

GRADUS

数学 III

解答册

循迹而上 · 拾级登顶

寻迹 · 梦叶子 编著

目录

第一章 计数原理	1
1.1 例题详解	1
1.2 知识点 (1) 分类加法与分步乘法的判别与应用详解	1
1.3 例题详解	3
1.4 知识点 (1) 特殊元素优先与相邻 (不相邻) 问题详解	3
1.5 知识点 (2) 捆绑法与插空法的易错辨析详解	4
1.6 例题详解	5
1.7 知识点 (1) 组合的基本应用详解	5
1.8 知识点 (2) 排列与组合的选用辨析详解	7
1.9 知识点 (3) 均匀分组与分配问题详解	7
1.10 例题详解	8
1.11 知识点 (1) 通项公式与指定项详解	8
1.12 知识点 (2) 二项式系数与系数的区分详解	9
第二章 概率	11
2.1 例题详解	11
2.2 知识点 (1) 古典概型与等可能前提详解	11
2.3 例题详解	13
2.4 知识点 (1) 条件概率详解	13
2.5 知识点 (2) 全概率与贝叶斯公式详解	14
2.6 例题详解	15
2.7 知识点 (1) 事件独立性的判定详解	16
2.8 知识点 (2) 二项分布与“至少一次”问题详解	17
2.9 例题详解	18
2.10 知识点 (1) 期望与方差详解	19

2.11 知识点 (2) 正态分布与 3σ 原则详解	20
第三章 统计	23
3.1 例题详解	23
3.2 知识点 (1) 普查与抽样调查详解	23
3.3 知识点 (2) 简单随机抽样与系统抽样详解	24
3.4 知识点 (3) 分层抽样详解	24
3.5 例题详解	25
3.6 知识点 (1) 直方图的读图与频率详解	25
3.7 知识点 (2) 由直方图估计统计量详解	26
3.8 例题详解	27
3.9 知识点 (1) 百分位数详解	28
3.10 知识点 (2) 四分位数与箱线图详解	28
3.11 例题详解	29
3.12 知识点 (1) 平均数、中位数与众数详解	29
3.13 知识点 (2) 方差与标准差详解	30
第四章 成对数据的统计分析	33
4.1 例题详解	33
4.2 知识点 (1) 相关关系与函数关系详解	33
4.3 知识点 (2) 散点图与“相关不等于因果”详解	34
4.4 知识点 (3) 相关系数详解	34
4.5 例题详解	35
4.6 知识点 (1) 回归直线与最小二乘法详解	36
4.7 知识点 (2) 残差与预测 (内插与外推) 详解	37
4.8 例题详解	38
4.9 知识点 (1) 列联表与期望频数详解	39
4.10 知识点 (2) 卡方统计量与临界值详解	39
第五章 数学建模与拓展	45
5.1 例题详解	45
5.2 知识点 (1) 建模四步与模型选择详解	46
5.3 知识点 (2) 现实定义域与模型检验详解	46

5.4	例题详解	47
5.5	知识点 (1) 数据形状决定模型家族详解	48
5.6	知识点 (2) 拟合、过拟合与检验详解	48
5.7	例题详解	50
5.8	知识点 (1) 矩阵的加减与数乘详解	51
5.9	知识点 (2) 矩阵乘法与线性方程组详解	51
5.10	例题详解	52
5.11	知识点 (1) 命题逻辑与反证法详解	52
5.12	知识点 (2) 数学史与新技术输出解读详解	53

第六章	卷末综合检测	55
------------	---------------	-----------

第一章 计数原理

1.1 例题详解

例题 1.1 解析

思考路径：判断分类还是分步——问“选完这一项，整件事完成了吗”。(1) 选一份饭就买好了一份，完成 → 分类加法；(2) 只选主食还没完成一餐，必须主食、荤菜、素菜都选齐 → 分步乘法。

规范解：(1) 三类方案互不重叠、合起来覆盖全部选择：

$$N = 3 + 4 + 2 = 9$$

答：共有 9 种选择。(2) 分三步：主食 3 种、荤菜 4 种、素菜 2 种：

$$N = 3 \times 4 \times 2 = 24$$

答：共有 24 种搭配。

易错点：(1) 是“或”（选其中一类即完成），(2) 是“且”（三类都要选）——判断标准不是“题目里数多不多”，而是“完成一件事需要几步”。

例题 1.2 解析

思考路径：组成四位数要依次确定千位、百位、十位、个位，分 4 步。但 0 不能做首位，所以第一步（千位）的选择数会“卡住”。

规范解：第 1 步（千位）：1、2、3 中选，3 种（0 无资格）；第 2 步（百位）：剩下 3 个数字任选，3 种；第 3 步（十位）：2 种；第 4 步（个位）：1 种。

$$N = 3 \times 3 \times 2 \times 1 = 18$$

答：共有 18 个。

易错点：若按 $4! = 24$ 算，就把 0 开头的“三位数”（如 0123）错当成了四位数。首位受限是数字组数题的第一道门槛。

1.2 知识点 (1) 分类加法与分步乘法的判别与应用详解

A1. C。

思路：先分类（直达 / 经乙地），再在“经乙地”类内分步。

解析：直达 2 种；经乙地 $4 \times 3 = 12$ 种；共 $2 + 12 = 14$ 种。**易错点：**直达与经乙地是“或”的关系用加法，别把直达的 2 也乘进去（ $4 \times 3 \times 2$ 就错了）。

A2. C。

思路：每个元素“在子集中 / 不在子集中”，4 步乘法。

解析： $2^4 = 16$ 。

A3. C。

思路：逐区涂色，相邻约束只影响下一步。

解析：A 有 3 种；B 不与 A 同色，2 种；C 不与 B 同色，2 种。 $N = 3 \times 2 \times 2 = 12$ 。**易错点：**C 只受 B 限制，不受 A 限制——不必与 A、B 都不同。

A4. B。

思路：分两步，主食与饮料各选一。

解析： $N = 5 \times 3 = 15$ 。

A5. B。

思路：任选 1 人，按性别分类相加。

解析： $N = 4 + 3 = 7$ 。**易错点：**只选 1 人是“加”，不是“乘”（ $4 \times 3 = 12$ 是“1 男 1 女”的两步问题）。

A6. C。

思路：任取 1 本，按科类分类相加。

解析： $N = 3 + 4 = 7$ 。

M1. B。

思路：检查分类是否“不重不漏”——0 根本没有资格做首位。

解析：首位只能取 1、2、3，小刚的三类已覆盖全部四位数， $3 \times 3 \times 2 \times 1 = 18$ 个；小红把“首位为 0”也当一类，而那 6 个“数”实际是三位数，多算了一类。

为什么错：选 A 的同学默认“4 个数字全排 4!”，把首位为 0 的 6 个结果（0123、0132、0213、0231、0312、0321）也当成四位数。“0 不能做首位”是数字组数第一铁律：数位问题必须先安排受限的首位。

C1. (1) 10000; (2) 5040; (3) 3439。

思路：(1) 每一位独立 10 选， 10^4 ；(2) 4 位互不相同，相当于从 10 个数字里依次取 4 个排好；(3) “至少含 1 个 0”用补集：总数 - 不含 0 的密码数。

解析：(1) $N = 10^4 = 10000$ 。(2) $N = P_{10}^4 = 10 \times 9 \times 8 \times 7 = 5040$ 。(3) 不含 0 的密码每一位从 1 9 选： $9^4 = 6561$ 。

$$N = 10000 - 6561 = 3439$$

易错点：第(3)问补集里“不含0”是 9^4 （每一位9种），不是 P_9^4 ——因为数字可以重复使用。可重复用“乘方”，不可重复用“排列/组合”，先分清这一点。

1.3 例题详解

例题 1.3 解析

思考路径：甲受“不在两端”限制，先安排甲；其余5人无限制，后全排。

规范解：第1步：甲站中间4个位置之一：4种；第2步：其余5人任意站剩余5个位置： $P_5^5 = 120$ 种。

$$N = 4 \times 120 = 480$$

答：共有480种站法。（检验：无限制 $6! = 720$ ，甲在最左 $5! = 120$ ，甲在最右 120 ， $720 - 240 = 480$ 。）

易错点：受限制元素（甲）必须优先安排；若先排其他人再补甲，往往要回头处理，容易重复或遗漏。

例题 1.4 解析

思考路径：“2名女生相邻”→捆绑法：把女生捆成1块，块与4名男生共5个对象全排，再乘块内2名女生的顺序。

规范解：第1步：女生块与4名男生共5个对象排列： $P_5^5 = 120$ 种；第2步：块内2名女生交换顺序： $P_2^2 = 2$ 种。

$$N = 120 \times 2 = 240$$

答：共有240种排法。

易错点：捆绑法两步缺一不可——“块外排列”与“块内排列”都要乘；只算 $5!$ 就漏了块内两种顺序。

1.4 知识点（1）特殊元素优先与相邻（不相邻）问题详解

A7. D。

思路：甲不能站排头，先安排甲（特殊元素优先）。

解析：甲站非排头的6个位置之一：6种；其余6人全排： $6! = 720$ 种。 $N = 6 \times 720 = 4320$ 。（或用补集 $7! - 6! = 5040 - 720 = 4320$ 。）

A8. C。

思路：甲、乙不相邻→插空法。

解析：先排其余4人： $P_4^4 = 24$ 种；4人形成5个空（含两端），甲、乙插入其中2个空： $P_5^2 = 5 \times 4 = 20$ 种。 $N = 24 \times 20 = 480$ 。**易错点：**插空要插进“空位”而不是“人旁边”，空数=其余人数+1；甲乙两人也有先后顺序（ P_5^2 而非 C_5^2 ）。

A9. C。

思路：4 人全排列。

解析： $N = 4! = 24$ 。

A10. B。

思路：从 4 个数字中取 2 个排成一列。

解析： $N = P_4^2 = 4 \times 3 = 12$ 。易错点：是“排列”不是“组合”——数字顺序不同代表不同两位数（12 与 21 是两个数）。

B1. (2022 · 新高考 II 卷 · 第 5 题) B。

思路：丙、丁相邻用捆绑；甲不站两端用插空（把甲插进中间两个空位）。

解析：先捆丙、丁： $2! = 2$ 种；捆块与乙、戊共 3 个对象排列： $3! = 6$ 种；甲不能站两端，只能插入这 3 个对象形成的 4 个空中的中间 2 个空：2 种。

$$N = 2 \times 6 \times 2 = 24$$

易错点：插空法里“含两端”与“不含两端”要看题目——本题甲不在两端，所以只插中间 2 个空；若题目要求“不相邻”才插全部空（含两端）。

1.5 知识点 (2) 捆绑法与插空法的易错辨析详解

M2. B。

思路：捆绑法只保证“A、B、C 三人相邻”，不保证块内“A、B 不相邻”——块内要再分类。

解析：满足“A、B、C 相邻且 A、B 不相邻”时，三人块内只能 C 居中： ACB 或 BCA ，共 2 种；块与 D、E 共 3 个对象全排： $3! = 6$ 种。 $N = 2 \times 6 = 12$ 。

为什么错：选 A 的同学直接套捆绑法，把块内当“任意 $3! = 6$ 种”——但块内 6 种中有 4 种让 A、B 挨在一起（ ABC 、 BAC 、 CAB 、 CBA ），违背附加条件。捆绑法只解决“相邻”这一条限制；块内若还有更细的约束，必须在块内再分类。

M3. B。

思路：甲不在两端 → 特殊元素优先；甲的合法位置是中间 3 个，不是“120 除以 2”。

解析：甲站中间 3 个位置之一：3 种；其余 4 人全排： $4! = 24$ 种。 $N = 3 \times 24 = 72$ 。（检验：补集 $5! - 2 \times 4! = 120 - 48 = 72$ 。）

为什么错：选 A 的同学以为“两端占 5 个位置的一半”，把 120 除以 2——但甲站两端的位置不是 2.5 个。甲不能站的位置是 2 个（排头、排尾），能站的是 3 个。受限制元素的位置数，要具体数出来，不能“平均”。

1.6 例题详解

例题 1.5 解析

思考路径：同一个“选4人”，看“换位置后结果变不变”——四个岗位不同，换人当岗结果变 → 排列；小组不分角色，换顺序结果不变 → 组合。

规范解：(1) 四个岗位互不相同，选人并分配岗位有顺序： $P_6^4 = 6 \times 5 \times 4 \times 3 = 360$ 种。(2) 组成小组不分角色： $C_6^4 = \frac{6 \times 5 \times 4 \times 3}{4!} = 15$ 种。

易错点：同一组人，一个是 360、一个是 15——差的正是“岗位”带来的顺序。判定标尺永远是“换位置变不变”。

例题 1.6 解析

思考路径：(1) 三组都叫“组”，没有名字 → 先分组、再除组间顺序；(2) 甲、乙、丙班有名字 → 直接分步乘。

规范解：(1) 先取再除重：

$$N = \frac{C_6^2 C_4^2 C_2^2}{3!} = \frac{15 \times 6 \times 1}{6} = 15$$

(2) 组有名字，无需除重：

$$N = C_6^2 C_4^2 C_2^2 = 15 \times 6 \times 1 = 90$$

答：(1) 15 种；(2) 90 种。

易错点：同样“每组 2 人”，给三组贴上甲、乙、丙的标签就多出 $3! = 6$ 倍。均匀分组不除重、或有名字的组误除重，是最常见的两个坑。

1.7 知识点 (1) 组合的基本应用详解

A11. A。

思路：选 3 人参加同一竞赛，不看顺序 → 组合。

解析： $C_8^3 = \frac{8 \times 7 \times 6}{3!} = 56$ 。

A12. C。

思路：正面分类多，用补集：总数 - 全是男生。

解析：总数 $C_{10}^3 = 120$ ，全男生 $C_6^3 = 20$ ，至少 1 名女生 $= 120 - 20 = 100$ 。

A13. B。

思路：对角线数 = 任意两点连线总数 - 边数。

解析： $C_8^2 - 8 = 28 - 8 = 20$ 。

A14. B。

思路：选 2 人不分角色 → 组合。

解析： $C_6^2 = \frac{6 \times 5}{2!} = 15$ 。

A15. B。

思路：解组合数方程。

解析： $C_n^2 = \frac{n(n-1)}{2} = 21$ ，即 $n(n-1) = 42$ ，得 $n = 7$ 。**易错点：** n 是正整数，解方程注意取正根。

B2. (2023 · 新高考 I 卷 · 第 13 题) 64。

思路：按选修门数分类，每类内部再分步（体育、艺术各选几门）。

解析：选 2 门时每类各 1 门： $C_4^1 C_4^1 = 16$ 种；选 3 门时：体育 1 艺术 2 为 $C_4^1 C_4^2 = 24$ ，体育 2 艺术 1 为 $C_4^2 C_4^1 = 24$ ，共 48 种。

$$N = 16 + 48 = 64$$

易错点：选 3 门时“每类至少 1 门”排除了 3 门全体育、3 门全艺术两类，别漏掉这个约束；且“选 2 门”与“选 3 门”是两类方案，用加法合并。

B3. (2023 · 全国甲卷 · 第 9 题) B。

思路：先确定“两天都参加”的那个人，再安排周六、周日各补 1 人。

解析：第 1 步：从 5 人中选 1 人两天都参加： $C_5^1 = 5$ 种；第 2 步：从余下 4 人中选 2 人分别安排到周六、周日： $P_4^2 = 4 \times 3 = 12$ 种。

$$N = 5 \times 12 = 60$$

易错点：周六、周日是“两个不同的日子”，补的 2 人谁去周六谁去周日有区别，必须用排列 P_4^2 而不是组合 C_4^2 ——用 C_4^2 会算成 30，正是陷阱选项。

B4. (2022 · 全国甲卷理 · 第 15 题) $\frac{6}{35}$ 。

思路：古典概型——总样本点用组合计数 C_8^4 ；“4 点共面”要把正方体里的共面情形找全，不只表面还有对角面。

解析：从 8 个顶点中任取 4 个，共 $C_8^4 = 70$ 种。4 点共面的情形：(1) 6 个表面（上、下、前、后、左、右）；(2) 6 个对角面（过两条相对棱或体对角线所在平面，如 AA_1C_1C 、 AB_1C_1D 等）。共 $6 + 6 = 12$ 种。

$$P = \frac{12}{70} = \frac{6}{35}$$

为什么错：只数 6 个表面会得 $\frac{6}{70} = \frac{3}{35}$ ——漏掉了对角面。正方体里的“4 点共面”必须“表面 + 对角面”两类都找全，这是几何计数题最容易丢项的地方。

易错点：本题表面考概率，实质考组合计数 C_8^4 与几何构型；概率的分子分母都必须用“数”的方式严谨数出。

1.8 知识点 (2) 排列与组合的选用辨析详解

M4. C。

思路：同一“选 2 人”，一个职务不同（有顺序）、一个代表不分先后（无顺序）。

解析：(1) 班长、副班长两个位置不同 $\rightarrow$ 排列 $P_5^2 = 5 \times 4 = 20$ ；(2) 代表不分先后 $\rightarrow$ 组合 $C_5^2 = 10$ 。

为什么错：选 A 的同学把两问都当组合——没看到“班长/副班长”是两张不同的椅子；选 B 的同学都当排列——“先选甲再选乙”和“先选乙再选甲”是同一个代表组合，分步会把它数成两次。**翻译实际问题时先问：换位置后结果变不变？**

1.9 知识点 (3) 均匀分组与分配问题详解

M5. B。

思路：两个“2 人组”无编号，分步相乘会把同一分组数两次。

解析：先分步得 $C_4^2 C_2^2 = 6 \times 1 = 6$ ，其中 $\{a, b\}|\{c, d\}$ 与 $\{c, d\}|\{a, b\}$ 是同一个分组，重复一倍，除以 $2!$ ： $\frac{6}{2} = 3$ 种。

为什么错：选 A 的同学只记得“分步相乘”，忘了组间无区别时要除重。判断口诀：**组没名字 $\rightarrow$ 除组数阶乘；组有名字 $\rightarrow$ 不用除。**（对照：分给两个不同的班级则答案是 6。）

C2. (1) 30；(2) 15；(3) 见下。

思路：(1) 直接给传送口分配，按“分步乘法”逐口选件；(2) 分堆时两个“2 件堆”无区别，要除 $2!$ ；(3) 把“分堆 $\times$ 分配”的重复路径挑出来。

解析：(1) A 口先选 2 件： $C_5^2 = 10$ ；B 口再从剩余 3 件选 2 件： $C_3^2 = 3$ ；C 口收最后 1 件： $C_1^1 = 1$ 。

$$N = 10 \times 3 \times 1 = 30$$

(2) 先分成 2、2、1 三堆，两堆 2 件无区别：

$$N = \frac{C_5^2 C_3^2 C_1^1}{2!} = \frac{10 \times 3 \times 1}{2} = 15$$

(3) 不能乘以 $3!$ 的原因：分堆时已经除以 $2!$ 抹掉了两个“2 件堆”的顺序；若再乘 $3!$ ，就把“这两堆 2 件互换”的重复路径又数回来了，会得到 $15 \times 6 = 90$ ，比正确答案大 3 倍。正确做法：1 件堆固定给 C (1 种)，两堆 2 件分给 A、B 两口有 $2! = 2$ 种：

$$N = 15 \times 2 = 30$$

与 (1) 一致。

易错点：这是“均匀分组除重 + 先分后配”的综合演练——分堆除的是“无名字的组间顺序”，分配时只乘“有名字的组间顺序”，两层顺序要分清。

1.10 例题详解

例题 1.7 解析

思考路径：写出通项 T_{r+1} ，把 $(2x)$ 与 $\left(-\frac{1}{x}\right)$ 原样代入，合并 x 的指数，令指数为 0 解 r 。

规范解：

$$T_{r+1} = C_6^r (2x)^{6-r} \left(-\frac{1}{x}\right)^r = C_6^r \cdot 2^{6-r} \cdot (-1)^r \cdot x^{6-r} \cdot x^{-r} = C_6^r \cdot 2^{6-r} \cdot (-1)^r \cdot x^{6-2r}$$

常数项要求 x 的指数为 0: $6 - 2r = 0$, 得 $r = 3$ 。

$$T_4 = C_6^3 \cdot 2^3 \cdot (-1)^3 = 20 \times 8 \times (-1) = -160$$

答：常数项为 -160 。

易错点： $r = 3$ 对应的是第 4 项 (T_{r+1})，不是第 3 项——这正是通项下标陷阱（见 M6）；另外别忘了 $(-1)^r$ 的符号。

例题 1.8 解析

思考路径：(1) 二项式系数最大只由 n 的奇偶决定， $n = 8$ 为偶数，中间一项最大；(2) “系数”带 2^r 因子，峰值会被往右推，要用相邻项系数比找峰值。

规范解：(1) $n = 8$ ，中间一项二项式系数最大：

$$T_5 = C_8^4 x^4 = 70x^4$$

(2) 设第 $r + 1$ 项系数为 $a_r = C_8^r \cdot 2^r$ ，考察相邻项之比：

$$\frac{a_{r+1}}{a_r} = \frac{C_8^{r+1} \cdot 2^{r+1}}{C_8^r \cdot 2^r} = \frac{8-r}{r+1} \cdot 2$$

当 $\frac{8-r}{r+1} \cdot 2 > 1$ 即 $r < 5$ 时系数递增， $r \geq 5$ 时递减。比较 $r = 4$ 与 $r = 5$: $a_4 = C_8^4 \cdot 2^4 = 70 \times 16 = 1120$, $a_5 = C_8^5 \cdot 2^5 = 56 \times 32 = 1792$ 。故系数最大的项为 $T_6 = 1792x^5$ 。

易错点：同一个 $n = 8$ ，二项式系数最大在 T_5 ($70x^4$)，系数最大却在 T_6 ($1792x^5$) ——“二项式系数”与“系数”是两码事，底数里的常数因子 2 会把峰值往右推。

1.11 知识点 (1) 通项公式与指定项详解

A16. B。

思路： $(1+x)^5$ 中 x^3 的系数即 C_5^3 。

解析： $C_5^3 = \frac{5 \times 4 \times 3}{3!} = 10$ 。

A17. C。

思路： $(1+x)^6$ 中 x^2 的系数即 C_6^2 。

解析: $C_6^2 = \frac{6 \times 5}{2!} = 15$ 。

A18. B。

思路: $(1+x)^n$ 展开共 $n+1$ 项。

解析: $n+1=10$, 故 $n=9$ 。易错点: 项数是 $n+1$ 不是 n ——别把“共 10 项”直接当成 $n=10$ 。

B5. (2022 · 新高考 I 卷 · 第 13 题) -28 。

思路: $\left(1 - \frac{y}{x}\right)(x+y)^8 = (x+y)^8 - \frac{y}{x}(x+y)^8$, 分别找 x^2y^6 项再相减。

解析: $(x+y)^8$ 中 x^2y^6 的系数为 $C_8^6 = 28$; $\frac{y}{x}(x+y)^8$ 中, 需 $(x+y)^8$ 提供 x^3y^5 , 其系数为 $C_8^5 = 56$, 再乘 $\frac{y}{x}$ 得 x^2y^6 。

$$28 - 56 = -28$$

易错点: 第二部分的处理—— $\frac{y}{x}$ 会把 x^3y^5 “降一个 x 、升一个 y ”变成 x^2y^6 , 对应的是 $(x+y)^8$ 里 x^3y^5 的项 ($r=5$), 不是 $r=6$ 。 r 取错一档, 符号和数值全错。

M6. B。

思路: T_{r+1} 的下标与 r 差 1—— $r=3$ 时对应 T_4 。

解析: $r=3$ 求得 $T_4 = C_6^3 \cdot 2^3 \cdot (-1)^3 = -160$, 数值正确, 但它是展开式的第 4 项, 不是第 3 项。

为什么错: 选 C 或 A 的同学, 要么把 $(-1)^3$ 的符号弄丢, 要么把“第 3 项”与“ $r=3$ ”划等号。记住: 第 $r+1$ 项对应 T_{r+1} , “第 5 项”要用 $r=4$; 求“常数项”“含 x^k 的项”时, 先解出 r 再回代 T_{r+1} , 最后说项数时记得 $+1$ 。

1.12 知识点 (2) 二项式系数与系数的区分详解

A19. B。

思路: 各项系数和 = 令 $x=1$ 。

解析: $(1-3 \times 1)^6 = (-2)^6 = 64$ 。易错点: 这是“系数和”不是“二项式系数和”——二项式系数和固定是 $2^6 = 64$, 这里碰巧都是 64; 若换 $(1-3x)^7$ 则系数和为 $(-2)^7 = -128$, 与二项式系数和 128 不同。

M7. B。

思路: 底数带常数因子 $(2x, 3y)$ 时, 该项系数 $= C_n^r \cdot 2^{n-r} \cdot 3^r$ 。

解析: 令 x, y 指数均为 3, 得 $r=3$;

$$C_6^3 \cdot 2^3 \cdot 3^3 = 20 \times 8 \times 27 = 4320$$

为什么错: 选 A 的同学只写了 $C_6^3 = 20$, 把藏在底数里的 2^3 和 3^3 全丢了。“二项式系数”是 C_n^r , “该项的系数”还要乘底数的常数因子——只写 C_n^r 的答案, 在 $(2x+3y)^n$ 这类题里一定

错。

本章综合详解

S1 解析

思考路径：三个小问共用“0 不能做首位”的限制。(1) 先定首位再排其余；(2) 按末位分三类（末位 0 / 末位 2 / 末位 4）；(3) “1 与 2 相邻”捆绑，注意块内两种顺序 + 块放哪 + 0 不能首位。

规范解：(1) 首位从 1、2、3、4 中选：4 种；其余 4 位全排： $4! = 24$ 。

$$N = 4 \times 24 = 96$$

(2) 末位为 0：其余 4 位任意排， $4! = 24$ 个；末位为 2 或 4（两类对称）：首位不能为 0，从其余 3 个非零数字选 1 个，3 种；中间 3 位全排， $3! = 6$ ；每类 $3 \times 6 = 18$ 个，两类共 36 个。

$$N = 24 + 36 = 60$$

(检验：总数 96，奇数末位为 1 或 3，各 $3 \times 6 = 18$ ，共 36， $96 - 36 = 60$ 。)(3) 把 1、2 捆成块（块内 $2! = 2$ 种），块与 0、3、4 共 4 个对象排列，0 不能首位。0 不在首位的 4 对象排列数 $= 4! - 3! = 24 - 6 = 18$ ，乘块内 2 种：

$$N = 18 \times 2 = 36$$

易错点：(2) 末位 2、4 两类都受“首位非 0”限制，容易漏算“首位排除 0”这一步；(3) 捆绑后 0 仍不能做首位——“块 + 0 + 3 + 4”排时 0 不能在首位，别把块的顺序和 0 的位置混在一起。

S2 解析

思考路径：二项式系数和是 2^n ，反解 n ；各项系数和令 $x = 1$ ； x^3 系数用通项。

规范解：(1) 二项式系数和 $2^n = 64$ ，故 $n = 6$ 。(2) 各项系数和：令 $x = 1$ ，

$$(1 + 2)^6 = 3^6 = 729$$

(3) 通项 $T_{r+1} = C_6^r \cdot 2^r x^r$ ，令 $r = 3$ ：

$$T_4 = C_6^3 \cdot 2^3 = 20 \times 8 = 160$$

答：(1) $n = 6$ ；(2) 729；(3) 160。

易错点：三个概念别混——二项式系数和固定 2^n （与底数无关）；各项系数和要令 $x = 1$ （与底数有关）；“某项系数”要乘底数的常数因子 2^r 。

第二章 概率

2.1 例题详解

例题 2.1 解析

思考路径：古典概型先查“等可能”。两枚骰子是两枚不同的骰子，结果用有序数对 (a, b) 表示，36 个样本点等可能；不能把“和值”当样本空间。

规范解：用有序数对 (a, b) 表示结果，共有 $6 \times 6 = 36$ 个样本点，且等可能。和为 3 的样本点： $(1, 2)$ 、 $(2, 1)$ ，共 2 个。

$$P = \frac{2}{36} = \frac{1}{18}$$

答：点数之和为 3 的概率为 $\frac{1}{18}$ 。

易错点：若把样本空间写成“和值 2 12 共 11 种”，则 $(1, 2)$ 与 $(2, 1)$ 被并成一个——但它们是两枚不同骰子的两个结果，这样计数不等可能。“等可能”是对样本点而言的，不是对“和值”而言的（详见 M1）。

例题 2.2 解析

思考路径：“至少命中 1 次”正面要枚举 4 种情况，转反面：一次都没中。每次脱靶概率 $1 - 0.6 = 0.4$ ，各次独立，4 次都脱靶为 0.4^4 。

规范解：

$$P(\text{至少命中 1 次}) = 1 - 0.4^4 = 1 - 0.0256 = 0.9744$$

答：至少命中 1 次的概率为 0.9744。

易错点： $0.4^4 = 0.0256$ （不是 0.16）；补集是“一次都没有”，对应“至少一次”的反面，别把两者搞反。

2.2 知识点 (1) 古典概型与等可能前提详解

A1. A。

思路：点数不小于 4 即 $\{4, 5, 6\}$ 。

解析： $P = \frac{3}{6} = \frac{1}{2}$ 。

A2. A。

思路：取球不看顺序，用组合计数。

解析：总数 $C_5^2 = 10$ ；恰 1 红 1 蓝 $C_2^1 C_3^1 = 6$ ； $P = \frac{6}{10} = \frac{3}{5}$ 。

A3. C。

思路：至少 2 枚正面 = 恰 2 枚或恰 3 枚。

解析： $\frac{C_3^2 + C_3^3}{2^3} = \frac{3+1}{8} = \frac{1}{2}$ 。

A4. A。

思路：古典概型，分子分母都数球的个数。

解析：总数 7 个球，红球 3 个， $P = \frac{3}{7}$ 。

A5. B。

思路：5 个数中偶数 2 个。

解析： $P = \frac{2}{5}$ 。

B1. (2022 · 新高考 I 卷 · 第 5 题) D。

思路：先数样本空间 ($C_7^2 = 21$)，再数互质对。

解析：28 的 7 个整数中取 2 个，总数 $C_7^2 = 21$ 对。互质对逐一列出：(2, 3), (2, 5), (2, 7), (3, 4), (3, 5), (3, 7), (3, 8), (4, 5), (4, 7), (5, 6), (5, 7), (6, 7), 共 14 对。

$$P = \frac{14}{21} = \frac{2}{3}$$

易错点：枚举互质对要“不重不漏”——按较小数从小到大枚举 (2 开头的 3 对、3 开头的 4 对、4 开头的 2 对、5 开头的 3 对、6 开头的 1 对、7 开头的 1 对)，别漏 (3, 8) 这类“大数与小奇数互质”的对。

B2. (2022 · 全国乙卷 · 第 13 题) $\frac{3}{10}$ (或 0.3)。

思路：从 5 人选 3 人是组合计数；甲、乙都入选，则剩下 1 个名额在其余 3 人中选。

解析：总选法 $C_5^3 = 10$ ；甲、乙都入选的选法为从丙、丁、戊中再选 1 人： $C_3^1 = 3$ 。

$$P = \frac{C_3^1}{C_5^3} = \frac{3}{10}$$

易错点：分子是“甲、乙都入选”的选法数 (固定两人，只选第 3 人)，不是“甲或乙”——若把分子算成 C_5^3 之类的会得 1，一眼假。也可枚举：甲乙丙、甲乙丁、甲乙戊共 3 个基本事件。

M1. C。

思路：查“等可能”前提——“和值”不是等可能样本点。

解析：正确做法：36 个有序数对等可能，和为 5 的样本点为 (1, 4), (2, 3), (3, 2), (4, 1) 共 4 个， $P = \frac{4}{36} = \frac{1}{9}$ 。

为什么错：选 A 的同学把“和值”当成等可能结果——和为 2 只有 (1, 1) 一个样本点，和为 5 有 4 个样本点，两者出现可能性不同，11 种和值并不等可能。古典概型必须先确认“等可能”是对谁说的：对样本点 (有序数对) 而言，不是对和值/点数集合而言。

2.3 例题详解

例题 2.3 解析

思考路径：已知“第一次摸到红球”，样本空间缩成“首红”——袋中剩 2 红 2 蓝。用“缩空间”法或公式法。

规范解：方法一（缩空间）：已知首红，袋中剩 2 红 2 蓝共 4 球，

$$P(\text{二红} | \text{首红}) = \frac{2}{4} = \frac{1}{2}$$

方法二（公式）：设 $A =$ “首红”， $B =$ “二红”。 $P(A) = \frac{3}{5}$ ， $P(A \cap B) = \frac{3}{5} \times \frac{2}{4} = \frac{3}{10}$ ，

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{3/10}{3/5} = \frac{1}{2}$$

答：第二次也摸到红球的概率为 $\frac{1}{2}$ 。

易错点：分子的 $P(A \cap B)$ 是“两次都红” ($\frac{3}{10}$)，不是“第二次红” ($\frac{3}{5}$)——后者忽略了“第一次必须是红球”这个条件（详见 M2）。

例题 2.4 解析

思考路径：任取一件，它可能来自甲厂也可能来自乙厂，两个来源的次品率不同，不能只看某一个厂——用全概率公式“先分类（来源），再求和”。

规范解：设 $A_1 =$ “来自甲厂”， $A_2 =$ “来自乙厂”， $B =$ “次品”。 A_1 、 A_2 两两互斥且并集为 Ω ，构成完备事件组。

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) = 0.7 \times 0.01 + 0.3 \times 0.03 = 0.007 + 0.009 = 0.016$$

答：任取一件是次品的概率为 0.016，即 1.6%。

易错点：完备事件组的每一项都不能漏——漏掉乙厂那一项会得到 0.007（明显偏小，见 M3）。

2.4 知识点（1）条件概率详解

A6. B。

思路：已知第一个是白球，样本空间缩到“剩 1 白 3 黑”。

解析： $P = \frac{3}{4}$ 。**易错点：**分母不能再用原来的 5 个球——条件已把袋子状态改了。

A7. A。

思路：条件概率公式，分子是交集。

解析： $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.1}{0.4} = 0.25$ 。

A8. B。

思路：已知第一次抽到次品后，样本空间缩到“剩 9 件、其中 2 件次品”。

解析： $P = \frac{2}{9}$ 。**易错点：**第一次抽到次品后次品剩 2 件、总数剩 9 件——分子分母都要跟着“缩空间”同步减。

B3. (2023 · 全国甲卷 · 第 6 题) A。

思路：先由并集公式求交集，再用条件概率。

解析：设 $A =$ “爱好滑冰”， $B =$ “爱好滑雪”。 $P(A) = 0.6$, $P(B) = 0.5$, $P(A \cup B) = 0.7$ 。
由 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ 得 $P(A \cap B) = 0.6 + 0.5 - 0.7 = 0.4$ 。

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.4}{0.5} = 0.8$$

易错点：条件概率的分母是 $P(B) = 0.5$ (条件“爱好滑雪”)，分子是交集 0.4；用 $P(A) = 0.6$ 做分母会得到 $\frac{2}{3}$ ，那是“爱好滑冰条件下也爱好滑雪”，方向反了。

M2. B。

思路：条件概率分子必须是“两次都红”(交集)，分母是“首红”。

解析： $P(\text{二红}|\text{首红}) = \frac{P(\text{首红} \cap \text{二红})}{P(\text{首红})} = \frac{\frac{3}{5} \times \frac{2}{4}}{\frac{3}{5}} = \frac{1}{2}$ (或直接缩空间：剩 2 红 2 蓝， $\frac{2}{4}$)。

为什么错：选 A 的同学把 $P(B|A)$ 当成了 $P(B)$ ——“第二次红”在不知道第一次结果时是 $\frac{3}{5}$ ，但知道第一次是红球后，袋中只剩 2 红 2 蓝，概率变为 $\frac{1}{2}$ 。 $P(B|A)$ 的分子是交集 $P(A \cap B)$ ，不是 $P(B)$ ；条件改变了样本空间。

2.5 知识点 (2) 全概率与贝叶斯公式详解

B4. (2023 · 天津卷 · 第 13 题) 0.05; $\frac{3}{5}$ 。

思路：设各盒球数为 $5n, 4n, 6n$ ；第一空三个盒子独立取球，概率相乘；第二空混合后按白球总数占比算。

解析：设甲、乙、丙三盒球数分别为 $5n, 4n, 6n$ ，总数为 $15n$ 。甲盒白球 $5n \times 60\% = 3n$ ，乙盒白球 $4n \times 75\% = 3n$ ，丙盒白球 $6n \times 50\% = 3n$ 。(1) 三盒各取 1 个球都是黑球：

$$P = 40\% \times 25\% \times 50\% = 0.05$$

(2) 混合后白球共 $3n + 3n + 3n = 9n$ 个：

$$P = \frac{9n}{15n} = \frac{3}{5}$$

易错点：第二空是“混合后任取 1 球”，直接数白球占比，不需要条件概率；三个盒子黑球比例不同，不能简单平均。

M3. B。

思路：抽到的零件可能来自甲也可能来自乙，用全概率公式，两项相加。

解析： $P(\text{次品}) = 0.6 \times 0.02 + 0.4 \times 0.04 = 0.012 + 0.016 = 0.028$ ，即 2.8%。

为什么错：选 A 的同学只盯着“次品率最高的 4%”，把乙机床的次品率当成了总次品率——漏掉了甲机床供货 60% 的大头占比。**全概率公式的完备事件组（甲、乙）每一项都不能漏：先按来源分类，类内相乘（占比 × 条件概率），类间相加。**

C1. (1) 0.01098; (2) 0.090。

思路：设 $A =$ “携带者”， $B =$ “检测阳性”。这是“由果溯因”——先求全概率 $P(B)$ ，再用贝叶斯公式。

解析： $P(A) = 0.001$ ， $P(B|A) = 0.99$ ， $P(B|\bar{A}) = 0.01$ 。(1) 全概率公式：

$$P(B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) = 0.001 \times 0.99 + 0.999 \times 0.01 = 0.00099 + 0.00999 = 0.01098$$

(2) 贝叶斯公式：

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} = \frac{0.00099}{0.01098} \approx 0.090$$

答：检测阳性概率为 0.01098；已知阳性确为携带者的概率约为 0.090。

易错点： $P(B)$ 的两项都不能漏（携带者阳性 + 非携带者误报）；阳性率约 1.098%，但真正患病只有约 9%——低患病率下，大多数“阳性”来自误报，这正是贝叶斯公式的警示。

2.6 例题详解

例题 2.5 解析

思考路径：先看集合关系（互斥？），再算概率等式（独立？）——两个层面分开判。

规范解： $A = \{1, 3, 5\}$ ， $B = \{2, 4, 6\}$ ， $A \cap B = \emptyset$ ，故 A 与 B **互斥**。 $P(A) = \frac{1}{2}$ ， $P(B) = \frac{1}{2}$ ， $P(A)P(B) = \frac{1}{4}$ ，而 $P(A \cap B) = 0$ ， $0 \neq \frac{1}{4}$ ，故 A 与 B **不独立**。答： A 与 B 互斥但不独立。

易错点：“一个数不可能既是奇数又是偶数”是互斥；但独立要的是概率等式 $P(AB) = P(A)P(B)$ 成立。当 $P(A), P(B) > 0$ 时，互斥与独立不可兼得——互斥给 $P(AB) = 0$ ，独立却要求 $P(AB) > 0$ 。

例题 2.6 解析

思考路径：核对二项分布三条件（固定 5 次、每次两结果、独立且 p 恒定）后， $X \sim B(5, 0.8)$ ，代入公式。

规范解：

$$P(X = 3) = C_5^3 (0.8)^3 (0.2)^2 = 10 \times 0.512 \times 0.04 = 0.2048$$

答：恰好命中 3 次的概率为 0.2048。

易错点：失败概率用 $1 - p = 0.2$ 的 $(n - k)$ 次方； $C_5^3 = 10$ 表示“哪 3 次命中”的选法，别漏。

2.7 知识点 (1) 事件独立性的判定详解

A9. A。

思路：验证 $P(A \cap B)$ 与 $P(A)P(B)$ 是否相等。

解析： $P(A) = \frac{3}{6} = \frac{1}{2}$, $P(B) = \frac{3}{6} = \frac{1}{2}$, $P(A \cap B) = \frac{3 \times 3}{36} = \frac{1}{4} = P(A)P(B)$, 故独立。

B5. (2024 · 新高考 I 卷 · 第 8 题) B。

思路：逐一验证独立判据 $P(A \cap B) = P(A)P(B)$ 。样本空间 $6 \times 6 = 36$ 个有序数对。

解析：各事件概率： $P(\text{甲}) = \frac{6}{36} = \frac{1}{6}$ (第一次是 1)； $P(\text{乙}) = \frac{6}{36} = \frac{1}{6}$ (第二次是 2)； $P(\text{丙}) = \frac{5}{36}$ (和为 8: (2,6), (3,5), (4,4), (5,3), (6,2))； $P(\text{丁}) = \frac{6}{36} = \frac{1}{6}$ (和为 7: (1,6), (2,5), (3,4), (4,3), (5,2), (6,1))。逐项验证：A: $\text{甲} \cap \text{丙} = \text{空}$ (和为 8 时第一次不可能取 1), $P = 0 \neq \frac{5}{216}$, 不独立。B: $\text{甲} \cap \text{丁} = \{(1,6)\}$, $P = \frac{1}{36}$; $P(\text{甲})P(\text{丁}) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$, 相等, **独立**。C: $\text{乙} \cap \text{丙} = \{(6,2)\}$, $P = \frac{1}{36} \neq \frac{5}{216}$, 不独立。D: $\text{丙} \cap \text{丁} = \text{空}$ (和 8 与和 7 不可能同时), $P = 0$, 不独立。

易错点：验证独立必须同时算“交集概率”与“乘积概率”并比较，不能凭直觉“感觉独立”；丙、丁事件的“和值”型样本点要写全。

B6. (2022 · 全国乙卷 · 第 10 题) D。

思路：三盘各结果相互独立。连胜两盘 $\iff$ 第二盘胜，且第一盘、第三盘至少胜一盘。设第二盘分别与甲、乙、丙比赛，写出三种安排下的连胜概率再比大小。

解析：设第二盘与丙比赛 (第一、三盘与甲、乙各赛一盘)，则

$$p_{\text{丙}} = p_3 [1 - (1 - p_1)(1 - p_2)] = p_3(p_1 + p_2 - p_1p_2)$$

同理，第二盘与乙比赛：

$$p_{\text{乙}} = p_2(p_1 + p_3 - p_1p_3)$$

第二盘与甲比赛：

$$p_{\text{甲}} = p_1(p_2 + p_3 - p_2p_3)$$

比较：

$$p_{\text{丙}} - p_{\text{乙}} = p_3(p_1 + p_2 - p_1p_2) - p_2(p_1 + p_3 - p_1p_3) = p_1(p_3 - p_2) > 0$$

$$p_{\text{乙}} - p_{\text{甲}} = p_2(p_1 + p_3 - p_1p_3) - p_1(p_2 + p_3 - p_2p_3) = p_3(p_2 - p_1) > 0$$

故 $p_{\text{丙}} > p_{\text{乙}} > p_{\text{甲}}$ ，该棋手在第二盘与丙比赛时连胜概率最大，答案 D。

为什么错：选 A 的同学以为“反正三盘都要打，次序不影响”——其实第二盘的位置最关键：第二盘若放胜率高的对手，能同时搭上“一二连胜”与“二三连胜”两条路径，收益最大；第二盘放胜率低的对手则白白浪费“承上启下”的一盘。选 B/C 的同学比较次序时少算了“第一、三盘至少胜一盘”的并集概率 (用 $p_i + p_j - p_i p_j$ 而不是简单相加)。**独立事件“至少一个”发生，要用并集概率 $p_i + p_j - p_i p_j$ ，不能只加一项或漏交叠。**

M4. B。

思路：互斥看集合 ($A \cap B = \emptyset$)，独立看概率等式。

解析： $A \cap B = \emptyset$ ，互斥成立； $P(AB) = 0$ ，而 $P(A)P(B) = \frac{1}{4} \neq 0$ ，独立不成立。

为什么错：选 A 的同学把“互斥”当成了“独立”——两个词都带“互”，但一个在空间层面（不能同时发生），一个在概率层面（互不影响）。当 $P(A), P(B) > 0$ 时，互斥与独立不可兼得：互斥给 $P(AB) = 0$ ，独立要求 $P(AB) = P(A)P(B) > 0$ ，直接矛盾。

2.8 知识点 (2) 二项分布与“至少一次”问题详解

A10. A。

思路：“至少一次 6 点”转补集：一次 6 点都没有。

解析：每次不是 6 点的概率为 $\frac{5}{6}$ ， $P = 1 - \left(\frac{5}{6}\right)^5$ 。

A11. B。

思路： $X \sim B(4, 0.7)$ ，求 $P(X = 2)$ 。

解析： $C_4^2(0.7)^2(0.3)^2 = 6 \times 0.49 \times 0.09 = 0.2646$ 。

A12. B。

思路： $X \sim B(3, 0.5)$ ，代入公式。

解析： $P(X = 2) = C_3^2(0.5)^2(0.5)^1 = 3 \times 0.125 = \frac{3}{8}$ 。

A13. C。

思路：硬币正面朝上概率 $p = \frac{1}{2}$ ， $X \sim B\left(4, \frac{1}{2}\right)$ 。

解析： $P(X = 2) = C_4^2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = 6 \times \frac{1}{16} = \frac{3}{8}$ 。

M5. B。

思路：判断模型——不放回不独立，用组合计数（超几何型）；有放回独立且 p 恒定，用二项分布。

解析：(1) 不放回：不是二项分布，

$$P = \frac{C_3^2 C_7^2}{C_{10}^4} = \frac{3 \times 21}{210} = \frac{63}{210} = \frac{3}{10}$$

(2) 有放回：每次次品率 $p = \frac{3}{10}$ 恒定且独立， $X \sim B\left(4, \frac{3}{10}\right)$ ，

$$P(X = 2) = C_4^2(0.3)^2(0.7)^2 = 6 \times 0.09 \times 0.49 = 0.2646$$

为什么错：选 A 的同学看到“抽 4 件”就套二项分布——不放回时第二次抽到次品的概率随第一次结果变化，各次不独立，三条件缺一条就不能用 $C_n^k p^k (1-p)^{n-k}$ 。先判断模型，再代公式：不放回 → 组合计数；有放回 → 二项分布。

M6. B。

思路：“至少命中 1 次”的反面是“4 次全没中”，补集一步到位。

解析： $1 - 0.3^4 = 1 - 0.0081 = 0.9919$ 。枚举四种情况（中 1、2、3、4 次）虽可得同样结果，但要算 4 项还易漏，不是最优。

为什么错：选 D 的同学把“至少 1 次”的反面算成了 0.3^4 —— 0.3^4 正是“一次都没中”的概率，而“至少 1 次”要的是它的补集 $1 - 0.3^4$ 。“至少一个”问题的万能钥匙：先算“一个都没有”，再用 1 减。

C2. (1) $C_{100}^2(0.04)^2(0.96)^{98}$; (2) 0.98; (3) 80%。

思路：对照组 100 人感染与否相互独立，每人感染概率 4% 恒定， $X \sim B(100, 0.04)$ 。

解析：(1) $P(X = 2) = C_{100}^2(0.04)^2(0.96)^{98}$ 。(2) 至少 1 人感染 $= 1 - P(X = 0) = 1 - 0.96^{100} \approx 1 - 0.0169 \approx 0.98$ 。(3) 有效率 $= \frac{0.04 - 0.008}{0.04} \times 100\% = \frac{0.032}{0.04} \times 100\% = 80\%$ 。

易错点：(2) 用补集， 0.96^{100} 表示 100 人全部不感染；疫苗有效率是“未接种感染率与接种感染率的差 $\div$ 未接种感染率”，分子是两者之差，别算成 $\frac{0.008}{0.04}$ （那是残留风险比例）。

2.9 例题详解

例题 2.7 解析

思考路径：命中次数 X 服从二项分布 $B(10, 0.2)$ ，直接套数字特征公式。

规范解：

$$E(X) = np = 10 \times 0.2 = 2, \quad D(X) = np(1 - p) = 10 \times 0.2 \times 0.8 = 1.6$$

答： $E(X) = 2$, $D(X) = 1.6$ 。

易错点：直觉校验——每次命中概率 0.2，射 10 次平均中 2 次，期望带 np 不带 $(1 - p)$ ；方差在 np 基础上再乘 $(1 - p)$ ，是 1.6 不是 2（见 M7）。

例题 2.8 解析

思考路径： $\mu = 10$, $\sigma = 0.2$ 。先把区间翻译成 $\mu \pm k\sigma$ 的形式，再用 3σ 原则与对称性。

规范解：(1) $9.8 = \mu - \sigma$, $10.2 = \mu + \sigma$ ：

$$P(9.8 < X < 10.2) = P(\mu - \sigma < X < \mu + \sigma) \approx 0.6827$$

(2) 曲线关于 $x = 10$ 对称，小于 μ 与大于 μ 的面积各占一半：

$$P(X < 10) = 0.5$$

答：(1) 约 0.6827; (2) 0.5。

易错点：正态分布记号 $N(10, 0.04)$ 中第二个参数是方差 σ^2 ，标准差 $\sigma = 0.2$ ——别把 0.04 当 σ 用。

2.10 知识点 (1) 期望与方差详解

A14. A。

思路： $X \sim B(10, 0.2)$ ，套 $E = np$ 、 $D = np(1 - p)$ 。

解析： $E = 10 \times 0.2 = 2$ ， $D = 10 \times 0.2 \times 0.8 = 1.6$ 。

A15. C。

思路：期望是取值按概率加权平均。

解析： $E(X) = 0 \times 0.2 + 1 \times 0.5 + 2 \times 0.3 = 0.5 + 0.6 = 1.1$ 。**易错点**：分布列三项概率和要等于 1（这里 $0.2 + 0.5 + 0.3 = 1$ ✓），加权时每一项是“取值 $\times$ 对应概率”。

B7. (2022 · 全国甲卷 · 第 19 题) (1) 0.6；(2) 分布列 X 取 0, 10, 20, 30，对应概率 0.16, 0.44, 0.34, 0.06， $E(X) = 13$ 。

思路：(1) 甲夺冠 = 甲至少胜 2 场，分“全胜”与“恰胜 2 场”两类，类内各场独立相乘、类间相加；(2) X 是乙学校总得分，乙在第 i 个项目获胜概率分别为 0.5, 0.6, 0.2，按 X 取值分类求分布列与期望。

解析：(1) 设甲在三个项目获胜概率为 0.5, 0.4, 0.8。全胜： $0.5 \times 0.4 \times 0.8 = 0.16$ ；恰胜 2 场（败第 1 场）： $0.5 \times 0.4 \times 0.8 = 0.16$ ；（败第 2 场）： $0.5 \times 0.6 \times 0.8 = 0.24$ ；（败第 3 场）： $0.5 \times 0.4 \times 0.2 = 0.04$ 。

$$P(\text{甲夺冠}) = 0.16 + 0.16 + 0.24 + 0.04 = 0.6$$

(2) 乙在三个项目获胜概率为 0.5, 0.6, 0.2， X 取 0, 10, 20, 30。

$$P(X = 0) = 0.5 \times 0.4 \times 0.8 = 0.16$$

$$P(X = 10) = 0.5 \times 0.4 \times 0.2 + 0.5 \times 0.6 \times 0.8 + 0.5 \times 0.4 \times 0.8 = 0.04 + 0.24 + 0.16 = 0.44$$

$$P(X = 20) = 0.5 \times 0.6 \times 0.8 + 0.5 \times 0.4 \times 0.2 + 0.5 \times 0.6 \times 0.2 = 0.24 + 0.04 + 0.06 = 0.34$$

$$P(X = 30) = 0.5 \times 0.6 \times 0.2 = 0.06$$

分布列为： X : 0 10 20 30 P : 0.16 0.44 0.34 0.06

$$E(X) = 0 \times 0.16 + 10 \times 0.44 + 20 \times 0.34 + 30 \times 0.06 = 4.4 + 6.8 + 1.8 = 13$$

为什么错：(1) 只算“全胜”漏掉“恰胜 2 场”三类，会得 0.16 而非 0.6——夺冠包含“三场赢两场”，不是“必须全赢”；(2) 写 $P(X = 10)$ 时，三场中哪一场是乙赢的必须逐一枚举，漏一种就错。**求分布列的关键动作**：把“取值对应的所有互斥情形”枚举全，类间用加法，每类内各场独立用乘法。

M7. B。

思路： $D(X) = np(1 - p)$ ，带上 $(1 - p)$ 。

解析： $D(X) = 10 \times 0.2 \times 0.8 = 1.6$ 。

为什么错：选 A 的同学把方差记成了 np —— $E(X) = np$ 不带 $(1-p)$ ， $D(X) = np(1-p)$ 带 $(1-p)$ 。**记法锚点：**期望是“平均成功次数”（直觉上 $10 \times 0.2 = 2$ ），方差是在此基础上再乘失败概率 $(1-p)$ 。也可用两点分布验证：一次试验成功时 $D = p(1-p)$ ， n 次独立试验方差相加得 $np(1-p)$ 。

2.11 知识点 (2) 正态分布与 3σ 原则详解

A16. A。

思路： $\mu = 170$ ， $\sigma = 5$ ， $[165, 175] = [\mu - \sigma, \mu + \sigma]$ 。

解析： $P \approx 0.6827$ 。**易错点：** $N(170, 25)$ 的 $\sigma = 5$ ，区间跨度是 $\pm\sigma$ 不是 $\pm 2\sigma$ ； $\pm 2\sigma$ 对应 $[160, 180]$ 。

B8. (2022 · 新高考 II 卷 · 第 13 题) 0.36。

思路：正态曲线关于 $\mu = 2$ 对称， $P(X > 2) = 0.5$ ； $P(X > 2.5)$ 比 $P(X > 2)$ 少了中间 $P(2 < X \leq 2.5)$ 这一段。

解析：

$$P(X > 2.5) = P(X > 2) - P(2 < X \leq 2.5) = 0.5 - 0.14 = 0.36$$

易错点：不能直接把 0.14 当答案——0.14 是 $[2, 2.5]$ 这段区间（右半侧靠近对称轴处）的概率；要求的是更右的尾部 $P(X > 2.5)$ ，需从右半侧总概率 0.5 中扣除这一段。

本章综合详解

S1 解析

思考路径：设 $A =$ “甲中”， $B =$ “乙中”， A, B 独立。三问分别用乘法公式、并集公式（或补集）、“恰一 = 甲中乙不中 + 甲不中乙中”。

规范解：(1) $P(A \cap B) = P(A)P(B) = 0.8 \times 0.7 = 0.56$ 。(2) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.8 + 0.7 - 0.56 = 0.94$ ，或用补集 $1 - 0.2 \times 0.3 = 1 - 0.06 = 0.94$ 。(3) $P(\text{恰一}) = 0.8 \times 0.3 + 0.2 \times 0.7 = 0.24 + 0.14 = 0.38$ 。（验证：都中 $0.56 +$ 恰一 $0.38 +$ 都不中 $0.06 = 1$ ，不重不漏。）

易错点：(2) 中 A, B 独立但不互斥， $P(A \cup B)$ 必须减去 $P(A \cap B)$ ——若误用互斥加法 $0.8 + 0.7 = 1.5$ ，概率超过 1，一眼假。

S2 解析

思考路径：(1) 组合计数；(2) 古典概型分子分母都数组组合数；(3) X 服从“不放回抽取”的超几何型分布，按女生人数分五种情况列分布列，再算期望。

规范解：(1) $C_{10}^4 = 210$ 种。(2) 恰 2 男 2 女：

$$P = \frac{C_6^2 C_4^2}{C_{10}^4} = \frac{15 \times 6}{210} = \frac{90}{210} = \frac{3}{7}$$

(3) X 可取 $0, 1, 2, 3, 4$:

$$P(X=0) = \frac{C_6^4}{C_{10}^4} = \frac{15}{210} = \frac{1}{14}, \quad P(X=1) = \frac{C_6^3 C_4^1}{C_{10}^4} = \frac{80}{210} = \frac{8}{21}$$

$$P(X=2) = \frac{C_6^2 C_4^2}{C_{10}^4} = \frac{90}{210} = \frac{3}{7}, \quad P(X=3) = \frac{C_6^1 C_4^3}{C_{10}^4} = \frac{24}{210} = \frac{4}{35}, \quad P(X=4) = \frac{C_4^4}{C_{10}^4} = \frac{1}{210}$$

(验证概率和: $\frac{15+80+90+24+1}{210} = \frac{210}{210} = 1$ 。)

$$E(X) = 0 \times \frac{1}{14} + 1 \times \frac{8}{21} + 2 \times \frac{3}{7} + 3 \times \frac{4}{35} + 4 \times \frac{1}{210} = \frac{80+180+72+4}{210} = \frac{336}{210} = \frac{8}{5} = 1.6$$

(对照超几何期望公式 $E = \frac{nM}{N} = \frac{4 \times 4}{10} = 1.6$, 一致。)

易错点: 第 (3) 问是不放回抽取, 不是二项分布——不能用 $B(4, 0.4)$; 分布列五项的概率和必须等于 1, 可用于自查; 期望通分到 210 再求和, 别跳步约分导致笔误。

第三章 统计

3.1 例题详解

例题 3.1 解析

思考路径：普查与抽样选谁，看三个触发条件：查了会不会坏、查完用不用得上、查全代价高不高。

规范解：(1) 抽样调查——5000 个逐一检测弹性费时费力，抽部分即可判断整批质量。(2) 普查——表彰名单本就齐全，直接统计又快又全。(3) 抽样调查——开瓶检测会破坏疫苗，只能抽样，用部分样本估计整批情况。

易错点：判断依据要落到“破坏性、代价、必要性”上，不要凭“数量大就一定抽样”的直觉，数量大但名单齐全时仍可普查。

例题 3.2 解析

思考路径：分层抽样各层抽取数 = 层人数占比 × 样本量，算完用 $\frac{n_i}{n} = \frac{N_i}{N}$ 反验。

规范解：总体 $N = 800 + 600 + 600 = 2000$ ，样本量 $n = 100$ 。(1) 高一： $\frac{800}{2000} \times 100 = 40$ (人)；高二： $\frac{600}{2000} \times 100 = 30$ (人)；高三： $\frac{600}{2000} \times 100 = 30$ (人)。检验： $40 + 30 + 30 = 100$ 。(2) $\frac{n_{高一}}{n} = \frac{40}{100} = 0.4 = \frac{800}{2000}$ ， $\frac{n_{高二}}{n} = \frac{30}{100} = 0.3 = \frac{600}{2000}$ ， $\frac{n_{高三}}{n} = 0.3 = \frac{600}{2000}$ 。

易错点：比例是“人数比”，不是“年级数比”——三层绝不是各抽三分之一。

3.2 知识点 (1) 普查与抽样调查详解

A1. C。

思路：普查适用条件——数量少、无破坏、名单齐全。

解析：灯泡寿命和种子发芽率检测有破坏性，全国中学生视力调查范围大、代价高，都应抽样；本班同学身高人数少、无破坏、直接统计即可，适合普查。

A2. B。

思路：抽样调查的核心优势。

解析：抽样调查省时省力，尤其适合破坏性检测（灯泡寿命、疫苗效期）与大范围调查（全国

人口), 结果带有一定误差但可用统计方法控制。A、C、D 说法均不成立。

3.3 知识点 (2) 简单随机抽样与系统抽样详解

A3. B。

思路: 简单随机抽样 = 等概率 + 随机决定选谁。

解析: A 是按固定间隔 (系统抽样); C 是分层抽样; D 是方便抽样 (无随机性); B 编号后随机抽取, 每个号码机会均等, 属于简单随机抽样。

A4. A。

思路: 间隔 $k = \frac{N}{n}$, 最后一个编号 = 首个体 + $(n-1)k$ 。

解析: $k = \frac{2000}{100} = 20$; 最后一个编号 = $5 + 99 \times 20 = 1985$ 。易错点: 共有 100 个个体, 间隔要乘 $100 - 1 = 99$, 不是 100。

3.4 知识点 (3) 分层抽样详解

A5. A。

思路: 分层抽样按人数比例抽取。

解析: 甲车间 $\frac{240}{600} \times 60 = 24$; 乙车间 $\frac{360}{600} \times 60 = 36$ 。易错点: 别把“两个车间”想成各抽 30——那样甲车间超代表。

M1. B。

思路: 分层抽样按人数比抽取, 先算每层占比再乘样本量。

解析: 高一 $\frac{5}{10} \times 120 = 60$; 高二 $\frac{3}{10} \times 120 = 36$; 高三 $\frac{2}{10} \times 120 = 24$ 。

为什么错: 选 A 的同学把“分层”理解成“平均分配”, 三层各 40——忽视了各层人数不同。选 C 的同学把比例 5:3:2 中的“份数”直接当人数 (50/30/20), 忘记了先换算成占总体比例。

分层抽样的铁律是 $\frac{n_i}{n} = \frac{N_i}{N}$ 。

M2. 不对。

规范解: 按人数比例抽取: 高一 $\frac{300}{600} \times 60 = 30$ (人); 高二 $\frac{200}{600} \times 60 = 20$ (人); 高三 $\frac{100}{600} \times 60 = 10$ (人)。

为什么错: 小明用“层数比” (三层各 $\frac{1}{3}$) 代替“人数比”, 结果人少的高三年级被“超代表” (抽 20 人 = 层内 20%, 而按比例应抽 10 人 = 层内 16.7%)。分层抽样的“按比例”三字, 指的是按各层人数占总体人数的比例, 不是“平均分配”。

C1. (1) 高一: $\frac{960}{2400} \times 120 = 48$ (人); 高二: $\frac{720}{2400} \times 120 = 36$ (人); 高三: $\frac{720}{2400} \times 120 = 36$ (人)。检验: $48 + 36 + 36 = 120$ 。(2) 住宿生: $\frac{1500}{2400} \times 120 = 75$ (人); 走读生: $\frac{900}{2400} \times 120 = 45$ (人)。检验: $75 + 45 = 120$ 。(3) 校门口拦截只覆盖“当时经过校门口”的学生: 时段选

择会系统性偏向（如放学时以走读生为主），未做到随机，样本代表性差。分层抽样按比例覆盖高一、高二、高三及住宿、走读各层，层内再随机抽取，兼顾层间差异，样本更有代表性。

3.5 例题详解

例题 3.3 解析

思考路径：频率 = 频数 ÷ 样本量；“面积和 = 1”即各频率相加为 1。

规范解：样本量 $n = 40$ ，组距为 5。(1) 各频率依次为 $\frac{4}{40} = 0.1$, $\frac{8}{40} = 0.2$, $\frac{12}{40} = 0.3$, $\frac{10}{40} = 0.25$, $\frac{6}{40} = 0.15$ 。(2) 不低于 55 kg 即 $[55, 60)$ 与 $[60, 65)$ 两组：频率 = $0.25 + 0.15 = 0.4$ ，人数 = $40 \times 0.4 = 16$ （人）。(3) 面积和 = $0.1 + 0.2 + 0.3 + 0.25 + 0.15 = 1$ 。

易错点：若用“高度相加”，各高度为 \$0.02, 0.04, 0.06, 0.05, 0.03\$，相加得 0.2，不等于 1——高度相加绝不是 1，面积相加才是 1。

例题 3.4 解析

思考路径：中位数看累积频率过 0.5 的区间；平均数用组中值加权。

规范解：累积频率依次为 \$0.1, 0.3, 0.6, 0.85, 1.0\$。(1) 0.5 落在 $[50, 55)$ ，中位数 $\approx 50 + \frac{0.5 - 0.3}{0.3} \times 5 = 50 + \frac{0.2}{0.3} \times 5 \approx 53.3$ (kg)。(2) 组中值分别为 \$42.5, 47.5, 52.5, 57.5, 62.5\$，

$$\bar{x} \approx 42.5 \times 0.1 + 47.5 \times 0.2 + 52.5 \times 0.3 + 57.5 \times 0.25 + 62.5 \times 0.15 = 4.25 + 9.5 + 15.75 + 14.375 + 9.375 = 53.25$$

(3) 频率最大的是 0.3，对应区间 $[50, 55)$ ，即众数区间。

易错点：中位数是用“面积”（频率）平分，不是用“区间个数”平分；平均数的权重是频率不是频数。

3.6 知识点 (1) 直方图的读图与频率详解

A6. B。

思路：频率 = 频数 ÷ 样本量。

解析：第二组频数 15， $15 \div 50 = 0.3$ 。

A7. C。

思路：直方图纵轴约定。

解析：纵轴是 $\frac{\text{频率}}{\text{组距}}$ ，面积 = 组距 × 高度 = 频率。

A8. B。

思路：频率和为 1。

解析： $1 - (0.15 + 0.25 + 0.35) = 0.25$ 。

M3. C。

思路：频率 = 组距 × 高度。

解析：第一组频率 = $10 \times 0.02 = 0.2$ ；第二组频率 = $10 \times 0.04 = 0.4$ ，是 0.2 的 2 倍。A、B 错（0.04、0.02 是高度不是频率）；D 错（两组频率之和是 $0.2 + 0.4 = 0.6$ ，不是 0.06）。

为什么错：选 A 或 B 的同学把“纵轴高度”当成了“频率”。条形图（横轴类别、纵轴数量）的惯性在这里失效：直方图的纵轴是 $\frac{\text{频率}}{\text{组距}}$ ，**高度 × 组距才是频率**。

M4. 不对。

规范解：第一个矩形频率 = $1 \times 0.3 = 0.3$ ；第二个矩形频率 = $4 \times 0.1 = 0.4$ 。频率反而是第二个更大。

为什么错：小华犯了“拿高度直接比频率”的错。**组距不一样时，高度不能直接比较；必须折算成面积（频率）再比。**这就是“面积才是频率”的含义。

3.7 知识点 (2) 由直方图估计统计量详解

A9. B。

思路：直方图中矩形面积即频率。

解析：面积 = 组距 × 高度 = 频率，题目已给出面积为 0.16，故频率为 0.16。**易错点：**纵轴高度才需要乘组距，面积本身就是频率，不必再乘组距。

A10. C。

思路：平均数 $\approx \sum(\text{组中值} \times \text{频率})$ 。

解析： $0.1 \times 10 + 0.2 \times 30 + 0.4 \times 50 + 0.3 \times 70 = 1 + 6 + 20 + 21 = 46$ 。

B1. C。

思路：先补全频数表（[1050, 1100) 的频数 = 100 减去其余五组），再逐项判断中位数、比例、极差、平均数。

解析：[1050, 1100) 频数 = $100 - (6 + 12 + 18 + 24 + 10) = 30$ 。A 错：前 3 组频数和 = $6 + 12 + 18 = 36 < 50$ ，故中位数落在 [1050, 1100)，不小于 1050 kg。B 错：低于 1100 kg 的稻田 = $6 + 12 + 18 + 30 = 66$ 块，占比 $66\% < 80\%$ 。C 对：极差最小为 $1150 - 950 = 200$ ，最大为 $1200 - 900 = 300$ ，介于 200 至 300 kg 之间。D 错：平均数 = $\frac{6 \times 925 + 12 \times 975 + 18 \times 1025 + 30 \times 1075 + 24 \times 1125 + 10 \times 1175}{100} = \frac{106700}{100} = 1067$ (kg)，超过 1000 kg。

易错点：B 项要算“低于 1100”而非“低于 1150”；D 项组中值别用区间左端点。

B2. (2022 · 新高考 II 卷 · 第 19 题) (1) 平均年龄 = $5 \times 0.001 \times 10 + 15 \times 0.002 \times 10 + 25 \times 0.012 \times 10 + 35 \times 0.017 \times 10 + 45 \times 0.023 \times 10 + 55 \times 0.020 \times 10 + 65 \times 0.017 \times 10 + 75 \times 0.006 \times 10 + 85 \times 0.002 \times 10 = 0.05 + 0.3 + 3 + 5.95 + 10.35 + 11 + 11.05 + 4.5 + 1.7 = 47.9$ (岁)。(2) 年龄位于 [20, 70) 的频率 = $(0.012 + 0.017 + 0.023 + 0.020 + 0.017) \times 10 = 0.089 \times 10 = 0.89$ ，故概率估计值为 0.89。(3) 设“选到的人年龄位于 [40, 50)”为事件 A，“选到的

人患这种疾病”为事件 B ，则 $P(A) = 0.16$ ， $P(B) = 0.001$ ， $P(A | B) = 0.023 \times 10 = 0.23$ 。由条件概率公式，

$$P(B | A) = \frac{P(A | B)P(B)}{P(A)} = \frac{0.23 \times 0.001}{0.16} = \frac{0.00023}{0.16} \approx 0.0014$$

易错点：(1) 用“组中值 $\times$ 频率”而不是“组中值 $\times$ 高度”——表中高度是 $\frac{\text{频率}}{\text{组距}}$ ，先乘组距 10 还原频率；(3) 第 (3) 问是条件概率的贝叶斯式求法，属第 6 章概率内容，可作为跨章综合训练。

为什么错：第 (1) 问最常见错误是漏乘组距 10 (把高度直接当频率)，导致平均值错成 4.79；第 (3) 问若只算 $P(A | B) \times P(B)$ 而不除以 $P(A)$ ，就漏了分母，这是跨章综合丢分的重灾区。记住：求“谁在什么条件下的概率”，先把已知的两个条件概率标清楚，再套条件概率公式。

C2. (1) 样本量 30。频率依次为 $\frac{3}{30} = 0.1$ ， $\frac{6}{30} = 0.2$ ， $\frac{9}{30} = 0.3$ ， $\frac{7}{30} \approx 0.23$ ， $\frac{5}{30} \approx 0.17$ 。面积和 $= 0.1 + 0.2 + 0.3 + 0.23 + 0.17 = 1$ 。(2) 累积频率 0.1, 0.3, 0.6, 0.5 落在 $[75, 100)$ ，中位数 $\approx 75 + \frac{0.5 - 0.3}{0.3} \times 25 = 75 + \frac{0.2}{0.3} \times 25 \approx 91.7 (\mu\text{g}/\text{m}^3)$ 。(3) 不低于 $100 \mu\text{g}/\text{m}^3$ 对应 $[100, 125)$ 、 $[125, 150)$ 两组，天数 $= 7 + 5 = 12$ (天)。由样本频率估计总体：该市 PM2.5 “轻度污染以上”的天数占比约 40%。

易错点：中位数是“面积平分”，不是“数据个数平分”；第 (3) 问直接由频数相加即可，不必再用频率折算。

3.8 例题详解

例题 3.5 解析

思考路径：先排序，再算 $i = \frac{p}{100} \times n$ ，按整数/非整数分情况定位。

规范解：排序得 51, 58, 61, 65, 66, 69, 72, 74, 77, 78, 80, 83, 84, 88, 90, 95 (1) 第 25 百分位数： $i = 0.25 \times 16 = 4$ ，恰为整数，取第 4 个与第 5 个数据的平均数： $\frac{65 + 66}{2} = 65.5$ 。(2) 第 60 百分位数： $i = 0.6 \times 16 = 9.6$ ，不是整数，取大于 9.6 的最小整数 10，即第 10 个数据，为 78。

易错点：不排序直接数必错； i 为整数时必须取相邻两个数的平均数。

例题 3.6 解析

思考路径：五数概括 = 最小值、 Q_1 、中位数、 Q_3 、最大值，全部由排序后的位置决定。

规范解：最小值 51，最大值 95。 $Q_1 = 65.5$ (见例题 3.5)。中位数： $i = 0.5 \times 16 = 8$ (整数)， $\frac{74 + 77}{2} = 75.5$ 。 Q_3 ： $i = 0.75 \times 16 = 12$ (整数)， $\frac{83 + 84}{2} = 83.5$ 。五数概括：51, 65.5, 75.5, 83.5, 95。箱子从 65.5 到 83.5，中位线在 75.5，箱子集中了约一半同学；95 明显远离 $Q_3 = 83.5$ ，可看作可能的离群点 (需结合考试背景判断)。

易错点： Q_1 、中位数、 Q_3 的位置计算要分别按 $i = 0.25n$ 、 $0.5n$ 、 $0.75n$ 判断整数性，不能只记一个。

3.9 知识点 (1) 百分位数详解

A11. C。

思路： $i = 0.8 \times 10 = 8$ 是整数，取第 8、9 个数据的平均数。

解析： $\frac{25+26}{2} = 25.5$ 。易错点： i 为整数要取平均，不是取第 8 个。

A12. B。

思路： $i = 0.25 \times 10 = 2.5$ 非整数，取大于 2.5 的最小整数 3，即第 3 个数据。

解析：第 3 个数据是 7。

A13. B。

思路： $i = 0.7 \times 15 = 10.5$ 非整数。

解析：取大于 10.5 的最小整数 11，即第 11 个数据。

M5. 不对。

规范解：先排序得 1, 2, 3, 4, 5, 6, 7, 8, 9, 10。 $i = 0.7 \times 10 = 7$ (整数)，取第 7 个与第 8 个数据的平均数： $\frac{7+8}{2} = 7.5$ 。

为什么错：小丽把“第 70 百分位数”理解成“第 7 个数据”，又没排序就直接数第 7 个 (6)。百分位数是“排序后处于 $p\%$ 位置的数”，排序是第一步；乱序数据不排序就定位，结果差之千里。

M6. 不对。

规范解： $i = 0.8 \times 5 = 4$ ，恰为整数，应取第 4 个与第 5 个数据的平均数： $\frac{20+25}{2} = 22.5$ 。

为什么错：该同学只背了“取第 i 个”，忽略了“ i 是整数时取相邻两个的平均”这一分支。百分位数算法分两支： i 非整数取第 $[i]$ 个， i 是整数取第 i 、 $i+1$ 个的平均——两支都要会。

3.10 知识点 (2) 四分位数与箱线图详解

A14. A。

思路： Q_1 即第 25 百分位数， $i = 0.25n$ 。

解析： $n = 7$ ， $i = 0.25 \times 7 = 1.75$ 非整数，取第 2 个数据，即 5。

A15. A。

思路：四分位距 $= Q_3 - Q_1$ 。

解析： $Q_3 - Q_1 = 30 - 20 = 10$ 。易错点：四分位距描述中间 50% 数据的跨度，用上四分位数减下四分位数，不是最大值减最小值。

3.11 例题详解

例题 3.7 解析

思考路径：平均数对极端值敏感，中位数抗极端值。先算两个指标，再判断“哪个指标服务哪个问题”。

规范解：(1) 平均数 $= \frac{10 \times 4000 + 48000}{11} = \frac{88000}{11} = 8000$ (元)。排序后第 6 个数据为 4000，中位数 = 4000 (元)。(2) 有误导性：8000 元是被 1 个 48000 元“拉高”的，11 人中有 10 人实际只有 4000 元，平均月薪严重偏离大多数人的现实。描述大多数员工收入应使用**中位数 4000 元**。

易错点：这是 M7 的落点——“平均数骗人”的机制是极端值把平均数往一边拽，中位数纹丝不动；谈“大多数”必须用中位数。

例题 3.8 解析

思考路径：先比平均（两人相同），再比方差（方差小 = 稳定）。

规范解：(1) $\bar{x}_{\text{甲}} = \frac{6+7+7+8+8+9}{6} = \frac{45}{6} = 7.5$ ； $\bar{x}_{\text{乙}} = \frac{3+6+7+8+9+12}{6} = \frac{45}{6} = 7.5$ 。(2) 甲的离均差 -1.5, -0.5, -0.5, 0.5, 0.5, 1.5, $s_{\text{甲}}^2 = \frac{2.25+0.25+0.25+0.25+0.25+2.25}{6} = \frac{5.5}{6} \approx 0.92$ 。乙的离均差 -4.5, -1.5, -0.5, 0.5, 1.5, 4.5, $s_{\text{乙}}^2 = \frac{20.25+2.25+0.25+0.25+2.25+20.25}{6} = \frac{45.5}{6} \approx 7.58$ 。 $s_{\text{甲}}^2 < s_{\text{乙}}^2$ ，甲的成绩更稳定。

易错点：平均数相同看方差，方差小 = 稳定；标准差与数据同单位，更贴近“平均偏离多远”的直觉。

3.12 知识点 (1) 平均数、中位数与众数详解

A16. B。

思路：平均数对极端值敏感，中位数抗极端值。

解析：原中位数 $\frac{3+4}{2} = 3.5$ ，平均数 = 4；加入 100 后中位数 $\frac{4+5}{2} = 4.5$ ，平均数 $= \frac{2+3+4+5+6+100}{6} = 20$ 。平均数变化剧烈，中位数基本不变。**易错点：**判断“哪个指标受影响大”看极端值是否参与平均运算——平均数每个数据都参与求和，中位数只看位置。

A17. B。

思路：偶数个数据中位数 = 中间两个的平均。

解析：第 10、11 个分别为 165、167，中位数 $= \frac{165+167}{2} = 166$ 。

B5. (2022 · 全国甲卷 (理) · 第 2 题) B。

思路：把散点图【文字化还原】为两组数据，分别排序后算中位数、平均数、标准差、极差，逐项对照。

解析：讲座前正确率排序：65%, 65%, 65%, 65%, 70%, 75%, 80%, 90%, 90%, 95%，中位数 =

$\frac{70\% + 75\%}{2} = 72.5\% > 70\%$, A 错。讲座后正确率: 90, 85, 80, 90, 85, 85, 95, 100, 85, 100 (%), 平均数 $= \frac{895}{10} = 89.5\% > 85\%$, B 对。讲座前数据明显更分散 (从 65% 到 95%), 标准差更大, C 说反了; 讲座前极差 $= 95\% - 65\% = 30\%$, 讲座后极差 $= 100\% - 80\% = 20\%$, D 说反了。答案 B。

易错点: (1) 中位数必须先排序再取中间; (2) 标准差、极差看“分散程度”可直观判断: 讲座后 10 个数据集中在 80% 100%, 明显比讲座前稳定; (3) 平均数按定义逐项求和, 别跳步。

M7. AB。

思路: 平均月薪远大于中位数, 是典型的右偏 (长尾在高收入端), 此时中位数代表“大多数”。

解析: A 对: 少数高收入把平均数拉高, 中位数更贴近大多数人的水平。B 对: 右偏数据有 $\bar{x} > M_e$ 。C 错: 两个指标描述不同侧面, 都不能直接当作每个人的收入。D 错: 平均数大小与数据质量无关。

为什么错: 选 C 的同学把“平均月薪”当作每个人的收入水平; 选 D 的同学把“平均数大”当作“好”。平均数与中位数的差恰恰是偏态的晴雨表——差得越远, 分布越偏, 越不能用平均数代表“大多数”。

3.13 知识点 (2) 方差与标准差详解

A18. A。

思路: $\bar{x} = 6$, 离均差 $-4, -2, 0, 2, 4$ 。

解析: $s^2 = \frac{16 + 4 + 0 + 4 + 16}{5} = \frac{40}{5} = 8$ 。

A19. B。

思路: 标准差是方差的算术平方根。

解析: $s = \sqrt{9} = 3$ 。

A20. C。

思路: 平均分看水平, 方差看稳定。

解析: 甲平均 80 大于乙平均 78 (水平更高); 乙方差 4 小于甲方差 16 (更稳定)。

A21. A。

思路: 所有数据加同一常数, 方差不变。

解析: 方差度量的是“离均差平方的平均”, 每个数据与平均数同步平移 +3, 离均差不变, 故方差仍为 4。**易错点:** 加常数方差不变, 乘常数方差变 k^2 倍——两条性质别混。

B3. (2023 · 新课标 I 卷 · 第 9 题) BD。

思路: 逐一考察平均数、中位数、标准差、极差, 判断“去掉最值后”各指标的变化方向。

解析: A 错: 如 1, 2, 2, 2, 2, 9 的平均数为 3, 而去掉两端后 2, 2, 2, 2 的平均数为 2, 不等。B

对: x_2, x_3, x_4, x_5 的中位数与 $x_1 \sim x_6$ 的中位数都是 $\frac{x_3 + x_4}{2}$ 。C 错: 去掉最小值和最大值后数据的波动性不大于原数据, 即标准差不大于原标准差, “不小于”说反了。D 对: $x_6 - x_1 \geq x_5 - x_2$, 即去掉两端后极差不大于原极差。

易错点: C 项方向——去掉极端值后数据更“集中”, 标准差应更小或相等, 不是更大。

B4. (2023·全国乙卷(理)·第17题) (1) $z_i = x_i - y_i$ 依次为 9, 6, 8, -8, 15, 11, 19, 18, 20, 12。
 $\bar{z} = \frac{9+6+8-8+15+11+19+18+20+12}{10} = \frac{110}{10} = 11$ 。离均差平方和 = $4 + 25 + 9 + 361 + 16 + 0 + 64 + 49 + 81 + 1 = 610$,

$$s^2 = \frac{610}{10} = 61$$

(2) $2\sqrt{\frac{s^2}{10}} = 2\sqrt{\frac{61}{10}} = 2\sqrt{6.1} \approx 2 \times 2.47 = 4.94$, 而 $\bar{z} = 11 > 4.94$, 所以认为甲工艺处理后的橡胶产品的伸缩率较乙工艺处理后的橡胶产品的伸缩率有显著提高。

易错点: z 可能为负 (第4号试验 -8), 先算 z_i 再求和, 别把 x_i, y_i 直接相加; 方差按 $\frac{1}{n}$ 口径 (本题给的是“样本方差”, 按题目给出的公式语境, $\frac{610}{10}$)。

为什么错: (1) 配对数据必须逐个做差 ($z_i = x_i - y_i$) 再统计, 很多同学把两组数据各自求平均再相减, 虽然本例 $\bar{x} - \bar{y}$ 也等于 11, 但遇到配对数据这种“先减后算”的路径是硬性规范, 写错中间步骤就要扣分; (2) 判断“显著提高”要严格走 $\bar{z} \geq 2\sqrt{\frac{s^2}{10}}$ 的比较, 别凭“平均数比 0 大就显著”的直觉——本题临界值约 4.94, $\bar{z} = 11$ 远超它, 结论才成立。

M8. ABC。

思路: 方差口径要分清“描述数据本身 ($\frac{1}{n}$)”与“用样本估计总体 ($\frac{1}{n-1}$)”。

解析: 离均差 -2, -1, 0, 1, 2, 平方和 = $4 + 1 + 0 + 1 + 4 = 10$ 。A: $\frac{10}{5} = 2$; B: $\frac{10}{4} = 2.5$; C: 高中描述一组数据统一用 $\frac{1}{n}$, 答案为 2; D 错, 两种口径用途不同, 不能随意换用。

为什么错: 选 D 的同学混淆了 n 与 $n-1$ 的适用场景。**描述这组数据本身用 $\frac{1}{n}$; 把这组数据当作样本去估计总体方差时用 $\frac{1}{n-1}$ (无偏修正)**。同一道题里先看题目问的是“这组数据的方差”还是“用样本估计总体方差”, 再决定分母。

本章综合详解

S1 解析

思考路径: 分层抽样求人数; 直方图“高度 $\times$ 组距 = 频率”求比例; 累积频率平分求中位数; 用样本频率估计总体。

规范解: (1) 总体 600 人, 样本量 60。男生 $\frac{360}{600} \times 60 = 36$ (人); 女生 $\frac{240}{600} \times 60 = 24$ (人)。(2) 睡眠时间低于 6 小时即 [4, 5)、[5, 6) 两组, 频率 = $0.02 \times 1 + 0.10 \times 1 = 0.12$, 即占比 12%。(3) 各区间频率依次为 0.02, 0.10, 0.28, 0.34, 0.20, 0.06, 累积频率 0.02, 0.12, 0.40, 0.74, 0.94, 1.00。0.5 落在 [7, 8), 中位数 $\approx 7 + \frac{0.5 - 0.40}{0.34} \times 1 = 7 + 0.29 \approx 7.3$ (小时)。(4) 睡眠充足频率 = $0.20 + 0.06 = 0.26$ 。估计全校高一睡眠充足人数 = $600 \times 0.26 = 156$ (人)。

易错点: (3) 中位数由“面积”定位, 分子是“还差多少面积” (0.5 - 0.40), 别写成组中值; (4)

用样本频率估计总体人数时乘的是总体人数 600，不是样本量 60。

S2 解析

思考路径：平均水平看平均数，稳定看方差；“水平高且稳定”要两个指标一起看。

规范解：(1) $\bar{x}_{\text{甲}} = \frac{85 + 90 + 92 + 88 + 95}{5} = \frac{450}{5} = 90$ ； $\bar{x}_{\text{乙}} = \frac{80 + 85 + 100 + 78 + 97}{5} = \frac{440}{5} = 88$ 。(2) 甲的离均差 $-5, 0, 2, -2, 5$ ， $s_{\text{甲}}^2 = \frac{25 + 0 + 4 + 4 + 25}{5} = \frac{58}{5} = 11.6$ 。乙的离均差 $-8, -3, 12, -10, 9$ ， $s_{\text{乙}}^2 = \frac{64 + 9 + 144 + 100 + 81}{5} = \frac{398}{5} = 79.6$ 。 $s_{\text{甲}}^2 < s_{\text{乙}}^2$ ，甲更稳定。(3) 推荐甲：平均分更高（90 大于 88）且方差更小（11.6 小于 79.6），整体水平高又稳定。

易错点：不要只看平均分——乙虽偶有高分（100），但方差大得多，稳定性差；评价“谁学得好”要“平均 + 方差”一起看。

第四章 成对数据的统计分析

4.1 例题详解

例题 4.1 解析

思考路径：函数关系看“是否完全确定”，相关关系看“有无趋势但不确定”，无关系看“互不影响”。

规范解：(1) 函数关系—— $I = \frac{U}{R}$ ，电阻一定时电流由电压完全确定。(2) 相关关系——身高高的人体重通常也重（趋势），但同一身高对应不同体重（不确定）。(3) 相关关系——气温越低供暖费越高（负相关趋势），但同温下不同住户费用不同。(4) 无关系——降雨量与彩票开奖号码互不影响。

易错点：(3) 别因为“反方向”就说没关系——负相关也是相关，方向相反而已。

例题 4.2 解析

思考路径：读散点图看方向和密疏；判断因果要看是否有共同原因或需要对照实验。

规范解：(1) 点云从左下到右上 → **正相关**；点较密集 → 线性相关**较强**（ r 应较大）。(2) 不能。相关不等于因果：锻炼时长与体质成绩可能同受“作息规律、重视健康、自律性强”等共同原因影响。要确认“锻炼提高成绩”的因果关系，应设计**随机对照实验**（随机分组、控制变量、比较前后差异），仅凭观察数据不能断言因果。

易错点：这是 M1 的核心——看到 r 大就喊“因为所以”，是本章最典型的错误。

4.2 知识点 (1) 相关关系与函数关系详解

A1. C。

思路：相关关系 = 有趋势但不确定。

解析：A 是函数关系 ($S = \pi r^2$)；B 是函数关系 ($I = \frac{U}{R}$)；D 是无关系；C 身高高体重通常重，但同一身高对应不同体重，是相关关系。

A2. B。

思路：函数关系 = 一个变量完全确定另一个变量。

解析：圆的半径与面积满足 $S = \pi r^2$ ，一一对应，是函数关系。A、C 是相关关系（有趋势但

不确定); D 是无关系。

4.3 知识点 (2) 散点图与“相关不等于因果”详解

A3. A。

思路：方向看斜向，强度看密疏。

解析：从左下到右上 → 正相关；点密集 → 线性相关较强。

A4. B。

思路：方向看斜向，强度看密疏。

解析：从左上到右下 → 负相关；点较分散 → 线性相关较弱。

M1. B。

思路：相关强只说明“同步变化”，不说明“谁导致谁”。

解析： $r = 0.88$ 是较强的正相关。A、C 都是因果断言，没有依据；B 正确；D 显然错。

为什么错：选 A 或 C 的同学把“相关”读成了“因果”。阅读量与成绩可能同受“学习习惯好、家长重视”等共同原因影响。记住：**相关系数只度量同步变化的方向与强度，因果关系需要对照实验或更严格的证据。**

C1. (1) $r \approx -0.82 < 0$ ，负相关； $|r| = 0.82 \geq 0.75$ ，线性相关较强。(2) 不合理。冰淇淋消费量与溺水人数负相关，但**相关不等于因果**：背后是共同的季节因素——夏天冰淇淋消费量大、下水游泳的人多溺水也多，冬天相反，两者同步反向变化。报道把相关当因果，犯了 M1 的错误。(3) 应做**对照实验或控制变量的研究**：固定游泳人数、气温等混杂因素，比较“冰淇淋消费量不同”人群的溺水率；或在实验条件下随机分组控制冰淇淋摄入。仅凭观察数据的相关关系，不能确认因果关系。

4.4 知识点 (3) 相关系数详解

A5. B。

思路：强度比绝对值，方向看符号。

解析： $|-0.8| = 0.8 > |0.6| = 0.6$ ，乙组线性相关更强。**易错点**：比较强度看 $|r|$ ，别看到负数就以为“弱”。

A6. B。

思路：强度比绝对值。

解析： $|-0.95| = 0.95 > |0.9| = 0.9$ ，乙组线性相关更强。

A7. B。

思路： r 只刻画线性相关。

解析： $r \approx 0$ 只说明没有**线性**相关关系，两个变量仍可能有曲线相关。**易错点：**别把“线性无关”升级成“完全无关”。

M2. B。

思路： r 只刻画线性相关。

解析： $y = (x - 3)^2$ 是抛物线关系，用直线去拟合会得到 $r \approx 0$ ，但两个变量明明有确定的曲线关系。

为什么错：选 A 的同学把“线性无关”升级成“完全无关”。 $|r|$ 接近 0 只能说“没有线性相关”，**不能排除曲线相关**——这就是为什么任何回归计算前必须先画散点图看形状。

B4. (2022 · 全国乙卷 (理) · 第 19 题) (1) 样本中 10 棵树木根部横截面积的平均值 $\bar{x} = \frac{0.6}{10} = 0.06$ (m^2)，材积量的平均值 $\bar{y} = \frac{3.9}{10} = 0.39$ (m^3)。据此估计该林区这种树木平均一棵的根部横截面积为 0.06 m^2 ，平均一棵的材积量为 0.39 m^3 。(2) 先求偏差平方和与协变差：

$$\sum_{i=1}^{10} (x_i - \bar{x})^2 = \sum x_i^2 - 10\bar{x}^2 = 0.038 - 10 \times 0.06^2 = 0.038 - 0.036 = 0.002$$

$$\sum_{i=1}^{10} (y_i - \bar{y})^2 = \sum y_i^2 - 10\bar{y}^2 = 1.6158 - 10 \times 0.39^2 = 1.6158 - 1.521 = 0.0948$$

$$\sum_{i=1}^{10} (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - 10\bar{x}\bar{y} = 0.2474 - 10 \times 0.06 \times 0.39 = 0.2474 - 0.234 = 0.0134$$

$$r = \frac{0.0134}{\sqrt{0.002 \times 0.0948}} = \frac{0.0134}{\sqrt{0.0001896}} \approx \frac{0.0134}{0.01377} \approx 0.973 \approx 0.97$$

(3) 由材积量与其根部横截面积近似成正比， $\frac{\bar{y}}{\bar{x}} = \frac{0.39}{0.06} = 6.5$ ，故总材积量估计值 $Y \approx 6.5 \times 186 = 1209$ (m^3)。

易错点：(1) 相关系数用“和与平方和”的简算公式，别逐对硬算 10 个离均差；(2) $\sum x_i^2 = 0.038$ 是平方和不是 $\sum x_i$ 的平方；(3) 第 (3) 问是“正比”估计：用样本均值比当比例常数，再乘总体根部横截面积总和。

为什么错：第 (2) 问最常见错误是把 $\sum (x_i - \bar{x})(y_i - \bar{y})$ 算成 $\sum x_i y_i$ (漏掉减 $n\bar{x}\bar{y}$ 这一项)， r 会算成 1.80 这种大于 1 的“错误值”还不自知—— r **必然落在 $[-1, 1]$ 内**，算出的 r **超过 1 立刻回头检查公式**。第 (3) 问若忘了“正比”而试图用第 (2) 问的 r 直接外推，就绕远了。

4.5 例题详解

例题 4.3 解析

思考路径：五步——算 $\bar{x}, \bar{y} \rightarrow$ 列偏差 $\rightarrow$ 求 $S_{xy}, S_{xx} \rightarrow$ 算 $b, a \rightarrow$ 用样本中心检验。

规范解： $\bar{x} = \frac{20 + 25 + 30 + 35 + 40}{5} = 30$ ， $\bar{y} = \frac{60 + 54 + 50 + 45 + 41}{5} = 50$ 。

$x_i - \bar{x}$	-10	-5	0	5	10
$y_i - \bar{y}$	10	4	0	-5	-9
乘积	-100	-20	0	-25	-90
$(x_i - \bar{x})^2$	100	25	0	25	100

$$S_{xy} = -100 - 20 + 0 - 25 - 90 = -235, \quad S_{xx} = 100 + 25 + 0 + 25 + 100 = 250,$$

$$b = \frac{-235}{250} = -0.94, \quad a = 50 - (-0.94) \times 30 = 50 + 28.2 = 78.2$$

回归方程 $\hat{y} = -0.94x + 78.2$ 。检验：把中心 $(30, 50)$ 代入， $-0.94 \times 30 + 78.2 = -28.2 + 78.2 = 50$ ，直线过样本中心。(2) $x = 32$: $\hat{y} = -0.94 \times 32 + 78.2 = -30.08 + 78.2 = 48.12 \approx 48.1$ (分钟)。

易错点： $b < 0$ 说明温度升高反应时间减少 (负相关)；斜率 b 的量纲是“分钟/摄氏度”，若算成正的就要检查公式。

例题 4.4 解析

思考路径：预测值 = 代入回归方程；残差 = 观测值 - 预测值； $|e_i|$ 最大者拟合最差； x 超出样本范围即外推。

规范解：(1) 各点预测值依次为 \$59.4, 54.7, 50, 45.3, 40.6\$ (如 $x = 20$ 时 $-0.94 \times 20 + 78.2 = 59.4$)，残差为

$$60 - 59.4 = 0.6, \quad 54 - 54.7 = -0.7, \quad 50 - 50 = 0, \quad 45 - 45.3 = -0.3, \quad 41 - 40.6 = 0.4$$

(2) $|e_i|$ 分别为 \$0.6, 0.7, 0, 0.3, 0.4\$，最大的是 0.7，即 $x = 25$ 这个点拟合最差 (也仅偏 0.7 分钟，整体拟合不错)。(3) $x = 60$ 已超出样本范围 $x \in [20, 40]$ ，属外推，没有数据支撑，不合理。

易错点：残差是“纵向距离”不是“错误”；拟合差的点可能是真实特殊样本，不能一删了之。外推是统计里最危险的动作。

4.6 知识点 (1) 回归直线与最小二乘法详解

A8. C。

思路：回归直线核心性质。

解析：回归直线必过样本中心 $(\bar{x}, \bar{y})$ 。

A9. A。

思路： $a = \bar{y} - b\bar{x}$ 。

解析： $a = 9 - 1.5 \times 4 = 3$ ，回归方程 $\hat{y} = 1.5x + 3$ 。**易错点：**别把 a 当 $\bar{y}$ ，要减去 $b\bar{x}$ 。

A10. A。

思路： $a = \bar{y} - b\bar{x}$ 。

解析: $a = 20 - 2 \times 6 = 8$, 回归方程 $\hat{y} = 2x + 8$ 。

M3. C。

思路: 回归直线唯一保证经过的是样本中心。

解析: $\hat{y} = bx + a$, 代入 $x = \bar{x}$ 得 $\hat{y} = b\bar{x} + \bar{y} - b\bar{x} = \bar{y}$, 故必过 $(\bar{x}, \bar{y})$ 。

为什么错: 选 A 的同学把初中学的“正比例函数 $y = kx$ 过原点”的印象搬过来了——回归直线带截距 a , 一般不过原点; 选 B、D 的同学以为回归线要“穿过数据点”——最小二乘线一般不过任何数据点, 它的使命是让所有点离它最近, 而不是穿过它们。

B1. (2023 · 天津卷 · 第 7 题) C。

思路: 相关系数定方向与强弱; 回归方程用于预测; 相关系数是样本指标, 换样本会变。

解析: $r = 0.8642 > 0$ 且 $|r|$ 较大 $\rightarrow$ 正相关较强, A、B 错。C: $\hat{y} = 0.7501 \times 7 + 0.6105 = 5.2507 + 0.6105 = 5.8612$ (cm), C 对。D: 样本不同相关系数一般不同, “一定为 0.8642” 错。

易错点: C 项“预测值”是平均值意义上的估计, 不是某个具体花朵的花瓣长度——这是回归预测的标准读法。

C2. (1) $\bar{x} = 4$, $\bar{y} = \frac{60 + 65 + 72 + 78 + 85}{5} = 72$ 。 $S_{xy} = (-2)(-12) + (-1)(-7) + 0 + 1 \times 6 + 2 \times 13 = 24 + 7 + 0 + 6 + 26 = 63$; $S_{xx} = 4 + 1 + 0 + 1 + 4 = 10$ 。 $b = \frac{63}{10} = 6.3$, $a = 72 - 6.3 \times 4 = 72 - 25.2 = 46.8$ 。回归方程 $\hat{y} = 6.3x + 46.8$ 。检验: (4, 72) 处 $6.3 \times 4 + 46.8 = 25.2 + 46.8 = 72$ 。(2) $x = 7$: $\hat{y} = 6.3 \times 7 + 46.8 = 44.1 + 46.8 = 90.9 \approx 91$ (分)。(3) $x = 6.5$: $\hat{y} = 6.3 \times 6.5 + 46.8 = 40.95 + 46.8 = 87.75$, 残差 $e = 80 - 87.75 = -7.75$ 。 $x = 12$ 超出样本范围 $[2, 6]$, 属于外推, 不能用该方程预测。

易错点: 回归方程只在样本范围内做内插预测才可靠; 残差为负说明回归线“高估”了该同学的成绩。

4.7 知识点 (2) 残差与预测 (内插与外推) 详解

A11. B。

思路: $x = 10$ 超出样本范围 $[1, 6]$ 。

解析: 外推无数据支撑, 预测不可靠。易错点: 方程“能算”不代表“能信”, 只在有数据的地方说话。

A12. A。

思路: 残差 = 观测值 - 预测值。

解析: $e_i = y_i - \hat{y}_i = 10 - 8.5 = 1.5$, 该点在回归线上方。

A13. B。

思路: 残差的定义与含义。

解析: 残差 $e_i = y_i - \hat{y}_i$ 是纵向距离, 可正可负; $|e_i|$ 越大拟合越差。A 错 (残差越大拟合越

差); C 错 (可正可负); D 错 (残差为零说明该点恰在直线上, 很常见)。

M4. B。

思路: $x = 100$ 超出 $[40, 60]$, 是外推。

解析: 回归方程只在样本 x 取值范围内预测可靠; $x = 100$ 处没有数据支撑, 预测不可靠。

为什么错: 选 A 的同学觉得“方程是万能的”。外推的荒谬常常很隐蔽: 样本外的规律可能完全改变 (比如 y 随 x 先增后减), 直线继续外延就会离谱。**内插可靠, 外推危险。**

4.8 例题详解

例题 4.5 解析

思考路径: 设“参加锻炼且未达标”为 a , “不锻炼且达标”为 $a + 10$, 用总人数方程解出 a , 再补全合计。

规范解: 设“参加锻炼且未达标”人数为 a , 则“不锻炼且达标”人数为 $a + 10$ 。总人数: $50 + a + (a + 10) + 20 = 120$, 解得 $a = 20$ 。故列联表为:

	达标	未达标	合计
参加锻炼	50	20	70
不锻炼	30	20	50
合计	80	40	120

三线校验: 行和 70, 50, 列和 80, 40, 总和 120。(2) 期望频数: $E(\text{参加, 达标}) = \frac{70 \times 80}{120} = \frac{140}{3} \approx 46.7$; $E(\text{参加, 未达标}) = \frac{70 \times 40}{120} = \frac{70}{3} \approx 23.3$; $E(\text{不锻炼, 达标}) = \frac{50 \times 80}{120} = \frac{100}{3} \approx 33.3$; $E(\text{不锻炼, 未达标}) = \frac{50 \times 40}{120} = \frac{50}{3} \approx 16.7$ 。

易错点: 期望频数可以是小数, 不用取整; 四格期望值也要满足行列合计对应 ($46.7 + 23.3 = 70$ 等)。

例题 4.6 解析

思考路径: 先求期望, 再逐格求 $\frac{(O - E)^2}{E}$ 求和, 最后与临界值比较。

规范解: (1) $E(\text{做操, 近视}) = \frac{120 \times 110}{200} = 66$; $E(\text{做操, 不近视}) = \frac{120 \times 90}{200} = 54$;
 $E(\text{不做, 近视}) = \frac{80 \times 110}{200} = 44$; $E(\text{不做, 不近视}) = \frac{80 \times 90}{200} = 36$ 。(2)

$$\chi^2 = \frac{(50 - 66)^2}{66} + \frac{(70 - 54)^2}{54} + \frac{(60 - 44)^2}{44} + \frac{(20 - 36)^2}{36} = \frac{256}{66} + \frac{256}{54} + \frac{256}{44} + \frac{256}{36} \approx 3.88 + 4.74 + 5.82 + 7.11 = 21.55$$

(3) $\chi^2 \approx 21.55 > 6.635 > 3.841$ 。与 $\alpha = 0.05$ 比: 拒绝原假设; 与 $\alpha = 0.01$ 比: 仍拒绝。结论: 在犯错误的概率不超过 0.01 的前提下, 认为“课间做眼保健操与近视状况有关联”——不做眼保健操的学生中近视比例明显更高。

易错点：比较方向是“大于临界值才拒绝”，不是小于； χ^2 大说明现实偏离“无关”假设越远，越敢说有关系。

4.9 知识点 (1) 列联表与期望频数详解

A14. B。

思路： $E = \frac{\text{行合计} \times \text{列合计}}{\text{总人数}}$ 。

解析： $\frac{60 \times 40}{200} = 12$ 。

A15. A。

思路：逐格公式 $\frac{(O - E)^2}{E}$ 。

解析： $\frac{(18 - 12)^2}{12} = \frac{36}{12} = 3$ 。

A16. B。

思路：期望频数的理论性。

解析：期望频数是“两变量无关”假设下的理论值，可以是小数；它不由观测数据直接读出，而是由行列合计推算。A、C、D 均错。

M5. ABC。

思路：期望频数公式 + 列联表三线校验。

解析：A 对： $E = \frac{\text{行合计} \times \text{列合计}}{\text{总人数}}$ 。B 对：期望频数就是“两变量无关”假设下应当出现的频数（乘法原理）。C 对：行合计、列合计、总合计三线必须对应。D 错：期望频数可为小数。

为什么错：选 D 的同学以为“人数必须是整数，期望也必须是整数”——期望频数是**理论值**，允许是小数（如 $E = 23.3$ ）。列联表里的**观测频数**必须是整数，期望频数不是。

4.10 知识点 (2) 卡方统计量与临界值详解

A17. B。

思路：原假设是“无关”， χ^2 大拒绝。

解析：A 错，原假设是“两变量无关”；B 对；C 错，“没有充分证据”不等于“证明无关”；D 错， χ^2 越小证据越弱。

A18. C。

思路： $\chi^2 = 2.6 < 3.841$ ，不拒绝原假设。

解析：规范说法是“没有充分证据认为两变量有关联”，不能写“无关”。

A19. (1) 列联表：

	感染	未感染	合计
接种	15	135	150
对照组	25	75	100
合计	40	210	250

(2) 期望频数: $E(\text{接种, 感染}) = \frac{150 \times 40}{250} = 24$; $E(\text{接种, 未感染}) = \frac{150 \times 210}{250} = 126$;
 $E(\text{对照, 感染}) = \frac{100 \times 40}{250} = 16$; $E(\text{对照, 未感染}) = \frac{100 \times 210}{250} = 84$ 。

$$\chi^2 = \frac{81}{24} + \frac{81}{126} + \frac{81}{16} + \frac{81}{84} \approx 3.375 + 0.643 + 5.063 + 0.964 = 10.04$$

(3) $10.04 > 3.841$, 拒绝原假设: 在犯错误的概率不超过 0.05 的前提下, 认为接种疫苗与是否感染有关——对照组感染比例 (25%) 明显高于接种组 (10%)。

易错点: 别把“接种组 150 人中 15 人”与“对照组 100 人中 25 人”的比例直接相减当结论, 必须走完列联表 + χ^2 + 临界值比较的完整流程。

A20. A。

思路: $\chi^2 = \sum \frac{(O - E)^2}{E}$ 。

解析: 每格期望 $E = \frac{50 \times 50}{100} = 25$, 四格相同, 故

$$\chi^2 = 4 \times \frac{(30 - 25)^2}{25} = 4 \times \frac{25}{25} = 4$$

B2. (2022 · 新高考 I 卷 · 第 20 题) (1) 统计量 (原卷记作 K^2 , 与 χ^2 等价):

$$K^2 = \frac{200 \times (40 \times 90 - 10 \times 60)^2}{100 \times 100 \times 50 \times 150} = \frac{200 \times 3000^2}{75000000} = 24$$

$24 > 6.635$, 所以有 99% 的把握认为患该疾病群体与未患该疾病群体的卫生习惯有差异。(2)

(i) 用 a, b, c, d 分别表示“不够良好”与“良好”、“病例”与“对照”四格频数, 即 $a = 40, b = 60, c = 10, d = 90$ 。由条件概率定义:

$$R = \frac{P(B | A)}{P(\bar{B} | A)} \cdot \frac{P(\bar{B} | \bar{A})}{P(B | \bar{A})} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}$$

而

$$\frac{P(A | B)}{P(\bar{A} | B)} \cdot \frac{P(\bar{A} | \bar{B})}{P(A | \bar{B})} = \frac{\frac{a}{a+c}}{\frac{c}{a+c}} \cdot \frac{\frac{d}{b+d}}{\frac{b}{b+d}} = \frac{a}{c} \cdot \frac{d}{b} = \frac{ad}{bc} = R$$

故 $R = \frac{P(A | B)}{P(\bar{A} | B)} \cdot \frac{P(\bar{A} | \bar{B})}{P(A | \bar{B})}$ 成立。(ii) $P(A | B) = \frac{40}{100} = 0.4$; $P(A | \bar{B}) = \frac{10}{100} = 0.1$ 。由 (i), $R = \frac{ad}{bc} = \frac{40 \times 90}{60 \times 10} = \frac{3600}{600} = 6$ 。即卫生习惯不够良好的人患该疾病的风险约为卫生习惯良好者的 6 倍。

易错点: (1) 分母各合计别抄错, $40 \times 90 - 60 \times 10 = 3000$ 别算成别的; (ii) $P(A | B)$ 是“病例组中不够良好的比例” ($40/100$), 不是全体中的比例。

为什么错：第 (2) 问的 R 定义式含三个条件概率，很多同学被“花括号”吓住直接放弃。其实把每一对 $\frac{P(\cdot | A)}{P(\cdot | \bar{A})}$ 用列联表频数代入，每一项都约成 $\frac{a}{b}$ 、 $\frac{d}{c}$ 之类的纯频数比，最后只需算 $\frac{ad}{bc}$ 。**记法：比值 R 就是“四格交叉相乘再相除”**——先约分再看形式，别硬背。

B3. (2023 · 全国甲卷 (理) · 第 19 题) (1) (超几何分布，可暂缓)： X 的可能取值为 0, 1, 2,

$$P(X=0) = \frac{C_{20}^2}{C_{40}^2} = \frac{190}{780} = \frac{19}{78}, \quad P(X=1) = \frac{C_{20}^1 C_{20}^1}{C_{40}^2} = \frac{400}{780} = \frac{40}{78}, \quad P(X=2) = \frac{190}{780} = \frac{19}{78}$$

$E(X) = 0 \times \frac{19}{78} + 1 \times \frac{40}{78} + 2 \times \frac{19}{78} = 1$ 。(2) (i) 将 40 只小鼠体重混合排序，第 20、21 个数分别为 23.2、23.6，故

$$m = \frac{23.2 + 23.6}{2} = 23.4$$

对照组体重不小于 23.4 g 的有 14 只 (24.6 至 28.3)，小于 23.4 的有 6 只；实验组体重不小于 23.4 的有 6 只 (23.6 至 26.0)，小于 23.4 的有 14 只。列联表为：

	体重小于 m	体重不小于 m	合计
对照组	6	14	20
实验组	14	6	20
合计	20	20	40

(ii) $\chi^2 = \frac{40 \times (6 \times 6 - 14 \times 14)^2}{20 \times 20 \times 20 \times 20} = \frac{40 \times 25600}{160000} = 6.4$ 。6.4 > 3.841，所以在犯错误的概率不超过 0.05 的前提下，认为药物对小鼠生长有抑制作用——实验组小鼠体重普遍偏小，与药物抑制生长一致。

易错点：中位数要先把 40 个数据合起来排序再取第 20、21 个的平均（两组内部已排序，但 23.2 在第 20 位、23.6 在第 21 位，需仔细数）； χ^2 用列联表公式或四格公式结果一致。

为什么错：(1) 中位数排序时两组数据是“各自从小到大排好”的，混排后 23.2 和 23.6 恰好在正中间——很多同学误把对照组中位数当总中位数，导致列联表全错；(2) χ^2 公式里 $(6 \times 6 - 14 \times 14)$ 的“交叉乘再相减”别漏括号、别写反顺序，算完要与临界值比较而不是直接下结论。

B5. (2022 · 全国甲卷 (文) · 第 17 题) (1) A 公司共调查 260 个班次，准点 240 次，准点概率估计值 $= \frac{240}{260} = \frac{12}{13}$ ；B 公司共调查 240 个班次，准点 210 次，准点概率估计值 $= \frac{210}{240} = \frac{7}{8}$ 。(2) $a = 240, b = 20, c = 210, d = 30, n = 500$,

$$K^2 = \frac{500 \times (240 \times 30 - 210 \times 20)^2}{260 \times 240 \times 450 \times 50} = \frac{500 \times 3000^2}{1404000000} \approx 3.205 > 2.706$$

3.205 > 2.706，所以有 90% 的把握认为甲、乙两城之间的长途客车是否准点与客车所属公司有关。

易错点：(1) 附公式中 a, b, c, d 对应四格频数 ($a = 240, b = 20, c = 210, d = 30$)，行列合计分别为 $a + b = 260$ 、 $c + d = 240$ 、 $a + c = 450$ 、 $b + d = 50$ ，抄错一个合计就全错；(2) 结论要用规范句式“有 90% 的把握认为……有关”。

M6. B。

思路： χ^2 大 $\rightarrow$ 证据强 $\rightarrow$ 拒绝“无关”。

解析： $5.02 > 3.841$ ，拒绝原假设，在犯错误的概率不超过 0.05 的前提下认为两变量有关联。

为什么错：选 C 的同学把方向记反了—— χ^2 越大，现实越偏离“无关”假设，越该拒绝它（原假设是“无关”）； χ^2 大不是“原假设成立”，恰恰相反。一句话记忆：卡方越大，越敢说有关系。

M7. B。

思路： χ^2 小 = 证据不足，不等于证明无关。

解析： $2.1 < 3.841$ ，不拒绝原假设，规范说法是“没有充分证据认为两变量有关联”。

为什么错：选 A、C 的同学把“没有充分证据”升级成“证明无关”。“没找到外星人”不等于“证明没有外星人”——证据的力度要跟结论的力度匹配。D 也是臆测，样本量是否不够要看设计，不能由此断言。

本章综合详解

S1 解析

思考路径：求中心点 $\rightarrow$ 列偏差算 $S_{xy}, S_{xx} \rightarrow$ 得回归方程 $\rightarrow$ 内插预测 $\rightarrow$ 算残差。

规范解：(1) $\bar{x} = \frac{34 + 38 + 42 + 46 + 50 + 54 + 58 + 62}{8} = \frac{384}{8} = 48$; $\bar{y} = \frac{68 + 74 + 80 + 86 + 92 + 98 + 104 + 110}{8} = \frac{712}{8} = 89$ 。(2) 偏差表： $x_i - \bar{x}$ 为 $-14, -10, -6, -2, 2, 6, 10, 14$; $y_i - \bar{y}$ 为 $-21, -15, -9, -3, 3, 9, 15, 21$ 。乘积之和 $S_{xy} = 294 + 150 + 54 + 6 + 6 + 54 + 150 + 294 = 1008$; $S_{xx} = 196 + 100 + 36 + 4 + 4 + 36 + 100 + 196 = 672$ 。

$$b = \frac{1008}{672} = 1.5, \quad a = 89 - 1.5 \times 48 = 17$$

回归方程 $\hat{y} = 1.5x + 17$ 。检验： $1.5 \times 48 + 17 = 89$ 。(3) $x = 47$: $\hat{y} = 1.5 \times 47 + 17 = 70.5 + 17 = 87.5$ (分)。(4) $x = 60$: $\hat{y} = 1.5 \times 60 + 17 = 107$ ，残差 $e = 96 - 107 = -11$ ，即回归线高估了该同学成绩 11 分。

易错点：第 (4) 问样本点 (60, 96) 处的“观测值”是 96，要代入回归方程求预测值 107 再作差，别把 96 当预测值。

S2 解析

思考路径：期望频数 $\rightarrow \chi^2 \rightarrow$ 与两个临界值比较 $\rightarrow$ 规范结论。

规范解：(1) $E(\text{自带, 支持}) = \frac{100 \times 120}{200} = 60$; $E(\text{自带, 反对}) = \frac{100 \times 80}{200} = 40$;
 $E(\text{不自带, 支持}) = 60$; $E(\text{不自带, 反对}) = 40$ 。(2)

$$\chi^2 = \frac{(70 - 60)^2}{60} + \frac{(30 - 40)^2}{40} + \frac{(50 - 60)^2}{60} + \frac{(50 - 40)^2}{40} = \frac{100}{60} + \frac{100}{40} + \frac{100}{60} + \frac{100}{40} = \frac{5}{3} + 2.5 + \frac{5}{3} + 2.5 = 8.33$$

(3) $8.33 > 6.635 > 3.841$ ：在 $\alpha = 0.05$ 下拒绝原假设；在 $\alpha = 0.01$ 下仍拒绝。结论：在犯错误的概率不超过 0.01 的前提下，认为“购物袋使用习惯与对限塑令的态度有关联”——自带购物袋的顾客中支持限塑令的比例 (70%) 明显高于不自带者 (50%)。

易错点：结论要用规范句式；比较时 χ^2 要同时与两个临界值比，显著性水平越严（ α 越小）临界值越大，同一 χ^2 可能在不同水平下结论不同——本题 8.33 足够大，两个水平下都拒绝。

第五章 数学建模与拓展

5.1 例题详解

例题 5.1 解析

思考路径：先看 T 的差，再看“高出室温的部分”的比——差递减、比恒定，就是指数模型。

规范解：(1) T 的差为 $-16, -12, -9$ ，在递减，不是线性。看 $T - 22$ ：64, 48, 36, 27，相邻比 $\frac{48}{64} = \frac{36}{48} = \frac{27}{36} = 0.75$ ，比恒定，故用**指数模型**。(2) 设 $T - 22 = c \cdot a^t$ 。 $t = 0$ 时 $T - 22 = 64$ ，得 $c = 64$ ；每 3 分钟乘一次 0.75，即 t 增加 3 时乘 0.75，所以

$$T - 22 = 64 \times (0.75)^{t/3}, \quad \text{即} \quad T = 22 + 64 \times \left(\frac{3}{4}\right)^{t/3}$$

(3) $t = 12$: $(0.75)^4 = \frac{81}{256} \approx 0.316$, $64 \times 0.316 \approx 20.2$, $T \approx 42.2$ ($^{\circ}\text{C}$); $t = 15$: $(0.75)^5 = \frac{243}{1024} \approx 0.237$, $64 \times 0.237 \approx 15.2$, $T \approx 37.2$ ($^{\circ}\text{C}$)。 (4) 检验：模型给出的温度单调下降且永远不低于室温 22°C ($(\frac{3}{4})^{t/3} \rightarrow 0$)，符合物理直觉； $t = 12, 15$ 都在实测时间附近，属内插，预测可靠。

易错点：判断模型要看“比”而不是“差”；指数底数是每 3 分钟乘 0.75，所以指数是 $\frac{t}{3}$ 而不是 t 。

例题 5.2 解析

思考路径：先定定义域（利润非负且 x 为自然数），再看对称轴是否落在可行域内，不在就取最近整数比较。

规范解：利润非负要求 $-x^2 + 15x - 44 \geq 0$ ，即 $x^2 - 15x + 44 \leq 0$ ，因式分解得 $(x - 4)(x - 11) \leq 0$ ，故 $4 \leq x \leq 11$ ；又 x 为自然数，定义域 $\{4, 5, \dots, 11\}$ 。对称轴 $x = \frac{15}{2} = 7.5$ ，但 7.5 不是自然数，卖不了“半台”打印机。比较 $x = 7$ 与 $x = 8$ ：

$$y(7) = -49 + 105 - 44 = 12, \quad y(8) = -64 + 120 - 44 = 12$$

答：卖出 **7 台或 8 台**时利润最大，最大利润 **12 万元**。

易错点：数学顶点 $x = 7.5$ 是不可取的——算完不检查“答案是不是个‘人话’（整数台、正数天）”，模型就在最后一步翻车。

5.2 知识点 (1) 建模四步与模型选择详解

A1. A。

思路：差为 3, 3, 3 恒定，对应线性。

解析： $y = 3x + 7$ ($x = 0$ 时 7，每步加 3)。

A2. A。

思路：周长 30，长 + 宽 = 15，面积 = 长 × 宽。

解析： $S = x(15 - x) = -(x - 7.5)^2 + 56.25$, $0 < x < 15$, $x = 7.5$ 时 $S_{\max} = 56.25$ (m^2)。易

错点：周长 30，长 + 宽 = 15，别把 30 当和；最大时是正方形（边长 7.5 m）。

A3. B。

思路：四步的顺序是审题、建模、求解、检验与回写。

解析：A、C、D 顺序都不完整或颠倒；B 完整且顺序正确。

A4. C。

思路：建模选模型要看数据形状，不是选最复杂。

解析：C 不恰当——“套用一个最复杂的模型”违背奥卡姆剃刀（解释力相当取更简单），也不看数据形状。

M1. B。

思路：建模 = 完整四步闭环，检验环节不能省。

解析：A、C、D 都是建模的正确动作；B “看到增长就套指数、不检验”是套模板——数据到底是“差稳定”（线性）还是“比稳定”（指数）要看数据说话。

为什么错：选 B 以外的同学没抓住“套模板”的本质。传统应用题的套路是“题目给公式让你代”，而建模情境往往是“没套路”的，**必须走完整四步，尤其是检验**——这正是本题的对抗点。

C1. (1) $T - 24$: 64, 48, 36, 27, 相邻比 $\frac{48}{64} = \frac{36}{48} = \frac{27}{36} = 0.75$ 恒定，故用**指数模型**。(2) $T - 24 = 64 \times (0.75)^{t/5}$, 即 $T = 24 + 64 \times \left(\frac{3}{4}\right)^{t/5}$ 。(3) $t = 25$: $(0.75)^5 = \frac{243}{1024} \approx 0.237$, $64 \times 0.237 \approx 15.2$, $T \approx 39.2$ ($^{\circ}\text{C}$)。 (4) $t \rightarrow +\infty$ 时 $\left(\frac{3}{4}\right)^{t/5} \rightarrow 0$, $T \rightarrow 24$ ($^{\circ}\text{C}$)，即最终趋向室温，与物理实际一致——检验通过。

易错点：先看“高出室温的部分”再判断模型，是这类“逼近平衡值”问题的标准手法；指数底数是每 5 分钟乘 0.75，指数写 $t/5$ 。

5.3 知识点 (2) 现实定义域与模型检验详解

A5. B。

思路：数值检验不过关。

解析：销量为负违背现实。应检查数据趋势（销量递增则斜率应为正），修正模型或重新拟合。

易错点：数学解“合理但不现实”时，先怀疑模型而不是硬采信。

A6. B。

思路：建利润函数，再查对称轴是否落在售价定义域 $[20, 50]$ 内。

解析：日利润 $L(x) = (x - 20)(100 - x) = -x^2 + 120x - 2000$ ，对称轴 $x = 60$ 不在 $[20, 50]$ 内，且开口向下，故 L 在 $[20, 50]$ 上单调递增， $x = 50$ 时 $L_{\max}$ 。**易错点：**对称轴不在定义域时最值在端点，别硬把 $x = 60$ 当答案——售价上限 50 元才是实际最优点。

M2. B。

思路：定义域是 $\{4, 5, \dots, 11\}$ 的自然数，对称轴 7.5 不在其中。

解析： $y(7) = y(8) = 12$ ，最大利润 12 万元。

为什么错：选 A 的同学直接采信数学顶点，忘了“台数必须是整数”。**现实定义域让最值点变了**——数学顶点 $x = 7.5$ 不可取，回到整数约束后最大值在 $x = 7$ 或 $x = 8$ 处取得，而不是 12.25 万元。

5.4 例题详解

例题 5.3 解析

思考路径：分别算差与比，哪个稳定用哪个；分段模型看断点衔接。

规范解：A：差为 3, 3, 3, 3 恒定，故用**线性**， $y = 3x + 1$ ($x = 1$ 时 4 ✓)。B：比为 2, 2, 2, 2 恒定，故用**指数**， $y = 3 \times 2^{x-1}$ ($x = 1$ 时 3， $x = 5$ 时 48 ✓)。C：前 1 kg 内 10 元，之后每 1 kg 加 4 元，故用**分段模型**： $y = 10$ ($0 < x \leq 1$)， $y = 10 + 4[x - 1]$ ($x > 1$)。衔接检验： $x = 1$ 时第一段 $y = 10$ ，第二段 $10 + 4 \times 0 = 10$ ，断点处两段取值相等 ✓。

易错点：数据都递增不代表都是线性——A 走“加”的路、B 走“乘”的路，增长模式不同，模型天差地别。模型选择要**实际算差和比**，不是看关键词。

例题 5.4 解析

思考路径：拟合好 $\neq$ 预测好。分别算残差平方和与预测值，再判断哪个模型更稳。

规范解：(1) 甲在三点处的预测值为 \$1.4, 1.8, 2.2\$，残差为 -0.2, 0.7, -0.2，残差平方和 = $0.04 + 0.49 + 0.04 = 0.57$ 。乙精确穿过三点，残差全为 0，残差平方和为 0\$。(2) 甲预测： $x = 3$ 时 $\hat{y}_{\text{甲}} = 0.4 \times 3 + 1.4 = 2.6$ (万件)；乙预测： $y_{\text{乙}} = -0.9 \times 9 + 2.2 \times 3 + 1.2 = -0.3$ (万件)——**销量是负数，明显荒谬**。(3) 甲更值得用：它对已知点残差稍大，但对新点预测合理稳定；乙“记住了”三个点，却在 $x = 3$ 处剧烈外摆，是典型的**过拟合**。

易错点：评判模型的标准不是“对已知数据拟合多完美”，而是“对新数据预测多靠谱”。残差平方和最小不一定是最好的模型——过拟合就是“拟合完美、预测崩坏”。

5.5 知识点 (1) 数据形状决定模型家族详解

A7. B。

思路：相邻比恒为 2。

解析： $\frac{6}{3} = \frac{12}{6} = \cdots = 2$ ，指数模型。

A8. A。

思路：相邻差恒为 3。

解析： $5 - 2 = 8 - 5 = \cdots = 3$ ，线性模型。

A9. C。

思路：先看一阶差，再看二阶差。

解析：一阶差 3, 5, 7 不恒定；二阶差 2, 2 恒定，对应二次模型。易错点：二阶差恒定才是二次，一阶差恒定是线性，别混淆。

A10. B。

思路：每年增长约 8%，即每年乘 1.08，比恒定。

解析： $N = N_0 \times 1.08^t$ ，指数模型。

M3. B。

思路：差稳定 vs 比稳定。

解析：甲组差 3, 3, 3，对应线性；乙组比 2, 2, 2，对应指数。

为什么错：选 A 或 C 的同学只看到“都递增”，没有区分增长模式。线性是“每步加同样多”，指数是“每步乘同样多”——要实际列差、列比，不能凭“都在长”下结论。

B1. (2023 · 新课标 I 卷 · 第 10 题) ACD。

思路：由 $L_p = 20 \lg \frac{p}{p_0}$ 反解出 p 的范围，再逐项比较。

解析： $L_1 \in [60, 90]$ 得 $p_1 \in [10^3 p_0, 10^{4.5} p_0]$ ； $L_2 \in [50, 60]$ 得 $p_2 \in [10^{2.5} p_0, 10^3 p_0]$ ； $L_3 = 40$ 得 $\frac{p_3}{p_0} = 10^2$ ，即 $p_3 = 100 p_0$ 。A： $p_1 \geq 10^3 p_0 \geq p_2$ ，对。B： $p_2 \leq 10^3 p_0 = 10 p_3$ ，故 $p_2 > 10 p_3$ 错。

C： $p_3 = 100 p_0$ ，对。D： $p_1 \leq 10^{4.5} p_0$ ，而 $100 p_2 \geq 100 \times 10^{2.5} p_0 = 10^{4.5} p_0$ ，故 $p_1 \leq 100 p_2$ ，对。

答案 ACD。

为什么错：(1) 声压级是“声压比的对数”，差 10 dB 对应声压变 10 倍——直接拿 dB 数比较大小会误判，必须反解出 p 的范围再比；(2) D 项用“ p_1 上界 $\leq 100 p_2$ 下界”验证，等号可同时达到，别因“临界太接近”而不敢选；(3) 多选漏选：每个选项独立核算一遍，C、D 都要逐个落笔。

5.6 知识点 (2) 拟合、过拟合与检验详解

A11. B。

思路：过拟合的特征——已知点全过、新点崩坏。

解析：模型把噪声也拟合进去了，应改用更简单的模型（如线性/二次），用残差与外推稳定性重新评估。

A12. B。

思路：残差有正有负，直接相加会抵消。

解析：平方让正负都变正、且放大离得远的点，才能反映真实拟合偏差。

A13. A。

思路：逐项平方再求和。

解析： $0.1^2 + (-0.2)^2 + 0.1^2 + (-0.3)^2 + 0.3^2 = 0.01 + 0.04 + 0.01 + 0.09 + 0.09 = 0.24$ 。**易错**

点：负残差平方后为正，别把 -0.2 、 -0.3 的符号带进和。

M4. B。

思路：指数模型 + 现实检验。

解析： $100 \times 1.05^{50} = 100 \times 11.467 \approx 1147$ （万人）。数学算对了，但人口受资源、空间约束，不可能长期指数增长——应结合现实修正为 S 型模型。

为什么错：选 A 的同学信奉“数学正确 = 结论正确”。指数增长可以解释初期的爆发，但**不能无限持续**——资源有限是现实给数学模型上的紧箍咒，检验环节就是专门把“数学漂亮、现实荒谬”的模型拦下来的。

C2. (1) 相邻比 $\frac{2}{1} = \frac{4}{2} = \dots = 2$ 恒定，故用**指数模型**， $N = 2^t$ （万辆， t 为年， $t = 0$ 时 1， $t = 4$ 时 16 ✓）。(2) $t = 7$ ： $N = 2^7 = 128$ （万辆）。(3) 不合理。128 万辆单车对应 200 万人口，人均约 0.64 辆，远超实际需求；且道路、停车资源有限，投放量不可能持续翻倍。**现实检验不过关**，应修正为 S 型（逻辑斯蒂）模型：初期指数增长、接近市场饱和时增长放缓。

B2. (2023 · 新课标 II 卷 · 第 19 题)

思路：先厘清两个概率的“看哪张直方图、算哪块面积”——漏诊率 $p(c)$ 是患病者直方图中指标 $\leq c$ 的面积（患病却被判阴性）；误诊率 $q(c)$ 是未患病者直方图中指标 $> c$ 的面积（没病却被判阳性）。再按 c 落在哪个组距内分段列式。

解析：(1) $p(c) = 0.5\% = 0.005$ 。患病者直方图中 $[95, 100]$ 的频率与组距之比为 0.002，且 $0.005 < 5 \times 0.002$ ，故 c 落在 $[95, 100]$ 内：

$$c = 95 + \frac{0.005}{0.002} = 97.5$$

此时误诊率（未患病者中指标 > 97.5 的面积）：

$$q(c) = (100 - 97.5) \times 0.01 + 5 \times 0.002 = 0.025 + 0.01 = 0.035 = 3.5\%$$

(2) 当 $c \in [95, 100]$ ： $p(c) = (c - 95) \times 0.002$ ， $q(c) = (100 - c) \times 0.01 + 5 \times 0.002$ ，故

$$f(c) = (c - 95) \times 0.002 + (100 - c) \times 0.01 + 0.01 = -0.008c + 0.82$$

当 $c \in (100, 105]$: $p(c) = 5 \times 0.002 + (c - 100) \times 0.012$, $q(c) = (105 - c) \times 0.002$, 故

$$f(c) = 0.01 + (c - 100) \times 0.012 + (105 - c) \times 0.002 = 0.01c - 0.98$$

$f(c)$ 在 $[95, 100]$ 上单调递减、在 $(100, 105]$ 上单调递增, 故最小值在 $c = 100$ 处:

$$f(100) = -0.008 \times 100 + 0.82 = 0.02$$

为什么错: (1) 漏诊与误诊方向搞反——漏诊看患病者、误诊看未患病者, 且一个是“ $\leq c$ ”的面积、一个是“ $> c$ ”的面积, 写反就会把两段公式张冠李戴; (2) 直方图纵轴是频率 $\div$ 组距, 某区间频率 = 高度 $\times$ 组距长度, 别把高度直接当概率; (3) 分段点 $c = 100$ 两段都要能取到, 求最值要比较 $f(100)$ 与两端点, 别只算一段。

5.7 例题详解

例题 5.5 解析

思考路径: 加法/数乘按对应元素; 乘法“行乘列求和”, 算完比较 AB 与 BA 。

规范解: $A + B = \begin{pmatrix} 1+4 & 3+1 \\ 2+1 & 0+5 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 3 & 5 \end{pmatrix}$; $2A = \begin{pmatrix} 2 & 6 \\ 4 & 0 \end{pmatrix}$ 。 AB : 第 1 行第 1 列 $1 \times 4 + 3 \times 1 = 7$; 第 1 行第 2 列 $1 \times 1 + 3 \times 5 = 16$; 第 2 行第 1 列 $2 \times 4 + 0 \times 1 = 8$; 第 2 行第 2 列 $2 \times 1 + 0 \times 5 = 2$ 。

$$AB = \begin{pmatrix} 7 & 16 \\ 8 & 2 \end{pmatrix}$$

BA : 第 1 行第 1 列 $4 \times 1 + 1 \times 2 = 6$; 第 1 行第 2 列 $4 \times 3 + 1 \times 0 = 12$; 第 2 行第 1 列 $1 \times 1 + 5 \times 2 = 11$; 第 2 行第 2 列 $1 \times 3 + 5 \times 0 = 3$ 。

$$BA = \begin{pmatrix} 6 & 12 \\ 11 & 3 \end{pmatrix}$$

$AB \neq BA$ ——矩阵乘法不满足交换律。

易错点: AB 与 BA 的每个元素都是“左行 $\times$ 右列”, 顺序不同结果就不同; 乘法还要求“左列数 = 右行数”。

例题 5.6 解析

思考路径: 系数矩阵 $\times$ 变量列向量 = 常数列向量; 求解用代入消元, 再用矩阵乘法验证。

规范解: 矩阵形式为

$$\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 12 \\ 7 \end{pmatrix}$$

由 $2x + y = 7$ 得 $y = 7 - 2x$, 代入 $3x + 2y = 12$: $3x + 14 - 4x = 12$, $-x = -2$, $x = 2$, $y = 3$ 。验证: $\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \times 2 + 2 \times 3 \\ 2 \times 2 + 1 \times 3 \end{pmatrix} = \begin{pmatrix} 12 \\ 7 \end{pmatrix}$ ✓。

易错点: 写成矩阵形式时系数别错位; 用矩阵乘法“还原”验证解是好习惯。

5.8 知识点 (1) 矩阵的加减与数乘详解

A14. A。

思路：同型矩阵对应元素相加。

解析：
$$\begin{pmatrix} 1+4 & 3+1 \\ 2+1 & 0+5 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 3 & 5 \end{pmatrix}。$$

A15. A。

思路：数乘是每个元素都乘这个数。

解析：
$$2 \times \begin{pmatrix} 1 & -3 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 2 & -6 \\ 0 & 8 \end{pmatrix}。$$
易错点：每个元素都要乘，且符号保留。

A16. A。

思路：对应元素相加列方程。

解析： $a+1=3$ 得 $a=2$ ； $1+c=2$ 得 $c=1$ ； $2+d=4$ 得 $d=2$ ； $b+3=5$ 得 $b=2$ 。

5.9 知识点 (2) 矩阵乘法与线性方程组详解

A17. A。

思路：行乘列求和。

解析：
$$\begin{pmatrix} 1 \times 2 + 2 \times 1 & 1 \times 0 + 2 \times 1 \\ 3 \times 2 + 4 \times 1 & 3 \times 0 + 4 \times 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 10 & 4 \end{pmatrix}。$$

A18. B。

思路：代入消元求解。

解析：由 $x-y=1$ 得 $x=1+y$ ，代入 $2(1+y)+3y=8$ ， $2+5y=8$ ， $y=\frac{6}{5}$ ， $x=\frac{11}{5}$ 。

M5. B。

思路：矩阵乘法是“行乘列求和”。

解析：正确结果 $AB = \begin{pmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}。$

为什么错：选 A 的同学把“矩阵加法”的规则（对应元素相加）迁移到了“矩阵乘法”上。乘法不是对应元素相乘，而是左矩阵的第 i 行与右矩阵的第 j 列对应相乘后求和。

M6. B、C。

思路：分别算 AB 与 BA 再比较。

解析：
$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}；BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}。AB \neq BA，且 BA 是零矩阵。$$

为什么错：选 A 的同学把“乘法满足交换律”的旧习惯（数的乘法）带进了矩阵。矩阵乘法

一般不满足交换律—— AB 与 BA 的运算顺序不同，结果可以完全不同（这里 BA 甚至把一切“抹平”为零矩阵）。

5.10 例题详解

例题 5.7 解析

思考路径：反证法——假设结论反面“ $x \leq 1$ 且 $y \leq 1$ ”成立，推出与条件矛盾。

规范解：反证。假设 $x \leq 1$ 且 $y \leq 1$ ，则 $x + y \leq 2$ ，与已知 $x + y > 2$ 矛盾。故假设不成立， $x > 1$ 或 $y > 1$ 。■

易错点：结论“ $x > 1$ 或 $y > 1$ ”的反面是“ $x \leq 1$ 且 $y \leq 1$ ”（否“或”变“且”），这是反证法中最容易写反的一步。

例题 5.8 解析

思考路径： R^2 高只说明“在直线假设下拟合得好”，模型类型对不对要看散点图。

规范解：（1）不能。散点图显示数据随施肥量先增后减，是抛物线形状——用直线去拟合弯的数据， R^2 再高也是“拟合错了形状”。（2）回到原理判断：放弃线性模型，改用二次或分段模型重新拟合，再回代检验残差。

易错点：这是 M7 的落点——对软件输出先问一句“散点图长什么样”。工具负责“算”，“选什么模型、结论怎么读”始终是人做的判断。

5.11 知识点（1）命题逻辑与反证法详解

A19. A。

思路：原命题真（ $x > 1$ 两边平方得 $x^2 > 1$ ）；逆否命题与原命题同真。

解析： $x > 1$ 时 $x^2 > 1$ ，真；逆否命题“若 $x^2 \leq 1$ ，则 $x \leq 1$ ”也为真。**易错点：**逆否命题要把条件和结论都否定并互换，别写成逆命题。

A20. C。

思路：条件与结论同时否定并互换。

解析：原命题“若 $x > 2$ ，则 $x^2 > 4$ ”，逆否命题为“若 $x^2 \leq 4$ ，则 $x \leq 2$ ”。**易错点：** $x > 2$ 的否定是 $x \leq 2$ ，不是 $x < 2$ 。

A21. A。

思路：反证法第一步是假设结论的反面成立。

解析：结论“质数有无穷多个”的反面是“质数只有有限个”，故第一步假设质数只有有限个，再推出矛盾。

5.12 知识点 (2) 数学史与新技术输出解读详解

A22. B。

思路：现实检验优先于 R^2 。

解析：营养早已耗尽，指数增长在现实上不可持续，输出不能直接用；应结合背景改用 S 型模型。

M7. B。

思路： R^2 高 $\neq$ 模型合适，先看散点图。

解析： $R^2 = 0.99$ 只能说明在“直线假设”下拟合得好；若散点图是弯的，直线假设本身错了， R^2 再高也不能采信。

为什么错：选 A、C 的同学有“工具崇拜”——觉得软件输出的数字大、可信度高。工具可以代替手算，但代替不了原理判断：选什么模型、结论怎么读，都是人做的判断。数据明明是弯的， R^2 再漂亮也是自欺欺人。

B3. (2022 · 全国甲卷 · 理 · 第 8 题) B。

思路：按“会圆术”公式求 s ，关键是求出弦长 AB 与“矢高” CD 。 C 是弦 AB 中点且 $CD \perp AB$ ，过圆心的弦的垂线必平分弦，故 O, C, D 三点共线。

解析： $OA = OB = 2$, $\angle AOB = 60^\circ$, $\triangle OAB$ 为等边三角形，故 $AB = 2$ 。连接 OC , C 是 AB 中点，则 $OC \perp AB$, $OC = \frac{\sqrt{3}}{2} \times AB = \sqrt{3}$ (等边三角形的高)。又 $CD \perp AB$ ，故 O, C, D 三点共线， $OD = OA = 2$ ，于是

$$CD = OD - OC = 2 - \sqrt{3}$$

代入公式：

$$s = AB + \frac{CD^2}{OA} = 2 + \frac{(2 - \sqrt{3})^2}{2} = 2 + \frac{7 - 4\sqrt{3}}{2} = \frac{11 - 4\sqrt{3}}{2}$$

选 B。

为什么错：(1) 漏掉“ O, C, D 三点共线”这一关键——圆心到弦中点的连线垂直于弦，而 CD 也垂直于弦，两条垂线是同一条，故 $CD = OD - OC$ ，把它想成“无关的两段”就算不出；(2) OC 是等边三角形的高 $\sqrt{3}$ ，不是 $\frac{AB}{2} = 1$ ，算错 CD 会顺错整个公式；(3) 公式里是 $\frac{CD^2}{OA}$ ，分母是 $OA = 2$ 不是 OC ，代入别抄错。

本章综合详解

S1 解析

思考路径：看相邻比，确定模型后外推预测，最后用“市场逐渐饱和”做现实检验。

规范解：(1) 相邻比 $\frac{8}{4} = \frac{16}{8} = \frac{32}{16} = 2$ 恒定，故用**指数模型**。 $N = 4 \times 2^{x-1}$ (万辆)，即 $N = 2^{x+1}$ ($x = 1$ 时 4, $x = 4$ 时 32 $\checkmark$)。 (2) 第 5 个月： $N = 2^6 = 64$ (万辆)；第 6 个月： $N = 2^7 = 128$ (万辆)。 (3) 第 10 个月： $N = 2^{11} = 2048$ (万辆)。目标用户约 5000 万、市场逐渐饱和，而该模型每两个月翻倍，第 12 个月就达 $2^{13} = 8192$ 万辆，逼近甚至超过用户总量

——**现实检验不过关**。应修正为 S 型（逻辑斯蒂）模型：初期指数增长，接近市场容量时增长放缓直至饱和。

易错点：指数模型适合刻画“初期爆发”，但不适合“长期预测”；检验环节要把“模型说了什么”与“现实是什么”对照，不合理就回头改模型。

S2 解析

思考路径：用“行对应原料、列对应产品”的消耗矩阵乘产量向量，得原料总需求；最后与库存比较做可行性判断。

规范解：(1) $M = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ （第 1 行对应 A 原料、第 2 行对应 B 原料；第 1 列对应甲产品、第 2 列对应乙产品）。(2) 产量列向量为 $\begin{pmatrix} 100 \\ 200 \end{pmatrix}$ ，则

$$M \begin{pmatrix} 100 \\ 200 \end{pmatrix} = \begin{pmatrix} 2 \times 100 + 1 \times 200 \\ 1 \times 100 + 3 \times 200 \end{pmatrix} = \begin{pmatrix} 400 \\ 700 \end{pmatrix}$$

即需 A 原料 400 kg、B 原料 700 kg。(3) B 原料需求 700 kg > 现有 500 kg，故按此计划**不可行**。需调整产量（如减少乙产品产量）或补充原料；调整后用矩阵乘法重新核算。

易错点：(1) 矩阵乘法“列对应”要一致——消耗矩阵的列数必须等于产量向量的行数，行对应原料、列对应产品，别把行列写反；(2) 可行性判断要逐项原料与库存比较，不能只看总需求量小于总库存就下结论。

第六章 卷末综合检测

一、单项选择题

Z1 解析

思路：总选法－全男生选法，或用分类加法计数原理。

规范解： $C_7^3 - C_4^3 = 35 - 4 = 31$ 。

易错点：“至少 1 名女生”用反面（全是男生）更快； $C_4^3 = 4$ 别算错。

Z2 解析

思路：列举和为 8 的等可能结果。

规范解：总结果 $6 \times 6 = 36$ 种。和为 8 的有 $(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)$ 共 5 种， $P = \frac{5}{36}$ 。

易错点： $(4, 4)$ 只有 1 种，别重复数成 2 种；和为 8 的组合共 5 种而非 6 种。

Z3 解析

思路：分层抽样按人数比例抽取。

规范解：高二占比 $\frac{2}{3+2+1} = \frac{1}{3}$ ，抽取 $60 \times \frac{1}{3} = 20$ （人）。

易错点：比例是“该层人数 ÷ 总人数”，三个年级分别对应 $\frac{3}{6}, \frac{2}{6}, \frac{1}{6}$ ，不是各抽 20 人。

Z4 解析

思路：方差性质——乘 k 变 k^2 倍，加 c 不变。

规范解：乘 3： $4 \times 3^2 = 36$ ；加 2 不变，故新方差为 36。

易错点：加常数不影响方差，只有乘除才改变；乘 k 方差乘 k^2 不是乘 k 。

Z5 解析

思路：回归直线必过样本中心，把点 $(3, 8)$ 代入。

规范解： $8 = 2 \times 3 + a$ ， $a = 2$ 。

易错点：这里 $(3, 8)$ 是样本中心，代入回归方程求 a ；别把 a 当成 $\bar{y}$ 。

二、多项选择题

Z6 解析

思路：数字特征性质——平均数、方差在线性变换下的变化，以及极差与平均数的关系。

规范解：A 错：平均数 10 不代表中位数也是 10（如 1, 10, 10, 10, 10, 10, 10, 19 的平均数恰为 10，中位数却是 10——但可构造反例使中位数不等于 10，故“一定”错）。B 对：最大值为 20，若极差 < 10 则最小值 > 10 ，全部数据都大于 10，平均数必大于 10，矛盾；故极差 ≥ 10 。C 对：平均数线性变换 $\bar{y} = 2 \times 10 + 1 = 21$ 。D 对：方差乘 $2^2 = 4$ 倍， $4 \times 4 = 16$ 。答案：BCD。

易错点：A 项“一定”过于绝对——中位数与平均数没有必然相等的关系；B 项用反证法判断极差下界。

Z7 解析

思路： r 的符号定方向、 $|r|$ 定强度；回归直线过中心点； $r \approx 0$ 只说明无线性相关；外推不可靠。

规范解：A 对： $r = -0.9 < 0$ 负相关， $|r| = 0.9$ 较强。B 对：回归直线必过 $(\bar{x}, \bar{y})$ 。C 错： r 接近 0 只说明无线性相关，可能有曲线相关。D 对：外推没有数据支撑，结果不一定可靠。答案：ABD。

易错点：C 是本章经典误区——“线性无关”不等于“无关”。

Z8 解析

思路：建模题——算相邻比判断模型，再做现实检验。

规范解： $\frac{140}{100} = 1.4$, $\frac{196}{140} = 1.4$, $\frac{274}{196} \approx 1.40$, $\frac{384}{274} \approx 1.40$ ，比约为 1.4 恒定，A 对。B 错：资源有限，指数增长不可能无限持续。C 对：指数模型 $N = 100 \times 1.4^t$ ， $t = 10$ 时 100×1.4^{10} 。D 对：长期应修正为 S 型模型。答案：ACD。

易错点：B 项把“模型能算”当成“现实会这样”——检验环节（现实约束）是建模四步中不可缺少的一步。

三、填空题

Z9 解析

思路：分类计数，先定百位（不能为 0）。

规范解：百位从 1, 2, 3, 4 中选，4 种；十位从剩余 4 个数字中选，4 种；个位从剩余 3 个中选，3 种。共 $4 \times 4 \times 3 = 48$ 个。

易错点：百位不能取 0；若先选两个非 0 再排序会漏算或重算，按“位数依次确定”最稳。

Z10 解析

思路：百分位数算法——排序 $\rightarrow i = \frac{p}{100} \times n \rightarrow$ 整数取相邻平均。

规范解： $n = 10$, $i = 0.7 \times 10 = 7$ 是整数，取第 7 个 (14) 与第 8 个 (16) 的平均数： $\frac{14 + 16}{2} = 15$ 。

易错点： i 为整数时必须取第 i 与第 $i + 1$ 个的平均，不能只取第 7 个 (14)。

Z11 解析

思路：独立事件，用对立事件：目标未击中的概率 = 两人都没命中。

规范解： $P = 1 - (1 - 0.8)(1 - 0.7) = 1 - 0.2 \times 0.3 = 1 - 0.06 = 0.94$ 。

易错点：不能直接把 0.8 与 0.7 相加（会重复计两人都命中的情形）；用“至少一个命中”的反面最稳。

四、解答题

Z12 解析

思路：频率 = 频数 ÷ 样本量；直方图高度 = 频率 ÷ 组距；中位数用累积频率平分；平均数用组中值加权。

规范解：(1) 频率依次为 $\frac{10}{100} = 0.1$, $\frac{25}{100} = 0.25$, $\frac{35}{100} = 0.35$, $\frac{20}{100} = 0.2$, $\frac{10}{100} = 0.1$ 。高度（频率 ÷ 组距 5）依次为 0.02, 0.05, 0.07, 0.04, 0.02。面积和 = $5 \times (0.02 + 0.05 + 0.07 + 0.04 + 0.02) = 5 \times 0.2 = 1$ ✓。(2) 累积频率 0.1, 0.35, 0.7, 0.9 落在 [10, 15]，中位数 $\approx 10 + \frac{0.5 - 0.35}{0.35} \times 5 = 10 + \frac{0.15}{0.35} \times 5 \approx 12.1$ （次）。(3) 组中值为 2.5, 7.5, 12.5, 17.5, 22.5，

$$\bar{x} \approx 2.5 \times 0.1 + 7.5 \times 0.25 + 12.5 \times 0.35 + 17.5 \times 0.2 + 22.5 \times 0.1 = 0.25 + 1.875 + 4.375 + 3.5 + 2.25 = 12.25$$

(4) 15 次及以上的人数 = 20 + 10 = 30，优秀率估计为 30%。

易错点：直方图高度是频率 ÷ 组距，不是频率；中位数用“还差多少面积” $(0.5 - 0.35)$ 除以该组面积 (0.35) 再乘组距。

Z13 解析

思路：组合计数 + 古典概型；“至少 1 名女生”用反面。

规范解：(1) $C_9^3 = \frac{9 \times 8 \times 7}{3 \times 2 \times 1} = 84$ （种）。(2) 恰有 1 名女生： $C_4^1 \times C_5^2 = 4 \times 10 = 40$ （种）， $P = \frac{40}{84} = \frac{10}{21}$ 。(3) 至少 1 名女生：反面是全是男生， $P = 1 - \frac{C_5^3}{C_9^3} = 1 - \frac{10}{84} = 1 - \frac{5}{42} = \frac{37}{42}$ 。

易错点：第(2)问是先选女生再选男生（乘法原理），别漏乘；第(3)问用对立事件“全是男生”更简洁，最后结果约分要彻底。

Z14 解析

思路：求回归方程 → 内插预测 → 列联表 + χ^2 独立性检验 → 外推判断。

规范解：(1) $\bar{x} = \frac{1+2+3+4+5}{5} = 3$, $\bar{y} = \frac{11+15+19+23+27}{5} = 19$ 。偏差： $x_i - \bar{x}$ 为 -2, -1, 0, 1, 2； $y_i - \bar{y}$ 为 -8, -4, 0, 4, 8。 $S_{xy} = 16 + 4 + 0 + 4 + 16 = 40$ ； $S_{xx} = 4 + 1 + 0 + 1 + 4 = 10$ 。

$$b = \frac{40}{10} = 4, \quad a = 19 - 4 \times 3 = 7$$

回归方程 $\hat{y} = 4x + 7$ 。检验： $4 \times 3 + 7 = 19$ ✓，过样本中心 (3, 19)。(2) $x = 6$ ： $\hat{y} = 4 \times 6 + 7 = 31$ （万元）。(3) 列联表为：

	销售额提升	销售额未提升	合计
投放短视频广告	45	15	60
未投放短视频广告	20	20	40
合计	65	35	100

期望频数： $E(\text{投放, 提升}) = \frac{60 \times 65}{100} = 39$ ； $E(\text{投放, 未提升}) = \frac{60 \times 35}{100} = 21$ ；
 $E(\text{未投放, 提升}) = \frac{40 \times 65}{100} = 26$ ； $E(\text{未投放, 未提升}) = \frac{40 \times 35}{100} = 14$ 。

$$\chi^2 = \frac{(45-39)^2}{39} + \frac{(15-21)^2}{21} + \frac{(20-26)^2}{26} + \frac{(20-14)^2}{14} = \frac{36}{39} + \frac{36}{21} + \frac{36}{26} + \frac{36}{14} \approx 0.92 + 1.71 + 1.38 + 2.57 = 6.59$$

(用公式 $\chi^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$ 验证： $a=45, b=15, c=20, d=20$, $ad-bc=900-300=600$, $\chi^2 = \frac{100 \times 600^2}{60 \times 40 \times 65 \times 35} = \frac{36000000}{5460000} \approx 6.59$ ✓) $6.59 > 3.841$, 拒绝原假设：

在犯错误的概率不超过 0.05 的前提下，认为投放短视频广告与销售额提升有关联——投放组销售额提升比例（75%）明显高于未投放组（50%）。(4) 不合理。 $x=20$ 已超出样本范围 $x \in [1, 5]$, 属于外推，没有数据支撑，预测结果不可靠。

易错点：第（1）问回归系数 $b=4$ 的量纲是“万元销售额/万元广告费”；第（3）问列联表四格要完整、三线合计要对应， χ^2 与临界值比较方向是“大于才拒绝”；第（4）问外推不可靠，不能用方程对范围外数据“自信满满”地预测。