On ADF Goodness of Fit Tests for Stochastic Processes

Yury A. Kutoyants
Université du Maine, Le Mans, FRANCE
e-mail: kutoyants@univ-lemans.fr

Abstract

We present several Goodness of Fit tests in the case of observations of diffusion and inhomogeneous Poisson processes. The tests studied are similar to the well known Cramèr-von Mises test of classical (i.i.d.) statistics. We propose linear transformations which make the limit distributions of statistics (under hypothesis) in all problems quite similar. Then this limit distribution is transformed in Brownian bridge or Wiener process. This allows us to construct the underlying tests asymptotically distribution free.

Keywords: Cramer-von Mises test, asymptotically distribution free tests, stochastic processes.

1. Introduction

We consider the construction of the *asymptotically distribution free* (ADF) *goodness-of-fit* (GoF) tests for three models of stochastic processes: *small noise* diffusion, ergodic diffusion and inhomogeneous Poisson processes. The basic hypothesis is supposed to be parametric.

Let us remind some well-known limits in the goodness of fit problems for the model of i.i.d. observations $X^n = (X_1, \ldots, X_n)$ with the distribution function (d.f.) $F(x)$. Suppose that we have to check the hypothesis: $\mathcal{H}_0 \quad : F(x) = F_0(x)$, where $F_0(x)$ is some known continuous distribution function.

Introduce the class of tests of asymptotic size $\alpha \in (0,1)$:

$$\mathcal{K}_\alpha = \left\{ \psi_n \quad : \quad \lim_{n\to\infty} \mathbf{E}_0 \psi_n = \alpha \right\}.$$

The Cramer-von Mises (C-vM) test $\hat\Psi_n = 1\!\!{\rm I}_{\{\Delta_n > c_\alpha\}}$ is based on the following statistics

$$\Delta_n = n \int_{-\infty}^{\infty} \left[\hat F_n(x) - F_0(x) \right]^2 \mathrm{d}F_0(x),$$

where $\hat F_n(x)$ is the empirical distribution function (EDF). It is known that the normalized difference $U_n(x) = \sqrt{n}\left(\hat F_n(x) - F_0(x)\right)$ (under hypothesis $\mathcal{H}_0$) converges in distribution to the stochastic process $B(F_0(x)), \quad x \in R$, where $B(s), 0 \le s \le 1$ is the Brownian bridge. This convergence provides the following remarcable property of the underlying statistics (we change the variables $s = F_0(x)$)

$$\Delta_n \Longrightarrow \Delta = \int_0^1 B(s)^2 \, \mathrm{d}s,$$

which make the test $\hat{\Psi}_n$ *asymptotically distribution free* (ADF), i.e., its limit distribution does not depend on the model $F_0(\cdot)$. Due to this convergence the constant c_α can be defined as solution of the equations $\mathbf{P}(\Delta > c_\alpha) = \alpha$.

If the basic hypothesis is composite parametric then the situation changes essentially. Suppose now that we have to check the hypothesis $\mathcal{H}_0 \quad : F(x) = F_0(\vartheta, x), \quad \vartheta \in \Theta$ where $\Theta = (\alpha, \beta)$, i.e.; the d.f. $F(x)$ belongs to a parametric family $\mathcal{F} = \{F_0(\vartheta, x), \quad \vartheta \in \Theta\}$. Introduce the statistics

$$\hat{\Delta}_n = n \int_{-\infty}^{\infty} \left[\hat{F}_n(x) - F_0\left(\hat{\vartheta}_n, x\right)\right]^2 \mathrm{d}F_0\left(\hat{\vartheta}_n, x\right),$$

where $\hat{\vartheta}_n$ is the maximum likelihood estimator (MLE) of the parameter ϑ. It is known that under conditions of regularity the process $\hat{U}_n(x) = \sqrt{n}\left(\hat{F}_n(x) - F_0\left(\hat{\vartheta}_n, x\right)\right)$ converges to the Gaussian process

$$U_n(x) \Longrightarrow U(t) = B(t) - \int_0^1 h(\vartheta, s)\, \mathrm{d}B(s) \int_0^t h(\vartheta, s)\, \mathrm{d}s.$$

Here $s = F_0(\vartheta, y), t = F_0(\vartheta, x)$, $h(\vartheta, s) = \mathrm{I}(\vartheta)^{-1/2} \dot{l}\left(\vartheta, F_\vartheta^{-1}(s)\right)$, $l(\vartheta, x) = \ln f(\vartheta, x)$ and dot means derivation w.r.t. ϑ.

It is easy to see that the limit distribution of the statistics $\hat{\Delta}_n$ depends strongly on the model $F_0(\cdot, \cdot)$ and on the unknown parameter ϑ. Therefore the problem of the choice of the threshold becames more difficult. It is possible to introduce a linear transformation $L[\cdot]$ of $U_n(x)$ such that the limit in distribution of $L[U_n](\cdot)$ is a Wiener process, then the GoF test constructed on the base of $L[U_n](\cdot)$ can be ADF (Khmaladze (1981)).

We consider the similar problem of the construction of ADF GoF tests for the mentioned three stochastic processes. We show that (after some transformations) the basic statistics in all three problems converge to the same limit process which can be written as follows

$$U(t) = w(t) - \int_0^1 h(s)\, \mathrm{d}w(s) \int_0^t h(s)\, \mathrm{d}s, \qquad \int_0^1 h(s)^2\, \mathrm{d}s = 1,$$

where $w(\cdot)$ is a Wiener process and $h(s) = h(\vartheta, t)$ is some function. Then we propose linear transformations of the process $U(\cdot)$ into Brownian bridge and Wiener process. This allow us to construct ADF tests for three models.

2. Results

A. *Diffusion process with small noise.*

Suppose that the observed process $X^T = (X_t, 0 \le t \le T)$ is solution of the stochastic differential equation

$$\mathrm{d}X_t = S(X_t)\, \mathrm{d}t + \varepsilon\sigma(X_t)\, \mathrm{d}W_t, \quad X_0 = x_0, \quad 0 \le t \le T,$$

where W_t is a Wiener process and the both functions $S(x)$ and $\sigma(x)$ have continuous bounded derivatives.

We have the hypothesis $\mathcal{H}_0$, $\quad S(x) = S(\vartheta, x)$, $\quad \vartheta \in \Theta = (a,b)$, i.e., this process has the stochastic differential

$$\mathrm{d}X_t = S(\vartheta, X_t)\,\mathrm{d}t + \varepsilon\sigma(X_t)\,\mathrm{d}W_t, \quad X_0 = x_0, \quad 0 \le t \le T.$$

Here $S(\vartheta, x)$ and $\sigma(x)$ are known strictly positive functions. We have to test this hypothesis in the asymptotics of *small noise* (as $\varepsilon \to 0$).

It is well-known that the solution X_t converges (uniformly in $t \in [0,T]$) to $x_t = x_t(\vartheta)$, solution of the ordinary differential equation

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = S(\vartheta, x_t), \qquad x_0, \quad 0 \le t \le T.$$

We suppose that the conditions of regularity are fulfilled. At particularly, the function $S(\vartheta, x) > 0$. The MLE $\hat\vartheta_\varepsilon$ is consistent and asymptotically normal. Moreover it admits the representation

$$\frac{\hat\vartheta_\varepsilon - \vartheta}{\varepsilon} = \frac{1}{\mathrm{I}(\vartheta)} \int_0^T \frac{\dot S(\vartheta, x_t)}{\sigma(x_t)}\mathrm{d}W_t + o(1), \quad \mathrm{I}(\vartheta) = \int_0^T \frac{\dot S(\vartheta, x_t)^2}{\sigma(x_t)^2}\mathrm{d}t.$$

Here and in the sequel dot means derivation w.r.t. ϑ.

Let us introduce the statistics

$$\delta_\varepsilon = \varepsilon^{-2} \int_0^T \left[X_t - x_t(\hat\vartheta_\varepsilon)\right]^2 \mathrm{d}t.$$

It can be shown that

$$\begin{aligned}\varepsilon^{-1}\left(X_t - x_t(\hat\vartheta_\varepsilon)\right) &= \varepsilon^{-1}\left(X_t - x_t(\vartheta)\right) - \varepsilon^{-1}\left(\hat\vartheta_\varepsilon - \vartheta\right)\dot x_t(\vartheta) + o(1) \\ &= x_t^{(1)}(\vartheta) - \mathrm{I}(\vartheta)^{-1} \int_0^T \frac{\dot S(\vartheta, x_t)}{\sigma(x_t)}\mathrm{d}W_t\, \dot x_t(\vartheta) + o(1).\end{aligned}$$

The O-U process $x_t^{(1)}$ and the derivative $\dot x_t(\vartheta)$ can be written as follows

$$x_t^{(1)} = S(\vartheta, x_t)\int_0^t \frac{\sigma(x_s)}{S(\vartheta, x_s)}\mathrm{d}W_s, \quad \dot x_t(\vartheta) = S(\vartheta, x_t)\int_0^t \frac{\dot S(\vartheta, x_s)}{S(\vartheta, x_s)}\mathrm{d}s.$$

Hence $u_\varepsilon(t) = (\varepsilon S(\vartheta, x_t))^{-1}\left(X_t - x_t(\hat\vartheta_\varepsilon)\right)$ converges to the process

$$u(t) = \int_0^t \frac{\sigma(\vartheta, x_s)}{S(\vartheta, x_s)}\,\mathrm{d}W_s - \mathrm{I}(\vartheta)^{-1}\int_0^T \frac{\dot S(\vartheta, x_s)}{\sigma(x_s)}\mathrm{d}W_s \int_0^t \frac{\dot S(\vartheta, x_s)}{S(\vartheta, x_s)}\mathrm{d}s.$$

The last step is to change the variables

$$\begin{aligned}U(s) &= \int_0^{sT} \frac{S(\vartheta, x_r)}{\sqrt{T}\sigma(x_r)}\,\mathrm{d}u(r) \\ &= \frac{W_{sT}}{\sqrt{T}} - \int_0^T \frac{\dot S(\vartheta, x_r)}{\sqrt{\mathrm{I}(\vartheta)}\,\sigma(x_r)}\mathrm{d}W_r \int_0^{sT} \frac{\dot S(\vartheta, x_r)}{\sqrt{T\mathrm{I}(\vartheta)}\,\sigma(x_r)}\mathrm{d}r \\ &= w(s) - \int_0^1 h(\vartheta, r)\,\mathrm{d}w(r) \int_0^s h(\vartheta, r)\,\mathrm{d}r, \quad \int_0^1 h(\vartheta, r)^2\,\mathrm{d}r = 1\end{aligned}$$

with obvious notation. Hence we obtained the first limit process $U(\cdot)$.

B. *Ergodic diffusion processes*

We observe a trajectory $X^T = (X_t, 0 \le t \le T)$ of diffusion process

$$\mathrm{d}X_t = S(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \quad X_0, \quad 0 \le t \le T,$$

where the function $\sigma(x)^2 > 0$ is known and we have to test the hypothesis: $\mathcal{H}_0 \qquad S(x) = S(\vartheta, x), \quad \vartheta \in \Theta$, i.e., this process has the stochastic differential

$$\mathrm{d}X_t = S(\vartheta, X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \quad X_0 = x_0, \quad 0 \le t \le T.$$

We suppose that the conditions of the existence of solution, ergodicity and regularity in the problem of estimation ϑ are fulfilled. are fulfilled. Let us denote $\hat f_T(x)$ local time estimator of the invariant density $f(\vartheta, x)$ and study the normalized difference $\zeta_T(\hat\vartheta_T, x) = \sqrt{T}\left(\hat f_T(x) - f(\hat\vartheta_T, x)\right)$:

$$\zeta_T(\hat\vartheta_T, x) = \sqrt{T}\left(\hat f_T(x) - f(\vartheta, x)\right) - \sqrt{T}\left(\hat\vartheta_T - \vartheta\right)\dot f(\vartheta, x) + o(1).$$

Here $\hat\vartheta_T$ is the maximum likelihood estimator (MLE) of ϑ. Remind that

$$\frac{\sqrt{T}\left(\hat f_T(x) - f(\vartheta, x)\right)}{2f(\vartheta, x)} \Longrightarrow \zeta(\vartheta, x) = \int_{-\infty}^{\infty} \frac{F(\vartheta, y) - 1\!\!\mathrm{I}_{\{y>x\}}}{\sigma(y)\sqrt{f(\vartheta, y)}}\,\mathrm{d}W(y),$$

where $W(\cdot)$ is double-sided Wiener process and

$$\sqrt{T}\left(\hat\vartheta_T - \vartheta\right) \Longrightarrow \frac{1}{\mathrm{I}(\vartheta)}\int_{-\infty}^{\infty} \frac{\dot S(\vartheta, y)}{\sigma(y)}\sqrt{f(\vartheta, y)}\,\mathrm{d}W(y).$$

Hence $\zeta_T\left(\hat\vartheta_T, x\right) \Rightarrow \hat\zeta(\vartheta, x)$, where

$$\hat\zeta(\vartheta, x) = \zeta(\vartheta, x) - \dot f(\vartheta, x)\int_{-\infty}^{\infty} \frac{\dot S(\vartheta, y)\sqrt{f(\vartheta, y)}}{2f(\vartheta, x)\,\mathrm{I}(\vartheta)\,\sigma(y)}\,\mathrm{d}W(y).$$

We show that (below $s = F(\vartheta, x)$

$$\int_{-\infty}^{x} \sigma(y) f(\vartheta, y)\,\mathrm{d}\hat\zeta(\vartheta, x) = w(s) - \int_0^1 h(\vartheta, t)\,\mathrm{d}w(t)\int_0^s h(\vartheta, t)\,\mathrm{d}t,$$
$$\int_0^1 h(\vartheta, t)^2\,\mathrm{d}t = 1.$$

Hence we obtain the similar limit process

C. *Inhomogeneous Poisson processes*

Suppose that we observe a periodic Poisson process $X_t, 0 \le t \le T$ of known period $\tau > 0$, $T = n\tau$. The mean and intensity functions we denote as

$\Lambda(t)$ and $\lambda(t)$ respectively. We have to test the hypothesis $\mathcal{H}_0: \quad \Lambda(t) = \Lambda(\vartheta, t), \quad \vartheta \in \Theta$, where $\Lambda(\vartheta, \cdot)$ is some known function. For simplicity of exposition we assume that . To construct the GoF test we first define the empirical mean function on one period

$$\hat{\Lambda}_n(r) = \frac{1}{n}\sum_{j=1}^{n}\left[X_{(j-1)\tau+r} - X_{(j-1)\tau}\right], \qquad r \in [0, \tau]$$

and study the process

$$\begin{aligned}\zeta_n\left(\hat{\vartheta}_n, r\right) &= \sqrt{n}\left(\hat{\Lambda}_n(r) - \Lambda(\hat{\vartheta}_n, r)\right) \\ &= \sqrt{n}\left(\hat{\Lambda}_n(r) - \Lambda(\vartheta, r)\right) - \sqrt{n}\left(\hat{\vartheta}_n - \vartheta\right)\dot{\Lambda}(\vartheta, r) + o(1),\end{aligned}$$

where $\hat{\vartheta}_n$ is the MLE. By the central limit theorem we have the convergence

$$\sqrt{n}\left(\hat{\Lambda}_n(r) - \Lambda(\vartheta, r)\right) \Longrightarrow W(\Lambda(\vartheta, r)) \sim \mathcal{N}(0, \Lambda(\vartheta, r)),$$

where $W(\cdot)$ is a Wiener process. For the MLE we have

$$\sqrt{n}\left(\hat{\vartheta}_n - \vartheta\right) = \frac{1}{\mathrm{I}(\vartheta)\sqrt{n}}\sum_{j=1}^{n}\int_0^{\tau}\frac{\dot{\lambda}(\vartheta, r)}{\lambda(\vartheta, r)}\,\mathrm{d}\left[X_{(j-1)\tau+r} - X_{(j-1)\tau}\right] + o(1).$$

Hence the process $u_n(r) = \Lambda(\hat{\vartheta}_n, \tau)^{-1/2}\zeta_n\left(\hat{\vartheta}_n, r\right)$ converges to the process

$$u(r) = \frac{W(\Lambda(\vartheta, r))}{\sqrt{\Lambda(\vartheta, \tau)}} - \frac{1}{\mathrm{I}(\vartheta)}\int_0^{\tau}\frac{\dot{\lambda}(\vartheta, v)}{\lambda(\vartheta, v)}\,\mathrm{d}W(\Lambda(\vartheta, v))\,\frac{\dot{\Lambda}(\vartheta, r)}{\sqrt{\Lambda(\vartheta, \tau)}}$$

and after the change of variables

$$s = \frac{\Lambda(\vartheta, v)}{\Lambda(\vartheta, \tau)}, \quad h(\vartheta, s) = \frac{\dot{\lambda}(\vartheta, v(s))}{\lambda(\vartheta, v(s))}\sqrt{\frac{\Lambda(\vartheta, \tau)}{\mathrm{I}(\vartheta)}}, \quad w(t) = \frac{W(\Lambda(\vartheta, r))}{\sqrt{\Lambda(\vartheta, \tau)}},$$

for $u_n(r)$ we obtain the same limit process

$$U(t) = w(t) - \int_0^1 h(\vartheta, s)\,\mathrm{d}w(s)\int_0^t h(\vartheta, s)^2\,\mathrm{d}s, \quad \int_0^1 h(\vartheta, s)^2\,\mathrm{d}s = 1. \quad (1)$$

D. *Linear transformation* We see that the limit basic statistics in all three problems has the same form (1). We propose two linear transfromations which allow to construct the ADF tests in these problems. The **first** one is to transform $U(\cdot)$ in Brownian bridge.

Introduce the process $b(t) = \int_0^t h(\vartheta, s)\,\mathrm{d}U(s)$ then

$$b(t) = \int_0^t h(\vartheta, s)\,\mathrm{d}w(s) - \int_0^1 h(\vartheta, s)\,\mathrm{d}w(s)\int_0^t h(\vartheta, s)^2\,\mathrm{d}s$$

It is easy to see that $b(t)$ is a Brownian bridge: $b(0) = b(1) = 0$,

$$\mathbf{E}b(t)\,b(s) = \int_0^{t\wedge s} h(\vartheta, v)^2 \,\mathrm{d}v - \int_0^t h(\vartheta, v)^2 \,\mathrm{d}v \int_0^s h(\vartheta, v)^2 \,\mathrm{d}v$$

Therefore if we change the time $\tau = \int_0^t h(\vartheta, v)^2 \,\mathrm{d}v$, then we have

$$\int_0^1 \left(\int_0^t h(\vartheta, s)\,\mathrm{d}U(s)\right)^2 h(\vartheta, t)^2 \,\mathrm{d}t = \int_0^1 B(\tau)^2 \,\mathrm{d}\tau.$$

We show that the tests based on the corresponding statistics, where $h(\vartheta, t)$ and $U(t)$ are replaced by their *empirical versions*, say, $h\left(\hat{\vartheta}_{\varepsilon,T,n}, t\right)$ are ADF.

The **second** solution is to transform $U(\cdot)$ in Wiener process (joint work with Kleptsyna and Liptser). Let us introduce the function $q(t,s)$ (solves the Fredholm equation) and the process M_t

$$q(t,s) - \int_0^t q(t,v)\,h(s)\,h(v)\,\mathrm{d}v = 1, \quad M_t = \int_0^t q(t,s)\,\mathrm{d}U(s).$$

We have

$$q(t,s) = 1 + N(t)^{-1} \int_0^t h(v)\,h(s)\,\mathrm{d}v, \qquad N(t) = \int_t^1 h(s)^2 \,\mathrm{d}s > 0.$$

Then it can be shown that M_t is martingale, which admits the representation

$$M_t = \int_0^t q(s,s)\,\mathrm{d}w_s, \quad \text{and} \quad w_t = \int_0^t q(s,s)^{-1}\,\mathrm{d}M_s.$$

with some Wiener process w_s. This means that

$$w_t = U(t) + \int_0^t \frac{1}{q(s,s)} \int_0^s q_t'(s,v)\,\mathrm{d}U(v)\,\mathrm{d}s = L(U)(t)$$

Therefore (in simbolic notation)

$$\hat{\delta}_T = \int_{-\infty}^{\infty} L\left[U_{\varepsilon,T,n}\right](x)^2 \,\mathrm{d}\nu_{\varepsilon,T,n}\left(\hat{\vartheta}, x\right) \Longrightarrow \int_0^1 w_t^2 \,\mathrm{d}t$$

with corresponding $\nu_{T,\varepsilon,n}\left(\hat{\vartheta}, x\right)$ (as $\varepsilon \to 0, T \to \infty, n \to \infty$) and the test

$$\hat{\psi}_T = 1\!\!\mathrm{I}_{\left\{\hat{\delta}_T > c_\alpha\right\}} \in \mathcal{K}_\alpha$$

is ADF. We discuss the relation between our approach and that of Khmaladze (1981).