第二章 随机变量及其分布

Instructor: 郝壮

haozhuang@buaa.edu.cn
School of Economics and Management
Beihang University

November 1, 2021

第二章随机变量及其分布

- §2.1 随机变量及其分布
- §2.2 随机变量的数学期望
- §2.3 随机变量的方差与标准差
- §2.4 常用离散分布
- §2.5 常用连续分布
- §2.6 随机变量函数的分布
- §2.7 分布的其他特征数

§2.1 随机变量及其分布

为了进行定量的数学处理, 需要引入表示随机现象数量化结果的变量

(1) 掷一颗骰子, 出现的点数 X: $1, 2, ..., 6.$

(2) n个产品中的不合格品个数 Y: $0, 1, 2, ..., n$

(3) 某商场一天内来的顾客数 Z: $0, 1, 2, ...$

(4) 某种型号电视机的寿命 $T : [0, +\infty)$

(5) 掷一枚硬币正反面 $X : 0, 1$

定义2.1.1

设 $\Omega = \{\omega\}$为某随机现象的样本空间, 称定义在Ω上的实值函数$X = X(\omega)$为随机变量.

- 随机变量常用大写字母X, Y, Z等表示, 其取值(随机变量的实现)用小写字母x, y, z等表示.
- 随机变量$X = X(\omega)$的自变量可以不是数, 但因变量一定是实数
- 区别于一般变量, 随机变量$X = X(\omega)$对应着分布函数(描述不同取值的概率)

注意点 (1)

(1) 随机变量$X = X(\omega)$是样本点ω的函数, 其定义域为Ω, 其值域为$R = (-\infty, +\infty)$

eg. 若 X表示掷一颗骰子出现的点数, 则 $\{X = 1.5\}$ 是不可能事件.

(2) 若 X为随机变量, 则$\{X = k\}, \{a < X \le b\}, ...$均为随机事件. 即 $\{a < X \le b\} = \{\omega : a < X(\omega) \le b\} \subset \boldsymbol{\Omega}$

注意点 (2)

(3) 注意以下一些表达式:

$$\{X = k\} = \{X \leq k\} - \{X < k\}$$
$$\{a < X \leq b\} = \{X \leq b\} - \{X \leq a\}$$
$$\{X > b\} = \Omega - \{X \leq b\}$$

(4) 同一样本空间可以定义不同的随机变量.

两类随机变量

- 若随机变量 X 可能取值的个数为有限个或可列个, 则称 X 为**离散随机变量**.
- 若随机变量 X 的可能取值充满某个区间(a, b), 则称 X 为**连续随机变量**.

哪些为离散随机变量? 哪些为连续随机变量?

(1) 掷一颗骰子, 出现的点数 X: $1, 2, ..., 6$.
(2) n个产品中的不合格品个数 Y: $0, 1, 2, ..., n$
(3) 某商场一天内来的顾客数 Z: $0, 1, 2, ...$
(4) 某种型号电视机的寿命 T : $[0, +\infty)$
(5) 掷一枚硬币正反面 X : $0, 1$

2.1.2 随机变量的分布函数

有没有分布是区分一般变量与随机变量的主要标志.

定义2.1.2 设X为一个随机变量, 对任意实数x, 称$F(x) = P(X \leq x)$ 为 X 的**分布函数**(cumulative distribution function, c.d.f.). 且称X 服从 $F(x)$, 记为 $X \sim F(x)$. ($F(x)$ 亦可写为 $F_X(x)$).

- $F(x)$ 亦可写为 $F_X(x)$
- 任一随机变量X(离散的或连续的)都有一个分布函数

定理2.1.1

(1) $F(x)$ 单调非减: 对任意$x_1 < x_2$, $F(x_1) \leq F(x_2)$
(2) 有界:$0 \leq F(x) \leq 1$,

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

(3) 右连续(continuous from right): 对任意的x_0, 有

$$\lim_{x \to x_0^+} F(x) = F(x_0),$$
$$F(x_0 + 0) = F(x_0)$$

证明见教材.

分布函数的基本性质

利用随机变量X的分布函数$F(x)$可以表示有关X的各种事件的概率

- $P(a < X \leq b) = F(b) - F(a)$
- $P(X = a) = F(a) - F(a-0)$
- $P(X \geq b) = 1 - F(b-0)$
- $P(X > b) = 1 - F(b)$
- $P(a < X < b) = F(b-0) - F(a)$
- $P(a \leq X \leq b) = F(b) - F(a-0)$
- $P(a \leq X < b) = F(b-0) - F(a-0)$

当$F(x)$在a与b处连续时, 有 $F(a-0) = F(a), F(b-0) = F(b)$

2.1.3 离散随机变量的分布列(distribution sequence)

离散随机变量可以用分布列来表示其分布.

定义2.1.3 设离散随机变量X的可能取值为:

$$x_1, x_2, \ldots, x_n, \ldots$$

称 $p_i = P(X = x_i), i = 1, 2, \ldots$ 为X的**概率分布列**, 简称**分布列**, 记为$X \sim \{p_i\}$.

分布列也可用表格形式表示:

X	x_1	x_2	...	x_n	...
P	p_1	p_2	...	p_n	...

分布列的基本性质

(1) $p_i \geq 0$, (非负性)

(2) $\sum\limits_{i=1} p_i = 1$ (正则性)

注意点 (1)

求离散随机变量的分布列应注意:

(1) 确定随机变量的所有可能取值;

(2) 计算每个取值点的概率.

注意点 (2)

由离散随机变量的**分布列**, 可写出离散随机变量的**分布函数**

$$F(x) = \sum_{x_i \le x} p(x_i)$$

对离散随机变量的分布函数应注意:

(1) $F(x)$是递增的阶梯函数;

(2) 其间断点均为右连续的;

(3) 其间断点即为X的可能取值点;

(4) 其间断点的跳跃高度是对应的概率值.

例2.1.1 已知随机变量X 的分布列如下:

X	0	1	2
P	$1/3$	$1/6$	$1/2$

求X 的分布函数.

例2.1.1 已知随机变量X 的分布列如下:

X	0	1	2
P	1/3	1/6	1/2

求X 的分布函数.

解:

$$F(x)=\begin{cases}0, & x<0.\\ 1/3, & 0\leqslant x<1.\\ 1/2, & 1\leqslant x<2.\\ 1, & x\geqslant 2.\end{cases}$$

例2.1.1 已知随机变量X 的分布列如下:

X	0	1	2
P	1/3	1/6	1/2

求X 的分布函数.

解:

$$F(x)=\begin{cases}0, & x<0.\\ 1/3, & 0\leqslant x<1.\\ 1/2, & 1\leqslant x<2.\\ 1, & x\geqslant 2.\end{cases}$$

练习: 画出其分布函数.

例2.1.2 已知 X 的分布函数如下,求 X 的分布列.

$$F(x)=\begin{cases}0, & x<0.\\ 0.4, & 0\leqslant x<1.\\ 0.8, & 1\leqslant x<2.\\ 1, & x\geqslant 2.\end{cases}$$

解:

例2.1.2 已知 X 的分布函数如下,求 X 的分布列.

$$F(x)=\begin{cases}0, & x<0.\\ 0.4, & 0\leqslant x<1.\\ 0.8, & 1\leqslant x<2.\\ 1, & x\geqslant 2.\end{cases}$$

解:

X	0	1	2
P	0.4	0.4	0.2

2.1.4 连续随机变量的密度函数

- 连续随机变量X的可能取值充满某个区间 (a, b).
- 因为对连续随机变量X, 有$P(X = x) = 0$, 所以无法仿离散随机变量用 $P(X = x)$来描述连续随机变量X的分布. 引入概率密度函数!
- 注意离散随机变量与连续随机变量的差别.

2.1.4 连续随机变量的密度函数

定义2.1.4 设随机变量X的分布函数为$F(x)$, 若存在非负可积函数 $p(x)$, 满足:

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

则称 X 为连续随机变量, 称 $p(x)$为**概率密度函数**(probability density function, p.d.f.), 简称密度函数.

密度函数的基本性质

满足(1) (2)的函数都可以看成某个连续随机变量的概率密度函数.

(1) $p(x) \geq 0$, (非负性)

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

注意点(1)

(1) $P(a \le X \le b) = \int_a^b p(x)dx$

(2) $F(x)$ 是 $(-\infty, +\infty)$ 上的连续函数

(3) $P(X = x) = F(x) - F(x-0) = 0$

注意点(2)

(4) 因为$P(X=a)=0, \forall a$,

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

不可能事件概率为0, 但概率为0的事件不一定是不可能事件

(5) 当$F(x)$ 在x点可导时, $p(x)=F'(x)$

当$F(x)$ 在x点不可导时, 可令$p(x)=0$ (令$p(x)=$任意有限正数也并不改变$F(x)$, 因为积分为0).

离散型概率分布 vs 连续型概率分布

离散型	连续型
分布列: $p_n = P(X = x_n)$	密度函数:$X \sim p(x)$
唯一	不唯一: 对单点, $p(x)$可取任意值
$F(x) = \sum\limits_{x_i \le x} P(X = x_i)$	$F(x) = \int_{-\infty}^{x} p(t)dt$
$F(a+0) = F(a)$	$p(a < X \le b) = F(b) - F(a)$
单点概率可不为0	$P(X = a) = 0, \forall a$
$F(x)$ 为阶梯函数 当X可取a时 $F(a-0) \neq F(a)$	$F(x)$ 为连续函数 $F(a-0) = F(a)$

课堂练习 (3')

$$X \sim p(x) = \begin{cases} ke^{-3x}, & x > 0. \\ 0, & x \le 0. \end{cases}$$

求 (1)常数k. (2) $F(x)$.

解:

课堂练习 (3')

$$X \sim p(x) = \begin{cases} ke^{-3x}, & x > 0. \\ 0, & x \leq 0. \end{cases}$$

求 (1)常数k. (2) $F(x)$.

解:

(1) 对非0取值求积分

$$\int_0^{+\infty} ke^{-3x}dx = 1 \Rightarrow -\frac{k}{3}e^{-3x}\Big|_0^{+\infty} = 1 \Rightarrow k = 3$$

(2)

$$F(x) = \begin{cases} 0, & x \leq 0. \\ 1 - e^{-3x}, & x > 0. \end{cases}$$

课堂练习 (2')

$$X \sim p(x) = \begin{cases} 1+x, & -1 \le x < 0. \\ 1-x, & 0 \le x < 1. \\ 0, & \text{otherwise.} \end{cases}$$

求 $F(x)$.

解:

$$X \sim p(x) = \begin{cases} 1+x, & -1 \le x < 0. \\ 1-x, & 0 \le x < 1. \\ 0, & \text{otherwise}. \end{cases}$$

求 $F(x)$.

解:

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

课堂练习 (5’)

设X与Y同分布, X的密度为

$$p(x)=\begin{cases}\frac{3}{8}x^2, & 0<x<2.\\ 0, & 其他\ .\end{cases}$$

已知事件 $A=\{X>a\}$和 $B=\{Y>a\}$独立, 且 $P(A\cup B)=3/4$, 求常数a.

解:

课堂练习 (5')

设X与Y同分布, X的密度为

$$p(x) = \begin{cases} \frac{3}{8}x^2, & 0 < x < 2. \\ 0, & \text{其他} . \end{cases}$$

已知事件 $A = \{X > a\}$和 $B = \{Y > a\}$独立, 且 $P(A \cup B) = 3/4$, 求常数a.

解: 因为 $P(A) = P(B)$, 且由A, B 独立, 得

$$P(A \cup B) = P(A) + P(B) - P(A)P(B) = 2P(A) - [P(A)]^2 = 3/4$$

从中解得$P(A) = 1/2$. 由此得$0 < a < 2$. 因此

$1/2 = P(A) = P(X > a) = \int_a^2 \frac{3}{8}x^2 dx = 1 - \frac{a^3}{8}$ 从中解

得$a = \sqrt[3]{4}$

课堂练习(2')

设 $X \sim p(x)$, 且 $p(-x) = p(x)$, $F(x)$是X的分布函数, 则对任意实数 $a > 0$, 有( )

① $F(-a) = 1 - \int_0^a p(x)dx$

② $F(-a) = \dfrac{1}{2} - \int_0^a p(x)dx$

③ $F(-a) = F(a)$

④ $F(-a) = 2F(a) - 1$

解:

设 $X \sim p(x)$, 且 $p(-x) = p(x)$, $F(x)$是X的分布函数, 则对任意实数 $a > 0$, 有( )

① $F(-a) = 1 - \int_0^a p(x)dx$

② $F(-a) = \frac{1}{2} - \int_0^a p(x)dx$

③ $F(-a) = F(a)$

④ $F(-a) = 2F(a) - 1$

解: ②

第4次作业

习题2.1中题目2, 7, 9, 10, 11, 12, 14, 15, 17, 18, 19

数学期望是随机变量取值的加权平均.

Motivation: 分赌本问题 Pascal, 1623-1662

甲乙两赌徒赌技相同, 各出赌注50元. 无平局, 谁先赢3局, 则获全部赌注. 当甲赢2局, 乙赢1局时, 中止了赌博. 问如何分赌本?

两种分法

1. 按已赌局数分:

   则甲分总赌本的$2/3$, 乙分总赌本的$1/3$

2. 按已赌局数和再赌下去的“期望” 分:

   因为再赌两局必分胜负, 共四种情况:
   甲甲, 甲乙, 乙甲, 乙乙
   所以甲分总赌本的$3/4$, 乙分总赌本的$1/4$

若按已赌局数和再赌下去的“期望” 分, 则甲的所得 X 是一个可能取值为0或100 的随机变量, 其分布列为:

X	0	100
P	$1/4$	$3/4$

甲的“期望” 所得是:$0 \times 1/4 + 100 \times 3/4 = 75$.

2.2.2 数学期望的定义

定义2.2.1 (离散情况) 设离散随机变量X的分布列为$p(x_i)=P(X=x_i), i=1,2,\ldots,n,\ldots$. 若级数$\sum_{i=1}^{\infty}x_i p(x_i)$绝对收敛(absolute convergence), 即

$$\sum_{i=1}^{\infty}|x_i|p(x_i)<\infty$$

则称

$$E(x)=\sum_{i=1}^{\infty}x_i p(x_i)$$

为X的**数学期望**, 或称**该分布的数学期望**, 简称**期望或均值**.

若级数$\sum_{i=1}^{\infty}|x_i|p(x_i)$不收敛, 则称$X$的数学期望不存在.

定义2.2.2 (连续情况) 设连续随机变量X的密度函数为$p(x)$, 若积分$\int_{-\infty}^{+\infty} xp(x)dx$绝对收敛, 即

$$\int_{-\infty}^{+\infty} |x|p(x)dx < \infty$$

则称

$$E(x) = \int_{-\infty}^{+\infty} xp(x)dx$$

为X 的**数学期望**, 或称为该分布$p(x)$的数学期望, 简称**期望**或均值.

若$\int_{-\infty}^{+\infty} |x|p(x)dx$不收敛, 则称$X$的数学期望不存在.

例:离散随机变量的期望

例2.2.1

X	-1	0	1	2
P	0.2	0.1	0.4	0.3

则

$$E(x) = -1 \times 0.2 + 0 \times 0.1 + 1 \times 0.4 + 2 \times 0.3 = 0.8$$

例 设X服从区间(a,b)上的均匀分布, 求 $E(X)$

解: X的密度函数为

$$p(x) = \begin{cases} \frac{1}{b-a}, a < x < b, \\ 0, otherwise \end{cases}$$

$$E(X) = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}$$

例:期望不存在的随机变量

例 柯西分布的密度函数为

$$p(x) = \frac{1}{\pi}\frac{1}{1+x^2}, -\infty < x < \infty$$

由级数不绝对收敛

$$\int_{-\infty}^{\infty} |x| \frac{1}{\pi}\frac{1}{1+x^2} dx = \infty$$

所以期望不存在

注意点

- 数学期望简称为**期望**.
- 数学期望又称为**均值**.
- 数学期望是一种**加权平均**.

定理2.2.1 设$Y = g(X)$ 是随机变量X的函数, 若$E(g(X))$存在, 则

$$E(g(X)) = \begin{cases} \sum\limits_i g(x_i)p(x_i) \\ \int_{-\infty}^{+\infty} g(x)p(x)dx \end{cases}$$

$p(x_i)$是分布列, 等于$P(X = x_i)$; $p(x)$是密度函数.

仅证明离散情况. 连续情况类似.

证明:$E\left(g\left(X\right)\right)=\sum_{x\in\Omega_X}g\left(x\right)P\left(X=x\right)$

证明: 令 $\Omega_X=support\left(X\right)=I$.

令 $Y=g\left(X\right)$. 所以 $\Omega_Y=support\left(Y\right)=g\left(I\right)$.

$$
\begin{aligned}
E\left(Y\right)&=\sum_{y\in g(I)}yP\left(Y=y\right)\\
&=\sum_{y\in g(I)}\left(y\sum_{\substack{x\in I\\ g(x)=y}}P\left(X=x\right)\right)\\
&=\sum_{x\in I}g\left(x\right)P\left(X=x\right)
\end{aligned}
$$

(最后一步可想像简单情况: Y有两个取值, 各对应X两个概率为0.25的取值)

例2.2.2 设随机变量 X 的概率分布为

X	0	1	2
P	$1/2$	$1/4$	$1/4$

求 $E(X^2+2)$.

解:

$$\begin{aligned} E(X^2+2) =& (0^2+2)\times 1/2 + (1^2+2)\times 1/4 + (2^2+2)\times 1/4 \\ =& 1 + 3/4 + 6/4 = 13/4 \end{aligned}$$

数学期望的性质

(1) $E(c) = c$

(2) $E(aX) = aE(X)$

(3) $E(g_1(X) + g_2(X)) = E(g_1(X)) + E(g_2(X))$

这些性质均可由定理2.2.1证明

例2.2.3 设

$$X \sim p(x) = \begin{cases} 2x, & 0 < x < 1. \\ 0, & \text{otherwise.} \end{cases}$$

求下列 X 的函数的数学期望. 1) $2X - 1$, 2) $(X - 2)^2$

解:

例2.2.3 设

$$X \sim p(x) = \begin{cases} 2x, & 0 < x < 1. \\ 0, & \text{otherwise.} \end{cases}$$

求下列 X 的函数的数学期望. 1) $2X-1$, 2) $(X-2)^2$

解:

1) $E(2X-1) = 1/3$,

2) $E(X-2)^2 = 11/6$.

- 数学期望反映了X取值的中心.
- 方差反映了X取值的离散程度.

2.3.1 方差与标准差的定义(Variance and standard deviation)

定义2.3.1 若 $E(X-E(X))^2$存在, 则称$E(X-E(X))^2$ 为 X 的**方差**, 记为

$$Var(X)=E(X-E(X))^2$$

称$\sigma_X=\sigma(X)=\sqrt{Var(X)}$ 为X 的**标准差**.

- 方差反映了随机变量相对其均值的偏离程度. 方差越大, 则随机变量的取值越分散.
- 标准差的量纲与随机变量的量纲相同.

2.3.2 方差的性质

性质 2.3.1 $Var(X) = E(X^2) - [E(X)]^2$.

性质 2.3.2 $Var(c) = 0$.

性质 2.3.3 $Var(aX + b) = a^2 Var(X)$.

- **课堂练习:** 运用方差定义证明上述三条性质 (3’)

例2.3.1 设

$$X \sim p(x) = \begin{cases} x, & 0 \le x < 1. \\ 2 - x, & 1 < x < 2. \\ 0, & \text{otherwise.} \end{cases}$$

求$E(X), Var(X)$

解:

例2.3.1 设

$$X \sim p(x) = \begin{cases} x, & 0 \leq x < 1. \\ 2-x, & 1 < x < 2. \\ 0, & \text{otherwise.} \end{cases}$$

求$E(X), Var(X)$

解： (1)

$$E(X) = \int_{-\infty}^{+\infty} xp(x)dx = \int_0^1 x \cdot xdx + \int_1^2 x(2-x)dx$$

$$= \frac{1}{3}x^3\Big|_0^1 + (x^2 - \frac{1}{3}x^3)\Big|_1^2 = 1$$

例2.3.1

解：(1)

$$E(X) = \int_{-\infty}^{+\infty} xp(x)dx = \int_0^1 x \cdot xdx + \int_1^2 x(2-x)dx$$
$$= \frac{1}{3}x^3\Big|_0^1 + (x^2 - \frac{1}{3}x^3)\Big|_1^2 = 1$$

(2)

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 p(x)dx$$
$$= \int_0^1 x^2 \cdot xdx + \int_1^2 x^2(2-x)dx = 7/6$$

所以, $Var(X) = E(X^2) - [E(X)]^2 = 7/6 - 1 = 1/6$

课堂练习

设

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

则方差 $Var(X)$=( )?

课堂练习

设

$$X \sim p(x) = \begin{cases} 1+x, & -1 \leq x < 0. \\ 1-x, & 0 < x \leq 1. \\ 0, & \text{其他}. \end{cases}$$

则方差 $Var(X)$=( )?

答案: $Var(X)$ = 1/6

随机变量的标准化(standardization)

设 $Var(X) > 0$, 令 $Y = \dfrac{X - E(X)}{\sqrt{Var(X)}}$. 则有 $E(Y) = 0$, $Var(Y) = 1$.

称 Y 为 X 的标准化.

2.3.3 切比雪夫不等式(Chebyshev's inequality)

定理2.3.1 设随机变量X的期望和方差存在, 则对任意正数ε, 有下面不等式成立

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{Var(X)}{\varepsilon^2}$$

切比雪夫不等式给出了大偏差发生概率的上界(upper bound), 方差越大, 上界越大.

切比雪夫不等式用于证明统计学和计量经济学中最重要的基础定律之一: 大数定律.

2.3.3 切比雪夫不等式(Chebyshev's inequality)

证明: 设X是一个连续随机变量, 其密度函数为$p(x)$.
记$E(X) = a$, 则

$$
\begin{aligned}
P(|X-a| \geq \varepsilon) &= \int\limits_{\{x:|x-a|\geq\varepsilon\}} p(x)dx \quad \text{所有outlier的概率之和} \\
&\leq \int\limits_{\{x:|x-a|\geq\varepsilon\}} \frac{(x-a)^2}{\varepsilon^2} p(x)dx \quad \text{因为} |X-a| \geq \varepsilon \\
&\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} (x-a)^2 p(x)dx \quad \text{扩大非负值的积分范围} \\
&= \frac{Var(X)}{\varepsilon^2}, Q.E.D.
\end{aligned}
$$

若方差为0, 则随机变量的取值集中在一点$(E(X))$上.

定理 2.3.2

若随机变量X的方差存在, 则$Var(X)=0$的充要条件是X几乎处处为某个常数a, 即

$$Var(X)=0 \Leftrightarrow P(X=a)=1$$

证明: 充分性: $P(X=a)=1 \Rightarrow Var(X)=0$

$Var(X)=E(X^2)-(E(X))^2=\int x^2p(x)dx-(\int xp(x)dx)^2=a^2-a^2=0$

必要性: $Var(X) = 0 \Rightarrow P(X = a) = 1$

(教材中证明难以理解, 在此用到Grimmett and Welsh (1986) pp107)

证明: 先证明对离散随机变量 $E(Y^2) = 0 \Rightarrow P(Y = 0) = 1$

$E(Y^2) = \sum_y y^2 P(Y = y) \geq 0$. 该等式只有在$P(Y = y) = 0, \forall y \neq 0$ 时成立.

又因为Y是离散随机变量, 所以$P(Y = 0) = 1$

现在, 令$Y = X - E(X)$, 则

$Var(X) = 0 \Rightarrow E(X - E(X))^2 = 0 \Rightarrow E(Y^2) = 0 \Rightarrow P(Y = 0) = 1 \Rightarrow P(X = E(X)) = 1$

连续情况下同理. $Q.E.D.$

第5次作业

习题2.2中题目5, 8, 12, 13, 14, 17, 19, 21

习题2.3中题目4, 7, 9, 10, 11, 12, 14

§2.4 常用离散分布

- 每个随机变量都对应一个分布.
- 不同的随机变量可以具有相同的分布.
- 常用的分布不多. 本节介绍常用的离散分布.

二项分布(binomial distribution)

记随机变量X为n重伯努里试验中“成功”(记为事件A)的次数, 则X可能的取值为$0, 1, \cdots, n$. 设 $P(A) = p$, 则X的分布列为

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

记为$X \sim b(n, p)$. (推导见后)

- 分布列所有取值概率之和恒为1.

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

二项分布的分布列推导(古典概型)

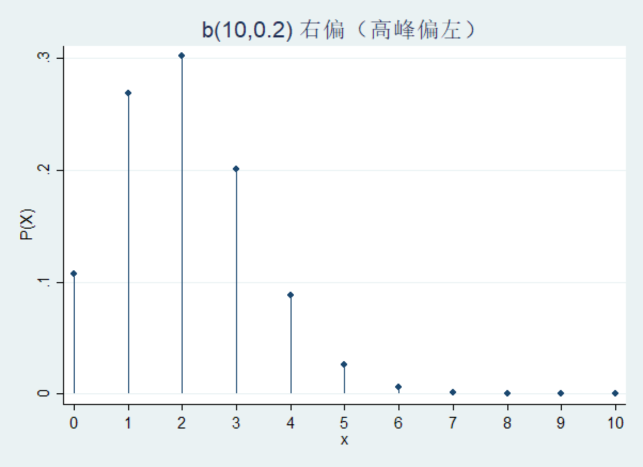
n重伯努里试验的基本结果可以记作$\omega = (\omega_1, \omega_2, ..., \omega_n)$, 样本点$\omega$中的元素$w_i$有两种取值$A$或$\overline{A}$, 所以样本空间$\Omega$中共有$2^n$个$\omega$ (各样本点概率之和等于1).

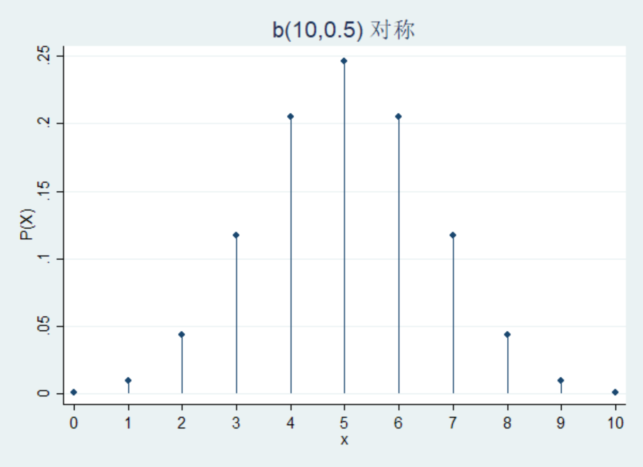
下面求事件$\{X = k\}$得概率. 若某个样本点$\omega = (\omega_1, \omega_2, ..., \omega_n) \in \{X = k\}$, 则$\omega_1, \omega_2, ..., \omega_n$中有$k$个$A$, $n - k$个$\overline{A}$, 所以由独立性

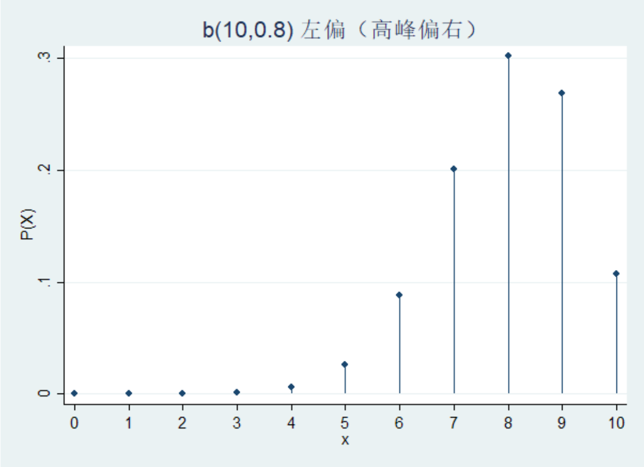
$$P(\omega) = p^k(1 - p)^{n-k}.$$

将全部这样的样本点概率相加

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


b(10,0.2) 右偏（高峰偏左）
P(X)
0
.1
.2
.3
0
1
2
3
4
5
6
7
8
9
10
x


b(10,0.5) 对称
P(X)
0
.05
.1
.15
.2
.25
0
1
2
3
4
5
6
7
8
9
10
x



教材错误: pp.85 图2.4.1(c), $P(X = 11)$的值应为0, 而非正数, 因为$X \in \{0, 1, ..., n\}$

```
// Binomial pdf

// b(10,0.2) b(10,0.5) b(10,0.8)

clear
set obs 11
help binomialp
gen b1_pdf=binomialp(10,_n-1,.2) //for each ob from 1 to 11 (indicated by _n), gen pdf.
gen b2_pdf=binomialp(10,_n-1,.5)
gen b3_pdf=binomialp(10,_n-1,.8)
gen x=_n-1



twoway dropline b1_pdf x, xlabel(#15) ytitle(P(X)) title( b(10,0.2) 右偏(高峰偏左))
twoway dropline b2_pdf x, xlabel(#15) ytitle(P(X)) title( b(10,0.5) 对称)
twoway dropline b3_pdf x, xlabel(#15) ytitle(P(X)) title( b(10,0.8) 左偏(高峰偏右))

twoway (dropline b1_pdf x) (dropline b2_pdf x) (dropline b3_pdf x)
```

$X \sim b(n,p)$的分布列为

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

所以

$$\begin{aligned}E(X)=&\sum_{k=0}^{n}k\binom{n}{k}p^k(1-p)^{n-k} \qquad (k=0\text{项等于}0)\\ =&np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}(1-p)^{(n-1)-(k-1)} \qquad (\text{构造多项式})\\ =&np[p+(1-p)]^{n-1}\\ =&np\end{aligned}$$

又因为

$$
\begin{aligned}
E(X^2) =& \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
=& \sum_{k=1}^{n} (k-1+1)k \binom{n}{k} p^k (1-p)^{n-k} \\
=& \sum_{k=1}^{n} (k-1)k \binom{n}{k} p^k (1-p)^{n-k} + \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\
=& n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{(n-2)-(k-2)} + np \\
=& n(n-1)p^2 + np
\end{aligned}
$$

所以$Var(X) = E(X^2) - [E(X)]^2 = np(1-p)$

伯努利分布(Bernoulli distribution)

$n=1$的二项分布 $b(1,p)$ 称为 $0-1$分布(two-point distribution),或伯努利分布.

$$P(X=x)=p^x(1-p)^{1-x}, x=0,1.$$

记为$X\sim b(1,p)$.

- 同二项分布的证明, 可得伯努利分布的期望和方差分别为p和$p(1-p)$

n重伯努利试验是n个相同的, 独立的伯努利试验的组成, 若将第i个伯努利试验中A出现的次数记为$X_i(X_i = 0, 1; i = 1, 2, \cdots, n)$, 则$X_i$相互独立, 且服从相同的两点分布$b(1, p)$.

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(二项分布的定义), 服从二项分布$b(n, p)$.

这就是二项分布$b(n, p)$和两点分布$b(1, p)$之间的联系.

即服从二项分布的随机变量是n个服从独立同为两点分布(IID)的随机变量之和.

例2.4.1 一批产品的合格率为0.8, 有放回地抽取4次, 每次一件, 则取得合格品件数X服从二项分布.

- 试验次数为 $n = 4$,
- “成功”即取得合格品的概率为 $p = 0.8$,
- 所以, $X \sim b(4, 0.8)$

思考: 若 Y 为不合格品件数, $Y \sim$?

例2.4.1 一批产品的合格率为0.8, 有放回地抽取4次, 每次一件, 则取得合格品件数X服从二项分布.

- 试验次数为 $n = 4$,
- "成功"即取得合格品的概率为 $p = 0.8$,
- 所以, $X \sim b(4, 0.8)$

思考: 若 Y 为不合格品件数, $Y \sim$?

- $Y \sim b(4, 0.2)$

例2.4.2 (课堂练习)

例2.4.2 设$X \sim b(2,p)$, $Y \sim b(4,p)$, 已知$P(X \geq 1) = 8/9$, 求 $P(Y \geq 1)$.

解:

例2.4.2 (课堂练习)

例2.4.2 设$X \sim b(2,p)$, $Y \sim b(4,p)$, 已知$P(X \geq 1) = 8/9$, 求 $P(Y \geq 1)$.

解: 由 $P(X \geq 1) = 8/9$, 知 $P(X = 0) = 1/9$.

例2.4.2 设$X \sim b(2,p)$, $Y \sim b(4,p)$, 已知$P(X \geq 1) = 8/9$, 求 $P(Y \geq 1)$.

解: 由 $P(X \geq 1) = 8/9$, 知 $P(X = 0) = 1/9$.

所以 $1/9 = P(X = 0) = (1-p)^2$,

例2.4.2 设$X \sim b(2,p)$, $Y \sim b(4,p)$, 已知$P(X \geq 1) = 8/9$, 求 $P(Y \geq 1)$.

解: 由 $P(X \geq 1) = 8/9$, 知 $P(X = 0) = 1/9$.

所以 $1/9 = P(X = 0) = (1-p)^2$,

从而解得: $p = 2/3$.

例2.4.2 (课堂练习)

例2.4.2 设$X \sim b(2,p)$, $Y \sim b(4,p)$, 已知$P(X \geq 1) = 8/9$, 求 $P(Y \geq 1)$.

解: 由 $P(X \geq 1) = 8/9$, 知 $P(X = 0) = 1/9$.

所以 $1/9 = P(X = 0) = (1-p)^2$,

从而解得: $p = 2/3$.

由此得: $P(Y \geq 1) = 1 - P(Y = 0) = 1 - (1-p)^4 = 80/81$

2.4.2 泊松分布 (Poisson distribution)

若随机变量 X 的概率分布为

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

则称X服从参数为λ的泊松分布$(\lambda>0)$, 记为 $X\sim P(\lambda)$.

泊松分布常用来描述单位计数, 如

- 一天来到商场的顾客数
- 单位面积玻璃上的气泡数
- 某地一天的犯罪数

分布列所有取值概率之和恒为1.

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\times\underbrace{\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}}_{(e^{\lambda}\text{的泰勒展开})}=e^{-\lambda}\times e^{\lambda}=1$$

e^x泰勒展开:

$$\begin{aligned}e^x&=f(0)\frac{x^0}{0!}+f'(0)\frac{x^1}{1!}+f''(0)\frac{x^2}{2!}+f'''(0)\frac{x^3}{3!}+f^{(4)}(0)\frac{x^4}{4!}+\cdots\\&=\frac{x^0}{0!}+\frac{x^1}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\cdots\\&=\sum_{n=0}^{\infty}\frac{x^n}{n!}\end{aligned}$$

泊松分布的期望和方差

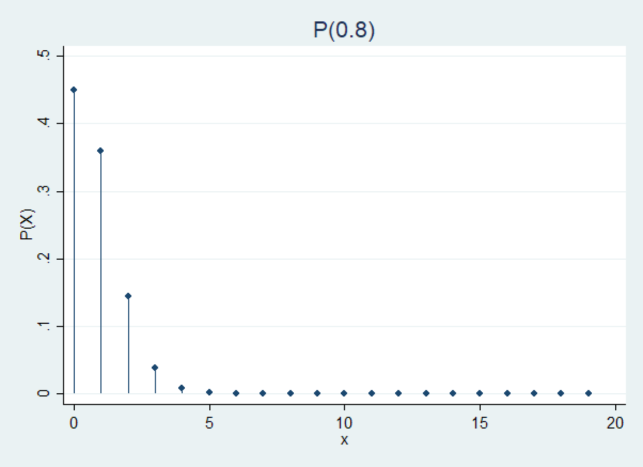
用类似二项分布计算期望和方差计算可得, 若随机变量 $X \sim P(\lambda)$, 则

$$E(X) = \lambda$$

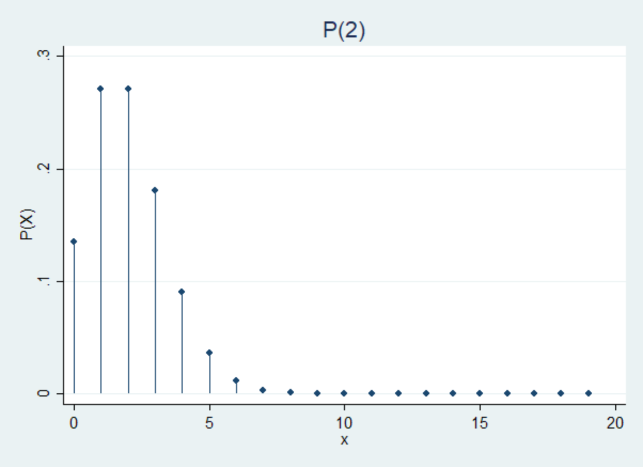
$$E(X^2) = \lambda^2 + \lambda$$

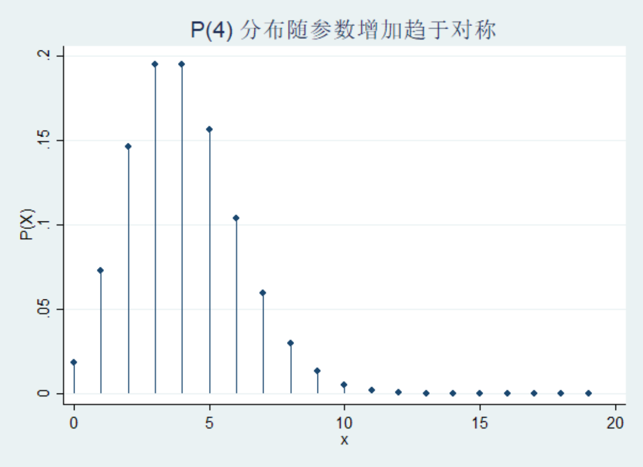
$$Var(X) = \lambda$$

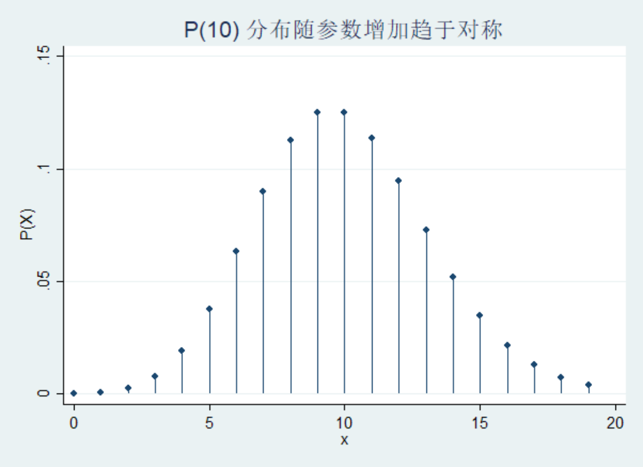
具体计算见教材.



教材错误: $pp86$ 表2.4.2, 图中空白处不为0, 而应为接近0的正数.


P(2)
P(X)
0
.1
.2
.3
0
5
10
15
20
x


P(4) 分布随参数增加趋于对称
P(X)
0
.05
.1
.15
.2
0
5
10
15
20
x


P(10) 分布随参数增加趋于对称
P(X)
0
.05
.1
.15
0
5
10
15
20
x

拓展: Stata 作图

```
clear
set obs 20
gen p1_pdf=poissonp(0.8,_n-1)
gen p2_pdf=poissonp(2,_n-1)
gen p3_pdf=poissonp(4,_n-1)
gen p4_pdf=poissonp(10,_n-1)
gen x=_n-1
```

```
twoway dropline p1_pdf x,  ytitle(P(X)) title( P(0.8) )
graph export "poisson1.png", as(png) replace
twoway dropline p2_pdf x,  ytitle(P(X)) title( P(2) )
graph export "poisson2.png", as(png) replace
twoway dropline p3_pdf x, ytitle(P(X)) title( P(4) 分布随参数增加趋于对称)
graph export "poisson3.png", as(png) replace
twoway dropline p4_pdf x, ytitle(P(X)) title( P(10) 分布随参数增加趋于对称)
graph export "poisson4.png", as(png) replace
```

例2.4.5 某商品月度销售量服从参数为8的泊松分布. 问月初需要多少库存在有90%的把握满足客户需求.

解: 以X表示商品的月销售量则$X \sim P(8)$. 那么满足要求的库存是使下试成立的最小正整数n.

$$P(X \leq n) \geq 0.9$$

查**附表1**得

$$P(X \leq 11) = 0.888$$

$$P(X \leq 12) = 0.936$$

故应进货12件. **(课堂练习查表)**

定理2.4.1 (二项分布的泊松近似) 在n重伯努里试验中, 记事件A在一次试验中发生的概率为 p_n. 如果当$n \to \infty$时, 有$np_n \to \lambda$, 则

$$P(X=k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}$$

或

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

注意:需要np_n不发散. **(下面给出证明)**

证明:记 $np_n = \lambda_n$, 即 $p_n = \lambda_n/n$, 我们可得

$$\begin{aligned}
&\binom{n}{k} p_n^k (1-p_n)^{n-k} \\
=&\frac{n(n-1)\ldots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k} \\
=&\frac{\lambda_n^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-k}
\end{aligned}$$

对固定的k有

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

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

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

对固定的k有

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

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

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

所以:

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

对任意$k(k = 0, 1, 2, \ldots)$ 成立. $Q.E.D.$

例2.4.6 已知某种疾病的发病率为0.001, 某单位共有5000. 问该单位患有这种疾病的人数不超过5人的概率为多少?

解: 设该单位患有该疾病的人数为X, 则有$X \sim b(5000, 0.001)$, 而我们所求的为

$$P(X \leq 5) = \sum_{k=0}^{5} \binom{5000}{k} 0.001^k 0.999^{5000-k}$$

这个概率的计算量很大.

例2.4.6 (二项分布的泊松分布近似)

例2.4.6 已知某种疾病的发病率为0.001, 某单位共有5000. 问该单位患有这种疾病的人数不超过5人的概率为多少?

解: 设该单位患有该疾病的人数为X, 则有$X \sim b(5000, 0.001)$, 而我们所求的为

$$P(X \leq 5) = \sum_{k=0}^{5} \binom{5000}{k} 0.001^k 0.999^{5000-k}$$

这个概率的计算量很大.

由于n很大, p很小, 且$\lambda = np = 5$.所以用泊松近似得

$$P(X \leq 5) \approx \sum_{k=0}^{5} \frac{5^k}{k!} e^{-5} = 0.616$$

2.4.3 超几何(hypergeometric)分布

超几何分布: 超几何分布对应于不返回抽样模型(sampling without replacement):

- N 个产品中有 M 个不合格品, 从中抽取n个, 不合格品的个数为X. 则X的概率分布列为

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

记为 $X\sim h(n,N,M)$.

其中$r=min\{M,n\}$,且$M\leq N$, $n\leq N$, N,M均为正整数.

易验证所有取值概率之和为1. 只需验证

$$\sum_{k=0}^{r} \binom{M}{k} \binom{N-M}{n-k} = \binom{N}{n}$$

其中$r = min\{M, n\}$, 且$M \leq N, n \leq N, N, M$均为正整数.

(习题1.2.1(5))

上式的意思是: 从N中取n的取法数量, 等于把N先分为任意两份$(M, N-M)$, 再在M中取k个, 在$N-M$中取$n-k$个, 再对所有k可能的取值相加.

超几何分布的期望与方差

利用上述逻辑和前述求期望和方差的方法可得,
若$X \sim h(n, N, M)$, 则

$$
\begin{aligned}
E(X) &= n\frac{M}{N} \\
Var(X) &= \frac{nM(N-M)(N-n)}{N^2(N-1)}
\end{aligned}
$$

超几何分布的二项分布近似

当$n \ll N$时, 不放回抽取(超几何分布)近似于有放回抽取(二项分布),

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \approx \binom{n}{k} p^k (1-p)^{n-k}, \text{其中 } p = \frac{M}{N}$$

几何分布: 记伯努利试验序列中, 每次试验中事件A发生的概率为p, 记X为事件首次出现时试验的次数, 则称X服从几何分布, 其分布列为

$$P(X=k)=(1-p)^{k-1}p, k=1,2,\ldots.$$

记为 $X \sim Ge(p)$.

易证明 (见教材)

- $E(X)=\dfrac{1}{p}$
- $Var(X)=\dfrac{1-p}{p^2}$

2.4.4 几何分布(geometric)

几何分布具有无记忆性, 即:$P(X > m+n|X > m) = P(X > n)$

- $P(X > n)$, 前n次伯努利试验没有出现事件A的概率
- $P(X > m+n|X > m)$在前m次伯努利试验没有出现事件A的前提下, 继续n次试验没有出现事件A的概率

证明用定义, 见教材.

负二项(negative binomial)分布(帕斯卡分布, Pascal distribution)

负二项分布: 伯努里试验序列中, 记每次试验中事件A发生的概率为p. 如果X为事件A第 r次出现时试验的次数, 则X可能的取值为$r, r+1, \ldots$. 称X服从负二项分布或帕斯卡分布, 其分布列为

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

记为$X \sim Nb(r, p)$.

- 上式由来: 最后一次一定是A, 故前$k-1$次试验A出现$r-1$次, 即二项分布 $\binom{k-1}{r-1}p^{r-1}(1-p)^{k-r}$, 再乘以最后一次$A$出现概率$p$
- 当 $r=1$时, 负二项分布退化成为几何分布

注意点

(1) 二项分布的随机变量是独立同分布$0-1$(伯努利分布)随机变量之和(之前已证明).

(2) 负二项随机变量是独立几何随机变量之和. 为什么?

负二项随机变量是独立几何随机变量之和

证明:

将第1个A出现的试验次数记为X_1,

从第1个A出现后算起, 将第2个A出现的试验次数记为X_2,...

从第$r-1$个A出现后算起, 将第r个A出现的试验次数记为X_r

则X_i独立同分布, 且服从几何分布$X_i \sim Ge(p)$, 此时
有$X = X_1 + X_2 + \ldots + X_i + \ldots + X_r \sim Nb(r, p)$ $Q.E.D$

常用离散分布的数学期望

- $0-1$ 分布的数学期望 $= p$
- 二项分布$b(n,p)$ 的数学期望 $= np$
- 几何分布$Ge(p)$ 的数学期望 $= 1/p$
- 泊松分布$P(\lambda)$ 的数学期望 $= \lambda$

常用离散分布的方差

- $0-1$ 分布的方差$= p(1-p)$
- 二项分布$b(n,p)$ 的方差$= np(1-p)$
- 几何分布$Ge(p)$ 的方差 $= (1-p)/p^2$
- 泊松分布 $P(\lambda)$的方差$= \lambda$

课堂练习(1')

例2.5.6 已知随机变量 X 服从二项分布，
且$E(X) = 2.4, Var(X) = 1.44$, 则参数 n, p 的值为多少?

解:

课堂练习(1')

例2.5.6 已知随机变量 X 服从二项分布,
且$E(X) = 2.4, Var(X) = 1.44$, 则参数 n, p 的值为多少?

解: 从 $2.4 = np$, $1.44 = np(1-p)$ 中解得$n = 6, p = 0.4$.

例2.5.7 设 X 表示 10 次独立重复射击命中目标的次数, 每次射中目标的概率为0.4, 则 $E(X^2)$的值为多少?

解:

例2.5.7 设 X 表示 10 次独立重复射击命中目标的次数, 每次射中目标的概率为0.4, 则 $E(X^2)$的值为多少?

解: 由题意, X服从二项分布.

因为 $E(X) = np = 4$, $Var(X) = np(1-p) = 2.4$, 所以

$E(X^2) = Var(X) + (E(X))^2 = 2.4 + 16 = 18.4$

第6次作业

习题2.4中题目4, 7, 9, 11, 13, 15, 17, 19

§2.5 常用的连续分布

- 正态分布(normal distribution)
- 均匀分布(uniform distribution)
- 指数分布(exponential distribution)
- 伽玛分布(Gamma distribution)
- 贝塔分布(Beta distribution)

2.5.1 正态分布(高斯分布)

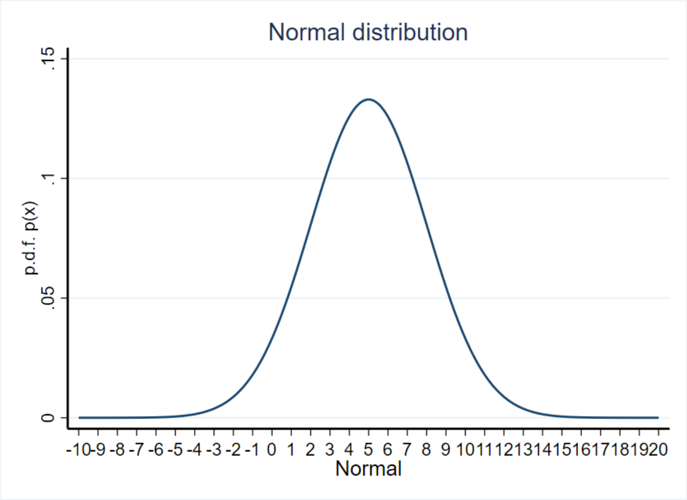
正态分布: 若随机变量X的密度函数为

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

则称X服从**正态分布**, 称X为**正态变量**, 记为$X \sim N(\mu, \sigma^2)$, 其中$\sigma > 0$, μ 是任意实数.

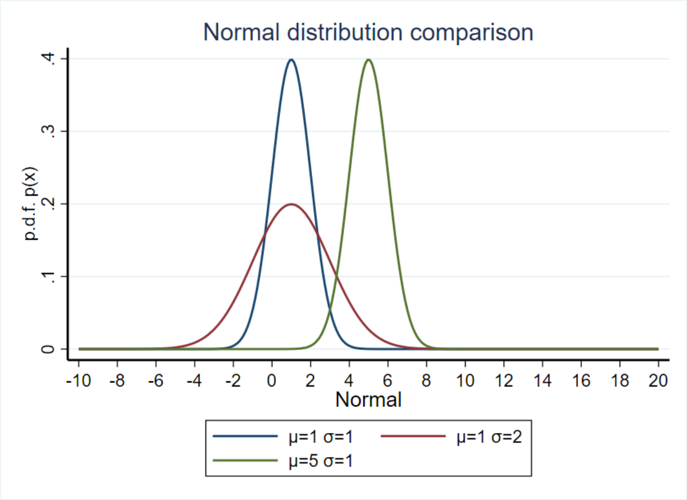
- μ是位置参数. $E(X) = \mu$.
- σ 是尺度参数. $Var(X) = \sigma^2$.

$X \sim N(5,3)$密度函数示意图(bell shaped)



拓展: Stata code

```
#delimit ;
graph twoway (function y=normalden(x,5,3), range(-10 20) lw(medthick)),
  title("Normal distribution")
  xtitle("Normal", size(medlarge)) ytitle("p.d.f. p(x)")
  xlabel(-10(1)20)
  xscale(lw(medthick)) yscale(lw(medthick))
  legend(label(1 "{&mu}=1 {&sigma}=2"))
  graphregion(fcolor(white));
 #delimit cr
 graph export normal1.png, as(png) replace
```


Normal distribution comparison
p.d.f. p(x)
0
.1
.2
.3
.4
-10
-8
-6
-4
-2
0
2
4
6
8
10
12
14
16
18
20
Normal
μ=1 σ=1
μ=1 σ=2
μ=5 σ=1

拓展: Stata code

```
#delimit ;
graph twoway (function y=normalden(x,1,1), range(-10 20) lw(medthick))
             (function y=normalden(x,1,2), range(-10 20) lw(medthick))
 (function y=normalden(x,5,1), range(-10 20) lw(medthick)),
  title("Normal distribution comparison")
  xtitle("Normal", size(medlarge)) ytitle("p.d.f. p(x)")
  xlabel(-10(2)20)
  xscale(lw(medthick)) yscale(lw(medthick))
  legend(label(1 "{&mu}=1 {&sigma}=1") label(2 "{&mu}=1 {&sigma}=2")
  label(3 "{&mu}=5 {&sigma}=1"))
  graphregion(fcolor(white));
 #delimit cr
 graph export normal2.png, as(png) replace
```

正态分布的性质

(1) $p(x)$ 关于μ 是对称的. 在μ 点$p(x)$ 取得最大值.

(2) 若 σ 固定, μ 改变, $p(x)$左右移动, 形状保持不变.

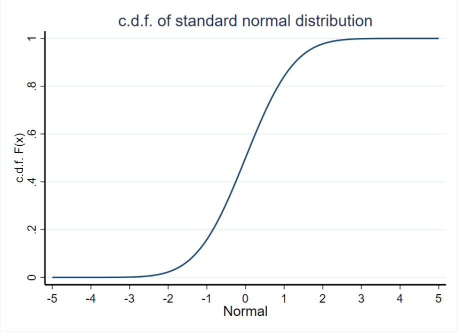
(3) 若μ 固定, σ改变, σ越大曲线越平坦; σ 越小曲线越陡峭.

分布函数

$X \sim N(\mu, \sigma^2)$的分布函数为

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

例: $X \sim N(0,1)$分布函数示意图:



拓展: Stata code

```
#delimit ;
graph twoway (function y=normal(x), range(-5 5) lw(medthick)),
  title("c.d.f. of standard normal distribution")
  xtitle("Normal", size(medlarge)) ytitle("c.d.f. F(x)")
  xlabel(-5(1)5)
  xscale(lw(medthick)) yscale(lw(medthick))
  legend(label(1 "{&mu}=0 {&sigma}=1"))
  graphregion(fcolor(white));
 #delimit cr
 graph export normal3.png, as(png) replace
```

标准正态分布N(0, 1)

称$\mu = 0, \sigma = 1$时的正态分布$N(0, 1)$为标准正态分布

- 密度函数记为 $\varphi(x)$,
- 分布函数记为 $\Phi(x)$.

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

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

- $\Phi(0) = \frac{1}{2}$; $\Phi(x) = 1 - \Phi(-x)$

$\Phi(x)$ 的计算

(1) $x \geq 0$时, 查标准正态分布函数表(附表2).

(2) $x < 0$时, 用$\Phi(x) = 1 - \Phi(-x)$

若 $X \sim N(0, 1)$, 则

(1) $P(X \leq a) = \Phi(a)$;
(2) $P(X > a) = 1 - \Phi(a)$;
(3) $P(a < X < b) = \Phi(b) - \Phi(a)$;
(4) 若$a \geq 0$, 则

$$\begin{aligned} P(|X| < a) =& P(-a < X < a) = \Phi(a) - \Phi(-a) \\ =& \Phi(a) - [1 - \Phi(a)] = 2\Phi(a) - 1 \end{aligned}$$

例2.5.1 (课堂练习查表2’)

例2.5.1 设$X \sim N(0,1)$, 求$P(X > -1.96)$, $P(|X| < 1.96)$ **解:**

例2.5.1 设$X \sim N(0,1)$, 求$P(X > -1.96)$, $P(|X| < 1.96)$ **解:**

$$\begin{aligned} P(X > -1.96) &= 1 - \Phi(-1.96) \\ &= 1 - (1 - \Phi(1.96)) = \Phi(1.96) \\ &= 0.975(\text{查表得}) \end{aligned}$$

$$P(|X| < 1.96) = 2\Phi(1.96) - 1 = 2 \times 0.975 - 1 = 0.95$$

服从标准正态分布的随机变量的某一个观察(实现)落在(-1.96, 1.96)上的概率为95%.

思考： 标准正态分布落在哪个对称区间上的上的概率为90% (或99%)?

(拓展：应用统计学中我们将利用上述结论构造估计量置信区间)

例2.5.2 设 $X \sim N(0,1)$, $P(X \leq b) = 0.9515$, $P(X \leq a) = 0.04947$, 求 a, b.

解:

例2.5.2 设 $X \sim N(0,1)$, $P(X \leq b) = 0.9515$, $P(X \leq a) = 0.04947$, 求 a, b.

解: $\Phi(b) = 0.9515 > 1/2$, 所以 $b > 0$,

反查表得: $\Phi(1.66) = 0.9515$, 故 $b = 1.66$

而 $\Phi(a) = 0.0495 < 1/2$, 所以 $a < 0$, $\Phi(-a) = 0.9505$,

反查表得: $\Phi(1.65) = 0.9505$, 故 $a = -1.65$

定理2.5.1 设$X \sim N(\mu, \sigma^2)$, $Y = \dfrac{X-\mu}{\sigma}$, 则 $Y \sim N(0,1)$.

证明:

$$F_Y(y) = P(Y \le y) = P\left(\frac{X-\mu}{\sigma} \le y\right) = P(X \le y\sigma + \mu) = F_X(\mu + y\sigma)$$

$$\begin{aligned} p_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(\mu + y\sigma) = p_X(\mu + y\sigma) \cdot \sigma \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\mu + y\sigma - \mu)^2}{2\sigma^2}} \cdot \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = \varphi(y) \end{aligned}$$

一般正态分布的标准化(重点)

定理2.5.1 设$X \sim N(\mu, \sigma^2)$, $Y = \dfrac{X-\mu}{\sigma}$, 则 $Y \sim N(0,1)$.

推论: 若 $X \sim N(\mu, \sigma^2)$, 则

- $F_X(x) = \Phi(\dfrac{x-\mu}{\sigma})$

课堂练习: 证明上述推论.

一般正态分布的标准化(重点)

定理2.5.1 设$X \sim N(\mu, \sigma^2)$, $Y = \dfrac{X-\mu}{\sigma}$, 则 $Y \sim N(0,1)$.

推论: 若 $X \sim N(\mu, \sigma^2)$, 则

- $F_X(x) = \Phi(\dfrac{x-\mu}{\sigma})$

课堂练习: 证明上述推论.

证明: (利用定理2.5.1, 将$X = \mu + Y\sigma$带入$F_X(x)$逆推即可)

$$F_X(x) = P(X \leq x) = P(\mu + Y\sigma \leq x) = P(Y \leq \frac{x-\mu}{\sigma}) = \Phi(\frac{x-\mu}{\sigma})$$

一般正态分布的标准化(重点)

推论: 若 $X \sim N(\mu, \sigma^2)$, 则

- $P(X \leq a) = \Phi(\dfrac{a-\mu}{\sigma})$
- $P(X > a) = 1 - \Phi(\dfrac{a-\mu}{\sigma})$
- $P(a < X \leq b) = \Phi(\dfrac{b-\mu}{\sigma}) - \Phi(\dfrac{a-\mu}{\sigma})$ (许多统计推断方法逻辑的基本出发点)

例2.5.3 设 $X \sim N(10,4)$, 求 $P(10 < X < 13)$和 $P(|X-10| < 2)$.

解:

$$P(10 < X < 13) = \Phi(1.5) - \Phi(0) = 0.9332 - 0.5 = 0.4332$$

$$P(|X-10| < 2) = P(8 < X < 12) = 2\Phi(1) - 1 = 0.6826$$

例2.5.4 设 $X \sim N(\mu, \sigma^2)$, $P(X \le -5) = 0.045$, $P(X \le 3) = 0.618$, 求μ 及σ.

解:

$$\begin{cases} \Phi(\dfrac{-5-\mu}{\sigma}) = 0.045 \\ \Phi(\dfrac{3-\mu}{\sigma}) = 0.618. \end{cases}$$

$\Rightarrow$

$$\begin{cases} \dfrac{5+\mu}{\sigma} = 1.69 \\ \dfrac{3-\mu}{\sigma} = 0.3. \end{cases}$$

$\Rightarrow$

$$\begin{cases} \mu = 1.76 \\ \sigma = 4. \end{cases}$$

课堂练习1

已知 $X \sim N(3, 2^2)$, 且$P(X > k) = P(X \leq k)$, 则 $k = ()$ **解:**

课堂练习1

已知 $X \sim N(3, 2^2)$, 且$P(X > k) = P(X \leq k)$, 则 $k = ()$ **解: 3**

课堂练习2

设$X \sim N(\mu, 4^2)$, $Y \sim N(\mu, 5^2)$, 记$p_1 = P(X \le \mu - 4)$, $p_2 = P(Y \ge \mu + 5)$, 则( )

①对任意的μ，都有 $p_1 = p_2$
②对任意的 μ，都有$p_1 < p_2$
③只个别的 μ，才有 $p_1 = p_2$
④对任意的 μ，都有 $p_1 > p_2$

解:

课堂练习2

设$X \sim N(\mu, 4^2)$, $Y \sim N(\mu, 5^2)$, 记$p_1 = P(X \leq \mu - 4)$, $p_2 = P(Y \geq \mu + 5)$, 则( )

①对任意的μ，都有 $p_1 = p_2$
②对任意的 μ，都有$p_1 < p_2$
③只个别的 μ，才有 $p_1 = p_2$
④对任意的 μ，都有 $p_1 > p_2$

解: ①

因为$P(X \leq a) = \Phi(\frac{a-\mu}{\sigma})$, $P(X > a) = 1 - \Phi(\frac{a-\mu}{\sigma})$.

课堂练习3

设$X \sim N(\mu, \sigma^2)$, 则随σ的增大, 概率 $P(|X-\mu|<\sigma)$ ( )

①单调增大　②单调减少

③保持不变　④增减不定

解:

课堂练习3

设$X \sim N(\mu, \sigma^2)$, 则随σ的增大, 概率 $P(|X-\mu|<\sigma)$ ( )

①单调增大　②单调减少
③保持不变　④增减不定

解: ③

因为$P(|X-\mu|<\sigma) = P(|\dfrac{X-\mu}{\sigma}|<1)$. 而$\dfrac{X-\mu}{\sigma} \sim N(0,1)$,
与σ 无关. $P(|\dfrac{X-\mu}{\sigma}|<1) = 2\Phi(1)-1 = 0.68$

正态分布的数学期望与方差

由定理**2.5.1**知, 若$X \sim N(\mu, \sigma^2)$, $Y = \dfrac{X-\mu}{\sigma}$, 则 $Y \sim N(0,1)$.

所以Y的期望为

$$E(Y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} te^{-\frac{t^2}{2}} dt$$

注意$te^{-\frac{t^2}{2}}$是奇函数, 故积分为0, 所以$E(Y)=0$. 所以$E(X)=\mu+\sigma\times 0=\mu$

又因为

$$
\begin{aligned}
Var(Y) =& E(Y^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2}} dt \\
=& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t d(e^{-\frac{t^2}{2}}) \\
=& \frac{1}{\sqrt{2\pi}} \left(-te^{-\frac{t^2}{2}} \Bigg|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \right) \quad \text{(分部定积分)} \\
=& \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \right) = \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} \quad \text{(高斯积分)} \\
=& 1
\end{aligned}
$$

所以 $Var(X) = Var(\mu + \sigma Y) = \sigma^2$

标准化随机变量

变换$Y=\dfrac{X-\mu}{\sigma}$具有一般性. 若随机变量X的期望为μ, 方差为σ^2, 则$Y=\dfrac{X-\mu}{\sigma}$的期望为0, 方差为1.

证明:

$$E(Y)=\frac{E(X)-\mu}{\sigma}=0$$

$$\begin{aligned}
Var(Y)=&Var\left(\frac{X-\mu}{\sigma}\right)\\
=&\frac{1}{\sigma^2}Var(X)\\
=&\frac{1}{\sigma^2}\sigma^2=1
\end{aligned}$$

正态分布的 3σ 原则

3σ **原则:** 虽然正态变量的取值范围是$(-\infty, \infty)$, 但有99.73%的概率会落在$(\mu - 3\sigma, \mu + 3\sigma)$内.

设$X \sim N(\mu, \sigma^2)$,

则$P(|X-\mu| < k\sigma) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$

所以

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

2.5.2 均匀分布

均匀分布: 若随机变量X的密度函数为

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

则称X服从区间(a,b)上的均匀分布, 记为$X \sim U(a,b)$. 其分布函数为

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

均匀分布的期望和方差

容易证明$E(X)=\dfrac{a+b}{2}, Var(X)=\dfrac{(b-a)^2}{12}$

$$E(X)=\int_a^b \frac{x}{b-a}\mathrm{d}x=\frac{b^2-a^2}{2(b-a)}=\frac{a+b}{2}$$

$$E\left(X^2\right)=\int_a^b \frac{x^2}{b-a}\mathrm{d}x=\frac{b^3-a^3}{3(b-a)}=\frac{a^2+ab+b^2}{3}$$

$$\mathrm{Var}(X)=E\left(X^2\right)-[E(X)]^2=\frac{a^2+ab+b^2}{3}-\frac{(a+b)^2}{4}=\frac{(b-a)^2}{12}$$

例2.5.5 $X \sim U(2,5)$. 现在对 X 进行三次独立观测, 试求至少有两次观测值大于3的概率.

解:

例2.5.5 $X \sim U(2,5)$. 现在对 X 进行三次独立观测, 试求至少有两次观测值大于3的概率.

解: $A=\{X>3\}$, 则 $P(A)=P(X>3)=2/3$. 设 Y 表示三次独立观测中A 出现的次数, 则 $Y \sim b(3,2/3)$, 所求概率为

$$\begin{aligned} P(Y \geq 2) =& P(Y=2)+P(Y=3) \\ =& C_3^2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)+C_3^3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^0 \\ =& 20/27 \end{aligned}$$

2.5.3 指数分布(exponential distribution)

指数分布: 若随机变量X的密度函数为

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

则称X服从指数分布, 记为$X \sim Exp(\lambda)$, 其中$\lambda > 0$. 其分布函数为

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

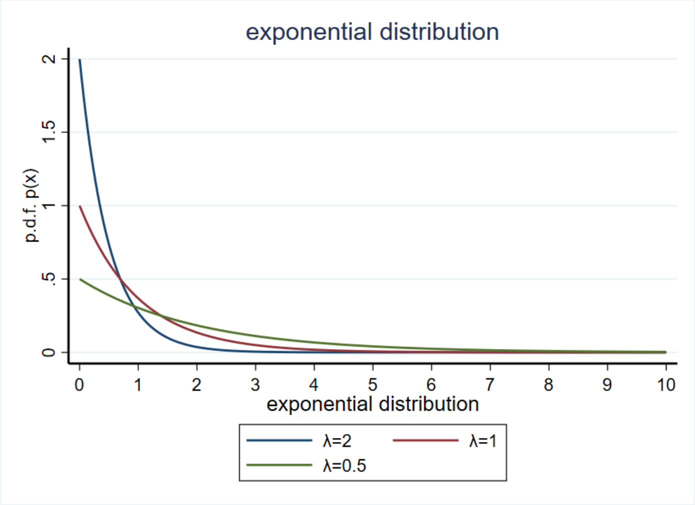
容易证明$E(X) = \dfrac{1}{\lambda}$, $Var(X) = \dfrac{1}{\lambda^2}$ (见后).

指数分布具有无记忆性, 即:$P(X > s + t | X > s) = P(X > t)$

(例：如寿命服从指数分布的产品使用s时间没发生故障, 则再使用t时间不发生故障的概率与已使用的s时常无关.)

2.5.3 指数分布(exponential distribution)

指数分布可用作描述“寿命”分布, 如电子元件的寿命, 动物的寿命等.



拓展: Stata code

```
#delimit ;
graph twoway (function y= exponentialden(0.5, x), range(0 10) lw(medthick))
             (function y= exponentialden(1,x), range(0 10) lw(medthick))
 (function y= exponentialden(2,x), range(0 10) lw(medthick)),
  title("exponential distribution")
  xtitle("exponential distribution", size(medlarge)) ytitle("p.d.f. p(x)")
  xlabel(0(1)10)
  xscale(lw(medthick)) yscale(lw(medthick))
  legend(label(1 "{&lambda}=2") label(2 "{&lambda}=1") label(3 "{&lambda}=0.5"))
  graphregion(fcolor(white));
 #delimit cr
 graph export  exponential.png , as(png) replace
```

指数分布的期望和方差

容易证明$E(X)=\dfrac{1}{\lambda}, Var(X)=\dfrac{1}{\lambda^2}$.

$$E(X)=\int_0^{\infty} x\lambda \mathrm{e}^{-\lambda x}\mathrm{d}x=\int_0^{\infty} x\mathrm{d}\left(-\mathrm{e}^{-\lambda x}\right)$$

$$=-x\mathrm{e}^{-\lambda x}\Big|_0^{\infty}+\int_0^{\infty}\mathrm{e}^{-\lambda x}\mathrm{d}x=-\frac{1}{\lambda}\mathrm{e}^{-\lambda x}\Big|_0^{\infty}=\frac{1}{\lambda}$$

$$E\left(X^2\right)=\int_0^{\infty} x^2\lambda \mathrm{e}^{-\lambda x}\mathrm{d}x=\int_0^{\infty} x^2\mathrm{d}\left(-\mathrm{e}^{-\lambda \mathrm{x}}\right)$$

$$=-x^2\mathrm{e}^{-\lambda x}\Big|_0^{\infty}+2\int_0^{\infty} x\mathrm{e}^{-\lambda x}\mathrm{d}x=\frac{2}{\lambda^2}$$

$$\mathrm{Var}(X)=E\left(X^2\right)-[E(X)]^2=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}$$

指数分布的无记忆性

指数分布具有无记忆性, 即: $P(X > s + t | X > s) = P(X > t)$

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

又因为

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

所以

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

2.5.4 伽玛分布(Gamma distribution)

称以下函数

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \alpha > 0$$

为伽玛函数. 伽玛函数具有以下性质

- $\Gamma(1) = 1$
- $\Gamma(1/2) = \sqrt{\pi}$ (高斯积分)
- $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ (分部积分)
- $\Gamma(n+1) = n\Gamma(n) = n!$ (当n为正整数时)

伽玛分布: 若随机变量X的密度函数为

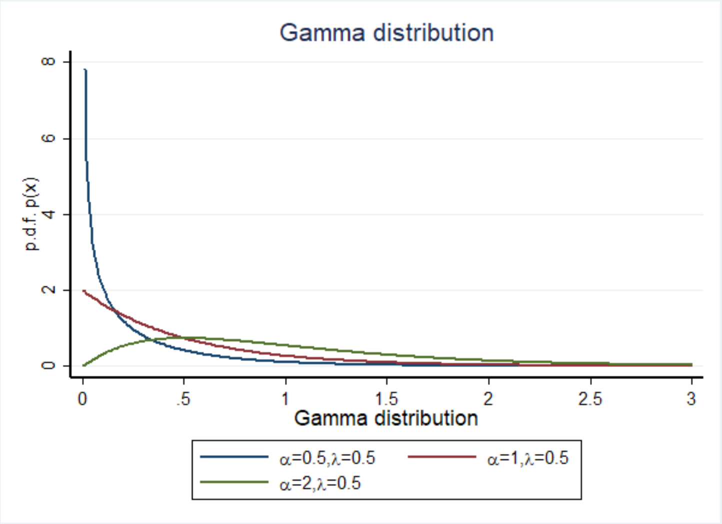
$$p(x) = \begin{cases} \dfrac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

则称X服从伽玛分布, 记为$X \sim Ga(\alpha, \lambda)$, 其中$\alpha > 0, \lambda > 0$.

$$E(X) = \frac{\alpha}{\lambda}$$

$$Var(X) = \frac{\alpha}{\lambda^2}$$

2.5.4 伽玛分布(Gamma distribution)



伽玛分布$Ga(\alpha,\lambda)$的特例

1. $Ga(\alpha,\lambda)$ 在 $\alpha=1$时退化为指数分布

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

2. $Ga(\alpha,\lambda)$ 在 $\alpha=n/2,\lambda=1/2$时是自由度为n的χ^2(卡方, Chi-square)分布

$$Ga(n/2,1/2)=\chi^2(n)$$

其密度函数为

$$p(x)=\begin{cases}\dfrac{1}{2^{n/2}\Gamma(\frac{n}{2})}x^{n/2-1}e^{x/2}, & x>0\\ 0, & x\le 0\end{cases}$$

若$X\sim\chi^2(n)$, 则$E(X)=n, Var(X)=2n$.

2.5.5 贝塔分布(Beta distribution)

贝塔分布:

$$p(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, 0 < x < 1$$

记为 $X \sim Be(a,b)$, 其中$a >, b > 0$.

称 $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx$ 为贝塔函数.

- 注意点

(1) 贝塔函数: $B(a,b)=B(b,a)$

(2) 贝塔函数和伽马函数关系:
$B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a+b)$

(3) 贝塔分布: $Be(1,1)=U(0,1)$ (均匀分布)

常用连续分布的数学期望

- 正态分布 $N(\mu, \sigma^2) : E(X) = \mu$
- 均匀分布 $U(a, b) : E(X) = (a + b)/2$
- 指数分布 $Exp(\lambda) : E(X) = 1/\lambda$
- 伽玛分布$Ga(\alpha, \lambda) : E(X) = \alpha/\lambda$
- 贝塔分布$Be(a, b) : E(X) = a/(a + b)$

常用连续分布的方差

- 正态分布 $N(\mu, \sigma^2)$ 的方差=σ^2
- 均匀分布 $U(a, b)$ 的方差 = $(b-a)^2/12$
- 指数分布 $Exp(\lambda)$ 的方差= $1/\lambda^2$
- 伽玛分布$Ga(\alpha, \lambda)$ 的方差= $\dfrac{\alpha}{\lambda^2}$
- 贝塔分布$Be(a, b)$ 的方差= $\dfrac{ab}{(a+b)^2(a+b+1)}$

课堂练习

设 $E(X)=\mu$, $Var(X)=\sigma^2$, 则对任意常数 C, 必有( ).

① $E[(X-C)^2]=E(X^2)-C^2$
② $E[(X-C)^2]=E[(X-\mu)^2]$
③ $E[(X-C)^2]<E[(X-\mu)^2]$
④ $E[(X-C)^2]\geq E[(X-\mu)^2]$

解:

课堂练习

设 $E(X)=\mu$, $Var(X)=\sigma^2$, 则对任意常数 C, 必有( ).

① $E[(X-C)^2]=E(X^2)-C^2$
② $E[(X-C)^2]=E[(X-\mu)^2]$
③ $E[(X-C)^2]<E[(X-\mu)^2]$
④ $E[(X-C)^2]\geq E[(X-\mu)^2]$

解: ④

思路: $E[(X-C)^2]=E[((X-\mu)+(\mu-C))^2]$, 展开消项即可得.

第7次作业

习题2.5中题目3, 6, 7, 8, 10, 15, 17, 19, 20, 22, 27, 30, 31, 32

设$y=g(x)$是定义在直线上的一个函数, X是一个随机变量, 那么$Y=g(X)$作为随机变量X的函数, 同样**也是一个随机变量**.

问题: 已知 X的分布, 求 $Y=g(X)$的分布.

例如: $Y_1(X)=4X+3$; $Y_2(X)=|X|$; $Y_3(X)=\pi X^2$.

2.6.1 离散随机变量函数的分布

当 X 为离散随机变量时, $Y = g(X)$为离散随机变量.

将$g(x_i)$一一列出, 再将相等的值合并即可.

例2.6.1 已知随机变量X的分布列如下，求$Y = X^2 + X$的分布列.

X	-2	-1	0	1	2
P	0.2	0.1	0.1	0.3	0.3

解:$Y = X^2 + X$的分布列为

$Y = X^2 + X$	2	0	0	2	6
P	0.2	0.1	0.1	0.3	0.3

即

Y	0	2	6
P	0.2	0.5	0.3

定理2.6.1 设 $X \sim p_X(x)$是连续随机变量, $y = g(x)$ 是 x 的严格单调函数, 记 $x = h(y)$ 为 $y = g(x)$的反函数, 且$h(y)$连续可导, 则$Y = g(X)$的密度函数为:

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

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

证明： 不妨设$y=g(x)$是增函数，意味$y=g(x)$最小可取$a=g(-\infty)$，最大可取$b=g(\infty)$.

$$F_Y(y)=\begin{cases} P(Y\le y)=0, & y<a \\ P(Y\le y), & a\le y\le b \\ P(Y\le y)=1, & y>b \end{cases}$$

当$a\le y\le b$时，由于$y=g(x)$ 是增函数，故反函数$x=h(y)$也是增函数

$$P(Y\le y)=P(g(X)\le y)=P(X\le h(y))=\int_{-\infty}^{h(y)} p_X(x)dx$$

所以:

$$
\begin{aligned}
p_Y(y) =& \frac{dF_Y(y)}{dy} = \frac{d}{dy}\int_{-\infty}^{h(y)} p_X(x)dx \\
=& \frac{d}{dy}\left(F_X(x)\Big|_{-\infty}^{h(y)}\right) \\
=& \frac{d}{dy}\left(F_X(h(y)) - 0\right) \\
=& p_X(h(y)) \cdot |h'(y)|
\end{aligned}
$$

综上,

$$
p_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} p_X(h(y)) \cdot |h'(y)|, & a < y < b. \\ 0, & \text{其他}. \end{cases}
$$

若$y = g(x)$是减函数, 则$h'(y) < 0$, 所以加上绝对值符号, 这时$a = g(\infty), b = g(-\infty)$.

例2.6.1 设 $X \sim p_X(x) = \dfrac{1}{\pi(1+x^2)}$, 求 $Y = e^X$ 的分布.

解: $y = e^x$ 单调可导, 反函数 $x = h(y) = lny$, $h'(y) = \dfrac{1}{y}$, 所以当 $y > 0$ 时,

$$p_Y(y) = p_X(h(y)) \cdot |h'(y)| = p_X(lny) \cdot |\frac{1}{y}| = \frac{1}{\pi y(1+(lny)^2)}$$

由此得

$$p_Y(y) = \begin{cases} \dfrac{1}{\pi y(1+(lny)^2)}, & y > 0. \\ 0, & \text{其他}. \end{cases}$$

正态变量的线性不变性

定理2.6.2 设 $X \sim N(\mu, \sigma^2)$, 则当$a \neq 0$时, $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

- 正态变量的线性变换仍为正态变量
- 由此得: 若 $X \sim N(\mu, \sigma^2)$, 则 $Y = (X - \mu)/\sigma \sim N(0, 1)$.

证明: 当$a > 0$时, $Y = aX + b$是增函数, 其反函数为 $X = (Y - b)/a$. 由**定理2.6.1**得,

$$\begin{aligned} p_Y(y) =& p_X(h(y)) \cdot |h'(y)| = p_X((y - b)/a) \cdot \frac{1}{a} \\ =& \frac{1}{\sqrt{2\pi} a\sigma} exp\left\{-\frac{(y - a\mu - b)^2}{2a^2\sigma^2}\right\} \end{aligned}$$

所以, $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

当$a < 0$时, 同理得证. $Q.E.D.$

对数正态分布(log-normal distribution)

定理2.6.3 设 $X \sim N(\mu, \sigma^2)$, 则 $Y = e^X$的服从

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

这个分布称为**对数正态分布**, 记为$Y \sim LN(\mu, \sigma^2)$

- 反之亦成立, 如果Y服从对数正态分布, 则$X = ln(Y)$服从正态分布.

对数正态分布(log-normal distribution)

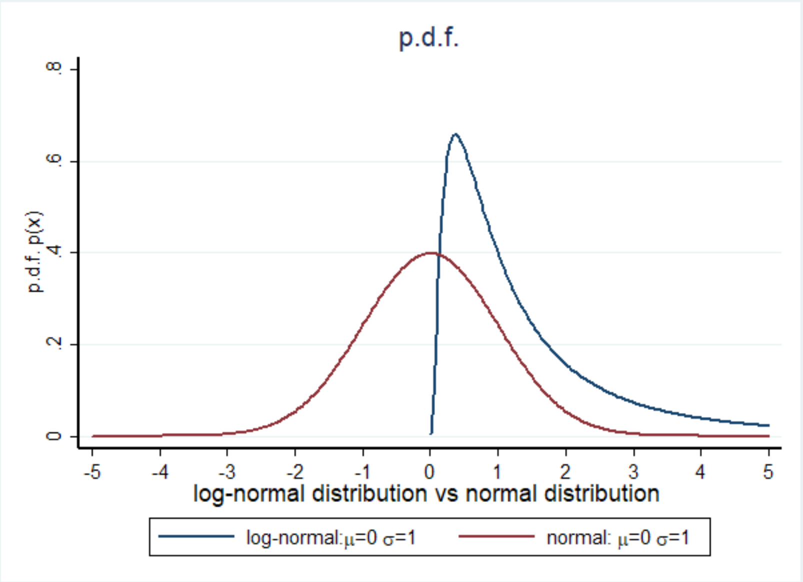
证明: $y = e^x$是严格增函数, 仅在$(0, \infty)$上取值, 反函数$x = ln(y)$连续可导, 故可用**定理2.6.1**.

当$y \leq 0$, $F_Y(y) = 0$, 所以$p_Y(y) = 0$

当$y > 0$,

$$\begin{aligned} p_Y(y) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(\ln y - \mu)^2}{2\sigma^2}\right\} \frac{1}{y} \\ &= \frac{1}{\sqrt{2\pi} y\sigma} \exp\left\{-\frac{(\ln y - \mu)^2}{2\sigma^2}\right\} \end{aligned}$$

正态分布和对数正态分布



伽玛分布的有用结论

定理2.6.4 设 $X \sim Ga(\alpha, \lambda)$, 则当$k > 0$ 时,
$Y = kX \sim Ga(\alpha, \lambda/k)$.

引理 设 $X \sim Ga(\alpha, \lambda)$, 则$2\lambda X \sim Ga(\alpha, 1/2) = \chi^2(2\alpha)$. 即任一伽玛分布可转化为卡方分布.

可用**定理2.6.1**证明, 见教材!

均匀分布的有用结论:

定理2.6.5 (一般分布和均匀分布间的关系): 设X的分布函数为$F_X(x)$, $X \sim F_X(x)$, 若$y = F_X(x)$为严格单调增的连续函数, 其反函数$x = F_X^{-1}(y)$存在, 则$Y = F_X(X)$服从$(0, 1)$上的均匀分布$\sim U(0, 1)$.

均匀分布的有用结论

证明: 求$Y=F_X(X)$的分布函数. 由于X的分布函数$F_X(x)$仅在$[0,1]$取值. 故

当$y<0$, 因为$\{F_X(X)\leqslant y\}$是不可能事件, 所以

$$F_Y(y)=P(Y\leqslant y)=P\left(F_X(X)\leqslant y\right)=0$$

当$0\leq y<1$,

$$\begin{aligned}F_Y(y)&=P(Y\leqslant y)=P\left(F_X(X)\leqslant y\right)\\&=P\left(X\leqslant F_X^{-1}(y)\right)=F_X\left(F_X^{-1}(y)\right)=y\end{aligned}$$

当$y\geq 1$, 因为$\{F_X(X)\leqslant y\}$是必然事件, 所以

$$F_Y(y)=P(Y\leqslant y)=P\left(F_X(X)\leqslant y\right)=1$$

所以$Y=F_X(X)$服从$(0,1)$上的均匀分布$\sim U(0,1)$. $Q.E.D.$

均匀分布的有用结论

任意连续随机变量X都可以通过分布函数$F(X)$和均匀分布随机变量U发生关系.

例: 若$X \sim Exp(\lambda)$, X的分布函数为$F(X) = 1 - e^{-\lambda x}$, 当$x$换为$X$后, 有

$$U = 1 - e^{-\lambda X}$$

或

$$X = \frac{1}{\lambda} ln \frac{1}{1-U}$$

后一式表明, 只要产生均匀分布的随机数, 就可以通过变换获得服从其他分布的随机数. 蒙特卡洛模拟方法中有重要应用!

$Y = g(X)$分布的其他求法

当定理2.6.1条件不满足时(如$g(x)$不严格单调), 可根据Y的分布函数定义($F_Y(y) = P(g(X) < y)$)去求解.

例2.6.3 设随机变量$X \sim N(0,1)$, 求$Y = X^2$的分布.

解: 先求$Y = X^2$的分布函数$F_Y(y)$.

由$Y = X^2 \geq 0$, 故当$Y \leq 0$时, 有$F_Y(y) = 0$, 所以$p_y(y) = 0$.
当$y > 0$时, 有

$$\begin{aligned} F_Y(y) =& P(Y < y) = P(X^2 \leq y) \\ =& P(-\sqrt{y} \leq X \leq \sqrt{y}) = 2\Phi(\sqrt{y}) - 1 \end{aligned}$$

所以$p.d.f.$为

$$p_Y(y) = \begin{cases} \varphi(\sqrt{y})y^{(-1/2)} = \dfrac{1}{\sqrt{2\pi}} y^{(-1/2)} e^{(-y/2)}, & y > 0 \\ 0, & y \leq 0. \end{cases}$$

$Y \sim \chi^2(1)$

第8次作业:

习题2.6中题目2, 5, 8, 10, 12, 14, 15

习题2.7中题目3, 5, 9, 10

§2.7 分布的其它特征数

- 矩 (moment)
- 变异系数 (coefficient of variation)
- 分位数 (quantile)
- 中位数 (median)
- 偏度系数 (coefficient of skewness)
- 峰度系数 (coefficient of kurtosis)

定义2.7.1

- k **阶原点矩(raw moments/moments about zero)**: $\mu_k = E(X^k), k = 1, 2, \ldots$

  注意: $\mu_1 = E(X)$.

- k **阶中心矩(central moments/moments about the mean)**: $\nu_k = E[X - E(X)]^k, k = 1, 2, \ldots$

  注意: $\nu_2 = Var(X)$.

因为$|X|^{k-1} \leq |X|^k + 1$, 故k阶矩存在时, $k-1$阶矩也存在(阶数更低时不会发散), 从而低于k的各阶矩都存在.

2.7.1 k 阶原点矩和中心矩

中心距和原点矩间的关系:

$$\nu_k = E(X - E(X))^k = E\left(X - \mu_1\right)^k = \sum_{i=0}^{k} \binom{k}{i} \mu_i \left(-\mu_1\right)^{k-i}$$

教材错误： pp121 例2.7.7中“前三阶中心矩”中一阶中心矩应永远为0, 教材中错误地将其写为了一阶原点矩$E(X) = a/10$.

2.7.2 变异系数 (coefficient of variation)

定义2.7.2 称$C_V = \frac{\sqrt{Var(X)}}{E(X)} = \frac{\sigma(X)}{E(X)}$ 为 X的变异系数.

- C_V 是无量纲的量
- 用于比较量纲不同的两个随机变量的波动大小

2.7.3 分位数(quantile)

定义2.7.3 设连续随机变量X的分布函数为$F(x)$, 密度函数为$p(x)$, 对任意$0 < p < 1$, 若 x_p 满足

$$F(x_p) = P(X \leq x_p) = \int_{-\infty}^{x_p} p(x)dx = p$$

则称 x_p 为此分布 p分位数.

- 亦称 x_p 为下侧 p分位数.

注意点

(1) 因为 X 小于等于 x_p 的可能性为 p, 所以 X 大于 x_p 的可能性为$1-p$.

(2) 对离散分布不一定存在p分位数.

上侧p分位数

若记 x_p' 为上侧 p分位数, 即$P(X \geq x_p') = 1 - F(x_p') = p$

则$x_p = x_{1-p}'$, $x_p' = x_{1-p}$

2.7.4 中位数(median)

定义2.7.4 称$p = 0.5$ 时的p分位数$x_{0.5}$ 为中位数. 即$x_{0.5}$满足

$$F(x_{0.5}) = P(X \leq x_{0.5}) = \int_{-\infty}^{x_{0.5}} p(x)dx = 0.5$$

中位数与均值

- 相同点: 都是反映随机变量的位置特征.
- 不同点: 含义不同.

思考: 讨论居民收入中位数和均值哪个更加合适地描述一国居民的收入水平(“被平均”)?

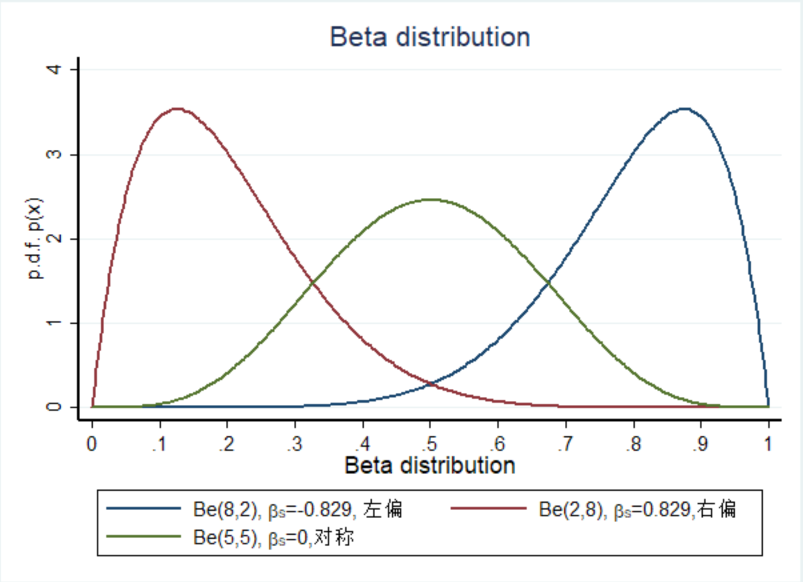
2.7.5 偏度系数 (coefficient of skewness)(略)

定义2.7.5 设随机变量X的前三阶矩存在, 则

$$\beta_S = \frac{\nu_3}{\nu_2^{3/2}} = \frac{E(X - E(X))^3}{[Var(X)]^{3/2}}$$

称为X的偏度系数, 简称偏度.

当$\beta_S > 0$时, 称该分布正偏(右偏); 当$\beta_S < 0$时, 称该分布负偏(左偏).


Beta distribution
p.d.f. p(x)
0
1
2
3
4
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1
Beta distribution
Be(8,2), βs=-0.829, 左偏
Be(2,8), βs=0.829,右偏
Be(5,5), βs=0,对称

算法见教材例2.7.7

2.7.6 峰度系数 (coefficient of kurtosis)(略)

定义2.7.6 设随机变量X的前四阶矩存在, 则

$$\beta_k = \frac{\nu_4}{\nu_2^2} - 3 = \frac{E(X - E(X))^4}{[Var(X)]^2} - 3$$

称为X的峰度系数, 简称峰度.

- $\beta_k > 0$表示标准化后比标准正态分布更尖, 尾部更粗
- $\beta_k < 0$表示标准化后比标准正态分布更平坦, 尾部更细

rarely used.

第8次作业:

习题2.6中题目2, 5, 8, 10, 12, 14, 15

习题2.7中题目3, 5, 9, 10