Example.10.1

EXAMPLE 10.1 A Single-Factor Analysis of Variance (Model I)

Nineteen pigs are assigned at random among four experimental groups. Each group is fed a different diet. The data are pig body weights, in kilograms, after being raised on these diets. We wish to ask whether pig weights are the same for all four diets.

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$.

H_A: The mean weights of pigs on the four diets are not all equal.

$\alpha = 0.05$

	Feed 1	*Feed 2*	*Feed 3*	*Feed 4*
	60.8	68.7	69.6	61.9
	67.0	67.7	77.1	64.2
	65.0	75.0	75.2	63.1
	68.6	73.3	71.5	66.7
	61.7	71.8		60.3
i	1	2	3	4
n_i	5	5	4	5
$\sum_{j=1}^{n_i} X_{ij}$	323.1	356.5	293.4	316.2
$\overline{X}_i$	64.62	71.30	73.35	63.24

EXAMPLE 10.1a Sums of Squares and Degrees of Freedom for the Data of Example 10.1.

	Feed 1	*Feed 2*	*Feed 3*	*Feed 4*
$\sum_{j=1}^{n_i} X_{ij}$	323.1	356.5	293.4	316.2
$\overline{X}_i$	64.62	71.30	73.35	63.24

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij} = 60.8 + 67.0 + 65.0 + \cdots + 63.1 + 66.7 + 60.3 = 1289.2$$

$$\overline{X} = \frac{1289.2}{19} = 67.8526$$

$$\text{Total SS} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij} - \overline{X})^2$$

$$= (60.8 - 67.8526)^2 + (67.0 - 67.8526)^2$$

$$+ \cdots + (66.7 - 67.8526)^2 + (60.3 - 6.8526)^2$$

$$= 49.7372 + 0.7269 + \cdots + 1.3285 + 57.0418 = 479.6874.$$

$$\text{total DF} = N - 1 = 19 - 1 = 18$$

$$\text{groups SS} = \sum_{i=1}^{k} n_i(\overline{X}_i - \overline{X})^2$$

$$= 5(64.62 - 67.8526)^2 + 5(71.30 - 67.8526)^2$$

$$+ 4(73.35 - 67.8526)^2 + 5(63.24 - 67.8526)^2$$

$$= 52.2485 + 59.4228 + 120.8856 + 106.3804 = 338.9372$$

$$\text{groups DF} = k - 1$$

$$\text{within-groups (error) SS} = \sum_{i=1}^{k}\left[\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2\right]$$

$$= (60.8 - 64.62)^2 + (67.0 - 64.62)^2$$

$$+ \cdots + (66.7 - 63.24)^2 + (60.3 - 63.24)^2$$

$$= 14.5924 + 5.6644 + \cdots + 11.9716 + 8.6436$$

$$= 140.7500$$

or, alternatively,

$$\text{within-groups (error) SS} = \text{Total SS} - \text{Groups SS}$$

$$= 479.6874 - 338.9373 = 140.7501.$$

$$\text{within-groups (error) DF} = \sum_{i=1}^{k}(n_i - 1)$$

$$= (5 - 1) + (5 - 1) + (4 - 1) + (5 - 1) = 15$$

or within-groups (error) DF $= N - k = 19 - 4 = 15$

or within-groups (error) DF = Total DF − Groups DF $= 18 - 3 = 15$.

Note: The quantities involved in the sum-of-squares calculations are carried to several decimal places (as computers typically do) to avoid rounding errors. All of these sums of squares (and the subsequent mean squares) have $(\text{kg})^2$ as units. However, for typographic convenience and ease in reading, the units for ANOVA computations are ordinarily not printed.

EXAMPLE 10.1b Sums of Squares and Degrees of Freedom for the Data of Example 10.1, Using Machine Formulas

	Feed 1	Feed 2	Feed 3	Feed 4	
i	1	2	3	4	
n_i	5	5	4	5	
$\sum_{j=1}^{n_i} X_{ij}$	323.1	356.5	293.4	316.2	
$\dfrac{\left(\sum_{j=1}^{n_i} X_{ij}\right)^2}{n_i}$	20878.7220	25418.4500	21520.8900	19996.4480	$\sum_{i=1}^{k} \dfrac{\left(\sum_{j=1}^{n_i} X_{ij}\right)}{n_i} = 87814.5500$

$\sum_i \sum_j X_{ij} = 1289.2$ total DF $= N - 1 = 19 - 1 = 18$

$\sum_i \sum_j X_{ij}^2 = 87955.30$ groups DF $= k - 1 = 4 - 1 = 3$

error DF $= N - k = 19 - 4 = 15$

$$C = \frac{\left(\sum_i \sum_j X_{ij}\right)^2}{N} = \frac{(1289.2)^2}{19} = 87475.6126$$

$$\text{total SS} = \sum_i \sum_j X_{ij}^2 - C = 87955.3000 - 87475.6126 = 479.6874$$

$$\text{groups SS} = \sum_{i=1}^{k} \frac{\left(\sum_{j=1}^{n_i}\right)^2}{n_i} - C = 87814.5500 - 87475.6126 = 338.9374$$

$$\text{error SS} = \text{total SS} - \text{groups SS} = 479.68747 - 338.9374 = 140.7500$$

EXAMPLE 10.1c The Conclusion of the ANOVA of Example 10.1, Using the Results of Either Example 10.1a or 10.1b

Summary of the Analysis of Variance

Source of variation	SS	DF	MS
Total	479.6874	18	
Groups	338.9374	3	112.9791
Error	140.7500	15	9.3833

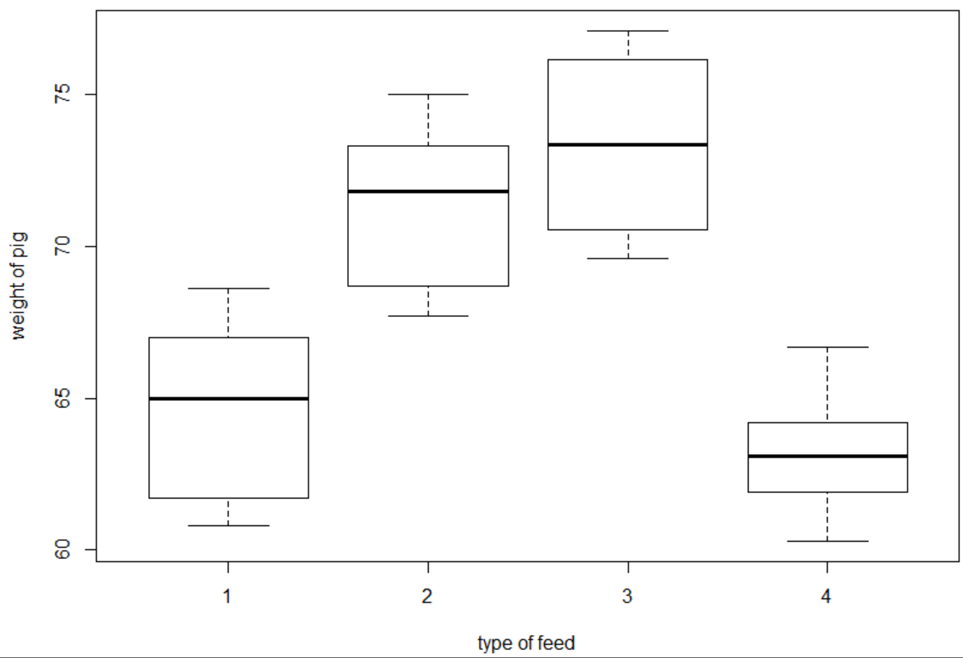
$$F = \frac{\text{groups MS}}{\text{error MS}} = \frac{112.9791}{9.3833} = 12.04$$

$F_{0.05(1),3,15} = 3.29$, so reject H_0.

$P < 0.0005$ [$P = 0.00029$]

R	
CODE	<pre>#ex 10.1 w1= c(60.8,67,65,68.6,61.7) w2= c(68.7,67.7,75,73.3,71.8) w3= c(69.6,77.1,75.2,71.5) w4= c(61.9,64.2,63.1,66.7,60.3) weight= c(w1, w2, w3, w4) n= c(length(w1), length(w2), length(w3), length(w4)) # ex10.1a sum1= sum(w1) sum2= sum(w2) sum3= sum(w3) sum4= sum(w4) sumx= c(sum(w1), sum(w2), sum(w3), sum(w4)) tot= sum(weight) mean1= mean(w1) mean2= mean(w2) mean3= mean(w3) mean4= mean(w4) meanx= c(mean(w1), mean(w2), mean(w3), mean(w4)) omean= mean(weight) sst= sum((weight-omean)^2) df= length(weight)-1 ssg= sum(n*(meanx-omean)^2) gdf= length(n)-1 sse= sum((w1-mean1)^2)+sum((w2-mean2)^2)+sum((w3-mean3)^2)+sum((w4-mean4)^2) edf= sum(n-1) # ex10.1b ss= sum(sumx^2/n) c= sum(tot)^2/length(weight) tss= sum(weight^2)-c gss= ss-c ess= tss-gss # ex10.1c msg= ssg/gdf mse= sse/edf f= msg/mse fp= qf(0.05, 3, 15, lower.tail = F) if(f>fp){</pre>

	`` answer= "reject H0" }else{ answer= "cannot reject H0" } pv= round(pf(f,3,15, lower.tail = F),5) print(c(answer, paste("p-value=", pv))) # method2 w1= c(60.8,67,65,68.6,61.7) w2= c(68.7,67.7,75,73.3,71.8) w3= c(69.6,77.1,75.2,71.5) w4= c(61.9,64.2,63.1,66.7,60.3) weight= c(w1, w2, w3, w4) n= c(length(w1), length(w2), length(w3), length(w4)) feed= rep(1:4, n) ex10.1= data.frame(weight,feed) ex10.1= transform(ex10.1, feed= factor(feed)) res=aov(weight~feed, data= ex10.1) residual= residuals(object=res) shapiro.test(residual) bartlett.test(weight~feed, data= ex10.1) summary(aov(weight~feed, data= ex10.1)) boxplot(weight~feed, xlab= "type of feed", ylab="weight of pig") ```
OUTPUT	``` > print(c(answer, paste("p-value=", pv))) [1] "reject H0" "p-value= 0.00028" > shapiro.test(residual) Shapiro-Wilk normality test data: residual W = 0.91042, p-value = 0.07534 > bartlett.test(weight~feed, data= ex10.1) Bartlett test of homogeneity of variances data: weight by feed Bartlett's K-squared = 0.47515, df = 3, p-value = 0.9243 > summary(aov(weight~feed, data= ex10.1)) Df Sum Sq Mean Sq F value Pr(>F) feed 3 338.9 112.98 12.04 0.000283 *** Residuals 15 140.8 9.38 --- ```



SAS

CODE

```
proc iml;
w1={60.8, 67, 65, 68.6, 61.7};
w2={68.7, 67.7, 75, 73.3, 71.8};
w3={69.6, 77.1, 75.2, 71.5};
w4={61.9, 64.2, 63.1, 66.7, 60.3};
weight= {60.8, 67, 65, 68.6, 61.7, 68.7, 67.7, 75, 73.3,
71.8, 69.6, 77.1, 75.2, 71.5, 61.9, 64.2, 63.1, 66.7, 60.3};
/************10.1a*************/
sum1= sum(w1);
sum2= sum(w2);
sum3= sum(w3);
sum4= sum(w4);
tot= sum(sum1, sum2, sum3, sum4);

mean1= mean(w1);
mean2= mean(w2);
mean3= mean(w3);
mean4= mean(w4);
ovmean= mean(weight);

sst= sum((weight-ovmean)#(weight-ovmean));
df= nrow(weight)-1;
ssg=
nrow(w1)*(mean1-ovmean)**2+nrow(w2)*(mean2-ovmean)**2+nrow(
```

```
w3)*(mean3-ovmean)**2+nrow(w4)*(mean4-ovmean)**2;
gdf= 3;
sse= sst-ssg;
edf= df-gdf;
/************10.1b************/
ss=   sum(sum1**2/nrow(w1),   sum2**2/nrow(w2),   sum3**2/
nrow(w3), sum4**2/nrow(w4));
c= tot**2/nrow(weight);
tss= sum(weight#weight)-c;
gss= ss-c;
ess= tss-ess;
/************10.1c************/
msg= ssg/gdf;
mse= sse/edf;
f= msg/mse;
fp=  finv(0.95,3,15);
if f > fp then answer= "reject H0";
else answer= "cannot reject H0";
pv= 1-cdf('F', f, 3, 15);
print answer pv;
run; quit;
/**********method2***********/
data ex10_1;
input feed$ weight @@;
cards;
F1 60.8 F1 67 F1 65 F1 68.6 F1 61.7
F2 68.7 F2 67.7 F2 75 F2 73.3 F2 71.8
F3 69.6 F3 77.1 F3 75.2 F3 71.5
F4 61.9 F4 64.2 F4 63.1 F4 66.7 F4 60.3
; run;
proc glm data= ex10_1 plot= diagnostics;
class feed;
model weight= feed;
means feed / hovtest=bartlett;
output out= res1 r= residual;
run;
proc univariate data= res1  normal;
var residual;
run;
proc anova data= ex10_1;
```

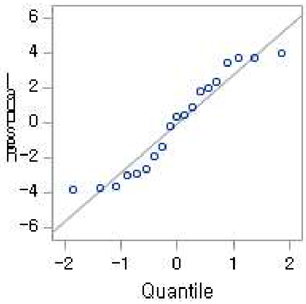
```
class feed;
model weight= feed;
run;
```

OUTPUT

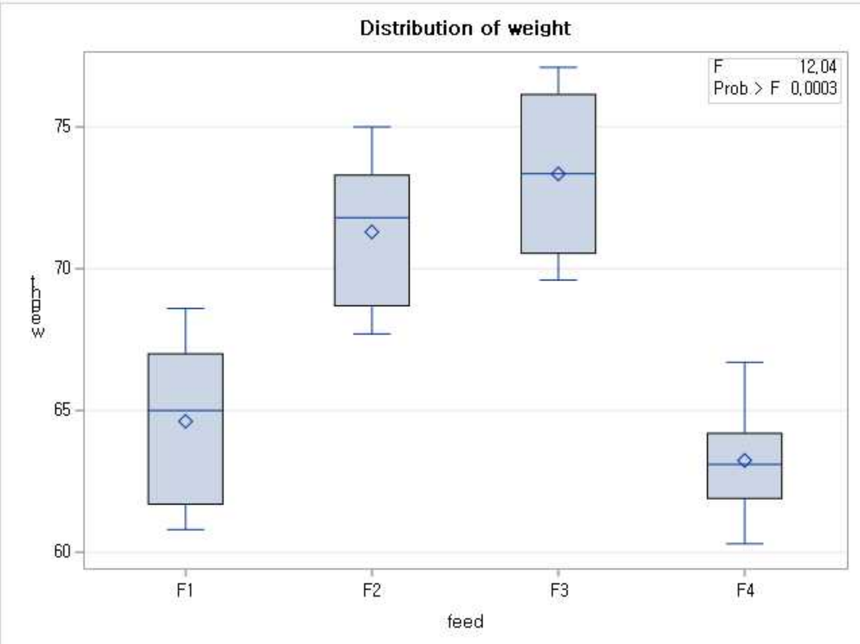
answer	pvalue
reject H0	0.000283

정규성 검정				
검정	통계량		p 값	
Shapiro-Wilk	W	0.910417	Pr < W	0.0753
Kolmogorov-Smirnov	D	0.13955	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.072413	Pr > W-Sq	0.2481
Anderson-Darling	A-Sq	0.5269	Pr > A-Sq	0.1610

Bartlett's Test for Homogeneity of weight Variance			
Source	DF	Chi-Square	Pr > ChiSq
feed	3	0.4752	0.9243



Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	338.9373684	112.9791228	12.04	0.0003
Error	15	140.7500000	9.3833333		
Corrected Total	18	479.6873684			



결과해석	네 종류의 먹이에 따른 돼지들의 평균 몸무게의 차이가 있는지를 알아보기 위해 19마리의 돼지를 네 종류의 먹이의 실험군에 무작위로 배정하였다. 이 때 먹이의 종류가 요인이 되고 먹이의 종류 1, 2, 3, 4가 먹이의 종류라는 요인의 각 처리가 된다. 각 처리수준에 따른 돼지들의 몸무게를 반응변수로 정하고 모든 먹이의 종류에 따라 돼지 몸무게의 모평균이 같은지 확인해보기 위해 정규성, 등분산성, 독립성을 만족한다는 가정(위의 정규성검정, 등분산검정결과, 랜덤화) 하에 일원분산분석(One-way ANOVA)를 실시했다. 귀무가설: 네 종류의 먹이에 따른 돼지들의 몸무게의 모평균은 모두 같다. 대립가설: 적어도 한 집단은 먹이의 종류에 따른 돼지들의 몸무게의 모평균이 다르다. 분산분석 결과 P-value= 0.00029로 매우 유의한 값이 나왔다. 따라서 귀무가설을 기각하여 적어도 네 종류의 먹이에 따라 돼지의 몸무게의 모평균에 미치는 영향은 모두 같지는 않다고 결론 내릴 수 있다. 결과의 상자그림을 보면 먹이 3과 먹이 4가 눈에 띄게 차이가 있어 보인다. 정확한 정보를 얻기 위해 다중비교를 통해 차이를 확인해 볼 수 있다.

Example.10.2

EXAMPLE 10.2 A Single-Factor Analysis of Variance for a Random-Effects Model (i.e., Model II) Experimental Design

A laboratory employs a technique for determining the phosphorus content of hay. The question arises: "Do phosphorus determinations differ among the technicians performing the analysis?" To answer this question, each of four randomly selected technicians was given five samples from the same batch of hay. The results of the 20 phosphorus determinations (in mg phosphorus/g of hay) are shown.

H_0: Determinations of phosphorus content do not differ among technicians.
H_A: Determinations of phosphorus content do differ among technicians.

$\alpha = 0.05$

	Technician 1	2	3	4
	34	37	34	36
	36	36	37	34
	34	35	35	37
	35	37	37	34
	34	37	36	35
Group sums:	173	182	179	176

$$\sum_i \sum_j X_{ij} = 710$$

$$\sum_i \sum_j X_{ij}^2 = 25234$$

$N = 20$

$$C = \frac{(710)^2}{20} = 25205.00$$

total SS $= 25234 - 25205.00 = 29.00$

$$\text{groups (i.e., technicians) SS} = \frac{(173)^2}{5} + \frac{a(182)^2}{5} + \frac{(179)^2}{5} + \frac{(176)^2}{5} - 25205.00$$

$$= 25214.00 - 25205.00 = 9.00$$

error SS $= 29.00 - 9.00 = 20.00$

Source of variation	SS	DF	MS
Total	29.00	19	
Groups (technicians)	9.00	3	3.00
Error	20.00	16	1.25

$$F = \frac{3.00}{1.25} = 2.40$$

$F_{0.05(1),3,16} = 3.24$

Do not reject H_0.

$0.10 < P < 0.25 \quad [P = 0.11]$

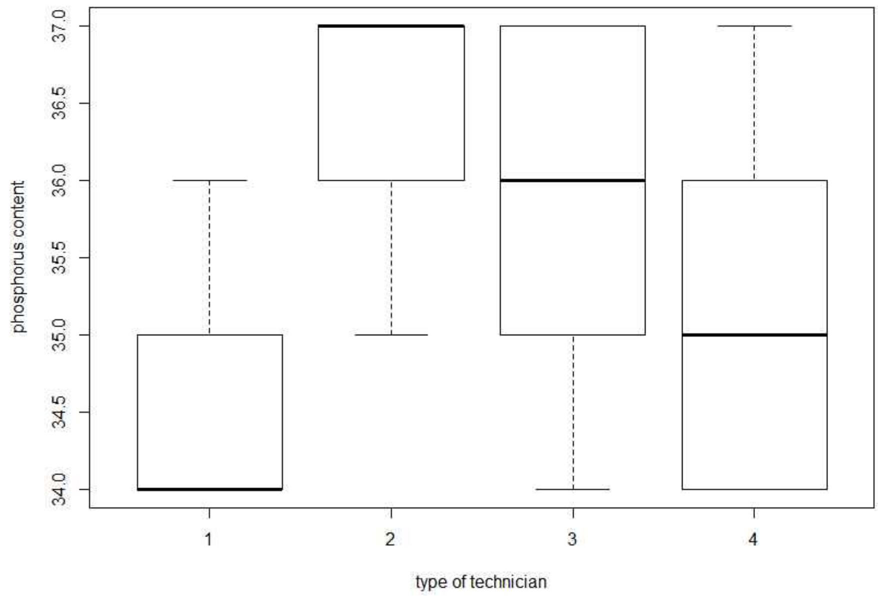
<table>
<tr><th colspan="2">R</th></tr>
<tr><td>CODE</td><td>

```
# ex10.2
p1= c(34,36,34,35,34)
p2= c(37,36,35,37,37)
p3= c(34,37,35,37,36)
p4= c(36,34,37,34,35)
phos= c(p1, p2, p3, p4)
n= c(length(p1), length(p2), length(p3), length(p4))
tech= rep(1:4, n)
ex10.2= data.frame(phos,tech)
ex10.2= transform(ex10.2, tech= factor(tech))

res=aov(phos~tech, data= ex10.2)
residual= residuals(object=res)
shapiro.test(residual)
bartlett.test(phos~tech, data= ex10.2)
summary(aov(phos~tech,data=ex10.2))
model = lm(phos ~tech,data=ex10.2)
anova(model)

boxplot(phos~tech, xlab = "type of technician", ylab = "phosphorus content")
```

</td></tr>
<tr><td>OUTPUT</td><td>

```
> shapiro.test(residual)

        Shapiro-Wilk normality test

data:  residual
W = 0.96599, p-value = 0.669

> bartlett.test(phos~tech, data= ex10.2)

        Bartlett test of homogeneity of variances

data:  phos by tech
Bartlett's K-squared = 1.0057, df = 3, p-value = 0.7999
> summary(aov(phos~tech,data=ex10.2))
            Df Sum Sq Mean Sq F value Pr(>F)
tech         3      9    3.00     2.4  0.106
Residuals   16     20    1.25
> anova(model)
Analysis of Variance Table

Response: phos
          Df Sum Sq Mean Sq F value Pr(>F)
tech       3      9    3.00     2.4 0.1059
Residuals 16     20    1.25
```

</td></tr>
</table>



SAS

CODE

```
data ex10_2;
input tech$ phos @@;
cards;
T1 34 T1 36 T1 34 T1 35 T1 34
T2 37 T2 36 T2 35 T2 37 T2 37
T3 34 T3 37 T3 35 T3 37 T3 36
T4 36 T4 34 T4 37 T4 34 T4 35
; run;
proc glm data= ex10_2 plot= diagnosti
class tech;
model phos= tech;
means tech / hovtest=bartlett;
output out= res2 r= residual;
run;
proc univariate data= res2  normal;
var residual;
run;
proc anova data= ex10_2;
class tech;
model phos= tech;
run;
```

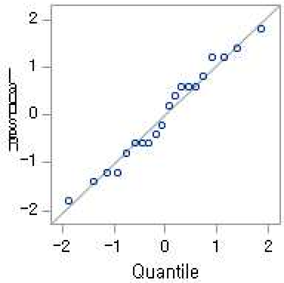
SAS와 R에서 random effect는 어떻게 지정하는지 알아두기 바랍니다.
https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_glm_details55.htm&docsetVersion=15.1&locale=en
https://stats.stackexchange.com/questions/228800/crossed-vs-nested-random-effects-how-do-they-differ-and-how-are-they-specified
http://lme4.r-forge.r-project.org/book/Ch2.pdf

OUTPUT

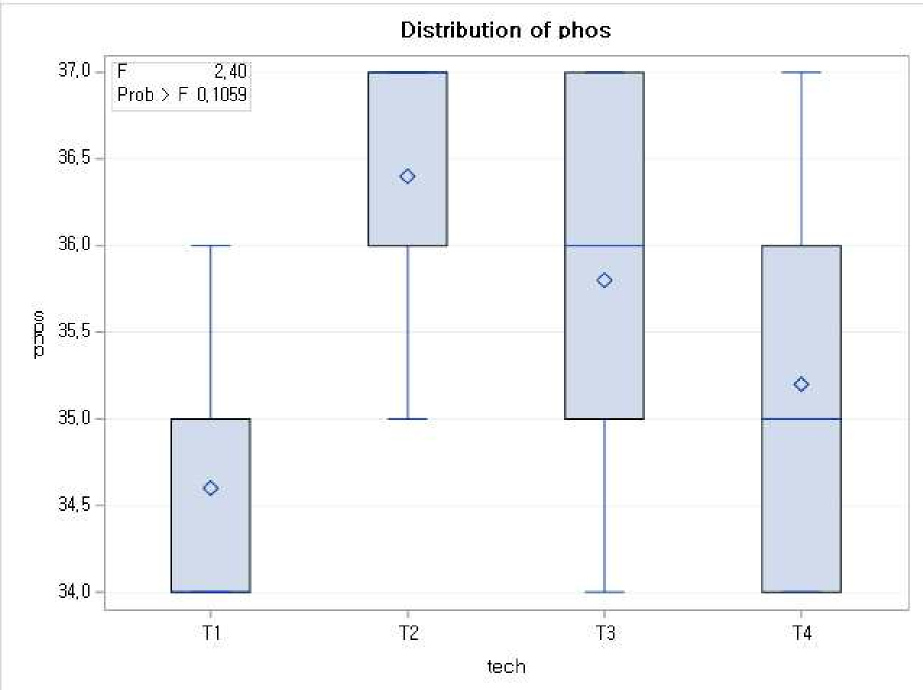
정규성 검정				
검정	통계량		p 값	
Shapiro-Wilk	W	0.965989	Pr < W	0.6690
Kolmogorov-Smirnov	D	0.120661	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.050536	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.288666	Pr > A-Sq	>0.2500

Bartlett's Test for Homogeneity of phos Variance

Source	DF	Chi-Square	Pr > ChiSq
tech	3	1.0057	0.7999



Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	9.00000000	3.00000000	2.40	0.1059
Error	16	20.00000000	1.25000000		
Corrected Total	19	29.00000000			



결과해석

연구실에서 건초의 인의 함유량을 확인해보기 위해 4명의 기술자들을 고용했다. 각각의 기술자들에 따라 측정하는 건초에 포함된 인의 평균 함유량에 차이가 있는지 확인해보려고 한다. 따라서 4명의 기술자들을 처리수준으로 정하고 각자의 기술자들에게 무작위로 건초를 5개씩 배정하여 건초에 포함된 인을 측정하게 한 후 그 결과를 가지고 분산분석을 실시하였다. 각 처리에 따라 실험단위를 랜덤하게 배정했고, 정규성검정, 등분산성검정 결과를 통해 정규성, 등분산성, 독립성 가정을 만족한다고 판단하였다.

random effect에 대한 귀무가설과 대립가설을 세우고 그에 따라서 해석해보기 바랍니다. 위에 적은 내용은 고정효과에 대한 해석입니다.

	귀무가설: 기술자들에 따라 측정하는 건초에 포함된 인의 함유량의 모평균이 모두 같다. 대립가설: 적어도 한 명의 기술자가 측정한 건초의 인의 함유량의 모평균은 다르다. 분석결과 p-value= 0.1059로 귀무가설을 기각할 충분한 증거를 얻지 못했다. 따라서 위의 분석결과로는 기술자들에 의해 측정된 건초의 인의 함유량의 모평균이 모두 다 동일하진 않다는 사실을 주장하기는 어렵다. 제시된 상자그림에서도 각 기술자들 별로 측정한 인의 함유량의 정도가 비슷해 보인다.

random effect에 대해서 귀무가설과 대립가설을 세우고 해석해야 됩니다.

Example.10.3

EXAMPLE 10.3 Welch's Test for an Analysis-of-Variance Experimental Design with Dissimilar Group Variances

The potassium content (mg of potassium per 100 mg of plant tissue) was measured in five seedlings of each of three varieties of wheat.

H_0: $\mu_1 = \mu_2 = \mu_3$.

H_A: The mean potassium content is not the same for seedlings of all three wheat varieties.

$\alpha = 0.05$

	Variety G	Variety A	Variety L	
	27.9	24.2	29.1	
	27.0	24.7	27.7	
	26.0	25.6	29.9	
	26.5	26.0	30.7	
	27.0	27.4	28.8	
	27.5	26.1	31.1	
i	1	2	3	
n_i	6	6	6	
ν_i	5	5	5	
$\overline{X}_i$	26.98	25.67	29.55	
s_i^2	0.4617	1.2787	1.6070	
$c_i = n_i/s_i^2$	12.9955	4.6923	3.7337	$C = \sum_i c_i = 21.4215$
$c_i\overline{X}_i$	350.6186	120.4513	110.3308	$\sum_i c_i\overline{X}_i = 581.4007$
$\dfrac{\left(1 - \dfrac{c_i}{C}\right)^2}{\nu_i}$	0.0309	0.1220	0.1364	$A = \sum_i \dfrac{\left(1 - \dfrac{c_i}{C}\right)^2}{\nu_i} = 0.2893$

$$\overline{X}_w = \frac{\sum_i c_i\overline{X}_i}{C} = \frac{581.4007}{21.4215} = 27.14$$

$$F' = \frac{\sum c_i(\overline{X}_i - \overline{X}_w)^2}{(k-1)\left[1 + \dfrac{2A(k-2)}{k^2-1}\right]}$$

$$= \frac{12.9955(26.98 - 27.14)^2 + 4.6923(25.67 - 27.14)^2 + 3.7337(29.55 - 27.14)}{(3-1)\left[1 + \dfrac{2(0.2893)(3-2)}{3^2-1}\right]}$$

$$= \frac{0.3327 + 10.1396 + 21.6857}{2(0.9268)} = \frac{32.4144}{1.8536} = 17.5$$

For critical value of F:

$$\nu_1 = k - 1 = 3 - 1 = 2$$

$$\nu_2 = \frac{k^2 - 1}{3A} = \frac{3^2 - 1}{3(0.2893)} = \frac{8}{0.8679} = 9.22$$

By harmonic interpolation in Appendix Table B.4 or by computer program:

$F_{0.05(1),2,9.22} = 4.22$. So, reject H_0.

$$0.0005 < P < 0.001 \quad [P = 0.0073]$$

<table>
<tr><th colspan="2">R</th></tr>
<tr><td>CODE</td><td>

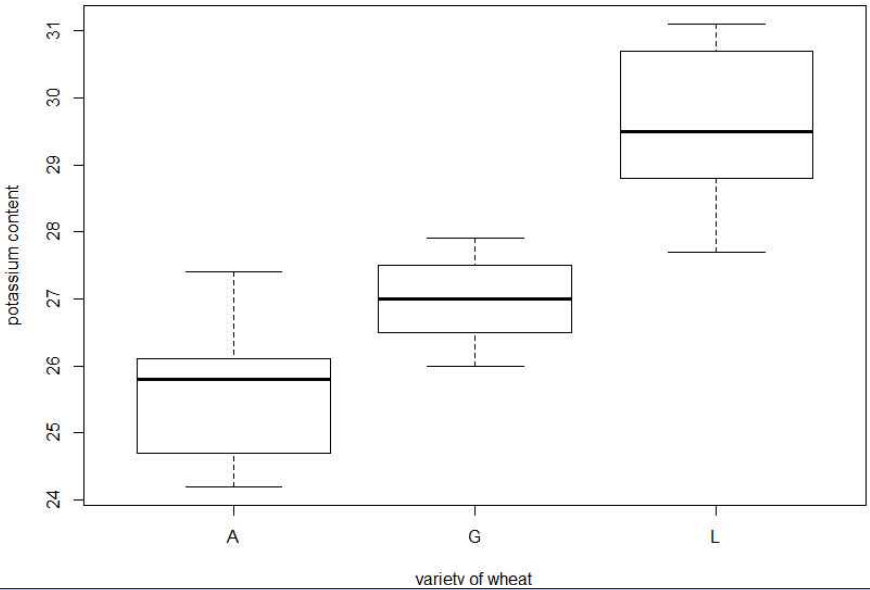
```
# ex10.3
library(onewaytests)
po1=c(27.9,27,26,26.5,27,27.5)
po2=c(24.2,24.7,25.6,26,27.4,26.1)
po3=c(29.1,27.7,29.9,30.7,28.8,31.1)
pot=c(po1, po2, po3)
wheat=rep(c("G","A","L"), c(6,6,6))
ex10.3= transform(data.frame(pot, wheat), wheat=factor(wheat))

model= lm(pot~wheat)
res= model$residuals

shapiro.test(res)
bartlett.test(pot~wheat,data= ex10.3)
welch.test(pot~wheat, data= ex10.3)
oneway.test(pot~wheat, data= ex10.3, var.equal = F)
oneway.test(pot~wheat, data= ex10.3, var.equal = T)

boxplot(pot~wheat, xlab = "variety of wheat", ylab = "potassium content")
```

</td></tr>
<tr><td>OUTPUT</td><td>

```
> shapiro.test(res)

        Shapiro-Wilk normality test

data:  res
W = 0.98293, p-value = 0.9751

> bartlett.test(pot~wheat,data= ex10.3)

        Bartlett test of homogeneity of variances

data:  pot by wheat
Bartlett's K-squared = 1.7512, df = 2, p-value = 0.4166
> welch.test(pot~wheat, data= ex10.3)

  Welch's Heteroscedastic F Test (alpha = 0.05)
-------------------------------------------------------------
  data : pot and wheat

  statistic  : 15.0096
  num df     : 2
  denom df   : 9.218254
  p.value    : 0.00126072

  Result     : Difference is statistically significant.
-------------------------------------------------------------
> oneway.test(pot~wheat, data= ex10.3, var.equal = F)

        One-way analysis of means (not assuming equal variances)

data:  pot and wheat
F = 15.01, num df = 2.0000, denom df = 9.2183, p-value = 0.001261
```

</td></tr>
</table>

```
> oneway.test(pot~wheat, data= ex10.3, var.equal = T)

        One-way analysis of means

data:  pot and wheat
F = 20.973, num df = 2, denom df = 15, p-value = 4.515e-05
```



SAS

CODE

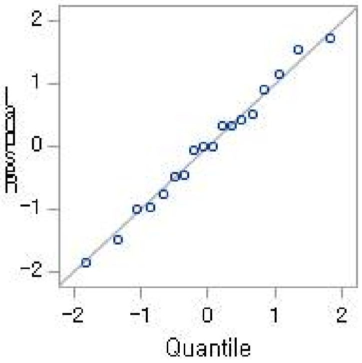
```
data ex10_3;
input wheat$ pota @@;
cards;
G 27.9 G 27 G 26 G 26.5 G 27 G 27.5
A 24.2 A 24.7 A 25.6 A 26 A 27.4 A 26.1
L 29.1 L 27.7 L 29.9 L 30.7 L 28.8 L 31.1
; run;
proc glm data= ex10_3 plot= diagnostics;
class wheat;
model pota= wheat;
means wheat / hovtest=bartlett;
output out= res3 r= residual;
run;
proc univariate data= res3  normal;
var residual;
run;
proc anova data= ex10_3;
class wheat;
model pota= wheat;
run;
```

OUTPUT

정규성 검정				
검정	통계량		p 값	
Shapiro-Wilk	W	0.965989	Pr < W	0.6690
Kolmogorov-Smirnov	D	0.120661	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.050536	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.288666	Pr > A-Sq	>0.2500

Bartlett's Test for Homogeneity of pota Variance			
Source	DF	Chi-Square	Pr > ChiSq
wheat	2	1.7512	0.4166

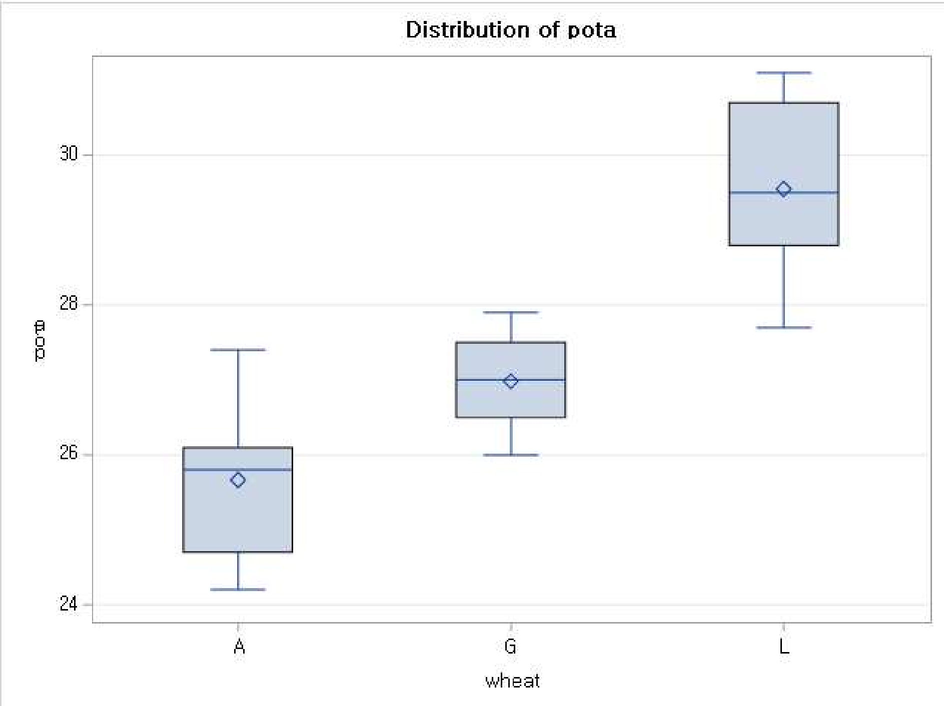
Welch's ANOVA for pota			
Source	DF	F Value	Pr > F
wheat	2.0000	15.01	0.0013
Error	9.2183		



- Welch's ANOVA 결과

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	46.80333333	23.40166667	20.97	<.0001
Error	15	16.73666667	1.11577778		
Corrected Total	17	63.54000000			

- 일반 ANOVA 결과



결과해석	3가지의 다른 종류의 밀의 묘목에 포함된 포타슘(칼륨)의 평균 함유량에 차이가 있는지 확인해보려고 한다. 따라서 3종류의 밀을 처리수준으로 정하고 각각 6그루의 묘목을 뽑아 포타슘을 측정한 후 그 결과를 가지고 분산분석을 실시하였다. 귀무가설: 3가지 종류의 밀의 묘목에 포함된 포타슘 함유량의 모평균은 모두 같다. 대립가설: 3가지 종류의 밀의 묘목에 포함된 포타슘 함유량의 모평균이 모두 같은 것은 아니다. 여기서 분산이 같지 않다고 가정을 하고 등분산성을 만족하지 못할 때 사용하는 분산분석방법인 Welch의 ANOVA를 실행해본 결과 p-value=0.0013으로 유의한 결과를 얻었다, 따라서 귀무가설을 기각하고 적어도 한 종류의 밀의 묘목에 포함된 포타슘 함유량의 모평균은 다른 종류의 밀들의 묘목에 포함된 포타슘 함유량의 모평균과 다르다고 결론을 내릴 수 있다. 그러나 실제로 등분산성검정을 해보면 위의 등분산검정 결과처럼 p-value= 0.4166으로 각 집단 간 분산의 차이가 없다는 결과를 확인할 수 있다. 따라서 일반 분산분석도 해봤지만 p-value가 매우 유의한 값이 나와 Welch's ANOVA와 같은 결론을 내릴 수 있다.