

# 程序报告

本题研究非参数估计中 Parzen 窗的实现。其中  $p(x) \sim 0.2N(-1, 1) + 0.8N(1, 1)$ ，分别从中取  $N=5, 10, 50, 100, 500, 1000, 5000, 10000$  八个样本数的样本，使得  $p_n(x) \approx p(x)$ 。

$$P(x) = \begin{cases} \frac{1}{a} & -\frac{1}{2}a \leq x \leq \frac{1}{2}a \\ 0 & otherwise \end{cases}$$

(a) 研究方窗，用不同的窗宽  $a$  来估计  $p(x)$ ，根据课件，不妨取  $a=0.25, 1, 4$  三个值，对应  $N=5, 10, 50, 100, 500, 1000, 5000, 10000$  画出估计图像如下图 1, 2, 3 所示。可以看到在窗宽  $a$  相同的情况下，采样数  $N$  越多，估计效果越好；在样本数  $N$  相同的情况下，窗宽  $a=4$  的估计效果最好， $a=0.25$  时估计值为较多尖峰函数叠加，使估计值变化幅度很大，而  $a=4$  时估计值较为平缓，分辨率降低，随  $N$  变化估计效果变化不明显。因此，窗宽较大或者较小，估计效果都不好，要根据样本数做折中。

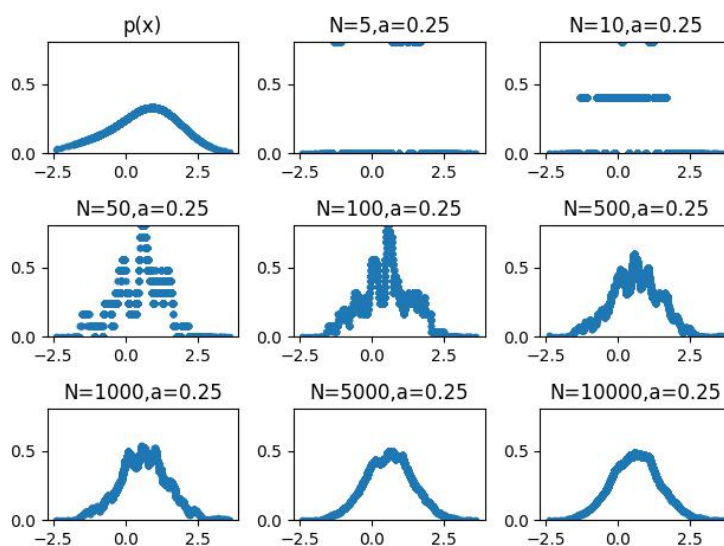


图 1

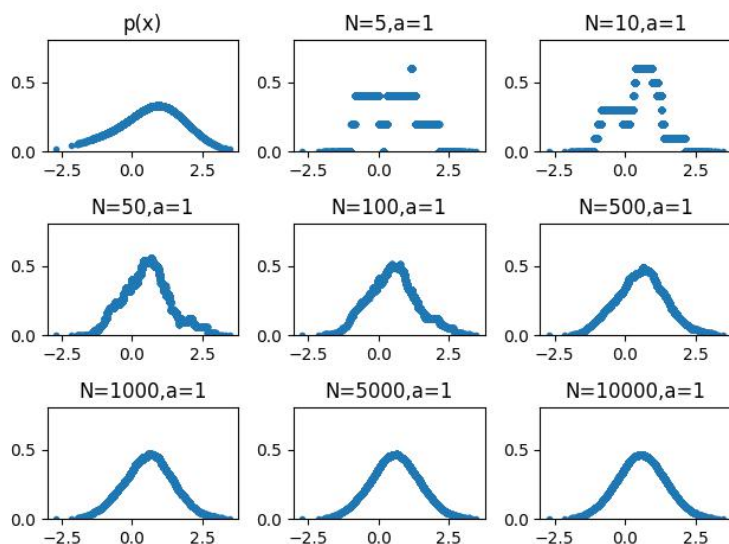


图 2

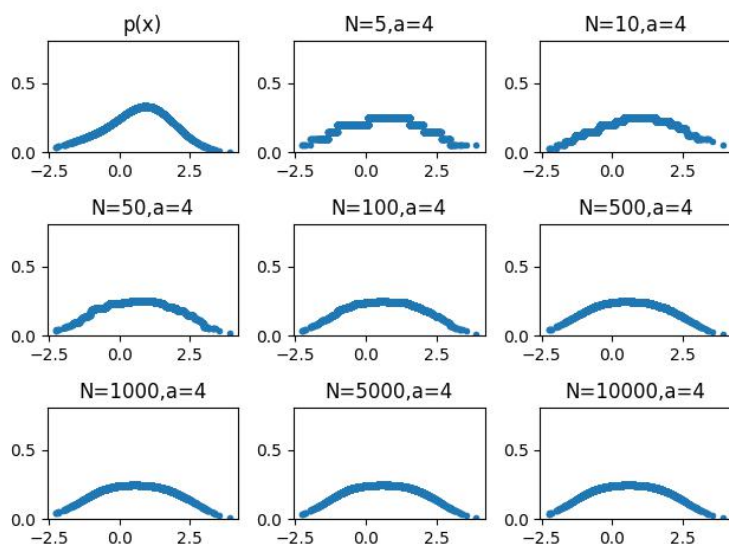


图 3

(b) 数值法计算  $E(p_n) = \int [p_n(x) - p(x)]^2 dx$ , 取足够小的  $dx$ , 即取足够多的  $N$ ,

用连加来近似  $E(p_n) = \sum [p_n(x) - p(x)]^2$ , 具体实现见代码。

(c) 求  $E(p_n)$  对不同  $N$  和  $a$  的期望和方差。每组  $N$ ,  $a$  进行 100 次试验, 每次随机采样。具体结果如下表 1, 2, 3 所示 (仅截取了前 24 次试验结果, 具体结果见 `expN_100.csv` 文件), 其中第  $i$  行第  $j$  列表示第  $j$  组的  $N$ ,  $a$  值的第  $i$  次的数值  $E(p_n)$ , 第 101 行为期望, 第 102 行为方差。可以看出在窗宽  $a$  相同的情况下, 期望和方差随着  $N$  增大而减小, 说明

估计效果变好，与(a)中图像结果相符；在样本数  $N$  相同的情况下，期望和方差随着窗宽  $a$  增大而减小，说明在  $a=0.25,1,4$  三个参数中， $a=4$  的估计效果最好。

(d) 在给定  $N$  的条件下，应该选取对应的期望值  $E(p_n)$  最小的时候的窗宽，此时是最优的。

表 1

$\alpha=0.25$

	N=5	N=10	N=50	N=100	N=500	N=1000	N=5000	N=10000
1	2448.933155	1018.884579	305.3340328	202.8152001	124.8901538	96.03711533	124.4002032	124.1069939
2	1924.812308	1040.637492	210.8342371	208.7734279	139.5151898	149.1766556	133.8654667	141.7664823
3	2128.995219	877.254061	500.6265048	266.6444711	152.8834812	144.959384	122.9804207	130.9224278
4	2854.723621	1203.495659	511.5709701	355.0656639	189.1379452	164.3443573	145.723864	135.7752801
5	2394.41288	2212.899971	522.0453634	292.4448911	201.8756871	180.0145627	148.4520315	138.3240802
6	2426.418607	1001.191735	312.3412331	210.289558	168.928705	156.3805482	140.2874286	132.2134526
7	3020.950571	949.6432788	285.9474411	189.2845559	149.1958292	151.386134	127.6565409	133.0209986
8	4643.317543	1839.372264	361.3894934	297.7653048	162.8575331	143.1820235	131.3811979	131.5873881
9	2342.963955	921.1539344	325.0759333	206.1755416	129.119605	122.1165494	121.04427	126.6177233
10	2572.134941	985.1988585	238.4436386	176.9275254	124.3087281	116.2207992	131.923525	129.5731099
11	1274.74483	1375.119329	202.1965334	106.3660445	100.2803776	116.5134474	130.5870345	124.6215712
12	2873.27014	956.1206832	721.9849066	358.7049268	178.8861472	160.4286917	124.3805108	125.5266634
13	2627.012582	1519.879541	527.7899362	317.8223112	153.7927625	149.6843714	129.2379038	147.6594754
14	3968.622034	1423.209241	311.1004971	365.9648769	142.9284662	134.1855402	119.3284258	118.137466
15	1593.746455	1501.442784	276.8108143	158.7681365	108.2176481	111.8562825	112.8052542	120.1739189
16	3864.329025	1726.84385	323.1157307	195.0623646	108.9469745	131.9379276	133.579444	128.5026289
17	2313.638264	836.6495193	375.3521428	209.1434333	110.490173	105.8611815	118.2036954	114.7221963
18	2168.957743	932.2732552	407.542126	218.1995865	163.5015727	174.7627006	127.4758068	133.0398807
19	1636.686743	1287.141192	336.2723951	202.3022158	166.5576774	128.8248582	136.981459	135.6913289
20	1433.453062	1366.26011	390.1884214	296.9275585	189.9082282	159.2412315	158.6063334	154.3950239
21	1562.154369	782.1307885	251.4760727	166.7065054	164.1998389	147.8106227	124.2504516	122.0349009
22	2513.478823	1623.180831	299.29929	236.2084187	164.1431112	149.6603295	139.8572276	125.0903427
23	4392.842391	1237.041822	673.239434	462.1323948	185.8134716	133.4063209	132.8616972	121.9175096
24	3291.739274	2170.357692	427.3241194	366.6122157	151.2894693	138.2849682	134.8024258	137.3115567
期望	2499.043993	1279.149425	381.0594762	265.8060691	156.3092179	139.9989113	132.2369847	131.3134712
方差	829881.3791	185281.7881	23381.65179	11395.53882	1016.406511	390.0616364	110.8916356	63.58852415

表 2

	a=1							
	N=5	N=10	N=50	N=100	N=500	N=1000	N=5000	N=10000
1	597.0555739	126.7514938	126.3058144	113.7256903	79.30559801	65.30146401	91.0928422	91.27445681
2	148.668028	106.2895311	44.26483327	72.85791919	89.36335954	102.4887096	98.0206337	103.9482524
3	568.2956471	123.3868457	205.760417	129.4317346	97.00717546	97.23632785	90.11514456	97.77348197
4	415.4885218	218.2153242	277.451456	211.9260581	129.6171571	117.4477774	105.715351	99.65713061
5	319.5365405	211.7255675	184.5555864	107.178885	123.884566	115.2880449	105.1641224	101.2831297
6	648.795872	241.3382364	116.8926987	100.7054606	117.7684887	112.7029553	104.6336047	98.20223489
7	562.3370964	279.6952172	142.4801338	93.68838734	102.2990622	106.0421569	93.53054957	98.82488487
8	845.6254645	456.5701991	117.6373353	156.4212307	112.7406924	103.3068344	97.6006208	98.52476939
9	768.4910003	121.3885959	138.1064876	84.88198449	86.61711673	88.8656644	90.00903181	94.72796289
10	679.79559	63.72123265	108.0809353	69.97206878	82.04929547	83.03510796	96.03161271	96.26888597
11	342.3226738	220.5666983	21.35614735	26.67232734	70.00629526	80.80228812	94.33074952	91.33977091
12	425.8775512	201.0325079	309.9283812	149.8606117	108.0749005	105.5117057	89.92260541	92.31284022
13	146.2372848	162.4088494	153.4846775	122.2915507	103.7230893	108.3726768	95.98270632	108.8914501
14	1284.76621	327.4725047	139.3752102	190.608297	93.12554155	94.85515237	88.49165546	87.90407653
15	219.4326031	267.3413598	102.0696809	84.2109029	72.47108107	80.15305581	82.89345049	90.32117875
16	672.4289234	252.4859149	112.4303901	84.89290698	74.46134079	90.57974914	98.56331437	94.9377114
17	402.5487444	354.8364498	188.3961356	95.03466276	75.15330577	78.27207603	87.68484264	85.75255892
18	255.3705334	133.044539	107.651987	124.4741688	118.9737619	124.2158139	93.40502997	97.63252762
19	334.2242553	253.6125848	134.0154746	76.43746782	101.4808723	90.19836729	101.9683814	101.427885
20	134.3478184	111.8929213	136.9740956	162.2011328	123.1477721	106.4127836	113.6850117	112.0453427
21	157.4015094	113.9513233	76.81751344	88.16655618	113.4388063	102.5225931	90.12018962	89.94521723
22	531.1991462	299.1210403	83.32333254	90.82832156	108.8260817	104.5299067	101.3855682	92.60383621
23	970.7367548	286.9401731	218.8218303	231.0042811	127.6185644	87.18117352	96.84417262	90.75457391
24	1093.341346	898.6711673	217.6108833	172.3202074	98.30735176	91.66681105	99.31558075	102.287317
期望	501.5644579	276.2602872	143.9700325	119.2973547	102.2925497	97.7032202	97.25679099	97.34544121
方差	92876.97144	33153.46916	4976.034549	2581.997122	411.9053451	182.5969295	52.29927247	31.02072359

表 3

	a=4							
	N=5	N=10	N=50	N=100	N=500	N=1000	N=5000	N=10000
1	34.18101314	51.4350497	33.60174831	38.45905566	35.00571928	33.8286431	33.22182613	33.04390713
2	34.96463931	29.81724831	36.27690828	30.55494191	31.70631677	31.60445973	32.11785185	32.22278789
3	29.38735902	31.96210111	35.00111246	32.3484569	32.16860268	34.24097394	33.62294698	32.85260666
4	69.42565626	49.26872492	34.74011022	33.01339547	32.71788343	33.21297235	33.23314805	32.9224892
5	96.98241171	49.96839005	33.4856878	33.00853322	31.72051278	32.78812968	32.9442605	32.63058124
6	48.87719389	38.34932408	32.55362615	31.83261627	32.97175291	32.33772181	33.26073417	33.14606423
7	27.0624679	28.73456003	29.81943867	32.97011605	33.47204215	33.25008345	32.84102681	32.69266471
8	34.48771632	31.9227568	33.31974931	32.1094344	32.24621934	33.06171347	33.08319361	32.75403857
9	34.99690077	49.33310762	34.10421346	31.84174376	32.77868721	32.45171404	32.70944591	32.92125319
10	37.92826567	26.13892109	41.98723384	35.37783923	32.90108363	31.98643157	32.36637708	32.49778126
11	112.2154054	70.7945508	42.32159917	35.95048445	32.42140293	33.27513343	32.79182211	32.72137092
12	30.41946895	26.16105378	29.18291076	32.96606856	30.84407463	31.95293138	32.7191712	32.61050509
13	46.6863178	44.47320368	33.0801666	32.69525845	34.81850563	34.51163817	33.39058086	32.85545411
14	35.1458024	31.37020249	33.54722337	31.54324237	35.12023855	34.52531261	33.08213979	33.12762352
15	26.06228556	39.06037787	39.18101553	34.25030892	32.38815605	32.24936102	32.21798964	32.24060017
16	31.31477195	30.77956731	32.3548872	32.65607384	32.8526982	32.68132347	32.91956391	33.07888798
17	27.42454768	31.23118161	32.86863816	33.40613729	33.79271726	34.39832887	33.41291809	33.01602877
18	26.91725397	37.26784131	30.99960339	30.33637306	32.48547295	31.82781007	32.32945357	32.73848371
19	31.79703486	35.63231143	36.06777187	34.47530769	33.08019241	32.4761663	32.74854726	33.1786811
20	37.81803457	37.36150693	41.68713093	35.00283502	32.88063452	32.7286983	33.40215095	33.04764259
21	25.14856206	27.95244481	39.26367169	34.41216782	34.73554458	33.15508723	33.5221314	32.99472987
22	30.49672179	30.44460346	34.48661681	35.33347142	32.35843185	32.60115117	32.39343645	32.7609893
23	62.7932431	56.56068593	36.5987588	33.5685937	32.14808129	33.64453638	33.20890836	33.29221661
24	47.0901485	37.84217606	33.268939	32.15212764	31.01696477	31.7691726	32.50851503	32.44091435
期望	42.6360126	38.41117308	33.63957764	33.09435032	32.78059034	32.77359021	32.82180259	32.78102584
方差	315.6243341	120.812091	12.95139922	7.331378356	1.614293229	0.894239126	0.233872717	0.128711953



$$P(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

(e) 将方窗替换为 Gaussian 窗，重复上述过程。用不同的窗宽  $a$  来估计  $p(x)$ ，根据课件，不妨取  $a=0.25, 1, 4$  三个值，对应  $N=5, 10, 50, 100, 500, 1000, 5000, 10000$  画出估计图像如下图 4, 5, 6 所示。可以看到在窗宽  $a$  相同的情况下，采样数  $N$  越多，估计效果不一定越好。在  $a=4$  的情况下， $N>500$  的采样值越多，图像越粗糙。在样本数  $N$  相同的情况下，窗宽  $a=4$  的估计效果最好， $a=0.25$  时估计值为较多尖峰函数叠加，使估计值变化幅度很大，而  $a=4$  时估计值较为平缓，分辨率降低，随  $N$  变化估计效果变化不明显。因此，窗宽较大或者较小，估计效果都不好，要根据样本数做折中。

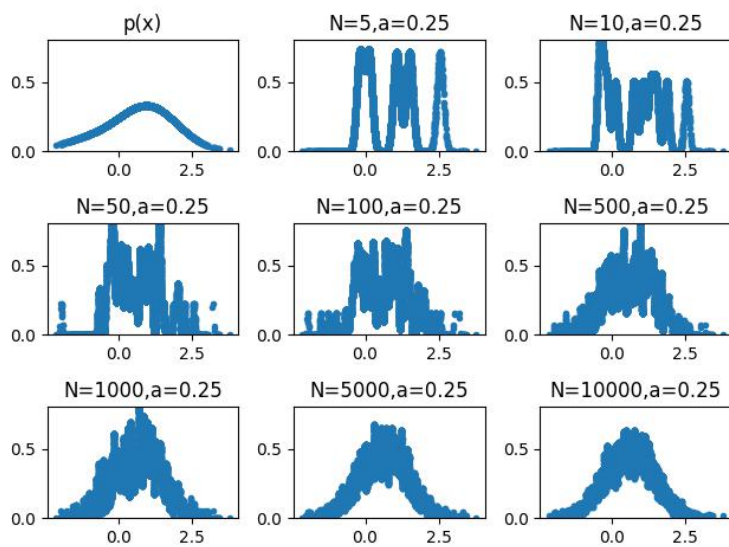


图 4

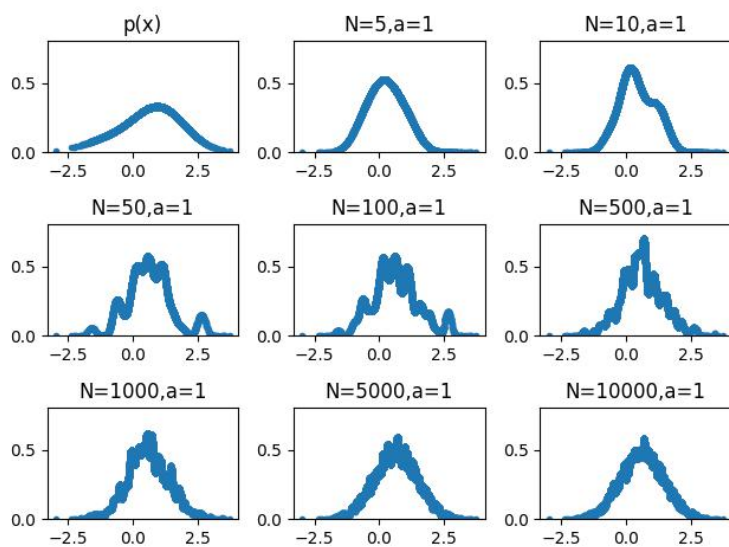


图 5

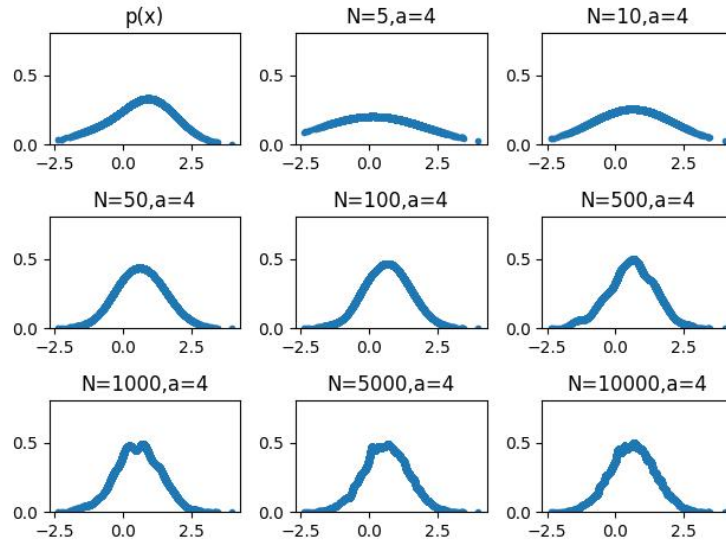


图 6

求 $E(p_n)$ 对不同  $N$  和  $a$  的期望和方差。每组  $N$ ,  $a$  进行 100 次试验，每次随机采样。具体结果如下表 4, 5, 6 所示（仅截取了前 24 次试验结果，具体结果见 `gauss_100.csv` 文件），其中第  $i$  行第  $j$  列表示第  $j$  组的  $N$ ,  $a$  值的第  $i$  次的数值 $E(p_n)$ ，第 101 行是期望，第 102 行是方差。可以看出在窗宽  $a=0.25$ , 1 的情况下，期望和方差随着  $N$  增大而减小，但是窗宽  $a=4$  的情况下，期望随着  $N$  的增大先减小后增大， $N=50$  时期望最小；在样本数  $N$  相同的情况下，期望和方差随着窗宽  $a$  增大而减小，说明在  $a=0.25, 1, 4$  三个参数中， $a=4$  的估计效果最好。

在给定  $N$  的条件下，应该选取对应的期望值 $E(p_n)$ 最小的时候的窗宽，此时是最优的。

由图和表可知，不同的窗函数相同的  $N$  和  $a$  参数也会有不同的表现。

表 4

	a=0.25							
	N=5	N=10	N=50	N=100	N=500	N=1000	N=5000	N=10000
1	1166.358415	572.7182717	515.4795119	696.6483805	255.1200077	229.5752083	192.5926009	188.8910753
2	941.9634449	894.3177101	810.6714092	412.5706325	364.5504208	269.3267645	207.7360002	195.2320867
3	1325.13921	2246.161369	612.2784858	443.2275283	207.8997039	215.6722072	197.0945076	201.1112107
4	1260.020964	681.8909068	361.1733955	366.6764247	317.4240425	268.9890913	195.296903	194.4719993
5	1327.624434	1108.145766	619.0545091	649.3291939	295.0115876	242.3160322	200.8048299	199.2588685
6	1587.087164	826.5606385	424.0119084	353.0860554	276.1123066	253.6141484	186.4459246	180.9774741
7	638.9421133	699.842741	377.6906532	454.0860554	344.9972733	295.6269927	210.7045342	193.676243
8	1572.138006	1187.685351	466.0996119	493.6222482	303.3157982	272.091884	215.875268	189.5312733
9	2134.929471	1141.797292	894.5035466	502.1670797	318.8955253	300.9925329	214.3388812	198.7371065
10	1306.631696	2871.80863	519.3833583	536.1143253	365.0562365	281.6028632	199.4449783	195.3170494
11	628.5129085	628.0367923	519.6104844	385.8045137	273.3716857	214.4620803	226.2370992	207.7062718
12	1495.702072	1638.135222	656.7363584	427.8575267	286.7483326	300.3116088	218.998987	183.998672
13	1534.073571	596.39626	606.579777	632.5114834	359.0936281	277.8784374	219.0450285	203.4223017
14	713.8666864	1348.771776	1026.779788	650.0423985	370.3920131	249.7179498	196.9813925	190.1634204
15	2588.504106	2515.005165	517.3255268	578.2940853	252.1552702	236.8968122	192.4371619	185.7694297
16	1121.068579	1792.548847	936.2918927	702.5866671	380.844836	274.8314332	210.8230958	201.9766184
17	894.2903279	727.6181388	370.750646	491.9060735	275.8910689	266.867687	215.3141312	201.5925571
18	995.0976467	1108.030331	327.2657077	420.9764446	251.1334169	275.0816676	199.0643101	193.3487979
19	570.9258719	911.2566463	527.822347	330.4019891	281.4149038	249.3818361	212.8003042	217.1023639
20	1774.395025	855.7862039	745.6437642	639.3322192	289.0976287	252.5549372	198.9381971	195.156761
21	1522.785625	1150.869561	595.6195726	648.3206182	356.961196	279.1113778	199.8851879	181.523478
22	1249.925174	1047.274629	444.415944	390.9337036	292.7899109	274.1385798	221.8572609	217.3077743
23	2356.975529	2396.530881	547.4903589	417.6227195	376.7318172	264.7531031	200.845431	181.7227726
24	1677.650849	1132.639404	787.8208397	495.0872592	275.121624	246.6688514	193.0009925	205.969665
期望	1570.798522	1273.375289	630.6838584	511.8806971	312.8471078	263.2928544	205.5574026	196.4887318
方差	567984.591	329203.158	34840.49082	17715.0351	2127.751051	751.0550651	158.3237126	99.13297489

表 5

	α=1							
	N=5	N=10	N=50	N=100	N=500	N=1000	N=5000	N=10000
1	306.6978223	159.9564061	164.9799028	325.170844	144.5334052	142.4560044	139.0537264	139.2575752
2	133.4393612	225.7082485	322.8042099	193.2377412	201.4700143	170.3738512	153.2267161	149.2399612
3	88.18671207	357.5991654	214.4685234	160.5446845	113.9348574	133.2227292	139.7779769	147.2962289
4	240.7959838	53.44396692	127.9562652	149.6858978	173.0645871	149.0576808	131.2081642	142.3376252
5	117.5107294	209.3593994	264.6488215	294.6341246	140.1074106	137.1826091	144.6909259	150.992453
6	642.4562046	79.78443783	70.77147312	77.68819678	133.261082	150.4100533	134.6549001	137.7698378
7	157.8112934	35.88576707	132.2652334	161.3234712	152.9094436	179.0760064	148.3188214	144.5092948
8	114.4986686	187.5395934	185.0233508	223.0602699	170.0496293	180.9033832	157.2061979	141.1369605
9	474.6202688	417.4123551	211.7340081	160.1618735	161.5921915	192.2695337	154.069334	148.8582935
10	138.9268129	636.3810456	204.431537	274.0279681	191.0143413	168.5264977	137.6418268	146.2208703
11	54.65058296	32.1025256	130.3612751	154.5499608	148.0462436	129.6059466	160.5843239	157.7618826
12	182.4410963	198.5429198	196.6210736	209.4653686	155.4814186	199.2509305	158.0633262	135.9286525
13	161.0206063	76.78809342	283.8507851	276.2843949	220.5642311	187.5358987	153.3269105	149.3444513
14	169.6653729	307.3073523	318.2834679	268.9010963	222.4309072	166.5326448	137.3567618	138.3971532
15	200.9435297	571.1419057	178.6116216	257.2248054	147.8382898	138.4865833	141.9232271	138.3762702
16	248.3140804	564.0100145	433.5980523	227.9538285	222.9709226	177.7987255	150.9745132	153.2836849
17	166.2108119	307.076545	84.41337386	177.527449	137.2989989	150.9172828	156.495845	146.7186295
18	127.0037292	234.55542	95.04288957	122.7002257	102.9240312	139.8916543	142.7600619	142.5696169
19	59.18408595	109.4916359	218.3871156	82.87437781	143.7583897	148.339949	160.0938407	169.002087
20	40.14750822	17.37793217	190.6270468	260.8657397	168.415124	160.4839496	146.2445771	144.3265911
21	162.3523323	229.5250508	253.2537624	333.1328195	219.5911547	181.1178268	144.7950458	137.1682492
22	147.9343483	107.4532195	150.8618418	148.5902872	160.1290566	164.7389889	161.8214475	159.4379719
23	307.484276	176.6201719	185.6473731	165.5713407	195.3775197	160.804761	130.9968692	133.1115384
24	234.3559135	182.5661642	180.847629	178.6400959	149.8968013	144.2123651	136.031546	153.923872
期望	242.4015735	265.6818323	214.8156015	209.6841273	172.7009011	162.6158853	148.3424584	146.8750468
方差	28916.42646	34244.855	10175.5006	5019.552672	1122.005441	480.8101844	131.0286316	76.15015079

表 6

	α=4							
	N=5	N=10	N=50	N=100	N=500	N=1000	N=5000	N=10000
1	66.05315877	11.04021814	31.50458753	95.84862395	98.31683234	116.3184605	128.0365768	128.725904
2	116.5537389	45.0602805	50.42415761	62.9710688	144.1904922	141.4400937	134.71453	136.8157263
3	84.26275547	15.12188051	40.8637163	58.15605703	84.92703938	99.15746316	121.1665735	130.0560073
4	112.6220891	32.14635596	15.59293334	67.03088459	127.1358202	117.8593893	115.5836512	127.9115072
5	75.60726588	38.28039877	69.06983627	145.7381319	104.2431391	115.3890538	128.6584558	137.7967754
6	93.2527337	25.11425323	20.20503851	22.66234336	86.47847594	119.656113	120.1884969	125.2906189
7	120.4323251	47.43076946	39.52054343	61.91039558	110.859753	132.8852804	136.2305931	133.8086325
8	88.24939073	29.08645973	37.62110484	74.16278719	137.4198177	152.0294362	141.6499618	129.7679876
9	76.99115091	59.38383918	34.9641892	76.75537422	124.4858888	151.4133833	136.6212566	133.4385584
10	75.95458493	5.892329075	62.07996246	121.0597465	122.6802443	130.213279	123.7934239	133.1241715
11	137.4509888	56.19787252	36.48502635	71.78694307	109.9712034	107.5840735	142.3132093	142.044727
12	75.5164967	26.91928068	40.41275432	104.5014835	124.8880887	172.2398171	141.3939455	125.3581542
13	91.8401845	29.99038568	43.48844554	121.4150596	164.8762178	159.9941343	140.3549386	135.142129
14	136.0382564	69.17949931	73.25719433	98.84500992	150.9075921	134.534347	125.6340395	127.6794558
15	77.94797425	25.61030655	43.70250594	93.90029365	117.7316117	118.4319999	131.0605462	125.6036715
16	70.20843331	34.75215716	118.6383518	106.2057554	159.0807267	144.0833946	137.4895646	141.3268938
17	73.70309634	19.83572806	21.47537207	64.28955089	106.2794798	127.644039	139.0494166	129.8439474
18	70.02868417	15.35097227	42.98757432	53.15973488	75.23379768	106.0691579	127.2117705	130.743957
19	90.95652248	22.57625457	59.04826196	40.29877184	101.9115149	120.3853521	148.6758445	155.0723412
20	81.99159679	37.22562403	16.66899705	104.8279182	125.1732886	134.7495935	134.1448729	133.098128
21	110.3595532	32.45258191	66.1042681	147.3629492	158.6316606	154.3854864	131.8516166	128.3956929
22	89.51770044	33.97784196	27.38780842	49.11478928	115.8008202	118.6412001	144.5328295	145.7461557
23	65.78756183	22.95176086	44.06174625	74.42377997	113.3464223	115.8137306	116.3296503	121.5220596
24	73.76189418	38.86603983	42.21905745	69.11904679	103.2717163	113.6372226	122.578128	140.5754496
期望	89.46402648	30.38877812	48.5579724	86.51891872	125.752432	131.7943118	133.3131194	134.2879474
方差	423.1484556	246.3061223	709.8176462	997.8538665	613.8004953	372.1586997	115.9383569	70.09763098

(f) 由上结果可知, Gaussian 窗  $N=10$ ,  $a=4$  的取值方法最优, 此时期望为 30.38, 方差为 246.31; 与方窗  $N>500$ ,  $a=4$  的取值方法相差不大, 此时期望为 32 左右, 方差为 1 左右。进一步说明不同窗函数下不同参数有不同的估计结果。