Lecture 9: Kernel (Variance Component) Tests and Omnibus Tests for Rare Variants

Timothy Thornton and Michael Wu

Summer Institute in Statistical Genetics 2019

Lecture Overview

1. Variance Component Tests
2. Omnibus Tests
3. Weights

Recall: Region Based Analysis of Rare Variants

- Single variant test is not powerful to identify rare variant associations
- Strategy: Region based analysis
  - Test the joint effect of rare/common variants in a gene/region while adjusting for covariates.

Major Classes of Tests

- Burden/Collapsing tests
- Supervised/Adaptive Burden/Collapsing tests
- Variance component (similarity) based tests
- Omnibus tests: hedge against difference scenarios

Variance component test

- Burden tests are not powerful, if there exist variants with different association directions or many non-causal variants
- Variance component tests have been proposed to address it.
- "Similarity" based test

C-alpha test

Neale BM, *et al.*(2011). *Plos Genet.*

- Case-control studies without covariates.
- Assume the jth variant is observed n_{j1} times, with r_{j1} times in cases.

	a	A	Total
Case	r_{j1}	r_{j2}	r
Control	s_{j1}	s_{j2}	s
Total	n_{j1}	n_{j2}	n

- Under H_0

$$r_{j1} \sim Binomial(n_{j1}, q) \quad (q = r/n)$$

C-alpha test

- Risk increasing variant:

$$r_{j1} - qn_{j1} > 0$$

- Risk decreasing variant:

$$r_{j1} - qn_{j1} < 0$$

- Test statistic:

$$T_\alpha = \sum_{j=1}^{p} (r_{j1} - qn_{j1})^2 - \sum_{j=1}^{p} n_{j1} q(1-q)$$

- This test is robust in the presence of the opposite association directions.

C-alpha test

- Weighting scheme

$$T_\alpha = \sum_{j=1}^{p} w_j (r_{j1} - q n_{j1})^2 - \sum_{j=1}^{p} w_j n_{j1} q(1-q)$$

- Test for the over-dispersion due to genetic effects
  - Neyman's $C(\alpha)$ test.

C-alpha test, P-value calculation

- Using normal approximation, since the test statistic is the sum of random variables.

$$T_\alpha = \sum_{j=1}^{p} (r_{j1} - q n_{j1})^2 - \sum_{j=1}^{p} n_{j1} q (1-q)$$

- Doesn't work well when p is small (or moderate).
  - P-value is computed using permutation.

C-alpha test

- C-alpha test is robust in the presence of the different association directions
- Disadvantages:
  - Permutation is computationally expensive.
  - Cannot adjust for covariates.

Sequence Kernal Association Test (SKAT)

Wu *et al.*(2010, 2011). *AJHG*

- Recall the original regression models:

$$\mu_i / logit(\mu_i) = \alpha_0 + \mathbf{X}_i^T \boldsymbol{\alpha} + \mathbf{G}_i^T \boldsymbol{\beta}$$

- Variance component test:
  - Assume $\beta_j \sim dist.(0, w_j^2 \tau)$.
  - $H_0 : \beta_1 = \cdots = \beta_p = 0 <=> H_0 : \tau = 0$.

Sequence Kernel Association Test (SKAT)

- $\beta_j \sim dist.(0, w_j^2\tau)$: $\tau = 0$ is on the boundary of the hypothesis.
- Score test statistic for $\tau = 0$:

$$Q_{SKAT} = (\mathbf{y} - \widehat{\boldsymbol{\mu}}_0)'\mathbf{K}(\mathbf{y} - \widehat{\boldsymbol{\mu}}_0),$$

- $\mathbf{K} = \mathbf{GWWG}'$: weighted linear kernel ($\mathbf{W} = diag[w_1, \ldots, w_p]$).

Sequence Kernel Association Test (SKAT)

- The C-alpha test is a special case of SKAT
  - With no covariates and flat weights:

$$Q_{SKAT} = \sum_{j=1}^{p} (r_{j1} - q n_{j1})^2$$

SKAT

- Q_{SKAT} is a weighted sum of single variant score statistics

$$Q_{SKAT} = (\mathbf{y} - \widehat{\boldsymbol{\mu}}_0)' \mathbf{GWWG}' (\mathbf{y} - \widehat{\boldsymbol{\mu}}_0)$$
$$= \sum_{j=1}^{p} w_j^2 [\boldsymbol{g}_j' (\mathbf{y} - \widehat{\boldsymbol{\mu}}_0)] = \sum_{j=1}^{p} w_j^2 U_j^2$$

  where $U_j = \sum_{i=1}^{n} g_{ij}(y_i - \widehat{\mu}_{0i})$.

- U_j is a score of individual SNP j only model:

$$\mu_i / logit(\mu_i) = \alpha_0 + \mathbf{X}_i^T \boldsymbol{\alpha} + g_{ij}\beta_j$$

SKAT

- Q_{SKAT} (asymptotically) follows a mixture of χ^2 distribution under the NULL.

$$
\begin{aligned}
Q &= (\mathbf{y} - \widehat{\mu}_0)' \mathbf{K} (\mathbf{y} - \widehat{\mu}_0) \\
&= (\mathbf{y} - \widehat{\mu}_0)' \widehat{\mathbf{V}}^{-1/2} \widehat{\mathbf{V}}^{1/2} \mathbf{K} \widehat{\mathbf{V}}^{1/2} \widehat{\mathbf{V}}^{-1/2} (\mathbf{y} - \widehat{\mu}_0) \\
&= \sum_{j=1}^{p} \lambda_j [\mathbf{u}_j' \widehat{\mathbf{V}}^{-1/2} (\mathbf{y} - \widehat{\mu}_0)]^2 \\
&\approx \sum_{j=1}^{p} \lambda_j \chi^2_{1,j}
\end{aligned}
$$

SKAT

- λ_j and $\mathbf{u}_j$ are eigenvalues and eigenvectors of $\mathbf{P}^{1/2}\mathbf{K}\mathbf{P}^{1/2}$. where $\mathbf{P} = \widehat{\mathbf{V}}^{-1} - \widehat{\mathbf{V}}^{-1}\widetilde{\mathbf{X}}(\widetilde{\mathbf{X}}'\widehat{\mathbf{V}}^{-1}\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\widehat{\mathbf{V}}^{-1}$ is the project matrix to account that $\boldsymbol{\alpha}$ is estimated.

SKAT: P-value calculation

- P-values can be computed by inverting the characteristic function using Davies' method (1973, 1980)
  - Characteristic function

$$\varphi_x(t) = E(e^{itx}).$$

  - Characteristic function of $\sum_{j=1}^{p} \lambda_j \chi^2_{1,j}$

$$\varphi_x(t) = \prod_{i=j}^{p} (1 - 2\lambda_j it)^{-1/2}.$$

  - Inversion Formula

$$P(X < u) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{Im[e^{-itu}\varphi_x(t)]}{t} \, \mathrm{d}t.$$

Small sample adjustment

Lee *et al.*(2012). *AJHG*

- When the sample size is small and the trait is binary, asymptotics does not work well.
- SKAT test statistic:

$$\begin{aligned} Q_{SKAT} &= (\mathbf{y} - \widehat{\boldsymbol{\mu}}_0)' \mathbf{K} (\mathbf{y} - \widehat{\boldsymbol{\mu}}_0) \\ &= \sum_{v=1}^{p} \lambda_v \eta_v^2, \end{aligned}$$

- η_vs are asymptotically independent and follow N(0,1).

Small sample adjustment

- When the trait is binary and the sample size is small:
  - $Var(\eta_v) < 1$.
  - η_vs are negatively correlated.

Small sample adjustment

- Mean and variance of the Q_{SKAT}

	Mean	Variance
Large Sample	$\sum \lambda_j$	$\sum \lambda_j^2$
Small Sample	$\sum \lambda_j$	$\sum \lambda_j \lambda_k c_{jk}$

- Adjust null distribution of Q_{SKAT} using the estimated small sample variance.

Small sample adjustment

- Variance adjustment is not enough to accurately approximate far tail areas.
- Kurtosis adjustment:
  - Estimate the kurtosis of Q_{SKAT} using parametric bootstrapping:
  - $\widehat{\gamma}$ (estimated kurtosis)
  - D.F. estimator: $\widehat{df} = 12/\hat{\gamma}$
  - Null distribution

$$(Q_{SKAT} - \sum \lambda_j^2) \frac{\sqrt{2\widehat{df}}}{\sqrt{\sum \lambda_j \lambda_k c_{jk}}} + \widehat{df} \sim \chi^2_{\widehat{df}}$$

Small sample adjustment

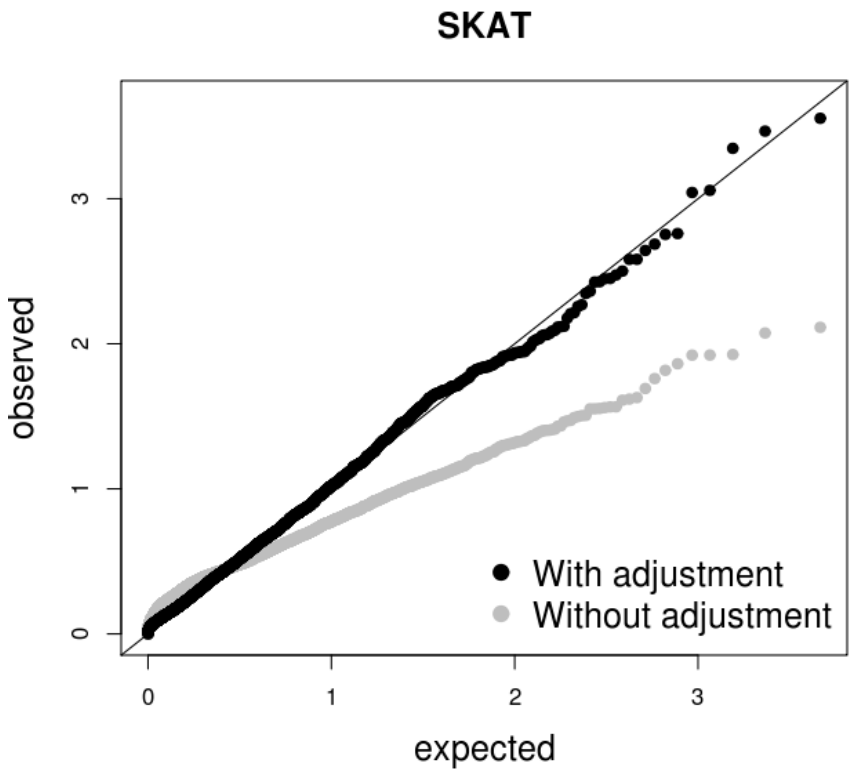


Figure: ARDS data (89 samples)

General SKAT

- General SKAT Model:

$$\mu_i / logit(\mu_i) = \alpha_0 + X_i \boldsymbol{\alpha} + h_i$$

  where $h_i \sim GP(0, \tau K)$.
- Kernel $K(\mathbf{G}_i, \mathbf{G}_{i'})$ measures genetic similarity between two subjects.

General SKAT

- Examples:
  - Linear kernel=linear effect

$$K(\mathbf{Z}_i, \mathbf{Z}_{i'}) = w_1^2 Z_{i1} Z_{i'1} + \cdots w_p^2 Z_{ip} Z_{i'p}$$

  - IBS Kernel (Epistatic Effect: SNP-SNP interactions)

$$K(\mathbf{Z}_i, \mathbf{Z}_j) = \frac{\sum_{k=1}^{p} w_k^2 IBS(Z_{ik}, Z_{jk})}{2p}$$

Omnibus Tests

- Questions:
  - Which group of variants test? I.e. what is the threshold for "rare"?
  - Which type of test should I use? Variance component or burden?
- Truth is unknown: depends on the situation
- Omnibus tests: work well across situation

Variable threshold (VT) test

- Most methods use a fixed threshold for rare variants: $\leq 0.5\%$, $\leq 1\%$, ... $\leq 5\%$?
- Choosing an appropriate threshold can have a huge impact on power: prefer to restrict analysis to meaningful variants

Variable threshold (VT) test

Price AL, Kryukov GV, *et al.*(2010) *AJHG*

- Find the optimal threshold to increase the power.
  - Weight:

$$w_j(t) = \begin{cases} 1 & \text{if} \quad maf_j \leq t \\ 0 & \text{if} \quad maf_j > t \end{cases}$$

  - $C_i(t) = \sum w_j(t) g_{ij}$
  - Test statistics:

$$Z_{max} = max_t Z(t)$$

    where $Z(t)$ is a Z-score of C_i.

P-value Calculations of Variable threshold (VT) test

- Price *et al.*proposed to use permutation to get a p-value
- Lin and Tang (2011) showed that the p-values can be calculated through numerical integration using normal approximation

Variable threshold (VT) test

- More robust than using a fixed threshold.
- Provide information on the MAF ranges of the causal variants.
- Lose power if there exist variants with opposite association directions.

SKAT vs. Collapsing

- Collapsing tests are more powerful when a large % of variants are causal and effects are in the same direction.
- SKAT is more powerful when a small % of variants are causal, or the effects have mixed directions.
- Both scenarios can happen when scanning the genome.
- Best test to use depends on the underlying biology.

  $\rightarrow$ Difficult to choose which test to use in practice.

We want to develop a unified test that works well in both situations. $\rightarrow$ Omnibus tests

Combine p-values of Burden and SKAT

Derkach A *et al.*(2013) *Genetic Epi*, 37:110-121

- Fisher method:

$$Q_{Fisher} = -2\log(P_{Burden}) - 2\log(P_{SKAT})$$

- Q_{Fisher} follows χ^2 with 4 d.f when these two p-values are independent
- Since they are not independent, p-values are calculated using resampling
- Mist (Sun et al. 2013) modified the SKAT test statistics to make them independent

Combine Test Statistics: Unified Test Statistics

Lee *et al.*(2012). *Biostatistics*

- Combined Test of Burden tests and SKAT

$$Q_\rho = (1-\rho)Q_{SKAT} + \rho Q_{Burden}, \quad 0 \le \rho \le 1.$$

- Q_ρ includes SKAT and burden tests.
  - $\rho = 0$: SKAT
  - $\rho = 1$: Burden

Derivation of the Unified Test Statistics

- Model:

$$g(\mu_i) = \mathbf{X}_i\boldsymbol{\alpha} + \mathbf{G}_i\boldsymbol{\beta}$$

  where β_j/w_j follows any arbitrary distribution with mean 0 and variance τ and the correlation among β_j's is ρ.
- Special cases:
  - SKAT: $\rho = 0$
  - Burden: $\rho = 1$
  - Combined: $0 \leq \rho \leq 1$

Derivation of the Unified Test Statsitics

- Q_ρ is a test statistic of the SKAT with $corr(\beta) = \mathbf{R}(\rho)$:
  - $\mathbf{R}(\rho) = (1-\rho)\mathbf{I} + \rho \underline{1}\underline{1}'$ (compound symmetric)
  - $\mathbf{K}_\rho = \mathbf{GWR}(\rho)\mathbf{WG}'$.

$$\begin{aligned} Q_\rho &= (\mathbf{y} - \hat{\boldsymbol{\mu}})' \mathbf{K}_\rho (\mathbf{y} - \hat{\boldsymbol{\mu}}) \\ &= (1-\rho) Q_{SKAT} + \rho Q_{Burden} \end{aligned}$$

Adaptive Test (SKAT-O)

- Use the smallest p-value from different ρs:

$$T = \inf_{0 \leq \rho \leq 1} P_\rho.$$

  where P_ρ is the p-value of Q_ρ for given ρ.

- Test statistic:

$$T = minP_{\rho_b}, \quad 0 = \rho_1 < \ldots < \rho_B = 1.$$

Adaptive Test (SKAT-O)

- Q_ρ is a mixture of two quadratic forms.

$$\begin{aligned} Q_\rho &= (1-\rho)(\mathbf{y}-\hat{\boldsymbol{\mu}})' GWWG'(\mathbf{y}-\hat{\boldsymbol{\mu}})' \\ &\quad + \rho(\mathbf{y}-\hat{\boldsymbol{\mu}})' GW\underline{1}\underline{1}'WG'(\mathbf{y}-\hat{\boldsymbol{\mu}}) \\ &= (1-\rho)(\mathbf{y}-\hat{\boldsymbol{\mu}})' K_1(\mathbf{y}-\hat{\boldsymbol{\mu}})' + \rho(\mathbf{y}-\hat{\boldsymbol{\mu}})' K_2(\mathbf{y}-\hat{\boldsymbol{\mu}}) \end{aligned}$$

- Q_ρ is asymptotically equivalent to

$$(1-\rho)\kappa + a(\rho)\eta_0,$$

where and $\eta_0 \sim \chi^2_1$, κ approximately follows a mixture of χ^2.

SKAT-O

- Q_ρ is the asymptotically same as the sum of two independent random variables.

$$(1-\rho)\kappa + a(\rho)\eta_0$$

  - $\eta_0 \sim \chi^2_1$
  - Approximate κ via moments matching.

- P-value of T:

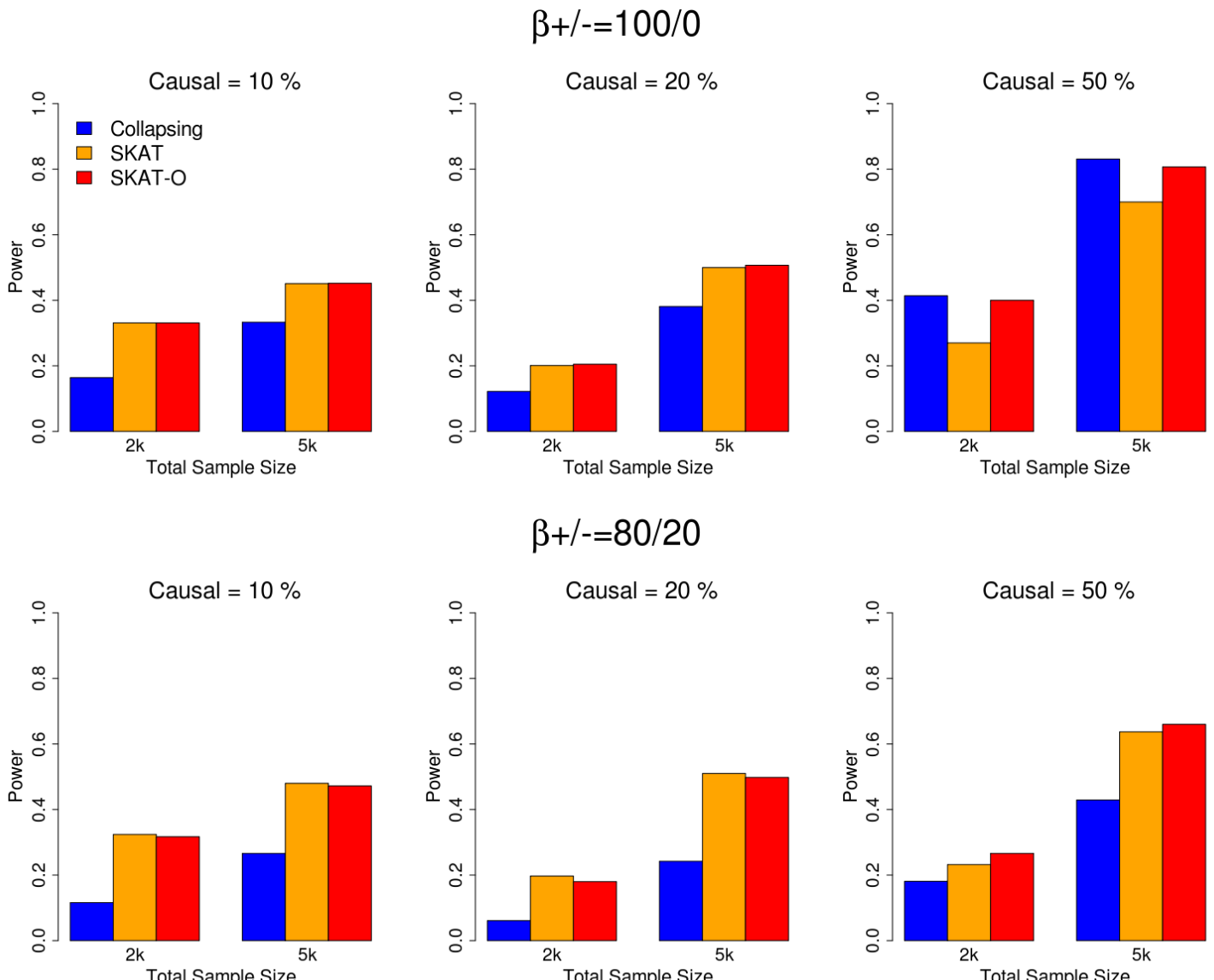
$$\begin{aligned}
&1 - Pr\left\{Q_{\rho_1} < q_{\rho_1}(T), \ldots, Q_{\rho_b} < q_{\rho_b}(T)\right\} \\
&= 1 - E\left[Pr\left\{(1-\rho_1)\kappa + a(\rho_1)\eta_0 < q_{\rho_1}(T), \ldots | \eta_0\right\}\right] \\
&= 1 - E\left[P\left\{\kappa < min\{(q_{\rho_v}(T)) - a(\rho_v)\eta_0)/(1-\rho_v)\} | \eta_0\right\}\right],
\end{aligned}$$

where $q_\rho(T) =$ quantile function of Q_ρ

Simulation

- Simulate sequencing data using COSI
- 3kb randomly selected regions.
- Percentages of causal variants = 10%, 20%, or 50%.
- $(\beta_j > 0)$% among causal variants = 100% or 80%.
- Three methods
  - Burden test with beta(1,25) weight
  - SKAT
  - SKAT-O

Simulation


β+/-=100/0
Causal = 10 %
Causal = 20 %
Causal = 50 %
Collapsing
SKAT
SKAT-O
Power
0.0
0.2
0.4
0.6
0.8
1.0
2k
5k
Total Sample Size
β+/-=80/20
Causal = 10 %
Causal = 20 %
Causal = 50 %

Simulation

- SKAT is more powerful than Burden test (Collapsing) when
  - Existence of $+/-\beta$s
  - Small percentage of variants are causal variants
- Burden test is more powerful than SKAT when
  - All βs were positive and a large proportion of variants were casual variants
- SKAT-O is robustly powerful under different scenarios.

Summary

- Region based tests can increase the power of rare variants analysis.
- Relative performance of rare variant tests depends on underlying disease models
- The combined test (omnibus test), e.g, SKAT-O, is robust and powerful in different scenarios

MAF based weighting

- It is generally assumed that rarer variants are more likely to be causal variants with larger effect sizes.
- Simple thresholding is widely used.

$$w(MAF_j) = \begin{cases} 1 & \text{if} \quad MAF_j < c \\ 0 & \text{if} \quad MAF_j \geq c \end{cases}$$

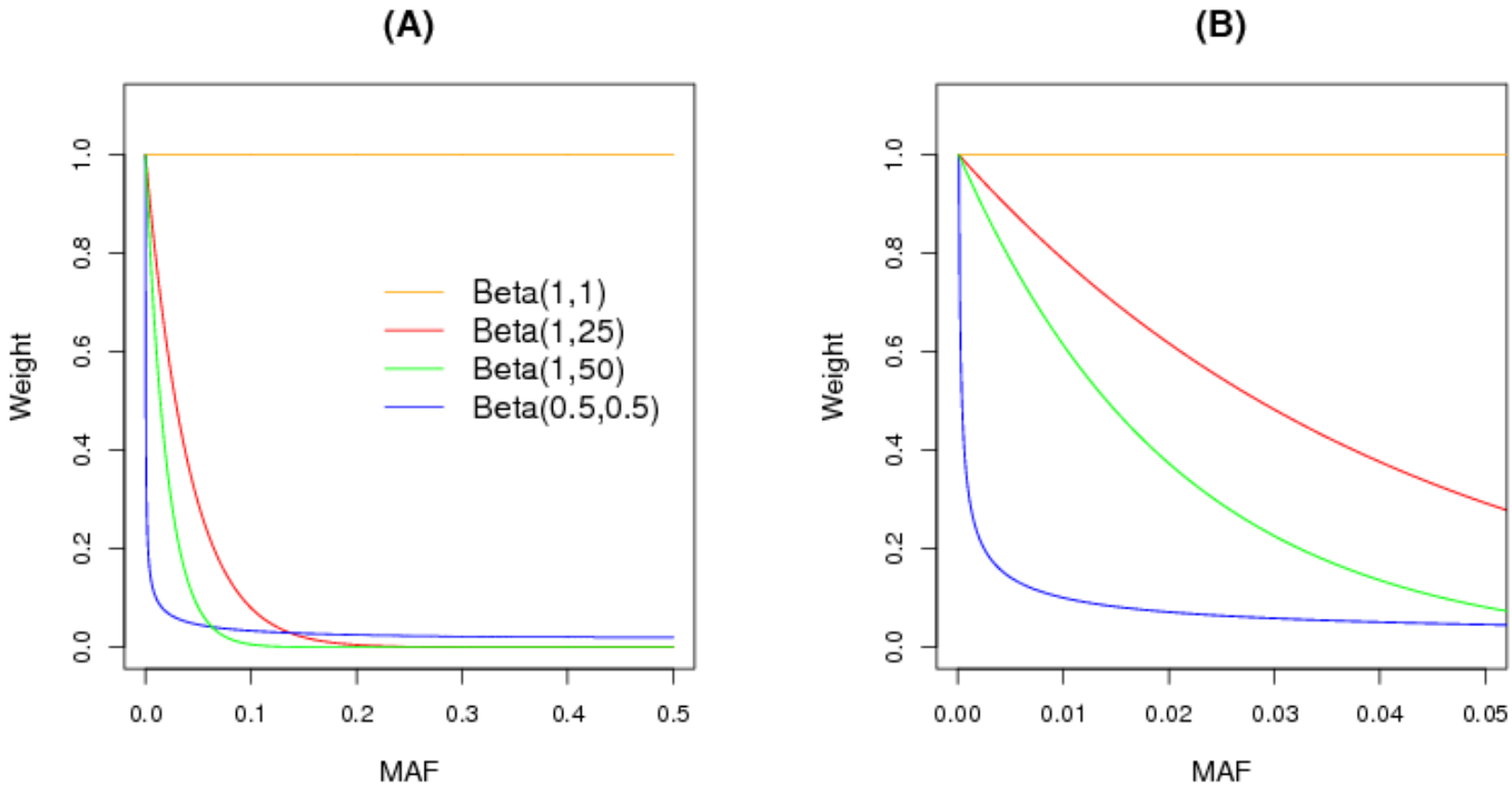
MAF based weighting

- Instead of thresholding, continuous weighting can be used to upweight rarer variants.
- Ex: Flexible beta density function.

$$w(MAF_j) = (MAF_j)^{\alpha-1}(1 - MAF_j)^{\beta-1}$$

  - $(\alpha = 0.5, \beta = 0.5)$: Madsen and Browning weight
  - $(\alpha = 1, \beta = 1)$: Flat weight

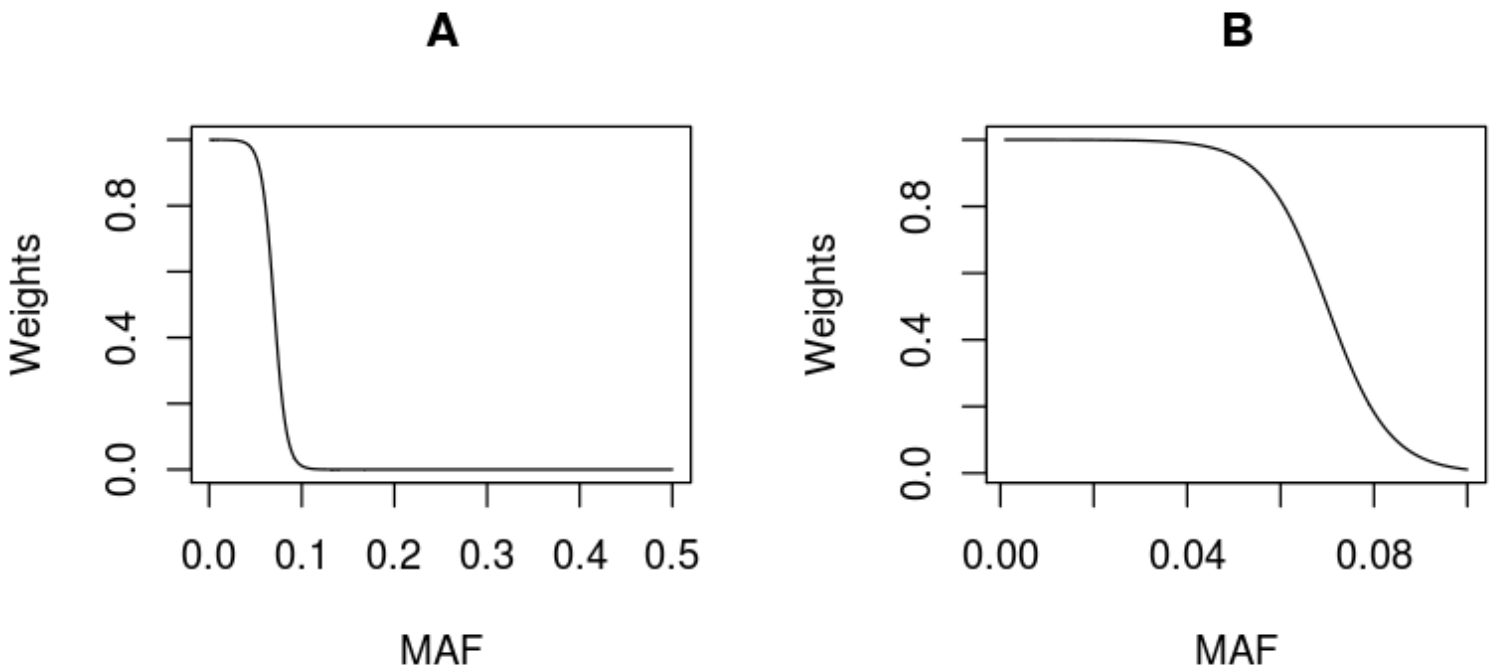
MAF based weighting- beta weight



MAF based weighting- logistic weight

- Soft-thresholding.

$$w(maf_j) = exp((\alpha - maf_j)\beta)/\{1 + exp((\alpha - maf_j)\beta\}$$



Weighting Using Functional information

- Variants have different functionalities.
  - Non-synonymous mutations (e.g. missense and nonsense mutations) change the amino-acid (AA) sequence.
  - Synonymous mutations do not change AA sequence.

Weighting Using Functional information

- Bioinformatic tools to predict the functionality of mutations.
  - Polyphen2 (http://genetics.bwh.harvard.edu/pph2/)
  - SIFT (http://sift.jcvi.org/)
- Test only functional mutations can increase the power.