第七次作业参考解答

1. 对一般的$X \sim \mathrm{B}(n,p)$,$\mathrm{E}(X) = np, \mathrm{Var}(X) = np(1-p)$且：

$$\mathrm{Skewness}(X) = \frac{1-2p}{\sqrt{np(1-p)}}.$$

若$np \in \mathbb{Z}$,则进一步有$\mathrm{Mode}(X) = \mathrm{Median}(X) = np.$

2. (2)指数分布、泊松分布、均匀分布的矩母函数分别为

$$M(s) = \frac{\lambda}{\lambda - s}$$
$$M(s) = e^{\lambda(e^s-1)}$$
$$M(s) = \frac{e^{sb} - e^{sa}}{s(b-a)}$$

(3)若X的矩母函数为$M_X(t)$,则

$$\mathrm{Skew}(X) = \frac{M_X^{(3)}(0) - 3M_X^{(1)}(0)M_X^{(2)}(0) + 2[M_X^{(1)}(0)]^3}{[M_X^{(2)}(0) - [M_X^{(1)}(0)]^2]^{\frac{3}{2}}},$$

$$\mathrm{Kurt}(X) = \frac{M_X^{(4)}(0) - 4M_X^{(3)}(0)M_X^{(1)}(0) + 6M_X^{(2)}(0)[M_X^{(1)}(0)]^2 - 3[M_X^{(1)}(0)]^4}{[M_X^{(2)}(0) - [M_X^{(1)}(0)]^2]^2}.$$

3. 指数随机变量的相加 $\frac{2}{3}e^{-2x} + 2e^{-3x}$

4. (1)对任意$t > 0$,可验证

$$\lim_{y \to +\infty} e^{ty} f_Y(y) dy = +\infty$$

故积分$\mathrm{E}(e^{tY})$发散；类似地可验证对$t < 0$,$\mathrm{E}(e^{tY})$不存在.

(2)记$X \sim \mathcal{N}(\mu, \sigma^2)$的矩母函数为$M_X(t)$,则$Y$的$n$阶原点矩形即$\mathrm{E}(Y^n) = \mathrm{E}(e^{nx}) = M_X(n)$.

5. 独立泊松变量相加仍然服从泊松分布

$$M_Y(s) = M_{X_1}(s) * M_{X_2}(s) = e^{(\lambda_1+\lambda_2)(e^s-1)}$$

6. 设第一次取到的断点的左侧占全部线段的比例为X,第二次取到的断点的左侧占剩余线段的比例为Y.则X, Y为独立的随机变量且都服从$[0,1]$上的均匀分布。故

$$f(x) = 1, 0 \le x \le 1, \quad f(y) = 1, 0 \le y \le 1$$

这样，二者的期望为$E(X) = \int_0^1 x\mathrm{d}x = \frac{1}{2}, E(Y) = \frac{1}{2}$。所求的剩余长度为$(1-X)(1-Y)$。

$$E((1-X)(1-Y)) = E(1 - X - Y + XY) = 1 - \frac{1}{2} - \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

7. 设X为矿工走到安全之处的时间，Y为矿工第一次选择的门的编号。那么有

$$E(X|Y=1)=2$$
$$E(X|Y=2)=3+E(X)$$
$$E(X|Y=3)=1+E(X)$$

根据重期望公式，有

$$E(X)=\sum_{y=1}^{3}E(X|Y=y)P(Y=y)=\frac{1}{3}(6+2E(x))$$

解得$E(X)=6$.

或者用类似几何分布的想法，选中第一个门可以视为成功。

设X为矿工走到安全之处的时间，在第Z次选择中选中了第一个门。则

$$P(Z=n)=\frac{1}{3}(\frac{2}{3})^{n-1}$$
$$E(X|Z=n)=\frac{3+1}{2}(n-1)+2=2n$$
$$E(X)=E(E(X|Z=n))=\sum_{n=1}^{\infty}(\frac{2}{3})^{n}n$$
$$=\lim_{n\to\infty}6-(2n+6)(\frac{2}{3})^{n}=6$$

8. 反复应用重期望公式和性质$\mathrm{E}[g(Y)X|Y]=g(Y)\mathrm{E}[X|Y]$得

$$\mathrm{E}[XY]=\mathrm{E}[\mathrm{E}[XY|X]]=\mathrm{E}[X\mathrm{E}[Y|X]]=\mathrm{E}[X^2],\mathrm{E}[Y]=\mathrm{E}[\mathrm{E}[Y|X]]=\mathrm{E}[X],$$

从而$\mathrm{Cov}(X,Y)=\mathrm{E}[XY]-\mathrm{E}[X]\mathrm{E}[Y]=\mathrm{Var}[X]$.

9. (1)仅要求对离散和连续的情形证明即可.对于Y为离散随机变量的情形,由X,Y独立知

$$\mathrm{E}[Y|X=x]=\sum_i y_i\mathrm{P}(Y=y_i|X=x)=\sum_i y_i\mathrm{P}(Y=y_i)=\mathrm{E}[Y].$$

从而$\mathrm{E}[Y|X]=\mathrm{E}[Y]$.

(2)反之不一定成立.令$X\sim\mathrm{U}(0,1),Y\sim\mathrm{U}(-X,X)$,通过计算验证$E[Y|X]=E[Y]=0$以及$X,Y$不独立.

画一个圆，要求(X,Y)落在$X^2+Y^2=1$的圆上，此时有

$$\mathrm{E}[Y|X]=0\quad\mathrm{E}[X]=0$$

但两者不独立，X的取值会影响Y的取值

10. 参考概率导论(Bertsekas,第2版)4.3.1节的证明(Page197),先证明$\mathrm{Cov}(\hat{Y},\tilde{Y})=0$,再在$Y=\hat{Y}+\tilde{Y}$两边取方差

即可.

$$\begin{aligned}\mathrm{Var}(Y) &= E[Y^2] - (E[Y])^2\\ &= E[\hat{Y}^2 + 2\hat{Y}\tilde{Y} + \tilde{Y}^2] - (E[\hat{Y}] + E[\tilde{Y}])^2\\ &= E[\hat{Y}^2] - E[\hat{Y}]^2 + E[\tilde{Y}^2] - E[\tilde{Y}]^2 + 2E[\hat{Y}\tilde{Y}] - 2E[\hat{Y}]E[\tilde{Y}]\\ &= Var[\hat{Y}] + Var[\tilde{Y}]\end{aligned}$$

11. (1)

$$\begin{aligned}\mathrm{Var}[Y|X] = \mathrm{E}[(Y-\mathrm{E}[Y|X])^2|X] &= E[Y^2 - 2Y\mathrm{E}[Y|X] + \mathrm{E}^2[Y|X]|X]\\ &= \mathrm{E}[Y^2|X] - 2\mathrm{E}[YE[Y|X]|X] + E^2[Y|X]\\ &= \mathrm{E}[Y^2|X] - 2\mathrm{E}^2[Y|X] + \mathrm{E}^2[Y|X] = \mathrm{E}[Y^2|X] - \mathrm{E}^2[Y|X]\end{aligned}$$

(2)

$$\begin{aligned}\mathrm{Var}[Y] &= \mathrm{E}[Y^2] - \mathrm{E}^2[Y]\\ &= \mathrm{E}[Y^2] - \mathrm{E}[\mathrm{E}^2[Y|X]] + \mathrm{E}[\mathrm{E}^2[Y|X]] - \mathrm{E}^2[Y]\\ &= \mathrm{E}[\mathrm{E}[Y^2|X] - \mathrm{E}^2[Y|X]] + \mathrm{E}[\mathrm{E}^2[Y|X]] - \mathrm{E}[\mathrm{E}^2[Y|X]]\\ &= \mathrm{E}[\mathrm{Var}[Y|X]] + \mathrm{Var}[\mathrm{E}[Y|X]]\end{aligned}$$

12. 对任意分布(X,Y),条件期望$\mathrm{E}[Y|X]$是在得到观测X下Y的最优预测(最小均方误差意义下),因为成立性质:

$$\mathrm{E}[Y|X] = \underset{g(X)}{\operatorname{argmin}}\, E[(Y-g(X))^2|X]$$

特别地,若$f_{(X,Y)}(x,y) = \dfrac{2}{\pi}\mathbf{1}_{\{x\geq 0, y\geq 0, x^2+y^2\leq 1\}}(x,y)$,计算得

$$\begin{aligned}f_X(x) &= \int_0^{\sqrt{1-x^2}} f(x,y)dy = \frac{2}{\pi}\sqrt{1-x^2}\mathbf{1}_{[0,1]}(x)\\ f_{Y|X}(y|x) &= \frac{1}{\sqrt{1-x^2}}\mathbf{1}_{[0,\frac{1}{\sqrt{1-x^2}}]}(y)\end{aligned}$$

$$\mathrm{E}[Y|X=x] = \int_{\mathbb{R}} y f_{Y|X}(y|x)\mathrm{d}y = \frac{\sqrt{1-x^2}}{2}, \mathrm{E}[Y|X=0.5] = \frac{\sqrt{3}}{4}.$$

13. (1)

$$\begin{aligned}E[(Y-(aX+b))^2] &= E[Y^2] - 2aE[XY] + a^2E[X^2] + 2abE[X] - 2bE[Y] + b^2\\ &= (\mu_2^2+\sigma_2^2) - 2a(\mu_1\mu_2 + \rho\sigma_1\sigma_2) + a^2(\mu_1^2+\sigma_1^2) + 2ab\mu_1 - 2b\mu_2 + b^2\\ &= (\mu_1^2+\sigma_1^2)a^2 + 2\mu_1 ab + b^2 - 2(\mu_1\mu_2+\rho\sigma_1\sigma_2)a - 2\mu_2 b + \mu_2^2 + \sigma_2^2\\ &\triangleq g(a,b)\end{aligned}$$

求偏导得到

$$\begin{aligned}\frac{\partial g(a,b)}{\partial a} &= 2[(\mu_1^2+\sigma_1^2)a + \mu_1 b - (\mu_1\mu_2 + \rho\sigma_1\sigma_2)] = 0\\ \frac{\partial g(a,b)}{\partial b} &= 2(\mu_1 a + b - \mu_2) = 0\end{aligned}$$

因而 $a = \rho\frac{\sigma_1}{\sigma_2}$，$b = \mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1}$ 时，是Y的最优线性预测。

(2)将第一小问中的a,b代入，得到均方误差为

$$E[(Y-(aX+b))^2]=\sigma_2^2(1-\rho^2)$$

可以知道当σ_2^2趋近于0或者ρ趋近于1的时候，该值接近于0。对应的实际解释为：本身随机变量的随机性比较小，或者观测变量与随机变量之间具有强相关性时，观测的均方误差会趋近于0。

(3)这道题实际上是让我们求$E[Y|X]$

$$E[Y|X]=\int_{-\infty}^{+\infty} f_{Y|X}(y|x)dy=\mu_2+\frac{\rho\sigma_1}{\sigma_2}(x-\mu_1)$$

而对于最优线性预测得到的$\hat{Y}=aX+b=\mu_2+\frac{\rho\sigma_1}{\sigma_2}(x-\mu_1)$，两者相同。

14. (1)

$$E(Y|N)=N\mu$$
$$E(Y)=E(E(Y|N)) \quad =E(N)\mu=a\mu$$

(2)

$$M_N(t)=E(e^{Nt})=\sum_{n=1}^{\infty}p_ne^{nt}$$

(3)

$$E(e^{Yt}|N)=\prod_{n=1}^{N}E(e^{X_nt})=M_X(t)^N$$
$$M_Y(t)=E(e^{Yt})=E(M_X(t)^N)=\sum_{n=1}^{\infty}p_nM_X(t)^n$$

(4) 此时X的矩母函数$M_X(t)=\frac{\lambda}{\lambda-t}, t<\lambda$.则$Y$的矩母函数为

$$\begin{aligned}M_Y(t)&=\sum_{n=1}^{\infty}(1-p)^{n-1}p(\frac{\lambda}{\lambda-t})^n=\frac{p}{1-p}\sum_{n=1}^{\infty}(\frac{(1-p)\lambda}{\lambda-t})^n\\&=\frac{p}{1-p}\frac{\frac{(1-p)\lambda}{\lambda-t}}{1-\frac{(1-p)\lambda}{\lambda-t}}\\&=\frac{\lambda p}{\lambda p-t}\end{aligned}$$

(5) 考察X_1+X_2的矩母函数，有

$$\begin{aligned}M_{X_1+X_2}(t)&=E(e^{(X_1+X_2)t})\\&=E(e^{X_1t})E(e^{X_2t})\\&=(\frac{\lambda}{\lambda-t})^2\end{aligned}$$

显然二者的分布不相同。

15. 例子1:设$X_i(i=1,2,\cdots)$相互独立且服从标准正态分布， $Y=X_1+X_2+\cdots+X_N$,个数$N\sim P(\lambda)$. 则

$$M_N(t)=e^{\lambda(e^t-1)}$$
$$M_X(t)=e^{\frac{t^2}{2}}$$
$$M_Y(t)=e^{\lambda(e^{\frac{t^2}{2}}-1)}$$

假设$Y\sim N(\mu,\sigma^2)$，则

$$M_Y(t)=e^{\frac{\sigma^2t^2}{2}+\mu t}$$
$$\lambda(e^{\frac{t^2}{2}}-1)=\frac{\sigma^2t^2}{2}+\mu t$$

而泊松分布的参数$\lambda>0$，关于t的指数函数不能写成有限多的多项式相加的形式。上述等式不成立。Y不是正态分布。

例子2: 取$X_i\overset{\text{i.i.d}}{\sim}\mathcal{N}(0,1),Y=\sum\limits_{i=1}^{N}X_i$,其中$X_i,N$彼此独立且

$$\mathrm{P}(N=1)=\mathrm{P}(N=2)=\frac{1}{2}.$$

由**Q14**结论计算得Y的矩母函数$M_Y(s)=\dfrac{1}{2}\exp(\dfrac{1}{2}s^2)+\dfrac{1}{2}\exp(s^2)$,可以验证其不是正态分布对应的矩母函数.

16. (1) $\mathrm{E}[X_n]=n\mathrm{E}[Y]=0,\mathrm{Var}[X_n]=n\mathrm{Var}[Y]=n.$

(2)参考one-dimensional Wiener process的性质.

作业总结：

- 典型分布的矩母函数。

  1. **泊松分布**（参数 λ）的矩母函数：

  $$M_X(t) = \exp\left(\lambda\left(e^t - 1\right)\right)$$

  2. **均匀分布**（在区间 $[a, b]$ 上的均匀分布）的矩母函数：

  $$M_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)} \quad (t \neq 0)$$

  3. **指数分布**（参数 λ）的矩母函数：

  $$M_X(t) = \frac{\lambda}{\lambda - t} \quad \text{for} \quad t < \lambda$$

  4. **高斯分布**（均值为 μ，方差为 σ^2）的矩母函数：

  $$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

- 矩母函数和均值方差的关系 1. **均值**（$\mu = \mathbb{E}[X]$）由矩母函数的导数计算：

  $$\mu = M_X'(0)$$

  2. **方差**（$\sigma^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$）由矩母函数的二阶导数计算：

  $$\sigma^2 = M_X''(0) - \left(M_X'(0)\right)^2$$

- 重期望公式的灵活应用：描述了条件期望与总体期望之间的关系：

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