

## سوال 1:

هدف از این کد مقایسه مقادیر تئوری و عملی میانگی و واریانس برای توزیع دوجمله ای توسط نمونه های تصادفی است.

مقایسه نمودارهای واریانس ها و میانگین های نظری و عملی به ارزیابی دقت شبیه سازی توزیع دوجمله ای کمک می کند.

همانطور که مشخص است، نمودارها نشان میدهند که مقادیر واریانس عملی و نظری که توسط نمونه های تصادفی بدست می آید، بسیار نزدیک به هم هستند و همینطور برای میانگین ها.

هرچه قدر تعداد نمونه ها بیسشتر باشد، مقدار عملی به تئوری بیسشتر میشود.

برای رسم نمودار از کتابخانه `matplotlib.pyplot` استفاده می کنیم. در ابتدا یک بازه برای رسم نمودار استفاده می کنیم که همین در واریانس و میانگین تئوری هم استفاده میشود. مقادیر تئوری طبق فرمول های اثبات شده بدست می آیند. برای مثال میانگین یک متغیر تصادفی `Bionomial` از رابطه  $n \cdot p$  بدست می آید. برای رسم نمودارها در ابتدا یک `figure` جدید تعریف میکنیم که میتوان سایز نمودارها را توسط آرگومان این تابع بدست آورد. آرگومان این تابع `plt.figure` است که دو پارامتر به عنوان طول و عرض که به `float` است میتوان به آن داد و تعریف کرد. بعد از میتوان از `plt.subplot` استفاده میکنیم که سه تا ورودی میگیرد که تعداد سطرها، ستونها و ایندکس برای رسم نمودار را میگیرد. از `plt.plot` میتوان برای رسم نمودار استفاده کرد به اینگونه که در ورودی اول محور `x` و در ورودی دوم محور `y` را به آن میدهیم تا برای ما تابع را رسم کند. این `plt.plot` تابع را از نقطه ای به نقطه دیگر رسم توسط خط رسم میکند. به عنوان آرگومان های دیگر این تابع میتوان از `label` و `color` هم برای بیسشتر کردن جزئیات تابع درنظر گرفت. از دیگر تابع برای رسم میتوان از `plt.scatter` استفاده کرد. این تابع برای هر مشاهده، یک نقطه ترسیم میکند. همانند `plt.plot` به آن محورهای مختصات را برای رسم میدهیم. و یک سری آرگومان های دیگر همانطور که بالا گفته شد!

## سوال 2:

در سوال دو از ما خواسته شده که از بین تقریب پواسون و نرمال کدام یک تقریب بهتری برای توزیع دوجمله ای است. چون احتمال ما نزدیک به صفر است، تقریب پواسون تخمین بهتری برای مسئله ما است.

در ابتدا باید یک حدی برای تعداد نمونه برداری در نظر گرفت که هرچه بیشتر باشد تخمین ما توسط پواسون دقیق و بهتر خواهد بود. در مرحله بعد باید با استفاده از کتابخانه **numpy**، باید تعدادی نمونه هایی از توزیع دوجمله ای توسط احتمال و تعداد فضای نمونه و تعداد نمونه برداری از توزیع درست کرد - **numpy.random.bionomial**

چون احتمال ما کوچک است اعدادی که در بالا تولید میشود کوچک خواهند بود. و برای بقیه احتمالها با تقریب خوبی صفر میشوند. پس یک حدی مثلا 2 یا 3 در نظر میگیریم و تعداد اینها را محاسبه میکنیم و سپس با تقسیم این تعدادها بر تعداد نمونه برداری ها، احتمال آنها را حساب میکنیم. سپس با توجه به چیزهایی که در سوال قبل گفته شد نمودار این توزیع دوجمله ای میتوان رسم کرد.

برای توزیع پواسون به ازای اون رنج اعداد مثلا 3 هرکدام از اعداد 0 تا 2 را به تابعی که در کتابخانه **scipy.stats.poisson** است به نام **poisson.pdf** میدهیم همراه با لامدا برای محاسبه **pdf** توزیع پواسون. برای نرمال هم به ازای همون رنج و به کمک **norm.pmf** و میانگین و انحراف معیار **pmf** توزیع نرمال را بدست می آوریم و اینها را رسم و مقایسه می کنیم -> تقریب پواسون بهتر است

### سوال 3:

برای بدست آوردن اینکه حداقل نمره فردی که جزء 10 درصد بالای کلاس است، چند است، از تابع **normal.ppf** استفاده میکنیم. به این ترتیب که آرگومان اول را 0.9 آرگومان دوم را میانگین و سومی را انحراف معیار میدهیم. برای بدست آوردن بازه نمرات بین چارک دوم و سوم از روند بالا استفاده میکنیم. برای مثال چارک دوم حداقل نمره در این بازه برابر با 0.25 اول بازه است. برای محاسبه قسمت سوم هم از **normal.cdf** استفاده می کنیم طوری که آرگومانهای اول را 80 و 90 و بعدی ها میانگین و انحراف معیار می باشند و این دو عدد را در نهایت از هم باید کم کنیم. برای بخش امتیازی:

عدد 100000 هم برای تعداد نمونه گیری ها در نظر میگیریم. برای درس فیزیک: حد نمرات بین 0 تا 20 در نظر گرفته شده با توزیع **uniform**. درس برنامه نویسی با توزیع نمایی: پارامتر **scale** در توزیع نمایی، گسترش را کنترل می کند. اگر می خواهیم امتیازات به طور گسترده تری متفاوت باشد، می توان مقیاس را افزایش داد. به عنوان مثال، شما ممکن است **scale=5** را برای توزیع کمی گسترده تر تنظیم میکنیم.

برای درس گسسته ☹️: پارامتر میانگین در توزیع پواسون ( $\lambda$ ) نشان دهنده میانگین تعداد رویدادها در یک بازه ثابت است. اگر می خواهیم میانگین امتیاز بالاتری را شبیه سازی کنیم، این پارامتر را میتوان افزایش داد. به عنوان مثال، می توان  $\lambda = 18$  را برای میانگین نمره 18 در نظر گرفت.

برای هر توزیع به اندازه تعداد نمونه گیری ها سمپل گرفته شد توسط توابع `np.random` و در آخر اینها را با هم جمع می کنیم. اگر نمودار جمع حاصله را رسم کنیم، مشاهده میشود که نمودار هیستوگرام خیلی به توزیع نرمال نزدیک است. برای بدست آوردن حدود محور  $x$  برای محوهای نمودار فعلی استفاده میشود. سپس توسط `normal.pdf` ، `pdf` را برای این رنج توسط میانگین و انحراف معیار آن لیست جمع شده از نمرات، را برای رسم محاسبه میکنیم. بعد از رسم مشاهده میشود که دو نمودار به یکدیگر بسیار نزدیک هستند.

سوال 4:

برای سوال 4 همانند سوال 2 رفتار میکنیم. اما این دفعه تقریب نرمال بهتر از تقریب پواسون خواهد بود. چون احتمال ما نزدیک به 0.5 است ☹️