

统计、统计案例

一、单选题

1. 下列结论正确的是 ()

① “ $a = \frac{1}{4}$ ”是“对任意的正数 x , 均有 $x + \frac{a}{x} \geq 1$ ”的充分非必要条件.

② 随机变量 ξ 服从正态分布 $N(2, 2^2)$, 则 $D(\xi) = 2$

③ 线性回归直线至少经过样本点中的一个.

④ 若 10 名工人某天生产同一零件, 生产的件数是 15, 17, 14, 10, 15, 17, 17, 16, 14, 12, 设其平均数为 a , 中位数为 b , 众数为 c , 则有 $c > b > a$

A. ③④ B. ①② C. ①③④ D. ①④

2. 下列说法错误的是 ()

A. 用相关指数 R^2 来刻画回归效果, R^2 越小说明拟合效果越好

B. 已知随机变量 $X \sim N(5, \sigma^2)$, 若 $P(X < 1) = 0.1$, 则 $P(X \leq 9) = 0.9$

C. 某人每次投篮的命中率为 $\frac{3}{5}$, 现投篮 5 次, 设投中次数为随机变量 Y . 则 $E(2Y + 1) = 7$

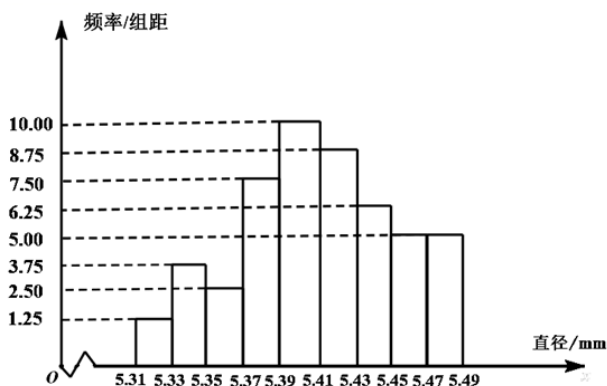
D. 对于独立性检验, 随机变量 K^2 的观测值 k 值越小, 判定“两分类变量有关系”犯错误的概率越大

3. 设一组样本数据 $x_1, x_2, \dots, x_n$ 的方差为 0.01, 则数据 $10x_1, 10x_2, \dots, 10x_n$ 的方差为 ()

A. 0.01 B. 0.1 C. 1 D. 10

4. 从一批零件中抽取 80 个, 测量其直径 (单位: mm), 将所得数据分为 9 组:

$[5.31, 5.33), [5.33, 5.35), \dots, [5.45, 5.47], [5.47, 5.49]$, 并整理得到如下频率分布直方图, 则在被抽取的零件中, 直径落在区间 $[5.43, 5.47)$ 内的个数为 ()



A. 10 B. 18 C. 20 D. 36

5. 在一组样本数据中, 1, 2, 3, 4 出现的频率分别为 p_1, p_2, p_3, p_4 , 且 $\sum_{i=1}^4 p_i = 1$, 则下面四种情形中, 对应样本的标准差最大的一组是 ()

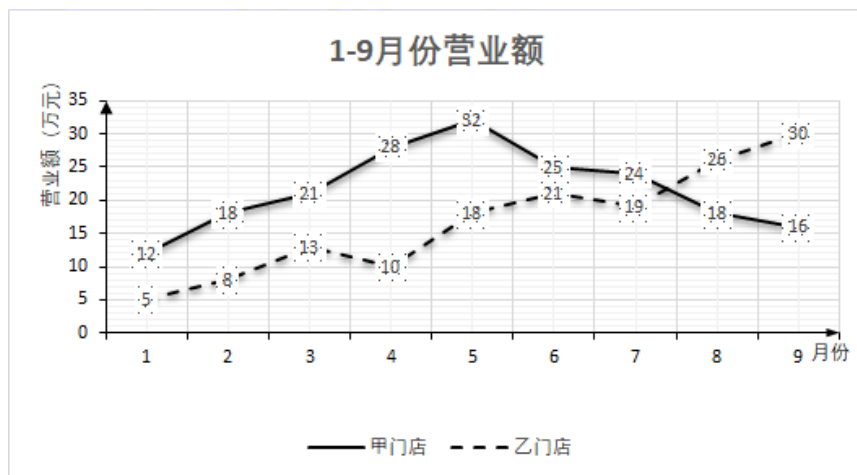
A. $p_1 = p_4 = 0.1, p_2 = p_3 = 0.4$

B. $p_1 = p_4 = 0.4, p_2 = p_3 = 0.1$

C. $p_1 = p_4 = 0.2, p_2 = p_3 = 0.3$

D. $p_1 = p_4 = 0.3, p_2 = p_3 = 0.2$

6. 某公司对旗下的甲、乙两个门店在 1 至 9 月份的营业额 (单位: 万元) 进行统计并得到如图折线图.



下面关于两个门店营业额的分析中, 错误的是 ()

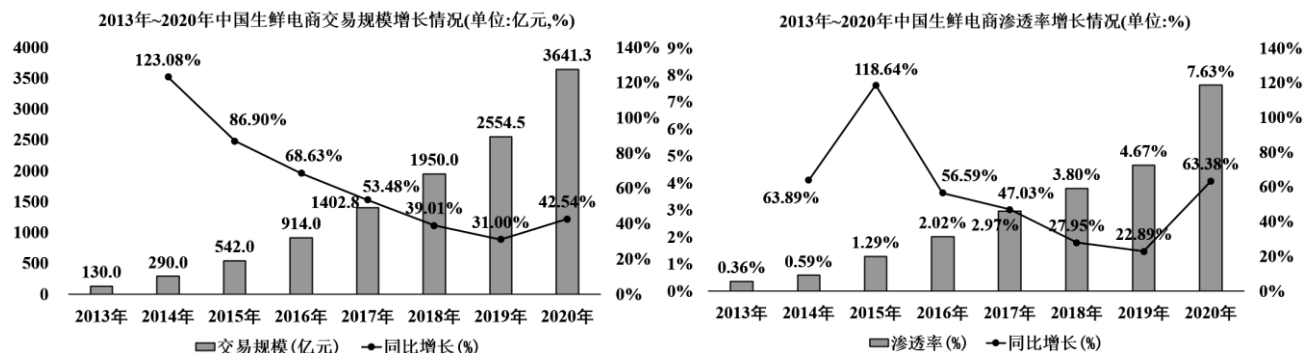
A. 甲门店的营业额折线图具有较好的对称性, 故而营业额的平均值约为 32 万元

B. 根据甲门店的营业额折线图可知, 该门店营业额的平均值在 $[20, 25]$ 内

C. 根据乙门店的营业额折线图可知, 其营业额总体是上升趋势

D. 乙门店在这 9 个月份中的营业额的极差为 25 万元

7. 近年来, 随着消费者习惯的变化, 吸引了更多的资本进入生鲜电商领域, 下表统计了 2013~2020 年中国生鲜电商交易规模增长情况与渗透率增长情况, 据此判断, 下列说法不正确的是 ()



A. 2019 年中国生鲜电商交易规模较 2018 年同比增长 31.00%, 同比增速较 2018 年进一步下滑

B. 2020 年生鲜电商交易规模同比增长的增速迎来回升

C. 2013-2020 年中国生鲜电商渗透率同比增长逐年上升

D. 可能受疫情催化的影响, 2020 年中国生鲜电商渗透率增速加快

8. 已知样本数据 x_1, x_2, x_3, x_4, x_5 , 该样本平均数为 4, 方差为 2, 现加入一个数 4, 得到新样本的平均数为 $\bar{x}$, 方差为 s^2 , 则 ()

A. $\bar{x} > 4, s^2 > 2$ B. $\bar{x} = 4, s^2 > 2$ C. $\bar{x} < 4, s^2 < 2$ D. $\bar{x} = 4, s^2 < 2$

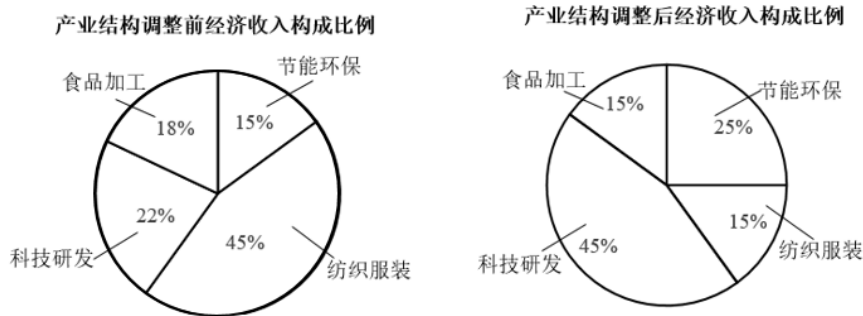
9. 对于数据组 $(x_i, y_i) (i=1, 2, 3, \dots, n)$, 如果由线性回归方程得到的对应于自变量 x_i 的估计值是 y_i , 那么将 $y_i - y_i$ 称为相应于点 (x_i, y_i) 的残差. 某工厂为研究某种产品产量 x (吨) 与所需某种原材料 y (吨) 的相关性, 在生产过程中收集 4 组对应数据 (x, y) 如下表所示:

x	3	4	5	6
y	2.5	3	4	m

根据表中数据, 得出 y 关于 x 的线性回归方程为 $y = 0.7x + a$, 据此计算出样本点处的残差为 -0.15 , 则表中 m 的值为 ()

A. 3.3 B. 4.5 C. 5 D. 5.5

10. 某经济开发区经过五年产业结构调整和优化, 经济收入比调整前翻了两翻, 为了更好的了解该开发区的经济收入变化情况, 统计了该开发区产业结构调整前后的经济收入构成比例, 得到如图所示的饼状图, 则下列选项正确的是 ()



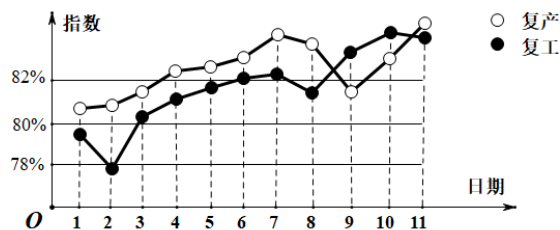
- ①产业结构调整后节能环保的收入与调整前的总收入一样多
- ②产业结构调整后科技研发的收入增幅最大
- ③产业结构调整后纺织服装收入相比调整前有所降低
- ④产业结构调整后食品加工的收入超过调整前纺织服装的收入

A. ②③ B. ③④ C. ①②③ D. ①②④

二、多选题

11. 我国新冠肺炎疫情进入常态化, 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 ()

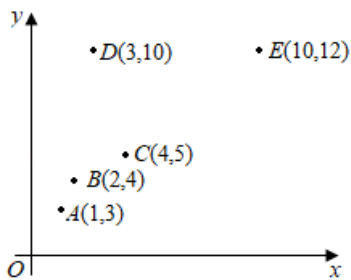
- A. 这 11 天复工指数和复产指数均逐日增加;
 B. 这 11 天期间, 复产指数增量大于复工指数的增量;
 C. 第 3 天至第 11 天复工复产指数均超过 80%;
 D. 第 9 天至第 11 天复产指数增量大于复工指数的增量;



12. 有一组样本数据 $x_1, x_2, \dots, x_n$, 由这组数据得到新样本数据 $y_1, y_2, \dots, y_n$, 其中 $y_i = x_i + c$ ($i=1, 2, \dots, n$), c 为非零常数, 则 ()

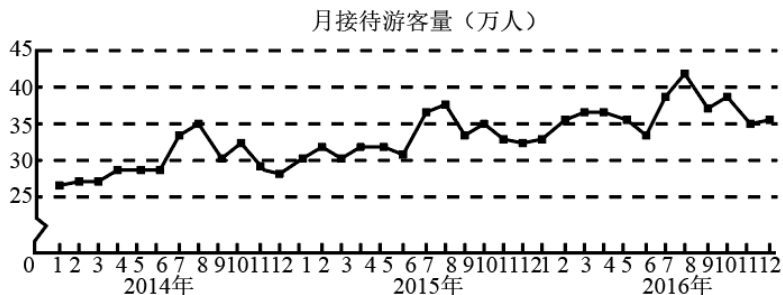
- A. 两组样本数据的样本平均数相同
 B. 两组样本数据的样本中位数相同
 C. 两组样本数据的样本标准差相同
 D. 两组样数据的样本极差相同

13. 有一散点图如图所示, 在 5 个 (x, y) 数据中去掉 $D(3, 10)$ 后, 下列说法不正确的是 ()



- A. 残差平方和变小
 B. 相关系数 r 变小
 C. 相关指数 R^2 变小
 D. 解释变量 x 与预报变量 y 的相关性变弱

14. 某城市为了解游客人数的变化规律, 提高旅游服务质量, 收集并整理了 2014 年 1 月至 2016 年 12 月期间月接待游客量(单位: 万人)的数据, 绘制了下面的折线图, 根据该折线图, 下列结论正确的是 ()

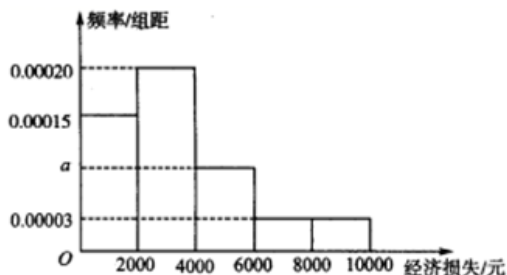


- A. 月接待游客量逐月增加
 B. 年接待游客量逐年增加

- C. 各年的月接待游客量高峰期大致在 7, 8 月份
D. 各年 1 月至 6 月的月接待游客量相对 7 月至 12 月, 波动性更小, 变化比较平稳

三、填空题

15. 有人批发黄豆 3000kg, 验得黄豆内混有少量豌豆, 两种豆子大小均匀、质量相等. 抽样取豆一把 226 颗, 数得豆内混有豌豆 3 颗, 则这批黄豆内混有豌豆约_____kg. (结果精确到个位数)
16. 2019 年, 习近平同志在福建、浙江等地调研期间, 提出了“绿水青山就是金山银山”的科学论断, 将绿色发展理念贯穿于治国理政思想之中. 为了响应中央号召, 某日深圳环保局随机抽查了本市市区汽车尾气排放污染物 x (单位: ppm) 与当天私家车路上行驶的时间 y (单位: 小时) 之间的关系, 从某主干路随机抽取 10 辆私家车, 根据测量数据的散点图可以看出 x 与 y 之间具有线性相关关系, 其回归直线方程为 $\hat{y} = 0.3x - 0.4$, 若该 10 辆车中有一辆私家车的尾气排放污染物为 6 (单位: ppm), 据此估计该私家车行驶的时间为_____小时.
17. 已知一组数据 $4, 2a, 3-a, 5, 6$ 的平均数为 4, 则 a 的值是_____.
18. 我国高铁发展迅速, 技术先进. 经统计, 在经停某站的高铁列车中, 有 10 个车次的正点率为 0.97, 有 20 个车次的正点率为 0.98, 有 10 个车次的正点率为 0.99, 则经停该站高铁列车所有车次的平均正点率的估计值为_____.
19. 由于受到网络电商的冲击, 某品牌的洗衣机在线下的销售受到影响, 承受了一定的经济损失, 现将 A 地区 200 家实体店该品牌洗衣机的月经济损失统计如图所示, 估算月经济损失的平均数为 m , 中位数为 n , 则 $m - n =$ _____.
20. 用线性回归模型求得甲、乙、丙 3 组不同的数据对应的 R^2 的值分别为 0.81, 0.98, 0.63, 其中_____ (填甲、乙、丙中的一个) 组数据的线性回归的效果最好.



四、解答题

21. 某厂接受了一项加工业务, 加工出来的产品(单位: 件)按标准分为 A, B, C, D 四个等级. 加工业务约定: 对于 A 级品、B 级品、C 级品, 厂家每件分别收取加工费 90 元, 50 元, 20 元; 对于 D 级品, 厂家每件要赔偿原料损失费 50 元. 该厂有甲、乙两个分厂可承接加工业务. 甲分厂加工成本费为 25 元/件, 乙分厂加工成本费为 20 元/件. 厂家为决定由哪个分厂承接加工业务, 在两个分厂各试加工了 100 件这种产品, 并统计了这些产品的等级, 整理如下:

甲分厂产品等级的频数分布表

等级	A	B	C	D
频数	40	20	20	20

乙分厂产品等级的频数分布表

等级	A	B	C	D
频数	28	17	34	21

- (1) 分别估计甲、乙两分厂加工出来的一件产品为 A 级品的概率；
- (2) 分别求甲、乙两分厂加工出来的 100 件产品的平均利润，以平均利润为依据，厂家应选哪个分厂承接加工业务？

22. 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100天空气中的PM2.5和SO₂浓度(单位: $\mu\text{g}/\text{m}^3$), 得下表:

SO ₂ PM2.5	[0,50]	(50,150]	(150,475]
[0,35]	32	18	4
(35,75]	6	8	12
(75,115]	3	7	10

(1) 估计事件“该市一天空气中PM2.5浓度不超过75, 且SO₂浓度不超过150”的概率;

(2) 根据所给数据, 完成下面的2×2列联表:

SO ₂ PM2.5	[0,150]	(150,475]
[0,75]		
(75,115]		

(3) 根据(2)中的列联表, 判断是否有99%的把握认为该市一天空气中PM2.5浓度与SO₂浓度有关?

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

$P(K^2 \geq k)$	0.050	0.010	0.001
k	3.841	6.635	10.828

23. 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？

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

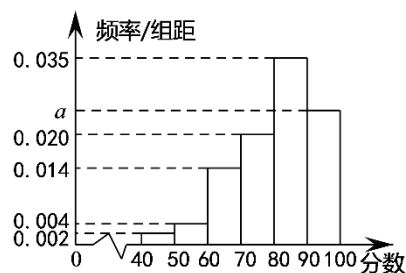
$P(K^2 \geq k)$	0.050	0.010	0.001
k	3.841	6.635	10.828

24. 2020 年新冠肺炎疫情期间, 某区政府为了解本区居民对区政府防疫工作的满意度, 从本区居民中随机抽取若干居民进行评分. 根据调查数据制成如下表格和频率分布直方图. 已知评分在 $[80, 100]$ 的居民有 600 人.

满意度评分	$[40, 60)$	$[60, 80)$	$[80, 90)$	$[90, 100]$
满意度等级	不满意	基本满意	满意	非常满意

(1) 求频率分布直方图中 a 的值及所调查的总人数;

(2) 定义满意指数 $\eta = \frac{\text{满意程度的平均分}}{100}$, 若 $\eta < 0.8$, 则防疫工作需要进行大的调整, 否则不需要大调整. 根据所学知识判断该区防疫工作是否需要进行大调整?



(3) 为了解部分居民不满意的原因, 从不满意的居民 (评分在 $[40, 50)$ 、 $[50, 60)$) 中用分层抽样的方法抽取 6 名居民, 倾听他们的意见, 并从 6 人中抽取 2 人担任防疫工作的监督员, 求这 2 人中仅有一人对防疫工作的评分在 $[40, 50)$ 内的概率.



25. 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的数量, 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 $(x_i, y_i)(i=1, 2, \dots, 20)$, 其中 x_i 和 y_i 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 $\sum_{i=1}^{20} x_i = 60$, $\sum_{i=1}^{20} y_i = 1200$, $\sum_{i=1}^{20} (x_i - \bar{x})^2 = 80$, $\sum_{i=1}^{20} (y_i - \bar{y})^2 = 9000$, $\sum_{i=1}^{20} (x_i - \bar{x})(y_i - \bar{y}) = 800$.

(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 $(x_i, y_i)(i=1, 2, \dots, 20)$ 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 $r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$, $\sqrt{2} \approx 1.414$.

26. 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	10.0	10.2	9.9	9.8	10.0	10.1	10.2	9.7
新设备	10.1	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 $\bar{x}$ 和 $\bar{y}$, 样本方差分别记为 s_1^2 和 s_2^2 .

(1) 求 $\bar{x}$, $\bar{y}$, s_1^2 , s_2^2 ;

(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 $\bar{y} - \bar{x} \geq 2\sqrt{\frac{s_1^2 + s_2^2}{10}}$, 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).



27. 盲盒里面通常装的是动漫、影视作品的周边, 或者设计师单独设计出来的玩偶. 由于盒子上没有标注, 购买者只有打开才会知道自己买到了什么, 因此这种惊喜吸引了众多年轻人, 形成了“盲盒经济”. 某款盲盒内可能装有某一套玩偶的 A、B、C 三种样式, 且每个盲盒只装一个.

(1) 若每个盲盒装有 A、B、C 三种样式玩偶的概率相同. 某同学已经有了 A 样式的玩偶, 若他再购买两个这款盲盒, 恰好能收集齐这三种样式的概率是多少?

(2) 某销售网点为调查该款盲盒的受欢迎程度, 随机发放了 200 份问卷, 并全部收回. 经统计, 有 30% 的人购买了该款盲盒, 在这些购买者当中, 女生占 $\frac{2}{3}$; 而在未购买者当中, 男生女生各占 50%. 请根据以上信

息填写表，并分析是否有95%的把握认为购买该款盲盒与性别有关？

	女生	男生	总计
购买			
未购买			
总计			200

参考公式： $K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$ ，其中 $n = a+b+c+d$ 。

参考数据：

$P(K^2 \geq k_0)$	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

(3) 该销售网点已经售卖该款盲盒6周，并记录了销售情况，如表：

周数 x	1	2	3	4	5	6
盒数 y	16	--	23	25	26	30

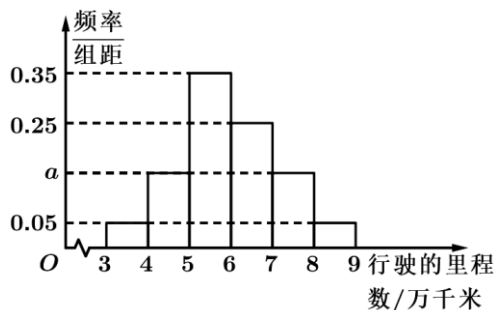
由于电脑故障，第二周数据现已丢失，该销售网点负责人决定用第4、5、6周的数据求线性回归方程，再用第1、3周数据进行检验。

①请用4、5、6周的数据求出 y 关于 x 的线性回归方程 $\hat{y} = bx + a$ ；（注： $\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$ ，

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

②若由线性回归方程得到的估计数据与所选出的检验数据的误差均不超过2盒，则认为得到的线性回归方程是可靠的，试问①中所得的线性回归方程是否可靠？

28. “碳达峰”“碳中和”成为今年全国两会热词,被首次写入政府工作报告.碳达峰就是二氧化碳的排放不再增长,达到峰值之后再慢慢减下去;碳中和是指在一定时间内直接或间接产生的温室气体排放总量通过植树造林、节能减排等方式,以抵消自身产生的二氧化碳排放量,实现二氧化碳“零排放”.2020年9月,中国向世界宣布了2030年前实现碳达峰,2060年前实现碳中和的目标.某城市计划通过绿色能源(光伏、风电、核能)替代煤电能源,智慧交通,大力发展新能源汽车以及植树造林置换大气中的二氧化碳实现碳中和.该城市某研究机构统计了若干汽车5年内所行驶的里程数(万千米)的频率分布直方图,如图.



- (1) 求 a 的值及汽车 5 年内所行驶里程的平均值(同一组中的数据用该组区间的中点值为代表).
- (2) 据“碳中和罗盘”显示:一辆汽车每年行驶 1 万千米的排碳量需要近 200 棵树用 1 年时间来吸收.根据频率分布直方图,该城市每一辆汽车平均需要多少棵树才能够达到“碳中和”?
- (3) 该城市为了减少碳排放量,计划大力推动新能源汽车,关于车主购买汽车时是否考虑对大气污染的因素,对 300 名车主进行了调查,这些车主中新能源汽车车主占 $\frac{1}{6}$,且这些车主在购车时考虑大气污染因素的占 20%,燃油汽车车主在购车时考虑大气污染因素的占 10%.根据以上统计情况,补全下面 2×2 列联表,并回答是否有 99% 的把握认为购买新能源汽车与考虑大气污染有关.

	考虑大气污染	没考虑大气污染	合计
新能源汽车车主			
燃油汽车车主			
合计			

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

$P(K^2 \geq k_0)$	0.10	0.025	0.010	0.005	0.001
k_0	2.706	5.024	6.635	7.879	10.828

29. 2021 年 2 月 1 日教育部办公厅关于加强中小学生手机管理工作的通知中明确“中小学生原则上不得将个人手机带入校园”，为此某学校开展了一项“你能否有效管控手机”调查，并从调查表中随机抽取 200 名学生(其中男、女生各占一半)的样本数据，其 2×2 列联表如下：

性别	能管控	不能管控	总计
男	30		
女			
总计	90		200

(1) 完成上述 2×2 列联表，并判断是否有 99.9% 的把握认为能否管控手机与性别有关？

(2) 若学生确因需要带手机进入校园需向学校有关部门报告，该校为做好这部分学生的手机管理工作，学校团委从能管控的学生中按样本中的比例抽取了 6 名学生组成一个团队.

①从该团队中选取 2 名同学作个人经验介绍，求选取的 2 人中恰有一名女生的概率.

②某老师根据以往学生自从玩手机导致成绩下降的数据构建了一个函数模型： $I(t) = \left(1 - \frac{1}{1 + e^{-2(t-30)}}\right)k$ ，其中 k 为没有玩手机时的原始成绩分数， $I(t)$ 是开始玩手机 t 天后的成绩，试根据该模型，求某学生自从玩手机后经过多少天成绩大约下滑到原来成绩的一半？

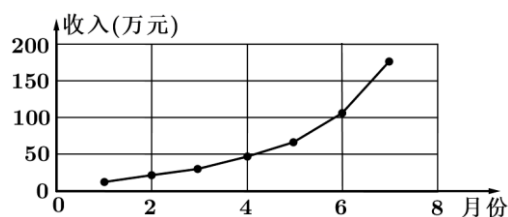
参考公式及数据： $K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$ ，其中 $n=a+b+c+d$.

$P(K^2 \geq k_0)$	0.050	0.010	0.001
k_0	3.841	6.635	10.828

30. 随着时代的发展,科技的进步,“网购”已经成为一种新的购物潮流,足不出户就可以在网上买到自己想要的东西,而且两三天就会送到自己的家门口,某研究机构调查了某地去年第一个月至第七个月的网店销售收入如表:

时间(月份)	1	2	3	4	5	6	7
收入(万元)	8	15	24	42	66	105	182

根据以上数据绘制散点图.



(1) 为了更深入的了解网购发展趋势,机构需要派出人员进行实地考察,拟从 A, B, C, D, E 五名员工中随机抽取 2 人前往,则 A, B 至少有一人被抽到的概率是多少?

(2) 根据散点图判断, $y = ax + b$ 与 $y = c \cdot d^x$ 哪一个适宜作为网店销售收入 y 关于月份 x 的回归方程类型?(给出判断即可,不必说明理由)并根据你判断的结果及表中的数据,求出 y 关于 x 的回归方程;(结果保留两位小数)

(3) 根据(2)中求得的回归方程预测 8 月份该地区的网店销售收入.

参考数据与参考公式:

$\sum_{i=1}^7 y_i$	$\sum_{i=1}^7 \lg y_i$	$\sum_{i=1}^7 x_i y_i$	$\sum_{i=1}^7 x_i u_i$	$10^{0.72}$	$10^{1.76}$
442	11.18	2512	50.93	5.25	57.54

其中 $u = \lg y$, $u_i = \lg y_i$, $\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}$, $\hat{a} = \bar{y} - \hat{\beta} \bar{x}$.

参考答案

1. D 2. A 3. C 4. B 5. B 6. A 7. C 8. D 9. B 10. D

11. CD 12. CD 13. BCD 14. BCD

15. 40 16. 1.4 17. 2 18. 0.98 19. 360 20. 乙

21. 【解析】(1) 由表可知, 甲厂加工出来的一件产品为 A 级品的概率为 $\frac{40}{100} = 0.4$, 乙厂加工出来的一件产品为 A 级品的概率为 $\frac{28}{100} = 0.28$;

(2) 甲分厂加工 100 件产品的总利润为 $40 \times (90 - 25) + 20 \times (50 - 25) + 20 \times (20 - 25) - 20 \times (50 + 25) = 1500$ 元,

所以甲分厂加工 100 件产品的平均利润为 15 元每件;

乙分厂加工 100 件产品的总利润为

$$28 \times (90 - 20) + 17 \times (50 - 20) + 34 \times (20 - 20) - 21 \times (50 + 20) = 1000 \text{ 元},$$

所以乙分厂加工 100 件产品的平均利润为 10 元每件.

故厂家选择甲分厂承接加工任务.

22. 【解析】(1) 由表格可知, 该市 100 天中, 空气中的 $PM_{2.5}$ 浓度不超过 75, 且 SO_2 浓度不超过 150 的天数有 $32 + 6 + 18 + 8 = 64$ 天,

所以该市一天中, 空气中的 $PM_{2.5}$ 浓度不超过 75, 且 SO_2 浓度不超过 150 的概率为 $\frac{64}{100} = 0.64$;

(2) 由所给数据, 可得 2×2 列联表为:

SO_2 $PM_{2.5}$	$[0, 150]$	$(150, 475]$	合计
$[0, 75]$	64	16	80
$(75, 115]$	10	10	20
合计	74	26	100

(3) 根据 2×2 列联表中的数据可得

$$K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} = \frac{100 \times (64 \times 10 - 16 \times 10)^2}{80 \times 20 \times 74 \times 26} = \frac{3600}{481} \approx 7.4844 > 6.635,$$

因为根据临界值表可知, 有 99% 的把握认为该市一天空气中 $PM_{2.5}$ 浓度与 SO_2 浓度有关.

23. 【解析】(1) 甲机床生产的产品中的一级品的频率为 $\frac{150}{200} = 75\%$,

乙机床生产的产品中的一级品的频率为 $\frac{120}{200} = 60\%$.

$$(2) K^2 = \frac{400(150 \times 80 - 120 \times 50)^2}{270 \times 130 \times 200 \times 200} = \frac{400}{39} > 10 > 6.635,$$

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.

24. 【解析】(1) 由频率分布直方图得: $(0.002 + 0.004 + 0.014 + 0.02 + 0.035 + a) \times 10 = 1$, 解得 $a = 0.025$,

设总共调查了 n 人, 则 $n = \frac{600}{(0.035 + 0.025) \times 10} = 1000$, 即调查的总人数为 1000 人;

(2) 由频率分布直方图知,

满意程度的平均分为 $\bar{x} = 45 \times 0.02 + 55 \times 0.04 + 65 \times 0.14 + 75 \times 0.2 + 85 \times 0.35 + 95 \times 0.25 = 80.7$,

所以, 满意指数 $\eta = \frac{80.7}{100} = 0.807 > 0.8$,

因此, 该区防疫工作不需要大的调整;

(3) 由题意可知, 评分在 $[40, 50)$ 、 $[50, 60)$ 的频率之比为 $\frac{0.02}{0.04} = \frac{1}{2}$,

所以, 所抽取的 6 人中评分在 $[40, 50)$ 的人数为 $6 \times \frac{1}{3} = 2$, 分别记为 a 、 b ,

评分在 $[50, 60)$ 的人数为 $6 \times \frac{2}{3} = 4$, 分别记为 A 、 B 、 C 、 D ,

抽取 2 人的基本事件为: ab 、 aA 、 aB 、 aC 、 aD 、 bA 、 bB 、 bC 、 bD 、 AB 、 AC 、 AD 、 BC 、 BD 、 CD , 共 15 个,

而仅有一人来自 $[40, 50)$ 的基本事件有: aA 、 aB 、 aC 、 aD 、 bA 、 bB 、 bC 、 bD , 共 8 个,

因此, 所抽取的 2 人中仅有一人对防疫工作的评分在 $[40, 50)$ 内的概率为 $P = \frac{8}{15}$.

25. 【解析】(1) 样区野生动物平均数为 $\frac{1}{20} \sum_{i=1}^{20} y_i = \frac{1}{20} \times 1200 = 60$,

地块数为 200, 该地区这种野生动物的估计值为 $200 \times 60 = 12000$

(2) 样本 $(x_i, y_i) (i=1, 2, \dots, 20)$ 的相关系数为

$$r = \frac{\sum_{i=1}^{20} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{20} (x_i - \bar{x})^2 \sum_{i=1}^{20} (y_i - \bar{y})^2}} = \frac{800}{\sqrt{80 \times 9000}} = \frac{2\sqrt{2}}{3} \approx 0.94$$

(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.

$$26. \text{【解析】} (1) \bar{x} = \frac{9.8+10.3+10+10.2+9.9+9.8+10+10.1+10.2+9.7}{10} = 10,$$

$$\bar{y} = \frac{10.1+10.4+10.1+10+10.1+10.3+10.6+10.5+10.4+10.5}{10} = 10.3,$$

$$s_1^2 = \frac{0.2^2+0.3^2+0+0.2^2+0.1^2+0.2^2+0+0.1^2+0.2^2+0.3^2}{10} = 0.036,$$

$$s_2^2 = \frac{0.2^2+0.1^2+0.2^2+0.3^2+0.2^2+0+0.3^2+0.2^2+0.1^2+0.2^2}{10} = 0.04.$$

$$(2) \text{依题意, } \bar{y} - \bar{x} = 0.3 = 2 \times 0.15 = 2\sqrt{0.15^2} = 2\sqrt{0.0225}, \quad 2\sqrt{\frac{0.036+0.04}{10}} = 2\sqrt{0.0076},$$

$$\bar{y} - \bar{x} \geq 2\sqrt{\frac{s_1^2+s_2^2}{10}}, \text{ 所以新设备生产产品的该项指标的均值较旧设备有显著提高.}$$

27. 【解析】(1) 由题意, 基本事件空间为:

$$\Omega = \{(A,A), (A,B), (A,C), (B,A), (B,B), (B,C), (C,A), (C,B), (C,C)\}$$

基本事件个数共 9 个.

恰好能收集齐这三种样式为事件 D , 则 $D = \{(B,C), (C,B)\}$, 事件个数共 2 个.

则恰好能收集齐这三种样式的概率 $P = \frac{2}{9}$;

(2) 2×2 列联表如下:

	女生	男生	总计
购买	40	20	60
未购买	70	70	140
总计	110	90	200

$$K^2 = \frac{200(40 \times 70 - 20 \times 70)^2}{110 \times 90 \times 60 \times 140} \approx 4.714,$$

又 $4.714 > 3.841$,

故有 95% 的把握认为购买该款盲盒与性别有关;

(3) ①由数据, 求得 $\bar{x} = 5$, $\bar{y} = 27$.

$$\hat{b} = \frac{(4-5)(25-27) + (5-5)(26-27) + (6-5)(30-27)}{(4-5)^2 + (5-5)^2 + (6-5)^2} = \frac{5}{2}.$$

$$\hat{a} = 27 - \frac{5}{2} \times 5 = 14.5.$$

$\therefore y$ 关于 x 的线性回归方程为 $\hat{y} = 2.5x + 14.5$;

②当 $x=1$ 时, $\hat{y} = 2.5 \times 1 + 14.5 = 17$, $|17 - 16| < 2$;

同样, 当 $x=3$ 时, $\hat{y} = 2.5 \times 3 + 14.5 = 22$, $|22 - 23| < 1$.

$\therefore$ ①中所得的线性回归方程是可靠的.

28. 【解析】(1) 由 $(0.05 + a + 0.35 + 0.25 + a + 0.05) \times 1 = 1$, 解得 $a = 0.15$.

设 $\bar{x}$ 为汽车 5 年内所行驶里程的平均值, 则

$$\bar{x} = 3.5 \times 0.05 + 4.5 \times 0.15 + 5.5 \times 0.35 + 6.5 \times 0.25 + 7.5 \times 0.15 + 8.5 \times 0.05 = 5.95 \text{ (万千米)}.$$

(2) 由 (1) 可知, 一辆汽车 1 年内所行驶里程的平均值为 $\frac{5.95}{5} = 1.19$ (万千米).

因为一辆汽车每年行驶 1 万千米的排碳量需要近 200 棵树用 1 年时间来吸收,

所以每一辆汽车平均需要 $1.19 \times 200 = 238$ (棵) 树才能够达到“碳中和”.

(3) 补全的 2×2 列联表如下:

	考虑大气污染	没考虑大气污染	合计
新能源汽车车主	10	40	50
燃油汽车车主	25	225	250
合计	35	265	300

$$\text{所以 } K^2 = \frac{300 \times (10 \times 225 - 25 \times 40)^2}{35 \times 265 \times 250 \times 50} \approx 4.04.$$

因为 $4.04 < 6.635$,

所以没有 99% 的把握认为购买新能源汽车与考虑大气污染有关.

29. 【解析】(1) 2×2 列联表如下

性别	能管控	不能管控	总计
男	30	70	100
女	60	40	100
总计	90	110	200

$$k^2 = \frac{200(30 \times 40 - 70 \times 60)^2}{100 \times 100 \times 90 \times 110} = \frac{200}{11} \approx 18.182 > 10.828,$$

所以有 99.9% 的把握认为能否管控手机与性别有关.

$$(2) \textcircled{1} \text{ 抽样比} = \frac{6}{90} = \frac{1}{15},$$

男生抽取 $30 \times \frac{1}{15} = 2$ 人, 设为 A, B , 女生抽取 $60 \times \frac{1}{15} = 4$ 人, 设为 a, b, c, d ,

从 6 人中选 2 人, 共有: $AB, Aa, Ab, Ac, Ad,$

$Ba, Bb, Bc, Bd, ab, ac, ad, bc, bd, cd$, 共 15 个基本事件,

设事件 A : 选取的 2 人中恰有一名女生, 事件 A 包含 8 个基本事件,

$$P(A) = \frac{8}{15}.$$

$\textcircled{2}$ 由题知: $\left(1 - \frac{1}{1+e^{-2(t-30)}}\right)k = \frac{1}{2}k$, 化简得到 $e^{-2(t-30)} = 1$, 解得 $t = 30$.

所以某学生自从玩手机后经过 30 天成绩大约下滑到原来成绩的一半.

30. 【解析】(1) 从 A, B, C, D, E 五名员工中随机抽取 2 人前往, 所有事件如下:

$(A, B), (A, C), (A, D), (A, E), (B, C), (B, D), (B, E), (C, D), (C, E), (D, E)$, 共 10 个,

其中 A, B 至少有一人的事件有: $(A, B), (A, C), (A, D), (A, E), (B, C), (B, D), (B, E)$, 共 7 个,

所以 A, B 至少有一人被抽到的概率是 $\frac{7}{10}$;

(2) 根据散点图知, 指数型 $y = c \cdot d^x$ 适合作为网店销售收入 y 关于月份 x 的回归方程类型.

由 $y = c \cdot d^x$ 两边取对数, 得 $\lg y = \lg c + x \cdot \lg d$, 而 $u = \lg y$, 设 $p = \lg c, q = \lg d$,

则回归方程即 $u = p + qx$, $p = \bar{u} - q \cdot \bar{x}$,

因为 $\bar{x} = \frac{1}{7} \cdot (1+2+\dots+6+7) = 4$, $\bar{u} = \frac{1}{7} \sum_{i=1}^7 \lg y_i = \frac{1}{7} \times 11.18 \approx 1.60$,

$\sum_{i=1}^7 x_i u_i = 50.93$, $\sum_{i=1}^7 x_i^2 = 1^2 + 2^2 + \dots + 7^2 = 140$,

所以 $q = \frac{\sum_{i=1}^7 x_i u_i - 7 \bar{x} \bar{u}}{\sum_{i=1}^7 x_i^2 - 7 \bar{x}^2} = \frac{50.93 - 4 \times 11.18}{140 - 7 \times 16} \approx 0.22$, 即 $p = 1.60 - 0.22 \times 4 = 0.72$,

所以 $p = \lg c = 0.72, q = \lg d = 0.22 = 1.76 \times \frac{1}{8}$,

即 $c = 10^{0.72} = 5.25, d = 10^{0.22} = 10^{1.76 \times \frac{1}{8}} = 57.54^{\frac{1}{8}}$

网店销售收入 y 关于月份 x 的回归方程为 $y = 5.25 \times 57.54^{\frac{x}{8}}$;

(3) $x = 8$ 时, $y = 5.25 \times 57.54 = 302.085$, 所以 8 月份该地区的网店销售收入大约为 302.085 (万元).