Séminaire Mathématique de Béjaïa (LaMOS)
Volume 9, 2010, pp. 45 - 51

8

Some statistical aspects to take into account when studying strong stability of stochastic models

Aicha BARECHE

Laboratoire de Modélisation et d'Optimisation des Systèmes (LAMOS)
Université de Béjaïa, Béjaïa 06000, Algérie
Tél. (213) 34 21 51 88

Résumé The aim of this work is to review and revisit some statistical techniques and to discuss their application to some stochastic models in the strong stability approximation sense.

Mots clés : Strong stability, Boundary kernel estimation, Heavy-tailed distribution, Quantile estimation, Risk measure.

8.1 Introduction

In [1], we considered the use of kernel density involving some boundary correction techniques for the study of the strong stability of the $M/M/1$ system. In this work, we reconsider these techniques, review some other techniques, give some related fields of such applications, and make a reflexion in the strong stability sense.

8.2 Kernel density estimation

One of the most popular and widely studied class of nonparametric estimators of a density f is the so called kernel class of estimators. If $X_1, ..., X_n$ is a sample from a random variable X having the probability density function (pdf) f and a distribution function (cdf) F, the Parzen-Rosenblatt kernel estimate [13, 11] of the density $f(x)$ for each point $x \in \mathbb{R}$ is given by

$$f_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K(\frac{x - X_j}{h_n}), \tag{8.1}$$

where K is a symmetric density function called kernel and h_n is the bandwidth.

8.3 Boundary bias correction

Classical symmetric kernels work well when estimating densities with non-bounded support. However, when these latter are defined on the positive real line $[0, \infty[$, without cor-

rection, the kernel estimates suffer from boundary effects since they have a boundary bias ($\mathbb{E}(f_n(0)) = \frac{1}{2}f(0) + o(h_n)$). In fact, using a fixed symmetric kernel is not appropriate for fitting densities with bounded supports as a weight is given outside the support.
Additionally, standard kernel methods yield wiggly estimation in the tail of the distribution (especially for heavy-tailed distributions) since the mitigation of the boundary bias leads to favor a small bandwidth which prevents pooling enough data.
Several techniques have been introduced to get a better estimation either on the border or in the tail. Some of them propose the use of particular kernels or bandwidths :

– Reflection method (mirror image modification) [14, 15] ;
– Boundary kernel method [12] ;
– Transformation method (transformed kernel) [10, 9] ;
– Pseudo-data method [5].

The main criticism addressed to these approaches is that a number of them are quite complicated (since implementation is a nontrivial exercise) and thus difficult to work with, both numerically and analytically. Also, they allow the corrected estimator to become negative.
In the last decade, to remedy these problems, other techniques propose the use of estimators based on flexible kernels (asymmetric kernels and smoothed histograms). They are very simple in implementation, free of boundary bias, always non-negative, their support matches the support of the probability density function to be estimated, and their rate of convergence for the mean integrated squared error is $0(n^{-4/5})$. We can cite :

8.3.1 Asymmetric kernel estimators

A simple idea for avoiding boundary effects is using a flexible kernel, which never assigns a weight out of the support of the density function and which corrects automatically and implicitly the boundary effects. The first category of the flexible kernels consists of the asymmetric kernels [3, 4] defined by the form

$$\hat{f}_b(x) = \frac{1}{n}\sum_{i=1}^{n} K(x,b)(X_i), \tag{8.2}$$

where b is the bandwidth and K is the asymmetric kernel.

Gamma kernels

The asymmetric kernel K can be taken as a Gamma density function K_G with the parameters $(x/b+1, b)$ [4] given by

$$K_G(\frac{x}{b}+1,b)(t) = \frac{t^{x/b}e^{-t/b}}{b^{x/b+1}\Gamma(x/b+1)}, \quad t \in [0,\infty], \tag{8.3}$$

where $\Gamma(.)$ is the Gamma function.

Beta kernels

The asymmetric kernel K can be taken as a beta density function $K_{\beta 1}$ with the parameters $(x/b, (1-x)/b)$ [3] such that

$$K_{\beta 1}(b,x)(t) = \frac{t^{b-1}(1-t)^{x-1}}{B(b,x)}, \quad t \in [0,1], \tag{8.4}$$

where $B(.)$ is the beta function.

8.3.2 Smoothed histograms

The second category of the flexible kernels is constituted of smoothed histograms [6, 7] defined by the form

$$\hat{f}_k(x) = k \sum_{i=0}^{+\infty} \omega_{i,k} p_{ki}(x), \tag{8.5}$$

where the random weights $\omega_{i,k}$ are given by

$$\omega_{i,k} = F_n(\frac{i+1}{k}) - F_n(\frac{i}{k}), \tag{8.6}$$

where F_n is the empirical distribution function, k is the smoothing parameter and $p_{ki}(.)$ can be taken as a Poisson distribution function with parameter kx,

$$p_{ki}(x) = e^{-kx}\frac{(kx)^i}{i!}, \quad i = 0, 1, ... \tag{8.7}$$

8.4 Heavy-tailed distributions

A class of distributions that is often used to capture the characteristics of highly-variable stochastic processes (i.e., more variable than the exponential distribution) is the class of heavy-tailed distributions. In the literature, different definitions of heavy-tailed like distributions exist.

Definition We refer to a distribution as heavy-tailed if its coefficient of variation is larger than the one of the exponential distribution.

Definition A distribution is heavy-tailed if its *complementary cumulative distribution* (CCDF), often referred to as the *tail*, $F^c(t) = 1 - F(t)$, where $F(t)$ is the CDF, decays slower than exponentially, i.e., there is some $\gamma > 0$ such that

$$\lim_{t\longrightarrow\infty}\exp(\gamma t)F^c(t)\longrightarrow\infty. \tag{8.8}$$

Definition A distribution has short tail if its CCDF $F^c(t)$, decays exponentially or faster, i.e., there is some $\gamma > 0$ such that

$$\lim_{t\longrightarrow\infty}\exp(\gamma t)F^c(t)\longrightarrow 0. \tag{8.9}$$

Definition A light-tailed distribution is one in which the extreme portion of the distribution (the part farthest away from the median) spreads out less far relative to the width of the center (middle 50) of the distribution than is the case for the normal distribution.

8.4.1 Kernel density estimation of heavy-tailed distributions

Kernel density estimation of heavy-tailed distributions has been studied by several authors. Some of them proposed estimators based on transformation of the original variable. [2] consider transformations based on the Champernowne generalized distribution.
An alternative way to obtain a kernel density estimator for heavy-tailed distributions is using the variable kernel estimator, which consists in selecting a different bandwidth parameter depending on the point where the density is being estimated [8].

8.5 Estimating quantiles and related risk measures

Quantiles often seem to be the natural thing to estimate in many applications when the underlying distribution is heavy-tailed. Furthermore, estimates of quantiles of the loss distribution in actuarial science and financial risk management are a fundamental part of the business. In this context, loss distribution is a positive variable which describes the probability distribution of payment to the insured. Quantile (called extreme Value-At-Risk (VaR) from the actuarial point of view) represents the worst loss ever on a target horizon that cannot be exceeded with a given level of confidence.

8.5.1 Estimating quantiles

Let X be a non-negative random variable admitting a continuous probability density function (pdf) f (called loss distribution in actuarial science and financial risk management), and cumulative distribution function (cdf) F_X. Given a probability level $p \quad (0 < p < 1)$, the pth quantile of X based on a random sample $X_1, ..., X_n$ is defined as

$$Q(X,p) = inf[x : F_X(x) \geq p] = F_X^{-1}(p). \tag{8.10}$$

Nonparametric estimation

For continuous distribution $q(\alpha) = F_X^{-1}(\alpha)$ (the quantile function is simply the inverse of the cumulative distribution function), thus, a natural idea would be to consider $\hat{q}(\alpha) = \hat{F}_X^{-1}(\alpha)$, for some nonparametric estimation of F_X.

Definition The empirical cumulative distribution function F_n, based on sample $X_1, ..., X_n$ is

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1(X_i \leq x). \tag{8.11}$$

Definition The kernel based cumulative distribution function, based on sample $X_1, ..., X_n$ is

$$\hat{F}_n(x) = \frac{1}{nh} \sum_{i=1}^{n} \int_{-\infty}^{x} k\left(\frac{X_i - t}{h}\right) dt = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right), \tag{8.12}$$

where $K(x) = \int_{-\infty}^{x} k(t)dt$, k being a kernel and h the bandwidth.

Smoothing nonparametric estimators

a) Implicit class Find a smooth estimator for F_X, and then find (numerically) the inverse. The α-quantile is defined as the solution of $F_X \circ q_X(\alpha) = \alpha$.

If $\hat{F}_n$ denotes a continuous estimate of F, then a natural estimate for $q_X(\alpha)$ is $\hat{q}_X(\alpha)$ such that $\hat{F}_X \circ \hat{q}_X(\alpha) = \alpha$, obtained using, for example, Gauss-Newton algorithm.

b) Explicit class Consider a linear combination of order statistics.

$$Q_n(p) = \sum_{i=1}^{n} \left[\int_{\frac{(i-1)}{n}}^{\frac{i}{n}} k\left(\frac{t-p}{h}\right) dt \right] X_{(i)} = \sum_{i=1}^{n} \left[K\left(\frac{\frac{i}{n} - p}{h}\right) - K\left(\frac{\frac{i-1}{n} - p}{h}\right) \right] X_{(i)}. \tag{8.13}$$

The idea is to give more weight to order statistics $X_{(i)}$ such that i is closed to pn.

Transforming observations

Given a random variable X, if H is a strictly increasing function, then the p-quantile of $H(X)$ is equal to $H(Q(X, p))$.

Thus, an idea can be to transform initial observations $X_1, ..., X_n$ into a sample $Y_1, ..., Y_n$ where $Y_i = H(X_i)$ taking values in $[0, 1]$, and then to use a beta-kernel based estimator. If $H : \mathbb{R} \longrightarrow [0, 1]$, then

$$Q_n(X, p) = H^{-1}(Q_n(Y, p)). \tag{8.14}$$

In theory, any transformation $H : \mathbb{R} \longrightarrow [0, 1]$ should work. But [2] suggested to set $Y_i = H(X_i)$ where H is the Champernowne generalized distribution which is suitable when modeling insurance claims, i.e. positive variables.

8.5.2 Risk measures

A very general class of risk measures can be defined as follows

$$\mathcal{R}_g(X) = \int_0^1 F_X^{-1}(1-u)dgu, \tag{8.15}$$

where g is a distortion function, i.e. increasing with $g(0) = 0$ and $g(1) = 1$.

8.6 Conclusion

In this paper, we presented some statistical techniques to take into consideration in the analysis of some stochastic processes. It will be interesting to apply these techniques in the study of strong stability of some stochastic models (especially when a law or its density function is unknown). For example in, queueing systems (like $G/M/1$ or $M/G/1$), reliability models and risk models, since they involve :

- computing a variation distance which depends on the unknown density to estimate ;
- computing a reliability function ($\bar{F} = 1 - F$) which depends on the cdf function F ;
- Computing a risk measure which depends on F^{-1} and often involve heavy-tailed distributions.

Références

1. A. Bareche, and D. Aïssani. Kernel density in the study of the strong stability of the $M/M/1$ queueing system. Operations Research Letters, 36 : 535–538, 2008.
2. T. Buch-Larsen, J.P. Nielsen, M. Guillen, and C. Bolance. Kernel density estimation for heavy-tailed distribution using the Champernowne transformation. Statistics, 6 :503-–518, 2005.
3. S.X. Chen. Beta kernel estimators for density functions. Computational Statistics and Data Analysis, 31 :131–145, 1999.
4. S.X. Chen. Probability Density Function Estimation Using Gamma Kernels. *Ann. Inst. Statis. Math.*, 52 :471–480, 2000.
5. A. Cowling, and P. Hall. On pseudodata methods for removing boundary effects in kernel density estimation. *J. Roy. Statist. Soc. Ser.*, B 58 :551–563, 1996.
6. N. Gawronski, and U. Stadtmüller. On Density Estimation by Means of Poisson's Distribution. *Scand. J. Statist.*, 7 :90–94, 1980.
7. N. Gawronski, and U. Stadtmüller. Smoothing histograms by means of lattice and continuous distributions. Metrika, 28 :155–164, 1981.
8. M.C. Jones. Variable kernel density estimation. Australian Journal of Statistics, 32 :361–371, 1990.
9. J.S. Marron, and D. Ruppert. Transformations to reduce boundary bias in kernel density estimation. *J. Roy. Statist. Soc. Ser.*, B 56 :653–671, 1994.
10. M.P. Wand, J.S. Marron and D. Ruppert. Transformations in density estimation (with dicussion). *J. Amer. Statist. Assoc.*, 86 :343–361, 1991.
11. E. Parzen. On estimation of a probability density function and mode. *Ann. Math. Stat.*, 33 :1065–1076, 1962.

12. J. Rice. Boundary modification for kernel regression. *Commun. Statist. Theory Meth.*, 13 :893–900, 1984.
13. M. Rosenblatt. Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.*, 27 :832–837, 1956.
14. E.F. Schuster. Incorporating support constraints into nonparametric estimation of densities. *Commun. Statist. Theory Meth.*, 14 :1123–1136, 1985.
15. B.W. Silverman. Density Estimation for Statistics and Data Analysis. Chapman and Hall, London, 1986.