Theoretical and Practical Considerations for Bridging Models of Learning and Assessment

Benjamin Deonovic, Gunter Maris

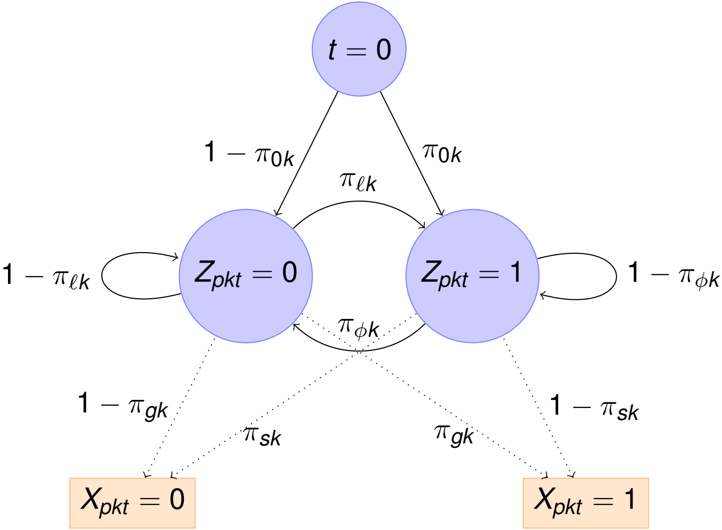
ACTNext

November 1-2, 2018

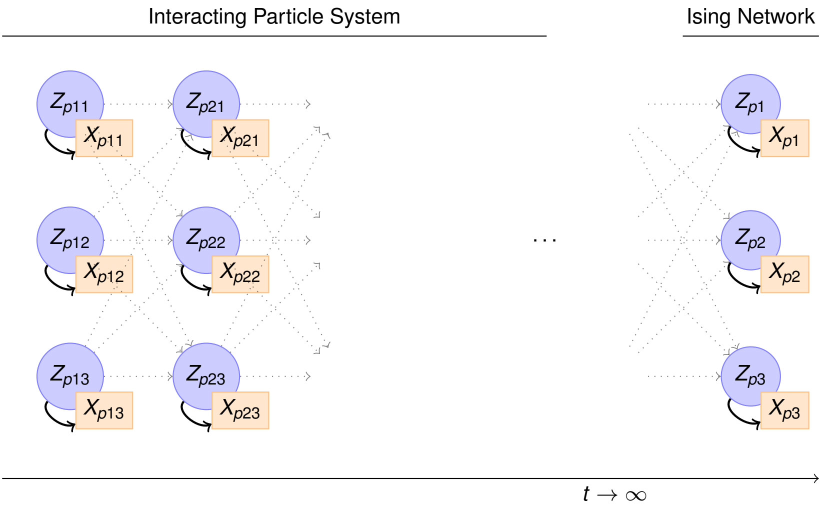
Learning meets Assessment

- Models for learning and models for assessment were developed and grew to utilize the salient features of their respective data sets.
- On the surface two popular and ubiquitous models in these fields, Bayesian Knowledge Tracing and Item Response Theory, appear quite different but we show there is an intimate relation between the two (Deonovic et al. 2018).
- Specifically we show that the equilibrium distribution of the response variable in the BKT model (under a particular reparameterization) follows a (4-1)PL IRT model.

Bayesian Knowledge Tracing



As a Network Model



Criticism

- No placeholder for education; learning as a ballistic rather than holistic model.
- Fundamental and well replicated psychological phenomenon not accounted for
  - positive manifold
  - Matthew effect

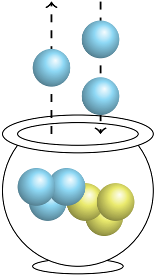
Help from Epidemiology: Model for Contagion

- In the early twentieth century, a very small number of discrete probability distributions were used in modeling (i.e. Poisson, Binomial) (Feller 1943).
- Studies carried out by various Biologists during this time period showed that these commonly used distributions did not adapt to all situations(Gurland 1958).
- Consequently, starting in 1920s, many biologists came up with novel probability distributions that were inspired by their specific fields of study. Some of these distributions were referred to as "contagious" distributions.

True Contagion

- Eggenberger-Pólya urn model (Eggenberger and Pólya 1923).
- The urn begins with a blue balls and b yellow balls. At time $t = 1, \ldots, n$ a ball is drawn at random from the urn and two balls of the same color are put back in.
- Let X_t be

$$X_t = \begin{cases} 1 & \text{if drawn ball is blue} \\ 0 & \text{otherwise} \end{cases}$$



True Contagion – Distribution

- Let $X = \sum_{t=1}^{n} X_t$ then Eggenberger and Pólya (1923) showed that in the limit

$$\lim_{n \to \infty} P(X = x) = \frac{\Gamma(a + x)}{x!\Gamma(a)} \left(\frac{b}{1 + b}\right)^{a} \left(\frac{1}{1 + b}\right)^{x}$$

which corresponds to a Negative Binomial Distribution.

Apparent Contagion

- Greenwood and Yule (1920) showed they could also fit data from biological phenomena better than the simple Poisson or Binomial by incorporating population features as a way to explain the clustering of events and the increase in variance.
- They introduce The Poisson-Gamma mixture distribution, i.e. a Poisson distribution where the intensity parameter is treated as a latent random variable from a Gamma distribution.
- This type of model was referred to as a "false" contagion or "apparent" contagion.

Poisson–Gamma Mixture

- Let $X|\lambda \sim \text{Poisson}(\lambda)$ and $\lambda \sim \text{Gamma}(r, (1-p)/p)$

$$
\begin{aligned}
f_X(x) &= \int_0^\infty f_{X|\lambda}(x|\lambda) f_\lambda(\lambda)\,\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\lambda^x}{x!} e^{-\lambda} \cdot \lambda^{r-1} \frac{e^{-\lambda(1-p)/p}}{\left(\frac{p}{1-p}\right)^r \Gamma(r)}\,\mathrm{d}\lambda \\
&= \frac{(1-p)^r p^{-r}}{x!\,\Gamma(r)} \int_0^\infty \lambda^{r+x-1} e^{-\lambda/p}\,\mathrm{d}\lambda \\
&= \frac{(1-p)^r p^{-r}}{x!\,\Gamma(r)}\, p^{r+x}\,\Gamma(r+x) \\
&= \frac{\Gamma(r+x)}{x!\;\Gamma(r)}\, p^k (1-p)^r.
\end{aligned}
$$

- This is the Negative Binomial distribution.

False Dichotomy

- However, the dichotomy between "true" contagion and "false" contagion is actually a false one.
- Coming from two opposing philosophical standpoints, both from an applied and a modeling perspective, Greenwood and Yule (1920) ("falsely contagious") and Eggenberger and Pólya (1923) ("truly contagious") ultimately arrived at the same distribution

Transitioning to Assessment

- The false dichotomy between these two perspectives also inspires another latent variable representation of the Eggenberger-Pólya urn model

$$
\begin{aligned}
P(X = x) &= \int_0^1 \binom{n}{x} \pi^x (1-\pi)^{n-x} \frac{\pi^{a-1}(1-\pi)^{b-1}}{B(a,b)} d\pi \\
&= \binom{n}{x} \frac{1}{B(a,b)} \int_0^1 \pi^{x+a-1}(1-\pi)^{n-x+b-1} d\pi \\
&= \binom{n}{x} \frac{B(x+a, n-x+b)}{B(a,b)}
\end{aligned}
$$

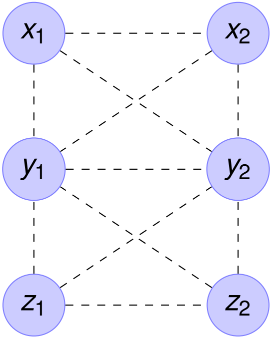
Network Pscyhometrics

- Past psychometric models for intelligence have focused on a factor analytic approach in which latent factors explain the variance in observed scores
- Spearman (1904) applied this methodology when developing his *g*-theory of intelligence and to explain the positive correlations between cognitive tasks.
- Recently, van der Maas et al. (2006) proposed an alternative view in which positive manifold is explained by the causal relationships between observed variables in a network.

Contagion as (not) learning

- Extending the network psychometric approach for decomposing ability as a network rather than a latent model two recent works construct a model for learning.
- Savi et al. (2018), in a model called wired cognition, utilizes the Ising model from statistical physics to describe a mechanism for how networks of skills and abilities for individual persons can grow
- Plak et al. (2018) extends the concepts of the Pólya urn to networks to describe a different mechanisms for the growth of abilities, one which is more tractable than the Ising model, yet retains many of the same desireable properties.

Pólya Networks



Joint Distribution

$$p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \frac{B(x_+ + y_+ + a, 2n_y - x_+ - y_+ + b)B(y_+ + 2 + a, 2n_y - y_+ - z_+ + b)}{B(y_+ + a, n_y - y_+ + b)B(a, b)}$$

$$= \int_0^1 \int_0^1 \left(\pi^{x_+}(1-\pi)^{n_x - x_+}\right) \frac{\pi^{y_+ + a - 1}(1-\pi)^{n_y - y + b - 1}}{B(y_+ + a, n_y - y_+ + b)} \times$$

$$\left(\gamma^{y_+}(1-\gamma)^{n_y - y_+}\right) \left(\gamma^{z_+}(1-\gamma)^{n_z - z_+}\right) \times$$

$$\frac{\gamma^{a-1}(1-\gamma)^{b-1}}{B(a, b)} d\pi d\gamma$$

Desirable Properties

- Describes the growth of ability and the acquisition of skills
- Has the positive manifold phenomenon built in
- Closed under marginalization
- Tractable

Future Directions

- What are testable hypothesis we can come up with from this model?

Acknowledgement

- Acknowledge
  - Gunter Maris
  - Simone Plak

- Thank you! Any questions?

References

- Deonovic, B., Yudelson, M., Bolsinova, M., Attali, M., and Maris, G. (2018). “Learning meets assessment”. In: *Behaviormetrika*.
- Eggenberger, F. and Pólya, G. (1923). “Über die Statistik verketteter Vorgänge”. In: *Zeitschrift für Angewandte Mathematik und Mechanik* 3.4, pp. 279–289.
- Feller, W. (1943). “On a General Class of "Contagious" Distributions”. In: *The Annals of Mathematical Statistics* 14.4, pp. 389–400.
- Greenwood, M. and Yule, G. U. (1920). “An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents”. In: *Journal of the Royal Statistical Society* 83.2, pp. 255–279.
- Gurland, J. (1958). “A Generalized Class of Contagious Distributions”. In: *Biometrics* 14.2, pp. 229–249.
- Gurland, J. (1963). *Some Families of Compund and Generalized Distributions*. MRC Technical Summary Report 380. Mathematics Research Center, United States Army.
- Plak, S., Maris, G. K. J., Waldorp, L. J., and Marsman, M. (2018). “The Contagion Mechanism in Cognitive Development: Introducing Pólya-Yule Network Models”. Master’s Thesis. University of Amsterdam.
- Pólya, G. (1930). “Sur quelques points de la théorie des probabilités”. fr. In: *Annales de l’institut Henri Poincaré* 1.2, pp. 117–161.

References (cont.)

- Savi, A. O., Marsman, M., van der Maas, H., and Maris, G. (2018). *The Wiring of Intelligence*.
- Spearman, C. (1904). “" General Intelligence," objectively determined and measured”. In: *The American Journal of Psychology* 15.2, pp. 201–292.
- van der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., and Raijmakers, M. E. (2006). “A dynamical model of general intelligence: the positive manifold of intelligence by mutualism.” In: *Psychological review* 113.4, p. 842.