Deep Generative Pattern-Set Mixture Models for Nonignorable Missingnes

Sahra Ghalebikesabi, Rob Cornish, Luke Kelly, Chris Holmes

July 11, 2021

Missingness Types

- **Missingness Completely at Random** (MCAR):

$$P(M|X) = P(M)$$

- **Missingness at Random** (MAR):

$$P(M|X) = P(M|X_{obs})$$

- **Missingness Not at Random** (MNAR)

Nonignorable Missingness Models [Little and Rubin, 2019]

- *Selection Models* [Little and Rubin, 2019]

$$p(X, M|\theta, \phi) = p(X|\theta)p(M|X, \phi),$$

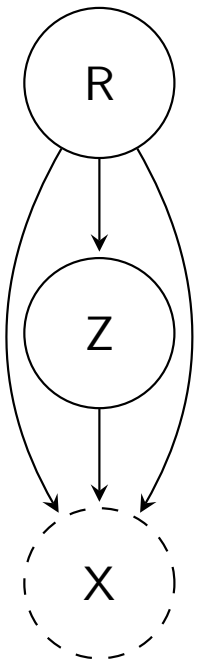
- *Pattern Mixture Models* [Little and Rubin, 2019]

$$p(X, M|\theta, \phi) = p(X|M, \theta)p(M|\phi).$$

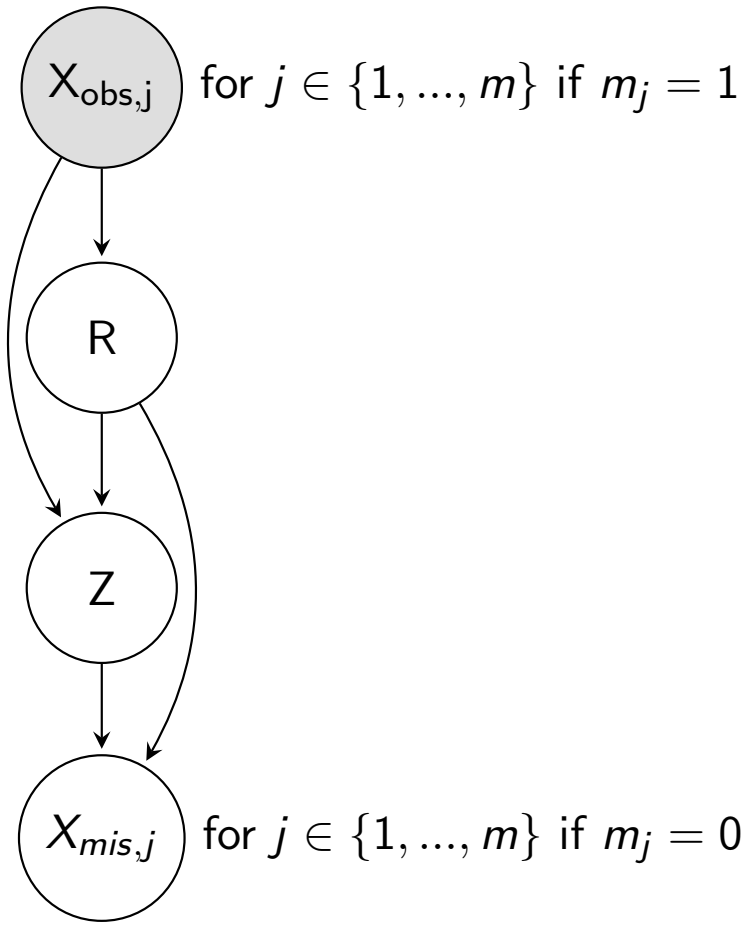
- *Pattern-Set Mixture Models* [Little, 1993]

$$p(X, R, M|\theta, \phi, \varphi) = p(X|R, \theta)p(M|X, R, \phi)p(R|\varphi)$$

HIVAE [Nazabal et al., 2020]

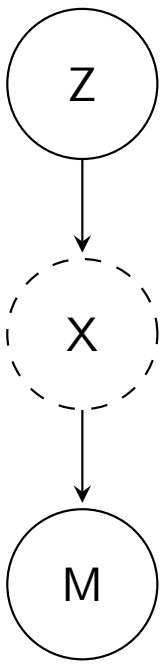


(a) Generative Model.

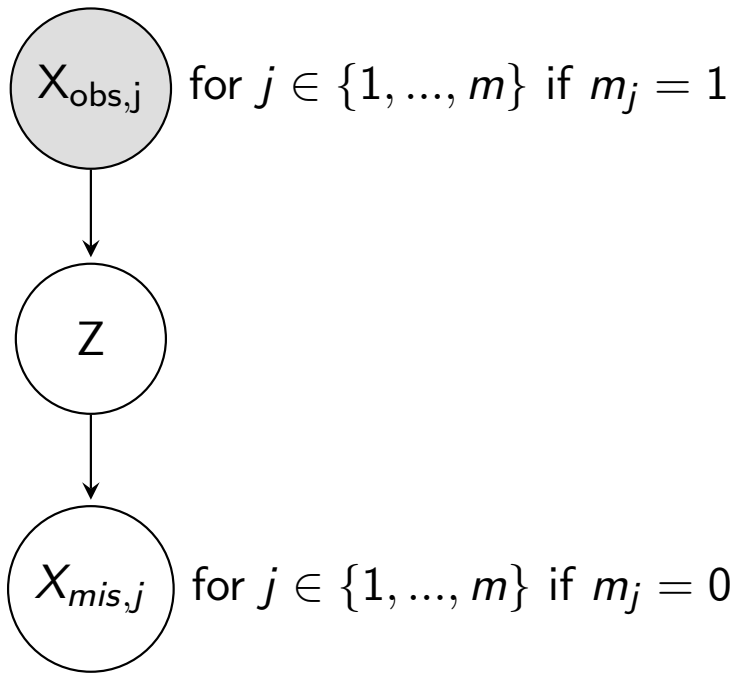


(b) Recognition Model.

not-MIWAE [Ipsen et al., 2021]

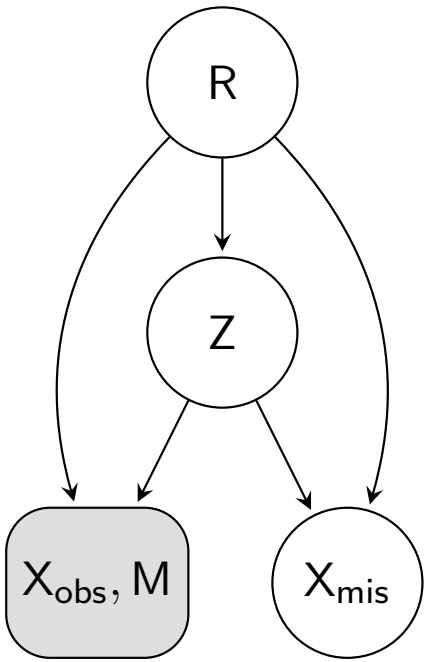


(a) Generative Model.

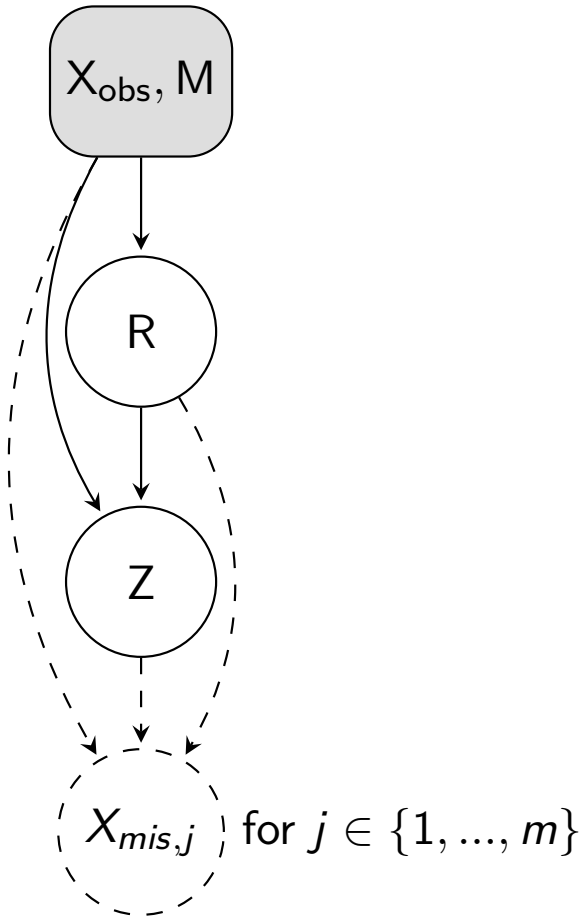


(b) Recognition Model.

Pattern-Set Mixture Model



(a) Generative Model.



(b) Recognition Model.

Probabilistic Semi-Supervision

Put simply, we sample each $x^i_{mis,j}$ from

$$y_{i,j} P_\theta(\widetilde{x}_{mis,j}) + (1 - y_{i,j}) \mathbb{1}(x^i_{obs,j}),$$

where y_j is an independent Bernoulli random variable with known success probability π' if $m_j = 1$, and with probability 1 otherwise

Results

Model	Adult		Letter		Wine	
	MCAR	MNAR	MCAR	MNAR	MCAR	MNAR
PSMVAE(a)	.2494	.4943	.0964	**.0835**	.0958	**.1034**
PSMVAE(b)	.2426	.5249	.0936	.0864	.0890	.1158
↳ K=10.000	**.2306**	.4981	**.0879**	.0854	**.0832**	.1069
↳ w/o M	.2450	**.4815**	.0941	.0908	.0885	.1167
DLGM	.2467	.5468	.0947	.0947	.0923	.1234
HIVAE	.2693	.4907	.1023	.0947	.0940	.1246
VAE	.2562	.5012	.1119	.1061	.1067	.1255
MIWAE	.2845	.6081	.1183	**.1024**	.1129	.1253
↳ K=10.000	**.2373**	.5872	**.1149**	.1242	**.0915**	.0803
not-MIWAE	.2374	**.5201**	.1153	.1192	.0928	**.0756**
GAIN	.2570	.5940	.1518	.1316	.1749	.1151
MICE	.2383	.5879	.1167	.1235	.0881	.0782

References

Roderick JA Little and Donald B Rubin. *Statistical analysis with missing data*, volume 793. John Wiley & Sons, 2019.

Roderick JA Little. Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association*, 88(421):125–134, 1993.

Alfredo Nazabal, Pablo M Olmos, Zoubin Ghahramani, and Isabel Valera. Handling Incomplete Heterogeneous Data using VAEs. *Pattern Recognition*, page 107501, 2020.

Niels Bruun Ipsen, Pierre-Alexandre Mattei, and Jes Frellsen. not-MIWAE: Deep Generative Modelling with Missing not at Random Data. In *International Conference on Learning Representations*, 2021.