Calculating Posterior Probability of the Maximum Contrast Using PROC IML

Akira Wakana, Banyu Pharmaceutical Co., Ltd., Tokyo, Japan
Isao Yoshimura, Tokyo University of Science, Tokyo, Japan
Chikuma Hamada, Tokyo University of Science, Tokyo, Japan

ABSTRACT
In this paper SAS® code for evaluating posterior probabilities when selecting a dose-response pattern from a set of candidate response patterns using contrast statistics in analyzing a phase II clinical trial is introduced. A method for selecting a response pattern with evaluation of the posterior probability that the specified response pattern is the true one has been considered. It is possible to evaluate the posterior probability by numerical integration based on either a multivariate t-distribution for a continuous variable or a multivariate normal distribution for a binary variable with large sample size. SAS code was developed using the ‘quad’ function in PROC IML. The method was illustrated on two datasets.

INTRODUCTION
Suppose that a dose-response clinical trial consisting of several doses of a compound is conducted to investigate dose-response relationship with the objective to determine a dose-response pattern. Pre-study information with respect to response pattern is not precise prior to the clinical trial in general. In this respect, the maximum contrast method formed by taking the maximum over multiple contrast statistics is useful to detect the existence of a dose-response. When a significant dose-response is demonstrated, it is possible to pursue further research for selecting a response pattern. A contrast-based approach for selecting the most plausible response pattern from a set of candidate patterns based on the maximum of observed contrast statistics has been proposed, that is, the response pattern corresponding to contrast coefficients which best fits data is specified. Then, it is necessary to assess the level of plausibility of the specified pattern. The posterior probability that the response pattern corresponding to the maximum of observed contrast statistics is the true one has been considered. SAS code is developed to evaluate posterior probabilities using ‘quad’ function in PROC IML for three- and four-group trials. This paper is organized as follows. In the next section, a method for identifying a response pattern using contrast statistics is described. The third section describes SAS PROC IML code to evaluate posterior probabilities. In the fourth section, examples are given. In the final section, conclusions are made.

PLAUSIBILITY OF THE SPECIFIED RESPONSE PATTERN
In this section a method for selecting a response pattern is explained. Denote a set of increasing dose levels by 1,2,...,a. Let Y ij (i=1,2,...,a; j=1,2,...,n a) be observations with sample means \( \bar{Y} = (\bar{Y}_1, \bar{Y}_2, ..., \bar{Y}_a) \), and assume that Y ij's are independent normal variables with mean \( \mu_i \) and a common variance \( \sigma^2 \). Let \( c_k \) (k=1,2,...,m) be the k-th vector of known constants subject to \( c_k'1=0 \). A contrast statistic corresponding to \( c_k \) is defined as

\[
T_k = \frac{c_k'\bar{Y}}{\sqrt{\sigma^2 c_k'\Sigma c_k}} \tag{1}
\]

where \( \Sigma \) is a correlation matrix of \( \bar{Y} \) and \( \sigma^2 \) is the usual estimator of \( \sigma^2 \). The maximum contrast method formed by taking the maximum over multiple contrast statistics, say \( T_{max}=max(T_1,T_2,...,T_a) \) is used to detect the existence of a dose-response. If \( T_{max}>c \) for an appropriate critical value \( c \) subject to \( Pr(T_{max}>c)=0.025 \) under the overall null hypothesis \( \mu_1=\mu_2=...=\mu_a \), a significant dose-response is concluded. Once a dose-response is shown, it is possible to pursue further research for selecting a response pattern. A contrast-based approach for selecting the most plausible response pattern from a set of candidate patterns based on the maximum of observed contrast statistics has been proposed. Bayesian inference provides a framework for evaluating the plausibility of the specified response pattern. Assume a non-informative prior for \( \mu=(\mu_1,\mu_2,...,\mu_a)' \) and \( \sigma^2 \). Let \( \tau=(\tau_1,\tau_2,...,\tau_m)' \) be contrast vector where \( \tau_k \) has the same form as the one given in (1) where \( \bar{Y} \) is replaced by \( \mu \). The posterior probability that the \( \tau_k \) takes the maximum over \( \tau_1,\tau_2,...,\tau_m \) is evaluated using a multivariate t-distribution. In this paper, only non-singular correlation matrix is assumed for simplicity. This posterior probability implies the plausibility of the specified response pattern. For binary variables, assume that the number of “success” counts have binomial distribution with probability \( \pi \) and sample size ni. When a non-informative prior is taken to be uniform, the posterior distribution of the contrast vector \( \lambda=(\lambda_1,\lambda_2,...,\lambda_m)' \) is a multivariate normal distribution with large ni where

\[ 
\lambda_k = c_k'\pi/\sqrt{c_k'\Sigma c_k} 
\]

with a variance covariance matrix \( \Sigma \). It is
possible to evaluate the posterior probability in the same fashion as that for continuous variables.

SAS PROC IML CODE
Here is SAS PROC IML code for evaluating the posterior probability for both cases of continuous and binary variables. The code below is used after the observed values of contrast statistics are obtained with the 'contrast' statement in PROC GLM for continuous variables. For binary variables, the number of subjects, the number of “success” counts, and the contrast coefficients need to be specified.

(A) FOR CONTINUOUS VARIABLES

proc iml;
start ni1(z2) global(n,df,c,t,zz1,zz2,v_c,eps);
  z=shape(.,2,1);z[1]=zz1;z[2]=z2;
p3=(1+(z-t)`*inv(v_c)*(z-t)/df)#(-df/2));
  return(p3);
finish ni1;
start ni2(z1) global(n,df,c,t,zz1,zz2,v_c,eps);
  zz1=z1;interval=-10||zz1;
call quad(p2,"NI1",interval) eps=eps;
  return(p2);
finish ni2;
start ni3(z3) global(n,df,c,t,zz1,zz2,v_c,eps);
  z=shape(.,3,1);z[1]=zz1;z[2]=zz2;z[3]=z3;
p3=(1+(z-t)`*inv(v_c)*(z-t)/df)#(-df/2));
  return(p3);
finish ni3;
start ni4(z2) global(n,df,c,t,zz1,zz2,v_c,eps);
  zz2=z2;interval=-10||zz1;
call quad(p2,"NI3",interval) eps=eps;
  return(p2);
finish ni4;
start ni5(z1) global(n,df,c,t,zz1,zz2,v_c,eps);
  zz1=z1;interval=-10||z1;
call quad(p5,"NI4",interval) eps=eps;
  return(p5);
finish ni5;
start c_poster(_n,_c,_t,_df) global(n,df,c,t,zz1,zz2,v_c,eps);
  n=_n;_c_=_c`;_t_=_t;df=_df;
  if nrow(n)=3 then v_y=inv(block(n[1],n[2],n[3]));
  if nrow(n)=4 then v_y=inv(block(n[1],n[2],n[3],n[4]));
  c_de=sqrt(diag(_c_`*_c_));c_st_=shape(.,nrow(_c_),ncol(_c_));
do i=1 to ncol(_c_);c_st_[,i]=_c_[,i]/c_de[i,i];end;
  if ncol(_c_)=2 then do;
    t=shape(.,ncol(_c_),1);c_st=shape(.,nrow(n),ncol(_c_));
    if max(_t_)=_t_[1] then do;T=t;c_st=c_st_end;end;
    else if max(_t_)=_t_[2] then do;
  end;
  else if ncol(_c_)=3 then do;
    t=shape(.,ncol(_c_),1);c_st=shape(.,nrow(n),ncol(_c_));
    if max(_t_)=_t_[1] then do;T=t;c_st=c_st_end;end;
    else if max(_t_)=_t_[2] then do;
    else if max(_t_)=_t_[3] then do;
  end;
  v_ct=nrow(v_c)*v_y*c_st;v_c=shape(.,nrow(v_ct),ncol(v_ct));
do i=1 to nrow(v_ct);do j=1 to ncol(v_ct);
    v_c[i,j]=v_c[i,j]/sqrt(v_c[i,i]*v_ct[j,j]);end;
  interval=10||20;eps=1e-6;
  if ncol(_c_)=2 then do;
    call quad(p1,"NI2",interval) eps=eps;
    p4=p1#exp(lgamma((df+2)/2)-gamma(df/2)-log(3.14159265)-log(df)
      -0.5*log(det(v_c)));end;
  if ncol(_c_)=3 then do;
    call quad(p1,"NI3",interval) eps=eps;
    p4=p1#exp(lgamma((df+3)/2)-gamma(df/2)-3/2*log(3.14159265)
      -3/2*log(df)-0.5*log(det(v_c)));end;
  return(p4);
finish c_poster;
Sample sizes, contrast coefficient vectors, observed values of contrast statistics, and degrees of freedom need to be specified as follows:

<four groups with three contrasts>
prob=c_poster({n1,n2,n3,n4},{cont1,cont2,cont3},{t1,t2,t3},df);

<four groups with two contrasts>
prob=c_poster({n1,n2,n3,n4},{cont1,cont2},{t1,t2},df);

<three groups with two contrasts>
prob=c_poster({n1,n2,n3},{cont1,cont2},{t1,t2},df);

An example of specification for the case of three groups with two contrasts is shown below.
prob=c_poster({84,82,83},{-1 0 1,-2 1 1},(6.689,6.368),246);

(B) FOR BINARY VARIABLES
proc iml;
start ni1(z1) global(cont,v_c,eps,p_st,p_st1,p_st2);
p2=pdf('NORMAL',z1,cont[1],1)#cdf('NORMAL',
    sqrt(v_c[2,2]/v_c[1,1])#(cont[2]+v_c[1,2]#(z1-cont[1]))),
sqrt(1-v_c[1,2]##2));
return(p2);
finish ni1;

start ni2(z1) global(cont,v_c,eps,p_st,p_st1,p_st2);
p_w=shape(.,2,1);p_w[1]=cont[2];p_w[2]=cont[3];
p_m=p_w+p_st1*inv(v_c[1,1])*(z1-cont[1]);
p2=pdf('NORMAL',z1,cont[1],1)#probnorm(sqrt(v_c[2,2]/v_c[1,1])#
    ((z1-p_m[1])/sqrt(p_st[1,1]),sqrt(v_c[2,2]/v_c[1,1])#
    (z1-p_m[2])/sqrt(p_st[2,2]))),
p_st[1,2]/sqrt(p_st[1,1])/sqrt(p_st[2,2]));
return(p2);
finish ni2;

start b_poster(_n,_x,_c) global(cont,v_c,eps,p_st,p_st1,p_st2);
    n=_n;_c_=_c`;_x=_x;
    if nrow(n)=3 then v_y=inv(block(n[1],n[2],n[3]));
    if nrow(n)=4 then v_y=inv(block(n[1],n[2],n[3],n[4]));
    p=shape(nrow(n),1);
    do i=1 to nrow(n);p[i]=sum(_x)/sum(n);end;
    pp=_x/n;
    _w=shape(.,ncol(_c_),1);
    do i=1 to ncol(_c_);
        if ncol(_c_)=2 then do;
            w=shape(.,ncol(_c_),1);
            c=shape(nrow(_c_),ncol(_c_));
        end;
        v_ct=shape(.,ncol(_c_),ncol(_c_));
        do i=1 to ncol(_c_);do j=1 to ncol(_c_);v_ct[i,j]=(c[i]#(v_y#(p`*(1-p)))*(c[j])#(v_y#(p`*(1-p)))*(c[j]));
            end;
        end;
        v_c=shape(.,ncol(_c_),ncol(_c_));
        do i=1 to ncol(_c_);do j=1 to ncol(_c_);v_c[i,j]=v_ct[i,j]/sqrt(v_ct[i,i]*v_ct[j,j]);end;
        end;
        interval=-10||20;eps=1e-6;
    end;
    if ncol(_c_)=2 then do;cont=shape(ncol(_c_),1);
    do i=1 to ncol(_c_);cont[i]=w[i]/sqrt(v_ct[i,i]);end;
    call quad(p1,"NI1",interval) eps=eps;
    if ncol(_c_)=3 then do;cont=shape(ncol(_c_),1);
    do i=1 to ncol(_c_);cont[i]=w[i]/sqrt(v_ct[i,i]);end;
p_st1=shape(.,2,1);p_st2[1,1]=v_c[1,2];p_st2[1,2]=v_c[1,3];
p_st2=shape(.,2,2);p_st2[1,1]=v_c[2,2];p_st2[1,2]=v_c[2,3];
Sample sizes, the number of "success" counts and contrast coefficients need to be specified as below. It is not necessary to specify the observed values of contrast statistics since they will be calculated automatically with this code.

<three groups with two contrasts>
\[ \text{prob} = \text{b_poster}\left(\{n_1,n_2,n_3\},\{x_1,x_2,x_3\},\{\text{cont}_1,\text{cont}_2\}\right) \]

<four groups with two contrasts>
\[ \text{prob} = \text{b_poster}\left(\{n_1,n_2,n_3,n_4\},\{x_1,x_2,x_3,x_4\},\{\text{cont}_1,\text{cont}_2\}\right) \]

<four groups with three contrasts>
\[ \text{prob} = \text{b_poster}\left(\{n_1,n_2,n_3,n_4\},\{x_1,x_2,x_3,x_4\},\{\text{cont}_1,\text{cont}_2,\text{cont}_3\}\right) \]

An example of specification for four groups with three contrasts is shown below.
\[ \text{prob} = \text{b_poster}\left(\{79,80,76,74\},\{40,51,51,56\},\{-3,-1,1,3,-5,-1,3,3,-3\}\right) \]

EXAMPLES

EXAMPLE 1: CONTINUOUS VARIABLE IN A PHASE II CLINICAL TRIAL
In this section, the method for a continuous variable is illustrated. A phase II clinical trial for a HMG-CoA reductase inhibitor in patients with hyperlipidemia (Saito et al., 2001) is used. This is a randomized, double-blind and parallel group trial in which approximately 250 patients were allocated to one of three active doses of low, middle and high. The primary variable is percent change from baseline in total cholesterol (TC) at the last visit during the 12 week treatment period. The summary statistics are shown in Table 1. The contrast coefficient vectors of (1,0,-1) and (2,-1,-1) are used to explore the response pattern. The observed values of contrast statistics corresponding to the contrast coefficient vectors above are \( t_1 = 6.689 \) and \( t_2 = 6.368 \), respectively. The \( p \)-value under the overall null hypothesis \( \mu_{\text{low}} = \mu_{\text{middle}} = \mu_{\text{high}} \) is less than 0.025, showing a significant dose-dependency. Then we proceed to assess the response pattern of the compound. The contrast coefficients of (1,0,-1) best fits data, implying that linear increasing is the most likely response pattern to the true one. Posterior probability implying the plausibility of the response pattern associated with (1,0,-1) is 0.731, showing that it is moderately plausible to select this pattern such that drug effect linearly increases with dose levels.
Table 1. Summary statistics in percent change from baseline for TC in patients with hyperlipidemia

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>84</td>
<td>82</td>
<td>83</td>
</tr>
<tr>
<td>Mean</td>
<td>-23.0</td>
<td>-29.1</td>
<td>-32.4</td>
</tr>
<tr>
<td>Pooled SD</td>
<td>9.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Trial results of TC in patients with hyperlipidemia

Table 2. Proportion of responders in patients with migraine

<table>
<thead>
<tr>
<th>Placebo</th>
<th>Low</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responder/N</td>
<td>40/79</td>
<td>51/80</td>
<td>51/76</td>
</tr>
<tr>
<td>Proportion(%)</td>
<td>50.6</td>
<td>63.8</td>
<td>67.1</td>
</tr>
</tbody>
</table>

Figure 2. Trial results of proportions of responders in patients with migraine
EXAMPLE 2: BINARY VARIABLE IN A PHASE II CLINICAL TRIAL

In this section, the method for a binary variable is illustrated. A phase II clinical trial for a 5-HT agonist in patients with migraine (Eletriptan Steering Committee in Japan, 2002) is used. This is a randomized, double-blind and parallel group trial in which approximately 310 patients were allocated to either placebo or one of three active doses of low, middle and high. The primary variable is proportion of responders at 2 hour postdose. The observed frequencies are shown in Table 2.

The contrast coefficient vectors of (-3,-1,1,3), (-5,-1,3,3) and (-3,1,1,1) are used to explore the response pattern. The observed values of contrast statistics corresponding to the contrast coefficient vectors above are $z_1=3.201$, $z_2=3.079$ and $z_3=2.910$, respectively. The $p$-value under the overall null hypothesis $\pi_{\text{placebo}}=\pi_{\text{low}}=\pi_{\text{middle}}=\pi_{\text{high}}$ is less than 0.025, showing a significant dose-dependency. Then we proceed to assess the response pattern of the compound. The contrast coefficients of (-3,-1,1,3) best fits data, implying that linear increasing is the most likely response pattern to the true one. Posterior probability implying the plausibility of the response pattern associated with (-3,-1,1,3) is 0.859, showing that it is plausible to select this pattern such that drug effect linearly increases with dose levels.

CONCLUSIONS

SAS code for evaluating posterior probability when selecting a dose-response pattern from a set of candidate response patterns using contrast statistics in analyzing a phase II clinical trial consisting of three or four doses was introduced. Posterior probabilities can be evaluated by numerical integration based on a multivariate normal or multivariate t-distribution. The code was developed for both continuous variables and binary variables. The method was illustrated on two data sets.

REFERENCES


CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Akira Wakana
Data Management & Biostatistics
Banyu Pharmaceutical Co., Ltd.
5-1 Nihombashi-Kabutocho, Chu-o-ku, Tokyo 103-0026, Japan
E-mail: akira_wakana@merck.com