Repeated Significance Tests for Distributions with Heavy Tails

Joseph Glaz and Vladimir Pozdnyakov

Department of Statistics
University of Connecticut
Storrs CT, 06269-4120, USA
(e-mail: glaz@uconnvm.uconn.edu, vladimir.pozdnyakov@uconn.edu)

Abstract. In this article we survey recent results on the development of nonparametric repeated significance tests for distributions with heavy tails. The implementation of these tests depends on the invariance theorem for partial sums of truncated independent and identically distributed random variables. We also discuss a method for evaluating the power function and the expected stopping times associated with these testing procedures.
Keywords: Repeated Significance Test, Heavy Tails, Invariance Principle.

1 Introduction

Repeated significance tests were introduced in [Armitage, 1958]. Major developments in the theory and applications of repeated significance tests have been reported among others in: [Dias and Garsia, 1999], [Hu, 1988], [Jennison and Turnbull, 2000], [Lai and Siegmund, 1977, 1979], [Lalley, 1983], [Lerche, 1986], [Sellke and Siegmund, 1983], [Sen, 1981, 1985, 2002], [Siegmund, 1982, 1985], [Takahashi, 1990], [Whitehead, 1997], [Woodroofe, 1979, 1982], and [Woodroofe and Takahashi, 1982]. Sequential tests and repeated significance tests have been developed mostly for normal or a t-distribution models ([Siegmund, 1985] and [Takahashi, 1990]. For independent and identically distributed (iid) observations from a distribution with a finite mean nonparametric repeated significance tests have been discussed in [Sen, 1981, 1985, 2002].

In this article we review recent results in [Glaz and Pozdnyakov, 2005] and [Pozdnyakov and Glaz, 2005] for repeated significance tests for iid observations from a continuous symmetric distribution with heavy tails and infinite variance and possibly no mean, as in the case of the Cauchy distribution. The repeated significance test developed in [Glaz and Pozdnyakov, 2005] is nonparametric in nature and is applicable to a class of stable distributions with a specified tail behavior. The method used in deriving this repeated significance extends the approach in [Sen, 1981] and [Sen, 1985]. It is based on the invariance principle for partial sums of truncated random variables ([Pozdnyakov, 2003]). Pozdnyakov and Glaz ([Pozdnyakov and Glaz, 2005]) introduce a sequential test that is a repeated significance test with a random target sample size. It depends on the gross rate of the sample variance of the tests statistics used in the repeated significance test. This test is fully nonparametric and its implementation does not depend on the asymptotic tail behavior of the underlying model for the observed data.

The article is organized as follows. In Section 2, we describe the repeated significance test for the median of a continuous symmetric distribution with heavy tails in a class of stable distributions with a specified tail behavior. We discuss how one selects a continuation region associated with this repeated significance test for specified significance level, initial sample size and target sample size. Power calculations and evaluation of expected stopping times are discussed in [Glaz and Pozdnyakov, 2005].

In Section 3 we describe a repeated significance test with a random target sample size. An application of this repeated significance test for a shift model is presented along with evaluation its performance. Concluding results are presented in Section 4.

2 A Repeated Significance Test for a Median

Let $\{X, X_i; i \geq 1\}$ be iid observations from a continuous distribution F symmetric about $-\infty < \theta < \infty$. Assume that $E(X^2) = \infty$. We present below a repeated significance test with initial sample size n_0 and target sample size N for testing

$$H_0 : \theta = 0 \ vs \ H_a : \theta \neq 0. \tag{1}$$

Assume that $H_0 : \theta = 0$ and F belongs to the domain of attraction of a stable distribution with exponent $0 < \gamma < 2$, i.e.

$$E\left(X^2 I_{(|X \leq t|)}\right) \sim t^{2-\gamma} L(t),$$

where L is a slowly varying function. Here and in what follows we denote by $u(t) \sim v(t)$ the asymptotic equivalence of two functions $u(t)$ and $v(t)$, in the sense that $\lim_{t\to\infty} [u(t)/v(t)] = 1$. The classical repeated significance test is based on a sequence of partial sums ([Siegmund, 1985]). The problem we encounter here is that the sequence of partial sums from a distribution with an infinite second moment does not converge to a Brownian motion. To overcome this difficulty we employ partial sums of truncated random variables. Let $\{d_n; n \geq 1\}$ be an increasing sequence of positive numbers such that

$$nP\left(|X| > d_n\right) \sim \gamma_n,$$

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

$$S_n = \sum_{i=1}^{n} X_i I_{(|X_i| \leq d_n)}. \tag{2}$$

Denote by

$$B_n = Var(S_n).$$

The classical invariance principle (Donsker Theorem, [Billingsley, 1995, p. 520]) is not applicable here as the sequence of truncated partial sums, $\{S_n; n \geq 1\}$, does not have independent increments . Let $\{W(t); 0 \leq t < \infty\}$ be the standard Brownian motion. The following invariance principle will be used to construct the continuation region for the repeated significance test:

Theorem 1 *(Pozdnyakov 2003) If the random variable X belongs to the Feller class, i.e.*

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

the average number of the excluded variables

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

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

$$\Big(0, 0\Big), \Big(\frac{B_1}{B_n}, \frac{S_1}{\sqrt{B_n}}\Big), ..., \Big(1, \frac{S_n}{\sqrt{B_n}}\Big).$$

Since B_n is unknown, following Sen ([Sen, 1981, p. 249]), we replace it with an almost sure equivalent sequence of sample variances:

$$A_n = \sum_{i=1}^{n} X_i^2 \mathrm{I}_{(|X_i|\leq d_n)} - \frac{S_n^2}{\sum_{i=1}^{n} \mathrm{I}_{(|X_i|\leq d_n)}}. \tag{3}$$

Let

$$\tau = \min\Big\{n_0 \leq n \leq N; |S_n| \geq b\sqrt{A_n}\Big\}$$

be the stopping time associated with the repeated significance test, where n_0 is the initial sample size and N is the target sample size. This test stops and rejects H_0, given in Equation (1), if and only if $\tau \leq N$. Its power function is given by:

$$\pi(\theta) = P_\theta\left(\tau \leq N\right) = 1 - \beta(\theta),$$

where

$$\beta(\theta) = P_\theta\left(|S_n| < b\sqrt{A_n}; n_0 \leq n \leq N\right),$$

is the probability of type II error function and $\left\{b_n = b\sqrt{A_n}; n_0 \leq n \leq N\right\}$ is a sequence of constants that determine its continuation region. The significance level of this repeated significance test is given by:

$$\begin{aligned} \alpha &= \pi(0) = P_0\left(\tau \leq N\right) = 1 - \beta(0) \\ &= P_0\left\{\max_{n_0\leq n\leq N}\left(\frac{|S_n|}{\sqrt{A_n}}\right) \geq b\right\}. \end{aligned}$$

To implement this test an accurate approximation for $b = b(\alpha, n_0, N)$ is needed. The following result is central for achieving this goal.

Proposition 1 *(Glaz and Pozdnyakov 2005) Assume that F belongs to the domain of attraction of a continuous symmetric stable distribution with exponent $0 < \gamma < 2$. Then the following results are true:*
1) F belongs to the Feller class.
2) The average number of the excluded terms $n\mathbf{P}(|X| > dn^{\delta}) \nearrow \infty$ whenever $1 - \gamma\delta > 0$. In particular, any $0 < \delta < 1/2$ guarantees it for all $0 < \gamma < 2$.
3) If $1 - \gamma\delta > 0$ and $\lim_{n_0, N \to \infty}(n_0/N) = c < 1$, then

$$\max_{n_0 \le n \le N} \frac{|S_n|}{\sqrt{A_n}} \xrightarrow{d} \sup_{[t_0 \le t \le 1]} \frac{|W(t)|}{\sqrt{t}},$$

where

$$t_0 = c^{1+(2-\gamma)\delta}. \tag{4}$$

In view of this result, let the truncating levels $d_n = dn^{\delta}$, where $d > 0$ and $0 < \delta < 1/2$. Let $c = n_0/N$, be the ratio of the initial and target sample sizes of the repeated significance test. Then, $b = b(\alpha, n_0, N)$ can be approximated by $b_{t_0}(\alpha)$ by solving

$$\mathbf{P}\left(\sup_{[t_0, 1]} \frac{|W(t)|}{\sqrt{t}} > b_{t_0}(\alpha) \right) = \alpha,$$

where t_0 is given in Equation (4). The algorithm in [De Long, 1981] is used to evaluate $b_{t_0}(\alpha)$.

Example 1 *Domain of attraction of a Cauchy distribution with location parameter θ and scale parameter 1.*

Let $\{X_i; i \ge 1\}$ be a sequence of iid observations from a distribution F in the class of distributions with a domain of attraction of a Cauchy distribution with location parameter θ and scale parameter 1. Assume that $H_0 : \theta = 0$ is true. We consider here truncation levels $d_n = n^{1/4}, n_0 \le n \le N$. In Table 1, the performance of the proposed repeated significance test is evaluated in terms of accuracy of achieving an assigned significance level α, for given values of n_0 and N. The theoretical critical values $b_{t_0}(\alpha)$ and the corresponding targeted significance levels have been obtained from [De Long, 1981]. The achieved significance levels were evaluated from a simulation with 10,000 trials.

Table 1. Simulated Significance Levels

n_0	N	t_0	$b_{t_0}(\alpha)$	*targeted* α	*simulated* α
100	303	1/4	2.7	.0503	.0541
100	303	1/4	3.3	.0098	.0094
100	754	1/12.5	2.6	.0989	.1012
30	91	1/4	2.7	.0503	.0638
30	91	1/4	3.3	.0098	.0167
30	226	1/12.5	2.6	.0989	.1119

For small values, the achieved significance levels are close to targeted significance levels even for a moderate value of $n_0 = 30$. For larger significance levels one has to use higher initial values to get accurate approximations. A value of $n_0 = 100$ produced accurate results even for $\alpha = .10$.

3 A Repeated Significance Test with Adaptive Target Sample Size

The implementation of the repeated significance test in the previous section requires specification of the asymptotic tail behavior of the distribution under the null hypothesis. In some applications this might not be known. To address this issue, in ([Pozdnyakov and Glaz, 2005]) we introduced a non-parametric repeated significance test with adaptive target sample size.

Let T_n be a sequence of test statistics associated with a repeated significance test. Let A_n be a sample variances of T_n. Define a stopping time $\mathcal{N}$ by

$$\mathcal{N} = \inf\{k \geq n_0 : \frac{A_k}{A_{n_0}} \geq \frac{1}{t_0}\}, \tag{5}$$

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 central to the implementation of the repeated significance test with adaptive target sample size.

Theorem 2 *(Pozdnyakov and Glaz 2005) Assume that the functional central limit theorem for the sequence $\{T_n\}$ holds, and there exists a sequence of numbers $B_n \nearrow \infty$ and $B_n/Var(T_n) \to 1$ as $n \to \infty$. If the sample variance A_n satisfies*

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

then

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

Theorem 2 is applied to a functional central limit theorem for a sequence of truncated partial sums to develop a repeated significance test with random sample size for the shift model. In what follows we describe this test and present a simulation study to evaluate its performance.

Let $\{X, X_i; i \geq 1\}$ be iid observations from a continuous distribution F symmetric about $-\infty < \theta < \infty$. We are interested in testing sequentially

$$H_0 : \theta = 0 \text{ vs } H_a : \theta \neq 0.$$

Define the sequence of truncated partial sums S_n as in Section 2, Equations (2). Note that the variances of the truncated partial sums satisfy the monotonicity condition needed in Theorem 2 and that one can employ the version of the sample variances given in Equation (3). It was shown in [Glaz and Pozdnyakov, 2005], that the conditions of Theorem **??** along with

$$\lim_{n\to\infty} \frac{n\mathbf{P}\big(|X| > d_n\big)}{\ln\ln(n)} = \infty \tag{6}$$

and

$$\ln\ln(B_n) = o(n), \tag{7}$$

imply that A_n is almost sure equivalent to the population variance B_n. Note that conditions (6) and (7) are not restrictive from the practical point of view.

Based on these results, the following repeated significance test with adaptive target sample size is developed. Let

$$\tau = \inf\left\{k \geq n_0 : |S_k| \geq b\sqrt{A_k}\right\}$$

be a stopping time, where n_0 is the initial sample size, and $\mathcal{N}$ is the adaptive target sample size defined by (5). The repeated significance test stops and rejects H_0 if and only if $\tau \leq \mathcal{N}$. Hence, $\tau \wedge \mathcal{N}$ is the stopping time associated with this repeated significance test.

The following class of heavy tail distributions will be used in evaluating the performance of this repeated significance test. We say that a random variable X has a *Cauchy*^{p} distribution iff

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

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 different tail behaviors and it is impossible to specify a deterministic target sample size in the repeated significance test based on the truncated sums S_n, discussed in [Glaz and Pozdnyakov, 2005], 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 levels are presented as top values in the table. The simulated values of $E(\tau \wedge \mathcal{N})$ are rounded to whole numbers and are presented as the bottom values in the table. These values are based on a simulation with 10000 trials.

Table 2. 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	.0097	.0104	.0093	.0076
		391	319	276	260
	2.7	.0508	.0465	.0473	.0463
		382	313	272	256
7.5	3.4	.0099	.0117	.0099	.0085
		729	544	439	397
	2.8	.0511	.0484	.0480	.0447
		711	533	429	391

4 Concluding Remarks

In this article we reviewed two recently developed repeated significance tests for distributions with heavy tails. For additional results and discussion the readers are referred to the articles [Glaz and Pozdnyakov, 2005] and [Pozdnyakov and Glaz, 2005]. We would like to note that currently the applications of these results are restricted to symmetric distributions. To extend these results to non symmetric distributions presents new challenges. The first step in this direction is to extend the invariance theorem that has been established in [Pozdnyakov, 2003]. The development of inference procedures following these repeated significance tests are also of great interest in applications.

References

[Armitage, 1958]P. Armitage. Numerical studies in the sequential estimation of a binomial parameter. *Biometrika*, pages 1-15, 1958.

[Billingsley, 1995]P. Billingsley. *Probability and Measure, 3nd ed.* Wiley-Interscience, New York, 1995.

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

[Dias and Garsia, 1999]R. Dias and N.L. Garcia. Approximations for non-symmetric truncated sequential and repeated significance tests. *Sequential Analysis*, pages 107-117, 1999.

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

[Hu, 1988]I. Hu. Repeated significance tests for exponential families. *Annals of Statistics*, pages 1643-1666, 1988.

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

[Lai and Siegmund, 1977]T.L. Lai and D. Siegmund. A nonlinear renewal theory with application to sequential analysis I. *Annals of Statistiscs*, pages 946-954, 1977.

[Lai and Siegmund, 1979]T.L. Lai and D. Siegmund. A nonlinear renewal theory with application to sequential analysis II. *Annals of Statistiscs*, pages 60-76, 1979.

[Lalley, 1983]S. Lalley. Repeated likelihood ratio tests for curved exponential families. *Z. Wahrsch. Verw. Gebiete*, pages 293-321, 1983.

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

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

[Pozdnyakov and Glaz, 2005]V. Pozdnyakov and J. Glaz. A nonparametric repeated significance test with adaptive target sample size. *Journal of Statistical Planning and Inference*, (to appear).

[Sellke and Siegmund, 1983]T. Sellke and D. Siegmund. Sequential analysis of the proportional hazards model. *Biometrika*, pages 315-326, 1983.

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

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

[Sen, 2002]P.Q. Sen. Repeated significance tests in frequency and time domains. *Sequential Analysis*, pages 249-283, 2002.

[Shiryaev, 1995]A.N. Shiryaev. *Probability.* Springer, New York, 1995.

[Siegmund, 1982]D. Siegmund. Large deviations for boundary crossing probabilities. *Annals of Probability*, pages 581-588, 1982.

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

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

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

[Woodroofe, 1979]M. Woodroofe. Repeated likelihood ratio tests. *Biometrika*, pages 453-463, 1979.

[Woodroofe, 1982]M. Woodroofe. *Nonlinear Renewal Theory in Sequential Analysis.* SIAM, Philadelphia, 1982.

[Woodroofe and Takahashi, 1982]M. Woodroofe and H. Takahashi. Asymptotic expansions for the error probabilities of some repeated significance tests. *Annals of Statistics*, pages 895-908, 1982.