بنام خدا

گزارش تمرین کامپیوتری سوم

آرشام لولوهری

۹۹۱۰۲۱۵۶

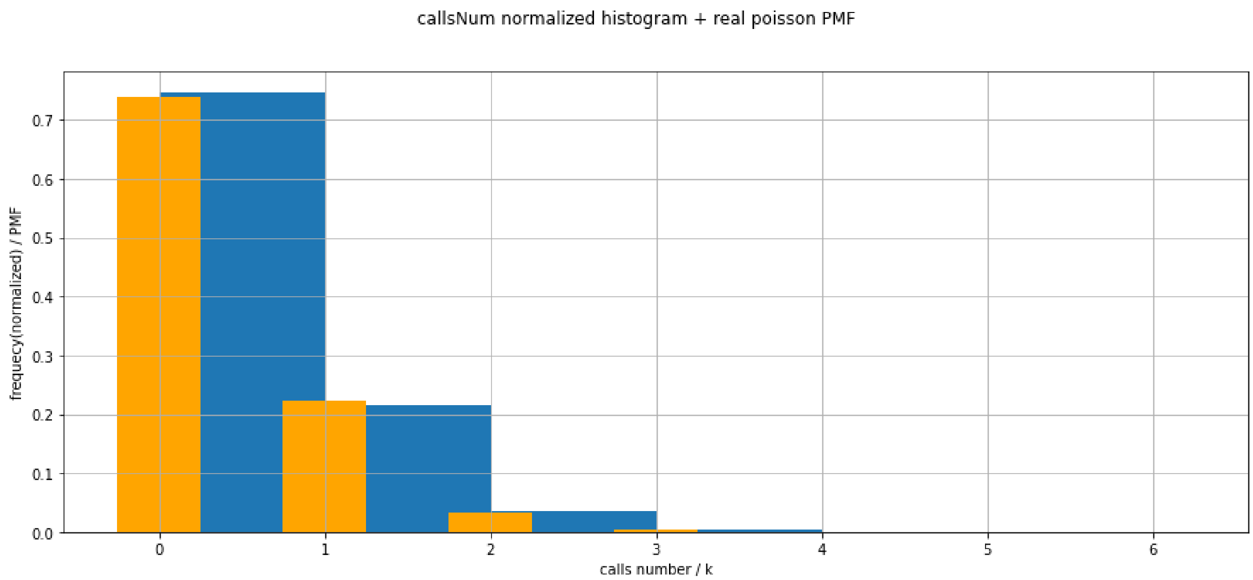
۱:

ابتدا پارامتر لاندای توزیع نمایی را ۰,۳ تعریف کردیم و تعداد کل نمونه ها را نیز در total ریخته ایم. تابع exponential پایتون، پارامتر ورودی را به صورت β میگیرد که معکوس لاندا است. تعداد نمونه ها نیز در ادامه به تابع داده شده است و بردار نمونه ها را در X ریخته ایم. lastCallTime زمان آخرین تماس از نمونه هاست. توجه داریم زمان تماس k ام، برابر با مجموع فواصل بین دو تماس، تا نمونه ی k ام است. به همین منظور تابع callTime ساخته شده که بردار نمونه ها، و تماس شماره ی index را میگیرد و زمان رخ دادن آن تماس را به خروجی میدهد. آخرین ثانیه (واحد زمان) ای که باید در فرایند ما تحلیل شود، کران بالای lastCallTime است (مثلا اگر آخرین تماس در زمان 2.6 رخ داده باشد، باید تا زمان ۳ تحلیل کنیم). بردار تعداد تماس ها در callsNum ساخته خواهد شد. المان صفرم این بردار، تعداد تماس ها در بازه ی صفر تا یک است، و به همین ترتیب برای بقیه المان ها. سپس در بردار callsTImes و با استفاده از تابع callTime، زمان رخ دادن تمام ۱۰۰۰ تماس را میریزیم. در حلقه for اول، تعیین میکنیم که تحلیل را برای بازه های زمانی از صفر تا آخرین ثانیه انجام میدهیم. در هر بازه زمانی، یک for بین تمام المان های X زده میشود تا ببینیم هر تماس در بازه مورد نظر قرار دارد یا نه. در newSum، زمان تماس j ام را توسط callsTimes میریزیم. اگر این زمان در بازه ی i تا i+1 قرار داشت (که در آخرین بازه

زمانی، میتواند برابر با i+1 نیز باشد)، مقدار callsNum در اندیس i ام خود، یک واحد زیاد میشود. بدین معنا که تعداد تماس های ما در آن بازه، یک واحد زیاد میشود. پس از این دو حلقه هم نمودار را رسم کرده ایم. ابتدا تعیین سایز نمودار و عنوانش، و سپسدستور hist. ابتدا یک maxNum تعریف میکنیم که ماکزیمم تعداد تماس های حاصله در یک بازه زمانی را نشان میدهد. این maxNum ، آخرین مقداری است که در محور افقی باید نشان داده شود. دسته ها را به صول یک واحد و از صفر تا maxNum میگیریم، طول هر دسته یک است و با density، نمودار را نرمالیزه میکنیم.

اگر فاصله بین رخ دادن های متوالی یک اتفاق، توزیع نمایی با پارامتر لاندا داشته باشد، همانطور که در کتاب نیز آمده، تعداد کل رخ دادن های آن اتفاق در واحد زمان، از توزیع پواسون با همان پارامتر لاندا تبعیت میکند. پس انتظار ما این است که در تعداد نمونه های زیاد، نمودار هیستوگرام رسم شده مشابه توزیع پواسون با پارامتر (میانگین) ۰,۳ شود.

با استفاده از poisson از scipy.stats و دستور pmf آن ، برای محدوده ی r از صفر تا ۶، PMF را در همان نمودار و با رنگ نارنجی رسم کرده ایم و میبینیم شباهت بسیاری بین دو نمودار وجود دارد (البته نمودار آبی، چون بیانگر فراوانی در طول بازه ی واحد است، اندکی نسبت به نمودار نارنجی رنگ شیفت خورده است تا فراوانی را در تمام طول بازه نشان دهد، ولی مقادیر متناظر دو نمودار، بسیار نزدیک اند)



:۲

۱– متغیر های Xi را با استفاده از scipy.stats.bernoulli تعریف کرده و n نمونه (که n از ۱ تا ۱۰۰۰۰ تغییر میکند) از آنها میگیریم. با استفاده از np.maximum، ماکزیمم المان به المان بردار ها را مطابق روابط Yi ها بدست آورده و Z را تعریف میکنیم. بدین ترتیب به سادگی و با دستورات np.mean,np.var میتوان میانگین و واریانس Z را در هر n محاسبه و ذخیره کرد.

۲و۳– توابع همانطور که گفته شده تعریف شده اند و با یک حلقه فور، برای عضو k ام آنها، میانگین k عضو اولِ (به ترتیب) E و Var قرار داده شده است.

۴– با استفاده از plt.bar و برای محدوده r از ۱ تا ۱۰۰۰۰ (n های مختلف)، هردو نمودار رسم شده است. نمودار میانگین(سمت چپ) به عددی در حدود

1.92 و نمودار واریانس به حدود 1.2 میل میکند. علت هردو در زیر به صورت تئوری نشان داده شده است. ضمنا توجه داریم به ازای n های کوچکتر، ممکن است بردار های E,Var ، با میانگین و واریانس واقعی Z فاصله محسوسی داشته باشند اما در بردار های بدست آمده در بخش های ۲ و ۳، ما میانگین این مقادیر را ریخته ایم. هرچه n بزرگتر شود، E,Var به مقدار تئوری نزدیکتر میشوند و در نتیجه با میانگین گیری از آنها نیز به EZ,Var(Z) نزدیک میشویم. پس انتظار میرود مقادیر حدودی 1.92 و 1.2 ، حدودا میانگین و واریانس Z باشند:

۲-۱۴

$$EZ = EY_1 + EY_2 + EY_3 \overset{\text{بدلیل تقارن}}{=\!=} 3EY_1$$

$$EY_1 = \sum_{y=0}^{1} y P_{Y_1}(y) = P_{Y_1}(1) = 1 - P(X_1 = 0 \text{ و } X_2 = 0)$$

$$p = 0.4$$

$$= 1 - p^2 = 1 - 0.36 = 0.64 \rightarrow EZ = 3 \times 0.64 = 1.92$$

$$EZ^2 = \sum_{z=0}^{3} z P_Z(z) = P_Z(1) + 4P_Z(2) + 9P_Z(3)$$

$P_Z(1) = 0$ (چون اگر یکی از X_iها ۱ شود، Z حداقل برابر با ۲ هست)

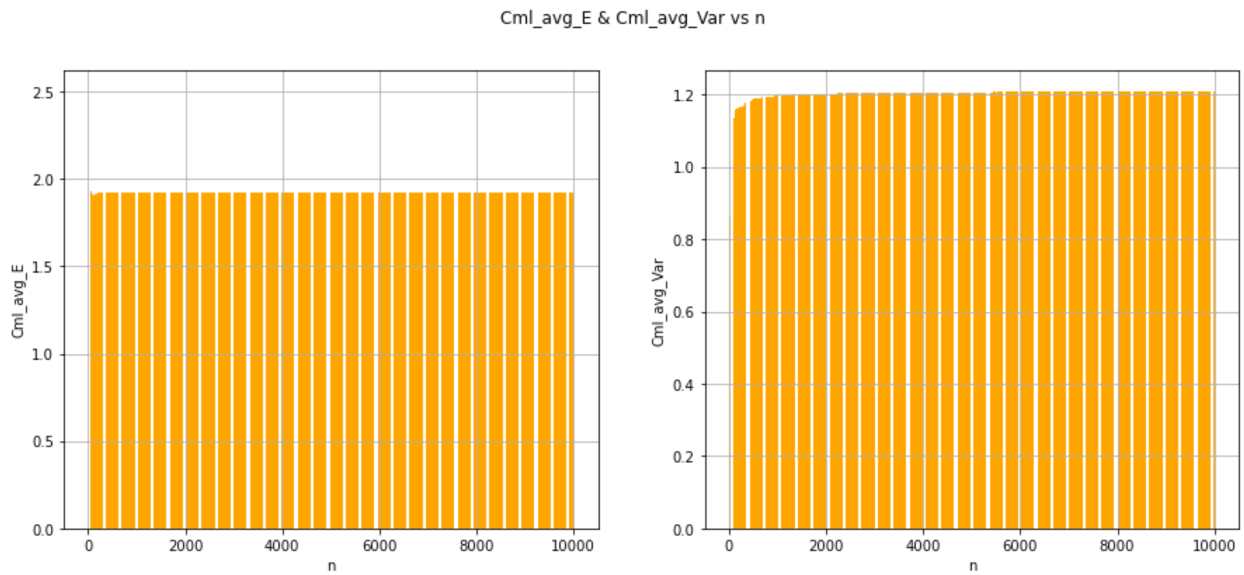
$P_Z(2) = 3 \times p(1-p)^2 \leftarrow$ احتمال اینکه دقیقاً دوتا از X_iها صفر باشد

$$P_Z(3) = \underbrace{3 \times p^2(1-p)}_{\text{دقیقاً دوتا از } X_i \text{ها یک باشند}} + \underbrace{p^3}_{\text{هر سه } X_1 \text{ و } X_2 \text{ و } X_3 \text{، یک باشند}}$$

جایگذاری $\Rightarrow EZ^2 = 4.896$

$$Var(Z) = EZ^2 - (EZ)^2 = 4.896 - (1.92)^2 = 1.2096$$

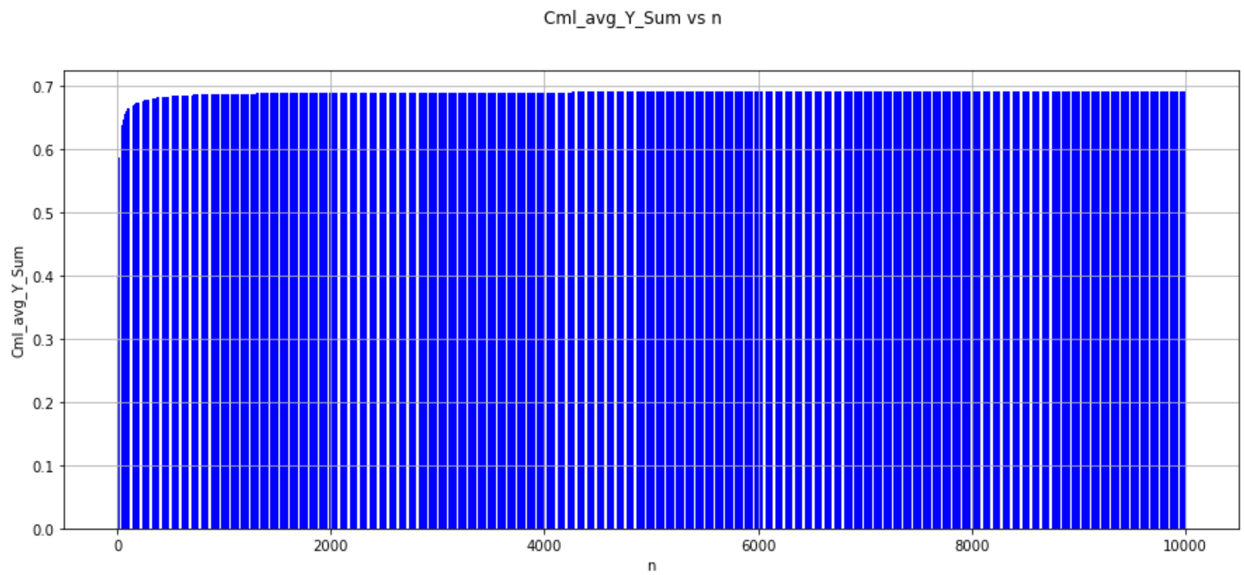
در ادامه هم نمودار های مربوطه آمده است:



۵- این کار در حلقه فور قسمت های قبل، انجام شده است.

۶و۷- این دو کار، همراه با بخش ۲ و ۳ و در همان حلقه فور انجام شده است(محل تمام قسمت های سوال با کامنت مشخص شده است). این کار برای جلوگیری از زدن حلقه های اضافی، و تسریع اجرای کد انجام شده است.

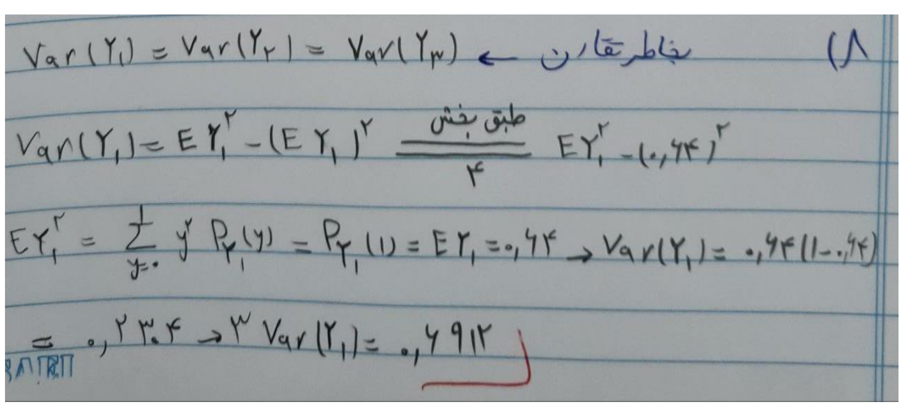
۸- این قسمت بعد از کد مربوط به رسم نمودار های بخش ۴ نوشته شده است. طرز رسم نمودار مشابه قبل است و حال، نمودار سوم به صورت زیر است:



میبینیم که باز هم همگرایی داریم اما مقدار نهایی با آنچه در Cml_avg_Var داشتیم فرق دارد و در حدود 0.69 است. علت تفاوت این است که شرط اینکه داشته باشیم:

Var(Y1+Y2+Y3)=Var(Y1)+Var(Y2)+Var(Y3)

این است که Yi ها مستقل باشند اما در اینجا به وضوح نیستند و کاملا به هم وابسته اند. در ادامه مقدار Var(Y1)+Var(Y2)+Var(Y3) محاسبه شده است:



(۸

$Var(Y_1) = Var(Y_2) = Var(Y_3) \leftarrow$ بخاطر تقارن

$Var(Y_1) = EY_1^2 - (EY_1)^2 \xrightarrow[\text{۴}]{\text{طبق بخش}} EY_1^2 - (۰,۶۴)^2$

$EY_1^2 = \sum_{y=0}^{1} y^2 P_{Y_1}(y) = P_{Y_1}(1) = EY_1 = ۰,۶۴ \rightarrow Var(Y_1) = ۰,۶۴(1-۰,۶۴)$

$= ۰,۲۳۰۴ \rightarrow ۳\, Var(Y_1) = ۰,۶۹۱۲$