

Medical Statistics: Exercises 1

Notes & Solutions

1) In the comparison of a new drug A with a standard drug B it is required that patients are assigned to drugs A and B in the proportions 3:1 respectively. Illustrate how this may be achieved for a group of 32 patients, and provide an appropriate randomization list. Comment on the rationale for selecting a greater proportion of patients for drug A.

(i) Need blocks of form AAAB (or of form AAAAAABB). There are 4 of form AAAB (and 28 of size 8). Using 1,2→AAAB; 3,4→AABA; 5,6→ABAA; 7,8→BAAA, 9,0→ignore, a sequence of random digits 7,1,4,2,0,1,8,1,2,4 gives

BAAA|AAAB|AABA|AAAB|AAAB|BAAA|AAAB|AAAB.

In R, to produce a random block of form AAAB do:

```
> sample(c(rep("A", 3), "B"))
[1] "A" "A" "B" "A"
```

and then repeat as often as necessary or build into a loop.

Alternatively, to get exact balance without blocks do:

```
> sample(c(rep("A", 24), rep("B", 8)))
[1] "B" "B" "A" "A" "A" "A" "A" "A" "B" "B" "A" "A" "A" "A" "A"
[15] "A" "A" "A" "A" "A" "A" "A" "A" "A" "B" "A" "B" "A" "A" "B"
[29] "A" "B" "A" "A"
```

There could be economic reasons for using more As than Bs, but more likely if B is the standard then there will be interest in efficacy and safety of the new treatment but this is likely to be known for the standard, as would be drop out rates, standard deviations etc. Having more patients on the new treatment protects against uncertainty in drop-out rates (or side effects) and consistency of response. Further, there will be more interest and enthusiasm amongst both patients and investigators if there is a greater chance of receiving the new



treatment and so easier to recruit centres and patients. This last reason is probably the most important in practice though not obviously 'statistical'.

- 2) The table below gives the age (≤ 55 / >55), gender (M/F), disease stage (I/II/III) of subjects entering a randomized controlled clinical trial at various intervals and who are to be allocated to treatment or placebo in approximately equal proportions immediately on entry.

order of entry	Age	Gender	Stage
1	≤ 55	F	III
2	≤ 55	M	III
3	≤ 55	M	I
4	≤ 55	F	I
5	>55	F	II
6	≤ 55	F	III
7	>55	F	I
8	>55	M	III
9	≤ 55	M	III
10	>55	F	III
11	≤ 55	F	III
12	≤ 55	M	I
13	>55	F	I

- i) Construct a randomization list for this group of subjects by a minimization method designed to achieve an overall balance between the factors.

order of entry	Age	Gender	Stage	First Run		Second Run	
				score for T	score for P	score for T	score for P
1	≤ 55	F	III	0★	0	0	0★
2	≤ 55	M	III	2	0★	0★	2
3	≤ 55	M	I	1★	2	2	1★
4	≤ 55	F	I	4	1★	1★	3
5	>55	F	II	1	1★	1	1★
6	≤ 55	F	III	4★	5	4★	5
7	>55	F	I	3★	4	3★	4
8	>55	M	III	4	3★	4	3★
9	≤ 55	M	III	6★	6	6	6★
10	>55	F	III	7	6★	6★	7
11	≤ 55	F	III	9	8★	10	8★
12	≤ 55	M	I	8	6★	6★	7
13	>55	F	I	6★	9	9	6★



The first subject has to be allocated randomly to T or P. The ★ indicates which of T or P is selected. Then for each subsequent subject it is easy to calculate the score for T and P as the total number of characteristics held in common between the new arrival and those subjects already allocated to that group. Two runs are presented above, one resulting from a choice of T for the first subject — this leads to a tied score for the 5th subject and P was [randomly] chosen, another tie for the 9th and T was [randomly] chosen. The second run with P selected first also leads to a tie on the 5th arrival and then the 9th.

ii) Cross-tabulate the treatment received with each [separate] factor.

Run 1:

	Age			Gender			Stage			
	≤55	>55	total	M	F	total	I	II	III	total
T	4	2	6	2	4	6	3	0	3	6
P	4	3	7	3	4	7	2	1	4	7
total	8	5	13	5	8	13	5	1	7	13

Run 2:

	Age			Gender			Stage			
	≤55	>55	total	M	F	total	I	II	III	total
T	4	2	6	2	4	6	3	0	3	6
P	4	3	7	3	4	7	2	1	4	7
total	8	5	13	5	8	13	5	1	7	13

Note that these are identical, as are essentially all possible runs (i.e. up to an interchange of T and P). Even with a different order of arrival of these patients the final allocations are not substantially different.



iii) Construct a list to allocate the subjects to treatment completely randomly without taking any account of any prognostic factor and compare the balance between treatment groups achieved on each of the factors.

Using Minitab and putting T and P into the first two cells of C1 and then using the Calc>Random Data>Sample from Columns... menu and then clicking the Sample with replacement option gives:

```
MTB > Sample 13 C1 c2;
SUBC> Replace.
MTB > print C2
```

Data Display

C2

T T T T P T T P T T P P T

	Age			Gender			Stage			
	≤55	>55	total	M	F	total	I	II	III	total
T	6	3	9	3	6	9	4	0	5	9
P	2	2	4	2	2	4	1	1	2	4
total	8	5	13	5	8	13	5	1	7	13

(Different randomisations will lead to different cross-tabulations.)

In **R** the function `sample(.)` with the `replace=TRUE` option gives the same facility:

```
> sample(c("T", "P"), 24, replace=TRUE)
[1] "T" "P" "T" "T" "T" "T" "P" "T" "P" "P" "P" "P" "P" "T" "P"
[15] "T" "T" "P" "P" "P" "T" "P" "T" "P" "P"
```



3) In a clinical trial of the use of a drug in twin pregnancies an obstetrician wishes to show a significant prolongation of pregnancy by use of the drug when compared to placebo. She assesses that the standard deviation of pregnancy length is 1.5 weeks, and considers a clinically significant increase in pregnancy length of 1 week to be appropriate.

- i) How many pregnancies should be observed to detect such a difference in a test with a 5% significance level and with 80% power?

Require a two-sided two sample t-test. Formula gives 35.3 per group and R, Minitab and programme POWER give 37 in each group (S-PLUS gives 36) **so 74 (or 72) pregnancies in total need to be observed.**

```
> power.t.test(sd=1.5,delta=1,power=0.8)
Two-sample t test power calculation
  n = 36.3058
  delta = 1
  sd = 1.5
  sig.level = 0.05
  power = 0.8
  alternative = two.sided
NOTE: n is number in *each* group
```

- ii) It is thought that between 40 and 60 pregnancies will be observed to term during the course of the study. What range of increases in length of pregnancy will the study have a reasonable chance (i.e. between 70% and 90%) of detecting?

Note that "40 to 60 in total" means 20 to 30 in each group.

Results produced by programme POWER below:

Results

Two Sample T test

Table of CRD calculations

Sample size group 1

: 20 : 25 : 30 :

```
-----
70 : 1.20670 : 1.07390 : 0.97708 :
75 : 1.27967 : 1.13884 : 1.03617 :
80 : 1.36103 : 1.21125 : 1.10205 :
85 : 1.45595 : 1.29572 : 1.17890 :
90 : 1.57545 : 1.40207 : 1.27566 :
-----
```

Rows are: power significance level = 0.05 standard deviation = 1.5



This will give an answer apparently accurate to about 6 seconds (since the working units are days and so they should be rounded to one (or at most two) decimal places.

In R, using the routine given in Task Sheet 3 we have

```
> group<-seq(20,30,by=5)
> power<-seq(0.70,0.90,by=0.05)
> group
[1] 20 25 30
> power
[1] 0.70 0.75 0.80 0.85 0.90
> delta<-matrix(nrow=5,ncol=3)
> for (i in 1:5) {
+   for (j in 1:3) {
+     delta[i,j]<-power.t.test(sd=1.5,power=power[i],
+     n=group[j])$delta
+   }
+ }
> options(digits=3)
> delta
      [,1] [,2] [,3]
[1,] 1.21 1.08 0.978
[2,] 1.28 1.14 1.038
[3,] 1.36 1.21 1.103
[4,] 1.46 1.30 1.180
[5,] 1.58 1.40 1.277
```

There are some numerical differences in these but only of the order of about 10 minutes.

