A NONPARAMETRIC REPEATED SIGNIFICANCE TEST WITH ADAPTIVE TARGET SAMPLE SIZE

VLADIMIR POZDNYAKOV AND JOSEPH GLAZ
DEPARTMENT OF STATISTICS, UNIVERSITY OF CONNECTICUT
STORRS, CT 06269-4120

Abstract. In this article a general result is derived that, along with a functional central limit theorem for a sequence of statistics, can be employed in developing a nonparametric repeated significance test with adaptive target sample size. This method is used in deriving a repeated significance test with adaptive target sample size for the shift model. The repeated significance test is based on a functional central limit theorem for a sequence of partial sums of truncated observations. Based on numerical results presented in this article one can conclude that this nonparametric sequential test performs quite well.
Key Words: Functional central limit theorem; Heavy tail distributions; Invariance principle; Repeated significance test; Sequential test; Trimming; Truncation.
AMS 2000 Subject Classification: Primary 62L12; Secondary 60F17.

1. Introduction

Repeated significance tests have been introduced in Armitage (1958) and since then have been investigated extensively. Recent developments in the theory and applications of repeated significance tests have been reported among others in: Chang, Hsiung, and Hwang (1999), Glaz and Pozdnyakov (2004), Gu and Lai (1998), Ho (1998), Jennison and Turnbull (2000), Kim (1994), Lee, Kim and Tsiatis (1996), Lerche (1986), Sen (1985 and 2002), Siegmund (1985), Takahashi (1990) and Whitehead (1997). Nonparametric repeated significance tests have been discussed in Glaz and Pozdnyakov (2004) and Sen (1981, 1985 and 1991).

In this article, based on a functional central limit theorem, we introduce a sequential test that is a repeated significance test with a random target sample size. This random target sample size adapts itself to the distribution of the observed data and it depends on the gross rate of the sample variance of the sequence of test statistics employed by the repeated significance test. The advantage in using a repeated significance test with an adaptive target sample size is that the resulting sequential test is fully nonparametric. Its implementation does not require the knowledge of the asymptotic tail behavior of the distribution of the observed data, a condition needed for developing the repeated significance test in Glaz and Pozdnyakov (2004).

The article is organized as follows. In Section 2, the role of a functional central limit theorem in developing a repeated significance test based on a given sequence of statistics is discussed. In Section 3, we prove a general result, that along with a functional central limit theorem for a sequence of statistics, can be employed in developing a repeated significance test with adaptive target sample size. In Section

4, the method developed in the previous sections is applied to a sequence of partial sums of truncated observations in the context the shift model. In Section 5, we present numerical results to evaluate the performance of the repeated significance test with adaptive target sample size developed in Section 4. In Section 6, concluding remarks related to two open problems, closely connected to results derived in this article, are given.

2. A Repeated Significance Test Based on a Functional Central Limit Theorem

Let $\{T_n, ; n \geq 1\}$ be a sequence of statistics with $E(T_n) = 0$, when a certain null hypothesis, about a parameter of the distribution of observed data, is in force. Let B_n be a sequence of numbers such that $B_n \sim Var(T_n)$, i.e. $B_n/Var(T_n) \to 1$ as $n \uparrow \infty$ and

$$B_n \uparrow \infty, \quad \frac{B_{n-1}}{B_n} \to 1, \text{ as } n \uparrow \infty. \tag{1}$$

We define $T_n(t)$ to be a random element of $C[0,1]$ obtained by linear interpolation between points

$$\Big(0, 0\Big), \Big(\frac{B_1}{B_n}, \frac{T_1}{\sqrt{B_n}}\Big), ..., \Big(1, \frac{T_n}{\sqrt{B_n}}\Big).$$

Assume that a functional central limit theorem holds for the sequence $\{T_n\}$:

$$T_n(t) \xrightarrow{d} W, \text{ as } n \uparrow \infty, \tag{2}$$

in the sense of $C[0,1]$ equipped with the uniform metric, where W is the standard Brownian motion.

Fix $0 < t_0 < 1$. Since

$$h(f) = \max_{t_0 \leq t \leq 1} \frac{f(t)}{\sqrt{t}} \leq \frac{1}{\sqrt{t_0}} ||f||,$$

$h(\cdot)$ is a continuous functional. Therefore, if $n_0, n \to \infty$ in such way that $B_n/B_{n_0} \to 1/t_0$ then

$$P\left(\max_{n_0 \leq k \leq n} \left|\frac{T_k}{\sqrt{B_k}}\right| > b\right) \longrightarrow \alpha(t_0, b),$$

where

$$\alpha(t_0, b) = P\left(\max_{t_0 \leq t \leq 1} \left|\frac{W(t)}{\sqrt{t}}\right| > b\right). \tag{3}$$

Moreover, if A_n is a sample variance of T_n, and $\lim_{n\to\infty} A_n/B_n = 1$ a.s., then

$$P\left(\max_{n_0 \leq k \leq n} \left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) \longrightarrow \alpha(t_0, b), \ as \ n_0 \to \infty. \tag{4}$$

Using the approach in Sen (1981, Chapter 9, p. 244) and the algorithm in De Long (1981), for a given significance level $\alpha = \alpha(t_0, b)$ and a value of the power function at a specified alternative, one can construct a repeated significance test with the associated stopping time

$$\tau_1 = \min\left\{n_0 \leq n \leq N; |T_n| \geq b\sqrt{A_n}\right\}, \tag{5}$$

where n_0 is the initial sample size and N is the target sample size. The repeated significance test stops and rejects the null hypothesis if and only if $\tau_1 \leq N$. Otherwise, the null hypothesis is accepted.

Lemma 1. *If (1) holds, then as $n_0 \to \infty$*

$$\frac{B_{n_0}}{B_N} \to t_0. \tag{6}$$

Proof. From the definition of N we have

$$\frac{B_{N-1}}{B_{n_0}} < \frac{1}{t_0} \leq \frac{B_N}{B_{n_0}},$$

or

$$\frac{B_{N-1}}{B_N} < \frac{1}{t_0} \frac{B_{n_0}}{B_N} \leq 1.$$

Hence (6) follows. ∎

Therefore, availability of a functional central limit theorem for the sequence $\{T_n, ; n \geq 1\}$, the conditions in (1) along with the $\lim_{n\to\infty} A_n/B_n = 1$ a.s., will ensure that the significance level of the repeated significance test, described above, is approximately equal to $\alpha(t_0, b)$.

Our main interest is in developing robust nonparametric repeated significance tests for the median of a family of distributions containing a class of distributions with heavy tails. In Glaz and Pozdnyakov (2004) a repeated significance tests has been derived for testing the median of a family of symmetric distributions whose tail behavior is asymptotically equivalent to that of a symmetric stable distribution with a specified exponent in the interval $(0, 2)$. The specified value of that exponent played a crucial role in determining the target sample size N and it is not clear how this can be accomplished otherwise. In the next Section we introduce the concept of a repeated significance test with a random sample size, that adapts itself to the distribution of the observed data.

3. The Adaptive Target Sample Size

In some applications the implementation of the repeated significance test, described in the previous section, requires additional assumptions on certain parameters of the distribution of the observed data. For example, the determination of the target sample size for the nonparametric repeated significance test in Glaz and Pozdnyakov (2004) requires specification of the asymptotic tail behavior of the distribution, under the null hypothesis. In practical applications this information might not be available. To address this issue we introduce here an adaptive target sample size.

Let A_n be a sample variance of T_n. Define a stopping time $\mathcal{N}$ by

$$\mathcal{N} = \inf\{k \geq n_0 : \frac{A_k}{A_{n_0}} \geq \frac{1}{t_0}\}, \tag{7}$$

where $0 < t_0 < 1$ is a design parameter. A repeated significance test with adaptive target sample size is defined as follows. At time $k \geq n_0$ observe T_k. Stop and reject H_0, if k is the smallest integer such that $A_k/A_{n_0} < 1/t_0$ and $|T_k| \geq b\sqrt{A_k}$. Otherwise, we stop monitoring at time $\mathcal{N}$ and accept H_a. The following result is of major importance for implementing the repeated significance test with adaptive target sample size.

Theorem 1. *Assume that the functional central limit theorem for the sequence* $\{T_n\}$ *holds, and the sequence of* B_n *satisfies (1). If the sample variance* A_n *satisfies*

$$\frac{A_n}{B_n} \to 1 \quad a.s., \tag{8}$$

then

$$P\left(\max_{n_0 \le k \le \mathcal{N}} \left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) \longrightarrow \alpha(t_0, b) \text{ as } n_0 \to \infty. \tag{9}$$

Proof. For any $\epsilon > -1$ define $N(\epsilon)$ by

$$N(\epsilon) = \inf\{k \ge n_0 : \frac{B_k}{B_{n_0}} \ge \frac{1+\epsilon}{t_0}\}.$$

First, estimate the probability in (9) from above. Fix $\epsilon > 0$. We have

$$\begin{aligned}
P\left(\max_{n_0 \le k \le \mathcal{N}} \left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) &\le P\left(\max_{n_0 \le k \le \mathcal{N}} \left|\frac{T_k}{\sqrt{A_k}}\right| > b, \mathcal{N} \le N(\epsilon)\right) \\
&\quad + P(\mathcal{N} > N(\epsilon)) \\
&\le P\left(\max_{n_0 \le k \le N(\epsilon)} \left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) + P(\mathcal{N} > N(\epsilon)) \\
&\le P\left(\max_{n_0 \le k \le N(\epsilon)} \left|\frac{T_k}{\sqrt{B_k}}\right| \sqrt{\frac{B_k}{A_k}} > b, \sup_{k \ge n_0} \frac{B_k}{A_k} \le 1+\epsilon\right) \\
&\quad + P\left(\sup_{k \ge n_0} \frac{B_k}{A_k} > 1+\epsilon\right) + P(\mathcal{N} > N(\epsilon)) \\
&\le P\left(\max_{n_0 \le k \le N(\epsilon)} \left|\frac{T_k}{\sqrt{B_k}}\right| > \frac{b}{\sqrt{1+\epsilon}}\right) \\
&\quad + P\left(\sup_{k \ge n_0} \frac{B_k}{A_k} > 1+\epsilon\right) + P(\mathcal{N} > N(\epsilon))
\end{aligned}$$

Note that almost sure consistency of A_n (8) implies (for example, Shiryaev 1995, p. 253) that

$$P\left(\sup_{k > n_0} \frac{B_k}{A_k} > 1+\epsilon\right) \to 0,$$

as $n_0 \to \infty$.

On the other hand, we also have

$$P(\mathcal{N} > N(\epsilon)) \to 0 \text{ as } n_0 \to \infty.$$

Indeed, as $n \to \infty$ with probability one

$$\frac{A_{n-1}}{A_n} = \frac{A_{n-1}}{B_{n-1}} \frac{B_n}{A_n} \frac{B_{n-1}}{B_n} \to 1.$$

Therefore, by the same arguments as in Lemma 1 we get that

$$\frac{A_{\mathcal{N}}}{A_{n_0}} \to \frac{1}{t_0} \quad \text{a.s. as } n_0 \to \infty.$$

Thus, we have

$$\frac{B_{\mathcal{N}}}{B_{n_0}} = \frac{B_{\mathcal{N}}}{A_{\mathcal{N}}} \frac{A_{n_0}}{B_{n_0}} \frac{A_{\mathcal{N}}}{A_{n_0}} \to \frac{1}{t_0} \quad \text{a.s. as } n_0 \to \infty,$$

which implies that

$$P(\mathcal{N} > N(\epsilon)) = P\left(\frac{B_{\mathcal{N}}}{B_{n_0}} > \frac{B_{N(\epsilon)}}{B_{n_0}}\right) \le P\left(\frac{B_{\mathcal{N}}}{B_{n_0}} > \frac{1+\epsilon}{t_0}\right) \to 0.$$

By the functional central limit theorem we find the following upper bound on the crossing probability in (9)

$$\lim_{n_0\to\infty} P\left(\max_{n_0\le k\le\mathcal{N}}\left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) \le \alpha\left(\frac{t_0}{1+\epsilon}, \frac{b}{\sqrt{1+\epsilon}}\right).$$

We now proceed to obtain similar estimates of the probability in (9) from below. Fix $0 < \epsilon < 1$. Then,

$$P\left(\max_{n_0\le k\le N(-\epsilon)}\left|\frac{T_k}{\sqrt{B_k}}\right| > b\sqrt{1+\epsilon}\right) \le$$

$$\begin{aligned}
&\le P\left(\max_{n_0\le k\le N(-\epsilon)}\left|\frac{T_k}{\sqrt{B_k}}\right| > b\sqrt{1+\epsilon}, \sup_{k\ge n_0}\frac{A_k}{B_k} \le 1+\epsilon\right) \\
&\quad + P\left(\sup_{k\ge n_0}\frac{A_k}{B_k} > 1+\epsilon\right) \\
&\le P\left(\max_{n_0\le k\le N(-\epsilon)}\left|\frac{T_k}{\sqrt{A_k}}\right|\sqrt{\frac{A_k}{B_k}} > b\sqrt{1+\epsilon}, \sup_{k\ge n_0}\frac{A_k}{B_k} \le 1+\epsilon\right) \\
&\quad + P\left(\sup_{k\ge n_0}\frac{A_k}{B_k} > 1+\epsilon\right) \\
&\le P\left(\max_{n_0\le k\le N(-\epsilon)}\left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) + P\left(\sup_{k\ge n_0}\frac{A_k}{B_k} > 1+\epsilon\right) \\
&\le P\left(\max_{n_0\le k\le N(-\epsilon)}\left|\frac{T_k}{\sqrt{A_k}}\right| > b, \mathcal{N} \ge N(-\epsilon)\right) \\
&\quad + P(\mathcal{N} < N(-\epsilon)) + P\left(\sup_{k\ge n_0}\frac{A_k}{B_k} > 1+\epsilon\right) \\
&\le P\left(\max_{n_0\le k\le\mathcal{N}}\left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) \\
&\quad + P(\mathcal{N} < N(-\epsilon)) + P\left(\sup_{k\ge n_0}\frac{A_k}{B_k} > 1+\epsilon\right)
\end{aligned}$$

By (8) we get that

$$P\left(\sup_{k>n_0}\frac{A_k}{B_k} > 1+\epsilon\right) \to 0,$$

as $n_0 \to \infty$. Moreover,

$$\begin{aligned}
P(\mathcal{N} < N(-\epsilon)) &= P\left(\frac{B_{\mathcal{N}}}{B_{n_0}} < \frac{B_{N(-\epsilon)}}{B_{n_0}}\right) = P\left(\frac{B_{\mathcal{N}}}{B_{n_0}} < \frac{B_{N(-\epsilon)-1}}{B_{n_0}}\frac{B_{N(-\epsilon)}}{B_{N(-\epsilon)-1}}\right) \\
&\le P\left(\frac{B_{\mathcal{N}}}{B_{n_0}}\frac{B_{N(-\epsilon)-1}}{B_{N(-\epsilon)}} < \frac{1-\epsilon}{t_0}\right) \to 0,
\end{aligned}$$

because with probability 1

$$\frac{B_{\mathcal{N}}}{B_{n_0}}\frac{B_{N(-\epsilon)-1}}{B_{N(-\epsilon)}} \to \frac{1}{t_0}.$$

Thus, by the functional central limit theorem we get

$$\lim_{n_0\to\infty} P\left(\max_{n_0\le k\le \mathcal{N}}\left|\frac{T_k}{\sqrt{A_k}}\right| > b\right) \ge \alpha\left(\frac{t_0}{1-\epsilon}, b\sqrt{1+\epsilon}\right).$$

Since $\epsilon > 0$ is arbitrary, and $\alpha(\cdot,\cdot)$ is continuous, we have (9). ■

4. An Application to Shift Model: Truncated Sums

In this section, by applying Theorem 1 to the functional central limit theorem for a sequence of truncated partial sums, a sequential testing procedure is developed for the shift model. We consider here the case of intermediate truncation, i.e., when the number of excluded observations is asymptotically negligible in comparison to the sample size. The use of a variable truncation level has certain benefits when information on the distribution of the data is not available. With this approach, we achieve a combination of robustness and effectiveness of the proposed testing procedure. If the distribution of the data is heavy tailed then a certain number of extreme observations is deleted and the resulting test statistic is asymptotically normal. If the distribution is well-behaved, for instance, it has a finite variance, then all the observations will be included in the observed sequence of partial sums, and as a result, we will obtain a sequential testing procedure that is based on the usual sequence of partial sums. Theorem 1 is especially useful here, as the variance of the observed sequence of partial sums grows faster than the sample size. Therefore, the target sample size in the classical repeated significance test cannot be deterministic without additional assumptions on the distribution of the observed data.

Let $\{X, X_i; i \ge 1\}$ be iid observations from a continuous distribution F symmetric about $-\infty < \theta < \infty$. We are interested in developing a sequential procedure, based on a sequence of truncated partial sums, for testing

$$H_0 : \theta = 0 \text{ vs } H_a : \theta \ne 0. \tag{10}$$

Let $\{d_n; n \ge 1\}$ be an increasing sequence of positive numbers such that

$$nP\left(|X| > d_n\right) \sim \gamma_n, \tag{11}$$

where $\gamma_n \nearrow \infty$, as $n \to \infty$. Define the truncated partial sums:

$$\tilde{S}_n = \sum_{i=1}^{n} X_i \mathrm{I}_{(|X_i|\le d_n)}. \tag{12}$$

Denote by

$$B_n = Var(\tilde{S}_n). \tag{13}$$

The following functional central limit theorem for $\{\tilde{S}_n\}$ was established in Pozdnyakov (2003).

Theorem 2. *(Pozdnyakov 2003) If the random variable X belongs to the Feller class, i.e.*

$$\limsup_{t\to\infty} \frac{t^2\mathbf{P}(|X| > t)}{\mathbf{E}\big(X^2\mathrm{I}_{|X|\le t}\big)} < \infty, \tag{14}$$

the average number of the excluded variables

$$n\mathbf{P}\big(|X| > d_n\big) \sim \gamma_n \nearrow \infty \tag{15}$$

and $lim_{n\to\infty} B_n/B_{n+1} = 1$*, then* $\tilde{S}_n(t) \xrightarrow{d} W(t)$ *in the sense* $(\mathcal{C}[0,1], \rho)$*, where* $\tilde{S}_n(t)$ *is the linear interpolation between points*

$$\Big(0,0\Big), \Big(\frac{B_1}{B_n}, \frac{\tilde{S}_1}{\sqrt{B_n}}\Big), ..., \Big(1, \frac{\tilde{S}_n}{\sqrt{B_n}}\Big).$$

Note that the variances of the truncated partial sums satisfy the monotonicity condition (1). A sample version that one can employ here is given by

$$A_n = \sum_{i=1}^{n} X_i^2 \mathrm{I}_{(|X_i| \le d_n)} - \frac{\tilde{S}_n^2}{\sum_{i=1}^{n} \mathrm{I}_{(|X_i| \le d_n)}}. \tag{16}$$

It was shown in Glaz and Pozdnyakov (2004), that the conditions of Theorem 2 with

$$\lim_{n\to\infty} \frac{n\mathbf{P}\big(|X| > d_n\big)}{\ln\ln(n)} = \infty \tag{17}$$

and

$$\ln\ln(B_n) = o(n), \tag{18}$$

imply that the sample variance A_n is almost sure equivalent to the population variance B_n. Note that the conditions (17) and (18) are not restrictive from the practical point of view.

These results allow us to develop a repeated significance test with adaptive target sample size. Let

$$\tau = \inf\Big\{k \ge n_0 : |\tilde{S}_k| \ge b\sqrt{A_k}\Big\} \tag{19}$$

be a stopping time, where n_0 is the initial sample size, and $\mathcal{N}$ is the adaptive target sample size defined by (7). The repeated significance test stops and rejects H_0 if and only if $\tau \le \mathcal{N}$. Therefore, $\tau \wedge \mathcal{N}$ is the stopping time associated with this repeated significance test.

We say that a random variable X has a *Cauchy*^{p} distribution iff

$$X \stackrel{d}{=} \mathrm{sign}(Y)|Y|^p,$$

where $p > 0$ and Y has a standard Cauchy distribution. If X has a *Cauchy*^{p} distribution, then it is symmetric and belongs to the Feller class for any $p > 0$. Moreover, $E(|X|^q) < \infty$, if $q < 1/p$.

To evaluate the performance of the proposed repeated significance test, we consider the following four distributions: Normal, Cauchy$^{1/2}$, Cauchy, and Cauchy^2. These distributions have very different tail behaviors and it is impossible to specify a deterministic target sample size in the repeated significance test based on the truncated sums $\tilde{S}_n$, discussed in Glaz and Pozdnyakov (2004), that guarantees a correct significance level α for all four distributions. Numerical results presented in Table 1 show that the introduction of an adaptive target sample size successfully addresses this problem. The truncation level $d_n = n^{1/4}$ was used. The design parameters corresponding to targeted values of $\alpha = .01$ and $.05$ were evaluated from the tables in De Long (1981). The simulated significance level is presented as the top value in the table. The simulated values of $E(\tau \wedge \mathcal{N})$ and $Var(\tau \wedge \mathcal{N})$, rounded to whole numbers, are presented as the bottom values in the table (in form $E(\tau \wedge \mathcal{N}) \pm \sqrt{Var(\tau \wedge \mathcal{N})}$). A simulation of 10000 trials was employed.

[Table 1 here]

We now present an example illustrating an algorithm for implementing this repeated significance test designed to achieve a specified level of power against a certain alternative. Suppose that the monitoring starts at $n_0 = 150$, the targeted significance level is set to be .05, the targeted power for Cauchy distribution at $\theta = .5$ is set to be .90, and the truncation level $d_n = n^{1/2}$.

As we have seen, the repeated significance test with adaptive target sample size has a correct probability of Type I error, regardless of the true distribution of the data. So, one can choose t_0 and b as if the data have a Cauchy distribution (with $\theta = 0$ under H_0 and with $\theta = .5$ under H_a). If the additional information on the distribution is available the repeated significance test with a deterministic target sample size N in Glaz and Pozdnyakov (2004, p. 94) can be employed. That test suggests that in order to obtain the targeted significance level and power, for the example under consideration, one should take $t_0 = 1/15$ and $b = 2.9$. Now, let us run the repeated significance test with adaptive target sample size with these design parameters. Theorem 1 implies that the probability of Type I error, α, of the test is about .05. Since in the Cauchy distribution case $A_N \approx A_{\mathcal{N}}$ (in the sense $A_N/A_{\mathcal{N}} \to 1$ a.s. as $n_0 \to \infty$), the power of the test with adaptive stopping stopping rule is going to be close to the power of the repeated significance test in Glaz and Pozdnyakov (2004). Indeed, a simulation of 10000 trials for the repeated significance test with adaptive target sample size resulted in observed power of .8961.

5. Comparison Between Repeated Significance Tests

In this section we present a comparison between a repeated significance test with adaptive target sample size, a classical repeated significance test that is based on the Donsker's Theorem (Billingsley 1995, p. 520) and a Cauchy score repeated significance test. If the distribution of X has a finite second moment, then the functional central limit theorem of Donsker is valid. Therefore, one can construct a repeated significance test with a finite target sample size (classical repeated significance test) based on a functional central limit theorem for the partial sums $S_n = X_1 + ... + X_n$. In Table 2, we present simulated values for the power function, $\hat{\pi}(\theta)$, for the classical repeated significance test (CRST) and the repeated significance test with adaptive target sample size (ARST), derived in Section 3. Simulation results are presented for the standard normal and Cauchy distributions and are based on 10000 trials. The design parameters $n_0 = 100, N = 750, t_0^{-1} = 7.5$ and $b = 2.8$ were used to achieve a targeted significance level of $\alpha = .0513$. For the adaptive repeated significance test a truncation level of $d_n = n^{.25}$ was employed. For normal data, this truncation has almost no effect. As a consequence, our sequential procedure does not lose any power in comparison with the classical repeated significance test.

If the distribution of the observed data is not normal and has heavy tails, for example assume it is Cauchy, then the adaptive repeated significance test outperforms the classical repeated significance test. In this case Donsker's theorem is not valid and the process associated with the self-normalized sums does not converge to the standard Brownian motion. As a result, the simulated probability of Type I error, $\widehat{\alpha} = \hat{\pi}(0)$, is far from the targeted one. Moreover, the power values show that the CRST fails to detect a deviation of the parameter θ from 0 at all. In contrast, our adaptive repeated significance test achieves a correct probability of Type I error, and the power function increases sharply with θ.

[Table 2 here]

One possible alternative to the classical repeated significance test is a repeated significance test based on Cauchy scores (CSRST). This sequential test employs partial sums of Cauchy scores:

$$\hat{S}_n = \frac{X_1}{1+X_1^2} + ... + \frac{X_n}{1+X_n^2},$$

rather than the standard partial sums S_n. Since Cauchy scores are almost surely bounded random variables, Donsker's theorem is applicable. This test performs well if the underlying distribution is indeed Cauchy. For the numerical example in Table 2, we get via simulation with 10000 trials, $\hat{\pi}(0) = .0544$ and $\hat{\pi}(.25) = .9865$.

The Cauchy score repeated significance test is a parametric procedure. It is designed specifically for the Cauchy model. Therefore, if the data are not Cauchy, this test will not perform well. For instance, assume that the distribution of X is a convolution of the standard Cauchy and the uniform distribution on the interval $[-5, 5]$. This distribution has the same asymptotic tail behavior as the standard Cauchy distribution, but the distortion of the shape of the distribution around 0 makes the Cauchy score repeated significance test less powerful. Numerical results presented in Table 3 show that the repeated significance test with adaptive target sample size outperforms the Cauchy score repeated significance test.

[Table 3 here]

6. Concluding Remarks

In this article, based on a functional central limit theorem for partial sums of truncated random variables and Theorem 1, a nonparametric repeated significance test with an adaptive target sample size has been derived. Numerical results for the shift model, given in Section 5, indicate that this test performs quite well. The main theoretical result of this article, Theorem 1, is quite general. Its use is not limited to a sequence of partial sums of truncated observations. To employ it in constructing repeated significance tests based on a sequence of statistics $\{T_n, ; n \geq 1\}$, a functional central limit theorem for $\{T_n, ; n \geq 1\}$ has to be valid. This suggests the following two open problems.

The first open problem is related to constructing a nonparametric repeated significance test based on partial sums of a trimmed sequence of observations. One can view trimming as an alternate approach to truncation for handling data modeled by heavy-tailed distributions. In that case a deterministic number of extreme observations is excluded. More specifically, let $\{X_{k,n} : 0 \leq k \leq n\}$ be the order statistics of $|X_1|, |X_2|, ..., |X_n|$,

$$|X_{1,n}| \leq ... \leq |X_{n,n}|.$$

The trimmed sum, $\bar{S}_n$, is defined as

$$\bar{S}_n = \sum_{i=1}^{n-k(n)} X_{i,n},$$

where $k(n)$ is a deterministic sequence. The cases when $k(n) \uparrow \infty$ and $k(n) = o(n)$ are referred to as intermediate trimming. If the number of the excluded terms $k(n) \sim \gamma_n$, where γ_n is average number of excluded terms in the truncated sums, then it is quite intuitive that under certain conditions the truncated and trimmed sums will be very close to each other (Hahn, Kuelbs, and Weiner 1991, p. 30). The trimmed sums have been studied extensively. In particular, the ordinary central limit theorem for symmetric distributions from the Feller class is established in Griffin and Pruitt (1987). However, to our best knowledge, a functional version of central limit theorem trimmed sums does not exist. Such result is needed if one wants to employ trimmed sums to develop a nonparametric repeated significance test with adaptive target sample size.

In a recent paper, Peng (2001) introduced a new estimator of the mean of a heavy tailed distribution with tail index from $(1, 2)$. It was shown that the estimator has an asymptotic normal distribution. The normalizing sequence for Peng's estimator is not proportional to $\sqrt{n}$. It depends on the tail index which is usually unknown in practice. Here too Theorem 1 could be employed to construct a nonparametric repeated significance test with adaptive target sample size. To achieve that, a functional central limit theorem for the sequence of estimators in Peng (2001) has to be established .

References

Armitage, P., 1975. Sequential Medical Trials, 2nd ed., Blackwell, Oxford.

Billingsley, P., 1995. Probability and Measure, 3nd Edition, Wiley-Interscience, New York.

Chang, I. S., Hsiung, C. A. and Hwang, L. C., 1999. Repeated significance tests of a multi-parameter in survival analysis. Statistica Sinica 9, 913-922.

De Long, D., 1981. Crossing probabilities for a square root boundary by a Bessel process, Communications in Statistics: Theory & Methods, Series A, 10, 2197-2213.

Glaz, J. and Pozdnyakov, V., 2005. A repeated significance test for distributions with heavy tails, Sequential Analysis 24, 77-98.

Griffin, P. and Pruitt, W., 1987. The central limit theorem for trimmed sums, Math. Proc. Cambridge Philos. Soc. 102, 329-349.

Gu, M. and Lai, T. L. 1998, Repeated significance testing with censored rank statistics in interim analysis of clinical trials. Statistica Sinica 8, 411-428.

Hahn, M., Kuelbs, J., and Weiner, D., 1991. Asymptotic behavior of partial sums: a more robust approach via trimming and self-normalization, Sum, trimmed sums and extremes, Progr. Probab. 23. 1-53.

Ho, C. H., 1998. Repeated significance tests on accumulating data of repairable systems. Communications in Statistics, Theory and Methods 27, 1181-1200.

Jennison, C. and Turnbull, B. W., 2000. Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRS, Boca Raton.

Kim, H. J., 1994. A repeated significance test in a linear model. Sequential Analysis 13, 113-126.

Lee, S. J., Kim, K. and Tsiatis, A. A., 1996. Repeated significance testing in longitudinal clinical trials. Biometrika 83, 779-789.

Lerche, H. R., 1986. An optimality property of the repeated significance test. Proceedings of the National Academy of Science USA 83, 1546-1548.

Peng, L., 2001. Estimating the mean of a heavy tailed distribution, Statistics & Probability Letters 52, 255-264.

Pozdnyakov, V., 2003. A Note on functional CLT for truncated sums, Statistics & Probability Letters 61, 277-286.

Sen, P. K., 1981. Sequential nonparametrics: invariance principles and statistical inference, Jonh Wiley & Sons, New York.

Sen, P. K., 1985. Theory and Applications of Sequential Nonparametrics. CBMS-NSF Regional Conference Series in Applied Mathematics 49, SIAM, Philadelphia.

Sen, P. K., 2002. Repeated significance tests in frequency and time domains. Sequential Analysis 21, 249-283.

Shiryaev, A.N., 1995. Probability, Springer, New York.

Siegmund, D., 1985. Sequential Analysis. Springer-Verlag, New York.

Takahashi, H., 1990. Asymptotic expansions for repeated sequential tests for the normal means. Journal of the Japanese Statistical Association 20, 51-60.

Whitehead, J., 1997. The Design and Analysis of Sequential Clinical Trials, 2nd ed., John Wiley & Sons Ltd., Chichester, West Sussex.

Table 1. Simulated Significance Levels and Expected Stopping Times, $n_0 = 100$, $d_n = n^{1/4}$

t_0^{-1}	b	$Normal$	$Cauchy^{1/2}$	$Cauchy$	$Cauchy^2$
4	3.3	.010	.010	.009	.008
		391 ± 52	319 ± 38	276 ± 36	260 ± 42
	2.7	.051	.047	.047	.046
		382 ± 69	313 ± 49	272 ± 43	256 ± 48
7.5	3.4	.010	.012	.010	.009
		729 ± 104	544 ± 70	439 ± 62	397 ± 69
	2.8	.051	.048	.048	.045
		711 ± 141	533 ± 92	429 ± 77	391 ± 79

Table 2. Simulated Significance Levels and Power,
$n_0 = 100$, $1/t_0 = 7.5$, $b = 2.8$, $\alpha = .0513$ and $d_n = n^{1/4}$ (for ARST)

	Normal		*Cauchy*	
θ	*CRST*	*ARST*	*CRST*	*ARST*
0	.0521	.0531	.0069	.0473
.05	.1872	.1885	.0082	.0751
.10	.6306	.6283	.0099	.1531
.15	.9445	.9357	.0111	.2847
.20	.9981	.9968	.0176	.4769
.25	1.0	.9999	.0204	.6682
.30	1.0	1.0	.0339	.8196

Table 3. Simulated Significance Levels and Power, Cauchy+U[-5;5]
$n_0 = 100$, $1/t_0 = 7.5$, $b = 2.8$, $\alpha = .0513$ and $d_n = n^{1/2}$ (for ARST)

θ	$CSRST$	$ARST$
0	.0532	.0507
.25	.0896	.1602
.50	.2239	.5082
.75	.4407	.8590
1.0	.7240	.9798