آمار و احتمال مهندسی

نیم‌سال اول ۹۸ ــ ۹۹

دانشکده‌ی مهندسی کامپیوتر — گردآورندگان: امیررضا کاظمی، روزبه مشکین، پدرام خرسندی، آیدا رمضانی

تمرین ۴ — توزیع توام، توزیع شرطی و حاشیه‌ای، نامساوی های احتمالاتی

توزیع توام، توزیع شرطی و حاشیه‌ای، نامساوی های احتمالاتی

مسئله‌ی ۱. دست گرمی

سه متغیر تصادفی Y_1، Y_2 و Y_3 از توزیع یکنواخت بین ۰ و ۱ پیروی می‌کنند. اگر این سه متغیر هر کدام طول یک چوپ را نشان دهند، احتمال ساخت مثلث با این سه قطعه چوب را بیابید.

حل. متغیر M ، ماکزیمم سه متغیر فوق؛ و Y_1 و Y_2 دو متغیر دیگر هستند. برای تشکیل مثلث کافیست $M \leqslant Y_1 + Y_2$ باشد. بنابراین،

$$P(M \leqslant Y_1 + Y_2) = \int_0^1 \int_0^m f_M(m) f_{Y_1|M}(y_1|m) P(Y_2 < m - y_1 | m, y_1) dy_1 dm$$
$$\rightarrow P(M \leqslant Y_1 + Y_2) = \int_0^1 \int_0^m m^3 \frac{1}{m}(m - y_1) dy_1 dm$$

$\triangleright$

مسئله‌ی ۲. دست گرمی ۲

سه متغیر تصادفی U_1، U_2 و U_3 از توزیع یکنواخت بین ۰ و ۱ پیروی می‌کنند. متغیرهای L و M به ترتیب مقادیر مینیمم و ماکزیمم این سه متغیر هستند.

الف

تابع توزیع تجمعی و چگالی احتمال L را بیابید.

ب

تابع توزیع تجمعی توام و چگالی احتمال توام L و M را بیابید.

حل.

الف

$$F_L(L < l) = 1 - F_L(L \geqslant l) = 1 - (1 - l)^3$$

در نتیجه، داریم:

$$f_L(l) = 3(1 - l)^2$$

ب

پیشامد $L \geqslant l, M \leqslant m$ معادل با این است که هر سه متغیر در بازه l تا m قرار گیرند.
همچنین میدانیم:

$$P(M < m) = P(M < m, L < l) + P(M < m, L \geqslant l)$$

حال با توجه به اینکه $P(M < m) = m^3$ است، توزیع CDF دو متغیر مذکور به شکل زیر است.

$$F_{M,L}(m, l) = m^3 - (m - l)^3$$

در نتیجه:

$$f_{M,L}(m, l) = 6(m - l)$$

▷

مسئله‌ی ۳. متغیر تصادفی های مستقل

الف

فرض کنید در حال تست کردن ۳ لامپ از شرکت های مختلف هستیم. تاریخ انقضای هر کدام از لامپ های P_1 ، P_2 و P_3 متغیر تصادفی های نمایی با امید ریاضی به ترتیب $\frac{1}{2}$، $\frac{1}{4}$، $\frac{1}{8}$ می‌باشد (بر حسب سال). با فرض اینکه تاریخ انقضای لامپ ها مستقل از هم باشند، اگر T متغیر تصادفی مدت زمانی باشد که هر ۳ لامپ روشن هستند تابع چگالی احتمال آن را بیابید.

ب

متغیر تصادفی X دارای تابع چگالی احتمال $f_X(x)$ می‌باشد. c را به گونه ای بیابید که $E_X[| X - c |]$ کمینه شود.

حل.

الف

راه حل بخش اول
از آنجا که متغیر تصادفی های P_1، P_2 و P_3 مستقل و همگی دارای توزیع نمایی هستند بنابراین خواهیم داشت:

$$P_1 = exp(2),\ P_2 = exp(4),\ P_3 = exp(8)$$

متغیر تصادفی T مدت زمانی است که هر ۳ لامپ روشن هستند بنابراین میتوان T را بر اساس P_1 ، P_2 و P_3 نوشت .

$$T = min(P_1, P_2, P_3)$$

حال برای محاسبه تابع توزیع چگالی از تابع توزیع تجمعی کمک میگیریم که با تعاریف بالا خواهیم داشت:

$$P(T \geqslant t) = P(P_1 \geqslant t, P_2 \geqslant t, P_3 \geqslant t)$$

$$P(T \geqslant t) = P(P_1 \geqslant t)P(P_2 \geqslant t)P(P_3 \geqslant t)$$

به صورت کلی برای توزیع توانی داریم:

$$P(X \geqslant x) = 1 - P(X \leqslant x) = 1 - F_X(x) = 1 - \int_0^x \lambda e^{-\lambda x}\, dx = e^{-\lambda x}$$

در نتیجه $P(T \geqslant t) = e^{-2t}e^{-4t}e^{-8t} = e^{-14t}$ و $F_T(t) = 1 - e^{-14t}$ و تابع توزیع چگالی یعنی همان $f_T(t)$ برابر می شود با:

$$f_T(t) = \frac{\partial F}{\partial t} = 14e^{-14t}.$$

ب

راه حل بخش دوم

$$E[\mid X-c \mid] = \int_{-\infty}^{\infty} \mid X-c \mid\, f_X(x)\, dx = \int_c^{\infty}(X-c)f_X(x)\, dx + \int_{-\infty}^{c}(c-X)f_X(x)\, dx = g(c)$$

$$\frac{\partial g}{\partial c} = 0 \implies \int_c^{\infty}(1)f_X(x)\, dx + \int_{-\infty}^{c}(1)f(X)\, dx = 0 \implies \int_c^{\infty} f_X(x) = \int_{-\infty}^{c} f_X(x)$$

$$\implies c = median(f(X))$$

$\triangleright$

مسئله‌ی ۴. توزیع توام

الف

متغیر تصادفی X را در نظر بگیرید به طوریکه $f_X(x) = kx^۲$ و $۰ \leqslant x \leqslant ۳$ و همچنین متغیر تصادفی Y به این صورت تعریف شده است $Y = X^۳$. حال مقادیر k ، $var(Y)$ را بیابید

ب

اگر برای متغیر تصادفی های مستقل X و Y داشته باشیم: $f_X(x) = Exp(x, ۱)$ و $f_Y(y) = Exp(y, ۱)$ تابع چگالی احتمال $\frac{U}{V}$ را محاسبه کنید به طوریکه $U = min(X, Y)$ و $V = max(X, Y)$

راهنمایی : $\int_x e^{-ax}\, dx = -\frac{(ax + ۱)e^{-ax}}{a^۲}$

حل.

الف

راه حل بخش اول

برای اینکه $kx^۲$ تابع چگالی احتمال باشد باید داشته باشیم:

$$\int_۰^۳ f_X(x)\, dx = ۱ \implies \int_۰^۳ kx^۲\, dx = ۱ \implies ۹k = ۱ \implies k = \frac{۱}{۹}$$

$$E[Y] = E[X^۳] = \int_۰^۳ x^۳ f_X(x)\, dx = \frac{۱}{۹}\int_۰^۳ x^۵\, dx = \frac{۲۷}{۲}$$

$$E[Y^۲] = E[X^۶] = \int_۰^۳ x^۶ f_X(x)\, dx = \frac{۱}{۹}\int_۰^۳ x^۸\, dx = ۳^۵$$

$$var(Y) = E[Y^۲] - E[Y]^۲ = \frac{۲۴۳}{۴}$$

ب

راه حل بخش دوم

$Z = \frac{U}{V}, ۰ \leqslant z \leqslant ۱$ قرار میدهیم . حال ابتدا تابع توزیع تجمعی $F_Z(z)$ را حساب میکنیم.

$$F_Z(z) = P(\frac{U}{V} \leqslant z) \cap P((X \leqslant Y) \cup (X > Y))$$

در تساوی بالا پرانتزی که با تعریف توزیع تجمعی اشتراک گرفته شده است شرط داشتن مینمم برای توزیع های X و Y است . حال با استفاده از خاصیت پخشی داریم :

$$F_Z(z) = (P(\frac{U}{V} \leqslant z) \cap (X \leqslant Y)) \cup (P(\frac{U}{V} \leqslant z) \cap (X > Y))$$

سمت راست تساوی بالا را دقت کنید که در آن دو عبارت با هم اجتماع شده اند ، اگر به هرکدام از این دو عبارت دقت کنیم درمیابیم که اگر شرط های $X \leqslant Y$ و $X > Y$ بخواهند برقرار باشند ، میتوان U و V را با X و Y جاگذاری کرد . بنابراین خواهیم داشت:

$$F_Z(z) = P(\frac{X}{Y} \leqslant z, X \leqslant Y) + P(\frac{Y}{X} \leqslant z, X > Y)$$

$$F_Z(z) = P(X \leqslant Yz, X \leqslant Y) + P(Y \leqslant Xz, X > Y)$$

$$F_Z(z) = \int_0^\infty \int_0^{yz} f_{XY}(x,y)\, dxdy + \int_0^\infty \int_0^{xz} f_{XY}(x,y)\, dydx$$

در این جا از دو طرف نسبت به z مشتق میگیریم و داریم:

$$f_Z(z) = \int_0^\infty y f_{XY}(x,y) dy + \int_0^\infty x f_{XY}(x,y)\, dx$$

با توجه به استقلال دو متغیر X و Y و توزیع آنها و محاسبه انتگرال فوق خواهیم داشت:

$$\implies f_Z(z) = \int_0^\infty y e^{-(yz+y)} dy + \int_0^\infty x e^{-(xz+x)} dx$$

با توجه به راهنمایی موجود در سوال داریم:

$$\int_x e^{-ax}\, dx = -\frac{(ax+1)e^{-ax}}{a^2}$$

$$\implies f_Z z = 2\frac{1}{(z+1)^2}$$

$$f_Z(z) = \begin{cases} \dfrac{2}{(1+z)^2} & 0 \leqslant z \leqslant 1 \\ 0 & ow \end{cases}$$

$\triangleright$

مسئله‌ی ۵. شمارش تخم مرغی

متغیر تصادفی N تعداد تخم مرغ های یک مزرعه است که از توزیع پوآسون با متغیر λ پیروی می‌کند. هر تخم مرغ با احتمال p به جوجه تبدیل می‌شود و شکستن تخم مرغ ها از هم مستقل است . متغیر تصادفی X تعداد تخم مرغ هایی هستند که به جوجه تبدیل می‌شوند و متغیر تصادفی Y تعداد آنهایی است که تبدیل نمی‌شوند . توزیع توام را برای دو متغیر تصادفی X و Y بدست آورید.

حل. توضیح کلی راه حل
ابتدا بدیهیاتی که از سوال قابل استنتاج است را می‌نویسیم:

$$X + Y = N,\ N\ Poisson(\lambda),\ X|N = Bin(N,p)$$

حال با استفاده از تعریف سوال را حل می‌کنیم:

$$P(X=i,Y=j) = \Sigma_{n=۰}^{\infty} P(X=i,Y=j|N=n)P(N=n)$$

$$= P(X=i,Y=j|N=i+j)P(N=i+j)$$

$$= P(X=i|N=i+j)P(N=i+j)$$

$$= \frac{(i+j)!}{i!j!}p^i(۱-p)^j e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}$$

$$= e^{-\lambda p}\frac{(\lambda p)^i}{i!}e^{-\lambda(۱-p)}\frac{(\lambda(۱-p))^j}{j!}$$

$\triangleright$

مسئله‌ی ۶. جمع متغیرهای تصادفی

فرض کنید x یک متغیر تصادفی uniform باشد به طوری که

$$f_X(x) = \begin{cases} ۱ & ۰ \leqslant x \leqslant ۱ \\ ۰ & \text{otherwise} \end{cases}$$

و y یک متغیر تصادفی exponential و مستقل از x باشد که

$$f_Y(y) = \begin{cases} e^{-y} & y \geqslant ۰ \\ ۰ & \text{otherwise} \end{cases}$$

همچنین فرض کنید

$$Z = X + Y$$

الف

تابع توزیع Z را محاسبه کنید.

ب

تابع توزیع تجمعی Z را محاسبه کنید.

حل. ابتدا تابع توزیع احتمال Z را محاسبه می‌کنیم. می‌دانیم تابع توزیع Z برابر است با convo-lution تابعهای توزیع X و Y. از طرفی با استفاده از شهود نموداری convolution می‌توانیم تشخیص دهیم که تابع توزیع Z برای $z < ۰$ برابرِ صفر است. حالا برای محاسبه تابع توزیع برای مقادیر بزرگتر از صفر داریم:

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\, dy \\
&= \int_{۰}^{\infty} f_X(z-y) e^{-y}\, dy
\end{aligned}
$$

از طرفی می‌دانیم زمانی $f_X(z-y)$ برابرِ ۱ است که

$$
\begin{aligned}
۰ \leqslant z-y \leqslant ۱ \equiv &\quad -۱ \leqslant y-z \leqslant ۰ \\
\equiv &\quad z-۱ \leqslant y \leqslant z
\end{aligned}
$$

بنابراین داریم:

$$
\begin{aligned}
f_Z(z) &= \int_{max(z-۱,۰)}^{z} e^{-y}\, dy \\
&= -e^{-y}|_{max(۰,z-۱)}^{z} \\
&= -e^{-z} + e^{-max(۰,z-۱)} \\
&= \begin{cases} e^{۱-z} - e^{-z} & z \geqslant ۱ \\ ۱ - e^{-z} & z < ۱ \end{cases}
\end{aligned}
$$

الف

طبق آنچه گفته شد پاسخ این قسمت برابر است با:

$$
f_Z(z) = \begin{cases} e^{۱-z} - e^{-z} & z \geqslant ۱ \\ ۱ - e^{-z} & z < ۱ \end{cases}
$$

ب

برای محاسبه تابع توزیع تجمعی ۳ حالت داریم. در حالت $z < ۰$ تابع توزیع تجمعی برابرِ صفر است. در حالت $۰ \leqslant z < ۱$:

$$
\begin{aligned}
F_Z(z) &= \int_{۰}^{z} ۱ - e^{-w}\, dw \\
&= z + e^{-w}|_{۰}^{z} \\
&= z + e^{-z} - ۱
\end{aligned}
$$

در حالت $z \geqslant ۱$ داریم:

$$
\begin{aligned}
F_Z(z) &= \int_{۱}^{z} e^{۱-w} - e^{-w}\, dw \\
&= -e^{۱-w}|_{۱}^{z} + e^{-w}|_{۱}^{z}
\end{aligned}
$$

$\triangleright$

مسئله‌ی ۷. جمع متغیرهای تصادفی گسسته

فرض کنید X یک متغیر تصادفی گسسته و $R_x = \{۰, ۱, ۲\}$. تابع توزیع X را اینگونه تعریف می‌کنیم:

$$P_X(x) = \begin{cases} ۱/۳ & x \in R_X \\ ۰ & \text{otherwise} \end{cases}$$

همچنین فرض کنید Y متغیر تصادفی گسسته‌ای باشد که مستقل از X است و $R_Y = \{۱, ۲\}$ تابع توزیع Y را اینگونه تعریف می‌کنیم:

$$P_Y(y) = \begin{cases} \frac{۱}{۳}y & y \in R_Y \\ ۰ & \text{otherwise} \end{cases}$$

تابع توزیعِ $Z = X + Y$ را بیابید.

حل. به طور کلی می‌توانیم تشخص دهیم که تابع توزیع Z تنها برای مقادیرِ

$$R_Z = \{۱, ۲, ۳, ۴\}$$

ناصفر است.
حالا تابع توزیع Z را در هر یک از این مقادیر محاسبه می‌کنیم.
برای $z = ۱$ داریم:

$$\begin{aligned} P_Z(۱) &= \sum_{y \in R_y} P_X(۱ - y) P_Y(y) \\ &= P_X(۱ - ۱) P_Y(۱) + P_X(۱ - ۲) P_Y(۲) \\ &= \frac{۱}{۳} \times \frac{۱}{۳} + ۰ \times \frac{۲}{۳} \\ &= \frac{۱}{۹} \end{aligned}$$

برای $z = ۲$ داریم:

$$\begin{aligned} P_Z(۲) &= \sum_{y \in R_Y} P_X(۲ - y) P_Y(y) \\ &= P_X(۲ - ۱) P_Y(۱) + P_X(۲ - ۲) P_Y(۲) \\ &= \frac{۱}{۳} \times \frac{۱}{۳} + \frac{۱}{۳} \times \frac{۲}{۳} \\ &= \frac{۱}{۳} \end{aligned}$$

در حالت $z = ۳$ داریم:

$$\begin{aligned} P_Z(۳) &= \sum_{y \in R_Y} P_X(۳ - y) P_Y(y) \\ &= P_X(۳ - ۱)P_Y(۱) + P_X(۳ - ۲)P_Y(۲) \\ &= \frac{۱}{۳} \times \frac{۱}{۳} + \frac{۱}{۳} \times \frac{۲}{۳} \\ &= \frac{۱}{۳} \end{aligned}$$

و در حالت $z = ۴$ داریم:

$$\begin{aligned} P_Z(۴) &= \sum y \in R_Y P_X(۴ - y) P_Y(y) \\ &= P_X(۴ - ۱)P_Y(۱) + P_X(۴ - ۲)P_Y(۲) \\ &= ۰ \times \frac{۱}{۴} + \frac{۱}{۳} \times \frac{۲}{۳} \\ &= \frac{۲}{۹} \end{aligned}$$

بنابراین، تابع توزیع متغیر تصادفی Z عبارتست از:

$$P_Z(z) = \begin{cases} \frac{۱}{۹} & z = ۱ \\ \frac{۱}{۳} & z = ۲ \\ \frac{۱}{۳} & z = ۳ \\ \frac{۲}{۹} & z = ۴ \\ ۰ & \text{otherwise} \end{cases}$$

$\triangleright$

مسئله‌ی ۸. تابع توزیع متغیرهای تصادفی مستقل

ثابت کنید متغیرهای $X_۱$ ، ... ، X_n مستقلند اگر و تنها اگر تابع توزیع توام آن‌ها به شکل زیر قابل بیان باشد.

$$f(x_۱, x_۲, ..., x_n) = \prod_{i=۱}^{n} g_i(x_i)$$

g_i تابعی مثبت است.

حل. با استفاده از استقرا قابل است که:

$$\int_{-\infty}^{\infty} g_i(x_i) dx_i = ۱$$

. برای اثبات حکم روی n استقرا میزنیم.
پایه استقرا: برای $n = ۲$ میدانیم، اگر $x_۱$ و $x_۲$ مستقل باشند، آنگاه؛

$$f(x_1, x_2) = f_1(x_1)f_2(x_2)$$

است. در نتیجه توزیع توام این دو متغیر به شکل $g_1(x_1)g_2(x_2)$ قابل بیان است.
همچنین اگر $f(x_1, x_2) = g_1(x_1)g_2(x_2)$ داریم:

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_1 = \int_{-\infty}^{\infty} g_2(x_2)g_1(x_1)dx_1 = g_2(x_2)$$
$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2 = \int_{-\infty}^{\infty} g_1(x_1)g_2(x_2)dx_2 = g_1(x_1)$$
$$\to f(x_1, x_2) = g_1(x_1)g_2(x_2) = f_1(x_1)f_2(x_2)$$

بنابراین دو متغیر x_1 و x_2 مستقلند. گام استقرا: فرض کنید متغیرهای $x_1, ..., x_n, x_{n+1}$ از یکدیگر مستقل باشند. حال با توجه به فرض استقرا؛

$$f(x_1, ..., x_n | x_{n+1}) = f(x_1, ..., x_n)f_{n+1}(x_{n+1}) = \prod_{i=1}^{n} g_i(x_i)f_{n+1}(x_{n+1})$$

همانطور که مشاهده شد فرم مذکور در صورت سوال قابل مشاهده است.
حال فرض میکنیم، $f(x_1, ..., x_n, x_{n+1}) = \prod_{i=1}^{n+1} g_i(x_i)$ است.
با انتگرالگیری روی x_{n+1} داریم:

$$f(x_1, ..., x_n) = \prod_{i=1}^{n} g_i(x_i)$$

باتوجه به فرض استقرا متغیرهای $x_1, ..., x_n$ از یکدیگر مستقلند. میتوان معادله‌ی بالا را فرم زیر نوشت.

$$f(x_1, ..., x_n, x_{n+1}) = f(x_1, ..., x_n)g_{n+1}(x_{n+1})$$

با انتگرالگیری روی x_1 تا x_n نتیجه میشود:

$$f_{n+1}(x_{n+1}) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, ..., x_n, x_{n+1})dx_1...dx_n = g_{n+1}(x_{n+1})$$
$$\to f(x_1, ..., x_n, x_{n+1}) = f(x_1, ..., x_n)f_{n+1}(x_{n+1})$$

در نتیجه متغیرهای مستقلند. $\triangleright$

مسئله‌ی ۹. نامساوی مار کوف

فرض کنید متغیر تصادفی X از یک توزیع $Binomial(n, p)$ می‌آید.

الف

اگر $p < \alpha < 1$ ، با استفاده از نابرابری مارکوف، یک کران بالا برای $P(X \geqslant \alpha n)$ بیابد.

ب

مقدار عددی این کران را برای $p = \frac{۱}{۲}$ و $\alpha = \frac{۳}{۴}$ بیابید.

حل. با توجه به اینکه X یک متغیر تصادفی مثبت و $E(X) = np$ است، می‌توانیم از نابرابری مارکوف استفاده کنیم.

الف

داریم:

$$P(X \geqslant \alpha n) \leqslant \frac{E(X)}{\alpha n} = \frac{pn}{\alpha n} = \frac{p}{\alpha}$$

ب

با جایگذاری $p = \frac{۱}{۲}$ و $\alpha = \frac{۳}{۴}$ داریم:

$$P(X \geqslant \frac{۳n}{۴}) \leqslant \frac{۲}{۳}$$

$\triangleright$

مسئله‌ی ۱۰. نامساوی چبیشف

فرض کنید متغیر تصادفی X از یک توزیع $Binomial(n, p)$ می‌آید.

الف

با فرض اینکه $p < \alpha < ۱$ و استفاده از نابرابری چبیشف، یک کران بالا برای $P(X \geqslant \alpha n)$ بیابید.

ب

مقدار کران بالا را برای $p = \frac{۱}{۲}$ و $\alpha = \frac{۳}{۴}$ تعیین کنید.

حل.

الف

می‌توانیم برای محاسبه کران بالا بنویسیم:

$$\begin{aligned} P(X \geqslant \alpha n) &= P(X - np \geqslant \alpha n - np) \\ &\leqslant P(|X - np| \geqslant n\alpha - np) \\ &\leqslant \frac{Var(X)}{(n\alpha - np)^{۲}} \\ &= \frac{p(۱ - p)}{n(\alpha - p)^{۲}} \end{aligned}$$

ب

با جایگذاری $p = \frac{۱}{۲}$ و $\alpha = \frac{۳}{۴}$ داریم:

$$P(X \geqslant \frac{۳n}{۴}) \leqslant \frac{۴}{n}$$

$\triangleright$

مسئله‌ی ۱۱. تراشه ها

حدود ۲ درصد از تراشه‌های RAM تولید شده در یک کارخانه خراب است. علی به ۵۰ عدد RAM سالم برای آزمایشگاهش نیاز دارد. علی باید چه تعداد تراشه RAM خریداری کند تا مطمئن باشد با احتمال حداقل ۹۹ درصد، حداقل ۴۹ تراشه سالم RAM دارد؟

حل. طبق نابرابری مارکوف داریم:

$$P(X \geqslant ۴۹) \leqslant \frac{E(X)}{۴۹}$$

از طرفی،

$$P(X \geqslant ۴۹) \geqslant \frac{۹۹}{۱۰۰}$$

پس

$$\begin{aligned} &\frac{E(X)}{۴۹} \geqslant \frac{۹۹}{۱۰۰} \\ &\Rightarrow \frac{N\frac{۹۸}{۱۰۰}}{۴۹} \geqslant \frac{۹۹}{۱۰۰} \\ &\Rightarrow \frac{۲N}{۱۰۰} \geqslant \frac{۹۹}{۱۰۰} \\ &\Rightarrow N \geqslant ۵۰ \end{aligned}$$

بنابراین حداقل ۵۰ قطعه RAM باید خریداری شود. $\triangleright$

موفق باشید