SEQUENTIAL STOPPING RULE FOR DETERMINING THE NUMBER OF REPLICATIONS WHEN SEVERAL MEASURES OF EFFECTIVENESS ARE OF INTEREST

Anthony J. Quinzi

October 2004

BACKGROUND

- **How many replications of a scenario is enough to estimate a mean performance parameter with a specified degree of accuracy and level of confidence?**
- **For single measure, use the following fact:**

If $X_1, \ldots, X_n$ is a random sample of size n from a normal population with mean μ, then

$$P\left(-t_{\alpha/2,n-1} \times \frac{s}{\sqrt{n}} \leq \left| \overline{X} - \mu \right| \leq t_{\alpha/2,n-1} \times \frac{s}{\sqrt{n}} \right) = 1 - \alpha$$

SEQUENTIAL STOPPING RULE

- Usual $100 \times (1-\alpha)\%$ confidence interval for μ

$$1 - \alpha \geq P(\left|\overline{X} - \mu\right| \leq \text{half - length}) \Rightarrow$$

$$1 - \alpha \geq P\left(\left|\frac{\overline{X} - \mu}{\mu}\right| \leq \frac{\gamma}{1 - \gamma}\right)$$

$$\left|\overline{X} - \mu\right| = \text{ absolute error}$$

$$\gamma = \left|\frac{\overline{X} - \mu}{\mu}\right| = \text{ relative error}$$

$$\gamma' = \frac{\gamma}{1 - \gamma} = \text{ adjusted relative error to achieve}$$

$$\text{a relative error of } \gamma$$

SEQUENTIAL STOPPING RULE

Problem: How many replications are sufficient to achieve a given precision (γ) with confidence $100 \times (1 - \alpha)\%$?

- **Law and Kelton (1982) suggest a sequential stopping rule for estimation of the mean μ**
- ***Step 1. Make n_0 replications of the simulation and set $n = n_0$.***
- ***Step 2. Compute $\overline{X}$ and the quantity***

$$\delta(n,\alpha) = t_{n-1,1-\alpha/2}\sqrt{\frac{s^2}{n}} \quad \text{where} \quad s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}$$

and the level of confidence is $100\times(1-\alpha)$ %, $0 < \alpha < 1$.

- ***Step 3. If*** $\delta(n,\alpha)/\left|\overline{X}\right| \leq \gamma'$, **where** $\gamma' = \gamma/(1+\gamma)$

use $\overline{X}$ as the point estimate and stop. Otherwise, Replace n by (n + 1), make one additional replication and go to step 2 to repeat the process.

SEQUENTIAL STOPPING RULE

THE NEED FOR MULTIPLE MOE

- **A study of key performance parameters for the Army's Future Combat System of Systems was conducted.**
- **Following question arose: How many replications are necessary to estimate a number of mean performance parameters simultaneously?**
- **Question was important in that the Caspian scenario being used required 60 hours of real time for each replication. Post processing was a huge effort and to meet the study deadline, analysts used the results from only n = 11 replications.**

SEQUENTIAL STOPPING RULE

THE CASE OF MULTIPLE MOE

- **Suppose $\mu_1, \ldots, \mu_k$ represent the means of *k* MOE.**
- **For each mean μ_s , form a $100 \times (1 - \alpha_s)\%$, where**

$$\sum_{s=1}^{k} \alpha_s = \alpha$$

- **Then the Bonferroni bound provides a lower bound for the joint probability that each of the *k* confidence intervals captures its respective mean.**

SEQUENTIAL STOPPING RULE

STOPPING RULE FOR MULTIPLE MEASURES

- **Perform the 3-step procedure outlined above for *each* measure**
- **Stop when *every* inequality in Step 3 of the above procedure holds.**

SEQUENTIAL STOPPING RULE

APPLICATION

- **A scaled-down version of the Caspian scenario (6 hours per replication) was used to test the 3-step procedure for multiple measures.**
- **MOE of interest were:**

  1) Friendly system losses

  2) Friendly individual soldier losses

  3) Threat system losses

  4) Threat individual soldier losses
- **Initial n_0 = 7 replications were run.**
- **Results of n = 30 replications were used.**

SEQUENTIAL STOPPING RULE

		Reference: Law & Kelton Ed. 3, pp. 513-514							
REP					Blue Losses Vehicles		Blue Losses Dismounts		Red Losses Vehicles
1					35		32		75
2					43		29		83
3					41		32		81
4					32		43		82
5					47		24		75
6					38		35		81
7					34		36		83
				delta(n, a)		delta(n, a)		delta(n, a)	
	d1	2.969	80% (Each .05)	6.04	CONTINUE	6.67	CONTINUE	3.94	STOP
	d2	3.707	92% (Each .02)	7.54	CONTINUE	8.33	CONTINUE	4.92	STOP
	d3	4.317	96% (Each .01)	8.78	CONTINUE	9.70	CONTINUE	5.73	STOP
	d4	5.959	99.2% (Each .002)	12.12	CONTINUE	13.39	CONTINUE	7.91	CONTINUE

SEQUENTIAL STOPPING RULE

RESULTS

In 30 replications, stopping rule not satisfied for measure (2), the mean number of friendly individual soldier losses. Variability due primarily to large increases in losses when a squad of soldiers was mounted and the platform received a catastrophic kill.

QUESTIONS FOR THE PANEL

1. Given that we are dealing with purely discrete distributions (numbers of losses) each of whose underlying distributions results from a large number of random draws in the model, should we really be considering a procedure based on the *t*-distribution which assumes population is Gaussian?

QUESTIONS FOR THE PANEL

2. Should we instead be trying to estimate the underlying discrete probability mass functions [cf. Chiu, S.T. (1991) "Bandwidth selection for kernel density estimation", *Ann. Stat. Vol. 19*, pp. 1883-1905] or use some other methodology so that we might be able to improve on the Bonferroni bounds?

QUESTIONS FOR THE PANEL

3. Are bootstrap methods really appropriate in a PURELY discrete context such as this?

4. Is this question crying for some sort of Bayesian approach?