# 概率论与数理统计考点总结

## 一、填空选择部分

1、事件与概率的运算

• 事件的差：$A-B=A\overline{B}$

• 德摩根律：$\overline{A\cup B}=\overline{A}\cap\overline{B}$，$\overline{A\cap B}=\overline{A}\cup\overline{B}$

• 减法公式：$P(A-B)=P(A)-P(AB)$

• 对立事件的概率：$P(\overline{A})=1-P(A)$

• 加法公式：$P(A\cup B)=P(A)+P(B)-P(AB)$

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

• 条件概率：$P(A|B)=\frac{P(AB)}{P(B)}$

2、古典概型与伯努利概型

• 古典概型：$P=\frac{m}{n}=\frac{\text{有利场合数}}{\text{总场合数}}$

• 伯努利概型：$n$次伯努利实验中事件$A$发生$k$次的概率为$\mathrm{C}_n^k p^k q^{n-k}$

3、分布函数与密度函数

• 分布函数

$$0\le F(x)\le 1,\qquad F(-\infty)=0,\qquad F(+\infty)=1$$

• 密度函数

$$p(x)\ge 0,\qquad \int_{-\infty}^{+\infty}p(x)\mathrm{d}x=1$$

• 对于连续型随机变量：$F'(x)=p(x)$

4、常见的随机变量

• 0-1 分布（两点分布，伯努利分布）：$X\sim B(1,p)$

$$P\{X=k\}=p^k q^{1-k},k=0,1$$

$$EX=p,\qquad DX=pq$$

- 二项分布：$X \sim B(n,p)$

$$P\{X=k\}=\mathrm{C}_n^k p^k q^{n-k}, k=0,1,\cdots,n$$

$$EX=np,\qquad DX=npq$$

- 泊松分布：$X \sim P(\lambda)$

$$P\{X=k\}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, k=0,1,2,\cdots$$

$$EX=\lambda,\qquad DX=\lambda$$

- 均匀分布：$X \sim U[a,b]$

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b\\ 0, & x<a \text{ or } x>b\end{cases}$$

$$EX=\frac{a+b}{2},\qquad DX=\frac{(b-a)^2}{12}$$

- 指数分布：$X \sim E(\lambda)$

$$p(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$$

$$EX=\frac{1}{\lambda},\qquad DX=\frac{1}{\lambda^2}$$

- 正态分布：$X \sim N(\mu,\sigma^2)$

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty<x<+\infty$$

$$EX=\mu,\qquad DX=\sigma^2$$

- 期望与方差的关系：$DX=EX^2-(EX)^2$

5、正态分布的性质

- $X \sim N(\mu,\sigma^2)$的密度函数关于$x=\mu$对称
- $X \sim N(\mu,\sigma^2)$，则$Y=\frac{X-\mu}{\sigma} \sim N(0,1)$
- 标准正态$N(0,1)$：$\varphi(x)=\varphi(-x)$，$\Phi(-x)=1-\Phi(x)$
- $X_k \sim N(\mu_k,\sigma_k^2)$且相互独立，则

$$a_1X_1+\cdots+a_nX_n+b \sim (a_1\mu_1+\cdots+a_n\mu_n+b, a_1^2\sigma_1^2+\cdots+a_n^2\sigma_n^2)$$

6、Chebyshev 不等式

$$P\{|X-EX|\geq\varepsilon\}\leq\frac{DX}{\varepsilon^2}$$
$$P\{|X-EX|<\varepsilon\}\geq 1-\frac{DX}{\varepsilon^2}$$

7、抽样分布定理（以下均在各随机变量相互独立的条件下）

- $X_k\sim N(0,1)$，则$\sum_{k=1}^{n}X_k^2\sim\chi^2(n)$
- $X\sim N(0,1)$, $Y\sim\chi^2(n)$，则$\frac{X}{\sqrt{Y/n}}\sim t(n)$
- $X\sim\chi^2(n), Y\sim\chi^2(m)$，则$\frac{X/n}{Y/m}\sim F(n,m)$
- $X_1,\cdots,X_n$是来自总体$N(\mu,\sigma^2)$的样本，则样本均值$\overline{X}$与样本方差$S^2$相互独立，

$$\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right),\qquad \frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

8、统计量、无偏性与有效性

- 统计量，若函数$T=T(X_1,\cdots,X_n)$中不含任何未知参数，则$T$是一个统计量
- 无偏性：估计量$\hat{\theta}$满足$E\hat{\theta}=\theta$

  注意：样本均值$\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k$是总体期望的无偏估计

  样本方差$S^2=\frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\overline{X}\right)^2$是总体方差的无偏估计

  但$\tilde{S}^2=\frac{1}{n}\sum_{k=1}^{n}\left(X_k-\overline{X}\right)^2$<u>不是</u>总体方差的无偏估计
- 有效性：两个无偏估计量$\hat{\theta}_1,\hat{\theta}_2$，若$D\hat{\theta}_1\leq D\hat{\theta}_2$，则称$\hat{\theta}_1$比$\hat{\theta}_2$有效
- $X$和$Y$相互独立，则

$$E(aX\pm bY\pm c)=aE(X)\pm bE(Y)\pm c$$
$$E(XY)=E(X)E(Y)$$
$$D(aX\pm bY\pm c)=a^2D(X)+b^2D(Y)$$

9、其他

（1）$A=\emptyset\Rightarrow P(A)=0,\quad P(A)=0\nRightarrow A=\emptyset$

$A=\Omega\Rightarrow P(A)=1,\quad P(A)=1\nRightarrow A=\Omega$

（2）互不相容：$AB=\emptyset$

相互独立：$P(AB)=P(A)P(B)$

两者没有任何关系

$\emptyset$和任何事件都相互独立，$\Omega$和任何事件都相互独立

（3）$A$和$B$相互独立，意味着$P(A|B)=P(A)$

以下四组事件相互独立是等价的：$\{A,B\}\Leftrightarrow\{\overline{A},B\}\Leftrightarrow\{A,\overline{B}\}\Leftrightarrow\{\overline{A},\overline{B}\}$

（4）两两独立：$A$和$B$相互独立、$A$和$C$相互独立、$B$和$C$相互独立

相互独立：两两独立基础上满足$P(ABC)=P(A)P(B)P(C)$

## 二、计算题部分

1、全概率公式和 Bayes 公式

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)\qquad\qquad P(B_i|A)=\frac{P(B_i)P(A|B_i)}{\sum_{i=1}^{n}P(B_i)P(A|B_i)}$$

2、随机变量函数的分布

已知连续型随机变量$X$的密度函数为$f_X(x)$，则$Y=aX+b$的密度函数为

$$F_Y(y)=P\{Y\le y\}=P\{aX+b\le y\}=\begin{cases}P\left\{X\le\dfrac{y-b}{a}\right\}=F_X\left(\dfrac{y-b}{a}\right),a>0\\P\left\{X\ge\dfrac{y-b}{a}\right\}=1-F_X\left(\dfrac{y-b}{a}\right),a<0\end{cases}$$

$$f_Y(y)=F_Y'(y)=\begin{cases}F_X'\left(\dfrac{y-b}{a}\right)=\dfrac{1}{a}f_X\left(\dfrac{y-b}{a}\right),a>0\\-F_X'\left(\dfrac{y-b}{a}\right)=-\dfrac{1}{a}f_X\left(\dfrac{y-b}{a}\right),a<0\end{cases}$$

附：

- $Y=\mathrm{e}^X$

$$F_Y=P\{Y\le y\}=P\{\mathrm{e}^X\le y\}=\begin{cases}P\{X\le\ln y\}=F_X(\ln y), & y>0\\0, & y\le 0\end{cases}$$

$$f_Y(y)=F_Y'(y)=\begin{cases}\dfrac{1}{y}\cdot f_X(\ln y), & y>0\\0, & y\le 0\end{cases}$$

- $Y=\ln X$

$$F_Y=P\{Y\le y\}=P\{\ln X\le y\}=P\{X\le\mathrm{e}^y\}=F_X(\mathrm{e}^y)$$

$$f_Y(y) = F'_Y(y) = P\{X \le \mathrm{e}^y\} = \mathrm{e}^y \cdot f_X(\mathrm{e}^y)$$

- $Y = |X|$

$$F_Y = P\{Y \le y\} = P\{|X| \le y\} = \begin{cases} P\{-y \le X \le y\} = F_X(y) - F_X(-y), & y \ge 0 \\ 0, & y < 0 \end{cases}$$

$$f_Y(y) = F'_Y(y) = \begin{cases} f_X(y) + f_X(-y), & y \ge 0 \\ 0, & y < 0 \end{cases}$$

- $Y = X^2$

$$F_Y = P\{Y \le y\} = P\{X^2 \le y\}$$

$$= \begin{cases} P\{-\sqrt{y} \le X \le \sqrt{y}\} = F_X(\sqrt{y}) - F_X(-\sqrt{y}), & y \ge 0 \\ 0, & y < 0 \end{cases}$$

$$f_Y(y) = F'_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} \cdot f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} \cdot f_X(-\sqrt{y}), & y > 0 \\ 0, & y \le 0 \end{cases}$$

3、边缘分布和数学期望

已知二维连续型随机变量$(X,Y)$的联合密度函数为$f(x,y)$，则

- 随机变量$X$的边缘分布为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y$$

- 随机变量$Y$的边缘分布为

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x$$

- 随机变量$Z = g(X,Y)$的数学期望为

$$E(Z) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$$

4、中心极限定理

- $X \sim B(n,p)$

$$P\{a < X \le b\} = P\left\{\frac{a-np}{\sqrt{npq}} < \frac{X-np}{\sqrt{npq}} \le \frac{b-np}{\sqrt{npq}}\right\} = \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right)$$

- 已知$EX_i = \mu$，$DX_i = \sigma^2$，则对于$X = \sum_{i=1}^{n} X_i$，$EX = n\mu$，$DX = n\sigma^2$

$$P\{a < X \le b\} = P\left\{\frac{a-EX}{\sqrt{DX}} < \frac{X-EX}{\sqrt{DX}} \le \frac{b-EX}{\sqrt{DX}}\right\} = \Phi\left(\frac{b-EX}{\sqrt{DX}}\right) - \Phi\left(\frac{a-EX}{\sqrt{DX}}\right)$$

5、参数估计

已知连续型随机变量$X$的密度函数为$p(x;\theta)$，$\theta$为未知参数，$\overline{X}$为样本均值

• 矩法估计

$$EX = \int_{-\infty}^{+\infty} xp(x;\theta)\mathrm{d}x$$

令$EX = \overline{X}$，解出$\hat{\theta}$

• 极大似然估计

$$L(\theta) = \prod_{i=1}^{n} p(X_i;\theta)$$

求对数似然$\ln L(\theta)$，令$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta} = 0$，解出$\hat{\theta}$

6、假设检验

$$H_0: \mu = \mu_0, \qquad H_1: \mu \ne \mu_0$$

（$\sigma^2$已知）

$$\left|\frac{\overline{X}-\mu_0}{\sqrt{\sigma^2/n}}\right| = (\text{计算具体的值})(> \text{or} <) u_{\frac{\alpha}{2}} = \left(u_{\frac{\alpha}{2}}\text{的值}\right)$$

（$\sigma^2$未知）

$$\left|\frac{\overline{X}-\mu_0}{\sqrt{S^2/n}}\right| = (\text{计算具体的值})(> \text{or} <) t_{\frac{\alpha}{2}}(n-1) = \left(t_{\frac{\alpha}{2}}(n-1)\text{的值}\right)$$

若“>”则拒绝$H_0$，若“<”则接受$H_0$，并结合题目给出最终结论。

注意：$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$