概率论与数理统计历年考点总结

Contents

1 第一章：随机事件与概率

1.1 基础概念（小白必看）

- 互斥（互不相容）：事件 A 和 B 不能同时发生。$AB = \emptyset$。
- 独立：事件 A 发生与否不影响 B 发生的概率。
- 完备事件组：一组事件 $B_1, B_2, ..., B_n$ 两两互斥，且它们的并集是整个样本空间（即必然有一个会发生）。

1.2 核心公式

- 加法公式：$P(A \cup B) = P(A) + P(B) - P(AB)$
  - 如果 A, B 互斥，则 $P(A \cup B) = P(A) + P(B)$。

例题（加法与独立性）：已知 $P(A) = 0.4, P(B) = 0.3$，且 A, B 相互独立，求 $P(A \cup B)$。 解答：
1. 因为独立，所以 $P(AB) = P(A)P(B) = 0.4 \times 0.3 = 0.12$。
2. $P(A \cup B) = P(A) + P(B) - P(AB) = 0.4 + 0.3 - 0.12 = 0.58$。

- 条件概率：$P(A|B) = \frac{P(AB)}{P(B)}$
  - 含义：在已知 B 发生的前提下，A 发生的概率。
- 乘法公式：$P(AB) = P(A)P(B|A)$
- 全概率公式：$P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$
  - 直观理解：要把一个复杂事件 A 发生的概率，拆分成在各种不同情况 B_i 下发生的概率之和。
- 贝叶斯公式：$P(B_j \mid A) = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$
  - 直观理解：执果索因。已知结果 A 发生了，问它是由原因 B_j 导致的概率是多少。

例题（贝叶斯）：某工厂有两台机器生产同一种产品，甲机器产量占60%，合格率为95%；乙机器产量占40%，合格率为90%。现从仓库中随机抽取一件产品，发现是不合格品，求这件产品是由乙机器生产的概率。

解答：
1. 设 A 为“抽到不合格品”，B_1 为“甲生产”，B_2 为“乙生产”。
2. 已知：$P(B_1) = 0.6, P(B_2) = 0.4$；$P(A|B_1) = 1 - 0.95 = 0.05, P(A|B_2) = 1 - 0.90 = 0.10$。
3. 由全概率公式，$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) = 0.6 \times 0.05 + 0.4 \times 0.10 = 0.07$。
4. 由贝叶斯公式，$P(B_2|A) = \frac{P(B_2)P(A|B_2)}{P(A)} = \frac{0.4\times 0.10}{0.07} \approx 0.571$。

- 独立性：A, B 独立 $\Leftrightarrow P(AB) = P(A)P(B)$

1.3 解题思路

1. 古典概型：利用排列组合 C_n^m 和 A_n^m 计算。注意“放回”与“不放回”。
2. 全概率与贝叶斯：
   - 第一步：定义事件。通常 A 是最后那个“结果”（如：抽到红球、产品合格），B_i 是导致结果的各种“途径”或“原因”（如：选了甲箱子、由乙机器生产）。
   - 第二步：列出已知的 $P(B_i)$（先验概率）和 $P(A|B_i)$（条件概率）。
   - 第三步：代入公式。如果题目问“…的概率是多少”，用全概率；如果问“已知…，求它是…的概率”，用贝叶斯。

2 第二章：一维随机变量及其分布

2.1 基础概念

- 离散型：变量取值是有限个或可列个（如：$1, 2, 3$）。用分布律 $P(X = x_i) = p_i$ 表示。
- 连续型：变量取值在一个区间内。用概率密度 $f(x)$ 表示。注意：连续型变量取某一个具体点的概率为0。
- 分布函数 $\boldsymbol{F(x)}$：表示 X 落在 $(-\infty, x]$ 区间的概率。它是概率的累加。

2.2 核心公式

- 分布函数：$F(x) = P(X \le x)$。
  - 性质：$0 \le F(x) \le 1$；单调不减；右连续；$F(-\infty) = 0, F(+\infty) = 1$。
- 密度函数 $\boldsymbol{f(x)}$：
  - 性质：$f(x) \ge 0$；$\int_{-\infty}^{+\infty} f(x)dx = 1$（归一性，常用于求未知常数）。
  - 概率计算：$P(a < X \le b) = F(b) - F(a) = \int_a^b f(x)dx$。
- 常见分布：
  - 二项分布 $\boldsymbol{B(n,p)}$：n 次独立重复试验中，事件发生 k 次的概率。

    例题：某射手射击命中率为0.8，独立射击5次，求至少命中4次的概率。 解答：$X \sim B(5, 0.8)$。$P(X \ge 4) = P(X = 4) + P(X = 5) = C_5^4(0.8)^4(0.2)^1 + C_5^5(0.8)^5(0.2)^0 \approx 0.737$。

  - 泊松分布 $\boldsymbol{P(\lambda)}$：通常用于描述单位时间内随机事件发生的次数。

    例题：某服务台每小时平均接到3次呼叫，求某小时内恰好接到2次呼叫的概率。 解答：$X \sim P(3)$。$P(X = 2) = \frac{3^2e^{-3}}{2!} \approx 0.224$。

  - 正态分布 $\boldsymbol{N(\mu, \sigma^2)}$：
    - 标准化：如果 $X \sim N(\mu, \sigma^2)$，则 $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$。
    - 查表：$P(X \le a) = \Phi\left(\frac{a-\mu}{\sigma}\right)$。

查表小贴士（小白必看）：
1. 什么是 $\boldsymbol{\Phi(x)}$ 表：它给出了标准正态分布 $N(0,1)$ 下，变量小于等于 x 的概率。
2. 如何查正数：比如查 $\Phi(1.24)$。在表格左侧找行 `1.2`，在上方找列 `0.04`，交叉点的值就是结果。

3. 如何查负数：考试给的表通常只有正数。利用对称性公式： $\mathbf{\Phi(-x) = 1 - \Phi(x)}$ 。
   - 例如：$\Phi(-1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$。
4. 区间概率：$P(a < Z < b) = \Phi(b) - \Phi(a)$。

计算器使用提醒：
- **Normal PD (Probability Density)**：算的是概率密度函数 $f(x)$ 的值，即钟形曲线在 x 点的高度。考试几乎不用这个。
- **Normal CD (Cumulative Distribution)**：算的是 区间概率 $P(a \le X \le b)$。
  - **Lower** (下限)：区间的左端点 a。
  - **Upper** (上限)：区间的右端点 b。
  - 结果：就是曲线下从 a 到 b 的 面积。如果你想算 $P(X \le x)$，只需把下限设为一个极小值（如 −9999），上限设为 x。

例题（正态分布）：已知某校学生身高 $X \sim N(170, 5^2)$（单位：cm），求身高在 165cm 到 175cm 之间的学生比例。

解答：
1. 标准化：我们要算 $P(165 < X < 175)$。把 X 变成 Z。
   - 左边：$\frac{165-170}{5} = -1$
   - 右边：$\frac{175-170}{5} = 1$
   - 得到：$P(-1 < Z < 1)$。
2. 展开公式：$P(-1 < Z < 1) = \Phi(1) - \Phi(-1)$。
3. 处理负数：因为表里没 -1，所以换成 $\Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1$。
4. 查表计算：
   - 在 $\Phi(x)$ 表中找到 $x = 1.0$ 对应的值为 0.8413。
   - 最终结果：$2 \times 0.8413 - 1 = 0.6826$。

2.3 解题思路
1. 求分布律/密度函数中的常数：利用所有概率之和为 1，或密度函数全区间积分为 1。
2. 求分布函数：
   - 离散型：把小于等于 x 的所有概率加起来。
   - 连续型：$F(x) = \int_{-\infty}^{x} f(t)dt$。如果是分段函数，记得分段积分并累加前面的部分。
3. 随机变量函数的分布：已知 X 的密度 $f_{X(x)}$，求 $Y = g(X)$ 的密度。
   - 分布函数法（万能）：先写出 $F_{Y(y)} = P(g(X) \le y)$，解出关于 X 的不等式，代入 X 的分布计算，最后对 y 求导得到 $f_{Y(y)}$。

3 第三章：多维随机变量及其分布

3.1 基础概念
- 联合分布：同时考虑两个变量 X 和 Y 的取值情况。
- 边缘分布：只考虑其中一个变量，不管另一个变量取什么值。
- 独立性：X 的取值不影响 Y 的取值分布。

3.2 核心公式
- 边缘密度：
  - $f_{X(x)} = \int_{-\infty}^{+\infty} f(x, y)dy$（对 y 积分，把 y "积掉"，剩下的就是 x 的分布）。
  - $f_{Y(y)} = \int_{-\infty}^{+\infty} f(x, y)dx$（对 x 积分，把 x "积掉"）。
- 独立性判断：X, Y 独立 $\Leftrightarrow f(x, y) = f_{X(x)} f_{Y(y)}$。
- 条件密度：$f_{X|Y}(x|y) = \frac{f(x,y)}{f_{Y(y)}}$。

例题（独立性）：已知 $f(x, y) = \begin{cases} e^{-(x+y)} x>0 \\ y>0 \\ 0 \quad \text{其他} \end{cases}$，判断 X, Y 是否独立。 解答：

1. $f_{X(x)} = \int_0^{+\infty} e^{-x} e^{-y} dy = e^{-x}[-e^{-y}]_0^{+\infty} = e^{-x}(x > 0)$。
2. 同理 $f_{Y(y)} = e^{-y}(y > 0)$。
3. 因为 $f(x, y) = e^{-x} e^{-y} = f_{X(x)} f_{Y(y)}$，所以 X, Y 相互独立。

例题（边缘密度）：已知 (X, Y) 的联合密度为 $f(x, y) = \begin{cases} 2 & 0<x<y<1 \\ 0 & \text{其他} \end{cases}$，求 $f_{X(x)}$。

解答：
1. 确定 x 的范围：由 $0 < x < y < 1$ 知 $0 < x < 1$。
2. 对 y 积分：固定 x，看 y 的范围。由 $x < y < 1$ 知 y 从 x 积到 1。
3. 计算：$f_{X(x)} = \int_x^1 2dy = [2y]_x^1 = 2(1 - x)$。

故 $f_{X(x)} = \begin{cases} 2(1-x) & 0<x<1 \\ 0 & \text{其他} \end{cases}$。

3.3 解题思路

1. 求边缘分布：
   - 关键点：确定积分限。一定要画出 $f(x, y)$ 非零的区域图。
   - 如果求 $f_{X(x)}$，就画一条垂直于 x 轴的直线，看它穿过区域的上下边界。
2. 判断独立性：先分别求出两个边缘密度，看它们的乘积是否等于题目给出的联合密度。
3. 两个随机变量之和的分布：$Z = X + Y$。
   - 卷积公式：$f_{Z(z)} = \int_{-\infty}^{+\infty} f(x, z - x) dx$。
   - 如果 X, Y 独立，则 $f_{Z(z)} = \int_{-\infty}^{+\infty} f_{X(x)} f_{Y(z-x)} dx$。

4 第四章：随机变量的数字特征

4.1 基础概念

- 数学期望 $\boldsymbol{E(X)}$：随机变量取值的"加权平均数"，代表分布的中心位置。
- 方差 $\boldsymbol{D(X)}$：随机变量偏离期望的程度，代表分布的"离散程度"或"波动大小"。
- 协方差与相关系数：衡量两个变量之间的线性相关程度。

4.2 核心公式

- 期望：
  - 离散型：$E(X) = \sum x_i p_i$
  - 连续型：$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$
  - 函数期望：$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$
- 方差：$D(X) = E(X^2) - [E(X)]^2$ （常用计算公式，比定义式好用）。
- 协方差：$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$。
  - 如何计算 $\boldsymbol{E(XY)}$：
    - 离散型：$E(XY) = \sum\sum x_i y_j P(X = x_i, Y = y_j)$。
    - 连续型：$E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f(x, y) dx dy$。
- 相关系数：$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$。

例题（协方差计算）：已知 (X, Y) 的联合分布律如下表，求 $\mathrm{Cov}(X, Y)$。

$X \setminus Y$	0	1
0	0.2	0.3
1	0.1	0.4

解答：

1. 先求边缘分布：$P(X=0)=0.2+0.3=0.5, P(X=1)=0.1+0.4=0.5$。$P(Y=0)=0.2+0.1=0.3, P(Y=1)=0.3+0.4=0.7$。
2. 计算期望：$E(X)=0\times0.5+1\times0.5=0.5$。$E(Y)=0\times0.3+1\times0.7=0.7$。
3. 计算 $E(XY)$：$E(XY)=(0\times0\times0.2)+(0\times1\times0.3)+(1\times0\times0.1)+(1\times1\times0.4)=0.4$。
4. 计算协方差：$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0.4-0.5\times0.7=0.4-0.35=0.05$。

- $|\rho_{XY}|\le 1$。如果 $|\rho|=1$，说明有线性关系；如果 $\rho=0$，说明不相关。

4.3 解题思路

1. 计算 $\boldsymbol{E(X)}$ 和 $\boldsymbol{D(X)}$：
   - 如果是常见分布，直接背公式：
     - $X\sim B(n,p)\Rightarrow E=np, D=np(1-p)$
     - $X\sim P(\lambda)\Rightarrow E=\lambda, D=\lambda$
     - $X\sim N(\mu,\sigma^2)\Rightarrow E=\mu, D=\sigma^2$
2. 利用性质（填空题常客）：
   - $E(aX+b)=aE(X)+b$
   - $E(X+Y)=E(X)+E(Y)$（无论是否独立都成立）
   - $D(aX+b)=a^2D(X)$
   - 如果 X,Y 独立，则 $D(X+-Y)=D(X)+D(Y)$。

例题：已知 $E(X)=2, D(X)=1, E(Y)=3, D(Y)=2$，且 X,Y 独立。求 $E(2X-3Y)$ 和 $D(2X-3Y)$。解答：
1. $E(2X-3Y)=2E(X)-3E(Y)=2(2)-3(3)=4-9=-5$。
2. $D(2X-3Y)=2^2D(X)+(-3)^2D(Y)=4(1)+9(2)=4+18=22$。

3. 判断不相关：计算 $\mathrm{Cov}(X,Y)$，看是否为 0。

5 第五章：大数定律及中心极限定理

5.1 核心公式

- 切比雪夫不等式：$P(|X-E(X)|\ge\varepsilon)\le\frac{D(X)}{\varepsilon^2}$

例题：已知 $E(X)=100, D(X)=10$，估计 $P(90<X<110)$ 的下限。解答：$P(90<X<110)=P(|X-100|<10)=1-P(|X-100|\ge10)$。由切比雪夫不等式，$P(|X-100|\ge10)\le\frac{10}{10^2}=0.1$。故 $P(90<X<110)\ge1-0.1=0.9$。

- 中心极限定理 **(CLT)**：n 足够大时，$\sum X_i\approx N(n\mu,n\sigma^2)$，或 $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\approx N(0,1)$。

例题：某保险公司有 10000 个客户，每个客户一年内出险概率为 0.005。求一年内出险人数超过 60 人的概率。解答：$X\sim B(10000,0.005)$。$E(X)=50, D(X)=10000\times0.005\times0.995=49.75$。由 CLT，$X\approx N(50,49.75)$。$P(X>60)=1-P(X\le60)=1-\Phi\left(\frac{60-50}{\sqrt{49.75}}\right)\approx1-\Phi(1.42)\approx1-0.9222=0.0778$。

5.2 解题思路

1. 概率估计：题目给出期望和方差，求某个偏差的概率上限，用切比雪夫。
2. 近似计算：求大量独立同分布随机变量之和落在某区间的概率，先标准化再查表。

6 第六、七章：数理统计基础与参数估计

6.1 基础概念

- 总体与样本：总体是研究对象的全体，样本是从中抽取的子集。我们通过样本来“推测”总体的特征。
- 参数估计：总体分布的形式已知（如正态分布），但其中的参数（如 μ,σ）未知，需要用样本数据去“猜”这些参数的值。

- "帽子"符号 $\hat{\boldsymbol{\theta}}$（小白必看）：
  - θ：代表总体的 真实参数（通常是未知的，像上帝视角才知道的值）。
  - $\hat{\theta}$：读作"theta hat"或"theta 拔"，代表参数的 估计值。
  - 直观理解：给参数戴个帽子，就表示这是我们根据手里的样本数据"猜"出来的结果，而不是那个绝对真实的真值。

6.2 核心公式

- 常用统计量：
  - 样本均值 $\overline{X} = \frac{1}{n}\sum X_i$
  - 样本方差 $S^2 = \frac{1}{n-1}\sum \left(X_i - \overline{X}\right)^2$ （注意分母是 $n-1$）。
- 三大抽样分布：
  - χ^2 分布：n 个独立标准正态变量的平方和。
  - t 分布：用于小样本、方差未知时的均值检验。
  - F 分布：用于比较两个总体的方差。
- 矩估计：
  - 核心思想：用样本的平均水平去估计总体的平均水平。
  - 做法：令 $E(X) = \overline{X}$，解出参数。

> 例题：设总体 $X \sim U(0,\theta)$，求 θ 的矩估计量。 解答：$E(X) = \frac{\theta}{2}$。令 $\frac{\theta}{2} = \overline{X} \Rightarrow \hat{\theta} = 2\overline{X}$。

- 最大似然估计 **(MLE)**：
  - 核心思想：在参数 θ 的所有可能取值中，哪一个能让当前这组样本出现的概率最大，就选哪一个。

> 例题（最大似然估计）：设总体 X 服从指数分布，其概率密度为 $f(x;\lambda) = \lambda e^{-\lambda x}(x > 0)$，其中 $\lambda > 0$ 为未知参数。$x_1, x_2, ..., x_n$ 是样本观测值，求 λ 的最大似然估计量。
>
> 解答：
> 1. 写出似然函数：$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum x_i}$。
> 2. 取对数：$\ln L(\lambda) = n \ln \lambda - \lambda \sum x_i$。
> 3. 求导：$\frac{d}{d\lambda} \ln L(\lambda) = \frac{n}{\lambda} - \sum x_i$。
> 4. 令导数为 0：$\frac{n}{\lambda} - \sum x_i = 0 \Rightarrow \hat{\lambda} = \frac{n}{\sum x_i} = \frac{1}{\overline{x}}$。
>
> 故 λ 的最大似然估计量为 $\frac{1}{\overline{X}}$。

6.3 解题思路

1. 最大似然估计 **(MLE)** 必考大题步骤：
   - 第一步：写出似然函数 $L(\theta) = f(x_1;\theta)f(x_2;\theta)...f(x_n;\theta)$。
   - 第二步：取对数 $\ln L(\theta)$ （为了把乘法变加法，方便求导）。
   - 第三步：求导 $\frac{d}{d\theta} \ln L(\theta)$。
   - 第四步：令导数为 0，解出 θ。这个解 $\hat{\theta}$ 就是最大似然估计量。
2. 区间估计：
   - 题目会给一个置信水平（如 95%）。
   - 根据总体是否正态、方差是否已知，选择合适的统计量（Z 或 t）。
   - 查表找到临界值，代入公式计算范围。

> 例题：从正态总体 $N(\mu, 2^2)$ 中抽取容量为 16 的样本，均值 $\overline{x} = 10$。求 μ 的 95% 置信区间（$z_{\{0.025\}} = 1.96$）。 解答：方差已知，用 Z 统计量。区间为 $\left[\overline{x} - z_{\{\frac{\alpha}{2}\}}\frac{\sigma}{\sqrt{n}}, \overline{x} + z_{\{\frac{\alpha}{2}\}}\frac{\sigma}{\sqrt{n}}\right]$。 代入：$\left[10 - 1.96 \times \frac{2}{\sqrt{16}}, 10 + 1.96 \times \frac{2}{\sqrt{16}}\right] = [10 - 0.98, 10 + 0.98] = [9.02, 10.98]$。

祝考试顺利！