Gludi M-novērtējumi robustajā statistikā

Māra Vēliņa

Latvijas Universitāte

27.10.2011

Saturs

Ievads

Mērķis: novērtēt lokācijas parametru ar robustas procedūras palīdzību

- Hubera M-novērtējums (1964) - robusts lokācijas parametra novērtējums
- Owen (1988) - empīriskās ticamības metode, pielietojama arī M-novertējumiem
- Hampel (2011) - gludais Hubera M-novērtējums

Procesā

Divu izlašu problemātika: empīriskas ticamības metode divu izlašu gludo Hubera novertējumu starpības novertēšanai
(Valeinis, Velina, Luta: abstract for ICORS 2011 conference)

M-novertējums

Pieņemsim, ka $X_1, X_2, \ldots, X_n \sim iid$, $X_1 \sim F$. M-novertējumu T_n definē konkretai funkcijai ρ, kur T_n apmierina vienādojumu

$$\sum_{i=1}^{n} \rho(X_i, t) = \int \rho(x, t) dF_n(x), \tag{1}$$

kur F_n - empīriskā sadalījuma funkcija.
Ja ρ ir diferencējama pēc t, tad (1) sasniedz minimumu pie sekojoša vienādojuma atrisinājuma:

$$\sum_{i=1}^{n} \psi(X_i, t) = 0,$$

kur $\psi(x, t) = \frac{\partial}{\partial t} \rho(x, t)$.

Piemēri

- Vidējā vērtība. $\psi(\mathrm{x}, \mathrm{t}) = \mathrm{x} - \mathrm{t}$, $\mathrm{T_n} = \bar{\mathrm{X}}$.
- Vislielakas ticamības (ML) novērtējums. $\psi(\mathrm{x}, \theta) = -\frac{\mathrm{d}}{\mathrm{d}\theta} \log \mathrm{f}(\mathrm{x}, \theta)$ varbūtību blīvuma funkciju klasei $\mathrm{f}(\mathrm{x}, \theta)$, dod $\mathrm{T_n}$ kā atrisinājumu no ticamības vienādojuma

$$\frac{\mathrm{d}}{\mathrm{d}\theta} \log \left(\prod_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{f}(\mathrm{X_i}, \theta) \right) = 0.$$

- Mediāna. $\psi(\mathrm{x}, \mathrm{t}) = \psi_0(\mathrm{x} - \mathrm{t})$, $\psi_0(\mathrm{z}) = \mathrm{k}\,\mathrm{sgn}(\mathrm{z})$, $\mathrm{k} > 0$.

Hubera novērtējums lokācijas parametram μ

Hubers (1964) apvienoja vienā novērtējumā labakās vidējas vērtības un mediānas īpašības.

Pieņemsim, ka F ir simetriska varbūtību blīvuma funkcija

$$\mathrm{f}_{\mu,\sigma}(\mathrm{x}) = \frac{1}{\sigma}\mathrm{f}\left(\frac{\mathrm{x}-\mu}{\sigma}\right),$$

pieņemsim, ka $\sigma = 1$. Lokācijas parametra μ M-novērtējumu definē kā sakni no vienādojuma

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \psi\left(\frac{\mathrm{X_i}-\mathrm{t}}{\sigma}\right) = 0. \tag{2}$$

Hubera M-novērtējumu definē sekojoša funkcija ψ no (2):

$$\psi_{\mathrm{k}}(\mathrm{x}) = \begin{cases} \mathrm{k}, & \mathrm{x} \geq \mathrm{k} \\ \mathrm{x}, & -\mathrm{k} \leq \mathrm{x} \leq \mathrm{k} \\ -\mathrm{k}, & \mathrm{x} < -\mathrm{k}. \end{cases} \tag{3}$$

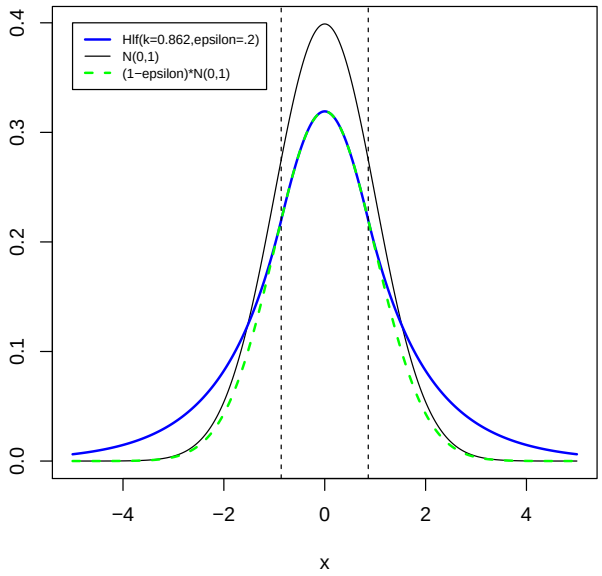
Hubera vismazāklabvēlīgais sadalījums

Hubera M-novērtējums ir lokācijas parametra vislielākās ticamības novērtējums sadalījumam Hubera vismazāklabvēlīgā sadalījumam (Hlf). Hlf bīvuma funkcija $\mathrm{f_k}$ ir

$$\mathrm{f_k(x)} = \begin{cases} (1-\epsilon)\phi(\mathrm{k})\exp(-\mathrm{k}(\mathrm{x}-\mathrm{k})), & \mathrm{x} > \mathrm{k} \\ (1-\epsilon)\phi(\mathrm{x}), & -\mathrm{k} \le \mathrm{x} \le \mathrm{k} \\ (1-\epsilon)\phi(\mathrm{k})\exp(\mathrm{k}(\mathrm{x}+\mathrm{k})), & \mathrm{x} \le -\mathrm{k}, \end{cases} \tag{4}$$

kur k un ϵ saista izteiksme $2\phi(\mathrm{k})/\mathrm{k} - 2\Phi(-\mathrm{k}) = \epsilon/(1-\epsilon)$ un ϕ, Φ ir standartnormālā blīvuma un sadalījuma funkcijas.

Hubera vismazāklabvelīgais sadalījums



(a)

(a) Hubera vismazāklabvelīgais sadalījums salīdzinājumā ar standartnormālo sadalījumu

Hubera motivācija:

- Neierobežotām ψ-funkcijām ir nevēlamas īpašības (nestabilas pret izlēcējiem);
- Aplūko k robežvērtības un ψ_{k} atbilstošos M-novērtējumus:
  - Ja $\mathrm{k} \to \infty$, tad ψ_{k} ir $\bar{\mathrm{X}}$;
  - Ja $\mathrm{k} \to 0$, tad ψ_{k} ir mediāna.
- k darbojas kā saskaņošanas konstante, kas nosaka robustuma pakāpi.
- Hubera novērtējumam ir minimax asimptotiskā dispersija sadalījuma funkciju klasē

$$(1-\epsilon)\phi(\mathrm{x}) + \epsilon \mathrm{h}(\mathrm{x}),$$

kur ϕ ir $\mathrm{N}(0,1)$ blīvuma funkcija, h - kāda simetriska sadalījuma blīvuma funkcija.

Skalēts lokācijas parametra novērtējums

Realitātē, $\sigma \neq 1$, un nav zināma, tāpēc σ jānovērtē ar kādu robustu procedūru. Bieži izmanto MAD (median absolute deviation) novērtējumu.

MAD

$$\mathrm{S_n} = \mathrm{MAD} = \mathrm{median}(|\mathrm{X_i} - \mathrm{median}(\mathrm{X_i})|).$$

MAD ir robusts novērtējums, darbojas apmierinoši pat datos ar izlēcēju piejaukumu (līdz pat 50% no datiem).

Gludais M-novērtējums (Hampel, 1996)

Visparīgai M-novertējumu funkciju klasei ψ definē

$$\tilde{\psi}(\mathrm{x}) = \int \psi(\mathrm{x}+\mathrm{u})\mathrm{dQ_n(u)}, \tag{5}$$

kur

- kur $\mathrm{Q_n}$ var izvēleties kā sākotnejā M-novertējuma sadalījumu n iid novērojumiem pie sākotnēji fiksēta sadalījuma.
- Gludināšanas pakāpe atkarīga no izlases apjoma n.
- $\mathrm{Q_n}$ var novērtēt ar $\mathrm{N}(0, \mathrm{V}/\mathrm{n})$, kur V ir M-novērtējuma asimptotiskā dispersija.
- Tātad: jāfiksē sadalījums, pie kura tiks aprēķināta asimptotiskā dispersija.
- Gludināšanu var pielietot jau gludiem M-novērtējumiem.

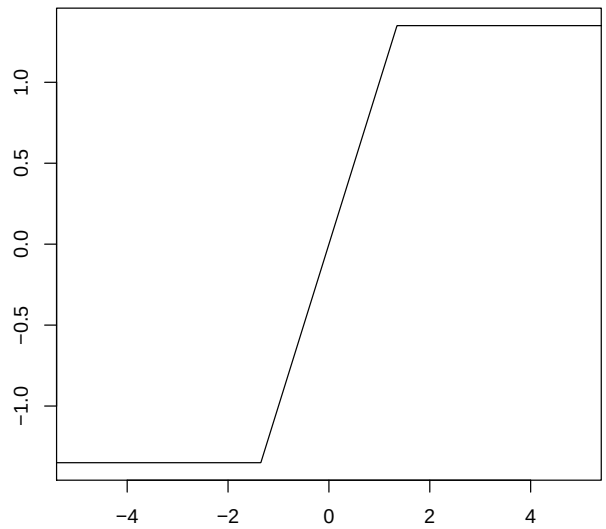
Gludais Hubera M-novērtējums

Gludā Hubera M-novērtējuma ψ-funkciju var izteikt analītiski (parasti nevar!) kā

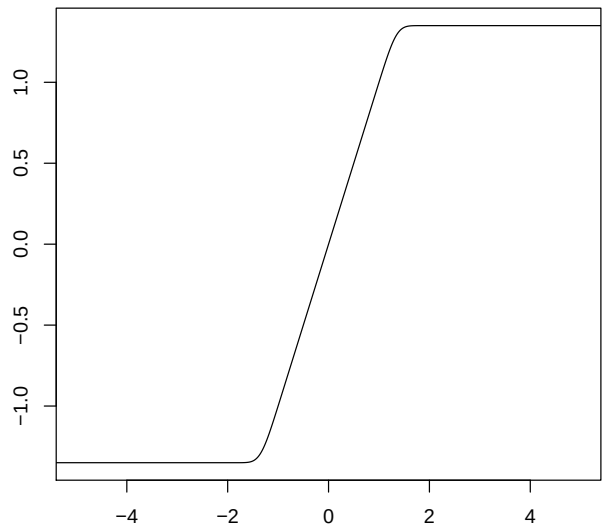
$$\begin{aligned}\tilde{\psi}_{\mathrm{k}}(\mathrm{x}) \quad = \quad & \mathrm{k}\Phi\left(\frac{\mathrm{x}-\mathrm{k}}{\sigma_{\mathrm{n}}}\right) - \mathrm{k}\left(1-\Phi\left(\frac{\mathrm{x}+\mathrm{k}}{\sigma_{\mathrm{n}}}\right)\right) \\ & +\mathrm{x}\,\Phi\left(\frac{\mathrm{x}+\mathrm{k}}{\sigma_{\mathrm{n}}}\right) - \Phi\left(\frac{\mathrm{x}-\mathrm{k}}{\sigma_{\mathrm{n}}}\right)\Bigg) \\ & +\sigma_{\mathrm{n}}\left(\phi\left(\frac{\mathrm{x}+\mathrm{k}}{\sigma_{\mathrm{n}}}\right) - \phi\left(\frac{\mathrm{x}-\mathrm{k}}{\sigma_{\mathrm{n}}}\right)\right), \qquad (6)\end{aligned}$$

kur $\sigma_{\mathrm{n}} = \sqrt{\mathrm{V}/\mathrm{n}}$, un Φ un ϕ ir $\mathrm{N}(0,1)$ blīvuma un sadalījuma funkcijas.

Piemērs



(b)



(c)

(a) Hubera novērtējuma ψ-funkcija;
(b) $\tilde{\psi}$ Gludā Hubera novērtējuma ψ-funkcija; k=1.35.

Empīriskās ticamības metode M-novērtējumiem

- Owen (1988) showed that EL method can be applied to certain M-estimators, including Huber estimator.
- Nonparametric Wilk's theorem applies thus EL based confidence intervals for Huber estimate can be obtained.
- Tsao, Zhu (2001) showed that EL based confidence intervals preserves robustness.

EL confidence bands for Huber estimator

Empirical likelihood ratio for parameter t

$$\mathrm{R(t)} = \sup\{\prod_{\mathrm{i=1}}^{\mathrm{n}} \omega_{\mathrm{i}} \sum_{\mathrm{i=1}}^{\mathrm{n}} \omega_{\mathrm{i}} \psi(\mathrm{X_i, t}) = 0, \omega_{\mathrm{i}} \geq 0, \sum_{\mathrm{i=1}}^{\mathrm{n}} \omega_{\mathrm{i}} = 1\}$$

is maximized by $\prod \omega_{\mathrm{i}}(\lambda)$, where

$$\omega_{\mathrm{i}}(\lambda) = \{\mathrm{n}(1 + \lambda \mathrm{Z_i})\}^{-1},$$

and $\mathrm{Z_i} = \psi(\mathrm{X_i, t})$ and λ follows from

$$\mathrm{n}^{-1} \sum \mathrm{Z_i}/(1 + \lambda \mathrm{Z_i}) = 0.$$

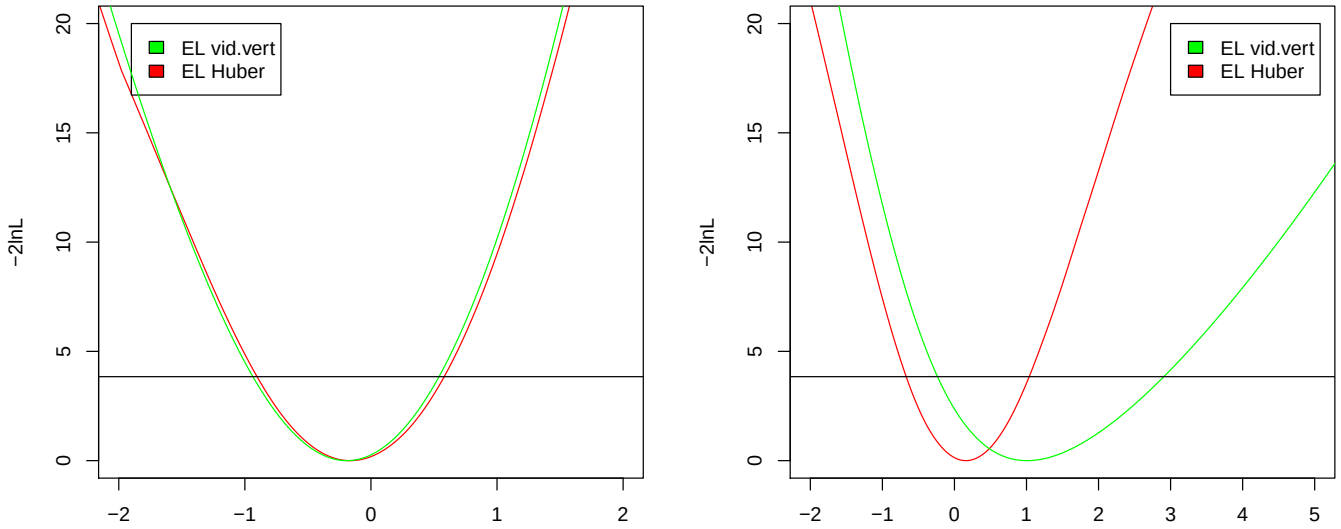


Figure: EL -2*ln,(a) $\mathrm{N}(0,3)$ (b) $0.95 * \mathrm{N}(0,3) + 0.05 * \mathrm{N}(20,3)$

Simulation results for one sample problem

Table: Huber estimation for location parameter and its EL confidence bands, alpha=0.05

	$N(0,3)$				$0.95 * N(0,3) + 0.05 * N(20,3)$			
sample		len		estimate		len		estimate
n=50	EL.huber	0.494	EL.huber	-0.055	EL.huber	1.706	EL.huber	0.159
	EL.mean	0.492	EL.mean	-0.064	EL.mean	3.14	EL.mean	1.008
	t-test	0.506	mean	-0.064	t-test	3.117	mean	1.008
	z-test	0.554	huber	-0.076	z-test	0.554	huber	0.159
	Bootstrap	0.497			Bootstrap	3.057		
n=20	EL.huber	0.667	EL.huber	-0.167	EL.huber	2.478	EL.huber	-0.441
	EL.mean	0.667	EL.mean	-0.167	EL.mean	4.894	EL.mean	0.498
	t-test	0.732	mean	-0.167	t-test	4.938	mean	0.498
	z-test	0.877	huber	-0.643	z-test	0.877	huber	-0.441
	Bootstrap	0.699			Bootstrap	4.583		
n=10	EL.huber	1.001	EL.huber	-0.067	EL.huber	4.303	EL.huber	-0.189
	EL.mean	1.001	EL.mean	-0.067	EL.mean	9.68	EL.mean	1.008
	t-test	1.239	mean	-0.067	t-test	11.494	mean	1.799
	z-test	1.24	huber	-0.201	z-test	1.24	huber	-0.189
	Bootstrap	1.039			Bootstrap	9.74		

Two sample EL problem

Consider empirical likelihood-based method for the difference of smoothed Huber estimators.
Given two independent samples X and Y with distribution functions F_1 and F_2, respectively, we have two unbiased estimating functions:

$$\mathrm{E}_{\mathrm{F}_1} \mathrm{w}_1(\mathrm{X}, \theta_0, \Delta) = 0, \quad \mathrm{E}_{\mathrm{F}_2} \mathrm{w}_2(\mathrm{Y}, \theta_0, \Delta) = 0,$$

where Δ is the parameter of interest and θ_0 is a nuisance parameter. Specifically, $\Delta = \theta_1 - \theta_0$ and

$$\mathrm{w}_1(\mathrm{X}, \theta_0, \Delta) = \tilde{\psi}\left(\frac{\mathrm{X} - \theta_0}{\hat{\sigma}_1}\right) \quad \mathrm{w}_2(\mathrm{Y}, \theta_0, \Delta) = \tilde{\psi}\left(\frac{\mathrm{Y} - \Delta + \theta_0}{\hat{\sigma}_2}\right),$$

where $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are scale estimators, and $\tilde{\psi}$ corresponds to the smoothed Huber estimator.

Novērtēšanas problēmas reāliem datiem

1. Kā novērtēt asimptotisko dispersiju V gludajā Hubera novērtējumā?
   - Fiksē sākotnējo sadalījumu, novērtē analītiski.
     Pieņemsim, ka ψ ir augoša. Dotam sadalījumam F, definē $\mu = \mu_0(\mathrm{F})$ kā atrisinājumu no $\mathrm{E_F}\psi(\mathrm{x} - \mu_0) = 0$. Kad $\mathrm{n} \to \infty$, $\hat{\mu} \to_{\mathrm{p}} \mu_0$. $\hat{\mu}$ sadalījums ir aptuveni
     $$\mathrm{N}(\mu_0, \nu/\mathrm{n}), \nu = \frac{\mathrm{E_F}(\psi(\mathrm{x} - \mu_0)^2)}{(\mathrm{E_F}(\psi'(\mathrm{x} - \mu_0))^2}.$$
     Rets gadījums - Hubera novērtējumam ν var atrast analītiski:
     $$\nu = \frac{(\Phi(\mathrm{x}) - \Phi(-\mathrm{k}))^2}{2[\mathrm{k}^2(1 - \Phi(\mathrm{k})) + \Phi(\mathrm{k}) - .5 - \mathrm{k}\phi(\mathrm{k})]}.$$
   - Literatūrā minēts $\mathrm{V} = \sqrt{\mathrm{n}}$, kas atbilst Hubera novērtējuma asimptotiskai dispersijai jauktam normālam sadalījumam ar $\epsilon = 0.2$ piesārņojumu unkonstanti $\mathrm{k} = 0.862$.
   - Aprēķināt ar simulāciju palīdzību

2 pakāpju novērtēšanas procedūra

1. Novērtēt Hubera novērtējuma (negludinātā) dispersiju:
   - Reāliem datiem: ar butstrapa palīdzību novērtēt negludā Hubera novērtējuma dispersiju;
   - Simulētiem datiem: simulēt N izlases, novērtēt $\mathrm{V_n}$.
2. Aprēķināt gludināto Hubera novērtējumu, lietojot novērtēto $\mathrm{V_n}$.

1. Piezīme. Lai novērtētu negludināto Hubera novērtējumu, jāizlemj, kā novērtēt skalēšanas parametru: MAD? SD? Const=1?
2. Piezīme: Hipotēze: Ir nozīme, vai $\mathrm{V_n}$ novērtē, vai lieto konstanti $\mathrm{V_n} = \sqrt{2.046/\mathrm{n}}$.

Analizējot reālo datu piemērus, secināts, ka vislabāko rezultātu sasniedz, skalēšanas parametru novērtējot ar MAD.

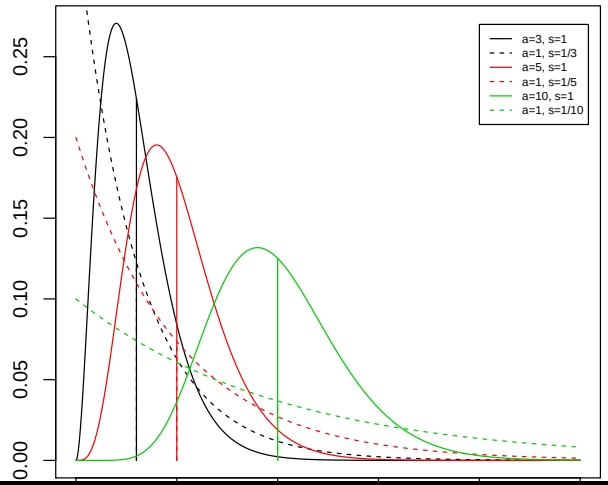
Table: Lokācijas parametra μ gludais Hubera novērtējums un p-vērtība. Hubera novērtējumam skalēšanas parametrs s , gludajam novērtējumam dispersijas V novērtējums MAD, alpha=0.05

k=0.862	t-test									
	var.equal=F	var.equal=T			s=1		s=MAD		s=sd	
Data	pval	pval	pval	mu	pval	mu	pval	mu	pval	mu
IQ dataset	0.122	0.016	0.052	11.719	0.054	12.794	0.047	11.503	0.044	11.930
Palestinians	0.005	0.002	0.002	-8.317	0.009	-8.391	0.002	-8.412	0.002	-8.415
data9-10	0.112	0.106	0.019	-6.801	0.228	-2.546	0.093	-3.420	0.041	-5.087
data10-11	0.632	0.626	0.620	0.705	0.568	0.800	0.537	0.871	0.596	0.753
data9-11	0.147	0.097	0.033	-6.097	0.332	-1.747	0.164	-2.550	0.064	-4.331
Marazzi1	0.192	0.000	0.023	-17.595	0.717	-0.209	0.665	-0.239	0.035	-6.302
Marazzi1*	0.953	0.933	0.951	0.912	0.987	0.010	0.885	0.075	0.494	0.720
Marazzi2	0.081	0.080	0.021	3.021	0.052	1.058	0.076	1.019	0.014	1.655
Marazzi3	0.886	0.886	0.886	-0.256	0.880	-0.144	0.118	-1.360	0.220	-1.220

Divu izlašu problēma. 1. piemērs.

Mērķis: simulēt izlašu pārus ar vienādām vidējām vērtībām μ_1, μ_2, aprēķināt $\mu_1 - \mu_2 = 0$ TI pārklājuma precizitāti.
Divi $\mathrm{Gamma(a, s)}$ sadalījumi. Izlases apjoms $\mathrm{n} = 50$, atkārtojumi $\mathrm{N} = 1000$.

- $\mathrm{F_1} = \mathrm{Gamma(a} = \sigma, \mathrm{s} = 1)$
- $\mathrm{F_2} = \mathrm{Gamma(a} = 1, \mathrm{s} = 1/\sigma)$

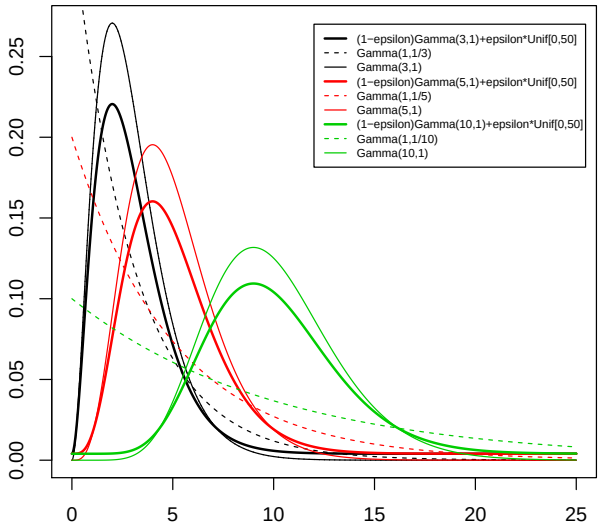


σ	huber.var1	huber.var2	t.test	EL	Huber1	Huber2
3	3.278517	7.266661	0.94	0.938	0.883	0.934
4	4.081278	12.67512	0.933	0.932	0.878	0.928
5	5.751523	20.15433	0.924	0.919	0.873	0.918
6	6.81938	28.87526	0.943	0.926	0.872	0.926
7	8.214858	39.22282	0.926	0.919	0.855	0.919
10	11.82479	81.09296	0.925	0.917	0.861	0.917
20	24.6246	318.3097	0.935	0.932	0.86	0.932

Divu izlašu problēma. 2. piemērs.

Divi $\mathrm{Gamma(a, s)}$ sadalījumi ar piesārņojumu, $\mathrm{n} = 50$, $\mathrm{N} = 1000$.

- $\mathrm{F_1} = (1 - \epsilon) * \mathrm{Gamma(a} = \sigma, \mathrm{s} = 1) + \mathrm{Unif}[0, 50]$
- $\mathrm{F_2} = \mathrm{Gamma(a} = 1, \mathrm{s} = 1/\sigma)$



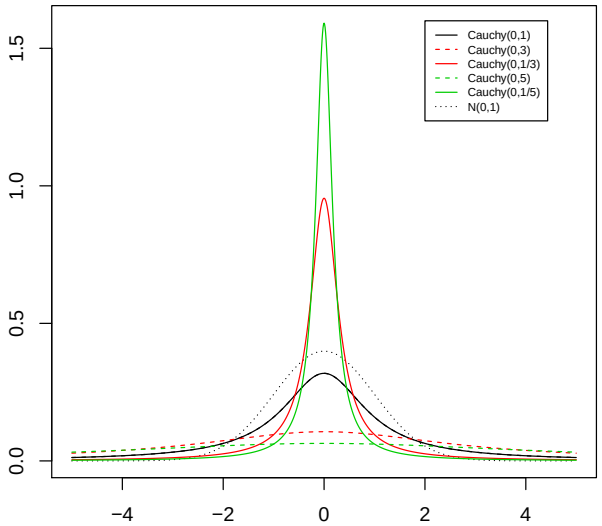
σ	huber.var1	huber.var2	t.test	EL	Huber1	Huber2
3	8.7	8.0	0.130	0.044	0.349	0.090
4	10.2	13.5	0.205	0.116	0.407	0.173
5	13.4	21.3	0.268	0.194	0.454	0.229
6	16.0	31.2	0.367	0.296	0.512	0.321
7	16.8	42.0	0.433	0.402	0.525	0.416
10	22.0	82.6	0.664	0.664	0.666	0.669
20	36.8	344.3	0.889	0.909	0.806	0.909
25	42.4	583.8	0.945	0.950	0.886	0.950
50	86.8	2133.5	0.924	0.869	0.906	0.869

Gamma(a,s) blīvuma funkciju grafiki; Simulācijas

Divu izlašu problēma. 3. piemērs.

Divi $\mathrm{Cauchy}(0, \mathrm{s})$ sadalījumi, $\mathrm{n} = 50$, $\mathrm{N} = 1000$.

- $\mathrm{F}_1 = \mathrm{Cauchy}(0, \mathrm{s} = \sigma)$
- $\mathrm{F}_2 = \mathrm{Cauchy}(0, \mathrm{s} = 1/\sigma)$



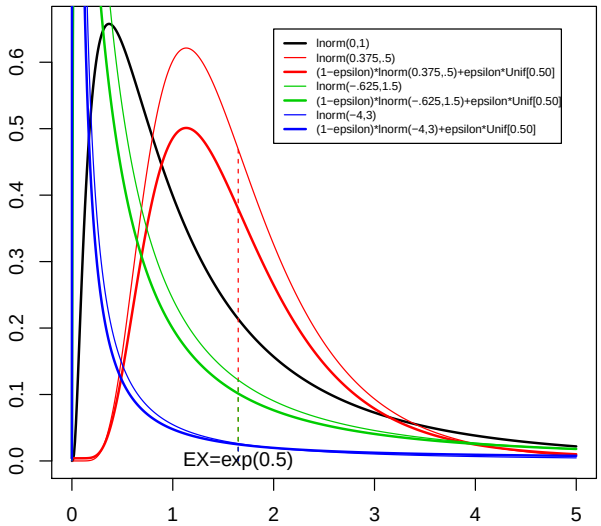
σ	huber.var1	huber.var2	t.test	EL	Huber1	Huber2
3	25.5	0.311	0.971	0.764	0.654	0.625
5	75.4	0.117	0.981	0.731	0.618	0.498
7	155.7	0.063	0.974	0.758	0.668	0.588
10	278.4	0.031	0.980	0.792	0.732	0.687
20	1146.9	0.008	0.975	0.931	0.918	0.911
25	1790.2	0.004	0.979	0.946	0.938	0.938
50	7324.7	0.001	0.974	0.993	0.992	0.992
100	28927.5	0.000	0.986	0.998	0.998	0.998

Cauchy(0,s) blīvuma funkciju grafiki; Simulācijas

Divu izlašu problēma. 4. piemērs.

Divi lognormālie $\mathrm{lnorm}(\mathrm{meanolg}, \mathrm{sdlog})$ sadalījumi, $\mathrm{n} = 50$, $\mathrm{N} = 1000$.

- $\mathrm{F}_1 = \mathrm{lnorm}(\mu, \sigma)$ $\mathrm{F}_2 = \mathrm{lnorm}(0, 1)$



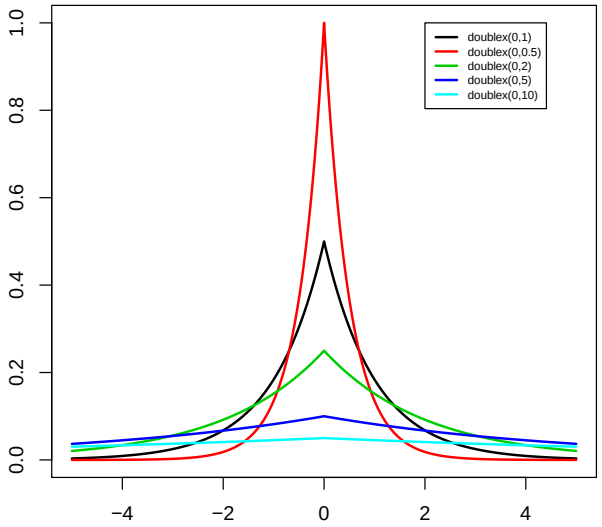
μ	σ	huber.var1	huber.var2	t.test	EL	Huber1	Huber2
-0.051	1.05	1.55	1.48	0.966	0.917	0.881	0.865
-0.625	1.50	1.18	1.57	0.915	0.880	0.503	0.370
-1.500	2.00	0.47	1.48	0.796	0.774	0.018	0.001
-2.625	2.50	0.09	1.48	0.635	0.633	0.010	0.010
-4.000	3.00	0.01	1.48	0.469	0.477	0.016	0.016
-5.625	3.50	0.00	1.48	0.302	0.319	0.016	0.016

Lognormāla sadalījuma blīvuma funkciju grafiki; Simulācijas

Divu izlašu problēma. 5. piemērs.

Divi dubulteksponenciālie sadalījumi $\mathrm{doublex}(0, \sigma)$ sadalījumi, $\mathrm{n} = 50$, $\mathrm{N} = 200$.

- $\mathrm{F}_1 = \mathrm{doublex}(0, \sigma)$
- $\mathrm{F}_2 = \mathrm{doublex}(0, 1)$

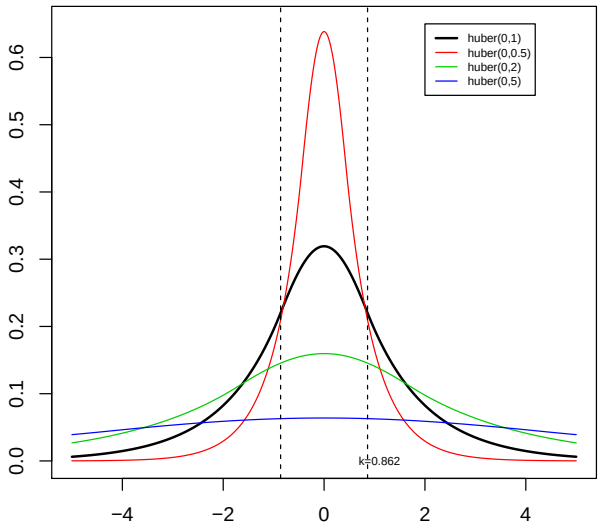


σ	huber.var1	huber.var2	t.test	EL	Huber1	Huber2
0.1	0.01	1.12	0.960	0.940	0.955	0.945
0.2	0.05	1.19	0.955	0.935	0.950	0.950
0.5	0.30	1.19	0.940	0.940	0.945	0.945
2	4.78	1.19	0.955	0.935	0.965	0.960
5	29.85	1.19	0.965	0.940	0.960	0.945
10	119.39	1.19	0.965	0.945	0.960	0.940

Divu izlašu problēma. 6. piemērs.

Divi Hubera vismazāklabvēlīgie sadalījumi $\mathrm{hlf}(0,\sigma)$, $\mathrm{n} = 50$, $\mathrm{N} = 1000$.

- $\mathrm{F}_1 = \mathrm{hlf}(0,\sigma)$
- $\mathrm{F}_2 = \mathrm{hlf}(0,1)$

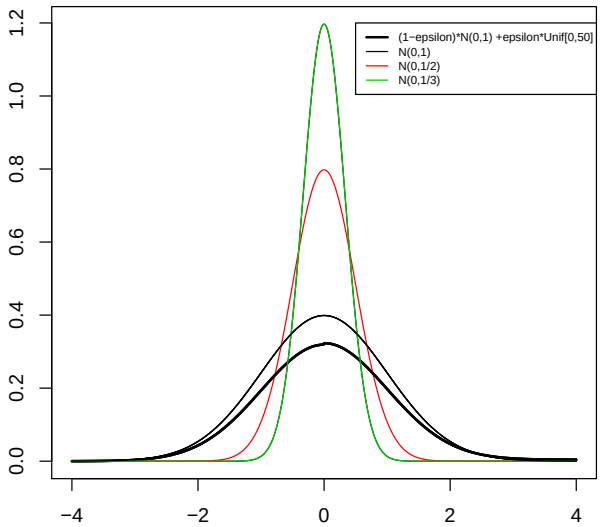


σ	huber.var1	huber.var2	t.test	EL	Huber1	Huber2
0.2	0.08	2.15	0.95	0.936	0.938	0.941
0.5	0.49	2.25	0.951	0.945	0.946	0.944
1	1.95	2.25	0.949	0.943	0.941	0.942
2	7.81	2.25	0.955	0.944	0.946	0.944
5	48.83	2.25	0.954	0.947	0.950	0.945
10	195.32	2.25	0.964	0.947	0.954	0.948

Divu izlašu problēma. 7. piemērs.

Divi Normālie sadalījumi ar piesārņojumu, $\mathrm{n} = 50$, $\mathrm{N} = 1000$.

- $\mathrm{F}_1 = (1 - \epsilon) * \mathrm{N}(0, 1) + \epsilon * \mathrm{Unif}[0, 50]$
- $\mathrm{F}_2 = \mathrm{N}(0, 1/\sigma)$



σ	huber.var2	huber.var1	t.test	EL	Huber1	Huber2
1	1.142	2.786	0.047	0.002	0.211	0.137
2	0.275	2.862	0.041	0.001	0.149	0.084
5	0.044	2.862	0.041	0.647	0.686	0.663
10	0.011	2.862	0.041	0.997	0.997	0.997
20	0.003	2.862	0.040	1.000	1.000	1.000
100	0.000	2.862	0.040	1.000	1.000	1.000

Normālā sadalījuma blīvuma funkciju grafiki; Simulācijas

Simulation results for two sample problem

Conisder two models:

- $Y_1 \sim (1-\epsilon)\text{Gamma}(\alpha = 5; \sigma = 1) + \epsilon\text{Uniform}[0; 50]$
- $Y_2 \sim \text{Gamma}(\alpha = 1; \sigma = 5)$

Table: Coverage accuray and average confidence interval lengths based on 1000 replicates, $n_1 = n_2 = 50$

	t.int		EL.hub1		EL.hub2		Boot1		Boot2	
	acc	ave	acc	len	acc	len	acc	len	acc	len
$\sigma = 5$	0.62	3.05	0.66	2.99	0.56	2.83	0.36	2.98	0.36	2.98
$\sigma = 6$	0.69	3.56	0.73	3.51	0.65	3.34	0.38	3.46	0.38	3.47
$\sigma = 7$	0.74	4.09	0.77	4.04	0.72	3.85	0.44	3.97	0.45	3.99
$\sigma = 8$	0.78	4.62	0.81	4.56	0.76	4.39	0.48	4.49	0.48	4.50
$\sigma = 9$	0.81	5.19	0.84	5.13	0.80	4.95	0.50	5.00	0.50	5.02

Thank you for your attention!