Chapter 7. Experiments with Two or More Factors

Section 7.1 Two-Factor Factorial Designs

Example 1 – Identify the response, factors and replications

- A study of smoking classifies subjects as nonsmokers, moderate smokers, or heavy smokers. Samples of 80 men and 80 women are drawn from each group. Each person reports the number of hours of sleep he or she gets on a typical night.

(factor A-smoker groups, factor B- Gender, rep=80)

b. The strength of concrete depends upon the formula used to prepare it. An experimenter compares 6 different mixtures. Nine specimens of concrete are poured from each mixture. Three of these specimens are subjected to 0 cycles of freezing and thawing, 3 are subjected to 100 cycles, and 3 specimens are subjected to 500 cycles. The strength of each specimen is then measured.

(factor A- formula mixtures, factor B- {subject to 0, 100, 500 cycles}, rep=3)

c. Four methods for teaching sign language are to be compared. Sixteen students in special education and sixteen students majoring in other areas are the subjects for the study. Within each group they are randomly assigned to the methods. Scores on a final exam are compared.

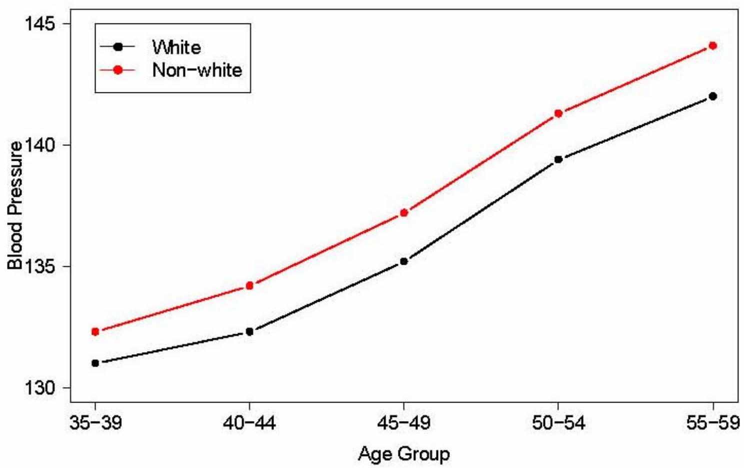
(factor A –methods of teaching sign language, factor B-majors, rep=16)

Example 2: In the course of a clinical trial of measures to prevent coronary heart disease, blood pressure measurements were taken on 12,866 men. Individuals were classified by age group and race. The means for systolic blood pressure are given in the following table:

	Age Group				
	35-39	40-44	45-49	50-54	55-59
White	131.0	132.3	135.2	139.4	142.0
Non-White	132.3	134.2	137.2	141.3	144.1

Note that we are not given raw data on these 12,866 men. The table above is the mean for each race/age combination. This means we can't use ANOVA. Well just use graphing and marginal means to describe the situation.

a. Plot the means with age on the x-axis and blood pressure on the y-axis. For each racial group connect the points for the different ages.



b. Describe the patterns you see. Does there appear to be a difference between the 2 racial groups? Does systolic blood pressure appear to vary with age? If so, how does it vary? Is there an interaction between race and age?

Model for a×b factorial design

- Mean Model: $Y_{ijk} = \mu_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ijk} \overset{iid}{\sim} N\left(0, \sigma^2\right)$

  $i = 1,\ldots,a$ (factor A), $j = 1,\ldots,b,$ (factor B),

  $k = 1,..,n.$ (replications).

Overall mean: $\bar{\mu}_{..} = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b}\mu_{ij}$

Row Treatment (factor A) mean: $\bar{\mu}_{i.} = \frac{1}{b}\sum_{j=1}^{b}\mu_{ij}, \ \mathrm{i} = 1,\ldots,\mathrm{a},$

Row effect: $\alpha_i = \bar{\mu}_{i.} - \bar{\mu}_{..}, \ \mathrm{i} = 1,\ldots,\mathrm{a}, \text{ and } \sum_{i=1}^{a}\alpha_i = 0.$

Column Treatment (factor B) mean: $\bar{\mu}_{.j} = \frac{1}{a}\sum_{i=1}^{a}\mu_{ij}, \ \mathrm{j} = 1,\ldots,\mathrm{b}.$

Column effect: $\beta_{\mathrm{j}} = \bar{\mu}_{.j} - \bar{\mu}_{..}, \ \mathrm{j} = 1,\ldots,\mathrm{b}, \text{ and } \sum_{j=1}^{b}\beta_{\mathrm{j}} = 0.$

Means Table

Factor Levels	1 - B	2 - B	…	b - B	Row Mean
1 - A	μ_{11}	μ_{12}	…	μ_{1b}	$\overline{\mu}_{1.}$
2 - A	μ_{21}	μ_{22}	…	μ_{2b}	$\overline{\mu}_{2.}$
…	…	…	…	…	…
a - A	μ_{a1}	μ_{a2}	…	μ_{ab}	$\overline{\mu}_{a.}$
Column Mean	$\overline{\mu}_{.1}$	$\overline{\mu}_{.2}$	…	$\overline{\mu}_{.b}$	$\overline{\mu}_{..}$

With or without interaction?

- check if $\mu_{ij} = \overline{\mu}_{..} + \alpha_i + \beta_j$,
- if it is ture, there is no interaction effect between factors A and B.
- if it is not ture, interaction effect exists and is defined as

$$(\alpha\beta)_{ij} = \mu_{ij} - \left(\overline{\mu}_{..} + \alpha_i + \beta_j\right) = \mu_{ij} - \overline{\mu}_{i.} - \overline{\mu}_{.j} + \overline{\mu}_{..}$$

Effect Model without interaction (additive model):

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

Effect Model with interaction (full model):

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Calculation of Interaction Effects

Table 7.1-3 Two Examples of Cell Means

Illustration 1

	Factor B			
Factor A	1	2	3	$\overline{\mu}_{i\bullet}$
1	1	2	6	3
2	3	4	8	5
$\overline{\mu}_{\bullet j}$	2	3	7	4

$\alpha_1 = 3 - 4 = -1$

$\alpha_2 = 5 - 4 = 1$

$\beta_1 = 2 - 4 = -2$

$\beta_2 = 3 - 4 = -1$

$\beta_3 = 7 - 4 = 3$

$(\alpha\beta)_{11} = 1 - 3 - 2 + 4 = 0$

$(\alpha\beta)_{12} = 2 - 3 - 3 + 4 = 0$

$(\alpha\beta)_{13} = 6 - 3 - 7 + 4 = 0$

$(\alpha\beta)_{21} = 3 - 5 - 2 + 4 = 0$

$(\alpha\beta)_{22} = 4 - 5 - 3 + 4 = 0$

$(\alpha\beta)_{23} = 8 - 5 - 7 + 4 = 0$

Illustration 2

	Factor B			
Factor A	1	2	3	$\overline{\mu}_{i\bullet}$
1	1	2	6	3
2	5	8	2	5
$\overline{\mu}_{\bullet j}$	3	5	4	4

$\alpha_1 = 3 - 4 = -1$

$\alpha_2 = 5 - 4 = 1$

$\beta_1 = 3 - 4 = -1$

$\beta_2 = 5 - 4 = 1$

$\beta_3 = 4 - 4 = 0$

$(\alpha\beta)_{11} = 1 - 3 - 3 + 4 = -1$

$(\alpha\beta)_{12} = 2 - 3 - 5 + 4 = -2$

$(\alpha\beta)_{13} = 6 - 3 - 4 + 4 = 3$

$(\alpha\beta)_{21} = 5 - 5 - 3 + 4 = 1$

$(\alpha\beta)_{22} = 8 - 5 - 5 + 4 = 2$

$(\alpha\beta)_{23} = 2 - 5 - 4 + 4 = -3$

Interaction Plot

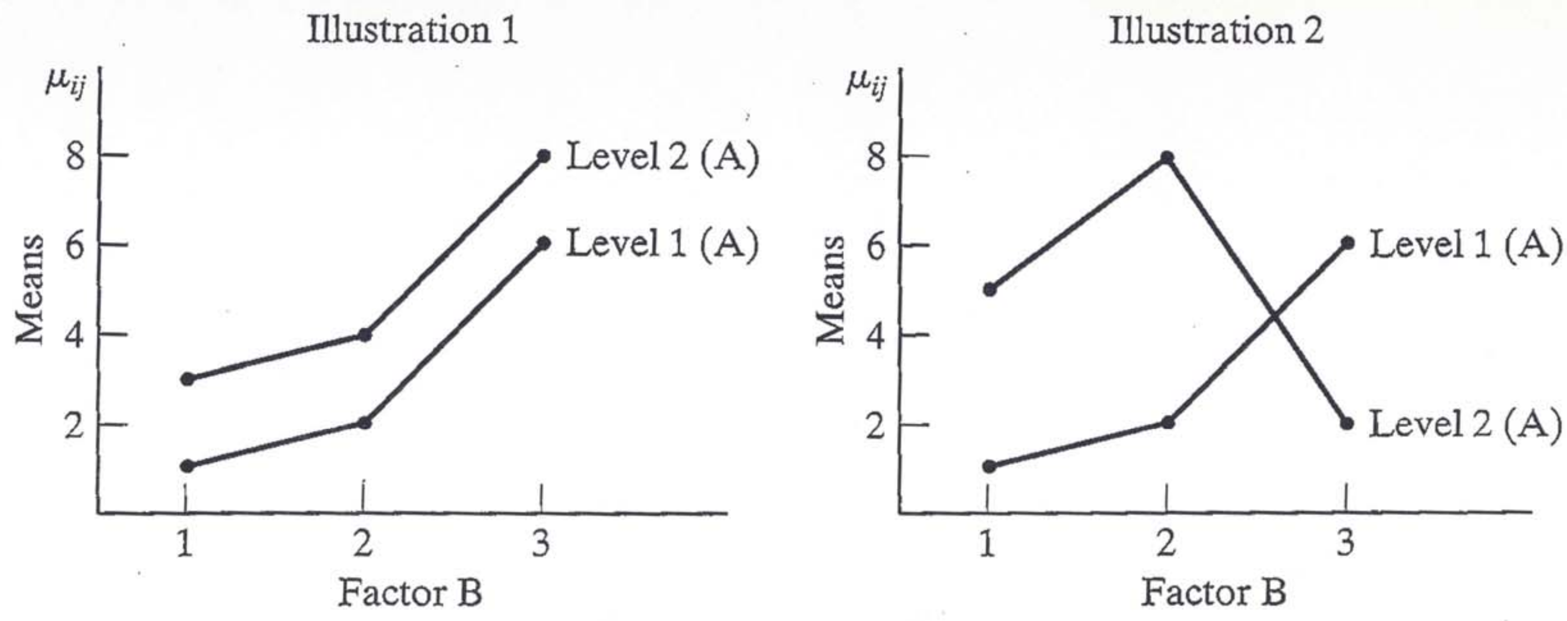


Figure 7.1-1 Population means for the two illustrations in Table 7.1-3

Observations with two-factor factorial design

Factor Levels	1 - B	2 - B	…	b - B
1 - A	$Y_{111}, \ldots, Y_{11n}$	$Y_{121}, \ldots, Y_{12n}$	…	$Y_{1b1}, \ldots, Y_{1bn}$
2 - A	$Y_{211}, \ldots, Y_{21n}$	$Y_{221}, \ldots, Y_{22n}$	…	$Y_{2b1}, \ldots, Y_{2bn}$
…	…	…	…	
a - A	$Y_{a11}, \ldots, Y_{a1n}$	$Y_{a21}, \ldots, Y_{a2n}$	…	$Y_{ab1}, \ldots, Y_{abn}$

Estimation of the means:

$$\hat{\mu}_{..} = \bar{Y}_{...} = \frac{1}{abn}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} Y_{ijk}, \qquad \hat{\mu}_{ij} = \bar{Y}_{ij.} = \frac{1}{n}\sum_{k=1}^{n} Y_{ijk},$$

$$\hat{\mu}_{i.} = \bar{Y}_{i..} = \frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n} Y_{ijk}, \qquad \hat{\mu}_{.j} = \bar{Y}_{.j.} = \frac{1}{an}\sum_{i=1}^{a}\sum_{k=1}^{n} Y_{ijk}$$

where $i = 1,.., a;\ j = 1,..., b.$

Example

- The Castle Bakery Company supplies wrapped Italian bread to a large number of supermarkets in a metropolitan area.
- An experimental study was made of the effects of heights of the shelf display (Factor A: bottom, middle, top) and the width of the shelf display (Factor B: regular, wide) on sales of this bakery's bread during the experimental period.
- Twelve supermarkets, similar in terms of sales volume and clients, were utilized in the study. The six treatments were assigned at random to two stores. Sales are presented in the following table.

	Factor B	
	Regular	**Wide**
Factor A		
bottom	47, 43	46, 40
middle	62, 68	67, 71
top	41, 39	42, 46

1. Deviation of response from the grand mean in (i, j) - th cell :

$$\left(Y_{ijk}-\overline{Y}_{...}\right)=\left(\overline{Y}_{ij.}-\overline{Y}_{...}\right)+\left(Y_{ijk}-\overline{Y}_{ij.}\right),$$

Take $\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\cdot)^2$ on both sides, SSTO = SSTR + SSE

2. Deviation Decomposition of the (i, j) - th cell mean :

$$\left(\overline{Y}_{ij.}-\overline{Y}_{...}\right)=\left(\overline{Y}_{i..}-\overline{Y}_{...}\right)+\left(\overline{Y}_{.j.}-\overline{Y}_{...}\right)+\left(\overline{Y}_{ij.}-\overline{Y}_{i..}-\overline{Y}_{.j.}+\overline{Y}_{...}\right)$$

$$\text{SSTR}=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\overline{Y}_{ij.}-\overline{Y}_{...}\right)^2=n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{Y}_{ij.}-\overline{Y}_{...}\right)^2$$

$$\text{SSA}=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\overline{Y}_{i..}-\overline{Y}_{...}\right)^2=bn\sum_{i=1}^{a}\left(\overline{Y}_{i..}-\overline{Y}_{...}\right)^2$$

$$\text{SSB}=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\overline{Y}_{.j.}-\overline{Y}_{...}\right)^2=an\sum_{j=1}^{b}\left(\overline{Y}_{.j.}-\overline{Y}_{...}\right)^2$$

$$\text{SSAB}=\text{SSTR}-\text{SSA}-\text{SSB}=\left(\text{SSTO}-\text{SSE}\right)-\text{SSA}-\text{SSB}$$

ANOVA for 2-factor factorial design

Source	SS	DF	MS	F
A	SSA	a-1	MSA = SSA / (a-1)	F_A = MSA / MSE
B	SSB	b-1	MSB = SSB / (b-1)	F_B = MSB / MSE
AB	SSAB	(a-1)(b-1)	MSAB = SSAB / ((a-1)(b-1))	F_{AB} = MSAB / MSE
Error	SSE	ab(n-1)	MSE = SSE/(ab(n-1))	
Total	SSTO	abn-1		

Expectations: $E(MSE) = \sigma^2,$

$$E(MSA) = \sigma^2 + \frac{bn}{(a-1)} \sum_{i=1}^{a} \alpha_i^2,$$

$$E(MSB) = \sigma^2 + \frac{an}{(b-1)} \sum_{j=1}^{b} \beta_j^2,$$

$$E(MSAB) = \sigma^2 + \frac{n}{(a-1)(b-1)} \sum_{j=1}^{b} (\alpha\beta)_{ij}^2.$$

Under $H_0 : \alpha_i = 0, \forall i = 1,\ldots,a,$

$$F_A \sim F((a-1), ab(n-1));$$

under $H_0 : \beta_j = 0, \forall j = 1,\ldots,b,$

$$F_B \sim F((b-1), ab(n-1));$$

under $H_0 : (\alpha\beta)_{ij} = 0, \forall i = 1,\ldots,a, j = 1,\ldots,b$

$$F_{AB} \sim F((a-1)(b-1), ab(n-1)).$$

Sampling Distribution of the Mean Estimators

Observed response:

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \overset{\text{ind.}}{\sim} N\left(\mu_{ij}, \sigma^2\right)$$

Means	Estimator	Sampling Distribution	Confidence Interval
(cell mean) μ_{ij}	$\bar{Y}_{ij\cdot} = \frac{1}{n}\sum_{k=1}^{n} Y_{ijk}$	$\sim N\left(\mu_{ij}, \frac{\sigma^2}{n}\right)$,	$\bar{Y}_{ij\cdot} \pm t_{\frac{\alpha}{2}}(ab(n-1))\sqrt{\frac{MSE}{n}}$
(trt A mean) $\bar{\mu}_{i\cdot}$	$\bar{Y}_{i\cdot\cdot} = \frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n} Y_{ijk}$	$\sim N\left(\bar{\mu}_{i\cdot}, \frac{\sigma^2}{bn}\right)$,	$\bar{Y}_{i\cdot\cdot} \pm t_{\frac{\alpha}{2}}(ab(n-1))\sqrt{\frac{MSE}{bn}}$
(trt B mean) $\bar{\mu}_{\cdot j}$	$\bar{Y}_{\cdot j\cdot} = \frac{1}{an}\sum_{i=1}^{a}\sum_{k=1}^{n} Y_{ijk}$	$\sim N\left(\bar{\mu}_{\cdot j}, \frac{\sigma^2}{an}\right)$,	$\bar{Y}_{\cdot j\cdot} \pm t_{\frac{\alpha}{2}}(ab(n-1))\sqrt{\frac{MSE}{an}}$

Follow-up Test and Simultaneous Comparison When Factors Do Not Interact

Pairwise Comparison

(1). Test $H_0 : \alpha_i = \alpha_{i'}$, vs. $H_1 : \alpha_i \neq \alpha_{i'}$,

C.I. for $(\bar{\mu}_{i.} - \bar{\mu}_{i'.})$: $\left(\bar{Y}_{i..} - \bar{Y}_{i'..}\right) \pm t_{\frac{\alpha}{2}}(ab(n-1))\sqrt{\frac{2 \cdot MSE}{bn}}$

(2). Test $H_0 : \beta_j = \beta_{j'}$, vs. $H_1 : \beta_j \neq \beta_{j'}$,

C.I. for $(\bar{\mu}_{.j} - \bar{\mu}_{.j'})$: $\left(\bar{Y}_{.j.} - \bar{Y}_{.j'.}\right) \pm t_{\frac{\alpha}{2}}(ab(n-1))\sqrt{\frac{2 \cdot MSE}{an}}$

Tukey's Simutaneous Comparison

(1). Test $H_0 : \alpha_i = \alpha_{i'}$, for any pair (i, i'), vs. H_1 : at least one pair $\alpha_i \neq \alpha_{i'}$,

Simultaneous C.I. for $(\bar{\mu}_{i.} - \bar{\mu}_{i'.})$: $\left(\bar{Y}_{i..} - \bar{Y}_{i'..}\right) \pm q_{\alpha}(a, ab(n-1))\sqrt{\frac{MSE}{bn}}, \forall (i, i')$

(2). Test $H_0 : \beta_j = \beta_{j'}$, for any pair (j, j'), vs. H_1 : at least one pair $\beta_j \neq \beta_{j'}$,

Simultaneous C.I. for $(\bar{\mu}_{.j} - \bar{\mu}_{.j'})$: $\left(\bar{Y}_{.j.} - \bar{Y}_{.j'.}\right) \pm q_{\alpha}(b, ab(n-1))\sqrt{\frac{MSE}{an}}, \forall (j, j')$

Multiple Comparison of Means When Factors Do Not Interact

Bonferroni Multiple Comparison (Simultaneous Confidence Interval)

(1). Test $H_0 : \alpha_i = \alpha_{i'}$, for any pair (i, i'), vs. H_1 : at least one pair $\alpha_i \neq \alpha_{i'}$,

S.C.I. for $(\bar{\mu}_{i\cdot} - \bar{\mu}_{i'\cdot})$: $\quad \left(\bar{Y}_{i\cdot\cdot} - \bar{Y}_{i'\cdot\cdot}\right) \pm t_{\frac{\alpha}{a(a-1)}}\left(ab(n-1)\right)\sqrt{\frac{2 \cdot MSE}{bn}}, \forall (i, i')$

(2). Test $H_0 : \beta_j = \beta_{j'}$, for any pair (j, j'), vs. H_1 : at least one pair $\beta_j \neq \beta_{j'}$,

S.C.I. for $(\bar{\mu}_{\cdot j} - \bar{\mu}_{\cdot j'})$: $\quad \left(\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdot j'\cdot}\right) \pm t_{\frac{\alpha}{b(b-1)}}\left(ab(n-1)\right)\sqrt{\frac{2 \cdot MSE}{an}}, \forall (j, j')$

Multiple Contrasts of Factor Level Means

(1). Contrast for Factor A means : $L_A = \sum_i a_i \mu_{i\cdot}$ subject to $\sum_i a_i = 0$.

(2). Contrast for Factor B means : $L_B = \sum_j b_j \mu_{\cdot j}$ subject to $\sum_j b_j = 0$.

(3). Scheffe' confidence limit for L_A : $\hat{L}_A \pm S_A \cdot se\left(\hat{L}_A\right)$,

where $\hat{L}_A = \sum_i a_i \bar{Y}_{i\cdot\cdot}$, $se^2\left(\hat{L}_A\right) = \frac{MSE}{bn}\sum_i a_i^2$, $S_A^2 = F_\alpha\left(a-1, ab(n-1)\right)$.

Analysis of Effects When Interactions Exist

(1). When important interactions exist, the analysis of factor effects generally must be based on the treatment means.

(2). One often compares the levels of one factor acroaa levels of the other factors, referring to the comparison of single effects. For example in a 2×3 factorial design, we compare individual cell means within levels of each factor, e.g. $\mu_{11} = \mu_{12} = \mu_{13}$, and $\mu_{21} = \mu_{22} = \mu_{23}$ or $\mu_{11} = \mu_{21}, \mu_{12} = \mu_{22}$, and $\mu_{13} = \mu_{23}$

(3). Test $H_0 : \mu_{ij} = \mu_{i'j'}$, for specific m pairs vs. H_1 : at least one pair $\mu_{ij} \neq \mu_{i'j'}$,

Bonferroni S.C.I. for $(\mu_{ij} - \mu_{i'j'})$: $\left(\bar{Y}_{ij\cdot} - \bar{Y}_{i'j'\cdot}\right) \pm t_{\frac{\alpha}{2m}}\left(ab(n-1)\right)\sqrt{\frac{2 \cdot MSE}{n}}$,

Test $H_0 : \mu_{ij} = \mu_{i'j'}$, for all pairs vs. H_1 : at least one pair $\mu_{ij} \neq \mu_{i'j'}$,

Tukey S.C.I. for $(\mu_{ij} - \mu_{i'j'})$: $\left(\bar{Y}_{ij\cdot} - \bar{Y}_{i'j'\cdot}\right) \pm q_{\alpha}\left(ab, ab(n-1)\right)\sqrt{\frac{MSE}{n}}, \forall (i, j) \neq (i', j')$.

(4). Multiple Contrasts of Treatment Means : $L = \sum_i \sum_j c_{ij} \mu_{ij}$, and $\sum_i \sum_j c_{ij} = 0$

where $\hat{L} = \sum_i \sum_j c_{ij} \bar{Y}_{ij\cdot}$, $\text{se}^2(\hat{L}) = \frac{MSE}{n} \sum_i \sum_j c_{ij}^2$,

Scheffe′ confidenc limits for L : $\hat{L} \pm S_L \cdot \text{se}(\hat{L}), S_L^2 = F_{\alpha}\left(ab-1, ab(n-1)\right)$

Example (Bakery)

- The Castle Bakery Company supplies wrapped Italian bread to a large number of supermarkets in a metropolitan area.
- An experimental study was made of the effects of heights of the shelf display (Factor A: bottom, middle, top) and the width of the shelf display (Factor B: regular, wide) on sales of this bakery's bread during the experimental period.
- Twelve supermarkets, similar in terms of sales volume and clientele, were utilized in the study. The six treatments were assigned at random to two stores. Sales are presented in the following table.

	Factor B	
	Regular	**Wide**
Factor A		
Bottom	47 43	46 40
Middle	62 68	67 71
Top	41 39	42 46

$$\text{Contrast}: \; c_1\alpha_1 + c_2\alpha_2 + \ldots + c_a\alpha_a \;\text{ with } \sum_{i=1}^{a} c_i = 0$$

$$\text{For example}: \; 0.5\alpha_1 - \alpha_2 + 0.5\alpha_3, \;\text{ or }\; \alpha_1 - \alpha_2 + \alpha_3 - \alpha_4$$

SAS code - proc glm for 2-way anova

```
/* Two-way ANOVA  (full model)*/
proc glm data=display;
   class height width;
   model sale = height width height*width;
  lsmeans height / stderr pdiff adjust=tukey;
   lsmeans width / stderr pdiff adjust=tukey;
   contrast "regular vs wide" width -1 1;
   contrast "middle vs the others" height 0.5 -1 0.5;
  run;
/* Two-way ANOVA  (reduced model)*/
proc glm data=display;
   class height width;
   model sale = height width;
   lsmeans height / stderr pdiff adjust=tukey;
   lsmeans width / stderr pdiff adjust=tukey;
   contrast "regular vs wide" width -1 1;
   contrast "middle vs the others" height 0.5 -1 0.5;
   run;
```

```
/*Interaction Plot  -  Method 1*/
proc sort data=display;
     by height width;       run;

proc means data=display noprint;
   by height width;
   var sale;
   output out=means mean=meany;
   run;

symbol1 v=r  i=join  c=black;
symbol2 v=w  i=join  c=blue;

proc gplot data=means;
   plot meany*height=width;     run;

/* Model Checking*/
```

Example Suppose you want to determine whether the brand of laundry detergent used and the temperature affects the amount of dirt removed from your laundry. To this end, you buy two different brand of detergent (" Super" and "Best") and choose three different temperature levels ("cold", "warm", and "hot"). Then you divide your laundry randomly into $6 \times r$ piles of equal size and assign each r piles into the combination of ("Super" and "Best") and ("cold","warm", and "hot"). In this example, we are interested in testing Null Hypotheses

$$\begin{aligned} H_{0D} &: \text{The amount of dirt removed does not depend on the type of detergent} \\ H_{0T} &: \text{The amount of dirt removed does not depend on the temperature} \end{aligned}$$

One says the experiment has **two factors** (Factor Detergent, Factor Temperature) at $a = 2$(Super and Best) and $b = 3$(cold,warm and hot) **levels**. Thus there are $ab = 3 \times 2 = 6$ different combinations of detergent and temperature. With each combination you wash $r = 4$ loads. r is called the number of **replicates**. This sums up to $n = abr = 24$ loads in total. The amounts Y_{ijk} of dirt removed when washing sub pile k ($k = 1, 2, 3, 4$) with detergent i ($i = 1, 2$) at temperature j ($j = 1, 2, 3$) are recorded in Table 1.

	Cold	Warm	Hot
Super	4,5,6,5	7,9,8,12	10,12,11,9
Best	6,6,4,4	13,15,12,12	12,13,10,13

Solution :

	Cold	Warm	Hot	m_D
Super	4,5,6,5 (5)	7,9,8,12 (9)	10,12,11,9 (10)	8
Best	6,6,4,4 (5)	13,15,12,12 (13)	12,13,10,13 (12)	10
m_T	5	11	11	9

```
/* SAS Code  -  Detergent Data*/

proc glm data = detergent;
   class brand temp;
   model dirt = brand temp brand*temp;
   run;

/* IF the interaction effect is significant */

proc glm data=detergent;
   class brand temp;
   model dirt = brand temp brand*temp;
   lsmeans  brand*temp
        / pdiff stderr adjust=TUKEY;
   run;
```

```
/*Interaction Plot  -  Method 2*/
proc glm data=detergent;
   class brand temp;
   model dirt = brand | temp;      /*  including interactions*/
   lsmeans brand * temp / out=lsm;
   output out=diag p=py r=ry;    /* p=predicted value, r=residual */
   run;

symbol1 v = b   i = join   c = black;
symbol2 v = s   i = join   c = blue;

proc gplot data=lsm;                    /* Interaction Plot */
   plot lsmean*temp=brand;       /* cell mean vs temp by brand */
   run;
proc plot data=diag;                    /* Diagnostic Plots */
   plot dirt*py;                       /* observed v. predicted */
   plot py*ry;       run;

proc univariate data=diag normal;   /* Model Checking*/
   var ry;
   qqplot;
   run;
```