



به نام خدا

دانشگاه صنعتی شریف

دانشکده مهندسی برق

گروه دکتر کرباسی - آمار و احتمال مهندسی

نیم سال دوم ۱۴۰۱ - ۱۴۰۰

تمرین عملی سری دوم

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

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 می باشد.

۳. برای سؤالات، باید روشی که استفاده کرده اید را توضیح و نتایجی که گرفته اید را ارائه دهید. این توضیحات می تواند در یک فایل pdf. و یا در یک فایل ipynb باشد.

۴. فایل های تحویلی شما دو بخش می باشند، یک بخش فایل zip. که شامل فایل ipynb. کد و گزارش شما می باشد، یک بخش هم کد های هر سوال به شکل جداگانه می باشند 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 کامنت گذاری کنید. در plot های خود عنوان، label و خط کشی های مناسب را اضافه کنید.

۶. در طول ترم امکان ارسال با تاخیر پاسخ همه ی تمرین تا سقف پنج روز و در مجموع دوازده روز وجود دارد. پس از گذشت این مدت، پاسخ های ارسال شده 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 می شود 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 می توانید از طریق Smmzdr@gmail.com یا Javadiamirhosein.2000@gmail.com مطرح نمایید.

۹. مهلت تحویل: جمعه ۲۶ فروردین ساعت ۲۳:۵۹

۱ جمع متغیر تصادفی

در این سوال قصد داریم که به جمع متغیرهای تصادفی پردازیم و با نمونه‌برداری از متغیرهای تصادفی هیستوگرام نرمالایز شده آن‌ها را رسم کنیم و شباهت این هیستوگرام نرمالایز شده (تابع جرم احتمال تجربی) با تابع جرم احتمال واقعی متغیرها را بررسی کنیم. دقت کنید که منظور ما از هیستوگرام نرمالایز شده این است که مقادیر داده‌های هر ستون بر تعداد کل داده‌ها تقسیم شده باشند تا مقادیر نرمالیزه شده بین ۰ تا ۱ را به خود اختصاص دهند.

۱. فرض کنید:

$$X_i \sim \text{Bernoulli}(p = 0.4)$$

از متغیرهای تصادفی X_1, X_2, \dots, X_{10} نمونه بگیرید و Y را به شکل زیر تعریف کنید:

$$Y_1 = \sum_{i=1}^{10} X_i$$

نمودار هیستوگرام نرمالایز شده متغیر Y_1 را رسم کنید. انتظار دارید این هیستوگرام متناظر با تابع جرم احتمال چه متغیر تصادفی‌ای با چه پارامترهایی باشد؟ تابع جرم احتمال این متغیر تصادفی را نیز رسم کنید و با تابع جرم احتمال تجربی Y_1 مقایسه کنید.

۲. فرض کنید:

$$Z_i \sim \text{Geometric}(p = 0.6)$$

از متغیرهای تصادفی Z_1, Z_2, Z_3 نمونه بگیرید و Y_2 را به شکل زیر تعریف کنید:

$$Y_2 = \sum_{i=1}^3 Z_i$$

مراحلی که در بخش ۱ انجام دادید را دوباره با Y_2 تکرار کنید.

توابع پیشنهادی: `np.random.geometric`, `scipy.stats.bernoulli` از کتابخانه‌های `numpy`, `scipy`

۲ متغیر تصادفی حدی!

عدد ثابت $\lambda = 5$ را در نظر بگیرید. متغیرهای تصادفی X_i و Y را به شکل زیر تعریف می‌کنیم:

$$X_i \sim \text{Bernoulli}\left(\frac{\lambda}{n}\right) \quad Y = \sum_{i=1}^n X_i$$

۱. به ازای $n = 10, 100, 1000, 10000, 100000$ نمونه از متغیر Y بگیرید و هیستوگرام نرمالایز شده آن را رسم کنیم، با افزایش n تابع جرم احتمال تجربی بدست آمده به تابع جرم احتمال چه متغیر تصادفی‌ای همگرا می‌شود؟ تابع جرم احتمال آن متغیر را نیز رسم کرده و مقایسه کنید.

۲. در قسمت قبل امیدریاضی و واریانس Y با افزایش n به چه مقادیری همگرا می‌شوند؟ (از نظر تئوری بررسی کنید)

توابع پیشنهادی: `scipy.stats.bernoulli` از کتابخانه‌ی `scipy`

۳ امید ریاضی و واریانس مجموع

در این بخش چند متغیر تصادفی را تعریف میکنیم سپس به بررسی امیدریاضی و واریانس مجموع آنها می‌پردازیم.

۱. فرض کنید

$$X_1 \sim \text{Binomial}(n = 5, p = 0.4)$$

$$X_2 \sim \text{Poisson}(\lambda = 1.6)$$

$$X_3 \sim \text{Geometric}(p = 0.1)$$

از متغیر تصادفی‌های X_1, X_2, X_3 ، $N = 10000$ نمونه بردارید و Y را به شکل زیر تعریف کنید.

$$Y = X_1 + X_2 + X_3$$

حال هیستوگرام Y را رسم کنید. همچنین میانگین و واریانس Y را به دست بیاورید و ارتباط این مقادیر را با میانگین و واریانس X_1, X_2, X_3 بررسی کنید.

۲. متغیر تصادفی

$$X_i \sim \text{Bernoulli}(p = 0.8)$$

را در نظر بگیرید. متغیر تصادفی Y را به شکل زیر تعریف میکنیم:

$$Y = \frac{1}{n} \sum_{i=1}^n X_i$$

به ازای سه حالت $N = 5, 40, 300$ از متغیر تصادفی Y ۱۰۰۰ نمونه تولید کنید و سه هیستوگرام بدست آمده را جداگانه رسم کنید. با افزایش n امیدریاضی و واریانس Y چگونه تغییر میکنند؟ چرا؟

توابع پیشنهادی: `np.random.binomial`, `np.random.poisson`, `np.random.geometric` از کتابخانه‌ی `numpy`

۴ پارادوکس روز تولد

احسان قصد دارد مهمانی ای برگزار کند و n نفر را نیز به این مهمانی دعوت خواهد کرد. احسان یک بازی تدارک دیده است. او از هر مهمان می‌خواهد که روز تولد خود را روی کاغذ بنویسد و به او بدهد. اگر حداقل دو نفر از مهمان‌ها متولد یک روز باشند احسان بازی را میبرد و مهمان‌ها باید هزینه تور اروپای احسان را بدهند. در غیر این صورت احسان باید هزینه سفر همه مهمان‌ها به زنوز را پیاده شود. از آنجایی که احسان باید به همه مهمان‌ها شام بدهد نمی‌خواهد تعداد مهمان‌ها را زیاد در نظر بگیرد تا در صورت رفتن به اروپا بیشتر خرید کند. برای همین میخواهیم به او کمک کنیم تا با کمترین تعداد مهمان با احتمال بالایی بازی را ببرد و رایگان به اروپا برود.

۱. 10000 آزمایش انجام دهید. در هر آزمایش قصد داریم تخمینی از n به دست بیاوریم، به این صورت که شروع به انتخاب تصادفی روزهای سال می‌کنیم و زمانی که روزی را انتخاب کردیم که قبلاً هم انتخاب شده بود تعداد انتخاب‌هایی که تا آن مرحله انجام داده بودیم را ذخیره میکنیم. سپس هیستوگرامی از تعداد انتخاب‌های این 10000 آزمایش رسم می‌کنیم.

۲. حال احتمال پیروزی احسان به ازای n مهمان را برای $1 \sim 365$ n را به دست بیاورید و رسم کنید. اگر احسان بخواهد حداقل با احتمال $\frac{1}{2}$ برنده شود باید چند مهمان دعوت کند؟