复习答疑课

考试说明

- 考试时间：未定
- 考试地点：
- 考试内容：1-6章
- 考试题型：
  - 填空(5)+选择(5)+计算(5)+应用(3)
- 成绩计算：
  - 平时成绩30% + 期末考试70%
- 可以带非记忆性计算器

第六章 参数估计和假设检验

- 点估计
  - 矩估计
  - 极大似然估计
  - 评价标准
- 区间估计
  - 一个正态总体的区间估计
- 假设检验
  - 一个正态总体的参数检验

2. 矩估计法

用相应的样本矩去估计总体矩

$$\begin{cases} A_1 = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} = E(X) = \mu \\ A_2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2 \end{cases}$$

$$\begin{cases} \hat{\mu} = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \\ \hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2 = S_n^2 \end{cases}$$

求极大似然估计的一般步骤

(1) 构造似然函数 $L(\theta)$

$$L(\theta) = L(x_1, \Lambda\ , x_n; \theta) = \prod_{i=1}^{n} p(x_i; \theta)$$

$$L(\theta) = L(x_1, \Lambda\ , x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

(2) 求似然函数 $L(\theta)$ 的最大值点

$$\frac{d}{d\theta} \ln L(\theta) = 0.$$

习题六

2. 设总体 $X \sim U(a, b)$，a, b 未知，求(1)参数a, b 的矩估计量.

解 由于$E(X) = \dfrac{a+b}{2}$，$D(X) = \dfrac{(b-a)^2}{12}$

$$E(X^2) = D(X) + E^2(X) = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2$$

$$\begin{cases} A_1 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i = \bar{X} = E(X) = \mu \\ A_2 = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2 \end{cases}$$

解得 $\hat{a}_{矩} = \bar{X} - \sqrt{3(A_2 - \bar{X}^2)}$

$$= \bar{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2},$$

$$\hat{b}_{矩} = \bar{X} + \sqrt{3(A_2 - \bar{X}^2)}$$

$$= \bar{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

或者

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = S_n^2$$

$$\overline{X} = \frac{a+b}{2}, \quad S_n^2 = \frac{(b-a)^2}{12}$$

$$\hat{a}_{矩} = \overline{X} - \sqrt{3}S_n \qquad \hat{b}_{矩} = \overline{X} + \sqrt{3}S_n$$

2.(2) 设$X \sim U[a,b]$;a,b未知，$x_1,\Lambda\ ,x_n$是一个样本值，

求：a,b的极大似然估计量。

解：X的概率密度为：

$$f(x;a,b)=\begin{cases}\dfrac{1}{b-a}, a\le x\le b;\\ 0 \quad ,其它\end{cases}$$

似然函数为

$$L(a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, a\le x_{(1)}, b\ge x_{(n)};\\ 0 \quad ,其它\end{cases}$$

似然函数只有当 $a < x_i < b, i = 1,2,\ldots, n$ 时才能获得最大值, 且 a 越大, b 越小, L 越大.

令
$$x_{\min} = \min\{x_1, x_2, \ldots, x_n\}$$
$$x_{\max} = \max\{x_1, x_2, \ldots, x_n\}$$

取
$$\hat{a} = x_{\min}, \quad \hat{b} = x_{\max}$$

则对满足 $a \le x_{\min} \le x_{\max} \le b$ 的一切 $a < b$,

都有
$$\frac{1}{(b-a)^n} \le \frac{1}{(x_{\max} - x_{\min})^n}$$

故 $\hat{a} = x_{\min}, \ \hat{b} = x_{\max}$

是 a , b 的极大似然估计值.

$$X_{\min} = \min\{X_1, X_2, \Lambda \ , X_n\}$$

$$X_{\max} = \max\{X_1, X_2, \Lambda \ , X_n\}$$

分别是 a , b 的极大似然估计量.

4.

设$X\sim G(p)$, $x_1,\cdots,x_n$是来自X的一个样本值，试求参数p与EX的极大似然估计.

解：X的分布律为：

$$P(X=k)=p(1-p)^{k-1},\quad k=1,2,\cdots$$

故似然函数为

$$L(p)=\prod_{i=1}^{n}p(1-p)^{x_i-1}=p^n(1-p)^{\sum\limits_{i=1}^{n}x_i-n},$$

令
$$\frac{d}{dp}\ln L(p)=\frac{n}{p}-\frac{\sum\limits_{i=1}^{n}x_i-n}{1-p}=0.$$

解得p的极大似然估计值

$$\hat{p}=\frac{\mathrm{n}}{\sum_{i=1}^{n}x_i}=\frac{1}{\overline{x}}$$

因为$EX=\dfrac{1}{p}$，故EX的极大似然估计为

$$E\hat{X}=\frac{1}{\hat{p}}=\overline{x}$$

点估计的评价标准

(1) 无偏性

设 $\hat{\theta}(X_1,\Lambda\ ,X_n)$是未知参数$\boldsymbol{\theta}$的估计量，若

$$E(\hat{\theta})=\theta$$

则称 $\hat{\boldsymbol{\theta}}$ 为$\boldsymbol{\theta}$ 的无偏估计 .

(2) 有效性

设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是参数$\boldsymbol{\theta}$的无偏估计量，若有

$$\boldsymbol{D}(\hat{\boldsymbol{\theta}}_1)<\boldsymbol{D}(\hat{\boldsymbol{\theta}}_2)$$

则称 $\hat{\boldsymbol{\theta}}_1$ 较 $\hat{\boldsymbol{\theta}}_2$ 有效 .

5. 设总体 X 的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0,\\ 0 & x\le 0\end{cases}$$ $\theta>0$ 为参数

$X_1,X_2,\Lambda\ ,X_n$ 为 X 的一个样本.

求 θ 的极大似然估计量, 并判断它是否无偏估计量.

解 由似然函数 $L(\theta)=\dfrac{1}{\theta^n}e^{-\frac{1}{\theta}\sum\limits_{i=1}^{n}x_i}$, $x_i>0$

$$\longrightarrow\quad \ln L(\theta)=-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}x_i$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta) = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2} \overset{令}{=} 0$$

$$\longrightarrow \quad \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

θ的极大似然估计量为 $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$

$$\Theta\ X \sim E\left(\frac{1}{\theta}\right) \qquad \therefore E(X) = \theta$$

故 $E(\hat{\theta}) = E(\bar{X}) = E(X) = \theta$

它是θ的无偏估计量.

区间估计

1. 置信区间定义

设 θ 是一个待估参数，给定 $\alpha > 0$,

若由样本 $X_1, X_2, \ldots X_n$ 确定的两个统计量

$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \Lambda\ , X_n), \hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \Lambda\ , X_n)$

$(\hat{\theta}_1 < \hat{\theta}_2)$ 满足

$$P(\hat{\theta}_1 \le \theta \le \hat{\theta}_2) = 1 - \alpha$$

则称区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 是 θ 的置信水平（置信度、置信概率）为 $1-\alpha$ 的置信区间.

求置信区间的一般步骤如下:

1. 明确问题, 是求什么参数的置信区间?

置信水平 $1-\alpha$ 是多少?

2. 寻找参数 θ 的一个良好的点估计

$$T(X_1, X_2, \ldots X_n)$$

3. 寻找一个待估参数 θ 和估计量 T 的函数 $S(T, \theta)$,且其分布为已知.

称 $S(T, \theta)$ 为枢轴量.

4. 对于给定的置信水平$1-\alpha$，根据$S(T,\theta)$的分布，确定常数a, b，使得

$$P(a \le S(T,\theta) \le b) = 1-\alpha$$

5. 对"$a \le S(T,\theta) \le b$"作等价变形,得到如下形式:

$$P\{\hat{\theta}_1 \le \theta \le \hat{\theta}_2\} = 1-\alpha$$

则$[\hat{\theta}_1, \hat{\theta}_2]$就是$\theta$的置信度为$1-\alpha$的置信区间.

置信区间常用公式

一个正态总体 $X \sim N(\mu, \sigma^2)$的情形

(1) 方差σ^2已知, μ的置信区间

$$\left[\bar{X} - \frac{\sigma}{\sqrt{n}} u_{\alpha/2},\ \bar{X} + \frac{\sigma}{\sqrt{n}} u_{\alpha/2}\right]$$

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

(2) 方差σ^2未知, μ的置信区间

$$(4)\ \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

$$\left[\bar{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1),\ \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right]$$

(3) 当 μ 已知时，方差 σ^2 的置信区间

取枢轴量 $Q=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$，

$\frac{X_i-\mu}{\sigma} : N(0,1)$

$$P\left(\chi^2_{1-\frac{\alpha}{2}}(n)<\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}<\chi^2_{\frac{\alpha}{2}}(n)\right)=1-\alpha$$

得 σ^2 的置信度为 $1-\alpha$ 置信区间为

$$\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)},\ \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right)$$

(4) 当μ未知时，方差σ^2的置信区间

选取 $K=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$

$$P\left(\chi^2_{1-\frac{\alpha}{2}}<\frac{(n-1)S^2}{\sigma^2}<\chi^2_{\frac{\alpha}{2}}\right)=1-\alpha$$

得 σ^2的置信区间为

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

15 设 $X \sim N(\mu, \sigma_0^2)$, σ_0^2 已知，样本容量 n 多大，才能使 μ 的置信度 $1-\alpha$ 的置信区间长度不大于常数 d？

$$[\bar{X} - \frac{\sigma_0}{\sqrt{n}} u_{\alpha/2},\ \bar{X} + \frac{\sigma_0}{\sqrt{n}} u_{\alpha/2}]$$

区间的长度为 $2u_{\frac{\alpha}{2}} \dfrac{\sigma_0}{\sqrt{n}} \le d$

$$n \ge (2u_{\frac{\alpha}{2}} \frac{\sigma_0}{d})^2$$

16. 随机地取某种炮弹9发做试验，得炮弹口速度的样本标准差为10.5(m/s). 设炮口速度服从正态分布. 求这种炮弹的炮口速度的标准差σ的置信度为0.95的置信区间.

解 σ^2的置信区间为

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

σ 的置信区间为

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

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

σ的置信度为0.95的置信区间为

$$\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}}, \sqrt{\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}\right) = \left(\frac{\sqrt{8}\times 10.5}{\sqrt{17.535}}, \frac{\sqrt{8}\times 10.5}{\sqrt{2.18}}\right)$$

$$= (7.1, 20.1)$$

其中α=0.05, n=9 查表知

$$\chi^2_{0.025}(8) = 17.535,\ \chi^2_{0.975}(8) = 2.180$$

假设检验步骤

(1) 建立假设

(2) 假设 H_0 为真,选择统计量

(3) 确定拒绝域

(4) 作出判断

U 检验法 (σ^2 已知)

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$U = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$	$\lvert U \rvert \geq u_{\frac{\alpha}{2}}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$U \leq -u_\alpha$
$\mu \leq \mu_0$	$\mu > \mu_0$		$U \geq u_\alpha$

T 检验法 (σ^2 未知)

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$	$\lvert T \rvert \geq t_{\frac{\alpha}{2}}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$T \leq -t_\alpha$
$\mu \leq \mu_0$	$\mu > \mu_0$		$T \geq t_\alpha$

（2）关于 σ^2 的检验　χ^2**检验法**(μ已知)

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n)$	$\chi^2\le\chi^2_{1-\frac{\alpha}{2}}(n)$ 或 $\chi^2\ge\chi^2_{\frac{\alpha}{2}}(n)$
$\sigma^2\ge\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\le\chi^2_{1-\alpha}(n)$
$\sigma^2\le\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\ge\chi^2_{\alpha}(n)$

χ^2检验法(μ 未知)

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$\chi^2\leq\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 或 $\chi^2\geq\chi^2_{\frac{\alpha}{2}}(n-1)$
$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\leq\chi^2_{1-\alpha}(n-1)$
$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\geq\chi^2_{\alpha}(n-1)$

14.[十三]　某种导线，要求其电阻的标准差不得超过 0.005(欧姆)。今在生产的一批导线中取样品 9 根，测得 s=0.007(欧姆)，设总体为正态分布。问在水平 $\alpha = 0.05$ 能否认为这批导线的标准差显著地偏大？

解：（1）提出 H_0：$\sigma \le 0.005$；H_1：$\sigma > 0.005$

（2）H_0 的拒绝域为 $(n-1)S^2 / 0.005^2 \ge \chi^2_{\alpha}(n-1)$

（3）n=9，$\alpha = 0.05$，S=0.007，由计算知

$$(n-1)S^2 / 0.005^2 = \frac{8 \times 0.007^2}{0.005^2} = 15.68 > \chi^2_{\alpha}(n-1)$$

查表 $\chi^2_{0.05}(8) = 15.507$

（4）故在 $\alpha = 0.05$ 下，拒绝 H_0，认为这批导线的标准差显著地偏大。

第一章 随机事件及其概率

- 概率的基本运算法则
- 全概率公式
- 贝叶斯公式

运算律

事件运算 $\xleftrightarrow{\text{对应}}$ 集合运算

- 吸收律　$A \cup \Omega = \Omega$　　$A \cap \Omega = A$

  $A \cup \varnothing = A$　　$A \cap \varnothing = \varnothing$

  $A \cup (AB) = A$　　$A \cap (A \cup B) = A$

- 重余律　$\overline{\overline{A}} = A$

- 幂等律　$A \cup A = A$　　$A \cap A = A$

- 差化积　$A - B = A\overline{B} = A - (AB)$

- 交换律 $A\cup B = B\cup A \qquad AB = BA$
- 结合律 $(A\cup B)\cup C = A\cup(B\cup C)$

  $(AB)C = A(BC)$
- 分配律 $(A\cup B)\cap C = (A\cap C)\cup(B\cap C)$

  $A\cup(BC) = (A\cup B)(A\cup C)$
- 反演律 $\overline{A\cup B} = \overline{A}\,\overline{B} \qquad \overline{AB} = \overline{A}\cup\overline{B}$

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

运算顺序： **逆交并差，括号优先**

1. 概率的性质

三条公理：

- 非负性： $P(A)\geq 0$
- 规范性： $P(\Omega)=1$

$P(\varnothing)=0$

基本性质

- 可列可加性： $P\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i)$

其中 $A_1, A_2, \cdots$ 为两两互斥事件，

加法公式

性质1 加法公式

若事件A, B互斥，则

$$P(A+B)=P(A)+P(B)$$

若事件$A_1, A_2, \cdots, A_n$两两互斥，则

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

性质4 广义加法公式

对任意两个事件A、B，有

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

性质2　逆事件公式

对任一事件A，有　$P(\overline{A}) = 1 - P(A)$

性质3　减法公式

设 A、B是两个事件，若 $A \subset B$，则有

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

$$P(B) \geq P(A)$$

- 对任意两个事件A, B, 有

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

推广:

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$$

一般:

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

右端共有 2^n-1 项.

定义 设A、B为两事件, $P(A)>0$, 则称

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

为事件 A 发生的条件下事件 B 发生的条件概率.

同理 $$P(A|B)=\frac{P(AB)}{P(B)}$$

称为在事件B发生的条件下事件A的条件概率.

性质 条件概率也是概率，故具有概率的性质：

- 非负性 $P(B|A) \ge 0$
- 规范性 $P(\Omega|A) = 1$
- 可列可加性 $P\left(\bigcup_{i=1}^{\infty} B_i \middle| A\right) = \sum_{i=1}^{\infty} P(B_i|A)$

概率的一些重要性质都适用于条件概率. 例如:

- $P(B_1 \cup B_2 | A) = P(B_1 | A) + P(B_2 | A) - P(B_1 B_2 | A)$
- $P(\overline{B}|A) = 1 - P(B|A)$
- $P(B_1 - B_2 | A) = P(B_1 | A) - P(B_1 B_2 | A)$

(2) 乘法公式

利用条件概率求积事件的概率即**乘法公式**

$$P(AB) = P(A)P\left(B \middle| A\right) \quad (P(A) > 0)$$

$$P(AB) = P(B)P\left(A \middle| B\right) \quad (P(B) > 0)$$

推广

$$P(A_1A_2\cdots A_n) = P(A_1)P\left(A_2 \mid A_1\right)\cdots P\left(A_n \mid A_1A_2\cdots A_{n-1}\right)$$

$$(P(A_1A_2\cdots A_{n-1}) > 0)$$

两事件独立的定义

定义 设 A , B 为两事件，若

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

则称事件 A 与事件 B 相互独立

❑ 四对事件 A, B; $A, \overline{B}$; $\overline{A}, B$; $\overline{A}, \overline{B}$
任何一对相互独立,则其它三对也相互独立

两事件相互独立的性质

- 若 $P(A)>0$, 则$P(B)=P(B|A)$

  若 $P(B)>0$, 则$P(A)=P(A|B)$

- 若 $P(A)>0$, $P(B)>0$,

  则 “事件 A 与 事件 B 相互独立” 和 “事件 A 与 事件 B 互斥”

  不能同时成立.

(2). 多个事件的独立性

将两事件独立的定义推广到三个事件：

定义 对于三个事件A、B、C，若

$$\begin{cases} P(AB)=P(A)P(B) \\ P(AC)=P(A)P(C) \\ P(BC)=P(B)P(C) \\ P(ABC)=P(A)P(B)P(C) \end{cases}$$

四个等式同时成立,则称事件A、B、C相互独立.

n个独立事件和的概率公式:

设事件 $A_1, A_2, \cdots, A_n$ 相互独立, 则

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - P(\overline{A_1}\,\overline{A_2} \cdots \overline{A_n})$$

$$= 1 - P(\overline{A_1})P(\overline{A_2}) \cdots P(\overline{A_n})$$

$\overline{A_1}, \overline{A_2}, \cdots, \overline{A_n}$

也相互独立

也就是说，n个**独立**事件至少有一个发生的概率等于1减去各自对立事件概率的乘积.

全概率公式

设S为随机试验的样本空间，$A_1,A_2,\ldots,A_n$是两两互斥的事件，且有$P(A_i)>0$，$i=1,2,\ldots,n$，$\bigcup_{i=1}^{n} A_i = S$，则对任一事件$B$，有

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

称满足上述条件的$A_1,A_2,\ldots,A_n$为完备事件组.

二. 贝叶斯公式

设$A_1,A_2,\ldots,A_n$是完备事件组，则对任一事件B，有

$$P(A_i \mid B)=\frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)} \qquad i=1,2,\Lambda\ ,n$$

该公式于1763年由贝叶斯(Bayes)给出. 它是在观察到事件B已发生的条件下，寻找导致B发生的每个原因的概率.

第二章 随机变量及其分布

- 随机变量
- 分布函数
- 离散型随机变量及分布
- 连续型随机变量及其分布
- 随机变量的函数

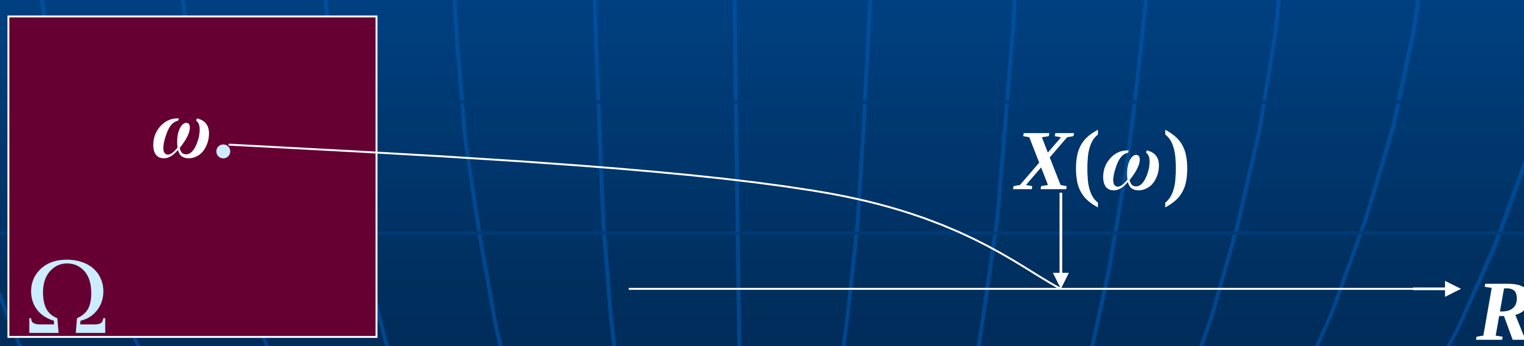
§2.1 随机变量及其分布函数

一.随机变量 (random variable)

定义 设 是试验E的样本空间, 若

$$\forall \omega \in \Omega \xrightarrow{\text{按一定法则}} \exists \text{ 实数 } X(\omega)$$

则称 $X(\omega)$ 为 上的 随机变量 (r.v.)

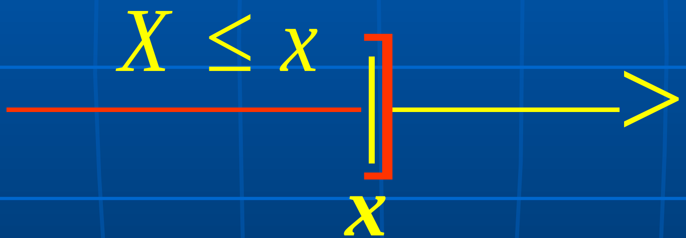


二. 随机变量的分布函数

定义 设 X 为 r.v., x 是任意实数,称函数

$$F(x)=P(X\le x),\quad -\infty<x<+\infty$$

为 X 的分布函数.



如果将 X 看作数轴上随机点的坐标,那么分布函数 $F(x)$ 的值就表示 X落在区间 $(-\infty,x]$ 的概率.

用分布函数计算 X 落在$(a,b]$里的概率:

$$P(a<X\le b)=P(X\le b)-P(X\le a)$$

$$=F(b)-F(a)$$

$$P(X>a)=P(X\ge a)=1-F(a)$$

分布函数的性质

(1) $F(x)$ 单调不减，即

$$\forall\ x_1 < x_2,\ \ F(x_1) \le F(x_2)$$

(2) $0 \le F(x) \le 1$

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

(3) $F(x)$ 右连续，即 $\lim\limits_{\boldsymbol{x} \to \boldsymbol{x}_0^+} \boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x}_0)$

三.离散型随机变量及分布律

定义 若随机变量 X 的可能取值是有限个或可列个, 则称 X 为离散型随机变量

描述X 的概率特性常用概率分布或分布律

即 $$P(X = x_k) = p_k,\ \ k = 1,2,\Lambda$$

或概率分布表

X	x_1	x_2	Λ	x_k	Λ
P	p_1	p_2	Λ	p_k	Λ

或 $$X \sim \begin{pmatrix} x_1 & x_2 & \Lambda & x_k & \Lambda \\ p_1 & p_2 & \Lambda & p_k & \Lambda \end{pmatrix}$$

分布律的性质

- $p_k \geq 0, \ k = 1,2,\Lambda$ ———— 非负性
- $\sum\limits_{k=1}^{\infty} p_k = 1$ ———— 归一性

用性质可以判断是否为分布律

离散型随机变量的分布函数

$$F(x) = P(X \le x) = \sum_{x_k \le x} p_k$$

$$p_k = P(X = x_k) = F(x_k) - F(x_{k-1})$$

其中 $x_{k-1} < x_k$.

$F(x)$ 是分段阶梯函数, 在 X 的可能取值 x_k 处发生间断, 间断点为第一类跳跃间断点,在间断点处有跃度 p_k.

四. 常见离散型随机变量的分布

1.超几何分布

例 设有 N 件产品，其中有 M 件次品，现从中任取 n 件，用 X 表示其中的次品数，求其分布律。

超几何公式

$$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k=0,1,2,\Lambda\ ,l.$$

$$l=\min(M,n)$$

超几何分布

2.几何分布

例 某射手连续向一目标射击，直到命中为止，已知他每发命中率是 p，求所需射击发数 X 的分布律.

$$P(X=k)=(1-p)^{k-1}\cdot p \qquad k=1,2,\Lambda\ \Lambda$$

3. 两点分布（0 – 1 分布）

X	0	1
P_k	$1-p$	p

$0<p<1$

或 $P(X=k)=p^k(1-p)^{1-k},\quad k=0,1$

4. 二项分布

n 重Bernoulli 试验中, X 是事件A 在 n 次试验中发生的次数 , $P(A)=p$,若

$$P_n(k)=P(X=k)=C_n^k p^k (1-p)^{n-k}, \quad k=0,1,\Lambda\ ,n$$

则称 X 服从参数为n, p 的二项分布，记作

$$X \sim B(n,p)$$

二项分布中最可能出现次数

当$(n+1)p=$整数时,在 $k=(n+1)p$与 $(n+1)p-1$ 处的概率取得最大值

当$(n+1)p\neq$整数时, 在 $k=[(n+1)p]$处的概率取得最大值

$[x]$ 表示不超过x 的最大整数

5. 泊松分布

若 $P(X=k)=e^{-\lambda}\dfrac{\lambda^k}{k!},\quad k=0,1,2,\Lambda$

其中 $\lambda>0$ 是常数，则称 X 服从参数为 λ 的泊松 (Poisson) 分布. 记作 $X\sim\pi(\lambda)$ 或 $P(\lambda)$

泊松分布中最可能出现次数

当 $\lambda=$ 整数时,在 λ 与 $\lambda-1$ 处的概率取得最大值

当 λ　　整数时, 在 $[\lambda]$ 处的概率取得最大值

Possion定理

设 $np_n = \lambda > 0$, 则对固定的 k

$$\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$$

$$k = 0,1,2,\Lambda$$

结论 二项分布的极限分布是 Possion 分布

若$X \sim B(n,p)$, 则当n 较大，p 较小, 则

$$C_n^k p^k (1-p)^{n-k} \approx e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0,1,2,\Lambda \qquad np = \lambda$$

$n > 10, p < 0.1$时近似效果较好

一. 连续型随机变量

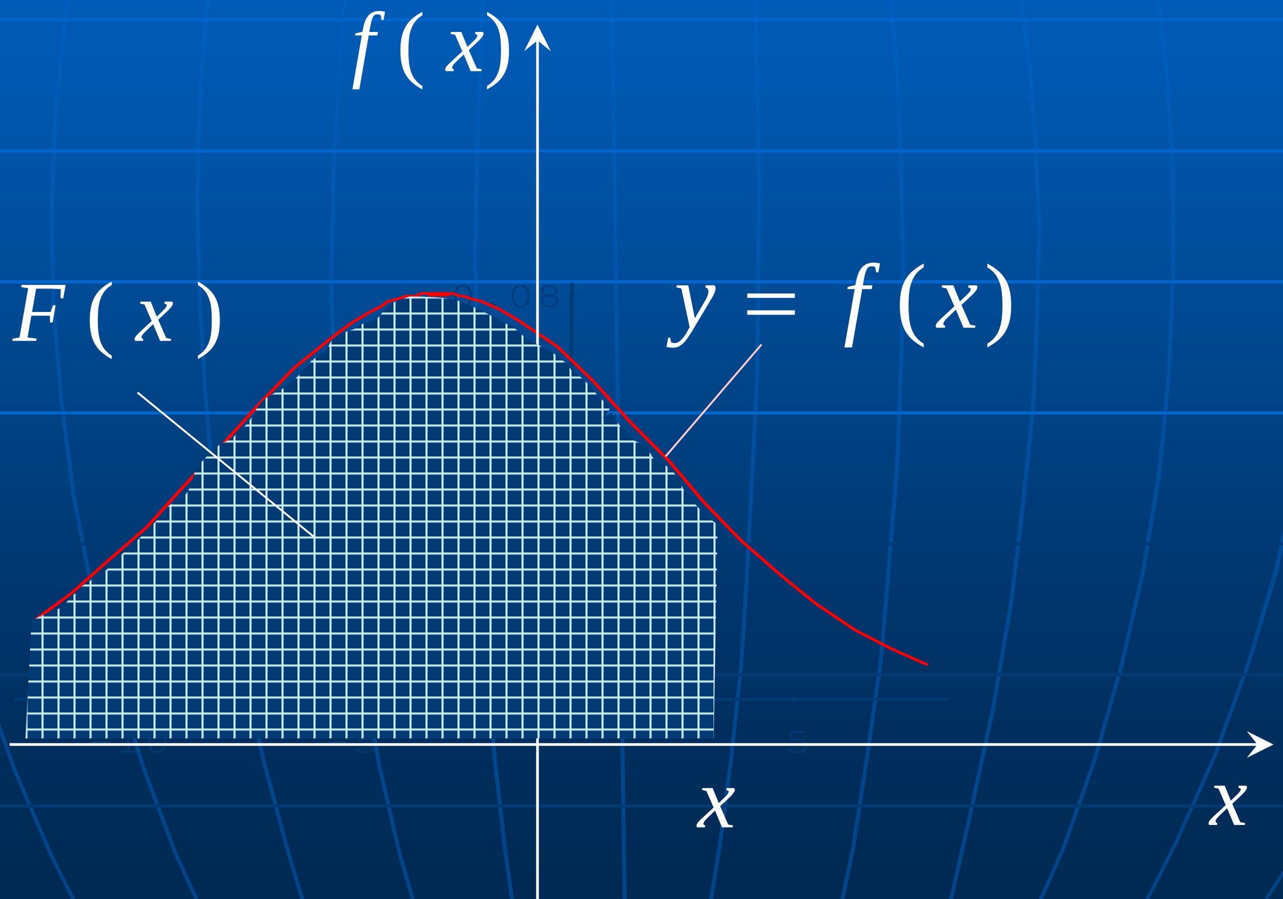
定义 设 X 是随机变量, 若存在一个非负可积函数 $f(x)$, 使得

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t \qquad -\infty < x < +\infty$$

其中$F(x)$是它的分布函数

则称 X 是 连续型 ***r.v.*** ，$f(x)$是它的概率密度函数(**p.d.f.**)，简记为***d.f.***

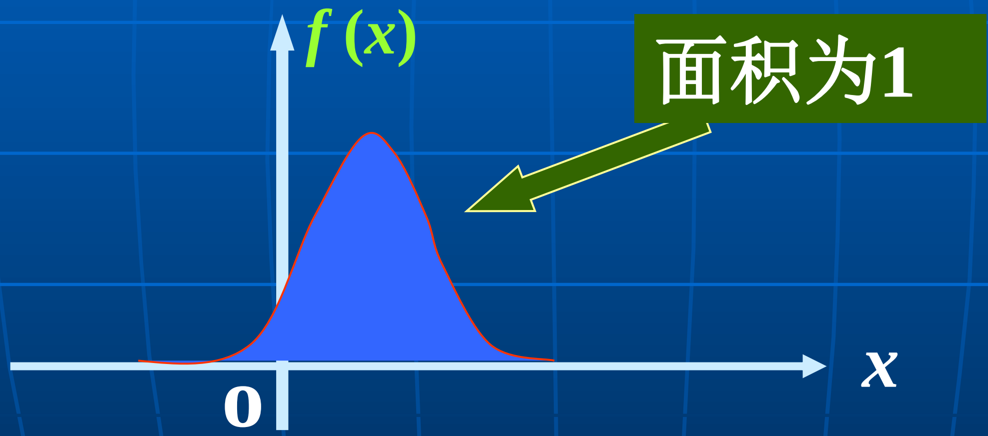
分布函数与密度函数 几何意义



p.d.f. $f(x)$的性质

- $f(x) \geq 0$
- $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = F(+\infty) = 1$

判定函数 $f(x)$是否为*r.vX*的概率密度函数的充要条件.



$$P\{X \in G\} = \int_G f(x)dx$$

- 在 $f(x)$ 的连续点处,

$$f(x) = F'(x)$$

(1) **均匀分布**　若 X 的 $d.f.$ 为

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

则称 X 服从区间(a,b)上的**均匀分布**

记作　　　$X\sim U(a,b)$

X 的分布函数为

$$F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\le x<b,\\ 1 & x\ge b\end{cases}$$

(2) **指数分布**

若 X 的$d.f.$ 为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & 其他\end{cases}\qquad \lambda>0\ 为常数$$

则称 X 服从 参数为 λ 的指数分布

记作 $X\sim E(\lambda)$

X 的分布函数为 $F(x)=\begin{cases}0, & x<0\\ 1-e^{-\lambda x}, & x\geq 0\end{cases}$

(3) 正态分布

若 X 的 $d.f.$ 为

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

μ, σ 为常数, $\sigma > 0$

则称 X 服从参数为 $\quad , \quad {}^2$ 的正态分布

记作 $X \sim N(\quad , \quad {}^2)$

亦称高斯(Gauss)分布

$f(x)$ 的性质：

- 图形关于直线 $x = \mu$ 对称, 即

$$f(\mu + x) = f(\mu - x)$$

在 $x = \mu$ 时, $f(x)$ 取得最大值 $\dfrac{1}{\sqrt{2\pi}\sigma}$

$$P(X \le \mu) = \frac{1}{2}$$

一种重要的正态分布

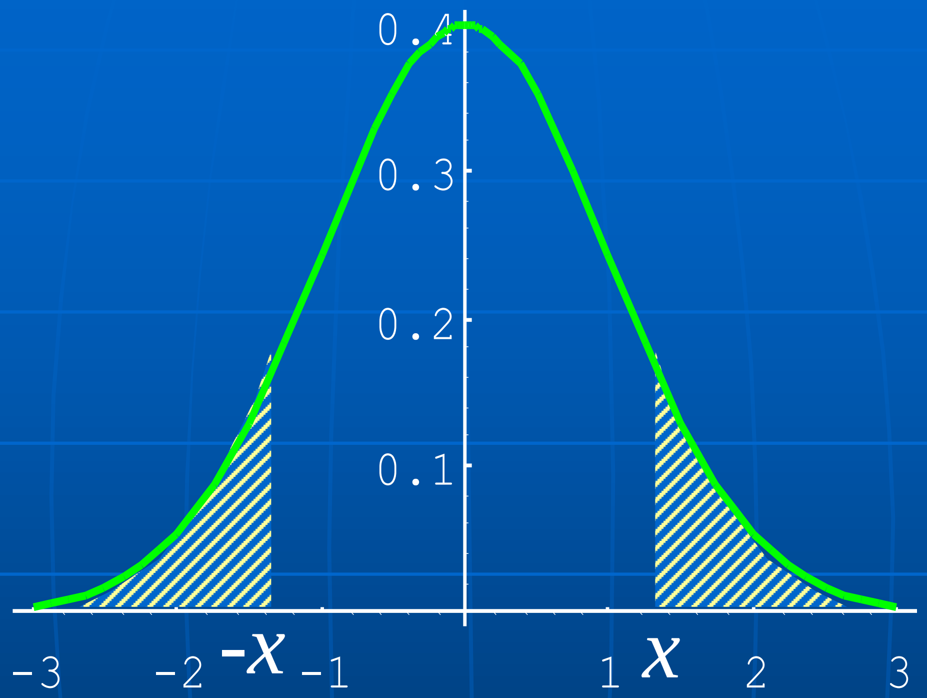
——标准正态分布$N(0,1)$

密度函数 $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \qquad -\infty<x<+\infty$

是偶函数，分布函数记为

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

其值有专门的表供查.



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

$$P(|X| < a) = 2\Phi(a) - 1$$

$$\Phi(0) = 0.5$$

设 $X \sim N(\mu, \sigma^2)$，则 $Y = \dfrac{X-\mu}{\sigma}$ **~N(0,1)**

$$P(a < X < b) = F(b) - F(a)$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(X > a) = 1 - F(a)$$

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

根据定理, 只要将标准正态分布的分布函数制成表，就可以解决一般正态分布的概率计算问题.

离散型随机变量函数的分布

设 r.v. X 的分布律为

$$P(X = x_k) = p_k, \qquad k = 1,2,$$

由已知函数 $g(x)$可求出 r.v. Y 的所有可能取值，则 Y 的概率分布为

$$P(Y = y_i) = \sum_{k:\, g(x_k)=y_i} p_k, \qquad i = 1,2,$$

连续型随机变量函数的分布

已知 X 的d.f. $f(x)$ 或分布函数
求 $Y = g(X)$ 的d.f.

方法:

(1) 从分布函数出发
(2) 用公式直接求d.f.

方法一 从分布函数出发

$$F_Y(y) = P(Y \le y) = P(g(X) \le y)$$

$$f_Y(y) = F_Y'(y)$$

对于连续型随机变量，在求Y=$g(X)$ 的分布时，关键的一步是把事件 $\{ g(X) \le y \}$ 转化为X在一定范围内取值的形式，从而可以利用 X 的分布来求 $P\{ g(X) \le y \}$.

方法二 用公式

定理 设连续型$r.v$ X具有概率密度 $f_X(x)$,

又设$y=g(x)$单调可导，其反函数为$x=g^{-1}(y)$,

则$Y=g(X)$是一个连续型$r.v$，其概率密度为

$$f_Y(y)=\begin{cases} f_X[g^{-1}(y)]\left|\dfrac{dg^{-1}(y)}{dy}\right|, & \alpha<y<\beta \\ 0, & \text{其它} \end{cases}$$

其中，$\alpha=\min\limits_{a\le x\le b} g(x),\ \beta=\max\limits_{a\le x\le b} g(x),$

第三章 多维随机变量及其分布

- 联合分布
- 边缘分布
- 随机变量的独立性
- 二维随机变量函数的分布

二维随机变量的联合分布函数

定义 设(X,Y)为二维 $r.v.$ 对任何一对实数(x,y)，事件

$$(X \le x) \cap (Y \le y) \quad \text{(记为}(X \le x, Y \le y)\text{)}$$

的概率 $P(X \le x, Y \le y)$定义了一个二元实函数 $F(x,y)$，称为二维 $r.v.(X,Y)$ 的分布函数，即

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

联合分布函数的性质

① $0 \le F(x, y) \le 1$

$$F(+\infty, +\infty) = 1$$

$$F(x, -\infty) = 0$$

$$F(-\infty, y) = 0$$

$$F(-\infty, -\infty) = 0$$

② 对每个变量单调不减

固定 x，对任意的 $y_1 < y_2$，

$$F(x, y_1) \le F(x, y_2)$$

固定 y，对任意的 $x_1 < x_2$，

$$F(x_1, y) \le F(x_2, y)$$

③ 对每个变量右连续

$$F(x_0, y_0) = F(x_0 + 0, y_0)$$

$$F(x_0, y_0) = F(x_0, y_0 + 0)$$

联合分布律

设(X,Y)的所有可能的取值为

$$(x_i, y_j), \qquad i, j = 1,2,\Lambda$$

则称

$$P(X = x_i, Y = y_j) = p_{ij}, \quad i, j = 1,2,\Lambda$$

为二维 $r.v.(X,Y)$ 的联合概率分布

也简称 概率分布 或 分布律

显然，$p_{ij} \geq 0, \qquad i, j = 1,2,\Lambda$

$$\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty} p_{ij} = 1$$

已知联合分布律可以求出其联合分布函数

二维离散 *r.v.*的联合分布函数

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

$$-\infty < x \ , \ y < +\infty .$$

§3.1 二维随机变量及其分布

3. 二维连续型随机变量

定义 设二维 $r.v.(X,Y)$ 的分布函数为 $F(x,y)$,若存在非负可积函数 $f(x,y)$,使得对于任意实数 x,y 有

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

则称 (X,Y) 为**二维连续型** $r.v.$

$f(x,y)$ 为 (X,Y) 的**联合概率密度函数**

简称**概率密度函数**简记 $p.d.f.$

联合密度的性质

(1) $f(x,y)\geq 0$

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dydx = 1$

(3) 在 $f(x,y)$ 的连续点处

$$\frac{\partial^2 F}{\partial x \partial y} = f(x,y)$$

(4) 若G 是平面上的区域，则

$$P\big((X,Y)\in G\big) = \iint_G f(x,y)dxdy$$

常用连续型二维随机变量分布

- 区域G上的均匀分布，记作$U(G)$

G是平面上的有界区域, 面积为 A

若$r.v.(X,Y)$的联合 $d.f.$ 为

$$f(x,y)=\begin{cases}1/A, & (x,y)\in G\\ 0, & \text{其他}\end{cases}$$

则称(X,Y)服从区域G上的均匀分布

若(X,Y)服从区域G上的均匀分布,

则 $\forall G_1 \subseteq G$, 设G_1的面积为A_1,

$$P\left((X,Y)\in G_1\right)=\frac{A_1}{A}$$

二维正态分布

若$r.v.(X,Y)$的联合$d.f.$为

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

$$-\infty<x<+\infty,-\infty<y<+\infty$$

则称(X,Y)服从参数为$\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho$的正态分布, 记作$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$

其中$\sigma_1,\sigma_2>0$, $-1<\rho<1$.

二维离散型随机变量的边缘分布

$$P(X=x_i)=\sum_{j=1}^{+\infty}p_{ij}\overset{\text{记作}}{=}p_{i\bullet},\qquad i=1,2,\Lambda$$

$$P(Y=y_j)=\sum_{i=1}^{+\infty}p_{ij}\overset{\text{记作}}{=}p_{\bullet j},\qquad j=1,2,\Lambda$$

由联合分布律可确定边缘分布律

二维连续型随机变量的边缘分布

$$F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(u,v)dvdu$$

$$F_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{+\infty} f(u,v)dudv$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy$$

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

已知联合密度可以求得边缘密度

结 论 （一）

二维正态分布的边缘分布是一维正态分布.

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

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

结 论 （二）

上述的两个边缘分布中的参数与二维正态分布中的常数 ρ 无关.

两随机变量独立的定义是：

设 X,Y是两个$r.v$，若对任意的x,y,有

$$P(X\le x,Y\le y)=P(X\le x)P(Y\le y)$$

则称X,Y相互独立 .

设 X,Y 是两个$r.v$，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 .

离散型

$\boldsymbol{X}$与$\boldsymbol{Y}$ 独立 $\Longleftrightarrow$ 对一切 $\boldsymbol{i}, \boldsymbol{j}$ 有

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

即

$$p_{ij} = p_{i\bullet} \cdot p_{\bullet j}$$

连续型

$\boldsymbol{X}$与$\boldsymbol{Y}$ 独立 $\Longleftrightarrow$ 对任何 x ,y 有

$$f(x,y) = f_X(x) f_Y(y)$$

二维随机变量（$\boldsymbol{X}$, $\boldsymbol{Y}$）相互独立, 则边缘分布完全确定联合分布

离散型二维 *r.v.*的函数

当(X, Y)为离散*r.v.*时, Z 也离散

$$Z = z_k = g(x_i, y_j)$$

$$P(Z = z_k) = \sum_{g(x_i, y_j) = z_k} P(X = x_i, Y = y_j) \quad k = 1, 2, \Lambda$$

二维连续r.v.函数的分布

问题 已知r.v.(X,Y)的密度函数，
求$Z=g(X,Y)$的密度函数.

方法 (1) 先求Z 的分布函数：

$$F_Z(z)=P(Z\le z)=P(g(X,Y)\le z)$$

$$=\iint\limits_{D_z} f(x,y)dxdy \qquad D_z:\{(x,y)\mid g(x,y)\le z\}$$

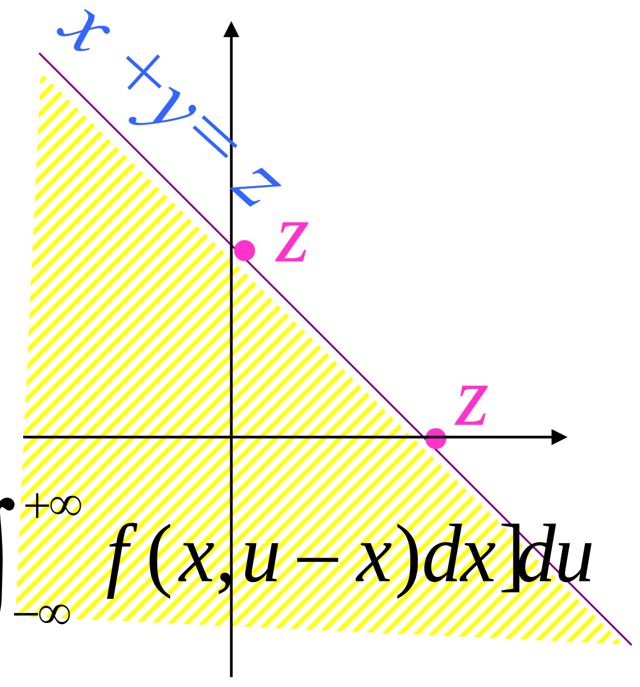
(2)再求Z的密度函数：$f_Z(z)=F_Z'(z)$

1. 和的分布：$Z = X + Y$

设(X, Y)的联合密度为 $f(x,y)$, 则

$$F_Z(z) = P(Z \le z) = P(X + Y \le z)$$

$$= \iint\limits_{x+y\le z} f(x,y)\,dxdy$$



$$= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-x} f(x,y)\,dy\right]dx = \int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty} f(x,u-x)\,dx\right]du$$

或 $$= \int_{-\infty}^{+\infty} dy \int_{-\infty}^{z-y} f(x,y)\,dx$$

$$-\infty < z < +\infty$$

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

或 $$f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)dy$$

特别地，若X ,Y相互独立，则

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

或 $$f_Z(z)=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)dy \overset{\text{记作}}{=} f_X(z)*f_Y(z)$$

称之为函数$f_X(z)$与$f_Y(z)$的卷积

正态随机变量的结论

如果随机变量 X 与 Y 相互独立，且

$$X \sim N\left(\mu_1, \ \sigma_1^2\right) \qquad Y \sim N\left(\mu_2, \ \sigma_2^2\right)$$

$$Z = X + Y,$$

则 $Z \sim N\left(\mu_1 + \mu_2, \ \sigma_1^2 + \sigma_2^2\right)$

2. 商的分布： $Z = X / Y$

问题

设 $(X, \quad Y)$是二维连续型随机变量，其联合密度函数为$f(x, \quad y)$，

计算随机变量 $Z = \dfrac{X}{Y}$ 的密度函数 $f_Z(z)$.

公式

$$f_Z(z) = \int_{-\infty}^{+\infty} |y| f(zy, \quad y) dy$$

3. M=max(X,Y)及N=min(X,Y)的分布

问题

设连续型随机变量X , Y 相互独立, X , Y的分布函数分别为 $F_X(x)$, $F_Y(y)$, $M=\max\{X,Y\}$, $N=\min\{X,Y\}$, 求 M ,N 的分布函数.

$$F_M(z)=P(\max\{X,Y\}\le z)=F_X(z)F_Y(z)$$

$$F_N(z)=P(\min\{X,Y\}\le z)$$

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

第四章 随机变量的数字特征

- 数学期望
- 方差
- 协方差与相关系数
- 大数定律
- 中心极限定理

1. 数学期望的定义

定义1 设 X 为离散 r.v. 其分布列为

$$P(X = x_k) = p_k, \quad k = 1,2,\Lambda$$

若无穷级数 $\sum_{k=1}^{+\infty} x_k p_k$ 绝对收敛，则称其和为 X 的数学期望，记作 $E(X)$，即

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

定义2 设连续 r.v. X 的 d.f. 为 $f(x)$

若广义积分

$$\int_{-\infty}^{+\infty} xf(x)dx$$

绝对收敛, 则称此积分为 X 的数学期望

记作 $E(X)$, 即

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

(1) $Y = g(X)$ 的数学期望

- 设离散 r.v. X 的概率分布为

$$P(X = x_i) = p_i, \quad i = 1, 2, \Lambda$$

若无穷级数 $\sum_{i=1}^{\infty} g(x_i) p_i$ 绝对收敛，则

$$E(Y) = \sum_{i=1}^{\infty} g(x_i) p_i$$

- 设连续 r.v. X 的 d.f. 为 $f(x)$

若广义积分 $\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛, 则

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

(2) 二维随机变量函数 $Z = g(X, Y)$的数学期望

- 设离散 r.v. (X, Y) 的概率分布为

$$P(X = x_i, Y = y_j) = p_{ij}, \quad i, j = 1, 2, \Lambda$$

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

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

❑ 设连续 r.v. (X, Y)的联合 d.f. 为

$$f(x, y)$$

若广义积分

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x, y) f(x, y) dxdy$$

绝对收敛，则

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

3. 数学期望的性质

(1) 设C是常数，则$E(C)=C$;

(2) 若C是常数，则$E(CX)=CE(X)$;

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

(4) 设X、Y独立，则 $E(XY)=E(X)E(Y)$.

注意:由$E(XY)=E(X)E(Y)$
不一定能推出X,Y独立

1. 方差概念

定义 若$E[X-E(X)]^2$存在, 则称其为随机变量 X 的**方差**, 记为$D(X)$ 或 $Var(X)$

即 $D(X)=E[X-E(X)]^2$

称 $\sqrt{D(X)}$ 为 X 的**均方差**或**标准差.**

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

2. 方差的性质

(1) 设C是常数,则$D(C)=0$;

(2) 若C是常数,则$D(CX)=C^2 D(X)$;

(3) 若X与Y 独立，则

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

推广：若$X_1,X_2,\ldots,X_n$相互独立,则

$$D[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} D(X_i)$$

$$D[\sum_{i=1}^{n} C_i X_i] = \sum_{i=1}^{n} C_i^2 D(X_i)$$

常见分布的数学期望和方差

分布	概率分布	期望	方差
0-1分布	$P(X=1)=p$ $P(X=0)=1-p$	p	$p(1-p)$
$B(n,p)$	$P(X=k)=C_n^k p^k (1-p)^{n-k}$ $k=0,1,2,\Lambda,n$	np	$np(1-p)$
$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2,\Lambda$	λ	λ

分布	概率密度	期望	方差
区间(a,b)上的均匀分布	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其它}\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & \text{其它}\end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

1. 协方差和相关系数的概念

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

$$\boldsymbol{Cov(X,Y)=E(XY)-E(X)E(Y)}$$

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

若 $\rho_{XY}=0$, 称 X,Y 不相关.

2. 协方差和相关系数的性质

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

(2) $\operatorname{cov}(aX,bY)=ab\operatorname{cov}(X,Y)$

(3) $\operatorname{cov}(X+Y,Z)=\operatorname{cov}(X,Z)+\operatorname{cov}(Y,Z)$

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

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

(6) $|\rho_{XY}| \le 1$

(7) $|\rho_{XY}| = 1 \Longleftrightarrow$ 存在常数a, b($a \neq 0$)，

使$P(Y=aX+b)=1$，

即X和Y以概率1线性相关.

X, Y 相互独立 $\underset{\not\Longleftarrow}{\Longrightarrow}$ X, Y 不相关

因为 $\rho_{XY} = 0 \Longleftrightarrow X, Y$ 不相关

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

若 (X, Y) 服从二维正态分布,

X, Y 相互独立 $\Longleftrightarrow X, Y$ 不相关

3. 矩和协方差矩阵

$E(X^k)$ —— X 的 k 阶原点矩

$E((X-E(X))^k)$ —— X 的 k 阶中心矩

$E(X^kY^l)$ —— X,Y 的 $k+l$ 阶混合原点矩

$E\left((X-E(X))^k(Y-E(Y))^l\right)$

—— X,Y 的 $k+l$ 阶混合中心矩

1. 切比雪夫不等式

设随机变量 X 的期望$E(X)$与方差 $D(X)$ 存在，则对于任意实数 $\varepsilon > 0$,

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

或 $$\boxed{P(|X - E(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}}$$

贝努里大数定律

设 n_A 是 n 次独立重复试验中事件 A 发生的次数, p 是每次试验中 A 发生的概率, 则

$\forall \varepsilon > 0$ 有

$$\lim_{n\to\infty} P\left(\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right) = 0$$

或

$$\lim_{n\to\infty} P\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) = 1$$

切比雪夫大数定律(平均数法则)

设 r.v. 序列 $X_1, X_2, \Lambda, X_n, \Lambda$ 相互独立，

且具有相同的数学期望和方差

$$E(X_k)=\mu, D(X_k)=\sigma^2, \quad k=1,2,\Lambda$$

则 $\forall \varepsilon > 0$ 有

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \geq \varepsilon\right)=0$$

或 $$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right)=1$$

定理一　林德伯格-列维中心极限定理

[独立同分布的中心极限定理]

设随机变量序列 $X_1, X_2, \Lambda\ , X_n, \Lambda$

独立同一分布, 且有期望和方差:

$$E(X_k) = \mu\,,\ D(X_k) = \sigma^2 > 0\,,\ k = 1,2,\Lambda$$

则对于任意实数 x,

$$\lim_{n\to\infty} P\left(\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \le x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \ = \Phi(x)$$

注 记 $Y_n = \dfrac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$

$$\lim_{n\to\infty} P\left(Y_n \le x\right) = \Phi(x)$$

即 n 足够大时，Y_n 的分布函数近似于标准正态随机变量的分布函数

$$\boldsymbol{Y_n \overset{\text{近似}}{\sim} N(0,1)}$$

$\sum_{k=1}^{n} X_k$ 近似服从 $N(n\mu, n\sigma^2)$

它表明:当n充分大时，n个具有期望和方差的独立同分布的$r.v$之和近似服从正态分布.

定理二

棣莫弗-拉普拉斯中心极限定理

[二项分布以正态分布为极限分布]

设 $Y_n \sim B(n, p)$，$0 < p < 1$，$n = 1,2,\ldots$

此定理表明，正态分布是二项分布的极限分布。当 n 充分大时，服从二项分布的随机变量 η_n 的概率计算可以转化为正态随机变量的概率计算：

$$P\{\eta_n = k\} \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(K-np)^2}{2npq}};$$

$$P\{a < \eta_n \leq b\} = P\left\{\frac{a-np}{\sqrt{npq}} < \frac{\eta_n - np}{\sqrt{npq}} \leq \frac{b-np}{\sqrt{npq}}\right\} \approx \phi\left(\frac{b-np}{\sqrt{npq}}\right) - \phi\left(\frac{a-np}{\sqrt{npq}}\right)$$

由于当 n 较大，p 又较小时，二项式分布的计算十分麻烦，因此，若用上面的近似公式计算将是非常简洁的。

第五章 数理统计的基本知识

- 总体与样本
- 统计量及查表

简单随机样本

若总体 X 的样本 (X_1, X_2, Λ, X_n) 满足:

(1) X_1, X_2, Λ, X_n 与 X 有相同的分布（代表性）

(2) X_1, X_2, Λ, X_n 相互独立（独立性）

则称 (X_1, X_2, Λ, X_n) 为简单随机样本.

一般,对有限总体,有放回抽样所得到的样本为简单随机样本,但使用不方便,常用不放回抽样近似代替.而代替的条件是

$$N / n \geq 10.$$

总体中个体总数　样本容量

定理：设 (X_1, X_2, Λ, X_n) 为来自总体X的简单随机样本，则有以下三条性质成立：

1）设总体 X 的分布函数为$F(x)$,则样本 (X_1, X_2, Λ, X_n) 的联合分布函数为

$$\boxed{F_{总}(x_1, x_2, \Lambda, x_n) = \prod_{i=1}^{n} F(x_i)}$$

2）若总体X的d.f.为 $f(x)$,则样本的联合 d.f.为

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

3）设总体 X 的均值和方差分别为μ和σ^2,则

$$E(X_k) = \mu, D(X_k) = \sigma^2, k = 1, 2, \ldots, n$$

常用的统计量

设 (X_1, X_2, Λ, X_n) 是来自总体 X 的容量为 n 的样本,称统计量

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$ 为样本均值

(2) $$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$ 为样本方差

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$ 为样本标准差

2) $$E(S_n^2) = \frac{n-1}{n}\sigma^2 \qquad E(S^2) = \sigma^2$$

(3) $A_k=\dfrac{1}{n}\sum_{i=1}^{n}X_i^k$ 为样本的k阶**原点矩**

(4) $B_k=\dfrac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^k$ 为样本的k阶**中心矩**

例如

$$A_1=\overline{X}$$

$$B_2=\frac{n-1}{n}S^2=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\overset{\triangle}{=}S_n^2$$

(2) χ^2-分布

χ^2分布是由正态分布派生出来的一种分布.

定义: 设 $X_1, X_2, \cdots, X_n$ 相互独立, 都服从正态分布N(0,1), 则称随机变量:

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

所服从的分布为自由度为 **n** 的 χ^2 分布.

记为 $\chi^2 \sim \chi^2(n)$

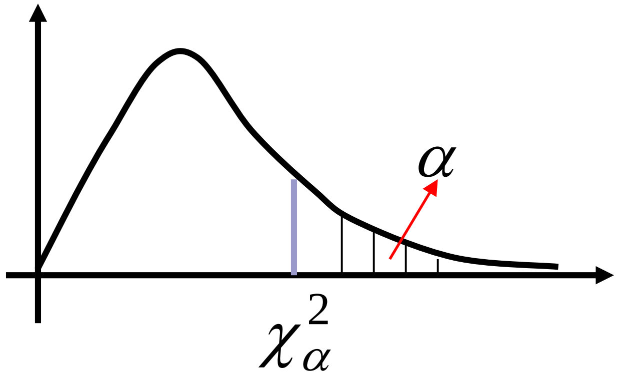
χ^2分布的性质：

$1^\circ \quad E\left(\chi^2(n)\right)=n,\ D\left(\chi^2(n)\right)=2n$

2° 若$X_1=\chi^2(n_1), X_2=\chi^2(n_2)$, X_1, X_2相互独立，

则 $X_1+X_2\sim\chi^2(n_1+n_2)$

$3^\circ \quad n\to\infty$时，$\chi^2(n)\to$ 正态分布



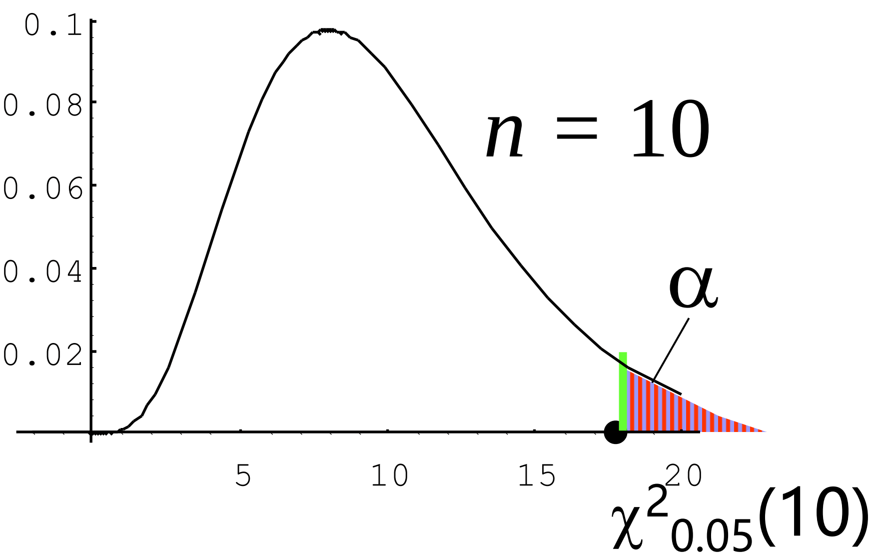
对于给定的$\alpha(0<\alpha<1)$，称满足条件：

$$P\{\chi^2 > \chi^2_\alpha(n)\} = \alpha$$

的点$\chi^2_\alpha(n)$为$\chi^2(n)$分布的上α分位点。

例如(P193, 表)

$$\chi^2_{0.05}(10) = 18.307$$
$$P\left(\chi^2(10) > 18.307\right) = 0.05$$



(3) *t* 分布 (Student 分布)

$X \sim N(0,1),\ Y \sim \chi^2(n),\ X, Y$独立，则称随机变量

$$\mathrm{T} = \frac{X}{\sqrt{Y/n}}$$

为服从自由度是n的$t-$分布，记作$T \sim t(n)$.

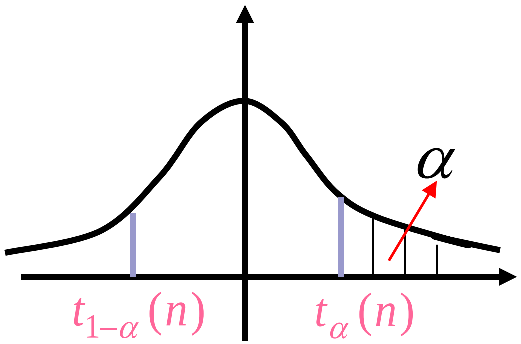
其密度函数为

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \qquad -\infty < t < \infty$$

对于给定的$\alpha(0<\alpha<1)$，称满足条件：

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

的点$t_{\alpha}(n)$为t分布的上α分位点。



由概率密度的对称性知：$t_{1-\alpha}(n) = -t_{\alpha}(n)$

(4) F 分布

若 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，$X, Y$ 独立，

则称随机变量

$$F = \frac{X / n_1}{Y / n_2}$$

为服从自由度是n_1, n_2 的 $F-$分布，

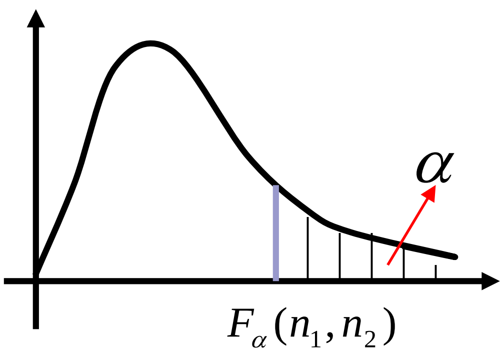
记作$F \sim F(n_1, n_2)$.

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

对于给定的$\alpha(0<\alpha<1)$，称满足条件：

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

的点$F_{\alpha}(n_1,n_2)$为F分布的上α分位点。



结论： $F_{1-\alpha}(n_1,n_2) = 1/F_{\alpha}(n_2,n_1)$

正态总体的抽样分布

（Ⅰ）一个正态总体

定理1 设X_1,Λ,X_n是总体$N(\mu,\sigma^2)$的样本，

$\overline{X},S^2$分别是样本均值与样本 方差，则有：

(1) $\overline{X}\sim N(\mu,\dfrac{\sigma^2}{n}).\qquad \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

(2) $\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$

(3) $\overline{X}$与S^2独立

(4) $\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$

一、导数的四则运算法则

$$\boxed{e^{\lambda}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}}$$

$(u\pm v)'=u'\pm v'$ $\quad (uv)'=u'v+uv'$ $\quad \left(\frac{u}{v}\right)'=\frac{u'v-uv'}{v^2}$

二、基本导数公式

(1) $(c)'=0$ $\quad$ (2) $x^{\mu}=\mu x^{\mu-1}$ $\quad$ (3) $(\sin x)'=\cos x$

(4) $(\cos x)'=-\sin x$ $\quad$ (5) $(\tan x)'=\sec^2 x$ $\quad$ (6) $(\cot x)'=-\csc^2 x$

(7) $(\sec x)'=\sec x\cdot\tan x$ $\quad$ (8) $(\csc x)'=-\csc x\cdot\cot x$

(9) $(e^x)'=e^x$ $\quad$ (10) $(a^x)'=a^x\ln a$ $\quad$ (11) $(\ln x)'=\frac{1}{x}$

(12) $\left(\log_a^{x}\right)' = \dfrac{1}{x\ln a}$ (13) $\left(\arcsin x\right)' = \dfrac{1}{\sqrt{1-x^2}}$ (14) $\left(\arccos x\right)' = -\dfrac{1}{\sqrt{1-x^2}}$

(15) $\left(\arctan x\right)' = \dfrac{1}{1+x^2}$ (16) $\left(\operatorname{arccot} x\right)' = -\dfrac{1}{1+x^2}$ (17) $\left(x\right)' = 1$ (18) $\left(\sqrt{x}\right)' = \dfrac{1}{2\sqrt{x}}$

七、基本积分公式

(1) $\int k dx = kx + c$ (2) $\int x^{\mu} dx = \dfrac{x^{\mu+1}}{\mu+1} + c$ (3) $\int \dfrac{dx}{x} = \ln|x| + c$

(4) $\int a^x dx = \dfrac{a^x}{\ln a} + c$ (5) $\int e^x dx = e^x + c$ (6) $\int \cos x dx = \sin x + c$

(7) $\int \sin x dx = -\cos x + c$ (8) $\int \dfrac{1}{\cos^2 x} dx = \int \sec^2 x dx = \tan x + c$

(9) $\int \dfrac{1}{\sin^2 x} = \int \csc^2 x dx = -\cot x + c$ (10) $\int \dfrac{1}{1+x^2} dx = \arctan x + c$

(11) $\int \frac{1}{\sqrt{1-x^2}}dx = \arcsin x + c$

$\int \tan x dx = -\ln|\cos x| + c$

$\int \cot x dx = \ln|\sin x| + c$

$\int \sec x dx = \ln|\sec x + \tan x| + c$

$\int \csc x dx = \ln|\csc x - \cot x| + c$

$\int \frac{1}{a^2+x^2}dx = \frac{1}{a}\arctan\frac{x}{a} + c$

$\int \frac{1}{x^2-a^2}dx = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + c$

$\int \frac{1}{\sqrt{a^2-x^2}}dx = \arcsin\frac{x}{a} + c$

$\int \frac{1}{\sqrt{x^2 \pm a^2}}dx = \ln\left|x+\sqrt{x^2 \pm a^2}\right| + c$

积分型	换元公式
$\int f(ax+b)dx = \frac{1}{a}\int f(ax+b)d(ax+b)$	$u = ax+b$
$\int f(x^{\mu})x^{\mu-1}dx = \frac{1}{\mu}\int f(x^{\mu})d(x^{\mu})$	$u = x^{\mu}$
$\int f(\ln x)\cdot\frac{1}{x}dx = \int f(\ln x)d(\ln x)$	$u = \ln x$
$\int f(e^x)\cdot e^x dx = \int f(e^x)d(e^x)$	$u = e^x$
$\int f(a^x)\cdot a^x dx = \frac{1}{\ln a}\int f(a^x)d(a^x)$	$u = a^x$
$\int f(\sin x)\cdot\cos x dx = \int f(\sin x)d(\sin x)$	$u = \sin x$
$\int f(\cos x)\cdot\sin x dx = -\int f(\cos x)d(\cos x)$	$u = \cos x$
$\int f(\tan x)\cdot\sec^2 x dx = \int f(\tan x)d(\tan x)$	$u = \tan x$
$\int f(\cot x)\cdot\csc^2 x dx = \int f(\cot x)d(\cot x)$	$u = \cot x$
$\int f(\arctan x)\cdot\frac{1}{1+x^2}dx = \int f(\arctan x)d(\arctan x)$	$u = \arctan x$
$\int f(\arcsin x)\cdot\frac{1}{\sqrt{1-x^2}}dx = \int f(\arcsin x)d(\arcsin x)$	$u = \arcsin x$

九、分部积分法公式

(1)形如$\int x^n e^{ax}dx$，令$u=x^n$，$dv=e^{ax}dx$

形如$\int x^n \sin xdx$令$u=x^n$，$dv=\sin xdx$

形如$\int x^n \cos xdx$令$u=x^n$，$dv=\cos xdx$

(2)形如$\int x^n \arctan xdx$，令$u=\arctan x$，$dv=x^n dx$

形如$\int x^n \ln xdx$，令$u=\ln x$，$dv=x^n dx$

(3)形如$\int e^{ax}\sin xdx$，$\int e^{ax}\cos xdx$令$u=e^{ax},\sin x,\cos x$均可。