

目 录

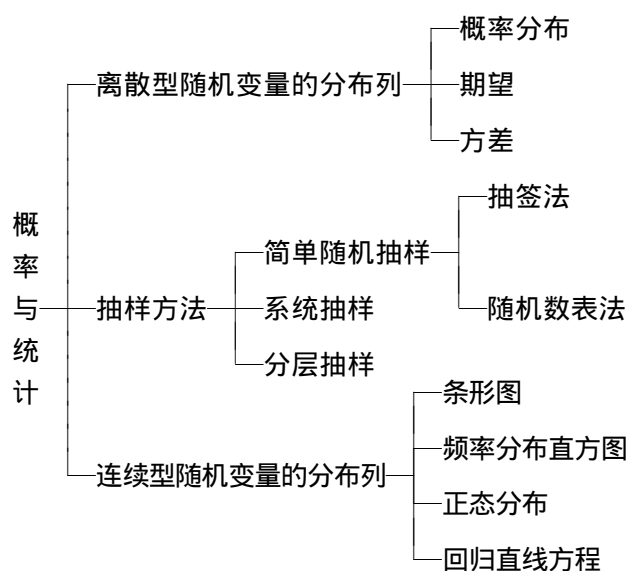
M U L U

第一章 概率与统计	1
1.1 离散型随机变量的分布列	2
1.2 离散型随机变量的期望与方差	7
1.3—1.4 抽样方法、总体分布的估计	12
1.5—1.6 正态分布与线性回归	15
本章知识及规律方法总结	18
第二章 极限	21
2.1 数学归纳法及其应用举例	22
2.2 数列的极限	26
2.3 函数的极限	29
2.4 极限的四则运算	32
2.5 函数的连续性	35
本章知识及规律方法总结	38
第三章 导数	42
3.1 导数的概念	43
3.2 几种常见函数的导数	46
3.3 函数的和、差、积、商的导数	48
3.4 复合函数的导数	51
3.5 对数函数与指数函数的导数	55
3.6 函数的单调性	58
3.7 函数的极值	62
3.8 函数的最大值与最小值	66
本章知识及规律方法总结	70
第四章 数系的扩充——复数	75
4.1 复数的概念	76
4.2 复数的运算	79
4.3 数系的扩充	82
本章知识及规律方法总结	84
参考答案	88
测试部分(另配)	



第一章 概率与统计

本章知识通览



本章学习目标

1. 了解随机变量、离散型随机变量的意义,会求出某些简单的离散型随机变量的分布列.
2. 了解离散型随机变量的期望、方差的意义,会根据离散型随机变量的分布列求出期望与方差.
3. 会用简单的随机抽样、系统抽样、分层抽样等常用的抽样方法从总体中抽取样本.
4. 会用样本频率分布去估计总体分布.
5. 了解正态分布的意义及主要性质.
6. 了解线性回归的方法.
7. 通过以抽样方法为内容的实习作业,培养学生运用统计方法解决问题的能力.
8. 了解假设检验的基本思想.

本章重点

1. 随机变量、离散型随机变量的分布列.
2. 离散型随机变量的期望与方差.
3. 抽样方法 ① 简单随机抽样 ② 系统抽样 ③ 分层抽样.
4. 总体分布的估计.
5. 正态分布的意义和性质.



本章难点

1. 离散型随机变量的期望的概念及意义.
2. 抽样方法中“各个个体被抽取的概率相等”的理解.
3. 正态分布及其意义.

本章学习注意事项

1. 本章内容与初中数学的统计初步、高中数学必修课的排列、组合及概率的内容有较密切的联系. 在学习中要注意复习有关内容, 应从整体的观点处理这部分知识.
2. 事件的概率着眼于随机现象的局部问题, 与此不同, 随机变量的概率分布及其期望、方差等则着眼于随机现象的整体和全局问题. 其中, 离散型随机变量的分布列给出了随机变量取所有可能值的概率, 期望反映了随机变量取值的平均水平, 方差反映了随机变量取值的集中与分散状况, 这些都是从整体和全局上来描述随机变量的.
3. 统计内容的实践性较强, 有一些习题带有“实习作业”的特点, 应充分重视这些习题, 以提高运用所学知识解决简单实际问题的能力和动手能力.
4. 本章里的统计计算较为复杂, 应学会运用科学计算器, 以提高解决问题的效率.

1.1 离散型随机变量的分布列

学法指导

了解随机变量的意义和分类, 判定什么是离散型随机变量; 会求出某些简单的离散型随机变量的分布列.

本节重点是离散型随机变量的分布列, 难点是随机变量、离散型随机变量、连续型随机变量概念的建立.

离散型随机变量这部分知识是随机事件的概率这一知识块的深化, 离散型随机变量的本质就是某些随机试验结果的数量化.

分布列的计算是概率部分计算的延伸, 以往讨论的是某一具体事件概率的计算, 分布列是讨论全部基本事件的概率计算, 正确计算的基础是对基本概念的理解, 因此在学习本部分内容时要重视上一章知识的复习及应用, 同时还要注意明确数学符号的含义.

ξ 代表随机变量, 是定义在随机试验全部结果上的函数. $\xi = k$ 是随机变量取 k 值, 表示某一个基

本事件. 又 $\xi \geq k, \xi > k, \xi \leq k, \xi < k$ 等均表示由一些基本事件组成的一般事件.

$P(\xi = k), P(\xi \geq k)$ 等表示事件的概率, 是取值于 $[0, 1]$ 上的一个实数;

$P(\xi = k) = p_k$ 表示事件 $\xi = k$ 时概率的值是 p_k ;

$\eta = a\xi + b$ 表示 η 是 ξ 的线性函数, ξ 是随某个试验结果而变的随机变量, 相应地, η 也随试验结果而变, 也是随机变量.

求分布列的步骤: 首先明确随机变量 ξ 取哪些值, 其次求 ξ 取每一个值的概率, 然后列成表格. 注意用分布列的两条性质检验所求的分布列的概率是否正确.

基础知识及重点难点

1. 随机变量

两点说明: (1) 课本在介绍随机变量的概念时, 随机试验可以作为不加定义的原始概念引入.

(2) 所谓随机变量, 即是随机试验的试验结果和实数之间的一个对应关系, 这种对应关系是人为建立起来的, 但又是客观存在的.

2. 离散型随机变量

说明: (1) 离散型随机变量 ξ 可能取的值为有



限个或至多可列个,这里的“可列”不易理解,所以课本用比较浅显的语言“按一定次序一一列出”来描述,比如 ξ 取 $1, 2, \dots, n, \dots$.

(2)教材中为了控制难度,所涉及的离散型随机变量可能取的值的个数多数是有限的.

3. 连续型随机变量

对于连续型随机变量,教科书无法给出更为严格的定义,只给出连续型随机变量的一种直观和朴素的描述.

4. 离散型随机变量的分布列

在离散型随机变量的分布列中,由于概率一定非负,而且一次试验的各种结果是彼此互斥的,全部结果之和为一必然事件,所以分布列中的概率值,具有下列两个性质:

$$(1) p_i \geq 0, i = 1, 2, \dots$$

$$(2) p_1 + p_2 + \dots = 1.$$

5. 几种常用的分布列及其求法

(1)单点分布:随机变量 ξ 只取一个值,它的分布列为 $P(\xi = a) = 1$.

(2)两点分布:随机变量 ξ 只取两个值,它的分布列为:

ξ	x_1	x_2
P	p	1 - p

(3)二项分布:①在 n 次独立重复试验中某个事件发生的次数 ξ 是一个随机变量.②如果在一次试验中某事件发生的概率为 p ,那么在 n 次独立重复试验中,这个事件恰好发生 k 次的概率为: $P(\xi = k) = C_n^k p^k q^{n-k}$,其中 $k = 0, 1, \dots, n, q = 1 - p$.它的分布列为:

ξ	0	1	2	...	k	...	n
P	$C_n^0 p^0 q^n$	$C_n^1 p^1 q^{n-1}$	$C_n^2 p^2 q^{n-2}$...	$C_n^k p^k q^{n-k}$...	$C_n^n p^n q^0$

(4)几何分布:在独立重复试验中,某事件第一次发生时所作试验次数 ξ 是一个取值为正整数的离散型随机变量.“ $\xi = k$ ”表示在第 k 次独立重复试验时,事件第一次发生, $P(\xi = k) = q^{k-1} p$,其中 $k = 1, 2, 3, \dots, q = 1 - p$,它的分布列为:

ξ	1	2	3	...	k	...
P	p	qp	q ² p	...	q ^{k-1} p	...

典型例题详解及解题方法透析

【例1】下列变量中是离散型随机变量的有

(1)在2005张已编号的卡片(从1号到2005号)中任取1张,被取出的号数为 ξ ;

(2)在2005张已编号的卡片(从1号到2005号)中任取3张,被取出的卡片的号数之和为 ξ ;

(3)连续不断射击,首次命中目标需要的射击次数 η ;

(4)某工厂加工的某种钢管,外径与规定的外径尺寸之差 η ;

(5)投掷一颗骰子,六面都刻上数字,所得的点数 η .

分析:由离散型随机变量的定义来看,不难知(1),(2),(3)正确,而(4)是连续随机变量,(5)中六面上的数字可为同一个数,故(4),(5)不是离散型随机变量.

解:(1),(2),(3).

点评:判断一个变量是否是随机的,主要是具体分析变量的某些值的出现是否确定,结果不能确定的是随机的,否则不是.判断一个随机变量是否是离散型随机变量,主要看此变量的取值是否是有限个,或虽是无限个,但可以按一定顺序列举出来的.

【例2】一个袋中装有6个同样大小的黑球,编号为1,2,3,4,5,6,现从中随机取出3个球,以 ξ 表示取出球的最大号码,求 ξ 的分布列.

分析:确定 ξ 的所有可能的取值,对于 ξ 的每一个取值,求出相应的概率即可.

解:随机变量 ξ 的取值为3,4,5,6.从袋中随机地取出3个球,包含的基本事件总数为 C_6^3 ,事件“ $\xi = 3$ ”包含的基本事件总数为 C_2^2 ,事件“ $\xi = 4$ ”包含的基本事件总数为 C_3^2 ,事件“ $\xi = 5$ ”包含的基本事件总数为 C_4^2 ,事件“ $\xi = 6$ ”包含的基本事件总数为 C_5^2 ,从而有

$$P(\xi = 3) = \frac{C_2^2}{C_6^3} = \frac{1}{20}, P(\xi = 4) = \frac{C_3^2}{C_6^3} = \frac{3}{20}, P(\xi = 5) = \frac{C_4^2}{C_6^3} = \frac{3}{10}, P(\xi = 6) = \frac{C_5^2}{C_6^3} = \frac{1}{2}.$$

随机变量 ξ 的分布列为

ξ	3	4	5	6
P	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{3}{10}$	$\frac{1}{2}$

点评:确定离散型随机变量 ξ 的分布列的关键是搞清 ξ 取每一个值对应的随机事件,进一步利用等可能事件的概率公式和排列、组合知识求出 ξ 取



每一个值的概率,从而获得 ξ 的分布列.

〔例3〕设随机变量 ξ 的分布列 $P\left(\xi = \frac{k}{5}\right) = ak$ ($k=1, 2, 3, 4, 5$).

(1) 求常数 a 的值;

(2) 求 $P\left(\xi \geq \frac{3}{5}\right)$;

(3) 求 $P\left(\frac{1}{10} < \xi < \frac{7}{10}\right)$.

分析: 利用分布列的性质(2)求解.

解: 题目所给分布列为:

ξ	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{5}{5}$
P	a	$2a$	$3a$	$4a$	$5a$

(1) 由 $a + 2a + 3a + 4a + 5a = 1$, 得 $a = \frac{1}{15}$.

(2) $P\left(\xi \geq \frac{3}{5}\right) = P\left(\xi = \frac{3}{5}\right) + P\left(\xi = \frac{4}{5}\right) + P(\xi = 1) = \frac{3}{15} + \frac{4}{15} + \frac{5}{15} = \frac{4}{5}$, 或 $P\left(\xi \geq \frac{3}{5}\right) = 1 - P\left(\xi \leq \frac{2}{5}\right) = 1 - \left(\frac{1}{15} + \frac{2}{15}\right) = \frac{4}{5}$.

(3) 因为 $\frac{1}{10} < \xi < \frac{7}{10}$ 只有 $\xi = \frac{1}{5}, \frac{2}{5}, \frac{3}{5}$ 满足,

故 $P\left(\frac{1}{10} < \xi < \frac{7}{10}\right) = P\left(\xi = \frac{1}{5}\right) + P\left(\xi = \frac{2}{5}\right) + P\left(\xi = \frac{3}{5}\right) = \frac{1}{15} + \frac{2}{15} + \frac{3}{15} = \frac{2}{5}$.

点评: 分布列有两条重要的性质: ① $p_i \geq 0, i = 1, 2, \dots$; ② $p_1 + p_2 + \dots = 1$.

随机变量并不一定要取整数值, 它的取值一般取决于实际问题, 并有其特定的含义, 因此, 可以是 R 中的任意值.

〔例4〕设某射手每次射击击中目标的概率是 0.8, 现在连续射击 4 次, 求击中目标的次数 ξ 的概率分布.

分析: 这是一个独立重复射击问题, 其击中目标的次数 ξ 的概率分布属二项分布, 可直接由二项分布列得出.

解: 在独立重复射击中, 击中目标的次数 ξ 的概率分布属二项分布 $B(n, p)$.

$n=4, p=0.8, P(\xi = k) = C_4^k (0.8)^k \cdot (0.2)^{4-k} (k=0, 1, 2, 3, 4)$.

, $P(\xi = 0) = C_4^0 (0.8)^0 (0.2)^4 = 0.0016$.

, $P(\xi = 1) = C_4^1 (0.8)^1 (0.2)^3 = 0.0256$.

, $P(\xi = 2) = C_4^2 (0.8)^2 (0.2)^2 = 0.1536$.

, $P(\xi = 3) = C_4^3 (0.8)^3 (0.2)^1 = 0.4096$.

, $P(\xi = 4) = C_4^4 (0.8)^4 (0.2)^0 = 0.4096$.

, ξ 的概率分布为:

ξ	0	1	2	3	4
P	0.0016	0.0256	0.1536	0.4096	0.4096

点评: 确定分布列的类型非常重要, 其中二项分布对应独立重复试验, 这一点是我们判断一个分布列是否为二项分布的标准.

〔例5〕在一个袋中装有 1 只红球和 9 只白球, 每次从袋中任取一球, 取后放回, 直到取到红球为止, 求取球次数 ξ 的分布列.

分析: 袋中虽然只有 10 个球, 由于每次取出一球, 取后又放回, 因此应注意以下几点:

(1) 一次取球有两个结果: 取红球 A 或取白球 \bar{A} , 且 $P(A) = 0.1$;

(2) 取球次数 ξ 可能取 $1, 2, \dots$

(3) 由于“取后放回”, 因此, 各次取球相互独立.

解: ξ 的所有可能取值为 $1, 2, \dots, n, \dots$. 令 A_k 表示第 k 次取到红球, 则由于各次取球相互独立, 且取到红球的概率为 $P = 0.1$, 于是得 $P(\xi = 1) = P(A_1) = 0.1$, $P(\xi = 2) = P(\bar{A}_1 \cdot A_2) = P(\bar{A}_1) \cdot P(A_2) = 0.9 \times 0.1, \dots, P(\xi = k) = P(\bar{A}_1 \cdot \bar{A}_2 \cdot \dots \cdot \bar{A}_{k-1} \cdot A_k) = P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \dots \cdot P(\bar{A}_{k-1}) \cdot P(A_k) = (1-p) \cdot (1-p) \cdot \dots \cdot (1-p)p = 0.9 \times 0.9 \times \dots \times 0.9 \times 0.1 = 0.9^{k-1} \times 0.1$.

因此所求的分布列为:

ξ	1	2	3	...	k	...
P	0.1	0.9×0.1	$0.9^2 \times 0.1$...	$0.9^{k-1} \times 0.1$...

点评: 此例进一步抽象可表述为: 在每次试验时, 若事件 A 发生的概率为 p , \bar{A} 发生的概率为 $q = 1 - p$, 则事件 A 首次发生的试验次数 ξ 是一个随机变量, 它的取值为 $1, 2, \dots, n, \dots$, 其分布列为的几何分布:

ξ	1	2	3	4	...	k	...
P	p	qp	q^2p	q^3p	...	$q^{k-1}p$...

知识应用及探究创新

〔例〕A, B 两个代表队进行乒乓球对抗赛, 每队三名队员, A 队队员是 A_1, A_2, A_3 , B 队队员是 $B_1,$



B_2, B_3 . 按以往多次比赛的统计, 对阵队员之间胜负概率如下:

对阵队员	A 队队员胜的概率	B 队队员胜的概率
A_1 对 B_1	$\frac{2}{3}$	$\frac{1}{3}$
A_2 对 B_2	$\frac{2}{5}$	$\frac{3}{5}$
A_3 对 B_3	$\frac{2}{5}$	$\frac{3}{5}$

现按表中对阵方式出场, 每场胜队得 1 分, 负队得 0 分. 设 A 队, B 队最后所得总分分别为 ξ, η . 求 ξ, η 的概率分布.

分析: 首先应分析出其中一个随机变量可能的取值, 再分析两个随机变量的关系, 然后利用两个相关的随机变量的有关性质或公式进行计算.

解: ξ 的可能取值为 3, 2, 1, 0, 由于 $\xi + \eta = 3$, 故 η 的相应可能取值为 0, 1, 2, 3.

$$P(\xi=3) = \frac{2}{3} \times \frac{2}{5} \times \frac{2}{5} = \frac{8}{75},$$

$$P(\xi=2) = \frac{2}{3} \times \frac{2}{5} \times \frac{3}{5} + \frac{1}{3} \times \frac{2}{5} \times \frac{2}{5} + \frac{2}{3} \times \frac{3}{5} \times \frac{2}{5} = \frac{28}{75}.$$

$$P(\xi=1) = \frac{2}{3} \times \frac{3}{5} \times \frac{3}{5} + \frac{1}{3} \times \frac{2}{5} \times \frac{3}{5} + \frac{1}{3} \times \frac{3}{5} \times \frac{2}{5} = \frac{2}{5}.$$

$$P(\xi=0) = \frac{1}{3} \times \frac{3}{5} \times \frac{3}{5} = \frac{3}{25}.$$

又因为 $\xi + \eta = 3$, 所以

$$P(\eta=0) = P(\xi=3) = \frac{8}{75}.$$

$$P(\eta=1) = P(\xi=2) = \frac{28}{75}.$$

$$P(\eta=2) = P(\xi=1) = \frac{2}{5}.$$

$$P(\eta=3) = P(\xi=0) = \frac{3}{25}.$$

综上, ξ, η 的概率分布列分别是

ξ	0	1	2	3
P	$\frac{3}{25}$	$\frac{2}{5}$	$\frac{28}{75}$	$\frac{8}{75}$

η	0	1	2	3
P	$\frac{8}{75}$	$\frac{28}{75}$	$\frac{2}{5}$	$\frac{3}{25}$

点评: 本题若不能注意到 $\xi + \eta = 3$, 也可以像求 ξ 的分布列那样去求 η 的分布列, 这种方法显然会增大计算量, 影响解答的准确性.

典型考题详解及考点透析

【例 1】(山东省 2005 年高考题) 袋中装有黑球和白球共 7 个, 从中任取两个球都是白球的概率为 $\frac{1}{7}$. 现有甲、乙两人从袋中轮流摸取 1 球, 甲先取, 乙后取, 然后甲再取……取后不放入, 直到两人中有一人取到白球时即终止. 每个球每一次被取出的机会是等可能的.

(1) 求袋中原有白球的个数;

(2) 用 ξ 表示取球终止时所需的取球次数, 求随机变量 ξ 的概率分布;

(3) 求甲取到白球的概率.

考点透析: 考查分布列及利用分布列解决实际问题的能力.

解: (1) 设袋中原有 n 个白球, 由题意知

$$\frac{1}{7} = \frac{C_n^2}{C_7^2} = \frac{\frac{n(n-1)}{2}}{\frac{7 \times 6}{2}} = \frac{n(n-1)}{7 \times 6}.$$

所以 $n(n-1) = 6$, 解得 $n=3$ (舍去 $n=-2$), 即袋中原有 3 个白球.

(2) 由题意 ξ 的可能取值为 1, 2, 3, 4, 5.

$$P(\xi=1) = \frac{3}{7}, P(\xi=2) = \frac{4 \times 3}{7 \times 6} = \frac{2}{7}, P(\xi=3) = \frac{4 \times 3 \times 3}{7 \times 6 \times 5} = \frac{6}{35}, P(\xi=4) = \frac{4 \times 3 \times 2 \times 3}{7 \times 6 \times 5 \times 4} = \frac{3}{35}, P(\xi=5) = \frac{4 \times 3 \times 2 \times 1 \times 3}{7 \times 6 \times 5 \times 4 \times 3} = \frac{1}{35}.$$

所以, 取球次数 ξ 的分布列为

ξ	1	2	3	4	5
P	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{6}{35}$	$\frac{3}{35}$	$\frac{1}{35}$

(3) 因为甲先取, 所以甲只可能在第 1 次、第 3 次和第 5 次取球, 记“甲取到白球”的事件为 A, 则 $P(A) = P(\xi=1)$ 或“ $\xi=3$ ”或“ $\xi=5$ ”.

因为事件“ $\xi=1$ ”、“ $\xi=3$ ”、“ $\xi=5$ ”两两互斥, 所以

$$P(A) = P(\xi=1) + P(\xi=3) + P(\xi=5) = \frac{3}{7} + \frac{6}{35} + \frac{1}{35} = \frac{22}{35}.$$



点评:准确找出随机变量 ξ 的取值,是解决此类问题的关键.

〔例2〕(江苏省2005年高考题)甲、乙两人各射击1次,击中目标的概率分别是 $\frac{2}{3}$ 和 $\frac{3}{4}$. 假设两人射击是否击中目标,相互之间没有影响;每次射击是否击中目标,相互之间没有影响.

(1) 求甲射击4次,至少有1次未击中目标的概率;

(2) 求两人各射击4次,甲恰好击中目标两次且乙恰好击中目标3次的概率;

(3) 假设某人连续两次未击中目标,则停止射击.问:乙恰好射击5次后,被中止射击的概率是多少?

考点透视:(1)利用逆向思维,考虑其对立事件的概率.

(2)分别求出甲、乙射击满足条件时的概率,再求出甲、乙两人同时满足条件时击中目标的概率.

(3)关键是分析清楚乙射击的三种情况,搞清分类和分步问题,然后准确计算.

解:设甲、乙两人各射击1次击中目标分别记为事件A、B. 则 $P(A) = \frac{2}{3}$, $P(B) = \frac{3}{4}$.

(1) 甲射击4次,全部击中目标的概率为

$$C_4^4 P^4(A) \cdot [1 - P(A)]^0 = \left(\frac{2}{3}\right)^4 = \frac{16}{81}.$$

所以,甲射击4次至少有1次未击中目标的概率为 $1 - \frac{16}{81} = \frac{65}{81}$.

答:甲射击4次,至少有1次未击中目标的概率为 $\frac{65}{81}$.

(2) 甲、乙两人各射击4次,甲恰好只击中目标两次的概率为

$$C_4^2 P^2(A) \cdot [1 - P(A)]^2 = 6 \times \left(\frac{2}{3}\right)^2 \times \left(\frac{1}{3}\right)^2 = \frac{8}{27};$$

乙恰好击中3次的概率为

$$C_4^3 \cdot P^3(B) \cdot [1 - P(B)]^1 = 4 \times \left(\frac{3}{4}\right)^3 \times \left(1 - \frac{3}{4}\right) = \frac{27}{64}.$$

所以甲、乙两人各射击4次,甲恰好击中两次且乙恰好击中目标3次的概率为 $\frac{8}{27} \times \frac{27}{64} = \frac{1}{8}$.

(或 $C_4^2 P^2(A) \cdot [1 - P(A)]^2 \cdot C_4^3 \cdot P^3(B) [1 - P(B)] = 6 \times \left(\frac{2}{3}\right)^2 \times \left(\frac{1}{3}\right)^2 \times 4 \times \left(\frac{3}{4}\right)^3 \times$

$$\left(1 - \frac{3}{4}\right) = \frac{1}{8}.)$$

(3) 乙恰好射击5次后,被中止射击,意味第三次射击击中目标,第4、5次射击未击中目标,第1、2次射击,至少有1次击中目标.

用“ \checkmark ”表示击中,用“ \times ”表示未击中目标,乙的命中情况如下表:

1	2	3	4	5
\checkmark	\checkmark	\checkmark	\times	\times
\checkmark	\times	\checkmark	\times	\times
\times	\checkmark	\checkmark	\times	\times

所以乙恰好射击5次后,被中止射击的概率为 $\left(\frac{3}{4}\right)^3 \times \left(\frac{1}{4}\right)^2 + \left(\frac{3}{4}\right)^2 \times \left(\frac{1}{4}\right)^3 + \left(\frac{3}{4}\right)^2 \times \left(\frac{1}{4}\right)^3 = \frac{45}{1024}$.

点评:(1)(2)可以直接使用独立重复试验的概率公式进行计算.

(3)中必须搞清楚乙被中止射击时的各种情况,然后利用分步或分类公式进行计算.

(3)也可以利用逆向思维来思考:

$$\begin{aligned} & \{1 - C_2^2 [1 - P(B)]^2\} P(B) \cdot [1 - P(B)]^2 \\ &= \left[1 - \left(\frac{1}{4}\right)^2\right] \times \frac{3}{4} \times \left(\frac{1}{4}\right)^2 \\ &= \frac{15}{16} \times \frac{3}{4} \times \frac{1}{16} \\ &= \frac{45}{1024}. \end{aligned}$$

随堂反馈

1. 下面给出四个随机变量:①某传呼台5分钟内接到的呼叫次数;②某人射击一次击中的环数;③公交车司机每天收取的费用;④一年内某棵树生长的高度.其中离散型随机变量的个数有().

A. 0个 B. 1个 C. 2个 D. 3个

2. 已知随机变量 ξ 的分布列为 $P(\xi = k) = \frac{1}{2^k}$ ($k=1, 2, \dots$), 则 $P(2 < \xi \leq 4)$ 等于().

A. $\frac{3}{16}$ B. $\frac{1}{4}$ C. $\frac{1}{16}$ D. $\frac{5}{16}$

3. 投掷一枚硬币,在第5次掷出后出现第一次反面朝上的概率是().

A. $\frac{1}{16}$ B. $\frac{1}{32}$ C. $\frac{1}{48}$ D. $\frac{1}{24}$



4. 某运动员在练习投篮时,每次投中的概率都是 0.6. 如果他投篮 5 次,恰好投中 2 次的概率是_____.

5. 已知随机变量 ξ 的分布列为:

ξ	-2	-1	0	1	2	3
P	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{4}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{12}$

若 $P(\xi^2 < x) = \frac{11}{12}$, 则实数 x 的取值范围是_____.

6. 设随机变量 ξ 的分布列为 $P(\xi = i) = \frac{i}{10}$ ($i = 1, 2, 3, 4$). 求:

(1) $P(\xi = 1 \text{ 或 } \xi = 2)$;

(2) $P\left(\frac{1}{2} < \xi < \frac{7}{2}\right)$.

巩固升华

1. 设随机变量 ξ 的概率分布为 $P(\xi = k) = \frac{c}{k+1}$ ($k=0, 1, 2, 3$), 则 $c =$ _____.

2. 如果 $\xi \sim B\left(15, \frac{1}{4}\right)$, 则使 $P(\xi = k)$ 最大的 k 值是().

A. 3 B. 4 C. 4 或 5 D. 3 或 4

3. 甲与乙两人掷硬币,甲用一枚硬币掷 3 次,记正面朝上的次数为 ξ ;乙用这枚硬币掷两次,记正面朝上的次数为 η . 规定:若 $\xi > \eta$, 则甲获胜;若 $\xi < \eta$, 则乙获胜. 分别求出甲和乙获胜的概率.

4. 在 10 件产品中有 2 件次品,连续抽 3 次,每次抽 1 个,求:

(1) 不放回抽样时,抽到次品数 ξ 的分布列.

(2) 放回抽样时,抽到次品数 η 的分布列.

5. 某射手有 5 发子弹,射击一次命中概率为 0.9, 如果命中就停止射击,否则一直到子弹用尽,求耗用子弹数 ξ 的分布列.

思维拓展

某商店采用“购物摸球中奖”促销活动,摸奖处袋中装有 10 个号码为 n ($1 \leq n \leq 10, n \in \mathbb{N}^*$), 重量为 $f(n) = n^2 - 9n + 21$ (克)的球,摸球方案见下表:

方案	摸奖办法	奖金
①	凡一次购物在 [50, 100] 元者,摸球一个,若球的重量小于该球的号码数,则中奖.	10 元
②	凡一次购物在 100 元以上者,同时摸出两球,若两球的重量相等,则中奖.	40 元

说明:凭购物发票到摸奖处,按规定方案摸奖;这些球以等可能性从袋中摸出,假定符合条件的顾客均参加摸奖.

试比较方案①与②的中奖概率的大小.

1.2 离散型随机变量的期望与方差

学法指导

了解离散型随机变量的期望、方差的意义,会根据离散型随机变量的分布列求出期望、方差.

本节的重点是离散型随机变量的期望与方差



的概念. 难点是离散型随机变量的期望与方差的定义引入及计算公式的给出.

数学期望和方差、标准差都是离散型随机变量最重要的特征数(或数字特征),它们分别反映了随机变量取值的平均水平、稳定程度、集中与离散的程度. 离散型随机变量的期望与方差都与随机变量的分布列有着密切关系,方差又与数学期望紧密相连. 求期望与方差的首要任务是求分布列,因此要弄清离散型随机变量的分布列、期望、方差之间的联系.

(1) 计算顺序:分布列→期望→方差;

(2) 概率的联系. 期望和方差是从整体和全局上分别描述随机变量取值的平均水平和平均偏离程度,它们都离不开基本概念.

基础知识及重点难点

1. 期望

(1) 概念分析

① 随机变量的数学期望表示了随机变量在随机试验中所取的平均值,又由于离散型随机变量的期望的计算是从它的概率分布出发的,因而期望就是离散型随机变量的概率平均值.

② 课本中给出的离散型随机变量的数学期望定义实质上是不严格的,所以课本中涉及的离散型随机变量所有可能取的不同值的个数是有限的,这个定义对于离散型随机变量取有限个值是成立的. 今后不作特别说明,离散型随机变量的取值均为有限个不同值.

(2) 根据离散型随机变量的期望的概念和意义,在实际应用中,我们可以用它来解决一些问题和作出科学的决策. 如本章引言中的商场促销.

(3) 关于随机变量的函数 $\eta = a\xi + b$ 的期望的计算公式的理解,关键是弄清 $\eta = a\xi + b = ax_i + b$ 的重要条件是 $\xi = x_i$, 从而有 $P(\eta = ax_i + b) = P(\xi = x_i)$ ($i=1, 2, \dots$), 由此可得到 η 的分布列,由期望的定义求得 η 的数学期望 $E\eta = aE\xi + b$.

2. 方差

(1) 方差的概念较难理解,因此课本采用与初中代数中介绍的一组数据的方差定义类比的方法,直接定义离散型随机变量 ξ 的方差. 方差体现了随机变量所取的值相对于它的期望的集中与离散、稳定与波动的程度. 它是继数学期望后的另一种随机变量的重要数字特征,在现实生活中有广泛的应用.

(2) 方差与标准差的计算较复杂,教材只要求能根据定义求出离散型随机变量的方差和标准差. 另外,为方便计算,课本上直接给出了两个计算方差的简单公式:

① $D(a\xi + b) = a^2 D\xi$; ② 如果 $\xi \sim B(n, p)$, 则 $D\xi = npq$.

这两个公式只要求会应用就行了.

典型例题详解及解题方法透析

[例1] 已知随机变量 ξ 的期望 $E\xi = 4$, 方差 $D\xi = 1$, 则 $\eta = 2\xi + 5$ 的期望 $E\eta =$ _____, 方差 $D\eta =$ _____.

分析: 直接利用期望和方差的性质即可.

解: $E(a\xi + b) = aE\xi + b$,

$E\eta = E(2\xi + 5) = 2E\xi + 5 = 2 \times 4 + 5 = 13.$

又 $D(a\xi + b) = a^2 D\xi$,

$D\eta = D(2\xi + 5) = 4D\xi = 4 \times 1 = 4.$

答案: 13 4.

点评: 期望与方差的性质非常重要,应用它可以解很多题目.

[例2] 先交 5 元钱,可以参加一次摸奖. 一袋中有同样大小的球 10 个,其中有 8 个标有 1 元钱, 2 个标有 5 元钱,摸奖者只能从中任取两个球,他所得的奖励是所抽两球的钱数之和. 试判断抽奖人是获利还是亏本.

分析: 抽到的两个球上的钱数之和 ξ 是个随机变量,其每一个 ξ 取值时所代表的随机事件的概率值是容易获得的,本题的目标是求参加摸奖的人获利或亏本 η 的数学期望. 由 ξ 与 η 关系为 $\eta = \xi - 5$, 利用公式 $E\eta = E\xi - 5$ 可获解答.

解: 设 ξ 为抽到的两球钱数之和,则 ξ 的可能取值如下:

$\xi = 2$ (抽到两个 1 元),

$\xi = 6$ (抽到 1 个 1 元, 1 个 5 元),

$\xi = 10$ (抽到 2 个 5 元).

$P(\xi = 2) = \frac{C_8^2}{C_{10}^2} = \frac{28}{45}, P(\xi = 6) = \frac{C_8^1 C_2^1}{C_{10}^2} =$

$\frac{16}{45}, P(\xi = 10) = \frac{C_2^2}{C_{10}^2} = \frac{1}{45}.$

$E\xi = 2 \times \frac{28}{45} + 6 \times \frac{16}{45} + 10 \times \frac{1}{45} = \frac{162}{45}.$

又设 η 为抽奖者获利或亏本的可能值,则 $\eta = \xi - 5$, 所以抽奖者获利或亏本的期望为:



$$E\eta = E\xi - 5 = \frac{162}{45} - 5 = -\frac{7}{5} = -1.4.$$

答 抽奖者平均每次亏 1.4 元.

点评 要分清楚是谁获利,不能忽视了先交 5 元才能参加这一抽奖,因此,不能只计算 $E\xi$. 最终 $E\eta$ 的结果为负值,说明摸奖者若重复这种抽奖,平均每摸一次要亏 1.4 元.

〔例 3〕甲、乙两个野生动物保护区有相同的自然环境,且野生动物的种类和数量也大致相等,而两个保护区每个季度发现违反保护条例的事件次数的分布列分别为:

甲保护区

ξ	0	1	2	3
P	0.3	0.3	0.2	0.2

乙保护区

η	0	1	2
P	0.1	0.5	0.4

试评定这两个保护区的管理水平.

分析:首先比较甲、乙两个保护区内一个季度发生的违规事件的次数的期望值是否相同,若相同,然后再看发生的违规次数的波动情况.

解:甲保护区的违规次数 ξ 的数学期望和方差分别是:

$$E\xi = 0 \times 0.3 + 1 \times 0.3 + 2 \times 0.2 + 3 \times 0.2 = 1.3,$$

$$D\xi = (0 - 1.3)^2 \times 0.3 + (1 - 1.3)^2 \times 0.3 + (2 - 1.3)^2 \times 0.2 + (3 - 1.3)^2 \times 0.2 = 1.21.$$

乙保护区的违规次数 η 的数学期望和方差分别是:

$$E\eta = 0 \times 0.1 + 1 \times 0.5 + 2 \times 0.4 = 1.3,$$

$$D\eta = (0 - 1.3)^2 \times 0.1 + (1 - 1.3)^2 \times 0.5 + (2 - 1.3)^2 \times 0.4 = 0.41.$$

答 因为 $E\xi = E\eta$, $D\xi > D\eta$, 所以两个保护区内每季度发生的违规平均次数是相同的,但乙保护区内的违规次数更集中和稳定,而甲保护区的违规次数则相对分散.

点评 数学期望仅体现了随机变量取值的平均值大小,但有时仅知道均值大小还是不够的,比如:两个随机变量的均值相等了,这就还需要知道随机变量的取值如何在均值周围变化,即计算其方差(或是标准差).方差大则说明随机变量取值分散性大,方差小则说明取值分散性小或者说取值比较集中、稳定.

〔例 4〕对一批产品进行检查,每次任取一件产品,检查后放回再任取一件产品,如此反复进行.如果发现次品,则认为这批产品不合格,立即停止检查;如果连续取出 5 件产品都是合格产品,则认为这批产品合格也停止检查.若每批产品的次品率为 0.2,求平均每批要抽多少件?

分析:因为抽到次品时即停止检查,连续取出 5 件产品都是合格品时,也停止检查,所以检查的件数 ξ 可能的取值为 1, 2, 3, 4, 5. 特别注意 $\xi = 5$ 时的概率的求法,按照题意,若要检查 5 次,则前 4 次一定都是合格品,而第 5 次不管抽出的产品是否合格,都要停止检查,故 $\xi = 5$ 时,概率 $P_5 = 0.8^4 \times (0.2 + 0.8) = 0.8^4$.

解:设每批产品要检查的件数为 ξ , ξ 为一随机变量,依题意 ξ 可能取的值为 1, 2, 3, 4, 5, 分布列为:

ξ	1	2	3	4	5
P	0.2	0.8×0.2	$0.8^2 \times 0.2$	$0.8^3 \times 0.2$	0.8^4

$$\text{则 } E\xi = 1 \times 0.2 + 2 \times 0.8 \times 0.2 + 3 \times 0.8^2 \times 0.2 + 4 \times 0.8^3 \times 0.2 + 5 \times 0.8^4 = 3.3616,$$

即平均每批抽查件数为 3.3616.

点评:此题由随机变量的期望公式入手解决.

〔例 5〕一个人有 n 把钥匙,其中只有一把能打开他的房门,他随意地进行试开,并将试开不对的除去,求打开房门所试开次数 ξ 的数学期望与方差.

分析:求 $P(\xi = k)$ 时,由题意知,前 $k - 1$ 次没打开锁,恰在第 k 次打开锁,可逐一从 $k = 1, 2, \dots, n$ 去分析、归纳.

解:设此人将 n 把钥匙排成一排,逐一试开,由于总的排法种数为 $n!$,以上每种排法是等可能的,所以每种排法的概率为 $\frac{1}{n!}$, ξ 的值为 $\{1, 2, 3, \dots, n\}$, 对 $k \in \{1, 2, 3, \dots, n\}$, 事件 $\xi = k$ 表示能打开门的那把钥匙放在第 k 个位置,这样的排列数为 $(n - 1)!$, 所以 ξ 的分布列为

$$P(\xi = k) = \frac{(n - 1)!}{n!} = \frac{1}{n} (k = 1, 2, 3, \dots, n).$$

$$E\xi = \sum_{k=1}^n \frac{k}{n} = \frac{1}{n} \sum_{k=1}^n k = \frac{n+1}{2}, E\xi^2 = \sum_{k=1}^n \frac{k^2}{n} = \frac{1}{n} \sum_{k=1}^n k^2 = \frac{(n+1)(2n+1)}{6}.$$

$$\text{所以 } D\xi = E\xi^2 - (E\xi)^2 = \frac{(n+1)(2n+1)}{6} -$$

$$\frac{(n+1)^2}{2^2} = \frac{n^2-1}{12}.$$

点评:个数较小时,可以逐个处理,但当个数较多时,要注意归纳总结,探讨规律,培养自己的能力.

知识应用及探索创新

【例】某寻呼台共有客户 3 000 人,若寻呼台准备了 100 份小礼品,邀请客户在指定时间来领取,假设任一客户去领奖的概率为 4%. 问寻呼台能否向每一个客户都发出领奖邀请?若能使每一位领奖人都得到礼品,寻呼台至少应准备多少礼品?

分析:可能来多少人,是一个随机变量.由于每人是否去领奖,相互间是独立的,因而随机变量服从二项分布,用数学期望来反映平均领奖人数,即能说明是否可行.

解:设来领奖的人数 $\xi = k (k = 0, 1, 2, \dots, 3\ 000)$, 所以 $P(\xi = k) = C_{3\ 000}^k (0.04)^k (1 - 0.04)^{3\ 000 - k}$, 则 $\xi \sim B(3\ 000, 0.04)$, 那么 $E\xi = 3\ 000 \times 0.04 = 120(\text{人}) > 100(\text{人})$.

答:寻呼台不能向每一位客户都发送领奖邀请.若要使每一位领奖人都得到礼品,寻呼台至少应准备 120 份礼品.

点评:数学期望反映了随机变量取值的平均水平,用它来刻画、比较和描述取值的平均情况,在一些实际问题中有重要的价值,因此,要想到用期望来解决这一问题.

典型考题详解及考点透析

【例 1】(浙江 2005 年高考题)袋子 A 和 B 中装有若干个均匀的红球和白球,从 A 中摸出一个红球的概率是 $\frac{1}{3}$,从 B 中摸出一个红球的概率为 p.

(1) 从 A 中有放回地摸球,每次摸出一个,有 3 次摸到红球即停止.

① 求恰好摸 5 次停止的概率;

② 记 5 次之内(含 5 次)摸到红球的次数为 ξ , 求随机变量 ξ 的分布列及数学期望 $E\xi$.

(2) 若 A, B 两个袋子中的球数之比为 1:2, 将 A, B 中的球装在一起后,从中摸出一个红球的概率是 $\frac{2}{5}$, 求 p 的值.

分析:(1) 求分布列及数学期望,本题是符合

几何分布的一个分布列,可以运用几何分布列公式.

(2) 这是一个二项分布列,可以运用二项分布公式.

(3) 利用概率基本定义.

$$\text{解: (1) ① } C_4^2 \times \left(\frac{1}{3}\right)^2 \times \left(\frac{2}{3}\right)^2 \times \frac{1}{3} = \frac{8}{81}.$$

② 随机变量 ξ 的取值为 0, 1, 2, 3.

由 n 次独立重复试验概率公式

$$P_n(k) = C_n^k p^k (1-p)^{n-k}, \text{ 得}$$

$$P(\xi=0) = C_5^0 \times \left(1 - \frac{1}{3}\right)^5 = \frac{32}{243},$$

$$P(\xi=1) = C_5^1 \times \frac{1}{3} \times \left(1 - \frac{1}{3}\right)^4 = \frac{80}{243},$$

$$P(\xi=2) = C_5^2 \times \left(\frac{1}{3}\right)^2 \times \left(1 - \frac{1}{3}\right)^3 = \frac{80}{243},$$

$$P(\xi=3) = 1 - \frac{32 + 80 \times 2}{243} = \frac{17}{81}.$$

ξ	0	1	2	3
P	$\frac{32}{243}$	$\frac{80}{243}$	$\frac{80}{243}$	$\frac{17}{81}$

ξ 的数学期望是

$$E\xi = \frac{32}{243} \times 0 + \frac{80}{243} \times 1 + \frac{80}{243} \times 2 + \frac{17}{81} \times 3 = \frac{131}{81}.$$

(2) 设袋子 A 中有 m 个球,则袋子 B 中有 2m

个球. 由 $\frac{\frac{1}{3}m + 2mp}{3m} = \frac{2}{5}$, 得 $p = \frac{13}{30}$.

点评:本题主要考查相互独立事件同时发生的概率和随机变量的分布列、数学期望等概念,同时考查学生的逻辑思维能力.

【例 2】(重庆市 2005 年高考题)在一次购物抽奖活动中,假设某 10 张券中有一等奖券 1 张,可获价值 50 元的奖品,有二等奖券 3 张,每张可获价值 10 元的奖品,其余 6 张没有奖.某顾客从此 10 张券中任抽 2 张,求

(1) 该顾客中奖的概率;

(2) 该顾客获得的奖品总价值 ξ (元)的概率分布列和期望 $E\xi$.

考点透视:综合考查概率、离散型随机变量分布列、期望及排列组合知识.

$$\text{解: (1) } P = \frac{C_4^1 C_6^1 + C_4^2}{C_{10}^2} = \frac{30}{45} = \frac{2}{3} \text{ 或 } P = 1 - \frac{C_6^2}{C_{10}^2} = 1 - \frac{15}{45} = \frac{2}{3}.$$



即该顾客中奖的概率为 $\frac{2}{3}$.

(2) ξ 的所有可能值为 0, 10, 20, 50, 60 (元).

$$P(\xi=0) = \frac{C_6^2}{C_{10}^2} = \frac{1}{3}, P(\xi=10) = \frac{C_3^1 C_6^1}{C_{10}^2} = \frac{2}{5},$$

$$P(\xi=20) = \frac{C_3^2}{C_{10}^2} = \frac{1}{15}, P(\xi=50) = \frac{C_6^1}{C_{10}^2} = \frac{2}{15},$$

$$P(\xi=60) = \frac{C_3^1}{C_{10}^2} = \frac{1}{15}.$$

故 ξ 有分布列:

ξ	0	10	20	50	60
P	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{15}$

$$\begin{aligned} \text{从而期望 } E\xi &= 0 \times \frac{1}{3} + 10 \times \frac{2}{5} + 20 \times \frac{1}{15} + 50 \\ &\times \frac{2}{15} + 60 \times \frac{1}{15} = 16. \end{aligned}$$

点评:(1) 分布列中概率行的各项和为 1 是检查分布列是否正确的一个重要手段.

(2) 要正确理解 ξ 的实际意义, 把 ξ 的值从小到大依次排列起来, 这样处理的好处是能做到不重不漏, 并且条理清楚.

随堂反馈

1. 若 ξ 是离散型随机变量, 则 $E(\xi - E\xi)$ 的值为().

- A. $E\xi$ B. 0 C. $2E\xi$ D. 无法求

2. 随机变量 ξ 的分布列是:

ξ	1	2	3
P	0.4	0.2	0.4

则 $D\xi$ 与 $E\xi$ 分别是().

- A. 2 和 0.8 B. 1.8 和 1
C. 2 和 2 D. 0.8 和 2

3. 若 $\xi \sim B(n, p)$, 且 $E\xi = 6$, $D\xi = 3$, 则 $P(\xi = 1)$ 的值为().

- A. $3 \cdot 2^{-2}$ B. 2^{-4}
C. $3 \cdot 2^{-10}$ D. 2^{-8}

4. 一盒中有 5 个正品和 2 个废品, 每次取 1 个产品, 取出后不再放回, 在取得正品前已取出的废品数为 ξ , 则 $E\xi =$ _____.

5. 设 ξ 是一个随机变量, 若 $D(10\xi + 5) = 10$, 则 $D\xi =$ _____.

6. 甲、乙两人在罚球线投球命中的概率分别

为 $\frac{1}{2}$ 和 $\frac{2}{5}$, 投中得 1 分, 投不中得 0 分.

(1) 甲、乙两人在罚球线各投球一次, 求两人得分之和 ξ 的数学期望;

(2) 甲、乙两人在罚球线各投球两次, 求这四次投球中至少有一次命中的概率.

巩固升华

1. 如果 ξ 是一个离散型随机变量, 那么下列命题中假命题是().

- A. ξ 取每一个可能值的概率是非负实数
B. ξ 取所有可能值的概率之和为 1
C. ξ 取某两个可能值的概率等于分别取其中每个值的概率之和
D. ξ 在某一范围内取值的概率大于它取这个范围内各个值的概率之和

2. 设随机变量 ξ 服从二项分布, 即 $\xi \sim B(n, p)$, 且 $E\xi = 3$, $p = \frac{1}{7}$, 则 $n =$ _____, $D\xi =$ _____.

3. 假设 1 部机器在 1 天内发生故障的概率为 0.2, 机器发生故障时, 全天停止工作. 若 1 周的 5 个工作日里无故障, 可获得利润 10 万元, 发生 1 次故障仍可获得利润 5 万元, 发生两次故障所获得利润为 0 万元, 发生 3 次或 3 次以上故障就要亏损 2 万元, 求 1 周的期望利润是多少?

4. 甲、乙两个盒子中, 各放有 5 个不同的电子元件, 已知: 甲盒子中有两个次品, 乙盒子中有 1 个次品, 其余均为正品, 若从甲、乙两个盒子中各取一个元件进行交换, 求交换后乙盒子中正品元件个数 ξ 的期望.

5. 在某地举办的射击比赛中, 规定每位射手射击 10 次, 每次一发, 记分的规则为: 击中目标一



次得 3 分,未击中目标得 0 分;并且凡参赛的射手一律另加 2 分.已知射手 A 击中目标的概率是 0.8,求射手 A 在比赛中得分的数学期望和方差.

思维拓展

某电器商经过多年的经验发现本店每个月售出某种电器的台数是一个随机变量,它的分布列是:

ξ	1	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

设每售出一台电器,获利 300 元,如果售不出而存积于仓库,则每台每月需花保养费 100 元,问电器商每月初购进多少台这种电器才能使月平均收益最大?

1.3—1.4 抽样方法、总体分布的估计

学法指导

会用随机抽样、系统抽样、分层抽样等常用的抽样方法从总体中抽取样本.理解什么是总体分布,会用样本频率分布去估计总体分布.了解累计频率分布的意义,会根据样本频率的分布求得其累计频率的分布.

本节重点是简单随机抽样、总体分布的概念和总体密度曲线的概念.难点是简单随机抽样的特点的概括,总体分布和总体密度曲线的概念的建立.

统计的基本思想方法是用样本估计总体,即用局部推断整体,这就要求样本应具有很好的代表性,而样本的良好客观代表性,则完全依赖抽样方法,弄清简单随机抽样、系统抽样和分层抽样的客观合理性,从而会在不同的情况下采用适当的抽样方法.

基础知识及重点难点

一、抽样方法

简单随机抽样是系统抽样和分层抽样的基础.三种抽样的共同点都是一种等概率抽样,体现了抽



样的公平性.可从它们各自的定义、特点分别理解三种抽样方法:

1. 简单随机抽样

(1) 它要求被抽取样本的总体的个数有限.这样便于对其中各个个体被抽取的概率进行分析.

(2) 它是从总体中逐个地进行抽取.这样,就便于在抽样实践中进行操作.

(3) 它是一种不放回式的抽样.这样便于进行有关的分析和计算.

2. 系统抽样

(1) 系统抽样适用于总体中的个体数较多的情况.因为这时采用简单随机抽样显得不方便.

(2) 系统抽样与简单随机抽样之间存在着密切联系.即在将总体中的个体均分后的每一段进行抽样时,采用的是简单随机抽样.

3. 分层抽样

(1) 分层抽样适用于总体由差异明显的几部分组成的情况;

(2) 在每一层抽样时,采用简单随机抽样或系统抽样.

二、总体分布的估计

用样本估计总体,是研究统计问题的一个基本思想方法.而对于总体分布,总是用样本的频率分布对它进行估计.对于每个个体所取不同数值较少的总体,常用条形图表示其样本分布,而对于每个个体所取不同数值较多或可以在实数区间内取值的总体,常用频率分布直方图表示其样本分布.

典型例题详解及解题方法透析

【例1】某公司在甲、乙、丙、丁四个地区分别有150个、120个、180个、150个销售点,公司为了调查产品销售的情况,需要从这600个销售点中抽取一个容量为100的样本,记这项调查为①;在丙地区中有20个特大型销售点,要从中抽取7个点调查其销售收入和售后服务等情况,记这项调查为②.则完成①②这两项调查宜采用的抽样方法依次是().

- A. 分层抽样法,系统抽样法
- B. 分层抽样法,简单随机抽样法
- C. 系统抽样法,分层抽样法
- D. 简单随机抽样法,分层抽样法

分析 特别要注意的是系统抽样与分层抽样的区别与联系,系统抽样要求均衡分成几部分,然后从每部分中用简单随机抽样的方法抽取相同数目

的样本.而分层抽样则是根据样本的差异分成几层,然后在各层中用简单随机抽样的方法按各层在总体中所占的比例进行抽样,不要求各部分抽取的样本数相同,但各层之间要有明显的差异.

答案 B.

点评:此类题目主要从基本概念入手考查,着重考查概念、方法和个体样本的数量.

【例2】某工厂中共有职工3000人,其中,中、青、老年职工的比例为5:3:2,从所有职工中抽取一个样本容量为400人的样本,应采取哪种抽样方法较合理?中、青、老年职工应分别抽取多少人?

分析 因为总体由三类差异明显的个体构成,所以应采用分层抽样的方法进行抽取.

解 应采取分层抽样方法,由样本容量为400,中、青、老年职工所占比例为5:3:2,

所以应抽取中年职工为 $400 \times \frac{5}{10} = 200$ (人);

应抽取青年职工为 $400 \times \frac{3}{10} = 120$ (人);

应抽取老年职工为 $400 \times \frac{2}{10} = 80$ (人).

点评:分层抽样在日常生活中应用较广,其抽取样本的步骤尤为重要,应牢记按照相应的比例去抽取.

【例3】对某公司生产的电子元件进行寿命追踪调查,情况如下表:

寿命/h	100~200	200~300	300~400	400~500	500~600
个数	20	30	80	40	30

- (1) 画出频率分布直方图;
- (2) 估计该公司电子原件寿命在400h以上的概率;
- (3) 估计总体的数学期望.

点评:画频率分布直方图,需先列出频率分布表,再由频率分布表求寿命在400h以上的电子元件的概率,进而利用数学期望公式计算求得总体的数学期望.

解:(1) 列出频率分布表,再由频率分布表画出频率分布直方图.

寿命/h	频数	频率	累积频率
100~200	20	0.10	0.10
200~300	30	0.15	0.25
300~400	80	0.40	0.65
400~500	40	0.20	0.85

500~600	30	0.15	1
---------	----	------	---

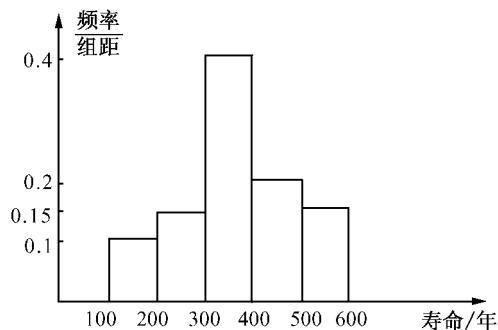


图 1-1 频率分布直方图

(2) 由频率分布表得寿命在 400 h 以上的电子元件出现的频率为 $0.2 + 0.15 = 0.35$.

(3) 样本的期望为:

$$\frac{100+200}{2} \times 0.10 + \frac{200+300}{2} \times 0.15 + \frac{300+400}{2} \times 0.4 + \frac{400+500}{2} \times 0.2 + \frac{500+600}{2} \times 0.15 = 365.$$

可以估计总体数学期望为 365 h.

点评:画频率分布直方图,只要求横轴的单位取得一致、纵轴单位取得一致即可,不要求两者单位相同.

知识应用及探究创新

[例]为了了解某大学一年级新生英语学习的情况,拟从 503 名大学一年级学生中抽取 50 名作为样本,如何采用系统抽样方法完成这一抽样?

分析:由题设条件可知总体的个数为 503,样本容量为 50,不能整除,可采用随机抽样的方法从总体中剔除 3 个个体,使剩下的个体数 500 能被样本容量 50 整除,然后再采用系统抽样方法.

解:第一步,将 503 名学生用随机方式编号为 1, 2, 3, ..., 503.

第二步,用抽签法或随机数表法,剔除 3 个个体,这样剩下 500 名学生,对剩下的 500 名学生重新编号,或采用补齐号码的方式.

第三步,确定分段间隔 $k, k = \frac{500}{50} = 10$ 将总体分为 50 个部分,每一部分包括 10 个个体,这时,第 1 部分的个体编号为 1, 2, ..., 10. 第 2 部分的个体编号为 11, 12, ..., 20. 依此类推,第 50 部分的个体编号为 491, 492, ..., 500.

第四步,在第 1 部分用简单随机抽样确定起始

的个体编号,例如是 5.

第五步,依次在第 2 部分,第 3 部分, ..., 第 50 部分,取出号码为 15, 25, ..., 495 这些数,这样得到一个容量为 50 的样本.

点评:总体中的每个个体,都必须等可能地入样,为了实现“等距”入样且又等概率,因此,应先剔除,再“分段”,后定起始位.采用系统抽样,是为了减少工作量,提高其可操作性,减少人为误差.

典型例题详解及考点透析

[例 1] (湖南省 2005 年高考题)一工厂生产了某种产品 16 800 件,它们来自甲、乙、丙 3 条生产线,为检查这批产品的质量,决定采用分层抽样的方法进行抽样,已知从甲、乙、丙 3 条生产线抽取的个体数组成一个等差数列,则乙生产线生产了 _____ 件产品.

分析:由总数及各层中抽取的样本数成等差数列,可求各层样本数.

解:设从甲、乙、丙抽取的个数分别为 $a - d, a, a + d$,乙生产了 x 件产品,则由 $\frac{3a}{16800} = \frac{a}{x}$ 得 $x = 5600$.

点评:关键是掌握分层抽样的步骤.

随堂反馈

1. 下面不属于抽样方法的是().

- A. 系统抽样法 B. 分层抽样法
C. 简单随机抽样 D. 分类抽样法

2. 一个总体中共有 10 个个体,用简单随机抽样的方法从中抽取一容量为 3 的样本,则某特定个体入样的概率是().

- A. $\frac{3}{C_{10}^3}$ B. $\frac{3}{10 \times 9 \times 8}$
C. $\frac{3}{10}$ D. $\frac{1}{10}$

3. 系统抽样又称为等距抽样,从 N 个个体中抽取 n 个个体为样本,先确定抽样间隔,即抽样间距 $k = \left[\frac{N}{n} \right]$ (取整数部分),从第一段 1, 2, ..., k 个号码中随机抽取一个入样号码 i_0 , 则 $i_0, i_0 + k, \dots, i_0 + (n-1)k$ 号码入样构成样本,所以每个个体的入样概率().

- A. 是相等的 B. 是不相等的
C. 与 i_0 有关 D. 与编号有关



4. 如果采用分层抽样从个体数为 N 的总体中抽取一个容量为 n 的样本,那么每个个体被抽到的概率等于_____.

5. 某人从湖中打了一网鱼,共 m 条,做上记号,再放入湖中,数日后,又打了一网鱼,共 n 条,其中 k 条有记号,由此估计湖中有鱼_____条.

6. 某班有 50 名学生,其中男生 30 人,女生 20 人,为了调查平均身高,准备抽取 $\frac{1}{10}$ 的学生进行测身高,应如何抽样?如果男、女生身高差异较大,又如何抽样?

4. 某校有在校高中生 1 600 人,其中高一学生 520 人,高二学生 500 人,高三学生 580 人.如果想通过抽查其中的 80 人,调查学生的消费情况,考虑学生的年级高低消费情况有明显的差别,而同一年级内消费情况差异较小,问应当采用怎样的抽样方法?高三学生中应抽查多少人?

巩固升华

1. 一个容量为 20 的样本,已知某组的频率为 0.25,则该组的频数为().

- A. 2 B. 5 C. 15 D. 80

2. 某校有老师 200 人,男学生 1 200 人,女学生 1 000 人,现用分层抽样的方法从所有师生中抽取一个容量为 n 的样本,已知从女生中抽取的人数为 80 人,则 $n =$ _____.

3. 某校为了了解学生的课外阅读情况,随机调查了 50 名学生,得到他们在某一天各自课外阅读所用时间的数据,结果用如图 1-2 表示,根据条形图求 50 名学生这一天平均每人的课外阅读时间.

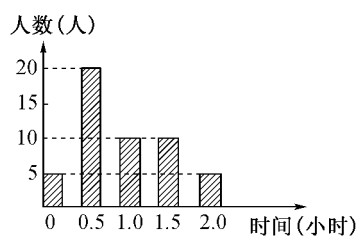


图 1-2



5. 为检测某种产品的质量,抽取了一个容量为30的样本,检测结果为一级品5件,二级品8件,三级品13件,次品4件.

- (1) 列出样本的频率分布表;
- (2) 画出表示样本频率分布的条形图;
- (3) 根据上述结果,估计此种产品为二级品或三级品的概率约是多少?

思维拓展

- (1) 列出样本的频率分布表(含累积频率);
- (2) 画出频率分布直方图和累积频率分布图;
- (3) 根据累积频率分布图,估计身高小于134厘米的人数占总人数的百分比.

1.5—1.6 正态分布与线性回归

学法指导

了解正态分布的意义及主要性质,了解线性回归的方法.

本节重点是正态分布的意义和性质,线性回归的基本思想和方法.本节难点是标准正态曲线的概念,正态分布的性质的抽象与概括,两个变量之间的线性关系的理解.

本部分内容是与生产实际联系得非常密切的,应重点掌握用生产过程中的质量控制图的意义,以及用概率事件原理来分析和计算 $\xi \sim N(\mu, \sigma^2)$ 的有关数据.

基础知识及重点难点

1. 正态分布

正态分布是概率统计中最重要的一种分布.对于正态分布函数,由于中学知识范围的限制,不必去深究它的来龙去脉,但对其函数图象即正态曲线,学生可以自己总结出正态曲线的性质.

对于抽象函数 $\Phi(x_0) = P(x < x_0)$,课本中没有给出具体的表达式,但其几何意义非常明显,即由正态曲线 $N(0, 1)$ 、 x 轴、直线 $x = x_0$ 所围成的图形的面积,再由 $N(0, 1)$ 的曲线关于 y 轴对称,可以得出等式中 $\Phi(-x_0) = 1 - \Phi(x_0)$,以及标准正态总体在任一区间 (a, b) 内取值概率 $P = \Phi(b) - \Phi(a)$.

对于任一正态总体 $N(\mu, \sigma^2)$,其取值小于 x 的概率 $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$.对于这个公式,课本中不加证明地给出,这表明对等式 $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ 的来由不作要求,只要会用它求正态总体 $N(\mu, \sigma^2)$ 在某个特定区间的概率即可.

2. 线性回归

(1) 回归分析是对具有相关关系的两个变量进行统计分析的方法,两个变量具有相关关系是回归分析的前提.

(2) 散点图是定义在具有相关关系的两个变量基础上的,对于性质不明确的两组数据,可先作散点图,在图上看它们有无关系,关系的密切程度,然后再进行相关回归分析.

(3) 求回归直线方程,首先应注意到,只有在散点图大致呈线性时,求出的回归直线方程才有实际意义,否则,求出的回归直线方程毫无意义.

典型例题详解及解题方法透析

[例1] 设 ξ 服从 $N(0, 1)$,求下列各式的值:

- (1) $P(\xi \geq 2.35)$; (2) $P(\xi < -1.24)$; (3) $P(|\xi| < 1.54)$.

分析:因为 ξ 服从标准正态分布,所以可以借助标准正态分布表,查出其值.但由于表中只列出 $x_0 \geq 0$ $P(\xi < x_0) = \Phi(x_0)$ 的情形,故需要转化成小于非负值 x_0 的概率.公式 $\Phi(-x_0) = 1 - \Phi(x_0)$. $P(a < \xi < b) = \Phi(b) - \Phi(a)$, $P(\xi \geq x_0) = 1 - P(\xi < x_0)$ 有其用武之地.

解:(1) $P(\xi \geq 2.35) = 1 - P(\xi < 2.35) = 1 - \Phi(2.35) = 1 - 0.9906 = 0.0094$;

(2) $P(\xi < -1.24) = \Phi(-1.24) = 1 - 0.8925 = 0.1075$;

(3) $P(|\xi| < 1.54) = P(-1.54 < \xi < 1.54) = \Phi(1.54) - \Phi(-1.54) = \Phi(1.54) - [1 - \Phi(1.54)] = 2\Phi(1.54) - 1 = 0.8764$.

点评:从本例可知,在标准正态分布表中只要