**14.4 用样本估计总体**

**【考点梳理】**

**考点一、平均数、中位数、众数的应用**

**考点二、极差、方差、标准差的计算与应用**

**考点三：各数据加减乘除对方差、平均数的影响**

**考点四：频率分布直方图中的方差、标准差**

**考点五：百分位数**

**考点六：用样本估计总体的综合问题**

**【知识梳理】**

**知识点一、用样本估计总体的集中趋势参数**

1.平均数

(1)均值：一般地、把总体中所有数据的算术平均数称为总体的均值，它通常可以代表总体的水平。

(2)平均数：如果给定的一组数是，，…，是则这组数的平均数为 即 。

(3)一般地，若取值为，，…，的频率分别为，，…，则其平均数为+++……+。

(4)如果将总体分为k层，第j层抽取的样本为，，……，第j层的样本量为，样本平均数为，j=1，2，…，k。记，则所有数据的样本平均数为。

2.众数

一般地，我们将一组数据中出现次数最多的那个数据叫作该组数据的众数。众数是一种刻画数据集中趋势的度量值。

3.中位数

一般地，将一组数据按照从小到大的顺序排成一列，如果数据的个数为奇数，那么排在正中间的数据就是这组数据的中位数；如果数据的个数为偶数，那么，排在正中间的两个数据的平均数即为这组数据的中位数。

**技巧归纳：平均数、中位数和众数的特点**

(1)平均数的大小与一组数据里每一个数据均有关系，任何一个数据的变化都会引起平均数的变化.容易受极端值的影响；

(2)中位数仅与数据排列位置有关，某些数据的变动对中位数没有影响，中位数可能在所给数据中（数据个数为奇数或数据个数为偶数且中间两数相等），也可能不在所给的数据中（数据个数为偶数且中间两数不相等）。当一组数据的个别数据变动较大时，可以用中位数描述其集中趋势

(3)众数只与这组数据中的部分数据有关，当一组数据中有多个数据多次重复出现时，众数往往更能反映问题。

4.频率直方图中的数字特征

(1)样本平均数可以用每个小矩形底边中点的横坐标与小矩形的面积的乘积之和近似代替。

(2)样本中位数可依据“中位数左边和右边的直方图的面积相等”来求出。

(3)样本众数可以用最高的矩形底边中点的横坐标近似代替。

**知识点二、用样本估计总体的离散程度参数**

1.极差我们把一组数据的最大值与最小值的差称为极差。

2.方差和标准差

假设一组数据是，，…，用表示这组数据的平均数，则为这组数据的方差。有时为了计算方差的方便，我们还把方差写成的形式。我们对方差开平方，取它的算

术平方根，即为这组数据的标准差。

3.总体方差(标准差)和样本方差(标准差)

①如果总体中所有个体的变量值分别为，，…，，总体平均数为，则称

为总体方差， 为总体标准差。

②如果总体的N个变量值中，不同的值共有k(k≤N)个，不妨记为，，…，，其中出现的频数为(i=1，2，…，k)，

则总体方差为 。

如果一个样本中个体的变量值分别为，，…，，样本平均数为，则称为样本方差， 为样本标准差。

3.标准差

因为方差与原始数据的单位不同，且平方后可能夸大了离差的程度，所以我们将方差的算术平方根称为这组数据的标准差.标准差刻画了数据的离散程度或波动幅度，标准差越大，数据的离散程度越大；标准差越小，数据的离散程度越小。

4.一般地，若取值为，，…，的频率分别为，，…，，则其方差为 。

5.分层抽样数据的方差

一般地，如果总体分为k层，第j层抽取的样本为，，……，第j层的样本量为，样本平均数为，样本方差为，=1,2，……，k。记。

**知识点三、百分位数**

1.k百分位数的特点

一般地，一组数据的k百分位数是这样一个值，它使得这组数据中至少有k%的数据小于或等于，且至少有(100-k)%的数据大于或等于

2.计算有n个数据的大样本的k百分位数的一般步骤

第1步，将所有数值按从小到大的顺序排列；

第2步，计算

第3步，如果结果为整数，那么k百分位数位于第k· 位和下一位数之间，通常取这两个位置上数值的平均数为k百分位数；

第4步，如果 不是整数，那么将其向上取整(即其整数部分加上1)，在该位置上的数值即为k百分位数。

3.四分位数常用的分位数有25百分位数，50百分位数(即中位数)和75百分位数。这三个分位数把一组由小到大排列后的数据分成四等份，因此称为四分位数。其中25百分位数称为下四分位数，75百分位数称为上四分位数。

**【题型归纳】**

**题型一、平均数、中位数、众数的应用**

1．（20-21高一下·河北·期末）某校举行校园歌手大赛，6位评委对某选手的评分分别为9.2，9.5，8.8，9.9，8.9，9.5，设该选手得分的平均数为*x*，中位数为*y*，众数为*z*，则（ ）

A． B． C． D．

【答案】A

【分析】根据平均数，中位数，众数的概念，分别求出，即可求出结果.

【详解】由题意可得，，，，

则．

故选：A.

2．（2024高一上·湖南邵阳·竞赛）如果将一组数据5、4、6、5、4、13、5依次重复写10次，会得到70个数组成的一组新数据，关于这组新数据的中位数、众数、平均数，下列说法正确的是（ ）

A．中位数和众数都是5 B．众数是10

C．中位数是4 D．中位数、平均数都是5

【答案】A

【分析】根据平均数，众数，中位数的定义和性质即可求解.

【详解】将这组数据从小到大的顺序排列为4，4，5，5，5，6，13，处于中间位置的那个数是5，

每个数字重复写10次，5依然处于中间位置，由中位数的定义可知，这组新数据的中位数是5，

这组新数据中出现次数最多的数是5，出现了30次，所以众数为5，故A正确，BC错误．

平均数，故D错误．

故选：A

3．（23-24高二上·四川成都·期中）已知甲、乙两支篮球队各6名队员某场比赛的得分数据（单位：分）从小到大排列为如下：甲队：7，12，12，20，，31；乙队：8，9，19，，25，28.这两组数据的中位数相等，且平均值也相等，则*x*和*y*的值分别为（ ）

A．2和3 B．0和2 C．0和3 D．2和4

【答案】C

【分析】

根据甲乙中位数相同求得*y*的值，再根据平均数相同即可求得*x*，即得答案.

【详解】由题意得甲的平均数为，

乙的平均数为，

而甲的中位数为，故乙的中位数为，即，

故，

故选：C

**题型二、极差、方差、标准差的计算与应用**

4．（23-24高一下·浙江宁波·期中）甲乙丙丁四位同学各掷5次骰子并记录点数，方差最大的是（ ）

甲：4 5 4 5 5 乙：4 2 3 4 3

丙：2 3 2 3 4 丁：6 1 2 6 1

A．甲 B．乙 C．丙 D．丁

【答案】D

【分析】根据条件，利用方差的定义，分别求出甲乙丙丁四位同学各掷5次骰子的点数的方差，即可求出结果.

【详解】由题知，所以，

，所以，

，所以，

，所以，

所以方差最大的是丁，

故选：D.

5．（2024·全国·模拟预测）已知总体划分为3层，通过分层随机抽样，各层抽取的样本容量分别为，，，样本平均数分别为，，，样本方差分别为，，，若，则（ ）

A．

B．

C．总体样本平均数

D．当时，总体方差

【答案】D

【分析】根据样本平均数以及方差的定义，即可判断A、B项；计算可判断C；根据分层抽样，总体方差的求解，计算即可得出D.

【详解】对于A、B项，由于样本容量与样本平均数、样本方差之间并不是成某种比例关系，所以选项A，B错误；

对于C项，设，

则总体样本平均数，所以选项C错误；

对于D项，当时，总体样本平均数，

所以总体方差，所以选项D正确.

故选：D．

6．（2024·山西朔州·一模）已知一组样本数据分别为，若这组数据的平均数为4，则数据，的方差为（ ）

A． B．8 C．16 D．24

【答案】D

【分析】先求出，再结合方差公式可得正确选项.

【详解】由题设有，故，

故，

故数据，的方差为：

，

故选：D.

**题型三：各数据加减乘除对方差、平均数的影响**

7．（2024·湖南长沙·二模）已知样本数据的平均数和标准差均为4，则数据的平均数与方差分别为（ ）

A． B． C． D．

【答案】B

【分析】根据样本数据同加上一个数和同乘以一个数后的新数据的平均值和方差的性质，即可求得答案.

【详解】由题意知样本数据的平均数和标准差均为4，则的方差为16，

则的平均数为，方差为，

故的平均数为，方差，

故选：B

8．（23-24高一下·浙江绍兴·期中）已知样本数据的平均数和标准差均为4，则数据的平均数与标准差分别为（ ）

A． B． C． D．

【答案】A

【分析】根据样本数据同加上一个数和同乘以一个数后的新数据的平均值和方差的性质即可求解.

【详解】由题意知，样本数据的标准差为，

所以样本数据的方差为16，

因为样本数据的平均数为和方差为，

所以的平均数为，

的方差为，

所以的标准差为，

故选：A.

9．（2023·全国·模拟预测）已知样本数据，，…，的平均数为，方差为，若样本数据，，…，的平均数为，方差为，则（ ）

A．5 B． C．1或5 D．或

【答案】C

【分析】由平均数和方差的运算性质即可求解.

【详解】由方差的性质，得，，…，的方差为，故，解得．

由平均数的性质，得，，…，的平均数为，故，

解得或5.

故选：C．

**题型四：频率分布直方图中的方差、标准差**

10．（2024·陕西西安·模拟预测）某校为了解在校学生对中国传统文化的传承认知情况，随机抽取了100名学生进行中国传统文化知识考试，并将这100名学生成绩整理得到如下频率分布直方图．根据此频率分布直方图（分成，，，，，六组），下列结论中不正确的是（ ）



A．图中的

B．若从成绩在，，内的学生中采用分层抽样抽取10名学生，则成绩在内的有3人

C．这100名学生成绩的中位数约为65

D．若同一组中的数据用该组区间的中点值作代表，则这100名学生的平均成绩约为68.2

【答案】C

【分析】根据频率分布直方图的特点逐个进行分析计算即可.

【详解】由，得，所以A正确；

这100名学生中成绩在，，内的频率分别为0.2，0.12，0.08，所以采用分层抽样抽取的10名学生中成绩在内的有人，故B正确；

根据频率分布直方图，可知这100名学生成绩的中位数在之间，设中位数为，则，所以，故C错误；

根据频率分布直方图的平均数的计算公式，可得，D正确．

故选：C

11．（22-23高一下·吉林长春·期末）某校1500名学生参加交通安全知识竞赛，随机抽取了100名学生的竞赛成绩（单位：分），成绩的频率分布直方图如图所示，则下列说法正确的是（ ）

A．频率分布直方图中的值为0.0045

B．估计这100名学生竞赛成绩的中位数为73

C．估计这100名学生竞赛成绩的众数为80

D．估计总体中成绩落在内的学生人数为525

【答案】D

【分析】先根据频率之和为1可得，进而可得每组的频率，再结合统计相关知识逐项分析判断即可得到结果.

【详解】对于选项A，因为，可得，故选项A错误；

对于选项B，可知每组的频率依次为，

设100名学生竞赛成绩的中位数为，则，

解得，故选项B错误；

对于选项C，因为的频率最大，所以这100名学生竞赛成绩的众数为75，故选项C错误；

对于选项D，总体中成绩落在内的学生人数为，故选项D正确.

故选：D.

12．（2023·天津河北·二模）某校举行知识竞赛，对全校参赛的1000名学生的得分情况进行了统计，把得分数据按，，，，分成5组，得到如图所示的频率分布直方图，则下列说法不正确的是（ ）



A．图中的*x*值为0.020 B．得分在的人数为400

C．这组数据的极差为50 D．这组数据的平均数的估计值为77

【答案】C

【分析】根据频率分布直方图中所有长方形的面积和为1，以及极值、频数以及平均数的计算，对每个选项进行逐一分析，即可判断和选择.

【详解】对于A，由，可解得，故选项A正确；

对于B，得分在80分及以上的人数的频率为，

故人数为，故选项B正确；

对于C，频率分布直方图无法看出这组数据的最大值和最小值，故选项C不正确；

对于D，这组数据的平均数的估计值为：，故选项D正确.

故选：C.

**题型五：百分位数**

13．（2024高一下·全国·专题练习）小明同学统计了他最近10次的数学考试成绩，得到的数据分别为92，85，87，91，95，90，88，83，98，96．则这组数据的分位数是（ ）

A．92 B．91.5

C．91 D．90

【答案】B

【分析】利用分位数的定义及求解方法直接计算即得.

【详解】将10次的数学考试成绩由小到大排序依次为83，85，87，88，90，91，92，95，96，98，

由，得这组数据的分位数是．

故选：B

14．（2024·天津·二模）某校举办了数学知识竞赛，把1000名学生的竞赛成绩（满分100分，成绩取整数）按，，，分成四组，并整理成如图所示的频率分布直方图，则下列说法正确的为（ ）



A．的值为0.015 B．估计这组数据的众数为80

C．估计这组数据的第60百分位数为87 D．估计成绩低于80分的有350人

【答案】C

【分析】利用频率分布直方图的性质可判定A，利用众数、百分位数的求法可判定B、C，根据频率分布直方图计算可估计总体判定D.

【详解】易知，解得，所以A错误；

由频率分布直方图可知众数落在区间，用区间中点表示众数即85，所以B错误；

由频率分布直方图可知前两组频率之和为，

前三组频率之和为，

故第60百分位数落在区间，设第60百分位数为，

则，解得，所以C正确；

成绩低于80分的频率为，所以估计总体有，故D错误.

故选：C.

15．（2024·安徽安庆·二模）在一次学科核心素养能力测试活动中，随机抽取了100名同学的成绩（评分满分为100分），将所有数据按，，，，，进行分组，整理得到频率分布直方图如图所示，则估计这次调查数据的第64百分位数为（ ）



A．80 B．78 C．76 D．74

【答案】B

【分析】借助百分位数的定义计算即可得.

【详解】由，

，

故这次调查数据的第64百分位数位于之间，

设这次调查数据的第64百分位数为，

则有，解得.

故选：B.

**题型六：用样本估计总体的综合问题**

16．（2024高一下·全国·专题练习）为了解学生的周末学习时间（单位：小时），高一年级某班班主任对本班40名学生某周末的学习时间进行了调查，将所得数据整理绘制出如图所示的频率分布直方图，根据直方图所提供的信息：

(1)求该班学生周末的学习时间不少于20小时的人数；

(2)①估计这40名同学周末学习时间的25%分位数；

②将该班学生周末学习时间从低到高排列，那么估计第10名同学的学习时长；

(3)如果用该班学生周末的学习时间作为样本去推断该校高一年级全体学生周末的学习时间，这样推断是否合理？说明理由．

【答案】(1)9人

(2)①8.75；②8.75小时

(3)不合理，理由见解析

【分析】（1）算出对应频率，乘以40即可得解；

（2）由百分位数的定义即可求解；

（3）判断选取的样本是否具有随机性即可判断.

【详解】（1）由图可知，该班学生周末的学习时间不少于20小时的频率为，

则40名学生中周末的学习时间不少于20小时的人数为．

（2）①学习时间在5小时以下的频率为，

学习时间在10小时以下的频率为，

所以25%分位数在区间内，则，

所以这40名同学周末学习时间的25%分位数为8.75．

②第10名是40名同学的25%，因而问题相当于求25%分位数，也就是估计第10名同学的学习时长为8.75小时．

（3）不合理，样本的选取只选在高一某班，不具有代表性．

17．（23-24高一下·福建泉州·期中）果切是一种新型水果售卖方式，商家通过对整果进行消洗、去皮、去核、冷藏等操作后，包装组合销售，在“健康消费”与“瘦身热潮”的驱动下，果切更能满足消费者的即食需求.

(1)统计得到10名中国果切消费者每周购买果切的次数依次为：1，7，4，7，4，6，6，3，7，5，求这10个数据的平均数与方差；

(2)统计600名中国果切消费者的年龄，他们的年龄均在5岁到55岁之间，按照，，，，分组，得到频率分布直方图.

①估计这600名中国果切消费者中年龄不小于35岁的人数；

②估计这600名中国果切消费者年龄的中位数及平均数（结果保留整数）.

【答案】(1)5；3.6

(2)①120；②24；25.

【分析】（1）根据平均数和方差的计算方法求得正确答案.

（2）①根据频率分布直方图求得正确答案；②根据中位数和平均数的求法计算求解.

【详解】（1），

.

（2）①600名中国果切消费者中年龄不小于35岁的人数为：

.

由，，可得，

所以，解得，

所以这600名中国果切消费者年龄的中位数为24.

②其平均数

.

18．（2024高一下·江苏·专题练习）“一带一路”是“丝绸之路经济带”和“21世纪海上丝绸之路”的简称.某市为了了解人们对“一带一路”的认知程度，对不同年龄和不同职业的人举办了一次“一带一路”知识竞赛，满分100分（90分及以上为认知程度高）.现从参赛者中抽取了*x*人，按年龄分成5组，第一组：，第二组：，第三组：，第四组：，第五组：，得到如图所示的频率直方图，已知第一组有6人.



(1)求*x*；

(2)从该市大学生、军人、医务人员、工人、个体户五种人中用分层抽样的方法依次抽取6人，42人，36人，24人，12人，分别记为1～5组，从这5个按年龄分的组和5个按职业分的组中每组各选派1人参加知识竞赛，分别代表相应组的成绩，年龄组中1～5组的成绩分别为93，96，97，94，90，职业组中1～5组的成绩分别为93，98，94，95，90.

①分别求5个年龄组和5个职业组成绩的平均数和方差；

②以上述数据为依据，评价5个年龄组和5个职业组对“一带一路”的认知程度.

【答案】(1)120

(2)①5个年龄组的平均数为94，方差为6，5个职业组的平均数为94，方差为6.8，②答案见解析

【分析】（1）根据第一组的频率和人数列出方程，求出；

（2）①利用平均数和方差公式计算即可；

②由①中求出的两组平均数和方差作出评价.

【详解】（1）根据频率直方图得第一组频率为，所以，所以.

（2）①5个年龄组的平均数为，

方差为，

5个职业组的平均数为，

方差为，

②评价：从平均数来看两组的认知程度相同，从方差来看年龄组的认知程度更好.

**【双基达标】**

**一、单选题**

19．（2024·江西·二模）从甲队60人、乙队40人中，按照分层抽样的方法从两队共抽取10人，进行一轮答题．相关统计情况如下：甲队答对题目的平均数为1，方差为1；乙队答对题目的平均数为1.5，方差为0.4，则这10人答对题目的方差为（ ）

A．0.8 B．0.675 C．0.74 D．0.82

【答案】D

【分析】根据分层抽样的均值与方差公式计算即可.

【详解】根据题意，按照分层抽样的方法从甲队中抽取人，

从乙队中抽取人，

这人答对题目的平均数为，

所以这人答对题目的方差为.

故选：D.

20．（2024高一下·全国·专题练习）已知一组数据从小到大排列为0，4，5，6，8，10，12，15，则这组数据的分位数、分位数分别是（ ）

A．5.5,10 B．5.5,12 C．6,11 D．6,10

【答案】C

【分析】根据给定的数据组，利用分位数、分位数的定义求解即得.

【详解】依题意，，则分位数为从小到大的第4个数据，因此分位数为6；

又，则分位数为第6个数据与第7个数据的平均数，即为.

故选：C

21．（2024高三·全国·专题练习）为了分析某次数学考试的情况，随机抽取了若干学生，将其考试成绩（单位：分）分组为[60，70），[70，80），[80，90），[90，100），[100，110），[110，120），[120，130），[130，140），[140，150]，并绘制成如图所示的频率分布直方图，据此可估计该次考试成绩的中位数为*m*，则*m*的值为（ ）

A．99.8 B．100.8 C．101.8 D．102.8

【答案】B

【分析】根据题意，先根据频率分布直方图求出考试成绩在内的频率，再结合中位数的定义与计算方法，即可求解.

【详解】考试成绩在内的频率为

，

该次考试成绩的中位数为，

则，解得.

故选：B.

22．（23-24高一下·甘肃·期中）某高校为宣扬中华文化，举办了“论语吟唱”的比赛，在比赛中，由A，B两个评委小组（各9人）给参赛选手打分．根据两个评委小组对同一名选手的打分绘制成如图所示折线图，则下列说法正确的是（ ）

A．B组打分的极差小于A组打分的极差 B．B组打分的中位数为75

C．A组的意见相对一致 D．A组打分的众数为50

【答案】C

【分析】对于A：根据折线图结合极差的定义分析判断；对于B：将数据按升序排列，结合中位数分析判断；对于C：根据方差的性质分析判断；对于D：根据题中数据结合众数的定义分析判断.

【详解】对于A：观察折线图可知，小组B的极差大于小组A的极差，故选项A错误；

对于B：小组B打分的分值按照从小到大排列为：36，55，58，62，66，68，68，70，75，

所以中间数为66，故中位数为66，故选项B错误；

对于C：小组A的打分成绩比较均匀，波动更小，故A小组意见相对一致，故选项C正确；

对于D：小组A打分的分值为：42，47，45，46，50，47，55，50，47，

所以小组A打分的分值的众数为47，故选项D错误．

故选：C．

23．（23-24高三下·四川泸州·阶段练习）已知某人收集一个样本容量为50的一组数据，并求得其平均数为70，方差为75，现发现在收集这些数据时，其中两个数据记录有误，一个错将80记录为60，另一个错将70记录为90，在对错误数据进行更正后，重新求得样本的平均数为，方差为，则（ ）

A． B．

C． D．

【答案】C

【分析】根据平均数与方差的定义判断．

【详解】因为，因此平均数不变，即，

设其他48个数据依次为，

因此，

，

，所以，

故选：C．

24．（23-24高一下·甘肃金昌·阶段练习）某校高一年级15个班参加朗诵比赛的得分如下（单位：分）：

则这组数据的分位数为（ ）

A．91 B．90 C．89.5 D．89

【答案】D

【分析】根据给定条件，利用30百分位数的定义直接求解.

【详解】依题意，，所以这组数据的分位数为从小到大排列的第5个数据89.

故选：D

25．（2024高一下·全国·专题练习）从某企业生产的某种产品中抽取100件，测量这些产品的一项质量指标值，由测量结果得到如下频数分布表：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 质量指标值分组 |  |  |   |   |  |
| 频数 | 6 | 26 | 38 | 22 | 8 |

(1)作出这些数据的频率分布直方图；

(2)估计这种产品质量指标值的平均数及方差（同一组中的数据用该组区间的中点值作代表）；

(3)根据以上抽样调查数据，能否认为该企业生产的这种产品符合“质量指标值不低于95的产品至少要占全部产品的”的规定？

【答案】(1)答案见解析

(2)100；104

(3)不能认为该企业生产的这种产品符合“质量指标值不低于95的产品至少要占全部产品的”的规定．



【分析】（1）根据题表中的数据直接绘制即可；

（2）由平均数公式、方差公式直接计算即可；

（3）计算出对应的频率，比较即可判断求解.

【详解】（1）频率分布直方图如图所示．

（2）质量指标值的样本平均数为

.

质量指标值的样本方差为

.

所以这种产品质量指标值的平均数的估计值为100，方差的估计值为104.

（3）质量指标值不低于95的产品所占比例的估计值为.

由于该估计值小于0.8，

故不能认为该企业生产的这种产品符合“质量指标值不低于95的产品至少要占全部产品的”的规定．

26．（23-24高二下·四川绵阳·阶段练习）文明城市是反映城市整体文明水平的综合性荣誉称号，作为普通市民，既是文明城市的最大受益者，更是文明城市的主要创造者，某市为提高市民对文明城市创建的认识，举办了“创建文明城市”知识竞赛，从所有答卷中随机抽取100份作为样本，将样本的成绩（满分100分，成绩均为不低于40分的整数）分成六段：，，得到如图所示的频率分布直方图.



(1)求频率分布直方图中的值；

(2)求样本成绩的第75百分位数；

(3)现从该样本成绩在与两个分数段内的市民中按分层抽样选取6人，求从这6人中随机选取2人，且2人的竞赛成绩之差的绝对值大于20的概率.

【答案】(1)

(2)84

(3)



【分析】（1）利用频率分布直方图每组小矩形的面积之和为1列式求解即可；

（2）利用频率分布直方图结合第75百分位数的计算公式求解即可；

（3）根据分层抽样的概念和古典概型的概率计算公式求解即可.

【详解】（1）因为频率分布直方图每组小矩形的面积之和为1，

所以，

解得.

（2）成绩落在内的频率为,

落在内的频率为,

设第75百分位数为，由，解得，

故第75百分位数为84.

（3）由图可知成绩在与的频率比为，

根据分层抽样在内选取2人,记为,

在内选取4人,记为，

从这6人中选取2人的所有选取方法:

,共15种，

其中2人成绩之差的绝对值大于20的选取方法:

共8种.

所以从这6人中随机选取2人，且2人的竞赛成绩之差的绝对值大于20的概率为.

**【高分突破】**

**一、单选题**

27．（2024·江苏南通·二模）一组样本数据删除一个数后，得到一组新数据：10，21，25，35，36，40.若这两组数据的中位数相等，则删除的数为（ ）

A．25 B．30 C．35 D．40

【答案】B

【分析】根据给定条件，利用中位数的定义求解即得.

【详解】依题意，新数据组有6个数，其中位数是，

显然原数据组有7个数，因此删除的数是中位数30.

故选：B

28．（2024·内蒙古呼和浩特·一模）已知样本数据的平均数为､方差为，若样本数据，的平均数为，方差为，则（ ）

A． B． C． D．

【答案】D

【分析】由平均数和方差的运算性质即可求解.

【详解】由方差的性质，得，，…，的方差为，

故，解得．由，可知.

由平均数的性质，得，，…，的平均数为，

故，

解得.

故选：D．

29．（2024·福建·模拟预测）某单位共有*A*、*B*两部门，1月份进行服务满意度问卷调查，得到两部门服务满意度得分的频率分布条形图如下.设*A*、*B*两部门的服务满意度得分的第75百分位数分别为，，方差分别为，，则（ ）



A．， B．，

C．， D．，

【答案】C

【分析】利用频率分布条形图可读出，，且*A*部门数据更为集中，即可得出结论.

【详解】根据频率分布条形图可知，，即；

显然*A*部门得分数据较*B*部门更为集中，其方差更小，即；

故选：C

30．（2024·江西赣州·一模）若一组样本数据的方差为，则样本数据的方差为（ ）

A．1 B．2 C．2.5 D．

【答案】C

【分析】

根据题意，结合方差的定义以及性质代入计算，即可得到结果.

【详解】设样本数据的平均数为，则，

设样本数据的平均数为，由，

则，所以

.

故选：C

**二、多选题**

31．（2024高一下·全国·专题练习）下列说法中正确的是（ ）

A．数据2，4，6，8的中位数是4，6

B．数据1，2，2，3，4，4的众数是2，4

C．一组数据的平均数、众数、中位数有可能是同一个数

D．8个数据的平均数为5，另3个数据的平均数为7，则这11个数据的平均数是

【答案】BCD

【分析】由数据的数字特征逐一判断各个选项即可得解.

【详解】数据2，4，6，8的中位数为，显然A错误；

由众数的定义可知，数据1，2，2，3，4，4的众数是2，4，

一组数据的平均数、众数、中位数有可能是同一个数，例如这组数据的每个数都相同的时候就满足，

8个数据的平均数为5，另3个数据的平均数为7，由加权平均数可知，这11个数据的平均数是

故B，C，D都正确．

故选：BCD.

32．（2024高一下·全国·专题练习）已知100个数据的75%分位数是9.3，则下列说法正确的是（ ）

A．这100个数据中至少有75个数小于或等于9.3

B．把这100个数据从小到大排列后，9.3是第75个数据

C．把这100个数据从小到大排列后，9.3是第75个数据和第76个数据的平均数

D．把这100个数据从小到大排列后，9.3是第74个数据和第75个数据的平均数

【答案】AC

【分析】根据百分位数的定义进行判断选择.

【详解】通过定义知A正确；

因为100×75%＝75为整数，所以第75个数据和第76个数据的平均数为75%分位数，是，则C正确．BD错误.

故选：AC

33．（2024·全国·模拟预测）在某市高三年级等行的一次数学期末考试中，为了解考生的成绩情况，随机抽取了50名考生的成绩，作出的频率分布直方图如图，成绩排在前的学生将获得“优秀学生”称号，则（ ）



A．估计该市考生的成绩低于60分的比例为

B．估计该市考生成绩的众数为60

C．估计该市考生成绩的平均数为70.6

D．估计该市82分以上的考生将获得“优秀学生”称号

【答案】AC

【分析】先利用直方图的性质，直接判断选项A， B， C；通过估算面积的方法，以样本频率估计总体概率来判断选项D.

【详解】对于A，由直方图的性质知：，

解得，所以样本成绩低于分的频率为；

所以以样本的频率估计总体概率，故A正确；

对于B，总体众数的估计值是直方图面积最大的直方块横坐标对应范围的中点值，

即众数为，故B错误；

对于C，总体平均数的估计值为各直方块中点值与各块面积的乘积的和，

，故C正确；

对于D，样本中分以上的占比，

估计总体荣获“优秀学生”的分数为：，故D错误.

故选：AC.

34．（23-24高一下·全国·课后作业）【多选】一组数据的平均值为5，方差为2，极差为7，中位数为6，记，，，，的平均值为，方差为，极差为，中位数为，则（ ）

A． B． C． D．

【答案】ACD

【分析】根据平均数、方差、极差、中位数定义及性质求解即可．

【详解】依题意，，，，.

故选：ACD

35．（2024·广东汕头·一模）某次数学考试后，为分析学生的学习情况，某校从某年级中随机抽取了名学生的成绩，整理得到如图所示的频率分布直方图.为进一步分析高分学生的成绩分布情况，计算得到这名学生中，成绩位于内的学生成绩方差为，成绩位于内的同学成绩方差为.则（ ）

参考公式：样本划分为层，各层的容量､平均数和方差分别为：、、；、、.记样本平均数为，样本方差为，.



A．

B．估计该年级学生成绩的中位数约为

C．估计该年级成绩在分及以上的学生成绩的平均数为

D．估计该年级成绩在分及以上的学生成绩的方差为

【答案】BCD

【分析】利用频率分布直方图中，所有直方图的面积之和为，列等式求出实数的值，可判断A选项；利用中位数的定义可判断B选项；利用总体平均数公式可判断C选项；利用方差公式可判断D选项.

【详解】对于A选项，在频率分布直方图中，所有直方图的面积之和为，

则，解得，A错；

对于B选项，前两个矩形的面积之和为，

前三个矩形的面积之和为，

设计该年级学生成绩的中位数为，则，

根据中位数的定义可得，解得，

所以，估计该年级学生成绩的中位数约为，B对；

对于C选项，估计成绩在分以上的同学的成绩的平均数为

分，C对；

对于D选项，估计该年级成绩在分及以上的学生成绩的方差为

，D对.

故选：BCD.

**三、填空题**

36．（2024高一下·全国·专题练习）从高三抽出50名学生参加数学竞赛，由成绩得到如图的频率分布直方图，则估计这50名学生成绩的分位数为 分．



【答案】

【分析】利用给定的频率分布直方图，借助频率估计即可.

【详解】依题意，前四个小矩形的面积之和为，

前五个小矩形的面积之和为，

因此分位数位于内，，

所以估计这50名学生成绩的分位数为分．

故答案为：

37．（2024高一下·全国·专题练习）某校从高一年级中随机抽取部分学生，将他们的期末数学测试成绩分成6组：加以统计，得到如图所示的频率分布直方图．据此统计，期末数学测试成绩不少于分位数的分数至少为 分．



【答案】74

【分析】直接由百分位数的定义进行估计即可.

【详解】因为，，

所以分位数应位于内．

由，得期末数学测试成绩不少于分位数的分数至少为74分．

故答案为：74.

38．（2024·全国·模拟预测）某小组5位同学各拋掷一枚正方体骰子，将正面向上的点数按从小到大的顺序记录下来，得到一组统计数据．已知这组数据的平均数为整数，最大值为6，中位数为3，方差为1.6，则这组数据的众数为 ．

【答案】3

【分析】根据题意将这组数据由小到大设出来，再根据方差为1.6可判断出平均数的取值有两种情况 ，对这两种情况分别讨论即可得出结果.

【详解】设这组数据为，则平均数大于等于，

假设平均数为3．因为，

所以平均数或，

若，则，

所以，解得；

若，则，

所以，此时无解．

所以这组数据的众数为3．

故答案为：3.

39．（23-24高三上·上海宝山·期末）在某次比赛中运动员五轮的成绩互不相等，记为，平均数为，若随机删去其中一轮的成绩，得到一组新数据，记为，平均数为，下面说法正确的是 .（写出所有正确选项）

①新数据的极差可能等于原数据的极差.

②新数据的中位数可能等于原数据的中位数.

③若，则新数据的方差一定大于原数据方差.

④若，则新数据的第40百分位数一定大于原数据的第40百分位数.

【答案】①②③

【分析】根据极差、中位数、平均数和方差的概念，以及百分位数的概念及计算方法，逐项判定，即可求解.

【详解】对于①，若随机删去任一轮的成绩，恰好不是最高成绩和最低成绩，

此时新数据的极差可能等于原数据的极差，所以①正确；

对于②，不妨假设，

当时，若随机删去的成绩是，

此时新数据的中位数等于原数据的中位数，所以②正确；

对于③，若，即删去的数据恰为平均数，

根据方差的计算公式，分子不变，分母变小，所以方差会变大，所以③正确；

对于④，若，即删去的数据恰为平均数，在按从小到大的顺序排列的5个数据中，

因为，此时原数据的分位数为第二数和第三个数的平均数；

删去一个数据后的4个数据，从小到大的顺序排列，可得，

此时新数据的分位数为第二个数，

显然新数据的分位数小于原数据的分位数，所以④错误.

故答案为：①②③.

**四、解答题**

40．（2024高二下·湖南娄底·学业考试）2023年10月22日，汉江生态城2023襄阳马拉松在湖北省襄阳市成功举行，志愿者的服务工作是马拉松成功举办的重要保障，襄阳市新时代文明实践中心承办了志愿者选拔的面试工作．现随机抽取了100名候选者的面试成绩，并分成五组：第一组，第二组，第三组，第四组，第五组，绘制成如图所示的频率分布直方图．已知第一、二组的频率之和为0.3，第一组和第五组的频率相同．



(1)估计这100名候选者面试成绩的平均数和第25百分位数；

(2)现从以上各组中用分层随机抽样的方法选取20人，担任本市的宣传者．若本市宣传者中第二组面试者的面试成绩的平均数和方差分别为62和40，第四组面试者的面试成绩的平均数和方差分别为80和70，据此估计这次第二组和第四组所有面试者的面试成绩的方差．

【答案】(1)，63

(2)



【分析】（1）由题意先求出，进一步结合平均数公式、百分位数的定义即可列式求解；

（2）首先算出抽样比，再根据加权平均公式以及方差的性质即可列式求解.

【详解】（1）由题意可知：，解得，

可知每组的频率依次为：0.05，0.25，0.45，0.2，0.05，

所以平均数为，

因为，

设第25百分位数为，则，则，

解得，故第25百分位数为63．

（2）设第二组、第四组面试者的面试成绩的平均数与方差分别为，

且两组频率之比为，

则第二组和第四组所有面试者的面试成绩的平均数，

第二组和第四组所有面试者的面试成绩的方差

．

故估计第二组和第四组所有面试者的面试成绩的方差是．

41．（2024高一下·全国·专题练习）文明城市是反映城市整体文明水平的综合性荣誉称号，作为普通市民，既是文明城市的最大受益者，更是文明城市的主要创造者.某市为提高市民对文明城市创建的认识，举办了“创建文明城市”知识竞赛，从所有答卷中随机抽取100份作为样本，将样本的成绩（满分100分，成绩均为不低于40分的整数）分成六段：，，，得到如图所示的频率分布直方图.



(1)求频率分布直方图中的值；

(2)求样本成绩的第75百分位数；

(3)已知落在的平均成绩是56，方差是7，落在的平均成绩为65，方差是4，求两组成绩的总平均数和总方差.

【答案】(1)；

(2)84；

(3)平均值62，方差23.

【分析】（1）利用每组小矩形的面积之和为1即可求得*a*的值.

（2）利用频率分布直方图结合第75百分位数的求法即可求得答案.

（3）根据平均数的计算公式可求得两组成绩的总平均数；根据由样本方差计算总体方差公式可求得两组成绩的总方差.

【详解】（1）由每组小矩形的面积之和为1得，

所以.

（2）成绩落在内的频率为，

落在内的频率为，

显然第75百分位数，由，解得，

所以第75百分位数为84.

（3）由频率分布直方图知，成绩在的市民人数为，

成绩在的市民人数为，所以；

由样本方差计算总体方差公式，得总方差为.

