A Simulation Study to Compare Gene Set Analysis Methods

Andrew James Pfeiffer

Thesis submitted for the degree of

Master of Philosophy

in

Statistics

at

The University of Adelaide

(School of Mathematical Sciences)

May 19, 2016

Contents

List of Tables

List of Figures

List of Acronyms

ALIGATOR Association LIst Go AnnoTatOR.

AUC area under the ROC curve.

BH Benjamini-Hochberg.

bp base pairs.

BY Benjamini-Yekutieli.

CATT Cochran-Armitage Trend Test.

DAVID the Database for Annotation, Visualization and Integrated Discovery.

dbSNP the Database of Short Genetic Variations.

DNA deoxyribonucleic acid.

DSL disease susceptibility locus.

EASE Expression Analysis Systematic Explorer.

ES enrichment score.

EVA Exploratory Visual Analysis.

FDR false discovery rate.

FET Fisher's exact test.

FPR false positive rate.

FWER family-wise error rate.

GLM general linear model.

GO gene ontology.

GSA gene set analysis.

GSE gene set enrichment.

GSEA Gene Set Enrichment Analysis.

GSEAPR GSEAPreranked.

GWA study genome-wide association study.

HapMap CEU Data genomic data collected from CEPH (Utah residents with ancestry from northern and western Europe) (abbreviation: CEU) as part of the International HapMap Project.

HWE Hardy-Weinberg Equilibrium.

IPA QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, `www.quiagen.com/ingenuity`).

KEGG Kyoto Encyclopaedia of Genes and Genomes.

LD linkage disequilibrium.

LE linkage equilibrium.

LS Li and Stephens.

MGSEA Modified GSEA.

MPEVA MinP-EVA.

NCBI the National Center for Biotechnology Information.

NES normalised enrichment score.

OR overrepresentation.

PANTHER Protein ANalysis THrough Evolutionary Relationships.

PARIS Pathway Analysis by Randomization Incorporating Structure.

RNA ribonucleic acid.

ROC receiver operating characteristic.

rs reference SNP.

SNP single nucleotide polymorphism.

the SRT the SNP Ratio Test.

TPR true positive rate.

Abstract

Genome-wide association studies (GWA studies) identify alleles that are associated with a disease. These allele variations are called single nucleotide polymorphisms (SNPs). However, GWA studies do not account for interaction between SNPs. Gene set analysis (GSA) is used in GWA studies to account for interaction. GSA methods map SNPs to gene sets and identify gene sets that are associated with a disease. Comprehensive reviews of GSA exist in the literature. However, these reviews do not compare specific methods or implement them on data.

In this thesis, we compare six GSA methods. We use seven factors highlighted by the reviews as important in GSA to compare these methods. For example, we analyse how each method accounts for parameters that could affect the analysis. These parameters include gene size and SNP interaction. We consider the null hypothesis tested by each method. We also analyse the sensitivity of methods to individual SNPs with small p-values. In contrast, the marginal effect of many SNPs that cause diseases is often small. The p-values of such SNPs need not be small.

We conduct a simulation study to compare four GSA methods. We investigate the sensitivity of these methods to SNPs with very small p-values. We use Manhattan plots to display gene sets that were assigned disparate p-values by different methods. We also use receiver operating characteristic curves to compare the performance of each method. Finally, we recommend a method that gave excellent performance.

Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

SIGNED: ..... DATE: .....

To my Father God, my Lord Jesus Christ, and the Holy Spirit. Thank you for foreknowing me, predestining me to be conformed to the image of Jesus, calling me, justifying me and glorifying me (Romans 8:29,30).

Acknowledgements

I begin by thanking my supervisors, Associate Professor Gary Glonek and Dr Jonathan Tuke. Gary and Jonathan have patiently taught me so much over the last two years. Of course, they have taught me so much about statistics. Yet even more importantly, they have also taught me many skills that I will continue to use, even when I am no longer using statistics in my career. Soft skills such as writing well, managing time well, being organised and communicating information clearly to others are not skills that students are taught explicitly. Yet they are skills that need to be learned to make the most of the short time that each of us has on this earth. I thank Gary and Jonathan for helping me to develop in these skills. Whether it be through taking notes and emailing them out after every meeting with a list of action points, learning how to use tools such as Evernote, OmniFocus and Sublime Text or growing in my understanding of programming languages such as LaTeX, Python, R and shell scripting, I have learned so much from both of you.

I especially thank Jonathan for rekindling my love for statistics. After completing Year 12 Mathematical Studies in 2006, I told myself that I would never study statistics again. Nonetheless, I remember going to lectures for Probability and Statistics II in 2011, feeling Jonathan's excitement for statistics, and being encouraged by the genuine care he had for students. He even gave students free lollipops when they came to his office to ask him a question. If it weren't for Jonathan, I would not be finishing a Master of Philosophy in statistics.

To my friends in the postgraduate room, thank you. The life of a postgraduate student is not always easy, but it has been both comforting and encouraging to share parts of this crazy ride with you.

To everyone at Adelaide University Evangelical Students (ES), thank you so much. ES is a Christian group on campus, passionate about sharing the good news of Jesus Christ with everyone, and God has used it to completely change my life while I have been at university. I recently attended the Leavers' Day and Dinner; we were given the opportunity to share something that we wanted to praise God for about ES, and some advice we'd like to offer continuing members of ES. And like many over-analytic perfectionists, I kept wondering if I said the right things. I posted the following on Facebook the morning after the dinner, and I repeat it here.

> First, I praise God for the staffworkers: Geoff, Reuben, Laura, Emily, Matthew, Mark, Oliver and Dave. As I am beginning to understand, vocational ministry involves many, many sacrifices, including a two-year apprenticeship (for many), three or more years at Bible college, and foregoing a secular career that would have probably paid considerably better. And yet, you've chosen the path of vocational ministry: you've chosen to share your lives with us, mentor us, teach us the Bible, help us to read it for ourselves, help us to teach it to other people, help us read it with our not-yet-Christian friends, train us in evangelism, and train us in leadership in other ways, whether it be by leading teams or serving on committees. I am so immensely thankful.
>
> For indeed, God created us to share the good news of Jesus Christ with other people wherever we go. As a uni student, one of my primary tasks has been to declare the praises of God and what he has done on the university campus - it's been my mission field. And the staffworkers have helped to train us up

to be missionaries on the university campus. But the awesome thing is, the way that they have trained us in Bible reading, evangelism, and serving the kingdom in other ways is not just limited to the university campus. We'll be able to use everything that they have taught us wherever we go in life: whether the workplace, vocational ministry, cross-cultural mission, and so on. In this sense, among many others, God has used the staffworkers, through ES, to change my life completely.

So how could I not want to get involved with such a brilliant group of people, whose vision is "reaching every uni student on North Terrace with the good news of Jesus Christ to present everyone mature in Him." This vision is the vision all of God's people need to have, just applied to the context of university. Hallelujah!! I'd also like to praise God for the staffworkers because they see people who are keen to be involved with ES, and give them opportunities to serve! I'm so grateful for the opportunity to lead Bible study groups for the last three years, lead the Bible talks team last year, and be the treasurer of ES this year. Bible study leading has stretched me, from the time when I nervously led my first BSG in March 2013 (just ask Tom Worley how nervous I was). It has given me an awesome opportunity to grow in my understanding of the Bible, a passion to teach it to other people, and a desire to mentor younger brothers in Christ in the faith. Leading the Bible Talks team and being Treasurer have been awesome opportunities to develop my administration and behind-the-scenes ministry skills, and to give back to ES in an awesome way. To those of you who, like me, are rather freaked out by things like walk-up evangelism, please don't feel like an inferior Christian. Of course, give walk-up a shot, because God can use even the most nervous Christian to bring people into the

kingdom. But know that serving in ministry behind the scenes is such a necessary part of ministry, to facilitate reaching the university campus with the gospel message.

So to the staffworkers, thank you for sharing your lives with us, for helping us to grow in our relationship with God in such a life-changing way, and for giving us opportunities to serve in ES.

Next, I praise God for my fellow Evangelical Students. (I don't understand why we are called ESers, because it doesn't make sense when the acronym is expanded, or ES students, because it's like "ATM machine" or "PIN number"). That's beside the point, of course. To my fellow ES members: thank you for being such welcoming and encouraging brothers and sisters in Christ. As someone who was bullied in primary school, high school, and the early parts of university, I cannot thank my fellow ES members enough for welcoming me and accepting me for who I am. It has been an honour, privilege and joy to be fellow soldiers for Christ with you all on campus. Whether it be sharing the highs and lows of evangelism, having deep theological or personal conversations (often on camps, but also on the university campus), or just chilling out and doing nothing with you, thank you. Thank you for rebuking me and pulling me up where necessary, for speaking the truth to me in love, and for being brilliant brothers and sisters in Christ. I have to say that I met most of my best friends at ES, and even though I am going into the workforce next year, God willing, I truly hope to stay in touch with you.

In particular, I'd like to thank Tom, Chloe and Hendré for serving on Exec with me this year. It's been a privilege to serve ES with you in this way. Tom and I were chatting last night

about how we worked so well as a team, and how everyone played their part and did what they needed to do. To see you all work so hard in ES on a volunteer basis is a huge testament to your passion for seeing God work on campus. Also, thanks to Chloe, Claire, Hayden and Lachlan, for being willing to serve on Exec next year. I know that under God, you will serve ES faithfully in your respective roles. To Irene, as last year's Treasurer, thank you for teaching me the ropes of treasury this year.

To Hayden, Daniel and George, thank you for serving with me on the Bible talks team last year. It was the first time ES ran the Bible talks team, and you all did what needed to be done exceptionally well. And thanks to Hayden for running the team so well this year - the fact that your work went unnoticed on the whole is a testament to how smoothly you ran everything. To my fellow Bible Study group leaders: Tom, Stephen, Sowmya and Renee (I apologise if I forgot anyone) - thank you for imparting your leadership wisdom as I was learning how to lead BSGs, and for allowing me to do the same. To the people that led O'Weeks, camps and Jesus Weeks over the last few years: Jane, Samantha, David, Jack, Shane, Ellen, Tom, Lachlan, Warwick and Timothy - thank you for all the hard work you put in to make these brilliant events happen for all of us. To the Exec committees over the last few years: Nicholas, Michael, Sasha, Kathryn, Ada, Sarah, Brendan, Liam, Irene, Madeleine and Ryan, thank you for your hard work, too. To the Faculty Leaders this year: Ryan, Naomi, Jack, Lydia, Liam and Bonnie, thank you for the work you put in this year to get the new Faculty system up and running. I know that it has been worth it, and it will continue to reap fruit. And to next year's faculty leaders: Hendre, Sowmya, Oli, Corinne and Warwick, I know that you will do everything to make new

people feel welcome at ES in your respective faculties. To Claire, Maggie, Declan and Phil: thank you for running the finances for the various events this year, and making my life as treasurer so much easier. To John, Lisa, Daniel, Abe and everyone else who came to the maths prayer group last year - thank you for joining me. It was a privilege to spend time with you on Thursday mornings in prayer to our awesome God. I still remember our awesome breakfast catchup at Aroma café at the end of last year.

Finally, to all my friends at ES. Thank you. Thank you for being some of the best friends I have ever had. ES has completely changed my life, and you have been a huge part of that, so thank you.

If you've read this far, then thank you. For those of you who are commencing university studies next year, or continuing members of ES, I encourage you to *get involved*. It's true, everyone says it. But the fact is, the period of time you have at university is brief, compared to the rest of your life. (Well, maybe not so brief for me). I doubt you will ever get another opportunity to be stretched and grown in your faith like you will at ES, or opportunities to learn how to evangelise. So try everything. If you're asked to be on a team, or lead a team, or do something for ES, say yes. Unless God tells you to say no, of course. Try doing walk-up evangelism, even if it freaks you out. Try helping to run a camp, or an event, or anything. And if you'd like to be involved more, just chat to a staffworker. They'd love to catch up with you for coffee and chat to you about how you can be more involved with ES. And even if you don't think you have time, *make time*. Like I said, it's a brief period of your life. Keep serving at your local church, of course, because that should be

the number one place you serve. But beyond that, serve as much as you possibly can with ES. You will *not* regret it.

One of the reasons that I love ES so much is that it is such a strategic ministry. Thousands and thousands of (mostly) young people are going onto campus to learn about various subjects. And it is during young adulthood when people are thinking for themselves about what it is they are going to believe, including strong Christians, nominal Christians who just go to church each Sunday because their parents do, and not-yet-Christians. The ministry that ES do can make the most of this melting pot of worldviews at university by presenting the gospel message. And if you don't know what the gospel message (the good news) is, then keep reading. Soon after God created mankind, we rebelled against him. We deserve God's judgment. But in his infinite mercy, God sent his son, Jesus Christ, into the world, to take the punishment for people's sins, so that anyone who believes in him can be in right relationship with God for all eternity. If you don't believe this good news yet, but you'd like to find out more, please just ask me.

Also, to those of you who think that ES is a bunch of weirdos, or are put off because you think we're pushy about the gospel message, please reconsider your scepticism. We just want everyone to know the good news about Jesus Christ, and we don't want anyone to fall under God's judgment. Forgive us for the times when we have been unloving in our portrayal of the gospel message, but please understand that we only share the gospel message because we love you.

Finally, to those of you who resonate with what I've said, and who are passionate about the ministry of ES, one more piece

> of advice. Seriously consider serving ES in full-time vocational ministry. You know how God has used the staffworkers in ES to change our lives, and it would be an amazing privilege to be used by God to change the lives of the next generation of university students. If you'd like to consider this challenge more, chat to one of the staffworkers about it, chat to me about it, or come along to CV Conference in 2016. Ultimately, nothing is more important than sharing the gospel message with everyone that does not know Jesus, and discipling those that do know Jesus, so that they can disciple others.
>
> In one sense, I'm sad that my time at ES has drawn to a close. Nonetheless, I know that now is the right time for this season, amazing as it has been, to come to an end. For the next few years, I will be working with the ATO in Adelaide. But soon, God willing, I will come back to serve ES as a ministry apprentice. And then, after getting some formal theological education and training at Bible College, I will come back to serve ES full-time, for as long as I am physically able. And I can't wait. Bring. It. On. Praise God for everything.

To my parents, thank you for so much. Thank you, first and foremost, for raising me to know my Lord, Jesus Christ. Thank you for supporting me throughout my university studies. Thank you for providing so much for me. Thank you for loving me so much. Thank you also to my mother for proofreading my thesis and finding approximately 15 typographical errors.

Finally, to my Father God, my Lord Jesus Christ, and the Holy Spirit. Thank you for foreknowing me, predestining me to be conformed to the image of Jesus, calling me, justifying me and glorifying me (Romans 8:29,30). Only by your grace do I exist; only by your grace do I have the skills that I needed to complete this thesis, and only by your grace did Jesus die for me, so that I can be in right

relationship with you. Please help me to keep giving my life to you, for your glory. Amen.