

To test software based on type I error, Power, and time&memory.

1. Simulate phenotype data.

1.1 Table of parameters

Data Type	Quantitative	Binary		
Number of causals	1K	10K		
h2	0(Type I error, in chr 13-22)	0.1	0.5	0.9
Model	GCTA	LDAC		
Number of Individuals	10K	70K		
Replicate Phenotypes	5 phenotypes			

1.2 To make SNP list in chr 1 to 12, for --extract

In a txt file, with one column of rsXXXXXX

1.3 One example of simulation(In the batch queue now)

In this case, it will generate 5 phenotypes based on:

70K individuals, GCTA model, h2 = 0.1, 5 phenotypes, 1000 causal SNPs in chr 1 to 12(13 to 22 are non-genetic).

```
${dir}/software/ldak5.XXX \ --make-phenos  
${dir}/type_1_error/Multi_Traits/Trait_qt_7Wan_GCTA_h01_K_1  
\ --bfile ${dir}/data_qc \ --ignore-weights YES \ --power -1 \  
\ --her 0.1 \ --num-phenos 5 \ --num-causals 1000 \ --extract  
${dir}/type_1_error/Multi_Traits/list_snps_1_to_12.txt
```

1.4 To simulate $2*2*3*2*2 = 48$ phenotype files in total

2. Run these on different software

Bolt-Imm, Bolt-Imm-inf, regenie, fastGWA(only for 10K individuals), plink, ldak

3. To compute Power based on results in chr 1 to 12;

4. To compute Type I error based on results in chr 13 to 22.

Note: when computing Power and Type_I_Error, I'll use the .effects file as the real answer.