第六章 概率论与数理统计

重要公式与结论

（1）$P(\overline{A})=1-P(A)$

（2）$P(A\cup B)=P(A)+P(B)-P(AB)$

$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)$

（3）$P(A-B)=P(A)-P(AB)$ （4）$P(A\overline{B})=P(A)-P(AB), P(A)=P(AB)+P(A\overline{B})$

（5）条件概率 $P(\cdot\mid B)$ 满足概率的所有性质

例如：$P(\overline{A_1}\mid B)=1-P(A_1\mid B)$

$P(A_1\cup A_2\mid B)=P(A_1\mid B)+P(A_2\mid B)-P(A_1A_2\mid B)$

$P(A_1A_2\mid B)=P(A_1\mid B)P(A_2\mid A_1B)$

（6）若 $A_1,A_2,\cdots,A_n$ 相互独立，则 $P\left(\bigcap_{i=1}^{n}A_i\right)=\prod_{i=1}^{n}P(A_i)$ $P\left(\bigcup_{i=1}^{n}A_i\right)=1-\prod_{i=1}^{n}(1-P(A_i))$

（7）互斥、互逆与独立性之间的关系：

A 与 B 互逆 $\Rightarrow$ A 与 B 互斥，但反之不成立，

A 与 B 互斥（或互逆）且均非零事件 $\Rightarrow$ A 与 B 不独立。

（8）若 $A_1,A_2,\cdots,A_m,B_1,B_2,\cdots,B_n$ 相互独立，则 $f(A_1,A_2,\cdots,A_m)$ 与 $g(B_1,B_2,\cdots,B_n)$ 也相互独立，其中 $f(\cdot),g(\cdot)$ 分别表示对相应事件作任意事件运算后所得的事件，另外，概率为 1（或 0）的事件与任何事件相互独立。

二、一维随机变量及其概率分布

重要公式与结论：

（1）$X\sim N(0,1)\Rightarrow \varphi(0)=\frac{1}{\sqrt{2\pi}},\Phi(0)=\frac{1}{2},$ （2）$X\sim N(\mu,\sigma^2)\Rightarrow\frac{X-\mu}{\sigma}\sim N(0,1)$ 且

$\Phi(-a)=P(X\le -a)=1-\Phi(a)$

（3）$X\sim E(\lambda)\Rightarrow P(X>s+t\mid X>s)=P(X>t)$ （4）$P(X\le a)=\Phi\left(\frac{a-\mu}{\sigma}\right)$

$X\sim G(p)\Rightarrow P(X=m+k\mid X>m)=P(X=k)$

（5）离散型随机变量的分布函数为阶梯间断函数；连续型随机变量的分布函数为连续函数，但不一定为处处可导函数。

（6）存在既非离散也非连续型随机变量。

三、二维随机变量及其概率分布

重要公式与结论：

（1）边缘密度公式：$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy,\ f_Y(x)=\int_{-\infty}^{+\infty}f(x,y)dx,$ （2）$P[(X,Y)\in D]=\iint_D f(x,y)dxdy$

（3）若 (X,Y) 服从二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 则有

1．$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$ 2. X 与 Y 相互独立 $\Leftrightarrow \rho=0$，即 X 与 Y 不相关

3．$C_1X+C_2Y\sim N(C_1\mu_1+C_2\mu_2,C_1^2\sigma_1^2+C_2^2\sigma_2^2+2C_1C_2\sigma_1\sigma_2\rho)$

4．X 关于 $Y=y$ 的条件分布为：$N\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2)\right)$

Y 关于 $X=x$ 的条件分布为：$N\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(y-\mu_1),\sigma_2^2(1-\rho^2)\right)$

（4）若 X 与 Y 独立，且分别服从 $N(\mu_1,\sigma_1^2),N(\mu_2,\sigma_2^2)$ 则

$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,0),C_1X+C_2Y\sim N(C_1\mu_1+C_2\mu_2,C_1^2\sigma_1^2+C_2^2\sigma_2^2)$

（5）若 X 与 Y 相互独立，$f(x)$ 与 $g(x)$ 为连续函数，则 $f(X)$ 与 $g(Y)$ 也相互独立。

四、随机变量的数字特征

重要公式与结论：

（1）$D(X)=E(X^2)-E^2(X)$　　（2）$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$

（3）$|\rho(X,Y)|\le 1$，且 1. $\rho(X,Y)=1\Leftrightarrow P(Y=aX+b)=1$，其中 $a>0$。

2. $\rho(X,Y)=-1\Leftrightarrow P(Y=aX+b)=1$，其中 $a<0$。

（4）下面 5 个条件互为充要条件：

$\rho(X,Y)=0\Leftrightarrow \operatorname{cov}(X,Y)=0$

$\Leftrightarrow E(XY)=E(X)E(Y)$

$\Leftrightarrow D(X+Y)=D(X)+D(Y)$

$\Leftrightarrow D(X-Y)=D(X)+D(Y)$

注：X 与 Y 独立为上述 5 个条件中任何一个成立的充分条件，但非必要条件。

3．注意

本讲重点是随机变量函数的数学期望、方差，并掌握下列常见分布的期望与方差。

分布	数学期望	方差
（1）0－1 分布 $B(1,P)$	P	$P(1-p)$
（2）二项分布 $B(n,P)$	np	$np(1-p)$
（3）Poisson 分布 $P(\lambda)$	λ	λ
（4）正态分布 $N(\mu,\sigma^2)$	μ	σ^2
（5）均匀分布 $U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
（6）指数分布 $E(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
（7）几何分布 $G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
（8）超几何分布 $H(N,M,n)$	$\frac{nM}{N}$	$\frac{nM}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$

二、 大数定律和中心极限定理

重要定律

切比雪夫不等式	$P\{\lvert X-E(X)\rvert\ge\varepsilon\}\le\frac{D(X)}{\varepsilon^2}$ 或 $P\{\lvert X-E(X)\rvert<\varepsilon\}\ge 1-\frac{D(X)}{\varepsilon^2}$
切比雪夫大数定律	设 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立，且 $E(X_i)=\mu,D(X_i)=\sigma^2(i=1,2,\cdots)$ 对任意正数 ε，有 $\lim\limits_{n\to\infty}P\left\{\left\lvert\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right\rvert<\varepsilon\right\}=1$
伯努利大数定律	设 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立，同 0－1 分布 $B(1,P)$，则对任意正数 ε，有 $\lim\limits_{n\to\infty}P\left\{\left\lvert\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right\rvert<\varepsilon\right\}=1$

辛钦大数定律	设 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立同分布，$EX_i=\mu,i=1,2,\cdots$，则对任意正数 ε，有 $\lim\limits_{n\to\infty}P\left\{\left\|\frac{1}{n}\sum\limits_{i=1}^{n}X_i-\mu\right\|<\varepsilon\right\}=1$
列维－林德伯格定理	设 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立同分布，$EX_i=\mu,D(X_i)=\sigma^2(\sigma\neq 0)$，$i=1,2,\cdots$，则有 $\lim\limits_{n\to\infty}P\left\{\frac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\le x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$
棣莫佛－拉普拉斯定理	设 $\eta_n\sim B(n,p)$，（即 $X_1,X_2,\cdots,X_n$ 相互独立且同服从 0－1 分布 $\eta_n=\sum\limits_{i=1}^{n}X_i$）则有 $\lim\limits_{n\to\infty}P\left\{\frac{\eta_n-np}{\sqrt{np(1-p)}}\le x\right\}=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$

六、数理统计的基本概率

重要公式与结论：

（1）对于 $\chi^2\sim\chi^2(n)$，有 $E(\chi^2(n))=n,D(\chi^2(n))=2n$

（2）对于 $T\sim t(n)$ 有 $E(T)=0,D(T)=\frac{n}{n-2}(n>2)$

（3）对于 $F\sim F(m,n)$，有 $\frac{1}{F}\sim F(n,m),F_{\frac{\alpha}{2}}(m,n)=\frac{1}{F_{1-\frac{\alpha}{2}}(n,m)}$

（4）对于任意总体 X，有 $E(\overline{X})=E(X),E(S^2)=D(X),D(\overline{X})=\frac{D(X)}{n}$。

七、参数估计

重要公式与结论

（1）$E(\overline{X})=E(X),E(S^2)=D(X)$，即 $\overline{X},S^2$ 分别为总体 $E(X),D(X)$ 的无偏估计量。

（2）由大数定律易知 $\overline{X},S^2$ 也分别是 $E(X),D(X)$ 的一致估计量。

（3）若 $E(\hat{\theta})=\theta,D(\hat{\theta})\to 0(n\to\infty)$，则 $\hat{\theta}$ 为 θ 的一致估计。

（4）$\hat{\theta}$ 为 θ 的矩估计，$g(x)$ 为连续函数，则 $g(\hat{\theta})$ 为 $g(\theta)$ 的矩估计。

（5）$\hat{\theta}$ 为 θ 的极大似然估计，$g(x)$ 为单调函数，则 $g(\hat{\theta})$ 为 $g(\theta)$ 的极大似然估计。

（6）$(\hat{\theta}_1,\hat{\theta}_2)$ 为 θ 的置信度是 $1-\alpha$ 的置信区间，$g(x)$ 为单调函数，则（$g(\hat{\theta}_1),g(\hat{\theta}_2)$ 或 $g(\hat{\theta}_2),g(\hat{\theta}_1)$）为 $g(\theta)$ 的置信度是 $1-\alpha$ 的置信区间。

八、假设检验

	原假设 H_0	H_0 下的统计量及分布	H_0 的拒绝域

一个正态总体	$\mu=\mu_0$（σ^2 未知）	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$	$\lvert u\rvert=\left\lvert\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right\rvert\geq u_{1-\frac{\alpha}{2}}$
	$\mu=\mu_0$（σ^2 已知）	$T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$	$\lvert t\rvert=\left\lvert\dfrac{\overline{x}-\mu_0}{S/\sqrt{n}}\right\rvert\geq t_{1-\frac{\alpha}{2}}(n-1)$
一个正态总体	$\sigma^2=\sigma_0^2$（μ 已知）	$W=\sum\limits_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma_0}\right)^2\sim\chi^2(n)$	$w=\sum\limits_{i=1}^{n}\left(\dfrac{x_i-\mu}{\sigma_0}\right)^2\geq\chi^2_{(1-\frac{\alpha}{2})}(n)$ 或 $w\leq\chi^2_{\frac{\alpha}{2}}(n)$
	$\sigma^2=\sigma_0^2$（μ 未知）	$W=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$w=\dfrac{(n-1)S^2}{\sigma_0^2}\geq\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 或 $w\leq\chi^2_{1-\frac{\alpha}{2}}(n-1)$
两个正态总体	$\mu_1-\mu_2=\delta$ （σ_1^2,σ_2^2 已知）	$U=\dfrac{\overline{X}_1-\overline{X}_2-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$	$\lvert u\rvert=\left\lvert\dfrac{\overline{x}_1-\overline{x}_2-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\right\rvert\geq u_{1-\frac{\alpha}{2}}$
	$\mu_1-\mu_2=\delta$ （σ_1^2,σ_2^2 已知但 $\sigma_1^2=\sigma_2^2$）	$T=\dfrac{\overline{X}_1-\overline{X}_2-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$ $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$	$t=\left\lvert\dfrac{\overline{x}_1-\overline{x}_2-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\right\rvert$ $\geq t_{1-\frac{\alpha}{2}}(n_1+n_2-2)$
	$\sigma_1^2=\sigma_2^2$ （μ_1,μ_2 未知）	$F=\dfrac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)$	$f=\dfrac{S_1^2}{S_2^2}\geq F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)$ 或 $f\leq F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$