23. Repeated Measures

Repeated Measures Example

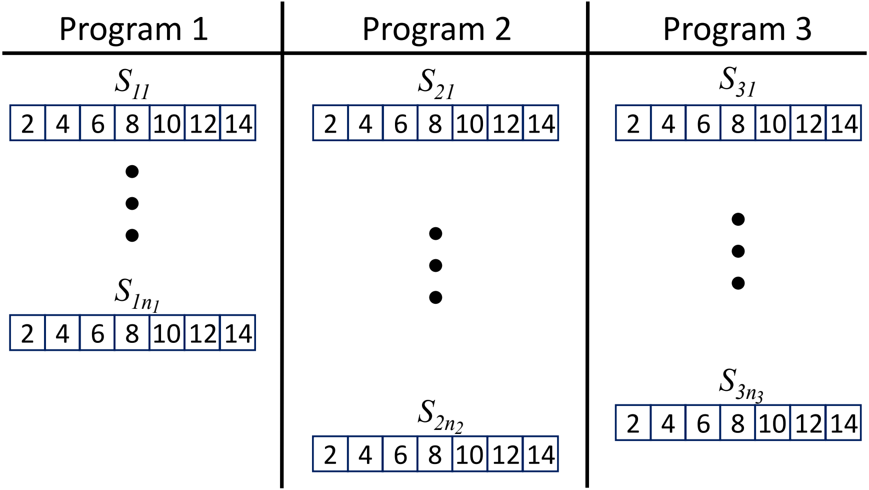
In an exercise therapy study, subjects were assigned to one of three weightlifting programs

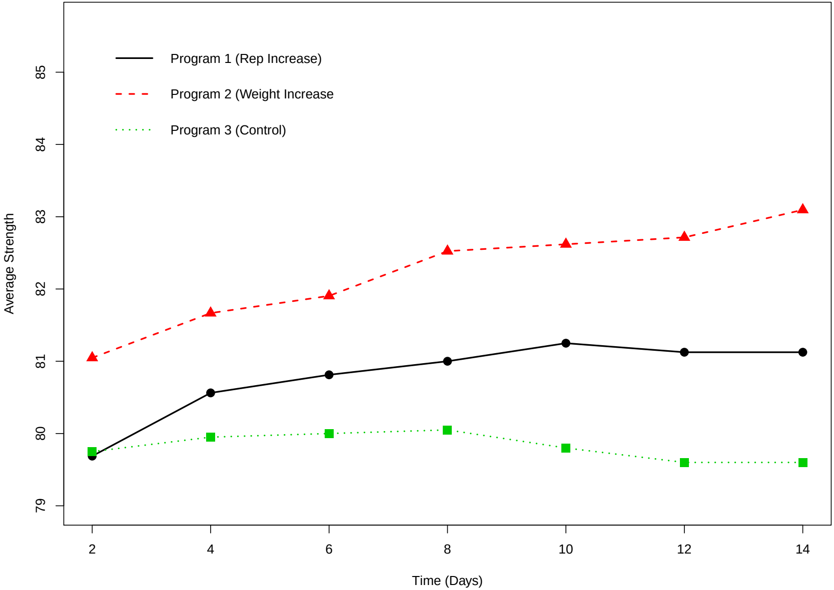
i=1: The number of repetitions of weightlifting was increased as subjects became stronger.

i=2: The amount of weight was increased as subjects became stronger.

i=3: Subjects did not participate in weightlifting.

- Measurements of strength (y) were taken on days 2, 4, 6, 8, 10, 12, and 14 for each subject.
- Source: Littel, Freund, and Spector (1991), SAS System for Linear Models.
- R code: RepeatedMeasures.R
- SAS code: RepeatedMeasures.sas


Program 1
Program 2
Program 3
S_{11}
2 4 6 8 10 12 14
S_{1n_1}
2 4 6 8 10 12 14
S_{21}
2 4 6 8 10 12 14
S_{2n_2}
2 4 6 8 10 12 14
S_{31}
2 4 6 8 10 12 14
S_{3n_3}
2 4 6 8 10 12 14



Program 1 (Rep Increase)
Program 2 (Weight Increase
Program 3 (Control)
Average Strength
79
80
81
82
83
84
85
Time (Days)
2
4
6
8
10
12
14

A Linear Mixed-Effects Model

Let y_{ijk} be the strength measurement for program i, subject j, and time point k. Suppose

$$y_{ijk} = \mu + \alpha_i + s_{ij} + \tau_k + \gamma_{ik} + e_{ijk},$$

where μ, $\alpha_1, \alpha_2, \alpha_3$, $\tau_1, \dots, \tau_7$, and $\gamma_{11}, \dots, \gamma_{37}$ are unknown real-valued parameters, and

$$s_{ij} \overset{iid}{\sim} N(0, \sigma_s^2) \text{ independent of } e_{ijk} \overset{iid}{\sim} N(0, \sigma_e^2).$$

Note that this model is the same model we would use for a split-plot experiment in which the whole-plot part of the experiment has a completely randomized design.

Subjects are the whole-plot experimental units, and measurement occasions within subject are treated like split-plot experimental units.

ANOVA Table

Source	DF
Program	$3-1$
Subject(Program)	$(16-1)+(21-1)+(20-1)$
Time	$7-1$
Program$\times$Time	$(3-1)(7-1)$
Time$\times$Subject(Program)	$(7-1)(57-3)$
C. Total	$57 \times 7 - 1$

The measurement occasions are not really split-plot experimental units because levels of the factor time (2, 4, . . ., 14) were not randomly assigned to measurement occasions.

Nonetheless, this split-plot model might be reasonable for some experiments where experimental units are measured repeatedly over time.

Average strength after $2k$ days on the ith program is

$$\begin{aligned} E(y_{ijk}) &= E(\mu + \alpha_i + s_{ij} + \tau_k + \gamma_{ik} + e_{ijk}) \\ &= \mu + \alpha_i + E(s_{ij}) + \tau_k + \gamma_{ik} + E(e_{ijk}) \\ &= \mu + \alpha_i + \tau_k + \gamma_{ik} \end{aligned}$$

for $i = 1, 2, 3$ and $k = 1, 2, \ldots, 7$.

The variance of any single observation is

$$\begin{aligned}
\mathrm{Var}(y_{ijk}) &= \mathrm{Var}(\mu + \alpha_i + s_{ij} + \tau_k + \gamma_{ik} + e_{ijk}) \\
&= \mathrm{Var}(s_{ij} + e_{ijk}) \\
&= \mathrm{Var}(s_{ij}) + \mathrm{Var}(e_{ijk}) \\
&= \sigma_s^2 + \sigma_e^2.
\end{aligned}$$

The covariance between any two different observations from the same subject is

$$
\begin{aligned}
\text{Cov}(y_{ijk}, y_{ij\ell}) &= \text{Cov}(\mu + \alpha_i + s_{ij} + \tau_k + \gamma_{ik} + e_{ijk}, \\
&\qquad \mu + \alpha_i + s_{ij} + \tau_\ell + \gamma_{i\ell} + e_{ij\ell}) \\
&= \text{Cov}(s_{ij} + e_{ijk}, s_{ij} + e_{ij\ell}) \\
&= \text{Cov}(s_{ij}, s_{ij}) + \text{Cov}(s_{ij}, e_{ij\ell}) \\
&\quad + \text{Cov}(e_{ijk}, s_{ij}) + \text{Cov}(e_{ijk}, e_{ij\ell}) \\
&= \text{Var}(s_{ij}) = \sigma_s^2.
\end{aligned}
$$

The correlation between y_{ijk} and $y_{ij\ell}$ is

$$\frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2} \equiv \rho.$$

Observations taken on different subjects are uncorrelated.

For the set of observations taken on a single subject, we have

$$\mathrm{Var}\left(\begin{bmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ij7} \end{bmatrix}\right) = \begin{bmatrix} \sigma_e^2+\sigma_s^2 & \sigma_s^2 & \cdots & \sigma_s^2 \\ \sigma_s^2 & \sigma_e^2+\sigma_s^2 & \cdots & \sigma_s^2 \\ \vdots & \vdots & \ddots & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 & \sigma_e^2+\sigma_s^2 \end{bmatrix}$$

$$= \sigma_e^2 \boldsymbol{I}_{7\times 7} + \sigma_s^2 \mathbf{1}\mathbf{1}'_{7\times 7}.$$

This is known as a *compound symmetric* covariance structure.

Using n_i to denote the number of subjects in the ith program, we can write this model in the form

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}.$$

To make things slightly easier to write, let

$$\boldsymbol{y}_{ij} = [y_{ij1}, y_{ij2}, y_{ij3}, y_{ij4}, y_{ij5}, y_{ij6}, y_{ij7}]'$$

and

$$\boldsymbol{e}_{ij} = [e_{ij1}, e_{ij2}, e_{ij3}, e_{ij4}, e_{ij5}, e_{ij6}, e_{ij7}]'$$

for all $i = 1, 2, 3$ and all $j = 1, \ldots, n_i$.

$$
\begin{bmatrix} \boldsymbol{y}_{11} \\ \boldsymbol{y}_{12} \\ \vdots \\ \boldsymbol{y}_{1n_1} \\ \boldsymbol{y}_{21} \\ \boldsymbol{y}_{22} \\ \vdots \\ \boldsymbol{y}_{2n_2} \\ \boldsymbol{y}_{31} \\ \boldsymbol{y}_{32} \\ \vdots \\ \boldsymbol{y}_{3n_3} \end{bmatrix}
=
\begin{bmatrix}
\mathbf{1}_{7\times 1} & \mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{0}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \mathbf{0}_{7\times 7} \\
\mathbf{1}_{7\times 1} & \mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{0}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \mathbf{0}_{7\times 7} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{1}_{7\times 1} & \mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{0}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \mathbf{0}_{7\times 7} \\
\mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} \\
\mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} \\
\mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{1}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \mathbf{0}_{7\times 7} & \boldsymbol{I}_{7\times 7} \\
\mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{1}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \mathbf{0}_{7\times 7} & \boldsymbol{I}_{7\times 7} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{1}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{0}_{7\times 1} & \mathbf{1}_{7\times 1} & \boldsymbol{I}_{7\times 7} & \mathbf{0}_{7\times 7} & \mathbf{0}_{7\times 7} & \boldsymbol{I}_{7\times 7}
\end{bmatrix}
\begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \\ \tau_6 \\ \tau_7 \\ \gamma_{11} \\ \gamma_{12} \\ \vdots \\ \gamma_{37} \end{bmatrix}
+
$$

$$\left[\boldsymbol{I}_{(n_1+n_2+n_3)\times(n_1+n_2+n_3)} \otimes \mathbf{1}_{7\times 1}\right] \begin{bmatrix} s_{11} \\ s_{12} \\ \vdots \\ s_{1n_1} \\ s_{21} \\ s_{22} \\ \vdots \\ s_{2n_2} \\ s_{31} \\ s_{32} \\ \vdots \\ s_{3n_3} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_{11} \\ \boldsymbol{e}_{12} \\ \vdots \\ \boldsymbol{e}_{1n_1} \\ \boldsymbol{e}_{21} \\ \boldsymbol{e}_{22} \\ \vdots \\ \boldsymbol{e}_{2n_2} \\ \boldsymbol{e}_{31} \\ \boldsymbol{e}_{32} \\ \vdots \\ \boldsymbol{e}_{3n_3} \end{bmatrix}$$

In this case,

$$\boldsymbol{G} = \text{Var}(\boldsymbol{u}) \;=\; \sigma_s^2 \boldsymbol{I}_{n.\times n.} \text{ and}$$

$$\boldsymbol{R} = \text{Var}(\boldsymbol{e}) \;=\; \sigma_e^2 \boldsymbol{I}_{(7n.)\times(7n.)},$$

where $n. = n_1 + n_2 + n_3$ is the total number of subjects.

$$\boldsymbol{\Sigma} = \text{Var}(\boldsymbol{y}) \;=\; \boldsymbol{ZGZ}' + \boldsymbol{R}$$

is a block diagonal matrix with one block of the form

$$\sigma_s^2 \mathbf{1}\mathbf{1}'_{7\times 7} + \sigma_e^2 \boldsymbol{I}_{7\times 7}$$

for each subject.

$\boldsymbol{G} = \mathrm{Var}(\boldsymbol{u}) = \sigma_s^2 \boldsymbol{I}$, $\boldsymbol{R} = \mathrm{Var}(\boldsymbol{e}) = \sigma_e^2 \boldsymbol{I}$, and

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \mathrm{Var}(\boldsymbol{y}) = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' + \boldsymbol{R} \\
&= \sigma_s^2 \begin{bmatrix} \mathbf{1}\mathbf{1}' & & & \\ & \mathbf{1}\mathbf{1}' & & \\ & & \ddots & \\ & & & \mathbf{1}\mathbf{1}' \end{bmatrix} + \sigma_e^2 \boldsymbol{I} \\
&= \begin{bmatrix} \sigma_e^2 \boldsymbol{I} + \sigma_s^2 \mathbf{1}\mathbf{1}' & & \\ & \ddots & \\ & & \sigma_e^2 \boldsymbol{I} + \sigma_s^2 \mathbf{1}\mathbf{1}' \end{bmatrix}.
\end{aligned}
$$

If predicting subject effects is not of interest and random subject effects are included only to introduce correlation among repeated measures on the same subject, we can work with an alternative expression of the same model by using the general linear model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$, where

$$\mathrm{Var}(\boldsymbol{y}) = \mathrm{Var}(\boldsymbol{\epsilon}) = \begin{bmatrix} \sigma_e^2\boldsymbol{I} + \sigma_s^2\mathbf{1}\mathbf{1}' & & \\ & \ddots & \\ & & \sigma_e^2\boldsymbol{I} + \sigma_s^2\mathbf{1}\mathbf{1}' \end{bmatrix}.$$

More generally, we can replace the mixed model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}$$

with the model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where

$$\mathrm{Var}(\boldsymbol{y}) = \mathrm{Var}(\boldsymbol{\epsilon}) = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{W} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{W} \end{bmatrix}.$$

- We can choose a structure for $\boldsymbol{W}$ that seems appropriate based on the design and the data.

- One choice for $\boldsymbol{W}$ is a compound symmetric matrix like we have considered previously.

- Another choice for $\boldsymbol{W}$ is an unstructured positive definite matrix.

- A common choice for $\boldsymbol{W}$ when repeated measures are equally spaced in time is the first-order autoregressive structure known as AR(1).

AR(1): First-Order Autoregressive Covariance Structure

$$
\boldsymbol{W} = \sigma^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 & \phi^6 \\ \phi & 1 & \phi & \phi^2 & \phi^3 & \phi^4 & \phi^5 \\ \phi^2 & \phi & 1 & \phi & \phi^2 & \phi^3 & \phi^4 \\ \phi^3 & \phi^2 & \phi & 1 & \phi & \phi^2 & \phi^3 \\ \phi^4 & \phi^3 & \phi^2 & \phi & 1 & \phi & \phi^2 \\ \phi^5 & \phi^4 & \phi^3 & \phi^2 & \phi & 1 & \phi \\ \phi^6 & \phi^5 & \phi^4 & \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}
$$

where $\sigma^2 \in (0, \infty)$ and $\phi \in (-1, 1)$ are unknown parameters.

In the next set of slides, we will see how to fit a variety of general linear models that might be appropriate for repeated measures experiments.