Paper 205-25

A Randomization Test: Somatic Cell Count Slope Comparisons of 55 California Dairies

Paul Johnson
University of California San Francisco

ABSTRACT

A SAS® macro for comparing two slope estimates using a randomization test is presented. A milking practice evaluation score was obtained for 55 Dairy Herd Improvement Association (DHIA) dairies. The category for general sanitation practices consisted of obtaining the total score for four questions. The median score over all dairy responses was used to provide a cut-off point in order to discriminate each dairy's response. Responses were recorded as being either below the median or above/equal to the median. Somatic Cell Count (SCC) data were also obtained for the dairies. The SCC values are an indicator to whether a cow has good or poor udder health. These counts were used as the dependent variable to examine the relationship between SCC and general sanitation for the two groups. Models were fit for each of the two groups and the slopes of the two models compared. A randomization test was constructed to examine whether or not the two slopes differed significantly from one another. This macro requires base SAS, SAS/STAT and SAS/GRAPH software to run.

INTRODUCTION

Goodger et al. (1993) obtained a milking practice evaluation score for 55 randomly selected California dairies. Four of the 48 questions asked at each dairy were used to obtain a general sanitation score (C1). The questions were:

1) Is there standing water, manure and mud in the alleyways (or pens)?
2) What efforts does the manager take to keep the cow yard clean and dry?
3) How frequently are the milking parlor floors and walls washed?
4) Are the milking machines rinsed, washed and sanitized between shifts?

The scores to the 4 questions were summed to give a total general sanitation score (C1) for each dairy. Each dairy's management practices were evaluated on four occasions. The median score for all C1 values was obtained and used to construct two groups. The median (M) was obtained by using the UNIVARIATE procedure of SAS (see base SAS, 1985). One group consisted of all observations below the median and the second those above (or equal to) the median.

There are two sources of somatic cells in milk. One source is due to the natural loss of secretory cells in the udder and the other source involves the production of leukocytes (white blood cells). A high somatic cell count (SCC) in milk means either the udder is injured mechanically or is infected by a disease organism. A dairy manager wishes to keep the dairy's SCC values to a minimum. The C1 scores and SCC values were measured for each of the 55 dairies on four occasions. These SCC scores were then used to obtain two linear models relating SCC to C1.

The method of least squares was used to find the parameter estimates for each of the two models (see Neter et al., 1985).

PROCEDURE

The two models developed, with $y =$ SCC and $x =$ C1, are:

Model 1: $y = \beta_{01} + \beta_{11} x$ for $x \le \mathrm{M}$.

Model 2: $y = \beta_{02} + \beta_{12} x$ for $x > \mathrm{M}$.

The β s are estimated using PROC REG (see SAS/STAT, 1993). The randomization test procedure to test for the difference between the slope estimates is used to test whether or not there exists a significant difference between the slopes. Manly (1997) discusses the use of the randomization test to test the difference between mean scores. This randomization test procedure tests the difference between two slopes. The procedure is:

a) The difference between the slope estimates is calculated via:

$$\mathrm{D}_1 = \hat{\beta}_{11} - \hat{\beta}_{12}$$

b) Randomly allocate the (x, SCC) pairs of observations between the two groups and calculate the new difference D_2.

c) Repeat step (b) a large number of times to find a sample of values from the distribution of D that occurs by allocating the pairs. This sample forms the randomization distribution.

d) If D_1 looks like a typical value from the randomization distribution then conclude that there is no significant change in slope moving from low general sanitation conditions to high general sanitation conditions. If we are performing a two-tailed test then if D_1 is unusually large or small then the data are unlikely to have arisen if the null hypothesis of equal slopes is true. It can then be concluded that the alternative hypothesis (the slopes differ) is more plausible. The null hypothesis to be tested is:

$$\mathrm{H}_0 \colon \mathrm{D} = 0$$

where $\mathrm{D} = \beta_{11} - \beta_{12}$.

The user needs to input the number of sample randomizations that need to be carried out (n_random). In the code (see below) n_random is set equal to 999, since then with the additional estimate of the difference obtained from the original data the total number of estimates used in constructing the p-value totals 1000.

The user also needs to input the value of 'alt', where:

1) alt = 1 indicates $\mathrm{H}_\mathrm{A} \colon \mathrm{D} > 0$ (one-tailed test);
2) alt = 2 indicates $\mathrm{H}_\mathrm{A} \colon \mathrm{D} < 0$ (one-tailed test); and
3) alt = 3 indicates $\mathrm{H}_\mathrm{A} \colon \mathrm{D} \ne 0$ (two-tailed test).

In the code (see below) 'alt' is set equal to 3 for a two-tailed test. To summarize the randomization results the p-value for a two-tailed test is obtained by calculating the proportion of all observed

D values that are either greater than or equal to the absolute value of D_1 or less than or equal to the negative of the absolute value of D_1. Including D_1 in the numerator and denominator is justified since if H_0 is true then D_1 is just another value from the randomization distribution.

If alt = 1 then the p-value is obtained by calculating the proportion of all observed D values that are greater than or equal to D_1.

If alt = 2 then the p-value is obtained by calculating the proportion of all observed D values that are less than or equal to D_1.

RESULTS

Summary statistics for the SCC values and C1 scores follow (see Table 1). The slope estimates are derived and the value for D_1 is computed.

Table 1. SCC and C1 Summary Statistics (overall and by model)

Overall

Variable	Mean	Standard Deviation
C1	48.75	12.25
SCC	248.83	115.31

Model 1:

Variable	Mean	Standard Deviation
C1	38.64	7.91
SCC	263.55	127.16

$$\hat{\beta}_{11} = -6.5593$$

Model 2:

Variable	Mean	Standard Deviation
C1	58.57	6.25
SCC	234.55	101.08

$$\hat{\beta}_{12} = -2.3315$$

$$D_1 = \hat{\beta}_{11} - \hat{\beta}_{12} = -6.5593 - (-2.3315) = -4.2278.$$

A two-tailed test was conducted with 999 randomizations (i.e., a total of 1000 D values were computed). It was found that 82 of the 1000 D values either fell below the value −4.2278 or above the value 4.2278 [40 of the values fell below −4.2278 and 42 fell above 4.2278]. The calculated p-value = (82+1)/(999+1) = 0.083. This is the significance level obtained for a two-sided test.

SAS/GRAPH (1991) is used to produce Figures 1, 2 and 3.

Figure 1 is a plot of the mean SCC observed values against the C1 scores for the two models. Figure 2 is a plot of the predicted mean SCC values against the C1 scores for the two models. The lines fit give a visual indication to whether or not there is a difference in slopes. Figure 3 is a bar chart showing the frequency distribution for the randomization results. Manly (1997) observes the fact that randomization tests and classical parametric tests tend to have similar power when the conditions for the parametric test are justified. The author continues to say that with data from non-standard distributions there is some evidence to suggest that randomization tests are more powerful than classical alternatives.

How many randomizations should be carried out? Marriott (1979) proposed that for a test of significance at the 5% level 1000 randomizations is a realistic minimum. For a test at the 1% level 5000 randomizations is a realistic minimum. Marriott was discussing Monte Carlo tests but the principle applies also to randomization tests.

Figure 1. Plot of the Mean SCC Observed Values and C1 Scores for the Two Models

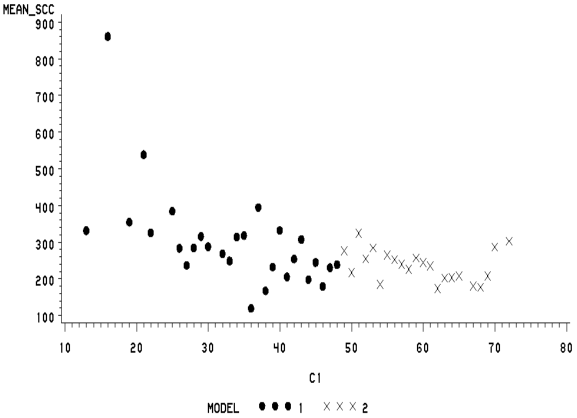


Figure 2. Plot of the Predicted Mean SCC Values and C1 Scores for the Two Models

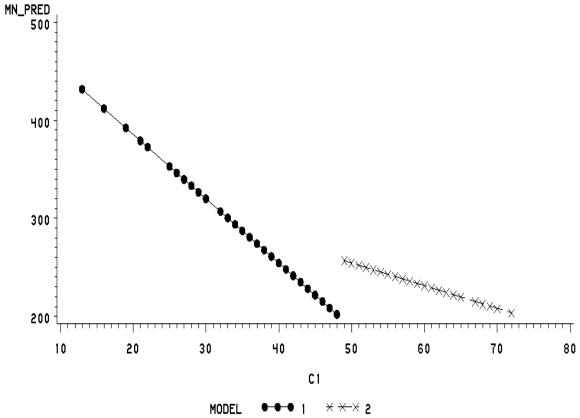
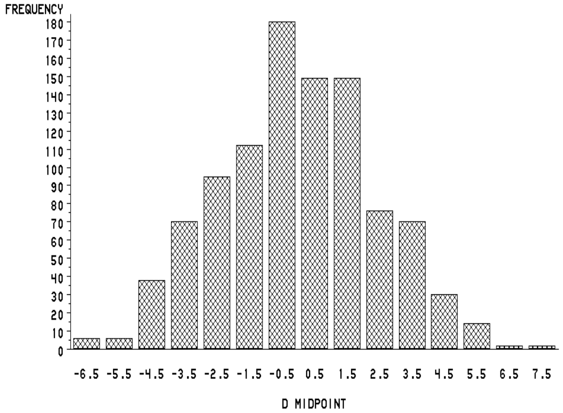


Figure 3. Frequency Distribution for the Randomization Results



SUMMARY

What does this mean in terms of the somatic cell count/general sanitation relationship for the 55 sampled dairies. At the α = 0.05 level of significance the p-value>α: hence we would have no reason to reject the null hypothesis. We would conclude that the slopes do not differ and that by moving from areas of relatively low sanitation to areas of relatively high sanitation the rate of decline in somatic cell count remains the same. At the α = 0.10 level of significance the p-value<α: hence we would have reason to reject the null hypothesis in favor of the alternative. It may be argued that in this situation we would only be interested in a one-sided alternative, namely $H_A: D < 0$. Then we would only be interested in whether the rate of decline in SCC values is less for higher general sanitation levels as compared to lower general sanitation levels. If this were the case then the p-value = (40+1)/(999+1) = 0.041 leading to the rejection of the null hypothesis at the α = 0.05 level of significance. The slope is not as steep for those observations with higher general sanitation conditions. What does this mean to the dairy farmer? All farms benefit by increases in general sanitation scores. Higher general sanitation scores will result in lower levels of SCC. However the level of reduction in SCC is steeper for those farms operating at low scores of general sanitation than those operating at relatively higher scores of general sanitation.

CODE

```
/*-----------------------------------------------------------------------------------*
|                                                                                   |
|  SAS Macro using a randomization test procedure to                                |
| compare  two slopes: Randomization Test procedure in                              |
|  linear regression. The one(2) tailed test for comparing the                      |
|  slopes of 2 models is conducted and n_random number of                           |
|  randomizations are carried out to compute an overall                             |
|  significance level. The user needs to input the number of                        |
|  randomizations to be carried out, and the value for ‘alt’.                       |
                                                                                    |
|  The null hypothesis to test for the equality of the 2 slopes:                    |
|                                                                                   |
|                    H0: slope1 - slope2 = 0                                        |
|                                                                                   |
|  against one of 3 alternative hypotheses:                                         |
|                                                                                   |
|  Ha: slope1 - slope2 > 0     (alt = 1)                                            |
|        slope1 - slope2 < 0    (alt = 2)                                           |
|        slope1 - slope2 =/= 0 (alt = 3)                                            |
|                                                                                   |
-------------------------------------------------------------------------------------*/
  libname old 'd:\btscc';
  OPTIONS NODATE NONUMBER PAGESIZE = 60
  LINESIZE =132;
    goptions cback=white
         colors=(black cyan yellow green blue magenta);

data ctrl;set old.btscc;

/*-----------------------------------------------------------------------------------*
|                                                                                   |
|  The data consists of the Somatic Cell Count (SCC) data                           |
|  and the category score #1 (C1) which describes the general                       |
|  sanitation conditions of the dairy farms. The SCC and C1                         |
|  scores were recorded on 55 randomly selected Dairy Herd                          |
|  Improvement Association (DHIA) dairies.                                          |
|                                                                                   |
|  The median of the C1 scores is used to separate the                              |
|  observations into 2 groups. Those values with a C1 score                         |
|  less than the median form one group and those with a C1                          |
|  score greater or equal to the median form the other. Linear                      |
|  models are fit to the 2 groups of data and slopes compared.                      |
|                                                                                   |
-------------------------------------------------------------------------------------*/

data ctrl;set ctrl;

proc univariate data = ctrl;var c1;
output out = univ median = median;
proc means data = ctrl;var scc c1;

data actrl;set ctrl;dummy = 'a'; data univ;set univ;dummy = 'a';

proc sort data = actrl;by dummy; proc sort data = univ;by dummy;

data ctrl;merge actrl univ;by dummy;drop dummy;
y = scc;x =c1;If c1 < median then model = 1;else model = 2;

proc sort data = ctrl;by model;
proc means noprint data = ctrl;by model;var scc c1;
output out = cctrl;

proc print data = cctrl;

data ctrl;set ctrl; keep y x model;

proc sort data = ctrl;by model;
proc reg data = ctrl outest = ma covout;by model;
model y = x; output out = ctrl2 pred = pred;
TITLE1 'Randomization-Test Results for Comparing the Slopes of
2 Models';

data mc;set ma;if _type_ = 'PARMS';beta = x;keep beta;

proc transpose data = mc out = me;var beta;

/*-------------------------------*
|  d = slope1 - slope2      |
-------------------------------*/

data me;set me;d= col1-col2;
proc print data = me;var d;
TITLE2 'Difference Estimates for Comparing the Slopes of two
Models';

data ctrl;set ctrl;ind=_N_;

/*-------------------------------------------------------------------------------*
|                                                                               |
|  ‘alt’ is set equal to 3 for a two-tailed test. n_random is                   |
|  set equal to 999. Together with the original d this                          |
|  provides for a total sampling frequency distribution of                      |
|  1000 values.                                                                 |
|                                                                               |
---------------------------------------------------------------------------------*/

%macro random;

 %let n_random=999;%let alt = 3;

       %do i=1 %to &n_random;

       data cb&i;set ctrl;drop ind x model;
        seed=floor(1000000000*(sqrt(time())-floor(sqrt(time()))));
         k=500*(ranuni(seed));
    proc sort data = cb&i;by k;

    data cb&i;set cb&i; ind=_N_;
    proc sort data = ctrl;by ind;
     proc sort data = cb&i;by ind;
      data acb&i; merge ctrl cb&i;by ind;

    proc sort data = acb&i;by model;
      proc reg data = acb&i outest = ma&i covout noprint;
      model y = x; by model;
       data mc&i;set ma&i;if _type_ = 'PARMS';
       keep x indx;indx = _N_;
        data mc&i;set mc&i;beta = x;keep beta;

     proc transpose data = mc&i out = me&i;var beta;
     data me&i;set me&i;d= col1-col2;
 %end;

   %do j = 2 %to &n_random;
   proc append base = me1 data = me&j;
   %end;

data me;set me;dummy = 'a'; d2 = d;drop d;
data me1;set me1;dummy = 'a';

proc sort data = me;by dummy;
proc sort data = me1;by dummy;

data mf;merge me me1;by dummy;
data mf;set mf;
 if (d >= d2 and &alt = 1) then ind_d =1;
    if (d < d2  and &alt = 1) then ind_d = 0;
        if (d <= d2 and &alt = 2) then ind_d =1;
           if (d > d2  and &alt = 2) then ind_d = 0;
       if (d >= abs(d2) and &alt = 3) then ind_d =1;
   if (d < abs(d2) and d > -1*abs(d2) and &alt = 3) then ind_d = 0;
 if (d <= -1*abs(d2) and &alt = 3) then ind_d =1;

proc means data = mf noprint;var ind_d;
output out = mdmc sum = sumlevel n = n_random;

proc sort data = mf;by d;
proc print data = mf;var d;
TITLE2 'Randomization Results for Testing for a Non-Zero
Difference Value';

data mdmc;set mdmc;p_value = (sumlevel+1)/(n_random+1);
n_2 = sumlevel;

%mend random;

%random;

proc print data = mdmc;var p_value n_random n_2;
TITLE2 'Significance Level (p_value) and Number of
Randomizations Carried Out';

proc sort data = ctrl;by model x;
proc means data = ctrl noprint;by model x;var y;
output out = ctrl3 mean = mean_scc;

data ctrl3;set ctrl3;c1=x;

/*-------------------------------------------------------------------------------*
|                                                                               |
|  Plots of Observed Mean and Predicted SCC values                              |
|                                                                               |
-------------------------------------------------------------------------------*/

 proc gplot data = ctrl3;
 plot mean_scc*c1=model/caxis = black;
 symbol1 color = black v=dot;
 symbol2 color = black v=x;
 Title1 'Plots of the Mean SCC Values (Observed) against the C1
Values for the 2 Models';

proc sort data = ctrl2;by model x;
proc means data = ctrl2 noprint;by model x;var pred;
output out = ctrl4 mean = mn_pred;

data ctrl4;set ctrl4;c1=x;

 proc gplot data = ctrl4;
 plot mn_pred*c1=model/caxis = black;
 symbol1 interpol = join color = black line=1 v=dot;
 symbol2 interpol = join color = black line = 2 v=x;
 Title1 'Plots of the Predicted Mean SCC Values against the C1
Values for the 2 Models';

/*------------------------------------------------------------------------*
|                                                                        |
|     Frequency Distribution of the Randomizations                       |
|                                                                        |
------------------------------------------------------------------------*/

proc gchart data = mf;vbar d;
Title1 'Frequency Distribution as a Chart Diagram for the
Randomization Results';

run;
```

REFERENCES

Goodger, W.J., Farver, T., Pelletier, J., Johnson, P., DeSnayer, G., and Galland, J. (1993), “The Association of Milking Management Practices with Bulk Tank Somatic Cell Counts,” *Preventive Veterinary Medicine*, 15(4), 235-251.

Manly, B.F.J. (1997), *Randomization, Bootstrap and Monte Carlo Methods in Biology*, New York: Chapman and Hall.

Marriott, F.H.C. (1979), Barnard’s Monte Carlo Tests: How many Simulations?” *Applied Statistics*, 28, 75-77.

Neter, J., Wasserman, W., and Kutner, M.H. (1985), *Applied Linear Statistical Models*, Homewood, Illinois: Richard D. Irwin, Inc.

SAS Institute. Inc. (1985), *SAS User’s Guide: Basics, Version 5 Edition*, Cary, NC: SAS Institute Inc.

SAS Institute. Inc. (1991), *SAS/GRAPH User's Guide, Version 6*, 3rd ed., Cary, NC: SAS Institute. Inc.

SAS Institute. Inc. (1993), *SAS/STAT User's Guide, Version 6, 4th ed., Volume 1 and Volume 2,* Cary, NC: SAS Institute Inc.

ACKNOWLEDGMENTS

SAS, SAS/STAT and SAS/GRAPH are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

CONTACT INFORMATION

Paul Johnson, P.O. Box 4146, Davis CA 95617-4146
E-Mail: JohnsonP12@prodigy.net