

[2021-03-02 Tue] Summary: simulated survival time and log-loss

Jonghyun Yun

Intro

This article provides analysis of the issue in the posterior predictive checking we discussed in *[2021-03-01 Mon]* meeting.

TL;DR

It is advised to switch signs of the latent distance terms. A new model gives acceptable results for posterior predictive checking of the response type. We may not want to report separate results for each response type. The computational method I implemented doesn't seem to be wrong. So the issue might be coming from the statistical model itself.

Models

The below is a list of models in consideration. They differ by the sign of the latent distance term. A baseline (base) model is one with no latent distance term.

np model: (-,+) = (correct,incorrect)

$$h_{i1}(t) = \lambda_{i1}(t) \exp(\beta_{i1} + \theta_{k1} - ||z_k - w_i||)$$

$$h_{i0}(t) = \lambda_{i0}(t) \exp(\beta_{i0} + \theta_{k0} + ||z_k - w_i||)$$

pn model: (+,-) = (correct,incorrect)

$$h_{i1}(t) = \lambda_{i1}(t) \exp(\beta_{i1} + \theta_{k1} + ||z_k - w_i||)$$

$$h_{i0}(t) = \lambda_{i0}(t) \exp(\beta_{i0} + \theta_{k0} - ||z_k - w_i||)$$

Histogram of survival times

You need to open this document in Github using the web browser to see PDF files. The simulated times are divided into two panes based on the observed response type. The np model didn't have the issue of simulating the max survival time (=31).

- [np model](#)
- [pn model](#)
- [base model](#)

Log-loss

Given the survival time t and MCMC samples, the conditional probability of “correct” response is estimated as

$$\hat{p}_{ik}^{(l)}(t) = \frac{h_{i1}^{(l)}(t)}{h_{i1}^{(l)}(t) + h_{i0}^{(l)}(t)},$$

where (l) denote the index of MCMC sample used to calculate the quantity. We use $Y_{ik} = 1$ to denote a correct response, and $Y_{ik} = 0$ an incorrect response. Then, $\hat{p}_{ik}^{(l)}(t)$ can be served as the prediction probability of Y_{ik} , which allows using classification performance metrics. For easiness of evaluation, we use the log-loss (over respondent and item)

$$-\frac{1}{L} \sum_{l=1}^L \left\{ y_{ik} \log \hat{p}_{ik}^{(l)}(t) + (1 - y_{ik}) \log (1 - \hat{p}_{ik}^{(l)}) \right\}.$$

The log-loss are concatenated over respondents for each item. On average, the np model gives smaller log-losses than the baseline for all correct responses, but the pn model fails to do so for both response types. In terms of “posterior density” and “likelihood”, these two models have similar fit, and they have better fit than the baseline model.

We calculated differences of concatenated log-loss: 1) base - np log-loss and 2) base - pn log-loss. So, the larger is the better. We provide a series of box plots for the differences over all items. The simulated survival time is used to predict the response in the 1st PDF. On the otherhand, the observed survival time is used in the 2nd PDF. They show similar patterns (the results is a bit more promising with the simulation though), so the simulation is not likely to be a “culprit”. I also didn't find any error in the MCMC sampler. Perhaps, this could be what we get from switching signs of the latent distance.

- [With simulated survival time](#)

- With observed survival time

Furthermore, tables for means and standard errors for the differences (with simulated survival time) are listed below. `pos_neg` and `se_pn` denote mean and standard error of base - pn log-loss. `neg_pos` and `np` denote those of base - np log-loss.

Log-loss differences

item	pos_neg	se_pn	neg_pos	se_np
1	-0.935	0.039	0.692	0.041
2	-0.944	0.038	0.636	0.040
3	-0.892	0.040	0.460	0.040
4	-0.938	0.038	0.511	0.040
5	-0.919	0.039	0.540	0.040
6	-0.882	0.040	0.326	0.038
7	-0.888	0.039	0.331	0.037
8	-0.854	0.041	0.221	0.035
9	-0.613	0.045	0.059	0.029
10	-0.771	0.043	0.350	0.040
11	-0.747	0.043	0.089	0.031
12	-0.606	0.045	0.034	0.026
13	-0.659	0.043	0.165	0.033
14	-0.593	0.045	0.049	0.028
15	-0.814	0.042	0.178	0.033
16	-0.328	0.043	-0.013	0.024
17	-0.413	0.045	0.118	0.032
18	-0.289	0.042	0.120	0.031
19	-0.412	0.044	0.087	0.029
20	-0.100	0.038	0.048	0.028
21	-0.942	0.038	0.575	0.039
22	-0.838	0.041	0.374	0.038
23	-0.800	0.041	0.349	0.038
24	-0.320	0.043	0.021	0.026
25	-0.768	0.041	0.331	0.038
26	-0.316	0.044	-0.017	0.023
27	-0.066	0.033	0.026	0.025
28	-0.267	0.044	0.095	0.030
29	-0.501	0.044	0.096	0.031

Continued on next page

Continued from previous page

item	pos_neg	se_pn	neg_pos	se_np
30	-0.046	0.033	0.137	0.032
31	-0.955	0.038	0.711	0.041
32	-0.933	0.038	0.643	0.040
33	-0.920	0.039	0.432	0.039
34	-0.425	0.044	-0.009	0.024
35	-0.510	0.043	0.065	0.028
36	-0.473	0.043	0.100	0.030
37	-0.135	0.038	0.008	0.025
38	-0.257	0.040	-0.005	0.024
39	-0.383	0.044	0.043	0.027
40	-0.125	0.037	0.167	0.033

Log-loss differences by response type

res	item	pos_neg	se_pn	neg_pos	se_np
incorrect	1	-0.083	0.014	-0.126	0.018
correct	1	-0.970	0.038	0.726	0.041
incorrect	2	-0.146	0.019	-0.318	0.034
correct	2	-0.995	0.037	0.696	0.037
incorrect	3	-0.164	0.015	-0.270	0.026
correct	3	-1.051	0.037	0.620	0.035
incorrect	4	-0.237	0.009	-0.398	0.024
correct	4	-1.033	0.037	0.634	0.035
incorrect	5	-0.178	0.016	-0.332	0.025
correct	5	-1.012	0.037	0.650	0.036
incorrect	6	-0.203	0.010	-0.318	0.024
correct	6	-1.109	0.035	0.542	0.031
incorrect	7	-0.200	0.013	-0.324	0.027
correct	7	-1.109	0.034	0.541	0.029
incorrect	8	-0.208	0.012	-0.292	0.023
correct	8	-1.174	0.034	0.474	0.029
incorrect	9	-0.107	0.016	-0.149	0.023
correct	9	-1.317	0.033	0.348	0.026
incorrect	10	-0.113	0.016	-0.161	0.024
correct	10	-1.078	0.038	0.589	0.036
incorrect	11	-0.167	0.014	-0.221	0.022

Continued on next page

Continued from previous page

res	item	pos_neg	se_pn	neg_pos	se_np
correct	11	-1.278	0.033	0.372	0.027
incorrect	12	-0.106	0.018	-0.154	0.025
correct	12	-1.325	0.032	0.305	0.018
incorrect	13	-0.100	0.018	-0.155	0.026
correct	13	-1.170	0.034	0.458	0.026
incorrect	14	-0.097	0.016	-0.143	0.024
correct	14	-1.329	0.033	0.335	0.022
incorrect	15	-0.179	0.014	-0.253	0.024
correct	15	-1.208	0.034	0.444	0.025
incorrect	16	-0.030	0.019	-0.065	0.024
correct	16	-1.466	0.032	0.184	0.016
incorrect	17	-0.018	0.020	-0.055	0.025
correct	17	-1.198	0.042	0.461	0.032
incorrect	18	0.015	0.021	-0.019	0.025
correct	18	-1.106	0.040	0.495	0.034
incorrect	19	-0.012	0.023	-0.067	0.027
correct	19	-1.223	0.030	0.401	0.023
incorrect	20	0.036	0.021	0.013	0.024
correct	20	-1.444	0.058	0.395	0.051
incorrect	21	-0.193	0.011	-0.401	0.032
correct	21	-1.010	0.036	0.664	0.035
incorrect	22	-0.132	0.016	-0.231	0.030
correct	22	-1.079	0.035	0.581	0.031
incorrect	23	-0.114	0.019	-0.205	0.029
correct	23	-1.081	0.033	0.576	0.031
incorrect	24	-0.020	0.019	-0.055	0.025
correct	24	-1.341	0.036	0.278	0.023
incorrect	25	-0.105	0.019	-0.185	0.026
correct	25	-1.079	0.033	0.573	0.033
incorrect	26	-0.029	0.018	-0.060	0.024
correct	26	-1.499	0.037	0.161	0.016
incorrect	27	0.043	0.022	0.009	0.026
correct	27	-1.326	0.029	0.217	0.021
incorrect	28	0.026	0.023	-0.013	0.026
correct	28	-1.242	0.040	0.456	0.033
incorrect	29	-0.055	0.019	-0.096	0.025

Continued on next page

Continued from previous page

res	item	pos_neg	se_pn	neg_pos	se_np
correct	29	-1.233	0.031	0.412	0.030
incorrect	30	0.076	0.023	0.045	0.025
correct	30	-0.907	0.044	0.786	0.045
incorrect	31	-0.141	0.006	-0.259	0.017
correct	31	-0.969	0.038	0.727	0.040
incorrect	32	-0.064	0.021	-0.249	0.038
correct	32	-0.989	0.036	0.700	0.037
incorrect	33	-0.218	0.013	-0.352	0.025
correct	33	-1.065	0.036	0.594	0.033
incorrect	34	-0.057	0.018	-0.094	0.024
correct	34	-1.414	0.032	0.221	0.016
incorrect	35	-0.062	0.019	-0.113	0.025
correct	35	-1.247	0.029	0.359	0.022
incorrect	36	-0.037	0.020	-0.089	0.026
correct	36	-1.201	0.030	0.415	0.025
incorrect	37	0.022	0.020	-0.007	0.025
correct	37	-1.552	0.045	0.146	0.021
incorrect	38	-0.012	0.020	-0.045	0.025
correct	38	-1.413	0.024	0.183	0.014
incorrect	39	-0.028	0.020	-0.066	0.025
correct	39	-1.318	0.035	0.331	0.024
incorrect	40	0.072	0.023	0.038	0.026
correct	40	-0.961	0.041	0.712	0.039