西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY

数学与统计学院
School of Mathematics and Statistics, NPU

概率论与数理统计

知识点串讲

概率论与数理统计

随机事件

概率

随机变量

分布函数

数字特征

理论基础

大数定律
中心极限定理

概率论 ⟺ 数理统计

X $\qquad f(X_1, X_2, \cdots X_n)$

统计量

三大分布

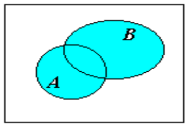
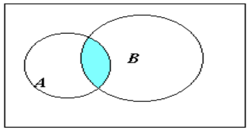
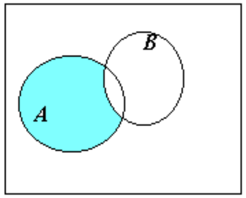
抽样分布

参数估计

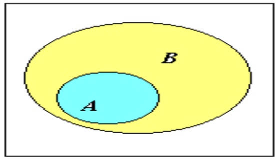
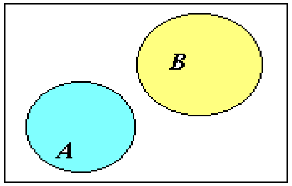
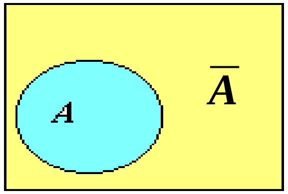
假设检验

知识模块之间的关系

1.2.1 随机事件间的运算

运算	符号	概率论	集合论	Venn图
和	$A \cup B$	事件A与B至少有一个发生	A与B的并集	
积	AB 或$A \cap B$	事件A与B同时发生	A与B的交集	
差	$A - B = A\overline{B}$	事件A发生而B不发生	A与B的差集	

1.2.2 随机事件间的关系

关系	符号	概率论	集合论	Venn图
包含	$A\subset B$	A发生则B必发生	A是B的子集	
等价	$A=B$	$A\subset B$ 且$B\subset A$	A与B相等	
互斥 (互不相容)	$AB=\varnothing$	事件A与B不能同时发生	A与B不相交	
对立 (互逆)	$\bar{A}$	A的对立事件	A的余集 ① $A\cup\bar{A}=\Omega$ ② $A\bar{A}=\varnothing$	

1.2.3 运算定律

1.交换律： (1) $A \cup B = B \cup A.$

(2) $AB = BA.$

2.结合律： (1) $(A \cup B) \cup C = A \cup (B \cup C).$

(2) $(AB)C = A(BC).$

3.分配律： (1) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$

★(2) $(AB) \cup C = (A \cup C)(B \cup C),$

$A \cup (BC) = (A \cup B)(A \cup C).$

1.2.4 对偶律(De Morgan定理)

(1) $\overline{A \cup B} = \bar{A} \cap \bar{B}.$

意义："A, B至少有一个发生"的对立事件是"A, B均不发生"。

(2) $\overline{A \cap B} = \bar{A} \cup \bar{B}.$

意义："A, B均发生"的对立事件是"A, B至少有一个不发生"。

证明

推广：
$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i}.$$

$$\overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i}.$$

1.2.5 其它一些性质

若 $A \subset B$，则 $A \cup B = B$，$AB = A$.

特别地，$A \cup \varnothing = A$, $A \cup \Omega = \Omega$

$$A\varnothing = \varnothing \qquad A\Omega = A$$

1.3.3 概率的统计定义

古典概型随机试验

1.古典概型定义

若随机试验 E 具有下列**两个**特征：

1) 有限性

样本空间Ω中，只有有限个样本点

即 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}.$

2) 等可能性

$\omega_1, \omega_2, \cdots, \omega_n$ 发生的可能性相等.

则称 E 所描述的概率模型为**古典概型**.

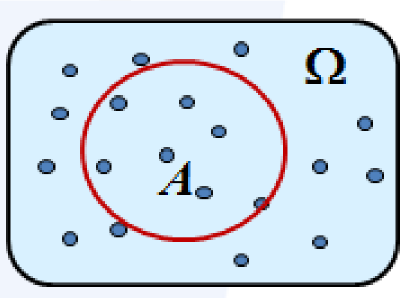
法国数学家
拉普拉斯
P.-S. Laplace
(1749-1827)

2. 古典概型中概率的计算公式

定义　　设**古典概型随机试验 E** 的样本空间Ω由n 个样本点构成 , A 为 E 的任意一个事件, 且包含 n_A个样本点 ,则事件 A 出现的概率记为:

$$P(A)=\frac{A\text{所包含样本点的个数}}{\Omega\text{所含样本点的总数}}=\frac{n_A}{n}$$

称此为**古典概型的概率** .



3. 常见的三种古典概型基本模型

(1) 摸球模型（有/无放回）

(2) 分配模型（有/无序）

(3) 随机取数模型（取后还原/不还原）

1.3.4几何概型

几何概型随机试验

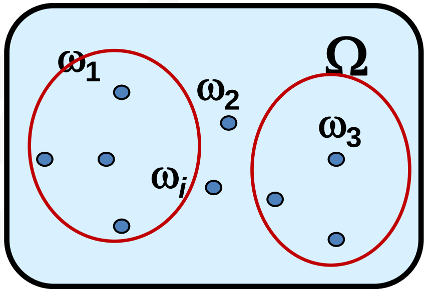
1. 定义

若试验E具有下列**两个**特征:

1)无限性: 样本空间Ω是几何空间中的一个区域,包含无穷多个样本点,每个样本点由区域Ω内的点的随机位置所确定,即 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\cdots\}.$

2)等可能性: 每个样本点落在Ω内几何度量相同的子区域内等可能的,

则称E所描述的概率模型为**几何概型**。



2.定义(几何概率的定义)

对于随机试验E，以$\mu(A)$表示事件A的几何度量，Ω为样本空间. 若 $0<\mu(\Omega)<+\infty$, 则对于任一事件A，其概率为
$$P(A)=\frac{\mu(A)}{\mu(\Omega)}.$$

一维

二维

三维

$$P(A)=\frac{A\text{的长度}}{\Omega\text{的长度}};\quad P(A)=\frac{A\text{的面积}}{\Omega\text{的面积}};\quad P(A)=\frac{A\text{的体积}}{\Omega\text{的体积}}.$$

1.定义 设E是随机试验，Ω是它的样本空间，对于E的每一事件A赋予一个**实数**，记作$P(A)$，若$P(A)$满足下列三条公理：

(1) **非负性：** 对于每一事件A，有 $P(A)\ge 0$;

(2) **规范性：** $P(\Omega)=1$;

(3) **可列可加性：** 对于两两互斥的事件A_1，A_2,$\cdots$，即$i\ne j$时，$A_iA_j=\varnothing$ $(i,j=1,2,\ldots)$,则有

$$P(A_1+A_2+\cdots+A_m+\cdots)$$
$$=P(A_1)+P(A_2)+\cdots+P(A_m)+\cdots.$$

则称$P(A)$为事件A的概率。

(3) 逆事件的概率： 对于任意事件A，有

$$P(\overline{A}) = 1 - P(A).$$

证 $A + \overline{A} = \Omega$, $A\overline{A} = \varnothing$,

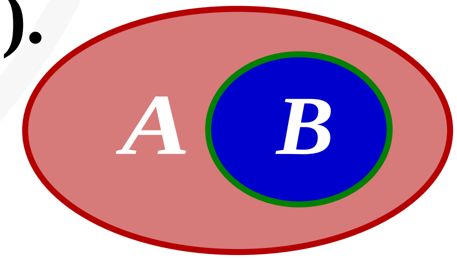
$\therefore P(A) + P(\overline{A}) = P(A + \overline{A}) = P(\Omega) = 1$,

即 $P(\overline{A}) = 1 - P(A)$.

(4) 若$A \supset B$，则 $P(A - B) = P(A) - P(B)$.

证 $\because B \subset A$, $\therefore A = A \cup B = B + (A - B)$.

又 $\because$ $B(A - B) = BA\overline{B} = \varnothing$,



$\therefore \quad P(A) = P(B) + P(A-B),$

即 $\quad P(A-B) = P(A) - P(B).$

一般的 $\quad P(A-B) = P(A) - P(AB)$

$\because A \supset AB$

$\therefore P(A-B) = P(A-AB) = P(A) - P(AB)$

推论1(单调性) 若$A \supset B$，则 $P(B) \leq P(A).$

证 由性质4，$P(A-B) = P(A) - P(B)$

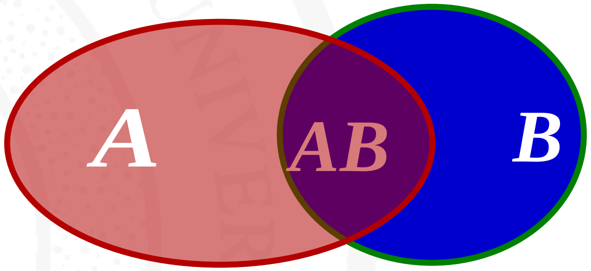
及 $P(A-B) \geq 0$，知命题成立.

(5) 概率的加法公式:

对于任意两个事件A, B，有

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

证 由图可得,



$$A \cup B = A + (B - AB),$$

且 $A \cap (B - AB) = \varnothing,$

故 $P(A \cup B) = P(A) + P(B - AB).$

又由性质 4 得

$$P(B - AB) = P(B) - P(AB),$$

因此得 $$P(A \cup B) = P(A) + P(B) - P(AB).$$

推论2 $P(A\cup B)\le P(A)+P(B).$

一般地， $P(\bigcup_{i=1}^{n}A_i)\le\sum_{i=1}^{n}P(A_i).$

推论3 设 A_1,A_2,A_3 是任意三个事件，则

$$P(A_1\cup A_2\cup A_3)$$
$$=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_2A_3)$$
$$-P(A_1A_3)+P(A_1A_2A_3).$$

一般地，对于任意n个事件A_1，$A_2,\cdots$，A_n，有

$$P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)-\sum_{1\le i<j\le n}P(A_iA_j)+$$
$$\sum_{1\le i<j<k\le n}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n).$$

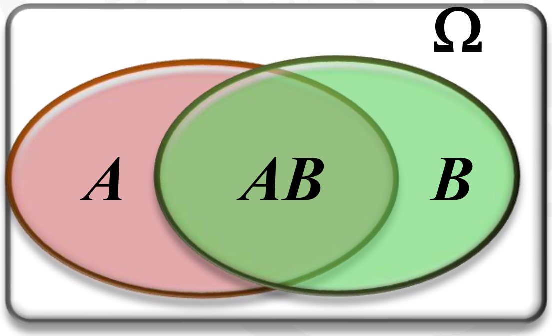
1.4.1 条件概率的定义

设A，B是两个事件，且$P(B) > 0$, 则称

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

为事件B发生的条件下，事件A发生的**条件概率**.

注



在古典概型条件下

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

1.4.2 条件概率的性质

(1)非负性： $0 \le P(A|B) \le 1;$

证 $\because AB \subset B \quad \therefore 0 \le P(AB) \le P(B)$

又 $\because P(B) > 0 \quad \therefore\ 0 \le \dfrac{P(AB)}{P(B)} \le 1$

即 $0 \le P(A|B) \le 1.$

(2)规范性： $P(\Omega|B) = 1,\ P(\varnothing|B) = 0;$

证 $\because \Omega B = B$

$$\therefore P(\Omega|B) = \frac{P(\Omega B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

(3) 可列可加性：

对于两两互斥的事件序列：$A_1, A_2, \cdots,$

有 $$P\left(\left(\sum_{k=1}^{\infty} A_k\right) \Big| B\right) = \sum_{k=1}^{\infty} P(A_k | B).$$

证 $$P\left(\left(\sum_{k=1}^{\infty} A_k\right) \Big| B\right) = \frac{P\left(\left(\sum_{k=1}^{\infty} A_k\right) B\right)}{P(B)}$$

$$= \frac{P\left(\sum_{k=1}^{\infty} (A_k B)\right)}{P(B)} = \frac{\sum_{k=1}^{\infty} P(A_k B)}{P(B)} = \sum_{k=1}^{\infty} P(A_k | B).$$

(4) 加法公式:

$$P((A_1 \cup A_2)|B) = P(A_1|B) + P(A_2|B) - P(A_1A_2|B).$$

证 $P((A_1 \cup A_2)B) = P(A_1B \cup A_2B)$

$$= P(A_1B) + P(A_2B) - P(A_1A_2B)$$

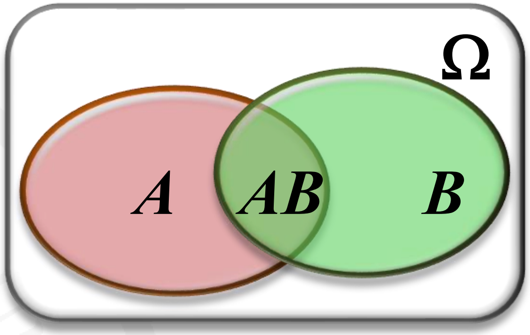
$$\therefore P((A_1 \cup A_2)|B) = \frac{P((A_1 \cup A_2)B)}{P(B)}$$

$$= \frac{P(A_1B) + P(A_2B) - P(A_1A_2B)}{P(B)}$$

$$= P(A_1 \mid B) + P(A_2 \mid B) - P(A_1A_2 \mid B).$$

(5) 逆事件的条件概率: $P(A|B) = 1 - P(\overline{A}|B).$

1.4.3 乘法公式



条件概率

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

$$\Downarrow$$

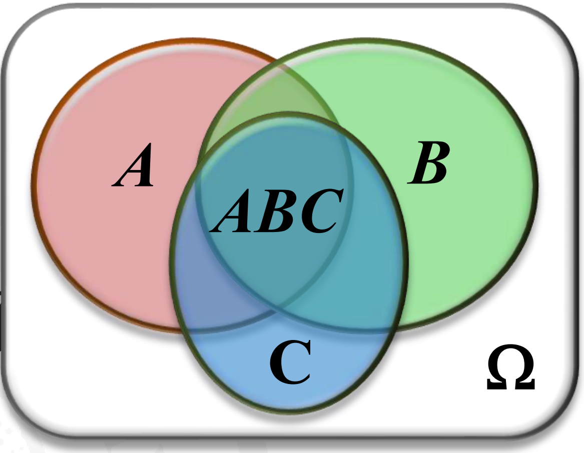
乘法公式

若$P(B) > 0$，则有 $P(AB) = P(B)P(A|B)$.

若$P(A) > 0$，则有 $P(AB) = P(A)P(B|A)$.

意义：积事件的概率等于某一事件的概率乘以另一事件在该事件发生条件下的**条件概率**。

乘法公式推广



设 A,B,C 为事件,且 $P(AB)>0$, 则有

$$P(ABC)=P(AB)P(C|AB)$$

$P(A)>0$

$P(B)>0$

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

简单 → 复杂 | 特殊 → 一般

乘法公式

若$P(B)>0$,则有 $P(AB)=P(B)P(A|B).$

若$P(A)>0$,则有 $P(AB)=P(A)P(B|A).$

乘法公式推广

设 $A_1,A_2\cdots A_n$是n个事件,若 $P(A_1A_2\cdots A_{n-1})>0$,

则 $P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots$

$\cdots P(A_n|A_1A_2\cdots A_{n-1}).$

复杂事件的概率 $\xrightarrow{\text{分解}}$ 简单事件概率的乘积

1.4.4 全概率公式与贝叶斯公式

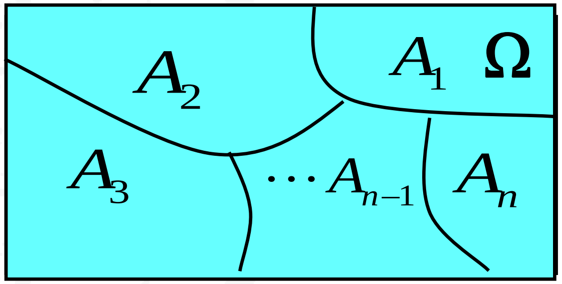
1. 样本空间的划分

定义 设Ω为试验E的样本空间，$A_1, A_2,\ldots, A_n$

为E的一组事件，若

(1) $A_iA_j = \varnothing, i \neq j, i, j = 1, 2, \cdots, n;$

(2) $\sum\limits_{i=1}^{n} A_i = \Omega.$



则称$A_1, A_2,\ldots, A_n$为Ω的一个划分。

化整为零

注： 1、同一样本空间的划分不唯一；

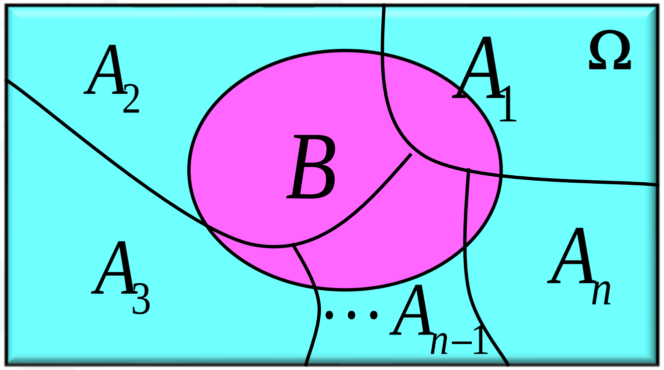
2、对于每次试验，$A_1, A_2,\ldots, A_n$中有且仅有一个发生。

(2)定理 设Ω为试验E的样本空间，B为E的事件，$A_1, A_2, \ldots, A_n$ 为Ω的一个划分，且$P(A_i)>0$ ($i=1,2,\ldots,n$)，则

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

称为全概率公式。（证明见分析）

注： $B \xrightarrow{\text{分解}} \sum_{i=1}^{n} BA_i$



3. 贝叶斯公式

B 结果 ⇒ A_i 原因

(1)分析

条件概率

$$P(A_i|B) = \frac{P(A_iB)}{P(B)}$$

乘法公式

$$= \frac{P(A_i)P(B|A_i)}{P(B)}$$

全概率公式

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

----贝叶斯公式

英国数学家 贝叶斯
Thomas Bayes
(1701-1761)

后验概率 ← 先验概率

结果

修正

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)}$$

贝叶斯公式 + 统计推断 ⇒ 贝叶斯统计

2. 两个事件的独立

(1) 定义 设 A, B 是两个事件，如果满足等式，

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

则称事件A, B**相互独立**，简称**独立**。

独立性：事件A的发生不影响事件B发生的概率。

独立 $\Leftrightarrow P(B|A)=P(B)$

$\Leftrightarrow P(AB)=P(A)P(B)$

(3)独立与互斥的关系

互斥是事件间本身的关系;

这是两个不同的概念.

两事件相互独立 $P(AB)=P(A)P(B)$

两事件互斥 $AB=\varnothing$

二者之间没有必然联系

两事件相互独立 $\not\longrightarrow$ 两事件互斥.

两事件互斥 $\not\longrightarrow$ 两事件相互独立.

可以证明： 特殊地，

当 $P(A)>0, P(B)>0$时， 有

A与B 独立 $\Rightarrow$ A与B不互斥 (相容),

或 A与B 互斥 $\Rightarrow$ A与B 不独立.

证 (1) 若A与B 独立, 则 $P(AB)=P(A)P(B).$

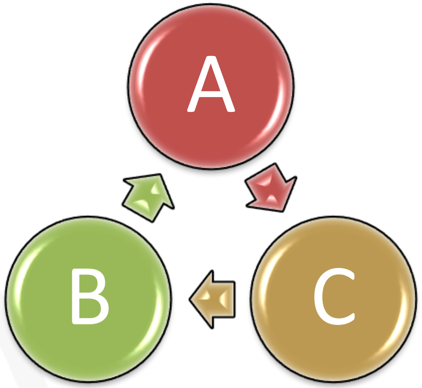
$\because\ P(A)>0,\ P(B)>0,$

$\therefore\quad P(AB)=P(A)P(B)>0,$

故 $AB\neq\varnothing,$

即 A与B 不互斥(相容).

3. 多个事件的独立性



(1) 三个事件的独立性

设A, B, C是三个事件，如果满足等式

两两相互独立

$$\left.\begin{aligned}&\left\{\begin{aligned}&P(AB)=P(A)P(B)\\&P(BC)=P(B)P(C)\\&P(AC)=P(A)P(C)\end{aligned}\right.\\&P(ABC)=P(A)P(B)P(C)\end{aligned}\right\}$$

相互独立

则称事件A, B, C 相互独立。

简单 → 复杂 | 特殊 → 一般

(2) n个事件的独立性

若对于任意 n ($n\geq2$) 个事件$A_1, A_2, \ldots, A_n$满足等式

$$\begin{cases} P(A_iA_j)=P(A_i)P(A_j) \\ P(A_iA_jA_k)=P(A_i)P(A_j)P(A_k) \\ \vdots \\ P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n) \end{cases}$$

两两相互独立

共C_n^2式子.

共$C_n^2+C_n^3+\cdots+C_n^n$
$=(1+1)^n-C_n^0-C_n^1$
$=2^n-1-n$个式子.

则称事件$A_1, A_2, \ldots, A_n$ 相互独立。

$A_1,A_2,\cdots,A_n$相互独立

$\overset{\Longrightarrow}{\not\Longleftarrow}$ $A_1,A_2,\cdots,A_n$两两相互独立.

分布函数的定义

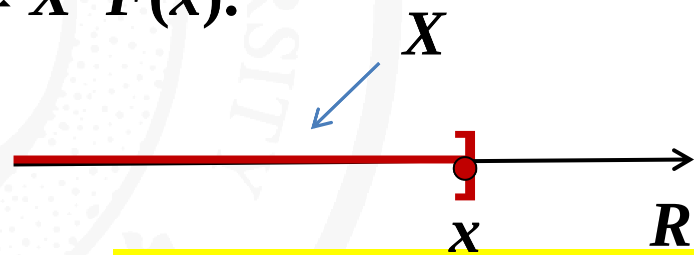
设X为随机变量，x为任意实数，称函数

$$F(x)=P\{X\le x\}\qquad -\infty<x<\infty$$

为随机变量X 的分布函数，记作 $X\sim F(x)$.

注

1° **实质：** 概率函数



2° $x\in R \Rightarrow F(x)\in[0,1]$

$$F(x)=P\{X\in[-\infty,x]\}$$

3° **几何意义**：表示随机点X落在$(-\infty, x]$的概率

分布函数的性质

(1) $0 \le F(x) \le 1,\ x \in R;$

(2) $F(x)$是单调不减的;

(3) $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0;$

$$F(+\infty) = \lim_{x \to +\infty} F(x) = 1;$$

(4) $F(x)$为右连续，即

$$\lim_{x \to x_0^+} F(x) = F(x_0),\quad x_0 \in R.$$

用分布函数表示随机事件的概率公式

$$F(x) \to P(A) = P\{X \in D\}$$

(1) $P\{X \le b\} = F(b);$

(2) $P\{a < X \le b\} = F(b) - F(a);$

(3) $P\{X > a\} = 1 - F(a);$

(4) $P\{X < b\} = F(b^-);$

(5) $P\{X = b\} = P\{X \le b\} - P\{X < b\} = F(b) - F(b^-);$

(6) $P\{a \le X < b\} = P\{X < b\} - P\{X < a\} = F(b^-) - F(a^-).$

3. 常见的离散型分布及其应用背景.

分布名称	记号	分布律	背景
退化分布 (单点分布)		$P\{X=c\}=1$	必然事件
两点分布 (或 0–1分布)	$X\sim B(1,p)$ $(0<p<1)$	$P\{X=k\}=p^k(1-p)^{1-k}$ $(k=0,1)$	伯努利事件

注：随机变量的所有可能取值范围

分布名称	记号	分布律	背景
离散型均匀分布		$P\{X=x_k\}=\dfrac{1}{n}$ $(k=1,2,\cdots,n)$	古典概型
二项分布	$X\sim B(n,p)$ $(0<p<1)$	$P\{X=k\}$ $=C_n^k p^k(1-p)^{n-k}$ $(k=0,1,\cdots,n)$	n重伯努利概型中，事件发生k次的概率
泊松分布	$X\sim P(\lambda)$ $(\lambda>0)$	$P\{X=k\}$ $=\dfrac{\lambda^k}{k!}e^{-\lambda}$ $(k=0,1,2,\cdots)$	稀有事件发生k次的概率

分布名称	记号	分布律	背景
几何分布		$P\{X=k\}$ $=(1-p)^{k-1}p$ $(k=1,2,\cdots)$	在n次伯努利试验中，A首次发生的试验次数为X.
超几何分布		$P\{X=k\}=\dfrac{C_M^k C_{N-M}^{n-k}}{C_N^n}$ $(k=0,1,\cdots,l)$ $l=\min\{M,n\})$ $n\le N, M<N$	设N件产品中有M件次品，从中任取n件，其中的次品数为X. (古典概型)

(1) 定义

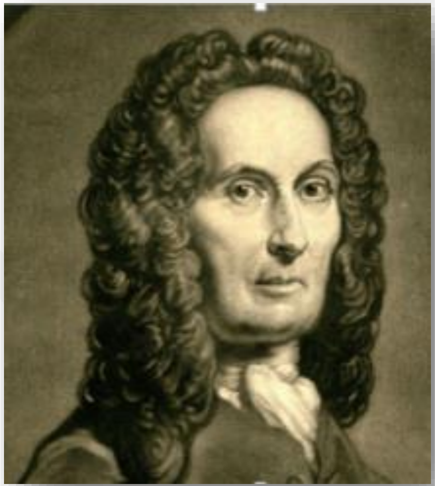
若随机变量 X的密度函数为

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

其中$-\infty < \mu < +\infty, \sigma > 0$为常数，则称 X

为服从参数μ和σ的正态分布[1]，记作 $X \sim N(\mu,\sigma^2)$.

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



法国数学家 棣莫佛
A. De Moivre
B. (1667-1754)

THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY · IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE AGRICULTURAL,AND ENGINEERING
IT ISAN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT

拓展阅读： 靳志辉，正态分布的前世今生

正态分布 $X \sim N(\mu,\sigma^2)$ $\xrightarrow{\mu=0,\ \sigma^2=1}$ 标准正态分布 $X \sim N(0,1)$

密度函数

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

分布函数

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

(3) 标准正态分布的性质

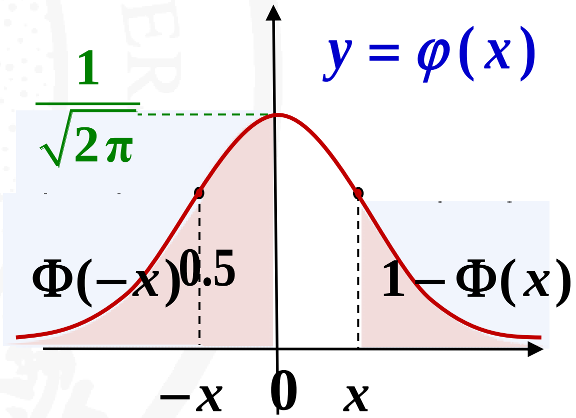
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t$$

1) $\varphi(x)$为偶函数;

2) $M=\max\varphi(x)=\varphi(0)=\frac{1}{\sqrt{2\pi}}$;

3) $\int_{-\infty}^{+\infty}\varphi(x)dx=1$ $\Rightarrow\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{x^2}{2}}dx=\sqrt{2\pi}$

$\Rightarrow\int_{-\infty}^{+\infty}\mathrm{e}^{-t^2}dt=\sqrt{\pi}$



4) $\Phi(0)=\int_{-\infty}^{0}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}dx=0.5$;

5) $\Phi(-x)=1-\Phi(x)$

(4) 正态分布概率的计算

$X \sim N(\mu, \sigma^2)$ $\qquad$ $P(X \le b)$

统计规律性 ⟹ 随机事件的概率

$$P(X \le b) = F(b) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{b} \boxed{e^{-\frac{(t-\mu)^2}{2\sigma^2}}} dt$$

?

2.2.1 多维随机变量及其分布

1. 二维随机变量的分布函数

$$F(x,y)=P\{X\le x, Y\le y\}.$$

意义：表示随机点(X,Y)落在平面区域$(-\infty,x]\times(-\infty,y]$的概率.

2. 二维离散型随机变量的分布律及分布函数

$$P\{X=x_i, Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots;$$

$$F(x,y)=\sum_{y_j\le y}\sum_{x_i\le x}p_{ij}.$$

$$P\{(X,Y)\in G\}=\sum_{(x_i,y_j)}\sum_{\in G}p_{ij},$$

3. 二维连续型随机变量的分布函数

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}p(u,v)\,\mathrm{d}u\,\mathrm{d}v.$$

其中$p(u,v)$为联合密度函数。

$$P\{(X,Y)\in G\}=\iint_{G}p(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

2.2.4 连续型随机变量的边缘分布

定义　若(X,Y)为二维连续型随机变量,设密度函数为$p(x,y)$,则X的边缘分布函数

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\left\{\underline{\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y}\right\}\mathrm{d}x$$

则X的边缘密度函数为

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}y$$

同理可得 Y的边缘密度函数

$$F_Y(y) = \int_{-\infty}^{y} [\int_{-\infty}^{+\infty} p(x, y) \mathrm{d} x] \mathrm{d} y,$$

$$p_Y(y) = \int_{-\infty}^{\infty} p(x, y) dx.$$

Y 的边缘密度函数.

2.2.5 随机变量的独立性

回忆：两事件 A,B独立的定义：

若 $P(AB)=P(A)P(B),$

则称事件 A,B独立.

设X,Y是两个随机变量,若对任意实数 $x,y,$

事件$\{X\le x\},\{Y\le y\}$是相互独立的,即

$$P\{X\le x,Y\le y\}=P\{X\le x\}P\{Y\le y\}$$

则称 X,Y是相互独立的.

用分布函数表示，即是

设X,Y是两个的随机变量，若对任意的x, y有

$$F(x,y)=F_X(x)F_Y(y)$$

则称X,Y相互独立.

它表明，两个随机变量(简记为$r.v$)相互独立时，它们的**联合分布函数**等于两个**边缘分布函数的乘积**．

对于**连续型随机变量**,上述定义等价于:

对于任意的x, y有

$$p(x,y)=p_X(x)p_Y(y)$$

成立,则称X,Y相互独立.

其中$p(x,y)$是X,Y的联合概率密度;

$p_X(x)$和$p_Y(y)$分别是X和Y的边缘概率密度.

事件的条件概率：$P(A|B)=P(AB)/P(B)$

1.离散型随机变量的条件分布

定义 设(X,Y)是二维离散型随机变量，对于固定的 j，若$P\{Y=y_j\}>0$,则称

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i;Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}}$$

其中 $i=1,2,\cdots$

为事件$\{Y=y_j\}$发生的条件下随机变量 X的条件分布律.

2. 连续型随机变量的条件分布

定义 设二维随机变量 (X,Y) 的联合密度函数为 $p(x,y)$, (X,Y) 关于 Y 的边缘密度函数为 $p_Y(y)$.若对于固定的 y, $p_Y(y)>0$, 则称 $\dfrac{p(x,y)}{p_Y(y)}$ 为在 $Y=y$ 的条件下 X 的**条件密度函数**,记为

$$p_{X|Y}(x|y)=\frac{p(x,y)}{p_Y(y)}.$$

$$\text{条件密度函数}=\frac{\text{联合密度函数}}{\text{边缘密度函数}}$$

称$\int_{-\infty}^{x} p_{X|Y}(x \mid y)\mathrm{d}x = \int_{-\infty}^{x} \frac{p(x,y)}{p_Y(y)}\mathrm{d}x$为在

$Y = y$的条件下,X的条件分布函数,记为

$$P\{X \le x \mid Y = y\} \text{ 或 } F_{X|Y}(x \mid y),$$

即$F_{X|Y}(x \mid y) = P\{X \le x \mid Y = y\} = \int_{-\infty}^{x} \frac{p(x,y)}{p_Y(y)}\mathrm{d}x$

同理可定义在$X = x$的条件下,Y的条件分布函数,

$$F_{Y|X}(y \mid x) = P\{Y \le y \mid X = x\} = \int_{-\infty}^{y} \frac{p(x,y)}{p_X(x)}\mathrm{d}y.$$

2. 公式法

定理 **(例2.18)** 设随机变量X具有概率密度$p_X(x)$，其中$-\infty < x < +\infty$.又设函数$f(x)$在（a,b）上可导，且恒有$f'(x) > 0$(或恒有$f'(x) < 0$),则$Y = f(X)$是连续型随机变量，其概率 密度为

$$p_Y(y) = \begin{cases} p_X[f^{-1}(y)] \cdot \left|[f^{-1}(y)]'\right|, & \alpha < y < \beta, \\ 0, & \text{其它}. \end{cases}$$

其中 $f^{-1}(y)$是 $f(x)$的反函数,(α,β)是$f^{-1}(y)$的定义域,

$$\left|[f^{-1}(y)]'\right| = \begin{cases} [f^{-1}(y)]', & \text{当 } f'(x) > 0\text{时}, \\ -[f^{-1}(y)]', & \text{当 } f'(x) < 0\text{时}. \end{cases}$$

2.3.3 常用结论

1. 离散型随机变量函数的分布律

若二维离散型随机变量 的联合分布律为

$$P\{X=x_i, Y=y_j\}=p_{ij}, i, j=1,2,\cdots$$

则随机变量函数 $Z=f(X,Y)$的分布律为

$$P\{Z=z_k\}=P\{f(X,Y)=z_k\}$$

$$=\sum_{z_k=f(x_i,y_j)} p_{ij} \qquad k=1,2,\cdots.$$

2. 连续型随机变量函数的分布

(1)$Z=X+Y, Z=X-Y$的分布

$$p_Z(z)=\int_{-\infty}^{+\infty}p(x,z-x)\mathrm{d}x=\int_{-\infty}^{+\infty}p(z-y,y)\mathrm{d}y$$

$$p_Z(z)=\int_{-\infty}^{+\infty}p(x,x-z)\mathrm{d}x=\int_{-\infty}^{+\infty}p(z+y,y)\mathrm{d}y$$

(2)$Z=\dfrac{X}{Y}$的分布

$$p_Z(z)=\int_{-\infty}^{+\infty}|y|\,p(yz,y)\mathrm{d}y$$

(3) $M=\max(X,Y)$ 及 $N=\min(X,Y)$ 的分布

当X，Y相互独立时，有

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

$$F_N(z)=1-[1-F_X(z)][1-F_Y(z)]$$

当X，Y相互独立且同分布时，有

$$F_M(z)=F^2(z)$$

$$F_N(z)=1-[1-F(z)]^2$$

推广： 一般地，设

$$M = \max\{X_1, X_2, \cdots X_n\},$$

$$N = \min\{X_1, X_2, \cdots X_n\},$$

则当 $X_1, X_2, \cdots, X_n$ 相互独立且同分布时，

有 $$F_M(z) = F^n(z)$$

$$F_N(z) = 1 - [1 - F(z)]^n$$

其中 $F(z) = P\{X_1 \le z\}.$

3. 其他常用结论

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

(2) $X_1 \sim P(\lambda_1), X_2 \sim P(\lambda_2)$, 且 $X_1 X_2$ 相互独立

$\Rightarrow X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$

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

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

(4) X,Y 相互独立且 $X \sim N(0,1), Y \sim N(0,1)$,

则 $Z = X / Y$ 服从柯西分布

$$p_Z(z) = \frac{1}{\pi}\frac{1}{1+z^2}$$

3.1 随机变量的数学期望

1. 数学期望

实数, 而非变量, 它是一种**加权平均**, 与一般的平均值不同, 它从本质上体现了随机变量 X 取可能值的**真正的平均值.**

$$E(X)=\sum_k x_k p_k.$$

$$E(X)=\int_{-\infty}^{+\infty} xp(x)\mathrm{d}x.$$

绝对收敛

常见离散型分布的数学期望小结

分布	分布律	$E(X)$
0–1 分布 $X \sim B(1, p)$	$P\{X=k\}=p^k(1-p)^{1-k}$ $k=0,1$	p
二项分布 $X \sim B(n, p)$	$P\{X=k\}=C_n^k p^k(1-p)^{n-k}$ $k=0,1,2,\dots,n$	np
泊松分布 $X \sim P(\lambda)$	$P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$, $k=0,1,2,\dots$	λ
几何分布	$P\{X=k\}=(1-p)^{k-1}p$ $k=1,2,\dots$	$\dfrac{1}{p}$

常见连续型分布的数学期望小结

分布名称	概率密度	$E(X)$
均匀分布	$p(x)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b]\\ 0, & \text{其他}\end{cases}$	$\dfrac{a+b}{2}$
正态分布	$p(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ
指数分布	$p(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$	$\dfrac{1}{\lambda}$
伽玛分布	$p(x)=\begin{cases}0, & x\le 0\\ \dfrac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}, & x>0\end{cases}$	$\dfrac{\alpha}{\beta}$

2. 随机变量函数的数学期望

$$E(Y)=E[f(X)]=\begin{cases}\sum_{k=1}^{\infty} f(x_k)p_k, X\text{为离散型}\\ \int_{-\infty}^{+\infty} f(x)p(x)\mathrm{d}x, X\text{为连续型}\end{cases}$$

$$E(Z)=E\left[f(X,Y)\right]=\begin{cases}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} f(x_i,y_j)p_{ij}\\ =\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)p(x,y)\mathrm{d}x\mathrm{d}y\end{cases}$$

3. 数学期望的性质

$$
\begin{cases}
1^0\ E(C)=C; \\
2^0\ E(CX)=C(X); \\
3^0\ E\left(\sum_{i=1}^{n} a_i X_i\right)=\sum_{i=1}^{n} a_i E(X_i); \\
4^0\ X,Y\text{独立} \Rightarrow E(XY)=E(X)E(Y).
\end{cases}
$$

数学期望 → 数字特征

3.2 方差与矩

1、方差的概念和意义

方差是一个体现随机变量*X* 取值分散程度的量. *D*(*X*)值大（小），表示*X* 取值分散程度大（集中），*E*(*X*) 的代表性差（好）

2、方差的计算

$$D(X)=E[X-E(X)]^2=E(X^2)-[E(X)]^2$$

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2 p_k,$$

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 p(x)\mathrm{d}x$$

3. 方差的性质

(1) $D(C)=0;$

(2) $D(CX)=C^2D(X);$

(3) X、Y相互独立，$D(X\pm Y)=D(X)+D(Y).$

(4)切比谢夫不等式

$$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}\Leftrightarrow P\{|X-\mu|<\varepsilon\}\geq 1-\frac{\sigma^2}{\varepsilon^2}.$$

(5). $D(X)=0\Leftrightarrow P\{X=C\}=1.$

4、常见分布的数学期望与方差

分布	分布律	$E(X)$	$D(X)$
0-1分布 $X \sim B(1, p)$	$P\{X=k\}=p^k(1-p)^{1-k}$ $k=0,1$	p	$p(1-p)$
二项分布 $X \sim B(n, p)$	$P\{X=k\}=C_n^k p^k(1-p)^{n-k}$ $k=0,1,2,\dots,n$	np	$np(1-p)$
泊松分布 $X \sim P(\lambda)$	$P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$ $K=0,1,2,\dots$	λ	λ
几何分布	$P\{X=k\}=(1-p)^{k-1}p$ $K=1,2,\dots$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$

常见分布的数学期望与方差

分布	密度函数	$E(X)$	$D(X)$
均匀分布 $X\sim U(a,b)$	$p(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其它}.\end{cases}$	$\dfrac{(a+b)}{2}$	$\dfrac{(a-b)^2}{12}$
正态分布 $X\sim N(u,\sigma^2)$	$p(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, $\sigma>0,-\infty<x<+\infty.$	u	σ^2
指数分布 $X\sim Exp(\lambda)$	$p(x)=\begin{cases}\lambda e^{-\lambda x} & x\ge 0\\ 0 & x<0\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$

5. 矩是随机变量的数字特征.

原点矩：$a_k = E(X^k) \quad (k = 1, 2, \cdots n)$

中心矩：$\mu_k = E\left[X - E(X)\right]^k.$

$a_1 = E(X);\qquad a_2 = E(X^2)$

$\mu_1 = E(X - EX) = 0;\quad \mu_2 = D(X) = a_2 - a_1^2$

3.3 协方差

1. 协方差与相关系数的定义及计算

X与Y的协方差：

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

$$=E(XY)-E(X)E(Y)$$

离散 $$\operatorname{cov}(X,Y)=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}x_i y_j p_{i,j}-E(X)E(Y),$$

连续 $$\operatorname{cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyp(x,y)\mathrm{d}x\mathrm{d}y-E(X)E(Y)$$

X与Y的相关系数：

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

$$X' = \frac{X-EX}{\sqrt{DX}}, Y' = \frac{Y-EY}{\sqrt{DY}}$$

$$\Rightarrow \begin{aligned} &E(X')=0,\\ &D(X')=1 \end{aligned}$$

2、协方差的性质

性质(1) $\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)$.

性质(2) $\operatorname{cov}(aX,bY)=ab\operatorname{cov}(X,Y)$, a, b为常数.

性质(3) $\operatorname{cov}(X_1+X_2,Y)=\operatorname{cov}(X_1,Y)+\operatorname{cov}(X_2,Y)$.

性质(4) 若X与Y独立, 则$\operatorname{cov}(X,Y)=0$.

性质(5) $D(X\pm Y)=D(X)+D(Y)\pm 2\operatorname{cov}(X,Y)$.

推广 $$D\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}D(X_i)+2\sum_{i<j}\sum\operatorname{cov}(X_i,X_j).$$

结论： $(X,Y)\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)\Rightarrow Cov(X,Y)=\rho$

3. 相关系数的意义与性质

衡量X与Y线性相关密切程度的指标

当$|\rho_{XY}|$接近1时,表明X,Y的线性关系联系较紧密.

当$|\rho_{XY}|$接近0时,X,Y线性相关的程度较差.

$\rho_{XY}=0$, 则称X和Y不相关.

性质(1) $|\rho_{XY}|\le 1.$

性质(2) $|\rho_{XY}|=1$的充要条件是,存在常数a,b

使$P\{Y=aX+b\}=1.$

性质(3) 相互独立 $\underset{\not\longleftarrow}{\longrightarrow}$ 不相关

特例： $(X,Y) \sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$

$$\rho = 0 \Leftrightarrow XY\text{相互独立} \Leftrightarrow \text{不相关}$$

4.不相关的充要条件

$$X,Y\ \text{不相关} \Leftrightarrow \rho_{XY} = 0;$$

$$\Leftrightarrow \operatorname{cov}(X,Y) = 0;$$

$$\Leftrightarrow E(XY) = E(X)E(Y);$$

$$\Leftrightarrow D(X+Y) = D(X) + D(Y).$$

4.1 大数定律

1. 频率的稳定性

随机现象在大量重复试验中呈现明显的统计规律性，即事件发生的频率具有稳定性.

$$\text{频率} \overset{n\to\infty}{\approx} \text{概率}$$

数学上如何准确刻画频率的稳定性？ 数列极限？

定理4.3 切比谢夫大数定律

设$X_1, X_2, \cdots, X_n, \cdots$是两两不相关的随机变量序列，每一随机变量都有有限的方差，并有公共的上界

$$D(X_1) \le C, D(X_2) \le C, \cdots, D(X_n) \le C, \cdots$$

则对任意的$\varepsilon > 0$，恒有

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

即 $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} EX_i$

定理4.5 辛钦大数定律

设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立,服从同一分布,且具有数学期望 $E(X_i)=\mu \quad (i=1,2,\cdots,n)$, 则对任意的 $\varepsilon>0$, 都有

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

$$\text{即} \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} E(X_i)$$

注1° 与切比谢夫大数定理相比, 不要求方差存在且有界.

2° 伯努利大数定理是辛钦大数定理的特例.

汇总

四个大数定理

切比谢夫大数定律

不相关、方差有界：$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}EX_i$

伯努利大数定律

独立同分布，$X \sim B(1,p)$ ：$\frac{\mu_n}{n} \xrightarrow{p} p$

泊松大数定律

独立不同分布，$X \sim B(1,p_k)$ ：$\frac{\mu_n}{n} \xrightarrow{p} \frac{1}{n}\sum_{k=1}^{n}p_k$

辛钦大数定律

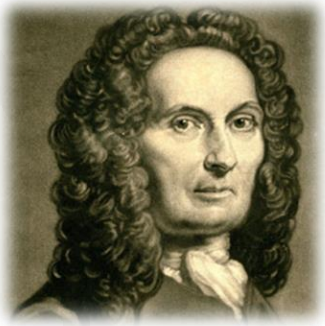
独立同分布：$\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{p} E(X_i)$

4.1 中心极限定理

定理4.7　棣莫佛-拉普拉斯(*De Movire - Laplace*)中心极限定理

设随机变量Y_n服从二项分布$B(n, p)$,则对任意x,有

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



A. De Moivre (1667-1754)

表 明　若$Y_n \sim B(n,p)$,

$$E(Y_n) = np$$
$$D(Y_n) = np(1-p)$$

则有$Y_n^* = \dfrac{Y_n - np}{\sqrt{np(1-p)}} \sim AN(0,1)$

P.-S. Laplace (1749-1827)

近似服从标准正态分布

De Moivre-Laplace中心极限定理

$$Y_n \sim B(n,p) \Rightarrow \frac{Y_n - np}{\sqrt{np(1-p)}} \sim AN(0,1)$$

$$\downarrow \qquad\qquad \downarrow$$

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

其中$\boxed{X_i \sim B(1,p)}$

$$\boxed{\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} \sim AN(0,1)}$$

特殊 $X_i \sim B(1,p)$ ⇒ 一般$X_i \sim F(x)$?

定理4.6 林德伯格-列维(Lindeberg-Levy)中心极限定理

设随机变量$X_1, X_2, \ldots, X_n$相互**独立**，服从**同一分布**，且存在

$$E(X_i) = \mu, \quad D(X_i) = \sigma^2 \neq 0 \quad (i=1, 2, \ldots, n)$$

则对任意x,有

$$\lim_{n\to\infty} P\left\{ \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq x \right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t = \Phi(x)$$

（证明略）

5.1 数理统计中的基本概念

基本概念: 总体 X (随机变量)

个体 $X_1, X_2, \cdots, X_n$ （随机变量）

（简单随机）样本 $(X_1, X_2, \cdots, X_n)$ （n维随机向量）

样本值 $(x_1, x_2, \cdots, x_n)$ （常量）

总体、样本、样本值的相互关系:

X 总体(理论分布)

估计

$(X_1, X_2, \cdots, X_n)$ 样本 → 样本值 $(x_1, x_2, \cdots, x_n)$

统计量: $f(X_1, X_2, \cdots, X_n)$ 随机变量

观测值: $f(x_1, x_2, \cdots, x_n)$ 常量

1、样本矩

样本均值 $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu = EX$

(1) $E(\bar{X}) = \mu$ (2) $D(\bar{X}) = \dfrac{1}{n}\sigma^2;$

样本方差 $S_n^{*2} = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \xrightarrow{p} \sigma^2 = DX$

(3) $E(S_n^2) = \dfrac{n-1}{n}\sigma^2;$ (4) $E(S_n^{*2}) = \sigma^2.$

样本矩 $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{p} \mu_k = E(X^k)$

2、次序统计量:

(1) 最大次序统计量$X_{(n)}$的分布密度为

$$p_{X_{(n)}}(x)=n[F(x)]^{n-1}p(x).$$

(2) 最小次序统计量$X_{(1)}$的分布密度为

$$p_{X_{(1)}}(x)=n[1-F(x)]^{n-1}p(x).$$

3、经验分布函数:

$$F_n(x)=\frac{\mu_n(x)}{n}\xrightarrow{p,a.e.}F(x)$$

$$nF_n(x)=\mu_n(x)=\sum_{i=1}^{n}I_i\sim B(n,F(x))$$

$$E[F_n(x)]=F(x),\quad D[F_n(x)]=\frac{F(x)[1-F(x)]}{n}$$

5.2常用统计分布

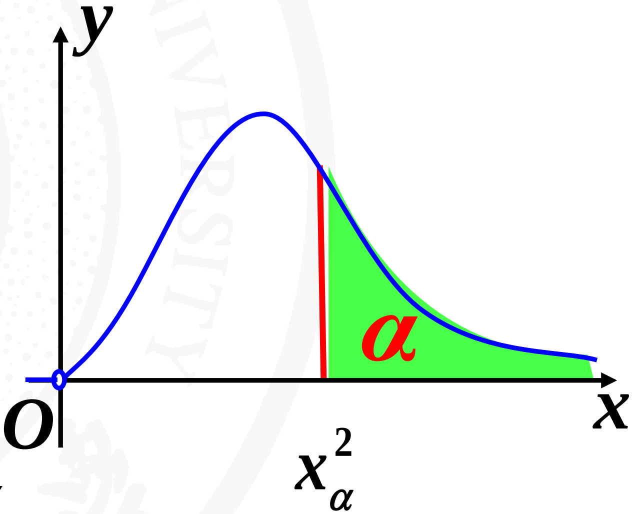
1.三大抽样分布:

χ^2 分布, t 分布, F 分布

的定义,性质.

2.概率分布的分位数概念. x_α

$$P\{X > x_\alpha\} = \alpha$$



$$\Rightarrow F(x_\alpha) = P(X \le x_\alpha) = 1 - \alpha$$

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

(1)χ^2 分布 $X_1, X_2, \cdots, X_n \sim N(0,1)$且$X_i$相互独立，

$$\chi_n^2 = X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$$

$$\Rightarrow E(\chi_n^2) = n, \quad D(\chi_n^2) = 2n$$

(2)t 分布 $X \sim N(0,1)$, $Y \sim \chi^2(n)$, 且 X, Y独立,

$$\text{则} T = \frac{X}{\sqrt{Y/n}} \sim t(n) \quad \Rightarrow E(T) = 0, \; D(T) = \frac{n}{n-2} (n > 2).$$

(3)F 分布

设$X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 且X, Y相互独立，

$$\text{则} F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)。$$

若$T \sim t(n)$, 则 $T^2 \sim F(1, n)$

分布	统计量	上侧分位数	性质
$N(0,1)$	U	μ_α	$\mu_{1-\alpha}=-\mu_\alpha$, $\Phi(\mu_\alpha)=1-\alpha$
$\chi^2(n)$	$\chi_n^2=\sum_{i=1}^{n}U_i^2$	$\chi_\alpha^2(n)$	$\chi_\alpha^2(n)\approx n+\sqrt{2n}\mu_\alpha$ $(n>60)$
$t(n)$	$T=\dfrac{U}{\sqrt{\chi_n^2/n}}$	$t_\alpha(n)$	$t_{1-\alpha}(n)=-t_\alpha(n)$, $t_\alpha(n)\approx\mu_\alpha(n>45)$
$F(n_1,n_2)$	$F=\dfrac{\chi_{n_1}^2/n_1}{\chi_{n_2}^2/n_2}$	$F_\alpha(n_1,n_2)$	$F_\alpha(n_1,n_2)=\dfrac{1}{F_{1-\alpha}(n_2,n_1)}$

5.3 抽样分布

抽样分布定理

1、单正态总体的抽样分布　$X \sim N(\mu,\sigma^2)$,

$$(1)\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \sigma^2/n)$$

$$(2)\ U = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$(3)\ \frac{nS_n^2}{\sigma^2} = \frac{(n-1)S_n^{*2}}{\sigma^2} \sim \chi^2(n-1)$$

$$(4)\ \frac{\bar{X}-\mu}{S_n^*/\sqrt{n}} = \frac{\bar{X}-\mu}{S_n/\sqrt{n-1}} \sim t(n-1).$$

2 两正态总体的抽样分布

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

(1) $\bar{X} \sim N(\mu_1,\dfrac{\sigma_1^2}{n_1})$ $\bar{Y} \sim N(\mu_2,\dfrac{\sigma_2^2}{n_2})$

$$\bar{X}-\bar{Y} \sim N(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2})$$

(2) 当 $\sigma_1^2=\sigma_2^2=\sigma^2$ 时,

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma(\dfrac{1}{n_1}+\dfrac{1}{n_2})} \sim N(0,1)$$

$$\frac{(n_1-1)S_1^{*2}}{\sigma_1^2}+\frac{(n_2-1)S_2^{*2}}{\sigma_2^2}\sim\chi^2(n_1+n_2-2)$$

$$\therefore T=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2),$$

其中 $S_w=\sqrt{\dfrac{(n_1-1)S_1^{*2}+(n_2-1)S_2^{*2}}{n_1+n_2-2}}.$

$$(3)\ \frac{(n_1-1)S_1^{*2}}{\sigma_1^2} \sim \chi^2(n_1-1), \quad \frac{(n_2-1)S_2^{*2}}{\sigma_2^2} \sim \chi^2(n_2-1),$$

$$\therefore F = \frac{S_1^{*2}/\sigma_1^2}{S_2^{*2}/\sigma_2^2} \sim F(n_1-1, n_2-1).$$

6.1 参数的点估计

$$X \sim F(x;\theta)$$

估计量 $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ **随机变量**

估计值 $\hat{\theta} = \hat{\theta}(x_1, x_2, \cdots, x_n)$ **常量**

两种求点估计的方法: $\begin{cases} \text{矩估计法} \\ \text{最大似然估计法} \end{cases}$

矩估计

(1) 计算总体矩。

$$E\left(X^k\right)=\int_{-\infty}^{+\infty}x^k dF\left(x;\ \theta_1,\ \theta_2,\cdots,\ \theta_m\right)=\alpha_k\left(\theta_1,\ \theta_2,\cdots,\ \theta_m\right)$$

(2) 用样本矩作为总体矩的估计，即令

$$\frac{1}{n}\sum_{i=1}^{n}X_i^k=\alpha_k\left(\hat\theta_1,\hat\theta_2,\ldots\hat\theta_m\right)\qquad\left(k=1,2,\cdots,m\right)$$

(3) 解该方程组得估计量

$$\hat\theta_k=\hat\theta_k\left(X_1,X_2,\cdots,X_n\right)\quad\left(k=1,2,\cdots,m\right)$$

估计值 $\hat\theta_k=\hat\theta_k\left(x_1,x_2,\cdots,x_n\right)$

最大似然估计

1° 求似然函数 $L(\theta)=\prod_{i=1}^{n} p(x_i;\ \theta)$ ；

2° 求出 $\ln L(\theta)$ 及似然方程

$$\frac{\partial \ln L(\theta)}{\partial \theta_i}\bigg|_{\theta=\hat{\theta}}=0 \quad (i=1,2,\cdots,\ m)$$

3° 解似然方程得到**最大似然估计值**

$$\hat{\theta}_i=\hat{\theta}_i(x_1,\ x_2,\cdots,\ x_m) \qquad (i=1,2,\cdots,\ m)$$

4° 最后得到**最大似然估计量**

$$\hat{\theta}_i=\hat{\theta}_i(X_1,\ X_2,\cdots,\ X_m) \quad (i=1,2,\cdots,\ m)$$

6.2估计量的评价标准

估计量的评选的三个标准 $\hat{\theta}(X_1, X_2 \cdots X_n)$

- 无偏性
- 有效性 ⟹ 最小方差无偏估计
- 相合性

无偏性 $\boxed{E\left(\hat{\theta}\right)=\theta} \Rightarrow \lim\limits_{n\to\infty} E\left(\hat{\theta}_n\right)=\theta$

有效性 $\boxed{E\left(\hat{\theta}\right)=\theta,}$ $\boxed{e\left(\hat{\theta}\right)=1} \Rightarrow \lim\limits_{n\to\infty} e(\hat{\theta})=1$

其中$e\left(\hat{\theta}\right)=\left(\dfrac{1}{nI\left(\theta\right)}\right)/D\left(\hat{\theta}\right)$

$$I\left(\theta\right)=E\left(\frac{\partial \ln p\left(x;\theta\right)}{\partial\theta}\right)^2=-E\left(\frac{\partial^2 \ln p\left(x;\ \theta\right)}{\partial\theta^2}\right)$$

相合性 $\lim\limits_{n\to\infty} P\{|\hat{\theta}_n-\theta|<\varepsilon\}=1$

$\Rightarrow$ $\lim\limits_{n\to\infty} E\left(\hat{\theta}_n\right)=\theta,$ $\lim\limits_{n\to\infty} D(\hat{\theta}_n)=0$

总结 $EX=\mu,\ DX=\sigma^2$

(1) $\bar{X}$是μ的矩估计，无偏估计，最小方差无偏估计，相合估计；

(2) S_n^2是σ^2的矩估计，渐近无偏估计；

(3) S_n^{*2}是σ^2的无偏估计，最小方差无偏估计，渐进有效估计；

若$X \sim N\left(\mu,\sigma^2\right) \Rightarrow$

$\hat{\mu}=\bar{X},\hat{\sigma}^2=S_n^2$为矩估计，最大似然估计

6.3 参数的区间估计

置信区间是一个随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$,它表示未知参数在该区间具有预先给定的置信度**1-*α***,

即对于任意的 $\theta \in \Theta$ ，有 $P(\hat{\theta}_1 \le \theta \le \hat{\theta}_2) = 1 - \alpha$.

求置信区间的步骤:

(1)构造含有样本和参数的统计量$W = W(X_1, X_2, \cdots X_n; \theta)$,
并确定其分布;

(2)给定置信度$1-\alpha$,利用统计量W的分布确定其范围,
使得$P(a \le W \le b) = 1 - \alpha$:

(3)由不等式$a \le W \le b$，反解出的取值范围$\hat{\theta}_1 \le \theta \le \hat{\theta}_2$,
即为置信区间$\left[\hat{\theta}_1, \hat{\theta}_2\right]$。

正态总体均值与方差的区间估计

1. 单个总体均值 μ 的置信区间

$$\begin{cases}(1)\ \sigma^2\text{已知，}\ U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1); & \left(\overline{X}\pm\dfrac{\sigma}{\sqrt{n}}u_{\alpha/2}\right).\\ (2)\ \sigma^2\text{未知，}\ T=\dfrac{\overline{X}-\mu}{S_n^*/\sqrt{n}}\sim t(n-1); & \left(\overline{X}\pm\dfrac{S_n^*}{\sqrt{n}}t_{\alpha/2}(n-1)\right).\end{cases}$$

2. 单个总体方差 σ^2 的置信区间

$$\chi^2=\frac{(n-1)S_n^{*2}}{\sigma^2}\sim\chi^2(n-1);\quad \left(\frac{(n-1)S_n^{*2}}{\chi^2_{\alpha/2}(n-1)},\frac{(n-1)S_n^{*2}}{\chi^2_{1-\alpha/2}(n-1)}\right).$$

其中$S_n^{*2}=\dfrac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)$

3. 两个总体均值差 $\mu_1-\mu_2$ 的置信区间

(1) σ_1^2 和 σ_2^2 均为已知，

$$\left(\overline{X}-\overline{Y}\pm u_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right).$$

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$，但 σ^2 为未知，

$$\left(\overline{X}-\overline{Y}\pm t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right).$$

其中 $S_w=\sqrt{\dfrac{(\sum_{i=1}^{n_1}X_i^2-n_1\overline{X}^2)+(\sum_{i=1}^{n_2}Y_i^2-n_2\overline{Y}^2)}{n_1+n_2-2}}$

4. 两个总体方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间

总体均值 μ_1, μ_2 为未知,

$$\left(\frac{S_1^{*2}}{S_2^{*2}}\frac{1}{F_{\alpha/2}(n_1-1, n_2-1)}, \frac{S_1^{*2}}{S_2^{*2}}\frac{1}{F_{1-\alpha/2}(n_1-1, n_2-1)}\right) \Leftrightarrow$$

$$\left(\frac{S_1^{*2}}{S_2^{*2}}F_{1-\alpha/2}(n_2-1, n_1-1), \frac{S_1^{*2}}{S_2^{*2}}F_{\alpha/2}(n_2-1, n_1-1)\right).$$

7.1 假设检验的基本概念

1、基本原理： 小概率推断原理

小概率事件$P(A)=\alpha\le 0.05$在一次试验中基本上不会发生。

2、基本概念

(1) 原假设与备择假设　　**(2) 检验统计量**

(3) 显著性水平

$$\alpha = P\{\text{小概率事件} \mid H_0\text{成立}\}$$

(4) 拒绝域

拒绝域W_1：拒绝原假设 H_0

的所有**样本值$x=(x_1, x_2, \cdots, x_n)$所组成的集合**.

拒绝原假设H_0的**检验统计量的取值范围**.

3、两类错误

第 I 类错误, 又叫**弃真错误**

$$\alpha = P\{\text{拒绝原假设}H_0 \mid H_0\text{为真}\}$$

第 II 类错误, 又叫**取伪错误**

$$\beta = P\{\text{接受}H_0 \mid H_0\text{不正确}\}$$

4、一般步骤

提出假设H_0/H_1

选择检验统计量 $T = f(X_1, X_2, \cdots, X_n) \overset{H_0}{\sim}$ 分布

确定拒绝域W $P\{x \in W_1 | H_0\} = P\{T \in B_\alpha | H_0\} = \alpha$

计算统计量的观察值 $t_0 = f(x_1, x_2, \cdots, x_n)$

作出推断 $\begin{cases} t_0 \in W \Rightarrow \text{拒绝} H_0 \\ t_0 \notin W \Rightarrow \text{接受} H_0 \end{cases}$

7.2 正态总体的均值和方差的假设检验

1. 根据实际问题的要求，提出待检验的假设H_0及备择假设H_1;

2. 选择适当的检验统计量,在 H_0成立的条件下,确定它的概率分布；$P(x \in W_1 | H_0) = P(U \in U_\alpha | H_0) = \alpha$

3. 给定显著性水平α，确定拒绝域W_1;

4. 根据样本观察值计算统计量的观察值;

5.根据统计量值是否落入拒绝域W_1内,作出拒绝或接受H_0的判断。

	原假设H_0	检验统计量	检验方法	拒绝域
1	$\mu=\mu_0$ (σ^2已知)	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$	U检验	$\lvert u\rvert \geq u_{\alpha/2}$
2	$\mu=\mu_0$ (σ^2未知)	$t=\dfrac{\overline{X}-\mu_0}{S_n^{*2}/\sqrt{n}}$	t检验	$\lvert t\rvert \geq t_{\alpha/2}(n-1)$
3	$\sigma^2=\sigma_0^2$ (μ未知)	$\chi^2=\dfrac{(n-1)S_n^{*2}}{\sigma_0^2}$	χ^2检验	$\chi_0^2 \geq \chi_{\alpha/2}^2(n-1)$或 $\chi_0^2 \leq \chi_{1-\alpha/2}^2(n-1)$

	原假设H_0	检验统计量	检验方法	拒绝域
5	$\mu_1=\mu_2$ （σ_1^2,σ_2^2已知）	$U=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	U检验	$\lvert u\rvert \geq u_{\alpha/2}$
6	$\mu_1=\mu_2$ （$\sigma_1^2=\sigma_2^2=\sigma^2$ 未知）	$t=\dfrac{\overline{X}-\overline{Y}}{S_w^*\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ $S_w^{*2}=\dfrac{(n_1-1)S_1^{*2}+(n_2-1)S_2^{*2}}{n_1+n_2-2}$	t检验	$\lvert t\rvert \geq t_{\alpha/2}(n_1+n_2-2)$
7	$\sigma_1^2=\sigma_2^2$ (μ_1,μ_2未知)	$F=\dfrac{S_1^{*2}}{S_2^{*2}}$	F检验	$F_0\geq F_{\alpha/2}(n_1-1,n_2-1)$或 $F_0\leq F_{1-\alpha/2}(n_1-1,n_2-1)$