



**HAL**  
open science

## Hypotheses Testing: Poisson Versus Stress-release

Serguei Dachian, Yury Kutoyants

► **To cite this version:**

Serguei Dachian, Yury Kutoyants. Hypotheses Testing: Poisson Versus Stress-release. Journal of Statistical Planning and Inference, 2009, 139 (5), pp.1668-1684. hal-00110604

**HAL Id: hal-00110604**

**<https://hal.science/hal-00110604v1>**

Submitted on 31 Oct 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hypotheses Testing: Poisson Versus Self-correcting

Sergueï DACHIAN

Laboratoire de Mathématiques

Université Blaise Pascal, 63177 Aubière Cedex, France

Yury A. KUTOYANTS

Laboratoire de Statistique et Processus

Université du Maine, 72085 Le Mans Cedex 9, France

October 31, 2006

## Abstract

We consider the problem of hypotheses testing with the basic simple hypothesis: observed sequence of points corresponds to stationary Poisson process with known intensity against a composite one-sided parametric alternative that this is a self-correcting point process. The underlying family of measures is locally asymptotically quadratic and we describe the behavior of score function, likelihood ratio and Wald tests in the asymptotics of large samples. The results of numerical simulations are presented.

AMS 1991 Classification: 62M05.

Key words: *Poisson process, self-correcting process, hypotheses testing, Wald test, likelihood ratio test, score function test.*

## 1 Introduction

The model of self-correcting point process was proposed in 1972 by Isham and Wescott [10] to describe a stationary sequence of events  $\{t_1, t_2, \dots\}$  which does not have the property of Poisson process of independence of increments on the disjoint intervals. To introduce this processes we denote by  $X =$

$\{X_t, t \geq 0\}$  the counting process, i.e.,  $X_t$  is equal to the number of events on the time interval  $[0, t]$ . Recall that for a stationary Poisson process with a constant intensity  $S > 0$  the increments of  $X$  on disjoint intervals are independent and distributed according to Poisson law

$$\mathbf{P} \{X_t - X_s = k\} = \frac{S^k (t-s)^k}{k!} e^{-S(t-s)}, \quad 0 \leq s < t, \quad k = 0, 1, \dots$$

Particularly,

$$\mathbf{P} \{X_{t+dt} - X_t > 0\} = S dt (1 + o(1)).$$

For self-correcting point process we have

$$\mathbf{P} \{X_{t+dt} - X_t > 0 \mid \mathcal{F}_t\} = S(t, X_t) dt (1 + o(1))$$

where  $\mathcal{F}_t$  is the  $\sigma$ -field generated by  $\{X_s, 0 \leq s \leq t\}$  and the intensity function

$$S(t, X_t) = a \psi(at - X_t), \quad t \geq 0.$$

Here  $a > 0$  and the function  $\psi(\cdot)$  satisfies the following conditions:

1.  $0 \leq \psi(x) < \infty$  for any  $x \in \mathbb{R}$ ,
2. there exists a positive constant  $c$  such that  $\psi(x) \geq c$  for any  $x > 0$ ,
3.  $\underline{\lim}_{x \rightarrow \infty} \psi(x) > 1$ , and  $\overline{\lim}_{x \rightarrow -\infty} \psi(x) < 1$ .

Self-correcting processes are called as well stress-release processes (see [4], p. 239). This class of processes is widely used as a good mathematical model for non-poissonian sequences of events. This model was found especially attractive in the description of earthquakes (see Ogata and Vere-Jones [23], Lu *at al.* [14]).

**Example 1.** Let

$$S(t, X_t) = \exp \{ \alpha + \beta(t - \varrho X_t) \}$$

where  $\beta > 0$ ,  $\varrho > 0$ . It is easy to see that the conditions 1–3 are fulfilled and the point process with such intensity function is self-correcting.

This model was studied by many authors (see the references in [4]). Particularly it was shown that under mild conditions there exists an invariant measure  $\mu$  and the law of large numbers (LLN)

$$\frac{1}{T} \int_0^T h(St - X_t) dt \longrightarrow \int h(y) \mu(dy) \quad (1)$$

is valid (see Vere-Jones and Ogata [17], Hayashi [9]), Zheng [24]. Here  $h(\cdot)$  is a continuous, integrable (w.r.t.  $\mu$ ) function and  $S > 0$  is the rate of the point process. For the model of Example 1 we have the LLN if  $\rho > 0$  and  $\beta > 0$ .

As the self-correcting model is an alternative for the stationary Poisson process, it is natural and important to test these two hypotheses by the observations  $\{t_1, t_2, \dots\}$  on the time interval  $[0, T]$ , i.e., to test

$$S(t, X_t) = S \quad \text{versus} \quad S(t, X_t) = a\psi(at - X_t).$$

Remind that the likelihood ratio in this problem has the following form

$$L(X^T) = \exp \left\{ \int_0^T \ln \frac{a\psi(at - X_{t-})}{S} [dX_t - S dt] - \int_0^T \left[ \frac{a\psi(at - X_t)}{S} - 1 - \ln \frac{a\psi(at - X_t)}{S} \right] S dt \right\},$$

where  $X_{t-}$  is the limit from the left of  $X_t$  at the point  $t$  [13]. Therefore, if the function  $a\psi(\cdot)/S$  is separated from 1 then the second integral in this representation tends to infinity and there are many consistent tests. Hence it is more interesting to compare tests in the situations when the alternatives are *contiguous*, i.e. the corresponding sequence of measures are contiguous. This corresponds well to Pitman's approach in hypotheses testing [19]. We can have such situations if  $\psi(\cdot) = S + o(1)$  with special rates  $o(1)$ . In this work we consider one of such models defined by the intensity function  $S(t, X_t) = S\psi(\vartheta(St - X_t))$  where  $\vartheta$  is a *small parameter* and  $\psi(0) = 1$ . We suppose that the function  $\psi(\cdot)$  is smooth and we can write

$$\begin{aligned} \int_0^T [\psi(\vartheta(St - X_t)) - 1 - \ln \psi(\vartheta(St - X_t))] S dt &= \\ &= \frac{\vartheta^2 \dot{\psi}(0)^2 S}{2} \int_0^T (St - X_t)^2 dt (1 + o(1)). \end{aligned}$$

It is easy to see that the rate  $\vartheta = \vartheta_T \rightarrow 0$  under hypothesis  $S(t, X_t) = S$  is  $\vartheta_T \sim T^{-1}$  because

$$\frac{1}{ST^2} \int_0^T (St - X_t)^2 dt = \int_0^1 W_T(s)^2 ds \implies \int_0^1 W(s)^2 ds$$

where  $W_T(s) = (ST)^{-1/2} (STs - X_{Ts}) \Rightarrow W(s)$ , and  $\{W(s), 0 \leq s \leq 1\}$  is Wiener process. Note that we put  $a = S$ , otherwise

$$\begin{aligned} & \frac{\dot{\psi}(0)^2 \vartheta_T^2}{2} \int_0^T (at - X_t)^2 dt = \\ &= \frac{\dot{\psi}(0)^2 \vartheta_T^2}{2} \int_0^T \left( (a - S)t + \sqrt{ST} \frac{St - X_t}{\sqrt{ST}} \right)^2 dt \\ &= \frac{\dot{\psi}(0)^2 \vartheta_T^2}{2} T \int_0^1 \left( (a - S)vT + \sqrt{ST} W_T(v) \right)^2 dv \\ &= \frac{\dot{\psi}(0)^2}{6} \vartheta_T^2 (a - S)^2 T^3 (1 + o(1)) \end{aligned}$$

Therefore, if  $a \neq S$ , then we have to take  $\vartheta_T = uT^{-3/2}$  and to test the simple hypothesis  $\mathcal{H}_0 : u = 0$  against  $\mathcal{H}_1 : u > 0$ . In this case the family of measures is LAN and the usual construction provides us *asymptotically uniformly most powerful test* (see, e.g., Roussas [20]). Note that according to (1) for any fixed alternative  $\vartheta > 0$  we have the convergence

$$\frac{1}{T} \int_0^T (St - X_t)^2 dt \longrightarrow \int y^2 \mu(dy)$$

which, of course, requires another normalization.

Therefore we consider the problem of hypotheses testing when under hypothesis  $\mathcal{H}_0$  the intensity function is a known constant  $S > 0$  (Poisson process) and the alternative  $\mathcal{H}_1$  is one-sided composite: self-correcting process with intensity function  $S(t, X_t) = S\psi(\vartheta_T(St - X_t))$ , where for convenience of notation we put  $\vartheta_T = u/S\dot{\psi}(0)T$  (we suppose that  $\dot{\psi}(0) > 0$ ). In this case the corresponding likelihood ratio  $Z_T(u)$  converges to the limit process

$$Z(u) = \exp \left\{ -u \int_0^1 W(s) dW(s) - \frac{u^2}{2} \int_0^1 W(s)^2 ds \right\},$$

i.e., the family of measures is *locally asymptotically quadratic* [12]. We study three tests: *score function test*, *likelihood ratio test*, *Wald test* and compare their power functions with the power function of the *Neyman-Pearson test*. Note that we calculate all limits under hypothesis (Poisson process) and we obtain the limit distributions of the underlying statistics under alternative (self-correcting process) with the help of Le Cam's Third Lemma. Therefore we do not use directly the conditions 1–3 given above.

The similar limit likelihood ratio process arises in the problem of hypotheses testing  $u = 0$  against  $u > 0$  for the time series

$$X_j = \left(1 - \frac{u}{n}\right) X_{j-1} + \varepsilon_j, \quad j = 1, \dots, n \rightarrow \infty,$$

where  $\varepsilon_j$  are i.i.d. random variables,  $\mathbf{E}\varepsilon_j = 0$ ,  $\mathbf{E}\varepsilon_j^2 = \sigma^2$ . The asymptotic properties of tests are described under hypothesis and alternatives by Chan and Wei [2] and Phillips [18]. Particularly, the limits of the power functions are given with the help of Ornstein-Uhlenbeck process

$$dY_s = -u Y_s ds + dW_s, \quad Y_0 = 0, \quad 0 \leq s \leq 1.$$

Then Swensen [22] compared these limit powers.

For the model of Example 1 the power functions (for local alternatives) was studied by Ogata and Vere-Jones [23] and by Luschgy [15], [16]. The limit likelihood ratio and tests are similar to that of the mentioned above time series problem. Remind as well that Feigin [6] noted that the same limit likelihood ratio arises in the problem of testing the simple hypothesis  $u = 0$  against one-sided alternative  $u > 0$  by observations

$$dX_t = -\frac{u}{T} X_t dt + dW_t, \quad X_0 = 0, \quad 0 \leq t \leq T \rightarrow \infty.$$

In our case we obtain similar limit expressions for the likelihood ratio and power functions and compare the errors of tests. The analytical considerations give us an asymptotic (for large values of  $u$ ) ordering of the tests. The numerical simulations of the tests show that for the small values of  $\varepsilon$  and for the moderate values of  $u$  the power functions of the likelihood ratio and Wald tests are indistinguishable (from the point of view of numerical simulations) of the Neyman-Pearson envelope. This interesting property was noticed (for  $\varepsilon = 0.05$ ) by Elliott *at al.* [5] on the base of  $2 \cdot 10^3$  simulations. In our work we obtain similar result having  $10^7$  simulations and we observe for the larger values of  $\varepsilon$  that the asymptotic ordering of the tests holds already for the moderate values of  $u$ .

A similar problem of hypotheses testing in the situation, when the alternative process is self-exciting [8] was considered in [3].

## 2 Score Function Test

We observe a trajectory  $X^T = \{X_t, 0 \leq t \leq T\}$  of a point process of intensity function  $S(\cdot, X_t)$  and consider the problem of testing the simple hypothesis

against close one sided composite alternative

$$\mathcal{H}_0 : \quad S(t, X_t) = S_*, \quad (2)$$

$$\mathcal{H}_1 : \quad S(t, X_t) = S_* \psi(\vartheta_T [S_* t - X_t]), \quad \vartheta_T > 0, \quad (3)$$

where  $\vartheta_T$  is a small parameter, the value  $S_*$  and the function  $\psi(\cdot)$  are known. The problem is regular in the following sense.

**Condition A.** *The function  $\psi(x)$ ,  $x \in \mathbb{R}$  is positive, continuously differentiable at the point  $x = 0$ ,  $\psi(0) = 1$  and  $\dot{\psi}(0) > 0$ .*

The rate of convergence  $\vartheta_T \rightarrow 0$  is chosen such that the likelihood ratio  $L(\vartheta_T, X^T)$  is asymptotically non degenerate. In the case  $\dot{\psi}(0) < 0$  we need to change just one sign in the test. This leads us to the reparametrization

$$\vartheta_T = \frac{u}{S_* \dot{\psi}(0) T}, \quad u \geq 0$$

and to the corresponding hypotheses testing problem

$$\mathcal{H}_0 : \quad u = 0, \quad (4)$$

$$\mathcal{H}_1 : \quad u > 0. \quad (5)$$

Therefore, we observe a Poisson process of intensity  $S_*$  under hypothesis  $\mathcal{H}_0$  and the point process under alternative  $\mathcal{H}_1$  has intensity function

$$S(t, X_t) = S_* + \frac{u}{T} (S_* t - X_t) + o(T^{-1/2}).$$

Let us fix  $\varepsilon \in (0, 1)$  and denote by  $\mathcal{K}_\varepsilon$  the class of test functions  $\phi_T(X^T)$  of asymptotic size  $\varepsilon$ , i.e., for  $\phi_T \in \mathcal{K}_\varepsilon$  we have

$$\lim_{T \rightarrow \infty} \mathbf{E}_0 \phi_T(X^T) = \varepsilon.$$

As usual,  $\phi_T(X^T)$  is the probability to accept the hypothesis  $\mathcal{H}_1$  having observations  $X^T$ . The corresponding power function is

$$\beta_T(u, \phi_T) = \mathbf{E}_u \phi_T(X^T), \quad u \geq 0.$$

Let us introduce the statistic

$$\begin{aligned} \Delta_T(X^T) &= \frac{1}{S_* T} \int_0^T (S_* t - X_{t-}) [dX_t - S_* dt] \\ &= \frac{X_T - (X_T - S_* T)^2}{2 S_* T}. \end{aligned} \quad (6)$$

This equality follows from the elementary representation (see, e.g. [11], Lemma 4.2.1) for the centered Poisson process  $\pi_t = X_t - S_* t$

$$\pi_T^2 = 2 \int_0^T \pi_{t-} d\pi_t + \pi_T + S_* T.$$

which obviously is equivalent to

$$\frac{1}{T} \int_0^T \pi_{t-} d\pi_t = \frac{\pi_T^2 - X_T}{2T}.$$

Define as well two random variables

$$\Delta(W) = \frac{1}{2} (1 - W(1)^2) = - \int_0^1 W(s) dW(s), \quad J(W) = \int_0^1 W(s)^2 ds,$$

where  $\{W(s), 0 \leq s \leq 1\}$  is standard Wiener process.

Remind that the likelihood ratio in this problem has the following form [13] ( $\gamma = S_* \dot{\psi}(0)$ )

$$\begin{aligned} L\left(\frac{u}{\gamma T}, X^T\right) &= \exp \left\{ \int_0^T \ln \psi\left(\frac{u}{\gamma T} (S_* t - X_{t-})\right) [dX_t - S_* dt] \right. \\ &\quad \left. - \int_0^T \left[ \psi\left(\frac{u}{\gamma T} (S_* t - X_t)\right) - 1 - \ln \psi\left(\frac{u}{\gamma T} (S_* t - X_t)\right) \right] S_* dt \right\}. \end{aligned} \quad (7)$$

Therefore the direct differentiation w.r.t.  $u$  at the point  $u = 0$  gives us the introduced above statistic

$$\frac{\partial}{\partial u} \ln L\left(\frac{u}{\gamma T}, X^T\right) \Big|_{u=0} = \Delta_T(X^T).$$

Below we denote

$$a_\varepsilon = \frac{1 - z_{\frac{1-\varepsilon}{2}}^2}{2} \quad \text{and} \quad h(u) = \sqrt{\frac{2u}{1 - e^{-2u}}},$$

where  $z_a$  is  $1 - a$  quantile of standard Gaussian law, i.e.,  $\mathbf{P}(\zeta > z_a) = a$ , for  $\zeta \sim \mathcal{N}(0, 1)$ .

We have the following result.

**Theorem 1** *Let the Condition  $\mathcal{A}$  be fulfilled, then the score function test*

$$\phi_T^*(X^T) = \mathbb{1}_{\{\Delta_T(X^T) > a_\varepsilon\}} \quad (8)$$

*belongs to the class  $\mathcal{X}_\varepsilon$  and for any  $u_* > 0$  its power function*

$$\beta_T(u_*, \phi_T^*) \rightarrow \beta^*(u_*) = \mathbf{P}\left\{|\zeta| \leq h(u_*) z_{\frac{1-\varepsilon}{2}}\right\}. \quad (9)$$

**Proof.** Under hypothesis  $\mathcal{H}_0$  the value  $X_T$  is a poissonian random variable with parameter  $S_*T$ . Therefore we have immediately

$$\frac{X_T}{S_*T} \longrightarrow 1, \quad \frac{X_T - ST}{\sqrt{S_*T}} \Longrightarrow W(1) \sim \mathcal{N}(0, 1)$$

and  $\Delta_T(X^T) \Longrightarrow \Delta(W)$  as  $T \rightarrow \infty$ . Hence

$$\mathbf{P}_0 \{ \Delta_T(X^T) > a_\varepsilon \} \longrightarrow \mathbf{P} \left\{ \Delta(W) > \frac{1 - z_{\frac{1-\varepsilon}{2}}^2}{2} \right\} = \mathbf{P} \{ |\zeta| < z_{\frac{1-\varepsilon}{2}} \} = \varepsilon.$$

This provides  $\phi_T^* \in \mathcal{K}_\varepsilon$ .

To study the power  $\beta_T(u_*, \phi_T^*)$  we would like to use the Third Le Cam Lemma [12], [21]. Therefore we need first to show the joint weak convergence

$$\mathcal{L}_0(\Delta_T, l_T(u)) \Longrightarrow \mathcal{L} \left( \Delta(W), u\Delta(W) - \frac{u^2}{2} J(W) \right) \quad (10)$$

where  $l_T(u) = \ln L \left( \frac{u}{\gamma^T}, X^T \right)$ .

To verify (10) we denote

$$l_T^*(u) = u \Delta_T(X^T) - \frac{u^2}{2} J_T(X^T),$$

where

$$J_T(X^T) = \frac{1}{S_* T^2} \int_0^T (S_* t - X_t)^2 dt$$

and show that

$$\mathcal{L}_0(l_T^*(u)) \Longrightarrow \mathcal{L} \left( u\Delta(W) - \frac{u^2}{2} J(W) \right). \quad (11)$$

Then (10) will follow from the convergence

$$l_T^*(u_T) - l_T(u_T) \rightarrow 0 \quad (12)$$

for any bounded sequence  $u_T$ .

**Lemma 1**

$$\mathcal{L}_0 \{ \Delta_T(X^T), J_T(X^T) \} \Longrightarrow \left( - \int_0^1 W(s) dW(s), \int_0^1 W(s)^2 ds \right). \quad (13)$$

**Proof.** Let us put  $W_T(s) = (S_*T)^{-1/2} \pi_{sT}$ ,  $s \in [0, 1]$ . Then

$$\mathbf{E}_0 W_T(s) = 0, \quad \mathbf{E}_0 [W_T(s_1) W_T(s_2)] = \min(s_1, s_2)$$

and we have

$$J_T(X^T) = \frac{1}{S_* T^2} \int_0^T \pi_t^2 dt = \int_0^1 W_T(s)^2 ds.$$

Using the standard arguments we verify (well-known fact) that for any collection  $\{s_1, \dots, s_k\}$  we have the weak convergence (as  $T \rightarrow \infty$ ) of the vectors

$$\left( W_T(s_1), \dots, W_T(s_k) \right) \Longrightarrow \left( W(s_1), \dots, W(s_k) \right).$$

Moreover the following estimate holds

$$\begin{aligned} (\mathbf{E}_0 |W_T(s_1)^2 - W_T(s_2)^2|)^2 &\leq \\ &\leq \mathbf{E}_0 |W_T(s_1) - W_T(s_2)|^2 \mathbf{E}_0 |W_T(s_1) + W_T(s_2)|^2 \leq 4 |s_2 - s_1|. \end{aligned}$$

Hence (see Gikhman and Skorohod [7], Section IX.7) we have the convergence (in distribution) of integrals

$$\int_0^1 W_T(s)^2 ds \Longrightarrow \int_0^1 W(s)^2 ds$$

and

$$\Delta_T(X^T) = \frac{1 - W_T(1)^2}{2} (1 + o(1)) \Longrightarrow \frac{1 - W(1)^2}{2} = - \int_0^1 W(s) dW(s).$$

It is easy to see that we have the same time the joint convergence too because from the given above proof it follows that for any  $\lambda_1, \lambda_2$

$$\lambda_1 W_T(1)^2 + \lambda_2 \int_0^1 W_T(s)^2 ds \Longrightarrow \lambda_1 W(1)^2 + \lambda_2 \int_0^1 W(s)^2 ds.$$

Therefore the Lemma 1 is proved.

Our goal now is to establish a slightly more strong than (12) relation

$$l_T(u_T) = u_T \Delta_T(X^T) (1 + o(1)) - \frac{u_T^2}{2} \int_0^1 W_T(s)^2 ds (1 + o(1)) \quad (14)$$

where  $o(1) \rightarrow 0$  for any sequence  $u_T \in \mathbb{U}_T$  with  $\mathbb{U}_T = \{u : 0 \leq u < \frac{\sqrt{S_*T}}{\ln T}\}$ .

We can write

$$\begin{aligned}
l_T^*(u) - l_T(u) &= \int_0^T \left[ -\frac{u W_T\left(\frac{t}{T}\right)}{\sqrt{S_* T}} - \ln \psi\left(\frac{-u W_T\left(\frac{t}{T}\right)}{\dot{\psi}(0) \sqrt{S_* T}}\right) \right] d\pi_t \\
&\quad - \int_0^T \left[ \frac{u^2 W_T\left(\frac{t}{T}\right)^2}{2S_* T} - \psi\left(\frac{-u W_T\left(\frac{t}{T}\right)}{\dot{\psi}(0) \sqrt{S_* T}}\right) + 1 + \ln \psi\left(\frac{-u W_T\left(\frac{t}{T}\right)}{\dot{\psi}(0) \sqrt{S_* T}}\right) \right] S_* dt \\
&\equiv u \delta_{1,T} - \frac{u^2}{2} \delta_{2,T}
\end{aligned}$$

with obvious notation. Remind that  $u > 0$ . Using Lengart inequality we obtain for the first term

$$\begin{aligned}
\mathbf{P}_0 \{|\delta_{1,T}| > a\} &\leq \frac{b}{a} \\
&\quad + \mathbf{P}_0 \left\{ \int_0^1 \left[ W_T(s) + \frac{\sqrt{S_* T}}{u} \ln \psi\left(\frac{-u W_T(s)}{\dot{\psi}(0) \sqrt{S_* T}}\right) \right]^2 ds > b \right\}
\end{aligned}$$

for any  $a > 0$  and  $b > 0$ . Now expanding the functions  $\psi(\cdot)$  we obtain

$$\psi\left(\frac{-u W_T(s)}{\dot{\psi}(0) \sqrt{S_* T}}\right) = 1 - \frac{u W_T(s)}{\dot{\psi}(0) \sqrt{S_* T}} \dot{\psi}\left(\frac{-\tilde{u} W_T(s)}{\dot{\psi}(0) \sqrt{S_* T}}\right)$$

where  $\tilde{u} \leq u$ . Introduce the set

$$\mathbb{C}_T = \left\{ \omega : \sup_{0 \leq s \leq 1} |W_T(s)| \leq \dot{\psi}(0) \sqrt{\ln T} \right\}$$

and note that for  $\omega \in \mathbb{C}_T$  we have the estimate

$$\sup_{u \in \mathbb{U}_T} \sup_{0 \leq s \leq 1} \frac{u |W_T(s)|}{\dot{\psi}(0) \sqrt{S_* T}} \leq \frac{1}{\sqrt{\ln T}}$$

Hence for all  $u \in \mathbb{U}_T$  on this set we can write

$$\sup_{0 \leq s \leq 1} \left| \dot{\psi}(0) - \dot{\psi}\left(\frac{-\tilde{u} W_T(s)}{\dot{\psi}(0) \sqrt{S_* T}}\right) \right| \leq \sup_{|v| \leq (\ln T)^{-1/2}} \left| \dot{\psi}(0) - \dot{\psi}(v) \right| = h_T \rightarrow 0$$

as  $T \rightarrow \infty$  because the derivative is continuous at the point  $v = 0$ .

Let us denote  $u_s = \frac{u W_T(s)}{\dot{\psi}(0) \sqrt{S_* T}}$ . Using the expansion of the logarithm

$$\ln(\psi(-u_s)) = \ln\left(1 - u_s \dot{\psi}\left(\frac{-\tilde{u}_s}{\dot{\psi}(0) \sqrt{S_* T}}\right)\right) = -\frac{u_s \dot{\psi}\left(\frac{-\tilde{u}_s}{\dot{\psi}(0) \sqrt{S_* T}}\right)}{1 - \tilde{u}_s \dot{\psi}\left(\frac{-\tilde{u}_s}{\dot{\psi}(0) \sqrt{S_* T}}\right)}.$$

we obtain the following estimate

$$\begin{aligned} \mathbf{P}_0 \left\{ \int_0^1 \left[ W_T(s) + \frac{\sqrt{S_* T}}{u} \ln \psi(-u_s) \right]^2 ds > b \right\} &\leq \mathbf{P}_0 \{ \mathbb{C}_T^c \} + \\ &+ \mathbf{P}_0 \left\{ \int_0^1 W_T(s)^2 \left( 1 - \frac{\dot{\psi}(-\tilde{u}_s)}{\dot{\psi}(0) (1 - \tilde{u}_s \dot{\psi}(-\tilde{u}_s))} \right)^2 ds > b, \mathbb{C}_T \right\}. \end{aligned}$$

Remind that  $W_T(s)$  is martingale, hence by Doob inequality we have

$$\mathbf{P}_0 \{ \mathbb{C}_T^c \} \leq \mathbf{P}_0 \left\{ |W_T(1)| > \dot{\psi}(0) \sqrt{\ln T} \right\} \leq \frac{1}{\dot{\psi}(0)^2 \ln T}.$$

For the second probability after elementary estimates we obtain

$$\begin{aligned} \mathbf{P}_0 \left\{ \int_0^1 W_T(s)^2 \left( 1 - \frac{\dot{\psi}(-\tilde{u}_s)}{\dot{\psi}(0) (1 - \tilde{u}_s \dot{\psi}(-\tilde{u}_s))} \right)^2 ds > b, \mathbb{C}_T \right\} &\leq \\ &\leq \mathbf{P}_0 \left\{ C \int_0^1 W_T(s)^2 ds \left( h_T^2 + \frac{1}{\ln T} \right) > b \right\} \leq \frac{C}{2b} \left( h_T^2 + \frac{1}{\ln T} \right) \end{aligned}$$

with some constant  $C > 0$ . Recall that by Tchebyshev inequality

$$\mathbf{P}_0 \left\{ \int_0^1 W_T(s)^2 ds > A \right\} \leq \frac{1}{2A}.$$

Therefore, if we take  $b = a^2$  then for any  $a > 0$

$$\mathbf{P}_0 \{ |\delta_{1,T}| > a \} \longrightarrow 0$$

as  $T \rightarrow \infty$ .

The similar arguments allow to prove the convergence

$$\mathbf{P}_0 \{ |\delta_{2,T}| > a \} \longrightarrow 0$$

too.

Therefore, the likelihood ratio  $Z_T(u) = L\left(\frac{u}{\gamma T}, X^T\right)$ ,  $u \geq 0$  is (under hypothesis  $\mathcal{H}_0$ ) *locally asymptotically quadratic* (LAQ) [12], because

$$Z_T(u) \implies Z(u) = \exp \left\{ -u \int_0^1 W(s) dW(s) - \frac{u^2}{2} \int_0^1 W(s)^2 ds \right\}. \quad (15)$$

Moreover, we have the convergence  $l_T^*(u_T) - l_T(u_T) \rightarrow 0$  for any bounded sequence of  $u_T \in \mathbb{U}_T$ . Note that the random function  $Z(u)$  is the likelihood ratio in the hypotheses testing problem

$$\begin{aligned}\mathcal{H}_0 &: & u &= 0, \\ \mathcal{H}_1 &: & u &> 0,\end{aligned}$$

by observations of Ornstein-Uhlenbeck process

$$dY(s) = -uY(s) ds + dW(s), \quad Y(0) = 0, \quad 0 \leq s \leq 1 \quad (16)$$

under hypothesis  $u = 0$ .

This limit for the likelihood ratio under alternative can be obtained directly as follows. Let us denote

$$Y_T(s) = \frac{X_{sT} - sS_*T}{\sqrt{S_*T}}, \quad 0 \leq s \leq 1.$$

Then using the representation

$$X_t = S \int_0^t \psi(\vartheta_T[S_*r - X_r]) dr + M_t$$

where  $M_t$  is local martingale and expansion of the function  $\psi(\cdot)$  at the vicinity of 0 we obtain the equation

$$Y_T(s) = -u \int_0^s \frac{\dot{\psi}(g_v)}{\psi(0)} Y_T(v) dv + V_T(s), \quad Y_T(0) = 0, \quad 0 \leq s \leq 1$$

where  $V_T(s)$  is local martingale and  $g_v = \frac{-\tilde{u}}{\psi(0)\sqrt{S_*T}} Y_T(v) \rightarrow 0$ . The central limit theorem for local martingales provides the convergence  $V_T(s) \Rightarrow W(s)$ . Hence the process (16) is the limit (in distribution) of  $Y_T(s)$ . Moreover from (7) we have

$$\Delta_T(X^T) = \frac{Y_T(1)}{2\sqrt{S_*T}} + \frac{1 - Y_T(1)^2}{2} \Rightarrow \frac{1 - Y(1)^2}{2}.$$

This limit of the statistic  $\Delta_T(X^T)$  follows from the Third Le Cam Lemma as well. Particularly, for any continuous bounded function  $H(\cdot)$

$$\begin{aligned}\mathbf{E}_u H(\Delta_T(X^T)) &= \mathbf{E}_0 [Z_T(u) H(\Delta_T(X^T))] \longrightarrow \\ &\longrightarrow \mathbf{E}_0 [Z(u) H(\Delta(W))] = \mathbf{E}_u H(\Delta(Y)),\end{aligned}$$

where

$$\Delta(Y) = - \int_0^1 Y(s) dY(s) = \frac{1 - Y(1)^2}{2}.$$

Hence under alternative ( $\vartheta_T = u_*/\gamma T$ ) we have the convergence

$$\beta_T(u_*, \phi_T^*) \longrightarrow \mathbf{P}_{u_*} \left\{ |Y(1)| \leq z_{\frac{1-\varepsilon}{2}} \right\} = \mathbf{P} \left\{ |W(1)| \leq z_{\frac{1-\varepsilon}{2}} \sqrt{\frac{2u_*}{1 - e^{-2u_*}}} \right\}$$

because

$$Y(1) = \int_0^1 e^{-u(1-s)} dW(s) \sim \mathcal{N} \left( 0, \frac{1 - e^{-2u_*}}{2u_*} \right)$$

This proves (9).

Theorem 1 is *asymptotic in nature*, and it is interesting to see the powers of the score function test for the moderate values of  $T$  and especially to compare them with the limit power functions. This can be done using numerical simulations.

We consider the model of Example 1 with  $S_* = 1$  and  $\psi(t) = e^t$ . This yields the intensity function

$$S(u, t, X_t) = \exp\left(\frac{u}{T} [t - X_t]\right), \quad u \geq 0, \quad 0 \leq t \leq T.$$

In Figure 1 we represent the power function of the score function test  $\phi_T^*$  of asymptotic size 0.05 given by

$$\beta_T(u, \phi_T^*) = \mathbf{P}_u \left\{ \Delta_T(X^T) > a_{0.05} \right\}, \quad 0 \leq u \leq 20,$$

for  $T = 100, 300$  and  $1000$ , as well as the limiting power function  $\beta^*(\cdot)$  given by the formula (9).

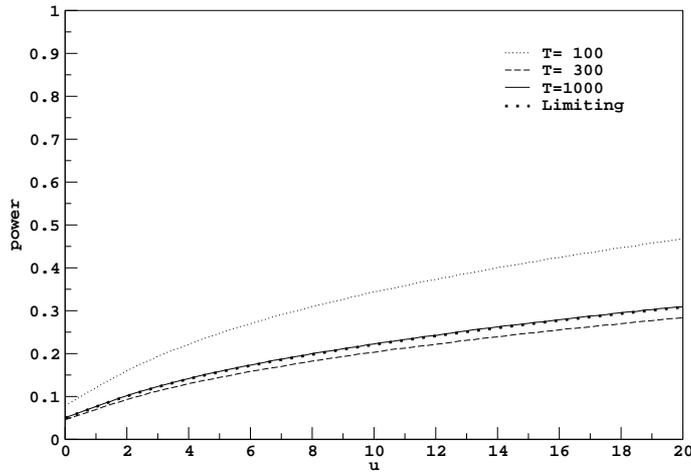


Fig. 1: Power of the score function test

The function  $\beta_T(\cdot, \phi_T^*)$  is estimated in the following way. We simulate (for each value of  $u$ )  $M = 10^6$  trajectories  $X_j^T, j = 1, \dots, M$  of self-correcting process of intensity  $S(u, t, X_t)$  and calculate  $\Delta_j = \Delta_T(X_j^T)$ . Then we calculate the empirical frequency of accepting the alternative hypothesis

$$\frac{1}{M} \sum_{j=1}^M \mathbb{1}_{\{\Delta_j > a_{0.05}\}} \approx \beta_T(u, \phi_T^*).$$

Note that for  $T = 1000$  the limiting power function is practically attained. Note also that for  $T = 100$  the size of the test is 0.079 which explains the position of the corresponding curve.

Remind that score-function test is locally optimal [1].

### 3 The Likelihood Ratio Test and the Wald Test

Let us study two other well-known tests: the *likelihood ratio test*  $\bar{\phi}_T$  based on the maximum of the likelihood ratio function and the *Wald test*  $\hat{\phi}_T$  based on the MLE  $\hat{\vartheta}_T$ .

Remind that the log-likelihood ratio formula is

$$\begin{aligned} \ln L(\vartheta, X^T) &= \int_0^T \ln \psi(\vartheta(S_*t - X_{t-})) [dX_t - S_* dt] \\ &\quad - \int_0^T [\psi(\vartheta(S_*t - X_{t-})) - 1 - \ln \psi(\vartheta(S_*t - X_{t-}))] S_* dt \end{aligned}$$

and the likelihood ratio test is based on the statistic

$$\delta_T(X^T) = \sup_{\vartheta \in \Theta} L(\vartheta, X^T),$$

where  $\Theta$  is the set of values of  $\vartheta$  under alternative. The test is given by the decision function

$$\bar{\phi}_T(X^T) = \mathbb{1}_{\{\delta_T(X^T) > \tilde{b}_\varepsilon\}}$$

where the threshold  $\tilde{b}_\varepsilon$  is chosen from the condition  $\bar{\phi}_T \in \mathcal{H}_\varepsilon$ .

Note that  $\delta_T(X^T) = L(\hat{\vartheta}_T, X^T)$  as well, where  $\hat{\vartheta}_T$  is the maximum likelihood estimator of the parameter  $\vartheta$ .

The reparametrization  $\vartheta = \vartheta_T = u/\gamma T$  reduces the problem (2)-(3) to (4)-(5) and we have to precise the region of *local alternatives*. In the traditional approach of *locally asymptotically uniformly most powerful tests* [20]

(regular case) to check the optimality of a test  $\phi_T$  we compare the power function  $\beta_T(u, \phi_T)$  with the power function of the Neyman-Pearson test on the compacts  $0 \leq u \leq K$  for any  $K > 0$ . For these values of  $u$  the alternatives are always *contiguous*. To consider the similar class of alternatives in our case is not reasonable because the constant  $\tilde{b}_\varepsilon$  became dependent of  $K$ . Indeed if we take the test function

$$\bar{\phi}_T(X^T) = \mathbb{1} \left\{ \sup_{0 < u \leq K} Z_T(u) > \tilde{b}_\varepsilon \right\}, \quad Z_T(u) = L \left( \frac{u}{\gamma T}, X^T \right),$$

then the condition  $\bar{\phi}_T \in \mathcal{H}_\varepsilon$  implies  $\tilde{b}_\varepsilon = \tilde{b}_\varepsilon(K)$ . Therefore we suppose that  $K = K_T = \frac{\sqrt{S_* T}}{\ln T} \rightarrow \infty$ .

Finally, we have the following hypotheses testing problem

$$\mathcal{H}_0 : \quad u = 0, \quad (17)$$

$$\mathcal{H}_1 : \quad u = u_* \in \mathcal{U}_T \quad (18)$$

Therefore, to study

$$\bar{\phi}_T(X^T) = \mathbb{1} \left\{ \sup_{u \in \mathcal{U}_T} Z_T(u) > \tilde{b}_\varepsilon \right\}$$

we need to describe the asymptotics of its errors under hypothesis  $\mathcal{H}_0$  and alternatives  $\mathcal{H}_1$  with  $\vartheta = \frac{u_*}{\gamma T}$ ,  $u_* \in \mathcal{U}_T$ .

Below

$$\Lambda(W) = \frac{\Delta(W)}{\sqrt{2J(W)}}.$$

**Theorem 2** *Let us suppose that condition  $\mathcal{A}$  is fulfilled and the value  $b_\varepsilon$  is solution of the equation*

$$\mathbf{P}(\Lambda(W) > b_\varepsilon) = \varepsilon. \quad (19)$$

*Then the test  $\bar{\phi}_T$  with  $\tilde{b}_\varepsilon = e^{b_\varepsilon^2}$  belongs to  $\mathcal{H}_\varepsilon$  and its power function converges to the following limit*

$$\beta(u_*, \bar{\phi}_T) \longrightarrow \hat{\beta}(u_*) = \mathbf{P}\{\Lambda(Y_{u_*}) > b_\varepsilon\},$$

where

$$\Lambda(Y_{u_*}) = \frac{\Delta(Y_{u_*})}{\sqrt{2J(Y_{u_*})}} = \frac{1 - Y_{u_*}(1)^2}{\sqrt{8J(Y_{u_*})}}.$$

and  $Y_{u_*} = \{Y_{u_*}(s), 0 \leq s \leq 1\}$  is Ornstein-Uhlenbeck process (16) with  $u = u_*$ .

**Proof.** The log-likelihood process  $l_T(u) = \ln Z_T(u)$  admits (under hypothesis  $\mathcal{H}_0$ ) the representation (14)

$$l_T(u) = u \Delta_T(X^T) (1 + \delta_{1,T}) - \frac{u^2}{2} J_T(X^T) (1 + \delta_{2,T}) \quad (20)$$

where  $\delta_{i,T} \rightarrow 0$  uniformly on  $u \in \mathcal{U}_T$ . Hence

$$\Lambda_T(X^T)^2 \equiv \sup_{u \in \mathcal{U}_T} l_T(u) \implies \frac{\Delta(W)^2}{2J(W)}$$

and we have

$$\mathbf{E}_0 \bar{\phi}_T(X^T) = \mathbf{P}_0 \left\{ \sup_{u \in \mathcal{U}_T} l_T(u) > b_\varepsilon^2 \right\} \longrightarrow \mathbf{P}(\Lambda(W) > b_\varepsilon) = \varepsilon.$$

Let us fix an alternative  $u = u_*$ . We have the convergence

$$\mathcal{L}_0 \{ \Lambda_T(X^T), l_T(u_*) \} \implies \mathcal{L} \left\{ \Lambda(W), u_* \Delta(W) - \frac{u_*^2}{2} J(W) \right\}. \quad (21)$$

The convergence (21) allows us to apply Third Le Cam's Lemma as follows: for any bounded continuous function  $H(\cdot)$

$$\begin{aligned} \mathbf{E}_{u_*} H(\Lambda_T(X^T)) &= \mathbf{E}_0 [Z_T(u_*) H(\Lambda_T(X^T))] \longrightarrow \\ &\longrightarrow \mathbf{E}_0 [Z(u_*) H(\Lambda(W))] = \mathbf{E}_{u_*} H(\Lambda(Y_{u_*})). \end{aligned}$$

Hence

$$\beta(u_*, \bar{\phi}_T) = \mathbf{P}_{u_*} \left\{ \sup_{u \in \mathcal{U}_T} l_T(u) > b_\varepsilon^2 \right\} \longrightarrow \mathbf{P}_{u_*} \{ \Lambda(Y_{u_*}) > b_\varepsilon \}.$$

This completes the proof of the theorem 2.

Let us note, that the threshold  $b_\varepsilon$  is given implicitly as the solution of the equation (19). In the following table we give some values of  $b_\varepsilon$  obtained using numerical simulations.

$\varepsilon$	0.01	0.02	0.03	0.04	0.05	0.1
$b_\varepsilon$	1.814	1.636	1.524	1.440	1.373	1.144

These thresholds are obtained by simulating  $M = 10^7$  trajectories on  $[0, 1]$  of a standard Wiener process, calculating for each of them the quantity  $\Lambda(W)$  and taking  $(1 - \varepsilon)M$ -th greatest between them.

The next test usually studied in such hypotheses testing problems is the Wald test

$$\hat{\phi}_T (X^T) = \mathbb{1}_{\{\gamma T \hat{\vartheta}_T \geq c_\varepsilon\}}$$

where  $\hat{\vartheta}_T$  is the maximum likelihood estimator of  $\vartheta$ .

Below

$$\Gamma(W) = \frac{\Delta(W)}{J(W)}.$$

**Theorem 3** *Let us suppose that condition  $\mathcal{A}$  is fulfilled and the value  $c_\varepsilon$  is solution of the equation*

$$\mathbf{P} (\Gamma(W) > c_\varepsilon) = \varepsilon. \quad (22)$$

*Then the test  $\hat{\phi}_T$  belongs to  $\mathcal{K}_\varepsilon$  and its power function for any alternative  $u_*$  converges to the following limit*

$$\beta (u_*, \hat{\phi}_T) \longrightarrow \hat{\beta} (u_*) = \mathbf{P} \{ \Gamma(Y_{u_*}) > c_\varepsilon \},$$

where

$$\Gamma(Y_{u_*}) = \frac{\Delta(Y_{u_*})}{J(Y_{u_*})} = -u_* + \frac{\int_0^1 Y_{u_*}(s) dW(s)}{J(Y_{u_*})}.$$

and  $Y_{u_*}$  is the same as in Theorem 2.

**Proof.** The proof follows immediately from the representation (20), because

$$\begin{aligned} \mathbf{P}_0^{(T)} \{ \gamma T \hat{\vartheta}_T \geq c_\varepsilon \} &= \mathbf{P}_0^{(T)} \left\{ \sup_{0 \leq u \leq c_\varepsilon} Z_T(u) < \sup_{u > c_\varepsilon, u \in \mathbb{U}_T} Z_T(u) \right\} \longrightarrow \\ &\longrightarrow \mathbf{P}_0 \left\{ \sup_{0 \leq u \leq c_\varepsilon} Z(u) < \sup_{u > c_\varepsilon} Z(u) \right\} = \mathbf{P} \{ \Gamma(W) > c_\varepsilon \} = \varepsilon \end{aligned}$$

and (under alternative  $u = u_*$ )

$$\begin{aligned} \mathbf{P}_{u_*}^{(T)} \{ \gamma T \hat{\vartheta}_T \geq c_\varepsilon \} &= \mathbf{P}_{u_*}^{(T)} \left\{ \sup_{0 \leq u \leq c_\varepsilon} Z_T(u) < \sup_{u > c_\varepsilon, u \in \mathbb{U}_T} Z_T(u) \right\} \longrightarrow \\ &\longrightarrow \mathbf{P}_{u_*} \left\{ \sup_{0 \leq u \leq c_\varepsilon} Z(u) < \sup_{u > c_\varepsilon} Z(u) \right\} = \mathbf{P} \{ \Gamma(Y_{u_*}) > c_\varepsilon \} = \hat{\beta} (u_*). \end{aligned}$$

As above, the threshold  $c_\varepsilon$  is given implicitly as the solution of the equation (22). In the following table we give some values of  $c_\varepsilon$  obtained using numerical simulations.

$\varepsilon$	0.01	0.02	0.03	0.04	0.05	0.1
$c_\varepsilon$	13.692	11.224	9.803	8.806	8.042	5.719

These thresholds are obtained by simulating  $M = 10^7$  trajectories on  $[0, 1]$  of a standard Wiener process, calculating for each of them the quantity  $\Gamma(W)$  and taking  $(1 - \varepsilon)M$ -th greatest between them.

## 4 Comparison of the Tests

Remind that all these three tests  $\phi_T^*$ ,  $\bar{\phi}_T$  and  $\hat{\phi}_T$  in regular (LAN) case are asymptotically equivalent to the Neyman-Pearson test  $\phi_{u,T}^\circ$  (with known alternative  $u$ ) and hence are asymptotically uniformly most powerful. In our singular situation all of them have different asymptotic behavior and therefore it is interesting to compare their limit power functions

$$\beta^*(u) = \mathbf{P}_u \{ \Delta(Y_u) > a_\varepsilon \}, \quad \bar{\beta}(u) = \mathbf{P}_u \left\{ \frac{\Delta(Y_u)}{\sqrt{2J(Y_u)}} > b_\varepsilon \right\},$$

$$\hat{\beta}(u) = \mathbf{P}_u \left\{ \frac{\Delta(Y_u)}{J(Y_u)} > c_\varepsilon \right\}, \quad \beta^\circ(u) = \mathbf{P}_u \left\{ u\Delta(Y_u) - \frac{u^2}{2}J(Y_u) > d_\varepsilon \right\}$$

of course, under condition that all of them belong to  $\mathcal{K}_\varepsilon$ . Our goal is to compare these quantities for the large values of  $u$ .

We have to study the distribution of the vector  $(\Delta(Y_u), J(Y_u))$ , where

$$\Delta(Y_u) = - \int_0^1 Y_u(s) dY_u(s), \quad J(Y_u) = \int_0^1 Y_u(s)^2 ds,$$

where  $Y_u$  is solution of the equation

$$dY_u(s) = -u Y_u(s) ds + dW(s), \quad Y_u(0) = 0, \quad 0 \leq s \leq 1.$$

Let us introduce the stochastic process  $y_v = \sqrt{u} Y_u\left(\frac{v}{u}\right)$ ,  $0 \leq v \leq u$  (this transformation was introduced by Luschgy [16]). Then we can write

$$dy_v = -y_v dv + dw_v, \quad y_0 = 0, \quad 0 \leq v \leq u,$$

where  $w_v = \sqrt{u} W\left(\frac{v}{u}\right)$  is a Wiener process and

$$\Delta(Y_u) = -u^{-1} \int_0^u y_v dy_v \equiv \frac{\Delta_u}{u}, \quad J(Y_u) = u^{-2} \int_0^u y_v^2 dv \equiv \frac{J_u}{u^2}.$$

in obvious notation. Further, the process  $y_v$  is ergodic with the density of the invariant law  $f(y) = e^{-y^2}/\sqrt{\pi}$ . Hence  $J_u \rightarrow \infty$  and

$$\frac{1}{u} \int_0^u y_v^2 dv \longrightarrow \frac{1}{2}.$$

Note that the distribution of the process  $y_v$  does not depend on  $u$ .

The constant  $d_\varepsilon = d_\varepsilon(u)$  because it is defined by the equation

$$\mathbf{P}_0 \left\{ u\Delta(W) - \frac{u^2}{2}J(W) > d_\varepsilon \right\} = \varepsilon.$$

For the large values of  $u$  this constant can be approximated as follows. We have (under hypothesis  $\mathcal{H}_0$ ) as  $u \rightarrow \infty$

$$\begin{aligned} \mathbf{P}_0 \left\{ u\Delta(W) - \frac{u^2}{2}J(W) > d_\varepsilon(u) \right\} &= \\ &= \mathbf{P}_0 \left\{ \int_0^1 W(s)^2 ds < -\frac{2d_\varepsilon(u)}{u^2} + \frac{2\Delta(W)}{u} \right\} \longrightarrow \\ &\longrightarrow \mathbf{P}_0 \left\{ \int_0^1 W(s)^2 ds < e_\varepsilon \right\} = \varepsilon, \end{aligned}$$

where the constant  $e_\varepsilon$  is defined by the last equality. For example, if we take  $\varepsilon = 0,05$  then the numerical simulation gives us the value  $e_{0,05} = 0,056$ . Therefore  $d_\varepsilon(u) = -0,5 e_\varepsilon u^2 (1 + o(1))$ . If we suppose that  $\varepsilon$  is small and try to solve the equation

$$\int_0^{e_\varepsilon} f_J(x) dx = \varepsilon$$

where  $f_J(x)$  is the density function of the integral  $J(W)$ , then we can easily see that  $f_J(0) = 0$  and all its derivatives  $f_J^{(k)}(0) = 0, k = 1, 2, \dots$ . Hence to see an approximative solution we need to calculate the large deviation probability of the following form (below  $r = s/\sqrt{e_\varepsilon}, E = e_\varepsilon^{-1/2} \rightarrow \infty$ ).

$$\mathbf{P}_0 \left\{ e_\varepsilon^{-1} \int_0^1 W(s)^2 ds < 1 \right\} = \mathbf{P}_0 \left\{ \int_0^E W(r)^2 dr < 1 \right\}.$$

Below we put  $d_\varepsilon(u) = -0,5 e_\varepsilon u^2$ .

We have the relations

$$\begin{aligned}\beta^*(u) &= \mathbf{P} \left\{ \Delta_u > u a_\varepsilon \right\} = \mathbf{P} \left\{ \int_0^u y_v dw_v < J_u - a_\varepsilon u \right\}, \\ \bar{\beta}(u) &= \mathbf{P} \left\{ \frac{\Delta_u}{\sqrt{2J_u}} > b_\varepsilon \right\} = \mathbf{P} \left\{ \int_0^u y_v dw_v < J_u - b_\varepsilon \sqrt{2J_u} \right\}, \\ \hat{\beta}(u) &= \mathbf{P} \left\{ \frac{\Delta_u}{J_u} > \frac{c_\varepsilon}{u} \right\} = \mathbf{P} \left\{ \int_0^u y_v dw_v < J_u - \frac{c_\varepsilon}{u} J_u \right\}, \\ \beta^\circ(u) &= \mathbf{P} \left\{ \Delta_u - \frac{J_u}{2} > d_\varepsilon \right\} = \mathbf{P} \left\{ \int_0^u y_v dw_v < \frac{1}{2} J_u + \frac{e_\varepsilon}{2} u^2 \right\}.\end{aligned}$$

Therefore the large values of  $u$  ( $J_u \sim u/2$ )

$$\frac{1}{2} J_u + \frac{e_\varepsilon}{2} u^2 > J_u - \frac{c_\varepsilon}{u} J_u > J_u - b_\varepsilon \sqrt{2J_u} > J_u - a_\varepsilon u,$$

and finally

$$\beta^*(u) < \bar{\beta}(u) < \hat{\beta}(u) < \beta^\circ(u).$$

These inequalities are in accord with [22].

Note that for small values of  $\varepsilon$  the constant  $a_\varepsilon$  is close to 0,5 (e.g.  $a_{0,05} = 0,498$ ,  $a_{0,01} = 0,49992$ ) and in this asymptotics the power of score-function test is

$$\beta^*(u) = \mathbf{P} \left\{ \int_0^u y_v dw_v < (0,5 - a_\varepsilon) u (1 + o(1)) \right\}.$$

Hence one can expect that in this case the score-function test has essentially smaller power than the others.

Now let us turn to numerical simulations of the limiting power functions. We aim to obtain the limiting power functions of all the three tests, as well as the Neyman-Pearson envelope, for the moderate values of  $u$  ( $u \leq 15$ ).

Note that for the score function test  $\beta^*(u)$  can be computed directly using (9). However the limiting power functions of the likelihood ratio and of the Wald tests are written as probabilities of some events related to Ornstein-Uhlenbeck process and can be obtained using numerical simulations.

For the likelihood ratio test we have

$$\bar{\beta}(u) = \mathbf{E}_u \mathbf{1}_{\{\Lambda(Y_u) > b_\varepsilon\}} = \mathbf{E}_0 Z(u) \mathbf{1}_{\{\Lambda(W) > b_\varepsilon\}}$$

where

$$Z(u) = \exp \left\{ u \Delta(W) - \frac{u^2}{2} J(W) \right\}.$$

So we simulate  $M = 10^7$  trajectories  $W_j = \{W_j(s), 0 \leq s \leq 1\}$ ,  $j = 1, \dots, M$  of a standard Wiener process and calculate for each of them the quantities  $\Delta_j = \Delta(W_j)$ ,  $J_j = J(W_j)$ ,  $\Lambda_j = \Delta_j/J_j$  and (for each value of  $u$ )  $Z_j(u) = \exp\left\{u\Delta_j - \frac{u^2}{2}J_j\right\}$ . Then we calculate the empirical mean

$$\frac{1}{M} \sum_{j=1}^M Z_j(u) \mathbb{1}_{\{\Lambda_j > b_\varepsilon\}} \approx \bar{\beta}(u).$$

For the Wald test we have similarly

$$\frac{1}{M} \sum_{j=1}^M Z_j(u) \mathbb{1}_{\{\Gamma_j > c_\varepsilon\}} \approx \hat{\beta}(u)$$

where  $\Gamma_j = \Delta_j/\sqrt{2J_j}$ .

Finally, in order to compute the Neyman-Pearson envelope, we first approximate (for each value of  $u$ ) the quantity  $d_\varepsilon = d_\varepsilon(u)$  by the  $(1 - \varepsilon)M$ -th greatest between the quantities  $\ln Z_j(u)$ , and then calculate

$$\frac{1}{M} \sum_{j=1}^M Z_j(u) \mathbb{1}_{\{\ln Z_j(u) > d_\varepsilon(u)\}} \approx \beta^\circ(u).$$

The results of these simulations for  $\varepsilon = 0.05$  are presented in Figure 2.

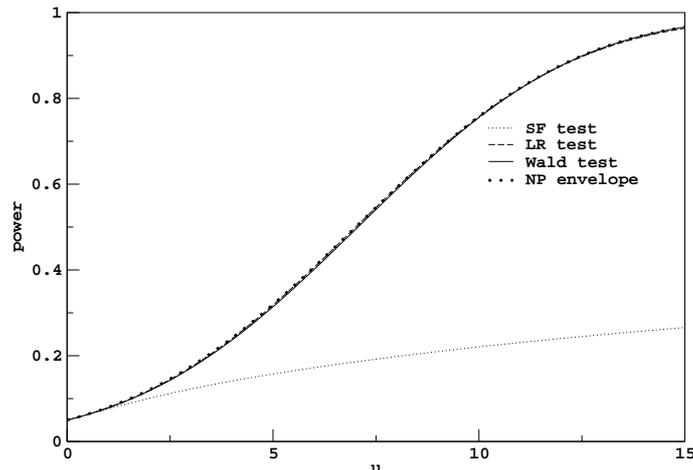


Fig. 2: Limiting powers for  $\varepsilon = 0.05$

Let us note here that in this case the power functions of the likelihood ratio test and of the Wald test are indistinguishable (from the point of view of numerical simulations) from the Neyman-Pearson envelope. This quite

surprising fact was already mentioned by Elliott *at al.* [5], who showed the similar pictures having  $2 \cdot 10^3$  simulations. As we see from Figure 2, with  $10^7$  simulations the curves are still indistinguishable. The situation is however different for bigger values of  $\varepsilon$ . The results of simulations for  $\varepsilon = 0.01, 0.05, 0.25$  and  $0.5$  are presented in Figure 3.

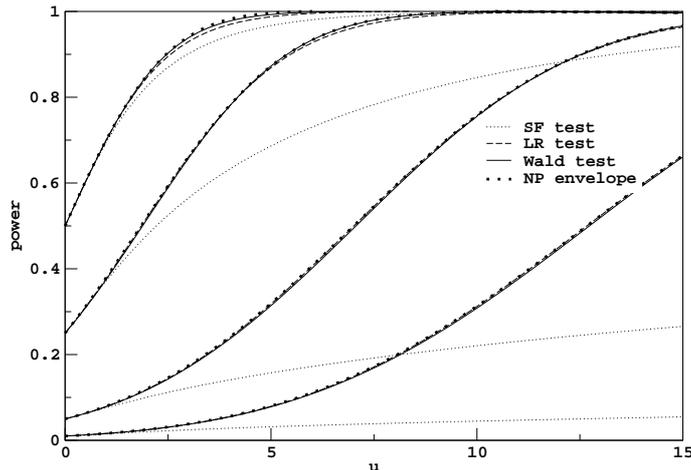


Fig. 3: Limiting powers for different values of  $\varepsilon$

One can note that for big values of  $\varepsilon$  (e.g.  $\varepsilon = 0.5$ ) the powers became more distinguishable, and that the asymptotically established ordering of the tests holds already for these moderate values of  $u$ . Note also that for the small values of  $\varepsilon$  (e.g.  $\varepsilon = 0.01$  and  $0.05$ ) the curve of score-function test is essentially lower as expected.

## 5 Discussion

**Remark 1.** Note that alternatives  $u = u_T \rightarrow \infty$  with  $\vartheta_{u_T} \rightarrow 0$  are local but not contiguous. That means that the corresponding sequences of measures  $(\mathbf{P}_{\vartheta_{u_T}}^{(T)}, \mathbf{P}_0^{(T)})$ ,  $T \rightarrow \infty$  are not contiguous. Particularly, the second integral in the likelihood ratio formula tends to infinity:

$$\int_0^T [\psi(\vartheta_{u_T}(S_*t - X_{t-})) - 1 - \ln \psi(\vartheta_{u_T}(S_*t - X_{t-}))] S_* dt \longrightarrow \infty.$$

In such situation the power function of any reasonable test tends to 1 and to compare tests we have to use, say, the large deviation principle. For example, the likelihood ratio test  $\phi_T^*$  is consistent for the *local far alternatives*  $\vartheta = \frac{v}{\sqrt{S_*T}}$ ,  $v \in [\nu, V]$  where  $0 < \nu < V < \infty$ . Indeed, under mild regularity

conditions we can write

$$\begin{aligned}
\mathbf{E}_v \phi_T^* (X^T) &= \mathbf{P}_0 \left\{ \sup_{\nu < v < V} L \left( \frac{v}{\sqrt{S_* T}}, X^T \right) > c_\varepsilon \right\} = \\
&= \mathbf{P}_0 \left\{ \sup_{\nu < v < V} \left[ \sqrt{S_* T} \int_0^1 \ln \psi (v W_T (s)) dW_T (s) - \right. \right. \\
&\quad \left. \left. - S_* T \int_0^1 [\psi (v W_T (s)) - 1 - \ln \psi (v W_T (s))] ds \right] > \ln c_\varepsilon \right\} = \\
&= \mathbf{P}_0 \left\{ \sup_{\nu < v < V} \left[ \frac{1}{\sqrt{S_* T}} \int_0^1 \ln \psi (v W_T (s)) dW_T (s) - \right. \right. \\
&\quad \left. \left. - \int_0^1 [\psi (v W_T (s)) - 1 - \ln \psi (v W_T (s))] ds \right] > \frac{\ln c_\varepsilon}{S_* T} \right\} \longrightarrow \\
&\longrightarrow \mathbf{P} \left\{ \inf_{\nu < v < V} \int_0^1 [\psi (v W (s)) - 1 - \ln \psi (v W (s))] ds > 0 \right\} = 1
\end{aligned}$$

because the function  $g(y) = y - 1 - \ln y > 0$  for  $y \neq 1$  and  $g(y) = 0$  iff  $y = 1$ .

**Remark 2.** Note, that we can construct asymptotically uniformly most powerful test if we change the statement of the problem in the following way. Let us fix some  $D > 0$  and introduce the stopping time

$$\tau_D = \inf \left\{ \tau : \int_0^\tau (S_* t - X_t)^2 S_* dt \geq D^2 \right\}.$$

Then we consider the problem of testing hypotheses

$$\begin{aligned}
\mathcal{H}_0 &: S(t, X_t) = S_*, \\
\mathcal{H}_1 &: S(t, X_t) = S_* \psi(\vartheta_D [S_* t - X_t]), \quad \vartheta_D = \frac{u}{\dot{\psi}(0) D} > 0
\end{aligned}$$

by observations  $X^{\tau_D} = \{X_t, 0 \leq t \leq \tau_D\}$  in the asymptotics  $D \rightarrow \infty$ . Now the likelihood ratio  $Z_{\tau_D}(u) = L\left(\frac{u}{\dot{\psi}(0)D}, X^D\right)$  will be LAN:

$$Z_{\tau_D}(u) \implies \exp \left\{ u \zeta - \frac{u^2}{2} \right\}, \quad \zeta \sim \mathcal{N}(0, 1)$$

and the test  $\hat{\phi}_{\tau_D} = \mathbb{1}_{\{\Delta_{\tau_D}(X^{\tau_D}) > z_\varepsilon\}}$  where

$$\Delta_{\tau_D}(X^{\tau_D}) = \frac{1}{D} \int_0^{\tau_D} (S_* t - X_{t-}) [dX_t - S_* dt]$$

is locally asymptotically uniformly most powerful.

The proof follows from the central limit theorem for stochastic integrals and the standard arguments (for LAN families).

**Remark 3.** Note that these problems of hypotheses testing are similar to the corresponding problems of hypotheses testing for diffusion processes. In particular, let the observed process  $X^T = \{X_t, 0 \leq t \leq T\}$  be diffusion

$$dX_t = \psi(-\vartheta_T X_t) dt + \sigma dW_t, \quad X_0 = 0, \quad 0 \leq t \leq T,$$

where the function  $\psi(0)=0$ , is continuously differentiable at the point 0 and  $\dot{\psi}(0) > 0$ . If we consider two hypotheses:  $\vartheta = 0$  and  $\vartheta > 0$  then the reparametrization

$$\vartheta_T = \frac{u \sigma}{\dot{\psi}(0) T}$$

provides local contiguous alternatives, i.e., the log-likelihood ratio in the problem

$$\begin{aligned} \mathcal{H}_0 : & \quad u = 0, \\ \mathcal{H}_1 : & \quad u > 0. \end{aligned}$$

has the limit:

$$\ln L \left( \frac{u \sigma}{\dot{\psi}(0) T}, X^T \right) \Longrightarrow -u \int_0^1 W(s) dW(s) - \frac{u^2}{2} \int_0^1 W(s)^2 ds.$$

The score function test based on the statistic

$$\Delta_T^*(X^T) = -\frac{1}{T} \int_0^T X_t dX_t,$$

the likelihood ratio test and the Wald test have the same asymptotic properties as those described in Theorems 1, 2 and 3 above.

For example, if  $\psi(x) = x$ , then we have the Wiener process (under hypothesis  $\mathcal{H}_0$ ) against ergodic Ornstein-Uhlenbeck process under alternative  $\mathcal{H}_1$ .

**Remark 4.** We supposed above that the derivative of the function  $\psi(x)$  at the point  $x = 0$  is not equal to 0, but sometimes it can be interesting to study the score function and the likelihood ratio test in the situations when the first  $k - 1$  derivatives with  $k \geq 2$  are null.

Let us consider a self-correcting process  $X^T = \{X_t, 0 \leq t \leq T\}$  with intensity function  $S_* \psi(\vartheta(S_* t - X_t))$  such that  $\psi(0) = 1$ ,  $\dot{\psi}(0) = 0$  and  $\ddot{\psi}(\cdot) \neq 0$

( $k = 2$ ). In this case the modifications have to be the following. Suppose that  $\dot{\psi}(0) > 0$ . To have LAQ family at the point  $\vartheta = 0$  we chose the reparametrization  $\vartheta = \vartheta_u$

$$\vartheta_u = \sqrt{\frac{2u}{\dot{\psi}(0)}} (S_*T)^{-3/4},$$

which provides the limit

$$\ln L(\vartheta_u, X^T) \Longrightarrow u \int_0^1 W(s)^2 dW(s) - \frac{u^2}{2} \int_0^1 W(s)^4 ds$$

Then in the hypotheses testing problem

$$\begin{aligned} \mathcal{H}_0 : & \quad u = 0, \\ \mathcal{H}_1 : & \quad u > 0 \end{aligned}$$

the score function test  $\hat{\psi}(X^T) = \mathbb{1}_{\{\Delta_T(X^T) > c_\varepsilon\}}$  is based on the statistic

$$\Delta_T(X^T) = \frac{1}{(S_*T)^{3/2}} \int_0^T (S_*t - X_t)^2 [dX_t - S_*dt].$$

It is easy to see that under  $\mathcal{H}_0$

$$\Delta_T(X^T) \Longrightarrow \frac{W(1)^3}{3} - \int_0^1 W(s) ds.$$

Hence to chose the threshold  $c_\varepsilon$  we have to solve the following equation

$$\frac{4}{3} \iint_{x^3 - y > 3c} \exp\left\{-2x^2 + 2xy - \frac{2}{3}y^2\right\} dx dy = \varepsilon$$

because  $(W(1), 3 \int_0^1 W(s) ds)$  is Gaussian vector.

The cases  $k > 2$  can be treated in a similar way.

## References

- [1] Capon, J. (1961) On asymptotic efficiency of locally optimum detectors, *IRE Transactions of Information Theory*, **7**, 67–71.
- [2] Chan, N.H. and Wei, C.Z. (1987) Asymptotic inference for nearly non-stationary AR(1) processes, *Ann. Statist.*, **15**, 1050–1063.

- [3] Dachian, S. and Kutoyants, Yu.A. (2006) Hypotheses Testing: Poisson versus self-exciting, *Scand. J. Statist.*, **33**, 391–408.
- [4] Daley, D.J. and Vere-Jones, D. (2003) *An introduction to the theory of point processes. vol. I. (2nd ed.)*, Springer-Verlag, New York.
- [5] Elliott, G., Rothenberg, T.J. and Stock, J.H. (1996) Efficient tests for an autoregressive unit root, *Econometrika*, **64**, 4, 813–836.
- [6] Feigin, P.D. (1979) Some comments concerning a curious singularity, *J. Appl. Probab.*, **16**, 440–444.
- [7] Gikhman, I.I. and Skorokhod, A.V. (1969) *Introduction to theory of random processes*, W.B. Saunders, Philadelphia.
- [8] Hawkes, A.G. (1972) Spectra for some mutually exciting point processes with associated variable, in *Stochastic point processes*, P.A.W. Lewis ed., Wiley, New York.
- [9] Hayashi, T. (1986) Laws of large numbers in self-correcting point processes, *Stoch. Proces. Appl.*, **23**, 319–326.
- [10] Isham, V. and Westcott, M. (1979) A self-correcting point processes, *Stoch. Proc. Appl.*, **8**, 335–347.
- [11] Kutoyants, Yu.A. (1984) *Parameter estimation for stochastic processes*, Heldermann, Berlin.
- [12] Le Cam, L. and Yang, G.L. (2000) *Asymptotics in statistics (2nd ed.)*, Springer-Verlag, New York.
- [13] Liptser, R.S. and Shiriyayev, A.N. (2001) *Statistics of random processes. II. Applications (2nd ed.)*, Springer-Verlag, New York.
- [14] Lu, C., Harte, D. and Bebbington, M. (1999) A linked stress release model for historical Japanese earthquakes: coupling among major seismic regions, *Earth Planets Space*, **51**, 907–916.
- [15] Luschgy, H. (1993) On a singularity ocuring in a self-correcting point process model, *Ann. Inst. Statist. Math.*, **45**, 3, 445–452.
- [16] Luschgy, H. (1994) Asymptotic inference for semimartingale models with singular parameter points, *J. Stoch. Plan. Inference*, **39**, 155–186.
- [17] Ogata, Y. and Vere-Jones, D. (1984) On the moments of a self-correcting point processes, *J. Appl. Prob.*, **21**, 335–342. Corr: *ibid* (1987), **24**, 291.

- [18] Phillips, P.C.B. (1987) Toward a unified asymptotic theory for autoregression, *Biometrika*, **74**, 535–547.
- [19] Pitman, E.J.G. (1948) *Lecture notes on nonparametric statistical inference*, Lectures given for the University of North Carolina, Institute of Statistics.
- [20] Roussas, G.G. (1972) *Contiguity of probability measures*, Cambridge University Press, Cambridge.
- [21] Strasser, H. (1985) *Mathematical theory of statistics*, Walter de Gruyter, Berlin.
- [22] Swensen, A.R. (1997) Contiguity in nonstationary time series, in *Festschrift for Lucien Le Cam*, Springer-Verlag, New York, 377–384.
- [23] Vere-Jones, D. and Ogata, Y. (1984) Inference for earthquake models: a self-correcting model, *Stoch. Proc. Appl.*, **17**, 337–347.
- [24] Zheng, X. (1991) Ergodic theorems for stress release proceses, *Stoch. Proc. Appl.*, **37**, 239–258.