第 2 章 随机变量及其分布

安徽财经大学

统计与应用数学学院

目录

1. 随机变量与分布函数
   - 随机变量的概念
   - 随机变量的分布函数
   - 离散型随机变量及其分布律
   - 连续型随机变量及其概率密度函数
2. 常见分布
3. 随机变量函数的分布

例 (2.1.1)

在由 5 人组成的科技活动小组中有 2 名女生和 3 名男生，从中任选两人参加全校百科知识竞赛．在这个试验中，我们用 W_1, W_2 表示女生，M_1, M_2, M_3 表示男生．为研究方便，我们将每一个试验结果 (样本点或基本事件) 用一个实数来对应．例如，将结果 $\omega_1 = (W_1, W_2)$ 对应于 2；结果 $\omega_2 = (W_1, M_1)$, $\omega_3 = (W_1, M_2)$, $\omega_4 = (W_1, M_3)$, $\omega_5 = (W_2, M_1)$, $\omega_6 = (W_2, M_2)$, $\omega_7 = (W_2, M_3)$ 均对应于 1；结果 $\omega_8 = (M_1, M_2)$, $\omega_9 = (M_1, M_3)$, $\omega_{10} = (M_2, M_3)$ 均对应于 0．不难发现，数值 $0, 1, 2$ 是 “选出的两人中女生的人数所可能取得的值的全体”．由此可见 “选取的两人中女生人数” 是一个变量．

例 (2.1.1)

如果我们用 X 来记这个变量，则 X 随着试验结果的不同而变化，即

$$X = \begin{cases} 2, & \omega = \omega_1, \\ 1, & \omega = \omega_2, \cdots, \omega_7, \\ 0, & \omega = \omega_8, \omega_9, \omega_{10}. \end{cases}$$

从函数的观点看，X 可以看成是定义在样本空间 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_{10}\}$ 上的一个函数. 通常记作 $X = X(\omega)$. 由于 ω 是随机出现的，因而 $X(\omega)$ 的取值也是随机的. 直观上讲这样的量就是随机变量.

定义 (2.1.1)

设随机试验的样本空间为 Ω, 如果对 Ω 中每一个元素 ω, 都有一个唯一确定的实数 $X(\omega)$ 与之对应, 则称 Ω 上的实值函数 $X = X(\omega)$ 是定义在 Ω 上的**随机变量**(Random variable).

定义 (2.1.1)

设随机试验的样本空间为 Ω, 如果对 Ω 中每一个元素 ω, 都有一个唯一确定的实数 $X(\omega)$ 与之对应, 则称 Ω 上的实值函数 $X = X(\omega)$ 是定义在 Ω 上的随机变量(Random variable).

引入随机变量后, 样本空间中的随机事件就可以通过随机变量表示出来. 如例题 2.1.1 中, X 表示选出来的女生数, 则事件 $\{(W_1, W_2)\}$ 可以表示为 $\{\omega \mid X(\omega) = 2\}$, 事件 $\{(M_1, M_2)\}$ 可以表示为 $\{\omega \mid X(\omega) = 0\}$. 这样通过随机变量, 我们就可以把对随机事件的研究转化为对随机变量的研究, 把求随机事件发生的概率转化为求随机变量取某些值或者在某一范围内取值的概率.

一般地我们以大写字母如 $X, Y, Z, \cdots$ 表示随机变量, 而以小写字母如 $x, y, z, \cdots$ 分别表示这些随机变量的取值.

1. 随机变量与分布函数
   - 随机变量的概念
   - 随机变量的分布函数
   - 离散型随机变量及其分布律
   - 连续型随机变量及其概率密度函数
2. 常见分布
3. 随机变量函数的分布

随机变量 X 是样本点 ω 所对应的一个实值函数，若 B 是某些实数所组成的集合，即 $B \subset \mathbb{R}$. 则 $\{X \in B\}$ 表示的就是一个随机事件:

$$\{\omega \mid X(\omega) \in B\} \subset \Omega.$$

为了研究随机变量的概率规律，只需要掌握 X 各种取值的概率. 由于

$$\{a < X \leqslant b\} = \{X \leqslant b\} - \{X \leqslant a\}, \ \{X > c\} = \Omega - \{X \leqslant c\},$$

因此对于任意实数 x, 只需要知道事件 $\{X \leqslant x\}$ 的概率就够了. 对应一个 $x \in \mathbb{R}$, 显然 $\{X \leqslant x\}$ 表示一个确定的事件，因而有唯一的概率 $P\{X \leqslant x\}$ 与之对应，这样就形成了一个函数关系，我们把这个函数称为 X 的分布函数，确切地我们有如下的定义.

定义 (2.1.2)

设 X 是随机变量，对任意的实数 x，称函数

$$F(x) = P(X \leqslant x),$$

为随机变量 X 的**分布函数**(Distribution function)，且称 X 服从分布 $F(x)$，记为 $X \sim F(x)$. 有时为了明确表示是随机变量 X 的分布函数，也把它记为 $F_X(x)$.

定义 (2.1.2)

设 X 是随机变量，对任意的实数 x，称函数

$$F(x) = P(X \leqslant x),$$

为随机变量 X 的分布函数(Distribution function)，且称 X 服从分布 $F(x)$，记为 $X \sim F(x)$. 有时为了明确表示是随机变量 X 的分布函数，也把它记为 $F_X(x)$.

显然，对于任意实数 x_1, x_2 $(x_1 < x_2)$，有

$$P(x_1 < X \leqslant x_2) = P(X \leqslant x_2) - P(X \leqslant x_1) = F(x_2) - F(x_1). \tag{1}$$

分布函数具有以下几个基本性质:

(1) $F(x)$ 为单调不减的函数.
事实上, 由 (1) 式, 对于任意实数 x_1, x_2 $(x_1 < x_2)$, 有 $\{X \leqslant x_1\} \subset \{X \leqslant x_2\}$. 由概率的单调性知:

$$F(x_1) = P(X \leqslant x_1) \leqslant F(x_2) = P(X \leqslant x_2).$$

分布函数具有以下几个基本性质:

(1) $F(x)$ 为单调不减的函数.
事实上, 由 (1) 式, 对于任意实数 x_1, x_2 $(x_1 < x_2)$, 有 $\{X \leqslant x_1\} \subset \{X \leqslant x_2\}$. 由概率的单调性知:

$$F(x_1) = P(X \leqslant x_1) \leqslant F(x_2) = P(X \leqslant x_2).$$

(2) 有界性. $0 \leqslant F(x) \leqslant 1$, 且

$$F(+\infty) = \lim_{x\to+\infty} F(x) = 1,\ F(-\infty) = \lim_{x\to-\infty} F(x) = 0.$$

分布函数具有以下几个基本性质:

(1) $F(x)$ 为单调不减的函数.
事实上, 由 (1) 式, 对于任意实数 x_1, x_2 $(x_1 < x_2)$, 有 $\{X \leqslant x_1\} \subset \{X \leqslant x_2\}$. 由概率的单调性知:

$$F(x_1) = P(X \leqslant x_1) \leqslant F(x_2) = P(X \leqslant x_2).$$

(2) 有界性. $0 \leqslant F(x) \leqslant 1$, 且

$$F(+\infty) = \lim_{x \to +\infty} F(x) = 1,\ F(-\infty) = \lim_{x \to -\infty} F(x) = 0.$$

(3) $F(x_0 + 0) = \lim\limits_{x \to x_0^+} F(x) = F(x_0)$, 即 $F(x)$ 是右连续的.

上述三个性质是随机变量的分布函数最本质的性质, 可以证明满足上述三个性质的实值函数一定可以作为某随机变量的分布函数.

概率论主要是利用随机变量来描述和研究随机现象，有了随机变量 X 的分布函数，有关 X 的各种事件的概率可以通过分布函数方便地表出，例如，对于任意的实数 a, b 有

$$
\begin{gathered}
P(a < X \leqslant b) = F(b) - F(a), \\
P(X = a) = F(a) - F(a-0), \\
P(X > b) = 1 - F(b), \\
P(X \geqslant b) = 1 - F(b-0), \\
P(a < X < b) = F(b-0) - F(a), \\
P(a \leqslant X \leqslant b) = F(b) - F(a-0), \\
P(a \leqslant X < b) = F(b-0) - F(a-0).
\end{gathered}
$$

例 (2.1.2)

验证函数

$$F(x) = \frac{1}{\pi}\left[\arctan x + \frac{\pi}{2}\right], -\infty < x < +\infty,$$

满足分布函数的三条基本性质 (因而可以作为一个分布函数, 通常把这个分布函数称为柯西分布函数).

例 (2.1.2)

验证函数

$$F(x)=\frac{1}{\pi}\left[\arctan x+\frac{\pi}{2}\right],-\infty<x<+\infty,$$

满足分布函数的三条基本性质 (因而可以作为一个分布函数, 通常把这个分布函数称为柯西分布函数).

在引进了随机变量和分布函数后我们就能利用高等数学的许多结果和方法来研究各种随机现象了, 它们是概率论的两个重要而基本的概念. 下面我们从离散型和连续型两种类别来更深入地研究随机变量及其分布函数.

定义 (2.1.3)

如果随机变量所有可能的取值为有限个或可数多个，则称这种随机变量为**离散型随机变量**.

定义 (2.1.3)

如果随机变量所有可能的取值为有限个或可数多个，则称这种随机变量为离散型随机变量.

容易知道，要掌握一个离散型随机变量 X 的统计规律，必须且只需知道 X 的所有可能取的值以及取每一个可能值的概率.
设离散型随机变量 X 所有可能的取值为 $x_i\ (i=1,2,\cdots)$，X 取各个可能值的概率为 p_i，即事件 $(X=x_i)$ 的概率

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

我们称 (2) 式为离散型随机变量 X 的分布列或分布律.

分布律也常用表格来表示.

表: 分布律

X	x_1	x_2	x_3	$\cdots$	x_i	$\cdots$
p_k	p_1	p_2	p_3	$\cdots$	p_i	$\cdots$

注意在表格形式表示分布列时, x_i 通常按从小到大的顺序排列.

注意到对于离散型随机变量，当 $i \neq j$ 时，$\{X = x_i\}$，$\{X = x_j\}$ 是互不相容的两个事件，由概率的性质容易推得，任一离散型随机变量的分布律 $\{ p_i \}$，都具有下述两个基本性质：

(1) 非负性：$p_i \geqslant 0,\ i = 1, 2, \cdots;$

(2) 正则性：$\sum\limits_{i=1}^{\infty} p_i = 1.$

反过来，能够证明：任意一个具有以上两个性质的数列 $\{ p_i \}$，一定可以作为某一个离散型随机变量的分布律.

对于离散型随机变量，容易知道分布律与分布函数通过如下关系相互唯一确定.

$$p_i = F(x_i) - F(x_i - 0),$$

$$F(x) = \sum_{x_i \leqslant x} P(X = x_i).$$

例 (2.1.3)

设离散型随机变量 X 的分布律为

$$P(X=k)=\frac{a}{N},\ k=1,2,\cdots,N.$$

试确定常数 a.

例 (2.1.3)

设离散型随机变量 X 的分布律为

$$P(X=k)=\frac{a}{N},\ k=1,2,\cdots,N.$$

试确定常数 a.

解

由分布律的正则性知

$$\sum_{k=1}^{N}\frac{a}{N}=N\cdot\frac{a}{N}=1,$$

因而 $a=1$.

例 (2.1.4)

一篮球运动员连续不断地进行投篮练习，直到第一次投中为止．假定这位运动员的投篮命中率为 $\frac{3}{4}$，试求：(1) 投篮次数 X 的分布列；(2) X 取偶数的概率．

例 (2.1.4)

一篮球运动员连续不断地进行投篮练习，直到第一次投中为止．假定这位运动员的投篮命中率为 $\frac{3}{4}$，试求：(1) 投篮次数 X 的分布列；(2) X 取偶数的概率．

解

(1) 易知 X 的可能取值为 $1, 2, 3, \cdots$，接下来确定 $P(X=k)$ 的概率．注意到 $(X=k)$ 表示前 $k-1$ 次投篮未中，第 k 次投篮命中．又因为一次投篮投中的概率为 $\frac{3}{4}$，未投中的概率为 $\frac{1}{4}$，所以

$$P(X=k) = \left(\frac{1}{4}\right)^{k-1}\left(\frac{3}{4}\right) = \frac{3}{4^k}, \quad k=1,2,\cdots,$$

即为所求的分布律．

(2) $P(X=\text{偶数}) = \sum_{k=1}^{+\infty} P(X=2k) = \sum_{k=1}^{+\infty} \frac{3}{4^{2k}} = \frac{1}{5}$.

例 (2.1.5)

设离散型随机变量 X 的分布律为

X	3	4	5
p_k	0.1	0.3	0.6

试求 X 的分布函数.

例 (2.1.5)

设离散型随机变量 X 的分布律为

X	3	4	5
p_k	0.1	0.3	0.6

试求 X 的分布函数.

解

注意到离散型随机变量的分布函数应是右连续的阶梯函数,
当 $x < 3$ 时, $F(x) = P(X \leqslant x) = P(\varnothing) = 0$;
当 $3 \leqslant x < 4$ 时, $F(x) = P(X \leqslant x) = P(X = 3) = 0.1$;
当 $4 \leqslant x < 5$ 时, $F(x) = P(X \leqslant x) = P(X = 3) + P(X = 4) = 0.4$;
当 $x \geqslant 5$ 时, $F(x) = P(X \leqslant x) = P(X = 3) + P(X = 4) + P(X = 5) = 1$.

解

于是 X 的分布函数可以表示

$$F(x)=\begin{cases}0, & x<3;\\ 0.1, & 3\leqslant x<4;\\ 0.4, & 4\leqslant x<5;\\ 1, & x\geqslant 5.\end{cases}$$

事实上，由离散型随机变量的特点可以知道离散型随机变量的分布函数应是右连续的阶梯函数，可能的取值点是 $F(x)$ 的跳跃点，跳跃高度就是其取该值的概率．所以在写离散型随机变量分布函数时，应将实数域以可能取值点分割成左闭右开区间．

离散型随机变量的特点是它的可能取值及其相应的概率能够逐一列出. 然而实际问题中许多的随机变量的取值可能取值连续地充满某个区间甚至是整条数轴, 如身高、体重、零件的尺寸、棉纤维的长度等. 这类随机变量就是所谓的连续型随机变量.

定义 (2.1.4)

设随机变量 X 的分布函数为 $F(x)$. 若存在一个非负可积函数 $f(x)$ 使得对于任意的实数 x 都有

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, \tag{3}$$

则称 X 为连续型随机变量. 称 $f(x)$ 为 X 的概率密度函数, 简称为概率密度或密度函数.

由连续型随机变量的定义可知连续型随机变量的分布函数 $F(x)$ 是连续函数. 由分布函数性质 $F(-\infty)=0$, $F(+\infty)=1$ 及单调不减性可知: $F(x)$ 是一条位于直线 $y=0$ 与 $y=1$ 之间的单调不减的连续曲线.

由定义 2.1.4 知道, $f(x)$ 具备如下性质:

(1) 非负性: $f(x) \geqslant 0$.

(2) 正则性: $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$.

(3) $P(a < X < b) = \int_a^b f(x)\mathrm{d}x$.

(4) 若 $f(x)$ 在点 x 连续, 则有 $F'(x) = f(x)$.

(5) 注意到连续型随机变量的分布函数是连续函数, 故有

$$P(X=a) = F(a) - F(a-0) = 0,$$

即连续型随机变量取任一特定值的概率为 0. 因而在计算连续型随机变量落在某一区间上的概率时, 不必区分该区间是否包含端点.

注: $P(X=a)=0$, 但 $\{X=a\}$ 不为不可能事件.

例 (2.1.6)

设随机变量 X 具有密度函数 $f(x)=\dfrac{C}{1+x^2},\quad -\infty<x<+\infty,$

求: (1) 常数 C; (2) $P(0\leqslant X<1)$.

例 (2.1.6)

设随机变量 X 具有密度函数 $f(x)=\dfrac{C}{1+x^2},\quad -\infty<x<+\infty,$

求: (1) 常数 C; (2) $P(0\leqslant X<1)$.

解

(1) 由密度函数的正则性知:

$$1=\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{C}{1+x^2}\mathrm{d}x=C\pi,$$

所以 $C=\dfrac{1}{\pi}$. 即 X 的密度函数为 $f(x)=\dfrac{1}{\pi(1+x^2)},\quad -\infty<x<+\infty.$

例 (2.1.6)

设随机变量 X 具有密度函数 $f(x)=\dfrac{C}{1+x^2},\quad -\infty<x<+\infty$,
求: (1) 常数 C; (2) $P(0\leqslant X<1)$.

解

(1) 由密度函数的正则性知:

$$1=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{C}{1+x^2}\mathrm{d}x=C\pi,$$

所以 $C=\dfrac{1}{\pi}$. 即 X 的密度函数为 $f(x)=\dfrac{1}{\pi(1+x^2)},\quad -\infty<x<+\infty$.

(2) $P(0\leqslant X<1)=\int_0^1\dfrac{1}{\pi(1+x^2)}\mathrm{d}x=\dfrac{1}{\pi}\arctan x\Big|_0^1=\dfrac{1}{\pi}\cdot\dfrac{\pi}{4}=\dfrac{1}{4}$.

例 (2.1.7)

设连续型随机变量 X 的分布函数为 $F(x)=\begin{cases}0, & x<0,\\ Ax^2, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$ 试求:

(1) 系数 A. (2) X 落在区间 $(0.3,0.7)$ 内的概率. (3) X 的密度函数.

例 (2.1.7)

设连续型随机变量 X 的分布函数为 $F(x)=\begin{cases}0, & x<0,\\ Ax^2, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$ 试求:

(1) 系数 A. (2) X 落在区间 $(0.3,0.7)$ 内的概率. (3) X 的密度函数.

解

(1) 由于 X 为连续型随机变量, 故 $F(x)$ 是连续函数, 因此有

$$1=F(1)=\lim_{x\to 1-0}F(x)=\lim_{x\to 1-0}Ax^2=A,$$

即 $A=1$, 于是有 $F(x)=\begin{cases}0, & x<0,\\ x^2, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$

例 (2.1.7)

设连续型随机变量 X 的分布函数为 $F(x)=\begin{cases}0, & x<0,\\ Ax^2, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$ 试求:

(1) 系数 A. (2) X 落在区间 $(0.3,0.7)$ 内的概率. (3) X 的密度函数.

解

(1) 由于 X 为连续型随机变量, 故 $F(x)$ 是连续函数, 因此有

$$1=F(1)=\lim_{x\to 1-0}F(x)=\lim_{x\to 1-0}Ax^2=A,$$

即 $A=1$, 于是有 $F(x)=\begin{cases}0, & x<0,\\ x^2, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$

(2) $P(0.3<X<0.7)=F(0.7)-F(0.3)=0.7^2-0.3^2=0.4.$

(3) X 的密度函数为 $f(x)=F'(x)=\begin{cases}2x, & 0\leqslant x<1,\\ 0, & \text{其他}.\end{cases}$

例 (2.1.8)

设连续型随机变量 X 的密度函数为 $f(x)=\begin{cases} x, & 0\leqslant x<1, \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其他}. \end{cases}$ 试求:

(1) $P(-1<X<0.5)$; (2) X 的分布函数 $F(x)$.

例 (2.1.8)

设连续型随机变量 X 的密度函数为 $f(x)=\begin{cases} x, & 0\leqslant x<1, \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其他}. \end{cases}$ 试求:

(1) $P(-1<X<0.5)$; (2) X 的分布函数 $F(x)$.

解

(1) $P(-1<X<0.5)=\int_{-1}^{0.5} f(x)\mathrm{d}x=\int_{0}^{0.5} x\mathrm{d}x=\left.\dfrac{x^2}{2}\right|_{0}^{0.5}=0.125;$

例 (2.1.8)

设连续型随机变量 X 的密度函数为 $f(x)=\begin{cases} x, & 0\leqslant x<1, \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其他}. \end{cases}$ 试求:

(1) $P(-1<X<0.5)$; (2) X 的分布函数 $F(x)$.

解

(1) $P(-1<X<0.5)=\int_{-1}^{0.5} f(x)\mathrm{d}x=\int_{0}^{0.5} x\mathrm{d}x=\left.\frac{x^2}{2}\right|_0^{0.5}=0.125$;

(2) 当 $x<0$ 时, $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=0$;

当 $0\leqslant x<1$ 时, $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\int_{0}^{x} t\mathrm{d}t=\frac{x^2}{2}$;

当 $1\leqslant x<2$ 时,

$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\int_{0}^{1} t\mathrm{d}t+\int_{1}^{x}(2-t)\mathrm{d}t=-\frac{x^2}{2}+2x-1$;

例 (2.1.8)

设连续型随机变量 X 的密度函数为 $f(x)=\begin{cases} x, & 0\leqslant x<1, \\ 2-x, & 1\leqslant x<2, \\ 0, & \text{其他}. \end{cases}$ 试求:

(1) $P(-1<X<0.5)$; (2) X 的分布函数 $F(x)$.

解

当 $x\geqslant 2$ 时, $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\int_0^1 t\mathrm{d}t+\int_1^2(2-t)\mathrm{d}t=1$;

综上所述, 得 X 的分布函数为

$$F(x)=\begin{cases} 0, & x<0; \\ \dfrac{x^2}{2}, & 0\leqslant x<1; \\ -\dfrac{x^2}{2}+2x-1, & 1\leqslant x<2; \\ 1, & \text{其他}. \end{cases}$$

例

设 $F_1(x)$ 与 $F_2(x)$ 为两个分布函数，其相应的概率密度 $f_1(x)$ 与 $f_2(x)$ 是连续函数，则必为概率密度的是 (　　).

A. $f_1(x)f_2(x)$

B. $2f_2(x)F_1(x)$

C. $f_1(x)F_2(x)$

D. $f_1(x)F_2(x)+f_2(x)F_1(x)$

例

设 $F_1(x)$ 与 $F_2(x)$ 为两个分布函数，其相应的概率密度 $f_1(x)$ 与 $f_2(x)$ 是连续函数，则必为概率密度的是（　　）.

A. $f_1(x)f_2(x)$

B. $2f_2(x)F_1(x)$

C. $f_1(x)F_2(x)$

D. $f_1(x)F_2(x)+f_2(x)F_1(x)$

解

检验密度函数的性质: $f_1(x)F_2(x)+f_2(x)F_1(x)\geqslant 0$,

$$\int_{-\infty}^{+\infty}[f_1(x)F_2(x)+f_2(x)F_1(x)]\,\mathrm{d}x=\int_{-\infty}^{+\infty}[F_1'(x)F_2(x)+F_2'(x)F_1(x)]\,\mathrm{d}x$$
$$=F_1(x)F_2(x)\Big|_{-\infty}^{+\infty}=1,$$

可知 $f_1(x)F_2(x)+f_2(x)F_1(x)$ 满足概率密度的两条基本性质. 故应选 D.

例

设 $F_1(x)$ 与 $F_2(x)$ 为两个分布函数，其相应的概率密度 $f_1(x)$ 与 $f_2(x)$ 是连续函数，则必为概率密度的是 ( ).

A. $f_1(x)f_2(x)$

B. $2f_2(x)F_1(x)$

C. $f_1(x)F_2(x)$

D. $f_1(x)F_2(x)+f_2(x)F_1(x)$

点评：本题将多个知识点综合在一起.

(1) 概率密度的性质: $f(x)\geqslant 0$, 对任何 $x\in(-\infty,+\infty)$, $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$.

(2) 分布函数的性质: $F(-\infty)=0,\quad F(+\infty)=1$;

(3) 分布函数与概率密度的关系: $f(x)=F'(x)$.

例

设随机变量 X 的概率密度满足 $f(-x)=f(x)$，$F(x)$ 是 X 的分布函数，则对任意的实数 a，下列式子中成立的是 (　　).

A. $F(-a)=1-\int_0^a f(x)\mathrm{d}x$

B. $F(-a)=\dfrac{1}{2}-\int_0^a f(x)\mathrm{d}x$

C. $F(-a)=F(a)$

D. $F(-a)=2F(a)-1$

例

设随机变量 X 的概率密度满足 $f(-x)=f(x)$，$F(x)$ 是 X 的分布函数，则对任意的实数 a，下列式子中成立的是 (　　).

A. $F(-a)=1-\int_0^a f(x)\mathrm{d}x$

B. $F(-a)=\dfrac{1}{2}-\int_0^a f(x)\mathrm{d}x$

C. $F(-a)=F(a)$

D. $F(-a)=2F(a)-1$

解

$F(-a)=\int_{-\infty}^{-a} f(x)\mathrm{d}x=\int_{-\infty}^{0} f(x)\mathrm{d}x+\int_0^{-a} f(x)\mathrm{d}x.$

因 $f(-x)=f(x)$，故 $f(x)$ 是偶函数，具有对称性，故 $\int_{-\infty}^{0} f(x)\mathrm{d}x=\dfrac{1}{2}$.

再令 $x=-t$，可得 $F(-a)=\dfrac{1}{2}+\int_0^a f(-t)\mathrm{d}(-t)=\dfrac{1}{2}-\int_0^a f(t)\mathrm{d}t.$

正确答案 B

例

设随机变量 X 的分布函数 $F(x)=\begin{cases}0, & x<0\\ \dfrac{1}{2}, & 0\leqslant x<1\\ 1-\mathrm{e}^{-x}, & x\geqslant 1\end{cases}$，则 $P\{X=1\}=($ $)$.

A. 0

B. $\dfrac{1}{2}$

C. $\dfrac{1}{2}-\mathrm{e}^{-1}$

D. $1-\mathrm{e}^{-1}$

例

设随机变量 X 的分布函数 $F(x)=\begin{cases}0, & x<0\\ \dfrac{1}{2}, & 0\leqslant x<1\\ 1-\mathrm{e}^{-x}, & x\geqslant 1\end{cases}$，则 $P\{X=1\}=($ $)$.

A. 0

B. $\dfrac{1}{2}$

C. $\dfrac{1}{2}-\mathrm{e}^{-1}$

D. $1-\mathrm{e}^{-1}$

解

由分布函数的性质, $P\{X=a\}=F(a)-F(a-0)$, 所以

$$P\{X=1\}=F(1)-F(1-0)=1-\mathrm{e}^{-1}-\frac{1}{2}=\frac{1}{2}-\mathrm{e}^{-1}.$$

1. 单点分布 (Single-point distribution)

若一个离散型随机变量的分布列为

$$P(X = a) = 1,$$

则称 X 服从单点分布或称 X 服从退化分布.

2. 两点分布 (Two-point distribution)

若随机变量 X 只可能取 a 与 b 两个值, 它的分布列为

$$P(X=a)=1-p,\ P(X=b)=p,\quad (0<p<1),$$

则称 X 服从参数为 p 的两点分布.

2. 两点分布 (Two-point distribution)

若随机变量 X 只可能取 a 与 b 两个值, 它的分布列为

$$P(X=a)=1-p,\ P(X=b)=p,\quad (0<p<1),$$

则称 X 服从参数为 p 的两点分布.
特别地, 当 $a=0$, $b=1$ 时, 两点分布也叫 $0-1$ 分布, 记作 $X\sim b(1,p)$,
其分布如表2所示.

表: $0-1$ 分布的分布律

X	0	1
p_k	$1-p$	p

$0-1$ 分布

上述 $0-1$ 分布的分布律也可以写成:

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

$0-1$ 分布是最常见的离散分布之一. 对于一个随机试验, 若它的样本空间只包含两个样本点, 即 $\Omega=\{\omega_1,\omega_2\}$, 我们总能在 Ω 上定义一个服从 $0-1$ 分布的随机变量

$$X=X(\omega)=\begin{cases}0, & \omega=\omega_1,\\ 1, & \omega=\omega_2.\end{cases}$$

用它来描述这个试验结果. 因此, 两点分布可以作为描述试验只包含两个基本事件的数学模型. 如在投篮 “命中” 与 “不中” 的概率分布; 产品抽验中 “合格品” 与 “不合格品” 的概率分布等等. 总之, 一个随机试验如果我们只关心某事件 A 出现与否, 则可用一个服从 $0-1$ 分布的随机变量来描述.

$0-1$ 分布

例 (2.2.1)

一箱中有 10 件产品，其中 8 件正品，2 件次品，从中任取 1 件．定义随机变量 X 如下

$$X = \begin{cases} 1, & \text{取到正品}, \\ 0, & \text{取到次品}. \end{cases}$$

则有

$$P(X=1) = 0.8,\ P(X=0) = 0.2,$$

即 X 服从 $0-1$ 分布．

3. 二项分布 (Binomial distribution)

若随机变量 X 的分布律为

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\quad k=0,1,2,\cdots,n,\quad 0<p<1.$$

则称 X 服从参数为 n,p 的二项分布，记作 $X\sim b(n,p)$. 而对于二项分布的第 k 项常记作 $b(k;n,p)$，即 $b(k;n,p)=\binom{n}{k}p^k(1-p)^{n-k}$. 不难看出，当 $n=1$ 时二项分布就是 $0-1$ 分布.

3. 二项分布 (Binomial distribution)

若随机变量 X 的分布律为

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ \ k=0,1,2,\cdots,n,\ \ 0<p<1.$$

则称 X 服从参数为 n,p 的二项分布，记作$X\sim b(n,p)$. 而对于二项分布的第 k 项常记作$b(k;n,p)$，即 $b(k;n,p)=\binom{n}{k}p^k(1-p)^{n-k}$. 不难看出，当 $n=1$ 时二项分布就是 $0-1$ 分布.

易知该分布列确实满足非负性和正则性. 事实上，$P(X=k)\geqslant 0$ 是显然的; 再由二项展开式知

$$\sum_{k=0}^{n}P(X=k)=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=[p+(1-p)]^n=1.$$

我们知道，$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ 恰好是 $[p+(1-p)]^n$ 二项展开式中出现 p^k 的那一项，这就是二项分布名称的由来.

回忆 n 重伯努利试验中事件 A 出现 k 次的概率计算公式

$$p_n(k)=\binom{n}{k}p^k(1-p)^{n-k}, k=0,1,2,\cdots,n.$$

可知, 若用 X 表示 n 重伯努利试验中事件 A 出现的次数, 则 $X\sim b(n,p)$. 因此, 二项分布可以作为描述 n 重伯努利试验中事件 A 出现次数的数学模型. 比如,

回忆 n 重伯努利试验中事件 A 出现 k 次的概率计算公式

$$p_n(k)=\binom{n}{k}p^k(1-p)^{n-k}, k=0,1,2,\cdots,n.$$

可知, 若用 X 表示 n 重伯努利试验中事件 A 出现的次数, 则 $X\sim b(n,p)$. 因此, 二项分布可以作为描述 n 重伯努利试验中事件 A 出现次数的数学模型. 比如,

(1) n 次射击中命中次数 $X\sim b(n,p)$, p 为每次射击的命中率.

(2) 随机抛掷硬币 n 次, 落地时出现 "正面" 次数 $Y\sim b(n,0.5)$.

(3) n 次有放回的随机摸球中, 摸到黑球个数 $Z\sim b(n,p)$, p 为袋中黑球的百分比.

(4) 从一批的产品中有放回任意抽取 n 件, 其中 "次品" 件数 $T\sim b(n,p)$, p 为产品中的次品率.

例 (2.2.2)

设某批产品的不合格率为 2%, 按某种抽样检验方案规定, 若在这批产品中任取 100 件, 经检验合格后发现不合格产品数不超过 3 件, 则可以认为这批产品合格而予以接收. 试求这批产品被接收的概率.

例 (2.2.2)

设某批产品的不合格率为 2%，按某种抽样检验方案规定，若在这批产品中任取 100 件，经检验合格后发现不合格产品数不超过 3 件，则可以认为这批产品合格而予以接收. 试求这批产品被接收的概率.

解

由于产品数量大，因此无放回的抽样检验也可以看成是有放回的，抽出的一件产品或是不合格品或是合格品，只有两种可能结果，因此抽出一件产品可以看成是 1 次伯努利试验，而且抽样可以看成是有放回的，因此每次抽样抽到不合格产品的概率为 0.02 保持不变；共抽取 100 件产品就相当于独立重复进行了 100 次伯努利试验. 记 X 表示 100 件产品中不合格品的个数，则

$$X \sim b(100, 0.02).$$

例 (2.2.2)

设某批产品的不合格率为 2%, 按某种抽样检验方案规定, 若在这批产品中任取 100 件, 经检验合格后发现不合格产品数不超过 3 件, 则可以认为这批产品合格而予以接收. 试求这批产品被接收的概率.

解

$$X \sim b(100, 0.02).$$

于是所求的产品被接收的概率为

$$P(X \leqslant 3) = \sum_{k=0}^{3} \binom{100}{k} 0.02^k (1-0.02)^{100-k}.$$

由于 $n = 100$ 较大, $p = 0.02$ 较小, 直接计算上述概率较为困难, 希望能够找到计算这个二项概率的近似公式. 1837 年法国科学家泊松 (Poisson) 经过研究得到了二项分布的逼近公式, 这就是下面的泊松定理.

定理 (2.2.1 泊松 (Poisson) 定理)

设 $np_n = \lambda$ ($\lambda > 0$ 是一常数, n 是任意正整数), 则对任意一固定的非负整数 k, 有

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

定理 (2.2.1 泊松 (Poisson) 定理)

设 $np_n = \lambda$ ($\lambda > 0$ 是一常数, n 是任意正整数), 则对任意一固定的非负整数 k, 有

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

证明.

由 $p_n = \lambda/n$, 有

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \left[1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right]\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)^{-k}.$$

□

证明.

对任意固定的 k, 当 $n \to \infty$ 时,

$$\left[1 \cdot \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right] \to 1,$$

$$\left(1-\frac{\lambda}{n}\right)^{n} \to e^{-\lambda}, \quad \left(1-\frac{\lambda}{n}\right)^{-k} \to 1.$$

故 $$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}.$$

由于 $\lambda = np_n$ 是常数, 所以当 n 很大时 p_n 必定很小, 因此, 上述定理表明当 n 很大 p 很小时, 有以下近似公式:

$$\binom{n}{k}p^k(1-p)^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda}, \tag{4}$$

其中 $\lambda = np$. □

二项分布的泊松近似

从下表可以直观地看出 (4) 式两端的近似程度.

表: 二项分布的泊松近似

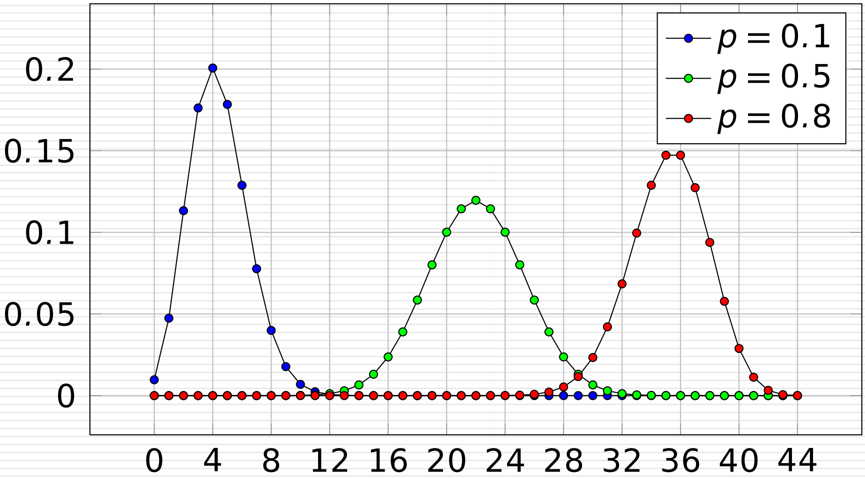
k	按二项分布公式直接计算				按泊松近似公式计算
	$n=10$ $p=0.1$	$n=20$ $p=0.05$	$n=40$ $p=0.025$	$n=100$ $p=0.01$	$\lambda=1\ (=np)$
0	0.349	0.358	0.363	0.366	0.368
1	0.385	0.377	0.372	0.370	0.368
2	0.194	0.189	0.186	0.185	0.184
3	0.057	0.060	0.060	0.061	0.061
4	0.011	0.013	0.014	0.015	0.015
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

二项分布的泊松近似

由上表可以看出，两者的结果是很接近的．在实际计算中，当 $n \geqslant 20$, $p \leqslant 0.05$ 时近似效果颇佳，而当 $n \geqslant 100$, $np \leqslant 10$ 时效果更好．$\frac{\lambda^k}{k!}e^{-\lambda}$ 的值可由数表查询计算得到 (见本书附录 1.1)

二项分布的泊松近似，常常被应用于研究稀有事件 (即每次试验中事件 A 出现的概率 p 很小)，当伯努利试验的次数 n 很大时，事件 A 发生的次数的分布．

二项分布 $n = 44$



例 (2.2.3)

某十字路口有大量汽车通过，假设每辆汽车在这里发生交通事故的概率为 0.001，如果每天有 5000 辆汽车通过这个十字路口，求发生交通事故的汽车数不少于 2 的概率.

例 (2.2.3)

某十字路口有大量汽车通过，假设每辆汽车在这里发生交通事故的概率为 0.001，如果每天有 5000 辆汽车通过这个十字路口，求发生交通事故的汽车数不少于 2 的概率.

解

设 X 表示发生交通事故的汽车数，则 $X \sim b(n,p)$，此处 $n = 5000$，$p = 0.001$，令 $\lambda = np = 5$，

$$
\begin{aligned}
P(X \geqslant 2) = 1 - P(X < 2) &= 1 - \sum_{k=0}^{1} P\{X = k\} \\
&= 1 - 0.999^{5000} - 5 \times (0.999)^{4999} \\
&= 1 - \frac{5^0}{0!} e^{-5} - \frac{5}{1!} e^{-5}.
\end{aligned}
$$

查表可得 $P(X \geqslant 2) = 1 - 0.00674 - 0.03369 = 0.95957$.

例 (2.2.4)

某校有 100 门电话分机. 根据实际使用情况, 在上班时间内每门分机有 4% 的时间需要使用外线. 假定各分机是否使用外线是相互独立的. 试问总机至少需要多少条外线才有 95% 以上的把握保证每门分机使用外线不需要等候?

例 (2.2.4)

某校有 100 门电话分机．根据实际使用情况，在上班时间内每门分机有 4% 的时间需要使用外线．假定各分机是否使用外线是相互独立的．试问总机至少需要多少条外线才有 95% 以上的把握保证每门分机使用外线不需要等候?

解

由于各分机是否需要使用外线是相互独立的，100 门分机是否使用外线可以看成是 100 次独立试验，每次试验只有两个可能结果：使用外线的概率为 0.04，不使用外线的概率为 0.96．设 $X =$“100 门分机中同时使用外线的门数”，则可知 $X \sim b(100, 0.04)$.

另设 x 表示需安装的外线数，则由题意知所求的 x 应满足

$$P(X \leqslant x) \geqslant 95\%.$$

解 $(P(X \leqslant x) \geqslant 95\%)$

注意到 n 较大，p 较小，$\lambda = np = 100 \times 0.04 = 4$，由泊松定理知

$$P(X \leqslant x) \approx \sum_{k=0}^{x} \frac{4^k}{k!} e^{-4} \geqslant 95\%.$$

查书后附表可知

$$\sum_{k=0}^{7} \frac{4^k}{k!} e^{-4} = 0.9489 < 0.95,$$

$$\sum_{k=0}^{8} \frac{4^k}{k!} e^{-4} = 0.9786 > 0.95.$$

故至少需要安装 8 条外线才能以 95% 以上的把握保证每门分机使用外线不必等候.

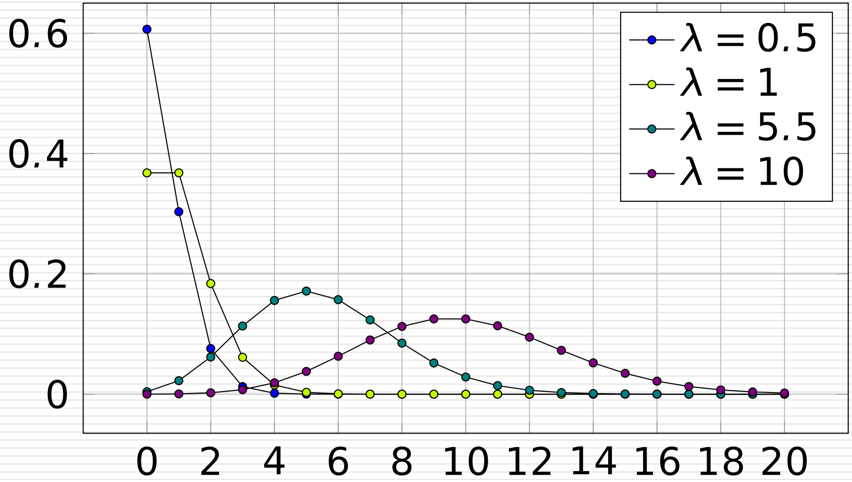
4. 泊松分布

泊松分布是 1837 年由法国数学家泊松 (Poisson) 首次提出的.
若随机变量 X 的分布列为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},\ k=0,1,2,\cdots, \tag{5}$$

其中参数 $\lambda>0$ 是常数, 则称 X 服从参数为 λ 的泊松分布(Poisson distribution), 记为$X\sim P(\lambda)$.

泊松分布



0.6
0.4
0.2
0
0 2 4 6 8 10 12 14 16 18 20
$\lambda = 0.5$
$\lambda = 1$
$\lambda = 5.5$
$\lambda = 10$

二项分布与泊松分布

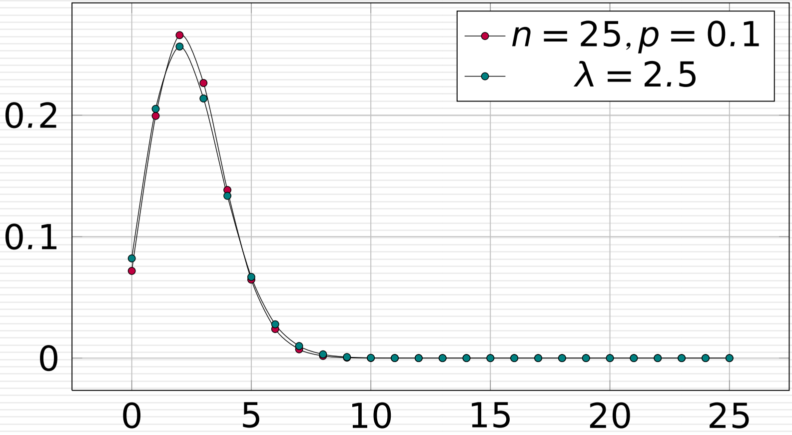


图: $b(n,p)$ 与 $P(\lambda)$ 近似 (n 大, p 小, $np = \lambda$)

泊松分布是一种常用分布，它常与单位时间 (或单位面积、单位产品等) 上的计数过程相联系，比如:

(1) 在单位时间内，电话总机接到用户呼唤的次数.

(2) 在单位时间内，一电路受到外界电磁波的冲击次数.

(3) 一段固定的时间间隔内由某块放射性物质放射出的 α 质点，到达某个计数器的质点数.

(4) 单位时间内来到某公共设施要求给予服务的顾客数 (这里的公共设施的意义可以是极为广泛的、诸如售货员机场跑道、电话交换台、医院等).

(5) 单位时间内，某支股票价格发生突变的次数.

(6) 单位时间内，某地区发生交通事故的次数.

都服从泊松分布. 因此泊松分布的应用面是十分广泛的，相应的分布函数值有数表可供查询 (见本书附录 1.1).

例 (2.2.5)

已知出租汽车每分钟到达总站的辆数 X 服从参数 $\lambda = 3$ 的泊松分布，试求
(1) 在某指定的一分钟内恰有 5 辆出租车到达总站的概率.
(2) 在某指定的一分钟内有多于 5 辆的出租车到达总站的概率.

例 (2.2.5)

已知出租汽车每分钟到达总站的辆数 X 服从参数 $\lambda = 3$ 的泊松分布，试求

(1) 在某指定的一分钟内恰有 5 辆出租车到达总站的概率.

(2) 在某指定的一分钟内有多于 5 辆的出租车到达总站的概率.

解

由题知 X 的分布列为 $P(X=k)=\dfrac{3^k}{k!}e^{-3},\ k=0,1,2,\cdots$.

(1) 在某指定的一分钟内恰有 5 辆出租车到达总站的概率为

$$P(X=5)=\frac{3^5}{5!}e^{-3}=0.1008.$$

(2) 在某指定的一分钟内有多于 5 辆的出租车到达总站的概率为

$$P(X>5)=P(X\geqslant 6)=1-P(X\leqslant 5)=1-\sum_{k=0}^{5}\frac{3^k}{k!}e^{-3}=0.0839.$$

例 (2.2.6)

由某商店过去的销售记录知道，某种商品每月的销售数可以用参数 $\lambda = 5$ 的泊松分布来描述．为了以 95% 以上的把握保证不脱销，问商店在月底至少应进某种商品多少件?

例 (2.2.6)

由某商店过去的销售记录知道，某种商品每月的销售数可以用参数 $\lambda = 5$ 的泊松分布来描述．为了以 95% 以上的把握保证不脱销，问商店在月底至少应进某种商品多少件?

解

设该商店每月销售这种商品数为 X，月底进货为 a 件，则当 $X \leqslant a$ 时不脱销，故有 $P(X \leqslant a) > 0.95$.

由于 $X \sim P(5)$，上式即为 $\sum\limits_{k=0}^{a} \frac{5^k}{k!} e^{-5} > 0.95$. 查表可知

$$\sum_{k=0}^{8} \frac{5^k}{k!} e^{-5} \approx 0.9319 < 0.95, \sum_{k=0}^{9} \frac{5^k}{k!} e^{-5} \approx 0.9682 > 0.95.$$

于是，这家商店只要在月底进货这种商品 9 件 (假定上个月没有存货)，就可以 95% 以上的把握保证这种商品在下个月不会脱销.

5. 几何分布

若一个随机变量 X 的分布列为

$$P(X=k)=pq^{k-1},\ k=1,2,3,\cdots,\ 0<p<1,\ q=1-p, \tag{6}$$

则称 X 服从参数为 p 的**几何分布**, 记为$X\sim Ge(p)$.
(6) 式的右端为几何级数的一般项, 这正是几何分布名称的由来, 容易验证它满足分布列的非负性和正则性.

实际中有许多的随机变量服从几何分布:

(1) 某产品的不合格率为 0.05, 则首次查到不合格品的检查次数 $X \sim Ge(0.05)$.

(2) 某射手的命中率为 0.9, 则首次命中目标时的射击次数 $Y \sim Ge(0.9)$.

(3) 同时抛掷两颗骰子, 则出现双六点时抛掷的次数 $Z \sim Ge(1/36)$.

(4) 在独立重复试验中, 首次成功出现时的试验总次数 $T \sim Ge(p)$, 其中 p 为一次试验中成功的概率.

可以证明几何分布满足如下的无记忆性，设 $X \sim Ge(p)$，则对于任意的正整数 m, n 有

$$P(X > m+n \mid X > m) = P(X > n).$$

证明.

因为

$$P(X > m) = 1 - P(X \leqslant m) = 1 - \sum_{k=1}^{m} pq^{k-1} = 1 - p\frac{1-q^m}{1-q} = q^m,$$

所以，

$$P(X > m+n \mid X > m) = \frac{P(X > m+n)}{P(X > m)} = \frac{q^{m+n}}{q^m} = q^n = P(X > n).$$

□

6. 超几何分布 *

若一个随机变量 X 的分布列为

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\quad k=0,1,\cdots,r$$

其中 $r=\min(n,M)$，且 $M\leqslant N$，$n\leqslant N$，n,N,M 均为正整数，则称 X 服从参数为 n,N,M 的**超几何分布**，记为 $X\sim h(n,N,M)$.

从一个有限总体中进行不放回抽样常会碰到超几何分布. 设一批产品共有 N 件，其中有 M 件不合格品，$N-M$ 件合格品. 现按无放回的方式抽取 n 个产品，记 X 为其中含有的不合格产品的件数，则由例 1.2.2 知 $X\sim h(n,N,M)$.

7. 巴斯卡分布 *

若一个随机变量 X 的分布列为

$$P(X=k)=\binom{k-1}{r-1}p^r q^{k-r},$$

其中: $k=r,r+1,r+2,\cdots$, $0<p<1$, $q=1-p$, $r\geqslant 1$ 为正整数, 则称 X 服从参数为 r,p 的巴斯卡分布, 记为$X\sim Nb(r,p)$. 显然当 $r=1$ 时即为几何分布.

巴斯卡分布是几何分布的一种推广. 可以看到, 在独立重复试验中, 第 r 次成功出现时的试验总次数 $X\sim Nb(r,p)$, 其中 p 为一次试验中成功的概率.

1. 随机变量与分布函数

2. 常见分布
   - 几种常见的离散型分布
   - 常见的连续型分布

3. 随机变量函数的分布

1. 均匀分布

称一个随机变量 X 服从 (a, b) 区间上的均匀分布(Uniform distribution), 若其概率密度函数为

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

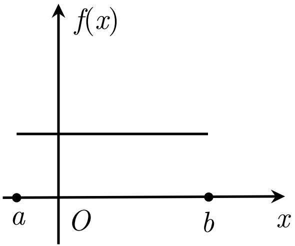
随机变量 X 服从 (a, b) 区间上的均匀分布通常记作$X \sim U(a, b)$, 其分布函数可以表示为

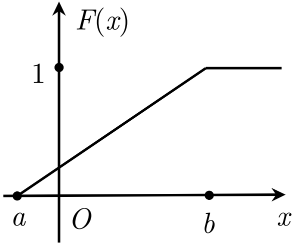
$$F(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{b-a}, & a \leqslant x < b, \\ 1, & x \geqslant b. \end{cases}$$

密度函数和分布函数的图像如下图所示.

1. 均匀分布

均匀分布的密度函数和分布函数的图像如下图所示.





均匀分布的背景可以看作是随机点 X 落在区间 (a, b) 上的位置. 均匀分布在实际中经常被使用, 如乘客在公共汽车站候车的时间, 近似计算中的舍入误差等常用均匀分布来描述.

例 (2.2.7)

已知秒表刻度的分划值为 0.2 秒. 如果计时精度取到邻近的整数值, 则使用该秒表时误差的绝对值大于 0.05 秒的概率是多少?

例 (2.2.7)

已知秒表刻度的分划值为 0.2 秒. 如果计时精度取到邻近的整数值, 则使用该秒表时误差的绝对值大于 0.05 秒的概率是多少?

解

设使用该秒表时的误差为 X, 则 $X \sim U(-0.1, 0.1)$, 其密度函数为

$$p(x) = \begin{cases} \dfrac{1}{0.2} = 5, & -0.1 < x < 0.1; \\ 0, & \text{其他}. \end{cases}$$

所求概率为

$$P(|X| > 0.05) = \int_{-0.1}^{-0.05} 5\mathrm{d}x + \int_{0.05}^{0.1} 5\mathrm{d}x = 0.5.$$

2. 指数分布

称随机变量 X 服从参数为 λ $(\lambda > 0)$ 的指数分布(Exponentially distribution), 若其密度函数为:

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

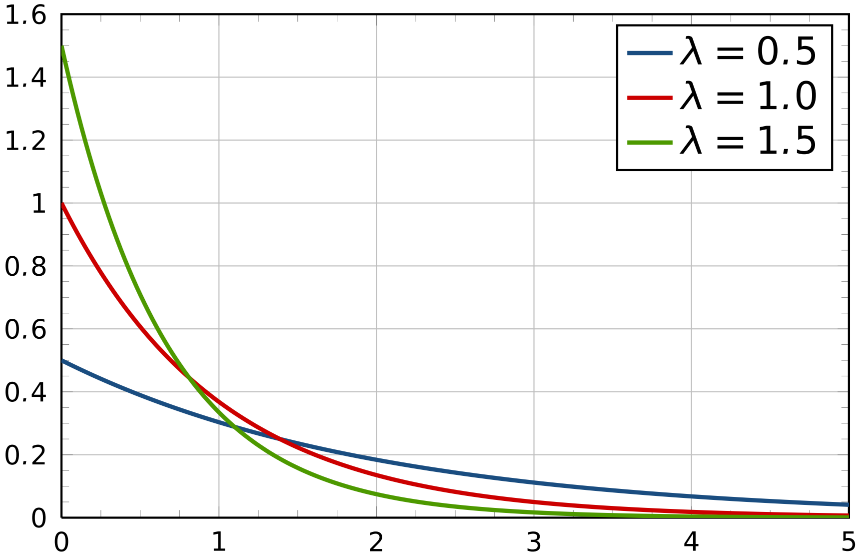
通常记作$X \sim Exp(\lambda)$.
因为

$$\int_0^x \lambda e^{-\lambda t}\,\mathrm{d}t = -\int_0^x e^{-\lambda t}\,\mathrm{d}(-\lambda t) = -e^{-\lambda t}\,\Big|_0^x = 1 - e^{-\lambda x},$$

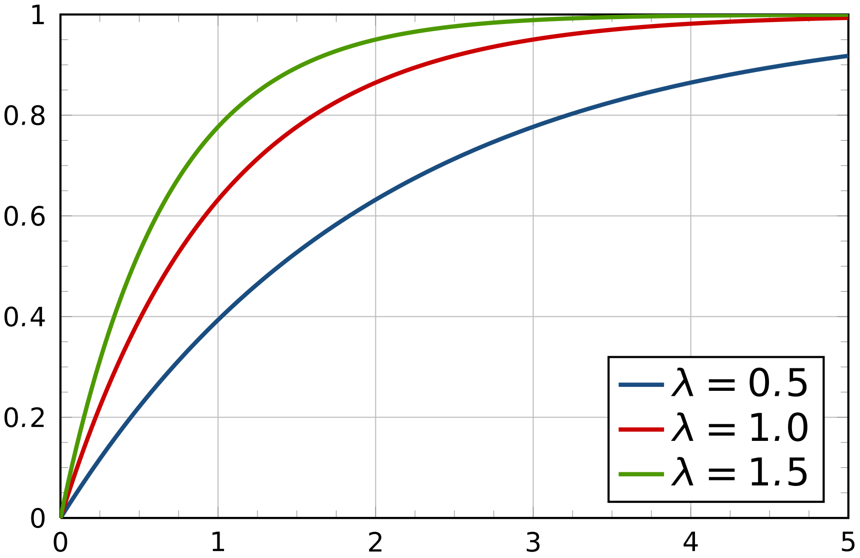
所以 X 的分布函数为

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

指数分布的密度函数



指数分布的分布函数



2. 指数分布

注意到指数分布的随机变量只可能取非负值，所以指数分布常被用来描述各种“寿命分布”，如：电子元件的寿命、动物的寿命、电话的通话时长、随机服务系统的服务时间等.

与几何分布一样，指数分布也具有所谓的“无记忆性”，即对于任意的实数 $s, t > 0$ 有

$$P(X > s + t \mid X > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t).$$

事实上，更进一步可以证明：具有无记忆性质的离散型分布一定为几何分布，具有无记忆性质的连续型分布一定为指数分布.

例 (2.2.8)

设顾客在某银行的窗口等待服务的时间 X(以分钟计算) 服从指数分布, 其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{5}e^{-\frac{1}{5}x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(1) 计算 $P(-1<X<10)$.

(2) 设某顾客在窗口等待服务, 若超过 10 分钟, 他就离开. 他一个月要到银行 5 次, 以 Y 表示他未等待服务而离开窗口的次数, 试求 $P(Y\geqslant 1)$.

例 (2.2.8)

设顾客在某银行的窗口等待服务的时间 X(以分钟计算) 服从指数分布，其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{5}e^{-\frac{1}{5}x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(1) 计算 $P(-1<X<10)$.
(2) 设某顾客在窗口等待服务，若超过 10 分钟，他就离开. 他一个月要到银行 5 次，以 Y 表示他未等待服务而离开窗口的次数，试求 $P(Y\geqslant 1)$.

解

(1) 根据密度函数性质有

$$P(-1<x<10)=\int_{-1}^{10}f(x)\mathrm{d}x=\int_{0}^{10}\frac{1}{5}e^{-\frac{1}{5}x}\mathrm{d}x=1-e^{-2}.$$

解

(1) 根据密度函数性质有

$$P(-1 < x < 10) = \int_{-1}^{10} f(x)\mathrm{d}x = \int_{0}^{10} \frac{1}{5} e^{-\frac{1}{5}x}\mathrm{d}x = 1 - e^{-2}.$$

(2) 记 $A =$“未等待服务而离开”, 则有

$$P(A) = P(X > 10) = 1 - P(X \leqslant 10) = e^{-2}.$$

由此可知: $Y \sim b(5, e^{-2})$. 于是

$$P(Y \geqslant 1) = 1 - P(Y = 0) = 1 - (1 - e^{-2})^5.$$

3. 正态分布

若连续型随机变量 X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<+\infty,$$

其中 μ,σ $(\sigma>0)$ 为常数，则称 X 服从参数为 μ,σ $(\sigma>0)$ 的正态分布(Normal distribution)，记为 $X\sim N(\mu,\sigma^2)$.

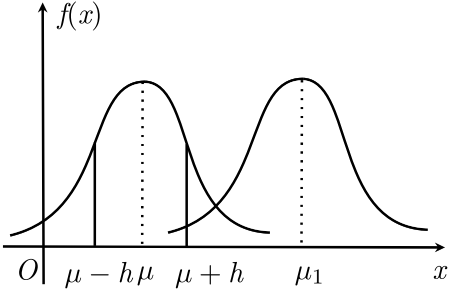
相应的分布函数为

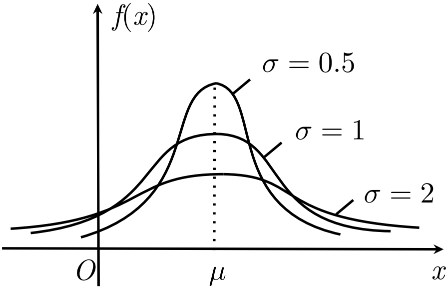
$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t.$$

其密度函数如下图所示.

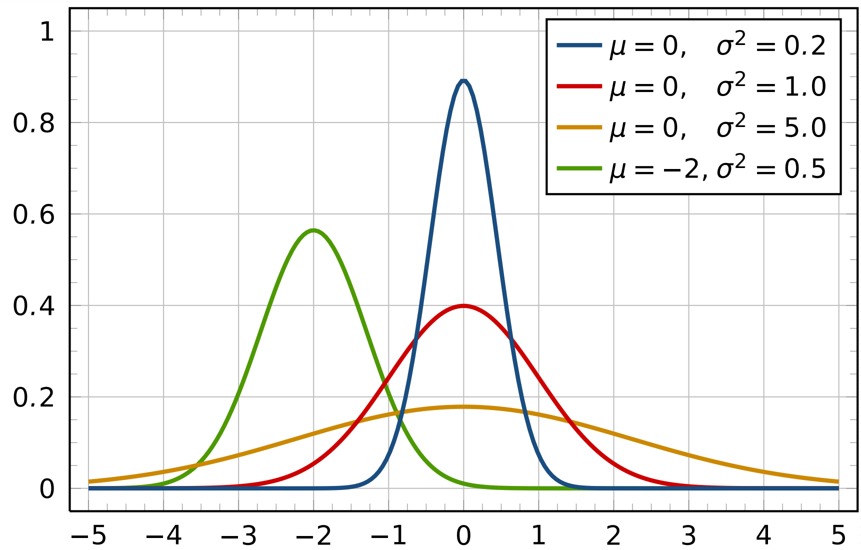
3. 正态分布

正态分布的密度函数如下图所示.

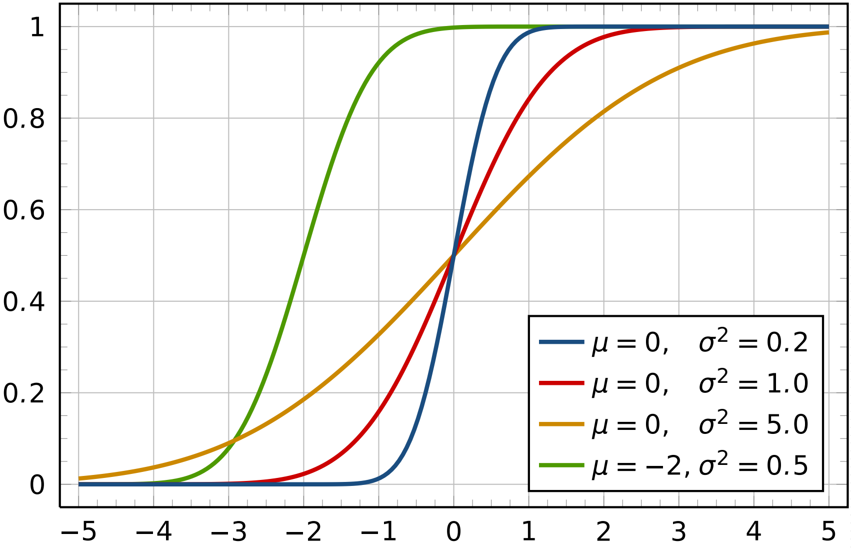




正态分布的密度函数


$\mu = 0, \quad \sigma^2 = 0.2$
$\mu = 0, \quad \sigma^2 = 1.0$
$\mu = 0, \quad \sigma^2 = 5.0$
$\mu = -2, \sigma^2 = 0.5$
1
0.8
0.6
0.4
0.2
0
−5
−4
−3
−2
−1
0
1
2
3
4
5

正态分布的分布函数


1
0.8
0.6
0.4
0.2
0
−5 −4 −3 −2 −1 0 1 2 3 4 5
$\mu = 0, \quad \sigma^2 = 0.2$
$\mu = 0, \quad \sigma^2 = 1.0$
$\mu = 0, \quad \sigma^2 = 5.0$
$\mu = -2, \sigma^2 = 0.5$

正态分布的密度函数具有如下特点:

(1) 曲线关于 $x=\mu$ 对称.

(2) 曲线在$x=\mu$处取到最大值, x 离 μ 越远, $f(x)$ 值越小. 这表明对于同样长度的区间, 当区间离 μ 越远, X 落在这个区间上的概率越小.

(3) 曲线在$\mu\pm\sigma$处有拐点.

(4) 曲线以 x 轴为渐近线.

(5) 若固定 μ, 当 σ 越小时图形越尖陡(图 2-2), 因而 X 落在 μ 附近的概率越大; 若固定 σ, μ 值改变, 则图形沿 x 轴平移, 而不改变其形状. 故称 σ 为精度参数, μ 为位置参数.

(2) 曲线在$x=\mu$处取到最大值

$$\left(\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)' = -\frac{(x-\mu)}{\sigma^2}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

(3) 曲线在$\mu\pm\sigma$处有拐点

$$\left(\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\right)'' = -\frac{1}{\sigma^2}\left(1-\frac{(x-\mu)^2}{\sigma^2}\right)\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

- 正态分布是概率论与数理统计中最重要的一种分布，德国数学家高斯 (Gauss) 在研究误差理论时首先用正态分布来刻画误差的分布，因此正态分布又称**高斯分布**.
- 一方面，在自然界和人类社会中，大量的随机变量都服从或近似服从正态分布. 如测量的误差、射击时弹着点距离目标的偏差、农作物的产量、工业产品的尺寸 (直径、长度等)、学生的身高、体重、学习成绩等.
- 一般说来若某一数量指标受很多随机因素的影响，而每种因素对该指标所起的作用又非常微小，那么这个指标就近似服从正态分布.
- 另一方面，正态分布具有很多优良的性质，许多分布都可以用正态分布来逼近，一些分布可由正态分布来导出.

补充积分

$$I=\int_0^{+\infty} e^{-x^2}\,\mathrm{d}x=\lim_{R\to+\infty}\int_0^R e^{-x^2}\,\mathrm{d}x=\frac{\sqrt{\pi}}{2}.$$

证明.

$$I^2=\lim_{R\to+\infty}\int_0^R e^{-y^2}\,\mathrm{d}y\int_0^R e^{-x^2}\,\mathrm{d}x=\lim_{R\to+\infty}\iint\limits_{\substack{0\leqslant x\leqslant R\\0\leqslant y\leqslant R}} e^{-\left(x^2+y^2\right)}\,\mathrm{d}x\,\mathrm{d}y,$$

$$\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^R e^{-r^2}r\,\mathrm{d}r\leqslant\iint\limits_{\substack{0\leqslant x\leqslant R\\0\leqslant y\leqslant R}} e^{-\left(x^2+y^2\right)}\,\mathrm{d}x\,\mathrm{d}y\leqslant\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^{\sqrt{2}R} e^{-r^2}r\,\mathrm{d}r,$$

$$\lim_{R\to+\infty}\mathit{Left}=\lim_{R\to+\infty}\mathit{Right}=\frac{\pi}{2}\cdot\frac{1}{2}=\frac{\pi}{4},\quad\therefore I^2=\frac{\pi}{4},\quad\therefore I=\frac{\sqrt{\pi}}{2}.$$

补充积分

$$\boxed{\int_0^{+\infty} e^{-x^2}\, \mathrm{d}x = \frac{\sqrt{\pi}}{2}}$$

$$\boxed{\int_{-\infty}^{+\infty} e^{-x^2}\, \mathrm{d}x = \sqrt{\pi}}$$

$$\boxed{\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, \mathrm{d}x = \sqrt{2\pi}}$$

标准正态分布

$\mu=0$, $\sigma=1$ 时的正态分布称为标准正态分布, X 服从标准正态分布通常记作 $X\sim N(0,1)$, 其密度函数由专用符号表示为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad -\infty<x<+\infty,$$

分布函数为

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,\mathrm{d}t,$$

下面讨论正态随机变量 X 落在任意区间 (a,b) 中的概率.

(1) 设 $X \sim N(0,1)$

由于标准正态分布函数 $\Phi(x)$ 不是初等函数, 故编制了相应的标准正态分布函数表 (见附录 1.2) 供计算查询.
注意到标准正态分布的密度函数 $\varphi(x)$ 是关于 y 轴对称的, 所以有

$$\Phi(-x) + \Phi(x) = 1. \tag{7}$$

于是可得 $X \sim N(0,1)$, 有

$$P(a < X < b) = \Phi(b) - \Phi(a); \tag{8}$$

$$P(|X| < a) = 2\Phi(a) - 1, \ \ a > 0. \tag{9}$$

(8) 式是显然的, (9) 式可由 (7) 式与 (8) 式得到.

例 (2.2.9)

设 $X \sim N(0,1)$, 试求如下概率:
(1) $P(|X| < 1)$; (2) 求 x 使 $P(|X| > x) = 0.10$.

例 (2.2.9)

设 $X \sim N(0,1)$, 试求如下概率:
(1) $P(|X| < 1)$; (2) 求 x 使 $P(|X| > x) = 0.10$.

解

(1) 由 (9) 知 $P(|X| < 1) = 2\Phi(1) - 1$, 查表得: $\Phi(1) = 0.8413$, 因此

$$P(|X| < 1) = 2\Phi(1) - 1 = 2 \times 0.8413 - 1 = 0.6826;$$

(2) 显然 $x > 0$, 于是有

$$P(|X| > x) = 1 - P(|X| \leqslant x) = 1 - [2\Phi(x) - 1] = 2 - 2\Phi(x) = 0.10,$$

于是有 $\Phi(x) = 0.95$. 反查标准正态分布函数表可得: $x = 1.65$.

(2) $X \sim N(\mu, \sigma^2)$

由于

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \mathrm{d}t = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}} \mathrm{d}u = \Phi\left(\frac{x-\mu}{\sigma}\right),$$

注意在第二个等号用到了变换替换 $u = \dfrac{t-\mu}{\sigma}$.

由此可得如下重要结论:

若 $X \sim N(\mu, \sigma^2)$, 则 $\dfrac{X-\mu}{\sigma} \sim N(0,1)$, 并称 $\dfrac{X-\mu}{\sigma}$ 是将随机变量 X 标准化. 所以有

$$P(a < X < b) = F(b) - F(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

例 (2.2.10)

设 $X \sim N(3,4)$, 试求如下概率:
(1) $P(X \leqslant 5)$; (2) $P(1 < X \leqslant 5)$; (3) $P(|X| < 1)$.

例 (2.2.10)

设 $X \sim N(3,4)$, 试求如下概率:
(1) $P(X \leqslant 5)$; (2) $P(1 < X \leqslant 5)$; (3) $P(|X| < 1)$.

解

$$P(X \leqslant 5) = \Phi\left(\frac{5-3}{2}\right) = \Phi(1) = 0.8413;$$

$$\begin{aligned} P(1 < X \leqslant 5) &= \Phi\left(\frac{5-3}{2}\right) - \Phi\left(\frac{1-3}{2}\right) = \Phi(1) - \Phi(-1) \\ &= 2\Phi(1) - 1 = 0.6826; \end{aligned}$$

$$\begin{aligned} P(|X| < 1) &= P(-1 \leqslant X \leqslant 1) = \Phi\left(\frac{1-3}{2}\right) - \Phi\left(\frac{-1-3}{2}\right) \\ &= \Phi(-1) - \Phi(-2) = 1 - \Phi(1) - [1 - \Phi(2)] \\ &= \Phi(2) - \Phi(1) = 0.1359. \end{aligned}$$

例 (2.2.11)

设 $X \sim N(\mu, \sigma^2)$，计算 $P(|X-\mu| \leqslant k\sigma)$，$(k = 1, 2, 3)$.

例 (2.2.11)

设 $X \sim N(\mu, \sigma^2)$，计算 $P(|X-\mu| \leqslant k\sigma)$，$(k=1,2,3)$.

解

当 $k=1$ 时，

$$\begin{aligned} P(|X-\mu| \leqslant \sigma) &= P(\mu-\sigma \leqslant X \leqslant \mu+\sigma) \\ &= \Phi(1)-\Phi(-1) = 2\Phi(1)-1 = 0.6826. \end{aligned}$$

同理可得

$$\begin{aligned} P(|X-\mu| \leqslant 2\sigma) = 2\Phi(2)-1 = 0.9544, \\ P(|X-\mu| \leqslant 3\sigma) = 2\Phi(3)-1 = 0.9974. \end{aligned}$$

由上可知，尽管正态变量的取值范围是 $(-\infty,+\infty)$，但它的值落在$(\mu-3\sigma, \mu+3\sigma)$内几乎是肯定的事，因此在实际问题中，基本上可以认为有 $|X-\mu|<3\sigma$. 这个性质被实际工作者称为正态分布的 "3σ 原则".

例

设随机变量 $X \sim N(\mu, \sigma^2)$ $(\sigma > 0)$，记 $p = P\{X \leqslant \mu + \sigma^2\}$, 则 (   )

A. p 随着 μ 的增加而增加.

B. p 随着 σ 的增加而增加.

C. p 随着 μ 的增加而减少.

D. p 随着 σ 的增加而减少.

例

设随机变量 $X \sim N(\mu, \sigma^2)$ $(\sigma > 0)$，记 $p = P\{X \leqslant \mu + \sigma^2\}$，则 (　　)

A. p 随着 μ 的增加而增加.

B. p 随着 σ 的增加而增加.

C. p 随着 μ 的增加而减少.

D. p 随着 σ 的增加而减少.

解

因为

$$P\{X \leqslant \mu + \sigma^2\} = P\left\{\frac{X-\mu}{\sigma} \leqslant \sigma\right\} = \Phi(\sigma),$$

所以概率 p 随着 σ 的增加而增加.

例

设 X_1, X_2, X_3 是随机变量，且 $X_1 \sim N(0,1)$，$X_2 \sim N(0,2^2)$，$X_3 \sim N(5,3^2)$，$p_j = P\{-2 \leqslant X_j \leqslant 2\}$ $(j=1,2,3)$，则 (　　)

A. $p_1 > p_2 > p_3$.

B. $p_2 > p_1 > p_3$.

C. $p_3 > p_1 > p_2$.

D. $p_1 > p_3 > p_2$.

例

设 X_1, X_2, X_3 是随机变量，且 $X_1 \sim N(0,1)$，$X_2 \sim N(0,2^2)$，$X_3 \sim N(5,3^2)$，$p_j = P\{-2 \leqslant X_j \leqslant 2\}$（$j=1,2,3$），则（ ）

A. $p_1 > p_2 > p_3$.

B. $p_2 > p_1 > p_3$.

C. $p_3 > p_1 > p_2$.

D. $p_1 > p_3 > p_2$.

解

$p_1 = P\{-2 \leqslant X_1 \leqslant 2\} = 2\Phi(2) - 1$;

$p_2 = P\{-2 \leqslant X_2 \leqslant 2\} = P\left\{-1 \leqslant \frac{X_2-0}{2} \leqslant 1\right\} = 2\Phi(1) - 1$;

$p_3 = P\{-2 \leqslant X_3 \leqslant 2\} = P\left\{-\frac{7}{3} \leqslant \frac{X_3-5}{3} \leqslant -1\right\} = \Phi(-1) - \Phi\left(-\frac{7}{3}\right)$.

因为 $\Phi(x)$ 为单调增函数，故 $p_1 = 2\Phi(2) - 1 > 2\Phi(1) - 1 = p_2$. 又因为 $p_2 = \Phi(1) - \Phi(-1) > 0.68$，故

$$p_2 = \Phi(1) - \Phi(-1) > 0.5 > \Phi(-1) > \Phi(-1) - \Phi\left(-\frac{7}{3}\right) = p_3.$$

例

设 $f_1(x)$ 为标准正态分布的概率密度，$f_2(x)$ 为 $[-1,3]$ 上均匀分布的概率密度，若

$$f(x)=\begin{cases} af_1(x), & x\leqslant 0 \\ bf_2(x), & x>0 \end{cases} \quad (a>0, b>0)$$

为概率密度，则 a, b 应满足 (　　)

A. $2a+3b=4$.

B. $3a+2b=4$.

C. $a+b=1$.

D. $a+b=2$.

解

要使 $f(x)$ 为概率密度，须满足概率密度的基本性质，故

$$1=\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = a\int_{-\infty}^{0} f_1(x)\mathrm{d}x + b\int_{0}^{+\infty} f_2(x)\mathrm{d}x$$

因 $f_1(x)$ 为标准正态分布的概率密度，故 $\int_{-\infty}^{0} f_1(x)\mathrm{d}x=\frac{1}{2}$. 因 $f_2(x)$ 为 $(-1,3)$ 上均匀分布的概率密度，即

$$f_2(x)=\begin{cases}\frac{1}{4}, & -1\leqslant x\leqslant 3,\\ 0, & \text{其他},\end{cases}$$

故

$$\int_{0}^{+\infty} f_2(x)\mathrm{d}x=\int_{0}^{3}\frac{1}{4}\,\mathrm{d}x=\frac{3}{4}.$$

所以 $1=a\cdot\frac{1}{2}+b\cdot\frac{3}{4}$，即 $2a+3b=4$.

4. $\Gamma-$ 分布

若一个随机变量 X 的密度函数为

$$f(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x>0,\lambda>0,\alpha>0,\\ 0, & x\leqslant 0,\end{cases}$$

则称 X 服从参数为 α, λ 的$\Gamma-$ 分布, 记作$X\sim Ga(\alpha,\lambda)$.

4. $\Gamma-$ 分布

若一个随机变量 X 的密度函数为

$$f(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x>0,\lambda>0,\alpha>0,\\ 0, & x\leqslant 0,\end{cases}$$

则称 X 服从参数为 α, λ 的$\Gamma-$ 分布, 记作$X\sim Ga(\alpha,\lambda)$.其中

$$\Gamma(x)=\int_0^{+\infty}t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t\quad(x>0),\text{或者}$$

$$\Gamma(x)=2\int_0^{+\infty}t^{2x-1}\mathrm{e}^{-t^2}\,\mathrm{d}t\quad(x>0).$$

$$\Gamma(1)=\int_0^{+\infty}\mathrm{e}^{-t}\,\mathrm{d}t=1,\quad \Gamma\left(\frac{1}{2}\right)=2\int_0^{+\infty}\mathrm{e}^{-t^2}\,\mathrm{d}t=\sqrt{\pi}.$$

4. $\Gamma-$ 分布

$\Gamma-$ 分布有两个常用的特例

(1) $\alpha = 1$ 时的 $\Gamma-$ 分布为**指数分布**，即

$$Ga(1, \lambda) = Exp(\lambda).$$

(2) $\alpha = n/2$, $\lambda = 1/2$ 时的 $\Gamma-$ 分布是自由度为 n 的 **χ^2 (卡方) 分布**，记作 $\chi^2(n)$，即

$$Ga\left(\frac{n}{2}, \frac{1}{2}\right) = \chi^2(n).$$

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

1. 随机变量与分布函数

2. 常见分布

3. 随机变量函数的分布
   - 离散型随机变量函数的分布
   - 连续型随机变量函数的分布

本节将讨论的问题是：设 $y = g(x)$ 为连续函数且当随机变量 X 分布已知时，如何求它的函数 $Y = g(X)$ 的分布？

例 (2.3.1)

设随机变量 X 分布列为

X	-1	0	1	2
P	0.2	0.1	0.3	0.4

试求 $Y = X - 1$; $Z = X^2$ 的分布列.

解

先求 $Y = X - 1$ 的分布列. 由于 X 的可能取值为 $-1, 0, 1, 2$, 所以 Y 的可能取值应为 $-2, -1, 0, 2$, 且取这些可能值的概率为

$$\begin{aligned}
P(Y=-2) = P(X-1=-2) = P(X=-1) = 0.2, \\
P(Y=-1) = P(X-1=-1) = P(X=0) = 0.1, \\
P(Y=0) = P(X=1) = 0.3, \\
P(Y=1) = P(X=2) = 0.4.
\end{aligned}$$

所以 $Y = X - 1$ 的分布列为:

Y	-2	-1	0	1
P	0.2	0.1	0.3	0.4

解

接下来求 $Z = X^2$ 的分布列. 由于 X 的可能取值为 $-1, 0, 1, 2$, 所以 Z 的可能取值应为 $0, 1, 4$, 且取这些可能值的概率为

$$P(Z=0) = P(X^2=0) = P(X=0) = 0.1,$$
$$P(Z=1) = P(X^2=1) = P(X=-1) + P(X=1) = 0.2 + 0.3 = 0.5,$$
$$P(Z=4) = P(X=4) = P(X=2) = 0.4.$$

所以 $Z = X^2$ 的分布列为:

Z	0	1	4
P	0.1	0.5	0.4

解

事实上，上述求离散型随机变量分布列也可以通过如下方式求得:

P	0.2	0.1	0.3	0.4
X	-1	0	1	2
$Y=X-1$	-2	-1	0	1
$Z=X^2$	1	0	1	4

只需注意到相同取值的概率应当相加，也可得 Y,Z 的分布列.

1. 随机变量与分布函数
2. 常见分布
3. 随机变量函数的分布
   - 离散型随机变量函数的分布
   - 连续型随机变量函数的分布

下面考虑 X 为连续型随机变量情形. 注意到此时 $Y=g(X)$ 通常也为连续型变量, 问题往往是: 已知 X 的密度函数 $f_X(x)$, 如何求 $Y=g(X)$ 的密度 $f_Y(y)$?

例 (2.3.2)

设随机变量 $X \sim U(0,1)$, $Y=e^X$, 试求 Y 的密度函数 $f_Y(y)$.

下面考虑 X 为连续型随机变量情形. 注意到此时 $Y=g(X)$ 通常也为连续型变量, 问题往往是: 已知 X 的密度函数 $f_X(x)$, 如何求 $Y=g(X)$ 的密度 $f_Y(y)$?

例 (2.3.2)

设随机变量 $X\sim U(0,1)$, $Y=e^X$, 试求 Y 的密度函数 $f_Y(y)$.

解

容易知道 Y 的取值范围 (Y 的密度函数 $f_Y(y)$ 不为 0 的 y 的范围) 是 $1<y<e$. 欲求 Y 的密度, 先求 Y 的分布函数 $F_Y(y)$.

当 $y<1$ 时, $F_Y(y)=P(Y\leqslant y)=P(e^X\leqslant y)=0$;

当 $y\geqslant e$ 时, $F_Y(y)=P(Y\leqslant y)=P(e^X\leqslant y)=P(X\leqslant \ln y)=1$;

当 $1<y<e$ 时,

$F_Y(y)=P(Y\leqslant y)=P(e^X\leqslant y)=P(X\leqslant \ln y)=\int_0^{\ln y}1\mathrm{d}x=\ln y$;

解

于是可得 Y 的分布函数为

$$F_Y(y) = \begin{cases} 0, & y < 1, \\ \ln y, & 1 \leqslant y < e, \\ 1, & y \geqslant e. \end{cases}$$

Y 的密度函数为

$$f_Y(y) = F'_Y(y) = \begin{cases} \dfrac{1}{y}, & 1 \leqslant y < e, \\ 0, & \text{其他}. \end{cases}$$

本例中我们通过先求 $Y = g(X)$ 的分布函数 $F_Y(y)$, 然后求其导数得到概率密度函数 $f_Y(y)$ 的方法称为分布函数法, 它是求解连续型随机变量函数概率密度最基本的方法.

下面我们将利用分布函数法得到一个重要结论.

定理 (2.3.1)

设 X 是连续型随机变量，其密度函数为 $f_X(x)$. $Y=g(X)$ 是另一个随机变量. 若函数 $y=g(x)$ 是严格单调函数，其反函数 $h(y)$ 有连续的导函数，则 $Y=g(X)$ 的密度可以表示为:

$$\boxed{f_Y(y)=\begin{cases} f_X[h(y)]\cdot|h'(y)|, & a<y<b, \\ 0, & \text{其他}. \end{cases}} \tag{10}$$

其中 $a=\min\{g(-\infty),g(+\infty)\}$, $b=\max\{g(-\infty),g(+\infty)\}$.

证明.

不妨设 $y=g(x)$ 是严格单调增函数，这时其反函数 $h(y)$ 也是严格单调增的，且 $h'(y)>0$. 记 $a=\min\{g(-\infty),g(+\infty)\}$, $b=\max\{g(-\infty),g(+\infty)\}$. 这意味着 $y=g(x)$ 仅在区间 (a,b) 上取值，于是

当 $y<a$ 时，$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y)=0$;

当 $y>b$ 时，$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y)=1$;

当 $a\leqslant y\leqslant b$ 时，有

$$\begin{aligned}F_Y(y)&=P(Y\leqslant y)=P(g(X)\leqslant y)\\&=P(X\leqslant h(y))=F_X(h(y))=\int_{-\infty}^{h(y)}f_X(x)\mathrm{d}x.\end{aligned}$$

□

证明.

由上可得 Y 的密度函数为

$$f_Y(y) = F'_Y(y) = \begin{cases} f_X[h(y)] \cdot |h'(y)|, & a < y < b, \\ 0, & \text{其他}. \end{cases}$$

同理可证当 $g(x)$ 为严格单调减时结论也成立. □

注: 若密度函数 $f_X(x)$ 在 $[c, d]$ 之外为零, 则只需假设 $g(x)$ 在 (c, d) 严格单调, 此时取 $a = \min\{g(c), g(d)\}$, $b = \max\{g(c), g(d)\}$, 定理结论依然成立.

例 (2.3.3)

设随机变量 $X \sim N(\mu, \sigma^2)$. 试证明当 $a \neq 0$ 时, X 的线性函数 $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

例 (2.3.3)

设随机变量 $X \sim N(\mu, \sigma^2)$. 试证明当 $a \neq 0$ 时, X 的线性函数 $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

证明.

X 的概率密度

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty,$$

当 $a > 0$ 时, $y = g(x) = ax + b$ 严格单调增, 其反函数

$$x = h(y) = \frac{y-b}{a}.$$

所以 $h'(y) = 1/a$. □

例 (2.3.3)

设随机变量 $X \sim N(\mu, \sigma^2)$. 试证明当 $a \neq 0$ 时, X 的线性函数 $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

证明.

由定理 2.3.1 知: $Y = aX + b$ 的概率密度为

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right), \quad -\infty < x < +\infty,$$

即

$$f_Y(y) = \frac{1}{a\sqrt{2\pi}\sigma} e^{-\frac{[y-(b+a\mu)]^2}{2(a\sigma)^2}}, \quad -\infty < y < +\infty,$$

从而 $Y \sim N(a\mu + b, a^2\sigma^2)$, 结论得证. 同理可证, 当 $a < 0$ 时结论也成立. □

这个例题表明正态随机变量的线性变换依然是正态随机变量.

例 (2.3.4)

由统计物理学知分子运动速度的绝对值 X 服从麦克斯韦 (Maxwell) 分布，其概率密度为

$$f(x)=\begin{cases}\dfrac{4x^2}{a^3\sqrt{\pi}}e^{-\frac{x^2}{a^2}}, & x>0, a>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中 $a>0$ 为常数，求分子动能 $Y=\dfrac{1}{2}mX^2$ (m 为分子质量) 的概率密度.

解

已知 $y = g(x) = \frac{1}{2}mx^2$, $f(x)$ 只在区间 $(0, +\infty)$ 上非零且 $g(x)$ 在此区间恒单调递增, 由 (10) 式, 得 Y 的概率密度为

$$\psi(y) = \begin{cases} \dfrac{4\sqrt{2y}}{m^{3/2}a^3\sqrt{\pi}} e^{-\frac{2y}{ma^2}}, & y > 0, \\ 0, & y \leqslant 0. \end{cases}$$

本章小结

(1) 随机变量的分布函数及其性质: $F(x) = P(X \leqslant x)$.
性质: 单调不减、有界性 $(0 \leqslant F(x) \leqslant 1)$ 和右连续性 $(F(x_0+0) = F(x_0))$.
$P(a < X \leqslant b) = F(b) - F(a)$.

(2) 两类重要的随机变量:
离散型随机变量: $F(x) = \sum_{x_i \leqslant x} P(X = x_i)$.
连续型随机变量: $F(x) = \int_{-\infty}^{x} f(x)\,\mathrm{d}x$.

本章小结

(3) 离散型型随机变量通常用分布列 $p_i = P(X = x_i)$ 来描述, 具有如下性质:

- 非负性: $p_i \geqslant 0$.
- 正则性: $\sum\limits_{i=1}^{\infty} p_i = 1$.

(4) 连续型随机变量常用密度函数 $f(x)$ 来描述, 具有如下性质:

- 非负性: $f(x) \geqslant 0$.
- 正则性: $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$.
- $P(a < X < b) = \int_a^b f(x)\mathrm{d}x$.
- 若 $f(x)$ 在点 x 连续, 则有 $F'(x) = f(x)$.

(5) 常见的离散型随机变量及分布列:

- 单点分布: $P(X=a)=1$.
- 两点分布: $P(X=a)=1-p$, $P(X=b)=p$.
- $0-1$ 分布 $b(1,p)$: $P(X=k)=p^k(1-p)^{1-k}$, $k=0,1$.
- 二项分布 $b(n,p)$: $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$, $k=0,1,\cdots,n$.
- 泊松分布 $P(\lambda)$: $P(X=k)=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$, $k=0,1,2,\cdots$, $\lambda>0$.
- 几何分布 $Ge(p)$: $P(X=k)=pq^{k-1}$, $q=1-p$, $k=1,2,\cdots$.

- **超几何分布** * $h(n, N, M)$:

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\ \ k=0,1,\cdots,r,\ \ r=\min(n,M).$$

- **巴斯卡分布** * $Nb(r, p)$:

$$P(X=k)=\binom{k-1}{r-1}p^r q^{k-r},\ \ r\geqslant 1,\ \ k=r,r+1,r+2,\cdots,$$

$$0<p<1,\ \ q=1-p.$$

(6) 常见的连续型分布及其密度函数.

- 均匀分布 $U(a,b)$:

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b;\\ 0, & \text{其他}.\end{cases}\qquad F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\leqslant x<b,\\ 1, & x\geqslant b.\end{cases}$$

- 指数分布 $Exp(\lambda)$:

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}\qquad F(x)=\begin{cases}1-e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

- 正态分布 $N(\mu,\sigma^2)$:

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\qquad F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt.$$

- 标准正态分布 $N(0,1)$: $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $\quad\Phi(x)=\int_{-\infty}^{x}\dfrac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,dt.$

(7) 随机变量的函数分布问题: 已知随机 X 分布, 求它的函数 $Y = g(X)$ 的分布?

- 分布函数法.
- 公式法.

定理

设 X 是连续型随机变量, 其密度函数为 $f_X(x)$. $Y = g(X)$ 是另一个随机变量. 若函数 $y = g(x)$ 是严格单调函数, 其反函数 $h(y)$ 有连续的导函数, 则 $Y = g(X)$ 的密度可以表示为:

$$f_Y(y) = \begin{cases} f_X[h(y)] \cdot |h'(y)|, & a < y < b, \\ 0, & \text{其他}. \end{cases}$$

其中 $a = \min\{g(-\infty), g(+\infty)\}$, $b = \max\{g(-\infty), g(+\infty)\}$.

例

设随机变量 X 服从正态分布 $N(1,\sigma^2)$，其分布函数为 $F(x)$，则对任意实数 x，有 (　　)

A. $F(x)+F(-x)=1$

B. $F(x+1)+F(x-1)=1$

C. $F(1+x)+F(1-x)=1$

D. $F(1-x)+F(x-1)=1$

例

设随机变量 X 服从正态分布 $N(1,\sigma^2)$，其分布函数为 $F(x)$，则对任意实数 x，有（　　）

A. $F(x)+F(-x)=1$

B. $F(x+1)+F(x-1)=1$

C. $F(1+x)+F(1-x)=1$

D. $F(1-x)+F(x-1)=1$

正确答案：C

例

设随机变量 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，且

$$P\{|X-\mu_1|<1\} > P\{|Y-\mu_2|<1\},$$

则必有 (　　)

A. $\sigma_1 < \sigma_2$

B. $\sigma_1 > \sigma_2$

C. $\mu_1 < \mu_2$

D. $\mu_1 > \mu_2$

例

设随机变量 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，且

$$P\{|X-\mu_1|<1\} > P\{|Y-\mu_2|<1\},$$

则必有（　　）

A. $\sigma_1 < \sigma_2$

B. $\sigma_1 > \sigma_2$

C. $\mu_1 < \mu_2$

D. $\mu_1 > \mu_2$

正确答案：A

例

设 $F_1(x)$ 与 $F_2(x)$ 为两个分布函数，其相应的概率密度 $f_1(x)$ 与 $f_2(x)$ 是连续函数，则必为概率密度的是（　　）.

A. $f_1(x)f_2(x)$

B. $2f_2(x)F_1(x)$

C. $f_1(x)F_2(x)$

D. $f_1(x)F_2(x)+f_2(x)F_1(x)$

例

设 $F_1(x)$ 与 $F_2(x)$ 为两个分布函数，其相应的概率密度 $f_1(x)$ 与 $f_2(x)$ 是连续函数，则必为概率密度的是（　　）.

A. $f_1(x)f_2(x)$

B. $2f_2(x)F_1(x)$

C. $f_1(x)F_2(x)$

D. $f_1(x)F_2(x)+f_2(x)F_1(x)$

正确答案： D

例

设随机变量 X 的概率密度为 $f(x)$, 则下列函数中是概率密度的是 (　　).

A. $f(2x)$
B. $f^2(x)$
C. $2xf(x^2)$
D. $3x^2f(x^3)$

例

设随机变量 X 的概率密度为 $f(x)$, 则下列函数中是概率密度的是 (　　).

A. $f(2x)$
B. $f^2(x)$
C. $2xf(x^2)$
D. $3x^2f(x^3)$

正确答案： D

例

设随机变量 X 的概率密度满足 $f(-x)=f(x)$, $F(x)$ 是 X 的分布函数, 则对任意的实数 a, 下列式子中成立的是 (　　).

A. $F(-a)=1-\int_0^a f(x)\mathrm{d}x$

B. $F(-a)=\frac{1}{2}-\int_0^a f(x)\mathrm{d}x$

C. $F(-a)=F(a)$

D. $F(-a)=2F(a)-1$

例

设随机变量 X 的概率密度满足 $f(-x)=f(x)$，$F(x)$ 是 X 的分布函数，则对任意的实数 a，下列式子中成立的是（　　）.

A. $F(-a)=1-\int_0^a f(x)\mathrm{d}x$

B. $F(-a)=\frac{1}{2}-\int_0^a f(x)\mathrm{d}x$

C. $F(-a)=F(a)$

D. $F(-a)=2F(a)-1$

正确答案： B