

بنام خدا

دانشکده‌ی مهندسی برق و کامپیوتر

درس آمار و احتمال مهندسی

تمرین کامپیوتری 2

مهلت تحویل: 24 دی

استاد: دکتر ربیعی

سر فصل مطالب

1 سوال 1: Basic Probability Theory

1 Discrete Uniform Distribution

1 Binomial Distribution

3 Poisson Distribution

3 Normal Distribution

4 سوال 2: Intro to SciPy

4 Binomial Distribution

4 Poisson Distribution

5 Normal Distribution

6 سوال 3: Birthday Problem

7 سوال 4: Central Limit Theorem

8 نکات تحویل

سوال 1 :

هدف این سوال بررسی توابع احتمال به کمک کتابخانه NumPy می باشد .

Discrete Uniform Distribution

1. شبیه سازی پرتاب تاس:

در این قسمت می خواهیم مساله پرتاب تاس را شبیه سازی کنیم . در ابتدا به کمک آموزشهای داده شده ، تاس را به صورت یک عدد رندوم بین 1 تا 7 ایجاد کنید . حال تابعی بنویسید که به ازای ورودی تعداد دفعات آزمایش (n) ، نتایج حاصل از پرتاب n بار تاس را در یک آرایه ذخیره کند .

(الف)

برای $n \in [100, 1000, 10000]$ ، Count plot مربوط به خروجی تابع نوشته شده در بالا را به صورت subplots رسم کنید .

(ب)

با افزایش n ، چه تغییری در نمودار ها ایجاد می شود ؟ آیا نتایج با آنچه که انتظار می رود همخوانی دارد ؟

Binomial Distribution

1. شبیه سازی پرتاب سکه سالم (Unbiased Coin):

در این بخش قصد داریم که مساله پرتاب سکه سالم (احتمال وقوع شیر یا خط برابر است) را به کمک تابع احتمال Binomial پیاده سازی کنیم .

(الف)

در مورد پارامتر های این تابع احتمال در [این لینک](#) تحقیق کنید .

(ب)

حال تابعی بنویسید که با گرفتن دو ورودی تعداد بار پرتاب سکه (n) و تعداد دفعات آزمایش (N) ، تعداد دفعات شیر / خط آمدن را برگرداند .

(ج)

برای $n = [2, 10]$ ، $N \in [100, 1000, 10000]$ ، Count plot مربوط به خروجی تابع نوشته

شده در بالا را به صورت subplots رسم کنید . (6 نمودار در مجموع)

(د)

خروجی بخش ج را با مقادیر بدست آمده از تئوری بررسی کنید. بیشترین احتمال بدست آمده ، مربوط به کدام تعداد شیر / خط در خروجی می باشد ؟

همچنین تاثیر تکرار آزمایش در همگرایی بخش شبیه سازی به مقدار تئوری را بررسی کنید.

2. شبیه سازی پرتاب سکه خراب (Biased Coin):

(الف)

همانند قسمت ب بخش قبل تابع مورد نظر را ایجاد کنید با این تفاوت که احتمال شیر آمدن در هر بار پرتاب سکه عدد p می باشد ($0.5 < p < 1$).

(ب)

مراحل قسمت های ج و د بخش قبل را برای $N = 10000, n = [2,10], p = [0.6,0.8]$ تکرار کنید .

3. تصمیم تصادفی با سکه خراب :

فرض کنید که سکه خرابی در اختیار دارید که با احتمال p ($0.5 < p < 1$) در هر بار شیر می آید .

(الف)

به کمک این سکه ، آزمایشی را طراحی کنید که بتوانید بین دو نفر ، یکی را به صورت تصادفی (احتمال 0.5) انتخاب کنید . (*راهنمایی : در آزمایش شما باید یک شرط خاص وجود داشته باشد که اگر تحقق یافت ، متوقف شوید)

(ب)

حال می خواهیم در قالب تئوری / شبیه سازی بررسی کنیم که به صورت میانگین ، چه تعداد باید سکه بیندازیم تا به شرط مطلوب برسیم .

1. به صورت میانگین بیان کنید که چه تعداد باید سکه بیندازیم تا شرط مطلوب محقق شود (*)

راهنمایی : از Expectation (امید ریاضی) استفاده کنید .

2. ابتدا تابعی بنویسید که به کمک آن آزمایش طراحی شده را پیاده سازی کنید . حال برای 10000

بار تکرار آزمایش ، میانگین تعداد انداختن سکه برای آنکه شرط مطلوب محقق شود را بدست آورید .

3. مقدار تئوری بدست آمده را با مقدار بدست آمده از شبیه سازی مقایسه کنید .

Poisson Distribution:

فرض کنید در یک صف مربوط به خرید بلیط سینما، در هر 10 دقیقه به طور متوسط 3.6 نفر قرار دارد و فرض کنید توزیع فوق از فرم Poisson تبعیت می کند.

(الف)

در مورد پارامتر های این تابع احتمال در [این لینک](#) تحقیق کنید .

(ب)

حال به کمک قسمت قبل، خروجی بدست آمده از تابع Poisson را برای 10000 داده در یک count plot رسم کنید. (*دقت کنید که محور x شما، تعداد افراد در صف (k) در هر 10 دقیقه می باشد).

(ج)

بیشترین تعداد تکرار مربوط به کدام k می باشد؟ احتمال مربوط به آن را به صورت تئوری نیز محاسبه کنید. از کدام k به بعد، احتمال وقوع بسیار ناچیز می شود؟

Normal Distribution:

(الف)

در مورد پارامتر های این تابع احتمال در [این لینک](#) تحقیق کنید .

(ب)

حال به کمک قسمت قبل، توزیع نرمالی با میانگین 100 و انحراف معیار 2 و برای 1000000 داده ایجاد کنید و سپس hist plot مربوط به آنرا بکشید و پارامتر kde را نیز روی True قرار دهید .

(ج)

به کمک کتابخانه NumPy، خواسته های زیر را گزارش دهید :

1. چند درصد دادگان در بازه $[-\sigma, \sigma]$ قرار گرفته اند.
2. چند درصد دادگان در بازه $[-2\sigma, 2\sigma]$ قرار گرفته اند.
3. چند درصد دادگان در بازه $[-3\sigma, 3\sigma]$ قرار گرفته اند.

سوال 2 :

در این قسمت قرار است به کمک کتابخانه SciPy، تحلیل های آماری دقیق تری ارائه دهیم (برای آشنایی با کتابخانه SciPy، قسمت سوم آموزش پایتون را نگاه کنید).

Binomial Distribution

فرض کنید محصول خروجی یک کارخانه تولید قطعات الکترونیکی به احتمال 12 درصد دچار نقص فنی (Defect Rate) می باشد . شما به عنوان یک بازرس تصمیم می گیرید تا 20 قطعه خروجی کارخانه را به صورت تصادفی انتخاب کنید و در صورتی که حداکثر 2 قطعه دارای نقص فنی باشد ، مجوز فروش را صادر کنید.

(الف)

احتمال صدور/عدم صدور مجوز را هم به صورت تئوری و هم به کمک کتابخانه SciPy محاسبه کنید .

(ب)

احتمال اینکه دقیقا 2 قطعه دچار نقص فنی باشد را هم به صورت تئوری و هم به کمک کتابخانه SciPy محاسبه کنید.

(ج)

نمودار احتمالاتی pmf(Probability mass function) را به صورت bar plot رسم کنید.

(د)

نمودار احتمالاتی cdf(Cumulative distribution function) را به صورت bar plot رسم کنید.

:Poisson Distribution

مساله مطرح شده مربوط به توزیع Poisson سوال 1 را در نظر بگیرید :

(الف)

میانگین و واریانس توزیع را با کمک کتابخانه SciPy بدست آورده و با مقداری تئوری مقایسه کنید.

(ب)

احتمال اینکه در بازه زمانی داده شده ، 7 نفر در صف قرار داشته باشند را با کمک کتابخانه SciPy و تئوری محاسبه کنید.

(ج)

نمودار pmf توزیع را برای تعداد صفر تا 10 نفر در صف به صورت bar plot رسم کنید .

(د)

نمودار cdf توزیع را برای تعداد صفر تا 10 نفر در صف به صورت bar plot رسم کنید .

:Normal Distribution

در یک کارخانه ، بطری های ادکلن با میانگین حجم 150 cc و انحراف معیار 2 cc پر می شوند.

(الف)

به کمک کتابخانه SciPy ، تعیین کنید چه درصدی از بطری های تولید شده توسط این کارخانه ، حجمی بیش تر از 153 cc دارند ؟

درصد بدست آمده را با مقدار تئوری مقایسه کنید.

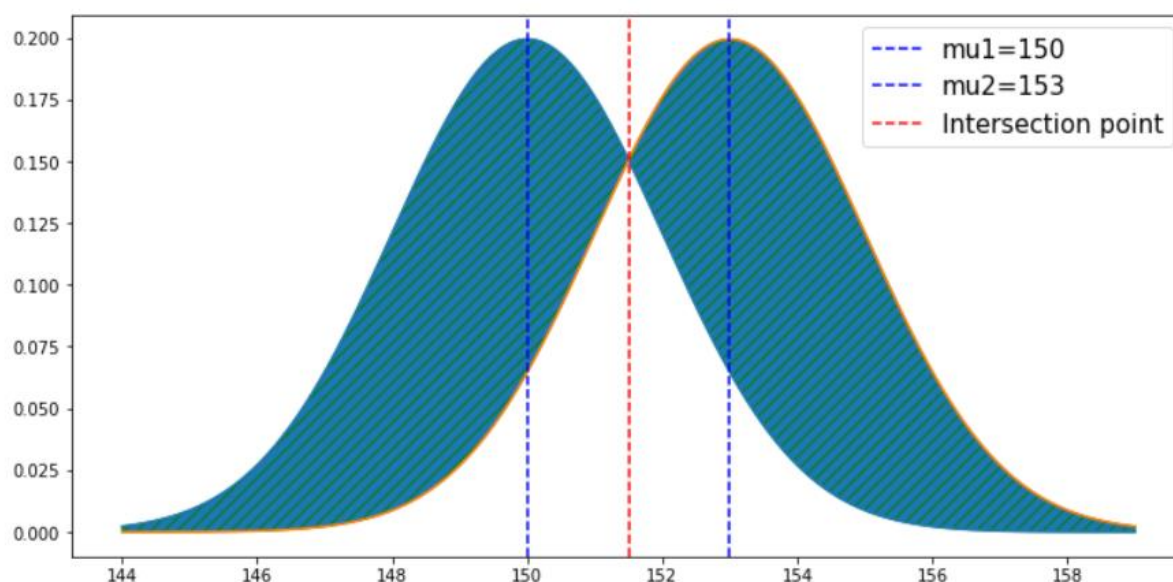
(ب)

به کمک کتابخانه SciPy ، تعیین کنید چه درصدی از بطری های تولید شده توسط این کارخانه ، حجمی بین 148 cc و 152 cc دارند ؟

درصد بدست آمده را با مقدار تئوری مقایسه کنید.

(ج)

مساحت هاشور خورده در شکل زیر را به کمک کتابخانه SciPy محاسبه کنید (انحراف معیار هر دو توزیع داده شده برابر 2 می باشد) .



شکل 1-3-3-2 : pdf مربوط به دو توزیع احتمالاتی

(د)

به کمک دستور [lineplot](#) در کتابخانه pdf, Seaborn مربوط به حجم بطری های ادکلن را برای بازه [144,156] و تعداد 1500 داده رسم کنید .

سوال 3 :

در این سوال قصد داریم مساله معروف تولد را شبیه سازی کنیم :

فرض کنید در یک جمع n نفره از دوستان خود قرار دارید . می خواهید از طریق شبیه سازی تعیین کنید که به ازای هر n ، احتمال اینکه حداقل دو نفر در جمع دارای تاریخ تولد یکسانی باشند را بدست آورید.

(الف)

برای شبیه سازی تاریخ تولد هر نفر ، ابتدا از [datetime](#) کتابخانه ، اولین روز سال 2022 را به عنوان مبدا زمانی در نظر بگیرید . سپس به کمک [timedelta](#) ، یک روز رندوم بین 0 تا 365 ایجاد کنید .

(* به کمک `datetime` مبدا زمانی را مشخص کنید و به کمک `timedelta` مدت زمانی (duration) در نهایت تابعی بنویسید که با گرفتن ورودی تعداد افراد (n) ، به تعداد n تاریخ تولد رندوم برگرداند .

(ب)

حال به کمک تابع نوشته شده در بخش قبل ، تابع جدیدی بنویسید که با گرفتن تعداد افراد (n) و تعداد دفعات آزمایش (N) ، به طور متوسط احتمال اینکه حداقل دو نفر دارای تاریخ تولد یکسانی باشند را برگرداند .
(* برای متوسط گیری ، تعداد کل بار هایی که حداقل دو نفر دارای تاریخ تولد یکسان هستند را تقسیم بر تعداد آزمایش می کنید)

(ج)

در نهایت به ازای تعداد افراد از 0 تا 100 ، مقادیر احتمالی بدست آمده را در یک نمودار رسم کنید.

(* حدود زمان لازم برای این شبیه سازی 3 دقیقه می باشد . صبور باشید !)

(د)

1. حداقل تعداد افرادی که لازم است در یک جمع حضور داشته باشند تا احتمال اینکه حداقل دو نفر دارای تولد یکسان باشند ، تقریباً برابر 0.5 باشد را از روی نمودار قسمت قبل اعلام کنید.
2. حداقل تعداد افرادی که لازم است در یک جمع حضور داشته باشند تا احتمال اینکه حداقل دو نفر دارای تولد یکسان باشند ، تقریباً برابر 1 باشد را از روی نمودار قسمت قبل اعلام کنید.

(ه)

قسمت 2 بخش د را با اصل لانه کبوتری پاسخ دهید . تناقض بوجود آمده را شرح دهید .

سوال 4 :

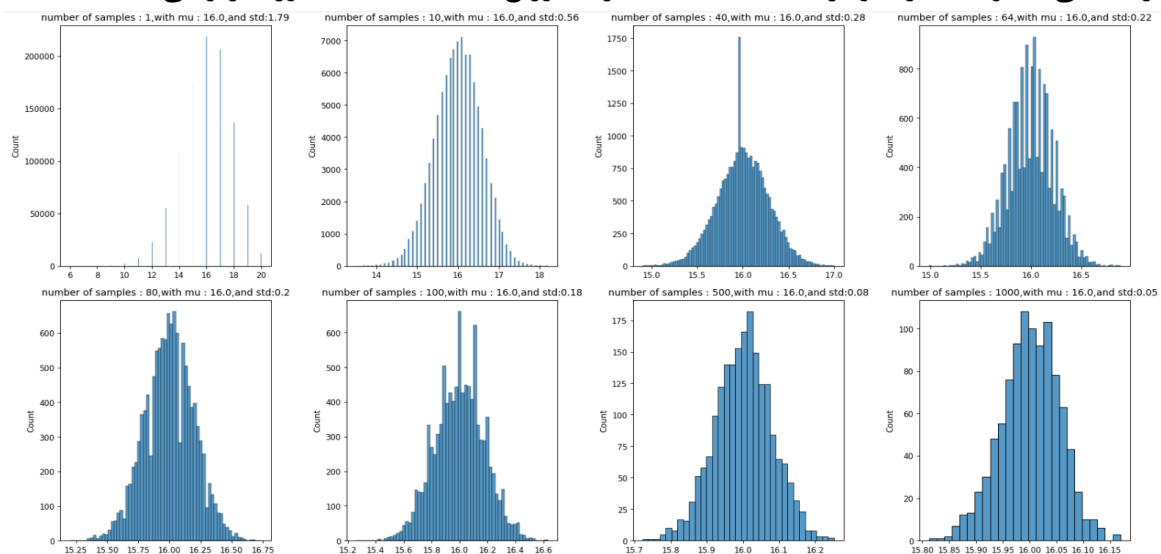
در این سوال به بررسی قضیه حد مرکزی (Central Limit Theorem) می پردازیم .

فرض کنید یک توزیع احتمالاتی دلخواه از یک جمعیت (Population) در اختیار دارید . در هر بار تعدادی مشاهده (Sample) (پارامتر مجهول) از جمعیت داده شده برداشت می کنید و میانگین آنها را محاسبه می کنید و به عنوان یک داده جدید آنرا لحاظ می کنید . این کار را تا جایی ادامه می دهید که از تمامی جمعیت اولیه نمونه برداری کرده باشید . (*مشاهدات برداشتی در هر بار با یکدیگر همپوشانی ندارند) .

(الف)

مراحل ذکر شده در بالا را برای سه توزیع نمایی (با پارامتر $\lambda = 2$) ، توزیع Poisson (با پارامتر $\mu = 3.5$) و توزیع Binomial (با پارامتر $p = 0.7, n = 20$) تکرار کنید . سپس برای هر یک از سه توزیع بالا ، هستوگرام داده های جدید بدست آمده را برحسب تعداد مشاهده برداشتی از جمعیت (n) که $n \in [1, 10, 40, 64, 80, 100, 500, 1000]$ می باشد را به صورت subplot رسم کنید .

*راهنمایی : نمونه ای از نمودار های بدست آمده برای توزیع binomial به صورت زیر می باشد :



شکل 1-4-1 subplot بدست آمده برای توزیع binomial

(ب)

نمودار های بدست آمده برای $n > 1$ با نمودار $n = 1$ مقایسه کنید . آیا از توزیع خاصی تبعیت می کنند ؟

با افزایش n ، چه در تغییری در نمودار های بدست آمده برای هر توزیع مشاهده می کنید ؟

(ج)

در ادامه برای هر کدام از سه توزیع بالا ، نحوه تغییر میانگین و واریانس دادگان جدید را بر حسب n رسم کنید . رابطه ای بین میانگین و واریانس توزیع جدید بدست آمده با توزیع پیشین بر حسب n ارائه دهید .

نکات تحویل :

1. کدهای نهایی تحویلی هر سوال را در یک فایل **ipynb** در نهایت قرار دهید (هر بخش و زیر بخش ها را با فرمت Markdown از هم جدا کنید) . (در نهایت 4 فایل **ipynb** آماده برای آپلود خواهید داشت)
2. تمامی شکل های خروجی خواسته شده در هر زیر بخش را با زیرنویس مربوط به آن زیربخش (به شکل های در صورت پروژه دقت کنید) مشخص کرده و در گزارش خود قرار دهید. همچنین در هر زیر بخش ، متناسب با مقدار خواسته شده توضیح و پاسخ دهید . در نهایت گزارش و کد های خود را به در قالب فایل zip و به فرمت **CA_num-Last_name-std_num** در صفحه درس آپلود کنید.
3. هدف از تمرین های کامپیوتری کمک به یادگیری شماست. بنابراین در صورت مشابهت بیش از حد در بخش های پروژه ، از شما نمره کم خواهد شد .
4. در صورتی که نسبت به پروژه سوال یا ابهامی داشتید ، از طریق ایمیل sh.vassef@ut.ac.ir یا در گروه تلگرامی با من در ارتباط باشید.

موفق باشید .