Elect. Comm. in Probab. **14** (2009), 372–381

ELECTRONIC COMMUNICATIONS in PROBABILITY

EXPONENTIAL INEQUALITIES FOR SELF-NORMALIZED PROCESSES WITH APPLICATIONS

VICTOR H. DE LA PEÑA
Department of Statistics, Columbia University, New York, NY 10027-6902
email: vp@stat.columbia.edu

GUODONG PANG
Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027-6902
email: gp2224@columbia.edu

Submitted April 4, 2009, *accepted in final form* September 4, 2009

AMS 2000 Subject classification: Primary: 60E15, 60G42, 60G44, 68M20; Secondary: 62F03, 60G40
Keywords: self-normalization, exponential inequalities, martingales, hypothesis testing, stochastic Traveling Salesman Problem

Abstract

We prove the following exponential inequality for a pair of random variables (A,B) with $B > 0$ satisfying the *canonical assumption*, $E[\exp(\lambda A - \frac{\lambda^2}{2}B^2)] \leq 1$ for $\lambda \in \mathbb{R}$,

$$P\left(\frac{|A|}{\sqrt{\frac{2q-1}{q}\left(B^2 + (E[|A|^p])^{2/p}\right)}} \geq x\right) \leq \left(\frac{q}{2q-1}\right)^{\frac{q}{2q-1}} x^{-\frac{q}{2q-1}} e^{-x^2/2}$$

for $x > 0$, where $1/p + 1/q = 1$ and $p \geq 1$. Applying this inequality, we obtain exponential bounds for the tail probabilities for self-normalized martingale difference sequences. We propose a method of hypothesis testing for the L^p-norm ($p \geq 1$) of A (in particular, martingales) and some stopping times. We apply this inequality to the stochastic TSP in $[0,1]^d$ ($d \geq 2$), connected to the CLT.

1 Introduction

Self-normalized stochastic processes are frequently found in statistical applications. They have the property of (in the standard form) being unit free and frequently eliminate or weaken moment assumptions. The prototypical example of a self-normalized process is Student's t-statistic, which is used in statistical analysis to test if the mean of a normally distributed sample has a value specified in a null hypothesis when the standard deviation of the underlying distribution is unknown. Let $\{X_i : i \geq 1\}$ be a sequence of i.i.d. normal random variables with mean 0 and variance σ^2. The sample mean $\bar{X}_n = \sum_{i=1}^n X_i/n$ and the sample variance $s_n^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2/(n-1)$. The t-statistic $T_n = \sqrt{n}\bar{X}_n/s_n$ has the Student t-distribution of freedom $n-1$, which converges to a standard normal random variable as $n \to \infty$. Some limit theorems and moment bounds have been proved for the t-statistic by observing that T_n is a function of self-normalized sums:

$$T_n = \frac{S_n/V_n}{\sqrt{(n-(S_n/V_n)^2)/(n-1)}},$$

where $S_n = \sum_{i=1}^n X_i$ and $V_n^2 = \sum_{i=1}^n X_i^2$. The limit distribution of the self-normalized sums S_n/V_n has been proved by Efron (1969) and Logan et al. (1973). Giné, Götze and Mason (1997) prove that T_n has a limiting standard normal distribution if and only if X_1 is in the domain of attraction of a normal law by making use of exponential and L_p bounds for the self-normalized sums S_n/V_n. Since the 1990s, there have been active developments of the probability theory of self-normalized processes. We refer to de la Peña, Lai and Shao (2009) for the comprehensive review of the state of the art of the theory and its applications in statistical inference. Here we make a contribution to this theory by proving a new exponential inequality for self-normalized processes (Theorem 2.1). We start by considering a pair of random variables (A,B) with $B > 0$ satisfying the following

$$\textit{Canonical Assumption:} \quad E\left[\exp\left(\lambda A - \frac{\lambda^2}{2}B^2\right)\right] \le 1, \quad \lambda \in \mathbb{R}. \tag{1.1}$$

This assumption is satisfied by a wide class of stochastic processes, for example, sums of conditionally symmetric random variables (Lemma A.4, de la Peña, Klass and Lai, 2007), discrete-time martingale difference sequences (Lemma B.1, Bercu and Touati, 2008), and continuous-time (local, sup) martingales (§3.5, Karatzas and Shreve, 1991).

De la Peña (1999) proves the following exponential bounds, connected to the law of large numbers (LLN), for such a pair (A,B):

$$P(A > x, B^2 \le y) \le e^{-x^2/2y}, \quad \text{for} \quad x \in \mathbb{R}, \quad y > 0,$$

and

$$P\left(\frac{A}{B^2} > x, \frac{1}{B^2} \le y\right) \le e^{-x^2/2y}, \quad \text{for} \quad x \in \mathbb{R}, \quad y > 0.$$

By the method of mixtures, de la Peña, Klass and Lai (2004) prove the following exponential bound for such a pair (A,B),

$$P\left(\frac{|A|}{\sqrt{2[B^2+(E[B])^2]}} > x\right) \le \sqrt{2}e^{-x^2/2}, \quad \text{for} \quad x \ge 0. \tag{1.2}$$

It is connected to the central limit theorem (CLT) and provides related control on the tail. Here we will prove a new exponential inequality for $|A|/\sqrt{((2q-1)/q)[B^2+(E[|A|^p])^{2/p}]}$ with $1/p + 1/q = 1$ and $p \ge 1$ (Theorem 2.1), which differs from (1.2) in two aspects. First, $(E[B])^2$ (or $E[B^2]$) in (1.2) is replaced by $(E[|A|^p])^{2/p}$. Second, when $E[A^2] = E[B^2]$ or $E[|A|] = E[B]$, our new inequality improves (1.2). Here we emphasize that one can not remove $(E[B])^2$ in the denominator of (1.2) to obtain exponential inequalities for the ratio $|A|/B$ with the upper bound independent of the distribution of B or its moments, see Remark 2.3.

These inequalities have provided moment bounds for a wide class of discrete-time martingales self-normalized by the square root of the conditional variance. In particular, in de la Peña, Lai and

Shao (2009), the inequality (1.2) has been applied to prove the necessary upper and lower bounds in Theorem 8.6 in order to obtain the moderate deviation result for self-normalized U-statistics in Theorem 8.5, but our new inequality in Theorem 2.1 will provide tighter upper and lower bounds as in Theorem 8.6 there. Moreover, the inequality (1.2) can be used to test if the expectation of the quadratic variation of martingales is equal to (or greater than) some specific value. Our new inequality in Theorem 2.1 provides a method of testing for the L^p-norm of martingales.
For the stochastic Traveling Salesman Problem (TSP) to find the length T_n of the shortest path connecting n independent points uniformly distributed over $[0,1]^d$ ($d \geq 2$), Steele (1981) and Rhee and Talagrand (1987,1989a,1989b) have shown that the random variable T_n is concentrated around its mean $E[T_n]$ by proving various upper bounds for $P(|T_n - E[T_n]| \geq t)$ for $t > 0$ and $d = 2$. The inequality in Theorem 2.1 gives us a new inequality for the stochastic TSP over $[0,1]^d$ ($d \geq 2$), connected to the CLT; see Theorem 5.1.
The new inequality is presented and proved in §2. Then, it is applied to obtain an exponential upper bound for the tail probability for self-normalized martingale difference sequences in §3. We propose a method of hypothesis testing for L^p-norms ($p \geq 1$) for martingales and stopping times in §4. In §5, we present the new inequality for the stochastic TSP.

2 Exponential Inequalities

In this section, we present and prove the new exponential inequality for self-normalized processes.

Theorem 2.1. *Let (A,B) be a pair of random variables with $B > 0$ in the probability space $(\Omega, \mathscr{F}, P)$ satisfying the canonical assumption* (1.1)*. Suppose $E[|A|^p] < \infty$ for some $p \geq 1$. Then for any $x > 0$ and for $q \geq 1$ such that $1/p + 1/q = 1$,*

$$P\left(\frac{|A|}{\sqrt{\frac{2q-1}{q}\left(B^2 + (E[|A|^p])^{2/p}\right)}} \geq x\right) \leq \left(\frac{q}{2q-1}\right)^{\frac{q}{2q-1}} x^{-\frac{q}{2q-1}} e^{-x^2/2}. \tag{2.1}$$

In particular, if $E[A^2] < \infty$,

$$P\left(\frac{|A|}{\sqrt{\frac{3}{2}(B^2 + E[A^2])}} \geq x\right) \leq \left(\frac{2}{3}\right)^{2/3} x^{-2/3} e^{-x^2/2}, \quad \textit{for} \quad x > 0, \tag{2.2}$$

and if $E[|A|] < \infty$,

$$P\left(\frac{|A|}{\sqrt{2\left(B^2 + (E[|A|])^2\right)}} \geq x\right) \leq 2^{-1/2} x^{-1/2} e^{-x^2/2}, \quad \textit{for} \quad x > 0. \tag{2.3}$$

Moreover, if B satisfies $E[B^2] = E[A^2] < \infty$, the upper bound becomes

$$\min\left\{2^{1/3}, (2/3)^{2/3} x^{-2/3}\right\} e^{-x^2/2}, \quad \textit{for} \quad x > 0.$$

Remark 2.1. *When $E[A^2] = E[B^2]$, the upper bound of the new inequality* (2.2) *is much tighter than that of* (1.2) *(with $(E[B])^2$ replaced by $E[B^2]$). The ratio of the two upper bounds (after rearranging the constant 2 and 3/2 in the denominators) is $x^{-2/3} e^{-x^2/12} / \sqrt{2} < 1$ for $x \geq 1$.*

Remark 2.2. *It is remarkable that this new inequality does not impose any restrictions on the distribution of A except that the pair (A,B) satisfies the canonical assumption and $E[|A|^p] < \infty$ $(p \geq 1)$, and that it gives enormous freedom for the choice of B. Moreover, it gives a similar upper bound as the classic upper bound for $A \stackrel{d}{=} \mathcal{N}(0,\sigma^2)$:*

$$P\left(\frac{|A|}{\sigma} \geq x\right) \leq \frac{2}{\sqrt{\pi}} x^{-1} e^{-x^2/2}, \quad x > 0,$$

because by letting $B = (E[A^2])^{1/2}$ in (2.2), we have

$$P\left(\frac{|A|}{\sigma} \geq x\right) \leq \min\left\{2^{1/3}, 2^{2/3}3^{-1/3}x^{-2/3}\right\} e^{-x^2/6}, \quad x > 0.$$

Remark 2.3. *We emphasize that it is impossible to remove $E[A^2]$ in (2.2) or $(E[B])^2$ in (1.2) to obtain an exponential bound without knowing the distribution of B or its moments. This is because, by Theorem 3.3 in de la Peña, Klass and Lai (2004),*

$$E[(|A|/B)^2] \leq 1 + x_0^2 + E\left[\log\left(|\log(B \vee B^{-1})| \vee e\right)\right],$$

where x_0 is some large constant, and Example 2.6 in de la Peña, Klass and Lai (2007) shows that this upper bound is sharp.

Proof. First of all, we establish the following identity, for any $C > 0$,

$$E\left[\left(\frac{C}{B^2 + C}\right)^{1/2} \exp\left(\frac{A^2}{2(B^2 + C)}\right)\right] \leq 1.$$

This holds because by the canonical assumption and Fubini's Theorem,

$$\begin{aligned}
1 &\geq \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}C^{-1/2}} e^{-\frac{\lambda^2}{2C^{-1}}} E\left[\exp\left(\lambda A - \frac{\lambda^2}{2}B^2\right)\right] d\lambda \\
&= E\left[C^{1/2} \exp\left(\frac{A^2}{2(B^2 + C)}\right) \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\lambda + \frac{A}{B^2+C})^2}{\frac{2}{B^2+C}}\right) d\lambda\right] \\
&= E\left[\left(\frac{C}{B^2 + C}\right)^{1/2} \exp\left(\frac{A^2}{2(B^2 + C)}\right)\right].
\end{aligned}$$

Let $G \in \mathscr{F}$ by any measurable set. Then, by Markov's inequality,

$$\begin{aligned}
P\left(\frac{|A|}{\sqrt{B^2 + C}} \geq x, G\right) &= P\left(\frac{|A|^2}{4(B^2 + C)} \geq \frac{x^2}{4}, G\right) \\
&\leq P\left(\frac{|A|^{1/2}}{(B^2 + C)^{1/4}} e^{\frac{A^2}{4(B^2+C)}} \geq x^{1/2} e^{x^2/4}, G\right) \\
&\leq x^{-1/2} e^{-x^2/4} E\left[\left(\frac{A^2}{B^2 + C}\right)^{1/4} e^{\frac{A^2}{4(B^2+C)}} \mathbf{1}_G\right]. \qquad (2.4)
\end{aligned}$$

Now, by Hölder's inequality,

$$E\left[\left(\frac{A^2}{B^2+C}\right)^{1/4} e^{\frac{A^2}{4(B^2+C)}}\mathbf{1}_G\right] = E\left[\frac{C^{1/4}}{|A|^{1/2}}\left(\frac{A^2}{B^2+C}\right)^{1/4} e^{\frac{A^2}{4(B^2+C)}}\frac{|A|^{1/2}}{C^{1/4}}\mathbf{1}_G\right]$$
$$\leq \left(E\left[\left(\frac{C}{B^2+C}\right)^{1/2}\exp\left(\frac{A^2}{2(B^2+C)}\right)\right]\right)^{1/2}\left(E\left[\frac{|A|}{C^{1/2}}\mathbf{1}_G\right]\right)^{1/2} \leq \left(E\left[\frac{|A|}{C^{1/2}}\mathbf{1}_G\right]\right)^{1/2}.$$

If $E[|A|^p] < \infty$ for $p > 1$, we can choose $C = (E[|A|^p])^{2/p}$ so that for p and q satisfying $1/p+1/q = 1$, again by Hölder's inequality,

$$E\left[\frac{|A|}{C^{1/2}}\mathbf{1}_G\right] \leq \left(E\left[\frac{|A|^p}{C^{p/2}}\right]\right)^{1/p} P(G)^{1/q} = P(G)^{1/q}.$$

This implies that

$$P\left(\frac{|A|}{\sqrt{B^2+(E[|A|^p])^{2/p}}} \geq x, G\right) \leq x^{-1/2}e^{-x^2/4}P(G)^{\frac{1}{2q}}. \tag{2.5}$$

Now letting $G = \left\{\frac{|A|}{\sqrt{B^2+(E[|A|^p])^{2/p}}} \geq x\right\}$, we obtain

$$P\left(\frac{|A|}{\sqrt{B^2+(E[|A|^p])^{2/p}}} \geq x\right) \leq x^{-\frac{q}{2q-1}}e^{-\frac{q}{2(2q-1)}x^2}.$$

This will imply (2.1).
When B satisfies $E[B^2] = E[A^2] < \infty$, we replace the argument in (2.4) by

$$P\left(\frac{|A|}{\sqrt{B^2+C}} \geq x, G\right) = P\left(\frac{|A|^2}{4(B^2+C)} \geq \frac{x^2}{4}, G\right)$$
$$\leq e^{-x^2/4}E\left[\exp\left(\frac{A^2}{4(B^2+C)}\right)\mathbf{1}_G\right] \leq e^{-x^2/4}\left(E[B^2]/C+1\right)^{1/4}P(G)^{1/4},$$

which implies that $P(G) \leq 2^{1/3}e^{-x^2/3}$ if we choose $C = E[B^2]$ and $G = \left\{\frac{|A|}{\sqrt{B^2+E[A^2]}} \geq x\right\}$.
We can take limit on both sides of (2.1) by the monotone convergence theorem as $q \to \infty$ (or $p \to 1$) to obtain (2.3). $\square$

3 Inequalities for Martingale Difference Sequences

Bercu and Touati (2008) prove that martingale difference sequences satisfy the canonical assumption in the form of the following lemma.

Lemma 3.1. *Let $\{X_i : i \geq 1\}$ be a martingale difference sequence with respect to the filtration $\mathbf{F} = \{\mathscr{F}_n : n \geq 1\}$ and suppose that $E[X_i^2] < \infty$ for all $i \geq 1$. Then for all $\lambda \in \mathbb{R}$,*

$$E\left[\exp\left(\lambda \sum_{i=1}^{n} X_i - \frac{\lambda^2}{2}\left(\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n} E[X_i^2|\mathscr{F}_{i-1}]\right)\right)\right] \leq 1.$$

By Lemma 3.1 and Theorem 2.1, we obtain the following theorem for martingale difference sequences. In time series analysis, Theorem 3.1 can also be used to establish useful bounds for moving average sequences since they can be regarded as martingale difference sequences.

Theorem 3.1. *Let $\{X_i : i \geq 1\}$ be a martingale difference sequence with respect to the filtration $\mathbf{F} = \{\mathscr{F}_n : n \geq 1\}$ and suppose that $E[X_i^2] < \infty$ for all $i \geq 1$. Let T be any stopping time with respect the filtration $\mathbf{F}$ and assume $T < \infty$ almost surely. Then for all $\lambda \in \mathbb{R}$,*

$$E\left[\exp\left(\lambda \sum_{i=1}^{T} X_i - \frac{\lambda^2}{2}\left(\sum_{i=1}^{T} X_i^2 + \sum_{i=1}^{T} E[X_i^2|\mathscr{F}_{i-1}]\right)\right)\right] \leq 1, \tag{3.1}$$

and for $x > 0$,

$$P\left(\frac{|\sum_{i=1}^{T} X_i|}{\sqrt{\frac{3}{2}\left(\sum_{i=1}^{T} X_i^2 + \sum_{i=1}^{T} E[X_i^2|\mathscr{F}_{i-1}] + E[\sum_{i=1}^{T} X_i^2]\right)}} \geq x\right) \leq \left(\frac{2}{3}\right)^{2/3} x^{-2/3} e^{-x^2/2}. \tag{3.2}$$

Proof. We first consider the sequences $\sum_{i=1}^{T\wedge n} X_i$, which can be written as $\sum_{i=1}^{n} X_i \mathbf{1}_{\{T \geq i\}}$, where $x \wedge y = \min\{x, y\}$ for $x, y \in \mathbb{R}$. Notice that $\{X_i \mathbf{1}_{\{T\geq i\}} : i \geq 1\}$ is also a martingale difference sequence with respect to the filtration $\mathbf{F}$ since $E[X_i \mathbf{1}_{\{T\geq i\}}|\mathscr{F}_{i-1}] = \mathbf{1}_{\{T\geq i\}} E[X_i|\mathscr{F}_{i-1}] = 0$. And $E[(X_i \mathbf{1}_{\{T\geq i\}})^2] \leq E[X_i^2] < \infty$. So by Lemma 3.1, we have

$$E\left[\exp\left(\lambda \sum_{i=1}^{n} X_i \mathbf{1}_{\{T\geq i\}} - \frac{\lambda^2}{2}\left(\sum_{i=1}^{n} X_i^2 \mathbf{1}_{\{T\geq i\}} + \sum_{i=1}^{n} E[X_i^2|\mathscr{F}_{i-1}]\mathbf{1}_{\{T\geq i\}}\right)\right)\right] \leq 1,$$

which implies that

$$E\left[\exp\left(\lambda \sum_{i=1}^{T\wedge n} X_i - \frac{\lambda^2}{2}\left(\sum_{i=1}^{T\wedge n} X_i^2 + \sum_{i=1}^{T\wedge n} E[X_i^2|\mathscr{F}_{i-1}]\right)\right)\right] \leq 1.$$

Then (3.1) follows by Fatou's lemma. Thus, we obtain (3.2) from (3.1) and Theorem 2.1. □

Remark 3.1. *Lemma 8.11 in de la Peña, Lai and Shao* (*2009*) *proves that the martingale difference sequence $\{X_i : i \geq 1\}$ in Theorem 3.1 satisfies*

$$P\left(\frac{|\sum_{i=1}^{n} X_i|}{\sqrt{\sum_{i=1}^{n}(X_i^2 + E[X_i^2|\mathscr{F}_{i-1}] + 2E[X_i^2])}} \geq x\right) \leq \sqrt{2} e^{-x^2/4}, \tag{3.3}$$

for $x > 0$. Our inequality in (3.2) *provides a tighter upper bound than the inequality* (3.3) *for $x \geq 1$.*

Remark 3.2. *As a special case when $\{X_i : i \geq 1\}$ is a sequence of independent random variables with mean 0 and $E[X_i^2] < \infty$ for each i,* (3.1) *and* (3.2) *hold with $E[X_i^2|\mathscr{F}_{i-1}]$ replaced by $E[X_i^2]$. If, in addition, $\sigma^2 = E[X_i^2] < \infty$ for each i, then for $x > 0$,*

$$P\left(\frac{|\sum_{i=1}^{T} X_i|}{\sqrt{\frac{3}{2}(\sum_{i=1}^{T} X_i^2 + \sigma^2 T + \sigma^2 E[T])}} \geq x\right) \leq \left(\frac{2}{3}\right)^{2/3} x^{-2/3} e^{-x^2/2}. \tag{3.4}$$

Lemma 8.8 in de la Peña, Lai and Shao (*2009*) *says that for such a sequence $\{X_i : i \geq 1\}$,*

$$P\left(\frac{|\sum_{i=1}^{n} X_i|}{\sqrt{5\sum_{i=1}^{n} E[X_i^2] + \sum_{i=1}^{n} X_i^2}} \geq x\right) \leq \sqrt{2} e^{-x^2/8}. \tag{3.5}$$

Our inequality (3.2) *provides a tighter upper bound than the inequality* (3.5) *for large x.*

Remark 3.3. *Lemma A.4 in de la Peña, Klass and Lai* (*2007*) *proves that for a sequence $\{X_i : i \geq 1\}$ of conditionally symmetric random variables adapted to the filtration $\mathbf{F} = \{\mathscr{F}_i : i \geq 1\}$,*

$$E\left[\exp\left(\lambda \sum_{i=1}^{n} X_i - \frac{\lambda^2}{2} \sum_{i=1}^{n} X_i^2\right)\right] \leq 1, \quad \lambda \in \mathbb{R}.$$

By Theorem 2.1, we obtain for $x > 0$,

$$P\left(\frac{|\sum_{i=1}^{n} X_i|}{\sqrt{\frac{3}{2}(\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n} E[X_i^2])}} \geq x\right) \leq \min\left\{2^{1/3}, (2/3)^{2/3} x^{-2/3}\right\} e^{-x^2/2}.$$

4 Applications to Hypothesis Testing

With the new inequality (2.1), we propose a method to test the L^p-norm ($p \geq 1$) for a random variable A. We first choose another positive random variable B such that the pair (A, B) satisfies the canonical assumption. The null hypothesis is $\mathbf{H}_0 : (E[|A|^p])^{1/p} \geq \mu$ and the alternative hypothesis is $\mathbf{H}_1 : (E[|A|^p])^{1/p} < \mu$. The null hypothesis will be rejected if

$$\frac{A}{\sqrt{\frac{2q-1}{q}(B^2 + (E[|A|^p])^{2/p})}} \geq x_\alpha,$$

where x_α is such that

$$P\left(\frac{A}{\sqrt{\frac{2q-1}{q}(B^2 + \mu^2)}} \geq x_\alpha\right) \leq \alpha, \quad \text{and} \quad \left(\frac{q}{2q-1}\right)^{2/3} x_\alpha^{-\frac{q}{2q-1}} e^{-x_\alpha^2/2} = \alpha.$$

As a special case when $p = 2$, we can test the variance of A with mean 0. A useful application is to test if the L^p-norms of martingales are equal to some specific values.

4.1 Testing for Stopping Times

For the sequence $\{X_i : i \geq 1\}$ and the stopping time T in Remark 3.2, the inequality (3.4) can be used to test if $E[T] = a$ ($E[T] \geq a$) for some constant $a > 0$. The null hypothesis is $\mathbf{H}_0 : E[T] = a$ (or $E[T] \geq a$) and the alternative hypothesis is $\mathbf{H}_1 : E[T] \neq a$ (or $E[T] < a$).

5 Application to the Stochastic TSP

In the stochastic modeling of the TSP, let $X_1, ..., X_n$ be i.i.d. uniformly distributed on $[0,1]^d$ ($d \geq 2$) and T_n be the shortest closed path through the n random points $X_1, ..., X_n$. It has been shown that the deviation of T_n from its mean $E[T_n]$ is remarkably small, see Steele (1981) and Rhee and Talagrand (1987, 1989a, 1989b). In particular, by Azuma's inequality, we have for $n \geq 2$,

$$P(|T_n - E[T_n]| \geq t) \leq \begin{cases} \exp(-t^2/(C \log n)), & \text{for} \quad d = 2, \\ \exp(-t^2/(Cn^{(d-2)/d})), & \text{for} \quad d > 2. \end{cases} \tag{5.1}$$

for some constant C, see §2.1 in Steele (1997).
It has been shown in Theorem 3 of Steele (1981) that $E[(T_n - E[T_n])^k] < c_k n^{(k/2)(1-d/2)}$ for all $k \geq 1$, $d \geq 2$ and $n \geq 1$, where c_k is some positive constant. In particular, $Var(T_n) < \infty$ for all n. Cerf et al. (1997) did numerical simulation for $d = 2$ to support that the variable $T_n - E[T_n]$ should have a Gaussian distribution as $n \to \infty$. Here we obtain the following upper bound, connected to the CLT.

Theorem 5.1. *For the stochastic TSP problem,*

$$P\left(\frac{|T_n - E[T_n]|}{\sqrt{\frac{3}{2}[Var(T_n) + \sum_{i=1}^n (d_i^2 + E[d_i^2|\mathscr{F}_{i-1}])]}} \geq t \right) \leq \left(\frac{2}{3}\right)^{2/3} t^{-2/3} \exp\left\{-\frac{t^2}{2}\right\}, \tag{5.2}$$

for $t > 0$, where $d_i = E[T_n|\mathscr{F}_i] - E[T_n|\mathscr{F}_{i-1}]$ for $1 \leq i \leq n$, and $\mathscr{F}_i = \sigma\{X_1, ..., X_i\}$, the σ-algebra generated by $X_1, ..., X_i$.

Proof. First, since T_n is $\mathscr{F}_n$ measurable, we can write

$$T_n - E[T_n] = \sum_{1 \leq k \leq n} d_k,$$

where $\{d_k : 1 \leq k \leq n\}$ is a martingale difference sequence. It can be easily checked that

$$||d_i||_\infty \leq C_1(d)(n - i + 1)^{-1/d}, \quad \text{and} \quad E[d_i^2|\mathscr{F}_{i-1}] \leq C_2(d)(n-i)^{-2/d}, \tag{5.3}$$

for all $1 \leq i \leq n$, where $|| \cdot ||_\infty$ is the essential supremum norm; see (2.8) in Steele (1997) and Corollary 5 in Rhee and Talagrand (1987). So $E[d_i^2] < \infty$ for all $1 \leq i \leq n$ and the conditions in Lemma 3.1 are satisfied.
Then, by Theorem 3.1, we have that for $t > 0$,

$$P\left(\frac{|T_n - E[T_n]|}{\sqrt{\frac{3}{2}\sum_{i=1}^n (d_i^2 + E[d_i^2|\mathscr{F}_{i-1}] + E[d_i^2])}} \geq t \right) \leq \left(\frac{2}{3}\right)^{2/3} t^{-2/3} e^{-t^2/2}.$$

Thus, (5.2) is obtained by noticing that

$$E\Big[\sum_{i=1}^n d_i^2\Big] = E\Big[\Big(\sum_{i=1}^n d_i\Big)^2\Big] = E[(T_n - E[T_n])^2] = Var(T_n).$$

□

Remark 5.1. *By replacing the denominator in* (5.2) *by its upper bound which can be derived from* (5.3)*, we obtain an inequality comparable to* (5.1)*. In particular, when* $d = 2$*, if we replace the denominator* $(3/2)[Var(T_n) + \sum_{i=1}^n (d_i^2 + E[d_i^2|\mathscr{F}_{i-1}])]$ *in* (5.2) *by its upper bound* $C \log n$ *for some constant* C*, we obtain*

$$P\left(\frac{|T_n - E[T_n]|}{\sqrt{C\log n}} \geq t\right) \leq \left(\frac{2}{3}\right)^{2/3} t^{-2/3} \exp\left\{-\frac{t^2}{2}\right\},$$

which implies that

$$P\left(|T_n - E[T_n]| \geq y\sqrt{n}\right) \leq \left(\frac{2}{3}\right)^{2/3} \left(\frac{y}{\sqrt{C(\log n)/\sqrt{n}}}\right)^{-2/3} \exp\left\{-\frac{y^2 n}{2C\log n}\right\}.$$

This inequality is related to the inequality (3.5) *in Rhee and Talagrand* (*1987*).

Remark 5.2. *In the proof of Theorem 5.1, by* (5.3)*, we have*

$$\begin{aligned} & \sum_{i=1}^n (d_i^2 + E[d_i^2|\mathscr{F}_{i-1}] + E[d_i^2]) \\ \leq \quad & C_1(d)\sum_{i=1}^n (n-i+1)^{-2/d} + 2C_2(d))\sum_{i=1}^n (n-i)^{-2/d} \equiv \frac{2}{3}C(d,n)^2. \end{aligned}$$

Then, for $t > 0$*,*

$$P\left(|T_n - E[T_n]| \geq t\right) \leq \left(\frac{2}{3}\right)^{2/3} C(d,n)^{2/3} t^{-2/3} \exp\left\{-\frac{t^2}{2C(d,n)^2}\right\}. \tag{5.4}$$

For $d = 2$*, Rhee and Talagrand* (*1989b*) *have proved the following sharp upper bound:*

$$P(|T_n - E[T_n]| \geq t) \leq K\exp(-t^2/K), \quad t > 0, \quad n \geq 1, \tag{5.5}$$

for some constant K *independent of* n*. When* $d = 2$*, the constant* $C(d,n)$ *in* (5.4) *is of order* $O(\sqrt{\log n})$*. So when* $t > K\sqrt{\log n}$ *for some constant* $K > 0$*, the bound in* (5.4) *is tighter than that in* (5.5)*.*

Acknowledgement

This research was supported by NSF grant DMS-05-05949. We thank Prof. Isaac Meijelson for helpful comments.

References

[1] B. Bercu and A. Touati. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.* (2008), 18, 1848–1869. MR2462551 MR2462551

[2] N.J. Cerf, J. Boutet de Monvel, O. Bohigas, O.C. Martin and A.G. Percus. The random link approximation for the Euclidean Traveling Salesman Problem. *Journal de Physique I*. (1997), 7, 117–136.

[3] V.H. de la Peña. A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* (1999), 27, 537–564. MR1681153 MR1681153

[4] V.H. de la Peña, M.J. Klass and T.L. Lai. Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. *Ann. Probab.*. (2004), Vol. 32, No.3A, 1902–1933. MR2073181 MR2073181

[5] V.H. de la Peña, M.J. Klass and T.L. Lai. Pseudo-maximization and self-normalized processes. *Probability Surveys*. (2007), Vol. 4,172–192. MR2368950 MR2368950

[6] V.H. de la Peña, T.L. Lai and Q.M. Shao. *Self-Normalized Processes: Limit Theory and Statistical Applications.* Springer. (2009). MR2488094 MR2488094

[7] V. Egorov. On the Growth Rate of Moments of Random Sums. *Preprint.* (1998).

[8] E. Giné, F. Götze and D. Mason. When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* (1997), 25, 1514–1531. MR1457629 MR1457629

[9] B. Efron. Student's t-test under symmetry conditions. *J. Amer. Statist. Assoc.* (1969), 64, 1278–1302. MR0251826 MR0251826

[10] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. 2nd ed. Springer. (1991). MR1121940 MR1121940

[11] B.F. Logan, C.L. Mallows, S.O. Rice and L.A. Shepp. Limit Distributions of Self-Normalized Sums. *Ann. Probab.* (1973), 1, 788–809. MR0362449 MR0362449

[12] W.T. Rhee and M. Talagrand. Martingale inequalities and NP-complete problems. *Mathematics of Operations Research.* (1987), 12, 177–181. MR0882849 MR0882849

[13] W.T. Rhee and M. Talagrand. Martingale inequalities, interpolation and NP-complete problems. *Mathematics of Operations Research.* (1989a), 13, 91–96. MR0984560 MR0984560

[14] W.T. Rhee and M. Talagrand. A sharp deviation inequality for the stochastic Traveling Salesman Problem. *Ann. Probab.* (1989b), 17, 1–8. MR0972767 MR0972767

[15] J.M. Steele. Complete convergence of short paths and Karp's algorithm for the TSP. *Mathematics of Operations Research.* (1981), 6, 374–378. MR0629637 MR0629637

[16] J.M. Steele. *Probability Theory and Combinatorial Optimization.* CBMS-NSF regional conference series in applied mathematics. (1997). MR1422018 MR1422018