Exchangeable Pairs and Multivariate Normal Approximation.

Elizabeth Meckes

Cornell Probability Summer School - 2007.

Stein's method: the characterizing operator

Given a probability measure $\mathbb{P}_0$ on $\mathbb{R}$, a <u>characterizing operator</u> T_0 for $\mathbb{P}_0$ is an operator on real-valued functions s.t.

1. $\int_{\mathbb{R}} T_0 f(x)\, d\mathbb{P}_0(x) = 0 \quad \forall f$ s.t. $\int_{\mathbb{R}} |T_0 f(x)|\, d\mathbb{P}_0(x) < \infty$.

2. If $\mathbb{P}$ is another p.m. on $\mathbb{R}$ s.t.

   $\int_{\mathbb{R}} T_0 f(x)\, d\mathbb{P}(x) = 0 \quad \forall f$ with $\int_{\mathbb{R}} |T_0 f(x)|\, d\mathbb{P}(x) < \infty$,

   then $\mathbb{P} = \mathbb{P}_0$.

<u>Idea</u>: Starting with a random variable W of interest and a probability measure $\mathbb{P}_0$ conjectured to approximate W well, find a characterizing operator T_0 for $\mathbb{P}_0$ and estimate $|\mathbb{E}\, T_0 f(W)|$ for a large class of f.

<u>Why</u> <u>should</u> this be enough?

Try to solve

$$T_0 U_0 = I - \mathbb{E}_0$$

(i.e., given f, find $U_0 f$ s.t. $T_0 U_0 f(x) = f(x) - \mathbb{E}_0 f$)

If you can do this, then

$$|\mathbb{E}\, T_0 U_0 f(W)| = |\mathbb{E} f(W) - \mathbb{E}_0 f|.$$

Note:

- $d_{T.V.}(X,Y) = \sup_f \frac{1}{2}|\mathbb{E}f(X) - \mathbb{E}f(Y)|$

  total variation

where the supremum is over continuous f vanishing at ∞ and bounded by 1.

- $d_W(X,Y) = \sup_f |\mathbb{E}f(X) - \mathbb{E}f(Y)|$

  Wasserstein

  where the supremum is over f with Lipschitz constant bd. by 1.

- $d_{L^*}(X,Y) = \sup_f |\mathbb{E}f(Y) - \mathbb{E}f(X)|$

  dual-Lipschitz

  where the class of f is bd. by 1 and 1-Lipschitz.

$\Rightarrow$ estimate $|\mathbb{E}\, T_0 U_0 f(W)|$ in terms of bounds on f and its derivatives to get rates of convergence in metrics like these.

Characterizing operators: Examples

* If $\mathbb{P}_0$ has a density w.r.t. Lebesgue measure on $\mathbb{R}$ or a subset of $\mathbb{R}$, then T_0 is usually a differential operator.

  If $\mathbb{P}_0$ is a measure on a discrete set (e.g $\mathbb{Z}$, $\mathbb{N}$, discrete circle ...) T_0 is usually a difference operator

- Gaussian: $\mathbb{P}_0 = \dfrac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx$ on $\mathbb{R}$

$$T_0 f(x) = f'(x) - x f(x)$$

- Exponential $\mathbb{P}_0 = e^{-x}\,dx$ on $\mathbb{R}^+$

$$T_0 f(x) = f'(x) - f(x)$$

- Poisson: $\mathbb{P}_0(n) = e^{-\lambda}\dfrac{\lambda^n}{n!}$ on $\mathbb{Z}^+$

$$T_0 f(k) = \lambda f(k+1) - k f(k).$$

So what??

Instead of trying to estimate $|\mathbb{E}f(W) - \mathbb{E}_0 f|$, now we want to estimate

$$|\mathbb{E}\, T_0 U_0 f(W)|.$$

How is that any easier?

<u>Exchangeable</u> <u>Pairs</u>:

From W, make a "small random change" to get W', such that (W, W') is an exchangeable pair: $(W, W') \overset{d}{=} (W', W)$

Use the pair (W, W') to approximate derivatives by difference quotients, and use the exchangeability and the way you got W' from W (possibly by conditioning on W) to analyze what you get.

Example: the classical C.L.T.

Let $\{X_i\}_{i=1}^n$ be i.i.d., $\mathbb{E}X_1 = 0$, $\mathbb{E}X_1^2 = 1$.

Let $W = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$. $\mathbb{P}_0 = \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$ on $\mathbb{R}$,

$$T_0 f(x) = f'(x) - x f(x)$$

$$U_0 f(x) = e^{x^2/2} \int_{-\infty}^{x} [f(t) - \mathbb{E}f(Z)] e^{-t^2/2} dt$$

1. $\mathbb{E}\, T_0 f(Z) = 0$ (integration by parts)
2. $T_0 U_0 f(x) = f(x) - \mathbb{E}f(Z)$ (easy to check)

A little calculus shows:

- $\|U_0 f\|_\infty \leq \sqrt{\frac{\pi}{2}} \|f - \mathbb{E}f(Z)\|_\infty$
- $\|(U_0 f)'\|_\infty \leq 2 \|f - \mathbb{E}f(Z)\|_\infty$
- $\|(U_0 f)''\|_\infty \leq 2 \|f'\|_\infty$.

We'll bound $|\mathbb{E}\, T_0 g(W)|$ (using exchangeable pairs) in terms of g, then use these bounds to take $g = U_0 f$ and get bounds in terms of f.

Exchangeable pair: pick $I \in \{1, \ldots, n\}$ uniformly, independent of $\{X_i\}_{i=1}^n$, and replace X_I with an independent copy X_I'. Then

$$W' = W - \frac{1}{\sqrt{n}} X_I + \frac{1}{\sqrt{n}} X_I' .$$

We want to estimate

$$|\mathbb{E}\, T_0 g(W)| = |\mathbb{E}(g'(W) - W g(W))|.$$

Start from (W, W') and try to get here:

$$\begin{aligned}
0 &= \mathbb{E}\left[(W'-W)(g(W')+g(W))\right] \\
&= \mathbb{E}\left[(W'-W)(g(W')-g(W)) + 2g(W)(W'-W)\right] \\
&= \mathbb{E}\left[(W'-W)^2 \underset{\sim}{g'(W)} + R + 2\underset{\sim}{g(W)}(W'-W)\right] \\
&= \mathbb{E}\left[g'(W)\,\mathbb{E}[(W'-W)^2 \mid W] - g(W)\,\mathbb{E}[2(W-W') \mid W] + R\right]
\end{aligned}$$

$$\mathbb{E}[W-W' \mid W] = \frac{1}{\sqrt{n}}\mathbb{E}[X_I - X_I' \mid W]$$
$$= \frac{1}{n^{3/2}} \sum_{i=1}^{n} \mathbb{E}[X_i - X_i' \mid W]$$
$$= \frac{1}{n} W$$

$$\mathbb{E}[(W'-W)^2 \mid W] = \frac{1}{n}\mathbb{E}[(X_I' - X_I)^2 \mid W]$$
$$= \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{E}[(X_i')^2 - 2X_i X_i' + X_i^2 \mid W]$$
$$= \frac{1}{n^2} \sum_{i=1}^{n} \left[1 + \mathbb{E}[X_i^2 \mid W]\right]$$
$$= \frac{2}{n} + \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{E}[X_i^2 - 1 \mid W]$$

$$\Rightarrow 0 = \mathbb{E}\left[\left(\frac{2}{n} + \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}[X_i^2 - 1 \mid W]\right) g'(W) - \frac{2}{n} W g(W) + R\right]$$

$$\Rightarrow 0 = \mathbb{E}\left[g'(W) - Wg(W) + g'(W)\left(\frac{1}{2n}\sum_{i=1}^{n}\mathbb{E}[X_i^2 - 1 \mid W]\right) + \frac{n}{2}R\right]$$

$$\Rightarrow \mathbb{E}T_0 g(W) = \mathbb{E}\left[-g'(W)\left(\frac{1}{2n}\sum_{i=1}^{n}\mathbb{E}[X_i^2 - 1 \mid W]\right) - \frac{n}{2}R\right]$$

$$\mathbb{E}\left|\frac{1}{2n}\mathbb{E}\left[\sum_{i=1}^{n}(X_i^2-1)\,\middle|\,W\right]\right|$$

$$\leq \frac{1}{2n}\mathbb{E}\left|\sum_{i=1}^{n}(X_i^2-1)\right|$$

$$\leq \frac{1}{2n}\left(\mathbb{E}\left[\sum_{i,j}(X_i^2-1)(X_j^2-1)\right]\right)^{1/2}$$

$$= \frac{1}{2n}\left(\mathbb{E}\left[\sum_i(X_i^4+1)\right]\right)^{1/2}$$

$$= \frac{1}{2\sqrt{n}}\sqrt{\mathbb{E}X_1^4-1}$$

Now, $R \doteq (W'-W)^2 g'(W) - (W'-W)(g(W')-g(W))$

$$= -(W'-W)\left[g(W')-g(W)-g'(W)(W'-W)\right]$$

$$\Rightarrow |R| \leq \frac{1}{2}\|g''\|_\infty |W'-W|^3$$

$$\mathbb{E}|W'-W|^3 = \frac{1}{n^{3/2}}\mathbb{E}|X_I'-X_I|^3$$

$$\leq \frac{8}{n^{3/2}}\mathbb{E}|X_1|^3$$

$$\Rightarrow \frac{n}{2}\mathbb{E}|R| \leq \frac{2}{\sqrt{n}}\|g''\|_\infty \mathbb{E}|X_1|^3$$

$$\Rightarrow \ |\mathbb{E}\, T_0 U_0 f(W)|$$

$$\le \frac{1}{2\sqrt{n}} \sqrt{\mathbb{E} X_1^4 - 1}\ \|(U_0 f)'\|_\infty + \frac{2}{\sqrt{n}} \mathbb{E}|X_1|^3 \|(U_0 f)''\|_\infty$$

$$\le \frac{1}{2\sqrt{n}} \sqrt{\mathbb{E} X_1^4 - 1} \cdot 2 \| f - \mathbb{E} f(Z)\|_\infty$$

$$+ \frac{2}{\sqrt{n}} \mathbb{E}|X_1|^3 \cdot 2 \| f' \|_\infty$$

$$\Rightarrow \ d_{L^*}(W, Z) \le \frac{1}{\sqrt{n}} \left[2\sqrt{\mathbb{E} X_1^4 - 1} + 4\,\mathbb{E}|X_1|^3 \right].$$

Example: Poisson-binomial trials

$\mathbb{P}_0(n) = e^{-\lambda} \dfrac{\lambda^n}{n!}$ for $n \ge 0$.

$$T_0 f(k) = \lambda f(k+1) - k f(k)$$

$$U_0 f(k) = \frac{(k-1)!}{\lambda^k} \sum_{\ell=0}^{k-1} \frac{\lambda^\ell}{\ell!} (f(\ell) - \mathbb{E}_0 f)$$

Easy to check:

1. $\mathbb{E}_0 T_0 f = \sum_{k=0}^{\infty} T_0 f(k) \dfrac{\lambda^k}{k!} e^{-\lambda} = 0$

2. $T_0 U_0 f(j) = f(j) - \mathbb{E}_0 f.$

Boundedness (Barbour & Eagleson):

For $f : \mathbb{Z}^+ \to [0,1]$,

$$|U_0 f(j)| \leq \min\{1, 1.4\lambda^{-1/2}\}$$

$$|U_0 f(j+1) - U_0 f(j)| \leq \frac{1-e^{-\lambda}}{\lambda}$$

Let $X_1, \ldots, X_n$ be i.i.d. with $\mathbb{P}(X_1 = 1) = p$, $\mathbb{P}(X_1 = 0) = 1-p$. Let $\lambda = pn$ and $W = \sum_{i=1}^n X_i$.

Want to estimate

$$|\mathbb{E} T_0 f(W)| = |\mathbb{E}[\lambda f(W+1) - W f(W)]|.$$

Same idea as last time. Exchangeable pair: pick $I \in \{1, \ldots, n\}$ at random, independent of $\{X_i\}$, and replace X_I with an independent copy X_I' : $W' = W - X_I + X_I'$.

Again, we'll start by writing down an anti-symmetric expression in W, W'.

$$0 = \mathbb{E}\left[f(W')\mathbb{1}(W = W'-1) - f(W)\mathbb{1}(W' = W-1)\right]$$

$$= \mathbb{E}\left[f(W+1)\mathbb{1}(W' = W+1) - f(W)\mathbb{1}(W' = W-1)\right]$$

$$= \mathbb{E}\left[\mathbb{P}[W' = W+1 \mid W] f(W+1) - \mathbb{P}[W' = W-1 \mid W] f(W)\right]$$

$$\mathbb{P}[W' = W+1 \mid W] = \left(\frac{n-W}{n}\right)p = \frac{\lambda}{n} - \frac{pW}{n}.$$

$$\mathbb{P}[W' = W-1 \mid W] = \left(\frac{W}{n}\right)(1-p) = \frac{W}{n} - \frac{pW}{n}.$$

$$\Rightarrow 0 = \mathbb{E}\left[\left(\frac{\lambda}{n} - \frac{pW}{n}\right)f(W+1) - \left(\frac{W}{n} - \frac{pW}{n}\right)f(W)\right]$$

$$\Rightarrow 0 = \mathbb{E}\left[\lambda f(W+1) - Wf(W) - pW(f(W+1) - f(W))\right].$$

$$\Rightarrow \mathbb{E}\left[\lambda f(W+1) - Wf(W)\right] = \mathbb{E}\left[T_o f(W)\right]$$

$$= p\,\mathbb{E}\left[W(f(W+1) - f(W))\right].$$

$$|\mathbb{E}\, T_o U_o g(W)| \le p \sup_k |U_o g(k+1) - U_o g(k)|\, \mathbb{E}[W]$$

$$\le \frac{\lambda}{n}\left(\frac{1-e^{-\lambda}}{\lambda}\right)\cdot \lambda$$

$$= \frac{\lambda(1-e^{-\lambda})}{n}, \quad \text{for } g : \mathbb{Z}^+ \to [0,1].$$

Abstract Normal Approximation Theorem

Typically, one doesn't go through the whole method each time as was done here, but proves an abstract approximation theorem. The following was proved by an argument similar to the one given for the classical C.L.T.

Theorem (Stein): Let (W, W') be an exchangeable pair of random variables s.t. $\mathbb{E}W = 0$, $\mathbb{E}W^2 = 1$, and suppose that there is a $\lambda \in (0,1)$ s.t.

$$\mathbb{E}[W - W' \mid W] = \lambda W.$$

Then for all $t \in \mathbb{R}$,

$$|\mathbb{P}(W \le t) - \Phi(t)| \le \frac{1}{\lambda}\sqrt{\mathrm{Var}(\mathbb{E}[(W'-W)^2 \mid W])} + \frac{1}{(2\pi)^{1/4}\lambda^{1/2}}\sqrt{\mathbb{E}|W'-W|^3}.$$

Multivariate Normal Approximation

(joint with Sourav Chatterjee)

Same idea as before: find a characterizing operator for multi-variate Gaussian.

Let $Z = (Z_1, \dots, Z_k)$ be a random vector with i.i.d. $N(0,1)$ components. Define T_o by

$$T_o f(x) = \Delta f(x) - x \cdot \nabla f(x)$$

Then $\mathbb{E}\, T_o f(Z) \equiv 0.$

Compare with the univariate: $T_o f(x) = f'(x) - x f(x)$.

$$U_o f(x) = \int_0^1 \frac{1}{2t} \left[\mathbb{E} f(\sqrt{t}\, x + \sqrt{1-t}\, Z) - \mathbb{E} f(Z) \right] dt;$$

$T_o U_o f(x) = f(x) - \mathbb{E} f(Z)$. Further:

- $\left\| \frac{\partial (U_o f)}{\partial x_i} \right\|_\infty \le \frac{1}{\sqrt{2}} \|f\|_\infty$
- $\left\| \frac{\partial^2 (U_o f)}{\partial x_i \partial x_j} \right\|_\infty \le \left\| \frac{\partial f}{\partial x_j} \right\|_\infty$
- $\left\| \frac{\partial^3 (U_o f)}{\partial x_i \partial x_j \partial x_l} \right\|_\infty \le \left\| \frac{\partial^2 f}{\partial x_j \partial x_l} \right\|_\infty$

- $\| \mathrm{Hess}_{U_o f}(x) \|_{H.S.} \le \sqrt{k}\, \| |\nabla f| \|_\infty$
- $\| \mathrm{Hess}_{U_o f}(x) \|_{op} \le \sqrt{\frac{2}{\pi}}\, \| |\nabla f| \|_\infty$

Abstract Normal Approximation Thm (Chatterjee, M.)

Let X, X' be random vectors in $\mathbb{R}^k$ with $\mathcal{L}(X) = \mathcal{L}(X')$, and let $Z = (Z_1, \dots, Z_k)$ be a standard Gaussian random vector. Suppose $\exists \lambda \in (0,1)$ such that

(i) $\mathbb{E}[X' - X \mid X] = -\lambda X$

(ii) $\mathbb{E}[(X_i' - X_i)(X_j' - X_j) \mid X] = 2\lambda \delta_{ij} + E_{ij}$

for some random variables E_{ij}.

Then for $g \in C^2(\mathbb{R}^k)$,

$$|\mathbb{E}g(X) - \mathbb{E}g(Z)| \leq \min\left\{ \frac{|g|_1}{2\lambda} \sum_{i,j} \mathbb{E}|E_{ij}|,\ \frac{\sqrt{k}\, \|\nabla g\|_\infty}{2\lambda} \mathbb{E}\Big(\sum_{i,j} E_{ij}^2\Big)^{\frac{1}{2}} \right\}$$

$$+ \frac{k^2 |g|_2}{6\lambda} \sum_i \mathbb{E}|X_i' - X_i|^3.$$

Examples

① A multi-variate C.L.T.

Let X be a random vector with i.i.d. components. Let $\{\theta_i\}_{i=1}^{k}$ be k orthonormal vectors in $\mathbb{R}^n$. Define the random vector W by

$$W_i = \langle \theta_i, X \rangle = \sum_{r=1}^{n} \theta_i^r X_r$$

for $1 \le i \le k$.

W is a rank k projection of X onto $\mathrm{span}\{\theta_1, \ldots, \theta_k\}$.

So long as the θ_i are "reasonable" and k isn't too big, then W should be approximately a standard normal random vector; i.e., have approximately independent Gaussian components.

Two Extensions of this

1. Let $X = (X_1, \ldots, X_n)$ have independent (not necessarily identically distributed) components with $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = 1$ $\forall i$. Then for $g \in C_c^2(\mathbb{R}^k)$,

$$|\mathbb{E}g(W) - \mathbb{E}g(Z)| \le \frac{\sqrt{k}}{2} \|\,|\nabla g|\,\|_\infty \sqrt{\max_{1\le i\le n} \mathbb{E}X_i^4 - 1}\left(\sum_{i=1}^{k} \|\theta_i\|_4^2\right)$$

$$+ \frac{4}{3} k^2 |g|_2 \left(\max_{1\le i\le n} \mathbb{E}|X_i|^3\right)\left(\sum_{i=1}^{k} \|\theta_i\|_3^3\right).$$

Special case:

Suppose $|\theta_i^r| = \frac{1}{\sqrt{n}}$ for each r, i; i.e., the θ_i are orthonormal vectors in directions of corners of the hypercube.

Then $\|\theta_i\|_4^2 = \left(\sum_{r=1}^{n} \frac{1}{n^2}\right)^{1/2} = \frac{1}{\sqrt{n}} \Rightarrow \sum_{i=1}^{k} \|\theta_i\|_4^2 = \frac{k}{\sqrt{n}}$

$\|\theta_i\|_3^3 = \sum_{r=1}^{n} \frac{1}{n^{3/2}} = \frac{1}{\sqrt{n}} \Rightarrow \sum_{i=1}^{k} \|\theta_i\|_3^3 = \frac{k}{\sqrt{n}}$

$\Rightarrow$ If $k = o(n^{1/4})$, then $W \mathrel{\dot\sim} Z$.

2. For $X = (X_1, \ldots, X_n)$ a finite exchangeable sequence and $\{\theta_i\}_{i=1}^k$ orthonormal vectors in $\mathbb{R}^n$ s.t. $\sum_{r=1}^n \theta_i^r = 0$ for each i. Then there are absolute constants a, b, c such that for $g \in C_c^2(\mathbb{R}^k)$,

$$|\mathbb{E}g(W) - \mathbb{E}g(Z)| \le ak|g|_1\left(\sqrt{|\mathbb{E}X_1X_2X_3X_4|} + \sqrt{|\mathbb{E}(X_1^2-1)(X_2^2-1)|}\right)$$
$$+ b|g|_1\sqrt{\mathbb{E}X_1^4}\left(\sum_{i=1}^k \|\theta_i\|_4\right)^2 + ck^2|g|_2\mathbb{E}|X_1|^3\left(\sum_{i=1}^k \|\theta_i\|_3^3\right).$$

Note: $|\mathbb{E}X_1X_2X_3X_4|$ small $\Leftrightarrow$ "weak dependence"

Same for $|\mathbb{E}(X_1^2-1)(X_2^2-1)|$.

<u>Idea of Proof of 2.</u>

Use the exchangeability of X:

Pick a random transposition $\tau = (I\,J)$ from S^n.

Let $X' = (X_{\tau(1)}, \ldots, X_{\tau(n)})$. Then

$$W_i' = W_i - \theta_i^I X_I + \theta_i^J X_I - \theta_i^J X_J + \theta_i^I X_J$$
$$= W_i + (\theta_i^J - \theta_i^I)(X_I - X_J).$$

$$\mathbb{E}[W_i' - W_i \mid X] = \frac{1}{n(n-1)} \sum_{r,s}{}' (\theta_i^r X_s - \theta_i^r X_r + \theta_i^s X_r - \theta_i^s X_s)$$

$$= \frac{1}{n(n-1)} \sum_{r,s} (\theta_i^r X_s - \theta_i^r X_r + \theta_i^s X_r - \theta_i^s X_s)$$

$$= \frac{1}{n(n-1)} \left[-2n \sum_r \theta_i^r X_r \right]$$

$$= -\frac{2}{n-1} W_i \qquad \Rightarrow \qquad \boxed{\lambda = \frac{2}{n-1}}.$$

Finding and estimating the E_{ij} is rather more involved, but along the same lines.

Hoeffding's Combinatorial CLT

Let $W = \sum_{i=1}^{n} a_{i\pi(i)}$ for π a random permutation and $A = (a_{ij})$ a fixed matrix with some normalization conditions. Then

$W \dot{\approx} Z$ for large n. (rate of convergence by Bolthausen)

Note: This can also be phrased in terms of matrices and projections: let $M(\pi)$ be the permutation matrix $m_{ij} = \begin{cases} 1 & \underline{\pi(j) = i} \\ 0 & \end{cases}$

Then $W = tr(AM)$. W is a rank 1 projection of $S^n \subseteq GL(n)$ onto span (A^t).

Multivariate version: Take $A_1, \dots, A_k$ s.t. $tr(A_i A_j^t) = (n-1)\delta_{ij}$ and with all row & column sums equal to zero. Define $W = (tr(A_1 M), \dots, tr(A_k M))$. Then for $g \in C_c^2(\mathbb{R}^k)$,

$$|\mathbb{E}g(W) - \mathbb{E}g(Z)| \le \frac{c_1 |g|_1}{\sqrt{n}} \left(k + \sum_{i,j=1}^{k} \left(\sum_{l,m=1}^{n} (A_i)_{lm}^2 (A_j)_{lm}^2 \right)^{1/2} \right)$$
$$+ |g|_2 \left(\frac{c_2 k^2}{n} \sum_{i=1}^{k} \sum_{l,m=1}^{n} |(A_i)_{lm}|^3 \right).$$

Continuous Version of the abstract normal approximation theorem:

Let X be a random vector in $\mathbb{R}^k$ and for each $\varepsilon > 0$, let X_ε be a random vector s.t. $\mathcal{L}(X) = \mathcal{L}(X_\varepsilon)$ and $\lim_{\varepsilon \to 0} X_\varepsilon = X$ a.s. Suppose $\exists\ \lambda \in (0,1)$, deterministic functions h and k with

$$\lim_{\varepsilon \to 0} \varepsilon^{-2} h(\varepsilon) = 0 = \lim_{\varepsilon \to 0} \varepsilon^{-2} k(\varepsilon),$$

measurable α, β with $\mathbb{E}|\alpha(X)| < \infty$, $\mathbb{E}|\beta(X)| < \infty$, and random variables E_{ij} such that

1. $\mathbb{E}\left[(X_\varepsilon - X)_i \,|\, X\right] = -\lambda \varepsilon^2 X_i + h(\varepsilon)\,\alpha(X)$
2. $\mathbb{E}\left[(X_\varepsilon - X)_i (X_\varepsilon - X)_j \,|\, X\right] = 2\lambda \varepsilon^2 \delta_{ij} + \varepsilon^2 E_{ij} + k(\varepsilon)\beta(X)$
3. $\mathbb{E}|X_\varepsilon - X|^3 = o(\varepsilon^2)$.

Then
$$d_W(X, Z) \le \min\left\{ \frac{1}{2\lambda} \sum_{i,j=1}^{n} \mathbb{E}|E_{ij}|,\ \frac{\sqrt{k}}{2\lambda} \mathbb{E}\Big(\sum_{i,j=1}^{n} E_{ij}^2\Big)^{1/2} \right\}.$$

If, further, $E_{ij} = E_i F_j + \delta_{ij} R_i$, then

$$d_W(X, Z) \le \frac{1}{\lambda\sqrt{2\pi}} \mathbb{E}\big(|E| \cdot |F|\big) + \frac{1}{\lambda} \mathbb{E}\Big(\sum_{i=1}^{n} |R_i|\Big).$$

Examples:

Let Y be a spherically symmetric random variable such that $\mathbb{E}Y_1^2 = 1$.
Assume $\exists$ a s.t. $\mathrm{Var}(|Y|^2) \leq a$.
Fix k, and let $P_k : \mathbb{R}^n \to \mathbb{R}^k$

$$(x_1, \dots, x_n) \longmapsto (x_1, \dots, x_k)$$

Then there is a c s.t.

$$d_W(P_k(Y), Z_k) \leq \frac{ck}{n}.$$

Exchangeable pair: small random rotation.
Fix ε, and let

$$R_\varepsilon = \begin{bmatrix} \sqrt{1-\varepsilon^2} & \varepsilon \\ -\varepsilon & \sqrt{1-\varepsilon^2} \end{bmatrix} \oplus I_{n-2}.$$

Let $U \in \mathcal{O}(n)$ be random, independent of Y.

Let $Y_\varepsilon = (U R_\varepsilon U^T) Y$. Then

(Y, Y_ε) is exchangeable by the spherical symmetry of Y.