۱- فرض کنید X و Y دو متغیر تصادفی باشند به گونه‌ای که $\mathbb{E}\{X|Y\} = -Y$. تعریف می‌کنیم: $Z = X + Y$.

الف) $\mathrm{Cov}(Z, Y)$ را حساب کنید. **[۴ نمره]**

$$Cov(Z,Y) = E[ZY] - E[Z]E[Y] = E[(X+Y)Y] - E[X+Y]E[Y]$$

$$E[X+Y] = E\big[E[X+Y|Y]\big] = E\big[E[X|Y] + E[Y|Y]\big] = E[-Y+Y] = E[0] = 0$$

$$E[(X+Y)Y] = E\big[E[(X+Y)Y|Y]\big] = E[Y\,E[X+Y|Y]] = E[Y(0)] = E[0] = 0$$

$$\rightarrow Cov(Z,Y) = 0$$

ب) اگر بدانیم X و Y متغیرهای تصادفی توأماً نرمال نیز هستند، بهترین تخمین Z برحسب X را با معیار LMSE و خطای متناظر با آن را به‌دست آورید. **[۴ نمره]**

$$E[Z|X] = E[X+Y|X] = X + E[Y|X]$$

$$(X,Y)\sim Normal \rightarrow \; f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi\sigma_Y^2(1-r^2)}}\exp\left(-\frac{1}{2\,\sigma_Y^2(1-r^2)}\left(y-\mu_Y - r\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)^2\right)$$

$$\rightarrow E[Y|X] = \int_{-\infty}^{\infty} y\, f_Y(y|x)dy \;= \mu_Y + r\frac{\sigma_Y}{\sigma_X}(X-\mu_X)$$

$$E[X|Y] = \mu_X + r\frac{\sigma_X}{\sigma_Y}(Y-\mu_Y) = -Y$$

$$\rightarrow r = -\frac{\sigma_Y}{\sigma_X}\;, \qquad \mu_X = -\mu_Y$$

$$\rightarrow E[Y|X] = \mu_y + r\frac{\sigma_Y}{\sigma_X}(X-\mu_X) = -\mu_X - \left(\frac{\sigma_Y}{\sigma_X}\right)^2(X-\mu_X)$$

$$\rightarrow E[Z|X] = \left(1 - \left(\frac{\sigma_Y}{\sigma_X}\right)^2\right)(X-\mu_X)$$

$$e = E[(Z - E[Z|X])^2] = E[(Y - E[Y|X])^2] =$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(Y + \mu_X + \left(\frac{\sigma_Y}{\sigma_X}\right)^2(X-\mu_X)\right)^2 f_{XY}(x,y)dx\,dy$$

که $f_{XY}(x,y)$ تابع توزیع توام متغیرهای تصادفی X و Y است:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}\exp\left(-\frac{1}{2(1-r^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2r(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]\right)$$

۲- فرض کنید X و Y دو متغیر تصادفی با توزیع مشترک (توأم) زیر باشند:

$$f_{XY}(x,y)=\begin{cases} x(y-x)e^{-y}, & 0\le x\le y<\infty \\ 0, & \text{در غیر این صورت} \end{cases}$$

الف) $f_{X|Y}(x|y)$ را محاسبه کنید. **[۴ نمره]**

$$f_Y(y)=\int_0^y x(y-x)e^{-y}dx=e^{-y}\left(\frac{1}{2}x^2y-\frac{1}{3}x^3\right)\ |_0^y=\frac{1}{6}y^3e^{-y}\ ,\qquad 0\le y<\infty$$

$$f_X(x|y)=\frac{f_{XY}(x,y)}{f_Y(y)}=\frac{6x(y-x)}{y^3}\ ,\qquad 0\le x\le y$$

ب) تخمین بهینه X برحسب Y با معیار LMSE و خطای متناظر با آن را بیابید. **[۶ نمره]**

$$E[X|Y=y]=\int_0^y x\ f_X(x|y)dx=\frac{6}{y^3}\int_0^y(x^2y-x^3)dx\ =\frac{6}{y^3}\left(\frac{1}{3}x^3y-\frac{1}{4}x^4\right)|_0^y$$

$$=\frac{6}{y^3}\left(\frac{1}{12}y^4\right)=\frac{y}{2}$$

$$\rightarrow E[X|Y]=\frac{1}{2}Y$$

$$e=E[(X-E[X|Y])^2]=E\left[\left(X-\frac{1}{2}Y\right)^2\right]=E\left[X^2-XY+\frac{1}{4}Y^2\right]=$$

$$\int_0^\infty\int_0^y x^3(y-x)e^{-y}\ dx\ \ dy-\int_0^\infty\int_0^y x^2y(y-x)e^{-y}\ dx\ \ dy+\frac{1}{4}\int_0^\infty\int_0^y xy^2(y-x)e^{-y}\ dx\ \ dy=$$

$$=\frac{1}{120}\int_0^\infty y^5e^{-y}\ dy=\frac{1}{120}(-y^5-5y^4-20y^3-60y^2-120y-120)e^{-y}|_0^y$$

$$=\frac{1}{120}(0-(-120))=1$$

۳- یک فروشگاه زنجیره‌ای قصد دارد دو شعبه در غرب و شرق تهران راه‌اندازی کند. تعداد کل مشتریان فروشگاه در هر روز یک متغیر تصادفی پوآسن با پارامتر a درنظر گرفته می‌شود. با آنالیز مسیریابی انجام‌شده و با درنظر گرفتن قانون اعداد بزرگ و کمی اغماض، مشخص شده هر یک از مشتریان با احتمال p شعبه غرب و با احتمال $q = 1 - p$، مستقل از سایرین، شعبه شرق را انتخاب می‌کند. اگر میزان خرید هر مشتری در این فروشگاه بزرگ مستقل از سایر مشتریان، متغیر تصادفی نمایی با پارامتر λ درنظر گرفته شود:

الف) میانگین و واریانس فروش کل در تهران را بیابید. **[۴ نمره]**

تعداد مشتری‌ها: $N \sim Poi(a)$

میزان خرید مشتری i-ام: $X_i \sim \mathrm{Exp}(\lambda)$

$$S = X_1 + X_2 + \cdots + X_N$$

S یک مجموع تصادفی است:

$$E[S] = E[X_i]E[N] = \frac{a}{\lambda}$$

$$Var[S] = E[N]Var[X_i] + Var[N](E[X_i])^2 = \frac{2a}{\lambda^2}$$

ب) تابع چگالی احتمال و میانگین فروش شعبه غرب را بیابید. **[۴ نمره]**

تعداد مشتریان شعبه غرب: W

$$P(W = w) = \sum_{n=0}^{\infty} P(W = w|N = n)P(N = n)$$

$$P(W = w|N = n) \sim Bin(n, p)$$

$$P(W = w) = \sum_{n=w}^{\infty} \binom{n}{w} p^w (1-p)^{n-w} e^{-a} \frac{a^n}{n!} =$$

$$= e^{-a} \sum_{n=w}^{\infty} \frac{n!}{w!\,(n-w)!} (ap)^w \big(a(1-p)\big)^{n-w} \frac{1}{n!} =$$

$$= \frac{e^{-a}}{w!} (ap)^w \sum_{n=w}^{\infty} \frac{1}{(n-w)!} \big(a(1-p)\big)^{n-w} = \frac{e^{-a}}{w!} (ap)^w e^{a(1-p)} = \frac{e^{-ap}(ap)^w}{w!} \sim Poi(ap)$$

مشابه با بخش الف:

$$E[W] = \frac{ap}{\lambda}\ , \qquad Var(W) = \frac{2ap}{\lambda^2}$$

۴- یک مغازه‌ی پیتزا فروشی n نوع پیتزای مختلف عرضه می‌کند. فرض کنید X مشتری در روز به این مغازه مراجعه می‌کنند که X یک متغیر تصادفی گسسته و نامنفی با تابع مولد ممان $M_X(s) = \mathbb{E}\{e^{sX}\}$ است. هر مشتری یکی از n نوع پیتزا را به‌صورت کاملاً تصادفی سفارش می‌دهد (هر کدام از n نوع پیتزا با احتمال برابر انتخاب می‌شوند). همچنین نوع پیتزای انتخاب‌شده توسط هر مشتری مستقل از تعداد مشتریان دیگر و نوع پیتزایی است که آن‌ها سفارش می‌دهند. تعیین کنید که به صورت میانگین چند نوع پیتزا در روز توسط این مغازه عرضه می‌شود؟ رابطه‌ی نهایی صرفاً برحسب n و تابع مولد ممان X است. **[۸ نمره]**

از متغیر تصادفی شاخص Y_i استفاده می‌کنیم که در صورت سفارش پیتزای نوع i برابر با 1 و در غیر این صورت برابر با 0 است.

در صورتی که x مشتری به مغازه مراجعه کنند:

$$P(Y_i = 0) = \left(1 - \frac{1}{n}\right)^x$$

$$P(Y_i = 1) = 1 - \left(1 - \frac{1}{n}\right)^x$$

اگر Y تعداد کل انواع پیتزای عرضه شده در یک روز باشد:

$$Y = Y_1 + Y_2 + \cdots + Y_n$$

$$E[Y|X = x] = n\left(1 - \left(1 - \frac{1}{n}\right)^x\right)$$

$$E[Y] = \sum_{x=0}^{\infty} E[Y|X = x]P(X = x) =$$

$$\sum_{x=0}^{\infty} n\left(1 - \left(1 - \frac{1}{n}\right)^x\right)P(X = x) =$$

$$n\sum_{x=0}^{\infty} P(X = x) - n\sum_{x=0}^{\infty}\left(1 - \frac{1}{n}\right)^x P(X = x)$$

از تعریف تابع مولد ممان داریم:

$$M_X(s) = E[e^{sX}] = \sum_{x=0}^{\infty} e^{sx}P(X = x)$$

$$\rightarrow \sum_{x=0}^{\infty}\left(1 - \frac{1}{n}\right)^x P(X = x) = M_X\left(\ln\left(1 - \frac{1}{n}\right)\right)$$

و از آنجایی که X یک توزیع گسسته است: $\sum_{x=0}^{\infty} P(X = x) = 1$

$$E[Y] = n\left(1 - M_X\left(\ln\left(1 - \frac{1}{n}\right)\right)\right)$$

۵- فرض کنید X، Y و Z متغیرهای تصادفی پیوسته و توأماً مستقل با توابع توزیع تجمعی $F_X(x) = h(x)$، $F_Y(y) = h^2(y)$ و $F_Z(z) = h^3(z)$ باشند. $\mathbb{Pr}\{X < Y < Z\}$ را به‌دست آورید. **[۸ نمره]**

$$P(X < Y < Z \mid Y = y) = P(X < y \ , y < Z) = P(X < y)P(Z > y) = F_X(y)\big(1 - F_Z(y)\big)$$
$$= h(y)(1 - h^3(y))$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = 2h(y)h'(y)$$

$$P(X < Y < Z) = \int_{-\infty}^{\infty} P(X < Y < Z|Y = y)f_Y(y)dy = \int_{-\infty}^{\infty} 2h^2(y)\big(1 - h^3(y)\big)h'(y)dy$$

با تغییر متغیر $u = h(y)$ داریم:

$$du = h'(y)dy$$

$$0 \le u \le 1$$

$$P(X < Y < Z) = \int_0^1 2u^2(1 - u^3)du = \frac{2}{3}u^3 - \frac{2}{6}u^6 \mid_0^1 = \frac{1}{3}$$

۶- فرض کنید میانگین ضریب هوشی مردم ایران برابر با ۱۰۰ باشد. نمونه‌ای ۶۴ نفره از دانشجویان دانشگاه تهران انتخاب و ضریب هوشی آن‌ها اندازه‌گیری شده است. میانگین ضریب هوشی این نمونه برابر با ۱۰۶ و انحراف معیار آن برابر با ۳ به‌دست آمده است.

الف) با طراحی یک آزمون فرض مناسب با $\alpha = 0.03$، تصمیم بگیرید آیا ضریب هوشی دانشجویان دانشگاه تهران از میانگین ضریب هوشی مردم ایران بیشتر است یا خیر؟ **[۴ نمره]**

$$\mathrm{H_0}:\ \mu = 100$$
$$\mathrm{H_A}:\ \mu > 100$$

$$\text{p-value} = \mathrm{P}(\overline{\mathrm{X}} > 106|\mu = 100) = \mathrm{P}\left(\mathrm{Z} > \frac{106-100}{\frac{3}{\sqrt{64}}}\right) = \mathrm{P}(\mathrm{Z} > 16) = 0 < 0.03$$

بنابراین فرض $\mathrm{H_0}$ را رد می‌کنیم.

ب) یک بازه‌ی اطمینان ۹۸ درصد برای میانگین ضریب هوشی دانشجویان دانشگاه تهران پیدا کنید. **[۴ نمره]**

$$\alpha = 1 - 0.98 = 0.02 \to 1 - \frac{\alpha}{2} = 0.99 \to z_{1-\frac{\alpha}{2}} = 2.33$$

بازه اطمینان برابر است با:

$$\left(106 - 2.33 \times \frac{3}{\sqrt{64}}\ ,\ 106 - 2.33 \times \frac{3}{\sqrt{64}}\right)$$

$$= \big(106 - 0.87\ ,\ 106 + 0.87\big) = (105.13,\ 106.87)$$