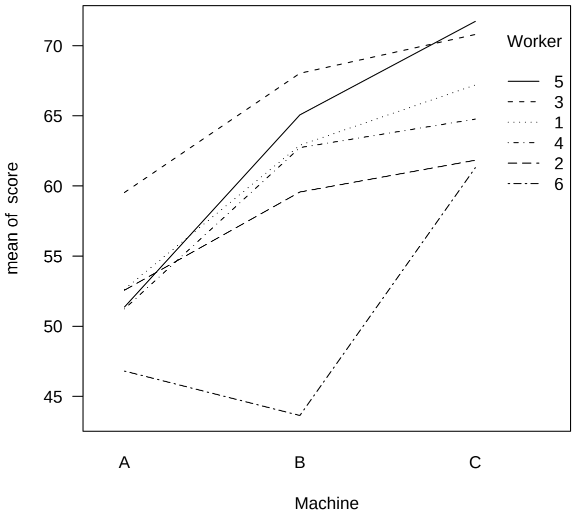
Stat 8311 Replicated blocked design

Six randomly selected workers use each of three different machine types in random order. Each worker uses each machine three times, giving replication. This can be called a *replicated block design*. Productivity scores are obtained.

Balanced data

```
> data(Machines, package = "nlme")
> attach(Machines)
> interaction.plot(Machine, Worker, score, las = 1)
```



First, we fit ignoring machines in workers. To save space, the input is shown but no output is given.

```
> options(contrasts = c("contr.SAS", "contr.poly"), digits = 8)
> library(lme4)
> m1 <- lmer(score ~ Machine + (1 | Worker), data = Machines)
```

Next, fit with machines nested within workers. This is a *two-level* model, written for machine j in subject i as

$$Y_{ij} = X_{ij}\beta + Z_{i,j}b_i + Z_{ij}b_{ij} + \varepsilon_{ij}$$

where X_{ij} is the design matrix for machine j in subject i and for the machine problem will specify the same mean for each observation; $Z_{i,j}$ and Z_{ij} are both columns of ones in this problem because both workers and machines in workers have one random effect.

```
> (m2 <- update(m1, ~. + (1 | with(Machines, Machine:Worker))))

Linear mixed-effects model fit by REML
Formula: score ~ Machine + (1 | Worker) + (1 | with(Machines, Machine:Worker))
   Data: Machines
    AIC    BIC  logLik MLdeviance REMLdeviance
 225.69 235.63 -107.84     225.46       215.69
```

```
Random effects:
 Groups                          Name        Variance Std.Dev.
 with(Machines, Machine:Worker) (Intercept) 13.88471 3.726219
 Worker                          (Intercept) 22.85039 4.780208
 Residual                                     0.92515 0.961847
number of obs: 54, groups: with(Machines, Machine:Worker), 18; Worker, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  66.2722     2.4847 26.6717
MachineA    -13.9167     2.1751 -6.3982
MachineB     -5.9500     2.1751 -2.7355

Correlation of Fixed Effects:
         (Intr) MachnA
MachineA -0.438
MachineB -0.438  0.500

> anova(m2, m1)

Data: Machines
Models:
m1: score ~ Machine + (1 | Worker)
m2: score ~ Machine + (1 | Worker) + (1 | with(Machines, Machine:Worker))
   Df     AIC     BIC   logLik   Chisq Chi Df Pr(>Chisq)
m1  4  301.753  309.709 -146.877
m2  5  235.461  245.406 -112.730 68.2928      1 < 2.22e-16
```

```
proc mixed data=machine method=REML;
      class Worker Machine;
      model Score = Machine/ alpha=.05;
      random Worker Worker(Machine);
   run;
                Covariance Parameter Estimates
Cov Parm            Estimate     Alpha       Lower       Upper
Worker               22.8584      0.05      7.6910      251.49
Worker(Machine)      13.9095      0.05      6.7031     44.2384
Residual              0.9246      0.05      0.6115      1.5601

           Fit Statistics
-2 Res Log Likelihood          215.7
AIC (smaller is better)        221.7
AICC (smaller is better)       222.2
BIC (smaller is better)        221.1

                      Solution for Fixed Effects
                              Standard
Effect     Machine  Estimate     Error     DF  t Value  Pr > |t|   Alpha
Intercept            66.2722    2.4858      5    26.66    <.0001    0.05
```

```
Machine    A        -13.9167     2.1770     10     -6.39     <.0001     0.05
Machine    B         -5.9500     2.1770     10     -2.73     0.0211     0.05
Machine    C               0          .      .         .          .        .

        Type 3 Tests of Fixed Effects
              Num      Den
Effect         DF       DF     F Value     Pr > F

Machine         2       10       20.58     0.000
```

More general models for the replicated block example

Suppose that subject i is allowed to have a separate random effect for each machine. In the balanced problem, each subject contributes nine observations, and

$$\mathbf{y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i b_i + \varepsilon_i, i = 1, \ldots, 6$$

If we use the R parameterization for fixed effects, then for this particular problem we can write

$$\mathbf{y}_i = \left[J_3 \otimes \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right] \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{pmatrix} + (J_3 \otimes I_3) \begin{pmatrix} b_{i1} \\ b_{i2} \\ b_{i3} \end{pmatrix} + \varepsilon_i, i = 1, \ldots, 6$$

with

$$b_i = \begin{pmatrix} b_{i1} \\ b_{i2} \\ b_{i3} \end{pmatrix} \sim \mathrm{N}(0, \Psi);\ \varepsilon_i \sim \mathrm{N}(0, \sigma^2 I)$$

where b_i is now 3×1 and Ψ is a 3×3 unknown positive definite matrix. There are many possible forms for Ψ, including the following four:

1. $\Psi = \sigma_b^2 I$. Each subject has a separate random effect for each machine, but all are from the same distribution. All random effects are equally variable. Only one parameter is estimated for the random effects. This is the model `m1` fit previously.

2. $\Psi = \sigma_1^2 I + \sigma_2^2 JJ'$, called compound symmetry, so the variances are $\sigma_1^2 + \sigma_2^2$, and the correlations are $\sigma_1^2/(\sigma_1^2 + \sigma_2^2)$. This case has two random effect parameters and *is identical to the nested model* `m2` *fit previously.*

3. Ψ is a general positive definite matrix, so it has three diagonal terms and three correlations, for a total of six random effect parameters. This is model `m3` below.

4. $\Psi = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$. In this case, each machine has its own variance and each subject has a different random effect for each machine. There are three random effect parameters. This is model `m4` below.

```
> (m3 <- update(m1, ~Machine + (0 + Machine | Worker)))

Linear mixed-effects model fit by REML
Formula: score ~ Machine + (0 + Machine | Worker)
   Data: Machines
    AIC    BIC  logLik MLdeviance REMLdeviance
```

```
 226.31 244.21 -104.16     216.61       208.31
Random effects:
 Groups   Name     Variance  Std.Dev. Corr
 Worker   MachineA 16.642335 4.07950
          MachineB 74.373057 8.62398  0.803
          MachineC 19.264406 4.38912  0.623 0.771
 Residual           0.924638 0.96158
number of obs: 54, groups: Worker, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  66.2722     1.8061  36.693
MachineA    -13.9167     1.5400  -9.037
MachineB     -5.9500     2.4462  -2.432

Correlation of Fixed Effects:
         (Intr) MachnA
MachineA -0.505
MachineB  0.362  0.331

> mA <- ifelse(Machines$Machine == "A", 1, 0)
> mB <- ifelse(Machines$Machine == "B", 1, 0)
> mC <- ifelse(Machines$Machine == "C", 1, 0)
> (m4 <- update(m1, ~Machine + (0 + mA | Worker) + (0 + mB |
+     Worker) + (0 + mC | Worker)))

Linear mixed-effects model fit by REML
Formula: score ~ Machine + (0 + mA | Worker) + (0 + mB | Worker) + (0 +      mC | Worker)
   Data: Machines
    AIC    BIC  logLik MLdeviance REMLdeviance
 229.65 241.58 -108.83     227.82       217.65
Random effects:
 Groups   Name Variance  Std.Dev.
 Worker   mA   16.639651 4.079173
 Worker   mB   74.392086 8.625085
 Worker   mC   19.266529 4.389365
 Residual       0.924649 0.961587
number of obs: 54, groups: Worker, 6; Worker, 6; Worker, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)  66.2722     1.8062  36.691
MachineA    -13.9167     2.4672  -5.641
MachineB     -5.9500     3.9639  -1.501

Correlation of Fixed Effects:
         (Intr) MachnA
MachineA -0.732
MachineB -0.456  0.334
```

Models:

Model	Covariance	SAS specification
m1	$b_{i1} = b_{i2} = b_{i3}$, $\text{var}(b_{ij}) = \sigma_b^2$	`random Worker;`
m2	$\text{var}(b_i) = \sigma_b^2 JJ'$	`random Worker Machine(Worker);`
m3	$\text{var}(b_i) =$ arbitrary pos. def. matrix	`random Machine/subject=worker type=UN;`
m4	$\text{var}(b_i) = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$	`random Machine/subject=worker type=UN(1);`

SAS specifications

```
proc mixed data=machine method=REML; /* ignore replications, m1*/
      class Worker Machine;
      model Score = Machine/ alpha=.05;
      random Worker;
   run;
proc mixed data=machine method=REML; /* machines in workers, m2*/
      class Worker Machine;
      model Score = Machine/ alpha=.05;
      random Worker Machine(Worker);
   run;
proc mixed data=machine method=REML; /* general covariance matrix, m3*/
      class Worker Machine;
      model Score = Machine/ alpha=.05;
      random Machine/subject=worker type=UN;
   run;
proc mixed data=machine method=REML; /* diagonal covariance amtrix, m4*/
      class Worker Machine;
      model Score = Machine/ alpha=.05;
      random Machine/subject=worker type=UN(1);
   run;
```