Distribution of empirical distribution generated by resample on Hilbert space R^S

— The influence of dimension number S —

Kiyotake KISHI

Let denote $\vartheta(\nu)$ be a density of an arbitrary measure ν, and $\{\vartheta(\mu_j)\}$ be the union of densities of measures $\{\mu_j\}$:(j=1,2,…, J) re-sampled from $\vartheta(\nu)$. Then metric $n(\vartheta(\mu_j))$ measured from $\vartheta(\nu)$ to $\vartheta(\mu_j)$:(j = 1,2,…, J) obeys approximately equal to Gaussian distribution in virtue of its theoretical distribution obeys χ(chi)-distribution[1].

Simulation: $\vartheta(\nu)$ designated as density of lognormal distribution with mean =0 and variance =1 and resample count J=30000.   **Table**; **S**: dimension number of R^S, **m**: mean metric of $n(\vartheta(\mu_j))$:(j =1,2,…, J), "6σBand-width": the length between m-3σ and m+3σ (i.e., 6σ) where σ^2 is a variance of $n(\vartheta(\mu_j))$:(j=1,2,…, J), "Hold-Rate": hold rate (or contain rate) of resample count within the range of $m\pm3\sigma$, $m\pm2\sigma$ and $m\pm\sigma$ (total count J=30000).

Table-1: sample size = 30,   (resample count **J** = 30000).

dimension :　**S**	501	250	125	62	31	16	8	4	3	2
mean-metric:　**m**	1.055	0.916	0.753	0.593	0.453	0.355	0.250	0.177	0.140	0.084
6σBand-width	0.168	0.240	0.312	0.341	0.352	0.372	0.355	0.383	0.391	0.458
m / 6σ	6.266	3.814	2.415	1.738	1.289	0.954	0.705	0.462	0.357	0.184
±3σHold-Rate (%)	99.73	99.74	99.74	99.70	99.57	99.64	99.15	99.89	99.58	100.0
±2σHold-Rate (%)	95.50	95.63	95.56	95.55	95.51	96.03	95.52	97.33	97.67	89.63
±σHold-Rate (%)	68.24	67.99	68.06	67.99	68.57	67.62	71.70	64.75	64.72	87.83
skewness	0.016	0.001	0.056	0.250	0.400	0.383	0.578	0.589	-0.057	1.412
kurtosis	-0.024	-0.047	-0.051	0.006	0.193	0.002	0.379	-0.364	0.209	1.241

Table-2: sample size = 200, (total integration of the density adjusted equal to 1, on the other previous paper[1] not).

dimension :　**S**	501	250	125	62	31	16	8	4	3	2
mean-metric:　**m**	0.640	0.496	0.378	0.284	0.205	0.147	0.094	0.059	0.046	0.028
6σBand-width(6σ)	0.134	0.139	0.141	0.145	0.151	0.166	0.183	0.167	0.149	0.130
m / 6σ	4.762	3.565	2.690	1.965	1.357	0.885	0.513	0.354	0.305	0.220
±3σHold-Rate (%)	99.73	99.69	99.71	99.73	99.65	99.62	99.73	99.15	99.28	98.92
±2σHold-Rate (%)	95.50	95.44	95.42	95.58	95.37	95.56	96.27	94.47	96.67	95.11
±σHold-Rate (%)	68.28	68.44	68.20	68.03	68.62	68.43	67.08	72.50	72.67	65.11
skewness	0.052	0.132	0.164	0.191	0.243	0.221	0.278	0.934	1.001	1.080
kurtosis	-0.047	0.066	0.060	0.025	0.082	0.117	-0.114	1.101	2.126	1.273

Table-3: sample size =1000, (total integration of the density adjusted equal to 1, on the other previous paper[1] not).

dimension :　**S**	501	250	125	62	31	16	8	4	3	2
mean-metric:　**m**	0.352	0.263	0.188	0.131	0.087	0.058	0.038	0.025	0.020	0.013
6σBand-width(6σ)	0.065	0.067	0.070	0.078	0.084	0.078	0.069	0.065	0.063	0.058
m / 6σ	5.441	3.942	2.682	1.682	1.040	0.745	0.549	0.391	0.317	0.220
±3σHold-Rate (%)	99.73	99.70	99.67	99.62	99.74	99.53	99.49	99.51	99.43	98.96
skewness	0.059	0.093	0.104	0.089	0.148	0.468	0.459	0.529	0.651	0.994
kurtosis	-0.039	0.018	0.044	0.115	-0.010	0.260	0.551	0.241	0.312	0.834

Reference

[1]Kiyotake,KISHI(2014): Both concentration and convergence of distribution of empirical distributions generated by resample on Hilbert space obeys sample size as small to large; Proceedings of 2014th Annual Meeting of Association of statistical Society of Japan, p.258. Tokyo, Japan.