

数理统计

上海财经大学 统计与管理学院





Contents

第一章 概率论基础

1. 随机事件与概率

2. 随机变量及其分布

3. 多元随机变量及其分布

4. 大数定律和中心极限定理



一、古典概型

1. 古典概型常用公式

(1) 条件概率

设 A, B 是两个事件, 若 $P(B) > 0$, 则称

$$P(A|B) = \frac{P(AB)}{P(B)}$$

为“在事件 B 发生下事件 A 发生的条件概率”, 简称条件概率。它满足概率的三条公理。



一、古典概型

(2)乘法公式

i) 二元

若 $P(B) > 0$, 则

$$P(AB) = P(B)P(A|B)$$

ii) n 元

若 $P(A_1 A_2 \cdots A_n) > 0$, 则

$$P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \cdots P(A_n|A_1 A_2 \cdots A_{n-1}) .$$



一、古典概型

(3) 全概率公式

设 B_1, B_2, \dots, B_n 互不相容, 且 $\bigcup_{i=1}^n B_i = \Omega$, 如果 $P(B_i) > 0, i = 1, 2, \dots, n$, 则对任一事件 A 有

$$P(A) = \sum_{i=1}^n P(B_i)P(A|B_i).$$

全概率公式提供了计算复杂事件概率的一条有效途径。

一、古典概型

(4) 贝叶斯公式

设 B_1, B_2, \dots, B_n 互不相容, 且 $\bigcup_{i=1}^n B_i = \Omega$, 如果 $P(A) > 0, P(B_i) > 0, i = 1, 2, \dots, n$, 则

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{i=1}^n P(B_i)P(A|B_i)}, i = 1, 2, \dots, n$$

在贝叶斯公式中, 诸 $P(B_i)$ 成为 B_i 的先验 (试验以前) 概率, 而诸 $P(B_i|A)$ 称为 B_i 的后验 (试验以后) 概率, 它表示在 “事件 A 发生” 这个新信息后, 对 B_i 的概率作出的修正。

一、古典概型

❖ **例1** (2012数1, 3) 设 A, B, C 是随机事件, A 与 C 互不相容, $P(AB) = \frac{1}{2}$, $P(C) = \frac{1}{3}$, 求 $P(AB|\bar{C})$ 。

❖ **解:** 由 A 与 C 互不相容得 $P(AC) = 0$, 从而 $P(ABC) = 0$. 于是

$$P(AB|\bar{C}) = \frac{P(AB\bar{C})}{P(\bar{C})} = \frac{P(AB) - P(ABC)}{1 - P(C)} = \frac{\frac{1}{2} - 0}{1 - \frac{1}{3}} = \frac{3}{4}.$$

一、古典概型

❖ **例2** 一盒晶体管中有8只合格品、2只不合格品。从中不返回地一只一只取出，试求第二次取出合格品的概率。

❖ **解：**记事件 A_i 为“第 i 次取出合格品”， $i = 1, 2$ 。
用全概率公式

$$\begin{aligned} P(A_2) &= P(A_1)P(A_2|A_1) + P(\bar{A}_1)P(A_2|\bar{A}_1) \\ &= \frac{8}{10} \times \frac{7}{9} + \frac{2}{10} \times \frac{8}{9} = \frac{4}{5}. \end{aligned}$$

注：本例可以推广如下：袋中 a 个红球， b 个黑球，不返回一只一只取出，求第 k 次取出红球概率。小球模型也适用于彩票中奖问题分析。

一、古典概型

❖ **例3** 甲、乙两人对同一目标进行射击，命中率分别为0.6,0.5。试在下列两种情况下，分别求“已知目标被命中，则甲射中”的概率：

(1)在甲、乙两人中随机地挑选一人，由他射击一次；

(2)甲、乙两人独立地各射击一次。

❖ **解：**(1) 记事件 A_1 为“选中甲”， A_2 为“选中乙”， B 为“命中目标”。

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^2 P(B|A_i)P(A_i)} = \frac{0.6 \times 0.5}{0.6 \times 0.5 + 0.5 \times 0.5} = \frac{6}{11}.$$

一、古典概型

(2)记事件 A 为“甲射中”， B 为“乙射中”，两人独立射击，故有 $P(AB) = P(A)P(B)$ ，则所求概

率为

$$P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A) + P(B) - P(A)P(B)} = \frac{0.6}{0.6 + 0.5 - 0.6 \times 0.5} = \frac{3}{4}.$$

一、古典概型

(5)独立性

i) **两个事件的独立性** 如果 $P(AB) = P(A)P(B)$, 则称事件 A 与 B 独立。否则称 A 与 B 不独立或相依。

ii) **多个事件的独立性** 设有 n 个事件 A_1, A_2, \dots, A_n , 如果对任意的 $1 \leq i < j < k < \dots \leq n$, 以下等式成立

$$\begin{cases} P(A_i A_j) = P(A_i)P(A_j), \\ P(A_i A_j A_k) = P(A_i)P(A_j)P(A_k), \\ \dots \\ P(A_1 A_2 \dots A_n) = P(A_1)P(A_2) \dots P(A_n), \end{cases}$$

则称此 n 个事件 A_1, A_2, \dots, A_n 相互独立。

一、古典概型

iii) **n 重独立重复试验** 假如一个试验重复进行 n 次，并各次试验间相互独立，则称其为 n 重独立重复试验。假如一个试验只可能有两个结果： A 与 \bar{A} ，则称其为**伯努利试验**。假如一个伯努利试验重复进行 n 次，并各次试验间相互独立，则称其为 n 重伯努利试验。

❖ **例4** 证明小概率事件是可能发生的。

❖ **证明：** 设小概率事件 A 发生的概率为 ε ，即 $P(A) = \varepsilon$ ， $\varepsilon > 0$ ，则重复 n 次都不发生概率为 $P(\bar{A})^n = (1 - \varepsilon)^n$ ，则发生概率为 $P = 1 - (1 - \varepsilon)^n \xrightarrow{n \rightarrow \infty} 1$ ，即必发生。

注：吃路边摊和乱穿马路的人们要注意了！思考举例事件可以发生，但发生的概率为0。



二、随机变量概念

1. 随机变量(Random Variable, RV)

定义在样本空间 Ω 上的实值函数 $X = X(\omega)$ 称为随机变量。

i) 仅取有限个或可列个值的随机变量称为离散随机变量；

ii) 取值充满某个区间 (a, b) 的随机变量称为连续随机变量，这里 a 可为 $-\infty$ ， b 可为 $+\infty$ 。

二、随机变量概念

2. 分布函数(Cumulative Distribution Function, CDF)

设 X 是一个随机变量, 对任意实数 x , 称 $F(x) = P(X \leq x)$ 为 X 的分布函数, 记为 $X \sim F(x)$. 分布函数具有如下三条基本性质:

i) **单调性** $F(x)$ 是单调非减函数, 即对任意的 $x_1 < x_2$, 有

$$F(x_1) \leq F(x_2).$$

ii) **有界性** 对任意的 x , 有 $0 \leq F(x) \leq 1$, 且

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

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

$$\lim_{x \rightarrow x_0^+} F(x) = F(x_0), \text{ 即 } F(x_0 + 0) = F(x_0).$$

反之, 可以证明: 具有上述三条性质的函数 $F(x)$ 一定是某一个随机变量的分布函数。

注: 分布函数有两种定义 $P(X \leq x)$ 与 $P(X < x)$, 区别在于性质分别为右连续和左连续。本课件采用 $P(X \leq x)$ 定义分布函数。

二、随机变量概念

3. 离散随机变量的概率分布列

若离散随机变量 X 的可能取值是 $x_1, x_2, \dots, x_n, \dots$, 则称 X 取 x_i 的概率

$$p_i = p(x_i) = P(X = x_i), i = 1, 2, \dots, n,$$

为 X 的概率分布列, 简称分布列。分布列也可以用列表方式来表示:

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

分布列 $p(x_i)$ 具有如下两条基本性质:

i) 非负性: $p(x_i) \geq 0$; ii) 正则性: $\sum_{i=1}^{+\infty} p(x_i) = 1$.

二、随机变量概念

4. 连续随机变量的概率密度函数

记连续随机变量 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，简称为密度函数。密度函数 $f(x)$ 具有如下两条性质：

i) 非负性： $f(x) \geq 0$;

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

注：密度函数有常用的两种表达方式 $f(x)$ 与 $p(x)$ ，目前更为常用 $f(x)$ 与对应大写相匹配，此外 $F'(x) = f(x)$ 。

二、随机变量概念

5.常用公式

设随机变量 X 的分布函数为 $F(x)$ ，则可用 $F(x)$ 表示下列概率：

i) $P(X \leq a) = F(a);$

ii) $P(x < a) = F(a - 0);$

iii) $P(X > a) = 1 - P(X \leq a) = 1 - F(a);$

iv) $P(X = a) = P(X \leq a) - P(X < a)$

$$= F(a) - F(a - 0)$$

v) $P(X \geq a) = 1 - P(X < a) = 1 - F(a - 0);$

vi) $P(|X| < a) = P(-a < X < a)$

$$= P(X < a) - P(X \leq -a)$$

$$= F(a - 0) - F(-a).$$

二、随机变量概念

❖ **例5** (2011数1,3) 设 $F_1(x)$, $F_2(x)$ 为两个分布函数, 其相应的概率密度 $f_1(x)$, $f_2(x)$ 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:** 因为 $F_1'(x) = f_1(x)$, $F_2'(x) = f_2(x)$, 于是

$$\begin{aligned} f_1(x)F_2(x) + f_2(x)F_1(x) &= F_1'(x)F_2(x) + F_2'(x)F_1(x) \\ &= [F_1(x)F_2(x)]', \end{aligned}$$

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

且 $f_1(x)F_2(x) + f_2(x)F_1(x) \geq 0$, 故选(D)。

取 F_1, F_2 服从 $[0, 1], [0, 2]$ 区间的均匀分布, 反证前三项不成立。

二、随机变量概念

❖ **例6** 设随机变量 X 的密度函数为

$$\varphi(x) = Ae^{-|x|}, \quad -\infty < x < +\infty,$$

试求：(1) 系数 A ； (2) $P\{0 < X < 1\}$ ；
(3) X 的分布函数.

❖ **解：** (1) 由密度函数的性质

$$1 = \int_{-\infty}^{+\infty} Ae^{-|x|} dx = 2A \int_0^{+\infty} e^{-x} dx = -2Ae^{-x} \Big|_0^{+\infty} = 2A$$

$$\text{解得 } A = \frac{1}{2}.$$

$$(2) P\{0 < X < 1\} = \int_0^1 \frac{1}{2} e^{-|x|} dx = -\frac{1}{2} e^{-x} \Big|_0^1 = \frac{1-e^{-1}}{2}.$$

二、随机变量概念

$$(3) \quad X \text{ 的分布函数 } F(x) = \int_{-\infty}^x \varphi(t) dt = \int_{-\infty}^x \frac{1}{2} e^{-|t|} dt.$$

$$\text{当 } x < 0 \text{ 时, } F(x) = \int_{-\infty}^x \frac{1}{2} e^t dt = \frac{1}{2} e^x;$$

$$\begin{aligned} \text{当 } x \geq 0 \text{ 时, } F(x) &= \int_{-\infty}^x \frac{1}{2} e^{-|t|} dt = \int_{-\infty}^0 \frac{1}{2} e^t dt + \int_0^x \frac{1}{2} e^{-t} dt \\ &= 1 - \frac{1}{2} e^{-x}; \end{aligned}$$

$$\text{故 } F(x) = \begin{cases} \frac{1}{2} e^x, & x < 0, \\ 1 - \frac{1}{2} e^{-x}, & x \geq 0. \end{cases}$$

三、随机变量的数字特征

1. 数学期望(Expectation)

设随机变量 X 的分布用分布列 $p(x_i)$ 或用密度函数 $f(x)$ 表示, 若

$$\begin{cases} \sum_i |x_i| p(x_i) < +\infty, \text{当} X \text{为离散随机变量,} \\ \int_{-\infty}^{+\infty} |x| f(x) dx < +\infty, \text{当} X \text{为连续随机变量,} \end{cases}$$

$$\text{则称 } E(X) = \begin{cases} \sum_i x_i p(x_i), & \text{当} X \text{为离散随机变量,} \\ \int_{-\infty}^{+\infty} x f(x) dx, & \text{当} X \text{为连续随机变量} \end{cases}$$

为 X 的数学期望, 简称期望或均值, 且称 X 的数学期望存在. 否则称数学期望不存在.

三、随机变量的数字特征

数学期望的性质:

(1) X 的某一函数 $g(X)$ 的数学期望为

$$E[g(X)] = \begin{cases} \sum_i g(x_i)p(x_i), & \text{在离散场合,} \\ \int_{-\infty}^{+\infty} g(x)f(x)dx, & \text{在连续场合;} \end{cases}$$

(2) 若 c 是常数, 则 $E(c) = c$;

(3) 线性组合不变性(思考: n 无穷大时成立吗?)

$$E\left(\sum_{i=1}^n k_i X_i\right) = \sum_{i=1}^n k_i E(X_i), \quad \forall k_i \in R, i = 1, 2, \dots, n$$

(4) 对任意常数 a , 有 $E(aX) = aE(X)$;

(5) 对任意的两个函数 $g_1(x)$ 和 $g_2(x)$, 有

$$E[g_1(X) \pm g_2(X)] = E[g_1(X)] \pm E[g_2(X)].$$

三、随机变量的数字特征

❖ **例7** (2009数1) 设随机变量 X 的分布函数为 $F(x) = 0.3\Phi(x) + 0.7\Phi\left(\frac{x-1}{2}\right)$, 其中 $\Phi(x)$ 为标准正态分布函数, 则 $E(X) = \underline{\hspace{2cm}}$.

❖ **解:** 随机变量 X 的概率密度为

$$f(x) = F'(x) = 0.3\Phi'(x) + 0.7\Phi'\left(\frac{x-1}{2}\right) = 0.3\varphi(x) + 0.35\varphi\left(\frac{x-1}{2}\right),$$

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = 0.3 \int_{-\infty}^{+\infty} x\varphi(x) dx + 0.35 \int_{-\infty}^{+\infty} x\varphi\left(\frac{x-1}{2}\right) dx.$$

注意到 $\int_{-\infty}^{+\infty} x\varphi(x) dx = 0$ (**思考: 为什么**) , 以及

$$\int_{-\infty}^{+\infty} x\varphi\left(\frac{x-1}{2}\right) dx \xrightarrow{\frac{x-1}{2}=u} \int_{-\infty}^{+\infty} (1+2u)\varphi(u) \cdot 2du = 2 \int_{-\infty}^{+\infty} \varphi(u) du = 2.$$

故 $E(X) = 0.3 \times 0 + 0.35 \times 2 = 0.7$. 其中 $\phi(x)$ 为标准正态密度函数.

注: 本题还有其他解法. 灵活使用正态分布性质.

三、随机变量的数字特征

❖ **例8** 某种灯管的使用寿命 T 是一个随机变量（单位：小时），其密度函数为

$$f(t) = \begin{cases} 0.001e^{-0.001t}, & t > 0, \\ 0, & t \leq 0, \end{cases}$$

试求该种灯管的平均寿命。

❖ **解：**

$$\begin{aligned} E(T) &= \int_0^{+\infty} t \cdot 0.001e^{-0.001t} dt \\ &= -te^{-0.001t} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-0.001t} dt \\ &= -\frac{1}{0.001} e^{-0.001t} \Big|_0^{+\infty} = 1000. \end{aligned}$$

故其平均维修时间为1000小时。

注：该分布为失效率参数 $\lambda=0.001$ 的指数分布。

三、分布的其他数字特征

❖ **例8-1** 设国际市场上对某种出口商品每年的需求量 X (单位: 吨) 是随机变量, 它服从区间 $[2000, 4000]$ 上的均匀分布, 每销售一吨商品, 可为国家赚取外汇3万元; 若销售不出, 则每吨商品需贮存费1万元. 问应组织多少货源, 才能使国家收益最大.

❖ **解:** 设 Y 表示国家收益, 假设组织货源 t 吨, 显然应要求 $2000 \leq t \leq 4000$, 则

$$Y = g(X, t) = \begin{cases} 3t, & X \geq t \\ 3X - (t - X), & X < t \end{cases}$$

又由题设知 X 的概率密度函数为

$$f_X(x) = \begin{cases} \frac{1}{2000}, & 2000 \leq x \leq 4000 \\ 0, & \text{其他} \end{cases}$$

于是 Y 的期望为

三、分布的其他数字特征

$$\begin{aligned} E(Y) &= \int_{-\infty}^{+\infty} g(x, t) f_X(x) dx \\ &= \frac{1}{2000} \left[\int_{2000}^t (4x - t) dx + \int_t^{4000} 3t dx \right] \\ &= \frac{1}{1000} (-t^2 + 7000t - 4 \times 10^6). \end{aligned}$$

由 $\frac{dE(Y)}{dt} = \frac{1}{1000} (-2t + 7000) = 0$, 得 $t = 3500$.

由题意知应组织3500吨货源, 才能使国家收益最大.

注: 需要注意, 这里 t 是确定性的参变量, 而 X, Y 是随机变量.

三、随机变量的数字特征

2. 方差

随机变量 X 对其期望 $E(X)$ 的偏差平方的数学期望（设其存在）

$$\text{Var}(X) = E[X - E(X)]^2$$

称为 X 的方差，方差的正平方根 $\sigma(X) = \sigma_X = \sqrt{\text{Var}(X)}$ 称为 X 的标准差。

方差性质：

- (1) $\text{Var}(X) = E(X^2) - [E(X)]^2$;
- (2) 若 c 是常数，则 $\text{Var}(c) = 0$;
- (3) 若 a, b 常数，则 $\text{Var}(aX + b) = a^2 \text{Var}(X)$;
- (4) 若随机变量 X 的方差存在，则 $\text{Var}(X) = 0$ 的充要条件是 X 几乎处处为某个常数 a ，即 $P(X = a) = 1$ 。

注： $\text{Var}(X)$ 与 $D(X)$ 用于表示方差，课件将统一采用 $\text{Var}(X)$ 。

三、随机变量的数字特征

3. 随机变量的标准化

对任意随机变量 X ，如果 X 的数学期望和方差存在，则称

$$Z = \frac{X - E(X)}{\sqrt{Var(X)}}$$

为 X 的标准化随机变量，此时有

$$E(Z) = 0, Var(Z) = 1.$$

注：若 $X \sim N(\mu, \sigma^2)$ ，则 $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$. 正态分布的标准化是数据分析常用的一种方法。 **标准化优势是什么？**

三、随机变量的数字特征

❖ **例9** 设 X 为随机变量, 具有密度函数 $f(x) = 2x, 0 \leq x \leq 1$, 试求 $E(X)$ 和 $Var(X)$.

❖ **解:** 因为 $E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \int_0^1 x \cdot 2x dx = \frac{2}{3}$,

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^1 x^2 \cdot 2x dx = \frac{1}{2},$$

$$\text{则 } Var(X) = E(X^2) - [E(X)]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}.$$

❖ **注:** 此为 $Beta(2,1)$ 分布, 由 $Beta$ 分布的数字特征可直接知:

$$E(X) = \frac{a}{a+b} = \frac{2}{2+1} = \frac{2}{3}$$

$$Var(X) = \frac{ab}{(a+b)^2(a+b+1)} = \frac{2 \times 1}{(2+1)^2 \times (2+1+1)} = \frac{1}{18}.$$

三、随机变量的数字特征

- ❖ **例10** (1)若 $X \sim U(1, n)$, 求 $E(X)$ 和 $Var(X)$;
(2)若 $X \sim disU(1, n)$, 求 $E(X)$ 和 $Var(X)$.

- ❖ **解:** (1)若 $X \sim U(1, n)$, 则 $f(x) = \frac{1}{n-1}$, $1 \leq x \leq n$.

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \int_1^n x \cdot \frac{1}{n-1} dx = \frac{1}{n-1} \cdot \frac{n^2-1}{2} = \frac{n+1}{2}$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_1^n x^2 \cdot \frac{1}{n-1} dx = \frac{1}{n-1} \cdot \frac{n^3-1}{3} = \frac{n^2+n+1}{3}$$

$$\text{故 } Var(X) = E(X^2) - [E(X)]^2 = \frac{n^2+n+1}{3} - \left(\frac{n+1}{2}\right)^2 = \frac{(n-1)^2}{12}$$

三、随机变量的数字特征

(2)若 $X \sim \text{dis}U(1, n)$, 则 $f(x) = \frac{1}{n}$, $x = 1, 2, \dots, n$.

$$E(X) = \sum_{x=1}^n x \cdot f(x) = \sum_{x=1}^n x \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$\begin{aligned} E(X^2) &= \sum_{x=1}^n x^2 \cdot f(x) = \sum_{x=1}^n x^2 \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{n(n+1)(2n+1)}{6} \\ &= \frac{(n+1)(2n+1)}{6} \end{aligned}$$

$$\text{故 } \text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$$

❖ **注:** 注意区别均匀和离散均匀数字特征差异.

四、常用离散分布

1.二项分布(Binomial)

(1)若 X 的概率分布列为

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

则称 X 服从二项分布, 记为 $X \sim b(n, p)$, 其中 $0 < p < 1$.

(2)背景: n 重伯努利试验中成功的次数 X 服从二项分布 $b(n, p)$, 其中 p 为一次伯努利试验中成功发生的概率. 如产品抽检抽取次品的概率等.

(3) $n = 1$ 时的二项分布 $b(1, p)$ 称为二点分布, 或称0-1分布. 因为当 $X \sim b(1, p)$ 时, X 可表示一次伯努利试验中成功的次数, 它只能取值0与1.



四、常用离散分布

(4)二项分布 $b(n, p)$ 的数学期望和方差分别是

$$E(X) = np, \quad \text{Var}(X) = np(1 - p).$$

(5)若 $X \sim b(n, p)$, 则 $Y = n - X \sim b(n, 1 - p)$, 其中 $Y = n - X$ 是 n 重伯努利试验中失败的次数。

四、常用离散分布

2. 泊松分布(Poisson)

(1) 若 X 的概率分布列为

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

则称 X 服从泊松分布，记为 $X \sim P(\lambda)$ ，（或记为 $Pois(\lambda)$ ，其中参数 $\lambda > 0$ 。）

(2) 背景：单位时间（或单位面积、单位产品等）上某稀有事件（这里稀有事件是指不经常发生的事件）发生的次数常服从泊松分布 $P(\lambda)$ ，其中 λ 为该稀有事件发生的强度。如单位时间到银行办理业务的顾客数。

四、常用离散分布

(3) 泊松分布 $P(\lambda)$ 的数学期望和方差分别是

$$E(X) = \lambda, \quad Var(X) = \lambda.$$

(4) **二项分布的泊松近似 (泊松定理)**

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 (与试验次数 n 有关), 如果当 $n \rightarrow +\infty$ 时, 有 $np_n \rightarrow \lambda$, 则

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



四、常用离散分布*

3.超几何分布

4.几何分布

5.负二项分布

(略)

四、常用离散分布

名称与记号	分布列	期望	方差
0-1分布 $b(1, p)$	$p_k = p^k(1-p)^{1-k}, k = 0, 1$	p	$p(1-p)$
二项分布 $b(n, p)$	$p_k = \binom{n}{k} p^k(1-p)^{n-k}, k = 0, 1, \dots, n.$	np	$np(1-p)$
泊松分布 $P(\lambda)$	$p_k = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, \dots$	λ	λ
超几何分布* $h(n, N, M)$	$p_k = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}, k = 0, 1, \dots, r, r = \min\{M, n\}$	$\frac{nM}{N}$	$\frac{nM(N-M)(N-n)}{N^2(N-1)}$
几何分布* $Ge(p)$	$p_k = (1-p)^{k-1}p, k = 1, 2, \dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
负二项分布* $Nb(r, p)$	$p_k = \binom{k-1}{r-1} (1-p)^{k-r} p^r, k = r, r+1, \dots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$

四、常用离散分布

❖ **例11** 设 X 服从参数为 $\sqrt{5}$ 的 $poisson$ 分布, 则使得 $P\{X = k\}$ 达到最大的 k 为多少?

❖ **解:** 由题知 $X \sim P(\sqrt{5})$,

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

$P\{X = k\}$ 达到最大, 即
$$\begin{cases} P\{X = k\} \geq P\{X = k + 1\} \\ P\{X = k\} \geq P\{X = k - 1\} \end{cases}$$

代入并化简可得

$$\begin{cases} \frac{\sqrt{5}^k}{k!} \geq \frac{\sqrt{5}^{k+1}}{(k+1)!} \\ \frac{\sqrt{5}^k}{k!} \geq \frac{\sqrt{5}^{k-1}}{(k-1)!} \end{cases}$$

解得 $\sqrt{5} - 1 \leq k \leq \sqrt{5}$, 又 k 为整数, 所以 $k = 2$.

四、常用离散分布

❖ **例12** (2008数1,3) 设随机变量 X 服从参数为1的泊松分布, 则 $P\{X = E(X^2)\} = \underline{\hspace{2cm}}$.

❖ **解:** 由 X 服从参数为1的泊松分布知

$$E(X) = \text{Var}(X) = 1,$$

因此

$$E(X^2) = \text{Var}(X) + [E(X)]^2 = 2.$$

所以 $P\{X = E(X^2)\} = P\{X = 2\} = \frac{1}{2}e^{-1}$, 应填 $\frac{1}{2}e^{-1}$.

五、常用连续分布

1.正态分布

(1)若 X 的密度函数（如下图）和分布函数分别为

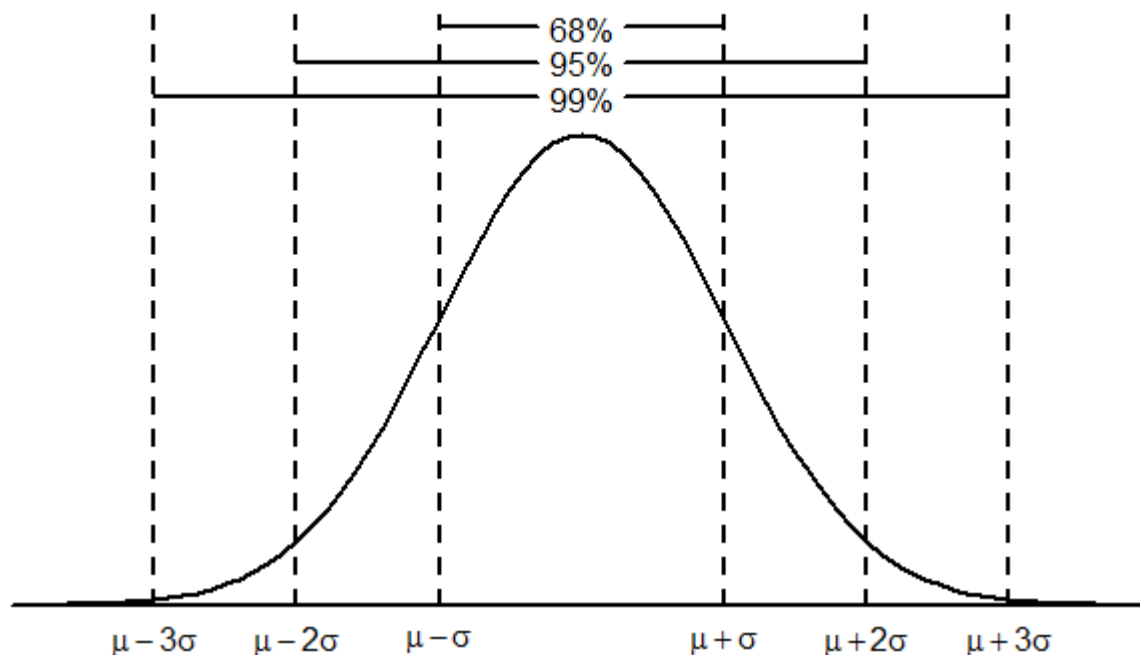
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty,$$

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

则称 X 服从正态分布，记作 $X \sim N(\mu, \sigma^2)$ ，其中参数
 $-\infty < \mu < \infty, \sigma > 0$.

五、常用连续分布

正态分布常用的概率



(2)背景：一个变量若是由大量微小的、独立的随机因素的叠加结果，则此变量一定是正态变量（服从正态分布的变量）。测量误差就是由量具偏差、测量环境的影响、测量技术的影响、测量人员的心理影响等等随机变量叠加而成的，所以测量误差常认为服从正态分布。最早由高斯研究误差理论发现。

五、常用连续分布

(3)关于参数 μ ：

- 正态分布的数学期望，即 $E(X) = \mu$ ，称 μ 为正态分布的位置参数(Location parameter).
- μ 是正态分布的对称中心，密度曲线在 μ 的两侧覆盖面积都为0.5，所以 μ 也是正态分布的中位数（见后面§ 2.7）.
- 若 $X \sim N(\mu, \sigma^2)$ ，则 X 在离 μ 愈近取值的可能性愈大，离 μ 愈远取值的可能性愈小.

关于参数 σ^2 ：

- 正态分布的方差，即 $Var(X) = \sigma^2$.
- σ 是正态分布的标准差， σ 愈小，正态分布愈集中； σ 愈大，正态分布愈分散. σ 又称为正态分布的尺度参数.
- 若 $X \sim N(\mu, \sigma^2)$ ，则其密度函数 $f(x)$ 在 $\mu \pm \sigma$ 处有两个拐点.

五、常用连续分布

(4)称 $\mu = 0, \sigma = 1$ 时的正态分布 $N(0,1)$ 为标准正态分布.记 Z 为标准正态变量, $\varphi(z)$ 和 $\Phi(z)$ 为标准正态分布的密度函数和分布函数. $\varphi(z)$ 和 $\Phi(z)$ 满足:

- $\varphi(-z) = \varphi(z)$;
- $\Phi(-z) = 1 - \Phi(z)$.对 $u > 0$, $\Phi(u)$ 的值有表可查。
- $P(|Z| < z) = \Phi(z) - \Phi(-z) = 2\Phi(z) - 1$.

(5)标准化变换: 若 $X \sim N(\mu, \sigma^2)$, 则

$Z = (X - \mu)/\sigma \sim N(0,1)$, 其中 $Z = (X - \mu)/\sigma$ 称为 X 的标准化变量。

五、常用连续分布

(6)若 $X \sim N(\mu, \sigma^2)$, 则对任意实数 a 与 b ,

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

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

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

可见, 涉及正态变量的概率计算, 一般先做标准化, 再查标准正态分布的分布函数表获得.

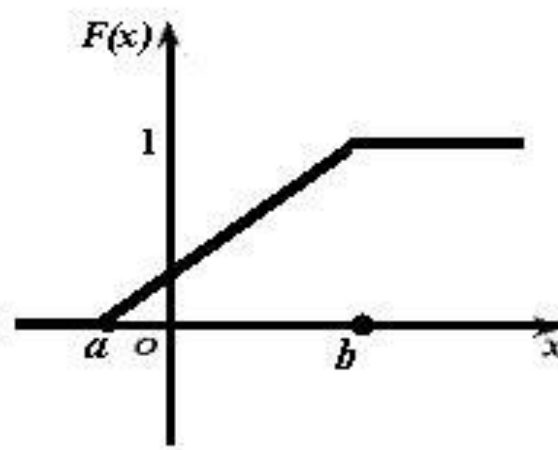
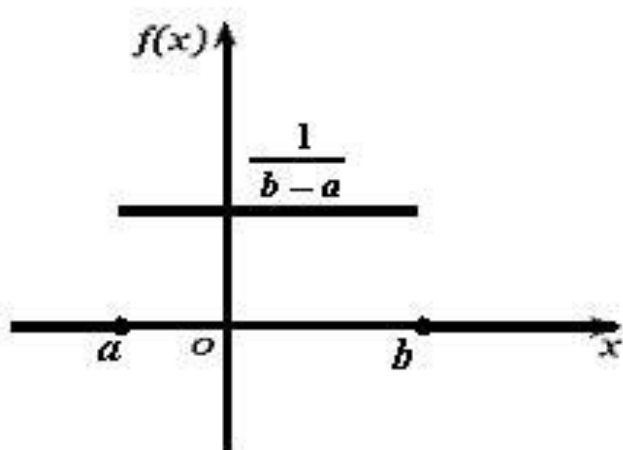
(7)正态分布的 3σ 原则: 设 $X \sim N(\mu, \sigma^2)$, 则

$$P(|X - \mu| < k\sigma) = \Phi(k) - \Phi(-k) = \begin{cases} 0.6826, k = 1 \\ 0.9545, k = 2 \\ 0.9973, k = 3. \end{cases}$$

五、常用连续分布

2. 均匀分布

(1) 若 X 的密度函数和分布函数 (如下图) 分别为



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

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

五、常用连续分布

(2)背景：向区间 (a, b) 随机投点，落点坐标 X 一定服从均匀分布 $U(a, b)$.这里“随机投点”是指：点落在任意相等长度的小区间上的可能性是相等的.

(3)均匀分布 $U(a, b)$ 的数学期望和方差分别是

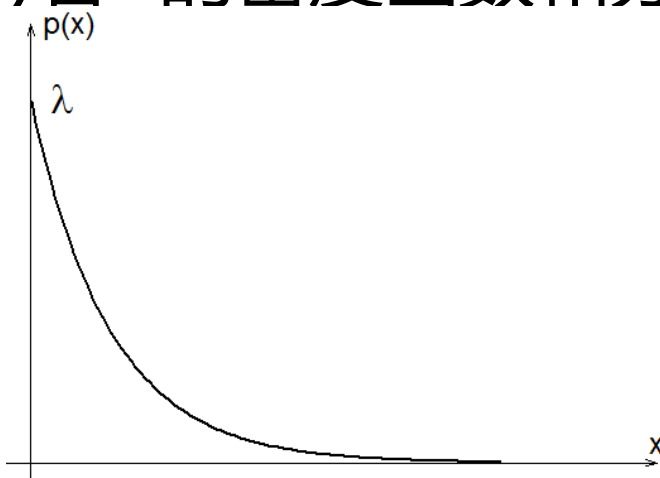
$$E(X) = \frac{a + b}{2}, \quad \text{Var}(X) = \frac{(b - a)^2}{12}$$

(4)称区间 $(0, 1)$ 上的均匀分布 $U(0, 1)$ 为标准均匀分布，它是导出其他分布随机数的桥梁.感兴趣的可查阅Ross著作*Simulation*。

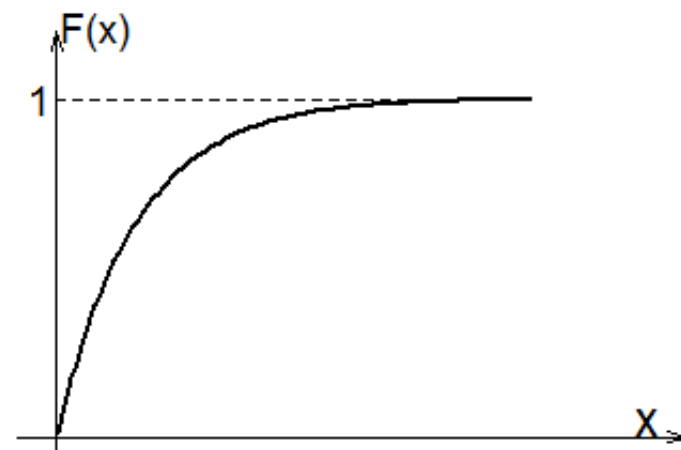
五、常用连续分布

3. 指数分布

(1) 若 X 的密度函数和分布函数（如下图）分别为



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



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

则称 X 服从指数分布，记作 $X \sim \text{Exp}(\lambda)$ ，其中参数 $\lambda > 0$ 。

五、常用连续分布

(2)背景：若一个元器件（或一台设备、或一个系统）遇到外来冲击时即告失效，则首次冲击来到的时间 X （寿命）服从指数分布.很多产品的寿命可认为服从或近似服从指数分布.

(3)指数分布 $Exp(\lambda)$ 的数学期望和方差分别为

$$E(X) = \frac{1}{\lambda}, Var(X) = \frac{1}{\lambda^2}$$

(4)指数分布的无记忆性：若 $X \sim Exp(\lambda)$ ，则对任意 $s > 0, t > 0$ ，有 $P(X > s + t | X > s) = P(X > t)$.



五、常用连续分布

名称与记号	密度函数 $f(x)$	期望	方差
正态分布 $N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, -\infty < x < +\infty$	μ	σ^2
均匀分布 $U(a, b)$	$\frac{1}{b-a}, a < x < b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $Exp(\lambda)$	$\lambda e^{-\lambda x}, x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
伽玛分布 $Ga(\alpha, \lambda)$	$\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, x \geq 0$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
$\chi^2(n)$ 分布	$\frac{x^{n/2-1} e^{-x/2}}{\Gamma(n/2) 2^{n/2}}, x \geq 0$	n	$2n$



五、常用连续分布*

名称与记号	密度函数 $f(x)$	期望	方差
贝塔分布 $Be(a, b)$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1},$ $0 < x < 1$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$
对数正态分布 $LN(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\},$ $x > 0$	$e^{\mu+\sigma^2/2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$
柯西分布 $Cau(\mu, \lambda)$	$\frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x - \mu)^2},$ $-\infty < x < +\infty$	不存在	不存在
韦布尔分布 $Wei(m, \eta)$	$F(x) = 1 - \exp\left\{-\left(\frac{x}{\eta}\right)^m\right\},$ $x > 0$	$\eta\Gamma(1 + \frac{1}{m})$	$\eta^2 \left[\Gamma\left(1 + \frac{2}{m}\right) - \Gamma^2\left(1 + \frac{1}{m}\right) \right]$

五、常用连续分布

❖ **例13** 设 K 服从 $(1,6)$ 上的均匀分布, 求方程 $x^2 + Kx + 1 = 0$ 有实根的概率.

❖ **解:** 方程 $x^2 + Kx + 1 = 0$ 有实根的充要条件是
 $\{K^2 - 4 \geq 0\} = \{K \leq -2\} \cup \{K \geq 2\}$

而 $K \sim U(1,6)$, 因此所求概率为

$$P(K \leq -2) + P(K \geq 2) = 0 + \int_2^6 \frac{1}{5} dx = \frac{4}{5}.$$

五、常用连续分布

❖ **例14** (2006数1,3) 设随机变量 X 服正态分布 $N(\mu_1, \sigma_1^2)$, Y 服从正态分布 $N(\mu_2, \sigma_2^2)$, 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) $\mu_1 > \mu_2$

❖ **解:** 由题设可得 $P\left\{\frac{|X - \mu_1|}{\sigma_1} < \frac{1}{\sigma_1}\right\} > P\left\{\frac{|Y - \mu_2|}{\sigma_2} < \frac{1}{\sigma_2}\right\}$,

则 $2\Phi\left(\frac{1}{\sigma_1}\right) - 1 > 2\Phi\left(\frac{1}{\sigma_2}\right) - 1$, 即 $\Phi\left(\frac{1}{\sigma_1}\right) > \Phi\left(\frac{1}{\sigma_2}\right)$,

其中 $\Phi(x)$ 是标准正态分布的分布函数.

又 $\Phi(x)$ 是严格单调增加函数, 所以 $\frac{1}{\sigma_1} > \frac{1}{\sigma_2}$, 即 $\sigma_1 < \sigma_2$. 选A.

五、常用连续分布

❖ **例15** 某校自主招生，共有10000人报考，假设考试成绩服从正态分布，且已知90分以上有359人，60分以下有1151人.现按考试成绩从高分到低分依次录用250人，试问被录取者中最低分为多少？

❖ **解：** 记 X 为考试成绩，则 $X \sim N(\mu, \sigma^2)$.由频率估计概率知

$$0.0359 = P(X > 90) = 1 - \Phi\left(\frac{90 - \mu}{\sigma}\right),$$

$$0.1151 = P(X < 60) = \Phi\left(\frac{60 - \mu}{\sigma}\right),$$

上面两式可改写为 $0.9641 = \Phi\left(\frac{90 - \mu}{\sigma}\right), 0.8849 = \Phi\left(\frac{\mu - 60}{\sigma}\right)$.

五、常用连续分布

再查表得 $\frac{90-\mu}{\sigma} = 1.8, \frac{\mu-60}{\sigma} = 1.2$

由此解得 $\mu = 72, \sigma = 10$. 设被录用者中最低分为 k , 则由

$$0.025 = P(X \geq k) = 1 - \Phi\left(\frac{k-72}{10}\right), \text{ 或 } 0.975 = \Phi\left(\frac{k-72}{10}\right),$$

查表得 $(k - 72)/10 \geq 1.96$, 从中解得 $k \geq 91.6$, 因此取被录用者中最低分为 91.6 分即可.

注: 当 $p < 0.5$ 时, 满足等式 $\Phi(x) = p$ 的 x 在标准正态分布函数表上不易查得, 故改写此式为 $\Phi(-x) = 1 - p > 0.5$, 即可查得 $-x$.

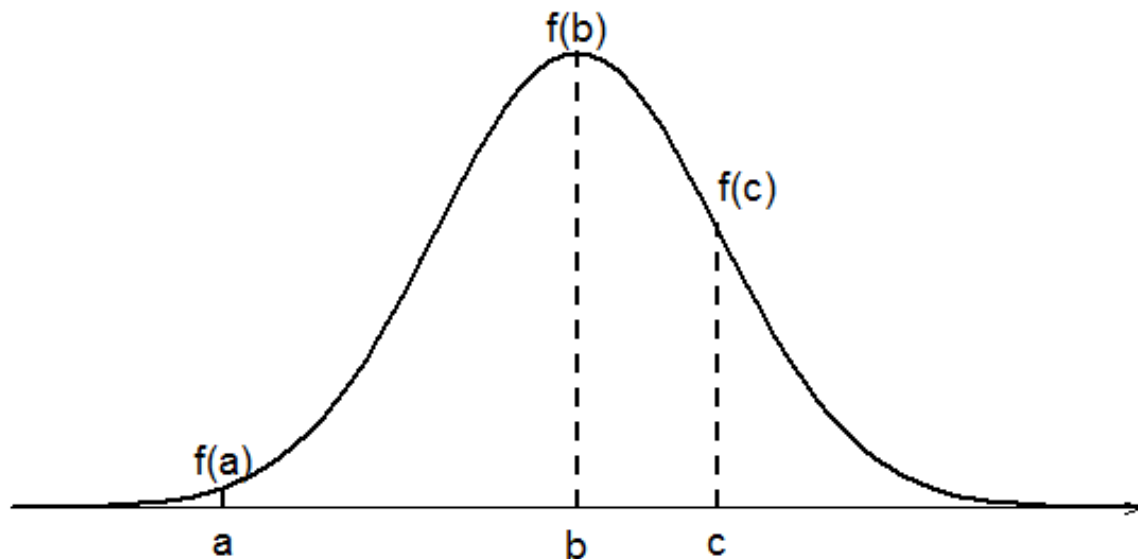
五、常用连续分布

❖ **例16** 设随机变量 $X \sim N(\mu, \sigma^2)$, $f(x)$ 为其密度函数, 且 a, b, c 满足: $a < b < c, f(a) < f(c) < f(b)$, 则

(A) $\frac{a+c}{2} < \mu < \frac{b+c}{2}$ (B) $\frac{a+b}{2} < \mu < \frac{a+c}{2}$

(C) $a < \mu < b$ (D) $b < \mu < c$

❖ **解:** 选A. 结合画图用排除法, 可得结论.



六、随机变量函数的分布

1. 设连续随机变量 X 的密度函数为 $f_X(x)$, $Y = g(X)$.

(1) 若 $y = g(x)$ 严格单调, 其反函数 $h(y)$ 有连续导数, 则 $Y = g(X)$ 的密度函数为

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

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

(2) 若 $y = g(x)$ 在不相重叠的区间 I_1, I_2, \dots 上逐段严格单调, 其反函数 $h_1(y), h_2(y), \dots$ 有连续导数, 则 $Y = g(X)$ 的密度函数为

$$f_Y(y) = \sum_i f_X(h_i(y))|h'_i(y)|$$

2. 正态变量的线性变换仍为正态变量 若 X 服从正态分布 $N(\mu, \sigma^2)$, 则当 $a \neq 0$ 时, 有 $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

六、随机变量函数的分布*

3.对数正态分布 (略)

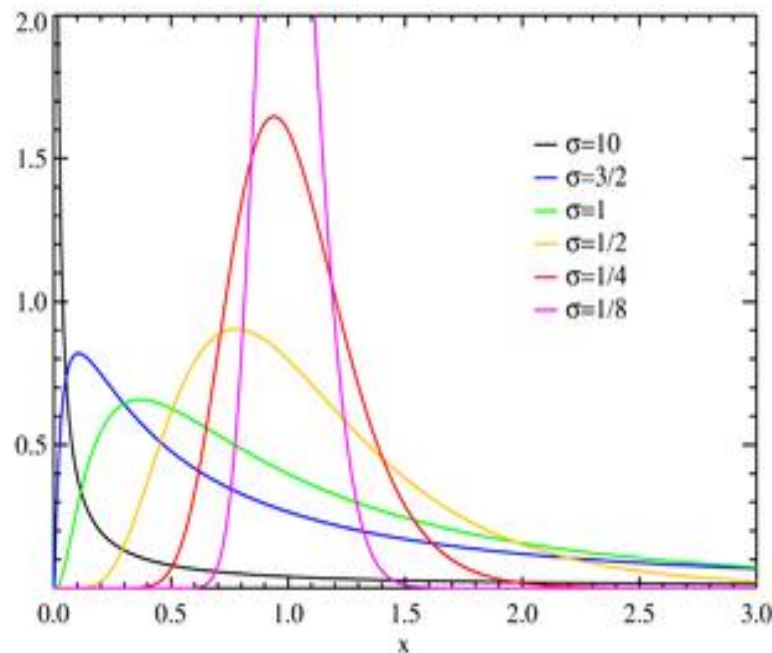
(1)若 X 的密度函数为

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

则称 X 服从对数正态分布, 记为 $X \sim LN(\mu, \sigma^2)$, 其中 $-\infty < \mu < +\infty, \sigma > 0$.对数正态分布的密度函数如图

(2)若 $X \sim LN(\mu, \sigma^2)$, 则 $E(X) = e^{\mu + \sigma^2/2}$,
 $Var(X) = (e^{\sigma^2} - 1) \cdot e^{2\mu + \sigma^2}$.

(3)若 $X \sim LN(\mu, \sigma^2)$, 则 $Y = \log X \sim N(\mu, \sigma^2)$



六、随机变量函数的分布

❖ **例17** 设随机变量 X 服从参数为2的指数分布，试证 $Y_1 = e^{-2X}$ 和 $Y_2 = 1 - e^{-2X}$ 都服从区间 $(0,1)$ 上的均匀分布。

❖ **证：** 因为 X 的密度函数为

$$f_X(x) = \begin{cases} 2e^{-2x}, & x > 0, \\ 0, & \text{其他}, \end{cases}$$

又因为 Y_1 的可能取值范围是 $(0,1)$ ，且 $y_1 = e^{-2x}$ 是严格单调减函数，其反函数为 $x = h(y_1) = -0.5\ln y_1$ ，所以 Y_1 的密度函数为

六、随机变量函数的分布

$$\begin{aligned} & f_{Y_1}(y) \\ = & \begin{cases} f_X(-0.5 \ln y_1) | -0.5/y_1 |, & 0 < y_1 < 1 \\ 0, & \text{其他} \end{cases} \\ = & \begin{cases} 1, & 0 < y_1 < 1, \\ 0, & \text{其他} \end{cases} \end{aligned}$$

即 $Y_1 \sim U(0,1)$. 又由前面第4题, 知 $Y_2 = 1 - e^{-2X} = 1 - Y_1$ 也服从区间 $(0,1)$ 上的均匀分布. 结论得证.

注: 若 $Z \sim F(x)$, 则 $F(Z) \sim U(0,1)$.

七、多维随机变量及其分布

1.联合分布函数 对任意的 n 个实数 x_1, x_2, \dots, x_n , n 个事件

$X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n$ 同时发生的概率

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

称为 n 维随机变量 (X_1, X_2, \dots, X_n) 的联合分布函数.

二维随机变量 (X, Y) 的联合分布函数 $F(x, y) = P(X \leq x, Y \leq y)$ 具有如下四条基本性质:

(1)单调性 $F(x, y)$ 分别对 x 或 y 是单调不减的.

(2)有界性 对任意的 x 和 y , 有 $0 \leq F(x, y) \leq 1$, 且

$$F(-\infty, y) = F(x, +\infty) = 0, \quad F(+\infty, +\infty) = 1.$$

(3)右连续性 对每个变量都是右连续的, 即

$$F(x+0, y) = F(x, y), \quad F(x, y+0) = F(x, y).$$

七、多维随机变量及其分布

(4)非负性 对任意的 $a < b, c < d$ 有

$$P(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c) \geq 0.$$

2.联合分布列

3.联合密度函数 如果存在二元非负函数 $f(x, y)$, 使得二维随机变量 (X, Y) 的分布函数 $F(x, y)$ 可表示为

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du,$$

则称 (X, Y) 为二维连续随机变量, 称 $f(x, y)$ 为 (X, Y) 的联合密度函数.

联合密度函数的基本性质:

(1)非负性: $f(x, y) \geq 0$; (2)正则性: $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1.$

七、多维随机变量及其分布*

4. 条件分布

(1) 离散型随机变量的条件分布

i) 条件分布列 对一切使 $P(Y = y_j) = p_{\cdot j} = \sum_{i=1}^{+\infty} p_{ij} > 0$ 的 y_j , 称

$$p_{i|j} = P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}, i = 1, 2, \dots$$

为给定 $Y = y_j$ 条件下 X 的条件分布列.

同理, 对一切使 $P(X = x_i) = p_{i\cdot} = \sum_{j=1}^{+\infty} p_{ij} > 0$ 的 x_i , 称

$$p_{j|i} = P(Y = y_j | X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i\cdot}}, j = 1, 2, \dots$$

为给定 $X = x_i$ 条件下 Y 的条件分布列.

七、多维随机变量及其分布*

ii) 条件分布函数 给定 $Y = y_j$ 条件下 X 的条件分布函数为

$$F(x|y_j) = \sum_{x_i \leq x} P(X = x_i | Y = y_j) = \sum_{x_i \leq x} p_{i|j},$$

给定 $X = x_i$ 条件下 Y 的条件分布函数为

$$F(y|x_i) = \sum_{y_j \leq y} P(Y = y_j | X = x_i) = \sum_{y_j \leq y} p_{j|i},$$

七、多维随机变量及其分布*

(2)连续随机变量的条件分布

i)条件密度函数

对于给定的 x , 若 $f_X(x) > 0$, 则称

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}, \quad -\infty < y < +\infty$$

为 Y 关于 $X = x$ 的**条件密度函数**.

类似地, 若 $f_Y(y) > 0$, 则称

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad -\infty < x < +\infty$$

为 X 关于 $Y = y$ 的**条件密度函数**.

注: 上式为条件概率的密度函数版本。

七、多维随机变量及其分布*

ii) 条件分布函数

对一切使 $f_Y(y) > 0$ 的 y , 给定 $Y = y$ 条件下 X 的条件分布函数为

$$F(x|y) = \int_{-\infty}^x f(u|y) du = \int_{-\infty}^x \frac{f(u, y)}{f_Y(y)} du.$$

类似对一切使 $f_X(x) > 0$ 的 x , 给定 $X = x$ 条件下 Y 的条件分布函数分别为

$$F(y|x) = \int_{-\infty}^y f(v|x) dv = \int_{-\infty}^y \frac{f(x, v)}{f_X(x)} dv.$$

注：条件密度函数同样满足密度函数的所有性质。

七、多维随机变量及其分布*

5. 多项分布

在 n 次独立重复试验中, 如果每次试验有 r 个可能结果: A_1, A_2, \dots, A_r , 且每次试验中 A_i 发生的概率为 $p_i = P(A_i), i = 1, 2, \dots, r. p_1 + p_2 + \dots + p_r = 1$. 记 X_i 为 n 次独立重复试验中 A_i 出现的次数, $i = 1, 2, \dots, r$. 则 (X_1, X_2, \dots, X_r) 服从**多项分布**, 又称 **r 项分布**, 记为 $M(n, p_1, p_2, \dots, p_r)$, 其联合分布列为

$$\begin{aligned} &P(X_1 = n_1, X_2 = n_2, \dots, X_r = n_r) \\ &= \frac{n!}{n_1! n_2! \dots n_r!} p_1^{n_1} p_2^{n_2} \dots p_r^{n_r} \end{aligned}$$

其中 $n = n_1 + n_2 + \dots + n_r$.

当 $r = 2$ 时, 即为二项分布.



七、多维随机变量及其分布*

6.二元均匀分布

设 (X, Y) 的联合密度函数为

$$f(x, y) = \begin{cases} \frac{1}{S(D)}, & (x, y) \in D \\ 0, & \text{其他} \end{cases}$$

其中 D 为一平面有界区域, $S(D)$ 为 D 的面积, 则称 (X, Y) 服从二元均匀分布, 常记为 $(X, Y) \sim U(D)$.

注: 本课件采用“多维”和“多元”表达同一概念。

七、多维随机变量及其分布

7.二元正态分布

如果二维随机变量 (X, Y) 的联合密度函数*为

$$f(x, y) = \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\}$$
$$-\infty < x, y < +\infty$$

则称 (X, Y) 服从二元正态分布, 记为 $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.
其中 $E(X) = \mu_1$, $E(Y) = \mu_2$; $Var(X) = \sigma_1^2$, $Var(Y) = \sigma_2^2$; $-1 \leq \rho \leq 1$.

七、多维随机变量及其分布*

❖ **例18** (2010数1,3) 设二维随机变量 (X, Y) 的概率密度为 $f(x, y) = Ae^{-2x^2+2xy-y^2}$, $-\infty < x < +\infty$, $-\infty < y < +\infty$ 求常数 A 以及条件概率密度 $f_{Y|X}(y|x)$.

❖ **解:** 设由联合概率密度的性质有

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = A \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-2x^2+2xy-y^2} dx dy \\ &= A \int_{-\infty}^{+\infty} e^{-x^2} dx \int_{-\infty}^{+\infty} e^{-(y-x)^2} d(y-x) \\ &= A \times \sqrt{\pi} \times \sqrt{\pi} = A\pi, \end{aligned}$$

得常数 $A = \frac{1}{\pi}$, 即 $f(x, y) = \frac{1}{\pi} e^{-2x^2+2xy-y^2}$.

七、多维随机变量及其分布*

$$\begin{aligned} \text{则 } f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy = \int_{-\infty}^{+\infty} \frac{1}{\pi} e^{-2x^2+2xy-y^2} dy \\ &= \frac{1}{\pi} e^{-x^2} \int_{-\infty}^{+\infty} e^{-(y-x)^2} d(y-x) = \frac{1}{\sqrt{\pi}} e^{-x^2}, \end{aligned}$$

因此

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{1}{\sqrt{\pi}} e^{-x^2+2xy-y^2}, \quad -\infty < x < +\infty, \\ -\infty < y < +\infty.$$

注：本题充分利用积分 $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$ 简化计算，思考该结论如何得到。

七、多维随机变量及其分布

❖ **例19** (2007数1,3) 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 2 - x - y, & 0 < x < 1, \quad 0 < y < 1, \\ 0, & \text{其他,} \end{cases}$$

(1) 求 $P\{X > 2Y\}$;

(2) 求 $Z = X + Y$ 的概率密度 $f_Z(z)$.

❖ **解:** (1) $P\{X > 2Y\} = \iint_{x>2y} f(x, y) dx dy$

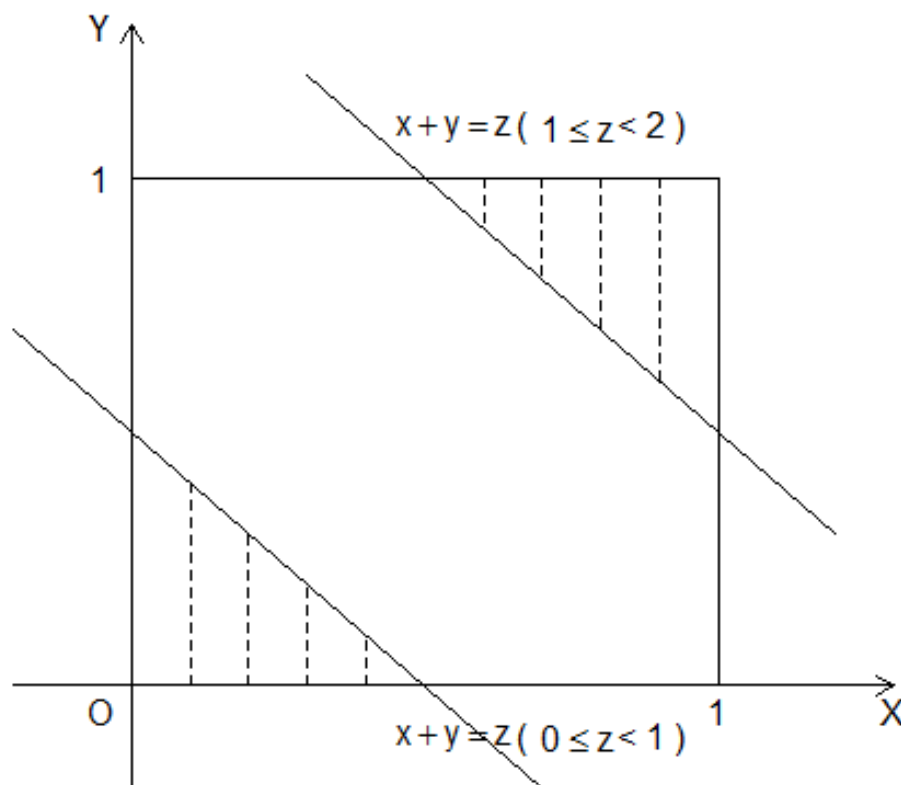
$$= \int_0^{\frac{1}{2}} dy \int_{2y}^1 (2 - x - y) dx = \frac{7}{24}.$$



七、多维随机变量及其分布

(2)先求Z的分布函数:

$$F_Z(z) = P\{X + Y \leq z\} = \iint_{x+y \leq z} f(x, y) dx dy.$$



七、多维随机变量及其分布

当 $z < 0$ 时, $F_Z(z)=0$;

$$\begin{aligned}\text{当 } 0 \leq z < 1 \text{ 时, } F_Z(z) &= \iint_{x+y \leq z} f(x, y) dx dy \\ &= \int_0^z dy \int_0^{z-y} (2-x-y) dx = z^2 - \frac{1}{3}z^3;\end{aligned}$$

$$\begin{aligned}\text{当 } 1 \leq z < 2 \text{ 时, } F_Z(z) &= 1 - \iint_{x+y > z} f(x, y) dx dy \\ &= 1 - \int_{z-1}^1 dy \int_{z-y}^1 (2-x-y) dx = 1 - \frac{1}{3}(2-z)^3;\end{aligned}$$

当 $z \geq 2$ 时, $F_Z(z) = 1$.

故 $Z = X + Y$ 的概率密度

$$f_Z(z) = F'_Z(z) = \begin{cases} 2z - z^2, & 0 < z < 1 \\ (2-z)^2, & 1 \leq z < 2 \\ 0, & \text{其他.} \end{cases}$$

八、边际分布与随机变量的独立性

1. 边际分布函数 若二维随机变量 (X, Y) 的联合分布函数为 $F(X, Y)$, 则称

$$F_X(x) = F(x, +\infty) = \lim_{y \rightarrow +\infty} F(x, y), -\infty < x < +\infty$$

为 X 的边际分布, 称

$$F_Y(y) = F(+\infty, y) = \lim_{x \rightarrow +\infty} F(x, y), -\infty < y < +\infty$$

为 Y 的边际分布.

2. 边际分布列 若二维离散随机变量 (X, Y) 的联合分布列为 $\{p_{ij}\}$, 则称 $p_{i\cdot} = \sum_{j=1}^{+\infty} p_{ij}$, $i = 1, 2, \dots$ 为 X 的边际分布列, 称 $p_{\cdot j} = \sum_{i=1}^{+\infty} p_{ij}$, $j = 1, 2, \dots$ 为 Y 的边际分布列.

注: 边际分布也称作边缘分布(Marginal distribution).

八、边际分布与随机变量的独立性

3. 边际密度函数 若二维连续随机变量 (X, Y) 的联合分布函数为 $f(x, y)$, 则称 $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy, -\infty < x < +\infty$ 为 X 的边际密度函数, 称

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx, -\infty < y < +\infty$$

为 Y 的边际密度函数.

4. 注解

- 由高维联合分布可以获得低维的边际分布, 反之不一定.
- 不同的联合分布可以有相同边际分布.
- 多项分布的边际分布仍为低维的多项分布或二项分布.
- 二维正态分布的边际分布为一维正态分布.

八、边际分布与随机变量的独立性

5. 随机变量间的独立性

(1) 设 n 维随机变量 (X_1, X_2, \dots, X_n) 的联合分布函数为 $F(x_1, x_2, \dots, x_n)$, 且 $F_{X_i}(x_i)$ 为 X_i 的边际分布函数. 如果对任意 n 个实数 x_1, x_2, \dots, x_n , 有

$$F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F_{X_i}(x_i)$$

则称 X_1, X_2, \dots, X_n 相互独立. 否则称 X_1, X_2, \dots, X_n 不相互独立, 或相关.

(2) 设 n 维离散随机变量 (X_1, X_2, \dots, X_n) 的联合分布列为 $P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$, 且 $P(X_i = x_i)$ 为 X_i 的边际分布列, $i = 1, 2, \dots, n$.

八、边际分布与随机变量的独立性

如果对其任意 n 个取值 x_1, x_2, \dots, x_n , 有

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i)$$

则称 X_1, X_2, \dots, X_n 相互独立的. 否则称 X_1, X_2, \dots, X_n 不相互独立, 或相关.

(3) 设 n 维连续随机变量 (X_1, X_2, \dots, X_n) 的联合密度函数为 $f(x_1, x_2, \dots, x_n)$, 且 $f_{X_i}(x_i)$ 为 X_i 的边际密度函数. 如果对任意 n 个实数 x_1, x_2, \dots, x_n , 有

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i)$$

则称 X_1, X_2, \dots, X_n 相互独立的. 否则称 X_1, X_2, \dots, X_n 不相互独立, 或相关.

八、边际分布与随机变量的独立性

❖ **例20** 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \frac{k}{2} x e^{-x(1+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

- (1) 求系数 k ;
- (2) 求 (X, Y) 关于 X 和关于 Y 的边缘密度;
- (3) 问 X 和 Y 是否相互独立.

❖ **解:** (1) 由联合概率密度的性质有

$$\begin{aligned} 1 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = \frac{k}{2} \int_0^{+\infty} \int_0^{+\infty} x e^{-x(1+y)} dx dy \\ &= \frac{k}{2} \int_0^{+\infty} x e^{-x} \left[\int_0^{+\infty} e^{-xy} dy \right] dx = \frac{k}{2}, \end{aligned}$$

得常数 $k = 2$.

八、边际分布与随机变量的独立性

$$(2) \text{由(1)得 } f(x, y) = \begin{cases} xe^{-x(1+y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases}$$

$$\begin{aligned} \text{则 } f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^{+\infty} xe^{-x(1+y)} dy \\ &= xe^{-x} \int_0^{+\infty} e^{-xy} dy = xe^{-x} \left[-\frac{1}{x} e^{-xy} \right]_{y=0}^{y=+\infty} = e^{-x}. \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f(x, y) dx = \int_0^{+\infty} xe^{-x(1+y)} dx \\ &= \frac{-1}{1+y} \int_0^{+\infty} x de^{-x(1+y)} \\ &= \frac{-1}{1+y} \left[xe^{-x(1+y)} \right]_{x=0}^{x=+\infty} - \int_0^{+\infty} e^{-x(1+y)} dx \\ &= \frac{-1}{1+y} \left[0 - \left(\frac{-1}{1+y} \right) e^{-x(1+y)} \right]_{x=0}^{x=+\infty} = \frac{1}{(1+y)^2}. \end{aligned}$$

八、边际分布与随机变量的独立性

$$\text{故 } f_X(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & x \leq 0 \end{cases}.$$

$$f_Y(y) = \begin{cases} \frac{1}{(1+y)^2}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

(3) $f(x, y) \neq f_X(x)f_Y(y)$, 所以 X 与 Y 不独立.

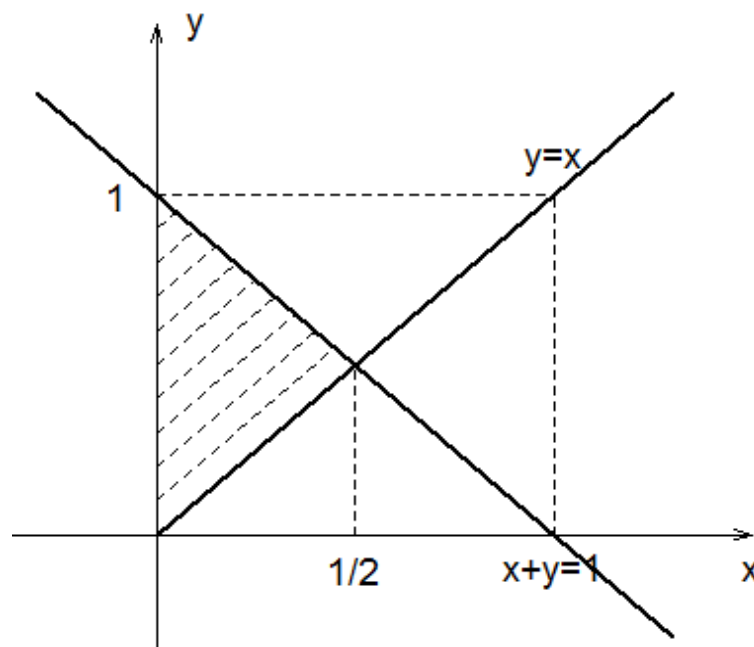
八、边际分布与随机变量的独立性

❖ 例21 设 (X, Y) 的联合概率密度为

$$f(x, y) = \begin{cases} 6x, & 0 \leq x \leq y, \\ 0, & \text{其他.} \end{cases}$$

则 $P\{X + Y \leq 1\} = \underline{\hspace{2cm}}$.

❖ **解:** $P\{X + Y \leq 1\} = \int_0^{\frac{1}{2}} \int_x^{1-x} 6x \, dy \, dx = \frac{1}{4}$, 故应填 $\frac{1}{4}$.



八、边际分布与随机变量的独立性

❖ **例22** 设 X_1, X_2, \dots, X_n 是独立同分布的正值随机变量, 证明:

$$E\left(\frac{X_1 + \dots + X_k}{X_1 + \dots + X_n}\right) = \frac{k}{n}.$$

❖ **证:** 记 $Y_j = X_j / \sum_{i=1}^n X_i$, 则诸 Y_j 同分布, 且由 $|Y_j| \leq 1$, 知 $E(Y_j)$ 存在且相等, $j = 1, 2, \dots, n$.

又因为 $1 = E\left(\sum_{j=1}^n Y_j\right) = nE(Y_1)$,

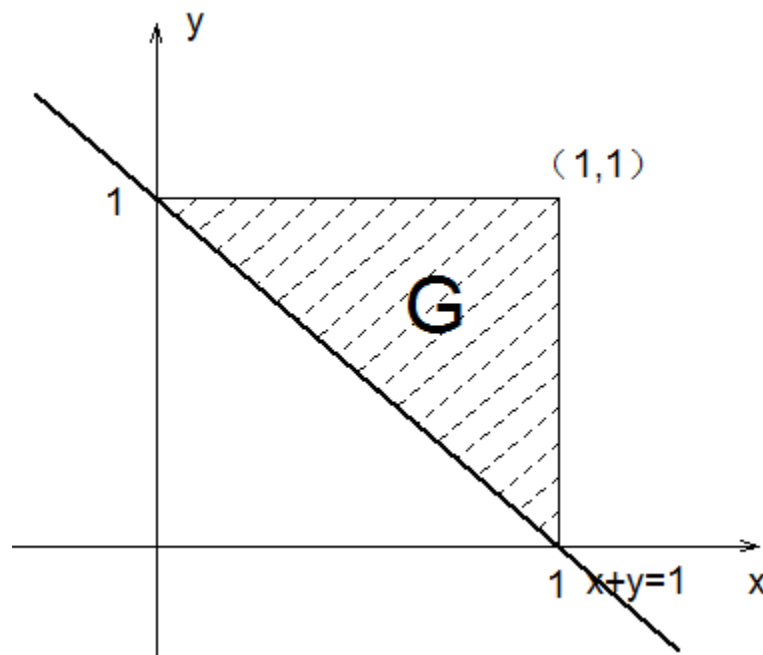
所以有 $E(Y_j) = \frac{1}{n}, j = 1, 2, \dots, n$. 由此得

$$E\left(\frac{X_1 + \dots + X_k}{X_1 + \dots + X_n}\right) = E(Y_1 + \dots + Y_k) = kE(Y_1) = \frac{k}{n}.$$

八、边际分布与随机变量的独立性

❖ **例23** (2001数4) 设随机变量 X 和 Y 的联合分布在以点 $(0,1)$, $(1,0)$, $(1,1)$ 为顶点的三角形区域上服从均匀分布, 试求 $U = X + Y$ 的方差.

❖ **解:** 令三角形区域 $G = \{(x, y) | 0 \leq x \leq 1, 1 - x \leq y \leq 1\}$, 显然 G 的面积为 $S_G = \frac{1}{2}$.





八、边际分布与随机变量的独立性

由均匀分布知 (X, Y) 的联合密度为

$$f(x, y) = \begin{cases} 2, & (x, y) \in G \\ 0, & \text{其他} \end{cases}$$

$$\begin{aligned} E(X + Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) f(x, y) dx dy \\ &= \iint_G 2(x + y) dx dy = 2 \int_0^1 dx \int_{1-x}^1 (x + y) dy = \frac{4}{3}. \end{aligned}$$

$$\begin{aligned} E(X + Y)^2 &= \iint_G 2(x + y)^2 dx dy \\ &= 2 \int_0^1 dx \int_{1-x}^1 (x + y)^2 dy = \frac{11}{6}. \end{aligned}$$

因此

$$Var(U) = E(X + Y)^2 - E^2(X + Y) = \frac{11}{6} - \left(\frac{4}{3}\right)^2 = \frac{1}{18}.$$

九、多维随机变量函数的分布

1. 最大值、最小值分布（极值分布）

设 (X_1, X_2, \dots, X_n) 是相互独立、同分布的 n 维连续随机变量，其共同的密度函数和分布函数分别为 $f(x)$ 和 $F(x)$ ，记

$$Y = \min\{X_1, X_2, \dots, X_n\}, Z = \max\{X_1, X_2, \dots, X_n\}.$$

则

$$F_Y(y) = 1 - [1 - F(y)]^n;$$

$$f_Y(y) = n[1 - F(y)]^{n-1}f(y);$$

$$F_Z(z) = [F(z)]^n;$$

$$f_Z(z) = n[F(z)]^{n-1}f(z).$$

九、多维随机变量函数的分布*

2. 变量变换法 (略) 若变换 $\begin{cases} u = g_1(x, y) \\ v = g_2(x, y) \end{cases}$, 存在唯一的反函数

$\begin{cases} x = x(u, v), \\ y = y(u, v), \end{cases}$ 且变换的雅可比行列式

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \left(\frac{\partial(u, v)}{\partial(x, y)} \right)^{-1} = \left(\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \right)^{-1} \neq 0.$$

则二维连续随机变量 (X, Y) 的函数 $\begin{cases} U = g_1(X, Y) \\ V = g_2(X, Y) \end{cases}$ 的联合密度函数
为 $f_{U,V}(u, v) = f_{X,Y}(x(u, v), y(u, v))|J|$.



九、多维随机变量函数的分布*

3.和的分布 卷积公式用于求随机变量 $Z = X + Y$ 的分布.

(1)离散场合的卷积公式 $Z = X + Y$ 的分布列为

$$P(Z = z_k) = \sum_{i=1}^{+\infty} P(X = x_i, Y = z_k - x_i)$$

$$\text{或} = \sum_{j=1}^{+\infty} P(X = z_k - y_j, Y = y_j).$$

当 X 与 Y 独立时,

$$P(Z = z_k) = \sum_{i=1}^{+\infty} P(X = x_i)P(Y = z_k - x_i)$$

$$\text{或} = \sum_{j=1}^{+\infty} P(X = z_k - y_j)P(Y = y_j).$$

其中诸 x_i 为 X 的取值, 诸 y_j 为 Y 的取值.

九、多维随机变量函数的分布*

(2)连续场合的卷积公式 $Z = X + Y$ 的密度函数为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x, z-x) dx$$

$$\text{或} = \int_{-\infty}^{+\infty} f_{X,Y}(z-y, y) dy.$$

当 X 与 Y 独立时,

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx$$

$$\text{或} = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy.$$

九、多维随机变量函数的分布

4.分布的可加性 若同一类分布的独立随机变量和的分布仍属于此类分布, 则称此类分布具有可加性. 以下一些常用分布具有可加性:

- (1) 二项分布: 若 $X \sim b(n, p)$, $Y \sim b(m, p)$, 且 X 与 Y 独立, 则 $Z = X + Y \sim b(n + m, p)$. 注意这里两个二项分布中的参数 p 要相同.
- (2) 泊松分布: 若 $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$, 且 X 与 Y 独立, 则 $Z = X + Y \sim P(\lambda_1 + \lambda_2)$.

九、多维随机变量函数的分布

(3)正态分布：若 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 且 X 与 Y 独立, 则 $Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

(4)伽玛分布：若 $X \sim Ga(\alpha_1, \lambda)$, $Y \sim Ga(\alpha_2, \lambda)$, 且 X 与 Y 独立, 则 $Z = X + Y \sim Ga(\alpha_1 + \alpha_2, \lambda)$.注：这里两个伽玛分布中的尺度参数 λ 要相同.

(5) χ^2 分布：若 $X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 且 X 与 Y 独立, 则 $Z = X + Y \sim \chi^2(n_1 + n_2)$.

注：可加性有助于计算更为复杂的抽样分布，详见后续章节。

九、多维随机变量函数的分布

5. 一些结论

- (1) 设 X_1, X_2, \dots, X_n 独立同分布, 都服从二点分布 $b(1, p)$, 则 $X_1 + X_2 + \dots + X_n$ 服从二项分布 $b(n, p)$.
- (2) 设 X_1, X_2, \dots, X_n 独立同分布, 都服从几何分布 $Ge(p)$, 则 $X_1 + X_2 + \dots + X_n$ 服从负二项分布 $Nb(n, p)$.
- (3) 设 X_1, X_2, \dots, X_n 独立同分布, 都服从指数分布 $Exp(\lambda)$, 则 $X_1 + X_2 + \dots + X_n$ 服从伽玛分布 $Ga(n, \lambda)$.

九、多维随机变量函数的分布

❖ **例24** 某种商品一周的需求量是一个随机变量，其密度函数为

$$f_1(t) = \begin{cases} te^{-t}, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

设每周的需求量是相互独立的，试求

(1) 两周需求量的密度函数 $f_2(x)$;

(2) 三周需求量的密度函数 $f_3(x)$.

九、多维随机变量函数的分布

❖ **解：** 记 X_i 为第 i 周的需求量， $i = 1, 2, 3$. 根据题意知 X_1, X_2, X_3 相互独立，且密度函数都为 $f_1(t)$. X_i 服从伽玛分布 $Ga(2, 1)$ ，所以由伽玛分布的可加性知

(1) $X_1 + X_2 \sim Ga(4, 1)$ ，其密度函数为

$$f_2(x) = \frac{x^3}{6} e^{-x}, \quad x > 0.$$

(2) $X_1 + X_2 + X_3 \sim Ga(6, 1)$ ，其密度函数为

$$f_3(x) = \frac{1}{\Gamma(6)} x^5 e^{-x} = \frac{1}{120} x^5 e^{-x}, \quad x > 0.$$

十、分布的其他数字特征

1. 二元随机变量函数的期望

设 (X, Y) 是二维随机变量, $z = g(x, y)$ 为实变量 x, y 的二元函数

(1)若 (X, Y) 是离散型随机变量, 其分布律为:

$P\{X = x_i, Y = y_i\} = p_{ij}, i, j = 1, 2, \dots$, 且

$\sum_i \sum_j g(x_i, y_i) p_{ij}$ 绝对收敛, 则

$$E(Z) = E[g(X, Y)] = \sum_i \sum_j g(x_i, y_i) p_{ij}.$$

(2)若 (X, Y) 是连续型随机变量, 其密度函数为

$f(x, y)$, 且 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$ 绝对收敛, 则

$$E(Z) = E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy.$$

十、分布的其他数字特征

2. k 阶矩

(1) 称 $\mu_k = E(X^k)$ 为 X 的 k 阶**原点矩**. 一阶原点矩就是数学期望;

(2) 称 $v_k = E(X - E(X))^k$ 为 X 的 k 阶**中心矩**. 二阶中心矩就是方差;

(3) 前 k 阶中心矩可用原点矩表示, 如

$$v_1 = 0;$$

$$v_2 = \mu_2 - \mu_1^2;$$

$$v_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3;$$

$$v_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4.$$

3. 变异系数 称比值 $C_v(X) = \frac{\sqrt{\text{Var}(X)}}{E(X)}$ 为 X 的变异系数. 变异系数是一个无量纲的量.

十、分布的其他数字特征

4.协方差 若 $E[(X - E(X))(Y - E(Y))]$ 存在, 则称

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))]$$

为 X 与 Y 的协方差.

- 当 $Cov(X, Y) > 0$ 时, 称 X 与 Y 正相关, 即 X 与 Y 同时增加或同时减少.
- 当 $Cov(X, Y) < 0$ 时, 称 X 与 Y 负相关, 即 X 增加 Y 减少, 或 X 减少 Y 增加.
- 当 $Cov(X, Y) = 0$ 时, 称 X 与 Y 不相关.

协方差的性质:

(1) $Cov(X, Y) = Cov(Y, X);$

(2) $Cov(X, Y) = E(XY) - E(X)E(Y) ;$

十、分布的其他数字特征

(3)若 X 与 Y 相互独立, 则 $Cov(X, Y) = 0$, 反之不然;

(4) $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$;

(5)对任意的常数 a , 有 $Cov(X, a) = 0$;

(6)对任意的常数 a, b , 有

$$Cov(aX, bY) = abCov(X, Y) ;$$

(7)对任意二维随机变量 (X, Y) , 有

$$Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X, Y).$$

对任意 n 个随机变量 X_1, X_2, \dots, X_n , 有

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i) + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} Cov(X_i, X_j).$$

十、分布的其他数字特征

5.相关系数 设 (X, Y) 是一个二维随机变量, 且
 $Var(X) > 0, Var(Y) > 0$.则称

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

为 X 与 Y 的 (线性) 相关系数.

十、分布的其他数字特征

相关系数的性质:

(1) $-1 \leq \text{Corr}(X, Y) \leq 1$.

(2) $\text{Corr}(X, Y)$ 与 $\text{Cov}(X, Y)$ 同号, 或同为零.

(3) $\text{Corr}(X, Y) = \text{Cov}(X^*, Y^*)$, 其中 X^*, Y^* 分别为 X, Y 的标准化随机变量.

(4) $\text{Corr}(X, Y) = \pm 1$ 的充要条件是 X 与 Y 间几乎处处有线性关系, 即存在 $a(a \neq 0)$ 与 b , 使得

$P(Y = aX + b) = 1$. 其中当 $\text{Corr}(X, Y) = 1$ 时, 有 $a > 0$; 当 $\text{Corr}(X, Y) = -1$ 时, 有 $a < 0$.

(5) 在二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 场合, 不相关与独立是等价的.

十、分布的其他数字特征

6. 随机向量的数学期望与协方差阵 记 n 维随机向量为 $\mathbf{X} = (X_1, X_2, \dots, X_n)'$, 以下假设所涉及的数学期望、方差、协方差均存在.

(1) 随机向量 \mathbf{X} 的数学期望向量为

$$E(\mathbf{X}) = (E(X_1), E(X_2), \dots, E(X_n))'.$$

(2) 随机向量 \mathbf{X} 的协方差阵为

$$\begin{aligned} & E \left[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))' \right] \\ &= \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \cdots & \text{Var}(X_n) \end{bmatrix} \end{aligned}$$

也记以上的协方差阵为 $\text{Cov}(\mathbf{X})$, 或记成 Σ .

十、分布的其他数字特征

(3) 随机向量 \mathbf{X} 的协方差阵 $Cov(\mathbf{X}) = (Cov(X_i, X_j))_{n \times n}$ 是一个对称的非负定矩阵.

(4) n 元正态分布 设 n 维随机向量 $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ 的协方差阵为 $\Sigma = Cov(\mathbf{X})$, 数学期望向量为 $\mathbf{a} = (a_1, a_2, \dots, a_n)'$. 又记 $\mathbf{x} = (x_1, x_2, \dots, x_n)'$, 则由密度函数

$$\begin{aligned} p(x_1, x_2, \dots, x_n) &= p(\mathbf{x}) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mathbf{a})' \Sigma^{-1}(\mathbf{x} - \mathbf{a})\right\} \end{aligned}$$

定义的分布称为 n 元正态分布, 记为 $\mathbf{X} \sim N(\mathbf{a}, \Sigma)$.

十、分布的其他数字特征

❖ **例25-1** (2012数1) 将长度为 $1m$ 的木棒随机地截成两段, 则两段长度的相关系数为_____.

❖ **解:** 设两段长度分别为 X, Y , 则由已知条件有 $X + Y = 1$, 则 $Y = -X + 1$, 所以 X, Y 的相关系数为 -1 .

❖ **例25-2** 设二维随机变量 (X, Y) 服从 $N(\mu, \mu; \sigma^2, \sigma^2; 0)$, 则 $E(XY^2) =$ _____.

❖ **解:** 由题可得, $Cov(X, Y) = 0$. 又, 在二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 场合, 不相关与独立是等价的. 因此 X 与 Y 独立. 故,

$$E(XY^2) = E(X)E(Y^2) = \mu(\mu^2 + \sigma^2)$$

十、分布的其他数字特征

❖ **例26** 设二维随机变量 (X, Y) 服从单位圆内的均匀分布, 其联合密度函数为

$$f(x, y) = \begin{cases} \frac{1}{\pi}, & x^2 + y^2 < 1, \\ 0, & x^2 + y^2 \geq 1. \end{cases}$$

试证: X 与 Y 不独立且 X 与 Y 不相关.

十、分布的其他数字特征

❖ **证：** 先求边际密度函数 $p_X(x)$ 和 $p_Y(y)$.

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{2}{\pi} \sqrt{1-x^2}, -1 < x < 1,$$

$$f_Y(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx = \frac{2}{\pi} \sqrt{1-y^2}, -1 < y < 1.$$

所以由 $f(x, y) \neq f_X(x)f_Y(y)$, 知 X 与 Y 不独立.

又因为 $f_X(x)$ 和 $f_Y(y)$ 在对称区间上是偶函数, 故 $E(X) = E(Y) = 0$, 从而

$$\text{Cov}(X, Y) = E(XY) = \frac{1}{\pi} \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy dy dx = 0,$$

所以 X 与 Y 不相关.

注：不相关并不意味着独立。

十、分布的其他数字特征

❖ **例27** (2012数1, 3) 设二维离散型随机变量 (X, Y) 的概率分布为

X/Y	0	1	2
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{3}$	0
2	$\frac{1}{12}$	0	$\frac{1}{12}$

(1) 求 $P\{X = 2Y\}$; (2) 求 $Cov(X - Y, Y)$.

❖ **解:** (1) $P\{X = 2Y\} = P\{X = 0, Y = 0\} +$

$$P\{X = 2, Y = 1\} = \frac{1}{4} + 0 = \frac{1}{4}.$$

(2) 易得 X, Y 的边缘分布律为



十、分布的其他数字特征

X	0	1	2
P	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

Y	0	1	2
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

又 XY 的所有可能取值为0,1,4, 且分布律为

XY	0	1	4
P	$\frac{7}{12}$	$\frac{1}{3}$	$\frac{1}{12}$

于是 $E(X) = \frac{2}{3}$, $E(Y) = 1$, $E(Y^2) = \frac{5}{3}$, $E(XY) = \frac{2}{3}$, 故

$$\begin{aligned} Cov(X - Y, Y) &= Cov(X, Y) - Cov(Y, Y) \\ &= E(XY) - E(X)E(Y) - [E(Y^2) - E^2(Y)] \\ &= \frac{2}{3} - \frac{2}{3} \times 1 - \left[\frac{5}{3} - 1^2 \right] = -\frac{2}{3}. \end{aligned}$$

十一、条件期望*

1. 条件数学期望 条件分布的数学期望（若存在）称为条件期望，其定义如下：

$$E(X|Y = y) = \begin{cases} \sum_i x_i p(X = x_i | Y = y), & (X, Y) \text{ 为二维离散随机变量,} \\ \int_{-\infty}^{+\infty} x f(x|y) dx, & (X, Y) \text{ 为二维连续随机变量,} \end{cases}$$
$$E(Y|X = x) = \begin{cases} \sum_j y_j p(Y = y_j | X = x), & (X, Y) \text{ 为二维离散随机变量,} \\ \int_{-\infty}^{+\infty} y f(y|x) dy, & (X, Y) \text{ 为二维连续随机变量,} \end{cases}$$

十一、条件期望*

- 条件期望具有数学期望的一切性质.
- 条件期望 $E(X|Y = y)$ 是 y 的函数, 记为 $g(y)$, 它是另一个随机变量 $g(Y) = E(X|Y)$ 的取值. $E(X|Y)$ 与 $E(X|Y = y)$ 虽都称为条件期望, 但含义不同.前者是特定的随机变量, 后者是其取值.

2.全期望公式 设 (X, Y) 是二维随机变量, 若 $E(X)$ 存在, 则 $E(X) = E[E(X|Y)]$.注意: 该公式中里面的期望是用条件分布 $f(x|y)$ 计算的, 外面的期望是用 y 的分布 $f(y)$ 计算的.

十一、条件期望*

3.二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的两个条件分布仍是正态分布, 即

$$X|Y = y \sim N\left(g_1(y), \sigma_1^2(1 - \rho^2)\right),$$

$$\text{其中 } g_1(y) = E(X|Y = y) = \mu_1 + \frac{\rho\sigma_1}{\sigma_2}(y - \mu_2).$$

$$Y|X = x \sim N\left(g_2(x), \sigma_2^2(1 - \rho^2)\right),$$

$$\text{其中 } g_2(x) = E(Y|X = x) = \mu_2 + \frac{\rho\sigma_2}{\sigma_1}(x - \mu_1).$$

可见, 在二维正态分布中, 一个变量的条件期望是另一个变量取值的线性函数, 常称为一元线性回归方程.

十一、条件期望*

❖ **例28** 设二维连续随机变量 (X, Y) 的联合密度函数为

$$f(x, y) = \begin{cases} x + y, & 0 < x, y < 1, \\ 0, & \text{其他.} \end{cases}$$

试求 $E(X|Y = 0.5)$.

❖ **解:** 先求条件密度函数 $f(x|y)$. 当 $0 < y < 1$ 时, $f_Y(y) = \int_0^1 (x + y) dx = 0.5 + y$. 所以

$$f(x|y = 0.5) = \begin{cases} x + 0.5, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$$

由此得

$$E(X|Y = 0.5) = \int_0^1 x \cdot (x + 0.5) dx = \frac{7}{12}.$$

十二、大数定律与中心极限定理

1. 随机变量序列的收敛性

(1) 依概率收敛 设 $\{X_n\}$ 为一随机变量序列, X 为一随机变量. 如果对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow +\infty} P\{|X_n - X| < \varepsilon\} = 1,$$

则称 $\{X_n\}$ 依概率收敛于 X , 记作 $X_n \xrightarrow{P} X$.

(2) 依概率收敛与服从大数定律间的关系 设 $\{X_n\}$ 为一随机变量序列, 记 $Y_n = \frac{1}{n} \sum_{i=1}^n X_i$, $E(Y_n) = \frac{1}{n} \sum_{i=1}^n E(X_i)$. 则 $\{X_n\}$ 服从大数定律等价于 $Y_n - E(Y_n) \xrightarrow{P} 0$.

(3) 依概率收敛的四则运算 如果 $X_n \xrightarrow{P} a$, $Y_n \xrightarrow{P} b$, 则有

十二、大数定律与中心极限定理

- i) $X_n \pm Y_n \xrightarrow{P} a \pm b;$
- ii) $X_n \times Y_n \xrightarrow{P} a \times b;$
- iii) $X_n/Y_n \xrightarrow{P} a/b (b \neq 0).$

(4) **按分布收敛、弱收敛** 设 $\{F_n(x)\}$ 是随机变量序列 $\{X_n\}$ 的分布函数列, $F(x)$ 为 X 的分布函数. 若对 $F(x)$ 的任一连续点 x , 都有 $\lim_{n \rightarrow \infty} F_n(x) = F(x)$, 则称 $\{F_n(x)\}$ 弱收敛于 $F(x)$, 记作 $F_n(x) \xrightarrow{W} F(x)$. 也称 $\{X_n\}$ 按分布收敛于 X , 记作 $X_n \xrightarrow{L} X$.

(5) 依概率收敛与按分布收敛间的关系

- i) $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{L} X.$
- ii) $X_n \xrightarrow{P} c \Leftrightarrow X_n \xrightarrow{L} c. (其中c为常数).$



十二、大数定律与中心极限定理

2. 大数定律

(1) **随机变量序列 $\{X_n\}$ 服从大数定律** 设 $\{X_n\}$ 为随机变量序列, 若对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i) \right| < \varepsilon \right\} = 1.$$

则称 $\{X_n\}$ 服从大数定律.

(2) **伯努利大数定律** 设 μ_n 为 n 重伯努利试验中事件 A 发生的次数, p 为每次试验中 A 出现的概率, 则对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1,$$

十二、大数定律与中心极限定理

这是第一个大数定律，它表明：事件发生的频率是依概率收敛与该事件的概率，这就是“频率稳定

于概率”的含义，也是“用频率去估计概率”的依据。

(3) **切比雪夫大数定律** 设 $\{X_n\}$ 为一列两两不相关的随机变量序列，若每个 X_i 的方差存在，且有共同的上界，即 $Var(X_i) \leq c, i = 1, 2, \dots$ ，则 $\{X_n\}$ 服从大数定律。

切比雪夫不等式 设 X 为随机变量，且有有限方差，则对任意正数 ε ，有

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

十二、大数定律与中心极限定理

(4) **马尔可夫大数定律** 对随机变量序列 $\{X_n\}$, 若有

$$\frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) \rightarrow 0 (n \rightarrow \infty)$$

则 $\{X_n\}$ 服从大数定律. 上式被称为**马尔可夫条件**.

(5) **辛钦大数定律** 设 $\{X_n\}$ 为一独立同分布的随机变量序列, 若 X_i 的数学期望存在, 则 $\{X_n\}$ 服从大数定律.

十二、大数定律与中心极限定理

3.中心极限定理

(1)中心极限定理 研究独立随机变量和的极限分布在什么条件下为正态分布的命题.

(2)林德伯格-莱维中心极限定理 $\{X_n\}$ 设是独立同分布的随机变量序列, 且 $E(X_i) = \mu, Var(X_i) = \sigma^2 > 0$. 记 $Y_n^* = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$.

则对任意实数 y , 有

$$\lim_{n \rightarrow \infty} P(Y_n^* \leq y) = \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt.$$

十二、大数定律与中心极限定理

(3)二项分布的正态近似

i)棣莫弗-拉普拉斯中心极限定理

设 n 重伯努利试验中, 事件 A 在每次试验中出现的概率为 $p(0 < p < 1)$, 记 μ_n 为 n 此试验中事件 A 出现的次数, 且记

$$Y_n^* = \frac{\mu_n - np}{\sqrt{np(1-p)}}.$$

则对任意实数 y , 有

$$\lim_{n \rightarrow +\infty} P(Y_n^* \leq y) = \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt.$$

十二、大数定律与中心极限定理*

ii) 近似中的修正 (略)

因为二项分布是离散分布，而正态分布是连续分布，所以用正态分布作为二项分布的近似计算中，作些修正可以提高精度。若 $k_1 < k_2$ 均为整数，一般先作如下修正后再用正态近似。

$$\begin{aligned} P(k_1 \leq \mu_n \leq k_2) &= P(k_1 - 0.5 < \mu_n < k_2 + 0.5) \\ &= \Phi\left(\frac{k_2 + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - 0.5 - np}{\sqrt{np(1-p)}}\right). \end{aligned}$$



十二、大数定律与中心极限定理

- ❖ **例29** 某保险公司多年的统计资料表明，在索赔户中被盗索赔户占20%，以 X 表示在随意抽查的100个索赔户中因被盗向保险公司索赔的户数。
- (1) 写出 X 分布列；
 - (2) 求被盗索赔户不少于14户且不多于30户的概率的近似值. (*P243 ex1*)

十二、大数定律与中心极限定理

❖ **解：** (1) X 服从 $n = 100$, $p = 0.2$ 的二项分布 $b(100, 0.2)$, 即

$$P(X = k) = \binom{n}{k} 0.2^k 0.8^{100-k}, k = 0, 1, 2, \dots, n.$$

(2) 利用棣莫弗-拉普拉斯中心极限定理, 有

$$P(14 \leq X \leq 30) = P(13.5 < X < 30.5)$$

$$\approx \Phi\left(\frac{30.5 - 100 \times 0.2}{\sqrt{100 \times 0.2 \times 0.8}}\right) - \Phi\left(\frac{13.5 - 100 \times 0.2}{\sqrt{100 \times 0.2 \times 0.8}}\right)$$

$$= \Phi(2.625) - \Phi(-1.625) = \Phi(2.625) - 1 + \Phi(1.625)$$

$$= 0.99565 - 1 + 0.948 = 0.9437.$$

这表明：被盗索赔户在14与30户之间的概率近似为0.9437.

十二、大数定律与中心极限定理

- ❖ **例30** 银行为支付某日即将到期的债券须准备一笔现金，已知这批债券共发放了500张，每张须付本息1000元，设持券人（1人1券）到期到银行领取本息的概率为0.4.问银行于该日应准备多少现金才能以99.9%的把握满足客户的兑换.
- ❖ **解：** 设 X 为该日到银行领取本息的总人数，则 $X \sim B(500, 0.4)$ ，所需支付现金为 $1000X$ ，为使银行能以99.9%的把握满足客户的兑换，设银行该日应准备现金 x 元，则 $P\{1000X \leq x\} \geq 0.999$.由棣莫弗—拉普拉斯中心极限定理知：



十二、大数定律与中心极限定理

$$\begin{aligned} P\{1000X \leq x\} &= P\left\{X \leq \frac{x}{1000}\right\} \\ &= P\left\{\frac{X - 500 \times 0.4}{\sqrt{500 \times 0.4 \times 0.6}} \leq \frac{\frac{x}{1000} - 500 \times 0.4}{\sqrt{500 \times 0.4 \times 0.6}}\right\} \\ &\approx \Phi\left(\frac{x - 20000}{2000\sqrt{30}}\right) \geq 0.999 = \Phi(3.1). \end{aligned}$$

即 $\frac{x - 20000}{2000\sqrt{30}} \geq 3.1$, 得 $x \geq 233958.798$.

因此银行于该日应准备234000元现金才能以99.9%的把握满足客户的兑换.

Thank You !

