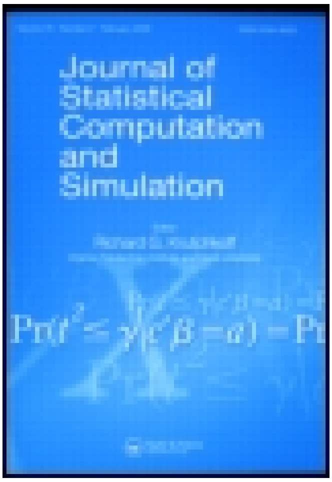

Journal of Statistical Computation and Simulation
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gscs20

Evidential inference for diffusion-type processes

Aniseh Dadgar[a], Khalil Shafie[b] & Mahdi Emadi[a]
[a] Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad 91775, Iran
[b] Department of Applied Statistics and Research Methods, University of Northern Colorado, Greeley, CO, USA
Published online: 19 Jan 2015.

Journal of Statistical Computation and Simulation, 2015
http://dx.doi.org/10.1080/00949655.2014.1002100

Taylor & Francis
Taylor & Francis Group

Evidential inference for diffusion-type processes

Aniseh Dadgar[a], Khalil Shafie[b*] and Mahdi Emadi[a]

[a]*Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad 91775, Iran;* [b]*Department of Applied Statistics and Research Methods, University of Northern Colorado, Greeley, CO, USA*

(*Received 9 April 2014; accepted 21 December 2014*)

This article analyses diffusion-type processes from a new point-of-view. Consider two statistical hypotheses on a diffusion process. We do not use a classical test to reject or accept one hypothesis using the Neyman–Pearson procedure and do not involve Bayesian approach. As an alternative, we propose using a likelihood paradigm to characterizing the statistical evidence in support of these hypotheses. The method is based on evidential inference introduced and described by Royall [Royall R. Statistical evidence: a likelihood paradigm. London: Chapman and Hall; 1997]. In this paper, we extend the theory of Royall to the case when data are observations from a diffusion-type process instead of iid observations. The empirical distribution of likelihood ratio is used to formulate the probability of strong, misleading and weak evidences. Since the strength of evidence can be affected by the sampling characteristics, we present a simulation study that demonstrates these effects. Also we try to control misleading evidence and reduce them by adjusting these characteristics. As an illustration, we apply the method to the Microsoft stock prices.

Keywords: diffusion-type processes; evidential paradigm; the law of likelihood; likelihood ratios; statistical evidence; misleading evidence

AMS Subject Classifications: 62A99; 62F03; 62M02

1. Introduction

Inference on stochastic processes involves analysis of dependent observations. The earliest work on statistics for stochastic processes was done by Grenander.[1] He continued to study this subject and extended much of statistical concepts and methods for stochastic processes. 'Abstract inference' is a complete collection of the result of these studies.[2] Hypothesis testing for stochastic processes is one of his most important works in this book. Some other earlier issues are Anderson and Goodman,[3] Billingsly,[4] Hajec,[5] Rao [6] and Lipster and Shiryeyv.[7]

Recently studies on inference for stochastic processes are excessive and is specialized to branches of stochastic processes, but in this paper we focus on diffusion processes. Suppose $W(t)$ is a Wiener process and consider the following stochastic differential equation (SDE):

$$\mathrm{d}X_t = a(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W(t).$$

*Corresponding author. Email: khalil.shafie@unco.edu

This equation under some conditions has a unique continuous strong solution [8] called diffusion-type process. Diffusion process is of the form

$$X_t = X_0 + \int_0^t a(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W(s).$$

For example, let X_t denote the coordinates of a particle in a liquid at an instant t. The velocity of the motion liquid at the point x and the instant t is equal to $a(t, x)$. Also suppose that fluctuational movement of this particle is a Wiener process with 0 mean, then X_t is a diffusion-type process.[8]

Diffusion-type process is used for model building in social, physical, engineering, financial and medical sciences. In economics and insurance marketing, diffusion process is a standard instrument for modelling the swing of price. Practically when we use diffusion process for modelling a real process, we meet unknown parameters which need to be estimated. This is done by observing continuous realizations of the process or from discrete sample data. Most of the time the process is not observable continuously because of restrictions on measuring instrument or inability of observations at all the time points. The idea of parameter estimation for discretely observed diffusion models is discussed in the early work by Le Breton.[9] Furthermore estimation of unknown parameter in diffusion processes is discussed in [10–12]. Pedersen [13] offered a method for approximating maximum likelihood estimator (MLE) for diffusion processes and discussed consistency and asymptotic normality of that estimator.[14] Ait-Sahalia [15] studied nonlinear diffusion processes and presented a method to estimate the transition density function of these processes. Parakasa Rao [8] presented several different methods for parameter estimation of diffusion processes and also studied some properties of estimators such as consistency and asymptotic normality. Sorensen [16] presented his PhD thesis about Inference for diffusion process. He discussed estimation of diffusion parameters, and approximation of the score function for discretely observation and designed likelihood approximations for volatility models (special case of diffusion). Pedersen [17] used the diffusion process for modelling the nitrous oxide emission rate from the soil surface. A Markov-chain Monte Carlo methodology for analysis of the diffusion model with application in finance is presented by Eraker.[18] Estimating multi-dimensional unknown parameter of diffusion processes is studied by Biby and Sorenson.[19] Ait Sahalia [20] presented an approximation approach for the maximum likelihood estimation for diffusion processes. The numerical techniques for obtaining MLE of diffusion processes are presented in [21]. Biby et al. [22] studied some methods for estimating parameters when marginal distribution and auto-correlation are given. Chen et al. [23] discussed a test for diffusion models based on kernel estimation. Boukhetala and Guidoum [24] designed a R package, say Sim.DiffProc, which provides a simulation of diffusion processes and different methods of simulation of solutions for SDEs of the Ito's type, in financial modelling and other areas of applications.

As seen, much effort has been done by statisticians to analyse diffusion processes and test or estimate the unknown parameters. But a very important question is 'What do data say about the parameters?'. Neither the question nor its answer is to be found in the above works. In statistical hypothesis, a very important object is the support of data for statistical hypotheses. When do data organize evidence in favour of one hypothesis vis-a-vis another? When does a given set of observation support one hypothesis over another? In this paper, we will deal with this question and show that measuring the statistical evidence can be a competitor for available methods to analyse diffusion processes.

The structure of this article is as follows. Section 2 introduces the evidential inference and extends this issue to diffusion processes. Section 3 considers two hypotheses for a parametric diffusion process and discusses the calculation of evidential index. Some specification of sampling can impress the strength of the statistical evidence, Section 4 studies these effects through simulation, Section 5 applies the method to the Microsoft stock prices and Section 6 concludes.

2. Evidential inference for diffusion processes

2.1. *Evidential inference*

Sciences look into statistics for help in interpreting data. Scientists expect from statistics to make a clear image from the collected real data such that they can evaluate support of observations from the hypothesis they think may be true. One of the inference method in classical statistics is hypothesis testing. The Neyman–Pearson approach is often used for hypothesis testing. Despite of its logical structure and its strong mathematical fund, using it leads to some defect in statistical methods.

(A) The result of a Neyman–Pearson test is to accept or reject a hypothesis over another. This method answers ‘Yes’ or ‘No’ and all different theorems in this approach are designed and proved with the default that our goal is to reject or accept the null hypothesis. This method cannot evaluate the measure of support of data from H_1 over H_2.

(B) In Neyman–Pearson test, the probability of type one error is specified before test by researcher. Thus, accuracy of this test is not characteristic of the result.

Generally two types of accuracy could be defined for a statistical test, pre-experimental accuracy and post-experimental accuracy. Pre-experimental accuracy is the sole characteristic of the test regardless of the observation determined by the researcher and can vary from a researcher to another. But post-experimental accuracy is determined after observing and collecting data. In the Neyman–Pearson method, whereas the probability of the type one error is determined before the test, and directly affects the test result (rejection or acceptance of the hypothesis), only pre-experimental accuracy is considered. So the result of the test is influenced by the researcher's decision about the size of the test.

(C) Neyman–Pearson test is designed to help with decision-making problem, when the situation induces us to choose between two actions.[25] Suppose that we want to analyse the two hypotheses H_1 and H_2, and by that we want to decide which of the two actions A or B to take? The best answer for this question is obtained by Neyman–Pearson test.

Now, consider this question; Which hypothesis is supported more by the observations, H_1 or H_2? And how strong is this support? Neyman–Pearson method does not provide any answer for this question. This essential question had not been answered, maybe because for half a century (1930–1990), all statistical methods were dominated by the popular Neyman–Pearson theory. What do statistical data as evidence say about our hypotheses? This question is answered in evidential inference.

2.2. *The law of likelihood*

Evidential inference uses the law of likelihood as the base for interpretation of statistical data as evidence. The law of likelihood was proposed for the first time by Hacking.[26] If one hypothesis H_1 implies that a random variable X takes the value x with $f_1(x)$ probability while another hypothesis H_2 expresses that this probability is $f_2(x)$, then the observation $X = x$ is evidence supporting H_1 over H_2 if $f_1(x) > f_2(x)$ and the likelihood ratio $R(x) = f_1(x)/f_2(x)$ measures the strength of the evidence.

Likelihood ratio has a basic and important role in statistical inference. Its application for MLE calculation and solving problem for hypothesis testing is obvious. In evidential inference likelihood ratio measures the strength of statistical evidence for supporting a hypothesis over another.[27] Likelihood ratio that is close to 1 represents a weak evidence, and extreme ratio (very large or very close to 0) represents a strong support. It is logical that a large number k be chosen and when the likelihood ratio is greater than k or less than $1/k$ we count data as strong evidence, and when it takes value between $1/k$ and k we conclude that evidence is weak. The values $k = 8$ and 32 have been proposed as benchmark by Royall.[27] Also Jeffreys [28] and Edward [29] have proposed similar criteria. In this paper, we use Royall's benchmark.

2.3. *Misleading evidence*

Consider a set of observation x and the likelihood ratio $R(H_1, H_2, x) = f_1(x)/f_2(x)$ for two hypotheses H_1 and H_2, R can take any value from 0 to infinity. Suppose in reality H_2 is true, if R takes a value greater than $k = 8$, then the data are a strong evidence to support H_1 over H_2, so the evidence is strong but misleading. For a large number of k if $R > k$ when H_2 is true, data are misleading evidence. Symmetrically when H_1 is true, if $R < 1/k$ evidences are misleading.

In statistical hypotheses, using data as evidence was proposed by Royall [27] for the first time. Some other earlier issues are: probability of observing misleading evidence,[30] likelihood methods for measuring statistical evidence,[31] interpreting statistical evidence by imperfect models,[32] statistical evidence in sampling.[33] DeSantis [34] studied determination of sample size using statistical evidence. Emadi et al. [35] compared record data and random observations based on statistical evidence. Statistical evidence for regression models is studied by Blume [36] and Blume et al.[37] Tompson [38] compiled a book in the nature of statistical evidence. Evidential inference for record data is considered in [39]. Kateria and Balakrishnan [40] deal with statistical evidence in contingency table analysis. Hoch and Blume [41] studied statistical evidence for a cost-effectiveness analysis. Lastly mixture models in view of evidential inference is studied in [42].

The Universal Bound. In evidential analysis observing strong misleading evidence is possible but fortunately the probability of observing this improper evidence is controllable. There exists an upper bound for this probability, and we are able to compute it. This upper bound is true for all random samples observed from any arbitrary distribution, and hence is called universal.[30]

THEOREM 2.1 *Suppose under H_i hypothesis $(i = 1, 2)$, X is distributed with f_i density function, then the probability of observing misleading evidence is lower than $1/k$. Namely:*

$$\text{(A) } P\left(\frac{f_1(X)}{f_2(X)} > k \mid H_2\right) \leq \frac{1}{k} \qquad \text{(B) } P\left(\frac{f_1(X)}{f_2(X)} < \frac{1}{k} \mid H_1\right) \leq \frac{1}{k}$$

For a proof, see Royall.[30]

This inequality can be used for a sample from a diffusion process because the base of its proof is Markov inequality.

2.4. *Likelihood ratio for diffusion processes*

This paper is an effort to characterize diffusion processes through likelihood paradigm. Consider two statistical hypotheses about an unknown parameter in a diffusion process, we do not use Neyman–Pearson approach or P-value procedure. As a complement for current techniques, we propose measuring and illustrating observation as statistical evidence. In this framework, we begin by calculating the Radon–Nikodym derivative for two hypotheses about diffusion processes, of course the likelihood ratio is our intention. Afterwards we attempt to find the empirical distribution of the likelihood ratio. Finally, the probability of misleading evidence and weak evidence are discussed. Here, we study models in which the diffusion coefficient does not depend on θ.

A very important problem in relation with stochastic processes is how to calculate the likelihood ratio. Because, in abstract spaces, there exists no natural invariant Lebesgue measure which can be used to define the likelihood ratio, we let the Radon–Nikodym derivative of a hypothetical measure with respect to another play the role of the likelihood ratio.[2]

Suppose $P^{\mathrm{T}}_{\theta_1}$, $P^{\mathrm{T}}_{\theta_2}$ denote two probability function corresponding to H_1 and H_2 for stochastic process $\{X_t : 0 \leq t \leq T\}$. The likelihood ratio $R(X_t, \theta_1, \theta_2) = (\mathrm{d}P^{\mathrm{T}}_{\theta_1}/\mathrm{d}P^{\mathrm{T}}_{\theta_2})(X_t)$ is the Radon–Nikodym derivative. The likelihood function is defined as $L(\theta) = (\mathrm{d}P^{\mathrm{T}}_{\theta}/\mathrm{d}\mu)(X_t)$ where μ is a σ-finite measure relative to which all measures P^{T}_{θ} are absolute continuous.

Here, we discuss the diffusion-type process X_t of the form

$$\mathrm{d}X_t = a(t, X_t, \theta)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t,$$

where W_t is a standard Wiener process and only the draft coefficient depends on the parameter and the diffusion coefficient does not. Now suppose the following hypotheses for θ, $H_1 : \theta = \theta_1$, $H_2 : \theta = \theta_2$. We want to use evidential inference approach to determine and compute the support of observation from these two hypotheses. Consider time interval $[0, T]$, we want to record the process in this interval. As said, the likelihood ratio for this process is obtained through Radon–Nikodym derivative as

$$\frac{\mathrm{d}P^{\mathrm{T}}_{\theta}}{\mathrm{d}P^{\mathrm{T}}_{W}} = \exp\left(\int_0^{\mathrm{T}} \frac{a(t, X_t, \theta)}{\sigma^2(t, X_t)}\,\mathrm{d}X_t - \frac{1}{2}\int_0^{\mathrm{T}} \frac{a^2(t, X_t, \theta)}{\sigma^2(t, X_t)}\,\mathrm{d}t\right).$$

For a proof, see [8]. By Ito's formula,[43] the first term in this expression changes to

$$\int_0^{\mathrm{T}} a(t, X_t)\,\mathrm{d}X_t = F(T, X_T) - \int_0^{\mathrm{T}} f(t, X_t)\,\mathrm{d}t,$$

where $F(t, x) = \int_0^x a(t, y)\,\mathrm{d}y$, and $f(t, x) = (\mathrm{d}/\mathrm{d}t)F(t, x) + \frac{1}{2}(\mathrm{d}/\mathrm{d}x)a(t, x)$.

3. Measuring statistical evidence

In this section, we present a method for calculating the empirical distribution of the likelihood ratio of two hypotheses on a diffusion process, using observations under both hypotheses. Also, we use empirical distribution to obtain the probability of strong evidence (strong support from true hypothesis), weak evidence and misleading evidence (strong support from wrong hypothesis).

First, we generate a realization of the process under H_1, using Sim.DiffProc package of R. Next, we compute the likelihood ratio. We repeat this process M times so for every repetition we have a likelihood ratio value. Finally, we obtain the empirical distribution of the likelihood ratio.

Example 3.1 Suppose X_t is a diffusion process with equation: $\mathrm{d}X_t = \theta X_t\,\mathrm{d}t + \mathrm{d}W_t$, where θ is an unknown parameter and we have statistical hypotheses $H_1 : \theta = 2$, $H_2 : \theta = 3$. We want to find the statistical evidence in support of these two hypotheses.

First, we simulate the process assuming $H_1 : \theta = 2$, and record the observations in interval $[0, 2]$, with the time increment of $\Delta t = 0.001$ between two observations. The sample path is shown in Figure 1. Next, we obtain the likelihood function for both hypotheses. Since

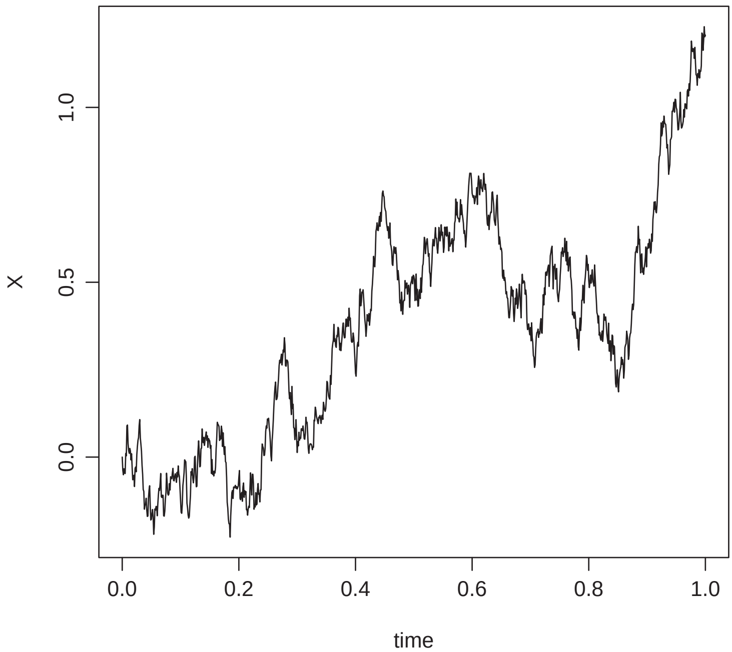

Figure 1. A realization of the diffusion process of the form $\mathrm{d}X_t = 2X_t\,\mathrm{d}t + \mathrm{d}W_t$.

$a(t, X_t, \theta) = \theta X_t$, and $\sigma(t, X_t) = 1$, the likelihood function is calculated as follows

$$L(\theta \mid X) = \exp\left[\int_0^{\mathrm{T}} \theta X_t\,\mathrm{d}X_t - \frac{1}{2}\int_0^{\mathrm{T}} (\theta X_t)^2\,\mathrm{d}t\right].$$

For computing stochastic integral of $\int_0^{\mathrm{T}} \theta X_t\,\mathrm{d}X_t$, using Ito's formula

$$F(t,x) = \frac{1}{2}\theta x^2, \quad \frac{\partial F(t,x)}{\partial t} = 0, \quad \frac{\partial a(t,x)}{\partial x} = \theta, \quad \text{and} \quad f(t,x) = \frac{\theta}{2}.$$

Now,

$$\int_0^{\mathrm{T}} \theta X_t\,\mathrm{d}X_t = F(T, X_T) - \int_0^{\mathrm{T}} f(t, X_t)\,\mathrm{d}t = \frac{1}{2}\theta X_T^2 - \int_0^{\mathrm{T}} \frac{\theta}{2}\,\mathrm{d}t,$$

and finally the likelihood function is

$$\exp\left[\frac{\theta}{2}(X_T^2 - T) - \frac{\theta^2}{2}\int_0^{\mathrm{T}} X_t^2\,\mathrm{d}t\right], \tag{1}$$

where X_T is the last observation. Since X_t is not a direct function of t, $\int_0^{\mathrm{T}} X_t^2\,\mathrm{d}t$ is not directly computable; so numerical method is necessary, and Riemann or Simpson approximation is useful. In this example, we use the Riemann method to approximate $\int_0^{\mathrm{T}} X_t^2\,\mathrm{d}t$ by $\sum_{t=0}^{N} X_t^2 \Delta t$, where $\Delta t = 0.001$. The results are $\log L(\theta_1) = 56.35$, $\log L(\theta_2) = 39.62$, and $R(\theta_1, \theta_2) = \exp(16.72) = 18255921$, so the observation are extremely strong evidence in favour of true hypothesis H_1.

Example 3.2 Consider the diffusion process in Example 1 with the same hypotheses. Now, we want to approximate the distribution of the likelihood ratio. First, all the above steps are repeated $M = 1000$ times; For each realization, we have a couple of likelihood function values $L(\theta_1, X)$

Table 1. Statistical evidence obtained from empirical distribution of the likelihood ratio.

	H_1	H_2
Strong	0.940	0.865
Misleading	0.054	0.129
Weak	0.006	0.006

and $L(\theta_2, X)$. Next, using these M couples we compute the likelihood ratios. Now, the empirical distribution of the likelihood ratio is computable. We obtain the probability of strong, weak and misleading supports.

Now all the above processes are repeated for a diffusion process with the same equation, except that $\theta = 3$. Thus again we have a distribution for the likelihood ratio with the difference that now H_2 is true. The results are given in Table 1. We can see from this table that when the data are produced with H_1 they are strong support for the true hypothesis (H_1) with the probability of 0.940, they are misleading with strong support from the wrong hypothesis with the probability of 0.054. Finally, the data are weak evidence (strong support for none of the hypotheses) with the probability of 0.006. The second column shows these probability for the data produced under H_2.

4. Simulation study

The strength of a realization of a diffusion process as statistical evidence can be affected by some sampling characteristics and/or hypotheses. In this section, we aim to study this effects by simulation. The items that we study their effects on strength of statistical evidence are: the difference between two simple hypotheses $\mid \theta_1 - \theta_2 \mid$, last time of sampling T, the time between observations or briefly sampling interval Δt, and the number of trajectories used for simulation M. At the end, we compare the Simpson approximation with the Riemann for the strength of the statistical evidence.

The criterion for evaluation is the strength of evidence and its stability. Whenever the probability of strong evidence is closer to 1 the evidence is stronger and more desirable. Weak evidence is undesirable and misleading evidence is the most unfavourable, thus smaller probability of them shows a stronger evidence.

Note that in order to see the effect of a parameter on our evidence, other parameters are intentionally kept fixed.

4.1. *Difference between two parameters*

Example 4.1 Suppose X_t is a diffusion process with equation: $\mathrm{d}X_t = \theta X_t \,\mathrm{d}t + \mathrm{d}W_t$, where θ is an unknown parameter. We let $\theta_1 = 5$ to be fixed and study the effect of changing θ_2 by a step increment from 1 to 10 on statistical evidence. The observations are generated under $H_1 : \theta = 5$ and $M = 300,\ T = 1,\ \Delta t = 0.01$, then we measure the strength of evidence. The results are given in Table 2. It is concluded that as distance between θ_1 and θ_2 increases, the probability of weak and misleading supports decreases, so we obtain stronger support for the true hypothesis. From Table 2, we can deduce for $|\theta_1 - \theta_2| \geq 3$ the evidence is adequate and stable, and we conclude that if $|\theta_1 - \theta_2| < 3$ we should set other components of sampling more strictly and accurately to have strong evidence.

Table 2. Statistical evidence for fixed $\theta_1 = 5$ and changing θ_2.

$\lvert\theta_1 - \theta_2\rvert$	0.25	0.5	1	1.5	2	3	4	5
Strong	0.00	0.02	0.02	0.07	0.63	0.95	0.98	0.99
Weak	0.03	0.07	0.08	0.11	0.13	0.05	0.02	0.01
Misleading	0.97	0.91	0.90	0.82	0.24	0.00	0.00	0.00

Table 3. Statistical evidence for different number of trajectories.

M	5	10	20	50	100	200	500	600	700
Strong	0.60	0.50	0.50	0.58	0.54	0.51	0.67	0.98	0.98
Weak	0.40	0.50	0.45	0.40	0.46	0.48	0.03	0.02	0.02
Misleading	0.00	0.00	0.05	0.02	0.01	0.00	0.00	0.00	0.00

4.2. *Number of trajectories*

Again consider the diffusion process in Example 3, with the hypotheses $H_1 : \theta = 2,\ H_2 : \theta = 3$. We study the stability of statistical evidence when we gradually increase the number of repetition in simulation from 5 to 700. Data are generated under $H_1 : \theta = 2,\ T = 1,\ \Delta t = 0.01$. The results are listed in Table 3. From this table, the statistical evidence for $M \geq 600$ is strong and stable.

4.3. *Sampling interval*

For the above diffusion process and hypotheses, we study the stability of statistical evidence with decreasing time interval between observations. We generated the data by simulation for $\theta = 2,\ T = 1,\ M = 300$. The result of simulation is given in Table 4.This table shows for $\Delta t \leq 0.001$, we have stable strong evidence.

4.4. *Last time of sampling*

The strength of statistical evidence depends on the end time of observing the process. We want to see the effect of T on strength of evidence. Here, the simulation is repeated for different T from 0.2 to 50 and $\theta = 2,\ \Delta t = 0.01,\ M = 300$. The evidential probabilities of different final time of observations are given in Table 5. According to these results, we conclude that $T > 2$ is suitable for strong stable evidence.

4.5. *The Simpson approximation*

Consider diffusion process $\mathrm{d}X_t = \theta X_t\,\mathrm{d}t + \mathrm{d}W_t$. As said in Section 2, we need numerical method for the likelihood ratio computation. Since the Simpson approximation is more accurate than the Riemann one, in this section, we obtain the integral by the Simpson method and compare the

Table 4. Statistical evidence for different sampling intervals Δt.

Δt	0.1	0.05	0.02	0.01	0.005	0.002	0.001	0.0005	0.0002	0.0001
Strong	0.02	0.70	0.82	0.84	0.84	0.89	0.83	0.87	0.92	0.88
Weak	0.22	0.26	0.16	0.14	0.16	0.10	0.16	0.12	0.07	0.10
Misleading	0.76	0.04	0.02	0.02	0.00	0.01	0.01	0.00	0.00	0.02

Table 5. Statistical evidence for different last time of sampling, *T*.

T	0.2	0.5	1.0	1.2	1.5	2	3	5	10	20	50
Strong	0.00	0.00	0.54	0.79	0.88	0.98	0.97	1.00	1.00	1.00	1.00
Weak	1.00	0.98	0.44	0.20	0.11	0.02	0.02	0.00	0.00	0.00	0.00
Misleading	0.00	0.02	0.02	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00

Table 6. Statistical evidence obtained from empirical distribution of the likelihood ratio, with Riemann and Simpson approximation.

	Riemann	Simpson
Strong	0.943	0.945
Misleading	0.007	0.005
Weak	0.005	0.005

results. For the Simpson approximation:

$$\int_0^{\mathrm{T}} f(x,t)\,\mathrm{d}t = \frac{\Delta t}{3}\left(f(X_0) + 4\sum_{i=1}^{N/2} f(X_{t_{2i-1}}) + 2\sum_{i=1}^{N/2-1} f(X_{t_{2i}}) + f(X_T)\right)$$

thus

$$\int_0^{\mathrm{T}} X_t^2\,\mathrm{d}t = \frac{\Delta t}{3}\left(X_0^2 + 4\sum_{i=1}^{N/2} X_{t_2 i-1}^2 + 2\sum_{i=1}^{N/2-1} X_{t_2 i}^2 + X_T\right)$$

We compute the likelihood ratio through both methods then compare the strength of statistical evidence. The results are given in Table 6. It shows that the strength of statistical evidence using the Simpson method is not different from that of the Riemann method.

5. Application on real data

Diffusion models are used intensively in mathematical finance for modelling interest rate, stock prices and option prices.[8, 16, 21] To demonstrate the methods discussed in this paper, we apply them to a real data set. We use daily Microsoft stock prices from 4 February to 25 July 2014 publicly available at www.nasdaq.com/symbol/msft/historical. The data are plotted in Figure 2. Let X_t be the daily stock price and $Y_t = X_t - X_0$. We fit a simple model of the form $\mathrm{d}Y_t = \theta Y_t\,\mathrm{d}t + \mathrm{d}W_t$ to the data. This model is chosen for the purpose of illustration and might not fit the data very well. The maximum likelihood approach is used first to estimate θ. To do this, the likelihood function (1) is maximized to obtain $\widehat{\theta}_{\mathrm{ML}} = 1.9$.

Now, we suppose the following two hypotheses for θ, $H_1 : \theta = 2$, $H_2 : \theta = 3$, we want to study these two hypotheses via both classical and evidential statistics approaches. Since the likelihood function is maximized at 1.9, we believe that H_1 is closer to the real θ. It is necessary to compute the likelihood value for both hypotheses, as in Example 3.1. The results are $\log L(\theta_1) = 26.377$ and $\log L(\theta_2) = 18.236$.

First, we test the hypotheses through the Neyman–Pearson method. For a test of the size of α, we need to specify a critical value C such that $\alpha = P(L(\theta_2)/L(\theta_1) > C \mid H_1)$. Then, if $L(\theta_2)/L(\theta_1) > C$ the hypothesis H_1 is rejected and accepted otherwise.

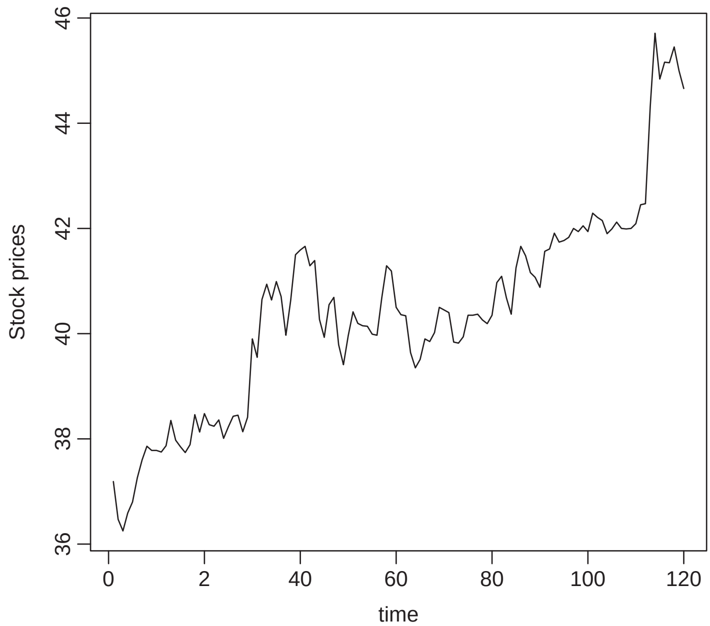

Figure 2. Daily observations of the Microsoft stock price.

For α equal to 0.01 and 0.05 the critical values are 0.37 and 0.18, respectively. These values are determined by simulation of the diffusion process under H_1 hypothesis. The result for the data is $R(\theta_2, \theta_1) = L(\theta_2)/L(\theta_1) = 0.0003$, so H_1 is not rejected at these significance levels.

Now, consider the hypotheses evidentially. For this, we need $R(\theta_1, \theta_2) = L(\theta_1)/L(\theta_2)$ and its distribution. We calculate the empirical distribution of $R(\theta_1, \theta_2)$, the statistical evidence is as follows, the probability of strong, misleading and weak evidence is 0.93, 0.00 and 0.07, respectively. Finally, the value of $R(\theta_1, \theta_2)$ for the observation is computed equal to 3431.073 which is so much greater than 8 or 32, and it is concluded that the observation support H_1 strongly.

6. Conclusion

To analyse diffusion processes much attempt has been made and numerous statisticians studied this area. The results are based upon Fisher approach, Neyman–Pearson theory or Bayesian procedure, but none of them are good enough to answer the question 'What do these data say?' when we have statistical hypotheses. We proposed using a likelihood paradigm to answer this question. We used the Radon–Nikodym derivative to obtain the likelihood ratio. Also we obtained the empirical distribution of likelihood ratio via simulation. Furthermore, we measured the support of data in favour of a hypothesis over another. This study could be a good complement for classical inference of diffusion processes.

Since some specification of sampling can impress the strength of the statistical evidence, we evaluated these effects through simulation studies. From the result of these simulations, we concluded that, although the misleading evidence is possible in this context, but fortunately we can control them, and by adjusting the items of sampling we can reduce the probability of observing misleading evidence desirably and therefore have favourably strong evidence.

In this paper, parametric evidential inference for diffusion processes was discussed, where only the draft coefficient depends on the parameter and the diffusion coefficient does not. The extension of the results to the more general diffusion models is a good field of study for future, but it may be more complicated because the computation of the likelihood function is not straightforward in general cases and some approximation methods are necessary.

Acknowledgments

The authors would like to thank the unknown referee for the critical comments and suggestions which helped to improve the quality of the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Grenander U. Stochastic processes and statistical inference. Ark Mat. 1950;1:195–277.
[2] Grenander U. Abstract inference. New York: Wiley; 1981.
[3] Anderson TW, Goodman LA. Statistical inference about Markov chains. Ann Math Stat. 1957;28:89–110.
[4] Billingsley P. Statistical inference for Markov processes. Chicago: University of Chicago Press; 1961.
[5] Hajek J. On linear statistical problems m stochastic processes. Cz Math J. 1962;12:404–444.
[6] Rao MM. Inference in stochastic processes, I. Theory Prob Appl. 1963;VTII:266–281.
[7] Liptser RS, Shiryayev AN. Statistics of random processes, Vol. 1. New York: Springer; 1977.
[8] Parakasa Rao BLS. Statistical inference for diffusion type processes. London: Arnold/New York: Oxford University Press; 1999.
[9] Le Breton A. Parameter estimation in a linear stochastic differential equation. In: J. Kožešnik, editor. Transactions of the Seventh Prague conference on information theory, statistical decision functions, random processes and of the 1974 European meeting of statisticians. Dordrecht: Springer, 1977. p. 353–366. .
[10] Kutoyants YA. Estimation of the drift coefficient parameter of a diffusion in the smooth case. Theory Probab Appl. 1977;22:399–406.
[11] Yoshida N. Estimation for diffusion processes from discrete observation. J Multivariate Anal. 1992;41:220–242.
[12] Hoffmann M. Adaptive estimation in diffusion processes. Stoch Process Appl. 1999;79:135–163.
[13] Pedersen AR. Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes. Bernoulli. 1995;1:257–279.
[14] Pedersen AR. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. Scand J Statist. 1995;22:55–71.
[15] Ait-Sahalia Y. Transition densities for interest rate and other nonlinear diffusions. J Financ. 1999;54:1361–1395.
[16] Sorensen H. Inference for diffusion processes and stochastic volatility models, Ph.D. thesis, University of Copenhagen; 2000 [cited 2015 Jan 13]. Available from: http://www.math.ku.dk/noter/filer/phd00hs.pdf
[17] Pedersen AR. Estimating the nitrous oxide emission rate from the soil surface by means of a diffusion model. Scand J Statist. 2000;27:385–403.
[18] Eraker B. MCMC analysis of diffusion models with application to finance. J Bus and Econom Statist. 2001;19:177–191.
[19] Bibby BM, Sorensen M. Simplified estimating functions for diffusion models with a high-dimensional parameter. Scand J Statist. 2001;28:99–112.
[20] Ait-Sahalia Y. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. Econometrica. 2002;70:223–262.
[21] Durham GB, Gallant AR. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. J Bus Econ Statist. 2002;20:297–338.
[22] Bibby BM, Skovgaard IM, Sorensen M. Diffusion-type models with given marginal distribution and autocorrelation function. Bernoulli. 2005;11:191–220.
[23] Chen SX, Gao J, Tang CY. A test for model specification of diffusion processes. Ann Stat. 2008;36:167–198.
[24] Boukhetala K, Guidoum A. Sim.DiffProc: A package for simulation of diffusion processes in r, Preprint submitted to Journal of Statistical Software; 2011 [cited 2015 Jan 13]. Available from: https://hal.archives-ouvertes.fr/hal-00629841
[25] Neyman J. A first course in probability and statistics. New York: Henry Holt and Company; 1950.
[26] Hacking I. Logic of statistical inference. New York: Cambridge University Press; 1965.
[27] Royall R. Statistical evidence: a likelihood paradigm. London: Chapman and Hall; 1997.
[28] Jeffreys H. Theory of probability. New York: Oxford University Press; 1961.
[29] Edwards AW. Likelihood. Baltimore (MD): Johns Hopkins University Press; 1992. expanded ed.
[30] Royall R. On the probability of observing misleading statistical evidence. JASA Theory Methods. 2000;95:760–780.

[31] Blume JD. Likelihood methods for measuring statistical evidence. Statist Med. 2002;21:2563–2599.
[32] Royall R, Tsou T-S. Interpreting statistical evidence by using imperfect models: robust adjusted likelihood functions. J R Statist Soc B. 2003;65:391–404. part 2.
[33] Bjornstad JF. Likelihood and statistical evidence in survey sampling. Stat Transit. 2003;6:23–31.
[34] De Santis F. Statistical evidence and sample size determination for Bayesian hypothesis testing. J Statist Plan Inference. 2004;124:121–144.
[35] Emadi M, Ahmadi J, Arghami NR. Comparison of record data and random observations based on statistical evidence. Statist Pap. 2006;48:1–21.
[36] Blume JD. How to choose a working model for measuring the statistical evidence about a regression parameter. Int Statist Rev. 2005;73:351–363.
[37] Blume JD, Su L, Olveda RM, McGarvey ST. Statistical evidence for GLM regression parameters: a robust likelihood approach. Stat Med. 2007;26:2919–2936.
[38] Thompson B. The nature of statistical evidence. New York: Springer; 2007.
[39] Arashi M, Emadi M. Evidential inference based on record data and inter-record times. Statistical Papers. 2008;49:291–301.
[40] Kateri M, Balakrishnan N. Statistical evidence in contingency tables analysis. J Stat Plan Infer. 2008;138:873–887.
[41] Hoch JS, Blume JD. Measuring and illustrating statistical evidence in a cost-effectiveness analysis. J Health Econ. 2008;27:476–495.
[42] Emadi M. Mixture models in view of evidential analysis. Communications in Statistics-Simulation and Computation. 2013;42:1824–1835.
[43] Kallianpur G. Stochastic filterinbg theory. New York: Springer; 1980.