

실험계획

Final Project

통계학전공 5526220 김성현

통계학전공 5638852 김유진

통계학전공 5639190 최지우

CONTENTS

01 실험계획의 세 가지 원리

02 반복이 있는 이원배치법 (가법 모형, 혼합 모형)

03 분할구 설계(Split - Plot Design)

04 최종 모형 선정

01

실험계획의 세 가지 원리

- 01. 반복
- 02. 랜덤
- 03. 블록

01 반복

02 랜덤

02 블록

반복

- 세 가지의 방법과 네 개의 소성 온도를 이용하여 두 가지 변수(방법, 온도)들이 리튬이온 배터리 수명에 어떤 영향을 끼치는지 알아보고자 함
- 요인 설계의 각 반복 수는 12개의 관측값으로, 실험자는 3번의 반복을 시행하고자 하는 것

랜덤

- 배터리를 만들기 전에 재료를 세 가지 방법 중 임의로 한 가지 방법을 선택하여 생산하고, 이후 남은 두 가지 방법 중 임의로 한 가지 방법을 선택하여 두 번째 생산을 함
- 실험에서 사용할 방법 순서를 랜덤배치하는 것

블록

- 배터리 수명의 차이가 방법과 온도의 차이인지에 대해서 알고자 함
- 외생변수라고 생각되는 반복을 블록하여 반복이 실험에 미칠 수 있는 영향을 상쇄

02

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

- 01. 선정 이유
- 02. 모형식
- 03. 분산분석표 및 해석
- 04. 잔차분석
- 05. 다중비교
- 06. 결과해석

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

01 선정 이유

02. 모형식

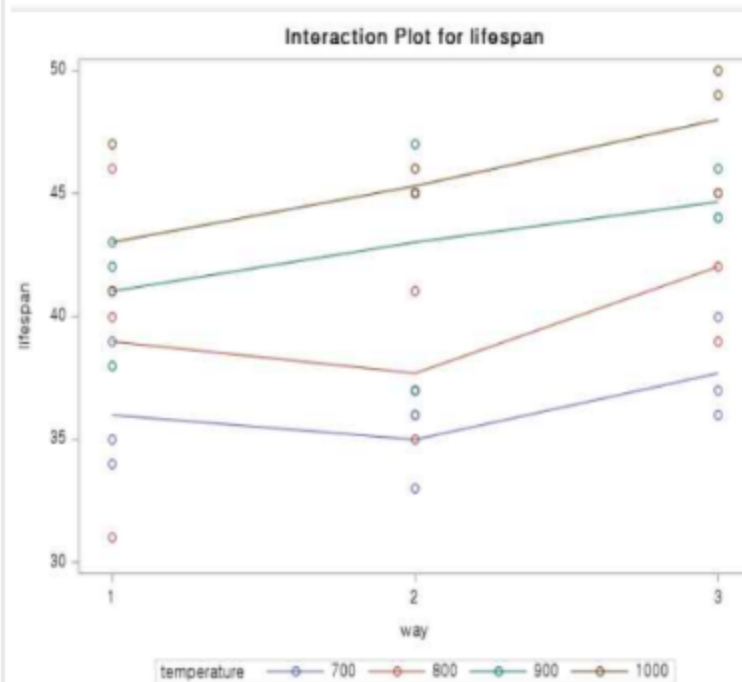
03. 분산분석표 및 해석

04. 잔차분석

05. 다중비교

06. 결과해석

선정 이유



Source	d.f	S.S	M.S.	F_0	P-value
방법	2	77.556	38.778	11.26	0.0093
온도	3	434.083	144.694	42.01	0.0002
방법*온도	6	20.667	3.444	0.28	0.9387

- 표면적으로만 보면 단순히 두 개의 요인에 반복이 있는 모형이라고 볼 수 있다고 생각
- 상호작용항에 대한 유의성 검정 결과, 유의확률이 0.9387로 상호작용항이 유의하지 않음
- 반복이 있는 이원배치법 중에서도 상호작용항이 없는 가법 모형을 첫 번째 모형으로 선택

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

01. 선정 이유

02 모형식

03. 분산분석표 및 해석

04. 잔차분석

05. 다중비교

06. 결과해석

모형식

$$y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk} \quad (i = 1, 2, 3, \quad j = 1, 2, 3, 4, \quad k = 1, 2, 3(\text{반복 수}))$$

μ : 배터리 수명의 모평균

α_i : 방법요인이 주는 효과

β_j : 온도요인이 주는 효과, ϵ_{ijk} : 방법 요인과 온도요인에서 얻은 k번째 측정값의 오차

$$\sum_{i=1}^3 \alpha_i = 0, \quad \beta_j \sim i.i.d. N(0, \sigma_\beta^2), \quad \epsilon_{ijk} \sim i.i.d. N(0, \sigma^2)$$

고정효과와 임의효과가 공존하는 혼합모형

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

01. 선정 이유

02. 모형식

03 분산분석표 및 해석

04. 잔차분석

05. 다중비교

06. 결과해석

분산분석표 및 해석

Source	d.f	S.S	M.S.	F_0	P-value
방법	2	77.556	38.778	3.74	0.0356
온도	3	434.083	144.694	13.94	<.0001
Error	30	311.333	10.378		
Total	35	822.972			

R-Square

0.622

1) 방법 요인에 대한 유의성 검정 가설

$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$, $H_1 : \text{적어도 하나의 } \alpha_i \text{ 는 } 0 \text{이 아니다. } (1 \leq i \leq 3)$

2) 온도요인에 대한 유의성 검정 가설

$H_0 : \sigma_{\beta}^2 = 0$, $H_1 : \sigma_{\beta}^2 > 0$

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

정규성

01. 선정 이유

02. 모형식

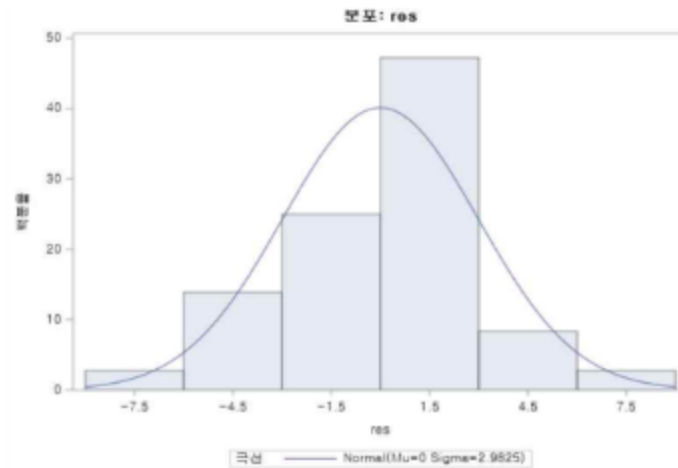
03. 분산분석표 및 해석

04 잔차 분석

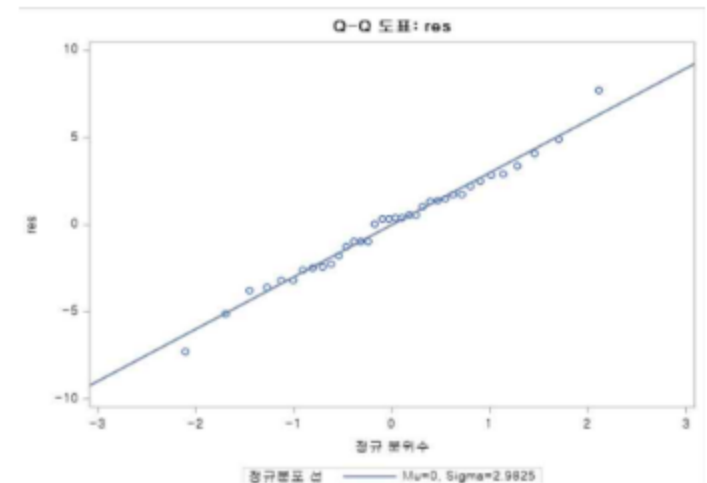
05. 다중 비교

06. 결과 해석

<histogram>



<qqplot>



<정규성 검정>

정규성 검정				
검정	통계량		p 값	
Shapiro-wilk	W	0.989	Pr < W	0.9707
Anderson-Darling	A-Sq	0.215	Pr > A-Sq	>0.2500

H_0 : 잔차가 정규분포를 따른다. H_1 : 잔차가 정규분포를 따르지 않는다.

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04 잔차 분석

05. 다중 비교

06. 결과 해석

등분산성

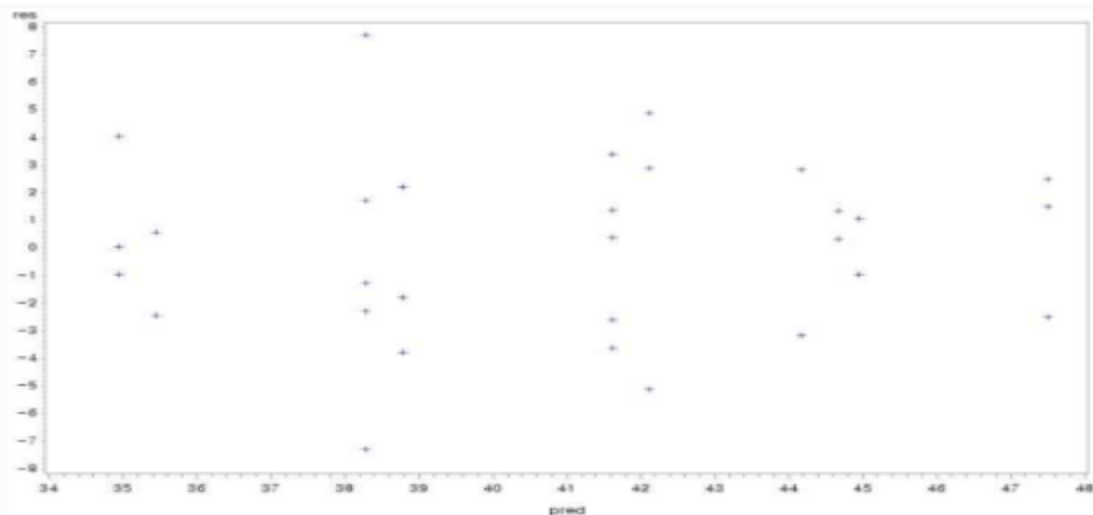
Levene's Test for Homogeneity of lifespan Variance					
ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
trt	11	3682.2	334.7	2.67	0.0216
Error	24	3014.5	125.6		

Brown and Forsythe's Test for Homogeneity of lifespan Variance					
ANOVA of Absolute Deviations from Group Medians					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
trt	11	50.8889	4.6263	0.72	0.7109
Error	24	154.7	6.4444		

Bartlett's Test for Homogeneity of lifespan Variance			
Source	DF	Chi-Square	Pr > ChiSq
trt	11	13.6321	0.2540

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{12}^2, \quad H_1 : \text{최소한 한 개의 } (i, j) \text{에 대하여 } \sigma_i^2 \neq \sigma_j^2, \quad (1 \leq i \neq j \leq 12)$$

독립성



반복이 있는 이원배치법 (가법 모형, 혼합 모형)

01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04. 잔차분석

05 다중 비교

06. 결과 해석

방법요인에 대한 다중 비교

<최소유의차검정> $i = 1, 2, 3$, $j = 1, 2, 3$

$LSD_{(i,j)}$	방법1	방법2	방법3
2.686	39.75	40.25	43.083

1) 방법1과 방법2의 경우	2) 방법2와 방법3의 경우	3) 방법1과 방법3의 경우
$ 39.75 - 40.25 = 0.5 < 2.686$	$ 40.25 - 43.083 = 2.833 > 2.686$	$ 39.75 - 43.083 = 3.333 > 2.686$

<본페로니 수정> $i = 1, 2, 3$, $j = 1, 2, 3$

MSD	방법1	방법2	방법3
3.335	39.75	40.25	43.083

1) 방법1과 방법2의 경우	2) 방법2와 방법3의 경우	3) 방법1과 방법3의 경우
$ 39.75 - 40.25 = 0.5 < 3.335$	$ 40.25 - 43.083 = 2.833 < 3.335$	$ 39.75 - 43.083 = 3.333 < 3.335$

<튜키의 검정> $i = 1, 2, 3$, $j = 1, 2, 3$, $i \neq j$

H_0 : 방법 i 와 방법 j 에 따른 배터리 수명이 같다. H_1 : 방법 i 와 방법 j 에 따른 배터리 수명이 다르다.

way	lifespan LSMEAN	LSMEAN Number
1	39,7500000	1
2	40,2500000	2
3	43,0833333	3

Least Squares Means for effect way Pr > t for H0: LSMean(i)=LSMean(j)			
Dependent Variable: lifespan			
i/j	1	2	3
1		0,9236	0,0429
2	0,9236		0,0959
3	0,0429	0,0959	

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04. 잔차분석

05 다중 비교

06. 결과 해석

온도요인에 대한 다중 비교

<최소유의차검정> $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4$

$LSD_{(i,j)}$	700°C	800°C	900°C	1000°C
3.101	36.222	39.556	42.889	45.444

1) 700°C와 800°C의 경우	2) 700°C와 900°C의 경우	3) 700°C와 1000°C의 경우
$ 36.222 - 39.556 = 3.334 > 3.101$	$ 36.222 - 42.889 = 6.667 > 3.101$	$ 36.222 - 45.444 = 9.222 > 3.101$
4) 800°C와 900°C의 경우	5) 800°C와 1000°C의 경우	6) 900°C와 1000°C의 경우
$ 39.556 - 42.889 = 3.333 > 3.101$	$ 39.556 - 45.444 = 5.888 > 3.101$	$ 42.889 - 45.444 = 2.555 < 3.101$

<본페로니 수정> $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4$

MSD	700°C	800°C	900°C	1000°C
4.290	36.222	39.556	42.889	45.444

1) 700°C와 800°C의 경우	2) 700°C와 900°C의 경우	3) 700°C와 1000°C의 경우
$ 36.222 - 39.556 = 3.334 < 4.290$	$ 36.222 - 42.889 = 6.667 > 4.290$	$ 36.222 - 45.444 = 9.222 > 4.290$
4) 800°C와 900°C의 경우	5) 800°C와 1000°C의 경우	6) 900°C와 1000°C의 경우
$ 39.556 - 42.889 = 3.333 < 4.290$	$ 39.556 - 45.444 = 5.888 > 4.290$	$ 42.889 - 45.444 = 2.555 < 4.290$

<튜키의 검정> $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4$, $i \neq j$

H_0 : 온도*i*와 온도*j*에 따른 배터리 수명이 같다. H_1 : 온도*i*와 온도*j*에 따른 배터리 수명이 다르다.

temperature	lifespan LSMEAN	LSMEAN Number
700	36,2222222	1
800	39,5555556	2
900	42,8888889	3
1000	45,4444444	4

Least Squares Means for effect temperature Pr > t for H0: LSMean(i)=LSMean(j)				
Dependent Variable: lifespan				
i/j	1	2	3	4
1		0,1478	0,0007	<,0001
2	0,1478		0,1478	0,0028
3	0,0007	0,1478		0,3501
4	<,0001	0,0028	0,3501	

반복이 있는 이원배치법 (가법 모형, 혼합 모형)

01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04. 잔차분석

05. 다중 비교

06 결과 해석

결과 해석

- 잔차는 정규성, 등분산성, 독립성을 모두 만족
- 방법요인에 대한 다중비교에서 최소유의차검정과 튜키의 검정의 결과를 수용
- 온도요인에 대한 다중비교의 모든 검정 결과에서 온도 간의 유의한 차이가 발견됨
- 반복이 있는 이원배치법(가법 모형, 혼합 모형) 모형의 경우 유의한 모형이라고 할 수 있음

03

분할구 설계 (Split-Plot Design)

- 01. 선정 이유
- 02. 모형식
- 03. 분산분석표 및 해석
- 04. 잔차분석
- 05. 다중비교
- 06. 결과해석

01 선정 이유

02 모형식

03. 분산분석표 및 해석

04. 잔차분석

05. 다중비교

06. 결과해석

선정 이유

- 리튬이온 배터리를 만드는 과정의 반복에 따라 리튬이온 배터리 수명이 달라진다는 점을 생각함
- 첫 번째 실험과 세 번째 실험의 배터리 수명 차이가 온전히 방법과 온도의 차이 때문인지 알 수 없게 됨
- 반복을 블록으로 하여 이 영향을 상쇄시킬 수 있는 분할구 설계를 선정

모형식

$$y_{ijk} = \mu + \rho_i + \tau_j + (\rho\tau)_{ij} + \gamma_k + \epsilon_{ijk} \quad (i = 1, 2, 3, j = 1, 2, 3, k = 1, 2, 3, 4)$$

$$\sum_{i=1}^3 \rho_i = 0, \sum_{j=1}^3 \tau_j = 0, \sum_{k=1}^4 \gamma_k = 0, \sum_{i=1}^3 (\rho\tau)_{ij} = \sum_{j=1}^3 (\rho\tau)_{ij} = 0, \epsilon_{ijk} \stackrel{iid}{\sim} N(0, \sigma^2)$$

ρ_i : 반복 (*block*) , τ_j : 방법 요인이 주는 효과 , γ_k : 온도요인이 주는 효과 , $(\rho\tau)_{ij}$: 주구오차

상호작용항은 유의하지 않으므로 상호작용 풀링을 통해 세구오차에 이를 포함시킴

분할구 설계
(Split - Plot Design)

01. 선정 이유

02. 모형식

03 분산분석표 및 해석

04. 잔차분석

05. 다중비교

06. 결과해석

분산분석표 및 해석

Source	DF	SS	MS	F_0	p-value
반복	2	128.389	64.194	4.28	0.1016
방법	2	77.556	38.778		
주구오차	4	36.278	9.069		
온도	3	434.083	144.694	23.68	<0.0001
세구오차	24	146.667	6.111		
Total	35	822.972			

R-Square

0.822

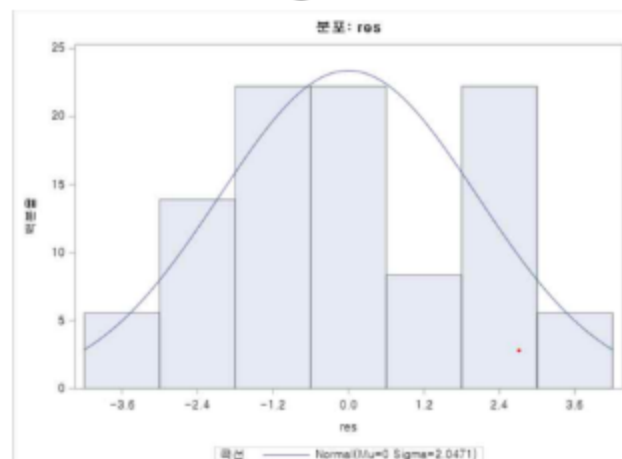
- 반복은 블록이므로 F검정은 실시하지 않음
- 방법요인에 대한 유의확률이 0.1016으로 0.05보다 크게 나옴 -> 다중비교에서 다시 언급
- 온도요인에 대한 효과는 유의함
- 결정계수는 0.822로 꽤 높음을 확인

분할구 설계 (Split - Plot Design)

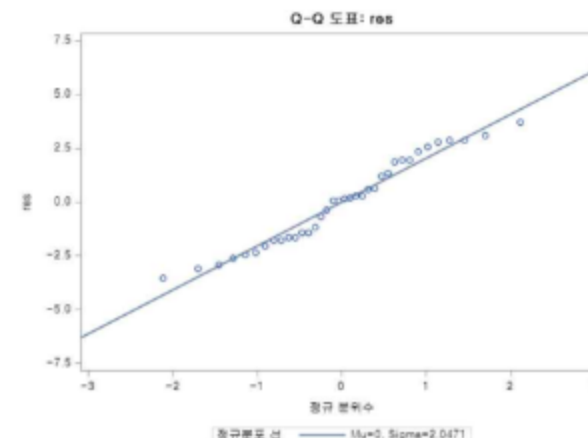
- 01. 선정 이유
- 02. 모형식
- 03. 분산분석표 및 해석
- 04 잔차 분석**
- 05. 다중 비교
- 06. 결과 해석

정규성

<histogram>



<qqplot>



<정규성 검정>

정규성 검정				
검정	통계량		p-value	
Shapiro-wilk	W	0.965	Pr < W	0.1654
Anderson-Darling	A-Sq	0.503	Pr > A-Sq	0.2018

H_0 : 잔차가 정규분포를 따른다. H_1 : 잔차가 정규분포를 따르지 않는다.

분할구 설계 (Split - Plot Design)

01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04 잔차 분석

05. 다중 비교

06. 결과 해석

등분산성

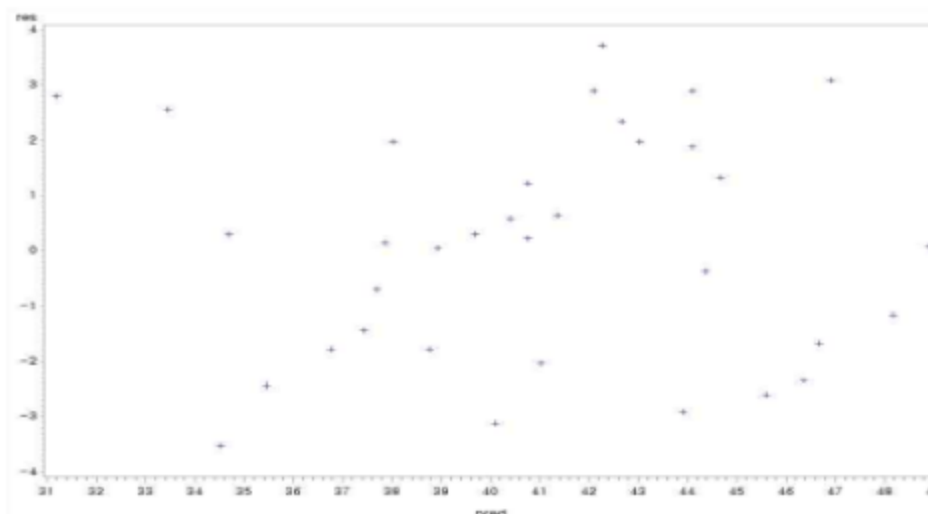
Levene's Test for Homogeneity of lifespan Variance					
ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
trt	11	248.7	22.61	1.57	0.1704
Error	24	344.9	14.369		

Brown and Forsythe's Test for Homogeneity of lifespan Variance					
ANOVA of Absolute Deviations from Group Medians					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
trt	11	9.201	0.837	0.25	0.9895
Error	24	79.833	3.326		

Bartlett's Test for Homogeneity of lifespan Variance			
Source	DF	Chi-Square	Pr > ChiSq
trt	11	5.8719	0.8818

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{12}^2, \quad H_1 : \text{최소한 한 개의 } (i, j) \text{에 대하여 } \sigma_i^2 \neq \sigma_j^2, \quad (1 \leq i \neq j \leq 12)$$

독립성



01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04. 잔차분석

05 다중 비교

06. 결과 해석

방법요인에 대한 다중 비교

<최소유의차검정> $i = 1, 2, 3$, $j = 1, 2, 3$

$LSD_{(i,j)}$	방법1	방법2	방법3
2.083	39.75	40.25	43.083

1) 방법1과 방법2의 경우	2) 방법2와 방법3의 경우	3) 방법1과 방법3의 경우
$ 39.75 - 40.25 = 0.5 < 2.083$	$ 40.25 - 43.083 = 2.833 > 2.083$	$ 39.75 - 43.083 = 3.333 > 2.083$

<본페로니 수정> $i = 1, 2, 3$, $j = 1, 2, 3$

MSD	방법1	방법2	방법3
2.597	39.75	40.25	43.083

1) 방법1과 방법2의 경우	2) 방법2와 방법3의 경우	3) 방법1과 방법3의 경우
$ 39.75 - 40.25 = 0.5 < 2.597$	$ 40.25 - 43.083 = 2.833 > 2.597$	$ 39.75 - 43.083 = 3.333 > 2.597$

<튜키의 검정> $i = 1, 2, 3$, $j = 1, 2, 3$, $i \neq j$

H_0 : 방법 i 와 방법 j 에 따른 배터리 수명이 같다. H_1 : 방법 i 와 방법 j 에 따른 배터리 수명이 다르다.

way	lifespan LSMEAN	LSMEAN Number
1	39,7500000	1
2	40,2500000	2
3	43,0833333	3

Least Squares Means for effect way Pr > t for H0: LSMean(i)=LSMean(j)			
Dependent Variable: lifespan			
i/j	1	2	3
1		0,8742	0,0081
2	0,8742		0,0255
3	0,0081	0,0255	

01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04. 잔차분석

05 다중 비교

06. 결과 해석

온도요인에 대한 다중 비교

<최소유의차검정> $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4$

$LSD_{(i,j)}$	700°C	800°C	900°C	1000°C
2.405	36.222	39.556	42.889	45.444

1) 700°C와 800°C의 경우	2) 700°C와 900°C의 경우	3) 700°C와 1000°C의 경우
$ 36.222 - 39.556 = 3.334 > 2.405$	$ 36.222 - 42.889 = 6.667 > 2.405$	$ 36.222 - 45.444 = 9.222 > 2.405$
4) 800°C와 900°C의 경우	5) 800°C와 1000°C의 경우	6) 900°C와 1000°C의 경우
$ 39.556 - 42.889 = 3.333 > 2.405$	$ 39.556 - 45.444 = 5.888 > 2.405$	$ 42.889 - 45.444 = 2.555 > 2.405$

<본페로니 수정> $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4$

MSD	700°C	800°C	900°C	1000°C
3.351	36.222	39.556	42.889	45.444

1) 700°C와 800°C의 경우	2) 700°C와 900°C의 경우	3) 700°C와 1000°C의 경우
$ 36.222 - 39.556 = 3.334 < 3.351$	$ 36.222 - 42.889 = 6.667 > 3.351$	$ 36.222 - 45.444 = 9.222 > 3.351$
4) 800°C와 900°C의 경우	5) 800°C와 1000°C의 경우	6) 900°C와 1000°C의 경우
$ 39.556 - 42.889 = 3.333 < 3.351$	$ 39.556 - 45.444 = 5.888 > 3.351$	$ 42.889 - 45.444 = 2.555 < 3.351$

<튜키의 검정> $i = 1, 2, 3, 4$, $j = 1, 2, 3, 4$, $i \neq j$

H_0 : 온도*i*와 온도*j*에 따른 배터리 수명이 같다. H_1 : 온도*i*와 온도*j*에 따른 배터리 수명이 다르다.

temperature	lifespan LSMEAN	LSMEAN Number
700	36,222222	1
800	39,555556	2
900	42,888889	3
1000	45,444444	4

Least Squares Means for effect temperature Pr > t for H0: LSMean(i)=LSMean(j)				
Dependent Variable: lifespan				
i/j	1	2	3	4
1		0,0402	<,0001	<,0001
2	0,0402		0,0402	0,0002
3	<,0001	0,0402		0,1539
4	<,0001	0,0002	0,1539	

분할구 설계 (Split - Plot Design)

01. 선정 이유

02. 모형식

03. 분산분석표 및 해석

04. 잔차분석

05. 다중 비교

06 결과 해석

결과 해석

- 잔차는 정규성, 등분산성, 독립성을 모두 만족
- 방법 요인과 온도요인에 대한 다중비교의 모든 검정 결과에서 각 방법 사이, 각 온도 사이에서 유의한 차이가 발견됨
- 분할구 설계의 경우 유의한 모형이라고 할 수 있음

04

최종 모형 선정

01. 최종 모형 선정

02. 최종 결론

최종 모형 선정

01 최종 모형 선정

02 최종 결론

최종 모형 선정

분할구 설계(Split plot design)로 선정

최종 결론

- 이원배치법은 CRD로 실행되면 모든 실험단위가 같은 확률로 조건에 배치되어야 함
- 우리가 얻고자 하는 결과는 요인의 수준이 랜덤 배치되는 실험단위가 달라지는 모형이어야 함
- 배터리 수명의 차이가 온전히 방법과 온도의 차이인지에 대해서 알고자 하는데 반복이 실험에 영향을 미칠 수도 있다고 생각하므로 이를 상쇄시킬 수 있게 반복을 블록하고자 함
- 반복을 블록으로 한 분할구 설계로 했을 때, 블록 효과 또한 유의함
- R-Square 값 또한 분할구 설계로 했을 때가 높음
- 반복을 블록 요인으로 한 분할구 설계를 최종 모형으로 선정함

감사합니다