A Two-Step Growth Mixture Model With Distributional Changes Over Time

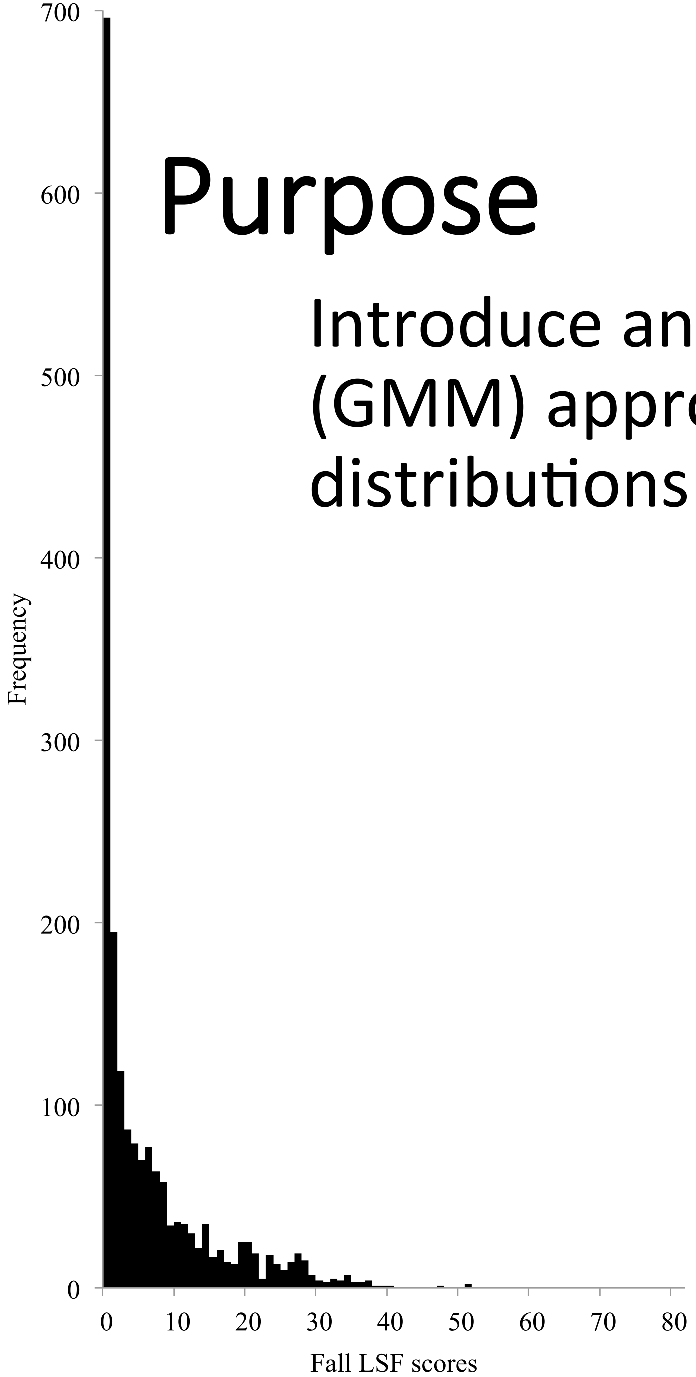
Joseph F. T. Nese
University of Oregon

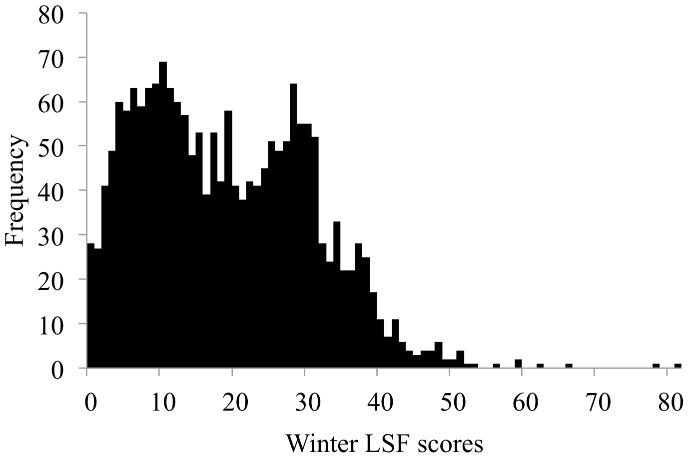
Akihito Kamata
Southern Methodist University

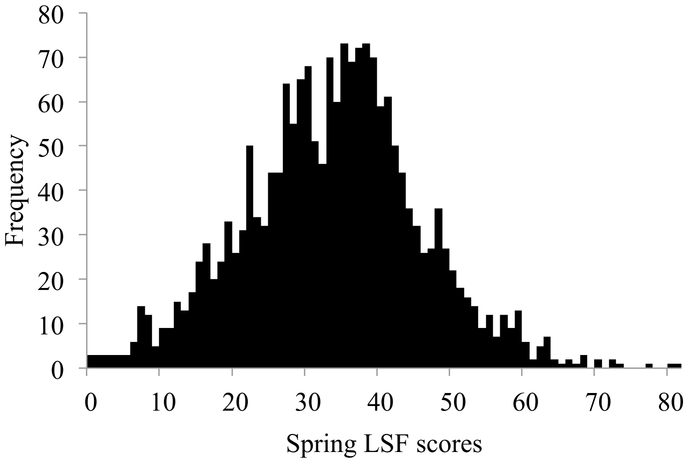
UNIVERSITY OF OREGON

Purpose

Introduce and apply a two-step growth mixture model (GMM) approach for modeling repeated measures with distributions changing over time.







Two-Step GMM Approach

I. Apply a 2-latent-class mixture zero-inflated Poisson (ZIP) regression model to the initial measurement occasion (entry) and related entry covariates to identify two classes:
   1) **Zero Class,** and
   2) **Zero & Above Class**.
II. Apply a GMM for count data to all measurement occasions for each of the two classes estimated in step-1.
   - We used the estimated class probabilities of the corresponding class as sampling weights (similar to propensity score weighting).

Step-1

- ZIP regression of initial measurement occasion (entry) on related entry covariates
  1) **Zero Class**: students who could only assume a zero score on initial occasion
  2) **Zero & Above Class**: students who could assume scores zero or higher

$$Y_i \sim \begin{cases} 0 & \text{with probability } p \\ Poisson(\lambda_i) & \text{with probability } 1-p \end{cases}$$

Where Y_i is the observed initial occasion score for the ith student

$$\ln(\lambda_i) = \begin{cases} b_0^{(1)} + b_1^{(1)} x_{1i} & \text{with probability } p \\ b_0^{(2)} + b_1^{(2)} x_{1i} & \text{with probability } 1-p \end{cases},$$

Where λ_i is an event rate for student *i*,
x_{1i} is the entry covariate for student *i*,
the numerical value in the parentheses in the superscript is an indicator of a latent class.
Intercept for $b_0^{(1)}$ was fixed at -15 to represent an extremely low log-rate such that the probablility of a count > 0 was essentially zero

Step-2

- The models for *J* latent classes can be written as

$$\ln(\boldsymbol{\lambda}_{ti}) = \begin{cases} \boldsymbol{\Lambda}^{(1)}\eta_{Si}^{(1)} & \text{with probability } p_2^{(1)} \\ \boldsymbol{\Lambda}^{(2)}\eta_{Si}^{(2)} & \text{with probability } p_2^{(2)} \\ \vdots & \vdots \\ \boldsymbol{\Lambda}^{(J)}\eta_{Si}^{(J)} & \text{with probability } p_2^{(J)}, \end{cases}$$

Zero Only class GMM

- Included observations with zero scores in the initial occasion (assumed observations with non-zero initial scores had zero likelihood of being in this class).

$$\eta_{Si}^{(j)} \sim N(\beta_1^{(j)}, \phi_{11}^{(j)}), \text{ and } \boldsymbol{\Lambda}^{(j)} = \begin{bmatrix} 0 \\ 1 \\ \lambda^{(j)} \end{bmatrix}$$

Here, $\beta_1^{(j)}$ is the mean change, $\phi_{11}^{(1)}$ is the variance of the change, and $\lambda^{(j)}$ is the estimated time score for the *j*th class

Zero & Above Class GMM

- Included all observations (assumed all observations had some likelihood of being in this class).

$$\boldsymbol{\eta}_i^{(j)} = \begin{bmatrix} \eta_{Ii}^{(j)} \\ \eta_{Si}^{(j)} \end{bmatrix} \sim N\left(\begin{bmatrix} \beta_0^{(j)} \\ \beta_1^{(j)} \end{bmatrix}, \begin{bmatrix} \phi_{00}^{(j)} & \\ \phi_{01}^{(j)} & \phi_{11}^{(j)} \end{bmatrix} \right), \text{ and } \boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & \lambda^{(j)} \end{bmatrix}, \text{ where } j = (1, \ldots, J).$$

Here, $\beta_0^{(j)}$ is the mean of intercept, $\beta_1^{(j)}$ is the mean of slope, $\phi_{00}^{(1)}$ is the variance of the intercept, $\phi_{11}^{(1)}$ is the variance of the slope, and $\lambda^{(j)}$ is the estimated time score for the *j*th class

Applied Example

- Sample
  - 1,911 kindergarten students in 2009-2010
- Measures
  - Entry covariate:
    - Letter Names Fluency (LNF)
    - Scale: names correct per minute (ncpm)
  - Repeated outcome:
    - Letter Sound Fluency (LSF)
    - Scale: sounds correct per minute (scpm)

Applied Results: Step 1

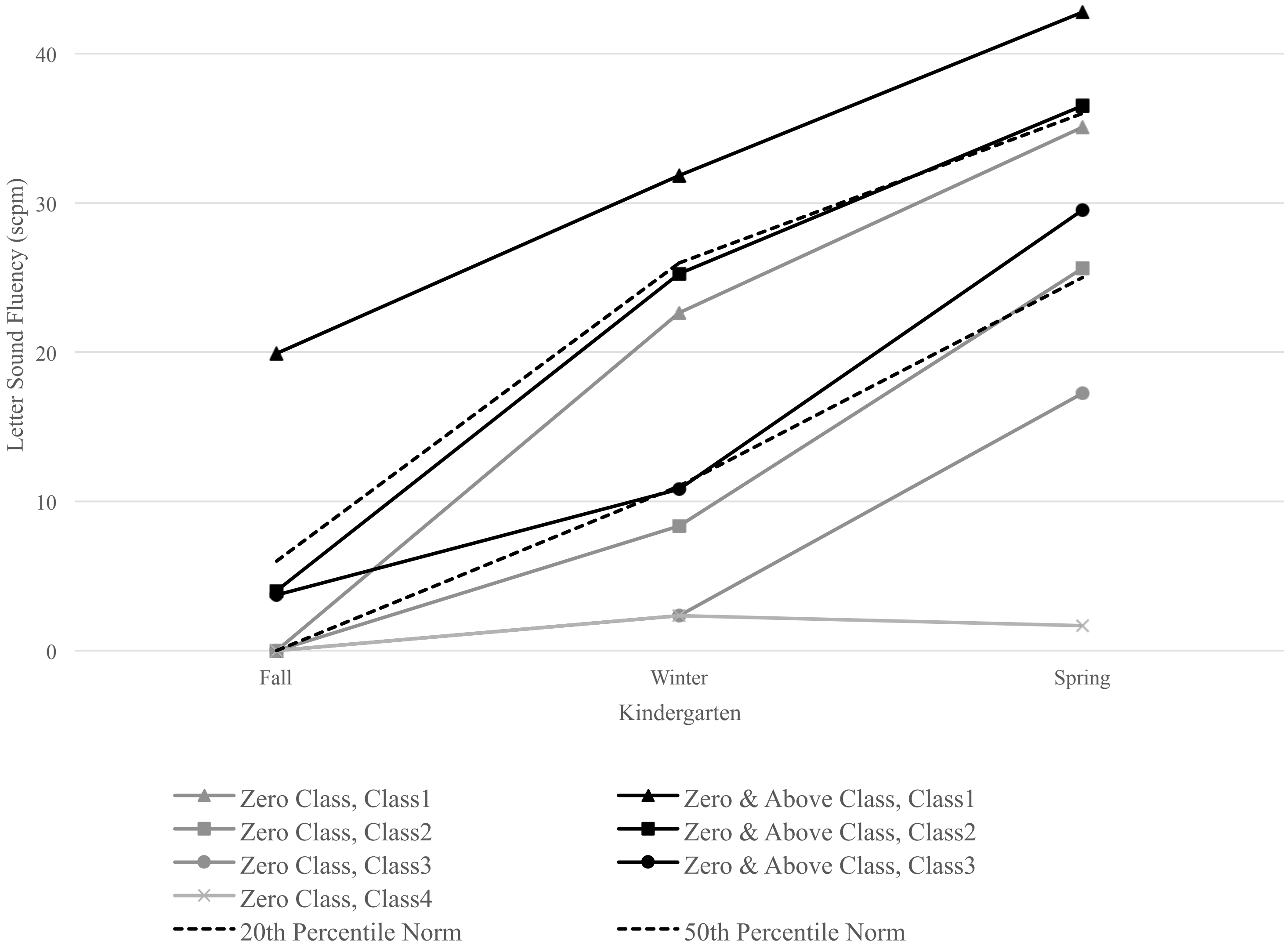
- Zero Class
  - 645.34 students (sum of the estimated class probabilities)
    - vs. 687 students (classify-analyze approach)
  - Intercept ≅ 0 scpm ($e^{-15} \cong 0$)
- Zero & Above Class
  - 1265.66 students (sum of the estimated class probabilities)
    - vs. 1224 students (classify-analyze approach)
  - Intercept = 2.36 scpm ($e^{0.875}$)

Applied Results: Step 2

Latent Classes								Latent Class n (%) [posterior probability]			
Step-1	Step-2	AIC	BIC	ABIC	Entropy	VLMR p-value	BLR p-value	1	2	3	4
Zero Class	1	11441.60	11455.24	11455.71	-	-	-	645.22			
	2	10247.60	10279.42	10257.19	.820	0.0000	0.0000	386.17 (59.9) [.97]	259.06 (40.2) [.91]		
	3	10072.76	10122.76	10087.84	.875	0.0000	0.0000	369.33 (57.2) [.97]	259.95 (40.3) [.91]	15.94 (2.5) [.89]	
	4	**9935.99**	**10004.17**	**9956.54**	**.800**	**0.0001**	**0.0000†**	**118.94 (18.4) [.78]**	**305.05 (47.3) [.90]**	**15.38 (2.4) [.90]**	**205.85 (31.9) [.93]**
Zero & Above Class	1	40958.26	40991.59	40972.53	-	-	-	1265.68			
	2	40079.89	40152.11	40110.81	.490	0.0000	0.0000†	726.28 (57) [.87]	539.37 (43) [.79]		
	3	**39838.48**	**39949.58**	**39886.05**	**.623**	**0.0000**	**0.0000†**	**300.93 (24) [.86]**	**492.47 (39) [.81]**	**472.26 (37) [.84]**	
	4	39768.07	39918.07	39832.29	.658	0.0035	0.0000†	542.29 (42) [.81]	251.20 (23) [.86]	440.05 (34) [.79]	32.11 (2) [.63]

Note. VLMR = Vuong-Lo-Mendell-Rubin likelihood ratio test. BLRT = parametric bootstrapped likelihood ratio test.

Results


40
30
20
10
0
Letter Sound Fluency (scpm)
Fall
Winter
Spring
Kindergarten
Zero Class, Class1
Zero Class, Class2
Zero Class, Class3
Zero Class, Class4
20th Percentile Norm
Zero & Above Class, Class1
Zero & Above Class, Class2
Zero & Above Class, Class3
50th Percentile Norm

Results

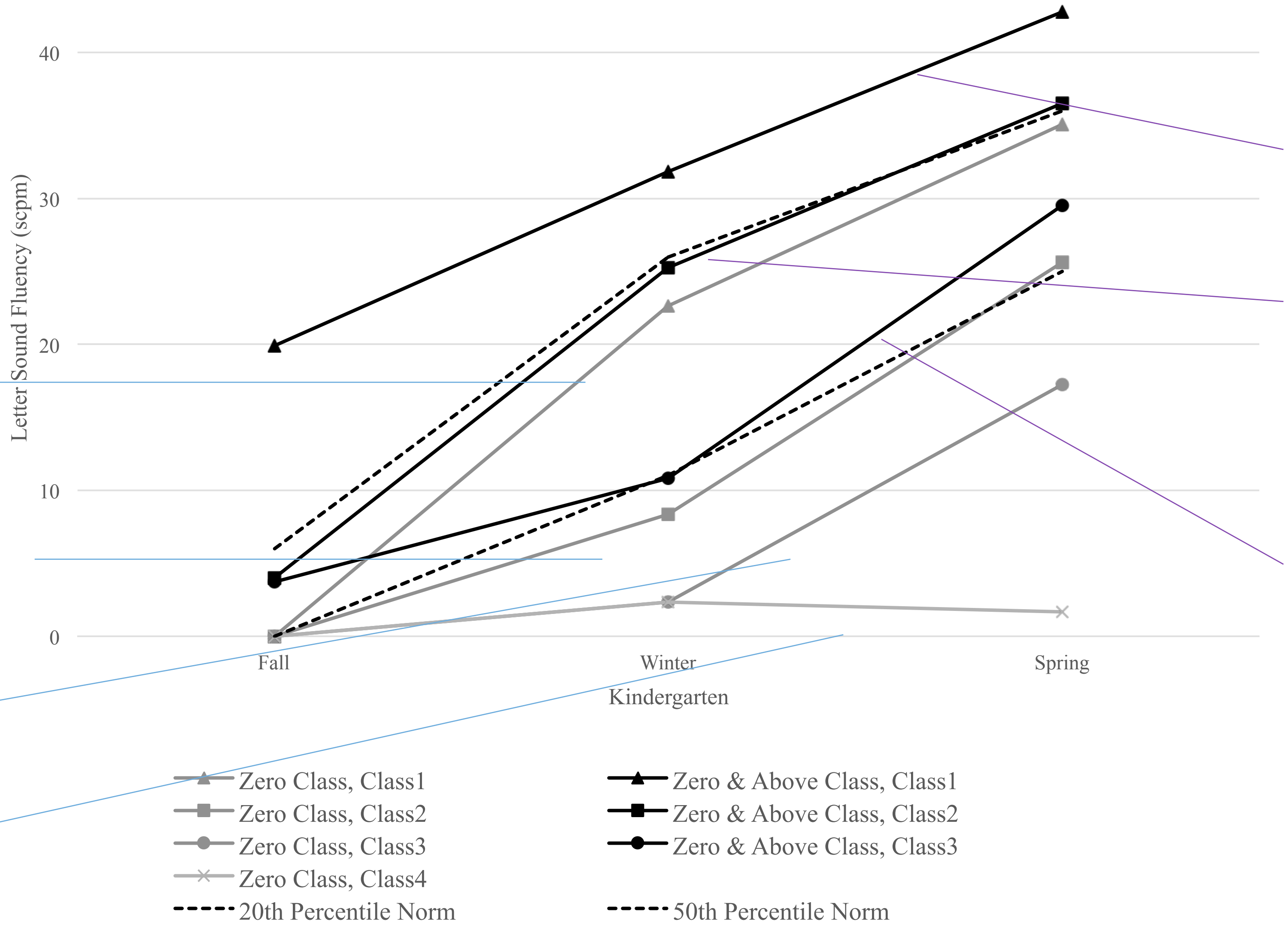
Zero Class

Very High Growth
- "Average" group, learning to read

High Growth
- "Low average" group, learning to read

Moderate Growth
- Not yet reading

No Growth
- Not learning to read



Zero & Above Class

High Intercept
- Ready to read words

Good Intercept / High Growth
- "Average" group, ready to read

Good Intercept / Good Growth
- "Low average" group, ready to learn to read words

Discussion

- ZI initial data and distributional changes over time is not interesting or novel in itself.
- Great potential lies in the method of distinguishing between students whom begin at zero and make meaningful gains and students whom begin at zero and do not.
- The value lies in demarcating these groups before the skill disparity between them becomes readily evident.

Questions

1) Do you have longitudinal data with similar distributional properties?

2) What are your reactions to the theoretical implications of the reading findings we presented?

3) Is there an approach to simplify our two-step approach into a single model?