آمار و احتمالات مهندسی

تمرین ششم - قضایای حدی

روژین و علی

تاریخ تحویل ۱۴۰۱/۱۰/۱۲

سؤال ۱.

فرض کنید دو متغیر تصادفی X و Y مستقل، با میانگین صفر و واریانس σ^2 هستند. Z و T را به صورت $Z = X + Y$ و $T = aX + \frac{1}{2}Y$ تعریف می‌کنیم. a را طوری بیابید که ضریب همبستگی بین Z و T برابر $\frac{1}{2}$ باشد.

پاسخ .

$$Var(Z) = Var(X+Y) = \underbrace{Var(X)}_{\sigma^2} + \underbrace{Var(Y)}_{\sigma^2} + 2\underbrace{Cov(X,Y)}_{0} = 2\sigma^2$$

$$Var(T) = Var(aX + \frac{1}{2}Y) = a^2 Var(X) + \frac{1}{4}Var(Y) + \underbrace{aCov(X,Y)}_{0} = (a^2 + \frac{1}{4})\sigma^2$$

$$\begin{aligned}
Cov(Z,T) &= E(ZT) - E(Z)E(T) \\
&= E(aX^2 + (a+\frac{1}{2})XY + \frac{1}{2}Y^2) - E(X+Y)E(aX + \frac{1}{2}Y) \\
&= aE(X^2) + (a+\frac{1}{2})\underbrace{E(XY)}_{E(X)E(Y)=0} + \frac{1}{2}E(Y^2) - (E(X)+E(Y))\underbrace{(aE(X) + \frac{1}{2}E(Y))}_{0} \\
&= aE(X^2) + \frac{1}{2}E(Y^2) \\
&= aE(X^2) + \frac{1}{2}E(Y^2) - \underbrace{a[E(X)]^2}_{0} - \underbrace{\frac{1}{2}[E(Y)]^2}_{0} \\
&= a\left(E(X^2) - [E(X)]^2\right) + \frac{1}{2}\left(E(Y^2) - [E(Y)]^2\right) \\
&= aVar(X) + \frac{1}{2}Var(Y) \\
&= (a + \frac{1}{2})\sigma^2
\end{aligned}$$

$$\Rightarrow Corr(Z,T) = \frac{Cov(Z,T)}{\sqrt{Var(Z)Var(T)}}$$

$$\frac{۱}{۲} = \frac{(a+\frac{۱}{۲})\sigma^۲}{\sqrt{۲(a^۲+\frac{۱}{۴})\sigma^۴}} \Rightarrow a = -۱ \pm \frac{\sqrt{۳}}{۲}$$

سؤال ۲.

تابع چگالی احتمال متغیرهای تصادفی X و Y برای $a > ۰$ به صورت زیر است:

$$f(x,y) = \begin{cases} \frac{۱}{a^۲} & ۰ < x,y < a \\ ۰ & \text{otherwise} \end{cases}$$

تابع توزیع انباشته متغیر تصادفی $W = max(\frac{X}{Y}, \frac{Y}{X})$ را بیابید.

پاسخ .

ابتدا $Z = \frac{X}{Y}$ را در نظر می‌گیریم. اگر $z \geq ۱$ باشد داریم:

$$\mathbb{P}(Z \leq z) = \mathbb{P}(X \leq zY) = \frac{۱}{a^۲}\int_۰^a\int_۰^a \chi_{x\leq zy}dydx = \frac{۱}{a^۲}\int_۰^a \min(a,zy)dy =$$

$$\frac{۱}{a^۲}(\int_۰^{\frac{a}{z}} zydy + \int_{\frac{a}{z}}^a ady) = \frac{۱}{a^۲}(\frac{a^۲}{۲z} + a^۲ - \frac{a^۲}{z}) = ۱ - \frac{۱}{۲z}$$

اگر $z \leq ۱$ باشد داریم:

$$\mathbb{P}(Z \leq z) = \mathbb{P}(X \leq zY) = \frac{۱}{a^۲}\int_۰^a\int_۰^a \chi_{x\leq zy}dydx = \frac{۱}{a^۲}\int_۰^a zydy = \frac{z}{۲}$$

با $w \geq ۱$ داریم:

$$\mathbb{P}(W \leq w) = \mathbb{P}(\frac{۱}{w} \leq Z \leq w) = \mathbb{P}(Z \leq w) - \mathbb{P}(Z < \frac{۱}{w}) = ۱ - \frac{۱}{۲w} - \frac{۱}{۲w} = ۱ - \frac{۱}{w}$$

و در نهایت برای تابع توزیع تجمعی داریم:

$$F_W(w) = \begin{cases} ۰ & w \leq ۱ \\ ۱ - \frac{۱}{w} & w \geq ۱ \end{cases}$$

سؤال ۳.

دو متغیر $Y \sim U(۰,۱)$ و $X \sim U(۰,۱)$ را دو متغیر تصادفی مستقل در نظر بگیرید.
اگر بدانیم متغیر تصادفی Z به صورت زیر تعریف می‌شود، تابع چگالی Z را بیابید.

$$Z = (-۲\ln X)^{\frac{۱}{۲}} cos(۲\pi Y)$$

پاسخ .

از متغیر کمکی $W = Y$ استفاده می‌کنیم و با جای‌گذاری آن در رابطه تعریف شده در صورت سوال، X را برحسب W و Z به دست می‌آوریم و داریم:

$$x_1 = e^{-\frac{(zsec(2\pi w))^2}{2}}$$

$$y_1 = w$$

می‌دانیم $f_{ZW}(z, w)$ از رابطه زیر به دست می‌آید:

$$f_{ZW}(z, w) = \sum_i |J(z, w)| f_{XY}(x_i, y_i)$$

پس ابتدا $J(z, w)$ را محاسبه می‌کنیم:

$$J(z, w) = \begin{vmatrix} \frac{\partial x_1}{\partial z} & \frac{\partial x_1}{\partial w} \\ \frac{\partial y_1}{\partial z} & \frac{\partial y_1}{\partial w} \end{vmatrix} = -zsec^2(2\pi w)e^{-\frac{(zsec(2\pi w))^2}{2}}$$

حال داریم:

$$f_{ZW}(z, w) = |z|sec^2(2\pi w)e^{-\frac{(zsec(2\pi w))^2}{2}}, -\infty < z < +\infty, 0 < w < 1$$

حال با انتگرال‌گیری بر روی مقادیر مختلف w می‌توانیم تابع چگالی Z را به دست بیاوریم:

$$f_z(z) = \int_0^1 f_{ZW}(z, w)dw = e^{-z^2/2} \int_0^1 |z|sec^2(2\pi w)e^{-\frac{(|z|tan(2\pi w))^2}{2}} dw$$

از متغیر کمکی $u = |z|tan(2\pi w)$ استفاده می‌کنیم. پس $du = 2\pi|z|sec^2(2\pi w)dw$. همین‌طور واضح است که $-\infty < u < +\infty$.

پس داریم:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \int_{-\infty}^{+\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, -\infty < z < +\infty$$

سؤال ۴.

یک بازیکن به طور همزمان یک تاس سالم و یک سکه سالم را پرتاب می‌کند. اگر سکه رو بیاید، ۲ برابر عدد تاس و اگر پشت بیاید، $\frac{1}{2}$ برابر عدد تاس امتیاز می‌گیرد. امید ریاضی امتیاز او را محاسبه کنید.

پاسخ .

فرض می‌کنیم R پیشامد رو آمدن سکه و N عدد روی تاس باشد. می‌دانیم N و R مستقل هستند. اگر امتیازی که برنده می‌شود را با W نشان دهیم داریم:

$$P(W = 2N|R) = 1 \qquad P(W = \frac{1}{2}N|R^c) = 1$$

بنابراین با توجه به رابطه‌ی $E(X)=\sum_{i=1}^{m}E(X|A_i)P(A_i)$ که با به کارگیری قضیه احتمال کل به آن می‌رسیم، داریم:

$$\begin{aligned}
E(W) &= E(W|R)P(R)+E(W|R^c)P(R^c)\\
&= E(2N|R)P(R)+E(\frac{1}{2}N|R^c)P(R^c)\\
&= E(2N)P(R)+E(\frac{1}{2}N)P(R^c)\\
&= 2E(N)P(R)+\frac{1}{2}E(N)P(R^c)\\
&= 2(\frac{1+2+3+4+5+6}{6})(\frac{1}{2})+\frac{1}{2}(\frac{1+2+3+4+5+6}{6})(\frac{1}{2})\\
&= \frac{35}{8}=4.375
\end{aligned}$$

سؤال ۵.

در یک سیستم نرم‌افزاری، دو سرور برای پاسخ به درخواست‌های کاربران وجود دارد. این درخواست‌ها با احتمال N به سرور اول ارسال می‌شوند که N خود یک متغیر تصادفی با توزیع $U(0,1)$ است. دو هزار درخواست از طرف کاربران ایجاد شده است. می‌دانیم هر دو هزار درخواست به سمت سرور اول ارسال شده‌اند. مطلوب است محاسبه:

الف) تابع چگالی احتمال شرطی متغیر تصادفی N

ب) امید ریاضی شرطی N

پ) احتمال اینکه دو هزار درخواست بعدی نیز به سرور اول ارسال شوند

پاسخ.

الف)
ابتدا X را تعداد درخواست‌های ارسال شده به سرور اول درنظر می‌گیریم. با توجه به قضیه بیز داریم:

$$f_{N|X}(n|x)=\frac{f_{X|N}(x|n)f_N(n)}{f_X(x)}$$

همچنین می‌دانیم:

$$f_{X|N}(x|n)=\binom{2000}{x}n^x(1-n)^{2000-x}$$

با جای‌گذاری در رابطه داریم:

$$f_{N|X}(n|2000)=\frac{n^{2000}}{f_X(x)}$$

می‌دانیم $f_{N|X}(n|2000)$ یک تابع چگالی احتمال است و رابطه زیر باید برقرار باشد:

$$f_X(x)=\int_0^1 n^{2000}dn=\frac{1}{2001}$$

و درنهایت داریم:

$$f_{N|X}(n|2000)=2001n^{2000}$$

ب)

$$E[N|X=2000]=\int_0^1 nf_{N|X}(n|2000)dn=\int_0^1 2001n^{2001}dn=\frac{2001}{2002}$$

پ)
ابتدا Y را تعداد درخواست‌های ارسال شده به سرور اول در دوهزار درخواست بعدی درنظر می‌گیریم و با توجه به قضیه بیز داریم:

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\int_۰^۱ f_{XY|N}(x,y|n) f_N(n) dn}{f_X(x)}$$

همچنین می‌دانیم:

$$f_{XY|N}(x,y|n) = \binom{۲۰۰۰}{x} n^x (۱-n)^{۲۰۰۰-x} \binom{۲۰۰۰}{y} n^y (۱-n)^{۲۰۰۰-y}$$

و در مرحله قبل به دست آوردیم که $f_X(x) = \frac{۱}{۲۰۰۱}$. پس داریم:

$$f_{Y|X}(۲۰۰۰|۲۰۰۰) = ۲۰۰۱ \int_۰^۱ n^{۴۰۰۰} dn = \frac{۲۰۰۱}{۴۰۰۱}$$

سؤال ۶.

می‌خواهیم یک برنامه را به صورت موازی بر روی یک پردازنده با دو هسته اجرا کنیم. هر هسته در این پردازنده دو حالت دارد؛ آزاد و مشغول. برای اجرا شدن این برنامه لازم است هر دو هسته در حالت آزاد باشند. برنامه برای شروع به اینصورت عمل می‌کند که به صورت یک در میان به هسته‌ها درخواست می‌دهد و اگر هر دو هسته به صورت پشت سر هم درخواست را قبول کنند (آزاد باشند)، برنامه اجرا می‌شود. احتمال اینکه هسته اول در هر درخواست آزاد باشد، p_A و احتمال اینکه هسته دوم در هر درخواست آزاد باشد، p_B است و می‌دانیم $p_B > p_A$.

حال اگر بخواهیم تعداد درخواست‌های مورد انتظار برای شروع برنامه کم‌ترین مقدار را داشته باشند، بهتر است ابتدا برنامه به کدام هسته درخواست بدهد؟ (توجه کنید به جواب نهایی بدون راه‌حل نمره‌ای تعلق نمی‌گیرد.)

پاسخ .

متغیر تصادفی N را تعداد درخواست‌های مورد نیاز در صورت شروع با هسته A و متغیر تصادفی M را تعداد درخواست‌های مورد نیاز در صورت شروع با هسته B در نظر می‌گیریم. همچنین f نشان‌دهنده آزاد بودن و b به معنای مشغول بودن است. حال داریم:

$$E[N] = E[N|f]p_A + E[N|b](۱-p_A)$$

$$E[N|f] = E[N|ff]p_B + E[N|fb](۱-p_B) = ۲ + (۱-p_B)E[N]$$

که در اینجا ff به معنای آزاد بودن هسته اول و دوم و fb به معنای آزاد بودن هسته اول و مشغول بودن هسته دوم در دو درخواست اول هستند.

همچنین داریم:

$$E[N|b] = ۱ + E[M]$$

با قرار دادن رابطه‌ها داریم:

$$E[N] = (۲ + (۱-p_B)E[N])p_A + (۱+E[M])(۱-p_A)$$

و با توجه به تقارن داریم:

$$E[M] = (۲ + (۱-p_A)E[M])p_B + (۱+E[N])(۱-p_B)$$

با توجه به شرط داده شده در سوال و کم کردن دو رابطه بالا از هم داریم: $E[N] < E[M]$

و این یعنی بهتر است از هسته A شروع کنیم.

سؤال ۷.

ندا می‌خواهد در N مسابقه شطرنج شرکت کند که N دارای توزیع هندسی با پارامتر s است. فرض کنید احتمال برد او در هر بازی، مستقل از سایر بازی‌ها، برابر p است. اگر T تعداد بازی‌هایی باشد که او می‌برد:

الف) میانگین و واریانس T را محاسبه کنید.

ب) تابع مولد گشتاور T را به دست آورید.

پاسخ .

الف)

از خواص امید ریاضی می‌دانیم:

$$E(T) = E(E(T|N))$$
$$E(T|N) = NE(p)$$

از دو خاصیت بالا نتیجه می‌شود:

$$E(T) = E(E(T|N)) = E(NE(p)) = E(N) \times p = \frac{p}{s}$$

برای محاسبهٔ واریانس، از لم زیر استفاده می‌کنیم.
حکم لم:

$$\text{Eve's Law}: \quad Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$$

اثبات لم:

$$E(Var(Y|X)) = E(E(Y^۲|X) - [E(Y|X)]^۲) = E(Y^۲) - E([E(Y|X)]^۲)$$
$$Var(E(Y|X)) = E([E(Y|X)]^۲) - [E(E(Y|X))]^۲ = E([E(Y|X)]^۲) - (E(Y))^۲$$

با جمع طرفین دو رابطهٔ بالا داریم:

$$E(Var(Y|X)) + Var(E(Y|X)) = E(Y^۲) - E([E(Y|X)]^۲) + E([E(Y|X)]^۲) - (E(Y))^۲$$
$$E(Var(Y|X)) + Var(E(Y|X)) = E(Y^۲) - (E(Y))^۲$$
$$E(Var(Y|X)) + Var(E(Y|X)) = Var(Y)$$

با استفاده از Eve's Law :

$$\begin{aligned} Var(T) &= E(Var(T|N)) + Var(E(T|N)) \\ &= E(Np(۱-p)) + Var(Np) \\ &= \frac{p(۱-p)}{s} + \frac{p^۲(۱-s)}{s^۲} \\ &= \frac{p(p(۱-s) + (۱-p)s)}{s^۲} \end{aligned}$$

ب) با توجه به اینکه برنده شدن ندا دارای توزیع برنولی با پارامتر p است، تابع مولد گشتاور T به صورت زیر به دست می‌آید:

$$\begin{aligned} E(e^{tT}) &= E(E(e^{tT}|N)) \\ &= E((pe^t+q)^N) \\ &= s\sum_{n=۰}^{\infty}(pe^t+q)^{n+۱}(۱-s)^n \\ &= s(pe^t+q)\sum_{n=۰}^{\infty}((pe^t+q)(۱-s))^n \end{aligned}$$

$$\xrightarrow[\text{همگرایی سری هندسی}]{\text{محاسبهٔ مقدار}} E(e^{tT}) = \frac{s(pe^t+q)}{۱-(۱-s)(pe^t+q)}$$

سؤال ۸.

می‌دانیم ایمیل‌ها با فرآیند پواسون با پارامتر λ دریافت می‌شوند. در نتیجه تعداد ایمیل‌ها در inbox در یک بازه زمانی به طول یک ساعت دارای توزیع $Poisson(\lambda)$ است. همچنین می‌دانیم تعداد ایمیل‌هایی که در بازه‌های زمانی مجزا می‌رسند، مستقل از هم هستند. فرض کنید X، Y و Z تعداد ایمیل‌های دریافتی یک روز خاص در بازه‌های زمانی ۹ صبح تا ۱۲ ظهر، ۱۲ ظهر تا ۶ بعد از ظهر و ۶ بعد از ظهر تا نیمه‌شب باشند.

الف) تابع جرم احتمال مشترک شرطی X، Y و Z را با شرط $X+Y+Z=۳۶$ به دست آورید.

ب) تابع جرم احتمال شرطی را برای $X+Y$ با شرط $X+Y+Z=۳۶$ به دست آورید.

پ) با توجه به قسمت «ب»، $E(X+Y|X+Y+Z=۳۶)$ را محاسبه کنید.

پاسخ .

الف) مجموع متغیرهای تصادفی مستقل X، Y و Z دارای توزیع پواسون است که پارامتر آن برابر مجموع پارامترهای X، Y و Z است.

$$\begin{aligned} T &= X+Y+Z \sim Poisson(۳\lambda+۶\lambda+۶\lambda) \\ T &= X+Y+Z \sim Poisson(۱۵\lambda) \end{aligned}$$

$$\begin{aligned} P(X=x,Y=y,Z=z|T=۳۶) &= \frac{P(X=x,Y=y,Z=z,T=۳۶)}{P(T=۳۶)} \\ &= \frac{\frac{e^{-۳\lambda}(۳\lambda)^{۳۶-y-z}}{(۳۶-y-z)!}\times\frac{e^{-۶\lambda}(۶\lambda)^y}{y!}\times\frac{e^{-۶\lambda}(۶\lambda)^z}{z!}}{\frac{e^{-۱۵\lambda}(۱۵\lambda)^{۳۶}}{۳۶!}} \\ &= \frac{۳۶!}{(۳۶-y-z)!y!z!}(\frac{۳}{۱۵})^{۳۶-y-z}(\frac{۶}{۱۵})^y(\frac{۶}{۱۵})^z \end{aligned}$$

ب) با کمی دقت متوجه می‌شویم تابع جرم احتمال مشترک شرطی که در قسمت «الف» حساب شد، دارای فرم توزیع چندجمله‌ای است:

$$\frac{n!}{k_۱!k_۲!k_۳!}p_۱^{K_۱}p_۲^{K_۲}p_۳^{K_۳}$$

اگر فرض کنیم $W=X+Y$ و $T=X+Y+Z$، آنگاه:

$$(W|T=۳۶) \sim Bin(۳۶, \underbrace{\frac{۳}{۱۵}}_{p_۱}+\underbrace{\frac{۶}{۱۵}}_{p_۲})$$

پ) با توجه به قسمت «ب» می‌دانیم دارای توزیع دوجمله‌ای با پارامترهای ۳۶ و $\frac{۹}{۱۵}$ است. پس:

$$E(W|T = ۳۶) = ۳۶ \times \frac{۹}{۱۵} = ۲۱/۶$$

سؤال ۹.

تمرین کامپیوتری سری ششم با موضوع «نمونه‌برداری» را می‌توانید از طریق این لینک [1] دریافت کنید.

- یک کپی از فایل مذکور با نام CA6_SID در گوگل درایو خود تهیه کنید.
- در فایل خود بخش‌هایی که به وسیله مستطیل مشخص شده‌اند را با کدهای مناسب جایگزین کنید.
- سوالاتی که به زبان فارسی و رنگ سفید مطرح شده‌اند را در همان سلول پاسخ دهید.
- فایل کد خود را با ایمیل omid.panakari.s@gmail.com با دسترسی Editor به اشتراک بگذارید.
- لینک فایل پاسخ خود را در بخش متنی جایگاه آپلود این تمرین در سامانه ایلرن قرار دهید.

هرگونه انتقاد، پیشنهاد یا نکته جانبی را می‌توانید از طریق یک سلول متنی در ابتدای فایل (قبل از سرفصل اصلی) به ما منتقل کنید.

پاسخ.

تکمیل شده‌ی فایل صورت سوال از طریق این لینک [2] در دسترس است.

[1] https://colab.research.google.com/drive/1X6Df5jymZhPclsvw0pjxCPSRgBwNe_Na?usp=sharing
[2] https://colab.research.google.com/drive/1EWDDDnWgQ_6LF_E6mGY_R-YFZklvo4lq?usp=sharing