'95춘계학술발표회 논문집
한국원자력학회

IAEA 사찰표본할당계산법에 대한 초기하분포 적용 연구

김현태 · 박완수 · 민경식 · 박찬식 · 곽은호
한국원자력연구소

요 약

IAEA(International Atomic Energy Agency, 국제원자력기구)에서는 사찰활동 수행시, 비복원추출(sampling without replacement)을 기술하는 초기하분포 대신 복원추출(sampling with replacement)을 기술하는 이항분포를 사용하여 표본크기를 계산하여 사찰방법들에 할당한다. 본 연구에서는 이항근사법이 사용되는 IAEA의 표본크기 할당계산결과와 이항근사법 대신 초기하분포를 적용한 IAEA 표본크기 할당계산결과를 비교검토하였다.

1. 서론

IAEA 사찰활동은 공산품 생산에 적용되는 품질관리 활동과는 달리 게임(game) 이론이 적용되는 상황이기 때문에 이항분포를 초기하분포의 근사(approximation)로 사용할 수 있는 범위를 넘는 경우를 대비하여 표본추출계획을 수립한다[1][2]. IAEA에서는 이항근사법의 사용에 따르는 제한점을 극복하기 위한 노력을 지속적으로 추진하고 있다[3]. 최근에는 기존의 이항근사법 보다 표본크기를 보다 정확하게 계산하는 개선된 이항근사법이 발표되었다[4]. 그러나 개선된 이항분포에 의한 표본크기 계산도 근사조건을 벗어나는 경우에는 기존의 이항근사법과 마찬가지로 초기하분포를 사용한 계산결과와 차이를 보인다[8]. IAEA에서는 사찰활동시에 두가지 또는 세가지 검증방법을 동일한 대상에 대해 적용하여, 예상되는 다양한 전용책략(diversion strategy)에 대비하고 있기 때문에 통계적으로 정확한 계산에 근거를 둔 사찰계획을 수립할 필요가 있다.

2. 사찰표본추출계획

표본추출검사는 검사 대상의 모든 요소를 하나 하나 검사하는 것이 아니라 모집단의 부분집합인 표본을 검사하여 모집단의 품질에 대한 합격/불합격 여부를 결정하는 것이다. IAEA의 사찰은 사찰대상에 있으리라고 예상되는 결함(defect)을 발견하려는 활동이다. 모집단의 요소 중에 원래 선언된 것이 아닌 것, 즉 결함이 있을 때 이를 빠른 시간 내에 탐지(timely detection)하자는 것이다. 여기서 결함이라는 것은 모집단에 있는 요소 중에 원래의 요소가 일부 손상되었거나, 다른 것으로 대치되었거나, 또는 없어진 것을 말한다. 적용되는 가설검정은 다음과 같다.

H_0: 모집단에서 전용이 발생하지 않았다.
H_1: 모집단에서 1 SQ의 전용이 발생했다.

1 SQ(Significant Quantity, 有意量)는 핵무기 1개를 제조할 수 있는 핵물질 量이다. IAEA의 핵사찰에서 분류되는 결함의 종류는 다음과 같다.

① 전체결함(gross defect) : 요소 전체가 결함
② 부분결함(partial defect): 요소의 일부가 결함
③ 경미결함(bias defect) : 요소의 극히 일부가 결함

모집단에 결함이 없다면 표본에 결함이 없을 것이므로 사찰관이 정상인 요소를 정상이라고 판단할 확률이 1 이라면 제 1 종 오류 확률 $\alpha = 0$ 이다. 모집단에 결함이 포함되어 있음에도 불구하고 이를 탐지하지 못할 확률인 제 2 종 오류 확률 β를 IAEA에서는 미탐지확률(non-detection probability)이라고 부른다. 따라서 합격판정계수 c = 0 인 표본추출계획을 수립하는 것이다.

3. 초기하분포와 이항분포

N을 모집단의 크기, D를 1 SQ에 해당되는 예상 결함수, n을 표본 크기라 할 때 표본에 d 개의 결함이 포함될 확률을 나타내는 초기하분포는 다음과 같다.

$$h(N,D,n,d) = \binom{D}{d}\binom{N-D}{n-d} \Big/ \binom{N}{n} \tag{1}$$

$N > 50$, $p = D/N \le 0.10$, $n \le D$ 일 때에는

$$h(N,D,n,d) \approx b(n,p,d) = \binom{n}{d} p^d (1-p)^{n-d} \tag{2}$$

$N > 50$, $f = n/N \le 0.10$, $D \le n$ 일 때에는

$$h(N,D,n,d) \approx b(D,p,d) = \binom{D}{d} f^d (1-f)^{D-d} \tag{3}$$

의 이항분포를 일반적으로 사용한다[5].

4. 미탐지확률과 표본크기

표본추출에 따른 오류(sampling error) 확률 β는 식 (4)와 같다. 결함인 요소를 결함

$$\beta = h(N,D,n,0) = \frac{(N-D)(N-D-1)\cdots(N-D-n+1)}{N(N-1)(N-2)\cdots(N-n+1)} \tag{4}$$

이라고 판단할 사찰관의 결함판별력이 완벽하다고 가정하면 식 (4)는 미탐지확률이 된다.

N과 D가 주어졌을 때, $\beta \ge h(N,D,n,0)$를 만족하는 n 중에서 최소값이 초기하분포에 의한 표본크기이다[6][7]. 식 (1), (4) 계산에 C언어의 long double이라는 데이타 형태를 사용했다[6]. Long double에서 표현가능한 부동소수의 범위는 $10^{\pm 4932}$ 이다.

$D \le n$ 일 경우 미탐지확률과 표본크기는 다음과 같다.

$$\beta \approx \left(1 - \frac{n}{N}\right)^D \tag{5}$$

$$n \approx N\left(1 - \beta^{\frac{1}{D}}\right) \tag{6}$$

식 (6)은 전체결함을 탐지하고자 할 때 적용되는 공식이다.

$n \le D$ 일 경우 이항근사법에 의한 미탐지확률과 표본크기는 다음과 같다.

$$\beta \approx (1 - \frac{D}{N})^n \tag{7}$$

$$n \approx \frac{\ln \beta}{\ln\left(1 - \frac{D}{N}\right)} \tag{8}$$

부분결함이나 경미결함에 의해 1 SQ가 전용된다면 D 가 커지게 되므로 표본크기 n 이 상대적으로 작아진다. 이 경우에는 $n \le D$가 되어 식 (8)을 적용할 수 있다.

5. 표본할당

부분결함을 탐지할 수 있는 탐지방법은 전체결함을 탐지할 수 있고, 경미결함을 탐지할 수 있는 탐지방법은 부분결함이나 전체결함을 탐지할 수 있다. 전체결함에 의해 1 SQ를 전용할 수 있는 요소의 수가 부분결함에 의해 1 SQ를 전용할 수 있는 요소의 수보다 적으며 부분결함에 의해 1 SQ를 전용할 수 있는 요소의 수가 경미결함에 의해 1 SQ를 전용할 수 있는 요소의 수보다 적기 때문에 다음과 같은 할당방식은 전용을 탐지하는데 매우 효율적이다[3]. 주어진 미탐지확률에 대해

① 식 (6)으로 전체결함을 탐지하기위한 표본크기 n을 구한다.
② 식 (8) 형태를 사용하여 부분결함을 탐지하기위한 표본크기 η_2의 최소값을 구한다.
③ 식 (8) 형태를 사용하여 경미결함을 탐지하기위한 표본크기 η_3의 최소값을 구한다.

①, ②, ③에서 구한 n, η_2, η_3에서 다음과 같은 방식으로 첫번째 표본할당 계산을 한다.

④ 전체결함을 탐지하는 방법에 할당되는 표본크기: bn1 = n - η_2
⑤ 부분결함을 탐지하는 방법에 할당되는 표본크기: bn2 = η_2 - η_3
⑥ 경미결함을 탐지하는 방법에 할당되는 표본크기: bn3 = η_3

전체결함을 탐지하는 방법을 적용하여 ④에서 구한 표본크기 bn1에 대해 미탐지확률 B1을 계산한다. B1 상황에서 부분결함을 탐지하는 방법을 적용하여 ⑤에서 구한 표본크기 bn2에 대해 미탐지확률 B2를 계산한다. B2 상황에서 경미결함을 탐지하는 방법을 적용하여 ⑥에서 구한 표본크기 bn3에 대해 미탐지확률 B3를 계산한다. 최초의 표본할당에 대한 미탐지확률은 Q = B1*B2*B3 이다. $\beta \ge$ Q 이면, 最適解는 ④, ⑤, ⑥에서 구한 bn1, bn2, 그리고 bn3이다. $\beta <$ Q 이면, 경미결함탐지방법이나 부분결함탐지방법의 표본크기를 증가시키어서 즉 전체결함탐지방법이나 부분결함탐지방법의 표본크기를 줄여서 B1, B2, B3, 그리고 Q = B1*B2*B3를 계산한 후 Q와 β를 비교하여 最適解여부에 따라 반복계산 여부를 결정한다.

본 연구에서는 근사식 (6), (8)을 사용한 결과와 식 (4)를 사용한 결과를 비교하였다.

6. 초기하분포와 이항근사법에 의한 표본할당크기 비교

입력 자료는 IAEA의 Safeguards Criteria를 고려하였다[2]. IAEA에서는 1 - β 를 탐지확률이라고 하며 이는 검정력이다. 격납/감시 (C/S, Containment and Surveillance), 봉인(seal) 정도에 따라 탐지확률을 다음과 같이 세 가지 경우로 구분하여 적용한다.

① 高探知確率 (high detection probability) = 90%
② 中探知確率 (medium detection probability) = 50%
③ 低探知確率 (low detection probability) = 20%

또한 플루토니움 탐지에 적용되는 탐지확률 95%를 계산에 사용했다. 전용에는 실제의 量보다 적다고 하는 방법(under-statement)과 실제의 量보다 많다고 하는 방법(over-statement)이 있다. 「표.1」은 over-statement 경우이다. 사용된 부호에 대한 설명은 다음과 같다.

미탐지확률 .... β
유의량 .... SQ
단위용기에 있는 핵물질 평균量 x
모집단에 있는 단위용기 수 .... N
i 번째 사찰방법의 상대표준편차 .... d*i*, *i*=1,2,3
이항근사법으로 계산하여 *i* 번째 사찰방법에 할당하는 표본 크기 .... bn*i*, *i*=1,2,3
초기하분포로 계산하여 *i* 번째 사찰방법에 할당하는 표본 크기 .... hn*i*, *i*=1,2,3

「표.1」에서 보여주는 바와 같이 이항근사법에 의한 계산결과는 초기하분포에 의한 계산결과와 차이를 보인다. 「표.1」의 첫번째 줄은 다음과 같다.

							이항근사계산			초기하계산		
β	SQ	x	N	d1	d2	d3	bn1	bn2	bn3	hn1	hn2	hn3
0.05	8.0	0.4	25	0.30	0.12	0.08	3	0	1	1	0	1

SQ와 x에서 예상 전체결함수는 SQ/x = 8.0/0.4 = 20 이다. 즉, 모집단에 있는 요소의 수가 25인데 그중 20개가 결함이라고 할 때, 미탐지확률이 5%를 만족하는 표본크기를 구하는 문제이다. 식 (6)을 적용하여 표본크기 n을 계산하면,

$$n = 25(\ 1 - 0.05^{1/20}\ ) \cong 25(\ 1- 0.861\ ) = 3.48 \cong\ 4$$

이나, 식 (4)를 적용하면

$$h(25,20,3,0) \cong 0.004, \quad h(25,20,2,0) \cong 0.033, \quad h(25,20,1,0) \cong 0.2 \text{ 이므로}$$

통계적으로 정확한 표본크기는 2 이다. 따라서 이항근사법을 사용하면 표본전체크기에서 over-sampling을 하게된다.

「표.1」의 두번째 줄은 다음과 같다.

							이항근사계산			초기하계산		
β	SQ	x	N	d1	d2	d3	bn1	bn2	bn3	hn1	hn2	hn3
0.05	8.0	0.4	50	0.30	0.12	0.08	2	3	2	1	4	1

모집단에 있는 요소의 수가 50인데 그중 20개가 결함이라고 할 때, 미탐지확률이 5%를 만족하는 표본크기를 구하는 문제이다. 식 (6)을 적용하여 표본크기 n을 계산하면,

$$n = 50(\ 1 - 0.05^{1/20}\ ) \cong 50(\ 1- 0.861\ ) = 6.95 \cong\ 7$$

이나, 식 (4)를 적용하면

h(50,20,7,0) ≅ 0.02, h(50,20,6,0) ≅ 0.04, h(50,20,5,0) ≅ 0.07 이므로

통계적으로 정확한 표본크기는 6 이다. 마찬가지로 이항근사법을 사용하면 표본전체크기에서 over-sampling을 하게된다.

전반적으로 이항분포에 의한 계산은 초기하분포에 의한 계산보다 전체표본크기에서 한개 내지 두개가 크다. 이항분포에 의한 계산은 주어진 미탐지확률 보다 더 낮은 미탐지확률에 대한 표본크기를 계산한다.

모집단의 크기가 크고 미탐지확률을 10% 이하로 할 때에는 아래의 경우와 같이 표본할당에서 상당한 차이를 보인다. 전체표본크기는 이항분포에 의한 계산이 각각 하나씩 더 크나 검증의 엄격한 면은 초기하분포에 의한 계산이 더 크므로 이항분포에 의한 계산은 under-sampling을 한다고 볼 수 있다.

							이항근사계산			초기하계산		
β	SQ	x	N	d1	d2	d3	bn1	bn2	bn3	hn1	hn2	hn3
0.05	8.0	0.4	500	0.30	0.12	0.08	12	29	29	11	25	33
0.10	8.0	0.4	500	0.30	0.12	0.08	11	22	22	10	19	25

위의 결과에서 이항근사법을 사용하면 over-sampling을 하게 되거나 under-sampling을 하게 되어 통계적 정확성에 대한 논쟁의 여지가 항상있게 된다.

7. 결 론

예전에는 대형컴퓨터에서 가능했던 계산을 PC에서 수행할 수 있게 됨에 따라 사찰계획에 관련되는 초기하분포계산을 수행있다. IAEA의 사찰활동에서는 초기하분포를 이항분포로 근사하는 조건이 충족된다고 볼 수 없기 때문에 통계적 정확성에 대한 논쟁의 여지가 항상있다. 따라서 보장조치 기술수준은 높이고 국가사찰을 정착시키기 위해서는 통계적 정확성에 대한 논쟁의 여지가 있는 이항근사법을 사용하는 계산법 대신 초기하분포를 사용하는 계산법을 개발할 필요가 있다.

참 고 문 헌

1. IAEA, "IAEA Safeguards: Statistical Concepts and Techniques," 4th rev. ed., IAEA document IAEA/SG/SCT/4 (1989).
2. IAEA, "Safeguards Criteria 1991-1995," IAEA Report 1990-11-21 (1990).
3. J. L. Jaech and M. Russell, "Algorithms to Calculate Sample Sizes for Inspection Sampling Plans," IAEA STR-261, Rev. 1 (1991).
4. J. L. Jaech, "An improved binomial approximation to the hypergeometric density function," *Journal of Nuclear Material Management*, 36-41 (1994).
5. V. K. Rohatgi, *Statistical Inference*, pp. 341-342, John Wiley & Sons, New York (1984).
6. 한국원자력연구소(서인석,노성기,안종성,김현태), "초기하분포의 확률함수 계산", 한국정보산업연합회 프로그램 등록번호: 94-01-12-3470 (1994).
7. 한국원자력연구소(서인석,곽은호,박완수,김현태), "초기하분포와 이항분포에 의한 표본크기, 오류 계산", 한국정보산업연합회 프로그램 등록번호: 94-01-12-3469 (1994).
8. 김현태, 명건식, "표본추출검사에 있어서 초기하분포와 이항근사법을 이용한 표본크기의 결정", '95.2 대한방사선방어학회에 제출

표.1

β	SQ	x	N	d1	d2	d3	이항근사계산 bn1	bn2	bn3	초기하계산 hn1	hn2	hn3
0.05	8.0	0.4	25	0.30	0.12	0.08	3	0	1	1	0	1
0.05	8.0	0.4	50	0.30	0.12	0.08	2	3	2	1	4	1
0.05	8.0	0.4	100	0.30	0.12	0.08	3	6	5	2	5	6
0.05	8.0	0.4	200	0.30	0.12	0.08	5	12	11	4	10	13
0.05	8.0	0.4	500	0.30	0.12	0.08	12	29	29	11	25	33
0.05	8.0	0.4	25	0.15	0.06	0.04	2	1	1	1	0	1
0.05	8.0	0.4	50	0.15	0.06	0.04	6	0	1	5	0	1
0.05	8.0	0.4	100	0.15	0.06	0.04	11	2	1	10	2	1
0.05	8.0	0.4	200	0.15	0.06	0.04	21	3	4	20	3	4
0.05	8.0	0.4	500	0.15	0.06	0.04	51	6	13	50	6	13
0.05	8.0	0.4	25	0.08	0.03	0.02	2	1	1	1	0	1
0.05	8.0	0.4	50	0.08	0.03	0.02	5	1	1	4	1	1
0.05	8.0	0.4	100	0.08	0.03	0.02	13	0	1	12	0	1
0.05	8.0	0.4	200	0.08	0.03	0.02	26	1	1	25	1	1
0.05	8.0	0.4	500	0.08	0.03	0.02	60	4	6	60	3	6
0.10	8.0	0.4	25	0.30	0.12	0.08	2	0	1	1	0	1
0.10	8.0	0.4	50	0.30	0.12	0.08	2	2	2	1	3	1
0.10	8.0	0.4	100	0.30	0.12	0.08	2	5	4	2	4	4
0.10	8.0	0.4	200	0.30	0.12	0.08	4	9	9	4	8	9
0.10	8.0	0.4	500	0.30	0.12	0.08	11	22	22	10	19	25
0.10	8.0	0.4	25	0.15	0.06	0.04	1	1	1	1	0	1
0.10	8.0	0.4	50	0.15	0.06	0.04	5	0	1	4	0	1
0.10	8.0	0.4	100	0.15	0.06	0.04	9	1	1	8	1	1
0.10	8.0	0.4	200	0.15	0.06	0.04	16	2	4	16	2	3
0.10	8.0	0.4	500	0.15	0.06	0.04	40	5	10	39	5	10
0.10	8.0	0.4	25	0.08	0.03	0.02	1	1	1	1	0	1
0.10	8.0	0.4	50	0.08	0.03	0.02	4	1	1	3	1	1
0.10	8.0	0.4	100	0.08	0.03	0.02	10	0	1	9	0	1
0.10	8.0	0.4	200	0.08	0.03	0.02	21	0	1	20	0	1
0.10	8.0	0.4	500	0.08	0.03	0.02	48	2	5	47	2	5
0.50	8.0	0.4	25	0.30	0.12	0.08	0	0	1	0	0	1
0.50	8.0	0.4	50	0.30	0.12	0.08	1	0	1	1	0	1
0.50	8.0	0.4	100	0.30	0.12	0.08	1	2	1	1	1	2
0.50	8.0	0.4	200	0.30	0.12	0.08	1	2	4	1	2	4
0.50	8.0	0.4	500	0.30	0.12	0.08	4	7	7	3	6	8
0.50	8.0	0.4	25	0.15	0.06	0.04	0	0	1	0	0	1
0.50	8.0	0.4	50	0.15	0.06	0.04	1	0	1	1	0	1
0.50	8.0	0.4	100	0.15	0.06	0.04	3	0	1	3	0	1
0.50	8.0	0.4	200	0.15	0.06	0.04	5	1	1	5	1	1
0.50	8.0	0.4	500	0.15	0.06	0.04	13	2	3	12	2	3
0.50	8.0	0.4	25	0.08	0.03	0.02	0	0	1	0	0	1
0.50	8.0	0.4	50	0.08	0.03	0.02	1	0	1	1	0	1
0.50	8.0	0.4	100	0.08	0.03	0.02	3	0	1	3	0	1
0.50	8.0	0.4	200	0.08	0.03	0.02	6	0	1	6	0	1
0.50	8.0	0.4	500	0.08	0.03	0.02	15	1	2	14	1	2
0.80	8.0	0.4	25	0.30	0.12	0.08	0	0	1	0	0	1
0.80	8.0	0.4	50	0.30	0.12	0.08	0	0	1	0	0	1
0.80	8.0	0.4	100	0.30	0.12	0.08	1	0	1	1	0	1
0.80	8.0	0.4	200	0.30	0.12	0.08	1	1	1	1	1	1
0.80	8.0	0.4	500	0.30	0.12	0.08	1	3	2	1	2	3
0.80	8.0	0.4	25	0.15	0.06	0.04	0	0	1	0	0	1
0.80	8.0	0.4	50	0.15	0.06	0.04	0	0	1	0	0	1
0.80	8.0	0.4	100	0.15	0.06	0.04	1	0	1	1	0	1
0.80	8.0	0.4	200	0.15	0.06	0.04	2	0	1	2	0	1
0.80	8.0	0.4	500	0.15	0.06	0.04	4	1	1	4	1	1
0.80	8.0	0.4	25	0.08	0.03	0.02	0	0	1	0	0	1
0.80	8.0	0.4	50	0.08	0.03	0.02	0	0	1	0	0	1
0.80	8.0	0.4	100	0.08	0.03	0.02	1	0	1	1	0	1
0.80	8.0	0.4	200	0.08	0.03	0.02	2	0	1	2	0	1
0.80	8.0	0.4	500	0.08	0.03	0.02	4	1	1	4	1	1