新高考下高中概率中的问题

薛美琴，马青莲，魏俊潮

扬州大学　江苏扬州

【摘要】本文主要研究了高考中常出现的概率期望和 K 值问题，主要为二项分布、离散型随机变量的分布和正态分布等模型的期望和一些简单应用，给出以上模型的一些常规方法。

【关键词】概率；期望；图像结合

Problems in the probability of high school under the new college entrance examination

Meiqin Xue, Qinglian Ma, Junchao Wei

Yangzhou University Yangzhou, Jiangsu

【**Abstract**】This paper mainly studies the probability expectation and K value problems that often occur in the college entrance examination, mainly the expectation and some simple applications of the binomial distribution, the distribution of discrete random variables and the normal distribution, and some conventional methods of the above models are given.

【**Keywords**】Probability; Expectation; Image Combination

在解决概率的相关问题时，或许可以尝试先分析出概率模型，而本文结合近几年的高考模拟题，探讨一些基本模型的期望问题、K 值问题、正态分布的一些简单运用的策略，以供参考。

1 有关K^2值问题

例 1、（2022 年新高考 I 卷，20）一医疗团队为研究某地的一种地方性疾病与当地居民的卫生习惯（卫生习惯分为良好和不够良好两种）的关系，在已患该疾病的病例中随机调查了 100 例（称为病例组），同时在未患该疾病的人群中随机调查了 100 人（称为对照组），得到如下表 1 数据：

表 1

	不够良好	良好
病例组	40	60
对照组	10	90

⑴能否有 99%的把握认为患该疾病群体与未患该疾病群体的卫生习惯有差异？

⑵从该地的人群中任选一人，A 表示事件“选到的人卫生习惯不够良好”，B 表示事件“选到的人患有该疾病”，$\frac{P(B|A)}{P(\bar{B}|A)}$与$\frac{P(B|\bar{A})}{P(\bar{B}|\bar{A})}$的比值是卫生习惯不够良好对患该疾病风险程度的一项度量指标，记该指标为 R.

(i)证明：R=$\frac{P(A|B)}{P(\bar{A}|B)}\cdot\frac{P(\bar{A}|\bar{B})}{P(A|\bar{B})}$;

(ii)利用该调查数据，给出$P(A|B),P(A|\bar{B})$的估计值，并利用(i)的结果给出 R 的估计值。

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

表 2

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

解：⑴设零事件H_0：患该疾病群体与未患该疾病群体的卫生习惯无差异

$$K^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}=\frac{200(40\times 90-60\times 10)^2}{100\times 100\times 50\times 150}=24>10.828$$

作者简介：薛美琴，女，在读硕士，研究方向：学科教学（数学）
马青莲，女，在读硕士，研究方向：学科教学（数学）
魏俊潮，男，扬州大学教授，博士，研究方向：代数环论，环上广义逆

查表可知$P(K^2 \geq 10.828) \leq 0.001$

所以假设不成立

故有 99%的把握认为患该疾病群体与未患该疾病群体的卫生习惯有差异

⑵(i)证明：

$$R=\frac{P(B|A)}{P(\bar{B}|A)}\cdot\frac{P(\bar{B}|\bar{A})}{P(B|\bar{A})}=\frac{\frac{P(AB)}{P(A)}\cdot\frac{P(\bar{B}\bar{A})}{P(\bar{A})}}{\frac{P(A\bar{B})}{P(A)}\cdot\frac{P(\bar{A}B)}{P(\bar{A})}}$$

$$=\frac{P(AB)\cdot P(\bar{B}\bar{A})}{P(A\bar{B})\cdot P(\bar{A}B)}=\frac{\frac{P(AB)}{P(B)}\cdot\frac{P(\bar{B}\bar{A})}{P(\bar{B})}}{\frac{P(A\bar{B})}{P(\bar{B})}\cdot\frac{P(\bar{A}B)}{P(B)}}$$

$=\frac{P(A|B)}{P(\bar{A}|B)}\cdot\frac{P(\bar{A}|\bar{B})}{P(A|\bar{B})}$（$P(B)$、$P(\bar{B})$都为正数）

所以 R 等式成立。

(ii)由表 1 可得：$P(A|B)=\frac{40}{100}=\frac{2}{5}, P(\bar{A}|B)=\frac{3}{5}$，$P(A|\bar{B})=\frac{10}{100}=\frac{1}{10}, P(\bar{A}|\bar{B})=\frac{9}{10}$

$$R=\frac{P(A|B)}{P(\bar{A}|B)}\cdot\frac{P(\bar{A}|\bar{B})}{P(A|\bar{B})}=\frac{\frac{2}{5}\times\frac{9}{10}}{\frac{3}{5}\times\frac{1}{10}}=6$$

评注：在本题中，将数据代入K^2的公式当中，从而进行判断，而做这类的题目的一般步骤为：

1、先假设所求事件的关系有无。

2、找出 a、b、c、d 所代表的数字进行代入。

3、分析结果，得出结论。

例 2：（江苏省南京市 2021 届高三 9 月学情调研，19）已知阅读达标情况如下表 1，随机抽取 100 名学生（男生 60 人，女生 40 人），如表 3。

表 3

是否达标 / 性别	不达标	达标
男生	36	24
女生	10	30

问：是否由 99%地把握认为课外阅读达标与性别有关？

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

表 4

$P(K^2 \geq k)$	0.050	0.025	0.010	0.005	0.001
k	3.841	5.024	6.635	7.879	10.828

解：假设事件H_0：课外阅读达标与性别无关。

$$K^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}=\frac{100(36\times30-24\times10)^2}{60\times40\times46\times54}=\frac{2450}{207}\approx11.836>6.635$$

查表可知$P(K^2 \geq 6.635) \leq 0.010$

所以假设不成立

故有 99%的把握认为课外阅读达标与性别有关。

2 数学期望

2.1 求二项分布的数学期望

例 3：（山东省 2021 届高三开学质量检测，20）假设每份样本的检验结果是阳性还是阴性是相互独立的，且每份样本是阳性样本的概率为 0.7，核酸试剂能把阳性样本检测出阳性结果的概率为 0.99（核酸检测存在阳性样本测不出来的情况，但不会把阴性检测呈阳性），求这 1000 份样本中检测呈阳性的份数的期望。

解：由题意得：每份检测出阳性的概率 P=0.7×0.99=0.693

所以 $X \sim B(1000, 0.693)$

所以 $E(X)=1000\times0.693=693$

评注：在本题中，分析出二项分布是解题的关键，考虑到由于检测结果只有两种情况：阴性、阳性，且检测结果相互独立，所以可以判断出该结构为二项分布，进而根据 $X \sim B(n,p)$，$E(X)=np$ 得出数学期望。

2.2 求离散型随机变量的数学期望

例 4、（江苏省南京市 2021 届高三 9 月学情调研，19）已知阅读达标情况如下表 1，随机抽取 100 名学生（男生 60 人，女生 40 人），如表 2。如果用这 100 名学生中男生和女生课外阅读“达标”的概率分布代替该校男生和女生课外阅读“达标”的概率，且每位学生是否“达标”相互独立。现从该校学生中随机抽取 3 人（2 男 1 女），设随机变量 X 表示“3 人中课外阅读达标的人数”，试求 X 的分布列和数学期望。

表 5

是否达标 / 性别	不达标	达标
男生	36	24
女生	10	30

解：记 A 为男生随机抽取一人，达标；B 为女生随便抽取一人，达标。

$$P(A)=\frac{24}{60}=\frac{2}{5}；\quad P(B)=\frac{30}{40}=\frac{3}{4}$$

且 X 的取值为：0，1，2，3

$$P(X=0)=\left(1-\frac{2}{5}\right)^2\times\left(1-\frac{3}{4}\right)=\frac{9}{100}$$

$$P(X=1)=C_2^1\cdot\frac{2}{5}\cdot\left(1-\frac{2}{5}\right)\cdot\left(1-\frac{3}{4}\right)+\left(1-\frac{2}{5}\right)^2\cdot\frac{3}{4}=\frac{39}{100}$$

$$P(X=2)=C_2^1\cdot\frac{2}{5}\cdot\left(1-\frac{2}{5}\right)\cdot\frac{3}{4}+\left(\frac{2}{5}\right)^2\cdot\left(1-\frac{3}{4}\right)=\frac{2}{5}$$

$$P(X=3)=\left(\frac{2}{5}\right)^2\times\frac{3}{4}=\frac{3}{25}$$

所以 X 的分布列为：

X	0	1	2	3
P	$\frac{9}{100}$	$\frac{39}{100}$	$\frac{2}{5}$	$\frac{3}{25}$

数学期望

$$E(X)=0\times\frac{9}{100}+1\times\frac{39}{100}+2\times\frac{2}{5}+3\times\frac{3}{25}=1.55$$

分析：对于复杂事件的概率，可以转化为简单的概率，先算出男生和女生的达标情况，并开始分析。注意：当 X=1 时，可以是两个男生中的一个，也可以是一个女生；当 X=2 时，则是可以都是男生，也可以是一个男生一个女生，这是本题容易忽略的地方，也是本题的关键。

评注：求离散型随机变量 X 的数学期望的一般步骤：

（1）先分析 X 的可取值，根据可取值求出对应的概率；

（2）根据所对应的概率，得到 X 的分布列；

（3）结合分布列，求出数学期望

此题第一问，至少一个需要调整，情况比较多，可以从反面算，即一个都不需要修理的概率，从而可以得出答案。

例 5：（山东省青岛市 2021 届高三第一学期期末学业水平测试）现有一个复原好的三阶魔方，白面朝上，只可以扭动最外侧的六个表面。某人按规定将魔方随机扭动两次，每次均按顺时针方向转动 90°，记顶面白色色块的个数为 X，求 X 的分布列及数学期望 E(X)。

解：X 的取值为：3，4，6，9.

$$P(X=3)=\frac{A_4^1}{6\times6}=\frac{1}{9}$$

$$P(X=4)=\frac{2\times A_4^1}{6\times6}=\frac{2}{9}$$

$$P(X=6)=\frac{A_4^1+2\times A_4^1\times2}{6\times6}=\frac{5}{9}$$

$$P(X=9)=\frac{A_2^1\times2}{6\times6}=\frac{1}{9}$$

所以 X 的分布列为：

X	3	4	6	9
P	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{5}{9}$	$\frac{1}{9}$

数学期望

$$E(X)=3\times\frac{1}{9}+4\times\frac{2}{9}+6\times\frac{5}{9}+9\times\frac{1}{9}=\frac{50}{9}$$

评注：准确列出随机变量，针对不同情况算出概率分布列，因为都是顺时针旋转，所以情况会比较清晰，其中 X=6 时，需要考虑四面的可以连续转两次，所以是A_4^1，另外也可以四面的转了一次，再加一次上底或者下底，同样还有顺序，所以是$2\times A_4^1\times2$。

2.3 正态分布的运用

例 6（多选题）：（江苏省常州市 2021 届高三第二学期学业水平监测期初联考）已知在数学测验中，某校学生的成绩服从正态分布 N（110，81），其中 90 分为及格线，则下列结论正确的有（附：随机变量ξ服从正态分布$N(\mu,\sigma^2)$，则$P(\mu-2\sigma<\xi<\mu+2\sigma=0.9545$）（ABD）

A. 该校学生成绩的期望为 110

B. 该校学生成绩的标准差为 9

C. 该校学生成绩的标准差为 81

D. 该校学生成绩及格率超过 95%

评注：因为该校学生的成绩服从正态分布 N（110，81），则期望$\mu=110$，方差为$\sigma^2=81$，所以标准差为 9，所以 A,B 正确，C 错误；因为$\mu-2\sigma=110-2\times9=92$，所以$P(\xi>90)>P(\xi\geq92)=P(\xi\geq\mu-2\sigma)>0.9545$,所以 D 正确。

例 7：（江苏省南京市、盐城市 2021 届高三第一次模拟考试，20）已知竞赛满分 100 分，大于等于 80 分为优秀，随机抽取 100 人的得分为样本，统计得到样本平均数 71，方差为 81。假设有 10 万人参加了该竞赛活动，得分 Z 服从正态分布 N（71，81）

（1）试估计这次竞赛活动，得分优秀者的人数

是多少万？

（2）“抽奖话费“活动，优秀者可抽奖两次，其余参加者抽奖一次。抽奖者点击抽奖按钮，即随机产生一个两位数（10，11，12，……，99），若产生的量为数的数字相同，则可奖励 40 元电话费，否则奖励 10 元电话费。假设参加竞赛活动的所有人均参加了抽奖活动，试估计这次活动奖励的电话费总额为多少万元。

参考数据：若 Z~N（μ，δ^2），则 P（$\mu-\delta<Z<\mu+\delta$）≈0.68

解：（1）由题意，得：P（71-9<Z<71+9）≈0.68

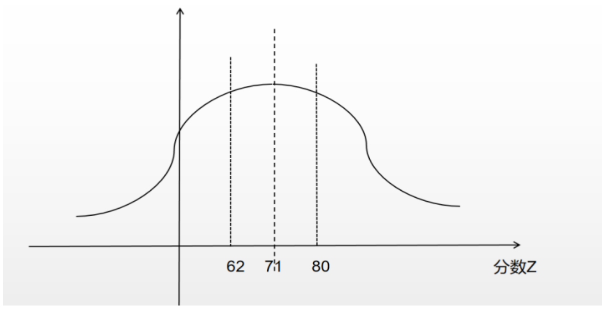

图 1

所以

$$P(Z\geq 80)=\frac{1}{2}\cdot\left[1-P(62<Z<80)\right]=0.16$$

获得优秀人数=0.16×100000=16000

（2）方法一：

设每位参加活动者获得的电话费为 X 元，则 X 的值为 10，20，40，50，80

$$P(X=10)=(1-0.16)\times\frac{81}{90}=\frac{756}{1000}$$

$$P(X=40)=(1-0.16)\times\frac{9}{90}=\frac{84}{1000}$$

$$P(X=20)=0.16\times\left(\frac{81}{90}\right)^2=\frac{1296}{10000}$$

$$P(X=80)=0.16\times\left(\frac{9}{90}\right)^2=\frac{16}{10000}$$

$$P(X=50)=0.16\times\frac{81}{90}\times\frac{9}{90}\times 2=\frac{288}{10000}$$

$$E(X)=10\times\frac{756}{1000}+40\times\frac{84}{1000}+20\times\frac{1296}{10000}$$
$$+80\times\frac{16}{10000}+50\times\frac{288}{10000}=15.08$$

所以总话费=15.08×10 万=150.8 万

方法二：

设抽奖一次获得的话费为 X 元。

$P(X=10)=\frac{81}{90}=\frac{9}{10}$，$P(X=40)=\frac{9}{90}=\frac{1}{10}$

所以抽奖一次获得电话费的期望值为

$$E(X)=10\times\frac{9}{10}+40\times\frac{1}{10}=13$$

因为获得优秀的人数为 1.6 万人

所以一共参加抽奖的次数为 10+1.6=11.6 万次

所以总话费=$11.6\times 13=150.8$(万元)

评注：（1）结合正态分布的图像规律，先求得超过 80 分的概率。

（2）可以从两个角度求出数学期望，法一则是算个人能获得的话费，值得注意的是当 X=50 时，P 有先后顺序。而法二则是计算一次获得话费的概率，但从优秀者的次数着手计算，所以法二很巧妙。

例 8：（江苏省苏州市 2021 届高三期初调研，20）已知将每科考生的原始分从高到低划分为 A,B,C,D,E 共 5 个等级，各等级人数所占比例分别为 15%，35%，35%，13%和 2%，并按给定的公式进行转换赋分。(1)略(2)假设该省此次高一学生生物学科原始 Y 服从正态分布 N（75.8，36）。若 $Y\sim N(\mu,\ \delta^2)$，令$\eta=\frac{Y-\mu}{\delta}$，则$\eta\sim N(0,1)$，请解决下列问题：

①若以此次高一学生生物学科原始分 C 等级的最低分为实施分层教学的划线分，试估计该划线分大约为多少分：（结果保留整数）

②先随机抽取了该省 800 名高一学生的此次生物学科的原始分，若这些学生的原始分相互独立，记ζ为被抽到的原始分不低于 71 分的学生人数，求$P(\zeta=\mathrm{k})$取得最大值时 k 的值。

附：若$\eta\sim N(0,1)$，则 $P(\mu\leq 0.8)\approx 0.788$，$P(\eta\leq 1.04)\approx 0.85$。

解：（2）①$\mu=75.8$，$\delta=6$

所以$\eta=\frac{Y-75.8}{6}$，则$Y=6\eta+75.8$

又因为$P(A)+P(B)+P(C)=0.85$

所以$P(Y\geq\chi)=0.85$

所以

$$P(6\eta+75.8\geq\chi)=P\left(\eta\geq\frac{\chi-75.8}{6}\right)=0.85$$

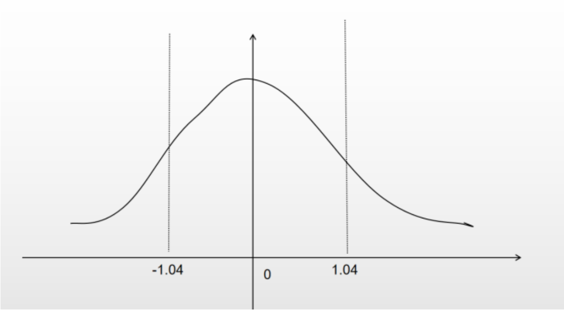

图 2

根据正态分布图像

$$P(\eta \le 1.04) = P(\eta \ge -1.04) \approx 0.85$$

所以 $\frac{\chi - 75.8}{6} = -1.04$

所以 $\chi = 69$

② $P(Y \ge 71) = P(6\eta + 75.8 \ge 71)$

$$= P\left(\eta \ge \frac{71 - 75.8}{6}\right)$$

$$= P(\eta \ge -0.8) = P(\eta \le 0.8) \approx 0.788$$

所以，每个学生生物统考原始分不低于 71 分的概率事件为 0.788

因为原始分相互独立，所以 $\zeta \sim B(800, 0.788)$

$$P(\zeta = \mathrm{k}) = C_{800}^{K} \cdot 0.788^{k} \cdot (1 - 0.788)^{800-k}$$

$$\begin{cases} P(\zeta = \mathrm{k}) \ge P(\zeta = k+1) \\ P(\zeta = k) \le P(\zeta = k-1) \end{cases}$$

解得： $630.188 \le k \le 631.188$

所以 $k = 631$

即 $k = 631$， $P(\zeta = k)$ 取得最大值。

评注：进行正态转化，并结合图形，算出概率，这是本题的关键。第二问 k 取得最大值的条件为 $\begin{cases} P(\xi = k) \ge P(\xi = k+1) \\ P(\xi = k) \ge P(\xi = k+1) \end{cases}$，求出整数的值。

3 总结

面对江苏改用全国卷后，概率这一部分的题，和往年会有不同，题目地位置也从 19 变成 18 再变成 20，再加上这一类题目材料往往比较多，这就需要提高信息提取和加工的能力，同时也需要在平时的运用中，加深对模型的理解，熟悉各种题型，理清概率计算公式的来龙去脉。希望对读者在解决此类问题时提供一些帮助。

参考文献

[1] 2021 年新高考全国卷 I 数学 18 题评析.伊翠红.

[2] 2022 全国高考模拟试卷汇编.

[3] 20222022 年新高考 I 卷.

收稿日期：2022 年 7 月 13 日
出刊日期：2022 年 9 月 12 日
引用本文：薛美琴，马青莲，魏俊潮，新高考下高中概率中的问题[J]. 中小学教育研究, 2022, 1(2) : 1-5.
DOI: 10.12208/j.jrpe.20220027

检索信息：RCCSE 权威核心学术期刊数据库、中国知网（CNKI Scholar）、万方数据（WANFANG DATA）、Google Scholar 等数据库收录期刊
版权声明：©2022 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。http://creativecommons.org/licenses/by/4.0/

OPEN ACCESS