Nonparametric Repeated Significance Tests with Random Sample Size

Joseph Glaz

University of Connecticut

June 2009

Outline of the Talk

- Introduction
- A robust nonparametric repeated significance test
- Repeated significance tests with random sample size
- Summary and future work

Introduction

- Recent References and Related Applications
  - Pozdnyakov (2003)
  - Glaz and Pozdnyakov (2005)
  - Pozdnyakov and Glaz (2007)
  - Guerriero, Pozdnyakov, Glaz and Willett (2009)
  - Computer network traffic
  - Telecommunication systems
  - High frequency financial data
  - Risk management and insurance data
  - Sensor networks

Introduction

Let $X_1, X_2, ..., X_n, ...$ be a sequence of independent and identically distributed (iid) observations from a continuous distribution F with median $-\infty < \theta < \infty$.
We are interested in testing $H_0 : \theta = 0$ vs $H_a : \theta \neq 0$ via a nonparametric sequential procedure. Suppose that we want to be sure that at most N observations will be needed to reach a decision.
Assuming that $\sigma^2 = Var(X_1) < \infty$, Sen (1981 and 1985) developed a a *repeated significance test* (**RST**) via the following Donsker's invariance theorem.

Let $\{X, X_i\}_{i\geq 1}$ be i.i.d random variables, $EX = \mu$, $VarX = \sigma^2$ and $S_n = X_1 + X_2 + ... + X_n$. A Functional Central Limit Theorem (Donsker's Theorem, see Billingsley (1995, p. 520)) implies that if $S_n(t)$ is the linear interpolation between points

$$\Big(0, 0\Big), \Big(\frac{1}{n}, \frac{S_1 - \mu}{\sigma\sqrt{n}}\Big), ..., \Big(1, \frac{S_n - n\mu}{\sigma\sqrt{n}}\Big)$$

then

$$S_n(t) \xrightarrow{d} W$$

in the sense $\mathcal{C}[0, 1]$ with uniform metric ρ where W is standard Brownian motion on $[0, 1]$.

A Robust Nonparametric RST

Let $X, X_1, X_2, ..., X_n, ...$ be a sequence of iid observations from a continuous symmetric distribution F with median $-\infty < \theta < \infty$. For F in the class of heavy tail distributions with an infinite variance and possibly no mean, Glaz and Pozdnyakov (2005) derived a repeated significance test testing $H_0 : \theta = 0$ vs $H_a : \theta \neq 0$. This repeated significance test is constructed as follows. Let $\{b_n\}_{n\geq 1}$ be an increasing sequence of positive numbers such that

$$n\mathbf{P}\left(|X| > b_n\right) \sim \gamma_n \nearrow \infty.$$

Denote by S_n^*

$$S_n^* = \sum_{i=1}^{n} X_i \mathrm{I}_{(|X_i| \leq b_n)}, \qquad (1)$$

the partial sums of a truncated sequence of observations and let $B_n = Var(S_n^*)$.

A Robust Nonparametric RST

Let

$$\tau = \min \left\{ n_0 \leq n \leq N; |S_n^*| \geq b_n \sqrt{A_n} \right\}$$

be a stopping time, where

$$A_n = \sum_{i=1}^{n} X_i^2 \mathrm{I}_{|X_i| \leq b_n} - \frac{S_n^{*2}}{\sum_{i=1}^{n} \mathrm{I}_{|X_i| \leq b_n}}, \qquad (2)$$

is a sequence of sample estimators of B_n, n_0 and N are the initial and the target sample size, respectively. The repeated significance test stops and rejects H_0 if and only if $\tau \leq N$. The power function of this test is given by

$$\begin{aligned} \beta(\theta) &= P_\theta(\tau \leq N) = 1 - P_\theta(\tau > N) \\ &= 1 - P_\theta\left(|S_n^*| < b_n \sqrt{A_n}; n_0 \leq n \leq N \right). \end{aligned} \qquad (3)$$

A Robust Nonparametric RST

Since the sequence of partial sums of truncated random variables, S_n^*, is not a process with independent increments, the classical Donsker functional central limit theorem cannot be used. However, in the case of symmetric distributions, $\{S_n^*\}$ is a martingale. The following analog of Donsker's theorem plays an important role in the implementation of the repeated significance test given above.

A Robust Nonparametric RST

Theorem (Pozdnyakov 2003) If the random variable X belongs to the Feller class:

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

the average number of the excluded observations

$$nP\left(|X| > b_n\right) \sim \gamma_n \nearrow \infty,$$

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

$$\left(0,0\right), \left(\frac{B_1}{B_n}, \frac{S_1^*}{\sqrt{B_n}}\right), ..., \left(1, \frac{S_n^*}{\sqrt{B_n}}\right).$$

A Robust Nonparametric RST

Glaz and Pozdnyakov (2005) show that for the problem at hand

$$\frac{A_n}{B_n} \longrightarrow 1 \text{ a.s.}.$$

Therefore, under H_0 if $B_{n_0}/B_N \longrightarrow t_0$ and $N \longrightarrow \infty$, then

$$\max\left\{\frac{|S_n|}{\sqrt{A_n}}; n_0 \leq n \leq N\right\} \xrightarrow{d} \sup\left\{\frac{W(t)}{\sqrt{t}}; t_0 \leq t \leq 1\right\}$$

and consequently,

$$\beta(0) = P_0\left(\max_{n_0 \leq n \leq N}\left\{\frac{|S_n^*|}{\sqrt{A_n}}\right\} \geq b_n\right) \longrightarrow \alpha,$$

where the constant $b_n = b_n(\alpha)$ is the critical value that determines the continuation region of the repeated significance test.

A Robust Nonparametric RST

Glaz and Pozdnyakov (2005) derived an approximation for $b_n(\alpha)$ for the class of symmetric stable continuous distributions with exponent $0 < \gamma < 2$, i.e.

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

where $L(t)$ is a slowly varying function. Feller (1971, p. 313) shows that the above condition is equivalent to

$$\lim_{t \to \infty} \frac{t^2 P(|X| > t)}{E\left(X^2 \mathrm{I}_{(|X| \leq t)}\right)} = \frac{2-\gamma}{\gamma}.$$

Based on the invariance principle in Pozdnyakov (2003), it follows that for $b_n = bn^{\delta}$, $0 < \gamma < 2$, $0 < \delta < 1/2$, $n_0, N \to \infty$ and $n_0/N \to c$, $0 < c < 1$

$$\max\left\{\frac{|S_n^*|}{\sqrt{A_n}}; n_0 \leq n \leq N\right\} \xrightarrow{d} \sup_{[c^{1+(2-\gamma)\delta}, 1]} \frac{|W(t)|}{\sqrt{t}}. \qquad (4)$$

A Robust Nonparametric RST

The constant $b_n(\alpha)$ can be approximated by $b_{t_0}(\alpha)$ by solving

$$P\left(\sup_{[c^{1+(2-\gamma)\delta},1]}\left\{\frac{|W(t)|}{\sqrt{t}}\right\}\geq b_{t_0}(\alpha)\right)=\alpha, \tag{5}$$

using the approach in De Long (1981). Numerical results in Glaz and Pozdnyakov (2005) show that the approximations for the critical value $b_n(\alpha)$ are good.

Approximation for the critical value

We now present numerical results for evaluating the approximation for the critical value $b_n(\alpha)$ for data from a Cauchy distribution. Let us assume that X has the Cauchy distribution, i.e. $\gamma = 1$. We consider the truncation level $b_n = n^{1/4}$, i.e. $\delta = 1/4$. In Table 1 simulation results are presented. For each case the number of simulations performed is 10,000. The theoretical critical values and the corresponding significance levels are taken from De Long (1981).

Table 1

n_0	N	t_0	$b_{t_0}(\alpha)$	theoretical 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

Table 1. Simulation Results for Probability of Type I Error

Approximation the power function of RST

Let $X, X_1, X_2, ..., X_n, ...$ be a sequence of independent and identically distributed observations from a continuous distribution F symmetric about the median $-\infty < \theta < \infty$ and $E\left(X^2\right) = \infty$. The power function of the RST is given by:

$$\begin{aligned} \beta(\theta) &= P_\theta(\tau \le N) = 1 - P_\theta(\tau > N) \\ &= 1 - P_\theta\left(|S_n^*| < b_n\sqrt{A_n}; n_0 \le n \le N\right). \end{aligned}$$

To approximate the power function an additional assumption has to be made:

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

where the constants $K > 0$ and $0 < \gamma < 2$ are known.

Approximation the power function

One can show that the power function of the RST can be approximated by

$$1 - P\left(\left|W(t) + \theta \frac{N^{[1-\delta(2-\gamma)]/2}}{K^{1/2} b^{(2-\gamma)/2}} t^{1/[1+\delta(2-\gamma)]}\right| < b_{t_0}(\alpha)\sqrt{t},\ \ t_0 \le t \le 1\right).$$

The power computation boils down to computing:

$$P\left(|W(t) + ct^{\rho}| < b\sqrt{t} \text{ for all } t \in [t_0, 1]\right),$$

where $1/2 < \rho < 1$. See Glaz and Pozdnyakov (2005).

Approximation the power function

Since

$$P\left(-b\sqrt{t} - ct^{\rho} < W(t) < b\sqrt{t} - ct^{\rho} \text{ for all } t \in [t_0, 1]\right)$$

$$= \int_{-b\sqrt{t_0}-ct_0^{\rho}}^{b\sqrt{t_0}-ct_0^{\rho}} P\left(-b\sqrt{t} - ct^{\rho} < W(t) < b\sqrt{t} - ct^{\rho}, t \in [t_0, 1] \middle| W(t_0) = x\right)$$

$$\times P\left(W(t_0) \in dx\right),$$

computing the power function is equivalent to solving the following problem.

Approximation the power function

Consider the domain

$$D = \left\{(x, y) : -t_0 \leq y \leq 1 - t_0, -b\sqrt{t_0 + y} - c(t_0 + y)^\rho < x \right.$$
$$\left. < b\sqrt{t_0 + y} - c(t_0 + y)^\rho\right\}.$$

Let $\tau_D(x, y)$ be the first time when the "degenerated" two-dimensional diffusion $(x_t, y_t) = (x + W(t), y + t)$ exits from the domain D, where (x, y) belongs to the interior of the domain D. What is the probability that a Brownian motion starting at point x and at time y will stay inside the curved boundaries, i.e.

$$P\left(y_{\tau_D(x,y)} = 1 - t_0\right)?$$

Approximation for the power function

The generating operator of the diffusion (x_t, y_t) is given by

$$\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial y}.$$

By Venttsel (1996, p. 333) the function

$$v(x,y) = P\left(y_{\tau_D(x,y)} = 1 - t_0\right)$$

is the unique solution of the PDE

$$\frac{1}{2}\frac{\partial^2 v}{\partial x^2}(x,y) + \frac{\partial v}{\partial y}(x,y) = 0 \quad (x,y) \in D,$$

that satisfies the following boundary conditions:

1. $v(\pm b(t_0 + y)^{1/2} - c(t_0 + y)^{\rho}, y) \equiv 0,$
2. $v(x, 1 - t_0) \equiv 1.$

We can solve this parabolic equation numerically which in turn will yield an approximation for the power function of the RST.

Approximation for the power function

In Table 2 we present approximations for the power function of the RST for the Cauchy distribution for various choices of μ computed via the Brownian motion approximation and by simulations .The initial sample size n is 100. The target sample size N_0 is 303. These choices correspond to the first row of Table 1. However, in this case we choose a higher truncating level $b_n = bn^{\delta} = 5n^{1/4}$.
The multiplier $b = 5$ is taken in order to get a good approximation by the Brownian motion with the nonlinear drift. Note that the multiplier b does not have an effect on the approximation for the probability of type I error .
For the Cauchy distribution $K = 2/\pi$ and $\gamma = 1$.

Approximation for the power function

μ	BM approximation (4 refinements)	BM approximation (5 refinements)	Simulation (1000 simulations)
0	.0508	.0505	.044 (.0503
.25	.1918	.1913	.183
.5	.5958	.5962	.557
.75	.9185	.9186	.907
1	.9948	.9948	.991

Table 2. Approximations for the Power Function .

A RST with a Random Stopping Time

Let $X, X_1, X_2, ..., X_n, ...$ be a sequence of iid observations from a continuous symmetric distribution F in the class of heavy tail distributions with an infinite variance and possibly no mean. Let $-\infty < \theta < \infty$, be the median F. For testing $H_0 : \theta = 0$ vs $H_a : \theta \neq 0$, Pozdnyakov and Glaz (2007) derived a repeated significance test with random target sample size. Let A_n be a sample variance of S_n^*, given in (2) and (1), respectively. Define a stopping time $\mathcal{N}$ by

$$\mathcal{N} = \inf\{k \geq n_0 : \frac{A_k}{A_{n_0}} \geq \frac{1}{t_0}\}, \tag{6}$$

where $0 < t_0 < 1$ is a design parameter. A repeated significance test with random target sample size is defined as follows.

A RST with a Random Stopping Time

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 random target sample size defined in (6). The repeated significance test stops and rejects H_0 if and only if $\tau \leq \mathcal{N}$. Therefore, $\tau \wedge \mathcal{N}$ is the stopping time associated with this test. The following plays a key role in implementing the repeated significance test with random target sample size.

Theorem (Pozdnyakov and Glaz 2007). Assume that the functional central limit theorem for the sequence $\{S_n^*\}$ holds, and the sequence of $B_n = Var(S_n^*)$ satisfies: $B_n \nearrow \infty$ and $B_n/B_{n-1} \to 1$, as $n \nearrow \infty$. If

$$\frac{A_n}{B_n} \to 1 \quad a.s., \tag{7}$$

then

$$P\left(\max_{n_0 \leq k \leq \mathcal{N}} \left|\frac{S_k^*}{\sqrt{A_k}}\right| > b\right) \longrightarrow \alpha(t_0, b) \text{ as } n_0 \to \infty. \tag{8}$$

A RST with a Random Stopping Time

For observations modeled by a distribution in the Feller class, such as the Cauchy distribution, Theorem 1 implies that a functional central limit theorem holds and the repeated significance test presented above can be carried out. Pozdnyakov and Glaz (2007) present numerical results indicating that this sequential test performs well. The advantage of using the repeated significance test with random target sample size over the one investigated in Glaz and Pozdnyakov (2005) is that for the design of the test we do not need to specify the asymptotic tail behavior of the heavy tail distribution. For power calculations one still needs to specify the asymptotic tail behavior.

A RST with a Random Stopping Time

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 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.

A RST with a Random Stopping Time

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 10,000 trials was employed.

A RST with a Random Stopping Time

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

A comparison between RST's

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.

A comparison between RST's

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.

A comparison between RST's

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

Summary and conclusions

Based on a functional central limit theorem for partial sums of truncated random variables and Theorem (Pozdnyakov and Glaz 2007), a nonparametric repeated significance test with an adaptive target sample size has been derived. Numerical results for the shift model indicate that this test performs quite well. Theorem (Pozdnyakov and Glaz 2007) 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.

Summary and conclusions

Recently, Guerriero, Pozdnyakov, Glaz and Willett (2009) derived a repeated significance test with a random target sample size that is controlled by the total available resources to carry out the sequential testing procedure. This test was applied to a decentralized sequential detection problem in a sensor communication network with communication constraints.

Glaz and Kenyon (1985) presented an approach used in developing median unbiased confidence intervals after the completion of a sequential testing procedure. It will be of interest to develop the algorithms needed to implement this methodology to repeated significance tests with random target sample size.

References

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

Billingsley, P. (1995). *Probability and Measure*, 3rd ed.,Wiley-Interscience, New York.

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

Glaz, J. and Kenyon, J. R. (1995). Approximating the characteristics of sequential tests. Probability and Mathematical Statistics, 15, 311-325.

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

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.

Guerriero, M., Pozdnyakov, V., Glaz, J. and Willett, P. Randomly truncated sequential hypothesis tests. Technical Report, Department of Statistics, University of Connecticut.

References

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.

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.

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

Pozdnyakov, V. and Glaz, J. (2007). A nonparametric sequential test for distributions with heavy tails. Journal of Statistical Planning and Inference, 137, 869-878.

References

Sen, P. K. (1981). *Sequential Nonparametrics: Invariance Principles and Statistical Inference*. John Wiley & Sons, New York.

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

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.