بسمه تعالی

آمار و احتمال مهندسی

نیم‌سال اول ۱۳۹۹-۱۴۰۰
مدرس: سید ابوالفضل مطهری

دانشکده مهندسی کامپیوتر

پاسخ تمرین چهارم

سوال اول)

ابه طور کلی می‌توانیم تشخص دهیم که تابع توزیع Z تنها برای مقادیرِ

$$R_Z = \{1, 2, 3, 4\}$$

ناصفر است.

حالا تابع توزیع Z را در هر یک از این مقادیر محاسبه می‌کنیم.

برای $z = 1$ داریم:

$$\begin{aligned} P_Z(1) &= \sum_{y \in R_y} P_X(1-y)P_Y(y) \\ &= P_X(1-1)P_Y(1) + P_X(1-2)P_Y(2) \\ &= \frac{1}{3} \times \frac{1}{3} + 0 \times \frac{2}{3} \\ &= \frac{1}{9} \end{aligned}$$

برای $z = 2$ داریم:

$$\begin{aligned} P_Z(2) &= \sum_{y \in R_Y} P_X(2-y)P_Y(y) \\ &= P_X(2-1)P_Y(1) + P_X(2-2)P_Y(2) \\ &= \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{2}{3} \\ &= \frac{1}{3} \end{aligned}$$

در حالت $z=3$ داریم:

$$\begin{aligned}P_Z(3) &= \sum_{y\in R_Y} P_X(3-y)P_Y(y)\\ &= P_X(3-1)P_Y(1)+P_X(3-2)P_Y(2)\\ &= \frac{1}{3}\times\frac{1}{3}+\frac{1}{3}\times\frac{2}{3}\\ &= \frac{1}{3}\end{aligned}$$

و در حالت $z=4$ داریم:

$$\begin{aligned}P_Z(4) &= \sum y\in R_Y P_X(4-y)P_Y(y)\\ &= P_X(4-1)P_Y(1)+P_X(4-2)P_Y(2)\\ &= 0\times\frac{1}{4}+\frac{1}{3}\times\frac{2}{3}=\frac{2}{9}\end{aligned}$$

بنابراین، تابع توزیع متغیر تصادفی Z عبارتست از:

$$P_Z(z)=\begin{cases}\frac{1}{9} & z=1\\ \frac{1}{3} & z=2\\ \frac{1}{3} & z=3\\ \frac{2}{9} & z=4\\ 0 & \text{otherwise}\end{cases}$$

سوال دوم

تعریف می‌کنیم:

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

ابتدا نشان می‌دهیم به ازای هر n در معادله $f(x_1,...,x_n)=\prod_{i=1}^{n} g_i(x_i)$ داریم:

$$\prod_{i=1}^{n} k_i = 1$$

اثبات:

$$1=\int_{-\infty}^{\infty} f(x_1,...,x_n)dx_1...dx_n=\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} g_1(x_1)...g_n(x_n)dx_1...dx_n=\prod_{i=1}^{n} k_i$$

سپس برای اثبات حکم روی n استقرا میزنیم.

پایه استقرا: برای $n=2$ میدانیم، اگر x_1 و x_2 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:

$$f_2(x_2)=\int_{-\infty}^{\infty} f(x_1,x_2)dx_1=\int_{-\infty}^{\infty} g_2(x_2)g_1(x_1)dx_1=g_2(x_2)\times\int_{-\infty}^{\infty} g_1(x_1)dx_1=k_1\times g_2(x_2)=g_2(x_2)/k_2$$
$$f_1(x_1)=\int_{-\infty}^{\infty} f(x_1,x_2)dx_2=\int_{-\infty}^{\infty} g_2(x_2)g_1(x_1)dx_1=g_1(x_1)\times\int_{-\infty}^{\infty} g_2(x_2)dx_2=k_2\times g_1(x_1)=g_1(x_1)/k_1$$
$$\to f(x_1,x_2)=g_1(x_1)g_2(x_2)=k_1k_2f_1(x_1)f_2(x_2)=f_1(x_1)f_2(x_2)$$

بنابراین دو متغیر x_1 و x_2 مستقلند.

قسمت اول: گام استقرا: ثابت می‌کنیم اگر برای n متغیر مستقل $x_1,...,x_n$ داشته باشیم $f(x_1,...,x_n)=\prod_{i=1}^{n} f_i(x_i)$ آنگاه برای $n+1$ متغیر مستقل $x_1,...,x_n,x_{n+1}$ خواهیم داشت: $f(x_1,...,x_n,x_{n+1})=\prod_{i=1}^{n+1} f_i(x_i)$. داریم:

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

(همانطور که مشاهده می‌شود فرم مذکور در صورت سوال قابل مشاهده است ($f_i=g_i$)).

قسمت دوم: گام استقرا: می‌خواهیم ثابت کنیم اگر: $f(x_1,...,x_n,x_{n+1})=\prod_{i=1}^{n+1} g_i(x_i)$ آنگاه متغیرهای $x_1,...,x_n,x_{n+1}$ از هم مستقلند.

با انتگرالگیری روی x_{n+1} به دست می‌آید:

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

بنابراین باتوجه به فرض استقرا متغیرهای $x_1,...,x_n$ از یکدیگر مستقلند(عدد ثابت k_{n+1} را می‌توان به هرکدام از توابع نسبت داد.)

بنابراین اکنون کافیست استقلال x_{n+1} را از بقیه متغیرها ثابت کنیم:

$$f_{n+1}(x_{n+1})=\int_{-\infty}^{\infty} f(x_1,...,x_n,x_{n+1})dx_1...dx_n=\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} g_1(x_1)...g_n(x_n)g_{n+1}(x_{n+1})dx_1...dx_n=$$
$$k_1k_2...k_ng_{n+1}(x_{n+1})=g_{n+1}(x_{n+1})/k_{n+1}\ (*)$$
$$f(x_1,...,x_n)=\int_{-\infty}^{\infty} f(x_1,...,x_n,x_{n+1})dx_{n+1}=\int_{-\infty}^{\infty} g_1(x_1)...g_n(x_n)g_{n+1}(x_{n+1})dx_{n+1}=$$
$$k_{n+1}\prod_{i=1}^{n} g_i(x_i)\ (**)$$
$$(*),(**)\to f(x_1,...,x_n,x_{n+1})=\prod_{i=1}^{n+1} g_i(x_i)=\prod_{i=1}^{n} g_i(x_i)\times g_{n+1}(x_{n+1})=f(x_1,...,x_n)f_{n+1}(x_{n+1})$$

در نتیجه همه‌ی متغیرهای مستقلند.

سوال سوم

الف) مانند حالت دو بعدی که توزیع یکنواخت به معنی این است که احتمال با مساحت متناسب است، در سه بعد، نقش مساحت را حجم ایفا می‌کند:

$$Volume(A) = \frac{4\pi}{3} r^3 \tag{۱}$$

بنابراین تابع توزیع PDF توأم این سه متغیر برابر است با:

$$f(x,y,z) = \begin{cases} \frac{1}{Volume(A)} & (x,y,z) \in Volume(A) \\ 0 & o.w. \end{cases} = \begin{cases} \frac{3}{4\pi r^3} & x^2+y^2+z^2 \leq r^2 \\ 0 & o.w. \end{cases}$$

ب) در اینجا توزیع توأم سه متغیر را داریم و به دنبال توزیع دوتای آن‌ها هستیم. بنابراین باید توزیع حاشیه‌ای را در حالتی که می‌خواهیم Z را از توزیع توأم حذف کنیم، محاسبه کنیم.

$$f_{X,Y}(x,y) = \int_{-\infty}^{\infty} f(x,y,z)\,dz = \frac{3}{4\pi r^3}\int_{-\sqrt{r^2-x^2-y^2}}^{\sqrt{r^2-x^2-y^2}} dz = \frac{3}{2\pi r^3}\sqrt{r^2-x^2-y^2}$$

ج) می‌توان از نتایج قبل استفاده کرد. مشابه قبل در حالتی که $-r \leq x \leq r$ تابع چگالی حاشیه‌ای X از تابع چگالی توأم X و Y برابر است با:

$$f_X(x) = \frac{3}{2\pi r^3}\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \sqrt{r^2-x^2-y^2}\,dy$$

سوال چهارم

۱.

$$\operatorname{cov}(X,Y|Z) = \mathbb{E}[(X-\mathbb{E}[X|Z])(Y-\mathbb{E}[Y|Z])|Z] =$$

$$\mathbb{E}[(XY - X\mathbb{E}[Y|Z] - Y\mathbb{E}[X|Z] + \mathbb{E}[X|Z]\mathbb{E}[Y|Z])|Z] =$$

$$\mathbb{E}[XY|Z] - \mathbb{E}[X\mathbb{E}[Y|Z]|Z] - \mathbb{E}[Y\mathbb{E}[X|Z]|Z] + \mathbb{E}[\mathbb{E}[X|Z]\mathbb{E}[Y|Z]|Z] =$$

$$\mathbb{E}[XY|Z] - \mathbb{E}[X|Z]\mathbb{E}[Y|Z] - \mathbb{E}[Y|Z]\mathbb{E}[X|Z] + \mathbb{E}[X|Z]\mathbb{E}[Y|Z] = \mathbb{E}[XY|Z] - \mathbb{E}[X|Z]\mathbb{E}[Y|Z]$$

۲. از عبارت ۱ نسبت به Z امید می‌گیریم. همچنین می‌دانیم که $\text{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

$$\mathbb{E}[\text{cov}(X,Y|Z)] = \mathbb{E}[\mathbb{E}[XY|Z]] - \mathbb{E}[\mathbb{E}[X|Z]\mathbb{E}[Y|Z]] = \mathbb{E}[XY] - \mathbb{E}[\mathbb{E}[X|Z]\mathbb{E}[Y|Z]]$$

$$\text{cov}(\mathbb{E}[X|Z],\mathbb{E}[Y|Z]) = \mathbb{E}[\mathbb{E}[X|Z]\mathbb{E}[Y|Z]] - \mathbb{E}[\mathbb{E}[X|Z]]\mathbb{E}[\mathbb{E}[Y|Z]] = \mathbb{E}[\mathbb{E}[X|Z]\mathbb{E}[Y|Z]] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\Rightarrow \mathbb{E}[\text{cov}(X,Y|Z)] + \text{cov}(\mathbb{E}[X|Z],\mathbb{E}[Y|Z]) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \text{cov}(X,Y)$$

۳. می‌دانیم که $\text{cov}(X,X) = \text{var}(X)$. کافیست در عبارت ۲ به جای Y قرار دهیم X.

$$\text{var}(X) = \text{cov}(X,X) = \mathbb{E}[\text{cov}(X,X|Z)] + \text{cov}(\mathbb{E}[X|Z],\mathbb{E}[X|Z]) = \mathbb{E}[\text{var}(X|Z)] + \text{var}[\mathbb{E}[X|Z]]$$

سوال پنجم

۱.

$$\mathbb{E}[\mathbb{E}[Y|X,Z]|Z=z] = \Sigma_x \mathbb{E}[Y|X=x,Z=z]P(X=x|Z=z) =$$

$$\Sigma_x \Sigma_y y \times P(Y=y|X=x,Z=z)P(X=x|Z=z) =$$

$$\Sigma_{x,y} y \times \frac{P(Y=y,X=x,Z=z)P(X=x,Z=z)}{P(X=x,Z=z)P(Z=z)} =$$

$$\Sigma_{x,y} y \times \frac{P(Y=y,X=x,Z=z)}{P(Z=z)} = \Sigma_y y \times \frac{P(Y=y,Z=z)}{P(Z=z)} = \Sigma_y y \times P(Y=y|Z=z) = \mathbb{E}[Y|Z=z]$$

۲.

$$\mathbb{E}[XY|X=x] = \mathbb{E}[xY|X=x] = x\mathbb{E}[Y|X=x]$$

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[XY|X]] = \mathbb{E}[X\mathbb{E}[Y|X]]$$

سوال ششم

متغیر a_k را نشان‌دهنده تغییر وضعیت (تغییر تعداد) باکتری k در هر مرحله تغییر نسل در نظر می‌گیریم:

$$a_k = \begin{cases} +1 & probability = p \\ 0 & probability = q \\ -1 & probability = 1-p-q \end{cases}$$

به دست می‌آید:

$$\mathbb{E}[a_k] = p + 0 - (1-p-q) = 2p+q-1$$

$$\mathbb{E}[a_k^2] = p + 0 + (1-p-q) = 1-q$$

$$\implies \operatorname{var}(a_k) = \mathbb{E}[a_k^2] - \mathbb{E}[a_k]^2 = 1-q-(2p+q-1)^2 = v$$

حال برای سادگی محاسبات، امید و واریانس تعداد باکتری‌ها در هر مرحله را به صورت شرطی نسبت به مرحله قبل به دست می‌آوریم:

$$\mathbb{E}[X_i|X_{i-1}] = \mathbb{E}[X_{i-1} + \Sigma_{k=1}^{X_{i-1}} a_k|X_{i-1}] = X_{i-1} + \mathbb{E}[\Sigma_{k=1}^{X_{i-1}} a_k] = X_{i-1} + \Sigma_{k=1}^{X_{i-1}} \mathbb{E}[a_k] =$$

$$X_{i-1} + \Sigma_{k=1}^{X_{i-1}}(2p+q-1) = X_{i-1} + (2p+q-1)X_{i-1} = (2p+q)X_{i-1}$$

$$\operatorname{var}[X_i|X_{i-1}] = \operatorname{Var}[X_{i-1} + \Sigma_{k=1}^{X_{i-1}} a_k|X_{i-1}] =$$

$$\operatorname{Var}[\Sigma_{k=1}^{X_{i-1}} a_k|X_{i-1}] = \operatorname{var}[\Sigma_{k=1}^{X_{i-1}} a_k] = \Sigma_{k=1}^{X_{i-1}} \operatorname{Var}[a_k] = \Sigma_{k=1}^{X_{i-1}} v = vX_{i-1}$$

حال برای محاسبه امید ریاضی و واریانس X_i (به صورت غیر شرطی) داریم:

$$\mathbb{E}[X_i] = \mathbb{E}[\mathbb{E}[X_i|X_{i-1}]] = \mathbb{E}[(2p+q)X_{i-1}] = (2p+q)\mathbb{E}[X_{i-1}] \implies \mathbb{E}[X_i] = (2p+q)^i\mathbb{E}[X_0] = (2p+q)^i X_0$$

$$\operatorname{Var}(X_i) = \operatorname{Var}(\mathbb{E}[X_i|X_{i-1}]) + \mathbb{E}[\operatorname{var}(X_i|X_{i-1})] =$$

$$\operatorname{Var}((2p+q)X_{i-1}) + \mathbb{E}[vX_{i-1}] = (2p+q)^2\operatorname{Var}(X_{i-1}) + v(2p+q)^{i-1}X_0$$

$$\Rightarrow \operatorname{Var}(X_i) = (2p+q)^2\operatorname{Var}(X_{i-1}) + v(2p+q)^{i-1}X_0$$

با حل رابطه بازگشتی بالا به دست می‌آید:

$$\mathrm{var}(X_i) = (2p+q)^{(2i)}\mathrm{var}(X_0) + vX_0\Sigma_{j=i-1}^{2i-2}(2p+q)^j = vX_0\Sigma_{j=i-1}^{2i-2}(2p+q)^j = vX_0(2p+q)^{i-1}\frac{1-(2p+q)^i}{1-2p-q}$$

موفق باشید.