Berry-Esseen Bounds for General Non-normal Approximation with Unbounded Exchangeable Pairs

Qi-Man Shao

Southern University of Science and Technology

Based on a joint work with Songhao Liu and Hao Shi

Let $W := W_n$ be the random variable of interest. Our goals are

(i) Identify the limiting distribution of W_n;

(ii) Suppose that $W_n \xrightarrow{d.} Y$. Find the error of the approximation. Especially, what is the Kolmogorov distance

$$\sup_x |P(W_n \le x) - P(Y \le x)|?$$

Let Y be a random variable with pdf $p(y)$, $-\infty < y < \infty$. Assume that $p(y) > 0$ and p is differentiable. Observe that

$$E\left\{\frac{\big(f(Y)p(Y)\big)'}{p(Y)}\right\} = \int_{-\infty}^{\infty} (f(y)p(y))' dy = 0. \tag{1}$$

- Stein's identity and equation (Stein, Diaconis, Holmes, Reinert (2004)):
  - Stein's identity:

$$E\left\{\frac{\big(f(Y)p(Y)\big)'}{p(Y)}\right\} = Ef'(Y) + Ef(Y)p'(Y)/p(Y) = 0.$$

  - If $Y \sim N(0,1)$, then $p'(Y)/p(Y) = -Y$ and the Stein identity reduces to

$$Ef'(Y) - EYf(Y) = 0.$$

- Stein's equation:

$$(f(y)p(y))'/p(y) = h(y) - Eh(Y)$$

- Stein's solution:

$$\begin{aligned} f(y) &= 1/p(y)\int_{-\infty}^{y}(h(t)-Eh(Y))p(t)dt \\ &= -1/p(y)\int_{y}^{\infty}(h(t)-Eh(Y))p(t)dt. \end{aligned}$$

- Properties of the solution (Chatterjee - Shao (2011)):

Let h be a measurable function and f_h be the Stein's solution. Under some regularity conditions on p

$$\|f_h\| \le C\|h\|, \;\; \|f_h'\| \le C\|h\|,$$

$$\|f_h\| \le C\|h'\|, \;\; \|f_h'\| \le C\|h'\|, \;\; \|f_h''\| \le C\|h'\|$$

▶ Identify the limiting distribution

Let $W := W_n$ be the random variable of interest. Our goal is to identify the limiting distribution of W_n with an error of approximation.

▶ Identify the limiting distribution

Let $W := W_n$ be the random variable of interest. Our goal is to identify the limiting distribution of W_n with an error of approximation.

- Exchangeable pair approach:

Let (W, W') be an exchangeable pair. Put $\Delta = W - W'$. Assume that

$$E(\Delta \mid W) = \lambda(g(W) + r_1(W)) \tag{2}$$

and

$$E(\Delta^2|W) = 2\lambda(1 + r_2(W)). \tag{3}$$

Let

$$p(y) = c_0 \exp\Big(- \int_0^y g(t)dt \Big),$$

where c_0 is the normalizing constant.

Let Y have pdf $p(y)$. Assume (2) and (3) are satisfied.

Chatterjee - Shao (2011) (also see Eichelsbacher - Löwe (2010)): Under some regularity conditions on g

- Assume that

$$E(|r_1| + |r_2|) + \frac{1}{\lambda} E|\Delta|^3 \to 0.$$

Then

$$W \xrightarrow{d.} Y .$$

Let Y have pdf $p(y)$. Assume (2) and (3) are satisfied.

Chatterjee - Shao (2011) (also see Eichelsbacher - Löwe (2010)):
Under some regularity conditions on g

- Assume that

$$E(|r_1| + |r_2|) + \frac{1}{\lambda} E|\Delta|^3 \to 0.$$

  Then

$$W \xrightarrow{d.} Y .$$

- If $|\Delta| \leq \delta$, then

$$\begin{aligned}
&\sup_x |P(W \geq x) - P(Y \geq x)| \\
&= O(1)\left(E(|r_1| + |r_2|) + \frac{1}{\lambda} E|\Delta|^3 + \delta\right)
\end{aligned}$$

Shao - Zhang (2019): Under some regularity conditions on g, we have

$$\sup_x |P(W \geq x) - P(Y \geq x)| = O(1)\Big(E(|r_1| + |r_2|) + \frac{1}{\lambda} E|E(\Delta\,\Delta^*|W)|, \Big)$$

where $\Delta^* = \Delta^*(W, W')$ is any random variable satisfying $\Delta^*(W, W') = \Delta^*(W', W)$ and $\Delta^* \geq |\Delta|$.

Recall

$$E(\Delta \mid W) = \lambda(g(W) + r_1(W)) \tag{2}$$

and

$$E(\Delta^2|W) = 2\lambda(1 + r_2(W)). \tag{3}$$

- Note: Condition (2) is always satisfied, while (3) is an assumption.

- Question: Can we have a more general result when

$$E(\Delta^2|W) = 2\lambda(v(W) + r_2(W))\ ? \tag{4}$$

▶ Towards a more general approximation

Let (W, W') be an exchangeable pair, let $\Delta = W - W'$. Assume

$$E(\Delta \mid W) = \lambda(g(W) + r_1(W)) \tag{2}$$

and

$$E(\Delta^2|W) = 2\lambda(v(W) + r_2(W)). \tag{4}$$

- Question: How to identify the limiting density function p?

Recall the Stein identity: If $Y \overset{pdf}{\sim} p$, then

$$E\left\{\frac{\big(f(Y)p(Y)\big)'}{p(Y)}\right\} = 0.$$

Clearly, it is also true that

$$E\left\{\frac{\big(f(Y)p(Y)v(Y)\big)'}{p(Y)}\right\} = 0 \quad \text{for any } v$$

i.e.

$$Ef'(Y)v(Y) + Ef(Y)\frac{(p(Y)v(Y))'}{p(Y)} = 0 \tag{5}$$

Also recall the fact that

$$
\begin{aligned}
0 &= E\Delta(f(W)+f(W')) \\
&= 2E\big(f(W)E(\Delta \mid W)\big) - E\Delta \int_{-\Delta}^{0} f'(W+t)dt \\
&= 2E\big(f(W)E(\Delta \mid W)\big) - E\big(f'(W)E(\Delta^2 \mid W)\big) \\
&\quad -E\Delta \int_{-\Delta}^{0} (f'(W+t)-f'(W))dt \\
&= -2\lambda\Big(Ef'(W)v(W) - Ef(W)g(W) + \text{error}\Big)
\end{aligned}
$$

Compared with (5), the limiting density function p should satisfy

$$
\frac{(p(y)v(y))'}{p(y)} = -g(y)
$$

or

$$
p(y) = \frac{c_0}{v(y)} \exp\Big(-\int_0^y \frac{g(t)}{v(t)}dt\Big). \tag{6}
$$

Assume (2) and (4) are satisfied. Let $Y \overset{pdf}{\sim} p$.

- Shao and Zhang (2016) (Also see Döbler (2015)): Under some regularity conditions on g and v, if

$$E|r_1(W)| + E|r_2(W)|/v(W) + \frac{1}{\lambda}E|\Delta|^3/v(W) \to 0,$$

then

$$W \xrightarrow{d.} Y$$

- When $|\Delta| \leq \delta$, a Kolmogorov distance was also proved.

▶ Berry-Esseen Bounds for General Approximation with Unbounded Exchangeable Pairs

Let (W, W') be an exchangeable pair, let $\Delta = W - W'$. Assume

$$E(\Delta \mid W) = \lambda(g(W) + r_1(W)), \;\; E(\Delta^2|W) = 2\lambda(v(W) + r_2(W)).$$

Let $Y \overset{pdf}{\sim} p$, where $p(y) = \frac{c_0}{v(y)} \exp\Big(- \int_0^y \frac{g(t)}{v(t)} dt\Big)$.

Liu - Shao - Shi (2022): Under some regularity conditions on g and v,

$$\begin{aligned} & |P(W \geq x) - P(Y \geq x)| \\ = \;\; & O(1)\Big(E|r_1| + E\frac{|r_2|}{v(W)} + \frac{1}{\lambda}E\frac{|E(\Delta\,\Delta^*|W)|}{v(W)} + \frac{1}{\lambda}E|\Delta|^3 U(W, W')\Big), \end{aligned}$$

where $\Delta^* \geq |\Delta|$, $U(W, W') = \sup_{t \in (W, W')} |v'(t)|/v^2(t)$.

- Pólya-Eggenberger Urn Model

An urn contains r red balls and w white balls at the beginning. A ball is drawn from the urn at random and this ball with other c balls of the same color is returned to the urn.
Let $X_j = 1$ if the jth drawn ball is red and $X_j = 0$ otherwise, and let

$$W_n = \frac{1}{n}\sum_{j=1}^{n} X_j.$$

It is known that

$$W_n \xrightarrow{d.} Y \sim \mathcal{B}(a, b),$$

where $a = r/c,\ b = w/c$, $\mathcal{B}(a, b)$ is the beta distribution with parameters a, b, i.e, the pdf is given by

$$p(y) = c_0 x^{a-1}(1-x)^{b-1},\ 0 < x < 1.$$

- Goldstein-Reinert (2013):

$$\sup_{\|h'\|\leq 1} |Eh(W_n) - Eh(Y)| = O(1/n)$$

- Liu - Shao - Shi (2022):

$$\sup_x |P(W_n \leq x) - P(Y \leq x)| = O(n^{-\min(1,a,b)})$$

The bound is sharp.

Construction of Exchangeable pair:

Assume $W = W(X_1, \cdots, X_n)$. For given $\{X_j, j \neq i\}$, let X_i' be a random sample from the conditional distribution of X_i given $\{X_j, j \neq i\}$. Let I be a uniform random variable over $\{1, \cdots, n\}$ independent of all other random variables. Define $W' = W(X_1, \cdots, X_I', \cdots, X_n)$. Then (W, W') is an exchangeable pair.

Outline of the proof. As proved in Döbler (2015),

$$E(\Delta|W) = \frac{1}{n(a+b+n-1)}\left((a+b)W - a\right)$$

$$E(\Delta^2|W) = \frac{2}{n(a+b+n-1)}\left\{W(1-W) + \frac{b-a}{2n}W + \frac{a}{2n}\right\}.$$

So,

$$\lambda = \frac{1}{n(a+b+n-1)}, \quad g(w) = (a+b)w - a, \quad v(w) = w(1-w).$$

▶ Mean-field Heisenberg Model

Let $\sigma_i \in \mathbb{S}^2$, $\mathbb{S}^2$ is three-dimensional sphere. The joint density function of $(\sigma_1, \cdots, \sigma_n)$ is given by

$$dP_{n,\beta} = C_{n,\beta} \exp\left(\frac{\beta}{2n} \sum_{1 \le i,j \le n} \langle \sigma_i, \sigma_j \rangle\right),$$

where $\beta > 0$.

Kirkparick - Meckes (2013, 2016):

(i) For $0 < \beta < 3$,

$$\frac{(3-\beta)^{1/2}}{\sqrt{n}} \sum_{j=1}^{n} \sigma_j \xrightarrow{d.} N(0, I_3)$$

(ii) For $\beta > 3$,

$$\sqrt{n}\left(\frac{\beta^2}{n^2\kappa^2} |\sum_{j=1}^{n} \sigma_j|^2 - 1\right) \xrightarrow{d.} N(0, B^2),$$

where κ is the solution to the equation $x = \beta(\coth(x) - 1/x)$.

(iii) For $\beta = 3$, let

$$W_n = \frac{\beta^2}{n^{3/2}} |\sum_{j=1}^{n} \sigma_j|^2$$

Then $W_n \xrightarrow{d.} Y$, where the density function of Y is given by

$$p(y) = c_0 y^{1/2} e^{-y^2/180}, \quad y > 0$$

- Shao-Zhang (2019): For $\beta > 3$, give a $O(n^{-1/2})$ Berry-Esseen bound

- Liu-Shao-Shi (2022): For $\beta = 3$,

$$\sup_x |P(W_n \le x) - P(Y \le x)| = O\Big(\frac{1}{\sqrt{n}}\Big)$$

by showing that

$$\begin{aligned} E(\Delta|W) &= \frac{6}{n^{3/2}}\Big(\frac{W^2}{45} - 3 + r_1\Big), \\ E(\Delta^2|W) &= \frac{12}{n^{3/2}}\big(2W + r_2\big). \end{aligned}$$

- Person's χ^2-Test

For n independent trials, with each trial leading to a unique classification over m classes. Let $(p_1, p_2, \cdots, p_m)$ be the nonzero classification probability. Let $(U_1, U_2, \cdots, U_m)$ be the observed numbers in each class. The Pearson's χ^2-test statistic is given by

$$W_n = \sum_{j=1}^{m} \frac{(U_j - np_j)^2}{np_j}.$$

It is well-known that

$$W_n \xrightarrow{d.} Y \sim \chi^2_{m-1}.$$

- Yarnold (1972):

$$\sup_x |P(W_n \le x) - P(Y \le x)| = O\Big(\frac{1}{n^{(m-1)/m}}\Big) \qquad (7)$$

- Götze - Ulyanov (2003): For $m \ge 6$,

$$\sup_x |P(W_n \le x) - P(Y \le x)| = O\Big(\frac{1}{n}\Big) \qquad (8)$$

- Gaunt-Pickett-Reinert (2017): For $h \in \mathcal{C}^5$,

$$|Eh(W_n) - Eh(Y)| = O\Big(\frac{1}{n}\Big)$$

- Liu-Shao-Shi (2022) (also Mann (1997)): Using Stein's method, for $m \ge 2$

$$\sup_x |P(W_n \le x) - P(Y \le x)| = O\Big(\frac{1}{\sqrt{n}}\Big)$$

- Open question: Can one use Stein's method to recover results (7) and (8)?


Thank You