

# 수의역학적 자료(獸醫疫學的 資料)에 대한 통계적 기술(統計的 記述)

조준형\* · 강영배\* · 이우용\*

## 제1절 개요(概要)

어떠한 질병의 존재(存在)와 본태(本態) 그리고 강도(強度)에 관한 증명(證明)은 여러가지 종류의 통계적자료(通計的 資料: statistical data)를 포함하게 된다. 이러한 자료들은 이환동물의 수(罹患動物數), 동물의 표본자료에 대한 물리적 계측(物理的 計測), 질병의 존재로 인한 한가지 이상의 생물학적 변수(生物學的 變數)에 대한 계측(計測) 그리고 기타 등의 형태로 나타날 수 있다.

질병에 대한 어떠한 보고서든 최소한도 통계적 증명(通計的 證明: statistical evidence)에 대한 기술적 제시(技術的 提示: descriptive presentation)를 포함하지 않으면 아니된다.

자료들을 나열하고 요약하는데 있어서는 통상적으로 사용되는 몇가지의 기본적인 방법과 계측 요령이 있다. 사용하고자 하는 기법의 선택은 주로 개재되어 있는 자료의 종류에 의존된다. 자료들은 두가지의 주요범주에 들어가게 되는데 즉, 불연속적 자료(不連續的 資料: categorical or discrete data)와 연속적 자료(連續的 資料: continuous or numerical data)자료이다.

불연속적 자료라는 것은 범주가 구별될 수 있는 자료로서 흔히 계정(計定: counts)의 형태를

취한다. 역학에서 볼 수 있는 불연속적 자료들은 대생적 자료(對生的 資料: dichotomous data)의 형태를 취할 수도 있는데 이러한 자료들은 예를들어 이병(罹病: diseased) 또는 비이병(非罹病: non-diseased), 감염(感染: infected) 또는 비감염(非感染: non-infected)과 같이 단지 두가지 값만을 갖는다.

연속적 자료라는 것은 그들이 비록 한정된 범주(限定範疇: defined categories)로 분류될 수 있다 하더라도 무한정적으로 재 분류되어 질수 있는 이론적 가능성을 갖고 있는 계측치(計測值: measurements)로 구성되어 있는 것이다. 예를들면 닭의 체중은 1.45 kg, 1.453 kg, 1.45327856 kg 등일수가 있는 것이다.

## 제2절 표(表)와 그림표(圖)

표 1은 하루에 대략 4,000마리의 닭이 팔리고 있는 큰 시장에서 무작위로 선택한 150마리의 닭의 생체중(生體重: liveweights)을 나타낸다.

이 형태에 있어서는 나열된 이들 숫자를 이해하기가 쉽지 않다. 어떻게 하면 우리가 그 숫자를 좀 더 알기 쉽게 만들 수 있을까? 아마도 우리들 대부분이 생각할 수 있는 첫번째 것은 이들 모든 수치를 더하여 150으로 나누어 평균(平均: mean) 즉, 표본집단 평균치(標本集團 平均值: sample average)를 구하는 것이다. 이렇게 함

\* 가축위생연구소

으로써 그 표본집단에서의 닭의 평균체중은 1.3824kg임을 알 수 있다. 이 숫자는 어떻게 유용한가? 그것만으로는 별로 유용하지 못하다. 예를들어 우리들은 “대부분의 닭의 체중은 1.38kg이다”라는 결론을 이끌어 낼 수는 없다.

가장 가벼운 닭의 무게는 1.01kg이고 가장 무거운 닭의 무게는 1.94kg이라는 정보를 추가함으로써 우리는 평균체중은 1.3824kg이며 그 표본의 범위(範圍 : range)는 0.93kg(1.94~1.01)이라고 말할 수 있을 것이다. 그러나 이러한 사항은 체중이 그 범위내에 균등히 분포되어 있을 가능성을 나타내는 것은 아니며 실제로는 약 반수의 닭은 범위의 하단(下端 : the low end)에 그리고 나머지는 상단(上端 : the upper end)에 있는 것이다. 다시 말하면 우리는 그 수치가 범위 전체에 걸쳐 어떻게 분포되어 있는가를 정확하게 알고자 한다. 이렇게 하는 가장 간단한 방법은 빈도표(頻度表 : frequency table)를 작성하는 것이다(표 2. 참조).

(집필자 주(註) : 여기에서는 ‘frequency’를 빈

도(頻度)라고 번역하여 빈도표, 빈도분포 등의 용어를 사용하고 있으나 어떤 책에도 도수(度數)라 번역하여 도수표, 도수분포 등으로 사용하기도 한다.)

제3열(第3列 : column 3)에 표시된 상대빈도(相對頻度 : relative frequencies)는 매 간격의 수치의 마리수를 표본집단의 닭의 전체 마리수로 나누어 백분율(百分率 : percentage)로 환산한 것이다. 예를들면 첫번째 간격의 상대빈도는;

$$(10/150) \times 100 = 6.7\% \text{이다.}$$

상대빈도의 열(列 : column)을 보아 내려가면 표본 닭의 17.3%에 해당하는 숫자가 체중 1.40과 1.49kg 사이에 있으며 반수이상(57.3%)의 닭이 1.20내지 1.49kg의 무게를 갖고 있는 것을 알 수 있다.

표에 있는 누적빈도(累積頻度 : cumulative frequencies)와 상대적 누적빈도(相對的 累積頻度 : relative cumulative frequencies)는 극한치(極限值 : the extremes) 또는 분포집단(分布集團)의 양단(兩端) (tails of the distribution)에 대

표 1. 시장에서 팔리는 닭 150마리의 표본집단에 대한 체중(kg)

1.40	1.09	1.74	1.48	1.82	1.09	1.52	1.41	1.83	1.22
1.34	1.68	1.25	1.65	1.14	1.33	1.06	1.71	1.17	1.51
1.36	1.34	1.03	1.24	1.06	1.12	1.15	1.57	1.38	1.40
1.39	1.31	1.50	1.10	1.45	1.34	1.38	1.35	1.49	1.58
1.25	1.42	1.64	1.57	1.53	1.18	1.39	1.34	1.13	1.23
1.17	1.88	1.30	1.27	1.01	1.63	1.47	1.23	1.48	1.48
1.37	1.42	1.22	1.47	1.31	1.05	1.61	1.41	1.17	1.45
1.43	1.22	1.40	1.14	1.53	1.25	1.02	1.30	1.35	1.37
1.69	1.37	1.11	1.30	1.05	1.19	1.36	1.63	1.44	1.29
1.35	1.59	1.94	1.51	1.78	1.37	1.11	1.38	1.53	1.44
1.47	1.39	1.55	1.76	1.43	1.37	1.67	1.36	1.31	1.41
1.36	1.26	1.17	1.15	1.79	1.46	1.35	1.29	1.50	1.26
1.36	1.41	1.36	1.32	1.08	1.28	1.33	1.29	1.42	1.50
1.32	1.39	1.20	1.68	1.20	1.35	1.56	1.57	1.37	1.27
1.25	1.38	1.56	1.60	1.74	1.40	1.11	1.60	1.21	1.44

표 2. 닭 150마리에 대한 개체별 체중(體重)의 빈도표(頻度表)

Grouped interval of chicken weights (kg)	Frequency(a)	Relative frequency (%)	Cumulative frequency(b)	Relative cumulative frequency (%)
1.00~1.09	10	(6.7)	10	(6.7)
1.10~1.19	16	(10.7)	26	(17.3)
1.20~1.29	21	(14.0)	47	(31.3)
1.30~1.39	39	(26.0)	86	(57.3)
1.40~1.49	26	(17.3)	112	(74.7)
1.50~1.59	17	(11.3)	129	(86.0)
1.60~1.69	11	(7.3)	140	(93.3)
1.70~1.79	6	(4.0)	146	(97.3)
1.80~1.89	3	(2.0)	149	(99.3)
1.90~1.99	1	(0.7)	150	(100.0)

a Number of values in each interval.

b Cumulative number of values up to the end of a particular interval.

한 질문응답에 유용하다. 예를들면 표본집단의 닭 중 17.3%에 해당하는 숫자는 체중 1.20kg 이하이며 14%(100~86)에 해당하는 수는 체중이 적어도 1.60kg이다.

표 2에 있어서의 정보는 또한 그림표로 나타낼 수도 있다. 빈도표는 종종 막대 그림표(histograms)라는 특별한 형태의 그림표로 표기되기도 한다(그림표기 생략).

막대 그림표에서 매 구획의 면적(每區劃面積 : area of each block)은 해당 간격의 상대빈도

와 비례한다. 다만 이 경우와 같이 급간의 간격(級間間隔 : class intervals)이 모두 같을 때에는 매 구획의 높이가 빈도에 비례한다.

킬로그램(kg) 단위에서 100분의 1에 가장 근사하게 측정하면 닭의 체중은 1.01부터 1.94kg의 범위에 놓이게 되는데 그 범위내에는 94개의 가능한 수치가 있다. 만일 우리가 그램(g) 단위에 가장 가깝게 체중을 측정한다면 그 범위내에는 940개의 가능한 수치가 있다.

(집필자 주(註) : 여기에서 범위(範圍 : range)

표 3. 153마리의 암소군에 있어서의 서로 다른 분만빈도

	Parturition number				
	0	1	2	3	4
Number of cows	26	38	47	24	18
Relative frequency(%)	0.17	0.25	0.31	0.16	0.12
Cumulative relative frequency(%)	0.17	0.42	0.73	0.88	1.00

라 하는 것은 어떠한 집단(集團 : population) 특히 표본집단(標本集團 : sample population)에 대한 모든 계측치(계측치 : measurements)중 최소치(最小值 : minimum value)와 최대치(最大值 : maximum value) 사이의 한계 된 구간을 의미한다.)

표 2에서 보는바와 같은 빈도표를 만들기 위해서는 계측치(計測值 : measurement)의 척도(尺度 : scale)에 따른 간격(間隔 : intervals)이라고 정의되는 여러개의 등급(等級 : classes)으로 나누어야 할 필요가 있다.

153마리의 암소군에 있어서의 서로 다른 분만 횟수(分娩回數 : different parturitions)를 나타내주고 있는 표 3은 하나의 실례이다.

이 자료에 대한 막대 그림표를 그리고자 하는 것은 의미가 없다. 비록 이 경우에는 표 자체가 자료를 표현하는 분명한 방법이긴 하지만 도표에 의한 또다른 가능한 표현방법이 아래와 같이 제시될 것이다.

평균 분만횟수(平均 分娩回數 : mean number of parturitions)를 산출하는데에 있어 그 자료를 이용할 수 있는데 ;

$$[(26 \times 0) + (38 \times 1) + (47 \times 2) + (24 \times 3) + (18 \times 4)] / 153 = 1.80$$

그러나 이것은 서로 다른 두개의 집단에 대하여 비교하기를 원하지 않은 한 유용한 정보자료는 아닌 것이다. 비록 그런 때라고 하더라도 두개의 집단에 대한 완전한 분만자료를 제시해 주는 것이 더욱 좋다.

### 제 3 절 막대 도표와 원 도표 (Bar and Pie Charts)

단지 두 가지의 가능한 수치만을 가지는 범주적 자료(範疇的資料 : categorical data)자료는 종종 불연속(不連續 : dichotomous) 자료라고 하며 우리는 주로 각각의 범주에 속하는 비율에 관심을 갖게 될 것이다.

범주적 변수(範疇的變數 : categorical var-

iables)에 대한 수치적 부호의 사용은 때로는 혼돈을 일으키지만 그러나 그것은 범주적 상태(範疇的 狀態 : categorical status)의 범주적 변수를 박탈하지는 않는다. 중요한 의문사항은 수치적 부호가 통상적인 면에 있어서의 숫자로서 여전히 해석되는지의 여부이다.

이러한 사항은 다음의 예에서 명확해 진다. 닭에 있어서 폐사를 일으키는 일반적인 세가지 원인은 살모넬라균증, 콕시디움증 그리고 뉴캐슬병이며 폐사된 닭 59마리 집단에 있어서의 폐사 발생 빈도가 표 4에 나타나 있다.

표에서 보는바와 같이 자료보존의 편의를 위하여 각 변수에 1, 2, 3 및 4의 색인번호(索引番號 : code numbers)를 주었다. 그러나 이 숫자들은 통상적으로 사용하는 수치를 나타내는 것은 아닌 것이다. 예를들면 2번(콕시디움증)은 1번(살모넬라증)보다 크다고 할 수는 없는 것 등이다. 그들은 단지 본래부호에 대한 단순한 번역인 셈이다. 그러므로 이들 색인번호 숫자에 대한 평균을 구하려고 하는 것은 어리석은 일 일 뿐이며 우리가 할 수 있는 최선의 것은 빈도표 또는 백분율표를 만드는 일이다.

표 4. 59마리의 닭 표본집단에 있어서의 폐사원인 빈도

Cause	Code	No. of deaths	Relative frequency (%)
Salmonellosis	(1)	12	0.20
Coccidiosis	(2)	7	0.12
Newcastle disease	(3)	30	0.51
Other	(4)	10	0.17

지적인 바와같이 막대 그림표(histogram)는 표 3에 있는 자료를 표시하는 방법으로 적합하지 않으며 이는 표 4에 대해서도 마찬가지로 적용된다. 이들 표에 있는 수의 자료들은 막대 도표(bar chart)나 원 도표(pie chart)에 그림으로 나타낼 수 있다. 그림 2는 표 3에 주어진 서로 다른

여러가지 분만횟수의 상대빈도를 나타내는 막대도표이다.

막대도표(bar chart)와 막대그림표(histog-ram)간의 차이에 주목하면 막대도표에서는 연결된 막대사이에 간격이 있어야 하며 이는 그 자료가 실제로 수평축에 표시된 불연속적인 수치만을 취하는 것을 강조하기 위해서이며 모든 막대(bar)는 정확히 똑같은 넓이로서 축 위에 중심을 잡으며 상대빈도에 비례하는 높이를 갖고 있다.

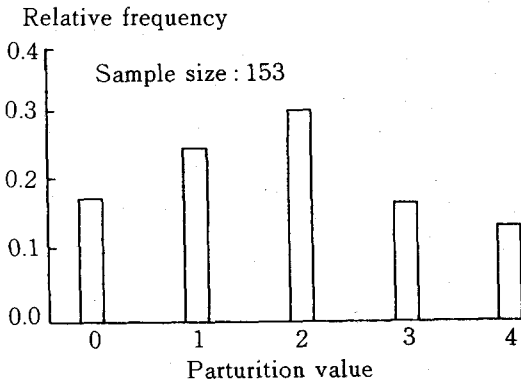


그림 2. 표 3으로부터 얻어진 분만자료에 대한 막대도표

닭의 병리에 대한 자료(표 4) 역시 막대도표(그림 3)로 나타내 보일 수 있다. 그러나 각각 짝을 이루는 수치(parity values)가 통상적이고 자연적인 순서를 갖는 그림 3에 있어서와는 달리 그림 3에서는 각각의 수치(예를들면 질병들)의 순서는 수평축에서 제멋대로이다. 자연적인 순서가 있을때는 그 순서가 고수되어야만 하며 자료가 범주적인 경우에는 순서가 임의로 선택될 수 있음을 기억하도록 한다.

그림 4에서와 같이 범주적 자료를 빈도의 감소 순서로 표시하는 것이 종종 도움이 될 수 있다.

순전히 범주적인(categorical) 자료에 대해서는 원도표(pie chart)가 막대도표(bar chart)에 대하여 통상 대체 사용되는 일반적 방법의 하나이다. 원도표는 범주(categories)의 갯수 만큼의 많은 부분(部分: sector)로 나뉘어진 원이다. 각 부분의 면적은 원의 중심에서 부분이 이루는

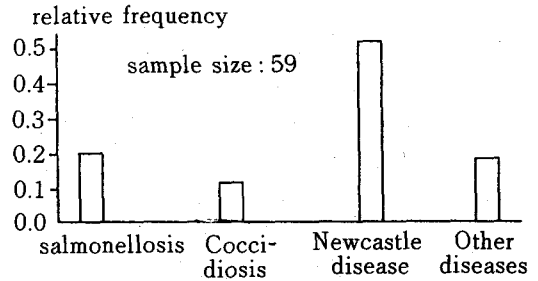


그림 3. 표 4로부터 얻어진 닭 폐사원인 자료에 대한 막대도표

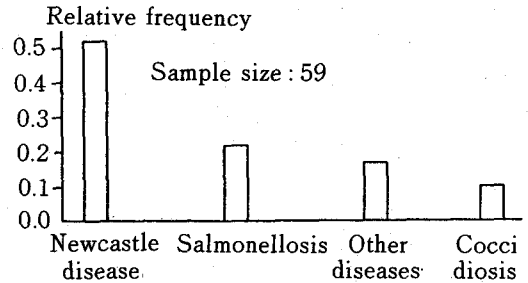


그림 4. 표 4의 자료를 대체한 막대도표

각도를 측정할 것과 상응하는 범주의 상대빈도에 비례한다. 모든 각도의 총계는 360도 이므로 상응하는 면적을 얻기 위해서는 여러가지 범주에서 정확한 비율로 360도를 나눌 필요가 있다.

예를들면 표 4로부터 살모넬라균증의 상대빈도는 0.20임을 알 수 있다. 그것에 상응하는 각도는;

$$360 \times 0.20 = 72 \text{이다.}$$

마찬가지로 콕시디움증과 뉴캐슬병에 상응하는 각도들은 각각 43도와 184도이다. 이러한 결과로 얻어진 원도표는 그림 6에 나타나 있다.

막대그림표, 원, 도표 그리고 막대도표에서는 표본의 크기가 항상 인용되어야 한다는 사실에 유의하도록 한다.

#### 제 4 절 변수(變數)에 의한 분류(分類)

지금까지 고찰된 모든 예(例: examples)는 어떤 동물의 단순집단(動物單純集團: a single population of animals)내에 있어서의 하나의 단

순변수(單純變數 : single variable)에 대한 관찰(觀察 : observations)만을 포함하였다. 그러나 그들 사이에 있어서의 가능한 차이를 조사하기 위해서는 그 집단을 몇개의 작은 집단(subgroup)으로 다시 나누어야 한다. 예를들면 소(牛 : cattle)는 성별(性別 : sex), 품종(品種 : breed), 지리적 위치(地理的 位置 : geographic location), 질병상태(疾病狀態 : disease status) 등으로 나눌 수 있다.

역학적조사연구(疫學的調查研究 : epidemiological investigations)에서는 분류가능변수(分類可能變數 : classificatory variables)는 범주적(categorical)인 것이며 통상적으로 요인(要因 : factors) 또는 결정요소(決定要素 : determinants)라고 인용된다.

만일 변수의 숫자가 적다면 실제의 수치적인 변수(實際 數值的變數 : true numerical variables)는 또한 변수의 수치 형태이거나 또는 급간격의 형태로서 분류를 위한 요인(分類要因 : classifying factors)으로 사용될 수 있다.

예를들면 표 3에서 제공되는 자료에 있는 각 동물은 분만횟수에 의해 분류될 수 있으므로 그 표본집단은 5개의 군으로 분류될 수 있으며 표 1

에 주어진 바와 같은 닭의 체중량 자료는 “생체중(生體重 : liveweights)”이라는 요인(要因 : factor)의 서로 다른 수준을 한정하기 위하여 표 2의 급간격(級間隔 : class intervals)을 사용하여 10개의 다른 체중군으로 분류될 수 있다.

요인(要因 : factors)의 선택과 각 요인별 수는 연구대상이 되는 집단에 대한 사전지식(事前知識 : prior knowledge)의 정도, 요인들에 대하여 기대되는 과학적 유의성(科學的 有意性 : scientific significance) 그리고 연구자에게 사용 가능한 방법(方法 : measures)에 따라 결정된다. 표 5는 돼지에 있어서의 회충감염(蛔蟲感染 : ascaris infections)에 관한 계수를, 3가지 요인 즉, 사양관리형태(飼養管理形態 : management system)-2개수준 : 축사내 사육(畜舍內 飼育 : raised indoors) 또는 축사의 사육(畜舍外 飼育 : raised outdoors), 돼지 개체별 분변에 있어서의 회충란 출현여부(蛔蟲卵 出現與否 : occurrence of ascaris eggs)-2개수준 : 있음(有 : present) 또는 없음(無 : absent) 그리고 도살후 개체별 돼지의 간에 있어서 관찰되는 백색병변의 정도(白色病變 程度 : degree of whitespot)-3개수준 : 없음(無 : absent), 경증(經症 : slight)

표 5. 돼지에 있어서 회충감염 실태조사표 : 합계를 함께 표시한 3요인표의 보기

Whitespot	Ascaris eggs	Management system		Any system
		Indoors	Outdoors	
Absent	Absent	503 *	112 *	615
	Present	141 *	38 *	179
	Total	644	150	794
Slight	Absent	231 *	75 *	306
	Present	87 *	30 *	117
	Total	318	105	423
Severe	Absent	79 *	32 *	111
	Present	71 *	17 *	88
	Total	150	49	199
Any whitespot condition	Absent	813	219	1032
	Present	299	85	384
	Total	1112	304	1416

\* Recorded data.

또는 중증(重症 : severe) 등을 나타내기 위하여 고안된 것이다.

예를들면 표 5에 있어서 엄밀히 말하자면 별표(asterisk)로 표시된 수치만이 꼭 필요한 것일 뿐이며 나머지 수치들(36개중 24개)은 보조적인 정보를 제공해 주는 것이다. 합계(合計 : marginal totals)의 사용여부는 개인적 판단에 따르는 문제인데 : 일반적으로 단일 완전한 표(complete table)가 요점을 명확히 하기보다 혼돈시킬 수 있다고 생각되는 경우에는 합계(合計 : totals)를 생략하는 것이 좋다.

표 6은 표 5로부터 유도될 수 있는 2요인 표(2-要因表 : twofactor tables) 중의 하나를 나타내고 있다.

복요인 표(複要因表 : multi-factor tables)에 있어서는 상대빈도(相對頻度 : relative frequencies)를 나타내기 위한 몇가지 선택사항(選擇事項 : option)이 언제나 존재한다. 예를들면 표 6에 있어서는 상이한 백색병변의 상대빈도가 회충란 이라는 요인의 각 수준에 주어져 있다.

반대로 회충란 수준의 빈도는 백색병변의 심한 강도(強度 : severity)에 있어서의 전 두수에 대한 상대치로 주어질 수 있거나 또는 백색병변-회충란 이라는 6개의 가능한 조합군에 있어서의 각각의 빈도가 표본집단내에 있는 폐지 총 두수에 의하여 상대적으로 산출될 수 있을 것이다.

선택사항(選擇事項 : option)은 만들고자 하는 요점에 따라 다르나 상대빈도가 명확하게 나타나도록 작성하여야 한다. 다른 연구자에 의하

표 6. 표 5로부터 유도된 2요인 표(2-要因表)

Whitespot	Ascaris eggs		Total
	Absent	Present	
Absent	615(57) <sup>a</sup>	179(47)	794(56)
Slight	306(30)	117(30)	423(30)
Severe	111(11)	88(23)	199(14)
Total	1032(100)	384(100)	1415(100)

a Figures in parentheses give the relative frequencies(%) of whitespot conditions.

여 제시된 표를 해석하는데 있어서는 표시되거나 논의되고 있는 상대빈도를 명확히 하도록 주의하여야 한다.

### 제 5 절 집단(集團)내 질병발생상황(疾病發生狀況)의 정량(定量)

집단에 있어서의 질병발생상황을 정량하기 위하여 사용되는 자료는 종종 대생적(對生的 : dichotomous)인데 즉, 동물은 병인체에 감염되어 있거나 또는 감염되어 있지 않은 것 중의 한가지이다. 그러한 자료는 흔히 역학적인 용어로 “율(率 : rate)”의 형태로 표시된다.

역학에 있어서 비율(比率 : rate)이라고 하는 것은 관찰기간 동안에 어떤 특별한 성상 또는 특성(特性 : particular characteristic) 즉, 일반적으로 감염, 질병 또는 질병과 관계있는 특성을 가지고 있거나 획득할 수 있는 위험성에 놓여 있는 총 개체수로 나눈 것으로 정의될 수 있다.

그러한 표현(表現 : expression)은 요인(要因 : factor)에 대한 배수로 표시되는데 집단의 특정 구성단위에 관련을 가지며 일반적으로 10의 배수(倍數 : multiple)로 표시된다.

비율(比率 : rate)은 일반적으로 십진법(十進法 : decimals), 백분율(百分率 : percentage) 또는 1000, 10,000과 같은 집단의 표준단위당 발생건수(發生件數 : events)로 표시된다. 이것은 질병발생에 대한 표준화된 측정방법이며 따라서 집단간 또는 집단내에 있어서 어느 시기에 대한

질병발생빈도의 비교를 가능케 하는 것이다.

비율(比率 : rate)에 있어서는 분자(分子 : numerator)가 항상 분모(分母 : denominator)에 포함되어 있으나 비(比 : ratio)에 있어서는 포함되어 있지 않음을 주목하도록 하여야 한다. 역학적인 “비율”(疫學的 “比率” : epidemiological “rate”)에 있어서는 항상 관찰기간이 정해져 있어야 한다.

잡단간 또는 집단내의 질병발생에 대한 정확한 비교는 분모(分母 : denominator)가 산출되지 않는 한 어렵다. 비교하기 위하여 현수분자(懸垂分子 : dangling numerators)를 사용하는 것은 역학연구자(疫學研究者 : epidemiologist)가 범할 수 있는 가장 큰 비도덕적 행위(非道德的 行爲 : crimes)의 하나이므로 가능한 피해야 한다.

예를들어 서로 다른 사양관리 조건하에서 사육한 동일 품종의 소 집단(牛集團 : cattle population)에 있어서 어떤 특정기간 동안에 어떤 특정 병인체에 의하여 감염된 건수를 비교하고 싶다고 가정하자, A군에 있어서는 1983년 6월에 문제의 병인체에 감염된 동물수가 5이었고, B군에 있어서는 같은 시기에 동일한 병인체에 감염된 동물수가 50이었다고 한다. 그러면 우리는 그 질병은 A군에서 보다는 B군에서 더욱 문제가 된 것이라고 잘못된 결론을 내릴 수가 있다.

우리는 각 군에 있어서 그 병인체에 감염될 수 있는 위험성에 노출된 동물수 즉, 모집단(母集團 : population) 또는 분모(分母 : denominator)를 알고 있지 못하였음에 주목해야 한다. 나중에 우리는 그해의 6월 동안에 질병이라는 위험에 노출될 수 있는 동물수가 A군에서는 100이었고 B군에서는 500이었다는 사실을 알게 되었다고 상상해 보자 그러면 각 군에 대한 비율(比率 : rates)은 A군이 25/100 또는 0.25 또는 25% 또는 1,000 두당 100두이다. 그리하여 실제의 상황은 그 질병에 대한 문제점이 A군에서 더 큰 것이었다.

수의역학(獸醫疫學 : veterinary epidemiol-

ogy)에서 사용되는 비율(比率 : rates)의 중요한 2가지 형태는 다음과 같다.

**이환율(罹患率 : morbidity rates)** : 한 집단 내에 있어서의 감염된 개체의 비례수(比例數 : proportion) 또는 감염될 수 있는 한 집단내에 있어서의 한 개체의 위험성을 예측하기 위해 사용된다(the proportion of affected individuals in a population or the risk of an individual in a population of becoming infected).

**폐사율(斃死率 : mortality rates)** : 한 집단내에 있어서의 죽은 동물의 비례(比例 : proportion)를 계측하는 것이다(the proportion of animals dying in a population).

**이환율(罹患率 : morbidity rates)**

이환율(罹患率 : morbidity rates)에는 동태이환율(動態罹患率 : incidence morbidity rate), 침습이환율(浸襲罹患率 : attack morbidity rate), 정태이환율(靜態罹患率 : prevalence morbidity rate) 그리고 비례이환율(比例罹患率 : proportional morbidity rate) 등이 포함된다.

(질필자 주(註) : 여기에서 “罹患率”이라는 용어 대신에 “罹患率” 또는 “疾病率”이라는 용어도 사용되고 있으며 이환율을 구분함에 있어서도 ‘incidence rate’를 “발생율(發生率)”, ‘attack rate’를 “발병율(發病率)”, ‘prevalence rate’를 “유행율(流行率)”로 표기하는 방법도 생각해 볼 수 있으나 우리는 병인체가 감염된 다음 그 질병이 발생되고 유행되는 시점과 시간에 관심을 더 두고자 “動態”와 “靜態”라는 용어를 사용하기로 하였음을 이해하기 바란다).

**동태이환율(動態罹患率 : incidence rate)**이라는 것은 어떤 특정 기간동안에 어떤 특정 집단에서 발생하는 질병에 대한 새로운 증례수 즉, 신환건수(新患件數 : number of new cases of a disease)를 그 특정 기간동안에 있어서의 그 집단 내에 있는 평균 개체수로 나눈 것이다.

예를들어 어느 검역장(檢疫場 : quarantine camp)에서 평균 4,000마리의 소 중에서 600마리



가 6월 한달동안에 우역(牛疫 : rinderpest) 증상을 나타내었다고 가정하자 6월중의 그 검역장(檢疫場 : quarantine camp)에 있어서의 우역의 동태이환율은  $600/4,000=0.05$  또는 15% 또는 1,000마리당 150마리의 새로운 발생이 있었다.

동태이환율은 특정 기간동안에 한 집단의 감수성 있는 한 개체가 어느 질병에 걸리는 위험성을 측정하는 한 방법이다. 그러므로 어떤 감수성 있는 동물이 6월 1일자에 검역장에 도입되었다면 6월말에 우역에 걸릴수 있는 기회는 15%이다. 동태이환율을 산출할 때는 모집단 즉, 분모(分母 : denominator)의 추정치에 흔히 문제점이 발생된다. 출생(出生 : births), 폐사(斃死 : deaths), 판매(販賣 : sales) 그리고 이동(移動 : movements) 등으로 인하여 가족의 수는 좀처럼 일정기간동안 일정한 수준을 유지하지 못하며 모집단에 있어서의 그러한 변동사항(變動事項 : fluctuations)은 동태이환율의 산출에 있어서 모집단 즉 분모(分母 : denominator)를 추정하는 데는 여러가지 방법이 있다. 이 중에는 일반적으로 연구 기간동안에 여러 간격으로 마리수를 측정하여 그 결과를 평균하는 방법이 포함되어 있다.

예를들어 앞에서의 예제에서 6월 초에 4,000마리가 있었고 2주째 말에는 100마리가 질병으로 죽었으며 그 달 말까지에는 300마리가 더 죽었다고 상상해 보자, 새로운 동물이 도입되거나 출생한 동물이 없었다고 가정하면 그 검역장내의 마리수는 관찰기간 시작시기에는 4,000, 중간시기에는 3,900 그리고 월말에는 3,600이 된다. 관찰 시작시기와 종료시기의 마리수를 취하여 그것을 평균해서 분모(分母 : denominator)를 산출할 수 있다. 즉,

$$(4,000+3,600)/2=3,800$$

상응하는 동태이환율(動態罹患率 : incidence rate)은,

$$600/3,800=0.158 \text{ 또는 } 15.8\% \text{가 된다.}$$

또 다른 방법으로 관찰 시작시기, 중간시기 그

리고 종료시기의 마리수를 취하여 그것을 평균하면,

$$(4,000+3,900+3,600)/3=3,833$$

그러하여 이 경우의 동태이환율은  $600/3,833=0.156$  또는 15.6%이다.

분모(分母 : denominator)를 산출하는 방법에 따라 동태이환율의 추정치에 약간의 차이가 발생된 것에 주목한다. 이 때문에 동태이환율을 비교하고자 할 때에는 분모를 산출하는 방법을 특정화시켜 놓아야 하며 전반에 걸쳐 동일한 방법이 적용되도록 하여야 한다. 동태이환율에 있어서 분모를 산출하기가 어렵기 때문에 이환율(罹患率 : morbidity rates)의 또다른 형태인 침습이환율(浸襲罹患率 : attack rate)이 사용되기도 한다.

침습이환율(浸襲罹患率 : attack rate)이라는 것은 어떤 특정기간 동안에 어떤 특정집단에 있어서 발생된 질병의 총 건수를 그 특정기간의 시작시에 있던 그 집단내의 개체의 총수로 나눈 것이다. 그러므로 분모는 관찰기간 전 기간(全期間)에 걸쳐 일정하게 유지된다. 따라서 앞에서의 예제에서 침습이환율은  $600/4,000=15\%$ 이다.

엄밀히 말하여 침습이환율에 대한 정의는 새로운 질병건수만이 아니고 표본집단 즉, 분자(分子 : numerator)내에 포함되는 질병의 총 건수로 되는 것이다. 어쨌든 침습이환율이 질병발생기간 중의 질병의 지행상황을 정량하기 위하여 통상적으로 사용되고 있는 것이다. 대부분의 예에서 문제의 질병이 발생하기 이전에는 질병건수가 없으므로 총 발생건수는 새로운 건수이며 그리하여 침습이환율(浸襲罹患率 : attack rate)은 동태이환율(動態罹患率 : incidence rate)의 변형된 형태(變型形態 : modified form)가 되는 것이므로 때로는 누적 동태이환율(累積動態罹患率 : cumulative incidence rate)이라고도 불리운다.

정태이환율(靜態罹患率 : prevalence rate)이라는 것은 어떤 특정시점에 있어서 어떤 특정집단에서 발생하는 질병의 총 건수를 그 특정시점에 있어서의 그 집단내에 있는 개체의 총수로 나

는 것이다.

예를들어 어떤 검역장내에 있는 4,000마리의 소 집단에 있어서 6월 18일자에 우역이 60건 검 색되었다고 상상해 보자, 그 검역장에 있어서의 6월 18일자 우역의 정태이환율은  $60/4,000=0.015$  또는 1.5% 또는 1,000마리당 15건으로 된다.

정태이환율은 어떤 특정시점에 있어서의 한 집단내에 존재하는 질병의 양(疾病量 : amount of disease)을 가리키는 횡단적 제측법(橫斷的計測法 : cross-sectional measure)임을 유의하여야 하며 따라서 시점 정태이환율(時點 靜態罹患率 : point prevalence)이라는 용어를 사용하기도 한다. 그러나 대규모 집단을 취급함에 있어서는 어떤 특정시점에 모든 개체를 조사할 수는 없으므로 시점 정태이환율을 구하기는 불가능 하다. 그러므로 일반적으로 정태이환율의 제측은 어느 기간에 걸쳐 수행되지 않으면 아니 되는데 이러한 때에는 기간 정태이환율(期間 靜態罹患率 : period prevalence)이라 알려져 있다.

정태이환율을 제측하는 기간은 가능한 한 짧아야 그 변수(變數 : parameter)의 정확도가 좋아진다. 그러나 만일 시간간격이 너무 길게되면 제 측을 시작한 이후 새로운 질병건수가 많이 발생 될 수 있다. 그러한 때에는 그 변수(變數 : parameter)는 시점 정태이환율(時點 靜態罹患率 : point prevalence)과 동태이환율(動態罹患率 : incidence) 등이 혼합된 상태이므로 정확도를 잃게 된다.

동태이환율(動態罹患率 : incidence)과 정태 이환율(靜態罹患率 : prevalence)이라는 용어는 흔히 혼동되거나 잘못 사용되고 있다. 혼돈은 통 상적으로 분모(denominator) 즉, 고려되어야 하는 실제집단을 정확하게 정의하지 못하였기 때문에 발생한다. 이러한 사항은 위험상황에 놓인 그 모 집단에 있어서 고스란히 무시되거나 전혀 고 려되지 못하는 결과를 초래할 수도 있는 것이다.

이러한 예의 것은 수의관청이나 연구소로 부터

의 보고서에서도 볼 수 있는데 그러한 보고서에 는 “동태이환율(incidence)”이라는 용어를, 진단 또는 분리시도된 총 건수에 대한 백분율로 특정 병인체의 진단이나 분리건수를 표현하는데 종종 사용하고 있는 것이다. 이러한 경우에 있어서의 분모(denominator)는 질병에 걸리는 위험성에 놓여 있는 개체의 집단은 아닌 것이며 그렇게 계 산된 율(率 : rate)은 비례이환율(比例罹患率 : proportional morbidity rate)과 유사하다.

비례이환율(比例罹患率 : proportional mortality rate)이라는 것은 어떤 특정기간 동안에 어떤 특정집단에 있어서의 어떤 특정질병 건수를 그 기간중에 그 집단에 있어서의 모든 질병의 총 건수로 나눈 것이다.

예를들어 우폐역(牛肺疫 : contagious bovine pleuropneumonia, CBPP)이 우군에서 발생하였다고 상상해 보자. 6개월의 기간동안에 18건의 우폐역을 포함하여 45건의 질병이 발생하였다. 그 6개월 동안에 그 우군에 있어서의 우폐역의 비례이환율(比例罹患率 : proportional morbidity rate)은  $18/45=0.4$  또는 40% 또는 질병 총 건수 1000건당 우폐역(CBPP) 400건으로 되는 것이다.

폐사율(斃死率 : mortality rates)

가장 통상적으로 사용되고 있는 폐사율(斃死率 : mortality rates)로는 조 폐사율(粗 斃死率 : crude death rate)과 원인-특이 폐사율(原因 特異 斃死率 : cause-specific death rate)이 있다.

(집필자 주(註) : 여기에서 “mortality rates” 또는 “death”를 번역함에 있어서 수의학적인 전문성에 비중을 두고 “폐사율”이라는 용어를 채택하기로 하였다.

사람에 대한 일반 의학적 표현에 있어서는 흔히 “사망율(死亡率 : death rate)”이라고 표현하기도 하며 약리학 또는 독성학, 실험동물학 등에 있어서는 흔히 “치사율(致死率 : fatality @ate or lethality rate)”이라는 의미로도 사용되고 있음을 유의하기 바란다.

**조 폐사율(粗 斃死率 : crude death rate)**이라는 것은 어떤 특정기간동안에 어떤 특정집단에 있어서의 총 폐사 건수를 그 특정기간동안의 그 집단내에 있는 평균 개체수로 나눈 것이다.

조 폐사율에 있어서의 분모는 동태이환율(動態罹患率 : incidence rate)에서와 같은 방법으로 추산될 수 있다. 분모의 산출방법은 항상 정의되어 있어야 하며 유의성 비교(比較 : comparisons)를 가능하게 하기 위해서는 전반에 걸쳐 똑같은 방법이 사용되도록 하여야 한다.

예 : 한 우군에 있어서 일년에 40마리의 폐사가 있었다고 가정하자. 그 우군의 동물수는 연도 초에 400이었으며 연중(mid-year)에 420, 그리고 연말에 390이었다. 그러면 그 우군의 평균크기는,

$$(400+390)/2=395, \text{ 또는}$$

$$(400+420+390)/3=403\text{이다.}$$

따라서 분모를 어떤 방법으로 계산하였나 그 사용방법에 따라 조 폐사율은,

$$40/395=0.101(10.1\%) \text{ 또는}$$

$$40/403=0.099(9.9\%) \text{로 될 수 있는 것이다.}$$

**원인-특이 폐사율(原因 特異 斃死率 : cause-specific death rate)**이라는 것은 쓸모가 있는 폐사율이며 어떤 특정기간 동안에 어떤 특정집단에서 어떤 특정원인에 의하여 발생한 총 폐사수를 그 기간동안에 그 집단내에 있어서의 평균개체수로 나눈 것으로 정의될 수 있다. 분모는 동태이환율(動態罹患率 : incidence rate) 또는 조 폐사율(粗 斃死率 : crude death rate)과 같은 방법으로 계산되며 계산시 똑같은 주의사항이 적용된다.

예 : 앞에서 언급한 우군에서 바베시아병(babesiosis)으로 20마리가 죽었다고 가정해 보면 그 우군에 있어서의 바베시아병(babesiosis)으로 인한 폐사율은,

$$20/395=0.051(5.1\%) \text{ 또는}$$

$$20/403=0.050(5.0\%) \text{로 될 수 있는 것이다.}$$

**기타 유용한 폐사율(其他 有用 斃死率 : other**

**useful mortality rates)**

**비례 폐사율(比例 斃死率 : proportional mortality rate)**이라는 것은 어떤 특정기간동안에 어떤 특정집단에 있어서 어떤 특정질병으로 인하여 발생한 한 총폐사수를 그 기간동안에 그 집단내에 있어서의 총폐사수로 나눈 것이다.

예 : 한 군의 폐사수 40중에서 바베시아병(babesiosis)으로 인한 것이 20이었다면 바베시아병으로 인한 비례폐사율은  $20/40=0.5$  또는 50%이다.

**증례 치사율(症例 致死率 : case fatality)**이라는 것은 어떤 특정기간동안에 어떤 특정집단에 있어서 어떤 특정질병으로 인하여 발생한 폐사수를 그 기간동안에 그 집단내에 있어서의 그 질병에 대한 증례수로 나눈 것이다.

예 : 그 군에서 50건의 바베시아병(babesiosis) 증례가 발생하였다고 가정해 보면 바베시아병으로 인한 증례치사율은  $20/50=0.4$  또는 40%이다.

위에 기술된 비율(比率 : rate)들은 역학적연구에서 가장 많이 사용되는 것들이다. 그밖에 비율(比率 : rates)에 대한 상세한 사항 즉, 산출하는 방법과 가능한 용도(可能用途 : potential uses)에 대하여는 슈바베 등(Schwabe et al, 1977)을 참조하는 것이 좋을 것이다.

**특정(特定 : (specific) 비율(比率 : rates)의 사용(使用 : use)**

역학에 있어서는 질병발생빈도에 대한 결정요소(決定要素 : determinants)의 영향을 연구하는데에 언제나 매달리게 된다. 이러한 사항은 앞에서 언급한 몇가지 율들에 대한 비교를 때때로 포함하게 되는데 즉, 결정요소(決定要素 : determinants)가 부가되거나 소거되기 전후의 기간에 있어서 동일집단내 또는 집단간 또는 결정요소의 발생빈도가 서로 다른 경우 또는 동일시점이나 일정기간에 걸쳐서의 율 등이 포함된다.

그러한 비교가 유효(有效 : valid)하기 위해서는 비교군(比較群 : comparison groups)들은 연

구하고자 하는 특정 결정요소에 있고(有 : presence), 없음(無 : absence) 또는 발생빈도(發生頻度 : frequency of occurrence)가 서로 달라야만 한다. 역학은 통상적으로 통제되지 않은 야외조건하에 있어서의 결정요소에 대한 연구를 포함하기 때문에 이러한 기준(基準 : criteria)을 만족시키기는 극히 어려운 것이다. 그럼에도 불구하고 만일 율이라는 것들이 비교하고자 하는 병인체(病因體 : disease agents)간 또는 숙주집단(宿主集團 : host population)간에 존재할 수 있는 서로 다른 특성을 무시해 버리는 형태로 표현된다면 그러한 율들은 지나치게 단순화(單純化 : over-simplified) 되거나 나아가서는 실제상황(實際狀況 : actual situation)을 허위(虛偽 : false)로 표현하게 될 위험마저 있는 것이다.

여러가지 서로 다른 특성을 고려함으로써 비율(比率 : rates)은 더욱 특이성이 있게 될 수 있으며 그들간의 비교가 더욱 확실성 있게 될 수 있다.

연구하고자 하는 병인체의 아종(亞種 : subspecies)이나 주(株 : strain)를 명확히 정의함으로써 또는 특정된 아종이나 주에 의하여 영향을 받는 그러한 개체들만을 분자 즉, 표본집단에 포함시키므로써 고려될 수 있다. 나이(年齡 : age), 품종(品種 : breed) 및 성별(性別 : sex)에 따른 숙주집단(宿主集團 : host populations)간의 특성의 차이점은 그러한 특성을 고려하여 산출함으로써 표현될 수 있는 것이다.

그리하여 예를들면 우리는 어떤 특정기간동안에 어떤 특정집단에 있어서 어떤 특정연령군(specified age group)에 속한 개체중에서 발생하는 어떤 질병의 새로운 증례수(新患件數 : new cases of a disease)를 그 기간동안에 그 집단에 있어서 그 특정연령군에 속한 평균개체수로 나눈 것으로 정의되는 연령-특이 동태이환율(年齡-特異 動態罹患率 : : age-specific incidence rate)을 계산할 수 있을 것이다.

바꾸어 생각해 보면 우리는 어떤 특정기간동안

에 어떤 특정집단에 있어서 어떤 특정품종에 속한 개체중에서 발생하는 어떤 질병의 새로운 증례수를 그 기간동안에 그 집단에 있어서 그 품종에 속한 평균개체수로 나눈 것으로 정의되는 품종-특이 동태이환율(品種-特異 動態罹患率 : breed-specific incidence rate)을 계산할 수도 있다. 우리는 또한 어떤 특정집단에 있어서 어떤 특정품종과 어떤 특정연령에 속한 개체중에서 발생하는 어떤 질병의 새로운 증례 총 수를 그 기간동안의 그 집단내의 특정연령과 품종에 속한 평균개체수로 나눈 것으로 정의되는 연령-품종-특이 동태이환율(年齡-品種-特異 動態罹患率 : : age-breed-specific incidence rate)까지도 계산할 수 있는 것이다.

동일한 절차가 기타 이환율(罹患率 : morbidity)과 폐사율(斃死率 : mortality)에 대하여도 적용될 수 있다. 표본집단 즉, 분자(分子 : numerator)와 모집단 즉, 분모(分母 : denominator)에 대한 적절한 정의를 사용하는 방법으로 대단히 많은 여러가지의 특정된 비율(比率 : rates)을 산출할 수 있다. 일반적인 원리로 비율(比率 : rates)은 자료가 허용되는 한 특이성(特異性 : specific)있게 만들어져야 하는데 그렇다고 갯수(個數 : numbers)를 너무 적게 하면 통계적 분석을 하기에는 특이성이 없게 된다.

만일 연령-품종 특이율(年齡-品種 特異性 : age-breed specific rate)이 계산될 수 있다면 통계분석을 위해서는 연령 또는 품종에 따른 비율을 계산하고 비교하는 것은 별로 이익이 없거나 아주 없는 것이다.

다음은 비교하는데 있어서 특정된 비율(比率 : rates)을 사용하는 이점을 설명하는 한 예이다. 우리는 질병의 공격수준(攻擊水準 : level of disease challenge), 환경적 조건(環境的 條件 : environmental conditions) 그리고 사양관리체계(飼養管理 體系 : system of management)가 대체로 동일한 조건에 있는 2개소의 동연안열(東沿岸熱 : East Coast fever, ECF) 전염지역

표 7. 2개의 소 집단에 있어서의 동연안열에 대한 품종-특이 동태이환율

Area	Breed	Number of cattle	Number of new cases of ECF	incidence (%)
A	Crossbred	400	97	24.3
	Holstein			
	East African Shorthorned Zebu	9,600	403	4.2
	Total	10,000	500	5.0
B	Crossbred	4,500	1,059	23.5
	Holstein			
	East African Shorthorned Zebu	10,500	441	4.2
	Total	15,000	1,500	10.0

(傳染地域 : endemic areas)에 있어서 진드기 방제계획(-防除計劃 : tick control programme)의 효과를 검정해 보고자 한다고 가정해 보자.

A지역에 있어서는 1개월의 연구기간 동안에 평균 10,000두의 소가 있었으며 그 기간동안에 그 집단중에서 500두가 ECF의 증상을 나타내었다. B지역에 있어서는 연구기간 동안에 평균 15,000두 중에서 1,500두가 증상을 나타내었다.

A지역에 있어서의 그 질병에 대한 조동태이환율(粗動態罹患率 : crude incidence rate)은  $500/10,000=5\%$ 이고, B지역에 있어서는  $1,500/15,000=10\%$ 이다. 그러므로 우리는 A지역에 있어서의 진드기 방제계획이 B지역에 있어서 보다 더 효과적이었다고 결론 지을 수 있다.

우리는 또한 A지역에 있어서의 소 집단(牛集團 : cattle population)은 400두의 교잡종 홀스타인(交雜種 - : crossbred Holstein)과 9,600두의 동아 단각우(東阿短角牛 : East Africa Shorthorned Zebu)로 구성되어 있는데 반하여 B지역에 있어서는 4,500두의 교잡종 홀스타인과 10,500두의 East Africa Shorthorned Zebus로 구성되어 있는 것을 발견하였다.

이제 우리는 표 7에 보이는 바와같이 품종-특이 동태이환율(品種-特異 動態罹患率 : breed specific incidence rate)을 계산해 낼 수 있다.

조동태이환율은 각각 5%와 10%인 반면에 두 지역에 있어서의 동아단각우(東阿短角牛 : East African Shorthorned Zebu)에 대한 품종-특이 동태이환율에는 차이가 없고, 교잡종 홀스타인에 대한율은 A지역보다 B지역에서 더 낮은 것에 주목하라. 두 지역간에 있어서의 조동태이환율의 차이는 보다 더 감수성이 높은 교잡종이 A지역에 있어서는 소 집단의 단지 4%를 차지하고 있는데 반하여 B지역에 있어서는 소 집단의 30%를 차지하고 있다는 사실에 근거하는 것이다.