Using computers to assist in discovering and proving new math theorems

Wei Sun

Concordia University
Montreal, Canada

based on joint work with Ze-Chun Hu, Ping Sun and Oliver Russell

Four-color theorem: first major theorem proved using a computer, K. Appel and W. Haken (1976)

Kepler conjecture: math theorem about sphere packing in $\mathbb{R}^3$.

T. Hales, ”A proof of the Kepler conjecture”, Annals of Math. (2005)

Riemann hypothesis: conjecture that the nontrivial zeros of the Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ all lie on the "critical line" $\mathrm{Re}\, s = \frac{1}{2}$.

The most extensive computer search (2020) for counterexamples of the Riemann hypothesis has verified it for $|\mathrm{Im}\, s| \leq 3.0001753328 \cdot 10^{12}$.

Variation comparison between infinitely divisible distributions and the normal distribution

Empirical rule: For a distribution of measurements that is approx. normal, it follows that the interval with end points

$$\mu \pm \sigma \text{ contains approx. } 68\% \text{ of measurements,}$$
$$\mu \pm 2\sigma \text{ contains approx. } 95\% \text{ of measurements,}$$
$$\mu \pm 3\sigma \text{ contains almost all of the measurements.}$$

What can we say about general (not necessarily symmetric) distributions?

Variation comparison inequality (Hu, Sun, Sun, 2023)

X: random variable with finite second moment, $Z \sim N(0,1)$.

$$P\left\{|X - E[X]| \le \sqrt{\mathrm{Var}(X)}\right\} > P\{|Z| \le 1\} \approx 0.6827$$

This inequality holds for most familiar infinitely divisible continuous distributions: Gamma, Laplace, Gumbel, Logistic, Pareto, (infinitely divisible) Weibull, log-normal, student's t, inverse Gaussian and F distributions.

Gamma distribution Let $\alpha, \beta > 0$ and $X_{\alpha,\beta}$ be Gamma random variable with pdf:

$$f_{\alpha,\beta}(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}, \quad x > 0.$$

Theorem 1

(Hu, Sun, Sun)

$$P\left\{|X_{\alpha,\beta} - E[X_{\alpha,\beta}]| \le \sqrt{\mathrm{Var}(X_{\alpha,\beta})}\right\} > P\{|Z| \le 1\} \approx 0.6827,$$

and

$$\inf_{\alpha,\beta} P\left\{|X_{\alpha,\beta} - E[X_{\alpha,\beta}]| \le \sqrt{\mathrm{Var}(X_{\alpha,\beta})}\right\} = P\{|Z| \le 1\}.$$

Sketch of proof By the central limit theorem,

$$\begin{aligned} & P\left\{|X_{\alpha,\beta} - E[X_{\alpha,\beta}]| \leq \sqrt{\mathrm{Var}(X_{\alpha,\beta})}\right\} \\ = \quad & P\{|X_{\alpha,\beta} - \alpha\beta| \leq \sqrt{\alpha}\beta\} \\ = \quad & P\{|X_{\alpha,1} - \alpha| \leq \sqrt{\alpha}\} \\ \to \quad & P\{|Z| \leq 1\} \text{ as } \alpha \to \infty. \end{aligned}$$

For $\alpha > 0$, define

$$t(\alpha) = P\{|X_{\alpha,1} - \alpha| \leq \sqrt{\alpha}\} = \int_{\max\{0,\alpha-\alpha^{\frac{1}{2}}\}}^{\alpha+\alpha^{\frac{1}{2}}} \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)} dy.$$

We need only show that $t(\alpha + 1) < t(\alpha)$.

It can be shown that for $\alpha > 1$, $t(\alpha+1) < t(\alpha)$ is equivalent to

$$\begin{aligned} & \int_0^{1+(\alpha+1)^{\frac{1}{2}}-\alpha^{\frac{1}{2}}} \left(1+\frac{y}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha} e^{-y} dy \\ < \quad & 1 \\ < \quad & \int_0^{1-(\alpha+1)^{\frac{1}{2}}+\alpha^{\frac{1}{2}}} \left(1+\frac{y}{\alpha-\alpha^{\frac{1}{2}}}\right)^{\alpha} e^{-y} dy \end{aligned}$$

It is a bit surprising this inequality is very delicate, which seems to be unknown in the literature.

For fixed α, $1-\left(1+\frac{y}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-y}$ is an increasing and concave function of y on $[0,1]$. It suffices to show that for $\alpha>1$,

$$\left\{1+4\left[(\alpha+1)^{\frac{1}{2}}-\alpha^{\frac{1}{2}}\right]\right\}\left(1+\frac{1}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-1}+2\left(1+\frac{1/2}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-1/2}<3.$$

For fixed α, $\left(1+\frac{y}{\alpha-\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-y}-1$ is an increasing and concave function of y on $[0,1]$. It suffices to show that for $\alpha>\left(\frac{15}{8}\right)^{2}$,

$$\left\{1-4[(\alpha+1)^{\frac{1}{2}}-\alpha^{\frac{1}{2}}]\right\}\left(1+\frac{1}{\alpha-\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-1}+2\left(1+\frac{1/2}{\alpha-\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-1/2}>3.$$

Define

$$\tau_+ := \frac{1}{\alpha+\alpha^{\frac{1}{2}}}, \quad \eta_+ := \frac{\tau_+}{2}.$$

We have

$$\begin{aligned}
&\left\{1+4[(\alpha+1)^{\frac{1}{2}}-\alpha^{\frac{1}{2}}]\right\}\left(1+\frac{1}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-1}+2\left(1+\frac{1/2}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha}e^{-1/2}\\
= \;&\left\{1+4[(\alpha+1)^{\frac{1}{2}}-\alpha^{\frac{1}{2}}]\right\}e^{-1+\alpha\ln(1+\tau_+)}+2e^{-\frac{1}{2}+\alpha\ln(1+\eta_+)}\\
< \;&\left\{1+4[(\alpha+1)^{\frac{1}{2}}-\alpha^{\frac{1}{2}}]\right\}e^{-1+\alpha\left(\tau_+-\frac{\tau_+^2}{2}+\frac{\tau_+^3}{3}-\frac{\tau_+^4}{4}+\frac{\tau_+^5}{5}\right)}\\
&+2e^{-\frac{1}{2}+\alpha\left(\eta_+-\frac{\eta_+^2}{2}+\frac{\eta_+^3}{3}-\frac{\eta_+^4}{4}+\frac{\eta_+^5}{5}\right)}.
\end{aligned}$$

Define

$$w := (\alpha+1)^{\frac{1}{2}} - \alpha^{\frac{1}{2}}.$$

By condition $\alpha > 1$, we have

$$0 < w < \frac{1}{2}, \quad 1 - w^2 > 0, \quad 1 + 2w - w^2 > 0.$$

$$\alpha = \frac{(1-w^2)^2}{4w^2}, \quad \tau_+ = \frac{4w^2}{(1-w^2)(1+2w-w^2)}, \quad \eta_+ = \frac{2w^2}{(1-w^2)(1+2w-w^2)}.$$

Define

$$P_+ := -1 + \alpha\left(\tau_+ - \frac{\tau_+^2}{2} + \frac{\tau_+^3}{3} - \frac{\tau_+^4}{4} + \frac{\tau_+^5}{5}\right),$$
$$Q_+ := -\frac{1}{2} + \alpha\left(\eta_+ - \frac{\eta_+^2}{2} + \frac{\eta_+^3}{3} - \frac{\eta_+^4}{4} + \frac{\eta_+^5}{5}\right).$$

We have

$$
\begin{aligned}
F_+ &:= 15(1-w^2)^3(1+2w-w^2)^5 P_+ \\
&= 2w(-15-135w-345w^2+190w^3+1735w^4+495w^5 \\
&\quad -3615w^6-716w^7+3615w^8+495w^9-1735w^{10} \\
&\quad +190w^{11}+345w^{12}-135w^{13}+15w^{14}).
\end{aligned}
$$

Set $w := \frac{1}{2(1+q^2)}$. We get

$$
\begin{aligned}
G_+ &:= 16384(1+q^2)^{14}\cdot\frac{F_+}{2w} \\
&= -1140603-17129046q^2-115786348q^4-468301840q^6 \\
&\quad -1267262160q^8-2427446688q^{10}-3393664576q^{12} \\
&\quad -3517163008q^{14}-2715321600q^{16}-1554209280q^{18}-649507840q^{20} \\
&\quad -192286720q^{22}-38154240q^{24}-4546560q^{26}-245760q^{28},
\end{aligned}
$$

which implies that P_+ is negative.

We have

$$\begin{aligned} H_+ &:= 30(1-w^2)^3(1+2w-w^2)^5 Q_+ \\ &= w(-30-255w-600w^2+410w^3+2900w^4+705w^5 \\ &\quad -5550w^6-1672w^7+5550w^8+705w^9-2900w^{10} \\ &\quad +410w^{11}+600w^{12}-255w^{13}+30w^{14}), \end{aligned}$$

and

$$\begin{aligned} I_+ &:= 8192(1+q^2)^{14}\cdot\frac{H_+}{w} \\ &= -1083048-16069911q^2-108024568q^4-435858040q^6 \\ &\quad -1178745360q^8-2259543408q^{10}-3165284416q^{12} \\ &\quad -3291555328q^{14}-2553515520q^{16}-1471031040q^{18}-619724800q^{20} \\ &\quad -185251840q^{22}-37171200q^{24}-4485120q^{26}-245760q^{28}, \end{aligned}$$

which implies that Q_+ is negative.

Thus,

$$
\begin{aligned}
& \left\{1+4[(\alpha+1)^{\frac{1}{2}}-\alpha^{\frac{1}{2}}]\right\}\left(1+\frac{1}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha} e^{-1} \\
& +2\left(1+\frac{1/2}{\alpha+\alpha^{\frac{1}{2}}}\right)^{\alpha} e^{-1/2}-3 \\
< \quad & (1+4w)e^{P_+}+2e^{Q_+}-3 \\
< \quad & (1+4w)\left(1+P_+ +\frac{P_+^2}{2}+\frac{P_+^3}{3!}+\frac{P_+^4}{4!}\right) \\
& +2\left(1+Q_+ +\frac{Q_+^2}{2}+\frac{Q_+^3}{3!}+\frac{Q_+^4}{4!}\right)-3 \\
:= \quad & R_+ .
\end{aligned}
$$

Define

$$L_+ := 9720000(1-w^2)^{12}(1+2w-w^2)^{20}w^{-3}R_+,$$

and

$$V_+ := -18014398509481984(1+q^2)^{62}L_+.$$

By virtue of Mathematica, we can show that R_+ is negative:

```
V+=23565171557938261664962395 +
 1985238765536369188253388462 q^2 +
 76017937191609745093093565184 q^4 +
 1815476155917282265018752272232 q^6 +
 30868042081839055982554050213660 q^8 +
 401897536051918258546845673711320 q^10 +
 4195397709111549929883773768957292 q^12 +
 36238699732610615067411794056699104 q^14 +
 265002286089679374723172860122766982 q^16 +
 1669237124849349342077586449716389470 q^18 +
 9179934813394932229977676436328785920 q^20 +
 44555295354320392501114611345123622400 q^22 +
 192537160208281140648975165919934835200 q^24 +
 746181252269526741637909507751082171520 q^26 +
 2609372572626683435719917787491018652160 q^28 +
 8276209631283583168755734561689661224960 q^30 +
 23913569456882144063241575623509484876800 q^32 +
```

```
631858518257554841619606681722926999909120 q^34 +
153171842040744452342444666253152790732800 q^36 +
341628452529444844018632398179833131991040 q^38 +
702775058219816773204544225960017412751360 q^40 +
1336281286191807830241756821296507838955520 q^42 +
2352931028757956298911312671496544634142720 q^44 +
3842851078067257537573706091171559356497920 q^46 +
5829597689354288278514821031866786159656960 q^48 +
8224048629268397888867021111129844469596160 q^50 +
10800344227796355322915422778734559920914432 q^52 +
13214925874596962589048294078754839901241344 q^54 +
15075437985869745487585690312752934894436352 q^56 +
16043253396277674458218536937392302561689600 q^58 +
15933483495160554733717977505944446676500480 q^60 +
14772181225346687498328707734409833005187072 q^62 +
12786642631638024827680853914736629691449344 q^64 +
```

```
10333607587858511627607426462208954269171712 q^66 +
7796135691442288431829216828566360534548480 q^68 +
5489455045169343972022956242977322445045760 q^70 +
3606053976460179205452292516224406577479680 q^72 +
2208802483012122361676530694278736018145280 q^74 +
1260683716848402925787749700503070572544000 q^76 +
669904634456217504368236284579198848204800 q^78 +
331081492020125941589702200450146757509120 q^80 +
152001230583526972614803279225239581491200 q^82 +
64734238129983061042604032362158017740800 q^84 +
25531661312772564369272230408940525977600 q^86 +
9307900274880934185285303838052660019200 q^88 +
3129622227458365558314112917099983667200 q^90 +
968033669852811173795309928049750835200 q^92 +
274640462267821497605095290057418342400 q^94 +
```

```
71224038857339584104332613373237657600 q^96 +
16816854347488682979216179348688076800 q^98 +
3598246400042953802386878316412928000 q^100 +
693860503839280523254707460767744000 q^102 +
119794916777653504670468143054848000 q^104 +
18371891299642536871251828277248000 q^106 +
2478647665316721166509051740160000 q^108 +
290656583441325139861690122240000 q^110 +
29171448597603811259616067584000 q^112 +
2455470675466333890778497024000 q^114 +
168582129968383522599075840000 q^116 +
9065459012974557989437440000 q^118 +
358068461185282678456320000 q^120 +
9236522547766697656320000 q^122 +
116733302341443256320000 q^124.
```

The variation comparison inequality

$$P\left\{|X-E[X]|\le\sqrt{\mathrm{Var}(X)}\right\}>P\{|Z|\le 1\}\approx 0.6827$$

also holds for Laplace, Gumbel, Logistic and Pareto distributions.

Weibull distribution The pdf of Weibull random variable with parameters $k, \lambda > 0$ is

$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Weibull distribution is infinitely divisible iff $k \in (0,1]$.

Theorem 2

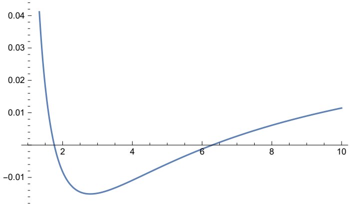
(Hu, Sun, Sun) Let $\lambda > 0$, $0 < k \leq 1$, $X_{\lambda,k}$ be a Weibull random variable with parameters λ and k. Then,

$$P\left\{|X_{\lambda,k} - E[X_{\lambda,k}]| \leq \sqrt{\mathrm{Var}X_{\lambda,k})}\right\} > P\{|Z| \leq 1\} \approx 0.6827.$$

Define

$$W_k := P\left\{|X_{\lambda,k} - E[X_{\lambda,k}]| \leq \sqrt{\mathrm{Var}(X_{\lambda,k})}\right\}.$$

Graph of the function $(W_k - 0.6827)$ for $k \in [1, 10]$:



Log-normal distribution The pdf of log-normal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ is

$$f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0,$$

Theorem 3

(Hu, Sun, Sun) Let $\mu \in \mathbb{R}$, $\sigma > 0$ and $X_{\mu,\sigma}$ be a log-normal random variable with parameters μ and σ. Then,

$$P\left\{|X_{\mu,\sigma} - E[X_{\mu,\sigma}]| \leq \sqrt{\mathrm{Var}(X_{\mu,\sigma})}\right\} > P\{|Z| \leq 1\} \approx 0.6827,$$

and

$$\inf_{\mu,\sigma} P\left\{|X_{\mu,\sigma} - E[X_{\mu,\sigma}]| \leq \sqrt{\mathrm{Var}(X_{\mu,\sigma})}\right\} = P\{|Z| \leq 1\}.$$

Theorem 4

(Hu, Sun, Sun) Let $\nu \geq 3$ and X_ν be a student's t-random variable with ν degrees of freedom. Then,

$$P\left\{|X_\nu - E[X_\nu]| \leq \sqrt{\mathrm{Var}(X_\nu)}\right\} > P\{|Z| \leq 1\} \approx 0.6827,$$

and

$$\inf_\nu P\left\{|X_\nu - E[X_\nu]| \leq \sqrt{\mathrm{Var}(X_\nu)}\right\} = P\{|Z| \leq 1\}.$$

Idea of proof: using the Gaussian hypergeometric function:

$$F(a,b;c;z) = \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \cdot \frac{z^j}{j!}, \quad |z| < 1,$$

where $(\alpha)_j := \alpha(\alpha+1)\cdots(\alpha+j-1)$ for $j \geq 1$.

The pdf of inverse Gaussian (also known as Wald) random variable $X_{\mu,\lambda}$ is given by

$$f_{\mu,\lambda}(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right), \quad x > 0.$$

Theorem 5

(Hu, Sun, Sun) Let $\mu, \lambda > 0$ and $X_{\mu,\lambda}$ be an inverse Gaussian random variable with parameters μ and λ. Then,

$$P\left\{|X_{\mu,\lambda} - E[X_{\mu,\lambda}]| \le \sqrt{\mathrm{Var}(X_{\mu,\lambda})}\right\} > P\{|Z| \le 1\} \approx 0.6827,$$

and

$$\inf_{\mu,\lambda} P\left\{|X_{\mu,\lambda} - E[X_{\mu,\lambda}]| \le \sqrt{\mathrm{Var}(X_{\mu,\lambda})}\right\} = P\{|Z| \le 1\}.$$

A useful tool: Let Φ be the cdf of Z. Denote the complementary error function by

$$\mathrm{erfc}(x)=\frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt,\quad x\in\mathbb{R}.$$

We have

$$\Phi(x)=\frac{1}{2}\mathrm{erfc}\left(-\frac{x}{\sqrt{2}}\right)$$

and the following asymptotic expansion

$$\begin{aligned}\mathrm{erfc}(x) &= \frac{1}{\sqrt{\pi}}e^{-x^2}\left(\frac{1}{x}-\frac{1}{2x^3}+\frac{1\cdot 3}{2^2x^5}-\cdots+(-1)^{n-1}\frac{(2n-3)!!}{2^{n-1}x^{2n-1}}\right)\\ &\quad+(-1)^n\frac{(2n-1)!!}{2^{n-1}\sqrt{\pi}}\int_x^\infty\frac{e^{-t^2}}{t^{2n}}dt.\end{aligned}$$

Theorem 6

(Hu, Sun, Sun) Let $d_1 \in \{1,2,3,4\}$, $5 \le d_2 \in \mathbb{N}$, X_{d_1,d_2} be an F-random variable with parameters d_1 and d_2. Then,

$$P\left\{|X_{d_1,d_2} - E[X_{d_1,d_2}]| \le \sqrt{\mathrm{Var}(X_{d_1,d_2})}\right\} > P\{|Z| \le 1\} \approx 0.6827,$$

and

$$\begin{aligned} &\inf_{d_2 \in \mathbb{N}} P\left\{|X_{d_1,d_2} - E[X_{d_1,d_2}]| \le \sqrt{\mathrm{Var}(X_{d_1,d_2})}\right\} \\ = \quad & P\left\{|\chi^2(d_1) - E[\chi^2(d_1)]| \le \sqrt{\mathrm{Var}(\chi^2(d_1))}\right\}. \end{aligned}$$

Sketch of proof Define

$$A := \frac{d_1}{d_1 + d_2\left(1+\sqrt{\frac{2(d_1+d_2)}{d_1(d_2-2)}}\right)^{-1}},$$

$$B := \frac{d_1}{d_1 + (d_2-2)\left(1+\sqrt{\frac{2(d_1+d_2-2)}{d_1(d_2-4)}}\right)^{-1}},$$

$$C := \frac{d_1}{d_1 + d_2\left(1-\sqrt{\frac{2(d_1+d_2)}{d_1(d_2-2)}}\right)^{-1}},$$

$$D := \frac{d_1}{d_1 + (d_2-2)\left(1-\sqrt{\frac{2(d_1+d_2-2)}{d_1(d_2-4)}}\right)^{-1}}.$$

Prove the following two inequalities:

$$2A^{\frac{d_1}{2}}(1-A)^{\frac{d_2}{2}} < d_2\int_A^B t^{\frac{d_1}{2}-1}(1-t)^{\frac{d_2}{2}-1}dt,$$
$$2C^{\frac{d_1}{2}}(1-C)^{\frac{d_2}{2}} > d_2\int_C^D t^{\frac{d_1}{2}-1}(1-t)^{\frac{d_2}{2}-1}dt.$$

Obtain delicate estimates through asymptotic expansions with the help of Mathematica.

For the case $d_1 \geq 5$, we need prove delicate combinatorial inequalities.

The variation comparison inequality

$$P\left\{|X - E[X]| \le \sqrt{\mathrm{Var}(X)}\right\} > P\{|Z| \le 1\} \approx 0.6827$$

also holds for the geometric distribution. But it should be modified with continuity correction when general discrete distributions are considered.

Remarks

Let X_3 be a Poisson random variable with parameter 3.

$$P\left\{|X_3 - E[X_3]| \le \sqrt{\mathrm{Var}(X_3)}\right\} = 0.6161 < 0.6827.$$

Let $B \sim N(0,1)$ be independent of X_3. Define

$$X_{B,\varepsilon} := \varepsilon B + X_3.$$

$$\lim_{\varepsilon \to 0} P\left\{|X_{B,\varepsilon} - E[X_{B,\varepsilon}]| \le \sqrt{\mathrm{Var}(X_{B,\varepsilon})}\right\} = 0.6161 < 0.6827.$$

Hence, the variation comparison inequality does not hold for all infinitely divisible continuous distributions.

For $n \in \mathbb{N}$, define

$$\nu_n(dx) = \frac{3n}{2} \cdot 1_{[1-\frac{1}{n}, 1+\frac{1}{n}]}(x)dx.$$

Let Y_n be a compound Poisson random variable with Lévy measure ν_n. Then, Y_n converges to X_3 in distribution as $n \to \infty$. Hence, Y_n does not satisfy the variation comparison inequality at least if n is large enough.

This simple example shows that the inequality might not hold even if Lévy measure is absolutely continuous w.r.t. Lebesgue measure.

Riemann zeta distribution

Riemann zeta function ζ is defined initially for complex numbers $z = u + iv$ where $u > 1$ by

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z} = \prod_{p \in \mathcal{P}} \frac{1}{1 - p^{-z}},$$

where $\mathcal{P}$ denotes the set of all prime numbers.

For $u \in \mathbb{R}$ with $u > 1$, define $\phi_u : \mathbb{R} \mapsto \mathbb{C}$ by

$$\phi_u(v) := \frac{\zeta(u + iv)}{\zeta(u + i0)}, \quad v \in \mathbb{R}.$$

Theorem 7

(Khintchine) For each $u > 1$, ϕ_u is the characteristic function of an infinitely divisible distribution.

Let X_u be an infinitely divisible random variable with Lévy measure

$$\nu_u := \sum_{p\in\mathcal{P}} \sum_{m=1}^{\infty} \frac{e^{ux}}{m} \delta_{-m\log(p)}(dx).$$

We assume that X_u does not have a Gaussian term.

Does the following variation comparison inequality hold?

$$P\left\{|X_u - E[X_u]| \leq \sqrt{\mathrm{Var}(X_u)}\right\} > P\{|Z| \leq 1\} \approx 0.6827$$

An interesting false conjecture

$\pi(x)$: the number of primes $\leq x$.

$\mathrm{li}(x) = \int_0^x \frac{dt}{\ln t}$.

It has been computed that $\pi(x) < \mathrm{li}(x)$ for all $x \leq 10^{25}$ and no value of x is known for which $\pi(x) > \mathrm{li}(x)$. However, in 1914 Littlewood proved that the difference $\pi(x) - \mathrm{li}(x)$ changes sign infinitely many times.

This example shows the danger of basing conjectures on numerical evidence!

Equivalence of Riemann hypothesis and infinite divisibility

For $t \geq 0$, define

$$\begin{aligned} g_\zeta(t) \quad := \quad & -4(e^{t/2}+e^{-t/2}-2)+\sum_{n\leq e^t}\frac{\Lambda(n)}{\sqrt{n}}(t-\log n) \\ & -\frac{t}{2}(\psi(1/4)-\log\pi)+\frac{1}{4}\left[e^{-t/2}\Phi(e^{-2t},2,1/4)-\Phi(1,2,1/4)\right], \end{aligned}$$

where $\psi(z)=\Gamma'(z)/\Gamma(z)$ is dgamma function,

$$\Lambda(n)=\begin{cases}\ln(p), & \exists \text{ prime } p \text{ and } k\in\mathbb{Z}^+,\ n=p^k,\\ 0, & \text{otherwise.}\end{cases}$$

is Mangoldt function and $\Phi(z,s,a):=\sum_{n=0}^{\infty}\frac{z^n}{(n+a)^s}$ is Hurwitz-Lerch zeta function. Define $g_\zeta(t)=g_\zeta(-t)$ for $t<0$.

The function g_ζ was originally introduced to study equivalent conditions for the RH in relation to Weil's positivity or Li's criterion, etc.

Theorem 8

(Nakamura and Suzuki, 2023) RH is true iff $exp(g_\zeta(t))$ is the characteristic function of an infinitely divisible distribution on $\mathbb{R}$.

Bell's theorem: If certain predictions of quantum theory are correct, then our world is non-local.

Experiments establish that our world is non-local. Very surprising, since non-locality is normally taken to be prohibited by the theory of relativity.

Bell's inequality

Random variables: Z^i_α, $i = 1, 2$, $\alpha = a, b, c$, taking only values ± 1. If $Z^1_\alpha = -Z^2_\alpha$, $\forall \alpha$ (perfectly anti-correlated), then

$$P(Z^1_a \neq Z^2_b) + P(Z^1_b \neq Z^2_c) + P(Z^1_c \neq Z^2_a) \geq 1.$$

Proof of Bell's theorem: EPR argument + Bell's inequality.

Inequalities involving Gaussian distributions are related to various fields: e.g., small-ball probabilities, zeros of random polynomials.

Royen (2014): Gaussian correlation inequality.

For any convex symmetric sets K, L in $\mathbb{R}^n$ and any centered Gaussian measure μ we have

$$\mu(K \cap L) \geq \mu(K)\mu(L).$$

Gaussian product inequality conjecture

Li and Wei (2012) For any non-negative real numbers y_j, $j = 1, \ldots, n$, and any n-dimensional real-valued centered Gaussian random vector $(X_1, \ldots, X_n)$,

$$E\left[\prod_{j=1}^{n} |X_j|^{y_j}\right] \geq \prod_{j=1}^{n} E[|X_j|^{y_j}].$$

Russell and Sun, opposite GPI, (2022) Let (X_1, X_2) be centered bivariate Gaussian random variables, $-1 < y_1 < 0$ and $y_2 > 0$. Then,

$$E[|X_1|^{y_1}|X_2|^{y_2}] \leq E[|X_1|^{y_1}]E[|X_2|^{y_2}],$$

and the equality sign holds if and only if X_1, X_2 are independent.

GPI $\Rightarrow$ 'real linear polarization constant' conjecture

For any $n \geq 2$, and any collection $x_1, \ldots, x_n$ of unit vectors in $\mathbb{R}^n$, there exists a unit vector $v \in \mathbb{R}^n$ such that

$$|\langle v, x_1 \rangle \cdot \cdots \cdot \langle v, x_n \rangle| \geq n^{-n/2}.$$

Hu, Lan and Sun (2019) For any 3-dimensional centered Gaussian random vector (X,Y,Z),

$$E\left[X^{2m}Y^{2m}Z^{2n}\right] \geq E[X^{2m}]E[Y^{2m}]E[Z^{2n}], \quad \forall m,n \in \mathbb{N}.$$

The equality holds if and only if X,Y,Z are independent.

Intrinsic connection between moments of Gaussian distributions and the Gaussian hypergeometric functions.

New combinatorial identities

Let $l, r \in \mathbb{N}$ satisfying $l \leq r$. Then we have

$$\sum_{i=0}^{l-1} \frac{\binom{2r}{i}\binom{l-1}{i}}{\binom{2r-l}{i}} = \frac{(2r)!}{2r!r!\binom{2r-l}{r}},$$

$$\sum_{i=0}^{l-1} \frac{\binom{l-1}{i}}{\binom{2r-i}{l}} = \frac{1}{2\binom{r}{l}}.$$

Connection with the classical Kummer's identity.

Russell and Sun (2022) Let $(X_1, \ldots, X_n)$ be a centered Gaussian random vector such that $E[X_iX_j] \geq 0$ for any $i \neq j$. Then,

$$E\left[\prod_{j=1}^{n} X_j^{2m_j}\right] \geq E\left[\prod_{j=1}^{k} X_j^{2m_j}\right] E\left[\prod_{j=k+1}^{n} X_j^{2m_j}\right], \quad \forall 1 \leq k \leq n-1,$$

and

$$E\left[\prod_{j=1}^{n} X_j^{2m_j}\right] \geq \prod_{j=1}^{n} E[X_j^{2m_j}].$$

Idea of proof: using the Isserlis-Wick formula.

SOS representation of GPI

Solve the GPI by showing that

$$E\left[\prod_{j=1}^{n} X_j^{2m_j}\right] - \prod_{j=1}^{n} E[X_j^{2m_j}]$$

has an SOS representation.

A sums-of-squares (SOS) representation of a polynomial is of the form $\sum_{i=1}^{p} f_i^2$, where the f_i's are real-coefficient polynomials.

Non-negative multivariate polynomial may not be SOS.

$P_{n,2d}$: all non-negative polynomials in n variables of degree at most $2d$.

$\sum_{n,2d}$: all polynomials in $P_{n,2d}$ that are SOS.

Hilbert (1888) $\sum_{n,2d} = P_{n,2d}$ if and only if

- $n = 1$ or $d = 1$ or $(n, d) = (2, 2)$.

Motzkin polynomial (1967) $M(x, y) = x^2y^4 + x^4y^2 + 1 - 3x^2y^2$.

$E\left[\prod_{j=1}^{n} X_j^{2m_j}\right]$ itself is an SOS.

Lemma Let $(X_1, \ldots, X_n)$ be a centered Gaussian random vector. Denote by Λ the covariance matrix of $(X_1, \ldots, X_n)$ and $c_{m_1,\ldots,m_n}$ the coefficient corresponding to the term $t_1^{2m_1} \cdots t_n^{2m_n}$ of the polynomial

$$G(t_1, \ldots, t_n) = \left(\sum_{k,l=1}^{n} \Lambda_{kl} t_k t_l \right)^{\sum_{j=1}^{n} m_j}, \quad t_1, \ldots, t_n \in \mathbb{R}.$$

Then,

$$E\left[\prod_{j=1}^{n} X_j^{2m_j} \right] = \frac{\prod_{j=1}^{n} (2m_j)!}{2^{\sum_{j=1}^{n} m_j} \left(\sum_{j=1}^{n} m_j \right)!} \cdot c_{m_1,\ldots,m_n}. \tag{1}$$

Let U_j, $1 \le j \le n$, be independent standard Gaussian random variables. Define

$$X_k = \sum_{j=1}^{n} x_{kj} U_j, \quad 1 \le k \le n,$$

where each $x_{kj} \in \mathbb{R}$, $1 \le k, j \le n$. Then, we have

$$\Lambda_{kk} = \sum_{j=1}^{n} x_{kj}^2, \quad 1 \le k \le n,$$
$$\Lambda_{kl} = \sum_{j=1}^{n} x_{kj} x_{lj}, \quad 1 \le k < l \le n. \tag{2}$$

Define

$$\begin{aligned} F_{m_1,\ldots,m_n}(\Lambda) &= E\left[\prod_{j=1}^{n} X_j^{2m_j}\right] - \prod_{j=1}^{n} E[X_j^{2m_j}] \\ &= E\left[\prod_{j=1}^{n} X_j^{2m_j}\right] - \prod_{k=1}^{n} [(2m_k - 1)!!\Lambda_{kk}^{m_k}]. \end{aligned}$$

By (1) and (2), it is easy to see that $F_{m_1,\ldots,m_n}(\Lambda)$ can be expressed as a polynomial of the x_{ij}'s, say $F_{m_1,\ldots,m_n}(x_{11},\ldots,x_{1n},\ldots,x_{n1},\ldots,x_{nn})$.

5D GPI using improved SOS method

Russell and Sun (2023) Let $y \in \left[\frac{1}{10}, \infty\right)$. For any centered Gaussian random vector $(X_1, X_2, X_3, X_4, X_5)$,

$$E[|X_1|^y X_2^2 X_3^2 X_4^2 X_5^2] \geq E[|X_1|^y] E[X_2^2] E[X_3^2] E[X_4^2] E[X_5^2].$$

Moment ratio inequality of bivariate Gaussian distribution

Define

$$\mathcal{S} := \{(1, m_3) : m_3 \geq 5\} \bigcup \{(2, m_3) : m_3 \geq 3\} \bigcup \{(m_2, m_3) : m_3 \geq m_2 \geq 3\},$$

$$r_{m_2,m_3} = (2m_2 + 1)(2m_3 + 1) + 1, \quad t_{m_2,m_3} = \frac{1}{r_{m_2,m_3} + \left(1 + \frac{1}{2m_2}\right)\left(1 + \frac{1}{2m_3}\right)},$$

and for $\frac{1}{r^2_{m_2,m_3}} < z \leq 1$,

$$H_{m_2,m_3}(z) = \frac{(m_2 + m_3 + 1)(r_{m_2,m_3} z - 1) + \sqrt{[(m_3 - m_2)(r_{m_2,m_3} z - 1)]^2 + (2m_2 + 1)^3 (2m_3 + 1)^3 z}}{r^2_{m_2,m_3} z - 1}.$$

Russell and Sun (2023) Let (X_2, X_3) be a centered Gaussian random vector. If $(m_2, m_3) \in \mathcal{S}$, then

$$\frac{\left|E[X_2^{2m_2+1} X_3^{2m_3+1}]\right|}{(2m_2+1)(2m_3+1)E[X_2^{2m_2} X_3^{2m_3}]} \leq \begin{cases} |\mathrm{Cov}(X_2, X_3)|, & \text{if } |\mathrm{Corr}(X_2, X_3)| \leq \sqrt{t_{m_2,m_3}}, \\ H_{m_2,m_3}([\mathrm{Corr}(X_2, X_3)]^2) \cdot |\mathrm{Cov}(X_2, X_3)|, & \text{if } \sqrt{t_{m_2,m_3}} < |\mathrm{Corr}(X_2, X_3)|. \end{cases}$$

The equality sign holds if and only if X_2 and X_3 are independent.