用样本估计总体知识点总结-高中数学必修3第二章

(一)用样本的频率分布估计总体分布

1.样本的频率分布直方图

　作频率分布直方图的步骤:



因为频率分布直方图中横轴表示“组距”,纵轴表示“频率/组距”,所以“小长方形的面积=组距×频率/组距=频率”,各小长方形的面积之和等于1.

2.频率分布折线图:连接频率分布直方图中各小长方形上端的中点,就得频率分布折线图.

3.总体密度曲线:随着样本容量的增加,作图时所分的组数增加,组距减小,相应的频率折线图会越来越接近于一条光滑曲线,即总体密度曲线.

4.茎叶图

(1)茎是指中间的一列数,叶是从茎的旁边生长出来的数.

(2)茎叶图表示数据有两个突出的优点:

其一是统计图上没有原始数据的损失,所有信息都可以从这个茎叶图中得到,其二是在比赛时随时记录,方便记录与表示.

(3)在样本数据较少时,用茎叶图表示数据的效果较好,但当样本数据较多时,茎叶图就显得不太方便了.

(二)用样本的数字特征估计总体的数字特征

1.平均数:一组数据的总和除以数据的个数所得的商就是平均数;它是频率分布直方图的“重心”.

2.中位数:如果将一组数据按从小到大或从大到小的顺序依次排列,当数据有奇数个时,处在中间的一个数就是这组数据

的中位数;当数据有偶数个时,处在中间的两个数的平均数就是这组数据的中位数.因此中位数不一定是样本数据中的数.在频率分布直方图中中位数左右两侧直方图的面积应该相等,由此在图中我们可以估计其近似值.

3.众数:在样本数据中出现次数最多的数叫做这组数据的众数.若一组样本数据中有两个或几个数据出现的最多且次数一样,这些数据都是这组数据的众数;若一组数据中每个数据出现的一样多,则认为这组数据没有众数.

4.方差:已知一组数据为则这组数据的方差,其中，用它来衡量数据围绕平均数的波动情况.

5.标准差:是方差的算术平方根.它比方差多开一次方,它的度量单位与样本数据的相同,有时用它更加方便.

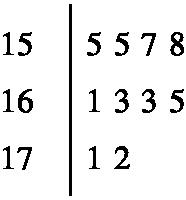
**理解总结**

样本频率分布直方图是考查热点,需注意频率分布直方图中横轴表示“组距”,纵轴表示“频率/组距”,所以“小长方形的面积=组距×(频率/组距)=频率”,各小长方形的面积之和等于1.

对于平均数、众数、中位数、方差等要会求,并能理解其意义.

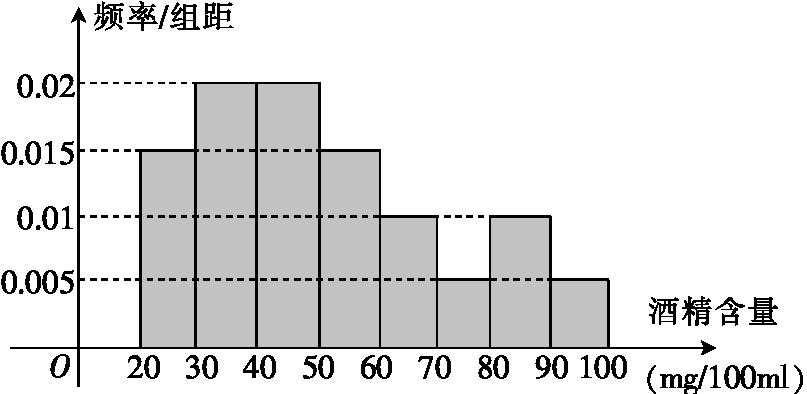
**高考导航**

例1 (1)右图是根据某校10位高一同学的身高(单位:cm)画出的茎叶图,其中左边的数字从左到右分别表示学生身高的百位数字和十位数字,右边的数字表示学生身高的个位数字,从图中可以得到这10位同学身高的中位数是 (　 )



(A)161 cm.　   (B)162 cm.　   (C)163 cm.　   (D)164 cm.

(2)根据《中华人民共和国道路交通安全法》规定:车辆驾驶员血液酒精浓度在20-80 mg/100 ml(不含80)之间,属于酒后驾车;血液酒精浓度在80 mg/100ml (含80)以上时,属醉酒驾车.据《法制晚报》报道,2010年3月15日至3月28日,全国查处酒后驾车和醉酒驾车共28800人,如图是对这28800人酒后驾车血液中酒精含量进行检测所得结果的频率分布直方图,则属于醉酒驾车的人数约为   .



**解题思路**

(1)理解茎叶图的结构和中位数的找法是解决本题的关键;(2)对于频率分布直方图,注意纵轴为频率与组距的比值.根据图中的量可求出醉酒驾车的频率,从而可知其人数.

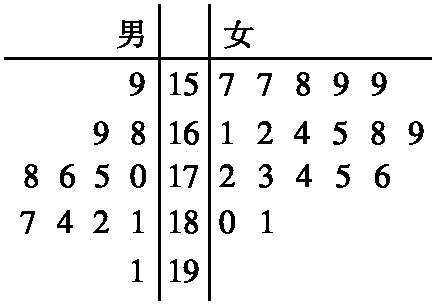
解析:(1)由给定的茎叶图可知,这10位同学身高的中位数为

.

(2)由给定的频率分布直方图可知,驾驶员血液酒精浓度在80 mg/100 ml(含80)以上的频率为(0.01+0.005)×10=0.15.故属于醉酒驾车的人数约为

28800×0.15=4320.

例2　第26届世界大学生夏季运动会于2011年8月12日到23日在深圳举行,运动会期间,为了搞好接待工作,组委会在某学院招募了12名男志愿者和18名女志愿者.将这30名志愿者的身高编成如下图所示的茎叶图(单位:cm):若身高在  
175 cm以上(包括175 cm)定义为“高个子”,身高在175 cm以下(不包括175 cm)定义为“非高个子”,且只有“女高个子”才担任“礼仪小姐”.



(1)如果用分层抽样的方法从“高个子”和“非高个子”中抽取5人,再从这5人中选2人,那么至少有一人是“高个子” 的概率是多少?

(2)若从所有“高个子”中选3名志愿者,用****表示所选志愿者中能担任“礼仪小姐”的人数,试写出的分布列,并求的数学期望.

**解题思路**

(1)首先根据分层抽样的规律找到所抽取的5人中“高个子”和“非高个子”的人数.对于“至少”问题可考虑其对立事件;(2)一共有12个高个子,其中有男志愿者8人,女志愿者4人,选出3名,其中能担任“礼仪小姐”的人数可取0,1,2,3;注意取每个值的意义,例如“”表示从12人中选3人,其中有1个“女高个子”,其余的为男志愿者.

解析**:**(1)根据茎叶图,“高个子”有12人,“非高个子”有18人,用分层抽样的方法,每个人被抽中的概率是,所以选中的“高个子”有

,“非高个子”有人.

用事件表示“至少有一名‘高个子’被选中”,则它的对立事件表示“没有一名‘高个子’被选中”,则.因此,至少有一人是“高个子”的概率是.

(2)依题意,一共有12个高个子,其中有男志愿者8人,女志愿者4人,则的取值为0,1,2,3.则

，，

，.

　　因此,的分布列如下:

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

.