شبیه سازی رایانه ای در فیزیک
تمرین ۶

علی ستاره کوکب
شماره دانشجویی: ۹۵۱۰۰۴۹۱

۳ آذر ۱۳۹۹

۱ مقدمه

در این سری از تمرین ها می خواهیم به بررسی شبکه اردوش-رنی [۱] و انتگرال گیری به روش مونت کارلو[۲] و الگوریتم متروپلیس [۳] برای تولید اعداد تصادفی بپردازیم.

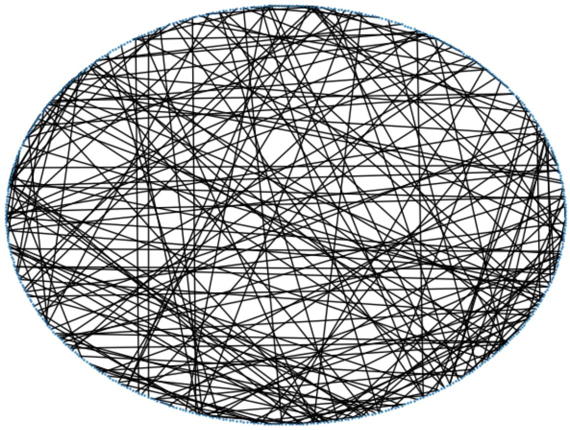
۲ شبکه اردوش-رنی

گراف اردوش رنی N راس دارد که هر دو راس به احتمال p می توانند به یکدیگر وصل باشند. برای ساختن این گراف از کتابخانه ی networx در پایتون استفاده می کنیم. در این کتابخانه دستور networx.fast_gnp_ random_graph برای این کار تعبیه شده است. این متد دو ورودی نیاز دارد. یکی تعداد راس های گراف مورد بررسی و دیگری احتمال وصل شدن دو راس به یکدیگر (p). در سوال میانگین تعداد یال ها داده شده است. برای آنکه p را بدست آوریم از رابطه ی ۱ استفاده می کنیم. این رابطه ارتباط میان میانگین تعداد یال ها و احتمال p را مشخص می کند.

$$p = \frac{\langle k \rangle}{N} \tag{۱}$$

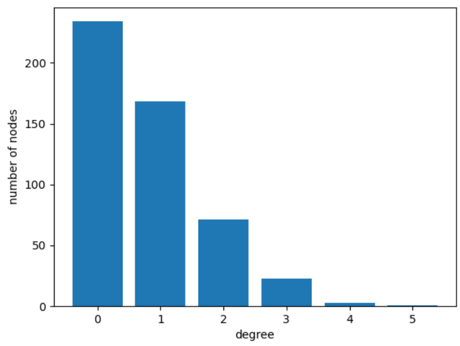
حال با داشتن این احتمال گراف مورد نظر را تولید می کنیم. در شکل ۱ نمایش یک گراف اردوش-رنی را می بینید. برای یافتن توزیع مرتبه ی راس ها از متد Graph name degree()

[۱] Erdős–Rényi
[۲] Monte Carlo
[۳] Metropolis

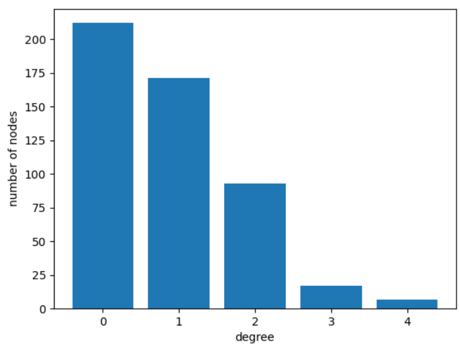


شکل ۱: نمایش یک گراف اردوش-رنی با 500 راس و میانگین تعداد یال های 0.8

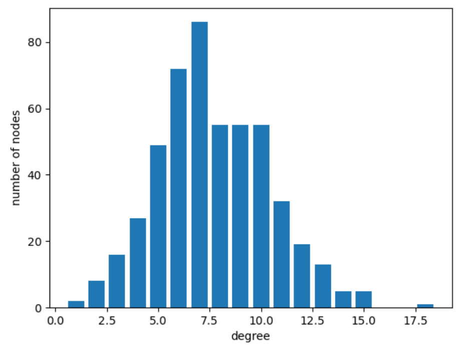
استفاده می کنیم و سپس با استفاده از تابع collections.counter تعداد راس ها با هر مرتبه را بدست می آوریم و سپس با استفاده از تابع matplotlib.bar تابع توزیع را رسم می کنیم که در شکل ۲ می بینید.
برای بدست آوردن توزیع ضریب خوشگی، از متد networkx.clutering استفاده می کنیم و سپس با استفاده از دستور hist، هیستوگرام توزیع را رسم می کنیم.(شکل ۳)



(آ)

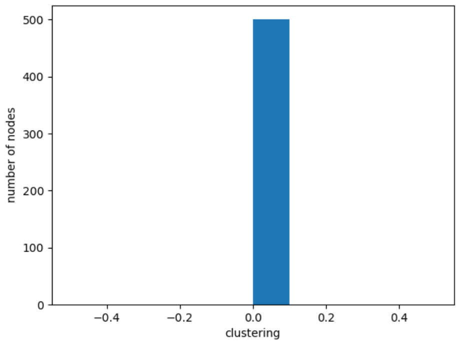


(ب)

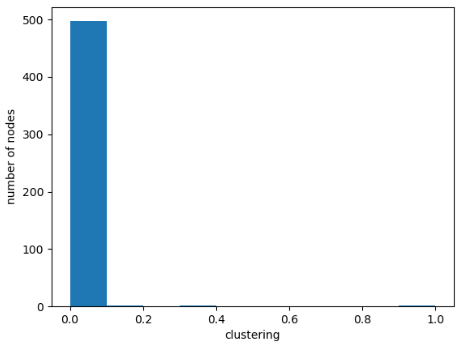


(ج)

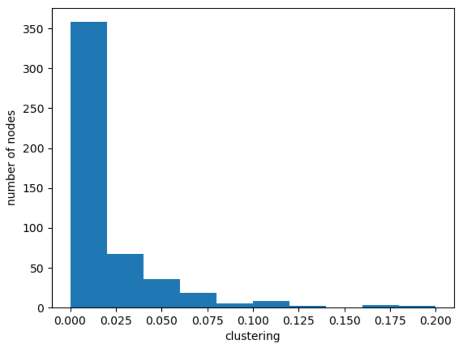
شکل ۲: آ) نمودار توزیع درجه رئوس برای میانگین یال ۸.۰ ب)نمودار توزیع درجه رئوس برای میانگین یال ۱ ج)نمودار توزیع درجه رئوس برای میانگین یال



(آ)



(ب)



(ج)

شکل ۳: آ) نمودار توزیع خوشه ها برای میاگین یال ۸.۰ب)نمودار توزیع خوشه ها برای میانگین یال ۱ ج)نمودار توزیع خوشه ها برای میانگین یال ۸

۳ انتگرال گیری مونت کارلو

در این تمرین می خواهیم انتگرال $\int_0^2 e^{-x^2}dx$ را با استفاده از روش مونت کارلو بدست آوریم. در قسمت اول به روش نمونه گیری ساده این کار را انجام می دهیم. برای این کار به کمک تابع numpy.random.uniform اعداد تصادفی با توزیع یکنواخت در بازه ی ۰ تا ۲ تولید می کنیم سپس میانگین تابع e^{-x^2} را روی این اعداد بدست می آوریم. این کار را برای تعداد مختلف نمونه گیری انجام می دهیم و نمودار خطا برحسب تعداد نمونه را رسم می کنیم. (شکل ۴) همانطور که می بینید خطا بصورت $\frac{1}{\sqrt{n}}$ کاهش می یابد. مقدار بدست آمده برای یک میلیون نمونه و ۱۰ بار تکرار برابر 0.8818 ∓ 0.0006 می باشد. در قسمت بعدی همین انتگرال گیری را با استفاده از نمونه گیری هوشمند بدست می آوریم. در این نمونه گیری به جای آنکه از تابع توزیع یکنواخت برای تولید اعداد تصادفی استفاده کنیم، از تابع توزیعی استفاده می کنیم که به تابع مورد انتگرال گیری نزدیک باشد. در این قسمت از تابع e^{-x} برای این کار استفاده می کنیم. دقت می کنیم که این تابع را در بازه ی انتگرال گیری نرمال کنیم. این تابع را با $g(x)$ نشان می دهیم.

$$g(x) = \begin{cases} \frac{1}{1-e^{-2}}e^{-x}, & \text{if } 0 \leq x \leq 2 \\ 0, & \text{otherwise} \end{cases} \tag{۲}$$

پس از نرمال کردن، با استفاده از تابع تبدیل، تابعی را که تابع توزیع یکنواخت را به این تابع توزیع تبدیل می کند را می یابیم. این تابع را به $h(x)$ نشان می دهیم.

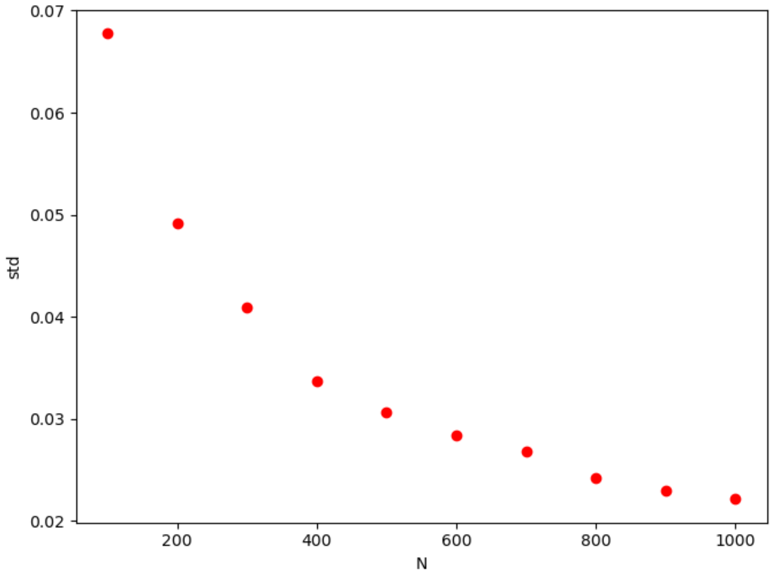
$$h(x) = -ln(1-(1-e^{-2}) \times \frac{x}{2})$$

حال با داشتن این تابع توزیع برای محاسبه ی انتگرال کافی است میانگین $\frac{e^{-x^2}}{g(x)}$ را روی این تابع توزیع بدست آوریم. حاصل بدست آمده برای این انتگرال برای یک میلیون نمونه و ده بار تکرار برابر 0.8820 ∓ 0.0003 می باشد. نمودار انحراف معیار برحسب تعداد نمونه در شکل ۵ آورده شده است.
پ

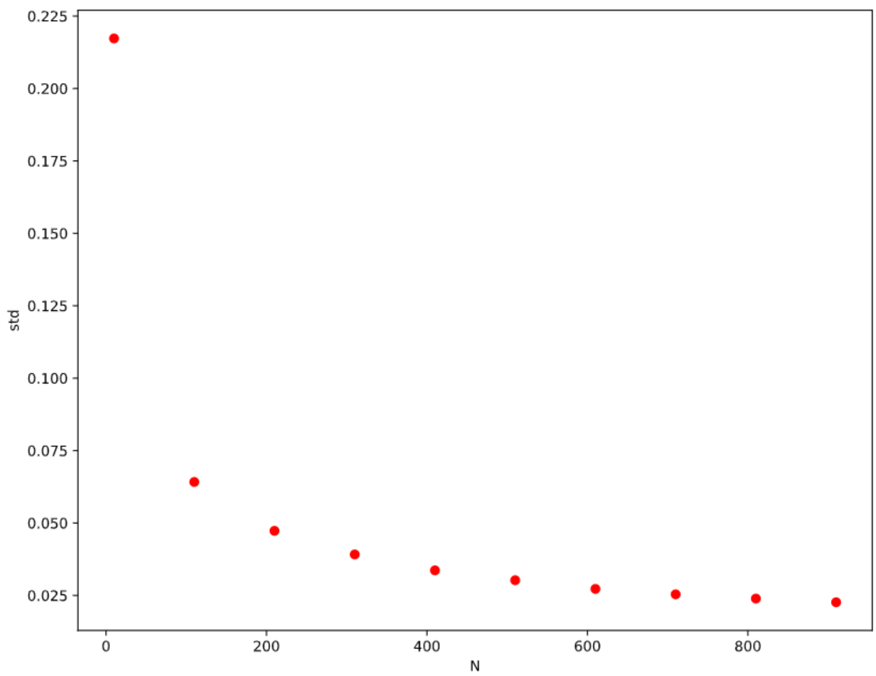
۴ انتگرال گیری مونت کارلو (امتیازی)

برای قسمت امتیازی نمونه گیری هوشمند را این بار با تابع $\frac{1}{x+1}$ انجام می دهیم.

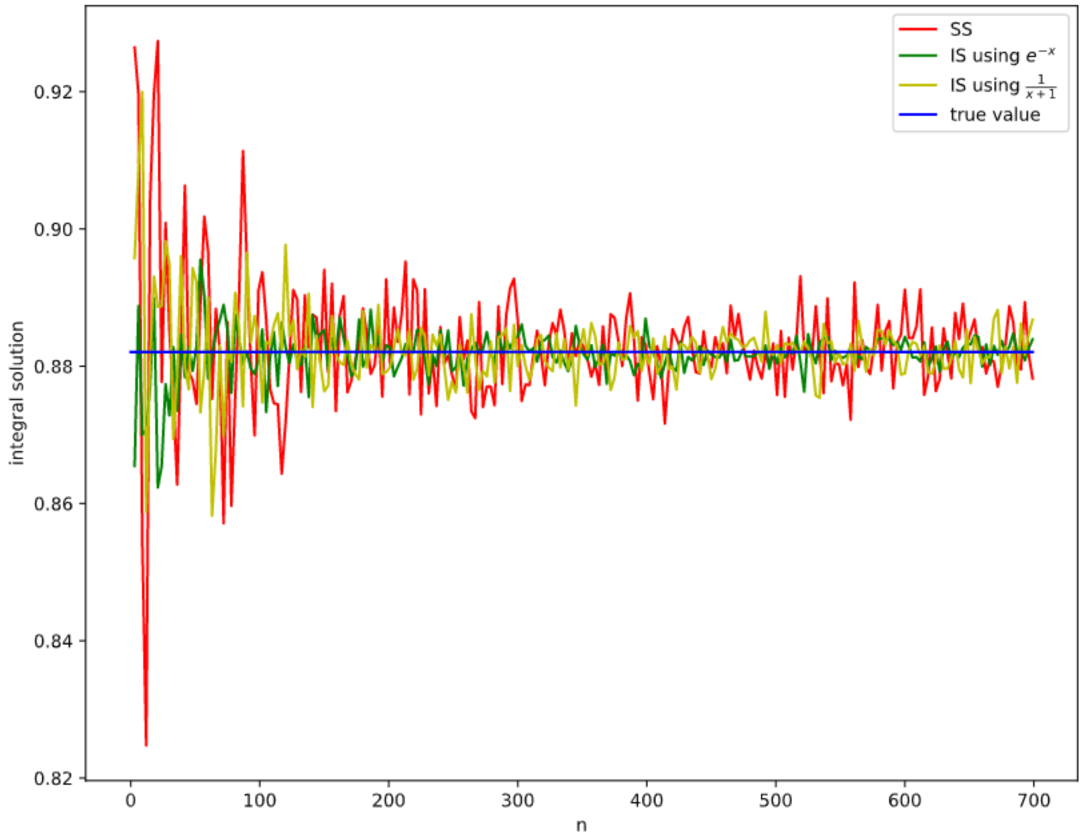
$$g(x) = \begin{cases} \frac{1}{ln3(1+x)}, & \text{if } 0 \leq x \leq 2 \\ 0, & \text{otherwise} \end{cases} \tag{۳}$$



شکل ۴: نمودار انحراف معیار برحسب تعداد نمونه برای نمونه گیری ساده



شکل ۵: نمودار انحراف معیار برحسب تعداد نمونه برای نمونه گیری هوشمند



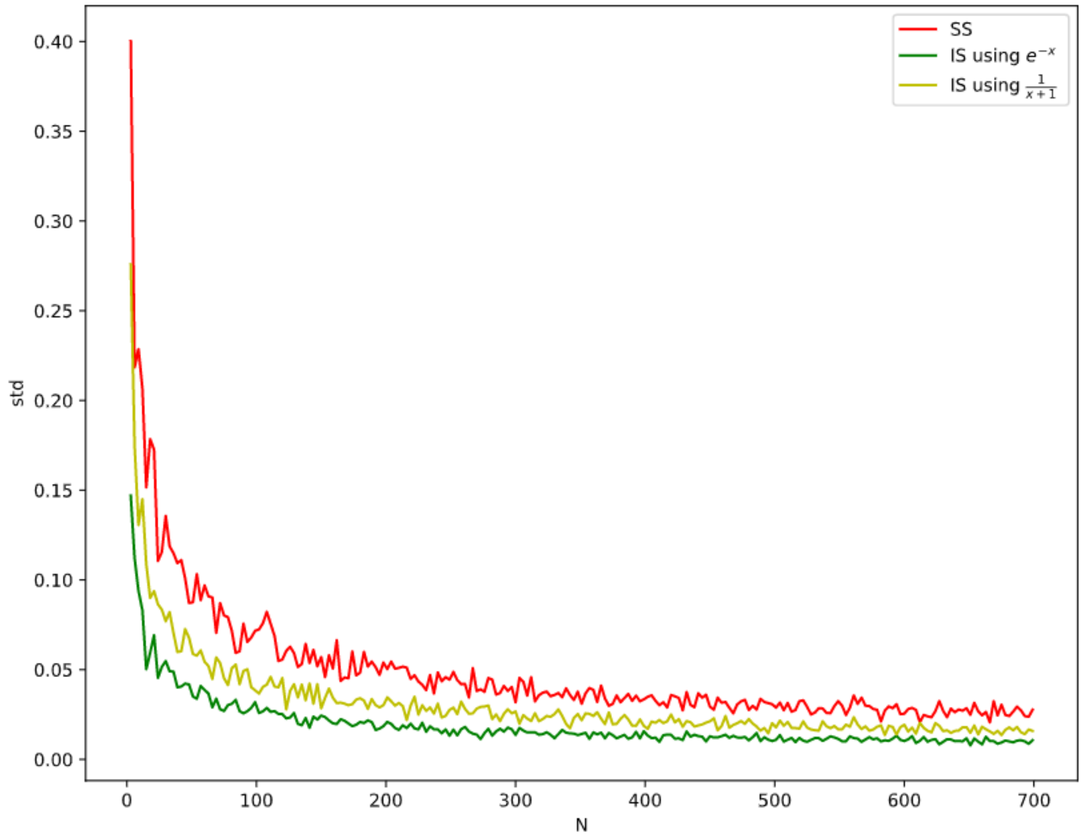
شکل ۶: مقایسه ی رفتار میل کردن به جواب انتگرال برای سه روش استفاده شده

تابع تبدیلی که برای بدست آوردن این تابع توزیع باید بکار ببریم برابر خواهد بود با:

$$h(x) = \frac{3}{2}x - 1$$

حال با داشتن تابع توزیع مشابه قسمت قبل حاصل انتگرال را بدست می آوریم. حاصل این انتگرال برای یک میلیون نمونه گیری و ده بار تکرار برابر 0.8820 ∓ 0.0003 می باشد. در شکل ۶ مقایسه بین این سه تابع انجام شده است است. خط افقی در این شکل برابر مقداری است که از نرم افزار Mathematica برای این انتگرال بدست آمده است. این مقدار برابر 0.8820813 می باشد. در شکل ۷ مقایسه میان افت انحراف معیار برحسب تعداد نمونه ها صورت گرفته است. در شکل ۸ نیز شکل خود این توابع را می بینید. در جدول های ۱ تا ۳ مقادیر خطای آماری ΔI، خطای از مقدار واقعی انتگرال ΔI_m و همچنین زمان مورد نیاز $t(S)$ برای اجرا را برای هر سه ی این توابع و برای نمونه های مختلف می بینید.

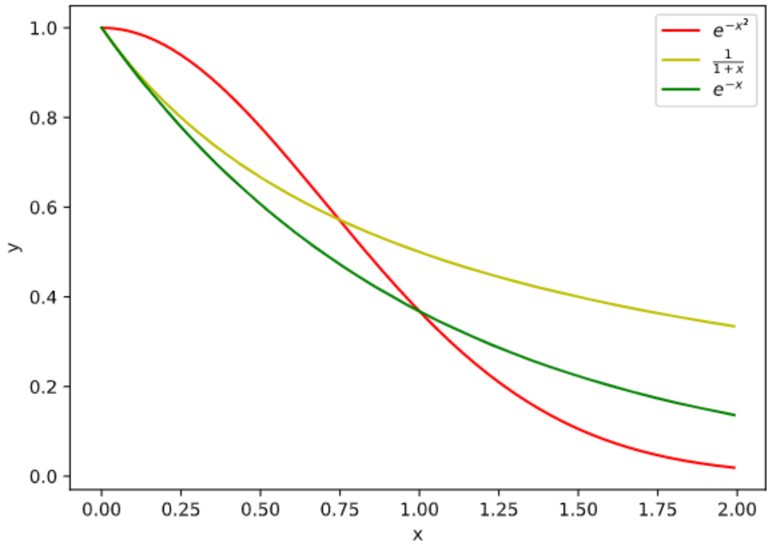
در شکل ۹ نمودار زمان اجرا برحسب تعداد نمونه ها را می بینید.



شکل ۷: نمودار مقایسه ی افت خطا ی آماری برای سه روش بکار برده شده

Table 1: Table of errors and run-time forsimple sampling

n	I	delta I	delta I_m	t(s)
100	0.881151	0.063672	0.00093	0.000272
200	0.887643	0.052498	0.005562	0.000405
300	0.886452	0.040061	0.00437	0.000594
400	0.880718	0.030521	0.001363	0.000798
500	0.884055	0.034513	0.001974	0.000982
600	0.886483	0.029753	0.004402	0.001175
700	0.886092	0.022013	0.004011	0.00136
800	0.886966	0.024063	0.004885	0.001561
900	0.881491	0.021824	0.000591	0.001729
1000	0.882988	0.021755	0.000906	0.00194



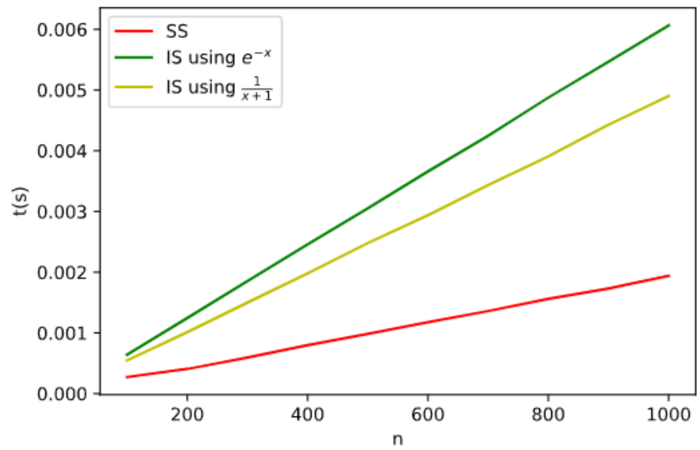
شکل ۸: شکل سه تابع مورد استفاده در این تمرین

Table 2: Table of errors and run-time for important sampling

100	0.881367	0.027177	0.000714	0.000642
200	0.882575	0.019598	0.000493	0.001245
300	0.882378	0.014737	0.000296	0.001855
400	0.882396	0.013307	0.000315	0.002456
500	0.88174	0.011737	0.000341	0.003051
600	0.881792	0.010983	0.000289	0.00366
700	0.882487	0.009843	0.000405	0.004248
800	0.882134	0.009495	5.3E-05	0.004875
900	0.882519	0.008948	0.000438	0.005467
1000	0.882043	0.008413	3.85E-05	0.007765

Table 3: Table of errors and run-time for bonus part

1-4	n	I	delta I	delta I_m	t(s)
	100	0.882827	0.040961	0.000746	0.000545
	200	0.881335	0.030392	0.000746	0.001011
	300	0.882267	0.024008	0.000186	0.0015
	400	0.882145	0.021867	6.36E-05	0.001982
	500	0.88239	0.01952	0.000309	0.002479
	600	0.882704	0.01737	0.000622	0.002938
	700	0.881997	0.016422	8.41E-05	0.003436
	800	0.881753	0.015166	0.000329	0.003905
	900	0.882279	0.01388	0.000197	0.004426
	1000	0.882396	0.013454	0.000315	0.0049



شکل ۹: مقایسه ی زمان اجرا برای سه روش بکار برده شده

۵ انتگرال گیری چندگانه

در این قسمت می خواهیم مرکز جرم کره ای توپر را بدست آوریم. در صورت سوال گفته شده که در بالاترین قسمت کره (قطب شمال) چگالی دو برابر پایین ترین قسمت کره (قطب جنوب) می باشد و رابطه ی چگالی بصورت خطی از قطب شمال تا جنوب تغییر می کند. با این توضیحات، شکل تابع توزیع بصورت $\rho = ax + 3aR$ می باشد که R شعاع کره می باشد. پارامتر a پارامتر آزاد می باشد و آن را برابر ۱ قرار می آوریم. بنابراین در رابطه ۴ شکل این تابع را می بینید.

$$\rho(x, y, z) = z + 3R \tag{۴}$$

حال از تعریف مرکز جرم می دانیم:

$$\vec{R} = \frac{\int \rho \vec{r} dv}{\int \rho dv} = \frac{1}{\int \rho dv}(\hat{x} \int \rho x dv + \hat{y} \int \rho y dv + \hat{z} \int \rho z dv)$$

بتابراین برای محاسبه ی مرکز جرم باید حاصل سه انتگرال سه گانه را بدست آوریم. برای اینکار مشابه روشی که برای انتگرال یک بعدی مونت کارلو ی یک بعدی بکار بردیم، نیاز است تا اعداد تصادفی در بازه ی انتگرال گیری تولید کنیم. در اینجا بازه ی انتگرال گیری ما بر روی حجم یک کره می باشد. به عبارت دیگر ما باید میانگین عبارت های داخل انتگرال بالا را رو ی نقاط این حجم بدست آوریم. برای آنکه بتوانیم تابع توزیع یکنواخت در یک کره به شعاع R تولید کنیم ابتدا تعدادی اعداد تصادفی در یک مکعب به طول ضلع $2R$ تولید می کنیم. این کار به آسانی با سه تابع توزیع یکنواخت در بازه ی $(-R, R)$ ممکن است. حال که این اعداد را تولید کردیم تنها آنهایی را نگه می داریم که شرط $x^2 + y^2 + z^2 \leq R^2$ را برآورده سازند. این شرط تضمین می کند که اعداد داخل کره می باشند. حال که نقاط مورد نیاز را بدست آورده ایم کافی است تا میانگین اانتگرال دهنده های بالا را بر روی این نقاط بدست آوریم. دقت می کنیم که یک فاکتور $\frac{4}{3}\pi R^3$ در حاصل میانگین گیری ضرب کنیم. علت این امر را در محاسبات زیر مشاهده می کنید.

$$\iiint f(x, y, z) = \frac{\int p(x, y, z) \iiint f(x, y, z) p(x, y, z)}{\int p(x, y, z)} = \int p(x, y, z) dv \times \langle f \rangle_{p(x)}$$

در عبارت بالا $p(x)$ تابع توزیع احتمال سه بعدی می باشد. حال اگر برای این مساله تابع توزیع بصورت یکنواخت باشد، حاصل انتگرال عبارت خط آخر برابر با حجم رویه ای است که روی آن انتگرال گیری انجام می شود که در اینجا برابر حجم کره می باشد. بنابراین برای یافتن مرکز جرم کافی است چهار انتگرال سه گانه را محاسبه کنیم یکی برای بدست آوردن جرم کل و سه تای دیگر برای یافتن مختصات مرکز جرم. در این جا ذکر این نکته ضروری است

که تقریبا نصف نقاطی که داخل مکعب تولید می شوند در داخل کره می افتند و عملا نصف نقاط تصادفی تولید شده در میانگین گیری قابل استفاده نیستند.در کد ارائه شده برای این تمرین شعاع کره برابر ۱ در نظر شده و میانگین گیری برای محاسبه ی انتگرال روی یک میلیون نقطه صورت گرفته. در انتها مختصات مرکز جرم تا سه رقم اعشار برابر با مقادیر زیر بدست آمده:

$$x_{cm} = 0, y_{cm} = 0, y_{cm} = 0.066$$

مقادیر دقیق مختصات مرکز جرم نیز برابر عبارات زیر می باشند:

$$x_{cm} = 0, y_c m = 0, z_c m = 0.066$$

۶ متروپلیس

در این تمرین می خواهیم به وسیله ی الگوریتم متروپلیس تابع توزیع احتمال گوسی را تولید کنیم. برای این کار مطابق الگوریتم ارائه شده در کتاب دکتر اجنهادی پیش می رویم. در ابتدای برنامه تابع گوسی را تعریف می کنیم. این تابع انحراف از معیار و متوسط را به عنوان ورودی می گیرد و تابع گوسی مورد نظر را تولید می کند. در اینجا نتایج برای یک تابع توزیع با میانگین صفر و انحراف از معیار ۴.۰ محاسبه شده است. در ابتدای الگوریتم، مکان اولیه را برابر صفر (میانگین تابع توزیع) قرار می دهیم. در هر مرحله یک عدد تصادفی بین منفی یک و یک تولید می کنیم و این عداد را در یک متغیر دیگر که با αنشان می دهیم ضرب می کنیم و حاصل این عملیات را در هر مرحله با مکان کنونی (x) جمع می کنیم تا مکان مرحله ی بعد (y)را بدست آوریم. حال احتمال وجود در مکان فعلی $p(x)$ و احتمال وجود در مکان بعدی $p(y)$ را بدست می آوریم و مینیمم نسبت $\frac{p(y)}{p(x)}$ و عدد یک را بدست آورده و آن را w می نامیم. حال w را با یک عدد تصادفی دیگر که در بازه ی صفر و یک تولید شده است مقایسه می کنیم. اگر این عدد از w کوچکتر بود x را برابر با y قرار می دهیم. علاوه بر این کار، در یک متغیر تعداد بار هایی که این شرط برقرار شده است را می شماریم. این کار را تکرار می کنیم تا اعداد تصادفی با تابع توزیع گوسی تولید شوند. به نسبت تعداد بار های برقراری شرط گفته شده به تعداد کل تکرار ها نرخ پذیرش می گوییم.
در قسمت بعدی سوال می خواهیم مقدار را بگونه تعیین کنیم که مقادیر مختلف برای نرخ پذیر حاصل شود. این کار را با آزمون و خطا انجام می دهیم به این شکل که برای مقادیر مختلف α الگوریتم را تکرار می کنیم و هر بار نرخ پذیرش را محاسبه می کنیم و اگر به مقدار دل خواهمان نرسیده بودیم، α را تغییر می دهیم و این کار را تکرار می کنیم تا نرخ پذیرش مطلوب حاصل شود. ای مقادیر را برای این مساله در جدول ۴ می بینید.α را به گونه تعیین کنیم که مقادیر

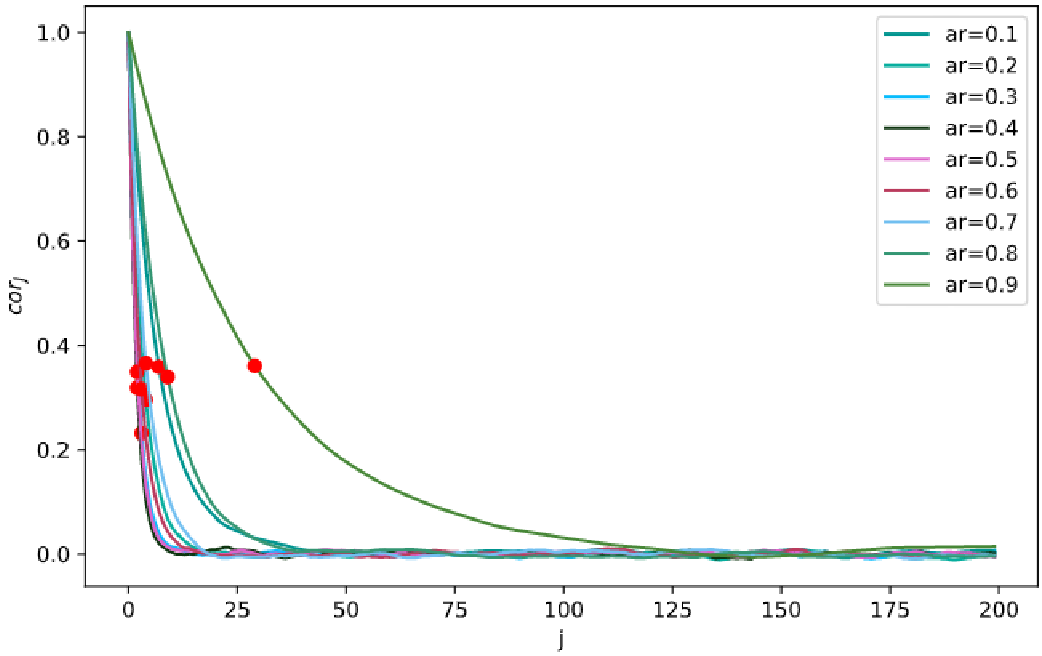
در قسمت بعدی می خواهیم هم بستگی اعداد تولید شده در این رشته را بررسی کنیم. برای این کار این می آییم هم بستگی را بین اعضای مختلف رشته با یکدیگر بررسی می کنیم. این

Table 4: alpha values corresponded to desired accepting ratios

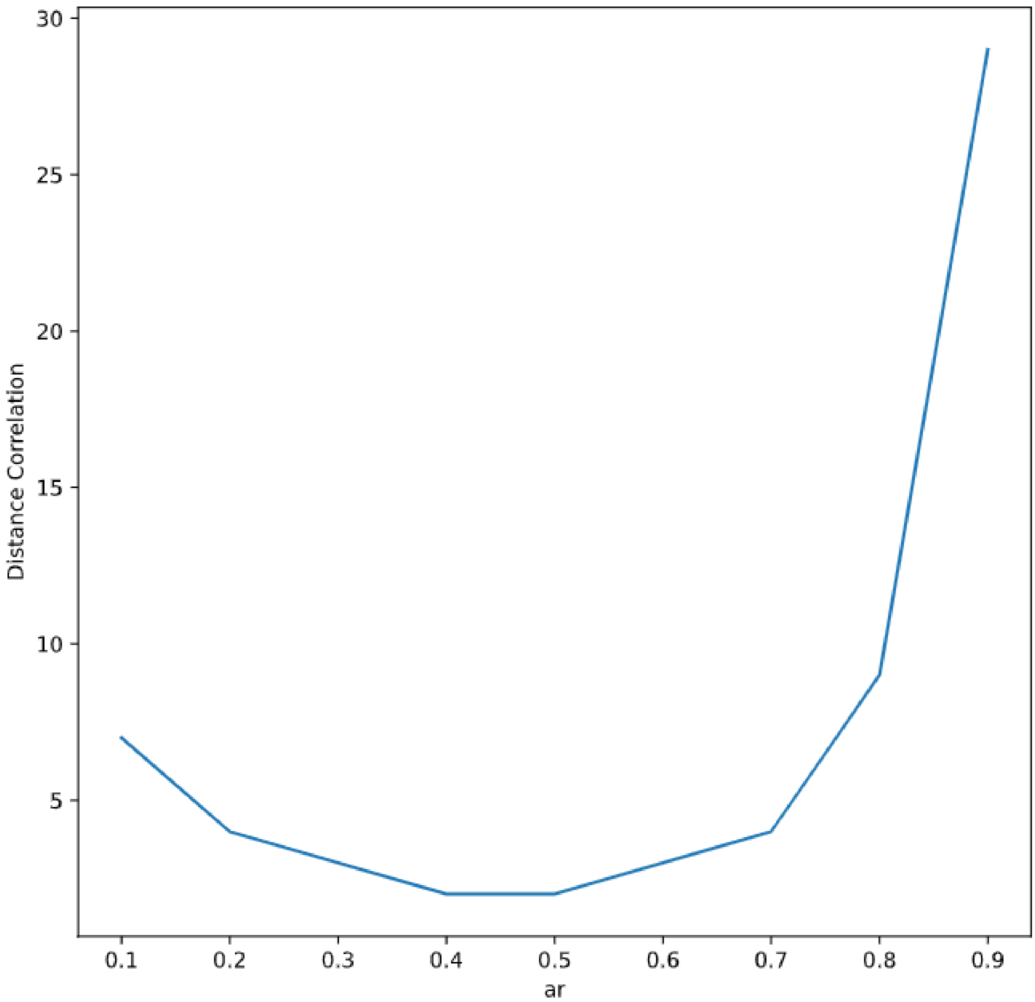
ar	alpha
0.1	6
0.2	3.2
0.3	2.1
0.4	1.55
0.5	1.15
0.6	0.86
0.7	0.63
0.8	0.4
0.9	0.2

کار در یک حلقه انجام می دهیم به این شکل که در هر مرحله از این حلقه رشته را یک عدد به جلو شیفت می دهیم و هم بستگی رشته ی حاصل شده را با رشته ی اصلی بررسی می کنیم. این کار را با تابع numpy corrcoef انجام می دهیم. این تابع دو آرایه را می گیرد و ماتریس کوواریانس را برای آنها بدست می آورد. همانطور که می دانیم مقادیر روی قطر اصلی این ماتریس برابر انحراف از معیار هر کدام از دو رشته ها می باشد و مقادیر روی قطر فرعی حاوی ضریب همبستگی میان این دو رشته می باشند. برای آنکه معیاری برای اینکه چه زمان می توان همبستگی میان اعضای این رشته را نادیده در نظر گرفت، پارامتری به عنوان طول همبستگی تعریف می کنیم. این طول بیانگر آن است که اگر به اندازه ی این طول میان دو عدد این رشته فاصله باشد، همبستگی میان این دوعدد کوچکتر از $\frac{1}{e}$ می شود و تقریبا می توان گفت که اگر تنها اعدادی را از دشته ی اصلی تولید شده نگه داریم که به اندازه ی این طول فاصله دارند، همبستگی میان اهضای رشته تقریبا وجود ندارد و اعداد تولید شده ی خوبی بدست آورده ایم. برای آنکه ظول همبستگی را بدست آوریم، نمودار همبستگی برحسب مقدار شیفت را رسم می کنیم و آنجا که مقدار همبستگی کوچکتر از $\frac{1}{e}$ شده را می یابیم و به عنوان طول همبستگی معرفی می کنیم. در شکل ۱۰ نمودار همبستگی برحسب شیفت را برای مقادیر مختلف نرخ پذیرش می بینید.

در ادامه نمودار همبستگی برحسب نرخ پذیرش را رسم می کنیم. (شکل ۱۱) همانطور که می بینید ای نمودار یک مینیمم در نقطه ای که نرخ پذیرش برابر 0.5 دارد و بنابراین بهترین نقطه برای تعیین αنقطه ای است که این نرخ برابر این مقدار می باشد. چرا که اگر نرخ پذیرش از این مقدار کمتر باشد، بدان معناست که الگوریتم ما تعداد اعداد تکراری زیادی پشت سر هم تولید می کند که طبیعتا مناسب یک رشته ی اعداد تصادفی نیست از طرفی اگر نرخ پذیرش بیشتر از این مقدار باشد، بدان معناست که گام های ما بیش از اندازه کوچک هستند و که سبب می شود تا اعداد تولید شده به یکدیگر نزدیک باشند که بازهم سبب بالا رفتن همبستگی میان اعداد این رشته می شود.



شکل ۱۰: نمودار همبستگی برای نرخ های پذیرش متفاوت بر حسب مقدار شیف رشته تصادفی



شکل ۱۱: نمودار طول همبستگی برحسب نرخ پذیرش