

A Combination Capture-Recapture and Line Transect Model in Clustered Population

Jinsik Choi¹⁾, Pyong Namkung²⁾

Abstract

In this paper, we present combined estimator of capture-recapture and line transect model using bivariate detection function and detection probability according to objects being in cluster population. Here bivariate detection function use distance and cluster size. The simulation shows that combined estimator approaches the more true value the larger size parameter. Therefore this estimator using the bivariate detection function is more efficient in estimate the population size and density by size parameter.

1. 서 론

자연상태에 있는 모집단의 구조와 동적 요소를 이해하기 위해서는 근본적으로 모집단의 크기를 알고 있어야 한다. 그러므로 자연상태에 있는 모집단, 특히 동물의 수를 추정하는 것은 일반적으로 생태적인 분석에 중요한 도구가 된다.

모집단의 크기나 생존율과 같은 모수를 추정하기 위한 방법으로는 포획·재포획 모형, 선횡단 모형과 적응집락 표본추출법, 공간예측 등 여러 가지 모형들이 제시되었는데 본 논문에서는 포획·재포획 모형(capture-recapture model)과 선횡단 모형(line transect model)을 중심으로 연구하고자 한다.

포획·재포획 모형과 선횡단 모형은 목표 모집단(target population)이 조사기간 동안에 출생이나 사망 그리고 새로이 유입되거나 전출되는 요인으로 인한 개체의 증감이 없다고 가정되었을 때 개체의 수 또는 모집단의 밀도를 추정하기 위한 것이다. 이 경우에 모집단의 크기는 조사기간 동안 상수로 취급하며 조사의 설계는 짧은 시간에 단편적으로 고려해야 한다. 그러나 이 가정은 비현실적인 경우가 많아 최근에는 모형의 가정을 실제 자연상태와 거의 같아지도록 완화시켜 모집단 모수에 대하여 보다 효율적이고 편의가 작은 추정량을 구할 수 있는 모형들이 주된 연구 대상이 되고 있다.

포획·재포획 모형과 선횡단 모형의 유사한 점은 개체를 포획하는 포획확률과 개체를 발견하게 되는 발견확률을 사용한다는 점이다. 포획확률과 발견확률은 개체들이 지니고 있는 특성 또는 고유의 이질성, 연구기간 동안의 날씨와 시간 같은 환경적 요인, 개체들이 집락을 이루고 있는 경우 등 여러 가지 요인에 의하여 변화하게 된다. 따라서 최근에는 포획·재포획 모형과 선횡단 모형을 결합하여 선횡단 모형의 가정 중 가장 기본적인 횡단선 위에서의 발견확률이 여러 가지 요인에 의하여 변화되는 경우에 발견확률을 추정하고, 이를 이용하여 모집단의 크기 또는 밀도를 추정

1) 217-071 강원도 속초시 노학동 산 244 동우대학 전자계산과 조교수

2) 110-745 서울시 종로구 명륜동 3-53 성균관대학교 통계학과 교수

하는 연구가 활발히 진행되고 있다.

선횡단 모형에서 한 명의 관측자가 관측하는 경우가 아니라 독립적인 여러 명의 관측자가 횡단선을 따라가며 개별적으로 개체를 관측하고 자료를 수집하게 되면 포획·재포획 모형의 원리를 이용하여 발견확률을 추정할 수가 있다. 이것이 두 모형을 결합하는 기본적인 원리가 되며 결국 두 모형을 결합한 형태는 선횡단 모형에서 횡단선 위에서 직접적으로 개체를 발견하게 되는 확률이 1이 된다는 가정이 위반되었을 경우 횡단선 위에서의 발견확률을 추정하게 하여 보다 효율적이고 편의가 작은 추정량을 제시하고 있다.

Alpizar-Jara(1997)가 제시한 추정량은 발견확률에 영향을 미치는 여러 가지 요인 중에서 단순히 횡단선으로부터 발견된 개체까지의 수직거리만을 고려한 형태의 포획·재포획 모형과 선횡단 모형을 결합한 추정량을 제시하였지만, 개체들이 집락을 이루고 있는 경우에 발견확률이 집락의 크기에 따라 변화되는 것에 대하여는 언급을 하지 않았다.

따라서 본 논문에서는 개체들이 집락의 형태를 갖는 모집단에 대해 거리뿐 아니라, 집락의 크기에 따른 발견확률을 Drummer와 Mcdonald(1987)가 제시한 거리와 집락크기의 이변량 발견함수를 최대 발견거리가 w 인 경우로 절삭시킨 이변량 발견함수를 사용하여 포획·재포획 모형과 선횡단 모형을 결합한 추정량을 제시하고자 한다.

2. 집락 모집단에서의 모수 추정

임의의 개체를 발견하는 것과 또 다른 임의의 개체를 발견하는 것이 통계적 독립사건이라는 가정은 개체들이 집단을 이루며, 연구지역 내에서 공간분포를 따르는 집락 모집단인 경우에는 명백히 위배된다. 이 경우에 임의의 개체를 발견하는 확률은 집단의 크기에 영향을 받게 될 것이다. 예를 들어 동물들이 무리를 이루고 있는 경우 발견하는 확률은 무리의 크기에 따라 영향을 받게 된다. 따라서 무리의 크기가 큰 경우는 동물을 발견하게 될 확률은 커지게 되므로 밀도의 추정량은 과대추정이 되고, 반대로 무리가 작은 경우는 발견하게 될 확률이 작아지므로 밀도는 과소추정된다.

이와 같이 모집단에 존재하는 개체들이 집락을 이루고 있는 경우 집락의 크기는 발견확률에 영향을 주게 되는데 이러한 상황을 '크기편의(size-biased)'라 한다. 그러므로 크기편의가 존재하는 상황에서는 개체 발견의 확률에서 집락의 크기를 고려해야 할 것이다. 즉, 선횡단 모형에서 모집단이 집락으로 구성되어 있는 경우 집락의 크기는 발견가능률(detectability)에 영향을 주게 되므로 보다 정확한 모집단의 밀도 또는 크기를 추정하기 위해서는 집락의 크기를 고려한 추정량이 되어야 한다.

우선 기본적인 선횡단 모형에서의 가정과 같이 횡단선은 연구지역에 랜덤하게 위치하며 임의의 한 집락을 발견하는 것은 다른 임의의 집락을 발견하는 사건과는 독립이라 가정하자. 또한 임의의 집락은 단 한 번만 세어지고 표본추출 동안에 집락은 이동하지 않는다고 하며 동시에 발견된 집락 또는 집단 내에 존재하는 모든 개체들을 셀 수 있다고 가정하자. 즉, 집락이 발견된 거리와 집락의 크기(집락 내에 존재하는 개체의 수)는 정확하게 측정이 되고 임의의 집락을 발견하는 확률은 횡단선으로부터 수직거리와 집락의 크기 s 에 영향을 받게 된다.

따라서 집락을 발견한 최대 발견거리 w 로 고정된 폭 $2w$ 에 대하여 모집단의 밀도 추정량은 다음과 같다.

$$\begin{aligned} \hat{D} &= \frac{\hat{E}(n_w)\hat{E}(s)}{A_w} \\ &= \frac{n_w\hat{E}(s)}{2wL\hat{P}_w} \end{aligned} \tag{1}$$

여기서, $\hat{E}(n_w)$ = 연구지역 내에서 평균 집락 수의 추정량
 $\hat{E}(s)$ = 연구지역 내에서 평균 집락크기의 추정량
 $\hat{E}(n_w)\hat{E}(s) = \hat{N}_w$ = 연구지역 내의 추정된 개체 수
 \hat{P}_w = 연구지역 내에서 임의의 집락을 발견하는 확률
 n_w = 연구지역 내에서 발견된 집락의 수
 L = 횡단선의 총 길이
 $A_w = 2wL$ = 연구지역의 면적.

식 (1)에서 임의의 집락을 발견할 확률 \hat{P}_w 는 발견된 집락 수 n_w 가 주어진 조건에서 추정된 확률이다. 만약 $E(\hat{P}_w | n_w) = E(\hat{P}_w)$ 로 n_w 와 \hat{P}_w 가 독립이라고 가정하면 P_w 의 불편추정량은 $(\hat{P}_w | n_w)$ 이 되며, 이 경우 n_w 와 \hat{P}_w 의 공분산은 0이 된다. 마찬가지로 $E(s)$ 는 n_w 과 발견된 거리가 주어진 조건에서 추정되며 $\hat{E}(s)$ 는 n_w 또는 \hat{P}_w 와 상관이 없게 된다. 그리고 집락의 표본 크기 n_w 는 조사기간 동안에 발견된 집락의 수이고 개체들의 수를 나타내는 s 는 관측된 각 집락 내에서 개별적인 개체들의 수를 셈하면 된다.

그러므로 개별적인 개체들의 밀도는 평균 집락크기의 추정량 $\hat{E}(s)$ 에 집락의 밀도 추정량 $\hat{D}(s)$ 의 곱인

$$\hat{D} = \hat{D}(s) \cdot \hat{E}(s) \tag{2}$$

으로 나타낼 수 있다.

이제 $g(x | s)$ 를 크기 s 의 집락이 주어졌을 때 수직거리 x 에서 집락을 발견할 확률에 대한 조건부 발견함수로 정의하면 이에 대응하는 확률밀도함수 $f(x | s)$ 는 다음과 같이 된다.

$$f(x | s) = \frac{g(x | s)}{wP_w} \tag{3}$$

식 (3)에서 분모를 $\mu_w = wP_w$ 라 하면 $f(x | s)$ 는 확률밀도함수이기 때문에

$$\mu_w = \int_0^w g(x | s) dx \quad (4)$$

가 된다. 따라서 확률밀도함수는

$$f(x | s) = \frac{g(x | s)}{\int_0^w g(x | s) dx} \quad (5)$$

가 되고 $x=0$ 인 횡단선 위에서 $g(0 | s) = 1$ 이라는 가정을 이용하면 다음과 같다.

$$\begin{aligned} f(0 | s) &= \frac{g(0 | s)}{\int_0^w g(x | s) dx} \\ &= \frac{1}{\int_0^w g(x | s) dx} \end{aligned} \quad (6)$$

$$P_w = \frac{\int_0^w g(x | s) dx}{w} \quad (7)$$

또한 발견된 집락의 수인 확률변수 n_w 는 연구지역 내에 존재하는 집락의 총 수 N_c 와 집락을 발견하는 발견확률 P_w 를 갖는 이항분포를 가정하면 모집단의 밀도 추정량은 다음과 같다.

$$\begin{aligned} \hat{D} &= \frac{\hat{E}(n_w) \hat{E}(s)}{2wL\hat{P}_w} \\ &= \frac{n_w \hat{f}(0 | s) \hat{E}(s)}{2L} \end{aligned} \quad (8)$$

여기서 모집단 밀도를 추정하기 위해서는 $E(s)$ 와 $f(0 | s)$ 를 추정하면 된다.

Buckland 등(1993)은 집락으로 구성된 모집단으로부터 얻어진 자료를 분석하는 방법을 제시하고 있는 데 집락의 경우 관심을 갖는 개체처럼 집락을 정의하고 집락의 크기를 보조변수로 척도화하여 집락모집단의 밀도를 추정할 수 있다고 제안하였다.

집락의 크기를 추정하는 방법은 다음과 같다.

첫째, 횡단선으로부터 발견된 거리가 먼 집락을 절삭(truncation)시키는 방법이다. 이 방법은 집락의 중심이 발견 최대거리 w 밖에 존재하면 이 집락의 크기를 무시하고 무조건 제외시키는 방법으로 발견거리를 절삭시키는 방법과 유사하며, 집락을 절삭시킨 거리가 절삭거리와 같을 필요는 없으나 절삭거리 선정에 주의를 해야 한다.

둘째, Quinn(1979, 1985)이 제안한 방법으로 집락의 크기를 사후층화 방법을 이용하여 층화시킨

후 모집단의 밀도를 추정하며 각 층에 $f(0)$ 를 추정하기 위한 적절한 표본크기가 요구된다.

셋째, Best와 Butterworth(1980)는 집락크기에 대한 평균거리 x 를 이용한 가중회귀 추정량을 제안하였다. 이 방법은 집락크기에 의한 크기편의에 대한 정보를 잃어버리는 경향이 존재한다.

넷째, 발견함수에 대한 모형으로 발견거리 x 와 크기변수 s 를 고려한 이변량 발견함수를 이용하는 방법으로 Drummer와 McDonald(1987)가 제안하였는데, Quang(1991)은 이 방법을 적용한 비모수적 방법을 제안하였다.

다섯째, 발견된 집락을 하나의 개체로 간주하고 재표본 추출방법(resampling method)을 이용하여 모집단의 밀도를 추정한다.

본 장에서는 네 번째 추정방법을 사용하여 포획·재포획 모형과 선횡단 모형을 결합한 형태의 추정량을 제시하고자 한다.

3. 발견확률이 집락크기와 발견거리에 영향을 받는 경우

집락크기를 고려한 모집단 밀도에 대해 보다 정도가 높은 추정량을 구하기 위하여는 집락크기와 집락의 발견거리를 동시에 고려해야 할 것이다. 그러나 집락크기는 집락을 발견하는 발견거리에 영향을 주게 되므로 평균집락 크기 $E(s)$ 의 불편추정량을 구하는데 어려움이 있기 때문에 분석이 보다 복잡해진다.

선횡단 모형에서 집락크기가 집락을 발견한 발견거리에 영향을 주는 경우에 대하여 Drummer와 McDonald(1987)는 집락의 크기와 발견거리가 모두 무한대인 경우를 가정한 지수분포족에서의 이변량 발견함수의 형태로 변형하여 사용하였다.

여기서는 Alpizar-Jara의 포획·재포획 모형과 선횡단 모형을 결합한 모형을 발견거리와 집락크기가 발견확률에 영향을 미치는 경우에 집락 모집단의 밀도를 추정하기 위하여 이변량 발견함수를 발견거리가 최대 w 가 되는 독립적인 두 명의 관측자가 횡단선을 따라 관측한 포획·재포획 모형과 선횡단 모형을 결합한 추정량을 고려하였다.

임의의 집락이 발견되는 확률은 발견되는 집락의 지리적 중심까지 횡단선으로부터의 수직거리와 발견되는 집락의 크기인 확률변수 S 값에 따라 달라지는 것을 고려하면 아래와 같은 가정을 설정할 수 있다.

- 1) 전체 길이 L 이 알려진 횡단선은 면적 A_w 의 연구지역에 랜덤하게 위치한다.
- 2) 확률변수 S 는 확률분포 $\pi(s), s > 0$ 을 가지면 연속확률밀도함수로 근사 가능하다.
- 3) 임의의 집락을 발견하는 것은 다른 집락을 발견하는 것과는 독립이고 하나의 집락은 오직 한 번만 세어지고 연구기간 동안 즉, 표본추출기간에 움직이거나 이동하지 않는다.
- 4) 임의의 집락을 발견하는 확률은 횡단선으로부터의 수직거리 x 와 확률변수 S 의 값인 s 의 함수가 된다. 여기서 확률변수 S 를 집락의 크기로 하자.

위의 가정에 따른 기호를

$$N_c = \text{연구지역 내에 존재하는 집락의 총 수}$$

$$n_w = \text{연구지역 내에서 발견된 집락의 수}$$

$x_i =$ 발견된 i 번째 집락의 지리적 중심까지의 수직거리 $i=1, \dots, n_w$

$s_i =$ 발견된 i 번째 집락의 크기 $i=1, \dots, n_w$

$E(s) =$ 평균 집락크기

$N_w = N_c E(s) =$ 연구지역내의 개별적 개체의 총 수

라 정의하고 (x_i, s_i) 는 측정오차 없이 관측된다고 가정하자.

그러면 관심을 갖는 모수는 연구지역 내에 존재하는 집락의 총 수 N_c 와 평균 집락크기 $E(s)$ 가 된다. 앞에서 정의한 가정에 따르면 이변량 발견함수

$$g(x, s), \quad 0 < g(x, s) \leq 1 \quad (9)$$

가 존재하게 되고 이 발견함수는 수직거리 x 와 집락크기 s 가 주어진 상태에서 집락을 발견하게 되는 확률로 다음을 가정할 수 있다.

- 1) 횡단선 위에 직접적으로 존재하는 집락이 발견되는 확률은 1이 된다. 즉, $g(0, s) = 1$ 이다.
- 2) $g(x, s)$ 는 x 에 대하여 비증가(nonincreasing) 함수가 된다.
- 3) $g(x, s)$ 는 s 에 대하여 비감소(nondecreasing) 함수가 된다.
- 4) 연구지역의 경계에서 임의의 집락을 발견하는 확률은 집락의 크기와 상관없이 근사적으로 0이 된다.

Drummer와 McDonald는 이변량 발견함수를 기본적인 선횡단 모형의 발견함수에 거리 x 만을 사용한 것이 아니라 크기를 비율 x/s^α 를 일변량 발견함수의 형태로 사용하기 위하여 집락의 크기 s 를 공변량(covariate)으로 삼입하여 사용하였다.

따라서 이변량 발견함수에서 집락의 크기 s 는 발견함수의 척도모수처럼 사용되고 모수 α 는 크기편의를 나타내는 모수로 발견확률에 영향을 미치게 되는 집락크기의 정도(degree)를 반영하고 있다. 그러므로 크기편의를 나타내는 α 의 값이 0이 되면 집락의 크기는 발견확률에 아무런 영향을 미치지 않는다는 것을 나타낸다. 즉, 이 경우 개체들은 집락을 이루고 있는 것이 아니라 개별적으로 연구지역 내에 위치하고 있다는 것이 된다. 그리고 크기모수 α 의 값이 커질수록 집락의 크기가 크다는 것으로 발견확률에 많은 영향을 미치게 된다.

이변량 발견함수를 갖는 모형에서 모수 추정은 일반적인 선횡단 모형의 추정절차를 고려하여 P_w 를 임의의 집락을 발견확률이라 정의하면

$$P_w = \frac{2L}{A_w} \int \int_0^w g(x, s) \pi(s) dx ds \quad (10)$$

가 되며, 연구지역의 면적이 $A_w = 2wL$ 이므로 발견확률은 다음과 같다.

$$P_w = \frac{1}{w} \int \pi(s) \int_0^w g(x, s) dx ds \quad (11)$$

임의의 집락을 발견하는 확률인 식 (11)에서 이변량 발견함수 $g(x, s)$ 는 집락 크기 s 와 발견거리 x 에 대하여 x/s^α 를 고려하므로 지수함수에서 다음과 같은 발견함수로 나타낼 수 있다.

$$\begin{aligned} \int_0^w g(x, s)dx &= \int_0^w g(x) \cdot s^\alpha dx \\ &= s^\alpha \cdot \int_0^w g(x)dx \end{aligned} \tag{12}$$

따라서 식 (12)에서 $\int_0^w g(x)dx$ 의 값은 상수가 되며 발견함수 $g(x)$ 는 모수의 함수로 발견함수가 EPS(exponential power series)모형인 경우의 이변량 발견함수는 다음과 같다.

$$g(x, s) = \beta e^{-\left(\frac{x}{\lambda s^\alpha}\right)^\alpha} \tag{13}$$

여기서, $\beta = g(0)$ 로 횡단선에서 개체를 발견하는 확률
 α ; 크기모수(size parameter)
 q ; 형태모수(shape parameter).

따라서 임의의 집락을 발견하게 되는 확률은 식 (11)과 식 (12)으로부터 다음과 같다.

$$\begin{aligned} P_w &= \frac{1}{w} \int \pi(s) \int_0^w g(x, s)dx ds \\ &= \frac{1}{w} \mu_\alpha \int_0^w g(x)dx \end{aligned} \tag{14}$$

여기서 $\mu_\alpha = \int s^\alpha \pi(s)ds < \infty$.

한편, 임의의 한 집락을 발견하는 것은 독립이라는 가정으로부터 발견된 연구지역 내에서 집락의 수 n_w 의 확률분포는 모수 N_c 와 P_w 인 이항분포를 따르게 되므로 N_c 의 최우추정량은 다음과 같다.

$$\begin{aligned} \hat{N}_c &= \frac{n_w}{\hat{P}_w} \\ &= \frac{n_w w}{\int_0^w g(x)dx} \frac{1}{\hat{\mu}_\alpha} \end{aligned} \tag{15}$$

그러므로 연구지역 내에 존재하는 개별적 개체의 총 수 N_w 는

$$\hat{N}_w = \hat{N}_c \cdot \hat{E}(s) \tag{16}$$

이고, $\hat{E}(s)$ 는 평균 집락크기의 추정량이므로 식 (15)를 이용하면 연구지역 내에 존재하는 총 개체의 추정량은

$$\hat{N}_w = \frac{n_w w}{\int_0^w g(x) dx} \cdot \frac{\hat{E}(s)}{\hat{\mu}_\alpha} \tag{17}$$

가 된다.

이 경우에 가중분포이론(weighted distribution, Patil과 Ord 1976)을 이용하면 지수족을 갖는 발견함수에 대하여 수직거리와 집락크기의 결합관측분포는

$$\begin{aligned} f^*(x, s) &= \frac{g(x, s)\pi(s)}{\int \int_0^w g(x, s)\pi(s) dx ds} \\ &= \frac{g(x, s)\pi(s)}{\int_0^w g(x) dx \int s^\alpha \pi(s) ds} \\ &= \frac{g(x, s)\pi(s)}{\mu_\alpha \int_0^w g(x) dx} \end{aligned} \tag{18}$$

이며 관측된 집락 크기의 분포는 다음과 같다.

$$\begin{aligned} \pi^*(s) &= \int_0^w f^*(x, s) dx \\ &= \frac{\int_0^w g(x, s)\pi(s) dx}{\mu_\alpha \int_0^w g(x) dx} \\ &= \frac{s^\alpha \pi(s)}{\mu_\alpha}, \quad s > 0 \end{aligned} \tag{19}$$

여기서 $\pi^*(s)$ 는 차수 α 의 일반적 크기편의 분포가 된다.

또한 발견이 독립이라는 가정하에 발견된 집락 수의 조건결합밀도함수는

$$\begin{aligned} f(x_1, \dots, x_{n_c}, s_1, \dots, s_{n_c} \mid n_w) \\ = \prod_{i=1}^{n_c} \frac{g(x_i, s_i) \pi(s_i)}{\mu_\alpha \int_0^w g(x) dx} \end{aligned} \tag{20}$$

이 되며, 식 (20)으로부터 $\int_0^w g(x)dx$ 와 α 를 추정하기 위하여는 $\mu_\alpha = \int s^\alpha \pi(s)ds$ 와 $\pi(s)$ 에 관한 정보가 필요하게 되므로 관측된 크기 s 를 조건으로 발견된 수직거리 x 를 고려하면 다음과 같다.

$$f^*(x | s) = \frac{g(x, s)}{s^\alpha \int_0^w g(x)dx} = \frac{g(x, s)}{s^\alpha \mu_w} \tag{21}$$

여기서 $\mu_w = \int_0^w g(x)dx$.

따라서 μ_w 와 α 는 조건결합밀도함수

$$\begin{aligned} f^*(x_1, \dots, x_{n_x} | s_1, \dots, s_{n_x}, n_w) \\ = \prod_{i=1}^{n_x} \frac{g(x_i, s_i)}{s_i^\alpha \int_0^w g(x)dx} \end{aligned} \tag{22}$$

을 이용하여 추정한다.

그러므로 임의의 집락을 발견하는 것은 다른 집락을 발견하는 것과는 독립이 되기 때문에 발견하는 집락의 수 n_w 는 모집단의 집락의 수 N_c 와 발견확률 P_w 를 갖는 이항분포를 따르므로 다음의 우도함수를 이용하여 추정한다.

$$\begin{aligned} L &= f^*(x_1, \dots, x_{n_x} | s_1, \dots, s_{n_x}, n_w) \\ &= f^*(x_1, \dots, x_{n_x} | s_1, \dots, s_{n_x}, n_w) f(n_w) \\ &= \prod_{i=1}^{n_x} \frac{g(x_i, s_i)}{s_i^\alpha \int_0^w g(x)dx} \binom{N_c}{n_w} P_w^{n_x} (1 - P_w)^{N_c - n_x} \end{aligned} \tag{23}$$

3.1 포획 · 재포획 모형과 선횡단 모형의 결합

집락 모집단의 경우에 모집단 밀도를 추정하기 위한 포획 · 재포획 모형과 선횡단 모형의 결합을 정의하기로 하자. 우선 두 명의 관측자가 독립적으로 동물을 관측한다고 가정하고 집락 모집단에 대하여 관측자 각각에 대한 두 개의 독립적인 발견함수 $g_1(x, s)$ 와 $g_2(x, s)$ 가 있다고 가정하자.

식 (23)의 우도함수를 일반화시키면 다음과 같다.

$$L = L_1 L_2 L_3 L_4 L_5 \tag{24}$$

여기서

$$\begin{aligned} L_1 &= f^{*11}(x_1, \dots, x_{n_{11}} \mid s_1, \dots, s_{n_{11}}, n_{11}) \\ &= \prod_{i=1}^{n_{11}} f^{*11}(x_i \mid s_1, \dots, s_{n_{11}}, n_{11}) \\ &= \prod_{i=1}^{n_{11}} \frac{g_1(x_i, s_i)g_2(x_i, s_i)}{s_{1i}^{a_1}s_{2i}^{a_2} \int_0^w g_1(x)g_2(x) dx} \end{aligned} \tag{25}$$

$$\begin{aligned} L_2 &= f^{*10}(x_1, \dots, x_{n_{10}} \mid s_1, \dots, s_{n_{10}}, n_{10}) \\ &= \prod_{i=1}^{n_{10}} f^{*10}(x_i \mid s_1, \dots, s_{n_{10}}, n_{10}) \\ &= \prod_{i=1}^{n_{10}} \frac{g_1(x_i, s_i)(1 - g_2(x_i, s_i))}{s_{1i}^{a_1}s_{2i}^{a_2} \int_0^w g_1(x) (1/s_{2i}^{a_2} - g_2(x)) dx} \end{aligned} \tag{26}$$

$$\begin{aligned} L_3 &= f^{*01}(x_1, \dots, x_{n_{01}} \mid s_1, \dots, s_{n_{01}}, n_{01}) \\ &= \prod_{i=1}^{n_{01}} f^{*01}(x_i \mid s_1, \dots, s_{n_{01}}, n_{01}) \\ &= \prod_{i=1}^{n_{01}} \frac{(1 - g_1(x_i, s_i))g_2(x_i, s_i)}{s_{1i}^{a_1}s_{2i}^{a_2} \int_0^w (1/s_{1i}^{a_1} - g_1(x))g_2(x) dx} \end{aligned} \tag{27}$$

$$\begin{aligned} L_4 &= f^*(n_{11}, n_{10}, n_{01} \mid s_1, \dots, s_{n_w}, n_w) \\ &= \binom{n_w}{n_{11}, n_{10}, n_{01}} \left[\frac{s_{1i}^{a_1}s_{2i}^{a_2} \int_0^w g_1(x)g_2(x) dx}{\mu_w^*} \right]^{n_{11}} \\ &\quad \times \left[\frac{s_{1i}^{a_1}s_{2i}^{a_2} \int_0^w g_1(x) (1/s_{2i}^{a_2} - g_2(x)) dx}{\mu_w^*} \right]^{n_{10}} \\ &\quad \times \left[\frac{s_{1i}^{a_1}s_{2i}^{a_2} \int_0^w (1/s_{1i}^{a_1} - g_1(x))g_2(x) dx}{\mu_w^*} \right]^{n_{01}} \end{aligned} \tag{28}$$

$$L_5 = \binom{N_c}{n_w} (P_w^*)^{n_w} (1 - P_w^*)^{N_c - n_w} \tag{29}$$

가 되며, 여기서

- $n_w = n_{11} + n_{10} + n_{01} =$ 최소한 한 명의 관측자가 발견한 집락의 수
- $n_{11} =$ 두 명의 관측자 모두가 발견한 집락의 수
- $n_{10} =$ 오직 첫 번째 관측자만 발견한 집락의 수
- $n_{01} =$ 오직 두 번째 관측자만 발견한 집락의 수

를 나타낸다.

또한 두 명의 관측자에 의하여 임의의 개체가 횡단선으로부터 발견되는 위치에 따른 확률은

$$\begin{aligned} \mu_w^* &= \int_0^w \left[1 - \prod_{i=1}^2 (1 - g_i(x, s)) \right] dx \\ &= \int_0^w (g_1(x, s) + g_2(x, s) - g_1(x, s)g_2(x, s)) dx \end{aligned} \tag{30}$$

이고 최소한 한 명의 관측자가 발견하는 집락의 발견확률은

$$\hat{P}_w^* = \frac{\mu_w^*}{w} \tag{31}$$

이므로 모집단 밀도 추정량은 다음과 같다.

$$\hat{D} = \frac{\hat{N}_w}{A_w} = \frac{n_w \hat{E}(s)}{2wL \hat{P}_w^*} \tag{32}$$

여기서

$$\begin{aligned} \hat{N}_c &= \frac{n_w}{\hat{P}_w^*} \\ A_w &= 2wL \\ \hat{N}_w &= \hat{N}_c \hat{E}(s) \end{aligned}$$

3.2 모수 추정절차

발견거리와 집락크기가 집락을 발견하는 확률에 영향을 주는 경우 모집단 밀도를 추정하기 위한 추정절차 또한 앞에서 고려한 방식과 동일한 절차를 적용할 수가 있다.

따라서 임의의 집락을 발견하는 발견확률을 두 명의 관측자 $h=1, 2$ 와 횡단선을 기준으로 한 k 개의 거리범주($j=1, 2, \dots, k$) 그리고 j 번째 거리범주에서 발견된 각각의 집락 크기 s_i ($i=1, \dots, n_{wj}$)에 따라서 다음과 같이 된다.

$$g_h(x_j, s_i) = g_h\left(\frac{x_j}{s_i}\right) = p_{hij}, \quad w_{j-1} < x < w_j$$

위 식에 주어진 확률은 발견된 집락의 크기에 따라 달라지게 될 것이다. 이 경우 집락 크기에 따른 모든 확률을 계산하기가 복잡해진다. 이를 단순화하기 위하여 발견된 집락의 크기를 동일한 거리범주 내에서는 각각의 집락크기 s_i 를 집락크기 s 의 가장 단순한 추정량으로 하면 각 거리범주내의 평균 집락크기

$$E_j(s) = \bar{s}_j = \frac{1}{n_{wj}} \sum_i^{n_{wj}} s_i$$

를 고려할 수 있다. 여기서 j 는 j 번째 거리범주를 나타내며, n_{wj} 는 j 번째 거리범주 내에서 발견된 집락의 수를 나타낸다.

이제 j 번째 거리범주 내에서 집락이 발견된 거리를 거리범주의 중간값인 $x_j = (w_{j-1} + w_j)/2$ 와 평균집락크기 \bar{s}_j 를 이용하면 발견확률의 추정량은 다음과 같다.

$$\hat{g}_h(x_j, \bar{s}_j) = p_{hj}$$

따라서 식 (24)의 우도함수는 식 (33)으로 나타낼 수 있다.

$$\begin{aligned} L &= \prod_{j=1}^k \binom{N_{wj}}{n_{wj}} p_{wj}^{n_{wj}} (1 - p_{wj})^{N_{wj} - n_{wj}} \\ &\times \prod_{j=1}^k \binom{n_{wj}}{n_{11j}, n_{10j}, n_{01j}} \left[\frac{p_{1j} p_{2j}}{p_{wj}} \right]^{n_{11j}} \left[\frac{p_{1j}(1 - p_{2j})}{p_{wj}} \right]^{n_{10j}} \\ &\times \left[\frac{(1 - p_{1j}) p_{2j}}{p_{wj}} \right]^{n_{01j}} \end{aligned} \tag{33}$$

또한 결합 모형을 고려한 모집단 크기의 추정량은 각 거리범주 내에서 집락의 밀도추정량에 평균 집락크기를 곱하여 구한다.

$$\hat{N}_{wj} = \hat{N}_{wj} \hat{E}_j(s) = \frac{n_{wj}}{\hat{p}_{wj}} \hat{E}_j(s) \tag{34}$$

여기서

$$\hat{p}_{wj} = \hat{p}_{1wj} + \hat{p}_{2wj} - \hat{p}_{1wj} \hat{p}_{2wj} \tag{35}$$

$$n_{wj} = n_{11j} + n_{10j} + n_{01j} \tag{36}$$

식 (35)의 \hat{p}_{wj} 는 거리계급 j 에서 최소한 한 명의 관측자에게서 집락을 발견하는 확률의 추정값이고, 식 (36)의 n_{wj} 는 거리계급 j 에서 최소한 한 명의 관측자에게 발견된 집락의 수를 나타낸다. 결국 전체 지역의 모집단 크기 추정량은 각 거리범주의 모집단 밀도추정량들의 합으로 구할 수 있다.

$$\hat{N}_w = \sum_{j=1}^k \hat{N}_{wj} \tag{37}$$

그러므로 모집단 밀도 추정량 \hat{D} 의 분산 추정량은 델타방법(delta method)을 이용하면 다음과 같이 근사 되어진다.

$$Var(\hat{D}) \approx \frac{Var(\hat{N}_w)}{A_w^2} \tag{38}$$

$$\begin{aligned} Var(\hat{N}_w) &\approx N_w^2 \left\{ \frac{Var(n_w)}{n_w^2} + \frac{Var\hat{E}(s)}{[\hat{E}(s)]^2} + \frac{Var(P_w^*)}{P_w^{*2}} \right\} \\ &= N_w^2 \{ [cv(n_w)]^2 + [cv(P_w^*)]^2 + [cv[E(s)]]^2 \} \end{aligned} \tag{39}$$

4. 모의 실험

모의실험을 위한 자료의 구성은 모집단의 집락크기가 포아송 분포를 따른다고 가정하고 집락크기 s 는 평균이 2가 되는 포아송 분포를 따르는 모집단을 구성하였다. 그러므로 집락크기의 평균과 분산은 각각 $E(s)=2$ 와 $Var(s)=E(s)=2$ 가 된다.

또한 각각의 개체들은 3장과 동일하게 5개의 거리범주(예, 0-5, 5-10, 10-15, 15-20, 20-25)를 고려하여 각 거리범주마다 $N_j=100$ 마리의 동물이 평균이 2인 포아송 분포를 이루며 존재하는 것으로 하여, 총 $N=500$ 마리의 동물이 전체지역 내에 $N_c=237$ 개의 집락을 이루고 있는 모집단을 구성하였다.

이 경우에 모집단 밀도 추정을 위해서는 집락의 밀도인 D_s 의 추정이 필요하게 되는데 이를 추정하기 위한 이변량 발견함수를 Pollock의 EPS모형을 사용하였다. 그리고 각 거리범주에서 이변량 발견함수의 발견확률을 정하기 위하여 두 명의 독립적인 관측자($h=1, 2$)에 대하여 척도모수 $\lambda=16, 20$ 이고 $\beta=0.8, 0.9$ 설정하여 모수 선택에 따라 8가지 시나리오를 선택하였으며 그것이 <표 1>에 나타나 있다. 이 경우 독립적인 두 명의 관측자에 대한 이변량 발견함수는 식 (13)에 의해 다음과 같다.

$$g_h(x, s) = \beta_h e^{-\left(\frac{x}{\lambda_h s^{\alpha}}\right)^{\alpha}}$$

또한 각각의 시나리오에서 크기모수 α 가 각각 0.6, 1.0, 1.4인 경우를 고려하였고 형태모수 q 가 2 와 3일 때 반복적으로 각각 500번의 모의실험을 실시하였으며 추정량의 효율성을 파악하기 위하여 상대편의(%)와 상대비율 표준오차(%) 그리고 $\sqrt{\text{평균제곱오차(RMSE)}}$ 를 구하였다.

<표 1> 시나리오별 EPS 이변량 발견함수의 모수

시나리오	β_1	β_2	λ_1	λ_2
1	0.9	0.8	20	20
2	0.8	0.8	20	20
3	0.9	0.9	20	20
4	0.9	0.8	16	20
5	0.9	0.8	20	16
6	0.9	0.8	16	16
7	0.8	0.8	16	16
8	0.9	0.9	16	16

<표 2>에는 $q=2$ 일 때 8가지 유형의 모의실험에 대하여 크기모수의 값이 각각 0.6, 1.0, 1.4인 경우의 결과를 보여주고 있으며 <표 3>에는 $q=3$ 일 때 8가지 시나리오에 대하여 각각의 크기모수 값에 따른 결과를 보여주고 있다.

모의실험 결과 <표 2>와 <표 3>을 살펴보면 집락 수의 추정량인 \hat{N}_c 와 모집단 개체 수의 추정량 \hat{N} 는 거의 대부분의 경우에 약간의 과소추정이 존재하는 것으로 나타났다.

우선 집락 수의 추정량 \hat{N}_c 는 $q=2$ 인 경우 집락 수 추정값의 평균들은 $\alpha=0.6$ 일 때 최소값 235.3에서 최대값 236.9로, $\alpha=1.0$ 일 때 최소값 236.1에서 최대값 237.2로, $\alpha=1.4$ 일 때 최소값 236.4에서 최대값 237.4로 나타났으며, 개체 수 추정량 \hat{N} 도 $\alpha=1.4$ 일 때 최소값 495.9에서 최대값 498.0으로 나타나 크기모수 α 의 값이 작을 때 보다 상대적으로 편의가 작은 것으로 나타났다.

또한 $q=3$ 인 경우에도 집락 수의 추정량 \hat{N}_c 는 집락 수 추정값의 평균이 $\alpha=1.4$ 일 때 236.1에서 238.3으로, 개체 수의 추정값의 평균이 495.3에서 499.4로 나타나 크기모수 α 의 값이 작을 때 보다 상대적으로 편의가 작다. 즉, 다른 결과보다는 상대적으로 과소추정이 적어진다. 그리고 상대비율 표준오차와 RMSE를 살펴보다더라도 크기모수 α 의 값이 클수록 상대적으로 상대비율 표준오차와 RMSE의 값들이 작아지는 것을 확인할 수 있다.

이러한 결과는 모집단내에 존재하는 집락 수와 개체의 추정은 집락의 크기모수 α 의 값이 커짐

에 따라 상대편의가 상대적으로 작아지는 것으로, 결국 추정된 집락의 수와 개체 수는 크기모수 α 가 클수록 전체 집락 수의 모의실험 설정값인 $N_c = 237$ 와 개체 수 $N=500$ 에 더욱 근접해 있음을 확인할 수 있다. 따라서 집락의 크기와 발견거리를 사용한 이변량 발견함수에서 집락의 크기가 추정량에 영향을 주고 있다는 것을 보여주고 있다.

또한 <그림 1>은 $q=2$ 일 때 500번의 모의실험 결과 추정된 집락 수 추정량 \hat{N}_c 의 상자그림이고, <그림 2>는 $q=2$ 일 때 개체 수 추정량 \hat{N} 의 상자그림이다. 그리고 <그림 3>은 $q=3$ 일 때 모의실험 결과 집락 수 추정량 \hat{N}_c 의 상자그림이고 <그림 4>는 $q=3$ 일 때 개체 수 추정량 \hat{N} 의 상자그림이다.

<그림 1>에서 <그림 4>까지의 변수명 S1에서 S8은 각각 시나리오 1에서 8까지를 나타내고 NC는 집락 수 추정량, NS는 개체 수 추정량을 06, 10, 14는 각각 $\alpha=0.6$, $\alpha=1.0$, $\alpha=1.4$ 를 나타낸다.

또한 집락 수의 추정량 \hat{N}_c 와 개체 수 추정량 \hat{N} 에 대한 분포를 나타낸 <그림 1>에서 <그림 4>의 상자그림을 보더라도 크기모수 α 의 값이 클 때 추정량의 분포가 밀집해 있다는 것을 확인할 수 있다.

집락의 크기를 나타내는 크기모수의 변화와 거리에 따른 이변량 발견함수의 형태에 따른 모집단의 크기 또는 밀도를 추정하기 위한 모의실험 결과를 종합해 보면, 크기 정도를 나타내는 모수 α 의 크기가 커짐에 따라 포획·재포획 모형과 선횡단 모형을 결합한 형태의 추정량이 보다 효율적임을 알 수 있다.

따라서 모집단이 집락을 이루고 있는 경우에 집락 또는 전체 개체의 밀도 또는 크기를 추정하기 위해서는 집락의 크기를 정확하게 추정하는 것이 모집단 전체의 크기 또는 밀도를 추정하는데 중요한 역할을 하게 되므로, 향후 평균 집락 크기를 정확하게 추정하는 방법에 대한 연구가 필요하다.

5. 결 론

전통적으로 선횡단 표본추출방법과 포획·재포획방법은 동물모집단의 밀도를 추정하기 위해 사용되어 왔으나 모집단의 추정량을 과대 또는 과소추정하는 경우가 나타나 이를 해결하기 위한 많은 방법들이 연구되었다.

일반적인 선횡단 모형의 추정량은 하나의 개체를 발견하는 것은 통계적 독립사건을 가정한 발견함수를 사용하였는데 특정한 종류의 동물이나 식물들은 집락을 이루고 있는 경우에 임의의 개체를 발견하는 확률은 집락의 크기에 영향을 받게 된다. 따라서 거리에 영향을 받는 발견확률을 고려하지만 집락 모집단이나 크기편의된 모집단에서 나타나는 집락의 크기에 따라 달라지게 되는 발견확률을 고려한 결합추정량을 제시하고 모의실험을 하였다.

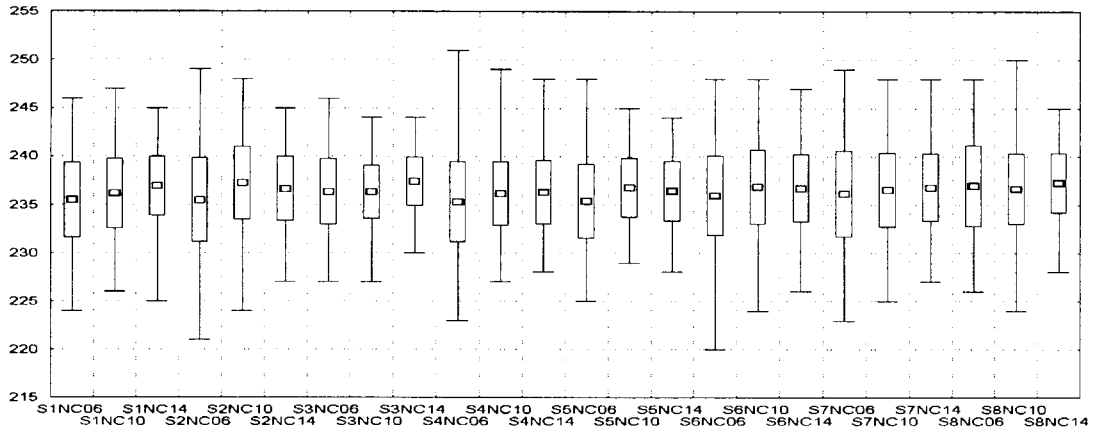
집락의 크기와 거리를 고려한 이변량 발견함수를 사용하여 포획·재포획 모형과 선횡단 모형을 결합한 추정량의 모의실험 결과는 크기모수 α 가 커질 수록 추정량이 진실된 값으로 접근해 가는 것으로 나타났다. 이는 이변량 발견함수를 사용하여 두 모형을 결합하는 경우에는 집락의 크기를

<표 2> $q=2$ 인 경우 시나리오별 모의실험 결과

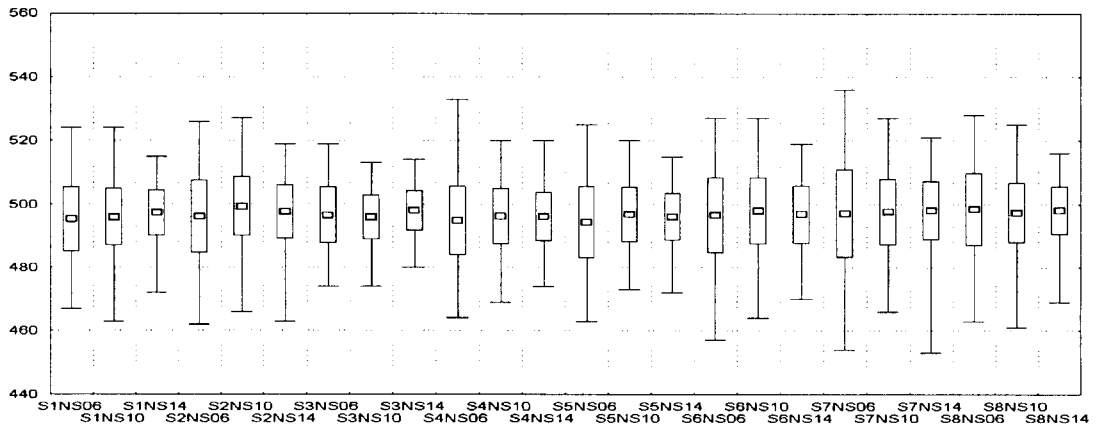
	시나리오	평균		상대편의(%)		상대비율 표준오차(%)		RMSE	
		\hat{N}_c	\hat{N}	\hat{N}_c	\hat{N}	\hat{N}_c	\hat{N}	\hat{N}_c	\hat{N}
$\alpha = 0.6$	1	235.5	495.2	-0.65	-0.97	0.074	0.091	4.19	11.27
	2	235.4	496.0	-0.67	-0.81	0.082	0.102	4.65	12.11
	3	236.3	496.4	-0.30	-0.72	0.064	0.079	3.51	9.57
	4	235.3	494.6	-0.74	-1.07	0.078	0.098	4.51	12.16
	5	235.3	494.3	-0.70	-1.15	0.073	0.101	4.21	12.71
	6	235.9	496.4	-0.46	-0.72	0.778	0.107	4.26	12.51
	7	236.1	497.0	-0.37	-0.59	0.085	0.124	4.58	14.24
	8	236.9	498.3	-0.03	-0.34	0.080	0.103	4.22	11.61
$\alpha = 1.0$	1	236.1	495.7	-0.37	-0.85	0.068	0.080	3.73	9.99
	2	237.2	499.1	0.08	-0.17	0.071	0.084	3.77	9.41
	3	236.3	495.7	-0.31	-0.85	0.052	0.062	2.85	8.13
	4	236.1	496.0	-0.39	-0.79	0.062	0.079	3.33	9.69
	5	236.7	496.7	-0.11	-0.66	0.058	0.078	3.08	9.35
	6	236.8	497.8	-0.08	-0.43	0.073	0.093	3.88	10.65
	7	236.5	497.5	-0.21	-0.49	0.073	0.093	3.89	10.70
	8	236.6	497.2	-0.16	-0.57	0.069	0.085	3.70	9.89
$\alpha = 1.4$	1	236.9	497.2	-0.04	-0.57	0.058	0.065	3.10	7.82
	2	236.6	497.5	-0.17	-0.49	0.063	0.077	3.37	8.93
	3	237.4	497.9	0.15	-0.43	0.048	0.056	2.57	6.64
	4	236.4	496.0	-0.32	-0.81	0.063	0.070	3.42	8.78
	5	236.4	495.9	-0.26	-0.81	0.060	0.114	3.22	8.51
	6	236.7	496.7	-0.14	-0.65	0.066	0.081	3.54	9.66
	7	236.8	497.9	-0.10	-0.43	0.066	0.082	3.51	9.46
	8	237.3	498.0	0.11	-0.40	0.059	0.069	3.14	7.97

<표 3> $q=3$ 인 경우 시나리오별 모의실험 결과

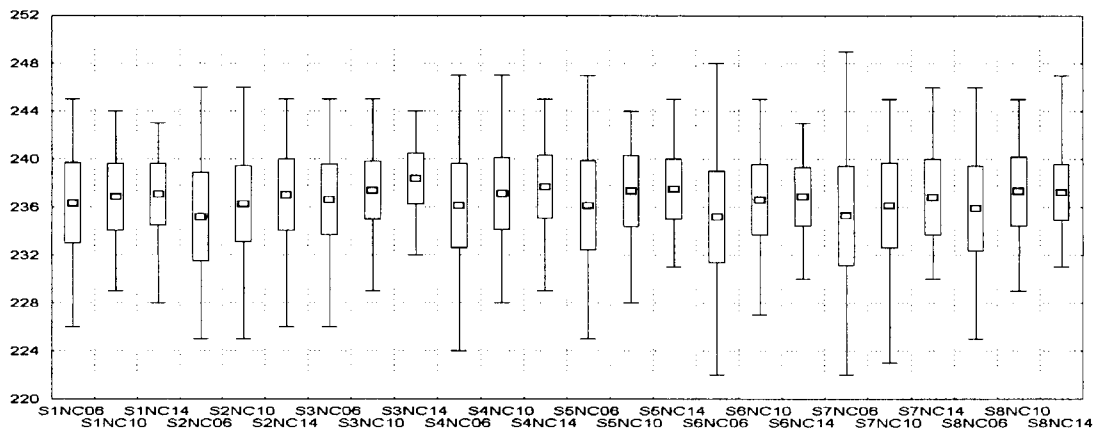
	시나리오	평균		상대편의(%)		상대비율 표준오차(%)		RMSE	
		\hat{N}_c	\hat{N}	\hat{N}_c	\hat{N}	\hat{N}_c	\hat{N}	\hat{N}_c	\hat{N}
$\alpha = 0.6$	1	236.3	496.0	-0.30	-0.80	0.063	0.079	3.42	9.67
	2	235.1	494.6	-0.78	-1.07	0.070	0.091	4.16	11.51
	3	236.6	496.4	-0.19	-0.72	0.056	0.067	2.97	8.31
	4	236.1	495.0	-0.39	-0.99	0.066	0.083	3.64	10.56
	5	236.1	495.8	-0.38	-0.84	0.070	0.086	3.86	10.52
	6	235.1	493.6	-0.79	-1.29	0.072	0.100	4.27	12.94
	7	235.2	495.4	-0.74	-0.92	0.079	0.110	4.54	13.11
	8	235.9	495.9	-0.48	-0.82	0.067	0.090	3.74	10.95
$\alpha = 1.0$	1	236.8	497.0	-0.08	-0.59	0.053	0.062	2.82	7.52
	2	236.2	496.9	-0.33	-0.63	0.060	0.0715	3.29	8.59
	3	237.3	497.9	0.14	-0.42	0.046	0.051	2.45	6.14
	4	237.1	497.6	0.03	-0.49	0.057	0.067	3.03	7.83
	5	237.3	497.9	0.13	-0.43	0.056	0.065	3.00	7.58
	6	236.6	496.6	-0.18	-0.67	0.056	0.072	3.33	8.73
	7	236.1	496.7	-0.38	-0.66	0.067	0.085	3.66	10.07
	8	237.3	497.6	0.12	-0.48	0.055	0.063	2.92	7.47
$\alpha = 1.4$	1	237.0	496.6	0.00	-0.67	0.049	0.056	2.29	7.10
	2	237.0	498.3	-0.01	-0.34	0.056	0.064	2.99	7.46
	3	238.3	499.4	0.57	-0.11	0.040	0.045	2.55	5.01
	4	236.1	495.3	-0.39	-0.99	0.050	0.060	2.71	6.79
	5	237.5	497.7	0.20	-0.46	0.048	0.056	2.57	6.63
	6	236.8	496.7	-0.08	-0.66	0.047	0.058	2.49	7.26
	7	236.8	498.1	-0.09	-0.38	0.060	0.068	3.20	7.83
	8	237.2	497.2	0.08	-0.56	0.044	0.050	2.34	6.31



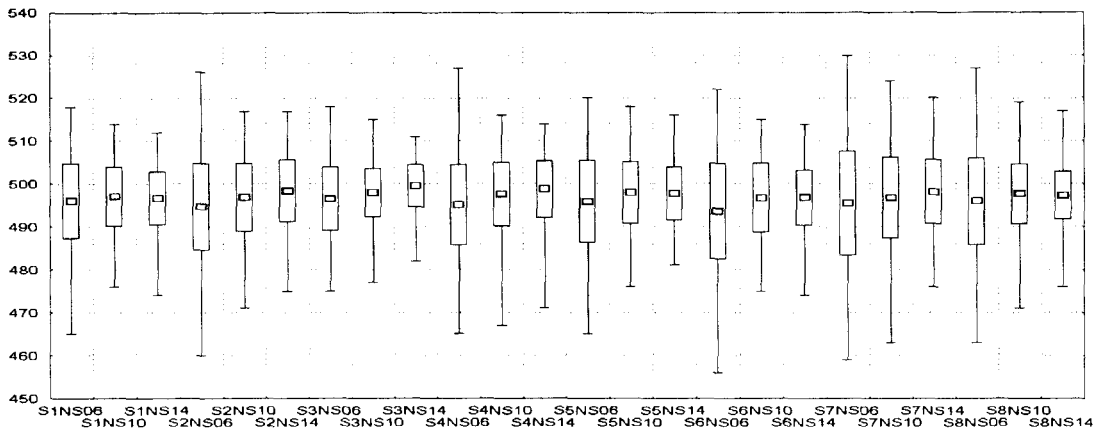
<그림 1> $q=2$ 인 경우 집락 수 추정량의 상자그림



<그림 2> $q=2$ 인 경우 개체 수 추정량의 상자그림



<그림 3> $q=3$ 인 경우 집락 수 추정량의 상자그림



<그림 4> $q=3$ 인 경우 개체 수 추정량의 상자그림

나타내는 α 에 따라 모집단 크기 추정량이 상당한 영향을 받는다는 것으로 향후 집락의 크기를 보다 정확히 추정하는 추정량에 대한 연구가 필요하다.

본 논문에서 제시한 추정 방법은 특히 항공조사(aerial survey)에서 여러 명의 관측자가 있는 경우에 선형단 방법과 포획·재포획 모형을 결합하여 모집단 크기를 추정하면 로버스트한 표본추출 방법이 된다. 그 이유는 동물들은 자기 등지에 있을 수도 있고, 나무그늘 아래에서 휴식을 취하고 있을 수도 있기 때문에 모든 동물들을 관측하기가 힘들 것이다. 일반적으로 동물을 발견하지 못하는 경우가 있어 모집단 크기를 과소추정하게 되므로 중심선 위의 모든 동물들이 공중에서 관측된다는 $g(0)=1$ 의 가정은 비현실적이므로 매우 효율적이다.

발견확률이 집락크기와 거리에 영향을 받는 경우에 집락의 크기를 추정하기 위하여 집락의 평균을 사용하였으나, 모집단이 집락을 이루고 있는 경우에는 모집단의 밀도 또는 크기를 추정하는데 보다 효율적인 방법이다. 모의실험에서는 이변량 발견함수의 모형을 Pollock의 EPS모형만을 사용하였으나 향후 Buckland 등(1993)이 제시한 보다 현실적인 발견함수 모형

$$\{\text{주함수}(x) \times [1 + \text{조정함수}(x)]\} \{ \text{key}(x) \times [1 + \text{series}(x)] \}$$

을 고려하여 비모수적 접근방법에 대한 연구와, 지역 내에 분포되어 있는 개체 또는 집락의 공간 분포를 추정하는데 이용할 수 있다.

참 고 문 헌

[1] Alpízar-Jara, R. (1997). *Assessing Assumption Violations in Line Transect Sampling*. Ph. D. Thesis. North Carolina State University, Raleigh, NC.

- [2] Best, P. B., and Butterworth, D. S.(1980) Report of the Southern Hemisphere Minke Whale assessment cruise, 1978/79., *Report of the International Whaling Commission*, 30, 257-83
- [3] Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. (1993). *Distance sampling : Estimating abundance of biological populations*. Chapman and Hall, London
- [4] Burnham, K. P., D. R. Anderson, and J. L. Laake. (1980). Estimation of density from line transect sampling of biological population. *Wildlife Monographs*, 72.
- [5] Drummer, T. D. and Mcdonald, L. L.(1987). Size bias in line transect sampling. *Biometrics*, 43, 13-21
- [6] Otis, D. L., K. P. Burnham, G.C. White, and D.R. Andersen. (1978). Statistical inference for capture data on closed animal populations. *Wildlife Monographs*. 62. 135pp.
- [7] Otto, M. C., and K. H. Pollock. (1990). Size bias line transect sampling : A field test. *Biometrics*, 46, 239-245.
- [8] Patil, G. P., and J. K. Ord. (1976). On size-biased sampling and related from invariant weighted distributions. *Sankhya B*, 38, 48-46
- [9] Patil, G. P., and C. R. Rao. (1978). Weighted distribution and size-biased sampling with applications to wildlife population and human families. *Biometrics*, 34, 179-89
- [10] Pollock, K. H. (1978). A family of density estimators for line transect sampling. *Biometrics*, 34, 475-478
- [11] Quinn, T. J. II. (1979). The effect of school structure on line transect estimators of abundance, in *Contemporary Quantitative Ecology and Related Econometrics*(eds G.P. Patil and M.L. Rosenzweig) International Co-operative Publishing House, Fairland, MD, USA
- [12] Quinn, T. J. II. (1985). Line transect estimators for school populations. *Fisheries Research*, 3, 189-99
- [13] Quang, P. X. (1991). A nonparametric approach to size-biased line transect sampling. *Biometrics*, 47, 269-279
- [14] Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters*, (2nd ed.), London : Griffin.
- [15] Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics* 42, 267-292
- [16] Seber, G. A. F. (1992). A review of estimating animal abundance II. *International Statistical Review*, 60, 2, pp. 129-166.
- [17] Thompson, S. K. (1992). *Sampling*. Wiley, NewYork.
- [18] White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. (1982). *Capture-recapture and Removal Methods for Sampling Closed Populations*. Los Alamos National Laboratory, LA 8787-NERP, Los Alamos, New Mexico.