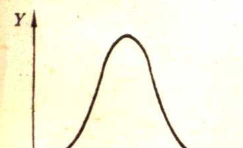
吴炯如

随机变量及其分布



福建教育出版社

随机变量及其分布

吴炯如编

福建教育出版社

内 容 提 要

本书用变量、函数的观点来研究随机现象，内容包括随机变量与分布函数、随机变量的数字特征、大数定律和中心极限定理三个部分。并配有适量的练习（附解答）。

本书可作为中学教师和具有高中程度的社会知青自学进修的读物，也可供综合大学和高等师范院校数学专业的学生阅读参考。

随机变量及其分布

吴炯如编

出版：福建教育出版社

发行：福建省新华书店

印刷：三明市印刷厂

787×1092毫米 1/32开本 3.625印张 74千字

1984年5月第一版 1984年5月第一次印刷

印数1—2,900

书号：7159·843 定价：0.38元

编者的话

如何研究随机现象的规律呢？我先将这种现象用随机变量来表达，进而用普通的实函数——随机变量的分布函数来研究它。当然有时并不需要那样全面地了解，有时要全面了解也很困难，因而就退一步只了解某些特征数字。事情往往是相辅相成的，这些数字的获得对分布律的探讨，也提供了不少的信息，在特殊的场合，甚至可以得到全部的分布律。

第三章研究概率论最本质的理论——大数定律和中心极限定理。这些理论给现实应用提供了依据。这里着重说明三个问题，（1）频率具有稳定性，使统计概率有了理论依据；（2）均值可靠性得到理论的证实；（3）大量的随机变量常常服从正态分布，因此对正态变量的研究就成为概率统计学的核心问题。

这三章是用变量、函数的观点来研究随机现象，而古典概率是一种概型一种特殊方法。

这些概率理论成为应用数学——数理统计的基础，没有这些理论，数理统计的许多方法是无法得到说明的。因此这部分虽然比较抽象，但不能缺少。尽管我们尽量避开更多的数学基础知识，只用数学分析为基础知识来研究，但由于水平的限制，不妥之处，在所难免，盼读者多多指正。

关炯如

一九八三年四月

目　录

§ 1　随机变量与分布函数……………………………（ 1 ）

1.1　离散型随机变量……………………………（ 1 ）
1.2　连续型随机变量……………………………（ 6 ）
1.3　多维随机变量及其分布……………………（ 17 ）
1.4　随机变量的函数……………………………（ 25 ）
练习一　………………………………………………（ 37 ）

§ 2　随机变量的字学特征……………………………（ 41 ）

2.1　数学期望……………………………………（ 42 ）
2.2　连续型随机变量的数学期望………………（ 49 ）
2.3　随机变量的数学期望的性质………………（ 51 ）
2.4　随机变量函数的数学期望…………………（ 54 ）
2.5　多维随机变量及多维随机变量函数的数学
期望…………………………………………（ 61 ）

§ 3　大数定律和中心极限定理………………………（ 65 ）

3.1　大数定律……………………………………（ 65 ）
3.2　中心极限定理………………………………（ 75 ）
练习二　………………………………………………（ 81 ）
练习一简解或解答……………………………………（ 82 ）
练习二简解或解答……………………………………（ 98 ）

§1 随机变量与分布函数

1.1 离散型随机变量

在随机现象中，有很多现象与数量有关．例如检验产品，关心的是抽样中出现的废品数；向车间供电，关心的是车间在某时刻正在工作的车床数；对电话，关心的是某段时间中，用户呼唤的次数和占线的时间．测量问题关心误差有多大；分子运动要关心分子运动的速度；水文站要考虑水位的高度，径流量的大小；灯泡厂要关心灯泡使用的寿命有多长；打靶者要关心每次打中的环数；水泥厂要考虑水泥的标号，是400号还是389号；跳高运动员要关心所跳的高度，等等都与数量有关，这些数量是会变化的，就叫变量．而且这些变量取各种不同值的机会有大有小，这种随机会大小而取这个值或取那个值的变量就是随机变量，粗略地说，按一定的概率规律取值的变量叫做随机变量．还有一些看起来与数量无关的随机现象，也常常可以数量化．例如掷硬币的随机试验，每次出现结果可能是正面也可能是反面，与数值好象没有关系，但可以用对应的办法让它与数量联系起来，比如说当出现正面时记为“1”，当出现反面时记为“0”，那么计算n 次投掷中出正面的次数，就等于计算“1”的个数；又如得到的信号“是脉冲的或非脉冲的”；“真与假”；

"好与坏"；"有与无"；等等都可用0与1来表记，来数量化。一般地，如果A为某个随机事件，常常可以通过如下示性函数使它与数量发生联系：

$$X_A=\begin{cases}1\text{（如果A发生），}\\0\text{（如果A不发生）.}\end{cases}$$

用数量ξ表记随机现象的每个结果。若ξ的取值个数是有限个，记作$x_1, x_2, \cdots, x_n$；或可数个，记作$x_1, x_2, \cdots, x_n, \cdots$。对ξ的每个取值（$\xi=x_i$，其中$i=1, 2, \cdots, n, \cdots$）这个事件都有确定的概率$P(\xi=x_i)$，就称ξ是离散型随机变量，常把$P(\xi=x_i)$简记为$p_i$或$P(x_i)$。今后还经常用下面的数学符号来表示：

$$\xi: \begin{pmatrix}x_1, & x_2, & \cdots, & x_n, & \cdots\\ p_1, & p_2, & \cdots, & p_n, & \cdots\end{pmatrix}\cdots\cdots\cdots\cdots\cdots\cdots(1)$$

其中事件$(\xi=x_1)$，$(\xi=x_2)$，…，$(\xi=x_n)$，…组成两两互斥的完备群；且$\sum\limits_i p_i=1$。(1)式也叫作离散型随机变量ξ的概率分布矩阵。

除了考虑$\xi=k$的概率外，经常还要考虑随机变量取若干个值的概率等于多少的问题。例如在打靶时，我们经常要考虑"至少射中一发"这一事件的概率。若设打靶发射n次，用随机变量ξ记射中的次数，"至少射中一发"这一事件就可表示为$\sum\limits_{k=1}^{n}(\xi=k)$，其概率就是$\sum\limits_{k=1}^{n}P(\xi=k)$。又如电话问题，为了考虑工作人员的工作负担量，也考虑这项工作的服务质量，就要研究呼唤次数ξ不超过5次，或超过5次的概率等于多少，也就是考虑事件$(\xi\leqslant5)$或$(\xi>5)$的概率。当然

也可以考虑事件$(5\leqslant\xi<10)$的概率等等．为了计算这些事件的概率我们只要知道对于任给的实数 x 事件$(\xi<x)$的概率，问题就都能解决．事实上，如果对任意实数 x，我们知道$P(\xi<x)$，那么

$$P(a\leqslant\xi<b)=P(\xi<b)-P(\xi<a),$$
$$P(\xi\geqslant b)=P(\xi<+\infty)-P(\xi<b).$$

由此我们引入分布函数的概念：

对于任给的实数x，令$F(x)=P(\xi<x)$，就称 $F(x)$ 为随机变量 ξ 的分布函数．有了分布函数的概念之后，计算事件概率问题，就变成研究普通的实函数问题．下面举例说明一些常见随机变量分布函数的求法．

1． 0～1分布

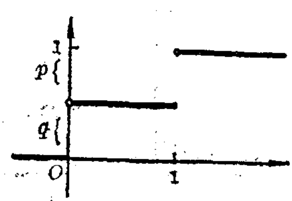
设随机变量ξ取值为0、1，取这些值的概率分别为 q，p，且$p+q=1$，写成概率分布矩阵如下：

$$\xi:\ \begin{pmatrix}0, & 1\\ q, & p\end{pmatrix},$$

这样的随机变量ξ称为服从0～1分布的随机变量．

由分布函数的定义，服从0～1分布的随机变量 ξ 的分布函数是：

$$F(x)=\begin{cases}0 & (x\leqslant 0),\\ q & (0<x\leqslant 1),\\ 1 & (x>1).\end{cases}$$



从分布函数的图象，可以看到$F(x)$具有以下性质：(1)$F(x)$是有界函数，下界为零，上界为1；(2)$F(x)$是单调不减函

数；(3)F(x)是左连续函数，在间断点所跳跃的高度，就是随机变量取该间断点值的概率。

2．二项分布

记ξ为n次贝努里试验中事件A出现的次数，每次试验中A发生的概率为p，得ξ的取值为0，1，2，…，n．则{ξ=k}，k⩽n，这一事件的概率为$C_n^k p^k q^{n-k}$ $(q=1-p)$用概率分布矩阵来表示如下：

$$\xi: \begin{pmatrix} 0, & 1, & \cdots, k, & \cdots, & n \\ q^n, & npq^{n-1}, & \cdots C_n^k p^k q^{n-k}, & \cdots, & p^n \end{pmatrix}$$

显然有：

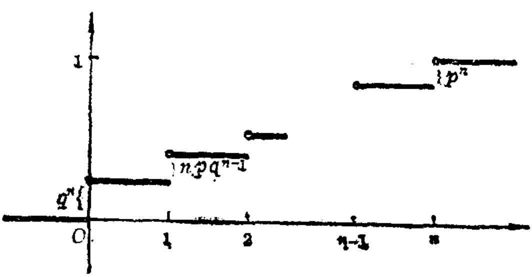
$$1=(p+q)^n=q^n+npq^{n-1}+\cdots+C_n^k p^k q^{n-k}+\cdots+p^n,$$

这样的随机变量ξ称为服从二项分布的随机变量。

由分布函数的定义，服从二项分布的随机变量ξ的分布函数是：

$$F(x)=\begin{cases} 0 & (x\leqslant 0), \\ \sum\limits_{i=0}^{k} C_n^i p^i q^{n-i} & (k<x\leqslant k+1,\ k=0,1,2,\cdots,n-1), \\ 1 & (x>n). \end{cases}$$

分布函数的图象：



由分布函数图象可以看到，它是有界函数，下界为零，上界为 1，非负不减，左连续的阶梯函数，每阶的高度就是随机变量取间断点值的概率。

3．普阿松分布：考虑热真空管中电子的发射，以ξ来表示单位时间内发射电子的个数，根据大量的实验，发现发射k个电子的概率为

$$P(\xi=k)=e^{-\lambda}\frac{\lambda^k}{k!},$$

其中λ为某正常数，ξ的概率分布矩阵为：

$$\begin{pmatrix} 0, & 1, & 2, & \cdots, & k, & \cdots \\ p_0, & p_1, & p_2, & \cdots, & p_k, & \cdots \end{pmatrix}$$

其中$p_k=e^{-\lambda}\frac{\lambda^k}{k!}$，且显然有$\sum\limits_{k=1}^{\infty}P_k=1$。这样的随机变量称为服从普阿松分布的随机变量。

由分布函数的定义，服从普阿松分布的随机变量ξ的分布函数是：

$$F(x)=\begin{cases} 0 & (x\leqslant 0), \\ \sum\limits_{i=0}^{k}e^{-\lambda}\frac{\lambda^i}{i!} & (i<x\leqslant i+1,\ i=0,1,2,\cdots,n,\cdots). \end{cases}$$

许多随机现象都可以用普阿松分布来描述：电话交换台所得呼唤数；纺纱机上的断头数；单位时间内某公共汽车站的乘客数；落到某区域的放射性物质的质点数；显微镜下落在某区域中的微生物的数目等等。因此这个分布是重要分布之一。

求离散型随机变量的分布函数，首先列出ξ的所有可能

取值和取这些值的概率，得出ξ的概率分布矩阵：

$$\begin{pmatrix} x_1, x_2, \cdots, & x_k, & \cdots, & x_n\cdots \\ p_1, p_2, \cdots, & p_k, & \cdots, & p_n\cdots \end{pmatrix}$$

然后对任意给定的x，找出使得$x_i<x$的一切x_i，最后将ξ 取这些x_i的概率值相加，即

$$F(x)=\sum_{x_i<x} P(\xi=x_i).$$

例如随机变量ξ的所有可能取值 为-1，-2，-3，…，-n，…. ξ的概率分布矩阵为：

$$\begin{pmatrix} \cdots -n, & -(n-1), & \cdots, & -3, & -2, & -1 \\ \cdots \frac{1}{2^n}, & \frac{1}{2^{n-1}}, & \cdots, & \frac{1}{8}, & \frac{1}{4}, & \frac{1}{2} \end{pmatrix}$$

这时ξ的分布函数应为

$$F(x)=\begin{cases} 1 & (x>-1), \\ \sum\limits_{i=k}^{\infty}\frac{1}{2^i} & (-k<x\leqslant -k+1,\ k\text{为任意自然数}). \end{cases}$$

1.2　连续型随机变量

一个随机变量，若可能取值于某一区间的任一点，这种随机变量不是离散型的随机变量。例如某地一月份平均温度ξ是一个随机变量，其取值范围是某一个区间；测量一个工件所得的尺寸，是一个随机变量，它的取值可能比工件准确尺寸小，也可能比工件准确尺寸大，总之取值于某一个区间，这类的随机变量取值是一个连续区间。首先就遇到一个困难，不可能把 ξ 的一切可能取值排列起来，其次这时随机变量的

任一个取值的概率该等于多少？例如测量一个工件的直径刚好是1.57828cm的概率是很小很小的，按概率的统计定义，应认为它的概率等于0．于是，对这类随机变量，就不能像离散型随机变量那样用概率分布矩阵来描述了．为了描述这种类型的随机变量，我们从数学分析中讨论不均匀金属棒的质量分布密度得到启发．

我们取不均匀金属棒的一端为坐标原点，用x表示棒上点的坐标，则(0，x)的质量m是x的函数．虽然棒上每一点的质量是0，但可以研究棒的质量分布情况，也就是可以讨论棒上某一小区间(x，x+Δx)的质量．若Δx很小时，那么这一小段的金属棒，可以近似看作均匀的，因此这一小段的平均密度是$\frac{\Delta m}{\Delta x}$，点x的密度就可以定义为

$$\lim_{\Delta x\to 0}\frac{\Delta m}{\Delta x}=p(x),$$

p(x)是x的函数．如果它是连续的（或只要求是可积的）那么要求从a到b这段棒的质量，便可用如下积分计算：

$$\int_a^b p(x)dx.$$

现在我们的问题与此类似，我们要考虑随机变量ξ的概率分布，ξ取某值x的概率为0，但可以考虑ξ在某一区间(x，x+Δx)取值的概率$P(x<\xi<x+\Delta x)$，仿照上面讨论，如果引入一个函数p(x)，叫做ξ在点x的概率密度：

$$p(x)=\lim_{\Delta x\to 0}\frac{P(x<\xi<x+\Delta x)}{\Delta x},$$

那么，要表示ξ在区间(a，b)上的概率便可由积分表示：

$$P(a<\xi<b)=\int_a^b p(x)dx.$$

由于概率非负，所以要求$p(x)\geqslant 0$。

一般地，若规定$F(x)=P(\xi<x)$，就有

$$F(x)=\int_{-\infty}^{x} p(t)dt.$$

这样分析之后，我们容易引入连续型随机变量及一般随机变量的定义如下：

连续型随机变量的定义：设随机变量ξ的分布函数为$F(x)$，若存在一个非负函数$p(x)$，且对于任何实数x，恒满足等式：

$$F(x)=\int_{-\infty}^{x} p(t)dt.$$

就称ξ是连续型随机变量，其中$p(x)$称为ξ的密度函数。

随机变量的定义：在随机现象中，其一切可能结果用变量ξ来描述，其取值个数可以有限，可以无限可数，也可以无限不可数，若对于任何实数x，事件$(\xi<x)$均有确定的概率，就称ξ是一个随机变量。用$F(x)$记$P(\xi<x)$，即$F(x)=P(\xi<x)$，就称$F(x)$是随机变量ξ的分布函数。今后为区别不同随机变量ξ、η的分布函数，经常加下标，以示区别，即记为$F_\xi(x)$，$F_\eta(x)$。

分布函数的性质

1．因为任何事件的概率都是介于0与1之间的数，所以对于任给的实数x，事件$(\xi<x)$的概率$P(\xi<x)=F(x)$也总在0与1之间：

$$0\leqslant F(x)\leqslant 1.$$

2．在x轴上任取两点x_1和x_2，并设$x_1<x_2$则：

$$P(x_1\leqslant\xi<x_2)=F(x_2)-F(x_1).$$

【证】$\because(\xi<x_1)+(x_1\leqslant\xi<x_2)=(\xi<x_2)$，

再由两个互斥事件概率运算的性质得：

$$P(x_1\leqslant\xi<x_2)=P(\xi<x_2)-P(\xi<x_1)$$
$$=F(x_2)-F(x_1).$$

3．F(x)是不减函数。

【证】：当$x_2>x_1$时由于

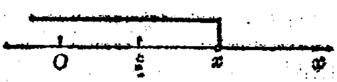
$$F(x_2)-F(x_1)=p(x_1\leqslant\xi<x_2)\geqslant 0,$$

所以 $$F(x_2)\geqslant F(x_1).$$

4． $$F(-\infty)=\lim_{x\to-\infty}F(x)=0,$$

$$F(+\infty)=\lim_{x\to+\infty}F(x)=1.$$

我们用几何直观的图解来说明这一特性，但不给以严格的证明，为此把随机变量看作ox轴上的随机点，随机试验的结果（ξ的取值）可用轴上的点来表示，这时，分布函数$F(x)=P(\xi<x)$就是在随机试验中、随机点ξ落在x左边（不包括x这一点）的概率。



为了说明$F(-\infty)=0$，我们可把点x沿x轴无限地向左移动，随机点ξ落在x左边，这一事件变成不可能事件，自然可认为这个事件的概率趋向于零，即$F(-\infty)=0$。

类似地，把点x无限地向右边移动，此时ξ落在x左边这一事件变成必然事件，因而$F(+\infty)=1$。

5．已知ξ是连续型随机变量，则

$$P(\xi=x_0)=0.$$

【证】 在x轴上，任取两点x_0和$x_0+\Delta x$，

令$\Delta x>0$，那么ξ落在$[x_0, x_0+\Delta x]$区间里的概率为：

$$P(x_0\leqslant\xi<x_0+\Delta x)=F(x_0+\Delta x)-F(x_0),$$

而$0\leqslant P(\xi=x_0)\leqslant P(x_0\leqslant\xi<x_0+\Delta x)$． $(*)$

易证当ξ为连续型随机变量时，$F(x)$是连续函数，所以当$\Delta x\to 0$时$F(x_0+\Delta x)-F(x_0)$趋于零．即上式$(*)$右端随着趋于零，这样就得：

$$P(\xi=x_0)=0.$$

由此可见，当计算连续型随机变量落在某一区间内的概率时，不论区间端点是开或闭，或半开半闭，其概率均相等，即

$$\begin{aligned}P(x_1<\xi<x_2)&=P(x_1\leqslant\xi\leqslant x_2)\\&=P(x_1\leqslant\xi<x_2)\\&=P(x_1<\xi\leqslant x_2).\end{aligned}$$

〔例1〕均匀分布：设连续型随机变量ξ的一切可能取值充满某一有限区间$[a, b]$，并且在该区间内任一点有相同的分布密度，即分布密度$p(x)$在区间$[a, b]$上为常量，这种分布叫做均匀分布。

试求在$[a, b]$上服从均匀分布的随机变量ξ的分布密度函数和分布函数：

设密度函数：$p(x)=\begin{cases}c\ (a\leqslant x\leqslant b),\\0\ (其它)。\end{cases}$

$$\because\quad \int_{-\infty}^{+\infty}p(x)dx=\int_a^b c\,dx=c(b-a)=1,$$

$$\therefore\qquad c=\frac{1}{b-a}.$$

得分布密度函数是：

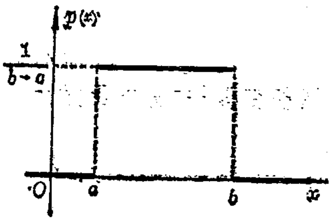
$$p(x)=\begin{cases}0 & (x\overline{\in}[a,b]),\\ \dfrac{1}{b-a} & (x\in[a,b]).\end{cases}$$

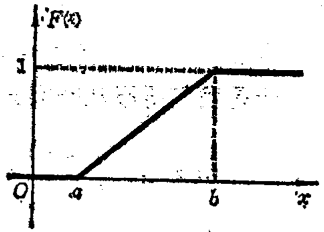
分布函数是：

$$F(x)=\begin{cases}0 & (x\leqslant a),\\ \int_{-\infty}^{a}0dt+\int_a^x\frac{1}{b-a}dx & (a<x\leqslant b),\\ 1 & (x>b).\end{cases}$$

$$=\begin{cases}0 & (x\leqslant a),\\ \dfrac{x-a}{b-a} & (a<x\leqslant b),\\ 1 & (x>b).\end{cases}$$

密度函数p(x)，分布函数F(x)的图象如下：





最后我们计算服从均匀分布的随机变量ξ落在区间〔a_1，b_1〕内的概率，这里〔a_1，b_1，〕是分布区间〔a，b〕的一部分，我们有

$$P(a_1\leqslant\xi\leqslant b_1)=\int_{a_1}^{b_1}\frac{1}{b-a}dx=\frac{b_1-a_1}{b-a}.$$

由此可见，所求概率只与区间〔a_1，b_1〕的长度b_1-a_1成正比与区间的位置无关。这与前面的几何概率相同，有时也称它为等概率分布。

均匀分布的实例：某公共汽车站每隔5分钟有一辆汽车通过。一位乘客对于汽车通过该站的时间完全不知道，他在任一时刻到达是等可能的，那么他的候车时间ξ是一个随机变量，且ξ是在区间〔0，5〕上的均匀分布的随机变量。

〔例2〕：指数分布：设连续随机变量ξ的分布函数为：

$$F(x)=\begin{cases}A+Be^{-\lambda x}(当x\geqslant 0),\\0\qquad(当x<0).\end{cases}$$

其中λ是常数且λ>0，则称ξ服从指数分布．求常数A、B的值及其分布密度。

【解】 ∵ $F(+\infty)=A=1$，

又 ∵ $F(0)=A+B=0$，

∴ $B=-1$。

于是我们得服从指数分布的随机变量ξ的分布函数为：

$$F(x)=\begin{cases}1-e^{-\lambda x}(x\geqslant 0),\\0\qquad(x<0).\end{cases}$$

分布密度为