概率论 第5讲 · 大数定律与中心极限定理 · 解析

> 对应 02-试题。每题给「**答案 / 思路 / 易错点 / 真题骨架**」。

核心公式（贴在眼前）

$$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2},\qquad \frac{\sum X_i-n\mu}{\sqrt{n}\,\sigma}\approx N(0,1),\qquad Y_n\sim B(n,p)\Rightarrow Y_n\approx N\big(np,\,np(1-p)\big).$$

A 组 · 基础

第 1 题

答案：（Ⅰ）见证明；（Ⅱ）$EX_n=1+\dfrac{1}{n}-\dfrac{1}{n^2}\to 1\neq 0$，**依概率收敛 ⇏ 期望收敛。**

思路：

（Ⅰ）对任意 $\varepsilon>0$，取 $n>\dfrac{1}{\varepsilon}$，则 $\dfrac{1}{n}<\varepsilon$。此时 $|X_n|\geq\varepsilon$ **只可能在** $X_n=n$ 时发生（$|X_n|=n$），故

$$P\{|X_n-0|\geq\varepsilon\}=P\{X_n=n\}=\frac{1}{n}\longrightarrow 0.$$

所以 $X_n\xrightarrow{P}0$。

（Ⅱ）

$$EX_n=\frac{1}{n}\left(1-\frac{1}{n}\right)+n\cdot\frac{1}{n}=1+\frac{1}{n}-\frac{1}{n^2}\xrightarrow{n\to\infty}1\neq 0.$$

结论：**依概率收敛不能推出期望收敛。**（一般地，需要"一致可积"或"有界"等额外条件。）

易错点：不要因为 $P\{X_n=n\}=\frac{1}{n}$ 很小就忽略它——**期望是"值 × 概率"**，$n\cdot\frac{1}{n}=1$ 恰好贡献了一个常数，这正是反例的精妙处。

第 2 题

答案：（Ⅰ）$\leq\dfrac{1}{8}$；（Ⅱ）$\geq\dfrac{7}{8}$。

思路：$\varepsilon=4$，

$$P\{|X-5|\geq 4\}\leq\frac{DX}{\varepsilon^2}=\frac{2}{16}=\frac{1}{8},\qquad P\{|X-5|<4\}=1-P\{|X-5|\geq 4\}\geq 1-\frac{1}{8}=\frac{7}{8}.$$

易错点：**方向**——"$\geq$"用上界 $\frac{DX}{\varepsilon^2}$；"$<$"用 $1-\frac{DX}{\varepsilon^2}$（对立事件）。

第 3 题

答案：≥ 0.9。

思路：$EX = np = 500$，$DX = np(1-p) = 1000 \times 0.25 = 250$。

$450 < X < 550 \iff |X - 500| < 50$，取 $\varepsilon = 50$：

$$P\{|X - 500| \geq 50\} \leq \frac{250}{50^2} = 0.1 \implies P\{450 < X < 550\} \geq 1 - 0.1 = 0.9.$$

易错点：二项方差是 $np(1-p) = 250$，**不是** $np = 500$。

第 4 题

答案：$\bar{X} \xrightarrow{P} 3.5$。

思路：各次投掷**相互独立**；X_i **同分布**（都在 $\{1, 2, \cdots, 6\}$ 上等可能）；期望

$$EX_i = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$$

存在。三点齐备，由**辛钦大数定律**得 $\bar{X} \xrightarrow{P} EX_i = 3.5$。

易错点：骰子的期望是 3.5（**不是** 3 或 4）；本题**用不着方差**，所以只能用辛钦（不是切比雪夫）。

第 5 题

答案：**C。**

思路：辛钦大数定律的三个条件——**相互独立**、**同分布**、**期望存在。**题干已给"相互独立"，只需再满足"同分布 + 期望存在"。

- （A）只说"期望相同"，**不保证同分布**（如一个取 ± 1、一个取 ± 2），✗；
- （B）（D）只说"同一**类型**分布"（离散型/连续型 ≠ 同一个分布），且**不保证期望存在**（如柯西分布连续但期望不存在），✗；
- （C）**同一泊松分布**：既同分布，期望 $= \lambda$ 又存在，✓。

易错点：「同一**类型**分布」≠「**同一个**分布」——这是本题最核心的文字陷阱。

第 6 题

答案：$\approx \Phi(1)$。

思路：$E(\sum X_i) = 100 \times 1 = 100$，$D(\sum X_i) = 100 \times 4 = 400$，标准差 $= \sqrt{400} = 20$。由中心极限定理

$$P\left\{\sum_{i=1}^{100} X_i \leq 120\right\} \approx \Phi\left(\frac{120 - 100}{20}\right) = \Phi(1).$$

易错点：**标准差**是 $\sqrt{n\sigma^2} = 20$（不是 $n\sigma^2 = 400$，也不是 $\sigma^2 = 4$）——标准化时除的是**标准差。**

B 组 • 中等

第 7 题 【真题骨架 2022•9】

答案：（Ⅰ）$EY_i=1$，$DY_i=4$；（Ⅱ）$\le \dfrac{4}{n\varepsilon^2}$。

思路：

（Ⅰ）$EY_i=EX_i^2=1$；

$$DY_i=EY_i^2-(EY_i)^2=EX_i^4-1^2=5-1=4.$$

（Ⅱ）$\bar{Y}=\dfrac{1}{n}\sum Y_i=\dfrac{1}{n}\sum X_i^2$，因 X_i 独立同分布 $\Longrightarrow$ Y_i 独立同分布：

$$E\bar{Y}=1,\qquad D\bar{Y}=\frac{1}{n^2}\sum_{i=1}^{n}DY_i=\frac{n\cdot 4}{n^2}=\frac{4}{n}.$$

由切比雪夫不等式

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i^2-1\right|\ge\varepsilon\right\}\le\frac{D\bar{Y}}{\varepsilon^2}=\frac{4}{n\varepsilon^2}.$$

易错点：

1. $D\bar{Y}$ 的 $\dfrac{1}{n^2}$ **不能丢**，否则会错答成 $\dfrac{4}{\varepsilon^2}$；
2. $D(X_i^2)=EX_i^4-(EX_i^2)^2$，用到**四阶矩** $\mu_4=5$。

真题：2022·9（单选 5 分）考的正是此结构（$DX_i^2=\mu_4-\mu_2^2$，界为 $\dfrac{\mu_4-\mu_2^2}{n\varepsilon^2}$）。

第 8 题

答案：$n\ge 90$。

思路：$E\bar{X}=20$，$D\bar{X}=\dfrac{DX}{n}=\dfrac{9}{n}$。要求"落在附近"，取下界：

$$P\{|\bar{X}-20|<1\}\ge 1-\frac{9/n}{1^2}=1-\frac{9}{n}\ge 0.9\Longrightarrow\frac{9}{n}\le 0.1\Longrightarrow n\ge 90.$$

易错点：$\bar{X}$ 的方差是 $\dfrac{\sigma^2}{n}=\dfrac{9}{n}$（**不是** 9）；"至少"故**向上取整**（本题恰为整数 90）。

第 9 题

答案：最多装 **99** 件。

思路：设第 i 件重量 X_i，$EX_i=10$，$DX_i=0.25$。记 $T_n=\sum_{i=1}^{n}X_i$。由中心极限定理

$$ET_n=10n,\qquad DT_n=0.25n,\qquad \sqrt{DT_n}=0.5\sqrt{n},$$

$$P\{T_n \le 1000\} \approx \Phi\left(\frac{1000-10n}{0.5\sqrt{n}}\right) > 0.977 = \Phi(2)$$

$$\Longrightarrow \frac{1000-10n}{0.5\sqrt{n}} \ge 2 \Longrightarrow 1000-10n \ge \sqrt{n}.$$

解不等式：令 $t=\sqrt{n}$，得 $10t^2+t-1000\le 0$，$t \le \dfrac{-1+\sqrt{40001}}{20} \approx 9.95$，故 $n \le 99.0$，即**最多 99 件**。

验证：$n=99$：$\dfrac{1000-990}{0.5\sqrt{99}} \approx \dfrac{10}{4.975} \approx 2.01 > 2$ ✓；$n=100$：$\dfrac{0}{5}=0$ ✗。

易错点：

- 标准差是 $\sqrt{n}\cdot 0.5$，**不是** $n\cdot 0.5$；方差是 $n\cdot 0.25$；
- $\Phi(2)=0.977$ 给出的是**关于 z 的单调性**：$\Phi(z)>0.977 \iff z>2$，方向别反；
- 答案是"最多"故**向下取整**（99，不是 100）。

第 10 题 【真题骨架 2020•8】

答案：$\approx \Phi(1)$。

思路：$\sum X_i \sim B(400,0.5)$，$E=400\times 0.5=200$，$D=400\times 0.5\times 0.5=100$，标准差 $=10$。由棣莫弗–拉普拉斯定理

$$P\left\{\sum_{i=1}^{400} X_i \le 210\right\} \approx \Phi\left(\frac{210-200}{10}\right)=\Phi(1).$$

易错点：二项的正态近似用**二项自己的期望方差**（np、$np(1-p)$），不要另算。

真题：2020·8（单选 4 分）结构完全一致（$n=100$、阈值 55，答案 $\Phi(1)$）。

第 11 题 【真题骨架 2025•9】

答案：$\approx 4e^{-3} \approx 0.199$。

思路：$T\sim B(300,0.01)$。$n=300$ 大、$p=0.01$ 很小，用**泊松近似**，$\lambda=np=3$，$T\approx P(3)$：

$$P\{T\le 1\} \approx P\{T=0\}+P\{T=1\} = e^{-3}\cdot\frac{3^0}{0!}+e^{-3}\cdot\frac{3^1}{1!} = e^{-3}(1+3)=4e^{-3}.$$

易错点：$\lambda=np=300\times 0.01=3$（**不是** $n=300$，也**不是** $p=0.01$）；$P\{T\le 1\}$ 是"0 或 1"，别只算一项。

真题：2025·9（单选 5 分）同结构（$n=20,p=0.1,\lambda=2$，答案 $3e^{-2}$，即选项 C）。

第 12 题

答案：$\leq \dfrac{1}{16}$。

思路：令 $Z = X - Y$。由 $EX = EY$ 得 $EZ = 0$。

$$\mathrm{Cov}(X, Y) = \rho_{XY}\sqrt{DX}\sqrt{DY} = \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2},$$

$$DZ = D(X - Y) = DX + DY - 2\,\mathrm{Cov}(X, Y) = 1 + 1 - 2 \times \frac{1}{2} = 1.$$

取 $\varepsilon = 4$：

$$P\{|X - Y| \geq 4\} = P\{|Z - EZ| \geq 4\} \leq \frac{DZ}{16} = \frac{1}{16}.$$

易错点：

- 方差**不能**直接相加（$DX + DY = 2$），必须**减去** $2\,\mathrm{Cov}$；
- $\mathrm{Cov} = \rho\sqrt{DX}\sqrt{DY}$，分母是**标准差之积**。

C 组 • 拔高

第 13 题

答案：**D**。

思路：$X_n \sim E(n) \Longrightarrow EX_n = \dfrac{1}{n}$（存在），$DX_n = \dfrac{1}{n^2}$。切比雪夫大数定律只差"**方差一致有界**"这一条（独立性题干已给）。逐项算方差：

选项	序列	$D(\cdot)$	是否有界
A	$\dfrac{1}{n}X_n$	$\dfrac{1}{n^2} \cdot \dfrac{1}{n^2} = \dfrac{1}{n^4}$	✓ ≤ 1
B	X_n	$\dfrac{1}{n^2}$	✓ ≤ 1
C	nX_n	$n^2 \cdot \dfrac{1}{n^2} = 1$	✓ $= 1$
D	n^2X_n	$n^4 \cdot \dfrac{1}{n^2} = n^2$	✗ $\to \infty$

故 D 不服从切比雪夫大数定律。

易错点：$D(aX_n) = a^2DX_n$——**系数要平方**；本题 D 项系数 n^2 平方后是 n^4，恰好把 $DX_n = \frac{1}{n^2}$ 抵消成 n^2（发散）。这是本类题的标准设计手法。

第 14 题 【切比雪夫大数定律】

证明：记 $\bar{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$。由期望的线性性 $E\bar{X}=\dfrac{1}{n}\sum EX_i$。由**两两不相关**（$\mathrm{Cov}(X_i,X_j)=0,\ i\neq j$）：

$$D\bar{X}=\frac{1}{n^2}D\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\left[\sum_{i=1}^{n}DX_i+2\sum_{i<j}\mathrm{Cov}(X_i,X_j)\right]=\frac{1}{n^2}\sum_{i=1}^{n}DX_i\le\frac{nC}{n^2}=\frac{C}{n}.$$

对 $\bar{X}$ 用切比雪夫不等式：

$$P\left\{\left|\frac{1}{n}\sum X_i-\frac{1}{n}\sum EX_i\right|\ge\varepsilon\right\}=P\{|\bar{X}-E\bar{X}|\ge\varepsilon\}\le\frac{D\bar{X}}{\varepsilon^2}\le\frac{C}{n\varepsilon^2}\xrightarrow{n\to\infty}0.\qquad\blacksquare$$

要点：

1. **「两两不相关」是用来消协方差项的**——没有它，$\sum DX_i$ 这一步不成立；
2. 「方差有界」保证 $\dfrac{C}{n}\to0$，这就是"**平均值稳定**"的来源；
3. 结论即 $\bar{X}\xrightarrow{P}\overline{EX}$，是**大数定律的原始形式**；伯努利、辛钦都是它的特例或并列版本。

易错点：把"不相关"误当"独立"——不相关只要求 $\mathrm{Cov}=0$（比独立弱），但**已足够**。

第 15 题

答案：$n_{\min}=102$。

思路：由中心极限定理

$$\sum_{i=1}^{n}X_i\overset{\text{近似}}{\sim}N(100n,\,100n),\qquad\sqrt{D}=10\sqrt{n}.$$

要求

$$P\left\{\sum_{i=1}^{n}X_i\ge10000\right\}\approx1-\Phi\left(\frac{10000-100n}{10\sqrt{n}}\right)\ge0.95.$$

即 $\Phi\left(\dfrac{10000-100n}{10\sqrt{n}}\right)\le0.05=\Phi(-1.645)$，故

$$\frac{10000-100n}{10\sqrt{n}}\le-1.645\implies10000-100n\le-16.45\sqrt{n}\implies100n-16.45\sqrt{n}-10000\ge0.$$

令 $t=\sqrt{n}$：$100t^2-16.45t-10000\ge0$，解得

$$t\ge\frac{16.45+\sqrt{16.45^2+4\,000\,000}}{200}\approx10.082\implies n\ge101.6.$$

故 $n_{\min}=102$。

验证：$n=101$：$z=\dfrac{10000-10100}{10\sqrt{101}}\approx -0.995>-1.645$ ✗；$n=102$：$z\approx -1.98\le -1.645$ ✓。

易错点：

- 是"**和 ≥ 10000**"，故用**右尾** $1-\Phi(\cdot)$，**不要**错写成 Φ；
- $\Phi(z)\le 0.05 \iff z\le -1.645$（**负号**易漏）；
- 标准差 $\sqrt{n\sigma^2}=\sigma\sqrt{n}$，**不是** σn。

第 16 题 【综合辨析】

答案：

（Ⅰ）$D\bar{X}=\dfrac{1}{n^2}\sum DX_i=\dfrac{\sigma^2}{n}$，由切比雪夫不等式

$$P\{|\bar{X}-\mu|\ge\varepsilon\}\le\frac{D\bar{X}}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}\xrightarrow{n\to\infty}0,$$

即 $\bar{X}\xrightarrow{P}\mu$。（这是**辛钦大数定律**的一个特例——多了"方差存在"的条件。）

（Ⅱ）由中心极限定理 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\overset{\text{近似}}{\sim}N(0,1)$，故

$$P\{|\bar{X}-\mu|<\varepsilon\}=P\left\{\left|\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right|<\frac{\varepsilon\sqrt{n}}{\sigma}\right\}\approx 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)-1.$$

当 $n\to\infty$ 时 $\dfrac{\varepsilon\sqrt{n}}{\sigma}\to+\infty$，$\Phi\to 1$，故该概率 $\to 1$。

（Ⅲ）**中心极限定理更精确**：

- **（Ⅱ）利用"和的分布趋正态"这一更深**的信息，给出的是**近似值**（可以很接近真值），但要求 n **充分大**、且 σ^2 **有限存在**；
- **（Ⅰ）只用到期望与方差，对任意 n 都能用、无条件普适，但只能给宽松的上界**（界 $\dfrac{\sigma^2}{n\varepsilon^2}$ 往往远大于真值），**给不出近似值。**

> **一句话**：**切比雪夫"普适但粗糙"，中心极限定理"精确但要求高"。** 前者证明"收敛"，后者刻画"收敛的速度与形态"。

易错点：Φ 的变量是 $\dfrac{\varepsilon\sqrt{n}}{\sigma}$（**除以标准差** σ，不是 σ^2）。

命题质量报告

一、考点覆盖

考点	题号	覆盖
依概率收敛的定义与反例	1	✅
切比雪夫不等式：直接套 + 方向	2	✅
切比雪夫不等式：二项分布	3	✅
辛钦大数定律：应用（求极限值）	4	✅
辛钦大数定律：**条件辨析**	5	✅
中心极限定理：iid 求和标准化	6	✅
2022•9 骨架（样本均值 + 四阶矩）	7	✅
反解样本量 n（切比雪夫）	8	✅
中心极限定理：应用建模（装箱）	9	✅
2020•8 骨架（0–1 变量求和）	10	✅
2025•9 骨架（泊松近似）	11	✅
切比雪夫 + 协方差（$Z = X - Y$）	12	✅
切比雪夫大数定律：**条件辨析**	13	✅
切比雪夫大数定律：**证明**	14	✅
中心极限定理：反解 n	15	✅
切比雪夫 vs 中心极限定理：精度/条件对比	16	✅

二、难度与用时

组	题量	难度	建议用时
A 基础	6（1–6）	★☆☆	22 分钟
B 中等	6（7–12）	★★☆	28 分钟
C 拔高	4（13–16）	★★★	20 分钟
合计	**16**		**70 分钟**

三、真题使用情况

- **骨架题**：

  - 第 7 题 ← **2022•9**（单选 5 分，「大数定律（切比雪夫不等式）」），数字原创；
  - 第 10 题 ← **2020•8**（单选 4 分，「中心极限定理」），数字原创；
  - 第 11 题 ← **2025•9**（单选 5 分，「泊松近似（二项分布）」），数字原创。
- **仓库例题改编**：第 4、5、9、13 题（`knowledge/4`、`exercise/4`、`knowledge/3 §切比雪夫`）。
- **本章十年 3 题全覆盖**（2020·8、2022·9、2025·9）。

四、仍未覆盖方向（诚实列出）

1. **依概率收敛的运算性质**（如 $X_n \xrightarrow{P} a,\ Y_n \xrightarrow{P} b \Rightarrow X_nY_n \xrightarrow{P} ab$）——考研几乎不单独考；
2. **李雅普诺夫中心极限定理**（不要求同分布，只要求某阶矩有界）——边缘内容，本讲未涉及；
3. **用切比雪夫不等式讨论估计量的"一致性"**——会在第 21 章（参数估计）中用到，届时回看 §2 即可；
4. **大数定律的建模应用**（如用频率估计概率、保险定价）——属应用题，非本讲主线。