```
%macro third_stage_sampling(N=,N1=,N2=,

                            f1=,f2=,f3h1=,f3h2=,

                            c=,c0h=,c1h=,c2h=,c3h=,reps=);


%let n1_y=ceil(%sysevalf(&N1*&f1));

%let n2_y=ceil(%sysevalf(&N2*&f2));

%do lyy=1 %to &reps;

data xc_population;

    set xc_population;

run;

proc sql noprint;

    create table district_list as

    select distinct district from xc_population;

quit;

proc sql noprint;

    create table district_site_list as

    select distinct district,site from xc_population;

quit;

proc surveyselect data=district_list noprint

    method=srs sampsize=&f1 out=first_stage_sampling seed=&lyy;

run;

data after_first_stage_sampling;

    merge district_site_list first_stage_sampling (in = yy1);

    by district;

    if yy1 = 1;

run;

proc surveyselect data=after_first_stage_sampling noprint

    method=srs out=second_stage_sampling (drop=SelectionProb

    SamplingWeight) sampsize=&f2 seed=&lyy;

    strata district;
```

```
run;

proc sort data=xc_population;by district site;run;

proc sort data=second_stage_sampling;by district site;run;

data after_second_stage_sampling;

    merge xc_population second_stage_sampling (in = yy2);

    by district site;

    if yy2 = 1;

run;

data after_second_stage_sampling_h1 after_second_stage_sampling_h2;

    set after_second_stage_sampling;

    if h=1 then output after_second_stage_sampling_h1;

    if h=2 then output after_second_stage_sampling_h2;

run;

proc surveyselect data=after_second_stage_sampling_h1 noprint

    method=srs out=third_stage_sampling_h1 (drop=SelectionProb

    SamplingWeight) samprate=&f3h1 seed=&lyy;

    strata district site;

run;

proc surveyselect data=after_second_stage_sampling_h2 noprint

    method=srs out=third_stage_sampling_h2 (drop=SelectionProb

    SamplingWeight) samprate=&f3h2 seed=&lyy;

    strata district site;

run;

data third_stage_sampling;

    set third_stage_sampling_h1 third_stage_sampling_h2;

run;

proc sql noprint;

    create table fcs_chosen_number1 as

    select count(h) as fcs from third_stage_sampling

    group by h;
```

```
quit;

proc transpose data=fcs_chosen_number1

                out=fcs_chosen_number

                (rename=(col1=fcs_h1 col2=fcs_h2) drop=_name_);

var fcs;

run;

proc sql noprint;

    create table fcs_total_number1 as

    select count(h) as total from xc_population

    group by h;

quit;

proc transpose data=fcs_total_number1

                out=fcs_total_number

                (rename=(col1=total_h1 col2=total_h2) drop=_name_);

var total;

run;

data w_h(keep=w1 w2);set fcs_total_number;

    w1=total_h1/sum(total_h1,total_h2);

    w2=total_h2/sum(total_h1,total_h2);

run;

data xc;

    set third_stage_sampling;

    seed=&lyy;

    rrt=rantbl(seed,0.6,0.4);

    if rrt=1 then rrt_ball='Red';else rrt_ball='White';

    if rrt_ball='Red' then do;

    if d1_real=1 then d1=1;

    if d1_real=2 then d1=2;

end;

else do;
```

```sas
        if birthdate=1 then d1=1;

        if birthdate=2 then d1=2;

end;

        keep district site h d1;

run;

proc sort data=xc;by district site h;run;

proc freq data=xc noprint;

        by district site h;

        table d1/out=d01a;

run;

proc transpose data=d01a out=d01b(drop=_name_);

        by district site h;

        id d1;

run;

data d01 (rename=(_1=count1 _2=count2)

                drop=_label_);

        if _n_=1 then set fcs_chosen_number;

        if _n_=1 then set fcs_total_number;

        set d01b;

        if _1=. then _1=0;

        if _2=. then _2=0;

        if _label_='总频数百分比' then delete;

        chosen_fcs=sum(_1,_2);

        percent1=_1/chosen_fcs;

run;

data d01_1_share;

        set d01 (keep=district site h count1

                chosen_fcs percent1 total_h1 total_h2 fcs_h1 fcs_h2);

        pij1=(percent1-(1-r)*rij)/r;

        sampling_ratio_h1=(fcs_h1/&n1_y/&n2_y)/(total_h1/&N1/&N2);
```

```
        sampling_ratio_h2=(fcs_h2/&n1_y/&n2_y)/(total_h2/&N1/&N2);

        if h=1 then nij3=chosen_fcs/sampling_ratio_h1;

        if h=2 then nij3=chosen_fcs/sampling_ratio_h2;

        nij3_pij1=nij3*pij1;

        if pij1<0 then pij1=0; if nij3_pij1<0 then nij3_pij1=0;

        if pij1>1 then pij1=1; if pij1=1 then nij3_pij1=nij3;

run;

proc means data=d01_1_share sum noprint;

        var nij3_pij1;

        class h district;

        output out=d01_1_p sum=sum_nij3_pij1;

run;

data d01_1_p;set d01_1_p(keep=h district sum_nij3_pij1);

        if h=. or district=. then delete;

        ni2_ni2_sum_nij3_pij1=&N2/&n2_y*sum_nij3_pij1;

run;

proc means data=d01_1_p sum noprint;

        class h;

        var ni2_ni2_sum_nij3_pij1;

        output out=d01_1_p sum=sum_ni2_ni2_sum_nij3_pij1;

run;

data d01_1_p;set d01_1_p(drop=_type_ _freq_);

        if _n_=1 then set fcs_total_number;

        if _n_=1 then set w_h;

        if h=. then delete;

        if h=1 then do;

            p_h_1=&N1/&n1_y/total_h1*sum_ni2_ni2_sum_nij3_pij1;

            w_p_h_1=w1*p_h_1;

        end;

        if h=2 then do;
```

```sas
            p_h_1=&N1/&n1_y/total_h2*sum_ni2_ni2_sum_nij3_pij1;

            w_p_h_1=w2*p_h_1;

        end;

run;

proc sql noprint;

    create table d01_p1 as

    select sum(w_p_h_1) as p1 from d01_1_p;

quit;

data d01_1_v3;set d01_1_share;

    stdijk1=pij1*(1-pij1);

run;

proc means data=d01_1_v3 sum noprint;

    var stdijk1;class h district;

    output out=d01_1_v3 sum=sum_stdijk1;

run;

data d01_1_v3;set d01_1_v3;

    if _type_<3 then delete;

    reciprocal_ni2_sum_stdijk1=1/&n2_y*sum_stdijk1;

    drop _type_ _freq_;

run;

proc means data=d01_1_v3 sum noprint;

    var reciprocal_ni2_sum_stdijk1;

    class h;

    output out=d01_1_v3 sum=sum_reciprocal_ni2_sum_stdijk1;

run;

data d01_1_v3;

    set d01_1_v3 (drop=_type_ _freq_);

    v3h=1/&n1_y*sum_reciprocal_ni2_sum_stdijk1;

run;

data d01_1_v2_pi;
```

```
set d01_1_share;run;
proc means data=d01_1_v2_pi sum noprint;
    var nij3 nij3_pij1;
    class h district;
    output out=d01_1_v2_pi sum=sum_nij3 sum_nij3_pij1;
run;
data d01_1_v2_pi(drop=_type_);
    set d01_1_v2_pi (drop=_freq_);
    if _type_<3 then delete;
    pih_1=sum_nij3_pij1/sum_nij3;
run;
proc sort data=d01_1_share;by h district;run;
data d01_1_v2;
    merge d01_1_share d01_1_v2_pi;
    by h district;
    pij_pi_square1=(pij1-pih_1)**2;
run;
proc means data=d01_1_v2 sum noprint;
    var pij_pi_square1;
    class h district;
    output out=d01_1_v2 sum=sum_pij_pi_square1;
run;
data d01_1_v2(drop=_type_);
set d01_1_v2(drop=_freq_);
    if _type_<3 then delete;
    stdij1=1/(&n2_y-1)*sum_pij_pi_square1;
run;
proc means data=d01_1_v2 sum noprint;
    var stdij1;
    class h;
```

```
        output out=d01_1_v2 sum=sum_stdij1;
run;
data d01_1_v2;set d01_1_v2 (drop=_type_ _freq_);
    if h=. then delete;
    v2h=(1/&n1_y)*sum_stdij1;
run;
data d01_1_v1;
    merge d01_1_p(keep=h p_h_1) d01_1_v2_pi(keep=h district pih_1);
    by h;
    p_pi_square1=(p_h_1-pih_1)**2;
run;
proc means data=d01_1_v1 sum noprint;
    var p_pi_square1;class h;
    output out=d01_1_v1 sum=sum_p_pi_square1;
run;
data d01_1_v1;set d01_1_v1(drop=_type_ _freq_);
    if h=. then delete;
    stdi1=1/(&n1_y-1)*sum_p_pi_square1;
    v1h=stdi1;
run;
data d01_1_v_p;answer_code='D01';
    if _n_=1 then set fcs_total_number;
    if _n_=1 then set fcs_chosen_number;
    if _n_=1 then set w_h;
    if _n_=1 then set d01_p1;
    merge d01_1_v1(keep=h v1h) d01_1_v2(keep=h v2h) d01_1_v3(keep=h v3h);
    by h;
    if h=1 then
        w_h_v_p1=w1*w1*(v1h/&n1_y*(1-&n1_y/&N1)+v2h/&n1_y/&n2_y*(1-&n2_y/&N2)+
                v3h/fcs_h1*((fcs_h1/&n1_y/&n2_y)/(total_h1/&N1/&N2)));
```

```
    if h=2 then

       w_h_v_p1=w2*w2*(v1h/&n1_y*(1-&n1_y/&N1)+v2h/&n1_y/&n2_y*(1-&n2_y/&N2)+

              v3h/fcs_h2*((fcs_h2/&n1_y/&n2_y)/(total_h2/&N1/&N2)));

run;

proc sql noprint;

    create table d01_1_v_p1 as

    select p1,sum(w_h_v_p1) as v_p1,

           p1-1.96*sqrt(calculated v_p1) as lower_limit,

           p1+1.96*sqrt(calculated v_p1) as upper_limit

    from d01_1_v_p;

quit;

data d01_1_ss;

    set d01_1_v_p;set d01_1_v_p1;

    var=v_p1;

    label v1h='σ 1h' v2h='σ 2h' v3h='σ 3h' p1='样本比例' upper_limit='95%上限'

          lower_limit='95%下限' answer_code='敏感问题编号' var='var(p)';

run;

data d01_ss1;

    set d01_1_ss(keep=answer_code total_h1 total_h2 fcs_h1 fcs_h2

                      w1 w2 h v1h v2h v3h var);

    label var='var(p_k)';

run;

data d01_ss2;set d01_ss1;

    n2h_bar=&N2;n1h=&N1;

    if h=1 then n3h_bar=total_h1/&N1/&N2;

         else n3h_bar=total_h2/&N1/&N2;

    nn2h_bar=((v2h-v3h/n3h_bar)/(v1h-v2h/n2h_bar)*(&c1h/&c2h))**0.5;

    nn3h_bar=(v3h/(v2h-v3h/n3h_bar)*(&c2h/&c3h))**0.5;

    if h=1 then

lamda_var_n=w1*sqrt(&c3h/v3h)*(nn2h_bar*nn3h_bar*(v1h-v2h/n2h_bar)+nn3h_bar*(v2h-v3h/
```

```
n3h_bar)+sqrt(v3h));
        else
lamda_var_n=w2*sqrt(&c3h/v3h)*(nn2h_bar*nn3h_bar*(v1h-v2h/n2h_bar)+nn3h_bar*(v2h-v3h/
n3h_bar)+sqrt(v3h));
    if h=1 then lamda_var_d=var/2+(w1**2)*(v1h/n1h);
        else lamda_var_d=var/2+(w2**2)*(v1h/n1h);
    if h=1 then lamda_cost_n=&c/2-&c0h;
        else lamda_cost_n=&c/2-&c0h;
    if h=1 then
lamda_cost_d=w1*(sqrt(&c1h)*sqrt(v1h-v2h/n2h_bar)+sqrt(&c2h)*sqrt(v2h-v3h/n3h_bar)+sqrt(
&c3h*v3h));
        else
lamda_cost_d=w2*(sqrt(&c1h)*sqrt(v1h-v2h/n2h_bar)+sqrt(&c2h)*sqrt(v2h-v3h/n3h_bar)+sqrt(
&c3h*v3h));
run;
proc sql noprint;
    create table d01_lamda as
    select answer_code,sum(lamda_var_n) as L_var_n, sum(lamda_var_d) as L_var_d,
            sum(lamda_cost_n) as L_cost_n, sum(lamda_cost_d) as L_cost_d,
            (calculated L_var_n)/(calculated L_var_d) as L_var,
            (calculated L_cost_n)/(calculated L_cost_d) as L_cost
    from d01_ss2 group by answer_code;
quit;
data d01_ss_&lyy;
    merge d01_ss2(drop=total_h1 total_h2 fcs_h1 fcs_h2
                    lamda_var_n lamda_var_d lamda_cost_n lamda_cost_d)
        d01_lamda (keep=answer_code L_var L_cost);
    by answer_code;
    if h=1 then nn1h_var=w1*sqrt(v3h)*L_var/nn2h_bar/nn3h_bar/sqrt(&c3h);
        else nn1h_var=w2*sqrt(v3h)*L_var/nn2h_bar/nn3h_bar/sqrt(&c3h);
```

```sas
    if h=1 then nn1h_cost=w1*sqrt(v3h)*L_cost/nn2h_bar/nn3h_bar/sqrt(&c3h);

        else nn1h_cost=w2*sqrt(v3h)*L_cost/nn2h_bar/nn3h_bar/sqrt(&c3h);

    code=&lyy;

run;

proc print data=d01_ss_&lyy;title "第 &lyy 次 D01 样本量";run;

%end;

%mend third_stage_sampling;

%third_stage_sampling    (N=60000,N1=10, N2=12,

                         f1=3,f2=4,f3h1=0.25,f3h2=0.25,

                         c=5*10**5,c0h=50000,c1h=30000,c2h=3000,c3h=30,reps=100);

run;

data d01_yn_ss;

    set d01_ss_1-d01_ss_100;

run;

data d1_ss_h1(rename=(h=h_h1    nn2h_bar=nn2h_bar_h1    nn3h_bar=nn3h_bar_h1

                      nn1_var=nn1_var_h1 nn1_cost=nn1_cost_h1 n3h_bar=n3h_bar_h1))

    d1_ss_h2(rename=(h=h_h2    nn2h_bar=nn2h_bar_h2    nn3h_bar=nn3h_bar_h2

                     nn1_var=nn1_var_h2 nn1_cost=nn1_cost_h2 n3h_bar=n3h_bar_h2)

                     drop=answer_code);

set d1_yn_ss;

    if h=1 then output d1_ss_h1;

    if h=2 then output d1_ss_h2;

run;

data d01_samplesize;

    merge d1_ss_h1 d1_ss_h2;

    by serial_number code;

    nn1_h1=max(nn1_var_h1,nn1_cost_h1);nn1_h2=max(nn1_var_h2,nn1_cost_h2);

    f1h1=ceil(nn1_h1);f1h2=ceil(nn1_h2);

    f2h1=ceil(nn2h_bar_h1);f2h2=ceil(nn2h_bar_h2);

    f3h1=ceil(nn3h_bar_h1)/n3h_bar_h1;f3h2=ceil(nn3h_bar_h2)/n3h_bar_h2;
```

run;