Heritability Analysis of Risk-Preference in Decision-making

*Dustin Pluta**

Abstract

We here present a Bayesian implementation of the popular Genome-wide Complex Trait Analysis (GCTA) model to quantify the influence of genetics (in the form of single nucleotide polymorphisms) on risky decision-making as measured by the cups task psychological experiment data adminstered to 209 healthy Chinese college students. The sensitivity analysis suggests a possibly significant heritability of risk preference. Due to the small sample size, the inference is largely affected by choice of prior, but the results do suggest that a follow-up analysis with a larger sample is merited.

Contents

1 Introduction **2**

2 Data **2**

3 Methods **3**
3.1 Statistical Definition of Heritability 3
3.2 Bayesian Inference for Heritability 4

4 Results **6**
4.1 Estimates of Heritability of Risk Count Differences 6
4.2 Model Diagnostics 7

5 Discussion **8**

6 Additional Figures **10**

References **16**

*All code, summary data, and supporting documentation is available at https://github.com/dspluta/STAT225. Contact email dpluta@uci.edu.

1 Introduction

In order to develop a more complete understanding of influences on cognition, current neuroscience studies are increasingly concerned with the analysis of multi-modal data sets, which include neuroimaging data, behavioral outcomes from pyschological tasks, and genetic information. With the advent of these complex data sets, there is a growing need for more refined and flexible statistical modeling frameworks that can accommodate data from a variety of sources and leverage the structure of the data as fully as possible. These studies are expensive to run and so often have limited sample sizes in the range of a few hundred to a few thousand subjects. This constrains the power and generality of the models that can be employed, thus it is of great benefit to be able to effectively and meaningfully incorporate prior results.

The most common approaches to heritability analyses rely on some form of the frequentist linear mixed effects (LME) model [Yang et al. (2011); ge2015massively]. The LME framework is intuitively appealing for modeling the aggregate effect of numerous genetic variables, often on the order of 100K distinct values per subject, but is limited in its flexibility, ease of use, and ability to incorporate prior beliefs. The LME model may also suffer from theoretical complications in deriving consistent tests and estimates, and from computational difficulties since the quantities of interest lie on the boundary of the parameter space under the null hypothesis of no effect.

Bayesian approaches have received much greater attention in recent years, and offer many advantages over frequentist methods. In particular, Bayesian heritability models provide a consistent inferential framework that is computationally simpler to describe and implement (albeit more expensive), and which can easily incorporate and compare a variety of prior beliefs regarding the true effect size. Moreover, Bayesian theory and modeling software is an extremely active area of current work, and Bayesian heritability analysis can avail itself of these many theoretical and computational tools.

2 Data

The subjects considered for this study consist of 209 healthy college students from Beijing Normal University and Sichuan Southwestern University.
The data were collected at laboratories at each of these universities over a 3-month time period. The students under consideration are a subset of the full study group, which consists of about 2000 subjects. The PI for this study is Gui Xue of the Cognition and Neuroscience Laboratory at Beijing Normal University.

Upon enrollment in the study, each subject provided a saliva sample from which genetic information was recorded in the form of approximately 500K SNPs. Each subject visited the lab on at least a weekly basis to undergo psychological testing and concurrent fMRI or EEG scanning. The psychological tests included the a decision-making task, working memory tests, reinforcement learning task, and resting state. The subjects also completed a variety of and behavioral personality inventory surveys.

We are here concerned only with the cups task experiment (previous work with this experiment by the PI is analyzed in Xue et al. (2008)). This experiment presents the subject with a group of "cups" from which one is to be selected. Each cup conceals a monetary reward for trials in the "gain domain", and a monetary loss for trials in the "loss domain." In each trial, one cup is designated as a sure win of \$1 (sure loss of \$1), and the remaining 3-5 cups contain a variable win (loss) of \$0-\$5. For each trial, the subjects are asked to make their choice of the no-risk cup vs one of the risky cups with the goal of maximizing their overall reward or minimizing their overall loss. The cups were shown until the subject's response, for up to 2.5s. After selecting a cup, the amount won or lost for the trial was shown for 0.5s. Experiment duration was 580s (9.5 min), with 140 trials given, with the ordering of 35 trials in the "gain" domain (subjects attempted to maximize reward), followed by 70 trials in the "loss" domain (subjects attempted to minimize loss), followed by 35 trials in the gain domain. Trial onsets were separated by approximately 4-5s. For the current data, the maximum response time across all subjects and trials was 1.5s, with a mean response time of 0.75s. The cups task trial design is illustrated in Figure 3.

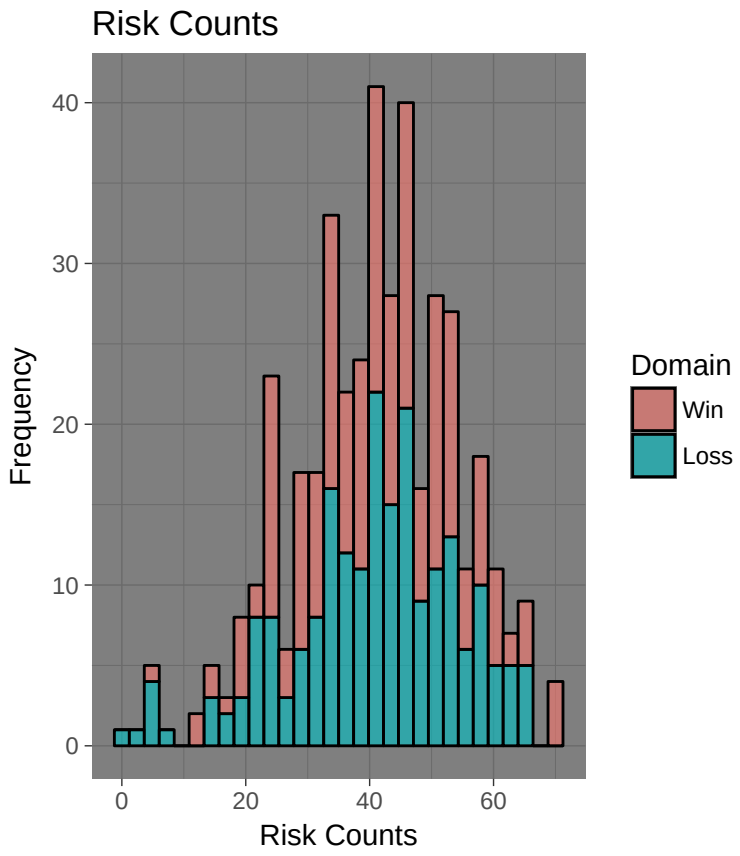


Figure 1: Distribution of risk counts by domain.

The intent of the experiment is to characteristics of a subject's risky decision-making, such as reaction time, aversion to risk, sensitivity to expected reward outcomes, and changes in behavior across the win/loss domains. The focus of the present analysis is risk preference in decision-making as measured by the number of times the risky option was chosen in the cups task experiment, specifically the heritability of risk preference in each domain, and the heritability of difference in risk preference across the domains. The distribution of risk counts by domain is given in Figure 1, and the distribution of differences of risk counts across domains in given in Figure 4. We note that the distribution of risk counts and risk count differences are both roughly normally distributed with a risk count mean of 41.3 and 40.7 for the gain and loss domains respectively, and with a mean of -0.6 for the risk count differences.

The total number of SNPs measured for each subject was approximately 500K. Using a simple selection procedure based on the significance of each SNP with the total number of risk counts, the SNPs are ordered by importance and the top 500 used to measure the genetic similarity of subjects. The SNPs were measured using the Affymetrix Genome-wide Human SNP Array. The genetic analysis software PLINK was used to filter any SNPs not meeting standard quality metrics (Purcell et al. 2007).

3 Methods

3.1 Statistical Definition of Heritability

In recent years, there has been substantial development of statistical models for heritability analysis. One of the simplest and most popular methods, relying on linear mixed effects models, has been packaged as

the Genome-wide Complex Trait Analysis (GCTA) algorithm. To derive the GCTA method, let Y be the scalar (continuous) response, let W be the $N \times S$ genetic data matrix for N subjects and S SNPs, and let $\varepsilon \sim N(0, \sigma_\varepsilon^2 I_N)$ be a vector of i.i.d. error terms. For convenience, we further assume the columns of W have been centered and scaled in a manner specified below. A natural model for this data is

$$Y = W\gamma + \varepsilon,$$

where $\gamma \sim N(0, \sigma_g^2 I_S)$ is a random coefficient vector of SNP effects. The resulting distribution of Y is

$$Y \sim N(0, \sigma_g^2 WW^T + \sigma_\varepsilon^2 I_N).$$

This leads to a more succinct form of the GCTA model

$$Y \sim N(0, \sigma_g^2 K + \sigma_\varepsilon^2 I_N),$$

where $K = WW^T$ is the genetic relationship matrix (GRM), which is an $N \times N$ matrix with ij-element calculated as

$$K_{ij} = \frac{1}{S} \sum_{s=1}^{S} \frac{(w_{is} - p_s)(w_{js} - p_s)}{2p_s(1 - p_s)}.$$

Here, w_{is} is the value of SNP s for subject i, and p_s is the frequency of the minor allele. The GRM can be understood as the genetic covariance matrix scaled by the expected heterozygosity.

From this parameterization, the heritability h^2 of the trait measured by Y is defined as

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_\varepsilon^2},$$

thus the primary target of inference is σ_ε^2. In practice the overall phenotypic variance $\sigma_g^2 + \sigma_\varepsilon^2$ is much easier to measure, and usually can be accurately estimated with the sample data or through additional data sources.

We are here concerned with estimating the influence of genetic factors on the sensitivity to win versus loss settings in the cups task experiment. We also consider strict counts of risks taken in each domain, which are nonnegative integer values, and can therefore be potentially modeled with a Poisson or negative binomial distribution.

3.2 Bayesian Inference for Heritability

The GCTA model was originally presented in the context of frequentist linear mixed effects (LME) models. This application of the LME model is sensible, but can present computational difficulties when estimating the variance components. More seriously, the power of the GCTA model is often insufficient for the sample and effect sizes that are common for current heritability studies.

The GCTA model can be easily adapted to a Bayesian framework, which greatly simplifies model fitting, and offers alternative estimates for h^2, which may perform better than the LME framework, or at least provides an alternative approach to compare the LME estimates.

The primary disadvantage of the Bayesian approach is the need to specify priors for the model parameters, and the increased computational cost brought by MCMC methods or other posterior approximation methods.

Considering a Bayesian version of the GCTA model, the most important interpretable parameter and primary quantity of interest is of course $0 \leq h^2 \leq 1$. The natural choice of prior for this quantity is therefore a

$Beta(\alpha, \beta)$ distribution, which is flexible enough to accommodate a variety of diffuse prior beliefs regarding the value of h^2. Since the model is parameterized in terms of σ_g^2 and σ_ε^2, we can induce a $Beta(\alpha, \beta)$ prior on h^2 by placing a $Gamma(\alpha, \theta)$ prior on σ_g^2 and an independent $Gamma(\beta, \theta)$ prior on σ_ε^2. By using HMC to calculate posterior estimates, we can avoid the need for conjugate distributions. Other priors for the variance components can certainly be considered, but care must be taken to ensure that the induced prior on h^2 places prior mass in appropriate regions. Specifically, for heritability studies with no strong prior beliefes, most of the mass should be placed near 0, with about 95% of the prior mass in the interval $[0, 0.5]$. The upper bound of this interval can easily be reduced to create more conservative priors.

The primary model employed here is

$$\begin{aligned} Y &\sim N(0, \sigma_g^2 K + \sigma_\varepsilon^2 I_N), \\ \sigma_g^2 &\sim Gamma(\alpha, \theta), \\ \sigma_\varepsilon^2 &\sim Gamma(\beta, \theta), \end{aligned}$$

with $\sigma_g^2 \perp \sigma_\varepsilon^2$, which induces the prior

$$h^2 \sim Beta(\alpha, \beta).$$

The hyperparameters α and β are chosen based on a desired prior mean and variance for h^2.
The present analysis uses a collection of values for these hyperparameters to assess the sensitivity of the results with respect to the choice of prior. The hyperparameter θ is chosen to give reasonably diffuse priors for σ_g^2 and σ_ε^2.

The prior mean of $\mu_0 = \alpha/(\alpha + \beta)$ of h^2 gives the relationship of the hyperparameters as

$$\beta = \frac{1 - \mu_0}{\mu_0}\alpha.$$

The variance component priors have variances given by

$$p(\sigma_g^2) \sim Gamma(\alpha, \theta) \Rightarrow Var_0(\sigma_g^2) = \frac{\alpha}{\theta^2}$$

$$p(\sigma_\varepsilon^2) \sim Gamma(\beta, \theta) \Rightarrow Var_0(\sigma_\varepsilon^2) = \frac{\beta}{\theta^2}.$$

To give a wide and roughly equal prior variance for σ_g^2 and σ_ε^2, we choose $\theta = 0.001$ for all models. To simplify the choice of model parameters, we also fix $\alpha = 1$ for all models, and choose β according to the desired prior mean. The models considered are given in Table 1. Note that, for a complex trait like risk-preference, a heritability of around 0.1 (10% variance explained) might be considered a lower bound on a practically significant level, as a very rough rule of thumb. A narrow-sense heritability (i.e. heritability from additive genetic effects) higher than 0.5 would be considered large and unexpected for this case. A highly heritable trait, such as height, has a broad-sense heritability (considering all possible genetic effects) upwards of 0.85.

All modeling and posterior inference was performed using the Bayesian modeling software Stan, as implemented by the package `rstan` (2016). This software provides a robust implementation of Hamiltonian Monte Carlo, which allows us to avoid worrying about issues of conjugacy and computational performance. Stan also provides many useful diagnostic checks and computational techniques to help ensure convergence of the sampler. Default sampler settings were used for all models: 2000 iterations with 10 leapfrog steps per iteration, a burn-in of 500, 4 chains initialized at different values, and no thinning.

Table 1: Model parameters for heritability of risk count differences.

	Prior.Mean.h2	alpha	beta	theta
Model 1	0.05	1	19	0.001
Model 2	0.10	1	9	0.001
Model 3	0.20	1	4	0.001
Model 4	0.50	1	1	0.001
All Genes	0.05	1	9	0.001

4 Results

4.1 Estimates of Heritability of Risk Count Differences

The inference results of the five models are given in Table 2 and Figure 4.1. Comparing the results of Models 1 - 4, which use the top 500 SNPs, we see that the inference results are fairly sensitive to the choice of prior mean. This is somewhat expected since the a sample size of 209 is low for this type of analysis.

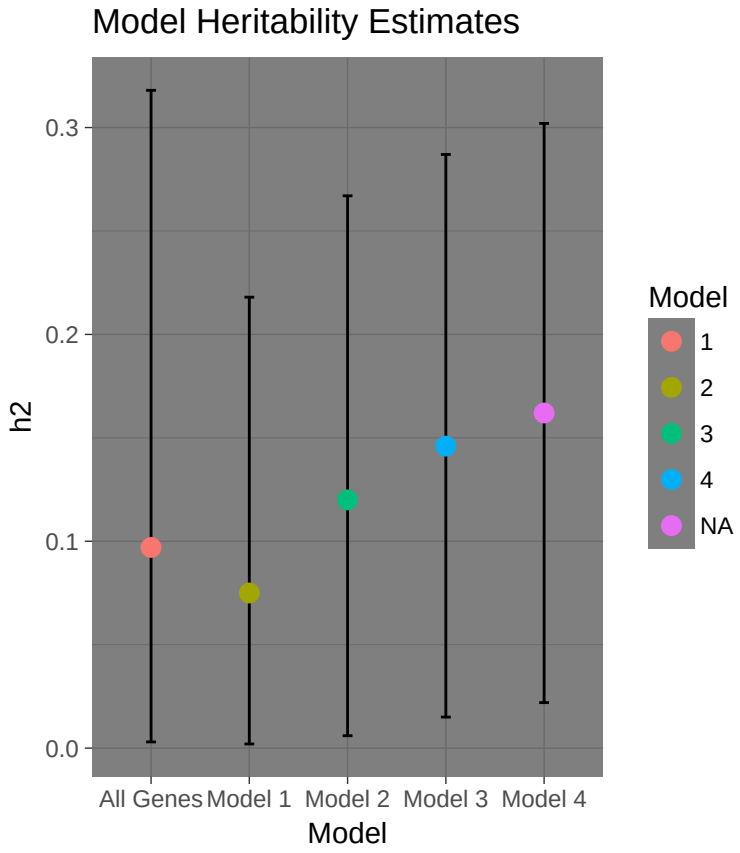
The posterior mean for h^2 for Model 1, the most conservative model with a prior mean of 0.05, is 0.075, with a 95% CI of $(0.002, 0.218)$. Model 2 has a prior mean of 0.1 and produces posterior mean 0.12 with a 95% CI $(0.006, 0.267)$. We cannot conclude a true practical effect from these estimates due to the excessively wide credibility intervals, but these results do at least verify that the model implementation and choice of prior are reasonable. The full study, of which the present sample is only a small subset, will contain upwards of 2000 subjects, and will thereby provide much greater precision in the posterior estimates.

The "All Genes" model, which uses 300K SNPs compared to the top 500 SNPs for the other models, uses the same prior as Model 2. We see that this model does produce a lower heritability estimate than Model 2, as expected. This gives some evidence that some SNPs have greater influence on others, and that an informed choice of of SNPs to consider can improve the inference results.

As an additional comparison, we also applied used the above model framework to estimate the heritability of total risks taken over the entire cups task experiment using the prior from Model 2. The estimated posterior mean heritability is $\hat{h}^2 = 0.066$ with 95% CI of $(0.002, 0.20)$, which is consistent with the risk count differences estimates.

Table 2: Comparison of model results.

h2_mean	SE	lower	upper	model
0.075	0.058	0.002	0.218	Model 1
0.120	0.070	0.006	0.267	Model 2
0.146	0.073	0.015	0.287	Model 3
0.162	0.074	0.022	0.302	Model 4
0.097	0.085	0.003	0.318	All Genes



\begin{figure} \caption{Posterior estimates for heritability of risk count differences, including posterior mean, standard error, and 95% credibility interval. Models 1 - 4 use the top 500 SNPs, while the All Genes model uses 300K SNPs.} \end{figure}

4.2 Model Diagnostics

Diagnostics of the HMC sampler for Model 2 are given in Figures 5 - 7.
The trace plots of samples for h^2 and the two variance components all show reasonable evidence of convergence. The ACF plots (not given) also indicate convergence, with very low autocorrelation at lag 20. The marginal posteriors and mean Metropolis acceptance rates do not indicate any excessively problematic sampling issues, although there are a handful of HMC iterations that resulted in a divergent transition (that is, the sampler moved far away from the expected high-density region of the posterior). The frequency of these divergent transitions was found to increase for increases in the prior mean of h^2, so a more detailed postmortem of these posterior samples may be in order.
It may be the case that a larger sample size will reduce the number and influence of these transitions. More importantly, the divergent transitions are within the high-density regions for the parameters at time of

divergence, so they should not be significantly distorting the resulting estimates.

Some posterior predictive checks are provided in Figures 8 & 9. Again, these diagnostics do not suggest any substantial problems with the posterior estimates, although we see from the posterior predictive densities that the model may be slightly underestimating the variance of the response data. Given the sample size, these diagnostics appear reasonable. With a larger sample these checks will allow for more detailed diagnostics.

5 Discussion

The results of the present analysis are restricted in power by the small sample size considered here and are therefore not conclusive in establishing significant heritability of risk count differences. However, given a moderate prior belief in heritability, even this small sample suggests that there may be significant genetic influences on decision-making and risk preferences. There may be sufficient evidence from previous studies to merit such as a belief. For example, Anokhin et al. (2009) found evidence of hertiability of risk-taking in adolescent twins, particulary for males. A similar twin study by Tuvblad et al. (2013) also found significant evidence of heritabilty of risk-preference in a gambling task. It may therefore be of value to consider how to incorporate these prior results into the present analysis to produce more informative priors.
The ability to include this prior information is a major advantage of the Bayesian approach, and, if carefully implemented, could offer much more reliable and stronger results than what is possible with the traditional GCTA method.

The present analysis shows that adapting the frequentist GCTA model to a Bayesian framework offers some substantial advantages, including a much easier implementation of the model fitting procedures, greater reliability in the estimates compared to the potentially difficult algorithms required for REML estimates in the LME model, and allows the use of the wide array of Bayesian diagnostic tools and theoretical results. Thus, at the very least, the Bayesian approach can serve as a useful check and comparison on the standard GCTA method.

The present sample is a small subset of the subjects in the Beijing study, which will contain more than 2000 subjects once the study is complete. The tools and code developed for this analysis can be easily applied to the whole data set, which will provide much more refined estimates of the heritability. Moreover, the code developed for this analysis can easily accommodate any scalar phenotype that can be well-approximated by a normal distribution. Other characteristics in the present data set that may be interesting candidates for heritability analysis include reaction time in the cups task, performance on the n-back working memory tasks, performance on the reinforcement learning task, and a variety of neuroimaging measures, such as the functional connectivity between specific brain regions.

As an additional advantage over the frequentist GCTA model, the Bayesian framework employed here can be more easily applied to multivariate phenotypes, such as functional connectivity matrices, and can also be used for the estimation of joint heritability of different scalar phenotypes, such as reaction time along with risk counts. These applications may potentially offer increases in power by leveraging the correlation of these characteristics. This approach is somewhat challenging to develop in a frequentist context, while Bayesian methods offer a much more unified and explicit framework to accomodate multivariate responses.

Considering future work, the modeling framework implemented here can also be adapted to account for data from exponential family distributions through the theory of generalized linear models. The recent work by Mair et al. (2015) give a summary of the developments of Bayesian GLMs for heritability analysis, including the generalized definition of heritability for these models. Given the wide variety of measurements in the present study, and the increasing interest in multi-modal imaging genetics data sets, these models may provide a useful direction for further research.

The potential downside of the Bayesian approach is of course the increased computational cost of approximating the posterior distribution with MCMC methods. However, there is extensive on-going work to develop more efficient algorithms and software. Given the fairly small sample sizes that are typical for neuroimaging studies, and the steady increase in computational power and availability of distributed systems, this computational

cost is not likely to be prohibitive in practice, particularly for scalar responses. For multivariate characteristics, the computational complexity may become intractable, but alternative frequentist methods are likely to be problematic in these instances as well.

6 Additional Figures

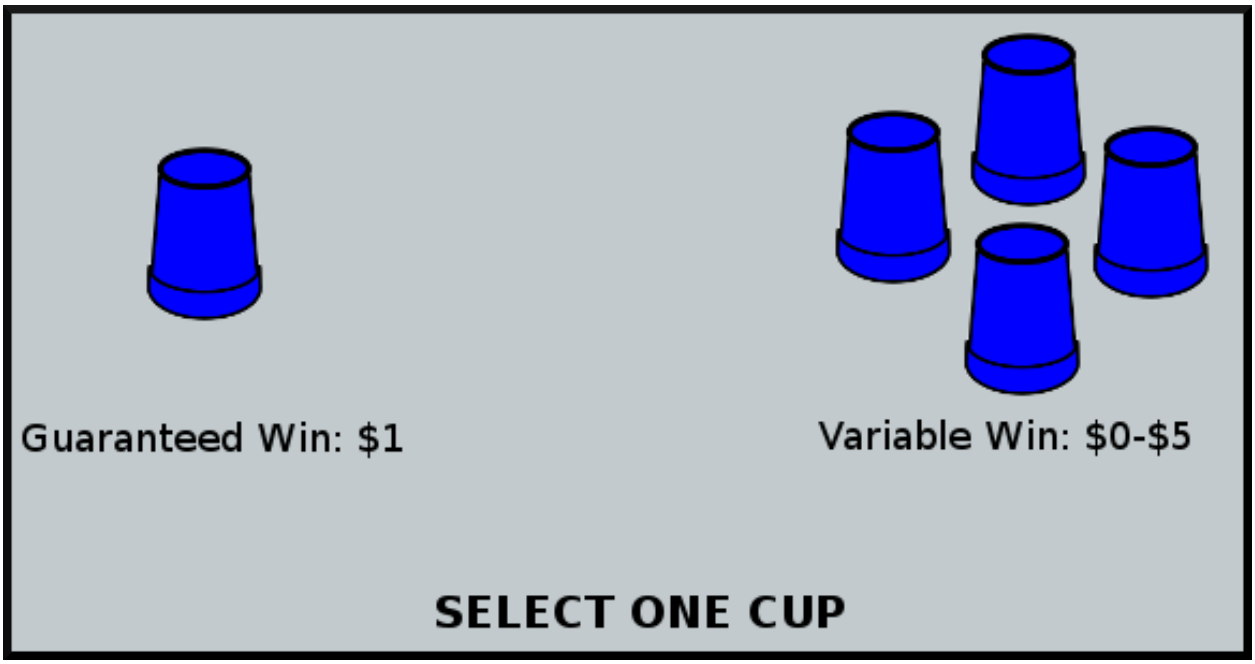


Figure 2: Example of a cups task trial.

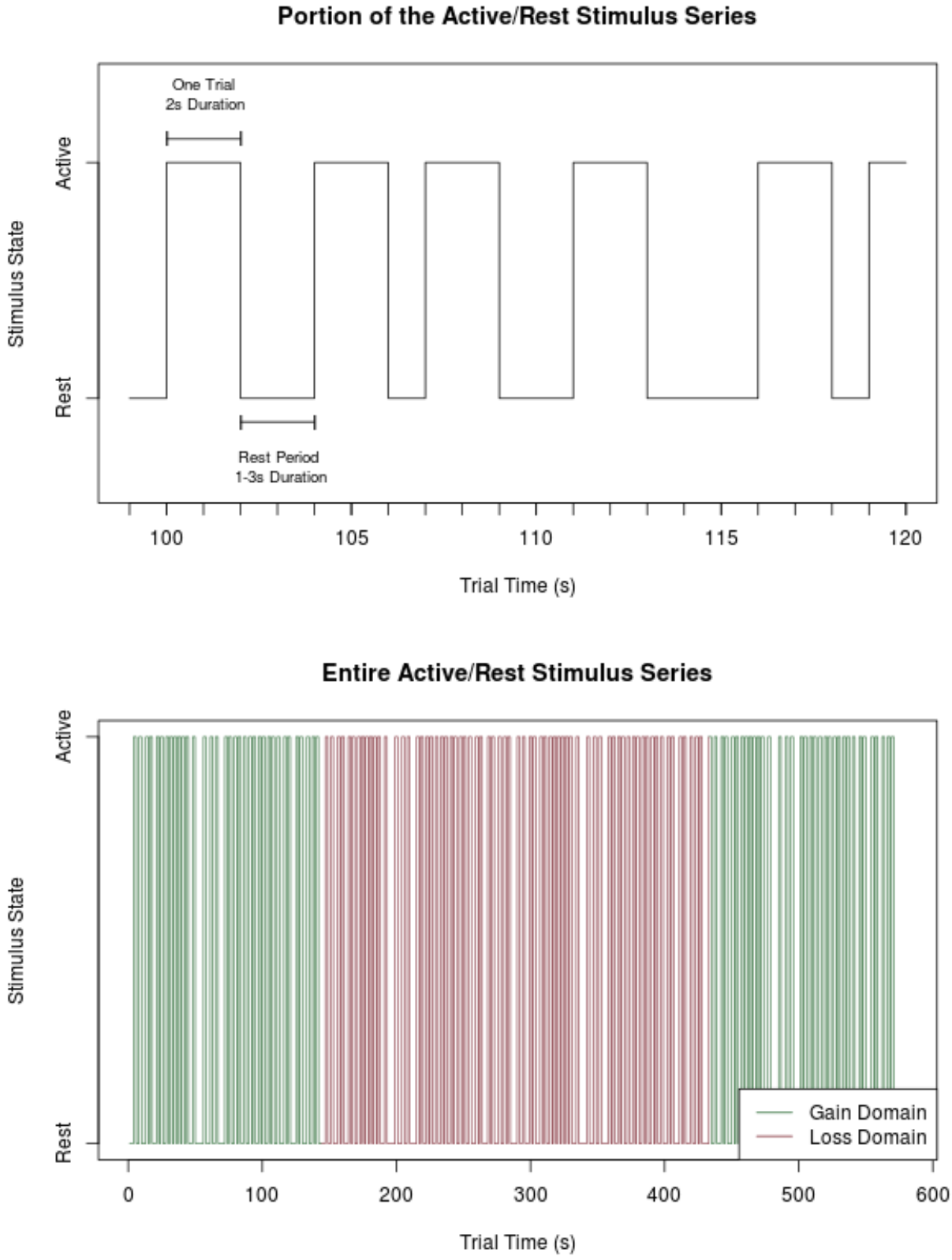


Figure 3: Stimulus presentation design for cups task experiment.

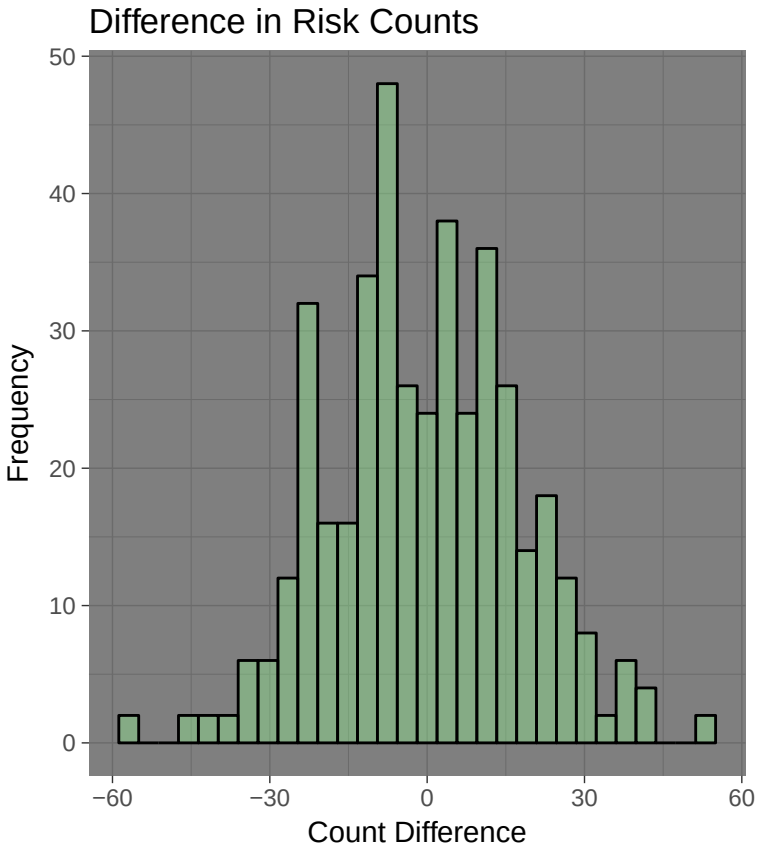


Figure 4: Distribution of differences risk counts across domains.

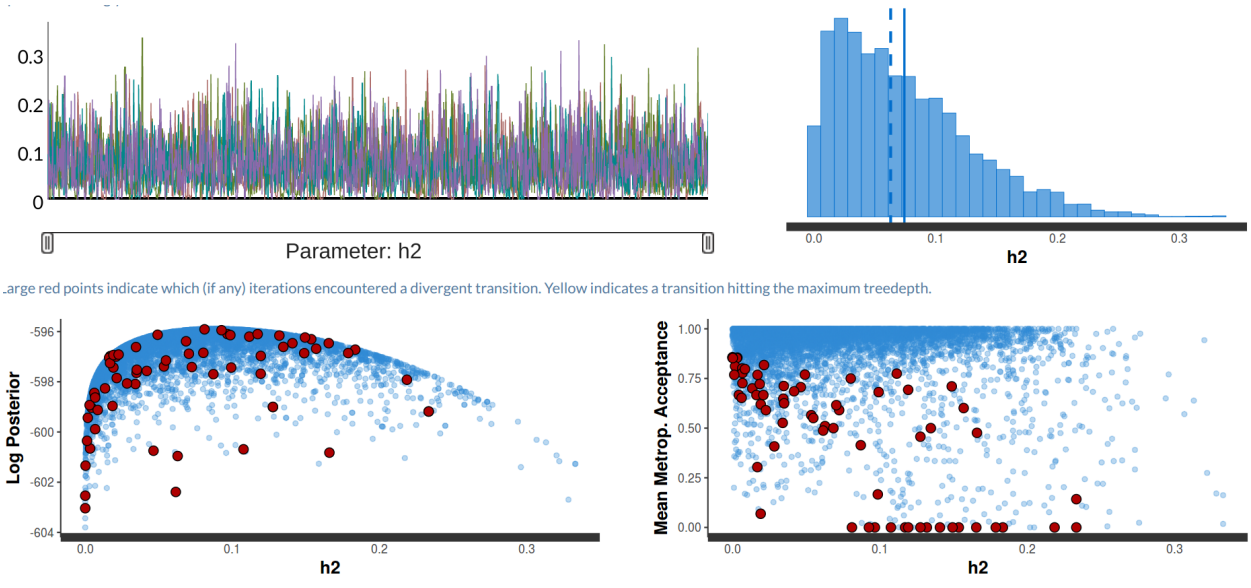


Figure 5: Diagnostics for h2 from Model 2.

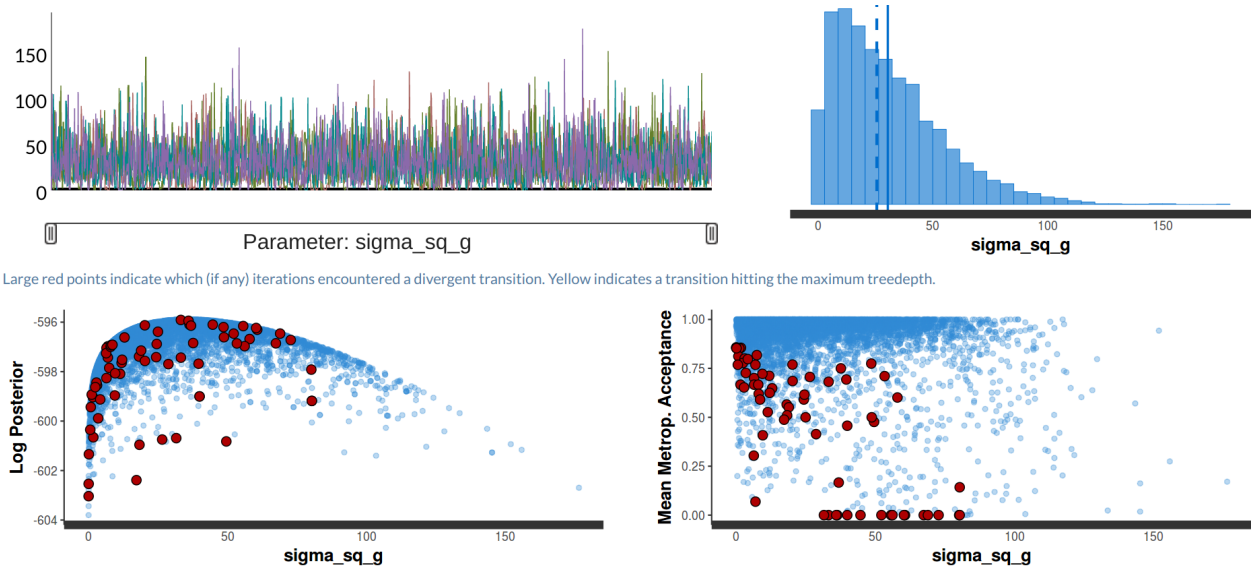


Figure 6: Diagnostics for genetic variance component from Model 2.

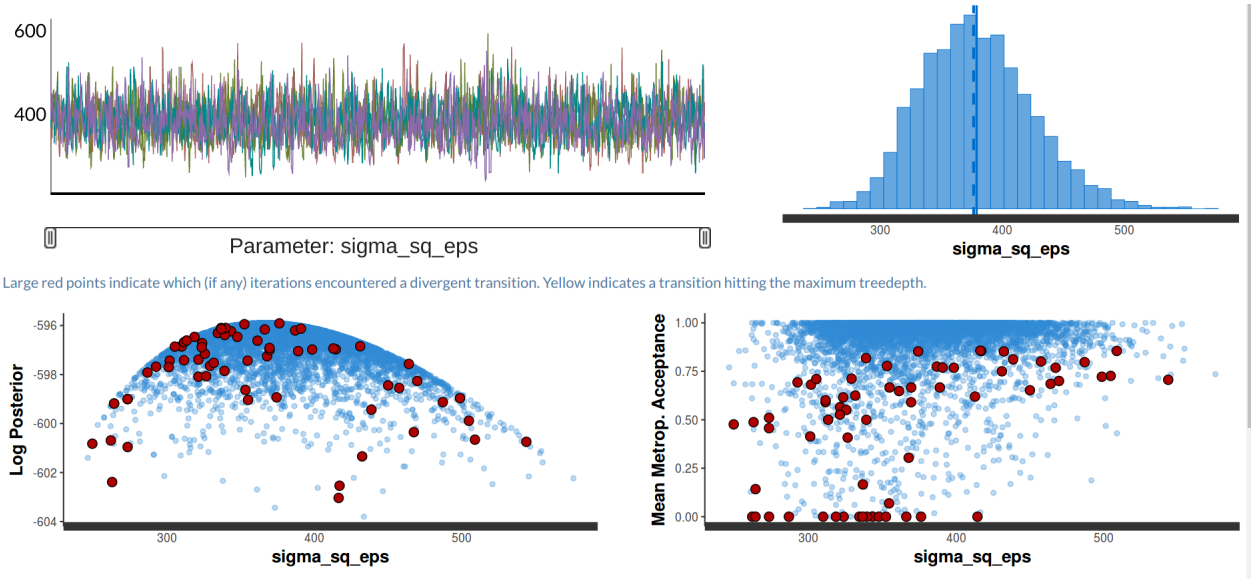


Figure 7: Diagnostics for error variance component from Model 2.

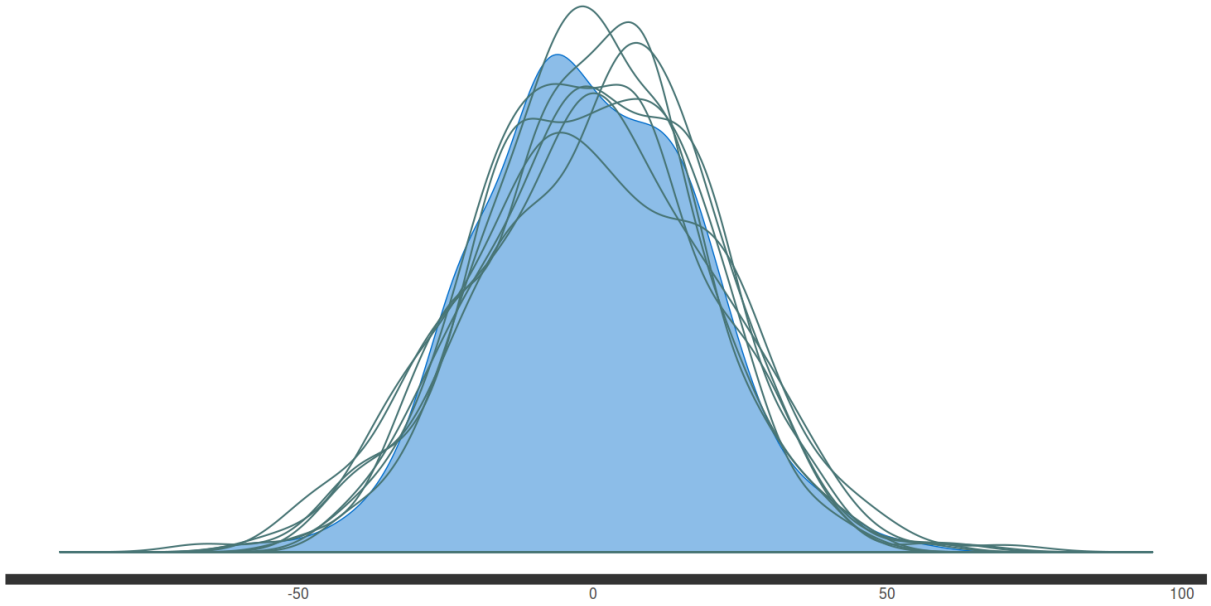


Figure 8: Posterior predictive densities for Model 2.

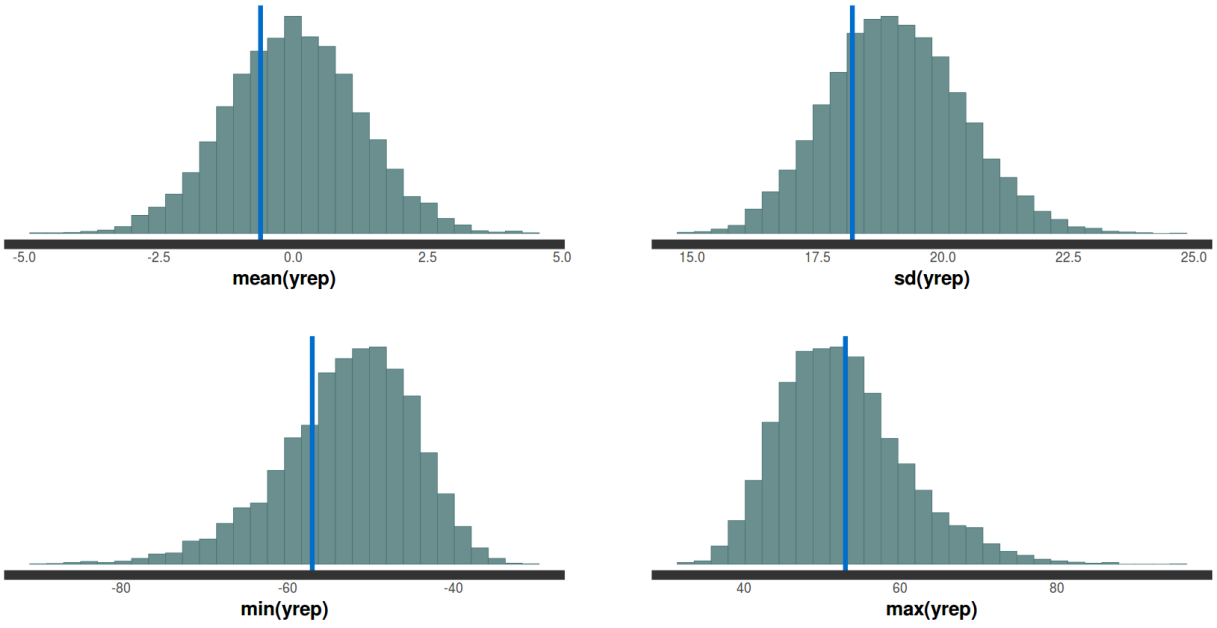


Figure 9: Empirical distributions of posterior predictive statistics for Model 2.

References

Anokhin, Andrey P, Simon Golosheykin, Julia Grant, and Andrew C Heath. 2009. "Heritability of Risk-Taking in Adolescence: A Longitudinal Twin Study." *Twin Research and Human Genetics* 12 (04). Cambridge Univ Press: 366–71.

Mair, Colette, Michael Stear, Paul Johnson, Matthew Denwood, Joaquin Prada Jimenez de Cisneros, Thorsten Stefan, and Louise Matthews. 2015. "A Bayesian Generalized Random Regression Model for Estimating Heritability Using Overdispersed Count Data." *Genetics Selection Evolution* 47 (1): 51. doi:10.1186/s12711-015-0125-5.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel AR Ferreira, David Bender, Julian Maller, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *The American Journal of Human Genetics* 81 (3). Elsevier: 559–75.

Stan Development Team. 2016. "RStan: The R Interface to Stan." http://mc-stan.org/.

Tuvblad, Catherine, Yu Gao, Pan Wang, Adrian Raine, Theodore Botwick, and Laura A Baker. 2013. "The Genetic and Environmental Etiology of Decision-Making: A Longitudinal Twin Study." *Journal of Adolescence* 36 (2). Elsevier: 245–55.

Xue, Gui, Zhonglin Lu, Irwin P Levin, Joshua A Weller, Xiangrui Li, and Antoine Bechara. 2008. "Functional Dissociations of Risk and Reward Processing in the Medial Prefrontal Cortex." *Cerebral Cortex* 19 (5). Oxford University Press: 1019–27.

Yang, Jian, S Hong Lee, Michael E Goddard, and Peter M Visscher. 2011. "GCTA: A Tool for Genome-Wide Complex Trait Analysis." *The American Journal of Human Genetics* 88 (1). Elsevier: 76–82.