ON THE ESTIMATION OF PROBABILITY OF INITIAL DISTRIBUTION DYNAMICS ON SAMPLE AT THE END OF INTERVAL

Elizbar Nadaraya[1], Petre Babilua[2], Grigol Sokhadze[3]

I. Javakhishvili Tbilisi State University, Tbilisi, Georgia
[1]*elizbar.nadaraya@tsu.ge*, [2]*petre.babilua@tsu.ge*, [3]*giasokhi1@i.ua*

Abstract. In this paper we formulate the problem of estimating the density of a random value which is the initial value of some dynamics. The dynamics is given in the form of a differential equation whose solution is observable at the end of an interval. Applying the technique of measure (distribution) transformation along the integral curve in combination with kernel type estimates, we present the procedure of density estimation. We call such a problem the problem of estimation by indirect observation data.

Let the differential equation of first order

$$y'(t) = f(t, y(t)) \tag{1}$$

be given on the interval $[0,T]$ and consider the Cauchy problem $y(0) = X$, where X is a random value with the unknown distribution density $p(x)$. Assume that the problem posed has for (1) a unique solution with probability $1 - y(t)$ which is certainly a random process. Observations of this process are accessible for a statistician only at the end of the interval – at. at the point T. These observations correspond to an inaccessible sample from X, i.e., we have the observations (sample) $y_1(T), y_2(T), \dots, y_n(T)$. Using these observations, it is required to estimate $p(x)$. If $f(t,x) \equiv 0$, then $y(t) = X$ and the problem reduces to a problem formulated in classical terms.

It is obvious that along the integral curve the behavior of the density $p(x)$ must be sufficiently regular and hence the reconstruction must be possible. Therefore it is this question that we will deal with in the first place.

Let $\{\Theta, \mathrm{F}\}$ be a measurable space, $\mathrm{M} = \mathrm{M}(\mathrm{F})$ be a space of real-valued σ-additive functions of sets on F with the norm

$$\|\mu\| = (Var\,\mu)(\Theta) = \mu^+(\Theta) - \mu^-(\Theta),$$

where $\mu = \mu^+ - \mu^-$ is – Khan's decomposition. M is a Banach space whose elements we call measures.

Let us consider the family of measures $\mu_t \in \mathrm{M}$, where t is a real-valued parameter $0 \le t \le T < \infty$. It is assumed that the function $t \mapsto \mu_t(A)$ is continuous for each fixed $A \in \mathrm{F}$. The convergence on sets from F is a weak convergence in M (see[1]). Therefore the considered family is bounded in M: $\sup_{0 \le t \le T} \|\mu_t\| < \infty$.

Suppose that for each $A \in \mathrm{F}$ the function $\mu_t(A)$ is differentiable with respect to a parameter $t \in (0,T)$ and has one-sided derivatives for $t = 0$ and $t = T$. Then a measure

$$\nu_t = \mu_t' : A \mapsto \frac{d\mu_t(A)}{dt}$$

is called a derivative of the measure μ_t with respect to the parameter. From the representation

$$\mu_{t_2} - \mu_{t_1} = \int_{t_1}^{t_2} \nu_\tau(d\tau)$$

we obtain continuities with respect to variation of the family μ_t:

$$\left\| \mu_{t_2} - \mu_{t_1} \right\| \le C \cdot \left| t_2 - t_1 \right|.$$

If the measure $\nu_t = \mu_t'$ is absolutely continuous with respect to μ_t, then the Radon-Nykodim density

$$\rho(t,x) = \frac{d\nu_t}{d\mu_t}(x)$$

is called the logarithmic derivative of the family of measures μ_t.

Let $\mathrm{B}(\mathsf{F})$ be the Banach space of bounded measurable functions with a uniform measure. The differentiability of μ_t in the sense described above is equivalent to the relation

$$\frac{d}{dt}\int_\Theta \varphi(x)\mu_t(dx) = \int_\Theta \varphi(x)\mu_t'(dx)$$

for each $\varphi \in \mathrm{B}(\mathsf{F})$, whereas the existence of the logarithmic derivative of the family of μ_t – to the relation

$$\frac{d}{dt}\int_\Theta \varphi(x)\mu_t(dx) = \int_\Theta \varphi(x)\rho(t,x)\mu_t(dx), \quad \varphi \in \mathrm{B}(\mathsf{F}).$$

For example, let μ be some measure, $\alpha(t,x)$ be continuously differentiable with respect to t, real-valued positive bounded function and $\alpha(T,x) \equiv 1$. Then the family of measures

$$\mu_t(A) = \int_A \alpha(t,x)\mu(dx), \quad t \in [0,T],$$

possesses a logarithmic derivative of the form

$$\rho(t,x) = \frac{\dfrac{\partial \alpha(t,x)}{\partial t}}{\alpha(t,x)}.$$

Clearly, hence we can write

$$\alpha(t,x) = e^{\int_0^t \rho(\tau,x)d\tau}, \quad 0 \le t \le T.$$

The following converse statement is important.

Lemma. *Let on the measurable space $\{\Theta, \mathsf{F}\}$ the family of measures $\mu_t \in \mathsf{M}$, $t \in [0,T]$, possess the logarithmic derivative $\rho(t,x)$ for which the following conditions are fulfilled:*

1) $\rho(t,x)$ *is continuous with respect to* t, μ_t *a. e.;*

2) $\rho(t,x)$ *is* σ *bounded in a sense that there exists a partitioning* $\Theta = \bigcup_{j=1}^{\infty} \Theta_j$ *such that* $\rho(t,x)$ *is bounded on each* $[0,T] \times \Theta_j$.

Then every two measures μ_s and μ_τ, $s,\tau \in [0,T]$ of the considered family are equivalent and

$$p(s,\tau,x) = \frac{d\mu_s}{d\mu_\tau}(x) = e^{\int_\tau^s \rho(t,x)dt}. \quad (2)$$

Proof. It suffices to show that for any $\varphi \in B(F)$ the expression

$$\int_\Theta \varphi(x)p(t,\tau,x)\mu_\tau(dx) = \psi(t,\tau)$$

does not depend on τ. Then for $t = \tau$, for the indicator function $\varphi(x) = I_A(x)$ we obtain (2). Let us show that $\frac{d}{d\tau}\psi(t,\tau) = 0$.

This formally follows from the equalities

$$\frac{\partial p(t,\tau,x)}{\partial \tau} = -\rho(\tau,x) \text{ and } \frac{d}{d\tau}\mu_\tau = \rho(\tau,x)\mu_\tau$$

and therefore we have only to verify the validity of term-by-term differentiation with respect to the parameter under the integral sign.

Let us consider the relation

$$\frac{1}{\varepsilon}\left\{\int_\Theta \varphi(x)p(t,\tau+\varepsilon,x)\mu_{\tau+\varepsilon}(dx) - \int_\Theta \varphi(x)p(t,\tau,x)\mu_\tau(dx)\right\} =$$

$$= \int_\Theta \varphi(x)p(t,\tau,x)\frac{\mu_{\tau+\varepsilon}(dx) - \mu_\tau(dx)}{\varepsilon} + \int_\Theta \varphi(x)\frac{p(t,\tau+\varepsilon,x) - p(t,\tau,x)}{\varepsilon}\mu_\tau(dx) +$$

$$+ \int_\Theta \varphi(x)\frac{p(t,\tau+\varepsilon,x) - p(t,\tau,x)}{\varepsilon}[\mu_{\tau+\varepsilon} - \mu_\tau](dx).$$

Let at first the function $\rho(t,x)$ be bounded on $[0,T]\times\Theta$. Then in each of the three summands the integrand is bounded and we can pass to the limit as $\varepsilon \to 0$. The last summand tends to zero since

$$\|\mu_{\tau+\varepsilon} - \mu_\tau\| \to 0 \text{ for } \varepsilon \to 0.$$

In the general case, under the considered conditions we can write $\Theta = \bigcup_{j=1}^{\infty}\Theta_j$ so that on each Θ_j the function $\rho(t,x)$ is bounded and then for each j

$$\frac{d}{d\tau}\int_{\Theta_j} \varphi(x)p(t,\tau,x)\mu_\tau(dx) = 0.$$

Let us return to the equation (1)

$$y'(t) = f(t,y(t)), \quad y(0) = X$$

with a random initial condition for X which possesses the unknown density $p(x)$.

Assume that

(f) $f(t,x)$ is a continuous function of its variables that has a continuous derivative $f'_x(t,x)$.

The solution of the problem (1) exists, is unique and is a random process with differentiable trajectories with probability 1. Let μ_t be the probability distribution of the process $y(t)$ at the point t. It is clear that

$$\mu_0(A) = \int_A p(t)dt, \quad A \in \mathrm{B}[0,T],$$

where $\mathrm{B}[0,T]$ denotes the Borel σ-algebra of subsets $[0,T]$.

Theorem. *Let $p(t) > 0$ and the condition (f) be fulfilled. Then the family of μ_t possesses the a logarithmic derivative.*

Proof. Denote by $S_{t\tau}$ the invertible evolutionary family related to the problem (1):

$$y(t) = S_{t\tau}y, \quad S_{tt} = I, \quad S_{t\tau} \circ S_{\tau s} = S_{ts}, \quad \frac{\partial S_{t\tau}y}{\partial t} = f(t, S_{t\tau}y), \quad 0 \le s \le \tau \le t \le T.$$

Then

$$\mu_t(A) = P(y(t) \in A) = P(S_{t0}X \in A) = P\left(X \in S_{t0}^{-1}A\right) =$$
$$= \int_{S_{t0}^{-1}A} p(s)ds = \int_A p\left(S_{t0}^{-1}\tau\right)\left(S_{t0}^{-1}\right)'(\tau)d\tau.$$

Therefore μ_t possesses a probability distribution density and

$$p_t(x) = p\left(S_{t0}^{-1}x\right)\frac{\partial\left(S_{t0}^{-1}x\right)}{\partial x}$$

or

$$p(x) = \frac{\partial S_{t0}x}{\partial x}p_t(S_{t0}x). \tag{3}$$

It is likewise easy to calculate the logarithmic derivative of the measure μ_t:

$$\rho(t,x) = \frac{\partial}{\partial t}\ln p\left(S_{t0}^{-1}x\right)\frac{\partial S_{t0}^{-1}x}{\partial x}.$$

Let the solution of the problem (1) be observed at the point T: $y_1(T), y_2(T), \ldots, y_n(T)$ be the corresponding sample. We must construct an estimate for the density $p(x)$ of a random value X.

Let $K(x)$ be the function possessing the following properties:

(k) $K(x)$ is a continuous, bounded, integrable, positive function defined on R and such that $\int_R K(x)dx = 1$.

Let us consider the kernel type estimate of the density $p_T(x)$

$$\hat{p}_T(x) = \frac{1}{nh_n}\sum_{j=1}^{n}K\left(\frac{x - y_j(T)}{h_n}\right), \tag{4}$$

where the sequence $\{h_n\}_{n=1}^{\infty}$ satisfies the condition:

(h) h_n is a positive, converging to zero, sequence of real-valued numbers, for which $nh_n \to \infty$.

Then, as is well known (see [2]), in the conditions (k) and (h), we have that $\hat{p}_T(x)$ is the consistent estimate for the density $p_T(x)$. Therefore, according to the formulas (3) and (4), for the estimate $p(t)$ we take

$$\hat{p}(x) = \frac{\partial(S_{T0}x)}{\partial x}\hat{p}_T(S_{T0}x). \tag{5}$$

If the integral flux corresponding to (1) is known – S_{t0}, then the formula (5) gives us the desired estimate. However the explicit form of this family is rarely constructed. Hence we have to look for other possibilities.

Let us consider the determined Cauchy problem

$$y'(t) = f(t, y(t)), \quad y_0 = y(0) = x$$

and construct for it the sequence $\varphi_n(x)$ which uniformly converges to a solution of the Cauchy problem – $y(t)$ (such is for example the Picard procedure: $\varphi_n(t) = x + \int_0^t f(s, \varphi_{n-1}(s))dx$, $n = 1,2,\ldots$, $\varphi_0(t) = y_0$). As an approximation to $S_{T0}x$ we can take $S_{T0}^n x = \varphi_n(T)$.

We finally come to a conclusion that

$$\hat{p}_n(x) = \frac{\partial(\varphi_n(T))}{\partial x} \hat{p}_T(\varphi_n(T)). \tag{6}$$

Theorem. *Let for the Cauchy problem* (1) *the conditions* (f), (k) *and* (h) *be fulfilled, where* X *is a random value with the unknown positive density* $p(x)$. *Then the consistent estimate of the density* $p(x)$ *is given by the formula* (6).

References

1. Danford N., Schwartz J.T. Linear operators. A general theory. *Mir Publ. House, Moscow*, 1962.
2. Absava R.M., Nadaraya E.A. Some problems of the theory of nonparametric estimation of functional characteristics of the observations distribution law. *Tbilisi University Press, Tbilisi*, 2005.