De la funció W de Lambert a la raó de Mills: tres articles de probabilitat

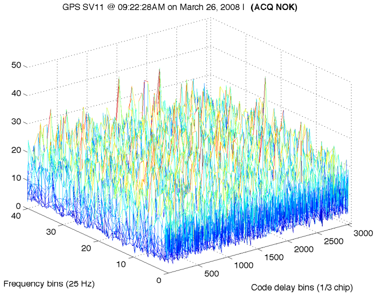
Frederic Utzet

11 de juny de 2019

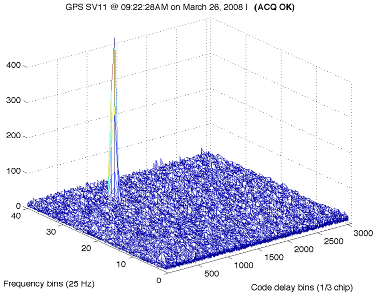
- A. GASULL, J. A. LÓPEZ–SALCEDO, FU,
  Maxima of Gamma random variables and other Weibull–like distributions and the Lambert W *function*,
  Test, **24** (2015) 714–733.
- A. GASULL, M. JOLIS, FU,
  On the norming constants for normal maxima,
  Journ. Mathem. Anal. Appl., **422** (2015) 376–396.
- A. GASULL, FU,
  Approximating Mills ratio,
  Journ. Mathem. Anal. Appl., **420** (2014) 1832–1853.

Un problema de processament de senyals

- Context d'un sistema de navegació global per satel·lit (GNSS), aplicat a localització d'exteriors o navegació espacial.
- Un receptor proporciona informació de la posició processant senyals emesos per satel·lits que orbiten al voltant de la Terra.
- Un dels punts claus del sistema és decidir quan es rep un senyal de determinat satel·lit.



Només soroll



Soroll i senyal

Descripció del soroll

El soroll està format per n (entre 10 i 10^6) variables aleatòries independents amb llei $\chi^2(m)$, amb m típicament 20. Aquesta distribució apareix perquè cada valor del soroll correspon a

$$X = Y_1 + \cdots Y_{m/2},$$

on cada Y_j és el mòdul al quadrat d'una variable normal complexa.

¿Hi ha senyal o només soroll?

Discriminar si es detecta un satèl·lit o no correspon a un test d'hipòtesis:

H_0: Només hi ha soroll contra H_1: Hi ha un senyal.

El test consisteix en calcular el valor màxim del soroll i mirar si passa d'un llindar:

$$\max\{X_1, \ldots, X_n\} > c,$$

on $X_1, \ldots, X_n$ són variables aleatòries independents amb llei $\chi^2(m)$.

Convergència dels valors extrems

$\{X_n,\ n \geq 1\}$ una successió de variables aleatòries independents, totes amb la mateixa distribució. Sigui

$$M_n = \max\{X_1, \ldots, X_n\}.$$

Teorema de Fisher-Tippet (1928). Sota certes condicions es poden trobar successions de nombres (constants normalitzadores) $\{a_n,\ n \geq 1\}$ i $\{b_n,\ n \geq 1\}$ tals que

$$\lim_n \frac{M_n - b_n}{a_n} = M, \text{ en distribució,}$$

on M és una variable aleatòria que només pot ser d'un dels tres tipus següents:

Les distribucions dels valors extrems

1. Gumbel: amb funció de distribució

$$\Lambda(x) = \exp\{-e^{-x}\}, \; x \in \mathbb{R}.$$

2. Frechet ($a > 0$):

$$F(x) = \begin{cases} 0, & \text{si } x \leq 0, \\ \exp\{-x^{-a}\}, & \text{si } x > 0. \end{cases}$$

3. Weibull ($a > 0$):

$$F(x) = \begin{cases} \exp\{-(-x)^{a}\}, & \text{si } x \leq 0, \\ 1, & \text{si } x > 0. \end{cases}$$

El cas del màxim de χ^2

Quan $X_1, \ldots, X_n$ són $\chi^2(m)$, o, més generalment, lleis de tipus Weibull, aleshores estem en el cas de Gumbel i les constants normalitzadores es calculen per:

$$b_n = F^{-1}(1 - n^{-1}) \quad \text{i} \quad a_n = \frac{1 - F(b_n)}{f(b_n)}$$

on F és la funció de distribució de la llei $\chi^2(m)$, i f la seva funció de densitat.

Però la funció F^{-1} no té expressió explícita, i d'altra banda, hi ha molt marge per triar les constants normalitzadores: successions asimptòticament equivalents a $\{a_n\}$ i $\{b_n\}$ també serveixen.

Les constants normalitzadores estàndard

Els llibres d'estadística de valors extrems contenen valors de $\{a_n\}$ i $\{b_n\}$, calculats per mètodes *ad hoc*, i acceptats per tothom. Es diuen les constants normalitzadores estàndard.

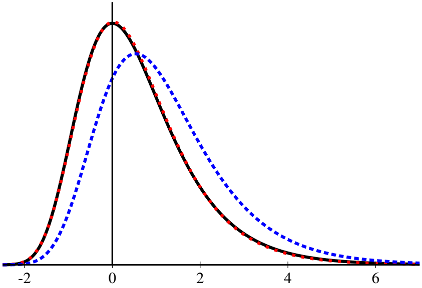
El cas més fàcil

$\{X_n,\ n \geq 1\}$ independents, totes amb funció de distribució

$$F(x) = 1 - e\,x\,e^{-x}, x \geq 1,$$

Les constants estàndard són

$$b_n' = \log n + \log\log n + 1 \quad \text{and} \quad a_n' = 1.$$



Línia vermella: densitat de Gumbel.
Línia blava: densitat de $(M_n - b_n)/a_n$ amb $n = 100$.

El nostre punt de partida

$$F(x) = 1 - e\,x\,e^{-x}, x \geq 1,$$

Volem calcular

$$b_n = F^{-1}(1 - 1/n).$$

El nostre punt de partida

$$F(x) = 1 - e\,x\,e^{-x}, x \geq 1,$$

Volem calcular

$$b_n = F^{-1}(1 - 1/n).$$

Això és fa amb la funció de Lambert!

La funció de Lambert

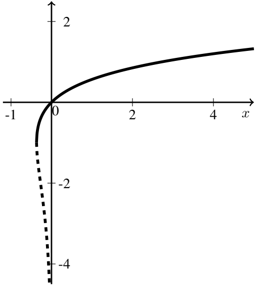
En el cas real, la funció de Lambert W es defineix implícitament a través de la solució real de l'equació

$$W(x)\, e^{W(x)} = x.$$

W té dues branques:

1. La principal, W_0, definida a $[-1/e, \infty)$

2. La branca negativa, W_{-1}, definida a $[-1/e, 0)$ i que compleix $W_{-1}(x) \leq -1$ i $\lim_{x\to 0^-} W_{-1}(x) = -\infty$.



Línia contínua: branca principal
Línia discontínua: branca negativa

Càlcul de les constants normalitzadores per la funció de Lambert

$$b_n = -W\big(-1/(en)\big).$$

Però per construcció, necessitem que $\lim_n b_n = \lim_n F^{-1}(1-n^{-1}) = \infty$. Aleshores hem d'utilitzar la branca negativa.

$$b_n = -W_{-1}\big(-1/(en)\big).$$

per calcular valors aproximats, utilitzarem el desenvolupament asimptòtic de la funció de Lambert (de Bruijn): per a $x \to 0^-$,

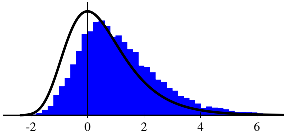
$$W_{-1}(x) = L_1(-x) - L_2(-x) + \frac{L_2(-x)}{L_1(-x)} + \frac{L_2(-x)(-2+L_2(-x))}{2L_1^2(-x)} + O\Big(\frac{L_2^3(-}{L_1^3(-}$$

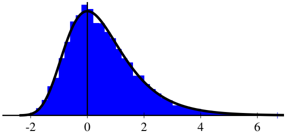
on

$$L_1(x) = \log x \quad \text{i} \quad L_2(x) = \log|\log x|.$$

Noves constants

Simulació de 10^4 màxims de 100 observacions de variables aleatòries amb funció de distribució $F(x) = 1 - e\,x\,e^{-x}$, normalitzats amb constants estàndards i noves constants. La línia contínua correspon a la densitat de la llei límit (Gumbel)





Amb constants estàndard

Amb noves constants

El teorema final

Let $X_1, \ldots, X_n$ be i.i.d. random variables with generalized Weibull distribution, with distribution function

$$F(x) = 1 - Kx^{\alpha} \exp\{-Cx^{\tau}\}, \; x \geq x_0,$$

with K, C, $\alpha > 0$ and $\tau \geq 1$.

1. Let $b_n = F^{-1}(1 - n^{-1})$ and $a_n = (C\tau b_n^{\tau-1} - \alpha/b_n)^{-1}$. Then, when $n \to \infty$,

$$F^n(a_n x + b_n) - \Lambda(x) = \begin{cases} g_1(x)/\log n + O(\log\log n/\log^2 n), & \text{if } \tau > 1, \\ g_2(x)/\log^2 n + O(\log\log n/\log^3 n), & \text{if } \tau = 1, \end{cases}$$

where $\Lambda(x) = \exp\{-e^{-x}\}$ is the distribution function of the Gumbel law, the functions g_1 and g_2 are bounded, and the big O terms depend on x.

El teorema, 2a. part

2. If $\widetilde{a}_n$ and $\widetilde{b}_n$ satisfy

$$\widetilde{b}_n^\tau = b_n^\tau + T_n \quad \text{and} \quad \widetilde{a}_n = 1/\big(C\tau\widetilde{b}_n^{\tau-1}\big) + O\big(1/\widetilde{b}_n^{2\tau-1}\big),$$

where $\lim_n \log n\, T_n^2 = 0$ and $\lim_n n\, T_n^2 = \infty$, then, when $n \to \infty$,

$$F^n(\widetilde{a}_n x + \widetilde{b}_n) - \Lambda(x) = h(x)T_n + O\big(1/\log n\big).$$

where the function h is bounded, and the big O term depends on x.

A. GASULL, M. JOLIS, FU,

On the norming constants for normal maxima

Journ. Mathem. Anal. Appl., **422** (2015) 376–396.

El cas de variables normals

Quan les variables X_n són normals també estem en el cas Gumbel. Sigui $M_n = \max\{X_1, \dots, X_n\}$,

$$\lim_n \frac{1}{a_n}(M_n - b_n) = G, \text{ en distribució,}$$

on G és una variable aleatòria de Gumbel amb funció de distribució

$$\Lambda(x) = \exp\{-e^{-x}\}, \; x \in \mathbb{R}.$$

Denotem per $\Phi(x)$ la funció de distribució d'una variable aleatòria normal estàndard i per $\phi(x)$ la seva densitat. La convergència anterior és equivalent a que per a tot $x \in \mathbb{R}$,

$$\lim_n \Phi^n(a_n x + b_n) = \Lambda(x).$$

El resultat de Peter Hall

En un remarcable article, Peter Hall (1979) demostrava que prenent b_n^* tal que

$$\frac{1}{\sqrt{2\pi}} \frac{1}{b_n^*} e^{-(b_n^*)^2/2} = \frac{1}{n} \quad \text{i} \quad a_n^* = 1/b_n^*,$$

tenim que per $n \geq 2$,

$$\frac{C'}{\log n} < \sup_{x \in \mathbb{R}} |\Phi^n(a_n^* x + b_n^*) - \Lambda(x)| < \frac{C}{\log n},$$

amb $C = 3$, que pot ser rebaixada a 0.91 quan $n \geq 10^6$. A més (i més important) aquesta velocitat de convergència no pot ser millorada amb cap altra seuccessió de constants normalitzadores.

D'aquesta manera, Hall donava una quantificació exacta de l'afirmació de Fisher i Tippet (1928) que *Pel cas de la distribució normal, l'aproximació a la distribució límit és extremadament lenta.*

La nostra contribució

Vam demostrar que si prenem

$$b_n = \Phi^{-1}\left(1 - \frac{1}{n}\right) \quad \text{and} \quad a_n^\circ = \frac{b_n}{1 + b_n^2}, \tag{1}$$

tenim:

Teorema. Donat $n_0 \geq 5$, per a qualsevol $n \geq n_0$, tenim

$$\sup_{x\in\mathbb{R}} \left|\Phi^n(a_n^\circ x + b_n) - \Lambda(x)\right| < \frac{C(n_0)}{\log n},$$

amb

$$C(n_0) = \begin{cases} 1, & \text{quan} \quad n_0 \leq 15 \\ \left(\dfrac{2}{3b_{n_0}^2} + \dfrac{1}{\sqrt{e}n_0}\right)\log(n_0) < 1 & \text{quan} \quad n_0 \geq 16. \end{cases}$$

A més, $\lim_{n_0\to\infty} C(n_0) = 1/3$.

Resultats numèrics

n_0	16	30	50	10^2	10^4	10^6	10^{10}	10^{20}	10^{100}
$C(n_0)$	0.90	0.75	0.67	0.60	0.45	0.41	0.38	0.36	0.34

Recordem que Hall havia estimat $C = 3$, que pot ser rebaixada a 0.91 quan $n \geq 10^6$.

Punt de vista aplicat

Hall proposa

$$\beta_n^* = (2\log n)^{1/2} - \log(4\pi \log n)/\left(2(2\log n)^{1/2}\right) \quad \text{and} \quad \alpha_n^* = 1/\beta_n^*.$$

Nosaltres, utilitzant una generalització de la funció de Lambert i estimacions fines de la raó de Mills proposem

$$\beta_n = \left(\log\left(n^2/(2\pi)\right) - \log\log\left(n^2/(2\pi)\right) + \frac{\log\left(\log(n^2) + 1/2\right) - 2}{\log\left(n^2/(2\pi)\right)} \right)^{1/2} \tag{2}$$

Que compleix

$$b_n = \beta_n + O\left(\frac{(\log\log n)^2}{(\log n)^{5/2}}\right), \quad n \to \infty.$$

i

$$\alpha_n = \frac{\beta_n}{1 + \beta_n^2}.$$

A. GASULL, FU

Approximating Mills ratio

Journ. Mathem. Anal. Appl., **420** (2014) 1832–1853.

La raó de Mills (Mills ratio)

La raó de Mills és la funció

$$f(x) = (1 - \Phi(x))/\phi(x) = e^{x^2/2} \int_x^\infty e^{-t^2/2}\,dt, \; x \geq 0,$$

on $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$ és la densitat d'una llei normal estàndard i $\Phi(x) = \int_{-\infty}^{x} \phi(t)\,dt$ la seva funció de distribució. Aquesta funció és molt més antiga que Mills (1928), i a través de

$$F(x) = e^{x^2} \int_x^\infty e^{-t^2}\,dt$$

i la relació

$$f(x) = \sqrt{2}\,F(x/\sqrt{2}),$$

la seva introducció pot ser atribuida a Laplace (1805), *Traité de Mécanique Céleste, Tome IV*, on analitza diferents hipòtesis sobre la refracció de la llum en l'atmòsfera.Laplace va donar molts dels resultats essencials, com l'expressió en fracció contínua i l'expansió asimptòtica.

Una funció molt estudiada

A més, com la funció F està relacionada amb la funció d'error i la funció gamma incompleta de paràmetre $1/2$, molts resultats han estat descoberts i redescoberts diverses vegades en diferents àmbits.

Algunes propietats

- f és completament monòtona: és de classe $\mathcal{C}^\infty$ i per a tot $n \geq 0$,

$$(-1)^n f^{(n)}(x) \geq 0, \ x > 0.$$

- Fracció contínua (Laplace)

$$f(x) = \cfrac{1}{x + \cfrac{1}{x + \cfrac{2}{x + \cfrac{3}{x + \cdots}}}}.$$

- Desenvolupament asimptòtic (Laplace)

$$f(x) \sim \frac{1}{x} - \frac{1}{x^3} + \frac{1\cdot 3}{x^5} - \frac{1\cdot 3\cdot 5}{x^7} + \cdots, \; x \to \infty.$$

Els dos primers termes són claus en l'article de Peter Hall.

- La sèrie de Taylor a 0 és convergent en $[0, \infty)$

$$f(x) = \frac{\sqrt{2\pi}}{2} - x + \frac{\sqrt{2\pi}}{4} x^2 - \frac{1}{3} x^3 + \frac{\sqrt{2\pi}}{16} x^4 - \frac{1}{15} x^5 + \cdots,$$

on el coeficient de x^n is

$$\begin{cases} \dfrac{\sqrt{2\pi}}{2\, n!!}, & \text{si } n \text{ és parell}, \\ -\dfrac{1}{n!!}, & \text{si } n \text{ és senar}, \end{cases}$$

Aproximació per funcions racionals

A partit de $f'(x) = xf(x) - 1$, la derivada n-èssima de f compleix

$$f^{(n)}(x) = P_n(x)f(x) - Q_n(x),$$

per certs polinomis P_n and Q_n, amb coeficients no negatius, d'ordres respectivament n i $n-1$. Tenim

Teorema

$$\frac{f^{(n)}(x)}{P_n(x)} = (-1)^n \left(f(x) - \frac{Q_n(x)}{P_n(x)} \right) < \frac{n!}{x^{2n+1}}.$$

A més, per a $x > 0$, quan $m \to \infty$,

$$\frac{Q_{2m}(x)}{P_{2m}(x)} \nearrow f(x) \quad \text{i} \quad \frac{Q_{2m+1}(x)}{P_{2m+1}(x)} \searrow f(x).$$

Recurrències per als polinomis P_n i Q_n

La demostració del teorema anterior es basa en les recurrències

$$P_n(x) = xP_{n-1}(x) + (n-1)P_{n-2}(x), \; n \geq 2,$$

amb condicions inicials $P_0(x) = 1$ i $P_1(x) = x$.

$$Q_n(x) = xQ_{n-1}(x) + (n-1)Q_{n-2}(x), \; n \geq 2,$$

amb condicions inicials $Q_0(x) = 0$ and $Q_1(x) = 1$. En particular, ambdós P_n i Q_n són mònics i amb coeficients no negatius.
Els primers polinomis són

$$P_1(x) = x, \; P_2(x) = x^2+1, \; P_3(x) = x^3+3x, \; P_4(x) = x^4+6x^2+3.$$

i

$$Q_1(x) = 1, \; Q_2(x) = x, \; Q_3(x) = x^2+2, \; Q_4(x) = x^3+5x.$$

Una propietat essencial

La demostració d'aquesta propietat (que és clau) es basa en que la integral $\int e^{-x^2/2}\, dx$ no es pot expressar en termes de funcions elementals, que es conseqüència d'un resultat de Liouville (1835), i que part d'un resultat molt més general.

Una aplicació: la funció $1/f$ és estrictament convexa

Del fet que

$$\left(\frac{1}{f(x)}\right)'' = \frac{2 + x^2 f^2(x) - f^2(x) - 3xf(x)}{f^3(x)} := \frac{g(x)}{f^3(x)}.$$

es suficient provar que

$$g(x) := 2 + x^2 f^2(x) - f^2(x) - 3xf(x) > 0.$$

Això es fa en dues passes: per a $x \in [0, 1]$ a partir de la sèrie de Taylor de f, i per a $x > 1$ utilitzant funcions racionals adients Q_n/P_n.

Estudi del segon cas (quan $x > 1$)

Utilitzem que

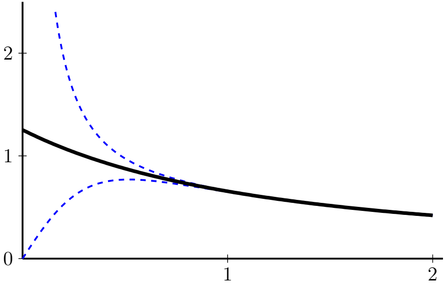
$$Q_{10}(x)/P_{10}(x) < f(x) < Q_{11}(x)/P_{11}(x),$$

on

$$\frac{Q_{10}(x)}{P_{10}(x)} = \frac{x^9 + 44x^7 + 588x^5 + 2640x^3 + 2895x}{x^{10} + 45x^8 + 630x^6 + 3150x^4 + 4725x^2 + 945},$$

$$\frac{Q_{11}(x)}{P_{11}(x)} = \frac{x^{10} + 54x^8 + 938x^6 + 6090x^4 + 12645x^2 + 3840}{x^{11} + 55x^9 + 990x^7 + 6930x^5 + 17325x^3 + 10395x},$$

Aproximació de la raó de Mills per funcions racionals



Línia contínua: Raó de Mills f.
Línia superior: funció racional Q_{11}/P_{11}
línia inferior: Q_{10}/P_{10}.

Afitació inferior de g

A l'expressió de $g(x) = 2 + x^2f^2(x) - f^2(x) - 3xf(x)$ canviem f per Q_{10}/P_{10} en els termes positius i per Q_{11}/P_{11} en els negatius. Obtenim una funció racional amb denominador positiu i numerador

$$\begin{aligned}N(x) =& x^{36} + 185x^{34} + 15388x^{32} + 761580x^{30} + 25019940x^{28} \\ & + 576522420x^{26} + 9601604100x^{24} + 117398708820x^{22} \\ & + 1059855272550x^{20} + 7043405005350x^{18} \\ & + 33995881448100x^{16} + 115607852356500x^{14} \\ & + 259703297525700x^{12} + 329529066520500x^{10} \\ & + 108511796893500x^{8} - 233411033740500x^{6} \\ & - 247669566519375x^{4} - 66176702274375x^{2} \\ & - 6584094720000.\end{aligned}$$

Darrer pas

Per tant n'hi ha prou demostrant que

$$N(x) > 0, \text{per a } x > 1.$$

Això es dedueix del fet que $N(1) > 0$ i que N no té arrels reals a $(1, \infty)$, ja que la seva arrel més gran està a l'interval $(93/100, 94/100)$.
Aquest darrer resultat es prova utilitzant el teorema de Sturm.
En el cas d'un polinomi, és senzill construir la seqüència de Sturn. Encara més, si el polinomi és a coeficients racionals, les condicions del teorema es poden comprovar analíticament.