Weighted multiple testing by convex optimization

Edgar Dobriban

Stanford/University of Pennsylvania

June 21, 2017

Overview

Background

Previous methods

New methods

Conclusion

Overview

Background

Previous methods

New methods

Conclusion

What I hope to accomplish in this talk

- Weighted testing has the potential to improve power and precision in modern data-centered research (e.g., genomics)
- However, current methods do not address imporant practitioner priorities
- Present Princessp, a weighting methodology using convex optimization

The potential of weighting

- Modern science: Test large number of hypotheses, only few of interest
  - genomics: millions of genetic variants $\rightarrow$ phenotype
  - neuroscience: stimulus $\rightarrow$ thousands of voxels
- State of the art: Treat all hypotheses equally
- Opportunity: Previous information
- Weighted testing: Improve power, precision

Example: Genome-Wide Association Studies

- Test associations between 500K-2million genetic variants (Single Nucleotide Polymorphisms SNPs) and phenotype (e.g., heart disease)
- Workhorse of genomics, hundreds of studies

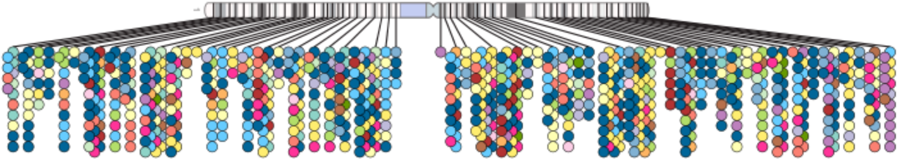


Figure : Chromosome 1 significant SNPs from GWAS catalog

Example: a Longevity GWAS. Fortney et al. (2015)

PLOS | GENETICS

Browse | Publish | About

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Genome-Wide Scan Informed by Age-Related Disease Identifies Loci for Exceptional Human Longevity

Kristen Fortney, Edgar Dobriban, Paolo Garagnani, Chiara Pirazzini, Daniela Monti, Daniela Mari, Gil Atzmon, Nir Barzilai, Claudio Franceschi, Art B. Owen, Stuart K. Kim

Published: December 17, 2015 • DOI: 10.1371/journal.pgen.1005728

Lessons learned from the longevity GWAS

- Weighting is useful in low-power settings
- Previous methods methods do not address practitioner priorities
- Need to develop new approaches

Overview

Background

Previous methods

New methods

Conclusion

Optimal weighting

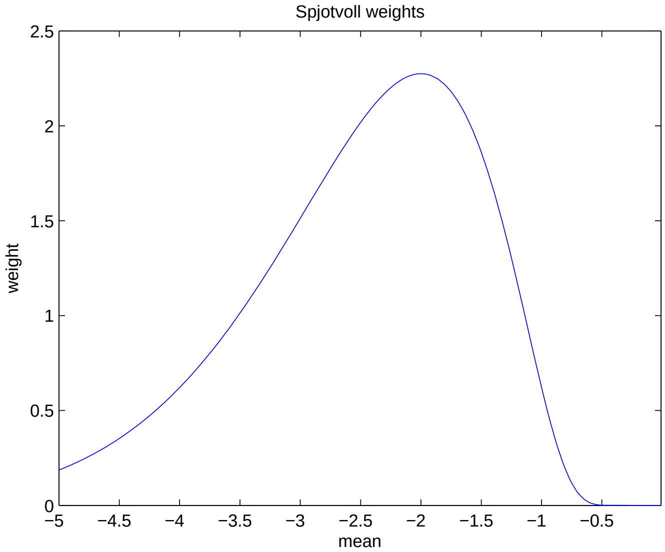
- Null hypothesis H_i: e.g., i-th SNP is not associated to longevity
- Weighted Bonferroni method:
  - reject H_i if $P_i \leq \alpha w_i$
  - weights $w_i \geq 0$, $\sum_{i=1}^J w_i = 1$
- Controls family-wise error rate (FWER), the probability of any errors
- Optimize expected number of discoveries (Spjøtvoll, 1972; Benjamini and Hochberg, 1997; Rubin et al., 2006; Roeder and Wasserman, 2009)

$$\max_{w \in \mathbb{R}^J} \sum_{i=1}^J \mathbb{P}_{H_i=1}(P_i \leq \alpha w_i)$$

$$\text{s.t. } w_i \geq 0, \sum_{i=1}^J w_i = 1.$$

Optimal weighting: Example

- Normal test $H_i : \mu_i = 0$ against $\mu_i < 0$. Plot $w_i = w(\mu_i)$



Practitioner priorities are not addressed

- Need not be monotone in the effect sizes μ_i
- Can be very close to zero—weighted p-values p_i/w_i are unstable
- Can't add further constraints (e.g., grouping)

Overview

Background

Previous methods

New methods

Conclusion

Princessp: New *princ*ipled *p*-value weights

- *Princessp*: new approach to weighted Bonferroni multiple testing
- Employs convex optimization
  1. Allows constraints—e.g., monotonicity, bounds
  2. Scales to very large problem sizes
- Key practitioner priorities are addressed

Princessp: New approach based on convex optimization

- $f_k : \mathbb{R}^J \to \mathbb{R}$, $k = 1, \dots, K$ convex, smooth.

$$\begin{aligned} \max_{w \in \mathbb{R}^J} \quad & \sum_{i=1}^{J} \mathbb{P}_{H_i=1}(P_i \le \alpha w_i) \\ \text{s.t.} \quad & f_k(w) \le 0, \quad k = 1, \dots, K \\ & Aw = b. \end{aligned}$$

- Key: ROC curve $\alpha \to \mathbb{P}_{H=1}(P \le \alpha)$ is often concave
- Previous formulations use critical values, non-convex (Westfall et al., 1998; Rubin et al., 2006; Roeder and Wasserman, 2009)
- The first time convexity is exploited for weighting

Well-known concave ROC curves

- One-sided tests in monotone likelihood ratio (MLR) families (e.g., Lehmann and Romano, 2005, p.101)
  - Natural param of continuous 1D exponential family
  - Non-centrality param in t, F, χ^2 distribution
- Princessp applicable

New results on concave ROC curves

Two-sided tests in continuous 1D exponential families.

- $T \sim \exp(\theta t - A(\theta))d\mu(t)$. Test $\theta = \theta_0$ against $\theta \neq \theta_0$. Reject if $|T - \theta_0| > c$.

Proposition

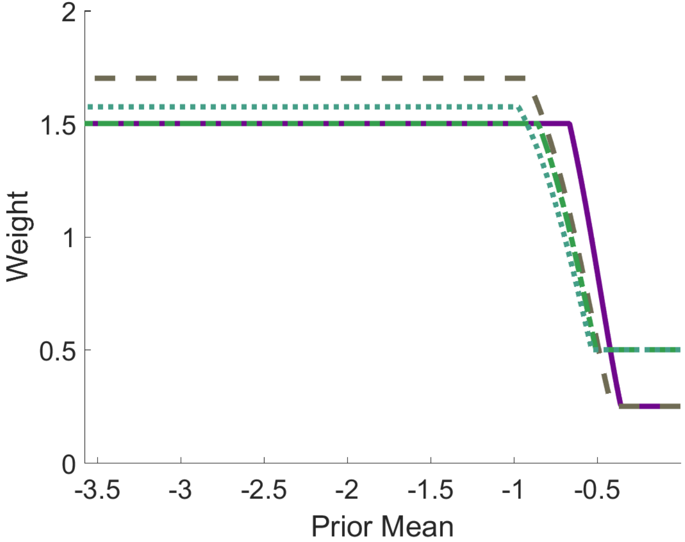
1. *If* $|\theta_1| > |\theta_0|$, *the ROC curve at* θ_1 *is concave.*
2. *If* $|\theta_1| > 2|\theta_0|$, *the ROC curve at* θ_1 *is strongly concave on* $(0, 1-\varepsilon)$, $\varepsilon > 0$.

(Strong concavity needed for numerics.)

Monotone bounded weights in Princessp framework

- $w_i \leq w_j$ if $|\mu_i| \leq |\mu_j|$
- $0 < l \leq w_i \leq u$
- In Dobriban (2016)
  1. Large-scale algorithm (subsampled interior point)
  2. GWAS example (weights work well)

Monotone bounded weights. $w_i = w(\mu_i)$



Overview

Background

Previous methods

New methods

Conclusion

Conclusion

- *Princessp*, a flexible weighting methodology using convex optimization
- Addresses practitioner priorities: constraints, large-scale applications
- Potential to improve power and precision

E. Dobriban, *Weighted mining of massive collections of p-values by convex optimization*, arXiv:1603.05334

Andreassen, O. A., Djurovic, S., Thompson, W. K., Schork, A. J., Kendler, K. S., ODonovan, M. C., Rujescu, D., Werge, T., van de Bunt, M., Morris, A. P., et al. (2013). Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *The American Journal of Human Genetics*, 92(2):197–209.

Benjamini, Y. and Hochberg, Y. (1997). Multiple hypotheses testing with weights. *Scandinavian Journal of Statistics*, 24(3):407–418.

Dobriban, E. (2016). Weighted mining of massive collections of p-values by convex optimization. *arXiv preprint arXiv:1603.05334*.

Dobriban, E., Fortney, K., Kim, S. K., and Owen, A. B. (2015). Optimal multiple testing under a Gaussian prior on the effect sizes. *Biometrika*, 102(4):753–766.

Fortney, K., Dobriban, E., Garagnani, P., Pirazzini, C., Monti, D., Mari, D., Atzmon, G., Barzilai, N., Franceschi, C., Owen, A. B., and Kim, S. K. (2015). Genome-wide scan informed by age-related disease identifies loci for exceptional human longevity. *PLoS Genet*, 11(12):e1005728.

Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses*. Springer Science & Business Media.

Li, L., Kabesch, M., Bouzigon, E., Demenais, F., Farrall, M., Moffatt, M. F., Lin, X., and Liang, L. (2013). Using eqtl weights to improve power for genome-wide association studies: a genetic study of childhood asthma. *Frontiers in Genetics*, 4(103).

Rietveld, C. A., Esko, T., Davies, G., Pers, T. H., Turley, P., Benyamin, B., Chabris, C. F., Emilsson, V., Johnson, A. D., Lee, J. J., et al. (2014). Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proceedings of the National Academy of Sciences*, 111(38):13790–13794.

Roeder, K. and Wasserman, L. (2009). Genome-wide significance levels and weighted hypothesis testing. *Statistical Science*, 24(4):398–413.

Rubin, D., Dudoit, S., and Van der Laan, M. (2006). A method to increase the power of multiple testing procedures through sample splitting. *Statistical Applications in Genetics and Molecular Biology*, 5(1):1–19.

Saccone, S. F., Hinrichs, A. L., Saccone, N. L., Chase, G. A., Konvicka, K., Madden, P. A., Breslau, N., Johnson, E. O., Hatsukami, D., Pomerleau, O., et al. (2007). Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 snps. *Human Molecular Genetics*, 16(1):36–49.

Spjøtvoll, E. (1972). On the optimality of some multiple comparison procedures. *The Annals of Mathematical Statistics*, 43(2):398–411.

Sveinbjornsson, G., Albrechtsen, A., Zink, F., Gudjonsson, S. A., Oddson, A., Másson, G., Holm, H., Kong, A., Thorsteinsdottir, U., Sulem, P., et al. (2016). Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nature Genetics*, 48:314–317.

Westfall, P. H., Krishen, A., and Young, S. S. (1998). Using prior information to allocate significance levels for multiple endpoints. *Statistics in Medicine*, 17(18):2107–2119.