概率论期末复习要点

随机事件与概率

随机试验与事件：随机试验是结果具有不确定性但**所有可能结果可事先确定**的试验。例如抛一枚硬币，可能结果为"正面"或"反面"。试验的所有可能结果组成**样本空间** S。样本空间的任何子集称为一个**事件**。若试验结果属于某事件，则称该事件发生。常见事件运算包括：**并事件**（$A \cup B$，发生A或B）、**交事件**（$A \cap B$，同时发生A和B）、**补事件**（A^c，A未发生）等。事件间关系可通过文氏图直观表示，如互斥事件（不可能同时发生）等性质。

概率的定义及公理：概率 $P(A)$ 表示事件 A 发生的可能性大小，取值范围在$0,1$之间。概率满足如下公理化定义：1) 非负性：$P(A) \geq 0$；2) 规范性：$P(S) = 1$（必然事件概率为1，空事件概率为0）；3) 可列可加性：对于两两互斥事件 A_i，有 $P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$。由此推得一些基本性质：如对立事件 A 与 A^c 满足 $P(A^c) = 1 - P(A)$；两个事件的并满足 加法公式：$P(A \cup B) = P(A) + P(B) - P(A \cap B)$。在古典概型（有限等可能模型）中，若样本空间共有 N 个等可能基本结果，事件 A 包含 m 个基本结果，则 $P(A) = m/N$。实际计算中常借助计数方法（排列组合）求概率。

> **例题**：抛掷两颗骰子，求事件"A：出现至少一个6" 的概率，以及事件"B：两颗骰子点数之和为7"的概率。并验证加法公式 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$。
> **解答**：样本空间 S 包含36个等可能基本结果（6×6 种点数组合）。事件A包括出现至少一个6的结果，有 $6 + 6 - 1 = 11$ 个（计算：第一颗是6或第二颗是6，重复计数了两个都是6的情况一次），故 $P(A) = 11/36$。事件B包括和为7的结果，共有6个（(1,6),(2,5),…,(6,1)），所以 $P(B) = 6/36 = 1/6$。事件 $A \cap B$ 表示至少一个6且和为7，只可能是组合(1,6)和(6,1)，共有2个结果，故 $P(A \cap B) = 2/36$。于是 $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{11}{36} + \frac{6}{36} - \frac{2}{36} = \frac{15}{36} = 5/12$。计算结果验证了加法公式。注意：事件A和B并非互斥，因此直接相加需减去重复计数的交集概率。

条件概率、全概率公式与贝叶斯公式

条件概率：在事件 B 已经发生的前提下事件 A 发生的概率称为条件概率，记为 $P(A|B)$。根据定义，若 $P(B) > 0$，有公式：$P(A|B) = \frac{P(A \cap B)}{P(B)}$. 直观理解：$P(A|B)$ 等于在 B 发生的情形中 A 同时发生的比例。由条件概率定义可推出**乘法公式**：$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$ 。**独立事件**则刻画为：如果 $P(A \cap B) = P(A)P(B)$（等价于 $P(A|B) = P(A)$ 且 $P(B|A) = P(B)$），则称事件A和B独立。注意：独立不等同于互斥，除非其中一个事件概率为0（例如互斥且非

不可能的事件必不独立，因为 $P(A\cap B)=0$ 但 $P(A)P(B)>0$）。在多个事件情形，若任意子集的交集概率等于各自概率之积，则这些事件**相互独立**。

全概率公式：若事件 $A_1,A_2,\ldots,A_n$ 构成了样本空间的一个划分（两两互斥且并集为S），则对于任一事件 B，有$P(B)=\sum_{i=1}^{n}P(B|A_i)P(A_i)$,这就是全概率公式。它将事件 B 的概率分解为若干情形（划分事件 A_i）下 B 的发生概率之和。全概率公式的应用关键是选取适当的情形划分，使得各条件概率 $P(B|A_i)$ 易于获得。

贝叶斯公式：在已知 $P(A_i)$ 和 $P(B|A_i)$ 的情况下，用于逆向计算 $P(A_i|B)$ 的公式称为贝叶斯公式。由条件概率和全概率公式推导可得，对于划分事件 $A_1,\ldots,A_n$：$P(A_i|B)=\frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n}P(B|A_j)P(A_j)}$. 贝叶斯公式提供了根据先验概率 $P(A_i)$ 和**似然 $P(B|A_i)$** 计算事件 A_i 在事件 B 发生时的**后验概率**的方法。在决策、医学检测等场景中常应用此公式进行推断。

> **例题**：（贝叶斯公式应用）设有两个箱子，箱I中有红球3个、白球2个，箱II中有红球1个、白球4个。先等概率随机选取一只箱子，从中随机抽出1个球。已知抽出的球是红色，求被抽取的箱子为箱I的概率。
> **解答**：设事件 A：选中箱I，事件 B：抽到红球。先计算相关概率：$P(A)=P(\text{选箱I})=1/2$，$P(A^c)=1/2$。条件概率：$P(B|A)=P(红|箱I)=3/5$，$P(B|A^c)=P(红|箱II)=1/5$。利用全概率公式先求红球被抽到的概率：$P(B)=P(B|A)P(A)+P(B|A^c)P(A^c)=\frac{3}{5}\cdot\frac{1}{2}+\frac{1}{5}\cdot\frac{1}{2}=\frac{3+1}{10}=\frac{2}{5}$. 然后套用贝叶斯公式求所求概率：$P(A|B)=\frac{P(B|A)P(A)}{P(B)}=\frac{\frac{3}{5}\cdot\frac{1}{2}}{2/5}=\frac{3}{4}=0.75$. 结果表明：在抽到红球的前提下，有 75% 的概率选取的箱子是箱I。

随机变量及其分布

随机变量的概念：随机变量 X 将试验的结果映射为一个实数，用于定量表示随机现象的结果。随机变量可以是离散型（取值为有限或可数多个）或**连续型**（取值充满某区间）。随机变量 X 的**分布函数**（累计分布函数）定义为 $F_X(x)=P(X\le x)$，它满足：$F_X(x)$ 非减、右连续，且 $\lim_{x\to-\infty}F_X(x)=0$，$\lim_{x\to+\infty}F_X(x)=1$。任意两个实数 $a<b$，有 $P(a<X\le b)=F_X(b)-F_X(a)$。

离散型随机变量的分布

离散型随机变量 X 的**概率质量函数**（PMF）列出各可能取值 x_i 及其概率 $p_i=P(X=x_i)$，通常称为**分布列**。分布列需满足 $\sum_i p_i=1$。离散随机变量的分布函数 $F_X(x)$ 表现为分段常值的阶跃函数：在每个可能取值处概率使 $F_X(x)$ 跳跃。常见离散分布包括：

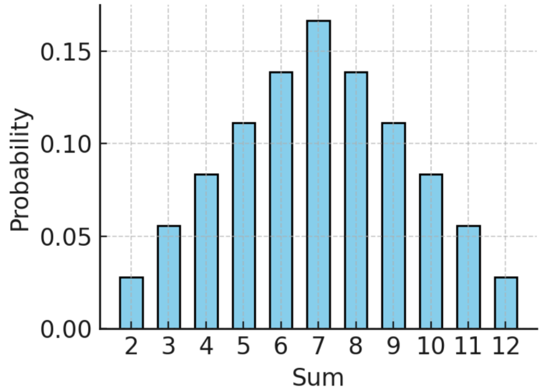
- **0-1分布（伯努利分布）**：表示一次成功概率为 p 的试验，取值{0,1}，$P(X=1)=p$，$P(X=0)=1-p$。**期望** $E(X)=p$，**方差** $Var(X)=p(1-p)$。

- **二项分布**：表示 n 次独立伯努利试验中成功次数，记作 $X \sim \mathrm{Bin}(n,p)$。概率公式：$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ k=0,1,\dots,n$。其期望 $E(X)=np$，方差 $Var(X)=np(1-p)$。
- **泊松分布**：表示单位时间内随机事件的发生次数，当事件发生率为 λ 且n大p小的极限情形可近似二项分布，记 $X \sim \mathrm{Poisson}(\lambda)$。概率质量：$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!},\ k=0,1,2,\dots$。期望 $E(X)=\lambda$，方差 $Var(X)=\lambda$。泊松分布常用于刻画**稀有事件**的计数。
- **几何分布**：表示伯努利试验中首次成功所需的试验次数。若记 X 为第一次成功所在的试验编号，则 $P(X=k)=(1-p)^{k-1}p,\ k=1,2,\dots$。期望 $E(X)=1/p$，方差 $Var(X)=(1-p)/p^2$。此分布的显著性质是**无记忆性**：对任意 $m,n\ge 1$，有 $P(X>m+n\mid X>m)=P(X>n)$。
- **超几何分布**：用于不放回抽样的情形。例如总共 N 件产品其中有 M 件不合格品，随机不放回抽取 n 件，设 X 为抽到不合格品的数量，则 X 服从超几何分布：$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$, 其中 $k=0,1,\dots,n$。期望 $E(X)=n\frac{M}{N}$，方差 $Var(X)=n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$。
- **负二项分布**：表示第 r 次成功出现时的试验总次数。它是几何分布的推广，若 X 表示获得第 r 个成功所需的伯努利试验次数，则 $P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r},\ k=r,r+1,\dots$。

> **例题**：抛掷两颗公平骰子，设随机变量 X 为两骰点数之和，求 X 的分布列并计算 $P(X\ge 9)$。
> **解答**：随机变量 X 所有可能取值为 2 至 12。利用等可能模型，可计算各点的概率：当和为 k 时，有 $k-1$ 种情况实现（例如和为5的情况：(1,4),(2,3),(3,2),(4,1)共4种），对于 $k=2,\dots,7$，$P(X=k)=\frac{k-1}{36}$；对称地，当 $k=7,\dots,12$，$P(X=k)=\frac{13-k}{36}$。具体列举如下：

> - $P(X=2)=\frac{1}{36}$
> - $P(X=3)=\frac{2}{36}$
> - $P(X=4)=\frac{3}{36}$
> - $P(X=5)=\frac{4}{36}$
> - $P(X=6)=\frac{5}{36}$
> - $P(X=7)=\frac{6}{36}$
> - $P(X=8)=\frac{5}{36}$
> - $P(X=9)=\frac{4}{36}$
> - $P(X=10)=\frac{3}{36}$
> - $P(X=11)=\frac{2}{36}$
> - $P(X=12)=\frac{1}{36}$。

> 据此分布列可计算 $P(X\ge 9)=P(X=9)+P(X=10)+P(X=11)+P(X=12)=\frac{4+3+2+1}{36}=\frac{10}{36}\approx 0.2778$。



> 上述分布列对应的概率质量分布如图所示。可以看出，X 的概率分布呈对称的**"三角形"**形状，峰值在7附近。考试中常见问题包括根据给定分布列计算概率，或验证分布列是否满足概率和为1等。

连续型随机变量的分布

连续型随机变量 Y 的分布由其**概率密度函数**（PDF） $f_Y(y)$ 描述。密度函数满足 $f_Y(y) \geq 0$ 且 $\int_{-\infty}^{+\infty} f_Y(y)\mathrm{d}y = 1$。对于任意区间 (a,b)，$P(a < Y \leq b) = \int_a^b f_Y(y)\mathrm{d}y$,而单点概率 $P(Y = c) = 0$（连续型随机变量取一个确值的概率为0）。由密度积分可得分布函数：$F_Y(x) = \int_{-\infty}^{x} f_Y(t)\mathrm{d}t$。常见连续分布包括：

- **均匀分布**：在区间 $[a,b]$ 上密度均匀为 $f(x) = \frac{1}{b-a}$（对于 $x \in [a,b]$，区间外为0）。记作 $U(a,b)$。期望 $E(X) = \frac{a+b}{2}$，方差 $Var(X) = \frac{(b-a)^2}{12}$。均匀分布是**几何概型**的连续对应。
- **指数分布**：密度 $f(x) = \lambda e^{-\lambda x}$，$x \geq 0$，参数 $\lambda > 0$。记作 $X \sim \mathrm{Exp}(\lambda)$。期望 $E(X) = 1/\lambda$，方差 $Var(X) = 1/\lambda^2$。指数分布具有无记忆性：$P(X > m+n \mid X > m) = P(X > n)$。它常用于描述随机寿命或等待时间。
- **正态分布**：又称高斯分布，密度函数$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)$，$-\infty < x < +\infty$,由均值 μ 和方差 σ^2 参数化，记作 $N(\mu,\sigma^2)$。正态分布曲线呈对称的"钟形"，平均值处概率密度最高。它的期望为 μ，方差为 σ^2。正态分布在自然和社会现象中非常常见，也是中心极限定理的极限分布。标准正态分布指 $\mu = 0, \sigma = 1$ 的情形，用 $\Phi(x)$ 表示其分布函数。
- **Gamma分布**：广义的指数族分布，形状参数 α、尺度参数 β。当 $\alpha = n$ 为正整数时，相当于 n 个独立指数分布之和（Erlang分布）。特别地，$\alpha = \frac{n}{2}$，$\beta = \frac{1}{2}$ 时，Gamma分布即**卡方分布**，其自由度为 n。
- **Beta分布**：定义在 $[0,1]$ 区间上的双参数分布，常用密度形式 $f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$（$B$ 是Beta函数）。当 $\alpha = \beta = 1$ 时为均匀分布，在贝叶斯统计中常作为概率的先验分布。

> **例题**：设连续随机变量 Y 的密度函数为
> $f_Y(y) = \begin{cases} 2y, & 0 \leq y \leq 1; \\ 0, & \text{其他}. \end{cases}$ 求：① 分布函数 $F_Y(x)$；② $P(0.2 < Y < 0.8)$。
> **解答**：首先验证 $f_Y(y)$ 的积分为1：$\int_0^1 2y\,dy = [y^2]_0^1 = 1$，符合密度性质。接着求分布函数：

> 当 $x<0$ 时 $F_Y(x)=0$；当 $0\le x\le 1$ 时，$F_Y(x)=\int_0^x 2y\,dy=x^2$；当 $x>1$ 时 $F_Y(x)=1$。因此$F_Y(x)=\begin{cases}0, & x<0;\\ x^2, & 0\le x\le 1;\\ 1, & x>1\,.\end{cases}$然后计算概率：$P(0.2<Y<0.8)=F_Y(0.8)-F_Y(0.2)=0.8^2-0.2^2=0.64-0.04=0.60$。即在该分布下随机变量有60%的概率取值落在0.2与0.8之间。

多维随机变量与联合分布

联合分布：对于两个或多个随机变量，可以定义在同一概率试验上的多维随机变量。例如二维随机变量 (X,Y) 表示两个分量 X 和 Y。它们的概率分布由**联合分布**给出。离散情形下，用**联合概率质量函数** $p_{X,Y}(x,y)=P(X=x,\ Y=y)$ 列出各组合情况的概率，并满足 $\sum_x\sum_y p_{X,Y}(x,y)=1$。连续情形下，用**联合密度函数** $f_{X,Y}(x,y)$（非负且双重积分为1）描述，其分布函数 $F_{X,Y}(a,b)=P(X\le a,\ Y\le b)$。边缘分布可由联合分布得到：离散情况下 $P(X=x)=\sum_y p_{X,Y}(x,y)$，$P(Y=y)=\sum_x p_{X,Y}(x,y)$；连续情况下 $f_X(x)=\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}y$（对 Y 积分），$f_Y(y)=\int f_{X,Y}(x,y)\mathrm{d}x$。**条件分布**则刻画在给定一个变量取值时另一个变量的分布：如离散情况下 $P(Y=y\,|\,X=x)=\frac{P(X=x,\ Y=y)}{P(X=x)}$。随机变量 X 和 Y **独立**当且仅当对于所有 x,y，$P(X=x,\ Y=y)=P(X=x)P(Y=y)$（连续时则 $f_{X,Y}(x,y)=f_X(x)f_Y(y)$）。独立性意味着任意关于 X 的事件和关于 Y 的事件也独立。

> **例题**：给定随机变量 X,Y 的联合分布（离散）如下表所示：
>
> | (X,Y) | $Y=0$ | $Y=1$ |
> |---|---|---|
> | $X=0$ | 0.3 | 0.2 |
> | $X=1$ | 0.2 | 0.3 |
>
> （表中各单元格给出 $P(X=i,\ Y=j)$）。试求：①边缘分布 $P(X=i),P(Y=j)$；②判断 X 与 Y 是否独立；③计算 $P(Y=0\,|\,X=0)$。
>
> **解答**：从表中直接求出边缘概率：
>
> - $P(X=0)=0.3+0.2=0.5,\quad P(X=1)=0.5$。
> - $P(Y=0)=0.3+0.2=0.5,\quad P(Y=1)=0.5$。
>
> 检查独立性：如果 X,Y 独立，则应有 $P(X=0,\ Y=0)=P(X=0)P(Y=0)=0.25$。但表中 $P(X=0,Y=0)=0.3\neq 0.25$。因此 $P(X=i,\ Y=j)$ 不等于 $P(X=i)P(Y=j)$，故 X 与 Y **不独立**。
>
> 条件概率计算：在 $X=0$ 发生的条件下，$Y=0$ 的条件概率为

> $$P(Y=0\mid X=0)=\frac{P(X=0,\ Y=0)}{P(X=0)}=\frac{0.3}{0.5}=0.6,$$
>
> 同理 $P(Y=1\mid X=0)=0.4$。这说明在 $X=0$ 时，Y 取0的可能性为60%。通过以上计算，我们掌握了联合分布转化为边缘分布、判断独立性以及求条件分布的方法。

数字特征（期望、方差、协方差、相关系数）

数学期望：数学期望（均值）反映随机变量取值的平均水平。定义为：离散型 X 的期望 $E(X)=\sum_x x\,P(X=x)$；连续型 X 的期望 $E(X)=\int_{-\infty}^{+\infty}x f_X(x)\mathrm{d}x$。期望具有**线性性质**：对任意常数 a,b，有 $E(aX+b)=aE(X)+b$。更一般地，任意函数 $g(X)$ 的期望 $E[g(X)]$ 等于离散情况下 $\sum_x g(x)P(X=x)$，连续情况下 $\int g(x)f_X(x)dx$。特别地，**若 X,Y 相互独立，则** $E(XY)=E(X)E(Y)$；即独立时乘积的期望等于期望的乘积（不独立时一般 $E(XY)\neq E(X)E(Y)$，偏差由协方差刻画）。多维随机变量的期望称为**向量的均值**（对每个分量求期望）。

方差与标准差：方差表示变量围绕均值波动的程度。定义 $Var(X)=E[(X-E(X))^2]$。计算公式为 $Var(X)=E(X^2)-[E(X)]^2$。离散型计算 $E(X^2)=\sum x^2P(X=x)$，连续型 $E(X^2)=\int x^2 f_X(x)dx$。标准差 $\sigma_X=\sqrt{Var(X)}$则是方差的正平方根，与原单位相同。方差的性质：常数的方差为0；对于常数 a,b，$Var(aX+b)=a^2Var(X)$。独立随机变量的方差具有可加性：若 X 与 Y 独立，则 $Var(X+Y)=Var(X)+Var(Y)$。

协方差与相关系数：协方差描述两个随机变量线性相关方向及程度。
定义 $Cov(X,Y)=E[(X-E(X))(Y-E(Y))]=E(XY)-E(X)E(Y)$。若 $Cov(X,Y)>0$，则 X,Y 正相关（一方偏大时另一方倾向偏大）；$Cov<0$ 则负相关；$Cov=0$ 则称**不相关**。注意，协方差为0并不一定表示独立（独立必然协方差0，但反之未必，除非在正态分布等特殊情形）。**相关系数** ρ_{XY} 是协方差的标准化量，定义为$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$. 相关系数取值范围在 $[-1,1]$。$\rho=1$或 -1 表示完全线性相关，此时 Y 几乎是 X 的线性函数；$\rho=0$ 则线性不相关。相关系数的绝对值越接近1，线性相关程度越强。

> **例题**：延续前述联合分布表格的例子，计算该分布下随机变量 X 和 Y 的期望、方差、协方差及相关系数。
> **解答**：根据先前计算，已知边缘分布 $P(X=1)=0.5, P(Y=1)=0.5$。首先求期望：
>
> 1. $E(X)=0\cdot P(X=0)+1\cdot P(X=1)=0\cdot 0.5+1\cdot 0.5=0.5$。
> 2. $E(Y)=0\cdot 0.5+1\cdot 0.5=0.5$。
>   然后计算二阶矩 $E(X^2),E(Y^2)$ 以及积的期望 $E(XY)$：
> 3. $E(X^2)=0^2\cdot 0.5+1^2\cdot 0.5=0.5;\quad E(Y^2)=0.5;$
> 4. $E(XY)=\sum_{x,y}xy\,P(X=x,Y=y)=0\cdot 0\cdot 0.3+0\cdot 1\cdot 0.2+1\cdot 0\cdot 0.2+1\cdot 1\cdot 0.3=0.3.$（仅 $x=1,y=1$ 项贡献非零积 $1\cdot 1\cdot 0.3$）。

> 由此得到方差和协方差：
> 5. $Var(X) = E(X^2) - [E(X)]^2 = 0.5 - 0.5^2 = 0.5 - 0.25 = 0.25; \quad Var(Y) = 0.25.$
> 6. $Cov(X,Y) = E(XY) - E(X)E(Y) = 0.3 - 0.5 \cdot 0.5 = 0.3 - 0.25 = 0.05.$
> 最后计算相关系数：
> 7. $\rho_{XY} = \dfrac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \dfrac{0.05}{\sqrt{0.25 \times 0.25}} = \dfrac{0.05}{0.25} = 0.2.$
> 结果表明 $E(X) = E(Y) = 0.5$，方差均为0.25，协方差0.05为正，表示 X, Y 存在轻微正相关性，对应的相关系数约为0.2。

大数定律与中心极限定理

大数定律

大数定律：揭示了大量重复试验的平均结果会稳定在理论概率或期望值附近。通俗地说，在随机试验中每次结果可能不同，但随着试验次数趋于无穷，多次试验结果的平均值几乎总接近于某个确定值（这一确定值通常就是期望或事件概率）。大数定律有不同形式：**弱大数定律**描述的是样本平均值依概率收敛于期望，即对于任意 $\varepsilon > 0$，$P\Big(|\overline{X}_n - \mu| < \varepsilon\Big) \to 1 \quad (n \to \infty)$,其中 $\overline{X} * n = \frac{1}{n} \sum *i = 1^n X_i$，$\mu = E(X_i)$。等价地，$P(|\overline{X}_n - \mu| \ge \varepsilon) \to 0$。弱大数定律通常要求 $Var(X_i)$ 有限，可利用切比雪夫不等式证明：由于 $E(\overline{X}_n) = \mu$，$Var(\overline{X}_n) = \frac{Var(X_i)}{n}$，由切比雪夫公式得到 $P\Big(|\overline{X}_n - \mu| \ge \varepsilon\Big) \le \frac{Var(X_i)/n}{\varepsilon^2} = \frac{Var(X_i)}{n\varepsilon^2}$,随 $n\to\infty$ 收敛至0，即证得 $\overline{X}_n$ 依概率收敛到 μ。**强大数定律**则进一步保证 $\overline{X}_n \to \mu$ 几乎必然收敛（概率为1的收敛）。总的来说，大数定律严格阐明了"频率的稳定性"：当试验次数很大时，频率将近似等于理论概率。

> **例题**：一枚硬币正面朝上的概率 $p = 0.5$。问：抛硬币多少次能够使正面出现的频率落在 0.5 ± 0.1 范围内的概率至少达到 95%？
> **解答**：这是应用大数定律的实际估计问题。设事件 A 为"正面频率偏离0.5超过0.1"。根据切比雪夫不等式：$P(|\overline{X}_n - 0.5| \ge 0.1) \le \frac{p(1-p)/n}{0.1^2} = \frac{0.25}{0.01\,n} = \frac{25}{n}$. 要使 $P(A) \le 0.05$（频率偏差超过0.1的概率不超过5%），只需令 $\frac{25}{n} \le 0.05$，解得 $n \ge 500$。也就是说，大约进行500次抛硬币实验，可以**保证**（以切比雪夫意义）频率落在 $[0.4, 0.6]$ 区间内的概率至少为0.95。这个结果与直观经验吻合：试验次数越多，频率越稳定地逼近理论概率0.5。

中心极限定理

中心极限定理（Central Limit Theorem, CLT）是概率论中最重要的极限定理之一。它揭示了**随机现象中的正态趋向**：在一定条件下，大量相互独立同分布的随机变量的平均值（或总和）经过适当标准化后，其分布逼近于标准正态分布。具体而言，设 $X_1, X_2, \ldots, X_n$ 独立同分布，具有期望 μ、方差 σ^2（均

有限）。令 $S_n = X_1 + \cdots + X_n$ 为总和，$\overline{X}_n = S_n/n$ 为样本均值，则当 n 很大时，$\dfrac{S_n - n\mu}{\sigma\sqrt{n}}$ 或 $\dfrac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$ 的分布近似 $N(0,1)$。用分布函数表述：对任意实数 x，$P\Big(\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}} \le x\Big) \xrightarrow[n\to\infty]{} \Phi(x)$,其中 $\Phi(x)$ 为标准正态分布的分布函数。中心极限定理解释了为何在大量独立随机因素作用下总体行为呈正态分布，这也是正态分布在统计推断中普适适用的理论依据。

中心极限定理在实际中的一个重要应用是**正态近似**：对于具有确定期望和方差的大样本分布，我们可以用正态分布来近似计算其概率。例如大数定律告诉我们二项分布 $\mathrm{Bin}(n,p)$ 的平均约为 np，而中心极限定理进一步说明其**形状**在 n 大时接近正态曲线。当 n 较大且 p 不极端时，可近似 $X \sim \mathrm{Bin}(n,p)$ 为正态 $N(np,;np(1-p))$。

> **例题**：某产品生产过程中次品率约为 5%。现随机抽检 100 件产品，设 X 表示其中次品的个数。试近似计算 $P(X \le 10)$ 和 $P(4 \le X \le 10)$。
> **解答**：因为每件产品次品的概率 $p = 0.05$，试验次数 $n = 100$ 独立，X 服从二项分布 $\mathrm{Bin}(100, 0.05)$。其期望 $E(X) = np = 5$，方差 $Var(X) = np(1-p) = 4.75$。由于 n 较大，可用正态近似：$X \approx N(5,\ 4.75)$.近似计算时，为提高准确性，可施加**连续性校正**：将离散事件 $X \le 10$ 转换为连续变量事件 $X < 10.5$。于是
>
> - $P(X \le 10) \approx P\Big(Z < \frac{10.5-5}{\sqrt{4.75}}\Big) = P(Z < 2.53) \approx 0.9943$。
> - 类似地，$P(4 \le X \le 10) \approx P(3.5 < X < 10.5) = P\Big(\frac{3.5-5}{\sqrt{4.75}} < Z < \frac{10.5-5}{\sqrt{4.75}}\Big) = P(-0.69 < Z < 2.53) = \Phi(2.53) - \Phi(-0.69)$.查标准正态分布表：$\Phi(2.53) \approx 0.9943$，$\Phi(-0.69) = 1 - \Phi(0.69) \approx 1 - 0.7549 = 0.2451$。相减得 $0.9943 - 0.2451 = 0.7492$。即约有 74.92% 的概率次品数在4到10之间。
>   通过中心极限定理的正态近似，我们快捷地估计了二项分布的区间概率。在期末考试中，常见题型包括让学生使用中心极限定理对二项分布或泊松分布进行近似计算，并说明步骤。建议熟悉标准正态分布表的查找方法，以及添加连续性校正处理离散到连续近似的技巧。