United States
Department of
Agriculture

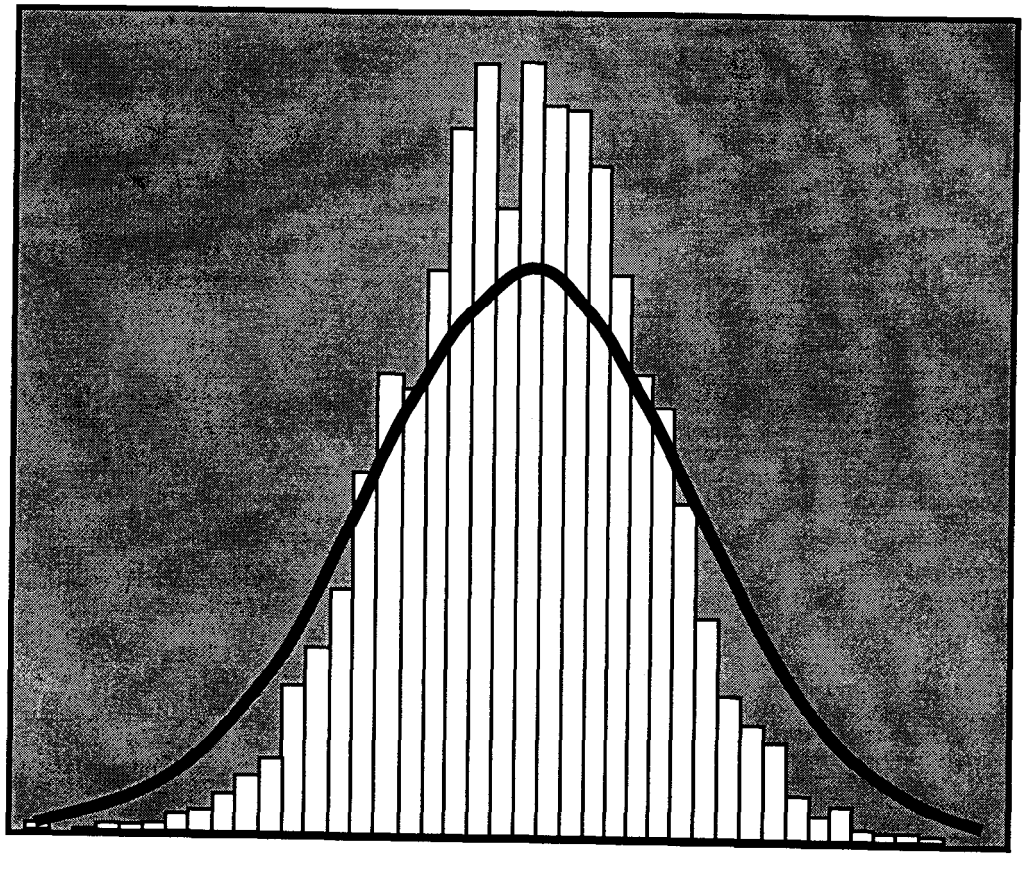
Forest Service

Forest
Products
Laboratory

General
Technical
Report
FPL-GTR-101

TT: A Program That Implements Predictor Sort Design and Analysis

Steve P. Verrill
David W. Green
Victoria L. Herian



Abstract

In studies on wood strength, researchers sometimes replace experimental unit allocation via random sampling with allocation via sorts based on nondestructive measurements of strength predictors such as modulus of elasticity and specific gravity. This report documents TT, a computer program that implements recently published methods to increase the sensitivity of such "predictor sort" experiments. The report consists of annotated keyboard sessions and computer output from runs of TT.

Keywords: program, wood, nondestructive, predictor, sorting

September 1997

Verrill, Steve P.; Green, David W.; Herian, Victoria L. 1997. TT: A program that implements predictor sort design and analysis. Gen. Tech. Rep. FPL-GTR-101. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 21 p.

A limited number of free copies of this publication are available to the public from the Forest Products Laboratory, One Gifford Pinchot Drive, Madison, WI 53705-2398. Laboratory publications are sent to hundreds of libraries in the United States and elsewhere.

The Forest Products Laboratory is maintained in cooperation with the University of Wisconsin.

TT: A Program That Implements Predictor Sort Design And Analysis

Steve P. Verrill, Mathematical Statistician
David W. Green, Supervisory Research General Engineer
Victoria L. Herian, Statistician

Introduction

TT is a computer program that implements the methods developed in Verrill and Green (1996). Currently, it is available in Solaris 1.x, Solaris 2.x, and DOS versions. The program can be obtained by sending a floppy disk to the authors, by e-mail, or via the World Wide Web[1]. The program can also be run over the Web at http://wwwl.fpl.fs.fed.us/ttweb.html.

We encourage you to run the program over the Web rather than on your computer. The Web version will always be up-to-date, you won't encounter out-of-memory errors, and the user interface is better.

To run TT on a PC, you need the DOS executable files tt .exe and dosxmsf.exe. These files are included with the test files analysis.dat, testpr6.dat, and testpr20.dat in a pkzip self-extracting archive, ttzip.exe. After you have obtained ttzip.exe, create a TT directory (e.g., mkdir c: \ttdir) and place ttzip.exe in that directory. Then, while in the directory, type ttzip, and the ttzip.exe archive will unpack itself. To run the TT program, you then simply type tt while in the TT directory. Alternatively, if you place the TT directory in your PC's path statement, you can run the program from anywhere in the directory tree.

This report walks you through the use of the program. It contains three sections: sample size calculations, specimen allocation, and analysis. The report consists of annotated keyboard sessions and computer output. Material printed by the program is flush left and set in typewriter font. Material that you need to type is indented, set in bold type, and followed by <Return> (to indicate the Return or Enter key). Annotations are set in *italics*. If you encounter difficulties in the course of running the program and cannot resolve them through a careful reading of this document, feel free to contact Steve Verrill at 608-231-9375 or by e-mail.

Sample Size Calculations

Sample Size Calculations: Simple – one factor, two levels

To begin the program, type tt:

tt <Return>

What name do you want for your results file?

[1]Steve Verrill, USDA Forest Service Forest Products Laboratory, One Gifford Pinchot Drive, Madison, WI 53705
steve@ws10.fpl.fs.fed.us
http://www1.fpl.fs.fed.us/papers.html

WARNING!!!
Any material that is currently in this file will be lost.
WARNING!!!

myresult <Return>

What do you want to do?

choose a sample size - 0
allocate specimens - 1
analyze results - 2

0 <Return>

What kind of experiment is it?

simple experiment, 1 factor, 2 levels --- 0
complex experiment, multiple factors, multiple levels --- 1

If you ignored the predictor sort nature of your experiment, you would use a standard t test to analyze the results of a simple experiment, and a standard analysis of variance to analyze the results of a complex experiment.

0 <Return>

This program will not tell you what sample size to use. Instead, you must give it information on the sizes of the differences that you want to be able to detect, on the variability of the response (e.g., modulus of rupture (MOR)), on the correlation between the predictor (e.g., modulus of elasticity (MOE)) and the response, on the significance level that you want to achieve, and on the sample sizes that you are considering.

Given this information, the program will calculate the probabilities that you will be able to detect the differences in which you are interested (power). If these probabilities are too low (say below .90), then you will have to find a predictor that is more correlated with the response, accept a larger significance level (say .10 rather than .05), accept a higher risk of not statistically detecting the differences in which you are interested, or be willing to consider larger sample sizes.

How many mean differences (diff) do you want to consider? (5 or fewer)

5 <Return>

What are the differences? (for example, .10 for a 10% difference in means)

.05 .1 .2 .3 .5 <Return>

How many coefficients of variation (CV) do you want to consider? (5 or fewer)
(The coefficent of variation of a property is 100 × (standard deviation)/mean.)

5 <Return>

What are the CVs? (for example, .20 for a 20% coefficient of variation)

.05 .1 .15 .2 .25 <Return>

What will be the significance level of your tests?

.10 - 1
.05 - 2
.01 - 3

2 <Return>

How many sample sizes do you want to consider? (5 or fewer)

3 <Return>

What are the sample sizes? (for example, 10 if you want 10 replicates for EACH treatment)

5 15 25 <Return>

What is the correlation value? (between -.99 and .99)

.7 <Return>

What power calculation approach do you want to take?

To use the power tables, the significance level must be .01 or .05. Also, the diff/CV ratio must lie between 0.0 and 3.0, and the number of replicates per treatment must lie between 2 and 7, or the diff/CV ratio must lie between 0.0 and 1.5, and the number of replicates per treatment must lie between 6 and 48. Otherwise, you cannot interpolate within the tables.

If the tables cannot be used, a non-central T approach is automatically taken.

table/paired --- 1
table/pooled --- 2
table/ANOCOV --- 3
non-central T --- 4

2 <Return>

The material below appears in the output file:

The correlation value is 0.70.

The number of replications for each treatment is 5.

Differences are on top. Coefficients of variation are along the left side. Power values are in the table.

	0.050	0.100	0.200	0.300	0.500
0.050	0.470	0.960	1.000*	1.000*	1.000*
0.100	0.160	0.470	0.960	1.000	1.000*
0.150	0.101	0.261	0.709	0.960	1.000*
0.200	0.080	0.160	0.470	0.800	0.990
0.250	0.070	0.122	0.344	0.622	0.960

A number followed by a * indicates a value that was obtained using a non-central T approach rather than a power table. This approach overestimates power for correlations greater than .8 and small sample sizes (e.g., 6 or fewer replicates). See Verrill and Green's 1996 paper for details.

The .261 entry in row 3 and column 3 of the table indicates that if we had 5 replicates for each of the two treatments, the coefficient of variation were 15% (the standard deviation were 15% of the mean), and the difference between the two treatments were 10% of the mean (for example, one treatment yielded a .95 response and the other yielded a 1.05 response), then we would have only a 26% chance of detecting that difference at a .05 significance level. From the following two tables, we see that we can up this chance to 68% by using 15 replicates per treatment, and to 88% by using ***25 replicates per treatment.***

The number of replications for each treatment is 15.

Differences are on top. Coefficients of variation are along the left side. Power values are in the table.

	0.050	0.100	0.200	0.300	0.500
0.050	0.944	1.000*	1.000*	1.000*	1.000*
0.100	0.440	0.944	1.000*	1.000*	1.000*

0.150	0.243	0.675	0.998	1.000*	1.000*
0.200	0.150	0.440	0.944	1.000	1.000*
0.250	0.115	0.320	0.813	0.986	1.000*

A number followed by a * indicates a value that was obtained using a non-central T approach rather than a power table. This approach overestimates power for correlations greater than .8 and small sample sizes (e.g., 6 or fewer replicates). See Verrill and Green's 1996 paper for details.

The number of replications for each treatment is 25.

Differences are on top. Coefficients of variation are along the left side. Power values are in the table.

	0.050	0.100	0.200	0.300	0.500
0.050	1.000	1.000*	1.000*	1.000*	1.000*
0.100	0.668	1.000	1.000*	1.000*	1.000*
0.150	0.388	0.878	1.000	1.000*	1.000*
0.200	0.219	0.668	1.000	1.000	1.000*
0.250	0.163	0.509	0.961	1.000	1.000*

A number followed by a * indicates a value that was obtained using a non-central T approach rather than a power table. This approach overestimates power for correlations greater than .8 and small sample sizes (e.g., 6 or fewer replicates). See Verrill and Green's 1996 paper for details.

Back to the conversation between the computer and the user:

Another power calculation approach? no - 0 yes - 1

0 <Return>

A different correlation value? no - 0 yes - 1

0 <Return>

Altered differences, CVs, and sample sizes? no - 0 yes - 1

0 <Return>

The program terminates.

Sample Size Calculations: Complex – multiple factors, multiple levels

To start the progmm, type tt:

tt <Return>

What name do you want for your results file?

WARNING!!!
Any material that is currently in this file will be lost.
WARNING!!!

myresult <Return>

What do you want to do?

choose a sample size - 0
allocate specimens - 1
analyze results - 2

0 <Return>

What kind of experiment is it?

simple experiment, 1 factor, 2 levels --- 0
complex experiment, multiple factors, multiple levels --- 1

If you ignored the predictor sort nature of your experiment, you would use a standard t ***test to analyze the results of a simple experiment, and a standard analysis of variance to analyze the results*** **of a** ***complex experiment.***

1 <Return>

This program will not tell you what sample size to use. Instead, you must give it information on the sizes of the differences that you want to be able to detect, on the variability of the response (e.g., MOR), on the correlation between the predictor (e.g., MOE) and the response, on the significance level that you want to achieve, on the numbers of factors and levels, and on the number of replicates that you are considering.

Given this information, the program will calculate the probabilities that you will be able to detect the differences in which you are interested (power). If these probabilities are too low (say below .90), then you will have to find a predictor that is more correlated with the response, accept a larger significance level (say .10 rather than .05), accept a higher risk of not statistically detecting the differences in which you are interested, or be willing to consider larger sample sizes.

NOTE: IF THE NUMBER OF REPLICATES IS SMALL, THEN THESE POWER CALCULATIONS WILL OVER-ESTIMATE THE POWER (AND THUS UNDERESTIMATE THE REQUIRED SAMPLE SIZE) UNLESS THE DATA IS ANALYZED VIA AN ANALYSIS OF COVARIANCE.

Do you already have values of the predictor (e.g., MOE) for all of your specimens?

no - 0
yes - 1

1 <Return>

What is the name of your data file? (It must contain two columns of data. The first column must contain specimen IDs. The IDs must not contain blanks. Only the first 20 characters of the IDs are retained. There must be at least one space between the two columns. The second column must contain the predictor values for the specimens. There may be no more than 5000 specimens.)

testpr6.dat <Return>

How many factors are there? (5 or fewer) (e.g., 3 for a three-way ANOVA)

3 <Return>

How many levels are there? (e.g., 3 2 2 for a 3x2x2 ANOVA) (There must be a total of 25 or fewer levels.)

3 2 2 <Return>

How many replicates are there per "cell"? (e.g., to have 5 replicates in a 3x2x2 ANOVA, a total of 5x3x2x2 specimens would be required. Since the program can handle at most 5000 specimens, given your proposed design, there must be 416 or fewer replicates per cell.)

2 <Return>

What is the starting value for the random number generator? (a positive integer less than 1000000000)

The "random" numbers that are generated are exactly determined by the starting value, Thus, if you want a different set of "random" numbers you must supply a different starting value.

86777 <Return>

The factors and the corresponding numbers of levels are given below. For tests of which factor do you want to estimate power values?

factor	number of levels
1	3
2	2
3	2

1 <Return>

Which kind of test do you want to consider?

a test for a difference between 2 PRE-SPECIFIED levels of the factor --- 0
a test for some difference among all of the levels of the factor --- 1

The PRE-SPECIFIED ***requirement is associated with a statistical subtlety that you might want to discuss with a statistician. It involves the issue of "multiple comparisons."***

1 <Return>

What are the hypothesized means as a fraction of the overall mean? (For example, if there are five levels for the factor, and you expect that level 1 will yield a mean that is about 80% of the overall average, levels 2, 3, and 4 will yield means that are approximately equal to the overall average, and level 5 will yield a mean that is about 120% of the overall average, then your response should be .8 1 1 1 1.2)

Obviously, if you already knew the ratios of the responses for the different levels, then you wouldn't need to do the experiment. Thus, these values must be approximate. However, they will ***enable you to get a feeling for whether your proposed sample sizes will be sufficient.***

.8 1 1.05 <Return>

How many correlation values do you want to consider? (5 or fewer)

2 <Return>

What are they? (between -.99 and .99)

.6 .7 <Return>

How many coefficients of variation do you want to consider? (5 or fewer)

5 <Return>

What are they? (e.g., .25 if the standard deviation of the response population is approximately 25% of its mean)

.05 .1 .15 .2 .25 <Return>

What power do you wish to achieve?

(e.g., .90 if you want to have a 90% chance of obtaining a statistically significant result given that the difference actually exists)

.9 <Return>

The following material appears in the output file:

The correlation value is 0.6000E+00

Power tables

The tables are produced for three significance levels – .10, .05, and .01.

The significance level is 0.10

The coefficients of variation and associated power values are

0.050 1.000
0.100 1.000
0.150 0.982
0.200 0.868
0.250 0.702

The significance level is 0.05

The coefficients of variation and associated power values are

0.050 1.000
0.100 1.000
0.150 0.956
0.200 0.771
0.250 0.567

The significance level is 0.01

The coefficients of variation and associated power values are

0.050 1.000
0.100 0.997
0.150 0.822
0.200 0.502
0.250 0.294

Replication tables

These tables give the number of replications of the current design that are needed to achieve the desired power.

A 1 next to a coefficient of variation value indicates that, given the variability indicated by the coefficient of variation, the current design is sufficient to yield the desired power; a 2 indicates that the current design will not yield sufficient power, but if the number of replicates were doubled, the desired power would be achieved; a 3 indicates that a doubling of the number of replicates would not yield the desired power, but a tripling would; and so on.

The tables are produced for three significance levels – .10, .05, and .01.

The desired power is 0.90

The significance level is 0.10

The coefficients of variation and associated replication values are

0.050 1
0.100 1
0.150 1
0.200 2
0.250 2

The significance level is 0.05

The coefficients of variation and associated replication values are

0.050 1
0.100 1
0.150 1
0.200 2
0.250 2

The significance level is 0.01

The coefficients of variation and associated replication values are

0.050 1
0.100 1
0.150 2
0.200 2
0.250 3

The correlation value is 0.7000E+00

Note that as correlation goes up, power goes up (so sample sizes ***could come down).***

Power tables

The significance level is 0.10

The coefficients of variation and associated power values are

0.050 1.000
0.100 1.000
0.150 0.995
0.200 0.929
0.250 0.791

The significance level is 0.05

The coefficients of variation and associated power values are

0.050 1.000
0.100 1.000
0.150 0.985
0.200 0.862
0.250 0.671

The significance level is 0.01

The coefficients of variation and associated power values are

0.050 1.000
0.100 1.000
0.150 0.914
0.200 0.632
0.250 0.390

Replication tables

These tables give the number of replications of the current design that are needed to achieve the desired power.

The desired power is 0.90

The significance level is 0.10

The coefficients of variation and associated replication values are

0.050 1
0.100 1
0.150 1
0.200 1
0.250 2

The significance level is 0.05

The coefficients of variation and associated replication values are

0.050 1
0.100 1
0.150 1
0.200 2
0.250 2

The significance level is 0.01

The coefficients of variation and associated replication values are

0.050 1
0.100 1
0.150 1
0.200 2
0.250 3

Back to the conversation between the computer and the user:

Would you like to perform additional calculations based on altered sample sizes, correlations, CVs, etc.? no - 0 yes - 1

0 <Return>

The program terminates.

Specimen Allocation

Given the values of a predictor, the program will perform the predictor sort allocation for you.

Specimen Allocation: Simple – one factor, two levels

To start the program, type tt:

tt <Return>

What name do you want for your results file?

WARNING!!!
Any material that is currently in this file will be lost.
WARNING!!!

myresult <Return>

What do you want to do?

choose a sample size - 0
allocate specimens - 1
analyze results - 2

1 <Return>

What kind of experiment is it?

simple experiment, 1 factor, 2 levels --- 0
complex experiment, multiple factors, multiple levels --- 1

0 <Return>

What is the name of your data file? (It must contain two columns of data. The first column must contain specimen IDs. The IDs must not contain blanks. Only the first 20 characters of the IDs are retained. There must be at least one space between the two columns. The second column must contain the predictor values for the specimens. There may be no more than 5000 specimens.)

testpr20.dat <Return>

Here is a listing of the testpr20.dat data file. The first column contains a specimen ID. These IDs must not contain blanks, and, of course, they need to be distinct. For this example the second column was randomly generated from a normal distribution with mean 0 and standard deviation 1. However, in general, it would contain measured values of some predictor (e.g., MOE).

1 -0.12528
2 0.12897
3 0.88499
4 -0.32207
5 0.87034

```
 6   -0.02332
 7   -1.05039
 8    0.45095
 9   -1.62132
10   -0.49165
11    0.04958
12   -0.50156
13   -0.99610
14   -0.34506
15    1.29276
16   -1.82928
17    1.11773
18    0.59903
19   -0.10677
20   -1.23812
```

End of the data file.

How many replicates are there per treatment? (Since the program can handle at most 5000 specimens, there must be 2500 or fewer replicates per treatment.)

10 <Return>

What is the starting value for the random number generator? (It must be a positive integer less than 1000000000.)

The "random" numbers that are generated are <u>exactly</u> determined by the starting value. Thus, if you want a different set of "random" numbers, you must supply a different starting value.

3344256 <Return>

Do you want to name the treatments? no - 0 yes - 1

1 <Return>

What are the names of the two treatments? You must type one name per line, and the names may contain no more than 10 characters.

himom1 <Return>
himom2 <Return>

The program terminates.

The following material appears in ***the output file:***

The ID of the specimen, the predictor value, and the treatment are

1	-0.125283+00	himom2
2	0.128973+00	himom1
3	0.88499E+00	himom1
4	-0.322073+00	himom1
5	0.87033E+00	himom2
6	-0.233166-01	himom1
7	-0.l0504E+0l	himom2
8	0.45095E+00	himom2
9	-0.16213E+01	himom1
10	-0.49165E+00	himom2
11	0.495833-01	himom2
12	-0.50156E+00	himom2
13	-0.99610E+00	himom1
14	-0.34506E+00	himom1
15	0.129283+01	himom2
16	-0.182933+01	himom2
17	0.11177E+0l	himom1
18	0.59903E+00	himom1
19	-0.10677E+00	himom2
20	-0.123813+01	himom1

As is required, the program splits the two lowest predictor values (specimens 9 and 16) between the two groups. Similarly, the two highest predictor values (specimens 15 and 17) are split between the two groups, And so on.

The allocation is presented in two equivalent forms: Given a specimen, what is the treatment? Given a treatment, what are the specimens?

In addition, a third table is provided that would be useful *if* ***you wanted to use this program to analyze the resulting data via an analysis of covariance.***

The ID of the predictor block, the ID of the treatment, the ID of the specimen, and the predictor value are

1	himom1	9	-0.l6213E+01
1	himom2	16	-0.18293E+01
2	himom1	20	-0.12381E+01
2	himom2	7	-0.10504E+01
3	himom1	13	-0.99610E+00
3	himom2	12	-0.50156E+00
4	himom1	14	-0.34506E+00
4	himom2	10	-0.49165E+00
5	himom1	4	-0.32207E+00
5	himom2	1	-0.12528E+00
6	himom1	6	-0.23316E-01
6	himom2	19	-0.l0677E+00
7	himom1	2	0.l2897E+00
7	himom2	11	0.49583E-01

```
 8   himom1   18   0.59903E+00
 8   himom2    8   0.45095E+00
 9   himom1    3   0.88499E+00
 9   himom2    5   0.87033E+00
10   himom1   17   0.11177E+01
10   himom2   15   0.12928E+01
```

The ID of the predictor block, the predictor value for the first treatment, and the predictor value for the second treatment are (this information would be needed in this form if you used this program to perform an analysis of covariance):

```
 1   -0.l6213E+01   -0.18293E+01
 2   -0.123811E+01  -0.10504E+01
 3   -0.99610E+00   -0.50156E+00
 4   -0.34506E+00   -0.49165E+00
 5   -0.32207E+00   -0.12528E+00
 6   -0.23316E-01   -0.10677E+00
 7    0.12897E+00    0.49583E-01
 8    0.59903E+00    0.45095E+00
 9    0.88499E+00    0.87033E+00
10    0.11177E+01    0.12928E+01
```

Specimen Allocation: Complex – multiple factors, multiple levels

To start the program, type tt:

```
tt <Return>
```

What name do you want for your results file?

```
WARNING!!!
Any material that is currently in this file will be lost.
WARNING!!!
```

```
myresult  <Return>
```

What do you want to do?

```
choose a sample size - 0
allocate specimens   - 1
analyze results      - 2
```

```
1 <Return>
```

What kind of experiment is it?

```
simple experiment, 1 factor, 2 levels --- 0
```

complex experiment, multiple factors, multiple levels --- 1

1 <Return>

What is the name of your data file? (It must contain two columns of data. The first column must contain specimen IDs. The IDs must not contain blanks. Only the first 20 characters of the IDs are retained. There must be at least one space between the two columns. The second column must contain the predictor values for the specimens. There may be no more than 5000 specimens.)

testpr6.dat <Return>

How many factors are there? (5 or fewer) (e.g., 3 for a three-way ANOVA)

3 <Return>

How many levels are there? (e.g., 3 2 2 for a 3x2x2 ANOVA) (There must be a total of 25 or fewer levels.)

3 2 2 <Return>

Do you want to name the levels? no - 0 yes - 1

1 <Return>

What are the names of the 3 levels of factor 1? You must type one name per line, and the names may contain no more than 10 characters.

small <Return>
medium <Return>
large <Return>

What are the names of the 2 levels of factor 2? You must type one name per line, and the names may contain no more than 10 characters.

cold <Return>
hot <Return>

What are the names of the 2 levels of factor 3? You must type one name per line, and the names may contain no more than 10 characters.

wet <Return>
dry <Return>

How many replicates are there per "cell"? (e.g., to have 5 replicates in a 3x2x2 ANOVA, a total of 5x3x2x2 specimens would be required.) (Since the program can handle at most 5000 specimens, given your proposed design, there must be 416 or fewer replicates per cell.)

2 <Return>

What is the starting value for the random number generator? (a positive integer less than 1000000000)

The "random" numbers that are generated are exactly determined by the starting value. Thus, if you want a different set of "random" numbers you must supply a different starting value.

432567 <Return>

The program terminates.

The allocation is presented in two equivalent forms: Given a "treatment," what is the specimen? Given a specimen, what is the "treatment?"

Note that there are 3x2x2 specimens per "block" and only two blocks. For the large sample theory described in Verrill and Green (1996) to apply, more replicates are needed. Thus, in this case, the data should be analyzed as a blocked analysis of variance, or, if the relationship between the predictor and the response is linear, as an analysis of covariance. Commercial programs to perform the necessary analyses are widely available.

The following material appears in the output file:

The ID of the predictor block, the levels of the factors, the ID of the specimen, and the predictor value are

1	small	cold	wet	119	-0.10677E+00
1	small	cold	dry	17	-0.10504E+01
1	small	hot	wet	121	-0.99060E+00
1	small	hot	dry	19	-0.l6213E+0l
1	medium	cold	wet	114	-0.34506E+00
1	medium	cold	dry	112	-0.50156E+00
1	medium	hot	wet	110	-0.49165E+00
1	medium	hot	dry	14	-0.32207E+00
1	large	cold	wet	113	-0.99610E+00
1	large	cold	dry	116	-0.18293E+01
1	large	hot	wet	11	-0.12528E+00
1	large	hot	dry	120	-0.12381E+01
2	small	cold	wet	111	0.49583E-01
2	small	cold	dry	123	0.99494E+00
2	small	hot	wet	122	-0.70177E-01
2	small	hot	dry	13	0.88499E+00
2	medium	cold	wet	118	0.59903E+00
2	medium	cold	dry	16	-0.23316E-01
2	medium	hot	wet	15	0.87033E+00
2	medium	hot	dry	117	0.11177E+01
2	large	cold	wet	18	0.45095E+00
2	large	cold	dry	115	0.l2928E+01
2	large	hot	wet	12	0.12897E+00
2	large	hot	dry	124	-0.93164E-01

The ID of the specimen, the predictor value, and the levels of the factors are

11	-0.12528E+00	large	hot	wet
12	0.12897E+00	large	hot	wet
13	0.88499E+00	small	hot	dry
14	-0.32207E+00	medium	hot	dry
15	0.87033E+00	medium	hot	wet
16	-0.23316E-01	medium	cold	dry
17	-0.l0504E+01	small	cold	dry
18	0.45095E+00	large	cold	wet
19	-0.16213E+01	small	hot	dry
110	-0.49165E+00	medium	hot	wet
111	0.49583E-01	small	cold	wet
112	-0.50156E+00	medium	cold	dry
113	-0.99610E+00	large	cold	wet
114	-0.34506E+00	medium	cold	wet
115	0.12928E+01	large	cold	dry
116	-0.18293E+01	large	cold	dry
117	0.11177E+01	medium	hot	dry
118	0.59903E+00	medium	cold	wet
119	-0.10677E+00	small	cold	wet
120	-0.l2381E+01	large	hot	dry
121	-0.99060E+00	small	hot	wet
122	-0.70177E-01	small	hot	wet
123	0.99494E+00	small	cold	dry
124	-0.93164E-01	large	hot	dry

Analysis

As currently written, TT can only analyze data from simple predictor sort experiments (two treatments). However, standard statistical packages can be used to analyze more complex predictor sort experiments. In particular, if the number of replicates is "large" and the data are not "blocked" by predictor value, you can divide the F statistic by $1 - r^2$ where r is the correlation between the predictor and the response. Standard F tables can then be used to evaluate the significance of the resulting F value. Alternatively, if the data are blocked by predictor value (the block value is included as a factor in the analysis of variance), you can simply use the standard output from the package. Finally, if the relationship between the predictor and the response is linear, you can perform an analysis of covariance. This is the optimal approach if the relationship between predictor and response is truly linear. However, if it deviates from linearity, the blocked analysis of variance is preferable.

To start the program, type tt:

```
tt <Return>
```

What name do you want for your results file?

WARNING!!!

Any material that is currently in this file will be lost.
WARNING!!!

myresult <Return>

What do you want to do?

choose a sample size - 0
allocate specimens - 1
analyze results - 2

2 <Return>

This program will only analyze simple l-factor, 2-level experiments. If your experiment is more complex, you need to use an analysis of variance package. See Verrill and Green's 1996 paper for details.

Continue with the analysis? no - 0 yes - 1

1 <Return>

What is the name of your data file?

(The first column must contain the response values for treatment 1. The second column must contain the matched response values for treatment 2. If you want to perform an analysis of covariance, you must also include two additional columns. The third column must contain the predictor values associated with the column one response values. The fourth column must contain the predictor values associated with the column two response values. There can be at most 2500 pairs of observations.)

analysis.dat <Return>

This is the data set described in section 3 of Verrill and Green (1996). It is listed in Table 56 of that paper.

Does the file include the predictor value data? no - 0 yes - 1

1 <Return>

Are you just interested in the analysis of covariance results? no - 0 yes - 1

0 <Return>

How do you want the correlation to be determined?

from the data - 0
experience - 1

(If the correlation is less than .95 and it is known from experience to within .05 or .10, there should be no problem. Alternatively, if the correlation is less than .95 and the number of replicates per treatment is at least 6, the correlation can be successfully estimated from the data. If neither of these conditions holds, then a paired tight t or an analysis of covariance approach should be taken rather than a pooled tight t approach.)

0 <Return>

The estimated correlation is 0.7220E+00.

The paired tight t value is 0.3612E+01.
There are 45 replicates.
The (two-sided) p value is less than .01.

The pooled tight t value is 0.2991E+01.
There are 45 replicates.
The (two-sided) p value is less than .01.

The analysis of covariance t value is 0.3030E+01.
There are 45 replicates.
The (two-sided) p value is 0.3224E-02.

The program terminates.

Reference

Verrill, S. and Green, D. (1996). Predictor sort sampling, tight t's, and the analysis of covariance: theory, tables, and examples. Research Paper FPL-RP-558. U.S. Department of Agriculture Forest Service Forest Products Laboratory, Madison, Wisconsin.

Acknowledgments

The authors would like to express their deep appreciation to Dr. Mark Durst of the Lawrence Berkeley Laboratory, and to Dr. Jerry Winandy and Mr. Rusty Dramm of the Forest Products Laboratory for evaluating the TT computer program. Their many valuable suggestions greatly improved the final product.