Chromatin architecture aberrations in prostate cancer and leukemia

by

James Hawley

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Graduate Department of Medical Biophysics
University of Toronto

Contents

Chapter 1

Hedging uncertainty in differential gene expression analyses with James-Stein estimators

J.R.H., and M.L. conceptualized the study. J.R.H. derived the statistical estimates and designed and conducted all the experiments. Figures were designed by J.R.H. The manuscript was written by J.H., and M.L. M.L. oversaw the study.

1.1 Abstract

1.2 Introduction

1.3 Results

- the two main approaches for reducing error in a model are to reduce the model variance or model bias - Figure 1.1 - here we attempt to decrease mean square error (MSE) by simultaneously increasing the bias and decreasing the variance in fold change coefficient estimators - derivation for the James-Stein (JS) estimator can be found in Appendix A.3 - Equation for the JS estimator can be seen in Figure 1.1b. - in theory, this may increase the error of some transcripts, but will decrease MSE for a set of transcripts in aggregate - Appendix A.5

- using RNA-seq data from a highly replicated yeast knockout (KO) experiment, we compared the

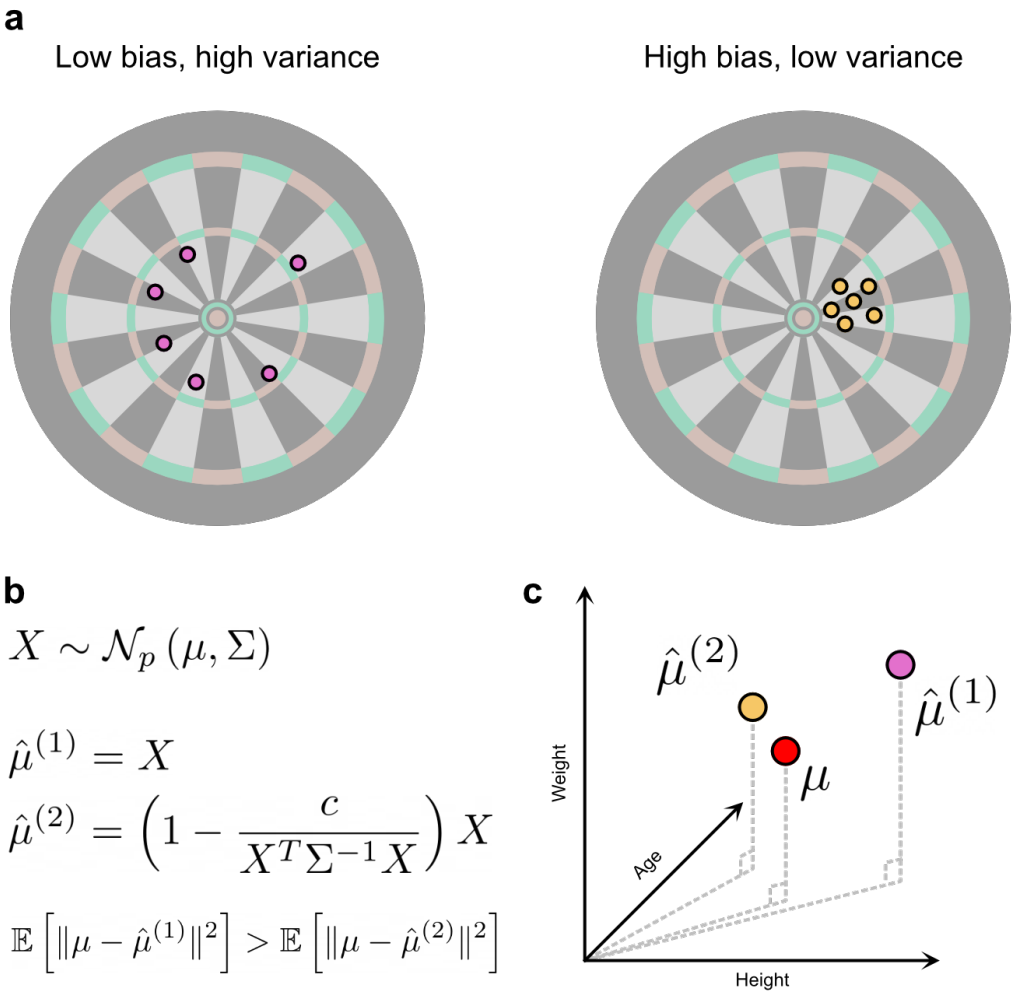

Figure 1.1: **Reducing the bias-variance tradeoff by combining information across multiple features. a.** Schematic of the bias-variance tradeoff for assessing model performance. Dartboard on the left shows low bias of the darts (mean is close to the bullseye) but a large variance. Dartboard on the right shows a high bias of the darts (mean is off-centre), but a small variance. **b.** For a p-variate normal distribution from which a single observation is made, the naive estimator has a higher MSE than the JS estimator, defined as $\hat{\mu}^{(2)}$. **c.** An analogy showing how the JS estimators work in theory. Trying to estimate the mean height, weight, and age for the entire population (μ) from a single person will give an estimate that is likely wrong ($\hat{\mu}^{(1)}$). Combining information from the three variables together can produce an estimate that is closer to the truth ($\hat{\mu}^{(2)}$).

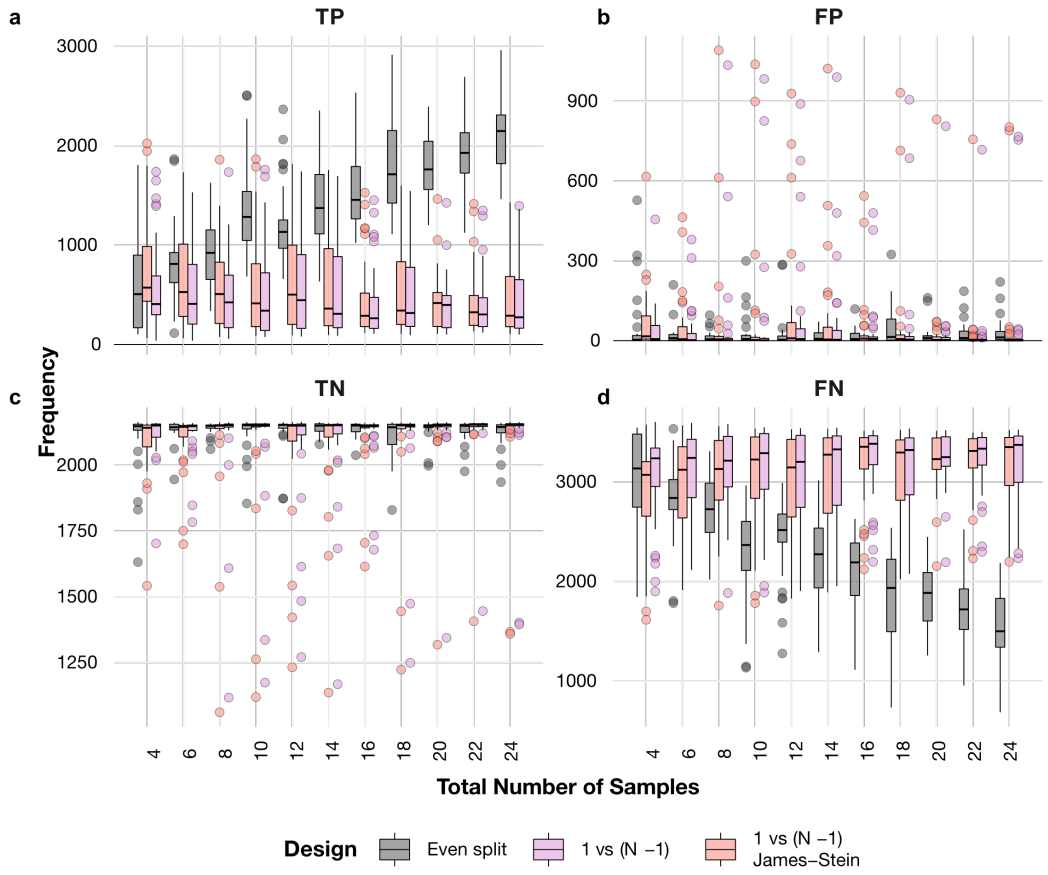

Figure 1.2: **Differential gene expression analysis of the entire yeast transcriptome with differently sized experimental designs.** Simulations ($n = 30$) using randomly selected samples which were then compared to the full dataset of 48 Δ Snf2 vs 48 WT to calculate TP (**a**), FP (**b**), TN (**c**), and FN (**d**).

statistical inference from differential gene expression analysis when using the ordinary least squares (OLS) and JS estimators - found that for small sample sizes ($n = 4$), the JS estimators predict the largest number of true positive (TP) and false positive (FP) as well as the smallest number of true negative (TN) and false negative (FN) on average - Student's unpaired t-test, $n = 30$, $p =$ - for larger sample sizes ($n \geq 6$), the JS estimators continue to identify more TP and FP, as well as fewer TN and FN than the OLS estimator, on average - these effects lessen with increasing sample size, as expected - both the OLS and JS methods with an unbalanced experimental design perform more poorly than the OLS method with a balanced experimental design, as expected

- to investigate where these changes in fold change estimates and p-values originates from, we investigate representative simulations - most genes in the Δ Snf2 model are under-expressed compared to the WT model

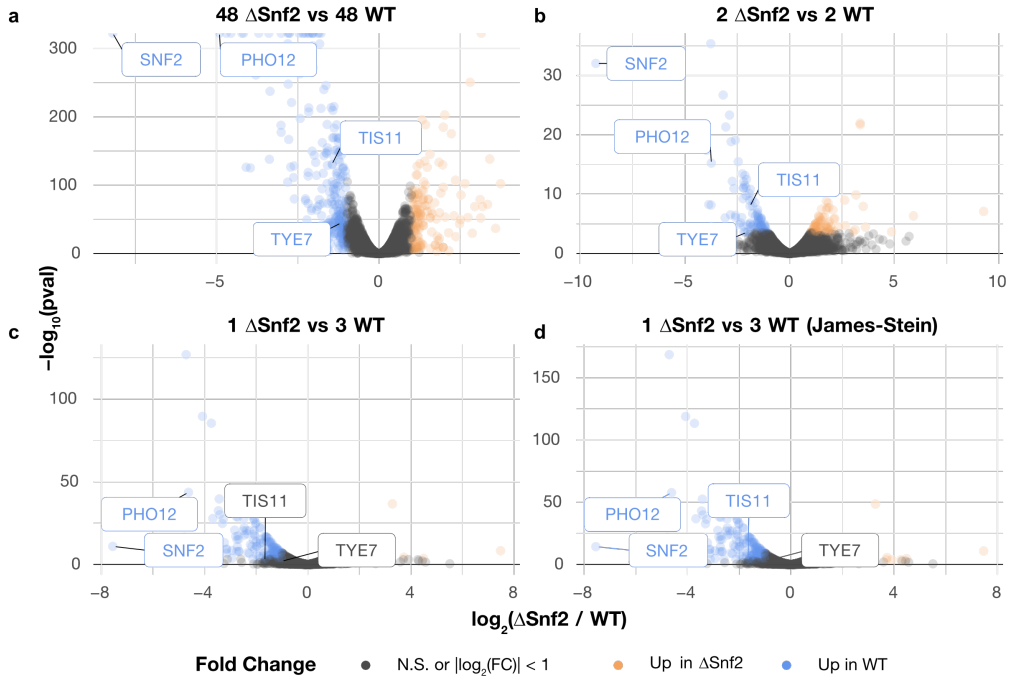

Figure 1.3: **Differential gene expression analysis of Δ Snf2 vs WT yeast cells using different sample sizes and experimental designs. a.** Volcano plot of differential expression results with OLS estimates in a highly replicated experiment consisting of 48 biological replicates of each condition. **b.** The same analysis as (**a**) using 4 samples in total, 2 Δ Snf2 and 2 WT samples. **c.** The same analysis as (**c**) using 1 Δ Snf2 and 3 WT. **d.** The same analysis as (**c**) using the JS method instead of OLS.

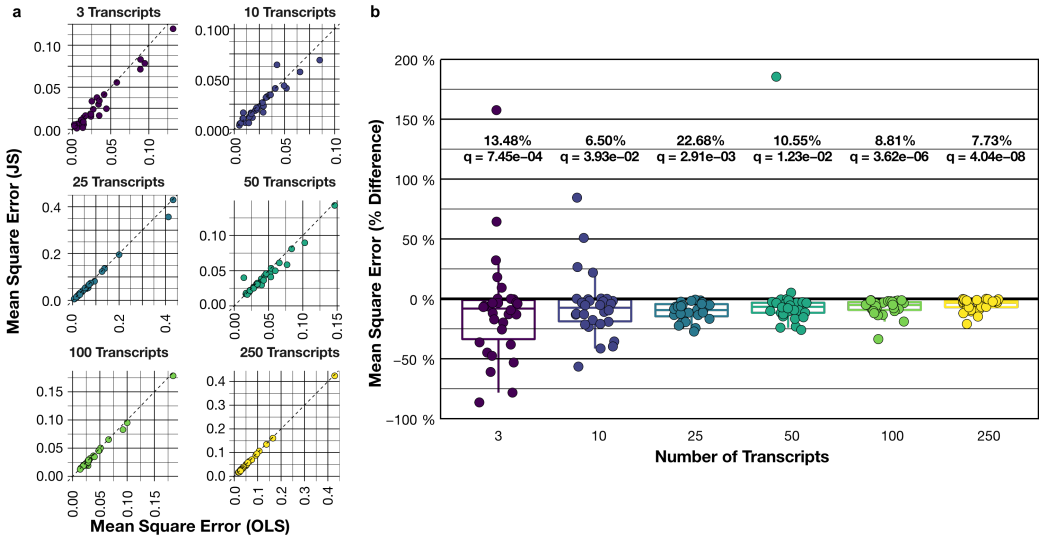

Figure 1.4: **Differential gene expression analysis focusing on a subset of trancsripts, not the entire transcriptome.** All experiments use 1 Δ Snf2 vs 5 WT samples (or vice versa). **a.** Comparison of the MSE of the JS estimates (y-axis) against the OLS estimates (x-axis). The total number of transcripts in each comparison is specified above each facet. **b.** Percent different in MSE between the JS and OLS estimates. One-sided, two-sample Student's t-test, $n = 30$, FDR multiple test corrections.

1.4 Discussion

1.5 Methods

Appendix A

Supplementary Material for Chapter 4

A.1 Differential expression analysis with Sleuth

The differential expression model employed in the Sleuth (v0.30.0) [1, 2] can be described as follows. Consider a set of transcripts, S, measured in N samples with an experimental design matrix, $X \in \mathbb{R}^{N \times p}$, where p is the number of covariates considered. Let Y_{si} be the natural log of the abundance of transcript s in sample i. Given the design matrix

$$X = [x_1^T; x_2^T; ...x_n^T], x_i \in \mathbb{R}^p$$

the abundance of transcripts can be modelled as a generalized linear model (GLM)

$$Y_{si} = x_i^T \beta_s + \epsilon_{si} \tag{A.1}$$

where $\epsilon_{si} \sim \mathcal{N}(0, \sigma_s^2)$ is the biological noise of transcript s in sample i and $B_s \in \mathbb{R}^p$ is the fixed effect of the covariates on the expression of transcript s.

Due to inferential noise from sequencing, each Y_{si} are not observed directly, but indirectly through the observed perturbations, D_{si}. This can be modelled as

$$D_{si}|Y_{si} = Y_{si} + \zeta_{si} \tag{A.2}$$

where $\zeta_{si} \sim \mathcal{N}(0, \tau_s^2)$ is the inferential noise of transcript s in sample i. Both biological and inferential noise for each transcript are independent and identically distributed (IID) and independent of each other. Namely:

$$\begin{aligned}
\mathbb{C}ov[\epsilon_{si}, \epsilon_{rj}] &= \sigma_s^2 \delta_{i,j} \delta_{s,r} \\
\mathbb{C}ov[\zeta_{si}, \zeta_{rj}] &= \tau_s^2 \delta_{i,j} \delta_{s,r} \\
\mathbb{C}ov[\epsilon_{si}, \zeta_{rj}] &= 0 \\
&\forall s, r \forall i, j
\end{aligned}$$

The abundances for transcript s in all N samples can then modelled as a multivariate normal distribution

$$D_s | Y_s \sim \mathcal{N}_N(X\beta_s, (\sigma_s^2 + \tau_s^2) I_N) \tag{A.3}$$

where $I_N \in \mathbb{R}^{N \times N}$ is the identity matrix.

The goal of the differential analysis is to estimate the $|S| \times p$ coefficients in $B_s \forall s \in S$, and to determine which coefficients differ significantly from 0. This is achieved through a Wald test or likelihood ratio test after estimating the inferential variance, τ_s^2, through bootstrapping and the biological variance, σ_s^2, through dispersion estimation and shrinkage.

The estimator for the differential effect is the OLS estimate:

$$\hat{\beta}_s = (X^T X)^{-1} X^T d_s$$

where d_s is the observed abundances given by

$$\begin{aligned}
d_{si} &= \ln\left(\frac{k_{si}}{\hat{f}_i} + 0.5\right) \\
\hat{f}_i &= \underset{s \in S^*}{\text{median}} \frac{k_{si}}{\sqrt[N]{\prod_{j=1}^{N} k_{sj}}}
\end{aligned}$$

where k_{si} is the estimated read count from the Kallisto package (v0.46.1) [3] for transcript s in

sample i and $\hat{f}_i$ is the scaling factor for sample i, calculated from the set of all transcripts that pass initial filtering, S^*.

A.2 Statistical moments of the ordinary least squares estimator

As shown in Supplementary Note 2 of [REF 1], the estimator is unbiased, Namely

$$\mathbb{E}\left[\hat{\beta}_s^{(OLS)}\right] = B_s \tag{A.4}$$

It can also be shown that, for a covariance matrix Σ,

$$\mathbb{V}\left[\hat{\beta}_s^{(OLS)}\right] = (X^T X)^{-1} X^T \Sigma X (X^T X)^{-1}$$

In the case where $\Sigma = (\sigma_s^2 + \tau_s^2) I_N$, this reduces to

$$\mathbb{V}\left[\hat{\beta}_s{}^{(OLS)}\right] = (\sigma_s^2 + \tau_s^2)(X^T X)^{-1}$$

Consider a simple experimental design where the only covariate of interest is the presence of a mutation. Then the design matrix, with the first column being the intercept and the second being the mutation status, looks like so:

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{(N+1)\times 2}$$

The variance of the OLS estimator is then

$$\mathbb{V}\left[\hat{\beta}_s{}^{(OLS)}\right] = \frac{(\sigma_s^2 + \tau_s^2)}{n_{mut} n_{nonmut}} \begin{bmatrix} n_{mut} & -n_{mut} \\ -n_{mut} & n_{mut} + n_{nonmut} \end{bmatrix}$$

Importantly, the estimate for the coefficient measuring the effect that the presence of the mutation has variance

$$\mathbb{V}\left[\beta_{s,mut}^{(OLS)}\right] = \frac{(\sigma_s^2 + \tau_s^2)(n_{mut} + n_{nonmut})}{n_{mut} n_{nonmut}}$$

When there is only 1 mutated sample, as per the motivation of this work, this reduces to

$$\mathbb{V}\left[\beta_{s,mut}^{(OLS)}\right] = \frac{(\sigma_s^2 + \tau_s^2)(1 + n_{nonmut})}{n_{nonmut}} \tag{A.5}$$

A.3 Derivation of the James-Stein estimator

For a p-variate normal distribution $Z \sim \mathcal{N}_p(\mu, \Sigma)$ where μ is unknown and Σ is known, if we observe a single realization of this distribution, z, we have the following theorem [4]:

Theorem 1 *The estimator $\hat{\mu}^{(0)} = z$, for any mean μ, does not minimize the MSE $\mathbb{E}\left[(\mu - \hat{\mu})^2\right]$ in the case that $m \geq 3$ and $\Sigma = I_N$. Namely, the estimator $\hat{\mu}^{(JS)} = \left(1 - \frac{b}{a+\|z\|^2}\right) z$ has a smaller MSE than $\hat{\mu}^{(0)}$ for sufficiently small b and large a.*

This result was generalized to non-singular covariance matrices that were not necessarily the identity matrix (Theorem 2 of [REF 5]):

Theorem 2 *Let $\hat{\mu}^{(JS)} = \left(1 - \frac{c}{z^T \Sigma^{-1} z}\right) z$. If $Tr(\Sigma) \geq 2\lambda_L$ where λ_L is the largest eigenvalue of the covariance matrix Σ and $0 \leq c \leq 2\left(\frac{Tr(\Sigma)}{\lambda_L} - 2\right)$, then $\hat{\mu}^{(JS)}$ is the minimax estimator for the mean μ.*

Consider the Sleuth model with the simple experimental design above:

$$D_s | Y_s \sim \mathcal{N}_p\left(\beta_{s,0} + \mathbb{I}_{mut}\beta_{s,1}, (\sigma_s^2 + \tau_s^2) I_N\right)$$

For the n_{nonmut} non-mutated samples, this is equivalent to

$$D_s | Y_s \sim \mathcal{N}_{n_{nonmut}}\left(\beta_{s,0}, (\sigma_s^2 + \tau_s^2) I_n\right)$$

which can be fit with the same model process that Sleuth employs. For the single mutated sample, this model is

$$D_s | Y_s \sim \mathcal{N}\left(\beta_{s,0} + \beta_{s,1}, \max\{\hat{\sigma}_s^2, \tilde{\sigma}_s^2\} + \hat{\tau}_s^2\right) \tag{A.6}$$

The covariance matrix is the same as the mutated samples, but the mean $\beta_{s,0} + \beta_{s,1}$ is unknown and we have a single observation of this distribution. Reparameterizing Equation (A.6) to consider every transcript in the single mutated sample can be written as follows:

$$\Delta \sim \mathcal{N}_{|S|}(\mathrm{B}_0 + \mathrm{B}_1, \Sigma) \tag{A.7}$$

where

$$\mathrm{B}_{i,s} = \beta_{s,i} \forall s \in S \tag{A.8}$$

$$\Sigma = \begin{bmatrix} \max\{\hat{\sigma}_1^2, \tilde{\sigma}_1^2\} + \hat{\tau}_1^2 & & 0 \\ & \ddots & \\ 0 & & \max\{\hat{\sigma}_{|S|}^2, \tilde{\sigma}_{|S|}^2\} + \hat{\tau}_{|S|}^2 \end{bmatrix} \tag{A.9}$$

We switch from using coefficients $\beta_{t,i}$ to $\mathrm{B}_{i,s}$ to avoid confusion, since $\beta_{t,i} \in \mathbb{R}^p$ (a p-dimensional vector for each covariate in the design) whereas $\mathrm{B}_{i,s} \in \mathbb{R}^{|S|}$ (an $|S|$-dimensional vector for only a single coefficient over all transcripts in S).

Observations of a single mutated sample from this model meet the criteria for the JS estimators. Let δ be a single observation of the distribution Δ. A JS estimator for the unknown effect coefficient, B_1, can be constructed.

$$\hat{\mathrm{B}}_1^{(JS)} = \left(1 - \frac{c}{(\delta - \hat{\mathrm{B}}_0)^T \Sigma^{-1} (\delta - \hat{\mathrm{B}}_0)}\right)(\delta - \hat{\mathrm{B}}_0) \tag{A.10}$$

where $\hat{\mathrm{B}}_0$ is the estimate obtained from the non-mutated samples for all transcripts $s \in S$.

It is simple to see that $\mathrm{Tr}(\Sigma) = \sum_{s \in S} \max\{\hat{\sigma}_s^2, \tilde{\sigma}_s^2\} + \hat{\tau}_s^2$ and that $\lambda_L = \max_{s \in S} \left\{\max\{\hat{\sigma}_s^2, \tilde{\sigma}_s^2\} + \hat{\tau}_s^2\right\}$.

A.4 Comparison between the OLS and James-Stein estimators

For a simple experimental design where the mutation status is the only coefficient the OLS estimator is given by:

$$\begin{bmatrix} \hat{\beta}_{s,0}^{(OLS)} \\ \hat{\beta}_{s,1}^{(OLS)} \end{bmatrix} = \hat{\beta}_s^{(OLS)} = (X^T X)^{-1} X^T d_s = \begin{bmatrix} \bar{d}_s^{(nonmut)} \\ d_s^{(mut)} - \bar{d}_s^{(nonmut)} \end{bmatrix}$$

Looking closely at the OLS estimator for the mutation coefficient, $\beta_{s,1}$, it is clear that it is given by:

$$\hat{\beta}_{s,1}^{(OLS)} = d_s^{(mut)} - \hat{\beta}_{s,0}^{(OLS)} = \delta_s - \hat{\beta}_{0,s} \tag{A.11}$$

which is used directly in the definition of the JS estimator in Equation (A.10). The JS estimator for B_1 can then be expressed simply as:

$$\hat{\mathrm{B}}_1^{(JS)} = \left(1 - \frac{c}{\left(\hat{\mathrm{B}}_1^{(OLS)}\right)^T \Sigma^{-1} \hat{\mathrm{B}}_1^{(OLS)}}\right) \hat{\mathrm{B}}_1^{(OLS)} \tag{A.12}$$

From this definition, it is easy to see that the JS estimate is colinear with the OLS estimate but uniformly shrunk towards 0.

For a more general experimental design, the above can be extended. Given an experimental design matrix

$$X \in \mathbb{R}^{n \times p}$$

where $n > p$, $\mathrm{rank}(X) = p$ and $\mathrm{rank}(X^*) = p - 1$ where $X^* \in \mathbb{R}^{(n-1)\times p}$ is the same design matrix but with one sample removed, a JS estimator for the linear coefficient uniquely specified by the one sample is given by

$$\hat{\mathrm{B}}_i^{(JS)} = \left(1 - \frac{c}{\left(\hat{\mathrm{B}}_i^{(OLS)}\right)^T \Sigma^{-1} \hat{\mathrm{B}}_i^{(OLS)}}\right) \hat{\mathrm{B}}_i^{(OLS)}$$

A.5 Statistical moments of the James-Stein estimator

A.5.1 Expected value of the James-Stein estimator

Due to the non-linear nature of the JS estimator, a Taylor expansion around B_1 can be used to approximate the expectation. Consider:

$$\hat{\mathrm{B}}_1^{(JS)} = \left(1 - \frac{c}{\left(\hat{\mathrm{B}}_1^{(OLS)}\right)^T \Sigma^{-1} \hat{\mathrm{B}}_1^{(OLS)}}\right) \hat{\mathrm{B}}_1^{(OLS)}$$

where

$$\begin{aligned}\hat{\mathrm{B}}_1^{(OLS)} &\sim N_{|S|}(\mathrm{B}_1, \Sigma) \\ \Sigma_{s,s} &= \left(\frac{n_{nonmut}+1}{n_{nonmut}}\right)(\sigma_t^2 + \tau_t^2) \\ \Sigma_{s,t} &= 0 \forall t \neq s\end{aligned}$$

Let $u = \Sigma^{-1/2}\hat{\mathrm{B}}_1^{(OLS)}$. Then

$$\begin{aligned}\mathbb{E}\left[\hat{\mathrm{B}}_1^{(JS)}\right] &= \mathbb{E}\left[\hat{\mathrm{B}}_1^{(OLS)}\right] - c\Sigma^{1/2}\mathbb{E}\left[\frac{u}{\|u\|^2}\right] \\ &= \mathrm{B}_1 - c\Sigma^{1/2}\mathbb{E}\left[\frac{u}{\|u\|^2}\right]\Sigma^{1/2}\end{aligned}$$

Expanding $\frac{u}{\|u\|^2}$ around $a = \Sigma^{-1/2}\mathrm{B}_1$ gives:

$$\begin{aligned}\mathbb{E}\left[\hat{\mathrm{B}}_1^{(JS)}\right] &= \mathrm{B}_1 - c\Sigma^{1/2}\mathbb{E}\left[\frac{a}{\|a\|^2} + \left(\frac{1}{\|a\|^2} - \frac{2}{\|a\|^4}aa^T\right)(u-a) + \mathcal{O}(\|u-a\|^2)\right] \\ &= \left(1 - \frac{c}{\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1}\right)\mathrm{B}_1 + \mathcal{O}(\|u-a\|^2)\end{aligned}$$

As long as the number of transcripts being considered, $|S|$, is not large, and that the true coefficient of variation is not large (i.e. that $\|u-a\|^2 \ll \|\mathrm{B}_1\|^2$), the Taylor approximation is close to

$$\mathbb{E}\left[\hat{\mathrm{B}}_1^{(JS)}\right] \approx \left(1 - \frac{c}{\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1}\right)\mathrm{B}_1 \tag{A.13}$$

Thus the JS estimator is an estimate of B_1 that is biased towards 0.

A.5.2 Variance of the James-Stein estimator

The MSE of the JS estimator is related to its variance.

$$\mathbb{E}\left[\|\hat{\mathrm{B}}_1^{(JS)} - \mathrm{B}_1\|^2\right] = \sum_{s\in S}\mathbb{E}\left[\left(\hat{\mathrm{B}}_{1,s}^{(JS)} - \mathrm{B}_{1,s}\right)^2\right] = \sum_{s\in S}\mathbb{V}\left[\hat{\mathrm{B}}_{1,s}^{(JS)}\right]$$

By [REF 5], $\mathbb{E}\left[\|\hat{\mathrm{B}}_1^{(JS)} - \mathrm{B}_1\|^2\right] \leq \mathbb{E}\left[\|\hat{\mathrm{B}}_1^{(OLS)} - \mathrm{B}_1\|^2\right]$. However, this does not imply that $\mathbb{V}\left[\hat{\mathrm{B}}_{1,s}^{(JS)}\right] \leq \mathbb{V}\left[\hat{\mathrm{B}}_{1,s}^{(OLS)}\right] \forall s \in S$. Some transcripts may have larger variances than the OLS estimator, but all transcripts in aggregate will have a smaller MSE. This is still desirable if the goal is to find if there is an effect on any transcripts in the set S, instead of a particular one within the set.

To calculate the variance for each individual transcript, a similar approach with Taylor expansions can be used, as above.

$$\begin{aligned}
&\mathbb{V}\left[\hat{\mathrm{B}}_1^{(JS)}\right] \\
&\approx \mathbb{E}\left[\hat{\mathrm{B}}_1^{(JS)}\left(\hat{\mathrm{B}}_1^{(JS)}\right)^T\right] - \left(1 - \frac{c}{\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1}\right)^2 \mathrm{B}_1\mathrm{B}_1^T \\
&= \Sigma^{1/2}\mathbb{E}\left[uu^T - \frac{2c}{u^Tu}uu^T + \left(\frac{c}{u^Tu}\right)^2 uu^T\right]\Sigma^{1/2} - \left(1 - \frac{c}{\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1}\right)^2 \mathrm{B}_1\mathrm{B}_1^T
\end{aligned}$$

where, again, $u = \Sigma^{-1/2}\hat{\mathrm{B}}_1^{(OLS)}$. Expanding about $a = \Sigma^{-1/2}\mathrm{B}_1$ yields:

$$\mathbb{V}\left[\hat{\mathrm{B}}_1^{(JS)}\right] = \left(1 - \frac{2c}{\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1}\right)\Sigma - \frac{2c}{\left(\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1\right)^2}\mathrm{B}_1\mathrm{B}_1^T + \mathcal{O}(\|u-a\|^4)$$

Under similar conditions of the number of transcripts under consideration, $|S|$, and $\|u-a\|^2$, we then have that

$$\mathbb{V}\left[\hat{\mathrm{B}}_1^{(JS)}\right] \approx \left(1 - \frac{2c}{\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1}\right)\Sigma - \frac{2c}{\left(\mathrm{B}_1^T\Sigma^{-1}\mathrm{B}_1\right)^2}\mathrm{B}_1\mathrm{B}_1^T \tag{A.14}$$

Again, a shrinkage factor multiplies Σ, providing the possibility of a smaller estimator variance than the OLS estimator. This smaller variance allows for more powerful statistical inferences when using a Wald test. Notably, B_1 is unknown, so the variance of the JS estimator is a function of both the mean and variance of the transcripts under consideration. This is in contrast to the OLS estimator, which is solely a function of the variance.

Glossary

3C chromatin conformation capture

AR androgen receptor

ChIP-seq chromatin immunoprecipitation sequencing

CPC-GENE Canadian Prostate Cancer Genome Network

crRNA CRISPR RNA

CRE *cis*-regulatory element

DEPMAP Cancer Dependency Map

DHS DNase I hypersensitive sites

FN false negative

FP false positive

FOX forkhead box

GLM generalized linear model

gRNA guide RNA

IID independent and identically distributed

JS James-Stein

kbp kilobase

KO knockout

MSE mean square error

mCRPC metastatic castration-resistant prostate cancer

OLS ordinary least squares

mRNA messenger RNA

PCa prostate cancer

RNAi RNA interference

RNA-seq RNA sequencing

shRNA small hairpin RNA

siRNA small interfering RNA

SNV single nucleotide variants

SV structural variant

TAD topologically associated domain

TCGA The Cancer Genome Atlas

TN true negative

TP true positive

TF transcription factor

tracrRNA trans-activating CPRISR RNA

UTR untranslated region

WGS whole genome sequencing

WT wild-type

References

1. Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential Analysis of RNA-Seq Incorporating Quantification Uncertainty. en. *Nature Methods* **14,** 687–690. ISSN: 1548-7105 (July 2017).

2. Yi, L., Pimentel, H., Bray, N. L. & Pachter, L. Gene-Level Differential Analysis at Transcript-Level Resolution. *Genome Biology* **19,** 53. ISSN: 1474-760X (Apr. 2018).

3. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-Optimal Probabilistic RNA-Seq Quantification. en. *Nature Biotechnology* **34,** 525–527. ISSN: 1546-1696 (May 2016).

4. Stein, C. *Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution* en. in *Contribution to the Theory of Statistics* **3** (University of California Press, Berkeley, California, USA, Dec. 1956), 197–206. ISBN: 978-0-520-31388-0.

5. Bock, M. E. Minimax Estimators of the Mean of a Multivariate Normal Distribution. en. *The Annals of Statistics* **3,** 209–218. ISSN: 0090-5364 (Jan. 1975).