دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر

تمرین سری دوم

درس آمار و احتمالات مهندسی

استاد امیرحائری

فروردین :)))))) ۹۷

مقدمه

سلام!

یه سری تغییر که اینطوری مشخص شده رو اضافه کردیم.

این تمرین دارای دو بخش است. بخش اول سوالات حل کردنی و بخش دوم سوالاتی با عنوان آزمایش‌های کامپیوتری است. برای بخش اول یک PDF ازتون می‌خواهیم که توش سوالات حل شده باشه و اسم فایلش باید به فرم

PS_HW2_PART1_9231020.pdf

باشه. اگه توی برگه دوست دارید بنویسید، باز در آخر یه فایل pdf بسازید و مثلا یه سری فایل jpg آپلود نکنید.

برای بخش دوم هم، هم فایل‌های مربوط به برنامه‌تون رو می‌خواهیم و هم یک گزارش که اونم به صورت pdf باید باشه و فرمت اسمش هم به صورت PS_HW2_CE_9231020.pdf بگذارید.

ما فایل‌های شما رو فقط و فقط از طریق مودل دریافت می‌کنیم و ارسال آن‌ها از طریق‌های دیگر به هیچ وجه مورد بررسی قرار نمی‌گیرد. البته کسایی که دسترسی به مودل دارند منظورمونه.

در صورتی که نیاز باشه خبری در مورد تمرین‌ها، کلاس تی‌ای و ... بدهیم سعی می‌کنیم هم از طریق این کانال این خبر رو انتقال بدهیم هم از طریق مودل. اکثرا مودل و ایمیل از طریق مودل

یه سری راهنمایی گذاشتیم که روش رو هایلایت نکردیم.

اگر سوالی در رابطه با تمرین‌ها داشتید می‌تونید از طریق این ایمیل از ما بپرسید.

Email Address: autps9697@gmail.com

قسمت سوالات حل کردنی

سوال ۱ (نامساوی مارکف)

الف)

برای هر متغیر تصادفی نامنفی X و هر $a > 0$ ثابت کنید:

$$P(X > a) \leq \frac{E(X)}{a}$$

که $E(X)$ برابر با امید ریاضی است.

ب)

با استفاده از قسمت قبل اثبات کنید به ازای هر متغیر تصادفی X و هر $a > 0$ رابطه زیر برقرار است.

$$P(|X - \mu| > a) \ \leq \frac{Var(X)}{a^2}$$

ج)

فرض کنید یک دستگاه داریم که یک دکمه دارد. در صورت فشار دادن آن، خروجی دستگاه یک متغیر تصادفی با توزیع $Normal(\mu = 0, \sigma = 2)$ به ما نشان می‌دهد. فرض کنید، علی یک بار این دکمه را فشار داده است و عدد خروجی را در جایی یادداشت کرده است. او ادعا می‌کند که عددی که یادداشت کرده است از 5 بزرگتر است. چقدر احتمال دارد علی راست بگوید؟

جواب شما باید به صورت زیر باشد:

احتمال اینکه علی راست بگوید، حداکثر فلان قدر (0.00123456) است.

$$P\left(\text{راستگویی}\right) \leq z$$

z را بیابید. (سعی کنید کوچک‌ترین z ای را بیابید که در رابطه بالا صدق می‌کند.)

سوال ۲

الف)

اثبات کنید:

$$E(aX+b)=aE(X)+b$$

ب)

در صورتی که X و Y مستقل باشند اثبات کنید:

$$E(XY)=E(X)E(Y)$$

ج)

اگر X و Y دو متغیر تصادفی با تابع چگالی احتمال توام f(x,y) باشند؛ ثابت کنید (a، b و c ضرایب ثابت هستند):

$$\sigma_{aX+bY+c}^2=a^2\sigma_X^2+b^2\sigma_Y^2+2ab\sigma_{XY}$$

د)

در صورتی که X و Y مستقل باشند اثبات کنید

$$\sigma_{aX-bY}^2=a^2\sigma_X^2+b^2\sigma_Y^2$$

سوال ۳

با توجه به تعریف کواریانس

$$Cov(X,Y) = E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big]$$

و با توجه به تعریف correlation:

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

الف)

ابتدا ثابت کنید رابطه زیر برقرار است:

$$Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 2 + 2\rho(X,Y)$$

همچنین ثابت کنید رابطه زیر برقرار است:

$$Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 2 - 2\rho(X,Y)$$

راهنمایی: از تعریف واریانس در سوال‌های قبل استفاده کنید.

ب)

اثبات کنید همیشه رابطه زیر برقرار است

$$-1 \leq \rho(X,Y) \leq 1$$

راهنمایی: توجه کنید که واریانس همیشه مثبت است. و از رابطه الف استفاده کنید.

ج)

اثبات کنید در صورتی که X و Y تابعی خطی به صورت زیر باشد:

$$Y = aX + b, a > 0$$

آنگاه

$$\rho(X,Y) = 1$$

د) اثبات کنید در صورتی که X و Y متغیرهایی تصادفی باشند که $\rho(X,Y) = 1$، آنگاه باید با هم رابطه خطی به صورت زیر داشته‌باشند:

$$Y = aX + b, a > 0$$

راهنمایی: یه متغیر تصادفی‌ای (یک ترکیب خطی از متغیرهای تصادفی X و Y) واریانسش صفر میشه، در نتیجه اون متغیر تصادفی باید یک عدد ثابت باشه. با توجه به این نشون بدید که X و Y باید چنین رابطه ای داشته باشند.

سوال ۴

فرض کنید سه متغیر تصادفی پیوسته به نام X,Y,Z داریم به طوری که:

$$P(X = Y) = P(X = Z) = P(Y = Z) = 0$$

مثلا ممکن است X میانگین دمای کلاس ۲۰۳ در هفته اخیر به ازای درجه سانتی‌گراد باشد. Y مقدار یک آلاینده در ش‌ه‌ر در ی‌ک مدت زمان خاص با یک واحد مشخص باشد و Z میانگین دمای کلاس ۱۰۷ در سه هفته اول تابستان امسال باشد. این متغیرهای تصادفی احتمال توام دارند. (البته بدیهی هست.)

حال اثبات کنید که امکان ندارد این شرایط برقرار باشد:

$$P(X > Y) = P(Y > Z) = P(Z > X) = 0.7$$

راهنمایی: از ۳ اصل احتمال برای این اثبات استفاده کنید. (جمع یه سری چیز بیشتر از یک میشه و تناقضه)

سوال ۵

میانگین و واریانس توزیع‌های زیر را بدست آورید:

الف)

توزیع برنولی

$$f(k;p) = \begin{cases} p & if\ k = 1 \\ 1-0 & if\ k = 0 \end{cases}$$

ب)

توزیع binomial

$$pr(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k} \qquad for\ k = 0, 1, 2, \ldots, n$$

ج)

توزیع پواسون

$$f(k;\ \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \quad :\ k\ =\ 0, 1, 2 \ldots$$

د)

توزیع نمایی

$$f(x;\ \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

سوال ۶

با توجه به رابطه زیر

$$p(x,y) = p(y)p(x|y)$$

الف) ثابت کنید برای دو متغیر تصادفی X، Y این رابطه صادق است.

$$E(X) = E\big(E(X|Y)\big)$$

ب) با توجه به رابطه بالا، به مساله زیر پاسخ دهید.

محسن قصد دارد یا یک فصل از کتاب احتمال و یا یک فصل از کتاب تاریخ خود را مطالعه کند.

اگر تعداد غلط های چاپی هر فصل از کتاب احتمال و کتاب تاریخ دارای توزیع پواسون با میانگین‌های به ترتیب ۲ و ۵ باشند.

و محسن به یک اندازه به این دو کتاب علاقه داشته باشد، تعداد غلط های چاپی که ممکن است محسن با آن مواجه شود (به طور میانگین) چقدر است؟

سوال ۷

علی در یک شرکت کار میکند و ساعتی 20 هزارتومن پول می‌گیرد. او برای رسیدن به محل کار، در ایستگاه تاکسی منتظر آمدن تاکسی است. با توجه به اینکه او زود به ایستگاه آمده است، کسی جز او در صف نیست. و او باید منتظر آمدن ۴ نفر بشود. فرض کنید این لحظه را لحظه $t = 0$ در نظر بگیریم. در $t = 5$ یعنی ۵ دقیقه بعد، یک نفر به ایستگاه می‌آید. در آن لحظه، علی از خود می‌پرسد آیا بهتر است که من اسنپ بگیرم و با آن به محل کار بروم و هزینه ۲۰ هزارتومان بپردازم (با این فرض که سرعت تاکسی و اسنپ برابر است و ماشین اسنپ بدون هیچ تاخیری شروع به حرکت می‌کند، ترافیک نیز در طول روز ثابت است)

یا اینکه صبر کنم تا دو نفر دیگر بیاید و من هزینه تاکسی را بدهم. که این هزینه ۵ هزارتومن می‌شود.

پس علی دو گزینه دارد.

یا صبر کند و سوار تاکسی شود و هزینه ۵ هزارتومن را بپردازد. در این صورت به ازای زمانی که منتظر تاکسی بوده است، نمی‌تواند کار کند و در نتیجه ضرر می‌کند.

یا اینکه همین الان اسنپ بگیرد با هزینه ۲۰ هزارتومن. در عوض وقت بیشتری کار می‌کند.

با توجه به اینکه زمان صبر کردن برای تاکسی یک متغیر تصادفی هست، امید این متغیر تصادفی را برای محاسبات خود در نظر بگیرید.

راهنمایی:

اولا فرض کنید که زمان آمدن هر مسافر از توزیع نمایی پیروی می‌کند.

$$f(x,\lambda) = \begin{cases} \lambda e^{-\lambda x}: x \geq 0 \\ 0: x \leq 0 \end{cases}$$

دو متغیر تصادفی t_1, t_2 فرض کنید. این دو متغیر تصادفی، به ترتیب میزان صبر کردن برای مسافر اول و مسافر دوم هستند.

حال زمان صبر کردن برابر با

$$T = t_1 + t_2$$

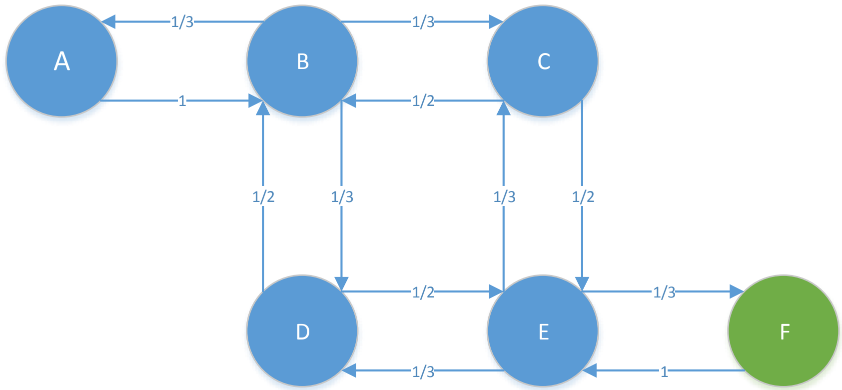
$$E(T) = E(t_1) + E(t_2)$$

با توجه به اینکه فرایند پواسن بی‌حافظه است، می‌توان دو متغیر t_1, t_2 را از هم مستقل و با توزیع یکسان دانست.

همچنین با توجه به زمان آمدن یک مسافر، مقدار $\frac{1}{\lambda} = 5$ را برای توزیع خود فرض می‌کنیم.

سوال ۸ (قدم‌زن تصادفی)

یک قدم زن تصادفی داریم که ابتدا در نقطه A است. و در هر ثانیه یک گام بر می‌دارد. به این صورت که با یک احتمال (که بر روی فلش ها نوشته شده است) به خانه های اطراف می‌رود. (مثلا اگر در خانه B باشد با احتمال 1/3 به A، 1/3 به C و 1/3 به D می‌رود.)



یک فرد در نقطه F منتظر این قدم‌زن است! او نمی‌داند که به حدودا چقدر باید منتظر A باشد.

الف)

به طور میانگین تعداد قدم‌های قدم‌زن تصادفی برای رسیدن از خانه A به خانه F را بیابید.

متغیر تصادفی T_F را اینگونه تعریف می‌کنیم که تعداد گام‌های لازم برای رسیدن به F است.

برای آشنایی بیشتر با اصطلاحات: X_i یک متغیر تصادفی است که میتواند مقادیر {A,B,C,D,E,F} را در این مساله بگیرد و احتمال بودن قدم زن تصادفی در گام iام را نشان می‌دهد.

با توجه به تعاریف بالا، خواسته این قسمت این است:

$$E(T_F|start = A) =?$$

راهنمایی: رابطه بین $E(T_F|start = A)$ و $E(T_F|start = B)$ و ... را بدست بیاورید. همچنین $E(T_F|start = F) = 0$ چون وقتی از F شروع می‌کنیم خوب قطعا رسیدیم دیگه. پس نباید صبر بکنیم.

ب)

در صورتی که بدانیم این قدم‌زن تصادفی پس از دو حرکت، در خانه C است، میانگین تعداد قدم‌ها برای رسیدن را بدست آورید.

$$E(T|\text{است } c \text{ پس از دو گام در})$$

ج) احتمال اینکه دقیقا بعد از ۴ گام قدم‌زن تصادفی به F برسد را حساب کنید.

د) احتمال اینکه دقیقا بعد از ۵ گام قدم‌زن تصادفی به F برسد را حساب کنید.

سوال ۹

- یک زندانی در سلولی با ۳ در زندانی شده است.
- در اول به تونلی منتهی می‌شود که بعد از دو روز او را به سلول برمی‌گرداند.
- در دوم به تونلی منتهی می‌شود که بعد از سه روز او را به سلول برمی‌گرداند.
- در سوم به بیرون راه دارد و باعث آزادی او می‌شود.

الف)

زندانی همیشه در ۱ و ۲ و ۳ را به ترتیب با احتمال های ۰,۵ و ۰,۳ و ۰,۲ انتخاب می‌کند.

بعد از چند روز انتظار داریم که زندانی به آزادی برسد؟

ب)

فرض کنید که زندانی از بین درهایی که تا الان انتخاب نکرده است یکی را با احتمال یکسان یکی را انتخاب می‌کند (اولش $1/3$ هست همه احتمالا) (به عنوان مثال اگر زندانی ابتدا در۱ را انتخاب کند زمانی که به سلول برمی‌گردد تنها از بین در ۳ و ۲ انتخاب می‌کند.)

در این حالت بعد از چند روز انتظار داریم تا زندانی به آزادی برسد؟

ج)

در حالت الف و ب واریانس روزهایی که زندانی به آزادی می‌رسد را محاسبه کنید؟ (راهنمایی: از امید شرطی استفاده کنید، هم برای $E(x)$ هم برای $E(x^2)$)

سوال ۱۰

تابع توزیع توام دو متغیر تصادفی در زیر داده شده است:

$$f(x_1\,,x_2\,) = \begin{cases} 6x_2. & 0 < x_2 < x_1 < 1 \\ 0 & elsewhere \end{cases}$$

الف)

تابع توزیع حاشیهای $f_{X_1}(x1)$ را محاسبه کنید. در ادامه ثابت کنید که این تابع میتواند تابع توزیع چگالی باشد.

ب)

با فرض این که $X_1 = 0/7$ باشد احتمال این که متغیر تصادفی $X_2 < 0/5$ باشد را محاسبه کنید.

تمرین های کامپیوتری

گزارش میخواد این بخش!

در این بخش می‌توانید از زبان‌های python، matlab و R استفاده کنید.

فایل گزارش به ازای این دو سوال را در یک pdf به فرم PS_HW2_CE_9231020.pdf بنویسید.

سوال اول

الف)

اول با استفاده از کتابخانه های دلخواه، n عدد از توزیع نرمال به صورت $Normal(\mu = 0, \sigma = 1)$ را در نظر بگیرید،

هیستوگرام این اعداد را بکشید. به ازای n={10,100,1000,10000}. اگه دوست دارید میتونید نرمال کنید این هیستوگرام ها رو یعنی مساحت ستون ها بشه یک فک کنم

ب)

حال نمودار توزیع نرمال را بکشید.

ج)

حال توزیع binomial(n,p) را در نظر بگیرید.

به ازای هر کدام از حالت های (n,p) زیر:

N	P
10	0.3
100	0.03
1000	0.003
10000	0.0003

از توزیع binomial(n,p) به تعداد k=10000 نمونه برداری کنید. و هیستوگرام این k عدد را بکشید.

آیا این هیستوگرام‌ها شبیه توزیع نرمال است؟ درصد داده هایی که در بازه [np*0.9, np*1.1] است آیا افزایش پیدا می‌کند؟ کمی توضیح دهید.

سوال دوم

الف)

با استفاده از توزیع پواسون با $\lambda = 3$ ، n=1000 نمونه تولید کنید. و هیستوگرام آن را بکشید.

ب)

حال به ازای مجموعه‌های زیر، نمونه تولید کنید و هیستوگرام آن را نیز بکشید. با استفاده از binomial(n,p).

N	P
10	0.3
100	0.03
1000	0.003
10000	0.0003

در نتیجه شما چهار هیستوگرام دارید. آیا این نمودارها شبیه هسیتوگرام قسمت الف هستند؟

خروجی مورد نظر

Computer_Experiment.py
PS_HW2_CE_9231020.pdf
PS_HW2_PART1_9231020.pdf