# 概率论与数理统计课程内容总结

#### 指导提纲

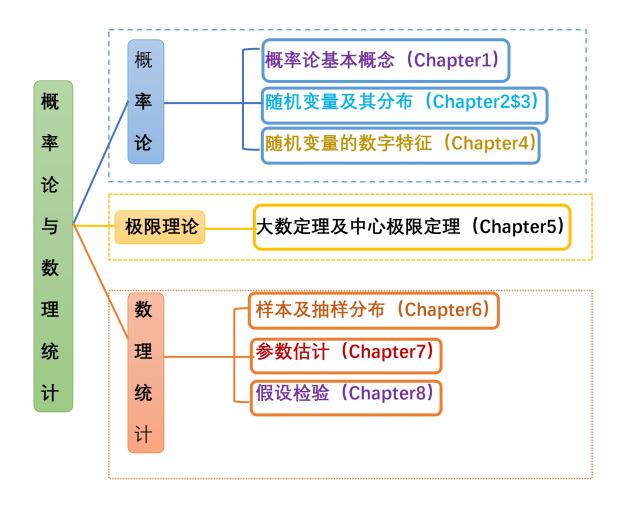
- ☞ 重要知识点回顾 (重点 I, 难点 D)
- ☞ 应试技巧和注意事项:解题规范(简答题写"解"开始,证明题写

"证明") 考试一律不用计算器, 带上有效证件

:

# 重要知识点回顾(重点 I, 难点 D)

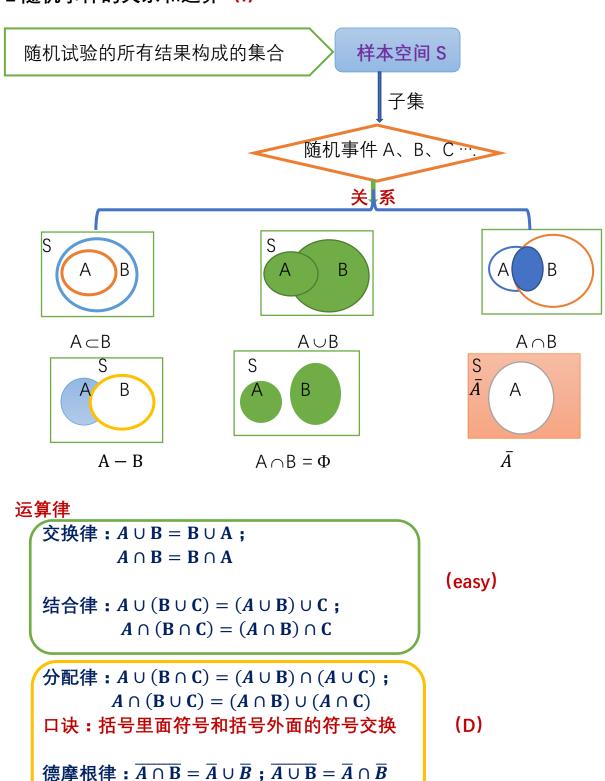
#### 三大模块:



#### 八章内容:

# Chapter1 概率论基本概念

## 1 随机事件的关系和运算(Ⅰ)



# 2 概率P(A)的 "3+6" 性质 (I)

#### 3条基本性质

非负性:  $P(A) \ge 0$ 

规范性: P(S) = 1

可列可加性:设 $A_1, A_2, ...$ 两两互斥,

 $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ 

推出

6条重要性质

不可能事件概率:  $P(\emptyset) = 0$ 

有限可加性: 设 $A_1, A_2, ..., A_n$ 两两互斥,

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

差事件概率: 若A C B, 则

P(B - A) = P(B) - P(A)

 $P(B) \ge P(A)$ 

概率不超过 1: P(A) ≤ 1

逆事件概率:  $P(\bar{A}) = 1 - P(A)$ 

加法公式:  $P(A \cup B) = P(A) + P(B) - P(AB)$ 

多个事件加法公式推广

# 3条件概率

定义:  $P(B|A) = \frac{P(AB)}{P(A)}$  (其中P(A) > 0) (1)

乘法公式:  $P(AB) = P(B \mid A)P(A), (P(A) > 0)$ 

# 3 大公式 (D)

# 全概率公式:

设随机试验 E 的样本空间为 S, A 为 E 的事件  $B_1, B_2, ..., B_n$ 为 S 的一个划分,且 $P(B_i) > 0$ ,则  $P(A) = \sum_{i=1}^n P(A|B_i)P(B_i)$ 

贝叶斯公式:全概率公式条件+P(A) > 0,有

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

# 4 独立性 (I)

定义及等价性:  $\frac{P(AB) = P(A)P(B) \Leftrightarrow}{P(B \mid A) = P(B), P(A) > 0}$ 

互推性: A与B, Ā与B, A与B, Ā与B独立性可以相互推出

# Chapter2(一维)随机变量及其分布

分布函数 F(x)

分布律 (离散型)

随机变量的分布

密度函数 (连续型)

随机变量函数的分布(离散型&连续型)

1 F(x)的定义和性质 (I) 定义域是整个实轴

$$F(x)$$
的定义:  $\forall x \in R$ ,  $F(x) = P\{X \le x\}$  (1)

x

**F**(x)的性质: 概率性:0

单调性:  $F(X_2) \ge F(X_1), X_2 > X_1$ 

概率性: $0 \le F(x) \le 1, F(-\infty) = 0, F(+\infty) = 1$  (1)

右连续性: F(x + 0) = F(x)

# 2 离散型随机变量的分布

分布律(离散型专属)(I)

$$P(X = x_k) = p_k \ (k = 1, 2, ...)$$
 或  $X \mid x_1 \quad x_2 \quad x_3 \quad ...$   $p_k \mid p_1 \quad p_2 \quad p_3 \quad ...$ 

其中:  $p_k \ge 0$ ,  $\sum_k p_k = 1$ 

# 三种重要分布的分布率

(0-1) 分布:

# 二项分布b(n,p):

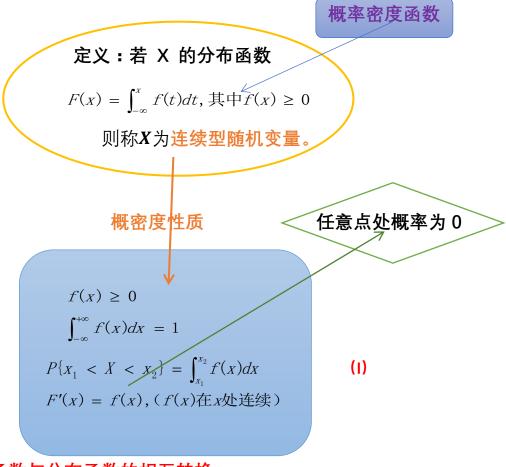
$$P\{x = k\} = C_n^k p^k (1 - p)^{n-k}, k = 0,1,..., n$$

泊松分布:  $\pi(\lambda)$ 

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

注:分布律与分布函数的相互转换及分布函数的特点

#### 3 连续型随机变量的分布(Ⅰ)



注:密度函数与分布函数的相互转换

**例题:**已知随机变量X的概率密度函数为

$$f(x) = \begin{cases} kx, & 0 < x < 2, \\ 0, & \text{#$\overline{c}$.} \end{cases}$$

试求:(1) k的值;(2) 分布函数F(x);(3)  $P{0 < x < 1}$ 。

#### 三种重要分 布的概率密度函数 (1)

均匀分布U(a,b):

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

指数分布:

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, x > 0 \\ 0, & \text{#th} \end{cases}$$

正态分布 $N(\mu, \sigma^2)$ :

正态分布密度函数图像的对称性:对称轴 $x = \mu$ 

#### 4 随机变量函数Y = g(X)的分布

#### 离散型随机变量的函数Y = g(X)的分布律

**例题:**已知随机变量 X 的分布律如下:

求
$$Y = X^2 - 1$$
的分布律 (easy)

# 连续型随机变量的函数Y = g(X)的密度函数 $f_V(y)$

分布函数法(万能方法): 先求分布函数, 然后求导得密度函数

定理法:(注意条件处处可导,严格单调)

设随机变量X的密度函数 $f_X(x)$ ,  $-\infty < x < +\infty$ , 函数g(X)处处可

导且恒有g'(X) > 0(< 0),则Y = g(X)的密度函数为

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)|, & \alpha < y < \beta \\ 0, & 其它 \end{cases}$$

其中,  $\alpha = \min\{g(-\infty), g(+\infty)\}, \beta = \max\{g(-\infty), g(+\infty)\}, h(y) = g^{-1}(x)_{0}$ 

**例 1 :** 设随机变量 X 服从(-1, 1)上的均匀分布,求(1)  $Y = X^2$ ;

(2) Z = |X|的概率密度函数。(注:不符合定理单调性条件)

**例 2:**设随机变量 X 服从参数为 1/2 的指数分布,证明 $Y = 1 - e^{-2X}$  在区间(0, 1)上服从均匀分布。**(符合定理单调性条件)** 

特别注意:若
$$X \sim N(\mu, \sigma^2)$$
,则 $\frac{X-\mu}{\sigma} \sim N(0, 1)$ 

# Chapter3 多维随机变量及其分布

#### (注意与一维随机变量之间的联系与区别)

分布函数 F(x,y)

分布律 (离散型)

密度函数 (连续型)

边缘分布 (独立性)(1)

条件分布

随机变量函数的分布(离散型&连续型)

二维随机变量分布函数 F(x,y)

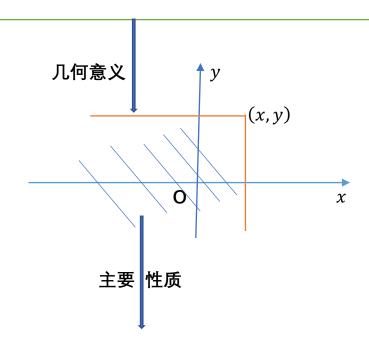
多维随机变量的分布

定义

设(X,Y)是二维随机变量, $\forall x,y \in R$ ,

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

称为(X,Y)的分布函数或称为X和Y的联合分布函数



- 1. F(x,y)关于x和y的不减函数;
- **2.**  $0 \le F(x, y) \le 1$

$$F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0$$
,  $x, y$  固定  $F(\infty, \infty) = 1$  ;

3.F(x,y)关于x右连续;关于y右连续;

### 二维离散型随机变量的分布律(I)

(X,Y)所有可能取的值: $(x_i,y_j)$ , i,j=1,2,3,...

记P
$$\{X = x_i, Y = y_j\} \triangleq p_{ij}, i, j = 1,2,3, ...,$$
 (\*\*)

易知
$$p_{ij} \ge 0$$
,  $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$ 

称之为

(X,Y)的分布律或X和Y的联合分布律

分布律的 表格法

				•			
Y	$x_1$	$x_2$	$x_3$		$x_i$		(从小到大排列)
$y_1$	$p_{11}$	$p_{21}$	$p_{31}$		$p_{i1}$		
$y_2$	$p_{12}$	$p_{22}$	$p_{32}$		$p_{i2}$	•••	
$y_3$	$p_{13}$	$p_{23}$	$p_{33}$		$p_{i3}$	•••	
$y_j$	$p_{1j}$	$p_{2j}$	$p_{3j}$	•••	$oldsymbol{p_{ij}}$	•••	
•							
	•						
	•						
(从小到大排列)							

#### 二维连续型随机变量的密度函数及分布函数(D)

随机变量(X,Y), 如果 $\forall x,y \in R$ ,有:

$$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x,y) dx dy$$
,其中  $f(x,y) \ge 0$ ,

称(X,Y)为二**维连续型随机变量,称**f(x,y)为(X,Y)的密度函数

或X和Y的联合密度函数。

注:二重积分及其几何意义——体积

*f(x,y*)的性质

(比较f(x)的性质)

**1.**  $f(x,y) \ge 0$ ;

2.  $F(\infty,\infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1;$ 

(x,y)

F(x,y)几何意义

- 3.  $P\{(x,y) \in G\} = \iint_G f(x,y) \, dx dy;$
- 4.  $\frac{\partial F(x,y)}{\partial x \partial y} = f(x,y)$ , (f(x,y)在(x,y)点连续)。

**例题:**设(X,Y)的密度函数

$$f(x,y) = \begin{cases} ke^{-2x-3y}, & x > 0, y > 0 \\ 0, & else \end{cases}$$

求: (1) k的值; (2) F(x,y); (3) $P\{Y \ge X\}$ ; (4)  $P\{3Y \ge 2X\}$ 

(提示:画出密度函数有效区域图)

#### 边缘分布 &独立性(I)

Case1: (X, Y)——离散 (easy)

(X, Y)分布律的 表格法

Y	<i>x</i> <sub>1</sub>	$x_2$	$x_3$		$x_i$		$P\{Y=y_j\}$	
$y_1$	$p_{11}$	$p_{21}$	$p_{31}$		$p_{i1}$	•••	$p_{\cdot 1}$	
$y_2$	$p_{12}$	$p_{22}$	$p_{32}$	••••	$p_{i2}$	•••	$p_{\cdot 2}$	
$y_3$	$p_{13}$	$p_{23}$	$p_{33}$	••••	$p_{i3}$		$p_{\cdot 3}$	
					1 1		•	
					1 1		•	
					1 1			
$y_{j}$	$p_{1j}$	$p_{2i}$	$p_{3i}$	•••	$p_{ij}$	•••	$p_{\cdot j}$	
•	'	,	,					
	.				1 1			
	[. ]						•	
<b>P</b> { <i>X</i>	$p_1$ .	$p_2$ .	$p_3$ .		$p_i$ .	•••	1	
$=x_i$		· · · · · · · · · · · · · · · · · · ·						
关于		同理,	关于 Y	的分布律	。(另列表格)			

Case2: (X, Y)—连续 (联合密度f(x, y))

关于X的(边缘)密度函数为:  $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$ 

关于Y的(边缘)密度函数为:  $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$ 

# **例题 1:**设(X,Y)的密度函数

$$f(x,y) = \begin{cases} 6e^{-2x-3y}, & x > 0, y > 0\\ 0, & else \end{cases}$$

求关于 X 和 Y 的边缘密度函数 $f_X(x)$ 和 $f_Y(y)$ 。

(提示:画出密度函数有效区域图)

解得: 
$$f(x,y) = \begin{cases} 6e^{-2x-3y}, x > 0, y > 0 \\ 0, & else \end{cases}$$

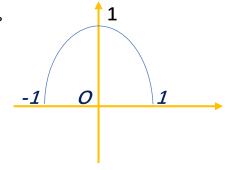
$$f_X(x) = \begin{cases} 2e^{-2x}, x > 0, \\ 0, & else \end{cases}$$

$$f_Y(y) = \begin{cases} 3e^{-3y}, y > 0, \\ 0, & else \end{cases}$$

易知: 
$$f(x,y) = f_X(x) \times f_Y(y)$$

**例题 2:设**(X,Y)在区域D: $\{(x,y)|0 \le y \le 1-x^2\}$ 上服从**均匀分布**,

求关于 X 和 Y 的边缘密度函数 $f_X(x)$ 和 $f_Y(y)$ 。



解得: 
$$f(x,y) = \begin{cases} 3/4, & 0 \le y \le 1 - x^2 \\ 0, & else \end{cases}$$

$$f_X(x) = \begin{cases} 3(1-x^2)/4 & -1 < x < 1, \\ 0, & else \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{3\sqrt{1-y}}{2}, & 0 < y < 1, \\ 0, & else \end{cases}$$

**易知:**
$$f(x,y) \neq f_X(x) \times f_Y(y)$$

注:观察两个例题的结果,有什么区别?

#### 独立性(I)

#### 相互独立

$$If \forall x, y \in R,$$
事件 $(X \le x)$ 与事件 $(Y \le y)$ 独立,then: 
$$F(x,y) = P\{(X \le x) \cap (Y \le y)\} \triangleq P\{X \le x, Y \le y\}$$
$$= P\{X \le x\}P\{Y \le y\} \text{ 即}$$
$$= F_X(x)F_Y(y) \longrightarrow \text{称}X, Y \text{ 相互独立}$$

#### Case1:离散型

(X, Y)分布律的 表格法

		-	-				
Y	$x_1$	$x_2$	$x_3$		$x_i$		$\mathbf{P}\{Y=y_j\}$
$y_1$	$p_{11}$	$p_{21}$	$p_{31}$		$p_{i1}$		$p_{\cdot 1}$
$\boldsymbol{y_2}$	$p_{12}$		$p_{32}$		$p_{i2}$		$p_{\cdot 2}$
$y_3$	$p_{13}$	$p_{23}$	$p_{33}$	••••	$p_{i3}$	•••	$p_{\cdot 3}$
							-
•							•
•							•
$y_j$	$p_{1j}$	$p_{2j}$	$p_{3j}$	•••	$p_{ij}$	1.1	$p_{\cdot j}$
•	-/						•
•							•
•	•						-
D(W							
<b>P</b> { <b>X</b>	$p_1$ .	$p_2$ .	$p_3$ .	••••	$p_i$ .	•••	
$=x_i$							

$$\forall i, j, P\{X = x_i, Y = y_i\} = P\{X = x_i\}P\{Y = y_i\}$$
 (每一个满足)

#### Case2:连续型

$$f(x,y) = f_X(x) \times f_Y(y)$$
 (几乎处处成立)

例题1相互独立,例题2不独立

# 条件分布

Case1: (X, Y)——离散型 (I)

直接根据条件概率 $P(B|A) = \frac{P(AB)}{P(A)} (P(A) > 0)$ 做就可以了。

(X,Y)的分布律: $P\{X = x_i, Y = y_j\} = p_{ij}, i, j = 1,2,3,...$ ,即:

Y	$x_1$	$x_2$	$x_3$		<i>x</i> <sub>i</sub>	$P\{Y=y_j\}$
$y_1$	$p_{11}$	$p_{21}$	$p_{31}$		<i>p</i> <sub>i1</sub>	$p_{\cdot 1}$
$y_2$	$p_{12}$	$p_{22}$	$p_{32}$	••••	$p_{i2}$	$p_{\cdot 2}$
$y_3$	$p_{13}$	$p_{23}$	$p_{33}$	••••	<i>p</i> <sub>i3</sub>	$p_{\cdot 3}$
	•					•
	•					•
	$p_{1j}$	$p_{2j}$	$p_{3j}$	•••	<b>p</b> <sub>ij</sub>	•
$y_j$	•					$p_{\cdot j}$
•	•					•
	•					•
	•					•
$\mathbf{P}\{X=x_i\}$	$p_1$ .	$p_2$ .	$p_3$ .	/	p <sub>i</sub>	1

如果  $P\{X = x_i\} > 0$  (i固定)

则: 
$$P\{Y = y_j | X = x_i\} = \frac{P\{X = x_i | Y = y_j\}}{P\{X = x_i\}} \quad (j = 1, 2, 3, ...)$$
$$= \frac{p_{ij}}{p_{i}} \quad (分布律)$$

同理  $P\{X = x_i | Y = y_j\} = \frac{p_{ij}}{p_{ij}}$  ( $i=1,2,3,...,p_{ij} > 0$ ). (分布律)

# 注:条件概率与条件分布律的区别。

**例题**: (X,Y)的分布律如下,

Y	0	1	2	3	
1	0	3/8	3/8	0	
3	1/3	0	0	1/8	

**求**:(1)  $P{X = 2|Y = 1}$ ;(2) 当Y = 1时,关于X的条件分布律。

#### Case2: (X, Y)——连续型. (D)

已知(X,Y)的密度函数f(x,y),  $x,y \in R$ 

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

#### 随机变量函数的分布

Case1:离散型 (easy) (I)

**例1:**已知二维离散型随机变量(X,Y)的概率分布律如下:

求:X+Y, X-2Y, XY, X/Y,  $X^2 + Y - 1$ ,  $\max\{X,Y\}$ ,  $\min\{X,Y\}$  的分布律。

Case2: 连续型 (D)

方法 1:公式法

设(X,Y)的概率密度函数为  $f(x,y), x,y \in R$ ,

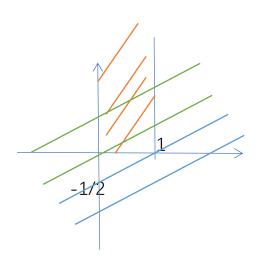
- (1) Z = X + Y的概率密度函数  $f_Z(z) = \int_{-\infty}^{+\infty} f(z y, y) dy$
- (2) Z = Y / X概率密度函数  $f_{Y/X}(z) = \int_{-\infty}^{+\infty} |x| f(x, xz) dx$
- (3) Z = XY 概率密度函数  $f_{XY}(z) = \int_{-\infty}^{+\infty} \frac{1}{|X|} f(x, \frac{Z}{X}) dx$

# 方法 2:分布函数法 (万能方法)

**例:** 已知(X,Y)的密度函数为:

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

求: Z=X-2Y 的概率密度。

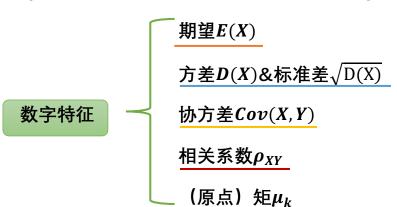


没法用现成公式,只能利用分布函数法,解题思路:

$$F_Z(z) = P\{Z \le z\} = P\{X - 2Y \le z\}$$

$$= \iint_{X-2Y \le z} f(x, y) dx dy \text{ (讨论 z 的取值)}$$

# Chapter4 随机变量的数字特征(So easy)



#### 期望E(X) (1)

#### 6个定义(了解)

Case1: X——离散型

定义 1:设离散型随机变量X的分布律为

$$P{X = x_k} = p_k, k = 1,2,...$$

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛,则级数 $\sum_{k=1}^{\infty} x_k p_k$ 的和为随机变量X

的数学期望(或均值),记为E(X)。即

$$E(X) = \sum_{k=1}^{\infty} x_k p_k$$

Case2: X——连续型

定义 2: 设连续型随机变量X的概率密度为f(x)

若积分 $\int_{-\infty}^{\infty} x f(x) dx$ **绝对收敛(绝对可积),则称** $\int_{-\infty}^{\infty} x f(x) dx$ 的值

为随机变量X的期望(或均值),记为E(X)。即

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$
。 别忘了我哦!

Case3: Y = g(X)(X 离散, g连续), E(Y) = ?

定义 3(定理):设Y = g(X), X为离散型随机变量且分布律为

$$P{X = x_k} = p_k, k = 1, 2, ...$$

若级数 $\sum_{k=1}^{\infty} g(x_k) p_k$ 绝对收敛,则

$$E(Y) = E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_{k,0}$$

Case4:Y = g(X)(X 连续, g连续), E(Y) = ?

定义 4(定理):设Y = g(X), X为连续型随机变量且概率密度为f(x)

若积分 $\int_{-\infty}^{\infty} g(x) f(x) dx$ **绝对收敛(绝对可积),**则

$$E(Y) = \int_{-\infty}^{\infty} g(x) f(x) dx_{\circ}$$

Case5: Z = g(X, Y)((X,Y))离散, g连续), E(Z) = ?

定义 5(定理):设Y = g(X,Y), (X,Y)为离散型且分布律为

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

若级数 $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} g(x_i, y_j) p_{ij}$ 绝对收敛,则

$$E(\mathbf{Z}) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} g(x_i, y_j) p_{ij}$$

Case6: Z = g(X, Y)((X,Y)连续, g连续), E(Z) = ?

定义 6(定理):设Y = g(X,Y), (X,Y)为连续型且概率密度为f(x,y)

若积分  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) dx dy$  **绝对收敛(绝对可积),**则

$$E(Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

性质 (I): 设X、Y为随机变量, C为常数, 有:

$$1, E(C) = C$$

$$2 \setminus E(CX) = CE(X)$$

$$E(aX + bY)$$

2. 
$$E(CX) = CE(X)$$
 
$$E(aX + bY)$$
3.  $E(X + Y) = E(X) + E(Y)$  
$$= aE(X) + bE(Y)$$

$$=aE(X) + bE(Y)$$

4、当X、Y相互独立时, E(XY) = E(X)E(Y) (注意条件)

例题 1:已知X为离散型随机变量且分布律如下, $Y = X^2$ ,求  $E(X), E(Y)_{\circ}$ 

例题 2:已知X为连续型随机变量且概率密度如下,求E(X)。

$$f(x) = \begin{cases} 1+x, & -1 \le x \le 0 \\ 1-x, & 0 < x < 1 \\ 0, & else \end{cases}$$

例题 3: 设(X,Y)的分布律如下, Z = X + Y, 求E(X), E(Z)。

Y	-1	0	1	2	
0	0	2/8	3/8	0	
1	1/8	1/8	0	1/8	

不求Y的密度函数,直接公式计算 例题 4:设随机变量(X,Y)的密度函数如下,求E(Y), $E(\frac{1}{XY})$ 

$$f(x,y) = \begin{cases} \frac{3}{2x^3y^2}, & \frac{1}{x} < y < x, x > 1\\ & 0, else \end{cases}$$

# 方差D(X)&标准差 $\sqrt{D(X)}$

定义:设X是一个随机变量, 若 $E\{[X - E(X)]^2\}$ 存在, 则称 $E\{[X - E(X)]^2\}$ 为X的方差, 记为D(X) 或 Var(X), 同时称 $\sqrt{D(X)}$ 为X的标准 差或均方差, 记成 $\sigma(X)$ 。

即:  $D(X) = E\{[X - E(X)]^2\} = E(X^2) - E^2(X)$ 

**性质**(I):设X、Y为随机变量,C为常数,有:

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

$$2 \setminus D(CX) = C^2D(X) \longrightarrow D(-X) = D(X)$$

3, 
$$D(X + Y) = D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\}$$

特殊地、当X、Y相互独立时、D(X+Y) = D(X) + D(Y)

$$4 \times D(X) = 0$$
的充要条件是 $X$ 以概率为  $1$  取常数 $E(X)$ ,即 
$$P\{X = E(X)\} = 1$$

# 期望和方差的重要结果:(1)

对象	b(1,p)	b(n,p)	$\pi(\lambda)$	U(a,b)	指数 分布	$N(\mu, \sigma^2)$	$\chi^2(n)$	$\bar{X}$	$S^2$
$E(\cdot)$	p	np	λ	$\frac{a+b}{2}$	θ	μ	n	μ	$\sigma^2$
D(·)	p(1-p)	np(1-p)	λ	$\frac{(a-b)^2}{12}$	$\theta^2$	$\sigma^2$	2 <i>n</i>	$\frac{\sigma^2}{n}$	/

# 协方差Cov(X,Y) (I)

定义:量 $E\{[X-E(X)][Y-E(Y)]\}$ 称为随机变量X与Y的协方差,记为Cov(X,Y).

#### Cov(X,Y)的计算公式变形:

①
$$Cov(X,Y) = E\{[X - E(X)][Y - E(Y)]\}$$
 我最简单、常用

- $(2) \operatorname{Cov}(X,Y) = E(XY) E(Y)E(Y)]$
- 3Cov(X,Y) = [D(X+Y) D(X) D(Y)]/2

# Cov(X,Y)的性质(线性性):

- $0. \operatorname{Cov}(X,X) = D(X)$
- 1. Cov(aX, bY) = abCov(X, Y)
- 2.  $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$

## 相关系数 $\rho_{XY}$ (I)

定义:  $\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$  称为随机变量X与Y的相关系数。

当  $\rho_{XY} = 0$ 时,称X和Y不相关。

# $\rho_{XY}$ 的性质:

- 1.  $|\rho_{XY}| \le 1$  (结果不要超出这个范围)
- $2. |\rho_{XY}| = 1$ 的充要条件是存在常数a, b 使

$$P\{Y = aX + b\} = 1$$

#### 不相关和独立的关系:

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{D}(X)}\sqrt{\text{D}(Y)}}$$

独立  $\longrightarrow$   $Cov(X,Y) = 0 \longrightarrow \rho_{XY} = 0 \longrightarrow$  不相关

**不相关 → 独立** (推不出,反例 P108 例 1)

**例题:**设(*X,Y*)的概率密度为 $f(x,y) = \begin{cases} 12y^2, 0 \le y \le x \le 1 \\ 0, else \end{cases}$ 

求 $\rho_{XY}$ 。(包含了前面讲的所有的数字特征,课后自己练习)

P108 例 1 (离散型的)

# (原点) 矩 $\mu_k$ (主要在第七章矩估计中使用)

定义:设X和Y是随机变量,若 $E(X^k)$ , k=1,2,...存在,称之为X的k阶原点矩(k阶矩),记成 $\mu_k=E(X^k)$ 。

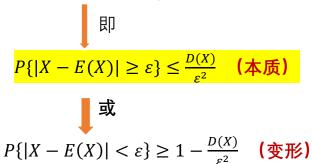
注:所有的数字特征本质就是计算期望。

# Chapter5 大数定理及中心极限定理

(切比雪夫不等式挪到本章)

#### 1 切比雪夫不等式:(I)

设随机变量X具有数学期望 $E(X)=\mu$ ,方差 $D(X)=\sigma^2$ ,则对于任意正数 $\epsilon$ ,不等式 $P\{|X-\mu|\geq \epsilon\}\leq \frac{\sigma^2}{\epsilon^2}$  成立。



特点:事件中的不等号与结论不等号反相时,第一反应切比 雪夫不等式。

# 切比雪夫不等式作用:1.估算概率;2.证明大数定律

2 大数定律:了解本质即可: $\bar{X} \stackrel{P}{\rightarrow} \mu$ ,  $f_n(A) \stackrel{P}{\rightarrow} \mu$ 

推广:  $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k \to \mu_k = E(X^k), n \to \infty$  (矩估计理论基础)

#### 3 中心极限定理(定理1和3)(用标准正态分布估算概率)(I)(D)

#### 独立同分布中心极限定理:

设 $X_1, X_2, \dots$ 是相互独立, 服从同一分布, 且具有数学期望和方差:

$$E(X_k) = \mu$$
,  $D(X_k) = \sigma^2 > 0$  ( $k = 1, 2, ...$ ), 前n 个随机变量之和

$$\sum_{k=1}^{n} X_k \sim N(n\mu, n\sigma^2) \quad (n \to \infty)$$

或 
$$\sum_{k=1}^{n} X_k$$
的标准化变量  $\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \sim N(\mathbf{0}, \mathbf{1})$   $(n \to \infty)$ 

或

 $\bar{X} \sim N(\mu, \sigma^2/n)$   $(n \to \infty)$ 

或

 $\overline{X}$ 的标准化变量  $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$   $(n \to \infty)$ 

主要用于求关于和或者算数平均的概率

# 棣 (dì) 莫弗—拉普拉斯定理:

设随机变量 $\eta_n \sim b(n,p)(0 ,则$ 

$$\frac{\eta_n - np}{\sqrt{np(1-p)}} \sim N(o,1), \quad n \to \infty$$

独立同分布中心极限定理的特殊情况,主要功能是计数

**例 1:**一加法器同时收到 20 个噪声电压 $X_i$  (i = 1, 2, ..., 20). 设它们 相互独立的随机变量,且都在区间(0,10)上服从均匀分布。求  $P\{\sum_{i=1}^{20} X_i > 105\}$ 的近似值(结果用 $\phi(\cdot)$ 表示)(求和问题)

相当于中心极限定理的标志性暗示

**例 2:**据调查某高校的大学生戴眼镜的概率为 0.2, 利用中心极限定理估计 10000 个学生中戴眼镜的学生数少于 2100 概率。(结果用 **()**表示) **(计数问题)** 

中心极限定理使用注意:1、结果近似值;

- 2、数据不要改动;
- 3、看清是求和还是记数

Chapter6 样本及抽样分布 (统计部分的根,特别重要)

# 重点:4个统计量+3大分布+2个定理

4 个统计量(样本数字特征)

设 $X_1, X_2, ..., X_n$ 是**来自**总体X的一个样本。

样本平均值 
$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

样本方差

$$S^{2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{i} - \overline{X})^{2} = \frac{1}{n-1} (\sum_{i=1}^{n} X_{i}^{2} - n \, \overline{X}^{2})$$

样本标准差

$$S = \sqrt{S^2} = \sqrt{\frac{1}{n-1} (\sum_{i=1}^n X_i^2 - n \, \overline{X}^2)}$$

样本k 阶

$$A_k = \frac{1}{n} \sum_{i=1}^n X_i^k$$
,  $k = 1, 2, ...$ 

注意: 联系随机变量的数字特征, 故也可称之为样本数字特征。

3 大统计量分布:

(一) χ<sup>2</sup>分布

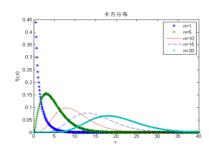
独立性

定义:  $\partial X_1, X_2, ..., X_n$  是来自总体N(0, 1) 的样本,则统计量

$$\chi^2 = \sum_{i=1}^n x_i^2$$

服从自由度为n的 $\chi^2$ 分布,记为 $\chi^2 \sim \chi^2(n)$ .

(密度函数代表曲线)



**性质(可加性):** 设 $\chi_1^2 \sim \chi^2(n_1)$ ,  $\chi_2^2 \sim \chi^2(n_2)$ , 且 $\chi_1^2$ 、 $\chi_2^2$ 相互独立,

则: $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$  (显而易见)

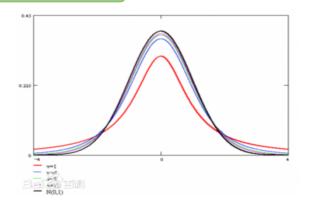
# (二) t 分布(Student 分布)

定义:设 $X \sim N(0,1)$ ,  $Y \sim \chi^2(n)$ , **且**X、Y相互独立,则统计量

$$t = \frac{X}{\sqrt{Y/n}}$$

服从自由度为n的t 分布, 记成 $t \sim t(n)$ .

$$\lim_{n\to\infty}f(x)=\varphi(x)$$



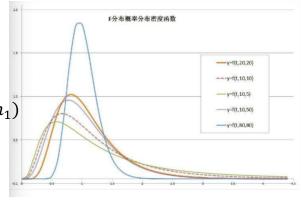
#### (三) F 分布

设 $U\sim\chi^2(n_1),\ V\sim\chi^2(n_2),$  且U、V相互独立,则统计量  $F=\frac{U/n_1}{V/n_2}$ 

服从自由度为 $(n_1, n_2)$ 的F分布,记成 $F \sim F(n_1, n_2)$ .

**密度函数图像**如右图

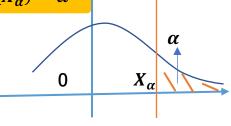
性质: 若 $F \sim F(n_1, n_2)$ , 则 $1/F \sim F(n_2, n_1)$ 



**分位点:**设X的密度函数为f(x), 给定 $\alpha$  (0 <  $\alpha$  < 1), 称满足

 $P\{X > X_{\alpha}\} = \int_{X_{\alpha}}^{\infty} f(x) dx = 1 - F(X_{\alpha}) = \alpha$ 

称 $X_{\alpha}$ 为该种分布的上 $\alpha$ 分位点



#### 2 个定理:

# 正态分布的样本均值和样本方差的分布())

 $N(\mu,\sigma^2)$ 的样本, $\overline{X}$ , $S^2$ 分别表示样本均值和样本方差,则有:

(1) 
$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$
;

(1) 
$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$
; (2)  $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ ;

(3) 
$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$
; (4)  $\overline{X}$ 与 $S^2$ 相互独立。

### 注意结论 对比联系

**定理 1 (二个正态总体)**: 设 $X_1, X_2, ..., X_{n_1}$ 与 $Y_1, Y_2, ..., Y_{n_2}$ 分别是

来自正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的样本,<mark>且这两个样本相互独立。</mark>

 $\overline{X}$  ,  $\overline{Y}$  ,  $S_1^2$ ,  $S_2^2$ 分别表示两个样本均值和样本方差,则:

(1) 
$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1) ;$$

其中
$$S_{\omega} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$
;

(3) 
$$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1-1,n_2-1)_{\circ}$$

# Chapter7参数估计

参数估计

点估计(矩估计&最大似然估计)

评选标准 (无偏性&有效性)

区间估计(双侧&单侧置信区间)

1点估计(矩估计&最大似然估计)

矩估计原理:  $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k \approx \mu_k = E(X^k)$  (n充分大时)

方法: "三步曲"

根据参数个数建立方程(组)→→解方程 ——戴帽子(加尖角符号)

注意:结果用样本表示,看清是量还是值

最大似然估计:似然函数(以一个参数为例)(I)(D)

若总体 X的概率密度函数为 $f(x;\theta),x \in R$ ,则似然函数

$$\mathbf{L}(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

若总体 X的分布律为 $P\{X = x_i\} = p(x_i; \theta)(i = 1, 2, ...n)$ ,则似然函数

$$\mathbf{L}(\theta) = \prod_{i=1}^{n} p(x_i; \theta)$$

方法 (一般情况): "三步曲"

构建似然函数或对数似然函数 求驻点(导数=0 或偏导数=0) 戴帽子(加尖角符合)

#### 2 评选标准 (无偏性&有效性)

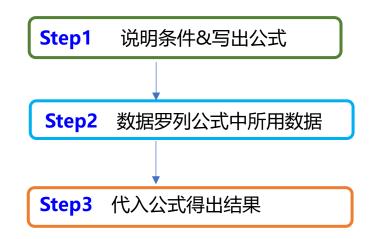
无偏性: $E(\widehat{\theta}) = \theta$  (?)

有效性:无偏的前提下, $D(\widehat{\theta})$ 越小越有效

#### 3区间估计(双侧、单侧)

重点:6种情况的置信区间的结果(I)(D)

计算步骤: "三步曲"



例: 为了比较天能和超威两款电瓶车电瓶寿命情况(单位:h),分别抽取天能和超威两款电瓶 10 只和 8 只新电瓶进行测试,测得寿命的样本均方差分别为 $\mathbf{s}_1^2$ =64, $\mathbf{s}_2^2$ =56。设两款电瓶的寿命分别服从 $\mathbf{N}(\mu_1, \sigma_1^2)$ 和 $\mathbf{N}(\mu_2, \sigma_2^2)$ ,其中 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ 均未知,求两款电瓶的方差比 $\sigma_1^2/\sigma_2^2$ 的置信水平为 95%的置信区间。

 $(F_{0.025}(9,7) = 4.20, F_{0.975}(9,7) = 0.21)$  (结果保留两位小数)

解:

**Step1** 因为 $\mu_1$ ,  $\mu_2$ 未知, 所以关于 $\sigma_1^2/\sigma_2^2$ 的置信水平为 $1-\alpha=$ 

95%的置信区间为: 
$$\left(\frac{S_1^2/S_2^2}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)}, \frac{S_1^2/S_2^2}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$$

**Step2** 
$$S_1^2 = 64$$
,  $S_2^2 = 56$ ,  $n_1 = 10$ ,  $n_2 = 8$ ,  $F_{0.025}(9,7) =$ 

$$4.2F_{0.975}(9,7) = 0.21$$

**Step3** 代入得: 
$$\left(\frac{S_1^2/S_2^2}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)}, \frac{S_1^2/S_2^2}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)}\right) = (0.27, 5.44)$$

# 单侧一般求单侧上、下限

# Chapter8 假设检验

重点:6种情况的检验统计量,拒绝域,接受域与置信区间的

关系 (I) (D)

检验方法: 拒绝域法、置信区间法 (接受域法)

#### 拒绝域法基本步骤:



# 置信区间法基本步骤:



例 2 设某厂所生产的某种细纱每缕支数服从正态分布  $N(\mu, 1.2^2)$ 。 现从该厂某日生产的一批产品中,随机抽 16 缕进行支数测量,求得样本标准差 s=2.1。问当天生产的细纱支数的方差有无显著变化?(已知  $\alpha=0.05$ ,  $\chi^2_{0.025}(15)=27.5$ ,  $\chi^2_{0.025}(16)=28.8$ ,计算结果保留两位小数。)

解:

#### 方法一(置信区间域法)

**Step1:** 根据题意,设  $H_0: \sigma^2 = 1.2^2 = \sigma_0^2$ ,  $H_1: \sigma^2 \neq 1.2^2 = \sigma_0^2$ 

Step2:  $\mu$ 未知 ,  $\therefore$  关于  $\sigma^2$  的置信水平为  $1-\alpha=0.95$  的双侧置信区间为:

$$(\frac{(n-1) S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1) S^2}{\chi^2_{1-\alpha/2}(n-1)})$$

**Step3:**  $\nabla n = 16, s = 2.1, \chi^2_{0.025}(15) = 27.5, \chi^2_{0.975}(15) = 6.3$ 

$$\therefore (\frac{(n-1) S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1) S^2}{\chi_{1-\alpha/2}^2(n-1)}) = (2.41, 10.5)$$

$$\sigma_0^2 = 1.2^2 = 1.44 \notin (2.41, 10.5)$$

Step4:  $H_1$ 为真,即有明显变化(不能少)

# 方法二(拒绝域法):

**Step1:** 根据题意,设  $H_0: \sigma^2 = 1.2^2 = \sigma_0^2$ ,  $H_1: \sigma^2 \neq 1.2^2 = \sigma_0^2$ 

**Step2:**  $\mu$ 未知, 选取检验统计量  $\chi^2 = \frac{(n-1) S^2}{\sigma_0^2}$ , 则拒绝域为:

$$\chi^{2} = \frac{(n-1) S^{2}}{\sigma_{0}^{2}} \ge \chi_{\alpha/2}^{2}(n-1) \quad \overrightarrow{\mathbb{E}}\chi^{2} = \frac{(n-1) S^{2}}{\sigma_{0}^{2}} \le \chi_{1-\alpha/2}^{2}(n-1)$$

Step3: 又 
$$n=16, s=2.1, \sigma_0^2=1.2^2$$
, 得  $\chi^2=\frac{(n-1) S^2}{\sigma_0^2}=45.94 \ge$ 

 $\chi^2_{0.025}(15) = 27.5$ ,落在拒绝域内。

Step4:  $H_1$ 为真,即有明显变化(不能少)

	对象	条件	枢轴量及分布	置信区间	被择假设 <i>H</i> 0	检验统计量	拒绝域
	μ	σ² 已知	$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$(\bar{X} \pm \frac{\sigma}{\sqrt{n}} Z_{\alpha/2})$	$\mu > \mu_0$ (右) $\mu < \mu_0$ (左) $\mu \neq \mu_0$	$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$	$z \ge Z_{\alpha}$ $z \le -Z_{\alpha}$ $ z  \ge Z_{\alpha/2}$
个正态	μ	σ² 未知	$t = \frac{\bar{X} - \mu}{s / \sqrt{n}} \sim t(n - 1)$	$(\bar{X} \pm \frac{s}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1))$	$\mu > \mu_0$ (右) $\mu < \mu_0$ (左) $\mu \neq \mu_0$	$t = \frac{\bar{X} - \mu_0}{s / \sqrt{n}}$	$t \ge t_{\alpha}(n-1)$ $t \le -t_{\alpha}(n-1)$ $ t  \ge t_{\alpha/2}(n-1)$
	$\sigma^2$	<b>μ</b> 未知	$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)})$	$\sigma^2 > \sigma_0^2$ (右) $\sigma^2 < \sigma_0^2$ (左) $\sigma^2 \neq \sigma_0^2$	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$	$\chi^{2} \ge \chi_{a}^{2} (n-1)$ $\chi^{2} \le \chi_{1-a}^{2} (n-1)$ $\chi^{2} \le \chi_{a/2}^{2} (n-1)$ $\chi^{2} \ge \chi_{1-a/2}^{2} (n-1)$
两	$\mu_1 - \mu_2$	$\sigma_1^2$ , $\sigma_2^2$ 已知	$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$	$((\bar{X} - \bar{Y}) \pm \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} Z_{\alpha/2})$	$\mu_1 - \mu_2 > \delta  (右)$ $\mu_1 - \mu_2 < \delta  (左)$ $\mu_1 - \mu_2 \neq \delta$	$Z = \frac{(\bar{X} - \bar{Y}) - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	,
个正态	$\mu_1 - \mu_2$	$\sigma_1^2$ , $\sigma_2^2$ 未知	$t = \frac{(\bar{X} - \bar{Y}) - \mu_1 - \mu_2}{S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$	$((\bar{X} - \bar{Y}) \pm S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} Z_{\alpha/2})$	$\mu_1 - \mu_2 > \delta  (右)$ $\mu_1 - \mu_2 < \delta  (左)$ $\mu_1 - \mu_2 \neq \delta$	$t = \frac{(\overline{X} - \overline{Y}) - \delta}{S_{\omega} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$t \ge t_{\alpha}(n_1 + n_2 - 2)$ $t \le -t_{\alpha}(n_1 + n_2 - 2)$ $ t  \ge t_{\alpha/2}(n_1 + n_2 - 2)$
体	$\frac{\sigma_1^2}{\sigma_2^2}$	<b>μ<sub>1</sub>,μ<sub>2</sub> 未</b> 知	$F = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$	$(\frac{S_1^2/S_2^2}{\frac{F_{\alpha}(n_1-1,n_2-1)}{2}}, \frac{S_1^2/S_2^2}{F_{1-\alpha/2}(n_1-1,n_2-1)})$	$\sigma_1^2 > \sigma_2^2$ (右) $\sigma_1^2 < \sigma_2^2$ (左) $\sigma_1^2 / \sigma_2^2 \neq 1$ $\sigma_1^2 \neq \sigma_2^2$	$F = S_1^2/S_2^2$	$F \ge F_{\alpha}(n_1 - 1, n_2 - 1)$ $F \le F_{1-\alpha}(n_1 - 1, n_2 - 1)$ $F \le F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)$ $F \ge F_{\alpha/2}(n_1 - 1, n_2 - 1)$

备注:(I)(D)(使用前先与课本核对,是否有误)