期末考试范围

- 期末考试范围：第一章到第八章（第十二章不考，但作业需要做）
- 重点章节：第二，第三，第四，第七，第八章
- 重点概念题：全概率公式，随机变量函数的分布，独立性，边缘分布，两个随机变量函数分布，协方差，相关系数，参数估计，假设检验，区间估计

常用公式&结论

$$P(AB)+P(A\overline{B})=P(A)$$

常用符号的定义

对任意实数a以及非负整数r，定义

$$\binom{a}{r}=\frac{a(a-1)\dots(a-r+1)}{r!},\quad \binom{a}{0}=1$$

特别的，当a为正整数，且$r\leq a$时，该式表示组合数，即有：

$$\binom{a}{r}=C_a^r$$

排列组合

$$A_n^m=n(n-1)(n-2)\cdots(n-m+1)=\frac{n!}{(n-m)!}$$

$$C_n^m=\frac{n!}{m!(n-m)!}$$

抽签原理

中签概率与抽签顺序无关。n个人不放回地先后从n个签中抽1张，其中m张中奖（m<n），则每个人中奖的概率均为$P=\frac{m}{n}$。

概率论的基本概念

事件的运算

- 若事件A、B互斥，则$P(AB)=0$，反之不成立

交换律

$$A \cup B = B \cup A, \quad A \cap B = B \cap A$$

结合律

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

分配律

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

德摩根律

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$
$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

概率

概率的定义（条件）

非负性

对于每一个事件A，有$P(A) \geq 0$

规范性

对于必然事件S，有$P(S) = 1$

可列可加性

设$A_1, A_2, \ldots$是两两互不相容的事件，即对于$A_iA_j = \varnothing, i \neq j, i, j = 1, 2, \ldots,$ 有：

$$P(A_1 \cup A_2 \cup \ \ldots) = P(A_1) + P(A_2) + \ldots$$

概率的重要性质

性质1

$$P(\varnothing) = 0$$

性质2（有限可加性）

若$A_1, A_2, \ldots A_n$是两两互不相容的事件，则有：

$$P(A_1 \cup A_2 \cup \ \ldots \ \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

性质3

若事件A, B满足$A \subset B$，则有：

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

性质4

对任一事件A有：

$$P(A) \leq 1$$

性质5（逆事件的概率）

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

性质6（加法公式）

对任意两事件A，B有：

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

古典概型

古典概型的特点：

1. 试验的样本空间只包含有限个元素
2. 试验中每个基本事件发生的可能性相同

古典概型事件概率的计算公式

若事件A包含k个基本事件，即$A = \{e_{i_1}\} \cup \{e_{i_2}\} \cup \ldots \cup \{e_{i_k}\}$，这里$i_1, i_2, \ldots, i_n$是$1, 2, \ldots, n$中某k个不同的数，则有：

$$P(A) = \sum_{j=1}^{k} P(\{e_{i_j}\}) = \frac{k}{n} = \frac{A\text{包含的基本事件数}}{S\text{中基本事件的总数}}$$

条件概率

定义

设A，B是两个事件，且$P(A) > 0$，称

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

为在事件A发生的条件下事件B发生的条件概率。

条件概率满足的条件

非负性

对每一事件B，有$P(B|A) \geq 0$

规范性

对于必然事件S，有$P(S|A) = 1$

可列可加性

设$B_1, B_2, ...$是两两互不相容的事件，则有：

$$P(\bigcup_{i=1}^{\infty} B_i|A) = \sum_{i=1}^{\infty} P(B_i|A)$$

乘法定理

设$P(A) > 0$，有：

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

上式称为乘法公式

乘法公式的推广

设A，B，C为事件，且$P(AB) > 0$，则有：

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

一般的，设$A_1, A_2, \ldots, A_n$为n个事件，$n \geq 2$，且$P(A_1A_2 \ldots A_{n-1}) > 0$，则有

$$P(A_1A_2 \ldots A_n) = P(A_n|A_1A_2 \ldots A_{n-1})P(A_{n-1}|A_1A_2 \ldots A_{n-2}) \ldots P(A_2|A_1)P(A_1)$$

全概率公式 重点

设试验E的样本空间为S，A为E的事件，$B_1, B_2, \ldots, B_n$为S的一个划分，且 $P(B_i) > 0, (i = 1, 2, \ldots, n)$，则：

$$\begin{aligned} P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_n)P(B_n) \\ &= \sum_{i=1}^{n} P(A|B_i)P(B_i) \end{aligned}$$

贝叶斯公式 重点

设试验E的样本空间为S，A为E的事件，$B_1, B_2, \ldots, B_n$为S的一个划分，且$P(A) > 0, P(B_i) > 0 (i = 1, 2, \ldots, n)$，则:

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

事实上有:

$$P(B_i|A) = \frac{P(AB_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}$$
$$\text{(利用全概率公式)}$$

注 贝叶斯公式可用于“已知结果判断原因”的情况

随机变量及其分布 重点

分布函数

定义 设ξ是随机变量，x是任意的实数，称

$$F(x) = P(\xi \leq x) \quad (-\infty < x < \infty)$$

为随机变量ξ的分布函数

注意

- $F(x)$是随机变量ξ取值不大于x的概率，即分布函数$F(x)$描述的是事件$\xi \leq x$的概率，它是定义于全体实数，以区间$[0, 1]$为值域的普通函数
- 对任意实数$a < b$，随机点落在区间$(a, b]$内的概率，显然有

$$P(a < \xi \leq b) = P(\xi \leq b) - P(\xi \leq a) = F(b) - F(a)$$

分布函数的性质

- $$0 \leq F(x) \leq 1$$
- 若$x_1 < x_2$，则$F(x_1) \leq F(x_2)$，即$F(x)$是不减函数
- $$F(-\infty) = \lim_{x \to -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \to +\infty} F(x) = 1$$

- $$F(x_0+0)=\lim_{t\to x_0^+}F(t)=F(x_0)$$

即$F(x)$是右连续的.

进一步还可以得到其他性质:

- $$P(\xi>x)=1-P(\xi\le x)=1-F(x)$$
- $$P(\xi=x)=F(x)-F(x-0)$$
- $$P(\xi<x)=P(\xi\le x)-P(\xi=x)=F(x-0)$$
- $$P(\xi\ge x)=1-P(\xi<x)=1-F(x-0)$$

离散型随机变量及其分布

离散型随机变量：随机变量的所有可能取值是有限多个或可列无限多个

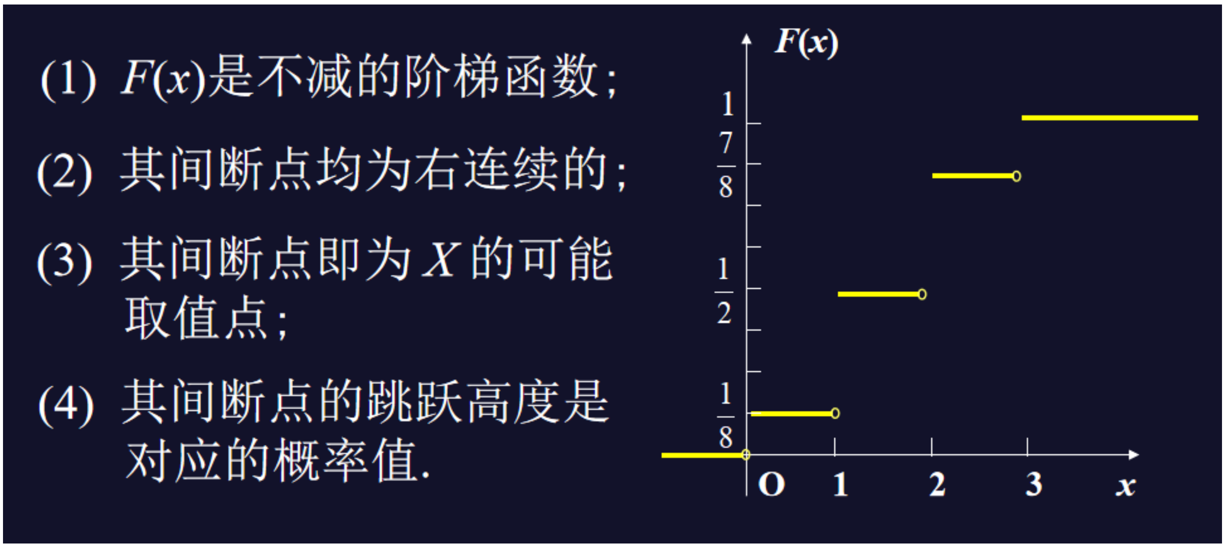
分布列

设ξ是一个离散型随机变量，取值为$x_1,x_2,\dots,x_n,\dots$,且 $P(\xi=x_k)=p_k(k=1,2,\dots)$，则称$\{p_k\}$为随机变量$\xi$的概率分布列，简称分布列.

根据概率的性质，分布列满足:

- $$p_k\ge 0,k=1,2,\dots$$
- $$\sum_{k=1}^{\infty}p_k=1$$

离散型随机变量的分布函数



常见的离散型随机变量

（0-1）分布（伯努利分布、两点分布）

随机变量X只可能取0、1两个值，其分布列为：

$$P(X=k)=p^k(1-p)^{1-k}$$
$$k=0,1 \quad (0<p<1)$$

二项分布

若试验E只有两种可能结果A和它的逆事件$\overline{A}$，则称E为伯努利试验.

设随机变量ξ表示n重伯努利试验中“成功”出现的次数，设每次成功的概率为p，则ξ的分布列为：

$$p_k=P(\xi=k)=C_n^k p^k(1-p)^{n-k}$$
$$(k=0,1,\ldots,n;0<p<1)$$

称ξ服从二项分布，记为：

$$\xi \sim b(n,p)$$

泊松分布

设随机变量X取值$0,1,2,\ldots$，且其分布列为：

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \quad k=0,1,2,\ldots$$
$$\text{其中}\lambda>0\text{是常数}$$

则称X服从参数为λ的泊松分布，记作：

$$X \sim \pi(\lambda)$$

泊松定理

设随机变量X服从二项分布$B(n,p_n)$，又设$\lambda>0$是一常数，n是任意正整数，若$np_n=\lambda$，则对任一固定的非负整数k有：

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

上述定理表明当n很大，p很小$(np=\lambda)$时有以下近似式：

$$C_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k e^{-\lambda}}{k!} \quad (\text{其中}\lambda=np)$$

即以n，p为参数的二项分布的概率值可以由参数为$\lambda=np$的泊松分布的概率值近似.

连续型随机变量及其分布

定义 对于随机变量X的分布函数$F(x)$，如果存在非负可积函数$f(x), x\in(-\infty,+\infty)$，使得对任意实数x，有

$$F(x)=\int_{-\infty}^{x} f(t)dt=P(X\leq x)$$

则称X为连续型随机变量，称$f(x)$为X的概率密度函数，简称为概率密度.

注

- 连续型随机变量的分布函数在R上连续
- 连续型随机变量取任一指定实数值a的概率均为0，即:

$$P(X=a)=0$$

- 对连续型随机变量X，有:

$$\begin{aligned}P(a\leq X\leq b)&=P(a<X\leq b)\\&=P(a\leq X<b)\\&=P(a<X<b)\end{aligned}$$

概率密度的性质

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

上述两条性质是判定一个$f(x)$是某随机变量X的概率密度的充要条件.

- 对于任意实数$x_1, x_2(x_1<x_2)$:

$$P\{x_1<X\leq x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)dx$$

- 若$f(x)$在点x处连续，则有:

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

常见的连续型随机变量

均匀分布

若随机变量X的概率密度为:

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

则称X在区间(a,b)上服从均匀分布，记作

$$X\sim U(a,b)$$

X的分布函数为

$$F(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\le x<b\\ 1, & x\ge b\end{cases}$$

指数分布

若随机变量X具有概率密度：

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0\\ 0, & x\le 0\end{cases}\qquad \text{其中}\theta>0\text{为常数}$$

则称X服从参数为θ的指数分布，其分布函数为：

$$F(x)=\begin{cases}1-e^{-\frac{x}{\theta}}, & x>0\\ 0, & x\le 0\end{cases}$$

说明 $E(X)=\theta$

性质

无记忆性

$$P\{X>x_0+\Delta x|X>x_0\}=P\{X>\Delta x\}$$

证明：

$$P\{X>x_0+\Delta x|X>x_0\}=\frac{P\{(X>x_0+\Delta x)\cap(X>x_0)\}}{P\{X>x_0\}}$$

$$=\frac{P\{X>x_0+\Delta x\}}{P\{X>x_0\}}=\frac{e^{-\frac{x_0+\Delta x}{\theta}}}{e^{-\frac{x_0}{\theta}}}=P\{X>\Delta x\}$$

另注

- 指数分布：唯一无记忆性的连续型分布
- 几何分布：唯一无记忆性的离散型分布

正态分布

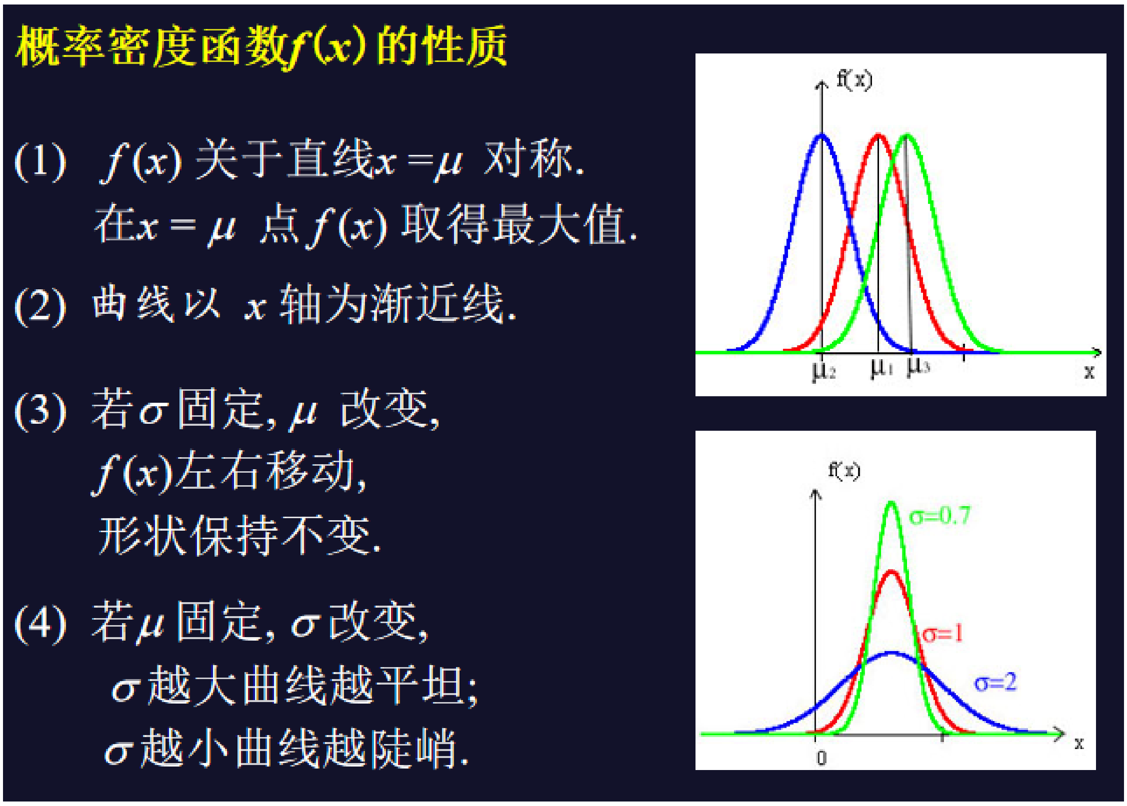
若连续型随机变量X的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty$$
$$\text{其中}\mu\text{和}\sigma(\sigma>0)\text{都是常数}$$

则称X服从参数为μ和σ的正态分布或高斯分布，记作

$$X\sim N(\mu,\sigma^2)$$

正态分布概率密度函数$f(x)$的性质

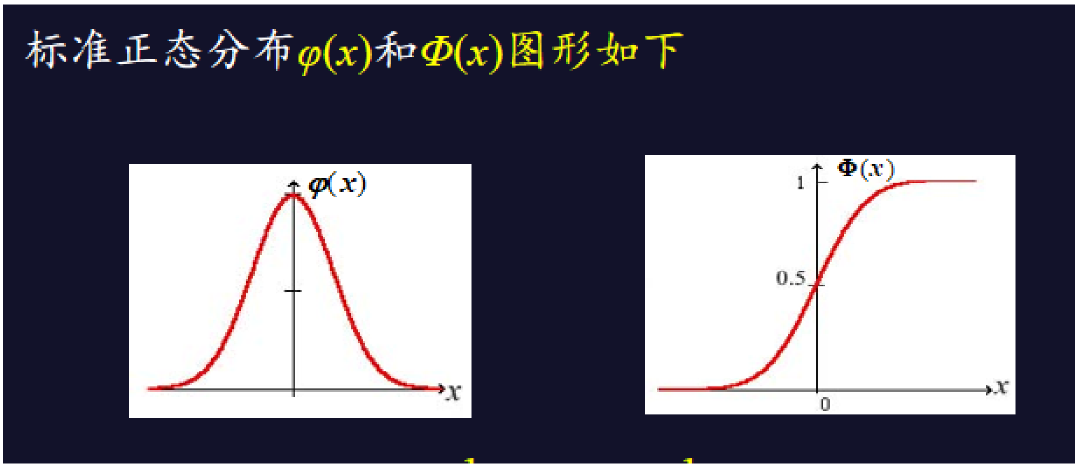


标准正态分布

当$\mu=0,\sigma=1$时，$N(0,1)$称为标准正态分布

$$\text{概率密度函数：}\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\quad -\infty<x<\infty$$

$$\text{分布函数：}\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt,\quad -\infty<x<\infty$$



$\varphi(x), \Phi(x)$**的性质**

- $$\varphi(0)=\frac{1}{\sqrt{2\pi}},\quad \Phi(0)=\frac{1}{2}$$
- $$\varphi(-x)=\varphi(x)$$
- $$\Phi(-x)=1-\Phi(x)$$
- 若正态分布$N(\mu,\sigma^2)$的分布函数为$F(x)$，则

$$F(x)=\Phi(\frac{x-\mu}{\sigma})$$

该性质表明，可以用标准正态分布的分布函数$\Phi(x)$计算其他正态分布的分布函数$F(x)$.

进而可以得到，若$X\sim N(\mu,\sigma^2)$，则

$$F(x)=P\{\frac{X-\mu}{\sigma}\le\frac{x-\mu}{\sigma}\}=\Phi(\frac{x-\mu}{\sigma})$$
$$P\{x_1<X\le x_2\}=F(x_2)-F(x_1)=\Phi(\frac{x_2-\mu}{\sigma})-\Phi(\frac{x_1-\mu}{\sigma})$$

$\Phi(x)$的计算

- $x\ge 0$ 时，查标准正态分布函数表.
- $x<0$时，用 $\Phi(x)=1-\Phi(-x)$.

若 $X\sim N(0,1)$，则

(1) $P(X\le a)=\Phi(a)$;

(2) $P(X>a)=1-\Phi(a)$;

(3) $P(a<X<b)=\Phi(b)-\Phi(a)$;

(4) 若$a\ge 0$，则

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

3σ规则

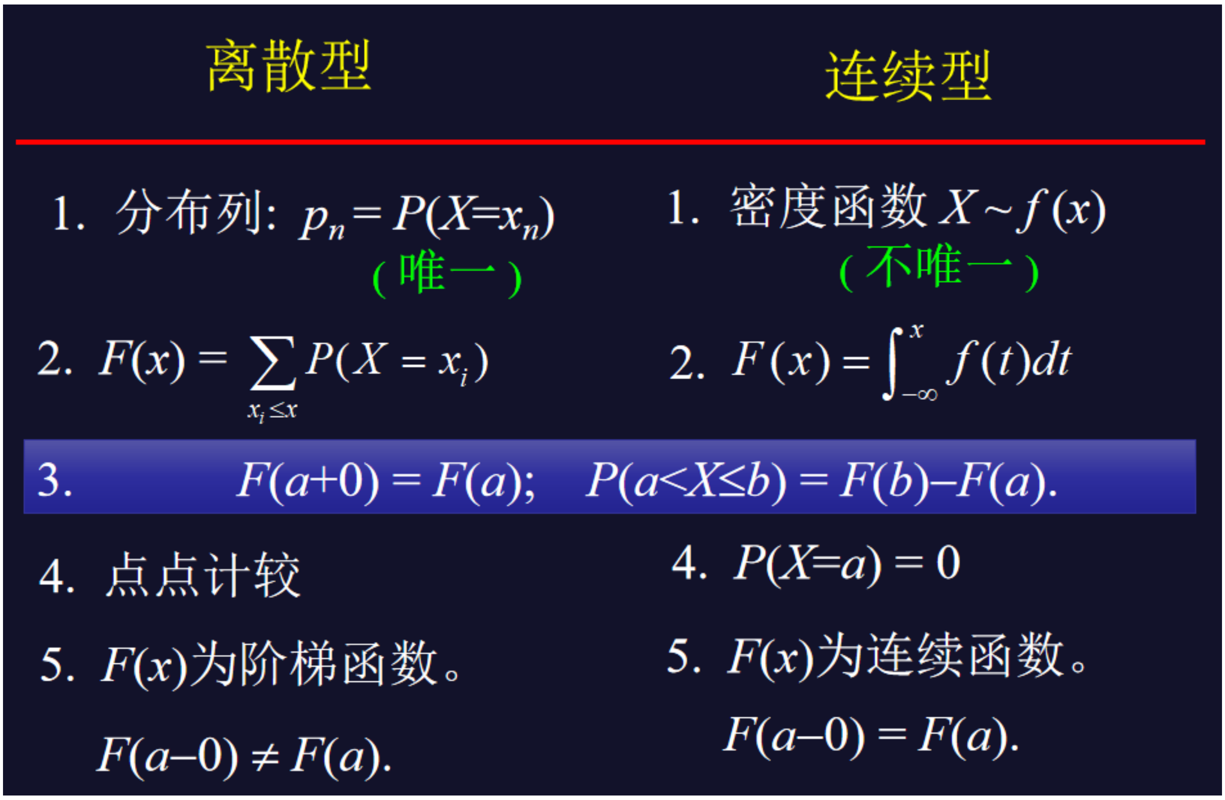
当$X \sim N(0,1)$时

$$\begin{aligned} P(|X| \leq 1) &= 0.6826 \\ P(|X| \leq 2) &= 0.9544 \\ P(|X| \leq 3) &= 0.9974 \end{aligned}$$

当$X \sim N(\mu, \sigma^2)$时

$$\begin{aligned} P(|X-\mu| \leq \sigma) &= 0.6826 \\ P(|X-\mu| \leq 2\sigma) &= 0.9544 \\ P(|X-\mu| \leq 3\sigma) &= 0.9974 \end{aligned}$$

离散型和连续型随机变量的比较



随机变量函数的分布 重点

若$X \sim N(\mu, \sigma^2)$，令$Y = \frac{X-\mu}{\sigma}$，则$Y \sim N(0,1)$.

X为连续型随机变量

若X为连续型随机变量，概率密度为$f_X(x)$，求$Y = g(X)$的概率密度$f_Y(x)$一般采用分布函数法：

1. 分布函数法

第一步：由X的分布函数求出Y的分布函数表达式

$$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=P\{X\in C_y\}$$

其中$C_y=\{x\mid g(x)\le y\}$

第二步：对分布函数求导数即得所求概率密度.

多维随机变量及其分布 重点

边缘分布 重点

- 离散型

$$P\{X=x_i\}=\sum_{j=1}^{\infty}p_{ij}=p_{i\cdot}$$

$$P\{Y=y_j\}=\sum_{i=1}^{\infty}p_{ij}=p_{\cdot j}$$

称$p_{i\cdot},p_{\cdot j}$为边缘分布律

- 连续型

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

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

称$f_X(x),f_Y(y)$为边缘概率密度

条件分布

###

相互独立的随机变量 重点

若随机变量X,Y相互独立，其等价条件为:

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

- 若连续型随机变量X,Y相互独立，其等价条件为:

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

- 若离散型随机变量X,Y相互独立，其等价条件为：

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{y=y_j\}$$

两个随机变量的函数的分布 重点

$Z=X+Y$的分布

设(X,Y)是二维连续型随机变量，它具有概率密度$f(x,y)$，则$Z=X+Y$仍为连续型随机变量，其概率密度为：

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

若X,Y相互独立，设(X,Y)关于X,Y的边缘密度分别为$f_X(x),f_Y(y)$，则上式可化为：

$$f_{X+Y}(z)=\int_{-\infty}^{\infty}f_X(z-y)f_Y(y)dy$$
$$\text{或}\quad f_{X+Y}(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx$$

上两式称为f_X,f_Y的卷积公式

正态随机变量的线性组合

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

$$Z=X+Y\sim N(\mu_1+\mu_2,\ \sigma_1^2+\sigma_2^2)$$
$$aX+bY\sim N(a\mu_1+b\mu_2,\ a^2\sigma_1^2+b^2\sigma_2^2)$$

更一般地，可以证明有限个相互独立的正态随机变量的线性组合仍然服从正态分布。

$M=\max\{X,Y\}$及$N=\min\{X,Y\}$的分布 常考

- $M=\max\{X,Y\}$的分布函数为：

$$F_{max}(z)=F_X(z)F_Y(z)$$

- $N=\min\{X,Y\}$的分布函数为：

$$F_{min}(z)=1-[1-F_X(z)][1-F_Y(z)]$$

随机变量的数字特征 重点

常用分布的数字特征总结

常用分布	分布律 / 概率密度	期望	方差
两点分布	$P(X=k)=p^k(1-p)^{1-k} \quad k=0,1$	p	pq
二项分布$B(n,p)$	$P(X=k)=C_n^k p^k(1-p)^{1-k}\ k=0,1,\dots,n$	np	npq
泊松分布$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!} \quad k=0,1,\dots$	λ	λ
均匀分布$U(a,b)$	$f(x)=\begin{cases}1/(b-a), & a<x<b \\ 0, & 其它\end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
指数分布	$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0 \\ 0, & 其它\end{cases}$	θ	θ^2
正态分布$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty<x<\infty$	μ	σ^2

纠正:

对于二项分布$B(n,p)$，其分布律为:

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

- 两点分布的方差推导:

$$E(X^2)=E(X)=p$$
$$D(X)=E(X^2)-E^2(X)=p-p^2=p(1-p)$$

- 指数分布的分布函数:

$$F(x)=1-e^{-\frac{x}{\theta}} \quad (x>0)$$

数学期望

数学期望简称期望，又称均值

离散型随机变量的数学期望

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

（级数绝对收敛）

连续型随机变量的数学期望

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

（积分绝对收敛）

随机变量函数的数学期望

设随机变量Y是随机变量X的函数：$Y=g(X)\quad$ (连续函数)

- X为离散型随机变量

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

- X为连续型随机变量

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

多随机变量函数的数学期望

离散型

定理 设 $g(X,Y)$ 为二元连续函数，

(1) 若 (X,Y) 是离散型，$P\{X=x_i,Y=y_j\}=p_{ij},i,j=1,2,\cdots$，

且级数 $\sum_i\sum_j g(x_i,y_j)p_{ij}$ 绝对收敛，则

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

(2) 若 (X,Y) 是连续型，其概率密度为 $f(x,y)$，

且级数 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$ 绝对收敛，则

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

连续型

设Z是随机变量X，Y的函数：$Z=g(X,Y)$（连续函数），若二维随机变量(X,Y)的概率密度为$f(x,y)$，则有

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

特别的，令$Z=X$，则有：

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy$$

由(X,Y)联合分布律或联合概率密度求X、Y的数学期望

对于二维随机变量(X,Y)，若$E(X)$、$E(Y)$都存在

(1) 若(X,Y)是离散型，$P\{X=x_i,Y=y_j\}=p_{ij},i,j=1,2,\cdots$,

则 $$E(X)=\sum_i x_i p_{i\cdot}=\sum_i\sum_j x_i p_{ij}$$

$$E(Y)=\sum_j y_j p_{\cdot j}=\sum_j\sum_i y_j p_{ij}$$

(2) 若(X,Y)是连续型，其概率密度为$f(x,y)$,

则 $$E(X)=\int_{-\infty}^{+\infty}xf_X(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)\mathrm{d}x\mathrm{d}y$$

$$E(Y)=\int_{-\infty}^{+\infty}yf_Y(y)\mathrm{d}y=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)\mathrm{d}x\mathrm{d}y$$

条件数学期望

- 离散型

$$E(X|Y=y_j)=\sum x_i p_{i|j}$$

- 连续型

$$E(X|Y=y)=\int_{-\infty}^{+\infty}xf_{X|Y}(x|y)dx$$

数学期望的性质

- $$E(C)=C$$
- $$E(CX)=CE(X)$$

- $$E(X \pm Y) = E(X) \pm E(Y)$$
- 设X, Y相互独立，则

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

注意 该结论反之不一定成立

方差

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

定义 设X是一个随机变量，若$E[(X-E(X)]^2$存在，称$E[(X-E(X)]^2$为X的方差. 记为$D(X)$或$Var(X)$，即

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

方差的算术平方根$\sqrt{D(X)}$称为X的标准差或均方差，记为$\sigma(X)$.

方差的计算

方差$D(X)$可以看做随机变量X的函数$g(X) = [X - E(X)]^2$的数学期望

1) X为离散型，分布律为$P\{X=x_k\}=p_k$

$$D(X) = \sum_k [x_k - E(X)]^2 p_k$$

2) X为连续型，概率密度$f(x)$

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

计算方差的简化公式:

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

方差的性质

设$D(X)$, $D(Y)$都存在，C为常数

(1) $D(C)=0$

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

(3) $D(X+Y)=D(X)+D(Y)+2E\left[X-E(X)\right]\left[Y-E(Y)\right]$

特别，若X,Y相互独立，则$D(X+Y)=D(X)+D(Y)$

(4) $D(X)=0$ 的充要条件是 $P\{X=E(X)\}=1$

对于(3)有：

$$D(X+Y)=D(X)+D(Y)+2Cov(X,Y)$$

当X,Y不相关时，$Cov(X,Y)=0$，则有$D(X+Y)=D(X)+D(Y)$

协方差与相关系数 重点

协方差

定义 设X和Y为两随机变量，若$E\{[X\text{-}E(X)][Y\text{-}E(Y)]\}$存在，则称之为随机变量$X$和$Y$的协方差，记为$Cov(X,Y)$或$\sigma_{XY}$，即

$$\sigma_{XY}=Cov(X,Y)=E\left[X-E(X)\right]\left[Y-E(Y)\right]$$

(1)若(X,Y)为离散型，$P\{X=x_i,Y=y_j\}=p_{ij}$，则

$$Cov(X,Y)=\sum_{i,j}[x_i-E(X)][y_j-E(Y)]p_{ij}$$

(2)若(X,Y)为连续型，其概率密度为$f(x,y)$，则

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

计算协方差的简单公式：

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

特别地，当$X=Y$时，

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

可见方差可看作协方差的一个特例，并且有

$$D(X+Y)=D(X)+D(Y)+2Cov(X,Y)$$

协方差的性质

协方差的性质：

(1) $Cov(X,Y)=Cov(Y,X)$

(2) $Cov(aX,bY)=abCov(X,Y)$ 其中a,b为常数

(3) $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

(4) 若X与Y独立，则 $Cov(X,Y)=0$

相关系数

定义： 设$D(X)>0, D(Y)>0$，称

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

为随机变量 X 和 Y 的**相关系数**.

也可表示为

$$\rho_{XY}=\frac{\sigma_{XY}}{\sqrt{\sigma_{XX}}\sqrt{\sigma_{YY}}}$$

相关系数的性质

- $|\rho_{XY}|\leq 1$
- $|\rho_{XY}|=1 \Longleftrightarrow$ 存在常数a,b,使得$P\{Y=aX+b\}=1$

说明

(1) 相关系数ρ_{XY}刻画了随机变量 Y 与X之间的"线性相关"程度: $|\rho_{XY}|$ 的值越接近于1， Y 与X 的线性相关程度越高； $|\rho_{XY}|$ 的值越接近于0， Y与X 的线性相关程度较弱.

当 $|\rho_{XY}|=1$， 称X与Y完全线性相关；

当 $\rho_{XY}=0$ ， 称X与Y不相关；

(2) 当ρ_{XY}=0时,只说明Y 与X之间没有线性关系,并不能说明Y 与X之间没有其他函数关系,从而不能推出Y 与X独立.

对"不相关性"和"相互独立"区别的说明

对随机变量X，Y，有如下事实等价：

1、$Cov(X,Y)=0$;

2、X与Y不相关；

3、$E(XY)=E(X)E(Y)$;

4、$D(X+Y)=D(X)+D(Y)$.

注意相互独立与不相关性是不同的两个概念， 独立性比不相关性的要求更强.

性质 如果随机变量X与Y互相独立，则X与Y不相关。（反之不成立）

相互独立 $\underset{\not\Leftarrow}{\Rightarrow}$ 不相关性

特例 二维正态随机变量X与Y相互独立与不相关是等价的

矩和协方差矩阵

原点矩&中心矩

定义 设X和Y是随机变量，k, l为正整数，若以下期望存在，称

$E(X^k)$ 为X的**k阶原点矩**（简称**k阶矩**）

$E\{[X-E(X)]^k\}$ 为X的 **k阶中心矩**

$E(X^kY^l)$ 为X和Y的**$k+l$ 阶混合矩**；

$E\{[X-E(X)]^k[Y-E(Y)]^l\}$为$X$和$Y$的**$k+l$ 阶混合中心矩**

显然，$E(X)$是X的**一阶原点矩**，$D(X)$是X的**二阶中心矩**，$\mathrm{Cov}(X,Y)$是X和Y的**二阶混合中心矩**.

协方差矩阵

定义 设n维随机变量$(X_1, X_2, \ldots, X_n)$的二阶混合中心矩

$$\sigma_{ij}=Cov(X_i,X_j)=E[X_i-E(X_i)][X_j-E(X_j)]$$

$$(i, j=1,2,\ldots,n)$$

都存在，则称矩阵

$$B=\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

为n维随机变量$(X_1,X_2,\ldots,X_n)$的**协方差矩阵，记为B.**

显然协方差矩阵的元素为方差和协方差，且$B^T=B$.

相关矩阵

若令$\rho_{ij} = \dfrac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$，则 ρ_{ij} 是X_i和X_j的相关系数，

此时称矩阵

$$R = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{pmatrix}$$

为n维随机变量$(X_1, X_2, \ldots, X_n)$的**相关矩阵**，记为R.

大数定律与中心极限定理

大数定律

切比雪夫不等式

设X为随机变量，$E(X) = \mu$，$D(X) = \sigma^2 < +\infty$，则对任意的正数$\varepsilon$，有

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

等价于：

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

说明

- 切比雪夫不等式只与随机变量X的期望和方差有关，与X的分布类型无关
- 切比雪夫不等式给出了随机变量偏离期望的概率的下界，但通常该下界比较“保守”

三个常用的大数定律

依概率收敛：

定义 设$X_1, X_2, \ldots, X_n, \ldots$是随机变量序列，$A$是常数若对任意$\varepsilon > 0$，有

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

则称$X_1, X_2, \ldots, X_n, \ldots$依概率收敛于$A$，记为$X_n \xrightarrow{P} A.$

注：依概率收敛表示当n无限增大时，对任意$\varepsilon>0$，事件$\{|X_n-A|<\varepsilon\}$发生的概率无限接近于1. 但并不排除事件$\{|X_n-A|\geq\varepsilon\}$发生，只是发生的可能性很小.

切比雪夫大数定律

定理 设$X_1, X_2, \ldots, X_n, \ldots$是相互独立随机变量序列，$l$是正常数，且$E(X_i) = \mu_i < +\infty$，$D(X_i) \leq l < +\infty$ $(i=1,2,\ldots)$，则对任意的$\varepsilon>0$，有

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

辛钦大数定律

设随机变量$X_1, X_2, \ldots, X_n, \ldots$相互独立，服从同一分布，且$E(X_k)=\mu, (k=1,2,\ldots)$，作前$n$个随机变量的算术平均

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k,$$

则对于任意正数ε，有

$$\lim_{n\to\infty} P\{|\bar{X} - \mu| < \varepsilon\} = \lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1.$$

说明

- 切比雪夫大数定律里随机变量序列**不要求是同分布的**，但是要求它们的方差有一致的上界.
- 辛钦大数定律里随机变量序列要求是同分布的，但**不要求它们的方差存在或有一致上界**.

伯努利大数定律

设 n次独立重复试验中事件A发生的次数为n_A，事件A在每次试验中发生的概率是$p(0<p<1)$，则对于任意正数$\varepsilon>0$，有

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

说明

伯努利大数定律表明事件发生的频率$\frac{n_A}{n}$依概率收敛于事件的概率p.

这个定理以严格的数学形式表达了频率的稳定性. 就是说n很大时, 事件发生的频率与概率有较大偏差的可能性很小. 由实际推断原理, 在实际应用中, 当试验次数很大时, 便可以用事件发生的频率来代替事件的概率.

中心极限定理

中心极限定理：随机变量和的分布收敛于正态分布的一类定理

三个常用的中心极限定理

独立同分布中心极限定理

定理 设随机变量$X_1, X_2, \dots, X_n, \dots$相互独立，服从同一分布, 且 $E(X_k)=\mu$, $D(X_k)=\sigma^2\neq 0$ $(k=1,2,\dots)$，则

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

也就是说当n充分大时,

$$Y_n=\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}=\frac{\frac{1}{n}\sum_{k=1}^{n} X_k-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$ 近似服从标准正态分布

说明

1)定理表明，独立同分布的随机变量之和 $\sum_{k=1}^{n} X_k$，当n充分大时，随机变量之和与其标准化变量分别有

$$\sum_{k=1}^{n} X_k \overset{\text{近似地}}{\sim} N(n\mu, n\sigma^2), \quad \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \overset{\text{近似地}}{\sim} N(0,1).$$

2)独立同分布中心极限定理的另一种形式可写为

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k \overset{\text{近似地}}{\sim} N(\mu, \sigma^2/n) \text{或} \quad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \overset{\text{近似地}}{\sim} N(0,1).$$

李雅普诺夫中心极限定理

定理 设$X_1, X_2, \ldots, X_n, \ldots$是相互独立随机变量序列，且$E(X_i) = \mu_i$, $D(X_i) = \sigma_i^2$ $(i=1,2,\ldots)$，

令 $B_n^2 = \sum_{i=1}^{n} \sigma_i^2$，若存在$\delta>0$，使得

$$\lim_{n\to\infty} \frac{1}{B_n^{2+\delta}} \sum_{i=1}^{n} E\left|X_i - \mu_i\right|^{2+\delta} = 0,$$

则对任意实数x，有

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\{Z_n \le x\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x)$$

其中 $Z_n = \dfrac{\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mu_k}{B_n} = \dfrac{\sum_{k=1}^{n} (X_k - \mu_k)}{B_n}$

德莫弗-拉普拉斯中心极限定理 重点 （参考教材第五版P126，例1,2,3）

设随机变量 $Y_n \sim B(n,p)$，其中 $0 < p < 1, n = 1,2,\ldots$

则对任意x，有

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

说明

定理表明，**正态分布是二项分布的极限分布**，事件发生的频数 Y_n 近似服从正态分布 $N(np, npq)$.

即 $Y_n \overset{\text{近似地}}{\sim} N(np, np(1-p))$ $\qquad \dfrac{Y_n}{n} \overset{\text{近似地}}{\sim} N(p, \dfrac{pq}{n})$

样本与抽样分布

基本概念

经验分布函数

定义：设 $(x_1, x_2, \ldots, x_n)$ 是总体 X 的一个容量为 n 的样本值. 先将 $x_1, x_2, \ldots, x_n$ 按自小到大的次序排列，并重新编号，设为 $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$.

则经验分布函数 $F_n(x)$ 为

$$F_n(x) = \begin{cases} 0, & 若 x \le x_{(1)}, \\ \dfrac{k}{n}, & 若 x_{(k)} < x \le x_{(k+1)}, \quad (k = 1, 2, \cdots, n-1) \\ 1, & 若 x > x_{(n)}. \end{cases}$$

性质

(1) $F_n(x)$ 为不减左连续函数；

(2) $\lim\limits_{x \to -\infty} F_n(x) = 0, \lim\limits_{x \to \infty} F_n(x) = 1$

$F_n(X)$ 的值表示在 n 次独立重复试验中事件“$X \le x$”发生的频率。

用样本推断总体

定理 设总体X的分布函数为$F(x)$，样本分布函数为$F_n(x)$，则当$n\to\infty$时$F_n(x)$依概率收敛于分布函数$F(x)$，即

$$P\{\lim_{n\to\infty}\sup_{-\infty<x<\infty}|F_n(x)-F(x)|=0\}=1.$$

由定理可知，当n充分大时,总体分布函数$F(x)$可用经验分布函数$F_n(x)$近似代替，这就是用样本来推断总体的理论依据.

样本数字特征

常用统计量

- 样本均值

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

- 样本方差

$$\begin{aligned}S^2&=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\\&=\frac{1}{n-1}(\sum_{i=1}^{n}X_i^2-n\overline{X}^2)\end{aligned}$$

注意 分母为$n-1$

- 样本标准差

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

- 样本k阶（原点）矩

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k\quad(k=1,2,\dots)$$

- 样本k阶中心矩

$$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k \quad (k=1,2,\ldots)$$

总结： 设$X_1, X_2, \ldots, X_n$是来自总体X的一个样本，$(x_1,x_2,\ldots,x_n)$是其观察值.

- 样本均值 $\bar{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$;
- 样本方差 $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2=\dfrac{1}{n-1}\left(\sum\limits_{i=1}^{n}X_i^2-n\bar{X}^2\right)$
- 样本标准差 $S=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2}$
- 样本k阶(原点)矩 $A_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^k \quad (k=1,2,\ldots)$
- 样本k阶中心矩 $B_k=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\bar{X})^k \quad (k=2,3,\ldots)$

计算观察值即用x_i代替X_i

定理 设$X_1,X_2,\ldots,X_n$是来自总体X的一个样本，若$D(X)=\sigma^2$，则

(1)$S^2=\dfrac{1}{n-1}\left(\sum\limits_{i=1}^{n}X_i^2-n\bar{X}^2\right)$;

(2)$E\left(S^2\right)=\sigma^2$. (即样本方差的期望等于总体方差)

正态总体的抽样分布

U统计量及其分布

设总体$X\sim N(\mu,\sigma^2)$, $X_1, X_2, \ldots, X_n$是它的一个样本，则样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是正态分布的线性组合，因此也服从正态分布，其期望和方差分别为

$$E(\bar{X})=E(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=E(X)=\mu,$$

$$D(\bar{X})=D(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n}D(X)=\frac{\sigma^2}{n}$$

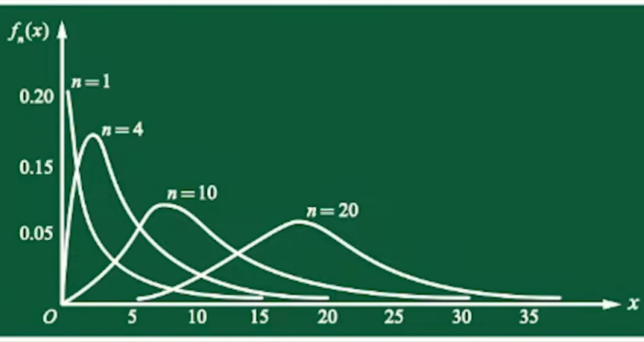
即 $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$.

定理 若总体 $X\sim N(\mu,\sigma^2)$, $X_1,X_2,\cdots,X_n$ 是它的一个样本，则统计量

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

称为U统计量.

χ^2统计量及其分布

定义	$\chi^2=X_1^2+X_2^2+\cdots+X_n^2$ 服从自由度为n的χ^2分布，记 $\chi^2\sim\chi^2(n)$，其中 $X_1,X_2,\cdots,X_n$ 相互独立，且 $X_i\sim N(0,1),(i=1,2,\cdots)$
图像	
期望与方差	$E(\chi^2)=n,\ D(\chi^2)=2n.$
性质	若 $X\sim\chi_1^2(n_1),Y\sim\chi_2^2(n_2)$， X,Y相互独立，则 $X+Y\sim\chi^2(n_1+n_2)$

性质：

χ^2分布的性质：

(1) 设 $\chi^2\sim\chi^2(n)$，则 $E(\chi^2)=n,D(\chi^2)=2n.$

(2) 当n充分大时，$\chi^2(n)$ 近似服从正态分布$N(n,2n)$

(3) 可加性，设 $\chi_1^2\sim\chi^2(n_1),\chi_2^2\sim\chi^2(n_2)$

且$\chi_1{}^2$与χ_2^2相互独立，则有 $\chi_1^2+\chi_2^2\sim\chi^2(n_1+n_2)$

定理 若总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是它的一个样本，则统计量

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

称为χ^2 统计量。

重要结论 （记忆，证明不要求）

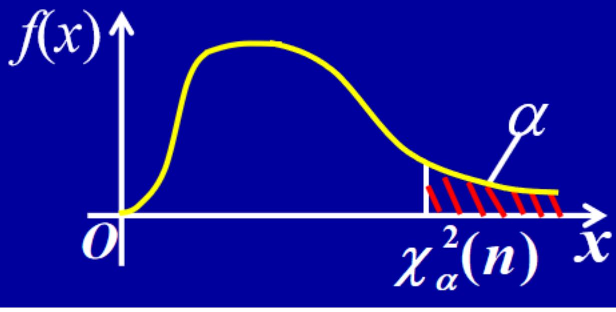
若总体$X\sim N(\mu,\sigma^2)$，$X_1,X_2,\ldots,X_n$是它的一个样本，则统计量

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

称为χ^2统计量

上侧α分位数

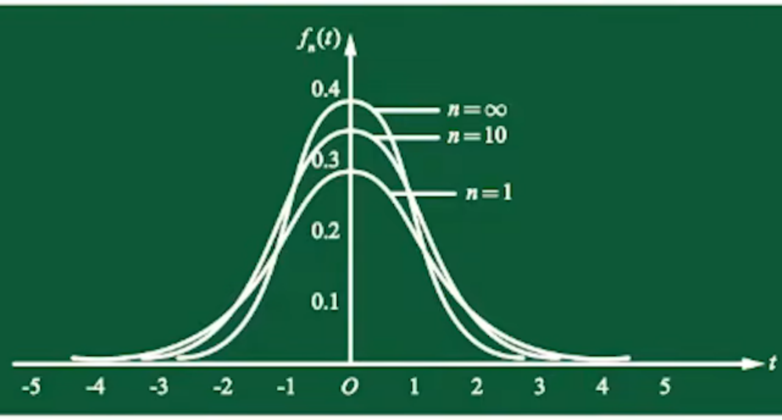
定义 设X是随机变量，对于给定的正数α $(0<\alpha<1)$，称满足$P\{X>\lambda\}=\alpha$的点λ为随机变量X的分布的**上侧α分位数**，或**上侧临界值**。称满足$P\{|X|>\lambda\}=\alpha$的点λ为随机变量X的分布的**双侧α分位数**，或**双侧临界值**。



注：当n充分大时，$\chi_\alpha^2(n)\approx\frac{1}{2}(u_\alpha+\sqrt{2n-1})^2$

u_α是标准正态分布的上α分位点。

T统计量及其分布

定义	设$X\sim N(0,1), Y\sim\chi^2(n)$，$X,Y$相互独立，则称$T=\frac{X}{\sqrt{Y/n}}$服从自由度为$n$的$t$分布，记$T\sim t(n)$.
图像	
性质	当$n>45$时，有$T=\frac{X}{\sqrt{Y/n}}\overset{近似}{\sim}N(0,1)$.

T分布的分位点

t分布的分位点

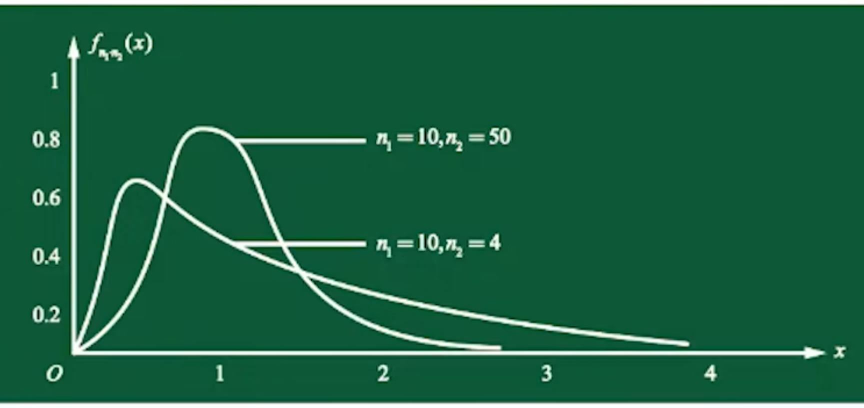
对给定α (0<α<1)，称满足 $P\{t>t_\alpha(n)\}=\int_{t_\alpha(n)}^{\infty}f(x)dt=\alpha$

的点$t_\alpha(n)$ 为$t(n)$分布的**上侧α 分位数**.

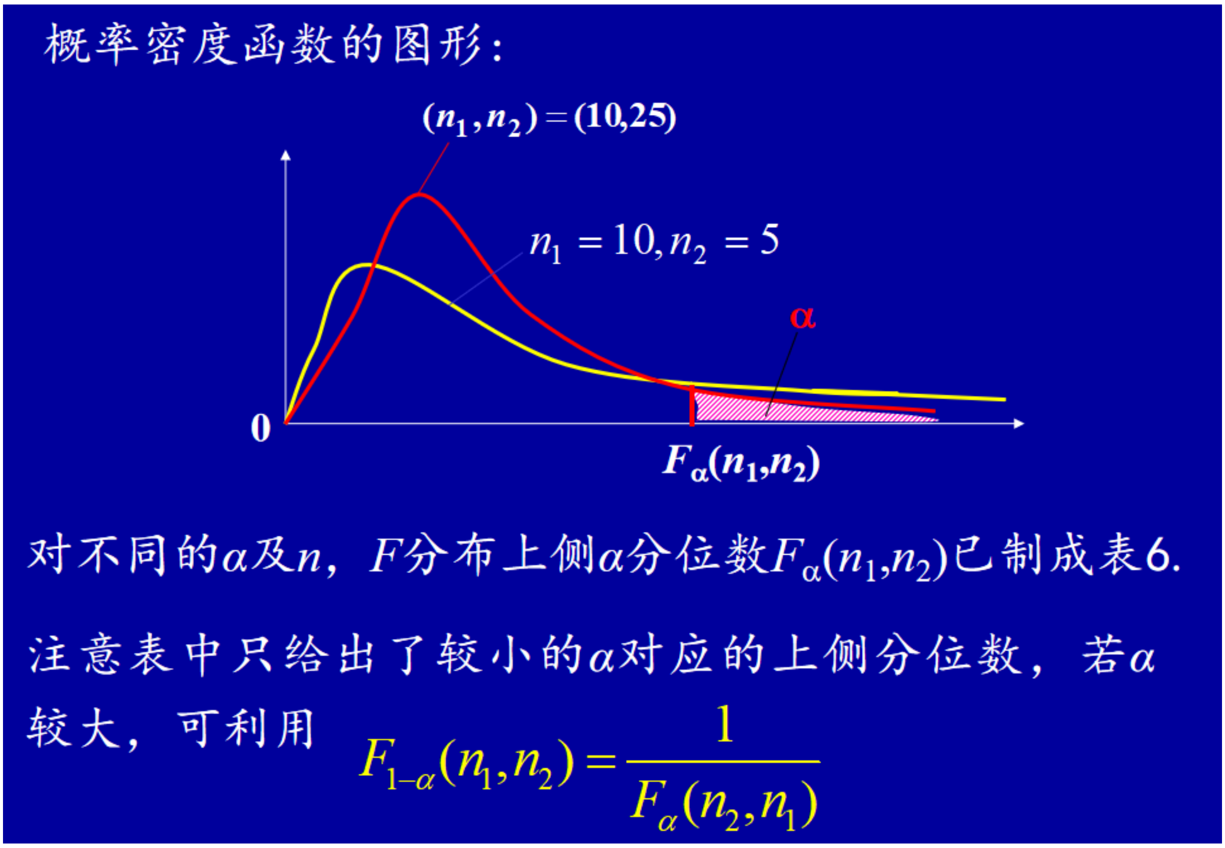
- 对不同的α及n，t分布的上侧α分位数已制成附表4.
- n>45时，无表可查，则可用标准正态分布近似
  $$t_\alpha(n)=u_\alpha, n>45$$
- 由于t分布的概率密度函数关于纵轴对称，因此实际中用得更多的是双侧分位数$t_{\alpha/2}(n)$，注意

$$P\{|t(n)|>t_{\alpha/2}(n)\}=\alpha \text{ 等价于 } P\{t(n)>t_{\alpha/2}(n)\}=\alpha/2$$

F统计量及其分布

定义	设$U\sim\chi_1^2(n_1), V\sim\chi_2^2(n_2)$，$U,V$相互独立，则称$F=\dfrac{U/n_1}{V/n_2}$服从自由度为$(n_1,n_2)$的$F$分布，记$F\sim F(n_1,n_2)$，其中$n_1$称为第一自由度，$n_2$称为第二自由度。
图像	
性质	$F\sim F(n_1,n_2)$，则$\dfrac{1}{F}\sim F(n_2,n_1)$.

F分布的上侧α分位数



正态总体样本均值与样本方差的分布

设$X_1, X_2, \ldots, X_n$来自正态总体$X \sim N(\mu, \sigma^2)$，$\overline{X}$表示样本均值，S^2表示样本方差，则近似有：

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

扩展 若$E(X) = \mu$，$D(X) = \sigma^2$，实际上有：

$$E(\overline{X}) = \mu$$
$$D(\overline{X}) = \frac{\sigma^2}{n}$$

Γ函数

补充知识：伽玛函数$\Gamma(\alpha)$

$$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}e^{-x}dx \quad 或\ \Gamma(\alpha)=2\int_0^{+\infty}x^{2\alpha-1}e^{-x^2}dx$$

作用：用来快速计算类似伽玛函数的积分

方法：$\Gamma(1)=\int_0^{+\infty}e^{-x}dx=1 \qquad \Gamma(\frac{1}{2})=2\int_0^{+\infty}e^{-x^2}dx=\sqrt{\pi}$

$\Gamma(s+1)=s\Gamma(s) \qquad \Gamma(n)=(n-1)!$

例如，$\int_0^{+\infty}x^2e^{-x}dx=\int_0^{+\infty}x^{3-1}e^{-x}dx=\Gamma(3)=2$

$$\int_{-\infty}^{+\infty}x^2e^{-x^2}dx=2\int_0^{+\infty}x^{2\times\frac{3}{2}-1}e^{-x^2}dx=\Gamma(\frac{3}{2})=\frac{1}{2}\Gamma(\frac{1}{2})=\frac{\sqrt{\pi}}{2}$$

参数估计 重点

点估计

点估计：构造合适的统计量，估计总体的未知参数

注意 估计值与估计量的区别

矩估计法

用样本均值估计总体均值：

设样本均值：

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

总体均值为μ，总体均值估计量为$\hat{\mu}$，记$\hat{\mu}=\overline{X}$

计算未知参数θ的矩估计套路

1. $\mu=E(X)=g(\theta)$
2. 令$\mu=g(\theta)=\overline{X}$，解得$\hat{\theta}=h(\overline{X})$，称为$\theta$的估计量
3. 用样本观测值求出$\overline{X}$，代入$h(\overline{X})$，得到θ的估计值

最大似然估计法

一个试验中，得到了一组样本观察值，则有理由相信这个观察值出现的概率最大。因此，考察未知参数θ取何值时，这组样本观察值出现的概率最大，那么就用这个值作为θ的（最大似然）估计值。

计算未知参数θ的最大似然估计套路

1. 写出似然函数（给定样本出现的概率）
2. 对似然函数取对数（方便求导）
3. 求导，令导函数等于0，反解出θ，得到的即为θ的估计值（如果是估计量则用随机变量表示）

求解极大似然估计的一般步骤：

(1) 由总体的分布写出似然函数$L(\theta)$;

(2)建立似然方程 $\dfrac{\partial L(\theta)}{\partial \theta_i} = 0$或 $\dfrac{\partial \ln L(\theta)}{\partial \theta_i} = 0$

(3)解上述似然方程。

似然估计的性质

如果$\hat{\theta}$为参数θ的极大似然估计量，又函数$g = g(\theta)$具有单值反函数，则$g(\hat{\theta})$是$g(\theta)$的极大似然估计量

估计量的评估标准

无偏性

定义 设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 是未知参数 θ 的估计量，如果

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

则称 $\hat{\theta}$ 是 θ 的**无偏估计量**.

记 $b_n=E(\hat{\theta})-\theta$，称 b_n 为估计量 $\hat{\theta}$ 的偏差.

当 $b_n\neq 0$，称 $\hat{\theta}$ 是 θ 的**有偏估计**.

若 $\lim\limits_{n\to\infty} b_n=0$，称 $\hat{\theta}$ 是 θ 的**渐近无偏估计**.

- 不论总体X服从什么分布，当总体k阶矩存在时，样本k阶原点矩A_k是总体k阶原点矩的无偏估计。特别地，样本均值$\overline{X}$是总体均值$E(X)$的无偏估计量。

区间估计

定义 设总体$X\sim F(x;\theta)$，θ是未知参数，$X_1,X_2,\ldots,X_n$是来自总体的样本，对于给定值$\alpha\,(0<\alpha<1)$，如果存在两个统计量 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)$

满足 $$P\{\hat{\theta}_1<\theta<\hat{\theta}_2\}\geq 1-\alpha$$

则称随机区间$(\hat{\theta}_1,\hat{\theta}_2)$是$\theta$的置信度为$1-\alpha$的置信区间，$1-\alpha$为置信水平，$\hat{\theta}_1$和$\hat{\theta}_2$分别称为置信下限和置信上限.

评价置信区间好坏标准:

(1)精度：$\hat{\theta}_2-\hat{\theta}_1$越小越好；

(2)置信度：越大越好.

- 区间估计要求根据样本给出未知参数的一个范围，并保证参数的真值以指定的较大概率属于这个范围。即以区间的形式给出未知参数的一个范围, 同时还给出此区间包含参数真值的可信程度。这种形式的估计称为区间估计。
- $1-\alpha$称为置信度或置信水平
- 置信度的理解：假设$1-\alpha=0.95$，则经过100次抽样得到的100个估计区间中约有95个包含了真实值。（以总体均值μ的估计为例，μ是确定的一个未知常数（不会变化），但每次抽样得到的样本均值$\overline{X}$都在变化）

构造置信区间的一般方法

构造置信区间的基本思想：在点估计的基础上，构造合适的含样本及待估参数的函数U，且已知U的分布，再根据给定的置信度导出待估参数置信区间.

➤**一般步骤：**

(1)明确是哪个参数的置信区间？置信水平$1-\alpha$是？

(2)选取未知参数θ的某个较优点估计$\hat{\theta}$；

(3) 构造待估参数θ和估计量$\hat{\theta}$的函数U，且其分布已知；

(4)对给定的置信水平$1-\alpha$，确定常数c和d，使

$$P\{c<U<d\}=1-\alpha$$

(5)对上式作恒等变形，化为$P\{\hat{\theta}_1<\theta<\hat{\theta}_2\}=1-\alpha$

则$(\hat{\theta}_1-\hat{\theta}_2)$就是$\theta$的置信水平为$1-\alpha$的置信区间.

注意 σ表示总体标准差，S表示样本标准差

单个正态总体参数的置信区间

参数	条件	置信度为$1-\alpha$的双侧置信区间
μ	σ^2已知	$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}\mu_{\alpha/2}, \overline{X}+\frac{\sigma}{\sqrt{n}}\mu_{\alpha/2}\right)$
	σ^2未知	$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1), \overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$
σ^2	μ未知	$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$

均值μ的置信区间

- 方差σ^2已知

由:

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

$$P\{|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}| < Z_{\alpha/2}\} = 1-\alpha$$

则有:

$$P\{\overline{X}-\frac{\sigma}{\sqrt{n}}Z_{\alpha/2} < \mu < \overline{X}+\frac{\sigma}{\sqrt{n}}Z_{\alpha/2}\} = 1-\alpha$$

即置信度为$1-\alpha$的置信区间为:

$$(\overline{X} \pm \frac{\sigma}{\sqrt{n}}Z_{\alpha/2})$$

- 方差σ^2未知

由：

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

$$P\{-t_{\alpha/2}(n-1)<T<t_{\alpha/2}(n-1)\}=1-\alpha$$

则有：

$$P\{\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)<\mu<\overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\}=1-\alpha$$

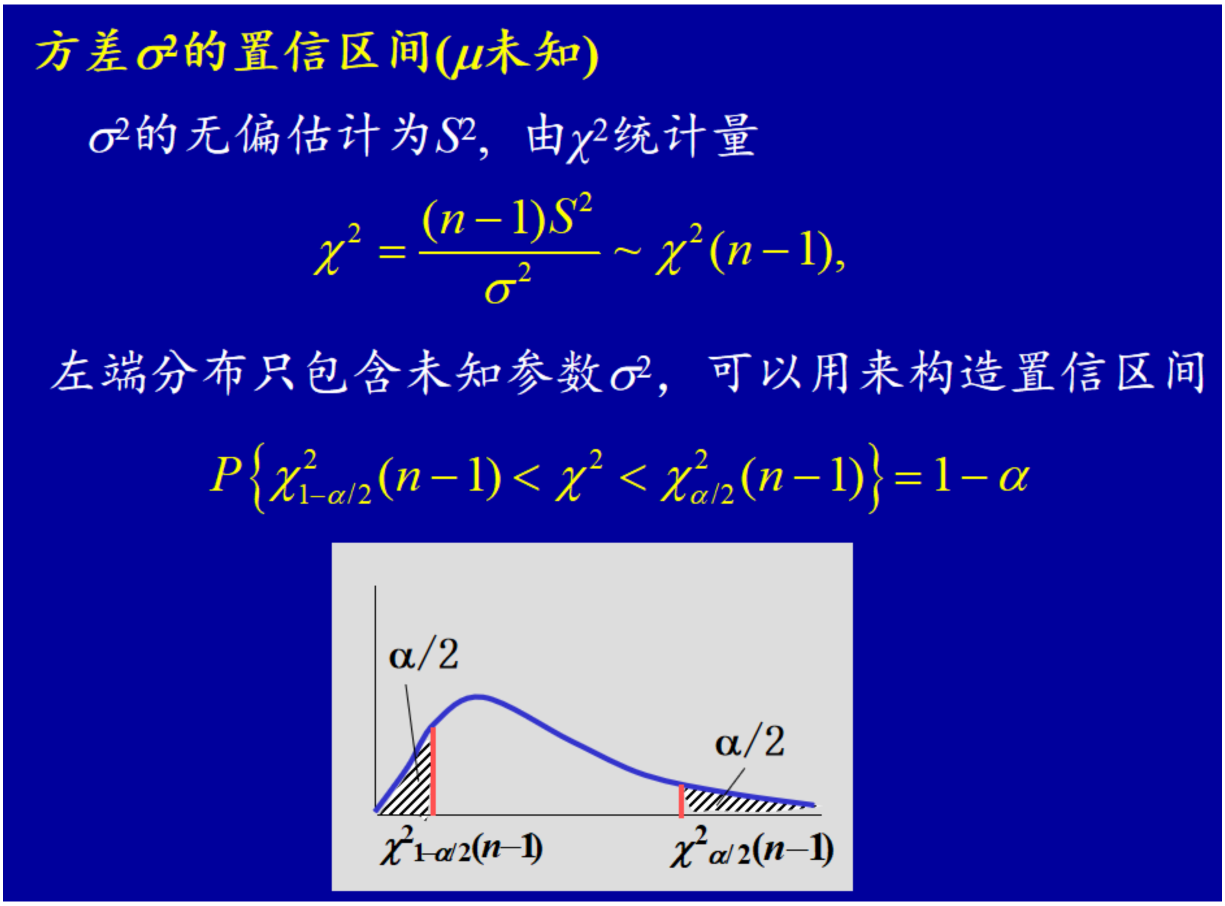
由此得到σ^2未知时，μ的置信度为$1-\alpha$的置信区间为：

$$(\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1),\ \overline{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1))$$

简写为：

$$\left(\overline{X}\pm\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$$

方差σ^2的置信区间



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

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

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

得到方差σ^2的一个置信度为$1-\alpha$的置信区间:

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

注：标准差σ的$1-\alpha$置信区间为$\left(\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{\alpha/2}(n-1)}}, \frac{\sqrt{n-1}S}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\right)$

假设检验 重点

假设检验的步骤

3.假设检验的步骤

(1)据题意写出原假设H_0和备择假设H_1；

(2)选择检验方法，写出检验统计量及其分布；

(3)根据给定的显著性水平确定拒绝域；

(4)计算检验统计量的值，做出推断.

假设检验情况分类

检验参数	条件	原假设与备择假设	检验法及检验统计量	拒绝域
μ	$\sigma^2=\sigma_0^2$ 已知	双边检验 $H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$ 单边检验 $H_0:\mu\le\mu_0,H_1:\mu>\mu_0$ $H_0:\mu\ge\mu_0,H_1:\mu<\mu_0$	U 检验 $U=\dfrac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}\sim N(0,1)$	$\lvert u\rvert\ge u_{\alpha/2}$ $u\ge u_\alpha$ $u\le -u_\alpha$
	σ^2未知	$H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$ $H_0:\mu\le\mu_0,H_1:\mu>\mu_0$ $H_0:\mu\ge\mu_0,H_1:\mu<\mu_0$	T 检验 $T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$	$\lvert t\rvert\ge t_{\alpha/2}(n-1)$ $t\ge t_\alpha(n-1)$ $t\le -t_\alpha(n-1)$
σ^2	μ 已知	$H_0:\sigma^2=\sigma_0^2,H_1:\sigma^2\neq\sigma_0^2$ $H_0:\sigma^2\le\sigma_0^2,H_1:\sigma^2>\sigma_0^2$ $H_0:\sigma^2\ge\sigma_0^2,H_1:\sigma^2<\sigma_0^2$	χ^2 检验 $\chi^2=\dfrac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$	$\chi^2\le\chi^2_{1-\alpha/2}(n)$或 $\chi^2\ge\chi^2_{\alpha/2}(n)$ $\chi^2\ge\chi^2_{\alpha}(n)$ $\chi^2\le\chi^2_{1-\alpha}(n)$
	μ 未知	$H_0:\sigma^2=\sigma_0^2,H_1:\sigma^2\neq\sigma_0^2$ $H_0:\sigma^2\le\sigma_0^2,H_1:\sigma^2>\sigma_0^2$ $H_0:\sigma^2\ge\sigma_0^2,H_1:\sigma^2<\sigma_0^2$	χ^2 检验 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$\chi^2\ge\chi^2_{\alpha/2}(n-1)$或 $\chi^2\le\chi^2_{1-\alpha/2}(n-1)$ $\chi^2\ge\chi^2_{\alpha}(n-1)$ $\chi^2\le\chi^2_{1-\alpha}(n-1)$