**十年（**2014**－**2023**）年高考真题分项汇编—概率统计解答题**

**目录**

[**题型一：二项式定理 1**](#_Toc140775634)

[**题型二：事件的频率与概率 3**](#_Toc140775635)

[**题型三：随机变量的分布列与期望、方差 8**](#_Toc140775636)

[**题型四：概率统计中的决策建议 44**](#_Toc140775637)

[**题型五：简单的随机抽样与用样本估计总体 49**](#_Toc140775638)

[**题型六：相关关系与回归分析 62**](#_Toc140775639)

[**题型七：独立性检验 67**](#_Toc140775640)

[**题型八：概率统计综合应用 74**](#_Toc140775641)

# 题型一：二项式定理

1．(2019·江苏·第24题)设.已知.

(1)求的值；(2)设，其中，求的值.

**【答案】**见解析

【解析】(1)因为，

所以，

．

因为，

所以，

解得．

(2)由(1)知，．





．

**解法一：**因为，所以，

从而．

**解法二：**

．

因为，所以．

因此．

2．(2016高考数学江苏文理科·第26题)(1)求的值；

(2)设，．

求证：．

**【答案】**(1)；(2)详见解析；

【官方解答】(1)；

(2)当时，结论显然成立．当时，



．

又因为，

所以，．

因此，

＝

＝

＝．

民间解答：(1)；

(2)对任意的，

① 当时，左边，右边，等式成立，

② 假设时命题成立，

即，

当时，

左边=

，

右边，

而，



因此，

因此左边=右边，

因此时命题也成立，

综合①②可得命题对任意均成立．

另解：因为，所以

左边

又由，知

，

所以，左边右边．

# 题型二：事件的频率与概率

1．(2022新高考全国I卷·第20题)一医疗团队为研究某地的一种地方性疾病与当地居民的卫生习惯(卫生习惯分为良好和不够良好两类)的关系，在已患该疾病的病例中随机调查了100例(称为病例组)，同时在未患该疾病的人群中随机调查了100人(称为对照组)，得到如下数据：

|  |  |  |
| --- | --- | --- |
|  | 不够良好 | 良好 |
| 病例组 | 40 | 60 |
| 对照组 | 10 | 90 |

(1)能否有99%的把握认为患该疾病群体与未患该疾病群体的卫生习惯有差异？

(2)从该地的人群中任选一人，*A*表示事件“选到的人卫生习惯不够良好”，*B*表示事件“选到的人患有该疾病”．与的比值是卫生习惯不够良好对患该疾病风险程度的一项度量指标，记该指标为*R*．

(ⅰ)证明：；

(ⅱ)利用该调查数据，给出的估计值，并利用(ⅰ)的结果给出*R*的估计值．

附，

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0．050 | 0．010 | 0．001 |
| *k* | 3．841 | 6．635 | 10．828 |

**【答案】**(1)答案见解析

(2)(i)证明见解析；(ii)；

解析：(1)由已知，

又，，

所以有99%的把握认为患该疾病群体与未患该疾病群体的卫生习惯有差异．

(2)(i)因为，

所以 所以，

(ii)由已知，，

又，， 所以

2．(2022高考北京卷·第18题)在校运动会上，只有甲、乙、丙三名同学参加铅球比赛，比赛成绩达到以上(含)的同学将获得优秀奖．为预测获得优秀奖的人数及冠军得主，收集了甲、乙、丙以往的比赛成绩，并整理得到如下数据(单位：m)：

甲：9．80，9．70，9．55，9．54，9．48，9．42，9．40，935，9．30，9．25；

乙：9．78，9．56，9．51，9．36，9．32，9．23；

丙：9．85，9．65，9．20，9．16．

假设用频率估计概率，且甲、乙、丙的比赛成绩相互独立．

(1)估计甲在校运动会铅球比赛中获得优秀奖的概率；

(2)设*X*是甲、乙、丙在校运动会铅球比赛中获得优秀奖的总人数，估计*X*的数学期望*E*(*X*)；

(3)在校运动会铅球比赛中，甲、乙、丙谁获得冠军的概率估计值最大？(结论不要求证明)

**【答案】**解析:(1)由频率估计概率可得

甲获得优秀的概率为0．4，乙获得优秀的概率为0．5，丙获得优秀的概率为0．5，故答案为0．4

(2)设甲获得优秀为事件*A*1,乙获得优秀为事件*A*2，丙获得优秀为事件*A*3

，



，



，

．

∴*X*的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *X* | 0 | 1 | 2 | 3 |
| *P* |  |  |  |  |

∴

(3)丙夺冠概率估计值最大．

因为铅球比赛无论比赛几次就取最高成绩．比赛一次，丙获得9．85的概率为，甲获得9．80的概率为，乙获得9．78的概率为．并且丙的最高成绩是所有成绩中最高的，比赛次数越多，对丙越有利．

3．(2020年高考课标Ⅰ卷理科·第19题)甲、乙、丙三位同学进行羽毛球比赛，约定赛制如下：累计负两场者被淘汰；比赛前抽签决定首先比赛的两人，另一人轮空；每场比赛的胜者与轮空者进行下一场比赛，负者下一场轮空，直至有一人被淘汰；当一人被淘汰后，剩余的两人继续比赛，直至其中一人被淘汰，另一人最终获胜，比赛结束．经抽签，甲、乙首先比赛，丙轮空．设每场比赛双方获胜的概率都为，

(1)求甲连胜四场的概率；

(2)求需要进行第五场比赛的概率；

(3)求丙最终获胜的概率．

**【答案】**(1)；(2)；(3)．

【解析】(1)记事件甲连胜四场，则；

(2)记事件为甲输，事件为乙输，事件为丙输，

则四局内结束比赛的概率为

，

所以，需要进行第五场比赛的概率为；

(3)记事件为甲输，事件为乙输，事件为丙输，

记事件甲赢，记事件丙赢，

则甲赢的基本事件包括：、、、

、、、、，

所以，甲赢概率为．

由对称性可知，乙赢的概率和甲赢的概率相等，

所以丙赢的概率为．

【点睛】本题考查独立事件概率的计算，解答的关键就是列举出符合条件的基本事件，考查计算能力，属于中等题．

4．(2019·全国Ⅱ·理·第18题)分制乒乓球比赛，每赢一球得分，当某局打成平后，每球交换发球权，先多得分的一方获胜，该局比赛结束．甲、乙两位同学进行单打比赛，假设甲发球时甲得分的概率为，乙发球时甲得分的概率为，各球的结果相互独立．在某局双方平后，甲先发球，两人又打了个球该局比赛结束．

求；

求事件“且甲获胜”的概率．

**【答案】**；.

**【官方解析】**

**就是平后，两人又打了个球该局比赛结束，则这个球均由甲得分，或者均由乙得分．因此．

**且甲获胜，就是平后，两人又打了个球该局比赛结束，且这个球的得分情况为：前两球是甲、乙各得分，后两球均为甲得分．

因此所求概率为

．

**【分析】**

本题首先可以通过题意推导出所包含的事件为“甲连赢两球或乙连赢两球”，然后计算出每种事件的概率并求和即可得出结果；

本题首先可以通过题意推导出所包含的事件为“前两球甲乙各得分，后两球均为甲得分”，然后计算出每种事件的概率并求和即可得出结果.

**【解析】**由题意可知，www.zqy.com所包含的事件为“甲连赢两球或乙连赢两球”，

所以www.zqy.com.

由题意可知，包含的事件为“前两球甲乙各得分，后两球均为甲得分”

所以www.zqy.com.

**【点评】**本题考查古典概型的相关性质，能否通过题意得出www.zqy.com以及www.zqy.com所包含的事件是解决本题的关键，考查推理能力，考查学生从题目中获取所需信息的能力，是中档题.

5．(本小题满分12分)甲、乙两台机床相互没有影响地生产某种产品，甲机床产品的正品率是0．9，乙机床产品的正品率是0．95．

(Ⅰ)从甲机床生产的产品中任取3件，求其中恰有2件正品的概率(用数字作答)；

(Ⅱ)从甲、乙两台机床生产的产品中各任取1件，求其中至少有1件正品的概率

(用数字作答)．

**【答案】**分析：本小题考查互斥事件、相互独立事件的概率等基础知识，及分析和解决实际问题的能力。

(I)任取甲机床的3件产品恰有2件正品的概率为



(II)解法一：记“任取甲机床的1件产品是正品”为事件A，“任取乙机床的1件产品是正品”为事件B。则任取甲、乙两台机床的产品各1件，其中至少有1件正品的概率为



解法二：运用对立事件的概率公式，所求的概率为



6．(2019·天津·理·第16题)设甲、乙两位同学上学期间，每天7：30之前到校的概率均为．假定甲、乙两位同学到校情况互不影响，且任一同学每天到校情况相互独立．

(Ⅰ)用表示甲同学上学期间的三天中7：30之前到校的天数，求随机变量的分布列和数学期望；

(Ⅱ)设为事件“上学期间的三天中，甲同学在7：30之前到校的天数比乙同学在7：30之前到校的天数恰好多2”，求事件发生的概率．

**【答案】**本小题主要考查离散型随机变量的分布列与数学期望，互斥事件和相互独立事件的概率计算公式等基础知识．考查运用概率知识解决简单实际问题的能力．满分13分．

(Ⅰ)解：因为甲同学上学期间的三天中到校情况相互独立，且每天7：30之前到校的概率均为，故，从而．

所以，随机变量的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 |
|  |  |  |  |  |

随机变量的数学期望．

(Ⅱ)解：设乙同学上学期间的三天中7：30之前到校的天数为，则，

且．由题意知事件与互斥，

且事件与，事件与均相互独立，从而由(Ⅰ)知



．

# 题型三：随机变量的分布列与期望、方差

1．(2022年高考全国甲卷数学(理)·第19题)甲、乙两个学校进行体育比赛，比赛共设三个项目，每个项目胜方得10分，负方得0分，没有平局．三个项目比赛结束后，总得分高的学校获得冠军．已知甲学校在三个项目中获胜的概率分别为0．5，0．4，0．8，各项目的比赛结果相互独立．

(1)求甲学校获得冠军的概率；

(2)用*X*表示乙学校的总得分，求*X*的分布列与期望．

**【答案】**(1)； (2)分布列见解析，．

【解析】(1)设甲在三个项目中获胜的事件依次记为，所以甲学校获得冠军的概率为





．

(2)依题可知，的可能取值为，所以，

,

，

，

．

即的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 10 | 20 | 30 |
|  | 0．16 | 0．44 | 0．34 | 0．06 |

期望．

2．(2021高考北京·第18题)在核酸检测中, “*k*合1” 混采核酸检测是指：先将*k*个人的样本混合在一起进行1次检测,如果这*k*个人都没有感染新冠病毒，则检测结果为阴性，得到每人的检测结果都为阴性，检测结束:如果这*k*个人中有人感染新冠病毒，则检测结果为阳性，此时需对每人再进行1次检测,得到每人的检测结果，检测结束．

现对100人进行核酸检测，假设其中只有2人感染新冠病毒，并假设每次检测结果准确．

(I)将这100人随机分成10组，每组10人，且对每组都采用“10合1”混采核酸检测．

(i)如果感染新冠病毒的2人在同一组，求检测的总次数;

(ii)已知感染新冠病毒的2人分在同一组的概率为．设*X*是检测的总次数，求*X*的

分布列与数学期望E(X)．

(II)将这100人随机分成20组，每组5人，且对每组都采用“5合1”混采核酸检测．设Y是检测的总次数，试判断数学期望E(*Y*)与(I)中E(X)的大小．(结论不要求证明)

**【答案】(1)①次；②分布列见解析；期望为；(2)．**

**解析：(1)①对每组进行检测，需要10次；再对结果为阳性的组每个人进行检测，需要10次；**

**所以总检测次数为20次；**

**②由题意，可以取20，30，**

**，，**

**则的分布列:**

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

**所以；**

**(2)由题意，可以取25，30，**

**两名感染者在同一组的概率为，不在同一组的概率为，**

**则．**

3．(2019·江苏·第25题)在平面直角坐标系*xOy*中，设点集，

令.从集合*Mn*中任取两个不同的点，用随机变量*X*表示它们之间的距离.

(1)当时，求*X*的概率分布；

(2)对给定的正整数，求概率(用表示).

**【答案】**见解析

【解析】(1)当时，的所有可能取值是．

*X*的概率分布为，

．

(2)设和是从中取出的两个点．

因为，所以仅需考虑的情况．

①若，则，不存在的取法；

②若，则，所以当且仅当，此时或，有2种取法；

③若，则，因为当时，，所以当且仅当，此时或，有2种取法；

④若，则，所以当且仅当，此时或，有2种取法．

综上，当时，*X*的所有可能取值是和，且

．

因此，．

4．(2014高考数学陕西理科·第21题)在一块耕地上种植一种作物，每季种植成本为1000元，此作物的市场价格和这块地上的产量具有随机性，且互不影响，其具体情况如下表：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 作物产量() |  |  |  | 作物市场价格(元/) |  |  |
| 概率 |  |  |  | 概率 |  |  |

(1)设表示在这块地上种植1季此作物的利润，求的分布列；

(2)若在这块地上连续3季种植此作物，求这3季中至少有2季的利润不少于2000元

的概率．

**【答案】**(1)详见解析；(2)

解析: ⑴设表示事件“作物产量为”，表示事件“作物市场价格为元/”，

由题设知，，

∵利润=产量×市场价格—成本，

∴有可能的取值为

，，

，．

，

，

，

∴的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

⑵设表示事件“第季利润不少于元”()，

由题意知，，相互独立，由⑴知，

()，

季的利润均不少于元的概率为

，

季中有季利润不少于元的概率为

，

所以，这季中至少有季的利润不少于元的概率为．

5．(2014高考数学重庆理科·第18题)一盒中装有9张各写有一个数字的卡片，其中4张卡片上的数字是1,3张卡片上的数字

是2,2张卡片上的数字是3，从盒中任取3张卡片．

(1)求所取3张卡片上的数字完全相同的概率;

(2)表示所取3张卡片上的数字的中位数，求的分布列(注：若三个数满足，则称为这三个数的中位数)．

**【答案】**(1)； (2)

解析：(1)由古典概型计算公式可得。

(2)的可能值为，由题意



，



所以的分布列为：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

从而。

6．(2023年新课标全国Ⅰ卷·第21题)甲、乙两人投篮，每次由其中一人投篮，规则如下：若命中则此人继续投籃，若末命中则换为对方投篮．无论之前投篮情况如何，甲每次投篮的命中率均为0．6，乙每次投篮的命中率均为0．8．由抽签确定第1次投篮的人选，第1次投篮的人是甲、乙的概率各为0．5．

(1)求第2次投篮的人是乙的概率；

(2)求第次投篮的人是甲的概率；

(3)已知：若随机变量服从两点分布，且，则．记前次(即从第1次到第次投篮)中甲投篮的次数为，求．

**【答案】**(1)

(2)

(3)

解析：(1)记“第次投篮的人是甲”为事件，“第次投篮的人是乙”为事件，

所以，

．

(2)设，依题可知，，则

，

即，

构造等比数列，

设，解得，则，

又，所以是首项为，公比为的等比数列，

即．

(3)因为，，

所以当时，，

故．

7．(2020江苏高考·第25题)甲口袋中装有个黑球和个白球，乙口袋中装有个白球．现从甲、乙两口袋中各任取一个球交换放入另一口袋，重复次这样的操作，记甲口袋中黑球个数为，恰有个黑球的概率为，恰有个黑球的概率为．

(1)求和；

(2)求与的递推关系式和的数学期望(用表示)．

**【答案】**(1)(2)

【解析】(1)，

，



(2)，

，

因此，

从而，

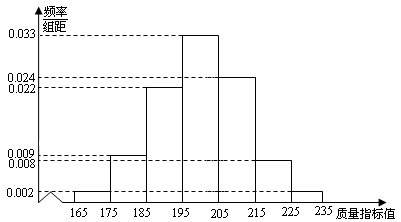
即．

又的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0 | 1 | 2 |
|  |  |  |  |

故．

8．(2014高考数学课标1理科·第18题)从某企业的某种产品中抽取500件,测量这些产品的一项质量指标值,由测量结果得如下频率分布直方图:



(1)求这500件产品质量指标值的样本平均数高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。和样本方差高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。(同一组数据用该区间的中点值作代表);

(2)由频率分布直方图可以认为,这种产品的质量指标值高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。服从正态分布高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。,其中高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。近似为样本平均数高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。,高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。近似为样本方差高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。．

(i)利用该正态分布,求高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。;

(ii)某用户从该企业购买了100件这种产品,记高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。表示这100件产品中质量指标值为于区间(187．8,212．2)的产品件数,利用(i)的结果,求高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。．

附:高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。．

若高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。~高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。,则高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。．

**【答案】**解析:(1)抽取产品质量指标值的样本平均数高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。和样本方差高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。分别为 高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。 高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。 (2)(ⅰ)由(1)知高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。~高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。,从而 高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。． (ⅱ)由(ⅰ)知,一件产品中质量指标值为于区间高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。的概率为高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。 依题意知高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。,所以高考资源网(ks5u.com),中国最大的高考网站,您身边的高考专家。． 考点:(1)频率分布直方图的绘制及应用;(2)离散型随机变量的均值及方差;(3)正态分布的应用;(4)数形结合思想 难度:C 备注:高频考点

8．(2014高考数学天津理科·第16题)某大学志愿者协会有名男同学,名女同学．在这名同学中,名同学来自数学学院,其余名同学来自物理、化学等其他互不相同的七个学院．现从这名同学中随机选取名同学,到希望小学进行支教活动(每位同学被选到的可能性相同)．

(Ⅰ)求选出的名同学是来自互不相同学院的概率;

(Ⅱ)设为选出的名同学中女同学的人数,求随机变量的分布列和数学期望．

**【答案】**(Ⅰ);(Ⅱ)详见解析．解析:(Ⅰ)设“选出的名同学是来自互不相同的学院”为事件,则．所以,选出的名同学是来自互不相同学院的概率为．(Ⅱ)随机变量的所有可能值为,则所以,随机变量的分布列是

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |

随机变量的数学期望．

9．(2014高考数学四川理科·第17题)一款击鼓小游戏的规则如下：每盘游戏都需要击鼓三次，每次击鼓要么出现一次音乐，要么不出现音乐；每盘游戏击鼓三次后，出现一次音乐获得10分，出现两次音乐获得20分，出现三次音乐获得100分，没有出现音乐则扣除200分(即获得-200分)．设每次击鼓出现音乐的概率为，且各次击鼓出现音乐相互独立．

(Ⅰ)设每盘游戏获得的分数为，求的分布列；

(Ⅱ)玩三盘游戏，至少有一盘出现音乐的概率是多少？

(Ⅲ)玩过这款游戏的许多人发现，若干盘游戏后，与最初的分数相比，分数没有增加反而减少了．请运用概率统计的相关知识分析分数减少的原因．

**【答案】**解析：(Ⅰ)可能的取值为：10，20，100，．根据题意，有

，　　　，，

所以的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | 10 | 20 | 100 |
| P |  |  |  |  |

(Ⅱ)设“第盘游戏没有出现音乐”为事件，则

．

所以“三盘游戏中至少有一次出现音乐”的概率为

．

因此，玩三盘游戏至少有一盘出现音乐的概率是．

(Ⅲ)的数学期望为．

　　　这表明，获得分数的均值为负，

　　　因此，多次游戏之后分减少的可能性更大．

10．(2014高考数学山东理科·第18题)乒乓球台面被球网分隔成甲、乙两部分，如图，甲上有两个不相交的区域，乙被划分为两个不相交的区域．某次测试要求队员接到落点在甲上的来球后向乙回球．规定：回球一次，落点在上记3分，在上记1分，其它情况记0分．对落点在上的来球，队员小明回球的落点在上的概率为，在上的概率为；对落点在上的来球，小明回球的落点在上的概率为，在上的概率为．假设共有两次来球且落在上各一次，小明的两次回球互不影响．求：

(Ⅰ)小明两次回球的落点中恰有一次的落点在乙上的概率；

(Ⅱ)两次回球结束后，小明得分之和的分布列与数学期望．



**【答案】**(Ⅰ)；(Ⅱ)见解析

解析：(Ⅰ)记为事件“小明对落点在上的来球回球的得分为分”()

则；

记为事件“小明对落点在上的来球回球的得分为分”()

则；

记为事件“小明两次回球的落点中恰有一次的落点在乙上”，

由题意，，由事件的独立性及互斥性得

()





．

所以小明两次回球的落点中恰有一次的落点在乙上的概率为．

(Ⅱ)由题意，随机变量可能的取值为．

由事件的独立性及互斥性得

；

；

；

；

；

．

可得随机变量的分布列为

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

所以，数学期望．

11．(2014高考数学辽宁理科·第18题)(本小题满分12分)

一家面包房根据以往某种面包的销售记录，绘制了日销售量的频率分布直方图，如图所示：

0.003

0

0.005

50

0.006

100

0.004

150

0.002

200

250

日销售量/个

将日销售量落入各组的频率视为概率，并假设每天的销售量相互独立．

(1)求在未来连续3天里，有连续2天的日销售量都不低于100个且另一天的日销售量低于50个的概率；

(2)用X表示在未来3天里日销售量不低于100个的天数，求随机变量X的分布列，期望及方差．

**【答案】**

(1)0．108；(2)1．8，0．72．

解析：(1)设A1表示事件“日销售量不低于100个”，A2表示事件“日销售量低于50个”

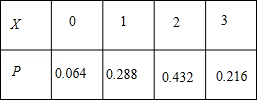
B表示事件“在未来连续3天里，有连续2天的日销售量都不低于100个且另1天的日销售量低于50个”，因此P(A1)=(0．006+0．004+0．002)×50=0．6，P(A2)=0．003×50=0．15，

P(B)=0．6×0．6×0．15×2=0．108，

(Ⅱ)X可能取的值为0，1，2，3，相应的概率为：

，，,，

随机变量X的分布列为



因为X～B(3，0．6)，所以期望E(X)=3×0．6=1．8，方差D(X)=3×0．6×(1﹣0．6)=0．72．

12．(2014高考数学江西理科·第22题)随机将这2n个连续正整数分成A,B两组,每组n个数,A组最小数为,最大数为;B组最小数为,最大数为,记

(1)当时,求的分布列和数学期望;

(2)令C表示事件与的取值恰好相等,求事件C发生的概率;

(3)对(2)中的事件C,表示C的对立事件,判断和的大小关系,并说明理由．

2014高考数学江西理科

**【答案】**(1)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 2 | 3 | 4 | 5 |
| P |  |  |  |  |

(2)当时,,当时 (3)当时,当时, 分析:(1)当时,将6个正整数平均分成A,B两组,不同的分组方法共有种,所有可能值为2,3,4,5．对应组数分别为4,6,6,4,对应概率为,,,,(2)和恰好相等的所有可能值为当和恰好相等且等于时,不同的分组方法有2种;当和恰好相等且等于时,不同的分组方法有2种;当和恰好相等且等于时,不同的分组方法有2种;当和恰好相等且等于时,不同的分组方法有2种;以此类推:和恰好相等且等于时,不同的分组方法有2种;所以当时, 当时(3)先归纳:当时,因此当时,即证当时,这可用数学归纳法证明． 当时,,利用阶乘作差可得大小． 解析:(1)当时,所有可能值为2,3,4,5．将6个正整数平均分成A,B两组,不同的分组方法共有种,所以的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 2 | 3 | 4 | 5 |
| P |  |  |  |  |

 (2)和恰好相等的所有可能值为 又和恰好相等且等于时,不同的分组方法有2种; 和恰好相等且等于时,不同的分组方法有2种; 和恰好相等且等于时,不同的分组方法有2种; 所以当时, 当时 (3)由(2)当时,因此 而当时,理由如下: 等价于① 用数学归纳法来证明: 当时,①式左边①式右边所以①式成立 假设时①式成立,即成立 那么,当时,①式左边  =①式右边 即当时①式也成立 综合得,对于的所有正整数,都有成立

13．(2014高考数学湖南理科·第17题)某企业有甲、乙两个研发小组,他们研发新产品成功的概率分别为和．现安排甲组研发新产品,乙组研发新产品．设甲、乙两组的研发相互独立．

(Ⅰ)求至少有一种新产品研发成功的概率;

(Ⅱ)若新产品研发成功,预计企业可获利润120万元;若新产品研发成功,预计企业可获利润100万元．求该企业可获利润的分布列和数学期望．

**【答案】**(1) (2)详见解析

解析:(1)解:设至少有一组研发成功的事件为事件且事件为事件的对立事件,则事件为一种新产品都没有成功,因为甲,乙成功的概率分别为,

则,再根据对立事件概率之间的公式可得

,所以至少一种产品研发成功的概率为．

(2)由题可得设该企业可获得利润为,则的取值有即,由独立试验的概率计算公式可得:

;;

;;

所以的分布列如下:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |

则数学期望．

|  |  |  |  |
| --- | --- | --- | --- |
| 年入流量 |  |  |  |
| 发电量最多可运行台数 | 1 | 2 | 3 |

14．(2014高考数学湖北理科·第20题)计划在某水库建一座至多安装3台发电机的水电站，过去50年的水文资料显示，水库年入流量

(年入流量：一年内上游来水与库区降水之和． 单位：亿立方米)都在40以上．其中，不足

80的年份有10年，不低于80且不超过120的年份有35年，超过120的年份有5年．将年入流

量在以上三段的频率作为相应段的概率，并假设各年的年入流量相互独立．

(Ⅰ)求未来4年中，至多有1年的年入流量超过120的概率；

(Ⅱ)水电站希望安装的发电机尽可能运行，但每年发电机最多可运行台数受年入流量限制，

并有如下关系:

若某台发电机运行，则该台年利润为5000万元；若某台发电机未运行，则该台年亏损800万

元，欲使水电站年总利润的均值达到最大，应安装发电机多少台？

**【答案】**(1)；(2)2．

解析：(1)依题意，，

，，

由二项分布，在未来4年中至多有1年入流量找过120的概率为：

．

(2)记水电站年总利润为(单位：万元)

①安装1台发电机的情形．

由于水库年入流量总大于40，所以一台发电机运行的概率为1，

对应的年利润，．

②安装2台发电机．

当时，一台发电机运行，此时，

因此，

当时，两台发电机运行，此时，

因此．由此得的分布列如下：

|  |  |  |
| --- | --- | --- |
|  | 4200 | 10000 |
|  | 0．2 | 0．8 |

所以．

③安装3台发电机．

依题意，当时，一台发电机运行，此时，

因此；

当时，两台发电机运行，此时，

此时，

当时，三台发电机运行，此时，

因此，

由此得的分布列如下：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 34 | 9200 | 15000 |
|  | 0．2 | 0．8 | 0．1 |

所以．

综上，欲使水电站年总利润的均值达到最大，应安装发电机2台．

15．(2014高考数学江苏·第25题)盒中共有9个球，其中有4个红球，3个黄球和2个绿球，这些球除颜色外完全相同．

(1)从盒中一次随机抽出2个球，求取出的2个球颜色相同的概率*P*；

(2)从盒中一次随机抽出4个球，其中红球、黄球、绿球的个数分别为*x*1，*x*2，*x*3，随机变量*X*表示*x*1，*x*2，*x*3的最大数，求*X*的概率分布和数学期望*E*(*X*)．

**【答案】**解析：(1)取出的2个颜色相同的球可能是2个红球、2个黄球或2个绿球，

所以

(2)随机变量的所有可能的取值为．

表示的随机事件是取到的4个球是4个红球，故；

表示的随机事件是取到的4个球是3个红球和1个其它颜色的球，或3个黄球和1个

其它颜色的球，故；

于是

所以随机变量的概率分布如下表：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 2 | 3 | 4 |
|  |  |  |  |

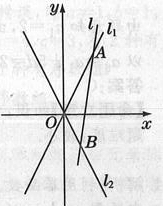
因此随机变量的数学期望

16．(2014高考数学福建理科·第18题)(本小题满分13分)

为回馈顾客，某商场拟通过摸球兑奖的方式对位顾客进行奖励． 规定：每位顾客从一个装有个标有面值的球的袋中一次性随机摸出２个球，球上所标的面值之和为该顾客所获的奖励额．

(1)若袋中所装的个球中有个所标的面值为元，其余个均为元，求①顾客所获的奖励额为的概率；②顾客所获的奖励额的分布列及数学期望；

(2)商场对奖励总额的预算是元，并规定袋中的个球只能由标有面值为元和元的两种球组成，或标有面值元和元的两种球组成． 为了使顾客得到的奖励总额尽可能符合商场的预算且每位顾客所获的奖励额相对均衡，请对袋中的个球的面值给出一个合适的设计，并说明理由．



**【答案】**解析：(I)设顾客获得的奖励额为．

①依题意，得，即顾客所获的奖励额为60元的概率是．

②依题意，随机变量的可能取值为．

．

得的分布列如下：

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

所以顾客所获的奖励额的期望为

(II)根据商场的预算，每个顾客的平均奖励额为元．所以，先寻找期望为60元的可能方案：

当球的面值为元和元时，若选择方案，因为元是面值之和的最大值，所以期望不可能为；若选择方案，因为元是面值之和的最小值，所以期望也不可能是．因此可能的方案是，记为方案．

当球的面值为元和元时，同理可排除、的方案，所以可能的方案是，记为方案．

以下是对两个方案的分析：

对于方案，即方案，设顾客所获的奖励额为，的可能取值为．



得的分布列如下：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

的期望为

的方差为

对于方案，即方案，设顾客所获得奖励额为，的可能取值为．

得的分布列如下：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

的期望为

的方差为

由于两种方案奖励额的期望都符合要求，但方案奖励额的方差要比方案的小，所以应该选择方案．即标有面值元和面值元的球各两个．

17．(2014高考数学大纲理科·第20题)设每个工作日甲、乙、丙、丁4人需使用某种设备的概率分别为0．6,0．5,,0．5,0．4，各人是否需使用设备相互独立．

(1)求同一工作日至少3人需使用设备的概率；

(2)X表示同一工作日需使用设备的人数，求X的数学期望．

**【答案】**(1)；(2)2

解析：记表示事件：同一工作日乙、丙中恰有人需使用设备，

表示事件：甲需使用设备

表示事件：丁需使用设备

表示事件：同一工作日至少3人需使用设备

(1)



所以



(2)的可能取值为，，，，

．



，

，

，



所以的分布列为

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 |
|  |  |  |  |  |  |

数学期望



．

18．(2014高考数学北京理科·第16题)李明在10场篮球比赛中的投篮情况如下(假设各场比赛相互独立)：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 场次 | 投篮次数 | 命中次数 | 场次 | 投篮次数 | 命中次数 |
| 主场1 | 22 | 12 | 客场1 | 18 | 8 |
| 主场2 | 15 | 12 | 客场2 | 13 | 12 |
| 主场3 | 12 | 8 | 客场3 | 21 | 7 |
| 主场4 | 23 | 8 | 客场4 | 18 | 15 |
| 主场5 | 24 | 20 | 客场5 | 25 | 12 |

(1)从上述比赛中随机选择一场, 求李明在该场比赛中投篮命中率超过0．6的概率．

(2)从上述比赛中随机选择一个主场和一个客场, 求李明的投篮命中率一场超过0．6, 一场不超过0．6的 概率．

(3)记是表中10个命中次数的平均数, 从上述比赛中随机选择一场, 记*X*为李明在这比赛中的命中次 数, 比较*E*(*X*)与 的大小(只需写出结论)

**【答案】**解析：(Ⅰ)根据投篮统计数据,在10场比赛中,李明投篮命中率超过0．6的场次有5场,

分别是主场2,主场3,主场5,客场2,客场4．

所以在随机选择的一场比赛中,李明的投篮命中率超过0．6的概率是0．5．

(Ⅱ)设事件为“在随机选择的一场主场比赛中李明的投篮命中率超过0．6”,

事件为“在随机选择的一场客场比赛中李明的投篮命中率超过0．6”,

事件为“在随机选择的一个主场和一个客场中,李明的投篮命中率一场超过0．6,一场不超过0．6”．

则独立．

根据投篮统计数据,







所以,在随机选择的一个主场和一个客场中,李明的投篮命中率一场超过,一场不超过0．6的概率为

(Ⅲ)

19．(2014高考数学安徽理科·第17题)甲乙两人进行围棋比赛，约定先连胜两局者直接赢得比赛，若赛完5局仍未出现连胜，则判定获胜局数多者赢得比赛． 假设每局甲获胜的概率为，乙获胜的概率为，各局比赛结果相互独立．

(Ⅰ)求甲在4局以内(含4局)赢得比赛的概率；

(Ⅱ)记X为比赛决出胜负时的总局数，求X的分布列和均值(数学期望)．

**【答案】**解：用表示“甲在4局以内(含4局)赢得比赛”，表示“第局甲获胜”，表示“第局乙获胜”，则，，．

(Ⅰ)



．

(Ⅱ)的可能取值是2,3,4,5．

，．

故的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |

．

20．(2015高考数学重庆理科·第17题)(本小题满分13分，(1)小问5分，(2)小问8分)

端午节吃粽子是我国的传统习俗．设一盘中装有10个粽子，其中豆沙粽2个，肉粽3个，白粽5个，这三种粽子的外观完全相同．从中任意选取3个。

(1)求三种粽子各取到1个的概率；

(2)设表示取到的豆沙粽个数，求的分布列与数学期望．

**【答案】**(1)；(2)分布列见解析，期望为．

分析：(1)本题属于古典概型，从10个棕子中任取3个，基本事件的总数为，其中事件“三种棕子各取1个”含基本事件的个数为，根据古典概型概率计算公式可计算得所求概率；(2)由于10个棕子中有2个豆沙棕，因此的可能值分别为，同样根据古典概型概率公式可得相应的概率，从而列出其分布列，并根据期望公式求得期望为．

解析：(1)令A表示事件“三个粽子各取到1个”，则由古典概型的概率计算公式有

；

(2)X的所有可能取值为0，1，2，且



综上知，X的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| X | 0 | 1 | 2 |
| P |  |  |  |

故(个)．

21．(2015高考数学天津理科·第16题)(本小题满分13分)为推动乒乓球运动的发展，某乒乓球比赛允许不同协会的运动员组队参加．现有来自甲协会的运动员3名，其中种子选手2名；乙协会的运动员5名，其中种子选手3名．从这8名运动员中随机选择4人参加比赛．

(Ⅰ)设为事件“选出的4人中恰有2名种子选手，且这2名种子选手来自同一个协会”求事件发生的概率；

(Ⅱ)设为选出的4人中种子选手的人数，求随机变量的分布列和数学期望．

**【答案】**(Ⅰ);

(Ⅱ)随机变量的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |



解析：(Ⅰ)由已知，有 所以事件发生的概率为．

(Ⅱ)随机变量的所有可能取值为



所以随机变量的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |

所以随机变量的数学期望

22．(2015高考数学四川理科·第17题)某市两所中学的学生组队参加辩论赛，中学推荐3名男生，2名女生，中学推荐了3名男生，4名女生，两校推荐的学生一起参加集训，由于集训后队员的水平相当，从参加集训的男生中随机抽取3人，女生中随机抽取3人组成代表队

(1)求A中学至少有1名学生入选代表队的概率．

(2)某场比赛前，从代表队的6名队员中随机抽取4人参赛，设表示参赛的男生人数，求得分布列和数学期望．

**【答案】**(1)A中学至少1名学生入选的概率为．

(2)X的分布列为：



X的期望为．

解析：(1)由题意，参加集训的男女生各有6名．

参赛学生全从B中抽取(等价于A中没有学生入选代表队)的概率为．

因此，A中学至少1名学生入选的概率为．

(2)根据题意，X的可能取值为1，2，3．

，

，

，

所以X的分布列为：



因此，X的期望为．

23．(2015高考数学陕西理科·第19题)(本小题满分12分)设某校新、老校区之间开车单程所需时间为，只与道路畅通状况有关，对其容量为的样本进行统计，结果如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| (分钟) | 25 | 30 | 35 | 40 |
| 频数(次) | 20 | 30 | 40 | 10 |

(Ⅰ)求的分布列与数学期望；

(Ⅱ)刘教授驾车从老校区出发，前往新校区做一个50分钟的讲座，结束后立即返回老校区，求刘教授从离开老校区到返回老校区共用时间不超过120分钟的概率．

**【答案】**(Ⅰ)分布列见解析，；(Ⅱ)．

分析：(Ⅰ)先算出的频率分布，进而可得的分布列，再利用数学期望公式可得数学期望；(Ⅱ)先设事件表示“刘教授从离开老校区到返回老校区共用时间不超过分钟”，再算出的概率．

解析：(Ⅰ)由统计结果可得的频率分步为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| (分钟) | 25 | 30 | 35 | 40 |
| 频率 | 0．2 | 0．3 | 0．4 | 0．1 |

以频率估计概率得的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 25 | 30 | 35 | 40 |
|  | 0．2 | 0．3 | 0．4 | 0．1 |

从而 (分钟)

(Ⅱ)设分别表示往、返所需时间，的取值相互独立，且与的分布列相同．设事件表示“刘教授共用时间不超过分钟”，由于讲座时间为分钟，所以事件对应于“刘教授在途中的时间不超过分钟”．

解法一：

．

解法二：



故．

24．(2015高考数学山东理科·第19题)若是一个三位正整数，且的个位数字大于十位数字，十位数字大于百位数字，则称为“三位递增数”(如等)．

在某次数学趣味活动中，每位参加者需从所有的“三位递增数”中随机抽取个数，且只能抽取一次．得分规则如下：若抽取的“三位递增数”的三个数字之积不能被整除，参加者得分；若能被整除，但不能被整除，得分；若能被整除，得分．

(I)写出所有个位数字是的“三位递增数” ；

(II)若甲参加活动，求甲得分的分布列和数学期望．

**【答案】**(Ⅰ)有：125，135，145，235，245，345；

(Ⅱ)X的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| X | 0 | -1 | 1 |
| P |  |  |  |



分析：(Ⅰ)明确“三位递增数”的含义，写出所有的三位符合条件的“三位递增数”；(Ⅱ)

解析：明确随机变量的所有可能取值及取每一个值的含义，结合组合的知识，利用古典概型求出的分布列和数学期望．

(Ⅰ)个位数是5的“三位递增数”有：125，135，145，235，245，345；

(Ⅱ)由题意知，全部“三位递增烽”的个数为

随机变量X的取值为：0，-1，1，因此

  , ,

所以X的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| X | 0 | -1 | 1 |
| P |  |  |  |

因此

25．(2015高考数学湖南理科·第20题)某商场举行有奖促销活动，顾客购买一定金额商品后即可抽奖，每次抽奖都从装有4个红球、6个白球的甲箱和装有5个红球、5个白球的乙箱中，各随机摸出1个球，在摸出的2个球中，若都是红球，则获一等奖；若只有1个红球，则获二等奖；若没有红球，则不获奖．

(1)求顾客抽奖1次能获奖的概率；

(2)若某顾客有3次抽奖机会，记该顾客在3次抽奖中获一等奖的次数为，求的分布列和数学期望．

**【答案】**(1)；(2)详见解析．

分析：(1)记事件｛从甲箱中摸出的1个球是红球｝，｛从乙箱中摸出的1个球是红球｝

｛顾客抽奖1次获一等奖｝，｛顾客抽奖1次获二等奖｝，｛顾客抽奖1次能获奖｝，则可知

与相互独立，与互斥，与互斥，且，，，再

利用概率的加法公式即可求解；(2)分析题意可知，分别求得，，，，即可知的概率分布及其期望．

解析：(1)记事件｛从甲箱中摸出的1个球是红球｝，｛从乙箱中摸出的1个球是红球｝

｛顾客抽奖1次获一等奖｝，｛顾客抽奖1次获二等奖｝，｛顾客抽奖1次能获奖｝，由题意，与相互独立，与互斥，与互斥，且，，，

∵，，∴，



，故所求概率为；(2)顾客抽奖3次独立重复试验，由(1)知，顾客抽奖1次获一等奖的概率为，∴，

于是，，，

，故的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 |
|  |  |  |  |  |

的数学期望为 ．

26．(2015高考数学湖北理科·第20题)(本小题满分12分)某厂用鲜牛奶在某台设备上生产两种奶制品．生产1吨产品需鲜牛奶2吨，使用设备1小时，获利1000元；生产1吨产品需鲜牛奶1．5吨，使用设备1．5小时，获利1200元．要求每天产品的产量不超过产品产量的2倍，设备每天生产两种产品时间之和不超过12小时．假定每天可获取的鲜牛奶数量W(单位：吨)是一个随机变量，其分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| W | 12 | 15 | 18 |
| P | 0．3 | 0．5 | 0．2 |

该厂每天根据获取的鲜牛奶数量安排生产，使其获利最大，因此每天的最大获利(单位：元)是一个随机变量．

(Ⅰ)求的分布列和均值；

(Ⅱ)若每天可获取的鲜牛奶数量相互独立，求3天中至少有1天的最大获利超过10000元的概率．

**【答案】**(Ⅰ)的分布列为：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 8160 | 10200 | 10800 |
|  | 0．3 | 0．5 | 0．2 |

；(Ⅱ)0．973．

解析：(Ⅰ)设每天两种产品的生产数量分别为，相应的获利为，

则有 (1)

第20题解答图1

第20题解答图2

第20题解答图3



目标函数为 ．

当时，(1)表示的平面区域如图1，三个顶点分别为．

将变形为，

当时，直线：在轴上的截距最大，

最大获利．

当时，(1)表示的平面区域如图2，三个顶点分别为．

将变形为，

当时，直线：在轴上的截距最大，

最大获利．

当时，(1)表示的平面区域如图3，

四个顶点分别为．

将变形为，

当时，直线：在轴上的截距最大，

最大获利．

故最大获利的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
|  | 8160 | 10200 | 10800 |
|  | 0．3 | 0．5 | 0．2 |

因此，

(Ⅱ)由(Ⅰ)知，一天最大获利超过10000元的概率，

由二项分布，3天中至少有1天最大获利超过10000元的概率为

27．(2015高考数学福建理科·第16题)某银行规定，一张银行卡若在一天内出现3次密码尝试错误，该银行卡将被锁定，小王到银行取钱时，发现自己忘记了银行卡的密码，但是可以确定该银行卡的正确密码是他常用的6个密码之一，小王决定从中不重复地随机选择1个进行尝试．若密码正确，则结束尝试；否则继续尝试，直至该银行卡被锁定．

(Ⅰ)求当天小王的该银行卡被锁定的概率；

(Ⅱ)设当天小王用该银行卡尝试密码次数为X，求X的分布列和数学期望．

**【答案】**(Ⅰ)；(Ⅱ)分布列见解析，期望为．

解析：(Ⅰ)设“当天小王的该银行卡被锁定”的事件为A，

则

(Ⅱ)依题意得，X所有可能的取值是1，2，3

又

所以X的分布列为



所以．

28．(2015高考数学北京理科·第16题)(本小题13分)，两组各有7位病人，他们服用某种药物后的康复时间(单位：天)记录如下：

组：10，11，12，13，14，15，16

组：12，13，15，16，17，14，

假设所有病人的康复时间互相独立，从，两组随机各选1人，组选出的人记为甲，组选出的人记为乙．

(Ⅰ)求甲的康复时间不少于14天的概率；

(Ⅱ)如果，求甲的康复时间比乙的康复时间长的概率；

(Ⅲ)当为何值时，，两组病人康复时间的方差相等？(结论不要求证明)

**【答案】**(Ⅰ)，(Ⅱ)，(Ⅲ)或

分析：针对甲有7种情况，康复时间不少于14天有3种情况，概率为；如果，甲、乙随机各取一人有49种情况，用列举法列出甲的康复时间比乙的康复时间长的情况有10种，概率为，由于A组数据为10，11，12，13，14，15，16；B组数据调整为，12，13，14，15，16，17，或12，13，14，15，16，17，，由于，两组病人康复时间的方差相等，即波动相同，所以或．

解析：(Ⅰ)甲有7种取法，康复时间不少于14天的有3种取法，所以概率；

(Ⅱ)如果，从，两组随机各选1人，组选出的人记为甲，组选出的人记为乙共有49种取法，甲的康复时间比乙的康复时间长的列举如下：(13，12)，(14，12)，(14，13)，(15，12)，(15，13)，(15,14),(16，12)(16，13),(16，15),(16,14)有10种取法，所以概率．

(Ⅲ)把B组数据调整为，12，13，14，15，16，17，或12，13，14，15，16，17，，可见当或时，与A组数据方差相等．(可利用方差公式加以证明，但本题不需要)

29．(2015高考数学安徽理科·第17题)(本小题满分12分)已知2件次品和3件正品混放在一起，现需要通过检测将其区分，每次随机检测一件产品，检测后不放回，直到检测出2件次品或者检测出3件正品时检测结束．

(Ⅰ)求第一次检测出的是次品且第二次检测出的是正品的概率；

(Ⅱ)已知每检测一件产品需要费用100元，设X表示直到检测出2件次品或者检测出3件正品时所需要的检测费用(单位：元)，求X的分布列和均值(数学期望)．

**【答案】**(Ⅰ)；(Ⅱ)．

分析：(Ⅰ)依据题目所给的条件可以先设“第一次检查出的是次品且第二次检测出的是正品”为事件．得出．(Ⅱ)的可能取值为．依此求出各自的概率，列出分布列，求出期望．

解析：(Ⅰ)记“第一次检查出的是次品且第二次检测出的是正品”为事件．

．

(Ⅱ)的可能取值为．

．

．

．

故的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

．

30．(2017年高考数学新课标Ⅰ卷理科·第19题)(12分)为了监控某种零件的一条生产线的生产过程，检验员每天从该生产线上随机抽取16个零件，并测量其尺寸(单位：)．根据长期生产经验，可以认为这条生产线正常状态下生产的零件的尺寸服从正态分布学科网 版权所有．

(1)假设生产状态正常，记表示一天内抽取的16个零件中其尺寸在学科网 版权所有之外的零件数，求学科网 版权所有及学科网 版权所有的数学期望；

(2)一天内抽检零件中，如果出现了尺寸在学科网 版权所有之外的零件，就认为这条生产线在这一天的生产过程可能出现了异常情况，需对当天的生产过程进行检查．

(ⅰ)试说明上述监控生产过程方法的合理性；

(ⅱ)下面是检验员在一天内抽取的16个零件的尺寸：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 9．95 | 10．12 | 9．96 | 9．96 | 10．01 | 9．92 | 9．98 | 10．04 |
| 10．26 | 9．91 | 10．13 | 10．02 | 9．22 | 10．04 | 10．05 | 9．95 |

经计算得学科网 版权所有，学科网 版权所有，其中学科网 版权所有为抽取的第学科网 版权所有个零件的尺寸，学科网 版权所有．

用样本平均数学科网 版权所有作为学科网 版权所有的估计值学科网 版权所有，用样本标准差学科网 版权所有作为学科网 版权所有的估计值学科网 版权所有，利用估计值判断是否需对当天的生产过程进行检查？剔除学科网 版权所有之外的数据，用剩下的数据估计学科网 版权所有和学科网 版权所有(精确到0．01)．

附：若随机变量学科网 版权所有服从正态分布学科网 版权所有，则学科网 版权所有，学科网 版权所有，学科网 版权所有．

**【答案】**(1),;(2)详见解析． 【分析】(1)根据题设条件知一个零件尺寸在之内的概率为,则零件的尺寸在之外的概率为,而,进而可以求出的数学期望．(2)(i)判断监控生产过程的方法的合理性,重点是考虑一天内抽取的个零件中,出现尺寸在之外的零件的概率大还是小,若小即合理;(ii)根据题设条件题出的估计值和的估计值,剔除之外的数据,算出剩下数据的平均数,即为的估计值,剔除之外的数据,剩下数据的样本方法,即为的估计值． 【解析】(1)抽取的一个零件的尺寸在之内的概率为0．9974,从而零件的尺寸在之外的概率为0．0026 故．因此． 的数学期望为． (2)(i)如果生产状态正常,一个零件尺寸在之外的概率只有,一天内抽取的16个零件中,出现尺寸在之外的零件的概率只有0．0408,发生的概率很小．因此一旦发生这种情况,就有理由认为这条生产线在这一天的生产过程可能出现了异常情况,需对当天的生产过程进行检查,可见上述监控生产过程的方法是合理的． (ii)由,得的估计值为,的估计值为,由样本数据可以看出有一个零件的尺寸在之外,因此需对当天的生产过程进行检查． 剔除之外的数据9．22,剩下数据的平均数为 因此的估计值为  剔除之外的数据,剩下数据的样本方差为,因此的估计值为．

31．(2017年高考数学天津理科·第16题)从甲地到乙地要经过个十字路口,设各路口信号灯工作相互独立,且在各路口遇到红灯的概率分别为．

(1)设表示一辆车从甲地到乙地遇到红灯的个数,求随机变量的分布列和数学期望;

(2)若有辆车独立地从甲地到乙地,求这辆车共遇到个红灯的概率．

**【答案】** 【解析】(1)随机变量学科网 版权所有的所有可能取值为． 学科网 版权所有, 学科网 版权所有, 学科网 版权所有, 学科网 版权所有 所以,随机变量学科网 版权所有的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 学科网 版权所有 |  |  |  |  |
| 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |

随机变量学科网 版权所有的数学期望学科网 版权所有． (2)设学科网 版权所有表示第一辆车遇到红灯的个数,学科网 版权所有表示第二辆车遇到红灯的个数,则所求事件的概率为学科网 版权所有 学科网 版权所有． 所以,这辆车共遇到个红灯的概率为学科网 版权所有．

32．(2017年高考数学山东理科·第18题)在心理学研究中,常采用对比试验的方法评价不同心理暗示对人的影响,具体方法如下:将参加试验的志愿者随机分成两组,一组接受甲种心理暗示,另一组接受乙中心理暗示,通过对比这两组志愿者接受心理暗示后的结果来评价两种心理暗示的作用,现有名男志愿者和名,从中随机抽取人接受甲种心理暗示,另人接受乙种心理暗示．

(I)求接受甲种心理暗示的志愿者中包含但不包含的频率．

(II)用表示接受乙种心理暗示的女志愿者人数,求的分布列与数学期望．

**【答案】**(I)说明: 学科网 版权所有(II)的分布列为

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 |
|  | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 |

的数学期望是说明: 学科网 版权所有． 【分析】(1)记接受甲种心理暗示的志愿者中包含但不包含的事件为,计算即得;(2)由题意知可以取的值为:0,1,2,3,4,利用超几何分布概率计算公式,得的分布列为

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 |
|  | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 | 说明: 学科网 版权所有 |

进一步计算的数学期望． 【解析】(I)记接受甲种心理暗示的志愿者中包含学科网 版权所有但不包含学科网 版权所有的事件为,则学科网 版权所有 (II)由题意知可取的值为:0,1,2,3,4,则 学科网 版权所有 学科网 版权所有 学科网 版权所有 学科网 版权所有 学科网 版权所有 因此的分布列为

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 |
|  | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |

的数学期望是学科网 版权所有 学科网 版权所有

33．(2017年高考数学课标Ⅲ卷理科·第18题)某超市计划按月订购一种酸奶,每天进货量相同,进货成本每瓶4元,售价每瓶6元,未售出的酸奶降价处理,以每瓶2元的价格当天全部处理完．根据往年销售经验,每天需求量与当天最高气温(单位:℃)有关．如果最高气温不低于25,需求量为500瓶;如果最高气温位于区间[20,25),需求量为300瓶;如果最高气温低于20,需求量为200瓶．为了确定六月份的订购计划,统计了前三年六月份各天的最高气温数据,得下面的频数分布表:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 最高气温 | [10,15) | [15,20) | [20,25) | [25,30) | [30,35) | [35,40) |
| 天数 | 2 | 16 | 36 | 25 | 7 | 4 |

以最高气温位于各区间的频率代替最高气温位于该区间的概率．

(Ⅰ)求六月份这种酸奶一天的需求量*X*(单位:瓶)的分布列;

(Ⅱ)设六月份一天销售这种酸奶的利润为*Y*(单位:元)．当六月份这种酸奶一天的进货量*n*(单位:瓶)为多少时,*Y*的数学期望达到最大值?

**【答案】**(Ⅰ)分布列略;(Ⅱ)*n*=300时,*Y*的数学期望达到最大值,最大值为520元． 【解析】(1)依题意可知的所有可能取值为 其中,, 所以的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

(2)①当时:,此时,当时取到． ②当学科网 版权所有时: 若学科网 版权所有,则学科网 版权所有, 若学科网 版权所有时,则学科网 版权所有 若学科网 版权所有时,则学科网 版权所有 学科网 版权所有的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |
| 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |

∴学科网 版权所有 此时,当时取到． ③当学科网 版权所有时,若学科网 版权所有,则学科网 版权所有 若学科网 版权所有时,则学科网 版权所有 若学科网 版权所有时,则学科网 版权所有 学科网 版权所有的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |
| 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |

∴学科网 版权所有(元)④当时,易知一定小于③的情况． 综上,当学科网 版权所有为学科网 版权所有瓶时,学科网 版权所有的数学期望达到最大值．

34．(2017年高考数学江苏文理科·第26题)已知一个口袋有学科网 版权所有个白球,学科网 版权所有个黑球(学科网 版权所有),这些球除颜色外全部相同．现将口袋中的球随机的逐个取出,并放入如图所示的编号为学科网 版权所有的抽屉内,其中第学科网 版权所有次取出的球放入编号为学科网 版权所有的抽屉学科网 版权所有．

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 学科网 版权所有 | 学科网 版权所有 |

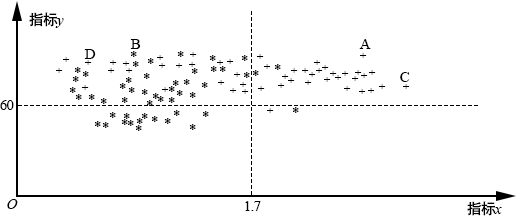
(1)试求编号为2的抽屉内放的是黑球的概率学科网 版权所有;(2)随机变量学科网 版权所有表示最后一个取出的黑球所在抽屉编号的倒数,学科网 版权所有是学科网 版权所有的数学期望,证明:学科网 版权所有

**【答案】**(1)学科网 版权所有(2)见解析 解析:解:(1)编号为2的抽屉内放的是黑球的概率学科网 版权所有为: 学科网 版权所有．  (2)随机变量 *X*的概率分布为:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *X* | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |  | 学科网 版权所有 |  | 学科网 版权所有 |
| *P* | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |  | 学科网 版权所有 |  | 学科网 版权所有 |

随机变量 *X*的期望为: 学科网 版权所有, 所以学科网 版权所有 学科网 版权所有 学科网 版权所有 学科网 版权所有 学科网 版权所有 学科网 版权所有 所以学科网 版权所有．

35．(2017年高考数学北京理科·第17题)为了研究一种新药的疗效,选名患者随机分成两组,每组各名,一组服药,另一组不服药．一段时间后,记录了两组患者的生理指标和的数据,并制成下图,其中“\*”表示服药者,“+”表示未服药者．



(Ⅰ)从服药的名患者中随机选出一人,求此人指标的值小于的概率;

(Ⅱ)从图中四人中随机选出两人,记学科网 版权所有为选出的两人中指标的值大于的人数,求学科网 版权所有的分布列和数学期望学科网 版权所有;

(Ⅲ)试判断这名患者中服药者指标数据的方差与未服药者指标数据的方差的大小．(只需写出结论)

**【答案】**(Ⅰ) ;(Ⅱ)详见解析;(Ⅲ)在这名患者中,服药者指标数据的方差大于未服药者指标数据的方差． 【解析】(Ⅰ)根据所给数据数出的个数,再除以就是概率;(Ⅱ)由图可知两人,根据超几何分布写出分布列,,,并求数学期望;(Ⅲ)方差表示数据的离散程度,波动越大,方差越大,波动越小,方差越小． 解:(Ⅰ)由图知,在服药的患者中,指标的值小于的有人,所以从服药的名患者中随机选出一人,此人指标的值小于的概率为． (Ⅱ)由图知,四人中,指标学科网 版权所有的值大于的有人:和． 所以学科网 版权所有的所有可能取值为． 学科网 版权所有． 所以学科网 版权所有的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| 学科网 版权所有 |  |  |  |
| 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 | 学科网 版权所有 |

故学科网 版权所有的期望学科网 版权所有． (Ⅲ)在这名患者中,服药者指标学科网 版权所有数据的方差大于未服药者指标学科网 版权所有数据的方差．

36．(2016高考数学天津理科·第16题)某小组共10人，利用假期参加义工活动．已知参加义工活动次数为1,2,3的人数分别为3,3,4．现从这10人中随机选出2人作为该组代表参加座谈会．

(Ⅰ)设为事件“选出的2人参加义工活动次数之和为4”，求事件发生的概率；

(Ⅱ)设为选出的2人参加义工活动次数之差的绝对值，求随机变量的分布列和数学期望．

**【答案】**(Ⅰ) (Ⅱ)分布列详见解析，*X*的数学期望为1

解析：(Ⅰ)设事件：选2人参加义工活动，次数之和为4



(Ⅱ)随机变量可能取值：0，1，2



|  |  |  |  |
| --- | --- | --- | --- |
|  | 0 | 1 | 2 |
|  |  |  |  |

．

37．(2016高考数学山东理科·第19题)(本小题满分12分)甲、乙两人组成“星队”参加猜成语活动，每轮活动由甲、乙各猜一个成语，在一轮活动中，如果两人都猜对，则“星队”得3分；如果只有一个人猜对，则“星队”得1分；如果两人都没猜对，则“星队”得0分．已知甲每轮猜对的概率是，乙每轮猜对的概率是；每轮活动中甲、乙猜对与否互不影响，各轮结果亦互不影响．假设“星队”参加两轮活动，求：

(I)“星队”至少猜对3个成语的概率；

(Ⅱ)“星队”两轮得分之和为的分布列和数学期望．

**【答案】**【解析】(Ⅰ)记事件A:“甲第一轮猜对”，记事件B：“乙第一轮猜对”，]

记事件C：“甲第二轮猜对”，记事件D：“乙第二轮猜对”，

记事件E：“‘星队’至少猜对3个成语”．

由题意，

由事件的独立性与互斥性，





 ,

所以“星队”至少猜对3个成语的概率为．

(Ⅱ)由题意，随机变量的可能取值为．

由事件的独立性与互斥性，得

 ,,

 ,

 ,

,．

可得随机变量的分布列为

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 6 |
|  |  |  |  |  |  |  |

所以数学期望．

38．(2016高考数学课标Ⅱ卷理科·第18题)(本题满分12分)某险种的基本保费为(单位：元)，继续购买该险种的投保人称为续保人，续保人的本年度的保费与其上年度的出险次数的关联如下：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 上年度出险次数 | 0 | 1 | 2 | 3 | 4 |  |
| 保费 |  |  |  |  |  |  |

设该险种一续保人一年内出险次数与相应概率如下：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 一年内出险次数 | 0 | 1 | 2 | 3 | 4 |  |
| 概率 | 0．30 | 0．15 | 0．20 | 0．20 | 0．10 | 0． 05 |

(I)求一续保人本年度的保费高于基本保费的概率；

(II)若一续保人本年度的保费高于基本保费，求其保费比基本保费高出的概率；

(III)求续保人本年度的平均保费与基本保费的比值．

**【答案】**(1)；(2)；(3)

【解析】(I)设表示事件：“一续保人本年度的保费高于基本保费”，则事件发生当且仅当一年内出险次数大于，故．

(II)设表示事件：“一续保人本年度的保费比基本保费高出”，则事件发生当且仅当一年内出险次数大于，故，

又

因此所求概率为．

(III)记续保人本年度的保费为，则的分布列为

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |



因此续保人本年度的平均保费与基本保费的比值为：．

# 题型四：概率统计中的决策建议

1．(2023年全国乙卷理科·第17题)某厂为比较甲乙两种工艺对橡胶产品伸缩率的处理效应，进行10次配对试验，每次配对试验选用材质相同的两个橡胶产品，随机地选其中一个用甲工艺处理，另一个用乙工艺处理，测量处理后的橡胶产品的伸缩率．甲、乙两种工艺处理后的橡胶产品的伸缩率分别记为，．试验结果如下：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 试验序号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 伸缩率 | 545 | 533 | 551 | 522 | 575 | 544 | 541 | 568 | 596 | 548 |
| 伸缩率 | 536 | 527 | 543 | 530 | 560 | 533 | 522 | 550 | 576 | 536 |

记，记样本平均数为，样本方差为．

(1)求，；

(2)判断甲工艺处理后的橡胶产品的伸缩率较乙工艺处理后的橡胶产品的伸缩率是否有显著提高(如果，则认为甲工艺处理后的橡胶产品的伸缩率较乙工艺处理后的橡胶产品的伸缩率有显著提高，否则不认为有显著提高)

**【答案】**(1)，；

(2)认为甲工艺处理后橡胶产品的伸缩率较乙工艺处理后的橡胶产品的伸缩率有显著提高．

解析：(1)，

，

，

 的值分别为: ，

故

(2)由(1)知:，，故有,

所以认为甲工艺处理后的橡胶产品的伸缩率较乙工艺处理后的橡胶产品的伸缩率有显著提高．

2．(2021年高考全国乙卷理科·第17题)某厂研制了一种生产高精产品的设备，为检验新设备生产产品的某项指标有无提高，用一台旧设备和一台新设备各生产了10件产品，得到各件产品该项指标数据如下：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 旧设备 | 9．8 | 10．3 | 100 | 102 | 9．9 | 9．8 | 10．0 | 10．1 | 10．2 | 9．7 |
| 新设备 | 101 | 10．4 | 10．1 | 10．0 | 10．1 | 10．3 | 10．6 | 10．5 | 10．4 | 10．5 |

旧设备和新设备生产产品的该项指标的样本平均数分别记为和，样本方差分别记为和．

(1)求，，，；

(2)判断新设备生产产品的该项指标的均值较旧设备是否有显著提高(如果，则认为新设备生产产品的该项指标的均值较旧设备有显著提高，否则不认为有显著提高)．

**【答案】**(1)；(2)新设备生产产品的该项指标的均值较旧设备有显著提高．

解析：(1)，

，

，

．

(2)依题意，，，

，所以新设备生产产品的该项指标的均值较旧设备有显著提高．

3．(2018年高考数学课标Ⅱ卷(理)·第18题)(12分)下图是某地区2000年至2016年环境基础设施投资额(单位：亿元)的折线图．

18

为了预测该地区2018年的环境基础设施投资额，建立了与时间变量的两个线性回归模型．根据2000年至2016年的数据(时间变量的值依次为)建立模型①：；根据2010年至2016年的数据(时间变量的值依次为)建立模型②：．

(1)分别利用这两个模型，求该地区2018年的环境基础设施投资额的预测值；

(2)你认为用哪个模型得到的预测值更可靠？并说明理由．

**【答案】**解析：(1)利用模型①，该地区2018年的环境基础设施投资额的预测值为

(亿元)．

利用模型②，该地区2018年的环境基础设施投资额的预测值为

(亿元)．

(2)利用模型②得到的预测值更可靠．

理由如下：

(i)从折线图可以看出，2000年至2016年的数据对应的点没有随机散布在直线上下，这说明利用2000年至2016年的数据建立的线性模型①不能很好地描述环境基础设施投资额，的变化趋势．2010年至2016年的数据对应的点位于一条直线的附近，这说明从2010年开始环境基础设施投资额的变化规律呈线性增长趋势，，利用2010年至2016年的数据建立的线性模型可以较好地描述2010年以后的环境基础设施投资额的变化趋势，因此利用模型②得到的预测值更可靠．

(ii)从计算结果看，相对于2016年的环境基础设施投资额220亿元，由模型①得到的预测值226．1亿元的增幅明显偏低，而利用模型②得到的预测值的增幅比较合理，说明利用模型②得到的预测值更可靠．

以上给出了2种理由，考生答出其中一种或其他合理理由均可得分．

4．(2018年高考数学课标卷Ⅰ(理)·第20题)(12分)某工厂的某种产品成箱包装，每箱200件，每一箱产品在交付用户之前要对产品作检验，如检验出不合格品，则更换为合格品．检验时，先从这箱产品中任取20件作检验，再根据检验结果决定是否对余下的所有产品作检验，设每件产品为不合格品的概率都为，且各件产品是否为不合格品相互独立．

(1)记20件产品中恰有2件不合格品的概率为,求的最大值点．

(2)现对一箱产品检验了20件，结果恰有2件不合格品，以(1)中确定的作为的值．已知每件产品的检验费用为2元，若有不合格品进入用户手中，则工厂要对每件不合格品支付25元的赔偿费用．

(i)若不对该箱余下的产品作检验，这一箱产品的检验费用与赔偿费用的和记为，求;

(ii)以检验费用与赔偿费用和的期望值为决策依据，是否该对这箱余下的所有产品作检验？

**【答案】**解析：(1)20件产品中恰有2件不合格品的概率为．

因此．

令，得．当时，；当时，．

所以的最大值点为．

(2)由(1)知，．

(i)令表示余下的180件产品中的不合格品件数，依题意知，，即．

所以．

(ii)如果对余下的产品作检验，则这一箱产品所需要的检验费为400元．

由于，故应该对余下的产品作检验．

5．(2021年新高考Ⅰ卷·第18题)某学校组织“一带一路”知识竞赛，有*A*，*B*两类问题，每位参加比赛的同学先在两类问题中选择一类并从中随机抽取一个问题回答，若回答错误则该同学比赛结束：若回答正确则从另一类问题中再随机抽取一个问题回答，无论回答正确与否，该同学比赛结束．*A*类问题中的每个问题回答正确得20分，否则得0分：*B*类问题中的每个问题回答正确得80分，否则得0分，己知小明能正确回答*A*类问题的概率为0．8，能正确回答*B*类问题的概率为0．6，且能正确回答问题的概率与回答次序无关．

(1)若小明先回答*A*类问题，记为小明的累计得分，求的分布列；

(2)为使累计得分期望最大，小明应选择先回答哪类问题？并说明理由．

**【答案】**解析：(1)由题可知，的所有可能取值为，，．

；

；

．

所以的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

(2)由(1)知，．

若小明先回答问题，记为小明的累计得分，则的所有可能取值为，，．

；

；

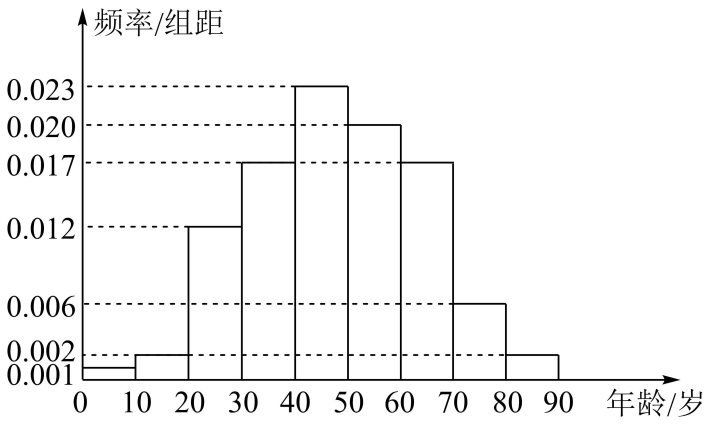
．

所以．

因为，所以小明应选择先回答类问题．

# 题型五：简单的随机抽样与用样本估计总体

1．(2022新高考全国II卷·第19题)在某地区进行流行病学调查，随机调查了100位某种疾病患者的年龄，得到如下的样本数据的频率分布直方图：



(1)估计该地区这种疾病患者的平均年龄(同一组中的数据用该组区间的中点值为代表)；

(2)估计该地区一位这种疾病患者的年龄位于区间的概率；

(3)已知该地区这种疾病的患病率为，该地区年龄位于区间的人口占该地区总人口的．从该地区中任选一人，若此人的年龄位于区间，求此人患这种疾病的概率．(以样本数据中患者的年龄位于各区间的频率作为患者的年龄位于该区间的概率，精确到0．0001)．

**【答案】**(1)岁；

(2)；

(3)．

解析：(1)平均年龄

(岁)．

(2)设{一人患这种疾病的年龄在区间}，所以

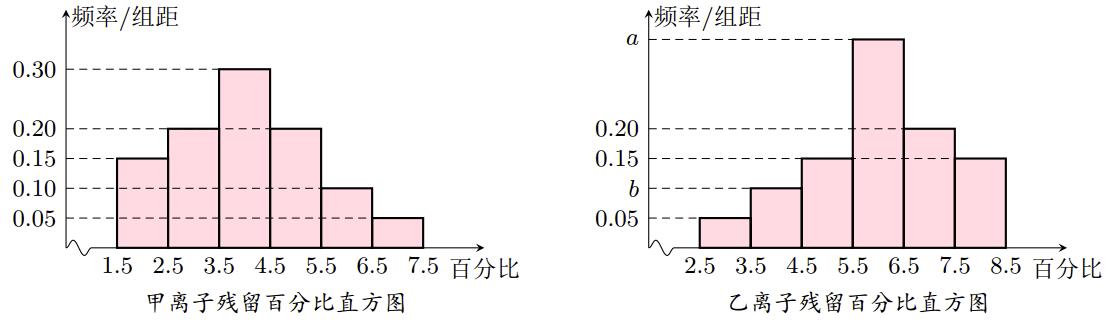
．

(3)设任选一人年龄位于区间，任选一人患这种疾病，

则由条件概率公式可得

．

2．(2019·全国Ⅲ·理·第17题)为了解甲、乙两种离子在小鼠体内的残留程度，进行如下试验：将200只小鼠随机分成两组，每组100只，其中组小鼠给服甲离子溶液，组小鼠给服乙离子溶液．每只小鼠给服的溶液体积相同、摩尔浓度相同．经过一段时间后用某种科学方法测算出残留在小鼠体内离子的百分比．根据试验数据分别得到如下直方图：



记为事件：“乙离子残留在体内的百分比不低于”，根据直方图得到的估计值为．

(1)求乙离子残留百分比直方图中的值；

(2)分别估计甲、乙离子残留百分比的平均值(同一组中的数据用该组区间的中点值为代表)．

**【答案】**(1)，；(2)，.00．

【官方解析】

(1)由已知得，故，．

(2)甲离子残留百分比的平均值的估计值为

．

乙离子残留百分比的平均值的估计值为

．

【点评】本题考查频率分布直方图的相关概念和频率分布直方图中平均数法人计算，属于基础题．

3．(2018年高考数学课标Ⅲ卷(理)·第18题)(12分)某工厂为提高生产效率，开展技术创新活动，提出了完成某项生产任务的两种生产方式，为比较两咱生产方式的效率，选取名工人，将他们随机分成两组，每组人，第一组工人用第一种生产方式，第二组工人用第二种生产方式．根据工人完成生产任务的工作时间(单位：)绘制了如下茎叶图：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 第一种生产方式 | | | | | | | | | |  | 第二种生产方式 | | | | | | | | | |
|  |  |  |  |  |  |  |  |  | 8 | 6 | 5 | 5 | 6 | 8 | 9 |  |  |  |  |  |
|  |  |  |  |  |  | 9 | 7 | 6 | 2 | 7 | 0 | 1 | 2 | 2 | 3 | 4 | 5 | 6 | 6 | 8 |
| 9 | 8 | 7 | 7 | 6 | 5 | 4 | 3 | 3 | 2 | 8 | 1 | 4 | 4 | 5 |  |  |  |  |  |  |
|  |  |  |  |  | 2 | 1 | 1 | 0 | 0 | 9 | 0 |  |  |  |  |  |  |  |  |  |

(1)根据茎叶图判断哪种生产方式的效率更高？并说明理由；

(2)求名工人完成生产任务所需时间的中位数，并将完成生产任务所需时间超过和不超过的工人数填入下面的列联表：

|  |  |  |
| --- | --- | --- |
|  | 超过 | 不超过 |
| 第一种生产方式 |  |  |
| 第二种生产方式 |  |  |

(3)根据(2)的列联表，能否有的把握认为两种生产方式的效率有差异？

附：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

**【答案】**【官方解析】(1)第二种生产方式的效率更高．

理由如下：

(i)由茎叶图可知：用第一种生产方式的工人中，有75%的工人完成生产任务所需时间至少80分钟，用第二种生产方式的工人中，有75%的工人完成生产任务所需时间至多79分钟．因此第二种生产方式的效率更高．．

(ii)由茎叶图可知：用第一种生产方式的工人完成生产任务所需时间的中位数为85．5分钟，用第二种生产方式的工人完成生产任务所需时间的中位数为73．5分钟．因此第二种生产方式的效率更高．

(iii)由茎叶图可知：用第一种生产方式的工人完成生产任务平均所需时间高于80分钟；用第二种生产方式的工人完成生产任务平均所需时间低于80分钟，因此第二种生产方式的效率更高．

(iv)由茎叶图可知：用第一种生产方式的工人完成生产任务所需时间分布在茎8上的最多，关于茎8大致呈对称分布；用第二种生产方式的工人完成生产任务所需时间分布在茎7上的最多，关于茎7大致呈对称分布，又用两种生产方式的工人完成生产任务所需时间分布的区间相同，故可以认为用第二种生产方式完成生产任务所需的时间比用第一种生产方式完成生产任务所需的时间更少，因此第二种生产方式的效率更高．

以上给出了4种理由，考生答出其中任意一种或其他合理理由均可得分．

(2)由茎叶图知．

列联表如下：

|  |  |  |
| --- | --- | --- |
|  | 超过 | 不超过 |
| 第一种生产方式 | 15 | 5 |
| 第二种生产方式 | 5 | 15 |

(3)由于

所以有99%的把握认为两种生产方式的效率有差异．

【民间解析】(1)法一：第二种生产方式效率更高，因为第二种多数数据集中在之间，第一种多数数据集中在之间，易知第一种完成任务的平均时间大于第二种，故第二种生产方式的效率更高。

法二：第一种生产方式完成任务的平均时间为





第二种生产完成任务的平均时间为







第一种生产方式完成任务的平均时间第二种生产方式完成任务的平均时间

所以第二种生产方式效率更高

(2)中位数为

|  |  |  |
| --- | --- | --- |
|  | 超过 | 不超过 |
| 第一种生产方式 | 15 | 5 |
| 第二种生产方式 | 5 | 15 |

(3)由(2)可计算得

所以有的把握认为两种生产方式的效率有差异．

点评：本题主要考查了茎叶图和独立性检验，考察学生的计算能力和分析问题的能力，贴近生活．

4．(2014高考数学广东理科·第17题)随机观测生产某种零件的某工厂25名工人的日加工零件数(单位：件)，获得数据如下：

30,42,41,36,44,40,37,37,25,45,29,43,31,36,49,34,33,43,38,42,32,34,46,39,36

根据上述数据得到样本的频率分布表如下：

|  |  |  |
| --- | --- | --- |
| 分组 | 频数 | 频率 |
| [25,30] | 3 | 0．12 |
| (30,35] | 5 | 0．20 |
| (35,40] | 8 | 0．32 |
| (40,45] |  |  |
| (45,50] |  |  |

(1)确定样本频率分布表中和的值；

(2)根据上述频率分布表，画出样本频率分布直方图；

(3)根据样本频率分布直方图，求在该厂任取4人，至少有1人的日加工零件数落在区间(30,35］的概率．

**【答案】**解：(1)

(2)先计算 频率/组距；然后作图即可

(3)由(1)知，任取一人，日加工零件数落在区间(30,35］的概率为

设该厂任取4人，没有人日加工零件数落在区间(30,35］的事件为，

则，所以

答：在该厂任取4人，至少有1人的日加工零件数落在区间(30,35］的概率为

5．(2015高考数学广东理科·第17题)(本小题满分12分)

某工厂36名工人的年龄数据如下表．

|  |  |  |  |
| --- | --- | --- | --- |
| 工人编号 年龄 | 工人编号 年龄 | 工人编号 年龄 | 工人编号 年龄 |
| 1 40 | 10 36 | 19 27 | 28 34 |
| 2 44 | 11 31 | 20 43 | 29 39 |
| 3 40 | 12 38 | 21 41 | 30 43 |
| 4 41 | 13 39 | 22 37 | 31 38 |
| 5 33 | 14 43 | 23 34 | 32 42 |
| 6 40 | 15 45 | 24 42 | 33 53 |
| 7 45 | 16 39 | 25 37 | 34 37 |
| 42 | 17 38 | 26 44 | 35 49 |
| 9 43 | 18 36 | 27 42 | 36 39 |

(1)用系统抽样法从36名工人中抽取容量为9的样本，且在第一分段里用随机抽样法抽到的年龄数据为44，列出样本的年龄数据；

(2)计算(1)中样本的平均值和方差；

(3)36名工人中年龄在与之间有多少人？所占的百分比是多少(精确到0．01％)？

**【答案】**解析：(1)依题所抽样本编号是一个首项为2，公差为4的等差数列，故其所有样本编号依次为2,6,10,14,18,22,26,30,34,对应样本的年龄数据依次为44,40,36,43,36,37,44,43,37;

(2)由平均值公式知：

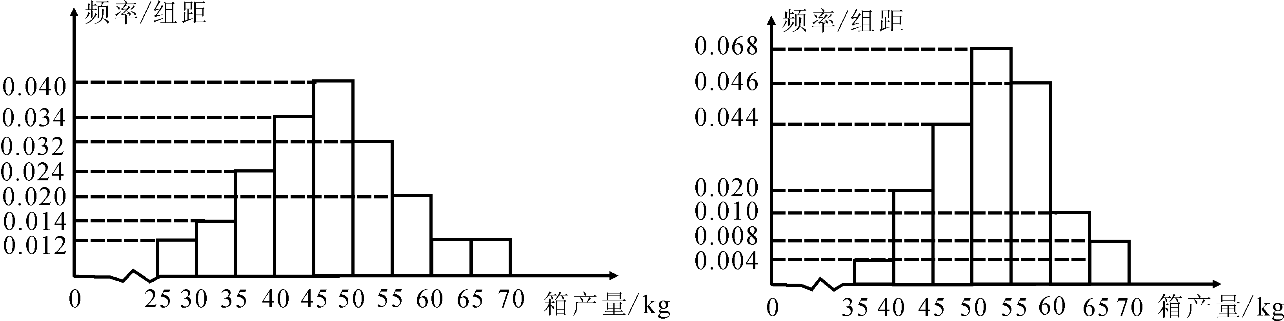
由方差公式知：

(3)因为

所以36名工人中年龄在和之间的人数等于区间的人数；

即40，40，41，…………，39，共23人．所占百分比为．

6．(2017年高考数学课标Ⅱ卷理科·第18题)(12分)淡水养殖场进行某水产品的新、旧网箱养殖方法的产量对比，收获时各随机抽取了100 个网箱，测量各箱水产品的产量(单位：kg)某频率直方图如下：

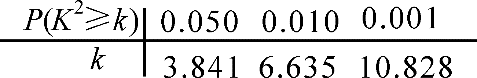


(1)设两种养殖方法的箱产量相互独立，记A表示事件：旧养殖法的箱产量低于50kg, 新养殖法的箱产量不低于50kg,估计A的概率；

(2)填写下面列联表，并根据列联表判断是否有99%的把握认为箱产量与养殖方法有关：

|  |  |  |
| --- | --- | --- |
|  | 箱产量＜50kg | 箱产量≥50kg |
| 旧养殖法 |  |  |
| 新养殖法 |  |  |

(3)根据箱产量的频率分布直方图，求新养殖法箱产量的中位数的估计值(精确到0．01)

**【答案】**(1)学科网 版权所有；

(2)有学科网 版权所有的把握认为箱产量与养殖方法有关；

(3)学科网 版权所有。

【命题意图】概率统计,独立检验等知识的综合运用

【基本解法】

(Ⅰ)旧养殖法的箱产量低于50kg的频率为0．012×5+0．014×5+0．024×5+0．034×5+0．040×5=0．62，由于两种养殖方法的箱产量相互独立，

于是P(A)=0．62×0．66=0．4092

(Ⅱ)旧养殖法的箱产量低于50kg的有100×0．62=62箱，不低于50kg的有38箱，新养殖法的箱产量不低于50kg的有100×0．66=66箱，低于50kg的有34箱，得到2×2列联表如下：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 箱产量<50kg | 箱产量≥50kg | 合计 |
| 旧养殖法 | 62 | 38 | 100 |
| 新养殖法 | 34 | 66 | 100 |
| 合计 | 96 | 104 | 200 |

所以



，所以有99%的把握认为箱产量与养殖方法有关。

(III)根据箱产量的频率分布直方图，新养殖法的箱产量不低于50kg的频率为0．038×5+0．046×5+0．010×5+0．008×5=0．66>0．50，不低于55kg的频率为0．046×5+0．010×5+0．008×5=0．32<0．50，于是新养殖法箱产量的中位数介于50kg到55kg之间，设新养殖法箱产量的中位数为x，则有

(55-x)×0．068**+**0．046×5+0．010×5+0．008×5=0．50

解得x=52． 3529

因此，新养殖法箱产量的中位数的估计值52． 35。

7．(2016高考数学四川理科·第16题)(12分)我国是世界上严重缺水的国家，某市政府为了鼓励居民节约用水，集合调整居民用水方案，拟确定一个合理的月用水标准 吨，一位居民的月用水量不超过的部分按平价收费，超出 的部分按议价收费，为了了解居民用水情况，通过抽样，获得了某年位居民每人的月均用水量(单位：吨)，将数据按照，，，分成组，制成了如图所示的频率分布直方图．



(1)求直方图中的值；

(2)设该市有万居民，估计全市居民中月均用水量不低于吨的人数．

(3)若该市政府希望使 的居民每月的用水量不超过标准吨，估计的值，并说明理由．

**【答案】**【官方解答】(1)由频率直方图知，月均 用水量在中的频率为

同理在中的频率分别为



由

解得

(2)由(1)知，为居民每人月均用水量不低于吨的的频率为

由以上样本的频率分布，可以估计全市万居民月均水量不低于吨的人数为



(3)因为前六组的频率之和为

而前五组的频率之和为

所以，由，解得

所以估计月用水量标准为吨时，的居民每月用水量不超过标准

【民间解析】(1)由题知组距

则有

解得

(2)设“月均用水量不低于吨的人数”为事件 ，则

所以人数为 万

(3)频率依次是

由图可知，月均用水量小于2．5吨的居民人数所占百分比为：



即的居民月均用水量小于2．5吨,

同理，88%的居民月均用水量小于3吨，故

假设月均用水量平均分布，则(吨)．

注：本次估计默认组间是平均分布，与实际可能会产生一定误差．

所以位于 之间，则

8．(2016高考数学北京理科·第16题)(本小题13分)三个班共有100名学生，为调查他们的体育锻炼情况，通过分层抽样获得了部分学生一周的锻炼时间，数据如下表(单位：小时)；

|  |  |
| --- | --- |
| A班 | 6  7  8 |
| B班 | 6 7 8  10 11 12 |
| C班 | 3  6  9  12 |

(I)试估计班的学生人数；

(II)从班和班抽出的学生中，各随机选取一人，班选出的人记为甲，班选出的人记为乙，假设所有学生的锻炼时间相对独立，求该周甲的锻炼时间比乙的锻炼时间长的概率；

(III)再从三个班中各随机抽取一名学生，他们该周的锻炼时间分别是7，9，8．25(单位：小时)，这3个新数据与表格中的数据构成的新样本的平均数记，表格中数据的平均数记为 ，试判断和的大小，(结论不要求证明)

**【答案】**(1)40人;(2);(3)拉低了平均值．

【官方解答】(Ⅰ)由题意知，抽出的名学生中，来自班的学生有名．

根据分层抽样方法，班的学生人数估计为．

(Ⅱ)设事件为“甲是现有样本中班的第个人”，，

事件为“乙是现有样本中班的第个人”，，

由题意可知，，；，．

,,．

设事件为“该周甲的锻炼时间比乙的锻炼时间长”．由题意知，





因此





(Ⅲ)．

【民间解答】⑴，C班学生40人

⑵在A班中取到每个人的概率相同均为

设班中取到第个人事件为

C班中取到第个人事件为

班中取到的概率为

所求事件为

则





⑶

三组平均数分别为总均值

但中多加的三个数据平均值为，比小，

故拉低了平均值

9．(2019·北京·理·第17题)改革开放以来，人们的支付方式发生了巨大转变．近年来，移动支付已成为主要支付方式之一．为了解某校学生上个月A，B两种移动支付方式的使用情况，从全校学生中随机抽取了100人，发现样本中A，B两种支付方式都不使用的有5人，样本中仅使用A和仅使用B的学生的支付金额分布情况如下：

|  |  |  |  |
| --- | --- | --- | --- |
| 交付金额(元)  支付方式 | (0，1000] | (1000，2000] | 大于2000 |
| 仅使用A | 18人 | 9人 | 3人 |
| 仅使用B | 10人 | 14人 | 1人 |

(Ⅰ)从全校学生中随机抽取1人，估计该学生上个月A，B两种支付方式都使用的概率；

(Ⅱ)从样本仅使用A和仅使用B的学生中各随机抽取1人，以*X*表示这2人中上个月支付金额大于1000元的人数，求*X*的分布列和数学期望；

(Ⅲ)已知上个月样本学生的支付方式在本月没有变化．现从样本仅使用A的学生中，随机抽查3人，发现他们本月的支付金额都大于2000元．根据抽查结果，能否认为样本仅使用A的学生中本月支付金额大于2000元的人数有变化？说明理由．

**【答案】**【解析】(Ⅰ)由题意可知，两种支付方式都使用的人数为：人，则全校学生中随机抽取1人，该学生上个月*A*，*B*两种支付方式都使用的概率．

(Ⅱ)由题意可知，

仅使用*A*支付方法的学生中，金额不大于1000的人数占，金额大于1000的人数占，

仅使用*B*支付方法的学生中，金额不大于1000的人数占，金额大于1000的人数占，

则*X*可能的取值为0，1，2．

，，，

所以*X*的分布列为：

|  |  |  |  |
| --- | --- | --- | --- |
| *X* | 0 | 1 | 2 |
|  |  |  |  |

其数学期望为．

(Ⅲ)我们不认为样本仅使用*A*的学生中本月支付金额大于2000元的人数有变化．理由如下：

随机事件在一次随机实验中是否发生是随机的，是不能预知的，随着试验次数的增多，频率越来越稳定于概率。学校是一个相对消费稳定的地方，每个学生根据自己的实际情况每个月的消费应该相对固定，出现题中这种现象可能是发生了“小概率事件”．(答案不唯一：小概率事件发生有理由认为人数发生了变化)

10．(2018年高考数学天津(理)·第16题)(本小题满分13分)已知某单位甲、乙、丙三个部门的员工人数分别为24，16，16．现采用分层抽样的方法从中抽取7人，进行睡眠时间的调查．

(1)应从甲、乙、丙三个部门的员工中分别抽取多少人？

(2)若抽出的7人中有4人睡眠不足，3人睡眠充足，现从这7人中随机抽取3人做进一步的身体检查．

(i)用*X*表示抽取的3人中睡眠不足的员工人数，求随机变量*X*的分布列与数学期望；

(ii)设*A*为事件“抽取的3人中，既有睡眠充足的员工，也有睡眠不足的员工”，求事件*A*发生的概率．

**【答案】**(1)解：由已知，甲、乙、丙三个部门的员工人数之比为3∶2∶2，由于采用分层抽样的方法从中抽取7人，因此应从甲、乙、丙三个部门的员工中分别抽取3人，2人，2人．

(2)(i)解：随机变量*X*的所有可能取值为0，1，2，3．



所以，随机变量*X*的分布列为

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *X* | 0 | 1 | 2 | 3 |
| *P* |  |  |  |  |

随机变量*X*的数学期望．

(ii)解：设事件为“抽取的3人中，睡眠充足的员工有1人，睡眠不足的员工有2人”；事件为“抽取的3人中，睡眠充足的员工有2人，睡眠不足的员工有1人”，则，且与互斥，由(i)知，，故．

所以，事件发生的概率为．

11．(2018年高考数学北京(理)·第17题)(本小题12分)电影公司随机收集了电影的有关数据，经分类整理得到下表：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 电影类型 | 第一类 | 第二类 | 第三类 | 第四类 | 第五类 | 第六类 |
| 电影部数 |  |  |  |  |  |  |
| 好评率 |  |  |  |  |  |  |

好评率是指：一类电影中获得好评的部数与该类电影的部数的比值．

假设所有电影是否获得好评相互独立．

(Ⅰ)从电影公司收集的电影中随机选取部，求这部电影是获得好评的第四类电影的概率；(Ⅱ)从第四类电影和第五类电影中各随机选取部，估计恰有部获得好评的概率；(Ⅲ)假设每类电影得到人们喜欢的概率与表格中该类电影的好评率相等，用“”表示第*k*类电影得到人们喜欢，“”表示第类电影没有得到人们喜欢()．写出方差的大小关系．

**【答案】**解：(Ⅰ)由题意知，样本中电影的总部数是，

第四类电影中获得好评的电影部数是200×0．25=50．

故所求概率为．

(Ⅱ)设事件为“从第四类电影中随机选出的电影获得好评”，

事件为“从第五类电影中随机选出的电影获得好评”．

故所求概率为

．

由题意知：估计为，估计为．

故所求概率估计为

(Ⅲ)．解：的分布列为

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |



的分布列为

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |



的分布列为

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |





的分布列为

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |





的分布列为

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |





的分布列为

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |







# 题型六：相关关系与回归分析

1．(2020年高考课标Ⅱ卷理科·第18题)某沙漠地区经过治理，生态系统得到很大改善，野生动物数量有所增加．为调查该地区某种野生动物数量，将其分成面积相近的200个地块，从这些地块中用简单随机抽样的方法抽取20个作为样区，调查得到样本数据(*xi*，*yi*)(*i*=1，2，…，20)，其中*xi*和*yi*分别表示第*i*个样区的植物覆盖面积(单位：公顷)和这种野生动物的数量，并计算得，，，，．

(1)求该地区这种野生动物数量的估计值(这种野生动物数量的估计值等于样区这种野生动物数量的平均数乘以地块数)；

(2)求样本(*xi*，*yi*)(*i*=1，2，…，20)的相关系数(精确到0．01)；

(3)根据现有统计资料，各地块间植物覆盖面积差异很大．为提高样本的代表性以获得该地区这种野生动物数量更准确的估计，请给出一种你认为更合理的抽样方法，并说明理由．

附：相关系数*r*=，≈1．414．

**【答案】**(1)；(2)；(3)详见解析

解析：(1)样区野生动物平均数为，

地块数为200，该地区这种野生动物的估计值为

(2)样本(*i*=1，2，…，20)的相关系数为



(3)由(2)知各样区的这种野生动物的数量与植物覆盖面积有很强的正相关性，

由于各地块间植物覆盖面积差异很大，从俄各地块间这种野生动物的数量差异很大，

采用分层抽样的方法较好地保持了样本结构与总体结构得以执行，提高了样本的代表性，

从而可以获得该地区这种野生动物数量更准确的估计．

【点晴】本题主要考查平均数的估计值、相关系数的计算以及抽样方法的选取，考查学生数学运算能力，是一道容易题．

2．(2022年高考全国乙卷数学(理)·第19题)某地经过多年的环境治理，已将荒山改造成了绿水青山．为估计一林区某种树木的总材积量，随机选取了10棵这种树木，测量每棵树的根部横截面积(单位：)和材积量(单位：)，得到如下数据：

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 样本号ｉ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 总和 |
| 根部横截面积 | 0．04 | 0．06 | 0．04 | 0．08 | 008 | 0．05 | 0．05 | 0．07 | 0．07 | 0．06 | 0．6 |
| 材积量 | 0．25 | 0．40 | 0．22 | 0．54 | 0．51 | 0．34 | 0．36 | 0．46 | 0．42 | 0．40 | 3．9 |

并计算得．

(1)估计该林区这种树木平均一棵的根部横截面积与平均一棵的材积量；

(2)求该林区这种树木的根部横截面积与材积量的样本相关系数(精确到0．01)；

(3)现测量了该林区所有这种树木的根部横截面积，并得到所有这种树木的根部横截面积总和为．已知树木的材积量与其根部横截面积近似成正比．利用以上数据给出该林区这种树木的总材积量的估计值．

附：相关系数．

**【答案】**(1)；

(2)

(3)

解析：【小问1详解】

样本中10棵这种树木的根部横截面积的平均值

样本中10棵这种树木的材积量的平均值

据此可估计该林区这种树木平均一棵的根部横截面积为，

平均一棵的材积量为

【小问2详解】





则

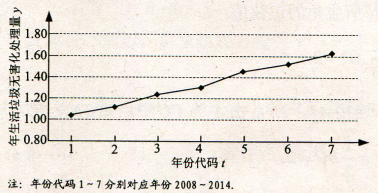
小问3详解】

设该林区这种树木的总材积量的估计值为，

又已知树木的材积量与其根部横截面积近似成正比，

可得，解之得．则该林区这种树木的总材积量估计为

3．(2016高考数学课标Ⅲ卷理科·第18题)下图是我国2008年至2014年生活垃圾无害化处理量(单位:亿吨)的折线图.



(Ⅰ)由折线图看出,可用线性回归模型拟合*y*与*t*的关系,请用相关系数加以说明;

(Ⅱ)建立*y*关于*t*的回归方程(系数精确到0.01),预测2016年我国生活垃圾无害化处理量.

参考数据:,,,.

参考公式:相关系数

回归方程中斜率和截距最小二乘估计公式分别为:,.

**【答案】**(Ⅰ)理由见解析;(Ⅱ)1.82亿吨.

【解析】(Ⅰ)由折线图中数据和附注中参考数据得,,,

,.

因为与的相关系数近似为0.99,说明与的线性相关程度相当高

从而可以用线性回归模型拟合与的关系.

(Ⅱ)由及(Ⅰ)得,

.

所以,关于的回归方程为:.

将2016年对应的代入回归方程得:.

所以预测2016年我国生活垃圾无害化处理量将约1.82亿吨.

4．(2014高考数学课标2理科·第19题)(本小题满分12分)

某地区2007年至2013年农村居民家庭人均纯收入*y*(单位：千元)的数据如下表：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 年份 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
| 年份代号*t* | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 人均纯收入*y* | 2．9 | 3．3 | 3．6 | 4．4 | 4．8 | 5．2 | 5．9 |

(Ⅰ)求*y*关于*t*的线性回归方程；

(Ⅱ)利用(Ⅰ)中的回归方程，分析2007年至2013年该地区农村居民家庭人均纯收入的变化情况，并预测该地区2015年农村居民家庭人均纯收入．

附：回归直线的斜率和截距的最小二乘估计公式分别为：

**【答案】**解析：(Ⅰ)



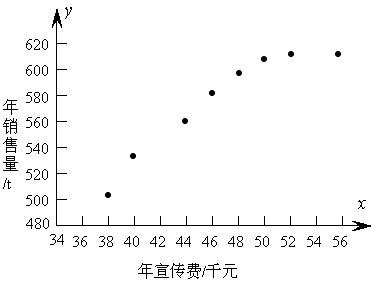
设回归方程为代入公式，经计算得：



所以，关于的回归方程为．

(Ⅱ)，2007年至2013年该区域人均纯收入稳步增长，预计到2015年，高地区人均纯收入(千元)，所以，预计到2015年，该地区人均纯收入约6800元左右．

5．(2015高考数学新课标1理科·第19题)某公司为确定下一年度投入某种产品的宣传费，需了解年宣传费(单位：千元)对年销售量(单位：)和年利润(单位：千元)的影响，对近8年的年宣传费和年销售量(=1,2，···，8)数据作了初步处理，得到下面的散点图及一些统计量的值。



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 46．6 | 56．3 | 6．8 | 289．8 | 1．6 | 1469 | 108．8 |

表中，。

(Ⅰ)根据散点图判断，与哪一个适宜作为年销售量关于年宣传费的回归方程类型？(给出判断即可，不必说明理由)

(Ⅱ)根据(Ⅰ)的判断结果及表中数据，建立关于的回归方程；

(Ⅲ)已知这种产品的年利率与、的关系为．根据(Ⅱ)的结果回答下列问题：

(i)年宣传费时，年销售量及年利润的预报值是多少？

(ii)年宣传费为何值时，年利率的预报值最大？

附：对于一组数据,，……，,其回归线的斜率和截距的最小二乘估计分别为：

、

**【答案】**(Ⅰ)适合作为年销售关于年宣传费用的回归方程类型；(Ⅱ)(Ⅲ)46．24

分析：(Ⅰ)由散点图及所给函数图像即可选出适合作为拟合的函数；(Ⅱ)令，先求出建立关于的线性回归方程，即可关于的回归方程；(Ⅲ)(ⅰ)利用关于的回归方程先求出年销售量的预报值，再根据年利率z与x、y的关系为z=0．2y-x即可年利润z的预报值；(ⅱ)根据(Ⅱ)的结果知，年利润z的预报值，列出关于的方程，利用二次函数求最值的方法即可求出年利润取最大值时的年宣传费用．

解析：

(Ⅰ)由散点图可以判断，适合作为年销售关于年宣传费用的回归方程类型．

(Ⅱ)令，先建立关于的线性回归方程，由于=，

∴=563-68×6．8=100．6．

∴关于的线性回归方程为，

∴关于的回归方程为．

(Ⅲ)(ⅰ)由(Ⅱ)知，当=49时，年销售量的预报值

=576．6，

．

(ⅱ)根据(Ⅱ)的结果知，年利润z的预报值

，

∴当=，即时，取得最大值．

故宣传费用为46．24千元时，年利润的预报值最大．……12分

# 题型七：独立性检验

1．(2023年全国甲卷理科·第19题)一项试验旨在研究臭氧效应．实验方案如下：选40只小白鼠，随机地将其中20只分配到实验组，另外20只分配到对照组，实验组的小白鼠饲养在高浓度臭氧环境，对照组的小白鼠饲养在正常环境，一段时间后统计每只小白鼠体重的增加量(单位：g)．

(1)设表示指定的两只小白鼠中分配到对照组的只数，求的分布列和数学期望；

(2)实验结果如下：

对照组的小白鼠体重的增加量从小到大排序为：

15．2 18．8 20．2 21．3 22．5 23．2 25．8 26．5 27．5 30．1

32．6 34．3 34．8 35．6 35．6 35．8 36．2 37．3 40．5 43．2

对照组的小白鼠体重的增加量从小到大排序为：

7．8 9．2 11．4 12．4 13．2 15．5 16．5 18．0 18．8 19．2

19．8 20．2 21．6 22．8 23．6 23．9 25．1 28．2 32．3 36．5

(i)求40只小鼠体重的增加量的中位数*m*，再分别统计两样本中小于*m*与不小于的数据的个数，完成如下列联表：

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 对照组 |  |  |
| 实验组 |  |  |

(ii)根据(i)中的列联表，能否有95%的把握认为小白鼠在高浓度臭氧环境中与正常环境中体重的增加量有差异．

附：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0．100 | 0．050 | 0．010 |
|  | 2．706 | 3．841 | 6．635 |

**【答案】**(1)分布列见解析，

(2)(i)；列联表见解析，(ii)能

解析：(1)依题意，的可能取值为，

则，，，

所以的分布列为：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

故．

(2)(i)依题意，可知这40只小白鼠体重增量的中位数是将两组数据合在一起，从小到大排后第20位与第21位数据的平均数，观察数据可得第20位为，第21位数据为，

所以，

故列联表为：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | 合计 |
| 对照组 | 6 | 14 | 20 |
| 实验组 | 14 | 6 | 20 |
| 合计 | 20 | 20 | 40 |

(ii)由(i)可得，，

所以能有的把握认为小白鼠在高浓度臭氧环境中与正常环境中体重的增加量有差异．

2．(2021年高考全国甲卷理科·第17题)甲、乙两台机床生产同种产品，产品按质量分为一级品和二级品，为了比较两台机床产品的质量，分别用两台机床各生产了200件产品，产品的质量情况统计如下表：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 一级品 | 二级品 | 合计 |
| 甲机床 | 150 | 50 | 200 |
| 乙机床 | 120 | 80 | 200 |
| 合计 | 270 | 130 | 400 |

(1)甲机床、乙机床生产的产品中一级品的频率分别是多少?

(2)能否有99%的把握认为甲机床的产品质量与乙机床的产品质量有差异?

附：

|  |  |  |  |
| --- | --- | --- | --- |
|  | 0．050 | 0．010 | 0．001 |
| *k* | 3．841 | 6．635 | 10．828 |

**【答案】**(1)75%；60%；

(2)能．

解析：(1)甲机床生产的产品中的一级品的频率为,

乙机床生产的产品中的一级品的频率为．

(2),

故能有99%的把握认为甲机床的产品与乙机床的产品质量有差异．

3．(2020年高考课标Ⅲ卷理科·第18题)某学生兴趣小组随机调查了某市100天中每天的空气质量等级和当天到某公园锻炼的人次，整理数据得到下表(单位：天)：

|  |  |  |  |
| --- | --- | --- | --- |
| 锻炼人次  空气质量等级 | [0，200] | (200，400] | (400，600] |
| 1(优) | 2 | 16 | 25 |
| 2(良) | 5 | 10 | 12 |
| 3(轻度污染) | 6 | 7 | 8 |
| 4(中度污染) | 7 | 2 | 0 |

(1)分别估计该市一天的空气质量等级为1，2，3，4的概率；

(2)求一天中到该公园锻炼的平均人次的估计值(同一组中的数据用该组区间的中点值为代表)；

(3)若某天的空气质量等级为1或2，则称这天“空气质量好”；若某天的空气质量等级为3或4，则称这天“空气质量不好”．根据所给数据，完成下面的2×2列联表，并根据列联表，判断是否有95%的把握认为一天中到该公园锻炼的人次与该市当天的空气质量有关？

|  |  |  |
| --- | --- | --- |
|  | 人次≤400 | 人次>400 |
| 空气质量好 |  |  |
| 空气质量不好 |  |  |

附：，

|  |  |  |  |
| --- | --- | --- | --- |
| *P*(*K*2≥*k*) | 0．050 | 0．010 | 0．001 |
| *k* | 3841 | 6．635 | 10．828 |

**【答案】**(1)该市一天的空气质量等级分别为、、、的概率分别为、、、；(2)；(3)有，理由见解析．

解析：(1)由频数分布表可知，该市一天的空气质量等级为的概率为，等级为的概率为，等级为的概率为，等级为的概率为；

(2)由频数分布表可知，一天中到该公园锻炼的人次的平均数为

(3)列联表如下：

|  |  |  |
| --- | --- | --- |
|  | 人次 | 人次 |
| 空气质量不好 |  |  |
| 空气质量好 |  |  |

，

因此，有的把握认为一天中到该公园锻炼的人次与该市当天的空气质量有关．

【点睛】本题考查利用频数分布表计算频率和平均数，同时也考查了独立性检验的应用，考查数据处理能力，属于基础题．

4．(2020年新高考全国Ⅰ卷(山东)·第19题)为加强环境保护，治理空气污染，环境监测部门对某市空气质量进行调研，随机抽查了天空气中的和浓度(单位：)，得下表：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  | 32 | 18 | 4 |
|  | 6 | 8 | 12 |
|  | 3 | 7 | 10 |

(1)估计事件“该市一天空气中浓度不超过，且浓度不超过”的概率；

(2)根据所给数据，完成下面的列联表：

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

(3)根据(2)中列联表，判断是否有的把握认为该市一天空气中浓度与浓度有关？

附：，

|  |  |
| --- | --- |
|  | 0．050 0．010 0．001 |
|  | 3．841 6．635 10．828 |

**【答案】**(1)；(2)答案见解析；(3)有．

解析：(1)由表格可知，该市100天中，空气中的浓度不超过75，且浓度不超过150的天数有天，

所以该市一天中，空气中的浓度不超过75，且浓度不超过150的概率为；

(2)由所给数据，可得列联表为：

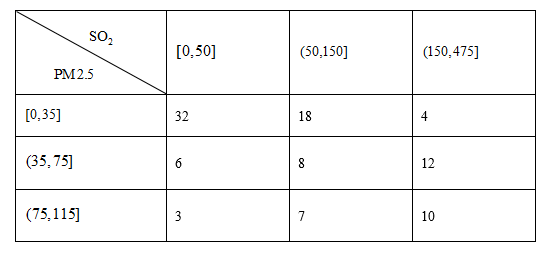
|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | 合计 |
|  | 64 | 16 | 80 |
|  | 10 | 10 | 20 |
| 合计 | 74 | 26 | 100 |

(3)根据列联表中的数据可得

，

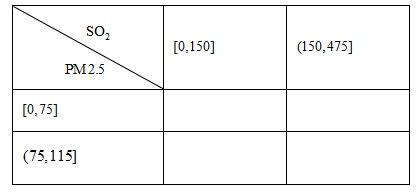
因为根据临界值表可知，有的把握认为该市一天空气中浓度与浓度有关．

5．(2020年新高考全国卷Ⅱ数学(海南)·第19题)为加强环境保护，治理空气污染，环境监测部门对某市空气质量进行调研，随机抽查了天空气中的和浓度(单位：)，得下表：



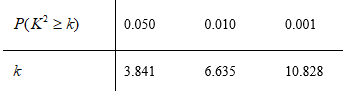
(1)估计事件“该市一天空气中浓度不超过，且浓度不超过”的概率；

(2)根据所给数据，完成下面的列联表：



(3)根据(2)中的列联表，判断是否有的把握认为该市一天空气中浓度与浓度有关？

附：，



**【答案】**(1)；(2)答案见解析；(3)有．

解析：(1)由表格可知，该市100天中，空气中的浓度不超过75，且浓度不超过150的天数有天，

所以该市一天中，空气中的浓度不超过75，且浓度不超过150的概率为；

(2)由所给数据，可得列联表为：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | 合计 |
|  | 64 | 16 | 80 |
|  | 10 | 10 | 20 |
| 合计 | 74 | 26 | 100 |

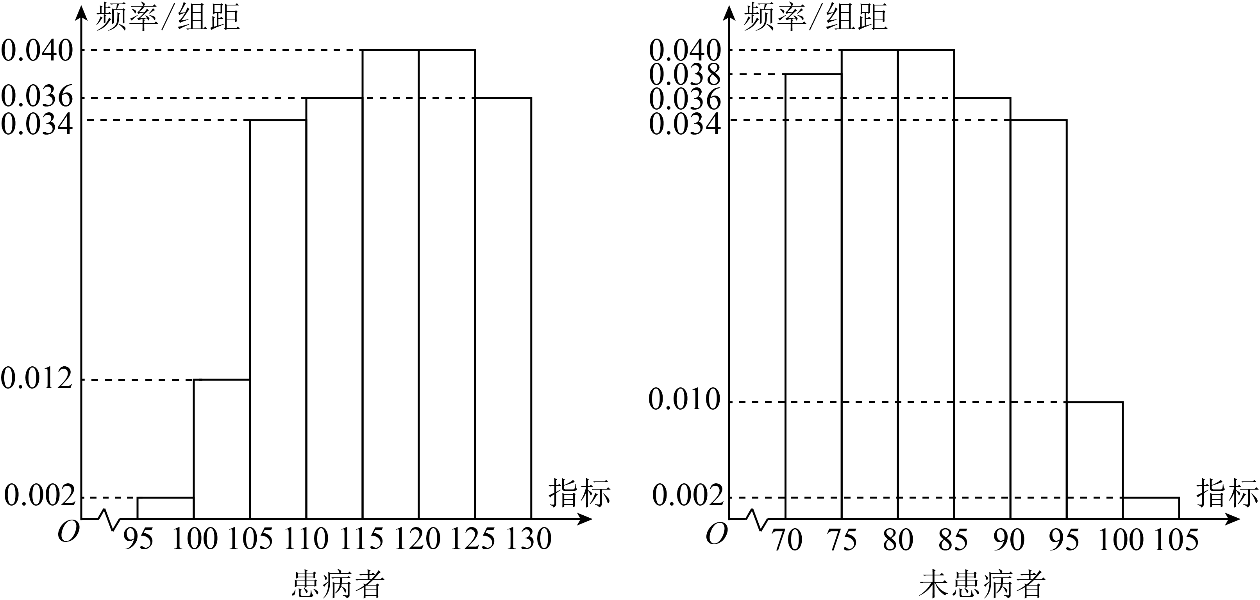
(3)根据列联表中的数据可得

，

因为根据临界值表可知，有的把握认为该市一天空气中浓度与浓度有关．

# 题型八：概率统计综合应用

1．(2023年新课标全国Ⅱ卷·第19题)某研究小组经过研究发现某种疾病患病者与未患病者的某项医学指标有明显差异，经过大量调查，得到如下的患病者和未患病者该指标的频率分布直方图：



利用该指标制定一个检测标准，需要确定临界值*c*，将该指标大于*c*人判定为阳性，小于或等于*c*的人判定为阴性．此检测标准的漏诊率是将患病者判定为阴性的概率，记为；误诊率是将未患病者判定为阳性的概率，记为．假设数据在组内均匀分布，以事件发生的频率作为相应事件发生的概率．

(1)当漏诊率％时，求临界值*c*和误诊率；

(2)设函数，当时，求的解析式，并求在区间的最小值．

**【答案】**(1)，；

(2)，最小值为．

解析：(1)

依题可知，左边图形第一个小矩形的面积为，所以，

所以，解得：，

．

(2)

当时，

 ；

当时，

 ,

故，

所以在区间的最小值为．

2．(2015高考数学新课标2理科·第18题)(本题满分12分)某公司为了解用户对其产品的满意度，从，两地区分别随机调查了20个用户，得到用户对产品的满意度评分如下：

地区：62 73 81 92 95 85 74 64 53 76

78 86 95 66 97 78 88 82 76 89

地区：73 83 62 51 91 46 53 73 64 82

93 48 65 81 74 56 54 76 65 79

(Ⅰ)根据两组数据完成两地区用户满意度评分的茎叶图，并通过茎叶图比较两地区满意度评分的平均值及分散程度(不要求计算出具体值，得出结论即可)；



(Ⅱ)根据用户满意度评分，将用户的满意度从低到高分为三个等级：

|  |  |  |  |
| --- | --- | --- | --- |
| 满意度评分 | 低于70分 | 70分到89分 | 不低于90分 |
| 满意度等级 | 不满意 | 满意 | 非常满意 |

记事件：“地区用户的满意度等级高于地区用户的满意度等级”．假设两地区用户的评价结果相互独立．根据所给数据，以事件发生的频率作为相应事件发生的概率，求的概率．

**【答案】**(Ⅰ)详见解析；(Ⅱ)．

解析：(Ⅰ)两地区用户满意度评分的茎叶图如下



通过茎叶图可以看出，A地区用户满意度评分的平均值高于B地区用户满意度评分的平均值；A地区用户满意度评分比较集中，B地区用户满意度评分比较分散．

(Ⅱ)记表示事件：“A地区用户满意度等级为满意或非常满意”；

表示事件：“A地区用户满意度等级为非常满意”；

表示事件：“B地区用户满意度等级为不满意”；

表示事件：“B地区用户满意度等级为满意”．

则与独立，与独立，与互斥，．

．

由所给数据得，，，发生的概率分别为，，，．故，

，，，故．

3．(2021年新高考全国Ⅱ卷·第21题)一种微生物群体可以经过自身繁殖不断生存下来，设一个这种微生物为第0代，经过一次繁殖后为第1代，再经过一次繁殖后为第2代……，该微生物每代繁殖的个数是相互独立的且有相同的分布列，设*X*表示1个微生物个体繁殖下一代的个数，．

(1)已知，求；

(2)设*p*表示该种微生物经过多代繁殖后临近灭绝概率，*p*是关于*x*的方程：的一个最小正实根，求证：当时，，当时，；

(3)根据你的理解说明(2)问结论的实际含义．

**【答案】**解析:(1)．

(2)设，

因为，故，

若，则，故．，

因为，，故有两个不同零点，且，且时，；时，；

故在，上为增函数，在上为减函数，若，因为在为增函数且，而当时，因为在上为减函数，故，

故为的一个最小正实根，

若，因为且在上为减函数，故1为的一个最小正实根，

综上，若，则．

若，则，故．此时，，故有两个不同零点，且，

且时，；时，；故在，上为增函数，在上为减函数，而，故，又，故在存在一个零点，且．所以为的一个最小正实根，此时，故当时，．

(3)意义：每一个该种微生物繁殖后代的平均数不超过1，则若干代必然灭绝，若繁殖后代的平均数超过1，则若干代后被灭绝的概率小于1．

4．(2019·全国Ⅰ·理·第21题)为治疗某种疾病，研制了甲、乙两种新药，希望知道哪种新药更有效，为此进行动物试验．试验方案如下：每一轮选取两只白鼠对药效进行对比试验．对于两只白鼠，随机选一只施以甲药，另一只施以乙药．一轮的治疗结果得出后，再安排下一轮试验．当其中一种药治愈的白鼠比另一种药治愈的白鼠多4只时，就停止试验，并认为治愈只数多的药更有效．为了方便描述问题，约定，对于每轮试验，若施以甲药的白鼠治愈且施以乙药的白鼠未治愈则甲药得1分，乙药得－1分；若施以乙药的白鼠治愈且施以甲药的白鼠未治愈则乙药得1分，甲药得－1分；若都治愈或都未治愈则两种药均得0分．甲、乙两种药的治愈率分别记为和，一轮试验中甲药的得分记为*X*．

(1)求*X*的分布列；

(2)若甲药、乙药在试验开始时都赋予4分，表示“甲药的累计得分为时，最终认为甲药比乙药更有效”的概率，则()，

其中，，．假设，．

(i)证明：为等比数列；

(ii)求，并根据的值解释这种试验方案的合理性．

**【答案】**(1)解：*X*的所有可能取值为，

．

所以的分布列为

|  |  |  |  |
| --- | --- | --- | --- |
| *X* |  | 0 | 1 |
| *P* |  |  |  |

(2)(i)由(1)得．

因此，故，即．

又因为，所以为公比为4，首项为的等比数列．

(ii)由(i)可得

.

由于，故，所以．

表示最终认为甲药更有效的概率，由计算结果可以看出，在甲药治愈率为0.5，乙药治愈率为0.8时，认为甲药更有效的概率为，此时得出错误结论的概率非常小，说明这种试验方案合理．

5．(2016高考数学课标Ⅰ卷理科·第19题)(本小题满分12分)某公司计划购买2台机器，该种机器使用三年后即被淘汰．机器有一易损零件，在购进机器时，可以额外购买这种零件作为备件，每个200元．在机器使用期间，如果备件不足再购买，则每个500元．现需决策在购买机器时应同时购买几个易损零件，为此搜集并整理了100台这种机器在三年使用期内更换的易损零件数，得下面柱状图：



以这100台机器更换的易损零件数的频率代替1台机器更换的易损零件数发生的概率，记表示2台机器三年内共需更换的易损零件数，表示购买2台机器的同时购买的易损零件数．

(I)求的分布列；

(II)若要求，确定的最小值；

(III)以购买易损零件所需费用的期望值为决策依据，在与之中选其一，应选用哪个？

**【答案】** (I)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|  |  |  |  |  |  |  |  |

(II) 19 (III)

【官方解答】(I)由柱状图并以频率代替概率可得，一台机器三年内需更换的易损零件数为8,9,10,11的概率分别为0．2,0．4,0．2,0．2．从而

，，



，

，

所以的分布列为

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|  |  |  |  |  |  |  |  |

(II) 由(I)得，，故的最小值为19

(III)记Y表示2台机器在购买易损零件上所需的费用(单位：元)

当时，



当时，

．

要令，，

则的最小值为19

可知当时所需要的费用的期望小于当时所需要的费用的期望∴故应选．

【民间解答】⑴ 每台机器更换的易损零件数为8，9，10，11

记事件为第一台机器3年内换掉个零件

记事件为第二台机器3年内换掉个零件

由题知，

设2台机器共需更换的易损零件数的随机变量为

则的可能的取值为16，17，18，19，20，21，22

















|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|  |  |  |  |  |  |  |  |

⑵要令，，

则的最小值为19．

⑶购买零件所需费用含两部分：

一部分为购买机器时购买零件的费用，另一部分为备件不足时额外购买的费用

当时，费用的期望为

当时，费用的期望为

所以应选用．