教 案

随机变量的数字特征

教学内容

随机变量的分布函数全面地反映了随机变量的统计规律，利用分布函数可以很方便地计算各种事件的概率。但在实际应用中，常常并不需要全面了解随机变量的变化情况，只需要知道一些能反映随机变量的特征的指标就能解决问题，这些指标便是数字特征。本节主要讲解以下内容：

（1） 数学期望、方差、协方差和相关系数等概念和计算方法；
（2） 随机变量的函数的数学期望的计算方法；
（3） 一些常见分布的数字特征的计算。

教学思路与要求

（1） 结合实际背景引出数学期望的概念，指出数学期望的性质，并结合实际例子讲解其计算方法，并进一步给出随机变量的函数的数学期望的计算方法；
（2） 结合实际背景引出方差与标准差的概念，指出方差的性质，并结合实际例子讲解其计算方法；
（3） 对几种常见分布计算它们的数学期望和方差；
（4） 结合实际背景引出协方差与相关系数的概念，指出它们的性质，并结合实际例子讲解其计算方法。
（5） 对于正态分布的数字特征，其重要性与常见性众所周知，计算也相对复杂，更需加以详细讲解。

教学安排

一．数学期望

我们先看一个例子。检验员每天从生产线取出n件产品进行检验。记ξ为每天检验出的次品数。若检验员检查了N天，记这N天出现$0,1,\cdots,n$件次品的天数分别为$x_0,x_1,\cdots,x_n$，则$x_0+x_1+\cdots+x_n=N$，且N天出现的总次品数为

$$0\cdot x_0+1\cdot x_1+\cdots+n\cdot x_n=\sum_{k=0}^{n}kx_k\text{ 。}$$

因此N天中平均每天出现的次品数为

$$\frac{\sum_{k=0}^{n}kx_k}{N}=\sum_{k=0}^{n}k\cdot\frac{x_k}{N}\text{ 。}$$

注意$\frac{x_k}{N}$就是N天中每天出现k件次品的频率，即$\{\xi=k\}$的频率。若记p_k为每天出现k件次品的概率，即$P(\xi=k)$，则由概率的统计意义，当N充分大时，$\frac{x_k}{N}$会在p_k附近摆动（$k=0,1,\cdots,n$），所以$\sum_{k=0}^{n}k\cdot\frac{x_k}{N}$就会在$\sum_{k=0}^{n}k\cdot p_k$附近摆动。因此从

统计意义上可以认为，$\sum_{k=0}^{n} k \cdot p_k$ 就是平均每天出现的次品数。

以此为背景，我们引入下面的定义。

定义 11.5.1 设离散型随机变量ξ的可能取值为$x_1, x_2, \cdots, x_n, \cdots$，且$\xi$取相应值的概率依次为$p_1, p_2, \cdots, p_n, \cdots$。若级数$\sum_{i=1}^{\infty} x_i p_i$绝对收敛，则称该级数的和为随机变量$\xi$的**数学期望**，简称**期望**，记为$E\xi$，即

$$E\xi = \sum_{i=1}^{\infty} x_i p_i \text{。}$$

此时也称ξ的数学期望存在。若级数$\sum_{i=1}^{\infty} |x_i| p_i$发散，则称$\xi$的数学期望不存在。

由数学期望的定义知，数学期望实质上是以概率为权的加权平均值，因此也常称为**均值**。我们在定义中需要级数绝对收敛，是因为数学期望应该与对随机变量取值的人为排序无关。只有当级数是绝对收敛时，才能保证收敛级数的和与求和次序无关。

对于连续型随机变量ξ，也应有数学期望的概念。如何得到呢？先做一个近似分析。设ξ的概率密度为$\varphi(x)$（假设$\varphi(x)$连续），在实轴上插入分点

$$x_0 < x_1 < \cdots < x_n \text{，}$$

则ξ落在$[x_i, x_{i+1}]$中的概率为（记$\Delta x_i = x_{i+1} - x_i$）

$$P(\xi \in [x_i, x_{i+1}]) = \int_{x_i}^{x_{i+1}} \varphi(x) dx \approx \varphi(x_i) \Delta x_i \quad (i = 0, 1, \cdots, n-1) \text{。}$$

这时，如下分布的离散型随机变量$\tilde{\xi}$就可以看作ξ的一种近似

$\tilde{\xi}$	x_0	x_1	$\cdots$	x_n
P	$\varphi(x_0)\Delta x_0$	$\varphi(x_1)\Delta x_1$	$\cdots$	$\varphi(x_n)\Delta x_n$

其数学期望为

$$E\tilde{\xi} = \sum_{i=0}^{n} x_i \varphi(x_i) \Delta x_i \text{，}$$

它近似地可看作ξ的平均值。可以想象，当分点在实轴上越来越密时，上述和式就会以$\int_{-\infty}^{+\infty} x\varphi(x) dx$为极限。由此为背景，我们给出下面的定义。

定义 11.5.2 设ξ是连续型随机变量，其概率密度为$\varphi(x)$。若$\int_{-\infty}^{+\infty} |x| \varphi(x) dx$收敛，则称$\int_{-\infty}^{+\infty} x\varphi(x) dx$的值为随机变量$\xi$的**数学期望**，简称**期望**，记为$E\xi$，即

$$E\xi = \int_{-\infty}^{+\infty} x\varphi(x) dx \text{。}$$

此时也称ξ的数学期望存在。若$\int_{-\infty}^{+\infty} |x| \varphi(x) dx$发散，则称$\xi$的数学期望不存在。

例 11.5.1 已知一箱中有产品 100 个，其中 10 个次品，90 个正品。从中任取 5 个，求这 5 个产品中次品数的期望值。

解 设ξ为任意取出 5 个产品中的次品数，则ξ可取值 0、1、2、3、4、5。

且易计算ξ的分布为

ξ	0	1	2	3	4	5
P	$\dfrac{C_{90}^5}{C_{100}^5}$	$\dfrac{C_{10}^1C_{90}^4}{C_{100}^5}$	$\dfrac{C_{10}^2C_{90}^3}{C_{100}^5}$	$\dfrac{C_{10}^3C_{90}^2}{C_{100}^5}$	$\dfrac{C_{10}^4C_{90}^1}{C_{100}^5}$	$\dfrac{C_{10}^5}{C_{100}^5}$

因此

$$E\xi=\sum_{k=0}^{5}kP(\xi=k)=\frac{\sum\limits_{k=0}^{5}kC_{90}^{5-k}C_{10}^{k}}{C_{100}^5}=0.5\text{。}$$

例 11.5.2 已知连续型随机变量ξ的概率密度为如下形式：

$$\varphi(x)=\begin{cases}ax^k, & 0<x<1,\\ 0, & \text{其它,}\end{cases}$$

其中$k>0, a>0$。又已知$E\xi=0.75$，求k和a的值。

解 由概率密度的性质得

$$1=\int_{-\infty}^{+\infty}\varphi(x)dx=\int_0^1 ax^k dx=\frac{a}{k+1},$$

所以$a=k+1$。又由已知

$$0.75=E\xi=\int_{-\infty}^{+\infty}x\varphi(x)dx=\int_0^1 ax^{k+1}dx=\frac{a}{k+2},$$

所以又成立$a=0.75(k+2)$。解方程组

$$\begin{cases}k+1=a\\ 0.75(k+2)=a\end{cases}$$

得$k=2$， $a=3$。

可以证明随机变量的数学期望有如下性质（假设以下涉及到的数学期望均存在）：

（1）设c是常数，则$Ec=c$。

（2）设ξ是随机变量，k是常数，则$E(k\xi)=kE\xi$。

（3）若ξ,η为两个随机变量，则$E(\xi+\eta)=E\xi+E\eta$。

因此，用归纳法可以得出，若$\xi_1,\xi_2,\cdots,\xi_n$为随机变量，则

$$E\left(\sum_{i=1}^{n}\xi_i\right)=\sum_{i=1}^{n}E\xi_i\text{。}$$

（4）若ξ,η为两个随机变量，满足$\xi\le\eta$（即对于每个$x\in\Omega$，成立$\xi(x)\le\eta(x)$），则

$$E\xi\le E\eta\text{。}$$

特别地

$$|E\xi|\le E|\xi|\text{。}$$

（5） 设随机变量ξ,η相互独立，则$E(\xi\eta)=E\xi\cdot E\eta$。

因此，若n个随机变量$\xi_1,\xi_2,\cdots,\xi_n$相互独立，则

$$E\left(\prod_{i=1}^{n}\xi_i\right)=\prod_{i=1}^{n}E\xi_i\text{。}$$

例 11.5.3 假设机场送客班车每次开出时有 20 名乘客，沿途有 10 个下客站。若到站时无乘客下车，则班车不停。假设每位乘客在各车站下车的机会是等可能的，且是否下车互不影响，求每班次停车的平均数。

解 用ξ表示班车的停车数。记

$$\xi_i=\begin{cases}1, & \text{在第}i\text{个车站有乘客下车,}\\ 0, & \text{在第}i\text{个车站无乘客下车,}\end{cases}\quad i=1,2,\cdots,10\text{，}$$

则$\xi=\xi_1+\xi_2+\cdots+\xi_{10}$。由于每位乘客在各车站下车的机会是等可能的，所以每个乘客在每站下车的概率为0.1，不下车的概率为0.9。而乘客是否下车是相互独立的，20 位乘客在第i站都不下车的概率就是0.9^{20}，即$P(\xi_i=0)=0.9^{20}$，所以$P(\xi_i=1)=1-0.9^{20}$。因此

$$E(\xi_i)=0\times0.9^{20}+1\times(1-0.9^{20})=1-0.9^{20}\text{，}\quad i=1,2,\cdots,10\text{。}$$

于是每班次停车的平均数，即ξ的数学期望为

$$E(\xi)=E(\xi_1+\xi_2+\cdots+\xi_{10})=E(\xi_1)+E(\xi_2)+\cdots+E(\xi_{10})=10\times(1-0.9^{20})\approx8.78\text{。}$$

二．随机变量的函数的数学期望

对于一维随机变量的函数的数学期望，有以下的计算方法：

定理 11.5.1 设ξ是随机变量，f是一元连续函数或单调函数。

（1）若ξ是离散型随机变量，其概率函数为$P(\xi=x_i)=p_i$（$i=1,2,\cdots$），则当$\sum_{i=1}^{\infty}|f(x_i)|p_i$收敛时，随机变量$\eta=f(\xi)$的数学期望存在，且

$$E\eta=Ef(\xi)=\sum_{i=1}^{\infty}f(x_i)p_i\text{；}$$

（2）若ξ是连续型随机变量，其概率密度为$\varphi(x)$，则当$\int_{-\infty}^{+\infty}|f(x)|\varphi(x)dx$收敛时，随机变量$\eta=f(\xi)$的数学期望存在，且

$$E\eta=Ef(\xi)=\int_{-\infty}^{+\infty}f(x)\varphi(x)dx\text{。}$$

此定理的证明从略。

例 11.5.4 设随机变量ξ服从参数为 0.5 的 Poisson 分布，求$\eta=\dfrac{1}{1+\xi}$的数学期望$E\eta$。

解 因为ξ服从参数为 0.5 的 Poisson 分布，所以

$$P(\xi=k)=\frac{(0.5)^k}{k!}e^{-0.5}\text{，}\quad k=0,1,2,\cdots\text{。}$$

由定理 11.5.1 得

$$E\eta = E\left(\frac{1}{1+\xi}\right) = \sum_{k=0}^{\infty}\frac{1}{1+k}\cdot\frac{(0.5)^k}{k!}e^{-0.5} = \frac{1}{0.5}e^{-0.5}\sum_{k=0}^{\infty}\frac{(0.5)^{k+1}}{(k+1)!}$$

$$= \frac{1}{0.5}e^{-0.5}\left(\sum_{n=0}^{\infty}\frac{(0.5)^n}{n!}-1\right) = \frac{2}{\sqrt{e}}(e^{0.5}-1) = 2\left(1-\frac{1}{\sqrt{e}}\right).$$

例 11.5.5 设随机变量ξ的概率密度为

$$\varphi(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

求$\eta = e^{-2\xi}$的数学期望$E\eta$。

解 由定理 11.5.1 得

$$E\eta = E(e^{-2\xi}) = \int_{-\infty}^{+\infty} e^{-2x}\varphi(x)dx$$

$$= \int_0^{+\infty} e^{-2x}e^{-x}dx = \int_0^{+\infty} e^{-3x}dx = \frac{1}{3}。$$

三．方差和标准差

在实际问题中，仅凭随机变量的数学期望（或平均值）常常并不能完全解决问题，还要考察随机变量的取值与其数学期望的之间的离散程度。例如，考察两个射击运动员的水平，自然会看他们的平均成绩，平均成绩好的，当然水平高些。但如果两个运动员的平均成绩相差无几，就要进一步看他们的成绩稳定性，即各次射击成绩与平均成绩的离散程度，离散程度越小，成绩越稳定。抽象地说就是，对于一个随机变量ξ，我们不但要考察其数学期望$E\xi$，还要考察$\xi - E\xi$。我们称$\xi - E\xi$为随机变量ξ的**离差**。显然，离差的数学期望为 0，即$E(\xi - E\xi) = 0$。因此，考虑离差的数学期望不能解决任何问题。我们自然会想到，这是由于$\xi - E\xi$的符号变化造成的。为了消除符号变化的影响，若使用$E|\xi - E\xi|$，却带来不便于计算的困难，因此在实际应用中常使用的是$E(\xi - E\xi)^2$，它易计算、实用且有效。

定义 11.5.3 设ξ是随机变量，若$E(\xi - E\xi)^2$存在，则称它为ξ的**方差**，记为$D\xi$。即

$$D\xi = E(\xi - E\xi)^2。$$

显然$D\xi \ge 0$。由定理 11.5.1 可知，关于方差有以下的计算公式：

（1）若ξ是离散型随机变量，其分布律为$P(\xi = x_i) = p_i$（$i = 1,2,\cdots$），则

$$D\xi = \sum_{i=1}^{\infty}(x_i - E\xi)^2 p_i。$$

（2）若ξ是连续型随机变量，其概率密度为$\varphi(x)$，则

$$D\xi = \int_{-\infty}^{+\infty}(x - E\xi)^2\varphi(x)dx。$$

注意在实际应用中，$D\xi$与随机变量ξ的量纲并不一致，为了保持量纲的一致性，常考虑$D\xi$的算术平方根，它称为ξ的**均方差**或**标准差**，记为σ_ξ或σ，即

$$\sigma_\xi = \sqrt{D\xi}。$$

均方差的量纲与随机变量ξ的量纲是一致的。

可以证明随机变量的方差如下性质（假设以下涉及到的方差均存在）：

(1) 设c是常数，则$D(c)=0$；反之，若随机变量ξ满足$D\xi=0$，则$P(\xi=E\xi)=1$。

(2) 设ξ是随机变量，k是常数，则$D(k\xi)=k^2D(\xi)$。

(3) 设ξ是随机变量，c是常数，则$D(\xi+c)=D(\xi)$。

(4) 若ξ和η为相互独立的随机变量，则$D(\xi+\eta)=D\xi+D\eta$。

因此，用归纳法可以得出，若随机变量$\xi_1,\xi_2,\cdots,\xi_n$相互独立，则

$$D\left(\sum_{i=1}^{n}\xi_i\right)=\sum_{i=1}^{n}D\xi_i\text{。}$$

注意，若ξ和η为相互独立的随机变量时，它们的差的方差为

$$D(\xi-\eta)=D[\xi+(-1)\eta]=D\xi+D[(-1)\eta]=D\xi+(-1)^2D\eta=D\xi+D\eta\text{。}$$

在实际计算方差时，常常用到下面的公式：

$$D\xi=E(\xi^2)-(E\xi)^2\text{。}$$

事实上，

$$\begin{aligned}D\xi&=E(\xi-E\xi)^2=E(\xi^2-2\xi E\xi+(E\xi)^2)\\&=E(\xi^2)-2E\xi\cdot E\xi+(E\xi)^2=E(\xi^2)-(E\xi)^2\text{。}\end{aligned}$$

例 11.5.6 设随机变量ξ服从参数为1的指数分布，随机变量

$$\eta=\begin{cases}-1, & \xi<1,\\ 0, & \xi=1,\\ 1, & \xi>1,\end{cases}$$

求$E\eta$和$D\eta$。

解 因为ξ服从参数为1的指数分布，则ξ的概率密度为

$$\varphi(x)=\begin{cases}e^{-x}, & x\ge 0,\\ 0, & x<0.\end{cases}$$

所以

$$P(\eta=-1)=P(\xi<1)=\int_0^1 e^{-x}dx=1-e^{-1}\text{，}$$

$$P(\eta=0)=P(\xi=1)=0$$

$$P(\eta=1)=P(\xi>1)=1-P(\xi\le 1)=1-(1-e^{-1})=e^{-1}\text{。}$$

于是

$$E\eta=(-1)\times P(\eta=-1)+0\times P(\eta=0)+1\times P(\eta=1)=2e^{-1}-1\text{。}$$

又因为

$$E(\eta^2)=(-1)^2\times P(\eta=-1)+0^2\times P(\eta=0)+1^2\times P(\eta=1)=1\text{，}$$

所以

$$D\eta=E(\eta^2)-(E\eta)^2=1-(2e^{-1}-1)^2=4(e^{-1}-e^{-2})\text{。}$$

例 11.5.7 设随机变量ξ服从$\left[-\frac{1}{2},\frac{1}{2}\right]$上的均匀分布，函数

$$f(x)=\begin{cases}\ln x, & x>0,\\ 0, & x\le 0.\end{cases}$$

求$\eta=f(\xi)$的数学期望与方差。

解 由于ξ服从$\left[-\frac{1}{2},\frac{1}{2}\right]$上的均匀分布，所以其概率密度为

$$\varphi(x)=\begin{cases}1, & -\frac{1}{2}\le x\le\frac{1}{2},\\ 0, & \text{其他}.\end{cases}$$

因此

$$E\eta=E[f(\xi)]=\int_{-\infty}^{+\infty}f(x)\varphi(x)dx$$

$$=\int_0^{\frac{1}{2}}\ln xdx=\left(x\ln x\right)\Big|_0^{\frac{1}{2}}-\int_0^{\frac{1}{2}}1dx=-\frac{1}{2}(1+\ln 2).$$

以及

$$E(\eta^2)=E[f^2(\xi)]=\int_{-\infty}^{+\infty}f^2(x)\varphi(x)dx$$

$$=\int_0^{\frac{1}{2}}(\ln x)^2dx=[x(\ln x)^2]\Big|_0^{\frac{1}{2}}-2\int_0^{\frac{1}{2}}\ln xdx=\frac{1}{2}(\ln 2)^2+\ln 2+1.$$

因此

$$D\eta=E(\eta^2)-(E\eta)^2=\frac{1}{4}(\ln 2)^2+\frac{1}{2}\ln 2+\frac{3}{4}。$$

四．几种常见分布的数学期望和方差。

（一）0−1 分布

设随机变量ξ服从 0−1 分布，且概率函数为

$$P(\xi=k)=p^k(1-p)^{1-k}，\quad k=0,1。$$

由定义

$$E\xi=0\times(1-p)+1\times p=p。$$

且由定理 11.5.1 知

$$E(\xi^2)=0^2\times(1-p)+1^2\times p=p。$$

于是

$$D\xi=E(\xi^2)-(E\xi)^2=p-p^2=p(1-p)。$$

这说明，若ξ服从参数p的 0−1 分布，则$E\xi=p$，$D\xi=p(1-p)$。

（二）二项分布

设随机变量ξ服从参数为n, p的二项分布，即$\xi\sim B(n,p)$。

首先说明服从二项分布的随机变量ξ可以看作是n个相互独立的$0-1$分布的随机变量的和。事实上，设在某个试验中事件A发生或着不发生，且A发生的概率为p，将这个实验独立地重复n次，构成一个n重 Bernoulli 试验。随机变量ξ就可以看作这个n重 Bernoulli 试验中事件A出现的次数，因此它服从二项分布$B(n,p)$。设随机变量ξ_i（$i=1,2,\cdots,n$）为

$$\xi_i=\begin{cases}1, & \text{第}i\text{次试验}A\text{发生},\\ 0, & \text{第}i\text{次试验}A\text{不发生},\end{cases}$$

则ξ_i服从参数p的$0-1$分布，且$\xi_1,\xi_2,\cdots,\xi_n$相互独立。因此$\xi=\sum_{i=1}^{n}\xi_i$，即ξ是n个相互独立的$0-1$分布的随机变量的和。

于是

$$E\xi = E\sum_{i=1}^{n}\xi_i = \sum_{i=1}^{n}E\xi_i = \sum_{i=1}^{n}p = np\text{。}$$

由于ξ_i（$i=1,2,\cdots,n$）相互独立，所以

$$D\xi = D\sum_{i=1}^{n}\xi_i = \sum_{i=1}^{n}D\xi_i = \sum_{i=1}^{n}p(1-p) = np(1-p)\text{。}$$

这说明，若$\xi \sim B(n,p)$，则$E\xi = np$，$D\xi = np(1-p)$。

（三） Poisson分布

设随机变量ξ服从参数为λ的Poisson分布，即$\xi \sim P(\lambda)$，则ξ的概率函数是

$$P(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0,1,2,\cdots\text{。}$$

由定义得

$$E\xi = \sum_{k=0}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!} = \lambda e^{-\lambda}e^{\lambda} = \lambda\text{。}$$

且由定理11.5.1得

$$E(\xi^2) = \sum_{k=0}^{\infty}k^2\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=2}^{\infty}(k-1)k\frac{\lambda^k}{k!}e^{-\lambda} + \sum_{k=1}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda}$$

$$= \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + \lambda = \lambda^2 + \lambda\text{。}$$

因此

$$D\xi = E(\xi^2) - (E\xi)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda\text{。}$$

这说明，若$\xi \sim P(\lambda)$，则$E\xi = \lambda$，$D\xi = \lambda$。

（四）均匀分布

设随机变量ξ服从区间$[a,b]$上的均匀分布，即$\xi \sim U[a,b]$，则ξ的概率密度为

$$\varphi(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b, \\ 0, & \text{其他}. \end{cases}$$

由定义

$$E\xi = \int_{-\infty}^{+\infty}x\varphi(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{a+b}{2}\text{。}$$

因为

$$E(\xi^2) = \int_{-\infty}^{+\infty}x^2\varphi(x)dx = \int_a^b x^2\frac{1}{b-a}dx = \frac{a^2+ab+b^2}{3},$$

所以

$$D\xi = E(\xi^2) - (E\xi)^2 = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}\text{。}$$

这说明，若$\xi \sim U[a,b]$，则$E\xi = \dfrac{a+b}{2}$，$D\xi = \dfrac{(b-a)^2}{12}$。

（五）指数分布

设随机变量ξ服从参数为λ的指数分布，即$\xi \sim E(\lambda)$，则ξ的概率密度为

$$\varphi(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

由定义

$$E\xi=\int_{-\infty}^{+\infty}x\varphi(x)dx=\int_{0}^{+\infty}\lambda xe^{-\lambda x}dx=-xe^{-\lambda x}\Big|_0^{+\infty}+\int_0^{+\infty}e^{-\lambda x}dx=\frac{1}{\lambda}\text{。}$$

因为

$$E(\xi^2)=\int_{-\infty}^{+\infty}x^2\varphi(x)dx=\int_0^{+\infty}\lambda x^2e^{-\lambda x}dx$$

$$=-x^2e^{-\lambda x}\Big|_0^{+\infty}+2\int_0^{+\infty}xe^{-\lambda x}dx=2\int_0^{+\infty}xe^{-\lambda x}dx=\frac{2}{\lambda^2}.$$

所以

$$D\xi=E(\xi^2)-(E\xi)^2=\frac{2}{\lambda^2}-\left(\frac{1}{\lambda}\right)^2=\frac{1}{\lambda^2}\text{。}$$

这说明，若$\xi\sim E(\lambda)$，则$E\xi=\dfrac{1}{\lambda}$，$D\xi=\dfrac{1}{\lambda^2}$。

（六）正态分布

设随机变量ξ服从参数μ，σ^2的正态分布，即$\xi\sim N(\mu,\sigma^2)$，则ξ的概率密度为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

由定义

$$E\xi=\int_{-\infty}^{+\infty}\frac{x}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\qquad\text{（令}t=\frac{x-\mu}{\sigma}\text{）}$$

$$=\int_{-\infty}^{+\infty}\frac{\sigma t+\mu}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\sigma\int_{-\infty}^{+\infty}\frac{t}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt+\mu\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\mu\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\mu\text{。}$$

且

$$D\xi=E(\xi-E\xi)^2=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}(x-\mu)^2e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$=\frac{-\sigma^2}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}(x-\mu)de^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$=\frac{-\sigma^2}{\sqrt{2\pi}\sigma}(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}\Bigg|_{-\infty}^{+\infty}+\frac{\sigma^2}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$=0+\sigma^2=\sigma^2\text{。}$$

这说明，若$\xi\sim N(\mu,\sigma^2)$，则$E\xi=\mu$，$D\xi=\sigma^2$。

例 11.5.8 设随机变量ξ，η相互独立，且都服从正态分布$N\left(0,\frac{1}{2}\right)$，求随机变量$|\xi-\eta|$的数学期望和方差。

解 由已知

$$E\xi=E\eta=0，\quad D\xi=D\eta=\frac{1}{2}。$$

令$\zeta=\xi-\eta$，则由ξ，η的相互独立性知，

$$E\zeta=E(\xi-\eta)=E\xi-E\eta=0，\quad D\zeta=D(\xi-\eta)=D\xi+D\eta=1，$$

且由定理 11.4.3 知，ζ服从正态分布$N(0,1)$。

因此

$$E|\xi-\eta|=E|\zeta|=\int_{-\infty}^{+\infty}|x|\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

$$=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}xe^{-\frac{x^2}{2}}dx=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}e^{-\frac{x^2}{2}}d\left(\frac{x^2}{2}\right)=\sqrt{\frac{2}{\pi}}.$$

因为

$$E(|\xi-\eta|^2)=E(\zeta^2)=D\zeta+(E\zeta)^2=1，$$

所以

$$D|\xi-\eta|=E(|\xi-\eta|^2)-(E|\xi-\eta|)^2=1-\frac{2}{\pi}。$$

五．协方差与相关系数

在引入这些数字特征之前，我们先介绍关于二维随机变量的函数的数学期望的计算方法。

定理 11.5.2 设(ξ,η)是二维随机变量，f是二元连续函数。

（1）若(ξ,η)是离散型随机变量，其分布为$P(\xi=x_i,\eta=y_j)=p_{ij}$ $(i,j=1,2,\cdots)$，则当$\sum_{i,j=1}^{\infty}|f(x_i,y_j)|p_{ij}$收敛时，随机变量$\zeta=f(\xi,\eta)$的数学期望存在，且

$$E\zeta=Ef(\xi,\eta)=\sum_{i,j=1}^{\infty}f(x_i,y_j)p_{ij}；$$

（2）若(ξ,η)是连续型随机变量，其联合概率密度为$\varphi(x,y)$，则当$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|f(x,y)|\varphi(x,y)dxdy$收敛时，随机变量$\zeta=f(\xi,\eta)$的数学期望存在，且

$$E\varsigma=Ef(\xi,\eta)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\varphi(x,y)dxdy。$$

此定理的证明从略。

例 11.5.9 设$D=\left\{(x,y)\,|\,0<x<2,0<y<\frac{x}{2}\right\}$，二元连续型随机变量$(\xi,\eta)$的联合概率密度为

$$\varphi(x,y)=\begin{cases}2xy, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

求$E(\xi\eta)$。

解　由定理 11.5.2 得

$$E(\xi\eta)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xy\varphi(x,y)dydy=\iint_D xy\varphi(x,y)dxdy$$

$$=\iint_D 2x^2y^2dxdy=\int_0^2dx\int_0^{\frac{x}{2}}2x^2y^2dy=\frac{8}{9}。$$

例 11.5.10 在长度为a的线段上任取两点P和Q。求线段PQ的长度的数学期望。

解　设点P和Q的坐标分别为ξ和η，则ξ和η都服从$[0,a]$上的均匀分布。由P，Q两点的任意性可知ξ与η相互独立，因而二维随机变量(ξ,η)的联合分布密度函数为

$$\varphi(x,y)=\begin{cases}\dfrac{1}{a^2}, & 0\le x\le a, 0\le y\le a,\\ 0, & \text{其他}.\end{cases}$$

于是，线段PQ的长度$|\xi-\eta|$的数学期望为

$$E(|\xi-\eta|)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|x-y|\varphi(x,y)dxdy$$

$$=\int_0^a\int_0^a|x-y|\frac{1}{a^2}dxdy=\frac{1}{a^2}\int_0^a\left[\int_0^a|x-y|dy\right]dx$$

$$=\frac{1}{a^2}\int_0^a\left[\int_0^x(x-y)dy+\int_x^a(y-x)dy\right]dx$$

$$=\frac{1}{a^2}\int_0^a\left(x^2-ax+\frac{a^2}{2}\right)dx=\frac{a}{3}.$$

对于二维随机变量(ξ,η)来说，数学期望$E\xi$和$E\eta$分别只反映了ξ，η各自的平均值，方差$D\xi$和$D\eta$分别只反映了ξ，η各自与平均值的偏差程度。它们并没有对ξ与η之间的相互关系提供任何信息。因此，人们希望有一个数字特征能在一定程度上反映这种信息。我们知道，若ξ与η相互独立，则必有$E(\xi-E\xi)(\eta-E\eta)=0$，因此当$E(\xi-E\xi)(\eta-E\eta)\ne 0$时，$\xi$与$\eta$必不相互独立，即有一定程度的联系。这使我们引入下面的概念。

定义 11.5.4 设(ξ,η)为二维随机变量。若$(\xi-E\xi)(\eta-E\eta)$的数学期望存在，则称$E(\xi-E\xi)(\eta-E\eta)$为ξ与η的**协方差**，记作$\mathrm{Cov}(\xi,\eta)$，即

$$\mathrm{Cov}(\xi,\eta)=E(\xi-E\xi)(\eta-E\eta)。$$

此时也称ξ与η的协方差存在。

若$D\xi>0$，$D\eta>0$，且$\mathrm{Cov}(\xi,\eta)$存在，则称$\dfrac{\mathrm{Cov}(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}}$为$\xi$与$\eta$的**相关系数**，记为$\rho(\xi,\eta)$或$\rho_{\xi\eta}$，即

$$\rho(\xi,\eta)=\frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}}=\frac{E(\xi-E\xi)(\eta-E\eta)}{\sqrt{D\xi}\sqrt{D\eta}}。$$

从协方差定义可以直接证明，协方差具有下列性质（假设以下涉及到的协方差等均存在）：

（1）$\mathrm{Cov}(\xi,\xi)=D\xi$。

（2）$\mathrm{Cov}(\xi,\eta)=\mathrm{Cov}(\eta,\xi)$。

（3）若c为常数，则$\mathrm{Cov}(\xi, c)=0$。

（4）若a, b为常数，则$\mathrm{Cov}(a\xi, b\eta)=ab\mathrm{Cov}(\xi, \eta)$。

（5）$\mathrm{Cov}(\xi_1+\xi_2, \eta)=\mathrm{Cov}(\xi_1, \eta)+\mathrm{Cov}(\xi_2, \eta)$。

（6）$\mathrm{Cov}(\xi, \eta)=E(\xi\eta)-E\xi E\eta$。

例如，性质（6）的证明如下：

$$\begin{aligned}\mathrm{Cov}(\xi,\eta)&=E(\xi-E\xi)(\eta-E\eta)\\&=E\left[\xi\eta-\xi E\eta-\eta E\xi+(E\xi)\cdot(E\eta)\right]\\&=E(\xi\eta)-(E\xi)\cdot(E\eta)-(E\xi)\cdot(E\eta)+(E\xi)\cdot(E\eta)\\&=E(\xi\eta)-(E\xi)\cdot(E\eta).\end{aligned}$$

还可以证明，相关系数具有下列性质：

（1）$\rho(\xi,\eta)=\rho(\eta,\xi)$。

（2）$|\rho(\xi,\eta)|\le 1$，且$|\rho(\xi,\eta)|=1$的充分必要条件是：存在常数$a\ne 0$与常数b，使得$P(\eta=a\xi+b)=1$。

（3）若ξ与η相互独立，则$\rho(\xi,\eta)=0$。

相关系数是反映两个随机变量线性相关程度的一个数字特征。由相关系数的性质（2）知道，若$|\rho(\xi,\eta)|=1$，则ξ与η之间以概率 1 成立线性关系。进一步的研究指出，ξ与η的线性关系随着$|\rho(\xi,\eta)|$的减小而减弱。当$\rho(\xi,\eta)=0$时，称ξ与η**不相关**，也就是ξ与η不线性相关。性质（3）说明，*若ξ与η相互独立，则它们不相关*。注意，不相关性一般并不能推出独立性（见例 11.5.11）。

例 11.5.11 已知二维离散型随机变量(ξ,η)的联合概率分布如下表所示，计算ξ与η的相关系数ρ，并判断ξ与η是否相互独立？

ξ \ η	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

解 易计算关于ξ，η的边缘分布都是

ξ（或η）	-1	0	1
P	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$

因此

$$E\xi=(-1)\times\frac{3}{8}+0\times\frac{1}{4}+1\times\frac{3}{8}=0\text{；}$$

$$E\xi^2=(-1)^2\times\frac{3}{8}+0^2\times\frac{1}{4}+1^2\times\frac{3}{8}=\frac{3}{4}\text{，}$$

$$D\xi=E\xi^2-(E\xi)^2=\frac{3}{4}-0^2=\frac{3}{4}\text{。}$$

同理$E\eta=0$，$E\eta^2=\frac{3}{4}$，$D\eta=\frac{3}{4}$。由于

$$\begin{aligned}E(\xi\eta)&=(-1)\times(-1)\times\frac{1}{8}+(-1)\times0\times\frac{1}{8}+(-1)\times1\times\frac{1}{8}\\&+0\times(-1)\times\frac{1}{8}+0\times0\times0+0\times1\times\frac{1}{8}\\&+1\times(-1)\times\frac{1}{8}+1\times0\times\frac{1}{8}+1\times1\times\frac{1}{8}=0\,.\end{aligned}$$

所以

$$\mathrm{Cov}(\xi,\eta)=E(\xi\eta)-(E\xi)\cdot(E\eta)=0\text{。}$$

于是

$$\rho=\frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}}=0\text{。}$$

这说明ξ与η不相关。

因为$P(\xi=0)=\frac{1}{4}$，$P(\eta=0)=\frac{1}{4}$，所以

$$P\{\xi=0,\eta=0\}=0\neq P\{\xi=0\}P\{\eta=0\}=\frac{1}{16}\,,$$

这说明ξ与η不相互独立。

对于二维正态分布，独立性与不相关性却是等价的。事实上，若$(\xi,\eta)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则其联合概率密度为

$$\varphi(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}\,,$$

则我们已经知道$\xi\sim N(\mu_1,\sigma_1^2)$，$\eta\sim N(\mu_2,\sigma_2^2)$，因此$D\xi=\sigma_1^2$，$D\eta=\sigma_2^2$。

又因为

$$\mathrm{Cov}(\xi,\eta)=E(\xi-E\xi)(\eta-\eta)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)\varphi(x,y)dxdy$$

$$=\int_{-\infty}^{+\infty}\frac{(x-\mu_1)(y-\mu_2)}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}dxdy\text{。}$$

作变换$s=\frac{x-\mu_1}{\sigma_1}$，$t=\frac{y-\mu_2}{\sigma_2}$得

$$\mathrm{Cov}(\xi,\eta)=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}st\cdot e^{-\frac{s^2-2\rho st+t^2}{2(1-\rho^2)}}dsdt$$

$$=\frac{\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}se^{-\frac{s^2}{2}}\left[\int_{-\infty}^{+\infty}\frac{t}{\sqrt{2\pi}\sqrt{1-\rho^2}}e^{-\frac{(t-\rho s)^2}{2(1-\rho^2)}}dt\right]ds$$

$$=\frac{\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\rho s^2e^{-\frac{s^2}{2}}ds=\rho\sigma_1\sigma_2\text{。}$$

于是

$$\rho(\xi,\eta)=\frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}}=\rho\text{。}$$

这说明，若$(\xi,\eta)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则$\xi$与$\eta$的相关系数为$\rho$。因此

$$\xi\text{与}\eta\text{不相关}\Leftrightarrow\rho(\xi,\eta)=0\Leftrightarrow\rho=0\Leftrightarrow\xi\text{与}\eta\text{相互独立。}$$

这就证明了：

定理 11.5.3 若(ξ,η)服从二维正态分布，则ξ与η不相关等价于ξ与η相互独立。

例 11.5.12 设$D=\left\{(x,y)\,|\,0<x<2,\,0<y<\dfrac{x}{2}\right\}$，二元连续型随机变量$(\xi,\eta)$的联合概率密度为

$$\varphi(x,y)=\begin{cases}2xy, & (x,y)\in D,\\ 0, & \text{其他}.\end{cases}$$

求$E(\xi-2\eta)$，$\mathrm{Cov}(\xi,\eta)$和$\rho(\xi,\eta)$。

解 在例 11.4.3 中已算得(ξ,η)关于ξ的边缘概率密度为

$$\varphi_\xi(x)=\begin{cases}\dfrac{x^3}{4}, & 0<x<2,\\ 0, & \text{其他}.\end{cases}$$

关于η的边缘概率密度为

$$\varphi_\eta(y)=\begin{cases}4y(1-y^2), & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

则

$$E\xi=\int_{-\infty}^{+\infty}x\varphi_\xi(x)dx=\int_0^2\frac{x^4}{4}dx=\frac{8}{5};$$

$$E\eta=\int_{-\infty}^{+\infty}y\varphi_\eta(y)dy=\int_0^1 4y^2(1-y^2)dy=\frac{8}{15};$$

所以

$$E(\xi-2\eta)=E\xi-2E\eta=\frac{8}{5}-2\times\frac{8}{15}=\frac{8}{15}\text{。}$$

又由例 11.5.9 知 $E(\xi\eta)=\dfrac{8}{9}$，所以

$$\mathrm{Cov}(\xi,\eta)=E(\xi\eta)-(E\xi)\cdot(E\eta)=\frac{8}{9}-\frac{8}{5}\times\frac{8}{15}=\frac{8}{225}\text{。}$$

由于

$$E\xi^2=\int_{-\infty}^{+\infty}x^2\varphi_\xi(x)dx=\int_0^2\frac{x^5}{4}dx=\frac{8}{3};$$

$$E\eta^2=\int_{-\infty}^{+\infty}y^2\varphi_\eta(y)dy=\int_0^1 4y^3(1-y^2)dy=\frac{1}{3},$$

所以

$$D\xi=E\xi^2-(E\xi)^2=\frac{8}{3}-\left(\frac{8}{5}\right)^2=\frac{8}{75};$$

$$D\eta=E\eta^2-(E\eta)^2=\frac{1}{3}-\left(\frac{8}{15}\right)^2=\frac{11}{225}\text{。}$$

因此

$$\rho(\xi,\eta)=\frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}}=\frac{8/225}{\sqrt{8/75}\sqrt{11/225}}\approx 0.492\text{。}$$

设 $\zeta=(\xi,\eta)$ 为二维随机变量。如果 $E\xi$ 和 $E\eta$ 都存在，则称二维向量 $(E\xi,E\eta)$ 为 ζ 的数学期望，记为

$$E\zeta=(E\xi,E\eta)\text{。}$$

称二阶矩阵

$$\begin{pmatrix}\mathrm{Cov}(\xi,\xi) & \mathrm{Cov}(\xi,\eta)\\ \mathrm{Cov}(\eta,\xi) & \mathrm{Cov}(\eta,\eta)\end{pmatrix}\text{，即}\begin{pmatrix}\mathrm{D}\xi & \mathrm{Cov}(\xi,\eta)\\ \mathrm{Cov}(\eta,\xi) & D\eta\end{pmatrix}$$

为二维随机变量 $\zeta=(\xi,\eta)$ 的**协方差矩阵**。可以证明，协方差矩阵是一个半正定矩阵。

六．习　　题

1，2，3，4．（1），（2），5，　7，8，9，11，12，13，16，17，18，19，21，23，24。