

표본수 산출법

서울대학교 의과대학 마취통증의학교실

안 원 식

*** 이 글에서 통계의 분류(표본수 산출 통계)에 대한 부분은 1999년 대한마취과학회 춘계 연수강좌에서 사용되었던 글(저자: 안원식)을 대한마취과학회의 허락을 받고 다시 실었음을 밝힙니다.

= 차례 =

- 1. 표본수 산출 이유
- 2. 표본수 산출 통계
- 3. 표본수 산출 방법 개괄
- 4. 마취과에서 많이 사용되는 통계에서 표본수 산출 방법
 - 4.1. 두 평균치 분석
 - 4.2. 카이제곱 검정

1. 표본수 산출 이유

의학연구에 있어서 어떤 집단 간에 서로 같고 다름에 대한 주장을 할 때가 있다. 이 때 다름에 대한 몇 가지 실례를 증거로 제시한다면, 서로 다른 것을 실제로 보였으므로 누구나 수긍할 수 있을 것이다. 그러나 서로 같다고 주장할 때 서로 같은 것을 몇 개 보였다고 하여도 이를 받아들이지 않을 수 있다. Fig. 1과 같은 가상 예를 함께 살펴보자. 상자 A와 상자 B가 속이 안 보이는 상태라고 가정하고 두 상자 안의 내용물이 같은지 여부를 알아보는 연구를 한다고 하자. 또한, 각각의 상자에는 10개가 있고 구성물은 주사위와 별만 있음은 알고 있다. 이 때 각각의 상자에서 1개나 2개의 물건을 꺼냈을 때 모두 주사위가 나왔다고 해서 두 상자의 내용이 같다고 주장한다면, 반론하는 입장에서는 당연히 ‘너무 조금 꺼내 보았기 때문에 우연히 그렇게 나온 것이고, 더 많은 구성물을 꺼내 본다면 다를 수도 있다’라고 주장할 것이다. 이 때 얼마나 더 많이 꺼내는 것이 좋을지에 대한 것은 주장하는 사람마다 다를 수 있지만 이 경우에는 6개 이상이면 명백히 다름을 알 수 있을 것이고, 이렇듯 어떤 개 수 이상으로 구성물을 꺼내서 두 군의 같음 여부를 판단할 때 관련 토론자들이 수긍할 수 있는

개수의 기준을 정하는 것이 표본수 산출의 근본 원리이다. 그런데, 상기 예에서 각 상자에 10개가 있는 것은 동일한데, 별이 A 상자에 1개만 있을 경우에는 몇 개를 뽑아야 할까? (Fig. 2) 또한, 각 상자에 1만개 정도의 주사위와 별이 있고, 주사위가 대부분을 차지하는 것은 알지만 별이 각각 몇 개 있을 때 두 상자의 같음은 어떻게 확인할 수 있을까? (Fig. 3) 이 때 몇 개의 구성물을 꺼내서 두 상자의 같고 다름을 알아보는 것이 타당할지 정하는 것이 표본수 산출이고, 의학연구에서 연구를 시작하기 전에 또는 연구결과를 투고하였을 때 적절한 표본수 산출 근거를 제시하라는 요구를 하는 이유인 것이다. 그래서 이번 강좌에서는 현재 많이 사용되는 표본수 구하는 방법에 대해 알아보고, 마취과 의학자에게 유용한 표본수 산출 책자를 소개하겠다.

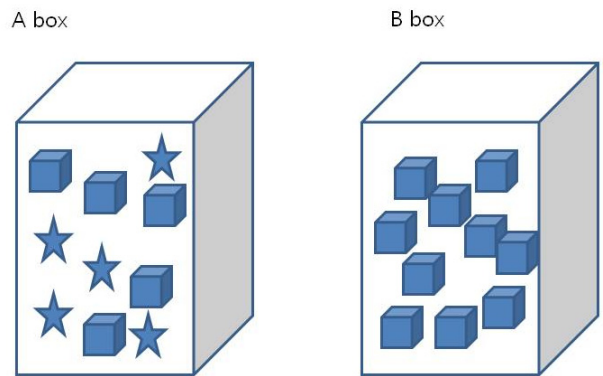


Fig. 1. Box A has five stars and five cubes. Box B has ten cubes. Suppose we would like to know whether the two boxes have the same elements if we do not know the inside of the boxes. When we pick one element out of Box A and one element out of Box B and the selected elements are one star and one cube, then we can say that the boxes have different elements. However, if the two elements are the same, the counter part may insist that the number to be selected is too small to decide whether the two boxes are the same or not. Then, how many elements should be selected?

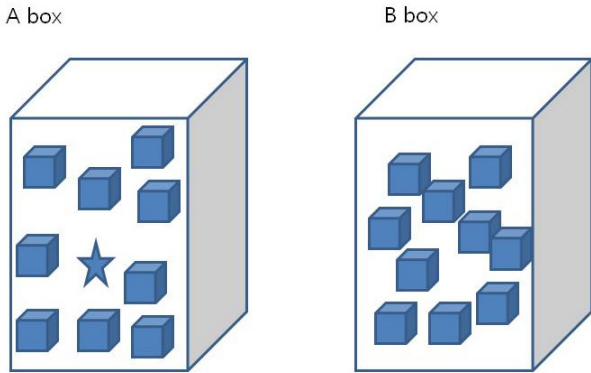


Fig. 2. Box A has one star and nine cubes. Box B has ten cubes. Suppose we would like to know whether the two boxes have the same elements if we do not know the inside of the boxes. When we pick one element out of Box A and one element out of Box B and the selected elements are one star and one cube, then we can say that the boxes have different elements. However, if the two elements are the same, the counter part may insist that the number to be selected is too small to decide whether the two boxes are the same or not. Then, how many elements should be selected?

2. 표본수 산출 통계

통계는 다양한 학문 분야에서 사용되고 있어서 매우 다양한 통계 기법이 있다. 그 중에서 마취과학에서 많이 사용하는 통계에 대해서 먼저 알아보고 이 중에서 어떤 것이 표본수 산출이 필요한 것인지 알아보자.

오늘날 통계적 사고와 통계적 기법의 활용은 과학이란 용어를 붙여둔 모든 학문분야에서 널리 사용되고 있는 추세이다. 통계학이 추구하는 바가 우리가 직면한 혼잡스럽고 불확정적인 상황하에서 합리적이고 과학적인 방법에 따라 숨겨져 있는 규칙성을 찾기 위한 최선의 의사결정이라고 볼 때 통계적 기법의 사용 증가는 당연한 결과라고 하겠다. 그러나 이러한 통계 내지 통계적 기법의 홍수를 면밀히 관찰해 보면 이용상에 많은 문제점을 갖고 있는 것도 사실이다. 통계적 기법을 활용하는 논문에서 부적절한 통계적 기법의 사용은 연구논문의 성취도를 떨어뜨리거나 결정적인 오류를 초래하게 된다.¹⁾ 이러한 오류에 빠지지 않고, 보다 효율적인 연구나 실험을 행하기 위해서는 어느 수준 이상의 통계 지식이 연구자에게 요구되는 상황이므로 마취과 의사들도 통계학 학습에 많은 금전적, 시간적 투자를 하고 있다. 또한 연구자 자신의 통계학 실력이 어느 수준이 되지 못한다면, 연구 디자인과 추측 통계기법의 활용등에 있어서 통계 전문가에게 자문을 구하여 적절한 통계를 사용하려고 노력하고 있다. 그렇지만 통계학 학습을 하려 해도 어느 부

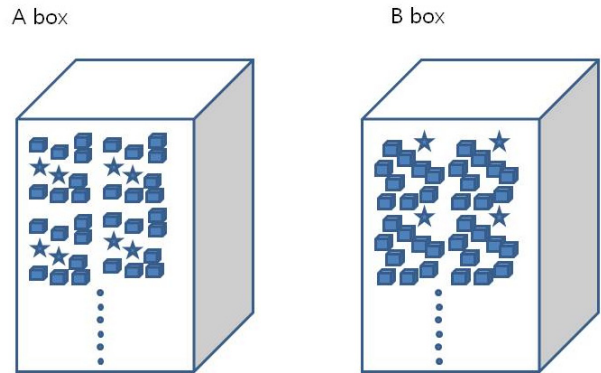


Fig. 3. Box A has 20% stars and 80% cubes. Box B has 10% stars and 90% cubes. Suppose we would like to know whether the two boxes have the same proportion of elements if we do not know the inside of the boxes. When we pick ten elements out of Box A and ten elements out of Box B and the selected elements have 18% and 12% star proportion each, then can we say that the boxes have different elements?

분을 어느 정도 공부해야 하는지도 잘 모르겠고, 통계 전문가와 상담을 하더라도 들을 때 뿐이고 비슷한 예가 나와도 다시 자문을 구해야 안심될 정도로 기초가 튼튼하지 못한 것이 대부분의 마취과 의사들의 통계학 실력이다. 이에 마취과 의사들이 많이 사용하는 통계학 분야들의 개관을 살펴봄으로써 통계 전문가와 상담 중에도 큰 골격은 알고서 상담할 수 있고, 통계학 학습을 할 때에도 큰 줄기는 알고서 공부할 수 있도록 하는데 도움이 되었으면 한다.

2.1. 마취과에서 많이 사용되는 통계학 분야

통계학의 여러 분야 중에서 마취과 의사들이 연구 논문을 쓸 때 주로 사용하는 통계적 기법으로는 표본(통계)조사론, 실험계획법, 기술통계기법, 추측통계기법을 들 수 있다.

2.1.1. 표본 조사론

표본 조사론이란, 관심의 대상이 되는 전체 모집단 중 일부의 부분집단을 선택하고 그 선택된 부분집단만을 대상으로 조사를 실시하여 그로부터 전체 모집단의 특성을 추정해 내는 통계조사방법이라고 정의된다. 실제 방법으로는 단순임의추출법, 층화임의추출법, 계통추출법, 집락추출법, 다단추출법 등이 사용되고 있다.^{2,3)}

2.1.2. 실험계획법

실험계획법이란, 시스템의 성격을 알기 위하여 실험을 수행할 때 가능한 한 최소의 노력으로 최대의 정보를 얻을 수 있고 합리적이고 통계적인 해석이 가능하도록 실험방법

을 결정하려는 통계방법이다. 실제 사용되는 방법은, 일원 배치법, 이원배치법, 다원배치법, 블록화계획, 지분실험계획, 분할구계획, 상관과 회귀, 공분산분석, 요인배치법, 교차법과 일부실시법, 직교배열표를 이용한 실험계획 등이 있다.^{4,5)}

2.1.3. 기술통계기법(Descriptive statistics)

기술통계기법이란, 조사된 자료를 요약 정리하여 나타내 주는 통계적 방법이다.⁶⁾ 여기에는 대표값(평균, 중앙값, 최빈값, 등분위수), 산포도(범위, 표준편차, 분산, 변이계수), 분율(proportion), 비율(rate) 등이 있다.^{7,8)}

2.1.4. 추측통계기법(Inferential statistics)

추측통계기법이란, 표본통계량을 기초로 모수에 대한 추정값을 얻는 것이다.⁹⁾ 여기에는 평균치의 비교, 분할표의 분석, 상관과 회귀, 비모수 통계 등 여러 응용 통계기법이 있다.¹⁰⁾

논문을 쓸 때 상기한 4가지 분야 중에서 표본 조사론과 실험계획법 분야가 가장 먼저 검토되어야 되는 분야이고 이곳에서 표본의 수를 포함한 적절한 설정이 이루어지면 많은 시간과 비용을 절약할 수 있고 보다 효율적인 논문이 될 수 있는 것이지만¹¹⁾ 현재 마취과에서는 주로 일원배치법 또는 이원배치법에 단순임의 추출법을 사용하여 연구를 진행하며, 이러한 방법론적인 기술이 마취과 원저에서 구체적으로 언급되지 않고 있으므로 이번 강좌에서는 제외하겠다. 또한 기술통계는 마취과 논문의 특성상 인구통계학적 자료(demographic data)를 처리하는데 주로 사용되고 있어 본 강좌에서 자세한 내용의 언급은 피하겠고, 다만, 많은 마취과 의사들이 가장 궁금해하고 있고 마취과 논문에서 통계를 사용했다면 빠짐 없이 언급되고 있는, ‘어떤 연구 모형에서 어떤 통계기법을 사용하는가’하는 추측통계기법의 선정 문제에 대해서 검토해 보고자 한다.

2.2. 자료의 유형과 추측통계기법¹²⁻¹⁴⁾

어떤 추측통계기법을 선택할 것인가를 결정하기 위해서는 우선 주어진 자료들이 어떠한 종류로 구성되어 있는가를 먼저 파악해야 한다. 그러기 위해서는 통계자료가 어떤 유형을 가질 수 있는지 유형에 따른 분류에 대해 알아볼 필요가 있겠다.

통계자료를 유형에 따라서 분류하는 방법에는 여러 가지가 있을 수 있겠으나 관측의 척도에 따른 분류와 기능에 따른 분류를 알아보면 다음과 같다.

2.2.1. 관측의 척도(scale)에 따른 분류

1) 명목(명칭) 척도 (nominal scale)

이 척도로 분류되는 변수는 측정된 현상을 순서가 없이 상호 배타적으로 서로 겹치지 않는 상태에서 범주(category)를 구분하는 의미만을 가진다. 이 척도는 특정 상태를 지칭해 주는 기능 외에는 아무런 의미를 가지지 못하는 것으로 숫자가 가지는 특성인 순서의 개념이나 간격의 개념 그리고 가감승제의 수학적 연산기능도 가지지 못하는 변수들이다.

예를 들어, 국소마취제 효과를 비교 검토 하는 논문에서 lidocaine, mepivacaine, bupivacaine 등을 사용하였다고 할 때나 흡입마취제의 심박수에 대한 영향을 평가하는데 있어서 enflurane, isoflurane, sevoflurane 등을 사용하였다고 할 때, 국소마취제나 흡입마취제라는 자료의 성격은 순서가 없고 또한 서로 연관된 것이 없이 독립적이므로 명목척도로 분류한다.

2) 서열(순위) 척도 (ordinal scale)

서열척도로 분류되는 변수들은 특정 상태를 지칭해 주는 명목척도로서의 기능뿐만 아니라 각 범주간의 대소관계, 즉 서열성에 관한 정보도 내포하고 있는 변수를 말한다. 명목척도와 마찬가지로 가감승제 같은 수학적 조작은 불가능하다. 가감승제를 할 수 없으므로 당연히 평균, 표준편차를 구할 수 없다.

예를 들어, double lumen tube intubation 후의 기도 점막의 손상도를 0, +, ++, +++로 나누거나, postdural puncture headache의 나이에 따른 변화를 연구하는데 있어서 나이로 20-30대(20-39세), 40대-50대(40-59세), 60-70대(60-79세)의 3개의 군으로 나누었다면 이들은 순서가 있으므로 서열척도로 분류한다. 또한, 통증 치료의 효과를 판정할 때 VAS (visual analogue scale)를 사용하였다면 이들이 비록 0, 1, 2, ..., 10의 형태를 취하더라도 통증의 정도가 등간격이 아니기 때문에 서열척도로 분류한다.

3) 구간(간격) 척도 (interval scale)

특정 상태의 지칭이나 대소관계개념 외에도 측정치간의 간격에 의미를 부여할 수 있을 때 이러한 변수를 구간척도라 한다. 수학적으로는 가감의 조작은 가능하지만, 승제의 조작은 불가능하다. 또한 절대 영(absolute zero)이 존재하지 않는다.

예를 들어, 체온을 섭씨나 화씨로 표시하게 되면 이들은 각각 일정한 간격을 가지나 절대 영을 가지지 못하고 (0℃가 온도가 없는 것이 아니고 절대온도로 표시하면 273 °K 임), 0℃ ≙ 0°F (0℃ = 32°F)이므로 이를 구간척도로 분류한다.

4) 비척도 (ratio scale)

구간척도에 비해 절대 영점을 가지기 때문에 수학적으로 가장 완전한 형태의 변수들로서 가감승제 등의 모든 수학적 조작이 가능하다.

예를 들어, 마취기 회로내의 압력을 mmHg나 cmH₂O로 표시하게 되면 이들은 각각 일정한 간격을 가지고 있고, 절대 영을 가지고 있다(0 mmHg = 0 cmH₂O).

5) 범주형 자료(categorical data)와 연속형 자료(continuous data)

흔히 1), 2)를 합쳐서 범주형 자료(categorical data, qualitative data)라고 하고, 3), 4)를 합쳐서 연속형 자료(continuous data, quantitative data)라고 한다.

2.2.2. 기능에 따른 분류

변수들의 기능에 따라 분류해 볼 수도 있는데, 여기에는 종속변수(dependent variable 또는 반응변수(responsive variable))와 독립변수(independent variable 또는 설명변수(explanatory variable))가 있다. 종속 변수란 그 행태에 관심을 갖고 행태를 규명하려는 변수이고, 독립변수란 종속변수의 행태를 규명하는데 도구로써 사용하는 변수를 말한다. 예를 들어, esmolol에 의한 혈압하강 효과를 볼 때, 혈압은 종속 변수이고, esmolol의 용량은 독립변수가 된다. 논문에서 그래프로 표시될 때 가로축에 있는 것이 독립변수이고, 세로축에 있는 것이 종속변수가 된다.

2.2.3. 마취과에서 많이 쓰는 추측통계기법

위에서 언급한 자료의 성격과 통계분석 절차 사이의 관계를 표로 나타내면 Table 1과 같다. 표에서 보는 바와 같이 자료의 유형에 따라 합당한 통계절차가 있으므로 통계기법 선정에는 자료의 성격을 파악하는 것이 제일 먼저 이루어져야 될 것이다.

마취과에서 사용하는 자료는, 독립변수가 범주형 자료(예, esmolol 30mg 정주군, 60mg 정주군, 90mg 정주군)이고 종속변수가 연속형자료(예, 혈압)인 것이 대부분이어서, 통계처

Table 1. 자료의 성격과 통계분석 절차의 관계

		독립변수	
		범주형	연속형
종속변수	범주형	분할표의 분석	로지스틱 회귀분석
	연속형	t-test, 분산분석	회귀분석

*비모수 통계 : 자료가 모수적 가정을 충족하지 못할 때 사용
 *상관관계 분석 : 큰 범주의 회귀 분석에 포함됨

리는 주로 평균치의 비교 방법인 t-test 또는 분산분석을 시행하고 있다.¹⁵⁾ 또한 독립변수와 종속변수가 모두 범주형인 경우 분할표의 분석을 사용하고, 두 변수 모두 연속형인 경우 회귀분석을 쓴다. 그 외에 독립변수가 연속형이고 종속변수가 범주형인 로지스틱 회귀분석이 있으나 마취과 영역에서 자주 사용되지 않으므로¹⁵⁾ 강좌에서는 제외하겠다. 이외에 자료가 모수적 가정을 충족하지 못할 때 사용하는 비모수 통계, 두 변수간에 선형관계의 정도를 분석하는 상관관계분석 등이 사용되고 있으므로 이들의 개관을 살펴 보고자 한다.

2.3. 추측 통계 기법

2.3.1. t-test와 분산분석(평균치의 비교)^{16,17)}

마취과 논문에서 가장 많이 사용되는 자료는, 독립변수가 범주형 자료이고 종속변수가 연속형자료이므로 주로 평균치의 비교 방법인 t-test 또는 분산분석을 시행하고 있다.¹⁵⁾ 이 중에서 두 군간의 평균치를 비교할 때는 t-test를 사용하고, 세 군이상 여러 군의 평균치를 비교할 때는 분산분석을 사용하고 있다.

1) t-test

두 개의 평균치에 대한 통계적 분석에서 가장 많이 쓰이고 있는 방법으로, 여기에는 unpaired t-test 와 paired t-test가 있는데 자료의 독립성 여부에 따라 선택적으로 사용된다. 자료의 독립성에 대한 이해를 위해 Fentanyl의 심박수에 대한 효과를 분석하는 예를 들어 보면 다음과 같다.

a. unpaired t-test

처치군 30명에는 Fentanyl 100mcg을 정주하고 대조군 30명에는 생리식염수를 정주한 5분 후에 심박수를 측정하여 두 군간에 심박수의 차이를 비교하는 연구

b. paired t-test

대상자 30명의 심박수를 Fentanyl 주사 전에 측정하고, Fentanyl 100mcg을 정주한 후에 심박수를 측정하여 심박수의 차이를 비교하는 연구

첫째 예에서는 처치군의 심박수와 대조군의 심박수에는 아무런 상관관계가 없는 독립적인 성질을 갖고 있고, 두 번째의 예에서 Fentanyl 주사 5분 후의 심박수는 주사 전에 측정한 심박수에 영향을 받는 의존적 현상(dependent phenomena)이 일어나므로 이 때에는 자료가 짝을 이루었다(paired)고 표현하고 paired t-test를 적용하게 된다.

자료의 독립성에 대한 이해를 돕기 위해 또다른 예를 들자면, Coarctation of Aorta 환자의 상지와 하지의 혈압을 비교하는 예에서 상하지의 혈압이 서로 독립적이라고 생각하여 대상 환자 모두의 상지혈압 평균과 하지혈압 평균을 구

해서 평균끼리 비교하는 것 보다는, 의존적 자료라고 판단하여 한 환자의 상지와 하지 혈압을 상호 연관시킨 후 자료를 처리하는 것이 좋을 것 같다.

한편 관찰된 표본수가 너무 작아 모집단의 분포를 가정할 수 없을 때이거나 자료의 분포가 정규분포가 아닌 경우에는 비모수 검정법을 사용하여야 한다. unpaired t-test에 상응하는 비모수 검정으로는 Wilcoxon rank sum test가 있고, paired t-test에 상응하는 비모수 검정으로는 Wilcoxon signed rank test가 있다.¹⁸⁾

2) 분산분석(Analysis of variance, ANOVA)¹⁹⁾

3개 이상의 평균치를 비교하는 방법으로 one-way ANOVA, two-way ANOVA, repeated measures ANOVA 등 여러 방법이 있고, 각각의 군을 개별 비교하기 위한 방법으로 다중비교(multiple comparison, 예; Tukey, Dunckan 등)를 사용하고 있다.

one-way ANOVA와 two-way ANOVA는 unpaired t-test에 상응하는 검사법이고, repeated measures ANOVA는 paired t-test에 상응하는 검사법이다. 흡입마취제 종류와 혈압의 관계를 연구하는 모델을 예로 들어 설명하면 다음과 같다.

a. one-way ANOVA

대상 환자를 3개의 군으로 나누어서 흡입마취제, enflurane, isoflurane, desflurane 1 MAC을 5분간 흡입 시킨 후 혈압의 변화를 비교하는 연구

b. two-way ANOVA

대상 환자를 6개의 군으로 나누어서 3군에는 흡입마취제, enflurane, isoflurane, desflurane 1 MAC을 5분간 흡입 시킨 후 혈압 변화를 측정하고, 3군에는 enflurane, isoflurane, desflurane 1 MAC을 15분간 흡입 시킨 후 혈압 변화를 측정하여 비교하는 연구

c. repeated measures ANOVA

대상 환자를 3개의 군으로 나누어서 흡입마취제, enflurane, isoflurane, desflurane 1 MAC을 5분간 흡입 시킨 후 혈압 변화를 측정하고, 연속해서 15분간 흡입 시킨 후 혈압 변화를 측정하여 비교하는 연구 (한 환자에서 5분, 15분의 측정이 모두 이루어지는 경우를 말함)

t-test와 마찬가지로 표본수가 너무 작아 모집단의 분포를 가정할 수 없는 경우에는 비모수 검정을 시행하게 된다. 상응하는 비모수 검정으로는 one-way ANOVA: Kruskal-Wallis test, two-way ANOVA: Friedman test, repeated measures ANOVA: 페이지 검정 등이 있다.¹⁸⁾

2.3.2. 분할표의 분석^{20,21)}

독립변수와 종속변수가 모두 범주형일 때 사용하는 범주

Table 2. 피어슨 카이제곱 검정에서 사용되는 분할표 예

	통증 발생 안함	통증 발생함	합계
대조군	5	45	50
처리군	45	5	50
합계	50	50	100

* 대조군 : lidocaine 전처치 없이 propofol을 정주한 군

* 처리군 : lidocaine 전처치 후 propofol을 정주한 군

형 자료의 분석에는 분할표의 분석, 로그선형 모형, 로지스틱 회귀분석, 관련도 등이 있는데, 주로 쓰이는 것은 분할표의 분석이므로 이것을 propofol 정주로 인한 통증이 lidocaine 전처치로 사라지나의 여부에 대한 연구 모델로 검토해 보자.

1) 피어슨의 카이제곱 검정

범주형 자료로 구성된 두 군간의 동질성 검정이나, 독립성 검정에 2 X 2 분할표를 사용하여 관찰된 표가 일어날 확률을 구하는 가장 대표적인 방법이 피어슨의 카이제곱 검정(흔히, 카이제곱 검정)이다. 예를 들면, 대조군 50명에는 propofol을 30mg 정주한 후 통증 발생 여부를 묻고, 처리군 50명에는 lidocaine 30mg을 먼저 정주한 후 propofol 30mg을 정주한 후 통증 발생 여부를 묻는 연구에서는, 대조군과 처리군, lidocaine의 전처리 여부라는 자료의 성질상 범주형 자료에 속하므로 Table 2와 같은 분할표를 이용하여 카이제곱 검정 통계량을 이용하여 독립성 여부를 검정한다.

상기의 검정법에서 보다 정확한 분포에 의한 통계량을 얻기 위해 ‘연속성의 수정’을 이용하는 경우가 자주 있다.

또한 다음의 두 경우에는 피셔의 정확한 검정(Fisher’s exact test)을 사용하는 것이 좋다.

a. 전체 표본 크기가 20 미만인 경우

b. 전체 표본 크기가 40 미만이고 한 cell의 최소 기대치가 5 이하일 경우

※ 학자에 따라서는 ‘전체 표본의 크기와 상관없이 한 cell의 최소 기대치가 5 이하일 경우’를 주장하는 경우도 있다.

2) 맥네마 검정(McNemar test)

수집된 범주형 자료가 짝을 이룬 범주형 자료(paired categorical data)인 경우에는 카이 제곱 검정을 사용하지 않고 맥네마 검정을 사용한다. 예를 들면, propofol 정주로 인한 통증이 lidocaine 전처치로 사라지나의 여부에 대한 연구를 계획할 때 대조군과 처리군을 한 사람에게서 얻는다면, 즉, 한 사람에게서 한 쪽 팔에 propofol 30mg을 정주한 후

Table 3. 맥네마 검정에서 사용되는 분할표 예

		대조군		합계
		통증발생함	통증발생안함	
처리군	통증발생함	14	1	15
	통증발생안함	25	10	35
	합계	39	11	50

* 대조군 : lidocaine 전처치 없이 propofol을 정주한 군
 * 처리군 : lidocaine 전처치 후 propofol을 정주한 군
 * 단, 대조, 처리가 한 사람에게서 이루어짐

통증 발생 여부를 묻고 다른쪽 팔에 lidocaine 30mg을 정주한 후 propofol을 정주하여 통증 발생 여부를 물었다면(먼저 투여한 propofol이 다음에 투여한 propofol의 통증에는 영향을 미치지 않는다고 가정하고), 여기서 얻은 자료는 서로 연관된 자료(어떤 사람이 통증에 민감하다면 두 경우 모두 통증을 호소할 가능성이 높고, 또 반대의 경우도 있을 수 있다.)이므로 맥네마 검정을 사용하고 분할표를 Table 3과 같이 작성한다.

3) I X J 분할표

분할표의 항수가 여러개이어도 대개는 카이제곱 통제량을 이용하면 된다.

2.3.3. 회귀분석^{22,23)}

독립변수와 종속변수가 모두 연속형일 때 사용하는 회귀분석은 관련된 변수들간에 상호 관련성을 찾으려고 할 때 주로 사용한다. 예를 들어 개심술 마취 후 혈소판의 감소 정도에 영향을 미치는 요인을 알기 위해 수술시간, cardiopulmonary bypass (이하, CPB로 약함) 시간, 출혈량 등과 술 후 혈소판 수치간에 회귀 분석을 하여

$$(술후 혈소판 수치) = (술전 혈소판 수치) - 0.03 (수술시간) - 0.2 (CPB 시간) - 0.1 (출혈량)$$

이라는 가상적인 회귀분석식을 얻었다면 혈액 검사를 해보지 않더라도 혈소판 수혈의 근거를 CPB 시간에 제일 큰 비중을 두어 미리 판단해 볼 수 있을 것이다. 이처럼 변수들간의 함수관계를 추구하는 통계적 방법을 회귀분석(regression analysis)이라고 한다. 회귀모형의 종류로는 단순회귀, 중회귀, 곡선회귀, 다항회귀, 비선형회귀 등 여러 종류가 있다.

1) 상관분석

상관분석은 엄밀한 의미에서는 회귀분석의 전단계로서

변수들 사이의 관계를 살펴보는 것이지만 넓은 의미에서는 회귀분석에 포함된다.

상관분석은 변수들 간의 선형관계의 정도를 분석하는 방법이다. 이를테면, 한 변수가 증가할 때 다른 변수가 얼마나 선형적으로 증가 또는 감소하는가를 조사할 필요가 있을 것이다. 회귀분석과는 달리 상관분석에서는 단지 관계의 존재여부와 정도를 분석하며 관련변수들이 구체적으로 어떠한 선형함수 관계를 갖는지는 조사하지 않는다.²⁴⁾

예를 들면, isoflurane 흡입농도(예, 0 MAC에서 3 MAC)와 cerebral blood flow (이하 CBF로 약함)와의 선형관계의 정도를 분석하는 것이 상관분석이다.

모수적 검정에서 상관의 측정치는 피어슨의 적률상관계수 r인데, 이 때 관측치들은 양변수가 모두 정규집단에서 추출되었다는 가정 아래 적어도 구간척도 이상으로 측정되어야 한다. 그러나, 이러한 정규성의 가정을 충족시킬 수 없거나 관측치가 명목척도나 서열척도에 의한 것이라면 비모수적 상관계수를 산출한다. 명목척도로 측정된 관측치들에 대해서는 분할계수(contingency coefficient)를 이용하여 상관계수를 구하고, 서열척도로 측정된 관측치들에 대해서는 스피어만 순위 상관계수를 이용하여 구한다.²⁵⁾

2) 상관계수(correlation coefficient, r)와 결정계수(coefficient of determination, r²)

상관계수는 단지 두 변수의 선형관계만을 나타내는 지표이다. 그러므로 상관계수가 0인 경우에 두 변수의 선형상관관계는 없지만 다른 관계는 가질 수 있다. 상관계수는 -1에서 +1 사이의 값을 가지며, 값이 +1에 가까울수록 강한 양의 상관관계를, -1에 가까울수록 강한 음의 상관관계를 나타내며, 값이 0에 가까울수록 선형관계는 약해진다. 예를 들면, enflurane 흡입농도와 CBF 그리고 isoflurane 흡입농도와 CBF와의 상관분석에서 enflurane의 상관계수가 +0.9가 나오고 isoflurane의 상관계수가 +0.2가 나왔다면(상관분석의 관점에서만 본다면) isoflurane보다 enflurane이 CBF를 증가시키는데 보다 밀접한 관련이 있다고 해석한다.

결정계수는 측정된 자료들이 회귀선에 의하여 설명되는 비율을 말하는데, 회귀선에 의하여 설명되는 변동이 기여하는 비율을 의미하므로 결정계수를 회귀선의 기여율이라고 부르기도 하며 r²의 범위는 0≤r²≤1이다. 모든 관측점들이 회귀선상에 위치한다면 r²=1이다. 이와 반대로, 회귀관계가 전혀 없어서 측정된 회귀선의 기울기가 영(0)이면 r²=0이 된다. 예를 들어, 앞의 개심술 후의 혈소판 추정의 회귀식에서 r²이 0.99라면, 술후 혈소판 수치를 직접 측정하지 않고 계산에 의해서 구해도 별 무리가 없다. 만약 r²이 0.39라면 60% 가량의 외부요인이 있으므로 이 회귀식에 의한 계산만으로는 술 후의 혈소판 수치의 정확한 예측에는

무리가 있고 단지 CPB 시간이 길 경우 혈소판이 낮을 가능성이 높다는 것을 예측할 수 있는 정도에 지나지 않을 것이다.

3) 회귀모형의 진단(diagnostic)

얻어진 자료를 이용하여 어떤 회귀식을 얻었다면 이러한 회귀모형이 타당한가 여부를 검토하는 통계적 절차가 필요한데 이를 회귀모형의 진단(diagnostic)이라고 말하고 다음과 같은 사항을 점검해야 한다.

- a. 잔차의 등분산성 검증
 - i. 산점도
 - ii. 편회귀플롯(partial regression plot)
- b. 잔차의 독립성 검증(오차의 자기상관 검토)
 - i. Durbin-Watson 검증
- c. 잔차의 정규성 검증
 - i. 정규확률플롯(normal probability plot)
- d. 다중 공선성(multicollinearity)의 검토
 - i. VIF(variance inflation factor)
 - ii. condition index
 - iii. variance proportion
- e. 특이점의 검출
 - i. leverage
 - ii. 잔차
- f. 영향이 큰 데이터의 파악(influential cases)
 - i. Cook's D 통계량
 - ii. DFFITS 통계량
 - iii. COVRATIO 통계량
 - iv. DFBETAS 통계량

4) 최적 회귀방정식의 선택

종속변수에 영향을 끼치는 독립변수의 수가 많은 경우 어떤 독립변수를 최종 모형 안에 포함시켜야 하느냐 하는 문제에서, 가능한 모든 독립변수를 회귀모형에 전부 포함시키는 경우에 결정계수 r^2 은 증가하지만, 독립변수의 수가 많아짐에 따라 상관관계가 커져서 다중공선성의 문제가 야기되어 회귀모형의 안정성 및 신뢰성을 떨어뜨리게 된다. 예를 들면, 앞에서 구한 혈소판 수치 예측 회귀식

$$(\text{술후 혈소판 수치}) = (\text{술전 혈소판 수치}) - 0.03 (\text{수술시간}) - 0.2 (\text{CPB 시간}) - 0.1 (\text{출혈량})$$

에서 CPB 시간이 길 경우 당연히 수술시간이 증가하여 수술시간도 마치 술후 혈소판 수치에 영향을 미치는 것으로 회귀식이 구해질 수 있으나 이것은 CPB 시간 증가에 의해

부차적으로 일어나는 현상일 수 있으므로 이를 검증(다중공선성의 검토)하여 항목에 넣을 것인가 여부를 결정해야 한다. 또한 중요한 독립변수만을 잘 선택해서 간단한 회귀모형을 만들게 되면 이해하기도 쉽고 예측모형을 임상에서 활용하기도 편하게 된다.

최적 회귀방정식을 선택하는 방법으로는 변수제거법(backward elimination method), 변수선택법(forward selection method), 단계적 방법(stepwise method)이 있다.

5) 타일 검증

회귀모형에서 기울기 모수에 대한 비모수적 검증법의 대표적인 방법이다.¹⁸⁾

2.3.4. 비모수 검증법^{18,26)}

비모수 검정을 설명하기 전에 이해를 돕기 위해 비모수 검증법의 예부터 검토해 보자.

1) 마취과에서 자주 쓰이는 비모수 검증법(상응하는 모수적 검증법)

- a. Wilcoxon rank sum test (unpaired t-test)
 - 예) 임신한 양 12마리를 6마리씩 두 군으로 나눠서, 한 군에는 lidocaine을 투여하고 다른 군에는 mepivacaine을 투여하여 1시간 후 fetal lamb에서의 lidocaine과 mepivacaine의 잔여용량을 비교연구하는 모델
- b. Wilcoxon signed rank test (paired t-test)
 - 예) 잡견 6마리에서 basal metHb 농도를 측정 한 후 NO 20 ppm을 20분간 흡입시킨 후 metHb을 다시 측정하여 증가여부를 연구하는 모델
- c. Kruskal-Wallis test (one-way ANOVA)
 - 예) VLDL(Very Low Density Lipoprotein), LDL(Low Density Lipoprotein), HDL(High Density Lipoprotein)이 각각 정상보다 높은 5명씩으로 구성된 3개의 환자군에서 일정량의 halothane으로 1시간 동안 마취했을 때 간 효소의 증가가 동일한가 여부를 연구하는 모델
- d. Friedman test (two-way ANOVA)
 - 예) VLDL (Very Low Density Lipoprotein), LDL (Low Density Lipoprotein), HDL (High Density Lipoprotein)이 각각 정상보다 높은 5명씩으로 구성된 6개의 환자군에서 일정량의 halothane으로 3군(VLDL, LDL, HDL)은 1시간동안 마취하고, 3군(VLDL, LDL, HDL)은 2시간 동안 마취했을 때 간 효소의 증가가 동일한가 여부를 연구하는 모델(각 군은 한 번만 대상이 됨)
- e. 페이저 검증(repeated measures ANOVA)
 - 예) VLDL (Very Low Density Lipoprotein), LDL (Low Density Lipoprotein), HDL (High Density Lipoprotein)이 각각

정상보다 높은 5명씩으로 구성된 3개의 환자군에서 일정량의 halothane으로 1시간 마취 후 간 효소를 측정 한 후 계속해서 2시간 동안 마취한 후 다시 간 효소치를 측정하여 증가가 동일한가 여부를 연구하는 모델(각 군은 두 번(1시간, 2시간) 연속해서 채혈함)

f. 타일 검정(회귀분석)

예) 간 이식 수술을 받은 8명의 환자에서 일정 용량의 propofol을 투여했을 때 투여 중단 후 의식 회복시간을 예측하는데 있어서 나이, 체중, 투여시간과 회복시간의 함수 관계를 연구하는 모델

2) 모수적 검정과 비모수적 검정

연구자가 얻은 연구자료들은 하나의 표본으로서 어떤 모집단에서 뽑혀 나왔고 그 모집단의 성격이나 분포(평균, 분산 등과 같은 모수)에 관하여는 몇 가지 사항이 가정되어 있으며 이를 전제로 통계적 추론이 이루어지는데 이를 ‘모수적 방법’(parametric method)이라고 한다. 예를 들면 어떤 통계기법은 정규분포하는 모집단으로부터 표본이 추출되었다는 가정을 근거로 하고, 어떤 통계기법은 두 표본이 동일한 분산을 가지는 두 모집단들로부터 추출되었다고 하는 가정에 근거를 두고 있다. 따라서 이러한 방법에 의하여 전개된 분석결과는 엄밀하게 “**단일 모집단(들)의 형태에 관**

한 가정들이 타당하다면, 연구자는라고 결론을 내린다” 와 같은 한정적 내용이 되는 것이다. 그러나 우리가 실제 접하는 자료는 그러한 가정들에 잘 맞지 않은 경우가 허다하다. 예컨대 정규분포와는 크게 다른 형태의 분포, 즉 좌우 대칭이 아닌 두터운 꼬리(heavy tail)를 가지고 있다든가, 모집단의 분포함수가 연속이 아닌 경우가 있을 수 있다. 그러므로 이 경우 정규분포나 연속성 등을 가정하고 전개하여 얻어진 통계적 분석결과는 오류의 가능성이 높다.

비모수적 통계기법(nonparametric statistical methods)은 모수에 관하여 특수한 분포를 전제하지 않거나, 평균치와 같은 어떤 특정값(parameter)이 아니고 분포형태 자체를 분석한다거나 혹은 통계적 검정에 이용되는 귀무가설하에서의 검정통계량의 분포와 모집단의 분포함수와를 전혀 무관(distribution-free method)하게 하여 분석하는 방법이다. 대부분의 비모수적 통계분석법은 어떤 가정들, 예를 들어 ‘관찰은 독립적이어야 한다’, ‘연구대상이 되는 변수는 그 기초가 되는 분포가 연속적이어야 한다’는 등이 있으나 이들 가정들은 모수적 방법에서의 가정들보다는 그 수가 적고 또 훨씬 약하다.

3) 비모수적 통계분석을 적용하게 되는 경우

- a. 수집한 자료가 명백하게 정규분포를 따르지 않는 경우

Table 4. 측정수준과 비모수 검정법

측정의 수준		척도		
		명목척도	서열척도	구간척도
단일표본		이항검정 chi-제곱검정	Kolmogorov-Smirnov검정 runs 검정	
두 표본	종속표본	McNemar 검정	부호검정 Wilcoxon의 부호화 순위검정	Walsh검정 무작위검정
	독립표본	Fisher의 직접확률법 chi-제곱검정	중앙치 검정 Mann-Whitney의 U검정 Kolmogorov-Smirnov검정 Wilcoxon 순위합검정 Moses검정	무작위검정
K 표본	종속표본	Cochran의 Q 검정	페이지 검정	
	독립표본	chi-제곱검정	중앙치검정 Kruskal-Wallis의 one-way ANOVA Friedman의 2-way ANOVA	
상관		분할계수: c	Spearman의 서열 상관계수 Kendall의 서열 상관계수 Kendall의 편서열 상관계수 Kendall의 일치도 계수	

- b. 표본수가 너무 적어 분포를 알 수 없는 경우로서 통상 한 집단의 표본수가 10 전후 또는 두 집단인 경우는 합하여 20 전후일 때
- c. 중간에 간이 분석 결과를 보고 싶을 때
- d. 자료의 관측 척도가 명목척도나 서열척도일 때

4) 비모수적 방법의 장점

- a. 대부분의 비모수적 방법으로 얻어지는 확률은 모집단의 분포에 관계없이 정확하다.
- b. 표본의 크기가 6 이하일 경우에는 비모수적 방법을 적용할 수 밖에 없다.
- c. 분포의 성격(분산 등)이 다른 몇 개의 모집단으로부터 구성된 표본을 다룰 때 모수적 방법의 어느 것도 비현실적인 가정을 전제하지 않고서는 적용할 수가 없는 경우가 있는데, 이런 경우에도 적용할 수 있는 비모수적 방법이 개발되어 있다.
- d. 선천적으로 서열을 이루고 있는 자료들도 다룰 수 있다.
- e. 비모수적 검정법은 계산이 쉽고, 간단하며 이용범위가 넓다.

5) 비모수적 방법의 단점

- a. 연구자료가 모수적 검정법의 모든 가정을 충족시킬 때 비모수검정법을 적용하면 정보의 손실이 일어나 자료를 낭비하는 결과를 초래한다. 즉, 동일한 검정력을 갖기 위해서는 표본수를 더 많이 취해야 한다.
- b. 표본의 크기가 커지면 오히려 계산이 복잡하고 어려워진다.

6) 측정의 수준과 비모수적 검정법

측정의 수준과 척도에 따라 비모수 검정법을 나누어 보면 Table 4와 같다.

2.4. 표본수 산출이 필요한 통계

각종 실험 연구에서 어떤 통계처리를 할 것인가를 고민할 때, Table 1에서 보는 바와 같이, 연구에서 어떤 변수가 사용되었는가를 살펴보고 적절한 분석 방법을 선택하는 것이 가장 빠른 길이라고 생각된다. 변수의 종류를 나누는 방법은 여럿 있겠지만 범주형 자료와 연속형 자료로 나누는 경우에, 독립변수와 종속변수가 모두 범주형인 경우 분할표의 분석을 사용하고, 두 변수 모두 연속형인 경우 회귀분석을 쓰며, 독립변수가 범주형이고, 종속변수가 연속형일 경우 평균치 비교(t-test, ANOVA)를 사용하면 된다. 또한 관찰된 표본수가 너무 작아 모집단의 분포를 가정할 수 없을 때나 모수적 검정에 필요한 가정을 충족시키지 못할 때는 비모수 검정을 사용하면 된다.

지금까지 마취과에서 많이 사용되고 있는 통계 기법 중

에서 어떤 상황에서 어떠한 통계들이 사용되는가를 다루는 추측 통계분야에 대해 간략하게 알아보았다. 이 밖에 많은 마취과 의사들이 통계학의 한 분야라고 생각하는 것으로서 SAS, SPSS 같은 응용 프로그램이 있다. 이것은 전통적인 통계학의 한 분야는 아니지만, 자료처리의 효율을 높이기 위해 또한 통계를 보다 빠른 시간에 해결하기 위한 보조 수단으로서 필요하다. 이렇듯 학술논문을 쓸 때 필수 불가결한 요소가 되고 있으므로 응용 프로그램의 사용법²⁷⁻³²⁾을 공부하면서 ‘통계를 공부하고 있다’는 말을 사용할 정도이다. 그러나, 이러한 학습은 통계학을 모르는 상태에서 이루어질 경우 사상누각이 될 가능성이 매우 높다. 그러므로 통계에 대한 기본적인 사항을 함께 학습해 나가는 것이 타당하리라 생각된다.

상기 통계 기법 중에서 기술 통계에서는 표본수 산출이 필요하지 않지만, 추측 통계의 모든 분야에서는 적절한 표본수 산출이 필요하다. 그러므로, 먼저 전반적인 표본수 산출 방법을 알아보겠다.

3. 표본수 산출 방법 개괄

앞에서 예로 든 연구 설계에서 표본수 산출은 다음과 같은 식으로 구할 수 있다.

$$\text{표본수} \cong (1 - \text{제 1종 오류}) \times (1 - \text{제 2종 오류}) \times (\text{분산}) / (\text{검정하려는 차이})$$

이 때 분산은 선행 다른 연구나 pilot study에서 구하게 되고, 제 1종 오류(alpha)나 제 2종 오류(beta), 검정하려는 차이는 연구자가 정하게 된다. Alpha나 beta는 관례적으로 0.05, 0.01 이나 0.2, 0.1을 각각 사용하지만 이러한 수치가 절대적인 것은 아니다. 또한, 검정하려는 차이는 연구 설계에 맞게 임상적으로나 학문적으로 합리적인 범위 이내이면 연구자가 자유롭게 정할 수 있다.³³⁾ 그러므로, 이러한 설정을 적절히 밝히지 않으면 서로 다른 기준으로 표본수를 산정할 수 있고, 연구 결과의 판단도 정 반대의 결과를 가져올 수 있으므로 연구를 디자인하고, 논문을 작성할 때 방법론에 기술해 주는 것이 좋겠다.

이제 각론으로 들어가 두 평균치 분석, 카이제곱 검정에 대해 표본수 산출법을 예를 들어 알아보겠다.

4. 마취과에서 많이 사용되는 통계에서 표본수 산출 방법

마취과에서 많이 사용되는 통계의 표본수 산출법을 실제 연구에 사용된 예를 들어 Cohen의 책에서 설명하는 방법을 이용하여 설명하겠다.³⁴⁾

4.1. 두 평균치 분석

연구 계획의 요약은 다음과 같음

연구제목: 일회용 압력 측정 장치(Autotransducer[®])의 길이와 사용시간에 따른 natural frequency (NF)와 damping coefficient (DC) 변화³⁵⁾

연구목적: 일회용 압력 측정 장치를 장기간 사용하거나 길이를 다르게 사용할 때 측정하고자 하는 압력을 얼마나 적절히 반영하는지 알아보고자 함

연구방법: 여러 종류의 길이로 준비한 일회용 압력 측정 장치를 장기간 사용하면서 수액을 투여한 상태에서 압력을 급격히 변화시켜(step input) 압력 측정기의 Natural frequency와 damping coefficient를 측정함

4.1.1 목표 시험 횟수

모든 설정 시험당 15회 이상, 총 225회

4.1.2 산출 근거

Autotransducer[®]를 이용한 비슷한 선행 연구가 없어 시험적으로 2주간의 자료를 받아 보았다. 각 시점과 길이에서의 NF와 DC의 평균과 표준편차는 각각 20~23±1.0~1.8Hz, 0.22~0.28±0.01~0.03이었고, 각 상황에서 필요한 검출 횟수는 NF 22Hz, DC 0.24 인지 여부에 대해 다음 절에 power analysis와 같다. 이를 고려하여 시험 횟수를 15주로 설정하였다.

4.1.3 power analysis

1) NF 평균치 추정에 대한 검정력

평균치 비교와 분산에 관한 식은 다음과 같은 식이 성립한다.

$$d = (m - c) / \sigma \quad (\text{참고문헌 34의 46쪽 식 2.3.3})$$

단, d: 검정력 분석을 위한 매개변수

m: 표본 평균

c: 검정하고자 하는 평균

σ : 표본 표준편차

이번 연구의 pilot test에서 나온 평균과 표준 편차를 고려하여 검정하고자 하는 평균 22Hz와 가장 차이가 많이 나고 표준 편차가 큰 96시간을 기준으로 하면 평균이 20Hz이고, 표준편차는 1.8초이다. 검정하고자 하는 평균이 22Hz이므로, 검정력 분석을 위한 매개변수 계산 값은,

$$\begin{aligned} d &= (m - c) / \sigma \\ &= (20-22) / 1.8 \\ &= -1.1 \end{aligned}$$

d는 차이를 알아보기 위한 매개변수이므로 절대값을 취하여 1.1로 하고, 연구 가정이 22Hz와 다른가 여부를 알아보는 기획이므로 양측검정을 취한다. 그래서, 제 1종 오류 alpha가 0.05, power 80%로 설정된 본 연구에서는, alpha=0.05, beta = 0.2, d = 1.1이 되어 참고문헌 55쪽 표 2.4.1에 의해 14.5회가 나오게 된다.

2) DC 평균치 추정에 대한 검정력

마찬가지 방법으로 이번 연구의 pilot test에서 나온 평균과 표준 편차를 고려하여 검정하고자 하는 평균 0.24와 가장 차이가 많이 나고 표준 편차가 큰 96시간을 기준으로 하면 평균이 0.28이고, 표준편차는 0.03초이다. 검정하고자 하는 평균이 0.24Hz이므로, 검정력 분석을 위한 매개변수 계산 값은,

$$\begin{aligned} d &= (m - c) / \sigma \\ &= (0.28-0.24) / 0.03 \\ &= 1.33 \end{aligned}$$

연구 가정이 0.24와 다른가 여부를 알아보는 기획이므로 양측검정을 취한다. 그래서, 제 1종 오류 alpha가 0.05, power 80%로 설정된 본 연구에서는, alpha=0.05, beta = 0.2, d = 1.33이 되어 참고문헌 55쪽 표 2.4.1에 의해 10회가 나오게 된다.

3) 총 시험 횟수 결정

NF 관련하여 14.5회를 시험하여야 하고, DC 와 관련하여 10회 이상을 시험하여야 하기에 15회를 시험 횟수로 정하였다.

4.2 카이제곱 검정

연구 개요

Sevoflurane 마취 유도 중 Fentanyl 투여 여부(saline control)가 다음 증상 발현 정도에 미치는 영향 평가³⁶⁾

1. 기침; 4 levels
2. excitation; 2 levels
3. apnea (mask ventilation 어려움); 2 levels
4. laryngospasm; 2 levels
5. secretion; 2 levels

4.2.1 contingency table에서 Effect size (W) 산출식

$$w = \sqrt{\sum_{i=1}^m \frac{(P_{1i} - P_{0i})^2}{P_{0i}}} = \sqrt{\frac{1}{N} \sum_{i=1}^m \frac{(n_{1i} - n_{0i})^2}{n_{0i}}}$$

(참고문헌 34의 216쪽 식 7.2.1)

- P_{0i} : 귀무 가설이 맞을 때 cell i의 비율
- P_{1i} : 대립 가설이 맞을 때 cell i의 비율
- m : cell의 수
- N : 총 수
- n_{0i} : 귀무 가설이 맞을 때 cell i의 수
- n_{1i} : 대립 가설이 맞을 때 cell i의 수

4.2.2 표본수 산출

각 항목에서 effect size를 0.40-0.78 정도로 설정하여 필요한 표본수는 최저 13명에서 최고 52명으로 나와 최소 표본수는 52명임. 대상 선정 후 누락을 고려하여 총 60명을 대상으로 연구를 시작하기로 함. 자세한 산출 과정은 다음 항목에 기술함.

4.2.3 표본수 산출 과정

1) 기침: 4 levels

- Summary of Pilot Study
- Fentanyl 투여군 25명
- Control 24명
- 총 49명

pilot study 결과

Cough	0	1	2	3	Total
Fentanyl	24	1	0	0	25
Control	14	2	6	2	24
Total	38	3	6	2	49

연관성이 없을 경우 분할표 예상치

Cough	0	1	2	3	Total
Fentanyl	19.4 (38x25/49)	1.5 (3x25/49)	3.1 (6x25/49)	1.0 (2x25/49)	25
Control	18.6 (38x24/49)	1.5 (3x24/49)	2.9 (6x24/49)	1.0 (2x24/49)	24
Total	38	3	6	2	49

상기 pilot test에서의 effect size (W)는 다음과 같이 계산됨.

$$\sqrt{\frac{1}{49} \times \left\{ \frac{(24 - 19.4)^2}{19.4} + \frac{(14 - 18.6)^2}{18.6} + \frac{(1 - 1.5)^2}{1.5} + \frac{(2 - 1.5)^2}{1.5} + \frac{(0 - 3.1)^2}{3.1} + \frac{(6 - 2.9)^2}{2.9} + \frac{(0 - 1)^2}{1} + \frac{(2 - 1)^2}{1} \right\}} = 0.47$$

- Type I error (alpha) = 0.05
- Type II error (beta) = 0.2 => Power 80%
- Degree of Freedom = 3 (<= (4-1) × (2-1))
- Effect size (W) = 0.47

effect size 0.40 일 때 68명, 0.50 일 때 44명임. (참고문헌 258쪽 표 7.4.6에 의해)

intrapolation에 의해, (0.47 - 0.40)/(0.50 - 0.40) × (44 - 68) + 68 = 51.2
minimum needed number is 52.

2) excitation: 2 levels

- Summary of Pilot Study
- Fentanyl 투여군 25명
- Control 24명
- 총 49명

pilot study 결과

Excitation	0	1	Total
Fentanyl	23	2	25
Control	4	20	24
Total	27	22	49

연관성이 없을 경우 분할표 예상치

Excitation	0	1	Total
Fentanyl	12.8 (27x25/49)	11.2 (22x25/49)	25
Control	13.2 (27x24/49)	10.8 (22x24/49)	24
Total	27	22	49

상기 pilot test에서의 effect size (W)는 다음과 같이 계산됨.

$$\sqrt{\frac{1}{49} \times \left\{ \frac{(23 - 12.8)^2}{12.8} + \frac{(2 - 11.2)^2}{11.2} + \frac{(4 - 13.2)^2}{13.2} + \frac{(20 - 10.8)^2}{10.8} \right\}} = 0.78$$

Type I error (alpha) = 0.05
 Type II error (beta) = 0.2 => Power 80%
 Degree of Freedom = 1 (<= (2-1) x (2-1))
 Effect size (W) = 0.78

effect size 0.70 일 때 16명, 0.80 일 때 12명임. (참고문헌 표 7.4.6에 의해)
 intrapolation에 의해, (0.78-0.70)/(0.80-0.70)×(12-16)+16=12.8
 minimum needed number is 13.

3) apnea (mask ventilation 어려움); 2 levels

Summary of Pilot Study
 Fentanyl 투여군 25명
 Control 24명
 총 49명

pilot study 결과

Apnea	0	1	Total
Fentanyl	24	1	25
Control	13	11	24
Total	37	12	49

연관성이 없을 경우 분할표 예상치

Apnea	0	1	Total
Fentanyl	18.9 (37x25/49)	6.1 (12x25/49)	25
Control	18.1 (37x24/49)	5.9 (12x24/49)	24
Total	37	12	49

상기 pilot test에서의 effect size (W)는 다음과 같이 계산됨.

$$\sqrt{\frac{1}{49} \times \left\{ \frac{(24 - 18.9)^2}{18.9} + \frac{(1 - 6.1)^2}{6.1} + \frac{(13 - 18.1)^2}{18.1} + \frac{(11 - 5.9)^2}{5.9} \right\}} = 0.48$$

Type I error (alpha) = 0.05
 Type II error (beta) = 0.2 => Power 80%
 Degree of Freedom = 1 (<= (2-1) x (2-1))
 Effect size (W) = 0.48

effect size 0.40 일 때 49명, 0.50 일 때 31명임. (참고문헌 표 7.4.6에 의해)

intrapolation에 의해, (0.48 - 0.40)/(0.50 - 0.40)×(31 - 49) + 49=34.6
 minimum needed number is 35.

4) laryngospasm; 2 levels

Summary of Pilot Study
 Fentanyl 투여군 25명
 Control 24명
 총 49명

pilot study 결과

spasm	0	1	Total
Fentanyl	19	6	25
Control	15	9	24
Total	34	15	49

연관성이 없을 경우 분할표 예상치

spasm	0	1	Total
Fentanyl	17.3 (34x25/49)	7.7 (15x25/49)	25
Control	16.7 (34x24/49)	7.3 (15x24/49)	24
Total	34	15	49

상기 pilot test에서의 effect size (W)는 다음과 같이 계산됨.

$$\sqrt{\frac{1}{49} \times \left\{ \frac{(19 - 17.3)^2}{17.3} + \frac{(6 - 7.7)^2}{7.7} + \frac{(15 - 16.7)^2}{16.7} + \frac{(9 - 7.3)^2}{7.3} \right\}} = 0.15$$

Type I error (alpha) = 0.05
 Type II error (beta) = 0.2 => Power 80%
 Degree of Freedom = 1 (<= (2-1) x (2-1))
 Effect size (W) = 0.40 (effect size 0.15를 검출한다는 것은 임상적으로 큰 의미가 없는 상황이므로, 통상적으로 사용되는 medium effect size level인 0.40을 채택함)

effect size 0.40 일 때 49명임. (참고문헌 표 7.4.6에 의해)
 minimum needed number is 49.

5) secretion: 4 levels

Summary of Pilot Study

Fentanyl 투여군 25명

Control 24명

총 49명

pilot study 결과

secretion	0	1	2	3	Total
Fentanyl	16	9	0	0	25
Control	5	16	2	1	24
Total	21	25	2	1	49

연관성이 없을 경우 분할표 예상치

secretion	0	1	2	3	Total
Fentanyl	10.7 (21x25/49)	12.8 (25x25/49)	1.0 (2x25/49)	0.5 (1x25/49)	25
Control	10.3 (21x24/49)	12.2 (25x24/49)	1.0 (2x24/49)	0.5 (1x24/49)	24
Total	21	25	2	1	49

상기 pilot test에서의 effect size (W)는 다음과 같이 계산됨.

$$\sqrt{\frac{1}{49} \times \left\{ \begin{array}{l} (16 - 10.7)^2 / 10.7 + (9 - 12.8)^2 / 12.8 + (0 - 1.0)^2 / 1.0 + (0 - 0.5)^2 / 0.5 \\ 0.5 + (5 - 10.3)^2 / 10.3 + (16 - 12.2)^2 / 12.2 + (2 - 1)^2 / 1 + (1 - 0.5)^2 / 0.5 \end{array} \right\}} = 0.47$$

Type I error (alpha) = 0.05

Type II error (beta) = 0.2 => Power 80%

Degree of Freedom = 3 (<= (4-1) x (2-1))

Effect size (W) = 0.47

effect size 0.40 일 때 68명, 0.50 일 때 44명임. (참고문헌 표 7.4.6에 의해)

intrapolation에 의해, (0.47-0.40)/(0.50-0.40)×(44-68)+68=51.2
minimum needed number is 52.

5. 맺음말

임상 연구에서 사용되고 있는 표본수 산출법이나 검정력 추정에 대해 논란이 있지만, 현행 제도권 하에서 임상 연구를 수행하려면, 윤리위원회를 통과하여야 하고, 윤리위원회 통과를 위해서는 현재 일반적으로 인정되는 표본수 산출법

을 알아야 한다. 이 글에서는 가능한 수식을 줄이거나 수식을 사용하더라도 매우 간단한 수식을 사용하여 두 군의 평균치 검정이나 카이제곱 검정을 하는 것을 알아보았다. 이러한 방법을 참고로 하여 그 밖에 마취과학에서 다루는 대부분의 연구 방법에 대해 이 글에서 사용한 책자나 영국에서 최근 출판한 다른 책(ISBN: 0-521-80916-9)을 참조하면 표본수, 검정력을 스스로 산출할 수 있을 것을 생각된다.³⁷⁾

감사의 글

이 글에 자신의 연구 기획을 신도록 허락해 준 김희수 선생님께 감사의 뜻을 전한다.

참 고 문 헌

1. Choi JH, Lee JC: Academic Thesis & Statistical Methods. 1st edition. Seoul, Free Academy 1990, pp 1-7.
2. Lee GO, Park JW: The study of Sample Research. 1st edition. Seoul, Korean National Open University Press. 1995, pp 3-21.
3. Bae KH, Lee TL: The study of Statistical Research. 1st edition. Seoul, Korean National Open University Press. 1995, pp 104-178.
4. Lim YC, Ahn BJ, Sung NK: Research Planning I. 1st edition. Seoul, Korean National Open University Press. 1995, pp 1-10.
5. Lim YC, Ahn BJ, Sung NK: Research Planning II. 1st edition. Seoul, Korean National Open University Press. 1995, pp iv-v.
6. Glantz SA: Primer of Biostatistics. 2nd ed. Roman, McGraw-Hill. 1987, pp 10.
7. Ahn YO, Yoo KY, Park BJ: Medical Statistics. 2nd edition. Seoul, Korean National Open University Press. 1996, pp 6-12.
8. Huh MH, Lee TL, Lim SS: Statistical Data Analysis I -SAS-. 1st edition. Seoul, Korean National Open University Press. 1994, pp 99-111.
9. Lim YC, Ahn BJ, Sung NK: Research Planning I. 1st edition. Seoul, Korean National Open University Press. 1995, pp 18.
10. Ahn YO, Yoo KY, Park BJ: Medical Statistics. 2nd edition. Seoul, Korean National Open University Press. 1996, pp vii-ix.
11. Lee YS, Kou H: Post-hoc Survey for Power of 119 Negative Results in Korean Journal of Anesthesiology. The Korean Journal of Anesthesiology 1999; 36: 286-292.
12. Lee TL, Kim BS, Lee YG: Categorical Data Analysis. 1st edition. Seoul, Korean National Open University Press. 1995, pp 2-5.
13. Aviva Petrie: Lecture Notes on Medical Statistics. 1st ed. London, Blackwell Scientific Publications. 1978, pp 1-3.
14. Yoo KY: Categorical Data Analysis for Medical-Public Health. 1st edition. Seoul, Seoul National University Press. 1996, pp 21-22.
15. Kou H, Gak IY, Kim GW and others: Present Statistical Status in Papers in the Korean Journal of Anesthesiology (1981-1990). The Korean Journal of Anesthesiology 1993; 26: 22-7.
16. Ahn YO, Yoo KY, Park BJ: Medical Statistics. 2nd edition. Seoul, Seoul National University Press. 1996, pp 59-86.

17. Yoo KY, Park BJ, Kim H and others: PC-SAS for Medicine & Public Health. 1st edition. Seoul, Hanul Academy. 1995, pp 81-100.
18. Song MS, Park CS: Non-Parametric Statistics. 1st edition. Seoul, Seoul National University Press. 1995.
19. Glantz SA, Slinker BK: Primer of Applied Regression and Analysis of Variance. 1st ed. Singapore, McGraw-Hill Book Co. 1990, pp 272-463.
20. Lee TL, Kim BS, Lee YG: Categorical Data Analysis. 1st edition. Seoul, Seoul National University Press. 1995.
21. Yoo KY: Categorical Data Analysis for Medical-Public Health. 1st edition. Seoul, v. 1996.
22. Park SH, Hwang SY: Regression Analysis. 1st edition. Seoul, Seoul National University Press. 1994.
23. Glantz SA, Slinker BK: Primer of Applied Regression and Analysis of Variance. 1st ed. Singapore, McGraw-Hill Book Co. 1990, pp 10-271.
24. Kagn GS, Kim SC, Kim JH and others: PC Statistics. 1st edition. Seoul, Free Academy. 1995, pp 218-223.
25. Ahn YO, Yoo KY, Park BJ: Medical Statistics. 2nd edition. Seoul, Seoul National University Press. 1996, pp 273.
26. Ahn YO, Yoo KY, Park BJ: Medical Statistics. 2nd edition. Seoul, Seoul National University Press. 1996, pp 243-6.
27. Huh MH, Lee TL, Lim SS: Statistical Data Analysis I -SAS-. 1st edition. Seoul, Korean National Open University Press. 1994.
28. Song IS, Lee TL, Kim SS: Statistical Data Analysis II -SPSS-. 1st edition. Seoul, Korean National Open University Press. 1994.
29. Choi BS: Introduction to PC SAS. 2nd edition. Seoul, Pakyoungsa. 1992.
30. Lee KE: SAS Application Methods-Data Processing & Statistical Analysis-. 1st edition. Seoul, Daeun Publishing Company. 1985.
31. Yoo KY, Park BJ, Kim H and others: PC-SAS for Medicine & Public Health. 1st edition. Seoul, Hanul Academy. 1995.
32. SAS/STAT User's guide, Version 6. 4th ed. Cary, SAS Institute Inc. 1990.
33. Kim H: Estimating sample size in clinical trials. Korean J Anesthesiol 2002; 42: 1-10.
34. Cohen J: Statistical power analysis for the behavioral sciences. 1988. 2nd ed. Lawrence Elbaum Associates. Hillsdale, New Jersey.
35. Lim KJ, Kim D, Ahn W, Lee NJ: The changes of natural frequency and damping coefficient of disposable pressure transducer according to the catheter length and time assigned. Korean J Anesthesiol 2007; 53: 759-66.
36. Lee J, Oh Y, Kim C, Kim S, Park H, Kim H: Fentanyl reduces desflurane-induced airway irritability following thiopental administration in children. Acta Anaesthesiol Scand 2006; 50: 1161-4.
37. Bausell RB, Li Y: Power analysis for experimental research: A practical guide for the biological, medical and social sciences. Cambridge, Cambridge University Press. 2002.