《概率论与数理统计》课程总结

第一章主要内容及要求：

1）熟练掌握事件的关系与运算法则：包含、交、并、差、互不相容、对立等关系和德摩根定律.会用事件的关系表示随机事件.

$B\supset A$ 或 $A\subset B$. $A\cup B,\ A\cap B=AB$,

$A-B=A\overline{B}=A-AB$ $A\cap B=\emptyset$

$A\cap B=\emptyset;\ A\cup B=S$

德·摩根律：$\overline{A\cup B}=\overline{A}\cap\overline{B}$，$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

$\overline{\bigcup_{\alpha} A_\alpha}=\bigcap_{\alpha}\overline{A}_\alpha$，$\overline{\bigcap_{\alpha} A_\alpha}=\bigcup_{\alpha}\overline{A}_\alpha$

2）掌握概率的定义及性质，会求常用的古典概型 中的概率；

(1) 若$A_1,A_2,\cdots$是两两互不相容事件，则

$P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots$

(2) 若$A_1,A_2,\cdots,A_n$是两两互不相容事件，则

$P(A_1\cup A_2\cup\cdots\cup A_n)$

$=P(A_1)+P(A_2)+\cdots+P(A_n)$

(3) $A\subset B\Rightarrow P(B-A)=P(B)-P(A)$

(4) $P(\overline{A})=1-P(A)$

(5) $P(A\cup B)=P(A)+P(B)-P(AB)$

(6) $P(B-A)=P(B)-P(AB)$

3）熟练运用条件概率的定义，乘法公式，全概率公式，事件的独立性及性质求概率.

(1) $P(A|B)=\dfrac{P(AB)}{P(B)}$;

(2) $P(AB)=P(A)P(B|A)$;

(3) $P(B)=\sum_{k=1}^{n}P(A_k)P(B|A_k)$;

(4) $P(A_k\mid B)=\dfrac{P(A_kB)}{P(B)}=\dfrac{P(A_k)P(B\mid A_k)}{\sum_{j=1}^{n}P(A_j)P(B\mid A_j)}$,

(5) $P(AB)=P(A)P(B)$.

(6) $\begin{cases}P(AB)=P(A)P(B)\\ P(BC)=P(B)P(C)\\ P(AC)=P(A)P(C)\\ P(ABC)=P(A)P(B)P(C)\end{cases}$

A ,B,C 相互独立

(7) 若随机事件 A 与 B 相互独立，则

$\overline{A}$ 与 B、A 与 $\overline{B}$、$\overline{A}$ 与 $\overline{B}$ 也相互独立.

(8) 若 $A_1,A_2,\cdots A_n$是相互独立的事件，则

$P(A_1\cup A_2\cup\cdots\cup A_n)=$

$=1-P(\overline{A_1}\ \overline{A_2}\cdots\overline{A_n})=1-P(\overline{A_1})P(\overline{A_2})\cdots P(\overline{A_n})$

第二章主要内容及要求：

1）掌握随机变量分布函数的定义及性质：

$F(x)=P\{X\le x\}$

$F(x)$ 是一个单调不减右连续的函数；$0\le F(x)\le 1$;

$F(-\infty)=0,\ F(\infty)=1$;

$P\{a<X\le b\}=F(b)-F(a)$

$P\{X=a\}=F(a)-F(a-0)$

2）掌握离散型随机变量分布率的定义和性质，会求离散型随机变量的分布率；

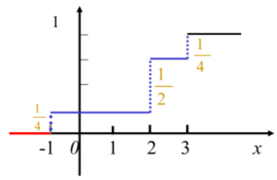
X	x_1	x_2	$\cdots,$	x_n	$\cdots$
P	p_1	p_2	$\cdots,$	p_n	$\cdots$

(1) 对任意的自然数 n，有 $p_n\ge 0$;

(2) $\sum_{n}p_n=1$.

3）会求离散型随机变量的分布函数；

X	-1	2	3
p_k	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$



4）掌握连续型随机变量概率密度的性质：会确定密度函数中的未知参数，掌握分布函数与概率密度的关系，会运用概率密度求连续型随机变量取值落在实轴某一区间上的概率.

(1) $F(x)=\int_{-\infty}^{x} f(t)dt$;

(2) $\int_{-\infty}^{\infty} f(x)dx=1$;

(3) $P\{x_1<X\le x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)dx$;

(4) $F'(x)=f(x)$. 当$f(x)$为连续函数时

5）理解伯努利试验，掌握两点分布及其概率背景；

$X\sim B(1,p)$

6）掌握二项分布的概率背景，即会把实际问题中服从二项分布的随机变量构设出来，运用有关公式求概率.

若 X 表示n重伯努利试验中成功出现的次数，

则 $X\sim B(n,p)$

$$P\{X=k\}=C_n^k p^k(1-p)^{n-k}\quad(k=0,\ 1,\ \cdots,\ n)$$

7）掌握泊松分布；

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

8）掌握均匀分布: $X\sim U[a,b]$

$$f(x)=\begin{cases}\dfrac{1}{b-a} & a\le x\le b\\ 0 & \text{其它}\end{cases}$$

9）掌握指数分布:

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & x\le 0.\end{cases}\qquad f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

10）掌握正态分布及其性质：理解一般正态分布函 数与标准正态分布函数的关系，会查表求概率，正态变量的线性变换仍然是正态变量

$X\sim N(0,\ 1)$:

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\qquad(-\infty<x<+\infty)$$

$$\Phi(-x)=1-\Phi(x)$$

$$F_X(x)=P\{X\le x\}=\Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$P\{a<X<b\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right).$$

若$X\sim N(\mu,\sigma^2)$,

有$Y=aX+b\sim N\left(a\mu+b,(a\sigma)^2\right)$.

11）会运用定理及先求分布函数法求随机变量函数的分布.

(1) $f_Y(y)=\begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta, \\ 0, & 其它. \end{cases}$

(2) 先求 $Y=g(X)$ 的分布函数

$$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=\int_{g(x)\le y} f_X(x)dx$$

利用 $Y=g(X)$ 的分布函数与密度函数 之间的关系求 $Y=g(X)$ 的密度函数 $f_Y(y)=F_Y'(y)$.

第三章主要内容及要求：

1）掌握二维离散型随机变量分布率的定义；会求二维离散型随机变量的分布率；

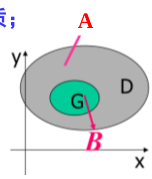
2）掌握二维连续型随机变量概率密度的性质：会运用概率密度求二维连续型随机变量取值落在平面某一区域上的概率.

$$P\{(X,Y)\in G\}=\iint_G f(x,y)dxdy.$$

3）掌握二维均匀分布的定义及性质；

$$f(x,\ y)=\begin{cases} \dfrac{1}{A} & (x,\ y)\in D \\ 0 & (x,\ y)\notin D \end{cases}$$

$$P\{(X,Y)\in G\}=\iint_G f(x,y)dxdy=\frac{B}{A}.$$



4）会求边缘分布率和边缘概率密度；

$$p_{i\cdot}=P\{X=x_i\}=\sum_j p_{ij} \qquad f_X(x)=\int_{-\infty}^{+\infty} f(x,\ y)dy$$

$$p_{\cdot j}=P\{Y=y_j\}=\sum_i p_{ij} \qquad f_Y(y)=\int_{-\infty}^{+\infty} f(x,\ y)dx$$

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_{i\cdot}$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	

5）掌握随机变量独立性的充分必要条件:

$$p_{ij}=p_{i\cdot}\,p_{\cdot j}$$

$$f(x,\ y)=f_X(x)f_Y(y)$$

6）会求二维随机变量函数的分布：

（1）一般情形

先求随机变量函数 $Z=g(X,\ Y)$ 的分布函数 $F_Z(z)$,

再求随机变量函数 $Z=g(X,\ Y)$ 的密度函数 $f_Z(z)=F_Z'(z)$,

（2）和的分布

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx$$

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy$$

（3）极值分布

$$F_{(n)}(x)=P\{X_{(n)}\le x\}=\prod_{i=1}^{n}F_i(x)$$

$$F_{(1)}(x)=P\{X_{(1)}\le x\}=1-\prod_{i=1}^{n}[1-F_i(x)]$$

7）掌握正态分布的性质：

如果随机变量 X_1， X_2， …， X_n 相互独立

$$X_i \sim N(\mu_i,\ \sigma_i^2)$$

令： $Z=\sum_{i=1}^{n} a_i X_i$，

则 $Z \sim N\left(\sum_{i=1}^{n} a_i \mu_i,\ \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$

第四章主要内容及要求：

1）熟练掌握期望定义和性质；

$$EX=\sum_{i=1}^{\infty} x_k p_k$$

$$EX=\int_{-\infty}^{+\infty} xf(x)dx$$

$$E(\sum_{i=1}^{n} a_i X_i)=\sum_{i=1}^{n} a_i EX_i$$

X,Y 不相关 $\Leftrightarrow EXY=EXEY.$

2）会求随机变量函数的数学期望；

设 Y=g(X), g(x) 是连续函数，

则 $EY=\sum_{k=1}^{\infty} p_k g(x_k)$

$$EY=\int_{-\infty}^{\infty} g(x)f(x)dx$$

若 $Z=g(X,Y)$

则 $EZ=\sum_{i,j=1}^{\infty} g(x_i,y_j)p_{ij}$

$$EZ=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy$$

3）熟练掌握方差的定义和性质；

$$DX=E(X-EX)^2=E(X^2)-(EX)^2$$

$$D(cX)=c^2 DX$$

$$D(aX+bY)=a^2DX+b^2DY+2abE(X-EX)(Y-EY)$$
$$=a^2DX+b^2DY+2abCOV(X,Y)$$

若 X,Y 不相关，则 $D(aX+bY)=a^2DX+b^2DY.$

4）掌握契比雪夫不等式

$$P\{|X-EX|\ge\varepsilon\}\le DX/\varepsilon^2;$$

5）熟记两点分布、二项分布、泊松分布、均匀分布、 正态分布、指数分布的期望值和方差值.

6）掌握协方差和相关系数的定义，不相关的定义及独立与不相关的关系；

$$COV(X,Y)=E(X-EX)(Y-EY)=EXY-EXEY$$

$$\rho_{XY}=\frac{COV(X,Y)}{\sqrt{DX}\sqrt{DY}}$$

若 $\rho_{XY}=0$，称 X,Y 不相关。

若X，Y独立，则 X , Y 不相关.(反之，不然）

第五章主要内容及要求：

1）掌握大数定律的定义；

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k-\frac{1}{n}\sum_{k=1}^{n} EX_k\right|<\varepsilon\right\}=1,$$

或 $\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k-\frac{1}{n}\sum_{k=1}^{n} EX_k\right|\ge\varepsilon\right\}=0,$

2）掌握辛钦、伯努利大数定律；

2）掌握独立同分布的中心极限定理和德莫佛-拉普拉斯定理；并会用这两个定理求概率；

$$\lim_{n\to\infty} P\{\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \Phi(x).$$

设随机变量 $\eta_n \sim B(n, p)(n = 1,2,\cdots)\ (0 < p < 1)$,

$$\lim_{n\to\infty} P\{\frac{\eta_n - np}{\sqrt{npq}} \le x\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \Phi(x).$$

第六章主要内容及要求：

1）掌握统计量的概念，会判断哪些样本的函数是统计量；

2）掌握正态总体的样本均值和样本方差的定义及其分布；

第七章主要内容及要求：

要会熟练运用矩法和极大似然法求估计量.

矩法求估计量的步骤：

(1) 求 $\mu_1 = EX$；

(2) 令 $A_1 = \mu_1$；

(3) 解上面方程，得 $\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n)$.

极大似然法求估计量的步骤：(一般情况下)

(1) 构造似然函数 $L(\theta)$：

$$L(\theta) = \prod_{i=1}^{n} f(x_i)\ (\text{连续型});$$

(2) 取对数：$\ln L(\theta)$;

(3) 令 $\dfrac{d\ln L}{d\theta} = 0$;

(4) 解似然方程得 θ 的极大似然估计量 $\hat{\theta}$.

说明：若似然方程（组）无解，或似然函数不可导， 此法失效，改用其它方法

例　设 $X \sim U[a,b]$；a,b 未知，$x_1,\cdots,x_n$ 是一个样本值，求：a,b 的极大似然估计量 .

分析：X 的概率密度为：$f(x;a,b) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b; \\ 0, & \text{其它} \end{cases}$

似然函数为

$$L(a,b) = \frac{1}{(b-a)^n} \quad (a < x_i < b,\ \ i = 1,\ 2,\ \cdots,\ n)$$

$$\ln L(a,b) = -n\ln(b-a),$$

$$\frac{\partial}{\partial a}\ln L(a,b) = \frac{n}{b-a} = 0;\quad \frac{\partial}{\partial b}\ln L(a,b) = -\frac{n}{b-a} = 0,$$

显然，似然方程组无解，但这不能说明不存在极大似然估计量，只是不能由似然方程组求解.

例（续）

解：将 $x_1,\cdots,x_n$ 按从小到大顺序排列成

$$x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)},$$

则 $L(a,b) = \begin{cases} \dfrac{1}{(b-a)^n}, & a \le x_{(1)} \le \cdots \le x_{(n)} \le b; \\ 0, & \text{其它} \end{cases}$

对于满足 $a \le x_{(1)} \le \cdots \le x_{(n)} \le b$ 的任意 a,b 有

$$L(a,b) = \frac{1}{(b-a)^n} \le \frac{1}{(x_{(n)} - x_{(1)})^n}$$

例（续）

即：$L(a,b)$在$a=x_{(1)}, b=x_{(n)}$时，

取最大值$(x_{(n)}-x_{(1)})^n$

故a,b的极大似然估计值为：

$$\hat{a}=x_{(1)}=\min x_i,\quad \hat{b}=x_{(n)}=\max x_i,$$

故a,b的极大似然估计量为：

$$\hat{a}=\min X_i,\quad \hat{b}=\max X_i.$$

求置信区间的步骤：

(1) 找一个样本的函数 $Z=Z(X_1,\cdots,X_n;\theta)$，它包含待估参数 θ，而不包含其它未知参 数.且Z的分布是已知的，不依 赖于未知参数 θ.

(2) 对给定的置信度 $1-\alpha$，确定常数 a,b,使
$P\{a<Z<b\}=1-\alpha$.

(3) 从不等式$a<Z(X_1,\cdots X_n;\theta)<b$得到等价的不等式
$\hat{\theta}_1<\theta<\hat{\theta}_2$,其中$\hat{\theta}_1=\hat{\theta}_1(X_1,\cdots X_n)$，$\hat{\theta}_2=\hat{\theta}_2(X_1,\cdots X_n)$
都是统计量.

(4) 随机区间 $[\hat{\theta}_1,\ \hat{\theta}_2]$ 是置信度为$1-\alpha$ 的置信区间.

一个正态总体未知参数的置信区间

待估参数		随机变量	随机变量的分布	双侧置信区间的上、下限
μ	σ^2已知	$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$N(0,\ 1)$	$\overline{X}\pm z_{\frac{\alpha}{2}}\cdot\dfrac{\sigma}{\sqrt{n}}$
	σ^2未知	$\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$	$t(n-1)$	$\overline{X}\pm t_{\frac{\alpha}{2}}(n-1)\cdot\dfrac{S}{\sqrt{n}}$
σ^2	μ已知	$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$	$\chi^2(n)$	$\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}$　$\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}$
	μ未知	$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2$	$\chi^2(n-1)$	$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}$　$\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$

第八章主要内容及要求：

1）单正态总体均值，方差的假设检验；

2）假设检验的步骤：

1. 根据实际问题的要求，提出原假设 H_0 及备择假设 H_1；
2. 选择适当的检验统计量,在H_0成立的条件下,确定它的概率分布;
3. 给定显著性水平 α,确定拒绝域W_1；
4. 根据样本观察值计算统计量的值;
5. 根据统计量值是否落入拒绝域W_1中，作出拒绝或者接受H_0的判断.

注：第一步靠理解题意，第二、三步背统计量(关键)；四、五步计算。