

# Randomized Complete Block Design

Đỗ Trọng Hợp

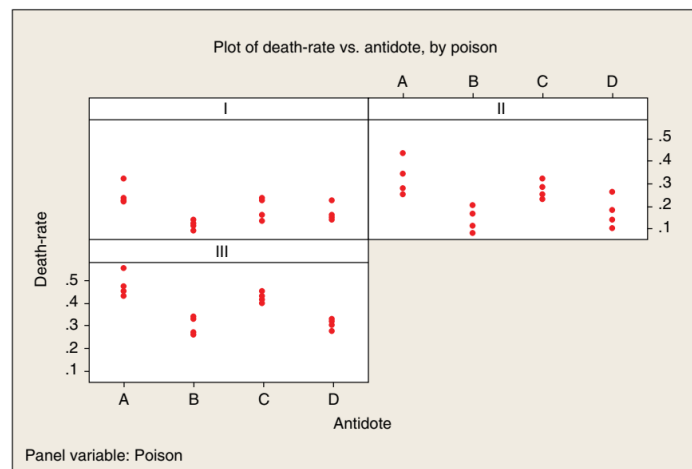
Khoa Khoa Học và Kỹ Thuật Thông Tin

Đại học Công Nghệ Thông Tin tp. Hồ Chí Minh

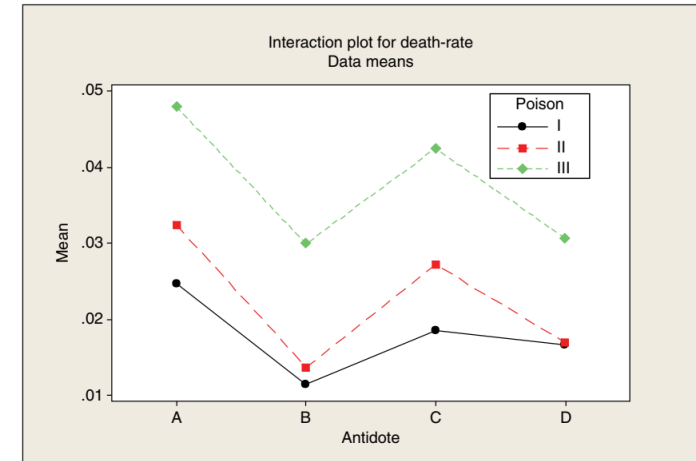
# Nhắc lại thí nghiệm Poison-Antidote (2 yếu tố)

**Table 5.2** Average Death Rates (units of hour<sup>-1</sup>) by Poisons and Antidotes.

Average Death Rate		Antidote			
Poison		A	B	C	D
I		.25	.12	.19	.17
II		.33	.14	.27	.17
III		.48	.30	.43	.31
Average		.35	.19	.30	.22



**Figure 5.2** Plot of Death Rates versus Antidotes, by Poisons.



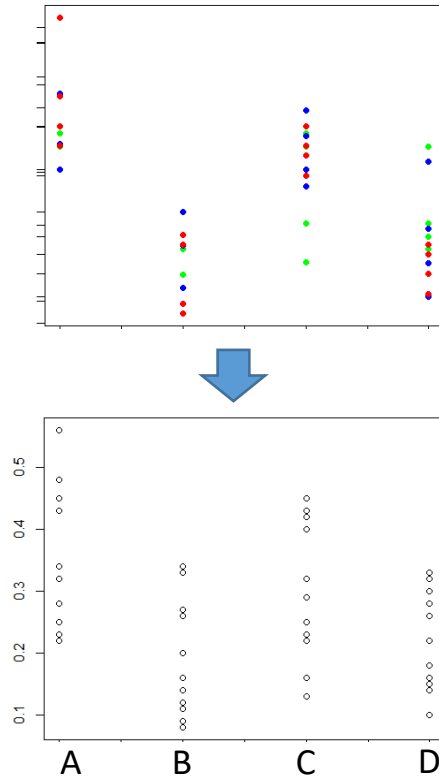
**Figure 5.4** Interaction Plot. Average death rate versus antidote, by poison.

**Table 5.3** ANOVA for Death Rate.

Source	DF	SS	MS	F	P
Poison	2	.349	.174	72.6	.000
Antidote	3	.204	.068	28.3	.000
Interaction	6	.016	.0026	1.1	.39
Error	36	.086	.0024		
Total	47	.655			

Có thể kiểm định sự khác nhau giữa các nhóm (ví dụ antidote) nếu dùng one-way ANOVA hay không?

# One way ANOVA



$$SS = SS_{Poison} + SS_{Antidote} + SS_{Interaction} + SSE$$

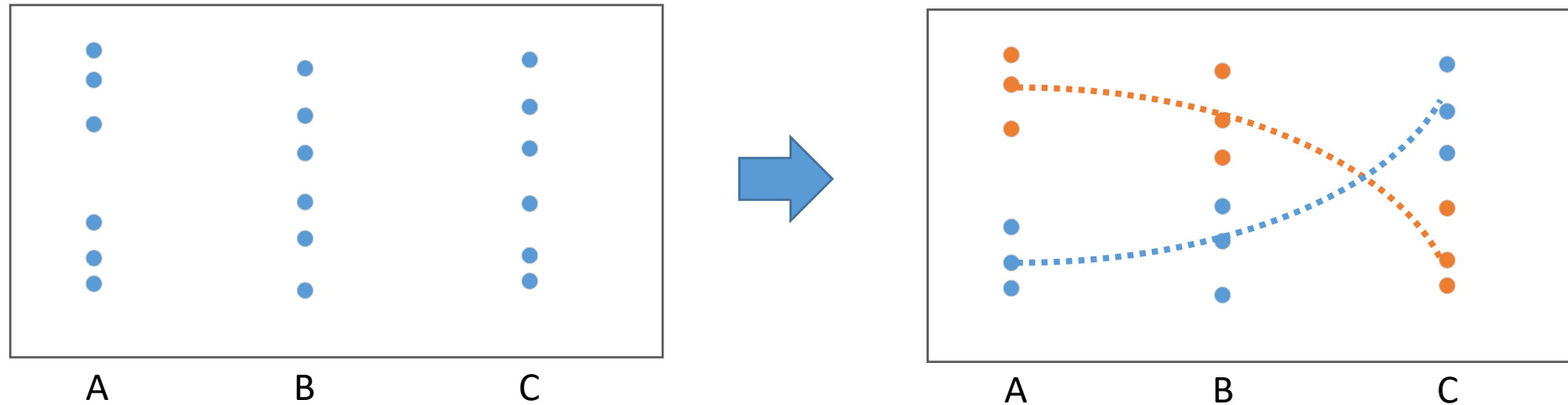
Source	DF	SS	MS	F	P
Poison	2	.349	.174	72.6	.000
Antidote	3	.204	.068	28.3	.000
Interaction	6	.016	.0026	1.1	.39
Error	36	.086	.0024		
Total	47	.655			

$$SS = SS_{Antidote} + SSE$$

Source	DF	SS	MS	F	P
Antidote	3	.204	.068	28.3	.000
Error	44	.451	.01025	6.6	.0009
Total	47	.655			

- Nếu dùng one way ANOVA, các biến thiên bên trong các nhóm A, B, C, D không thể được phân biện
- Tất cả các biến thiên bên trong các nhóm A, B, C, D đều được cho là lỗi ngẫu nhiên
- MSE sẽ lớn hơn, dẫn đến F-value sẽ nhỏ hơn, dẫn đến P-value lớn hơn giá trị thật, dẫn đến khả năng chấp nhận giả thuyết H0 sai (false negative) (tức là độ nhạy của kiểm định bị giảm)

# Source of variance within group

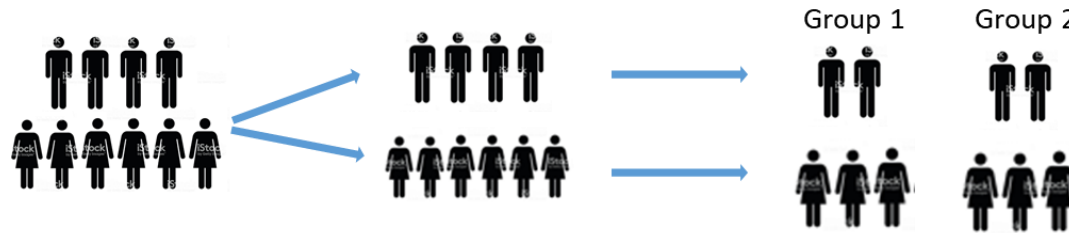


- Biến thiên bên trong 1 group có thể gây ra bởi
  - Ảnh hưởng của yếu tố khác (trường hợp thí nghiệm nhiều yếu tố)
  - Sự không đồng nhất của experiment unit trong mỗi nhóm (thí nghiệm 1 yếu tố)
  - Biến thiên ngẫu nhiên
- Để phân tích chính xác, ta cần tách các nguồn gây ra các biến thiên

# Blocking

- Thí nghiệm thuốc tăng lực. Đối tượng là 1 nhóm không đồng nhất gồm nam nữ lẫn lộn. Chia ngẫu nhiên thành 2 nhóm để thử thuốc thật và thuốc giả để so sánh.

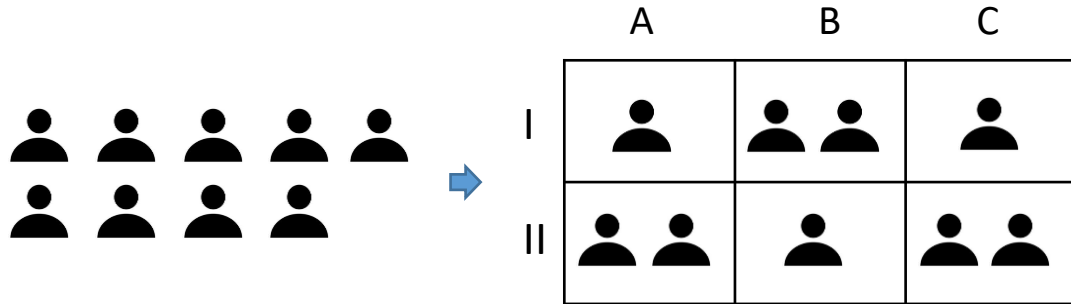
- Do nam thường chạy nhanh hơn nữ, kết quả sẽ bị ảnh hưởng nếu 1 nhóm có nhiều nam hơn đáng kể.



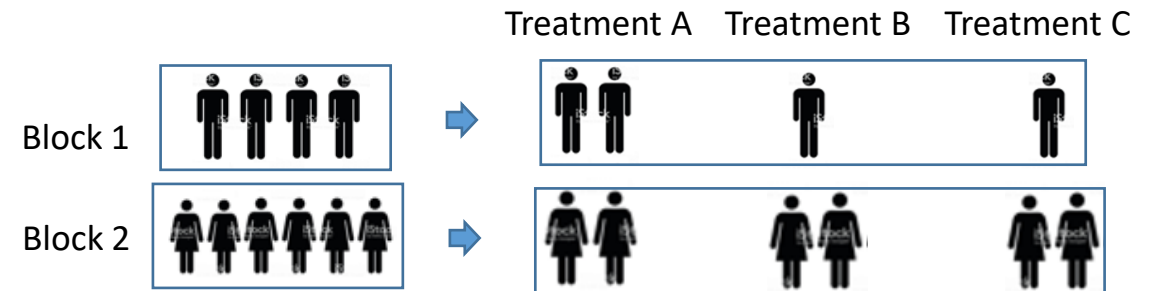
- Sửa lại thiết kế: tách nam và nữ thành 2 nhóm (block), sau đấy chọn ngẫu nhiên những thành viên nam và nữ từ 2 block này vào 2 group thử thuốc.
- Nguyên tắc blocking:
  - Tách các E unit thành các khối (block) dựa theo yếu tố (factor) có khả năng ảnh hưởng đến kết quả.
  - Áp dụng ngẫu nhiên các treatment lên các E unit trong mỗi khối.

# Blocking vs multiple factor experiment

- Với multiple factor experiment, toàn bộ đối tượng thí nghiệm là một nhóm đồng nhất.
- Mỗi cell (kết hợp của các yếu tố) là một treatment.
- Các đối tượng sẽ được đưa một cách ngẫu nhiên vào các cell này.
- Các yếu tố chính là yếu tố đang được quan tâm phân tích.



- Với blocking, đối tượng thí nghiệm là một nhóm không đồng nhất. Tức là các đối tượng thí nghiệm khác nhau bởi yếu tố blocking.
- Toàn bộ đối tượng thí nghiệm sẽ được chia thành các khối (block).
- Các đối tượng trong mỗi khối sẽ được đưa một cách ngẫu nhiên vào các nhóm nhận các treatment khác nhau.
- Yếu tố blocking không phải là yếu tố đang được quan tâm phân tích mà thường là yếu tố có khả năng gây nhiễu cho kết quả.

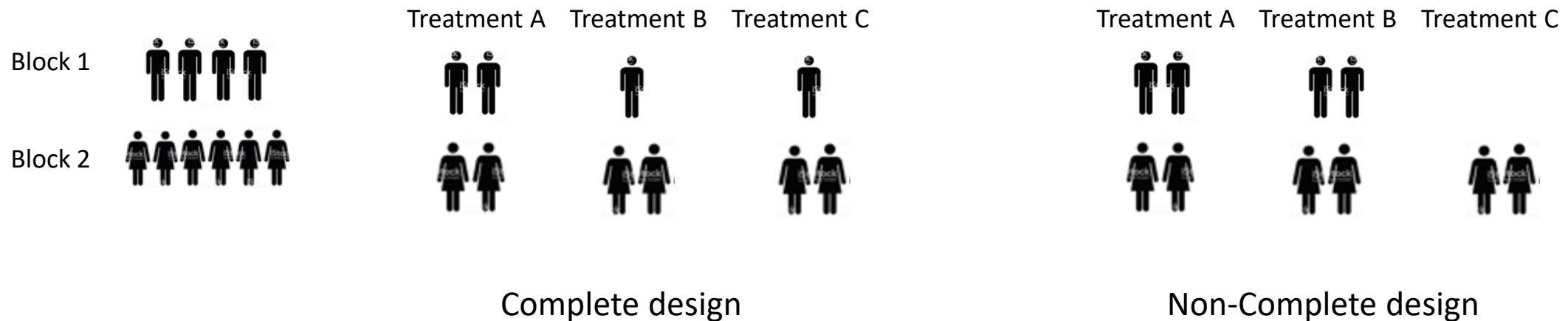


# Blocking vs randomization

- Khi yếu tố gây nhiễu có thể kiểm soát được thì tạo khối
  - Giới tính ảnh hưởng đến kết quả chạy (thử nghiệm thuốc tăng lực) → tạo khối nam và nữ để cho kết quả không có thiên vị. Lưu ý ở đây ta không quan tâm ảnh hưởng của giới tính nên giới tính chỉ là yếu tố tạo khối chứ không phải yếu tố chính. Mục tiêu thí nghiệm là tìm ảnh hưởng của thuốc lên con người nói chung và ảnh hưởng này sẽ được đánh giá qua kết quả của cả nam và nữ.
  - Giống lúa ảnh hưởng đến năng suất (thử nghiệm phân bón) → tạo khối cho mỗi giống lúa. Ở đây ta không quan tâm ảnh hưởng của giống lúa nên giống lúa chỉ là yếu tố tạo khối chứ không phải yếu tố chính. Mục tiêu của thí nghiệm là xem phân bón nào có hiệu quả hơn với cây lúa nói chung.
- Khi yếu tố gây nhiễu không được biết hoặc biết mà thể kiểm soát được thì dùng ngẫu nhiên
  - Stress và nhiều yếu tố chưa biết khác ảnh hưởng đến huyết áp. Những yếu tố này không dễ để kiểm soát. Do đó cần chọn ngẫu nhiên các đối tượng thử thuốc trong thí nghiệm thuốc kiểm soát huyết áp. Số lượng đối tượng cần thiết để cân bằng các yếu tố gây nhiễu sẽ nhiều hơn so với khi tạo khối.

# Randomized Complete Block design

- **Randomized block design** là thiết kế mà các E unit được nhóm thành các khối. Các treatments được áp dụng ngẫu nhiên cho các E unit trong mỗi khối.
- **Completely randomized block design** là thiết kế mà mỗi treatment được áp dụng ít nhất 1 lần trong mỗi khối.

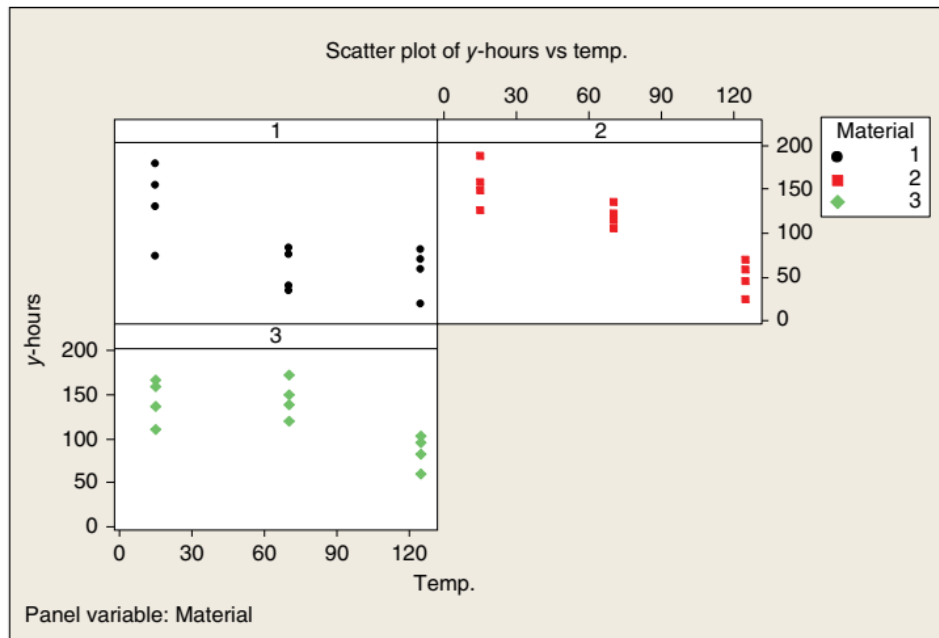


Treatment C không được áp dụng lần nào trong block 1.



# Example 1: battery experiment

- Thử nghiệm ảnh hưởng của nhiệt độ lên thời gian sử dụng của pin
- Đối tượng thử nghiệm gồm 36 viên pin
- Các viên pin này không đồng nhất vì chúng thuộc 3 nhà sản xuất khác nhau (mỗi nhà sx dùng 1 loại vật liệu để sx pin)
- Các viên pin được chia thành 3 block. Sau đấy các viên pin trong mỗi block sẽ được áp dụng ngẫu nhiên 1 treatment (là 1 loại nhiệt độ).

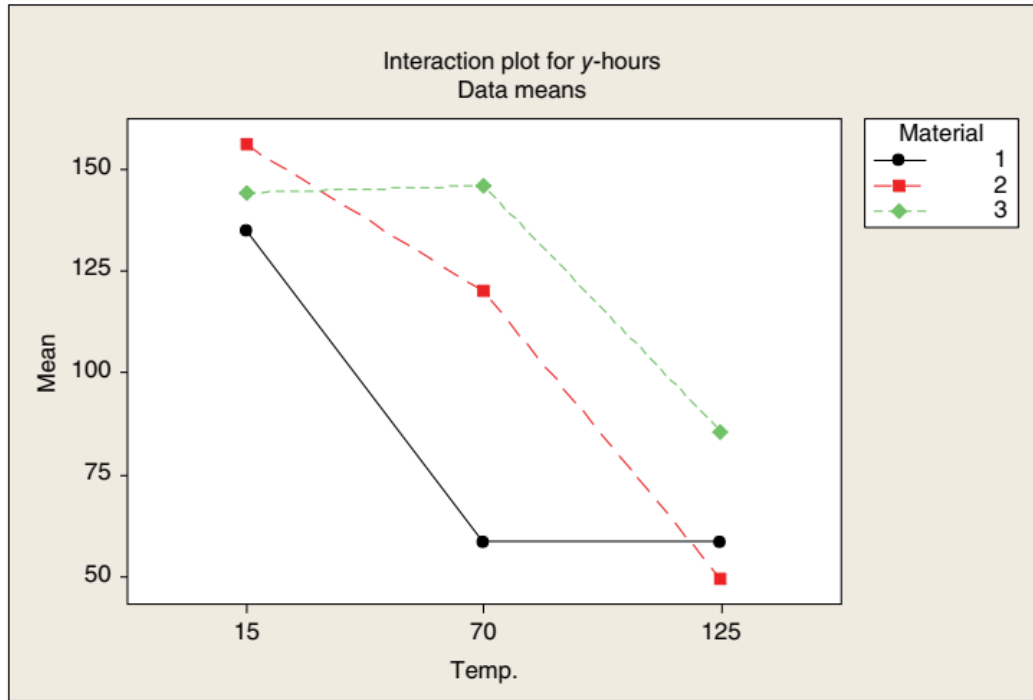


**Figure 6.1** Scatter Plot of Battery Lifetimes versus Temperature, by Material.

**Table 6.1** Battery Lifetimes by Material and Temperature.

Material	Temperature (°F)		
	15	70	125
1	130	34	20
	74	80	82
	155	40	70
	180	75	58
2	150	136	25
	159	106	58
	188	122	70
	126	115	45
3	138	174	96
	168	150	82
	110	120	104
	160	139	60

Table 5.1, p. 176 of Montgomery (2001), reproduced by permission of John Wiley & Sons.



**Figure 6.2** Interaction Plot of Material/Temperature Means.

b là số (block) hàng, t là số treatment (cột), r là số replication  
 $b \times t \times r$  là tổng số đơn vị thí nghiệm

**Table 6.2** ANOVA Structure for Randomized Complete Block Design with Replication.

Source	df	SS	MS	F	P
Blocks (B)	$b - 1$				
Treatments (T)	$t - 1$				
B $\times$ T interaction	$(b - 1)(t - 1)$				
Error	$bt(r - 1)$				
Total	$btr - 1$				

- Dùng two way ANOVA như trường hợp thí nghiệm nhiều yếu tố

# Two way ANOVA (kiểm định cho các ảnh hưởng)

$SS_T = \sum_k \sum_j \sum_i (x_{ijk} - \bar{x})^2$	$df_T = n - 1$	$MS_T = SS_T/df_T$
$SS_A = mc \sum_i (\bar{x}_i - \bar{x})^2$	$df_A = r - 1$	$MS_A = SS_A/df_A$
$SS_B = mr \sum_j (\bar{x}_j - \bar{x})^2$	$df_B = c - 1$	$MS_B = SS_B/df_B$
$SS_{AB} = m \sum_j \sum_i (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2$	$df_{AB} = (r - 1)(c - 1)$	$MS_{AB} = SS_{AB}/df_{AB}$
$SS_W = \sum_k \sum_j \sum_i (x_{ijk} - \bar{x}_{ij})^2$	$df_W = n - rc$	$MS_W = SS_W/df_W$

Test Desired	Null Hypothesis $H_0$	Equivalent $H_0$	Statistical Test
Effect of blocking factor	$\mu_1 = \mu_2 = \dots = \mu_i$ for all $i$	$\alpha_i = 0$ for all $i$	$\frac{MS_A}{MS_W} \sim F(df_A, df_W)$
Effect of treatment factor	$\mu_1 = \mu_2 = \dots = \mu_j$ for all $j$	$\beta_j = 0$ for all $j$	$\frac{MS_B}{MS_W} \sim F(df_B, df_W)$
Effect of interaction		$\delta_{ij} = 0$ for all $i, j$	$\frac{MS_{AB}}{MS_W} \sim F(df_{AB}, df_W)$

- F-test kiểm định 3 giả thuyết không sau:
  - $H_0$ : mean của các mức của yếu tố hàng (blocking) đều bằng nhau (tức là các ảnh hưởng  $\alpha_i$  đều bằng 0)
  - $H_0$ : mean của các mức của yếu tố cột (treatment) đều bằng nhau (tức là các ảnh hưởng  $\beta_j$  đều bằng 0)
  - $H_0$ : không có tương tác giữa hàng và cột (tức là các ảnh hưởng tương tác  $\delta_{ij}$  đều bằng 0)

**Table 6.1** Battery Lifetimes by Material and Temperature.

Material	Temperature (°F)		
	15	70	125
1	130	34	20
	74	80	82
	155	40	70
	180	75	58
2	150	136	25
	159	106	58
	188	122	70
	126	115	45
3	138	174	96
	168	150	82
	110	120	104
	160	139	60

Table 5.1, p. 176 of Montgomery (2001), reproduced by permission of John Wiley & Sons.

**Table 6.3** ANOVA for Battery Experiment.

Source	DF	SS	MS	F	P
Material	2	10684	5342	7.91	.002
Temp	2	39119	19559	29.0	.000
Interaction	4	9614	2403	3.56	.019
Error	27	18235	675		
Total	35	77647			

$$SS = SS_{Material} + SS_{Temp} + SS_{Interaction} + SSE$$

- Do ảnh hưởng của Interaction có ý nghĩa, ta không thể lấy kết quả trung bình tại mỗi nhiệt độ để so sánh và đưa ra kết luận về ảnh hưởng của nhiệt độ được
- Nếu không tạo block thì phân tích có gì khác?
  - $SSE = ?$        $MSE = ?$        $F_{Temp} = ?$

# Ước lượng tỷ lệ pin đạt yêu cầu

**Table 6.3** ANOVA for Battery Experiment.

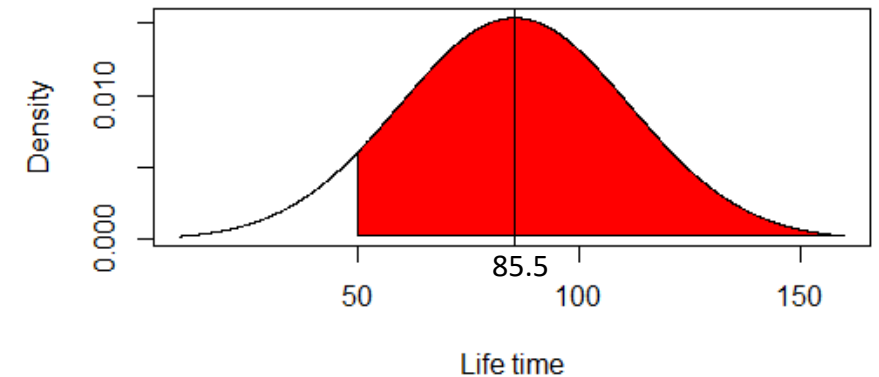
Source	DF	SS	MS	F	P
Material	2	10684	5342	7.91	.002
Temp	2	39119	19559	29.0	.000
Interaction	4	9614	2403	3.56	.019
Error	27	18235	675		
Total	35	77647			

- Pin gọi là đạt yêu cầu nếu có thể sử dụng ít nhất 50h tại nhiệt độ 125°F
- Tính tỷ lệ pin làm bằng vật liệu loại 3 (M3) đạt yêu cầu
- Với **mức tin cậy thấp**, ta có thể giả sử thời gian sử dụng ltime của pin M3 là biến ngẫu nhiên có phân phối chuẩn với mean = 85.5h (từ bảng 6.1) và độ lệch chuẩn  $s = \sqrt{675} = 26h$  (từ bảng ANOVA)
- Do đó, tỷ lệ pin đạt chuẩn là  $P(ltime > 50) = 91\%$

**Table 6.1** Battery Lifetimes by Material and Temperature.

Material	Temperature (°F)		
	15	70	125
1	130	34	20
	74	80	82
	155	40	70
	180	75	58
2	150	136	25
	159	106	58
	188	122	70
	126	115	45
3	138	174	96
	168	150	82
	110	120	104
	160	139	60

Table 5.1, p. 176 of Montgomery (2001), reproduced by permission of John Wiley & Sons



# Ước lượng tỷ lệ pin đạt yêu cầu

**Table 6.3** ANOVA for Battery Experiment.

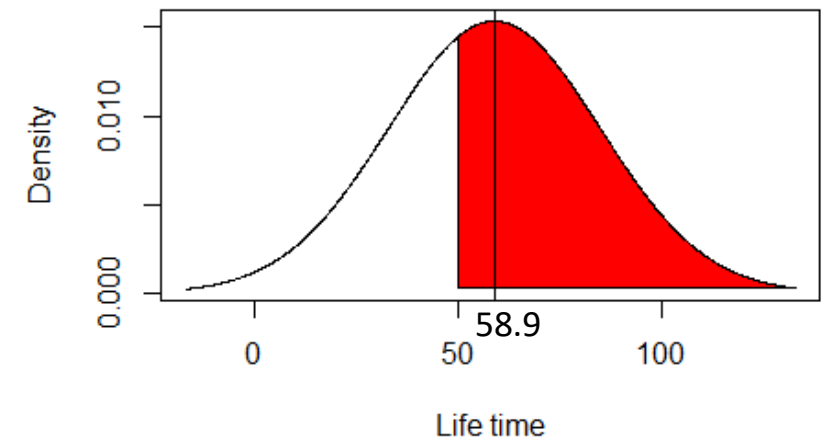
Source	DF	SS	MS	F	P
Material	2	10684	5342	7.91	.002
Temp	2	39119	19559	29.0	.000
Interaction	4	9614	2403	3.56	.019
Error	27	18235	675		
Total	35	77647			

- Khoảng tin cậy 95% của thời gian sử dụng trung bình của pin M3 là:
- $$C.I = 85.5 \pm t_{.025}(27) * \sqrt{\frac{MSE}{4}} = 85.5 \pm 26.65 = [58.9; 112.2]$$
- Với mức tin cậy 97.5%, mean của thời gian sử dụng sẽ  $\geq 58.9$
- Tính  $P(\text{time} > 50)$  của biến ngẫu nhiên có phân phối chuẩn với mean = 58.9h và độ lệch chuẩn  $s = \sqrt{675} = 26h$  (từ bảng ANOVA)
- Với **mức tin cậy 97.5%**, tỷ lệ pin đạt chuẩn **thấp nhất** là  $P(\text{time} > 50) = 63\%$

**Table 6.1** Battery Lifetimes by Material and Temperature.

Material	Temperature (°F)		
	15	70	125
1	130	34	20
	74	80	82
	155	40	70
	180	75	58
2	150	136	25
	159	106	58
	188	122	70
	126	115	45
3	138	174	96
	168	150	82
	110	120	104
	160	139	60

Table 5.1, p. 176 of Montgomery (2001), reproduced by permission of John Wiley & Sons.



Tính xác suất phân phối chuẩn dung z-table

## STANDARD NORMAL DISTRIBUTION

$z$ -distribution

$$X \sim N(\mu, \sigma) \longrightarrow Z \sim N(0, 1)$$

$$Z = \frac{x - \mu}{\sigma}$$

z-score: number of standard deviations from the mean

$$P(<X) = P(<Z)$$





LESS THAN  
*Cumulative*



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
.	.	.	.	.	.	.	.	.	.	.
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952

Giá trị trên bảng z là xác suất  $P(<Z)$

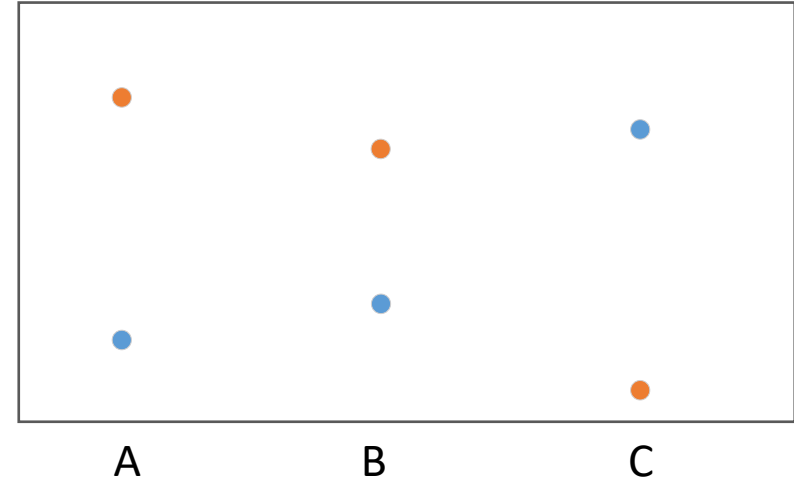
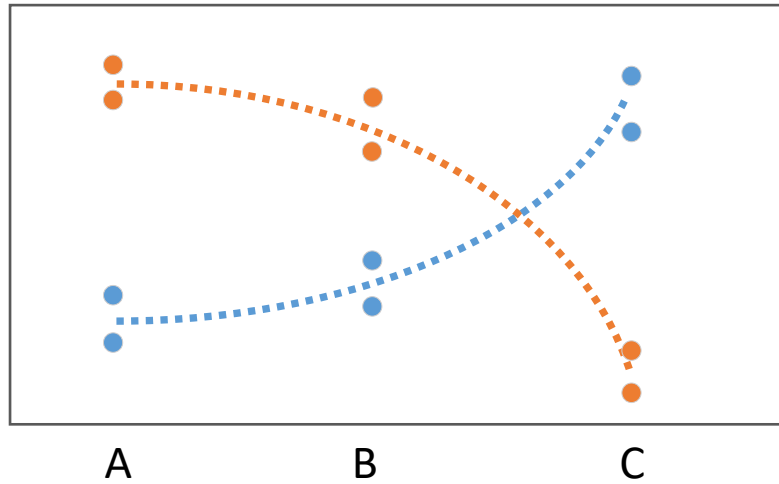


# Tính xác suất phân phối chuẩn dùng z-table

- Phân phối chuẩn với mean = 85.5h (từ bảng 6.1) và  $s = \sqrt{675} = 26$ . Ta có  $P(\text{time} > 50) = 91\%$
- $z = \frac{50 - 85.5}{26} = -1.37$ 
  - $P(\text{time} > 50) = P(Z > -1.37)$
  - $P(Z < -1.37) = 0.09 \rightarrow P(Z > -1.37) = 0.91$
- Tính  $P(\text{time} > 50)$  với mean = 58.9h và 26
- $z = \frac{50 - 58.9}{26} = -0.34$
- $P(Z < -0.34) = 0.37$
- $P(\text{time} > 50) = P(Z > -0.34) = 63\%$

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
-0.1	.46017	.45620	.45224	.44828	.44433	.44034	.43640	.43251	.42858	.42465
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-1	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592

# With vs without replication



- Khi không có replication, ta không thể phân biệt biến thiên gây ra bởi interaction hay do ngẫu nhiên
- Do đó khi không có replication, biến thiên gây ra bởi interaction (nếu có) được gộp chung với biến thiên ngẫu nhiên

# Two Factor ANOVA with vs without Replications

Without replication

$SS_T = \sum_j \sum_i (x_{ij} - \bar{x})^2$	$df_T = n - 1$	$MS_T = SS_T/df_T$
$SS_A = c \sum_i (\bar{x}_i - \bar{x})^2$	$df_A = r - 1$	$MS_A = SS_A/df_A$
$SS_B = r \sum_j (\bar{x}_j - \bar{x})^2$	$df_B = c - 1$	$MS_B = SS_B/df_B$
$SS_E = \sum_j \sum_i (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2$	$df_E = (r - 1)(c - 1)$	$MS_E = SS_E/df_E$

$$SS_T = SS_A + SS_B + SS_E$$

With replication

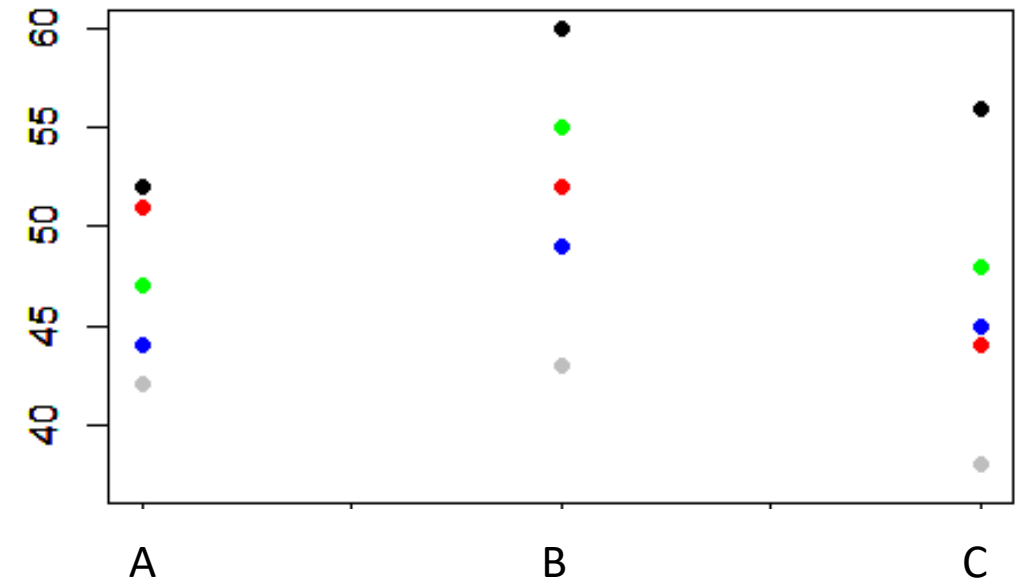
$SS_T = \sum_k \sum_j \sum_i (x_{ijk} - \bar{x})^2$	$df_T = n - 1$	$MS_T = SS_T/df_T$
$SS_A = mc \sum_i (\bar{x}_i - \bar{x})^2$	$df_A = r - 1$	$MS_A = SS_A/df_A$
$SS_B = mr \sum_j (\bar{x}_j - \bar{x})^2$	$df_B = c - 1$	$MS_B = SS_B/df_B$
$SS_{AB} = m \sum_j \sum_i (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2$	$df_{AB} = (r - 1)(c - 1)$	$MS_{AB} = SS_{AB}/df_{AB}$
$SS_W = \sum_k \sum_j \sum_i (x_{ijk} - \bar{x}_{ij})^2$	$df_W = n - rc$	$MS_W = SS_W/df_W$

$$SS_T = SS_A + SS_B + SS_{AB} + SS_W$$

## Example 2: Randomized Complete Block Design without replication

- Thí nghiệm 1 yếu tố với 3 treatment A, B, C.
- 15 đối tượng thí nghiệm không đồng nhất được chia thành 5 khối

Block	Treatment		
	A	B	C
I	52	60	56
II	47	55	48
II	44	49	45
IV	51	52	44
V	42	43	38

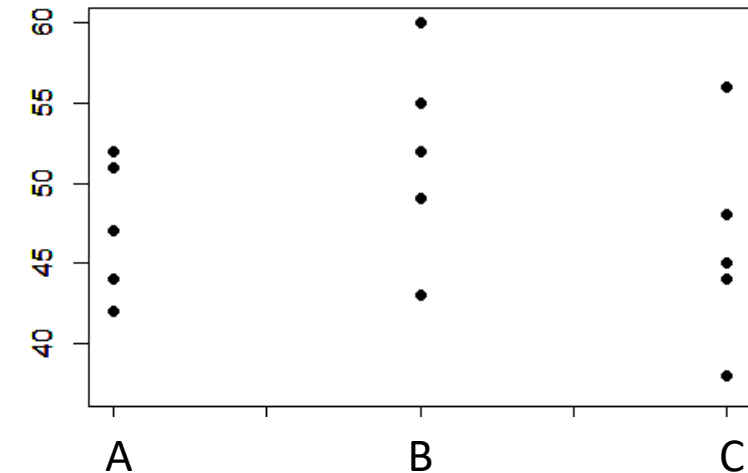


```
> my_data <- read.csv(file.choose(), head = TRUE, sep = ",")
```

```
> av <- aov(Result ~ Treatment + Block, data=my_data)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treatment	2	89.2	44.60	7.624	0.014023 *
Block	4	363.6	90.90	15.538	0.000768 ***
Residuals	8	46.8	5.85		

- Kết quả cho thấy có sự khác nhau giữa các treatment và có sự khác nhau giữa các block
- Nếu không tạo block, ta không biết tồn tại sự biến thiên bởi block và phân tích dữ liệu dùng oneway ANOVA. Khi đó, tất cả biến thiên bên trong mỗi treatment đều được xem là biến thiên ngẫu nhiên (lỗi)

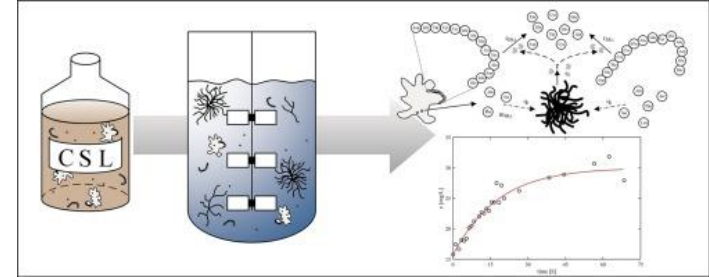


```
> av <- aov(Result ~ Treatment, data=my_data)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treatment	2	89.2	44.6	1.304	0.307
Residuals	12	410.4	34.2		

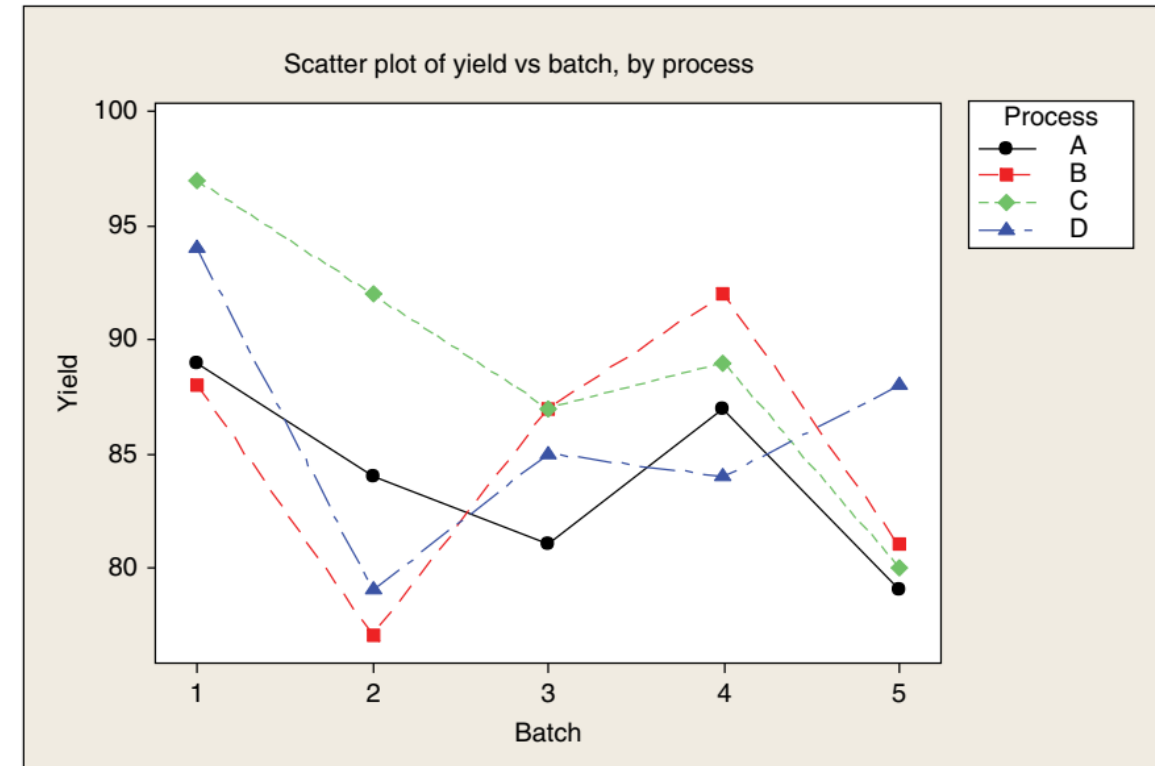
## Example 3: RCB design without replication (penicillin experiment)

- Thí nghiệm ảnh hưởng của 1 yếu tố chính lên sản lượng penicillin
  - Yếu tố chính: batch of corn steep liquor (CSL)
  - Yếu tố blocking: production process
- Mỗi lô CSL được chia thành 4 phần, mỗi phần được đem sản xuất penicillin dưới 1 quy trình khác nhau
- Thí nghiệm được thực hiện dựa theo một số kiến thức có từ trước
  - Sản lượng penicillin sẽ thay đổi theo từng lô CSL (do đó ta đặt lô CSL làm yếu tố chính để khảo sát)
  - Không có interaction giữa từng lô CSL và production process (**nên không cần replication**)
- Dữ liệu thí nghiệm sẽ được phân tích để kiểm tra 2 giả thuyết
  - Có sự khác nhau giữa các lô CSL hay không
  - Có sự khác nhau giữa các production process hay không



## Example 3: RCB design without replication (penicillin experiment)

- Nhìn vào biểu đồ, ta có thể dự đoán rằng
  - Có sự khác nhau giữa các CSL batch (ví dụ batch 1 cho kết quả tốt hơn batch 5 khá rõ)
  - Không có sự khác nhau giữa các process (ví dụ process C cho kết quả cao với 2 lô đầu, nhưng kết quả kém ở 3 lô sau, còn process B kém ở 2 lô đầu, nhưng lại tốt ở 3 lô sau)



# Two way ANOVA without replication

- Từ bảng ANOVA

- Có sự khác nhau ( $\alpha = 0.05$ ) giữa các lô CSL
- Không có sự khác nhau giữa các process

**Table 6.6** Two-Way ANOVA: Yield versus Batch, Process.

<i>Source</i>	<i>DF</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>
Batch	4	264	66.0	3.50	.04
Process	3	70	23.3	1.24	.34
Error	12	226	18.8		
Total	19	560			

- Do không có sự khác nhau giữa các process, các biến thiên của kết quả bên trong mỗi batch đều là biến thiên ngẫu nhiên (Error). Tuy nhiên, khi sử dụng Two-way ANOVA, ta đã tách biến thiên ngẫu nhiên (Error) khỏi biến thiên bởi process (vốn không tồn tại).
- Để có phân tích chính xác hơn, ta cần phân tích kết quả dùng One-way ANOVA. Phân tích này xem tất cả biến thiên bên trong mỗi batch đều là biến thiên ngẫu nhiên (Error)



# Phân tích lại dùng One-way ANOVA

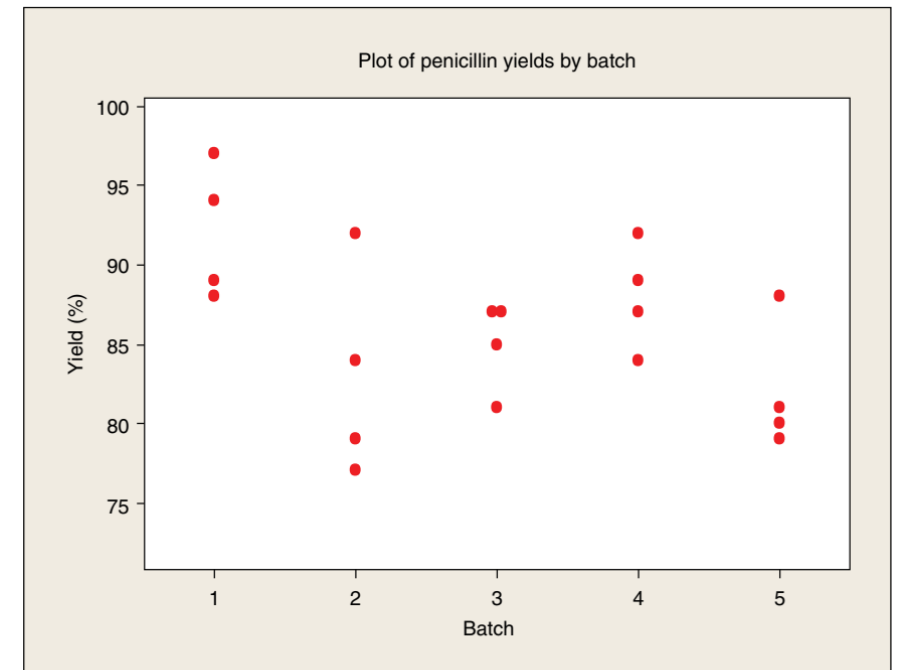
- Với One-way ANOVA, ta không phân biệt kết quả của từng process và xem tất cả biến thiên bên trong mỗi batch đều là biến thiên ngẫu nhiên
- SSE ở One-way ANOVA sẽ là tổng của SSE và  $SS_{Batch}$  ở bảng Two-way ANOVA

**Table 6.6** Two-Way ANOVA: Yield versus Batch, Process.

Source	DF	SS	MS	F	P
Batch	4	264	66.0	3.50	.04
Process	3	70	23.3	1.24	.34
Error	12	226	18.8		
Total	19	560			

**Table 6.7** One-Way ANOVA: Penicillin Yield versus Batch.

Source	DF	SS	MS	F	P
Batch	4	264	66.0	3.34	.04
Error	15	296	19.7		
Total	19	560			



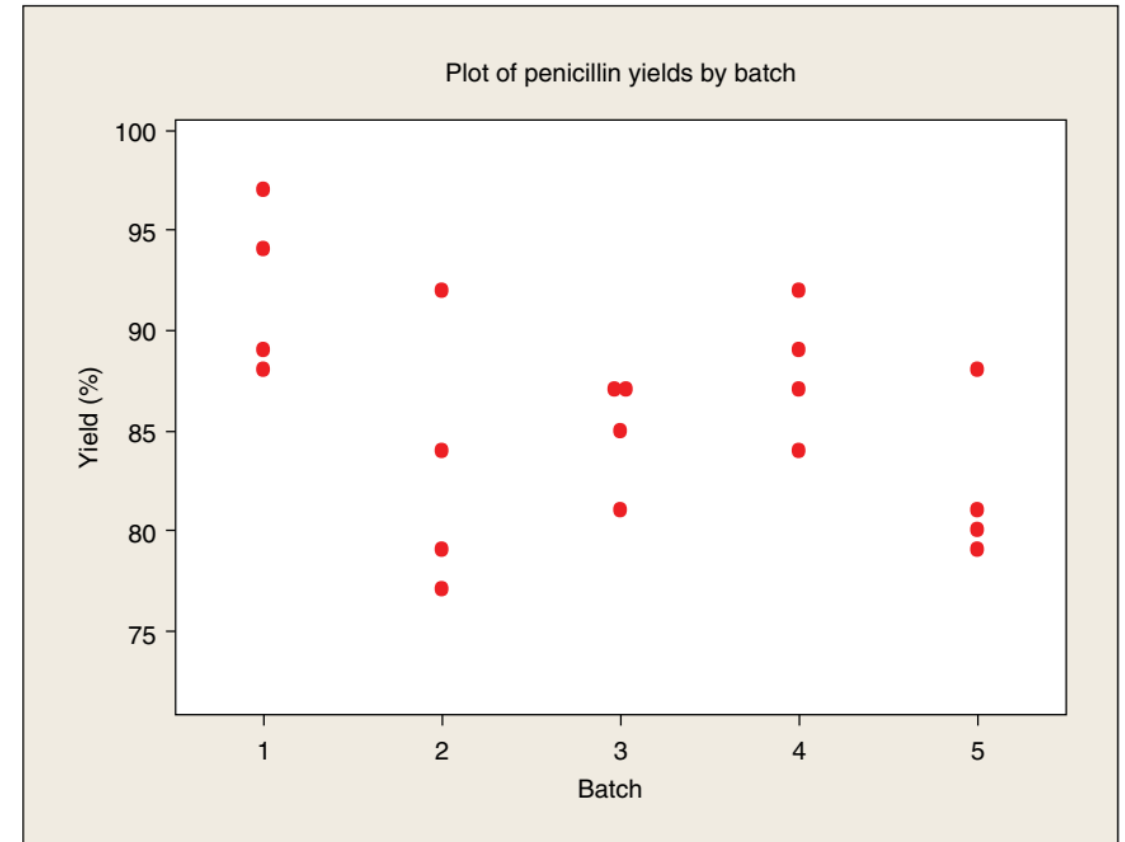
**Figure 6.5** Penicillin Yields by Batch.

# Components of variation

**Table 6.7** One-Way ANOVA: Penicillin Yield versus Batch.

Source	DF	SS	MS	F	P
Batch	4	264	66.0	3.34	.04
Error	15	296	19.7		
Total	19	560			

- Ước lượng effect của từng lô CSL thế nào?



**Figure 6.5** Penicillin Yields by Batch.

# Components of variation

Source	df	SS	MS=SS/df	F
Treatments	k-1	$SSA = \sum n(y_{\bar{i}} - y_{\bar{..}})^2$	$MSA = SSA/df$	$MSA/MSE$
Error	k(n-1)	$SSE = \sum \sum (y_{ji} - y_{\bar{.i}})^2$	$MSE = SSE/df$	
Total	n*k-1	$SS = \sum \sum (y_{ji} - y_{\bar{..}})^2$		

**Table 6.7** One-Way ANOVA: Penicillin Yield versus Batch.

Source	DF	SS	MS	F	P
Batch	4	264	66.0	3.34	.04
Error	15	296	19.7		
Total	19	560			

- Sản lượng penicillin trong thí nghiệm sẽ theo mô hình

$$y_{ij} = \mu + b_i + e_{ij}$$

- $b_i$  là ảnh hưởng **ngẫu nhiên** của mỗi lô CSL.
- Biến ngẫu nhiên  $b_i$  có variance  $\sigma_b^2$
- $e_{ij}$  là ảnh hưởng **ngẫu nhiên** của từng sub-batch của mỗi lô.
- Biến ngẫu nhiên  $e_{ij}$  có variance  $\sigma_e^2$

- Ta có

- $E(MSE) = \sigma_e^2$
- $E(MSA) = \sigma_e^2 + n * \sigma_b^2$

Suy ra:  $\sigma_e = \sqrt{19.7} = 4.4$

Estimated  $\sigma_b^2 = \frac{MSA - MSE}{4} = \frac{66.0 - 19.7}{4} = 11.6.$

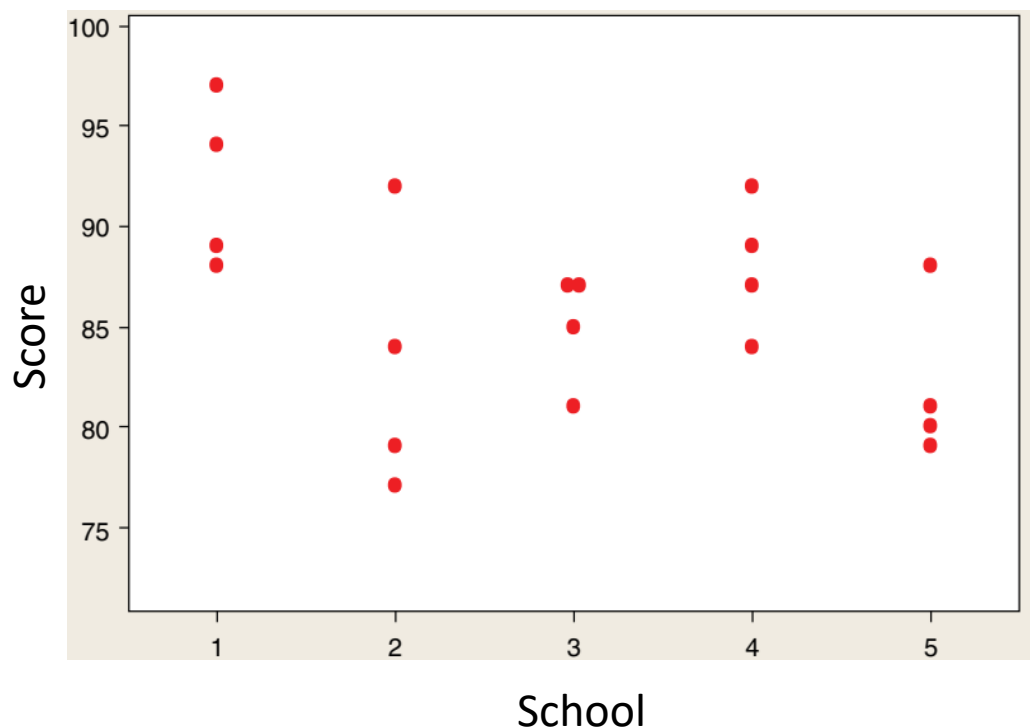
# Random effect model

- Chọn ngẫu nhiên k trường học trong cả nước để kiểm tra trình độ tiếng Anh.
- Chọn ngẫu nhiên n học sinh (cùng độ tuổi) ở mỗi trường để làm bài kiểm tra.

- Điểm số của mỗi học sinh trong thí nghiệm sẽ theo mô hình

$$y_{ij} = \mu + b_i + e_{ij}$$

- $b_i$  là ảnh hưởng **ngẫu nhiên** của mỗi trường.
- $e_{ij}$  là ảnh hưởng **ngẫu nhiên** của từng học sinh ở mỗi trường
- $b_i$  có variance  $\sigma_b^2$
- $e_{ij}$  có variance  $\sigma_e^2$
- Ta có
  - $E(MSW) = \sigma_e^2$
  - $E(MSB) = \sigma_e^2 + n * \sigma_b^2$



# True replication

## Heat chamber



- Xét 2 quy trình làm thí nghiệm
  - Với mỗi loại pin, chỉnh buồng nhiệt ở mức 15°F. Cho đồng thời 4 viên pin vào buồng nhiệt và đo thời gian sử dụng của từng viên pin. Sau đó chỉnh buồng nhiệt sang mức 70, 125°F và lặp lại quy trình trên với từng lô 4 viên pin.
  - Với mỗi loại pin, thiết lập buồng nhiệt tại 15°F. Cho 1 viên pin vào buồng nhiệt và đo thời gian sử dụng. Sau đó, cho tiếp lần lượt 3 viên pin nữa vào buồng nhiệt và lần lượt đo thời gian sử dụng từng viên pin. Sau đó chỉnh nhiệt độ sang 70, 125°F và lặp lại quy trình trên.
- Đơn vị thí nghiệm (experiment unit) được định nghĩa là đối tượng nhận treatment một cách độc lập. Với quy trình đầu tiên, cả 4 viên pin được đưa vào buồng nhiệt cùng lúc, tức là cả 4 viên pin đồng thời nhận treatment. Do đó experiment unit ở đây là 4 viên pin (chứ không phải từng viên pin).
- Như vậy ở quy trình đầu tiên, mỗi treatment (là một kết hợp của treatment factor và blocking factor) chỉ có 1 replication. Response của mỗi replication sẽ là kết quả trung bình của 4 viên pin.
- Do đó bảng ANOVA cho dữ liệu thu được từ quy trình đầu tiên sẽ không có interaction

**Table 6.1** Battery Lifetimes by Material and Temperature.

Material	Temperature (°F)		
	15	70	125
1	134.75	57.25	57.5
2	155.75	119.75	49.5
3	144	145.75	85.5

Table 5.1, p. 176 of Montgomery (2001), reproduced by permission of John Wiley & Sons.

**Table 6.8** ANOVA for Modified Battery Experiment.

Source	DF	SS	MS	F	P
Material	2	2626	1313	2.24	.22
Temp	2	9784	4895	8.34	.04
Error	4	2348	587		
Total	8	5953			

# Efficiency of RCBD vs. CRD

- Hiệu quả của việc chia khối được tính như sau:

$$Eff = \frac{(r - 1)MS_{Block} + r(c - 1)MSE}{(rc - 1)MSE}$$

r là số block  
c là số treatment

- Ví dụ

Source	DF	SS	MS	F	P
Batch	4	264	66.0	3.50	.04
Process	3	70	23.3	1.24	.34
Error	12	226	18.8		

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Treatment	2	89.2	44.60	7.624	0.014023	*
Block	4	363.6	90.90	15.538	0.000768	***
Residuals	8	46.8	5.85			

$$Eff = \frac{(4 - 1)23.3 + 4(5 - 1)18.8}{(4 * 5 - 1)18.8} = 1.04$$

$$Eff = \frac{(5 - 1)90.9 + 5(3 - 1)5.85}{(5 * 3 - 1)5.85} = 5.15$$

# RCBD with **one** missing data element

- Bước 1: thay thế số liệu bị mất bởi giá trị

$$x_{ij} = \frac{r\Sigma_i + c\Sigma_j - \Sigma}{(r-1)(c-1)} \text{ trong đó } \Sigma_i \text{ là tổng dòng } i \text{ và } \Sigma_j \text{ là tổng cột } j$$

- Ví dụ

		Herbicide dosage									
3		5	10	15	20	25	30				
4	Field 1	9.7		8.3	6.3	6.7	5.8	36.8	r	4	
5	Field 2	13.0	20.2	6.4	6.0	5.2	3.6	54.4	c	6	
6	Field 3	15.6	17.7	6.1	4.0	3.1	4.1	50.6			
7	Field 4	7.1	11.8	1.7	0.5	0.8	0.2	22.1			
8		45.4	49.7	22.5	16.8	15.8	13.7	163.9			

$$x_{12} = \frac{4(36.8) + 6(49.7) - 163.9}{(4-1)(6-1)} = 18.77$$

- Giá trị thay thế sẽ đóng góp tối thiểu vào sai số ngẫu nhiên MSE

# RCBD with **one** missing data element

- Bước 2: tính bảng ANOVA với bảng số liệu mới (đã cập nhật số liệu bị mất bằng số thay thế)

- Bước 3: cập nhật lại bảng ANOVA

- $df_{total}^* = df_{total} - 1$

- $df_{Error}^* = df_{Error} - 1$

- $SS_{Block}^* = SS_{Block} - \frac{(\sum_j - (r-1)x_{ij})^2}{r(r-1)}$

- $SS_{Treatment}^* = SS_{Treatment} - \frac{(\sum_i - (c-1)x_{ij})^2}{c(c-1)}$

- Các giá trị có liên quan được tính lại

	AD	AE	AF	AG	AH	AI	AJ
3	ANOVA				Alpha	0.05	
4	Sources	SS	df	MS	F	P value	F crit
5	Blocks	125.6678	3	41.88926	16.62101	4.92E-05	3.287382
6	Groups	603.3698	5	120.674	47.88157	1.08E-08	2.901295
7	Error	37.80389	15	2.520259			
8	Total	766.8415	23				
9							
10	Sources	SS					
11	Blocks	3.63					
12	Groups	108.4267					
13							
14	ANOVA				Alpha	0.05	
15	Sources	SS	df	MS	F	P value	F crit
16	Blocks	122.0378	3	40.67926	15.06484	0.000116	3.343889
17	Groups	494.9431	5	98.98862	36.65868	1.47E-07	2.958249
18	Error	37.80389	14	2.700278			
19	Total	654.7848	22				



# Blocking in CO emission experiment

- Thí nghiệm ảnh hưởng của 2 yếu tố đến lượng khí thải CO
  - Ethanol concentration (x1) : 0.1, 0.2, 0.3
  - Air-fuel ratio (x2): 14, 15, 16
- Có tổng cộng 9 treatment. Có 18 experiment unit (?). Có 2 replication cho mỗi treatment.
- Giả sử thí nghiệm được thực hiện trong 2 ngày. Ngày có thể là yếu tố blocking. Ở ngày thứ nhất thực hiện thí nghiệm với 9 treatment theo thứ tự ngẫu nhiên. Ngày thứ 2 thực hiện thí nghiệm với 9 treatment theo thứ tự ngẫu nhiên.
- Thí nghiệm có thể được thực hiện với 2 động cơ. Động cơ có thể là yếu tố blocking. Mỗi động cơ sẽ được thử với 9 treatment.

CO emissions		Air-fuel ratio		
		14	15	16
Ethanol concentration	0.1	61.8	67.4	66.2
		65.8	72	68.4
	0.2	78.5	80	65.5
		80.5	81.6	69.5
	0.3	89.9	74.6	58
		93.7	78	60

# Latin Square design

- Dùng để block 2 yếu tố
- Các treatment (đơn lẻ hoặc kết hợp) được bố trí theo 2 hướng: hàng và cột
- Mỗi treatment chỉ xuất hiện 1 lần trong hàng và cột do đó giảm được nguồn sai số do bias của yếu tố hàng và cột
- Ví dụ thí nghiệm 4 loại phân bón A, B, C, D. 4 loại phân này sẽ được trồng thử 16 lượt. Mỗi lượt trồng tại 1 ô đất và 1 giống lúa. Loại phân bón là yếu chính, loại lúa và loại đất là yếu tố blocking. Nếu không tạo block, phân loại A có thể có kết quả tốt hơn chỉ vì giống lúa 4 năng suất cao hơn, hoặc phân loại D có kết quả tốt hơn vì đất thịt cho năng suất cao hơn.

	Lúa 1	Lúa 2	Lúa 3	Lúa 4
Đất sỏi	1 - B	2 - C	3 - B	4 - A
Đất cát	5 - C	6 - C	7 - B	8 - A
Đất phèn	9 - B	10 - A	11 - C	12 - A
Đất thịt	13 - D	14 - D	15 - D	16 - D

CRD (không blocking)

	Lúa 1	Lúa 2	Lúa 3	Lúa 4
Đất sỏi	1 - A	2 - C	3 - B	4 - D
Đất cát	5 - A	6 - D	7 - B	8 - C
Đất phèn	9 - A	10 - B	11 - D	12 - C
Đất thịt	13 - A	14 - D	15 - B	16 - C

RCBD (block loại đất)

	Lúa 1	Lúa 2	Lúa 3	Lúa 4
Đất sỏi	1 - A	2 - C	3 - B	4 - D
Đất cát	5 - B	6 - D	7 - A	8 - C
Đất phèn	9 - C	10 - B	11 - D	12 - A
Đất thịt	13 - D	14 - A	15 - C	16 - B

Latin square (block loại đất và giống lúa)

# Latin Square design

- Thí nghiệm thời gian khô của 4 loại vữa A, B, C, D (tức là có 4 treatment).
  - 4 loại vữa được thử nghiệm qua 16 lần xây thử nghiệm.
  - Việc xây được thực hiện trong 4 ngày
  - Mỗi lần xây thử nghiệm được thực hiện vào 1 buổi trong ngày (sáng, trưa, chiều, tối).
  - 2 yếu tố blocking là ngày và buổi có thể ảnh hưởng kết quả do độ ẩm thay đổi theo từng ngày và từng buổi trong ngày.

	Ngày 1	Ngày 2	Ngày 3	Ngày 4
Sáng	1 - A	2 - C	3 - B	4 - D
Trưa	5 - A	6 - D	7 - B	8 - C
Chiều	9 - A	10 - B	11 - D	12 - C
Tối	13 - A	14 - D	15 - B	16 - C

RCBD (chỉ block buổi)

	Ngày 1	Ngày 2	Ngày 3	Ngày 4
Sáng	1 - A	2 - C	3 - B	4 - D
Trưa	5 - B	6 - D	7 - A	8 - C
Chiều	9 - C	10 - B	11 - D	12 - A
Tối	13 - D	14 - A	15 - C	16 - B

Latin square

# Latin square design: Gasoline additives and car emissions

- Thí nghiệm ảnh hưởng của 4 loại chất phụ gia Control, A1, A2, A3
- Có 4 tài xế và 4 chiếc xe
- Thực hiện 16 lượt chạy thử.
- Tài xế và xe là 2 yếu tố blocking vì có khả năng ảnh hưởng kết quả
- Với thiết kế Latin square, mỗi chất phụ gia sẽ được thử nghiệm với cả 4 tài xế trên 4 chiếc xe khác nhau. Do đó khi so sánh kết quả trung bình của 2 chất phụ gia thì ảnh hưởng của tài xế và xe nếu có sẽ được cân bằng.
- Ở đây ta giả sử không có interaction giữa các yếu tố

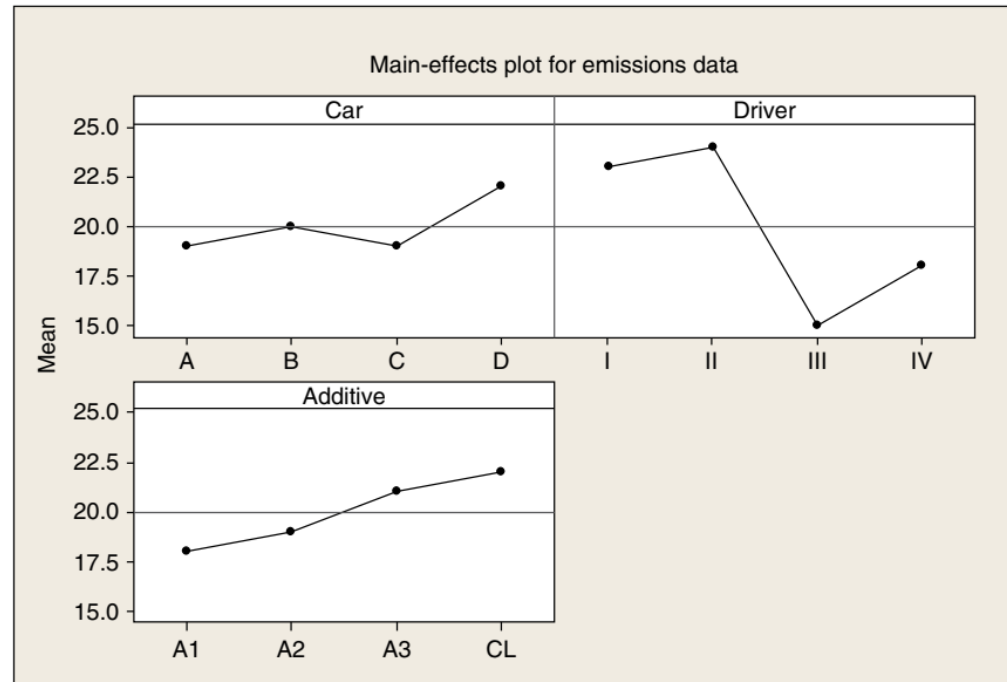
**Table 7.1** 4 × 4 Latin Square Design for Additive Experiment.<sup>a</sup>

Driver	Car			
	A	B	C	D
I	A1	A2	A4	A3
II	A4	A3	A1	A2
III	A2	A4	A3	A1
IV	A3	A1	A2	A4

<sup>a</sup>Cell entries denote the additive assigned to a given car/driver combination.

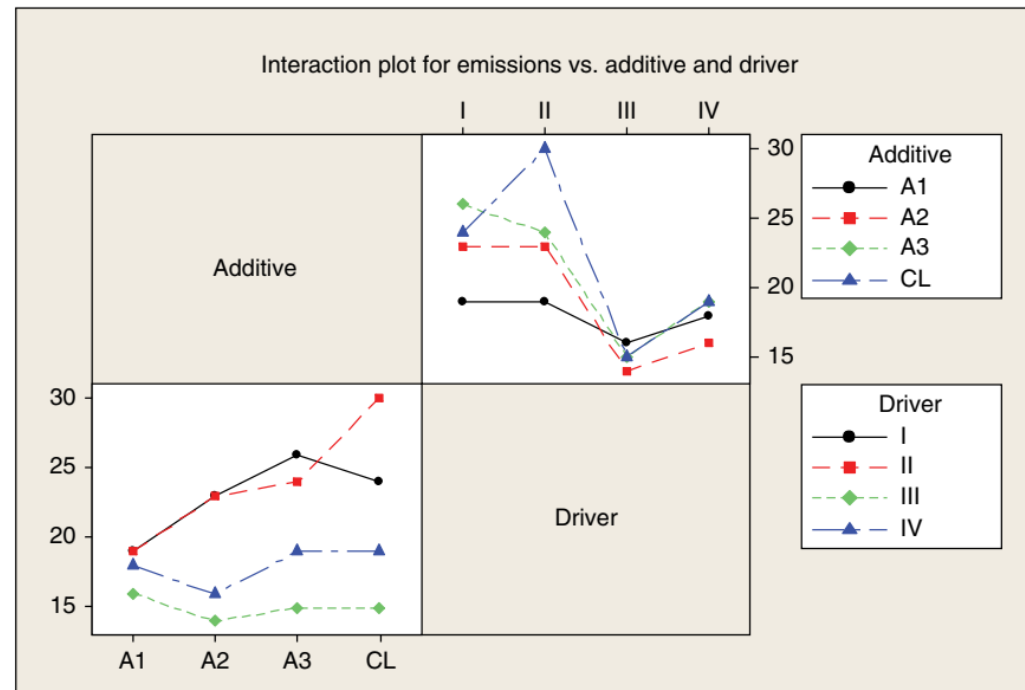
**Table 7.2** Results of Car Emission Latin Square Experiment.<sup>a</sup>

Driver	Car			
	A	B	C	D
I	A1	CL	A2	A3
	19	24	23	26
II	A2	A3	A1	CL
	23	24	19	30
III	CL	A2	A3	A1
	15	14	16	16
IV	A3	A1	CL	A2
	19	18	19	16



**Figure 7.1** Average Emissions by Car, Driver, and Additive.

- Nhận xét
  - Do không xét được interaction nên ta có thể đi thẳng vào main-effect plot
  - Có sự khác nhau rõ rệt giữa các driver (III và IV cho ít khí thải hơn I và II)
  - Các chất phụ gia A1, A2, A3 có kết quả trung bình thấp hơn (tốt hơn) control. Nhưng ta cần xem kết quả này có ý nghĩa hay không.



**Figure 7.2** Interaction Plot of Emissions Data by Additive and Driver.

- Ở biểu đồ emission vs additive, dường như không có sự khác biệt giữa các chất phụ gia ở hai tài xế III và IV, nhưng lại có sự khác biệt rõ rệt giữa các chất phụ gia ở hai tài xế I và II. Có thể đây là kết quả của interaction giữa tài xế và chất phụ gia nhưng ta không thể kiểm tra được (nên chỉ có thể xem như biến thiên ngẫu nhiên).

**Table 7.3** ANOVA for Emissions Experiment.

<i>Source</i>	<i>DF</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>
Car	3	24	8.0	1.5	.31
Driver	3	216	72.0	13.5	.004
Additive	3	40	13.3	2.5	.16
Error	6	32	5.3		
Total	15	312			

(Rõ ràng không có khác biệt giữa car)

**Table 7.4** Reduced ANOVA of Emissions Data.

<i>Source</i>	<i>DF</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>
Driver	3	216	72.0	11.6	.002
Additive	3	40	13.3	2.14	.17
Error	9	56	6.2		
Total	15	312			

# Split-unit (or Split-plot) design

- Split-unit (hay split-plot) là 1 trường hợp đặc biệt của factorial design
- Có 2 hoặc nhiều factor cần được khảo sát
- Một số factor khó (hoặc tốn nhiều tiền) để thay đổi các mức. Các factor này sẽ được áp dụng cho những đơn vị thí nghiệm kích thước lớn và kết quả là số lượng đơn vị thí nghiệm cho factor này sẽ không nhiều. Mỗi đơn vị thí nghiệm cho các factor này gọi là 1 plot.
- Những factor dễ thay đổi các mức hơn sẽ được áp dụng ngẫu nhiên bên trong các đơn vị thí nghiệm (unit/plot) lớn. Tức là các đơn vị thí nghiệm (unit/plot) lớn sẽ được chia nhỏ ra để tạo thành các đơn vị thí nghiệm nhỏ hơn cho các factor dễ thay đổi.
- Như vậy split-unit/split-plot là thí nghiệm gồm nhiều loại đơn vị thí nghiệm với kích thước khác nhau dành cho các factor khác nhau. Đơn vị thí nghiệm nhỏ sẽ nằm trong đơn vị thí nghiệm lớn.



# Ví dụ thí nghiệm split-plot

- Thí nghiệm 2 yếu tố

- Có 3 mức tưới nước
- Có 4 giống bắp
- Có 6 lô đất



4	4	1	3	2	2
2	1	4	1	4	1
3	3	2	2	1	4
1	2	3	4	3	3

- Rõ ràng yếu tố nước tưới không thể áp dụng **độc lập** trên những phần đất quá nhỏ. Ở đây ta giả sử mức nước tưới có thể áp dụng độc lập (và ngẫu nhiên) trên từng lô đất. Nên mỗi lô đất là 1 đơn vị thí nghiệm cho yếu tố này.
- Với yếu tố giống bắp ta có thể áp dụng độc lập trên những phần đất nhỏ hơn. Do đó trên mỗi lô đất (1 đơn vị thí nghiệm của yếu tố nước tưới) ta chia thành nhiều phần (mỗi phần là một đơn vị thí nghiệm nhỏ hơn) để áp dụng độc lập (và ngẫu nhiên) các giống bắp.
- 2 quy trình gán ngẫu nhiên các treatment được thực hiện một cách độc lập
- Yếu tố nước tưới gọi là whole-plot factor, yếu tố giống bắp gọi là split-plot factor

## Ví dụ thí nghiệm split-plot

- Thí nghiệm độ chống ăn mòn của kim loại khi mạ
  - Yếu tố Coating có 4 mức
  - Yếu tố Heat có 3 mức
  - Có 24 thanh kim loại để làm thí nghiệm

**Table 7.6** Corrosion-Resistance Experiment Data.

Temp. (°F)	Heat	Coating			
		1	2	3	4
360	1	67	73	83	89
	6	33	8	46	54
370	2	65	91	87	86
	5	140	142	121	150
380	3	155	127	147	212
	4	108	100	90	153

- Áp dụng Heat độc lập cho từng thanh kim loại sẽ rất tốn kém (mỗi lần áp dụng cần cả quy trình nung lò và để nguội). Do đó để tiết kiệm chi phí thì 4 thanh kim loại sẽ được nung cùng lúc tại 1 nhiệt độ. Tức là experiment unit cho yếu tố Heat là từng nhóm 4 thanh kim loại. Như vậy sẽ có 6 experiment unit được áp dụng ngẫu nhiên 3 mức nhiệt độ.
- Trong từng nhóm 4 thanh kim loại ta áp dụng ngẫu nhiên 4 loại coating. Như vậy mỗi thanh kim loại là một experiment unit cho yếu tố coating.
- Như vậy heat là whole-plot factor và coating là split-plot factor. Thí nghiệm có 2 experiment unit (eu): nhóm 4 thanh kim loại là eu cho yếu tố heat, từng thanh kim loại là eu cho yếu tố coating.

**Table 7.7** ANOVA Structure for Corrosion-Resistance Experiment.

<i>Source</i>	<i>df</i>	<i>Main unit</i>
Temp	2	
Error1	3	
-----	-----	-----
Coating	3	Subunit
Temp × Coating	6	
Error2	9	
Total	23	

**Table 7.8** ANOVA for Corrosion-Resistance Split-Unit Experiment.

<i>Source</i>	<i>DF</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>
Temp	2	26 519	13 260	2.8	.21
Error1	3	14 440	4813	38.7	.00
-----	-----	-----	-----	-----	-----
Coating	3	4289	1430	11.5	.002
Temp*coating	6	3270	545	4.4	.024
Error2	9	1121	125		
Total	23	49 639			