

R version 2.4.1 (2006-12-18)
 Copyright (C) 2006 The R Foundation for Statistical Computing
 ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
 You are welcome to redistribute it under certain conditions.
 Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
 Type 'contributors()' for more information and
 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
 'help.start()' for an HTML browser interface to help.
 Type 'q()' to quit R.

[Previously saved workspace restored]

```

Loading lattice() and nlme()
Loading required package: graphics
Loading required package: grDevices
Loading required package: stats
> invisible(options(echo = TRUE))
> #####
> # Flay et al. (1995) The television, school, and family smoking prevention
> # and cessation project. Preventive Medicine, 24, 29-40.
> #####
> # data available from http://www.biostat.harvard.edu/~fitzmaur/ala
> # first 43 lines of the file explain the dataset, skipped
> tvsfp.df <- read.table(file = "./data/tvsfp.txt", skip = 44, header = FALSE,
+   col.names = c("school", "class", "sctx", "tvtx", "prethks", "postthks"))
> tvsfp.df$school <- factor(tvsfp.df$school)
> tvsfp.df$class <- factor(tvsfp.df$class)
> tvsfp.df$sctx <- factor(tvsfp.df$sctx, levels=c(0,1),labels=c("sr.n", "sr.y"))
> tvsfp.df$tvtx <- factor(tvsfp.df$tvtx, levels=c(0,1),labels=c("tv.n", "tv.y"))
> dim(tvsfp.df) # size of the data
[1] 1600    6
> tvsfp.df[1:7, ]
  school class sctx tvtx prethks postthks
1    403 403101 sr.y tv.n      2        3
2    403 403101 sr.y tv.n      4        4
3    403 403101 sr.y tv.n      4        3
4    403 403101 sr.y tv.n      3        4
5    403 403101 sr.y tv.n      3        4
6    403 403101 sr.y tv.n      4        3
7    403 403101 sr.y tv.n      2        2
> table(tvsfp.df$class) # classes nested within schools

193101 194101 194102 194103 194104 194105 194106 196101 196102 197101 197102
   26    11    10    15    12    12    10    21    10    17    19
197103 197104 198101 198102 198103 199101 199102 199103 199104 199105 199106
   2     4    21    16    15    13     2    14    13     1    12
401101 401102 402101 402102 403101 403102 404101 404102 404103 405101 405102
  18    21    17    16    20     3    11     9     5    15    16
405103 407101 407102 407103 408101 408102 408103 408104 409101 409102 409103
  21    21    27    17     3     2    15     7    24    22    15
409104 410101 410102 410103 410104 411101 411102 412101 412102 412103 412104
  19     7    12     8     6    11     7     7     8     3     5
412105 412106 414101 414102 414104 414105 414106 415101 415102 415103 415104
   8     3     8     9     8     2    11     9    21    13    12

```

```

415105 505101 505102 505103 505104 505105 505106 505107 506102 506103 506105
   12   14    9    6    11    10    12    11    6    6    14
506107 506110 506111 506112 506113 506114 506115 506116 507101 507102 507103
   10   14    2    3    2    4    3    6    10   14   11
507104 507105 507106 507107 508101 508102 508103 508104 509101 509102 509103
   6    22    4    7    22   18   23   19   14    5   14
509104 509105 509106 509107 509108 510101 510102 510103 510104 510105 510106
  28   15   15   11   12   11   11   15   17   24   20
510107 513101 513102 513103 513104 514101 514102 514103 514104 514105 514106
   15    6    8   14    5   11    9   17   14   11   17
514107 515101 515102 515103 515104 515105 515106 515107 515108 515109 515110
   15    6    7   14   10   14   18    4    9    7   12
515111 515112 515113
   15   11   10
> #
> # group data according to the lme() model, single level of nesting:
> # Classes nested within schools. Treatment conditions as outer.
> #
> tvsfp.df <- groupedData(postthks ~ prethks | school/class, data = tvsfp.df,
+   outer= list(~ schtx, ~ tvtx),
+   labels = list(x = "pre-intervention THINKS scores",
+     y = "post-intervention THINKS scores"),
+   units = list(x = "(0 - 6)", y = "(0 - 6)"))
> attach(tvsfp.df)      # variables of data.frame directly available
> png("tvsfp_by_class.png", height = 700, width = 800)
> plot(tvsfp.df)       # by individual classes, unit of analysis in lme()
> dev.off(dev.cur())   # _by_class.png saved when graphics device is closed
null device
      1
> png("tvsfp_by_sch.png", height = 700, width = 800)
> # pick the colors you like, see http://html-color-codes.com/
> # color names found at http://www.computerhope.com/htmcolor.htm
> trellis.par.get()$superpose.symbol
$alpha
[1] 1 1 1 1 1 1 1

$cex
[1] 0.8 0.8 0.8 0.8 0.8 0.8 0.8

$col
[1] "#0080ff"   "#ff00ff"   "darkgreen" "#ff0000"   "orange"    "#00ff00"
[7] "brown"

$font
[1] 1 1 1 1 1 1 1

$pch
[1] 1 1 1 1 1 1 1

$fill
[1] "transparent" "transparent" "transparent" "transparent" "transparent"
[6] "transparent" "transparent"

> tc <- trellis.par.get()$superpose.symbol$col
> tc[1:2] <- c("#0000A0", "#800080") # dark blue, dark purple
> trellis.par.set(superpose.symbol = list(col = tc))
> #
> # plot by schools, grouped by the TV intervention condition. Blue is
> #
> xyplot(postthks ~ prethks | school, data = tvsfp.df,
+   groups = tvtx,
+   panel = function(x, y, ...)
+     { panel.superpose(x, y, ...); panel.loess(x, y, ...) },

```

```
+ strip = strip.custom(style = 1, strip.names = TRUE, strip.levels = TRUE,
+                       par.strip.text = list(cex = 0.8) ),
+ key = list(columns = 2, text = list(paste(c("TV programs: ", "")),
+ unique(tvsvfp.df$tvtx)), points = Rows(sps, 1:3) ) )
+ )
```

There were 50 or more warnings (use warnings() to see the first 50)

```
> dev.off(dev.cur())
```

```
null device
```

```
1
```

```
> tapply(postthks, list(tvtx, schtx), mean)
```

```
sr.n sr.y
```

```
tv.n 2.361045 2.968421
```

```
tv.y 2.538462 2.822454
```

```
> tapply(postthks, list(tvtx, schtx), sd)
```

```
sr.n sr.y
```

```
tv.n 1.295708 1.404498
```

```
tv.y 1.437354 1.312373
```

```
> tapply(postthks, tvtx, mean)
```

```
tv.n tv.y
```

```
2.649189 2.674593
```

```
> tapply(postthks, schtx, mean)
```

```
sr.n sr.y
```

```
2.449223 2.895151
```

```
> tvlme1 <- lme(postthks ~ prethks + schtx*tvtx, random = ~ 1 | school/class, data = tvsvfp.df)
```

```
> summary(tvlme1)
```

Linear mixed-effects model fit by REML

Data: tvsvfp.df

	AIC	BIC	logLik
	5389.335	5432.332	-2686.668

Random effects:

Formula: ~1 | school

(Intercept)

StdDev: 0.1965707

Formula: ~1 | class %in% school

(Intercept) Residual

StdDev: 0.2542862 1.265818

Fixed effects: postthks ~ prethks + schtx * tvtx

	Value	Std.Error	DF	t-value	p-value
(Intercept)	1.7019852	0.12543005	1464	13.569198	0.0000
prethks	0.3053628	0.02589132	1464	11.794021	0.0000
schtxsr.y	0.6413260	0.16094730	24	3.984696	0.0005
tvtxtv.y	0.1820802	0.15724056	24	1.157972	0.2583
schtxsr.y:tvtxtv.y	-0.3309400	0.22458559	24	-1.473559	0.1536

Correlation:

(Intr) prthks schtx. tvtxt.

prethks	-0.442			
schtxsr.y	-0.634	0.015		
tvtxtv.y	-0.645	0.008	0.501	
schtxsr.y:tvtxtv.y	0.448	0.005	-0.716	-0.700

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-2.49874554	-0.69757192	-0.01721255	0.68240730	3.14602046

Number of Observations: 1600

Number of Groups:

school class %in% school

28 135

```
> intervals(tvlme1)
```

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	1.4559434	1.7019852	1.9480270
prethks	0.2545748	0.3053628	0.3561509
schtxsr.y	0.3091471	0.6413260	0.9735049
tvtxtv.y	-0.1424483	0.1820802	0.5066088
schtxsr.y:tvtxtv.y	-0.7944619	-0.3309400	0.1325819

attr(,"label")
[1] "Fixed effects:"

Random Effects:

Level: school

	lower	est.	upper
sd((Intercept))	0.1033422	0.1965707	0.3739038

Level: class

	lower	est.	upper
sd((Intercept))	0.1650072	0.2542862	0.3918706

Within-group standard error:

	lower	est.	upper
	1.220870	1.265818	1.312420

> anova(tvlm1)

	numDF	denDF	F-value	p-value
(Intercept)	1	1464	2240.0362	<.0001
prethks	1	1464	136.7953	<.0001
schtx	1	24	17.6286	0.0003
tvtx	1	24	0.0313	0.8610
schtx:tvtx	1	24	2.1714	0.1536

> ###
> # wrong p-values in simple regression without cluster, Fitzmaurice: table 17.4
> ###
> tvlm2 <- lm(postthks ~ prethks + schtx*tvtx, data = tvsfp.df)
> summary(tvlm2)

Call:

lm(formula = postthks ~ prethks + schtx * tvtx, data = tvsfp.df)

Residuals:

Min	1Q	Median	3Q	Max
-3.60252	-0.95216	0.01357	0.84558	4.16450

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.66126	0.08436	19.693	< 2e-16 ***
prethks	0.32518	0.02585	12.578	< 2e-16 ***
schtxsr.y	0.64055	0.09210	6.955	5.14e-12 ***
tvtxtv.y	0.19871	0.08996	2.209	0.0273 *
schtxsr.y:tvtxtv.y	-0.32162	0.13025	-2.469	0.0136 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.301 on 1595 degrees of freedom
Multiple R-Squared: 0.117, Adjusted R-squared: 0.1148
F-statistic: 52.85 on 4 and 1595 DF, p-value: < 2.2e-16

> anova(tvlm2) # wrong denominator SS, causing inflated p-values.
Analysis of Variance Table

Response: postthks

Df	Sum Sq	Mean Sq	F value	Pr(>F)
----	--------	---------	---------	--------

```
prethks 1 254.90 254.90 150.5691 < 2.2e-16 ***
schtx 1 91.82 91.82 54.2381 2.835e-13 ***
tvtx 1 0.82 0.82 0.4853 0.48613
schtx:tvtx 1 10.32 10.32 6.0971 0.01364 *
Residuals 1595 2700.21 1.69
```

```
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # Note: F tests differ from t-tests b/c of inbalanced design
```

```
> ###
> # wrong p-values even with Error(class) to find a better denominator SS
> ###
> tvaov1 <- aov(postthks ~ schtx*tvtx, data=tvsfp.df)
> summary(tvaov1)
```

```
      Df Sum Sq Mean Sq F value    Pr(>F)
schtx  1  79.37   79.37 42.6795 8.647e-11 ***
tvtx    1   0.22    0.22  0.1158  0.73370
schtx:tvtx 1 10.44   10.44  5.6112  0.01796 *
Residuals 1596 2968.05    1.86
```

```
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> tvaov2 <- aov(postthks ~ schtx*tvtx + Error(class), data=tvsfp.df)
> summary(tvaov2)
```

Error: class

```
      Df Sum Sq Mean Sq F value    Pr(>F)
schtx  1  79.37   79.37 22.6261 5.117e-06 ***
tvtx    1   0.22    0.22  0.0614  0.80472
schtx:tvtx 1 10.44   10.44  2.9747  0.08693 .
Residuals 131 459.54    3.51
```

```
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Error: Within

```
      Df Sum Sq Mean Sq F value Pr(>F)
Residuals 1465 2508.52    1.71
```

```
> ##
> # Show the model matrix
> ##
> tvsfpl <- tvsfp.df[tvsfp.df$class == "197102" | tvsfp.df$class == "198101", ]
> model.matrix( postthks ~ prethks + tvtx*schtx, data = tvsfpl)
```

```
(Intercept) prethks tvtxtv.y schtxsr.y tvtxtv.y:schtxsr.y
193          1          1          0          0          0
194          1          0          0          0          0
195          1          2          0          0          0
196          1          4          0          0          0
197          1          2          0          0          0
198          1          3          0          0          0
199          1          3          0          0          0
200          1          3          0          0          0
201          1          3          0          0          0
202          1          1          0          0          0
203          1          3          0          0          0
204          1          2          0          0          0
205          1          2          0          0          0
206          1          3          0          0          0
207          1          2          0          0          0
208          1          2          0          0          0
209          1          2          0          0          0
210          1          1          0          0          0
211          1          2          0          0          0
```

218	1	1	1	1	1
219	1	2	1	1	1
220	1	2	1	1	1
221	1	4	1	1	1
222	1	1	1	1	1
223	1	3	1	1	1
224	1	5	1	1	1
225	1	3	1	1	1
226	1	2	1	1	1
227	1	2	1	1	1
228	1	3	1	1	1
229	1	3	1	1	1
230	1	3	1	1	1
231	1	0	1	1	1
232	1	2	1	1	1
233	1	3	1	1	1
234	1	3	1	1	1
235	1	3	1	1	1
236	1	1	1	1	1
237	1	2	1	1	1
238	1	5	1	1	1

```
attr("assign")
```

```
[1] 0 1 2 3 4
```

```
attr("contrasts")
```

```
attr("contrasts")$tvtx
```

```
[1] "contr.treatment"
```

```
attr("contrasts")$schtx
```

```
[1] "contr.treatment"
```

```
> ranef(tvlme1)
```

```
Level: school
```

```
(Intercept)
```

```
193 0.032872451
```

```
194 0.055932853
```

```
196 0.038411248
```

```
[... snipped ...]
```

```
513 -0.146524361
```

```
514 0.099681608
```

```
515 -0.067687760
```

```
Level: class %in% school
```

```
(Intercept)
```

```
193/193101 0.0550098232
```

```
194/194101 -0.0427131919
```

```
194/194102 0.0138410076
```

```
194/194103 -0.2570226849
```

```
194/194104 0.0202312303
```

```
194/194105 0.0866510391
```

```
[... snipped ...]
```

```
515/515111 0.3447466725
```

```
515/515112 0.0838291349
```

```
515/515113 -0.2892764561
```

```
> # garbage collection
```

```
> rm(tc, tvaov1, tvsfpl, tvlme1)
```

```
> proc.time()
```

```
[1] 15.848 0.208 17.641 0.000 0.000
```

```
>
```