2020 高考统计概率大题必刷热点题型

1.（2020•临川区校级模拟）2018 年反映社会现实的电影《我不是药神》引起了很大的轰动，治疗特种病的创新药研发成了当务之急．为此，某药企加大了研发投入，市场上治疗一类慢性病的特效药品 A 的研发费用 x（百万元）和销量 y（万盒）的统计数据如表：

研发费用 x（百万元）	2	3	6	10	13	15	18	21
销量 y（万盒）	1	1	2	2.5	3.5	3.5	4.5	6

（1）根据数据用最小二乘法求出 y 与 x 的线性回归方程 $y=\hat{b}x+\hat{a}$（系数用分数表示，不能用小数）；

（2）该药企准备生产药品 A 的三类不同的剂型 A_1，A_2，A_3，并对其进行两次检测，当第一次检测合格后，才能进行第二次检测．第一次检测时，三类剂型 A_1，A_2，A_3 合格的概率分别为 $\frac{1}{2}$，$\frac{3}{4}$，$\frac{3}{5}$，第二次检测时，三类剂型 A_1，A_2，A_3 合格的概率分别为 $\frac{4}{5}$，$\frac{2}{3}$，$\frac{2}{3}$．两次检测过程相互独立，设经过两次检测后 A_1，A_2，A_3 三类剂型合格的种类数为 X，求 X 的分布列与数学期望．

附：（1）$\hat{b}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{xy}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2},\hat{a}=\overline{y}-\hat{b}\overline{x}$（2）$\sum_{i=1}^{8}x_iy_i=347,\sum_{i=1}^{8}x_i^2=1308$．

2.（2020•湖北模拟）一只蚂蚁在如图所示的棱长为1米的正四面体的棱上爬行，每次当它到达四面体顶点后，会在过此顶点的三条棱中等可能的选择一条棱继续爬行（包含来时的棱），已知蚂蚁每分钟爬行 1 米，$t=0$时蚂蚁位于点A处.

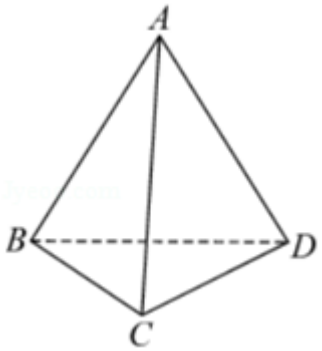
（1）2 分钟末蚂蚁位于哪点的概率最大；

（2）记第n分钟末蚂蚁位于点A，B，C，D的概率分别为P_n（A），P_n（B），P_n（C），P_n（D）.

①求证：P_n（B）$=P_n$（C）$=P_n$（D）；

②辰辰同学认为，一段时间后蚂蚁位于点A、B、C、D的概率应该相差无几，请你通过计算 10 分钟末蚂蚁位于各点的概率．解释辰辰同学观点的合理性.

附：$(\frac{1}{3})^9\approx5\times10^{-5},(\frac{1}{3})^{10}\approx1.7\times10^{-5}$，$(\frac{1}{2})^9\approx1.9\times10^{-9},(\frac{1}{3})^{10}\approx9.8\times10^{-4}$.

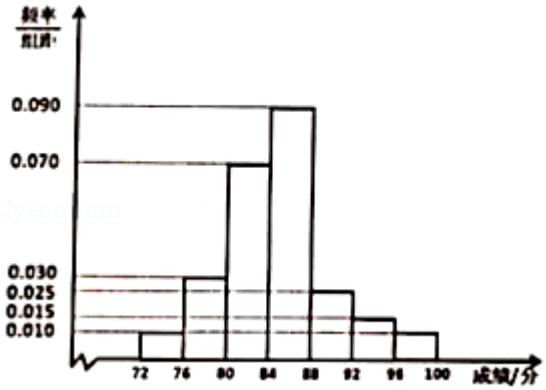


3.（2020•江西模拟）年前某市质监部门根据质量管理考核指标对本地的500家食品生产企业进行考核，然后通过随机抽样抽取其中的50家，统计其考核成绩（单位：分），并制成如图频率分布直方图.

（1）求这 50 家食品生产企业考核成绩的平均数 $\bar{x}$（同一组中的数据用该组区间的中点值为代表）及中位数 a（精确到0.01)

（2）该市质监部门打算举办食品生产企业质量交流会，并从这 50 家食品生产企业中随机抽取 4 家考核成绩不低于 88 分的企业发言，记抽到的企业中考核成绩在[92，100]的企业数为 X，求 X 的分布列与数学期望

（3）若该市食品生产企业的考核成绩 X 服从正态分布 $N(\mu,\sigma^2)$ 其中“近似为 50 家食品生产企业考核成绩的平均数 $\bar{x},\sigma^2$ 近似为样本方差，s^2，经计算得 $s^2=27.68$，利用该正态分布，估计该市 500 家食品生产企业质量管理考核成绩高于 90.06 分的有多少家？（结果保留整数）. 附参考数据与公式：$\sqrt{27.68}\approx 5.26$，$X-N(\mu,\sigma^2)$ 则 $P(\mu-\sigma<X,,\mu+\sigma)\approx 0.6827$，$P(\mu-2\sigma<X,,2\mu+\sigma)\approx 0.9545$. $P(\mu-3\sigma<X,,\mu+3\sigma)\approx 0.9973$.



4.（2020•厦门一模）小明和爸爸玩亲子游戏，规则如下：袋中装有 3 个大小相同的球，1 个白球，2 个红球，每次摸出一个球，记下颜色后放回，若摸出白球，则下一次由原摸球人继续摸球；若摸出红球，则下一次由对方摸球，规定摸球 m 次，最后一次由谁摸球就算谁获胜，第一次由小明摸球.

（1）求前 3 次摸球中小明恰好摸 2 次的概率；

（2）设第 n 次 $(n,, m)$ 由小明摸球的概率为 P_n，则 $P_1=1$.

(i) 求 P_4；

(ii) 在 $m=19$ 与 $m=20$ 之中选其一，小明应选哪个？（只写结果，不必说明理由！）

5.（2020•石家庄模拟）某工厂为生产一种精密管件研发了一台生产该精密管件的车床，该精密管件有内外两个口径，监管部门规定“口径误差”的计算方式为：管件内外两个口径实际长分别为 $a(mm)$， $b(mm)$，标准长分别为 $\overline{a}(mm),\overline{b}(mm)$，则“口径误差”为 $|a-\overline{a}|+|b-\overline{b}|$，只要“口径误差”不超过 $0.2min$ 就认为合格，已知这台车床分昼、夜两个独立批次生产．工厂质检部在两个批次生产的产品中分别随机抽取 40 件作为样本，经检测其中昼批次的 40 个样本中有 4 个不合格品，夜批次的 40 个样本中有 10 个不合格品．

(I) 以上述样本的频率作为概率，在昼夜两个批次中分别抽取 2 件产品，求其中恰有 1 件不合格产品的概率；

(II) 若每批次各生产 1000 件，已知每件产品的成本为 5 元，每件合格品的利润为 10 元；若对产品检验，则每件产品的检验费用为 2.5 元；若有不合格品进入用户手中，则工厂要对用户赔偿，这时生产的每件不合格品工厂要损失 25 元．以上述样本的频率作为概率，以总利润的期望值为决策依据，分析是否要对每个批次的所有产品作检测？

6.（2020•广东一模）一支担负勘探任务的队伍有若干个勘探小组和两类勘探人员，甲类人员应用某种新型勘探技术的精准率为 0.6，乙类人员应用这种勘探技术的精准率为$a(0<a<0.4)$．每个勘探小组配备 1 名甲类人员与 2 名乙类人员，假设在执行任务中每位人员均有一次应用这种技术的机会且互不影响，记在执行任务中每个勘探小组能精准应用这种新型技术的人员数量为ξ．

（1）证明：在ξ各个取值对应的概率中，概率$P(\xi=1)$的值最大．

（2）在特殊的勘探任务中，每次只能派一个勘探小组出发，工作时间不超过半小时，如果半小时内无法完成任务，则重新派另一组出发．现在有三个勘探小组$A_i(i=1，2，3)$可派出，若小组A_i能完成特殊任务的概率t；$t_i=P(\xi=i)(i=1，2，3)$，且各个小组能否完成任务相互独立．试分析以怎样的先后顺序派出勘探小组，可使在特殊勘探时所需派出的小组个数的均值达到最小．

7.（2020•濮阳一模）2020 年 1 月 10 日，引发新冠肺炎疫情的$COVID-9$病毒基因序列公布后，科学家们便开始了病毒疫苗的研究过程．但是类似这种病毒疫苗的研制需要科学的流程，不是一朝一夕能完成的，其中有一步就是做动物试验．已知一个科研团队用小白鼠做接种试验，检测接种疫苗后是否出现抗体．试验设计是：每天接种一次，3 天为一个接种周期．已知小白鼠接种后当天出现抗体的概率为$\frac{1}{2}$，假设每次接种后当天是否出现抗体与上次接种无关．

（Ⅰ）求一个接种周期内出现抗体次数kk的分布列；

（Ⅱ）已知每天接种一次花费 100 元，现有以下两种试验方案：

①若在一个接种周期内连续 2 次出现抗体即终止本周期试验，进行下一接种周期，试验持续三个接种周期，设此种试验方式的花费为X元；

②若在一个接种周期内出现 2 次或 3 次抗体，该周期结束后终止试验，已知试验至多持续三个接种周期，设此种试验方式的花费为Y元．

比较随机变量X和Y的数学期望的大小．

8．（2020•河南模拟）某中医药研究所研制出一种新型抗癌药物，服用后需要检验血液是否为阳性，现有$n(n\in N^*)$份血液样本，每个样本取到的可能性均等，有以下两种检验方式：（1）逐份检验，则需要检验n次；（2）混合检验，将其中$k(k\in N^*$，$2\leqslant k\leqslant n)$份血液样本分别取样混合在一起检验，若结果为阴性，则这k份的血液全为阴性，因而这k份血液样本只需检验一次就够了：若检验结果为阳性，为了明确这k份血液究竟哪份为阳性，就需要对这k份再逐份检验，此时这k份血液的检验次数总共为$k+1$次．假设在接受检验的血液样本中，每份样本的检验结果是阳性还是阴性都是相互独立的，且每份样本是阳性的概率为$p(0<p<1)$．

（1）假设有6份血液样本，其中只有两份样本为阳性，若采取逐份检验的方式，求恰好经过两次检验就能把阳性样本全部检验出来的概率．

（2）现取其中的$k(k\in N^*$，$2\leqslant k\leqslant n)$份血液样本，记采用逐份检验的方式，样本需要检验的次数为ξ_1；采用混合检验的方式，样本需要检验的总次数为ξ_2．

（ⅰ）若$E\xi_1=E\xi_2$，试运用概率与统计的知识，求p关于k的函数关系$p=f(k)$；

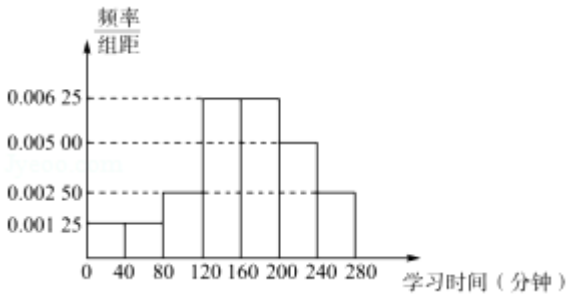
（ⅱ）若$p=1-\frac{1}{\sqrt[4]{e}}$，采用混合检验的方式需要检验的总次数的期望比逐份检验的总次数的期望少，求k的最大值．$(ln4=1.386$，$ln5=1.609$，$ln6=1.792$，$ln7=1.946$，$ln8=2.079$，$ln9=2.197)$

9.（2020•湖北模拟）2020 年寒假期间新冠肺炎肆虐，全国人民众志成城抗击疫情．某市要求全体市民在家隔离，同时决定全市所有学校推迟开学．某区教育局为了让学生“停课不停学”，要求学校各科老师每天在网上授课，每天共 280 分钟，请学生自主学习．区教育局为了了解高三学生网上学习情况，上课几天后在全区高三学生中采取随机抽样的方法抽取了 100 名学生进行问卷调查，为了方便表述把学习时间在 $(0，120]$ 分钟的学生称为 A 类，把学习时间在 $(120，200]$ 分钟的学生称为 B 类，把学习时间在 $(200，280]$ 分钟的学生称为 C 类，随机调查的 100 名学生学习时间的人数频率分布直方图如图所示：以频率估计概率回答下列问题：

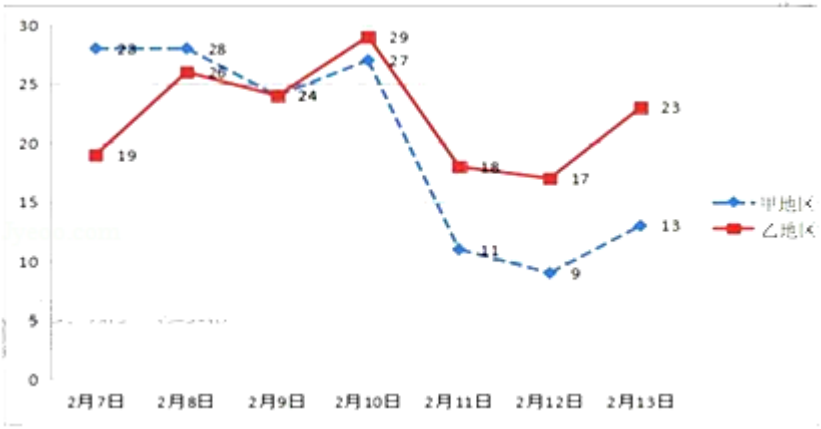
（1）求 100 名学生中 A，B，C 三类学生分别有多少人？

（2）在 A，B，C 三类学生中，按分层抽样的方法从上述 100 个学生中抽取 10 人，并在这 10 人中任意邀请 3 人电话访谈，求邀请的 3 人中是 C 类的学生人数的分布列和数学期望；

（3）某校高三（1）班有 50 名学生，某天语文和数学老师计划分别在 $19:00-19:40$ 和 $20:00-20:40$ 在线上与学生交流，由于受校园网络平台的限制，每次只能 30 个人同时在线学习交流．假设这两个时间段高三（1）班都有 30 名学生相互独立地随机登录参加学习交流．设 ξ 表示参加语文或数学学习交流的人数，当 ξ 为多少时，其概率最大．



10.（2020•东莞市模拟）在党中央的正确领导下，通过全国人民的齐心协力，特别是全体一线医护人员的奋力救治，二月份“新冠肺炎”疫情得到了控制．甲、乙两个地区采取防护措施后，统计了从 2 月 7 日到 2 月 13 日一周的新增“新冠肺炎”确诊人数，绘制成如图折线图：



（1）根据图中甲、乙两个地区折线图的信息，写出你认为最重要的两个统计结论；

（2）治疗“新冠肺炎”药品的研发成了当务之急，某药企计划对甲地区的 A 项目或乙地区的 B 项目投入研发资金，经过评估，对于 A 项目，每投资十万元，一年后利润是 1.38 万元、1.18 万元、1.14 万元的概率分别为 $\frac{1}{6}$、$\frac{1}{2}$、$\frac{1}{3}$；对于 B 项目，利润与产品价格的调整有关，已知 B．项目产品价格在一年内进行 2 次独立的调整，每次价格调整中，产品价格下调的概率都是 $p(0<p<1)$，记 B 项目一年内产品价格的下调次数为 ξ，每投资十万元，ξ 取 0、1、2 时，一年后相应利润是 1.4 万元、1.25 万元、0.6 万元．记对 A 项目投资十万元，一年后利润的随机变量为 ξ_1，记对 B 项目投资十万元，一年后利润的随机变量为 ξ_2．

(i) 求 ξ_1，ξ_2 的概率分布列和数学期望 $E\xi_1$，$E\xi_2$；

(ii) 如果你是投资决策者，将做出怎样的决策？请写出决策理由．

11.（2020•葫芦岛一模）2020 年是具有里程碑意义的一年我们将全面建成小康社会，实现第一个百年奋斗目标；2020 年也是脱贫攻坚决战决胜之年.（总书记二O二O年新年贺词）截至 2018 年底，中国农村贫困人口从 2012 年的 9899 万人减少至 1660 万人，贫困发生率由 2012 年的10.2% 下降至 2018 年的1.7% ；连续 7 年每年减贫规模都在 1000 万人以上；确保到 2020 年农村贫困人口实现脱贫，是我们党立下的军令状，脱贫攻坚越到最后时刻，越要响鼓重锤．某贫困地区截至 2018 年底，按照农村家庭人均年纯收入 8000 元的小康标准，该地区仅剩部分家庭尚未实现小康．现从这些尚未实现小康的家庭中随机抽取 50 户，得到这 50 户家庭 2018 年的家庭人均年纯收入的频率分布直方图.

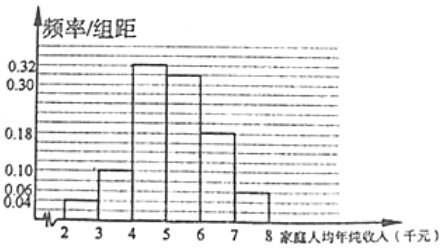
（1）将家庭人均纯年收入不足 5000 元的家庭称为“特困户”若从这 50 户家族中再取出 10 户调查致贫原因，求这 10 户中含有“特困户“的户数 X 的数学期望；

（2）2019 年 7 月，为估计该地能否在 2020 年全面实现小康，统计了该地当时最贫困的一个家庭 2019 年 1 至 6 月的人均月纯收入如表：

月份 /2019（时间代码 x)	1	2	3	4	5	6
人均月纯收入（元)	275	365	415	450	470	485

由散点图及相关性分析发现：家庭人均月纯收入 y 与时间代码 x 之间具有较强的线性相关关系，由此估计该家庭 2020 年能实现小康生活，但 2020 年 1 月突如其来的新冠肺炎疫情影响了奔小康的进展，该家庭 2020 年第一季度 (1，2，3 月份）每月的人均月纯收入人均只有 2019 年 12 月的预估值的$\frac{2}{3}$，为加快脱贫进程，政府加大扶贫力度，拟从 2020 年 3 月份起，以后每月的增长率为 a，为了使 2020 年该家庭顺利迈入小康生活，则 a 至少应为多少？（保留小数点后两位数字）；

①可能用到的数据：$\sum_{i=1}^{6} x_i y_i = 9310, \sqrt{45^2 + 4 \times 120 \times 4} \approx 62.81, 1.15^{10} \approx 4.05$ ②参考公式：线性回归方程 $\hat{y} = \hat{b}x + \hat{a}$ 中，$\hat{b} = \dfrac{\sum_{i=1}^{6} x_i y_i - 6\overline{xy}}{\sum_{i=1}^{6} x_i^2 - 6\overline{x}^2}, \hat{a} = \overline{y} - \hat{b}\overline{x}$ $(1+a)^n \approx 1 + C_n^1 a + C_n^2 a^2 + C_n^3 a^3 (m..10, |a| < 0.15)$



12.（2020•湛江一模）我国全面二孩政策已于 2016 年 1 月 1 日起正式实施．国家统计局发布的数据显示，从 2012 年到 2017 年，中国的人口自然增长率变化始终不大，在 5‰上下波动（如图）为了了解年龄介于 24 岁至 50 岁之间的适孕夫妻对生育二孩的态度如何，统计部门按年龄分为 9 组，每组选取 150 对夫妻进行调查统计有生育二孩意愿的夫妻数，得到如表：

年龄区间	[24，26]	[27，29]	[30，32]	[33，35]	[36，38]	[39，41]	[42，44]	[45，47]	[48，50]
有意愿数	80	81	87	88	84	83	83	70	66

（1）设每个年龄区间的中间值为 x，有意愿数为 y，求样本数据的线性回归直线方程，并求该模型的相关系数 r（结果保留两位小数

（2）从[24， 26]，[33， 35]，[39， 41]，[45， 47]，[48， 50]这五个年龄段中各选出一对夫妻（能代表该年龄段超过半数夫妻的意愿）进一步调研，再从这 5 对夫妻中任选 2 对夫妻．设其中不愿意生育二孩的夫妻数为 X，求 X 的分布列和数学期望．

（参考数据和公式：$r=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2\sum_{i=1}^{n}(y_i-\overline{y})^2}}$，$\hat{b}=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$，$\hat{a}=\overline{y}-\hat{b}\overline{x}$，$\sum_{i=1}^{9}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^{9}x_iy_i-\overline{x}\sum_{i=1}^{9}y_i$，$\sum_{i=1}^{9}x_iy_i=26340$，$\sqrt{224640}\approx473.96$．）



13.（2020•呼和浩特一模）检验中心为筛查某种疾病，需要检验血液是否为阳性，对 $n(n\in N^*)$ 份血液样本，有以下两种检验方式：①逐份检验，需要检验 n 次；②混合检验，即将其中 $k(k\in N^*$ 且 $k..2)$ 份血液样本分别取样混合在一起检验，若检验结果为阴性，这 k 份的血液全为阴性，因而这 k 份血液样本只要检验一次就够了，如果检验结果为阳性，为了明确这 k 份血液究竟哪几份为阳性，再对这 k 份再逐份检验，此时这 k 份血液的检验次数总共为 $k+1$ 次．假设在接受检验的血液样本中，每份样本的检验结果是阳性还是阴性都是独立的，且每份样本是阳性结果的概率为 $p(0<p<1)$．

（Ⅰ）假设有 5 份血液样本，其中只有 2 份样本为阳性，若采用逐份检验方式，求恰好经过 2 次检验就能把阳性样本全部检验出来的概率；

（Ⅱ）现取其中 $k(k\in N^*$ 且 $k..2)$ 份血液样本，记采用逐份检验方式，样本需要检验的总次数为 ξ_1，采用混合检验方式，样本需要检验的总次数为点 ξ_2．当 $p=1-\frac{1}{\sqrt[4]{e}}$ 时，根据 ξ_1 和 ξ_2 的期望值大小，讨论当 k 取何值时，采用逐份检验方式好？（参考数据：$ln2\approx0.69$，$ln3\approx1.10$，$ln5\approx1.61$，$e\approx2.72$，$e^2\approx7.39$，$e^3\approx20.09$．）

14.（2020•德州一模）医院为筛查某种疾病，需要血检，现有 $n(n\in N^*)$ 份血液样本，有以下两种检验方式：

方式一：逐份检验，需要检验 n 次；

方式二：混合检验，把每个人的血样分成两份，取 $k(k..2)$ 个人的血样各一份混在一起进行检验，如果结果是阴性，那么对这 k 个人只作一次检验就够了；如果结果是阳性，那么再对这 k 个人的另一份血样逐份检验，此时这 k 份血液的检验次数总共为 $k+1$ 次.

（1）假设有 6 份血液样本，其中只有 2 份样本为阳性，若采用逐份检验的方式，求恰好经过 3 次检验就能把阳性样本全部检验出来的概率；

（2）假设在接受检验的血液样本中，每份样本的检验结果是阳性还是阴性都是相互独立的，且每份样本是阳性结果的概率为 $p(0<p<1)$．现取其中 $k(k\in N^*$ 且 $k..2)$ 份血液样本，记采用逐份检验方式，样本需要检验的总次数为 X_1，采用混合检验方式，样本需要检验的总次数为 X_2.

①运用概率统计的知识，若 $EX_1=EX_2$，试求 p 关于 k 的函数关系式 $p=f(k)$；

②若 $p=1-e^{-\frac{1}{5}}$，且采用混合检验方式可以使得样本需要检验的总次数的期望值比逐份检验的总次数期望值更少，求 k 的最大值.

参考数据：$ln11\approx 2.3978$，$1n12\approx 2.4849$，$ln13\approx 2.5649$.

15.（2020•宜昌模拟）目前，新冠病毒引发的肺炎疫情在全球肆虐，为了解新冠肺炎传播途径，采取有效防控措施，某医院组织专家统计了该地区 500 名患者新冠病毒潜伏期的相关信息，数据经过汇总整理得到如图所示的频率分布直方图（用频率作为概率）．潜伏期不高于平均数的患者，称为“短潜伏者”，潜伏期高于平均数的患者，称为“长潜伏者”.

（1）求这 500 名患者潜伏期的平均数（同一组中的数据用该组区间的中点值作代表），并计算出这 500 名患者中“长潜伏者”的人数；

（2）为研究潜伏期与患者年龄的关系，以潜伏期是否高于平均数为标准进行分层抽样，从上述 500 名患者中抽取 300 人，得到如下列联表，请将列联表补充完整，并根据列联表判断是否有97.5% 的把握认为潜伏期长短与患者年龄有关：

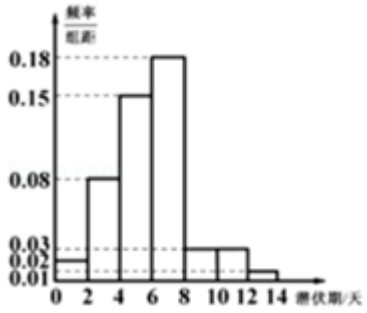
	短潜伏者	长潜伏者	合计
60 岁及以上	90	____	____
60 岁以下	____	____	140
合计	____	____	300

（3）研究发现，有 5 种药物对新冠病毒有一定的抑制作用，其中有 2 种特别有效，现在要通过逐一试验直到把这 2 种特别有效的药物找出来为止，每一次试验花费的费用是 500 元，设所需要的试验费用为 X，求 X 的分布列与数学期望 X．

附表及公式：

$P(K^2..k_0)$	0.15	0.10	0.05	0.025	0.010	0.005	0.001
k_0	2.072	2.706	3.841	5.024	6.635	7.879	10.828

$$K^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

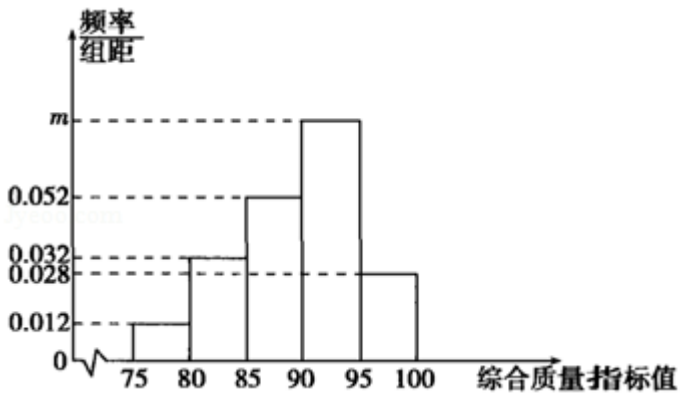


16.（2020•山东模拟）第五届中国“互联网＋”大学生创新创业大赛在浙江大学开幕.10月 15 日，第五届中国“互联网＋”大学生创新创业大赛全国总决赛闭幕．在此次全国总决赛中，高教主赛道共评选出 60 个金奖项目（含 5 个港澳台金奖项目），“青年红色筑梦之旅”赛道、职教赛道和国际赛道各评选出 15 个金奖项目．某公司迅速引进其中的两个项目，在甲，乙两个生产车间进行试验，其中，甲车间只有一条生产线，乙车间有两条生产线，公司在试运营过程中随机选取了甲车间生产的部分产品，测得相应的综合质量指标值为t．根据国家规定，综合质量指标值$t..90$为优等品；$80,,t<90$为合格品；$t<80$为不合格品．甲车间生产的产品综合质量指标值的频率分布直方图如图．

（1）已知每件产品的生产成本与综合质量指标值t的关系式为$P(t)=0.5t$，若优等品的出售价格为 100 元，合格品的出售价格为 80 元，不合格品需要全部销毁，每件产品的销毁费用为 20 元，将每件产品的利润用综合质量指标值t的函数$Q(t)$来表示；

（2）根据预测，甲车间生产线成功生产产品的概率为p，每年可获得利润$\frac{700}{3}$万元；乙车间每条生产线成功生

产产品的概率为$\frac{1}{2}(e^{p}-1)$，每年可获得利润 300 万元，记每年两个车间的利润之和为X，则当p为何值时，X的期望最大？



17.（2020•宁德模拟）某市旅游局为尽快恢复受疫情影响的旅游业，准备在本市的景区推出旅游一卡通（年卡）．为了更科学的制定一卡通的有关条例，市旅游局随机调查了 2019 年到本市景区旅游的 1000 个游客的年旅游消费支出（单位：百元），并制成如图频率分布直方图：

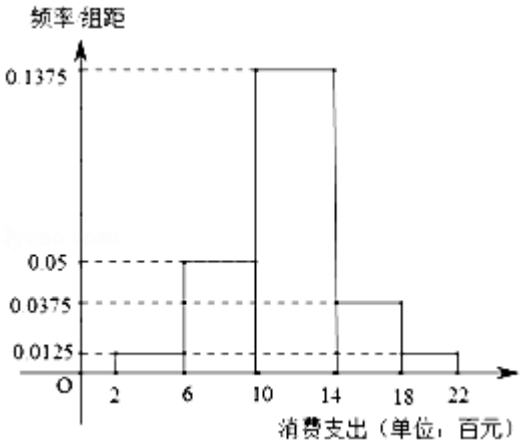
由频率分布直方图，可近似地认为到本市景区旅游的游客，其旅游消费支出服从正态分布 $N(\mu,\ 3.2^2)$，其中 μ 近似为样本平均数 $\bar{x}$（同一组数据用该组区间的中点值作代表）．

（1）若 2019 年到本市景区旅游游客为 500 万人，试估计 2019 年有多少游客在本市的年旅游消费支出不低于 1820 元；

（2）现依次抽取 n 个游客，假设每个游客的旅游消费支出相互独立，记事件 A 表示“连续 3 人的旅游消费支出超出 μ”．若 P_n 表示 $\overline{A}$ 的概率，$P_n = aP_{n-1} + \frac{1}{4}P_{n-2} + bP_{n-3}(n..3,\ a,\ b$ 为常数），且 $P_0 = P_1 = P_2 = 1$．

(i) 求 $P_3,\ P_4$ 及 $a,\ b$；

(ii) 判断并证明数列 $\{P_n\}$ 从第三项起的单调性，试用概率统计知识解释其实际意义．（参考数据：$P(\mu-\sigma<X<\mu+\sigma)\approx 0.6826,\ P(\mu-2\sigma<X<\mu+2\sigma)\approx 0.9544,\ P(\mu-3\sigma<X<\mu+3\sigma)\approx 0.9973$）



18.（2020 春•武昌区校级期中）湖北省从 2021 年开始将全面推行新高考制度，新高考“3+1+2”中的“2”要求考生从政治、化学、生物、地理四门中选两科，按照等级赋分计入高考成绩，等级赋分规则如下：高考政治、化学、生物、地理四门等级考试科目的考生原始成绩从高到低划分为 A，B，C，D，E 五个等级，确定各等级人数所占比例分别为15%，35%，35%，13%，2%，等级考试科目成绩计入考生总成绩时，将 A 至 E 等级内的考生原始成绩，依照等比例转换法分别转换到[86，100]、[71，85]、[56，70]、[41，55]、[30，40]五个分数区间，得到考生的等级分，等级转换分满分为 100 分．具体转换分数区间如表：

等级	A	B	C	D	E
比例	15%	35%	35%	13%	2%
赋分区间	[86，100]	[71，85]	[56，70]	[41，55]	[30，40]

而等比例转换法是通过公式计算：$\frac{Y_2-Y}{Y-Y_1}=\frac{T_2-T}{T-T_1}$

其中 Y_1、Y_2 分别表示原始分区间的最低分和最高分，T_1、T_2 分别表示等级分区间的最低分和最高分，Y 表示原始分，T 表示转换分，当原始分为 Y_1、Y_2 时，等级分分别为 T_1、T_2，假设小明同学的生物考试成绩信息如表：

考试科目	考试成绩	成绩等级	原始分区间	等级分区间
生物	75 分	B 等级	[69，84]	[71，85]

设小明转换后的等级成绩为 T，根据公式得：$\frac{84-75}{75-69}=\frac{85-T}{T-71}$，所以 $T=76.6\approx77$（四舍五入取整），小明最终生物等级成绩为 77 分．已知某学校学生有 60 人选了政治，以期中考试成绩为原始成绩转换该学校选政治的学生的政治等级成绩，其中政治成绩获得 A 等级的学生原始成绩统计如表：

成绩	90	86	81	80	79	78	75
人数	1	2	1	1	2	1	1

（1）从政治成绩获得A等级的学生中任取 3 名，求至少有 2 名同学的等级成绩不小于 93 分的概率；

（2）从政治成绩获得A等级的学生中任取 4 名，设 4 名学生中等级成绩不小于 93 分人数为ξ，求ξ的分布列和期望．

19．（2020•重庆模拟）手工艺是一种生活态度和对传统的坚持，在我国有很多手工艺品制作村落，村民的手工技艺世代相传，有些村落制造出的手工艺品不仅全国闻名，还大量远销海外．近年来某手工艺品村制作的手工艺品在国外备受欢迎，该村村民成立了手工艺品外销合作社，为严把质量关，合作社对村民制作的每件手工艺品都请 3 位行家进行质量把关，质量把关程序如下：(i) 若一件手工艺品 3 位行家都认为质量过关，则该手工艺品质量为A级；(ii) 若仅有 1 位行家认为质量不过关，再由另外 2 位行家进行第二次质量把关，若第二次质量把关这 2 位行家都认为质量过关，则该手工艺品质量为B级，若第二次质量把关这 2 位行家中有 1 位或 2 位认为质量不过关，则该手工艺品质量为C级；(iii) 若有 2 位或 3 位行家认为质量不过关，则该手工艺品质量为D级．已知每一次质量把关中一件手工艺品被 1 位行家认为质量不过关的概率为$\frac{1}{3}$，且各手工艺品质量是否过关相互独立．

（1）求一件手工艺品质量为B级的概率；

（2）若一件手工艺品质量为A，B，C级均可外销，且利润分别为 900 元，600 元，300 元，质量为D级不能外销，利润记为 100 元．

①求 10 件手工艺品中不能外销的手工艺品最有可能是多少件；

②记 1 件手工艺品的利润为X元，求X的分布列与期望．

20.（2020•永州三模）为丰富学生课外生活，某市组织了高中生钢笔书法比赛，比赛分两个阶段进行：第一阶段由评委给出所有参赛作品评分，并确定优胜者；第二阶段为附加赛，参赛人员由组委会按规则另行确定．数据统计员对第一阶段的分数进行了统计分析，这些分数 X 都在 $[70，100)$ 内，在以组距为 5 画分数的频率分布直方图（设“$\frac{\text{频率}}{\text{组距}}=Y$”）时，发现 Y 满足 $Y=\begin{cases}\frac{8n-109}{300}, n\le 16\\ \frac{1}{15}-k\frac{1}{20-n}, n>16\end{cases}$，$n\in N^*$，$5n\le X<5(n+1)$．

（1）试确定 n 的所有取值，并求 k；

（2）组委会确定：在第一阶段比赛中低于 85 分的参赛者无缘获奖也不能参加附加赛；分数在 $[95，100)$ 的参赛者评为一等奖；分数在 $[90，95)$ 的同学评为二等奖，但通过附加赛有 $\frac{1}{11}$ 的概率提升为一等奖；分数在 $[85，90)$ 的同学评为三等奖，但通过附加赛有 $\frac{1}{7}$ 的概率提升为二等奖（所有参加附加赛的获奖人员均不降低获奖等级）．已知学生 A 和 B 均参加了本次比赛，且学生 A 在第一阶段评为二等奖．

(i) 求学生 B 最终获奖等级不低于学生 A 的最终获奖等级的概率；

(ii) 已知学生 A 和 B 都获奖，记 A，B 两位同学最终获得一等奖的人数为 ξ，求 ξ 的分布列和数学期望．

21.（2020•重庆模拟）新型冠状病毒属于 β 属的冠状病毒，人群普遍易感，病毒感染者一般有发热咳嗽等临床表现，现阶段也出现无症状感染者. 基于目前的流行病学调查和研究结果，病毒潜伏期一般为 $1-14$ 天，大多数为 $3-7$ 天．为及时有效遏制病毒扩散和蔓延，减少新型冠状病毒感染对公众健康造成的危害，需要对与确诊新冠肺炎病人接触过的人员进行检查．某地区对与确诊患者有接触史的 1000 名人员进行检查，检查结果统计如下：

	发热且咳嗽	发热不咳嗽	咳嗽不发热	不发热也不咳嗽
确诊患病	200	150	80	30
确诊未患病	150	150	120	120

（1）能否在犯错率不超过 0.001 的情况下，认为新冠肺炎密切接触者有发热症状与最终确诊患病有关.

临界值表：

$P(K^2>k)$	0.10	0.05	0.010	0.005	0.001
k	2.706	3.841	6.645	7.879	10.828

（2）在全国人民的共同努力下，尤其是全体医护人员的辛勤付出下，我国的疫情得到较好控制，现阶段防控重难点主要在境外输入病例和无症状感染者（即无相关临床表现但核酸检测或血清特异性免疫球蛋白 M 抗体检测阳性者）．根据防控要求，无症状感染者虽然还没有最终确诊患 2019 新冠肺炎，但与其密切接触者仍然应当采取居家隔离医学观察 14 天，已知某人曾与无症状感染者密切接触，而且在家已经居家隔离 10 天未有临床症状，若该人员居家隔离第 k 天出现临床症状的概率为 $(\frac{1}{2})^{k-10}$，$(k=11$，12，13，$14)$，两天之间是否出现临床症状互不影响，而且一旦出现临床症状立刻送往医院核酸检查并采取必要治疗，若 14 天内未出现临床症状则可以解除居家隔离，求该人员在家隔离的天数（含有临床症状表现的当天）ξ 的分布列以及数学期望值.（保留小数点后两位）

22.（2020•新乡二模）某公司准备上市一款新型轿车零配件，上市之前拟在其一个下属 4S 店进行连续 30 天的试销，定价为 1000 元/件.

（1）设日销售 40 个零件的概率为 p，记 5 天中恰有 2 天销售 40 个零件的概率为 z，写出 z 关于 p 的函数关系式，并求 z 的极大值点 p_0.

（2）试销结束后统计得到该 4S 店这 30 内的日销售量（单位：件）的数据如表：

日销售量	40	60	80	100
频数	9	12		

其中，有两个数据未给出．试销结束后，这款零件正式上市，每件的定价仍为 1000 元，但生产公司对该款零件不零售，只提供零件的整箱批发，大箱每箱有 55 件，批发价为 550 元/件；小箱每箱有 40 件，批发价为 600 元/件，以这 30 天统计的各日销售量的频率作为试销后各日销售量发生的概率.

该 4S 店决定每天批发两箱，若同时批发大箱和小箱，则先销售小箱内的零件，同时根据公司规定，当天没销售出的零件按批发价的 9 折转给该公司的另一下属 4S 店，假设日销售量为 80 件的概率为 $\frac{p_0}{2}$，其中 P_0 为（1）中 z 的极大值点.

(i) 设该 4S 店批发两大箱，当天这款零件的利润为随机变量 X；批发两小箱，当天这款零件的利润为随机变量 Y，求 EX 和 EY；

(ii) 以日利润的数学期望作为决策依据，该 4S 店每天应该按什么方案批发零件？

23.（2020•山东模拟）“爱国，是人世间最深层、最持久的情感，是一个人立德之源、立功之本.”在中华民族几千年绵延发展的历史长河中，爱国主义始终是激昂的主旋律．爱国汽车公司拟对“东方红”款高端汽车发动机进行科技改造，根据市场调研与模拟，得到科技改造投入 x（亿元）与科技改造直接收益 y（亿元）的数据统计如下：

x	2	3	4	6	8	10	13	21	22	23	24	25
y	13	22	31	42	50	56	58	68.5	68	67.5	66	66

当 $0<x,17$ 时，建立了 y 与 x 的两个回归模型，模型①：$\hat{y}=4.1x+11.8$；模型②：$\hat{y}=21.3x-14.4$．当 $x>17$ 时， y 与 x 满足的线性回归方程为 $\hat{y}=-0.7x+a$.

（1）根据下列表格中的数据，比较当 $0<x,17$ 时模型①和模型②的相关指数 R^2，从而选择拟合精度更高、更

可靠的模型，并据此预测当“东方红”款高端汽车发动机科技改造的投入为 17 亿元时的直接收益；

回归模型	模型①	模型②
回归方程	$\hat{y}=4.1x+11.8$	$\hat{y}=21.3\sqrt{x}-14.4$
$\sum_{i=1}^{n}(y_i-\overline{y_i})^2$	182.4	79.2

（2）为鼓励科技创新，当科技改造投人不少于 20 亿元时，国家给予公司补贴收益 10 亿元，以回归方程为预测

依据，比较科技改造投人 17 亿元与 20 亿元时公司实际收益的大小；

（3）科技改造后，“东方红”款高端汽车发动机的热效率 X 大幅提高， X 服从正态分布 $N(0.52，\ 0.01^2)$，公司对科技改造团队的奖励方案如下：若发动机的热效率不超过 50%，则不予奖励；若发动机的热效率超过 50% 但不超过 53%，则每台发动机奖励 2 万元；若发动机的热效率超过 53%，则每台发动机奖励 5 万元．求

每台发动机获得奖励的数学期望.

附：刻画回归效果的相关指数 $R^2=1-\dfrac{\sum_{i=1}^{n}(y_i-\overline{y_i})^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}$，$\sqrt{17}\approx 4.1$．用最小二乘法求线性回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 的系数公式：$\hat{b}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2}=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$；$\hat{a}=\hat{y}-\hat{b}\overline{x}$．随机变量 ξ 服从正态分布 $N(\mu,\sigma^2)$，则 $P(\mu-\sigma<\xi<\mu+\sigma)=0.6827$，$P(\mu-2\sigma<\xi<\mu+2\sigma)=0.9545$.

24.（2020•运城一模）在创建“全国文明卫生城”过程中，运城市“创城办”为了调查市民对创城工作的了解情况，进行了一次创城知识问卷调查（一位市民只能参加一次），通过随机抽样，得到参加问卷调查的100人的得分统计结果如表所示：.

组别	[30，40)	[40，50)	[50，60)	[60，70)	[70，80)	[80，90)	[90，100)
频数	2	12	20	25	24	13	4

（1）由频数分布表可以大致认为，此次问卷调查的得分 $Z-N(\mu,198)$，μ 近似为这100人得分的平均值（同一组中的数据用该组区间的中点值作代表），利用该正态分布，求 $P(38.2<Z,,80.2)$；

（2）在（1）的条件下，“创城办”为此次参加问卷调查的市民制定如下奖励方案：

①得分不低于 μ 的可以获赠2次随机话费，得分低于 μ 的可以获赠1次随机话费；

②每次获赠的随机话费和对应的概率为：

赠送话费的金额（单位：元）	20	50
概率	$\frac{3}{4}$	$\frac{1}{4}$

现有市民甲参加此次问卷调查，记 X（单位：元）为该市民参加问卷调查获赠的话费，求 X 的分布列与数学期望.

附：参考数据与公式：$\sqrt{198}\approx 14$，若 $X\sim N(\mu,\sigma^2)$，则 $P(\mu-\sigma<X,,\mu+\sigma)=0.6826$；$P(\mu-2\sigma<X,,\mu+2\sigma)=0.9544$，$P(\mu-3\sigma<X,,\mu+3\sigma)=0.9974$.

25.（2020•太原一模）新冠病毒是一种通过飞沫和接触传播的变异病毒，为筛查该病毒，有一种检验方式是检验血液样本相关指标是否为阳性，对于a份血液样本，有以下两种检验方式：一是逐份检验，则需检验n次．二是混合检验，将其中k份血液样本分别取样混合在一起，若检验结果为阴性，那么这k份血液全为阴性，因而检验一次就够了；如果检验结果为阳性，为了明确这k份血液究竟哪些为阳性，就需要对它们再逐份检验，此时k份血液检验的次数总共为$k+1$次．某定点医院现取得 4 份血液样本，考虑以下三种检验方案：方案一，逐个检验；方案二，平均分成两组检验；方案三，四个样本混在一起检验．假设在接受检验的血液样本中，每份样本检验结果是阳性还是阴性都是相互独立的，且每份样本是阴性的概率为$P=\frac{2\sqrt{2}}{3}$．

（Ⅰ）求把 2 份血液样本混合检验结果为阳性的概率；

（Ⅱ）若检验次数的期望值越小，则方案越“优”．方案一、二、三中哪个最“优”？请说明理由．

26．（2020•赣州模拟）春节期间爆发的新型冠状病毒 $(COVID-19)$ 是新中国成立以来感染人数最多的一次疫情，一个不知道自己已感染但处于潜伏期的甲从疫区回到某市过春节，回到家乡后与朋友乙、丙、丁相聚过，最终乙、丙、丁也感染了新冠病毒、可以肯定的是乙受甲感染的，丙是受甲或乙感染的，假设他受甲和受乙感染的概率分别是 0.6 和 0.4，丁是受甲、乙或丙感染的，假设他受甲、乙和丙感染的概率分别是 0.2、0.4 和 0.4．在这种假设之下，乙、丙、丁中直接受甲感染的人数为 X ．

（1）求 X 的分布列和数学期望；

（2）该市在发现在本地出现新冠病毒感染者后，迅速采取应急措施，其中一项措施是各区必须每天及时上报新增疑似病例人数，A 区上报的连续 7 天新增疑似病例数据是“总体均值为 3，中位数 4”， B 区上报的连续 7 天新增疑似病例数据是“总体均值为 2，总体方差为 3“．设 A 区和 B 区连续 7 天上报新增疑似病例人数分别为 x_1 ， x_2 ， ……， x_7 和 y_1 ， y_2 ， ……， y_7 ， x_i 和 $y_i(1\leqslant i\leqslant 7, i\in N)$ 分别表示 A 区和 B 区第 i 天上报新增疑似病例人数（x_i 和 y_i 均为非负）．记 $M=max\{x_1$ ， x_2 ， ……， $x_7\}$ ， $N=max\{y_1$ ， y_2 ， ……， $y_7\}$ ．

①试比较 M 和 N 的大小；

②求 M 和 N 中较小的那个字母所对应的 7 个数有多少组？

27.（2020•山东模拟）2019 年由袁隆平团队研发的第三代杂交水稻 10 月 21 日至 22 日首次公开测产，经测产专家组评定，最终亩产为 1046.3 公斤，第三代杂交水稻的综合优势可以推动我国的水稻生产向更加优质、高产、绿色和可持续方向发展．某企业引进一条先进的食品生产线，计划以第三代杂交水稻为原料进行深加工，创建一个新产品，已知该产品的质量以某项指标值 $k(k\in[70,100])$ 为衡量标准，其质量指标的等级划分如表：

质量指标值 k	$90\leqslant k\leqslant 100$	$85\leqslant k<90$	$80\leqslant k<85$	$75\leqslant k<80$	$70\leqslant k<75$
产品等级	废品	合格	良好	优秀	良好

为了解该产品的生产效益，该企业先进行试生产，从中随机抽取了 1000 件产品，测量了每件产品的指标值，得到产品质量指标值 k 的频率分布直方图如图．

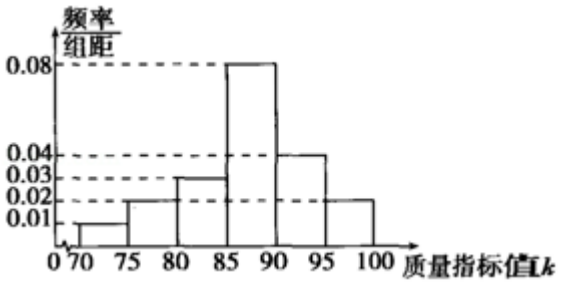
（1）若从质量指标值不小于 85 的产品中利用分层抽样的方法抽取 7 件产品，并采集相关数据进行分析，然后从这 7 件产品中任取 3 件产品，求质量指标值 $k\in[90,\ 95)$ 的件数 X 的分布列及数学期望；

（2）若将频率视为概率，从该产品中有放回地随机抽取 3 件产品，记“抽出的产品中至少有 1 件为合格或合格以上等级“为事件 A，求事件 A 发生的概率；

（3）若每件产品的质量指标值 k 与利润 y（单位：元）的关系如表所示 $(1<t<4)$：

质量指标值 k	$90\leqslant k\leqslant 100$	$85\leqslant k<90$	$80\leqslant k<85$	$75\leqslant k<80$	$70\leqslant k<75$
利润 y（元）	$-e^{t}$	t	$2t$	$4t$	$3t$

请问生产该产品能否盈利？若不能，试说明理由；若能，试确定 t 为何值时，利润达到最大（参考数值：$ln2\approx 0.7$，$ln3\approx 1.1$，$ln5\approx 1.6$）．



28.（2020•聊城一模）2020 年春节期间，武汉市爆发了新型冠状病毒肺炎疫情，在党中央的坚强领导下，全国人民团结一心，众志成城，共同抗击疫情．某中学寒假开学后，为了普及传染病知识，增强学生的防范意识，提高自身保护能力，校委会在全校学生范围内，组织了一次传染病及个人卫生相关知识有奖竞赛（满分 100 分），竞赛奖励规则如下，得分在[70，80)内的学生获三等奖，得分在[80，90)内的学生获二等奖，得分在[90，100]内的学生获一等奖，其他学生不得奖．教务处为了解学生对相关知识的掌握情况，随机抽取了 100 名学生的竞赛成绩，并以此为样本绘制了如下样本频率分布直方图．

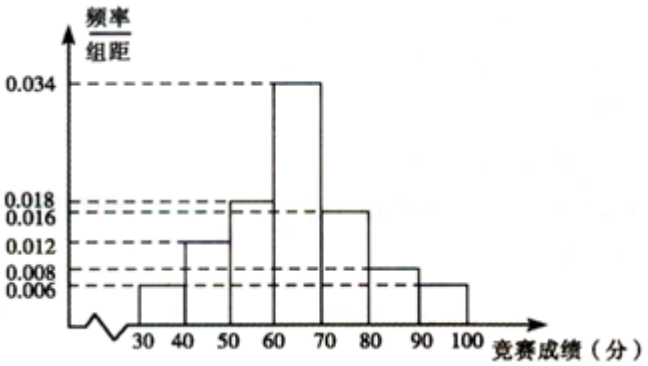
（1）现从该样本中随机抽取两名学生的竞赛成绩，求这两名学生中恰有一名学生获奖的概率；

（2）若该校所有参赛学生的成绩 X 近似服从正态分布 $N(\mu,\sigma^2)$，其中 $\sigma\approx15$，μ 为样本平均数的估计值，利用所得正态分布模型解决以下问题：

(i) 若该校共有 10000 名学生参加了竞赛，试估计参赛学生中成绩超过 79 分的学生数（结果四舍五入到整数）；

(ii) 若从所有参赛学生中（参赛学生数大于 10000）随机抽取 3 名学生进行座谈，设其中竞赛成绩在 64 分以上的学生数为 ξ，求随机变量 ξ 的分布列和均值．

附：若随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，则 $P(\mu-\sigma<X,,\mu+\sigma)\approx0.6827$，$P(\mu-2\sigma<X,,\mu+2\sigma)\approx0.9544$，$P(\mu-3\sigma<X,,\mu+3\sigma)\approx0.9973$．



29.（2020•青岛模拟）某网络购物平台每年 11 月 11 日举行“双十一”购物节，当天有多项优惠活动，深受广大消费者喜爱.

（1）已知该网络购物平台近 5 年“双十”购物节当天成交额如表：

年份	2015	2016	2017	2018	2019
成交额（百亿元）	9	12	17	21	27

求成交额 y（百亿元）与时间变量 x（记 2015 年为 $x=1$，2016 年为 $x=2$，……依此类推）的线性回归方程，并预测 2020 年该平台“双十一”购物节当天的成交额（百亿元）；

（2）在 2020 年“双十一”购物节前，某同学的爸爸、妈妈计划在该网络购物平台．上分别参加 A、B 两店各一个订单的“秒杀”抢购，若该同学的爸爸、妈妈在 A、B 两店订单“秒杀”成功的概率分别为 p、q，记该同学的爸爸和妈妈抢购到的订单总数量为 X．

(i) 求 X 的分布列及 $E(X)$；

(ii) 已知每个订单由 $k(k..2,k\in N^*)$ 件商品 W 构成，记该同学的爸爸和妈妈抢购到的商品 W 总数量为 Y，假设 $p=\dfrac{7\sin\dfrac{\pi}{k}}{4k}-\dfrac{\pi}{k^2},q=\dfrac{\sin\dfrac{\pi}{k}}{4k}$，求 $E(Y)$ 取最大值时正整数 k 的值．

附：回归方程 $\hat{y}=\hat{b}x+a$ 中斜率和截距的最小二乘估计公式分别为：$\hat{b}=\frac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2}=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$；

$a=\overline{y}-\hat{b}\overline{x}$.

30.（2020•深圳二模）2020 年初，新冠肺炎疫情袭击全国，某省由于人员流动性较大，成为湖北省外疫情最严重的省份之一，截至 2 月 29 日，该省已累计确诊 1349 例患者（无境外输入病例）.

（1）为了解新冠肺炎的相关特征，研究人员从该省随机抽取 100 名确诊患者，统计他们的年龄数据，得如表的频数分布表：

年龄	[10，20]	(20，30]	(30，40]	(40，50]	(50，60]	(60，70]	(70，80]	(80，90]	(90，100]
人数	2	6	12	18	22	22	12	4	2

由频数分布表可以大致认为，该省新冠肺炎患者的年龄服从正态分布 $N(\mu，15.2^2)$，其中 μ 近似为这 100 名患者年龄的样本平均数（同一组中的数据用该组区间的中点值作代表）. 请估计该省新冠肺炎患者年龄在 70 岁以上 $(..70)$ 的患者比例；

（2）截至 2 月 29 日，该省新冠肺炎的密切接触者（均已接受检测）中确诊患者约占 10%，以这些密切接触者确诊的频率代替 1 名密切接触者确诊发生的概率，每名密切接触者是否确诊相互独立. 现有密切接触者 20 人，为检测出所有患者，设计了如下方案：将这 20 名密切接触者随机地按 $n(1<n<20$ 且 n 是 20 的约数）个人一组平均分组，并将同组的 n 个人每人抽取的一半血液混合在一起化验，若发现新冠病毒，则对该

组的 n 个人抽取的另一半血液逐一化验，记 n 个人中患者的人数为 X_n，以化验次数的期望值为决策依据，试确定使得 20 人的化验总次数最少的 n 的值.

参考数据：若 $Z \sim N(\mu,\sigma^2)$，则 $P(\mu-\sigma<Z<\mu+\sigma)=0.6826$，$P(\mu-2\sigma<Z<\mu+2\sigma)=0.9544$，$P(\mu-3\sigma<Z<\mu+3\sigma)=0.9973$，$0.9^4\approx0.66$，$0.9^5\approx0.59$，$0.9^{10}\approx0.35$.

参考答案与试题解析

1.（2020•临川区校级模拟）2018 年反映社会现实的电影《我不是药神》引起了很大的轰动，治疗特种病的创新药研发成了当务之急．为此，某药企加大了研发投入，市场上治疗一类慢性病的特效药品 A 的研发费用 x（百万元）和销量 y（万盒）的统计数据如表：

研发费用 x（百万元）	2	3	6	10	13	15	18	21
销量 y（万盒）	1	1	2	2.5	3.5	3.5	4.5	6

（1）根据数据用最小二乘法求出 y 与 x 的线性回归方程 $y=\hat{b}x+\hat{a}$（系数用分数表示，不能用小数）；

（2）该药企准备生产药品 A 的三类不同的剂型 A_1，A_2，A_3，并对其进行两次检测，当第一次检测合格后，才能进行第二次检测．第一次检测时，三类剂型 A_1，A_2，A_3 合格的概率分别为 $\frac{1}{2}$，$\frac{3}{4}$，$\frac{3}{5}$，第二次检测时，三类剂型 A_1，A_2，A_3 合格的概率分别为 $\frac{4}{5}$，$\frac{2}{3}$，$\frac{2}{3}$．两次检测过程相互独立，设经过两次检测后 A_1，A_2，A_3 三类剂型合格的种类数为 X，求 X 的分布列与数学期望．

附：（1）$\hat{b}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{xy}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2},\hat{a}=\overline{y}-\hat{b}\overline{x}$（2）$\sum_{i=1}^{8}x_iy_i=347,\sum_{i=1}^{8}x_i^2=1308$．

【分析】（1）求出样本中心，求出回归直线方程的斜率，求和求解 $\hat{a}$，即可得到回归直线方程．

（2）药品 A 的三类剂型 A_1、A_2、A_3 经过两次检测后合格分别为事件 B_1、B_2、B_3，求出概率，X 可取 0，1，2，3 求出概率，得到 X 的分布列，然后求解期望．

【解答】解：（1）由题意可知 $\overline{x}=\dfrac{2+3+6+10+21+13+15+18}{8}=11$，$\overline{y}=\dfrac{1+1+2+2.5+6+3.5+3.5+4.5}{8}=3$，

由公式 $\hat{b}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{xy}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2}=\dfrac{347-8\times11\times3}{1308-8\times11^2}=\dfrac{83}{340}$，

$\hat{a}=\overline{y}-\hat{b}\overline{x}=3-\dfrac{83}{340}\times11=\dfrac{107}{340}$，

$\therefore y=\dfrac{83}{340}x+\dfrac{107}{340}$．

（2）药品 A 的三类剂型 A_1、A_2、A_3 经过两次检测后合格分别为事件 B_1、B_2、B_3，则 $p(B_1)=\dfrac{1}{2}\times\dfrac{4}{5}=\dfrac{2}{5}$，

$P(B_2)=\dfrac{3}{4}\times\dfrac{2}{3}=\dfrac{1}{2}$，$P(B_3)=\dfrac{3}{5}\times\dfrac{2}{3}=\dfrac{2}{5}$，

由题意，X 可取 0，1，2，3，$p(X=0)=p(\overline{B_1}\,\overline{B_2}\,\overline{B_3})=(1-\dfrac{2}{5})^2(1-\dfrac{1}{2})=\dfrac{9}{50}$，

$p(X=1)=p(B_1\overline{B_2}\,\overline{B_3}+\overline{B_1}B_2\overline{B_3}+\overline{B_1}\,\overline{B_2}B_3)=(1-\frac{2}{5})^2\cdot\frac{1}{2}+(1-\frac{2}{5})\cdot\frac{2}{5}\cdot(1-\frac{1}{2})\cdot 2=\frac{21}{50}$，

$p(X=2)=p(B_1B_2\overline{B_3}+B_1\overline{B_2}B_3+\overline{B_1}B_2B_3)=(\frac{2}{5})^2\cdot(1-\frac{1}{2})+(1-\frac{2}{5})\cdot\frac{2}{5}\cdot\frac{1}{2}\cdot 2=\frac{8}{25}$，

$p(X=3)=p(B_1B_2B_3)=(\frac{2}{5})^2\cdot\frac{1}{2}=\frac{2}{25}$，

$\therefore X$ 的分布列为

X	0	1	2	3
p	$\frac{9}{50}$	$\frac{21}{50}$	$\frac{8}{25}$	$\frac{2}{25}$

$\therefore X$ 的期望为：$EX=0\times\frac{9}{50}+1\times\frac{21}{50}+2\times\frac{8}{25}+3\times\frac{2}{50}=\frac{59}{50}$，

【点评】本题考查离散型随机变量的分布列，以及期望的求法，考查分析问题解决问题的能力，是中档题.

2.（2020•湖北模拟）一只蚂蚁在如图所示的棱长为 1 米的正四面体的棱上爬行，每次当它到达四面体顶点后，会在过此顶点的三条棱中等可能的选择一条棱继续爬行（包含来时的棱），已知蚂蚁每分钟爬行 1 米，$t=0$ 时蚂蚁位于点 A 处.

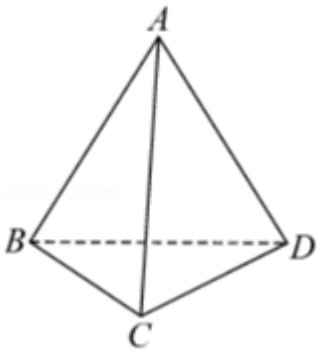
（1）2 分钟末蚂蚁位于哪点的概率最大；

（2）记第 n 分钟末蚂蚁位于点 A，B，C，D 的概率分别为 P_n（A），P_n（B），P_n（C），P_n（D）.

①求证：P_n（B）$=P_n$（C）$=P_n$（D）；

②辰辰同学认为，一段时间后蚂蚁位于点 A、B、C、D 的概率应该相差无几，请你通过计算 10 分钟末蚂蚁位于各点的概率．解释辰辰同学观点的合理性.

附：$(\frac{1}{3})^9\approx 5\times 10^{-5},(\frac{1}{3})^{10}\approx 1.7\times 10^{-5}$，$(\frac{1}{2})^9\approx 1.9\times 10^{-9},(\frac{1}{3})^{10}\approx 9.8\times 10^{-4}$.



【分析】（1）由题可知，在 1 钟末蚂蚁位于 A、B、C、D 点的概率分别为 0，$\frac{1}{3}$，$\frac{1}{3}$，$\frac{1}{3}$，由此求出 2 分钟末位于 A 点的概率和位于 B 的概率及位于 C、D 的概率，从而得以 2 分钟末蚂蚁位于 A 点的概率最大.

（2）①记第 n 分钟末蚂蚁位于 A、B、C、D 点的概率分别为 P_n（A）、P_n（B）、P_n（C）、P_n（D），则

$B_{n+1}=\frac{1}{3}(A_n+C_n+D_n)=\frac{1}{3}(1-B_n)$，同理：$C_{n+1}=\frac{1}{3}(1-C_n)$，相减得$B_{n+1}-C_{n+1}=-\frac{1}{3}(B_n-C_n)B_n-C_n=(B_1-C_1)\cdot(-\frac{1}{3})^{n-1}$，推导出$B_n=C_n=D_n$，由此能证明$P_n$（B）$=P_n$（C）$=P_n$（D）.

②数列$\{A_n-\frac{1}{4}\}$是公比为$-\frac{1}{3}$的等比数列，推导出 10 分钟末蚂蚁位于A、B、C、D点的概率相差无几，第$n(n>10)$分钟末蚂蚁位于A、B、C、D点的概率之差将会更小，所以辰辰的话合理.

【解答】解：（1）解：由题可知，在 1 钟末蚂蚁位于A、B、C、D点的概率分别为 0，$\frac{1}{3}$，$\frac{1}{3}$，$\frac{1}{3}$

故 2 分钟末位于A点的概率$P(A)=\frac{1}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{3}$，

位于B的概率等于$P(B)=\frac{1}{3}\cdot\frac{1}{3}+\frac{1}{3}\cdot\frac{1}{3}=\frac{2}{9}$；

同理，位于C、D的概率也等于$\frac{2}{9}$，

2 分钟末蚂蚁位于A点的概率最大.

（2）①证明：记第n分钟末蚂蚁位于A、B、C、D点的概率分别为P_n（A）、P_n（B）、P_n（C）、P_n（D），

则$B_{n+1}=\frac{1}{3}(A_n+C_n+D_n)=\frac{1}{3}(1-B_n)$，

同理：$C_{n+1}=\frac{1}{3}(1-C_n)$，相减得$B_{n+1}-C_{n+1}=-\frac{1}{3}(B_n-C_n)B_n-C_n=(B_1-C_1)\cdot(-\frac{1}{3})^{n-1}$，

又$B_1=C_1=\frac{1}{3}$，$B_n-C_n=0$，$B_n=C_n$，同理可得$C_n=D_n$，$\therefore B_n=C_n=D_n$，

故P_n（B）$=P_n$（C）$=P_n$（D）.

②解：$\because A_{n+1}=\frac{1}{3}(1-A_n)$，$\therefore A_{n+1}-\frac{1}{4}=-\frac{1}{3}(A_n-\frac{1}{4})$，

$\therefore$数列$\{A_n-\frac{1}{4}\}$是公比为$-\frac{1}{3}$的等比数列，

$A_1-\frac{1}{4}=-\frac{1}{4}$ $A_n-\frac{1}{4}=(-\frac{1}{4})(-\frac{1}{3})^{n-1}$，$A_n=\frac{1}{4}+(-\frac{1}{4})(-\frac{1}{3})^{n-1}$，

$\therefore A_{10}=\frac{1}{4}+(-\frac{1}{4})(-\frac{1}{3})^9$，同理$B_{10}=\frac{1}{4}+(\frac{1}{12})(-\frac{1}{3})^9$，

$\therefore A_{10}-B_{10}=(\frac{1}{4}+(-\frac{1}{4})(-\frac{1}{3})^9)-(\frac{1}{4}+(\frac{1}{12})(-\frac{1}{3})^9)=(-\frac{1}{3})^{10}\approx1.7\times10^{-5}$，

又$B_n=C_n=D_n$，$\therefore$10 分钟末蚂蚁位于A、B、C、D点的概率相差无几，

第$n(n>10)$分钟末蚂蚁位于A、B、C、D点的概率之差将会更小，所以辰辰的话合理.

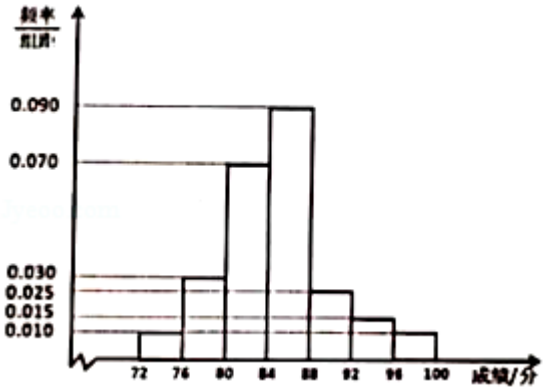
【点评】本题考查概率的求法与证明，考查相互独立事件概率的乘法公式、等比数列的性质等基础知识，考查运算求解能力，是中档题.

3．（2020•江西模拟）年前某市质监部门根据质量管理考核指标对本地的500家食品生产企业进行考核，然后通过随机抽样抽取其中的50家，统计其考核成绩（单位：分），并制成如图频率分布直方图．

（1）求这 50 家食品生产企业考核成绩的平均数 $\bar{x}$（同一组中的数据用该组区间的中点值为代表）及中位数 a（精确到 0.01)

（2）该市质监部门打算举办食品生产企业质量交流会，并从这 50 家食品生产企业中随机抽取 4 家考核成绩不低于 88 分的企业发言，记抽到的企业中考核成绩在 $[92，100]$ 的企业数为 X，求 X 的分布列与数学期望

（3）若该市食品生产企业的考核成绩 X 服从正态分布 $N(\mu,\sigma^2)$ 其中“近似为 50 家食品生产企业考核成绩的平均数 $\bar{x},\sigma^2$ 近似为样本方差，s^2，经计算得 $s^2=27.68$，利用该正态分布，估计该市 500 家食品生产企业质量管理考核成绩高于 90.06 分的有多少家？（结果保留整数）．附参考数据与公式：$\sqrt{27.68}\approx5.26$，$X-N(\mu,\sigma^2)$ 则 $P(\mu-\sigma<X,,\mu+\sigma)\approx0.6827$，$P(\mu-2\sigma<X,,2\mu+\sigma)\approx0.9545$．$P(\mu-3\sigma<X,,\mu+3\sigma)\approx0.9973$．



【分析】（1）利用频率分布直方图的性质能求出这 50 家食品生产企业考核成绩的平均数和中位数．

（2）这 50 家食品生产企业中考核成绩不低于 88 分的企业有 10 家，其中考核成绩在 $[92，100]$ 内的企业有 5 家，X 的可能取值为 0，1，2，3，4，分别求出相应的概率，能求出 X 的分布列和 $E(X)$．

（3）由题意得 $X\sim N(84.80,27.68)$，由此能求出估计该市 500 家食品生产企业质量管理考核成绩高于 90.06 分的有多少家．

【解答】解：（1）这 50 家食品生产企业考核成绩的平均数为：

$\bar{x}=74\times0.04+78\times0.12+82\times0.28+86\times0.36+90\times0.10+94\times0.06+98\times0.04=84.80$（分），

由频率分布直方图得 $a\in[84，88]$ 内，

$\therefore0.04+0.12+0.28+0.09\times(a-84)=0.5$，

解得中位数 $a\approx 84.67$（分）.

（2）这 50 家食品生产企业中考核成绩不低于 88 分的企业有 $50\times(0.1+0.06+0.04)=10$ 家，

其中考核成绩在 [92，100] 内的企业有 $50\times(0.06+0.04)=5$ 家，

$\therefore X$ 的可能取值为 0，1，2，3，4，

$$P(X=0)=\frac{C_5^4}{C_{10}^4}=\frac{1}{42},$$

$$P(X=1)=\frac{C_5^1C_5^3}{C_{10}^4}=\frac{5}{21},$$

$$P(X=2)=\frac{C_5^2C_5^2}{C_{10}^4}=\frac{10}{21},$$

$$P(X=3)=\frac{C_5^3C_5^1}{C_{10}^4}=\frac{5}{21},$$

$$P(X=4)=\frac{C_5^4}{C_{10}^4}=\frac{1}{42},$$

$\therefore X$ 的分布列为：

X	0	1	2	3	4
P	$\frac{1}{42}$	$\frac{5}{21}$	$\frac{10}{21}$	$\frac{5}{21}$	$\frac{1}{42}$

$$E(X)=0\times\frac{1}{42}+1\times\frac{5}{21}+2\times\frac{10}{21}+3\times\frac{5}{21}+4\times\frac{1}{21}=2.$$

（3）由题意得 $X\sim N(84.80,27.68)$，

$\therefore \mu+\approx 84.80+27.68=90.06$，

$\therefore P(X>\mu+\sigma)\approx\frac{1}{2}-\frac{0.6827}{2}\approx 0.1587$，$\therefore 50\times 0.1587\approx 79$（家），

$\therefore$ 估计该市 500 家食品生产企业质量管理考核成绩高于 90.06 分的有 79 家.

【点评】本题考查平均数、中位数、离散型随机变量的分布列、数学期望的求法，考查正态分布的应用，考查频率分布直方图、古典概型等基础知识，考查运算求解能力，是中档题.

4.（2020•厦门一模）小明和爸爸玩亲子游戏，规则如下：袋中装有 3 个大小相同的球，1 个白球，2 个红球，每次摸出一个球，记下颜色后放回，若摸出白球，则下一次由原摸球人继续摸球；若摸出红球，则下一次由对方摸球，规定摸球 m 次，最后一次由谁摸球就算谁获胜，第一次由小明摸球.

（1）求前 3 次摸球中小明恰好摸 2 次的概率；

（2）设第 n 次 $(n,, m)$ 由小明摸球的概率为 P_n，则 $P_1=1$.

(i) 求 P_4；

(ii) 在 $m=19$ 与 $m=20$ 之中选其一，小明应选哪个？（只写结果，不必说明理由！）

【分析】（1）前 3 次摸球中小明恰好摸 2 次，分 2 种情况：小明第 1 次摸到白球，第 2 次摸到红球或者小明第 1 次摸到红球之后，之后爸爸也摸到红球．再结合相互独立事件的概率求出 2 种情况对应的概率，将其相加即可得解；

（2）(i) 第 4 次由小明摸球，共分 4 种情况，然后结合相互独立事件的概率求出 4 种情况对应的概率，将其相加即可得解；

(ii) 小明先摸球，所以选择奇数相对获胜概率更大．

【解答】解：（1）前 3 次摸球中小明恰好摸 2 次，即小明第 1 次摸到白球，第 2 次摸到红球或者小明第 1 次摸到红球之后，之后爸爸也摸到红球，

$\therefore P=\frac{1}{3}\times\frac{2}{3}+\frac{2}{3}\times\frac{2}{3}=\frac{2}{3}$.

（2）(i) 第 4 次由小明摸球的情况有以下几种情况：

①前 3 次小明都摸到白球；

②小明第 1 次摸到红球，之后两次爸爸分别摸到白球和红球；

③小明第 1 次摸到红球，之后爸爸也摸到红球，再由小明摸到白球；

④小明第 1 次摸到白球，第 2 次摸到红球，之后爸爸摸到红球；

$P_4=\frac{1}{3}\times\frac{1}{3}\times\frac{1}{3}+\frac{2}{3}\times\frac{1}{3}\times\frac{2}{3}+\frac{2}{3}\times\frac{2}{3}\times\frac{1}{3}+\frac{1}{3}\times\frac{2}{3}\times\frac{2}{3}=\frac{13}{27}$.

(ii) 应该选择 $m=19$.

【点评】本题考查相互独立事件的概率求法，考查学生的逻辑推理能力和运算能力，属于基础题．

5．（2020•石家庄模拟）某工厂为生产一种精密管件研发了一台生产该精密管件的车床，该精密管件有内外两个口径，监管部门规定“口径误差”的计算方式为：管件内外两个口径实际长分别为 $a(mm)$，$b(mm)$，标准长分别为 $\bar{a}(mm),\bar{b}(mm)$，则“口径误差”为 $|a-\bar{a}|+|b-\bar{b}|$，只要“口径误差”不超过 $0.2min$ 就认为合格，已知这台车床分昼、夜两个独立批次生产．工厂质检部在两个批次生产的产品中分别随机抽取 40 件作为样本，经检测其中昼批次的 40 个样本中有 4 个不合格品，夜批次的 40 个样本中有 10 个不合格品．

(I) 以上述样本的频率作为概率，在昼夜两个批次中分别抽取 2 件产品，求其中恰有 1 件不合格产品的概率；

(II) 若每批次各生产 1000 件，已知每件产品的成本为 5 元，每件合格品的利润为 10 元；若对产品检验，则每件产品的检验费用为 2.5 元；若有不合格品进入用户手中，则工厂要对用户赔偿，这时生产的每件不合

格品工厂要损失 25 元．以上述样本的频率作为概率，以总利润的期望值为决策依据，分析是否要对每个批次的所有产品作检测？

【分析】(I) 先求出昼夜两批次产品各自的不合格率，再分 2 种情况，并结合相互独立事件的概率求解即可；(II) 先求出昼夜两批次各 1000 件产品中合格品的利润，再分不检验和检验 2 种情形，分别求出相应的总利润，比较大小后，即可得解．

【解答】解：(I) 以样本的频率作为概率，则昼批次产品的不合格率为 $\frac{4}{40}=\frac{1}{10}$，夜批次产品的不合格率为 $\frac{10}{40}=\frac{1}{4}$，

在昼夜两个批次中分别抽取 2 件产品，恰有 1 件不合格产品，分 2 种情况：

不合格产品在昼批次中，概率为 $P_1=C_2^1\cdot\frac{1}{10}\cdot\frac{9}{10}\times C_2^2\cdot(\frac{3}{4})^2=\frac{81}{800}$，

不合格产品在夜批次中，概率为 $P_2=C_2^2\cdot(\frac{9}{10})^2\times C_2^1\cdot\frac{1}{4}\cdot\frac{3}{4}=\frac{243}{800}$，

故所求的概率为 $P=P_1+P_2=\frac{81}{200}$．

(II) 这批产品中合格品的利润为 $(1000\times\frac{9}{10}+1000\times\frac{3}{4})\times5=16500$，

若不检验，则总利润为 $W_1=16500-(1000\times\frac{1}{10}+1000\times\frac{1}{4})\times25-10000=-2250$，

若检验，则总利润为 $W_2=16500-2000\times(5+2.5)=1500$，

$\therefore W_2>W_1$，

故需要对每个批次的所有产品作检测．

【点评】本题考查相互独立事件的概率、数学期望的实际应用，考查学生将理论知识与实际生活相联系的能力和运算能力，属于基础题．

6．（2020•广东一模）一支担负勘探任务的队伍有若干个勘探小组和两类勘探人员，甲类人员应用某种新型勘探技术的精准率为 0.6，乙类人员应用这种勘探技术的精准率为 $a(0<a<0.4)$．每个勘探小组配备 1 名甲类人员与 2 名乙类人员，假设在执行任务中每位人员均有一次应用这种技术的机会且互不影响，记在执行任务中每个勘探小组能精准应用这种新型技术的人员数量为 ξ．

（1）证明：在 ξ 各个取值对应的概率中，概率 $P(\xi=1)$ 的值最大．

（2）在特殊的勘探任务中，每次只能派一个勘探小组出发，工作时间不超过半小时，如果半小时内无法完成任务，则重新派另一组出发．现在有三个勘探小组 $A_i(i=1，2，3)$ 可派出，若小组 A_i 能完成特殊任务的概率 t；$t_i=P(\xi=i)(i=1，2，3)$，且各个小组能否完成任务相互独立．试分析以怎样的先后顺序派出勘探

小组，可使在特殊勘探时所需派出的小组个数的均值达到最小.

【分析】（1）每个勘探小组共有 3 名人员，故ξ的所有可能取值为 0，1，2，3，再依据相互独立事件的概率求出每个ξ的取值所对应的概率，并用作差法逐一比较

$P(\xi=1)$与$P(\xi=0)$、$P(\xi=2)$、$P(\xi=3)$的大小关系即可得证；

（2）先根据（1）中的结论比较$P(\xi=2)$和$P(\xi=3)$的大小，可得到$t_1>t_2>t_3$，故而可猜想出结论，再进行证明．证明时，设三个小组$A_i(i=1，2，3)$按照某顺序派出，该顺序下三个小组能完成特殊任务的概率依次为p_1，p_2，p_3，记在特殊勘探时所需派出的小组个数为η，则$\eta=1$，2，3，然后求出η的分布列和数学期望，只需证明数学期望$E(\eta)=3-2p_1-p_2+p_1p_2..3-2t_1-t_2+t_1t_2$成立即可，这一过程采用的是作差法，其中用到了因式分解的相关技巧.

【解答】解：（1）由已知，ξ的所有可能取值为 0，1，2，3，

$P(\xi=0)=(1-0.6)(1-a)^2=0.4(1-a)^2$，

$P(\xi=1)=0.6(1-a)^2+(1-0.6)C_2^1a(1-a)=0.2(1-a)(3+a)$，

$P(\xi=2)=0.6C_2^1a(1-a)+(1-0.6)a^2=0.4a(3-2a)$，

$P(\xi=3)=0.6a^2$.

$\because 0<a<0.4$，

$\therefore P(\xi=1)-P(\xi=0)=0.2(1-a)(1+3a)>0$，

$P(\xi=1)-P(\xi=2)=0.2(3a^2-8a+3)>0$，

$P(\xi=1)-P(\xi=3)=-0.2(4a^2+2a-3)>0$，

$\therefore$概率$P(\xi=1)$的值最大.

（2）由（1）可知，当$0<a<0.4$时，有$t_1=P(\xi=1)$的值最大，

且$t_2-t_3=P(\xi=2)-P(\xi=3)=0.2a(6-7a)>0$，

$\therefore t_1>t_2>t_3$，

$\therefore$应当以A_1，A_2，A_3的顺序派出勘探小组，可使在特殊勘探时所需派出的小组个数的均值达到最小，即优先派出完成任务概率大的小组可减少所需派出的小组个数的均值.

证明如下：

假定 p_1，p_2，p_3 为 t_1，t_2，$t_3(t_1>t_2>t_3)$ 的任意一个排列，即若三个小组 $A_i(i=1，2，3)$ 按照某顺序派出，该顺序下三个小组能完成特殊任务的概率依次为 p_1，p_2，p_3，记在特殊勘探时所需派出的小组个数为 η，则 $\eta=1，2，3$，且 η 的分布列为

η	1	2	3
P	p_1	$(1-p_1)p_2$	$(1-p_1)(1-p_2)$

$\therefore$ 数学期望 $E(\eta)=p_1+2(1-p_1)p_2+3(1-p_1)(1-p_2)=3-2p_1-p_2+p_1p_2$

下面证明 $E(\eta)=3-2p_1-p_2+p_1p_2..3-2t_1-t_2+t_1t_2$ 成立，

$\because(3-2p_1-p_2+p_1p_2)-(3-2t_1-t_2+t_1t_2)$

$=2(t_1-p_1)+(t_2-p_2)+p_1p_2-p_1t_2+p_1t_2-t_1t_2$

$=2(t_1-p_1)+(t_2-p_2)+p_1(p_2-t_2)+t_2(p_1-t_1)$

$=(2-t_2)(t_1-p_1)+(1-p_1)(t_2-p_2)$

$..(1-p_1)(t_1-p_1)+(1-p_1)(t_2-p_2)$

$=(1-p_1)[(t_1+t_2)-(p_1+p_2)]..0$，

$\therefore$ 按照完成任务概率从大到小的 A_1，A_2，A_3 的先后顺序派出勘探小组，可使在特殊勘探时所需派出的小组个数的均值达到最小．

【点评】本题考查相互独立事件的概率、离散型随机变量的分布列和数学期望，以及期望的实际应用等，考查学生对数据的分析能力和运算能力，属于难题．

7．（2020•濮阳一模）2020 年 1 月 10 日，引发新冠肺炎疫情的 $COVID-9$ 病毒基因序列公布后，科学家们便开始了病毒疫苗的研究过程．但是类似这种病毒疫苗的研制需要科学的流程，不是一朝一夕能完成的，其中有一步就是做动物试验．已知一个科研团队用小白鼠做接种试验，检测接种疫苗后是否出现抗体．试验设计是：每天接种一次，3 天为一个接种周期．已知小白鼠接种后当天出现抗体的概率为 $\frac{1}{2}$，假设每次接种后当天是否出现抗体与上次接种无关．

（Ⅰ）求一个接种周期内出现抗体次数 kk 的分布列；

（Ⅱ）已知每天接种一次花费 100 元，现有以下两种试验方案：

①若在一个接种周期内连续 2 次出现抗体即终止本周期试验，进行下一接种周期，试验持续三个接种周期，

设此种试验方式的花费为 X 元；

②若在一个接种周期内出现 2 次或 3 次抗体，该周期结束后终止试验，已知试验至多持续三个接种周期，设此种试验方式的花费为 Y 元.

比较随机变量 X 和 Y 的数学期望的大小.

【分析】（Ⅰ）由题意可知，随机变量 k 服从二项分布 $B(3,\frac{1}{2})$，求出概率，得到 k 的分布列.

（Ⅱ）①设一个接种周期的接种费用为 ξ 元，则 ξ 可能的取值为 200，300，然后求解概率与期望，

②随机变量 Y 可能的取值为 300，600，900，求出概率与期望，即可判断 $E(X)>E(Y)$.

【解答】解：（Ⅰ）由题意可知，随机变量 k 服从二项分布 $B(3,\frac{1}{2})$，故 $P(k)=C_3^k(\frac{1}{2})^k(\frac{1}{2})^{3-k}(k=0,1,2,3)$.

则 k 的分布列为

k	0	1	2	3
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

（Ⅱ）①设一个接种周期的接种费用为 ξ 元，则 ξ 可能的取值为 200，300，

因为 $P(\xi=200)=\frac{1}{4}$，$P(\xi=300)=\frac{3}{4}$，

所以 $E(\xi)=200\times\frac{1}{4}+300\times\frac{3}{4}=275$.

所以三个接种周期的平均花费为 $E(X)=3E(\xi)=3\times275=825$.

②随机变量 Y 可能的取值为 300，600，900，

设事件 A 为“在一个接种周期内出现 2 次或 3 次抗体”，由（Ⅰ）知，$P(A)=\frac{3}{8}+\frac{1}{8}=\frac{1}{2}$.

所以 $P(Y=300)=P(A)=\frac{1}{2}$，$P(Y=600)=[1-P(A)]\times P(A)=\frac{1}{4}$，$P(Y=900)=[1-P(A)]\times[1-P(A)]\times1=\frac{1}{4}$，

所以 $E(Y)=300\times\frac{1}{2}+600\times\frac{1}{4}+900\times\frac{1}{4}=525$.

所以 $E(X)>E(Y)$.

【点评】本题考查二项分布以及离散型随机变量的分布列与数学期望，是中档题.

8.（2020•河南模拟）某中医药研究所研制出一种新型抗癌药物，服用后需要检验血液是否为阳性，现有 $n(n\in N^*)$ 份血液样本，每个样本取到的可能性均等，有以下两种检验方式：（1）逐份检验，则需要检验 n 次；（2）混合检验，将其中 $k(k\in N^*$，$2\leqslant k\leqslant n)$ 份血液样本分别取样混合在一起检验，若结果为阴性，则这 k 份的血液全为阴性，因而这 k 份血液样本只需检验一次就够了：若检验结果为阳性，为了明确这 k 份血液究竟哪份为阳性，就需要对这 k 份再逐份检验，此时这 k 份血液的检验次数总共为 $k+1$ 次．假设在接受检验的

血液样本中，每份样本的检验结果是阳性还是阴性都是相互独立的，且每份样本是阳性的概率为 $p(0<p<1)$．

（1）假设有 6 份血液样本，其中只有两份样本为阳性，若采取逐份检验的方式，求恰好经过两次检验就能把阳性样本全部检验出采的概率．

（2）现取其中的 $k(k\in N^*,\ 2\leqslant k\leqslant n)$ 份血液样本，记采用逐份检验的方式，样本需要检验的次数为 ξ_1；采用混合检验的方式，样本需要检验的总次数为 ξ_2．

（ⅰ）若 $E\xi_1=E\xi_2$，试运用概率与统计的知识，求 p 关于 k 的函数关系 $p=f(k)$；

（ⅱ）若 $p=1-\frac{1}{\sqrt[4]{e}}$，采用混合检验的方式需要检验的总次数的期望比逐份检验的总次数的期望少，求 k 的最大值．$(ln4=1.386,\ ln5=1.609,\ ln6=1.792,\ ln7=1.946,\ ln8=2.079,\ ln9=2.197)$

【分析】（1）设“恰好经过两次检验就能把阳性样本全部检验出来”为事件 A，利用古典概型的概率公式求解即可．

（2）（ⅰ）由题意知 $E\xi_1=k$，ξ_2 取值的可能有 1，$k+1$，求出概率，即可求解期望，列出关系式，推出 p 关于 k 的函数关系 $p=f(k)$；

（ⅱ）由题意知，$E\xi_1>E\xi_2$，得到 $k(1-p)^k>1$．结合 $p=1-e^{-\frac{1}{4}}$，推出 $(e^{-\frac{1}{4}})^k>\frac{1}{k}$，利用对数的运算法则，通过构造法，结合函数的导数，判断函数的单调性转化求解 k 的最大值．

【解答】解：（1）设“恰好经过两次检验就能把阳性样本全部检验出来”为事件 A，则 $P(A)=\frac{A_2^2A_4^4}{A_6^6}=\frac{1}{15}$，

即恰好经过两次检验就能把阳性样本全部检验出来的事件的概率为 $\frac{1}{15}$．

（2）（ⅰ）由题意知 $E\xi_1=k$，ξ_2 取值的可能有 1，$k+1$，$p(\xi_2=1)=(1-p)^k$，$p(\xi_2=k+1)=1-(1-p)^k$，

所以 $E\xi_2=(1-p)^k+(k+1)[1-(1-p)^k]=k+1-k(1-p)^k$，

由 $E\xi_1=E\xi_2$，得 $k=k+1-k(1-p)^k$，即 $(1-p)^k=\frac{1}{k}$，所以 $p=1-(\frac{1}{k})^{\frac{1}{k}}$，

所以 p 关于 k 的函数关系 $p=1-(\frac{1}{k})^{\frac{1}{k}}(k\in N^*,2\leqslant k\leqslant n)$．

（ⅱ）由题意知，$E\xi_1>E\xi_2$，所以 $k>1+k-k(1-p)^k$，即 $k(1-p)^k>1$．

所以 $\frac{1}{k}<(1-p)^k$，又 $p=1-e^{-\frac{1}{4}}$，所以 $(e^{-\frac{1}{4}})^k>\frac{1}{k}$，

两边同时取对数，得 $-\frac{k}{4}>-lnk$，即 $lnk-\frac{k}{4}>0$，

设 $f(x)=lnx-\frac{x}{4}$，则 $f'(x)=\frac{1}{x}-\frac{1}{4}$，易知函数 $f(x)$ 在 $(4,+\infty)$ 上单调递减，

f（8）$=ln8-2=2.079-2=0.079>0$，$f(9)=ln9-\frac{9}{4}=2.197-2.25<0$，

所以 k 的最大值为 8．

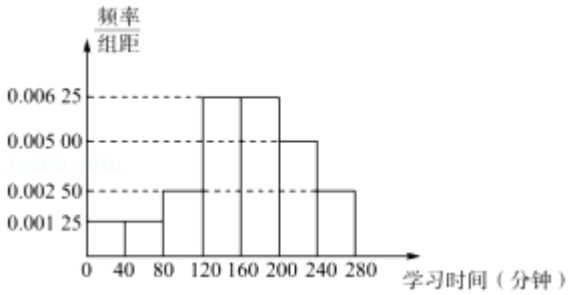
【点评】本题考查离散型随机变量的期望的求法，古典概型概率的求法，构造法的应用，函数的导数的应用，考查分析问题解决问题的能力．

9．（2020•湖北模拟）2020 年寒假期间新冠肺炎肆虐，全国人民众志成城抗击疫情．某市要求全体市民在家隔离，同时决定全市所有学校推迟开学．某区教育局为了让学生“停课不停学”，要求学校各科老师每天在网上授课，每天共 280 分钟，请学生自主学习．区教育局为了了解高三学生网上学习情况，上课几天后在全区高三学生中采取随机抽样的方法抽取了 100 名学生进行问卷调查，为了方便表述把学习时间在 $(0，120]$ 分钟的学生称为 A 类，把学习时间在 $(120，200]$ 分钟的学生称为 B 类，把学习时间在 $(200，280]$ 分钟的学生称为 C 类，随机调查的 100 名学生学习时间的人数频率分布直方图如图所示：以频率估计概率回答下列问题：

（1）求 100 名学生中 A，B，C 三类学生分别有多少人？

（2）在 A，B，C 三类学生中，按分层抽样的方法从上述 100 个学生中抽取 10 人，并在这 10 人中任意邀请 3 人电话访谈，求邀请的 3 人中是 C 类的学生人数的分布列和数学期望；

（3）某校高三（1）班有 50 名学生，某天语文和数学老师计划分别在 $19:00-19:40$ 和 $20:00-20:40$ 在线上与学生交流，由于受校园网络平台的限制，每次只能 30 个人同时在线学习交流．假设这两个时间段高三（1）班都有 30 名学生相互独立地随机登录参加学习交流．设 ξ 表示参加语文或数学学习交流的人数，当 ξ 为多少时，其概率最大．



【分析】（1）利用分层抽样方法能求出 100 名学生中 A，B，C 三类学生分别有多少人．

（2）$A:B:C=20:50:30$，故从 A 类学生中抽 2 人，B 类学生中抽 5 人，C 类学生中抽 3 人，设邀请的三人中是 C 类的学生人数为 X，则 X 可取 0，1，2，3，分别求出相应的概率，由此能求出 X 的分布列和 EX．

（3）学生随机独立参加语文或数学在线辅导所包含怕基本哈拉叫数为 $(C_{50}^{30})^2$，当 $\xi=k$ 时，由韦恩图，得只

参加语文辅导的人数为$k-30$，只参加数学辅导的人数为$k-30$，语文和数学都参加辅导的人数为$60-k$，事件$(\xi=k)$所包含的基本事件总数为$C_{50}^{30}C_{30}^{k-30}C_{30}^{k-30}$，$P(\xi=k)=\frac{C_{50}^{30}C_{30}^{k-30}C_{30}^{k-30}}{(C_{50}^{30})^2}=\frac{C_{30}^{k-30}C_{30}^{k-30}}{C_{50}^{30}}$最大，则$\begin{cases}P(\xi=k)..P(\xi=k+1)\\P(\xi=k)..P(\xi=k-1)\end{cases}$，由此能求出结果.

【解答】解：（1）A类学生有：$(0.00125\times80+0.0025\times40)\times100=20$人，

B类学生有：$0.00625\times80\times100=50$人，

C类学生有：$(0.005\times40+0.0025\times40)\times100=30$人.

（2）$A:B:C=20:50:30$，故从A类学生中抽2人，B类学生中抽5人，C类学生中抽3人，

设邀请的三人中是C类的学生人数为X，则X可取0，1，2，3，

$$P(X=0)=\frac{C_7^3}{C_{10}^3}=\frac{7}{24},$$

$$P(X=1)=\frac{C_3^1C_7^2}{C_{10}^3}=\frac{21}{40},$$

$$P(X=2)=\frac{C_3^2C_7^1}{C_{10}^3}=\frac{7}{40},$$

$$P(X=3)=\frac{C_3^3}{C_{10}^3}=\frac{1}{120},$$

$\therefore X$的分布列为：

X	0	1	2	3
P	$\frac{7}{24}$	$\frac{21}{40}$	$\frac{7}{40}$	$\frac{1}{120}$

$$\therefore EX=0\times\frac{7}{24}+1\times\frac{21}{24}+2\times\frac{7}{40}+3\times\frac{1}{120}=\frac{9}{10}.$$

（3）学生随机独立参加语文或数学在线辅导所包含怕基本哈拉叫数为$(C_{50}^{30})^2$，

当$\xi=k$时，由韦恩图，得只参加语文辅导的人数为$k-30$，只参加数学辅导的人数为$k-30$，

语文和数学都参加辅导的人数为$60-k$，

事件$(\xi=k)$所包含的基本事件总数为$C_{50}^{30}C_{30}^{k-30}C_{30}^{k-30}$，

$$\therefore P(\xi=k)=\frac{C_{50}^{30}C_{30}^{k-30}C_{30}^{k-30}}{(C_{50}^{30})^2}=\frac{C_{30}^{k-30}C_{30}^{k-30}}{C_{50}^{30}}\text{最大},$$

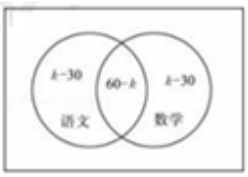
则 $\begin{cases}P(\xi=k)\ge P(\xi=k+1)\\P(\xi=k)\ge P(\xi=k-1)\end{cases}$，

$\therefore\begin{cases}C_{30}^{k-30}C_{30}^{k-30}\ge C_{30}^{k-20}C_{30}^{k-30}\\C_{20}^{k-30}C_{20}^{k-30}\ge C_{30}^{k-30}C_{30}^{k-30}\end{cases}$，

$\therefore\begin{cases}(k-29)^2\ge(60-k)(50-k)\\(61-k)(51-k)\ge(k-30)^2\end{cases}$，

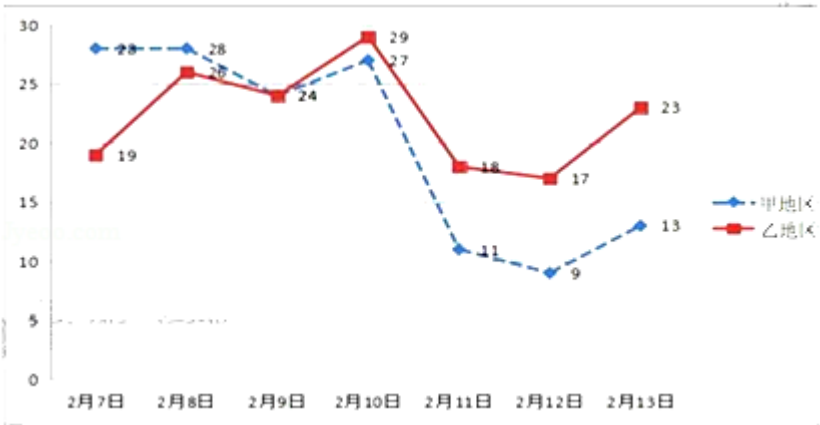
解得 $41\frac{27}{52}\le k\le 42\frac{27}{52}$，

$\because k\in N^*$，$\therefore k=42$．



【点评】本题主要考查频率分布直方图、古典概型、离散型随机变量分布列、超几何分布等基础知识与基本技能．意在考查考生的数据处理能力、数学建模能力、运算求解能力和应用意识．

10．（2020•东莞市模拟）在党中央的正确领导下，通过全国人民的齐心协力，特别是全体一线医护人员的奋力救治，二月份“新冠肺炎”疫情得到了控制．甲、乙两个地区采取防护措施后，统计了从 2 月 7 日到 2 月 13 日一周的新增“新冠肺炎”确诊人数，绘制成如图折线图：



（1）根据图中甲、乙两个地区折线图的信息，写出你认为最重要的两个统计结论；

（2）治疗“新冠肺炎”药品的研发成了当务之急，某药企计划对甲地区的 A 项目或乙地区的 B 项目投入研发资金，经过评估，对于 A 项目，每投资十万元，一年后利润是 1.38 万元、1.18 万元、1.14 万元的概率分别为 $\frac{1}{6}$、$\frac{1}{2}$、$\frac{1}{3}$；对于 B 项目，利润与产品价格的调整有关，已知 B．项目产品价格在一年内进行 2 次独立的调整，每次价格调整中，产品价格下调的概率都是 $p(0<p<1)$，记 B 项目一年内产品价格的下调次数

为ξ，每投资十万元，ξ取 0、1、2 时，一年后相应利润是 1.4 万元、1.25 万元、0.6 万元．记对A项目投资十万元，一年后利润的随机变量为ξ_1，记对B项目投资十万元，一年后利润的随机变量为ξ_2．

(i)求ξ_1，ξ_2的概率分布列和数学期望$E\xi_1$，$E\xi_2$；

(ii)如果你是投资决策者，将做出怎样的决策？请写出决策理由．

【分析】（1）根据图中甲、乙两个地区折线图的信息，得到①甲地区比乙地区的新增人数的平均数低．②甲地区比乙地区的方差大．

（2）(i)求出ξ_1的概率分布列，得到$E\xi_1=1.2$，由题意得$\xi\sim B(2,p)$，求出ξ的概率分布列，再由由题意得下调次数ξ和利润ξ_2的关系求出ξ_2的概率分布列和$E\xi_2=-0.5p^2-0.3p+1.4$．

(ii)当$E\xi_1<E\xi_2$，解得$0<p<\frac{2}{5}$，当$E\xi_1=E\xi_2$时，$p=\frac{2}{5}$．当$E\xi_1>E\xi_2$时，$\frac{2}{5}<p<1$，从而当$0<p<\frac{2}{5}$时，投资B项目；当$p=\frac{2}{5}$时，两个项目都可以；当$\frac{2}{5}<p<1$时，投资A项目．

【解答】解：（1）根据图中甲、乙两个地区折线图的信息，得到：

①甲地区比乙地区的新增人数的平均数低．

②甲地区比乙地区的方差大．

（2）(i)由题意得ξ_1的概率分布列为：

ξ_1	1.38	1.18	1.14
P	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

$\therefore E\xi_1=1.38\times\frac{1}{6}+1.18\times\frac{1}{2}+1.14\times\frac{1}{3}=1.2$，

由题意得$\xi\sim B(2,p)$，即ξ的概率分布列为：

ξ	0	1	2
P	$(1-p)^2$	$2p(1-p)$	p^2

由题意得下调次数ξ和利润ξ_2的关系为：

ξ	0	1	2
ξ_2	1.4	1.25	0.6

$\therefore\xi_2$的概率分布列为：

ξ_2	1.4	1.25	0.6
P	$(1-p)^2$	$2p(1-p)$	p^2

$\therefore E\xi_2 = 1.4\times(1-p)^2 + 1.25\times 2p(1-p) + 0.6\times p^2 = -0.5p^2 - 0.3p + 1.4$.

(ii) 当 $E\xi_1 < E\xi_2$，得 $1.2 < -0.5p^2 - 0.3p + 1.4$，即 $5p^2 + 3p - 2 < 0$，

解得 $0 < p < \frac{2}{5}$，

当 $E\xi_1 = E\xi_2$ 时，$p = \frac{2}{5}$．当 $E\xi_1 > E\xi_2$ 时，$\frac{2}{5} < p < 1$，

$\therefore$ 当 $0 < p < \frac{2}{5}$ 时，投资 B 项目；当 $p = \frac{2}{5}$ 时，两个项目都可以；

当 $\frac{2}{5} < p < 1$ 时，投资 A 项目．

【点评】本题考查离散型随机变量的分布列、数学期望的求法及应用，考查二项分布、折线图性质等基础知识，考查运算求解能力，是中档题．

11．（2020•葫芦岛一模）2020 年是具有里程碑意义的一年我们将全面建成小康社会，实现第一个百年奋斗目标；2020 年也是脱贫攻坚决战决胜之年．（总书记二 O 二 O 年新年贺词）截至 2018 年底，中国农村贫困人口从 2012 年的 9899 万人减少至 1660 万人，贫困发生率由 2012 年的10.2% 下降至 2018 年的1.7%；连续 7 年每年减贫规模都在 1000 万人以上；确保到 2020 年农村贫困人口实现脱贫，是我们党立下的军令状，脱贫攻坚越到最后时刻，越要响鼓重锤．某贫困地区截至 2018 年底，按照农村家庭人均年纯收入 8000 元的小康标准，该地区仅剩部分家庭尚未实现小康．现从这些尚未实现小康的家庭中随机抽取 50 户，得到这 50 户家庭 2018 年的家庭人均年纯收入的频率分布直方图．

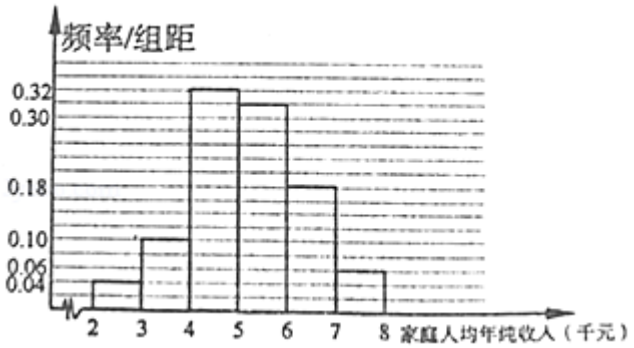
（1）将家庭人均纯年收入不足 5000 元的家庭称为“特困户”若从这 50 户家族中再取出 10 户调查致贫原因，求这 10 户中含有“特困户“的户数 X 的数学期望；

（2）2019 年 7 月，为估计该地能否在 2020 年全面实现小康，统计了该地当时最贫困的一个家庭 2019 年 1 至 6 月的人均月纯收入如表：

月份 /2019（时间代码 x)	1	2	3	4	5	6
人均月纯收入（元)	275	365	415	450	470	485

由散点图及相关性分析发现：家庭人均月纯收入 y 与时间代码 x 之间具有较强的线性相关关系，由此估计该家庭 2020 年能实现小康生活，但 2020 年 1 月突如其来的新冠肺炎疫情影响了奔小康的进展，该家庭 2020 年第一季度(1，2，3 月份）每月的人均月纯收入人均只有 2019 年 12 月的预估值的 $\frac{2}{3}$，为加快脱贫进程，政府加大扶贫力度，拟从 2020 年 3 月份起，以后每月的增长率为 a，为了使 2020 年该家庭顺利迈入小康生活，则 a 至少应为多少？（保留小数点后两位数字）；

①可能用到的数据：$\sum_{i=1}^{6} x_i y_i = 9310, \sqrt{45^2 + 4 \times 120 \times 4} \approx 62.81, 1.15^{10} \approx 4.05$ ②参考公式：线性回归方程 $\hat{y} = \hat{b}x + \hat{a}$ 中，$\hat{b} = \frac{\sum_{i=1}^{6} x_i y_i - 6\overline{xy}}{\sum_{i=1}^{6} x_i^2 - 6\overline{x}^2}, \hat{a} = \overline{y} - \hat{b}\overline{x}$，$(1+a)^n \approx 1 + C_n^1 a + C_n^2 a^2 + C_n^3 a^3 (m..10, |a| < 0.15)$



【分析】（1）由频率分布直方图求出各个收人段的户数，共计 50 户，求出家庭人均年收入不足 5000 元的特困户有：$2+5+16=23$ 户；判断 X 满足参数为 50，23，10 的超几何分布，求出期望 $EX = 10 \times \frac{23}{50} = 4.6$ 户；这 10 户中含有“特困户”的户数 X 的数学期望为 4.6（户）；

（2）从 2020 年 3 月份起，每月的增长率为 a，设从开始 3 月份到 12 月份的月纯收入之和为 S_{10}，得到 $\frac{(1+a)10-1}{a}..14$，

显然 S_{10} 是以 a 为自变量的增函数，然后通过①当 $a..0.15$ 时，②当 $0 < a < 0.15$ 时，求解即可．

【解答】（1）解：由频率分布直方图可知，

家庭人均年收入在 $[2000, 3000)$ 元的家庭数为：$0.04 \times 50 = 2$ 户；

家庭人均年收入在 $[3000, 4000)$ 元的家庭数为：$0.10 \times 50 = 5$ 户；

家庭人均年收入在 $[4000, 5000)$ 元的家庭数为：$0.32 \times 50 = 16$ 户；

家庭人均年收入在 $[5000, 6000)$ 元的家庭数为：$0.30 \times 50 = 15$ 户；

家庭人均年收入在 $[6000, 7000)$ 元的家庭数为：$0.18 \times 50 = 9$ 户；

家庭人均年收入在 $[7000, 8000)$ 元的家庭数为：$0.06 \times 50 = 3$ 户；

共计 50 户，其中家庭人均年收入不足 5000 元的特困户有：$2+5+16=23$ 户；

由题意：X 满足参数为 50，23，10 的超几何分布，所以 $EX=10\times\frac{23}{50}=4.6$ 户；

即这 10 户中含有“特困户”的户数 X 的数学期望为 4.6（户）；

（2）令 $x=12$，则可知 2019 年 12 月份该家庭人均月纯收入为 40 $x12+270=750$（元），

由此可知 2020 年第一季度(1 月份，2 月份，3 月份）该家庭人均月纯收入为 750 $x\frac{2}{3}=500$（元），

∵从 2020 年 3 月份起，每月的增长率为 a，设从开始 3 月份到 12 月份的月纯收入之和为 S_{10}，

则 $S_{10}=500+500$ $x(1+a)+500$ $x(1+a)^2+\ldots+500$ $x(1+a)^9=\frac{500x[1-(1+a)10]}{1-(1+a)}=\frac{500x[(1+a)10-1]}{a}$，

由题意应有：$\therefore 500+500+\frac{500x[(1+a)10-1]}{a}..8000$，

即：$\frac{(1+a)10-1}{a}..14$，

显然 S_{10} 是以 a 为自变量的增函数，$\therefore\frac{(1+a)10-1}{a}$ 是以 a 为自变量的增函数；

①当 $a..0.15$ 时，$\frac{(1+a)10-1}{a}..\frac{1.1510-1}{0.15}\approx\frac{3.05}{0.15}=20\frac{1}{3}>14$ 显然成立，

②当 $0<a<0.15$ 时，$(1+a)^{10}\approx1+C_{10}^1a+C_{10}^2a^2+C_{10}^3a^3=1+10a+45a^2+120a^3$

$\therefore$ $\frac{(1+a)10-1}{a}..14$ $10+45a+120a^2$厖$14\Leftrightarrow120a^2+45a-4$ $0\Leftrightarrow a?\frac{-45-\sqrt{452+4x120x4}}{240}$（舍）或 $a..\frac{-45+\sqrt{452+4x120x4}}{240}\approx\frac{-45+62.81}{240}=\frac{17.81}{240}\approx0.074$

$\therefore$为了使 2020 年该家庭顺利迈入小康生活，则 a 至少应为 0.08(8%)．

【点评】本题考查随机变量的期望的求法，回归方程的应用，考查转化思想以及计算能力，是中档题．

12．（2020•湛江一模）我国全面二孩政策已于 2016 年 1 月 1 日起正式实施．国家统计局发布的数据显示，从 2012 年到 2017 年，中国的人口自然增长率变化始终不大，在 5‰上下波动（如图）为了了解年龄介于 24 岁至 50 岁之间的适孕夫妻对生育二孩的态度如何，统计部门按年龄分为 9 组，每组选取 150 对夫妻进行调查统计有生育二孩意愿的夫妻数，得到如表：

年龄区间	[24，26]	[27，29]	[30，32]	[33，35]	[36，38]	[39，41]	[42，44]	[45，47]	[48，50]
有意愿数	80	81	87	88	84	83	83	70	66

（1）设每个年龄区间的中间值为 x，有意愿数为 y，求样本数据的线性回归直线方程，并求该模型的相关系数 r（结果保留两位小数

（2）从[24， 26]，[33， 35]，[39， 41]，[45， 47]，[48， 50]这五个年龄段中各选出一对夫妻（能代表该年龄段超过半数夫妻的意愿）进一步调研，再从这 5 对夫妻中任选 2 对夫妻．设其中不愿意生育二孩的夫妻数为 X，求 X 的分布列和数学期望．

（参考数据和公式：$r=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2\sum_{i=1}^{n}(y_i-\overline{y})^2}}$，$\hat{b}=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$，$\hat{a}=\overline{y}-\hat{b}\overline{x}$，$\sum_{i=1}^{9}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^{9}x_iy_i-\overline{x}\sum_{i=1}^{9}y_i$，$\sum_{i=1}^{9}x_iy_i=26340$，$\sqrt{224640}\approx473.96$．）



【分析】（1）根据图中数据代入线性回归方程公式和相关系数公式求出结果即可；

（2）根据题意，在[24， 26]，[33， 35]，[39， 41]这三个年龄段中，超过半数的夫妻由生育二孩的意愿，在[45， 47]，[48， 50]这两个年龄段中，超过半数的夫妻没有生育二孩的意愿，

所以从 5 对夫妻中任选 2 对，其中不愿意生育二孩的夫妻数 $X=0$，1，2，再求出 X 的概率和分布列，求出数学期望即可．

【解答】解：（1）根据题意可得，$\overline{x}=\dfrac{25+28+31+34+37+40+43+46+49}{7}=37$，

$\sum_{i=1}^{9}y_i=720$，$\overline{y}=80$，$\sum_{i=1}^{9}x_iy_i=26340$，

$\sum_{i=1}^{9}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^{9}x_iy_i-\overline{x}\sum_{i=1}^{9}y_i=26340-37\times720=-300$，

又 $\sum_{i=1}^{9}(x_i-\overline{x})^2=540$，$\sum_{i=1}^{9}(y_i-\overline{y})^2=416$，

所以 $\sum_{i=1}^{9}(x_i-\overline{x})^2\sum_{i=1}^{9}(y_i-\overline{y})^2=540\times416=224640$，

所以 $\hat{b}=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}=\dfrac{-300}{540}=-\dfrac{5}{9}$，

$\hat{a}=\overline{y}-\hat{b}\overline{x}=80+\dfrac{5}{9}\times37\approx100.56$，

故线性回归方程为：$y=-\frac{5}{9}x+100.56$，

相关系数$r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}=\frac{-300}{\sqrt{224640}}\approx\frac{-300}{473.96}\approx-0.63$；

（2）根据题意，在[24，26]，[33，35]，[39，41]这三个年龄段中，超过半数的夫妻由生育二孩的意愿，在[45，47]，[48，50]这两个年龄段中，超过半数的夫妻没有生育二孩的意愿，

所以从5对夫妻中任选2对，其中不愿意生育二孩的夫妻数$X=0$，1，2，

其中$P(X=0)=\frac{C_2^0C_3^2}{C_5^2}=0.3$，$P(X=1)=\frac{C_2^1C_3^1}{C_5^2}=0.6$，$P(X=2)=\frac{C_2^2}{C_5^2}=0.1$，

X的分布列如下：

X	0	1	2
P	0.3	0.6	0.1

$EX=0\times0.3+1\times0.6+2\times0.1=0.8$．

【点评】本题考查了求线性回归方程和相关系数，考查了超几何分布求分布列和数学期望，考查运算能力和实际应用能力，中档题．

13．（2020•呼和浩特一模）检验中心为筛查某种疾病，需要检验血液是否为阳性，对$n(n\in N^*)$份血液样本，有以下两种检验方式：①逐份检验，需要检验n次；②混合检验，即将其中$k(k\in N^*$且$k..2)$份血液样本分别取样混合在一起检验，若检验结果为阴性，这k份的血液全为阴性，因而这k份血液样本只要检验一次就够了，如果检验结果为阳性，为了明确这k份血液究竟哪几份为阳性，再对这k份再逐份检验，此时这k份血液的检验次数总共为$k+1$次．假设在接受检验的血液样本中，每份样本的检验结果是阳性还是阴性都是独立的，且每份样本是阳性结果的概率为$p(0<p<1)$．

（Ⅰ）假设有5份血液样本，其中只有2份样本为阳性，若采用逐份检验方式，求恰好经过2次检验就能把阳性样本全部检验出来的概率；

（Ⅱ）现取其中$k(k\in N*$且$k..2)$份血液样本，记采用逐份检验方式，样本需要检验的总次数为ξ_1，采用混合检验方式，样本需要检验的总次数为点ξ_2．当$p=1-\frac{1}{\sqrt[4]{e}}$时，根据$\xi_1$和$\xi_2$的期望值大小，讨论当$k$取何值时，采用逐份检验方式好？（参考数据：$ln2\approx0.69$，$ln3\approx1.10$，$ln5\approx1.61$，$e\approx2.72$，$e^2\approx7.39$，$e^3\approx20.09$．）

【分析】（1）记恰好经过2次检验就能把阳性样本全部检验出来为A事件，求解概率．

（2）$E(\xi_1)=k$，ξ_2 的取值为 1，$k+1$，求出概率与期望，通过 $k+1-ke^{\frac{k}{4}}>k$，即 $lnk-\frac{k}{4}<0$．设 $f(x)=lnx-\frac{x}{4}$，利用导数与函数的单调性，转化求解即可．

【解答】解：（1）记恰好经过 2 次检验就能把阳性样本全部检验出来为 A 事件，

则 $P(A)=\frac{A_2^2}{A_5^2}=\frac{1}{10}$．

（2）$E(\xi_1)=k$，ξ_2 的取值为 1，$k+1$，

计算 $P(\xi_2=1)=(1-p)^k$，$P(\xi_2=k+1)=1-(1-p)^k$，

所以 $E(\xi_2)=(1-p)^k+(k+1)(1-(1-p)^k)=k+1-k(1-p)^k$，

又 $p=1-e^{-\frac{1}{4}}$，$E(\xi_2)=k+1-ke^{\frac{k}{4}}$，所以 $k+1-ke^{\frac{k}{4}}>k$，

即 $lnk-\frac{k}{4}<0$．

设 $f(x)=lnx-\frac{x}{4}$，$f'(x)=\frac{1}{x}-\frac{1}{4}=\frac{4-x}{4x}$，$x>0$，

当 $x\in(0,4)$ 时，$f'(x)>0$，$f(x)$ 在 $(0,4)$ 上单调递增；

当 $x\in(4,+\infty)$ 时，$f'(x)<0$，$f(x)$ 在 $(4,+\infty)$ 上单调递减．

且 f（8）$=ln8-2=3ln2-2>0$，$f(9)=ln9-\frac{9}{4}=2ln3-\frac{9}{4}<0$，

所以 k 的取值大于等于 9 时采用逐份检验方式好．

【点评】本题考查期望的求法，函数的导数的应用，考查转化思想以及计算能力，是中档题．

14.（2020•德州一模）医院为筛查某种疾病，需要血检，现有 $n(n\in N^*)$ 份血液样本，有以下两种检验方式：

方式一：逐份检验，需要检验 n 次；

方式二：混合检验，把每个人的血样分成两份，取 $k(k..2)$ 个人的血样各一份混在一起进行检验，如果结果是阴性，那么对这 k 个人只作 – 一次检验就够了；如果结果是阳性，那么再对这 k 个人的另一份血样逐份检验，此时这 k 份血液的检验次数总共为 $k+1$ 次．

（1）假设有 6 份血液样本，其中只有 2 份样本为阳性，若采用逐份检验的方式，求恰好经过 3 次检验就能把阳性样本全部检验出来的概率；

（2）假设在接受检验的血液样本中，每份样本的检验结果是阳性还是阴性都是相互独立的，且每份样本是阳性结果的概率为 $p(0<p<1)$．现取其中 $k(k\in N^*$ 且 $k..2)$ 份血液样本，记采用逐份检验方式，样本需要检验的总次数为 X_1，采用混合检验方式，样本需要检验的总次数为 X_2．

①运用概率统计的知识，若 $EX_1=EX_2$，试求 p 关于 k 的函数关系式 $p=f(k)$；

②若 $p=1-e^{-\frac{1}{5}}$，且采用混合检验方式可以使得样本需要检验的总次数的期望值比逐份检验的总次数期望值更少，求 k 的最大值.

参考数据：$ln11\approx2.3978$，$1n12\approx2.4849$，$ln13\approx2.5649$.

【分析】（1）记恰好经过 3 次检验就能把阳性样本全部检验出来为 A 事件，利用古典概型、排列组合能求出恰好经过 3 次检验就能把阳性样本全部检验出来的概率.

（2）① X_1 的取值为 k，$P(X_1=k)=1$，$EX_1=k$，X_2 的取值为 1，$k+1$，$P(X_2=1)=(1-p)^k$，$P(X_2=k+1)=1-(1-p)^k$，从而 $E(X_2)=k+1-k(1-p)^k$，由 $EX_1=EX_2$，能求出 p.

② $p=1-e^{-\frac{1}{5}}$，$EX_2=k+1-ke^{-\frac{k}{5}}<k$，从而 $lnk-\frac{k}{5}>0$，设 $f(x)=lnx-\frac{x}{5}$，则 $f'(x)=\frac{1}{x}-\frac{1}{5}=\frac{5-x}{5x}$，$x>0$，利用导数性质能求出 k 的最大值.

【解答】解：（1）记恰好经过 3 次检验就能把阳性样本全部检验出来为 A 事件，

则恰好经过 3 次检验就能把阳性样本全部检验出来的概率为 P（A）$=\frac{C_2^1C_4^1A_2^2}{A_6^3}=\frac{2}{15}$.

（2）① X_1 的取值为 k，$P(X_1=k)=1$，$\therefore EX_1=k$，

X_2 的取值为 1，$k+1$，

$P(X_2=1)=(1-p)^k$，$P(X_2=k+1)=1-(1-p)^k$，

$\therefore E(X_2)=(1-p)^k+(k+1)[1-(1-p)^k]=k+1-k(1-p)^k$，

由 $EX_1=EX_2$，得 $k=k+1-k(1-p)^k$，

$\therefore p=1-(\frac{1}{k})^{\frac{1}{k}}$，$(k\in N^*$，$k..2)$.

② $p=1-e^{-\frac{1}{5}}$，$EX_2=k+1-ke^{-\frac{k}{5}}<k$，$\therefore lnk-\frac{k}{5}>0$，

设 $f(x)=lnx-\frac{x}{5}$，则 $f'(x)=\frac{1}{x}-\frac{1}{5}=\frac{5-x}{5x}$，$x>0$，

当 $x\in(0,5)$ 时，$f'(x)>0$，$f(x)$ 在 $(0,5)$ 上单调递增，

$x\in(5,+\infty)$ 时，$f'(x)<0$，$f(x)$ 在 $(5,+\infty)$ 单调递减，

且 $f(12)=ln2-2$，$.2>0$，$f(13)=ln13-2.6<0$，

$\therefore k$ 的最大值为 12.

【点评】本题考查概率、离散型随机变量的分布列、数学期望的求法及应用，考查古典概型、导数性质、函数的单调性等基础知识，考查运算求解能力，是中档题.

15.（2020•宜昌模拟）目前，新冠病毒引发的肺炎疫情在全球肆虐，为了解新冠肺炎传播途径，采取有效防控措施，某医院组织专家统计了该地区 500 名患者新冠病毒潜伏期的相关信息，数据经过汇总整理得到如图所示的频率分布直方图（用频率作为概率）．潜伏期不高于平均数的患者，称为“短潜伏者”，潜伏期高于平均数的患者，称为“长潜伏者”.

（1）求这 500 名患者潜伏期的平均数（同一组中的数据用该组区间的中点值作代表），并计算出这 500 名患者中“长潜伏者”的人数；

（2）为研究潜伏期与患者年龄的关系，以潜伏期是否高于平均数为标准进行分层抽样，从上述 500 名患者中抽取 300 人，得到如下列联表，请将列联表补充完整，并根据列联表判断是否有97.5% 的把握认为潜伏期长短与患者年龄有关：

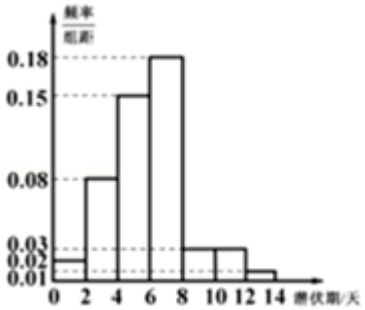
	短潜伏者	长潜伏者	合计
60 岁及以上	90	70	____
60 岁以下	____	____	140
合计	____	____	300

（3）研究发现，有 5 种药物对新冠病毒有一定的抑制作用，其中有 2 种特别有效，现在要通过逐一试验直到把这 2 种特别有效的药物找出来为止，每一次试验花费的费用是 500 元，设所需要的试验费用为 X，求 X 的分布列与数学期望 X .

附表及公式：

$P(K^2..k_0)$	0.15	0.10	0.05	0.025	0.010	0.005	0.001
k_0	2.072	2.706	3.841	5.024	6.635	7.879	10.828

$$K^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$



【分析】（1）由频率分布直方图中的数据可求出平均数和频率，用频率乘以 500 即可得到频数；

（2）根据 2×2 列联表中的已知数据和列联表本身的性质，补充完整表格，再由 K^2 的公式计算出观测值，并与附表中的数据进行对比即可作出判断；

（3）所需要的试验费用 X 所有可能的取值为 1000，1500，2000，然后结合排列组合与概率逐一计算出每个 X 的取值所对应的概率即可得到分布列，进而求得数学期望．

【解答】解：（1）平均数 $x=(0.02\times1+0.08\times3+0.15\times5+0.18\times7+0.03\times9+0.03\times11+0.01\times13)\times2=6$，

这 500 名患者中“长潜伏者”的频率为 $(0.18+0.03+0.03+0.01)\times2=0.5$，

所以“长潜伏者”的人数为 $500\times0.5=250$ 人．

（2）由题意补充后的列联表如下，

	短潜伏者	长潜伏者	合计
60 岁及以上	90	70	160
60 岁以下	60	80	140
合计	150	150	300

则 k^2 的观测值为 $k=\dfrac{300\times(90\times80-60\times70)^2}{150\times150\times160\times140}=\dfrac{75}{14}\approx5.357>5.024$，

经查表，得 $P(k^2..5.024)\approx0.025$，

所以有 97.5% 的把握认为潜伏期长短与年龄有关．

（3）由题意知，所需要的试验费用 X 所有可能的取值为 1000，1500，2000，

$P(X=1000)=\dfrac{A_2^2}{A_5^2}=\dfrac{1}{10}$，

$P(X=1500)=\dfrac{C_2^1C_3^lA_2^2+A_3^3}{A_5^3}=\dfrac{3}{10}$，

$P(X=2000)=\dfrac{C_2^1A_3^1A_3^2C_2^1}{A_5^4}=\dfrac{3}{5}$（或 $P(X=2000)=\dfrac{C_2^1C_3^1A_3^2}{A_5^3}=\dfrac{36}{60}=\dfrac{3}{5}$），

所以 X 的分布列为

X	1000	1500	2000
P	$\dfrac{1}{10}$	$\dfrac{3}{10}$	$\dfrac{3}{5}$

数学期望 $E(X)=1000\times\dfrac{1}{10}+1500\times\dfrac{3}{10}+2000\times\dfrac{3}{5}=1750$（元）．

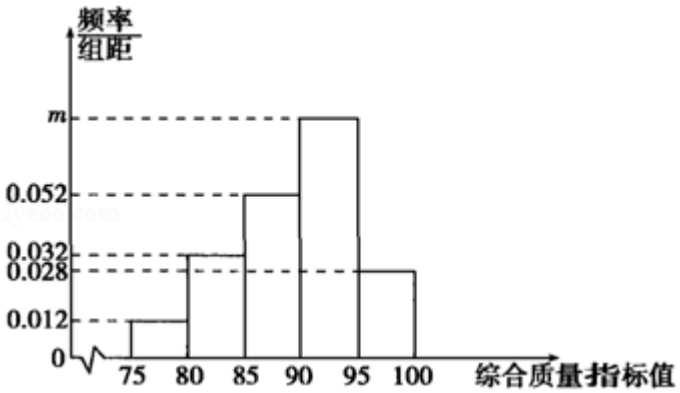
【点评】本题考查频率分布直方图、独立性检验和离散型随机变量的分布列与期望，考查学生对数据的分析能力和灵活运用知识的能力，属于基础题.

16.（2020•山东模拟）第五届中国“互联网＋”大学生创新创业大赛在浙江大学开幕.10月 15 日，第五届中国“互联网＋”大学生创新创业大赛全国总决赛闭幕．在此次全国总决赛中，高教主赛道共评选出 60 个金奖项目（含 5 个港澳台金奖项目），“青年红色筑梦之旅”赛道、职教赛道和国际赛道各评选出 15 个金奖项目．某公司迅速引进其中的两个项目，在甲，乙两个生产车间进行试验，其中，甲车间只有一条生产线，乙车间有两条生产线，公司在试运营过程中随机选取了甲车间生产的部分产品，测得相应的综合质量指标值为t．根据国家规定，综合质量指标值$t..90$为优等品；$80,, t<90$为合格品；$t<80$为不合格品．甲车间生产的产品综合质量指标值的频率分布直方图如图.

（1）已知每件产品的生产成本与综合质量指标值t的关系式为$P(t)=0.5t$，若优等品的出售价格为 100 元，合格品的出售价格为 80 元，不合格品需要全部销毁，每件产品的销毁费用为 20 元，将每件产品的利润用综合质量指标值t的函数$Q(t)$来表示；

（2）根据预测，甲车间生产线成功生产产品的概率为p，每年可获得利润$\frac{700}{3}$万元；乙车间每条生产线成功生

产产品的概率为$\frac{1}{2}(e^{p}-1)$，每年可获得利润 300 万元，记每年两个车间的利润之和为X，则当p为何值时，X的期望最大？



【分析】（1）根据t的取值范围不同，得到产品的合格性不同，分三段分别写出对应的$Q(t)$即可得解；

（2）随机变量X的可能取值为$\frac{700}{3}$，300，$\frac{1600}{3}$，600，$\frac{2500}{3}$，然后根据独立事件的概率逐一表示出每个X的取值所对应的概率，由概率和为 1 可求得$e^{p}-1$与p之间的等量关系，将其代入数学期望中，由$0<p,,1$，从而得解.

【解答】解：（1）当$75,, t<80$时，$Q(t)=-0.5t-20$；

当 $80\le t<90$ 时，$Q(t)=80-0.5t$；

当 $90\le t\le 100$ 时，$Q(t)=100-0.5t$．

$$\therefore Q(t)=\begin{cases}-0.5t-20, & 75\le t<80\\ 80-0.5t, & 80\le t<90\\ 100-0.5t, & 90\le t\le 100\end{cases}.$$

（2）随机变量 X 的可能取值为 $\frac{700}{3}$，300，$\frac{1600}{3}$，600，$\frac{2500}{3}$，

$P(X=\frac{700}{3})=p\times[1-\frac{1}{2}(e^p-1)]^2$，

$P(X=300)=(1-p)\times C_2^1\times\frac{1}{2}(e^p-1)\times[1-\frac{1}{2}(e^p-1)]$，

$P(X=\frac{1600}{3})=p\times C_2^1\times\frac{1}{2}(e^p-1)\times[1-\frac{1}{2}(e^p-1)]$，

$P(X=600)=(1-p)\times[\frac{1}{2}(e^p-1)]^2$，

$P(X=\frac{2500}{3})=p\times[\frac{1}{2}(e^p-1)]^2$．

由 $P(X=\frac{700}{3})+P(X=300)+P(X=\frac{1600}{3})+P(X=600)+P(X=\frac{2500}{3})=1$，化简整理后得 $e^p-1=4p$．

数学期望 $E(X)=\frac{700}{3}\times p\times[1-\frac{1}{2}(e^p-1)]^2+300\times(1-p)\times C_2^1\times\frac{1}{2}(e^p-1)\times[1-\frac{1}{2}(e^p-1)]+\frac{1600}{3}\times p\times C_2^1\times\frac{1}{2}(e^p-1)\times[1-\frac{1}{2}(e^p-1)]+600\times(1-p)\times[\frac{1}{2}(e^p-1)]^2+\frac{2500}{3}\times p\times[\frac{1}{2}(e^p-1)]^2=300(e^p-1)+\frac{700}{3}p=300\times 4p+\frac{700}{3}p=\frac{4300}{3}p$．

$\because 0<p\le 1$，$\therefore$ 当 $p=1$ 时，$E(X)$ 最大，为 $\frac{4300}{3}$ 万元．

【点评】本题考查分段函数的应用、离散型随机变量的数学期望及其应用，试题本身难度不大，但计算量特别大，考查学生对数据的分析能力和运算能力，属于中档题．

17．（2020•宁德模拟）某市旅游局为尽快恢复受疫情影响的旅游业，准备在本市的景区推出旅游一卡通（年卡）．为了更科学的制定一卡通的有关条例，市旅游局随机调查了 2019 年到本市景区旅游的 1000 个游客的年旅游消费支出（单位：百元），并制成如图频率分布直方图：

由频率分布直方图，可近似地认为到本市景区旅游的游客，其旅游消费支出服从正态分布 $N(\mu,3.2^2)$，其中 μ 近似为样本平均数 $\bar{x}$（同一组数据用该组区间的中点值作代表）．

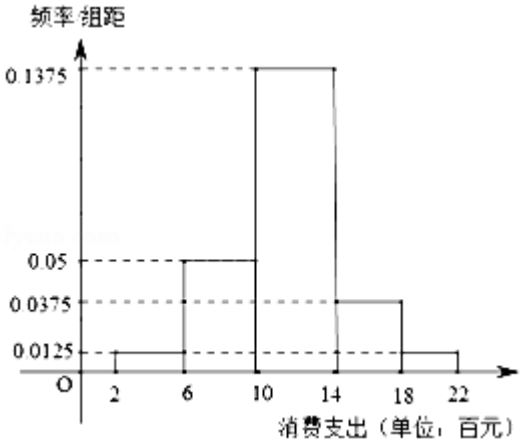
（1）若 2019 年到本市景区旅游游客为 500 万人，试估计 2019 年有多少游客在本市的年旅游消费支出不低于 1820 元；

（2）现依次抽取 n 个游客，假设每个游客的旅游消费支出相互独立，记事件 A 表示“连续 3 人的旅游消费支出超出 μ”．若 P_n 表示 $\bar{A}$ 的概率，$P_n=aP_{n-1}+\frac{1}{4}P_{n-2}+bP_{n-3}(n\ge 3,\ a,\ b$ 为常数），且 $P_0=P_1=P_2=1$．

(i) 求 P_3，P_4 及 a，b；

(*ii*) 判断并证明数列 $\{P_n\}$ 从第三项起的单调性，试用概率统计知识解释其实际意义．（参考数据：

$P(\mu-\sigma<X<\mu+\sigma)\approx 0.6826$，$P(\mu-2\sigma<X<\mu+2\sigma)\approx 0.9544$，$P(\mu-3\sigma<X<\mu+3\sigma)\approx 0.9973$)



【分析】（1）由直方图可得 $\bar{x}$，即可得到 μ，结合已知的 $\sigma=3.2$，可知旅游费用支出不低于 1820 元的概率为 $P(x..\mu+2\sigma)$，求得概率后乘以 500 得答案．

（2）(*i*) 先由题意求得 P_3 与 P_4 的值，再列关于 a，b 的方程组求解 a，b 的值；

(*ii*) 由 $P_n=\frac{1}{2}P_{n-1}+\frac{1}{4}P_{n-2}+\frac{1}{8}P_{n-3}(n..3)$，利用作差法可得从第三项起数列 $\{P_n\}$ 单调递减．其实际意义为随着抽查人数 n 的增加，事件“不连续 3 人的旅游费用支出超出 μ”的可能性会越来越小．（即最终会出现连续 3 人的旅游费用支出超出 μ 这一事件）．

【解答】解：（1）直方图可得 $\bar{x}=(0.0125\times4+0.05\times8+0.1375\times12+0.375\times16+0.125\times20)\times4=11.8$，

$\because\ \mu=\bar{x}=11.8$，$\sigma=3.2$，$\mu+2\sigma=1820$ 元，

$\therefore$ 旅游费用支出不低于 1820 元的概率为 $P(x..\mu+2\sigma)=\frac{1-P(\mu-2\sigma<x<\mu+2\sigma)}{2}=\frac{1-0.9544}{2}=0.0228$，

$\therefore 500\times0.022=11.4$，

估计 2019 年有 11.4 万的游客在本市的年旅游费用支出不低于 1820 元．

（2）(*i*) $P_3=1-\frac{1}{8}=\frac{7}{8}$，$P_4=1-\frac{2+1}{16}=\frac{13}{16}$，

由 $\begin{cases}P_3=aP_2+\frac{1}{4}P_1+bP_0\\P_4=aP_3+\frac{1}{4}P_2+bP_1\end{cases}$，即 $\begin{cases}\frac{7}{8}=a+\frac{1}{4}+b\\\frac{13}{16}=\frac{7}{8}a+\frac{1}{4}+b\end{cases}$，

解得 $\begin{cases}a=\frac{1}{2}\\b=\frac{1}{8}\end{cases}$；

(*ii*) 数列 $\{P_n\}$ 从第三项起单调递减．

$P_n=\frac{1}{2}P_{n-1}+\frac{1}{4}P_{n-2}+\frac{1}{8}P_{n-3}(n..3)$，

故 $P_{n+1}-P_n=(\frac{1}{2}P_n+\frac{1}{4}P_{n-1}+\frac{1}{8}P_{n-2})-(\frac{1}{2}P_{n-1}+\frac{1}{4}P_{n-2}+\frac{1}{8}P_{n-3})=\frac{1}{2}P_n-\frac{1}{4}P_{n-1}-\frac{1}{8}P_{n-2}-\frac{1}{8}P_{n-3}$

$=\frac{1}{2}(\frac{1}{2}P_{n-1}+\frac{1}{4}P_{n-2}+\frac{1}{8}P_{n-3})-\frac{1}{4}P_{n-1}-\frac{1}{8}P_{n-2}-\frac{1}{8}P_{n-3}=-\frac{1}{16}P_{n-3}$，

又 $P_n>0$，$\therefore -\frac{1}{16}P_{n-3}<0$，

即从第三项起数列$\{P_n\}$单调递减.

由此，可知随着抽查人数n的增加，事件“不连续 3 人的旅游费用支出超出μ”的可能性会越来越小.（即最终会出现连续 3 人的旅游费用支出超出μ这一事件）.

【点评】本小题主要考查频率分布直方图、平均数、正态分布、随机事件的概率、数列及其性质等基础知识，考查运算求解能力、数据处理能力、应用意识，考查分类与整合思想、统计思想、化归与转化思想.

18.（2020 春•武昌区校级期中）湖北省从 2021 年开始将全面推行新高考制度，新高考“3+1+2”中的“2”要求考生从政治、化学、生物、地理四门中选两科，按照等级赋分计入高考成绩，等级赋分规则如下：高考政治、化学、生物、地理四门等级考试科目的考生原始成绩从高到低划分为A，B，C，D，E五个等级，确定各等级人数所占比例分别为15%，35%，35%，13%，2%，等级考试科目成绩计入考生总成绩时，将A至E等级内的考生原始成绩，依照等比例转换法分别转换到[86，100]、[71，85]、[56，70]、[41，55]、[30，40]五个分数区间，得到考生的等级分，等级转换分满分为 100 分．具体转换分数区间如表：

等级	A	B	C	D	E
比例	15%	35%	35%	13%	2%
赋分区间	[86，100]	[71，85]	[56，70]	[41，55]	[30，40]

而等比例转换法是通过公式计算：$\frac{Y_2-Y}{Y-Y_1}=\frac{T_2-T}{T-T_1}$

其中Y_1、Y_2分别表示原始分区间的最低分和最高分，T_1、T_2分别表示等级分区间的最低分和最高分，Y表示原始分，T表示转换分，当原始分为Y_1、Y_2时，等级分分别为T_1、T_2，假设小明同学的生物考试成绩信息如表：

考试科目	考试成绩	成绩等级	原始分区间	等级分区间
生物	75 分	B等级	[69，84]	[71，85]

设小明转换后的等级成绩为T，根据公式得：$\frac{84-75}{75-69}=\frac{85-T}{T-71}$，所以$T=76.6\approx 77$（四舍五入取整），小明最终生物等级成绩为 77 分．已知某学校学生有 60 人选了政治，以期中考试成绩为原始成绩转换该学校选

政治的学生的政治等级成绩，其中政治成绩获得 A 等级的学生原始成绩统计如表：

成绩	90	86	81	80	79	78	75
人数	1	2	1	1	2	1	1

（1）从政治成绩获得 A 等级的学生中任取 3 名，求至少有 2 名同学的等级成绩不小于 93 分的概率；

（2）从政治成绩获得 A 等级的学生中任取 4 名，设 4 名学生中等级成绩不小于 93 分人数为 ξ，求 ξ 的分布列和期望.

【分析】（1）先根据等级分计算公式将等级分 93 分换算成原始分，从而确定等级分不小于 93 分的人数，再根据古典概型概率的计算方法求解；

（2）由题意，这是一个超几何分布问题，可根据（1）中的结果，按照超几何分布的规律，来计算 ξ 的分布列及期望.

【解答】解：（1）设等级分 93 分对应的原始分为 X，

由题意得：$\dfrac{100-93}{93-86}=\dfrac{90-X}{X-75}$，解得 $X=82.5$，

所以，A 等级的学生中，等级分不小于 93 分的有 3 人.

设事件 $M=$ “至少有两名学生等级分不小于 93 分”，

$\therefore n(M)=C_3^2C_6^1+C_3^3=19$，$n(\Omega)=C_9^3=84$，

故 $P(M)=\dfrac{n(M)}{n(\Omega)}=\dfrac{19}{84}$.

（2）由题意 ξ 的可能取值为 0，1，2，3.

$\therefore P(\xi=0)=\dfrac{C_3^0C_6^4}{C_9^4}=\dfrac{5}{42}$，

$P(\xi=1)=\dfrac{C_3^1C_6^3}{C_9^4}=\dfrac{10}{21}$，

$P(\xi=2)=\dfrac{C_3^2C_6^2}{C_9^4}=\dfrac{5}{14}$，

$P(\xi=3)=\dfrac{C_3^3C_6^1}{C_9^4}=\dfrac{1}{21}$.

故 ξ 的分布列为：

P	0	1	2	3

ξ	$\frac{5}{42}$	$\frac{10}{21}$	$\frac{5}{14}$	$\frac{1}{21}$

故期望 $EX=0\times\frac{5}{42}+1\times\frac{10}{21}+2\times\frac{5}{14}+3\times\frac{1}{21}=\frac{4}{3}$.

故 ξ 的期望为 $\frac{4}{3}$.

【点评】本题考查古典概型概率的计算、离散型随机变量的分布列和期望的计算方法．同时考查学生数学建模、数据分析、数学运算及逻辑推理等数学核心素养．属于中档题．

19．（2020•重庆模拟）手工艺是一种生活态度和对传统的坚持，在我国有很多手工艺品制作村落，村民的手工技艺世代相传，有些村落制造出的手工艺品不仅全国闻名，还大量远销海外．近年来某手工艺品村制作的手工艺品在国外备受欢迎，该村村民成立了手工艺品外销合作社，为严把质量关，合作社对村民制作的每件手工艺品都请 3 位行家进行质量把关，质量把关程序如下：(i) 若一件手工艺品 3 位行家都认为质量过关，则该手工艺品质量为 A 级；(ii) 若仅有 1 位行家认为质量不过关，再由另外 2 位行家进行第二次质量把关，若第二次质量把关这 2 位行家都认为质量过关，则该手工艺品质量为 B 级，若第二次质量把关这 2 位行家中有 1 位或 2 位认为质量不过关，则该手工艺品质量为 C 级；(iii) 若有 2 位或 3 位行家认为质量不过关，则该手工艺品质量为 D 级．已知每一次质量把关中一件手工艺品被 1 位行家认为质量不过关的概率为 $\frac{1}{3}$，且各手工艺品质量是否过关相互独立．

（1）求一件手工艺品质量为 B 级的概率；

（2）若一件手工艺品质量为 A，B，C 级均可外销，且利润分别为 900 元，600 元，300 元，质量为 D 级不能外销，利润记为 100 元．

①求 10 件手工艺品中不能外销的手工艺品最有可能是多少件；

②记 1 件手工艺品的利润为 X 元，求 X 的分布列与期望．

【分析】（1）利用独立重复实验的关键求解一件手工艺品质量为 B 级的概率．

（2）①求出一件手工艺品质量为 D 级的概率，设 10 件手工艺品中不能外销的手工艺品可能是 ξ 件，得到二项分布，通过概率的比值，判断 10 件手工艺品中不能外销的手工艺品最有可能是 2 件．

②求出一件手工艺品质量为 A 级的概率，一件手工艺品质量为 B 级的概率，一件手工艺品质量为 C 级的概率，一件手工艺品质量为 D 级的概率，得到 X 的分布列，然后求解期望即可．

【解答】解：（1）一件手工艺品质量为 B 级的概率为 $C_3^1\times\frac{1}{3}\times(1-\frac{1}{3})^2\times(1-\frac{1}{3})^2=\frac{16}{81}$.

（2）①由题意可得一件手工艺品质量为 D 级的概率为 $C_3^2\times(\frac{1}{3})^2\times(1-\frac{1}{3})+C_3^3\times(\frac{1}{3})^3=\frac{7}{27}$，

设 10 件手工艺品中不能外销的手工艺品可能是 ξ 件，则 $\xi\sim B(10,\frac{7}{27})$，

则 $P(\xi=k)=C_{10}^{k}(\frac{7}{27})^{k}(\frac{20}{27})^{10-k}$，$\frac{P(\xi=k+1)}{P(\xi=k)}=\frac{C_{10}^{k+1}(\frac{7}{27})^{k+1}(\frac{20}{27})^{9-k}}{C_{10}^{k}(\frac{7}{27})^{k}(\frac{20}{27})^{10-k}}=\frac{70-7k}{20k+20}$.

由 $\frac{70-7k}{20k+20}>1$ 得 $k<\frac{50}{27}$，所以当 $k=1$ 时，$\frac{P(\xi=2)}{P(\xi=1)}>1$，即 $P(\xi=2)>P(\xi=1)$，

由 $\frac{70-7k}{20k+20}<1$ 得 $k>\frac{50}{27}$，所以当 $k..2$ 时，$P(\xi=k+1)<P(\xi=k)$，

所以当 $k=2$ 时，$P(\xi=k)$ 最大，即 10 件手工艺品中不能外销的手工艺品最有可能是 2 件.

②由上可得一件手工艺品质量为 A 级的概率为 $(1-\frac{1}{3})^{3}=\frac{8}{27}$，一件手工艺品质量为 B 级的概率为 $\frac{16}{81}$，

一件手工艺品质量为 C 级的概率为 $C_3^1\times\frac{1}{3}\times(1-\frac{1}{3})^{2}\times[C_2^1\times\frac{1}{3}\times(1-\frac{1}{3})+(\frac{1}{3})^{2}]=\frac{20}{81}$，

一件手工艺品质量为 D 级的概率为 $\frac{7}{27}$，

所以 X 的分布列为

X	900	600	300	100
P	$\frac{8}{27}$	$\frac{16}{81}$	$\frac{20}{81}$	$\frac{7}{27}$

则期望为 $E(X)=900\times\frac{8}{27}+600\times\frac{16}{81}+300\times\frac{20}{81}+100\times\frac{7}{27}=\frac{13100}{27}$.

【点评】本题考查离散型随机变量的分布列以及期望的求法，独立重复实验的应用，考查转化思想以及计算能力，是中档题.

20.（2020•永州三模）为丰富学生课外生活，某市组织了高中生钢笔书法比赛，比赛分两个阶段进行：第一阶段由评委给出所有参赛作品评分，并确定优胜者；第二阶段为附加赛，参赛人员由组委会按规则另行确定．数据统计员对第一阶段的分数进行了统计分析，这些分数 X 都在 $[70，100)$ 内，在以组距为 5 画分数的频率分布直方图（设“$\frac{\text{频率}}{\text{组距}}=Y$”）时，发现 Y 满足 $Y=\begin{cases}\frac{8n-109}{300}, n\,..\,16\\ \frac{1}{15}-k\frac{1}{20-n}, n>16\end{cases}$，$n\in N^{*}$，$5n\,..\,X<5(n+1)$.

（1）试确定 n 的所有取值，并求 k；

（2）组委会确定：在第一阶段比赛中低于 85 分的参赛者无缘获奖也不能参加附加赛；分数在 $[95，100)$ 的参赛者评为一等奖；分数在 $[90，95)$ 的同学评为二等奖，但通过附加赛有 $\frac{1}{11}$ 的概率提升为一等奖；分数在 $[85，90)$ 的同学评为三等奖，但通过附加赛有 $\frac{1}{7}$ 的概率提升为二等奖（所有参加附加赛的获奖人员均不降低获奖等级）．已知学生 A 和 B 均参加了本次比赛，且学生 A 在第一阶段评为二等奖．

(i) 求学生 B 最终获奖等级不低于学生 A 的最终获奖等级的概率；

(*ii*) 已知学生 A 和 B 都获奖，记 A，B 两位同学最终获得一等奖的人数为 ξ，求 ξ 的分布列和数学期望.

【分析】（1）X 在 $[70，100)$ 内，按组距为 5 可分成 6 个小区间，分别是 $[70，75)$，$[75，80)$，$[80，85)$，$[85，90)$，$[90，95)$，$[95，100)$，由 $70,, X<100$，由 $5n,, X<5(n+1)$，$n\in N^*$，能求出 n 的对值和 k.

（2）(*i*) 由于参赛学生很多，可以把频率视为概率，学生 B 的分数属于区间 $[70，75)$，$[75，80)$，$[80，85)$，$[85，90)$，$[90，95)$，$[95，100)$ 的概率分别是 $\frac{3}{60},\frac{11}{60},\frac{19}{60},\frac{14}{60},\frac{11}{60},\frac{2}{60}$，用符号 A_{ij}（或 B_{ij}）表示学生 A（或 B）在第一轮获奖等级为 i，通过附加赛最终获奖等级为 j，其中 $j,, i(i，\ j=1，2，3)$，记 $W=$ “学生 B 最终获奖等级不低于学生 A 的最终获奖等级”，由此能求出学生 B 最终获奖等级不低于学生 A 的最终获奖等级的概率.

(*ii*) 学生 A 最终获得一等奖的概率是 $P(A_{21})=\frac{1}{11}$，学生 B 最终获得一等奖的概率是 $P(B_{1'}+B_{21'})=\frac{2}{27}+\frac{11}{27}\square\frac{1}{11}=\frac{1}{9}$，$\xi$ 的可能取值为 0，1，2，分虽求出相应的概率，由此能求出 ξ 的分布列和 $E\xi$.

【解答】解：（1）根据题意，X 在 $[70，100)$ 内，按组距为 5 可分成 6 个小区间，

分别是 $[70，75)$，$[75，80)$，$[80，85)$，$[85，90)$，$[90，95)$，$[95，100)$，

$\because 70,, X<100$，由 $5n,, X<5(n+1)$，$n\in N^*$，

$\therefore n=14$，15，16，17，18，19，

每个小区间对应的频率值分别是 $P=5Y=\begin{cases}\frac{8n-109}{60},n=14,15,16\\ \frac{1}{3}-5k\square\frac{1}{20-n},n=17,18,19\end{cases}$.

$\frac{3}{60}+\frac{11}{60}+\frac{19}{60}+1-5k(\frac{1}{3}+\frac{1}{2}+1)=1$，解得 $k=\frac{3}{50}$，

$\therefore n$ 的对值是 14，15，16，17，18，19，$k=\frac{3}{50}$.

（2）(*i*) 由于参赛学生很多，可以把频率视为概率，

由（1）知，学生 B 的分数属于区间 $[70，75)$，$[75，80)$，$[80，85)$，$[85，90)$，$[90，95)$，$[95，100)$ 的概率分别是：

$\frac{3}{60},\frac{11}{60},\frac{19}{60},\frac{14}{60},\frac{11}{60},\frac{2}{60}$，

我们用符号 A_{ij}（或 B_{ij}）表示学生 A（或 B）在第一轮获奖等级为 i，通过附加赛最终获奖等级为 j，

其中 $j,, i(i，\ j=1，2，3)$，

记 $W=$ “学生 B 最终获奖等级不低于学生 A 的最终获奖等级”，

则 $P(W)=P(B_1+B_{21}+B_{22}A_{22}+B_{32}A_{22})$

$= P(B_1) + P(B_{21}) + P(B_{22})P(A_{22}) + P(B_{32})P(A_{22})$

$= \frac{2}{60} + \frac{11}{60} \times \frac{1}{11} + \frac{11}{60} \times \frac{10}{11} \times \frac{10}{11} + \frac{14}{60} \times \frac{1}{7} \times \frac{10}{11} = \frac{51}{220}$.

(ii) 学生 A 最终获得一等奖的概率是 $P(A_{21}) = \frac{1}{11}$，

学生 B 最终获得一等奖的概率是 $P(B_{1'} + B_{21'}) = \frac{2}{27} + \frac{11}{27} \times \frac{1}{11} = \frac{1}{9}$，

$P(\xi = 0) = (1 - \frac{1}{11})(1 - \frac{1}{9}) = \frac{80}{99}$，

$P(\xi = 1) = \frac{1}{11} \times (1 - \frac{1}{9}) + (1 - \frac{1}{11}) \times \frac{1}{9} = \frac{18}{99}$，

$P(\xi = 2) = \frac{1}{11} \times \frac{1}{9} = \frac{1}{99}$，

$\therefore \xi$ 的分布列为：

ξ	0	1	2
P	$\frac{80}{99}$	$\frac{18}{99}$	$\frac{1}{99}$

$E\xi = 0 \times \frac{80}{99} + 1 \times \frac{18}{99} + 2 \times \frac{1}{99} = \frac{20}{99}$.

【点评】本题考查概率、离散型随机变量的分布列、数学期望的求法，考查分布列、数学期望等基础知识，考查运算求解能力，是中档题.

21.（2020•重庆模拟）新型冠状病毒属于 β 属的冠状病毒，人群普遍易感，病毒感染者一般有发热咳嗽等临床表现，现阶段也出现无症状感染者. 基于目前的流行病学调查和研究结果，病毒潜伏期一般为1–14天，大多数为3–7天. 为及时有效遏制病毒扩散和蔓延，减少新型冠状病毒感染对公众健康造成的危害，需要对与确诊新冠肺炎病人接触过的人员进行检查. 某地区对与确诊患者有接触史的 1000 名人员进行检查，检查结果统计如下：

	发热且咳嗽	发热不咳嗽	咳嗽不发热	不发热也不咳嗽
确诊患病	200	150	80	30
确诊未患病	150	150	120	120

（1）能否在犯错率不超过 0.001 的情况下，认为新冠肺炎密切接触者有发热症状与最终确诊患病有关.

临界值表：

$P(K^2 > k)$	0.10	0.05	0.010	0.005	0.001
k	2.706	3.841	6.645	7.879	10.828

（2）在全国人民的共同努力下，尤其是全体医护人员的辛勤付出下，我国的疫情得到较好控制，现阶段防控重难点主要在境外输入病例和无症状感染者（即无相关临床表现但核酸检测或血清特异性免疫球蛋白 M 抗体检测阳性者）．根据防控要求，无症状感染者虽然还没有最终确诊患 2019 新冠肺炎，但与其密切接触者仍然应当采取居家隔离医学观察 14 天，已知某人曾与无症状感染者密切接触，而且在家已经居家隔离 10 天未有临床症状，若该人员居家隔离第 k 天出现临床症状的概率为 $(\frac{1}{2})^{k-10}$，$(k=11，12，13，14)$，两天之间是否出现临床症状互不影响，而且一旦出现临床症状立刻送往医院核酸检查并采取必要治疗，若 14 天内未出现临床症状则可以解除居家隔离，求该人员在家隔离的天数（含有临床症状表现的当天）ξ 的分布列以及数学期望值．（保留小数点后两位）

【分析】（1）根据表格中数据，列出确诊、未确诊中发热和不发热的人数，填写 2×2 列联表，根据列联表中数据，计算 K^2，对照临界值表得出结论即可．

（2）根据题意 ξ 的可能取值为 11，12，13，14，计算出对应的概率，列出 ξ 的分布列，求出 ξ 的数学期望值即可．

【解答】解：（1）由表可得，患者有发热症状与确诊的 2╳2 列联表如下：

	发热	不发热	合计
确诊	350	110	460
未确诊	300	240	540
合计	650	350	1000

由公式可得：$K^2=\frac{1000\times(350\times240-300\times110)^2}{460\times540\times650\times350}=\frac{10404000}{226044}\approx46.02>10.828$；

故在犯错率不超过 0.001 的情况下，有把握认为新冠肺炎密切接触者有发热症状与最终确诊患病有关．

（2）由题可知，随机变量 ξ 可以取值：11，12，13，14；

其分布列为：

ξ	11	12	13	14
P	$\frac{1}{2}$	$\frac{1}{2}\times\frac{1}{4}=\frac{1}{8}$	$\frac{1}{2}\times\frac{3}{4}\times\frac{1}{8}=\frac{3}{64}$	$\frac{1}{2}\times\frac{3}{4}\times\frac{7}{8}=\frac{21}{64}$

其数学期望为：$E(\xi)=\frac{781}{64}\approx12.20$．

【点评】本题考查了列联表与独立性检验的应用问题以及离散型随机变量的概率分布列问题，属于中档题．

22.（2020•新乡二模）某公司准备上市一款新型轿车零配件，上市之前拟在其一个下属 4S 店进行连续 30 天的试销，定价为 1000 元/件.

（1）设日销售 40 个零件的概率为 p，记 5 天中恰有 2 天销售 40 个零件的概率为 z，写出 z 关于 p 的函数关系式，并求 z 的极大值点 p_0.

（2）试销结束后统计得到该 4S 店这 30 内的日销售量（单位：件）的数据如表：

日销售量	40	60	80	100
频数	9	12		

其中，有两个数据未给出．试销结束后，这款零件正式上市，每件的定价仍为 1000 元，但生产公司对该款零件不零售，只提供零件的整箱批发，大箱每箱有 55 件，批发价为 550 元/件；小箱每箱有 40 件，批发价为 600 元/件，以这 30 天统计的各日销售量的频率作为试销后各日销售量发生的概率.

该 4S 店决定每天批发两箱，若同时批发大箱和小箱，则先销售小箱内的零件，同时根据公司规定，当天没销售出的零件按批发价的 9 折转给该公司的另一下属 4S 店，假设日销售量为 80 件的概率为 $\frac{p_0}{2}$，其中 P_0 为（1）中 z 的极大值点.

(i) 设该 4S 店批发两大箱，当天这款零件的利润为随机变量 X；批发两小箱，当天这款零件的利润为随机变量 Y，求 EX 和 EY；

(ii) 以日利润的数学期望作为决策依据，该 4S 店每天应该按什么方案批发零件？

【分析】（1）设日销售 40 个零件的概率为 p，记 5 天中恰有 2 天销售 40 个零件的概率为 z，写出 z 关于 p 的函数关系式 $z=C_5^2p^2(1-p)^3=10p^2(1-p)^3$，$0<p<1$，再由 $z'=10p(1-p)^2(2-5p)$，利用导数性质能求出 z 的极大值点 p_0.

（2）日销售量为 80 件的概率为 $\frac{p_0}{2}=\frac{1}{5}$，日销售量为 100 的概率为 $1-\frac{3}{10}-\frac{2}{5}-\frac{1}{5}=\frac{1}{10}$，

(i) 批发两大箱，则批发成本为 60500 元，分别求出当日销售量为 40 件、60 件、80 件、100 件时的利润，由此能求出 EX；若批发两小箱，则批发成本为 48000 元，分别求出当日销售量为 40 件、60 件、80 件或 100 件时的利润，由此能求出 EY.

(ii) 当 4S 店批发一大箱和一小箱时，成本为 54250 万元，当天这款零件的利润为随机变量 ξ，分别求出当日销售量为 40 件、60 件、80 件、100 件时的利润，求出 $E\xi$，由 $EY<E\xi<EX$，得到以日利润的数学期望作为决策依据，该 4S 店每天应该按批发两大箱.

【解答】解：（1）由题意可得 $z=C_5^2p^2(1-p)^3=10p^2(1-p)^3$，$0<p<1$，

$z'=10[2p(1-p)^3-3p^2(1-p)^2]=10p(1-p)^2(2-5p)$，

当$0<p<\frac{2}{5}$时，$z'>0$，当$\frac{2}{5}<p<1$时，$z'<0$，

$\therefore p_0=\frac{2}{5}$.

（2）由题意得日销售量为 80 件的概率为$\frac{p_0}{2}=\frac{1}{5}$，

日销售量为 100 的概率为$1-\frac{3}{10}-\frac{2}{5}-\frac{1}{5}=\frac{1}{10}$，

(i)批发两大箱，则批发成本为 60500 元，

当日销售量为 40 件时，利润为：$40\times1000-60500+70\times550\times90\%=1.415$（万元），

当日销售量为 60 件时，利润为：$60\times1000-60500+50\times550\times90\%=2.425$（万元），

当日销售量为 80 件时，利润为：$80\times1000-60500+30\times550\times90\%=3.435$（万元），

当日销售量为 100 件时，利润为：$100\times1000-60500+10\times550\times90\%=4.445$（万元），

$\therefore EX=1.415\times\frac{3}{10}+2.425\times\frac{2}{5}+3.435\times\frac{1}{5}+4.445\times\frac{1}{10}=2.526$（万元）.

若批发两小箱，则批发成本为 48000 元，

当日销售量为 40 件时，利润为：$40\times1000-48000+40\times600\times90\%=1.36$（万元），

当日销售量为 60 件时，利润为：$60\times1000-48000+20\times600\times90\%=2.3075$（万元），

当日销售量为 80 件或 100 件时，利润为：$80\times1000-48000=3.2$（万元），

$\therefore EY=1.36\times\frac{3}{10}+2.28\times\frac{2}{5}+3.2\times\frac{3}{10}=2.28$（万元）.

(ii)当 4S 店批发一大箱和一小箱时，成本为 54250 万元，当天这款零件的利润为随机变量ξ，

当日销售量为 40 件时，利润为：$40\times1000-54250+55\times550\times90\%=1.2975$（万元），

当日销售量为 60 件时，利润为：$60\times1000-54250+35\times550\times90\%=2.3075$（万元），

当日销售量为 80 件时，利润为：$80\times1000-54250+15\times550\times90\%=3.3175$（万元），

当日销售量为 100 件时，利润为：$100\times1000-54250=4.075$（万元），

$\therefore E\xi=1.2975\times\frac{3}{10}+2.3075\times\frac{2}{5}+3.3175\times\frac{1}{5}+4.075\times\frac{1}{10}=2.38325$（万元）.

$\therefore EY<E\xi<EX$，$\therefore$以日利润的数学期望作为决策依据，该 4S 店每天应该按批发两大箱.

【点评】本题考查概率、离散型随机变量的分布列、数学期望的求法及应用，考查导数性质等基础知识，考查运算求解能力，是中档题.

23.（2020•山东模拟）“爱国，是人世间最深层、最持久的情感，是一个人立德之源、立功之本.”在中华民族几千年绵延发展的历史长河中，爱国主义始终是激昂的主旋律．爱国汽车公司拟对“东方红”款高端

汽车发动机进行科技改造，根据市场调研与模拟，得到科技改造投入 x（亿元）与科技改造直接收益 y（亿元）的数据统计如下：

x	2	3	4	6	8	10	13	21	22	23	24	25
y	13	22	31	42	50	56	58	68.5	68	67.5	66	66

当 $0<x,,17$ 时，建立了 y 与 x 的两个回归模型，模型①：$\hat{y}=4.1x+11.8$；模型②：$\hat{y}=21.3x-14.4$．当 $x>17$ 时，y 与 x 满足的线性回归方程为 $\hat{y}=-0.7x+a$．

（1）根据下列表格中的数据，比较当 $0<x,,17$ 时模型①和模型②的相关指数 R^2，从而选择拟合精度更高、更

可靠的模型，并据此预测当“东方红”款高端汽车发动机科技改造的投入为 17 亿元时的直接收益；

回归模型	模型①	模型②
回归方程	$\hat{y}=4.1x+11.8$	$\hat{y}=21.3\sqrt{x}-14.4$
$\sum_{i=1}^{n}(y_i-\overline{y_i})^2$	182.4	79.2

（2）为鼓励科技创新，当科技改造投人不少于 20 亿元时，国家给予公司补贴收益 10 亿元，以回归方程为预测

依据，比较科技改造投人 17 亿元与 20 亿元时公司实际收益的大小；

（3）科技改造后，“东方红”款高端汽车发动机的热效率 X 大幅提高，X 服从正态分布 $N(0.52,\ 0.01^2)$，公司对科技改造团队的奖励方案如下：若发动机的热效率不超过 50%，则不予奖励；若发动机的热效率超过 50% 但不超过 53%，则每台发动机奖励 2 万元；若发动机的热效率超过 53%，则每台发动机奖励 5 万元．求每台发动机获得奖励的数学期望．

附：刻画回归效果的相关指数 $R^2=1-\dfrac{\sum_{i=1}^{n}(y_i-\overline{y_i})^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}$，$\sqrt{17}\approx 4.1$．用最小二乘法求线性回归方程 $\hat{y}=\hat{b}x+\hat{a}$ 的系数公式：$\hat{b}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum_{i=1}^{n}x_i^2-n\overline{x}^2}=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$；$\hat{a}=\hat{y}-\hat{b}\overline{x}$．随机变量 ξ 服从正态分布 $N(\mu,\sigma^2)$，则 $P(\mu-\sigma<\xi<\mu+\sigma)=0.6827$，$P(\mu-2\sigma<\xi<\mu+2\sigma)=0.9545$．

【分析】（1）由表格中的数据，结合刻画回归效果的相关指数，可得结论；

（2）求得样本中心点，可得当 $x>17$ 亿元时，y 与 x 满足的线性回归方程，令 $x=20$，可得所求大小；

（3）由正态分布的计算公式，以及数学期望公式，可得所求值．

【解答】解：（1）由表格中的数据，有$182.4>79.2$，即$\frac{182.4}{\sum_{i=1}^{7}(y_i-\overline{y})^2}>\frac{79.2}{\sum_{i=1}^{7}(y_i-\overline{y})^2}$，

所以模型①的R^2小于模型②，说明回归模型②刻画的拟合效果更好．

所以当$x=17$亿元时，科技改造直接收益的预测值为：

$\hat{y}=21.3\times\sqrt{17}-14.4=72.93$（亿元）．

（2）由已知可得：$\overline{x}-20=\frac{0.5+2+3.5+4+5}{5}=3$，

$\therefore\overline{x}=23$，$\overline{y}-60=\frac{8.5+8+7.5+6+6}{5}=7.2$，$\therefore\overline{y}=67.2$，

$\therefore a=\overline{y}+0.7\overline{x}=67.2+0.7\times23=83.3$，

$\therefore$当$x>17$亿元时，y与x满足的线性回归方程为：$\hat{y}=-0.7x+83.3$，

$\therefore$当$x=20$亿元时，科技改造直接收益的预测值$\hat{y}=-0.7\times20+83.3=69.3$，

$\therefore$当$x=20$亿元时，实际收益的预测值为$69.3+10=79.3$亿元>72.93亿元，

$\therefore$技改造投入 20 亿元时，公司的实际收益的更大．

（3）$\because P(0.52-0.02<X<0.52+0.02)=0.9544$，

$P(X>0.50)=\frac{1+0.9544}{2}=0.9772$，$P(X,,0.5)=\frac{1-0.9544}{2}=0.0228$，

$\because P(0.52-0.1<X<0.52+0.1)=0.6826$，

$\therefore P(X>0.53)=\frac{1-0.6826}{2}=0.1587$，

$\therefore P(0.50<X,,0.53)=0.9772-0.1587=0.8185$，

设每台发动机获得的奖励为Y（万元），则Y的分布列为：

Y	0	2	5
P	0.0228	0.8185	0.1587

$\therefore$每台发动机获得奖励的数学期望为$E(Y)=0\times0.0228+2\times0.8185+5\times0.1587=2.4305$（万元）．

【点评】本题考查线性回归方程的求法和运用，考查离散型随机变量的分布列、数学期望的求法，考查运算求解能力，是中档题．

24．（2020•运城一模）在创建“全国文明卫生城”过程中，运城市“创城办”为了调查市民对创城工作的了解情况，进行了一次创城知识问卷调查（一位市民只能参加一次），通过随机抽样，得到参加问卷调查的100 人的得分统计结果如表所示：．

组别	[30， 40)	[40， 50)	[50， 60)	[60， 70)	[70， 80)	[80， 90)	[90， 100)
频数	2	12	20	25	24	13	4

（1）由频数分布表可以大致认为，此次问卷调查的得分 $Z-N(\mu,198)$，μ 近似为这 100 人得分的平均值（同一组中的数据用该组区间的中点值作代表），利用该正态分布，求 $P(38.2<Z,,80.2)$；

（2）在（1）的条件下，“创城办”为此次参加问卷调查的市民制定如下奖励方案：

①得分不低于 μ 的可以获赠 2 次随机话费，得分低于 μ 的可以获赠 1 次随机话费；

②每次获赠的随机话费和对应的概率为：

赠送话费的金额（单位：元）	20	50
概率	$\frac{3}{4}$	$\frac{1}{4}$

现有市民甲参加此次问卷调查，记 X （单位：元）为该市民参加问卷调查获赠的话费，求 X 的分布列与数学期望.

附：参考数据与公式：$\sqrt{198}\approx14$，若 $X\sim N(\mu,\sigma^2)$，则 $P(\mu-\sigma<X,,\mu+\sigma)=0.6826$；$P(\mu-2\sigma<X,,\mu+2\sigma)=0.9544$，$P(\mu-3\sigma<X,,\mu+3\sigma)=0.9974$.

【分析】（1）先算出每个区间上的频率，然后套公式求出 μ 的值．然后根据正态分布密度曲线的性质

（2）显然得分超过与不超过 μ 的概率各占 0.5，根据题意可知，所得话费分别为 20 元、40 元、50 元、70 元、100 元，然后分别算出它们的概率，即可得到分布列，求出期望.

【解答】解：（1）$\mu=35\times0.02+45\times0.12+55\times0.20+65\times0.25+75\times0.24+85\times0.13+95\times0.04=66.2$.

故 $Z\sim N(66.2,198)$，易知 $\sigma=\sqrt{198}\approx14$．∴ $P(Z,,80.2)=1-\frac{1-P(66.2-14<Z,,66.2+14)}{2}=1-\frac{1-0.6826}{2}=0.8413$.

又 $P(Z,,38.2)=\frac{1-P(66.2-2\times14<Z,,66.2+2\times14)}{2}=\frac{1-0.9544}{2}=0.0228$.

故 $P(38.2<Z$翏$80.2)=P(Z\ \ 80.2)-P(Z?\ 38.2)=0.8185$.

（2）由题意一位市民得分高于或低于 μ 的概率各为 $\frac{1}{2}$．该市民参与活动获赠话费 X 的可能取值为 20，40，50，70，100.

∴ $P(X=20)=\frac{1}{2}\times\frac{3}{4}=\frac{3}{8}$；$P(X=40)=\frac{1}{2}\times\frac{3}{4}\times\frac{3}{4}=\frac{9}{32}$；$P(X=50)=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8}$；$P(X=70)=\frac{1}{2}\times\frac{3}{4}\times\frac{1}{4}\times2=\frac{3}{16}$；$P(X=100)=\frac{1}{2}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{32}$.

故 X 的分布列为：

X	20	40	50	70	100
P	$\frac{3}{8}$	$\frac{9}{32}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{32}$

所以 $EX=20\times\frac{3}{8}+40\times\frac{9}{32}+50\times\frac{1}{8}+70\times\frac{3}{16}+100\times\frac{1}{32}=41.25$.

故该市民参与活动获赠话费的数学期望为 41.25 元.

【点评】本题考查了频率直方图的识图，二项分布、正态分布的性质及应用，同时也考查了一般情况下分布列的求法．属于中档题．

25．（2020•太原一模）新冠病毒是一种通过飞沫和接触传播的变异病毒，为筛查该病毒，有一种检验方式是检验血液样本相关指标是否为阳性，对于 a 份血液样本，有以下两种检验方式：一是逐份检验，则需检验 n 次．二是混合检验，将其中 k 份血液样本分别取样混合在一起，若检验结果为阴性，那么这 k 份血液全为阴性，因而检验一次就够了；如果检验结果为阳性，为了明确这 k 份血液究竟哪些为阳性，就需要对它们再逐份检验，此时 k 份血液检验的次数总共为 $k+1$ 次．某定点医院现取得 4 份血液样本，考虑以下三种检验方案：方案一，逐个检验；方案二，平均分成两组检验；方案三，四个样本混在一起检验．假设在接受检验的血液样本中，每份样本检验结果是阳性还是阴性都是相互独立的，且每份样本是阴性的概率为 $P=\frac{2\sqrt{2}}{3}$.

（Ⅰ）求把 2 份血液样本混合检验结果为阳性的概率；

（Ⅱ）若检验次数的期望值越小，则方案越“优”．方案一、二、三中哪个最“优”？请说明理由．

【分析】（Ⅰ）该混合样本阴性的概率是 $(\frac{2\sqrt{2}}{3})^2=\frac{8}{9}$，根据对立事件原理，能求出阳性的概率．

（Ⅱ）方案一：逐个检验，检验次数为 4，方案二：每组 2 个样本检验时，若阴性则检测次数为 1，概率为 $\frac{8}{9}$，若阳性，则检测次数为 3，概率为 $\frac{1}{9}$，设方案二的检验次数记为 ξ，则 ξ 的可能取值为 2，4，6，求出分布列，得到 $E(\xi)=\frac{22}{9}$，方案三：混在一起检验，设方案三的检验次数记为 η，η 的可能取值为 1，5，由分布列求出 $E(\eta)=\frac{149}{81}$，从而选择方案三最“优”．

【解答】解：（Ⅰ）该混合样本阴性的概率是 $(\frac{2\sqrt{2}}{3})^2=\frac{8}{9}$，

根据对立事件原理，阳性的概率为 $1-\frac{8}{9}=\frac{1}{9}$.

（Ⅱ）方案一：逐个检验，检验次数为 4，

方案二：由（Ⅰ）知，每组 2 个样本检验时，若阴性则检测次数为 1，概率为 $\frac{8}{9}$，

若阳性，则检测次数为 3，概率为 $\frac{1}{9}$，

设方案二的检验次数记为 ξ，则 ξ 的可能取值为 2，4，6，

其分布列为：

ξ	2	4	6
P	$\frac{64}{81}$	$\frac{16}{81}$	$\frac{1}{81}$

$\therefore E(\xi)=2\times\frac{64}{81}+4\times\frac{16}{81}+6\times\frac{1}{81}=\frac{22}{9}$，

方案三：混在一起检验，设方案三的检验次数记为η，η的可能取值为1，5，

其分布列为：

η	1	5
P	$\frac{64}{81}$	$\frac{17}{81}$

$E(\eta)=1\times\frac{64}{81}+5\times\frac{17}{81}=\frac{149}{81}$，

$\because E(\eta)<E(\xi)<4$，故选择方案三最“优”.

【点评】本题考查概率的求法，考查对立事件概率计算公式、相互独立事件概率计算公式、古典概型等基础知识，考查运算求解能力，是中档题.

26.（2020•赣州模拟）春节期间爆发的新型冠状病毒$(COVID-19)$是新中国成立以来感染人数最多的一次疫情，一个不知道自己已感染但处于潜伏期的甲从疫区回到某市过春节，回到家乡后与朋友乙、丙、丁相聚过，最终乙、丙、丁也感染了新冠病毒、可以肯定的是乙受甲感染的，丙是受甲或乙感染的，假设他受甲和受乙感染的概率分别是0.6和0.4，丁是受甲、乙或丙感染的，假设他受甲、乙和丙感染的概率分别是0.2、0.4和0.4．在这种假设之下，乙、丙、丁中直接受甲感染的人数为X．

（1）求X的分布列和数学期望；

（2）该市在发现在本地出现新冠病毒感染者后，迅速采取应急措施，其中一项措施是各区必须每天及时上报新增疑似病例人数，A区上报的连续7天新增疑似病例数据是“总体均值为3，中位数4”，B区上报的连续7天新增疑似病例数据是“总体均值为2，总体方差为3“．设A区和B区连续7天上报新增疑似病例人数分别为x_1，x_2，……，x_7和y_1，y_2，……，y_7，x_i和$y_i(1\leqslant i\leqslant 7,i\in N)$分别表示$A$区和$B$区第$i$天上报新增疑似病例人数($x_i$和$y_i$均为非负)．记$M=max\{x_1,x_2,\ldots\ldots,x_7\}$，$N=max\{y_1,y_2,\ldots\ldots,y_7\}$．

①试比较M和N的大小；

②求M和N中较小的那个字母所对应的7个数有多少组？

【分析】（1）记事件C：“丙受甲感染”，事件D：“丁受甲感染”，则P（C）$=0.6$，P（D）$=0.2$，X的取值为1，2，3，分别求出相应的概率，由此能求出X的分布列和EX．

（2）(i) 对于 B 区，由 $(y_1-2)^2+(y_2-2)^2+(y_3-2)^2+(y_4-2)^2+(y_5-2)^2+(y_6-2)^2+(y_7-2)^2=21$，知 $(y_i-1)^2\leqslant 21$，$(i=1,2,\ldots,7)$，从而求出 $N\leqslant 6$，当 y_1，y_2，y_3，…，y_7 中有一个取 6，有一个取 2，其余取 1 时，$N=6$，对于 A 区，当 $x_1=x_2=x_3=0$ 时，$x_4=x_5=x_6=4$，$x_7=9$ 时，满足“总体均值为 3，中位数为 4”，由此能求出 $N<M$．

(ii) 当 $N=6$ 时，y_1，y_2，…，y_7 只有两种情况，①有一个是 6，有五个是 1 或 3，有一个是 2，②有一个是 6，有一个是 1 或 3，有一个是 0 或 4，其余是 2，由此能求出结果．

【解答】解：（1）记事件 C：“丙受甲感染”，事件 D：“丁受甲感染”，

则 P（C）$=0.6$，P（D）$=0.2$，

X 的取值为 1，2，3，

$P(X=1)=P(\bar{C}\bar{D})=0.4\times 0.8=0.32$，

$P(X=2)=P(C\bar{D})+P(\bar{C}D)=0.6\times 0.8+0.4\times 0.2=0.56$，

$P(X=3)=P(CD)=0.6\times 0.2=0.12$，

$\therefore X$ 的分布列为：

X	1	2	3
P	0.32	0.56	0.12

$EX=1\times 0.32+2\times 0.56+3\times 0.12=1.8$．

（2）(i) 对于 B 区，由 $(y_1-2)^2+(y_2-2)^2+(y_3-2)^2+(y_4-2)^2+(y_5-2)^2+(y_6-2)^2+(y_7-2)^2=21$，

知 $(y_i-1)^2\leqslant 21$，$(i=1,2,\ldots,7)$，

$\because y_i$ 是非负整数，$\therefore |y_i-2|\leqslant 4$，$\therefore y_i\leqslant 6$，$\therefore N\leqslant 6$，

当 y_1，y_2，y_3，…，y_7 中有一个取 6，有一个取 2，其余取 1 时，$N=6$，

对于 A 区，当 $x_1=x_2=x_3=0$ 时，$x_4=x_5=x_6=4$，

$x_7=9$ 时，满足“总体均值为 3，中位数为 4”，此时 $m=9$，

$\therefore N<M$．

(ii) 当 $N=6$ 时，y_1，y_2，…，y_7 只有两种情况，

①有一个是 6，有五个是 1 或 3，有一个是 2，

②有一个是 6，有一个是 1 或 3，有一个是 0 或 4，其余是 2，

对于①，共有 $C_7^1C_6^5\times2^5=1344$ 组，

对于②，共有 $C_7^1C_6^1C_5^1\times2^2=840$ 组，

故共有：$1344+840=2184$ 组．

【点评】本题考查概率、离散型随机变量的分布列、数学期望的求法，考查互斥事件概率加法公式和相互独立事件概率加法公式等基础知识，考查运算求解能力，是中档题．

27．（2020•山东模拟）2019 年由袁隆平团队研发的第三代杂交水稻 10 月 21 日至 22 日首次公开测产，经测产专家组评定，最终亩产为 1046.3 公斤，第三代杂交水稻的综合优势可以推动我国的水稻生产向更加优质、高产、绿色和可持续方向发展．某企业引进一条先进的食品生产线，计划以第三代杂交水稻为原料进行深加工，创建一个新产品，已知该产品的质量以某项指标值 $k(k\in[70,100])$ 为衡量标准，其质量指标的等级划分如表：

质量指标值 k	90≤k≤100	85≤k<90	80≤k<85	75≤k<80	70≤k<75
产品等级	废品	合格	良好	优秀	良好

为了解该产品的生产效益，该企业先进行试生产，从中随机抽取了 1000 件产品，测量了每件产品的指标值，得到产品质量指标值 k 的频率分布直方图如图．

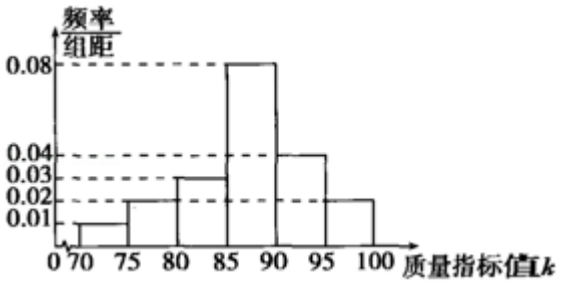
（1）若从质量指标值不小于 85 的产品中利用分层抽样的方法抽取 7 件产品，并采集相关数据进行分析，然后从这 7 件产品中任取 3 件产品，求质量指标值 $k\in[90，95)$ 的件数 X 的分布列及数学期望；

（2）若将频率视为概率，从该产品中有放回地随机抽取 3 件产品，记“抽出的产品中至少有 1 件为合格或合格以上等级“为事件 A，求事件 A 发生的概率；

（3）若每件产品的质量指标值 k 与利润 y（单位：元）的关系如表所示 $(1<t<4)$：

质量指标值 k	90≤k≤100	85≤k<90	80≤k<85	75≤k<80	70≤k<75
利润 y（元）	$-e^t$	t	$2t$	$4t$	$3t$

请问生产该产品能否盈利？若不能，试说明理由；若能，试确定 t 为何值时，利润达到最大（参考数值：$ln2\approx0.7$，$ln3\approx1.1$，$ln5\approx1.6$）．



【分析】（1）由频率分布直方图和分层抽样求出抽取的 7 件产品中，$k\in[85，90)$ 的有 4 件，$k\in[90，95)$ 的有 2 件，$k\in[95，100)$ 的有 1 件，从这 7 件产品中，任取 3 件，质量指标值 $k\in[90，95)$ 的件数 X 的所有可能取值为 0，1，2，分别求出相应的概率，由此能求出 X 的分布列和 $E(X)$.

（2）设事件 A 的合格率为 P（A），则根据概率分布直方图求出一件产品为合格或合格以上等级的概率，由此能求出事件 A 发生的概率.

（3）由频率分布直方图可得该产品的质量指标值 k 与利润 y（元）的关系，从而求出每件产品的利润 $y=-0.3e^{t}+1.25t$，$(1<t<4)$，则 $y'=-0.3e^{t}+1.25$，利用导数性质能求出生产该产品能够实现盈利，当 $t=ln\frac{25}{6}=1.4$ 时，每件产品的利润取得最大值为 0.5 元.

【解答】解：（1）由频率分布直方图得指标值不小于 85 的产品中，

$k\in[85，90)$ 的频率为 $0.08\times5=0.4$，

$k\in[90，95)$ 的频率为 $0.04\times5=0.2$，

$k\in[95，100]$ 的频率为 $0.02\times5=0.1$，

$\therefore$ 利用分层抽样抽取的 7 件产品中，$k\in[85，90)$ 的有 4 件，

$k\in[90，95)$ 的有 2 件，$k\in[95，100)$ 的有 1 件，

从这 7 件产品中，任取 3 件，质量指标值 $k\in[90，95)$ 的件数 X 的所有可能取值为 0，1，2，

$$P(X=0)=\frac{C_5^3}{C_7^3}=\frac{2}{7},$$

$$P(X=1)=\frac{C_2^1C_5^2}{C_7^3}=\frac{4}{7},$$

$$P(X=2)=\frac{C_2^2C_5^1}{C_7^3}=\frac{1}{7},$$

$\therefore X$ 的分布列为：

X	0	1	2

P	$\frac{2}{7}$	$\frac{4}{7}$	$\frac{1}{7}$

$E(X)=0\times\frac{2}{7}+1\times\frac{4}{7}+2\times\frac{1}{7}=\frac{6}{7}$.

（2）设事件 A 的合格率为 P（A），则根据概率分布直方图得：

一件产品为合格或合格以上等级的概率为 $p=1-(0.04+0.02)\times5=0.7$，

∴事件 A 发生的概率 P（A）$=C_3^3\times0.7^0\times0.3^3=0.973$.

（3）由频率分布直方图可得该产品的质量指标值 k 与利润 y（元）的关系与表所示 $(1<t<4)$，

质量指标值 k	90,, k 100	85,, $k<90$	80,, $k<85$	75,, $k<80$	70,, $k<75$
利润 y	$-e^t$	t	$2t$	$4t$	$3t$
P	0.3	0.4	0.15	0.1	0.05

∴每件产品的利润：

$y=-0.3e^t+0.4t+0.3t+0.4t+0.15t=-0.3e^t+1.25t$，$(1<t<4)$，

则 $y'=-0.3e^t+1.25$，

令 $y'=-0.3e^t+1.25=0$，解得 $t=ln\frac{25}{6}$，

∴当 $t\in(1,ln\frac{25}{6})$ 时，$y'>0$，函数 $y=-0.3e^t+1.25$ 单调递增，

当 $t\in(ln\frac{25}{6}$，4) 时，$y'<0$，函数 $y=-0.3e^t+1.25t$，单调递减，

∴当 $t=ln\frac{25}{6}$ 时，y 取最大值，为 $-0.3e^{ln\frac{25}{6}}+1.25\times ln\frac{25}{6}=-\frac{3}{10}\times\frac{25}{6}+\frac{5}{4}\times(2ln5-ln2-ln3)=0.5$，

∴生产该产品能够实现盈利，当 $t=ln\frac{25}{6}=1.4$ 时，每件产品的利润取得最大值为 0.5 元.

【点评】本题考查离散型随机变量的分布列、数学期望、利润最大值的求法，考查频率分布直方图、分层抽样、导数性质等基础知识，考查运算求解能力，是中档题.

28.（2020•聊城一模）2020 年春节期间，武汉市爆发了新型冠状病毒肺炎疫情，在党中央的坚强领导下，全国人民团结一心，众志成城，共同抗击疫情．某中学寒假开学后，为了普及传染病知识，增强学生的防范意识，提高自身保护能力，校委会在全校学生范围内，组织了一次传染病及个人卫生相关知识有奖竞赛（满分 100 分），竞赛奖励规则如下，得分在[70，80) 内的学生获三等奖，得分在[80，90) 内的学生获二等奖，得分在[90，100] 内的学生获一等奖，其他学生不得奖．教务处为了解学生对相关知识的掌握情况，随机抽取了 100 名学生的竞赛成绩，并以此为样本绘制了如下样本频率分布直方图.

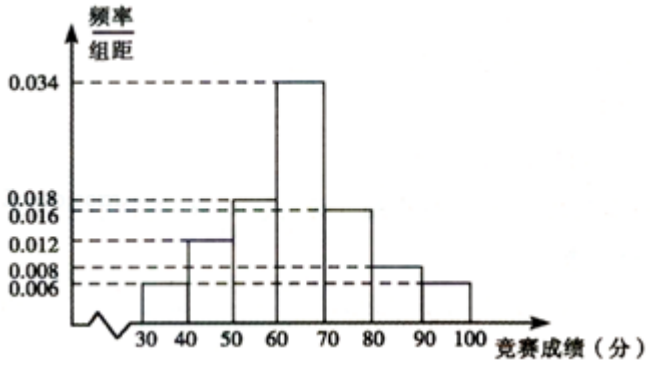
（1）现从该样本中随机抽取两名学生的竞赛成绩，求这两名学生中恰有一名学生获奖的概率；

（2）若该校所有参赛学生的成绩 X 近似服从正态分布 $N(\mu,\sigma^2)$，其中 $\sigma\approx15$，μ 为样本平均数的估计值，利用所得正态分布模型解决以下问题：

(i) 若该校共有 10000 名学生参加了竞赛，试估计参赛学生中成绩超过 79 分的学生数（结果四舍五入到整数）；

(ii) 若从所有参赛学生中（参赛学生数大于10000) 随机抽取 3 名学生进行座谈，设其中竞赛成绩在 64 分以上的学生数为 ξ，求随机变量 ξ 的分布列和均值.

附：若随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，则 $P(\mu-\sigma<X,,\mu+\sigma)\approx0.6827$，$P(\mu-2\sigma<X,,\mu+2\sigma)\approx0.9544$，$P(\mu-3\sigma<X,,\mu+3\sigma)\approx0.9973$.



【分析】（1）由样本频率分布直方图得样本中获一等奖的 6 人，获二等奖的 8 人，获三等奖的 16 人，共 30 人获奖，70 人没有获奖，从该样本中随机抽取两名学生的竞赛成绩，基本事件总数为 $n=C_{100}^2$，设“抽取的两名学生中恰有一名学生获奖”为事件 A，则事件 A 包含的基本事件的个数为 $m=C_{70}^1C_{30}^1$，由此能求出这两名学生中恰有一名学生获奖的概率.

（2）(i) 由样本频率分布直方图得样本平均数的估计值，得到所有参赛学生的成绩 X 近似值服从正态分布 $N(64,15^2)$，由此能估计参赛学生中成绩超过 79 分的学生数.

(ii) 由 $\mu=64$，得 $P(X>64)=\frac{1}{2}$，得到随机变量 $\xi\sim B(3,\frac{1}{2})$，由此能求出 ξ 的分布列和 $E(\xi)$.

【解答】解：（1）由样本频率分布直方图得样本中获一等奖的 6 人，获二等奖的 8 人，

获三等奖的 16 人，共 30 人获奖，70 人没有获奖，

从该样本中随机抽取两名学生的竞赛成绩，基本事件总数为 $n=C_{100}^2$，

设“抽取的两名学生中恰有一名学生获奖”为事件 A，

则事件 A 包含的基本事件的个数为 $m=C_{70}^1C_{30}^1$，

$\therefore$ 这两名学生中恰有一名学生获奖的概率 p（A）$=\frac{C_{70}^1C_{30}^1}{C_{100}^2}=\frac{14}{33}$.

（2）由样本频率分布直方图得样本平均数的估计值：

$\mu=35\times0.006\times10+45\times0.012\times10+55\times0.018\times10+65\times0.034\times10+75\times0.016\times10+85\times0.008\times10+95\times0.006\times10=64$，

所有参赛学生的成绩 X 近似值服从正态分布 $N(64,\ 15^2)$，

$(i)\because \mu+\sigma=79$，$\therefore P(X>79)\approx\dfrac{1-0.6827}{2}=0.15865$，

$\therefore$ 估计参赛学生中成绩超过 79 分的学生数为：

$0.15868\times10000\approx1587$（人）.

(ii) 由 $\mu=64$，得 $P(X>64)=\dfrac{1}{2}$，

即从所有参赛学生中随机抽取 1 名学生，该生竞赛成绩在 64 分以上的概率为 $\dfrac{1}{2}$，

$\therefore$ 随机变量 $\xi\sim B(3,\dfrac{1}{2})$，

$P(\xi=0)=C_3^0(\dfrac{1}{2})^3=\dfrac{1}{8}$，

$P(\xi=1)=C_3^1(\dfrac{1}{2})(\dfrac{1}{2})^2=\dfrac{3}{8}$，

$P(\xi=2)=C_3^2(\dfrac{1}{2})^2(\dfrac{1}{2})=\dfrac{3}{8}$，

$P(\xi=3)=C_3^3(\dfrac{1}{2})^3=\dfrac{1}{8}$，

$\therefore \xi$ 的分布列为：

ξ	0	1	2	3
P	$\dfrac{1}{8}$	$\dfrac{3}{8}$	$\dfrac{3}{8}$	$\dfrac{1}{8}$

$\therefore E(\xi)=0\times\dfrac{1}{8}+1\times\dfrac{3}{8}+2\times\dfrac{3}{8}+3\times\dfrac{1}{8}=\dfrac{3}{2}$.

【点评】本题考查概率、离散型随机变量的分布列、数学期望的求法，考查频率分布直方图、二项分布的性质等基础知识，考查运算求解能力，是中档题.

29.（2020•青岛模拟）某网络购物平台每年 11 月 11 日举行“双十一”购物节，当天有多项优惠活动，深受广大消费者喜爱.

（1）已知该网络购物平台近 5 年“双十”购物节当天成交额如表：

年份	2015	2016	2017	2018	2019
成交额（百亿元）	9	12	17	21	27

求成交额 y（百亿元）与时间变量 x（记 2015 年为 $x=1$，2016 年为 $x=2$，……依此类推）的线性回归方

程，并预测 2020 年该平台“双十一”购物节当天的成交额（百亿元）；

（2）在 2020 年“双十一”购物节前，某同学的爸爸、妈妈计划在该网络购物平台．上分别参加 A、B 两店各一个订单的“秒杀”抢购，若该同学的爸爸、妈妈在 A、B 两店订单“秒杀”成功的概率分别为 p、q，记该同学的爸爸和妈妈抢购到的订单总数量为 X．

(i) 求 X 的分布列及 $E(X)$；

(ii) 已知每个订单由 $k(k..2,k\in N^*)$ 件商品 W 构成，记该同学的爸爸和妈妈抢购到的商品 W 总数量为 Y，假设 $p=\dfrac{7\sin\dfrac{\pi}{k}}{4k}-\dfrac{\pi}{k^2},q=\dfrac{\sin\dfrac{\pi}{k}}{4k}$，求 $E(Y)$ 取最大值时正整数 k 的值．

附：回归方程 $\hat{y}=\hat{b}x+a$ 中斜率和截距的最小二乘估计公式分别为：$\hat{b}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum\limits_{i=1}^{n}x_i^{\ 2}-n\overline{x}^2}=\dfrac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}$；

$a=\overline{y}-\hat{b}\overline{x}$．

【分析】（1）计算 $\overline{x}$、$\overline{y}$，求出系数 $\hat{b}$ 和 a，写出线性回归方程，利用方程计算 $x=6$ 时 $\hat{y}$ 的值即可；

（2）(i) 由题意知随机变量 X 的可能取值，计算对应的概率值，写出分布列，求出数学期望值；

(ii) 根据题意求出 $E(Y)$ 的解析式，利用换元法和求导法计算 $E(Y)$ 取最大值时正整数 k 的值．

【解答】解：（1）由题意，计算 $\overline{x}=\dfrac{1}{5}\times(1+2+3+4+5)=3$，$\overline{y}=\dfrac{1}{5}\times(9+12+17+21+27)=17.2$，

$\sum\limits_{i=1}^{5}x_iy_i=1\times9+2\times12+3\times17+4\times21+5\times27=303$；

$\sum\limits_{i=1}^{5}x_i^{\ 2}=1^2+2^2+3^2+4^2+5^2=55$，

所以 $\hat{b}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\overline{x}\overline{y}}{\sum\limits_{i=1}^{n}x_i^{\ 2}-n\overline{x}^2}=\dfrac{303-5\times3\times17.2}{55-5\times3^2}=\dfrac{45}{10}=4.5$；

$a=\overline{y}-\hat{b}\overline{x}=17.2-4.5\times3=3.7$，

所以 y 与 x 的线性回归方程是 $\hat{y}=4.5x+3.7$，

当 $x=6$ 时，$\hat{y}=4.5\times6+3.7=30.7$，

所以预测 2020 年该平台“双十一”购物节当天的成交额为 30.7 百亿元；

（2）(i) 由题意知，随机变量 X 的可能取值为 0，1，2；

计算 $P(X=0)=(1-p)(1-q)$，

$P(X=1)=(1-p)q+(1-q)p$，

$P(X=2)=pq$；

所以 X 的分布列为：

X	0	1	2
P	$(1-p)(1-q)$	$(1-p)q+(1-q)p$	pq

计算数学期望值为 $E(X)=0\times(1-p)(1-q)+1\times(1-p)q+(1-q)p+2\times pq=p+q$；

(ii) 因为 $Y=kX$，所以 $E(Y)=kE(X)=k(p+q)=k(\frac{7\sin\frac{\pi}{k}}{4k}-\frac{\pi}{k^2}+\frac{\sin\frac{\pi}{k}}{4k})=2\sin\frac{\pi}{k}-\frac{\pi}{k}$；

令 $t=\frac{1}{k}\in(0,\ \frac{1}{2}]$，设 $f(t)=2\sin\pi t-\pi t$，则 $E(Y)=f(t)$；

因为 $f'(t)=2\pi\cos\pi t-\pi=2\pi(\cos\pi t-\frac{1}{2})$，且 $\pi t\in(0,\ \frac{\pi}{2}]$；

所以当 $t=\frac{1}{k}\in(0,\frac{1}{3})$ 时，$f'(t)>0$，$f(t)$ 单调递增；

当 $t=\frac{1}{k}\in(\frac{1}{3},\ \frac{1}{2})$ 时，$f'(t)<0$，$f(t)$ 单调递减；

所以当 $t=\frac{1}{k}=\frac{1}{3}$，即 $k=3$ 时，$f(t)_{,,}f(\frac{1}{3})=\sqrt{3}-\frac{\pi}{3}$；

所以 $E(Y)$ 取最大值时正整数 k 的值为 3．

【点评】本题主要考查了概率与随机变量的分布列和数学期望的计算问题，也考查了运用概率统计知识解决简单实际问题的能力，是中档题．

30．（2020•深圳二模）2020 年初，新冠肺炎疫情袭击全国，某省由于人员流动性较大，成为湖北省外疫情最严重的省份之一，截至 2 月 29 日，该省已累计确诊 1349 例患者（无境外输入病例）．

（1）为了解新冠肺炎的相关特征，研究人员从该省随机抽取 100 名确诊患者，统计他们的年龄数据，得如表的频数分布表：

年龄	[10，20]	(20，30]	(30，40]	(40，50]	(50，60]	(60，70]	(70，80]	(80，90]	(90，100]
人数	2	6	12	18	22	22	12	4	2

由频数分布表可以大致认为，该省新冠肺炎患者的年龄服从正态分布 $N(\mu,\ 15.2^2)$，其中 μ 近似为这 100 名患者年龄的样本平均数（同一组中的数据用该组区间的中点值作代表）．请估计该省新冠肺炎患者年龄在 70 岁以上 $(..70)$ 的患者比例；

（2）截至 2 月 29 日，该省新冠肺炎的密切接触者（均已接受检测）中确诊患者约占10%，以这些密切接触者确诊的频率代替 1 名密切接触者确诊发生的概率，每名密切接触者是否确诊相互独立．现有密切接触

者 20 人，为检测出所有患者，设计了如下方案：将这 20 名密切接触者随机地按 $n(1<n<20$ 且 n 是 20 的约数）个人一组平均分组，并将同组的 n 个人每人抽取的一半血液混合在一起化验，若发现新冠病毒，则对该组的 n 个人抽取的另一半血液逐一化验，记 n 个人中患者的人数为 X_n，以化验次数的期望值为决策依据，试确定使得 20 人的化验总次数最少的 n 的值.

参考数据：若 $Z\sim N(\mu,\sigma^2)$，则 $P(\mu-\sigma<Z<\mu+\sigma)=0.6826$，$P(\mu-2\sigma<Z<\mu+2\sigma)=0.9544$，$P(\mu-3\sigma<Z<\mu+3\sigma)=0.9973$，$0.9^4\approx0.66$，$0.9^5\approx0.59$，$0.9^{10}\approx0.35$.

【分析】（1）先由题目给的数据根据公式求出均值，根据正态分布密度函数的性质即可年龄在 70 岁以上的患者比例，即 $P(Z..70)$.

（2）20 人平均分成 n 组，可知 n 应该可以整除 20 方可，并且每组内患者的人数服从二项分布；然后算出每种分组中化验次数的期望，横向比较，取最小值即可获解.

【解答】解：（1）$\mu=\dfrac{2\times15+6\times25+12\times35+18\times45+22\times55+22\times65+12\times75+4\times85+2\times95}{100}=54.8$

所以 $P(54.8-15.2<X<54.8+15.2)=P(39.6<X<70)=0.6826$.

$P(Z..70)=\dfrac{1-P(39.6<Y<70)}{2}=\dfrac{1-0.6826}{2}=0.1587=15.87\%$.

所以该省新冠肺炎患者年龄在 70 岁以上 $(..70)$ 的患者比例为 15.87%.

（2）由题意，每名密切接触者确诊为新冠脑炎的概率均为 $\dfrac{1}{10}$，n 的可能取值为 2，4，5，10. 且 $X_n\sim B(n,\dfrac{1}{10})$

对于某组 n 个人，化验次数 Y 的可能取值为 1，$n+1$.

$P(Y=1)=(\dfrac{9}{10})^n$，$P(Y=n+1)=1-(\dfrac{9}{10})^n$

所以 $E(Y)=1\square(\dfrac{9}{10})^n+(n+1)[1-(\dfrac{9}{10})^n]=n+1-n(\dfrac{9}{10})^n$，

则 20 人的化验总次数为 $f(n)=\dfrac{20}{n}[n+1-n(\dfrac{9}{10})^n]=20[1+\dfrac{1}{n}-(\dfrac{9}{10})^n]$，

经计算 f（2）$=13.8$，f（4）≈11.8，f（5）≈12.2，$f(10)\approx15$.

所以，当 $n=4$ 时符合题意，即按 4 人一组检测，可是化验总次数最少.

【点评】本题是一道应用题，考查了正态分布密度函数的性质及应用，同时考查了二项分布、分布列的数学期望的求法，第二问也是一个优化方案问题，难度不大，需正确理解题意.