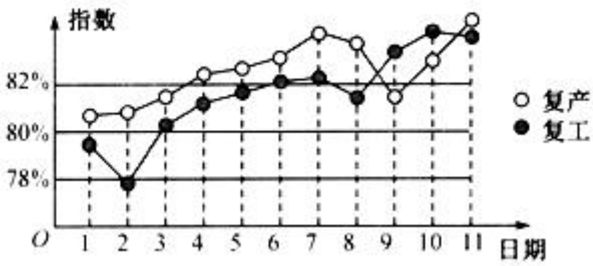
# 专题十一　概率统计

## 一、单项选择

**1.（菏泽一模 4）**2020 年 5 月我国抗击新冠肺炎疫情工作取得阶段性胜利，各地有序推进复工复产，下面是某地连续 11 天复工复产指数折线图，下列说法正确的是（　　）




A．这 11 天复工指数和复产指数均逐日增加

B．这 11 天期间，复产指数的极差大于复工指数的极差

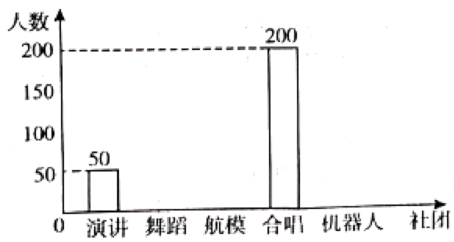
C．第 3 天至第 11 天复工复产指数均超过 80%

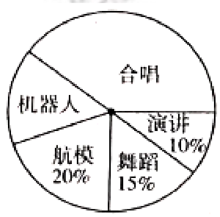
D．第 9 天至第 11 天复工指数的增量大于复产指数的增量

**2.（德州一模 4）**《史记》卷六十五《孙子吴起列传第五》中有这样一道题：齐王与田忌赛马，田忌的上等马劣于齐王的上等马，优于齐王的中等马，田忌的中等马劣于齐王的中等马，优于齐王的下等马，田忌的下等马劣于齐王的下等马，现两人进行赛马比赛，比赛规则为：每匹马只能用一次，每场比赛双方各出一匹马，共比赛三场．每场比赛中胜者得 1 分，否则得 0 分．若每场比赛之前彼此都不知道对方所用之马，则比赛结束时，田忌得 2 分的概率（　　）

A．$\frac{1}{3}$　　B．$\frac{2}{3}$　　C．$\frac{1}{6}$　　D．$\frac{1}{2}$

**3.（2021•临沂一模 4）**某学校组建了演讲、舞蹈、航模、合唱、机器人五个社团，全校 3000 名学生每人都参加且只参加其中一个社团，校团委从这 3000 名学生中随机选取部分学生进行调查，并将调查结果绘制了如下不完整的两个统计图：




则选取的学生中参加机器人社团的学生数为（　　）

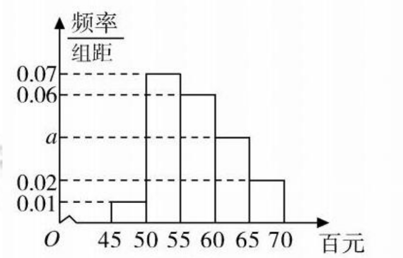
A．50　　B．75　　C．100　　D．125

**4.（聊城一模 5）**2021 年 2 月 25 日，全国脱贫攻坚总结表彰大会在北京隆重举行，会上习近平总书记庄严

宣告，我国脱贫攻坚取得了全面胜利，同时要切实做好巩固拓展脱贫攻坚成果同乡村振兴有效衔接各项工作．某县扶贫办积极响应党的号召，准备对 A 乡镇的三个脱贫村进一步实施产业帮扶．现有“特色种养”、“庭院经济”、“农产品加工”三类帮扶产业，每类产业中都有两个不同的帮扶项目，若要求每个村庄任意选取一个帮扶项目(不同村庄可选取同一个项目)，那么这三个村庄所选项目分别属于三类不同帮扶产业的概率为

A．$\frac{2}{9}$　　B．$\frac{1}{6}$　　C．$\frac{1}{3}$　　D．$\frac{2}{5}$

**5.（济宁一模 6）**为了解某贫困地区实施精准扶贫后的成果，现随机抽取了该地区部分人员，调查了 2020 年其人均纯收入状况.经统计，这批人员的年人均纯收入数据(单位:百元)全部介于 45 至 70 之间.将数据分成 5 组,并得到如图所示的频率分布直方图.现采取分层抽样的方法,从[55,60)，[60,65)，[65,70)这三个区间中随机抽取 6 人，再从 6 人中随机抽取 3 人，则这三人中恰有 2 人年人均纯收入位于[60,65)的概率是

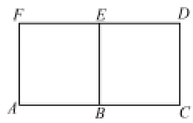


第6题图

A. $\frac{9}{10}$　　B. $\frac{3}{5}$

C. $\frac{9}{20}$　　D. $\frac{1}{5}$

**6.（济南一模 4）**环保部门为降低某社区在改造过程中产生的扬尘污染，决定对全部街道采取洒水降尘作业。该社区街道的平面结构如图所示（线段代表街道），洒水车随机选择 A,B,C,D,E,F 中的一点驶入进行作业，则选择的驶入点使洒水车能够不重复地走遍全部街道的概率为



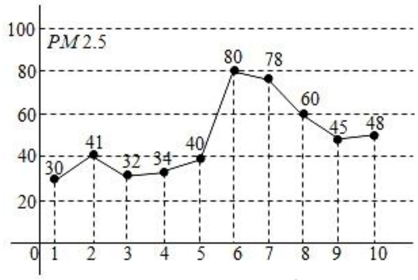

A. $\frac{1}{6}$　　B. $\frac{1}{3}$　　C. $\frac{1}{2}$　　D. $\frac{2}{3}$

**7.（菏泽一模 6）**菏泽万达商场在春节前开展商品促销活动，顾客凡购物金额满 50 元，则可以从“福”字、春联和灯笼这三类礼品中任意免费领取一件，若有 4 名顾客都领取一件礼品，则他们中有且仅有 2 人领取的礼品种类相同的概率是（　　）

A．$\frac{5}{9}$　　B．$\frac{4}{9}$　　C．$\frac{8}{9}$　　D．$\frac{9}{16}$

**二、多项选择**

**8．（日照一模 9）**PM2.5 是衡量空气质量得重要指标，我国采用世卫组织得最宽值限定值，即 PM2.5 日均值在 35$\mu g/m^3$ 以下空气质量为一级，在 35～75$\mu g/m^3$ 空气质量为二级，超过 75$\mu g/m^3$ 为超标.如图是某地 12 月 1 日至 10 日的 PM2.5（单位：$\mu g/m^3$）的日均值，则下列说法正确的是




A. 这 10 天中有 3 天空气质量为一级

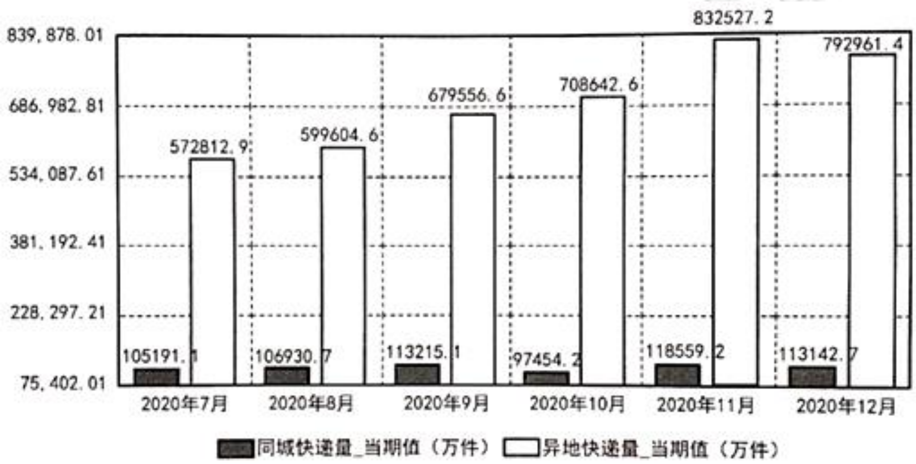
B.从 6 日到 9 日 PM2.5 日均值逐渐降低

C.这 10 天中 PM2.5 日均值的中位数是 55

D.这 10 天中 PM2.5 日均值的平均数是 45

**9．（2021•淄博一模 9）**快递行业作为邮政业的重要组成部分，具有带动产业领域广、吸纳就业人数多、经

济附加值高、技术特征显著等特点．它将信息传递、物品递送、资金流通和文化传播等多种功能融合在一起，关联生产、流通、消费、投资和金融等多个领域，是现代社会不可替代的基础产业．如图是国家统计局公布的 2020 年下半年快递运输量情况，请根据图中信息选出正确的选项（　　）




A．2020 年下半年，每个月的异地快递量都是同城快递量的 6 倍以上

B．2020 年 10 月份异地快递增长率小于 9 月份的异地快递增长率

C．2020 年下半年，异地快递量与月份呈正相关关系

D．2020 年下半年，同城和异地快递量最高均出现在 11 月

**10.（德州一模 9）**2020 年是全面实现小康社会目标的一年，也是全面打赢脱贫攻坚战的一年，某研究性学习小组调查了某脱贫县的甲、乙两个家庭，对他们过去 6 年（2014 年到 2019 年）的家庭收入情况分别进行统计，得到这两个家庭的年人均纯收入（单位：百元/人）茎叶图．对甲、乙两个家庭的年人均纯收入（以下分别简称“甲”“乙”）情况的判断，正确的是（　　）

| 甲 | | 乙 |
| --- | --- | --- |
| 8 7 7 6 | 3 | 4 6 8 9 |
| 2 0 | 4 | 0 1 |

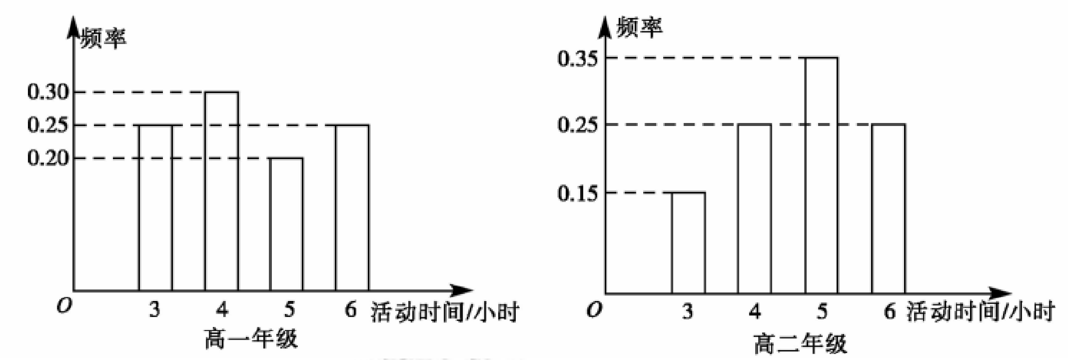
A．过去的 6 年，“甲”的极差小于“乙”的极差

B．过去的 6 年，“甲”的平均值小于“乙”的平均值

C．过去的 6 年，“甲”的中位数小于“乙”的中位数

D．过去的 6 年，“甲”的平均增长率小于“乙”的平均增长率

**11.（聊城一模 10）**某学校为了解高一、高二学生参加体育活动的时间情况，分别统计了这两个年级学生某周的活动时间，并制成了如图所示的条形图进行比较．则下列说法中正确的是




A．高二年级学生周活动时间的众数比高一年级的大
B．高二年级学生周活动时间的平均值比高一年级的小
C．高二年级学生周活动时间的中位数比高一年级的大
D．高二年级学生周活动时间的方差比高一年级的小

**12.（烟台一模 12）**骰子通常作为桌上游戏的小道具，最常见的骰子是六面骰，它是一个质地均匀的正方体，六个面上分别写有数字 1，2，3，4，5，6，现有一款闯关游戏，共有 4 关，规则如下：在第 n 关要抛掷六面骰 n 次，每次观察向上面的点数并做记录，如果这 n 次抛掷所出现的点数之和大于 $2^n+n$，则算闯过第 n 关，n=1,2,3,4，假定每次闯关互不影响，则

A.直线挑战第 2 关并过关的概率为$\frac{7}{12}$

B.连续挑战前两关并过关的的概率为$\frac{5}{24}$

C.若直接挑战第 3 关，设 A=“三个点数之和等于 15”，B=“至少出现一个 5 点”，则$P(A|B)=\frac{1}{13}$

D.若直接挑战第 4 关，则过关的概率是$\frac{35}{1296}$

## 三、填空

**13.（滨州一模 13）**某公司对近 5 年的年广告支出 *x*（单位：万元）与年利润 *y*（单位：万元）进行了初步统计如表所示：

| 年广告支出 *x* | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 年利润 *y* | 5 | 6 | *a* | 8 | 10 |

由上表中数据求得年广告支出 *x* 与年利润 *y* 满足线性回归方程 $\hat{y}=1.2\hat{x}+3.6$，则 *a* 的值为_____.

**14.（泰安一模 14）**某产品的广告费用 *x* 与销售额 *y* 的统计数据如表

| 广告费用 *x*（万元） | 4 | 2 | 3 | 5 |
|---|---|---|---|---|
| 销售额 *y*（万元） | 49 | 26 | 39 | 54 |

根据上表可得回归方程 $\hat{y}=\hat{b}\hat{x}+\hat{a}$ 中的 $\hat{b}$ 为 9.4，据此模型预报广告费用为 6 万元时销售额为_______.

15．（2021•临沂一模 15）蟋蟀鸣叫可以说是大自然优美、和谐的音乐，蟋蟀鸣叫的频率 $y$（每分钟鸣叫的次数）与气温 $x$（单位：℃）存在着较强的线性相关关系．某地研究人员根据当地的气温和蟋蟀鸣叫的频率得到了如下数据：

| $x$（℃） | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
|---|---|---|---|---|---|---|---|
| $y$（次数/分钟） | 24 | 28 | 31 | 39 | 43 | 47 | 54 |

利用如表中的数据求得回归直线方程为$\hat{y}=\hat{b}x+\hat{a}$，若利用该方程知，当该地的气温为 30℃时，蟋蟀每分钟鸣叫次数的预报值为 68，则$\hat{b}$的值为_____.

**16.（烟台一模 14）**2021 年 2 月 25 日，全国脱贫攻坚总结表彰大会在北京举行，习近平总书记庄严宣告我国脱贫攻坚战取得了全面胜利，已知在党委政府精准扶贫政策下，字 2017 年起某地区贫困户第 x 年的年人均收入 y（单位：万元）的统计数据如下表：

| 年份 | 2017 | 2018 | 2019 | 2020 |
|---|---|---|---|---|
| 年份编号 x | 1 | 2 | 3 | 4 |
| 年人均收入 y | 0.6 | 0.8 | 1.1 | 1.5 |

根据上表可得回归方程$\hat{y}=\hat{b}x+\hat{a}$中的$\hat{b}$为 0.3，据此模型预报该地区贫困户 2021 年的年人均收入为_______。（单位：万元）.

**17.（青岛一模 15）**某驾驶员培训学校为对比了解“科目二”的培训过程采用大密度集中培训与周末分散培训两种方式的效果，调查了 105 名学员，统计结界为:接受大密度集中培训的 55 个学员中有 45 名学员一次考试通过，接受周末分散培训的学员一次考试通过的有 30 个.根据统计结果，认为“能否一次考试通过与是否集中培训有关”犯错误的概率不超过____________

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

| P（$k^2\geq k$） | 0.05 | 0.025 | 0.010 | 0.001 |
|---|---|---|---|---|
| k | 3.841 | 5.024 | 6.635 | 10.828 |

**四、解答**

**18.（潍坊一模 20）**在对人体的脂肪含量和年龄之间的关系的研究中，科研人员获得了一些年龄和脂肪含量的简单随机样本数据$(x_i, y_i)(i=1, 2, \cdots, 20, 25<x_i<65)$，其中 $x_i$ 表示年龄，$y_i$ 表示脂肪含量，并计算

得到 $\sum\limits_{i=1}^{20} x_i^2 = 48280$，$\sum\limits_{i=1}^{20} y_i^2 = 15480$，$\sum\limits_{i=1}^{20} x_i y_i = 27220$，$\overline{x} = 48$，$\overline{y} = 27$，$\sqrt{22} \approx 4.7$.

（1）请用相关系数说明该组数据中 $y$ 与 $x$ 之间的关系可用线性回归模型进行拟合，并求 $y$ 关于 $x$ 的线性回归方程 $y = a + \hat{b}x$（$a$，$\hat{b}$ 的计算结果保留两位小数）；

（2）科学健身能降低人体脂肪含量，下表是甲，乙两款健身器材的使用年限（整年）统计表：

| 使用年限 台数 款式 | 5 年 | 6 年 | 7 年 | 8 年 | 合计 |
| --- | --- | --- | --- | --- | --- |
| 甲款 | 5 | 20 | 15 | 10 | 50 |
| 乙款 | 15 | 20 | 10 | 5 | 50 |

某健身机构准备购进其中一款健身器材，以使用年限的频率估计概率，请根据以上数据估计，该机构选择购买哪一款健身器材，才能使用更长久？

参考公式：相关系数

$$r = \frac{\sum\limits_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i - \overline{y})^2}} = \frac{\sum\limits_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sqrt{\sum\limits_{i=1}^{n} x_i^2 - n\overline{x}^2}\sqrt{\sum\limits_{i=1}^{n} y_i^2 - n\overline{y}^2}};$$

对于一组具有线性相关关系的数据$(x_i，y_i)(i=1，2，\cdots，n)$，其回归直线 $y = \hat{b}x + a$ 的斜率和截距的最小二乘估计分别为：$\hat{b} = \dfrac{\sum\limits_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}$，$a = \overline{y} - \hat{b}\overline{x}$.

**19.（德州一模 19）**2021 年春晚首次采用“云”传播，“云”互动形式，实现隔空连线心意相通，全球华人心连心“云团圆”，共享新春氛围，“云课堂”亦是一种真正完全突破时空限制的全方位互动性学习模式，某市随机抽取 200 人对“云课堂”倡议的了解情况进行了问卷调查，记 $Y$ 表示了解，$N$ 表示不了解，统计结果如表所示：

（表一）

| 了解情况 | $Y$ | $N$ |
|---|---|---|
| 人数 | 140 | 60 |

（表二）

| | 男 | 女 | 合计 |
|---|---|---|---|
| $Y$ | 80 | | |
| $N$ | | 40 | |
| 合计 | | | |

（1）请根据所提供的数据，完成上面的 2×2 列联表（表二），并判断是否有 99%的把握认为对“云课堂”倡议的了解情况与性别有关系；

（2）用样本估计总体，将频率视为概率，在男性市民和女性市民中各随机抽取 4 人，记“4 名男性中恰有 3 人了解云课堂倡议”的概率为 $P_1$，“4 名女性中恰有 3 人了解云课堂倡议”的概率为 $P_2$，试求出 $P_1$ 与 $P_2$，并比较 $P_1$ 与 $P_2$ 的大小．

附：临界值参考表的参考公式

| $P（K^2\geqslant k_0）$ | 0.10 | 0.05 | 0.025 | 0.010 | 0.005 | 0.001 |
|---|---|---|---|---|---|---|
| $k_0$ | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 | 10.828 |

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

**20.（2021•临沂一模 19）**党中央、国务院高度重视新冠病毒核酸检测工作，中央应对新型冠状病毒感染肺炎疫情工作领导小组会议作出部署，要求尽力扩大核酸检测范围，着力提升检测能力，根据统计发现，疑似病例核酸检测呈阳性的概率为 $p(0<p<1)$。现有 4 例疑似病例，分别对其取样检测，既可以逐个化验，也可以将若干个样本混合在一起化验，混合样本中只要有病毒，则化验结果呈阳性，若混合样本呈阳性，则需将该组中备用的样本再逐个化验；若混合样本呈阴性，则判定该组各个样本均为阴性，无需再化验现有以下三种方案：

方案一：4 个样本逐个化验

方案二：4 个样本混合在一起化验；

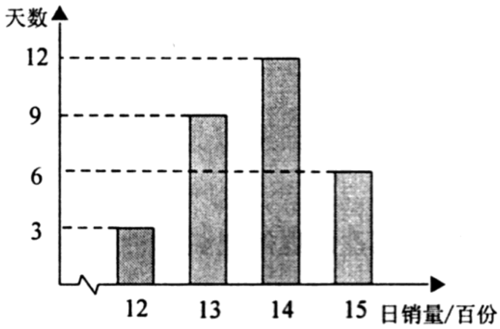
方案三：4 个样本均分为两组，分别混合在一起化验.

在新冠肺炎爆发初期，由于检测能力不足，化验次数的期望值越小，则方案越“优”

(1)若 $p=\frac{1}{3}$,按方案一，求 4 例疑似病例中恰有 2 例呈阳性的概率；

(2)若 $p=\frac{1}{10}$，现将该 4 例疑似病例样本进行化验，试比较以上三个方案中哪个最“优”，并说明理由.

**21.（烟台一模 20）**某品牌餐饮企业为满足人们餐饮需求、丰富产品花色、提高企业竞争力，研发了一款新产品，该产品每份成本 60 元，售价 80 元，产品保质期为两天，若两天内未售出，则产品过期报废.由于烹制工艺复杂，该产品在最初推广阶段由企业每两天统一生产、集中配送一次，该企业为决策每两天的产量，选取旗下的直营连锁店进行试销，统计并整理连续 30 天的日销量（单位：百份），假定该款新产品每日销量互相独立，得到右侧柱状图：




（1）记两天中销售该新产品的总份数为 ξ（单位：百份），求 ξ 的分布列和数学期望；

（2）以该新产品两天内获得利润较大为决策依据，在每两天生产配送 27 百份、28 百份两种方案中应选择哪种？

**22.（聊城一模 20）**为了对学生进行劳动技术教育，培养正确的劳动观点和态度，养成自立、自强、艰苦奋斗的思想作风，加强理论联系实际，使学生掌握一定的生产知识和劳动技能，某学校投资兴建了甲、乙两个加工厂，生产同一型号的小型电器，产品按质量分为 A，B，C 三个等级，其中 A，B 等级的产品为合格品，C 等级的产品为次品．质监部门随机抽取了两个工厂的产品各 100 件，检测结果为：甲厂合格品 75 件，甲、乙两厂次品共 60 件．

(1)根据所提供的数据完成下面的 2×2 列联表，并判断是否有 95%的把握认为产品的合格率与生产厂家有关?

| | 合格品 | 次品 | 合计 |
|---|---|---|---|
| 甲厂 | | | |
| 乙厂 | | | |
| 合计 | | | 200 |

(2)每件产品的生产成本为 30 元，每件 A，B 等级的产品出厂销售价格分别为 60 元，40 元，C 等级的产品必须销毁，且销毁费用为每件 4 元．若甲、乙两厂抽到的产品中各有 10 件为 A 级产品，用样本的频率代替概率，分别说明甲，乙两厂是否盈利．

附：$K^2=\dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n=a+b+c+d$．

| $P(K^2\geqslant k_0)$ | 0.100 | 0.050 | 0.010 | 0.005 |
|---|---|---|---|---|
| $k_0$ | 2.706 | 3.841 | 6.635 | 7.879 |

**23.（2021•淄博一模 20）**某市会展公司计划在未来周组织 5 天广场会展. 若会展期间有风雨天气，则暂停该天会展，根据该市气象台预报得知，未来一周从周一到周五的 5 天时间内出现风雨天气情况的概率是:前 3 天均为 $\dfrac{1}{2}$，后 2 天均为 $\dfrac{4}{5}$ (假设每一天出现风雨天气与否是相互独立的).

(1)求未来一周从周一到周五 5 天中至少有一天暂停会展的概率;

(2)求这次会展活动展出的平均天数. (结果精确到 0.1)

**24.（青岛一模 20）**某商场每年都会定期答谢会员，允许年度积分超过指定积分的会员参加特价购物赠券活动，今年活动的主题为“购物三选一，正清暖心里”，符合条件的会员可以特价购买礼包 A（十斤肉类），礼包 B（十斤蔬菜）和礼包 C（十斤鸡蛋）三类特价商品中的任意一类，并且根据购买的礼包不同可以获取价值不等的代金券，根据以往经验得知，会员购买礼包 A 和礼包 B 的概率均为$\frac{2}{5}$。

（1）预计今年有 400 名符合条件的会员参加活动，求商场为此活动需要准备多少斤鸡蛋合理；

（2）在促销活动中，若有甲、乙、丙三位会员同时参加答谢活动，各人购买礼包互相独立，已知购买礼包 A 或购买礼包 B 均可以获得 50 元商场代金券，购买礼包 C 可以获得 25 元商场代金券，设 Y 是三人获得代金券金额之和，求 Y 的分布列和数学期望。

**25.（济宁一模 19）**垃圾分类收集处理是一项利国利民的社会工程和环保工程．搞好垃圾分类收集处理，可为政府节省开支，为国家节约能源，减少环境污染，是建设资源节约型社会的一个重要内容．为推进垃圾分类收集处理工作，$A$ 市通过多种渠道对市民进行垃圾分类收集处理方法的宣传教育，为了解市民能否正确进行垃圾分类处理，调查机构借助网络进行了问卷调查，并从参与调查的网友中抽取了 200 人进行抽样分析，得到如下列联表（单位人）：

| | 能正确进行垃圾分类 | 不能正确进行垃圾分类 | 总计 |
|---|---|---|---|
| 55 岁及以下 | 90 | 30 | 120 |
| 55 岁以上 | 50 | 30 | 80 |
| 总计 | 140 | 60 | 200 |

（1）根据以上数据，判断是否有 90%的把握认为 $A$ 市能否正确进行垃圾分类处理与年龄有关？

（2）将频率视为概率，现从 $A$ 市 55 岁及以下的市民中里随机抽样的方法每次抽取 1 人，共抽取 3 次．记被抽取的 3 人中“不能正确进行垃圾分米”的人数为 $X$，若每次抽取的结果是相互独立的，求随机变量 $X$ 的分布列和均值 $E$（$X$）．

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

| $P$（$K^2\geqslant k_0$） | 0.15 | 0.10 | 0.05 | 0.025 |
|---|---|---|---|---|
| $k_0$ | 2.072 | 2.706 | 3.841 | 5.024 |

**26.（滨州一模 20）**国家发展改革委、住房城乡建设部于 2017 年发布了《生活垃圾分类制度实施方案》，规定 46 个城市在 2020 年底实施生活垃圾强制分类，垃圾回收、利用率要达 35%以上．截至 2019 年底，这 46 个重点城市生活垃圾分类的居民小区覆盖率已经接近 70%．武汉市在实施垃圾分类之前，从本市人口数量在两万人左右的 320 个社区中随机抽取 50 个社区，对这 50 个社区某天产生的垃圾量（单位：吨）进行了调查，得到如下频数分布表，并将人口数量在两万人左右的社区垃圾数量超过 28 吨/天的确定为“超标”社区：

| 垃圾量 $X$ | [12.5，15.5） | [15.5，18.5） | [18.5，21.5） | [21.5，24.5） | [24.5，27.5） | [27.5，30.5） | [30.5，33.5] |
|---|---|---|---|---|---|---|---|
| 频数 | 5 | 6 | 9 | 12 | 8 | 6 | 4 |

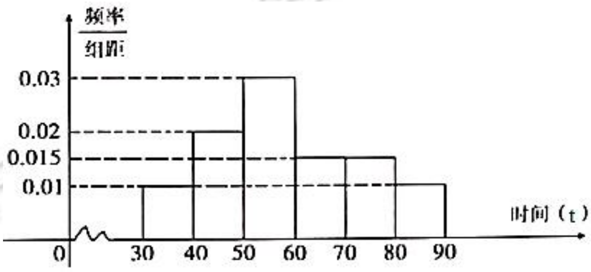
（1）通过频数分布表估算出这 50 个社区这一天垃圾量的平均值 $\overline{x}$（精确到 0.1）；

（2）若该市人口数量在两万人左右的社区这一天的垃圾量大致服从正态分布 $N(\mu，\sigma^2)$，其中 $\mu$ 近似为（1）中的样本平均值 $\overline{x}$，$\sigma^2$ 近似为样本方差 $s^2$，经计算得 $s=5.2$．请利用正态分布知识估计这 320 个社区中“超标”社区的个数．

（3）通过研究样本原始数据发现，抽取的 50 个社区中这一天共有 8 个“超标”社区，市政府决定对这 8 个“超标”社区的垃圾来源进行跟踪调查．现计划在这 8 个“超标”社区中任取 5 个先进行跟踪调查，设 $Y$ 为抽到的这一天的垃圾量至少为 30.5 吨的社区个数，求 $Y$ 的分布列与数学期望．

（参考数据：$P(\mu-\sigma<X\leqslant\mu+\sigma)\approx0.6827$；$P(\mu-2\sigma<X\leqslant\mu+2\sigma)\approx0.9545$；$P(\mu-3\sigma<X\leqslant\mu+3\sigma)\approx0.9974$）

**27.（泰安一模 20）**某市为了了解本市初中生周末运动时间，随机调查了 3000 名学生，统计了他们的周末运动时间，制成如图所示的频率分布直方图.




（1）按照分层抽样，从[40，50）和[80，90）中随机抽取了 9 名学生．现从已抽取的 9 名学生中随机推荐 3 名学生参加体能测试．记推荐的 3 名学生来自[40，50）的人数为 $X$，求 $X$ 的分布列和数学期望；

（2）由频率分布直方图可认为：周末运动时间 $t$ 服从正态分布 $N$（μ，σ$^2$），其中，μ 为周末运动时间的平均数 $\overline{t}$，σ 近似为样本的标准差 $s$，并已求得 $s\approx 14.6$．可以用该样本的频率估计总体的概率，现从本市所有初中生中随机抽取 12 名学生，记周末运动时间在（43.9，87.7]之外的人数为 $Y$，求 $P$（$Y=3$）（精确到 0.001）．

参考数据 1：当 $t\sim N$（（μ，σ$^2$）时，$P$（μ - σ＜$t$≤μ+σ）＝0.6826，$P$（μ - 2σ＜$t$≤μ+2σ）＝0.9544，$P$（μ - 3σ＜$t$≤μ+3σ）＝0.9974；

参考数据 2：$0.8185^9=0.1649$；$0.1815^3=0.0060$.

28.（日照一模 20）为加强进口冷链食品监管，某省于 2020 年底在全省建立进口冷链食品集中监管专仓制度，在口岸、目的地市或县（区、市）等进口冷链食品第一入境点，设立进口冷链食品集中监管专仓，集中开展核酸检测和预防性全面消毒工作，为了进一步确定某批进口冷冻食品是否感染病毒，在入关检疫时需要对其采样进行化验，若结果呈阳性，则有该病毒；若结果呈阴性，则没有该病毒，对于 $n,(n\in N^*)$ 份样本，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 k 份样本分别取样混合在一起，若检验结果为阴性，那么这 k 份全为阴性，因而检验一次就够了；如果检验结果为阳性，为了明确这 k 份究竟哪些为阳性，就需要对它们再次取样逐份检验，则 k 份检验的次数共为 k+1 次，若每份样本没有该病毒的概率为 $\sqrt{p}(0<p<1)$，而且样本之间是否有该病毒是相互独立的

(1)若 $p=\dfrac{2}{3}$，求 2 份样本混合的结果为阳性的概率

(2)若取得 4 份样本，考虑以下两种检验方案

方案一：采用混合检验；

方案二：平均分成两组，每组 2 份样本采用混合检验.

若检验次数的期望值越小，则方案越“优”，试问方案一、二哪个更“优”？请说明理由.

**29.（菏泽一模 20）**随着生活质量的提升，家庭轿车保有量逐年递增，方便之余却加剧了交通拥堵和环保问题，绿色出行引领时尚，共享单车进驻城市．菏泽市有统计数据显示，2020 年该市共享单车用户年龄等级分布如图 1 所示，一周内市民使用单车的频率分布扇形图如图 2 所示．若将共享单车用户按照年龄分为“年轻人”（20 岁～39 岁）和“非年轻人”（19 岁及以下或者 40 岁及以上）两类，将一周内使用的次数为 6 次或 6 次以上的经常使用共享单车的称为“单车族”，使用次数为 5 次或不足 5 次的称为“非单车族”．

已知在“单车族”中有$\frac{5}{6}$是“年轻人”．

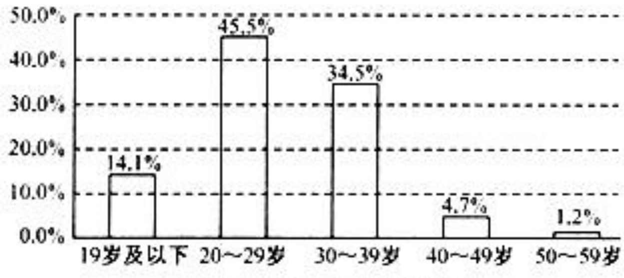


图1 共享单车用户年龄等级分布

图2 共享单车使用频率分布

（1）现对该市市民进行“经常使用共享单车与年龄关系”的调查，采用随机抽样的方法，抽取一个容量为 400 的样本，请你根据图表中的数据，补全下列 2×2 列联表，并判断是否有 95%的把握认为经常使用共享单车与年龄有关？

使用共享单车情况与能力列联表

| | 年轻人 | 非年轻人 | 合计 |
|---|---|---|---|
| 单车族 | | | |
| 非单车族 | | | |
| 合计 | | | |

（2）若将（1）中的频率视为概率，从该市市民中随机任取 3 人，设其中既是“单车族”又是“非年轻人”的人数为随机变量 $X$，求 $X$ 的分布列与期望．

参考数据：独立性检验界值表

| $P(K^2\geqslant k_0)$ | 0.15 | 0.10 | 0.05 | 0.0025 | 0.01 |
|---|---|---|---|---|---|
| $k_0$ | 2.072 | 2.706 | 3.841 | 5.024 | 6.635 |

其中，$n=a+b+c+d$，$K^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$（注：保留三位小数）．

**30.（济南一模 21）**某机构为研究考生物理成绩与数学成绩之间的关系，从一次考试中随机抽取 11 名考生的数据，统计如下表：

| 数学成绩 $x$ | 46 | 65 | 79 | 89 | 99 | 109 | 110 | 116 | 123 | 134 | 140 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 物理成绩 $y$ | 50 | 54 | 60 | 63 | 66 | 68 | 0 | 70 | 73 | 76 | 80 |

（1)由表中数据可知，有一位考生因物理缺考导致数据出现异常，剔除该组数据后发现，考生物理成绩 y 与数学成绩 x 之间具有线性相关关系，请根据这 10 组数据建立 y 关于 x 的回归直线方程，并估计缺考考生如果参加物理考试可能取得的成绩；

（2)已知参加该次考试的 10000 名考生的物理成绩服从正态分布 $N(\mu,\sigma^2)$,用剔除异常数据后的样本平均值作为 μ 的估计值，用剔除异常数据后的样本标准差作为 σ 的估计值，估计物理成绩不低于 75 分的人数 Y 的期望.

附：参考数据：

| $\sum_{i=1}^{11}x_i$ | $\sum_{i=1}^{11}y_i$ | $\sum_{i=1}^{11}x_iy_i$ | $\sum_{i=1}^{11}x_i^2$ | $\sum_{i=1}^{11}(y_i-\bar{y})^2$ | $\frac{2\ 586}{8\ 326}$ |
|---|---|---|---|---|---|
| 1 110 | 660 | 68 586 | 120 426 | 4 770 | 0.31 |

上表中的 x;表示样本中第 i 名考生的数学成绩，y;表示样本中第 i 名考生的物理成绩，$\bar{y}=\frac{1}{11}\sum_{i=1}^{11}y_i$.

参考公式：①对于一组数据：$u_1,u_2,\cdots,u_n$,其方差：$s^2=\frac{1}{n}\sum_{i=1}^{n}(u_i-\bar{u})^2=\frac{1}{n}\sum_{i=1}^{n}u_i^2-\bar{u}^2$.

②对于一组数据（$u_1,v_1),(u_2,v_2),\cdots,(u_n,v_n)$,其回归直线$\hat{v}=\hat{a}+\hat{b}u$ 的斜率和截距的最小二乘估计分别为：$\hat{b}=\frac{\sum_{i=1}^{n}u_iv_i-n\bar{u}\bar{v}}{\sum_{i=1}^{n}u_i^2-n\bar{u}^2}$，$\hat{a}=\bar{v}-\hat{b}\bar{u}$.

③若随机变量 ε 服从 $N(\mu,\sigma^2)$,则 $P(\mu-\sigma<\xi<\mu+\sigma)\approx0.683$,

$P(\mu-2\sigma<\xi<\mu+2\sigma)\approx0.955$,　$P(\mu-3\sigma<\xi<\mu+3\sigma)\approx0.997$.

# 专题十一　概率统计

## 一、单项选择

**1.**【答案】C

【解析】第 8 天比第 7 天的复工指数和复产指数均低，$A$ 错；

这 11 天期间，复产指数的极差小于复工指数的极差：两者最高差不多，但最低的复工指数比复产指数低得多，$B$ 错；

第 3 天至第 11 天复工复产指数均超过 80%，$C$ 正确；

第 9 天至第 11 天复工指数的增量小于复产指数的增量，$D$ 错误．

故选：$C$．

**2.**【答案】C

【解析】每匹马只能用一次，每场比赛双方各出一匹马，共比赛三场．每场比赛中胜者得 1 分，否则得 0 分．

设田忌的上等马、中等马、下等马分别为 $A$，$B$，$C$，

齐王的上等马、中等马、下等马分别为 $a$，$b$，$c$，

所有的基本事件有 6 种，分别为：

（$Aa$，$Bb$，$Cc$），（$Aa$，$Bc$，$Cb$），（$Ab$，$Ba$，$Cb$），（$Ab$，$Bc$，$Cb$），（$Ac$，$Bb$，$ca$），（$Ac$，$Ba$，$Cb$），

比赛结束时，田忌得 2 分的基本事件为：（$Ab$，$Bc$，$Ca$），只有 1 种，

∴比赛结束时，田忌得 2 分的概率 $P=\frac{1}{6}$．

故选：$C$．

3．【答案】B

【分析】由条形统计图得共抽到 50 名同学演讲，由扇形统计图片得抽到的学生中演讲同学占 10%，从而求出一共抽取的学生数为 500 人，再求出抽到的学生中合唱学生占 40%，由此能求出选取的学生中参加机器人社团的学生数．

【解析】由条形统计图得抽到 50 名同学演讲，

由扇形统计图片得抽到的学生中演讲同学占 10%，

∴一共抽取的学生数为：$n=\frac{50}{10\%}=500$（人），

∴抽到的学生中合唱学生占：$\frac{200}{500}\times 100\%=40\%$，

∴选取的学生中参加机器人社团的学生数为：

500（1﹣40%﹣10%﹣15%﹣20%）＝75（人）.

故选：$B$.

**4.**【答案】A

**5.**【答案】D

【解析】因为在频率分布直方图中，所有小矩形面积之和为 1，所以在[60,65)区间的频率为 0.2，则 $a=0.04$，采用分层抽样，从[55,60),[60,65),[65,70)这三个区间中，随机抽取 6 人，则在[55,60)中抽取 3 人，在[60,65)中抽取 2 人,在[65,70)中抽取 1 人。从 6 人中再抽取 3 人，恰有 2 人位于[60,65)的概率为

$$P(X=2)=\frac{C_2^2C_4^1}{C_6^3}=\frac{1}{5}$$

**6.**【答案】B

【解析】仅能从 E 或 B 点出发不重复地走遍全城，所以 $P=\frac{2}{C_6^1}=\frac{1}{3}$，故选 B。

**7.**【答案】B

【解析】菏泽万达商场在春节前开展商品促销活动，顾客凡购物金额满 50 元，则可以从“福”字、春联和灯笼这三类礼品中任意免费领取一件，

若有 4 名顾客都领取一件礼品，则基本事件总数 $n=3^4=81$，

其中他们中有且仅有 2 人领取的礼品种类相同包含的基本事件个数：

$m=C_4^2\cdot A_3^3=36$，

则他们中有且仅有 2 人领取的礼品种类相同的概率是 $P=\frac{m}{n}=\frac{36}{81}=\frac{4}{9}$.

故选：$B$.

**二、多项选择**

8.【答案】AB

【解析】这 10 天中第一天，第三天和第四天共 3 天空气质量为一级，所以 A 正确；
从图可知从 6 日到 9 日 PM2.5 日均值逐渐降低，所以 B 正确；
由图可知，这 10 天中 PM2.5 日均值的中位数是 $\frac{41+45}{2}=43$，所以 C 不正确；
这 10 天中 PM2.5 日均值的平均数为 $50+\frac{1}{10}(-20-9-18-16-10+30+28+10-5-2)=48.8$ 所以 D 错，故选 AB。

9.【答案】BD

【分析】选项 $A$ 只需计算每个月异地快递量是否是同城快递量的 6 倍以上，即可判定；选项 $B$ 只需计算 10 月份异地快递增长率与 9 月份的异地快递增长率进行比较即可；选项 $C$ 与 $D$ 只需观察图表，根据图

中信息即可判定．

【解答】解：$A$ 选项：2020 年下半年，每个月的异地快递量都是同城快递量的 6 倍以上，

而 7 月份异地快递量都是同城快递量的$\frac{572812.9}{105191.1}\approx 5.5<6$，故选项 $A$ 不正确；

选项 $B$：由图可知 10 月份异地快递增长率为$\frac{708642.6-679556.6}{679556.6}\approx 0.043=4.3\%$，

9 月份份异地快递增长率为$\frac{679556.6-599604.6}{599604.6}\approx 0.133=13.3\%$，$13.3\%>4.3\%$，

2020 年 10 月份异地快递增长率小于 9 月份的异地快递增长率，故选项 $B$ 正确；

选项 $C$：由图可知随着月份的增长，异地快递量在 7 月到 11 月都是增长，而 12 月开始下降，不是正相关，

故选项 $C$ 不正确；

选项 $D$：由图可知，2020 年下半年，同城和异地快递量最高均出现在 11 月，故选项 $D$ 正确．

故选：$BD$．

**10.**【答案】ACD

【解析】对于 $A$，甲的极差为 $42-36=6$，乙的极差为 $41-34=7$，

所以“甲”的极差小于“乙”的极差，$A$ 正确；

对于 $B$，甲的平均数是$\frac{1}{6}\times(36+37+37+38+40+42)=\frac{230}{6}$，

乙的平均数为$\frac{1}{6}\times(34+36+38+39+40+41)=\frac{228}{6}$，

所以“甲”的平均值大于“乙”的平均值，$B$ 错误；

对于 $C$，甲的中位数是$\frac{1}{2}\times(37+38)=37.5$，

乙的中位数是$\frac{1}{2}\times(38+39)=38.5$，

所以，“甲”的中位数小于“乙”的中位数，$C$ 正确；

对于 $D$，过去 6 年甲的平均增长率为：$\sqrt[6]{\frac{42}{36}}-1$；

乙的平均增长率为：$\sqrt[6]{\frac{41}{34}}-1$，且$\frac{42}{36}<\frac{41}{34}$，

所以“甲”的平均增长率小于“乙”的平均增长率，$D$ 正确．

故选：$ACD$．

**11.**【答案】ACD

**12.**【答案】ACD

【解析】A 选项，总的基本事件为 36，点数之和大于 6 的总共有 21 种，$\therefore p=\frac{21}{36}=\frac{7}{12}$，A 正确；

B 选项，挑战第一关并过关的概率$p_1=\frac{1}{2}$，前两关过关的概率$p=\frac{1}{2}\times\frac{7}{12}=\frac{7}{24}$，$\therefore$B 错误

C 选项，事件 B 包含基本事件：$C_3^1\times C_5^1\times C_5^1+C_3^2\times C_5^1+1=91$ 个，

事件 A 包含的基本事件：$A_3^3+1=7$，$P(A|B)=\frac{7}{91}=\frac{1}{11}$，C 正确

D 选项，直接挑战第四关，总的基本事件数：$6^4=1296$，点数之和超过 20 的有以下几类

（5,5,5,6）4 个，（5,5,6,6）6 个，（5,6,6,6）4 个，（6,6,6,6）1 个，（4,6,6,6）4 个，（3,6,6,6）4 个，（4,5,6,6）12 个，总计 53 个，$P=\frac{35}{1296}$，D 正确。

**三、填空**

**13.**【答案】7

【解析】$\overline{x}=\frac{1+2+3+4+5}{5}=3$，

$\overline{y}=\frac{5+6+a+8+10}{5}=\frac{29+a}{5}$，

线性回归方程$\hat{y}=1.2\hat{x}+3.6$经过样本中心，所以$\frac{29+a}{5}=1.2\times3+3.6$，

解得 $a=7$.

故答案为：7.

**14.**【答案】65.5 万元

【解析】$\because\overline{x}=\frac{4+2+3+5}{4}=3.5$，

$\overline{y}=\frac{49+26+39+54}{4}=42$，

∵数据的样本中心点在线性回归直线上，

回归方程$\hat{y}=\hat{b}\hat{x}+\hat{a}$中的$\hat{b}$为 9.4，

$\therefore 42=9.4\times3.5+a$，

$\therefore\hat{a}=9.1$，

∴线性回归方程是 $y=9.4x+9.1$，

∴广告费用为 6 万元时销售额为 $9.4\times6+9.1=65.5$，

故答案为：65.5 万元.

15.【答案】5

【分析】先求得样本中心点为（$\overline{x}$，$\overline{y}$），再把样本中心点和（30，68）均代入线性回归方程，解方程组即可.

【解答】$\overline{x}=\frac{1}{7}\times$（21+22+23+24+25+26+27）＝24，$\overline{y}=\frac{1}{7}\times$（24+28+31+39+43+47+54）＝38，

∴样本中心点为（24，38），

∴$38=\hat{b}\times24+\hat{a}$①，

∵当该地的气温为 30℃时，蟋蟀每分钟鸣叫次数的预报值为 68，

∴$68=\hat{b}\times30+\hat{a}$②，

由①②解得，$\hat{b}=5$．

故答案为：5．

**16.**【答案】1.75

【解析】由题$\overline{x}$=2.5，$\overline{y}=1$，将（2.5,1）带入 y=0.3x+a 中，得$a=0.25$，

2021 年对应的年份编号 x 为 5，将 x=5 代入 y=0.3x+0.25 中，得 y=1.75

**17.**【答案】0.025

【解析】由题意，可作出 2×2 列联表如下：

| 项目 | 大密度集中培训 | 周末分散培训 | 总计 |
|---|---|---|---|
| 一次通过考试 | 45 | 30 | 75 |
| 未一次通过考试 | 10 | 20 | 30 |
| 总计 | 55 | 50 | 105 |

故$k^2=\frac{105(45\times20-30\times10)^2}{55\times50\times75\times30}\approx6.109>5.024$，因此“能否一次考试通过与是否集中培训有关”犯错误的概率不超过 0.025.

**四、解答**

**18.**【解析】（1）$\overline{x}^2=2304,\overline{y}^2=729,\sum_{i=1}^{20}x_iy_i-20\overline{x}\,\overline{y}=1300,\sum_{i=1}^{20}x_i^2-20\overline{x}^2=2200,$

$\sum_{i=1}^{n}y_i^2-20\overline{y}^2=900,$

$$r=\frac{\sum_{i=1}^{20}x_iy_i-20\overline{x}\,\overline{y}}{\sqrt{\sum_{i=1}^{20}x_i^2-20\overline{x}^2}\sqrt{\sum_{i=1}^{n}y_i^2-20\overline{y}^2}}\approx0.92,$$

因为 $y$ 与 $x$ 的相关系数接近 1，所以 $y$ 与 $x$ 之间具有较强的线性相关关系，可用线性回归模型进行拟合；

由题意可得，$\hat{b}=\dfrac{\sum\limits_{i=1}^{20}(x_i-\bar{x})(y_i-\bar{y})}{\sum\limits_{i=1}^{20}(x_i-\bar{x})^2}=\dfrac{\sum\limits_{i=1}^{20}x_iy_i-20\bar{x}\bar{y}}{\sum\limits_{i=1}^{20}x_i^2-20\bar{x}^2}=\dfrac{13}{22}\approx 0.591$，

$\hat{a}=\bar{y}-\hat{b}\bar{x}=27-0.591\times 48\approx -1.37$，

所以 $\hat{y}=0.59x-1.37$.

（2）以频率估计概率，设甲款健身器使用年限为 X（单位：年）

| $X$ | 5 | 6 | 7 | 8 |
|---|---|---|---|---|
| $p$ | 0.1 | 0.4 | 0.3 | 0.2 |

$E(X)=5\times 0.1+6\times 0.4+7\times 0.3+8\times 0.2=6.6$，

设乙款健身器使用年限为 Y（单位：年）

| $Y$ | 5 | 6 | 7 | 8 |
|---|---|---|---|---|
| $p$ | 0.3 | 0.4 | 0.2 | 0.1 |

$E(Y)=5\times 0.3+6\times 0.4+7\times 0.2+8\times 0.1=6.1$，

因为 $E(X)>E(Y)$，所以该机构购买甲款健身器材更划算.

**19.【解析】**（1）根据题意填写列联表，如下：

| | 男 | 女 | 合计 |
|---|---|---|---|
| $Y$ | 80 | 60 | 140 |
| $N$ | 20 | 40 | 60 |
| 合计 | 100 | 100 | 200 |

根据表中数据，计算 $K^2=\dfrac{200\times(80\times 40-20\times 60)^2}{100\times 100\times 140\times 60}=\dfrac{200}{21}\approx 9.524>6.635$，

对照临界值表知，有 99%的把握认为对“云课堂”倡议的了解情况与性别有关系；

（2）用样本估计总体，将频率视为概率，

根据列联表得出男性了解“云课堂”倡议的概率为 $\dfrac{80}{100}=\dfrac{4}{5}$，

女性了解“云课堂”倡议的概率为 $\dfrac{60}{100}=\dfrac{3}{5}$，

所以计算概率 $P_1=C_4^3\cdot\left(\dfrac{4}{5}\right)^3\cdot\dfrac{1}{5}=\dfrac{256}{625}$，

概率 $P_2=C_4^3\cdot\left(\dfrac{3}{5}\right)^3\cdot\dfrac{2}{5}=\dfrac{216}{625}$，

所以 $P_1>P_2$.

**20.**

19.解：(1)用 $X$ 表示 4 例疑似病例中化验呈阳性的人数，则随机变量 $X\sim B(4,\frac{1}{3})$ ……………… 2 分

由题意可知：$P(X=2)=C_4^2\cdot(\frac{1}{3})^2\cdot(1-\frac{1}{3})^2=\frac{8}{27}$ ……………… 4 分

(2)方案一：若逐个检验，则检验次数为 4. ……………… 5 分

方案二：混合一起检验，记检验次数为 $X$，则 $X=1,5$.

$P(X=1)=(1-\frac{1}{10})^4=\frac{81\times81}{10000}=\frac{6561}{10000}$,

$P(X=5)=1-P(X=1)=\frac{3439}{10000}$

$E(X)=1\times\frac{6561}{10000}+5\times\frac{3439}{10000}=\frac{23756}{10000}$ ……………… 7 分

方案三：每组的两个样本混合在一起化验，若结果呈阴性，则检测次数为 1，

其概率为 $(1-\frac{1}{10})^2=\frac{81}{100}$; ……………… 8 分

若结果呈阳性，则检测次数为 3，其概率为 $1-\frac{81}{100}=\frac{19}{100}$ ……………… 9 分

设方案三检测次数为随机变量 $Y$，则 $Y=2,4,6$.

$P(Y=2)=\frac{81}{100}\times\frac{81}{100}=\frac{81\times81}{10000}=\frac{6561}{10000}$,

$P(Y=4)=\frac{81}{100}\times\frac{19}{100}\times2=\frac{2\times81\times19}{10000}=\frac{3078}{10000}$,

$P(Y=6)=\frac{19}{100}\times\frac{19}{100}=\frac{361}{10000}$, ……………… 10 分

则 $E(Y)=2\times\frac{81\times81}{10000}+4\times\frac{2\times81\times19}{10000}+6\times\frac{19\times19}{10000}=\frac{27600}{10000}$ ……………… 11 分

由 $E(X)<E(Y)<4$,

知方案二最优 ……………… 12 分

**21.**

解：（1）根据题意可得，$\xi$ 的所有可能取值为 $24,25,26,27,28,29,30$.

$P(\xi=24)=\frac{1}{10}\times\frac{1}{10}=\frac{1}{100}$, $P(\xi=25)=\frac{1}{10}\times\frac{3}{10}\times2=\frac{3}{50}$,

$P(\xi=26)=\frac{1}{10}\times\frac{2}{5}\times2+\frac{3}{10}\times\frac{3}{10}=\frac{17}{100}$, $P(\xi=27)=\frac{1}{10}\times\frac{1}{5}\times2+\frac{3}{10}\times\frac{2}{5}\times2=\frac{7}{25}$,

$P(\xi=28)=\frac{3}{10}\times\frac{1}{5}\times2+\frac{2}{5}\times\frac{2}{5}=\frac{7}{25}$，　　$P(\xi=29)=\frac{2}{5}\times\frac{1}{5}\times2=\frac{4}{25}$，

$P(\xi=30)=\frac{1}{5}\times\frac{1}{5}=\frac{1}{25}$，　　……………………………………………5 分

$\xi$ 的分布列如下：

| $\xi$ | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|
| $P$ | $\frac{1}{100}$ | $\frac{3}{50}$ | $\frac{17}{100}$ | $\frac{7}{25}$ | $\frac{7}{25}$ | $\frac{4}{25}$ | $\frac{1}{25}$ |

$E(\xi)=24\times\frac{1}{100}+25\times\frac{3}{50}+26\times\frac{17}{100}+27\times\frac{7}{25}+28\times\frac{7}{25}+29\times\frac{4}{25}+30\times\frac{1}{25}=27.4$.

……………………………………7 分

（2）当每两天生产配送 27 百份时，利润为

$(24\times20-3\times60)\times\frac{1}{100}+(25\times20-2\times60)\times\frac{3}{50}+(26\times20-1\times60)\times\frac{17}{100}$

$+27\times20\times(1-\frac{1}{100}-\frac{3}{50}-\frac{17}{100})=514.4$ 百元.　　…………………………9 分

当每两天生产配送 28 百份时，利润为

$(24\times20-4\times60)\times\frac{1}{100}+(25\times20-3\times60)\times\frac{3}{50}+(26\times20-2\times60)\times\frac{17}{100}$

$+(27\times20-1\times60)\times\frac{7}{25}+28\times20\times\frac{12}{25}=492.8$ 百元.　　…………………………11 分

由于 $514.4>492.8$，所以选择每两天生产配送 27 百份.　　………………………12 分

**22.**【解析】(1)2×2 列联表如下

| | 合格品 | 次品 | 合计 |
|---|---|---|---|
| 甲厂 | 75 | 25 | 100 |
| 乙厂 | 65 | 35 | 100 |
| 合计 | 140 | 60 | 200 |

……………………………………………………………………………………………2 分

因为 $K^2=\frac{200\times(75\times35-65\times25)^2}{100\times100\times140\times60}\approx2.38<3.841$，　…………………………4 分

所以没有 95%的把握认为产品的合格率与生产厂家有关. ……………………………6 分

(2)对于甲厂，抽到的 100 件产品中有 A 等级产品 10 件，B 等级产品 65 件，C 等级产品 25 件，设生产一件产品的利润为 X 元，则 X 可能取得的值为 30，10，－34，X 的分布列为

| $X$ | 30 | 10 | －34 |
|---|---|---|---|
| $P$ | 0.1 | 0.65 | 0.25 |

因为 E(Y)=30×0.1+10×0.65+(－34)×0.25=1>0，

所以甲厂能盈利. ……………………………………………………………………9 分

对于乙厂，抽到的 100 件产品中有 A 等级产品 10 件，B 等级产品 55 件，C 等级产品 35 件，设生产一件产品的利润为 Y 元，则 Y 可能取得的值为 30，10，－34，Y 的分布列为

| Y | 30 | 10 | －34 |
|---|---|---|---|
| $P$ | 0.1 | 0.55 | 0.35 |

因为 $E\left(Y\right)=30\times0.1+10\times0.55+\left(-34\right)\times0.35=-3.4<0$，

所以乙厂不能盈利. ……………………………………………………………………12 分

**23.**

20.（12 分）解：（1）记“未来一周从周一到周五5天中至少有一天暂停会展”为事件 $A$，则事件 $\overline{A}$ 表示未来一周5天展出会展，于是

$$P(\overline{A})=(1-\frac{1}{2})^3(1-\frac{4}{5})^2=\frac{1}{200},$$ ………………………………1 分

$$P(A)=1-P(\overline{A})=1-\frac{1}{200}=\frac{199}{200},$$

所以，未来一周从周一到周五5天中至少有一天暂停会展的概率是 $\frac{199}{200}$. ……2 分

（2）设随机变量 $X$ 表示会展展出的天数，则 $x=0,1,2,3,4,5$

于是，$P(x=0)=(\frac{1}{2})^3(\frac{4}{5})^2=\frac{2}{25}$， ……………………………………3 分

$$P(x=1)=C_3^1(\frac{1}{2})^3C_2^2(\frac{4}{5})^2+C_3^3(\frac{1}{2})^3C_2^1(\frac{4}{5})(\frac{1}{5})=\frac{7}{25},$$ ………………5 分

$$P(x=2)=C_3^2(\frac{1}{2})^3C_2^2(\frac{4}{5})^2+C_3^1(\frac{1}{2})^3C_2^1(\frac{4}{5})(\frac{1}{5})+C_3^3(\frac{1}{2})^3C_2^2(\frac{1}{5})^2=\frac{73}{200},$$ …7 分

$$P(x=3)=C_3^3(\frac{1}{2})^3C_2^2(\frac{4}{5})^2+C_3^2(\frac{1}{2})^3C_2^1(\frac{4}{5})(\frac{1}{5})+C_3^1(\frac{1}{2})^3C_2^2(\frac{1}{5})^2=\frac{43}{200},$$ …9 分

$$P(x=4)=C_3^3(\frac{1}{2})^3C_2^1(\frac{4}{5})(\frac{1}{5})+C_3^2(\frac{1}{2})^3C_2^2(\frac{1}{5})^2=\frac{11}{200},$$ ……………11 分

由（1）知，$P(x=5)=P(\bar{A})=\dfrac{1}{200}$，

所以 $E(x)=\sum_{i=0}^{5} iP(x=i)=1\times\dfrac{7}{25}+2\times\dfrac{73}{200}+3\times\dfrac{43}{200}+4\times\dfrac{11}{200}+5\times\dfrac{1}{200}=1.9$，

即这次会展活动展出的平均天数是1.9天．  ……………………………12 分

**24.【解析】**（1）会员购买礼包 C 的概率为 $1-\dfrac{2}{5}\times 2=\dfrac{1}{5}$，

∴准备鸡蛋：$400\times\dfrac{1}{5}\times 10=800$（斤），

（2）Y 的所有可能取值为：150，125，100，75，

$P(Y=150)=\left(\dfrac{4}{5}\right)^3=\dfrac{64}{125}$，$P(Y=125)=C_3^1\cdot\dfrac{1}{5}\times\left(\dfrac{4}{5}\right)^2=\dfrac{48}{125}$

$P(Y=100)=C_3^2\cdot\left(\dfrac{1}{5}\right)^2\times\dfrac{4}{5}=\dfrac{12}{125}$，$P(Y=75)=\left(\dfrac{1}{5}\right)^3=\dfrac{1}{125}$

∴Y 的分布列如下：

| $Y$ | 150 | 125 | 100 | 75 |
|---|---|---|---|---|
| $P$ | $\dfrac{64}{125}$ | $\dfrac{48}{125}$ | $\dfrac{12}{125}$ | $\dfrac{1}{125}$ |

$\therefore E(Y)=150\times\dfrac{64}{125}+125\times\dfrac{48}{125}+100\times\dfrac{12}{125}+75\times\dfrac{1}{125}=\dfrac{384}{5}+48+\dfrac{48}{5}+\dfrac{3}{5}=135$．

**25.【解析】**（1）根据以上数据，$K^2$ 的观测值 $k=\dfrac{200\times(90\times30-50\times30)^2}{140\times60\times120\times80}\approx3.571>2.706$，

∴有 90%的把握认为 $A$ 市能否正确进行垃圾分类处理与年龄有关．

（2）由题意可得：$X\sim B(3,\dfrac{1}{4})$，

∴$P(X=k)=C_3^k\times(\dfrac{1}{4})^k(1-\dfrac{1}{4})^{3-k}$，$k=0$，1，2，3，

$P(X=0)=\dfrac{27}{64}$，$P(X=1)=\dfrac{27}{64}$，$P(X=2)=\dfrac{9}{64}$，$P(X=3)=\dfrac{1}{64}$．

可得：随机变量 $X$ 的分布列：

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | $\dfrac{27}{64}$ | $\dfrac{27}{64}$ | $\dfrac{9}{64}$ | $\dfrac{1}{64}$ |

均值 $E(X)=3\times\frac{1}{4}=\frac{3}{4}$.

**26.【解析】**(1)由频数分布表得：$\overline{x}=\frac{14\times5+17\times6+20\times9+23\times12+26\times8+29\times6+32\times}{50}=22.76\approx22.8$.

所以这 50 个社区这一天垃圾量的平均值为 22.8 吨.

（2）由（1）知 $\mu=22.8$，$\because s=5.2$，$\sigma=s=5.2$，

$\therefore P(X>28)=P(X>\mu+\sigma)=\frac{1-0.6827}{2}=0.15865$，

$\because 320\times0.15865=50.768\approx51$，

所以这 320 个社区中“超标”社区的个数为 51.

（3）由频数分布表知：8 个“超标”社区中这一天的垃圾量至少为 30.5 吨的社区有 4 个，

所以 $Y$ 的可能取值为 1，2，3，4，

$P(Y=1)=\frac{C_4^1C_4^4}{C_8^5}=\frac{1}{14}$，$P(Y=2)=\frac{C_4^2C_4^3}{C_8^5}=\frac{3}{7}$，$P(Y=3)=\frac{C_4^3C_4^2}{C_8^5}=\frac{3}{7}$，$P(Y=4)=\frac{C_4^4C_4^1}{C_8^5}=\frac{1}{14}$，

所以 $Y$ 的分布列为：

| $Y$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P$ | $\frac{1}{14}$ | $\frac{3}{7}$ | $\frac{3}{7}$ | $\frac{1}{14}$ |

$\therefore E(Y)=1\times\frac{1}{14}+2\times\frac{3}{7}+3\times\frac{3}{7}+4\times\frac{1}{14}=\frac{5}{2}$.

**27.【解析】**（1）根据分层抽样，从[40，50）中抽取 6 人，在[80，90）中抽取 3 人，

随机变量 $X$ 的可能取值为 0，1，2，3，

$P(X=0)=\frac{C_6^0C_3^3}{C_9^3}=\frac{1}{84}$，

$P(X=1)=\frac{C_6^1C_3^2}{C_9^3}=\frac{3}{14}$，

$P(X=2)=\frac{C_6^2C_3^1}{C_9^3}=\frac{15}{28}$，

$P(X=3)=\frac{C_6^3C_3^0}{C_9^3}=\frac{5}{21}$，

则 $X$ 的分布列为：

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|

| P | $\frac{1}{84}$ | $\frac{3}{14}$ | $\frac{15}{28}$ | $\frac{5}{21}$ |
|---|---|---|---|---|

$E(X)=0\times\frac{1}{84}+1\times\frac{3}{14}+2\times\frac{15}{28}+3\times\frac{5}{21}=2$.

（2）$\mu=\overline{t}=35\times0.01\times10+45\times0.02\times10+55\times0.03\times10+65\times0.015\times10+75\times0.015\times10+85\times0.01\times10=58.5$，

又因为 $43.9=58.5-14.6=\mu-\sigma$，$87.7=58.5+2\times14.6=\mu+2\sigma$，

所以 $P(43.9<t\leqslant87.7)=P(\mu-\sigma<t\leqslant\mu+2\sigma)=\frac{0.6826+0.9544}{2}=0.8185$，

所以 $P(t\leqslant43.9$ 或 $t>87.7)=1-0.8185=0.1815$，

则 $Y\sim B(12,0.1815)$，

所以 $P(Y=3)=C_{12}^{3}\times0.1815^3\times0.8185^9=220\times0.0060\times0.1649\approx0.218$.

28.【解析】(1)该混合样本阴性的概率是 $\left(\sqrt{p}\right)^2=\frac{2}{3}$

据对立事件可得，阳性的概率为 $1-\frac{2}{3}=\frac{1}{3}$

(2)方案一:混在一起检验，方案一的检验次数记 X，则 X 的可能取值为 1,5

$P(X=1)=(\sqrt{p})^4=p^2$，$P(X=5)=1-p^2$

其分布列为:

| X | 1 | 5 |
|---|---|---|
| P | $p^2$ | $1-p^2$ |

则 $E(X)=5-4p^2$

方案二:由题意分析可知，每组 2 份样本混合检验时，若阴性则检测次数为 1,概率为 $\left(\sqrt{p}\right)^2=p$，若阳性，则检测次数为 3，概率为 $1-p$，方案二的检验次数记为 Y，则 Y 的可能取值为 2, 4, 6，

$P(Y=2)=p^2, P(Y=4)=C_2^1p(1-p)$，$P(Y=6)=(1-p)^2$

其分布列为:

| $Y$ | 2 | 4 | 6 |
|---|---|---|---|
| $P$ | $p^2$ | $2p-2p^2$ | $(1-p)^2$ |

则 $E(Y)=2p^2+4(2p-2p^2)+6(1-p)^2=6-4p$

$E(Y)-E(X)=6-4p-(5-4p^2)=4p^2-4p+1$

当 $p\neq\frac{1}{2}$ 时，可得 $E(X)<E(Y)$，所以方案一更“优”

当 $p=\frac{1}{2}$ 时，可得 $E(X)=E(Y)$，所以方案一、二一样

**29.**【解析】（1）补全的列联表如下：

| | 年轻人 | 非年轻人 | 合计 |
|---|---|---|---|
| 单车族 | 200 | 40 | 240 |
| 非单车族 | 120 | 40 | 160 |
| 合计 | 320 | 80 | 400 |

$\therefore K^2=\frac{400(200\times40-120\times40)^2}{240\times160\times320\times80}\approx4.167>3.841$，

即有 95%的把握可以认为经常使用共享单车与年龄有关．

（2）由（1）的列联表可知，既是“单车族”又是“非年轻人”占样本总数的频率为$\frac{40}{400}\times100\%=10\%$，

即在抽取的用户中既是“单车族”又是“非年轻人”的概率为 0.1，

$\because X\sim B(3,0.1)$，$X=0$，1，2，3，

$\therefore P(X=0)=(1-0.1)^3=0.729$，

$P(X=1)=C_3^1\times0.1\times(1-0.1)^2=0.243$，

$P(X=2)=C_3^2\times0.1^2\times(1-0.1)=0.027$，

$P(X=3)=0.1^3=0.001$，

$\therefore X$ 的分布列为：

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P$ | 0.729 | 0.243 | 0.027 | 0.001 |

∴$X$ 的数学期望 $E$（$X$）＝0×0.729+1×0.243+2×0.027+3×0.001＝0.3.

**30.**【解析】

（1）设根据剔除后数据建立的 $y$ 关于 $x$ 的回归直线方程为 $\hat{y}=\hat{b}x+\hat{a}$，

剔除异常数据后的数学平均分为 $\dfrac{1110-110}{10}=100$，

剔除异常数据后的物理平均分为 $\dfrac{660-0}{10}=66$，

则 $\hat{b}=\dfrac{68586-110\times0-10\times66\times100}{120426-110^2-10\times100^2}=\dfrac{2586}{8326}\approx0.31$，

则 $\hat{a}=66-0.31\times100=35$，

所以 所求回归直线方程为 $\hat{y}=0.31x+35$．

又物理缺考考生的数学成绩为 110,

所以 估计其可能取得的物理成绩为 $\hat{y}=0.31\times110+35=69.1$．

（2）由题意知 $\mu=66$，

因为 $\sum\limits_{i=1}^{11}y_i^2=\sum\limits_{i=1}^{11}(y_i-\overline{y})^2+11\overline{y}^2=4770+11\times(\dfrac{660}{11})^2=44370$，

所以 $\sigma=\sqrt{\dfrac{1}{10}\times44370-66^2}=\sqrt{81}=9$，

所以 参加该次考试的 10000 名考生的物理成绩服从正态分布 $N(66,9^2)$，

则物理成绩不低于 75 分的概率为 $\dfrac{1-0.683}{2}=0.1585$，

由题意可知 $Y\sim B(10000,0.1585)$，

所以 物理成绩不低于 75 分的人数 $Y$ 的期望

$EY=10000\times0.1585=1585$．