Single Factor Model

$Y_{ij} = \mu + \tau_j + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim NID(0, \sigma_\varepsilon^2)$

Example)

Source	DF	SS	MS	F-value	P-value
Factor	2	4608.2	2304.1	39.99	0.0001
Error	9	518.5	57.6		
Total	11	5126.7			

$$E\left(\frac{\text{Sum of Square of Factor or Error}}{\text{degree of freedom for Factor or Error}}\right) = E(\text{ MS : Mean Square }) : \text{Expected Mean Square, EMS}$$

① Fixed model : $\sum_{j=1}^{k} \tau_j = \sum_{j=1}^{k} (\mu_j - \mu) = 0$

Source	df	EMS
Factor	k-1	$\sigma_\varepsilon^2 + n\phi_\tau$
Error	k(n-1)	σ_ε^2

where $\phi_\tau = \frac{1}{k-1} \sum_{j=1}^{k} \tau_j^2$

k : the number of treatment

n : the number of replication at each treatment

$H_0 : {}^\forall\tau_j = 0 \text{ for } j = 1,2,\cdots,n$

Test statistics : $F = \frac{MS_{factor}}{MS_{error}}$

$$E(F) \approx E\left(\frac{MS_{factor}}{MS_{error}}\right) = \frac{\sigma_\varepsilon^2 + n\phi_\tau}{\sigma_\varepsilon^2}$$

$$= 1 + n \cdot \frac{\phi_\tau}{\sigma_\varepsilon^2} \xrightarrow{H_0 :\, {}^\forall\tau_j = 0} 1 : H_0 \text{ isn't rejected.}$$

즉, F test는 순수한 Factor의 효과 (ϕ_τ)를 검정함.

② Random model : $\tau_j \sim NID(0, \sigma_\tau^2)$

Source	df	EMS
Factor	k-1	$\sigma_\varepsilon^2 + n\sigma_\tau^2$
Error	k(n-1)	σ_ε^2

k : the number of treatment

n : the number of replication at each treatment

$H_0 : \sigma_\tau^2 = 0$

Test statistics : $F = \frac{MS_{factor}}{MS_{error}}$

$$E(F) \approx E\left(\frac{MS_{factor}}{MS_{error}}\right) = \frac{\sigma_\varepsilon^2 + n\sigma_\tau^2}{\sigma_\varepsilon^2}$$

$$= 1 + n \cdot \frac{\sigma_\tau^2}{\sigma_\varepsilon^2} \xrightarrow{H_0 :\, \sigma_\tau^2 = 0} 1 : H_0 \text{ isn't rejected.}$$

즉, F test는 순수한 Factor의 효과 (σ_τ^2)를 검정함

① Fixed model in single factor model

분산 분석: Y 대 A

요인	유형	수준	값	
A	고정됨	3	60, 90, 120	← $k = 3$

Y에 대한 분산 분석

출처	DF	SS	MS	F	P	
A	2	4608.2	2304.1	39.99	0.000	← $\frac{2304.1}{57.6} = 39.99$
오차	9	518.5	57.6			
총계	11	5126.7				

S = 7.59020   R-제곱 = 89.89%   R-제곱(수정) = 87.64%

	출처	분산 성분	오차 항	각 항에 대한 기대 평균 제곱(제한적 모형 사용)	
1	A		2	(2) + 4 Q[1]	← $\sigma_\varepsilon^2 + n\phi_\tau$ ↔ $Q[1] = \phi_\tau$, $n = 4$
2	오차	57.61		(2)	← σ_ε^2 ↔ $(2) = \sigma_\varepsilon^2$

② Random model in single factor model

분산 분석: Y 대 A

요인	유형	수준	값	
A	랜덤	3	60, 90, 120	← $k = 3$

Y에 대한 분산 분석

출처	DF	SS	MS	F	P	
A	2	4608.2	2304.1	39.99	0.000	← $\frac{2304.1}{57.6} = 39.99$
오차	9	518.5	57.6			
총계	11	5126.7				

S = 7.59020   R-제곱 = 89.89%   R-제곱(수정) = 87.64%

	출처	분산 성분	오차 항	각 항에 대한 기대 평균 제곱(제한적 모형 사용)	
1	A	561.62	2	(2) + 4 (1)	← $\sigma_\varepsilon^2 + n\sigma_\tau^2$ ↔ $(1) = \sigma_\tau^2$, $n = 4$
2	오차	57.61		(2)	← σ_ε^2 ↔ $(2) = \sigma_\varepsilon^2$

Two Factor Model

$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{k(ij)},\ \ \varepsilon_{k(ij)} \sim NID(0, \sigma_\varepsilon^2)$ where $i = 1,2,\cdots,a,\ \ j = 1,2,\cdots,b,\ \ k = 1,2,\cdots,n$

① Fixed model
Assumptions : $\sum_{i=1}^{a} A_i = 0, \sum_{j=1}^{b} B_j = 0 \ \rightarrow\ \sum_{i=1}^{a} AB_{ij} = 0, \sum_{j=1}^{b} AB_{ij} = 0$
$H_{01} : {}^{\forall}A_i = 0,\ \ H_{02} : {}^{\forall}B_j = 0,\ \ H_{03} : {}^{\forall}AB_{ij} = 0$

② Random model
Assumptions : $A_i \sim NID(0, \sigma_A^2), B_j \sim NID(0, \sigma_B^2) \ \rightarrow\ AB_{ij} \sim NID(0, \sigma_{AB}^2)$
$H_{01} : \sigma_A^2 = 0,\ \ H_{02} : \sigma_B^2 = 0,\ \ H_{03} : \sigma_{AB}^2 = 0$

③ Mixed model
Assumptions : $\sum_{i=1}^{a} A_i = 0, B_j \sim NID(0, \sigma_B^2) \ \rightarrow\ \sum_{i=1}^{a} AB_{ij} = 0, \sum_{j=1}^{b} AB_{ij} \neq 0$
$H_{01} : {}^{\forall}A_i = 0,\ \ H_{02} : \sigma_B^2 = 0,\ \ H_{03} : \sigma_{AB}^2 = 0$

Source	df	EMS		
		Fixed	Random	Mixed
A_i	a-1	$\sigma_\varepsilon^2 + nb\phi_A$	$\sigma_\varepsilon^2 + n\sigma_{AB}^2 + nb\sigma_A^2$	$\sigma_\varepsilon^2 + n\sigma_{AB}^2 + nb\phi_A$
B_j	b-1	$\sigma_\varepsilon^2 + na\phi_B$	$\sigma_\varepsilon^2 + n\sigma_{AB}^2 + na\sigma_B^2$	$\sigma_\varepsilon^2 + na\sigma_B^2$
AB_{ij}	(a-1)(b-1)	$\sigma_\varepsilon^2 + n\phi_{AB}$	$\sigma_\varepsilon^2 + n\sigma_{AB}^2$	$\sigma_\varepsilon^2 + n\sigma_{AB}^2$
$\varepsilon_{k(ij)}$	ab(n-1)	σ_ε^2	σ_ε^2	σ_ε^2

Source	Test Statistics		
	Fixed model	Random model	Mixed model
A_i	$\frac{MS_A}{MS_E} = \frac{\sigma_\varepsilon^2 + nb\phi_A}{\sigma_\varepsilon^2}$	$\frac{MS_A}{MS_{AB}} = \frac{\sigma_\varepsilon^2 + n\sigma_{AB}^2 + nb\sigma_A^2}{\sigma_\varepsilon^2 + n\sigma_{AB}^2}$	$\frac{MS_A}{MS_{AB}} = \frac{\sigma_\varepsilon^2 + n\sigma_{AB}^2 + nb\phi_A}{\sigma_\varepsilon^2 + n\sigma_{AB}^2}$
B_j	$\frac{MS_B}{MS_E} = \frac{\sigma_\varepsilon^2 + na\phi_B}{\sigma_\varepsilon^2}$	$\frac{MS_A}{MS_{AB}} = \frac{\sigma_\varepsilon^2 + n\sigma_{AB}^2 + na\sigma_B^2}{\sigma_\varepsilon^2 + n\sigma_{AB}^2}$	$\frac{MS_B}{MS_E} = \frac{\sigma_\varepsilon^2 + na\sigma_B^2}{\sigma_\varepsilon^2}$
AB_{ij}	$\frac{MS_{AB}}{MS_E} = \frac{\sigma_\varepsilon^2 + n\phi_{AB}}{\sigma_\varepsilon^2}$	$\frac{MS_{AB}}{MS_E} = \frac{\sigma_\varepsilon^2 + n\sigma_{AB}^2}{\sigma_\varepsilon^2}$	$\frac{MS_{AB}}{MS_E} = \frac{\sigma_\varepsilon^2 + n\sigma_{AB}^2}{\sigma_\varepsilon^2}$

* **EMS Rule**에 따라, 순수한 Factor의 효과만을 검정할 수 있도록 검정 통계량이 구성된다.
* 순수한 Factor의 효과만을 검정할 수 있도록 검정 통계량이 구성될 수 없는 경우가 발생할 수 있다.
  이런 경우에는 유사 F 검정(Pseudo F test)를 통해 검정한다.

① Fixed model in two factor model

요인	유형	수준	값	
A	고정됨	3	60, 90, 120	$\leftarrow a = 3$
B	고정됨	2	127, 220	$\leftarrow b = 2$

Y에 대한 분산 분석

출처	DF	SS	MS	F	P	
A	2	4608.17	2304.08	99.46	0.000	$\leftarrow \frac{2304.08}{23.17} = 99.46$
B	1	96.33	96.33	4.16	0.088	$\leftarrow \frac{96.33}{23.17} = 4.16$
A*B	2	283.17	141.58	6.11	0.036	$\leftarrow \frac{141.58}{23.17} = 6.11$
오차	6	139.00	23.17			
총계	11	5126.67				

S = 4.81318 R-제곱 = 97.29% R-제곱(수정) = 95.03%

	출처	분산 성분	오차 항	각 항에 대한 기대 평균 제곱(제한적 모형 사용)		
1	A		4	(4) + 4 Q[1]	$\leftarrow \sigma_\varepsilon^2 + nb\phi_A$	$\leftrightarrow$ $Q[1] = \phi_A$, $n = 2$, $b = 2$
2	B		4	(4) + 6 Q[2]	$\leftarrow \sigma_\varepsilon^2 + na\phi_B$	$\leftrightarrow$ $Q[2] = \phi_B$, $n = 2$, $a = 3$
3	A*B		4	(4) + 2 Q[3]	$\leftarrow \sigma_\varepsilon^2 + n\phi_{AB}$	$\leftrightarrow$ $Q[3] = \phi_{AB}$, $n = 2$
4	오차	23.17		(4)	$\leftarrow \sigma_\varepsilon^2$	$\leftrightarrow$ $(4) = \sigma_\varepsilon^2$

② Random model in two factor model

요인	유형	수준	값	
A	랜덤	3	60, 90, 120	$\leftarrow a = 3$
B	랜덤	2	127, 220	$\leftarrow b = 2$

Y에 대한 분산 분석

출처	DF	SS	MS	F	P	
A	2	4608.17	2304.08	16.27	0.058	$\leftarrow \frac{2304.08}{141.58} = 16.27$
B	1	96.33	96.33	0.68	0.496	$\leftarrow \frac{96.33}{141.58} = 0.68$
A*B	2	283.17	141.58	6.11	0.036	$\leftarrow \frac{141.58}{23.17} = 6.11$
오차	6	139.00	23.17			
총계	11	5126.67				

S = 4.81318 R-제곱 = 97.29% R-제곱(수정) = 95.03%

	출처	분산 성분	오차 항	각 항에 대한 기대 평균 제곱(제한적 모형 사용)		
1	A	540.625	3	(4) + 2 (3) + 4 (1)	$\leftarrow \sigma_\varepsilon^2 + n\sigma_{AB}^2 + nb\sigma_A^2$	$\leftrightarrow$ $(1) = \sigma_A^2$, $n = 2$, $b = 2$
2	B	-7.542	3	(4) + 2 (3) + 6 (2)	$\leftarrow \sigma_\varepsilon^2 + n\sigma_{AB}^2 + na\sigma_B^2$	$\leftrightarrow$ $(2) = \sigma_B^2$, $n = 2$, $a = 3$
3	A*B	59.208	4	(4) + 2 (3)	$\leftarrow \sigma_\varepsilon^2 + n\sigma_{AB}^2$	$\leftrightarrow$ $(3) = \sigma_{AB}^2$, $n = 2$
4	오차	23.167		(4)	$\leftarrow \sigma_\varepsilon^2$	$\leftrightarrow$ $(4) = \sigma_\varepsilon^2$

③ Mixed model in two factor model

요인	유형	수준	값	
A	고정됨	3	60, 90, 120	$\leftarrow a = 3$
B	랜덤	2	127, 220	$\leftarrow b = 2$

Y에 대한 분산 분석

출처	DF	SS	MS	F	P	
A	2	4608.17	2304.08	16.27	0.058	$\leftarrow \frac{2304.08}{141.58} = 16.27$
B	1	96.33	96.33	4.16	0.088	$\leftarrow \frac{96.33}{23.17} = 4.16$
A*B	2	283.17	141.58	6.11	0.036	$\leftarrow \frac{141.58}{23.17} = 6.11$
오차	6	139.00	23.17			
총계	11	5126.67				

S = 4.81318   R-제곱 = 97.29%   R-제곱(수정) = 95.03%

	출처	분산 성분	오차 항	각 항에 대한 기대 평균 제곱(제한적 모형 사용)		
1	A		3	(4) + 2 (3) + 4 Q[1]	$\leftarrow \sigma_\varepsilon^2 + n\sigma_{AB}^2 + nb\phi_A$	$\leftrightarrow$ $Q[1] = \phi_A$, $n = 2$, $b = 2$
2	B	12.19	4	(4) + 6 (2)	$\leftarrow \sigma_\varepsilon^2 + na\sigma_B^2$	$\leftrightarrow$ $(2) = \sigma_B^2$, $n = 2$, $a = 3$
3	A*B	59.21	4	(4) + 2 (3)	$\leftarrow \sigma_\varepsilon^2 + n\sigma_{AB}^2$	$\leftrightarrow$ $(3) = \sigma_{AB}^2$, $n = 2$
4	오차	23.17		(4)	$\leftarrow \sigma_\varepsilon^2$	$\leftrightarrow$ $(4) = \sigma_\varepsilon^2$

Pseudo F test in three factor model

요인	유형	수준	값
D	랜덤	2	1, 2
O	랜덤	3	A, B, C
G	고정됨	3	2, 4, 6

Thickness에 대한 분산 분석

출처	DF	SS	MS	F	P	
D	1	0.00100	0.00100	0.34	0.621	
O	2	0.11207	0.05604	18.77	0.051	
G	2	1.57317	0.78659	56.58	0.001 x	← Pseudo F test
D*O	2	0.00597	0.00299	9.19	0.002	
D*G	2	0.01134	0.00567	2.29	0.218	
O*G	4	0.04284	0.01071	4.32	0.093	
D*O*G	4	0.00991	0.00248	7.62	0.001	
오차	18	0.00585	0.00033			
총계	35	1.76216				

x 정확한 F-검정이 아님. ← Pseudo F test

S = 0.0180278 R-제곱 = 99.67% R-제곱(수정) = 99.35%

	출처	분산 성분	오차 항	각 항에 대한 기대 평균 제곱(제한적 모형 사용)	
1	D	-0.00011	4	(8) + 6 (4) + 18 (1)	
2	O	0.00442	4	(8) + 6 (4) + 12 (2)	
3	G		*	(8) + 2 (7) + 4 (6) + 6 (5) + 12 Q[3]	← Pseudo F test
4	D*O	0.00044	8	(8) + 6 (4)	
5	D*G	0.00053	7	(8) + 2 (7) + 6 (5)	
6	O*G	0.00206	7	(8) + 2 (7) + 4 (6)	
7	D*O*G	0.00108	8	(8) + 2 (7)	
8	오차	0.00033		(8)	

* 합성 검정.

합성 검정에 대한 오차 항

출처	오차 DF	오차 MS	오차 평균 제곱의 합성
3 G	4.18	0.01390	(5) + (6) - (7)

$$\begin{aligned} E(MS) &= \sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2 + 6\sigma_{DG}^2 + 12\phi_G \\ &= E(MS_{DG}) + E(MS_{OG}) - E(MS_{DOG}) \\ &= (\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2) + (\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 6\sigma_{DG}^2) - (\sigma_\varepsilon^2 + 2\sigma_{DOG}^2) = \sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2 + 6\sigma_{DG}^2 \end{aligned}$$

F test statistics for G : $\dfrac{MS_G}{MS} = \dfrac{\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2 + 6\sigma_{DG}^2 + 12\phi_G}{\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2 + 6\sigma_{DG}^2}$ ← 이런 식으로 분모를 만들어서 검정함.