一. 复习

(1) 期望

$E[X] = \sum_x x f(x)$，当 X 取值自然数时：$E[X] = \sum_{n=0}^{\infty} P(X>n)$

证：$E[X] = \sum_{n=1}^{\infty} nP(X=n) = \sum_{n=1}^{\infty} n(P(X>n-1) - P(X>n)) = \sum_{n=0}^{\infty} P(X>n)$

(cauchy-schwarz) $(E[XY])^2 \leq E[X^2]E[Y^2]$

(2) 方差. 协方差

$Var(X) = E(X-EX)^2 = \min_a E(X-a)^2$ $\quad E[X] = \arg\min_a E(X-a)^2$

$Cov(X,Y) = E[(X-EX)\cdot(Y-EY)] = E[XY] - E[X]\cdot E[Y]$

(cauchy-schwarz) $|Cov(X,Y)| \leq \sqrt{Var(X)Var(Y)}$

$var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} var(X_i) + 2\sum_{i<j} cov(X_i, X_j)$

协方差的双线性性：$cov(aX+bY, Z) = a\,cov(X,Z) + b\,cov(Y,Z)$

$cov(X, aY+bZ) = a\,cov(X,Y) + b\,cov(X,Z)$

(3) 条件期望/分布

(X,Y) $\quad P(Y=y)>0$

$f_{X|Y}(x|y) = P(X=x|Y=y) = \dfrac{P(X=x, Y=y)}{P(Y=y)}$

$F_{X|Y}(x|y) = P(X \leq x|Y=y)$

$E[X|Y=y] = \sum_x x f_{X|Y}(x|y) \in R$ $\quad E[X|Y]$ r.v. $\quad E[E[Y|X]] = E[Y]$

如何理解条件期望？投影

$E[Y|X] = \arg\min_{\varphi(X)} E[(Y-\varphi(X))^2]$ —— X是已知的信息，此时对Y作出的最好估计

X, Y独立时，$E[Y|X] = E[Y] = \arg\min_a E[(Y-a)^2]$

• (X,Y)为联合离散型随机向量. X, Y二阶矩存在. 记 $\varphi(X) = E[Y|X]$. 若 g 为可测函数且 $g(X)$ 二阶矩存在. 证：$E(Y-\varphi(X))^2 \leq E(Y-g(X))^2$.

证：利用 $E(g(X)h(Y)|X) = g(X)E(h(Y)|X)$

$E(Y-g(X))^2 = E[((Y-\varphi(X))+(\varphi(X)-g(X)))^2] = E(Y-\varphi(X))^2 + \overbrace{E(\varphi(X)-g(X))^2}^{\geq 0}$

$+2E[(Y-\varphi(X))\cdot(\varphi(X)-g(X))]$

$E[(Y-\varphi(X))\cdot(\varphi(X)-g(X))]=E[E[(Y-\varphi(X))\cdot(\varphi(X)-g(X))|X]]$

$\overset{(*)}{=}E[(\varphi(X)-g(X))\cdot(\underbrace{E[Y|X]-\varphi(X)}_{=0})]=0$

区分: $F_X(\overset{r.v}{X})$. $F_X(x)$. $F_Y(\overset{r.v}{X})$

$F_{X,Y}(x,y)\Rightarrow F_X(x).F_Y(y).F_{Y|X}(y|x).F_{X|Y}(x|y)$

但 $F_X(x).F_Y(y)\not\Rightarrow F_{X,Y}(x,y)$

$F_X(x).F_{Y|X}(y|x)\Rightarrow F_{X,Y}(x,y)$

(4) 母函数 $G_X(s)=\sum_{k=0}^{\infty}P(X=k)s^k$ $=E[s^X]$, X取非负整数值.

$X_1,\cdots,X_n$ 相互独立. $Y=\sum_{i=1}^{n}X_i$. X_i母函数$G_i(s)$, 则 $G_Y(s)=\prod_{i=1}^{n}G_i(s)$

(5) 常见分布 $f(x)$. $E[X]$. $var(X)$. 母函数.

(i) $X\sim B(n,p)$ $P(X=k)=C_n^k p^k q^{n-k}$ $k=0,1,\cdots,n$

$X_1\sim B(n_1,p)$ $X_2\sim B(n_2,p)$ $X_1\perp\!\!\!\perp X_2$ $\Rightarrow X_1+X_2\sim B(n_1+n_2,p)$

(ii) $X\sim G(p)$ $P(X=k)=q^{k-1}\cdot p$, $k=1,2,\cdots$

无记忆性: $P(X=k+n|X>k)=P(X=n)$

(iii) $X\sim P(\lambda)$ $P(X=k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!}$ $k=0,1,2\cdots$

$X_1\sim Poi(\lambda_1)$ $X_2\sim Poi(\lambda_2)$ $X_1\perp\!\!\!\perp X_2\Rightarrow X_1+X_2\sim Poi(\lambda_1+\lambda_2)$

(iv) $X\sim f(r,p)$ $P(X=k)=C_{k-1}^{r-1}p^r q^{k-r}$ $k=r,r+1,\cdots$

$X=X_1+\cdots+X_r$ $X_1,\cdots,X_r$独立 $X_i\sim G(p)$ X_i表示第$i-1$次成功到第i次成功所需次数.

(6) 随机游走 (☆变换)

二. 习题

示性函数:

$I_A(w)=\begin{cases}1, & w\in A\\ 0, & w\notin A\end{cases}$ $P(I_A=1)=P(A)=E(I_A)$

配对问题(点数相同. 拿到自己伞. 置换)是否(是:1, 否:0). 计数

1. 电梯n层, $m(<n)$人均匀随机停, 记X为电梯停的次数. 求$E[X]$.

解: 令$I_i=\{1$, 第i层有人停 $i=1,2,\cdots,n$

$\quad\quad\quad \lfloor 0$, 否则

$X=\sum_{i=1}^{n} I_i \qquad E[X]=E[\sum_{i=1}^{n} I_i]=\sum_{i=1}^{n} E[I_i]=\sum_{i=1}^{n} E[1-I_i^c]$

$\qquad\qquad\qquad\qquad\qquad =\sum_{i=1}^{n}[1-(1-\frac{1}{n})^m]=n(1-(1-\frac{1}{n})^m)$

2.

(35分) S_n表示从$[n]=\{1,2,\ldots,n\}$到$[n]$双射全体, 从S_n中(均匀地)随机选取一个σ, 定义不动点数为$X(\sigma)=\#\{k:\sigma(k)=k\}$, 对换数为$Y(\sigma)=\#\{(i,j):\sigma(i)=j,\sigma(j)=i,i<j\}$. 回答

(i)详细给出一个有关的概率空间.

(ii)X与Y是否独立? 说明理由.

(iii)计算X的分布列.

(iv)求Y的期望.

解: (i) $\Omega=S_n=\left\{\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}: i_k\in[n]\text{且各不相同}\right\}$

$\qquad F=\{S_n\text{中元素所有可能并}\}\cup\{\phi\}$

$\qquad \forall A\in F:\ P(A)=\frac{|A|}{n!}$

(ii) $P(X(\sigma)=n)=\frac{1}{n!} \qquad P(Y(\sigma)=1)>0$ 但 $P(X(\sigma)=n, Y(\sigma)=1)=0 \Rightarrow$ 不独立.

(iii) 同上课所讲"拿伞问题"

(iv) 令 $I_{ij}=\begin{cases} 1 & \sigma(i)=j,\sigma(j)=i,\ i<j \\ 0 & \text{否则}\end{cases} \qquad Y=\sum_{i<j} I_{ij} \qquad E(I_{ij})=1\cdot\frac{(n-2)!}{n!}=\frac{1}{n(n-1)}$

$\qquad E(Y)=E(\sum_{i<j} I_{ij})=\sum_{i<j}E(I_{ij})=\frac{n(n-1)}{2}\cdot\frac{1}{n(n-1)}=\frac{1}{2}$

3.

(20分) 给定$b>a>0$, 离散随机变量X取值于区间$[a,b]$, 试回答

(i)证明$\mathrm{Var}(X)\le\frac{1}{4}(b-a)^2$;

(ii)当X变化时, 找出并验证乘积$\mathbf{E}[X]\mathbf{E}[1/X]$的取值范围.

证: (i) $\mathrm{Var}(X)=\min_{a} E(X-a)^2\le E(X-\frac{a+b}{2})^2\le\frac{(b-a)^2}{4}$

(ii) $E(X)E(\frac{1}{X})\ge(E(\sqrt{X}\cdot\sqrt{\frac{1}{X}}))^2=1$ 取等: 例如 $X=\frac{a+b}{2}$

$\qquad \mathrm{cov}(X,\frac{1}{X})=E(X\cdot\frac{1}{X})-E(X)E(\frac{1}{X})=1-E(X)E(\frac{1}{X})$

$$|cov(X,\tfrac{1}{X})| \le \sqrt{var(X)var(\tfrac{1}{X})} \le \sqrt{\frac{(b-a)^2}{4}\cdot\frac{(\frac{1}{b}-\frac{1}{a})^2}{4}} = \frac{(b-a)^2}{4ab}$$

$$1-E(X)E(\tfrac{1}{X}) \ge -\frac{(b-a)^2}{4ab} \Rightarrow E(X)E(\tfrac{1}{X}) \le \frac{(a+b)^2}{4ab}$$ 取等 $P(X=a)=\frac{1}{2}$, $P(X=b)=\frac{1}{2}$

4.

(15分) ζ小盆友有N块积木，N服从参数为λ的泊松分布，δ小盆友独立地以1/2概率拿走每一块. 若δ小盆友的积木块数为K，求$\mathrm{E}[K]$ 和$\mathrm{E}[N|K]$.

解: $E[K]=E[E[K|N]]=E[\frac{1}{2}N]=\frac{1}{2}E[N]=\frac{1}{2}\lambda$

$$E[N|K=k]=\sum_{n=0}^{k} n\cdot f_{N|K}(n|k)$$

$$f_{N|K}(n|k)=\frac{P(K=k,N=n)}{P(K=k)}=\frac{P(K=k|N=n)P(N=n)}{\sum_{n=0}^{+\infty}P(K=k|N=n)P(N=n)}=\frac{C_n^k\cdot(\frac{1}{2})^n\cdot\frac{\lambda^n}{n!}e^{-\lambda}}{\sum_{n=0}^{+\infty}C_n^k\cdot(\frac{1}{2})^n\cdot\frac{\lambda^n}{n!}e^{-\lambda}}$$

$$=\frac{(\frac{\lambda}{2})^{n-k}}{(n-k)!}e^{-\frac{\lambda}{2}}$$

$$E[N|K=k]=\sum_{n=0}^{k}\underbrace{n}_{n-k+k}\cdot\frac{(\frac{\lambda}{2})^{n-k}}{(n-k)!}e^{-\frac{\lambda}{2}}=k+\frac{\lambda}{2}$$

$$E[N|K]=K+\frac{\lambda}{2}$$

5.

设$S_N=X_1+\cdots+X_N$为N个相互独立随机变量之和，其中每个随机变量等概率地取值$1,2,\ldots,m$. 求

(1)S_N概率母函数; (2) 关于k的序列$P(S_N\le k)$的母函数; (3) 又设N为参数为$p\in(0,1)$的几何分布, 且N与$\{X_j: j=1,2,\ldots,\}$ 独立, 试回答(2)中问题.

解：(1) $G_{S_N}(s)=E[s^{S_N}]=(E[s^{S_1}])^N=(\frac{1}{m}(s+s^2+\cdots+s^m))^N=(\frac{1-s^{m+1}}{m(1-s)})^N$

(2) $G(s)=\sum_{k=0}^{\infty}P(S_N\le k)\cdot s^k=\sum_{k=0}^{\infty}(\sum_{j=0}^{k}P(S_N=j))\cdot s^k=\sum_{j=0}^{\infty}(\sum_{k=j}^{\infty}s^k)P(S_N=j)=\sum_{j=0}^{\infty}\frac{s^j}{1-s}P(S_N=j)$

$$=\frac{G(S_N)}{1-s}=\frac{(1-s^{m+1})^N}{(1-s)^{N+1}\cdot m^N}$$

(3) $G_N(s)=\sum_{n=1}^{+\infty}P(N=n)s^n=\sum_{n=1}^{+\infty}p(1-p)^{n-1}\cdot s^n=\frac{p}{1-p}\sum_{n=1}^{+\infty}(s(1-p))^n=\frac{p}{1-p}\cdot\frac{s(1-p)}{1-s(1-p)}=\frac{ps}{1-s(1-p)}$

$$G_{S_N}(s)=G_N(G_X(s))=\frac{p\cdot\frac{1-s^{m+1}}{m(1-s)}}{1-\frac{s(1-s^m)}{m(1-s)}(1-p)}=\cdots$$

$P(S_N\le k)$的母函数: $\frac{G_{S_N}(s)}{1-s}=\cdots$

6.

(20分) 直线上简单随机游动$S_n=\sum_{k=1}^n X_k$，$S_0=0$, 这里$P(X_1=1)=p$, $P(X_1=-1)=1-p$, $0<p<1$. 记$S_0,S_1,\dots,S_n$中互不相同的值个数为R_n. 试证明

(i) $\mathbf{P}(R_n=R_{n-1}+1)=\mathbf{P}(S_1\cdots S_n\neq 0)$;

(ii) 当$n\to\infty$时, $\frac{1}{n}\mathbf{E}[R_n]\to \mathbf{P}(S_k\neq 0,\ \forall k\geq 1)$;

(iii) $\mathbf{P}(S_k\neq 0,\ \forall k\geq 1)=|2p-1|$.

证: (i) $R_n=R_{n-1}+1$ 即S_n取值与$S_0,\cdots,S_{n-1}$不同.

对路径$(X_1,\cdots,X_n)$作变换 $\varphi:(X_1,\cdots,X_n)\longrightarrow(X_n,\cdots,X_1)$ 一一对应

$A=\{(X_1,\cdots,X_n): R_n=R_{n-1}+1\}$ $\quad B=\{(X_1,\cdots,X_n): S_1,\cdots,S_n\neq 0\}$ $\quad \varphi(A)=B$

$P(A)=P(\varphi(A))=\varphi(B)$ $\quad P(R_n=R_{n-1}+1)=P(S_1,\cdots,S_n\neq 0)$

(ii) $E[R_n]=1+E[\sum_{k=1}^n 1_{\{R_k=R_{k-1}+1\}}]=1+\sum_{k=1}^n P(R_k=R_{k-1}+1)=1+\sum_{k=1}^n P(S_1,\cdots,S_k\neq 0)$

$\frac{1}{n}E[R_n]=\frac{1}{n}(1+\sum_{k=1}^n P(S_1,\cdots,S_k\neq 0))\longrightarrow P(S_k\neq 0,\forall k\geq 1)$ 当$n\to+\infty$时

(iii) ① $p>\frac{1}{2}$ 即$p>q$. 先考虑如下问题: 若$S_1=1$. $V_i=\min\{n:S_n=i\}$ 求$P(V_0>V_N)$

令$p_i=P(S_1=i, V_0>V_N)$. 则: $p_0=0$, $p_N=1$ $\Rightarrow p_i=p\cdot p_{i+1}+(1-p)\cdot p_{i-1}$

$p(p_{i+1}-p_i)=q(p_i-p_{i-1})\Rightarrow p_{i+1}-p_i=\frac{q}{p}(p_i-p_{i-1})=(\frac{q}{p})^{n-1}(p_1-p_0)$

$\Rightarrow p_N-p_0=\sum_{k=1}^N(p_k-p_{k-1})=\sum_{k=1}^N(\frac{q}{p})^{k-1}\cdot p_1=\frac{1-(\frac{q}{p})^N}{1-\frac{q}{p}}\cdot p_1 \quad\Rightarrow p_1=\frac{1-\frac{q}{p}}{1-(\frac{q}{p})^N}$

$P(S_1=1,S_k>0,\forall k)=P(S_1=1,V_0>V_k,\forall N)=1-\frac{1-p}{p}=\frac{2p-1}{p}$

$P(S_1=-1,S_k<0,\forall k)=0$ $\quad$ §5.3(6) corollary $\quad$ P(访问过正轴)$=\min\{1,\frac{p}{q}\}$

$\Rightarrow P(S_k\neq 0,\forall k)=p\cdot\frac{2p-1}{p}=2p-1$

② $p\leq q$时同理.

7. $\{\overset{r.v.}{X_k}, k\geq 1\}$ 独立同分布，且与 $\overset{r.v.}{N}$ 独立，$Y=\sum_{k=1}^{N} N_k$，证：$E[Y]=E[N]\cdot E[X]$，

$Var(Y)=E[N]\cdot Var(X)+Var(N)\cdot(E(X))^2$

证：记 X_k 的母函数为 $F(s)$，N 的母函数为 $G(s)$，则 Y 的母函数 $Y(s)=G(F(s))$

$Y'(s)=G'(F(s))\cdot F'(s)$，　$Y''(s)=G''(F(s))\cdot(F'(s))^2+G'(F(s))\cdot F''(s)$

$E[Y]=Y'(1)=G'(F(1))F'(1)=G'(1)F'(1)=E[N]\cdot E[X]$

$Var(Y)=Y''(1)+Y'(1)-(Y'(1))^2$

$=G''(1)(F'(1))^2+G'(1)F''(1)+G'(1)F'(1)-(G'(1)F'(1))^2$

$=G'(1)(F''(1)+F'(1)-(F'(1))^2)+(F'(1))^2(G''(1)+G'(1)-(G'(1))^2)$

$=E[N]\cdot Var(X)+(E[X])^2+Var(N)$

8. 直线上简单随机游走 $S_n=\sum_{k=1}^{n}X_k$，$S_0=0$。X_i 独立同分布，$P(X_i=1)=p$，$P(X_i=-1)=1-p$，$0<p<1$。

(1) 求 $E[S_n]$，$Var(S_n)$，$Cov(S_m,S_n)$

(2) Y 服从 $G(p)$ 且与 $\{X_i\}$ 独立，求 $Var(S_Y)$

(3) 对正整数 k，求 S_{n+k} 关于 S_n 的条件分布列 $f_{S_{n+k}|S_n}$ 与条件期望 $E[S_{n+k}|S_n]$。

解：(1) $E[S_n]=E[\sum_{k=1}^{n}X_k]=\sum_{k=1}^{n}E[X_k]=n(p\cdot 1+(1-p)\cdot(-1))=(2p-1)\cdot n$

$Var(S_n)=\sum_{i=1}^{n}Var(X_i)=\sum_{i=1}^{n}(E[X_i^2]-(E[X_i])^2)=n(1-(2p-1)^2)=4np(1-p)$

$Cov(S_m,S_n)=\sum_{i=1}^{m}\sum_{j=1}^{n}Cov(X_i,X_j)=\sum_{i=1}^{\min(m,n)}Var(X_i)=4\min\{m,n\}p(1-p)$

(2) 由 T7，$Var(S_Y)=E[Y]\cdot Var[X_1]+Var[Y]\cdot(E[X_1]^2)$

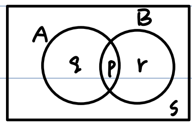
$=\frac{1}{p}\cdot 4p(1-p)+\frac{1-p}{p^2}\cdot(2p-1)^2=\frac{(1-p)(8p^2-4p+1)}{p^2}$

(3) $P(S_{n+k}=a+b\mid S_n=a)=P(S_k=b\mid S_0=0)=\begin{cases}C_k^{\frac{k+b}{2}}p^{\frac{k+b}{2}}(1-p)^{\frac{k-b}{2}} & 2\mid k+b\\ 0 & 2\nmid k+b\end{cases}$

$E[S_{n+k}|S_n]=E[S_n+X_{n+1}+\cdots+X_{n+k}|S_n]=S_n+\sum_{i=n+1}^{n+k}E[X_i]=S_n+k(2p-1)$

9. 试证明 $|P(A\cap B)-P(A)P(B)|\le \frac{1}{4}$ 并讨论等号成立的条件.

证:



① $p=P(A\cap B)$ $\quad q=P(A\setminus B)$ $\quad r=P(B\setminus A)$

$s=P(A^c\cap B^c)$ $\quad p+q+r+s=1$

$|P(A\cap B)-P(A)P(B)| = |p-(p+q)(p+r)| = |p(1-p-q-r)-qr| = |ps-qr|$

$\le \max\{ps,qr\} \le \frac{1}{4}\max\{(p+s)^2,(q+r)^2\}\le \frac{1}{4}$.

取等:($p=s=\frac{1}{2}$ 或 $q=r=\frac{1}{2}$) 且 ($ps=0$ 或 $qr=0$)

② $E[I_A]=p(A)$ $\quad E[I_B]=p(B)$ $\quad E[I_AI_B]=E[I_{AB}]=p(AB)$

$|E[I_AI_B]-E[I_A]E[I_B]|=|cov(I_A,I_B)|\le\sqrt{var(I_A)var(I_B)}=\sqrt{p(A)(1-p(A))\cdot p(B)\cdot(1-p(B))}\le\frac{1}{4}$