

평균과 비율 검정에서 표본 크기와 검정력 계산의 구현

이창선¹ · 강희모² · 심송용³

¹²³한림대학교 금융정보통계학과

접수 2011년 11월 22일, 수정 2011년 12월 21일, 게재확정 2011년 12월 26일

요약

많은 조사에서 표본 크기는 유의수준 또는 제1종 오류확률만을 고려하여 결정하였으나 최근에는 다양한 분야를 중심으로 제1종 오류확률뿐만 아니라 제2종 오류확률 또는 검정력을 함께 고려하여 표본 크기를 결정하는 경우가 늘어나고 있다. 이런 경향은 표본을 많이 얻을 수 없는 연구에서 더욱 뚜렷하다. 본 연구에서는 모평균과 모비율에 대한 검정에서 제1종 오류뿐만 아닌 제2종 오류를 고려한 경우 필요한 표본 크기를 결정하는 과정을 살펴보고 이를 웹사이트를 통해 공개하였다. 또한 주어진 표본 크기와 유의수준에 의한 검정력 계산도 함께 공개하였다.

주요용어: 검정력, 표본 크기, 효과 크기.

1. 서론

통계학은 지식기반사회의 학문으로 여러 연구 분야뿐만 아니라 실생활에서도 그 응용범위가 더욱 넓어지고 있다. 이러한 배경에서 학문 분야를 포함한 사회 여러 부분의 많은 사람들에게 있어 통계학에 대한 이해는 필수적인 것이 되고 있다. 디지털 혁명이라고 명명되고 있을 정도로 많은 정보가 발생하며 통계학은 대량의 데이터를 수집하여 축적하고 이를 과학적으로 분석하여 의사결정에 활용하는 하나의 학문 분야로서 그 중요성이 강조되고 있다 (Kang과 Sim, 2003; 김진경 등, 2008).

많은 조사나 실험에서 필요한 표본 크기를 결정할 때 제1종 오류확률 (α)만 고려하여 계산한다. α 를 고려한 연구는 Ko와 Kim (2010)의 비모수적 검정방법에서 표본 크기에 대해 연구나 Kwak 등 (2010)의 임상시험에서 베이지안 방법을 이용한 표본 크기에 대한 연구 등을 참고할 수 있다. 하지만 의학 및 생물통계와 많은 임상실험에서 필요한 표본의 수는 제2종 오류확률 ($1 - \beta$) 또는 검정력 (β)를 α 와 함께 고려하는 경우가 많으며 (Rosner, 2010 등) 검정력을 고려한 표본 크기 결정은 Cohen (1988)을 참고하기로 한다.

하지만 제 2종 오류 또는 검정력을 함께 고려한 경우의 표본 크기를 계산하는 것은 다소 복잡한 계산을 거쳐야 하며, 특히 표본 크기가 먼저 얻어졌을 경우에는 특정한 대립가설의 값에서 검정력이 얼마나 되는지 계산은 따로 제시되는 경우가 드물고 Cohen (1988)을 포함한 대부분의 연구는 필요한 표본 크기를 계산하여 표로 제시하고 있다. 하지만 표라는 제약 때문에 모든 경우에 대한 표본 크기나 검정력을 제시할 수 없는 한계를 가지고 있다.

본 연구에서는 이러한 계산을 웹사이트에서 시행하고 계산결과를 웹에서 바로 확인할 수 있도록 중간 계산을 구현하였다. 사이트의 주소는

¹ (200-702) 강원도 춘천시 옥천동 1번지, 한림대학교 대학원 금융정보통계학과, 대학원생.

² (200-702) 강원도 춘천시 옥천동 1번지, 한림대학교 금융정보통계학과, 겸임부교수.

³ 교신저자: (200-702) 강원도 춘천시 옥천동 1번지, 한림대학교 금융정보통계학과, 교수.

E-mail: sysim@hallym.ac.kr

<http://pluto.hallym.ac.kr/zsize> 또는 <http://jupiter.hallym.ac.kr/zsize>

이며 사용 언어는 PHP이다 (루크 웰링, 2009; 정진호, 2000; 최완규, 2002).

인터넷 환경의 발달로 통계적 계산을 웹의 CGI (Common Gateway Interface)나 PHP 등과 같은 웹 기반 언어로 구현이 가능하며 대화형 솔루션을 제공한 연구는 편리한 접근성과 대화형 그래프 등을 제공한다 (Kang과 Sim, 2000; Kang과 Sim, 2003; Sim과 Lee, 1999).

본 연구에서는 모평균과 모비율에 대한 가설검정에서

- 유의수준 (또는 제1종 오류) 뿐만 아니라 검정력을 함께 고려한 경우의 필요한 표본 크기에 대한 계산과
- 유의수준과 표본 크기를 고려한 경우 검정력 계산

을 웹에 구현하였다. 웹 사이트에 제1종의 오류와 제2종의 오류를 고려한 표본 크기 계산은 사용자에게 편리하도록 단측 검정과 양측 검정인 경우 모두 구현하였으나 본 논문에서는 단측 검정인 경우 양측 검정에서 유의수준에 2를 곱하여 계산할 수 있으므로 양측 검정만 소개하였다.

제2절에서는 모평균 μ 에 대한 검정에서 유의수준과 검정력을 고려할 때 필요한 표본 크기와 유의수준과 표본 크기를 고려할 경우 검정력 계산을 소개하였고, 제3절에서는 모비율 p 에 대한 검정에 대하여 제2절에서 소개된 것과 동일한 사용자 인터페이스로 구현하였다.

2. 평균 검정에서 필요한 표본 크기와 검정력

귀무가설이 $H_0 : \mu = \mu_0$ 이고 대립가설이 양측 또는 단측 가설인 경우 모평균 μ 에 대한 검정에서 표본 크기를 얻고자 할 때는 다음과 같이 두 가지 형태로

- 분산, 유의수준, 귀무가설, 대립가설 및 대립가설에서의 검정력, 또는
- 효과 크기 (effect size), 유의수준 및 검정력을

필요한 매개변수를 입력하였을 때 표본 크기를 계산하며, 검정력은 아래와 같이 두 가지 형태로

- 귀무가설, 대립가설, 분산, 표본 크기 및 유의수준 또는
- effect size, 표본 크기 및 유의수준을

필요한 매개변수를 입력한 경우 계산된다. 여기서 effect size는 귀무가설 하에서 평균값과 대립가설 하에서 평균값의 차이를 표준편차로 보정한 것으로 $|\mu_a - \mu_0|/\sigma$ 이며 평균차가 표준편차의 몇 배정도인지 정해주는 역할을 하며, μ_a 는 대립가설에서 모평균 값이다 (Cohen, 1988).

2.1. 단일 표본 모평균 μ 에 대한 검정

X_1, X_2, \dots, X_n 이 정규분포 $N(\mu, \sigma^2)$ 의 확률표본이고 σ^2 이 알려져 있을 때 표본 크기에 대하여 알아보자. 귀무가설 $H_0 : \mu = \mu_0$ 와 대립가설 $H_1 : \mu \neq \mu_0$ 인 평균검정에서 유의수준 α 와 대립가설하에서의 모평균 값 μ_a 에서의 검정력이 최소한 β 가 되기 위해 필요한 표본 크기 n 은

$$n = \frac{(z_{\alpha/2} + z_{1-\beta})^2 \cdot \sigma^2}{(\mu_a - \mu_0)^2} \quad (2.1)$$

이고, 위의 표본 크기를 사용하였을 때 모평균이 μ_a 일 때 이 검정의 검정력 β 는

$$\beta = \Pr \left\{ Z > -z_{\alpha/2} + \frac{\sqrt{n \cdot (\mu_a - \mu_0)^2}}{\sigma} \right\} \quad (2.2)$$

가 된다. 또한 effect size (e.s; δ)로 표본 크기를 구하면

$$n = \frac{(z_{\alpha/2} + z_{1-\beta})^2}{\delta^2}, \quad \delta = \frac{|\mu_a - \mu_0|}{\sigma} \quad (2.3)$$

이며, 검정력은

$$\beta = \Pr \left\{ Z > -z_{\alpha/2} + \sqrt{n\delta^2} \right\} \tag{2.4}$$

로 계산된다. 위의 결과를 사용한 표본 크기는 검정력과 effect size 및 유의수준의 특정한 값에 대해서는 Cohen (1988)에서 표로 정리되어 있다. 하지만 많은 연구자들은 특정한 값을 입력하였을 때 표본 크기나 검정력을 원하기 때문에 이 표의 사용은 일부 한계점이 있다.

따라서 본 연구에서는 사용자가 특정한 값들을 입력하였을 때 표본 크기와 검정력을 계산해주며 역으로 주어진 표본 크기와 유의수준을 입력하면 검정력을 계산한다.

단일 표본 모평균 μ 에 대한 양측 검정은

- 모평균 μ 에 대한 양측 검정에서 표본 크기 계산은 귀무가설, 대립가설, 분산, 유의수준, 검정력을 입력하고 (그림 2.1(a)) 계산하면 결과는 effect size와 표본 크기가 나타난다 (그림 2.1(b)).

<div style="text-align: center;">2. two-sided</div> <p>유의수준(α)와 검정력(β), effect size를 알고 있을 경우의 양측 표본수 계산</p> <p>유의 수준(α) <input type="text" value="0.05"/></p> <p>검정력(β) <input type="text" value="0.9"/></p> <p>귀무가설($\mu = \mu_0$) <input type="text" value="10"/></p> <p>대립가설($\mu = \mu_1$) <input type="text" value="9"/></p> <p>분산(σ^2) <input type="text" value="5"/></p> <p style="text-align: center;"><input type="button" value="표본수 계산"/></p>	<div style="text-align: center;">단일 표본수 계산(two-sided)</div> <p>유의 수준(α) = 0.05, 검정력(β) = 0.9, $\mu_0 = 10$, $\mu_1 = 9$, 분산(σ^2) = 5인 경우</p> <p>effect size는 0.2 이고,</p> <p>계산된 결과는 262.6856 입니다.</p> <p>따라서 표본수는 263 개입니다.</p>
--	---

(a) 필요한 매개변수 입력

(b) 가설을 이용한 표본 크기 계산 결과

그림 2.1 단일 표본 양측 검정에서 표본 크기 계산 및 결과 (가설 이용)

- 모평균 μ 에 대한 양측 검정에서 표본 크기 계산은 effect size, 유의수준, 검정력을 입력하고 (그림 2.2(a)) 계산하면 결과는 표본 크기가 나타난다 (그림 2.2(b)).

<div style="text-align: center;">2. two-sided</div> <p>유의수준(α)와 검정력(β), effect size를 알고 있을 경우의 양측 표본수 계산</p> <p>유의 수준(α) <input type="text" value="0.05"/></p> <p>검정력(β) <input type="text" value="0.9"/></p> <p>effect size <input type="text" value="0.3"/></p> <p style="text-align: center;"><input type="button" value="표본수 계산"/></p>	<div style="text-align: center;">단일 표본수 계산(two-sided)</div> <p>유의 수준(α) = 0.05, 검정력(β) = 0.9, effect size = 0.3 인 경우</p> <p>계산된 결과는 116.7491 입니다.</p> <p>따라서 표본수는 117 개입니다.</p>
---	---

(a) 필요한 매개변수 입력

(b) e.s를 이용한 표본 크기 계산 결과

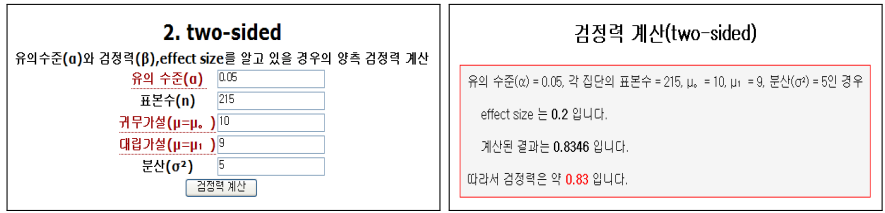
그림 2.2 단일 표본 양측 검정에서 표본 크기 계산 및 결과 (effect size 이용)

- 모평균 μ 에 대한 양측 검정에서 검정력 계산은 귀무가설, 대립가설, 분산, 유의수준, 표본 크기를 입력하고 (그림 2.3(a)) 계산하면 결과는 effect size와 검정력 나타난다 (그림 2.3(b)).
- 모평균 μ 에 대한 양측 검정에서 검정력 계산은 effect size, 유의수준, 표본 크기를 입력하고 (그림 2.4(a)) 계산하면 결과는 검정력 나타난다 (그림 2.4(b)).

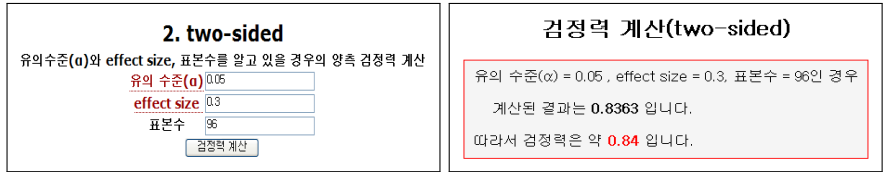
단일 표본 단측검정인 경우 대립가설의 방향에 상관없이 표본 크기 및 검정력은 식 (2.1) – (2.4)에서 $\alpha/2$ 대신 α 를 사용하며, 사용법은 양측검정과 마찬가지로이다. 구현 결과는 위와 같은 사이트에서 제공된다.

2.2. 독립인 두 표본의 모평균 $\mu_1 - \mu_2$ 에 대한 검정

X_1, X_2, \dots, X_{n_1} 이 서로 독립이며, 정규분포 $N(\mu_1, \sigma_1^2)$ 의 확률표본이고, Y_1, Y_2, \dots, Y_{n_2} 가 서로 독립이며, 정규분포 $N(\mu_2, \sigma_2^2)$ 의 확률표본이라 하자. 또한 X_i 와 Y_j 들도 서로 독립이다. 여기서 $n_1 = n_2 = n$ 및 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 인 경우에 대해서 알아보기로 한다.



(a) 가설을 이용한 검정력 계산 (b) 가설을 이용한 검정력 계산 결과
 그림 2.3 단일 표본 양측 검정에서 검정력 계산 및 결과 (가설 이용)



(a) e.s를 이용한 검정력 계산 (b) e.s를 이용한 검정력 계산 결과
 그림 2.4 단일 표본 양측 검정에서 검정력 계산 및 결과 (effect size 이용)

이 경우 귀무가설 $H_0 : \mu_1 - \mu_2 = 0$ 와 대립가설 $H_1 : \mu_1 - \mu_2 \neq 0$ 에 대한 가설검정에서 유의수준 α , 대립가설하에서 두 그룹의 모평균이 각각 μ_{a1}, μ_{2a} 인 경우 최소 β 만큼의 검정력이 있으려면 필요한 표본 크기 n 은

$$n = \frac{(z_{\alpha/2} + z_{1-\beta})^2 \cdot 2\sigma^2}{(\mu_{2a} - \mu_{1a})^2} \tag{2.5}$$

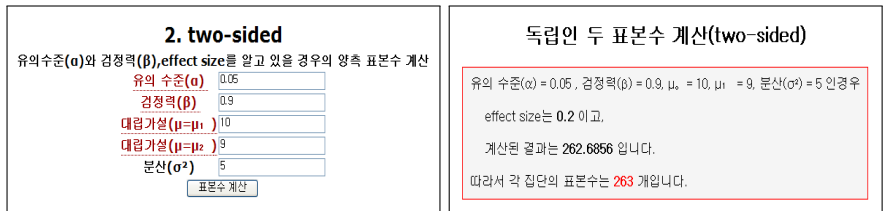
이다. 위의 대립가설하에서 평균값과 표본 크기 n 에서 이 검정의 검정력 β 는

$$\beta = \Pr \left\{ Z > -z_{\alpha/2} + \sqrt{\frac{n \cdot (\mu_{1a} - \mu_{2a})^2}{2\sigma^2}} \right\} \tag{2.6}$$

이다. 이표본 평균검정에서의 effect size $\delta = (\mu_{2a} - \mu_{1a})/\sigma$ 로 주어지므로 식 (2.5) 및 식 (2.6)에 $(\mu_{2a} - \mu_{1a})/\sigma$ 대신 δ 를 사용하면 된다.

독립인 두 표본의 모평균 $\mu_1 - \mu_2$ 에 대한 양측 검정은

- 모평균 $\mu_1 - \mu_2$ 에 대한 양측 검정에서 표본 크기 계산은 귀무가설, 대립가설, 분산, 유의수준, 검정력을 입력하고 (그림 2.5(a)) 계산하면 결과는 effect size와 표본 크기가 나타난다 (그림 2.5(b)).



(a) 가설을 이용한 표본 크기 계산 (b) 가설을 이용한 표본 크기 계산 결과
 그림 2.5 독립인 두 표본 양측 검정에서 표본 크기 계산 및 결과 (가설 이용)

- 모평균 $\mu_1 - \mu_2$ 에 대한 양측 검정에서 표본 크기 계산은 effect size, 유의수준, 검정력을 입력하고 (그림 2.6(a)) 계산하면 결과는 표본 크기가 나타난다 (그림 2.6(b)).

<p style="text-align: center;">2. two-sided</p> <p>유의수준(α)와 검정력(β), effect size를 알고 있을 경우의 양측 표본수 계산</p> <p>유의 수준(α) <input type="text" value="0.05"/></p> <p>검정력(β) <input type="text" value="0.9"/></p> <p>effect size <input type="text" value="0.3"/></p> <p style="text-align: center;"><input type="button" value="표본수 계산"/></p>	<p style="text-align: center;">독립인 두 표본수 계산(two-sided)</p> <p>유의 수준(α) = 0.05, 검정력(β) = 0.9, effect size = 0.3 인 경우</p> <p>계산된 결과는 116.7491 입니다.</p> <p>따라서 각 집단의 표본수는 117 개입니다.</p>
---	--

(a) e.s를 이용한 표본 크기 계산

(b) e.s를 이용한 표본 크기 계산 결과

그림 2.6 독립인 두 표본 양측 검정에서 표본 크기 계산 및 결과 (effect size 이용)

- 모평균 $\mu_1 - \mu_2$ 에 대한 양측 검정에서 검정력 계산은 귀무가설, 대립가설, 분산, 유의수준, 표본 크기를 입력하고 (그림 2.7(a)) 계산하면 결과는 effect size와 검정력 나타난다 (그림 2.7(b)).

<p style="text-align: center;">2. two-sided</p> <p>유의수준(α)와 검정력(β), effect size를 알고 있을 경우의 양측 검정력 계산</p> <p>유의 수준(α) <input type="text" value="0.05"/></p> <p>표본수(n) <input type="text" value="215"/></p> <p>대립가설($\mu = \mu_1$) <input type="text" value="10"/></p> <p>대립가설($\mu = \mu_2$) <input type="text" value="9"/></p> <p>분산(σ^2) <input type="text" value="5"/></p> <p style="text-align: center;"><input type="button" value="검정력 계산"/></p>	<p style="text-align: center;">검정력 계산(two-sided)</p> <p>유의 수준(α) = 0.05, 각 집단의 표본수 = 215, $\mu_1 = 10$, $\mu_2 = 9$, 분산(σ^2) = 5인 경우</p> <p>effect size 는 0.2 입니다.</p> <p>계산된 결과는 0.8346 입니다.</p> <p>따라서 검정력은 약 0.83 입니다.</p>
---	--

(a) 가설을 이용한 검정력 계산

(b) 가설을 이용한 검정력 계산 결과

그림 2.7 독립인 두 표본 양측 검정에서 검정력 계산 및 결과 (가설 이용)

- 모평균 $\mu_1 - \mu_2$ 에 대한 양측 검정에서 검정력 계산은 effect size, 유의수준, 표본 크기를 입력하고 (그림 2.8(a)) 계산하면 결과는 검정력 나타난다 (그림 2.8(b)).

<p style="text-align: center;">2. two-sided</p> <p>유의수준(α)와 effect size, 표본수를 알고 있을 경우의 양측 검정력 계산</p> <p>유의 수준(α) <input type="text" value="0.05"/></p> <p>effect size <input type="text" value="0.3"/></p> <p>표본수 <input type="text" value="117"/></p> <p style="text-align: center;"><input type="button" value="검정력 계산"/></p>	<p style="text-align: center;">검정력 계산(two-sided)</p> <p>유의 수준(α) = 0.05, effect size = 0.3, 각 그룹의 표본수 = 117인 경우</p> <p>계산된 결과는 0.9006 입니다.</p> <p>따라서 검정력은 약 0.9 입니다.</p>
---	--

(a) e.s를 이용한 검정력 계산

(b) e.s를 이용한 검정력 계산 결과

그림 2.8 독립인 두 표본 양측 검정에서 검정력 계산 및 결과 (effect size 이용)

독립 이 표본 단측검정인 경우 대립가설의 방향에 상관없이 표본 크기 및 검정력은 식 (2.5) 및 식 (2.6)에서 $\alpha/2$ 대신 α 를 사용하며, 사용법은 양측검정과 마찬가지로이다. 구현 결과는 위와 같은 사이트에서 제공된다.

3. 비율 p 에 대한 검정에서의 표본 크기와 검정력

모비율 p 에 대한 검정에서 표본 크기는

- 귀무가설, 대립가설, 분산, 유의수준, 검정력
을 입력하였을 때 계산하며, 검정력도 아래와 같이
- 귀무가설, 대립가설, 분산, 표본 크기, 유의수준
을 입력한 경우 계산된다.

3.1. 단일 표본 모비율 p 에 대한 검정

귀무가설 $H_0 : p = p_0$ 와 대립가설 $H_1 : p \neq p_0$ 에 대한 검정에서 유의수준 α , 대립가설하에서의 모비율 값 p_a 에서의 검정력이 최소한 β 가 되기 위해 필요한 표본 크기 n 은

$$n = \frac{\left(z_{\alpha/2} \sqrt{p_0(1-p_0)} + z_{1-\beta} \sqrt{p_a(1-p_a)} \right)^2}{(p_a - p_0)^2} \tag{3.1}$$

이며, 위의 표본 크기 n 을 사용할 때 모비율이 p_a 일 때의 검정력 β 는

$$\beta = \Pr \left\{ Z > \frac{-z_{\alpha/2} \sqrt{p_0(1-p_0)} + \sqrt{n(p_a - p_0)^2}}{\sqrt{p_a(1-p_a)}} \right\} \tag{3.2}$$

이다.

단일 표본 모비율 p 에 대한 양측 검정에서 표본 크기와 검정력의 계산은

- 모비율 p 에 대한 양측 검정에서 표본 크기 계산은 귀무가설, 대립가설, 유의수준, 검정력을 입력하고 (그림 3.1(a)) 계산하면 표본 크기가 나타난다 (그림 3.1(b)).

<p style="text-align: center;">2. two-sided</p> <p>유의수준(α)와 검정력(β), 귀무가설($p=p_0$), 대립가설($p=p_1$)를 알고 있을 경우의 양측 표본수 계산</p> <p>유의 수준(α) <input style="width: 50px;" type="text" value="0.05"/></p> <p>검정력(β) <input style="width: 50px;" type="text" value="0.9"/></p> <p>귀무가설($p=p_0$) <input style="width: 50px;" type="text" value="0.3"/></p> <p>대립가설($p=p_1$) <input style="width: 50px;" type="text" value="0.4"/></p> <p style="text-align: center;"><input type="button" value="표본수 계산"/></p>	<p style="text-align: center;">단일 표본수 계산(two-sided)</p> <p style="font-size: small;">유의 수준(α) = 0.05, 검정력(β) = 0.9, $p_0 = 0.3$, $p_1 = 0.4$인 경우</p> <p style="text-align: center;">계산된 결과는 232.8669 입니다.</p> <p style="text-align: center;">따라서 표본수는 233 개입니다.</p>
--	---

(a) 표본 크기 계산

(b) 표본 크기 계산 결과

그림 3.1 단일 표본 양측 검정에서 표본 크기 계산 및 결과

- 모비율 p 에 대한 양측 검정에서 검정력 계산은 귀무가설, 대립가설, 유의수준, 표본 크기를 입력하고 (그림 3.2(a)) 계산하면 검정력 나타난다 (그림 3.2(b)).

<p style="text-align: center;">2. two-sided</p> <p>유의수준(α)와 표본수(n), 귀무가설($p=p_0$), 대립가설($p=p_1$)를 알고 있을 경우의 양측 검정력 계산</p> <p>유의 수준(α) <input style="width: 50px;" type="text" value="0.05"/></p> <p>표본수 (n) <input style="width: 50px;" type="text" value="191"/></p> <p>귀무가설($p=p_0$) <input style="width: 50px;" type="text" value="0.3"/></p> <p>대립가설($p=p_1$) <input style="width: 50px;" type="text" value="0.4"/></p> <p style="text-align: center;"><input type="button" value="검정력 계산"/></p>	<p style="text-align: center;">검정력 계산(two-sided)</p> <p style="font-size: small;">유의 수준(α) = 0.05, 표본수 = 191, $p_0 = 0.3$, $p_1 = 0.4$인 경우</p> <p style="text-align: center;">계산된 결과는 0.8383 입니다.</p> <p style="text-align: center;">따라서 검정력은 약 0.84 입니다.</p>
---	--

(a) 검정력 계산

(b) 검정력 계산 결과

그림 3.2 단일 표본 양측 검정에서 검정력 계산 및 결과

비율에 대한 일표본 단측검정인 경우 대립가설의 방향에 상관없이 표본 크기 및 검정력은 식 (3.1) 및 식 (3.2)에서 $\alpha/2$ 대신 α 를 사용하며, 사용법은 양측검정과 마찬가지로이다. 구현 결과는 위와 같은 사이트에서 제공된다.

3.2. 독립인 두 표본의 모비율 차 $p_1 - p_2$ 에 대한 검정

X_1, X_2, \dots, X_{n_1} 이 서로 독립이며, 이항분포 $B(n_1, p_1)$ 에서 얻어진 표본이고, Y_1, Y_2, \dots, Y_{n_2} 가 서로 독립이며, 이항분포 $B(n_2, p_2)$ 에서 얻어진 표본이라고 하자. 또한 X_i 와 Y_j 들도 서로 독립이라 가정한다. 귀무가설 $H_0 : p_1 - p_2 = 0$ 과 대립가설 $H_1 : p_1 - p_2 \neq 0$ 인 모비율 차에 대한 검정에서 유의수준 α , 대립가설 하에서 각 표본의 성공확률이 각각 p_{1a} 및 p_{2a} 일 때의 검정력이 최소한 β 가 되기 위해 필요한 표본 크기는

$$n = \frac{\left(z_{\alpha/2} \sqrt{2p_{10}(1-p_{10})} + z_{1-\beta} (\sqrt{p_{1a}(1-p_{1a})} + \sqrt{p_{2a}(1-p_{2a})}) \right)^2}{(p_{1a} - p_{2a})^2} \tag{3.3}$$

이며 (p_{10} 는 두 그룹의 공통 성공확률), 검정력 β 는

$$\beta = \Pr \left\{ Z > \frac{-z_{\alpha/2} \sqrt{2p_{10}(1-p_{10})} + \sqrt{n(p_{1a} - p_{2a})^2}}{\sqrt{p_{1a}(1-p_{1a})} + \sqrt{p_{2a}(1-p_{2a})}} \right\} \tag{3.4}$$

이다.

독립인 두 표본 모비율 차 $p_1 - p_2$ 에 대한 양측 검정에서 표본 크기와 검정력의 계산은

- 모비율 p 에 대한 양측 검정에서 표본 크기 계산은 귀무가설, 대립가설, 유의수준, 검정력을 입력하고 (그림 3.3(a)) 계산하면 표본 크기가 나타난다 (그림 3.3(b)).

<p style="text-align: center;">2. two-sided</p> <p>유의수준(α)와 검정력(β), 두 그룹의 공통의 성공확률($p=p_0$), 그룹 1의 성공확률($p=p_1$), 그룹 2의 성공확률($p=p_2$)를 알고 있을 경우의 양측 표본수 계산</p> <p>유의 수준(α) <input type="text" value="0.05"/></p> <p>검정력(β) <input type="text" value="0.9"/></p> <p>두 그룹의 공통의 성공확률($p=p_0$) <input type="text" value="0.3"/></p> <p>그룹 1의 성공확률($p=p_1$) <input type="text" value="0.4"/></p> <p>그룹 2의 성공확률($p=p_2$) <input type="text" value="0.5"/></p> <p style="text-align: center;"><input type="button" value="표본수 계산"/></p>	<p style="text-align: center;">독립인 두 표본수 계산(two-sided)</p> <p>유의 수준(α) = 0.05, 검정력(β) = 0.9, $p_0 = 0.3$, $p_1 = 0.4$인 경우, $p_2 = 0.5$인 경우</p> <p>계산된 결과는 644.5542입니다.</p> <p>따라서 각 집단의 표본수는 645개입니다.</p>
---	--

(a) 표본 크기 계산

(b) 표본 크기 계산 결과

그림 3.3 독립인 두 표본 양측 검정에서 표본 크기 계산 및 결과

- 모비율 p 에 대한 양측 검정에서 검정력 계산은 귀무가설, 대립가설, 유의수준, 표본 크기를 입력하고 (그림 3.4(a)) 계산하면 검정력 나타난다 (그림 3.4(b)).

<p style="text-align: center;">2. two-sided</p> <p>유의수준(α)와 검정력(β), 두 그룹의 공통의 성공확률($p=p_0$), 그룹 1의 성공확률($p=p_1$), 그룹 2의 성공확률($p=p_2$)를 알고 있을 경우의 양측 검정력 계산</p> <p>유의 수준(α) <input type="text" value="0.05"/></p> <p>각 그룹의 표본수 <input type="text" value="645"/></p> <p>두 그룹의 공통의 성공확률($p=p_0$) <input type="text" value="0.3"/></p> <p>그룹 1의 성공확률($p=p_1$) <input type="text" value="0.4"/></p> <p>그룹 2의 성공확률($p=p_2$) <input type="text" value="0.5"/></p> <p style="text-align: center;"><input type="button" value="검정력 계산"/></p>	<p style="text-align: center;">검정력 계산(two-sided)</p> <p>유의 수준(α) = 0.05, 각 집단의 표본수 = 645, $p_0 = 0.3$, $p_1 = 0.4$, $p_2 = 0.5$인 경우</p> <p>계산된 결과는 0.9002입니다.</p> <p>따라서 검정력은 약 0.9입니다.</p>
---	---

(a) 검정력 계산

(b) 검정력 계산 결과

그림 3.4 독립인 두 표본 양측 검정에서 검정력 계산 및 결과

비율차에 대한 단측검정인 경우 대립가설의 방향에 상관없이 표본 크기 및 검정력은 식 (3.3) 및 식 (3.4)에서 $\alpha/2$ 대신 α 를 사용하며, 사용법은 양측검정과 마찬가지로이다. 구현 결과는 위와 같은 사이트에서 제공된다.

4. 결론

표본 (sample)의 특성 또는 표본이 내포하고 있는 정보를 바탕으로 모집단 (population)에 관한 통계적 추론을 하게 된다. 이때 표본이 모집단에 대한 대표성을 갖기 위해서는 신뢰할 수 있고 정밀성을 제공하는 표본 크기와 검정력의 계산이 필요하다. 만일 표본이 너무 작으면 모집단의 중요한 정보를 잃을 수 있고, 반대로 표본이 너무 크면 시간이나 비용 등을 낭비할 수 있다. 또한 사회 여러 분야에서 실험 결과의 일관성 (consistency)을 유지하는 방법으로 메타 분석 (meta analysis)이 사용되고 있으며, 이때 여러 조건에 맞는 표본 크기 계산이 반드시 필요하다.

이에 본 연구는 실험결과의 신뢰성과 정밀성을 제공하는 표본 크기와 검정력 계산을 소개하였다. 여기서 모분산 σ^2 를 알고 있다는 가정 하에 가설을 입력하거나 Cohen (1988)이 제시한 effect size를 이용하여 평균 검정에서 단측 검정과 양측 검정의 표본 크기와 검정력이 계산되도록 구현하였고, 비율 검정도 가설에 대한 단측 검정과 양측 검정의 표본 크기와 검정력이 웹에 계산되도록 구현하였다.

참고문헌

- 김진경, 박진호, 박헌진, 이재준, 전홍석, 황진수 (2008). <통계학-엑셀을 이용한 분석>, 자유아카데미, 경기.
- 루크 웰링 (2009). <(성공적인 웹 프로그래밍) PHP와 MySQL>, 정보문화사, 서울.
- 정진호 (2000). <PHP Web-DB programming Guide>, 동일출판사, 서울.
- 최완규 (2002). <Professional PHP4>, 정보문화사, 서울.
- Cohen. J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd Ed., Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ.
- Kang, H. and Sim, S. (2000). Implementation of estimation and inference on the web. *Communications of the Korean Statistical Society*, **7**, 913-926.
- Kang, H. and Sim, S. (2003). Regression and correlation analysis via dynamic graphs. *Communications of the Korean Statistical Society*, **10**, 695-705.
- Ko, H. and Kim, D. (2010). Sample size comparison for two independent populations. *Journal of the Korean Data & Information Science Society*, **21**, 1243-1251.
- Kwak, S. G., Kim, D. H., Shin, I. H., Kim, H. G. and Kim, S. G. (2010). Two Bayesian methods for sample size determination in clinical trials. *Journal of the Korean Data & Information Science Society*, **21**, 1343-1351.
- Rosner, B. (2010). *Fundamentals of biostatistics*, 7/e, Brooks/Cole Cengage Learning, Boston, MA.
- Sim, S. and Lee, K. W. (1999). Approximation of binomial distribution via dynamic graphics. *Communications of the Korean Statistical Society*, **6**, 821-829.

An implementation of the sample size and the power for testing mean and proportion

Changsun Lee¹ · Heemo Kang² · Songyong Sim³

¹²³Department of Finance Information Statistics, Hallym University

Received 22 November 2011, revised 21 December 2011, accepted 26 December 2011

Abstract

There are cases when the sample size is determined based not only on the significance level but also on on the power or type II error. In this paper, we implemented the sample size and the power calculation when both the significance level and power for testing means in normal distributions and proportions in binomial distributions. The implementation is available on a web site. Alternately, we also calculate the power for a given effect size, type I error probability and sample size.

Keywords: Effect size, power, sample size.

¹ Graduate student, Department of Finance Information Statistics, Hallym University, Chuncheon, Kangwon 200-702, Korea.

² Adjunct associate professor, Department of Finance Information Statistics, Hallym University, Chuncheon, Kangwon 200-702, Korea.

³ Corresponding author: Professor, Department of Finance Information Statistics, Hallym University, Chuncheon, Kangwon 200-702, Korea. E-mail: sysim@hallym.ac.kr