Dealing with correlated tests

Yurii Aulchenko
Yurii [dot] Aulchenko [at] gmail [dot] com

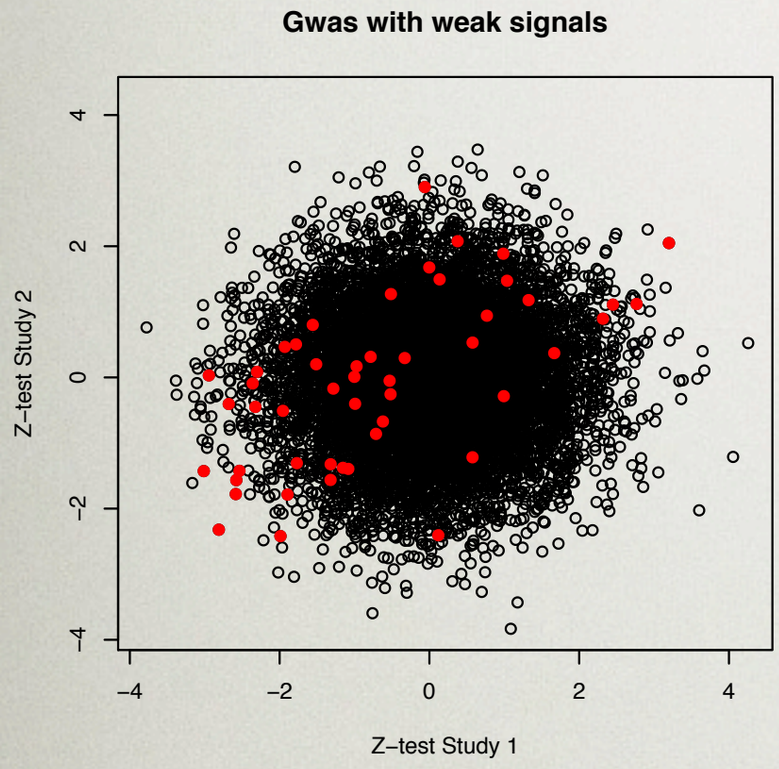
Standard scenario

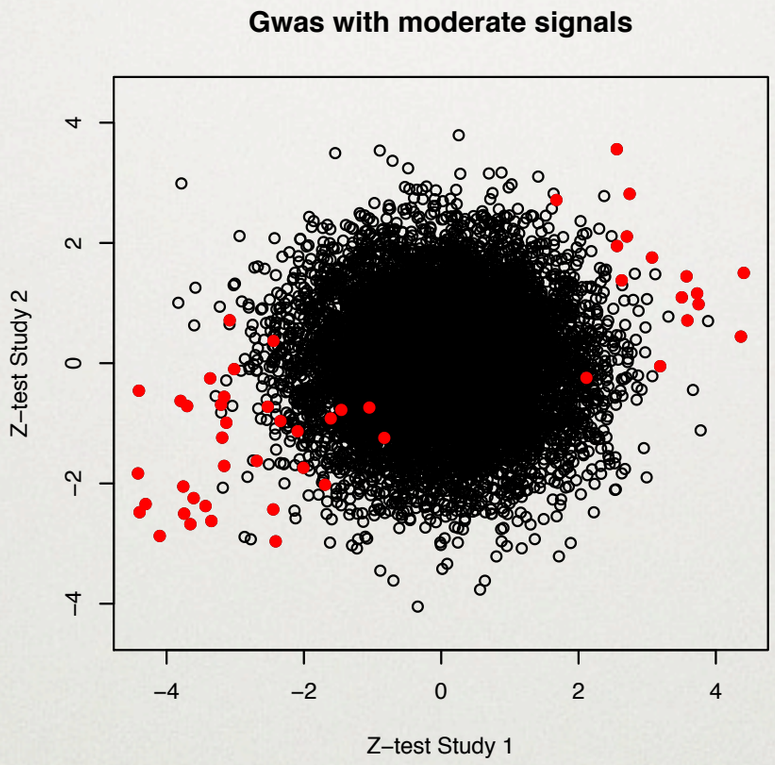
- You run GWAS analysis of a single trait
- The sample was genotyped using 500k SNP chip and imputed using HapMap panel to 2.5×10^6 variants
- What is your threshold *p*-value to claim genome-wide significance?

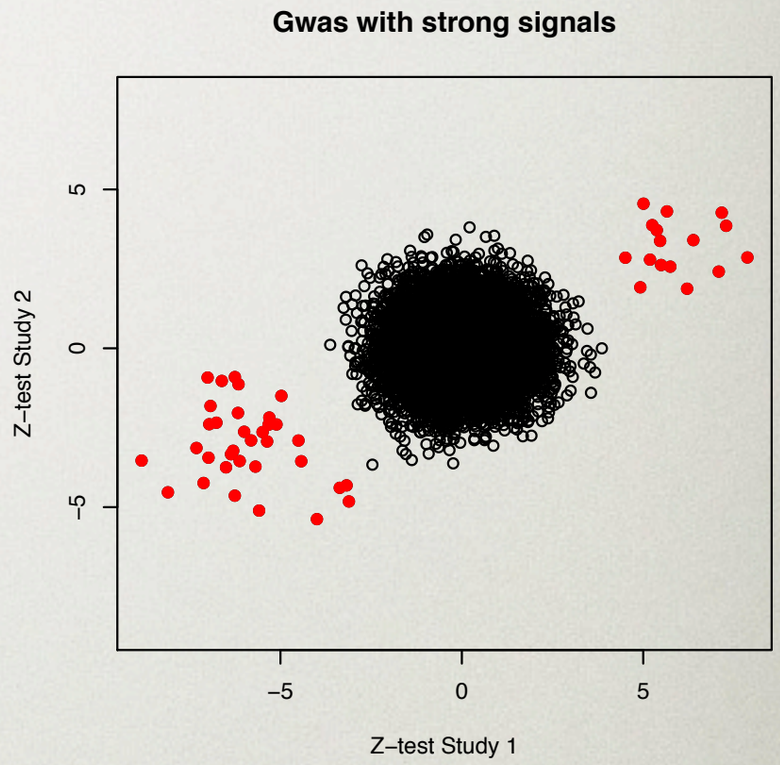
Standard scenario

- You run GWAS analysis of a single trait
- The sample was genotyped using 500k SNP chip and imputed using HapMap panel to 2.5×10^6 variants
- What is your threshold *p*-value to claim genome-wide significance?
- *p*-values $< 5 \times 10^{-8}$ are "significant"

Meta-GWAS of two studies







What is significance threshold?

- You analyzed 4 phenotypes (e.g. HDL, LDL, TC, TG)

What is significance threshold?

- You analyzed 4 phenotypes (e.g. HDL, LDL, TC, TG)
- You have analyzed 22,000 phenotypes ('omics' scenario)

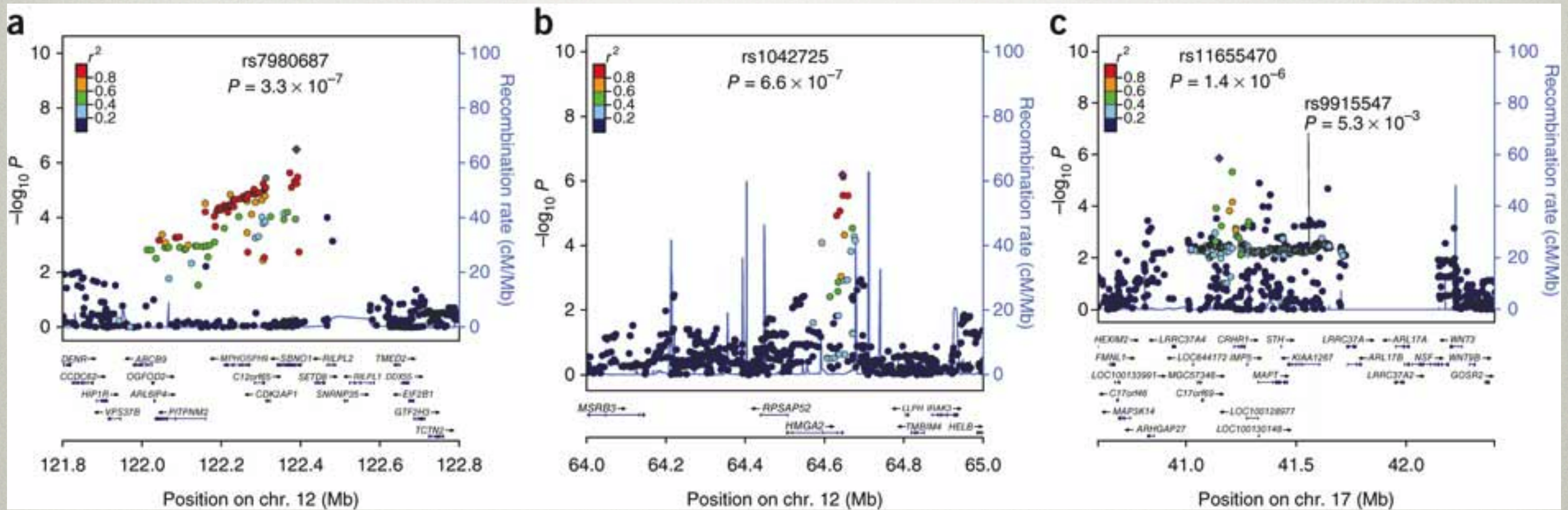
What is significance threshold?

- You analyzed 4 phenotypes (e.g. HDL, LDL, TC, TG)
- You have analyzed 22,000 phenotypes ('omics' scenario)
- You analyzed multiple SNPs in a region, and would like to have regional *p*-value

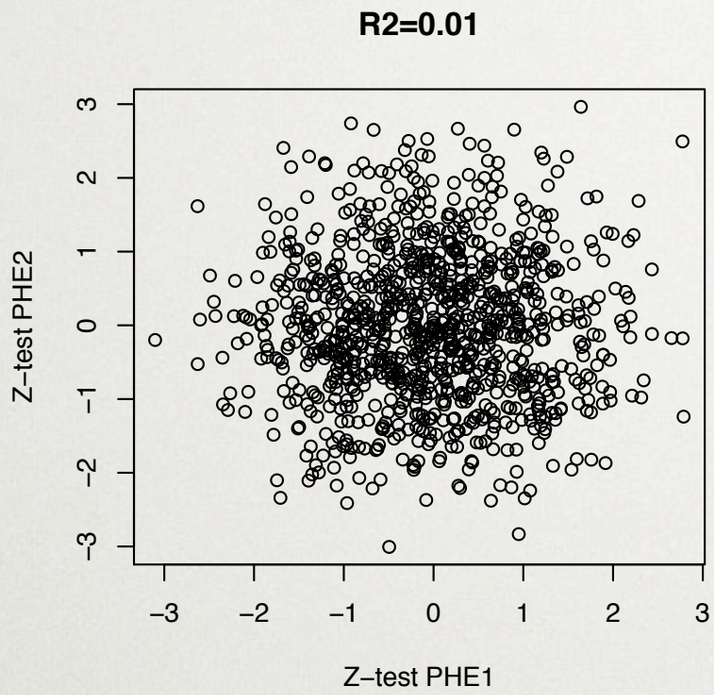
What is significance threshold?

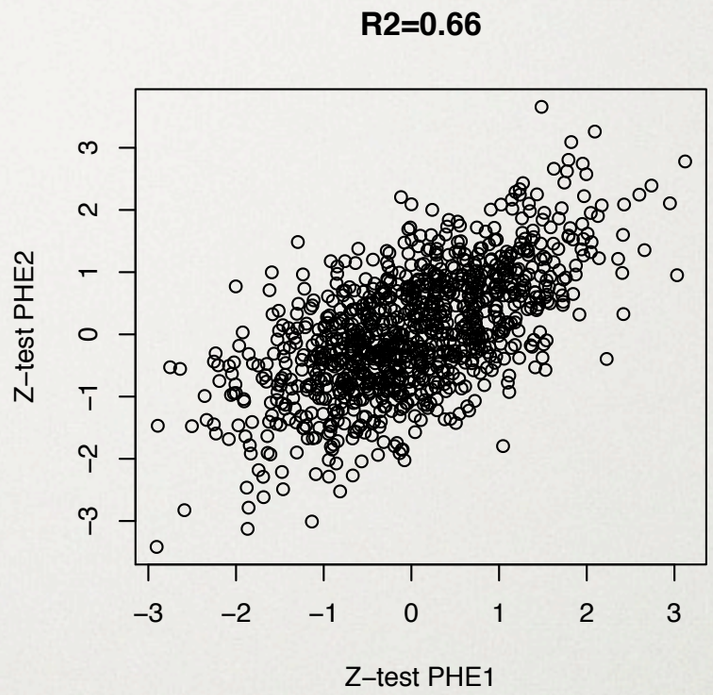
- You analyzed 4 phenotypes (e.g. HDL, LDL, TC, TG)
- You have analyzed 22,000 phenotypes ('omics' scenario)
- You analyzed multiple SNPs in a region, and would like to have regional *p*-value
- You did GWAS using several different models (e.g. additive and genotypic)

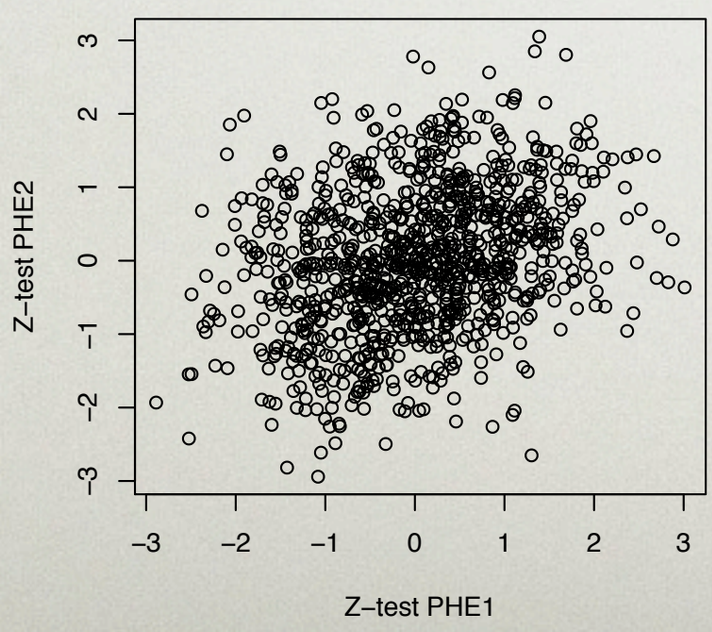
Regional associations

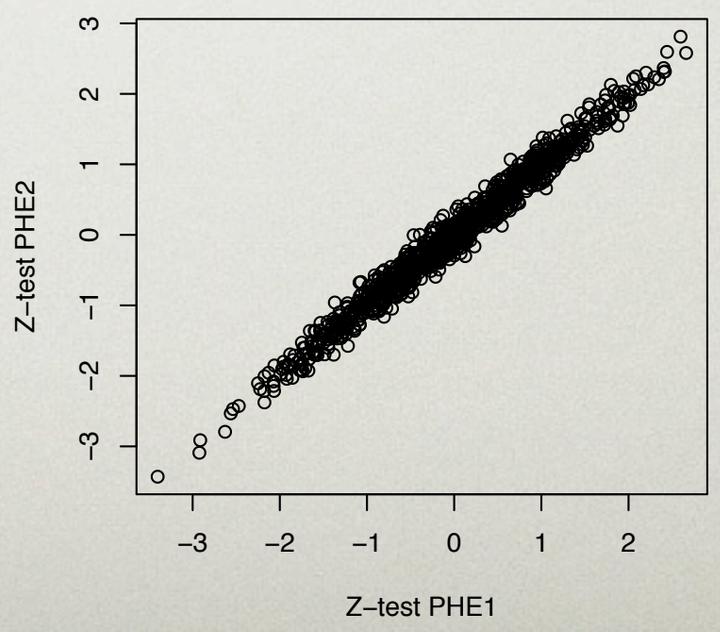


GWAS of two (correlated) traits

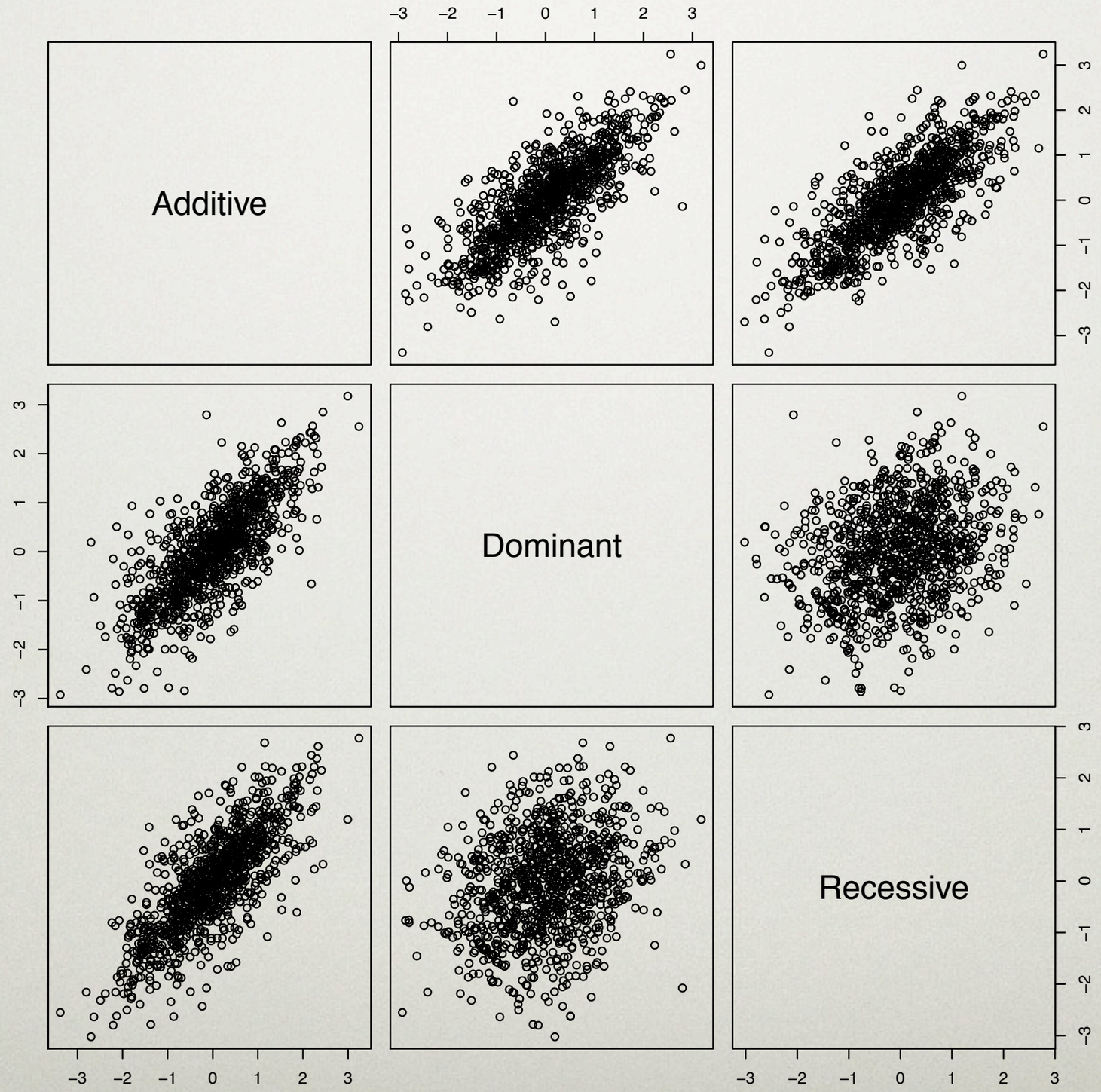








GWAS USING DIFFERENT MODELS



Empirical p-values

- Empirical techniques to derive null distribution of the test statistic (and thus approximation to exact *p*-value)
- Typically: permute the phenotypes, repeat analysis, ... 1000s of times

Empirical p-values

- Empirical techniques to derive null distribution of the test statistic (and thus approximation to exact *p*-value)
- Typically: permute the phenotypes, repeat analysis, ... 1000s of times
- If your GWAS analysis takes 5 minutes, deriving empirical thresholds will take few days

Empirical p-values

- Empirical techniques to derive null distribution of the test statistic (and thus approximation to exact *p*-value)
- Typically: permute the phenotypes, repeat analysis, ... 1000s of times
- If your GWAS analysis takes 5 minutes, deriving empirical thresholds will take few days
- ... some “single” analyses do take days!

Empirical P-values

Am J Hum Genet. 2005 Mar;76(3):399-408. Epub 2005 Jan 11.

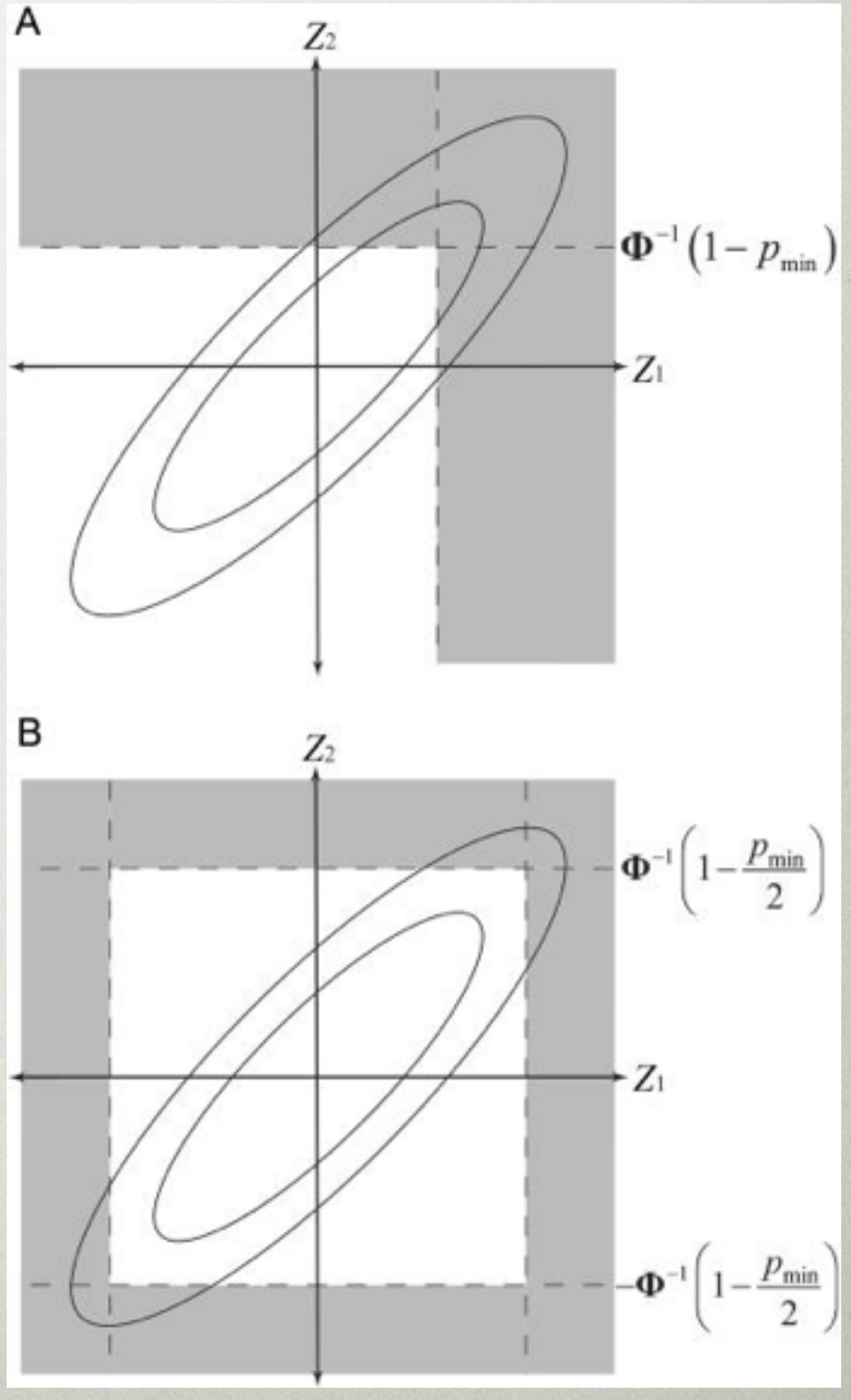
Rapid simulation of P values for product methods and multiple-testing adjustment in association studies.

Seaman SR, Müller-Myhsok B.

- Very smart speed-up was suggested by SSR & BMM
- Addresses very wide range of “typical” analysis scenarios
- It could be that ...
  - your scenario does not fall into “typical” ones
  - your data are not permutable (e.g. in structured populations)

P-ACT (*Conneely, Boehnke, 2007*)

- P-value Adjusted for Correlated Tests
- The idea is the the distribution of the Z-statistic from correlated tests follow multivariate normal distribution, characterized by some correlation matrix
- Hence the "overall" *p*-value can be computed as an integral over this distribution



P-ACT

$$P_{\text{ACT}} = \begin{cases} 1 - P[\max(Z_1, \ldots, Z_L) < \Phi^{-1}(1 - P_{\min})] & \text{for one-sided tests} \\ 1 - P\left[\max(|Z_1|, \ldots, |Z_L|) < \Phi^{-1}\left(1 - \frac{P_{\min}}{2}\right)\right] & \text{for two-sided tests} \end{cases},$$

- Sanity checks passed:

P-ACT

$$P_{ACT} = \begin{cases} 1 - P[\max(Z_1, \ldots, Z_L) < \Phi^{-1}(1 - P_{min})] & \text{for one-sided tests} \\ 1 - P\left[\max(|Z_1|, \ldots, |Z_L|) < \Phi^{-1}\left(1 - \frac{P_{min}}{2}\right)\right] & \text{for two-sided tests} \end{cases},$$

- Sanity checks passed:
  - If tests are not correlated, doing P-ACT becomes equivalent to Bonferroni/Sidak correction

P-ACT

$$P_{ACT} = \begin{cases} 1 - P[\max(Z_1, \ldots, Z_L) < \Phi^{-1}(1 - P_{min})] & \text{for one-sided tests} \\ 1 - P\left[\max(|Z_1|, \ldots, |Z_L|) < \Phi^{-1}\left(1 - \frac{P_{min}}{2}\right)\right] & \text{for two-sided tests} \end{cases},$$

- Sanity checks passed:
  - If tests are not correlated, doing P-ACT becomes equivalent to Bonferroni/Sidak correction
  - If statistics are perfectly correlated, P-ACT is equivalent to single-test *p*-value

Estimating S

- How do you know *S* (the correlation matrix for Z)?
- Different models on the same data and analysis of multiple traits: estimable directly from the analysis results
- Analysis of multiple SNPs: Conneely and Boehnke demonstrated that ***S*** is proportional to the genotypic correlation matrix

Simulations: multiple SNPs

Type I Error Rate and Power When 20 *HNF1A* SNPs Are Tested for Association with Binary Traits

				One Binary Trait Tested						Five Binary Traits Tested		
				On Additive Model			On Three Models			On Additive Model		
Disease SNP	MAF	r^2_{total}[a]	r^2_{max}[b]	$P_{Šidák}$	P_{ACT}	P_{perm}	$P_{Šidák}$	P_{ACT}	P_{perm}	$P_{Šidák}$	P_{ACT}	P_{perm}
None (type I error)	...	...	...	.0301	.0503	.0507	.0247	.0500	.0508	.0259	.0495	.0502
Most common SNP	.48	.88	.78	.899	.927	.925	.859	.911	.910	.806	.857	.859
Moderately frequent SNP	.20	.93	.19	.419	.535	.538	.338	.482	.484	.280	.385	.377
Least common SNP	.04	.91	.79	.878	.916	.915	.811	.874	.874	.686	.772	.773
SNP least predicted by others	.05	.42	.35	.387	.475	.476	.296	.401	.402	.220	.304	.299

[a] r^2_{total} = Proportion of variance in disease SNP allele count explained by the other 19 SNPs.
[b] r^2_{max} = Maximum pairwise r^2 between disease SNP and each of the other 19 SNPs.

Simulations: multiple traits

Type I Error Rate and Power When 10 Correlated Quantitative Traits Are Tested for Association

	10 Traits Tested for Association with									
	One SNP and a Covariate						20 Correlated *HNF1A* SNPs			
	Type I Error Rate			Power			Type I Error Rate		Power	
Trait Correlation Structure	$P_{\text{Šidák}}$	P_{ACT}	P_{perm}	$P_{\text{Šidák}}$	P_{ACT}	P_{perm}	$P_{\text{Šidák}}$	P_{ACT}	$P_{\text{Šidák}}$	P_{ACT}
Independent traits	.0498	.0499	.0496	.819	.819	.816	.0325	.0514	.780	.821
Equicorrelated traits	.0302	.0502	.0503	.826	.880	.878	.0216	.0507	.778	.852
Autocorrelated traits	.0393	.0494	.0495	.820	.842	.839	.0274	.0499	.777	.833
Independent blocks of traits	.0386	.0497	.0501	.824	.850	.848	.0264	.0501	.779	.836
Negatively correlated blocks of traits	.0327	.0496	.0500	.825	.870	.868	.0234	.0503	.779	.846
Five binary and five quantitative traits	.0341	.0491	.0488	.825	.864	.860	.0263	.0517	.781	.844

Summary P-ACT

- Approximates exact *p*-value very well
- Is computationally much faster than permutations

- **Caution:** P-ACT requires integration over high-D multivariate normal. Numerically, the results become not stable/reliable when the Z-values are very large and/or there are too many dimensions

Simes-type methods addressing situation

Am J Hum Genet. 2011 March 11; 88(3): 283–293. PMCID: PMC3059433
doi: 10.1016/j.ajhg.2011.01.019

GATES: A Rapid and Powerful Gene-Based Association Test Using Extended Simes Procedure

Miao-Xin Li,[1,2,3] Hong-Sheng Gui,[1] Johnny S.H. Kwan,[1] and Pak C. Sham[1,2,3,*]

PLoS Genet. 2013 January; 9(1): e1003235. PMCID: PMC3554627
Published online 2013 January 24. doi: 10.1371/journal.pgen.1003235

TATES: Efficient Multivariate Genotype-Phenotype Analysis for Genome-Wide Association Studies

Sophie van der Sluis,[1,*] Danielle Posthuma,[1,2,3] and Conor V. Dolan[4,5]

Simes/GATES/TATES

Given p - ascending vector of (correlated) p-values, define overall p_G as

$$P_G = Min\left(\frac{m_e p_{(j)}}{m_{e(j)}}\right),$$

where m_e is the effective number of independent p values among the m SNPs and $m_{e(j)}$ is the effective number of independent p-values among the top j SNPs. The value of m_e is estimated to be equal to

$$M - \sum_{i=1}^{M} [I(\lambda_i > 1)(\lambda_i - 1)] \quad \lambda_i > 0$$

where $I(x)$ is an indicator function and λ_i is the i^{th} eigenvalue of the p value correlation coefficient matrix $[\rho_{i,j}]$ of SNP-based statistic tests

Summary

- Ideally: empirical *p*-values. Best tool in class is WGPERMER (Stephan Ripke, Bertram Muller-Myhsok)
- If not, consider P-ACT. This is easily implemented in R. Do test the stability of the results!
- If not, consider Simes/GATES/TATES. Easily implemented in R. The methods are new: do sanity checks.