MATHEMATICS IN ENGINEERING, SCIENCE AND AEROSPACE
MESA - www.journalmesa.com
Vol. 2, No. 3, pp. 243-250, 2011

A note on optimality of hypothesis testing

Albert Vexler[1] and Gregory Gurevich[2]

[1] Department of Biostatistics,
State University of New York at Buffalo,
Buffalo, NY 14214, USA.
E-mail: avexler@buffalo.edu

[2] Department of Industrial Engineering and Management,
SCE- Shamoon College of Engineering,
Beer Sheva 84100, Israel

⋆ *Corresponding Author*. E-mail: avexler@buffalo.edu

Abstract. Commonly, in accordance with a given risk-function of hypothesis testing, investigators try to derive an optimal property of a test. This paper demonstrates that criteria for which a given test is optimal can be declared by the structure of this test, and hence almost any reasonable test is optimal. In order to establish this conclusion, the principle idea of the fundamental lemma of Neyman and Pearson is applied to interpret the goodness of tests, as well as retrospective and sequential change point detections are considered in the context of the proposed technique.Aside from that, the present article evaluates a specific classification problem that corresponds to measurement error effects in occupational medicine.

1 Introduction

Without loss of generality, we can say that the principle idea of the proof of the fundamental lemma of Neyman and Pearson is based on the trivial inequality

$$(A-B)\,(I\{A \geq B\}-\delta) \geq 0, \tag{1.1}$$

for all A, B and $\delta \in [0,1]$ ($I\{\cdot\}$ is the indicator function). Thus, for example, if we would like to classify i.i.d. observations $\{X_i,\, i=1,\ldots,n\}$ regarding the hypotheses:

$$\mathbf{H_0} : X_1 \text{ is from a density function } f_0 \ \ \textbf{versus} \ \ \mathbf{H_1} : X_1 \text{ is from a density function } f_1,$$

2010 **Mathematics Subject Classification**: Primary: 62F03; Secondary: 62L10.
Keywords: Change Point; Hypotheses Testing; Limit of Detection; Optimality.

then the likelihood ratio test (i.e. we should reject H_0 iff $\prod_{i=1}^{n} f_1(X_i)/f_0(X_i) \geq H$, where H is a threshold) is uniformly most powerful. This classical proposition directly follows from the expected value under H_0 of (1.1) with $A = \prod_{i=1}^{n} f_1(X_i)/f_0(X_i)$, $B = H$ and δ that is considered as any decision rule based on the observed sample. However, in the case when we have a given test-statistic S_n instead of the likelihood ratio $\prod_{i=1}^{n} f_1(X_i)/f_0(X_i)$, the inequality (1.1) yields $(S_n - H)I\{S_n \geq H\} \geq (S_n - H)\delta$, for any decision rule $\delta \in [0,1]$. Therefore, the test-rule $S_n \geq H$ has also an optimal property following the application of (1.1). The problem is to obtain a reasonable interpretation of the optimality.

The present paper proposes that the structure of any given retrospective or sequential test can substantiate type-(1.1) inequalities that provide optimal properties of that test. Although this approach can be applied easily to demonstrate an optimality of tests, presentation of that optimality in the classical operating-characteristics (e.g. the power/significance level of tests) is the complicated issue.

In order to present the proposed methodology, different kinds of hypothesis testing are considered in this article. Consequently, the sections of the paper are supplied with brief introductions related to the corresponding problem statements. The paper proceeds as follows. Section 2 considers the retrospective change point problem. This section demonstrates that the Shiryayev-Roberts approach applied to the change point detection yields the averagely most powerful procedure. A sequential testing is evaluated in Section 3. Here the focus on the inequality (1.1) leads us to a non-asymptotic optimality of a well known sequential procedure, for which only an asymptotic result of optimality has been proved in the literature. Section 4 introduces a specific classification problem. Although the issue, which corresponds to limits of instrumentation of epidemiologic studies, has wide practical meaning, the stated problem has not been well addressed in the statistical literature. Section 5 mentions a remark and short conclusion of the paper.

2 Retrospective change point detection

The issues of the so-called retrospective (non-sequential) change point problem arise in experimental and mathematical sciences (epidemiology, quality control, etc.). In this section we will focus on a sequence of previously obtained independent observations in an attempt to detect whether they all share the same distribution. Note that, commonly, the change point detection problem is evaluated under assumption that if a change in the distribution did occur, then it is unique and the observations after the change all have the same distribution, which differs from the distribution of the observations before the change (e.g. Page, 1954; Sen and Srivastava, 1975; Gombay and Horvath, 1994; Harel et al., 2008; Gurevich, 2006, 2007; Gurevich and Vexler, 2010).

Let $X_1, \ldots, X_n$ be independent observations with density functions $g_1, \ldots, g_n$, respectively. We want to test the null hypothesis

$$\mathbf{H_0} : g_i = f_0 \text{ for all } i = 1, \ldots, n$$

versus the alternative

$$\mathbf{H_1} : g_1 = \ldots = g_{\nu-1} = f_0 \neq g_\nu = \ldots = g_n = f_1, \quad \nu \text{ is unknown.}$$

The maximum likelihood estimation of the change point parameter ν applied to the likelihood ratio $\prod_{i=\nu}^{n} f_1(X_i)/f_0(X_i)$ leads us to the well accepted in the change point literature CUSUM test (e.g. Page, 1954; Gurevich and Vexler, 2005): we should reject H_0 iff

$$\max_{1 \leq k \leq n} \prod_{i=k}^{n} \frac{f_1(X_i)}{f_0(X_i)} \geq H,$$

where $H > 0$ is a threshold. Alternatively, following Vexler (2006), the test based on the Shiryayev-Roberts approach has the folowing form: we reject H_0 iff

$$\frac{1}{n} R_n \equiv \frac{1}{n} \sum_{k=1}^{n} \prod_{i=k}^{n} \frac{f_1(X_i)}{f_0(X_i)} \geq H. \tag{2.1}$$

Consider an optimal meaning of the test (2.1). Let P_k and E_k $(k = 0, \ldots, n)$ respectively denote probability and expectation conditional on $\nu = k$ (the case $k = 0$ corresponds to H_0). By virtue of (1.1) with $A = R_n/n$ and $B = H$, it follows that

$$\left(\frac{1}{n} R_n - H \right) I\{R_n \geq nH\} \geq \left(\frac{1}{n} R_n - H \right) \delta. \tag{2.2}$$

Without loss of generality, we assume that $\delta = 0, 1$ is any decision rule based on the observed sample $\{X_i, i = 1, \ldots, n\}$ such that if $\delta = 1$, then we reject H_0. Since

$$\begin{aligned}
\frac{1}{n} E_0(R_n \delta) &= \frac{1}{n} \sum_{k=1}^{n} E_0 \left(\prod_{i=k}^{n} \frac{f_1(X_i)}{f_0(X_i)} \delta \right) = \frac{1}{n} \sum_{k=1}^{n} \int \cdots \int \prod_{i=k}^{n} \frac{f_1(x_i)}{f_0(x_i)} \delta \prod_{i=1}^{n} f_0(x_i) \prod_{i=1}^{n} dx_i \\
&= \frac{1}{n} \sum_{k=1}^{n} \int \cdots \int I\{\delta = 1\} \prod_{i=1}^{k-1} f_0(x_i) \prod_{i=k}^{n} f_1(x_i) \prod_{i=1}^{n} dx_i \\
&= \frac{1}{n} \sum_{k=1}^{n} P_k\{\delta = 1\},
\end{aligned}$$

derivation of H_0-expectation of the left and right side of (2.2) directly provides the following proposition.

Proposition 1. The test (2.1) is the averagely most powerful test, i.e.

$$\begin{aligned}
&\frac{1}{n} \sum_{k=1}^{n} \left(P_k \left\{ \frac{1}{n} R_n \geq H \right\} - H P_0 \left\{ \frac{1}{n} R_n \geq H \right\} \right) \\
&\geq \frac{1}{n} \sum_{k=1}^{n} \left(P_k\{\delta \text{ declares rejection of } H_0\} - H P_0\{\delta \text{ declares rejection of } H_0\} \right),
\end{aligned}$$

for any decision rule $\delta \in [0, 1]$ based on the observations $\{X_i, i = 1, \ldots, n\}$.

3 Sequential Shiryayev-Roberts change point detection

In this section, we suppose that independent observations $X_1, X_2, \ldots$ are surveyed sequentially. Let $X_1, \ldots, X_{\nu-1}$ be each distributed according to a density function f_0, whereas $X_\nu, X_{\nu+1}, \ldots$ have a density function f_1, where $1 \leq \nu \leq \infty$ is unknown. The case $\nu = \infty$ corresponds to the situation when all observations are distributed according to f_0. In this case, while considering the notations P_∞ and E_∞, we denote probability and expectation, respectively, when all observations are from the same distribution. The formulation of sequential change point detection conforms to raising an alarm as soon as possible after the change and to avoid false alarms. Efficient detection methods for this stated problem are based on CUSUM and Shiryayev-Roberts stopping rules. The CUSUM policy is: we stop sampling of Xs and report that a change in distribution of X has been detected at the first time $n \geq 1$

that $\max_{1\le k\le n}\prod_{i=k}^{n} f_1(X_i)/f_0(X_i) \ge H$, for a given threshold H. The Shiryaev-Roberts procedure is defined by the stopping time

$$N_H = \inf\{n \ge 1 : R_n \ge H\}, \tag{3.1}$$

where the Shiryayev-Roberts test-statistic R_n is

$$R_n = \sum_{k=1}^{n}\prod_{i=k}^{n}\frac{f_1(X_i)}{f_0(X_i)}. \tag{3.2}$$

The sequential CUSUM detection procedure has a non-asymptotic optimal property (Moustakides, 1986). At that, for the Shiryaev-Roberts procedure an asymptotic (as $H \to \infty$) optimality has been shown (Pollak, 1985). (Although Yakir (1997) attempted to prove a non-asymptotic optimality of the Shiryaev-Roberts procedure, Mei (2006) has pointed out the inaccuracy of the Yakir's proofs.) In order to demonstrate optimality of the Shiryaev-Roberts detection scheme, Pollak (1985) has proved an asymptotic closeness of the expected loss using a Bayes rule for the considered change problem with a known prior distribution of ν to that using the rule N_H. However, in the context of simple application of the proposed methodology, the procedure (3.1) itself declares loss functions for which that detection policy is optimal. That is, setting $A = R_{\min(N_H,n)}$ and $B = H$ in (1.1) leads to

$$\left(R_{\min(N_H,n)} - H\right)\left(I\left\{R_{\min(N_H,n)} \ge H\right\} - \delta\right) \ge 0, \tag{3.3}$$

for all $\delta \in [0,1]$. Since $\left\{R_{\min(N_H,n)} \ge H\right\} = \{N_H \le n\}$,

$$\begin{aligned}&\left(R_{\min(N_H,n)} - H\right)\left(I\left\{R_{\min(N_H,n)} \ge H\right\} - \delta\right)\\ &= \sum_{k=1}^{n}(R_k - H)(1-\delta)I\{N_H = k\} + (R_n - H)(-\delta)I\{N_H > n\} \ge 0.\end{aligned} \tag{3.4}$$

It is clear that (3.4) can be utilized to present an optimal property of the detection rule N_H. However, here, for simplicity, noting that every summand in the left side of the inequality (3.4) is non-negative, we focus only on $(R_n - H)(-\delta)I\{N_H > n\} \ge 0$. Thus, if τ is a stopping time (assuming that for all k, the event $\{\tau \le k\}$ is measurable in the σ-algebra generated by $X_1,\dots,X_k$) and δ is defined by $\delta = I\{\tau \le n\}$ then

$$E_\infty\left[(H - R_n)I\{\tau \le n, N_H > n\}\right] \ge 0. \tag{3.5}$$

By virtue of the definition (3.2), we obtain from (3.5) that

$$\begin{aligned}&H\left(P_\infty\{N_H > n\} - P_\infty\{\min(\tau,N_H) > n\}\right)\\ &-\sum_{k=1}^{n}\left(P_k\{N_H > n\} - P_k\{\min(\tau,N_H) > n\}\right) \ge 0.\end{aligned} \tag{3.6}$$

Therefore,

$$\begin{aligned}&H\sum_{n=1}^{\infty}\left(P_\infty\{N_H > n\} - P_\infty\{\min(\tau,N_H) > n\}\right)\\ &-\sum_{n=1}^{\infty}\sum_{k=1}^{n}\left(P_k\{N_H > n\} - P_k\{\min(\tau,N_H) > n\}\right) \ge 0,\end{aligned} \tag{3.7}$$

where

$$\sum_{n=1}^{\infty}\sum_{k=1}^{n}\left(P_k\{N_H>n\}-P_k\{\min(\tau,N_H)>n\}\right)=\sum_{k=1}^{\infty}\sum_{n=k}^{\infty}\left(P_k\{N_H>n\}\right. \tag{3.8}$$

$$\left.-P_k\{\min(\tau,N_H)>n\}\right)=\sum_{k=1}^{\infty}\left(E_k(N_H-k+1)^+-E_k(\min(\tau,N_H)-k+1)^+\right)$$

($a^+=aI\{a\geq 0\}$). Let us summarize (3.7) with (3.8) in the next proposition.

Proposition 2. The Shiryaev-Roberts stopping time (3.1) satisfies

$$\sum_{n=1}^{\infty}E_n(N_H-n+1)^+ + (-HE_\infty(N_H))$$
$$=\min_{\tau}\left[\sum_{n=1}^{\infty}E_n(\min(\tau,N_H)-n+1)^+ + (-HE_\infty(\min(\tau,N_H)))\right].$$

Note that, $E_\infty(\tau)$ corresponds to the average run length to false alarm of a stopping rule τ, and hence small values of $-E_\infty(\tau)$ are privileged, whereas small values of $E_n(\tau-n+1)^+$ are also preferable (because $E_n(\tau-n+1)^+$ relates to fallibility of the sequential detection in the case $\nu=n$). Obviously, if $\nu<\infty$ then $\min(\tau,N_H)$ detects that $\nu<\infty$ faster than the stopping time N_H.

4 Testing for Limit of Detection

The following problem of testing has been introduced by Vexler *et al.* (2006). Assume that independent observations $\{Z_i,\ i=1,\ldots,n\}$ satisfy

$$Z_i=x_iI\{x_i\geq d\}+\varepsilon_iI\{x_i<d\}, \tag{4.1}$$

where x_i,ε_i are some real random variables and d is a fixed (non-random) threshold value. For the sake of clarity of exposition, we assume that $\{x_i,\ i\geq 1\}$ are i.i.d. random variables with a density function f_x, and $\{x_i,\ i\geq 1\}$ are independent of i.i.d. random variables $\{\varepsilon_i,\ i\geq 1\}$ which have a density function f_ε. We focus on the problem of testing for homogeneity of the observed sample, i.e.

$$\begin{aligned}&\mathbf{H_0}: Z_1,\ldots,Z_n \text{ are each distributed according to a density } f_x(u),\ \textbf{versus}\\&\mathbf{H_1}: \text{for all } i=1,\ldots,n\text{: } Z_i \text{ is distributed according to a density}\\&\qquad f_Z(u;d)\equiv f_x(u)I\{u\geq d\}+f_\varepsilon(u)P\{x_1<d\};\ d \text{ is unknown.}\end{aligned} \tag{4.2}$$

An important application of this stated problem corresponds to the so-called limit of detection issue in environmental epidemiology (Vexler et al. 2008). In this case, the model (4.1) and testing for (4.2) can be considered in the context of exposure measurement bound by a limit of detection (e.g. Schisterman *et al.*, 2006). Frequently in epidemiological studies, data can include noise values that are effect of exposure quantification compromised when the measurement process of a biomarker is subject to a lower threshold.

In this section, we denote the probability and expectation for a given d as P_d and E_d, respectively. The case $d=-\infty$ corresponds to the hypothesis H_0. In this case, $P_{-\infty}$ and $E_{-\infty}$ denote probability and expectation, respectively, when we observe $\{Z_i=x_i,\ i=1,\ldots,n\}$. When d is known, the classical test-statistic for (4.2) is the likelihood ratio

$$\Lambda_n(d) \equiv \prod_{i=1}^{n} \frac{f_Z(Z_i;d)}{f_x(Z_i)}. \tag{4.3}$$

In (4.2), the parameter d is unknown, and therefore we construct an estimator of d based on the maximum likelihood method. To this end, we arrange the sequence $\{Z_i, i = 1,\dots,n\}$ in decreasing order: $\infty = Z_{(0:n)} > Z_{(1:n)} \geq Z_{(2:n)} \geq \dots \geq Z_{(n:n)} > Z_{(n+1:n)} = -\infty$. Since

$$1 = \sum_{k=1}^{n+1} I\{Z_{(k-1:n)} \geq d > Z_{(k:n)}\} \text{ and hence } \Lambda_n(d) = \sum_{k=1}^{n+1} \Lambda_n(d) I\{Z_{(k-1:n)} \geq d > Z_{(k:n)}\},$$

we define the maximum likelihood estimator of $\Lambda_n(d)$ in the form $\Lambda_n = \max_k \Lambda_n(Z_{(k-1:n)})$ (d is estimated by $Z_{(k-1:n)}$, where $k = \arg\max_l \prod_i f_Z(Z_i; Z_{(l-1:n)})$). Now, following Section 2, we denote the Shiryayev-Roberts test-statistic $R_n = \sum_k \Lambda_n(Z_{(k-1:n)})$. Formally, we obtain

$$\begin{aligned} R_n &\equiv \sum_{k=2}^{n} \Lambda_n(Z_{(k-1:n)}) + \prod_{i=1}^{n} \frac{f_\varepsilon(Z_i)}{f_x(Z_i)} = \sum_{k=2}^{n} \prod_{i=1}^{n} \frac{f_Z(Z_i; Z_{(k-1:n)})}{f_x(Z_i)} + \prod_{i=1}^{n} \frac{f_\varepsilon(Z_i)}{f_x(Z_i)} \\ &= \sum_{k=2}^{n} \prod_{i=1}^{n} \frac{f_Z(Z_i; Z_{k-1})}{f_x(Z_i)} + \prod_{i=1}^{n} \frac{f_\varepsilon(Z_i)}{f_x(Z_i)} = \sum_{k=2}^{n} \Lambda_n(Z_{k-1}) + \prod_{i=1}^{n} \frac{f_\varepsilon(Z_i)}{f_x(Z_i)}, \end{aligned} \tag{4.4}$$

where without ties: $Z_{(1:n)} > Z_{(2:n)} > \dots > Z_{(n:n)}$ and $\sum_2^1 = 0$. Note that, (4.4) is a very simple representation of the test statistic. Without additional assumptions dealing with bounds for the unknown d, the application of the test statistic in the form of $\max_{\hat{d}} \Lambda_n(\hat{d})$ is a very complex problem, which strongly depends on the type of density functions of the stated problem.

The proposed test for the problem (4.2) is: we fix a threshold H and reject H_0 iff

$$\frac{1}{n} R_n \geq H \tag{4.5}$$

(H can depend on n). Suppose $\delta = 0,1$ is any decision rule based on the observed sample $\{Z_i, i = 1,\dots,n\}$ such that the event $\{\delta = 1\}$ recommends rejection of H_0. Rewrite the inequality (1.1) in the form of

$$E_{-\infty}\left(\frac{1}{n}R_n - H\right) I\{R_n \geq nH\} \geq E_{-\infty}\left(\frac{1}{n}R_n - H\right)\delta. \tag{4.6}$$

In accordance with the definition (4.4), we consider

$$\begin{aligned} E_{-\infty}\left(\sum_{k=2}^{n} \Lambda_n(Z_{k-1})\delta\right) &= \sum_{k=2}^{n} E_{-\infty}\left(E_{-\infty}\left(\Lambda_n(Z_{k-1})\delta | Z_{k-1}\right)\right) \\ &= \sum_{k=2}^{n} E_{-\infty}\left(\frac{f_Z(Z_{k-1}; Z_{k-1})}{f_x(Z_{k-1})} \int \cdots \int \prod_{\substack{i=1 \\ i\neq k-1}}^{n} \frac{f_Z(z_i; Z_{k-1})}{f_x(z_i)} \delta \right. \\ &\qquad \left. \times \prod_{\substack{i=1 \\ i\neq k-1}}^{n} f_x(z_i) \prod_{\substack{i=1 \\ i\neq k-1}}^{n} dz_i \right) \\ &= \sum_{k=2}^{n} E_{-\infty}\left(\frac{f_Z(Z_{k-1}; Z_{k-1})}{f_x(Z_{k-1})} P_{d=Z_{k-1}}\{\delta = 1 | Z_{k-1}\}\right). \end{aligned} \tag{4.7}$$

Since, under H_0, Z_i, $i = 1,\dots,n$ are independent identically $P\{x_1 < u\}$-distributed random variables, applying (4.7) to the inequality (4.6) directly yields the next proposition.

Proposition 3. The test (4.5) is the averagely most powerful test with respect to distribution functions of x_1 and ε_1, i.e.

$$\frac{1}{n}\sum_{k=2}^{n}\int \frac{f_Z(u;u)}{f_x(u)} P_{d=u}\left\{\frac{1}{n}R_n \geq H | Z_{k-1}=u\right\} dP\{x_1<u\} + \frac{1}{n}P_{d=\infty}\left\{\frac{1}{n}R_n \geq H\right\}$$
$$-H\,P_{d=-\infty}\left\{\frac{1}{n}R_n \geq H\right\}$$
$$\geq \frac{1}{n}\sum_{k=2}^{n}\int \frac{f_Z(u;u)}{f_x(u)} P_{d=u}\{\delta \text{ declares rejection of } H_0 | Z_{k-1}=u\}\, dP\{x_1<u\}$$
$$+\frac{1}{n}P_{d=\infty}\{\delta \text{ declares rejection of } H_0\} - H\,P_{d=-\infty}\{\delta \text{ declares rejection of } H_0\},$$

for any decision rule $\delta \in [0,1]$ based on the observations $\{Z_i,\, i=1,\ldots,n\}$.

5 Remark and Conclusion

5.1 The case, where distributions of observations are known up to parameters.

In the context of statistical testing, when distribution functions of observations are known up to parameters, one of the approaches dealing with estimation of unknown parameters is the mixture technique. To be specific, consider the change point problem introduced in Section 2. Assume that the density function $f_1(u) \equiv f_1(u;\theta)$, where θ is the unknown parameter. In this case, we can transform the test (2.1) in accordance with the mixture methodology (e.g. Krieger *et al.,* 2003). That is, we have to choose a prior Θ and pretend that $\theta \sim \Theta$. Thus, the mixture Shiryayev-Roberts statistic has form

$$R_n^{(1)} \equiv \frac{1}{n}\sum_{k=1}^{n}\int \prod_{i=k}^{n} \frac{f_1(X_i;u)}{f_0(X_i)} d\Theta(u),$$

and hence the consequential transformation of the test (2.1) has the following property

$$\frac{1}{n}\sum_{k=1}^{n}\int P_{\nu=k}\left\{\frac{1}{n}R_n^{(1)} \geq H \middle| \{X_j\}_{j\geq\nu} \text{ are from } f_1(X_i;u)\right\} d\Theta(u)$$
$$-\,H\,P_0\left\{\frac{1}{n}R_n^{(1)} \geq H\right\}$$
$$\geq\ \frac{1}{n}\sum_{k=1}^{n}\int P_{\nu=k}\left\{\delta \text{ declares rejection of } H_0 \middle| \{X_j\}_{j\geq\nu} \text{ are from } f_1(X_i;u)\right\} d\Theta(u)$$
$$-\,H\,P_0\{\delta \text{ declares rejection of } H_0\},$$

for any decision rule $\delta \in [0,1]$ based on the observations $\{X_i,\, i=1,\ldots,n\}$. In this case the optimality formulated in Proposition 1 is integrated over values of the unknown parameter θ.

5.2 The case where one wishes to investigate the optimality of a given test in an unfixed context.

In this situation, inequalities similar to (1.1) can be obtained by focusing on the structure of a test. These inequalities report an optimal property of the test. However, translation of that property in terms of the quality of tests is the issue.

References

[1] Gombay, E. and Horvath, L. (1994). An application of the maximum likelihood test to the change-point problem. *Stochastic Processes and their Applications* **50**, 161-171.
[2] Gurevich, G. (2006). Nonparametric AMOC changepoint tests for stochastically ordered alternatives *Commun. Statist. Theor. Meth.* **35**, 887-903.
[3] Gurevich, G. (2007). Retrospective parametric tests for homogeneity of data. *Commun. Statist. Theor. Meth.* **36**, 2841-2862.
[4] Gurevich, G. and Vexler, A. (2005). Change point problems in the model of logistic regression. *Journal of Statistical Planning and Inference* **131**, 313331.
[5] Gurevich, G. and Vexler, A. (2010). Retrospective change point detection: from parametric to distribution free policies. *Commun. Statist. Simul. Comput.* **39**, 899-920.
[6] Harel, O., Schisterman, E. F., Vexler, A. and Ruopp M. D. (2008). Monitoring quality control: can we get better data? *Epidemiology* **19**, 621-627.
[7] Krieger, A. M., Pollak, M. and Yakir, B. (2003). Surveillance of a simple linear regression. *J. Amer. Statist. Assoc.* **98**, 456-469.
[8] Mei, Y. (2006). Comments on "A note on optimal detection of a change in distribution" by Benjamin Yakir, The Annals of Statistics (1997) **25** 2117-2126. *Ann. Statist.* **34**, 1570-1576.
[9] Moustakides, G. V. (1986). Optimal stopping times for detecting changes in distributions. *Ann. Statist.* **14**, 1379-1387.
[10] Page, E. S. (1954). Continuous inspectation schemes. *Biometrika* **41**, 100-114.
[11] Pollak, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.* **13**, 206-227.
[12] Pollak, M. (1987). Average run lengths of an optimal method of detecting a change in distribution. *Ann. Statist.* **15**, 749-779.
[13] Schisterman, E. F., Vexler, A., Whitcomb, B. W. and Liu, A. (2006). The limitations due to exposure detection limits for regression models. *American Journal of Epidemiology* **163**, 374-383.
[14] Sen, A. and Srivastava, M. S. (1975). On tests for detecting change in mean. *Ann. Statist.* **1**, 98-108.
[15] Vexler, A. (2006). Guaranteed testing for epidemic changes of a linear regression model. *Journal of Statistical Planning and Inference* **136**, 3101-3120.
[16] Vexler, A., Liu, A., Eliseeva, E. and Schisterman, E.F. (2008). Maximum likelihood ratio tests for comparing the discriminatory ability of biomarkers subject to limit of detection. *Biometrics.* **64**, 895-903.
[17] Vexler, A., Wu, C., Liu, A., Whitcomb, B. W. and Schisterman, E. F. (2006). An extension of a change point problem. *Statistics.* **43:3**, 213-225.
[18] Yakir, B. (1997). A note on optimal detection of a change in distribution. *Ann. Statist.* **25**, 2117-2126.