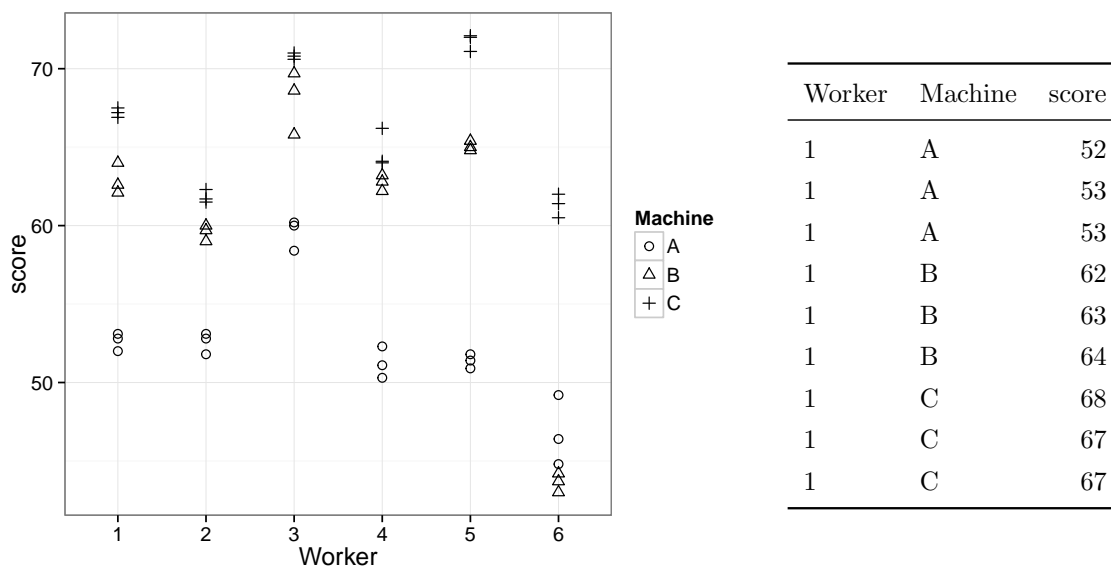


Aufgabe 1

(a) Machen Sie sich zunächst mit dem `Machines`-Datensatz aus dem R-Paket `nlme` vertraut (`?Machines`).

Die folgende Abbildung veranschaulicht den `Machines`-Datensatz und zeigt den Produktivitätsscores auf drei Maschinen getrennt nach Arbeiter (links) sowie Daten für Arbeiter 1 (rechts).



- (b) Handelt es sich um ein gekreuztes oder ein genestetes Design? Begründen Sie Ihre Antwort.
- (c) Sie möchten nun den Produktivitätsscore in Abhängigkeit der Maschine und der Arbeiter modellieren. Warum bieten sich dazu sowohl feste als auch zufällige Effekte an?
- (d) Stellen Sie ein Modell mit festen Effekten für die Maschine und zufälligen Effekten für die Arbeiter auf. Geben Sie die Modellgleichung für eine Beobachtung y_{ijk} an. Dabei ist y_{ijk} der Score für Arbeiter i an Maschine j in der k -ten Wiederholung. Vergessen Sie dabei nicht Verteilungsannahmen anzugeben.
- (e) Schreiben Sie das Modell aus (d) für den Arbeiter i auf, also die Modellgleichung für den Vektor $\mathbf{y}_i = (y_{i11}, \dots, y_{i33})'$.
- (f) Schreiben Sie die Modellgleichung aus (d) und (e) in Matrixnotation für den gesamten Datensatz auf. Geben Sie auch die Kovarianzmatrizen für die zufälligen Effekte sowie für die Residuen an. Leiten Sie daraus die Kovarianzmatrix der Zielvariable ab.
- (g) Sie schätzen nun das eben betrachtete Modell in R und erhalten folgenden Output:

```
print(fm1Machine <- lme(score ~ -1 + Machine, random = ~ 1 | Worker, data = Machines))

Linear mixed-effects model fit by REML
Data: Machines
Log-restricted-likelihood: -143.4391
```

```

Fixed: score ~ -1 + Machine
MachineA MachineB MachineC
52.35556 60.32222 66.27222

Random effects:
Formula: ~1 | Worker
      (Intercept) Residual
StdDev:    5.146552 3.161647

Number of Observations: 54
Number of Groups: 6

```

Die Prognosen für die zufälligen Effekte sind:

```

ranef(fm1Machine)

(Intercept)
6 -8.70711058
2 -1.59425968
4 -0.06931564
1  1.21035769
3  6.21174760
5  2.94858062

```

Interpretieren Sie die Parameter aus dem Modell sowie die zufälligen Effekte für Arbeiter 6 und für Arbeiter 3. Welche Leistung erwartet man laut Modell für Arbeiter 6 an Maschine B?

- (h) Stellen Sie ein Modell mit einem festen Effekt für den Maschinentyp und zufälligen Effekten für die Interaktionen aus Arbeitern und Maschinen auf (ohne zufällige Effekte für die Arbeiter). Nehmen Sie an, dass die zufälligen Interaktionseffekte innerhalb jedes Arbeiters korreliert sein können. Für die Residuen wird Unabhängigkeit angenommen. Geben Sie die Modellgleichung für Arbeiter i an, also das Modell für den Vektor \mathbf{y}_i .
- (i) Wenn man das Modell aus (h) in R schätzt, erhält man folgenden Output:

```

print(fm2Machine <- lme(score ~ - 1 + Machine, random = ~ Machine - 1 | Worker, data = Machines))

Linear mixed-effects model fit by REML
Data: Machines
Log-restricted-likelihood: -104.1556
Fixed: score ~ -1 + Machine
MachineA MachineB MachineC
52.35556 60.32222 66.27222

Random effects:
Formula: ~Machine - 1 | Worker
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev   Corr
MachineA 4.0792807 MachnA MachnB
MachineB 8.6252908 0.803

```

```
MachineC 4.3894795 0.623 0.771
Residual 0.9615766

Number of Observations: 54
Number of Groups: 6
```

Für die zufälligen Effekte erhält man:

```
ranef(fm2Machine)

      MachineA      MachineB      MachineC
6 -5.5916013 -16.5838056 -5.0300737
2  0.1838844  -0.8033247 -4.2822823
4 -1.0238706  2.3284558 -1.4146283
1  0.3119890  2.5532237  0.9302963
3  6.9692218  7.7793500  4.4735105
5 -0.8496233  4.7261008  5.3231775
```

Interpretieren Sie die Korrelation zwischen den zufälligen Effekten. Welche Leistung erwartet man nach Modell (h) für Arbeiter 6 an Maschine B?

Aufgabe 2

In dieser Aufgabe soll der Datensatz `Oxide` aus dem R-Paket `nlme` ausgewertet werden. Die Zielgröße ist die Dicke der Oxidschicht (**Thickness**) auf Siliziumscheiben (**Wafer**). Dazu wurden zufällig acht Parzellen (**Lot**) ausgewählt. Aus jeder Parzelle wurden drei Siliziumscheiben gezogen, bei denen je an drei unterschiedlichen Stellen (**Site**) die Dicke der Oxidschicht gemessen wurde. Das Ziel der Studie ist es, die Varianzkomponenten zu schätzen, die zu den verschiedenen Ebenen (Parzellen, Scheiben) gehören, um die Ursachen für die Variabilität in der Ablagerung von Oxid zu evaluieren ¹. Im folgenden sehen Sie die Beobachtungen für die erste Parzelle.

| Lot | Wafer | Site | Thickness |
|-----|-------|------|-----------|
| 1 | 1 | 1 | 2006 |
| 1 | 1 | 2 | 1999 |
| 1 | 1 | 3 | 2007 |
| 1 | 2 | 1 | 1980 |
| 1 | 2 | 2 | 1988 |
| 1 | 2 | 3 | 1982 |
| 1 | 3 | 1 | 2000 |
| 1 | 3 | 2 | 1998 |
| 1 | 3 | 3 | 2007 |

- Handelt es sich um ein gekreuztes oder ein genestetes Design? Begründen Sie Ihre Antwort.
- Sie möchten nun die Dicke der Oxidschicht in Abhängigkeit der Parzellen und Scheiben modellieren. Warum sollte man zufällige Effekte verwenden?
- Stellen Sie ein Modell für die Dicke der Oxidschicht y_{ijk} gemessen an der k -ten Stelle an der j -ten Scheibe in der i -ten Parzelle auf. Vergessen Sie nicht die Modellannahmen anzugeben.

¹Pinheiro, J. C.; Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer, New York

(d) Wenn man das Modell in R schätzt, erhält man u.a. folgenden Output:

```
print(fm1Oxide <- lme(Thickness ~ 1, random = ~ 1 | Lot / Wafer, data = Oxide))

Linear mixed-effects model fit by REML
Data: Oxide
Log-restricted-likelihood: -227.011
Fixed: Thickness ~ 1
(Intercept)
2000.153

Random effects:
Formula: ~1 | Lot
(Intercept)
StdDev: 11.39768

Formula: ~1 | Wafer %in% Lot
(Intercept) Residual
StdDev: 5.988802 3.545341

Number of Observations: 72
Number of Groups:
Lot Wafer %in% Lot
8 24
```

Vergleichen Sie die geschätzten Varianzkomponenten für die Ebene der Parzellen und der Scheiben. Auf welcher Ebene gibt es mehr Variabilität?

(e) Die Prognosen für die zufälligen Effekte ist :

```
[1] "Lot"
 1 2 3 4 5 6 7 8
-3.46 -11.22 0.87 -4.47 13.46 19.41 -8.20 -6.39
[1] "Wafer %in% Lot"
 1/1 1/2 1/3 2/1 2/2 2/3 3/1 3/2 3/3 4/1
6.55 -11.96 4.46 0.66 -0.83 -2.92 1.47 -0.62 -0.62 -0.01
 4/2 4/3 5/1 5/2 5/3 6/1 6/2 6/3 7/1 7/2
3.27 -4.49 -4.43 3.03 5.12 11.73 2.18 -8.56 -1.75 -0.56
 7/3 8/1 8/2 8/3
0.04 -0.09 1.40 -3.07
```

Welche Dicke der Oxidschicht erwartet man laut Modell für Parzelle 8? Welche Dicke der Oxidschicht erwartet man laut Modell für Scheibe 2 aus Parzelle 8?

(f) Berechnen Sie jeweils für marginale und konditionale Sichtweise die Kovarianz zweier Messungen an der gleichen Scheibe, sowie die Kovarianz zweier Messungen in der gleichen Parzelle für unterschiedliche Scheiben.