7. Analysis of Variance for Balanced Two-Factor Experiments

An Example Two-Factor Experiment

Researchers were interested in studying the effects of 2 diets (low fiber, high fiber) and 3 drugs (D1, D2, D3) on weight gained by Yorkshire pigs. A total of 12 pigs were assigned to the 6 diet-drug combinations using a balanced and completely randomized experimental design. Pigs were housed in individual pens, injected with their assigned drugs once per week, and fed their assigned diets for a 6-week period. The amount of weight gained during the 6-week period was recorded for each pig.

A Model for the Data

For $i = 1, 2;\ j = 1, 2, 3;$ and $k = 1, 2$; let y_{ijk} denote the weight gain of the k^{th} pig that received diet i and drug j, and suppose

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad (i = 1, 2;\ j = 1, 2, 3;\ k = 1, 2)$$ where

$$\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3, \gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23}$$

are unknown real-valued parameters and

$$\epsilon_{111}, \epsilon_{112}, \epsilon_{121}, \epsilon_{122}, \epsilon_{131}, \epsilon_{132}, \epsilon_{211}, \epsilon_{212}, \epsilon_{221}, \epsilon_{222}, \epsilon_{231}, \epsilon_{232} \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

for some unknown $\sigma^2 > 0$.

Model in Matrix and Vector Form

$$
\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{131} \\ y_{132} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{231} \\ y_{232} \end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \end{bmatrix}
+
\begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{231} \\ \epsilon_{232} \end{bmatrix}
$$

$$
\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})
$$

Same Model with Cell Means Parameterization

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{131} \\ y_{132} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \\ y_{231} \\ y_{232} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{131} \\ \epsilon_{132} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \\ \epsilon_{231} \\ \epsilon_{232} \end{bmatrix}$$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

We could consider a sequence of progressively more complex models for the response mean that lead up to our full cell means model.

1. $E(y_{ijk}) = \mu$
2. $E(y_{ijk}) = \mu + \alpha_i$
3. $E(y_{ijk}) = \mu + \alpha_i + \beta_j$
4. $E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij} \iff E(y_{ijk}) = \mu_{ij}$

Matrices with Nested Column Spaces

$$\boldsymbol{X}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{X}_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \boldsymbol{X}_3 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \text{ and}$$

Matrices with Nested Column Spaces

$$X_4 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ or}$$

Matrices with Nested Column Spaces

$$
\boldsymbol{X}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.
$$

ANOVA Table for Our Two-Factor Example

Source	Sum of Squares	DF
$\text{Diets}\|1$	$\boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}$	$2 - 1 = 1$
$\text{Drugs}\|1, \text{Diets}$	$\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}$	$4 - 2 = 2$
$\text{Diets} \times \text{Drugs}\|1, \text{Diets}, \text{Drugs}$	$\boldsymbol{y}'(\boldsymbol{P}_4 - \boldsymbol{P}_3)\boldsymbol{y}$	$6 - 4 = 2$
Error	$\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_4)\boldsymbol{y}$	$12 - 6 = 6$
C. Total	$\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y}$	$12 - 1 = 11$

ANOVA Table for Our Two-Factor Example

Source	Sum of Squares	DF
Diets	$\boldsymbol{y}'(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{y}$	$2 - 1 = 1$
Drugs	$\boldsymbol{y}'(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{y}$	$4 - 2 = 2$
Diets $\times$ Drugs	$\boldsymbol{y}'(\boldsymbol{P}_4 - \boldsymbol{P}_3)\boldsymbol{y}$	$6 - 4 = 2$
Error	$\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_4)\boldsymbol{y}$	$12 - 6 = 6$
C. Total	$\boldsymbol{y}'(\boldsymbol{I} - \boldsymbol{P}_1)\boldsymbol{y}$	$12 - 1 = 11$

The Diet-Drug Dataset

```
> d
   diet drug weightgain
1     1    1       41.3
2     1    1       43.7
3     1    2       40.9
4     1    2       39.2
5     1    3       37.4
6     1    3       37.9
7     2    1       36.8
8     2    1       34.6
9     2    2       33.6
10    2    2       34.3
11    2    3       35.8
12    2    3       35.1
```

R Code and Output for Two-Factor ANOVA

```
> d$diet=factor(d$diet)
> d$drug=factor(d$drug)
>
> a=d$diet
> b=d$drug
> y=d$weightgain
```

R Code and Output for Two-Factor ANOVA

```
> x1=matrix(1,nrow=nrow(d),ncol=1)
> x1
      [,1]
 [1,]    1
 [2,]    1
 [3,]    1
 [4,]    1
 [5,]    1
 [6,]    1
 [7,]    1
 [8,]    1
 [9,]    1
[10,]    1
[11,]    1
[12,]    1
```

R Code and Output for Two-Factor ANOVA

```
> x2=cbind(x1,model.matrix(~0+a))
> x2
   x1 a1 a2
1   1  1  0
2   1  1  0
3   1  1  0
4   1  1  0
5   1  1  0
6   1  1  0
7   1  0  1
8   1  0  1
9   1  0  1
10  1  0  1
11  1  0  1
12  1  0  1
```

R Code and Output for Two-Factor ANOVA

```
> x3=cbind(x2,model.matrix(~0+b))
> x3
   x1 a1 a2 b1 b2 b3
1   1  1  0  1  0  0
2   1  1  0  1  0  0
3   1  1  0  0  1  0
4   1  1  0  0  1  0
5   1  1  0  0  0  1
6   1  1  0  0  0  1
7   1  0  1  1  0  0
8   1  0  1  1  0  0
9   1  0  1  0  1  0
10  1  0  1  0  1  0
11  1  0  1  0  0  1
12  1  0  1  0  0  1
```

R Code and Output for Two-Factor ANOVA

```
> x4=model.matrix(~0+b:a)
> x4
   b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
1      1     0     0     0     0     0
2      1     0     0     0     0     0
3      0     1     0     0     0     0
4      0     1     0     0     0     0
5      0     0     1     0     0     0
6      0     0     1     0     0     0
7      0     0     0     1     0     0
8      0     0     0     1     0     0
9      0     0     0     0     1     0
10     0     0     0     0     1     0
11     0     0     0     0     0     1
12     0     0     0     0     0     1
```

R Code and Output for Two-Factor ANOVA

```
> library(MASS)
> proj=function(x){
+   x%*%ginv(t(x)%*%x)%*%t(x)
+ }
>
> p1=proj(x1)
> p2=proj(x2)
> p3=proj(x3)
> p4=proj(x4)
> I=diag(rep(1,12))
```

R Code and Output for Two-Factor ANOVA

```
> SumOfSquares=c(
+ t(y)%*%(p2-p1)%*%y,
+ t(y)%*%(p3-p2)%*%y,
+ t(y)%*%(p4-p3)%*%y,
+ t(y)%*%(I-p4)%*%y,
+ t(y)%*%(I-p1)%*%y)
>
> Source=c(
+ "Diet|1",
+ "Drug|1,Diet",
+ "Diet x Drug|1,Diet,Drug",
+ "Error",
+ "C. Total")
```

R Code and Output for Two-Factor ANOVA

```
> data.frame(Source,SumOfSquares)
                      Source SumOfSquares
1                     Diet|1     76.00333
2                Drug|1,Diet     14.82000
3 Diet x Drug|1,Diet,Drug     12.28667
4                      Error      7.36000
5                   C. Total    110.47000
```

R Code and Output for Two-Factor ANOVA

```
> o=lm(weightgain~diet+drug+diet:drug,data=d)
>
> anova(o)
Analysis of Variance Table

Response: weightgain
          Df Sum Sq Mean Sq F value    Pr(>F)
diet       1 76.003  76.003 61.9592 0.0002226 ***
drug       2 14.820   7.410  6.0408 0.0365383 *
diet:drug  2 12.287   6.143  5.0082 0.0525735 .
Residuals  6  7.360   1.227
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

What do the F-tests in this ANOVA table test?

Recall the null hypothesis for F_j is true if and only if

$$\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}\boldsymbol{\beta} = 0.$$

We have the following equivalent conditions

$$\begin{aligned}
\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}\boldsymbol{\beta} = 0 \quad &\iff \quad \boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)'(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}\boldsymbol{\beta} = 0 \\
&\iff \quad ||\ (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}\boldsymbol{\beta}\ ||^2 = 0 \\
&\iff \quad (\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}\boldsymbol{\beta} = \mathbf{0} \\
&\iff \quad \boldsymbol{C}\boldsymbol{\beta} = \mathbf{0},
\end{aligned}$$

where $\boldsymbol{C}$ is any full-row-rank matrix with the same row space as $(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}$.

What do the F-tests in this ANOVA table test?

Let's take a look at $(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}$ for each test in the ANOVA table.

When computing $(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}$, we can use any model matrix $\boldsymbol{X}$ that specifies one unrestricted treatment mean for each of the six treatments.

The entries in any rows of $(\boldsymbol{P}_{j+1} - \boldsymbol{P}_j)\boldsymbol{X}$ are coefficients defining linear combinations of the elements of the parameter vector $\boldsymbol{\beta}$ that corresponds to the chosen model matrix $\boldsymbol{X}$.

Our Choice for $\boldsymbol{X}$ and $\boldsymbol{\beta}$

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix}$$

ANOVA Diet Test

```
> x=x4
> fractions((p2-p1)%*%x)
   b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
1   1/6   1/6   1/6  -1/6  -1/6  -1/6
2   1/6   1/6   1/6  -1/6  -1/6  -1/6
3   1/6   1/6   1/6  -1/6  -1/6  -1/6
4   1/6   1/6   1/6  -1/6  -1/6  -1/6
5   1/6   1/6   1/6  -1/6  -1/6  -1/6
6   1/6   1/6   1/6  -1/6  -1/6  -1/6
7  -1/6  -1/6  -1/6   1/6   1/6   1/6
8  -1/6  -1/6  -1/6   1/6   1/6   1/6
9  -1/6  -1/6  -1/6   1/6   1/6   1/6
10 -1/6  -1/6  -1/6   1/6   1/6   1/6
11 -1/6  -1/6  -1/6   1/6   1/6   1/6
12 -1/6  -1/6  -1/6   1/6   1/6   1/6
```

ANOVA Diet Test

$$(\boldsymbol{P}_2 - \boldsymbol{P}_1)\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0} \Longleftrightarrow \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{0},$$

where

$$\boldsymbol{C}\boldsymbol{\beta} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} = \bar{\mu}_{1\cdot} - \bar{\mu}_{2\cdot}.$$

ANOVA Drug Test

```
> fractions((p3-p2)%*%x)
   b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
1    1/3  -1/6  -1/6   1/3  -1/6  -1/6
2    1/3  -1/6  -1/6   1/3  -1/6  -1/6
3   -1/6   1/3  -1/6  -1/6   1/3  -1/6
4   -1/6   1/3  -1/6  -1/6   1/3  -1/6
5   -1/6  -1/6   1/3  -1/6  -1/6   1/3
6   -1/6  -1/6   1/3  -1/6  -1/6   1/3
7    1/3  -1/6  -1/6   1/3  -1/6  -1/6
8    1/3  -1/6  -1/6   1/3  -1/6  -1/6
9   -1/6   1/3  -1/6  -1/6   1/3  -1/6
10  -1/6   1/3  -1/6  -1/6   1/3  -1/6
11  -1/6  -1/6   1/3  -1/6  -1/6   1/3
12  -1/6  -1/6   1/3  -1/6  -1/6   1/3
```

ANOVA Drug Test

```
> p3p2x=(p3-p2)%*%x
>
> fractions(p3p2x[1,]-p3p2x[3,])
b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
  1/2  -1/2     0   1/2  -1/2     0
>
> fractions(p3p2x[1,]-p3p2x[5,])
b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
  1/2     0  -1/2   1/2     0  -1/2
```

ANOVA Drug Test

$$(\boldsymbol{P}_3 - \boldsymbol{P}_2)\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0} \iff \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{0},$$

where

$$\begin{aligned}
\boldsymbol{C}\boldsymbol{\beta} &= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} \\
&= \begin{bmatrix} \bar{\mu}_{\cdot 1} - \bar{\mu}_{\cdot 2} \\ \bar{\mu}_{\cdot 1} - \bar{\mu}_{\cdot 3} \end{bmatrix}.
\end{aligned}$$

ANOVA Test for Diet $\times$ Drug Interactions

```
> fractions((p4-p3)%*%x)
   b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
1   1/3  -1/6  -1/6  -1/3   1/6   1/6
2   1/3  -1/6  -1/6  -1/3   1/6   1/6
3  -1/6   1/3  -1/6   1/6  -1/3   1/6
4  -1/6   1/3  -1/6   1/6  -1/3   1/6
5  -1/6  -1/6   1/3   1/6   1/6  -1/3
6  -1/6  -1/6   1/3   1/6   1/6  -1/3
7  -1/3   1/6   1/6   1/3  -1/6  -1/6
8  -1/3   1/6   1/6   1/3  -1/6  -1/6
9   1/6  -1/3   1/6  -1/6   1/3  -1/6
10  1/6  -1/3   1/6  -1/6   1/3  -1/6
11  1/6   1/6  -1/3  -1/6  -1/6   1/3
12  1/6   1/6  -1/3  -1/6  -1/6   1/3
```

ANOVA Test for Diet × Drug Interactions

```
> p4p3x=(p4-p3)%*%x
>
> fractions(2*(p4p3x[1,]-p4p3x[3,]))
b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
    1    -1     0    -1     1     0
>
> fractions(2*(p4p3x[1,]-p4p3x[5,]))
b1:a1 b2:a1 b3:a1 b1:a2 b2:a2 b3:a2
    1     0    -1    -1     0     1
>
```

ANOVA Test for Diet × Drug Interactions

$$(\boldsymbol{P}_4 - \boldsymbol{P}_3)\boldsymbol{X\beta} = \boldsymbol{0} \iff \boldsymbol{C\beta} = \boldsymbol{0},$$

where

$$\begin{aligned}
\boldsymbol{C\beta} &= \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \\ \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{bmatrix} \\
&= \begin{bmatrix} \mu_{11} - \mu_{12} - \mu_{21} + \mu_{22} \\ \mu_{11} - \mu_{13} - \mu_{21} + \mu_{23} \end{bmatrix}.
\end{aligned}$$

ANOVA for Balanced Two-Factor Experiments

The diet-drug experiment is *balanced* in the sense that every treatment (defined by a diet-drug combination) has the same number of experimental units.

Each experimental unit provided a single response measurement (weight gain), so the resulting diet-drug dataset is *balanced* in the sense that each treatment has the same number of independent, constant variance observations of the response.

Due to this balance, the tests for diets, drugs, and diets $\times$ drugs in the ANOVA table turn out to be exactly the same as the tests for diet main effect, drug main effects, and diet $\times$ drug interactions we discussed previously as tests of $\boldsymbol{C\beta} = \mathbf{0}$.