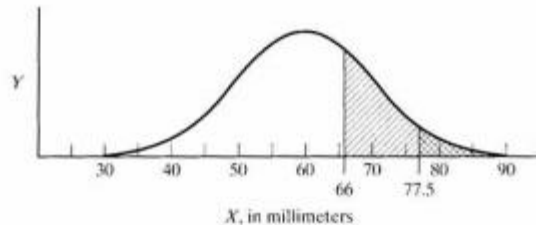


Example.6.1a

EXAMPLE 6.1a Calculating Proportions of a Normal Distribution of Bone Lengths, Where $\mu = 60$ mm and $\sigma = 10$ mm



1. What proportion of the population of bone lengths is larger than 66 mm?

$$Z = \frac{X_i - \mu}{\sigma} = \frac{66 \text{ mm} - 60 \text{ mm}}{10 \text{ mm}} = 0.60$$

$$P(X_i > 66 \text{ mm}) = P(Z > 0.60) = 0.2743 \text{ or } 27.43\%$$

2. What is the probability of picking, at random from this population, a bone larger than 66 mm? This is simply another way of stating the quantity calculated in part (1). The answer is 0.2743.
3. If there are 2000 bone lengths in this population, how many of them are greater than 66 mm?

$$(0.2743)(2000) = 549$$

4. What proportion of the population is smaller than 66 mm?

$$P(X_i < 66 \text{ mm}) = 1.0000 - P(X_i > 66 \text{ mm}) = 1.0000 - 0.2743 = 0.7257$$

5. What proportion of this population lies between 60 and 66 mm? Of the total population, 0.5000 is larger than 60 mm and 0.2743 is larger than 66 mm. Therefore, $0.5000 - 0.2743 = 0.2257$ of the population lies between 60 and 66 mm. That is, $P(60 \text{ mm} < X_i < 66 \text{ mm}) = 0.5000 - 0.2743 = 0.2257$.

6. What portion of the area under the normal curve lies to the right of 77.5 mm?

$$Z = \frac{77.5 \text{ mm} - 60 \text{ mm}}{10 \text{ mm}} = 1.75$$

$$P(X_i > 77.5 \text{ mm}) = P(Z > 1.75) = 0.0401 \text{ or } 4.01\%$$

7. If there are 2000 bone lengths in the population, how many of them are larger than 77.5 mm?

$$(0.0401)(2000) = 80$$

8. What is the probability of selecting at random from this population a bone measuring between 66 and 77.5 mm in length?

$$P(66 \text{ mm} < X_i < 77.5 \text{ mm}) = P(0.60 < Z < 1.75) = 0.2743 - 0.0401 = 0.2342$$

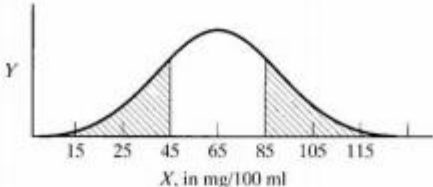
R

CODE

```
#ex6.1a
no1=round(pnorm(66, 60, 10, lower.tail = F), 4)
no1
no3=round(no1*2000)
no3
1-no1 == round(pnorm(66,60,10),4)
no4=round(pnorm(66,60,10),4)
```

	<pre> no4 no5=round(pnorm(66,60,10)-pnorm(60,60,10),4) no5 no6=round(pnorm(77.5,60,10, lower.tail = F),4) no6 stand=function(x) (x-60)/10 round(pnorm(stand(77.5),0,1, lower.tail = F),4) no7=round(no6*2000) no7 no8=round(pnorm(77.5,60,10)-pnorm(66,60,10),4) no8 round(pnorm(stand(77.5),lower.tail = T)-pnorm(stand(66), lower.tail = T),4) </pre>
OUTPUT	<pre> > #ex6.1a > no1=round(pnorm(66, 60, 10, lower.tail = F), 4) > no1 [1] 0.2743 > no3=round(no1*2000) > no3 [1] 549 > 1-no1 == round(pnorm(66,60,10),4) [1] TRUE > no4=round(pnorm(66,60,10),4) > no4 [1] 0.7257 > no5=round(pnorm(66,60,10)-pnorm(60,60,10),4) > no5 [1] 0.2257 > no6=round(pnorm(77.5,60,10, lower.tail = F),4) > no6 [1] 0.0401 > stand=function(x) (x-60)/10 > round(pnorm(stand(77.5),0,1, lower.tail = F),4) [1] 0.0401 > no7=round(no6*2000) > no7 [1] 80 > no8=round(pnorm(77.5,60,10)-pnorm(66,60,10),4) > no8 [1] 0.2342 > round(pnorm(stand(77.5),lower.tail = T)-pnorm(stand(66), lower.tail = T),4) [1] 0.2342 </pre>
SAS	
CODE	<pre> data ex6_1a; a1= round(1-cdf('normal',66, 60,10), .0001); a3= round(a1 * 2000); a4= round(1-a1, .0001); a5= round(cdf('normal', 66,60,10) - cdf('normal', 60,60,10), .0001); a6= round(1-cdf('normal', 77.5, 60,10), .0001); a7= round(a6 * 2000); a8= round(cdf('normal', 77.5,60,10) - cdf('normal', 66,60,10), .0001); run; proc print data=ex6_1a; </pre>

	run;							
OUTPUT	OBS	a1	a3	a4	a5	a6	a7	a8
	1	0.2743	549	0.7257	0.2257	0.0401	80	0.2342
결과해석	위 자료는 뼈의 길이에 대한 분포를 나타낸 것으로 평균이 60(mm)이고 표준편차가 10(mm)인 정규분포를 따르고 있다. 뼈의 길이가 특정 길이의 뼈가 해당 모집단에서 얼마나 차지하는지 비율을 확인하기 위해서 위와 같이 구할 수 있다.							

Example.6.1b	
EXAMPLE 6.1b Calculating Proportions of a Normal Distribution of Sucrose Concentrations, Where $\mu = 65 \text{ mg/100 ml}$ and $\sigma = 25 \text{ mg/100 ml}$	
	
1. What proportion of the population is greater than 85 mg/100 ml?	
$Z = \frac{(X_i - \mu)}{\sigma} = \frac{85 \text{ mg/100 ml} - 65 \text{ mg/100 ml}}{25 \text{ mg/100 ml}} = 0.8$	
$P(X_i > 85 \text{ mg/100 ml}) = P(Z > 0.8) = 0.2119 \text{ or } 21.19\%$	
2. What proportion of the population is less than 45 mg/100 ml?	
$Z = \frac{45 \text{ mg/100 ml} - 65 \text{ mg/100 ml}}{25 \text{ mg/100 ml}} = -0.80$	
$P(X_i < 45 \text{ mg/100 ml}) = P(Z < -0.80) = P(Z > 0.80) = 0.2119$	
That is, the probability of selecting from this population an observation less than 0.80 standard deviations below the mean is equal to the probability of obtaining an observation greater than 0.80 standard deviations above the mean.	
3. What proportion of the population lies between 45 and 85 mg/100 ml?	
$\begin{aligned} P(45 \text{ mg/100 ml} < X_i < 85 \text{ mg/100 ml}) &= P(-0.80 < Z < 0.80) \\ &= 1.0000 - P(Z < -0.80) \\ &\quad \text{or } Z > 0.80) \\ &= 1.0000 - (0.2119 + 0.2119) \\ &= 1.0000 - 0.4238 \\ &= 0.5762 \end{aligned}$	
R	
CODE	<pre>#ex6.1b z=function(x, m, s) (x-m)/s no1=round(pnorm(z(85,65,25), lower.tail = F),4) no1 no2=round(pnorm(z(45,65,25)),4) no2 no3= 1-(no1+no2)</pre>

	<pre>no3 round(pnorm(z(85,65,25),lower.tail = T)-pnorm(z(45,65,25), lower.tail = T),4)</pre>								
OUTPUT	<pre>> #ex6.1b > z=function(x, m, s) (x-m)/s > no1=round(pnorm(z(85,65,25), lower.tail = F),4) > no1 [1] 0.2119 > no2=round(pnorm(z(45,65,25)),4) > no2 [1] 0.2119 > no3= 1-(no1+no2) > no3 [1] 0.5762 > round(pnorm(z(85,65,25),lower.tail = T)-pnorm(z(45,65,25), lower.tail = T),4) [1] 0.5763</pre>								
SAS									
CODE	<pre>data ex6_1b; a1= round(1-cdf('normal', 85, 65, 25), .0001); a2= round(cdf('normal', 45, 65, 25), .0001); a3= round(cdf('normal', 85,65,25) - cdf('normal', 45,65,25), .0001); run; proc print data=ex6_1b; run;</pre>								
OUTPUT	<table><tr><th>OBS</th><th>a1</th><th>a2</th><th>a3</th></tr><tr><td>1</td><td>0.2119</td><td>0.2119</td><td>0.5763</td></tr></table>	OBS	a1	a2	a3	1	0.2119	0.2119	0.5763
OBS	a1	a2	a3						
1	0.2119	0.2119	0.5763						
결과해석	<p>수크로오스(Sucrose)의 농도의 분포는 평균이 65(mg/100ml) 이고 표준편차가 25(mg/100ml)인 정규분포를 따르고 정규분포는 좌우대칭이므로 평균을 기준으로 대칭인 값들이 각각 크고 작을 확률은 서로 같다.</p>								

Example.6.2

EXAMPLE 6.2 Proportions of a Sampling Distribution of Means

1. A population of one-year-old children's chest circumferences has $\mu = 47.0$ cm and $\sigma = 12.0$ cm, what is the probability of drawing from it a random sample of nine measurements that has a mean larger than 50.0 cm?

$$\sigma_{\bar{X}} = \frac{12.0 \text{ cm}}{\sqrt{9}} = 4.0 \text{ cm}$$

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{50.0 \text{ cm} - 47.0 \text{ cm}}{4.0 \text{ cm}} = 0.75$$

$$P(\bar{X} > 50.0 \text{ cm}) = P(Z > 0.75) = 0.2266$$

2. What is the probability of drawing a sample of 25 measurements from the preceding population and finding that the mean of this sample is less than 40.0 cm?

$$\sigma_{\bar{X}} = \frac{12.0 \text{ cm}}{\sqrt{25}} = 2.4 \text{ cm}$$

$$Z = \frac{40.0 \text{ cm} - 47.0 \text{ cm}}{2.4 \text{ cm}} = -2.92$$

$$P(\bar{X} < 40.0 \text{ cm}) = P(Z < -2.92) = P(Z > 2.92) = 0.0018$$

3. If 500 random samples of size 25 are taken from the preceding population, how many of them would have means larger than 50.0 cm?

$$\sigma_{\bar{X}} = \frac{12.0 \text{ cm}}{\sqrt{25}} = 2.4 \text{ cm}$$

$$Z = \frac{50.0 \text{ cm} - 47.0 \text{ cm}}{2.4 \text{ g}} = 1.25$$

$$P(\bar{X} > 50.0 \text{ cm}) = P(Z > 1.25) = 0.1056$$

Therefore, $(0.1056)(500) = 53$ samples would be expected to have means larger than 50.0 cm.

R

CODE

```
#ex6.2
sdX=12/sqrt(9)
sdX
z=function(x, m, s) (x-m)/s
z(50,47,sdX)
no1=round(pnorm(z(50,47,sdX), lower.tail = F),4)
no1
sdX2=12/sqrt(25)
round(z(40,47,sdX2),2)
no2=round(pnorm(z(40,47,sdX2)),4)
no2
no3=round(pnorm(z(50,47,sdX2),lower.tail = F),4)
round(no3*500)
```

OUTPUT	<pre>> #ex6.2 > sdx=12/sqrt(9) > sdx [1] 4 > z=function(x, m, s) (x-m)/s > z(50,47,sdx) [1] 0.75 > nol=round(pnorm(z(50,47,sdx), lower.tail = F),4) > nol [1] 0.2266 > sdx2=12/sqrt(25) > round(z(40,47,sdx2),2) [1] -2.92 > no2=round(pnorm(z(40,47,sdx2)),4) > no2 [1] 0.0018 > no3=round(pnorm(z(50,47,sdx2),lower.tail = F),4) > round(no3*500) [1] 53</pre>																
SAS																	
CODE	<pre>data ex6_2; n1=9; sx1=12/sqrt(n1); a1= round(1-cdf('normal', 50,47,sx1), .0001); n2=25; sx2= 12/sqrt(n2); a2= round(cdf('normal', 40,47,sx2), .0001); a3= 500 * round(1-cdf('normal', 50, 47, sx2), .001); run; proc print data=ex6_2; run;</pre>																
OUTPUT	<table><tr><th>OBS</th><th>n1</th><th>sx1</th><th>a1</th><th>n2</th><th>sx2</th><th>a2</th><th>a3</th></tr><tr><td>1</td><td>9</td><td>4</td><td>0.2266</td><td>25</td><td>2.4</td><td>.0018</td><td>53</td></tr></table>	OBS	n1	sx1	a1	n2	sx2	a2	a3	1	9	4	0.2266	25	2.4	.0018	53
OBS	n1	sx1	a1	n2	sx2	a2	a3										
1	9	4	0.2266	25	2.4	.0018	53										
결과해석	<p>위의 문제에서 보면 1세 아기들의 가슴둘레는 평균이 47(cm)이고 표준편차가 12(cm)인 정규분포를 따른다고 가정하고 있다. n명의 1세 아기들을 표본으로 뽑아 평균 가슴둘레의 분포를 알기 위해 알려진 모분산에 n의 제곱근을 나눠준 표준오차를 사용한다.</p>																

Example.6.3

EXAMPLE 6.3 The Calculation of the Standard Error of the Mean, $s_{\bar{X}}$

The Following are Data for Systolic Blood Pressures, in mm of Mercury, of 12 Chimpanzees.

121	$n = 12$
125	$\bar{X} = \frac{1651 \text{ mm}}{12} = 137.6 \text{ mm}$
128	
134	$SS = 228,111 \text{ mm}^2 - \frac{(1651 \text{ mm})^2}{12}$
136	$= 960.9167 \text{ mm}^2$
138	
139	$s^2 = \frac{960.9167 \text{ mm}^2}{11} = 87.3561 \text{ mm}^2$
141	
144	$s = \sqrt{87.3561 \text{ mm}^2} = 9.35 \text{ mm}$
145	
149	$s_{\bar{X}} = \frac{s}{\sqrt{n}} = \frac{9.35 \text{ mm}}{\sqrt{12}} = 2.7 \text{ mm or}$
151	
$\sum X = 1651 \text{ mm}$	
$\sum X^2 = 228,111 \text{ mm}^2$	$s_{\bar{X}} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{87.3561 \text{ mm}^2}{12}} = \sqrt{7.2797 \text{ mm}^2} = 2.7 \text{ mm}$

R

CODE	<pre>#ex6.3 ex63=c(121,125,128,134,136,138,139,141,144,145,149,151) n=length(ex63) xbar=mean(ex63) ss=function(x) sum(x^2)-(sum(x))^2/length(x) sos=ss(ex63) sos varx=sos/(n-1) varx sdx=sqrt(varx) sdx sdxbar=sqrt(varx/n) sdxbar</pre>
OUTPUT	<pre>> #ex6.3 > ex63=c(121,125,128,134,136,138,139,141,144,145,149,151) > n=length(ex63) > xbar=mean(ex63) > ss=function(x) sum(x^2)-(sum(x))^2/length(x) > sos=ss(ex63) > sos [1] 960.9167 > varx=sos/(n-1) > varx [1] 87.35606 > sdx=sqrt(varx) > sdx [1] 9.346446 > sdxbar=sqrt(varx/n) > sdxbar [1] 2.698087</pre>
SAS	
CODE	<pre>data ex6_3;</pre>

	<pre>input bp @@; cards; 121 125 128 134 136 138 139 141 144 145 149 151 ; run; proc univariate data=ex6_3; run;</pre>																												
OUTPUT	<table><tr><th colspan="4">적용</th></tr><tr><td>N</td><td>12</td><td>가중합</td><td>12</td></tr><tr><td>평균</td><td>137.583333</td><td>관측값 합</td><td>1651</td></tr><tr><td>표준 편차</td><td>9.34644642</td><td>분산</td><td>87.3560606</td></tr><tr><td>왜도</td><td>-0.3827996</td><td>첨도</td><td>-0.6559762</td></tr><tr><td>제곱합</td><td>228111</td><td>수정 제곱합</td><td>960.916667</td></tr><tr><td>변동계수</td><td>6.79329842</td><td>평균의 표준 오차</td><td>2.69808668</td></tr></table>	적용				N	12	가중합	12	평균	137.583333	관측값 합	1651	표준 편차	9.34644642	분산	87.3560606	왜도	-0.3827996	첨도	-0.6559762	제곱합	228111	수정 제곱합	960.916667	변동계수	6.79329842	평균의 표준 오차	2.69808668
적용																													
N	12	가중합	12																										
평균	137.583333	관측값 합	1651																										
표준 편차	9.34644642	분산	87.3560606																										
왜도	-0.3827996	첨도	-0.6559762																										
제곱합	228111	수정 제곱합	960.916667																										
변동계수	6.79329842	평균의 표준 오차	2.69808668																										
결과해석	<p>12마리의 침팬지들의 수축기혈압(systolic blood pressures)를 측정한 결과 평균이 137.6(mm)이고 표준편차가 9.35(mm)라는 값을 구할 수 있었다.</p>																												

Example.6.4

EXAMPLE 6.4 Hypothesis Testing of $H_0: \mu = 0$ and $H_A: \mu \neq 0$

The variable, X_i , is the weight change of horses given an antibiotic for two weeks. The following measurements of X_i are those obtained from 17 horses (where a positive weight change signifies a weight gain and a negative weight change denotes a weight loss):

2.0, 1.1, 4.4, -3.1, -1.3, 3.9, 3.2, -1.6, 3.5
1.2, 2.5, 2.3, 1.9, 1.8, 2.9, -0.3, and -2.4 kg.

For these 17 data, the sample mean (\bar{X}) is 1.29 kg. Although the population variance (σ^2) is typically not known, for the demonstration purpose of this example, σ^2 is said to be 13.4621 kg². Then the population standard error of the mean would be

$$\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{13.4621 \text{ kg}^2}{17}} = \sqrt{0.7919 \text{ kg}^2} = 0.89 \text{ kg}$$

and

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{1.29 \text{ kg} - 0}{0.89 \text{ kg}} = 1.45.$$

Using Table B.2,

$$P(\bar{X} \geq 1.29 \text{ kg}) = P(Z \geq 1.45) = 0.0735$$

and, because the distribution of Z is symmetrical,

$$P(\bar{X} \leq -1.29 \text{ kg}) = P(Z \leq -1.45) = 0.0735.$$

Therefore,

$$\begin{aligned} P(\bar{X} \geq 1.29 \text{ kg or } \bar{X} \leq -1.29 \text{ kg}) \\ &= P(Z \geq 1.45 \text{ or } Z \leq -1.45) \\ &= 0.0735 + 0.0735 = 0.1470. \end{aligned}$$

As $0.1470 > 0.05$, do not reject H_0 .

R

CODE

```
#ex6.4
x=c(2.0,1.1,4.4,-3.1,-1.3,3.9,3.2,-1.6,3.5,1.2,2.5,2.3,1.9,1.8,2.9,-0.3,-2.4)
mean(x)
svar=13.4621
ssd=round(sqrt(svar/length(x)),2)
z=function(w) (w-0)/ssd
zstd=round(z(1.29),2)
pnorm(zstd,lower.tail = F)
pnorm(-zstd, lower.tail = T)
t=pnorm(zstd,lower.tail = F)+pnorm(-zstd, lower.tail = T)
t>0.05
t.test(x)
```

OUTPUT	<pre> > #ex6.4 > x=c(2.0,1.1,4.4,-3.1,-1.3,3.9,3.2,-1.6,3.5,1.2,2.5,2.3,1.9,1.8,2.9,-0.3,-2.4) > mean(x) [1] 1.294118 > svar=13.4621 > ssd=round(sqrt(svar/length(x)),2) > z=function(w) (w-0)/ssd > zstd=round(z(1.29),2) > pnorm(zstd,lower.tail = F) [1] 0.07352926 > pnorm(-zstd, lower.tail = T) [1] 0.07352926 > pnorm(zstd,lower.tail = F)==pnorm(-zstd, lower.tail = T) [1] TRUE > t=pnorm(zstd,lower.tail = F)+pnorm(-zstd, lower.tail = T) > t>0.05 [1] TRUE > t.test(x) One Sample t-test data: x t = 2.3606, df = 16, p-value = 0.03127 alternative hypothesis: true mean is not equal to 0 95 percent confidence interval: 0.1319352 2.4563001 sample estimates: mean of x 1.294118 </pre>				
SAS					
CODE	<pre> proc iml; x1={2.0,1.1,4.4,-3.1,-1.3,3.9,3.2,-1.6,3.5,1.2,2.5,2.3,1.9, 1.8,2.9,-0.3,-2.4}; xbar= mean(x1); popv=13.4621; stderr=sqrt(popv/nrow(x1)); zpu= 1-cdf('normal',1.29, 0, stderr) ; zpl= cdf('normal',-1.29, 0, stderr); p = zpu+zpl; print stderr p ; </pre>				
OUTPUT	<table border="1"> <thead> <tr> <th>stderr</th><th>p</th></tr> </thead> <tbody> <tr> <td>0.889881</td><td>0.1471612</td></tr> </tbody> </table>	stderr	p	0.889881	0.1471612
stderr	p				
0.889881	0.1471612				
결과해석	<p>말들의 몸무게의 변화가 있는지를 확인하기 위해 17마리의 말들을 뽑아 가설검정을 하려고 한다. 여기서는 평균이 표본평균 1.29보다 크거나 -1.29보다 작을 확률을 더한 값이 유의수준 0.05보다 크므로 말들의 몸무게의 변화가 없다는 귀무가설 기각할 수 없다. 따라서 말들의 몸무게의 변화는 없다고 볼 수 있다.</p>				

value 2 p - .

t.test()
가 .

Example.6.5

EXAMPLE 6.5 Probability of Rejecting a True Null Hypothesis

Hypothetical outcomes of testing the same null hypothesis for 2500 random samples of the same size from the same population (where the samples are taken with replacement).

	If H_0 is true	If H_0 is false	Row total
If H_0 is rejected	100	450	550
If H_0 is not rejected	1900	50	1950
Column total	2000	500	2500

Probability that H_0 is rejected if H_0 is true = $100/2000 = 0.05$.

Probability that H_0 is true if H_0 is rejected = $100/550 = 0.18$.

R

CODE	<pre>#ex6.5 type1err=100/2000 type1err type2err=100/550 type2err</pre>
OUTPUT	<pre>> #ex6.5 > type1err=100/2000 > type1err [1] 0.05 > type2err=100/550 > type2err [1] 0.1818182</pre>

SAS

CODE	<pre>data ex6_5; input ho\$ rej\$ size @@; cards; T REJECT 100 T NOTREJECT 1900 F REJECT 450 F NOTREJECT 50 ; run; proc freq data=ex6_5; table ho * rej /norow nocol nopercnt; weight size; run;</pre>
------	--

OUTPUT	테이블 ho * rej			
		rej		
	ho	NOTREJEC	REJECT	합계
	F	50	450	500
	T	1900	100	2000
	합계	1950	550	2500
결과해석	귀무가설이 참인데 귀무가설을 기각할, 즉 제 1종오류를 범할 확률은 0.05, 귀무가설이 거짓인데도 귀무가설을 기각하지 않는, 제 2종오류를 범할 확률은 0.18을 차지하고 있다.			

Example.6.6	
CODE	<p>EXAMPLE 6.6 Confidence Limits for the Mean</p> <p>For the 17 data in Example 6.4, $\bar{X} = 1.29$ kg and $\sigma_{\bar{X}} = 0.89$ kg. We can calculate the 95% confidence limits for μ using Equations 6.13 and 6.14 and $Z_{0.05(2)} = 1.96$:</p> $L_1 = \bar{X} - Z_{\alpha(2)}\sigma_{\bar{X}}$ $= 1.29 \text{ kg} - (1.96)(0.89 \text{ kg})$ $= 1.29 \text{ kg} - 1.74 \text{ kg} = -0.45 \text{ kg}$ $L_2 = \bar{X} + Z_{\alpha(2)}\sigma_{\bar{X}}$ $= 1.29 \text{ kg} + (1.96)(0.89 \text{ kg})$ $= 1.29 \text{ kg} + 1.74 \text{ kg} = 3.03 \text{ kg}.$ <p>So, the 95% confidence interval could be stated as</p> $P(-0.45 \text{ kg} \leq \mu \leq 3.03 \text{ kg}).$ <p>Note that the μ_0 of Example 6.4 (namely 0) is included between L_1 and L_2, indicating that H_0 is not rejected.</p>
	R
CODE	<pre>#ex6.6 x=c(2.0,1.1,4.4,-3.1,-1.3,3.9,3.2,-1.6,3.5,1.2,2.5,2.3,1.9,1.8,2.9,-0.3,-2.4) xbar=round(mean(x),2) svar=13.4621 ssd=round(sqrt(svar/length(x)),2) z0.05=round(qnorm(0.025, lower.tail = F),2) z0.05 l1=xbar-z0.05*ssd l1 l2=xbar+z0.05*ssd l2</pre>

OUTPUT	<pre> > #ex6.6 > x=c(2.0,1.1,4.4,-3.1,-1.3,3.9,3.2,-1.6,3.5,1.2,2.5,2.3,1.9,1.8,2.9,-0.3,-2.4) > xbar=round(mean(x),2) > svar=13.4621 > ssd=round(sqrt(svar/length(x)),2) > z0.05=round(qnorm(0.025, lower.tail = F),2) > z0.05 [1] 1.96 > l1=xbar-z0.05*ssd > l1 [1] -0.4544 > l2=xbar+z0.05*ssd > l2 [1] 3.0344 </pre>				
SAS					
CODE	<pre> proc iml; x1={2.0,1.1,4.4,-3.1,-1.3,3.9,3.2,-1.6,3.5,1.2,2.5,2.3,1.9, 1.8,2.9,-0.3,-2.4}; xbar= mean(x1); popv=13.4621; stderr=sqrt(popv/nrow(x1)); lowerlimit=xbar-1.96*stderr; upperlimit=xbar+1.96*stderr; print lowerlimit upperlimit; run; quit; </pre>				
OUTPUT	<table border="1"> <thead> <tr> <th>lowerlimit</th><th>upperlimit</th></tr> </thead> <tbody> <tr> <td>-0.450049</td><td>3.0382845</td></tr> </tbody> </table>	lowerlimit	upperlimit	-0.450049	3.0382845
lowerlimit	upperlimit				
-0.450049	3.0382845				
결과해석	<p>예제 6.4에서 구해본 말들의 몸무게 변화에 차이가 있는지에 대한 검정을 다른 방법으로 확인해보기 위해 표본평균에 대한 신뢰구간 범위를 구해봤더니 [-0.45, 3.04]가 나왔다. 0이 구한 범위 안에 속하므로 귀무가설을 기각할 수 없다. 위 방법을 통해서도 같은 결과를 얻을 수 있다는 것을 알 수 있다.</p>				

Example.6.7

EXAMPLE 6.7 The Heights of the First 70 Graduate Students in My Biostatistics Course

Height (X_i) (in.)	Observed Frequency (f_i)	Cumulative Frequency (cum. f_i)	$f_i X_i$ (in.)	$f_i X_i^2$ (in. ²)
63	2	2	126	7,938
64	2	4	128	8,192
65	3	7	195	12,675
66	5	12	330	21,780
67	4	16	268	17,956
68	6	22	408	27,744
69	5	27	345	23,805
70	8	35	560	39,200
71	7	42	497	35,287
72	7	49	504	36,288
73	10	59	730	53,290
74	6	65	444	32,856
75	3	68	225	16,875
76	2	70	152	11,552
$\Sigma f_i =$ $n = 70$			$\Sigma f_i X_i =$ 4,912 in.	$\Sigma f_i X_i^2 =$ 345,438 in. ²

$$SS = \Sigma f_i X_i^2 - \frac{(\Sigma f_i X_i)^2}{n} = 345,438 \text{ in.}^2 - \frac{(4,912 \text{ in.})^2}{70} = 755.9429 \text{ in.}^2$$

$$s^2 = \frac{SS}{n - 1} = \frac{755.9429 \text{ in.}^2}{69} = 10.9557 \text{ in.}^2$$

Figure 6.9

R

CODE	<pre>#ex6.7 height=c(63,64,65,66,67,68,69,70,71,72,73,74,75,76) frq=c(2,2,3,5,4,6,5,8,7,7,10,6,3,2) fh=height*frq fhs=height^2*frq ss=sum(fhs)-(sum(fh)^2)/sum(frq) ss sdx=ss/(sum(frq)-1) sdx</pre>
OUTPUT	<pre>> #ex6.7 > height=c(63,64,65,66,67,68,69,70,71,72,73,74,75,76) > frq=c(2,2,3,5,4,6,5,8,7,7,10,6,3,2) > fh=height*frq > fhs=height^2*frq > ss=sum(fhs)-(sum(fh)^2)/sum(frq) > ss [1] 755.9429 > sdx=ss/(sum(frq)-1) > sdx [1] 10.95569</pre>
SAS	
CODE	<pre>proc iml;</pre>

	<pre>height={63,64,65,66,67,68,69,70,71,72,73,74,75,76}; frq={2,2,3,5,4,6,5,8,7,7,10,6,3,2}; hf=height # frq; hsf = (height#height) # frq ; ss= sum(hsf) - sum(hf)**2/sum(frq); sv= ss/ sum(frq); print sv ; run; quit;</pre>		
OUTPUT	<table><tr><td>SV</td></tr><tr><td>10.799184</td></tr></table>	SV	10.799184
SV			
10.799184			
결과해석	생물통계학 수업을 듣는 70명의 대학원생들의 표본표준편차를 구하면 위와 같은 방법으로 구할 수 있다.		

Example.7.1	
<p>EXAMPLE 7.1 The Two-Tailed t Test for Difference between a Population Mean and a Hypothesized Population Mean</p> <p>Body temperatures (measured in °C) of 25 intertidal crabs placed in air at 24.3°C: 25.8, 24.6, 26.1, 22.9, 25.1, 27.3, 24.0, 24.5, 23.9, 26.2, 24.3, 24.6, 23.3, 25.5, 28.1, 24.8, 23.5, 26.3, 25.4, 25.5, 23.9, 27.0, 24.8, 22.9, 25.4.</p> $H_0: \mu = 24.3^\circ\text{C}$ $H_A: \mu \neq 24.3^\circ\text{C}$ $\alpha = 0.05$ $n = 25$ $\bar{X} = 25.03^\circ\text{C}$ $s^2 = 1.80(^{\circ}\text{C})^2$ $s_{\bar{X}} = \sqrt{\frac{1.80(^{\circ}\text{C})^2}{25}} = 0.27^\circ\text{C}$ $t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{25.03^\circ\text{C} - 24.3^\circ\text{C}}{0.27^\circ\text{C}} = \frac{0.73^\circ\text{C}}{0.27^\circ\text{C}} = 2.704$ $\nu = 24$ $t_{0.05(2), 24} = 2.064$ <p>As $t > t_{0.05(2), 24}$, reject H_0 and conclude that the sample of 25 body temperatures came from a population whose mean is not 24.3°C.</p> $0.01 < P < 0.02 [P = 0.012]^*$	
R	
CODE	<pre>#ex7.1 temp=c(25.8,24.6,26.1,22.9,25.1,27.3,24,24.5,23.9,26.2,24.3,24.6,23.3,25.5,28.1,24.8, 23.5,26.3,25.4,25.5,23.9,27,24.8,22.9,25.4) m0=24.3 n=25 xbar=round(mean(temp),2)</pre>

	<pre>xbar varx=round(var(temp),2) sdxbars=round(sqrt(varx)/sqrt(n),2) sdxbars t=(xbar-m0)/sdxbars t tp=qt(0.025,24, lower.tail = F) t>tp t.test(temp,mu=24.3, alternative = "two.sided")</pre>						
OUTPUT	<pre>> #ex7.1 > temp=c(25.8,24.6,26.1,22.9,25.1,27.3,24,24.5,23.9,26.2,24.3,24.6,23.3,25.5,28.1,24.8,23.5 > m0=24.3 > n=25 > xbar=round(mean(temp),2) > xbar [1] 25.03 > varx=round(var(temp),2) > sdxbars=round(sqrt(varx)/sqrt(n),2) > sdxbars [1] 0.27 > t=(xbar-m0)/sdxbars > t [1] 2.703704 > tp=qt(0.025,24, lower.tail = F) > t>tp [1] TRUE > t.test(temp,mu=24.3, alternative = "two.sided") One Sample t-test data: temp t = 2.7128, df = 24, p-value = 0.01215 alternative hypothesis: true mean is not equal to 24.3 95 percent confidence interval: 24.47413 25.58187 sample estimates: mean of x 25.028</pre>						
SAS							
CODE	<pre>data ex7_1; input temp @@; cards; 25.8 24.6 26.1 22.9 25.1 27.3 24 24.5 23.9 26.2 24.3 24.6 23.3 25.5 28.1 24.8 23.5 26.3 25.4 25.5 23.9 27 24.8 22.9 25.4 ; run; proc ttest data=ex7_1 h0=24.3 ; var temp; run; proc univariate data=ex7_1 mu0=24.3; var temp; run;</pre>						
OUTPUT	<table><tr><th>DF</th><th>t Value</th><th>Pr > t </th></tr><tr><td>24</td><td>2.71</td><td>0.0121</td></tr></table>	DF	t Value	Pr > t	24	2.71	0.0121
DF	t Value	Pr > t					
24	2.71	0.0121					

	위치모수 검정: $\mu_0=24.3$			
	검정	통계량	p 값	
	스튜던트의 t	t	2.712769	Pr > t 0.0121
결과해석	조간대 해안에 서식하는 게들의 체온의 평균이 24.3인가에 대해 가설검정을 실시한 결과, p-value가 0.0121정도로 유의수준 0.05 하에서 귀무가설을 기각하였다. 따라서 게들의 체온의 평균은 24.3이 아니다.			

Example.7.2																									
CODE	<p>EXAMPLE 7.2 A Two-Tailed Test for Significant Difference between a Population Mean and a Hypothesized Population Mean of Zero</p> <p>Weight change of twelve rats after being subjected to a regimen of forced exercise. Each weight change (in g) is the weight after exercise minus the weight before.</p> <table> <tr><td>1.7</td><td>$H_0: \mu = 0$</td></tr> <tr><td>0.7</td><td>$H_A: \mu \neq 0$</td></tr> <tr><td>-0.4</td><td>$\alpha = 0.05$</td></tr> <tr><td>-1.8</td><td>$n = 12$</td></tr> <tr><td>0.2</td><td>$\bar{X} = -0.65 \text{ g}$</td></tr> <tr><td>0.9</td><td>$s^2 = 1.5682 \text{ g}^2$</td></tr> <tr><td>-1.2</td><td>$s_{\bar{X}} = \sqrt{\frac{1.5682 \text{ g}^2}{12}} = 0.36 \text{ g}$</td></tr> <tr><td>-0.9</td><td>$t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{-0.65 \text{ g}}{0.36 \text{ g}} = -1.81$</td></tr> <tr><td>-1.8</td><td>$\nu = n - 1 = 11$</td></tr> <tr><td>-1.4</td><td>$t_{0.05(2),11} = 2.201$</td></tr> <tr><td>-1.8</td><td>Since $t < t_{0.05(2),11}$, do not reject H_0.</td></tr> <tr><td>-2.0</td><td>$0.05 < P < 0.10 [P = 0.098]$</td></tr> </table> <p>Therefore, we conclude that the exercise does not cause a weight change in the population from which this sample came.</p>	1.7	$H_0: \mu = 0$	0.7	$H_A: \mu \neq 0$	-0.4	$\alpha = 0.05$	-1.8	$n = 12$	0.2	$\bar{X} = -0.65 \text{ g}$	0.9	$s^2 = 1.5682 \text{ g}^2$	-1.2	$s_{\bar{X}} = \sqrt{\frac{1.5682 \text{ g}^2}{12}} = 0.36 \text{ g}$	-0.9	$t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{-0.65 \text{ g}}{0.36 \text{ g}} = -1.81$	-1.8	$\nu = n - 1 = 11$	-1.4	$t_{0.05(2),11} = 2.201$	-1.8	Since $ t < t_{0.05(2),11}$, do not reject H_0 .	-2.0	$0.05 < P < 0.10 [P = 0.098]$
1.7	$H_0: \mu = 0$																								
0.7	$H_A: \mu \neq 0$																								
-0.4	$\alpha = 0.05$																								
-1.8	$n = 12$																								
0.2	$\bar{X} = -0.65 \text{ g}$																								
0.9	$s^2 = 1.5682 \text{ g}^2$																								
-1.2	$s_{\bar{X}} = \sqrt{\frac{1.5682 \text{ g}^2}{12}} = 0.36 \text{ g}$																								
-0.9	$t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{-0.65 \text{ g}}{0.36 \text{ g}} = -1.81$																								
-1.8	$\nu = n - 1 = 11$																								
-1.4	$t_{0.05(2),11} = 2.201$																								
-1.8	Since $ t < t_{0.05(2),11}$, do not reject H_0 .																								
-2.0	$0.05 < P < 0.10 [P = 0.098]$																								
R																									
CODE	<pre>#ex7.2 weight=c(1.7,0.7,-0.4,-1.8,0.2,0.9,-1.2,-0.9,-1.8,-1.4,-1.8,-2.0) m0=0 n=12 xbar=round(mean(weight),2) xbar varx=round(var(weight),2) sdxbar=round(sqrt(varx)/sqrt(n),2) sdxbar t=(xbar-m0)/sdxbar t tp=qt(0.025,11, lower.tail = F) abs(t)<tp t.test(weight,alternative = "two.sided")</pre>																								

OUTPUT	<pre>> #ex7.2 > weight=c(1.7,0.7,-0.4,-1.8,0.2,0.9,-1.2,-0.9,-1.8,-1.4,-1.8,-2.0) > m0=0 > n=12 > xbar=round(mean(weight),2) > xbar [1] -0.65 > varx=round(var(weight),2) > sdxbar=round(sqrt(varx)/sqrt(n),2) > sdxbar [1] 0.36 > t=(xbar-m0)/sdxbar > t [1] -1.805556 > tp=qt(0.025,11, lower.tail = F) > abs(t)<tp [1] TRUE > t.test(weight,alternative = "two.sided") One Sample t-test data: weight t = -1.7981, df = 11, p-value = 0.09964 alternative hypothesis: true mean is not equal to 0 95 percent confidence interval: -1.4456548 0.1456548 sample estimates: mean of x -0.65</pre>																		
SAS																			
CODE	<pre>data ex7_2; input wc @@; cards; 1.7 0.7 -0.4 -1.8 0.2 0.9 -1.2 -0.9 -1.8 -1.4 -1.8 -2.0 ; run; proc ttest data=ex7_2 h0=0; var wc; run; proc univariate data=ex7_2 mu0=0; var wc; run;</pre>																		
OUTPUT	<table><tr><th>DF</th><th>t Value</th><th>Pr > t </th></tr><tr><td>11</td><td>-1.80</td><td>0.0996</td></tr></table> <table><tr><th colspan="4">위치모수 검정: Mu0=0</th></tr><tr><th>검정</th><th>통계량</th><th colspan="2">p 값</th></tr><tr><td>스튜던트의 t</td><td>t</td><td>-1.79807</td><td>Pr > t 0.0996</td></tr></table>	DF	t Value	Pr > t	11	-1.80	0.0996	위치모수 검정: Mu0=0				검정	통계량	p 값		스튜던트의 t	t	-1.79807	Pr > t 0.0996
DF	t Value	Pr > t																	
11	-1.80	0.0996																	
위치모수 검정: Mu0=0																			
검정	통계량	p 값																	
스튜던트의 t	t	-1.79807	Pr > t 0.0996																
결과해석	운동을 한 후 쥐들의 몸무게의 변화가 있을지에 대해 가설검정을 실시하였다. 12마리의 쥐들을 뽑아서 실시하였더니 p-value가 0.0996으로 유의수준 0.05하에서 귀무가설을																		

	기각할 근거를 얻지 못하였다. 따라서 12마리의 표본으로는 운동을 실시한 후 쥐들의 몸무게 변화가 있다고 판단하기는 어렵다.
--	---

Example.7.3

EXAMPLE 7.3 A One-Tailed t Test for the Hypotheses $H_0: \mu \geq 0$ and $H_A: \mu < 0$

The data are weight changes of humans, tabulated after administration of a drug proposed to result in weight loss. Each weight change (in kg) is the weight after minus the weight before drug administration.

$$\begin{array}{ll}
 0.2 & n = 12 \\
 -0.5 & \bar{X} = -0.61 \text{ kg} \\
 -1.3 & s^2 = 0.4008 \text{ kg}^2 \\
 -1.6 & s_{\bar{X}} = \sqrt{\frac{0.4008 \text{ kg}^2}{12}} = 0.18 \text{ kg} \\
 -0.7 & \\
 0.4 & t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{-0.61 \text{ kg}}{0.18 \text{ kg}} = -3.389 \\
 -0.1 & \\
 0.0 & \nu = n - 1 = 11 \\
 -0.6 & t_{0.05(1), 11} = 1.796. \\
 -1.1 & \text{If } t \leq -t_{0.05(1), 11}, \text{ reject } H_0. \\
 -1.2 & \text{Conclusion: reject } H_0. \\
 -0.8 &
 \end{array}$$

$$0.0025 < P(t \leq -3.389) < 0.005 [P = 0.0030]$$

We conclude that the drug does cause weight loss.

R

	<pre> #ex7.3 weight=c(0.2,-0.5,-1.3,-1.6,-0.7,0.4,-0.1,0,-0.6,-1.1,-1.2,-0.8) ttest=function(x, m0, n) { m0=m0 n=length(x) xbar=round(mean(x),2) sdxb=round(sqrt(var(x)/n),2) tp=qt(0.05, n-1, lower.tail = F) t=(xbar-m0)/sdxb answer=c(t, tp) return(answer) } ttest(weight,0,12) t.test(weight, alternative = "less") </pre>
CODE	

OUTPUT	<pre>> #ex7.3 > weight=c(0.2,-0.5,-1.3,-1.6,-0.7,0.4,-0.1,0,-0.6,-1.1,-1.2,-0.8) > ttest=function(x, m0, n) { + m0=m0 + n=length(x) + xbar=round(mean(x),2) + sdxbar=round(sqrt(var(x)/n),2) + tp=qt(0.05, n-1, lower.tail = F) + t=(xbar-m0)/sdxbar + answer=c(t, tp) + return(answer) + } > ttest(weight,0,12) [1] -3.388889 1.795885 > t.test(weight, alternative = "less") One Sample t-test data: weight t = -3.3285, df = 11, p-value = 0.003364 alternative hypothesis: true mean is less than 0 95 percent confidence interval: -Inf -0.2801098 sample estimates: mean of x -0.6083333</pre>						
SAS							
CODE	<pre>%macro test(dt, m, x, var1); proc ttest data=&dt h0=&m sides=&x; var &var1; run; proc univariate data=&dt; var &var1; run; %mend; data ex7_3; input wc2 @@; cards; 0.2 -0.5 -1.3 -1.6 -0.7 0.4 -0.1 0 -0.6 -1.1 -1.2 -0.8 ; run; %test(ex7_3, 0, 1, wc2);</pre>						
OUTPUT	<table><tr><th>DF</th><th>t Value</th><th>Pr < t</th></tr><tr><td>11</td><td>-3.33</td><td>0.0034</td></tr></table>	DF	t Value	Pr < t	11	-3.33	0.0034
DF	t Value	Pr < t					
11	-3.33	0.0034					

	적용			
	N	12	가중합	12
	평균	-0.6083333	관측값 합	-7.3
	표준 편차	0.633114	분산	0.40083333
	왜도	0.11659779	첨도	-1.0489817
	제곱합	8.85	수정 제곱합	4.40916667
	변동계수	-104.07353	평균의 표준 오차	0.18276427
결과해석	체중감량 효과가 있다고 제안된 약을 복용한 사람들의 몸무게 변화가 있는지에 대해 가설검정을 실시한 결과 p-value가 0.0030으로 매우 유의한 값이 나와 귀무가설을 기각할 수 있다. 따라서 그 약은 체중감소를 시키는 효과가 있다고 판단할 수 있다.			

Example.7.4	
<p>EXAMPLE 7.4 The One-Tailed t Test for the Hypotheses $H_0: \mu \leq 45$ sec and $H_A: \mu > 45$ sec</p> <p>Dissolving times (in sec) of a drug in gastric juice: 42.7, 43.4, 44.6, 45.1, 45.6, 45.9, 46.8, 47.6.</p> <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> $H_0: \mu \leq 45 \text{ sec}$ $H_A: \mu > 45 \text{ sec}$ $\alpha = 0.05$ $n = 8$ $\bar{X} = 45.21 \text{ sec}$ $SS = 18.8288 \text{ sec}^2$ $s^2 = 2.6898 \text{ sec}^2$ </div> <div style="width: 45%;"> $s_{\bar{X}} = 0.58 \text{ sec}$ $t = \frac{45.21 \text{ sec} - 45 \text{ sec}}{0.58 \text{ sec}} = 0.36$ $\nu = 7$ $t_{0.05(1),7} = 1.895$ If $t \geq t_{0.05(1),7}$, reject H_0. Conclusion: do not reject H_0. </div> </div> <p style="text-align: center;">$P(t \geq 0.36) > 0.25 [P = 0.36]$</p> <p>We conclude that the mean dissolving time is not greater than 45 sec.</p>	
R	
CODE	<pre>#ex7.4 time=c(42.7,43.4,44.6,45.1,45.6,45.9,46.8,47.6) ttest(time,45,8) t.test(time,mu=45, alternative = "greater")</pre>

OUTPUT	<pre>> #ex7.4 > time=c(42.7,43.4,44.6,45.1,45.6,45.9,46.8,47.6) > ttest(time,45,8) [1] 0.362069 1.894579 > t.test(time,mu=45, alternative = "greater") One Sample t-test data: time t = 0.36647, df = 7, p-value = 0.3624 alternative hypothesis: true mean is greater than 45 95 percent confidence interval: 44.11393 Inf sample estimates: mean of x 45.2125</pre>																																		
SAS																																			
CODE	<pre>data ex7_4; input time @@; cards; 42.7 43.4 44.6 45.1 45.6 45.9 46.8 47.6 ; run; %test(ex7_4, 45,u, time);</pre>																																		
OUTPUT	<table><tr><th>DF</th><th>t Value</th><th>Pr > t</th></tr><tr><td>7</td><td>0.37</td><td>0.3624</td></tr></table> <table><tr><th colspan="4">적률</th></tr><tr><td>N</td><td>8</td><td>가중합</td><td>8</td></tr><tr><td>평균</td><td>45.2125</td><td>관측값 합</td><td>361.7</td></tr><tr><td>표준 편차</td><td>1.64006751</td><td>분산</td><td>2.68982143</td></tr><tr><td>왜도</td><td>-0.1750017</td><td>첨도</td><td>-0.6411801</td></tr><tr><td>제곱합</td><td>16372.19</td><td>수정 제곱합</td><td>18.82875</td></tr><tr><td>변동계수</td><td>3.62746477</td><td>평균의 표준 오차</td><td>0.57985143</td></tr></table>	DF	t Value	Pr > t	7	0.37	0.3624	적률				N	8	가중합	8	평균	45.2125	관측값 합	361.7	표준 편차	1.64006751	분산	2.68982143	왜도	-0.1750017	첨도	-0.6411801	제곱합	16372.19	수정 제곱합	18.82875	변동계수	3.62746477	평균의 표준 오차	0.57985143
DF	t Value	Pr > t																																	
7	0.37	0.3624																																	
적률																																			
N	8	가중합	8																																
평균	45.2125	관측값 합	361.7																																
표준 편차	1.64006751	분산	2.68982143																																
왜도	-0.1750017	첨도	-0.6411801																																
제곱합	16372.19	수정 제곱합	18.82875																																
변동계수	3.62746477	평균의 표준 오차	0.57985143																																
결과해석	<p>위액에 의해 약이 녹는 시간의 평균이 45초 이상인지 알아보기 위해 가설검정을 실시한 결과 p-value가 0.36으로 귀무가설을 기각할 충분한 근거를 얻지 못했다. 따라서 아직은 약이 위액에 녹는 속도의 평균이 45초 이상이라고 판단할 수 없다.</p>																																		

...

Example.7.5

EXAMPLE 7.5 Computation of Confidence Intervals and Confidence Limits for the Mean, Using the Data of Example 7.1

(a) At the 95% confidence level:

$$\bar{X} = 25.03^{\circ}\text{C}$$

$$s_{\bar{X}} = 0.27^{\circ}\text{C}$$

$$t_{0.05(2), 24} = 2.064$$

$$\nu = 24$$

$$\begin{aligned} 95\% \text{ confidence interval} &= \bar{X} \pm t_{0.05(2), 24} s_{\bar{X}} \\ &= 25.03^{\circ}\text{C} \pm (2.064)(0.27^{\circ}\text{C}) \\ &= 25.03^{\circ}\text{C} \pm 0.56^{\circ}\text{C} \end{aligned}$$

$$95\% \text{ confidence limits: } L_1 = 25.03^{\circ}\text{C} - 0.56^{\circ}\text{C} = 24.47^{\circ}\text{C}$$

$$L_2 = 25.03^{\circ}\text{C} + 0.56^{\circ}\text{C} = 25.59^{\circ}\text{C}$$

(b) At the 99% confidence level:

$$t_{0.01(2), 24} = 2.797$$

$$\begin{aligned} 99\% \text{ confidence interval} &= \bar{X} \pm t_{0.01(2), 24} s_{\bar{X}} \\ &= 25.03^{\circ} \pm (2.797)(0.27^{\circ}\text{C}) \\ &= 25.03^{\circ}\text{C} \pm 0.76^{\circ}\text{C} \end{aligned}$$

$$99\% \text{ confidence limits: } L_1 = 25.03^{\circ}\text{C} - 0.76^{\circ}\text{C} = 24.27^{\circ}\text{C}$$

$$L_2 = 25.03^{\circ}\text{C} + 0.76^{\circ}\text{C} = 25.79^{\circ}\text{C}$$

In both parts (a) and (b), the hypothesized value, $\mu_0 = 24.3^{\circ}\text{C}$ in Example 7.1, lies outside the confidence intervals. This indicates that H_0 would be rejected using either the 5% or the 1% level of significance.

R

CODE

```
#ex7.5
ci=function(x, alpha){
  alpha=alpha
  n=length(x)
  v=n-1
  mu=mean(x)
  sx=round(sqrt(var(x)/n),2)
  tp=round(qt(alpha,v,lower.tail = F),3)
  lowerlimit=round(mu-sx*tp,2)
  upperlimit=round(mu+sx*tp,2)
  answer=c(lowerlimit, upperlimit)
  return(answer)
}
ci(temp, 0.025)
ci(temp,0.005)

cit=t.test(temp, mu=24.3, alternative = "two.sided")
cit$conf.int
```

OUTPUT	<pre>> #ex7.5 > ci=function(x, alpha){ + alpha=alpha + n=length(x) + v=n-1 + mu=mean(x) + sx=round(sqrt(var(x)/n),2) + tp=round(qt(alpha,v,lower.tail = F),3) + lowerlimit=round(mu-sx*tp,2) + upperlimit=round(mu+sx*tp,2) + answer=c(lowerlimit, upperlimit) + return(answer) + } > ci(temp, 0.025) [1] 24.47 25.59 > ci(temp,0.005) [1] 24.27 25.78 > > cit=t.test(temp, mu=24.3, alternative = "two.sided") > cit\$conf.int [1] 24.47413 25.58187 attr(,"conf.level") [1] 0.95</pre>												
SAS													
CODE	<pre>data ex7_5; input temp @@; cards; 25.8 24.6 26.1 22.9 25.1 27.3 24 24.5 23.9 26.2 24.3 24.6 23.3 25.5 28.1 24.8 23.5 26.3 25.4 25.5 23.9 27 24.8 22.9 25.4 ; run; proc means data= ex7_5 clm alpha=0.05; var temp; run; proc means data= ex7_5 clm alpha=0.01; var temp; run;</pre>												
OUTPUT	<table><tr><th colspan="2">분석 변수: temp</th></tr><tr><th>Lower 95% CL for Mean</th><th>Upper 95% CL for Mean</th></tr><tr><td>24.4741312</td><td>25.5818688</td></tr></table> <table><tr><th colspan="2">분석 변수: temp</th></tr><tr><th>Lower 99% CL for Mean</th><th>Upper 99% CL for Mean</th></tr><tr><td>24.2774120</td><td>25.7785880</td></tr></table>	분석 변수: temp		Lower 95% CL for Mean	Upper 95% CL for Mean	24.4741312	25.5818688	분석 변수: temp		Lower 99% CL for Mean	Upper 99% CL for Mean	24.2774120	25.7785880
분석 변수: temp													
Lower 95% CL for Mean	Upper 95% CL for Mean												
24.4741312	25.5818688												
분석 변수: temp													
Lower 99% CL for Mean	Upper 99% CL for Mean												
24.2774120	25.7785880												

결과해석	계의 평균 체온에 대해 95%, 99%의 신뢰구간을 위의 결과와 같이 구했다. 95%, 99% 신뢰구간 모두 24.3을 포함하고 있지 않으므로 여기서도 귀무가설을 기각할 수 있다는 것을 알 수 있다. 그리고 신뢰수준이 높을수록 신뢰구간의 간격이 더 넓어진다는 것을 알 수 있다.
------	---

Example.7.6	
CODE	<p>EXAMPLE 7.6 Prediction Limits for Additional Sampling from the Population Sampled in Example 7.1</p> <p>From Example 7.1, which is a sample of 25 crab body temperatures,</p> $n = 25, \bar{X} = 25.03^{\circ}\text{C}, \text{ and } s^2 = 1.80(^{\circ}\text{C})^2.$ <p>(a) If we intend to collect 8 additional crab body temperatures from the same population from which the 25 data in Example 7.1 came, then (by Equation 7.7) we can be 95% confident that the mean of those 8 data will be within this prediction interval:</p> $\begin{aligned} 25.03^{\circ}\text{C} \pm t_{0.05(2),24} \sqrt{\frac{1.80(^{\circ}\text{C})^2}{8} + \frac{1.80(^{\circ}\text{C})^2}{2}} \\ = 25.03^{\circ}\text{C} \pm 2.064(0.545^{\circ}\text{C}) \\ = 25.03^{\circ}\text{C} \pm 1.12^{\circ}\text{C}. \end{aligned}$ <p>Therefore, the 95% prediction limits for the predicted mean of these additional data are $L_1 = 23.91^{\circ}\text{C}$ and $L_2 = 26.15^{\circ}\text{C}$.</p> <p>(b) If we intend to collect 1 additional crab body temperature from the same population from which the 25 data in Example 7.1 came, then (by Equation 7.8) we can be 95% confident that the additional datum will be within this prediction interval:</p> $\begin{aligned} 25.03^{\circ}\text{C} \pm t_{0.05(2),24} \sqrt{1.80(^{\circ}\text{C})^2 + \frac{1.80(^{\circ}\text{C})^2}{2}} \\ = 25.03^{\circ}\text{C} \pm 2.064(1.368^{\circ}\text{C}) \\ = 25.03^{\circ}\text{C} \pm 2.82^{\circ}\text{C}. \end{aligned}$ <p>Therefore, the 95% prediction limits for this predicted datum are $L_1 = 22.21^{\circ}\text{C}$ and $L_2 = 27.85^{\circ}\text{C}$.</p>
	R
CODE	<pre>#ex7.6 ci=function(x, alpha, a){ alpha=alpha a=a n=length(x) v=n-1 mu=mean(x) sx=round(var(x)/n,2) sx2=round(var(x)/a,2) tp=round(qt(alpha,v,lower.tail = F),3) lowerlimit=round(mu-sqrt(sx+sx2)*tp,2) upperlimit=round(mu+sqrt(sx+sx2)*tp,2) answer=c(lowerlimit, upperlimit) return(answer) }</pre>

	ci(temp, 0.025, 8) ci(temp, 0.025, 1)											
OUTPUT	<pre>> #ex7.6 > ci=function(x, alpha, a){ + alpha=alpha + a=a + n=length(x) + v=n-1 + mu=mean(x) + sx=round(var(x)/n,2) + sx2=round(var(x)/a,2) + tp=round(qt(alpha,v,lower.tail = F),3) + lowerlimit=round(mu-sqrt(sx+sx2)*tp,2) + upperlimit=round(mu+sqrt(sx+sx2)*tp,2) + answer=c(lowerlimit, upperlimit) + return(answer) + }</pre> <pre>> ci(temp, 0.025, 8) [1] 23.90 26.16 > ci(temp, 0.025, 1) [1] 22.21 27.85</pre>											
SAS												
CODE	<pre>proc iml; use ex7_1; read all; xba=mean(temp); ss=var(temp); t1= -tinv(0.025, 24); low1=xba-t1*sqrt(ss/8+ss/25); up1=xba+t1*sqrt(ss/8+ss/25); low2=xba-t1*sqrt(ss/1+ss/25); up2=xba+t1*sqrt(ss/1+ss/25); print low1 up1 low2 up2 ; run; quit;</pre>											
OUTPUT	<table><tr><th>low1</th><th>up1</th><th>low2</th><th>up2</th></tr><tr><td>23.903087</td><td>26.152913</td><td>22.203812</td><td>27.852188</td></tr></table>				low1	up1	low2	up2	23.903087	26.152913	22.203812	27.852188
low1	up1	low2	up2									
23.903087	26.152913	22.203812	27.852188									
결과해석	ex7.5에서 구한 신뢰구간에 추가로 표본을 추가하여 다시 신뢰구간을 구해보면 신뢰구간의 간격이 더 넓어지는 것을 알 수 있다. 그리고 더 적은 수의 표본을 추가할수록 신뢰구간의 간격이 더 넓어진다는 것도 확인할 수 있다.											

<http://www.math.wpi.edu/saspdf/qc/chap6.pdf>

EXAMPLE 7.7 Determination of Sample Size Needed to Achieve a Stated Precision in Estimating a Population Mean, Using the Data of Example 7.3

If we specify that we wish to estimate μ with a 95% confidence interval no wider than 0.5 kg, then $d = 0.25$ kg, $1 - \alpha = 0.95$, and $\alpha = 0.05$. From Example 7.3 we have an estimate of the population variance: $s^2 = 0.4008 \text{ kg}^2$.

Let us guess that a sample of 40 is necessary; then,

$$t_{0.05(2), 39} = 2.023.$$

So we estimate (by Equation 7.7):

$$n = \frac{(0.4008)(2.023)^2}{(0.25)^2} = 26.2.$$

Next, we might estimate $n = 27$, for which $t_{0.05(2), 26} = 2.056$, and we calculate

$$n = \frac{(0.4008)(2.056)^2}{(0.25)^2} = 27.1.$$

Therefore, we conclude that a sample size greater than 27 is required to achieve the specified confidence interval.

R

CODE	<pre>#ex7.7 samplesize=function(x, d, n, alpha){ d=d n=n alpha=alpha v=n-1 tp=round(qt(alpha,v,lower.tail = F),3) n=round((var(x)*tp^2)/d^2,1) return(n) } samplesize(weight, 0.25, 40, 0.025) samplesize(weight, 0.25, 27, 0.025)</pre>

OUTPUT	<pre>> #ex7.7 > samplesize=function(x, d, n, alpha){ + d=d + n=n + alpha=alpha + v=n-1 + tp=round(qt(alpha,v,lower.tail = F),3) + n=round((var(x)*tp^2)/d^2,1) + return(n) + } > samplesize(weight, 0.25, 40, 0.025) [1] 26.2 > samplesize(weight, 0.25, 27, 0.025) [1] 27.1</pre>

SAS

CODE	<pre>proc iml; use ex7_3; read all; ss=var(wc2); d= 0.25;</pre>

	<pre>t1=-tinv(0.025, 39); n1= (ss*(t1**2))/d**2; t2= -tinv(0.025, 26); n2=(ss*(t2**2))/d**2; print n1 n2; run; quit;</pre>					
OUTPUT	<table><tr><th>n1</th><th>n2</th></tr><tr><td>26.238733</td><td>27.097624</td></tr></table>	n1	n2	26.238733	27.097624	
n1	n2					
26.238733	27.097624					
결과해석	어느정도 정밀한 신뢰구간을 얻기위해 필요한 최소 표본 수를 구하려면 위의 공식을 이용해 임의의 표본 수를 대입한 후에 나온 값이 대입한 표본 수보다 큰 값이 나올 때까지 구하면 필요한 최소 표본 수를 구할 수 있다.					

Example.7.8		
<p>EXAMPLE 7.8 Estimation of Required Sample Size to Test $H_0: \mu = \mu_0$</p> <p>How large a sample is needed to reject the null hypothesis of Example 7.2 when sampling from the population in that example? We wish to test at the 0.05 level of significance with a 90% chance of detecting a population mean different from $\mu_0 = 0$ by as little as 1.0 g. In Example 7.2, $s^2 = 1.5682 \text{ g}^2$.</p> <p>Let us guess that a sample size of 20 would be required. Then, $\nu = 19$, $t_{0.05(2), 19} = 2.093$, $\beta = 1 - 0.90 = 0.10$, $t_{0.10(1), 19} = 1.328$, and we use Equation 7.8 to calculate</p> $n = \frac{1.5682}{(1.0)^2} (2.093 + 1.328)^2 = 18.4.$ <p>We now use $n = 19$ as an estimate, in which case $\nu = 18$, $t_{0.05(2), 18} = 2.101$, $t_{0.10(1), 18} = 1.330$, and</p> $n = \frac{1.5682}{(1.0)^2} (2.101 + 1.330)^2 = 18.5.$ <p>Thus, we conclude that a new sample of at least 19 data may be taken from this population to test the above hypotheses with the specified α, β, and δ.</p>		
R		
CODE	<pre>#ex7.8 pwrsample=function(x, n, d, alpha, beta){ alpha=alpha beta=beta v=n-1 tp=round(qt(alpha/2,v,lower.tail = F),3) tp2=round(qt(beta, v, lower.tail = F),3) n=round((var(x)/d^2)*(tp+tp2)^2,1) return(n) } pwrsample(weight, 20, 1,0.05,0.1) pwrsample(weight, 19, 1,0.05,0.1) library(pwr) delta0=1/sqrt(1.5682)</pre>	

	<pre> pwr.t.test(n=,d=delta,sig.level = 0.05, power=0.9, type="one.sample", alternative = "two.sided") pwr.t.test(n=,d=delta,sig.level = 0.05, power=0.9, type="one.sample", alternative = "greater") </pre>
OUTPUT	<pre> > #ex7.2 > weight=c(1.7,0.7,-0.4,-1.8,0.2,0.9,-1.2,-0.9,-1.8,-1.4,-1.8,-2.0) > #ex7.8 > pwr.sample=function(x, n, d, alpha, beta){ + alpha=alpha + beta=beta + v=n-1 + tp=round(qt(alpha/2,v,lower.tail = F),3) + tp2=round(qt(beta, v, lower.tail = F),3) + n=round((var(x)/d^2)*(tp+tp2)^2,1) + return(n) + } > pwr.sample(weight, 20, 1,0.05,0.1) [1] 18.4 > pwr.sample(weight, 19, 1,0.05,0.1) [1] 18.5 > pwr.t.test(n=,d=delta0,sig.level = 0.05, power=0.9, type="one.sample", alternative = "two.sided") one-sample t test power calculation n = 18.50579 d = 0.7985448 sig.level = 0.05 power = 0.9 alternative = two.sided > pwr.t.test(n=,d=delta0,sig.level = 0.05, power=0.9, type="one.sample", alternative = "greater") one-sample t test power calculation n = 14.89189 d = 0.7985448 sig.level = 0.05 power = 0.9 alternative = greater </pre>
SAS	
CODE	<pre> proc iml; use ex7_2; read all; xba=mean(wc); vx=var(wc); d=1; v1= 19; beta=0.1; ta1=-tinv(0.025,v1); tb1=-tinv(0.1,v1); n1=round((vx/d**2)*(ta1+tb1)**2, .1); v2=18; ta2=-tinv(0.025,v2); tb2=-tinv(0.1,v2); n2=round((vx/d**2)*(ta2+tb2)**2, .1); print n1 n2; run; quit; </pre>

OUTPUT	<table border="1"> <tr> <th>n1</th><th>n2</th></tr> <tr> <td>18.4</td><td>18.5</td></tr> </table>	n1	n2	18.4	18.5
n1	n2				
18.4	18.5				
결과해석	ex7.2에서 12마리의 표본을 가지고는 쥐들의 몸무게에 변화가 있다고 보기 어려웠었다. 이에 귀무가설을 기각하기 위해 어느 정도의 표본이 필요한지를 알고 싶을 때 위의 공식을 이용하여 구할 수 있다. 여기서는 18.4가 나왔으므로 바로 위의 정수인 19마리가 귀무가설을 기각하는데 필요한 최소 표본 수이다.				

Example.7.9	
<p>EXAMPLE 7.9 Estimation of Minimum Detectable Difference in a One-Sample t Test for $H_0: \mu = \mu_0$</p> <p>In the two-tailed test of Example 7.2, what is the smallest difference (i.e., difference between μ and μ_0) that is detectable 90% of the time using a sample of 25 data and a significance level of 0.05?</p> <p>Using Equation 7.9:</p> $\delta = \sqrt{\frac{1.5682}{25}}(t_{0.05(2),24} + t_{0.10(1),24})$ $= (0.25)(2.064 + 1.318)$ $= 0.85 \text{ g.}$	
R	
CODE	<pre>#ex7.9 delta=function(x, n, alpha, beta){ n=n v=n-1 alpha=alpha beta=beta tp=round(qt(alpha/2,v,lower.tail = F),3) tp2=round(qt(beta, v, lower.tail = F),3) delta=round(sqrt(var(x)/n)*(tp+tp2),2) return(delta) } delta(weight,25,0.05,0.1)</pre>
OUTPUT	<pre>> #ex7.9 > delta=function(x, n, alpha, beta){ + n=n + v=n-1 + alpha=alpha + beta=beta + tp=round(qt(alpha/2,v,lower.tail = F),3) + tp2=round(qt(beta, v, lower.tail = F),3) + delta=round(sqrt(var(x)/n)*(tp+tp2),2) + return(delta) + } > delta(weight,25,0.05,0.1) [1] 0.85</pre>
SAS	
CODE	<pre>proc iml; use ex7_2; read all;</pre>

	<pre>xb=mean(wc); vx=var(wc); n=25; v=n-1; ta=-tinv(0.025,v); tb=-tinv(0.1,v); delta=round(sqrt(vx/n)*(ta+tb),.01); print delta; run; quit;</pre>		
OUTPUT	<table><tr><td>delta</td></tr><tr><td>0.85</td></tr></table>	delta	0.85
delta			
0.85			
결과해석	90%의 power로 귀무가설 하의 평균과 실제 평균의 가장 작은 차이는 위의 공식과 같이 구하며 ex7.2의 쥐들의 무게 변화의 평균의 Minimum Detectable Difference는 0.85g이라는 결과를 얻을 수 있다.		

Example.7.10	
<div> <p>EXAMPLE 7.10 Estimation of the Power of a One-Sample t Test for $H_0: \mu = \mu_0$</p> <p>What is the probability of detecting a true difference (i.e., a difference between μ and μ_0) of at least 1.0 g, using $\alpha = 0.05$ for the hypotheses of Example 7.2, if we run the experiment again using a sample of 15 from the same population?</p> <p>For $n = 15$, $v = 14$; $\alpha = 0.05$, $t_{0.05(2),14} = 2.145$, $s^2 = 1.5682 \text{ g}^2$, and $\delta = 1.0 \text{ g}$; and we use Equation 7.12 to find</p> $t_{\beta(1),14} = \frac{1.0}{\sqrt{\frac{1.5682 \text{ g}^2}{15}}} - 2.145$ $= 0.948.$ <p>Consulting Appendix Table B.3 tells us that, for $t_{\beta(1),14} = 0.948$, $0.10 < \beta < 0.25$, so we can say that the power would be $0.75 < 1 - \beta < 0.90$. Alternatively, by considering 0.948 to be a normal deviate and consulting Appendix Table B.2, we conclude that $\beta = 0.17$ and that the power of the test is $1 - \beta = 0.83$. (The exact probabilities, by computer, are $\beta = 0.18$ and power = 0.82.)</p> </div>	
R	
CODE	<pre> #ex7.10 beta=function(x, n, d, alpha){ n=n v=n-1 d=d alpha=alpha tp=round(qt(alpha/2, v, lower.tail = F),3) beta=round(d/(sqrt(var(x)/n))-tp, 3) return(beta) } beta(weight,15,1,0.05) </pre>

delta0가 Rcode

	<pre>pwr.t.test(n=15, d=delta0, sig.level = 0.05, power = , type = "one.sample", alternative = "two.sided")</pre>
OUTPUT	<pre>> #ex7.10 > beta=function(x, n, d, alpha){ + n=n + v=n-1 + d=d + alpha=alpha + tp=round(qt(alpha/2, v, lower.tail = F),3) + beta=round(d/(sqrt(var(x)/n))-tp, 3) + return(beta) + } > beta(weight,15,1,0.05) [1] 0.948 > pwr.t.test(n=15, d=delta0, sig.level = 0.05, power = , type = "one.sample", alternative = "two.sided") One-sample t test power calculation n = 15 d = 0.7985448 sig.level = 0.05 power = 0.8199455 alternative = two.sided</pre>

pwr.t.test(n=15, d=0.798545,
sig.level=0.05,
power=,
type="one.sample",
alternative="two.sided")

SAS			
CODE	<pre>proc iml; use ex7_2; read all; xba=mean(wc); vx=var(wc); n=15; v= n-1; d=1; ta=-tinv(0.025,v); tb=round(d/sqrt(vx/n)-ta, .001); print tb; run; quit;</pre>		
OUTPUT	<table><tr><td>tb</td></tr><tr><td>0.948</td></tr></table>	tb	0.948
tb			
0.948			
결과해석	<p>ex7.2 가설검정의 검정력을 추정하기 위해 위와 같은 공식을 이용해 t값을 구하고 그에 대응하는 확률 값을 찾으면 검정력은 0.17라는 값을 얻을 수 있다.</p>		

```
PROC POWER ;
onesamplemeans test=t
mean = 1
stddev = 1.252278
ntotal = 15
power=.;
RUN;
```

0.948 t .
power .
power 0.82 .

Example.7.

EXAMPLE 7.11 A One-Tailed Test for the Hypotheses $H_0: \sigma^2 \leq 1.5 \text{ sec}^2$ and $H_A: \sigma^2 > 1.5 \text{ sec}^2$, Using the Data of Example 7.4

$$\begin{aligned}
 SS &= 18.8288 \text{ sec}^2 \\
 \nu &= 7 \\
 s^2 &= 2.6898 \text{ sec}^2 \\
 \chi^2 &= \frac{SS}{\sigma_0^2} = \frac{18.8288 \text{ sec}^2}{1.5 \text{ sec}^2} = 12.553 \\
 \chi_{0.05, 7}^2 &= 14.067
 \end{aligned}$$

Since $12.553 < 14.067$, H_0 is not rejected.

$$0.05 < P < 0.10 \quad [P = 0.084]$$

We conclude that the variance of dissolving times is no more than 1.5 sec^2 .

R

CODE	<pre>#ex7.11 time=c(42.7,43.4,44.6,45.1,45.6,45.9,46.8,47.6) ss=function(x) round(sum(x^2)-(sum(x))^2/length(time),4) ss(time) sigma0=1.5 chi=ss(time)/sigma0 chip=qchisq(0.05, 7, lower.tail = F) chi<chip library("EnvStats") varTest(time, sigma.squared = sigma0, alternative = "greater")</pre>
OUTPUT	<pre>> #ex7.11 > time=c(42.7,43.4,44.6,45.1,45.6,45.9,46.8,47.6) > ss=function(x) round(sum(x^2)-(sum(x))^2/length(time),4) > ss(time) [1] 18.8288 > sigma0=1.5 > chi=ss(time)/sigma0 > chip=qchisq(0.05, 7, lower.tail = F) > chi<chip [1] TRUE > varTest(time, sigma.squared = sigma0, alternative = "greater") Chi-Squared Test on Variance data: time Chi-squared = 12.552, df = 7, p-value = 0.08379 alternative hypothesis: true variance is greater than 1.5 95 percent confidence interval: 1.338492 Inf sample estimates: variance 2.689821</pre>

SAS

CODE	<pre>proc iml; use ex7_4; read all; ss=18.8288; v0=1.5; v=7; vx=var(time); chi=ss/v0; chip=cinv(0.95, v); if chi < chip then answer="H0 is not rejected"; else answer="H0 is rejected";</pre>
------	--

	<pre>print answer; run; quit;</pre>
OUTPUT	<div style="border: 1px solid black; padding: 5px; margin-bottom: 5px; background-color: #e6f2ff;">answer</div> <div style="border: 1px solid black; padding: 5px;">H0 is not rejected</div>
결과해석	약이 위액에 녹는 시간에 대한 분산이 1.5보다 큰 지 알아보기 위해 위와 같은 단일 카이제곱 검정을 이용하여 가설검정을 실시하였더니 p-value가 0.084로 귀무가설을 기각할 충분한 근거가 되지 못한다. 따라서 약의 용해 시간에 대한 분산이 1.5보다 크다고 할 수 없다.

Example.7.12	
<p>EXAMPLE 7.12 Estimation of Required Sample Size to Test $H_0: \sigma^2 \leq \sigma_0^2$ versus $H_A: \sigma^2 > \sigma_0^2$</p> <p>How large a sample is needed to reject $H_0: \sigma^2 \leq 1.50 \text{ sec}^2$, using the data of Example 7.11, if we test at the 0.05 level of significance and with a power of 0.90? (Therefore, $\alpha = 0.05$ and $\beta = 0.10$.)</p> <p>From Example 7.11, $s^2 = 2.6898 \text{ sec}^2$. As we have specified $\sigma_0^2 = 1.75 \text{ sec}^2$, $\sigma_0^2/s^2 = 0.558$.</p> <p>To begin the iterative process of estimating n, let us guess that a sample size of 30 would be required. Then,</p> $\frac{\chi_{0.90, 29}^2}{\chi_{0.05, 29}^2} = \frac{19.768}{42.557} = 0.465.$ <p>Because $0.465 < 0.558$, our estimate of n is too low. So we might guess that $n = 50$ is required:</p> $\frac{\chi_{0.90, 49}^2}{\chi_{0.05, 49}^2} = \frac{36.818}{66.339} = 0.555.$ <p>Because 0.555 is a little less than 0.558, $n = 50$ is a little too low and we might guess $n = 55$, for which $\chi_{0.90, 54}^2/\chi_{0.05, 54}^2 = 41.183/70.153 = 0.571$.</p> <p>Because 0.571 is greater than 0.558, our estimate of n is high, so we could try $n = 51$, for which $\chi_{0.90, 50}^2/\chi_{0.05, 50}^2 = 37.689/67.505 = 0.558$.</p> <p>Therefore, we estimate that a sample size of at least 51 is required to perform the hypothesis test with the specified characteristics.</p>	
R	
CODE	<pre>#ex7.12 varsize=function(x,n,alpha,beta){ n=n v=n-1 alpha=alpha beta=beta chipa=round(qchisq(alpha, v, lower.tail = F),3) chipb=round(qchisq(beta, v, lower.tail = T),3) ratio=round(chipb/chipa,3) return(ratio) } sigma0=1.5 varsize(time, 30, 0.05, 0.1) sigma0/var(time)</pre>

	<pre>varsize(time, 50, 0.05, 0.1)</pre>
OUTPUT	<pre>> #ex7.12 > varsize=function(x,n,alpha,beta){ + n=n + v=n-1 + alpha=alpha + beta=beta + chipa=round(qchisq(alpha, v, lower.tail = F),3) + chipb=round(qchisq(beta, v, lower.tail = T),3) + ratio=round(chipb/chipa,3) + return(ratio) + } > sigma0=1.5 > varsize(time, 30, 0.05, 0.1) [1] 0.465 > sigma0/var(time) [1] 0.5576578 > varsize(time, 50, 0.05, 0.1) [1] 0.555</pre>
SAS	
CODE	<pre>%macro sample(n); proc iml; var1=2.6898; var2=1.5; alpha=0.05; beta=0.1; vx= var1; v0 =var2; vratio= v0/vx; chia=cinv(1-alpha, &n-1); chib=cinv(beta, &n-1); ratio= chib/chia; if vratio > ratio then ans="need more sample"; else ans="n is enough"; print ans; run;quit; %mend; %sample(30); %sample(50); %sample(51);</pre>
OUTPUT	<div><div>ans</div><div>need more sample</div></div> <div><div>ans</div><div>need more sample</div></div>

	<div>ans</div> <div>n is enough</div>
결과해석	ex7.11에서 시행한 일표본 분산에 대한 검정에 대해 귀무가설을 기각하기 위한 최소 표본 수를 구하기 위해서 위와 같이 카이제곱 값들의 비와 분산비를 비교해 얻을 수 있다. n=51일 때 처음으로 분산비보다 큰 값이 나오므로 귀무가설을 기각하는데 필요한 최소 표본 수는 51이다.

Example.7.13

EXAMPLE 7.13 Two-Tailed Nonparametric Testing of Symmetry Around the Median, Using the Data of Example 6.7 and the Wilcoxon Test of Section 9.5

H_0 : The population of data from which this sample came is distributed symmetrically around its median.

H_A : The population is not distributed symmetrically around its median.

$n = 70$; median = $X_{(70+1)/2} = X_{35.5} = 70.5$ in.

X (in.)	d (in.)	f	$ d $ (in.)	Rank of $ d $	Signed rank of $ d $	$(f)(\text{Signed rank})$
63	-7.5	2	7.5	69.5	-69.5	-139
64	-6.5	2	6.5	67.5	-67.5	-135
65	-5.5	3	5.5	64	-64	-192
66	-4.5	5	4.5	57.5	-57.5	-287.5
67	-3.5	4	3.5	48.5	-48.5	-194
68	-2.5	6	2.5	35.5	-35.5	-213
69	-1.5	5	1.5	21.5	-21.5	-107.5
70	-0.5	8	0.5	8	-8	-64
71	0.5	7	0.5	8	8	56
72	1.5	7	1.5	21.5	21.5	160.5
73	2.5	10	2.5	35.5	35.5	355
74	3.5	6	3.5	48.5	48.5	291
75	4.5	3	4.5	57.5	57.5	172.5
76	5.5	2	5.5	64	64	128
		70				

$$T_- = 1332$$

$$T_+ = 1163$$

$$T_{0.05(2),70} = 907 \text{ (from Appendix Table B.12)}$$

As neither T_- nor $T_+ < T_{0.05(2),70}$, do not reject H_0 . [$P > 0.50$]

R

#ex7.13

x=seq(63.,76)

frq=c(2,2,3,5,4,6,5,8,7,7,10,6,3,2)

ex=data.frame(x,frq)

ex[rep(1:nrow(ex), times=ex\$frq),1:2]

new=ex[rep(1:nrow(ex), times=ex\$frq),0:1]

new=as.numeric(new)

wilcox.test(new, mu = 70.5, paired = F, alternative = "two.sided", correct = T)

CODE

OUTPUT	<pre> > #ex7.13 > x=seq(63.,76) > frq=c(2,2,3,5,4,6,5,8,7,7,10,6,3,2) > ex=data.frame(x,frq) > ex[rep(1:nrow(ex), times=ex\$frq),1:2] x frq 1 63 2 1.1 63 2 2 64 2 2.1 64 2 3 65 3 3.1 65 3 3.2 65 3 4 66 5 4.1 66 5 4.2 66 5 4.3 66 5 4.4 66 5 5 67 4 5.1 67 4 5.2 67 4 5.3 67 4 6 68 6 6.1 68 6 6.2 68 6 6.3 68 6 6.4 68 6 . . . 12 74 6 12.1 74 6 12.2 74 6 12.3 74 6 12.4 74 6 12.5 74 6 13 75 3 13.1 75 3 13.2 75 3 14 76 2 14.1 76 2 > new=ex[rep(1:nrow(ex), times=ex\$frq),0:1] > new=as.numeric(new) > wilcox.test(new, mu = 70.5, paired = F, alternative = "two.sided", correct = T) Wilcoxon signed rank test with continuity correction data: new V = 1153, p-value = 0.6011 alternative hypothesis: true location is not equal to 70.5 </pre>
SAS	
CODE	<pre> data ex7_13; input height @@; cards; 63 63 64 64 65 65 65 66 66 66 66 66 67 67 67 68 68 68 68 68 68 69 69 69 69 69 </pre>

	<pre>70 70 70 70 70 70 70 70 70 71 71 71 71 71 71 71 72 72 72 72 72 72 72 73 73 73 73 73 73 73 73 73 73 74 74 74 74 74 74 75 75 75 76 76 ; run; proc univariate data= ex7_13 normal mu0=70.5; var height; run;</pre>																									
OUTPUT	<table><tr><th colspan="5">위치모수 검정: Mu0=70.5</th></tr><tr><th colspan="2">검정</th><th>통계량</th><th colspan="2">p 값</th></tr><tr><td>스튜던트의 t</td><td>t</td><td>-0.83054</td><td>Pr > t </td><td>0.4091</td></tr><tr><td>부호</td><td>M</td><td>0</td><td>Pr >= M </td><td>1.0000</td></tr><tr><td>부호 순위</td><td>S</td><td>-89.5</td><td>Pr >= S </td><td>0.6026</td></tr></table>	위치모수 검정: Mu0=70.5					검정		통계량	p 값		스튜던트의 t	t	-0.83054	Pr > t	0.4091	부호	M	0	Pr >= M	1.0000	부호 순위	S	-89.5	Pr >= S	0.6026
위치모수 검정: Mu0=70.5																										
검정		통계량	p 값																							
스튜던트의 t	t	-0.83054	Pr > t	0.4091																						
부호	M	0	Pr >= M	1.0000																						
부호 순위	S	-89.5	Pr >= S	0.6026																						
결과해석	<p>ex6.7의 자료를 통해 모집단이 중위수를 중심으로 대칭적으로 분포되어 있지 않다는 것을 확인하기 위해 윌콕슨 부호순위 검정을 한 결과 p-value가 약 0.6 정도로 귀무가설을 기각할 수 없다. 따라서 ex6.7의 자료는 모집단이 중위수를 중심으로 대칭적으로 분포되어 있지 않다고 할 수 없다.</p>																									