第 3 节 二项分布与超几何分布（★★★）

强化训练

1.（★★）某人参加一次考试，共 3 道题，至少答对其中 2 道才能合格，若他答对每道题的概率均为 0.6，则他能合格的概率为_____.

答案：0.648

解析：答 3 题可看成三次独立试验，答对每题概率相同满足重复性，所以这是独立重复试验，故答对题目的个数服从二项分布，可用其概率分布计算所求概率，

设该人答对的题的个数为随机变量 X，则 $X \sim B(3,0.6)$，

所以他能合格的概率为 $P(X \ge 2) = P(X = 2) + P(X = 3) = C_3^2 \times 0.6^2 \times (1 - 0.6) + C_3^3 \times 0.6^3 = 0.648$.

2.（2021 • 天津卷 • ★★）甲、乙两人在每次猜谜活动中各猜一个谜语，若一方猜对且另一方猜错，则猜对的一方获胜，否则本次平局. 已知每次活动中，甲、乙猜对的概率分别为 $\frac{5}{6}$ 和 $\frac{1}{5}$，且每次活动中甲、乙猜对与否互不影响，各次活动也互不影响，则 1 次活动中，甲获胜的概率为_____；3 次活动中，甲至少获胜 2 次的概率为_____.

答案：$\frac{2}{3}$；$\frac{20}{27}$

解析：1 次活动中甲获胜，即甲猜对且乙未猜对，其概率为 $\frac{5}{6} \times (1 - \frac{1}{5}) = \frac{2}{3}$；3 次活动中，甲获胜的次数 $X \sim B(3, \frac{2}{3})$，

所以甲至少获胜 2 次的概率为 $P(X \ge 2) = P(X = 2) + P(X = 3) = C_3^2 \times (\frac{2}{3})^2 \times (1 - \frac{2}{3}) + C_3^3 \times (\frac{2}{3})^3 = \frac{20}{27}$.

3.（2022 • 福州模拟 • ★★★）为了保障我国民众的身体健康，某产品在进入市场前必须进行两轮检测，只有两轮检测都合格才能进行销售，否则不能销售. 已知该产品第一轮检测不合格的概率为 $\frac{1}{9}$，第二轮检测不合格的概率为 $\frac{1}{10}$，两轮检测是否合格相互之间没有影响，若产品可以销售，则每件产品获利 40 元，若产品不能销售，则每件亏损 80 元，已知一箱中有尚未检测的 4 件产品，记该箱产品总共获利 X 元，则 $P(X \ge -80) =$（　　）

（A）$\frac{96}{625}$　（B）$\frac{256}{625}$　（C）$\frac{608}{625}$　（D）$\frac{209}{625}$

答案：C

解析：4 件产品中能够销售的件数服从二项分布，先求一件产品可以销售的概率，

由题意，一件产品可以销售的概率为$(1-\frac{1}{9})\times(1-\frac{1}{10})=\frac{4}{5}$，所以该 4 件产品中，可以销售的件数$Y\sim B(4,\frac{4}{5})$，

该箱产品的获利X与可以销售的件数Y有关，找到此关系，即可把$P(X\geq -80)$化为Y的取值概率来算，

由题意，$X=40Y-80(4-Y)=120Y-320$，所以$P(X\geq -80)=P(120Y-320\geq -80)=P(Y\geq 2)$

$=1-P(Y=0)-P(Y=1)=1-C_4^0\times(1-\frac{4}{5})^4-C_4^1\times\frac{4}{5}\times(1-\frac{4}{5})^3=\frac{608}{625}$.

4．（2023•新高考Ⅱ卷•★★★★）（多选）在信道内传输 0，1 信号，信号的传输相互独立. 发送 0 时，收到 1 的概率为$\alpha(0<\alpha<1)$，收到 0 的概率为$1-\alpha$；发送 1 时，收到 0 的概率为$\beta(0<\beta<1)$，收到 1 的概率为$1-\beta$. 考虑两种传输方案：单次传输和三次传输. 单次传输是指每个信号只发送 1 次，三次传输是指每个信号重复发送 3 次. 收到的信号需要译码，译码规则如下：单次传输时，收到的信号即为译码；三次传输时，收到的信号中出现次数多的即为译码（例如，若依次收到 1，0，1，则译码为 1）.（　　）

（A）采用单次传输方案，若依次发送 1，0，1，则依次收到 1，0，1 的概率为$(1-\alpha)(1-\beta)^2$

（B）采用三次传输方案，若发送 1，则依次收到 1，0，1 的概率为$\beta(1-\beta)^2$

（C）采用三次传输方案，若发送 1，则译码为 1 的概率为$\beta(1-\beta)^2+(1-\beta)^3$

（D）当$0<\alpha<0.5$时，若发送 0，则采用三次传输方案译码为 0 的概率大于采用单次传输方案译码为 0 的概率

答案：ABD

解析：A 项，由题意，若采用单次传输方案，则发送 1 收到 1 的概率为$1-\beta$，发送 0 收到 0 的概率为$1-\alpha$，所以依次发送 1，0，1，则依次收到 1，0，1 的概率为$(1-\beta)(1-\alpha)(1-\beta)=(1-\alpha)(1-\beta)^2$，故 A 项正确；

B 项，采用三次传输方案，发送 1 即独立重复发送 3 次 1，

由于是依次收到 1，0，1，顺序不能交换，所以分别计算每次对应的概率，相乘即可，

由题意，依次收到 1，0，1 的概率为$(1-\beta)\beta(1-\beta)=\beta(1-\beta)^2$，故 B 项正确；

C 项，采用三次传输方案，由 B 项的分析过程可知若发送 1，则收到 1 的个数$X\sim B(3,1-\beta)$，

而译码为 1 需收到 2 个 1，或 3 个 1，

所以译码为 1 的概率为$P(X=2)+P(X=3)=C_3^2(1-\beta)^2\beta+C_3^3(1-\beta)^3=3(1-\beta)^2\beta+(1-\beta)^3$，故 C 项错误；

D 项，若采用单次传输方案，则发送 0 译码为 0 的概率为$1-\alpha$；

若采用三次传输方案，则发送 0 等同于发 3 个 0，收到 0 的个数$Y\sim B(3,1-\alpha)$，

且译码为 0 的概率为$P(Y=2)+P(Y=3)=C_3^2(1-\alpha)^2\alpha+C_3^3(1-\alpha)^3=3(1-\alpha)^2\alpha+(1-\alpha)^3$，

要比较上述两个概率的大小，可作差来看，

$3(1-\alpha)^2\alpha+(1-\alpha)^3-(1-\alpha)=(1-\alpha)[3(1-\alpha)\alpha+(1-\alpha)^2-1]=(1-\alpha)(1-2\alpha)\alpha$，

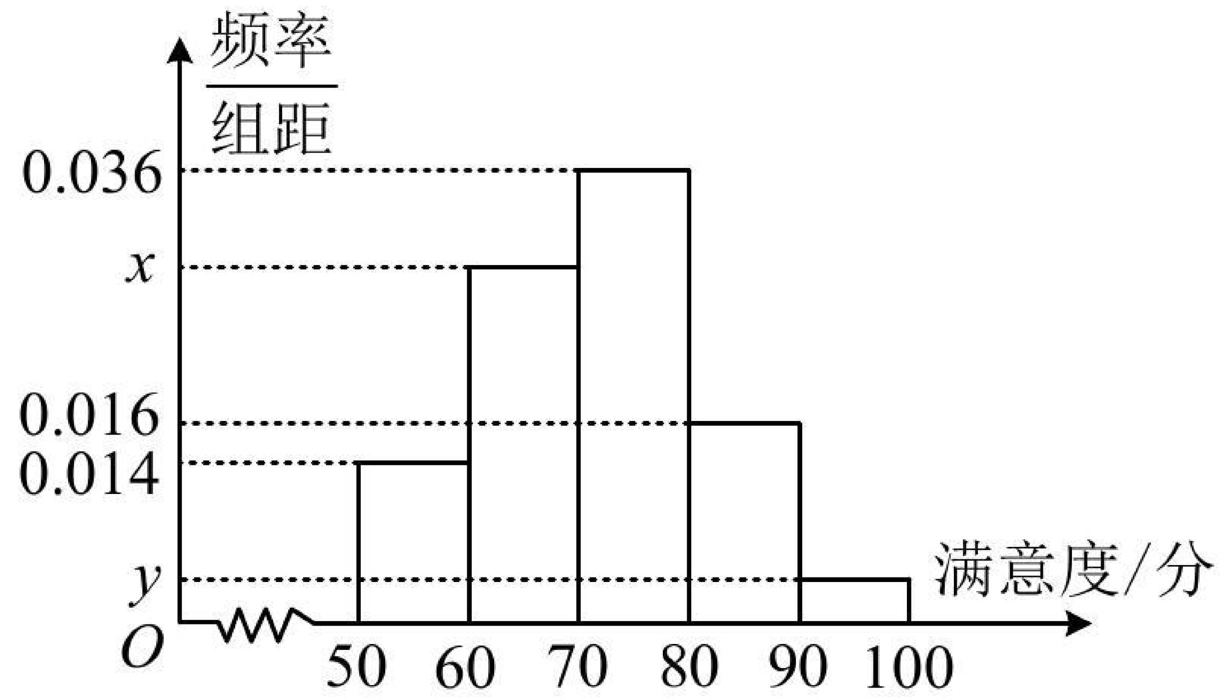
因为$0<\alpha<0.5$，所以$3(1-\alpha)^2\alpha+(1-\alpha)^3-(1-\alpha)=(1-\alpha)(1-2\alpha)\alpha>0$，

从而$3(1-\alpha)^2\alpha+(1-\alpha)^3>1-\alpha$，故 D 项正确.

5．（2023•成都七中模拟•★★★）随着人民生活水平的不断提高，“衣食住行”愈发被人们所重视，其中对饮食的要求也越来越高. 某地区为了解当地餐饮情况，随机抽取了 100 人对该地区的餐饮情况进行了

问卷调查. 请根据下面尚未完成并有局部污损的频率分布表和频率分布直方图解决下列问题.

组别	分组	频数	频率
第1组	$[50,60)$	14	0.14
第2组	$[60,70)$	m	
第3组	$[70,80)$	36	0.36
第4组	$[80,90)$		0.16
第5组	$[90,100]$	4	n



（1）求 m，n，x，y 的值；

（2）若将满意度在 80 分以上的人群称为“美食客”，将频率视为概率，用样本估计总体，从该地区中随机抽取 3 人，记其中“美食客”的人数为 ξ，求 ξ 的分布列和期望.

解：（1）（表中给出了 $[90,100]$ 这组的频数，可求频率 n，进而求得频率分布直方图中的 y）

由表中数据可知 $n=\dfrac{4}{100}=0.04$，所以频率分布直方图中 $y=\dfrac{0.04}{10}=0.004$，

（到此只有 $[60,70)$ 这一组的频率不知道了，可用频率和为 1 来求）

$[60,70)$ 这一组的频率为 $1-0.14-0.36-0.16-0.04=0.3$，所以 $m=100\times0.3=30$， $x=\dfrac{0.3}{10}=0.03$.

（2）由题意，样本的 100 人中“美食客”的频率为 $0.16+0.04=0.2$，

（由于是从该地区中抽取 3 人，该地区总人数很多，所以看成 3 重伯努利试验，用二项分布处理）

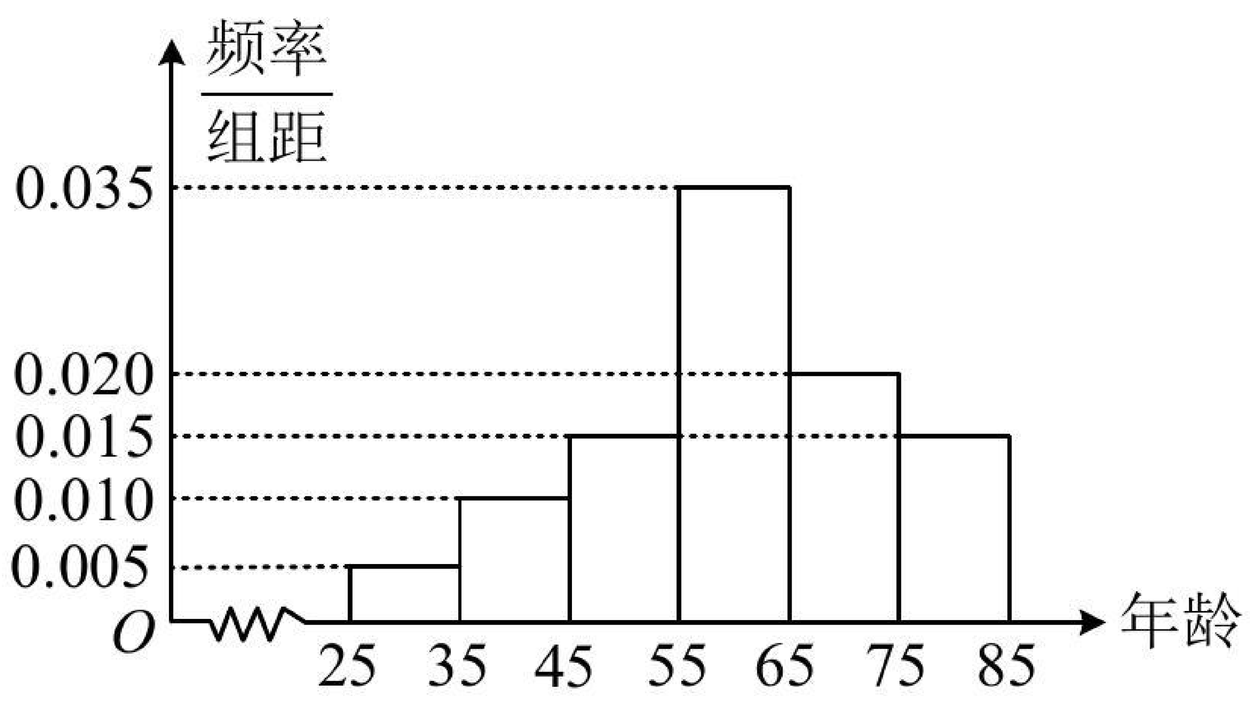
随机变量 $\xi\sim B(3,0.2)$，所以 $P(\xi=0)=\mathrm{C}_3^0\times(1-0.2)^3=0.512$， $P(\xi=1)=\mathrm{C}_3^1\times0.2\times(1-0.2)^2=0.384$，

$P(\xi=2)=\mathrm{C}_3^2\times0.2^2\times(1-0.2)=0.096$， $P(\xi=3)=\mathrm{C}_3^3\times0.2^3=0.008$，所以 ξ 的分布列为：

ξ	0	1	2	3
P	0.512	0.384	0.096	0.008

因为 $\xi\sim B(3,0.2)$，所以 ξ 的期望 $E(\xi)=3\times0.2=0.6$.

6.（2023·青海模拟·★★★）2021 年国庆期间，某县书画协会在县宣传部门的领导下组织了国庆书画展，参展的 200 幅书画作品反映了该县人民在党的领导下进行国家建设中的艰苦卓绝，这些书画作品的作者的年龄都在 $[25,85]$ 内，根据统计结果，得到如图所示的频率分布直方图：



（1）求这 200 位作者年龄的平均数 $\bar{x}$ 和方差 s^2；（同一组数据用该区间的中点值作代表）

（2）县委宣传部从年龄在 $[35,45)$ 和 $[65,75)$ 的作者中，用按比例分配的分层抽样方法抽取 6 人参加县委组织的表彰大会，现要从 6 人中选出 3 人作为代表发言，设这 3 位发言者的年龄落在区间 $[35,45)$ 的人数为 X，

求 X 的分布列和数学期望.

解：（1）由图可知，$\bar{x}=30\times0.05+40\times0.1+50\times0.15+60\times0.35+70\times0.2+80\times0.15=60$，

$s^2=(30-60)^2\times0.05+(40-60)^2\times0.1+(50-60)^2\times0.15+(60-60)^2\times0.35+(70-60)^2\times0.2+(80-60)^2\times0.15=180$.

（2）（先求出样本中 $[35,45)$ 和 $[65,75)$ 这两组的人数，才能找到选出的 6 人的组成情况）

由频率分布直方图可知 $[35,45)$ 这一组有 $200\times0.1=20$ 人，$[65,75)$ 这一组有 $200\times0.2=40$ 人，

所以抽取的 6 人中 $[35,45)$ 有 2 人，$[65,75)$ 有 4 人，

（再从这 6 人中取 3 人，取到 $[35,45)$ 的人数服从超几何分布，可用古典概率求其分布列）

故 X 可能的取值有 0，1，2，且 $P(X=0)=\dfrac{C_4^3}{C_6^3}=\dfrac{1}{5}$，$P(X=1)=\dfrac{C_4^2C_2^1}{C_6^3}=\dfrac{3}{5}$，$P(X=2)=\dfrac{C_4^1C_2^2}{C_6^3}=\dfrac{1}{5}$，

所以 X 的分布列为：

X	0	1	2
P	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

故 X 的期望 $E(X)=0\times\dfrac{1}{5}+1\times\dfrac{3}{5}+2\times\dfrac{1}{5}=1$.

【反思】求期望也可直接用超几何分布的期望公式 $E(X)=n\cdot\dfrac{M}{N}$，本题的期望即为 $E(X)=3\times\dfrac{2}{6}=1$.

7.（2023・江西模拟・★★★）党的二十大是全党全国各族人民迈上全面建设社会主义现代化国家的新征程、向第二个百年奋斗目标进军的关键时刻召开的一次十分重要的大会，认真学习宣传和全面贯彻落实党的二十大精神，是当前和今后一个时期的首要政治任务和头等大事. 某校计划举行党的二十大知识竞赛，对前来报名者进行初试，初试合格者进入正赛. 初试有备选题 6 道，从备选题中随机挑出 4 道题进行测试，至少答对 3 道题者视为合格. 已知甲、乙两人报名参加初试，在这 6 道题中甲能答对 4 道，乙能答对每道题的概率均为 $\dfrac{2}{3}$，且甲、乙两人各题是否答对相互独立.

（1）分别求甲、乙两人进入正赛的概率；

（2）记甲、乙两人中进入正赛的人数为 ξ，求 ξ 的分布列及 $E(2\xi-1)$.

解：（1）甲能进入正赛即选出的 4 道题中，至少有 3 道是会答的，可分两类：答对 3 道，有 $C_4^3C_2^1$ 种，答对 4 道，有 C_4^4 种，总的方法数为 C_6^4，故甲能进入正赛的概率 $p_1=\dfrac{C_4^3C_2^1+C_4^4}{C_6^4}=\dfrac{3}{5}$，

（对于乙，答 4 道题可看成 4 重伯努利试验，答对题目的个数应服从二项分布）

设抽出的 4 道题中，乙能正确回答的个数为 X，则 $X\sim B(4,\dfrac{2}{3})$，

所以乙能通过测试的概率 $p_2=P(X\geq3)=P(X=3)+P(X=4)=C_4^3\times(\dfrac{2}{3})^3\times(1-\dfrac{2}{3})+C_4^4\times(\dfrac{2}{3})^4=\dfrac{16}{27}$.

（2）由题意，ξ 可能的取值为 0，1，2，且 $P(\xi=0)=(1-p_1)(1-p_2)=\dfrac{22}{135}$，

$P(\xi=1)=p_1(1-p_2)+(1-p_1)p_2=\dfrac{13}{27}$，$P(\xi=2)=p_1p_2=\dfrac{16}{45}$，所以 ξ 的分布列为：

ξ	0	1	2
P	$\frac{22}{135}$	$\frac{13}{27}$	$\frac{16}{45}$

从而 $E(\xi)=0\times\frac{22}{135}+1\times\frac{13}{27}+2\times\frac{16}{45}=\frac{161}{135}$，故 $E(2\xi-1)=2E(\xi)-1=\frac{187}{135}$.

8．（2023 • 广东模拟 • ★★★★）某次射击比赛规定：每位参赛者最多有两次射击机会，第一次射击击中靶标，立即停止射击，得 4 分；第一次未击中靶标，继续进行第二次射击，若击中靶标，则得 3 分，若未击中靶标，则得 2 分. 现有 12 人参加该射击比赛，假设每人两次射击击中靶标的概率分别为 m，0.5，凡是击中了靶标的，称该参赛者过关，否则不过关，每人过关的概率为 p.

（1）求 p；（用 m 表示）

（2）设这 12 人中恰有 9 人过关的概率为 $f(p)$，求 $f(p)$ 取最大值时 p 和 m 的值；

（3）在（2）的结果下，求这 12 人射击比赛中所得总分的平均数.

解：(1)（过关可以第一次击中，也可以第二次击中，而不过关只能两次都不击中，故用对立事件求 p）

由题意，$p=1-(1-m)(1-0.5)=\frac{1+m}{2}$.

（2）（12 人参赛看成 12 重伯努利试验，过关的人数服从二项分布）

记 12 人中过关的人数为 X，则 $X\sim B(12,p)$，所以 $f(p)=P(X=9)=\mathrm{C}_{12}^{9}p^{9}(1-p)^{3}=220p^{9}(1-p)^{3}$，

（要研究此函数何时取得最大值，可求导分析）

$f'(p)=220[9p^{8}(1-p)^{3}+p^{9}\cdot 3(1-p)^{2}\cdot(-1)]=660p^{8}(1-p)^{2}(3-4p)$，

所以 $f'(p)>0\Leftrightarrow 0<p<\frac{3}{4}$，$f'(p)<0\Leftrightarrow \frac{3}{4}<p<1$，故 $f(p)$ 在 $(0,\frac{3}{4})$ 上单调递增，在 $(\frac{3}{4},1)$ 上单调递减，

所以当 $f(p)$ 取得最大值时，$p=\frac{3}{4}$，又由（1）知 $p=\frac{1+m}{2}$，所以此时 $m=2p-1=\frac{1}{2}$.

（3）（要求 12 人所得总分的平均数，可先求 1 人得分的期望）

设 1 位参赛者的得分为 Y，则 Y 可能的取值有 4，3，2，且 $P(Y=4)=m=0.5$，

$P(Y=3)=(1-m)\times 0.5=0.25$，$P(Y=2)=(1-m)(1-0.5)=0.25$，所以 $E(Y)=4\times 0.5+3\times 0.25+2\times 0.25=3.25$，

故这 12 人射击比赛中所得总分的平均数为 $12\times 3.25=39$.

9．（2023 • 四省联考 • ★★★★）一个池塘里的鱼的数目记为 N，从池塘里捞出 200 尾鱼，并给鱼作上标记，然后把鱼放回池塘里，过一小段时间后再从池塘里捞出 500 尾鱼，X 表示捞出的 500 尾鱼中有标识的鱼的数目.

（1）若 $N=5000$，求 X 的数学期望；

（2）已知捞出的 500 尾鱼中 15 尾有标识，试给出 N 的估计值.（以使得 $P(X=15)$ 最大的 N 的值作为 N 的估计值）

解：(1)（数字较大，先求分布列再算期望很繁琐，故直接代公式 $E(X)=n\cdot\frac{M}{N}$ 求期望）

由题意，X 服从超几何分布，且 $N=5000$，$M=200$，$n=500$，所以 $E(X)=n\cdot\frac{M}{N}=500\times\frac{200}{5000}=20$.

（2）（要分析何时 $P(X=15)$ 最大，需计算 $P(X=15)$，从题意来看，应按超几何分布求此概率，三个关键参数中， $M=200$， $n=500$，N 的情况未知，先结合抽取结果初步分析 N 的范围）

捞出的 500 尾鱼中有 15 尾有标识，则无标识的有 485 尾，结合总共有 200 尾有标识，故可推断 $N\ge 685$，

所以 $P(X=15)=\dfrac{C_{200}^{15}C_{N-200}^{485}}{C_N^{500}}$，（此式随 N 的变化而改变，由于 N 只能取正整数，故看成数列，要分析数列的最大项，可用相邻项作差或作商比较，此处涉及组合数计算，作商更方便）

记 $a_N=\dfrac{C_{200}^{15}C_{N-200}^{485}}{C_N^{500}}$， 则 $a_{N+1}=\dfrac{C_{200}^{15}C_{N-199}^{485}}{C_{N+1}^{500}}$， 所以 $\dfrac{a_{N+1}}{a_N}=\dfrac{C_{200}^{15}C_{N-199}^{485}}{C_{N+1}^{500}}\cdot\dfrac{C_N^{500}}{C_{200}^{15}C_{N-200}^{485}}$

$$=\frac{C_{N-199}^{485}\cdot C_N^{500}}{C_{N+1}^{500}\cdot C_{N-200}^{485}}=\frac{\dfrac{(N-199)!}{(485!)\times(N-684)!}\cdot\dfrac{N!}{(500!)\times(N-500)!}}{\dfrac{(N+1)!}{(500!)\times(N-499)!}\cdot\dfrac{(N-200)!}{(485!)\times(N-685)!}}=\frac{(N-199)(N-499)}{(N-684)(N+1)},$$

令 $\dfrac{a_{N+1}}{a_N}>1$ 可得：$\dfrac{(N-199)(N-499)}{(N-684)(N+1)}>1$，解得：$685\le N<\dfrac{99985}{15}\approx 6665.7$，结合 $N\in\mathbf{N}^*$ 可得 $685\le N\le 6665$，

此时 $a_{N+1}>a_N$； 同理，令 $\dfrac{a_{N+1}}{a_N}<1$ 可得 $N\ge 6666$，此时， $a_{N+1}<a_N$；

所以 $a_{685}<a_{686}<a_{687}<a_{688}<\cdots<a_{6666}>a_{6667}>a_{6668}>\cdots$，故当 a_N 取得最大值时， $N=6666$，即 N 的估计值为 6666.