7. 가설검정
Hypothesis testing

2017/4/24

7.1 Introduction

- 가설검정(Hypothesis Testing)의 목적

  어느 표본을 분석해 표본의 모집단에 관해 어떤 결론을 내릴 수 있도록 도와준다

We can make a decision on the population (parameters) based on the results of hypothesis testing

- 가설- 하나 또는 그 이상의 모집단에 대한 기술.

  모집단의 모수(parameter)에 관련

Hypothesis-description(s) about the population (parameters)

*Basic logic of statistical hypothesis testing

- A(observed data) -> B (research hypothesis)
- -B -> -A
- Null hypothesis(-B) : No difference (Ho)
- Alternative hypothesis (B): different pop's (Ha)
- Type I error : prob of rejecting right null hypo = Pr (reject Ho | Ho is true) α
- Type II error : prob of accepting wrong null hypo = Pr (Not reject Ho | Ha is true) β
- Power = 1- β (prob of detecting the diff which is real)

- 4 cases of the hypothesis testing

Results of the Hypo. Test	Truth	
	H_0 : True	H_0 : False (Ha:True)
Accept H_0	Right ($1-\alpha$)	Type II error(β)
Reject H_0	Type I error(α)	Right (power=$1-\beta$)

- 가설검정 단계(Procedures of hypothesis testing)

1) 자료(Data) -> to understand the characteristics of the data

2) 가정(Assumptions of the model)

   정규성, 등분산성, 독립성 등 (Normality, homogeneity (same variances), independence, … etc.)

3) 가설(Hypotheses)

   ① 영가설: null hypotheses, H_O– 'No diff' we want to reject to prove the research hypothesis

   ② 대립가설: alternative hypothesis H_A – research hypo to be proven

   ex) $H_o: \mu = 50,\ H_A: \mu \neq 50$

   ❖ 증명하고자 하는 것은 대립가설이다. (We want to have a statistical support to claim the alternative hypothesis.)

   ❖ 영가설을 기각하지 못했다는 것이 영가설을 증명했다는 것은 아니다. ('we cannot reject the null hypo'='no evidence of rejecting the null hypo' ≠ 'null hypo is right')

4) 검정통계량(Test Statistic)

- 자료에서 계산됨(depends on observed data)
- determines rejecting the null hypo
- $$\text{test stat}=\frac{\text{observed stat}-\text{param from the null}}{\text{SE of the stat}} \qquad z=\frac{\bar{x}-\mu_o}{\sigma/\sqrt{n}}$$

  $=$ standardized diff between observed stat and the param from the null

5) 검정통계량의 분포 (dist'n of the test stat under the null dist'n)

6) 결정규칙 (decision rule)

- 기각역(rejection region), 비기각역(nonrejection region)
- 검정통계량이 기각역에 들어오면 영가설을 기각하고 그렇지 않으면 기각하지 않는다. (If the value of the test stat is in the rejection region then we reject the null hypo, otherwise we cannot reject the null hypo.)

❖ 유의수준(lever of significance, α)

- Porb of rejecting right null hypo.-> Type 1 error

		Ho is	
		Correct	Wrong
decision	Fail to reject H_O	Right decision	Type II error
	Reject H_O	Type I error	Right decision

7) 검정통계량의 계산(Calculate Test Statistic)

8) 통계적 결정(Statistical Decision)

9) 결론(Conclusion)

영가설 기각->대립가설이 증명됨. (Reject H_O-> Evidence that H_A is true)

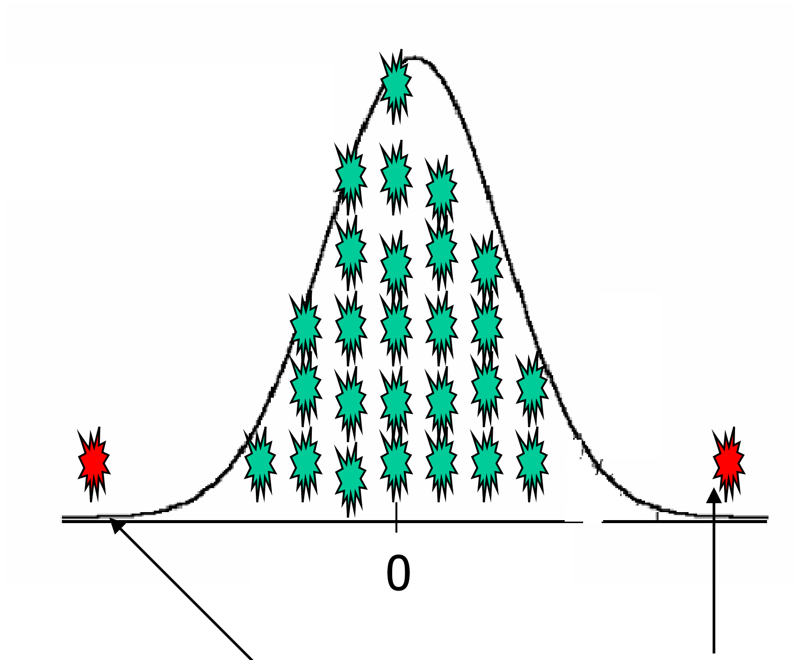
영가설을 기각하지 못함 -> 대립가설이 사실이라는 증거가 없음. 영가설은 사실일 수 있음. (Not reject H_O -> We don't have an evidence that H_A is true, H_O can be right)

P-value (1)

- 연구목적 : 관심변수의 (모)평균이 두 집단에서 다르다. (Research Interest: two means are different)

- $\bar{Y}_1$ 첫 번째 집단에서의 표본 평균 (sample mean from the first pop)
- $\bar{Y}_2$ 두 번째 집단에서의 표본 평균 (sample mean from the second pop)
- 만약 두 집단에서의 모평균이 같다고 하면 (If we assume that the two means are the same)
- 두 표본 평균은 비슷할 것이다. (then two sample means would be similar)
- 표본평균의 차이를 반복적으로 구해보면 (If we calculate the difference of the two means)

P-value (2)



통계적으로 대단히 일어나기 어려운 사건

Events with small proabilities

P-value (3)

- P-value = 두 집단의 평균이 같다고 가정했을 때 우리의 자료, 혹은 더 차이가 나는 자료를 얻을 확률
- Prob(observing our data or larger difference when Ho is true)
- 작은 p-value : 위의 확률이 작다 (small p-value)

➔ 통계적으로 가능하지 않은 일이 일어났다.

➔ 두 집단의 평균이 같다는 가정에 문제가 있다. (We reject Ho)

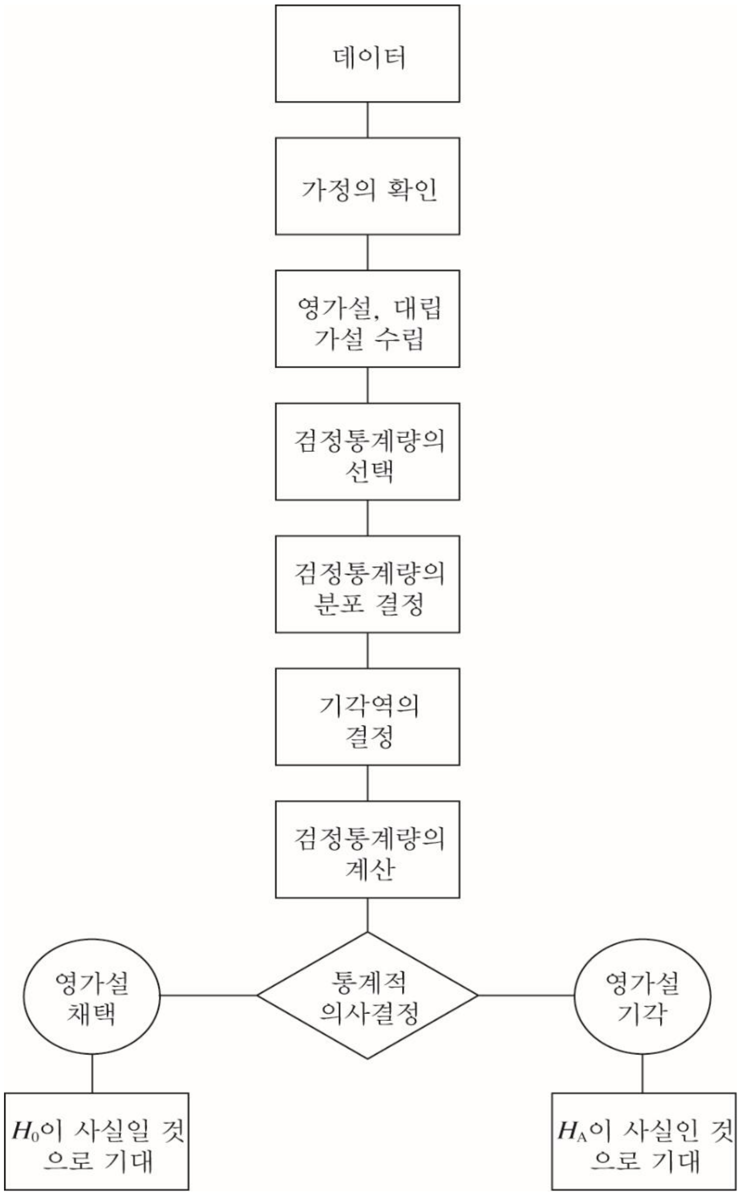
➔ 두 집단의 평균은 같지 않다고 결론 내린다.(We conclude that the means are different)

P-value (3)

- 작지 않은 p-value : 두 집단의 평균이 같다고 가정하면 우리의 자료를 관측할 확률이 작지 않다. (p-value which is not small: observing our data under the null hypo is likely to happen)

➔두 집단의 평균이 같다는 가정에 문제가 없다. (No problem to assume that the means are the same)

양쪽검정, 한쪽검정 (one or two-sided tests)



Data evaluation

Assumptions check

Generate Hypothesis

Choose Test stat

Dist'n of test stat under Ho

Rule of the test (Rejection region)

Calculate test stat

Stat test

Reject Ho -> conclude that Ha is true

Not reject Ho ->conclude that Ho can be true

7.2 모집단 평균에 대한 가설 검정 (Hypo test about the mean)

i. 모집단이 분산이 알려진 정규분포를 따르는 경우 (Sample from normal and known variance)

$$\Rightarrow \quad z = \frac{\bar{x} - \mu_o}{\sigma / \sqrt{n}}$$

<Ex 7.2.1>

모집단의 평균 연령이 30세가 아닌지 확인하고자 한다. We want to know that pop mean is very different from 30

1) data: n=10. $\bar{x} = 27$

2) assumption: normal with $\sigma^2 = 20$

3) hypotheses: $H_o : \mu = 30, \; H_A : \mu \neq 30$

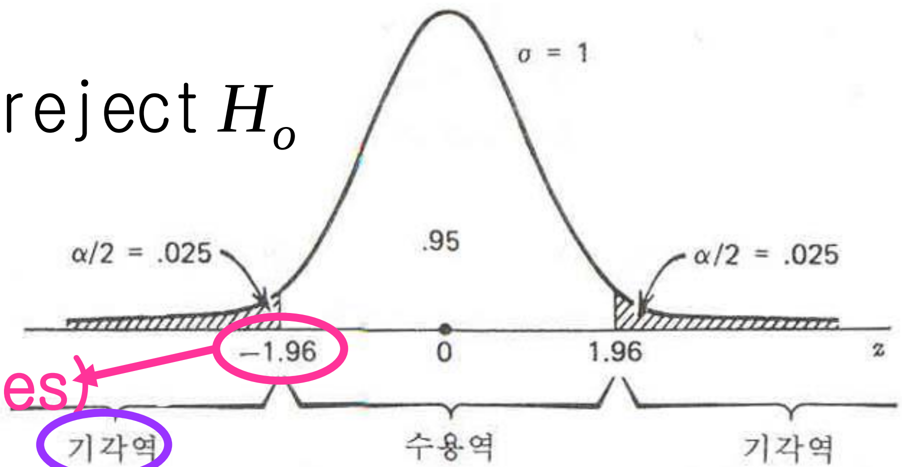
4) Test stat: $$z = \frac{\bar{x} - \mu_o}{\sigma / \sqrt{n}}$$

5) dist'n of the test stat– Under the assumptions and Ho Test Stat ~N(0, 1)

6) Decision rule:

$$\alpha = 0.05$$

If $z \leq -1.96$, or $z \geq 1.96$ then reject H_o



기각값, 임계값(critical values)

critical region

Reject and accept regions

7) Calculate test stat:

$$z = \frac{\bar{x} - \mu_o}{\sigma / \sqrt{n}} = \frac{27 - 30}{\sqrt{20/10}} = \frac{-3}{1.4142} = -2.12$$

8) stat'l decision:

-2.12 is in the rejection region-> we reject the null hypo -> TS is significantly different from 0 by $\alpha = 0.05$

9) conclusion: We conclude that the mean age is different from 30.

10) p-value: Prob(observing our data or larger difference when Ho is true)

one-sided p-value=pnorm(-2.12)=1-pnorm(2.12) = 0.017, two-sided p-value=.0340

- Testing H_O using Confidence Interval

$$\text{CI of } \mu = 27 \pm 1.96\sqrt{20/10} = 24.23,\ \ 29.77$$

-> CI does not include 30 => we reject H_O

If two-sided $100(1-\alpha)\%$ CI does not include the parameter value => we reject (Ho:parameter = the value) by the α level.

If .. include … => we do not reject

- 단측가설 검정(One-sided test)

<Ex 7.2.2> 만약 '$\mu \neq 30$'이 아닌 '$\mu < 30$'이라고 결론을 내릴 수 있는지 관심이 있다고 가정하자. 가설 검정의 10단계를 설명하라.

mean of age. Can we claim that the pop mean is less than 30?

1) data

2) assumptions

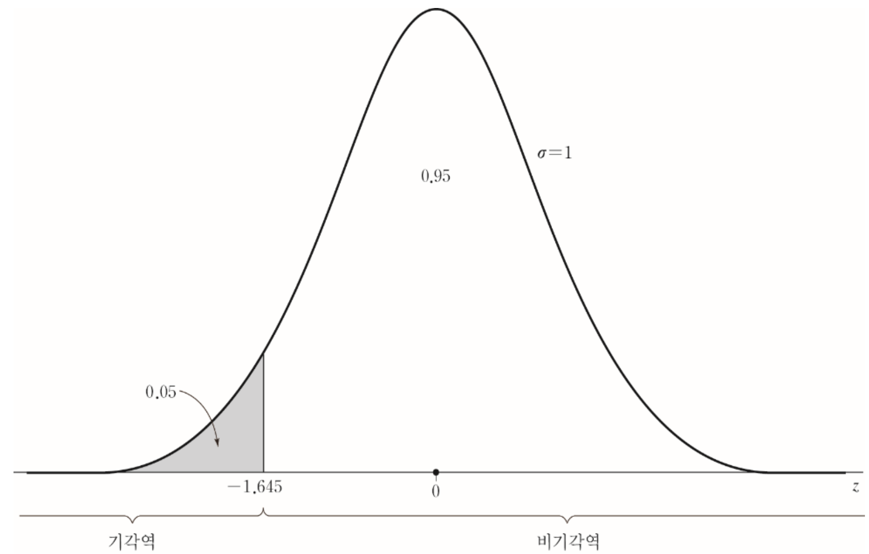
3) Hypothesis : $H_o : \mu \geq 30, \;\; H_A : \mu < 30$

4) Test stat: $z = \dfrac{\bar{x} - \mu_o}{\sigma / \sqrt{n}}$

5) Distribution of the test stat

6) Decision rule: $\alpha = 0.05$

If $z \le -1.645$, then reject H_0



7) Calculating test stat:

$$z = \frac{27-30}{\sqrt{20/10}} = -2.12$$

8) stat'l decision:

−2.12<−1.645 => reject Ho

9) Conclusion: We conclude that the mean is less than 30

10)p-value= 0.017

ii. 모집단이 정규분포이나 분산을 모르는 경우 (Sampling from normal pop : unknown variance)

$$t = \frac{\bar{x} - \mu_o}{s/\sqrt{n}}$$

<Ex 7.2.3>

환자들이 첫번째 MRI 촬영까지 걸린 시간을 조사 (measured # of days to MRI scanning for 20 patients)

1) data:

환자	시간(일)	환자	시간(일)	환자	시간(일)	환자	시간(일)
1	14	6	0	11	28	16	14
2	10	7	10	12	24	17	9
3	18	8	4	13	24	18	20
4	26	9	8	14	2	19	10
5	12	10	21	15	3	20	15

2) assumptions: random sample from normal pop. Variance is not known.

3) hypo: $H_o: \mu = 15,\ H_A: \mu \neq 15$

4) Test stat:

$$t = \frac{\bar{x} - \mu_o}{s/\sqrt{n}}$$

5) dist'n of the test stat: t-dist with df=n-1 under Ho

6) Decision rule: $\alpha = 0.05$

2-sided test

qt(0.025,df=19)=-2.093

Reject Ho if TS < -2,093 or > 2.093

7) Cal test stat:

$$t = \frac{13.6-15}{8.312/\sqrt{20}} = \frac{-1.4}{1.85862} = -0.7532$$

8) stat'l decision:

$-0.7532 > -2.093$ => cannot reject H_O

9) conclusion: We cannot claim that the mean is different from 15.

10) p-value

```
> pt(-0.7532,df=19)*2
[1] 0.4605616
```

iii. 모집단이 정규분포를 따르지 않는 경우(Random sample from non-normal pop): CLT

$$z = \frac{\bar{x} - \mu_o}{s/\sqrt{n}}$$

<Ex 7.2.4>

157명 환자들의 혈압평균은 146mmHg, 표본편차는 27이라고 한다. 이때 모집단의 평균 혈압이 140mmHg보다 크다고 할 수 있는지에 관심이 있다고 한다. 가설검정을 수행하라.

SBP for 157 patients. Sample mean=146 mmHg, sample sd=27. Is pop mean > 140?

1) data: $\bar{x} = 146, \; s = 27$

2) assumption: sample from non-normal pop

3) hypo: $H_o : \mu \le 140, \; H_A : \mu > 140$

4) Test stat:

$$z = \frac{\bar{x} - \mu_o}{s/\sqrt{n}}$$

5) dist'n of test stat: By CLT, approximately normal dist under Ho

6) Decision rule: $Z > 1.645. \quad \alpha = 0.05$

7) Cal test stat:

$$z = \frac{146 - 140}{27/\sqrt{157}} = \frac{6}{2.1548} = 2.78$$

8) stat'l decision:

2.78>1.645 => Reject H_O

9) conclusion: Mean SBP is greater than 140. (p=0.0027)

10) p-value

1-pnorm(2.78)=1-0.9973=0.027

* 7.3 두 정규분포 모집단의 평균비교 (hypo test: diff of two means)

- hypotheses

$$(1)\ \ H_o: \mu_1 - \mu_2 = 0,\ \ H_A: \mu_1 - \mu_2 \neq 0$$

$$(2)\ \ H_o: \mu_1 - \mu_2 \geq 0,\ \ H_A: \mu_1 - \mu_2 < 0$$

$$(3)\ \ H_o: \mu_1 - \mu_2 \leq 0,\ \ H_A: \mu_1 - \mu_2 > 0$$

i. 모집단이 분산이 알려져있는 정규분포를 따르는 경우 (From normal pop with known variances)

$$\Rightarrow \quad z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}}}$$

<Ex 7.3.1>

12명의 환자군과 15명의 대조군의 BMI를 비교. 모분산은 1, 1.5라고 알려져 있으며, 표본평균: 환자군 4.5mg, 대조군 3.4mg. 두 그룹의 모평균이 같은지 가설검정을 수행하라.

mean diff of BMI between pts and normal controls

1) data: 12 pts, 15 normal controls

$$\bar{x}_1 = 4.5,\ \ \bar{x}_2 = 3.4$$

2) assumptions: two independent samples from two independent normal pop with variance=1 and 1.5 for pts and controls, respectively.

3) hypo: $$H_o: \mu_1 - \mu_2 = 0, \rightarrow H_o: \mu_1 = \mu_2$$
$$H_A: \mu_1 - \mu_2 \neq 0, \quad H_A: \mu_1 \neq \mu_2$$

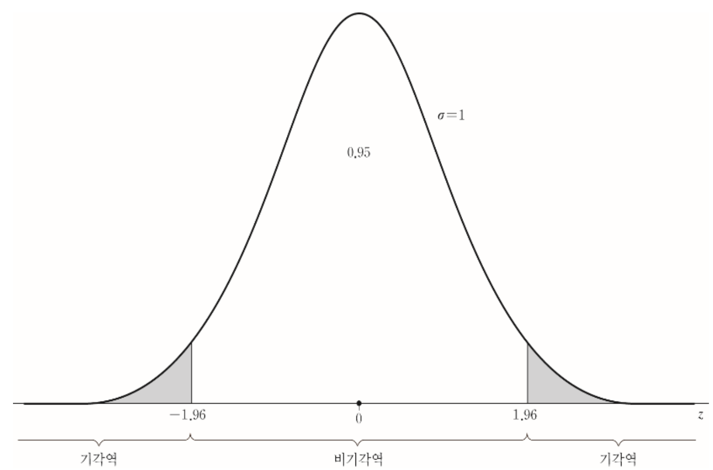
4) Test stat: $$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$$

5) dist'n of test stat: N(0,1) under H_O

6) Decision rule: $\alpha = 0.05$

7) Cal test stat:

$$z = \frac{(4.5 - 3.4) - 0}{\sqrt{\frac{1}{12} + \frac{1.5}{15}}} = 2.57$$



8) stat'l decision:

2.57>1.96 => Reject H_O

9) conclusion: means are diff

10) p-value

> pnorm(-2.57)*2 [1] 0.01016985

95% CI of : (0.26, 1.94) 0 is not included, significantly different ii.모집단이 정규분포를 따르나 모분산을 모르는 경우 (From normal pop with unknown variances)

(1) 모분산이 같은 경우 (same variances)

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

<Ex 7.3.2>

환자 10명과 정상인 10명. 환자군이 대조군에 비하여 혈압평균이 낮다고 할 수 있는지 유의수준 0.05에서 검정하라.

SBP for 10 pts and 10 controls. Can we say that mean SBP of pt is lower than that of controls?

1) data:

pts	120	115	130	131	111	117	90	114	150	170
control	60	150	130	180	153	135	121	119	130	130

2) assumptions: two indep random samples from two normal pop with unknown variances

3) hypo: $H_0 : \mu_D \geq \mu_C$, vs $H_A : \mu_D < \mu_C$

4) Test stat:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p{}^2}{n_1} + \frac{s_p{}^2}{n_2}}} \quad \left(s_p{}^2 = \frac{(n_1 - 1)s_1{}^2 + (n_2 - 1)s_2{}^2}{n_1 + n_2 - 2} \right)$$

5) dist'n of test stat: t dist'n with df= $n_1 + n_2 - 2$ =18 under Ho

6) Decision rule: qt(0.05,18) = −1.734064

7) Cal test stat: $\bar{x}_D = 124.8$, $s_D = 22.215$, $\bar{x}_C = 130.8$,

$s_C = 30.82$, $s_P^2 = \frac{(9)(22.215)^2 + (9)(30.82)^2}{9+9} = 721.7333$,

$$t = \frac{(124.8 - 130.8) - 0}{\sqrt{\frac{721.733}{10} + \frac{721.733}{10}}} = -0.4994.$$

8) stat'l decision: $-0.4994 > -1.7341$ => cannot reject H_O

9) conclusion: we cannot reject Ho .

We cannot say that mean SBP of pt is lower than that of controls

10) p-value

pt(−0.4994, df=18)=0.31177

(2) 모분산이 같지 않은 경우 (diff variances)

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

<Ex 7.3.3>

두 모집단으로부터 뽑은 독립 확률표본의 표본평균 및 표본표준편차는 다음과 같다고 한다.

$$n_1 = 15, \quad \bar{x}_1 = 19.16, \quad s_1 = 5.29$$
$$n_2 = 30, \quad \bar{x}_2 = 9.53, \quad s_2 = 2.69$$

두 모평균이 다른지 관심이 있다고 할 때, 가설검정을 수행하라. (Interested in equality of the two means)

1) data:

2) assumptions: two indep random samples from normal pop with unknown variances (unequal)

3) hypo: $H_0: \mu_1 = \mu_2, \quad \text{vs} \quad H_A: \mu_1 \neq \mu_2$

4) Test stat: $t' = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

5) dist'n of TS: $t'_{1-\alpha/2} = \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$

6) Dec rule: $w_1 = 1.8656, w_2 = 0.2412,$ qt(0.975, df=14), qt(0.975, df=29) $t_1 = 2.1448, t_2 = 2.0452$

$$t'_{1-\alpha/2} = \frac{1.8656(2.1448) + 0.2412(2.0452)}{1.8656 + 0.2412} = 2.133$$

7) Cal TS: $t' = \frac{(19.16 - 9.53) - 0}{\sqrt{\frac{(5.29)^2}{15} + \frac{(2.69)^2}{30}}} = \frac{9.63}{1.4515} = 6.63$

8) stat'l dec: 6.63>2.133 -> reject Ho

9) conclusion: Two means are different

iii. 모집단이 정규분포를 따르지 않는 경우 (Sample from non-normal pop)

Apply CLT for large sample data

<Ex 7.3.4> 두 그룹 A, B에서 IgG의 농도를 측정하였다. 그룹 A의 IgG 농도가 그룹 B에 비하여 더 높은지 관심이 있다고 한다. (Interested in IgG of group A > IgG of group B), $\alpha = 0.01$

1) data:

Group	Mean IgG Level (ml/unit)	Sample Size	Standard Deviation
A	59	53	45
B	46	54	34

2) assumptions: two indep random samples from normal pop, sample sizes are big enough.

3) hypo: $H_0 : \mu_A - \mu_B \leq 0,$ vs $H_A : \mu_A - \mu_B > 0$

4) Test stat: $t' = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

5) dist'n of TS: Z by CLT

6) Dec rule: qnorm(0.99)=2.33

7) Cal TS $t = \frac{59-46}{\sqrt{\frac{(45)^2}{53} + \frac{(34)^2}{54}}} = 1.6837$

8) stat'l dec: 1.68< 2.33 cannot reject H_O

9) conclusion: We cannot claim "IgG of group A > IgG of group B"

10) p-value: pnorm(1.6837,lower=F)=0.04611

Homework

- 1-11

7.4 짝비교(Paired Comparisons)

- Samples from non-indep observations

  ex) observation pre/post treatment

- Test stat

$$\Rightarrow \quad t = \frac{\bar{d} - \mu_o}{s_{\bar{d}}}$$

- 이론적 배경

$\mathrm{E}(X - Y) = E(X) - E(Y)$

$Var(X - Y) = Var(X) + Var(Y) - 2Cov(X, Y)$

$< Var(X) + Var(Y)$ if X and Y are positively correlated

<Ex 7.4.1>

새롭게 개발한 치료방법이 역류성식도염 환자에게 효과가 있는지 확인을 하고자 한다. 짝비교 검정을 이용하여 검정하라.

Comparisons between before and after treatments (After−Before)

1) data:

치료 전 (%)	22	60	96	10	3	50	33	69	64	18	0	34
치료 후 (%)	63	91	59	37	10	19	41	87	86	55	88	40

$$d_i \quad 41\ \ 31\ -37\ \ 27\ \ 7\ -31\ \ 8\ \ 18\ \ 22\ \ 37\ \ 88\ \ 6$$

2) $d_i \sim$ iid $N(\cdot,\cdot)$

3) hypo: $H_0: \mu_d \le 0 \quad \text{vs} \quad H_A: \mu_d > 0$

4) TS: $t = \dfrac{\bar{d} - \mu_{d0}}{s_{\bar{d}}}$

5) dist'n TS: t dist'n with df=n−1 under H_O

6) Dec rule: qt(0.95, df=11)=1.7959

7) Cal TS

$$\bar{d} = \frac{\sum d_i}{n} = \frac{(41) + (31) + (-37) + \cdots + (6)}{12} = \frac{217}{12} = 18.0833$$

$$s^2{}_d = \frac{\sum(d_i - \bar{d})^2}{n-1} = \frac{n\sum d_i^2 - \left(\sum d_i\right)^2}{n(n-1)} = \frac{12(15771) - (217)^2}{(12)(11)}$$

$$= 1076.992$$

$$t = \frac{18.0833 - 0}{\sqrt{1076.992/12}} = \frac{18.0833}{9.473614} = 1.908807$$

8) stat'l dec: 1.9088>1.7959=> reject H_O

9) conclusion: The new treatment is effective to treat the symptom.

10) p-value

pt(1.908807,df=11,lower=F)=0.041

95% CI of :

$$\bar{d} \pm t_{1-\frac{\alpha}{2}} \cdot s_{\bar{d}} = 18.0833 \pm 2.2010\sqrt{\frac{1076.992}{12}}$$

$$= 18.075 \pm 20.851 \;\Rightarrow\; (-2.776, 38.926)$$

7.5 모집단 비율의 가설검정
Hypo test: proportion from one pop

→ $$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

<Ex 7.5.1> 300명 검사 25명 공복혈당장애. 공복혈당장애를 앓고 있는 사람의 모비율은 6.3% 이상인지 관심이 있다고 할 때, 가설검정을 수행하라.

Impaired fasting glucose (IFG) : 24 out of 300. Interested in p > 6.3%

1) Data: $\hat{p} = \frac{25}{301} = \mathbf{0.083}$

2) assu. $n\hat{p} > 5$ and $n(1-\hat{p}) > 5$ -> CLT

3) Hypo. $H_0 : p \leq 0.063$ vs $H_A : p > 0.063$

$$p_0 = 0.063, \sigma_p = \sqrt{(0.063)(0.937)/300}$$

4) TS: $z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$

5) dist'n of TS: asymptotically standard normal under H_O

6) Dec rule: reject Ho if Z > 1.645

7) Cal TS: $z = \frac{0.083-0.063}{\sqrt{\frac{(0.063)(0.937)}{300}}} = 1.4257$

8) sta'l dec: 1.4257<-1.645 => cannot reject H_O

9) conclusion: We don't have enough evidence of "p>6.3%"

10) p-value: pnorm(1.4257, lower=F)= 0.07697

7.6 두 모비율 차이에 대한 검정 (hypo test: diff of two proportions)

<Ex 7.6.1>

30명의 남성, 50명의 여성 조사, 12명 남성과 25명 여성이 누난증후군. 여성이 남성보다 누난증후군을 앓고 있는 사람의 비율이 높은지 관심.

compare the proportion of patients between male and female. Research hypothesis is that female's prevalence is higher.

1) data:12/30 for male, 25/50 for female

2) ass: Independently sampled

3) hypothesis: $H_0: \ p_f \le p_m$ vs $H_A: \ p_f > p_m$

4) Test statistics $z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)_0}{\hat{\sigma}_{\widehat{p_1} - \widehat{p_2}}}$

5) dist'n TS: $n_f p_f > 5$, $n_f(1 - p_f) > 5$, $n_m p_m > 5$, $n_m(1 - p_m) > 5$

-> asymptotically standard normal under H_O

6) Dec rule: Reject Ho if TS > 1.645 $\alpha = 0.05$,

7) Cal TS: $\hat{p}_f = \frac{25}{50} = 0.5$, $\hat{p}_m = \frac{12}{30} = 0.4$, $\bar{p} = \frac{12+25}{50+30} = 0.4625$ $z = \frac{(0.5-0.4)}{\sqrt{\frac{(04625)(0.5375)}{50} + \frac{(0.4625)(0.5375)}{30}}} = 0.8684$

8) Stat dec: 0.8684<1.645=> cannot reject H_O

9) conclusion: We failed to prove that female's prevalence is higher.

10) p-value: pnorm(0.8684, lower=F)=0.1926

7.7 모집단 분산의 가설검정 (hypo test: variance of one pop)

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

<Ex 7.7.1>

정규분포 가정. 20명 실험자의 표본분산 670.81. 모분산이 600? 유의수준 0.05

Normality is assumed. n=20. Sample variance =670.81 pop variance=600? 95% confidence level

1) data: $s^2 = 670.81$

2) ass: random sample from normal pop

3) hypo: $H_o : \sigma^2 = 600, \; H_A : \sigma^2 \neq 600$

4) Test stat: $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

5) dist'n of TS: χ^2 with df=n−1 under H_O

6) Dec rule: qchisq(.975, df=19), qchisq(.025, df=19) 8.9065와 32.8523

7) Cal TS: $\chi^2 = \frac{19(670.81)}{600} = 21.2423$

8) stat'l dec: 8.9065 <21.2423<32.8523 => cannot reject H_O

9) conclusion: No evidence that the variance is not 600(p>0.05) "sample variance is not significantly diff from 600"

10) 카이제곱 분포가 대칭이 아니므로 양쪽 검정의 정확한 p-값을 정의하기가 복잡함. 한쪽 검정의 경우는 훨씬 간담함.

7.8 두 모분산 비에 대한 가설검정 (hypo test: ratio of two variances)

- 분산비검정(Variance Ratio Test):
  Homogeneity of two pop variances
  Variance ratio = 1 or not

$$\Rightarrow \quad V.R = \frac{{s_1}^2}{{s_2}^2}$$

<Ex 7.8.1>

정규분포 가정, 표본표준편차가 각각 30, 12, 표본의 크기는 각각 6이다. 모집단 A의 모분산이 모집단 B의 모분산보다 크다고 가정할 수 있는지

Pop A, B : Normally distributed. n_A=n_B=6, Sample sd=30, 12. We are interested in “Pop var of A > Pop Var of B”

1) data: 두 모집단의 표본표준편차는 각각 30, 12이다

2) ass: random samples from normal pop

3) hypo: $H_o : \sigma_A{}^2 \leq \sigma_B{}^2,\ H_A : \sigma_A{}^2 > \sigma_B{}^2$

4) TS: $V.R = \dfrac{s_A{}^2}{s_B{}^2}$

5) dist'n of TS: F-dist with df= $n_1 - 1, n_2 - 1$ under H_O

6) Dec rule: α는 0.05 qf(0.95,df_1=5,df_2=5) 5.05

7) Cal TS: $\dfrac{s_1^2}{s_2^2} = \dfrac{(30)^2}{(12)^2} = 6.25$

8) Stat dec: 6.25<5.05=> reject H_O

9) con: We conclude that the variance of pop A is larger than that of pop B

10) p-value : pf(6.25,df_1=5, df_2=5,lower=F) =0.0328

7.9 2종의 오류와 검정력 (Type II error and power)

- 2종의 오류=틀린 영가설을 받아들일 확률
  Type II error : prob of accepting wrong null hypo = Pr (Not reject Ho | Ha is true)
- 검정력=2종의 오류가 일어나지 않을 확률 (실제로 효과가 있능 때 효과를 찾아 낼 확률)
  Power: prob of detecting the diff which is real
- 연구목적을 달성하기 위해서는 검정력이 큰 방법을 사용하여야 한다.
- 검정력은 표본수와 관련이 있다.

<Ex. 7.9.1> $\alpha = 0.05$ H_0 : $\mu = 17.5$, vs H_A : $\mu \neq 17.5$
σ^2 =3.6, n=100. 대립가설이 참, $\mu = 16.5$. 검정력?
(Derive the power when $\mu = 16.5$)

Sol) $\bar{x}_U = \mu_0 + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$, $\bar{x}_L = \mu_0 - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$

$\bar{x}_U = 17.50 + 1.96\frac{(3.6)}{(10)} = 17.50 + 1.96(0.36) = 17.50 + 0.7056 = 18.21$, $\bar{x}_L = 17.50 - 0.7056 = 16.79$

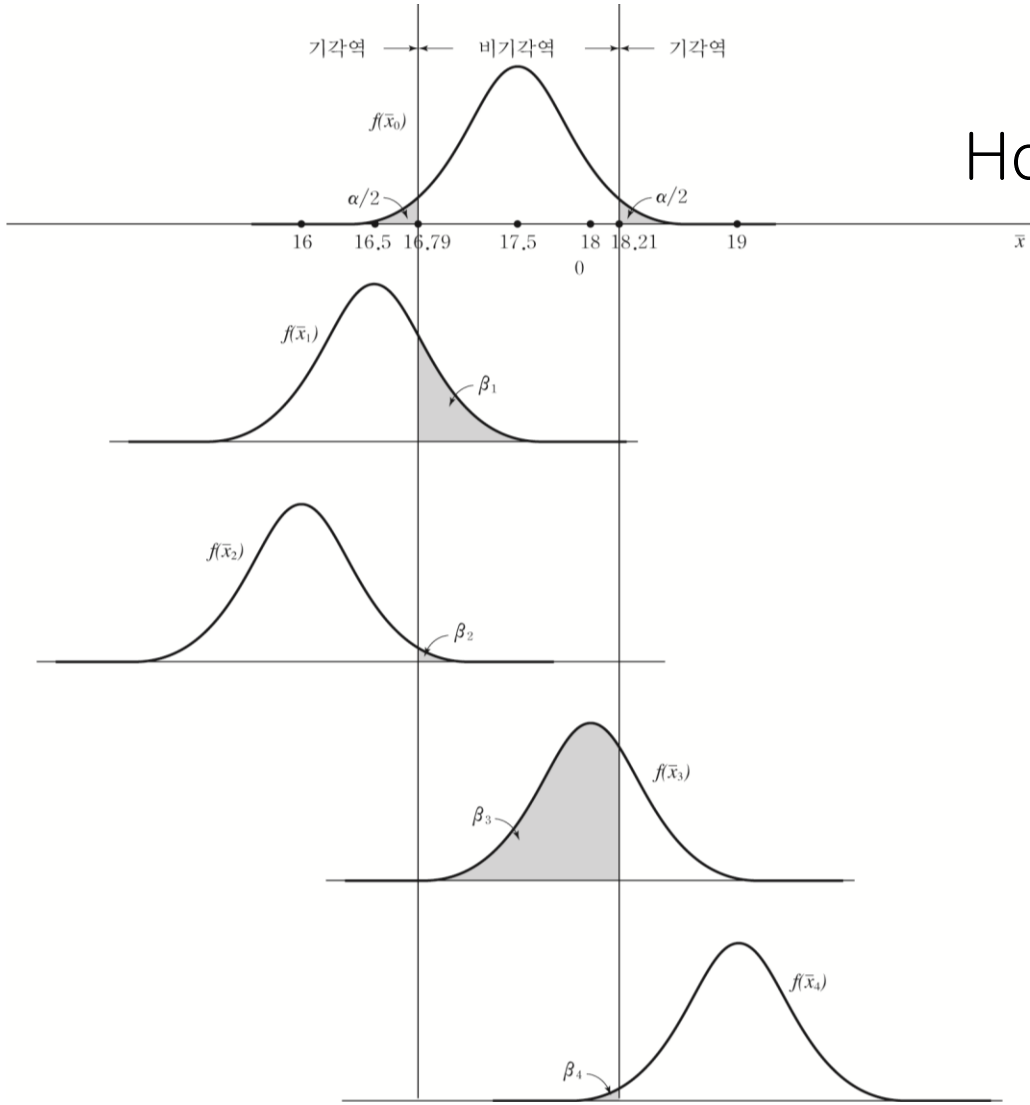
Rule: Do not reject H_0 if $\bar{x}_L < \bar{x} < \bar{x}_U$,

Type II error =P(Not reject $H_0 | H_A$ is true)

=P($\bar{x}_L < \bar{x} < \bar{x}_U | H_A$ is true)

$= \mathrm{P}\left(\frac{16.79-16.5}{0.36} \leq \frac{\bar{x}-\mu_a}{\sigma/\sqrt{n}} \leq \frac{18.21-16.5}{0.36}\right) = P\left(\frac{0.29}{0.36} \leq z \leq\right.$

Stat'l dec. Ha | Ho | Ha

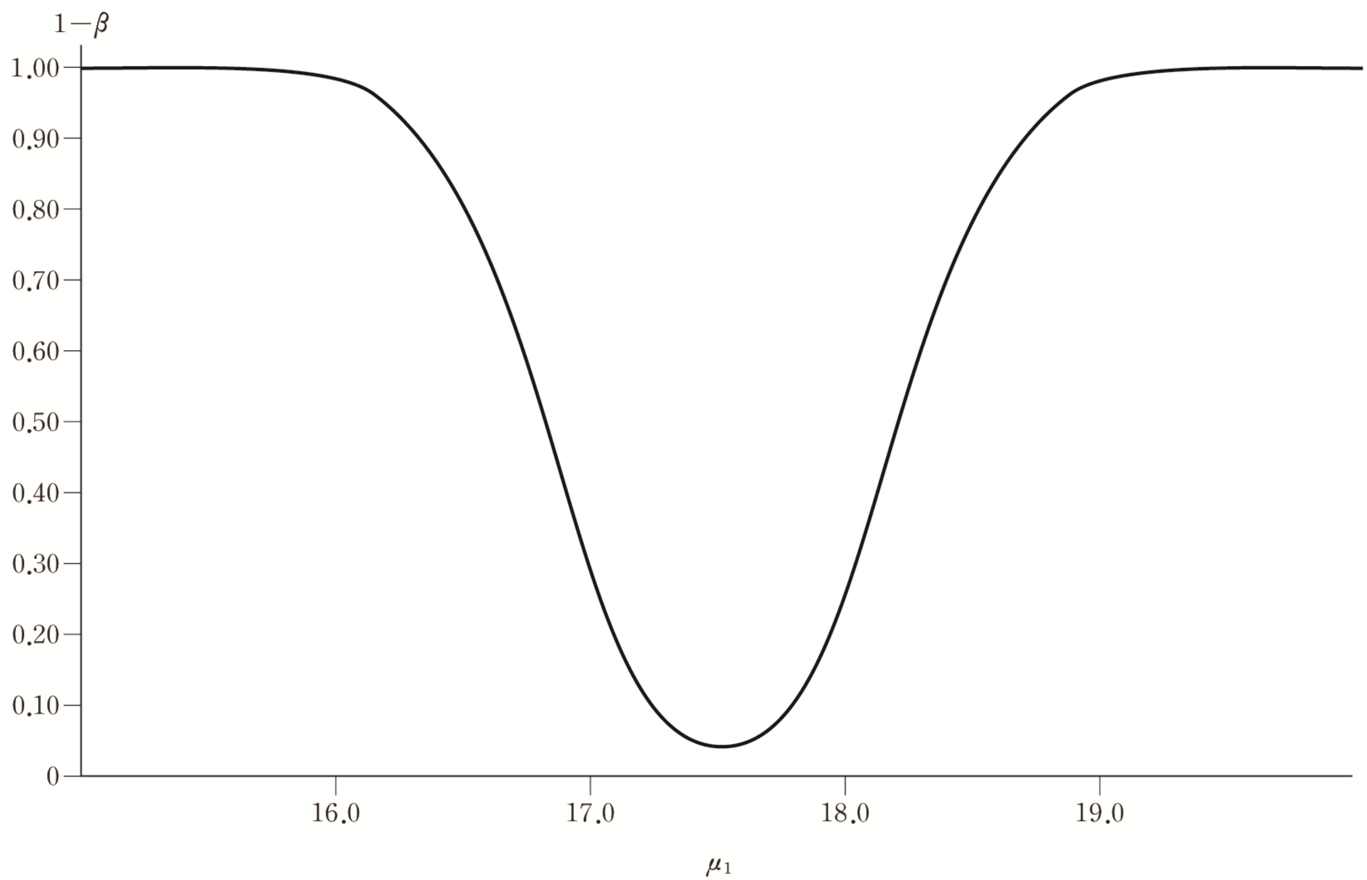


Ho 하에서 TS의 분포

μ_1	β	$1-\beta$
16.0	0.0141	0.9859
16.5	0.2102	0.7898
17.0	0.7198	0.2802
18.0	0.7198	0.2802
18.5	0.2102	0.7898
19.0	0.0141	0.9859

```
> mu1=seq(16,19,0.5)
> mu1
[1] 16.0 16.5 17.0 17.5 18.0 18.5 19.0
>
> power=1-(pnorm((18.21-mu1)/0.36)-
pnorm((16.79-mu1)/0.36))
> power
[1] 0.98589826 0.78975144 0.28022261 0.04858424
0.28022261 0.78975144   0.98589826
```

Ex. 7.9.1의 검정력 곡선



Challenge: draw this plot using only R

단측검정의 검정력 곡선

<Ex. 7.9.2> $\alpha = 0.01, n = 20, \ \sigma = 15$

$$H_0 : \mu \geq 65, \qquad \text{vs} \ \ H_A : \mu < 65$$

대립가설이 참, $\mu = 55$. 검정력? (Derive the power when $\mu = 55$)

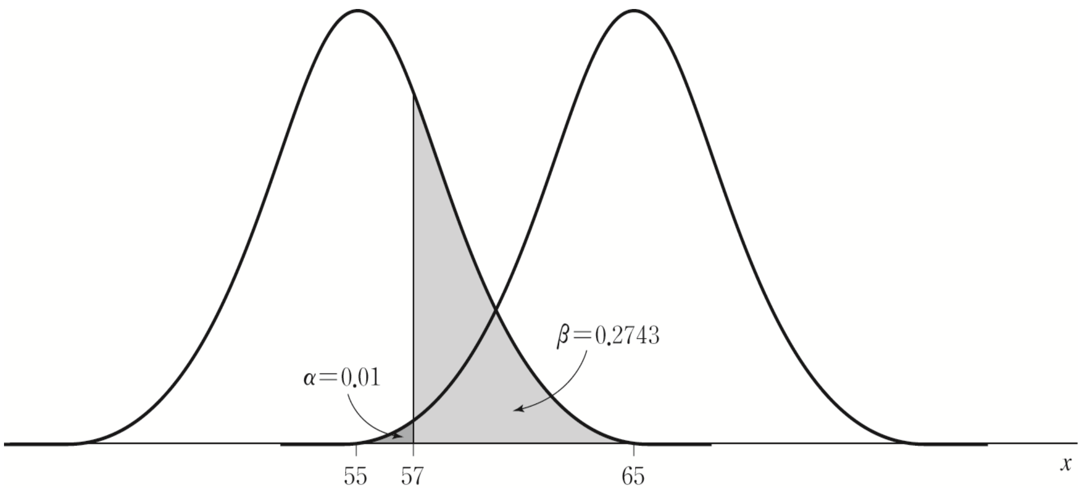
Sol) 임계치 (critical value) $65 - 2.33\left(\frac{15}{\sqrt{20}}\right) = 57$

Type II error =P(Not reject $H_0 | H_\mathrm{A}$ is true)

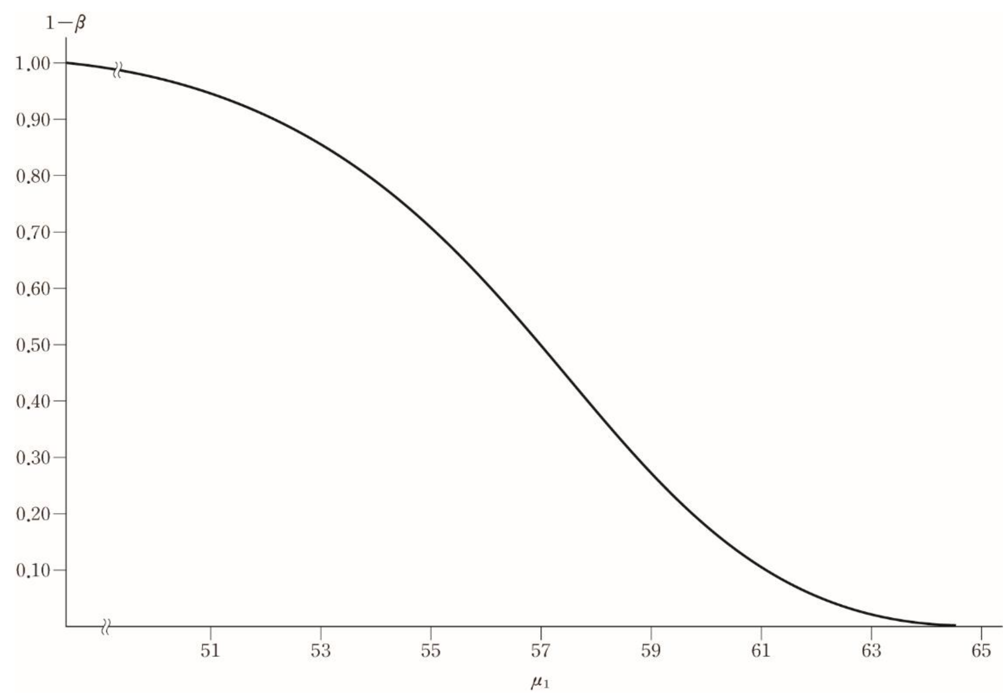
$$= P(\bar{x} > 57 | \mu = 55) = P\left(\frac{\bar{x} - 55}{\frac{15}{\sqrt{20}}} > \frac{57 - 55}{\frac{15}{\sqrt{20}}}\right)$$

$$= P(z > 0.60) = 1 - 0.7257 = 0.2743$$

$\mu = 55$일 때 β



[예제 7.9.2]의 검정력곡선



7.10 2종의 오류를 고려한 표본크기 결정 (Type II error and sample size)

<Ex 7.10.1> [예제 7.9.2]

$$H_0 : \mu \geq 65, \qquad \text{vs } H_A : \mu < 65$$

σ=5, α =0.01, 만약 2종 오류(β)를 0.05라 한다면 1종 오류와 2종 오류를 모두 만족할 수 있는 표본의 숫자는 얼마나 커야 하는가?

What is n to guarantee α and β?

Sol) let C be the critical Value, then

$$P\left(\mathrm{Z} \leq \frac{C-\mu_0}{\sigma/\sqrt{n}}\right) = \alpha,\ \ P\left(\mathrm{Z} \geq \frac{C-\mu_1}{\sigma/\sqrt{n}}\right) = \beta,\ \mathrm{Z} \sim N(0,1)$$

$$\Longrightarrow \frac{C-\mu_0}{\frac{\sigma}{\sqrt{n}}} = z_\alpha = -z_{1-\alpha},\ \ \frac{C-\mu_1}{\sigma/\sqrt{n}} = z_{1-\beta}$$

$$\Longrightarrow \mathrm{C} = \mu_0 - z_{1-\alpha}\frac{\sigma}{\sqrt{n}} = \mu_1 + z_{1-\beta}\frac{\sigma}{\sqrt{n}}$$

$$n = \left\lfloor \frac{\left(z_{1-\alpha} + z_{1-\beta}\right)\sigma}{(\mu_0 - \mu_1)} \right\rfloor^2$$

$\mu_0 = 65$, $\mu_1 = 55$, $\sigma = 15$. $\alpha = 0.01, z_{1-\alpha} =$ 2.33, $\beta = 0.05, z_{1-\beta}$=1.645

$$n = \left\lfloor \frac{(2.33+1.645)(15)}{(65-55)} \right\rfloor^2 = 35.55 \ \Longrightarrow 36$$

$$C = 65 - 2.33\left(\frac{15}{\sqrt{36}}\right) = 59.175$$