第二章：随机变量

2.1 随机变量及分布函数

随机变量的定义

定义 1.1

设 E 是随机试验, Ω 是其样本空间. 如果对每个 $\omega \in \Omega$, 总有一个 $X(\omega)$ 与之对应, 则称 Ω 上的实值函数 $X(\omega)$ 为 E 的一个随机变量.

- **随机变量是一个函数**, 定义在样本空间 Ω 上, 取值在实数轴上. **(概率是定义在 Ω 上吗?)**
- 与通常的函数不同, 它的自变量是随机试验的结果, 其具有随机性. **因此随机变量的取值也具有随机性.** 这是随机变量与一般函数的最大区别.

随机变量的例子

例 1.1

抛掷一枚均匀硬币, 观察币值是否朝上. 若记

$$\{\omega_1\} = \{\text{币值面朝上}\}, \quad \{\omega_2\} = \{\text{币值面朝下}\},$$

则样本空间 $\Omega = \{\omega_1, \omega_2\}$. 于是, 试验有两个可能的结果: ω_1 和 ω_2. 引入随机变量

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

对样本空间不同的元素 ω_1 和 ω_2, 随机变量 $X(\omega)$ 取不同的值 1 和 0. 由于试验结果的出现是随机的, 所以 $X(\omega)$ 的取值也是随机的, 值域为 $\{0, 1\}$.

随机变量的例子

例 1.2

观察一部电梯一年内出现故障的次数. 记录

$$\{\omega_i\} = \{\text{电梯一年内发生 } i \text{ 次故障}\}, i = 0, 1, 2, \cdots ,$$

则样本空间 $\Omega = \{\omega_i,\ i = 0, 1, 2, \cdots\}$. 引入随机变量

$$X(\omega_i) = i,\ i = 0, 1, 2, \cdots ,$$

可使样本空间中的每个元素 ω_i 都与一个非负整数 i 对应, $i = 0, 1, 2, \cdots$. 由于试验结果的出现是随机的, 所以 $X(\omega)$ 的取值也是随机的, 值域为 $\{0, 1, 2, \cdots\}$.

随机变量的例子

例 1.3

对某只灯泡做寿命试验, 观察其使用寿命值. 记 ω 为该只灯泡的使用寿命 (单位: 小时), 则试验的样本空间

$$\Omega = \{\omega,\ 0 \le \omega < +\infty\}.$$

对每个 $\omega \in \Omega$, 可取 $X(\omega) = \omega$ 与之对应, 这样就建立了样本空间 Ω 与区间 $[0, +\infty)$ 之间的对应关系. 由于试验结果的出现是随机的, 所以 $X(\omega)$ 的取值也是随机的, 值域为 $[0, +\infty)$.

随机变量的定义

若将 $X(\omega)$ 简记为 X, 对实数集的每个子集 L, 将 $\{\omega : \omega \in \Omega$ 且 $X(\omega) \in L\}$ 简记为 $\{X \in L\}$, 则可用随机变量 X 描述随机事件.

- 在例 1.1 中, 可用 $\{X = 1\}$ 表示事件 $\{\omega_1\}$, $\{X = 0\}$ 表示事件 $\{\omega_2\}$.
- 在例 1.2 中, 可用 $\{X \leq 5\}$ 表示事件 {电梯在一年内出现故障的次数不超过 5 次}.
- 在例 1.3 中, 可用 $\{X \geq 1000\}$ 表示事件 {灯泡使用寿命大于等于 1000 小时}.

这时, 可称随机变量 $X = 1$ 的概率为 $P\{X = 1\}$; $X \leq 5$ 的概率为 $P\{X \leq 5\}$; $X \geq 1000$ 的概率为 $P\{X \geq 1000\}$.

分布函数的定义

定义 1.2

设 X 是一随机变量, 称函数

$$F(x) = P\{X \leq x\}$$

为 X 的分布函数. 且称 X 服从 $F(x)$, 有时也记为 $F_X(x)$.

性质 1.1

对任意实数 $a < b$, 总有

$$P\{a < X \leq b\} = F(b) - F(a).$$

分布函数的定义

证明: 将事件 $\{a < X \leq b\}$ 分解

$$\{a < X \leq b\} = \{X \leq b\} - \{X \leq a\}.$$

因为 $a < b$, 所以

$$\{X \leq a\} \subset \{X \leq b\}.$$

于是

$$\begin{aligned} P\{a < X \leq b\} &= P\big\{\{X \leq b\} - \{X \leq a\}\big\} \\ &= P\{X \leq b\} - P\{X \leq a\} \\ &= F(b) - F(a). \end{aligned}$$

□

分布函数的例子

例 1.4

投三次硬币 设 H, T 表示一次投币试验中"正面、反面"朝上. 设随机变量 X 表示正面朝上的次数.

样本点	*HHH*	*HHT*	*HTH*	*THH*
X	3	2	2	2
样本点	*TTH*	*THT*	*HTT*	*TTT*
X	1	1	1	0

求 X 的分布函数.

分布函数的例子

解: X 的取值范围为 $\{0, 1, 2, 3\}$, 比样本空间 Ω 简单.

X	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

X 认为 HHT, HTH, THH 都是 $X = 2$ 这个事件. 由上表可知, X 的分布函数为:

$$F_X(x) = \begin{cases} 0, & -\infty < x < 0 \\ 0.125, & 0 \leq x < 1 \\ 0.5, & 1 \leq x < 2 \\ 0.875 & 2 \leq x < 3 \\ 1, & 3 \leq x < \infty \end{cases}$$

□

分布函数的性质

定理 1.1

任一分布函数 $F(x)$ 具有如下三条性质:

(1) 单调性. $F(x)$ 是定义在整个实数轴 $(-\infty, \infty)$ 上的单调非减函数, 即对任意的 $x_1 < x_2$, 都有 $F(x_1) < F(x_2)$.

(2) 有界性. 对任意的 x, 有 $0 \le F(x) \le 1$, 且

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

(3) 右连续性. $F(x)$ 是 x 的右连续函数, 即对任意的 x_0, 有

$$\lim_{x\to-x_0^{+0}} F(x) = F(x_0) \Rightarrow F(x_0+0) = F(x_0).$$

分布函数性质证明

证明: (1) 由分布函数定义, 当 $x_1 < x_2$ 时, 不妨设 $x_2 = x_1 + c, c$ 为任一正数, 则

$$F(x_1) = P(X \le x_1) = P(X \le x_2 - c) < P(X \le x_2) = F(x_2).$$

(2) 由分布函数定义可知, $0 \le F(x) = P(X \le x) \le 1, F(x)$ 有界. 再由性质 (1) 可知, $F(x)$ 单调, 则单调有界函数必存在极限, 即对任意的整数 m 和 n, 有

$$\lim_{x \to -\infty} F(x) = \lim_{m \to -\infty} F(x), \quad \lim_{x \to \infty} F(x) = \lim_{n \to \infty} F(x)$$

都存在.

分布函数性质证明

再由概率的可列可加性知:

$$
\begin{aligned}
1 &= P(-\infty < X < \infty) = P(\bigcup_{i=-\infty}^{\infty} \{i-1 < X \leq i\}) \\
&= \sum_{i=-\infty}^{\infty} P(i-1 < X \leq i) = \lim_{m\to-\infty, n\to\infty} \sum_{i=m}^{n} P(i-1 < X \leq i) \\
&= \lim_{n\to\infty} F(x) - \lim_{m\to-\infty} F(x).
\end{aligned}
$$

两个在 $[0,1]$ 中的数相减等于 1, 只能是

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

分布函数性质证明 (续)

(3) 因为 $F(x)$ 单调有界非降函数, 故任一 x_0 的右极限 $F(x_0+)$ 必存在. 再证其右连续, 需证对任意单调下降的数列 $x_1 > x_2 > \cdots > x_n > \cdots > x_0$, 当 $x_n \to x_0 (n \to \infty)$ 时, 证明 $\lim_{n\to\infty} F(x_n) = F(x_0)$ 成立. 因为

$$\begin{aligned}
F(x_1) - F(x_0) &= P(x_0 < X \leq x_1) = P\Big(\bigcup_{i=1}^{\infty}\{x_{i+1} < X \leq x_i\}\Big) \\
&= \sum_{i=1}^{\infty} P(x_{i+1} < X \leq x_i) = \sum_{i=1}^{\infty}[F(x_i) - F(x_{i+1})] \\
&= \lim_{n\to\infty}[F(x_1) - F(x_n)] = F(x_1) - \lim_{n\to\infty} F(x_n).
\end{aligned}$$

由此,

$$F(x_0) = \lim_{n\to\infty} F(x_n) = F(x_0 + 0).$$

□

分布函数确定随机变量

上述定理是说:

- 如果 $F(x)$ 是随机变量 X 的分布函数, 则 $F(x)$ 满足上述 (1)-(3) 性质.
- 也可以证明, **如果一个函数 $F(x)$ 满足上述 (1)-(3) 性质, 则必存在一个随机变量 X 有分布函数 $F(x)$** (需要引进测度论的知识才能证明).

由分布函数也可确定随机变量:

- 若 $F(x)$ 是连续函数, 则 X 为**连续随机变量**.
- 若 $F(x)$ 是阶梯函数, 则 X 为 **离散随机变量**.
- 若 $F(x)$ 既不是阶梯函数, 又不是连续函数, 则随 X 既不是连续的又不是离散的, 称其为**奇异随机变量**.

分布函数确定随机变量的例子

例 1.5

设 c 是一个常数, 且满足

$$G_c(x) = \begin{cases} 0, & x < 0 \\ 1 - ce^{-3x}, & x \geq 0 \end{cases}$$

当 c 取何值时, $G_c(x)$ 是一个分布函数?

解: 要使得 $G_c(x)$ 是一个分布函数, 需满足性质 (1) – (3). 首先, 无论 c 取何值, 均有 $G_c(x)$ 右连续, 且

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

性质 (2), (3) 均满足.

分布函数确定随机变量: 例子 (续)

此外, 考虑 $G^{'}(x)$,

$$G_c^{'}(x) = \begin{cases} 0, & x < 0 \\ 3ce^{-3x}, & x > 0 \end{cases}$$

分析其导数

- 当 $c < 0$ 时, $G_c^{'}(x)$ 是减函数, 不满足性质 (1).
- 当 $c > 1$ 时, $G_c(0) = 1 - c < 0$, 不满足概率性质.
- 当 $0 \leq c \leq 1$ 时, $G_c^{'}(x)$ 是非降函数, 满足性质 (1).

所以, 当 $0 \leq c \leq 1$ 时, $G_c(x)$ 是分布函数. 注意, 当 $c = 0$ 时, $G_c(x)$ 是阶梯函数; 当 $c = 1$ 时, $G_c(x)$ 是连续函数; 当 $0 < c < 1$ 时, $G_c(x)$ 既不是阶梯函数又不是连续函数, 即 $G_c(x)$ 对应的随机变量是奇异的. □

2.2 离散随机变量

离散随机变量: 定义

定义 2.1

设 X 是一个离散随机变量, 如果 X 的所有可能取值是 $x_1, x_2, \cdots, x_n, \cdots$, 则称 X 取 x_i 的概率

$$p_i = f(x_i) = P(X = x_i), i = 1, 2, \cdots, n, \cdots$$

为 X 的**概率质量函数** (*probability mass function, p.m.f.*) 或概率分布列, 简称分布列. 其中 $p_k \geq 0$, 且 $\sum_{i=1}^{\infty} p_i = 1$. 分布列可以用表格表示:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

离散随机变量: 例子

例 2.1

电子线路中装有两个并联继电器. 假设这两个继电器是否接通具有随机性, 且彼此独立. 已知每个继电器接通的概率为 0.8, 记 X 为线路中接通的继电器的个数. 求 (1) X 的概率分布; (2) 线路接通的概率.

解: (1) X 取 0, 1, 2 三个值. 记 $A_i = \{\text{第 } i \text{ 个继电器接通}\}$, $i = 1, 2$. 注意到两个继电器是否接通是相互独立的, 于是 A_1 与 A_2 相互独立, 且 $P(A_1) = P(A_2) = 0.8$. $\{X = 0\}$ 表示两个继电器都没接通, 则

$$P\{X = 0\} = P(\bar{A}_1\bar{A}_2) = P(\bar{A}_1) \cdot P(\bar{A}_2) = 0.2 \times 0.2 = 0.04.$$

离散随机变量: 例子

类似地, 可得

$$
\begin{aligned}
P\{X=1\} &= P(A_1\bar{A}_2 \cup \bar{A}_1 A_2) = P(A_1\bar{A}_2) + P(\bar{A}_1 A_2) \\
&= P(A_1)P(\bar{A}_2) + P(\bar{A}_1)P(A_2) \\
&= 0.8 \times 0.2 + 0.2 \times 0.8 = 0.32, \\
P\{X=2\} &= P(A_1A_2) = P(A_1)P(A_2) = 0.8 \times 0.8 = 0.64.
\end{aligned}
$$

(2) 因为此电路并联, 所以只要一个继电器接通, 整个线路就接通. 于是,

$$
P\{X \geq 1\} = P\{X=1\} + P\{X=2\} = 0.32 + 0.64 = 0.96.
$$

例 2.2

3 个人随机进入编号为 1, 2, 3, 4 的 4 个房间中, 每个房间容纳的人数不限. 设 X 代表有人的房间的最大房号，求 X 的分布律.

解: 不易直接求 $P(X=k), k=1,2,3,4$, 先求

$$P(X \le k) = k^3/4^3.$$

故

$$P(X=4) = P(X \le 4) - P(X \le 3) = \frac{4^3 - 3^3}{4^3} = \frac{37}{64};$$

离散随机变量: 例子

类似地,

$$P(X=3)=P(X\le 3)-P(X\le 2)=\frac{3^3-2^3}{4^3}=\frac{19}{64};$$

$$P(X=2)=P(X\le 2)-P(X\le 1)=\frac{2^3-1^3}{4^3}=\frac{7}{64};$$

$$P(X=1)=\frac{1^3}{4^3}=\frac{1}{64}.$$

由此, X 的分布列为

X	1	2	3	4
P	$\frac{1}{64}$	$\frac{7}{64}$	$\frac{19}{64}$	$\frac{37}{64}$

$\square$

两点分布

定义 2.2

若随机变量 X 只可能取 0 或 1 两个值, 其概率分布为

$$P\{X=1\}=p, \quad P\{X=0\}=q,$$

其中 $0<p<1, q=1-p$, 则称 X 服从参数为 p 的 **两点分布**. 或 $(0-1)$ 分布, 记为 $X \sim B(1,p)$.

对任何只有两种可能结果的随机试验 E, 令 $\Omega=\{\omega_1,\omega_2\}$ 表示其样本空间, 总可以在 Ω 上定义一个服从两点分布的随机变量

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

来描述随机试验的结果.

两点分布: 例子

例 2.3

200 件产品中, 有 *196* 件是正品, *4* 件是次品, 今从中随机抽取一件, 若规定

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

则

$$P\{X=1\} = \frac{196}{200} = 0.98, \quad P\{X=0\} = \frac{4}{200} = 0.02.$$

于是, X 服从参数为 *0.98* 的两点分布, 即 $X \sim B(1, 0.98)$.

定义 2.3

设试验 E 有两种可能的结果 A 与 $\bar{A}$, 且

$$P(A) = p\ (0 < p < 1),$$

将试验 E 独立重复 n 次, 称为 n 重伯努利试验. 设 X 是 n 重贝努利试验中事件 A 发生的次数, 则它的分布律为:

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

则称 X 服从参数为 n, p 的二项分布, 记为 $X \sim B(n, p)$.

二项分布

- 容易验证:

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

二项概率恰好是n **次二项式** $[p+(1-p)]^n$ **的展开式中的第** $k+1$ **项**, 这正是其名称由来.

- 伯努利试验是从现实许多的随机现象中抽象出来的一种基本概率模型.
- 在一批产品的质量检查中, 若检查结果分为" 合格" 和" 不合格" 两种, 采用**放回抽样**, 则检查 n 件产品就是 n 次伯努利试验.

二项分布: 例子

例 2.4

某特效药的临床有效率为 0.95, 今有 10 人服用, 问至少有 8 人治愈的概率是多少?

解: 设 X 为 10 人中被治愈的人数, 则 $X \sim B(10, 0.95)$, 且

$$
\begin{aligned}
P(X \geq 8) &= P(X=8) + P(X=9) + P(X=10) \\
&= \binom{10}{8} 0.95^8 0.05^2 + \binom{10}{9} 0.95^9 0.05 + \binom{10}{10} 0.95^{10} \\
&\approx 0.0746 + 0.3151 + 0.5987 = 0.9884.
\end{aligned}
$$

10 人中至少有 8 人治愈的概率为 0.9884. □

二项分布: 例子

例 2.5

一种 40 瓦的灯泡. 规定其使用寿命超过 2000 小时的微正品, 否则为次品. 已知有很大一批这样的灯泡, 其次品率为 0.2. 现从该灯泡中随机抽取 20 只做寿命试验, 问这 20 只灯泡中恰有 k 只是次品的概率是多少?

解: 这是一个不放回抽样问题, 但由于样本量很大, 抽出灯泡的数量相对于灯泡总数相对较小, 可当作放回抽样处理. 将观测一只灯泡的使用寿命是否超过 2000 小时看成一次试验, 观测 20 只灯泡相当于做 20 次伯努利试验. 记随机变量 X 为 20 只灯泡中次品的只数, 则 $X \sim B(20, 0.2)$, 此时

$$P\{X = k\} = \binom{20}{k} (0.2)^k (0.8)^{20-k}, k = 0, 1, \cdots, 20.$$

二项分布: 与两点分布的关系

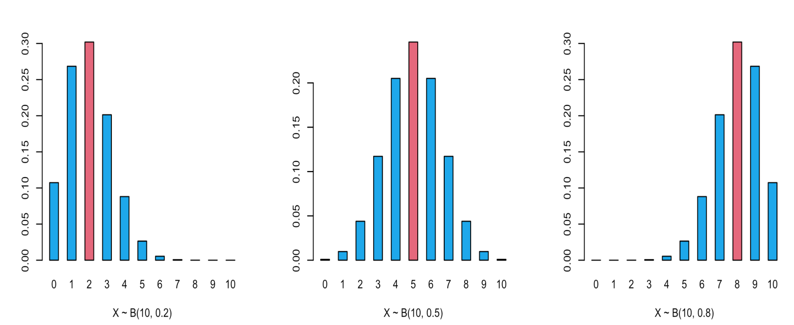
- 两点分布用来描述**一次伯努利试验**中成功 (记为 A) 的次数.
- 如果进行 n **个相同的、独立的伯努利试验**, 记第 i 个伯努利试验中 A 出现的次数为 X_i $(i = 1, 2, \cdots, n)$.
- 由于每个伯努利试验是独立的, 故其产生的 n 个随机变量 X_i 也是独立的, 且 $X_i \sim B(1, p)$, 则其和

$$X = X_1 + X_2 + \cdots + X_n$$

  就是 n 重伯努利试验中 A 出现的总次数, 它服从二项分布 $B(n, p)$.
- **服从二项分布的随机变量总可分解成 n 个独立同分布为两点分布的随机变量之和.**

二项分布: n 固定, p 改变

考察当 n 固定 ($n = 10$) 时, 比较不同 p 值 ($p = 0.2, 0.5, 0.8$) 对分布的影响.



从图中可以看出: (1) 位于均值 np 附近的概率较大; (2) 随着 p 的增加, 分布的峰逐渐右移 (右偏 → 对称 → 左偏).

泊松分布: 定义

定义 2.4

泊松 (*Poisson*) 分布的概率分布列是

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

其中参数 $\lambda>0$, 记为 $X\sim\mathcal{P}(\lambda)$.

$\star$ 容易验证,

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}e^{\lambda}=1.$$

常见的例子: (1) 一天内, 来到某商场的顾客数. (2) 在单位时间内, 一电路受到外界电磁波的冲击次数. (3) 在一定时期内, 某种放射性物质放射出来的 α-粒子数.

泊松分布: 例子

例 2.6

某床单厂生产的每条床单上含有疵点的个数 X 服从参数 $\lambda = 1.5$ 的泊松分布. 质量检验部分规定: 床单上无疵点或只有一个疵点的为一等品, 有 2 个疵点到 4 个疵点的为二等品, 5 个以上疵点的为次品. 试求该床单厂生产的床单为一等品、二等品和次品的概率.

解: 由 $X \sim \mathcal{P}(1.5)$ 及概率的可加性, 得

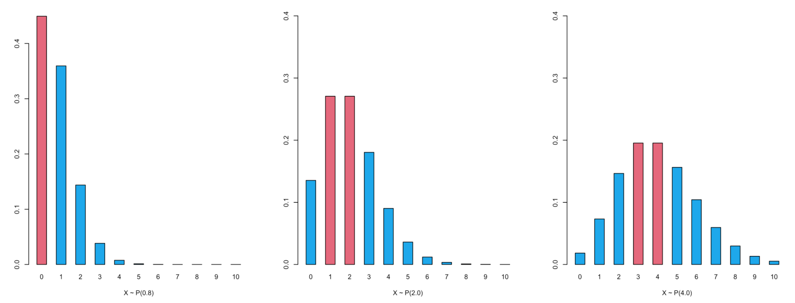
$$P\{\text{床单为一等品}\} = P\{X \leq 1\} = e^{-1.5}\left(\frac{1.5^0}{0!} + \frac{1.5^1}{1!}\right) = 0.558,$$

$$P\{\text{床单为二等品}\} = P\{2 \leq X \leq 4\} = 0.424,$$

$$\begin{aligned} P\{\text{床单为次品}\} &= 1 - P\{\text{床单为一等品}\} - P\{\text{床单为二等品}\} \\ &= 0.018. \end{aligned}$$

泊松分布: λ 变化

比较不同 λ 值 ($\lambda = 0.8, 2.0, 4.0$) 对分布的影响.



从图中可以看出: (1) 位于 λ 附近的概率较大; (2) 随着 λ 的增加, 分布的峰逐渐对称.

泊松分布: 与二项分布关系

定理 2.1

在 n 重伯努利试验中, 记事件 A 在一次试验中发生的概率为 p_n (与试验次数 n 有关), 如果当 $n \to \infty$ 时, 有 $np_n \to \lambda$, 则

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

证明: 设 $np_n = \lambda_n, p_n = \lambda_n/n$, 可得

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \Big(\frac{\lambda_n}{n}\Big)^k \Big(1-\frac{\lambda_n}{n}\Big)^{n-k}$$

$$= \frac{\lambda^k}{k!} \Big(1-\frac{1}{n}\Big)\Big(1-\frac{2}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big)\Big(1-\frac{\lambda_n}{n}\Big)^{n-k}.$$

泊松分布: 与二项分布关系 (续)

对固定的 k, 有

$$\lim_{n\to\infty} \lambda_n = \lambda,$$

$$\lim_{n\to\infty} \left(1 - \frac{\lambda_n}{n}\right)^{n-k} = e^{-\lambda},$$

$$\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) = 1.$$

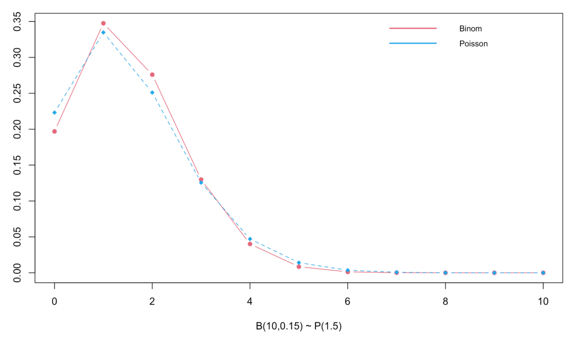
从而

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

由于泊松定理在 $np_n \to \lambda$ 的条件下获得, 在实际计算中, 当 n 很大, p 很小, 乘积 $\lambda = np$ 大小适中时, 二项分布可利用泊松分布近似. □

泊松分布: 与二项分布关系 (图)

图中**红色线**表示二项分布 $B(10, 0.15)$, **蓝色线**表示泊松分布 $\mathcal{P}(1.5)$.



显然从图中可以看出, 当 $np = 1.5 = \lambda$ 时, 近似程度很好.

泊松分布: 泊松定理的例子

例 2.7

某公司生产的一种产品, 根据历史记录知, 该产品的次品率为 0.01, 问该种产品 300 件中次品数大于 5 的概率是多少?

解: 把检验每件产品看作一次试验, 它有两个可能的结果: $A = \{正品\}$ 和 $\bar{A} = \{次品\}$. 检验 300 件产品做 300 次伯努利试验. 记 X 表示次品数, 则 $X \sim B(300, 0.01) \sim \mathcal{P}(3)$. 得到

$$P\{X > 5\} = \sum_{k=6}^{300} \binom{300}{k} 0.01^k 0.99^{300-k}$$
$$\approx \sum_{k=6}^{300} \frac{e^{-3}}{k!} 3^k \approx \sum_{k=6}^{\infty} \frac{e^{-3}}{k!} 3^k.$$

查表得 $P\{X > 5\} \approx 0.08$. $\square$

例 2.8

甲每天收到的电子邮件数服从泊松分布 $\mathcal{P}(\lambda)$, 且每封电子邮件被过滤掉的概率是 0.2. 试问:

(a) 当有 n 封电子邮件发给甲, 计算他见到 k 封的概率 h_k;

(b) 计算甲每天见到电子邮件数的分布;

(c) 已知甲看到了自己的 k 封电子邮件, 计算他有 m 封被过滤掉的概率;

(d) 甲每天见到的邮件数和被滤掉的邮件数是否独立.

解: (a) 当有 n 封电子邮件发给甲, 他见到其中 k 封为事件 $Y = k|X = n$, n 为固定值, 则

$$h_k = P(Y = k|X = n) = \binom{n}{k} 0.8^k 0.2^{n-k}, k = 0, 1, \cdots, n.$$

泊松分布: 泊松与二项混合分布的例子 (续)

(b) 由全概率公式

$$
\begin{aligned}
P(Y=k) &= \sum_{n=k}^{\infty} P(Y=k|X=n)P(X=n) \\
&= \sum_{n=k}^{\infty} \binom{n}{k} 0.8^k 0.2^{n-k} \frac{\lambda^n}{n!} e^{-\lambda} \\
&= \frac{0.8^k}{k!} e^{-\lambda} \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} 0.2^{n-k} \frac{\lambda^n}{n!} \\
&= \frac{(0.8\lambda)^k}{k!} e^{-\lambda} \sum_{n=k}^{\infty} \frac{(0.2\lambda)^{n-k}}{(n-k)!} \\
&= \frac{(0.8\lambda)^k}{k!} e^{-\lambda} e^{0.2\lambda} = \frac{(0.8\lambda)^k}{k!} e^{-0.8\lambda},
\end{aligned}
$$

$k = 0, 1, 2, \cdots$. 因此 $Y \sim \mathcal{P}(0.8\lambda)$.

泊松分布: 泊松与二项混合分布的例子 (续)

同理，若设随机变量 M 表示每天过滤掉的邮件数, 则 $M \sim \mathcal{P}(0.2\lambda)$.

(c) 甲看到 k 封电子邮件, 有 m 封被过滤掉的事件为 $X = m + k | Y = k$，这时 k 是固定值, m 是变化的, 则

$$P(X = m + k | Y = k) = \frac{P(Y = k | X = m + k) P(X = m + k)}{P(Y = k)}$$

$$= \frac{\binom{m+k}{k} 0.8^k 0.2^m \frac{\lambda^{(m+k)}}{(m+k)!} e^{-\lambda}}{\frac{(0.8\lambda)^k}{k!} e^{-0.8\lambda}} = \frac{(0.2\lambda)^m}{m!} e^{-0.2\lambda},$$

$m = 0, 1, 2, \cdots$.

(d) 由 (b), (c) 间可以看出, $P(M = m | Y = k) = P(M = m)$, 即甲每天见到的邮件数和被滤掉的邮件数是独立的. □

2.3 连续型随机变量

连续随机变量

如何定义连续随机变量 X 的 $f_X(x) = P(X = x)$?

对任一 $\varepsilon > 0$,

$$P(X = x) \leq P(x - \varepsilon < X \leq x) = F_X(x) - F_X(x - \varepsilon).$$

由于 $F_{(}x)$ 在 x 处连续, 则

$$P(X = x) \leq \lim_{\varepsilon \to 0} F_X(x) - F_X(x - \varepsilon) = 0.$$

由概率非负性可知, $P(X = x) = 0, \forall x \in \mathbb{R}$. 由离散随机变量的分布函数

$$F_X(x) = \sum_{k \leq x} f_X(k)$$

可类推**连续随机变量的分布函数是将求和变成求积分**.

连续随机变量: 定义

定义 3.1

设随机变量 X 的分布函数为 $F(x)$, 如果存在实数轴上的一个非负可积函数 $f(x)$, 使得对任意实数 x, 有

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

则称 $f(x)$ 为 X 的概率密度函数 (*probability density function, p.d.f.*), 简称密度函数或密度. 在 $F(x)$ 导数存在的点上有 $F'(x) = f(x)$.

⋆ 从定义来看, 可能有不存在密度函数的情况, 要严格定义需利用测度论知识: 连续随机变量 X 存在密度函数, 当且仅当其分布函数是**绝对连续**的.

连续随机变量: 性质

性质 3.1

密度函数 $f(x)$ 满足下列性质:

(1) 非负性. $f(x) \geq 0$.

(2) 正则性. $\int_{-\infty}^{\infty} f(x)dx = 1$.

上述性质是密度函数必须具有的性质, 也可由上述两条性质判断某个函数是否为密度函数.

性质 3.2

若非负函数 $f(x)$ 是随机变量 X 的概率密度函数, 则 X 取值于任一区间 $(a, b]$ 的概率可表示为

$$P\{a < X \leq b\} = \int_a^b f(x)dx.$$

连续随机变量: 例子

例 3.1 (通过密度函数求分布函数)

设 $f_X(x)$ 为随机变量 X 的密度函数, 且

$$f_X(x) = \begin{cases} 1+x, & -1 \le x < 0 \\ 1-x, & 0 \le x < 1 \\ 0, & \text{其他} \end{cases}$$

求 X 的分布函数 $F_X(x)$.

解: 分情况讨论: (1) $x < -1$ 时, $F_X(x) = 0$.

(2) 当 $-1 \le x < 0$ 时,

$$F_X(x) = \int_{-\infty}^{x} f(t)dt = \int_{-1}^{x} (1+t)dt = t+\frac{t^2}{2}\Big|_{t=-1}^{x} = \frac{x^2}{2}+x+\frac{1}{2}.$$

连续随机变量: 例子 (续)

(3) 当 $0 \le x < 1$ 时,

$$F_X(x) = \int_{-\infty}^{x} f(t)dt = \int_{-1}^{0}(1+t)dt + \int_{0}^{x}(1-t)dt = -\frac{x^2}{2} + x + \frac{1}{2}.$$

(4) 当 $x \ge 1$ 时,

$$F_X(x) = \int_{-\infty}^{x} f(t)dt = \int_{-1}^{0}(1+t)dt + \int_{0}^{x}(1-t)dt = \frac{1}{2} + \frac{1}{2} = 1.$$

由此, 分布函数为

$$F_X(x) = \begin{cases} 0 & x < -1 \\ \frac{x^2}{2} + x + \frac{1}{2}, & -1 \le x < 0 \\ -\frac{x^2}{2} + x + \frac{1}{2}, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

离散与连续随机变量的差别

- 离散随机变量 X 在可能取值的点 $x_1, x_2, \cdots, x_n, \cdots$ 上的概率都不为 0, 但连续随机变量 Y 在 $(-\infty, \infty)$ 上任意点 a 的概率恒为 0 ($P(Y=a) = \int_a^a f(x)dx = 0$).
- **不可能事件的概率为 0, 但概率为 0 的事件不一定是不可能事件**.
- **必然事件的概率为 1, 但概率为 1 的事件不一定为必然事件.**
- 对于连续随机变量 Y, 事件 $\{a \le Y \le b\}$ 剔除 $x=a$ 或 $x=b$ 不影响其概率:

$$\begin{aligned} P(a \le Y \le b) &= P(a < Y \le b) \\ &= P(a \le Y < b) = P(a < Y < b). \end{aligned}$$

- 但离散随机变量 X 此性质不存在, 必须**点点计较**.

离散与连续随机变量的差别 (续)

- 由于有时改变密度函数 $f(x)$ 的值并不改变积分值, **故一个连续随机变量的密度函数可能不唯一**. 如

$$f_1(x) = \begin{cases} \frac{1}{a}, & 0 \le x \le a \\ 0, & \text{其他} \end{cases} \qquad f_2(x) = \begin{cases} \frac{1}{a}, & 0 < x < a \\ 0, & \text{其他} \end{cases}$$

  它们都是 $(0, a)$ 上**均匀分布**的密度函数.

- 注意 $f_1(x)$ 与 $f_2(x)$ 虽然不同, 但

$$P(f_1(x) \neq f_2(x)) = P(X = 0) + P(X = a) = 0.$$

  称这两个函数**几乎处处相等**(almost everywhere, a.s.). 其意义是: 剔除概率为 0 点后的事件这两个函数相等. 记为 $f_1(x) = f_2(x)$ $a.s.$ P.

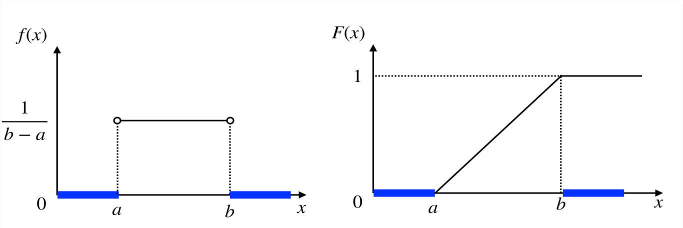
均匀分布: 定义

定义 3.2

设随机变量 X 的密度函数, 分布函数为

$$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b,\\ 0, & \text{其他}.\end{cases}\qquad F(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\le x<b\\ 1, & x\ge b.\end{cases}$$

则称 X 服从区间 (a,b) 上的 **均匀分布**, 记作 $X\sim U(a,b)$.



伽马分布: 伽马函数

定义 3.3

称以下函数

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

为**伽马函数**, 其中 $\alpha > 0$.

性质 3.3

伽马函数的性质如下:

(1) $\Gamma(1) = 1, \Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(2) $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$.

(3) 若 n 为自然数, $\Gamma(n + 1) = n\Gamma(n) = n!$.

证明: (1) 当 $\alpha = 1$ 时, $\int_0^\infty e^{-x}dx = 1$. 当 $\alpha = 1/2$ 时,

$$\begin{aligned}\Gamma(1/2) &= \int_0^\infty x^{-\frac{1}{2}}e^{-x}dx = \int_0^\infty (t^2)^{-1/2}e^{-t^2}2tdt \quad (x = t^2, t > 0)\\ &= 2\int_0^\infty e^{-t^2}dt\end{aligned}$$

令 $I = \int_0^\infty e^{-t^2}dt$, 首先考虑

$$\begin{aligned}I^2 &= \int_0^\infty e^{-x^2}dx \int_0^\infty e^{-y^2}dy = \int_0^\infty \int_0^\infty e^{-(x^2+y^2)/2}dxdy\\ &= \int_0^{\pi/2}\int_0^\infty e^{-r^2}rdrd\theta = \frac{\pi}{4}.\end{aligned}$$

所以, $I = \sqrt{\pi}/2 \Rightarrow \Gamma(1/2) = \sqrt{\pi}$.

伽马分布: 伽马函数

(2) 等式左边

$$\begin{aligned}\Gamma(\alpha+1) &= \int_0^\infty x^\alpha e^{-x} dx \\ &= -\int_0^\infty x^\alpha de^{-x} \\ &= -\left(e^{-x}x^\alpha\Big|_0^\infty - \int_0^\infty e^{-x} dx^\alpha\right) \quad (\text{分部积分}) \\ &= \alpha \int_0^\infty x^{\alpha-1} e^{-x} dx = \alpha\Gamma(\alpha).\end{aligned}$$

(3) 令 $\alpha = n \in N$ 时,

$$\Gamma(n+1) = n\Gamma(n) = n!$$

$\square$

伽马分布: 定义

定义 3.4

设随机变量 X 的密度函数为

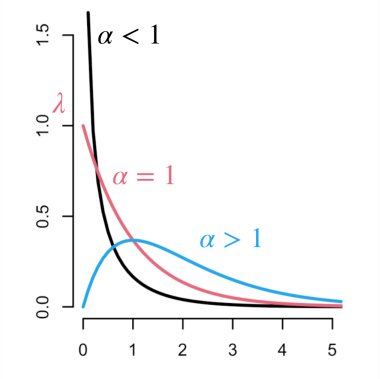
$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

称 X 服从 伽马分布, 记作 $X \sim Ga(\alpha, \lambda)$. 其中 $\alpha > 0$ 为形状参数, $\lambda > 0$ 为尺度参数.

伽马分布: 定义 (续)

固定 λ, 不同 α 的伽马分布的密度曲线图:

- 当 $0 < \alpha < 1$ 时, $f(x)$ 严格下降函数, $x = 0$ 处有奇异点.
- 当 $\alpha = 1$ 时, $f(x)$ 严格下降函数, $x = 0$ 处 $f(x) = 0$.
- 当 $\alpha \geq 0, f(x)$ 单峰函数, α 越大越对称.



伽马分布: 两个特例

伽马分布有两个特例值得注意:

(1) 当 $\alpha = 1$ 时, **伽马分布 $Ga(1, \lambda)$ 就是指数分布 $E(\lambda)$.**

(2) 当 $\alpha = n/2, \lambda = 1/2$ 时, 伽马分布是自由度为 n 的**卡方分布**, 记作 $\chi^2(n)$, 即**伽马分布 $Ga(n/2, 1/2) = \chi^2(n)$**, 其密度函数为

$$f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, x \geq 0.$$

这里 n 是 χ^2 分布的唯一参数.

指数分布: 定义

定义 3.5

设随机变量 X 的密度函数, 分布函数为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0, & x < 0. \end{cases}$$

称 X 服从**指数分布**, 记作 $X \sim E(\lambda)$, $\lambda > 0$.

指数分布是一种**偏态分布**, 常被用作各种“寿命”分布. 容易验证:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} \lambda e^{-\lambda x} dx = 1.$$

指数分布: 无记忆性

定理 3.1

如果随机变量 $X \sim E(\lambda)$, 则对任意 $s > 0, t > 0$, 有

$$P(X > s + t | X > s) = P(X > t).$$

证明: 因为 $X \sim E(\lambda)$, 所以 $P(X > s) = e^{-\lambda s}, s > 0$. 又因为

$$\{X > s + t\} \subset \{X > s\}.$$

于是有

$$\begin{aligned} P(X > s + t | X > s) &= \frac{P(X > s + t)}{P(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = P(X > t). \end{aligned}$$

$\square$

指数分布: 例子

例 3.2

设某电子管的使用寿命 X (单位: 小时) 服从参数 $\lambda = 0.0002$ 的指数分布, 求电子管的使用寿命超过 *3000* 小时的概率.

解: 因为 $X \sim E(0.0002)$, 则

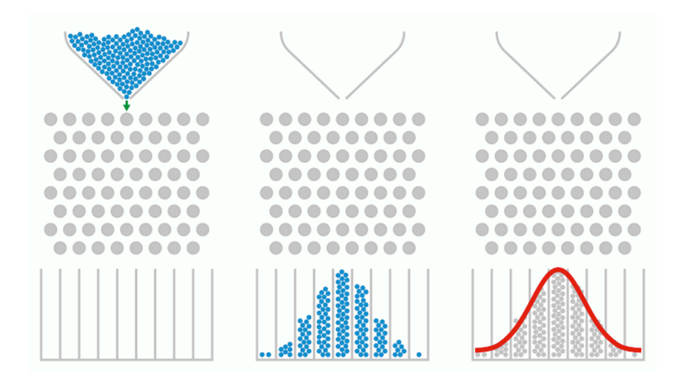
$$
\begin{aligned}
P\{X > 3000\} &= \int_{3000}^{\infty} 0.0002 \times e^{-0.0002x} dx \\
&= e^{-0.6} = 0.5488,
\end{aligned}
$$

即电子管的使用寿命超过 3000 小时的概率为 0.5488. $\square$

正态分布: 引入

正态分布反映**无数微小误差因素共同影响的总效果.** 连续量的测量误差一般服从正态分布.

高尔顿板试验:



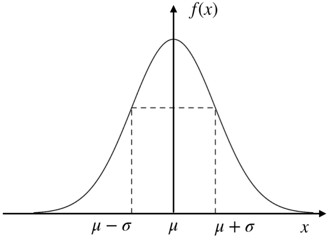
正态分布: 定义

定义 3.6

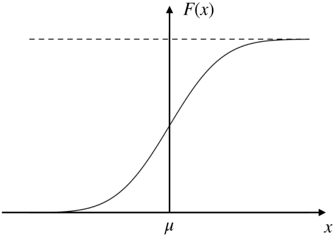
设随机变量 X 的密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2\sigma^2}(x-\mu)^2 \right\}, -\infty < x < \infty,$$

称 X 服从正态分布, 记作 $X \sim \mathcal{N}(\mu, \sigma^2)$, 参数为 $\mu, \sigma(\sigma > 0)$.



密度函数



分布函数

正态分布: 定义

显然 $f(x) \geq 0$, 它在 $(-\infty, \infty)$ 上的积分为 1 吗?

- 令 $t = (x-\mu)/\sigma$, 可得

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \\ &= \frac{1}{2\sqrt{\pi}} \left(\int_{-\infty}^{0} e^{-\frac{t^2}{2}} dt + \int_{0}^{\infty} e^{-\frac{t^2}{2}} dt \right)\end{aligned}$$

- 再令 $u = t^2/2, t < 0$, $v = t^2/2, t > 0$, 可得

$$\begin{aligned}&\frac{1}{2\sqrt{\pi}} \left(\int_{\infty}^{0} -u^{-1/2} e^{-u} du + \int_{0}^{\infty} v^{-1/2} e^{-v} dw \right) \\ =&\frac{1}{2\sqrt{\pi}} \times 2 \int_{0}^{\infty} x^{-1/2} e^{-x} dx = \frac{\Gamma(1/2)}{\sqrt{\pi}} = 1.\end{aligned}$$

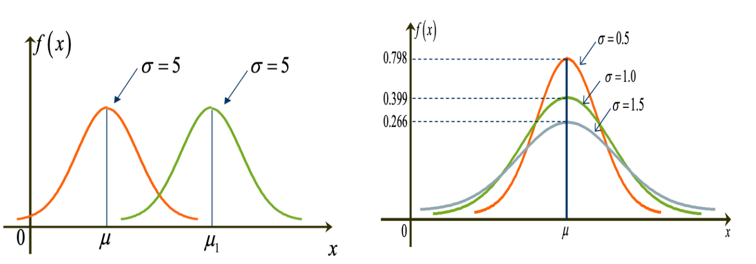
正态分布: 定义 (续)

- $f(x)$ 是一条**钟形曲线**, 中间高, 两边低.
  - ⋆ 左右关于 $x = \mu$ 对称 μ 是正态分布的中心.
  - ⋆ 在 $x = \mu$ 附近取值的可能性大, 取值为 $1/(\sqrt{2\pi}\sigma)$.
  - ⋆ 当 $|x| \to \infty$ 时, 曲线以 x 轴为渐近线.
- 正态分布的分布函数为

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \exp\left\{ -\frac{(t-\mu)^2}{2\sigma^2} \right\} dt.$$

它是一条光滑上升的 S 形曲线.

正态分布: 位置-尺度参数



- 左图: 如果固定 σ, 改变 μ 的值, 则图形沿 x 轴平移, 而不改变形状. 此时称 μ 为**位置参数**.
- 右图: 如果固定 μ, 改变 σ 的值, 则分布的位置不改变, 但σ **的值越小, 曲线呈" 高而瘦", 分布较集中;** σ **的值越大, 曲线呈" 矮而胖", 分布较分散.** 此时称 σ 为**尺度参数**.

正态分布: 标准正态分布

定义 3.7

称 $\mu=0,\sigma=1$ 时的正态分布 $\mathcal{N}(0,1)$ 为**标准正态分布**. 记标准正态变量为 U, 则其密度 $\phi(u)$, 分布函数 $\Phi(u)$ 分别为:

$$\phi(u)=\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}},\ \Phi(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u}e^{-\frac{u^2}{2}}dt.$$

其中 $-\infty<\mu<\infty$.

性质 3.4

$\Phi(U)$ 有如下性质:

(1) $\Phi(-u)=1-\Phi(u)$.

(2) $\Phi(U>u)=1-\Phi(u)$.

(3) $P(a<U<b)=\Phi(b)-\Phi(a)$.

正态分布: 正态变量的标准化

定理 3.2

若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则

$$U = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1).$$

证明: 记 X 和 U 的分布函数分别为 $F_X(x)$ 和 $F_U(u)$, 则

$$F_U(u) = P(U \leq u) = P(\frac{X-u}{\sigma} \leq u) = P(X \leq \mu+\sigma u) = F_X(\mu+\sigma u).$$

因为正态分布函数严格且处处可导, 记 X 与 U 的密度函数分别为 $f_X(x)$ 与 $f_U(u)$, 则

$$f_U(u) = \frac{d}{du}F_X(\mu+\sigma u) = f_X(\mu+\sigma u)\cdot\sigma = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}.$$

因此, $U = (X-\mu)/\sigma \sim \mathcal{N}(0,1)$. □

正态分布: 正态变量的标准化

定理 3.3

若随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则对任意的 $a, b(a < b)$

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

证明: 由概率密度函数的性质及变量变换 $(x-\mu)/\sigma = u$ 得

$$\begin{aligned} P\{a < X \leq b\} &= \frac{1}{\sqrt{2\pi}\sigma}\int_a^b \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} e^{-u^2/2}dy \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(b-\mu)/\sigma} e^{-u^2/2}dy - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(a-\mu)/\sigma} e^{-u^2/2}dy \\ &= \Phi((b-\mu)/\sigma) - \Phi((a-\mu)/\sigma). \end{aligned}$$

正态分布: 正态变量的标准化

定理 3.4

因为 $\Phi(-\infty)=0$ 和 $\Phi(\infty)=1$, 则

$$P\{X\leq b\}=P\{-\infty<X\leq b\}=\Phi\left(\frac{b-\mu}{\sigma}\right);$$

$$P\{X>a\}=P\{a<X<\infty\}=1-\Phi\left(\frac{a-\mu}{\sigma}\right).$$

例 3.3

已知某台机器生产的螺栓长度 X (单位: 厘米) 服从参数 $\mu=10.05, \sigma=0.06$ 的正态分布. 规定螺栓长度在 10.05 ± 0.12 内为合格品, 试求螺栓为合格品的概率.

正态分布: 正态变量的标准化例子

解: 设 $X \sim N(10.05, 0.06^2)$, 记

$$a = 10.05 - 0.12, \quad b = 10.05 + 0.12,$$

则

$$\begin{aligned} P\{a \le X \le b\} &= P\{a < X \le b\} \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \\ &= \Phi(2) - \Phi(-2) \\ &= \Phi(2) - [1 - \Phi(2)] \\ &= 2\Phi(2) - 1 \\ &= 2 \times 0.9772 - 1 = 0.9544. \end{aligned}$$

$\square$

正态分布: 正态变量的标准化例子

例 3.4

假设某地区成年男性的身高 (单位: 厘米) $X \sim N(170, 7.69^2)$, 求该地区成年男性的身高超过 175 厘米的概率.

解: 设 $X \sim N(170, 7.69^2)$, 且 $\{X > 175\}$ 表示该地区成年男性的身高超过 175 厘米, 计算

$$\begin{aligned} P\{X > 175\} &= P\{175 < X < \infty\} \\ &= 1 - \Phi\left(\frac{175 - 170}{7.69}\right) \\ &= 1 - \Phi(0.65) \\ &= 1 - 0.7422 = 0.2758. \end{aligned}$$

$\square$

正态分布: α 分位数

定义 3.8

若 $X \sim \mathcal{N}(0,1)$, 若 z_α 满足条件:

$$P(X \leq z_\alpha) = \Phi(z_\alpha) = \alpha, 0 < \alpha < 1.$$

则称 z_α 为标准正态分布的 α 下分位点.

性质 3.5

标准正态分布的下 α 分位点有如下性质:

(1) $z_\alpha = -z_{1-\alpha}$.

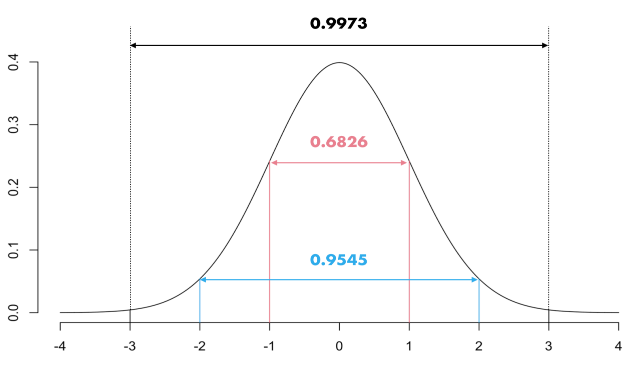
(2) $P(X > z_{1-\alpha}) = \alpha$.

(3) $P(|X| > z_{1-\alpha/2}) = \alpha, P(|X| \leq z_{1-\alpha/2}) = \alpha$.

正态分布: α 分位数 (续)

- 当 $k = 1, 2, 3$ 时

$$P(|X - \mu| < \sigma) = 2\Phi(1) - 1 = 0.6826,$$
$$P(|X - \mu| < 2\sigma) = 2\Phi(2) - 1 = 0.9545,$$
$$P(|X - \mu| < 3\sigma) = 2\Phi(3) - 1 = 0.9973.$$



正态分布: 例子

例 3.5

设 $X \sim \mathcal{N}(\mu, \sigma^2)$, 试用分位数表示常数 a, b.

(1) $\mu = 0, \sigma = 1, P(-X < a) = 0.025.$

(2) $\mu = 1, \sigma = 2, P(|X - 1| \leq b) = 0.75.$

解: (1) 由分位数定义可知:

$$P(-X < a) = P(X > -a) = 1 - P(X \leq -a) = 1 - \Phi(\frac{-a-0}{1}) = 0.025$$

$$\Rightarrow \Phi(\frac{-a-0}{1}) = 0.975 \Rightarrow a = -z_{0.975} = z_{0.025}.$$

(2) 同理,

$$P(|X - 1| \leq b) = P(\left|\frac{X-1}{2}\right| \leq \frac{b}{2}) = 2\Phi(\frac{b}{2}) - 1 = 0.75$$

$$\Rightarrow \Phi(\frac{b}{2}) = 0.875 \Rightarrow b = 2z_{0.875}.$$

正态分布: 例子

例 3.6

设中国人身高服从均值为 *1.75* 米, 方差为 *0.05* 米的正态分布. 试问公共汽车门至少多高才能保证需要低头通过的人数不超过 *0.5%?*

解: 设随机变量 X 为中国人身高, 则 $X \sim \mathcal{N}(1.75, 0.05^2)$. 由题意可知, 设门至少需要 h 米, 则

$$P(X \geq h) \leq 0.005 \Rightarrow P(X \leq h) \geq 0.995.$$

计算

$$\begin{aligned}
P(X \leq h) &= P(\frac{X-1.75}{0.05} \leq \frac{h-1.75}{0.05}) \geq 0.995 \\
\Rightarrow & \Phi(\frac{h-1.75}{0.05}) \geq 0.995 \Rightarrow \frac{h-1.75}{0.05} \geq z_{0.995} = 2.58 \\
\Rightarrow & h \geq 0.05 \times 2.58 + 1.75 = 1.879.
\end{aligned}$$

2.4 随机变量函数的分布

分布函数法

若 X 为随机变量, 则 $Y = g(X)$ 也是随机变量. 对任意的集合 A,

$$P(Y \in A) = P(g(X) \in A), X \in \mathcal{X}, Y \in \mathcal{Y}.$$

考虑 $y = g(x)$, $g(x)$ **是一个从 X 的样本空间 $\mathcal{X}$ 到新的随机变量 Y 的样本空间 $\mathcal{Y}$ 的一个映射:**

$$g(x) : \mathcal{X} \to \mathcal{Y} \Rightarrow g^{-1}(A) = \{x \in \mathcal{X} : g(x) \in A\}.$$

此时, 对任意集合 $A \subset \mathcal{Y}$,

$$\begin{aligned} P(Y \in A) = P(g(X) \in A) &= P(\{x \in \mathcal{X} : g(x) \in A\}) \\ &= P(X \in g^{-1}(A)). \end{aligned}$$

将按照分布函数定义求解的方法称为**分布函数法**.

离散随机变量函数的分布

若 X 是离散随机变量, 则 $\mathcal{X}$ 是可数的. 随机变量 Y 的样本空间 $\mathcal{Y}=\{y:y=g(x),x\in\mathcal{X}\}$ 也是可数的, 由此 Y 是一个离散随机变量.

$$P(Y=y)=\sum_{x\in g^{-1}(y)}P(X=x),y\in\mathcal{Y}.$$

其中, 若 $y\notin\mathcal{Y},f_Y(y)=0$.

例 4.1

设 X 具有以下的分布律:

X	-1	0	1	2
P	0.2	0.3	0.1	0.4

求 $Y=X^2-1$ 的分布列.

离散随机变量函数的分布

解: 求 $Y = 2X - 1$ 的分布列为

X	-1	0	1	2
$X^2 - 1$	0	-1	0	3
P	0.2	0.3	0.1	0.4

再整理可知:

X	-1 和 1	0	2
$X^2 - 1$	0	-1	3
P	0.3	0.3	0.4

$\square$

离散随机变量函数的分布例子

例 4.2

在应用上认为: 单位时间内, 一个地区发生火灾的次数服从泊松分布. 设某城市一个月内发生火灾的次数 $X \sim \mathcal{P}(5)$, 试求随机变量 $Y = |X-5|$ 的概率分布.

解: 由 X 所有可能取值的集合为 $\{0, 1, 2, \cdots\}$, 其对应的概率分布为

$$P\{X = k\} = \frac{e^{-5}5^k}{k!}, \quad k = 0, 1, 2, \cdots$$

及 $Y = |X-5|$ 可能的取值为 $\{0, 1, 2, \cdots\}$,

- 当 $0 < i \leq 5$ 时, $k = 5 + i$ 和 $k = 5 - i$ 使得 $|k-5| = i$;
- 当 $i = 0$ 或 $i \geq 6$ 时, 只有 $k = 5 + i$ 使得 $|k-5| = i$.

离散随机变量函数的分布例子

于是, 随机变量 Y 取值为 i 的概率,

(1) 当 $i = 1, 2, 3, 4, 5$ 时,

$$q_i = P\{Y = i\} = \left[\frac{5^{5-i}}{(5-i)!} + \frac{5^{5+i}}{(5+i)!}\right]e^{-5},$$

(2) 当 $i = 0, 6, 7, \cdots$ 时,

$$q_i = P\{Y = i\} = \frac{5^{5+i}}{(5+i)!}e^{-5}.$$

$\square$

连续随机变量函数的分布

若 $X, Y = g(X)$ 是连续随机变量, 则 Y 的分布函数为

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(g(X) \le y) \\ &= P(\{x \in \mathcal{X} : g(x) \le y\}) \\ &= \int_{\{x \in \mathcal{X} : g(x) \le y\}} f_X(x)dx. \end{aligned}$$

利用**变上限积分求导**: 当 $f(x)$ 连续时,

$$\frac{d}{dx}\int_a^x f(t)dt = f(x);$$

$$\frac{d}{dx}\int_a^{u(x)} f(t)dt = f(u(x))u'(x);$$

$$\frac{d}{dx}\int_{u(x)}^{v(x)} f(t)dt = f[v(x)]v'(x) - f[u(x)]u'(x).$$

例 4.3

设随机变量 X 有密度函数

$$f_X(x) = \begin{cases} |x|, & -1 < x < 1, \\ 0, & \text{其他}. \end{cases}$$

求随机变量 $Y = 2X + 1$ 的概率密度函数.

解: 利用分布函数法

$$\begin{aligned} F_Y(y) &= P\{Y \le y\} = P\{2X + 1 \le y\} \\ &= P\left\{X \le \frac{y-1}{2}\right\} = \int_{-\infty}^{\frac{y-1}{2}} f_X(x)dx, \end{aligned}$$

再由 $f_Y(y) = F'_Y(y)$, 得

$$
\begin{aligned}
f_Y(y) &= f_X\left(\frac{y-1}{2}\right) \cdot \left(\frac{y-1}{2}\right)' \\
&= \begin{cases} \left|\frac{y-1}{2}\right| \times \frac{1}{2}, & -1 < \frac{y-1}{2} < 1, \\ 0, & \text{其他}, \end{cases} \\
&= \begin{cases} \frac{|y-1|}{4}, & -1 < y < 3, \\ 0, & \text{其他}. \end{cases}
\end{aligned}
$$

$\square$

例 4.4 (正态分布及卡方分布的关系)

令 $X \sim \mathcal{N}(0,1), Y = X^2$. 求 Y 的概率密度函数.

解: 利用分布函数法

$$F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\} = P\{-\sqrt{y} \le X \le \sqrt{y}\}$$
$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

则

$$f_Y(y) = F_Y'(y) = \frac{1}{\sqrt{2\pi}} e^{-y/2} \frac{1}{2\sqrt{y}} - \frac{1}{\sqrt{2\pi}} e^{-y/2} (-\frac{1}{2\sqrt{y}})$$
$$= \frac{1}{2^{1/2}\Gamma(1/2)} \cdot y^{-1/2} e^{-y/2}, \quad y > 0,$$

即 $Y \sim \chi^2(1)$. □

例 4.5

设随机变量 $X \sim U[-\pi/2, \pi/2]$, 求 $Y = 2\cos X$ 的概率密度函数.

解: 利用分布函数法 (1) 当 $y < 0$ 时, $F_Y(y) = 0$. (2) 当 $y \geq 2$ 时, $F_Y(y) = 1$. (3) 当 $0 \leq y < 2$ 时,

$$\begin{aligned} F_Y(y)P\{Y \leq y\} &= P\{2\cos X \leq y\} \\ &= P\{-\frac{\pi}{2} \leq X \leq -\arccos\frac{y}{2} \text{ 或 } \arccos\frac{y}{2} \leq X \leq \frac{\pi}{2}\} \\ &= 1 - \frac{2}{\pi}\arccos\frac{y}{2}. \end{aligned}$$

得到

$$f_Y(y) = \frac{2}{\pi\sqrt{4-y^2}}, \quad 0 \leq y < 2.$$

例 4.6

某台机床加工一种零件，假定加工后零件的长度 X (单位：毫米) 服从正态分布 $N(52.5, 0.1^2)$，当零件长度在质量规格限 52.5 ± 0.3 范围内时，零件为合格品，售出后可获利润 a 元；否则不合格品，损失为 b 元 (视利润 $-b$ 元). 若用 $Y = g(X)$ 表示零件的利润，则

$$Y = g(X) = \begin{cases} a, & 52.2 \leq X \leq 52.8, \\ -b & ,\text{其他}. \end{cases}$$

Y 可能取值为 a 和 $-b$，则

$$P\{Y = a\} = P\{52.2 \leq X \leq 52.8\} = 0.9974.$$

$$P\{Y = -b\} = 1 - P\{Y = a\} = 0.0026.$$

公式法: 背景

考虑 X 和 $Y = g(X)$ 的样本空间:

$$\mathcal{X} = \{x : f_X(x) > 0\},\ \mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\}.$$

(1) 当 $g(x)$ 为**严格单增函数**时 $(u > v \Rightarrow g(u) > g(v))$,

$$\begin{aligned}\{x \in \mathcal{X} : g(x) \le y\} &= \{x \in \mathcal{X} : g^{-1}(g(x)) \le g^{-1}(y)\} \\ &= \{x \in \mathcal{X} : x \le g^{-1}(y)\}.\end{aligned}$$

此时

$$\begin{aligned}F_Y(y) &= \int_{\{x \in \mathcal{X} : x \le g^{-1}(y)\}} f_X(x)dx \\ &= \int_{-\infty}^{g^{-1}(y)} f_X(x)dx = F_X(g^{-1}(y)).\end{aligned}$$

公式法: 背景

(2) 当 $g(x)$ 为**严格单减函数**时 $(u < v \Rightarrow g(u) > g(v))$,

$$
\begin{aligned}
\{x \in \mathcal{X} : g(x) \le y\} &= \{x \in \mathcal{X} : g^{-1}(g(x)) \ge g^{-1}(y)\} \\
&= \{x \in \mathcal{X} : x \ge g^{-1}(y)\}.
\end{aligned}
$$

此时

$$
\begin{aligned}
F_Y(y) &= \int_{\{x \in \mathcal{X} : x \le g^{-1}(y)\}} f_X(x) dx \\
&= \int_{g^{-1}(y)}^{\infty} f_X(x) dx = 1 - F_X(g^{-1}(y)).
\end{aligned}
$$

公式法: 背景

定理 4.1

令 $X \sim F_X(x), Y = g(X), x \in \mathcal{X}, y \in \mathcal{Y}$,

a. 若 $g(\cdot)$ 是 $\mathcal{X}$ 上的一个单增函数, 则

$$F_Y(y) = F_X(g^{-1}(y)), y \in \mathcal{Y}.$$

b. 若 $g(\cdot)$ 是 $\mathcal{X}$ 上的一个单减函数, 则

$$F_Y(y) = 1 - F_X(g^{-1}(y)), y \in \mathcal{Y}.$$

公式法: 背景例子

例 4.7 (均匀分布-指数分布关系)

设 $X \sim U(0,1)$, 即 $F_X(x) = x$. 试求 $Y = g(X) = -\log X$ 的分布函数.

解: 因为 $Y = -\log X$ 是单减函数, 且

$$g^{-1}(y) = e^{-y}, y > 0.$$

利用上述公式:

$$F_Y(y) = 1 - F_X(g^{-1}(y)) = 1 - F_X(e^{-y}) = 1 - e^{-y}, y > 0.$$

显然, $y \leq 0$ 时, $F_Y(y) = 0$, 即 $Y \sim Exp(1)$. $\square$

定理 4.2

令 $X \sim f_X(x), Y = g(X), x \in \mathcal{X}, y \in \mathcal{Y}$, 其中 $g(\cdot)$ 是单调函数. 设 $f_X(x)$ 在 $\mathcal{X}$ 上连续, $g^{-1}(y)$ 在 $\mathcal{Y}$ 上有连续导数, 则 Y 的概率密度为

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right|, & y \in \mathcal{Y} \\ 0, & \text{其他}. \end{cases}$$

证明: 根据定理 4.1 和求导链式法则

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y), & \text{若 } g(\cdot) \text{ 是单增的;} \\ -f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y), & \text{若 } g(\cdot) \text{ 是单减的.} \end{cases}$$

□

公式法: 例子

例 4.8 (对数正态分布)

证明: 设 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $Y = e^X$ 的概率密度函数为

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}y\sigma} \exp\left\{ -\frac{(\ln y-\mu)^2}{2\sigma^2} \right\}, & y > 0 \\ 0, & y \le 0. \end{cases}$$

解: $y = e^x$ 是严格单增函数, 其反函数为 $g^{-1}(y) = \ln y$, 当 $y > 0$ 时, 由公式法

$$f_Y(y) = f_X(\ln y)\frac{1}{y} = \frac{1}{\sqrt{2\pi}y\sigma} \exp\left\{ -\frac{(\ln y - \mu)^2}{2\sigma^2} \right\}.$$

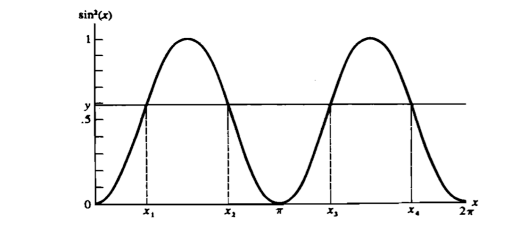
此分布称为**对数正态分布**, 记为 $\mathcal{LN}(\mu, \sigma^2)$, 其中 μ 成为对数均值, σ^2 称为对数方差. □

综合练习

例 4.9

设 $X \sim U(0, 2\pi)$, 试利用分布函数法和公式法分别求 $Y = \sin^2(X)$ 的分布.

解: 考虑 $Y = \sin^2(X)$ 的图像:



$$P(Y \le y) = P(X \le x_1) + P(x_2 \le X \le x_3) + P(X \ge x_4)$$

$$P(X \le x_1) = P(X \ge x_4),\ P(x_2 \le X \le x_3) = 2P(x_2 \le X \le \pi).$$

综合练习: 分布函数法

利用**分布函数法**, 当 $0 \le y \le 1$ 时,

$$
\begin{aligned}
P(Y \le y) &= 2P(0 \le X \le x_1) + 2P(x_2 \le X \le \pi) \\
&= 2\left(\int_0^{\arcsin\sqrt{y}} \frac{1}{2\pi}dx + \int_{\pi-\arcsin\sqrt{y}}^{\pi} \frac{1}{2\pi}dx\right) \\
&= \frac{1}{\pi}(\arcsin\sqrt{y} + \pi - \pi + \arcsin\sqrt{y}) = \frac{2\arcsin\sqrt{y}}{\pi}.
\end{aligned}
$$

求其密度为

$$
f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{2}{\pi}\frac{1}{\sqrt{1-y}}\frac{1}{2\sqrt{y}} = \frac{1}{\pi\sqrt{y(1-y)}},
$$

其中, $0 \le y \le 1$.

综合练习: 公式法

利用**公式法**, 将支撑集 $(0, 2\pi)$ 分割成

$$A_0 = \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}, 2\pi\}, A_1 = (0, \frac{\pi}{2}), A_2 = (\pi, \frac{3}{2}\pi);$$

$$A_3 = (\frac{\pi}{2}, \pi), A_4 = (\frac{3}{2}\pi, 2\pi).$$

则

$$\begin{aligned} f_Y(y) &= 2f_X(\arcsin\sqrt{y})\frac{d}{dy}(\arcsin\sqrt{y}) \\ &\quad - 2f_X(\pi - \arcsin\sqrt{y})\frac{d}{dy}(\pi - \arcsin\sqrt{y}) \\ &= 2\left(\frac{1}{2\pi}\frac{1}{\sqrt{1-y}}\frac{1}{2\sqrt{y}} - \frac{1}{2\pi}\frac{-1}{\sqrt{1-y}}\frac{1}{2\sqrt{y}}\right) \\ &= \frac{1}{\pi\sqrt{y(1-y)}}, 0 \le y \le 1. \end{aligned}$$