

보건학 연구에서 요구되는 적정 표본크기 결정 방법

김 호

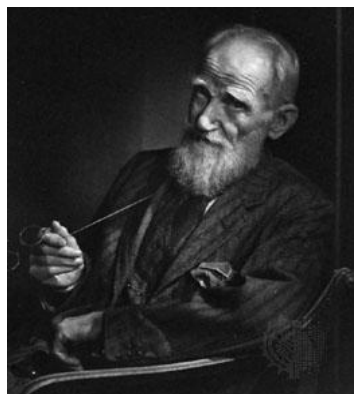
서울대학교 보건대학원

들어가며



- “세상에는 세가지 거짓말이 있다. 거짓말, 새빨간 거짓말, 그리고 통계.(There are three kinds of lies: lies, damned lies, and statistics)”

– 벤자민 디즈레일리(Benjamin Disraeli)



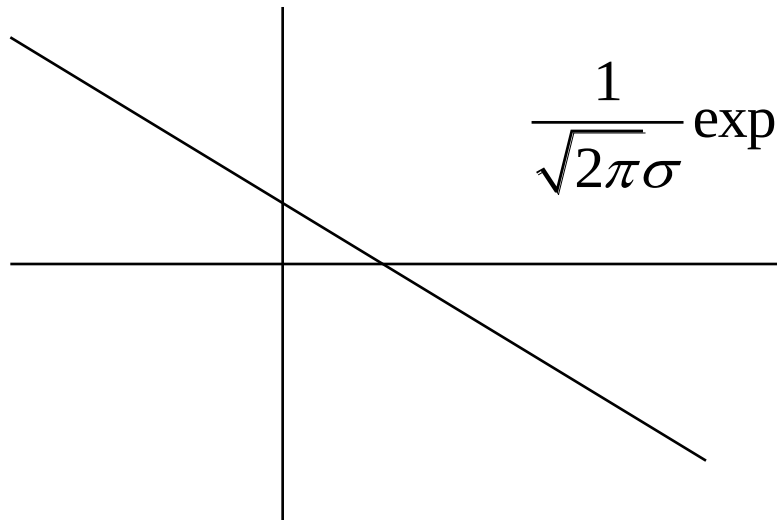
- “진정으로 교육받은 사람의 징표는 그가 통계에 의해 깊은 감동은 받았다는 것이다. (It is the mark of a truly intelligent person to be moved by statistics.)”

✦ 조지 버나드 쇼 (George Bernard Shaw)

Terminology

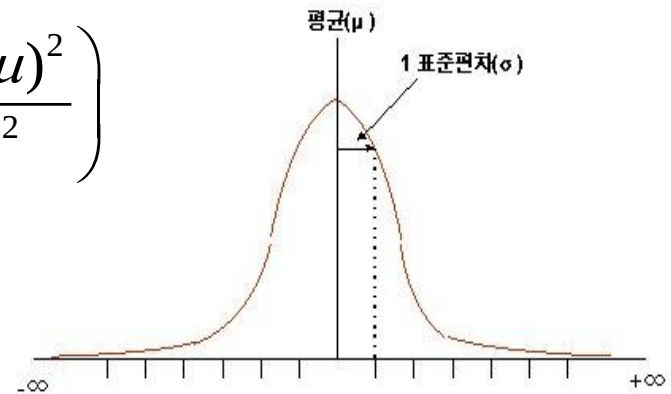
- **Parameter:** a constant in a model, or a constant that wholly characterizes function or a probability distribution

$$Y = a + b x$$



$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$N(\mu, \sigma^2)$$



- 관심모수 : 연구의 가설을 수학적인 모수로 표시해야 함

- 두 집단에서 평균비교 $d = \mu_1 - \mu_2$

- 두 집단의 비율비교 $r = \frac{p_1}{p_2}$

$$OR = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)}$$

- 연구가설

- 두 집단에서 평균비교

귀무가설

$$H_0 : d = \mu_1 - \mu_2 = 0$$

대립가설 (양측검정)

$$H_a : d = \mu_1 - \mu_2 \neq 0$$

대립가설 (단측검정)

$$H_a : d = \mu_1 - \mu_2 > 0$$

혹은

$$H_a : d = \mu_1 - \mu_2 < 0$$

- 연구가설

- 두 집단의 비율비교

귀무가설

$$H_0 : r = \frac{p_1}{p_2} = 1$$

$$H_0 : OR = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} = 1$$

대립가설 (양측검정)

$$H_0 : r = \frac{p_1}{p_2} \neq 1$$

$$H_0 : OR = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} \neq 1$$

대립가설 (단측검정)

$$H_0 : r = \frac{p_1}{p_2} > 1$$

$$H_0 : OR = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} > 1$$

혹은

$$H_0 : r = \frac{p_1}{p_2} < 1$$

$$H_0 : OR = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} < 1$$

Terminology

분위수 : quantile

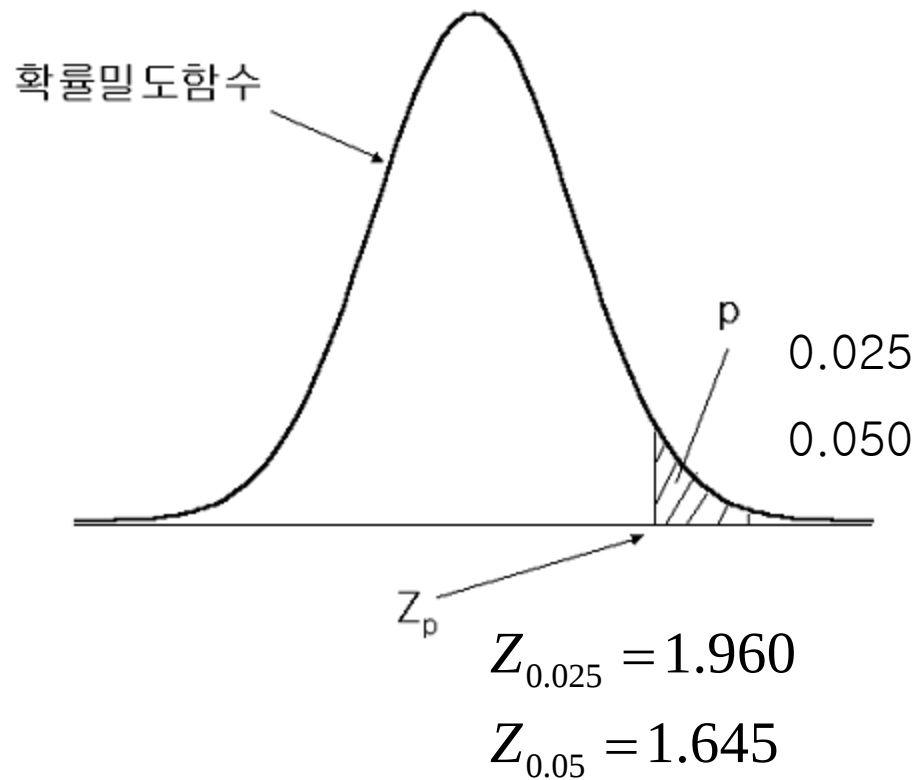


표 4. 표준정규분포 하에서의 Z값

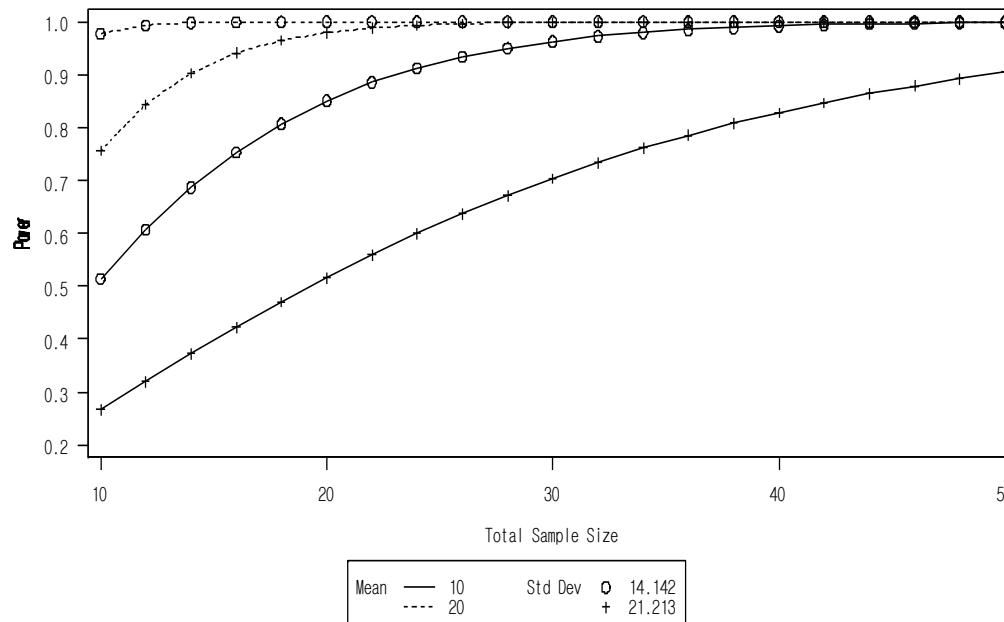
제1종 오류 (α)	Z_α		제2종 오류 (β)	검정력 (power)	Z_β
	단측	양측			
0.005	2.576	2.813			
0.010	2.326	2.576			
0.025	1.960	2.248			
0.050	1.645	1.960	0.05	0.95	1.645
0.100	1.282	1.645	0.10	0.90	1.282
0.200	0.842	1.282	0.20	0.80	0.842
0.300	0.524	1.036			
0.400	0.253	0.842			
0.500	0.000	0.674			

- $A(\text{얻은 자료}) \rightarrow B(\text{연구가설})$
- $\neg B \rightarrow \neg A$
- 귀무가설 ($\neg B$) : 두 집단에 차이가 없다. (H_0)
- 대립가설 (B) : 두 집단에 차이가 있다. (H_a)
- 일종의 오류 : 옳은 귀무가설을 기각할 확률
 $= \Pr(\text{reject } H_0 \mid H_0 \text{ is true}) \quad \alpha$
- 이종의 오류 : 틀린 귀무가설을 받아들이는 확률
 $= \Pr(\text{Not reject } H_0 \mid H_a \text{ is true}) \quad \beta$
- $\text{Power} = 1 - \beta$ (있는 차이를 발견할 확률)

❖검정방법 A(모수적인 방법, 검정력=90%)의 검정방법 B (비모수적인 방법, 검정력=70%)보다 더 큰 검정력을 주었다.

-> 실제로 차이가 있을 때 A 방법을 100번 실시했을 때 90번의 경우 차이가 있다고(귀무가설 기각) 결정하였고 B 방법을 실시하였을 때는 100번 중 70번 귀무가설을 기각하였다.

-> A가 더 좋은 방법 ! (실제차이=?)



Get Motivated <예시1>

	<i>Trt A</i>	<i>Trt B</i>	
<i>+</i>	n_{11} 52	n_{12} 48	n_{1+}
<i>-</i>	n_{21} 48	n_{22} 52	n_{2+}
	n_{+1}	n_{+2}	n_{++}

$$\chi^2 = \frac{(n_{11} - n_{+1}n_{1+} / n_{++})^2}{v_{11}}, \quad v_{11} = \frac{n_{1+}n_{2+}n_{+1}n_{+2}}{n_{++}^2 (n_{++} - 1)}$$

$$\chi^2 = \frac{(52 - 100 \times 100 / 200)^2}{(100 \times 100 \times 100 \times 100) / (200^2 \times 199)} = 0.32, \quad p > 0.05$$

- $n_{ij}^* = 100 \times n_{ij}$ 라고 하고, χ^2 를 다시 계산하면

$$\chi^{*2} = 100^2 / 100 \times \chi^2 = 32.00, p < 0.01$$

- 두 예에서 비율은 정확히 같음에도 불구하고 통계적 유의성은 상당히 다르다. ???
- 전통적 통계적 가설 검정의 유의성은 표본수에 크게 의존한다.
- 통계적 유의성이 없었던 경우라도 표본수를 크게 하면 유의성을 볼 수 있다.
- 표본수(실험의 비용)와 통계적 유의성(실험의 효용성)의 균형을 맞추는 것이 요구됨
- 최소의 비용으로 효과를 증명하고 싶다.

통계학에서의 표본수 계산

- 표본조사의 경우
 - 목적 : 추정 (estimation)
 - 도구 : 표본오차
 - 예 : 여론조사
- 임상시험의 경우
 - 목적 : 검정 (testing)
 - 도구 : 제1종의 오류, 제2종의 오류
 - 예 : 임상시험

- 단순임의 추출(simple random sampling)에서
- N : 모집단의 크기, n : 표본의 크기라면

$$\hat{\mu} = \bar{y} = \sum_{i=1}^n y_i / n$$

$$Var(\bar{y}) = \frac{\sigma^2}{n} \cdot \left(\frac{N-n}{N-1} \right)$$

$$1.96\sqrt{Var(\bar{y})} \approx 2\sqrt{\frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}} = B : 95\% \text{ 신뢰구간 (표준오차)}$$

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2}, \quad D = B^2 / 4$$

- 만약 y_i 가 0 혹은 1의 값을 가지게 되면,
 $\bar{y}$ 는 비율이 되고,
 이 경우 $n = \frac{Npq}{(N-1)D + pq}$ 가 된다.

예1) $N=2000$, 95% 신뢰수준, $B=0.05$ 이라면 n 은 ?

>> 사전정보가 없다면 $p=q=0.5$ 대입

$$D = B^2 / 4 = 0.05^2 / 4 = .000625$$

$$n = \frac{2000 \times 0.5 \times 0.5}{1999 \times .000625 + 0.5 \times 0.5} = 333.56$$

최소한 334명의 표본이 필요하다.

예제. 인구 10,000,000인 도시에서 유병율이 약 20%로 예상되는 질병의 정확한 유병율을 95% 신뢰구간이 $\pm 2\%$ 정도로 추정하고 싶다면 표본의 크기를 얼마나 하여야 하는가 ?

$$D = B^2 / 4 = 0.02^2 / 4 = .0001$$

$$n = \frac{10,000,000 \times 0.2 \times 0.8}{(10,000,000 - 1) \times .0001 + 0.2 \times 0.8} = 1599.74$$

예제. 인구 500,000인 지역에서 유병율이 약 20%로 예상되는 질병의 정확한 유병율을 95% 신뢰구간이 $\pm 2\%$ 정도로 추정하고 싶다면 표본의 크기를 얼마나 하여야 하는가 ?

$$D = B^2 / 4 = 0.02^2 / 4 = .0001$$

$$n = \frac{500,000 \times 0.2 \times 0.8}{(500,000 - 1) \times .0001 + 0.2 \times 0.8} = 1594.90$$

예제. 인구 10,000,000인 도시에서 유병율이 약 20%로 예상되는 질병의 정확한 유병율을 95% 신뢰구간이 $\pm 1\%$ 정도로 추정하고 싶다면 표본의 크기를 얼마로 하여야 하는가 ?

$$D = B^2 / 4 = 0.01^2 / 4 = .000025$$

$$n = \frac{10,000,000 \times 0.2 \times 0.8}{(10,000,000 - 1) \times .0001 + 0.2 \times 0.8} = 6395.91$$

예제. 인구 10,00,000인 지역에서 유병율이 약 30%로 예상되는 질병의 정확한 유병율을 95% 신뢰구간이 $\pm 3\%$ 정도로 추정하고 싶다면 표본의 크기를 얼마로 하여야 하는가 ?

$$D = B^2 / 4 = 0.03^2 / 4 = .000225$$

$$n = \frac{10,00,000 \times 0.3 \times 0.7}{(10,000,000 - 1) \times .000225 + 0.3 \times 0.7} = 933.246$$

연구대상수의 계산 시 고려할 사항

- 어떠한 연구설계인가? (연구설계의 형태)
- 처치 후 결과는 언제 나타나는가?
(처치효과의 지연기간)
- 처치의 결과는 이산형인가 연속형인가?
(결과의 척도)
- 연구기간을 얼마나 길게 할 것인가?
(예상되는 추적기간)
- 중도에 탈락하는 사람은 얼마나 될 것인가?
(예상되는 추적손실)
- 연구자의 지시대로 치료를 잘 받는가?
(예상되는 처치의 비순응율)
- 할당비는 얼마로 할 것인가? (시험군 : 비교군)
- 시험집단의 수는 몇 개로 할 것인가?

연구설계의 형태

1) 고정형 연구설계 (Fixed sample size design)

임상시험을 시작하기 전에 연구자가 필요한 연구대상수를 확정하는 방법.

2)연속형 연구설계(Sequential design)

시험군과 비교군의 관찰된 결과의 차이가 미리 설정된 유의한 범위를 벗어날 때까지 계속해서 연구대상을 모집하여 진행하는 방법.

결과변수의 척도

- 연속형 변수

- 비척도(**Ratio Scale**) : 혈당량, 혈압, 적혈구수, 무게(mg, lb등), 부피(cc, m³ 등), 비율, 시간의 길이
- 간격척도(**Interval scale**) : 온도(°C, °F), Circular scale(1년, 24시간, 방위)

- 범주형변수

- 순위척도(**Ordinal scale**) : 암의 진행정도(제1기, 제2기, 제3기), 당뇨의 정도(+, ++, +++) 등
- 명목척도(**Nominal scale**): 생사여부, 질병발생여부, 종교의 종류 등

가설 설정

- 1) 귀무가설(null hypothesis, H_0) : 시험의 효과가 없다
- 2) 대립가설(alternative hypothesis, H_a) 증명하고자 하는 가설.

예) 새로 개발된 항암제의 치료율(P_t)이 기존 약의 치료율(P_c) 보다 높을 것이라는 연구목적을 실험하고자 할 때

$$(H_a : P_t > P_c, \text{ 또는 } P_t - P_c > 0, \text{ 단측가설})$$

새로 개발된 소염제와 기존 소염제의 치료 후 관절염의 재발율에 차이가 있는 지를 보고자 할 때

$$(H_a : P_t \neq P_c, \text{ 또는 } P_t - P_c \neq 0, \text{ 양측가설})$$

- 동일한 α 와 β 를 사용하였을 경우 단측보다도 양측의 경우가 더 많은 연구대상수를 필요

- 표2. 가설검정시 발생 가능한 4가지 상황

표본을 이용한 가설검정결과	모집단의 진실	
	H_0 참	H_0 거짓 (H_a 참)
H_0 채택	옳음 ($1-\alpha$)	제2종 오류(β)
H_0 기각	제1종 오류(α)	옳음 (검정력= $1-\beta$)

오류의 비교

- H_0 : 두 약의 효과의 차이가 없다.
- H_a : 두 약의 효과에 차이가 있다.
- 제 1 종의 오류 = $P(\text{차이가 있다.} \mid \text{실제는 차이가 없다})$
- 제 2 종의 오류 = $P(\text{차이가 없다.} \mid \text{실제는 차이가 있다})$

• 표3. 제1종 오류와 제 2종 오류의 설정기준

α 와 β 의 설정	적용기준
$\alpha = \beta$	치료군과 비교군이 모두 새로운 처치 방법이며, 처치비용이 비슷하고, 모두 안전하다고 생각할 만한 충분한 이유가 있을 때.
$\alpha > \beta$	치료군의 비용이 비교적 저렴하고, 적용하기 쉬우며, 알려져 있는 심각한 부작용이 없고 또한 비교군의 처치 방법이 없는 경우.
$\alpha < \beta$	치료법이 새로운 방법이고, 비용이 많이 들고, 심각한 부작용을 일으키는데 반하여 비교군의 처치방법은 이미 널리 사용되고 있으며, 안전하고 효과적이라고 알려져 있는 경우.

이분형 결과에 대한 비교 (비율을 비교하는 경우)

- 두 비율의 비교

$$H_0 : P_1 = P_2 \text{ (} P_1 < P_2 \text{)} \text{ vs. } H_a : P_1 > P_2$$

$$\text{검정통계량 } Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{2\bar{p}\bar{q}/n}}, \quad \bar{p} = (\hat{p}_1 + \hat{p}_2)/2, \quad \bar{q} = 1 - \bar{p}$$

Reject H_0 if $Z > Z_\alpha$

P_1, P_2, Q_1, Q_2 : 실제 관심이 있는 미지의 고정된 수(모수)

$\hat{p}_1, \hat{p}_2, \hat{q}_1, \hat{q}_2$: 모수를 추정하기 위한 추정치

$$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

$$= P\left(Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{2\bar{p}\bar{q}/n}} > Z_\alpha \mid P_1 = P_2 \right)$$

$$= P(Z > Z_\alpha \mid Z \sim N(0,1))$$

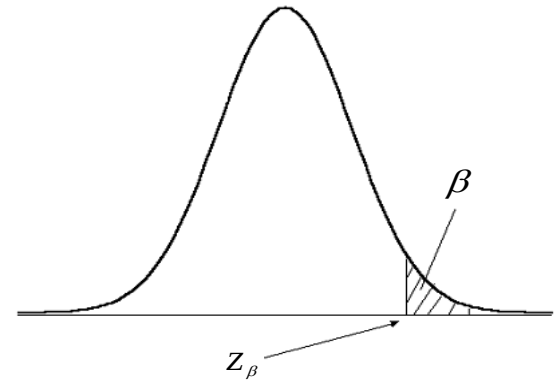
$$\text{power} = P(\text{Reject } H_0 \mid H_a \text{ is true})$$

$$= P\left(Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{2\bar{p}\bar{q}/n}} > Z_\alpha \mid P_1 - P_2 = d \right)$$

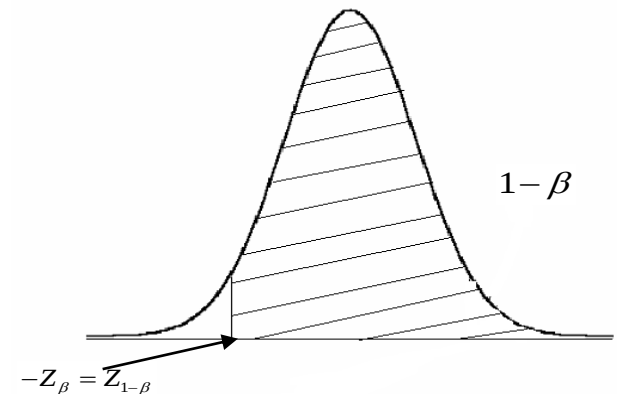
$$E(\hat{p}_1 - \hat{p}_2 \mid H_a) = d$$

$$\text{Var}(\hat{p}_1 - \hat{p}_2 \mid H_a) = (P_1Q_1 + P_2Q_2)/n$$

$$\begin{aligned}
& * P \left(\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{2\bar{p}\bar{q}/n}} > Z_\alpha \mid d > 0 \right) \\
& = P(\hat{p}_1 - \hat{p}_2 - (P_1 - P_2) > Z_\alpha \sqrt{2\bar{p}\bar{q}/n} - (P_1 - P_2) \mid d > 0) \\
& = P \left(\frac{\hat{p}_1 - \hat{p}_2 - (P_1 - P_2)}{\sqrt{(P_1Q_1 + P_2Q_2)/n}} > \frac{Z_\alpha \sqrt{2\bar{p}\bar{q}/n} - (P_1 - P_2)}{\sqrt{(P_1Q_1 + P_2Q_2)/n}} \mid d > 0 \right) \\
& = p \left(Z > \frac{Z_\alpha \sqrt{2\bar{p}\bar{q}/n} - (P_1 - P_2)}{\sqrt{(P_1Q_1 + P_2Q_2)/n}} \mid d > 0 \right) \\
& = \text{Power} = 1 - \beta
\end{aligned}$$



$$\begin{aligned}
& * \frac{Z_\alpha \sqrt{2\bar{p}\bar{q}/n} - (P_1 - P_2)}{\sqrt{(P_1Q_1 + P_2Q_2)/n}} = -Z_\beta \\
& n = \frac{\left(Z_\alpha \sqrt{2\bar{p}\bar{q}} + Z_\beta \sqrt{(P_1Q_1 + P_2Q_2)} \right)^2}{d^2}
\end{aligned}$$



예) 신약의 부작용 발생률이 기존 약의 부작용 발생률(30%로 알려져 있다고 하자)보다 5% 이상 작다는 것을 증명하고자 하는 연구에서의 적절한 표본수는 ? 6개월 동안 추적손실은 5%이며 환자의 순응도는 90%라고 가정한다. 일종의 오류=5%, 검정력=80%

$$Z_{\alpha} = 1.645, Z_{\beta} = 0.840 (\text{단측검정})$$

$$\bar{P} = (0.30 + 0.25) / 2 = 0.275 \text{ 이며}$$

$$\left(\frac{1.645\sqrt{2 \times 0.275 \times 0.725} + 0.840\sqrt{0.3 \times 0.7 + 0.25 \times 0.75}}{0.05} \right)^2 = 983.90$$

표5. 두집단의 비율 비교시의 표본수표(단형할당의 경우 각 집단별 표본수)

upper number : $\alpha = 0.05$ (one-tailed) or $\alpha = 0.10$ (two-tailed) ; $\beta = 0.20$

middle number : $\alpha = 0.025$ (one-tailed) or $\alpha = 0.05$ (two-tailed) ; $\beta = 0.20$

lower number : $\alpha = 0.005$ (one-tailed) or $\alpha = 0.01$ (two-tailed) ; $\beta = 0.20$

Smaller of P1 and P2 ^a	Expected difference between P1 and P2									
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
.05	342	110	59	38	27	21	17	13	11	9
	434	140	75	49	35	27	21	17	14	12
	581	187	100	65	46	35	28	22	19	15
.10	539	156	78	48	33	25	19	15	12	10
	685	199	99	62	43	31	24	19	16	13
	916	266	133	82	56	42	32	25	21	17
.15	712	197	95	57	38	287	21	16	13	11
	904	250	120	72	49	35	27	21	17	14
	1210	334	161	96	65	47	35	28	22	18
.20	860	231	108	64	42	30	23	17	14	11
	1093	293	138	81	54	38	29	22	18	14
	1462	392	184	108	72	51	38	29	23	19
.25	984	258	119	69	45	32	24	18	14	11
	1249	328	152	88	58	41	30	23	18	14
	1672	439	203	117	77	54	40	30	24	19
.30	1083	280	128	73	47	33	24	18	14	11
	1375	356	162	93	60	42	31	23	18	14
	1840	476	217	124	80	56	41	31	24	19
.35	1157	295	133	75	48	33	24	18	14	11
	1469	375	169	96	61	42	31	23	18	14
	1966	502	226	128	82	56	41	30	23	18
.40	1206	305	136	76	48	33	24	17	13	10
	1532	387	173	97	61	42	30	22	17	13
	2050	518	231	129	82	56	40	29	22	17

추적 손실과 순응도를 고려하면 각 처리군 별로

$$n = \frac{984}{(1-0.05) \times 0.90} = 1150.88 \cong 1151 \text{ 의 표본수를 얻는다.}$$

```
> power.prop.test(power=.80,p1=.30,p2=.25,  
alternative=c("one.sided"))
```

Two-sample comparison of proportions power calculation

n = 985.071

p1 = 0.3

p2 = 0.25

sig.level = 0.05

power = 0.8

alternative = one.sided

NOTE: n is number in *each* group

```
proc power;  
    twosamplefreq test=pchi  
    proportiondiff = 0.05  
    refproportion = 0.25  
    power=0.80  
    sides=1  
    ntotal=. ;  
run;
```

The POWER Procedure

Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements

Distribution	Asymptotic normal
Method	Normal approximation
Number of Sides	1
Reference (Group 1) Proportion	0.25
Proportion Difference	0.05
Nominal Power	0.8
Null Proportion Difference	0
Alpha	0.05
Group 1 Weight	1
Group 2 Weight	1

Computed N Total

Actual	N
Power	Total
0.800	1972

이형할당인 경우

- 만약 $n_t = \lambda n_c$ 이라면 $\bar{P} = (P_c + \lambda P_t) / (1 + \lambda)$ 이라고 하고

$$n_c = \frac{\left(Z_\alpha \sqrt{\bar{P}\bar{Q}(\lambda+1)/\lambda} + Z_\beta \sqrt{P_c Q_c + P_t Q_t / \lambda} \right)^2}{\Delta_A^2}$$

- $\lambda = 1$ 이면 단형할당과 동일

연속형 변수의 비교

예) 새로운 관절염 치료제의 치료효과에 대한 임상실험을 실시한다고 하자.

치료효과는 2주간 치료 후 혈중 Prostag-landing 양이 평균 10, 표준편차 2 이면 치료가 된 것으로 간주한다. 치료 후 두 집단의 혈중 Prostaglandin 양의 변동이 20% 미만이면 두 치료제의 효과는 동등한 것으로 간주한다. 단측 검정으로 연구 대상수를 구하시오. 또 동일한 가정으로 양측검정의 결과와 비교하시오.

결과의 척도 :Prostaglandin농도 (연속형)

$$n_c = \frac{2(Z_\alpha + Z_\beta)^2 \sigma^2}{(\mu_c - \mu_t)^2}$$

$$\Delta_A = \mu_1 - \mu_2 = 2\mu gm / dl(10 \times 0.2) , \sigma = 2.0\mu gm / dl$$

$$Z_\alpha = 1.645 , Z_\beta = 1.282$$

$$n_t = n_c = \frac{2 \times (1.645 + 1.282)^2 \times 2.0^2}{2^2} = 17.13 \cong 18$$

$$\text{Effective Size } (=E/S) = (\mu_1 - \mu_2) / \sigma$$

• 표 6. 두 집단의 평균 비교시의 표본수표
(단형할당의 경우 각 집단별 표본수)

One-tailed $\alpha =$ Two-tailed $\alpha =$	0.005 0.01			0.025 0.05			0.05 0.10		
$\beta =$ E/S ^a	0.05	0.10	0.20	0.05	0.10	0.20	0.05	0.10	0.20
.10	3563	2977	2377	2599	2102	1570	2165	1713	1237
.15	1584	1323	1038	1155	934	698	962	762	550
.20	891	744	584	650	526	393	541	428	309
.25	570	476	374	416	336	251	346	274	198
.30	396	331	260	289	234	174	241	190	137
.40	223	186	146	162	131	98	135	107	77
.50	143	119	93	104	84	63	87	69	49
.60	99	83	65	72	58	44	60	48	34
.70	73	61	48	53	43	32	44	35	25
.80	56	47	36	41	33	25	34	27	19
.90	44	37	29	32	26	19	27	21	15
1.00	36	30	23	26	21	16	22	<u>17</u>	12

- 표6을 이용하면 $E/S=2/2=1$, $n=17$
한편 양측검정을 실시하면 $z_{\alpha}=1.960$ 이므로

$$n_t = n_c = \frac{2 \times (1.960 + 1.282)^2 \times 2.0^2}{2^2} = 21.02 \cong 21$$

표6를 이용하면 $n=21$

```
> power.t.test(delta=2,sd=2,sig.level=0.05,power=0.9)
```

Two-sample t test power calculation

n = 22.02110

delta = 2

sd = 2

sig.level = 0.05

power = 0.9

alternative = two.sided

NOTE: n is number in *each* group

```
>
```

```
power.t.test(delta=2,sd=2,sig.level=0.05,power=0.9,altern  
ative=c("one.sided"))
```

Two-sample t test power calculation

n = 17.84713

delta = 2

sd = 2

sig.level = 0.05

power = 0.9

alternative = one.sided

NOTE: n is number in *each* group

```
proc power;  
    twosamplemeans test=diff  
    meandiff = 2  
    stddev = 2  
    power=0.90  
    sides=1  
    npergroup=. ;  
run;
```

The POWER Procedure

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution	Normal
Method	Exact
Number of Sides	1
Mean Difference	2
Standard Deviation	2
Nominal Power	0.9
Null Difference	0
Alpha	0.05

Computed N Per Group

Actual	N Per
Power	Group
0.902	18

```
proc power;
```

```
    twosamplemeans test=diff
```

```
    meandiff = 2
```

```
    stddev = 2
```

```
    power=0.90
```

```
    sides=2
```

```
    npergroup=. ;
```

```
run;
```

The POWER Procedure

Two-sample t Test for Mean Difference

Fixed Scenario Elements

Distribution	Normal
Method	Exact
Number of Sides	2
Mean Difference	2
Standard Deviation	2
Nominal Power	0.9
Null Difference	0
Alpha	0.05

Computed N Per Group

Actual	N Per
Power	Group
0.912	23

One-sample and paired tests

```
> power.t.test(n=10,delta=10, sd=10*sqrt(2),type="paired")
```

Paired t test power calculation

n = 10

delta = 10

sd = 14.14214

sig.level = 0.05

power = 0.514318

alternative = two.sided

$$PRE \sim N(d, 10^2)$$

$$POST \sim N(d + 10, 10^2)$$

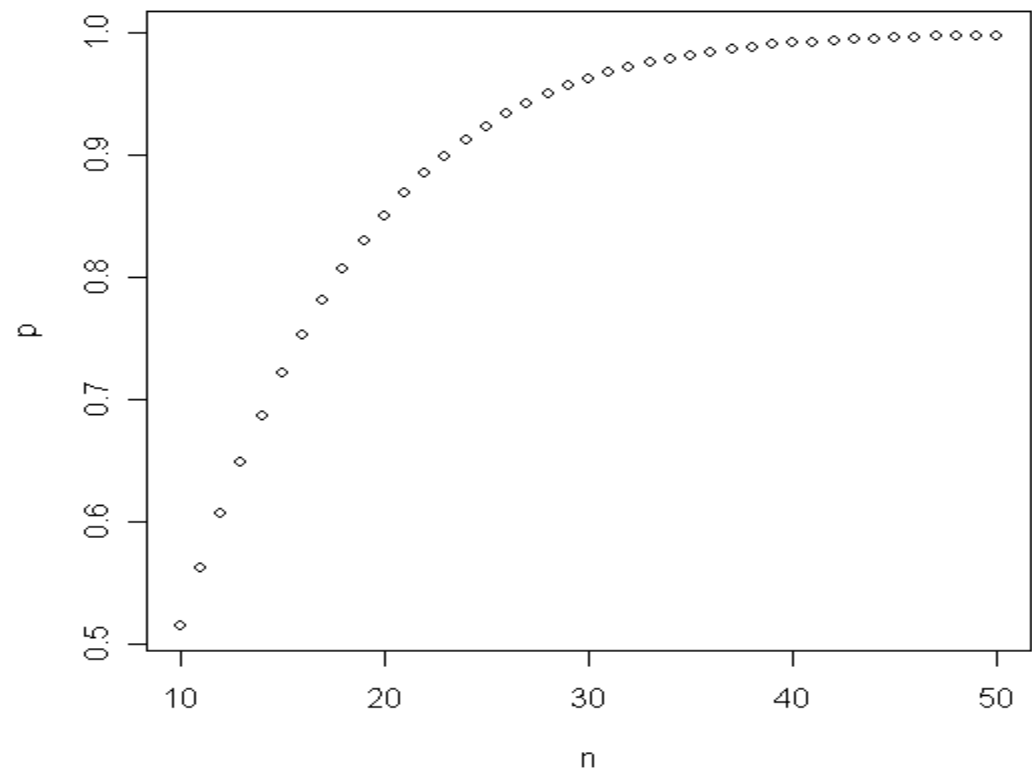
$$POST - PRE \sim N(10, (\sqrt{2} \times 10)^2)$$

NOTE: n is number of *pairs*, sd is std.dev. of *differences* within pairs

```
> n=10:50
```

```
> p=power.t.test(n,delta=10, sd=10*sqrt(2),type="paired")$power
```

```
> plot(n,p)
```



```
proc power;  
    onesamplemeans  
    means = 10  
    stddev =14.142  
    power=.  
    ntotal=10 ;  
run;
```

The POWER Procedure

One-sample t Test for Mean

Fixed Scenario Elements

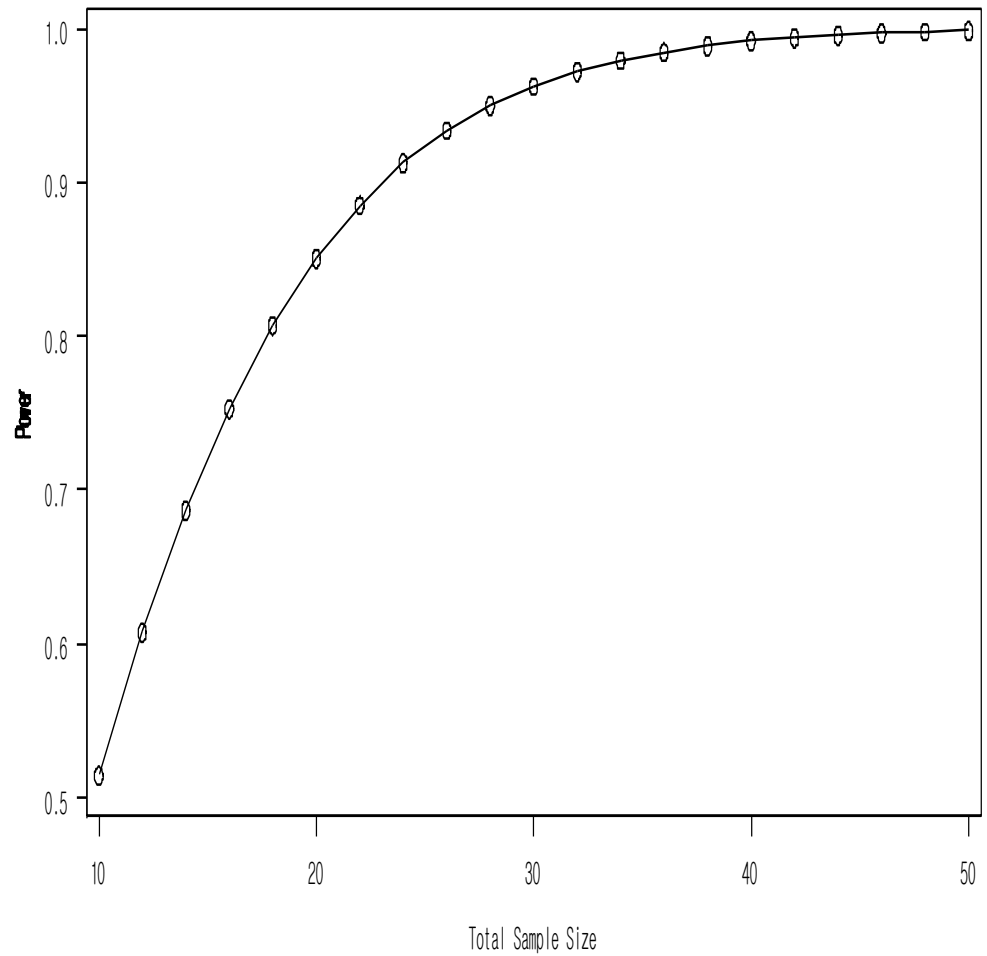
Distribution	Normal
Method	Exact
Mean	10
Standard Deviation	14.142
Total Sample Size	10
Number of Sides	2
Null Mean	0
Alpha	0.05

Computed Power

Power

0.514

```
proc power;  
  onesamplemeans  
  means = 10  
  stddev =14.142  
  power=.  
  ntotal=10 ;  
  plot x=n min=10 max=50 ;  
run;
```



```

proc power;

  onesamplemeans

  means = 10 20

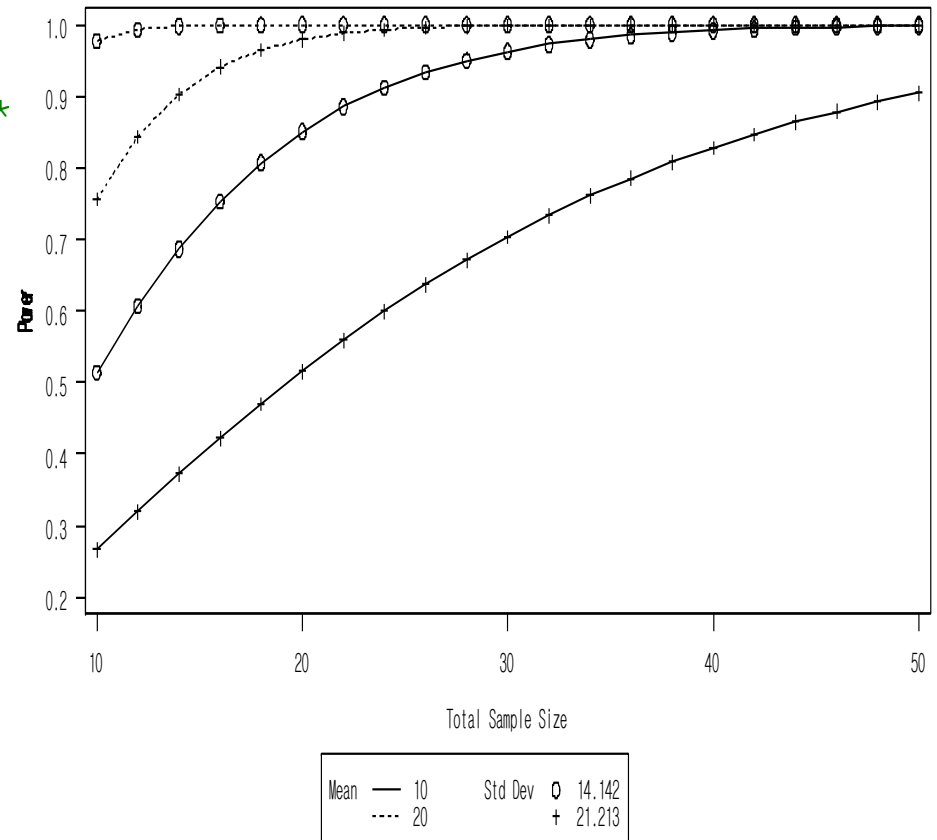
  stddev =14.142 21.213 /*
  power=.

  ntotal=10 ;

  plot x=n min=10 max=50 ;

run;

```



시간형 결과의 비교 (생존분석 이용) 모수적 방법 (지수분포 이용)

- 총 사건수 $D = 4 \frac{(Z_{\alpha} + Z_{\beta})^2}{(\log(\theta))^2}$, $\Delta =$ 위험비
- 예) 검정력 90%를 가지고 위험비 2.0을 95% 유의수준의 양쪽 검정으로 유의하게 감지할 수 있는 검정을 위해서는
는
$$4 \frac{(1.96 + 1.282)^2}{(\log(2.0))^2} = 87.50 \cong 88 \text{ event}$$
가 필요하다.

즉 실험군 비교군 각각 44개의사건이 필요하게 된다.
표7 과 동일한 결과

시간형 결과의 비교 (생존분석 이용) 비모수적 방법 (로그순위 검정 이용)

- 단형할당의 경우 총 사건수

$$D = 4 \frac{(Z_{\alpha} + Z_{\beta})^2 (\theta + 1)^2}{(\theta - 1)^2}$$

- 예) 검정력 90%를 가지고 위험비 1.75를 95% 유의수준의 양쪽 검정으로 위험비 1로부터 유의하게 감지할 수 있는 검정을 위해서는

$$D = \frac{(1.96 + 1.282)^2 (1.75 + 1)^2}{(1.75 - 1)^2} = 141 \text{ event} < \text{표 7} >$$

- 만약 30%가 실험 종료까지 사건이 발생하지 않는다면 (event free) (중도절단비, censoring rate=30%), 총표본수는 $141/0.30=202$ 로 주어진다.

- 표7. 로그 순위 검정을 이용한 비교를 위해 필요한 (두 집단) 총표본수 표 (위:모수적, 아래:비모수적 방법)

Δ	양측	$\beta=0.1$		$\beta=0.2$	
		$\alpha=0.05$	$\alpha=0.10$	$\alpha=0.05$	$\alpha=0.10$
1.25		844	688	630	496
		852	694	636	500
1.50		256	208	192	150
		262	214	196	154
1.75		134	110	100	80
		<u>142</u>	116	106	84
2.00		<u>88</u>	72	66	52
		94	78	70	56
2.25		64	52	48	38
		72	58	54	42
2.50		50	42	28	30
		58	48	42	34

연구대상수를 최소화하는 방법

1) 연속변수를 사용한다.

예를 들면 혈압을 연속형(mmHg) vs. 이분형(고혈압 여부)

2) 변수를 보다 정밀하게 측정한다.

분산 줄이기

3) 짝 측정치를 사용하라.

4) 이형 할당법을 사용한다.

5) 보다 일반적인 결과를 사용한다.

참고문헌

- 1)Piantadosi, S. (1997), Clinical Trials, A methodological perspectives, John Wiley & Sons, New York.
- 2)Machin, D. et al. (1987), Sample size tables for clinical studies, Blackwell Sciences, London.
- 3)Schein-Chung Chow and Jen-pei Liu (1998), Design and analysis of clinical trials: Concepts and methodologies, Wiley
- 4)Schuster, J.J. (1993), Practical handbook of sample size guidelines for clinical trials, CRC Press, Florida.
- 5)김 호 (2002), 적절한 연구대상수의 산출, 대한마취과학회지: 42(1), 1-10.
- 6)김 호 (2008), “튜토리얼 : R과 SAS를 이용한 표본수 및 검정력 계산” 한국보건정보통계학회지; 33(1): 11-20
- 7)박애경, 김 호 (2007), 유전체 연관연구에서의 검정력 및 연구대상수 계산 고찰, 예방의학회지; 40(2):114-121.

감사합니다.

hokim@snu.ac.kr

<http://plaza.snu.ac.kr/~hokim>