Higher Dimensional Quasi-Power Theorem and Berry–Esseen Inequality

Clemens Heuberger

Alpen-Adria-Universität Klagenfurt
Joint work with Sara Kropf (Academia Sinica, Taipei)

AofA'16, Kraków, 2016-07-07

Prologue: Classical Central Limit Theorem

Prologue: Classical Central Limit Theorem

Ω_j: sequence of independent and identically distributed ($\sim \Omega$) random variables; $\mathbb{E}(\Omega) = \mu$, $0 < \mathbb{V}(\Omega) = \sigma^2 < \infty$,
$S_n = \Omega_1 + \cdots + \Omega_n$.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Prologue: Classical Central Limit Theorem

Ω_j: sequence of independent and identically distributed ($\sim \Omega$) random variables; $\mathbb{E}(\Omega) = \mu$, $0 < \mathbb{V}(\Omega) = \sigma^2 < \infty$,
$S_n = \Omega_1 + \cdots + \Omega_n$.

Theorem (Classical Central Limit Theorem)

$$\lim_{n\to\infty} \mathbb{P}\Big(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\Big) = \frac{1}{2\pi}\int_{-\infty}^{x} \exp\Big(-\frac{1}{2}t^2\Big)\, dt.$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Prologue: Classical Central Limit Theorem

Ω_j: sequence of independent and identically distributed ($\sim \Omega$) random variables; $\mathbb{E}(\Omega) = \mu$, $0 < \mathbb{V}(\Omega) = \sigma^2 < \infty$,
$S_n = \Omega_1 + \cdots + \Omega_n$.

Theorem (Classical Central Limit Theorem)

$$\lim_{n\to\infty} \mathbb{P}\Big(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\Big) = \frac{1}{2\pi}\int_{-\infty}^{x} \exp\Big(-\frac{1}{2}t^2\Big)\, dt.$$

Proof:

- Characteristic function

$$\varphi_{S_n}(t) = \mathbb{E}(e^{itS_n}) = \big(\mathbb{E}(e^{it\Omega})\big)^n = \varphi_\Omega^n(t).$$

Prologue: Classical Central Limit Theorem

Ω_j: sequence of independent and identically distributed ($\sim \Omega$) random variables; $\mathbb{E}(\Omega) = \mu$, $0 < \mathbb{V}(\Omega) = \sigma^2 < \infty$,
$S_n = \Omega_1 + \cdots + \Omega_n$.

Theorem (Classical Central Limit Theorem)

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \frac{1}{2\pi}\int_{-\infty}^{x} \exp\left(-\frac{1}{2}t^2\right) dt.$$

Proof:

- Characteristic function

$$\varphi_{S_n}(t) = \mathbb{E}(e^{itS_n}) = \left(\mathbb{E}(e^{it\Omega})\right)^n = \varphi_\Omega^n(t).$$

- Rescaled characteristic function

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \varphi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right)\exp\left(-\frac{it\mu\sqrt{n}}{\sigma}\right)$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Prologue: Classical Central Limit Theorem (cont.)

- Rescaled characteristic function

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \varphi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right)\exp\left(-\frac{it\mu\sqrt{n}}{\sigma}\right)$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Prologue: Classical Central Limit Theorem (cont.)

- Rescaled characteristic function

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \varphi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right)\exp\left(-\frac{it\mu\sqrt{n}}{\sigma}\right)$$
$$= \exp\left(n\log\varphi_\Omega\left(\frac{t}{\sigma\sqrt{n}}\right) - \frac{it\mu\sqrt{n}}{\sigma}\right).$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Prologue: Classical Central Limit Theorem (cont.)

- Rescaled characteristic function

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \varphi_{S_n}\Big(\frac{t}{\sigma\sqrt{n}}\Big)\exp\Big(-\frac{it\mu\sqrt{n}}{\sigma}\Big)$$
$$= \exp\Big(n\log\varphi_\Omega\Big(\frac{t}{\sigma\sqrt{n}}\Big) - \frac{it\mu\sqrt{n}}{\sigma}\Big).$$

- Use $\varphi_\Omega(0) = 1$, $\varphi'_\Omega(0) = i\mu$, $\varphi''_\Omega(0) = -\mathbb{E}(\Omega^2) = -(\sigma^2+\mu^2)$:

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Prologue: Classical Central Limit Theorem (cont.)

- Rescaled characteristic function

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \varphi_{S_n}\Big(\frac{t}{\sigma\sqrt{n}}\Big)\exp\Big(-\frac{it\mu\sqrt{n}}{\sigma}\Big)$$
$$= \exp\Big(n\log\varphi_\Omega\Big(\frac{t}{\sigma\sqrt{n}}\Big) - \frac{it\mu\sqrt{n}}{\sigma}\Big).$$

- Use $\varphi_\Omega(0) = 1$, $\varphi'_\Omega(0) = i\mu$, $\varphi''_\Omega(0) = -\mathbb{E}(\Omega^2) = -(\sigma^2+\mu^2)$:

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \exp\Big(-\frac{t^2}{2} + O\Big(\frac{1}{\sqrt{n}}\Big)\Big).$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Prologue: Classical Central Limit Theorem (cont.)

- Rescaled characteristic function

$$\begin{aligned}\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) &= \varphi_{S_n}\Big(\frac{t}{\sigma\sqrt{n}}\Big)\exp\Big(-\frac{it\mu\sqrt{n}}{\sigma}\Big)\\ &= \exp\Big(n\log\varphi_\Omega\Big(\frac{t}{\sigma\sqrt{n}}\Big)-\frac{it\mu\sqrt{n}}{\sigma}\Big).\end{aligned}$$

- Use $\varphi_\Omega(0)=1$, $\varphi'_\Omega(0)=i\mu$, $\varphi''_\Omega(0)=-\mathbb{E}(\Omega^2)=-(\sigma^2+\mu^2)$:

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t)=\exp\Big(-\frac{t^2}{2}+O\Big(\frac{1}{\sqrt{n}}\Big)\Big).$$

- For $n\to\infty$: characteristic function of standard normal distribution.

Prologue: Classical Central Limit Theorem (cont.)

- Rescaled characteristic function

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \varphi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right)\exp\left(-\frac{it\mu\sqrt{n}}{\sigma}\right) = \exp\left(n\log\varphi_\Omega\left(\frac{t}{\sigma\sqrt{n}}\right) - \frac{it\mu\sqrt{n}}{\sigma}\right).$$

- Use $\varphi_\Omega(0) = 1$, $\varphi'_\Omega(0) = i\mu$, $\varphi''_\Omega(0) = -\mathbb{E}(\Omega^2) = -(\sigma^2+\mu^2)$:

$$\varphi_{(S_n-\mu n)/(\sigma\sqrt{n})}(t) = \exp\left(-\frac{t^2}{2} + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

- For $n \to \infty$: characteristic function of standard normal distribution.
- Lévy continuity theorem: convergence in distribution. $\square$

Quasi-Power-Theorem

$\{\mathbf{\Omega}_n\}_{n\geq 1}$: sequence of m-dimensional real random vectors;
$\tau > 0$

Quasi-Power-Theorem

$\{\boldsymbol{\Omega}_n\}_{n\geq 1}$: sequence of m-dimensional real random vectors;
$\tau > 0$

Theorem (Quasi-Power-Theorem; H.–Kropf 2016)

Assume moment generating function of shape

$$M_n(\mathbf{s}) := \mathbb{E}(e^{\langle \boldsymbol{\Omega}_n, \mathbf{s}\rangle}) = e^{u(\mathbf{s})\phi_n + v(\mathbf{s})}(1 + O(\kappa_n^{-1})),$$

uniformly for $\mathbf{s} \in \mathbb{C}^m$, $\|\mathbf{s}\| \leq \tau$ *where*

Quasi-Power-Theorem

$\{\mathbf{\Omega}_n\}_{n\geq 1}$: sequence of m-dimensional real random vectors;
$\tau > 0$

Theorem (Quasi-Power-Theorem; H.–Kropf 2016)

Assume moment generating function of shape

$$M_n(\mathbf{s}) := \mathbb{E}(e^{\langle \mathbf{\Omega}_n, \mathbf{s}\rangle}) = e^{u(\mathbf{s})\phi_n + v(\mathbf{s})}(1 + O(\kappa_n^{-1})),$$

uniformly for $\mathbf{s} \in \mathbb{C}^m$, $\|\mathbf{s}\| \leq \tau$ *where*

1. $u(\mathbf{s})$ *and* $v(\mathbf{s})$ *analytic for* $\|\mathbf{s}\| \leq \tau$ *and independent of* n;

Quasi-Power-Theorem

$\{\boldsymbol{\Omega}_n\}_{n\geq 1}$: sequence of m-dimensional real random vectors;
$\tau > 0$

Theorem (Quasi-Power-Theorem; H.–Kropf 2016)

Assume moment generating function of shape

$$M_n(\mathbf{s}) := \mathbb{E}(e^{\langle \boldsymbol{\Omega}_n, \mathbf{s}\rangle}) = e^{u(\mathbf{s})\phi_n + v(\mathbf{s})}(1 + O(\kappa_n^{-1})),$$

uniformly for $\mathbf{s} \in \mathbb{C}^m$, $\|\mathbf{s}\| \leq \tau$ *where*

1. $u(\mathbf{s})$ *and* $v(\mathbf{s})$ *analytic for* $\|\mathbf{s}\| \leq \tau$ *and independent of* n;
2. *non-singular Hessian* $H_u(\mathbf{0})$ *of* u;

Quasi-Power-Theorem

$\{\mathbf{\Omega}_n\}_{n\geq 1}$: sequence of m-dimensional real random vectors;
$\tau > 0$

Theorem (Quasi-Power-Theorem; H.–Kropf 2016)

Assume moment generating function of shape

$$M_n(\mathbf{s}) := \mathbb{E}(e^{\langle \mathbf{\Omega}_n, \mathbf{s}\rangle}) = e^{u(\mathbf{s})\phi_n + v(\mathbf{s})}(1 + O(\kappa_n^{-1})),$$

uniformly for $\mathbf{s} \in \mathbb{C}^m$, $\|\mathbf{s}\| \leq \tau$ *where*

1. $u(\mathbf{s})$ *and* $v(\mathbf{s})$ *analytic for* $\|\mathbf{s}\| \leq \tau$ *and independent of* n*;*
2. *non-singular Hessian* $H_u(\mathbf{0})$ *of* u*;*
3. $\lim_{n\to\infty} \phi_n = \infty$*;*

Quasi-Power-Theorem

$\{\mathbf{\Omega}_n\}_{n\geq 1}$: sequence of m-dimensional real random vectors;
$\tau > 0$

Theorem (Quasi-Power-Theorem; H.–Kropf 2016)

Assume moment generating function of shape

$$M_n(\mathbf{s}) := \mathbb{E}(e^{\langle \mathbf{\Omega}_n, \mathbf{s}\rangle}) = e^{u(\mathbf{s})\phi_n + v(\mathbf{s})}(1 + O(\kappa_n^{-1})),$$

uniformly for $\mathbf{s} \in \mathbb{C}^m$, $\|\mathbf{s}\| \leq \tau$ *where*

1. $u(\mathbf{s})$ *and* $v(\mathbf{s})$ *analytic for* $\|\mathbf{s}\| \leq \tau$ *and independent of* n;
2. *non-singular Hessian* $H_u(\mathbf{0})$ *of* u;
3. $\lim_{n\to\infty} \phi_n = \infty$;
4. $\lim_{n\to\infty} \kappa_n = \infty$.

Quasi-Power-Theorem

$\{\mathbf{\Omega}_n\}_{n\geq 1}$: sequence of m-dimensional real random vectors;
$\tau > 0$

Theorem (Quasi-Power-Theorem; H.–Kropf 2016)

Assume moment generating function of shape

$$M_n(\mathbf{s}) := \mathbb{E}(e^{\langle \mathbf{\Omega}_n, \mathbf{s}\rangle}) = e^{u(\mathbf{s})\phi_n + v(\mathbf{s})}(1 + O(\kappa_n^{-1})),$$

uniformly for $\mathbf{s} \in \mathbb{C}^m$, $\|\mathbf{s}\| \leq \tau$ *where*

1. $u(\mathbf{s})$ *and* $v(\mathbf{s})$ *analytic for* $\|\mathbf{s}\| \leq \tau$ *and independent of* n*;*
2. *non-singular Hessian* $H_u(\mathbf{0})$ *of* u*;*
3. $\lim_{n\to\infty} \phi_n = \infty$;
4. $\lim_{n\to\infty} \kappa_n = \infty$.

Then, the distribution of $\mathbf{\Omega}_n$ *is asymptotically normal with speed of convergence* $O(\phi_n^{-1/2})$, ...

Quasi-Power-Theorem, continued

Theorem (Quasi-Power-Theorem, continued)

Then, the distribution of $\mathbf{\Omega}_n$ *is asymptotically normal with speed of convergence* $O(\phi_n^{-1/2})$, *i.e.,*

$$\sup_{\mathbf{x}\in\mathbb{R}^m}\left|\mathbb{P}\left(\frac{\mathbf{\Omega}_n - \operatorname{grad} u(\mathbf{0})\phi_n}{\sqrt{\phi_n}} \le \mathbf{x}\right) - \Phi_{H_u(\mathbf{0})}(\mathbf{x})\right| = O\left(\frac{1}{\sqrt{\phi_n}}\right),$$

where

$$\Phi_\Sigma(\mathbf{x}) = \frac{1}{(2\pi)^{m/2}\sqrt{\det\Sigma}}\int_{\mathbf{y}\le\mathbf{x}} \exp\left(-\frac{1}{2}\mathbf{y}^\top\Sigma^{-1}\mathbf{y}\right)\,d\mathbf{y}$$

(distribution function of the m-dimensional normal distribution with mean $\mathbf{0}$ *and variance-covariance matrix* Σ*).*

Quasi-Power-Theorem, continued

Theorem (Quasi-Power-Theorem, continued)

Then, the distribution of $\mathbf{\Omega}_n$ is asymptotically normal with speed of convergence $O(\phi_n^{-1/2})$, i.e.,

$$\sup_{\mathbf{x}\in\mathbb{R}^m} \left| \mathbb{P}\left(\frac{\mathbf{\Omega}_n - \operatorname{grad} u(\mathbf{0})\phi_n}{\sqrt{\phi_n}} \le \mathbf{x} \right) - \Phi_{H_u(\mathbf{0})}(\mathbf{x}) \right| = O\left(\frac{1}{\sqrt{\phi_n}} \right),$$

where

$$\Phi_\Sigma(\mathbf{x}) = \frac{1}{(2\pi)^{m/2}\sqrt{\det \Sigma}} \int_{\mathbf{y}\le\mathbf{x}} \exp\left(-\frac{1}{2}\mathbf{y}^\top \Sigma^{-1} \mathbf{y} \right) d\mathbf{y}$$

(distribution function of the m-dimensional normal distribution with mean $\mathbf{0}$ *and variance-covariance matrix* Σ*).*

Dimension 1: Hwang, 1998; Dimension 2: H. 2007.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Typical Application: Meromorphic Perturbation

- $\mathbf{\Omega}_n$: random variables in $\mathbb{N}^m$

$$\mathbb{P}(\mathbf{\Omega}_n = \mathbf{k}) = \frac{a_{n\mathbf{k}}}{a_n} \text{ with } a_n = \sum_{\mathbf{k}} a_{n\mathbf{k}}.$$

- Bivariate generating function

$$F(z, \mathbf{u}) = \sum_{n,\mathbf{k}} a_{n\mathbf{k}} z^n \mathbf{u}^{\mathbf{k}}.$$

- Probability generating function p_n of $\mathbf{\Omega}_n$:

$$p_n(\mathbf{u}) = \frac{[z^n]F(z, \mathbf{u})}{[z^n]F(z, \mathbf{1})}.$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Typical Application: Meromorphic Perturbation (cont.)

- Assumption: $F(z, \mathbf{u})$ has unique dominant simple pole at $z = \rho(\mathbf{u})$ for some analytic ρ for $\mathbf{u} \approx \mathbf{1}$.
- Cauchy's formula ("singularity analysis"):

$$[z^n]F(z,\mathbf{u}) = \frac{1}{2\pi i}\oint_{|z| \text{ small}} \frac{F(z,\mathbf{u})}{z^{n+1}}\,dz$$

$$= -\frac{1}{\rho(\mathbf{u})^{n+1}} \operatorname{Res}_{z=\rho(\mathbf{u})} F(z,\mathbf{u}) + \frac{1}{2\pi i}\oint_{|z| \text{ larger}} \frac{F(z,\mathbf{u})}{z^{n+1}}\,dz$$

$$= \frac{C(\mathbf{u})}{\rho(\mathbf{u})^n}\Bigl(1 + O\Bigl(\frac{1}{\kappa^n}\Bigr)\Bigr) \qquad \kappa > 1;\ C(\mathbf{u}) \text{ analytic.}$$

- Moment generating function ($e^{\mathbf{s}} = (e^{s_1}, \dots, e^{s_m})$):

$$M_n(\mathbf{s}) = p_n(e^{\mathbf{s}}) = \frac{[z^n]F(z,e^{\mathbf{s}})}{[z^n]F(z,\mathbf{1})} = \frac{C(e^{\mathbf{s}})}{C(\mathbf{1})}\Bigl(\frac{\rho(\mathbf{1})}{\rho(e^{\mathbf{s}})}\Bigr)^n\Bigl(1 + O\Bigl(\frac{1}{\kappa^n}\Bigr)\Bigr)$$

$$= \exp\Bigl(n\log\Bigl(\frac{\rho(\mathbf{1})}{\rho(e^{\mathbf{s}})}\Bigr) + \log\Bigl(\frac{C(e^{\mathbf{s}})}{C(\mathbf{1})}\Bigr)\Bigr)\Bigl(1 + O\Bigl(\frac{1}{\kappa^n}\Bigr)\Bigr).$$

- Quasi-power theorem applies (check variability condition!).

Example (cf. Drmota 1997)

Context-free grammar G:

- non-terminal symbols S and T,
- terminal symbols $\{a, b, c\}$,
- starting symbol S,
- rules $P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}$.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Example (cf. Drmota 1997)

Context-free grammar G:

- non-terminal symbols S and T,
- terminal symbols $\{a, b, c\}$,
- starting symbol S,
- rules $P = \{S \to aSbS, S \to bT, T \to bS, T \to cT, T \to a\}$.

e.g.

$$S \to aSbS$$

Example (cf. Drmota 1997)

Context-free grammar G:

- non-terminal symbols S and T,
- terminal symbols $\{a, b, c\}$,
- starting symbol S,
- rules $P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}$.

e.g.

$$S \to aSbS \to abTbaSbS$$

Example (cf. Drmota 1997)

Context-free grammar G:

- non-terminal symbols S and T,
- terminal symbols $\{a, b, c\}$,
- starting symbol S,
- rules $P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}$.

e.g.

$$S \to aSbS \to abTbaSbS \to abcTbabTbbT$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Example (cf. Drmota 1997)

Context-free grammar G:

- non-terminal symbols S and T,
- terminal symbols $\{a, b, c\}$,
- starting symbol S,
- rules $P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}$.

e.g.

$$S \to aSbS \to abTbaSbS \to abcTbabTbbT \to abcababab ba.$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Example (cf. Drmota 1997)

Context-free grammar G:

- non-terminal symbols S and T,
- terminal symbols $\{a, b, c\}$,
- starting symbol S,
- rules $P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}$.

e.g.

$$S \to aSbS \to abTbaSbS \to abcTbabTbbT \to abcabababba.$$

- $\mathbf{\Omega}_n = (x_1, x_2) \iff$ word of length n in language has $x_1, x_2, n - x_1 - x_2$ terminal symbols a, b, c.

Example (cf. Drmota 1997)

Context-free grammar G:

- non-terminal symbols S and T,
- terminal symbols $\{a, b, c\}$,
- starting symbol S,
- rules $P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}$.

e.g.

$$S \to aSbS \to abTbaSbS \to abcTbabTbbT \to abcabababba.$$

- $\mathbf{\Omega}_n = (x_1, x_2) \iff$ word of length n in language has $x_1, x_2, n - x_1 - x_2$ terminal symbols a, b, c.
- Moment generating function:

$$\mathbb{E}(e^{\langle \mathbf{\Omega}_n, \mathbf{s} \rangle}) = \frac{y_n(e^{\mathbf{s}})}{y_n(\mathbf{1})}$$

with y_n of the shape

$$y_n(\mathbf{u}) = C(\mathbf{u}) f(\mathbf{u})^{-n} n^{-3/2} (1 + O(1/n)).$$

Interlude: Classical Berry–Esseen Inequality (Idea)

Goal: Translate bounds on the difference of characteristic functions to bounds on the difference of distribution functions.

Interlude: Classical Berry–Esseen Inequality (Idea)

Goal: Translate bounds on the difference of characteristic functions to bounds on the difference of distribution functions.

Rough idea:

- Lévy's theorem (X: continuous one-dimensional random variable):

$$\mathbb{P}(a \leq X \leq b) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt.$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Interlude: Classical Berry–Esseen Inequality (Idea)

Goal: Translate bounds on the difference of characteristic functions to bounds on the difference of distribution functions.

Rough idea:

- Lévy's theorem (X: continuous one-dimensional random variable):

$$\mathbb{P}(a \le X \le b) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt.$$

- Taking difference (X, Y: continuous one-dimensional random variables, expectation exists):

$$\mathbb{P}(a \le X \le b) - \mathbb{P}(a \le Y \le b) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \left(\varphi_X(t) - \varphi_Y(t)\right) dt.$$

Interlude: Classical Berry–Esseen Inequality (Idea)

Goal: Translate bounds on the difference of characteristic functions to bounds on the difference of distribution functions.

Rough idea:

- Lévy's theorem (X: continuous one-dimensional random variable):

$$\mathbb{P}(a \leq X \leq b) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt.$$

- Taking difference (X, Y: continuous one-dimensional random variables, expectation exists):

$$\mathbb{P}(a \leq X \leq b) - \mathbb{P}(a \leq Y \leq b) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \left(\varphi_X(t) - \varphi_Y(t)\right) dt.$$

- Note: $\varphi_X(0) = 1 = \varphi_Y(0) \Rightarrow$ integrand bounded.

Interlude: Classical Berry–Esseen Inequality

Theorem

Let X, Y be one-dimensional random variables, F_Y' bounded. Then for $T > 0$:

$$\|F_X - F_Y\|_\infty \leq \frac{1}{\pi} \int_{-T}^{T} \left| \frac{\varphi_X(t) - \varphi_Y(t)}{t} \right| dt + \frac{24}{\pi} \frac{\|F_Y'\|_\infty}{T}.$$

Towards Higher Dimensional Berry–Esseen Inequality

- Higher dimensional Lévy's theorem ($\mathbf{X}$: continuous m-dimensional random variable):

$$\mathbb{P}_{\mathbf{X}}(\mathbf{a} \leq \mathbf{X} \leq \mathbf{b}) =$$

$$\frac{1}{(2\pi)^m} \lim_{T_1 \to \infty} \cdots \lim_{T_m \to \infty} \int_{|\mathbf{t}| \leq \mathbf{T}} \prod_{k=1}^{m} \left(\frac{e^{-ia_k t_k} - e^{-ib_k t_k}}{it_k} \right) \varphi_{\mathbf{X}}(\mathbf{t}) \, d\mathbf{t}.$$

Towards Higher Dimensional Berry–Esseen Inequality

- Higher dimensional Lévy's theorem (**X**: continuous m-dimensional random variable):

$$\mathbb{P}_{\mathbf{X}}(\mathbf{a} \leq \mathbf{X} \leq \mathbf{b}) = \frac{1}{(2\pi)^m} \lim_{T_1 \to \infty} \cdots \lim_{T_m \to \infty} \int_{|\mathbf{t}| \leq \mathbf{T}} \prod_{k=1}^{m} \left(\frac{e^{-ia_k t_k} - e^{-ib_k t_k}}{it_k} \right) \varphi_{\mathbf{X}}(\mathbf{t})\, d\mathbf{t}.$$

- Taking difference

$$\varphi_{\mathbf{X}}(\mathbf{t}) - \varphi_{\mathbf{Y}}(\mathbf{t})$$

does not remove the singularity at $\mathbf{0}$ (denominator $t_1 \ldots t_m$).

Towards Higher Dimensional Berry–Esseen Inequality

- Higher dimensional Lévy's theorem ($\mathbf{X}$: continuous m-dimensional random variable):

$$\mathbb{P}_{\mathbf{X}}(\mathbf{a} \leq \mathbf{X} \leq \mathbf{b}) =$$

$$\frac{1}{(2\pi)^m} \lim_{T_1 \to \infty} \cdots \lim_{T_m \to \infty} \int_{|\mathbf{t}| \leq \mathbf{T}} \prod_{k=1}^{m} \Big(\frac{e^{-ia_k t_k} - e^{-ib_k t_k}}{it_k} \Big) \varphi_{\mathbf{X}}(\mathbf{t}) \, d\mathbf{t}.$$

- Taking difference

$$\varphi_{\mathbf{X}}(\mathbf{t}) - \varphi_{\mathbf{Y}}(\mathbf{t})$$

does not remove the singularity at $\mathbf{0}$ (denominator $t_1 \ldots t_m$).

- Need operator Λ such that $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) = 0$ if $t_j = 0$ for one j.

Constructing Λ

Wanted:

- $\Lambda(\varphi_{\mathbf{x}})(\mathbf{t}) = 0$ if $t_j = 0$ for one j.

Constructing Λ

Wanted:

- $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) = 0$ if $t_j = 0$ for one j.
- $\Lambda(\varphi_{\mathbf{X}})$ must make probabilistic sense.
  Building blocks of shape $\varphi_{\mathbf{X}}(t_1, \ldots, t_j, 0, \ldots, 0)$: characteristic function of marginal distribution (ignore some components).
  Apply Lévy's theorem on these blocks.

Constructing Λ

Wanted:

- $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) = 0$ if $t_j = 0$ for one j.
- $\Lambda(\varphi_{\mathbf{X}})$ must make probabilistic sense.
  Building blocks of shape $\varphi_{\mathbf{X}}(t_1, \ldots, t_j, 0, \ldots, 0)$: characteristic function of marginal distribution (ignore some components).
  Apply Lévy's theorem on these blocks.
- $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t})$ decays exponentially in every variable.
  Integrate over all t_j, do not want asymptotic contributions from tails.

Constructing Λ

Wanted:

- $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) = 0$ if $t_j = 0$ for one j.
- $\Lambda(\varphi_{\mathbf{X}})$ must make probabilistic sense.
  Building blocks of shape $\varphi_{\mathbf{X}}(t_1, \dots, t_j, 0, \dots, 0)$: characteristic function of marginal distribution (ignore some components).
  Apply Lévy's theorem on these blocks.
- $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t})$ decays exponentially in every variable.
  Integrate over all t_j, do not want asymptotic contributions from tails.

Example: $m = 3$, $h(0,0,0) = 1$:

$$\begin{aligned}\Lambda(h)(t_1, t_2, t_3) = h(t_1, t_2, t_3) &- h(t_1, t_2, 0)h(0, 0, t_3) \\ &- h(t_1, 0, t_3)h(0, t_2, 0) - h(0, t_2, t_3)h(t_1, 0, 0) \\ &+ 2h(t_1, 0, 0)h(0, t_2, 0)h(0, 0, t_3)\end{aligned}$$

Constructing Λ

Wanted:

- $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) = 0$ if $t_j = 0$ for one j.
- $\Lambda(\varphi_{\mathbf{X}})$ must make probabilistic sense.
  Building blocks of shape $\varphi_{\mathbf{X}}(t_1, \ldots, t_j, 0, \ldots, 0)$: characteristic function of marginal distribution (ignore some components).
  Apply Lévy's theorem on these blocks.
- $\Lambda(\varphi_{\mathbf{X}})(\mathbf{t})$ decays exponentially in every variable.
  Integrate over all t_j, do not want asymptotic contributions from tails.

Example: $m = 3,\ h(0,0,0) = 1$:

$$\begin{aligned}\Lambda(h)(t_1, t_2, 0) = h(t_1, t_2, 0) &- h(t_1, t_2, 0)h(0, 0, 0) \\ &- h(t_1, 0, 0)h(0, t_2, 0) - h(0, t_2, 0)h(t_1, 0, 0) \\ &+ 2h(t_1, 0, 0)h(0, t_2, 0)h(0, 0, 0)\end{aligned}$$

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \ldots, m\}$,

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \ldots, m\}$,
- $h\colon \mathbb{C}^m \to \mathbb{C}$,

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \ldots, m\}$,
- $h\colon \mathbb{C}^m \to \mathbb{C}$,
- Π_m: set of partitions of $[m]$,

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \ldots, m\}$,
- $h \colon \mathbb{C}^m \to \mathbb{C}$,
- Π_m: set of partitions of $[m]$,
- $\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \ldots, m\}$,
- $h \colon \mathbb{C}^m \to \mathbb{C}$,
- Π_m: set of partitions of $[m]$,
- $\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,
- $\psi_K(h)(t_1, \ldots, t_m) = h(t_1[1 \in K], \ldots, t_m[m \in K])$,

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \dots, m\}$,
- $h\colon \mathbb{C}^m \to \mathbb{C}$,
- Π_m: set of partitions of $[m]$,
- $\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,
- $\psi_K(h)(t_1, \dots, t_m) = h(t_1[1 \in K], \dots, t_m[m \in K])$,

then set

$$\Lambda(h) := \sum_{\alpha \in \Pi_m} \mu_\alpha \prod_{K \in \alpha} h \circ \psi_K.$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \dots, m\}$,
- $h \colon \mathbb{C}^m \to \mathbb{C}$,
- Π_m: set of partitions of $[m]$,
- $\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,
- $\psi_K(h)(t_1, \dots, t_m) = h(t_1[1 \in K], \dots, t_m[m \in K])$,

then set

$$\Lambda(h) := \sum_{\alpha \in \Pi_m} \mu_\alpha \prod_{K \in \alpha} h \circ \psi_K.$$

Examples:

- $m = 1$: $\Lambda(h)(t) = h(t)$,

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Esseen Inequality: Operator Λ—Definition

Let

- $[m] := \{1, \dots, m\}$,
- $h \colon \mathbb{C}^m \to \mathbb{C}$,
- Π_m: set of partitions of $[m]$,
- $\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,
- $\psi_K(h)(t_1, \dots, t_m) = h(t_1[1 \in K], \dots, t_m[m \in K])$,

then set

$$\Lambda(h) := \sum_{\alpha \in \Pi_m} \mu_\alpha \prod_{K \in \alpha} h \circ \psi_K.$$

Examples:

- $m = 1$: $\Lambda(h)(t) = h(t)$,
- $m = 2$: $\Lambda(h)(t_1, t_2) = h(t_1, t_2) - h(t_1, 0)h(0, t_2)$.

Berry–Esseen Inequality: Operator Λ—Property

$\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,

$$\Lambda(h) := \sum_{\alpha \in \Pi_m} \mu_\alpha \prod_{K \in \alpha} h \circ \psi_K .$$

Lemma

For $J \subsetneq [m]$ and $h(\mathbf{0}) = 0$:

$$\Lambda(h) \circ \psi_J = 0.$$

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Esseen Inequality: Operator Λ—Property

$\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,

$$\Lambda(h) := \sum_{\alpha \in \Pi_m} \mu_\alpha \prod_{K \in \alpha} h \circ \psi_K.$$

Lemma

For $J \subsetneq [m]$ and $h(\mathbf{0}) = 0$:

$$\Lambda(h) \circ \psi_J = 0.$$

Sketch of proof:

- $\mu_\alpha = \mu(\alpha, [m])$ (Möbius function of the poset Π_m).

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Esseen Inequality: Operator Λ—Property

$\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,

$$\Lambda(h) := \sum_{\alpha \in \Pi_m} \mu_\alpha \prod_{K \in \alpha} h \circ \psi_K .$$

Lemma

For $J \subsetneq [m]$ and $h(\mathbf{0}) = 0$:

$$\Lambda(h) \circ \psi_J = 0.$$

Sketch of proof:

- $\mu_\alpha = \mu(\alpha, [m])$ (Möbius function of the poset Π_m).
- Collect equal terms.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Esseen Inequality: Operator Λ—Property

$\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha|-1)!$ for $\alpha \in \Pi_m$,

$$\Lambda(h) := \sum_{\alpha\in\Pi_m} \mu_\alpha \prod_{K\in\alpha} h \circ \psi_K.$$

Lemma

For $J \subsetneq [m]$ and $h(\mathbf{0}) = 0$:

$$\Lambda(h) \circ \psi_J = 0.$$

Sketch of proof:

- $\mu_\alpha = \mu(\alpha, [m])$ (Möbius function of the poset Π_m).
- Collect equal terms.
- Coefficients vanish by Weisner's theorem.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Berry–Essen Inequality

$\mathbf{X}$, $\mathbf{Y}$: m-dimensional random variables; $\mathbf{Y}$ continuous.

$$A_j = \sup_{\mathbf{y}\in\mathbb{R}^m} \frac{\partial F_{\mathbf{Y}}(\mathbf{y})}{\partial y_j}, \qquad B_j = \sum_{k=1}^{j} \left\{ {j \atop k} \right\} k!\ .$$

Berry–Essen Inequality

X, **Y**: m-dimensional random variables; **Y** continuous.

$$A_j = \sup_{\mathbf{y}\in\mathbb{R}^m} \frac{\partial F_{\mathbf{Y}}(\mathbf{y})}{\partial y_j}, \qquad B_j = \sum_{k=1}^{j} \left\{ {j \atop k} \right\} k!\ .$$

Theorem (Berry–Esseen Inequality; H.-Kropf 2016)

Let $T > 0$ be fixed. Then

$$\sup_{\mathbf{z}\in\mathbb{R}^m} |F_{\mathbf{X}}(\mathbf{z}) - F_{\mathbf{Y}}(\mathbf{z})| \le \frac{2}{(2\pi)^m} \int_{\|\mathbf{t}\|\le T} \left| \frac{\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})}{\prod_{\ell=1}^{m} t_\ell} \right| d\mathbf{t}$$
$$+ 2 \sum_{\emptyset\ne J\subsetneq[m]} B_{m-|J|} \sup_{\mathbf{z}_J\in\mathbb{R}^J} \left| F_{\mathbf{X}_J}(\mathbf{z}_J) - F_{\mathbf{Y}_J}(\mathbf{z}_J) \right|$$
$$+ \frac{2}{T} \left(\sqrt[3]{\frac{32}{\pi(1-(\frac{3}{4})^{1/m})}} + \frac{1}{2}\pi \right) \sum_{j=1}^{m} A_j.$$

Berry–Essen Inequality

$\mathbf{X}$, $\mathbf{Y}$: m-dimensional random variables; $\mathbf{Y}$ continuous.

$$A_j = \sup_{\mathbf{y}\in\mathbb{R}^m} \frac{\partial F_{\mathbf{Y}}(\mathbf{y})}{\partial y_j}, \qquad B_j = \sum_{k=1}^{j} \left\{ {j \atop k} \right\} k!\ .$$

Theorem (Berry–Esseen Inequality; H.-Kropf 2016)

Let $T > 0$ be fixed. Then

$$\sup_{\mathbf{z}\in\mathbb{R}^m} |F_{\mathbf{X}}(\mathbf{z}) - F_{\mathbf{Y}}(\mathbf{z})| \le \frac{2}{(2\pi)^m} \int_{\|\mathbf{t}\|\le T} \left| \frac{\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})}{\prod_{\ell=1}^{m} t_\ell} \right| d\mathbf{t}$$
$$+ 2 \sum_{\emptyset \ne J \subsetneq [m]} B_{m-|J|} \sup_{\mathbf{z}_J \in \mathbb{R}^J} \left| F_{\mathbf{X}_J}(\mathbf{z}_J) - F_{\mathbf{Y}_J}(\mathbf{z}_J) \right|$$
$$+ \frac{2}{T} \left(\sqrt[3]{\frac{32}{\pi(1-(\frac{3}{4})^{1/m})}} + \frac{1}{2}\pi \right) \sum_{j=1}^{m} A_j.$$

Dimension 2: Sadikova 1966.

Quasi-Power Theorem: Sketch of Proof

- Bound difference of characteristic functions by

$$|\Lambda(\varphi_{\mathbf{X}})(\mathbf{s}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{s})| \leq \exp\Big(-\frac{\sigma}{4}\|\mathbf{s}\|^2 + O(\|\mathbf{s}\|)\Big) O\Big(\frac{\|\mathbf{s}\|^3 + \|\mathbf{s}\|}{\sqrt{\phi_n}}\Big)$$

σ : smallest eigenvalue of $\Sigma = H_{\mathbf{u}}(\mathbf{0})$.

Quasi-Power Theorem: Sketch of Proof

- Bound difference of characteristic functions by

$$|\Lambda(\varphi_{\mathbf{X}})(\mathbf{s})-\Lambda(\varphi_{\mathbf{Y}})(\mathbf{s})| \le \exp\Big(-\frac{\sigma}{4}\|\mathbf{s}\|^2+O(\|\mathbf{s}\|)\Big)O\Big(\frac{\|\mathbf{s}\|^3+\|\mathbf{s}\|}{\sqrt{\phi_n}}\Big)$$

  σ : smallest eigenvalue of $\Sigma = H_{\mathbf{u}}(\mathbf{0})$.

- If $t_1, \ldots, t_k$ are small:

$$\Big|\frac{\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})}{t_1 \ldots t_k}\Big| \le \sup_{|t_1|,\ldots,|t_k|\le 1}\Big|\frac{\partial^k}{\partial t_1 \ldots \partial t_k}\big(\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})\big)\Big|$$

Quasi-Power Theorem: Sketch of Proof

- Bound difference of characteristic functions by

$$|\Lambda(\varphi_{\mathbf{X}})(\mathbf{s})-\Lambda(\varphi_{\mathbf{Y}})(\mathbf{s})| \le \exp\Big(-\frac{\sigma}{4}\|\mathbf{s}\|^2+O(\|\mathbf{s}\|)\Big)O\Big(\frac{\|\mathbf{s}\|^3+\|\mathbf{s}\|}{\sqrt{\phi_n}}\Big)$$

  σ : smallest eigenvalue of $\Sigma = H_{\mathbf{u}}(\mathbf{0})$.

- If $t_1, \ldots, t_k$ are small:

$$\Big|\frac{\Lambda(\varphi_{\mathbf{X}})(\mathbf{t})-\Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})}{t_1\ldots t_k}\Big| \le \sup_{|t_1|,\ldots,|t_k|\le 1}\Big|\frac{\partial^k}{\partial t_1\ldots\partial t_k}\big(\Lambda(\varphi_{\mathbf{X}})(\mathbf{t})-\Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})\big)\Big|$$

- Use Cauchy's formula to estimate derivative.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ

Quasi-Power Theorem: Sketch of Proof

- Bound difference of characteristic functions by

$$|\Lambda(\varphi_{\mathbf{X}})(\mathbf{s})-\Lambda(\varphi_{\mathbf{Y}})(\mathbf{s})| \le \exp\Big(-\frac{\sigma}{4}\|\mathbf{s}\|^2+O(\|\mathbf{s}\|)\Big)\,O\Big(\frac{\|\mathbf{s}\|^3+\|\mathbf{s}\|}{\sqrt{\phi_n}}\Big)$$

  σ : smallest eigenvalue of $\Sigma = H_{\mathbf{u}}(\mathbf{0})$.

- If $t_1, \ldots, t_k$ are small:

$$\Big|\frac{\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})}{t_1 \ldots t_k}\Big| \le \sup_{|t_1|,\ldots,|t_k|\le 1}\Big|\frac{\partial^k}{\partial t_1 \ldots \partial t_k}\big(\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})\big)\Big|$$

- Use Cauchy's formula to estimate derivative.
- Contribution of large variables $t_{k+1}, \ldots, t_m$ bounded by exponential decay.

ALPEN-ADRIA UNIVERSITÄT KLAGENFURT | WIEN GRAZ