概率论与数理统计总复习

安徽财经大学

统计与应用数学学院

目录

1. 利用全概率公式, 求复杂事件的概率 (参考例 1.4.4 和例 1.4.5); 利用贝叶斯公式, 确定引起事件 A 发生的诸多原因事件中, 哪一个出现的可能性最大. (参考例 1.4.6 和例 1.4.7)
2. 由随机变量的密度函数, 求密度函数中的未知参数, 随机变量落入某一区间的概率和随机变量的分布函数. 也要会由分布函数求密度函数. (参考例 2.1.6, 例 2.1.7, 例 2.1.8 和习题 2.1(8))
3. 会求二维随机变量的联合分布律 (函数) 和边际 (缘) 分布律 (边际密度函数), 会判断两个随机变量是否相互独立. (参考习题 3.1(9), 习题 3.2(3), 例 3.3.2)
4. 会求随机变量的数学期望和方差, 会求两个随机变量的协方差和相关系数. (参考复习题四解答题 (5))
5. 利用独立同分布中心极限定理或棣莫弗 · 拉普拉斯中心极限定理, 近似计算某一事件发生的概率. (参考例 5.4.1 和例 5.4.2)
6. 掌握 χ^2- 分布、$t-$ 分布和 $F-$ 分布的构造, 会验证统计量服从何种分布? (参考例 6.3.3 和例 6.3.4)
7. 利用矩估计法或极大似然估计法对总体分布中的未知参数进行估计. (参考习题 7.1 中的 (2) 和 (3))

1. 概率的性质, 条件概率, 概率的乘法公式, 事件的独立性
1. 伯努利试验中事件 A 出现 k 次的概率
2. 常见离散型随机变量及其分布列, 连续型随机变量及其密度函数和分布函数.
2. 正态随机变量的标准化
3. 随机变量函数的分布
4. 常见分布的期望和方差
4. 期望、方差、协方差和相关系数的性质
5. 切比雪夫不等式
6. 常用的抽样分布
7. 统计量的判断、统计量的无偏性和有效性

第 1 章 随机事件与概率

- 随机试验, 样本空间, 随机事件, 互不相容事件与对立事件.
- 概率的公理化定义, 概率的频率定义, 概率的统计定义, 概率的古典定义, 概率的几何定义.
- 概率的性质, 有限可加性, 单调性, 减法公式, 加法公式.
- 条件概率, 乘法公式, 全概率公式, 贝叶斯公式.
- 事件的独立性, n 重伯努利试验.

1. 随机试验, 样本空间, 随机事件

- **随机事件的关系**:
  包含: $A \subset B$; 相等: $A = B$; 互不相容: $A \cap B = \varnothing$.
- **随机事件的运算**:
  和: $A \cup B$; 交: $A \cap B$ (或者 AB); 差: $A - B$ (或者 $A\overline{B}$); 对立事件: $\overline{A}$
- **随机事件的运算律**: 与集合的运算律相仿
  交换律: $A \cup B = B \cup A, A \cap B = B \cap A$.
  结合律: $A \cup (B \cup C) = (A \cup B) \cup C, A \cap (B \cap C) = (A \cap B) \cap C$.
  分配律: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
  对偶律: $\overline{A \cup B} = \overline{A} \cap \overline{B}, \overline{A \cap B} = \overline{A} \cup \overline{B}$.

符号	概率论	集合论
Ω	样本空间，必然事件	全集
$\varnothing$	不可能事件	空集
ω	样本点	元素
$\{\omega\}$	基本事件	单点集
A	随机事件	子集
$\overline{A}$	事件 A 的对立事件	A 的补集
$A \subset B$	事件 A 是事件 B 的子事件	A 是 B 的子集
$A = B$	事件 A 与事件 B 相等	A 与 B 相等
$A \cup B$	事件 A 与 B 至少有一个发生	A 与 B 的并集
$A \cap B$	事件 A 与事件 B 同时发生	A 与 B 的交集
$A \cap B = \varnothing$	事件 A 与事件 B 互不相容	A 与 B 的交为空集
$A - B$	事件 A 发生且事件 B 不发生	A 与 B 的差

2. 概率的基本性质

- 概率应满足的三个公理:
  非负性公理: $P(A) \geqslant 0$;
  规范性公理: $P(\Omega) = 1$;
  可列可加性公理: 对于两两互不相容的可列无穷多个事件 $A_1, A_2, \cdots, A_n, \cdots$, 有

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

- 常用性质: $P(\varnothing) = 0$;

2. 概率的基本性质

- **有限可加性:** 若 $A_1, A_2, \cdots, A_n$ 为两两互不相容事件, 则有 $P(\bigcup\limits_{k=1}^{n} A_k) = \sum\limits_{k=1}^{n} P(A_k)$.
- **减法公式:** $P(B-A) = P(B) - P(AB)$, 特别地: $A \subset B$ 有 $P(B-A) = P(B) - P(A)$.
- **加法公式:** 对于任意两个事件 A, B 有 $P(A \cup B) = P(A) + P(B) - P(AB)$.
- 对立事件的概率关系: $P(\overline{A}) = 1 - P(A)$.

3. 概率的计算

- 频率; 古典概率 $P(A) = \dfrac{m}{n}$; 几何概率 $P(A) = \dfrac{S_A}{S_\Omega}$.

4. 条件概率及其三大公式

- **条件概率:** $P(A \mid B) = \dfrac{P(AB)}{P(B)} \quad (P(B) > 0).$
- **乘法公式:** $P(AB) = P(B)P(A \mid B),$
- **全概率公式:** $A_1, A_2, \cdots, A_n$ 是 Ω 的一个分割, 则有

$$P(B) = P(\bigcup_{i=1}^{n}(A_iB)) = \sum_{i=1}^{n} P(A_iB) = \sum_{i=1}^{n} P(A_i)P(B \mid A_i).$$

- **贝叶斯 (Bayes) 公式:**

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum\limits_{j=1}^{n} P(B \mid A_j)P(A_j)}, \quad i = 1, 2, \cdots, n.$$

5. 独立性

- 两个事件独立: $P(AB) = P(A)P(B)$; 多个事件的独立.
- 伯努利 (Bernoulli) 试验; n 重独立重复试验; n 重伯努利试验.
- n 重伯努利试验中, A 出现次的概率:

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

第 2 章 随机变量及其分布

- 随机变量的分布函数及其性质.
- 离散型型随机变量的分布列及其性质, 连续型随机变量的密度函数 $f(x)$ 及其性质.
- 常见的离散型随机变量及分布列, 常见的连续型随机变量的密度函数和分布函数.
- 用分布函数法或公式法求随机变量函数的分布.

随机变量及其分布

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

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

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

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

(4) 连续型随机变量常用密度函数 $f(x)$ 来描述, 具有如下性质:

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

(5) 常见的离散型随机变量及分布列:

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

(6) 常见的连续型分布及其密度函数.

- 均匀分布 $U(a,b)$:

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

- 指数分布 $Exp(\lambda)$:

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

- 正态分布 $N(\mu,\sigma^2)$:

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

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

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

- 分布函数法.
- 公式法.

定理

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

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

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

第 3 章 多维随机变量

- 联合分布函数的定义和性质，二维离散型随机变量的联合分布列，二维连续型随机变量的联合密度 $f(x, y)$ 及其性质，
- 对于二维随机变量，联合分布函数唯一确定两个边际分布函数，联合密度函数唯一确定两个边际密度函数，联合分布列唯一确定两个边际分布列，反之不成立. 当 X, Y 相互独立时，边际分布可以唯一确定联合分布.
- 两个随机变量相互独立的充要条件.
- 两个独立的随机变量和的分布，卷积公式.
- 有限个相互独立的正态随机变量的线性组合 (组合系数不全为零) 仍服从正态分布.

第 3 章 多维随机变量

(1) 常见的二维随机变量: 二维离散型随机变量和二维连续型随机变量.

(2) 二维随机变量的分布函数: $F(x, y) = P\{X \leqslant x, Y \leqslant y\}$, 联合分布函数的性质: 单调不减性、有界性、右连续性、非负性.

(3) 二维离散型随机变量的联合分布列描述, 有性质: 非负性: $p_{ij} \geqslant 0$; 规范性: $\sum\limits_{i,j} p_{ij} = 1$.

(4) 二维连续型随机变量的联合密度 $f(x, y)$ 描述, 有性质

- 非负性: $f(x, y) \geqslant 0, \ (-\infty < x, y < +\infty)$.
- 正则性: $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(u, v) \mathrm{d}u \mathrm{d}v = 1$.
- $P((X, Y) \in D) = \iint_D f(x, y) \mathrm{d}x \mathrm{d}y$.
- 若 $f(x, y)$ 在点 (x, y) 处连续, 则 $f(x, y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$.

(5) 边际分布函数、边际分布列和边际密度函数

$$F_X(x) = \lim_{y\to+\infty} F(x,y) = F(x,+\infty),$$

$$F_Y(y) = \lim_{x\to+\infty} F(x,y) = F(+\infty,y).$$

$$p_{i\cdot} = P(X = x_i) = \sum_j P(X = x_i, Y = y_j) = \sum_j p_{ij},\quad i = 1,2,\cdots,$$

$$p_{\cdot j} = P(Y = y_j) = \sum_i P(X = x_i, Y = y_j) = \sum_i p_{ij},\quad j = 1,2,\cdots,$$

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x} = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y.$$

$$f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x.$$

(6) 随机变量 X, Y 相互独立的三个充要条件:

- $F(x, y) = F_X(x)F_Y(y)$.
- 对于二维离散型变量: $P\{X = x_i, Y = y_j\} = P\{X = x_i\}P\{Y = y_j\}$.
- 对于二维连续型变量: $f(x, y) = f_X(x)f_Y(y)$.

(7) 两个随机变量函数的分布. 分布函数法, 或者用卷积公式求 $Z = X + Y$ 的分布.

$$\boxed{f_Z(z) = \int_{-\infty}^{+\infty} f(z-y, y)\mathrm{d}y,} \quad \boxed{f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x)\mathrm{d}x.}$$

- 卷积 (Convolution) 公式
当 X 和 Y 相互独立时, 设 (X, Y) 关于 X, Y 的边际概率密度分别为 $f_X(x)$, $f_Y(y)$, 则有

$$\boxed{f_Z(z) = f_X * f_Y = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)\mathrm{d}y = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x}$$

- 若 X, Y 相互独立且 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，则

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

- $M = \max(X, Y)$ 及 $N = \min(X, Y)$ 的分布

$$F_M(z) = F_X(z)F_Y(z).$$

$$F_N(z) = 1 - (1 - F_X(z))(1 - F_Y(z)).$$

第 4 章 随机变量的数字特征

- 随机变量的期望和方差的定义、性质和常用的计算公式.
- 常见分布的定义、期望和方差.
- 两个随机变量的协方差和相关系数的定义、性质、独立与不相关之间的关系.
- 了解随机变量的矩及协方差矩阵.

随机变量数字特征的定义

- 期望的定义

$$E(X)=\sum_{k=1}^{+\infty}x_kp_k;\quad E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x;$$

- 方差的定义

$$D(X)=\sum_{k=1}^{+\infty}[x_k-E(X)]^2p_k;\quad D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2f(x)\mathrm{d}x.$$

- 协方差的定义:

$$\operatorname{cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\};$$

- 相关系数的定义

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}.$$

期望和方差的性质

(1) $E(c) = c$, $D(c) = 0$, 其中 c 是常数;

(2) $E(cX) = cE(X)$, $D(cX) = c^2 D(X)$;

(3) $E(X + Y) = E(X) + E(Y)$, $E(X - Y) = E(X) - E(Y)$;

(4) X, Y 相互独立 $\Rightarrow$ X, Y 不相关 $\Leftrightarrow E(XY) = E(X)E(Y)$
$\Leftrightarrow D(X \pm Y) = D(X) + D(Y) \Leftrightarrow \rho = 0 \Leftrightarrow \mathrm{cov}(X, Y) = 0.$

(5) 数学期望表示的是随机变量取值的“加权平均”.

(6) 方差是描述随机变量的取值偏离其数学期望的程度.

协方差和相关系数的性质

(1) $\operatorname{cov}(X, X) = D(X)$;

(2) $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$;

(3) 若 X 与 Y 相互独立, 则 $\operatorname{cov}(X, Y) = 0$;

(4) 若 a, b 为常数, 则 $\operatorname{cov}(aX, bY) = ab \cdot \operatorname{cov}(X, Y)$;

(5) $\operatorname{cov}(X_1 + X_2, Y) = \operatorname{cov}(X_1, Y) + \operatorname{cov}(X_2, Y)$;

(6) 若 a 为常数, 则 $\operatorname{cov}(a, X) = 0$.

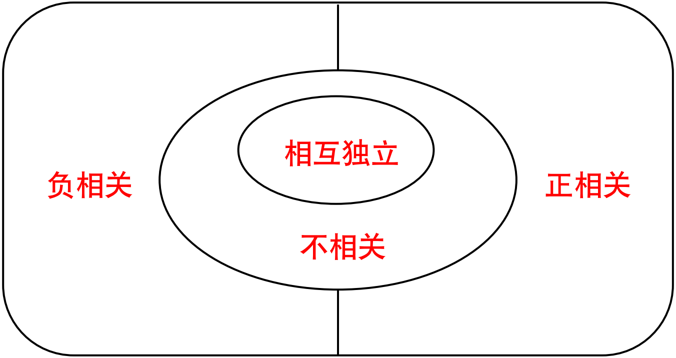
(7) $|\rho_{XY}| \leqslant 1$; $|\rho_{XY}| = 0$ 时称 X, Y 不相关;

$$|\rho_{XY}| = 1 \Leftrightarrow P\{Y = aX + b\} = 1, \ (a \neq 0).$$

(8) X, Y 的相关系数描述了 X, Y 之间线性关系的紧密程度.

(9) 若 (X, Y) 服从二维正态分布, 那么 X 和 Y 相互独立的充要条件是 $\rho = 0$, 即 X 与 Y 不相关.

相互独立与线性无关、线性相关之间的关系



常用的计算公式

- 方差计算的常用公式: $D(X)=E(X^2)-E^2(X)$.
- 方差与协方差的关系式: $D(X)=\mathrm{cov}(X,X)$

$$D(X\pm Y)=D(X)+D(Y)\pm 2\mathrm{cov}(X,Y).$$

- 协方差的计算公式: $\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)$;
- 柯西 – 许瓦兹不等式: $[E(XY)]^2\leqslant E(X^2)E(Y^2)$.

分布	分布律或密度函数	期望	方差
$b(1,p)$	$P(X=k)=p^k(1-p)^{1-k},\ k=0,1$	p	$p(1-p)$
$b(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{n-k},\ k=0,1,\cdots,n$	np	$np(1-p)$
$P(\lambda)$	$P(X=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\ k=0,1,2,\cdots,\ \lambda>0$	λ	λ
$Ge(p)$	$P(X=k)=p(1-p)^{k-1},\ k=1,2,\cdots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$U(a,b)$	$f(x)=\frac{1}{b-a},\ a<x<b;\ f(x)=0,$ 其他	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$Exp(\lambda)$	$f(x)=\lambda\mathrm{e}^{-\lambda x},\ x>0;\ f(x)=0,\ x\leqslant 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},\ -\infty<x<+\infty$	μ	σ^2

随机变量函数的数学期望

(1) 若 X 的分布列为 $P(X=x_k)=p_k$, $k=1,2,\cdots$, 则 $Y=g(X)$ 的期望

$$E(Y)=E[g(X)]=\sum_{k=1}^{+\infty} g(x_k)p_k.$$

(2) X 的概率密度为 $f(x)$, 则有 $Y=g(X)$ 的期望为

$$E(Y)=E[g(X)]=\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x.$$

(3) 设 (X, Y) 的联合分布列为 $P\{X=x_i, Y=y_j\}=p_{ij}$, $(i,j=1,2,\cdots)$ 时, 则有 $Z=g(X, Y)$ 的数学期望为

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

(4) 当 (X, Y) 的概率密度为 $f(x, y)$ 时, 则有 $Z=g(X, Y)$ 的数学期望为

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

第 5 章 大数定律与中心极限定理

- 会用契比雪夫不等式估计随机变量取值落在某个区间的概率.
- 了解依概率收敛和依分布收敛的定义.
- 理解契比雪夫大数定律、伯努利大数定律、辛钦大数定律的内涵,注意三大定律条件的不同.
- 掌握独立同分布的中心极限定理和棣莫弗 · 拉普拉斯中心极限定理,会用中心极限定理近似计算事件发生的概率.

定理 (契比雪夫不等式)

设随机变量 X 存在有限方差 $D(X)$, 则有对任意 $\varepsilon > 0$, 有

$$P\{|X - E(X)| \geqslant \varepsilon\} \leqslant \frac{D(X)}{\varepsilon^2}.$$

$$P\{|X - E(X)| < \varepsilon\} \geqslant 1 - \frac{D(X)}{\varepsilon^2}.$$

定理 (契比雪夫不等式的另一种形式)

设随机变量 X 的期望为 μ, 存在有限方差 σ^2, $k > 0$, 则

$$P\{|X - \mu| \geqslant k\sigma\} \leqslant \frac{1}{k^2}.$$

$$P\{|X - \mu| \geqslant 2\sigma\} \leqslant \frac{1}{4}, P\{|X - \mu| \geqslant 3\sigma\} \leqslant \frac{1}{9}, P\{|X - \mu| \geqslant 4\sigma\} \leqslant \frac{1}{16}.$$

定义 (依概率收敛)

设 $\{X_n\}$ 为一随机变量序列，X 为一随机变量，如果对于任意的 $\varepsilon > 0$ 有

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

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

定义 (依分布收敛)

设一个随机变量序列 $\{X_n\}$ 的分布函数列为 $\{F_n(x)\}$，$F(x)$ 为随机变量 X 的分布函数，若在 $F(x)$ 的任意连续点 x 都有

$$\lim_{n\to\infty} F_n(x) = F(x),$$

则称 $\{F_n(x)\}$ 弱收敛于 $F(x)$，记作 $F_n(x) \xrightarrow{W} F(x)$. 也称随机变量序列 $\{X_n\}$ 依分布收敛于 X，记作$X_n \xrightarrow{L} X$.

定理 (契比雪夫 (Chebyshev) 大数定律)

设 $X_1, X_2, \cdots$ 是两两不相关的随机变量序列, 各有数学期望 $E(X_1)$, $E(X_2), \cdots$ 及方差 $D(X_1), D(X_2), \cdots$, 并且对于所有 $i = 1, 2, \cdots$ 都有 $D(X_i) < C$, (C 是与 i 无关的常数), 则 $X_1, X_2, \cdots$ 服从大数定律, 简记为 $X_n \sim LLN$. 即对任给 $\varepsilon > 0$, 有

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

定理 (伯努利 (Bernoulli) 大数定律)

设 μ_A 是 n 次独立重复试验中事件 A 发生的次数. p 是事件 A 在每次试验中发生的概率, 则对于任意正数 $\varepsilon > 0$, 有

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_A}{n} - p\right| < \varepsilon\right\} = 1.$$

定理 (马尔可夫大数定律)

对随机变量序列 $X_1, X_2, \cdots$, 若马尔可夫条件 $\frac{1}{n^2}D(\sum_{i=1}^{n} X_i) \to 0$成立, 则序列 $X_1, X_2, \cdots$ 服从大数定律, 即对任意的 $\varepsilon > 0$, 有

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

定理 (辛钦 (Khinchin) 大数定律)

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立, 服从同一分布, 且具有数学期望 $E(X_k) = \mu$ $(k = 1, 2, \cdots)$, 则序列 $X_n \sim LLN$. 即对任给 $\varepsilon > 0$, 有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1.$$

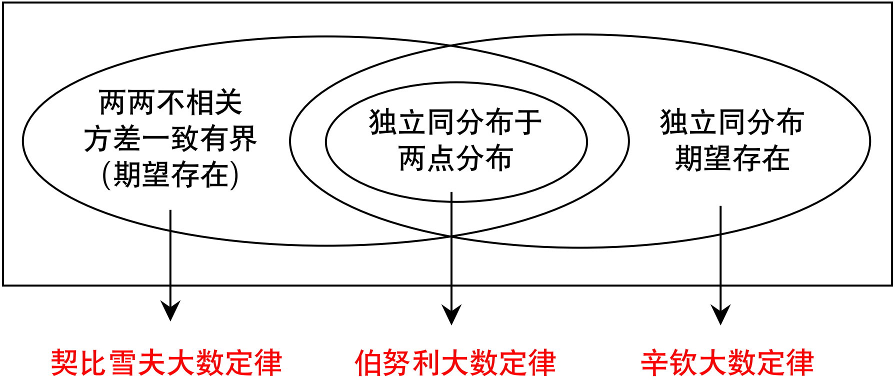


图: 三个大数定律的条件关系

定理 (独立同分布的中心极限定理)

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，服从同一分布，且具有数学期望 $E(X_k) = \mu$ 和方差 $D(X_k) = \sigma^2 > 0, (k = 1, 2, \cdots)$. 记

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}},$$

则对于任意实数 y 满足

$$\lim_{n\to\infty} P\{Y_n \leqslant y\} = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t = \Phi(y),$$

即: $\{X_n\}$ 服从中心极限定理，简记为 $X_n \sim CLT$.

当 n 充分大时，近似地有 $Y_n \sim N(0, 1)$，或者，当 n 充分大时，近似地有

$$\sum_{k=1}^{n} X_k \sim N(n\mu, n\sigma^2), \text{ 或 } \frac{1}{n}\sum_{k=1}^{n} X_k \sim N(\mu, \frac{\sigma^2}{n}).$$

定理 (棣莫佛 – 拉普拉斯 (De Moivre-Laplace) 中心极限定理)

设随机变量 X 服从参数为 $n, p\ (0 < p < 1)$ 的二项分布，则对于任意的 x，有

$$\lim_{n\to\infty} P\left\{\frac{X-np}{\sqrt{np(1-p)}} \leqslant x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\mathrm{d}t = \Phi(x).$$

第 6 章　数理统计基础

- 数理统计中的一些基本概念，如总体、样本、样本容量、简单随机抽样、经验分布函数、统计量等概念.
- 两个重要统计量：样本均值 $\bar{\xi}=\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$ 与样本方差
$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2.$$
- 数理统计中常用三大抽样分布：χ^2- 分布、$t-$ 分布、$F-$ 分布的定义、构造、密度函数的图像、性质、查分布表确定上 α 分位点.
- 抽样分布定理，正态总体统计量样本均值与样本方差的分布.

定理 (6.3.1)

若 $\xi_1,\xi_2\cdots,\xi_n$ 相互独立, 且都服从标准正态分布 $N(0,1)$, 则

$$\chi^2=\sum_{i=1}^{n}\xi_i^2\sim\chi^2(n).$$

定理 (6.3.2)

若 $\xi\sim N(0,1),\eta\sim\chi^2(n)$, 且 ξ 与 η 相互独立, 则

$$t=\frac{\xi}{\sqrt{\eta/n}}\sim t(n).$$

定理 (6.3.3)

设 $\xi\sim\chi^2(n_1),\eta\sim\chi^2(n_2)$, 且 ξ 和 η 相互独立, 则

$$F=\frac{\xi/n_1}{\eta/n_2}\sim F(n_1,n_2).$$

分位数

定义 (分位数)

(1) χ^2- 分布的上 α 分位数 $\chi_\alpha^2(n)$:

$$P\left\{\xi > \chi_\alpha^2(n)\right\} = \alpha.$$

(2) $t-$ 分布的上 α 分位数 $t_\alpha(n)$:

$$P\left\{t > t_\alpha(n)\right\} = \alpha.$$

(3) $F-$ 分布的上 α 分位数 $F_\alpha(n_1, n_2)$:

$$P\left\{F > F_\alpha\left(n_1, n_2\right)\right\} = \alpha.$$

$$t_{1-\alpha}(n) = -t_\alpha(n), \quad F_{1-\alpha}\left(n_1, n_2\right) = \frac{1}{F_\alpha\left(n_1, n_2\right)}.$$

χ^2- 分布的性质

(1) 若 $\xi \sim \chi^2(n_1), \eta \sim \chi^2(n_2)$，且 ξ, η 相互独立，则

$$\xi + \eta \sim \chi^2(n_1 + n_2).$$

(2) 若 $\xi \sim \chi^2(n)$，则 $E(\xi) = n, D(\xi) = 2n$.

$t-$ 分布的性质

(1) $t-$ 分布的密度函数的图像关于纵轴对称且 $\lim_{x\to\infty} f(x) = 0$.

(2) 当 n 充分大时 $t-$ 分布近似于标准正态分布.

(3) 若 $\xi \sim t(n)$，则 $E(\xi) = 0$, $(n > 1)$; $D(\xi) = \dfrac{n}{n-2}(n > 2)$.

$F-$ 分布的性质

(1) 若 $\xi \sim t(n)$，则 $\xi^2 \sim F(1, n)$.

(2) 若 $F \sim F(n_1, n_2)$，则 $\dfrac{1}{F} \sim F(n_2, n_1)$.

定理 (Fisher 定理, 抽样分布定理)

设总体 $\xi \sim N\left(\mu, \sigma^2\right)$, $\xi_1, \xi_2, \cdots, \xi_n$ 是来自正态总体 ξ 的一个简单样本, $\overline{\xi}$, S^2 分别为该样本的样本均值与样本方差, 则有

(1) $\overline{\xi} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$; (2) $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$; (3) $\overline{\xi}$ 与 S^2 相互独立.

表: 单个正态总体 $\xi \sim N(\mu, \sigma^2)$

样本函数	分布
$U = \dfrac{\overline{\xi} - \mu}{\sigma/\sqrt{n}}$	$N(0,1)$
$t = \dfrac{\overline{\xi} - \mu}{S/\sqrt{n}}$	$t(n-1)$
$\chi^2 = \dfrac{\sum\limits_{i=1}^{n}(\xi_i - \mu)^2}{\sigma^2}$	$\chi^2(n)$
$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} = \dfrac{\sum\limits_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2}{\sigma^2}$	$\chi^2(n-1)$

表: 双正态总体, $\xi \sim N\left(\mu_1, \sigma_1^2\right)$, $\eta \sim N\left(\mu_2, \sigma_2^2\right)$, ξ, η 独立

样本函数	分布
$U=\dfrac{\bar{\xi}-\bar{\eta}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$N(0,1)$
$T=\dfrac{\bar{\xi}-\bar{\eta}-(\mu_1-\mu_2)}{\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ (当 $\sigma_1^2=\sigma_2^2$ 时)	$t(n_1+n_2-2)$
$F=\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$	$F(n_1-1, n_2-1)$

第 7 章 参数估计

- 会用矩估计法和极大似然估计法对未知参数进行估计.
- 会从估计量的无偏性、和有效性一致性三个方面来讨论衡量估计量的优良性.

- 矩估计的基本思想是用样本的经验分布和样本矩作为总体相应矩的估计, 以样本矩的函数作为总体相应矩的同一函数的估计.

设总体 $X \sim F(x,\theta_1,\cdots,\theta_k)$, 其中 $(\theta_1,\cdots,\theta_k)\,,k \geqslant 1$ 为 k 个未知参数, 假定总体 X 的 k 阶矩存在, 则其 $r(r \leqslant k)$ 阶矩也存在. 即

$$m_r(\theta_1,\cdots,\theta_k) = EX^r = \int_{-\infty}^{+\infty} x^r \mathrm{d}F(x,\theta_1,\cdots,\theta_k)\,, \quad r = 1,2,\cdots,k$$

记 r 阶样本矩为 $A_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r, \quad r = 1,2,\cdots,k$. 得到 k 个方程

$$m_r(\theta_1,\cdots,\theta_k) = A_r, r = 1,2,\cdots,k.$$

求解方程组, 得到 $\theta_1,\cdots,\theta_k$ 的一组解 $\hat{\theta}_j = \hat{\theta}_j(X_1,X_2,\cdots,X_n)$, $j = 1,2,\cdots,k$. 称 $\hat{\theta}_j(X_1,X_2,\cdots,X_n)$ 为参数 θ_j 的矩估计量, $\hat{\theta}_j(x_1,x_2,\cdots,x_n)$ 称为参数 θ_j 的矩估计值, $(j = 1,2,\cdots,k)$.

以下分别从离散型和连续型总体两种情形来介绍极大似然估计法.

(1) 设总体 X 为离散型随机变量, 其分布列为 $P(X=x)=p(x,\theta)$, 其中 $\theta\in\Theta$ 为待估的未知参数. $(x_1,x_2,\cdots,x_n)$ 为样本 $(X_1,X_2,\cdots,X_n)$ 的一组观测值, 极大似然原理就是选取 θ 的估计值 $\hat{\theta}=\hat{\theta}\left(x_1,x_2,\cdots,x_n\right)$ 使得似然函数 $L(\theta)=\prod_{i=1}^{n} p\left(x_i,\theta\right)$ 达到最大值.

$$\prod_{i=1}^{n} p(x_i,\hat{\theta})=\max_{\theta\in\Theta} L(\theta)=\max_{\theta\in\Theta}\prod_{i=1}^{n} p\left(x_i,\theta\right).$$

(2) 设总体 X 为连续型随机变量, 其概率密度为 $f(x,\theta),\theta\in\Theta$ 为待估的未知参数. 则样本 $(X_1,X_2,\cdots,X_n)$ 的联合密度为 $\prod_{i=1}^{n} f(x_i,\theta)$, $(x_1,x_2,\cdots,x_n)$ 为一组样本观测值, 极大似然原理就是选取 θ 的估计值 $\hat{\theta}=\hat{\theta}\left(x_1,x_2,\cdots,x_n\right)$ 使得似然函数 $L(\theta)=\prod_{i=1}^{n} f(x_i,\theta)$ 达到最大值.

$$\prod_{i=1}^{n} f(x_i,\hat{\theta})=\max_{\theta\in\Theta} L(\theta)=\max_{\theta\in\Theta}\prod_{i=1}^{n} f\left(x_i,\theta\right).$$

由此可见，不管总体是离散型还是连续型，我们只需知道其分布列或概率密度函数，就可以得到一个似然函数 $L(\theta)$，其中 θ 为任意 k 维的参数向量 $\theta=(\theta_1\cdots,\theta_k)$；然后根据极大似然原理来寻找 θ 的估计值 $\hat{\theta}=\hat{\theta}\left(x_1,x_2,\cdots,x_n\right)$ 使得

$$L\left(\hat{\theta};x_1,\cdots,x_n\right)=\sup_{\theta\in\Theta}L\left(\theta;x_1,\cdots,x_n\right). \tag{7.1.4}$$

如果存在 $\hat{\theta}\left(X_1,X_2,\cdots,X_n\right)$ 使得 (7.1.4) 式成立，则称之为 θ 的极大似然估计量，记为 $\hat{\theta}_{\mathrm{MLE}}$，相应的估计值 $\hat{\theta}\left(x_1,x_2,\cdots,x_n\right)$ 为 θ 的极大似然估计值. 上述这种求参数 θ 的估计量的方法称为极大似然估计法.

由于 $\ln L(\theta)$ 是 $L(\theta)$ 的单增函数, 故 $L(\theta)$ 与 $\ln L(\theta)$ 在同一点取得极大值, 于是 (7. 1.4) 式可转化为

$$\ln L\left(\hat{\theta}; x_1, \cdots, x_n\right) = \sup_{\theta\in\Theta} \ln L\left(\theta; x_1, \cdots, x_n\right). \tag{7.1.5}$$

为求 (7.1.5) 式的解, 根据多元微分学知识, 若 Θ 为开集, 似然函数 $L(\theta)$ 关于 θ 的偏导数存在, 则 $\theta = (\theta_1, \cdots, \theta_k)$ 的极大似然估计 $\hat{\theta} = \left(\hat{\theta}_1, \cdots, \hat{\theta}_k\right)$ 一定为方程组

$$\frac{\partial}{\partial\theta_i} L\left(\theta_1, \cdots, \theta_k\right) = 0, i = 1, 2, \cdots, k \tag{7.1.6}$$

或

$$\frac{\partial}{\partial\theta_i} \ln L\left(\theta_1, \cdots, \theta_k\right) = 0, i = 1, 2, \cdots, k \tag{7.1.7}$$

的解. (7.1.6) 式或 (7.1.7) 式称为似然方程.

1. 无偏性

定义 (7.1.1)

设 $\hat{\theta}=\hat{\theta}\left(X_1,\cdots,X_n\right)$ 为未知参数 $\theta\in\Theta$ 的估计量. 如果对一切 $\theta\in\Theta$, 都有

$$E\hat{\theta}=\theta.$$

则称 $\hat{\theta}$ 为参数 θ 的无偏估计量 (unbiased estimator).

2. 有效性

定义 (7.1.2)

设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都为参数 θ 的无偏估计量，若对任意参数 $\theta \in \Theta$ 和固定的样本容量 n，有

$$D(\hat{\theta}_1) \leqslant D(\hat{\theta}_2),$$

则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效.

3. 一致性

定义 (7.1.3)

设 $\hat{\theta}_n(X_1, X_2, \cdots, X_n)$ 为未知参数 θ 的估计量, 若 $\forall \varepsilon > 0$, 有

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| \geqslant \varepsilon\right) = 0 \text{ 或 } \lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| < \varepsilon\right) = 1$$

成立, 则称 $\hat{\theta}_n$ 为参数 θ 的一致估计量 (uniformly estimator).

根据辛钦大数定律, 我们可证明样本均值为总体均值的一致估计量; 样本方差为总体方差的一致估计量; 样本的 k 阶矩为总体 k 阶矩的一致估计量.

典型例题

例 (1.1)

设某工厂有甲、乙、丙 3 个车间生产同一种产品，产量依次占全厂的 45%，35%，20%，且各车间的次品率分别为 4%，2%，5%，现在从一批产品中检查出 1 个次品，问该次品是由哪个车间生产的可能性最大？

解

设 A_1, A_2, A_3 表示产品来自甲、乙、丙三个车间，B 表示产品为“次品”的事件，易知 A_1, A_2, A_3 是样本空间 Ω 的一个划分，且有

$$P(A_1) = 0.45, P(A_2) = 0.35, P(A_3) = 0.2,$$

$$P(B \mid A_1) = 0.04, P(B \mid A_2) = 0.02, P(B \mid A_3) = 0.05.$$

解

由全概率公式得

$$\begin{aligned} P(B) &= P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + P(A_3)P(B \mid A_3) \\ &= 0.45 \times 0.04 + 0.35 \times 0.02 + 0.2 \times 0.05 = 0.035. \end{aligned}$$

由贝叶斯公式得

$$\begin{aligned} P(A_1 \mid B) &= (0.45 \times 0.04)/0.035 = 0.514, \\ P(A_2 \mid B) &= (0.35 \times 0.02)/0.035 = 0.200, \\ P(A_3 \mid B) &= (0.20 \times 0.05)/0.035 = 0.286. \end{aligned}$$

由此可见，该次品由甲车间生产的可能性最大.

例 (1.2)

某电子设备厂所用的晶体管由三家元件制造厂提供. 已知第一、二、三厂的次品率分别为 0.02, 0.01, 0.03, 又知三个厂提供晶体管的份额分别为 0.15, 0.80, 0.05, 设三个厂的产品是同规格的 (无区别标志), 且均匀的混合在一起. 求:

(1) 在混合的晶体管中随机的取一支是次品的概率;

(2) 在取出一支是次品的条件下, 它是由第二厂生产的概率是多少?

解

设 A 表示“取得的一支是次品”，B_i 表示“取得的一件产品是由第 i 厂生产的”，$i=1,2,3$. 则 B_1 、B_2 、B_3 构成了一个完备事件组. 由题意知, $P(A \mid B_1)=0.02, P(A \mid B_2)=0.01, P(A \mid B_3)=0.03, P(B_1)=0.15, P(B_2)=0.8, P(B_3)=0.05,$

(1) 由全概率公式得:

$$P(A)=\sum_{i=1}^{3} P(B_i)\,P(A \mid B_i)=0.15\times0.02+0.8\times0.01+0.05\times0.03=0.0125.$$

(2) 由贝叶斯公式得:

$$P(B_2 \mid A)=\frac{P(B_2)\,P(A \mid B_2)}{\sum_{i=1}^{3} P(B_i)\,P(A \mid B_i)}=\frac{0.8\times0.01}{0.0125}=\frac{16}{25}.$$

例 (2.1)

设连续型随机变量 X 的密度函数为

$$f(x)=\begin{cases} Ae^{-5x}, & x>0 \\ 0, & x\leqslant 0 \end{cases},$$

求:(1) 常数 A; (2) $P\{X>0.2\}$; (3) X 的分布函数.

解

(1) 因为 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_0^{+\infty} Ae^{-5x}\mathrm{d}x=1$, 所以 $A=5$.

(2) $P\{X>0.2\}=\int_{0.2}^{+\infty} 5e^{-5x}\mathrm{d}x=e^{-1}$.

(3) $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\begin{cases} 1-e^{-5x}, & x\geqslant 0; \\ 0, & x<0. \end{cases}$

例 (2.2)

连续型随机变量 X 的分布密度函数为

$$f(x)=\begin{cases} kx^2, & 0<x<1; \\ 0, & \text{其他}. \end{cases}$$

(1) 确定参数 k; (2) 求 X 的分布函数;

(3) 求 $P\{X\leqslant\frac{1}{2}\}$, $P\{-1<x<\frac{1}{3}\}$, $P\{\frac{1}{4}<X\leqslant 1\}$.

解

(1) 由密度函数的正则性知

$$1=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_0^1 kx^2\,\mathrm{d}x=\frac{k}{3},$$

所以 $k=3$. 即 X 的密度函数为 $f(x)=\begin{cases} 3x^2, & 0<x<1; \\ 0, & \text{其他}; \end{cases}$

解

(2) 当 $x<0$ 时，$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\int_{-\infty}^{x} 0\ \mathrm{d}t=0$;
当 $0\leqslant x<1$ 时，$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\int_{-\infty}^{0} 0\ \mathrm{d}t+\int_{0}^{x} 3t^2\ \mathrm{d}t=x^3$;
当 $x\geqslant 1$ 时，$F(x)=1$;

所以 X 的分布函数为 $F(x)=\begin{cases} 0, & x<0; \\ x^3, & 0\leqslant x<1; \\ 1, & x\geqslant 1; \end{cases}$

(3) $P\left(X\leqslant \frac{1}{2}\right)=F\left(\frac{1}{2}\right)=\left(\frac{1}{2}\right)^3=\frac{1}{8}$;

$P\left(-1<X<\frac{1}{3}\right)=F\left(\frac{1}{3}\right)-F(-1)=\frac{1}{27}-0=\frac{1}{27}$;

$P\left(\frac{1}{4}<X\leqslant 1\right)=F(1)-F\left(\frac{1}{4}\right)=1-\frac{1}{64}=\frac{63}{64}$.

例 (3.1)

随机变量 (X, Y) 的概率密度为 $f(x,y)=\begin{cases} k\,\mathrm{e}^{-(3x+4y)}, & x>0, y>0; \\ 0, & \text{其他}. \end{cases}$

(1) 确定常数 k; (2) 求 (X, Y) 的分布函数;
(3) 求 $P\{0<X\leqslant 1, 0<Y\leqslant 2\}$;
(4) 求 $f_X(x), f_Y(y)$; (5) X 与 Y 是否相互独立?

解 ((1) 由联合概率密度函数的完备性得)

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y\,\mathrm{d}x = \int_{0}^{+\infty}\int_{0}^{+\infty} k\mathrm{e}^{-3x-4y}\,\mathrm{d}x\,\mathrm{d}y$$
$$= k\int_{0}^{+\infty}\mathrm{e}^{-3x}\,\mathrm{d}x\int_{0}^{+\infty}\mathrm{e}^{-4y}\,\mathrm{d}y = \frac{k}{12} = 1,$$

所以, $k=12$.

解

(2) 由联合分布函数的定义，当 $x \leqslant 0$ 或 $y \leqslant 0$ 时，$f(x, y) = 0$，所以 $F(x, y) = 0$；当 $x \geqslant 0$ 且 $y \geqslant 0$ 时，

$$
\begin{aligned}
F(x, y) &= \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t) \mathrm{d}s \, \mathrm{d}t = \int_{0}^{x} \int_{0}^{y} 12 \mathrm{e}^{-3s-4t} \, \mathrm{d}s \, \mathrm{d}t \\
&= \left(1 - \mathrm{e}^{-3x}\right) \left(1 - \mathrm{e}^{-4y}\right);
\end{aligned}
$$

综上，分布函数的表达式为

$$
F(x, y) = \begin{cases} \left(1 - \mathrm{e}^{-3x}\right) \left(1 - \mathrm{e}^{-4y}\right), & x > 0, y > 0; \\ 0, & \text{其他}; \end{cases}
$$

(3)

$$
\begin{aligned}
P\{0 < X \leqslant 1, 0 < Y \leqslant 2\} &= \int_{0}^{1} \int_{0}^{2} f(x, y) \mathrm{d}y \, \mathrm{d}x \\
&= \left(1 - \mathrm{e}^{-3}\right) \left(1 - \mathrm{e}^{-8}\right).
\end{aligned}
$$

解

(4) 由边际密度函数的定义知

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y = \begin{cases} \int_0^{+\infty} 12\mathrm{e}^{-3x-4y}\,\mathrm{d}y, & x>0; \\ 0, & \text{其他}; \end{cases}$$

$$= \begin{cases} 3\mathrm{e}^{-3x}, & x>0; \\ 0, & \text{其他}; \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x = \begin{cases} \int_0^{+\infty} 12\mathrm{e}^{-3x-4y}\,\mathrm{d}x, & y>0; \\ 0, & \text{其他}; \end{cases}$$

$$= \begin{cases} 4\mathrm{e}^{-4y}, & y>0; \\ 0, & \text{其他}; \end{cases}$$

(5) 由于 $f(x,y) = f_X(x)\cdot f_Y(y)$, 所以 X 与 Y 相互独立.

例 (3.2)

假设随机变量 U 在区间 $[-2,2]$ 上服从均匀分布，随机变量

$$X=\begin{cases}-1, & \text{若 } U\leqslant -1;\\ 1, & \text{若 } U>-1.\end{cases},\quad Y=\begin{cases}-1, & \text{若 } U\leqslant 1;\\ 1, & \text{若 } U>1.\end{cases}$$

试求 X 和 Y 的联合概率分布，判断 X 与 Y 是否独立.

解

由于 U 在 $[-2,2]$ 上服从均匀分布，所以 U 的密度函数为

$$f(x)=\begin{cases}\frac{1}{4}, & -2<x<2;\\ 0, & \text{其他};\end{cases}$$

(X,Y) 的所有可能取值为 $(-1,-1)$, $(-1,1)$, $(1,-1)$, $(1,1)$,

解

$$
\begin{aligned}
P\{X=-1, Y=-1\} =& P\{U \leqslant -1, U \leqslant 1\} = P\{U \leqslant -1\} \\
&= \int_{-2}^{-1} \frac{1}{4} \, \mathrm{d}x = \frac{1}{4}, \\
P\{X=-1, Y=1\} =& P\{U \leqslant -1, U > 1\} = 0, \\
P\{X=1, Y=-1\} =& P\{U > -1, U \leqslant 1\} = P\{-1 < U \leqslant 1\} \\
&= \int_{-1}^{1} \frac{1}{4} \, \mathrm{d}x = \frac{1}{2}, \\
P\{X=1, Y=1\} =& P\{U > -1, U > 1\} = P\{U > 1\} = \int_{1}^{2} \frac{1}{4} \, \mathrm{d}x = \frac{1}{4}.
\end{aligned}
$$

解

所以 (X, Y) 的联合概率分布如下:

$X \backslash Y$	-1	1
-1	$1/4$	0
1	$1/2$	$1/4$

因为 $P(X=-1, Y=1)=0 \neq P(X=-1)P(Y=1)=\frac{1}{4}\cdot\frac{1}{4}$, 所以 X 与 Y 不独立.

例 (4.1)

已知随机变量 X 和 Y 分别服从正态分布 $N(1,3^2)$ 和 $N(0,4^2)$，且 X 与 Y 的相关系数 $\rho_{XY}=-1/2$，设 $Z=\dfrac{X}{3}+\dfrac{Y}{2}$.

(1) 求 Z 的数学期望 $E(Z)$ 和方差 $D(Z)$.

(2) 求 X 与 Z 的相关系数 ρ_{XZ}.

(3) 问 X 与 Z 是否相互独立，为什么？

解

(1) 由题意：$E(Z)=E\left(\dfrac{X}{3}+\dfrac{Y}{2}\right)=\dfrac{1}{3}E(X)+\dfrac{1}{2}E(Y)=\dfrac{1}{3}$;

$$
\begin{aligned}
D(Z)=D\left(\frac{X}{3}+\frac{Y}{2}\right)&=\frac{1}{9}D(X)+\frac{1}{4}D(Y)+2\operatorname{cov}\left(\frac{X}{3},\frac{Y}{2}\right)\\
&=\frac{1}{9}\cdot 3^2+\frac{1}{4}\cdot 4^2+2\times\frac{1}{3}\times\frac{1}{2}\operatorname{cov}(X,Y)\\
&=5+\frac{1}{3}\sqrt{D(X)}\cdot\sqrt{D(Y)}\cdot\rho_{XY}=3;
\end{aligned}
$$

解

(2) X 与 Z 的协方差:

$$\begin{aligned}\operatorname{cov}(X, Z) &= \operatorname{cov}\left(X, \frac{X}{3}+\frac{Y}{2}\right) = \operatorname{cov}\left(X, \frac{X}{3}\right) + \operatorname{cov}\left(X, \frac{Y}{2}\right) \\ &= \frac{1}{3}\operatorname{cov}(X, X) + \frac{1}{2}\operatorname{cov}(X, Y) \\ &= \frac{1}{3}\times 3^2 + \frac{1}{2}\sqrt{D(X)}\cdot\sqrt{D(Y)}\cdot\rho_{XY} = 0.\end{aligned}$$

故 X 与 Z 的相关系数

$$\rho_{XZ} = \frac{\operatorname{cov}(X, Z)}{\sqrt{D(X)}\cdot\sqrt{D(Z)}} = 0;$$

(3) 由于 (X, Z) 不一定服从二维正态分布, 故由 $\rho_{XZ} = 0$ 不能确定 X 与 Z 是否相互独立.

例 (4.2)

设 (X, Y) 的概率密度函数为 $f(x,y)=\begin{cases}\dfrac{3xy}{16}, & 0\leqslant x\leqslant 2, 0\leqslant y\leqslant x^2;\\ 0, & \text{其他}.\end{cases}$

求: (1) $f_X(x)$, $f_Y(y)$; (2) $E(X)$, $E(Y)$, $D(X)$, $D(Y)$; (3) $\operatorname{cov}(X, Y)$ 与 ρ_{XY}.

解

(1) 由边际密度函数的定义

$$f_X(x)=\int_0^{x^2}\frac{3xy}{16}\,\mathrm{d}y=\frac{3x^5}{32}, 0\leqslant x\leqslant 2;$$

$$f_Y(y)=\int_{\sqrt{y}}^{2}\frac{3xy}{16}\,\mathrm{d}x=\frac{3y(4-y)}{32}, 0\leqslant y\leqslant 4;$$

$$\therefore f_X(x)=\begin{cases}\dfrac{3x^5}{32}, & 0\leqslant x\leqslant 2;\\ 0, & \text{其他}.\end{cases}\quad f_Y(y)=\begin{cases}\dfrac{3y(4-y)}{32}, & 0\leqslant y\leqslant 4;\\ 0, & \text{其他}.\end{cases}$$

解

(2) 由期望与方差的计算公式

$$E(X)=\int_0^2 x\cdot 3x^5/32\cdot \mathrm{d}x=12/7;$$
$$E(Y)=\int_0^4 y\cdot 3y(4-y)/32\cdot \mathrm{d}y=2.$$

又因为

$$E\left(X^2\right)=\int_0^2 x^2\cdot 3x^5/32\cdot \mathrm{d}x=3;$$
$$E\left(Y^2\right)=\int_0^4 y^2\cdot 3y(4-y)/32\cdot \mathrm{d}y=24/5;$$

所以

$$D(X)=3-(12/7)^2=3/49;\ D(Y)=24/5-2^2=4/5.$$

解

(3) 由协方差计算公式

$$E(XY) = \int_0^2 \int_0^{x^2} xy \cdot 3xy/16 \cdot \mathrm{d}x\,\mathrm{d}y = 32/9;$$

所以

$$\mathrm{cov}(X, Y) = E(XY) - E(X)E(Y) = 8/63;$$

$$\rho_{XY} = \frac{\mathrm{cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{4\sqrt{15}}{27} \approx 0.574.$$

例 (5.1)

茶叶用机器袋装，每袋的净重与标准重量相比存在一定的偏差，设每袋净重为随机变量，其期望值为 100 克，标准差为 10 克，一箱内装 200 袋，求一箱茶叶净重超过 20.5 千克的概率．(计算结果用标准正态分布函数值表示)

解

设箱内第 i 袋茶叶重量为 X_i 克，$(i = 1, 2, \cdots, 200)$，显然 $X_1, X_2, \cdots, X_{200}$ 独立同分布，由题意知 $E(X_i) = 100$，$D(X_i) = 100$．由独立同分布中心极限定理得，$\sum_{i=1}^{200} X_i$ 近似服从 $N(20000, 20000)$，所以

$$P\left(\sum_{i=1}^{200} X_i > 20500\right) = P\left(\frac{\sum_{i=1}^{200} X_i - 20000}{100\sqrt{2}} > \frac{20500 - 20000}{100\sqrt{2}}\right)$$

$$\approx 1 - \Phi(3.54).$$

例 (5.2)

某市年满 22 岁的居民中 20% 受过高等教育. 今从年满 22 岁的居民中随机抽取 10000 人, 求受过高等教育的人数在 1960 和 2040 之间的概率. $(\Phi(1) = 0.8413)$

解

设受高等教育的人数为 X, 则 $X \sim B\left(10000, \frac{1}{5}\right)$

$$E(X) = 10000 \times \frac{1}{5} = 2000, D(X) = 10000 \times \frac{1}{5} \times \frac{4}{5} = 1600.$$

由棣莫弗 · 拉普拉斯中心极限定理知: X 近似服从 $N(2000, 40^2)$,

$$P(1960 < X < 2040) \approx \Phi\left(\frac{2040 - 2000}{40}\right) - \Phi\left(\frac{1960 - 2000}{40}\right)$$
$$= 2\Phi(1) - 1 = 0.6826.$$

例 (6.1)

设 ξ_1,ξ_2,ξ_3,ξ_4 为来自总体 $N\left(1,\sigma^2\right)\ (\sigma>0)$ 的简单随机样本, 求统计量 $\dfrac{\xi_1-\xi_2}{|\xi_3+\xi_4-2|}$ 的分布.

解

由正态分布的性质知 $\xi_1-\xi_2\sim N\left(0,2\sigma^2\right)$, $\xi_3+\xi_4\sim N\left(2,2\sigma^2\right)$. 所以,

$$\frac{\xi_1-\xi_2}{\sqrt{2}\sigma}\sim N(0,1),\quad \frac{\xi_3+\xi_4-2}{\sqrt{2}\sigma}\sim N(0,1),\ \left(\frac{\xi_3+\xi_4-2}{\sqrt{2}\sigma}\right)^2\sim\chi^2(1),$$

由于 $\dfrac{\xi_1-\xi_2}{\sqrt{2}\sigma}$ 与 $\left(\dfrac{\xi_3+\xi_4-2}{\sqrt{2}\sigma}\right)^2$ 相互独立, 从而由 $t-$ 分布构造, 有

$$\frac{\xi_1-\xi_2}{|\xi_3+\xi_4-2|}=\frac{\xi_1-\xi_2}{\sqrt{2}\sigma}\Bigg/\sqrt{\left(\frac{\xi_3+\xi_4-2}{\sqrt{2}\sigma}\right)^2}\sim t(1).$$

例 (6.2)

设总体 ξ 服从标准正态分布 $N(0,1)$, $\xi_1, \xi_2, \cdots, \xi_n$ 是来自总体 ξ 的一个简单随机样本, 试问统计量

$$\eta = \left(\frac{n}{5} - 1\right) \sum_{i=1}^{5} \xi_i^2 \Big/ \sum_{i=6}^{n} \xi_i^2, \quad n > 5.$$

服从何种分布?

解

因为 $\xi_i \sim N(0,1)$, $\sum\limits_{i=1}^{5} \xi_i^2 \sim \chi^2(5)$, $\sum\limits_{i=6}^{n} \xi_i^2 \sim \chi^2(n-5)$, 且 $\sum\limits_{i=1}^{5} \xi_i^2$ 与 $\sum\limits_{i=6}^{n} \xi_i^2$ 相互独立, 所以

$$\eta = \frac{\sum\limits_{i=1}^{5} \xi_i/5}{\sum\limits_{i=6}^{n} \xi_i^2/(n-5)} \sim F(5, n-5).$$

例 (7.1)

设总体 X 的密度函数为 $f(x,\theta)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1\\ 0, & \text{其他}\end{cases}$,$\theta>-1$，求参数 θ 的矩估计.

解

因为

$$E(X)=\int_{-\infty}^{+\infty}xf(x,\theta)\mathrm{d}x=\int_{0}^{1}(\theta+1)x^{\theta+1}\ \mathrm{d}x=\frac{\theta+1}{\theta+2},$$

令 $E(X)=\overline{X}$, 解得

$$\theta=\frac{2\overline{X}-1}{1-\overline{X}},$$

于是参数 θ 的矩估计为

$$\hat{\theta}=\frac{2\overline{X}-1}{1-\overline{X}}.$$

例 (7.2)

设总体 X 的密度函数为 $f(x,\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0,\\ 0, & \text{其他}.\end{cases}$, $\theta>0$, 求参数 θ 的极大似然估计.

解

参数 θ 的极大似然函数为 $L(\theta,x_1,\cdots,x_n)=\prod\limits_{i=1}^{n} f(x_i,\theta)=\theta^{-n}e^{-\frac{1}{\theta}\sum\limits_{i=1}^{n}x_i}$,

则对数似然函数为 $\ln L(\theta,x_1,\cdots,x_n)=-n\ln\theta-\frac{1}{\theta}\sum\limits_{i=1}^{n}x_i$. 令

$$\frac{\mathrm{d}\ln L(\theta,x_1,\cdots,x_n)}{\mathrm{d}\theta}=\frac{\mathrm{d}\left(-n\ln\theta-\frac{1}{\theta}\sum\limits_{i=1}^{n}x_i\right)}{\mathrm{d}\theta}=0,$$

解得 $\theta=\overline{x}$, 于是参数 θ 的极大似然估计量为 $\hat{\theta}=\overline{X}$.