21世纪保险精算系列教材

精算师考试用书

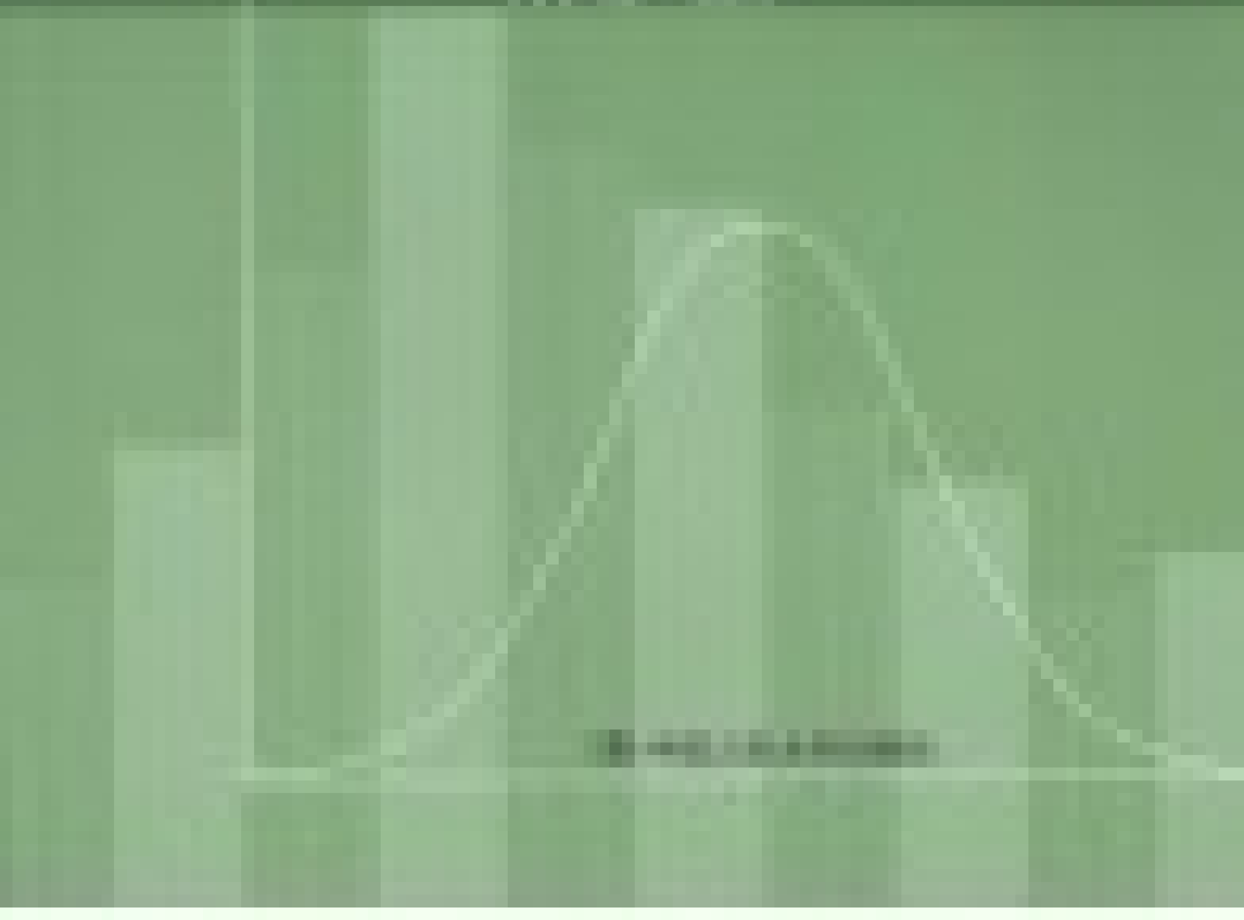
中国人民大学风险管理与精算中心主编

风险理论

肖争艳 编著

中国人民大学出版社

风险理论



21 世纪保险精算系列教材

精算师考试用书

中国人民大学风险管理与精算中心主编

风险理论

肖争艳　编著

中国人民大学出版社

图书在版编目（CIP）数据

风险理论/肖争艳编著.
北京：中国人民大学出版社，2008
（21世纪保险精算系列教材）
精算师考试用书
ISBN 978-7-300-09344-4

Ⅰ.风…
Ⅱ.肖…
Ⅲ.保险学-风险理论-教材
Ⅳ.F840

中国版本图书馆CIP数据核字（2008）第069590号

21世纪保险精算系列教材
精算师考试用书
中国人民大学风险管理与精算中心主编
风险理论
肖争艳 编著

出版发行	中国人民大学出版社		
社　　址	北京中关村大街31号	**邮政编码**	100080
电　　话	010－62511242（总编室）		010－62511398（质管部）
	010－82501766（邮购部）		010－62514148（门市部）
	010－62515195（发行公司）		010－62515275（盗版举报）
网　　址	http://www.crup.com.cn		
	http://www.ttrnet.com(人大教研网)		
经　　销	新华书店		
印　　刷	北京雅艺彩印有限公司		
规　　格	170 mm×228 mm　16开本	**版　　次**	2008年5月第1版
印　　张	18	**印　　次**	2008年5月第1次印刷
字　　数	317 000	**定　　价**	32.00元

总　序

从1775年英国公平人寿最早将运用数学工具为产品定价的专门人员命名为精算师以来，精算师职业在国际上已有200多年的发展历史。这一职业最早在人寿和养老金业务中发挥作用，之后逐步向非寿险、健康保险、社会保障等领域扩展。20世纪以后，精算师的职业进一步延伸到银行、投资、公司财务、金融工程等领域。精算师职业领域的扩展与精算职业组织的发展和精算教育水平的提高密切相关。1848年后欧美一些国家陆续成立的精算师协会以及国际精算师协会，为提高全球精算教育标准做出了贡献。例如，国际精算师协会早在1998年就公布了初级精算教育标准，要求2005年后加入国际精算师协会的成员在精算教育标准上符合国际教育标准。2007年，国际精算师协会再次公布了重新修订的初级精算教育标准及教育大纲。国际上著名的精算师职业组织，包括北美寿险精算师协会、北美非寿险精算师协会、英国精算师协会等，也从2000年后陆续对其精算教育标准和精算师考试体系进行改革，强调精算学与统计学、金融学、投资学、会计学、经济学等学科的融合，强调精算学科培养复合型风险管理人才的目标。

我国精算教育和精算师职业发展起步较晚，1992年后才陆续引入北美寿险精算师考试、英国精算师考试、日本精算师考试、北美非寿险精算师考试等，2000年后，中国精算师考试体系逐步建立起来。目前，中国精算师考试的考点已增加到15个。2006年12月，民政部批准中国精算师协会正式筹备成立。中国精算师协会的成立，必将进一步推动中国精算教育和精算师职业的发展，也迫切要求对当前的精算教育体系和精算师考试体系进行必要的改革，以尽快向国际精算师协会发布的精算教育标准看齐。

中国人民大学统计学院是国内较早开展风险管理与精算教育的大学之一。1992年统计学院就开始招收风险管理与精算专业方向的硕士研究生，1993年开始招收该方向的本科生，1996年招收了该专业方向的第一批博士研究生。2004年，经教育部批准备案，统计学院设立了独立的风险管理与精算学硕士学位点和博士学位点，标志着在风险管理与精算人才培养上，形成了学士、硕士、博士多层次、专业化的人才培养教育体系。其专业课程设置完全与国际接轨，涵盖了北美、英国和中国精算师初级课程考试的基本内容，教学大纲紧跟国际精算师协会公布的精算教育指南，同时根据学科发展的国际趋势，每年重新修订课程和教学大纲。在研究方面，设立了中国人民大学风险管理与精算中心。多年来，在寿险风险管理与精算、非寿险特别是汽车保险风险管理与精算、养老金、社会保障等领域取得了很多有影响的成果，进一步促进了风险管理与精算教育的发展。为适应我国精算教育改革与发展的需要，并体现与国际精算师协会的精算教育标准接轨，中国人民大学风险管理与精算中心精心组织编写了一套精算学系列教材，分两个阶段完成。第一阶段涵盖精算师考试初级课程的全部专业课内容，包括《金融数学》、《风险理论》、《寿险精算学》、《非寿险精算学》、《精算中常用的统计模型》5本教材，其中每本教材包含大量的练习题和解答，共10本。第二阶段涵盖精算师考试高级课程的全部内容，分寿险、非寿险、养老金、健康保险、社会保障、投资等不同系列。这套教材一方面可以满足各高校精算专业的教学需求，另一方面也可以作为参加各类精算师资格考试学员的学习参考资料，同时，也可以作为对精算学科有兴趣的同仁了解和学习精算的参考书。

这套教材的特点，一是在内容上涵盖了北美寿险、北美非寿险、英国、中国精算师考试最新的内容，同时紧跟国际精算师协会提出的精算教育标准，涵盖了国际精算教育大纲的基本内容；二是为了便于读者自学和教师讲授，我们为第一部每本教材编写了学习辅导用书，辅导书中包括学习要点、教材习题解答和一部分补充练习题及其解答等；三是在写法上，力求把精算学的数理理论与实务结合起来，注意精算数学背后的实践意义，努力从实际意义上解释各种数学关系。

本套教材凝结了中国人民大学风险管理与精算中心全体教师的心血，特别是王晓军、孟生旺、黄向阳、王燕、肖争艳、肖宇谷等老师，他们为本套教材的编写付出了极大的艰辛，统计学院部分硕士研究生和本科生对辅导用书中的习题解答和答案进行了验证，感谢他们为本套教材做出的贡献，同时也感谢中国人民大学出版社的编辑们为本书的出版付出的辛勤劳动。

袁卫

前 言

保险的基本功能是分散风险。因此，如何定义和测量风险是保险精算学中一个重要内容。风险理论就是使用统计学和数学等研究工具，对经济管理活动中的损失风险和经营风险进行定量的刻画，建立相关风险模型来研究风险的性质，并为现实的保险经营进行有效的风险分析和控制提供技术支持的一门学科。

本书兼顾了理论体系的完整性和与精算师考试的一致性。一方面，尽量做到理论体系具有完整性，基本涵盖风险理论的主要内容和重要模型。另一方面，作为精算师资格考试的阅读教材，本书在内容的取舍上基本与中国精算师资格考试中关于风险理论方面的考试大纲相符。此外，在编写过程中还参考了 SOA、英国和澳大利亚精算师资格考试的资料以及国内出版的部分风险理论的教材。

全书分为四部分，第一部分是预备知识，介绍本书所用到的概率统计的基本知识。第二部分是短期风险模型，讨论短期内单个保单的理赔额分布和保单组合的总理赔额分布。第三部分是长期风险模型，讨论保险公司长期内盈余的变化规律。第四部分是效用与风险决策理论，从效用角度讨论保险人和被保险人在面临风险时如何进行最优决策。

风险理论可以看作统计学和数学在保险精算学中的应用。因此，阅读本书的读者应具有统计学和高等数学的基本知识。作为一门应用性很强的学科，为了加强对风险理论的基本原理的深入理解，读者还需要具备使用 Matlab，Excel 等软件进行计算的能力。为此，本书的部分例子使用数值算法进行求解，并提供了相应的 Matlab 程序供读者参考。为了方便读者学习，本书还设计了一定数量的例

题和习题，并给出了所有习题的解答过程。此外，本书还提供了四套模拟题供读者准备精算师考试时使用。

本书在编写过程中得到了许多人的大力支持和帮助，凝结了许多人的劳动成果。感谢王晓军、孟生旺、王燕、黄向阳和肖宇谷等同仁的鼓励、支持和帮助；中国人民大学统计学院风险管理与精算专业的本科生房文晶、司亚娟、谢嫣和陈涵参与了本书的习题解答和模拟题解答的编写，硕士研究生魏晓东、杨俏、徐柳和王晓静认真阅读和校对了所有的书稿。另外，在本书初稿作为讲义的教学过程中，风险管理与精算专业2004级和2005级本科班同学提出了不少修改意见，使本书的质量得到进一步提高。在此向他们表示诚挚的谢意。

由于编者水平有限，书中的缺点错误在所难免，敬请读者批评指正。

编著者

目　　录

第1章

预备知识

一个精算模型总是试图描述未来可能发生的随机性损失。损失的随机性有三个含义：损失发生的次数是随机的；损失发生的时间是随机的；损失的大小是随机的。对一份特定的保单而言，这三种随机性至少满足一种。而在概率论里，这三种随机性都能用相应的随机变量来刻画。为了后面叙述的方便，本章将对本书经常用到的概率论基本知识进行简要的回顾。

1.1 概率空间、随机变量及分布

随机试验是概率论中的基本概念。事件可以看作随机试验的一种结果。试验的结果事先不能准确地预言，但具有三种特征：

（1）可以在相同的条件下重复进行；

（2）每次试验的结果不止一个，但预先知道试验所有可能的结果；

（3）每次试验前不能确定哪个结果会出现。

记随机试验的基本结果为 ω，称为样本点。随机试验所有可能的结果组成的集合称为试验的样本空间，记为 Ω。Ω 中的样本点也被称为基本事件，样本空间 Ω 称为必然事件，空集 ϕ 为不可能事件。Ω 的子集 A 由基本事件组成，通常称为事件。在实际问题中，人们一般不对样本空间的所有子集都感兴趣，而是关心某些事件及其发生的可能性大小。我们可用下面的概念来描述上述问题。

定义 1—1—1 设 Ω 是一个样本空间（或任意一集合），F 是 Ω 中某些子集组成集合。如果满足：

(1) $\Omega \in F$；

(2) 若 $A \in F$，则 $A^c \in F$；

(3) 若 $A_n \in F$ ($n=1, 2, \cdots$)，则 $\bigcup_{n=1}^{\infty} A_n \in F$；

则称 F 为 σ 代数，(Ω, F) 为可测空间，F 中元素称为事件。

如果 F 是一个 σ 代数，可以证明 $\phi \in F$；若 $A_n \in F$ ($n \geqslant 1$)，则 $\bigcap_{n=1}^{\infty} A_n \in F$。

定义 1—1—2 设 (Ω, F) 是可测空间，$P(\cdot)$是定义在 F 上的实值函数。如果：

(1) 任意 $A \in F$，$P(A) \geqslant 0$；

(2) $P(\Omega)=1$；

(3) 对两两不相容的事件 A_1，A_2，$\cdots$（即 $A_i \cap A_j = \phi$，$i \neq j$）

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

则称 P 是 (Ω, F) 上的概率，(Ω, F, P) 是概率空间，$P(A)$ 称为事件 A 的概率。

事件 A 的概率可理解为事件 A 发生可能性的度量。由定义可见，事件的概率具有如下性质：

(1) $P(\phi)=0$；

(2) $0 \leqslant P(A) \leqslant 1$；

(3) 若 $A \in F$，$B \in F$，且 $A \subset B$，则 $P(A) \leqslant P(B)$；

(4) 若 $A \in F$，$B \in F$，则 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$；

(5) 若 $A_n \in F$ ($n \geqslant 1$)，则 $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leqslant \sum_{i=1}^{\infty} P(A_i)$；

(6) 若 $A_n \in F$ ($n \geqslant 1$)，且 $A_n \subset A_{n+1}$ ($n \geqslant 1$)，则 $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} P(A_n)$；

(7) 若 $A_n \in F$ ($n \geqslant 1$)，且 $A_{n+1} \subset A_n$ ($n \geqslant 1$)，则 $P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} P(A_n)$。

定义 1—1—3 设(Ω, F, P) 是概率空间，X 是定义在 Ω 上的取值于实数域 R 的函数。如果对任意实数 $x \in R$，$(\omega | X(\omega) \leqslant x) \in F$，则称 $X(\omega)$ 是 F 上的随机变量。并称 $F_X(x) = P(X \leqslant x)$ 为 X 的累积分布函数，有时简称为分布函数。

可以证明分布函数 $F(x)$ 满足下列性质：

(1) $0 \leqslant F(x) \leqslant 1$；

(2) $F(x)$ 递增；

(3) $F(x)$ 右连续，即 $\lim\limits_{x\to x_0^+} F(x)=F(x_0)$；

(4) $\lim\limits_{x\to-\infty} F(x)=0$，$\lim\limits_{x\to+\infty} F(x)=1$。

随机变量可以用来描述现实中我们感兴趣但又不能预先知道的事件。如被保险人的损失次数、索赔额和保险公司的盈余额。在概率论中，常用的随机变量有两种类型：离散型随机变量和连续型随机变量。离散型随机变量常用来描述被保险人的理赔次数的分布。离散型随机变量的概率特性除了用分布函数表示外，还可以用分布律来描述（$p_k=P(N=k)$，$k=0$，1，2，…）。例如，某投保汽车险的驾驶员在一年之内发生索赔次数 N 的分布如下：

$$p_0=P(N=0)=0.8;\ p_1=P(N=1)=0.1;\ p_2=P(N=2)=0.075$$

$$p_3=P(N=3)=0.025;\ p_k=P(N=k)=0,\ k\geqslant 4$$

常见的离散分布有泊松分布、负二项分布、几何分布、二项分布等。我们将在本书附录中给出这些分布的性质。

连续型随机变量常被用来描述损失额的大小，连续型分布的概率特征还可以用分布密度函数来刻画。

定义 1—1—4　设(Ω, F, P) 是概率空间，X 是随机变量，$F_X(x)$ 是其分布函数，如果 $F'_X(x)$ 可积，则存在函数 $f(x)$ 使得

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

$f(x)$ 称为连续型随机变量 X 的分布密度函数。

在实际问题中，对于某些随机试验的结果有时需要用多个随机变量来描述。例如，为了研究保险公司在一段时期内总理赔额的性质，不仅需要知道每次理赔额的大小，而且要知道这段时期内的理赔次数。一般来说，随机向量 $X=(X_1, \cdots, X_n)$ 的分布不仅与 $X_1, \cdots, X_n$ 的分布有关，还依赖于 $X_1, \cdots, X_n$ 之间的相互关系。

定义 1—1—5　对于 n 维随机向量 $X=(X_1, \cdots, X_n)$，它的 n 维分布函数（或联合分布函数）定义为：

$$F(x_1, \cdots, x_n)=P(X_1\leqslant x_1, \cdots, X_n\leqslant x_n)$$

如果 $f(x_1, \cdots, x_n)=\dfrac{\partial^n F}{\partial x_1\cdots\partial x_n}$ 对所有 $(x_1, \cdots, x_n)\in R^n$ 都存在，则称 $X=(X_1, \cdots, X_n)$ 是连续型随机向量，函数 $f(x_1, \cdots, x_n)$ 是 $X=(X_1, \cdots, X_n)$ 联合密度函数，并且

$$F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f(t_1,\cdots,t_n)\mathrm{d}t_1\cdots\mathrm{d}t_n$$

如果 $X=(X_1, \cdots, X_n)$ 的所有可能取值是有限个 n 维向量或可列无穷个 n 维向量，则称 $X=(X_1, \cdots, X_n)$ 是离散型随机向量，其联合分布列为：

$$P(X_1=x_1, \cdots, X_n=x_n)$$

定义 1—1—6 设 $F(x_1, \cdots, x_n)$ 是 $X=(X_1, \cdots, X_n)$ 的联合分布函数，$1\leqslant k_1<k_2<\cdots<k_n\leqslant n$，则 $X_1, \cdots, X_n$ 的边际分布 $F_{k_1,\cdots,k_n}(x_{k_1}, \cdots, x_{k_n})$ 定义为

$$F_{k_1,\cdots,k_n}(x_{k_1}, \cdots, x_{k_n})$$
$$=F(\infty, \cdots, \infty, x_{k_1}, \infty, \cdots, \infty, x_{k_2}, \infty, \cdots, \infty, x_{k_n}, \infty, \cdots, \infty)$$

特别地，X_k 的边际分布 $F_{X_k}(x_k)$ 为：

$$F_{X_k}(x_k)=F(\infty, \cdots, x_k, \cdots, \infty)$$

用来描述损失大小的随机变量 X 的密度函数 $f(x)$ 通常具有下列特征：

(1) 非负性。损失额应该都是非负的，因此 $P(x\geqslant 0)=1$；

(2) 损失额应该是连续变化的，因而 $f(x)$ 是连续的。

(3) 损失额越小的保险事故发生的可能较大，损失额越大的保险事故发生的可能性较小，但不可以忽略。从直观上说，损失分布密度函数的尾部较厚（见图 1—1—1）。

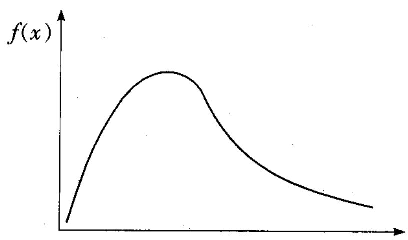


图 1—1—1 典型的损失额分布密度图

常用于描述损失分布的连续型随机变量有指数分布、伽玛分布、对数正态分布、韦伯分布、帕累托分布等。本书附录将列举一些常见分布的性质。

1.2 随机变量的数字特征

随机变量的概率分布（分布函数或密度函数）包含了随机变量的全部信息。但是在很多情况下，人们并不需要获得一个随机变量的全部信息，只需要了解它的某些特征值就可以了，如均值、最大值、最小值、与均值的离散程度、分布的对称情况。这些信息通常可以由随机变量的数字特征来反映。

定义 1—1—7 (1) 取值于 $\{x_j\}$ 的离散型随机变量 X 的 k 阶原点矩定义为：

$$\mu_K = \sum_j x_j^k P(X = x_j) = \sum_j x_j^k p_j$$

如果 $\sum_j |x_j|^k p_j < \infty$ 。

(2) 连续型随机变量 X 的 k 阶原点矩定义为：

$$\mu_k{}' = E(X^k) = \int_{-\infty}^{\infty} x^k \mathrm{d}F(x) = \int_{-\infty}^{\infty} x^k f(x)\mathrm{d}x$$

如果 $\int_{-\infty}^{\infty} |x|^k \mathrm{d}F(x) < \infty$ ，这里 $F(x)$ 为分布函数。

特别地，$\mu_1 = E(X)$ 称为 X 的数学期望，或均值，通常用 μ 来表示。μ 用来描述平均水平。在精算模型中，设每张保单实际赔付额为 X，则期望值 $E(X)$ 通常被称为纯保费，它是保费定价的基础。

定义 1—1—8 取值于 $\{x_j\}$ 的离散型随机变量 X 的 k 阶中心矩定义为：

$$\mu_k = \sum_j (x_j - \mu)^k P(X = x_j)$$

如果 $\sum_j |x_j - \mu|^k p_j < \infty$ 。

(2) 连续型随机变量 X 的 k 阶中心矩定义为：

$$\mu_k = E((X-\mu)^k) = \int_{-\infty}^{\infty} (x-\mu)^k \mathrm{d}F(x) = \int_{-\infty}^{\infty} (x-\mu)^k f(x)\mathrm{d}x$$

特别地，当 $k=2$ 时，μ_2 称为 X 的方差，记为 $\mathrm{Var}(X)$ 或者 σ^2。计算方差还有一个简单的公式：

$$\mu_2 = E(X-\mu)^2 = E(X^2) - \mu^2$$

方差的平方根 σ 称为 X 的标准差，记为 $\mathrm{Std}(X)$ 或 σ。它与随机变量 X 具有相同的量纲。

除了均值和方差外，还经常使用以下指标来描述随机变量的分布特征：

1) 变异系数 (coefficient of variation)：$r=\sigma/\mu$；

2) 偏度：$r_3 = \mu_3/\sigma^3$；

3) 峰度：$k = \mu_4/\sigma^4$ 。

方差和标准差、变异系数都可以用来描述分布的离散程度。方差和标准差以均值为中心计算分布的离散程度，它们的优点是直观、容易理解。但是，一方面，其数值的大小取决于随机变量 X 本身水平高低的影响，也就是说与 X 的均值大小有关。X 的绝对水平高的，离散程度的测度值自然也就大；绝对水平小的，离散程度的测度值自然也就小。另一方面，它们与随机变量 X 的计量单位

相同，采用不同计量单位计量的随机变量，其离散程度的测度值也不同。因此，用方差和标准差无法比较对于平均水平不同或计量单位不同的随机变量的离散程度。

变异系数则从相对的角度来观察差异和离散程度。由于变异系数等于标准差除以均值，分子和分母具有相同的量纲，因此消除了量纲的影响。在比较两个不同分布的差异程度时，变异系数比标准差更好一些。变异系数越大，说明分布的离散程度越大；变异系数越小，说明分布的离散程度越小。

偏度和峰度这两个指标用来描述分布的形状。偏度是对分布偏斜方向和程度的测度。当分布对称时，离差三次方后正负离差可以相互抵消，因此偏度$r_3=0$。当分布不对称时，正负离差不能抵消，就形成了正的或负的偏态。$r_3>0$，说明正向偏差值较大，可以判断分布函数是正偏或者是右偏的。当 $r_3<0$ 时，说明负离差数值较大，可判断为负偏或者左偏。r_3 的绝对值越大，说明偏斜程度越大。

峰度是对分布密度为平峰或尖峰程度的测度。峰度的值常与标准正态分布相比较而言。可以证明，如果随机变量服从标准正态分布，则峰度值为 3。若 $k>3$ 说明分布比正态分布更尖，为尖峰分布；当 $k<3$ 时说明分布比正态分布更平，为平峰分布。

对于多维随机变量（X，Y），除了讨论各随机变量的数学期望和方差以外，还需要讨论随机变量之间的相互关系的数字特征。

定义 1—1—9 称

$$E(X-E(X))E(Y-E(Y))$$

为随机变量 X 和 Y 的协方差，记为 $\mathrm{Cov}(X, Y)$。称

$$\rho_{XY}=\frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\ \mathrm{Var}(Y)}}$$

为随机变量 X 和 Y 的相关系数。

从定义 1—1—9 可以推出协方差和相关系数具有如下简单性质：

（1） $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)-2\mathrm{Cov}(X, Y)$

（2） $\mathrm{Cov}(X, Y)=E(XY)-E(X)E(Y)$

（3） $|\rho_{XY}|\leqslant 1$

（4） $|\rho_{XY}|=1$ 的充要条件是，存在常数 a，b，使得 $P(Y=a+bX)=1$。

从性质（4）可以看出 ρ_{XY}是一个可以用来表示 X，Y 之间线性关系紧密程度的量，当$|\rho_{XY}|$较大时，通常说 X，Y 线性相关程度较好；当$|\rho_{XY}|$较小时，通常说 X，Y 线性相关程度较差；当 $\rho_{XY}=0$ 时，则称 X 和 Y 不相关。