DOWNLOAD PDF BINARY TIME SERIES

Chapter 1 : 3 Ways to Read a Binary Clock - wikiHow

Regression models for binary time series received less attention and are confined to methods developed for a small number of repeated observations in a longitudinal context, see Diggle et al. () or Fitzmaurice et al. () for a detailed discussion.

Accurate network modeling is critical to the design of network protocols. This problem is exacerbated in the wireless domain, where fading events create ext This problem is exacerbated in the wireless domain, where fading events create extreme burstiness of delays, losses, and errors on wireless links. In this paper, we describe the data preconditioning modeling technique that is capable of capturing the statistical characteristics of wired and wireless network traces. Our main contributions are methodologies created to quantify the accuracy of network models, methodology to choose the most accurate model for a given network and characteristic of interest e. An application of higher order crossings by Kedem, Eric Slud - Biometrika , " A relation between the graphical appearance of a stationary time series and its autocorrelations is discussed. The use of counts of visual features termed higher order crossings is proposed in testing goodness of fit of hypothesized models against very general alternatives. Attention is focused on a Attention is focused on a specific example of a new type of nonparametric test statistic arising from iterative feature extraction procedures called extractors. Hyndman , " I consider models for binary time series, starting with autoregression models and then developing generalizations of them which allow nonparametric additive covariates. I show that several apparently di erent binary AR 1 models are equivalent. Three possible nonparametric additive regression m Three possible nonparametric additive regression models which allow for autocorrelation are considered; one is a generalization of an ARX model, the other two are generalizations of a regression model with AR errors. One of the models is applied to two data sets: The fitted models show that stock transaction occurrences are more likely if there have been large transactions in the previous time period. They also show that the Southern Oscillation Index does not provide a strong predictor of rainfall occurrence in Melbourne, contrary to current meteorological practice. ARX models, autocorrelated errors; autocorrelation; binary time series; generalized additive model; general Hunter , " The problem arises in fitting models involving Gaussian latent variables to environmental data. Spectral estimators and least-squares fits Spectral estimators and least-squares fits of auto- and cross-covariances are found to be of similar efficiency for fitting models to rainfall and solar radiation data. Fourier transform, latent variable, multivariate time series, rainfall, solar radiation 1 Introduction Environmental variables such as temperature can be modelled as Gaussian processes, whereas others, such as rainfall and solar radiation, are far from Gaussian but can possibly be transformed to normality. In particular, many models have been proposed for rainfall, based either on point processes Rodriguez-Iturbe et al.

Chapter 2 : javascript - D3JS: Efficient display of binary timeseries - Stack Overflow

I need information relating to logistic regression with binary time series. I have series data, it's series. My response variable is binary (1 or 0) and the covariate is numeric. I want to.

So this is a binary-valued classification problem i. This was not a very straight-forward problem to tackle because it seemed like there two possible strategies to employ. Extract features from the time series like its mean, maximum, minimum, and other differential features. Use a k-NN approach. For a given time series example that you want to predict, find the most similar time series in the training set and use its corresponding output as the prediction. I tried both of these strategies and the latter produced the best results. However this approach is not as simple as it may seem. This is because finding a good similarity measure between time series is a very non-trivial task. Finding a Similarity Measure A naive choice for a similarity measure would be Euclidean distance. The following example will show why this choice is not optimal. Consider the following of 3 time series. This is the problem with using the Euclidean distance measure. It often produced pessimistic similarity measures when it encounters distortion in the time axis. The way to deal with this is to use dynamic time warping. Dynamic Time Warping Dynamic time warping finds the optimal non-linear alignment between two time series. The Euclidean distances between alignments are then much less susceptible to pessimistic similarity measurements due to distortion in the time axis. There is a price to pay for this, however, because dynamic time warping is quadratic in the length of the time series used. Dynamic time warping works in the following way. We want to find a path through this matrix that minimizes the cumulative distance. This path then determines the optimal alignment between the two time series. It should be noted that it is possible for one point in a time series to be mapped to multiple points in the other time series. This can be implemented via the following python function. As you can see, our results have changed from when we only used the Euclidean distance measure. If you are performing dynamic time warping multiple times on long time series data, this can be prohibitively expensive. However, there are a couple of ways to speed things up. The first is to enforce a locality constraint. This way, only mappings within this window are considered which speeds up the inner loop. It can be implemented with the following function. Classification and Clustering Now that we have a reliable method to determine the similarity between two time series, we can use the k-NN algorithm for classification. The following is the 1-NN algorithm that uses dynamic time warping Euclidean distance. In this algorithm, for every time series in the test set, a search must be performed through all points in the training set so that the most similar point is found. Given that dynamic time warping is quadratic, this can be very computationally expensive. We can speed up classification using the LB Keogh lower bound. Computing LB Keogh is much less expensive than performing dynamic time warping. In this way we are eliminating many unnecessary dynamic time warping computations. We will use a window size of 4. Although the code is sped up with the use of the LB Keogh bound and the dynamic time warping locality constraint, it may still take a few minutes to run. In this algorithm, the number of clusters is set apriori and similar time series are clustered together. References The vast majority of research in this area is done by Dr. Proudly powered by Pelican , which takes great advantage of Python.

Chapter 3 : r - Forecasting binary time series - Cross Validated

I have a binary time series: We have data (0=didn't happen, 1=happened) for one-hour period in 90 days. I want to forecast after these 90 days, where the next 1 will happen, and also Extend this provision for next one month.

While the main focus of time series analysis is on the magnitude of temporal increments, a significant piece of information is encoded into the binary projection i. In this paper we provide further evidence of this by showing strong nonlinear relations between binary and non-binary properties of financial time series. These relations are a novel quantification of the fact that extreme price increments occur more often when most stocks move in the same direction. We then introduce an information-theoretic approach to the analysis of the binary signature of single and multiple time series. Through the definition of maximum-entropy ensembles of binary matrices and their mapping to spin models in statistical physics, we quantify the information encoded into the simplest binary properties of real time series and identify the most informative property given a set of measurements. Our approach allows us to connect spin models, simple stochastic processes, and ensembles of time series inferred from partial information. Export citation and abstract Content from this work may be used under the terms of the Creative Commons Attribution 3. Any further distribution of this work must maintain attribution to the author s and the title of the work, journal citation and DOI. Introduction In large systems, the observed dynamics or activity of each unit can be represented by a discrete time series providing a sequence of measurements of the state of that unit. One of the main challenges researchers are faced with is that of extracting meaningful information from the high-dimensional multiple time series characterizing all the elements of a complex system [1 â€" 9]. Traditionally, the main object of time series analysis is the characterization of patterns in the amplitude of the increments of the quantities of interest. Other studies have also documented that the binary projections of various financial [16] and neural [17] time series exhibit non-trivial dynamical features that resemble those of the original data. All these results suggest that binary projections indeed retain a non-trivial piece of information about the original time series, and call for a deeper analysis of the problem. Being binary, the sign of the increments is much more robust to noise than the increments themselves. Moreover, it is scale-invariant i. Binary time series can also be analyzed with the aid of much simpler mathematical models than required by non-binary data several examples of such models will be provided in this paper. Finally, as we show later on, in multiple financial time series the total binary increment of a given cross-section measures the instantaneous level of synchronization i. Motivated by the above considerations, in this paper we further study, both empirically and theoretically, the relationship between weighted time series and their binary projections. We first provide robust empirical evidence of novel relationships between binary and non-binary properties of real financial time series. We show that the average daily increment and average daily coupling of an empirical set of multiple time series are strongly and nonlinearly related to the corresponding average increment of the binary projections of the same time series. These empirical relations quantify in a novel way the strong correlations existing between the increments of individual stocks and the overall level of synchronization among all stocks in the market. Building on this evidence, we then introduce a formalism to analytically characterize random ensembles of single and multiple time series with desired constraints. This maximum-entropy approach is widely used in many areas, from neuroscience [21] to social network analysis [22] and more recently network science in general [23] , where it is known under the name of exponential random graph ERG formalism. In the case of interest here, we introduce ensembles of maximum-entropy binary matrices that represent projections of single and multiple binary time series, subject to a set of desired constraints defined as simple empirical measurements. We discuss the main differences between our matrix ensembles and other techniques in time series analysis, including other ensembles of random matrices encountered in random matrix theory [24 â€" 28]. Our approach leads to a family of analytically solved null models that allow us to quantify the amount of information encoded in the chosen constraints, i. After applying the approach to the financial time series in our analysis, we compare the informativeness of various measured properties and show that different properties are more relevant for different time series and temporal windows. Finally, and most importantly, we show that our approach is able to reproduce and mathematically characterize the observed nonlinear relationships between binary and non-binary properties of real time series. The rest of the paper is organized as follows. In section 2 we describe the data and provide empirical evidence of the relationships that motivate our work. In section 3 we introduce our theoretical formalism in its general form. In section 4 we apply the formalism to single time series, while in section 5 we apply it to single cross-sections temporal snapshots of multiple time series. Finally, in section 6 we consider our method in its full extent and apply it to entire spans of multiple time series, for different financial markets around the globe. We end with our conclusions in section 7. For each index, we restrict our sample to the maximal group of stocks that are traded continuously throughout the selected period. Correspondingly, we construct time series of log-returns where each entry represents the increment r i t as defined in equation 1. Finally, we take the sign x i t of each log-return r i t to obtain an additional, binarized set of time series as in equation 2. We will refer to the binarized time series as the binary projection of the original time series. In figure 1 we show a simple example of a weighted time series, along with the corresponding binary projection. The multiple time series of r i t and x i t are the main objects of our analysis throughout the paper. Note that, while the use of log-returns rather than simple returns i. Standard image High-resolution image Export PowerPoint slide The main reason for choosing the daily frequency is to achieve an optimal level of structural compatibility between the data and the models we introduce later. As we discuss in detail in section 3 , our models are binary, i. An increment of 0 is not admitted in the models: In financial markets, this is the case for daily or lower frequency. Indeed, a zero return value occurs in less than 0. Higher-frequency data feature an increasing percentage of zero returns, a property that calls for an extension of the models considered here. It should be noted that other types of binary time series, different from the type considered here, can also be defined. Most notably, binary time series can indicate the occurrence of an event in a time period, i. For such binary time series, correlations may not be very informative when measuring a dependence between the dichotomous variables. To confront this gap, in recent years new methods have been introduced, like the auto-persistence function and auto-persistence graph [29]. In these methods, the dependence structure among the observations is described in terms of conditional probabilities, rather than correlations. Although throughout this paper we will be entirely focusing on binary time series that naturally descend from the original signed time series of fluctuations, it is interesting to notice that our approach can be extended, with slight modifications, to time series as well. To this end, one needs to re-express all quantities in terms of a binary variable binary variable, and adapt our approach accordingly. For each index and for each day t in the sample, we first calculate the average over all stocks weighted return, that we denote as and define as Note that the above expression does not depend on the particular stock i, but it does depend on time t. Our unconventional choice of the symbol to denote an average over stocks is to avoid confusion with temporal averages, which will be denoted by the more usual bar later in the paper. Similarly, we calculate the corresponding average binary return , defined as In figure 2 we plot as a function of for all days of various 1 year intervals and for the three indices separately. We find a strong nonlinear dependency between the two quantities. Note that the average binary return is bound between and by construction, but the average weighted return is unbounded from both sides. While there are in principle infinite values of that are consistent with the same value of , we observe a tight relationships between the two quantities. This relationship can be fitted by a one-parameter curve of the form the theoretical justification for this functional form will be given in section 6 , where a is in general different for different years and different indices. Still, as we show later, for a given year and market the average weighted return of any day t is to a large extent predictable out of sample from the average binary return of the same day, once a is known for instance by fitting the above curve to the data for a past time window. In section 6 we will also show that the nonlinear character of the observed relations is a genuine signature of correlation in the data, as an uncorrelated null model shows a completely linear behaviour. Here each point corresponds to one day in the time interval of trading days approximately one year. The red line represents the best fit with the function , whose use is theoretically justified later in section 6. There is another empirical relationship, involving a higher-order quantity. In figure 3 we plot as a function of the average binary return , for the same data as in figure 2. Again, we find a strong nonlinear dependency, where for a given value of the average binary return of day t there is a typical value of the average coupling among all stocks in the same day. The relationship can be fitted by a one-parameter curve that diverges at. As we show in section 6 , an uncorrelated null model would yield a different, parabolic curve with no divergences. Again, this means that the empirical trend is due to genuine correlations, whose nature will be clarified later on in the paper. Standard image High-resolution image Export PowerPoint slide There are even more examples of dependencies that we can find between binary and non-binary properties in the data. However, in one way or another all these relationships, including that shown in figure 3 , ultimately derive from equation 5. For this reason, we refrain from showing redundant results and focus on the empirical findings discussed so far. The above analysis indicates that the binary signature of financial time series contains relevant information about the original data. Rather than exploring the practical aspects of this possibility of reconstruction of the original signal from its binary projection, in this paper we are interested in understanding the origin of this behaviour and providing a simple data-driven model of it. This is an important type of correlation between the magnitude of log-returns of individual time series and the level of synchronization common sign of the increments of all stocks in the market. Maximum-entropy matrix MEM ensembles Having established that the binary projections of real time series contain non-trivial information, in the rest of the paper we introduce a theory of binary time series aimed, among other things, at reproducing the observed nonlinear relationships showed in figures 2 and 3. An analogous approach is widely used, e. Moreover, we provide an analytical maximum-likelihood method to find the optimal values of the paramaters governing the ensembles, which is again similar in spirit to a method that has been recently introduced for networks [32 â€” 34]. Being entropy-based, our approach automatically allows us to measure the amount of information encoded into the observed properties chosen as constraints, i. It also allows us to identify, given a set of measured properties, which ones are more informative and which ones can be discarded, as we show on specific financial examples. Our framework turns out to reproduce the observed nonlinear relationships very well, thus providing a simple mathematical explanation and functional form for the plots shown in the previous section. We incidentally note that, despite the available variety of refined and advanced techniques in time series analysis [36], how one can quantify in the sense of statistical ensembles how much information is actually encoded into any given, measurable property of a time series is still not fully understood. While most studies, starting from the celebrated work by Kolmogorov about the algorithmic complexity of sequences of symbols [37], have addressed the quantification of the information content of a single time series, much less is known about the information encoded in the measured value of a given time series property which, necessarily, involves the idea of an entire ensemble of time series consistent with the measured value itself. Our approach can provide an answer to such a question, by associating an absolute level of uncertainty entropy to each observable of an empirical set of time series. In relative terms, this also allows us to compare the information content of different properties of a time series, thereby indicating which measured property is the most informative about the original time series. As a final consideration, it is worth mentioning that the MEM ensembles that we introduce are clearly related to and, depending on the specification, potentially overlapping with some ensembles that are well studied by random matrix theory [38 â€” 43]. However, our approach is different since we generate ensembles of matrices whose probability distributions are determined by the kind of partial information empirically measured constraint about the real system. This formalism allows us to relate the probabilistic structure of each matrix ensemble with the choice of the original observed property, or constraint. Similarly, since our matrices represent multiple time series, we are able to connect the various ensembles to simple stochastic processes induced by the associated matrix probabilities and, again, to the chosen empirical property specifying the ensembles themselves. Exponential random matrices ERMs We first analytically characterize the properties of families of randomized matrices. This procedure is analogous to e. However, we will modify it to accommodate matrices, as opposed to or non-negative matrices that describe binary and weighted networks, respectively. Each such matrix can represent N synchronous time series, all of duration T for instance, if applied to a set of multiple financial time series, N refers to the number of stocks and T to the number of time steps. Let denote a generic matrix in the ensemble, and x i t its entry ,. Let be the particular real matrix that we observe. In other words, our ensemble is composed of all possible matrices of the same type as , and includes itself. For any data-dependent property R, we will consider the value obtained when R is measured on the particular matrix.

Chapter 4 : Binary Watch | eBay

I have a binary time series with 1 when the car is not moving, and 0 when the car is moving. I want to make a forecast for a time horizon up to 36 hours ahead and for each hour.

Chapter 5 : Binary time series - Cross Validated

I would like to ask some clarifications on the method: www.nxgvision.com My problem is to forecast a binary time series one period ahead. I have a time series bin2 of length

Chapter 6 : Solved: how to model binary outcome with time series predi - SAS Support Communities

Binary Time Series Modeling With Application to Adhesion Frequency Experiments Ying HUNG, Veronika ZARNITSYNA, Yan ZHANG, Cheng ZHU, and C. F. Jeff WU Repeated adhesion frequency assay is the only published method for measuring the kinetic rates of cell adhesion.

Chapter 7 : Time series models with binary outcome - Statalist

Abstract. We consider the general regression problem for binary time series where the covariates are stochastic and time dependent and the inverse link is any differentiable cumulative distribution function.

Chapter 8 : CiteSeerX â€” Citation Query Binary Time Series

Regression models for binary time series received less attention and are conï¬‚ned to methods developed for a small number of repeated observations in a longitudinal context, see Diggle et al. () or Fitzmaurice et al. () for a detailed discussion.

Chapter 9 : R: forecasting a binary time series using the VLMC package - Stack Overflow

Get Textbooks on Google Play. Rent and save from the world's largest eBookstore. Read, highlight, and take notes, across web, tablet, and phone.