概率分布

知识梳理

一般地，对于随机试验样本空间 Ω 中的每个样本点 ω，都有唯一的实数 $X(\omega)$ 与之对应，则称 X 为**随机变量**(random variable). 通常用大写英文字母 X, Y, Z(或小写希腊字母 ξ, η, ζ)等表示随机变量，而用小写英文字母 x, y, z(加上适当下标)等表示随机变量的取值.

植树成活的树苗数、抛掷骰子向上的点数……像这种取值为离散的数值的随机变量称为**离散型随机变量**(discrete random variable).而接听电话的时长、降雨量……取值为连续的实数区间,具有这种特点的随机变量称为**连续型随机变量**(continuous random variable).

一般地，随机变量 X 有 n 个不同的取值，它们分别是 x_1，x_2，…，x_n，且

$$P(X=x_i)=p_i,\ i=1,\ 2,\ \cdots,\ n, \tag{①}$$

称①为随机变量 X 的**概率分布列**，简称为 X 的分布列. ①也可以用表 8-2-2 的形式来表示.

表 8-2-2

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

我们将表 8-2-2 称为随机变量 X 的**概率分布表**. 它和①都叫作随机变量 X 的**概率分布**.

显然，这里的 $p_i(i=1,\ 2,\ \cdots,\ n)$满足条件

$$p_i\geqslant 0,$$

$$p_1+p_2+\cdots+p_n=1.$$

在例 3 中,随机变量 X 只取两个可能值 0 和 1. 像这样的例子还有很多,例如:在射击中,只考虑“命中”与“不命中”;对产品进行检验时,只关心“合格”与“不合格”. 我们把这一类概率分布称为 **0－1 分布**或**两点分布**,并记为 $X\sim 0-1$ 分布或 $X\sim$ 两点分布. 此处“$\sim$”表示“服从”.

一般地，随机变量 X 的概率分布如表 8 - 2 - 8 所示，

表 8 - 2 - 8

X	x_1	x_2	$\cdots$	x_n
概率 p	p_1	p_2	$\cdots$	p_n

其中 $p_i \geqslant 0$，$i = 1, 2, \cdots, n$，$p_1 + p_2 + \cdots + p_n = 1$，我们将

$$p_1 x_1 + p_2 x_2 + \cdots + p_n x_n$$

称为随机变量 X 的**均值**(mean)或**数学期望**(mathematical expectation)，记为 $E(X)$ 或 μ.

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

其中，$p_i \geqslant 0, i=1, 2, \cdots, n$，$p_1+p_2+\cdots+p_n=1$，则 $(x_i-\mu)^2$ ($\mu=E(X)$) 描述了 x_i ($i=1, 2, \cdots, n$) 相对于均值 μ 的偏离程度，故

$$(x_1-\mu)^2 p_1+(x_2-\mu)^2 p_2+\cdots+(x_n-\mu)^2 p_n$$

（其中 $p_i \geqslant 0$，$i=1, 2, \cdots, n$，$p_1+p_2+\cdots+p_n=1$）刻画了随机变量 X 与其均值 μ 的平均偏离程度，我们将其称为离散型随机变量 X 的**方差**，记为 $D(X)$ 或 σ^2. 即

$$D(X)=\sigma^2=(x_1-\mu)^2 p_1+(x_2-\mu)^2 p_2+\cdots+(x_n-\mu)^2 p_n.$$

方差也可用公式 $D(X)=\sum_{i=1}^{n}x_i^2p_i-\mu^2$ 计算，有兴趣的同学可以尝试证明.

随机变量 X 的方差也称为 X 的概率分布的方差，X 的方差 $D(X)$ 的算术平方根称为 X 的**标准差**，即

若随机变量 X 的分布列为

$$P(X=k)=C_n^k p^k q^{n-k},$$

其中 $0<p<1$, $p+q=1$, $k=0, 1, 2, \cdots, n$,则称 X 服从参数为 n, p 的**二项分布**(binomial distribution),记作 $X \sim B(n, p)$. 其概率分布如表 8 - 2 - 19 所示.

表 8 - 2 - 19

X	0	1	2	$\cdots$	n
P	$C_n^0 p^0 q^n$	$C_n^1 p q^{n-1}$	$C_n^2 p^2 q^{n-2}$	$\cdots$	$C_n^n p^n q^0$

一般地，当 $X \sim B(n, p)$ 时，

$$E(X) = np,$$

$$D(X) = np(1-p),$$

$$\sigma = \sqrt{np(1-p)}.$$

对一般情形，一批产品共 N 件，其中有 M 件不合格品，随机取出的 n 件产品中，不合格品数 X 的概率分布如表 8-2-22 所示.

表 8-2-22

X	0	1	2	$\cdots$	l
P	$\frac{C_M^0 C_{N-M}^n}{C_N^n}$	$\frac{C_M^1 C_{N-M}^{n-1}}{C_N^n}$	$\frac{C_M^2 C_{N-M}^{n-2}}{C_N^n}$	$\cdots$	$\frac{C_M^l C_{N-M}^{n-l}}{C_N^n}$

其中 $l=\min\{n, M\}$.

一般地，若一个随机变量 X 的分布列为

$$P(X=r)=\frac{C_M^r C_{N-M}^{n-r}}{C_N^n}, \qquad ①$$

其中 $r=0, 1, 2, 3, \cdots, l, l=\min\{n, M\}$，则称 X 服从**超几何分布**(hypergeometric distribution)，记为 $X \sim H(n, M, N)$，并将 $P(X=r)=\frac{C_M^r C_{N-M}^{n-r}}{C_N^n}$ 记为 $H(r; n, M, N)$.

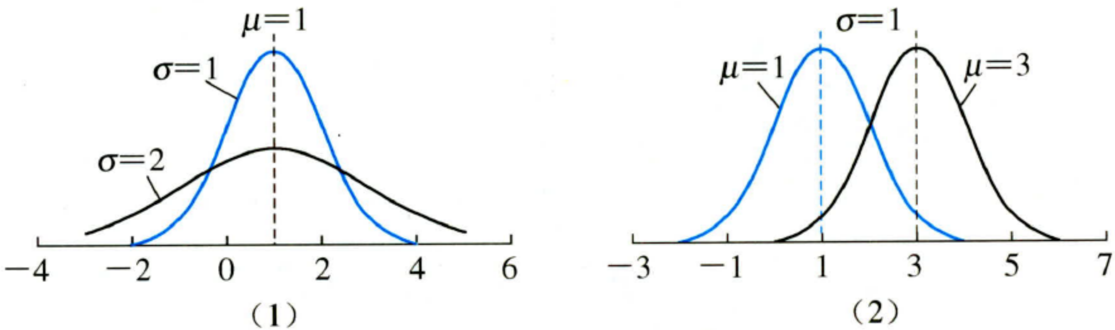
一般地，当 $X \sim H(n, M, N)$ 时，

$$E(X)=\sum_{k=0}^{l} kP_k=\frac{nM}{N},$$

其中 $l=\min\{n, M\}$.

函数 $P(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$（$x\in\mathbf{R}$）的图象与上述曲线非常吻合，我们将该函数的图象称为**正态密度曲线**. 这里有两个参数 μ 和 σ，其中 $\sigma>0$，$\mu\in\mathbf{R}$.

如图 8-3-3，不同的 μ 和 σ 对应着不同的正态密度曲线.



正态密度曲线具有如下特征：

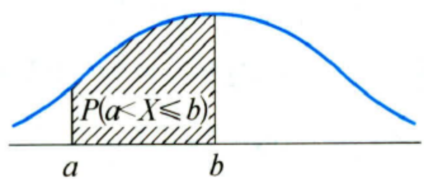
(1) 当 $x<\mu$ 时，曲线上升；当 $x>\mu$ 时，曲线下降；当曲线向左右两边无限延伸时，以 x 轴为渐近线.

(2) 曲线关于直线 $x=\mu$ 对称.

(3) σ 越大，曲线越扁平；σ 越小，曲线越尖陡.

(4) 在曲线下方和 x 轴上方范围内的区域面积为 1.

设 X 是一个随机变量，若对任给区间 $(a, b]$，$P(a<X\leqslant b)$ 是正态密度曲线下方和 x 轴上 $(a, b]$ 上方所围成的图形的面积(图 8－3－4)，则称随机变量 X 服从参数为 μ 和 σ^2 的**正态分布**(normal distribution)，简记为 $X\sim N(\mu, \sigma^2)$.



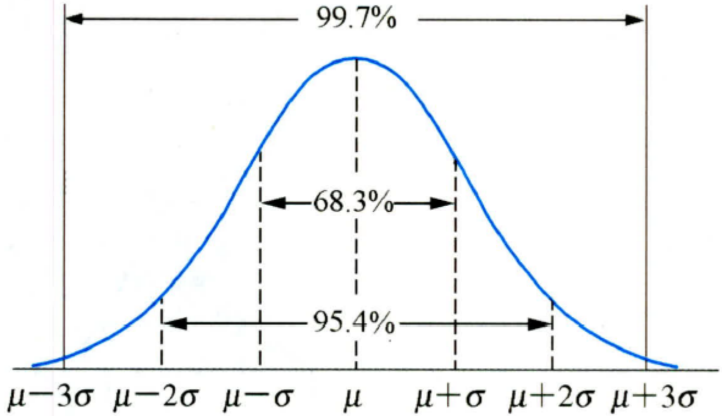


图 8-3-5

落在区间 $(\mu-\sigma, \mu+\sigma)$ 内的概率约为 68.3%；

落在区间 $(\mu-2\sigma, \mu+2\sigma)$ 内的概率约为 95.4%；

落在区间 $(\mu-3\sigma, \mu+3\sigma)$ 内的概率约为 99.7%.

事实上，μ 就是随机变量 X 的均值，σ^2 就是随机变量 X 的方差，它们分别反映 X 取值的平均大小和稳定程度.