



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

تمرین سری دوم و سوم آمار و احتمال مهندسی

برای تحویل تمرین به نکات زیر توجه فرمایید.

- پاسخ سوالات بخش تشریحی به همراه گزارش مختصری از سولات قسمت بخش آزمایشات کامپیوتری را در فایل pdf در تحت عنوان HW23_StudentNumber.pdf قرار دهید.
- پاسخ سوالات تشریحی می تواند به صورت تایپی و یا عکس خوانا از نسخه دستنویس باشد.
- گزارش برای سوالات بخش کامپیوتری، شامل شرح مختصر و نتیجه گرفته شده به همراه تصویری از اجراست.
- کدهای مربوط به بخش کامپیوتری را در در فولدری جداگانه قرار داده و همراه با فایل pdf با نام HW23_StudentNumber.zip فشرده سازی کرده و بارگذاری نمایید.
- تمرین خود را قبل از زمان مشخص شده در مودل یا lms آپلود کنید.
- نمره تمرین های مشابه نمره ای بر تعداد افراد دیگر در کپی تقسیم خواهد شد .
- در صورت وجود هرگونه ابهام می توانید از طریق ایمیل با تدریس یارهای درس در ارتباط باشید .

۱- فرض کنید X و Y دو متغیر با توزیع یکنواخت و یکسان در بازه‌ی $[0, 1]$ باشند. مقدار $E(X|X>Y)$ را به دست آورید.

۲- اگر X یک متغیر تصادفی با چگالی احتمال $f_X(x) = \frac{1}{\sqrt{4\pi}} e^{-\frac{x^2}{4}}$ و $Y=AX$ که در آن A یک متغیر تصادفی باینری با توزیع $P\{A=0\}=3/4$ و $P\{A=1\}=1/4$ و مستقل از متغیر تصادفی X باشد. مقدار ضریب همبستگی دو متغیر X و Y را به دست آورید.

۳- فرض کنید :

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{3}(x+y) & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{otherwise.} \end{cases}$$

مقدار واریانس: $V(2X-3Y+8)$ را به دست آورید.

۴- بردن نمرات دانشجویان بر روی نمودار بدین صورت انجام می‌شود که ابتدا فرض می‌شود نمرات از توزیع نرمال با پارامترهای μ و σ^2 به دست آمده‌اند. سپس بر اساس نمرات دانشجویان برآوردی از μ و σ^2 به دست می‌آید و آنگاه نمرات A, B, C, D و F به ترتیب به دانشجویان زیر اختصاص داده می‌شود.

به دانشجویانی که نمره‌ی آنها بیش از $\mu + \sigma$ باشد نمره A

به دانشجویانی که نمره‌ی آنها بین μ و $\mu + \sigma$ باشد نمره B

به دانشجویانی که نمره‌ی آنها بین μ و $\mu - \sigma$ باشد نمره C

به دانشجویانی که نمره‌ی آنها بین $\mu - \sigma$ و $\mu - 2\sigma$ باشد نمره D

و به دانشجویانی که نمره‌ی آنها کمتر از $\mu - 2\sigma$ باشد نمره F تعلق می‌گیرد.

با روش بالا هریک از نمرات فوق به چند درصد از دانشجویان کلاس تعلق می‌گیرد؟

نکته: از جدول توزیع تجمعی نرمال استاندارد شده (Standard Normal Cumulative Distribution) واقع در صفحه ۷۳۵ کتاب یا [wiki](#) استفاده نمایید.

۵- در یک شهر احتمال قطعی برق در هر روز از روزهای دیگر مستقل و برابر ۰.۰۲ است، اگر ۶۰۰ روز این شهر مورد بررسی قرار گیرد، الف) با چه احتمالی دقیقاً ۳ روز قطع برق وجود دارد؟ ب) با چه احتمالی بیشتر از ۱۰ روز و کمتر مساوی ۲۵ روز قطعی برق رخ می‌دهد؟

*** با استفاده از هر دو توزیع نرمال و پواسون تقریب بزنید.

۶- فرض کنید X یک متغیر تصادفی با توزیع نرمال و میانگین ۱۰ است، اگر $P(X > 20) = 0.2$ باشد، $Var(X)$ را به دست آورید؟

۷- مدت زمانی (بر حسب ساعت) که لازم است تا یک قطعه تعمیر شود دارای توزیع نمایی با میانگین ۴ است.

(الف) احتمال اینکه مدت تعمیر بیش از ۴ ساعت طول بکشد چقدر است؟

(ب) احتمال شرطی اینکه مدت زمان تعمیر حداقل ۱۰ ساعت طول بکشد به شرط اینکه بیش از ۹ ساعت از زمان تعمیر گذشته باشد، چقدر است؟

۸- طول عمر یک رادیو بر حسب سال دارای توزیع نمایی با میانگین ۸ سال است. (الف) اگر فردی یک رادیو دست دوم

خریده باشد که تا الان ۱۰ سال کار کرده است، احتمال اینکه رادیو ۸ سال دیگر کار کند چقدر است؟ (ب) اگر رادیو

دست دوم تا الان ۱۰۰ سال کار کرده باشد، احتمال اینکه ۸ سال دیگر کار کند چقدر است؟

۹- اگر X دارای توزیع دو جمله ای به صورت زیر باشد، توزیع $Y = X^2$ را بیابید.

$$f(x) = \begin{cases} \binom{3}{x} \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{3-x} & x = 0, 1, 2, 3 \\ 0 & o.w \end{cases}$$

۱۰- اگر X دارای توزیع زیر باشد، توزیع $Y = X^2$ را بیابید.

$$f(x) = \begin{cases} \frac{1+x}{2} & -1 < x < 1 \\ 0 & o.w \end{cases}$$

۱۱- ادعا شده است که میانگین برق مصرفی فروردین ماه یک ناحیه تهران دست کم ۱۴۰۰ کیلو وات ساعت بوده است.

بدین منظور یک نمونه تصادفی به تعداد ۵۰۰ خانوار از منطقه انتخاب شده که میانگین و انحراف معیار برق مصرفی

آن‌ها ۱۳۵۲ و ۲۷۷ کیلو وات ساعت است. سطح معناداری ۱٪ را در نظر گرفته و صحت ادعا را با توجه به یک

آزمون دو طرفه بررسی کنید.

۱۲- می خواهیم دو الگوریتم جستجوی تصادفی گراف را به لحاظ زمان اجرا با یکدیگر مقایسه نماییم. این الگوریتم ها را

چند بار اجرا نموده ایم. زمانهای اجرا بر حسب ثانیه در جدول زیر مشخص شده اند. آیا می توان نتیجه گرفت که

الگوریتم دوم به طور متوسط پیچیدگی زمانی کمتری دارد؟ (از آزمون فرضیه یک طرفه مناسب استفاده نمایید و تمام جزییات را بیان نمایید. $\alpha = 0.05$)

ALG1	۲	۳	۱	۵	۱	۳	۳
ALG2	۳	۴	۱	۱	۳	-	-

۱۳- روی بردهای الکتریکی یک کارخانه قید شده است که توان مصرفی آنها به طور میانگین ۲۰۰ وات با انحراف معیار ۱۵ وات است. یک مرکز کنترل کیفیت نمونه های ۲۰ تایی از این بردها را تهیه می کند. اگر میانگین توان مصرفی این نمونه ها بین ۱۸۰ تا ۲۲۰ بود به کارخانه گواهی صحت اطلاعات می دهد و در غیر این صورت ادعای کارخانه تایید نمی شود.

(الف) در اینجا عملاً از چه آزمون فرضیه ای استفاده شده است. (فرضیه های آن را بنویسید)
(ب) خطای نوع اول و دوم را تعریف نمایید و بیان نمایید چگونه همزمان می توانید خطای نوع اول و دوم را کاهش دهید؟

(ج) خطای نوع اول در این آزمون برابر با چه مقداری است؟
(د) خطای نوع دوم را درحالتی که میانگین توزیع توان مصرفی برابر با ۲۳۰ وات برای مرکز کنترل کیفیت غیر قابل قبول باشد، به دست آورید.

۱۴- طراح یک الگوریتم رمز نگاری ادعا دارد به ازای پیام های ۱۰۰۰ کاراکتری میانگین زمان رمز کردن پیام برابر با ۸۰ نانو ثانیه و انحراف معیار ۲۰ نانو ثانیه است. طراح یک سیستم ارسال ایمیل این الگوریتم را برای رمز نمودن ایمیل های سیستم خود در صورتی انتخاب می کند، که اگر برای یک نمونه ۳۵ تایی از پیام های ۱۰۰۰ کاراکتری میانگین زمان رمز کردن پیام کمتر از ۸۷ نانو ثانیه باشد.

(الف) این شیوه ارزیابی عملاً معادل چه آزمون فرضیه ای است. (فرضیه های آن را بنویسید)
(ب) خطای نوع اول و دوم را تعریف نمایید و بیان نمایید چگونه همزمان می توانید خطای نوع اول و دوم را کاهش دهید؟

(ج) خطای نوع اول در این آزمون برابر با چه مقداری است؟
(د) خطای نوع دوم را درحالتی که اگر میانگین زمان رمز کردن الگوریتم ۹۰ نانو ثانیه غیر قابل قبول باشد، به دست آورید.

۱۵- عملیات جستجو توسط موتور جستجوی A برای یک نمونه ۳۰ تایی از سوالات دارای میانگین ۳۴۰ میلی ثانیه و انحراف معیار ۵۰ میلی ثانیه و برای موتور جستجوی B برای یک نمونه ۳۲ تایی از سوالات دارای میانگین ۳۸۰ میلی ثانیه و انحراف معیار ۴۸ میلی ثانیه است. آیا می توان با استفاده از آزمون فرضیه دو طرفه بیان نمود که موتور جستجوی A به طور متوسط سریعتر از موتور جستجوی B است. سطح معناداری آزمون را ۱/۰ در نظر بگیرید.

۱۶- اگر دو متغیر تصادفی X_1 و X_2 از یکدیگر مستقل باشند و هر دو دارای توزیع نمایی با پارامتر یکسان باشند تابع توزیع $Y = X_1 - X_2$ را بیابید.

آزمایشات کامپیوتری

۱- با استفاده از شبیه سازی کامپیوتری توزیع پواسون $X \sim \text{Pois}(6)$ را ترسیم نمایید. سپس با استفاده از توزیع دوجمله‌ای (Binomial distribution) این توزیع را تقریب بزنید. این کار را برای تعداد نمونه ۵۰۰ و ۵۰۰۰۰ تکرار کرده و نمودار هیستوگرام نرمال شده آن‌ها را با توزیع اصلی مقایسه نمایید.
راهنمایی:

$$\lim_{n \rightarrow \infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{\lambda^x e^{-\lambda}}{x!}$$

where $\lambda = np$

مثلاً توزیع پواسون $X \sim \text{Pois}(6)$ را می‌توانید با استفاده از $X \sim \text{binomial}(100, 0.06)$ تقریب زد.

۲- سه توزیع دوجمله‌ای $X_1 \sim \text{binomial}(100, 0.3)$ ، $X_2 \sim \text{binomial}(100, 0.5)$ و $X_3 \sim \text{binomial}(200, 0.5)$ را در نظر بگیرید. با استفاده از شبیه سازی کامپیوتری نشان دهید $X_1 + X_2$ و $X_2 + X_3$ نیز دارای توزیع دوجمله‌ای هستند و پارامترهای مربوط به هریک را بیابید.

موفق باشید.

a.mahdi@aut.ac.ir