

随机变量的期望与方差

一、 随机变量的期望

1. 离散型随机变量的数学期望

数学期望的定义及表示

一般地，设一个离散型随机变量 X 所有可能的取值是 $x_1, x_2, \dots, x_n$ ，这些值对应的概率是 $p_1, p_2, \dots, p_n$ ，则 $E(x) = x_1p_1 + x_2p_2 + \dots + x_np_n$ ，叫做这个离散型随机变量 X 的均值或数学期望（简称期望）。

数学期望的意义

离散型随机变量的数学期望刻画了这个离散型随机变量的平均取值水平。

数学期望的性质

X 为随机变量且 a, b 为常数，记 $Y = aX + b$ ，则 $E(Y) = aE(X) + b$ 。

例1

1. 已知随机变量 ξ 分布列为（如表所示）：设 $\eta = 2\xi + 1$ ，则 η 的数学期望 $E\eta$ 的值是_____。

ξ	-1	0	1
P	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$

练1

2. 已知随机变量 ξ 的分布列如下，那么 ξ 的期望 $E\xi =$ _____，设 $\eta = 2\xi + 1$ ，则 η 的期望 $E\eta =$ _____。

ξ	-1	0	1
P	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$

3. 已知随机变量 ξ 和 η ，其中 $\eta = 4\xi - 2$ ，且 $E\eta = 7$ ，若 ξ 的分布列如下表，则 n 的值为_____。

ξ	1	2	3	4
P	$\frac{1}{4}$	m	n	$\frac{1}{12}$

2. 常见分布列模型的期望

🔔 两点分布

若随机变量 X 服从参数为 p 的两点分布, 则 $E(X) = p$.

🔔 二项分布

若 $X \sim B(n, p)$, 则 $E(X) = np$.

🔔 超几何分布

若离散型随机变量 X 服从参数为 N, M, n 的超几何分布, 则 $E(X) = \frac{nM}{N}$.

【补充说明】

二项分布与超几何分布的期望公式证明难度较大, 主体是利用公式 $kC_n^k = nC_{n-1}^{k-1}$, 感兴趣的同学可以挑战一下.

🔔 例2

4. 随机变量 X 服从二项分布: $X \sim B(10, 0.5)$, 则它的期望 $E(X) = ()$.

- A. 0.5 B. 2.5 C. 5 D. 10

🔔 练2

5. 若随机变量 $X \sim B(2, p)$, 随机变量 $Y \sim B(3, p)$, 若 $P(X = 2) = \frac{4}{9}$, 则 $E(2Y + 1)$ 的值为 _____.

6. 在篮球比赛中, 罚球命中1次得1分, 不中得0分. 如果运动员甲罚球命中的概率是0.8, 记运动员甲罚球1次的得分为 X , 则 $E(X)$ 等于().

- A. 0.2 B. 0.4 C. 0.8 D. 1

3. 期望的实际应用

🔔 超几何分布问题

🔔 例3

7. 一个口袋中装有大小相同的2个白球和3个黑球.

- (1) 采取放回抽样方式, 从中摸出两个球, 求两球恰好颜色不同的概率.
(2) 采取不放回抽样方式, 从中摸出两个球, 求摸得白球的个数的分布列与期望.

8. 在8件获奖作品中，有3件一等奖，有5件二等奖，从这8件作品中任取3件．
- (1) 求取出的3件作品中，一等奖多于二等奖的概率．
- (2) 设 X 为取出的3件作品中一等奖的件数，求随机变量 X 的分布列和数学期望．

练3

9. 十八届四中全会明确提出“以法治手段推进生态文明建设”，为响应号召，某市红星路小区的环保人士向该市政府部门提议“在全市范围内禁放烟花、炮竹”．为此，红星路小区的环保人士对该小区年龄在 $[15, 75)$ 的市民进行问卷调查，随机抽查了50人，并将调查情况进行整理后制成下表：

年龄(岁)	$[15, 25)$	$[25, 35)$	$[35, 45)$	$[45, 55)$	$[55, 65)$	$[65, 75)$
频数	6	10	12	12	5	5
赞成人	3	6	10	6	4	3

- (1) 请估计红星路小区年龄在 $[15, 75)$ 的市民对“禁放烟花、炮竹”的赞成率和被调查者的年龄平均值；
- (2) 若从年龄在 $[55, 65)$ 、 $[65, 75)$ 的被调查者中各随机选取两人进行追踪调查，记被选4人中不赞成“禁放烟花、炮竹”的人数为 ξ ，求随机变量 ξ 的分布列和数学期望．
10. 某区选派7名队员代表本区参加全市青少年围棋锦标赛，其中3名来自A学校且1名为女棋手，另外4名来自B学校且2名为女棋手．从这7名队员中随机选派4名队员参加第一阶段的比赛．
- (1) 求在参加第一阶段比赛的队员中，恰有1名女棋手的概率．
- (2) 设 X 为选出的4名队员中A、B两校人数之差的绝对值，求随机变量 X 的分布列和数学期望．
11. 甲、乙两人参加某种选拔测试．规定每人必须从备选的6道题中随机抽出3道题进行测试，在备选的6道题中，甲答对其中每道题的概率都是 $\frac{3}{5}$ ，乙只能答对其中的3道题．答对一题加10分，答错一题(不答视为答错)得0分．
- (1) 求乙得分的分布列和数学期望．
- (2) 规定：每个人至少得20分才能通过测试，求甲、乙两人中至少有一人通过测试的概率．
12. 一个袋子中装有大小形状完全相同的编号分别为1, 2, 3, 4, 5的5个红球与编号分别为1, 2, 3, 4的4个白球，从中任意取出3个球．
- (1) 求取出的3个球颜色相同且编号是三个连续整数的概率．
- (2) 求取出的3个球中恰有2个球编号相同的概率．
- (3) 记 ξ 为取出3个球中编号的最大值，求 ξ 的分布列与数学期望．

独立重复试验

例4

13. 甲、乙两人按如下规则进行射击比赛，双方对同一目标轮流射击，若一方未击中，另一方可继续射击，甲先射，直到有人命中目标或两人总射击次数达4次为止．若甲击中目标的概率为 $\frac{2}{3}$ ，乙击中目标的概率为 $\frac{1}{2}$ ．
- (1) 求甲在他第二次射击时击中目标的概率．
- (2) 求比赛停止时，甲、乙两人射击总次数 X 的分布列和期望．
14. 某商场搞促销，规定顾客购物达到一定金额可抽奖，最多有三次机会，每次抽中，可依次分别获得20元、30元、50元奖金，顾客每次抽中后，可以选择带走所得奖金，结束抽奖；也可以选择继续抽奖，若有任何一次没有抽中，则连同前面所得奖金也全部归零，结束抽奖，设顾客甲第一次、第二次、第三次抽中的概率分别为 $\frac{3}{4}$ ， $\frac{2}{3}$ ， $\frac{1}{2}$ ，选择继续抽奖的概率均为 $\frac{1}{2}$ ，且每次是否抽中互不影响．
- (1) 求顾客甲第一次抽中，但所得奖金为零的概率．
- (2) 设该顾客所得奖金总数为 X ，求随机变量 X 的分布列和数学期望．

练4

15. 某学校社团招聘工作人员，设置A、B两组测试项目供应聘人员选择，甲、乙、丙、丁四人参加应聘，其中甲、乙、丙三人各自独立参加A组测试，已知甲、乙两人各自通过测试的概率均为 $\frac{1}{2}$ ，丙通过测试的概率为 $\frac{3}{5}$ ．丁参加B组测试，已知B组共有6道试题，丁会做其中的4道题．丁只能且必须选择4道题作答，答对3道题则竞聘成功．
- (1) 求丁应聘成功的概率．
- (2) 记测试通过的总人数为 ξ ，求 ξ 的分布列和期望．
16. 现有4个人去参加某娱乐活动，该活动有甲、乙两个游戏可供参加者选择．为增加趣味性，约定：每个人通过掷一枚质地均匀的骰子决定自己去参加哪个游戏，掷出点数为1或2的人去参加甲游戏，掷出点数大于2的人去参加乙游戏．
- (1) 求这4个人中恰有2人去参加甲游戏的概率；
- (2) 求这4个人中去参加甲游戏的人数大于去参加乙游戏的人数的概率；
- (3) 用 X, Y 分别表示这4个人中去参加甲、乙游戏的人数，记 $\xi = |X - Y|$ ，求随机变量 ξ 的分布列与数学期望 $E\xi$ ．
17. 在盒子里有大小相同，仅颜色不同的乒乓球共10个，其中红球4个，白球3个，蓝球3个．现从中任取一球确定颜色后放回盒子里，再取下一个球．重复以上操作，最多取3次，过程中如果取出蓝色球则不再取球．求：
- (1) 最多取两次就结束的概率．
- (2) 整个过程中恰好取到2个白球的概率．

(3) 设取球的次数为随机变量 X ，求 X 的分布列和数学期望。

二、离散型随机变量的方差

1. 离散型随机变量的方差

🔔 方差的定义及表示

一般地，设一个离散型随机变量 X 所有可能取的值为 $x_1, x_2, \dots, x_n$ ，这些值对应的概率是 $p_1, p_2, \dots, p_n$ ，

则 $D(X) = (x_1 - E(x))^2 \cdot p_1 + (x_2 - E(x))^2 \cdot p_2 + \dots + (x_n - E(x))^2 \cdot p_n$ 叫做这个离散型随机变量 X 的方差。

$D(X)$ 的算术平方根 $\sqrt{D(X)}$ 叫做离散型随机变量 X 的标准差。

🔔 方差的意义

离散型随机变量的方差（标准差）反映了离散随机变量的取值相对于期望的平均波动的大小（离散程度）。

🔔 方差的性质

对于 $a, b \in \mathbf{R}$ ，定义随机变量 $\eta = a\xi + b$ （ ξ 为随机变量），则 $D(\eta) = a^2 D(\xi)$ 。

🔔 方差的另一种求法

$$D(X) = E(X^2) - E^2(X) .$$

🔔 例5

18. 已知随机变量的分布列为：

ξ	-1	0	2
P	x	$\frac{1}{3}$	y

若 $E(\xi) = \frac{1}{3}$ ，则 $x + y = \underline{\hspace{2cm}}$ ； $D(\xi) = \underline{\hspace{2cm}}$ 。

🔔 练5

19. 已知 ξ 的分布列如下：

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

并且 $\eta = 2\xi + 3$, 则方差 $D\eta = (\quad)$.

A. $\frac{179}{36}$

B. $\frac{143}{36}$

C. $\frac{299}{72}$

D. $\frac{227}{72}$

20. 若 ξ 是离散型随机变量, $P(\xi = x_1) = \frac{2}{3}$, $P(\xi = x_2) = \frac{1}{3}$, 且 $x_1 < x_2$, 又已知 $E\xi = \frac{4}{3}$,

$D\xi = \frac{2}{9}$, 则 $x_1 + x_2$ 的值为 $(\quad)$.

A. $\frac{5}{3}$

B. $\frac{7}{3}$

C. 3

D. $\frac{11}{3}$

2. 常见分布模型的方差

🔔 两点分布

若随机变量 X 服从参数为 p 的两点分布, 则 $D(X) = p(1 - p)$.

🔔 二项分布

若 $X \sim B(n, p)$, 则 $D(X) = np(1 - p)$.

🔔 超几何分布

若离散型随机变量 X 服从参数为 N, M, n 的超几何分布, 则 $D(X) = \frac{n(N - n)(N - M)M}{N^2(N - 1)}$

🔔 例6

21. 已知随机变量 $X \sim B(5, \frac{1}{3})$, 则 $E(X) = \underline{\hspace{2cm}}$, $D(X) = \underline{\hspace{2cm}}$.

22. 随机变量 $Y \sim B(n, p)$, 且 $E(Y) = 3.6$, $D(Y) = 2.16$, 则 $(\quad)$.

A. $n = 4, p = 0.9$

B. $n = 9, p = 0.4$

C. $n = 18, p = 0.2$

D. $n = 36, p = 0.1$

🔔 练6

23. 已知随机变量 $\xi \sim B(n, p)$, 若 $E\xi = 4$, $\eta = 2\xi + 3$, $D\eta = 3.2$, 则 $P(\xi = 2) = \underline{\hspace{2cm}}$. (结果用数字表示)

24. 已知某随机变量 X 的分布列如下 $(p, q \in \mathbf{R})$:

X	1	-1
P	p	q

且 X 的数学期望 $E(X) = \frac{1}{2}$, 那么 X 的方差 $D(X) = \underline{\hspace{2cm}}$.

三、 方差的实际应用

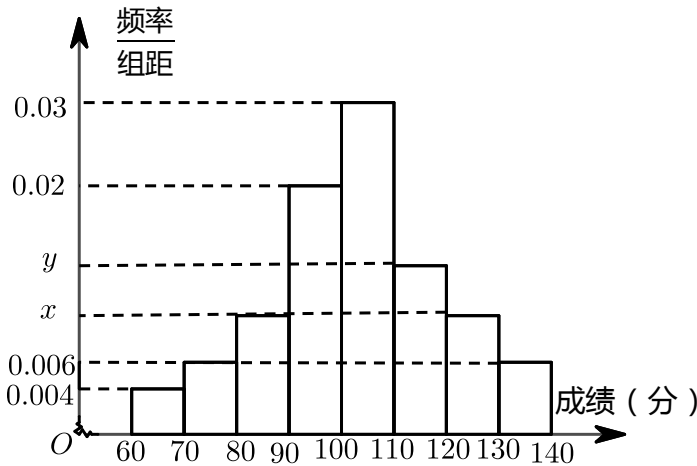
期望反应了随机变量取值的平均水平，而方差反映了随机变量的波动情况．

因此，我们在实际问题的决策时，需要权衡期望与方差孰轻孰重：若平均取值水平比较重要，那么首选决策依据是随机变量的期望；若稳定情况比较重要，那么首选决策依据是随机变量的方差．

在进行优选时一定要对数据进行综合分析，有时候往往需要两者兼顾，当两组随机变量的期望一致时，那么需要考察方差进行进一步筛选．

例7

25. 某学校为了解该校高三年级学生数学学习情况，对广一模考试数学成绩进行分析，从中抽取了 n 名学生的成绩作为样本进行统计，该校全体学生的成绩均在 $[60, 140)$ ，按照 $[60, 70)$ ， $[70, 80)$ ， $[80, 90)$ ， $[90, 100)$ ， $[100, 110)$ ， $[110, 120)$ ， $[120, 130)$ ， $[130, 140)$ 的分组作出频率分布直方图如图(a)所示，样本中分数在 $[70, 90)$ 内的所有数据的茎叶图如图(b)所示．根据上级统计计划出预选分数线，有下列分数与可能被录取院校层次对照表为表(c)．



图(a)

7 2 5 8
8 1 3 5 8 9

图(b)

分数	$[50, 85)$	$[85, 110)$	$[110, 150)$
可能被录取院校层次	专科	本科	重本

图(c)

- (1) 求 n 和频率分布直方图中的 x ， y 的值．
- (2) 根据样本估计总体的思想，以事件发生的频率作为概率，若在该校高三年级学生中任取3人，求至少有一人是可能录取为重本层次院校的概率．
- (3) 在选取的样本中，从可能录取为重本和专科两个层次的学生中随机抽取3名学生进行调研，用 ξ 表示所抽取的3名学生中为重本的人数，求随机变量 ξ 的分布列和数学期望．

练7

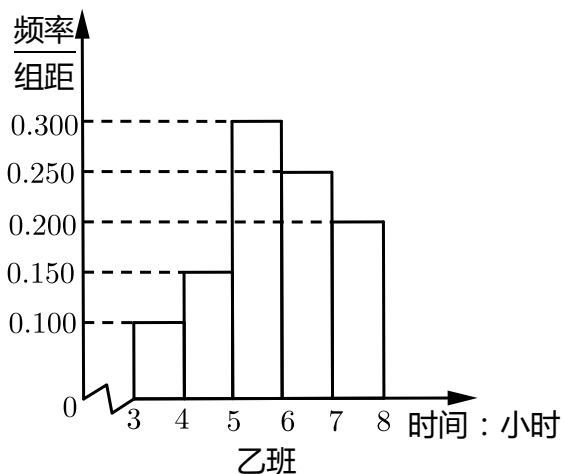
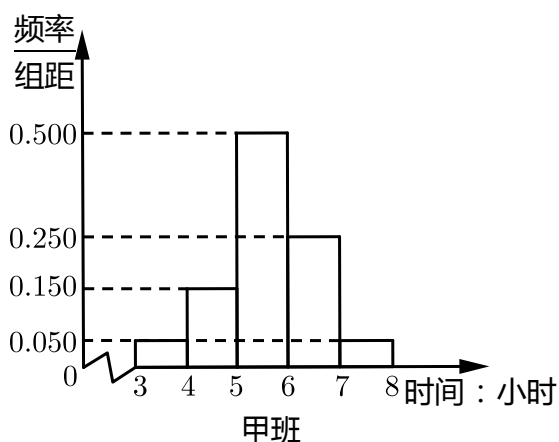
26. 为了解学生暑假阅读名著的情况，一名教师对某班级的所有学生进行了调查，调查结果如下表

阅读名著的本数(本)	1	2	3	4	5

男生人数 (人)	1	4	3	2	2
女生人数 (人)	0	1	3	3	1

- (1) 从这班学生中任选一名男生, 一名女生, 求这两名学生阅读名著本书之和为4的概率?
- (2) 若从阅读名著不少于4本的学生中任选4人, 设选到的男生人数为 X , 求随机变量 X 的分布列和数学期望;
- (3) 判断男学生阅读名著本数的方差 s_1^2 与女学生阅读名著本数的方差 s_2^2 的大小 (只需写出结论).

27. 在全民抗击新冠肺炎疫情期间, 北京市开展了“停课不停学”活动, 此活动为学生提供了多种网络课程资源以供选择使用. 活动开展一个月后, 某学校随机抽取了高三年级的甲、乙两个班级进行网络问卷调查, 统计学生每天的学习时间, 将样本数据分成 $[3, 4)$, $[4, 5)$, $[5, 6)$, $[6, 7)$, $[7, 8]$ 五组, 并整理得到如图所示的频率分布直方图:



- (1) 已知该校高三年级共有600名学生, 根据甲班的统计数据, 估计该校高三年级每天学习时间达到5小时及以上的学生人数.
- (2) 已知这两个班级各有40名学生, 从甲、乙两个班级每天学习时间不足4小时的学生中随机抽取3人, 记从甲班抽到的学生人数为 X , 求 X 的分布列和数学期望.
- (3) 记甲、乙两个班级学生每天学习时间的方差分别为 D_1 , D_2 , 试比较 D_1 与 D_2 的大小. (只需写出结论)

四、综合问题

28. 某城市为鼓励人们绿色出行, 乘坐地铁, 地铁公司决定按照乘客经过地铁站的数量实施分段优惠政策, 不超过30站的地铁票价如下表:

乘坐站数 x	$0 < x \leq 10$	$10 < x \leq 20$	$20 < x \leq 30$
票价 (元)	3	6	9

现有甲、乙两位乘客同时从起点乘坐同一辆地铁，已知他们乘坐地铁都不超过30站．甲、乙乘坐不超过10站的概率分别为 $\frac{1}{4}$ ， $\frac{1}{3}$ ；甲、乙乘坐超过20站的概率分别为 $\frac{1}{2}$ ， $\frac{1}{3}$ ．

(1) 求甲、乙两人付费相同的概率．

(2) 设甲、乙两人所付费用之和为随机变量 X ，求 X 的分布列和数学期望．

29. 为了丰富学生的课余生活，促进校园文化建设，我校高二年级通过预赛选出了6个班（含甲、乙）进行经典美文诵读比赛决赛．决赛通过随机抽签方式决定出场顺序，求：

(1) 甲、乙两班恰好在前两位出场的概率．

(2) 决赛中甲、乙两班之间的班级数记为 x ，求 x 的分布列和均值．

30. 设甲、乙两位同学上学期间，每天7：30之前到校的概率均为 $\frac{2}{3}$ ，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．

(1) 用 X 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 X 的分布列和数学期望．

(2) 设 M 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 M 发生的概率．