

Applications of Zero-inflated Model for Single-cell RNA-seq Data

Fanny Perradeau

2016-06-14

Contents

Supervised Differential Expression Analysis	1
Unsupervised Differential Expression Analysis	1
Imputation	1

Supervised Differential Expression Analysis

We want to find genes differentially expressed between two cell types. Stat: read counts from single cell RNASeq. We know Y and X. Model: ZINB. Parameters of interest β .

Write equation

Steps:

1. Filter low qv: number of reads $> 15,000$, perc aligner $> 50\%$
2. Estimate β_π by fitting ZINB (pscl then team JP zinb)
3. Estimate β_μ by fitting NB (?)
4. DE with β_π and β_μ : edgeR?, DESeq?
5. metric to compare method
6. Plot MD, heatmaps (look at Sandrine's slides)

Unsupervised Differential Expression Analysis

We want to identify clusters and find genes differentially expressed between found clusters. Different possibilities: one against all, pairwise, hierarchical (closest to closest). For the moment, limma + voom. We want zinb instead.

Steps:

1. read vignette clusterExperiment
2. Hierarchical clusters using clusterExperiment. Use function catbestfeatures().

Imputation

We want a function to impute Y using zinb (equation) from Y. In scone function, arg imputation. For the moment, only possibility is imputation=identity. Then scone would perform normalization and compute metrics to say if normalization worked well