Introductory Sequential Methodologies with Applications: An Appreciation

Nitis Mukhopadhyay
www.stat.uconn.edu/~nitis

Department of Statistics
University of Connecticut-Storrs

The 24th BASS Conference and Workshop 2017

Monday, October 23, 2017

The Hotel Indigo
Savannah Georgia

!!! Greetings from Connecticut !!!

The Layout

1. Introduction and Key Ingredients:

- Statistical tests or confidence interval estimation with a predetermined sense of accuracy
- Some sense of optimality in place
- Appropriate sample size determination
- Appropriate data collection strategies
- Gather data sequentially, step-by-step, and “learn” as we proceed

Some Key References:

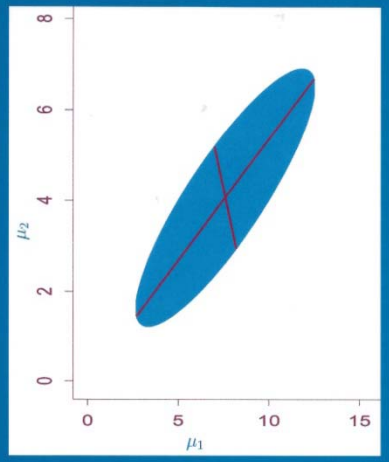
- Mahalanobis (1940)
- Stein (1945,1949)
- Wald (Wiley: 1947)
- Mukhopadhyay and Solanky (Dekker: 1994, Chapter 2)

- Ghosh, Mukhopadhyay, and Sen (Wiley: 1997, Chapter 6)
- Mukhopadhyay, Datta, and Chattopadhyay, eds. (Dekker: 2004)
- Mukhopadhyay and de Silva (CRC: 2009)

Note 1: The last-mentioned book includes a large number of standard methodologies with accompanying computer programs in executable forms

Note 2: We will implement some of them in a number illustrations to follow

Sequential Methods and Their Applications



Nitis Mukhopadhyay
Basil M. de Silva

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

2. Sequential Probability Ratio Test (SPRT)

X: A response variable with its probability distribution $f(x;\theta)$

X may be discrete or continuous, θ in Θ unknown

We may want to test

$H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_1$, $\theta_1 > \theta_0$ and θ_0, θ_1 from Θ.

With n independent responses $x_1, x_2, ..., x_n$, Neyman-Pearson theory will provide the most powerful (MP) test: Reject H_0 iff

$\frac{\Pi_{i=1}^{n} f(x_i;\theta_1)}{\Pi_{i=1}^{n} f(x_i;\theta_0)} \geq k$, that is the ratio is "large".

The cut-off $k(> 0)$ is chosen by the fixed size of the test, α.

We may also require power $1 - \beta$ with preassigned α, β. Then?

Let us look at a "Normal Mean" example.

2.1. Normal Mean Example

$X_1, ..., X_n$ are random samples from $N(\theta, \sigma^2)$ with θ unknown, but σ^2 known. We want to test

$H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_1$, $\theta_1 > \theta_0$.

Detect effect size "$\theta_1 - \theta_0$" efficiently!

The **N-P MP** level α test would reject H_0 iff $\overline{X}$ is "large",

that is iff $\overline{X} > \theta_0 + z_\alpha \sigma / \sqrt{n}$

where z_α is the upper $100\alpha\%$ point of $N(0, 1)$ distribution.

We further demand that type-II error must be $\leq \beta$, then the optimal fixed-sample-size n must be determined as:

$$n = \left(\frac{(z_\alpha + z_\beta)\, \sigma}{\theta_1 - \theta_0} \right)^2 \equiv n^*, \text{ say.}$$

SPRT will provide much more economical n.

2.2. Wald's SPRT for General Distribution f

Having fixed α, β we determine $A = \frac{1-\beta}{\alpha}, B = \frac{\beta}{1-\alpha}$. Then, keep observing $X_1, X_2, ...$ one by one in a sequential manner until the likelihood ratio

$$\Lambda_j = \frac{\Pi_{i=1}^{j} f(x_i; \theta_1)}{\Pi_{i=1}^{j} f(x_i; \theta_0)}$$

goes outside the interval (B, A) first time.

Here, "A" indicates "large" and "B" indicates "small".

Then, we determine required number of observations N as:

$N \equiv$ first integer $j(\geq 1)$ such that $\Lambda_j \geq A$ or $\Lambda_j \leq B$;

We will accept H_a if $\Lambda_N \geq A$; accept H_0 if $\Lambda_N \leq B$.

Although $P(N < \infty) = 1$ w.p.1, N ought to be truncated if it asks to go beyond n^* observations in the normal example.

2.3. Optimality

Wald and Wolfowitz (1948) proved the following remarkable optimality property of SPRT among all tests (**including** the **N-P MP** test) with comparable α, β (included in Mukhopadhyay and de Silva 2009, p. 43).

Optimal property of the SPRT: Consider testing $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_1$.Among all tests with type-I error probability $\leq \alpha$, type-II error probability $\leq \beta$, $\alpha + \beta \leq 1$, with $E_\theta[N] < \infty$ when $\theta = \theta_0, \theta_1$, Wald's SPRT with error probabilities α, β has the minimum $E_\theta[N]$ when $\theta = \theta_0, \theta_1$.

2.4. Truncation of the General SPRT

Suppose we can observe at most K observations only. Then we will implement the **truncated SPRT** as follows:

Keep observing $X_1, X_2, ...$ one by one in a sequential manner checking whether the corresponding Λ_j goes outside (B, A) according to sampling stategy. If a decision is made with N smaller than K, then N is te sample size. if no decision is made even when the K^{th} response is observed, then we stop at $N = K$: Decide to accept $H_0(H_a)$ if $\Lambda_K \leq (>)\frac{1}{2}(A + B)$.

In the normal problem, we should pick $K = n^*$ given earlier!

2.5. Back to Normal Distribution Illustration

Illustration Normal with θ and σ^2

$$\Lambda_n = \exp\left\{-(\theta_0 - \theta_1)\sigma^{-2}\sum_{i=1}^{n} X_i - n(\theta_1^2 - \theta_0^2)\sigma^{-2}/2\right\}$$

$$N \equiv \text{first } n(\geq 1) \text{ such that } \sum_{i=1}^{n} X_i \geq c_1 \ln A + nc_2,$$

$$\text{or } \sum_{i=1}^{n} X_i \leq c_1 \ln B + nc_2, \text{with } c_1 = \sigma^2(\theta_1 - \theta_0)^{-1},$$

$$c_2 = (\theta_0 + \theta_1)/2.$$

$$\text{Accept } H_a(H_0) \text{ if } \sum_{i=1}^{n} X_i \geq c_1 \ln A + nc_2 (\leq c_1 \ln B + nc_2).$$

Data Simulations with 10,000 Replication

$\theta_0 = 1, \theta_1 = 2, \alpha = 0.05, \beta = 0.1$

$H_0 : \theta = 1$ versus $H_a : \theta = 2$

$\theta_0 = 1, \theta_1 = 2, \alpha = 0.05, \beta = 0.1$

with $\sigma^2 = 4$, $n^* = 34.27$;

with $\sigma^2 = 16$, $n^* = 137.08$.

Implement **Seq03.exe** from Mukhopadhyay and de Silva (2009)

In what follows, I take liberty in showing some screen shots in order to illustrate how a job runs.

```
Select one of the following programs:

       1   SimAB:  Simulation of alpha + beta for a test
       2   SPRT:   SPRT for given or simulated data
       3   OC:     Computation of OC and ASN functions
       0   Exit

       Enter your choice :2


SPRT for H0:theta = theta0  versus H1:theta = theta1

*************************************************************
*                                                           *
*    ID#  DISTRIBUTION         Theta       Known           *
*                          (Parameter1) Parameter2          *
*                                                           *
*     1   Normal                mean    variance          *
*     2   Gamma                scale       shape          *
*     3   Weibull              scale       shape          *
*     4   Erlang           exp. mean  k(integer)           *
*     5   Exponential           mean                       *
*     6   Poisson               mean                       *
*     7   Bernoulli          P(Success)                    *
*     8   Geometric          P(Success)                    *
*     9   Negative Binomial     mean           k           *
*                                                           *
*************************************************************


Input distribution  ID# : 1

Enter theta0 = 1
      theta1 = 2
      alpha  = 0.05
      beta   = 0.1
```

```
Do you want to change the input parameters (Y or N)? n


   Would you like to truncate the SPRT when
   sample size reaches K (Y or N)? n


   Do you want to simulate data or input real data?
   Type S (Simulation) or D (Real Data) : s


   Do you want to store simulated sample sizes (Y or N)? n

   Number of replications for a simulation? 10000

   Enter a positive integer(<10) to initialize
             the random number sequence : 8

   For the data simulation, input
        mean of the normal distribution: 1
                                variance: 4
```

```
**********************************************************
SPRT:  H0:theta =  1.00   versus   H1:theta =  2.00 at
       alpha =0.05 and beta =0.10  where theta is the
       mean of normal distribution with variance =   4.0


       Results for simulated data from
       Normal distribution with mean =  1.00
       -------------------------------------


       Number of Simulations          =  10000
       Number of times H0 is accepted =   9611

       Average sample size (n bar)    =  18.75
       Std. dev. of n (s)             =  14.18
       Minimum sample size (min n)    =      2
       Maximum sample size (max n)    =    130
       ****************************************
```

Table 1. Simulated results with 10,000 replications for untruncated SPRT of $H_0 : \theta = 1$ vs. $H_a : \theta = 2$ for $N(\theta, \sigma^2)$ with $\sigma^2 = (4, 16)$ known, preassigned $\alpha = 0.05$, $\beta = 0.1$, true $\theta = (1, 1.5, 2)$.

σ^2	true θ	n^*	$\overline{n}$	$s_{\overline{n}}$	$p_{H_0}\%$	Min	Max	Q_1	Q_2	Q_3
4	1	34.27	18.75	0.142	96.11	2	130	9	15	24
	1.5	34.27	32.91	0.267	55.89	2	297	14	25	44
	2	34.27	22.43	0.154	7.30	2	147	12	18	29
16	1	137.08	69.52	0.549	95.66	6	571	32	53	90
	1.5	137.08	116.25	0.957	55.75	7	825	49	87	153
	2	137.08	81.86	0.568	9.04	7	526	41	66	106

Table 2. Simulated results with 10,000 replications for truncated SPRT test of $H_0 : \theta = 1$ vs. $H_a : \theta = 2$ for $N(\theta, \sigma^2)$ with $\sigma^2 = (4, 16)$ known, preassigned $\alpha = 0.05$,$\beta = 0.1$, true $\theta = (1, 1.5, 2)$. Truncation at K.

σ^2	true θ	K	$\overline{n}$	$s_{\overline{n}}$	Min	Max	Q_1	Q_2	Q_3	p_{H_0} %	n_T	$p_{H_0}\_T$ %
4	1	35	17.15	0.099	2	35	9	15	24	93.79	1146	72.69
	1.5	35	23.72	0.107	2	35	14	25	35	55.77	3368	49.73
	2	35	20.00	0.098	2	35	12	18	29	10.84	1522	25.95
16	1	138	63.78	0.393	6	138	31	53	91	93.51	1017	71.98
	1.5	138	88.05	0.427	6	138	49	87	138	56.17	2916	51.13
	2	138	74.25	0.385	8	138	42	66	105	11.71	1377	26.94

The **Operations Characteristic** (OC) function:

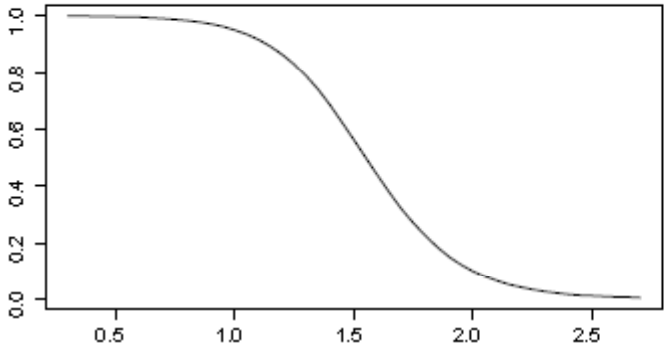
$L(\theta) =$ Probability of accepting H_0 when θ is true,

$L(\theta_0) = 1 - \alpha,$ and $L(\theta_1) = \beta.$

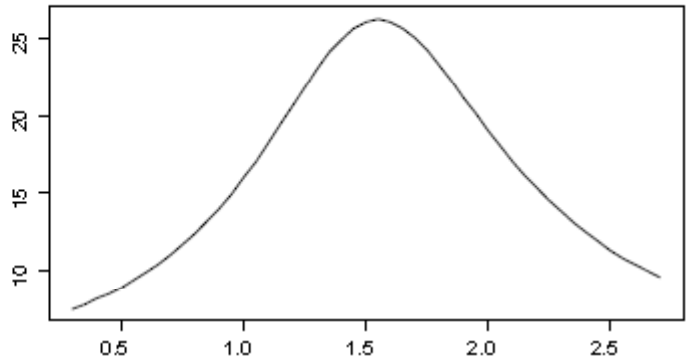
The **Average Sample Number** (ASN) function:

$E_\theta[N]$ when θ is true

Next, we show some plots for the OC function $[(\theta, L(\theta))]$ and ASN function $[(\theta, E_\theta[N])]$:



a: OC function



b: ASN function

Fig. 1: $H_0 : \theta = 1$ vs. $H_a : \theta = 2$ for $N(\theta, 16), \alpha = 0.05, \beta = 0.1$

Real Data Illustration:

"Bodyfat" dataset from Statlib (lib.stat.cmu.edu/datasets/), submitted by Johnson (1996). Data include percentage of body fat determined by underwater weighing and various body circumference measurements for 252 men.

We llooked at response variable (X): "**Density determined from underwater weighing**".

Some descriptive statistics of density_underwater are shown:

Min	Q_1	Median	Mean	Q_3	Max	s.d.
0.995	1.041	1.055	1.056	1.070	1.109	0.019

We looked at histogram and Q-Q plot and boxplot. These looked rather normal distribution. Shapiro-Wilk normality test gave p–value = 0.6571.

We treated this dataset (X) as our population: $N(\theta, 0.019^2)$

We set out to test: $H_0 : \theta = 1.055$ vs $H_a : \theta = 1.060$ with $\alpha = 0.05$, $\beta = 0.1$, and $K = 124 (= n^*)$.
The SPRT terminated with $N = 24$ observations, and our decision was to accept H_0 upon termination.

How about SPRT's in non-normal distributions?

In an extended version of the draft paper (with Yan Zhuang, expected Ph.D. August 2018), we have illustrations with simulations real datasets from

- Gamma (unknown scale, known shape): Body dimensions (Heinz et al., 2003),
- Lognormal (location unknown, scale known): Rosner's FEV (maximum forced expiratory volume in one second) data.

2.6. Remarks: Sequential Tests Under Unknown Nuisance Parameters and Other Situations:

Test on normal with unknown variance: Use program Seq04.exe (Mukhopadhyay and de Silva, 2009): Sequential *t*-test or Chi-square test or *F*-test …

Discrete distributions (Binomial, Poisson, Negative binomial, Zero-inflated)

Unequal sampling at stages, perhaps according to some distribution (Mukhopadhyay and de Silva, 2005)

Two-sample or multi-sample tests

Scenarios where the response $\boldsymbol{X}$ may be multivariate …

3. Fixed-Width Confidence Interval

3.1. Normal Mean Example

$X_1, ..., X_n$ are independent, distributed as $N(\theta, \sigma^2)$, where θ, σ^2 are both unknown. Usual $(1-\alpha)$ confidence interval for θ is:

$$I_n = \left[\overline{X}_n \pm t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}\right],$$

where $t_{n-1,\alpha/2}$ is the upper $50\alpha\%$ point of the t-distribution with $(n-1)$ degrees of freedom, $\overline{X}_n$ is the sample mean, and S_n is the sample standard deviation.

What should be n?

Length of the confidence interval I_n is $K_n = 2t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}$, which is random that might be (very) large for fixed n even if unknown σ is "small". That is, I_n may be too wide for practical use.

Therefore, it may be necessary to construct a confidence interval J_n for θ which satisfies two requirements:

1) $P\{\theta \in J_n\} \geq 1 - \alpha$;

2) Length of $J_n \leq 2d, d(> 0)$ is preassigned half-width.

This J_n is called $(1 - \alpha)$ fixed-width confidence interval for θ. Such J_n can not be constructed by any fixed-sample-size procedure (Dantzig 1940) if n is predetermined and fixed.

The fixed-width confidence interval $J_n = [\overline{X}_n \pm d]$ for θ **also** has **preset** $(1-\alpha)$ **confidence** if the required sample size n is determined as follows:

$$n \geq z^2_{\alpha/2}\sigma^2/d^2 \equiv n^*, \text{ say.}$$

This n^* is called the **optimal fixed sample size** had σ^2 been known. But, magnitude of n^* remains unknown!

Some adaptive estimation methods for n^* are provided next.

3.1.1. Two-Stage Sampling

Mahalanobis (1940), Stein (1945,1949)

Begin with $m(\geq 2)$ **pilot** observations $X_1, ..., X_m$ and obtain the sample variance S_m^2 from $\{X_1, X_2, ..., X_m\}$ to estimate σ^2. Recall that n^* was $z_{\alpha/2}^2 \sigma^2/d^2$. Let the **final sample size**:

$$N = \max\left\{m, \left\langle t_{m-1,\alpha/2}^2 S_m^2/d^2 \right\rangle + 1\right\},$$

where $\langle u \rangle$ = the largest integer $< u$.

Example 1: Suppose $m = 20$, and $t_{m-1,\alpha/2}^2 S_m^2/d^2 = 12.82$.

Then, $N = \max\{20, < 12.82 > +1\} = 20 \Rightarrow N = 20$.

Example 2: Suppose $m = 20$, and $t_{m-1,\alpha/2}^2 S_m^2/d^2 = 63.27$.

Then, $N = \max\{20, < 63.27 > +1\} = 64 \Rightarrow N = 64$.

Generally speaking, we proceed as follows:

a) if $N = m$, then we already have enough data in pilot stage.
In this case, final dataset is $\{X_1, X_2, ..., X_m\}$;

b) if $N > m$, then we have too few data in pilot stage. Then,
we gather $(N - m)$ new observations in the second stage.
In this case, final dataset is $\{X_1, ..., X_m, X_{m+1}, ..., X_N\}$.

Final fixed-width confidence interval for θ: $J_N = [\overline{X}_N \pm d]$.
where $\overline{X}_N$ is the sample mean from final dataset $\{X_1, ..., X_N\}$.

Big Result (Consistency or Exact Consistency): For fixed θ, σ, α and d, one has: $P_{\theta,\sigma}\{\theta \in J_N = [\overline{X}_N \pm d]\} \geq 1 - \alpha$.

This fundamental construction of adaptive sampling in two-steps **changed everything** after Wald's construction of the SPRT.

Simulations were carried out using the program Seq06.exe from Mukhopadhyay and de Silva (2009)

Table 3a. Simulations of two-stage estimation strategy with $10,000$ replications with $\alpha = 0.1, m = 10$

α	d	n^*	$\overline{n}$	$s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{p}$	$s_{\overline{p}}$
0.1	1.500	19.239	24.314	0.110	1.264	0.903	0.003
	1.000	43.289	54.272	0.252	1.254	0.906	0.003
	0.800	67.639	83.978	0.392	1.242	0.902	0.003
	0.600	120.246	149.965	0.703	1.247	0.904	0.003
	0.400	270.554	336.704	1.580	1.244	0.894	0.003
	0.200	1082.217	1346.401	6.318	1.244	0.898	0.003

Table 3b. Simulations of two-stage estimation strategy with $10,000$ replications with $\alpha = 0.05, m = 10$

α	d	n^*	$\overline{n}$	$s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{p}$	$s_{\overline{p}}$
0.05	1.500	27.317	37.001	0.171	1.354	0.952	0.002
	1.000	61.463	81.718	0.382	1.330	0.949	0.002
	0.800	96.036	128.128	0.592	1.334	0.950	0.002
	0.600	170.732	229.777	1.093	1.346	0.950	0.002
	0.400	384.146	517.077	2.466	1.346	0.950	0.002
	0.200	1536.584	2051.753	9.630	1.335	0.952	0.002

Table 3c. Simulations results of two-stage estimation strategy with $10,000$ replications with $\alpha = 0.01, m = 10$

α	d	n^*	$\overline{n}$	$s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{p}$	$s_{\overline{p}}$
0.01	1.500	47.181	75.382	0.357	1.598	0.991	0.001
	1.000	106.158	169.157	0.795	1.593	0.989	0.001
	0.800	165.872	264.773	1.230	1.596	0.992	0.001
	0.600	294.884	470.074	2.222	1.594	0.990	0.001
	0.400	663.490	1062.446	5.022	1.601	0.990	0.001
	0.200	2653.959	4258.497	20.121	1.605	0.991	0.001

Stein's two-stage solution, though path-breaking, also came with some baggage of its own:

a) **Oversampling**: $E[N/n^*] > t^2_{m-1,\alpha/2}/z^2_{\alpha/2} (> 1)$.
Not too disturbing!

b) **Asymptotic Oversampling**: $E[N/n^*] \to t^2_{m-1,\alpha/2}/z^2_{\alpha/2}$ (> 1) as $d \to 0$. This can be disturbing!

Note that σ^2 was estimated only once through pilot data. Appropriate modifications were proposed and implemented by Mukhopadhyay (1980) and Mukhopadhyay and Duggan (1997) in order to gain **substantial efficiency** by drastically **reducing the extent of oversampling** via right choices of m under two-stage adaptive sampling methods.

Why not estimate σ^2 **sequentially, multiple times**?

Table 4. Asymptotic Oversampling Percentage $t^2_{m-1,\alpha/2}/z^2_{\alpha/2}$ values for $\alpha = 0.1, 0.05, 0.01$, $m = (10, 20)$ in the Two-Stage Procedure

	$m = 10$	$m = 20$
α	oversampling%	oversampling%
0.10	1.2420 (24.20%)	1.1051(10.51%)
0.05	1.3321 (33.21%)	1.1404(14.04%)
0.01	1.5918 (59.18%)	1.2336(23.36%)

Real Data Illustration:

We used the infection risk of hospital infection data from Kutner et al. (2005) on 113 hospitals in the US, 1975-76. We treated the 113 observed values of infection risk as population data. Normality checks went smoothly.

Table 5. Single run to estimate the mean of infection risk with $\alpha = 0.05$, $m = 10$ for two stage procedure

d	n^*	n	$\overline{X}_n$	S_n^2	95% CI
0.6	19.02	25	4.26	1.57	$(3.66, 4.86)$
0.5	27.38	35	4.31	1.97	$(3.81, 4.81)$
0.4	42.79	55	4.40	1.97	$(4.00, 4.80)$

True mean infection risk 4.355 lies within each conf interval.

3.1.2. Purely Sequential Sampling

Anscombe (1952), Ray (1957), and Chow and Robbins (1965)

Begin with $m(\geq 2)$ **pilot** observations $X_1, ..., X_m$ and obtain one additional observation at-a-time giving rise to $\left(n, \overline{X}_n, S_n^2\right)$, successively for $n = m, m+1, ...$ as needed.

Recall that n^* was $z_{\alpha/2}^2 \sigma^2 / d^2$. Let the **final sample size:**

$$N = \min\left\{n \geq m : n \geq z_{\alpha/2}^2 S_n^2 / d^2\right\}.$$

In this case, final dataset is $\{X_1, ..., X_m, ..., X_N\}$.

Final fixed-width confidence interval for θ: $J_N = [\overline{X}_N \pm d]$.

with $\overline{X}_N$, the sample mean from final dataset $\{X_1, ..., X_N\}$.

Properties:

$P_{\theta,\sigma}(N < \infty) = 1.$

Asymptotic Consistency: $\lim_{d \to 0} P_{\theta,\sigma}\{\theta \in J_N\} = 1 - \alpha.$

Asymptotic Efficiency: $\lim_{d \to 0} E_{\theta,\sigma}[N/n^*] = 1.$

This methodology is also **asymptotically second-order efficient** (Ghosh and Mukhopadhyay 1981).

Next, we show some simulated results.

Table 6. Simulated results of purely sequential estimation strategy with $10,000$ replications, $\alpha = 0.05, m = 10$

α	d	n^*	$\overline{n}$	$s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{p}$	$s_{\overline{p}}$
0.05	1.500	27.317	25.423	0.082	0.931	0.923	0.003
	1.000	61.463	59.531	0.124	0.969	0.941	0.002
	0.800	96.036	94.469	0.147	0.984	0.943	0.002
	0.600	170.732	169.660	0.190	0.994	0.948	0.002
	0.400	384.146	383.407	0.276	0.998	0.948	0.002
	0.200	1536.584	1536.069	0.558	1.000	0.952	0.002

Table 7. Simulated results of purely sequential estimation strategy with $10,000$ replications, $\alpha = 0.05, m = 20$

α	d	n^*	$\overline{n}$	$s_{\overline{n}}$	$\overline{n}/n^*$	$\overline{p}$	$s_{\overline{p}}$
0.05	1.500	27.317	27.104	0.062	0.992	0.942	0.002
	1.000	61.463	59.686	0.120	0.971	0.941	0.002
	0.800	96.036	94.553	0.145	0.985	0.943	0.002
	0.600	170.732	169.677	0.189	0.994	0.948	0.002
	0.400	384.146	383.364	0.276	0.998	0.948	0.002
	0.200	1536.584	1536.060	0.557	1.000	0.949	0.002

Real Data Illustration:

Table 8. Single run to estimate the mean infection risk with $\alpha = 0.05, m = 10$ for purely sequential procedure

d	n^*	N	$\overline{\overline{X}}_N$	S_n^2	95% CI
0.6	19.02	20	4.20	1.87	(3.60, 4.80)
0.5	27.38	24	4.20	1.55	(3.70, 4.70)
0.4	42.79	48	4.46	1.94	(4.06, 4.86)

True mean infection risk 4.355 lies within each confidence interval.

3.2. Remarks: Sequential Fixed-Size Confidence Regions and Other Situations:

Discrete distributions (Binomial, Poisson, Negative binomial, Zero-inflated): Based on the MLE

Many other continuous distributions (Negative exponential, Exponential, Gamma, Lognormal …): Based on the MLE

Multivariate and regression simultaneous confidence and point estimation problems (under other loss functions)

Unequal variances (treatment allocations)

Multiple comparisons …

Thank You

Any Question or Comment?