به نام خدا

دانشگاه صنعتی شریف

دانشکده مهندسی برق

گروه دکتر کرباسی ــ آمار و احتمال مهندسی
نیم سال دوم ۱۴۰۱ ــ ۱۴۰۰

تمرین عملی سری سوم

لطفاً به نکات زیر توجه بفرمایید: (رعایت نکردن این موارد باعث کاهش نمره می‌شود.)

۱. نتایج و پاسخ های خود را در یک فایل با فرمت zip به نام HW2_StudentID_Name در سایت Quera قرار دهید. همچنین فایل پایتون خود را به همان نام در قسمت مخصوص به خود آپلود کنید.

۲. کسب نمره کامل در هر سؤال مستلزم تحویل **کدها** و **توضیحات** می‌باشد.

۳. برای سؤالات، باید روشی که استفاده کرده‌اید را توضیح و نتایجی که گرفته‌اید را ارائه دهید. این توضیحات می‌تواند در یک فایل pdf. و یا در یک فایل ipynb. باشد.

۴. فایل های تحویلی شما دو بخش میباشند، یک بخش فایل zip. که شامل فایل ipynb. کد و گزارش شما میباشد، یک بخش هم کد های هر سوال به شکل جداگانه میباشند که باید در فرمت py. در سامانه کوئرا در کنار فایل zip. آپلود شوند.(برای مثال اگر تمرین شامل ۳ سوال بود، باید علاوه بر فایل zip. که تحویل مصحح میشود، ۳ فایل py. در سامانه کوئرا در محل بارگذاری مشخص شده آپلود کنید.)

۵. کدهای خود را خوانا بنویسید و کامنت‌گذاری کنید. در plot های خود عنوان، label و خط‌کشی‌های مناسب را اضافه کنید.

۶. در طول ترم امکان ارسال با تاخیر پاسخ همه‌ی تمارین تا سقف پنج روز و در مجموع دوازده روز وجود دارد. پس از گذشت این مدت، پاسخ‌های ارسال‌شده پذیرفته نخواهند بود. همچنین، به ازای هر روز تأخیر غیر مجاز بیست درصد از نمره تمرین به صورت ساعتی کسر خواهد شد.

۷. کدهای شما تماماً باید توسط خودتان نوشته شده باشند. هرگونه استفاده از کد دیگران به هر شکل ممکن، تقلب محسوب می‌شود و نمره تمرین کامپیوتری جاری صفر خواهد شد. پس در هیچ صورت کدهای خود را برای دیگران ارسال نکنید.

۸. ابهام یا اشکالات خود را مي توانيد از طريق smmzdr@gmail.com يا javadiamirhosein.2000@gmail.com مطرح نمایید.

۹. مهلت تحویل: **نیمه شب جمعه ۲۶ فروردین**

---

۱ تلفنچی نمایی

فرض کنید شما در مرکز مخابرات نشسته‌اید و به زمان‌هایی که یک تماس تلفنی برقرار می‌شود نگاه می‌کنید. با بررسی دقیق‌تر این تماس‌ها می‌بینید که فاصله‌ی بین زمان‌هایی که یک تماس برقرار می‌شود از توزیع نمایی با پارامتر 0.3 پیروی می‌کند و این فاصله‌ها بین تماس‌های مختلف مستقل از هم هستند. چون لحظه‌ی برقراری تماس‌ها تصادفی است و شما از توزیع که روی فاصله ی زمانی بین تماس‌ها وجود دارد آگاهید، سوالی برای شما پیش می‌آید و آن توزیع حاکم بر تعداد تماس‌ها در واحد زمان است. برای این منظور شما به یک شبیه سازی روی می‌آوردید. از متغیر تصادفی $X \sim \text{Exponential}(0.3)$ 1000 نمونه می‌گیرید و فرض می‌کنید این‌ها فاصله‌ی بین تماس‌های مختلف هستند. حال به بازه‌های زمان به طول ۱ نگاه می‌کنید و تعداد تماس‌های گرفته شده در هر بازه ی زمانی به طول واحد را یادداشت می‌کنید. یعنی به باز ه های $[0, 1]$، $[1, 2]$ تا انتها (تا جایی که تماسی گرفته شده باشد) نگاه کرده و در هر بازه تعداد تماس‌ها را خواهید شمرد. سپس هیستوگرام نرمالیزه‌شده تعداد تماس‌ها در بازه‌های به طول واحد را رسم می‌کنید. انتظار دارید این هیستوگرام متناظر با تابع جرم احتمال چه متغیر تصادفی با چه پارامترهایی باشد؟ تابع جرم احتمال آن متغیر تصادفی را همراه این هیستوگرام بر روی یک plot ترسیم کنید.

توابع پیشنهادی: np.random.exponential از کتاب‌خانه‌ی numpy

۲ جمع ماکسیمم ها

متغیر تصادفی زیر را در نظر بگیرید:

$$X_i \sim Bernoulli(p = 0.4)$$

متغیر های تصادفی Z, Y_1, Y_2, Y_3 را به شکل زیر تعریف میکنیم:

$$Y_1 = max(X_1, X_2)$$
$$Y_2 = max(X_2, X_3)$$
$$Y_3 = max(X_3, X_1)$$
$$Z = Y_1 + Y_2 + Y_3$$

به ازای $n = 1 : 10000$ از X_1, X_2, X_3 n نمونه بگیرید و Z, Y_1, Y_2, Y_3 را تشکیل دهید.

۱. امیدریاضی و واریانس Z به ازای n های مختلف را در دو آرایه E و Var ذخیره کنید.

۲. آرایه Cml_avg_E را تعریف کرده و میانگین k عضو ابتدایی E را در عضو kام آن ذخیره کنید.

۳. آرایه Cml_avg_Var را تعریف کرده و میانگین k عضو ابتدایی Var را در عضو kام آن ذخیره کنید.

۴. نمودار Cml_avg_E و Cml_avg_Var رابر حسب n رسم کنید. آیا با افزایش n واریانس و امیدریاضی تجربی به مقدار خاصی میل میکنند؟ این مقادیر را بصورت تئوری نیز حساب کنید و با نتیجه بدست آمده مقایسه کنید.

۵. آرایه های Var_Y1, Var_Y2, Var_Y3 را تعریف کنید و واریانس های Y_1, Y_2, Y_3 را به ازای n های مختلف در آنها ذخیره کنید.

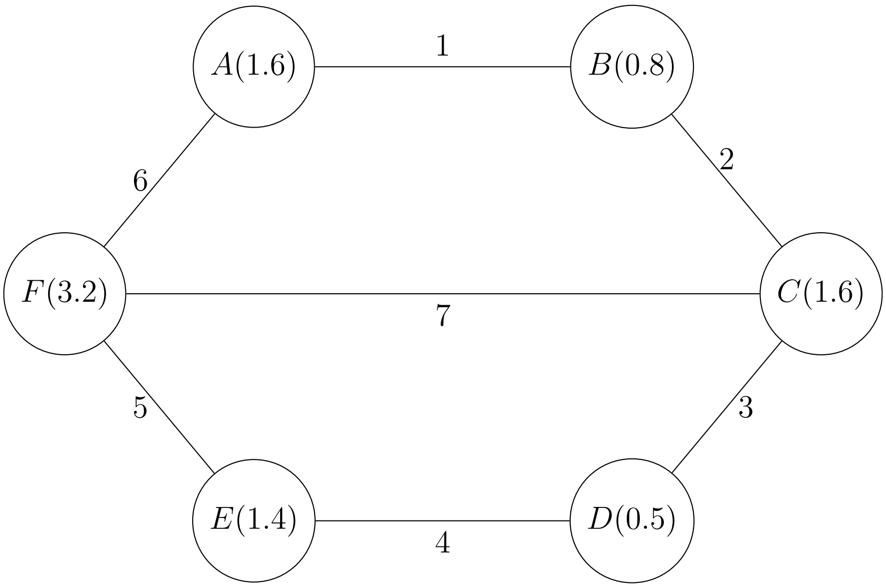
۶. آرایه Var_Y_Sum را جمع سه آرایه ای که در قسمت قبل تعریف کردید قرار دهید.

۷. آرایه Cml_avg_Y_Sum را تعریف کرده و میانگین k عضو ابتدایی Var_Y_Sum را در عضو kام آن ذخیره کنید.

۸. نمودار Cml_avg_Y_Sum را بر حسب n رسم کنید. آیا با افزایش n جمع واریانس های تجربی Y_1, Y_2, Y_3 به مقدار خاصی میل میکند؟ اگر بله این مقدار را با مقداری که در قسمت ۴ بدست آوردید مقایسه کرده و در صورت متفاوت بودن آن را توجیه کنید.

۳ جمع رئوس (امتیازی)

گراف زیر را در نظر بگیرید. هر کدام از رئوس این گراف یک عدد مخصوص دارد و هدف ما پیدا کردن مجموع عددهای رئوس این گراف در هر راس از گراف است. در اینجا هر راس می‌تواند با راس همسایه تبادل اطلاعات کند ولی رئوس حافظه محدودی دارند و می‌توانند فقط تعداد محدودی عدد در خود ذخیره کنند. یال‌ها به دلخواه شماره‌گذاری شده‌اند و عدد هر راس داخل آن در پرانتز نوشته شده‌است.



یک ایده برای تخمین مجموع مقادیر رئوس این است که هر راس نمونه‌ی تصادفی از متغیر تصادفی نمایی با با نرخ عدد آن راس تولید کند و نمونه‌ی خود را با راس همسایه تبادل کند و پس از آن دو راس همسایه نمونه‌ی خود را با مینیمم نمونه‌ی خود و همسایه جایگزین کنند. پس از چند مرحله همه رئوس مینیمم نمونه‌های تصادفی گره‌های مختلف را میبینند و الگوریتم همگرا می‌شوند و گره‌ها میتوانند مقدار مینیمم را ذخیره کنند. حال می‌دانیم مینیمم n متغیر تصادفی نمایی با نرخ‌های λ_i، یک متغیر تصادفی نمایی با نرخ $\sum_{i=1}^{n} \lambda_i$ است. پس اگر تعداد مناسبی این الگوریتم را تکرار کنیم، همه رئوس به چند نمونه iid از متغیر نمایی با پارامتر مجموع مطلوب ما دسترسی دارند و می‌توانند با تخمین Maximum Likelihood پارامتر نرخ متغیر نمایی را تخمین بزنند.

۱. می‌توانید یال‌ها را شماره‌گذاری کنید یا از شماره‌گذاری پیشنهادی استفاده کنید. در هر راند به ترتیب، یالی را انتخاب کنید و بین دو نمونه‌ی رئوس دو سر یال مینیمم‌گیری کنید و در صورت نیاز نمونه‌ی رئوس را به روز رسانی کنید.

۲. حداقل تعداد راند مورد نیاز برای این که مطمئن شویم نمونه‌ی مینیمم به همه‌ی گره‌ها رسیده است چه قدر است؟

۳. الگوریتم بالا را اجرا کنید و مجموع اعداد رئوس را با حداکثر خطای 0.1 به دست بیاورید.