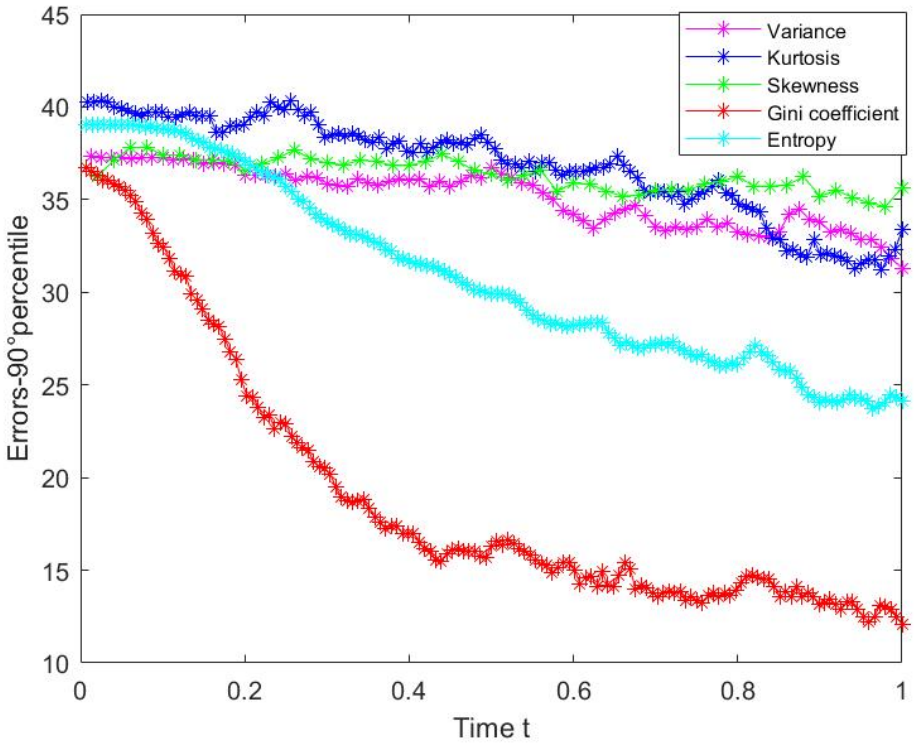

FIGURE 3.37: Prediction errors (in absolute value) considering Beta with $\alpha = \beta = 2$ and the conditioning on the 90^{th} percentile of the distribution.

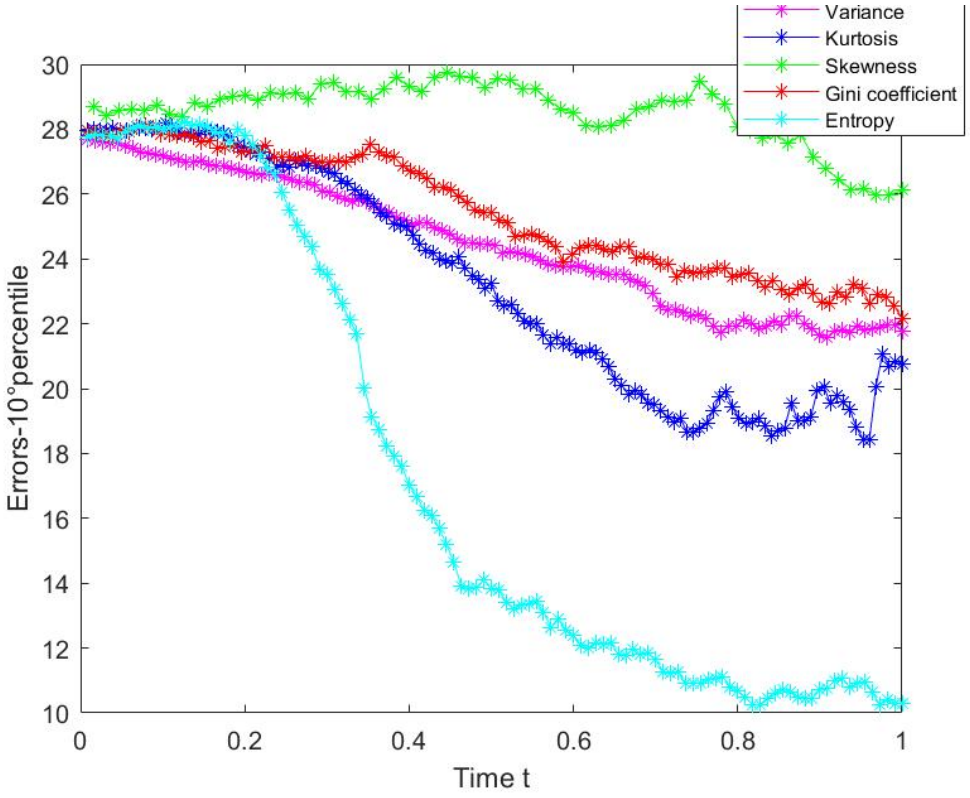

FIGURE 3.38: Prediction errors for Beta distribution with $\alpha = 1$ and $\beta = 0.5$ conditioned on the 10^{th} percentile of the statistical indicator distributions.

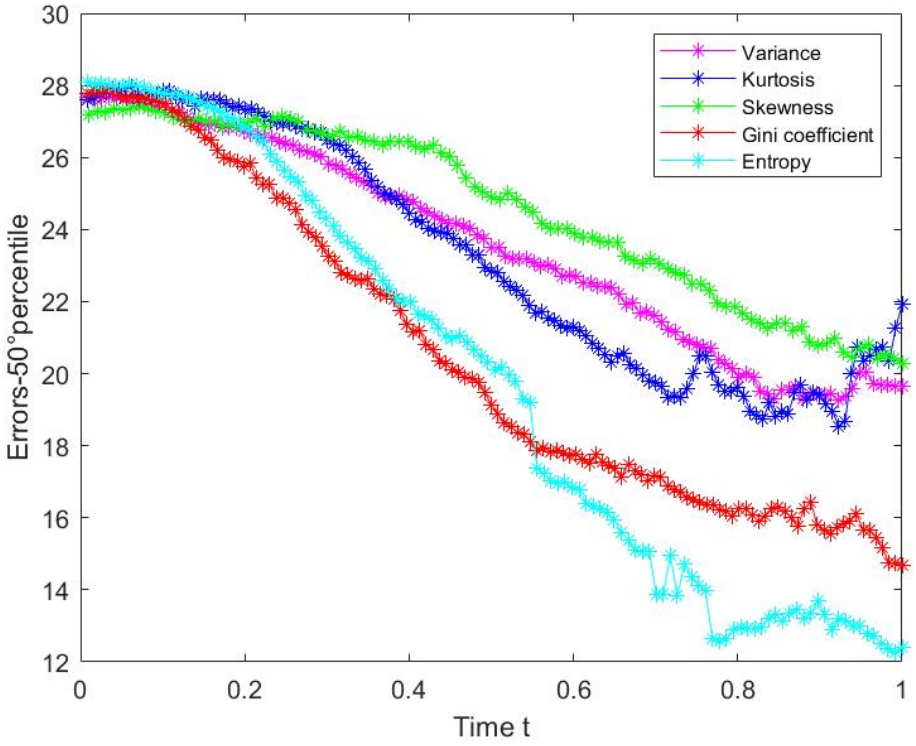

FIGURE 3.39: Final comparison between variance (purple line), kurtosis (blue line), skewness (green line), Gini coefficient (red line) and Shannon entropy (light blue line) considering Beta distribution with $\alpha = 1$ and $\beta = 0.5$ and a conditioning on the 50^{th} percentile.

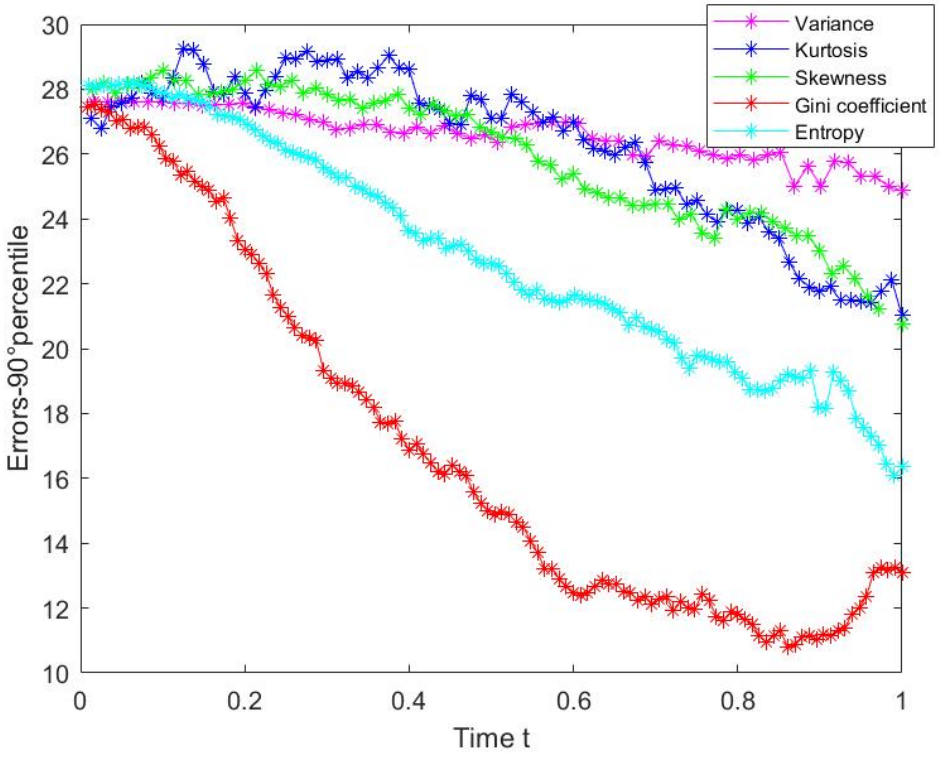

FIGURE 3.40: Prediction errors as above, but in this figure we take into account the 90^{th} percentile.

Conclusions and future research

Our goal is to provide the most accurate prediction of the failure times of simulated stochastic systems in a context of rational expectations. We therefore focus on the reliability of this framework, investigating the optimal statistical measure to obtain successful predictions depending on the initial conditions of the systems considered.

We have presented a theoretical setting followed by two different frameworks of application that both exploit rational expectations in the context of the reliability of systems with interconnected components. The main elements of this research are therefore the time and the statistical measures on which the analysis is based. We explore two perspectives on this problem with results focused on different aspects: in the first, the focus is on errors that depend directly on the values of the given descriptive statistics; in the second, we relate directly to the temporal evolution.

In the first model we make a transversal analysis over time: we emphasize the contribution of the statistical indicator, showing how the error varies as a function of the measure considered aggregated over time.

Instead, in the second model, we emphasize the importance of the time factor on different scenarios of the statistical indicators, based on the percentiles we condition them on.

Unless otherwise stated, all the results presented in this thesis are new and original, and encourage the application of rational expectations to our simulated data.

In both models we compare our research field with the existing literature and can thus evaluate the robustness of our results, the proposed methods, and the computational efficiency of our algorithms.

We examine how the paths of the errors are influenced by several indicators, the

different levels of these indicators, and initial weight distributions used in the study. In both presented models we have successfully demonstrated that it is possible to obtain a prediction of failure times of stochastic systems which improves for certain combinations of indicators and initial distributions and with the passage of time.

The graphs show five characteristics of the α distribution (variance, kurtosis, skewness, Gini coefficient and Shannon entropy) considering five different initial weight distributions. They confirm the significant gain in prediction accuracy, taking into account indicators that synthetize the systems.

Depending on the initial distribution, Beta distribution with $\alpha = \beta = 0.5$ and the one with $\alpha = 1$ and $\beta = 3$ provide the best results and the lowest errors for all indicators studied. Notably, the symmetric case with $\alpha = \beta = 2$ is the worst, returning very high errors. This suggests that we are able to identify and capture the characteristics of the distributions with the most predictive power. We are also able to understand for which values of the indicators we obtain the greatest gain in the use of rational expectations by recording more information and providing high precision in predicting the residual failure times.

Applying this new method in reliability theory, we have managed to implement many strategies and the best one, depending on the distribution examined and the type of errors taken into consideration, is the one that allows us to minimize the prediction errors and obtain a greater gain. To our knowledge, this has not been done before.

In conclusion, we have proposed a model for describing rational expectations in the context of reliability theory, and we have validated the effectiveness of our model through two different extensive simulation procedures.

In future research further implementations of the model will have to be developed by increasing the complexity of the simulation procedure. The problem can be addressed in other rational expectations models and should certainly be attempted.

For example, we can modify the reallocation rule by proposing instead of the proportional one, a uniform reallocation rule, a reallocation rule threshold based or a rule that depends on the trajectories of previous failures. We can do the same with

the failure rule by changing the conditions under which systems fail in our theoretical model.

We will be able to extend the analysis by comparing additional statistical indicators: the Frosini index for a comparison with the Gini coefficient, the Bienaymé–Chebyshev inequality, the Pearson index to capture the dependence between the components, and Goodman and Kruskal's index, also for the correlation between the components. We can also use other distances belonging to the Shannon family of entropies such as the Kullback-Leibler divergence, the Jeffreys distance, the K divergence or the Jensen difference.

We can also consider a distance measure to check which combination of parameters will allow us to get the lowest error.

The discipline of predicting failure times is still very active and useful. There are therefore several areas whre future research could seek improvement.

One intention there is to study specifically how the evolution of the distribution of the components can affect our results. We would like to investigate a possible regularity in the random path of the components and the way it changes for different exogenously assigned distribution. The finite Markov chain imbedding technique could be useful here, as it would allow us to imbed the pattern of our components into the framework of a finite Markov chain in order to carry out further research in this area. In this way, in fact, we could reduce the computational complexity associated with the various methods usually used to find the probability distribution of lifetimes of systems and understand whether there is a regularity that binds the components of our systems.

We will be able to analyse in-depth the importance of components within a system, assuming that the structure of the systems is known. There are many measures that can be used, such as the Barlow-Proschan importance, the Hwang index, the Fussell-Vesely importance, or the B-importance.

Again, with regard to this field, a network could be built on the connection existing between the components by considering of the probability with which they subsequently fail.

We can also integrate cluster analysis by clustering our predictors over different time windows using a dendrogram to show in which periods the prediction was most effective.

It is also our intention to apply the simulation procedure to real data to model systems with interconnected components. Numerical experiments in the economic and financial field will follow.

Bibliography

[1] Navarro, J., Ruiz, J. M., & Sandoval, C. J. (2005). A note on comparisons among coherent systems with dependent components using signatures. Statistics and Probability Letters, 72(2), 179-185.

[2] Khaledi, B. E., & Shaked, M. (2007). Ordering conditional lifetimes of coherent systems. Journal of Statistical Planning and Inference, 137(4), 1173-1184.

[3] Navarro, J., Samaniego, F. J., Balakrishnan, N., & Bhattacharya, D. (2008). On the application and extension of system signatures in engineering reliability. Naval Research Logistics (NRL), 55(4), 313-327.

[4] Navarro, J., & Spizzichino, F. (2010). Comparisons of series and parallel systems with components sharing the same copula. Applied Stochastic Models in Business and Industry, 26(6), 775-791.

[5] Di Crescenzo, A., & Pellerey, F. (2011). Improving series and parallel systems through mixtures of duplicated dependent components. Naval Research Logistics (NRL), 58(5), 411-418.

[6] Navarro, J., del Águila, Y., Sordo, M. A., & Suárez-Llorens, A. (2013). Stochastic ordering properties for systems with dependent identically distributed components. Applied Stochastic Models in Business and Industry, 29(3), 264-278.

[7] Navarro, J., Pellerey, F., & Di Crescenzo, A. (2015). Orderings of coherent systems with randomized dependent components. European Journal of Operational Research, 240(1), 127-139.

[8] Gupta, N., Misra, N., & Kumar, S. (2015). Stochastic comparisons of residual lifetimes and inactivity times of coherent systems with dependent identically distributed components. European Journal of Operational Research, 240(2), 425-430.

[9] Azaron, A., Katagiri, H., Sakawa, M., & Modarres, M. (2005). Reliability function of a class of time-dependent systems with standby redundancy. European Journal of Operational Research, 164(2), 378-386.

[10] Borgonovo, E., Aliee, H., Glaß, M., & Teich, J. (2016). A new time-independent reliability importance measure. European Journal of Operational Research, 254(2), 427-442.

[11] Parsa, M., Di Crescenzo, A., & Jabbari, H. (2018). Analysis of reliability systems via Gini-type index. European Journal of Operational Research, 264(1), 340-353.

[12] Çekyay, B., & Özekici, S. (2010). Mean time to failure and availability of semi-Markov missions with maximal repair. European Journal of Operational Research, 207(3), 1442-1454.

[13] Zarezadeh, S., Mohammadi, L., & Balakrishnan, N. (2018). On the joint signature of several coherent systems with some shared components. European Journal of Operational Research, 264(3), 1092-1100.

[14] Oe, S., Soeda, T., & Nakamizo, T. (1980). A method of predicting failure or life for stochastic systems by using autoregressive models. International Journal of Systems Science, 11(10), 1177-1188.

[15] Bhattacharya, S. K. (1967). Bayesian approach to life testing and reliability estimation. Journal of the American Statistical Association, 62(317), 48-62.

[16] El-Sayyad, G. M. (1967). Estimation of the parameter of an exponential distribution. Journal of the Royal Statistical Society: Series B (Methodological), 29(3), 525-532.

[17] Canfield, R. V. (1970). A Bayesian approach to reliability estimation using a loss function. IEEE Transactions on Reliability, 19(1), 13-16.

[18] Varian, H. R. (1975). A Bayesian approach to real estate assessment. Studies in Bayesian econometric and statistics in Honor of Leonard J. Savage, 195-208.

[19] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. Journal of the American Statistical Association, 81(394), 446-451.

[20] Basu, A. P., & Ebrahimi, N. (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. Journal of Statistical Planning and Inference, 29(1-2), 21-31.

[21] Van Noortwijk, J. M., Cooke, R. M., & Kok, M. (1995). A Bayesian failure model based on isotropic deterioration. European Journal of Operational Research, 82(2), 270-282.

[22] Gunawan, S., & Papalambros, P. Y. (2006). A Bayesian approach to reliability-based optimization with incomplete information. Journal of Mechanical Design, 128(4), 909-918.

[23] Mastran, D. V. (1976). Incorporating component and system test data into the same assessment: a Bayesian approach. Operations Research, 24(3), 491-499.

[24] Mastran, D. V., & Singpurwalla, N. D. (1978). A Bayesian estimation of the reliability of coherent structures. Operations Research, 26(4), 663-672.

[25] Barlow, R. E. (1985). Combining component and system information in system reliability calculation. In Probabilistic Methods in the Mechanics of Solids and Structures (pp. 375-383). Springer, Berlin, Heidelberg.

[26] Martz, H. F., Wailer, R. A., & Fickas, E. T. (1988). Bayesian reliability analysis of series systems of binomial subsystems and components. Technometrics, 30(2), 143-154.

[27] Martz, H. F., & Wailer, R. A. (1990). Bayesian reliability analysis of complex series/parallel systems of binomial subsystems and components. Technometrics, 32(4), 407-416.

[28] Peng, W., Huang, H. Z., Xie, M., Yang, Y., & Liu, Y. (2013). A Bayesian approach for system reliability analysis with multilevel pass-fail, lifetime and degradation data sets. IEEE Transactions on Reliability, 62(3), 689-699.

[29] Kim, M. J., Jiang, R., Makis, V., & Lee, C. G. (2011). Optimal Bayesian fault prediction scheme for a partially observable system subject to random failure. European Journal of Operational Research, 214(2), 331-339.

[30] Aktekin, T., & Caglar, T. (2013). Imperfect debugging in software reliability: A Bayesian approach. European Journal of Operational Research, 227(1), 112-121.

[31] Jiang, Z., & Skorupski, W. (2018). A Bayesian approach to estimating variance components within a multivariate generalizability theory framework. Behavior Research Methods, 50(6), 2193-2214.

[32] Da Costa Bueno, V., & do Carmo, I. M. (2007). Active redundancy allocation for a k-out-of-n: F system of dependent components. European Journal of Operational Research, 176(2), 1041-1051.

[33] Eryilmaz, S. (2012). On the mean residual life of a k-out-of-n: G system with a single cold standby component. European Journal of Operational Research, 222(2), 273-277.

[34] Wang, Y., Li, L., Huang, S., & Chang, Q. (2012). Reliability and covariance estimation of weighted k-out-of-n multi-state systems. European Journal of Operational Research, 221(1), 138-147.

[35] Zhang, N., Fouladirad, M., & Barros, A. (2019). Reliability-based measures and prognostic analysis of a K-out-of-N system in a random environment. European Journal of Operational Research, 272(3), 1120-1131.

[36] Barlow and Heidtmann: Barlow, R. E., & Heidtmann, K. D. (1984). Computing k-out-of-n system reliability. IEEE Transactions on Reliability, 33(4), 322-323.

[37] Jain and Gopal: Jain, S. P., & Gopal, K. (1985). Recursive algorithm for reliability evaluation of k-out-of-n: G system. IEEE Transactions on Reliability, 34(2), 144-150.

[38] Rai et al: Rai, S., Sarje, A. K., Prasad, E. V., & Kumar, A. (1987). Two recursive algorithms for computing the reliability of k-out-of-n systems. IEEE Transactions on Reliability, 36(2), 261-265.

[39] Pham, H., & Upadhyaya, S. J. (1988). The efficiency of computing the reliability of k-out-of-n systems. IEEE Transactions on Reliability, 37(5), 521-523.

[40] Hecht, M., & Hecht, H. (2000, March). Use of importance sampling and related techniques to measure very high reliability software. In 2000 IEEE Aerospace Conference. Proceedings, (4), 533-546.

[41] Li, X., & Zuo, M. J. (2002). On the behaviour of some new ageing properties based upon the residual life of k-out-of-n systems. Journal of Applied Probability, 426-433.

[42] Sarhan, A. M. (2005). Reliability equivalence factors of a parallel system. Reliability Engineering & System Safety, 87(3), 405-411.

[43] Asadi, M., & Bayramoglu, I. (2006). The mean residual life function of a k-out-of-n structure at the system level. IEEE Transactions on Reliability, 55(2), 314-318.

[44] Eryilmaz, S. (2011). Dynamic behavior of k-out-of-n: G systems. Operations Research Letters, 39(2), 155-159.

[45] Van Gemund, A. J., & Reijns, G. L. (2012). Reliability analysis of k-out-of-n systems with single cold standby using Pearson distributions. IEEE Transactions on Reliability, 61(2), 526-532.

[46] Zhang, Y. (2020). Reliability analysis of randomly weighted k-out-of-n systems with heterogeneous components. Reliability Engineering & System Safety, 205, 107184.

[47] Xie, M., & Pham, H. (2005). Modeling the reliability of threshold weighted voting systems. Reliability Engineering & System Safety, 87(1), 53-63.

[48] Li, W., & Zuo, M. J. (2008). Reliability evaluation of multi-state weighted k-out-of-n systems. Reliability Engineering & System Safety, 93(1), 160-167.

[49] Eryilmaz, S., & Bozbulut, A. R. (2014). Computing marginal and joint Birnbaum, and Barlow–Proschan importances in weighted-k-out-of-n: G systems. Computers & Industrial Engineering, 72, 255-260.

[50] Eryilmaz, S. (2014). Multivariate copula based dynamic reliability modeling with application to weighted-k-out-of-n systems of dependent components. Structural Safety, 51, 23-28.

[51] Eryilmaz, S. (2019). (k1, k2,. . . , km)-out-of-n system and its reliability. Journal of Computational and Applied Mathematics, 346, 591-598.

[52] Zhang, Y. (2018). Optimal allocation of active redundancies in weighted k-out-of-n systems. Statistics & Probability Letters, 135, 110-117.

[53] Zhang, Y. (2020). Reliability analysis of randomly weighted k-out-of-n systems with heterogeneous components. Reliability Engineering & System Safety, 205, 107184.

[54] Milczek, B. (2003). On the class of limit reliability functions of homogeneous series-"k-out-of-n" systems. Applied Mathematics and Computation, 137(1), 161-176.

[55] Ge, G., & Wang, L. (1990). Exact reliability formula for consecutive-k-out-of-n: F systems with homogeneous Markov dependence. IEEE Transactions on Reliability, 39(5), 600-602.

[56] Boland, P. J., & El-Neweihi, E. (1998). Statistical and information based (physical) minimal repair for k out of n systems. Journal of Applied Probability, 35(3), 731-740.

[57] Taghipour, S., & Kassaei, M. L. (2015). Periodic inspection optimization of a k-out-of-n load-sharing system. IEEE Transactions on Reliability, 64(3), 1116-1127.

[58] Sheu, S. H., Liu, T. H., Tsai, H. N., & Zhang, Z. G. (2019). Optimization issues in k-out-of-n systems. Applied Mathematical Modelling, 73, 563-580.

[59] Friedman, M. (1957). Introduction to "A theory of the consumption function". A theory of the consumption function (pp. 1-6). Princeton University Press.

[60] Chow, G. C. (1991). Rational versus adaptive expectations in present value models. Econometric Decision Models (pp. 269-284). Springer, Berlin, Heidelberg.

[61] Muth, J. F. (1961). Rational expectations and the theory of price movements. Econometrica: Journal of the Econometric Society, 29(3), 315-335.

[62] Lucas Jr, R. E. (1972). Expectations and the Neutrality of Money. Journal of economic theory, 4(2), 103-124.

[63] Sargent, T. J., Fand, D., & Goldfeld, S. (1973). Rational expectations, the real rate of interest, and the natural rate of unemployment. Brookings Papers on Economic Activity, 1973(2), 429-480.

[64] Sargent, T. J., & Wallace, N. (1975). "Rational" expectations, the optimal monetary instrument, and the optimal money supply rule. Journal of Political Economy, 83(2), 241-254.

[65] Barro, R. J. (1976). Rational expectations and the role of monetary policy. Journal of Monetary Economics, 2(1), 1-32.

[66] Maddock, R., & Carter, M. (1982). A child's guide to rational expectations. Journal of Economic Literature, 20(1), 39-51.

[67] Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica: Journal of the Econometric Society, 50(5), 1269-1286.

[68] Blanchard, O. J., & Watson, M. W. (1982). Bubbles, rational expectations and financial markets (No. w0945). National Bureau of Economic Research.

[69] Delcey, T., & Sergi, F. (2019). The Efficient Market Hypothesis and Rational Expectations. How Did They Meet and Live (Happily?) Ever After.

[70] Becker, O., Leitner, J., & Leopold-Wildburger, U. (2007). Heuristic modeling of expectation formation in a complex experimental information environment. European Journal of Operational Research, 176(2), 975-985.

[71] Atıcı, F. M., Ekiz, F., & Lebedinsky, A. (2014). Cagan type rational expectation model on complex discrete time domains. European Journal of Operational Research, 237(1), 148-151.

[72] Sanyal, S., Shah, V., & Bhattacharya, S. (1997). Framework of a software reliability engineering tool. In Proceedings 1997 High-Assurance Engineering Workshop, 114-119.

[73] Krishnamurthy, S., & Mathur, A. P. (1997). On the estimation of reliability of a software system using reliabilities of its components. In Proceedings The Eighth International Symposium on Software Reliability Engineering, 146-155.

[74] Gokhale, S. S., Lyu, M. R., & Trivedi, K. S. (1998). Reliability simulation of component-based software systems. In Proceedings Ninth International Symposium on Software Reliability Engineering, 192-201.

[75] Yacoub, S., Cukic, B., & Ammar, H. H. (2004). A scenario-based reliability analysis approach for component-based software. IEEE Transactions on Reliability, 53(4), 465-480.

[76] Andersen, J. V., & Sornette, D. (2005). Predicting failure using conditioning on damage history: demonstration on percolation and hierarchical fiber bundles. Physical Review E, 72(5), 056124.

[77] Sornette, D., & Andersen, J. V. (2006). Optimal prediction of time-to-failure from information revealed by damage. Europhysics Letters, 74(5), 778.

[78] Keilis-Borok, V., Shebalin, P., Gabrielov, A., & Turcotte, D. (2004). Reverse tracing of short-term earthquake precursors. Physics of the Earth and Planetary Interiors, 145(1-4), 75-85.

[79] Shebalin, P., Keilis-Borok, V., Zaliapin, I., Uyeda, S., Nagao, T., & Tsybin, N. (2004). Advance short-term prediction of the large Tokachi-oki earthquake, September 25, 2003, M= 8.1 A case history. Earth, Planets and Space, 56(8), 715-724.

[80] Jiang, Y., Lin, J., Cukic, B., & Menzies, T. (2009, November). Variance analysis in software fault prediction models. In 2009 20th International Symposium on Software Reliability Engineering (pp. 99-108). IEEE.

[81] Twomey, J. M., & Smith, A. E. (1998). Bias and variance of validation methods for function approximation neural networks under conditions of sparse data. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 28(3), 417-430.

[82] Jondeau, E., Zhang, Q., & Zhu, X. (2019). Average skewness matters. Journal of Financial Economics, 134(1), 29-47.

[83] Larasati, A., Hajji, A. M., & Dwiastuti, A. (2019). The relationship between data skewness and accuracy of Artificial Neural Network predictive model. IOP Conference Series: Materials Science and Engineering, 523(1), 012070.

[84] Reijns, G. L., & Van Gemund, A. J. (2007). Reliability analysis of hierarchical systems using statistical moments. IEEE Transactions on Reliability, 56(3), 525-533.

[85] Amari, S. V., Pham, H., & Misra, R. B. (2012). Reliability characteristics of k-out-of-n warm standby systems. IEEE Transactions on Reliability, 61(4), 1007-1018.

[86] Ramberg, J. S., Dudewicz, E. J., Tadikamalla, P. R., & Mykytka, E. F. (1979). A probability distribution and its uses in fitting data. Technometrics, 21(2), 201-214.

[87] Kinateder, H., & Papavassiliou, V. G. (2019). Sovereign bond return prediction with realized higher moments. Journal of International Financial Markets, Institutions and Money, 62, 53-73.

[88] Ooghe, H., & Spaenjers, C. (2010). A note on performance measures for business failure prediction models. Applied Economics Letters, 17(1), 67-70.

[89] Parzen, E. (1962). On estimation of a probability density function and mode. The Annals of Mathematical Statistics, 33(3), 1065-1076.

[90] Jiang, Y., Lin, J., Cukic, B., & Menzies, T. (2009, November). Variance analysis in software fault prediction models. In 2009 20th International Symposium on Software Reliability Engineering, 99-108.

[91] Zhang, M. H., Shen, X. H., He, L., & Zhang, K. S. (2018). Application of differential entropy in characterizing the deformation inhomogeneity and life prediction of low-cycle fatigue of metals. Materials, 11(10), 1917.

[92] Bal, J., Cheung, Y., & Hsu-Che, W. (2013). Entropy for business failure prediction: An improved prediction model for the construction industry. Advances in Decision Sciences, 2013, 1-13.