1 "Machines" Data Set - Example

? describes an experiment whereby the productivity of six randomly chosen workers are assessed three times on each of three machines, yielding the 54 observations tabulated below.

Observation	Worker	Machine	score	Observation	Worker	Machine	score
1	1	A	52.00	28	4	B	63.20
2	1	A	52.80	29	4	B	62.80
3	1	A	53.10	30	4	B	62.20
4	2	A	51.80	31	5	B	64.80
5	2	A	52.80	32	5	B	65.00
6	2	A	53.10	33	5	B	65.40
7	3	A	60.00	34	6	B	43.70
8	3	A	60.20	35	6	B	44.20
9	3	A	58.40	36	6	B	43.00
10	4	A	51.10	37	1	C	67.50
11	4	A	52.30	38	1	C	67.20
12	4	A	50.30	39	1	C	66.90
13	5	A	50.90	40	2	C	61.50
14	5	A	51.80	41	2	C	61.70
15	5	A	51.40	42	2	C	62.30
16	6	A	46.40	43	3	C	70.80
17	6	A	44.80	44	3	C	70.60
18	6	A	49.20	45	3	C	71.00
19	1	B	62.10	46	4	C	64.10
20	1	B	62.60	47	4	C	66.20
21	1	B	64.00	48	4	C	64.00
22	2	B	59.70	49	5	C	72.10
23	2	B	60.00	50	5	C	72.00
24	2	B	59.00	51	5	C	71.10
25	3	B	68.60	52	6	C	62.00
26	3	B	65.80	53	6	C	61.40
27	3	B	69.70	54	6	C	60.50

Table 1: Machines Data , Pinheiro Bates

(Overall mean score = 59.65, mean on machine A = 52.35 , mean on machine B = 60.32, mean on machine C = 66.27)

The 'worker' factor is modelled with random effects(u_i), whereas the 'machine' factor is modelled with fixed effects (β_j). Due to the repeated nature of the data, interaction effects between these factors are assumed to be extant, and shall be examined accordingly. The interaction effect in this case (τ_{ij}) describes whether the effect of changing from one machine to another is different for each worker. The productivity score y_{ijk} is the kth observation taken on worker i on machine j, and is formulated as follows;

$$y_{ijk} = \beta_j + u_i + \tau_{ij} + \epsilon_{ijk} \tag{1}$$
$$u_i \sim N(0, \sigma_u^2), \epsilon_{ijk} \sim N(0, \sigma^2), \tau_i \sim N(0, \sigma_\tau^2)$$

The 'nlme' package is incorporated into the R programming to perform linear mixed model calculations. For the 'Machines' data, **?** use the following code, with the hierarchical structure specified in the last argument.

```
lme(score~Machine, data=Machines, random=~1|Worker/Machine)
```

The output of the R computation is given below.

```
Linear mixed-effects model fit by REML
  Data: Machines
  Log-restricted-likelihood: -107.8438
  Fixed: score ~ Machine
(Intercept)    MachineB    MachineC
  52.355556    7.966667   13.916667

Random effects:
 Formula: ~1 | Worker
        (Intercept)
StdDev:     4.78105

 Formula: ~1 | Machine %in% Worker
        (Intercept)  Residual
StdDev:    3.729532 0.9615771

Number of Observations: 54 Number of Groups:
             Worker Machine %in% Worker
                  6                  18
```

The crucial pieces of information given in the programme output are the estimates of the intercepts for each of the three machines. Machine A, which is treated as a control case, is estimated to have an intercept of 52.35. The intercept estimates for machines B and C are found to be 60.32 and 66.27 (by adding the values 7.96 and 13.91 to 52.35 respectively). Estimate for the variance components are also given; $\sigma_u^2 = (4.78)^2$, $\sigma_\tau^2 = (3.73)^2$ and $\sigma_\epsilon^2 = (0.96)^2$.