A Class of Purely Sequential Minimum Risk Point Estimation Methodologies with Second-Order Properties

Jun Hu
(Joint work with Prof. Nitis Mukhopadhyay)

Department of Mathematics and Statistics
University of Vermont, Burlington

jun.hu@uvm.edu

June 12, 2019

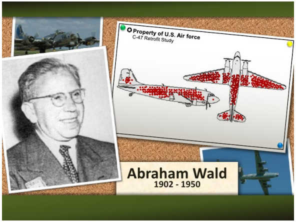
Outline

1. Introduction
   1.1. Sequential Analysis
   1.2. Minimum Risk Point Estimation (MRPE)
   1.3. Estimators for σ
2. Sequential MRPE
   2.1 Methodologies
   2.2 Asymptotics
3. Illustrations
4. Simulated Performances
5. Future Work

Selected Reference

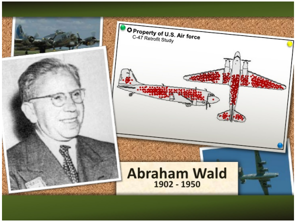
1.1. Sequential Analysis

- Sequential analysis is founded and developed by Abraham Wald during World War II.



1.1. Sequential Analysis

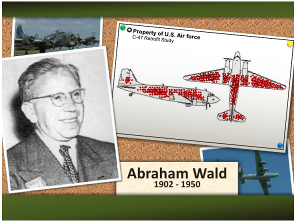
- Sequential analysis is founded and developed by Abraham Wald during World War II.



- The sample size is not predetermined.

1.1. Sequential Analysis

- Sequential analysis is founded and developed by Abraham Wald during World War II.



- The sample size is not predetermined.
- One observation is recorded at a time successively until termination.

1.2. Minimum Risk Point Estimation (MRPE)

- Originally formulated in Robbins (1959).

1.2. Minimum Risk Point Estimation (MRPE)

- Originally formulated in Robbins (1959).
- Assuming $X_1, ..., X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, with μ and σ^2 both unknown.

1.2. Minimum Risk Point Estimation (MRPE)

- Originally formulated in Robbins (1959).
- Assuming $X_1, ..., X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, with μ and σ^2 both unknown.
- Loss function:

$$L_n \equiv L_n(\mu, \overline{X}_n) = A(\overline{X}_n - \mu)^2 + cn, \tag{1}$$

where $A(> 0)$ and $c(> 0)$ are both known.

- Originally formulated in Robbins (1959).
- Assuming $X_1, ..., X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, with μ and σ^2 both unknown.
- Loss function:

$$L_n \equiv L_n(\mu, \overline{X}_n) = A(\overline{X}_n - \mu)^2 + cn, \tag{1}$$

  where $A(> 0)$ and $c(> 0)$ are both known.
- Risk function:

$$R_n(c) \equiv E_{\mu,\sigma}[L_n(\mu, \overline{X}_n)] = A\sigma^2 n^{-1} + cn. \tag{2}$$

1.2. Minimum Risk Point Estimation (MRPE)

- Optimal fixed sample size:

$$n^* \equiv n(c) = \sigma\sqrt{A/c}. \tag{3}$$

1.2. Minimum Risk Point Estimation (MRPE)

- Optimal fixed sample size:

$$n^* \equiv n(c) = \sigma\sqrt{A/c}. \tag{3}$$

- Minimum risk:

$$R_{n^*}(c) = 2cn^*. \tag{4}$$

1.2. Minimum Risk Point Estimation (MRPE)

- Optimal fixed sample size:

$$n^* \equiv n(c) = \sigma\sqrt{A/c}. \tag{3}$$

- Minimum risk:

$$R_{n^*}(c) = 2cn^*. \tag{4}$$

- NO fixed-sample-size procedure.

Solutions

- Two-stage: Stein (1945,1949)
- Purely sequential: Robbins (1959), Starr (1966)
- Three-stage: Mukhopadhyay (1990)
- Accelerated sequential: Mukhopadhyay and Solanky (1991), Mukhopadhyay (1996)

Solutions

- Two-stage: Stein (1945,1949)
- Purely sequential: Robbins (1959), Starr (1966)
- Three-stage: Mukhopadhyay (1990)
- Accelerated sequential: Mukhopadhyay and Solanky (1991), Mukhopadhyay (1996)

1.3. Estimators for σ

- σ is unknown.
- A general arbitrary estimator, assumed positive w.p.1.,

$$W_n \equiv W_n(X_1, ..., X_n).$$

Conditions on W_n

C1 *Independence*: $\overline{X}_n$ and $\{W_k; 2 \leq k \leq n\}$ are distributed independently for all $n \geq 2$.

Conditions on W_n

C1 *Independence*: $\overline{X}_n$ and $\{W_k; 2 \leq k \leq n\}$ are distributed independently for all $n \geq 2$.

C2 *Convergence in probability*: $W_n \overset{P_{\mu,\sigma}}{\rightarrow} \sigma$ as $n \rightarrow \infty$.

Conditions on W_n

C1 *Independence*: $\overline{X}_n$ and $\{W_k; 2 \leq k \leq n\}$ are distributed independently for all $n \geq 2$.

C2 *Convergence in probability*: $W_n \overset{P_{\mu,\sigma}}{\to} \sigma$ as $n \to \infty$.

C3 *Asymptotic normality*: $\sqrt{n}(\sigma^{-1} W_n - 1) \overset{\mathscr{L}}{\to} N(0, \delta^2)$ as $n \to \infty$.

Conditions on W_n

C1 *Independence*: $\overline{X}_n$ and $\{W_k; 2 \leq k \leq n\}$ are distributed independently for all $n \geq 2$.

C2 *Convergence in probability*: $W_n \overset{P_{\mu,\sigma}}{\to} \sigma$ as $n \to \infty$.

C3 *Asymptotic normality*: $\sqrt{n}(\sigma^{-1} W_n - 1) \overset{\mathscr{L}}{\to} N(0, \delta^2)$ as $n \to \infty$.

C4 *Uniform continuity in probability*: For every $\varepsilon > 0$, there exists a large ν and small $\gamma > 0$ for which $\forall n \geq \nu$,

$$P_{\mu,\sigma}\left(\max_{0 \leq k \leq n\gamma} |W_{n+k} - W_n| \geq \varepsilon \right) < \varepsilon.$$

C5 *Kolmogorov's inequality*: For every $\varepsilon > 0$, and some $2 \leq n_1 \leq n_2$, with $r \geq 2$,

$$P_{\mu,\sigma}\left(\max_{n_1 \leq n \leq n_2} |W_n - \sigma| \geq \varepsilon\right) \leq \varepsilon^{-r} E_{\mu,\sigma}[|W_{n_1} - \sigma|^r].$$

C5 *Kolmogorov's inequality*: For every $\varepsilon > 0$, and some $2 \leq n_1 \leq n_2$, with $r \geq 2$,

$$P_{\mu,\sigma}\left(\max_{n_1 \leq n \leq n_2} |W_n - \sigma| \geq \varepsilon\right) \leq \varepsilon^{-r} E_{\mu,\sigma}[|W_{n_1} - \sigma|^r].$$

C6 *Order of central absolute moments*: For $n \geq 2$ and $r \geq 2$,

$$E_{\mu,\sigma}[|W_n - \sigma|^r] = O(n^{-r/2}).$$

C5 *Kolmogorov's inequality*: For every $\varepsilon > 0$, and some $2 \leq n_1 \leq n_2$, with $r \geq 2$,

$$P_{\mu,\sigma}\left(\max_{n_1 \leq n \leq n_2} |W_n - \sigma| \geq \varepsilon\right) \leq \varepsilon^{-r} E_{\mu,\sigma}[|W_{n_1} - \sigma|^r].$$

C6 *Order of central absolute moments*: For $n \geq 2$ and $r \geq 2$,

$$E_{\mu,\sigma}[|W_n - \sigma|^r] = O(n^{-r/2}).$$

C7 *Wiener's condition*: $E_{\mu,\sigma}[\sup_{n \geq 2} W_n] < \infty$.

2.1 Methodologies

- Hu and Mukhopadhyay (2019)

- Hu and Mukhopadhyay (2019)
- Stopping rules:

$$\mathcal{P} : N_{\mathcal{P}} \equiv N_{\mathcal{P}}(c) = \inf\{n \geq m(\geq 2) : n \geq \sqrt{A/c}(W_n + n^{-\lambda})\}, \quad (5)$$

where $\lambda(> \frac{1}{2})$ is held fixed.

- Hu and Mukhopadhyay (2019)
- Stopping rules:

$$\mathcal{P} : N_{\mathcal{P}} \equiv N_{\mathcal{P}}(c) = \inf\{n \geq m(\geq 2) : n \geq \sqrt{A/c}(W_n + n^{-\lambda})\}, \quad (5)$$

where $\lambda(> \frac{1}{2})$ is held fixed.

- $P_{\mu,\sigma}\{N_{\mathcal{P}} < \infty\} = 1$ and $N_{\mathcal{P}} \uparrow \infty$ w.p.1 as $c \downarrow 0$.

- Upon $\{N_{\mathcal{P}}, X_1, ..., X_m, X_{m+1}, ..., X_{N_{\mathcal{P}}}\}$:

$$\overline{X}_{N_{\mathcal{P}}} \equiv N_{\mathcal{P}}^{-1} \Sigma_{j=1}^{N_{\mathcal{P}}} X_j. \tag{6}$$

- *Risk Efficiency:*

$$\xi_{\mathcal{P}}(c) = \frac{R_{N_{\mathcal{P}}}(c)}{R_{n^*}(c)} = \frac{1}{2} E_{\mu,\sigma}[N_{\mathcal{P}}/n^*] + \frac{1}{2} E_{\mu,\sigma}[n^*/N_{\mathcal{P}}];$$

- *Risk Efficiency:*

$$\xi_{\mathcal{P}}(c) = \frac{R_{N_{\mathcal{P}}}(c)}{R_{n^*}(c)} = \frac{1}{2}E_{\mu,\sigma}[N_{\mathcal{P}}/n^*] + \frac{1}{2}E_{\mu,\sigma}[n^*/N_{\mathcal{P}}];$$

- *Regret:*

$$\omega_{\mathcal{P}}(c) = R_{N_{\mathcal{P}}}(c) - R_{n^*}(c) = cE_{\mu,\sigma}[(N_{\mathcal{P}} - n^*)^2/N_{\mathcal{P}}].$$

- **Asymptotic First-Order Efficiency**:

$$\lim_{c \to 0} \mathrm{E}_{\mu,\sigma}[N_{\mathcal{P}}/n^*] = 1. \quad (7)$$

- **Asymptotic First-Order Efficiency**:

$$\lim_{c \to 0} \mathrm{E}_{\mu,\sigma}[N_{\mathcal{P}}/n^*] = 1. \quad (7)$$

- **Asymptotic First-Order Risk Efficiency**:

$$\lim_{c \to 0} \xi_{\mathcal{P}}(c) = 1, \quad (8)$$

where $\xi_{\mathcal{P}}(c) = \frac{1}{2} E_{\mu,\sigma}[N_{\mathcal{P}}/n^*] + \frac{1}{2} E_{\mu,\sigma}[n^*/N_{\mathcal{P}}]$.

- **Asymptotic First-Order Efficiency**:

$$\lim_{c \to 0} \mathrm{E}_{\mu,\sigma}[N_{\mathcal{P}}/n^*] = 1. \qquad (7)$$

- **Asymptotic First-Order Risk Efficiency**:

$$\lim_{c \to 0} \xi_{\mathcal{P}}(c) = 1, \qquad (8)$$

where $\xi_{\mathcal{P}}(c) = \frac{1}{2} E_{\mu,\sigma}[N_{\mathcal{P}}/n^*] + \frac{1}{2} E_{\mu,\sigma}[n^*/N_{\mathcal{P}}]$.

- **Asymptotic Second-Order Risk Efficiency**:

$$\omega_{\mathcal{P}}(c) = \delta^2 c + o(c) \text{ as } c \to 0, \qquad (9)$$

with δ^2 coming from (C3).

Illustrations

- What kinds of W_n would satisfy (C1)-(C7)?

Illustrations

- What kinds of W_n would satisfy (C1)-(C7)?
- Consider W_n what involves only

$$\mathbf{Y}_n = (X_1 - X_n, X_2 - X_n, ..., X_{n-1} - X_n).$$

Illustration 0: Sample Standard Deviation

- Robbins (1959): $W_n \equiv S_n$.

Illustration 0: Sample Standard Deviation

- Robbins (1959): $W_n \equiv S_n$.
- Stopping rule:

$$\mathcal{P} \equiv \mathcal{P}_0 : N_{\mathcal{P}_0} \equiv N_{\mathcal{P}_0}(c) = \inf\{n \geq m(\geq 2) : n \geq \sqrt{A/c}(S_n + n^{-\lambda})\}.$$

Illustration 0: Sample Standard Deviation

- Robbins (1959): $W_n \equiv S_n$.
- Stopping rule:

$$\mathcal{P} \equiv \mathcal{P}_0 : N_{\mathcal{P}_0} \equiv N_{\mathcal{P}_0}(c) = \inf\{n \geq m(\geq 2) : n \geq \sqrt{A/c}(S_n + n^{-\lambda})\}.$$

- The regret expansion:

$$\delta^2 = \frac{1}{2} \Rightarrow \omega_{\mathcal{P}_0}(c) = \frac{1}{2}c + o(c).$$

Illustration 1: Gini's Mean Difference (GMD)

- Gini (1914,1921):

$$\text{GMD:} \quad g_n = \binom{n}{2}^{-1} \Sigma\Sigma_{1\le i<j\le n} |X_i - X_j|.$$

Illustration 1: Gini's Mean Difference (GMD)

- Gini (1914,1921):

$$\text{GMD:} \quad g_n = \binom{n}{2}^{-1} \Sigma\Sigma_{1 \le i < j \le n} |X_i - X_j|.$$

- $W_n \equiv G_n = \frac{\sqrt{\pi}}{2} g_n$

Illustration 1: Gini's Mean Difference (GMD)

- Gini (1914,1921):

$$\text{GMD:} \quad g_n = \binom{n}{2}^{-1} \Sigma\Sigma_{1\le i<j\le n}|X_i - X_j|.$$

- $W_n \equiv G_n = \frac{\sqrt{\pi}}{2} g_n$
- Stopping rule:

$$\mathcal{P} \equiv \mathcal{P}_1 : N_{\mathcal{P}_1} \equiv N_{\mathcal{P}_1}(c) = \inf\{n \ge m(\ge 2) : n \ge \sqrt{A/c}(G_n + n^{-\lambda})\}.$$

Illustration 1: Gini's Mean Difference (GMD)

- Gini (1914,1921):

$$\text{GMD:} \quad g_n = \binom{n}{2}^{-1} \Sigma\Sigma_{1\le i<j\le n} |X_i - X_j|.$$

- $W_n \equiv G_n = \frac{\sqrt{\pi}}{2} g_n$
- Stopping rule:

$$\mathcal{P} \equiv \mathcal{P}_1 : N_{\mathcal{P}_1} \equiv N_{\mathcal{P}_1}(c) = \inf\{n \ge m(\ge 2) : n \ge \sqrt{A/c}(G_n + n^{-\lambda})\}.$$

- The regret expansion:

$$\delta^2 = \frac{\pi + 6\sqrt{3} - 12}{3} \approx 0.511 \Rightarrow \omega_{\mathcal{P}_1}(c) = 0.511c + o(c).$$

Illustration 2: Mean Absolute Deviation (MAD)

- Mean absolute deviation:

$$\text{MAD}: \quad m_n = n^{-1}\Sigma_{i=1}^{n}|X_i - \overline{X}_n|.$$

Illustration 2: Mean Absolute Deviation (MAD)

- Mean absolute deviation:

$$\text{MAD}: \quad m_n = n^{-1}\Sigma_{i=1}^{n}|X_i - \overline{X}_n|.$$

- $W_n \equiv M_n = \sqrt{\frac{\pi n}{2(n-1)}} m_n$

Illustration 2: Mean Absolute Deviation (MAD)

- Mean absolute deviation:

$$\text{MAD}: \quad m_n = n^{-1}\Sigma_{i=1}^{n}|X_i - \overline{X}_n|.$$

- $W_n \equiv M_n = \sqrt{\frac{\pi n}{2(n-1)}} m_n$
- Stopping rule:

$$\mathcal{P} \equiv \mathcal{P}_2 : N_{\mathcal{P}_2} \equiv N_{\mathcal{P}_2}(c) = \inf\{n \geq m(\geq 2) : n \geq \sqrt{A/c}(M_n + n^{-\lambda})\}.$$

Illustration 2: Mean Absolute Deviation (MAD)

- Mean absolute deviation:

$$\text{MAD}: \quad m_n = n^{-1}\Sigma_{i=1}^{n}|X_i - \overline{X}_n|.$$

- $W_n \equiv M_n = \sqrt{\frac{\pi n}{2(n-1)}}m_n$
- Stopping rule:

$$\mathcal{P} \equiv \mathcal{P}_2 : N_{\mathcal{P}_2} \equiv N_{\mathcal{P}_2}(c) = \inf\{n \geq m(\geq 2) : n \geq \sqrt{A/c}(M_n + n^{-\lambda})\}.$$

- The regret expansion:

$$\delta^2 = \frac{\pi - 2}{2} \approx 0.571 \Rightarrow \omega_{\mathcal{P}_2}(c) = 0.571c + o(c).$$

Table 1. Simulations from $N(5,4)$ with $A = 100, m = 10, \lambda = 2$ under 1000 runs implementing $\mathcal{P}_0 - \mathcal{P}_2$

n^*	$100c$	$\mathcal{P}$	$\overline{n}$	$s(\overline{n})$	$\widehat{\xi}$	$s(\widehat{\xi})$	δ^2	$\widehat{\omega}/c$
50	16	$\mathcal{P}_0$	50.012	0.1671	0.9880	0.003340	0.5	0.593131
		$\mathcal{P}_1$	50.313	0.1703	0.9879	0.003377	0.511	0.612431
		$\mathcal{P}_2$	50.259	0.1778	0.9872	0.003339	0.571	0.666431
100	4	$\mathcal{P}_0$	99.955	0.2408	0.9932	0.002404	0.5	0.599650
		$\mathcal{P}_1$	100.335	0.2347	0.9943	0.002306	0.511	0.561200
		$\mathcal{P}_2$	100.332	0.2495	0.9939	0.002327	0.571	0.636125
200	1	$\mathcal{P}_0$	200.012	0.3325	0.9969	0.001660	0.5	0.561100
		$\mathcal{P}_1$	200.255	0.3380	0.9971	0.001661	0.511	0.580400
		$\mathcal{P}_2$	200.026	0.3562	0.9962	0.001659	0.571	0.643800
400	0.25	$\mathcal{P}_0$	399.931	0.4588	0.9983	0.001146	0.5	0.531200
		$\mathcal{P}_1$	400.282	0.4508	0.9984	0.001114	0.511	0.514000
		$\mathcal{P}_2$	400.232	0.4873	0.9985	0.001145	0.571	0.598800

Accelerated Sequential MRPE Saving Sampling Operations

- Given the pilot sample size $m \geq 2, 0 < \rho \leq 1$ and $k \geq 1$, an integer, consider the following stopping rule:

$$T \equiv T(c) = \inf \left\{ n \geq 0 : m + kn \geq \rho \sqrt{A/c} \left[W_{m+kn} + (m+kn)^{-\lambda} \right] \right\}.$$

Accelerated Sequential MRPE Saving Sampling Operations

- Given the pilot sample size $m \geq 2, 0 < \rho \leq 1$ and $k \geq 1$, an integer, consider the following stopping rule:

$$T \equiv T(c) = \inf \left\{ n \geq 0 : m + kn \geq \rho \sqrt{A/c} \left[W_{m+kn} + (m+kn)^{-\lambda} \right] \right\}.$$

- The final sample size is then given by

$$N \equiv N(c) = \lfloor \rho^{-1}(m + kT) \rfloor + 1,$$

where $\lfloor u \rfloor$ means the largest integer smaller than u.

Accelerated Sequential MRPE Saving Sampling Operations

- Given the pilot sample size $m \geq 2, 0 < \rho \leq 1$ and $k \geq 1$, an integer, consider the following stopping rule:

$$T \equiv T(c) = \inf\left\{n \geq 0 : m + kn \geq \rho\sqrt{A/c}\left[W_{m+kn} + (m+kn)^{-\lambda}\right]\right\}.$$

- The final sample size is then given by

$$N \equiv N(c) = \lfloor \rho^{-1}(m + kT) \rfloor + 1,$$

where $\lfloor u \rfloor$ means the largest integer smaller than u.

- Operational time reduced by approximately $100(1 - k^{-1}\rho)\%$.

Selected References

- Gini, C. (1914). *L'Ammontare la Composizione della Ricchezza delle Nazioni*, Boca, Torino.
- Gini, C. (1921). Measurement of Inequality of Incomes, *Economic Journal* 31: 124-126.
- Hu, J. and Mukhopadhyay, N. (2019) Second-Order Asymptotics in a Class of Purely Sequential Minimum Risk Point Estimation (MRPE) Methodologies, *Japanese Journal of Statistics and Data Science* 2: 81-104.
- Mukhopadhyay, N. (1996). An Alternative Formulation of Accelerated Sequential Procedures with Applications to Parametric and Nonparametric Estimation, *Sequential Analysis 15*: 253-269.

- Mukhopadhyay, N., and Hu, J. (2017). Confidence Intervals and Point Estimators for a Normal Mean Under Purely Sequential Strategies Involving Gini's Mean Difference and Mean Absolute Deviation, *Sequential Analysis* 36: 210-239.
- Mukhopadhyay, N. and Solanky, T. K. S. (1991). Second Order Properties of Accelerated Stopping Times with Applications in Sequential Estimation, *Sequential Analysis 10*: 99-123.
- Robbins, H. (1959). Sequential Estimation of the Mean of a Normal Population, in *Probability and Statistics*, H. Cramér volume, Ulf Grenander, ed., pp. 235-245, Uppsala: Almquist & Wiksell.
- Starr, N. (1966). On the Asymptotic Efficiency of a Sequential Procedure for Estimating the Mean, *Annals of Mathematical Statistics 37*: 1173-1185.

- Stein, C. (1945). A Two Sample Test for a Linear Hypothesis Whose Power Is Independent of the Variance, *Annals of Mathematical Statistics* 16: 243-258.
- Stein, C. (1949). Some Problems in Sequential Estimation (Abstract), *Econometrica* 17: 77-78.


Thank
you