概率论与数理统计

主讲：杨宇明

2021 年 12 月 31 日

概率论与数理统计专题讲座

目录

1 概率论

概率核心问题:

1. 古典概率
2. 概率的公理化定义
3. 随机变量刻画概率模型的方法
   a 一维
   b 多维
4. 数字特征
5. 近似计算与理论推导

从计算的角度看

有三个核心问题:

1. 概率怎么算?
2. 算概率的东西怎么算?
3. 概率分布能算什么?

1.1 概率论的基本概念

主要内容

目录

概率论的基本概念

1. 基本概念
   (a) 事件 ⇔ 集合
   (b) 事件的关系与运算
2. 概率计算
   (a) 直接计算
   (b) 间接计算

事件的关系和运算

记号	集合论	概率论
Ω	空间，全集	样本空间，必然事件
Φ	空集	不可能事件
A	Ω 的子集	事件
$\overline{A}$	A 的余集	A 的对立事件
$\omega \in A$	ω 属于集合 A	事件 A 发生
$\omega \notin A$	ω 不属于集合 A	事件 A 不发生
$A \subset B$	A 是 B 的子集	A 发生，则 B 必发生
$A = B$	A 与 B 相等	A 与 B 等价
$A \cup B$	A 与 B 的和	A 与 B 至少有一个发生
$A \cap B$	A 与 B 的交	A 与 B 同时发生
$A \cap B = \Phi$	A 与 B 无公共元素	A 与 B 互不相容
$A \cap B = \Phi, A \cup B = \Omega$	A 与 B 互为余集	A 与 B 对立
$A - B$	A 与 B 的差集	A 发生且 B 不发生

本章易混淆概念区分:

1. $P(A)$,$P(A|B)$ 与 $P(AB)$

$P(A)$: 试验前预测事件 A 发生的概率

$P(A|B)$: 试验中当事件 B 已发生时，预测事件 A 发生的概率

$P(AB)$: 试验前预测事件 A, B 同时发生的概率

一个简单的例子:

设袋中有 3 个白球，2 个红球，现从袋中任意抽取两次，每次取一个，取后不放回 设 A={第二次取到红球},B={第一次取到红球}. 考虑此例中 $P(A), P(A|B), P(AB)$?

本章易混淆概念区分:

2. 互不相容事件、对立事件、相互独立事件

互不相容事件: 若 $AB = \Phi$, 称 A, B 为互不相容或互斥事件, 即事件 A, B 不可能同时发生。

对立事件: 若 $A \cap B = \Phi, A \cup B = \Omega$, 称 A 与 B 互为对立事件

相互独立事件: $P(AB) = P(A)P(B)$

1、直接计算 $P(A) = \frac{m}{n}$

古典概型:

1. 仅有有限多个基本事件;
2. 每个基本事件发生的可能性相等。

概率直接计算方法:
设 E 为古典概型试验, $A_i, i = 1, 2, \ldots, n$ 是基本事件, 则由

$$P(A) = \frac{A\text{所含的基本事件个数}}{\text{基本事件个数}} = \frac{A\text{所含的含样本点的数目}}{\text{样本空间的样本点总数}}$$

所确定的概率称为事件 A 的古典概率。

2、间接计算

1. 条件概率 $P(A|B) = \frac{P(AB)}{P(B)}$
2. 乘法公式 $P(AB) = P(B)P(A|B)$(蕴含相互独立公式)
3. 全概率公式 $P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$
4. 贝叶斯公式 $P(B_j|A) = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$

5. 加法公式 $P(A\cup B)=P(A)+P(B)-P(AB)$

6. 单调性公式 $P(B-A)=P(B)-P(AB)$

典型例题:

(2012) 设 A,B,C 是随机事件, A,C 互不相容, $P(AB)=\dfrac{1}{2}, P(C)=\dfrac{1}{3}$, 则 $P(AB|\overline{C})$=______。

$$
\begin{aligned}
\text{解：}P(AB|\overline{C})=&\frac{P(AB\overline{C})}{P(\overline{C})}\\
=&\frac{P(AB)-P(ABC)}{1-P(C)}\\
=&\frac{\frac{1}{2}-0}{1-\frac{1}{3}}\\
=&\frac{3}{4}
\end{aligned}
$$

典型例题:

例.(2014)

设事件 A,B 相互独立, 且 $P(B)=0.5, P(A-B)=0.3$, 则 $P(B-A)=(\quad)$

(a) 0.1, (b) 0.2, (c) 0.3, (d) 0.4

例.(2015)

设 A,B 为任意两个随机事件，则 ( )

(a) $P(AB)\le P(A)P(B)$ (b) $P(AB)\ge P(A)P(B)$

(c) $P(AB)\le \dfrac{P(A)+P(B)}{2}$ (d) $P(AB)\ge \dfrac{P(A)+P(B)}{2}$

典型例题:

例.(2017)

设 A,B,C 为三个随机事件, A 与 C 相互独立, B 与 C 相互独立, 则 $A\cup B$ 与 C 相互独立的充要条件是 ( )

(a) A 与 B 相互独立　(b) A 与 B 互不相容

(c) AB 与 C 相互独立　(d) AB 与 C 互不相容

例.(2018)

A,B 独立, A,C 独立, $BC=\phi, P(A)=P(B)=0.5, P(AC|AB\cup C)=0.25$,
则 $P(C)=$ __________

典型例题

例 6.

n 个战士各有一支自己使用的枪，外形完全一样。一次夜间紧急集合，每人随机拿了一支枪，求

a. 至少一人拿到自己的枪的概率。

b. 恰有 k 个人拿到自己的枪的概率

分析: 在这个问题中, 有一类事件的概率是容易算的, 即第 i 个人拿到自己的枪。(基本事件概率一般好算)

解：设 $A=\{$至少一人拿到自己的枪$\}$

$B=\{$恰有 k 个人拿到自己的枪$\}$

$A_i=\{$第 i 个人拿到自己的枪$\}$，$i=1,\cdots,n$

$$a.\ P\left(\bigcup_{i=1}^{n}A_i\right)=S_1-S_2+S_3-\cdots+(-1)^{n+1}S_n$$

$$\text{其中：}S_1=\sum_{i=1}^{n}P(A_i)$$

$$S_2=\sum_{1\le i<j\le n}P(A_iA_j)$$

$$S_3=\sum_{1\le i<j<k\le n}P(A_iA_jA_k)$$

$$\cdots$$

$$S_n=P(A_1A_2\cdots A_n)$$

$$P\left(\bigcup_{i=1}^{n}A_i\right)=C_n^1\frac{1}{n}-C_n^2\frac{(n-2)!}{n!}+\cdots+(-1)^{n+1}C_n^n\frac{1}{n!}=\sum_{i=1}^{n}\frac{(-1)^{i+1}}{i!}$$

b.分析：事件 B={恰有 k 个人拿到自己的枪} 等价于于先从 n 个人中选 k 个，让他们都拿到自己的枪，剩下 $n-k$ 个人则都不能拿到自己的枪

$$\begin{aligned}P\left(B\right)=&C_n^k\frac{(n-k)!}{n!}\cdot\left(1-\sum_{i=1}^{n-k}\frac{(-1)^{i+1}}{i!}\right)\\=&\frac{1}{k!}\sum_{i=1}^{n-k}\frac{(-1)^{i+1}}{i!}\end{aligned}$$

典型例题

例: 不断抛掷两颗骰子，设 A = {两颗骰子点数之和为 5}，B= {两颗骰子点数之和为 7}，求 A 在 B 之前发生的概率。

解：设 C={A 在 B 之前发生}

考虑第一次抛掷的三种结果:

D_1={第一次抛掷出现 A}　　　　则 D_1, D_2, D_3 构成样本空间

D_2={第一次抛掷出现 B}

D_3={第一次抛掷 A, B 均未出现}

的有限划分,故使用全概率公式有: $P(C) = P(C|D_1)P(D_1)+P(C|D_2)P(D_2)+P(C|D_3)P(D_3)$

$$\begin{aligned}P(C)=&P(C|D_1)P(D_1)+P(C|D_2)P(D_2)+P(C|D_3)P(D_3)\\=&1\times P(A)+0\times P(B)+P(C)P(\bar{A}\bar{B})\\=&1\times\frac{1}{9}+0\times\frac{1}{6}+P(C)\left(1-\frac{1}{9}-\frac{1}{6}\right)\\=&\frac{1}{9}+\frac{13}{18}P(C)\\\Rightarrow&\frac{5}{18}P(C)=\frac{1}{9}\\\Rightarrow&P(C)=\frac{2}{5}\end{aligned}$$

¶ 理解基本概念，概念是讨论的基础；

¶ 要正确表示出事件；

¶ 能够分析清楚事件之间的关系；

¶ 能够按照分析出的关系选取正确的公式计算概率。

1.2 随机变量的分布

主要内容

目录

随机变量的分布

三种刻画概率分布的方法：

1. 分布函数（普适刻画方法）
2. 分布律（离散型）
3. 概率密度（连续型）

六个个常见分布：

离散型	连续型
$0-1$ 分布	$U(a,b)$
$B(n,p)$	$EXP(\lambda)$
$P(\lambda)$	$N(\mu,\sigma^2)$

能够做到：

1. 分布函数 $\rightleftharpoons$ 分布律（离散型）
2. 分布函数 $\rightleftharpoons$ 概率密度（连续型）
3. 根据问题，判断其是否是常见分布

本章易混淆概念区分：

1. 概率密度函数与分布函数的性质
2. 连续型随机变量的分布函数一定是连续函数，但密度函数未必是连续函数。

3. 取值连续的随机变量不一定是连续型随机变量。例:

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

典型例题:

(2011) 设 $F_1(x), F_2(x)$ 为两个分布函数，其相应的概率密度为连续函数 $f_1(x), f_2(x)$，则下列必为概率密度的是 ( )

(1) $f_1(x)f_2(x)$ (2) $2f_2(x)F_1(x)$

(3) $f_1(x)F_2(x)$ (4) $f_1(x)F_2(x)+f_2(x)F_1(x)$

分析：需满足密度函数的非负性和规范性

例:设 $F_1(x), F_2(x)$ 为两个分布函数，其对应的概率密度函数分别为 $f_1(x)$ 与 $f_2(x)$，试判断以下哪个函数是 (联合) 概率密度函数，并说明理由。

(1) $2f_1(x)-f_2(x), x\in R$ 不一定

(2) $2f_1(x)f_1(y), -\infty<x\le y<+\infty$ 是

(3) $f_1(x)F_2(x)+f_2(x)F_1(x), x\in R$ 是

典型例题:

(2013)

设随机变量 $Y\sim E(1)$, α 大于零且为常数, 则$P\{Y\le a+1|Y>a\}=$ ______

分析：指数分布有无后效性

$$\begin{aligned}P\{Y\le a+1|Y>a\}=&P\{Y\le 1\}\\ =&F(1)\\ =&1-e^{-1}\end{aligned}$$

典型例题:

(2004) 设随机变量 $X \sim E(\lambda)$，则 $P\{X > \sqrt{DX}\} =$ ____.

(2006) 设随机变量 $X \sim N(\mu_1, \sigma_1^2), Y \sim N((\mu_2, \sigma_2^2)$，且 $P\{|X - \mu_1| < 1\} > P\{|Y - \mu_2| < 1\}$ 则 (　)

(1)　$\sigma_1 < \sigma_2$，　(2)　$\sigma_1 < \sigma_2$，　(3)　$\mu_1 > \mu_2$，　(4)　$\mu_1 > \mu_2$

(2016) 设随机变量$X \sim N(\mu, \sigma^2), (\sigma > 0)$, 记 $p = P\{X \le \mu + \sigma^2\}$，则 ()

(1)　p 随着 μ 的增加而增加　(2)　p 随着 σ 的增加而增加

(3)　p 随着 μ 的增加而减少　(4)　p 随着 σ 的增加而减少

分析：考察正态分布概率计算

$$P\{X \le x\} = \Phi(\frac{x - \mu}{\sigma})$$

典型例题:

(2010) 设 $f_1(x)$ 为标准正态分布的概率密度，$f_2(x)$ 为 [-1,3] 上的均匀分布的概率密度，若 $f(x) = \begin{cases} af_1(x), & x \le 0 \\ bf_2(x), & x > 0 \end{cases}$ $(a > 0, b > 0)$ 为概率密度，则 a, b 应满足 (　):

(1)　$2a + 3b = 4$,　(2)　$3a + 2b = 4$,

(3)　$a + b = 1$,　(4)　$a + b = 2$

(2013) 设 X_1, X_2, X_3 是随机变量，且 $X_1 \sim N(0,1), X_2 \sim N(0, 2^2), X_3 \sim N(5, 3^2), p_i = P(-2 \le X_i \le 2)(i = 1, 2, 3)$，则 (　)

(1)　$p_1 > p_2 > p_3$,　(2)　$p_2 > p_1 > p_3$,

(3)　$p_3 > p_1 > p_2$,　(4)　$p_1 > p_3 > p_2$

典型例题:

求分布函数的例子

设随机变量 X 的绝对值不大于 1

$$P\{X = -1\} = 1/8, P\{X = 1\} = 1/4$$

当 $-1 < X < 1$ 时，X 在区间 $(-1, 1)$ 内的任一子区间取值的概率与该子区间长度成正比，试求 X 的分布函数。

分析：

a 根据随机变量值域可直接写出分布函数等于 0 和 1 的两段

b 随机变量 X 在区间 (-1,1) 上类似均匀分布

$$F(x) = \begin{cases} 0 & x < -1 \\ \frac{7+5x}{16} & -1 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$ 中间概率只剩 5/8, 所以 (x+1)/2*5/8+1/8

1.3 多维随机变量与数字特征

主要内容

目录

多维随机变量的概率计算问题

两个核心问题:

1. 概率怎么算？
2. 算概率的东西怎么算？

利用：

1. 第一章的概念、性质和公式；
2. 联合分布函数计算；
3. 联合分布律计算；
4. 联合概率密度计算；
5. 边缘分布计算；
6. 独立性计算；
7. 条件分布计算；
8. 常见分布特殊计算；

提醒：概率论中，如果没有思路，试着把分布找出来

随机变量函数的数学期望 (这是第四章的核心概念!!!)

求随机变量函数的数学期望的主要方法:

利用:

1. 定义算 (先求出函数的分布);
2. 期望的性质来算，特别是性质 $E(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} E(X_k)$
3. 一元函数的期望公式来算 $g(X)$ 的期望，

   即: $Eg(X) = \begin{cases} \sum_k g(x_k)p_k \\ \int g(x)f(x)dx \end{cases}$
4. 二元函数的期望公式来算 $g(X,Y)$ 的期望，即:

   $E(G(X,Y)) = \begin{cases} \sum_i \sum_j g(x_i,y_j)p_{ij} \\ \int\int g(x,y)f(x,y)dxdy \end{cases}$
5. 常见随机变量数字特征公式算;

请注意: 大多数情况下，求得随机变量或向量的分布是关键!

典型问题

例: 设随机变量 X,Y 相互独立，且 $X \sim N(0,1)$，Y 的概率分布为 $P\{Y=0\} = P\{Y=1\} = 0.5$，求 $P\{XY < 1.96\}$?

分析: Y 只有两个取值。可否分情况讨论? 怎么讨论? 解: 按全概率公式:

$$\begin{aligned} P\{XY \le 1.96\} =& P\{XY \le 1.96, Y=0\} + P\{XY \le 1.96, Y=1\} \\ =& P\{0 \le 1.96, Y=0\} + P\{X \le 1.96, Y=1\} \\ =& 0.5 + P\{X \le 1.96\}P\{Y=1\} \\ =& 0.5 + 0.975/2 = 0.9875 \end{aligned}$$

进一步，求 $P\{XY \le z\}$? 怎么计算?

典型例题

例：设随机变量 X,Y 相互独立，且 $X\sim N(0,1)$，Y 的概率分布为 $P\{Y=0\}=P\{Y=1\}=0.5$，求 $P\{XY<z\}$?

分析：Y 只有两个取值。还是分两种情况讨论。

解：按全概率公式：

$$\begin{aligned}P\{XY\le z\}=&P\{XY\le z,Y=0\}+P\{XY\le z,Y=1\}\\=&P\{0\le z,Y=0\}+P\{X\le z,Y=1\}\\=&\begin{cases}0.5\Phi(z), & z<0\\0.5+0.5\Phi(z), & z\ge 0\end{cases}\end{aligned}$$

若记 $F_Z(z)$ 为随机变量 $Z=XY$ 的分布函数，则 $F_Z(z)$ 间断点个数为? 答案：一个

典型例题:

例：设随机变量 X 与 Y 相互独立，X 的概率分布为 $P(X=i)=\frac{1}{3}(i=-1,0,1)$, 且 Y 的概率密度为

$$f_Y(y)=\begin{cases}1, & 0\le y<1,\\0, & \text{其它}\end{cases}$$

记 $Z=X+Y$. (1) $P\{Z\le\frac{1}{2}|X=0\}$; (2) 求 Z 的概率密度 $f_Z(z)$。

- $P\{X=0\}>0$，根据条件概率定义算；
- 分布函数求导！

解：(1) 由条件概率定义：

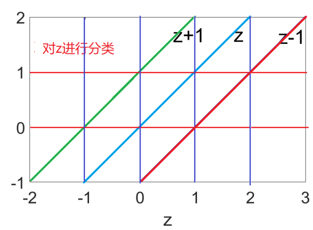
$$
\begin{aligned}
P\{Z \le \frac{1}{2}|X=0\} =& P\{X+Y \le \frac{1}{2}|X=0\} \\
=& P\{Y \le \frac{1}{2}|X=0\} \\
=& P\{Y \le \frac{1}{2}\} \\
=& \frac{1}{2}
\end{aligned}
$$

解：(2) 先求分布函数，再求导

$$
\begin{aligned}
F_Z(z) =& P\{Z \le z\} \\
=& P\{X+Y \le z\} \\
=& \sum_{i=-1}^{1} P\{X+Y \le z, X=i\} \\
=& P\{-1+Y \le z, X=-1\} + P\{Y \le z, X=0\} \\
& + P\{1+Y \le z, X=1\} \\
=& \frac{1}{3}(P\{Y-1 \le z\} + P\{Y \le z\} + P\{Y+1 \le z\}) \\
=& \frac{1}{3}\left[F_Y(z+1) + F_Y(z) + F_Y(z-1)\right]
\end{aligned}
$$

解：(2) 再求导

$$
\begin{aligned}
f_Z(z) =& F_Z^{'}(z) \\
=& \frac{1}{3}\left[f_Y(z+1)+f_Y(z)+f_Y(z-1)\right] \\
=& \begin{cases} \frac{1}{3}, & -1<z<2 \\ 0, & \text{其它} \end{cases}
\end{aligned}
$$



1) 横坐标用于讨论 z2) 直线用来讨论概率密度自变量

一个公式： $\boxed{f_Z(z)=\sum_{k=1}^{n} p\{X=x_k\} f_Y(z-x_k)}$

典型例题:

例: 设 A, B 为随机事件，且$P(A)=\dfrac{1}{4}, P(B|A)=\dfrac{1}{3}, P(A|B)=\dfrac{1}{2}$, 令

$$
X=\begin{cases} 1, & A\text{发生} \\ 0, & A\text{不发生} \end{cases} \quad Y=\begin{cases} 1, & B\text{发生} \\ 0, & B\text{不发生} \end{cases}
$$

求二维随机变量 (X, Y) 的概率分布；

分析：X, Y 都是 0-1 分布，(X, Y) 是二维离散型随机变量。考虑其概率分布即考虑

1. (X, Y) 在平面上可取几个点?

2. 每个点出现的概率是多少?

解：(X,Y) 可取四个点：$(0,0),(0,1),(1,0),(1,1)$ 根据定义

$$\begin{aligned}
P\{X=0,Y=0\}=&P(\bar{A}\bar{B})\\
P\{X=0,Y=1\}=&P(\bar{A}B)\\
P\{X=1,Y=0\}=&P(A\bar{B})\\
P\{X=1Y=1\}=&P(AB)
\end{aligned}$$

问题的关键是计算：$P(AB),P(\bar{A}B),P(A\bar{B}),P(\bar{A}\bar{B})$

由已知

$$\begin{aligned}
&P(AB)=P(A)P(B|A)=\frac{1}{12}=P(B)P(A|B)\\
\Rightarrow &P(B)=P(AB)/P(A|B)=\frac{1}{6}\\
\Rightarrow &P(A\cup B)=P(A)+P(B)-P(AB)=\frac{1}{3}\\
\Rightarrow &P(A\bar{B})=P(A)-P(AB)=\frac{1}{6}\\
\Rightarrow &P(\bar{A}B)=P(B)-P(AB)=\frac{1}{12}\\
\Rightarrow &P(\bar{A}\bar{B})=1-P(A\cup B)=\frac{2}{3}
\end{aligned}$$

所以 (X,Y) 的联合分布律为：

$$\begin{aligned}
P\{X=0,Y=0\}=&P(\bar{A}\bar{B})=\frac{2}{3}\\
P\{X=0,Y=1\}=&P(\bar{A}B)=\frac{1}{12}\\
P\{X=1,Y=0\}=&P(A\bar{B})=\frac{1}{6}\\
P\{X=1Y=1\}=&P(AB)=\frac{1}{12}
\end{aligned}$$

即：

X \ Y	0	1
0	2/3	1/12
1	1/6	1/12

如何求 X 和 Y 的相关系数 ρ_{XY}？

典型例题:

例： 已知随机变量 X,Y 以及 XY 的分布律如下表所示

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

Y	0	1	2
P	1/3	1/3	1/3

XY	0	1	2	4
P	7/12	1/3	0	1/12

求: (1) $P\{X=2Y\}$, (2) $cov(X-Y,Y)$ 与 ρ_{XY}

分析:

1. 二维离散型随机变量一般用联合分布律来刻画其概率分布;
2. 根据分布律，可进行任何关于 (X,Y) 的讨论;
3. 在取值有限的情况下，可以用列表的方法求解，比较直观。

解： 将随机变量 (X,Y) 联合分布律列成表格

X \ Y	0	1	2	$p_{i\cdot}$
0	1/4	0	1/4	1/2
1	0	1/3	0	1/3
2	1/12	0	1/12	1/6
$p_{\cdot j}$	1/3	1/3	1/3	1

$$
\begin{aligned}
P\{X=2Y\} =& P\{X=0,Y=0\}+P\{X=2,Y=1\} \\
=& 1/4+0 \\
=& 1/4
\end{aligned}
$$

解： (2) 分析：

1. X, Y 的期望和方差由定义或计算公式直接可算；
2. $cov(X-Y, Y) = cov(X, Y) - D(Y)$；
3. $\rho_{XY} = \dfrac{E(XY) - E(X)E(Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$；
4. 计算关键：$E(XY)$，可利用二元函数期望求解。

$$\begin{aligned} E(XY) =& 1 \times 1 \times 1/3 + 2 \times 2 \times 1/12 \\ =& 2/3 \end{aligned}$$

其它从略

典型例题

例： 设随机变量 X 的概率密度为 $f(x) = \begin{cases} \dfrac{1}{9}x^2, & 0 < x < 3 \\ 0, & \text{其他} \end{cases}$，令随机变量

$$Y = \begin{cases} 2, & X \le 1 \\ X, & 1 < X < 2 \\ 1, & X \ge 2 \end{cases}$$

求： (1) 求 Y 的分布函数； (2) 求概率 $P\{X \le Y\}$。

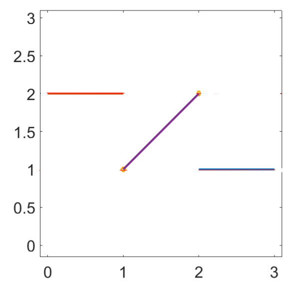
分析：

1. Y 的取值是分段的，分段讨论是关键；
2. 可以用图像辅助讨论；

解：(1)Y 的取值范围为 $[1, 2]$，故 当 $y < 1$ 时，$F(y) = P\{Y \le y\} = 0$，当 $y \ge 2$ 时，$F(y) = P\{Y \le y\} = 1$

当 $1 \leq y < 2$ 时

$$\begin{aligned}F(y) =& P\{Y \leq y\} = P\{Y = 1\} + P\{1 < Y \leq y\} \\ =& P\{2 \leq X \leq 3\} + P\{1 < X \leq y\} \\ =& \int_2^3 \frac{x^2}{9} dx + \int_1^y \frac{x^2}{9} dx = \frac{18 + y^3}{27}\end{aligned}$$



所以 Y 的分布函数为:

$$F(y) = \begin{cases} 0, & y < 1 \\ \dfrac{18 + y^3}{27}, & 1 \leq y < 2 \\ 1, & y \geq 2 \end{cases}$$

解: (2) 由:

$$Y = \begin{cases} 2, & X \leq 1 \\ X, & 1 < X < 2 \\ 1, & X \geq 2 \end{cases}$$

可知 (根据图像同样也可得到该结果，似乎更容易)

$$\{X \leq Y\} = \{1 < X < 2\} \cup \{X \leq 1\}$$

所以

$$
\begin{aligned}
P\{X \le Y\} =& P\{1 < X < 2\} + P\{X \le 1\} \\
=& \int_1^2 \frac{x^2}{9} dx + \int_0^1 \frac{x^2}{9} dx \\
=& \frac{8}{27}
\end{aligned}
$$

典型例题

例： 设随机变量 Y 的分布函数 $F(y)$ 连续且严格单调增加，而随机变量 $X \sim U(0,1)$，令 $Z = F^{-1}(X)$，则 Z 与 Y 有相同分布

分析：$F(y)$ 连续且严格单调增加，意味着 $F^{-1}(X) \in (-\infty, +\infty)$ 证明：

$$
\begin{aligned}
F_Z(z) &= P\{Z \le z\} \\
&= P\{F^{-1}(X) \le z\} (\text{这里 } F^{-1}(X) \in (-\infty, +\infty)\text{，不需要讨论 z}) \\
&= P\{X \le F(z)\} \\
&= F(z) \text{因为 } F(z) \in (0,1), X \sim U(0,1)
\end{aligned}
$$

这表明 Z 与 Y 有相同的分布函数。 这个方法通常用来由均匀分布生成其它类型的分布。

典型例题

例： 设随机变量 X 的分布函数 $F(x)$ 连续且严格单调增加，而随机变量 $Y = F(X)$，求 Y 的分布函数?

分析：求分布函数一般要讨论随机变量的值域 解： 由分布函数性质，$Y \in [0,1]$，所以 $F_Y(y) = P\{Y \le y\}$ 当 $y < 0$ 时，$F_Y(y) = 0$ 当 $y \ge 1$ 时，$F_Y(y) = 1$ 当 $0 \le y < 1$ 时，$F_Y(y) = P\{F(X) \le y\} = P\{X \le F^{-1}(y)\} = y$

所以

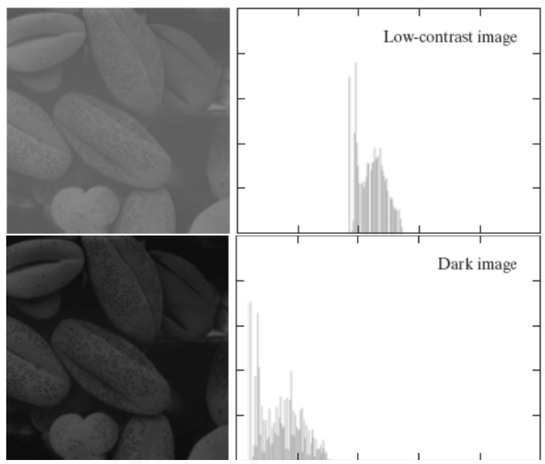
$$F_Y(y) = \begin{cases} 0, & y < 0 \\ y, & 0 \leq y < 1 \\ 1, & y \leq 1 \end{cases}$$

这表明 Y 服从均匀分布。 这个方法可以用来做图像的灰度变换。

灰度变换的例子

变换前

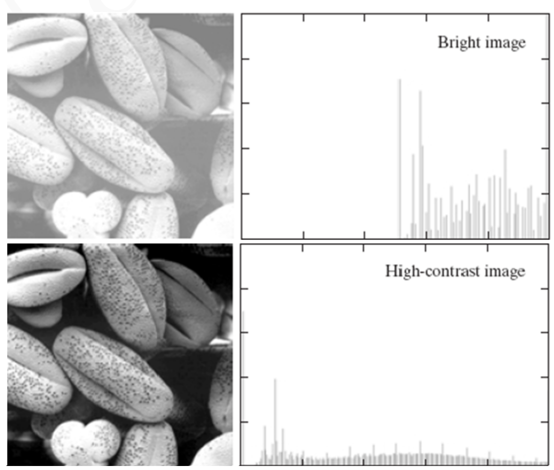
变换前



灰度变换的例子

变换前

$\boxed{\text{变换后}}$



典型问题

例： 二维随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=Ae^{-2x^2+2xy-y^2},(x,y)\in R^2$$

求常数 A 及概率 $P_{Y|X}\{Y<0|X=1\}$?

分析：由归一性可得 A，由条件概率密度可算条件概率 解：由归一性 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$，得

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

解：由条件概率密度概念可知

$$P_{Y|X}\{Y<0|X=1\}=\int_{-\infty}^{0}f_{Y|X}(y|1)dy$$

而

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\frac{1}{\pi}e^{-x^2}\int_{-\infty}^{+\infty}e^{-(y-x)^2}dy=\frac{1}{\sqrt{\pi}}e^{-x^2}$$

从而

$$f_{Y|X}(y|1)=\frac{f(1,y)}{f_X(1)}=\frac{1}{\sqrt{\pi}}e^{-(y-1)^2}$$

所以

$$P_{Y|X}\{Y<0|X=1\}=\int_{-\infty}^{0}\frac{1}{\sqrt{\pi}}e^{-(y-1)^2}dy=\Phi(-\sqrt{2})$$

典型问题

例：　二维随机变量 (X, Y) 的联合概率密度为

$$f(x,y)=\begin{cases}3x, & 0\le x\le 1, 0<y<x\\ 0 & \text{其它}\end{cases}$$

求概率 $P\{Y<\frac{1}{8}|X<\frac{1}{4}\}$?

请先思考：这个条件概率该用什么方法算？ 接下来 $P\{Y<\frac{1}{8}, X<\frac{1}{4}\}$ 怎么算？ 接下来 $P\{X<\frac{1}{4}\}$ 怎么算？

典型例题

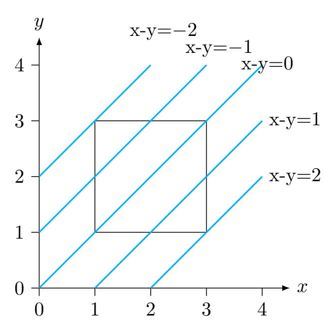
例：　设二维随机变量 (X, Y) 在矩形

$$G=\{(x,y)|1\le x\le 3, 1\le y\le 3\}$$

服从均匀分布，求 $Z=X-Y$ 的概率密度。

分析：

- 先求分布函数，再对其求导
- 画图像，讨论 $\{x-y<z\}$ 在平面上对应区域
- 一定要把关键的 $x-y=z$ 的图像画出来



解：(1)Z 的取值范围为 $[-2,2]$，故

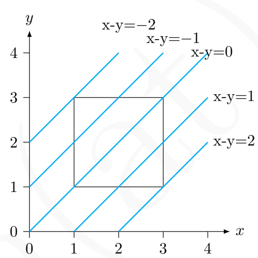
$$\text{当 } z < -2 \text{ 时，} F(z) = P\{Z \leq z\} = 0 \text{ 当 } y \geq 2 \text{ 时，} F(z) = P\{Z \leq z\} = 1$$

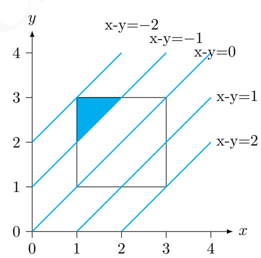
当 $-2 \leq z < 0$ 时

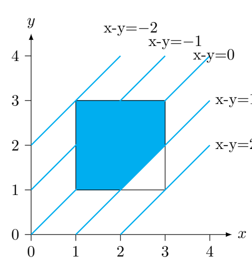
$$\begin{aligned} F(z) =& P\{Z \leq z\} = P\{Z \leq -2\} + P\{-2 < Z \leq z\} \\ =& P\{-2 < Z \leq z\} \\ =& \frac{(2+z)^2/2}{4} \end{aligned}$$

当 $0 \leq z < 2$ 时

$$\begin{aligned} F(z) =& P\{Z \leq z\} = P\{Z \leq -2\} + P\{-2 < Z \leq z\} \\ =& P\{-2 < Z \leq z\} \\ =& 1 - \frac{(2-z)^2/2}{4} \end{aligned}$$







所以 Z 的分布函数为:

$$F(z)=\begin{cases}0, & z<-2\\ \dfrac{(2+z)^2}{8}, & -2\le z<0\\ 1-\dfrac{(2-z)^2}{8}, & 0\le z<2\\ 1, & z\ge 2\end{cases}$$

所以 Z 的密度函数为:

$$f(z)=F^{'}(z)=\begin{cases}\dfrac{2+z}{4}, & -2\le z<0\\ \dfrac{2-z}{4}, & 0\le z<2\\ 0, & z<-2\text{或}z\ge 2\end{cases}$$

典型例题

例: 设随机变量 X 的概率密度为 $f(x)=\begin{cases}2^{-x}ln2, & x>0\\ 0, & x\le 0\end{cases}$, 对 X 进行独立重复观测，直到第二个大于 3 的观测值出现时停止，记 Y 为观测次数.

a 求 Y 的概率分布;

b 求数学期望 EY。

分析:

- (1) 显然与贝努利试验有关!
- 实际上就是某个事件直到第二次出现时所进行的实验次数的概率分布(不必要纠结于大于几! 抑或小于几! 或其它!)
- (2) 的关键在于 (1)，计算也很关键

解: (1) 首先: $p=P\{X>3\}=\int_3^{+\infty}f(x)dx=\int_3^{+\infty}2^{-x}ln2dx=\frac{1}{8}$ 其次，由题意可知 Y 可以取的值为: $2,3,\cdots,k,\cdots$ 事件 $\{Y=k\}$ 意味着在前

$k-1$ 次观测中，恰好有一次观测值大于 3(位置不确定)，而第 k 次观测值也大于 3 所以，Y 的分布律为 $P\{Y=k\}=(k-1)(1-p)^{k-2}p^2, k=2,3,\cdots,k,\cdots$
(2) 由期望定义

$$
\begin{aligned}
E(Y) =& \sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2}p^2 \\
=& p^2 \left(\sum_{k=0}^{\infty} x^k\right)'' \Big|_{x=1-p} \\
=& p^2 \left(\frac{1}{1-x}\right)'' \Big|_{x=1-p} = \frac{2}{p} = 16
\end{aligned}
$$

典型例题

例： 设二维随机变量 (X,Y) 在区域$D=\{(x,y)|0<x<1,x^2<y<\sqrt{x}\}$ 上服从均匀分布，令

$$
U=\begin{cases} 1, & X \le Y \\ 0, & X > Y \end{cases}
$$

a 写出 (X,Y) 的概率密度；

b 问 U 与 X 是否相互独立，并说明理由；

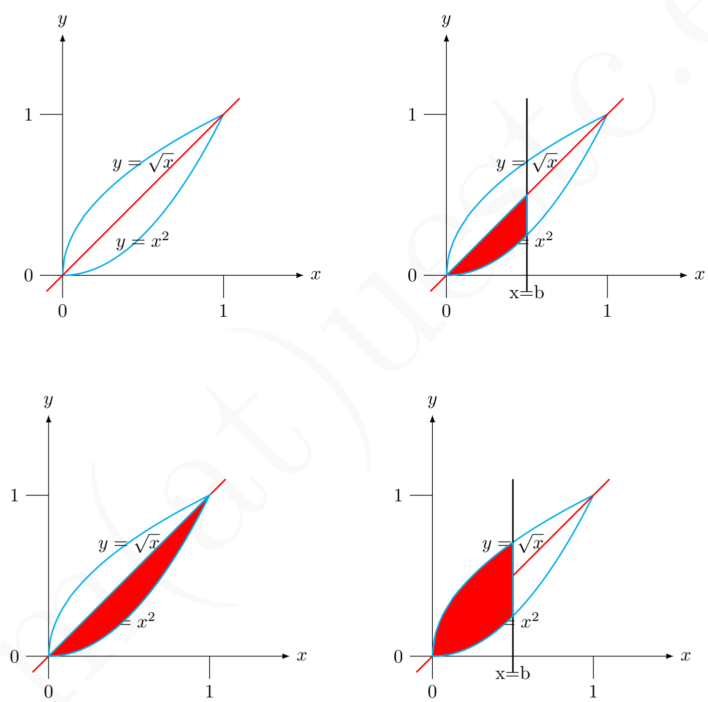
c 求 $Z=U+X$ 的分布函数 $F(z)$.

分析:

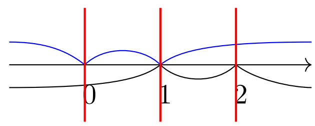
- (1) 计算出 $S(D)$，由均匀分布定义容易给出！
- 两个随机变量独立性怎么判定？
- 类似的问题前面解决过！

解: (a),$S(D)=\iint_D d\sigma=1/3$ 所以:

$$f(x,y)=\begin{cases}1/S(D), & (x,y)\in D\\ 0, & \text{其它}\end{cases}$$
$$=\begin{cases}3, & (x,y)\in D\\ 0, & \text{其它}\end{cases}$$



(b), 设 $0<a<1, 0<b<1$, 则 $P\{U\le a, X\le b\}=P\{X>Y, X\le b\}=\frac{3}{2}b^2-b^3$ $P\{U\le a\}=P\{X>Y\}=\frac{1}{2}, P\{X\le b\}=2b^{\frac{3}{2}}-b^3$ $P\{U\le a, X\le b\}\ne P\{U\le a\}P\{X\le b\}$ $\therefore U$ 与 X 不独立



解：(c),

$$
\begin{aligned}
F(z) =& P\{U+X\le z\}\\
=& P\{U+X\le z, U=0\}+P\{U+X\le z, U=1\}\\
=& \begin{cases} 0, & z<0\\ \dfrac{3}{2}z^2-z^3, & 0\le z<1\\ \frac{1}{2}, & z\ge 1\end{cases} + \begin{cases} 0, & z<1\\ 2(z-1)^{\frac{3}{2}}-\dfrac{3}{2}(z-1)^2, & 1\le z<2\\ \frac{1}{2}, & z\ge 2\end{cases}\\
=& \begin{cases} 0, & z<1\\ \frac{3}{2}z^2-z^3, & 0\le z<1\\ \frac{1}{2}+2(z-1)^{\frac{3}{2}}-\frac{3}{2}(z-1)^2, & 1\le z<2\\ 1, & z\ge 2\end{cases}
\end{aligned}
$$

例： 设随机变量 X 的概率密度为

$$
f(x)=\begin{cases} 0.5, & -1<x<0\\ 0.25, & 0\le x<2\\ 0, & \text{其他}\end{cases}
$$

令 $Y=X^2, F(x,y)$ 为二维随机变量 (X,Y) 的分布函数，求

a 求 Y 的概率密度 $f_Y(y)$；

b 求 $cov(X,Y)$;

c 求 $F(-0.5,4)$.

分析：

- (a) 先求分布函数，再对其求导 (画图像，对 Y 的值域进行分类)；
- (b) 都是跟 X 有关，先简化，再计算！

- (c) 即计算 $P\{X \leq -0.5, Y < 4\}$，但它只与 X 有关，不必求联合分布函数

解：(1)Y 的取值范围为 $[0,4]$，故

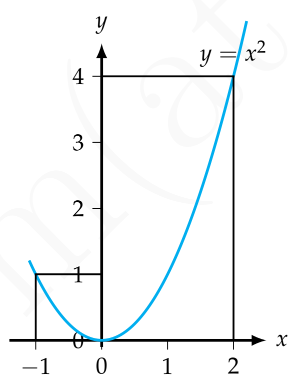
当 $y<0$ 时，$F(y)=P\{Y\leq y\}=0$ 当 $y\geq 4$ 时，$F(y)=P\{Y\leq y\}=1$

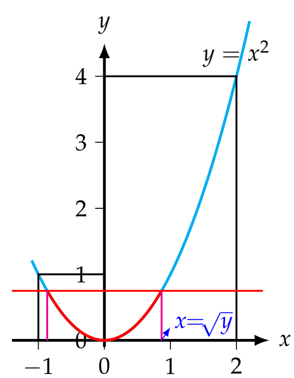
当 $0\leq y<1$ 时

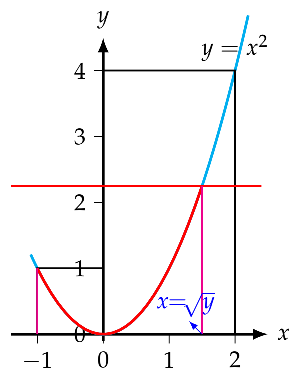
$$\begin{aligned} F(y) =& P\{Y\leq y\} = P\{-\sqrt{y}\leq X\leq \sqrt{y}\} \\ =& \int_{-\sqrt{y}}^{0} \frac{1}{2}dx + \int_{0}^{\sqrt{y}} \frac{1}{4}dx = \frac{3\sqrt{y}}{4} \end{aligned}$$

当 $1\leq y<4$ 时

$$\begin{aligned} F(y) =& P\{Y\leq y\} = P\{-1\leq X\leq \sqrt{y}\} \\ =& \int_{-1}^{0} \frac{1}{2}dx + \int_{0}^{\sqrt{y}} \frac{1}{4}dx = \frac{1}{2} + \frac{1}{4}\sqrt{y} \end{aligned}$$







综上所述:Y 的分布函数为

$$F(y)=\begin{cases}0, & y<0\\ \frac{3}{4}\sqrt{y}, & 0\le y<1\\ \frac{1}{2}+\frac{1}{4}\sqrt{y}, & 1\le z<4\\ 1, & y\ge 4\end{cases}$$

所以:Y 的概率密度为

$$f_Y(y)=F^{'}(y)=\begin{cases}\frac{3}{8\sqrt{y}}, & 0\le y<1\\ \frac{1}{8\sqrt{y}}, & 1\le z<4\\ 0, & \text{其它}\end{cases}$$

(2) 由协方差的计算公式:

$$cov(X,Y)=cov(X,X^2)=E(X^3)-E(X)E(X^2)$$

又由一元函数期望的计算公式:

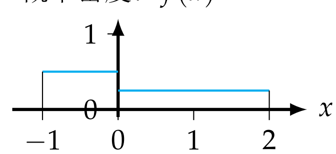
$$E(X)=\int_{-\infty}^{\infty}xf_X(x)dx=\int_{-1}^{0}\frac{x}{2}dx+\int_{0}^{2}\frac{x}{4}dx=\frac{1}{4}$$
$$E(X^2)=\int_{-\infty}^{\infty}x^2f_X(x)dx=\int_{-1}^{0}\frac{x^2}{2}dx+\int_{0}^{2}\frac{x^2}{4}dx=\frac{5}{6}$$
$$E(X^3)=\int_{-\infty}^{\infty}x^3f_X(x)dx=\int_{-1}^{0}\frac{x^3}{2}dx+\int_{0}^{2}\frac{x^3}{4}dx=\frac{7}{8}$$

所以:$cov(X,Y)=\frac{7}{8}-\frac{1}{4}\times\frac{5}{6}=\frac{2}{3}$

(3) 根据二维随机变量分布函数的定义:

$$\begin{aligned}F(-\frac{1}{2},4)=&P\{X\le-\frac{1}{2},Y\le4\}\\=&P\{X\le-\frac{1}{2},X^2\le4\}\\=&P\{-2\le X\le-\frac{1}{2}\}\\=&\int_{-2}^{-\frac{1}{2}}f_X(x)dx\\=&\int_{-1}^{-\frac{1}{2}}\frac{1}{2}dx\\=&\frac{1}{4}\end{aligned}$$

概率密度：$f(x)$



1. 设随机变量 X,Y 相互独立,且 EX,EY 存在,记 $U=max\{X,Y\},V=min\{X,Y\}$，则 $E(UV)=($ $)$

   (1) $EU\cdot EV$, (2) $EX\cdot EY$, (3) $EU\cdot EY$, (4) $EX\cdot EV$

2. 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______。

3. 设连续型随机变量 X_1, X_2 相互独立，且方差均存在，X_1, X_2 的概率密度分别为 $f_1(x), f_2(x)$，随机变量 Y_1 的概率密度为$f_{Y_1}(y) = \dfrac{1}{2}(f_1(y) + f_2(y))$，随机变量 $Y_2 = \dfrac{1}{2}(X_1 + X_2)$ 则 (　)

   (1) $EY_1 > EY_2, DY_1 > DY_2$　(2) $EY_1 = EY_2, DY_1 = DY_2$
   (3) $EY_1 = EY_2, DY_1 < DY_2$　(4) $EY_1 = EY_2, DY_1 > DY_2$

1. 显然 $E(Y_2) = \dfrac{1}{2}(E(X_1) + E(X_2)), D(Y_2) = \dfrac{1}{4}(D(X_1) + D(X_2))$ 关键是 $E(Y_1), D(Y_1)$

$$E(Y_1) = \int_{-\infty}^{\infty} \frac{y}{2}(f_1(y) + f_2(y))dy = \frac{1}{2}(E(X_1) + E(X_2))$$
$$E(Y_1^2) = \int_{-\infty}^{\infty} \frac{y^2}{2}(f_1(y) + f_2(y))dy = \frac{1}{2}(E(X_1^2) + E(X_2^2))$$
$$\begin{aligned}\therefore D(Y_1) =& \frac{1}{2}(E(X_1^2) + E(X_2^2)) - [\frac{1}{2}(E(X_1) + E(X_2))]^2 \\ =& \frac{1}{4}(D(X_1) + D(X_2)) + \frac{1}{4}(E(X_1^2) + E(X_2^2) - 2E(X_1)E(X_2))\end{aligned}$$
$$E[(X_1 - X_2)^2] = E(X_1^2) + E(X_2^2) - 2E(X_1)E(X_2) \geq 0$$

例 随机试验 E 有两两不相容的结果 A_1, A_2, A_3，且三种结果发生的概率均为 $\frac{1}{3}$，将试验 E 独立重复做 2 次，X 表示 2 次试验中 A_1 发生的次数，Y 表示 2 次试验中 A_2 发生的次数，则 X 与 Y 的相关系数为

解法 (1): X, Y 的联合分布律为

X \ Y	0	1	2	$p_{i\cdot}$
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
$p_{\cdot j}$	4/9	4/9	1/9	1

解法 *(2)*: $X \sim B(2,1/3), Y \sim B(2,1/3)$ $EX = EY = 2/3, DX = DY = 4/9$ $E(XY) = 1\times 1\times P\{X=1, Y=1\} = 2/9$, $\rho = -1/2$

解法 *(3)*:

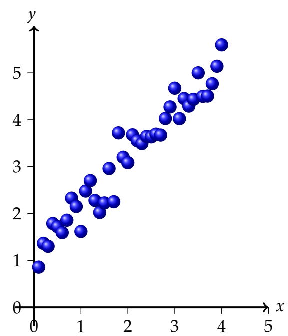
$$\begin{aligned}
\frac{cov(X,Y)}{\sqrt{DX}\sqrt{DY}} =& \frac{cov(X, 2-X-Z)}{\sqrt{DX}\sqrt{DY}} \\
=& \frac{-cov(X,X) - cov(X,Z)}{\sqrt{DX}\sqrt{DY}} \\
=& \frac{-cov(X,X) - cov(X,Y)}{\sqrt{DX}\sqrt{DY}} \\
\therefore 2\frac{cov(X,Y)}{\sqrt{DX}\sqrt{DY}} =& \frac{-cov(X,X)}{\sqrt{DX}\sqrt{DX}} = -1 \\
\therefore \rho = -1/2
\end{aligned}$$

例: 设随机变量 X 的分布函数 $F(x) = 0.5\Phi(x) + 0.5\Phi(\frac{x-4}{2})$，其中 $\Phi(x)$ 为标准正态分布函数，则 $E(X)$?

$f(x) = F^{'}(x) = 0.5\varphi(x) + 0.25\varphi(\frac{x-4}{2})$ 所以:

$$\begin{aligned}
EX =& \int_{-\infty}^{+\infty} xf(x)dx \\
=& \int_{-\infty}^{+\infty} 0.5x\varphi(x)dx + \int_{-\infty}^{+\infty} 0.25x\varphi(\frac{x-4}{2})dx \\
=& \int_{-\infty}^{+\infty} 0.25x\varphi(\frac{x-4}{2})dx \\
=& \int_{-\infty}^{+\infty} 0.5(2t+4)\varphi(t)dt \\
=& 2 \\
=& \int_{-\infty}^{+\infty} 0.5x\varphi(x)dx + \int_{-\infty}^{+\infty} 0.5\frac{x}{\sqrt{2\pi}2}e^{\frac{(x-4)^2}{2\times 4}}dx = 2
\end{aligned}$$

一个概念的讨论：为什么相关系数 ρ 反应的是线性关系？



考虑: $\alpha X+\beta$ 近似 Y 产生的均方误差:$MSE=E(Y-\alpha X-\beta)^2$。

均方误差 $(mean-square-error, MSE)$ 是反映估计量与被估计量之间差异程度的一种度量

$$MSE=E(Y^2+\alpha^2X^2+\beta^2+2\alpha\beta X-2\alpha XY-2\beta Y)$$

为了让 MSE 取到最小值，需要解驻点方程组

$$\begin{cases} \frac{\partial MSE}{\partial \alpha} = & 2E(X^2)\alpha+2E(X)\beta-2E(XY)=0 \\ \frac{\partial MSE}{\partial \beta} = & 2E(X)\alpha+2\beta-2E(Y)=0 \end{cases}$$

得到

$$\begin{cases} \alpha = & \dfrac{cov(X,Y)}{D(X)} \\ \beta = & E(Y)-\dfrac{cov(X,Y)}{D(X)}E(X) \end{cases}$$

再次代回 MSE 化简整理可得

$$\underset{min}{MSE}=(1-\rho_{\mathrm{XY}}^2)D(Y)$$

- 这意味着 $|\rho|$ 越大，用 $\alpha X+\beta$ 近似 Y 产生的最小均方误差:$MSE=E(Y-\alpha X-\beta)^2$ 就越小;
- 当 $|\rho|=1$ 时，均方误差为 0，表明 Y 与 X 有良好的线性关系。
- 当 $\rho=0$ 时，均方误差等于 $D(Y)$，这意味着用 X 的一个线性表达去近似 Y 产生的均方误差完全与 X 无关，仅仅取决于随机变量 Y 自身的波动。可以认为 Y 与 X 没有线性相关。

例: 试证：若 X,Y 都是只取两个值的随机变量，则 X,Y 不相关时，X,Y 一定相互独立。

证明：设 (X,Y) 的联合分布律为

Y \ X	c	d	$p_{i\cdot}$
a	p_{11}	p_1-p_{11}	p_1
b	p_2-p_{11}	$1-p_1-p_2+p_{11}$	$1-p_1$
$p_{\cdot j}$	p_2	$1-p_2$	1

若 X,Y 不相关，则有 $EXY=EX\cdot EY$，即:

$$\begin{aligned}&acp_{11}+ad(p_1-p_{11})+bc(p_2-p_{11})+bd(1-p_1-p_1+p_{11})\\=&(ap_1+b(1-p_1))(cp_2+d(1-p_2))\\=&acp_1p_2+adp_1(1-p_2)+bc(1-p_1)p_2+bd(1-p_1)(1-p_2)\end{aligned}$$

展开并消去左侧和右侧相等的项得到:

$$\begin{aligned}&acp_{11}-adp_{11}-bcp_{11}+bdp_{11}\\=&acp_1p_2-adp_1p_2-bcp_1p_2+bdp_1p_2\end{aligned}$$

整理得到:

$$(ac-ad-bc+bd)p_{11}=(ac-ad-bc+bd)p_1p_2$$

即得到: $p_{11} = p_1 p_2$ ，这意味着:

$$P\{X=a, Y=c\} = P\{X=a\}P\{Y=c\}$$

类似的方法可以证明:

$$P\{X=a, Y=d\} = P\{X=a\}P\{Y=d\}$$
$$P\{X=b, Y=c\} = P\{X=b\}P\{Y=c\}$$
$$P\{X=b, Y=d\} = P\{X=b\}P\{Y=d\}$$

这就证明了 X, Y 是相互独立的。

例:将 n 封不同的信的 n 张信笺与 n 个信封进行随机匹配，记 N 为匹配成对数，求 N 的望?

分析:

- 分布律虽然是可以求出来的，但期望不好算（错配!）
- 可考虑用期望性质算

解:设 X_i 为第 $i(i=1,2,\cdots,n)$ 个信封与信笺匹配成对数，则

$$N = \sum_{i=1}^{n} X_i$$

根据抽签公平性: $X_i(i=1,2,\cdots,n)$ 同分布且都服从 0−1 分布，分布律为:

X_i	0	1
p_i	1-1/n	1/n

$$E(X_i) = 1/n \quad (i=1,2,\cdots,n)$$
$$EN = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = 1$$

柯西-施瓦兹（Cauchy-Schwarz）不等式:

对两个随机变量 X, Y，若 $E(X^2), E(Y^2)$ 存在，证明

$$E^2(XY) \le E(X^2)E(Y^2).$$

分析: 柯西-施瓦兹不等式常用构造一元二次实函数证明。 证明: 考虑 t 的

一元二次函数

$$\begin{aligned} f(t) =& E[(X+tY)^2] \\ =& E(X^2) + 2tE(XY) + t^2E(Y^2) \geq 0 \\ \text{故} \quad \Delta =& 4E^2(XY) - 4E(X^2)E(Y^2) \leq 0 \end{aligned}$$

即

$$E^2(XY) \leq E(X^2)E(Y^2).$$

总结：

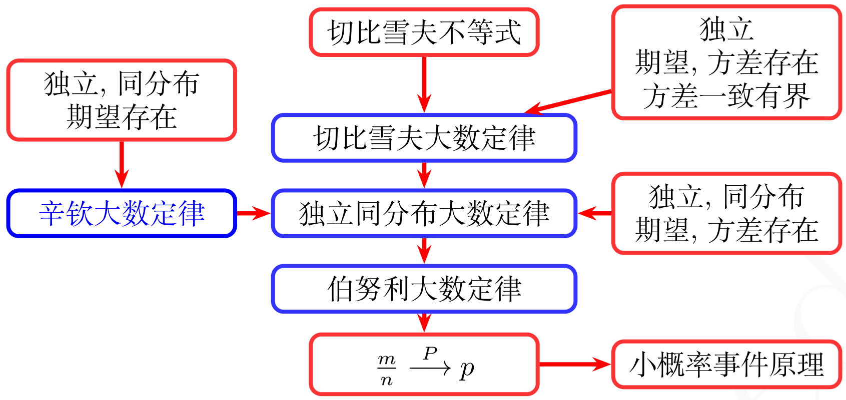
- 分布函数、概率密度和分布律刻画了随机变量的概率分布，一般来讲是计算概率的关键；
- 条件分布和独立性讨论了两个随机变量之间的关系，讨论关系一般就是指这两种情形；
- 遇到复杂问题时最好分情况讨论，当加一个”限制”在问题上时，问题往往会被简化；
- 即便抽象能力很强，使用图形辅助解决问题往往让问题变得形象，因此更容易理解。
- 最好熟悉基本结论，基本方法，问题不管怎么变化，基本的东西不会变。

1.4 大数定律与中心极限定理

主要内容

目录

大数定律



切比雪夫（Chebyshev）不等式

设随机变量 X 的数学期望 $E(X)$ 和方差 $D(X)$ 都存在, 则对于任意的 $\varepsilon > 0$，有

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

或者

$$P\left\{\left|X - E(X)\right| < \varepsilon\right\} > 1 - \frac{D(X)}{\varepsilon^2}$$

a 估计概率

b 其估计是很粗略的

c 理论推导

例：$X \sim N(\mu, \sigma^2)$，令 $\varepsilon = 3\sigma$，则有：$P\left\{\left|X - \mu\right| < 3\sigma\right\} > 1 - \frac{\sigma^2}{(3\sigma)^2} = \frac{8}{9} \approx 0.8889$

大数定律在信号检测中的应用

若发送端发出的信号为确定信号

$$s_t = a\cos(\omega t + \varphi)$$

实际中的噪声假设为加性高斯白噪声

$$X_t \sim N(0, \sigma^2)$$

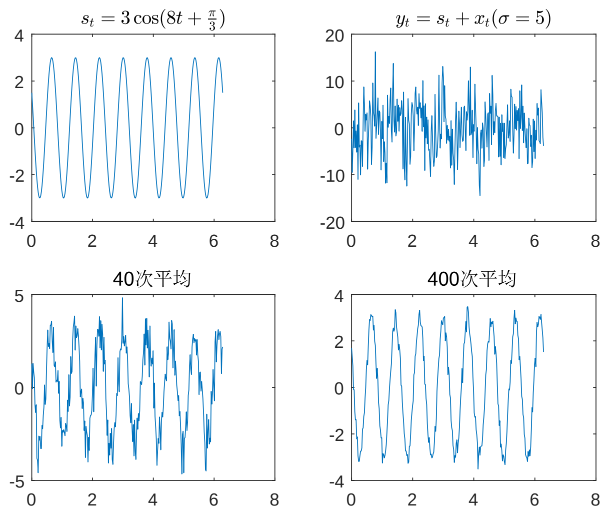
则接收端收到的信号为

$$Y_t = s_t + X_t \sim N(s_t, \sigma^2)$$

但若进行 n 次收发，将每次收到的信号叠加求算术平均

$$\frac{1}{n}\sum_{i=1}^{n} Y_t^{(i)} = s_t + \frac{1}{n}\sum_{i=1}^{n} X_t^{(i)} \sim N(s_t, \frac{\sigma^2}{n})$$

仿真实例



中心极限定理的意义

若随机变量序列 $\{X_k\}$，$k = 1, 2, \cdots$ 服从中心极限定理，则

$$\lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k - \sum\limits_{k=1}^{n} E(X_k)}{\sqrt{\sum\limits_{k=1}^{n} D(X_k)}} \le x\right\} = \Phi(x)$$

即 $\sum\limits_{k=1}^{n} X_k$ 的标准化随机变量的极限分布为标准正态分布。

正态分布常见的原因

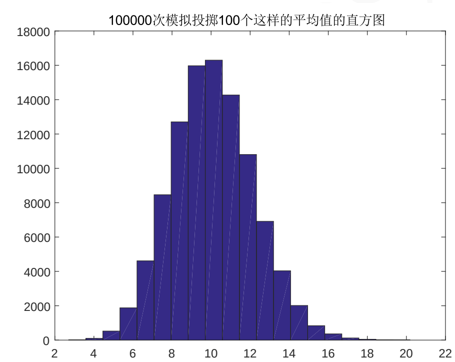
统计学里面，正态分布最常见。男女身高、寿命、血压、考试成绩、测量误差等等，都属于正态分布。

一般来讲，测量某个物理量的时候，不可避免会有误差产生，产生误差的原因是多方面的，有测量仪器引起，也有观察者者引起，还有测试条件的随机扰动引起，等等。所有这些因素，每个都可以产生一定的误差，我们观察到的实际上是一个“总的误差”，也就是说，这个误差是一个大量相互独立的随机误差的累加效应综合而成的随机变量，根据中心极限定理，这种误差通常服从正态分布。

非同分布的例子

假设现在有 100 个奇形怪状的均匀六面骰子，并且第 i 个六面骰子上的点数分别为 1,1,2,3,3,i.

记 X_i 为第 i 个骰子掷出的点数，令 $S = \frac{1}{100}\sum_{i=1}^{100} X_i$ 对 S 做模拟



独立同分布中心极限定理的应用

1 求随机变量之和 $S_n = \sum\limits_{k=1}^{n} X_k$ 取值的概率。例如产品检验问题

2 已知 $S_n = \sum\limits_{k=1}^{n} X_k$ 取值的概率，反求 n。例如产品测重问题

3 数理统计中大样本推断的理论基础。

由棣莫弗-拉普拉斯中心极限定理，设随机变量 $X \sim B(n,p)$ 则对于足够大的 $n(np \geq 5, n(1-p) \geq 5$, 茆诗松, 统计手册)，有

$$
\begin{aligned}
&P\{m_1 < X \leq m_2\} \\
=&P\left\{\frac{m_1-np}{\sqrt{np(1-p)}} < \frac{X-np}{\sqrt{np(1-p)}} \leq \frac{m_2-np}{\sqrt{np(1-p)}}\right\} \\
\approx&\Phi\left(\frac{m_2-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{m_1-np}{\sqrt{np(1-p)}}\right)
\end{aligned}
$$

棣莫佛—拉普拉斯中心极限定理的应用

1 计算在大 n 情况下二项分布概率的近似值。

2 已知在大 n 情况下二项分布在某范围内取值的概率，求该范围。

3 近似计算与用频率估计概率的有关问题。

中心极限定理在保险中的应用

例:

某保险公司有 10000 个同阶层的人参加人寿保险，每人每年付 600 元保险费，在一年内一个人死亡的概率为 0.006，死亡时，其家属可向保险公司领得 5 万元赔偿金。试问：保险公司亏本的概率有多大? 保险公司每年利润大于 250 万的概率是多少?

分析：设 X_i 表示保险公司支付给第户的赔偿金（单位：万元），则在实际情况下，可近似认为 X_i 独立同分布，服从两点分布

X_i	0	5
p_k	0.994	0.006

保险公司的全部赔偿金为 $\sum_{i=1}^{10000} X_i$

2 数理统计学

主要内容

目录

数理统计思想、方法与模式

- 首先，研究对象全体的某数量指标 X 被看成是具有某一分布的随机变量；
- 其次，通过随机抽样的方式得到总体的一个代表性的样本，所以样本是一组随机变量；
  - 本课程重点考虑 X 的分布类型已知时，如何通过代表性的个体所含的信息去推断这些未知参数的信息；
- 最后，利用得到的样本的信息去对总体的未知信息做出推断。
  - 由于样本的随机性，需要利用概率论度量推断中的不确定性；
  - 对样本所含信息进行整理就得到统计量，统计量是随机变量。

统计推断的一般模式：样本 X 的分布或概率密度 $f(x, \theta)$ 依赖于未知参数 θ，只知道 θ 属于某一集合 Θ，但不知它取 Θ 中的何值。

统计推断的任务，就是依据所抽得的样本 X 去作出某种有关 θ 之值的论断，例如对 θ 的值作一估计，或判断它是否落在 Θ 的某个指定子集 Θ_1 之内等.

2.1 抽样分布定理

主要内容

目录

四大常见统计分布及其构造

假设涉及到的随机变量相互独立

1. 标准正态分布

$$X \sim N(\mu, \sigma^2) \Longrightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$$

2. χ^2 分布

$$X_i \sim N(0,1), i=1,2,\cdots,n \Longrightarrow \chi^2 = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$$

3. t 分布

$$X \sim N(0,1), Y \sim \chi^2(n) \Longrightarrow T = \frac{X}{\sqrt{Y/n}} \sim t(n)$$

4. F 分布

$$X \sim \chi^2(n_1), Y \sim \chi^2(n_2) \Longrightarrow F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$$

典型例题

1. 设总体 $\xi \sim N(0,\sigma^2), \xi_1, \xi_2, \cdots, \xi_{15}$ 为简单随机样本，试判断统计量

$$T = \frac{\sum\limits_{i=1}^{10} (-1)^i \xi_i}{\sqrt{2}\sqrt{\sum\limits_{i=11}^{15} \xi_i^2}}$$

服从什么分布，并给出理由。

2. 设总体 $X \sim N(0,1), Y \sim N(0,4)$ 相互独立 X_1, X_2, X_3, X_4 和 $Y_1, Y_2, Y_3, \cdots, Y_9$ 分别为来自总体 X 和 Y 的样本，确定统计量

$$Z = \frac{3(X_1 + X_2 + X_3 + X_4)}{\sqrt{\sum\limits_{i=1}^{9} Y_i^2}}$$

的分布

典型例题

3. 设 $X_1, X_2, \cdots, X_{n+m}$ 是来自总体 $X(X \sim N(0,\sigma^2))$ 的容量为 $n+m$ 的简单随机样本。

$$Y_1 = \frac{\sqrt{m}\sum\limits_{i=1}^{n} X_i}{\sqrt{n}\sqrt{\sum\limits_{i=n+1}^{n+m} X_i^2}}, \quad Y_2 = \frac{m\sum\limits_{i=1}^{n} X_i^2}{n\sum\limits_{i=n+1}^{n+m} X_i^2}$$

问 Y_1, Y_2 服从什么分布？说明理由

4. 设总体 $X \sim N(0,\sigma^2), X_1, X_2, X_3, X_4$ 为其样本，试确定 $\dfrac{(X_1-X_2)^2}{(X_3+X_4)^2}$ 的分布.

2.2 参数估计

典型例题

5. 假设抛硬币的结果记为一个整体 X，对这个总体进行抽样，得到一个样本 $X_1, X_2, \cdots, X_n$，这里 X_i 表示第 i 次抛硬币时正面出现的次数。记 $\overline{X}$ 为样本均值，用切比雪夫不等式估计至少抛多少次 $\overline{X}$ 落入区间 $[0.4, 0.6]$ 的概率为 0.9? 改用中心极限定理来计算这个问题，需抛的次数又是多少?

6. 设 X_1, X_2, X_3, X_4, X_5 为来自正态总体 $X \sim N(0,1)$ 的样本，确定统计量
$$\frac{3(X_1-X_2)^2}{2(X_3^2+X_4^2+X_5^2)}$$
的分布。

主要内容

目录

参数估计

学习目标

1. 掌握点估计的两种估计方法:
   - 矩估计;
   - 极大似然估计.

   估计量的记号中，^ 表示 $\hat{\theta}$ 仅仅是 θ 的估计，而非真值。

2. 掌握判别优良估计量的三个准则:
   - 无偏性;

- 有效性;
- 相合性.

3. 掌握单个正态总体的各种置信区间

典型例题

1. 设 $X_1, X_2, \cdots, X_n$ 为来自正态总体 $X \sim N(0, \sigma^2)$ 的样本, 求 σ 的 (1) 矩估计量;(2)极大似然估计量。
2. 已知幼儿身高在正常情况下服从正态分布，现从某一幼儿园 5 岁至 6 岁的幼儿中随机抽查了 9 人, 其身高 (以 cm 为单位) 分别为 115, 120, 131, 115, 109, 115, 115, 105, 110, 试在置信度 95% 条件下，求 5 岁至 6 岁的幼儿身高方差的置信区间。

典型例题

例: 设测量误差服从零均值的正态分布，进行 4 次独立测量，各次误差为 0.1, -0.08, 0.04, -0.07。求测量误差方差的极大似然估计值。解：设测量误差 $X \sim N(0, \sigma^2), x_1 = 0.1, x_2 = -0.08, x_3 = 0.04, x_4 = -0.07$ 为样本观测值。似然函数为:

$$L(\sigma) = \prod_{i=1}^{4} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} = \frac{1}{4\pi^2\sigma^4} e^{-\frac{0.01145}{\sigma^2}}$$

$$\ln L(\sigma) = -2\ln 2\pi - 4\ln\sigma - \frac{0.01145}{\sigma^2}$$

令

$$\frac{d\ln L(\sigma)}{d\sigma} = -\frac{4}{\sigma} + \frac{0.0229}{\sigma^3} = 0$$

得极大似然估计值 $\hat{\sigma}^2 = 0.005725$ （对 σ 求导和对 σ^2 求导得到的估计值结果一样!）

典型例题

5. 已知总体 X 服从几何分布

$$p\{X=x\}=(1-p)^{x-1}p, x=1,2,\cdots$$

其中，p 为未知参数，$0<p<1$。设 $X_1,X_2,\cdots,X_n$ 为总体 X 的一个样本，求参数 p 的极大似然估计。

6. 对样本容量为 n 的样本，求密度函数

$$f(x;\alpha)=\begin{cases}\dfrac{2}{\alpha^2}(\alpha-x), & 0<x<\alpha;\\ 0, & \text{others.}\end{cases}$$

中参数 α 的矩估计量。

典型例题

例: 设 $X_1,X_2,\cdots$ 独立同分布，服从 $[0,\theta]$ 上的均匀分布，定义

$$Y=\max\{X_1,\cdots,X_n\}$$

则 Y 是 θ 的相合估计量。

证明：对任意的 $0<\epsilon<\theta$

$$\begin{aligned}P\{|Y-\theta|\leq\epsilon\}=&1-P\{|Y-\theta|\geq\epsilon\}\\=&1-P\{|X_1-\theta|\geq\epsilon,\cdots,|X_n-\theta|\geq\epsilon\}\\=&1-P\{|X_1-\theta|\geq\epsilon\}\cdots P\{|X_n-\theta|\geq\epsilon\}\\=&1-(1-\epsilon)^n\\\to&1\quad(n\to\infty)\end{aligned}$$

2.3 假设检验

主要内容

目录

第八章假设检验

学习目标

1. **理解假设检验的目的和基本检验思路**
2. **熟练掌握单个正态总体的四种双侧假设检验方法**
3. **了解单侧假设检验方法**

假设检验的基本步骤

1. 提出原假设：根据实际问题提出原假设 H_0 和备择假设 H_1 ;
2. 建立检验统计量：寻找一个参数的良好估计量，并在 H_0 成立的条件下，建立一个不带任何未知参数的统计量 U 作为检验统计量，确定 U 的分布 (或近似分布);
3. 确定 H_0 的拒绝域：根据对立假设 H_1 的形式确定 H_0 的拒绝域;
4. 对 H_0 作判断：根据样本值算出检验统计量的统计值 u ，判断 u 是否落在拒绝域，以确定拒绝或接受 H_0 。

1. 设一种电子元件的寿命 X（以小时计）服从正态分布，μ, σ^2 均未知。现测得 $n = 16$ 只元件的寿命，计算得：

$$x = 241.5, \bar{s} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} = 98.73$$

问：是否有理由认为元件的平均寿命为 225（小时）? $(\alpha = 0.1, u_{0.05} = 1.645, t_{0.05}(15) = 1.7531)$

2. 已知某电子元件的长度服从正态分布，且方差为 0.01。从一批次的产品中任取 10 个，测得 $s^2 = 0.012$, 则在显著性水平 $\alpha = 0.05$ 下，这批电子元件的精度是否正常?

典型例题

2、解:设元件长度 $X \sim N(\mu, \sigma^2)$，根据题意要求，需作检验:

$$H_0 : \sigma^2 = \sigma_0^2 = 0.01 \quad H_1 : \sigma^2 \neq \sigma_0^2$$

由于 μ 未知，故采用 χ^2 检验法，当 H_0 成立时:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$$

拒绝域为

$$\chi^2 > \chi^2_{\frac{\alpha}{2}}(n-1) \quad 或 \quad \chi^2 < \chi^2_{1-\frac{\alpha}{2}}(n-1)$$

由已知: $n = 10, (n-1)s^2 = 9s^2 = 0.108, \chi^2 = 10.8$ 查表得: $\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{0.025}(9) = 19.023$,$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{0.975}(9) = 2.7$ 因: $\chi^2_{0.975}(9) < \chi^2 < \chi^2_{0.025}(9)$ 故不能拒绝原假设 H_0，在显著性水平 0.05 下，可以认为车床保持了原加工精度。

2.4 一元回归分析

主要内容

目录

第九章一元回归分析

学习目标

(A) 能用相关系数检验法检验两个变量之间是否有显著的线性关系

(B) 熟练掌握求一元线性回归中截距和斜率的最小二乘估计方法

(C) 熟练掌握求一元线性回归方差的估计方法

(D) 会做非线性回归的线性化处理

典型例题

1. 在考察硝酸钠的可溶程度时，对一系列不同温度观察它在 100 毫升的水中溶解的硝酸钠的重量，得到一组观测值 $(x_i, y_i), i = 1, 2, \cdots, 9$ 计

算得：

$$\sum_{i=1}^{9} x_i = 234, \sum_{i=1}^{9} y_i = 234811.3, \sum_{i=1}^{9} x_i^2 = 10144$$
$$\sum_{i=1}^{9} y_i^2 = 76218.1339, \sum_{i=1}^{9} x_i y_i = 24628.6$$

从理论上推测，温度 x_i 与溶解得硝酸钠的重量 Y_i 之间有关系式：$Y_i = a + bx_i + \epsilon_i, i = 1, 2, \cdots, 9$。式子中 $\epsilon_1, \epsilon_2, \cdots, \epsilon_9$ 相互独立，均服从正态分布 $N(0, \sigma^2)$

a. 求未知参数 a,b, 的最小二乘估计值和 σ^2 的无偏估计值；

b. 检验线性回归是否显著 $(\alpha = 0.01)$?（结果保留四位有效数字）

自由度 / α	6	7
0.01	0.834	0.798
0.005	0.870	0.836

典型例题

2. 某种金属的抗拉强度 Y 与硬度 X 存在相关关系，现测得 20 对数据 $(x_i, y_i), i = 1, 2, \cdots, 20$，算得

$$\sum_{i=1}^{20} x_i = 606, \sum_{i=1}^{20} y_i = 210.5$$

$$\sum_{i=1}^{20} x_i^2 = 23748, \sum_{i=1}^{20} y_i^2 = 2664.25, \sum_{i=1}^{20} x_i y_i = 7805.5$$

a. 用相关系数检验法检验 X 与 Y 线性关系是否显著?$(\alpha = 0.01, R_{0.01}(18) = 0.561)$；

b. 给出 Y 与 X 的经验线性回归方程。