A Maximum Likelihood Estimators of Selected State-Dependent Distributions

Without loss of generality, we provide the MLE for the distribution parameters of the u-th latent state in the m-th iteration. To simplify the notation, we let $\hat{z}_{ijl,u}^{(m)} = \mathbb{E}(\mathbb{1}_{\{Z_{ijl}=u\}}|\mathbf{Y}^*, \mathbf{T}^*, \mathbf{\Phi}^{(m-1)})$.

- Gamma(k, θ) (shape-scale parametrization)

$$k^{(m)} = k \quad \text{such that} \quad \sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} \left\{ \log\Gamma(k) + k\log(\theta^{(m-1)}) - (k-1)\log(y_{ijl}) + \frac{1}{\theta^{(m-1)}} y_{ijl} \right\} = 0$$

$$\theta^{(m)} = \frac{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} y_{ijl}}{k^{(m)} \sum_{i,j,l} \hat{z}_{ijl,u}^{(m)}}$$

- Laplace(μ, θ)

$$\mu^{(m)} = \text{weighted median of } y_{ijl} \text{ with weights } \hat{z}_{ijl,u}^{(m)}$$

$$\theta^{(m)} = \frac{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} |y_{ijl} - \mu^{(m)}|}{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)}}$$

- Log-Normal(μ, σ^2)

$$\mu^{(m)} = \frac{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} \log(y_{ijl})}{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)}}$$

$$(\sigma^2)^{(m)} = \frac{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} (\log(y_{ijl}) - \mu^{(m)})^2}{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)}}$$

- Normal(μ, σ^2)

$$\mu^{(m)} = \frac{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} y_{ijl}}{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)}}$$

$$(\sigma^2)^{(m)} = \frac{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} (y_{ijl} - \mu^{(m)})^2}{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)}}$$

- von Mises(μ, κ)

$$\mu^{(m)} = \arctan\left(\frac{\sum_{j,k} \hat{z}_{ijl,u}^{(m)} \sin(y_{ijl})}{\sum_{j,k} \hat{z}_{ijl,u}^{(m)} \cos(y_{ijl})} \right)$$

$$\kappa^{(m)} = \kappa \quad \text{such that} \quad A(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)} = \frac{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)} \cos(y_{ijl} - \mu^{(m)})}{\sum_{i,j,l} \hat{z}_{ijl,u}^{(m)}}$$

where $I_0(\kappa)$ and $I_1(\kappa)$ are the modified Bessel function of the first kind of orders 0 and 1 respectively.

B Controlled Study: Anomaly Indices from Individual-Specific Models with Fewer States

Table B1: UAH-DriveSet: Anomaly index in each dimension of telematics observations, computed from individual-specific CTHMMs with 5 states for each driver and averaged over 100 random initializations. Note: S stands for Secondary Road and M stands for Motorway; Z and Y are longitudinal and lateral accelerations, respectively.

Driver	Trip	Y_KF	Z_KF	Speed
D1	S-Normal 1	0.0025	0.0014	0.0000
	S-Normal 2	0.0027	0.0032	0.0000
	S-Aggressive	0.0112	**0.0104**	0.0000
	S-Drowsy	0.0043	**0.0083**	0.0000
	M-Normal	0.0135	0.0020	0.0002
	M-Aggressive	0.0032	*0.0079*	0.0010
	M-Drowsy	0.0013	0.0034	0.0000
D2	S-Normal 1	0.0031	0.0015	0.0000
	S-Normal 2	0.0047	0.0000	0.0000
	S-Aggressive	0.0016	**0.0317**	0.0000
	S-Drowsy	0.0258	0.0078	0.0000
	M-Normal	0.0063	0.0019	0.0019
	M-Aggressive	0.0129	**0.0158**	0.0108
	M-Drowsy	0.0180	0.0043	0.0001
D3	S-Normal 1	0.0065	0.0029	0.0000
	S-Normal 2	0.0121	0.0105	0.0000
	S-Aggressive	0.0114	**0.0436**	0.0000
	S-Drowsy	0.0114	0.0062	0.0000
	M-Normal	0.0067	0.0038	0.0119
	M-Aggressive	0.0103	**0.0305**	0.0404
	M-Drowsy	0.0150	0.0010	0.0000
D4	S-Normal 1	0.0056	0.0021	0.0000
	S-Normal 2	0.0103	0.0022	0.0000
	S-Aggressive	0.0112	**0.0168**	0.0000
	S-Drowsy	0.0235	0.0059	0.0000
	M-Normal	0.0035	0.0030	0.0070
	M-Aggressive	0.0242	**0.0275**	0.0000
	M-Drowsy	0.0107	0.0019	0.0000
D5	S-Normal 1	0.0000	0.0000	0.0000
	S-Normal 2	0.0000	0.0000	0.0044
	S-Aggressive	0.0067	**0.0116**	0.0000
	S-Drowsy	0.0040	0.0013	0.0000
	M-Normal	0.0033	0.0030	0.0011
	M-Aggressive	0.0088	**0.0138**	0.0000
	M-Drowsy	0.0046	0.0009	0.0000

Table B2: UAH-DriveSet: Anomaly index in each dimension of telematics observations, computed from individual-specific CTHMMs with 10 states for each driver and averaged over 100 random initializations. Note: S stands for Secondary Road and M stands for Motorway; Z and Y are longitudinal and lateral accelerations, respectively.

Driver	**Trip**	**Y_KF**	**Z_KF**	**Speed**
D1	S-Normal 1	0.0076	0.0032	0.0000
	S-Normal 2	0.0033	0.0000	0.0000
	S-Aggressive	0.0064	*0.0056*	0.0000
	S-Drowsy	0.0053	**0.0057**	0.0000
	M-Normal	0.0023	0.0004	0.0010
	M-Aggressive	0.0042	**0.0064**	0.0039
	M-Drowsy	0.0018	0.0032	0.0000
D2	S-Normal 1	0.0026	0.0003	0.0000
	S-Normal 2	0.0053	0.0003	0.0000
	S-Aggressive	0.0006	**0.0174**	0.0000
	S-Drowsy	0.0245	0.0051	0.0000
	M-Normal	0.0051	0.0022	0.0019
	M-Aggressive	0.0104	**0.0136**	0.0060
	M-Drowsy	0.0151	0.0038	0.0000
D3	S-Normal 1	0.0049	0.0028	0.0000
	S-Normal 2	0.0073	0.0082	0.0000
	S-Aggressive	0.0092	**0.0301**	0.0000
	S-Drowsy	0.0073	0.0043	0.0000
	M-Normal	0.0062	0.0017	0.0016
	M-Aggressive	0.0078	**0.0283**	0.0138
	M-Drowsy	0.0083	0.0008	0.0000
D4	S-Normal 1	0.0046	0.0028	0.0000
	S-Normal 2	0.0068	0.0023	0.0000
	S-Aggressive	0.0082	**0.0145**	0.0000
	S-Drowsy	0.0184	0.0054	0.0000
	M-Normal	0.0023	0.0015	0.0028
	M-Aggressive	0.0112	**0.0239**	0.0002
	M-Drowsy	0.0062	0.0005	0.0000
D5	S-Normal 1	0.0001	0.0003	0.0000
	S-Normal 2	0.0030	0.0027	0.0046
	S-Aggressive	0.0056	**0.0147**	0.0000
	S-Drowsy	0.0042	0.0019	0.0000
	M-Normal	0.0009	0.0025	0.0010
	M-Aggressive	0.0105	**0.0112**	0.0000
	M-Drowsy	0.0056	0.0011	0.0000

Table B3: UAH-DriveSet: Anomaly index in each dimension of telematics observations, computed from individual-specific CTHMMs with 15 states for each driver and averaged over 100 random initializations. Note: S stands for Secondary Road and M stands for Motorway; Z and Y are longitudinal and lateral accelerations, respectively.

Driver	**Trip**	**Y_KF**	**Z_KF**	**Speed**
D1	S-Normal 1	0.0105	0.0043	0.0000
	S-Normal 2	0.0055	0.0015	0.0000
	S-Aggressive	0.0085	*0.0060*	0.0000
	S-Drowsy	0.0090	**0.0062**	0.0000
	M-Normal	0.0033	0.0012	0.0001
	M-Aggressive	0.0049	**0.0092**	0.0027
	M-Drowsy	0.0021	0.0038	0.0000
D2	S-Normal 1	0.0022	0.0006	0.0000
	S-Normal 2	0.0043	0.0002	0.0000
	S-Aggressive	0.0017	**0.0197**	0.0003
	S-Drowsy	0.0226	0.0048	0.0000
	M-Normal	0.0032	0.0019	0.0011
	M-Aggressive	0.0082	**0.0163**	0.0019
	M-Drowsy	0.0146	0.0041	0.0000
D3	S-Normal 1	0.0047	0.0021	0.0000
	S-Normal 2	0.0072	0.0056	0.0000
	S-Aggressive	0.0099	**0.0216**	0.0000
	S-Drowsy	0.0075	0.0040	0.0000
	M-Normal	0.0057	0.0020	0.0049
	M-Aggressive	0.0057	**0.0225**	0.0004
	M-Drowsy	0.0080	0.0006	0.0000
D4	S-Normal 1	0.0046	0.0031	0.0000
	S-Normal 2	0.0083	0.0015	0.0000
	S-Aggressive	0.0098	**0.0143**	0.0000
	S-Drowsy	0.0169	0.0032	0.0000
	M-Normal	0.0022	0.0020	0.0021
	M-Aggressive	0.0146	**0.0187**	0.0002
	M-Drowsy	0.0059	0.0011	0.0000
D5	S-Normal 1	0.0002	0.0009	0.0000
	S-Normal 2	0.0022	0.0022	0.0043
	S-Aggressive	0.0065	**0.0128**	0.0000
	S-Drowsy	0.0056	0.0027	0.0000
	M-Normal	0.0011	0.0031	0.0005
	M-Aggressive	0.0102	**0.0072**	0.0000
	M-Drowsy	0.0059	0.0011	0.0000

C Real-Data Analysis: Summaries of Logistic Models from Individual-Specific Models

Table C1: Model summary of logistic GLMs, with raw or normalized anomaly indices as covariates. Note: x1, x2 and x3 denote indices for speed, longitudinal and lateral accelerations respectively.

	Estimate	**Std. Error**	**z value**	**p value**	
(Intercept)	-3.395	0.137	-24.825	< 2e-16	***
x1	-23.146	20.594	-1.124	0.261	
x2	51.219	8.630	5.935	2.93e-09	***
x3	63.827	9.416	6.779	1.21e-11	***

	Estimate	**Std. Error**	**z value**	**p value**	
(Intercept)	-4.316	0.206	-20.961	<2e-16	***
norm x1	0.309	0.557	0.555	0.579	
norm x2	2.770	0.298	9.305	<2e-16	***
norm x3	2.903	0.292	9.957	<2e-16	***

Significance codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' ', 1

Table C2: Feature importance of logistic XGBoost, with raw or normalized anomaly as covariates. Note: x1, x2 and x3 denote indices for speed, longitudinal and lateral accelerations respectively.

	Gain	**Cover**	**Frequency**
x3	0.557	0.492	0.447
x2	0.382	0.401	0.392
x1	0.061	0.106	0.161

	Gain	**Cover**	**Frequency**
norm x3	0.537	0.530	0.389
norm x2	0.430	0.414	0.465
norm x1	0.033	0.056	0.146

D Real-Data Analysis: Pooled Model Analysis with Alternative Training Sets

Table D1: Comparison of anomaly indices (%) in each telematics response dimension for claimed and no-claim rentals, computed from pooled CTHMMs trained on samples with 0% and 100% claim rates.

Trip-Basis

	0%		**100%**	
	claimed	no claim	claimed	no claim
speed	0.0716	0.0995	0.0462	0.0523
longitudinal	0.5405	0.6549	0.2473	0.3008
lateral	0.3048	0.3517	0.2178	0.2384

Rental-Basis

	0%		**100%**	
	claimed	no claim	claimed	no claim
Average				
speed	0.0713	0.0968	0.0433	0.0559
longitudinal	0.6246	0.8882	0.2680	0.3936
lateral	0.3190	0.3801	0.2219	0.2711
Maximum				
speed	32.3730	17.2720	25.2180	12.0900
longitudinal	35.8500	20.8000	28.2700	14.8470
lateral	33.9200	18.8010	27.2260	13.9027

Table D2: 5-fold cross-validation ROC-AUC of logistic GLMs with different sets of covariates. Set 1: maximum anomaly indices, Set 2: right tail of the anomaly index empirical distribution, Set 3: right tail and the number of trips as an offset. Note: first column indicates the claim rate of the training set the pooled CTHMM is fitted on.

	Covariate Sets		
	Set 1	**Set 2**	**Set 3**
0% Training Set	0.6905	0.7509	0.7861
100% Training Set	0.6962	0.7286	0.7669