

Bradley C. Martin and Jeffrey H. Dorfman

The presenters will briefly describe the uses and applications of the bootstrap technique for estimating confidence intervals and mean variances. The authors will also present the SAS® code used to estimate bootstrapped CIs and mean variances accompanied with a step by step descriptions of the essential SAS steps. Several bootstrap methods will be presented: the single variable non-parametric bootstrap, the single variable bias corrected bootstrap, the nonparametric and bias corrected bootstrap methods for estimating CIs and mean variances for two variable ratios.

```

libname res 'c:\research\data\';
options ps=94 ls=120;
data bootstrp;
    set res.hosp_ept;

*-----*
output true means of population for bias corrected bootstrap
*-----*

proc means noprint; var nc_yrs nclos elig_yrs;
    output out=trumean mean=mncyr mnclos melig;

*-----*
set number of iterations equal to n_iterat
change ci limits when iteration change
*-----*
run;
proc iml;
    use bootstrp;
    read all var (nc_yrs nclos elig_yrs) into util;
    n_iterat=1000; /* N_ITERAT = # OF ITERATIONS */
    n=nrow(util);
    y=j(n,3,0); /* CREATE Y MATRIX */
    do k=1 to n_iterat; /* BEGIN BOOTSTRAP OF N ITERATIONS */
        do i=1 to n;
            pick=floor(n#uniform(0))+1; /* RANDOMLY SELECT N OBS */
            y[i,]=util[pick,]; /* WITH REPLACEMENT */
        end;
        ysum=y[+,]; /* SUM OF VARIABLES */
        ymean=ysum/n; /* MEAN OF VARIABLES */
        z=z // ymean; /* BUILD MATRIX OF BOOTSTRAPPED MEANS */
    end;
    zz = z[+,]/n_iterat; /* MEAN OF BOOTSTRAPPED MEANS */
    print zz;
    create bsout from z; /* CREATE SAS DATA SET */
    append from z;
    create iterate from n_iterat [colname='n-iterat']; /* SAS DATA SET */
    append from n_iterat;
    quit;

*-----*
ITERATE IS SAS DATA SET CONTAINING NUMBER OF ITERATIONS
*-----*
*-----*
Z MATIX CONTAINS BOOTSTRAPPED MEANS
BSOUT IS SAS DATA SET CONTAINING N_ITERATED BOOTSTRAPPED MEANS
*-----*
*-----*
!!!!!!END OF BOOTSTRAP!!!!!!
*-----*

```

```

*-----*
SORT FIRST VARIABLE:    NC_YRS
*-----*
data V1A;
    set bsout;
        label col1="B.S. MEANS FOR NC_YRS";
        label col2="B.S. MEANS FOR NCLOS";
        label col3="B.S. MEANS FOR ELIG_YRS";
    proc means mean var std;
    proc sort; by col1;
*-----*
CALCULATE LOWER (C1) AND UPPER (C2) 95% CONFIDENCE LIMITS
*-----*
proc iml;
    use V1A;
    read all var {col1} into bsort;
    use trumean;
    read all var {mncyr} into mncyr;
    c1=bsort[26,1];
    c2=bsort[975,1];
    print c1 "LOWER 95% BOOTSRAP CI FOR NC_YRS";
    print c2 "UPPER 95% BOOTSRAP CI FOR NC_YRS";
    n_iterat=nrow(bsort);
    do i= 1 to n_iterat; /* DETERMINE POSITION OF TRUE MEAN AMONG */
        if mncyr>=bsort[i,1] then p0=i/n_iterat; /* BOOTSTAP MEANS */
    end; /* P0 */
    print p0;
    create pout from p0 [colname='p0']; /* CREATE SAS DATA SET */
    append from p0; /* TO OUTPUT P0 */
quit;

```

```

*-----*
COMPUTE CDF (Z VALUE) OF 95% CI
TO ARRIVE AT BIAS CORRECTED CI
*-----*
data biasc1;
    merge pout iterate;
    z0=probit(p0);
    zmin=probit(0.025);
    zmax=probit(0.975);
    z1=2*z0+zmin;
    z2=2*z0+zmax;
    q1=round(probnorm(z1)*n_iterat);
    q2=round(probnorm(z2)*n_iterat);
    proc print;
proc iml;
    use biasc1;
    read all var {q1 q2} into q;
    use V1A;
    read all var {col1} into bsort;
        q1=q[1,1];
        q2=q[1,2];
    bc1=bsort[q1,1];
    bc2=bsort[q2,1];
    print bc1 "BIAS CORRECTED LOWER 95% CI FOR NC_YRS";
    print bc2 "BIAS CORRECTED UPPER 95% CI FOR NC_YRS";

```

```

*-----*
CIS FOR 2 VARIABLE CASE
*-----*
*-----*
SUM NCLOS / SUM ELIG YEARS
*-----*
data v2a;
    set bsout;
    hratio=col2/col3;
    label hratio="B.S. MEANS FOR SUMNCLOS/SUMELIG";
    proc means mean var std;
    proc sort; by hratio;
*-----*
CALCULATE LOWER (C1) AND UPPER (C2) 95% CONFIDENCE LIMITS
*-----*
proc iml;
    use v2a;
    read all var {hratio} into bsort;
    use trumean;
    read all var {mnclos melig} into mu;
    c1=bsort[26,1];
    c2=bsort[975,1];
    print c1 "LOWER 95% BOOTSRAP CI";
    print c2 "UPPER 95% BOOTSRAP CI";
    n_iterat=nrow(bsort);
    do i= 1 to n_iterat; /* DETERMINE MEDIAN OF */
        if mu[1,1]/mu[1,2]>=bsort[i,1] then p0=i/n_iterat; /* BOOTSTAP MEANS */
    end; /* P0 */
print p0;
create pout2 from p0 [colname='p0']; /* CREATE SAS DATA SET */
append from p0; /* TO OUTPUT P0 */
quit;

*-----*
COMPUTE CDF (Z VALUE) OF 95% CI
TO ARRIVE AT BIAS CORRECTED CI
*-----*
data biasc2v;
    merge pout2 iterate;
    z0=probit(p0);
    zmin=probit(0.025);
    zmax=probit(0.975);
    z1=2*z0+zmin;
    z2=2*z0+zmax;
    q1=round(probnorm(z1)*n_iterat);
    q2=round(probnorm(z2)*n_iterat);
    proc print;
proc iml;
    use biasc2v;
    read all var {q1 q2} into q;
    use v2a;
    read all var {hratio} into bsort;
    q1=q[1,1];
    q2=q[1,2];
    bc1=bsort[q1,1];
    bc2=bsort[q2,1];
    print bc1 "BIAS CORRECTED LOWER 95% CI";
    print bc2 "BIAS CORRECTED UPPER 95% CI";
run;

```

SAS and SAS/IML are registered trademarks or trademarks of SAS Institute, Inc in the USA and other countries.