



USER MANUAL VERSION 1.0

JOHN R. GARBE, LI MA, AND YANG DA

Department of Animal Science

University of Minnesota

Contact: yda@umn.edu

January 24, 2007

1	Introduction	3
2	Statistical power and sample size for QTL detection	5
2.1	Power calculation	3
2.2	Sample size calculation	4
2.3	Application to genome scan	5
2.4	Control of parameter values	5
3	Statistical power and sample size for candidate gene detection	6
3.1	Power calculation	6
3.2	Sample size calculation	6
3.3	Application to genome-wide association studies.....	6
3.4	Control of parameter values	6
4	Statistical background of power and sample size	8
Figure 1.	Power calculation for QTL detection	4
Figure 2.	Sample size calculation for QTL detection	5
Figure 3.	Power calculation for candidate gene detection	7
Figure 4.	Sample size calculation for candidate gene detection	8
5	References	8

1. INTRODUCTION

The SPSSQTL program is designed to calculate the statistical power and sample size for detecting bi-allelic QTL and candidate gene effects. The statistical formulae implemented by SPSSQTL are those for detecting additive and dominance QTL effects under the F-2 and reciprocal backcross (RBC) designs (London and Da, 2006) and those for detecting epistasis effects under the F-2 design (Mao and London, 2005). For detecting candidate gene effects, the statistical formulae are the same as those for QTL detection except that the marker-QTL recombination frequency is set to zero. With the SPSSQTL program, statistical power or sample size can be calculated for any given set of parameter values. For QTL detection, statistical power is a function of experiment design, sample size, type-I error, marker-QTL recombination frequency, and QTL effect size or heritability, and sample size is a function of type-II error (or statistical power), type-I error, marker-QTL recombination frequency, and QTL effect size or heritability. For candidate gene detection, statistical power and sample size are functions of the same sets of parameters except marker-QTL recombination frequency and experiment design.

2. STATISTICAL POWER AND SAMPLE SIZE FOR QTL DETECTION

2.1 Power Calculation

The power calculation for QTL detection requires the following input items in the SPSSQTL dialog box (Figure 1):

- 1) Select an experimental design, F-2 or RBC (only F-2 is available if “Calculate epistasis effects” is activated)
- 2) Select “Calculate power”
- 3) Enter sample size
- 4) Enter type-I error
- 5) Check whether to calculate for detecting epistasis effects. Power or sample size will not be calculated for detecting epistasis effects if the “Calculate epistasis effect” box is left unchecked.
- 6) Enter marker-QTL recombination frequency
- 7) Enter heritability values.

Figure 1. Power calculation for QTL detection.

2.2 Sample Size Calculation

The sample size calculation involves the same items in the dialog box (Figure 2) except items 2) and 3), i.e.,

- 2) Select “Calculate sample size”
- 3) Enter the value of the target statistical power in the ‘Power’ field.

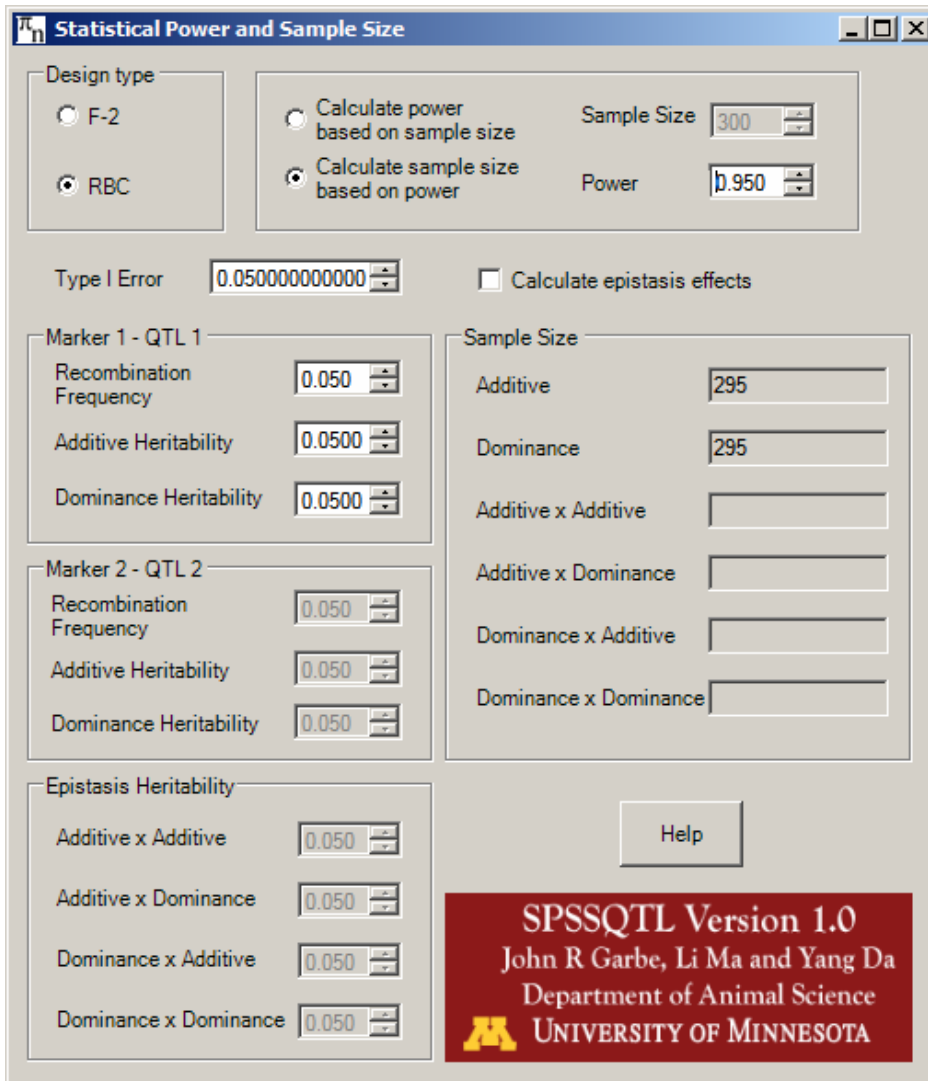


Figure 2. Sample size calculation for QTL detection

2.3 Application to genome scan

The power and sample size calculations were derived for a single statistical test of one effect. To apply these results to a genome scan using a number of markers, the type-I error value should be adjusted using a multiple testing correction.

2.4 Control of parameter values

The sum of all heritabilities should be less than or equal to 1. Each recombination frequency must be less than or equal to 0.5. The values of '0' or '1' should not be used for type-I error or statistical power.

3 STATISTICAL POWER AND SAMPLE SIZE FOR CANDIDATE GENE DETECTION

3.1 Power Calculation

The power calculation for candidate gene detection is the same as that for QTL detection with two exceptions: the marker-QTL recombination frequency must be set zero and the selection of experimental design is unnecessary. This calculation requires the following input items in the SPSSQTL dialog box (Figure 3):

- 1) Enter “0” for both recombination frequencies
- 2) Select “Calculate power”
- 3) Enter sample size
- 4) Enter type-I error
- 5) Check whether to calculate power for detecting epistasis effects
- 6) Enter heritability values

3.2 Sample Size Calculation

The sample size calculation involves the same items in the dialog box (Figure 4) except the following two items,

- 2) Select “Calculate sample size”
- 3) Enter the value of the target statistical power in the ‘Power’ field.

3.3 Application to genome-wide association study

The power and sample size calculations were derived for a single statistical test of one effect. To apply these results to a genome-wide association study using a number of markers, the type-I error value should be adjusted using a multiple testing correction.

3.4 Control of parameter values

Parameter limits are the same as for QTL detection, i.e., the sum of all heritabilities should be less than or equal to 1, each recombination frequency must be less than or equal to 0.5, and the values of ‘0’ or ‘1’ should not be used for type-I error or statistical power.

Statistical Power and Sample Size

Design type
 F-2
 RBC

Calculate power based on sample size
 Calculate sample size based on power

Sample Size: 300
Power: 0.950

Type I Error: 0.0500000000000
 Calculate epistasis effects

Marker 1 - QTL 1
Recombination Frequency: 0.000
Additive Heritability: 0.0500
Dominance Heritability: 0.0500

Marker 2 - QTL 2
Recombination Frequency: 0.000
Additive Heritability: 0.050
Dominance Heritability: 0.050

Epistasis Heritability
Additive x Additive: 0.050
Additive x Dominance: 0.050
Dominance x Additive: 0.050
Dominance x Dominance: 0.050

Power
Additive: 0.998817181594
Dominance: 0.998817181594
Additive x Additive: 0.998817181594
Additive x Dominance: 0.998817181594
Dominance x Additive: 0.998817181594
Dominance x Dominance: 0.998817181594

Help

SPSSQTL Version 1.0
John R Garbe, Li Ma and Yang Da
Department of Animal Science
UNIVERSITY OF MINNESOTA

Figure 3. Power calculation for candidate gene detection.

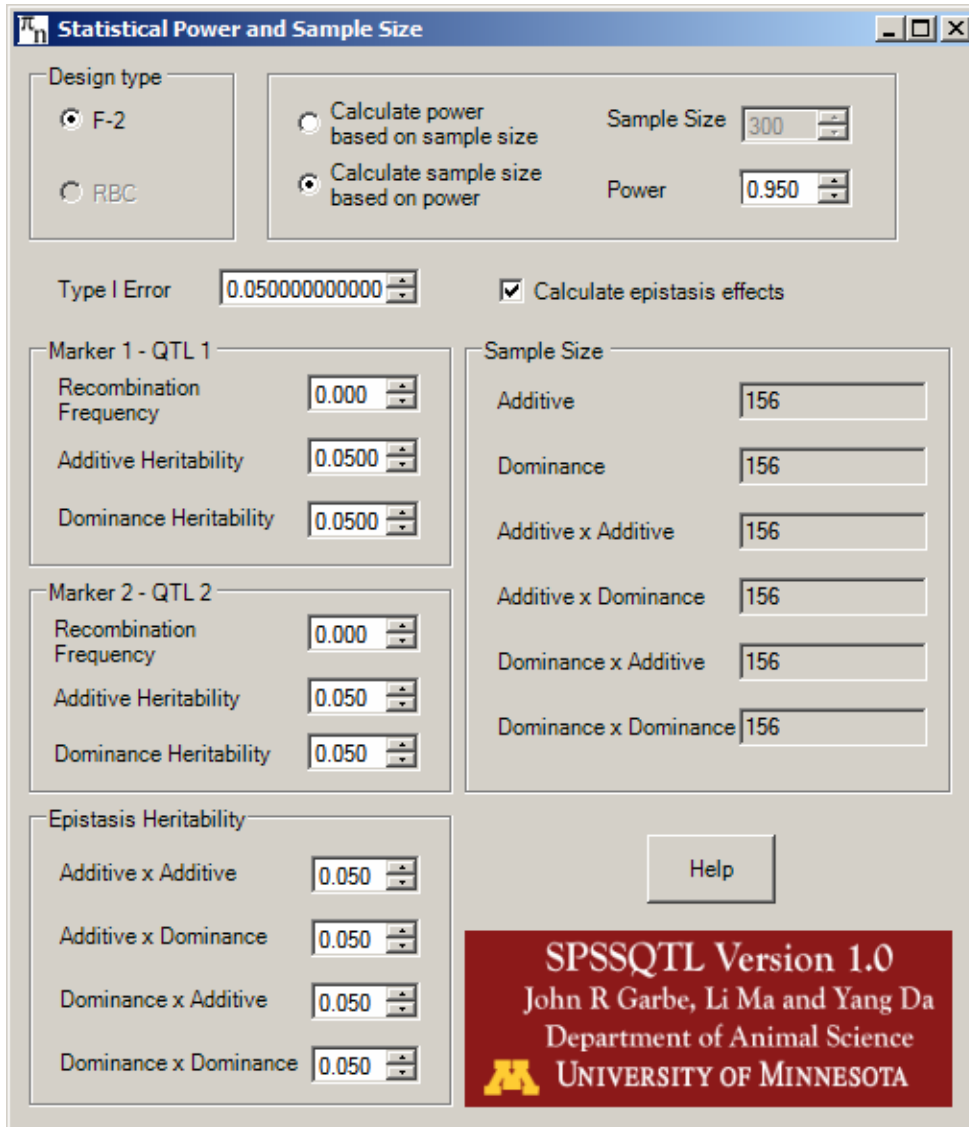


Figure 4. Sample size calculation for candidate gene detection.

4. STATISTICAL BACKGROUND OF POWER AND SAMPLE SIZE

4.1 Formulae for statistical power and sample size

Statistical power (π) is the probability that an effect is detected when the effect is present, commonly denoted by $\pi = 1 - \beta$, where β is the type II error, i.e., the probability of false ‘negatives’. A standardized normal distribution denoted by $N(0,1)$ is assumed for deriving the!Statistical power (π) is the probability that an effect is detected when the effect is present, commonly denoted by $\pi = 1$ statistical power of each design. The general expression for π is:

$$\pi = 1 - \beta = \text{Prob}(Z \geq z_i) \quad (1)$$

where Z is a $N(0,1)$ normal variable, and z_i = the ordinate of the standardized normal curve corresponding to the type-II error of β , and z_i can be expressed in terms of QTL parameters as

$$z_i = z_{\alpha/2} - \lambda_i(N_i)^{1/2} \quad (2)$$

where λ_i is a function of effect size, recombination frequencies, and experimental design for QTL detection. The formula for calculating sample size is:

$$N_i = (Z_{\alpha/2} + Z_{\beta})^2 / \lambda_i^2 \quad (3)$$

For QTL mapping, the statistical power defined by Equations (1-2) is a function of experiment design, sample size, effect size, type-I error, and marker-QTL recombination frequency or frequencies. The sample size defined by Equation (3) is a function of the same parameters except that sample size is replaced by type-II error (or statistical power). For candidate gene testing, the same parameters are involved except recombination frequencies and experimental design.

Interpretation of the estimated power and sample size

The estimated power and sample size in all case by SPSSQTL should be considered approximate and minimal rather than exact, because many factors affecting power and sample size such as unequal allele frequencies and correlated individuals are not considered by SPSSQTL.

4 REFERENCES

London N and Da Y. Statistical power and sample size requirement for QTL detection. In: *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production*, edited by Valente BD, Rossi de Moraes O, and Ventura RV. Belo Horizonte, Brazil: Instituto Prociência. 2006, vol. 22: article#18.

Mao Y and Da Y. Statistical power for detecting epistasis QTL effects under the F-2 design. *Genetics, Selection, Evolution*. 37:129-50, 2005.