概率统计复习提纲

第一部分　随机事件和概率

一、基本概念

事件：随机事件、不可能事件、必然事件。事件的运算：并（和）、交（积）、差、逆（对立事件）以及事件的运算律。事件的关系：包含、互不相容、相互独立。概率、频率、条件概率。概率的公理化定义。

二、基本公式：

概率的一般加法公式： $P(A\cup B)=P(A)+P(B)-P(AB)$。

加法公式的几个变形：

1、$P(A)=1-P(\overline{A})$；

2、若 A_1，A_2，・・・，A_n，・・・互不相容，则 $P(\bigcup_j A_j)=\sum_j P(A_j)$；

3、若 $A\supset B$，则 $P(A-B)=P(A)-P(B)$，一般情况下有 $P(A-B)=P(A)-P(AB)$。

乘法公式：

1、一般情况有 $P(AB)=P(B)P(A|B)$（$P(B)>0$ 时）；

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)\cdots P(A_n|A_1A_2\cdots A_{n-1})\text{（}P(A_1A_2\cdots A_{n-1})>0\text{时）。}$$

2、若 A_1，A_2，・・・，A_n 相互独立，则 $P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n)$。

全概率公式：　$P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$

贝叶斯公式：　$P(B_k|A)=\dfrac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n}P(B_i)P(A|B_i)}$。

全概率公式和贝叶斯公式成立的条件： 1）$P(B_k)>0$；2）$B_1\cup B_2\cup\cdots\cup B_n=S$

注意：　有无穷多个 B_j 的时候也有完全类似的全概率公式和贝叶斯公式

三、基本题型（计算）

1、简单应用乘法公式、加法公式求概率；

例如：在独立性的假定或者互斥的假定下，已知 $P(A)$，$P(B)$，$P(A\cup B)$ 中的两个求第三个等。

2、古典概型；

解这类题目的时候注意：一是注意组合一定不能有任何顺序；排列就一定要考虑所有的顺序。例如在考虑“任取 5 个球至少取到两个红球”的时候，不能“先取两个红球再任取 3 个球”。“任取 6 只鞋至少有一双”的时候，不能“先取一双再任取 4 只”。

考虑事件 A 所包含的样本点数时，采用“能够使得 A 发生的所有方法有多少”的思维方式计算。重点是“所有”两个字。

3、应用全概率公式和贝叶斯公式分别求概率和条件概率；

首先，公式要正确。其次，定义事件时不能“定义 B_j 为在宿舍、图书馆、食堂找到卡片”，而应该“定义 $B_j(i=1,2,3)$ 分别为在宿舍、图书馆、食堂丢失卡片”。再次，不能把条件概率与积事件的概率混淆。不能定义 D=“已经找到卡片的条件下，是在宿舍找到”！这个不是事件，题目要求的是条件概率 P（在宿舍丢失卡片|找到卡片）。

4、应用独立性求概率：

5、电路问题：注意我在例题中强调的，不要分“路径”而要“剪线”分成串联和并的嵌套。

伯努利概型----要满足下列条件：一是每次试验只考虑两个互逆事件 A、$\overline{A}$；二是各次试验要独立（即互不影响），三是 A、$\overline{A}$ 在各次试验中的概率不变。

射击问题----要根据已知条件判断是否符合“互不影响（独立）、命中率不变”等特点。

6、条件概率问题----特别注意不要把条件概率与两个事件的积事件的概率，也不要犯上述“全概率公式问题”中的错误。

第二部分 随机变量（一维和二维）

一、基本概念：

随机变量（离散型和连续型），分布函数，密度函数，概率分布，联合概率分布与边缘概率分布，联合密度与边缘密度，随机变量的独立性，随机变量的数学期望、方差、原点矩和中心矩。

一维和二维概率分布的性质；一维分布函数的性质；一维和二维密度的性质；

二、常见分布：

（0—1）分布，二项分布，泊松分布，均匀分布，指数分布，正态分布。这些分布的密度（或概率分布）、数学期望与方差。

三、常见的题目类型：

1、利用（一维和二维）概率分布、一维分布函数、（一维和二维）密度函数的性质求待定常数。

分别应用性质 $\sum_{i=1}^{+\infty} p_i = 1$ 和 $\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} p_{ij} = 1$；$F(+\infty)=1,\ F(-\infty)=0$ 以及 $F(x)$ 右连续；$\int_{-\infty}^{+\infty} f(x)dx = 1$，$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,\ y)dxdy = 1$。

2、求离散型随机变量（一维和二维）的概率分布。----参加书上的例题

3、已知连续型随机变量的分布函数求密度函数：

$$f(x)=\begin{cases} F'(x), & F'(x)\text{存在，}\\ 0, & F'(x)\text{不存在。}\end{cases}$$

4、已知随机变量的概率分布、分布函数或密度函数求概率。离散型：$P(X\in D)=\sum_{x_i\in D}^{+\infty} P(X=x_i)$；

分布函数：$P(a<X\le b)=F(b)-F(a)$；连续型：$P(a<X<b)=\int_a^b f(x)dx$

5、已知一维随机变量的概率分布或密度函数求分布函数。离散型见第 2.3 节的例题 1 的方法；

连续型用公式：$F(x)=\int_{-\infty}^{x} f(t)dt$

5、已知随机变量的概率分布（一维和二维）或密度函数（一维），求它的函数的概率分布或密度函数。

随机变量 X 的密度函数时：直接利用 X 的密度函数连续型用公式：$f(x)$ 用求概率的思想先解关于 x 的不等式 $g(x)\le y$，然后把 $F_Y(y)=\int_{g(x)\le y} f(x)dx$ 写成可以“形式”对 y 求导的式子。最后形式求导得到函数 $Y=g(X)$ 的密度。

6、已知联合分布（概率分布、分布函数或密度函数）求边缘分布（离散型和连续型）和条件分布（离散型）。----参见书上的例题。

7、已知联合分布（概率分布、分布函数或密度函数）判断两个分量的独立性（离散型和连续型）。----参见书上的例题和 PPT 上补充的方法。

8、求数学期望与方差及相关系数、函数的数学期望（含离散型和连续型、一维和二维）。----注意两点：一是要复习《高等数学》中定积分的计算方法。二是带密度的时候，由于我们的密度常常有多个表达式，一定要注意。要先把积分分解成几个部分的积分使得每一部分密度只有一个表达式，然后才代表达式。

第三部分 统计

一、基本概念： 总体，简单随机样本，样本均值，样本方差，样本矩，χ^2-分布，t-分布，自由度，统计量，矩估计，似然函数，最大似然估计，参数估计的无偏性、有效性和一致性，区间估计及置信度，原假设和备择假设，拒绝域，显著性水平，第一类错误。

二、定理、公式与常见的题目类型

抽样分布：

定理. 设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu,\sigma^2)$的样本, 有

(1) $\overline{X}\sim N(\mu,\sigma^2/n)$ (定理2);　(2) $\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$;

（3）$\overline{X}$与S^2独立（定理3）；　（4）$\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$ （定理4）。

1、应用上述定理解不等式、计算概率（参见例题和作业题）。
2、判断参数估计的无偏性、有效性等。----归结为计算期望和方差。
3、求参数的矩估计。----注意：矩法中的方程是我们人为建立的，要用“用…..替换……”或“令……”这样的句子叙述。不能用“因为”、“所以”等语句叙述。
4、求参数的最大似然估计。----一是不要把似然函数写成$L(\theta)=\prod_{i=1}^{n} f(x)$。二是“求导数并令导数等于 0”的方法只有当似然函数不在参数空间的边界取最大值的时候才能用。
5、当$\sigma^2>0$已知时，总体分布$N(\mu,\ \sigma^2)$中参数μ的假设检验（双边，左边，右边）。
6、当$\sigma^2>0$未知时，总体分布$N(\mu,\ \sigma^2)$中参数μ的假设检验（双边，左边，右边）。
7、总体分布$N(\mu,\ \sigma^2)$中参数$\sigma^2>0$的假设检验（双边，左边，右边）。

做假设检验的题目要注意：一是题目问的是“关于随机变量的分布的一个判断对不对？”二是先看

是关于均值还是关于方差的假设检验问题，如果是关于均值的假设检验问题再看方差是否已知。最后看是左边、右边还是双边。这样就不会用错方法了。