On Goodness-of-Fit Testing for Continuous Time Stochastic Processes

(Poisson and Diffusion)

Yury A. Kutoyants

Université du Maine, Le Mans

23.03.2007

Poisson Processes

We observe a periodic Poisson process $X^T = \{X_t, 0 \le t \le T\}$ of intensity function $\lambda(\cdot)$ and consider the following hypotheses testing problem:

$$\mathscr{H}_0 \qquad : \qquad \lambda(t) \equiv \lambda_*(t), \quad t \ge 0$$

where $\lambda_*(t)$ is known periodic function of period τ, against

$$\mathscr{H}_1 \qquad : \qquad \lambda(t) \ne \lambda_*(t), \quad t \ge 0,$$

but $\lambda(t)$ is always τ-periodic. Let us suppose that $T = n\tau$ and denote $X_j(t) = X_{\tau(j-1)+t} - X_{\tau(j-1)}$, $j = 1, \ldots, n$. Put

$$\hat{\Lambda}_n(t) = \frac{1}{n} \sum_{j=1}^{n} X_j(t) \to \Lambda(t) = \int_0^t \lambda(s)\,\mathrm{d}s.$$

The GoF tests of C-vM and K-S type can be based on the statistics

$$W_n^2 = \Lambda_* (\tau)^{-2}\, n \int_0^\tau \left[\hat{\Lambda}_n (t) - \Lambda_* (t)\right]^2 \mathrm{d}\Lambda_* (t),$$

$$D_n = \Lambda_* (\tau)^{-1/2} \sup_{0\leq t\leq \tau} \sqrt{n} \left|\hat{\Lambda}_n (t) - \Lambda_* (t)\right|.$$

It can be shown that

$$W_n^2 \Longrightarrow \int_0^1 W (s)^2 \,\mathrm{d}s, \qquad D_n \Longrightarrow \sup_{0\leq s\leq 1} |W (s)|$$

where $W(\cdot)$ is a Wiener process. Hence these statistics are asymptotically distribution-free and the tests

$$\psi_n \left(X^T\right) = 1_{\{W_n^2 > c_\alpha\}} \in \mathcal{K}_\alpha, \qquad \phi_n \left(X^T\right) = 1_{\{D_n > d_\alpha\}} \in \mathcal{K}_\alpha$$

where $\mathcal{K}_\alpha$ is the classe of tests of asymptotic size α.

These tests are uniformly consistent against any alternative of the type

$$\mathscr{H}_\rho = \left\{\Lambda\left(\cdot\right): \ \left\|\Lambda\left(\cdot\right) - \Lambda_*\left(\cdot\right)\right\| \geq \rho\right\}$$

but are not consistent against

$$\hat{\mathscr{H}}_\rho = \left\{\Lambda\left(\cdot\right): \ \left\|\lambda\left(\cdot\right) - \lambda_*\left(\cdot\right)\right\| \geq \rho\right\}$$

because they *can not see* the intensities like

$$\lambda_k\left(t\right) = \lambda_*\left(t\right) + c_\rho \cos\left(kt\right), \quad k = 1, 2; \ldots$$

For example, for the power function $\beta\left(\bar{\psi}, \lambda\right) = \mathbf{E}_\lambda \bar{\psi}$ we have

$$\inf_{\lambda(\cdot) \in \hat{\mathscr{H}}_\rho} \beta\left(\bar{\psi}, \lambda\right) \leq \inf_k \beta\left(\bar{\psi}, \lambda_k\right) \rightarrow \alpha$$

Let us consider the problem of testing alternatives like $\hat{\mathscr{H}}_\rho$ even with $\rho = \rho_T \to 0$, but we suppose that the functions $\lambda(\cdot)$ are sufficiently smooth. The intensity $\lambda_*(\cdot)$ we transform to constant using the time change

$$t = \int_0^s \lambda_*(v)\,\mathrm{d}v, \qquad 0 \le t \le T^* = \int_0^T \lambda_*(v)\,\mathrm{d}v$$

and put $\tau = 1$. Hence

$$\mathscr{H}_0 \quad : \qquad\qquad \lambda(t) = 1.$$

The alternative is

$$\mathscr{H}_1 \quad : \qquad \lambda(\cdot) \in \Lambda_T = \left\{\lambda(\cdot)\ : \|\lambda(\cdot) - 1\| \ge \rho_T,\ \left\|\lambda^{(\sigma)}(\cdot)\right\| \le R\right\}.$$

Let us put $\lambda(t) = 1 + \vartheta(t)$ and introduce the trigonometric orthonormal base $\{\varphi_i(\cdot), i \in \mathbb{Z}\}$ in the space $\mathcal{L}_2(0,1)$ as

$$\varphi_0(1) = 1, \quad \varphi_i(t) = \sqrt{2}\cos(2\pi i t), \quad \varphi_{-i}(t) = \sqrt{2}\sin(2\pi i t),$$

where $i > 0$. Then

$$\vartheta(t) = \sum_{i\in\mathbb{Z}} \vartheta_i\ \varphi_i(t), \qquad \vartheta_i = \int_0^1 \vartheta(t)\ \varphi_i(t)\ \mathrm{d}t$$

and Λ_T corresponds to Θ_T (with condition $\inf_t \vartheta(t) \geq -1$)

$$\Theta_T = \left\{\vartheta\ :\ \sum_{i\in\mathbb{Z}} \vartheta_i^2 \geq \rho_T^2, \quad \sum_{i\in\mathbb{Z}} (2\pi\,|i|)^{2\sigma}\ \vartheta_i^2 \leq R^2\right\}.$$

Our goal is to minimize the second type error $\gamma\left(\bar\psi,\vartheta\right)=1-\mathbf{E}_\vartheta\,\bar\psi$ uniformly on alternative:

$$\sup_{\vartheta\in\Theta_T}\gamma\left(\hat\psi,\vartheta\right)=\inf_{\bar\psi\in\mathcal{K}_\alpha}\ \sup_{\vartheta\in\Theta_T}\gamma\left(\bar\psi,\vartheta\right)+o\left(1\right).$$

The test $\hat\psi\in\mathcal{K}_\alpha$ we call asymptotically minimax.

Let us introduce the statistic

$$t_w = \sum_{i=-m}^{m} w_i \ \left(X_i^2 - 1 \right),$$

where

$$w_i = z^2 \left(1 - \left| \frac{i}{m} \right|^{2\sigma} \right), \qquad z = \left(2 \sum_{i=-m}^{m} \left[1 - \left| \frac{i}{m} \right|^{2\sigma} \right]^2 \right)^{-1/4},$$

and

$$X_i = \frac{1}{\sqrt{T}} \int_0^T \phi_i(t) \left[\mathrm{d}X_t - \mathrm{d}t \right].$$

Here $\phi_i(\cdot)$ is periodic prolongation of $\varphi_i(\cdot)$ on R_+ and

$$m = \left(\frac{R^2 c_1(\sigma)}{c_2(\sigma)} \right)^{\frac{1}{2\sigma}} \rho_T^{-\frac{1}{\sigma}} \longrightarrow \infty$$

Theorem 1. *(Ingster, Kutoyants) Let* $\sigma > 1/4$*, then the test*

$$\hat{\psi} = 1_{\{t_w > z_\alpha\}} \in \mathcal{K}_\alpha$$

is asymptotically minimax. Its power function admits the representations

$$\inf_{\vartheta \in \Theta_T} \beta\left(\hat{\psi}, \vartheta\right) = \mathbf{P}\left\{\zeta > z_\alpha - u_T\right\} + o\left(1\right)$$

where $\zeta \sim \mathcal{N}\left(0, 1\right)$ *and*

$$u_T = d\left(\sigma, R\right)\, T\, \rho_T^{\frac{4\sigma+1}{2\sigma}}\, \left(1 + o\left(1\right)\right).$$

The constants

$$c_1(\sigma) = \frac{4\sigma}{2\sigma+1}, \quad c_2(\sigma) = \frac{2^{2\sigma+2}\ \pi^{2\sigma}\ \sigma}{(2\sigma+1)(4\sigma+1)},$$
$$c_3(\sigma) = \frac{8\sigma^2}{(2\sigma+1)(4\sigma+1)}$$

and

$$u_T \sim d(\sigma, R)\ T\ r_T^{2+\frac{1}{2\sigma}}, \quad d(\sigma, R) = \frac{c_3(\sigma)^{1/2}\ c_2(\sigma)^{1/\sigma}}{c_1(\sigma)\ R^{1/2\sigma}}\ .$$

The most interesting case is $u_T \to 1$. Then the separation rate

$$\rho_T = (d(\sigma, R)\ T)^{-\frac{2\sigma}{4\sigma+1}}$$

To prove this theorem we need to prove two different type results. The first one is to establish the lower bound on the errors: for all tests $\bar\psi \in \mathcal{K}_\alpha$

$$\sup_{\vartheta\in\Theta_T} \gamma(\bar\psi,\vartheta) \geq 1-\Phi\left(z_\alpha-u_T\right)+o\left(1\right),$$

and the second is to show that for the test $\hat\psi$ we have the asymptotic equalities

$$\sup_{\vartheta\in\Theta_T} \gamma(\hat\psi,\vartheta) = 1-\Phi\left(z_\alpha-u_T\right)+o\left(1\right),$$

Diffusion Processes

Let $X^T = \{X_t, 0 \le t \le T\}$ be an observation of solution of some SDE and we would like to know if this SDE is of the following form

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

where the trend $S_*\left(\cdot\right)$ and diffusion coefficient $\sigma\left(\cdot\right)^2$ are known functions. Our goal is to construct a test (called *Goodness-of-Fit*) which can answer to this question. We study such tests in two types of asymptotics: *small noise* ($\sigma \to 0$) and *large samples* ($T \to \infty$).

Small Noise Asymptotics

Suppose that the observed process $X^\varepsilon = \{X_t, 0 \le t \le T\}$ is solution of the SDE

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

If $\varepsilon \to 0$ then the stochastic process X^ε converges to the deterministic function $\{x_t, 0 \le t \le T\}$, solution of the ordinary DE

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

Our goal is to construct the GoF tests (C-vM and K-S type) for this model.

The basic hypothesis is simple:

$$\mathscr{H}_0 \quad : \qquad x_t = x_t^*, \quad 0 \leq t \leq T, \qquad \frac{\mathrm{d}x_t^*}{\mathrm{d}t} = S_*\left(x_t^*\right)$$

and the alternative is

$$\mathscr{H}_1 \quad : \qquad \left\{S\left(\cdot\right) : \left\|x_\cdot - x_\cdot^*\right\| \geq \rho\right\}, \qquad \rho > 0$$

Here x_t^* is solution x_t under hypothesis $\mathscr{H}_0$. Introduce two statistics

$$W_\varepsilon^2 = \int_0^T \left(\frac{X_t - x_t^*}{\tau\,\varepsilon\, S_*\left(x_t^*\right)^2}\right)^2 \sigma\left(x_t^*\right)^2\,\mathrm{d}t, \quad \tau = \tau\left(T\right)$$

$$D_\varepsilon = \sup_{0\leq t\leq T}\left|\frac{X_t - x_t^*}{\sqrt{\tau}\,\varepsilon\, S_*\left(x_t^*\right)}\right|, \qquad \tau\left(s\right) = \int_0^s \left(\frac{\sigma\left(x_t^*\right)}{S_*\left(x_t^*\right)}\right)^2 \mathrm{d}t$$

The *C-vM and K-S type tests*

$$\psi_\varepsilon\left(X^\varepsilon\right)=1_{\{W_\varepsilon^2>c_\alpha\}}\in\mathcal{K}_\alpha,\qquad \phi_\varepsilon\left(X^\varepsilon\right)=1_{\{D_\varepsilon>d_\alpha\}}\in\mathcal{K}_\alpha$$

with the constants c_α, d_α defined by the equations

$$\mathbf{P}\left\{\int_0^1 W\left(s\right)^2\ \mathrm{d}s>c_\alpha\right\}=\alpha,\qquad \mathbf{P}\left\{\sup_{0\leq s\leq 1}\left|W\left(s\right)\right|>d_\alpha\right\}=\alpha,$$

where $W\left(\cdot\right)$ is standard Wiener process. The both tests are *distribution free* and uniformly consistent against any alternative $\mathscr{H}_\rho$.

Let us consider local (contiguous) alternatives of the following form

$$\mathrm{d}X_t = S_*\left(X_t\right)\,\mathrm{d}t + \varepsilon\,\frac{h\left(X_t\right)\sigma\left(X_t\right)^2}{\sqrt{\tau}\,S_*\left(X_t\right)}\,\mathrm{d}t + \varepsilon\,\mathrm{d}W_t,\quad 0\leq t\leq T.$$

and denote $h_*\left(v\right) = h\left(x_{s\left(vu_T\right)}\right),\quad 0\leq v\leq 1$, where $s\left(\tau\right)$ is inverse to $\tau\left(s\right)$. We have the convergence

$$W_\varepsilon^2 \Longrightarrow \int_0^1\left[W\left(s\right)+\int_0^s h_*\left(v\right)\mathrm{d}v\right]^2\,\mathrm{d}s,$$

$$D_\varepsilon \Longrightarrow \sup_{0\leq s\leq 1}\left|W\left(s\right)+\int_0^s h_*\left(v\right)\mathrm{d}v\right|$$

The power function

$$\beta\left(\psi_\varepsilon, h\right) \to \mathbf{P}\left\{\int_0^1 \left[W\left(s\right) + \int_0^s h_*\left(v\right) \mathrm{d}v\right]^2 \mathrm{d}s > c_\alpha\right\}$$

Note that these tests are not uniformly consistent against alternatives of the form

$$\mathscr{H}_\rho^* \quad : \quad \left\{S\left(\cdot\right) : \left\|S\left(\cdot\right) - S_*\left(\cdot\right)\right\| \geq \rho\right\}, \qquad \rho > 0$$

because for the functions like $S_k\left(x\right) = S_*\left(x\right) + h_k\left(x\right)$

$$h_k\left(x\right) = c_\rho \cos\left(k\left(x - x_0\right)\right), \quad k = 1, 2, \ldots$$

we have

$$\inf_{S(\cdot)\in\mathscr{H}_\rho} \beta\left(\psi_\varepsilon, h\right) \leq \inf_k \beta\left(\psi_\varepsilon, h_k\right) \to \alpha$$

Chi Square Test

The basic hypothesis as before is simple

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

and the alternative we write as

$$\mathrm{d}X_t = S_*\left(X_t\right)\mathrm{d}t + h\left(X_t\right)\sigma\left(X_t\right)\sqrt{\left|S_*\left(X_t\right)\right|}\mathrm{d}t + \varepsilon\sigma\left(X_t\right)\mathrm{d}W_t, \quad X_0 = x_0,$$

where

$$h\left(x\right) = \left(x_T^* - x_0\right)^{-1/2}\; h_*\left(\frac{x - x_0}{x_T^* - x_0}\right), \qquad \frac{\mathrm{d}x_t^*}{\mathrm{d}t} = S_*\left(x_t^*\right),\; x_0^* = x_0.$$

Therefore, the alternative is entirely defined by the function $h_*\left(\cdot\right)$.

We suppose that $h_*(\cdot) \in \mathbb{H}_r$:

$$\mathbb{H}_r = \left\{ h(\cdot) : \quad \|h(\cdot)\| \geq r_\varepsilon, \quad \left\| h^{(k)}(\cdot) \right\| \leq R \right\},$$

where $\|\cdot\|$ is $\mathcal{L}_2(0,1)$-norm and $r = r(\varepsilon) \to 0$. We call a test $\hat{\psi}_\varepsilon$ asymptotically minimax in $\mathcal{K}_\alpha$ if

$$\sup_{h_* \in \mathbb{H}_r} \gamma\left(\hat{\psi}_\varepsilon, h_*\right) = \inf_{\bar{\psi}_\varepsilon \in \mathcal{K}_\alpha} \sup_{h_* \in \mathbb{H}_r} \gamma\left(\bar{\psi}_\varepsilon, h_*\right) + o(1).$$

Here $\gamma\left(\bar{\psi}_\varepsilon, h_*\right) = 1 - \mathbf{E}_{h_*} \bar{\psi}_\varepsilon = 1 - \beta\left(\bar{\psi}_\varepsilon, h_*\right)$

Let us introduce orthonormal base in $\mathcal{L}_2\left(0,1\right)$: $\varphi_0\left(y\right)=1$,

$$\varphi_j\left(y\right)=\sqrt{2}\ \sin\left(2\pi jy\right),\quad j>0,\quad \varphi_j\left(y\right)=\sqrt{2}\ \cos\left(2\pi jy\right),\quad j<0$$

and write

$$h_*\left(y\right)=\sum_{j\in\mathbb{Z}}\vartheta_j\,\varphi_j\left(y\right),\quad \vartheta_j=\int_0^1 h_*\left(y\right)\,\varphi_j\left(y\right)\,\mathrm{d}y.$$

The set $\mathbb{H}_r$ can be rewritten in terms of $\vartheta=\{\vartheta_j\}$ as follows

$$\Theta_r=\left\{\vartheta:\quad \sum_{j\in\mathbb{Z}}\vartheta_j^2\geq r_\varepsilon^2,\qquad \sum_{j\in\mathbb{Z}}\left(2\pi j\right)^{2k}\vartheta_j^2\leq R^2\right\}.$$

The Chi-square test we construct with the help of the random variables

$$X_j^\varepsilon = \int_0^T \varphi_j\left(\frac{X_t - x_0}{x_T^* - x_0}\right) \frac{\sqrt{|S_*(X_t)|}}{\sqrt{x_T^* - x_0}\,\sigma(X_t)} \left[\mathrm{d}X_t - S_*(X_t)\,\mathrm{d}t\right],$$

Let us put

$$q_w = \sum_{|j|\leq m} w_j \left[\left(X_j^\varepsilon\right)^2 - 1\right], \qquad m = \left(\frac{R^2 c_1(\sigma)}{c_2(\sigma)}\right)^{\frac{1}{2\sigma}} r_\varepsilon^{-\frac{1}{\sigma}} \longrightarrow \infty$$

where

$$w_j = z^2\left(1 - \left|\frac{j}{m}\right|^{2\sigma}\right), \qquad z = \left(2\sum_{j=-m}^{m}\left[1 - \left|\frac{j}{m}\right|^{2\sigma}\right]^2\right)^{-1/4},$$

Theorem 2. *Let $h_*(\cdot) \in \mathbb{H}_r$ with $k \geq 1$, then the test*

$$\hat{\psi} = 1_{\{q_w > z_\alpha\}} \in \mathcal{K}_\alpha$$

and is asymptotically minimax. Its second type error admits the representations

$$\sup_{\vartheta \in \Theta_T} \gamma\left(\hat{\psi}, \vartheta\right) = 1 - \Phi\left(z_\alpha - u_\varepsilon\right) + o\left(1\right)$$

where

$$u_\varepsilon = d\left(\sigma, R\right)\, \varepsilon^{-1}\, r_\varepsilon^{\frac{4\sigma+1}{2\sigma}}\ \left(1 + o\left(1\right)\right).$$

The test statistic can be simplified as follows, let us put

$$Y_j^\varepsilon = \int_0^T \varphi_j\left(\frac{x_t^* - x_0}{x_T^* - x_0}\right)\frac{\sqrt{|S_*\left(x_t^*\right)|}}{\sqrt{x_T^* - x_0}\,\sigma\left(X_t\right)}\ \left[\mathrm{d}X_t - S_*\left(X_t\right)\mathrm{d}t\right],$$

and note that under hypothesis the random variables Y_j^ε are independent and $Y_j^\varepsilon \sim \mathcal{N}\left(0,1\right)$. The test statistic is

$$Q_w = \sum_{|j|\leq m} w_j\left[\left(Y_j^\varepsilon\right)^2 - 1\right],$$

where w_j and m are the same as before. Then the test

$$\tilde{\psi} = 1_{\{Q_w > z_\alpha\}} \in \mathcal{K}_\alpha$$

has the same asymptotic properies as $\hat{\psi}$ above.

Ergodic Diffusion

Suppose that the observed process is one dimensional ergodic diffusion

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

i.e., there exists an invariant probability distribution $F_S\left(x\right)$ such that

$$\frac{1}{T}\int_0^T g\left(X_t\right)\,\mathrm{d}t \longrightarrow \int_{-\infty}^{\infty} g\left(x\right) f_S\left(x\right)\,\mathrm{d}x = \mathbf{E}_S g\left(\xi\right).$$

where

$$f_S\left(x\right) = \frac{1}{G\left(S\right)\sigma\left(x\right)^2}\,\exp\left\{2\int_0^x \frac{S\left(y\right)}{\sigma\left(y\right)^2}\,\mathrm{d}y\right\}.$$

The basic hypothesis $\mathscr{H}_0$ is simple: $S\left(x\right) = S_*\left(x\right)$.

We propose two different type of tests. The **First** one is based on the following two statistics

$$W_T^2 = \frac{1}{T^2\, \mathbf{E}_{S_*}\left[\sigma\left(\xi\right)^2\right]} \int_0^T \left[X_t - X_0 - \int_0^t S_*\left(X_v\right)\, \mathrm{d}v\right]^2 \mathrm{d}t,$$

$$D_T = \frac{1}{\sqrt{T\, \mathbf{E}_{S_*}\left[\sigma\left(\xi\right)^2\right]}} \sup_{0\le t\le T} \left|X_t - X_0 - \int_0^t S_*\left(X_v\right)\, \mathrm{d}v\right|$$

It can be shown that under hypothesis $\mathscr{H}_0$

$$W_T^2 \Longrightarrow \int_0^1 W\left(s\right)^2 \mathrm{d}s, \qquad D_T \Longrightarrow \sup_{0\le s\le 1} \left|W\left(s\right)\right|.$$

Hence the C-vM and K-S type tests

$$\psi_T\left(X^T\right) = 1_{\left\{W_T^2 > c_\alpha\right\}} \in \mathcal{K}_\alpha, \qquad \phi_T\left(X^T\right) = 1_{\{D_T > d_\alpha\}} \in \mathcal{K}_\alpha$$

with the same constants c_α and d_α.

Remind that $\mathbf{E}_{S_*} S_*\left(\xi\right) = 0$. These tests are consistent against any alternative of the type

$$\mathscr{H}_1 = \left\{S\left(\cdot\right) : \mathbf{E}_{S_*} S\left(\xi\right) \neq 0\right\}.$$

Let us denote

$$\mathscr{H}_\rho = \left\{h\left(\cdot\right) : \mathbf{E}_{S_*} h\left(\xi\right) \geq \rho\right\}$$

where $\rho > 0$.

The contiguous alternatives we introduce by the SDE

$$\mathrm{d}X_t = S\left(X_t\right)\,\mathrm{d}t + \frac{h\left(X_t\right)}{\sqrt{T}}\,\mathrm{d}t + \sigma\left(X_t\right)\mathrm{d}W_t, \quad X_0, \quad 0 \le t \le T.$$

We put $\rho_h = \mathbf{E}_{S_*} h\left(\xi\right)\left(\mathbf{E}_{S_*}\left[\sigma\left(\xi\right)^2\right]\right)^{-1/2}$. The limits for the both statistics under alternative $\mathscr{H}_\rho$ are

$$W_T^2 \Rightarrow \int_0^1 \left[W\left(s\right) + \rho_h\, s\right]^2 \mathrm{d}s, \quad D_T \Rightarrow \sup_{0\le s\le 1}\left|W\left(s\right) + \rho_h\, s\right|$$

We can compare these tests with the Neyman-Pearson test

$$\hat{\psi}_T\left(X^T, h\right) = 1_{\{Z_T(h) > c_\alpha(h)\}}$$

for simple alternatives. Here $Z_T\left(h\right)$ is the LR function.

Its power function ($\zeta \sim \mathcal{N}(0,1)$)

$$\beta\left(\hat{\psi}_T, h\right) \to \mathbf{P}\left\{\zeta > z_\alpha - \sqrt{\mathrm{I}(h)}\right\}, \quad \mathrm{I}(h) = \mathbf{E}_{S_*}\left(\frac{h(\xi)}{\sigma(\xi)}\right)^2$$

The least favorable alternative corresponds to

$$\inf_{h(\cdot)\in\mathscr{H}_\rho} \mathrm{I}(h) = \rho_h^2, \qquad \mathscr{H}_\rho = \left\{h(\cdot) : \mathbf{E}_{S_*} h(\xi) \geq \rho\right\}.$$

Hence

$$\inf_{h(\cdot)\in\mathscr{H}_\rho} \beta\left(\hat{\psi}_T, h\right) \to \mathbf{P}\left\{\zeta > z_\alpha - \rho_h\right\},$$

Let us introduce statistic

$$V_T\left(X^T\right)=\left(T\mathbf{E}_{S_*}\left[\sigma\left(\xi\right)^2\right]\right)^{-1/2}\left[X_T-X_0-\int_0^T S_*\left(X_t\right)\mathrm{d}t\right]$$

and the corresponding test

$$\tilde\psi_T\left(X^T\right)=1_{\{V_T(X^T)>z_\alpha\}}\in\mathcal{K}_\alpha.$$

Then its power function

$$\beta\left(\tilde\psi_T,h\right)=\mathbf{E}_{S_h}\tilde\psi_T\left(X^T\right)\rightarrow\beta\left(\tilde\psi,h\right)=\mathbf{P}\left\{\zeta>z_\alpha-\rho_h)\right\}$$

and the least favorable means

$$\inf_{h(\cdot)\in\mathscr{H}_\rho}\beta\left(\tilde\psi,h\right)=\mathbf{P}\left\{\zeta>z_\alpha-\rho_h\right\}$$

Hence this test is AUMP.

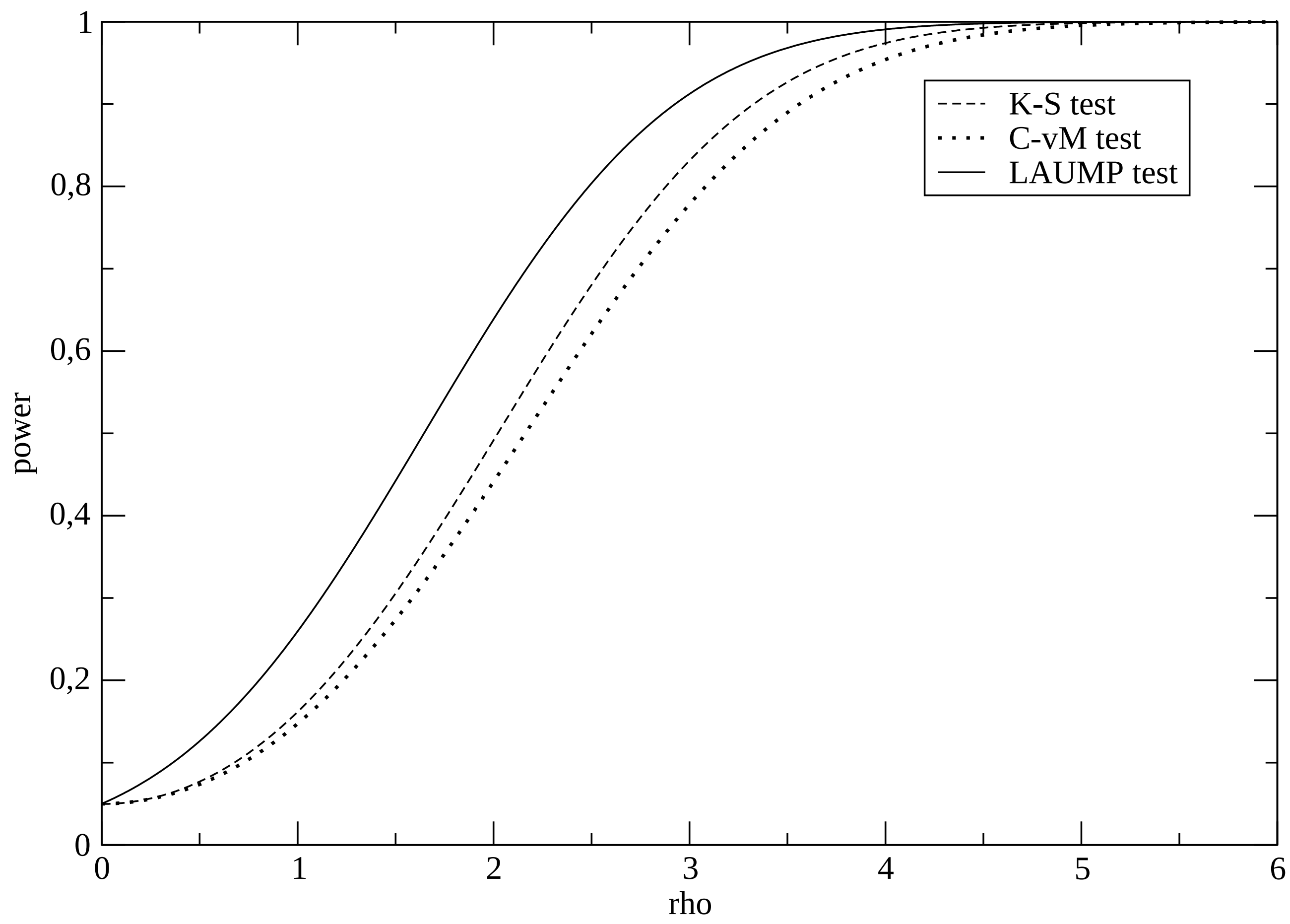
We can compare these three tests by their limit powers

$$(\mathrm{C-vM})\quad \beta\left(\psi,\rho\right)=\mathbf{P}\left\{\int_0^1\left[W\left(s\right)+\rho\, s\right]^2\mathrm{d}s>c_\alpha\right\},$$

$$(\mathrm{K-S})\quad \beta\left(\phi,\rho\right)=\mathbf{P}\left\{\sup_{0\leq s\leq 1}\left|W\left(s\right)+\rho\, s\right|>d_\alpha\right\},$$

$$(\mathrm{AUMP})\quad \beta\left(\tilde{\psi},\rho\right)=\mathbf{P}\left\{\zeta>z_\alpha-\rho\right\}$$

The results of simulation we find on the next slide.



1
0,8
0,6
0,4
0,2
0
power
0
1
2
3
4
5
6
rho
K-S test
C-vM test
LAUMP test

The **Second** type of tests is a direct analogue of the classical Cramér-von Mises and Kolmogorov-Smirnov tests based on empirical distribution and density functions:

$$\hat{F}_T\left(x\right) = \frac{1}{T}\int_0^T 1_{\{X_t<x\}}\,\mathrm{d}t, \qquad \hat{f}_T\left(x\right) = \frac{2}{T}\int_0^T 1_{\{X_t<x\}}\,\mathrm{d}X_t$$

Remind that both of them are unbiased:

$$\mathbf{E}_S\hat{F}_T\left(x\right) = F_S\left(x\right), \qquad \mathbf{E}_S\hat{f}_T\left(x\right) = f_S\left(x\right),$$

admit the representations

$$\eta_T\left(x\right) = -\frac{2}{\sqrt{T}}\int_0^T \frac{F_S\left(X_t\wedge x\right) - F_S\left(X_t\right)F_S\left(x\right)}{f_S\left(X_t\right)}\,\mathrm{d}W_t + o\left(1\right),$$

$$\zeta_T\left(x\right) = -\frac{2f_S\left(x\right)}{\sqrt{T}}\int_0^T \frac{1_{\{X_t>x\}} - F_S\left(X_t\right)}{f_S\left(X_t\right)}\,\mathrm{d}W_t + o\left(1\right)$$

The Cramér-von Mises and Kolmogorov-Smirnov type test statistics are

$$W_T^2 = T \int_{-\infty}^{\infty} \left[\hat{F}_T(x) - F_{S_*}(x) \right]^2 \, \mathrm{d}F_{S_*}(x),$$

$$D_T = \sup_x \sqrt{T} \left| \hat{F}_T(x) - F_{S_*}(x) \right|$$

and

$$v_T^2 = T \int_{-\infty}^{\infty} \left[\hat{f}_T(x) - f_{S_*}(x) \right]^2 \, \mathrm{d}F_{S_*}(x),$$

$$d_T = \sup_x \sqrt{T} \left| \hat{f}_T(x) - f_{S_*}(x) \right|$$

respectively. Unfortunately, all these statistics are not distribution-free even asymptotically and the choice of the corresponding thresholds for the tests is much more complicate.

The Cramér-von Mises and Kolmogorov-Smirnov type tests based on these statistics are

$$\Psi_T\left(X^T\right) = 1_{\left\{W_T^2 > C_\alpha\right\}}, \qquad \Phi_T\left(X^T\right) = 1_{\left\{D_T > D_\alpha\right\}},$$

$$\psi_T\left(X^T\right) = 1_{\left\{v_T^2 > c_\alpha\right\}}, \qquad \phi_T\left(X^T\right) = 1_{\left\{d_T > d_\alpha\right\}}$$

with corresponding constants belong to $\in \mathcal{K}_\alpha$.

The contiguous alternatives can be introduced by the following way

$$S(x) = S_*(x) + \frac{h(x)}{\sqrt{T}}.$$

Then we obtain for the Cramér-von Mises statistics the limits

$$W_T^2 \Rightarrow \int \left[2\mathbf{E}_{S_*} \left(\left[1_{\{\xi < x\}} - F_{S_*}(x) \right] \int_0^\xi h(s)\ \mathrm{d}s \right) + \eta(x) \right]^2 \mathrm{d}F_{S_*}(x),$$

$$v_T^2 \Rightarrow \int \left[2\mathbf{E}_{S_*} \int_\xi^x h(s)\ \mathrm{d}s + \zeta(x) \right]^2 \mathrm{d}F_{S_*}(x).$$

where $\eta(\cdot)$ and $\zeta(\cdot)$ are limit gaussian processes:

$$\sqrt{T}\left(\hat{F}_T(x) - F_{S_*}(x)\right) \Rightarrow \eta(x), \quad \sqrt{T}\left(\hat{f}_T(x) - f_{S_*}(x)\right) \Rightarrow \zeta(x)$$

Partially Observed Processes

Let us consider the partially observed system

$$\mathrm{d}Y_t = -a\,Y_t\,\mathrm{d}t + b\,\mathrm{d}V_t, \quad Y_0,$$
$$\mathrm{d}X_t = c\,Y_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \quad X_0, \quad 0 \le t \le T$$

where $a > 0, b > 0, c > 0$. We observe $X^T = \{X_t, 0 \le t \le T\}$ and we have to construct C-vM and K-S type tests to check if this model corresponds well to X^T. Remind that $m_t = \mathbf{E}_0\left(Y_t | X_s, 0 \le s \le t\right)$ satisfies the Kalman-Bucy equations:

$$\mathrm{d}m_t = c\,m_t\,\mathrm{d}t + \frac{c\,\gamma_t}{\sigma^2}\left[\mathrm{d}X_t - c\,m_t\,\mathrm{d}t\right], \quad m_0$$
$$\frac{\mathrm{d}\gamma_t}{\mathrm{d}t} = -2a\,\gamma_t - \frac{c^2\,\gamma_t^2}{\sigma^2} + b^2, \quad \gamma_0, \quad 0 \le t \le T$$

Let us introduce the statistics

$$W_T^2 = \frac{1}{T^2\sigma^2} \int_0^T \left[X_t - X_0 - c \int_0^t m_s \,\mathrm{d}s \right]^2 \mathrm{d}t,$$

$$D_T = \frac{1}{\sigma\sqrt{T}} \sup_{0 \le t \le T} \left| X_t - X_0 - c \int_0^t m_s \,\mathrm{d}s \right|.$$

Then under hypothesis $\mathscr{H}_0$ we have

$$W_T^2 = \int_0^1 W\left(s\right)^2 \,\mathrm{d}s, \qquad D_T = \sup_{0 \le s \le 1} \left| W\left(s\right) \right|.$$

and the tests

$$\psi_T\left(X^T\right) = 1_{\left\{W_T^2 > c_\alpha\right\}}, \qquad \phi_T\left(X^T\right) = 1_{\left\{D_T > d_\alpha\right\}}$$

with corresponding constants belong to $\in \mathcal{K}_\alpha$.

These tests are consistent against any alternative of the type

$$\mathrm{d}Y_t = -a\, Y_t\, \mathrm{d}t + b\, \mathrm{d}V_t, \quad Y_0,$$
$$\mathrm{d}X_t = c\, Y_t\, \mathrm{d}t + h\left(X_t\right) \mathrm{d}t + \sigma\, \mathrm{d}W_t, \quad X_0, \quad 0 \le t \le T.$$

The contigous alternatives can be introduced by the equations

$$\mathrm{d}Y_t = -a\, Y_t\, \mathrm{d}t + b\, \mathrm{d}V_t, \quad Y_0,$$
$$\mathrm{d}X_t = c\, Y_t\, \mathrm{d}t + \frac{h\left(X_t\right)}{\sqrt{T}} \mathrm{d}t + \sigma\, \mathrm{d}W_t, \quad X_0, \quad 0 \le t \le T.$$

Then $\left(r = \left(a^2 + b^2 c^2 \sigma^{-2}\right)^{1/2}\right)$

$$W_T^2 = \int_0^1 \left[W\left(s\right) + \frac{1}{\sigma T} \int_0^{sT} \left[1 - \frac{c^2}{\sigma^2} e^{-r(sT - v)} \right] h\left(X_v\right)\, \mathrm{d}v \right]^2 \mathrm{d}s$$

Limit=open question.

References

[1] Dachian, S. and Kutoyants, Yu.A. (2006) *Hypotheses testing: Poisson versus self-exciting*, Scand. J. Statist., 33, 391-408.

[2] Dachian, S. and Kutoyants, Yu.A. (2007) *On the goodness-of-fit tests for some continuous time Processes*, Statistical Models and Methods for Biomedical and Technical Systems, F.Vonta et al. (Eds), Birkhäuser, Boston, 395-413.

[3] Ingster, Yu.I. and Kutoyants, Yu.A. (2007) *Nonparametric hypothesis testing for an intensity of Poisson process*, (in prep.).

[4] Kutoyants, Yu.A. (2007) *Goodness-of-fit tests for perturbed dynamical systems*, (in preparation).

[5] Negri, I. and Nishiyama, Y. (2007) *Goodness of fit test for ergodic diffusion processes*, (submitted).