Sequential Estimation of the Mean of a Class of Skewed Distributions

Mohamed Tahir

Abstract

In this paper, we propose a sequential procedure \((t, \hat{\mu}_t)\) for estimating the mean, \(\mu\), of a class of skewed probability density functions, subject to the loss function \(L_a = a^2(\hat{\mu}_t - \mu)^2 + t\), where \(a\) is a given positive number, \(t\) is a stopping time of the type proposed by Robbins (1959) and \(\hat{\mu}_t\) is a bias-corrected estimator of \(\mu\). We provide a second-order asymptotic expansion, as \(a \to \infty\), for the regret with respect to the loss \(L_a\). For the Pareto and Skew-uniform distributions, the proposed sequential procedure \((t, \hat{\mu}_t)\) performs better than the procedure \((t, \bar{X}_t)\), in the sense that it has a lower asymptotic regret. Moreover, the regret is negative for large values of \(a\) under the Gamma, Pareto, Rayleigh and Skew-uniform distributions. Using the loss considered by Chow and Yu (1981) and Martinsek (1988), we propose a bias-corrected estimator of \(\mu\) and provide a second-order asymptotic expansion, as \(a \to \infty\), for the incurred regret.

1 Department of Mathematics, University of Sharjah, United Arab Emirates. E-mail: mtahir@sharjah.ac.ae

Article Info: Received: January 3, 2014. Revised: March 1, 2014. Published online: March 31, 2014.
**Mathematics Subject Classification:** 62F10; 62L10; 62L12

**Keywords:** Asymptotic expansion, bias-corrected estimator, excess over the stopping boundary, loss function, regret, skewness, stopping time.

## 1 Introduction

Let $X_1, X_2, \ldots$ be independent random variables with common probability density function $f_\theta(x)$, where the value of $\theta$ is unknown, but lies in some interval $\Omega \subset (-\infty, \infty)$. Suppose that $X_1, X_2, \ldots$ are to be observed sequentially up to stage $n$ at a cost of one unit per observation and that when observation is terminated, the population mean

$$\mu = \int_{-\infty}^{\infty} x f_\theta(x) \, dx$$

is estimated by an appropriate estimator, $\hat{\mu}_n$, and the loss incurred is of the form

$$L_a(\hat{\mu}_n, \theta) = a^2 (\hat{\mu}_n - \mu)^2 + n,$$

where $a$ is a known positive number, determined by the cost of estimation relative to the cost of a single observation. Robbins (1959) proposed the sequential procedure $(t, \bar{X}_t)$, which stops the sampling process after observing $X_1, \ldots, X_t$ and estimates $\mu$ by $\hat{\mu}_t = \bar{X}_t$, where

$$t = \inf \left\{ n \geq m_a : n > a \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}{n}} \right\}$$

with $m_a$ being a positive integer.

Let $\mathcal{C}$ denote the class of skewed probability density functions, $f_\theta(x)$, $\theta \in \Omega$, for which the skewness is independent of $\theta$. This class contains, among
others, the density functions of the following distributions:

1- GAMMA(α, θ): the Gamma distribution with known shape parameter α and scale parameter β = θ. Its density function is

\[ f_\theta (x) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{x/\theta}, \quad x > 0 \]

and its skewness is \( \gamma = \frac{2}{\sqrt{\alpha}} \).

2- PARETO(α, θ): the Pareto distribution with known shape parameter α > 0 and scale parameter β = θ. Its density function is

\[ f_\theta (x) = \frac{\alpha \theta^\alpha}{x^{\alpha+1}}, \quad x \geq \theta \]

and its skewness is \( \gamma = \frac{2(1+\alpha)^2}{\alpha - 2} \sqrt{\frac{\alpha-2}{\alpha}} \) for \( \alpha > 3 \).

3- RAYLEIGH(θ): the Rayleigh distribution with shape parameter α = θ. Its density function is

\[ f_\theta (x) = \frac{x}{\theta^2} e^{-x^2/2\theta^2}, \quad x > 0 \]

and its skewness is \( \gamma = \frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}} \).

4- SKEWUNIFORM(λ, θ): the Skew-uniform distribution with known λ and unknown θ. Its density function is

\[ f_\theta (x) = \frac{1}{\theta^2} [\max \{ \min \{ \lambda x, \theta \}, -\theta \} + \theta], \quad -\theta < x < \theta \]

for \(-\theta < x < \theta\) and its skewness is \( \gamma = \frac{2\lambda(5\lambda^2-9)}{5(3-\lambda^2)^{3/2}} \) for \(-\sqrt{3} < \lambda < \sqrt{3} \).

In this paper, we propose a bias-corrected estimator \( \hat{\mu}_t \) of \( \mu \) and provide a second-order asymptotic expansion, as \( a \to \infty \), for the regret \( r_a(t, \hat{\mu}_t) \) with respect to the loss defined by (1). It is seen that the asymptotic regret is negative for the Gamma, Pareto, Rayleigh and Skew-uniform distributions. We also provide second-order asymptotic expansion, as \( a \to \infty \), for the regret with respect to the more general loss function considered by Chow and Yu (1981) and Martinsek (1988).
In the Normal case, Starr and Woodroofe (1969) showed that $r_a(t, \bar{X}_a) = O(1)$ as $a \to \infty$. Woodroofe (1977) showed that $r_a(t, \bar{X}_a) = 0.5 + o(1)$ as $a \to \infty$ if $m_a \geq 4$.

For the Gamma and Poisson cases, Starr and Woodroofe (1972) and Vardi (1979) obtained bounded regret using stopping times different from the one in (2). For the distribution-free case, Ghosh and Mukhopadhyay (1979) and Chow and Yu (1981) established asymptotic risk efficiency based on (2) under certain conditions. Tahir (1989) proposed a class of bias-reduction estimators of the mean of the one-parameter exponential family and provided a second order approximation for the regret.

2 Preliminary Notes

Let $t$ be as in (2). Martinsek (1988) indicated that

$$E[\bar{X}_a] = \mu - \frac{\gamma}{2a} + o\left(\frac{1}{a}\right)$$

as $a \to \infty$, provided that $E[|X_1|^{8+p}] < \infty$ for some $p > 0$, where $\gamma$ denotes the population skewness; that is, $\gamma = \sigma^{-3}E[(X_1 - \mu)^3]$, where $\sigma$ is the population standard deviation. Thus, $\bar{X}_a$ is an asymptotically biased estimator of $\mu$ if $f_\theta(x) \in \mathcal{C}$. Consider the bias-corrected estimator

$$\hat{\mu}_n = \bar{X}_a + \frac{\gamma}{2a}$$

for $n \geq 1$. Then, $E[\hat{\mu}_n] = \mu + o(1)$ as $a \to \infty$, by (3).

In order to define the regret incurred by the sequential procedure $(t, \hat{\mu}_t)$ under the loss (1), we first assume that $X_1, \ldots, X_n$ have been observed sequentially up to a predetermined stage $n$ from a population with density function $f_\theta(x) \in \mathcal{C}$. The risk incurred by estimating $\mu$ by (4), subject to the loss (1), is
\[
R_a(n, \theta) = E[L_a(n, \hat{\mu}_n)]
\]
\[
= E[a^2 (\bar{X}_n - \mu)^2] + \frac{a^2 \gamma}{a^2} E[(\bar{X}_n - \mu)] + \frac{\gamma^2}{4} + n
\]
\[
= \frac{a^2 \sigma^2}{n} + \frac{\gamma^2}{4} + n,
\]
This risk is minimized with respect to \(n\) by choosing \(n\) as the greatest integer less than or equal to \(n_a = a\sigma\). The minimum risk is
\[
R^*_a(\theta) = R_a(n_a, \theta) = 2a\sigma + \frac{\gamma^2}{4}
\]
for \(a > 0\). Since \(\sigma\) is unknown, there is no fixed-sample-size procedure that attains the minimum risk in practice. Therefore, we propose to use the sequential procedure \((t, \hat{\mu}_t)\), where \(t\) be as in (2). The performance of this procedure is measured by its regret, which is defined below.

**Definition 2.1** The regret of the procedure \((t, \hat{\mu}_t)\) under the loss (2) is defined as
\[
r_a(t, \hat{\mu}_t) = E[L_a(t, \hat{\mu}_t)] - R^*_a(\theta) = E[a^2 (\hat{\mu}_t - \mu)^2 + t] - 2a\sigma - \frac{\gamma^2}{4}
\]
for \(a > 0\).

The stopping time \(t\) in (2) can be rewritten as
\[
t = \inf \left\{ n \geq m_a : n \left( \frac{V_n}{n} \right)^{-1/2} > a \right\},
\]
where
\[
V_n = \sum_{i=1}^{n} (X_i - \bar{X}_n)^2
\]
for \(n \geq 1\). Let \(U_a = t(V_t/t)^{-1/2} - a\) denote the excess over the stopping boundary. Chang and Hsiung (1979) showed that \(U_a\) converges in distribution to a random variable \(U\) as \(a \to \infty\).
Lemma 2.2. Let $t$ be as in (2). Then, $\frac{t}{a} \to \sigma$ w.p.1 as $a \to \infty$. Moreover, if $E[|X_1|^{8+p}] < \infty$ for some $p > 0$, then

$$E[t] = a + \nu - 0.5 - \frac{3}{8} \sigma^4 (\kappa - 1) + o(1)$$

as $a \to \infty$, where $\nu = E[U]$ is the asymptotic mean of the excess over the boundary and $\kappa = \sigma^{-4} E[(X_1 - \mu)^4]$ is the population kurtosis.

Proof: The first assertion follows from Lemma 1 of Chow and Robbins (1965). The second assertion is adopted from Chang and Hsiung (1979).

3 Main Results

3.1 Asymptotic regret under the loss (1)

Let $X_1, X_2, \ldots$ be as in Section 1. The following theorem provides a second-order asymptotic expansion for the regret in (5).

Theorem 3.1. Let $t$ be defined by (2) with $m_a$ being such that $\delta \sqrt{a} \leq m_a = o(a)$ as $a \to \infty$ for some $\delta > 0$. For any probability density function $f_\theta(x) \in \mathcal{C}$ with respect to which $E[|X_1|^{8+p}] < \infty$ for some $p > 0$,

$$r_a(t, \hat{\mu}_t) = 2.75 - 0.75\kappa + 2\gamma^2 - \frac{\gamma}{2} + o(1)$$

as $a \to \infty$.

Proof: Substituting (4) in (5) yields

$$r_a(t, \hat{\mu}_t) = E[a^2 (\bar{X}_t - \mu)^2 + t - 2a \sigma] + a\gamma E[(\bar{X}_t - \mu)]$$

$$= r_a(t, \bar{X}_t) + a\gamma E[(\bar{X}_t - \mu)]$$

for $a > 0$. Moreover,

$$a E[(\bar{X}_t - \mu)] = -\gamma/2 + o(1) \quad \text{and} \quad r_a(t, \bar{X}_t) = 2.75 - 0.75\kappa + 2\gamma^2 + o(1)$$
as \( a \to \infty \), by (3) and Martinsek (1983). Take the limit as \( a \to \infty \) in (7) and use (8) to complete the proof.

The distributions considered in Tables 1-5 in Section 4 below are positively skewed, except for the Skew-uniform distribution with \(-\sqrt{3} < \lambda < -\frac{3}{\sqrt{5}}\) and Skew-Laplace distribution with \( \lambda = 0.5 \). For Table 1, the minimum value of \( \rho^* \) is \( \frac{75}{28} \approx 2.68 \), which is attained when \( \alpha = 49 \). The tables show that

1- the sequential procedure \( (t, \hat{\mu}_t) \) is a clear improvement over the procedure \( (t, \overline{X}_t) \) since its asymptotic regret is lower, except for the Skew-uniform distribution with \( \lambda = -1.4 \).

2- the asymptotic regret of the procedure \( (t, \hat{\mu}_t) \) under the PARETO(5, \( \theta \)) and SKEWUNIFORM(\( \lambda \), \( \theta \)) distributions is negative; which means that, for large values of \( a \) that the procedure \( (t, \hat{\mu}_t) \) performs better for these distributions than the best fixed-sample-size procedure.

### 3.2 Asymptotic regret under a more general loss function

Let \( X_1, X_2, \ldots \) be as in Section 1 and suppose that the loss function for estimating \( \mu \) is of the form considered by Chow and Yu (1981) and Martinsek (1988); that is,

\[
L_a(\mu_n^*, \theta) = a^2 \sigma^{2\beta-2} (\mu_n^* - \mu)^2 + n
\]  

for \( a > 0 \), where \( \beta \) is a given positive number and \( \mu_n^* \) is an estimator of \( \mu \). If \( \theta \) is estimated by \( \mu_n^* = \overline{X}_n \), Martinsek (1988) proposed to use the stopping time

\[
\]
\[ T = \inf \left\{ n \geq m_a : n > a \left( \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \right)^{\beta/2} \right\} \]  

(10)

and showed that the regret of the procedure \( (T, \bar{X}_T) \) under the loss (9) is

\[
r^*_a(T, \bar{X}_T) = E[a^2 \sigma^{2\beta-2} (\bar{X}_T - \mu)^2 + T] - 2a \sigma \beta \\
= 3 \beta - \frac{\beta^2}{4} + \left( \frac{\beta^2}{4} - \beta \right) \kappa + (\beta^2 + \beta) \gamma^2 + o(1) 
\]

(11)
as \( a \to \infty \), provided that \( E[|X_1|^{8+p}] < \infty \) for some \( p > 0 \). Straightforward calculations yield that, for large values of \( a \),

1) \( r^*_a(T, \bar{X}_T) \) is negative under the Gamma distribution with \( \alpha = 0.5 \) if \( 0 < \beta < 0.1 \).

2) \( r^*_a(T, \bar{X}_T) \) is negative under the Pareto distribution with \( \alpha = 5 \) if \( 0 < \beta < 1.24 \).

Martinsek (1988) also indicated that

\[
E[\bar{X}_T] = \mu - \frac{\beta \gamma}{2a\sigma^{\beta-1}} + o\left( \frac{1}{a} \right) 
\]

(12)
as \( a \to \infty \). Thus, if the distribution of \( X_1 \) is not symmetric, then \( \bar{X}_T \) is biased for large values of \( a \).

**Proposition 3.2:** Suppose that \( \gamma \) does not depend on \( \theta \) and let

\[
\mu^*_n = \bar{X}_n + \frac{\beta \gamma}{2a^{1/\beta}} n^{1-1/\beta} 
\]

for \( n \geq 1 \), where \( \beta > 1 \). Let \( T \) be defined by (10) with \( m_a \) being such that \( \delta \sqrt{a} \leq m_a = o(a) \) as \( a \to \infty \) for some \( \delta > 0 \). For any probability density function \( f_\theta(x) \in \mathcal{C} \) with respect to which \( E[|X_1|^{8+p}] < \infty \) for some \( p > 0 \),

\[
E[\mu^*_n] = \mu + o(1) \text{ as } a \to \infty. 
\]

**Proof:** For \( a > 0 \),

\[
aE[\mu^*_T - \mu] = aE[\bar{X}_T - \mu] + \frac{\beta \gamma}{2} E\left[ \left( \frac{T}{a} \right)^{(1-1/\beta)} \right] 
\]

(13)

Next, \( E[(T/a)^{(1-1/\beta)}] \to \sigma^{-1}\beta \text{ as } a \to \infty \) if \( \beta > 1 \), by the fact that \( T/a \to \sigma^{\beta} \).
w.p.1 as \( a \to \infty \) and (2.2) of Martinsek (1983). Taking the limit as \( a \to \infty \) in (13), using this fact and (12) yields the desired result.

Let \( r^*_a(T, \mu^*_T) \) denote the regret of the biased-corrected procedure \((T, \mu^*_T)\) under the loss (9). Then,

\[
r^*_a(T, \mu^*_T) = E[a^2 \sigma^{2\beta-2} (\bar{X}_T - \mu)^2 + T - 2a\sigma^\beta] + \beta\gamma \sigma^{2\beta-2} a^{2-1/\beta} E \left[ \frac{1}{T^{1-1/\beta}} (\bar{X}_T - \mu) \right] + \frac{\gamma^2 \sigma^{2\beta-2}}{4} E \left[ \frac{a^{2-2/\beta}}{T^{2-2/\beta}} \right]
\]

(14)

\[
= r^*_a(T, \bar{X}_T) + \beta\gamma \sigma^{2\beta-2} E \left[ \frac{a^{1-1/\beta}}{T^{1-1/\beta}} a (\bar{X}_T - \mu) \right] + \frac{\gamma^2 \sigma^{2\beta-2}}{4} E \left[ \frac{a^{2-2/\beta}}{T^{2-2/\beta}} \right]
\]

**Lemma 3.3:** Let \( T \) be as in (3.2) with \( \beta > 1 \). If \( E[|X_1|^{8+p}] < \infty \) for some \( p > 0 \), then

\[
E \left[ \frac{a^{1-1/\beta}}{T^{1-1/\beta}} a (\bar{X}_T - \mu) \right] = \frac{2(\beta-1)}{\sigma^{2\beta+1}} - \frac{\beta\gamma}{2\sigma^{2(\beta-1)}} o(1)
\]

as \( a \to \infty \).

**Proof:** First, observe that

\[
E \left[ \frac{a^{1-1/\beta}}{T^{1-1/\beta}} a (\bar{X}_T - \mu) \right] = E \left[ \left( \frac{a^{1-1/\beta}}{T^{1-1/\beta}} - \frac{1}{\sigma^{\beta-1}} \right) a (\bar{X}_T - \mu) \right] + \frac{1}{\sigma^{\beta-1}} a E[\bar{X}_T - \mu]
\]

(15)

for \( a > 0 \). Moreover,

\[
a E[\bar{X}_T - \mu] = -\frac{\beta\gamma}{2\sigma^{\beta-1}} + o(1)
\]

(16)

as \( a \to \infty \), by (12). Next, expand \( g(y) = 1/y^{1-1/\beta} \) at \( y = \sigma^\beta \), substitute \( y = a/T \) and multiply by \( a(\bar{X}_T - \mu) \) to obtain

\[
\left( \frac{a^{1-1/\beta}}{T^{1-1/\beta}} - \frac{1}{\sigma^{\beta-1}} \right) a (\bar{X}_T - \mu) = \left( \frac{1}{\beta} - 1 \right) T^{1-1/\beta} \left( \frac{T}{a} - \sigma^\beta \right) a (\bar{X}_T - \mu),
\]

(17)

where \( T \) is a random variable such that \( |T - \sigma^\beta| \leq |T/a - \sigma^\beta| \). Next, rewrite \( T \) in as \( T = \inf\{ n \geq m_n: n(V_n/n)^{\beta/2} > a \} \), where \( V_n \) is as in (6), and let
Sequential estimation of the mean

\[ U_a^* = T \left( \frac{V_T}{T} \right)^{-\beta/2} - a \]

denote the excess over the stopping boundary. Expanding \( h(y) = y^{-\beta/2} \) at \( y = \sigma^2 \), substituting \( y = V_T/T \) and multiplying by \( T \) yields

\[ T \left( \frac{V_T}{T} \right)^{-\beta/2} = \frac{T}{\sigma^\beta} - \frac{\beta}{2\sigma^{\beta+2}} (V_T - T\sigma^2) + \frac{\beta(\beta + 2)}{8\lambda_T^{\beta/2+2}} \frac{(V_T - T\sigma^2)^2}{T} \]

for \( a > 0 \), where \( \lambda_T \) is a random variable between \( V_T/T \) and \( \sigma^2 \). Furthermore, write \( V_T = \sum_{i=1}^{T} (X_i - \mu)^2 - T(X_T - \mu)^2 \) to obtain

\[ U_a^* = \frac{T}{\sigma^\beta} - a - \frac{\beta}{2\sigma^{\beta+2}} (W_T - T\sigma^2) + \frac{\beta}{2\sigma^{\beta+2}} T(X_T - \mu)^2 + \frac{\beta(\beta + 2)}{8\lambda_T^{\beta/2+2}} \frac{(V_T - T\sigma^2)^2}{T} \]

for \( a > 0 \), where \( W_T = \sum_{i=1}^{T} (X_i - \mu)^2 \). It follows from easily that

\[ \frac{T}{a} - \sigma^\beta = \frac{\sigma^\beta}{a} (U_a^* - \xi_T) + \frac{\beta}{2a^2} (W_T - T\sigma^2) \]

for \( a > 0 \), where

\[ \xi_T = \frac{\beta}{2\sigma^{\beta+2}} T(X_T - \mu)^2 + \frac{\beta(\beta + 2)}{8\lambda_T^{\beta/2+2}} \frac{(V_T - T\sigma^2)^2}{T} \].

Substituting (18) in (17) yields

\[ \left( a^{1-1/\beta} \frac{1}{T^{1-1/\beta} - \frac{1}{\sigma^{\beta-1}}} \right) a(X_T - \mu) = \left( 1 - \frac{1}{\beta} \right) \sigma^\beta T^{1/\beta-2}(U_a^* - \xi_T)(X_T - \mu) \]

\[ + \left( 1 - \frac{1}{\beta} \right) \frac{\beta}{2\sigma^2} T^{1/\beta-2} (W_T - T\sigma^2)(X_T - \mu) \]

\[ = \left( 1 - \frac{1}{\beta} \right) \sigma^\beta I_1(a) + \frac{1 - \beta}{2\sigma^2} I_2(a), \]

say. Let \( S_n = X_1 + \cdots + X_n, \ n \geq 1 \). Then,
\[ E[I_1(a)] = E \left[ \frac{T^{1/\beta-2}}{T} (U_a - \xi_T) (S_T - \mu T) \right] = -\frac{\sigma^\beta}{\sqrt{a\sigma^\beta}} E \left[ \left( \frac{a}{T} \right) T^{1/\beta-2} \left( S_T - \mu T \right) \right] \]

\[ \leq \frac{\sqrt{\sigma^\beta}}{\sqrt{a}} \sqrt{E[(U_a - \xi_T)^2]} \sqrt{E \left[ T^{2/\beta-4} \left( \frac{a}{T} \right)^2 \left( S_T - \mu T \right)^2 \right]} \]

\[ \leq \frac{1}{\sqrt{a}} \sqrt{2 \sigma^\beta E[U_a^2] + 2 \sigma^\beta E[\xi_T^2]} \sqrt{E \left[ T^{2/\beta-4} \left( \frac{a}{T} \right)^2 \left( S_T - \mu T \right)^2 \right]} \]

\[ \rightarrow 0 \quad (20) \]

as \( a \to \infty \), by Hölder’s inequality, the fact that \( T_\ast \to \sigma^\beta \) (\( |T_\ast - \sigma^\beta| \leq |T/a - \sigma^\beta| \to 0 \) w.p.1 since \( T/a \to \sigma^\beta \), as in the first assertion of Lemma 1), \( S_T - \mu T \) converges in distribution to a Standard Normal random variable by Anscombe’s theorem, the facts that \( E[U_a^2] \to E[U^2] < \infty \) and \( E[\xi_T^2] = O(1) \) as \( a \to \infty \) and (2.3), (2.8) and (2.9) of Martinsek (1983).

To evaluate \( E[I_2(a)] \), observe that

\[ I_2(a) = \frac{2a\sigma^\beta}{T} T^{1/\beta-2} \left( W_T - T \sigma^2 \right) \left( S_T - \mu T \right) \frac{a}{a\sigma^\beta} = 2\sigma^\beta \frac{a}{T} T^{1/\beta-2} \left( \frac{W_T - \sigma^2 T}{\sqrt{a\sigma^\beta}} + \frac{S_T - \mu T}{\sqrt{a\sigma^\beta}} \right)^2 \]

\[ - 2\sigma^\beta \frac{a}{T} T^{1/\beta-2} \left( \frac{W_T - \sigma^2 T}{\sqrt{a\sigma^\beta}} \right)^2 - 2\sigma^\beta \frac{a}{T} T^{1/\beta-2} \left( \frac{S_T - \mu T}{\sqrt{a\sigma^\beta}} \right)^2 \]

\[ \rightarrow 2\sigma^{1-2\beta} (2Z)^2 - 2\sigma^{1-2\beta} Z^2 - 2\sigma^{1-2\beta} Z^2 = 4\sigma^{1-2\beta} Z^2 \quad (21) \]

as \( a \to \infty \), by Anscombe’s theorem and the fact that \( T_\ast \to \sigma^\beta \) w.p.1 as \( a \to \infty \), where \( Z \) is a random variable having the Standard Normal distribution. Thus,

\[ E[I_2(a)] = 4\sigma^{1-2\beta} + o(1) \quad (22) \]

as \( a \to \infty \), by (21) and (2.3) and (2.4) of Martinsek (1983). Taking expectation in (19) and using (20) and (22) yields

\[ E \left[ \left( \frac{a^{1-1/\beta}}{T^{1-1/\beta}} - \frac{1}{\sigma^{\beta-1}} \right) a (X_T - \mu) \right] = \frac{2(1-\beta)}{\sigma^{2\beta+1}} + o(1) \quad (23) \]

as \( a \to \infty \). The lemma follows by taking the limit, as \( a \to \infty \), in (15) and using (23) and (16).
Theorem 3.4: Let $T$ be defined by (3.2) with $m_a$ being such that $\delta \sqrt{a} \leq m_a = o(a)$ as $a \to \infty$ for some $\delta > 0$ and $\beta > 1$. Then, for any probability density function $f_\theta(x) \in C$ with respect to which $E[|X_i|^{8+p}] < \infty$ for some $p > 0$,

$$r^*(T, \mu^*_T) = 3\beta - \frac{\beta^2}{4} + \left(\frac{\beta^2}{4} - \beta\right)\kappa + (\beta^2 + \beta)\gamma^2 + \frac{2\beta(\beta - 1)\gamma}{\sigma^3} - \frac{\beta^2\gamma^2}{2} + \frac{\gamma^2}{4} + o(1)$$

as $a \to \infty$.

Proof: The theorem follows by taking the limit, as $a \to \infty$, in (14) and using (11), Lemma 3.3 and the fact that

$$E\left[\frac{a^{2-2/\beta}}{T^{2-2/\beta}}\right] = \frac{1}{\sigma^{2\beta-2}} + o(1)$$

as $a \to \infty$ if $\beta > 1$, by the fact that $T/a \to \sigma^\beta$ w.p.1 as $a \to \infty$ (see the first assertion of Lemma 2.2) and (2.2) of Martinsek (1983).

4 Tables

The tables below show the values of $\rho$ and $\rho^*$ for certain skewed distributions, where $\rho^* = \rho - \frac{\gamma^2}{2}$ is the asymptotic regret incurred by the procedure $(t, \hat{\mu}_t)$ and $\rho = 2.75 - 0.75\kappa + 2\gamma^2$ represents the asymptotic regret incurred by the procedure $(t, \bar{X}_t)$.

**Table 1: GAMMA($\alpha, \theta$) with known $\alpha$**

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\kappa$</th>
<th>$\rho$</th>
<th>$\rho^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{2}{\sqrt{\alpha}}$</td>
<td>$\frac{6}{\alpha}$</td>
<td>$2.75 + \frac{3.5}{\alpha}$</td>
<td>$2.75 + \frac{3.5}{\alpha} - \frac{1}{\sqrt{\alpha}}$</td>
</tr>
</tbody>
</table>
Table 2: PARETO(5, θ)

<table>
<thead>
<tr>
<th>γ</th>
<th>κ</th>
<th>ρ</th>
<th>ρ*</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \frac{2(1 + \alpha)}{\alpha - 3} \sqrt{\frac{\alpha - 2}{\alpha}} = 4.6476 ]</td>
<td>[ \frac{6(\alpha^3 + \alpha^2 - 6\alpha - 2)}{\alpha(\alpha - 3)(\alpha - 4)} + 3 = 73.8 ]</td>
<td>-9.4</td>
<td>-11.7238</td>
</tr>
</tbody>
</table>

Table 3: RAYLEIGH(θ)

<table>
<thead>
<tr>
<th>γ</th>
<th>κ</th>
<th>ρ</th>
<th>ρ*</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \frac{2\sqrt{\pi}(\pi - 3)}{(4 - \pi)^{3/2}} = 0.6311 ]</td>
<td>[ 3 - \frac{6\pi^2 - 24\pi + 16}{(4 - \pi)^2} = 3.2451 ]</td>
<td>1.11245</td>
<td>0.7969</td>
</tr>
</tbody>
</table>

Table 4: SKEW-UNIFORM(λ, θ) with λ = -1.4 and λ = 1.35

<table>
<thead>
<tr>
<th>γ</th>
<th>κ</th>
<th>ρ</th>
<th>ρ*</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ = [ \frac{2\lambda(5\lambda^2 - 9)}{5(3 - \lambda^2)^{3/2}} ]</td>
<td>κ = [ \frac{2\lambda(5\lambda^2 - 9)}{5(3 - \lambda^2)^{3/2}} ]</td>
<td>[ -29.9109 ]</td>
<td>[ -29.6997 ]</td>
</tr>
</tbody>
</table>

γ < 0 if \( \lambda \in \left(-\sqrt{3}, -\frac{3}{\sqrt{5}} \right) \cup \left(0, \frac{3}{\sqrt{5}} \right) \) and γ > 0 if \(-\frac{3}{\sqrt{5}} < \lambda < 0\)

κ > 0 if \(-\sqrt{3} < \lambda < \sqrt{3}\) and κ < 0 if \(-1.4 < \lambda < -\sqrt{3}\) or \(-1.4 < \lambda < -3\)
5 Conclusion

We have proposed a bias-corrected estimator of the mean of a class of skewed probability density functions and provided a second-order asymptotic expansion for the regret under the squared error loss. The results indicate that the proposed procedure performs better than the best fixed-sample-size procedure when the observations are taken from the Gamma, Pareto, Rayleigh or Skew-uniform distribution. For a more general loss function, we have proposed bias-corrected estimator of the mean and provided a second-order asymptotic expansion for the incurred regret.

References


