

浅谈随机变量的数字特征

朱先阳

铜仁学院大数据学院, 贵州 铜仁

收稿日期: 2024年9月8日; 录用日期: 2024年10月9日; 发布日期: 2024年10月18日

摘要

随机变量的数字特征是概率论和统计学中的重要概念, 它们提供了关于随机变量分布的关键信息, 同时也是学生不易掌握和应用的一个知识点。在投资组合优化、市场分析、金融建模、信号处理等多领域中有广泛应用, 在解决实际问题扮演着至关重要的作用, 然而学生对这类实际问题常常感到无从下手, 究其原因在于对其概念“知其然, 不知其所以然”, 使得数学概念与实际脱节。本文在概率论与数理统计教材的基础上, 对随机变量的几个重要的数字特征所表达的实际意义进行分析, 使之与生活相通, 与实际相符, 达到学以致用目的。

关键词

期望, 方差, 协方差, 相关系数

A Brief Introduction to the Numerical Characteristics of Random Variables

Xianyang Zhu

School of Data Science, Tongren University, Tongren Guizhou

Received: Sep. 8th, 2024; accepted: Oct. 9th, 2024; published: Oct. 18th, 2024

Abstract

The numerical characteristics of random variables are important concepts in probability theory and statistics, which provide the key information about the distribution of random variables, and are also a knowledge point that is difficult for students to grasp and apply. It is widely used in portfolio optimization, market analysis, financial modeling, signal processing and other fields, and plays a vital role in solving practical problems, however, students often feel helpless about this kind of practical problem, the reason lies in the concept of “knowing it’s so, but don’t know why it’s so”, making mathematical concept out of touch with reality. Based on the teaching materials of

probability theory and mathematical statistics, this paper analyzes the practical significance expressed by several important numerical characteristics of random variables, so as to make them correspond with life and reality, to achieve the purpose of application.

Keywords

Expectation, Variance, Covariance, Correlation Coefficient

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在现实生活中,人们常常会遇到一些涉及随机性的事件,例如掷骰子的点数、购物车中的商品数量等等,这些事件的结果是不可预测的,这种随机现象各种结果的变量就是随机变量。数学上,随机变量表示随机试验各种结果的实值单值函数。随机事件不论与数量是否直接有关,都可以数量化,即都能用数量化的方式表达,它是概率论与数理统计学中的重要概念之一。要全面了解一个随机变量,不但要知道它取哪些值,而且要知道它取这些值的规律,即要掌握它的概率分布。概率分布可以由分布函数 $F(x)$ 刻画,表示落在区间 $(-\infty, x)$ 上的概率,能全面地描述了随机变量的统计规律性,但在实际问题中,有时很难求出随机变量的分布函数或者不需要知道随机变量的一切统计特性,而只需要知道随机变量的某些特征。例如在分析某校学生英语四级水平时,只要计算该校的平均成绩和计算该校每位学生的考试成绩与平时成绩的偏离大小,便可以对该校的学生英语四级水平做出比较客观的判断,这种能表示随机变量某些方面特征的数就是随机变量的数字特征。另外,注意到许多重要分布会含 1 到 3 个参数,而这些参数与数字特征重合或关系密切,因此只要知道分布的类型,通过数字特征就能完全确定分布函数。由此可见,随机变量的数字特征的研究具有理论上和实际上的重要意义。

2. 数学期望

“期望”在我们日常生活中常指有根据的希望,而在概率论中,数学期望源于历史上一个著名的分赌本问题。

问题:赌技相当的甲、乙两人进行一场赌局,各出赌注 50 法郎,每局无平局。谁先赢三局,则得全部的 100 法郎。当甲赢了两局,乙赢了一局时,赌博因故终止。问如何分这 100 法郎才算公平?

首先大家都明白:每人 50 法郎对甲不公平;全部给甲对乙不公平。那么比较合理的分法是,按一定的比例甲多分些,乙少分些。所以问题的交点在于按怎样的比例来分。我们考虑两种分法:

1) 甲得 100 法郎的 $\frac{2}{3}$, 乙得 100 法郎的 $\frac{1}{3}$ 。这是基于已赌的局数:甲赢了两局,乙只赢了一局。

2) 假设再赌下去,则甲最终所得 X 为一随机变量,其可能取值可能取值为 0 或 100。再赌两局,此赌局结束,结果一定是下面的情况之一:甲甲、甲乙、乙甲、乙乙(记录名字表示获胜一局)。在这四种情况中,有三种使甲获胜,只有一种情况(乙乙)下乙获胜。因为赌技相当,所以甲获得 100 法郎的可能性为 $\frac{3}{4}$, 获得 0 法郎的可能性为 $\frac{1}{4}$, 则 X 的分布列为:

X	0	100
p	$\frac{1}{4}$	$\frac{3}{4}$

所以, 甲的“期望”所得应为 $0 \times \frac{1}{4} + 100 \times \frac{3}{4} = 75$ (法郎), 乙得 25 法郎。这种分法不仅考虑了已赌的局数, 还包含了对再赌下去的一种“期望”, 它比第一种的分法更合理。

在上面的例子中, 我们解决问题的方法就是找到如果继续赌下去, 甲可以获得的“平均”钱数。这个“平均”是考虑到了每种回报所出现频率的加权平均。根据概率的定义, 当实验次数不断增加时, 频率会稳定在概率周围波动。因此我们直接用随机变量 X 的概率来进行加权平均的计算。

2.1. 离散型随机变量的数学期望

定义 1 [1] [2] 设离散型随机变量 X 的分布律为

$$P\{X = x_k\} = p_k, k = 1, 2, \dots,$$

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛, 则称级数 $\sum_{k=1}^{\infty} x_k p_k$ 为 X 的**数学期望**, 记作 $E(X)$, 或 EX 。

注 定义中数列 $\{x_k p_k\}$ 中 x_k 是随机变量 X 的取值, 不可能要求 X 按 $x_1, x_2, \dots$ 的顺序逐个去取, 当数列 $\{x_k p_k\}$ 中项的顺序发生变化时, 级数 $\sum_{k=1}^{\infty} x_k p_k$ 可能收敛到不同的数值, 而级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛就能保证无论数列各项的顺序如何变化, 级数和是同一个确定的数。随机变量 X 的数学期望 $E(X)$ 就是 X 的加权平均值, 代表了 X 的取值的平均水平。

2.2. 连续型随机变量的数学期望

连续型随机变量取任意一个数的概率都为 0, 那么对于加权平均值的计算, 就不能简单地用取值乘以概率相加了。

设连续型随机变量 X 的概率密度为 $f(x)(a \leq x \leq b)$ 。

1) 将 $[a, b]$ 任意分为 n 个小区间 $[x_i, x_{i+1}]$ ($i = 1, 2, \dots, n$), 区间长度为 $\Delta x_i = x_{i+1} - x_i$, 其中 $x_1 = a, x_{n+1} = b$;
2) 任取 $\xi_i \in [x_i, x_{i+1}]$ 作为 X 在 $[x_i, x_{i+1}]$ 上取值的代表, 又 X 在 $[x_i, x_{i+1}]$ 取值的概率为图 4.1 中阴影所示小曲边梯形的面积, 现用矩形面积 $f(\xi_i)\Delta x_i$ 近似代替, 则 X 在 $[x_i, x_{i+1}]$ 取值与相应概率的积近似等于 $\xi_i f(\xi_i)\Delta x_i$;

3) 作和 $\sum_{i=1}^n \xi_i f(\xi_i)\Delta x_i$, 则为随机变量 X 期望值的近似值;

4) 无限细分小区间, 令 $\lambda = \max_{1 \leq i \leq n} \{\Delta x_i\}$ 趋于 0, 取极限 $\lim_{\lambda \rightarrow 0} \sum_{i=1}^n \xi_i f(\xi_i)\Delta x_i$, 若极限存在, 则该极限即为连续型随机变量 X 的加权平均值, 也就是数学期望。

看上面四步, 恰为定积分定义, 当 X 取遍实数, 那么有

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx。$$

定义 2 [1] [2] 设连续型随机变量 X 的概率密度为 $f(x)$ 。若广义积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛, 则称 $\int_{-\infty}^{+\infty} xf(x)dx$ 为随机变量 X 的**数学期望**, 记作 $E(X)$, 或 EX 。即

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx。$$

2.3. 实际例子

中国体彩推出一种福利彩票, 一个号码对应一张彩票, 每 50 万张设一个开奖组, 中奖概率及奖金如下:

金额(元)	0	10	50	500	5000	50000	500000
概率 P	0.9	0.09	0.009	0.0009	0.00009	0.000009	0.0000009

问：每张彩票售价多少时可以确保体彩中心不会亏损？
分析：要保证不亏损的话，价格不能低于每张彩票平均获得的金额，也就是概率中的数学期望。

$$E(X)=\sum_{k=1}^{\infty}x_kp_k=3.15。$$

3. 方差

先看这样一个例子：
对袋装糖的包装机，希望每袋糖的重量都是 500 g，实际上不可能，每袋糖的重量是随机变量 X 。当然要求 $E(X)=500\text{ g}$ ，不然或买方吃亏，或卖方吃亏。现在有两台包装机，包装的每袋糖的重量都保证了期望值为 500 g，但是一台包装机包装的每袋糖的重量或者重了 10 g 左右，或者轻了 10 g 左右；另一台包装机包装的每袋糖或者重了 2 g 左右，或者轻了 2 g 左右。比较两台包装机，显而易见，后一台包装机的性能较好，因为尽管二者包装的每袋糖的重量都有波动，然后后者与期望值的偏差程度要小。
我们将如何来构建刻画随机变量取值偏离程度的指标？
再以射击为例进行讨论。有甲、乙两名射手，每次射击命中的环数分别为随机变量 X,Y ，它们的分布律如下表 1、表 2 所示。

Table 1. The probability distribution of the random variable X
表 1. 随机变量 X 的概率分布

X	8	9	10
P{X = K}	0.2	0.6	0.2

Table 2. The probability distribution of the random variable Y
表 2. 随机变量 Y 的概率分布

Y	8	9	10
P{Y = K}	0.1	0.8	0.1

求得 $E(X)=E(Y)=9(\text{环})$ 。可见从均值的角度分不出谁的射击技术更高，于是通常的想法就是看谁命中的环数比较集中于平均值附近，即命中的环数偏离均值的大小程度。
1) 将随机变量的取值与期望值直接作差的加权平均。如对 X 计算得到：
$$(8-9)\times0.2+(9-9)\times0.6+(10-9)\times0.2=0。$$

对 Y 计算也如此。因为出现正负项刚好抵消，这样做不能说明问题。
2) 为克服正负抵消的缺点，采用将随机变量的取值与期望值直接作差的绝对值的加权平均，但这样往往会增加计算的不便。
3) 综合以上两种想法，由于要衡量的是偏离程度，而不是偏离多少，所以想到用随机变量的取值与期望值作“差”的平方的加权平均，则对 X,Y 计算得到

$$(8-9)^2\times0.2+(9-9)^2\times0.6+(10-9)^2\times0.2=0.4；$$

$$(8-9)^2 \times 0.1 + (9-9)^2 \times 0.8 + (10-9)^2 \times 0.1 = 0.2。$$

如此可知，乙的射击技术更高。

定义 3 [3] [4] 设 X 是一随机变量，若 $E[X - E(X)]^2$ 存在，则称其为 X 的方差，记为 $D(X)$ ，即 $D(X) = E[X - E(X)]^2$ 。

称 $\sqrt{D(X)}$ 为随机变量 X 的标准差或均方差，记为 $\sigma(X)$ 。

注 1) 按此定义，方差是随机变量 X 的函数 $g(X) = [X - E(X)]^2$ 的数学期望，所以得到：若 X 是一离散型随机变量，其分布律为 $P\{X = x_k\} = p_k (k=1, 2, \dots)$ ，则

$$D(X) = \sum_{k=1}^{\infty} [x_k - E(X)]^2 p_k；$$

若 X 是一连续型随机变量，其概率密度为 $f(x)$ ，则

$$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx。$$

2) 方差反映了随机变量离数学期望的平均偏离程度。如果随机变量 X 的期望与方差分别为 $E(X), D(X)$ ，那么对任意的 $\varepsilon > 0$ ，事件 $\{|X - E(X)| \geq \varepsilon\}$ 发生的概率 $P\{|X - E(X)| \geq \varepsilon\}$ 应该与 $D(X)$ 有一定的关系。大致来说，如果 $D(X)$ 越大，那么 $P\{|X - E(X)| \geq \varepsilon\}$ 也会大一些，把这个直觉严格化，就是著名的切比雪夫不等式。

4. 协方差

对于二维随机变量 (X, Y) ， X, Y 本身就是一维随机变量，可以定义其数学期望和方差，这在前面已经进行了讨论。同时，由于 X, Y 也是同一随机实验的结果，一般具有依存关系，而其数学期望和方差都没有反映它们之间的联系。

在变量之间，线性关系是一种既简单又重要的关系，对随机变量来说也是一样。协方差，就是对随机变量之间线性关系强弱给以量的刻画的数字特征。

定义 4 [5] [6] 设 (X, Y) 是二维随机变量。若 $E[(X - E(X))(Y - E(Y))]$ 存在，则称它为 X 和 Y 的协方差，记作 $\text{cov}(X, Y)$ ，即

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

注 1) “协方差”名称的由来：“协”即“协同”的意思。 X 的方差是 $(X - E(X))$ 与 $(X - E(X))$ 的乘积的期望，现在把一个 $(X - E(X))$ 换成 $(Y - E(Y))$ ，其形式接近方差，又有 X, Y 两随机变量的参与，由此得出协方差的名称；而且方差是协方差的特例，事实上有，

$$\text{cov}(X, X) = E[(X - E(X))(X - E(X))] = D(X)。$$

2) 若 (X, Y) 是二维离散型随机变量，其联合分布律为 $P\{X = x_i, Y = y_j\} = p_{ij}$ ， $i, j = 1, 2, \dots$ ，则有

$$\text{cov}(X, Y) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} [x_i - E(X)][y_j - E(Y)] p_{ij}。$$

若 (X, Y) 是二维连续型随机变量，其概率密度为 $f(x, y)$ ，则有

$$\text{cov}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [x - E(X)][y - E(Y)] f(x, y) dx dy。$$

3) 协方差 $\text{cov}(X, Y)$ 是描述 X 与 Y 线性关系强弱的指标。

事实上, 对上面的(3)解释如下:

a) 假若随机变量 (X, Y) 所有可能的取值都在一条直线上, 那么 $(E(X), E(Y))$ 也在此直线上。我们以 $(E(X), E(Y))$ 为原点, 再作直角坐标系, 则 $(X - E(X))$ 和 $(Y - E(Y))$ 的取值要么同号, 要么异号, 这样使得 $E[(X - E(X))(Y - E(Y))]$ 的值的绝对值很大(相对于本身的可能取值, 因为没有项被抵消)。此时 X 与 Y 之间线性关系最强。

b) 假若随机变量 (X, Y) 所有可能的取值有的在直线上, 有的在此直线附近, 那么点 $(E(X), E(Y))$ 也在此直线上或此直线附近。我们以 $(E(X), E(Y))$ 为原点, 再作直角坐标系, 则 $(X - E(X))$ 和 $(Y - E(Y))$ 的取值要么大多数都同号, 要么大多数都异号, 这样使得 $E[(X - E(X))(Y - E(Y))]$ 的值的绝对值比较大(相对于本身的可能取值, 因为有正负项被抵消, 但不多)。此时 X 与 Y 之间线性关系较强。

c) 假若随机变量 (X, Y) 所有可能的取值在 XOY 平面上分的较散, 使得在以 $(E(X), E(Y))$ 为原点的新直角坐标系中, 四个象限中均有存在, 则 $(X - E(X))$ 和 $(Y - E(Y))$ 的取值乘积有正、有负, 这样使得 $E[(X - E(X))(Y - E(Y))]$ 的值的绝对值小, 而且有可能接近于 0(相对于本身的可能取值, 因为有正负项被抵消的多)。此时 X 与 Y 之间线性关系弱。

5. 相关系数

在前面讨论的协方差可以刻画随机变量 X 与 Y 之间线性关系的强弱, 同时从中可知也有其不足之处:

1) 协方差 $\text{cov}(X, Y)$ 的绝对值越大, X 与 Y 之间线性关系就越强, 但多大才算大? 给不出明确的量的指标;

2) 协方差是有单位的, 同一个随机试验用 kg 作单位与用 g 作单位计算协方差, 其结果在数值上相差 10^6 倍, 但这时不能说改变了线性关系。

假设我们对随机变量 X 与 Y 进行标准化, 即

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}}, \quad Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}},$$

得到随机变量 X^* 与 Y^* , 使得 $E(X^*) = E(Y^*) = 0$, $D(X^*) = D(Y^*) = 1$, 且 X^* 与 Y^* 没有单位。这样计算出 X^* 与 Y^* 的协方差就不会因单位的改变而改变, 而且 X 与 Y 之间的线性关系也不会因标准化而改变。因此有

$$\begin{aligned} \text{cov}(X^*, Y^*) &= E(X^* Y^*) - E(X^*) E(Y^*) = E(X^* Y^*) \\ &= E\left[\frac{X - E(X)}{\sqrt{D(X)}} \cdot \frac{Y - E(Y)}{\sqrt{D(Y)}}\right] = \frac{E[(X - E(X))(Y - E(Y))]}{\sqrt{D(X)} \sqrt{D(Y)}} \\ &= \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}}. \end{aligned}$$

定义 5 [7][8] 设 (X, Y) 为二维随机变量, 协方差 $\text{cov}(X, Y)$ 存在, 且 $D(X) > 0$, $D(Y) > 0$, 则称数值 $\frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}}$ 为随机变量 X 与 Y 的**相关系数**, 记作 ρ_{XY} , 即

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}}.$$

定理 1 [1]-[8] 设 $D(X) > 0$, $D(Y) > 0$, ρ_{XY} 为 X 与 Y 的相关系数, 则

- 1) 若 X 与 Y 相互独立, 则 $\rho_{XY} = 0$;
- 2) $|\rho_{XY}| \leq 1$, 即 $[\text{cov}(X, Y)]^2 \leq D(X)D(Y)$;
- 3) $|\rho_{XY}| = 1$ 的充分必要条件是 X 与 Y 以概率 1 线性相关。即存在常数 a, b 使得 $P\{Y = aX + b\} = 1$ 。
证略。

注 由此定理及前面对协方差刻画线性关系的讨论可知: 相关系数 ρ_{XY} 同样地是反映随机变量 X 与 Y 之间线性关系强弱的指标, 比协方差更方便, 所以确切地说, 相关系数应该称为线性相关系数。

a) $|\rho_{XY}|$ 愈接近 1, 使得 $|\text{cov}(X, Y)|$ 愈接近取得最大, 则 X 与 Y 之间愈接近线性关系;

b) $|\rho_{XY}| = 1$ 时, $|\text{cov}(X, Y)|$ 取得最大, 则 X 与 Y 之间依概率 1 线性相关;

c) $|\rho_{XY}|$ 愈接近 0, $|\text{cov}(X, Y)|$ 也愈接近 0, 则 X 与 Y 之间的线性关系就愈弱;

d) $\rho_{XY} = 0$ 时, 即 $\text{cov}(X, Y) = 0$, X 与 Y 不存在线性关系, 则称 X 与 Y **不相关**。由此可知, 当 X 与 Y 相互独立, 即什么关系都没有, 那当然线性关系也没有, 所以 X 与 Y 不相关; 反之, 若 X 与 Y 不相关, 仅仅是不存在线性关系, 但可能存在其他的非线性关系, 得不出 X 与 Y 相互独立。更确切地说, X 与 Y 不相关应该称作线性不相关。

6. 应用

例 1. 据统计, 一位 40 岁的健康者, 在 5 年内活着或自杀的概率为 p , 保险公司开办 5 年人寿保险, 每人保费为 a 元, 若 5 年内非自杀死亡, 公司赔偿 b 元, b 应如何定才能使公司期望获益; 若有 m 人参保, 公司可期望获益多少?

解: 设 X_i 表示公司从第 i 个参保者身上获得的收益, 则

X_i	a	$a - b$
p	p	$1 - p$

公司期望获益 $EX_i = ap + (a - b)(1 - p) > 0$, 即 $a < b < a(1 - p)^{-1}$ 。

m 个人参保, 获益

$$X = \sum_{i=1}^m X_i, \quad EX = \sum_{i=1}^m EX_i = ma - mb(1 - p)。$$

例 2. 某公司生产的机器无故障工作时间 X 有密度函数(单位: 万小时)

$$f(x) = \begin{cases} \frac{1}{x^2}, & x \geq 1; \\ 0, & \text{其他。} \end{cases}$$

公司每售出一台机器可获利 1600 元, 若机器在售出 1.2 万小时之内出现故障, 则予以更换, 这时每台亏损 1200 元; 若在 1.2 到 2 万小时之内出现故障, 则予以维修, 由公司负担维修费 400 元; 若在使用 2 万小时以上出现故障, 则用户自己负责。求该公司售出每台机器的平均获利。

分析: 这里涉及到两个随机变量: 无故障工作时间 X , 售出每台机器获得的利润 Y 。所求为 EY , X 的分布已知, Y 与 X 存在一定的函数关系已知。

解决方法: 求随机变量函数的数学期望, 关键是建立 Y 与 X 的函数关系。

解: 由题意可知

$$Y = g(X) = \begin{cases} -1200, & 0 \leq X \leq 1.2; \\ 1200, & 1.2 < X < 2; \\ 1600, & X \geq 2. \end{cases}$$

则

$$\begin{aligned} EY &= E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx \\ &= \int_0^{1.2} (-1200)\frac{1}{x^2}dx + \int_{1.2}^2 1200\frac{1}{x^2}dx + \int_2^{+\infty} 1600\frac{1}{x^2}dx \\ &= 1000. \end{aligned}$$

例 3. 某人有一笔资金，可投入两个项目：房产和商业，其收益都与市场有关。若把未来市场划分为好、中、差三个等级，其发生的概率分别是 0.2, 0.7 和 0.1 通过调查，该投资者认为投资房产的收益 X (万元) 和投资商业的收益 Y (万元) 的分布分别为：

X	11	3	-3	Y	6	4	-1
P	0.2	0.7	0.1	P	0.2	0.7	0.1

请问：该投资者如何投资为好？

解：计算可得 $EX = 4.0$ ， $EY = 3.9$ ， $DX = 15.4$ ， $DY = 3.29$ ，可知投资房产与商业的收益接近，但投资房产的稳定性，也就是风险是投资商业的将近 5 倍，相比较而言，投资商业为好。

注：如果 EX ， EY 差异较大，可以选择组合投资。

7. 总结

本文从随机变量数字特征的具体意义出发，说明其与生活相通，避免遇到实际问题无从下手。明白了知道数学期望概念，为何还得学习方差；了解协方差，又必须有相关系数。例如，做投资组合优化，就是确定一组投资项目的最优投资比例(或者各项目的最优投资额)，在该投资组合的总回报率的方差不超过某个可接受的值的约束下(即在可接受的风险水平下)，使得总回报率的期望值最大(即投资回报最大)；或者在投资组合的总回报率的期望值不低于某个所要求的值的约束下(即在所要求的投资回报水平下)，使得总回报率的方差最小(即投资风险最小)。这里所说的“最优”，可以是指在一定期望投资回报水平下使得风险最小，或者是指在一定风险水平下使得投资回报最大。在 20 世纪 50 年代，Harry Markowitz 研究了一定期望投资回报水平下使得方差最小的最优投资比例问题，Harry Markowitz 在该问题上取得的研究成果以及关于投资的其他研究成果，使他荣获 1991 年诺贝尔经济奖。当然，我们本还可以举出一些实例来凸显理解概念的重要性，不过这些在教材中已经得到了体现，也就不再重复。

基金项目

铜仁学院博士科研启动基金项目(trxyDH2220)；铜市科研[2023]42 号。

参考文献

[1] 韩旭里, 谢永钦, 等. 概率论与数理统计(修订版) [M]. 上海: 复旦大学出版社, 2009.
[2] 邢家省, 马健. 概率统计教程[M]. 北京: 机械工业出版社, 2015.
[3] 曹飞龙. 概率论与数理统计[M]. 北京: 高等教育出版社, 2012.
[4] 王红蔚. 概率论与数理统计[M]. 郑州: 郑州大学出版社, 2015.
[5] 李博纳, 赵志文. 概率论与数理统计[M]. 北京: 北京大学出版社, 2017.
[6] 耿素云, 张立昂. 概率统计[M]. 北京: 北京大学出版社, 2016.
[7] 谢国瑞, 等. 概率论与数理统计[M]. 北京: 高等教育出版社, 2002.
[8] 陈希孺. 概率论与数理统计[M]. 合肥: 中国科技大学出版社, 2009.