Multivariate approximation in total variation

A. D. Barbour, M. J. Luczak and **A. Xia***

*Department of Mathematics and Statistics

The University of Melbourne, VIC 3010

29 May, 2015

Background Information

- Two corner-stone distributions
  - Continuous: normal and the central limit theorem, in the Kolmogorov distance
  - Discrete: Poisson and the law of small numbers, in the total variation

Poisson approximation

- The Stein–Chen method (Chen, 1975)
- Barbour and Hall (1984)

$$\frac{1}{32}\min\left\{\frac{1}{\lambda},1\right\}\sum_{i=1}^{n}p_i^2 \le d_{TV}(\mathcal{L}(W),\mathrm{Pn}(\lambda)) \le \frac{1-e^{-\lambda}}{\lambda}\sum_{i=1}^{n}p_i^2.$$

- The message: Poisson has only one parameter, so to improve on the quality of approximation, more parameters should be introduced.

Improvements

Translated Poisson (Röllin 2005), perturbation (Barbour and X. 1999, Barbour, Čekanavičius and X. 2007), polynomial birth-death (Brown and X. 2001), zero biasing with birth-death process (Goldstein and X. 2006), **discretised normal** (Roy 2003, Chen and Leong (2010), and Fang (2014)).

Multivariate approximation

- Normal: Gotze 1991, Goldstein and Rinott 1996, Rinott and Rotar (1996), ...
- Discrete version?
  - Multivariate Poisson (Barbour 1988, Roos 1999)
  - How to introduce more parameters?
    - Multivariate Poisson with negative correlation?
    - Transform a multivariate Poisson?
    - Discretise multivariate normal?
      - Stein's identity?
    - How to extract Stein's constants?

One dimension

- Stein's equation for Poisson(λ):

$$\lambda g(i+1) - ig(i) = f(i) - \mathrm{Pn}(\lambda)(f).$$

- Birth-death process interpretation (Barbour 1988): by taking $g(i) = h(i) - h(i-1)$,

$$\mathcal{A}h(i) := \lambda(h(i+1)-h(i))+i(h(i-1)-h(i)) = f(i)-\mathrm{Pn}(\lambda)(f).$$

  - $\mathcal{A}$ is the generator of the birth-death process with birth rate λ and each individual has lifetime of exp(1), independent of others.

– The solution of the Stein equation is

$$h(i) = -\int_0^\infty [\mathbb{E}f(Z_i(t)) - \mathrm{Pn}(\lambda)(f)]dt,$$

where Z_i is a birth-death process with generator $\mathcal{A}$ and initial value $Z_i(0) = i$.

– Stein's constants are from estimates of $\|\Delta h\| := \sup_i |h(i+1) - h(i)|$ and $\|\Delta^2 h\| := \sup_i |\Delta h(i+1) - \Delta h(i)|$.

Discrete CLT: Goldstein and X. 2006

$$\mathcal{A}h(i) := \alpha_i(h(i+1) - h(i)) + \beta_i(h(i-1) - h(i)),$$

where

$$\alpha_i = \begin{cases} \sigma^2, & i \geq \kappa, \\ \sigma^2 + \mu - i & i \leq \kappa - 1, \end{cases} \qquad \beta_i = \begin{cases} \sigma^2 + i - \mu & i \geq \kappa, \\ \sigma^2 & i \leq \kappa - 1. \end{cases}$$

- $\mathcal{A}$ is the generator of the bilateral birth-death processes with "birth" rates $\{\alpha_i : i \in \mathbf{Z}\}$ and "death" rates $\{\beta_i : i \in \mathbf{Z}\}$.
- Similarly, h can be represented in terms of $Z_i(t), t \geq 0$, the Markov process with generator $\mathcal{A}$ and initial value i.
- If $\mu - \sigma^2$ is an integer and $\kappa = \mu - \sigma^2$, the states $< \kappa$ are transient: translated Poisson.

Fang 2014

Works on

$$\sigma^2 f'(s) - (s-\mu)f(s) = h(s) - \mathbb{E}h(Z_{\mu,\sigma^2}),$$

where h is constant on $[i-0.5, i+0.5)$ for $i \in \mathbb{Z}$, so can extract Stein's factors for discretised normal from those of normal approximation.

Multivariate

- d-dimensional Ornstein-Uhlenbeck operator

$$\mathcal{A}h(w) = \frac{n}{2}\mathrm{Tr}(\sigma^2 D^2 h(w)) + Dh^T(w)A(w - nc),\ w \in \mathbb{R}^d$$

  on twice differentiable functions $h:\ \mathbb{R}^d \to \mathbb{R}$.

  - A is a matrix whose eigenvalues all have negative real parts.
  - σ^2 is a positive definite symmetric matrix.
  - The equilibrium distribution is $\mathcal{N}_d(nc, n\Sigma)$ with

$$A\Sigma + \Sigma A^T + \sigma^2 = 0.$$

    * $\Sigma = \int_0^\infty e^{At}\sigma^2 e^{A^T t} dt$.

Discrete version

- Consider $\mathcal{DN}_d(nc, n\Sigma)$ as discretised version of $\mathcal{N}_d(nc, n\Sigma)$: for $(i_1, \ldots, i_d) \in \mathbb{Z}^d$,

$$\begin{aligned} &\mathcal{DN}_d(nc, n\Sigma)\{(i_1, \ldots, i_d)\} \\ &= \mathcal{N}_d(nc, n\Sigma)([i_1 - 0.5, i_1 + 0.5) \times \cdots \times [i_d - 0.5, i_d + 0.5)). \end{aligned}$$

- Fix $c \in \mathbb{R}^d$. For $h: \ \mathbb{Z}^d \to \mathbb{R}$, define the generator

$$\mathcal{A}_n h(w) = \frac{n}{2} \mathrm{Tr}\left(\sigma^2 \Delta^2 h(w)\right) + \Delta h^T(w) A(w - nc), \ w \in \mathbb{Z}^d.$$

  - A, σ^2: as above.
  - $\Delta_j h(w) = h(w + e^{(j)}) - h(w)$, $\Delta^2_{jk} h(w) = \Delta_j(\Delta_k h)(w)$, $1 \le j, k \le d$.

Markov population process

- X_n has transition rates $ng^J(n^{-1}w)$ from w to $w+J$, $w \in \mathbb{Z}^d,\ J \in \mathcal{J}$, where $\mathcal{J}$ is a finite subset of $\mathbb{Z}^d$.
  - g^J is twice continuously differentiable.
  - Define $F(\xi) = \sum_{J \in \mathcal{J}} J g^J(\xi)$, we assume

    $$\frac{d\xi}{dt} = F(\xi)$$

    has an equilibrium point c so $F(c) = 0$ and $A = DF(c)$.
  - This Markov population process has the equilibrium distribution Π_n.

Modification of the Markov population process

- X_n has small chance to "drift away" from its "centre" nc but if it happens, ...

- $\mathcal{DN}_d(nc, n\Sigma)$ essentially distributes in a region of radius $n\delta$ around nc:

$$I_{n,\delta} = \left\{ \xi \in \mathbb{Z}^d : \sqrt{(\xi - nc)^T \Sigma^{-1} (\xi - nc)} \leq n\delta \right\}.$$

- We can focus on approximating any arbitrary $\mathbb{Z}^d$-valued random element W in $I_{n,\delta}$ first and then add in the errors outside this region

- We modify X_n to keep it in $I_{n,\delta}$

Modification (2)

- Modify X_n to X_n^δ with transition rates from w to $w+J$:

$$ng_\delta^J(n^{-1}w) = \begin{cases} ng^J(n^{-1}w), & \text{if } w - nc, w + J - nc \\ & \in I_{n,\delta}, \\ 0, & \text{otherwise.} \end{cases}$$

- If X_n^δ starts in $I_{n,\delta}$, then it stays in $I_{n,\delta}$.
- We assume conditions to ensure it is irreducible so it has the unique equilibrium distribution Π_n^δ

Π_n^δ approximation

- Generator

$$\mathcal{A}_n^\delta(w) = n \sum_{J \in \mathcal{J}} g_\delta^J(n^{-1}w)\{h(w+J) - h(w)\}, \ w \in \mathbb{Z}^d$$

- Stein's equation:

$$\mathcal{A}_n^\delta h_B(w) = 1_B(w) - \Pi_n^\delta(B), \ B \subset I_{n,\delta}.$$

- The solution h_B can be written as

$$h_B(w) = -\int_0^\infty (\mathbb{P}(X_n^\delta(t) \in B | X_n^\delta(0) = w) - \Pi_n^\delta(B))dt.$$

- Estimates:

$$\begin{aligned} \|h_B(w)\| &\le c_0 \ln n, \\ \|\Delta h_B(w)\| &\le c_1 n^{-1/2} \ln n; \\ \|\Delta^2 h_B(w)\| &\le c_2 n^{-1} \ln n \end{aligned}$$

for w in a $n\delta/4$-neighbourhood of nc and n sufficiently large.

- For any $\mathbb{Z}^d$ valued random element W, $B \subset I_{n,\delta}$, we have

$$
\begin{aligned}
&\mathbb{P}(W \in B) - \Pi_n^\delta(B) \\
&= \mathbb{E}\{(1_B(W) - \Pi_n^\delta(B))1_{W \in I_{n,\delta'}}\} - \Pi_n(B)\mathbb{P}(W \notin I_{n,\delta'}) \\
&= \mathbb{E}\{\mathcal{A}_n^\delta h_B(W) 1_{W \in I_{n,\delta'}}\} - \Pi_n(B)\mathbb{P}(W \notin I_{n,\delta'}),
\end{aligned}
$$

hence

$$
\begin{aligned}
& d_{TV}(\mathcal{L}(W), \Pi_n^\delta) \\
\leq \quad & \sup_{B \subset I_{n,\delta}} \left| \mathbb{E}\{\mathcal{A}_n^\delta h_B(W) 1_{W \in I_{n,\delta'}}\} \right| + \underbrace{\mathbb{P}(W \notin I_{n,\delta'})}_{\text{Bienaymé-Chebyshev ineq.}} .
\end{aligned}
$$

Multivariate in total variation

Assume

- $\mathbb{E}|W - nc|^2 \le Vn$,
- $d_{TV}(\mathcal{L}(W), \mathcal{L}(W + e^{(j)})) \le \varepsilon_1$, $1 \le j \le d$,
- and

$$|\mathbb{E}\{\mathcal{A}_n h(W)\} 1_{[|W-nc| \le c_1 n]}| \\ \le c_2 \underbrace{\|h\|}_{\sim c \ln n} \varepsilon_{20} + c_3 \underbrace{n^{1/2} \|\Delta h\|}_{\sim c \ln n} \varepsilon_{21} + c_4 \underbrace{n \|\Delta^2 h\|}_{\sim c \ln n} \varepsilon_{22},$$

where h is the solution for $\mathcal{A}_n^\delta$, $\|h\|$, $\|\Delta h\|$ and $\|\Delta^2 h\|$ are estimates around an ηn neighbourhood of nc.

Then

$$d_{TV}(\mathcal{L}(W), \mathcal{DN}_d(nc, n\Sigma))$$
$$\leq O\left\{\ln n \left(\underbrace{n^{-1/2}}_{\text{diff of } \mathcal{DN}_d \text{ and } \Pi_n} + \varepsilon_1 + \varepsilon_{20} + \varepsilon_{21} + \varepsilon_{22} \right)\right\}.$$