About the Pediatric Pain data

This data set consists of up to four observations on 64 elementary school children. The response is the length of time the child can tolerate keeping their arm in very cold water.

Two trials were made during a first visit, and two more trials were made during a second visit two weeks later. During each visit, the first trial used the dominant arm, and the second trial used the non-dominant arm.

There is missing data on six students because of arms in casts and similar reasons; some are missing data on one arm for both visits, some for only one visit; some are missing data from the second visit altogether.

Subjects were asked what they were thinking about during the first two trials, and their response was classified into two coping style groups: those thinking about the experiment (attenders) and those thinking about other things (distracters).

Before the fourth trial, a randomized treatment was assigned, consisting of a ten-minute counseling session, where advice was given either to attend or to district, or no advice was given.

Scientific interest lies both in the main effects of treatment and coping strategy and in possible interactions between treatment and coping strategy. In particular, subjects for whom the advice treatment matched the coping style were expected to do better (have longer times) than those with mismatched advice treatment or no advice treatment.

- How could you test for overall differences in the mean response between coping style (ignoring treatment) using methods you already know about? You may have to take a subset of the data or use summary statistics instead of the raw data. How might you handle the missing data?

- What if you were instead interested in how the tolerance time changes over time (if it inccreases or decreases), and if that depends on coping style?

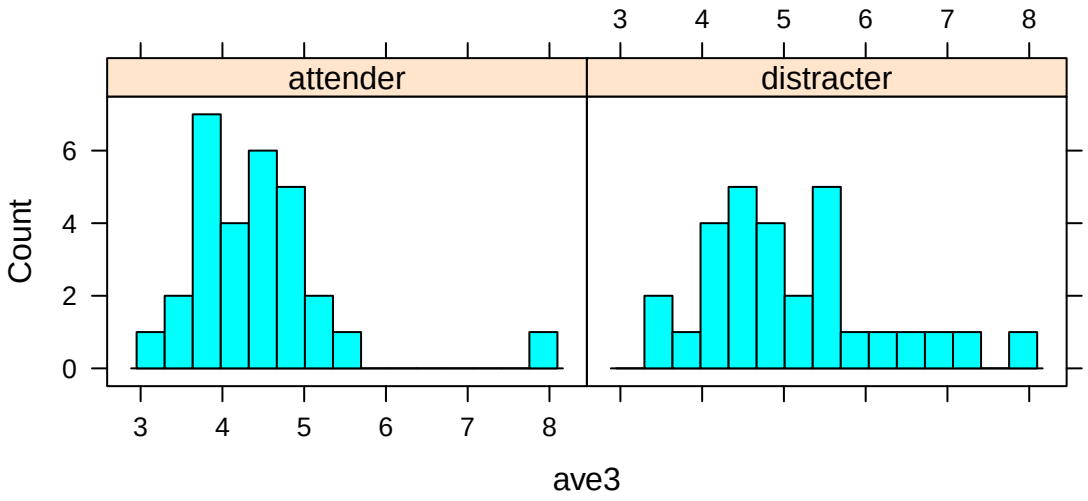
- How could you test for differences between treatments (ignoring coping style)? Again, you might have to manipulate the data before applying the test.

- How could you test for an interaction between treatment and coping style? Specifically consider those who were given advice to distract. Did the effect of this treatment depend on the coping style the student used previously?

Testing means, by coping style

Read in data and remove any with missing data

```
> pp <- read.delim("http://rem.ph.ucla.edu/rob/mld/data/tabdelimiteddata/pain.txt")
> pp <- subset(pp, !is.na(l2paintol))
> ppl <- cast(id ~ ., value = "l2paintol", length, data = pp)
> names(ppl)[2] <- "n"
> id.ok <- ppl$id[ppl$n == 4]
> pp <- subset(pp, id %in% id.ok)

> pp3 <- subset(pp, trial <= 3)
> ppw <- cast(id + cs + treatment ~ ., value = "l2paintol", data = pp3,
+     fun.aggregate = mean)
> names(ppw)[4] <- "ave3"

> plot(histogram(~ave3 | cs, data = ppw, nint = 15, type = "count"))
```




```
> cast(cs ~ ., data = ppw, value = "ave3", function(x) c(n = length(x),
+     mean = mean(x), sd = sd(x)))

          cs  n     mean       sd
1   attender 29 4.472964 0.879483
2 distracter 29 5.076170 1.093887
```

```
> t.test(ave3 ~ cs, data = ppw)

        Welch Two Sample t-test

data:  ave3 by cs
t = -2.3143, df = 53.531, p-value = 0.02452
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.12586495 -0.08054762
sample estimates:
  mean in group attender mean in group distracter
                4.472964                 5.076170
```

Testing slopes, by coping style

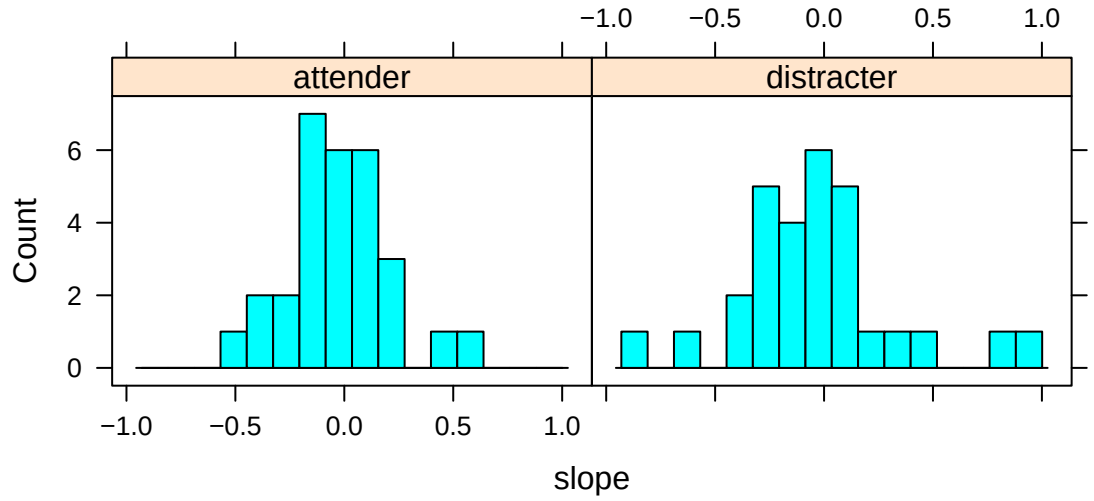
How do we get a linear model fit and the slope?

```
> coef(lm(l2paintol ~ trial, data = pp))

(Intercept)       trial
 4.85179622 -0.03150331
```

We split the data by id and do that to each part, merging the result with the original data to keep the coping style information.

```
> ppid <- split(pp, pp$id)
> regs <- sapply(ppid, function(x) coef(lm(l2paintol ~ trial, data = x)))
> regs <- as.data.frame(t(regs))
> names(regs) <- c("intercept", "slope")
> regs$id <- as.numeric(rownames(regs))
> regs <- merge(regs, ppw[, 1:3])
> head(regs)

  id intercept       slope         cs     treatment
1  1  5.157234 -0.37392629   attender        attend
2  2  4.941397 -0.20226956 distracter      distract
3  3  3.721983 -0.13184438   attender no directions
4  5  4.427293  0.14114682   attender        attend
5  6  7.028407 -0.29019996 distracter        attend
6  7  4.255284  0.04324618 distracter no directions
```

```
> plot(histogram(~slope | cs, data = regs, nint = 16, type = "count"))
```




```
> cast(cs ~ ., data = regs, function(x) c(n = length(x), mean = mean(x),
+     sd = sd(x)), value = "slope", fill = NA)

          cs  n        mean        sd
1   attender 29 -0.01737223 0.2399295
2 distracter 29 -0.04563439 0.3721908

> t.test(slope ~ cs, data = regs)

        Welch Two Sample t-test

data:  slope by cs
t = 0.3437, df = 47.845, p-value = 0.7326
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1370865  0.1936109
sample estimates:
  mean in group attender mean in group distracter
             -0.01737223              -0.04563439
```

A paired t-test

```
> ppw <- cast(id + cs + treatment ~ trial, value = "l2paintol",
+     data = pp)
> head(ppw)
```

```
  id         cs     treatment        1        2        3        4
1  1   attender        attend 4.361066 5.142005 3.836934 3.549669
2  2 distracter      distract 4.814038 4.598127 3.987321 4.343408
3  3   attender no directions 3.584963 3.321928 3.614710 3.047887
4  5   attender        attend 4.638653 4.549053 4.961160 4.971773
5  6 distracter        attend 6.773073 6.424754 6.099716 5.914086
6  7 distracter no directions 4.306700 4.244126 4.555816 4.346957
```

Two equivalent ways:

```
> t.test(ppw$`4`, ppw$`2`, paired = TRUE)
```

```
        Paired t-test

data:  ppw$`4` and ppw$`2`
t = 0.547, df = 57, p-value = 0.5865
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.1809929  0.3170310
sample estimates:
mean of the differences
             0.06801907
```

```
> ppw$diff <- ppw$`4` - ppw$`2`
> t.test(ppw$diff)
```

```
        One Sample t-test

data:  ppw$diff
t = 0.547, df = 57, p-value = 0.5865
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.1809929  0.3170310
sample estimates:
 mean of x
0.06801907
```

Difference of Differences

```
> ppa <- subset(ppw, cs == "attender" & treatment != "distract")
> t.test(diff ~ treatment, data = ppa, var.equal = TRUE)

        Two Sample t-test

data:  diff by treatment
t = 0.4516, df = 17, p-value = 0.6573
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5639146  0.8710289
sample estimates:
       mean in group attend mean in group no directions
                  0.18194159                  0.02838446

> ppd <- subset(ppw, cs == "distracter" & treatment != "attend")
> t.test(diff ~ treatment, data = ppd, var.equal = TRUE)

        Two Sample t-test

data:  diff by treatment
t = 2.6448, df = 17, p-value = 0.01702
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.2819996 2.5060545
sample estimates:
     mean in group distract mean in group no directions
                  0.7153376                  -0.6786894
```