ISE 315: Engineering Statistics

Lecture 26: Factorial Diagnostics and Applications

Instructor: Mansur M. Arief, PhD
Industrial and Systems Engineering, KFUPM

Office: 22-219 — Email: mansur.arief@kfupm.edu.sa

Based on Montgomery & Runger, Applied Statistics and Probability for Engineers, 6th Ed.

Lecture 26

Factorial Diagnostics and Applications (Chapter 9, §9.8–§9.13)

Lecture 26 Outline

- Residual analysis and model adequacy (§9.8, Procedure 9.2)

Lecture 26 Outline

- Residual analysis and model adequacy (§9.8, Procedure 9.2)
- One observation per cell (§9.9)

Lecture 26 Outline

- Residual analysis and model adequacy (§9.8, Procedure 9.2)
- One observation per cell (§9.9)
- Three or more factors: the general factorial (§9.10)

Lecture 26 Outline

- Residual analysis and model adequacy (§9.8, Procedure 9.2)
- One observation per cell (§9.9)
- Three or more factors: the general factorial (§9.10)
- The 2^k workhorse: 2^2 and the sign table (§9.11, Procedure 9.3)

Lecture 26 Outline

- Residual analysis and model adequacy (§9.8, Procedure 9.2)
- One observation per cell (§9.9)
- Three or more factors: the general factorial (§9.10)
- The 2^k workhorse: 2^2 and the sign table (§9.11, Procedure 9.3)
- Single replicates, center points, blocking, fractions (§9.12)

Lecture 26 Outline

- Residual analysis and model adequacy (§9.8, Procedure 9.2)
- One observation per cell (§9.9)
- Three or more factors: the general factorial (§9.10)
- The 2^k workhorse: 2^2 and the sign table (§9.11, Procedure 9.3)
- Single replicates, center points, blocking, fractions (§9.12)
- Factorial thinking in engineering practice (§9.13)

Residuals in a Factorial Model

Section 9.8

Every conclusion from Lecture 25 rests on the model (9.1): independent normal errors with constant variance σ^2 in every cell. **Before trusting the F tests, check the residuals.**

Residuals in a Factorial Model

Section 9.8

Every conclusion from Lecture 25 rests on the model (9.1): independent normal errors with constant variance σ^2 in every cell. **Before trusting the F tests, check the residuals.**

The fitted value for any observation in a factorial model is its *cell mean*, so the residual is

$$e_{ijk} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \bar{y}_{ij\cdot} \tag{9.12}$$

Residuals in a Factorial Model

Section 9.8

Every conclusion from Lecture 25 rests on the model (9.1): independent normal errors with constant variance σ^2 in every cell. **Before trusting the F tests, check the residuals.**

The fitted value for any observation in a factorial model is its *cell mean*, so the residual is

$$e_{ijk} \;=\; y_{ijk} - \hat{y}_{ijk} \;=\; y_{ijk} - \bar{y}_{ij\cdot} \tag{9.12}$$

Built-in check: the residuals must sum to zero *within every cell*. Use this as a free arithmetic check on your cell means.

Residuals in a Factorial Model

Section 9.8

Every conclusion from Lecture 25 rests on the model (9.1): independent normal errors with constant variance σ^2 in every cell. **Before trusting the F tests, check the residuals.**

The fitted value for any observation in a factorial model is its *cell mean*, so the residual is

$$e_{ijk} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \bar{y}_{ij\cdot} \tag{9.12}$$

Built-in check: the residuals must sum to zero *within every cell*. Use this as a free arithmetic check on your cell means.

What we look for: normality of the errors, constant spread across fitted values and across factor levels, and isolated outliers. The formal recipe is Procedure 9.2, next slide.

Textbook §9.8: Chapter 9 ((9.12) onward)

Procedure 9.2: Diagnostics for a Factorial Experiment

1. Compute every fitted value ($\bar{y}_{ij\cdot}$) and residual e_{ijk} by (9.12). Residuals must sum to zero within every cell.
2. **Normal probability plot** of all *abn* residuals: points near a straight line support normality; curvature or heavy tails do not.
3. **Residuals vs. fitted values**: spread should be roughly constant. A funnel widening with the fitted value $\Rightarrow$ non-constant variance; consider transforming ($\log y$ or $\sqrt{y}$) and re-analyzing.
4. **Residuals vs. each factor**: similar spread at every level. One level with visibly larger scatter $\Rightarrow$ that factor also affects the *variance*.
5. Investigate any isolated large residual — check the data record before deleting anything.
6. Read the interaction plot alongside the ANOVA: parallel lines with significant F_0^{AB}, or wildly non-parallel lines with an insignificant one, usually signal an underpowered experiment.

Residual Check: the Picking Study from Lecture 25

Example 9.5

Recall (supply-chain study, Table 9.6): picking time vs. layout $\times$ shift, 3×3, $n = 4$; all three effects significant. Do the assumptions hold?

Residual Check: the Picking Study from Lecture 25

Example 9.5

Recall (supply-chain study, Table 9.6): picking time vs. layout $\times$ shift, 3×3, $n = 4$; all three effects significant. Do the assumptions hold?

Step 1 — Residuals by (9.12). For the layout-1 morning cell, $\bar{y}_{11\cdot} = 134.75$, and the observations were 130, 155, 74, 180:

$$e_{111} = 130 - 134.75 = -4.75, \qquad e_{112} = 155 - 134.75 = 20.25,$$

$$e_{113} = 74 - 134.75 = -60.75, \qquad e_{114} = 180 - 134.75 = 45.25$$

Residual Check: the Picking Study from Lecture 25

Example 9.5

Recall (supply-chain study, Table 9.6): picking time vs. layout $\times$ shift, 3×3, $n = 4$; all three effects significant. Do the assumptions hold?

Step 1 — Residuals by (9.12). For the layout-1 morning cell, $\bar{y}_{11\cdot} = 134.75$, and the observations were 130, 155, 74, 180:

$$e_{111} = 130 - 134.75 = -4.75, \qquad e_{112} = 155 - 134.75 = 20.25,$$

$$e_{113} = 74 - 134.75 = -60.75, \qquad e_{114} = 180 - 134.75 = 45.25$$

Check: $-4.75 + 20.25 - 60.75 + 45.25 = 0.$ ✓

Residual Check: the Picking Study from Lecture 25

Example 9.5

Recall (supply-chain study, Table 9.6): picking time vs. layout $\times$ shift, 3×3, $n=4$; all three effects significant. Do the assumptions hold?

Step 1 — Residuals by (9.12). For the layout-1 morning cell, $\bar{y}_{11\cdot} = 134.75$, and the observations were 130, 155, 74, 180:

$$e_{111} = 130 - 134.75 = -4.75, \qquad e_{112} = 155 - 134.75 = 20.25,$$

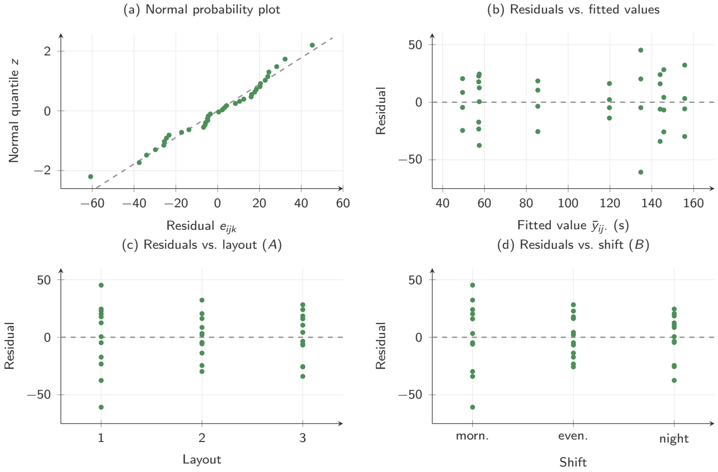
$$e_{113} = 74 - 134.75 = -60.75, \qquad e_{114} = 180 - 134.75 = 45.25$$

Check: $-4.75 + 20.25 - 60.75 + 45.25 = 0.$ ✓

Repeating for all nine cells gives 36 residuals, ranging from -60.75 to 45.25 — both extremes from the same layout-1 morning cell. Now the four plots.

The Four Diagnostic Panels

Figure 9.6, steps 2–4 of Procedure 9.2



Reading the Panels, and the Red Flags

Steps 2–6 of Procedure 9.2

This study: (a) the normal plot is acceptably straight — the two largest residuals sit at the ends but not far off the line. (b) No funnel against fitted values. (c), (d) Spread is broadly similar across levels, mildly larger for layout 1 and the morning shift — both driven by the same noisy cell.

Reading the Panels, and the Red Flags

Steps 2–6 of Procedure 9.2

This study: (a) the normal plot is acceptably straight — the two largest residuals sit at the ends but not far off the line. (b) No funnel against fitted values. (c), (d) Spread is broadly similar across levels, mildly larger for layout 1 and the morning shift — both driven by the same noisy cell.

Conclusion: assumptions adequately met; the ANOVA conclusions stand. No transformation needed. The layout-1 morning cell (range 74 to 180 s) is the first place to look if the study is repeated.

Reading the Panels, and the Red Flags

Steps 2–6 of Procedure 9.2

This study: (a) the normal plot is acceptably straight — the two largest residuals sit at the ends but not far off the line. (b) No funnel against fitted values. (c), (d) Spread is broadly similar across levels, mildly larger for layout 1 and the morning shift — both driven by the same noisy cell.

Conclusion: assumptions adequately met; the ANOVA conclusions stand. No transformation needed. The layout-1 morning cell (range 74 to 180 s) is the first place to look if the study is repeated.

Red flags to memorize:

- Funnel in (b) $\Rightarrow$ non-constant variance $\Rightarrow$ transform ($\log y$, $\sqrt{y}$), re-analyze.
- Curvature or heavy tails in (a) $\Rightarrow$ non-normal errors; outliers $\Rightarrow$ investigate, don't delete blindly.
- One factor level with much larger scatter in (c)/(d) $\Rightarrow$ the factor affects the variance too.

One Observation per Cell

Section 9.9

The problem. Error df $= ab(n-1)$ come entirely from replication. With $n = 1$: $df_E = 0$, every "cell mean" is the observation itself, every residual in (9.12) is zero, and SS_{AB} and SS_E collapse into one quantity. No F test of the interaction.

One Observation per Cell

Section 9.9

The problem. Error df $= ab(n-1)$ come entirely from replication. With $n=1$: $df_E = 0$, every "cell mean" is the observation itself, every residual in (9.12) is zero, and SS_{AB} and SS_E collapse into one quantity. No F test of the interaction.

The workaround. *Assume* the interaction away: if $(\tau\beta)_{ij} = 0$, then $Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$ — exactly the RCBD structure of Chapter 8. The former interaction line, with $(a-1)(b-1)$ df, becomes the error line; both main effects can be tested.

One Observation per Cell

Section 9.9

The problem. Error df $= ab(n-1)$ come entirely from replication. With $n = 1$: $df_E = 0$, every "cell mean" is the observation itself, every residual in (9.12) is zero, and SS_{AB} and SS_E collapse into one quantity. No F test of the interaction.

The workaround. *Assume* the interaction away: if $(\tau\beta)_{ij} = 0$, then $Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$ — exactly the RCBD structure of Chapter 8. The former interaction line, with $(a-1)(b-1)$ df, becomes the error line; both main effects can be tested.

The price. A real interaction gets folded silently into MS_E: tests conservative at best, misleading at worst. Diagnostic: *Tukey's single-df test of nonadditivity*. If it rejects, transform or replicate.

One Observation per Cell

Section 9.9

The problem. Error df $= ab(n-1)$ come entirely from replication. With $n = 1$: $df_E = 0$, every "cell mean" is the observation itself, every residual in (9.12) is zero, and SS_{AB} and SS_E collapse into one quantity. No F test of the interaction.

The workaround. *Assume* the interaction away: if $(\tau\beta)_{ij} = 0$, then $Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$ — exactly the RCBD structure of Chapter 8. The former interaction line, with $(a-1)(b-1)$ df, becomes the error line; both main effects can be tested.

The price. A real interaction gets folded silently into MS_E: tests conservative at best, misleading at worst. Diagnostic: *Tukey's single-df test of nonadditivity*. If it rejects, transform or replicate.

Practical rule. Going from $n = 1$ to $n = 2$ buys an honest $\hat{\sigma}^2$, a valid interaction test, and residuals you can plot. Whenever feasible, run $n \geq 2$.

Three or More Factors: the General Factorial

Section 9.10, Table 9.8

With A, B, C at a, b, c levels and n replicates: $N = abcn$ runs, *seven* effect sources plus error — 3 main effects, 3 two-factor interactions, 1 three-factor interaction.

Three or More Factors: the General Factorial

Section 9.10, Table 9.8

With A, B, C at a, b, c levels and n replicates: $N = abcn$ runs, *seven* effect sources plus error — 3 main effects, 3 two-factor interactions, 1 three-factor interaction.

Source	**df**	**Source**	**df**
A	$a-1$	AB	$(a-1)(b-1)$
B	$b-1$	AC	$(a-1)(c-1)$
C	$c-1$	BC	$(b-1)(c-1)$
ABC	$(a-1)(b-1)(c-1)$	Error	$abc(n-1)$
Total			$abcn-1$

Interaction df = product of the main-effect df involved. Every F still uses MS_E; interactions are read before main effects.

Three or More Factors: the General Factorial

Section 9.10, Table 9.8

With A, B, C at a, b, c levels and n replicates: $N = abcn$ runs, *seven* effect sources plus error — 3 main effects, 3 two-factor interactions, 1 three-factor interaction.

Source	df	Source	df
A	$a-1$	AB	$(a-1)(b-1)$
B	$b-1$	AC	$(a-1)(c-1)$
C	$c-1$	BC	$(b-1)(c-1)$
ABC	$(a-1)(b-1)(c-1)$	Error	$abc(n-1)$
Total			$abcn-1$

Interaction df = product of the main-effect df involved. Every F still uses MS_E; interactions are read before main effects.

The catch: run count grows multiplicatively ($3^5 = 243$ cells) — hence the 2^k designs.

Quick Check: Three-Factor df

Q. A three-factor factorial has $a = 3$, $b = 2$, $c = 4$, $n = 2$. The df for the ABC interaction and for error are:

(A) $df_{ABC} = 9, \quad df_E = 24$

(B) $df_{ABC} = 6, \quad df_E = 24$

(C) $df_{ABC} = 9, \quad df_E = 47$

(D) $df_{ABC} = 24, \quad df_E = 24$

Quick Check: Three-Factor df

Q. A three-factor factorial has $a = 3$, $b = 2$, $c = 4$, $n = 2$. The df for the ABC interaction and for error are:

(A) $df_{ABC} = 9, \quad df_E = 24$

(B) $df_{ABC} = 6, \quad df_E = 24$

(C) $df_{ABC} = 9, \quad df_E = 47$

(D) $df_{ABC} = 24, \quad df_E = 24$

Answer: (B). $df_{ABC} = (a-1)(b-1)(c-1) = 2\times1\times3 = 6$; $df_E = abc(n-1) = 24 \times 1 = 24$; total $abcn - 1 = 47$. (A) multiplies the wrong terms; (C) confuses df_E with df_T; (D) mistakes the cell count abc for the interaction df.

The 2^k Workhorse

Section 9.11

Why 2^k Designs?

Section 9.11

The idea. Study k factors, each at only *two* levels, coded low ($-$) and high ($+$). $2^5 = 32$ combinations instead of $3^5 = 243$.

Why 2^k Designs?

Section 9.11

The idea. Study k factors, each at only *two* levels, coded low $(-)$ and high $(+)$. $2^5 = 32$ combinations instead of $3^5 = 243$.

Two levels per factor is the minimum that still detects:

- which factors move the response (main effects),
- which factors interact (two- and higher-order interactions).

Why 2^k Designs?

Section 9.11

The idea. Study k factors, each at only *two* levels, coded low ($-$) and high ($+$). $2^5 = 32$ combinations instead of $3^5 = 243$.

Two levels per factor is the minimum that still detects:

- which factors move the response (main effects),
- which factors interact (two- and higher-order interactions).

The price: two levels cannot detect curvature between them — center points (later today) repair that cheaply.

Why 2^k Designs?

Section 9.11

The idea. Study k factors, each at only *two* levels, coded low ($-$) and high ($+$). $2^5 = 32$ combinations instead of $3^5 = 243$.

Two levels per factor is the minimum that still detects:

- which factors move the response (main effects),
- which factors interact (two- and higher-order interactions).

The price: two levels cannot detect curvature between them — center points (later today) repair that cheaply.

Where they live: screening process variables, tuning simulation parameters, structuring red-team test campaigns, and multivariate product tests — the workhorse of engineering experimentation.

The 2^2 Design: Notation

Table 9.9

Factors A and B, each at two levels: four cells, n replicates each, $N = 4n$ runs.

The 2^2 Design: Notation

Table 9.9

Factors A and B, each at two levels: four cells, n replicates each, $N = 4n$ runs.

Label	A	B	**Meaning**
(1)	$-$	$-$	both factors low
a	$+$	$-$	A high, B low
b	$-$	$+$	A low, B high
ab	$+$	$+$	both factors high

The 2^2 Design: Notation

Table 9.9

Factors A and B, each at two levels: four cells, n replicates each, $N = 4n$ runs.

Label	A	B	**Meaning**
(1)	$-$	$-$	both factors low
a	$+$	$-$	A high, B low
b	$-$	$+$	A low, B high
ab	$+$	$+$	both factors high

Two conventions in one label: a lowercase letter appears when that factor is at its *high* level; (1) means no factor is high. And each label also names the *total* of the n replicates at that combination.

2^2 Effects, Contrasts, and Sums of Squares

The effect of A = average response at high A (runs a, ab; $2n$ observations) minus average at low A (runs (1), b):

$$\text{Effect}_A = \frac{a + ab - (1) - b}{2n}, \quad \text{Effect}_B = \frac{b + ab - (1) - a}{2n}, \quad \text{Effect}_{AB} = \frac{(1) + ab - a - b}{2n} \tag{9.13}$$

2^2 Effects, Contrasts, and Sums of Squares

The effect of A = average response at high A (runs a, ab; $2n$ observations) minus average at low A (runs (1), b):

$$\text{Effect}_A = \frac{a + ab - (1) - b}{2n}, \quad \text{Effect}_B = \frac{b + ab - (1) - a}{2n}, \quad \text{Effect}_{AB} = \frac{(1) + ab - a - b}{2n} \tag{9.13}$$

Each numerator — a signed sum of the four treatment totals — is that effect's **contrast**. The AB contrast compares the A effect at high B with the A effect at low B: the non-parallelism of the interaction plot.

2^2 Effects, Contrasts, and Sums of Squares

The effect of A = average response at high A (runs a, ab; $2n$ observations) minus average at low A (runs (1), b):

$$\text{Effect}_A = \frac{a + ab - (1) - b}{2n}, \quad \text{Effect}_B = \frac{b + ab - (1) - a}{2n}, \quad \text{Effect}_{AB} = \frac{(1) + ab - a - b}{2n} \tag{9.13}$$

Each numerator — a signed sum of the four treatment totals — is that effect's **contrast**. The AB contrast compares the A effect at high B with the A effect at low B: the non-parallelism of the interaction plot.

Each effect carries *one* degree of freedom, and one formula gives every sum of squares:

$$SS_{\text{effect}} = \frac{(\text{Contrast})^2}{4n}, \qquad df = 1 \tag{9.14}$$

2^2 Effects, Contrasts, and Sums of Squares

The effect of A = average response at high A (runs a, ab; $2n$ observations) minus average at low A (runs (1), b):

$$\text{Effect}_A = \frac{a + ab - (1) - b}{2n}, \quad \text{Effect}_B = \frac{b + ab - (1) - a}{2n}, \quad \text{Effect}_{AB} = \frac{(1) + ab - a - b}{2n} \tag{9.13}$$

Each numerator — a signed sum of the four treatment totals — is that effect's **contrast**. The AB contrast compares the A effect at high B with the A effect at low B: the non-parallelism of the interaction plot.

Each effect carries *one* degree of freedom, and one formula gives every sum of squares:

$$SS_{\text{effect}} = \frac{(\text{Contrast})^2}{4n}, \qquad df = 1 \tag{9.14}$$

$SS_E = SS_T - SS_A - SS_B - SS_{AB}$ with $df_E = 4(n-1)$ — Lecture 25's machinery simplified by the ± 1 coding.

2^2 Worked Example: Pricing Page $\times$ Onboarding Flow

Example 9.6

Setup. A startup tests two pricing pages (A: current $-$, redesigned $+$) and two onboarding flows (B: current $-$, guided $+$) on signup conversion (%); $n = 2$ weekly cohorts of 200 visitors each. Treatment totals: $(1) = 60, \quad a = 76, \quad b = 48, \quad ab = 72; \quad SS_T = 248.$

2^2 Worked Example: Pricing Page $\times$ Onboarding Flow

Example 9.6

Setup. A startup tests two pricing pages (A: current $-$, redesigned $+$) and two onboarding flows (B: current $-$, guided $+$) on signup conversion (%); $n = 2$ weekly cohorts of 200 visitors each. Treatment totals: $(1) = 60, \quad a = 76, \quad b = 48, \quad ab = 72; \quad SS_T = 248.$

Step 1 — Contrasts and effects by (9.13), with $2n = 4$:

$$\text{Contrast}_A = 76 + 72 - 60 - 48 = 40, \qquad \text{Effect}_A = 40/4 = 10.0$$

2^2 Worked Example: Pricing Page $\times$ Onboarding Flow

Example 9.6

Setup. A startup tests two pricing pages (A: current $-$, redesigned $+$) and two onboarding flows (B: current $-$, guided $+$) on signup conversion (%); $n = 2$ weekly cohorts of 200 visitors each. Treatment totals: $(1) = 60, \quad a = 76, \quad b = 48, \quad ab = 72; \quad SS_T = 248.$

Step 1 — Contrasts and effects by (9.13), with $2n = 4$:

$$\text{Contrast}_A = 76 + 72 - 60 - 48 = 40, \qquad \text{Effect}_A = 40/4 = 10.0$$

$$\text{Contrast}_B = 48 + 72 - 60 - 76 = -16, \qquad \text{Effect}_B = -16/4 = -4.0$$

2^2 Worked Example: Pricing Page $\times$ Onboarding Flow

Example 9.6

Setup. A startup tests two pricing pages (A: current $-$, redesigned $+$) and two onboarding flows (B: current $-$, guided $+$) on signup conversion (%); $n = 2$ weekly cohorts of 200 visitors each. Treatment totals: $(1) = 60, \quad a = 76, \quad b = 48, \quad ab = 72; \quad SS_T = 248.$

Step 1 — Contrasts and effects by (9.13), with $2n = 4$:

$$\text{Contrast}_A = 76 + 72 - 60 - 48 = 40, \qquad \text{Effect}_A = 40/4 = 10.0$$

$$\text{Contrast}_B = 48 + 72 - 60 - 76 = -16, \qquad \text{Effect}_B = -16/4 = -4.0$$

$$\text{Contrast}_{AB} = 60 + 72 - 76 - 48 = 8, \qquad \text{Effect}_{AB} = 8/4 = 2.0$$

2^2 Worked Example: Pricing Page $\times$ Onboarding Flow

Example 9.6

Setup. A startup tests two pricing pages (A: current $-$, redesigned $+$) and two onboarding flows (B: current $-$, guided $+$) on signup conversion (%); $n = 2$ weekly cohorts of 200 visitors each. Treatment totals: $(1) = 60, \quad a = 76, \quad b = 48, \quad ab = 72; \quad SS_T = 248.$

Step 1 — Contrasts and effects by (9.13), with $2n = 4$:

$$\text{Contrast}_A = 76 + 72 - 60 - 48 = 40, \qquad \text{Effect}_A = 40/4 = 10.0$$

$$\text{Contrast}_B = 48 + 72 - 60 - 76 = -16, \qquad \text{Effect}_B = -16/4 = -4.0$$

$$\text{Contrast}_{AB} = 60 + 72 - 76 - 48 = 8, \qquad \text{Effect}_{AB} = 8/4 = 2.0$$

The redesigned page is worth about $+10$ points of conversion; guided onboarding *lowers* it by about 4; the interaction estimate is small.

2^2 Worked Example: Sums of Squares and Error

Steps 2–3

Step 2 — Sums of squares by (9.14), with $4n = 8$:

$$SS_A = \frac{40^2}{8} = 200, \qquad SS_B = \frac{(-16)^2}{8} = 32, \qquad SS_{AB} = \frac{8^2}{8} = 8$$

2^2 Worked Example: Sums of Squares and Error

Steps 2–3

Step 2 — Sums of squares by (9.14), with $4n = 8$:

$$SS_A = \frac{40^2}{8} = 200, \qquad SS_B = \frac{(-16)^2}{8} = 32, \qquad SS_{AB} = \frac{8^2}{8} = 8$$

Step 3 — Error by subtraction:

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = 248 - 200 - 32 - 8 = 8$$

2^2 Worked Example: Sums of Squares and Error

Steps 2–3

Step 2 — Sums of squares by (9.14), with $4n = 8$:

$$SS_A = \frac{40^2}{8} = 200, \qquad SS_B = \frac{(-16)^2}{8} = 32, \qquad SS_{AB} = \frac{8^2}{8} = 8$$

Step 3 — Error by subtraction:

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = 248 - 200 - 32 - 8 = 8$$

$$df_E = N - 2^k = 8 - 4 = 4, \qquad MS_E = \frac{8}{4} = 2$$

2^2 Worked Example: F Tests and Recommendation

Step 4

Each effect has $df = 1$, so $MS_{\text{effect}} = SS_{\text{effect}}$. Critical value $f_{0.05,1,4} = 7.71$.

2^2 Worked Example: F Tests and Recommendation

Step 4

Each effect has $df = 1$, so $MS_{\text{effect}} = SS_{\text{effect}}$. Critical value $f_{0.05,1,4} = 7.71$.

Interaction first: $F_0^{AB} = \dfrac{8}{2} = 4.0 < 7.71 \quad \Rightarrow$ not significant. The two page elements act additively.

2^2 Worked Example: F Tests and Recommendation

Step 4

Each effect has $df = 1$, so $MS_{\text{effect}} = SS_{\text{effect}}$. Critical value $f_{0.05,1,4} = 7.71$.

Interaction first: $F_0^{AB} = \dfrac{8}{2} = 4.0 < 7.71 \quad \Rightarrow$ not significant. The two page elements act additively.

Main effects: $F_0^{A} = \dfrac{200}{2} = 100 > 7.71; \quad F_0^{B} = \dfrac{32}{2} = 16 > 7.71 \quad \Rightarrow$ both significant.

2^2 Worked Example: F Tests and Recommendation

Step 4

Each effect has $df = 1$, so $MS_{\text{effect}} = SS_{\text{effect}}$. Critical value $f_{0.05,1,4} = 7.71$.

Interaction first: $F_0^{AB} = \dfrac{8}{2} = 4.0 < 7.71 \Rightarrow$ not significant. The two page elements act additively.

Main effects: $F_0^{A} = \dfrac{200}{2} = 100 > 7.71$; $F_0^{B} = \dfrac{32}{2} = 16 > 7.71 \Rightarrow$ both significant.

Recommendation: deploy a — redesigned pricing with the *current* onboarding — worth about $10 + 4 = 14$ points of conversion over the worst combination.

2^2 Worked Example: F Tests and Recommendation

Step 4

Each effect has $df = 1$, so $MS_{\text{effect}} = SS_{\text{effect}}$. Critical value $f_{0.05,1,4} = 7.71$.

Interaction first: $F_0^{AB} = \dfrac{8}{2} = 4.0 < 7.71 \quad \Rightarrow$ not significant. The two page elements act additively.

Main effects: $F_0^{A} = \dfrac{200}{2} = 100 > 7.71; \quad F_0^{B} = \dfrac{32}{2} = 16 > 7.71 \quad \Rightarrow$ both significant.

Recommendation: deploy a — redesigned pricing with the *current* onboarding — worth about $10 + 4 = 14$ points of conversion over the worst combination.

Startup note. Eight cohort-weeks answered three questions at once: the pricing effect, the onboarding effect, and whether they interfere. Two sequential A/B tests would have spent the same traffic and left the interaction question unanswered.

Scaling Up: the General 2^k Formulas

With k factors and n replicates: $N = n\,2^k$ runs and $2^k - 1$ effects (for $k = 3$: three main effects, three two-factor interactions, one three-factor interaction).

Scaling Up: the General 2^k Formulas

With k factors and n replicates: $N = n\,2^k$ runs and $2^k - 1$ effects (for $k = 3$: three main effects, three two-factor interactions, one three-factor interaction).

Two formulas replace every special case:

$$\text{Effect} = \frac{\text{Contrast}}{n\,2^{k-1}} \tag{9.15}$$

Scaling Up: the General 2^k Formulas

With k factors and n replicates: $N = n\,2^k$ runs and $2^k - 1$ effects (for $k = 3$: three main effects, three two-factor interactions, one three-factor interaction).

Two formulas replace every special case:

$$\text{Effect} = \frac{\text{Contrast}}{n\,2^{k-1}} \tag{9.15}$$

$$SS_{\text{effect}} = \frac{(\text{Contrast})^2}{n\,2^k}, \qquad df = 1 \tag{9.16}$$

Scaling Up: the General 2^k Formulas

With k factors and n replicates: $N = n\,2^k$ runs and $2^k - 1$ effects (for $k = 3$: three main effects, three two-factor interactions, one three-factor interaction).

Two formulas replace every special case:

$$\text{Effect} = \frac{\text{Contrast}}{n\,2^{k-1}} \tag{9.15}$$

$$SS_{\text{effect}} = \frac{(\text{Contrast})^2}{n\,2^k}, \qquad df = 1 \tag{9.16}$$

For $k = 2$ the denominators reduce to $2n$ and $4n$, matching (9.13) and (9.14).

Scaling Up: the General 2^k Formulas

With k factors and n replicates: $N = n\,2^k$ runs and $2^k - 1$ effects (for $k = 3$: three main effects, three two-factor interactions, one three-factor interaction).

Two formulas replace every special case:

$$\text{Effect} = \frac{\text{Contrast}}{n\,2^{k-1}} \tag{9.15}$$

$$SS_{\text{effect}} = \frac{(\text{Contrast})^2}{n\,2^k}, \qquad df = 1 \tag{9.16}$$

For $k = 2$ the denominators reduce to $2n$ and $4n$, matching (9.13) and (9.14).

Common exam error: keeping $4n$ in the SS denominator when $k > 2$. The denominator is $n\,2^k$ — for a 2^3 with $n = 2$, that is 16, not 8.

The Sign Table for 2^3

Table 9.10

Rule. Columns A, B, C code the factor levels. Each interaction column is the *entrywise product* of its parents: $AB = A \times B$, $ABC = A \times B \times C$.

The Sign Table for 2^3

Table 9.10

Rule. Columns A, B, C code the factor levels. Each interaction column is the *entrywise product* of its parents: $AB = A \times B$, $ABC = A \times B \times C$.

Run	A	B	C	AB	AC	BC	ABC
(1)	−	−	−	+	+	+	−
a	+	−	−	−	−	+	+
b	−	+	−	−	+	−	+
ab	+	+	−	+	−	−	−
c	−	−	+	+	−	−	+
ac	+	−	+	−	+	−	−
bc	−	+	+	−	−	+	−
abc	+	+	+	+	+	+	+

The Sign Table for 2^3

Table 9.10

Rule. Columns A, B, C code the factor levels. Each interaction column is the *entrywise product* of its parents: $AB = A \times B$, $ABC = A \times B \times C$.

Run	A	B	C	AB	AC	BC	ABC
(1)	−	−	−	+	+	+	−
a	+	−	−	−	−	+	+
b	−	+	−	−	+	−	+
ab	+	+	−	+	−	−	−
c	−	−	+	+	−	−	+
ac	+	−	+	−	+	−	−
bc	−	+	+	−	−	+	−
abc	+	+	+	+	+	+	+

To build any contrast: read down the effect's column, attach each sign to the matching treatment total, add. For AC:

$$\text{Contrast}_{AC} = (1) - a + b - ab - c + ac - bc + abc; \quad \text{then } SS_{AC} \text{ by (9.16)}$$

Procedure 9.3: Analyzing a 2^k Factorial

1. Write the sign table for the 2^k treatment combinations (Table 9.10 for $k = 3$); record the treatment totals.
2. For each of the $2^k - 1$ effects: multiply the column signs by the treatment totals and sum $\Rightarrow$ the contrast.
3. Compute each effect from (9.15) and each SS from (9.16). The SS denominator is $n\,2^k$ for *every* effect.
4. Compute SS_T from the individual observations (9.4); then $SS_E = SS_T - \sum SS_{\text{effect}}$, with $df_E = N - 2^k$.
5. Build the ANOVA table: every effect has $df = 1$, so $MS = SS$; test each $F_0 = MS_{\text{effect}}/MS_E$ against $f_{\alpha,1,N-2^k}$. Read interactions before main effects, as always.

2^3 Worked Example: Truck Fuel Savings

Example 9.7

Setup. A fleet engineer studies % fuel saving under three two-level factors: A aerodynamic fairing; B low-rolling-resistance tires; C speed governor. Each configuration on $n = 2$ matched routes, $N = 16$. Treatment totals: $(1) = 3, \quad a = 17, \quad b = 9, \quad ab = 27, \quad c = 7, \quad ac = 21, \quad bc = 13, \quad abc = 31; \quad SS_T = 356.$

2^3 Worked Example: Truck Fuel Savings

Example 9.7

Setup. A fleet engineer studies % fuel saving under three two-level factors: A aerodynamic fairing; B low-rolling-resistance tires; C speed governor. Each configuration on $n = 2$ matched routes, $N = 16$. Treatment totals: $(1) = 3, \quad a = 17, \quad b = 9, \quad ab = 27, \quad c = 7, \quad ac = 21, \quad bc = 13, \quad abc = 31; \quad SS_T = 356.$

Steps 1–2 — One contrast, carefully. Reading down the A column of Table 9.10:

$$\text{Contrast}_A = -(1)+a-b+ab-c+ac-bc+abc$$

2^3 Worked Example: Truck Fuel Savings

Example 9.7

Setup. A fleet engineer studies % fuel saving under three two-level factors: A aerodynamic fairing; B low-rolling-resistance tires; C speed governor. Each configuration on $n = 2$ matched routes, $N = 16$. Treatment totals: $(1) = 3, \quad a = 17, \quad b = 9, \quad ab = 27, \quad c = 7, \quad ac = 21, \quad bc = 13, \quad abc = 31; \quad SS_T = 356.$

Steps 1–2 — One contrast, carefully. Reading down the A column of Table 9.10:

$$\text{Contrast}_A = -(1)+a-b+ab-c+ac-bc+abc = -3+17-9+27-7+21-13+31 = 64$$

2^3 Worked Example: Truck Fuel Savings

Example 9.7

Setup. A fleet engineer studies % fuel saving under three two-level factors: A aerodynamic fairing; B low-rolling-resistance tires; C speed governor. Each configuration on $n = 2$ matched routes, $N = 16$. Treatment totals: $(1) = 3, \quad a = 17, \quad b = 9, \quad ab = 27, \quad c = 7, \quad ac = 21, \quad bc = 13, \quad abc = 31; \quad SS_T = 356.$

Steps 1–2 — One contrast, carefully. Reading down the A column of Table 9.10:

$$\text{Contrast}_A = -(1)+a-b+ab-c+ac-bc+abc = -3+17-9+27-7+21-13+31 = 64$$

Repeat for the other six columns: $\text{Contrast}_B = 32$, $\text{Contrast}_C = 16$, $\text{Contrast}_{AB} = 8$, $\text{Contrast}_{AC} = \text{Contrast}_{BC} = \text{Contrast}_{ABC} = 0$.

2^3 Worked Example: Effects, SS, and Error

Steps 3–4

Step 3. With $n\,2^{k-1} = 8$ (9.15) and $n\,2^{k} = 16$ (9.16):

Effect	A	B	C	AB	AC	BC	ABC
Contrast	64	32	16	8	0	0	0
Effect $=$ Contrast$/8$	8.0	4.0	2.0	1.0	0	0	0
$SS =$ Contrast$^2/16$	256	64	16	4	0	0	0

2^3 Worked Example: Effects, SS, and Error

Steps 3–4

Step 3. With $n\,2^{k-1} = 8$ (9.15) and $n\,2^{k} = 16$ (9.16):

Effect	A	B	C	AB	AC	BC	ABC
Contrast	64	32	16	8	0	0	0
Effect $=$ Contrast$/8$	8.0	4.0	2.0	1.0	0	0	0
$SS =$ Contrast$^2/16$	256	64	16	4	0	0	0

Read the effects in the response's own units: the fairing is worth about 8 percentage points of saving, the tires about 4, the governor about 2.

2^3 Worked Example: Effects, SS, and Error

Steps 3–4

Step 3. With $n\,2^{k-1} = 8$ (9.15) and $n\,2^k = 16$ (9.16):

Effect	A	B	C	AB	AC	BC	ABC
Contrast	64	32	16	8	0	0	0
Effect $=$ Contrast$/8$	8.0	4.0	2.0	1.0	0	0	0
$SS =$ Contrast$^2/16$	256	64	16	4	0	0	0

Read the effects in the response's own units: the fairing is worth about 8 percentage points of saving, the tires about 4, the governor about 2.

Step 4 — Error:

$$SS_E = 356 - (256+64+16+4+0+0+0) = 16, \qquad df_E = N - 2^k = 16-8 = 8, \qquad MS_E = 2$$

2^3 Worked Example: F Tests and Conclusion

Step 5

Each effect has $df = 1$, so $F_0 = SS/MS_E = SS/2$, against $f_{0.05,1,8} = 5.32$:

Effect	A	B	C	AB	AC	BC	ABC
F_0	128	32	8	2	0	0	0
vs. 5.32	sig	sig	sig	ns	ns	ns	ns

2^3 Worked Example: F Tests and Conclusion

Step 5

Each effect has $df = 1$, so $F_0 = SS/MS_E = SS/2$, against $f_{0.05,1,8} = 5.32$:

Effect	A	B	C	AB	AC	BC	ABC
F_0	128	32	8	2	0	0	0
vs. 5.32	sig	sig	sig	ns	ns	ns	ns

Conclusion. All three main effects significant; no interaction is. The three measures act independently, so their savings *add*: fitting all three is expected to save about $8 + 4 + 2 = 14\%$ of fuel.

2^3 Worked Example: F Tests and Conclusion

Step 5

Each effect has $df = 1$, so $F_0 = SS/MS_E = SS/2$, against $f_{0.05,1,8} = 5.32$:

Effect	A	B	C	AB	AC	BC	ABC
F_0	128	32	8	2	0	0	0
vs. 5.32	sig	sig	sig	ns	ns	ns	ns

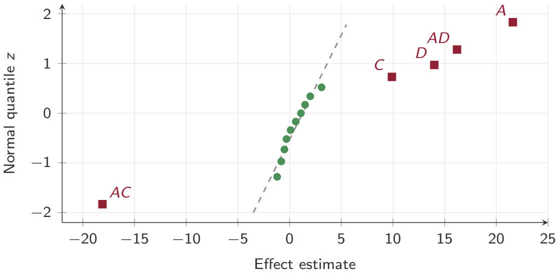
Conclusion. All three main effects significant; no interaction is. The three measures act independently, so their savings *add*: fitting all three is expected to save about $8 + 4 + 2 = 14\%$ of fuel.

Engineering note. The additivity is itself information: fairing, tires, and governor attack different loss mechanisms (drag, rolling resistance, speed profile). A large AB would have pointed to a shared mechanism — physics you can go look for.

Single Replicates and the Normal Plot of Effects

Section 9.12, Figure 9.8

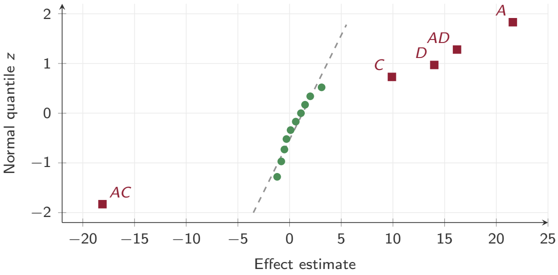
For $k = 5$, $n = 1$ is already 32 runs — but then $df_E = N - 2^k = 0$: no F test. **Sparsity of effects**: in real systems only a few effects are large; the rest behave like noise, so they line up on a normal probability plot, and the *active* effects fall off the line.



Single Replicates and the Normal Plot of Effects

Section 9.12, Figure 9.8

For $k = 5$, $n = 1$ is already 32 runs — but then $df_E = N - 2^k = 0$: no F test. **Sparsity of effects**: in real systems only a few effects are large; the rest behave like noise, so they line up on a normal probability plot, and the *active* effects fall off the line.



Single-replicate 2^4: ten effects on the line (noise) are pooled as error; A, D, AD, C, AC are active — and the active set contains *interactions*.

Center Points: a Curvature Check

Section 9.12

The blind spot. A 2^k design observes only the *corners* of a k-dimensional box. It cannot tell whether the response is curved between them.

Center Points: a Curvature Check

Section 9.12

The blind spot. A 2^k design observes only the *corners* of a k-dimensional box. It cannot tell whether the response is curved between them.

The fix. Add n_C runs (typically 3–5) at the **center point**: every factor at the midpoint of its two levels. Two things at once:

- Genuine replicates of one condition $\Rightarrow$ pure-error estimate of σ^2 without disturbing the factorial balance.
- Curvature check: compare average corner response $\bar{y}_F$ with average center response $\bar{y}_C$. Small $|\bar{y}_F - \bar{y}_C| \Rightarrow$ linear picture adequate; large $\Rightarrow$ the surface bends, plan a follow-up with more levels.

Center Points: a Curvature Check

Section 9.12

The blind spot. A 2^k design observes only the *corners* of a k-dimensional box. It cannot tell whether the response is curved between them.

The fix. Add n_C runs (typically 3–5) at the **center point**: every factor at the midpoint of its two levels. Two things at once:

- Genuine replicates of one condition $\Rightarrow$ pure-error estimate of σ^2 without disturbing the factorial balance.
- Curvature check: compare average corner response $\bar{y}_F$ with average center response $\bar{y}_C$. Small $|\bar{y}_F - \bar{y}_C| \Rightarrow$ linear picture adequate; large $\Rightarrow$ the surface bends, plan a follow-up with more levels.

Do not confuse center points with replicates of the corners — they sit at the midpoint, a condition the factorial itself never visits.

Blocking and Confounding in 2^k

Section 9.12

Why block? The 2^k runs may not fit under uniform conditions — one shift, one batch, one day of stable traffic. Splitting into blocks controls the nuisance variation (same motive as the RCBD of Chapter 8).

Blocking and Confounding in 2^k

Section 9.12

Why block? The 2^k runs may not fit under uniform conditions — one shift, one batch, one day of stable traffic. Splitting into blocks controls the nuisance variation (same motive as the RCBD of Chapter 8).

The price: confounding. One effect's contrast becomes identical to the block contrast — the two cannot be told apart. Standard choice: sacrifice the *highest-order interaction*, the effect least likely to be real.

Blocking and Confounding in 2^k

Section 9.12

Why block? The 2^k runs may not fit under uniform conditions — one shift, one batch, one day of stable traffic. Splitting into blocks controls the nuisance variation (same motive as the RCBD of Chapter 8).

The price: confounding. One effect's contrast becomes identical to the block contrast — the two cannot be told apart. Standard choice: sacrifice the *highest-order interaction*, the effect least likely to be real.

Example: 2^3 **in two blocks of 4**, confounding ABC. Split by the sign of the ABC column in Table 9.10:

- Block 1 (ABC sign $-$): $(1),\ ab,\ ac,\ bc$
- Block 2 (ABC sign $+$): $a,\ b,\ c,\ abc$

Blocking and Confounding in 2^k

Section 9.12

Why block? The 2^k runs may not fit under uniform conditions — one shift, one batch, one day of stable traffic. Splitting into blocks controls the nuisance variation (same motive as the RCBD of Chapter 8).

The price: confounding. One effect's contrast becomes identical to the block contrast — the two cannot be told apart. Standard choice: sacrifice the *highest-order interaction*, the effect least likely to be real.

Example: 2^3 in two blocks of 4, confounding ABC. Split by the sign of the ABC column in Table 9.10:

- Block 1 (ABC sign $-$): $(1),\ ab,\ ac,\ bc$
- Block 2 (ABC sign $+$): $a,\ b,\ c,\ abc$

What survives: all three main effects and all three two-factor interactions are still cleanly estimated. Only ABC is lost to the block difference.

Fractional Replication: 2^{k-p} Designs

Section 9.12

The problem. $2^7 = 128$ runs is impractical for screening — and by sparsity, most runs buy estimates of high-order interactions that are almost surely negligible.

Fractional Replication: 2^{k-p} Designs

Section 9.12

The problem. $2^7 = 128$ runs is impractical for screening — and by sparsity, most runs buy estimates of high-order interactions that are almost surely negligible.

Half-fraction 2^{k-1}. Run only the half chosen by a *generator*. A 2^{5-1} with defining relation $I = ABCDE$ keeps the 16 runs (of 32) with a $+$ sign in the $ABCDE$ column.

Fractional Replication: 2^{k-p} Designs

Section 9.12

The problem. $2^7 = 128$ runs is impractical for screening — and by sparsity, most runs buy estimates of high-order interactions that are almost surely negligible.

Half-fraction 2^{k-1}. Run only the half chosen by a *generator*. A 2^{5-1} with defining relation $I = ABCDE$ keeps the 16 runs (of 32) with a $+$ sign in the $ABCDE$ column.

The price: aliasing. In a fraction, every effect shares its sign column with a partner: with $I = ABCDE$, $A = BCDE$, $AB = CDE$, ... The data estimate the alias pair's *sum*; only engineering judgment (usually sparsity) says which member to credit.

Fractional Replication: 2^{k-p} Designs

Section 9.12

The problem. $2^7 = 128$ runs is impractical for screening — and by sparsity, most runs buy estimates of high-order interactions that are almost surely negligible.

Half-fraction 2^{k-1}. Run only the half chosen by a *generator*. A 2^{5-1} with defining relation $I = ABCDE$ keeps the 16 runs (of 32) with a $+$ sign in the $ABCDE$ column.

The price: aliasing. In a fraction, every effect shares its sign column with a partner: with $I = ABCDE$, $A = BCDE$, $AB = CDE$, ... The data estimate the alias pair's *sum*; only engineering judgment (usually sparsity) says which member to credit.

Smaller fractions 2^{k-p}: p generators cut runs by 2^p, with denser aliasing. **Industrial workflow:** screen many factors with a small fraction $\rightarrow$ identify the few active ones (normal plot) $\rightarrow$ follow up with a full factorial on those alone.

Quick Check: Screening under a Run Budget

Example 9.8

Q. A red team must screen $k = 6$ binary configuration factors of an LLM deployment (safety fine-tune, prompt hardening, output filter, retrieval guard, temperature cap, tool-use gate) for their effect on jailbreak success rate. The evaluation pipeline can afford 16 runs. Which design fits, and what is the trade-off?

(A) 2^6 full factorial; no trade-off.

(B) 2^{6-2} quarter-fraction; effects are aliased in groups.

(C) 2^{6-1} half-fraction; no aliasing.

(D) One-way ANOVA with 6 treatments.

Quick Check: Screening under a Run Budget

Example 9.8

Q. A red team must screen $k = 6$ binary configuration factors of an LLM deployment (safety fine-tune, prompt hardening, output filter, retrieval guard, temperature cap, tool-use gate) for their effect on jailbreak success rate. The evaluation pipeline can afford 16 runs. Which design fits, and what is the trade-off?

(A) 2^6 full factorial; no trade-off.

(B) 2^{6-2} quarter-fraction; effects are aliased in groups.

(C) 2^{6-1} half-fraction; no aliasing.

(D) One-way ANOVA with 6 treatments.

Answer: (B). $2^{6-2} = 2^4 = 16$ runs — exactly the budget; the price is aliasing (with resolution-IV generators, main effects are clean under sparsity, but two-factor interactions alias each other). (A) needs 64 runs. (C) needs 32 *and* still has aliasing. (D) ignores the factorial structure entirely.

Factorial Thinking in Engineering Practice

Section 9.13

AI safety testing. An evaluation campaign is a factorial the moment it crosses model versions with test conditions. Regression risk lives on the *interaction* line: a version that improves the average while getting worse on one scenario family is a version-by-scenario interaction, not a main effect. Reports quoting only per-version averages report main effects — ask to see the cell means. Many switches $\rightarrow$ fractional designs, aliasing stated in the test plan.

Factorial Thinking in Engineering Practice

Section 9.13

AI safety testing. An evaluation campaign is a factorial the moment it crosses model versions with test conditions. Regression risk lives on the *interaction* line: a version that improves the average while getting worse on one scenario family is a version-by-scenario interaction, not a main effect. Reports quoting only per-version averages report main effects — ask to see the cell means. Many switches → fractional designs, aliasing stated in the test plan.

Supply chains and transportation. Pilots span heterogeneous conditions: layouts across shifts, routing policies across terminals. Treating the condition as a second factor — not averaging over it — separates a defensible pilot from an anecdote. Nuisance conditions (day of week, vessel arrivals) → blocking, confounded with a chosen high-order interaction. The policy-by-site interaction says where a policy will and will not transfer.

Factorial Thinking in Engineering Practice

Section 9.13

AI safety testing. An evaluation campaign is a factorial the moment it crosses model versions with test conditions. Regression risk lives on the *interaction* line: a version that improves the average while getting worse on one scenario family is a version-by-scenario interaction, not a main effect. Reports quoting only per-version averages report main effects — ask to see the cell means. Many switches → fractional designs, aliasing stated in the test plan.

Supply chains and transportation. Pilots span heterogeneous conditions: layouts across shifts, routing policies across terminals. Treating the condition as a second factor — not averaging over it — separates a defensible pilot from an anecdote. Nuisance conditions (day of week, vessel arrivals) → blocking, confounded with a chosen high-order interaction. The policy-by-site interaction says where a policy will and will not transfer.

Building your own startup. Sequential A/B testing is OFAT; multivariate testing is a factorial, and the 2^2 of today is its smallest honest unit. Constraint: traffic — a 2^3 splits visitors eight ways, so effects must be large or cohorts long. Fractions and single replicates with a normal plot are the standard compromises.

Two Cautions Before You Use Any of This

Caution 1: significant $\neq$ worth acting on. With enough data, an effect of 0.05 units clears the F test even when the engineering tolerance is ± 2 units. Compare effect *estimates* — (9.15) gives them in the response's own units — against practical thresholds, not only against critical values.

Two Cautions Before You Use Any of This

Caution 1: significant $\neq$ worth acting on. With enough data, an effect of 0.05 units clears the F test even when the engineering tolerance is ± 2 units. Compare effect *estimates* — (9.15) gives them in the response's own units — against practical thresholds, not only against critical values.

Caution 2: the four standard factorial mistakes (exams and practice alike):

1. Using $4n$ in the SS denominator when $k > 2$ — it is $n\,2^k$ (9.16).
2. Mis-signing a contrast — recheck the sign-table column (Table 9.10).
3. Forgetting that $n = 1$ leaves $df_E = 0$ — use the normal plot of effects.
4. Reading center points as corner replicates — they sit at the midpoint.

Two Cautions Before You Use Any of This

Caution 1: significant $\neq$ worth acting on. With enough data, an effect of 0.05 units clears the F test even when the engineering tolerance is ± 2 units. Compare effect *estimates* — (9.15) gives them in the response's own units — against practical thresholds, not only against critical values.

Caution 2: the four standard factorial mistakes (exams and practice alike):

1. Using $4n$ in the SS denominator when $k > 2$ — it is $n\,2^k$ (9.16).
2. Mis-signing a contrast — recheck the sign-table column (Table 9.10).
3. Forgetting that $n = 1$ leaves $df_E = 0$ — use the normal plot of effects.
4. Reading center points as corner replicates — they sit at the midpoint.

Slowing down at exactly these four points prevents most wrong answers.

Final Exam: Chapter 9 Focus

What you must be able to do:

1. Run the full two-factor ANOVA by hand: Procedure 9.1, interaction tested first ((9.10), then (9.11)).
2. Check adequacy with residuals $e_{ijk} = y_{ijk} - \bar{y}_{ij\cdot}$ (9.12) and the four panels of Procedure 9.2.
3. Count df for any factorial, including three factors and the $n = 1$ case.
4. Build contrasts from the sign table; compute effects (9.15) and SS (9.16); finish the 2^k ANOVA (Procedure 9.3).
5. Explain, in one sentence each, what blocking confounds and what a fraction aliases.

Final Exam: Chapter 9 Focus

What you must be able to do:

1. Run the full two-factor ANOVA by hand: Procedure 9.1, interaction tested first ((9.10), then (9.11)).
2. Check adequacy with residuals $e_{ijk} = y_{ijk} - \bar{y}_{ij\cdot}$ (9.12) and the four panels of Procedure 9.2.
3. Count df for any factorial, including three factors and the $n = 1$ case.
4. Build contrasts from the sign table; compute effects (9.15) and SS (9.16); finish the 2^k ANOVA (Procedure 9.3).
5. Explain, in one sentence each, what blocking confounds and what a fraction aliases.

And interpret in plain language: a significant interaction means the conclusion must be stated level by level — from cell means, never from row means alone.

Lecture 26 Summary

- Residuals (9.12): deviation from the cell mean; sum to zero per cell. Procedure 9.2: normal plot, vs. fits, vs. each factor; funnel $\Rightarrow$ transform.

Lecture 26 Summary

- Residuals (9.12): deviation from the cell mean; sum to zero per cell. Procedure 9.2: normal plot, vs. fits, vs. each factor; funnel $\Rightarrow$ transform.
- $n = 1$: no error df; assume additivity (RCBD form), check with Tukey's nonadditivity test.

Lecture 26 Summary

- Residuals (9.12): deviation from the cell mean; sum to zero per cell. Procedure 9.2: normal plot, vs. fits, vs. each factor; funnel $\Rightarrow$ transform.
- $n = 1$: no error df; assume additivity (RCBD form), check with Tukey's nonadditivity test.
- General factorial: interaction df are products of main-effect df; run count grows fast — hence 2^k.

Lecture 26 Summary

- Residuals (9.12): deviation from the cell mean; sum to zero per cell. Procedure 9.2: normal plot, vs. fits, vs. each factor; funnel $\Rightarrow$ transform.
- $n = 1$: no error df; assume additivity (RCBD form), check with Tukey's nonadditivity test.
- General factorial: interaction df are products of main-effect df; run count grows fast — hence 2^k.
- 2^k: contrasts (9.13) and the sign table; Effect $=$ Contrast$/(n\,2^{k-1})$ (9.15), $SS =$ Contrast$^2/(n\,2^k)$ (9.16); Procedure 9.3.

Lecture 26 Summary

- Residuals (9.12): deviation from the cell mean; sum to zero per cell. Procedure 9.2: normal plot, vs. fits, vs. each factor; funnel $\Rightarrow$ transform.
- $n = 1$: no error df; assume additivity (RCBD form), check with Tukey's nonadditivity test.
- General factorial: interaction df are products of main-effect df; run count grows fast — hence 2^k.
- 2^k: contrasts (9.13) and the sign table; Effect $=$ Contrast$/(n\,2^{k-1})$ (9.15), $SS =$ Contrast$^2/(n\,2^k)$ (9.16); Procedure 9.3.
- Single replicates: normal plot of effects. Center points: curvature check. Blocking: confound the highest-order interaction. Fractions 2^{k-p}: aliasing for run count.

Lecture 26 Summary

- Residuals (9.12): deviation from the cell mean; sum to zero per cell. Procedure 9.2: normal plot, vs. fits, vs. each factor; funnel $\Rightarrow$ transform.
- $n = 1$: no error df; assume additivity (RCBD form), check with Tukey's nonadditivity test.
- General factorial: interaction df are products of main-effect df; run count grows fast — hence 2^k.
- 2^k: contrasts (9.13) and the sign table; Effect $=$ Contrast$/(n\,2^{k-1})$ (9.15), $SS =$ Contrast$^2/(n\,2^k)$ (9.16); Procedure 9.3.
- Single replicates: normal plot of effects. Center points: curvature check. Blocking: confound the highest-order interaction. Fractions 2^{k-p}: aliasing for run count.
- Applications: eval grids, operational pilots, multivariate tests — the interaction is usually the finding that matters.

Lecture 26 Summary

- Residuals (9.12): deviation from the cell mean; sum to zero per cell. Procedure 9.2: normal plot, vs. fits, vs. each factor; funnel $\Rightarrow$ transform.
- $n = 1$: no error df; assume additivity (RCBD form), check with Tukey's nonadditivity test.
- General factorial: interaction df are products of main-effect df; run count grows fast — hence 2^k.
- 2^k: contrasts (9.13) and the sign table; Effect $=$ Contrast$/(n\,2^{k-1})$ (9.15), $SS =$ Contrast$^2/(n\,2^k)$ (9.16); Procedure 9.3.
- Single replicates: normal plot of effects. Center points: curvature check. Blocking: confound the highest-order interaction. Fractions 2^{k-p}: aliasing for run count.
- Applications: eval grids, operational pilots, multivariate tests — the interaction is usually the finding that matters.

Thank you for a great semester. Best of luck on the final.