一．随机事件和概率

1、概率的定义和性质

（1）概率的公理化定义

设Ω为样本空间，A为事件，对每一个事件A都有一个实数P(A)，若满足下列三个条件：

1° 0≤P(A)≤1，

2° P(Ω) =1

3° 对于两两互不相容的事件A_1，A_2，…有

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

常称为可列（完全）可加性。

则称P(A)为事件A的概率。

（2）古典概型（等可能概型）

1° $\Omega=\{\omega_1,\omega_2\cdots\omega_n\}$，

2° $P(\omega_1)=P(\omega_2)=\cdots P(\omega_n)=\frac{1}{n}$。

设任一事件A，它是由$\omega_1,\omega_2\cdots\omega_m$组成的，则有

$P(A)=\{(\omega_1)\cup(\omega_2)\cup\cdots\cup(\omega_m)\}$

$=P(\omega_1)+P(\omega_2)+\cdots+P(\omega_m)$

$$=\frac{m}{n}=\frac{A\text{所包含的基本事件数}}{\text{基本事件总数}}$$

2、五大公式（加法、减法、乘法、全概、贝叶斯）

（1）加法公式

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

当P(AB)＝0时，P(A+B)=P(A)+P(B)

（2）减法公式

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

当B⊂A时，P(A-B)=P(A)-P(B)

当A=Ω时，P($\overline{B}$)=1- P(B)

（3）条件概率和乘法公式

定义 设A、B是两个事件，且P(A)>0，则称$\frac{P(AB)}{P(A)}$为事件A发生条件下，事件B发生的条件概率，记为$P(B/A)=\frac{P(AB)}{P(A)}$。

条件概率是概率的一种，所有概率的性质都适合于条件概率。

（4）全概公式

设事件$B_1,B_2,\cdots,B_n$满足

1° $B_1,B_2,\cdots,B_n$两两互不相容，$P(B_i)>0(i=1,2,\cdots,n)$，

2° $A\subset\bigcup_{i=1}^{n}B_i$，

则有

$P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\cdots+P(B_n)P(A|B_n)$。

此公式即为全概率公式。

（5）贝叶斯公式

设事件B_1，B_2，…，B_n及A满足

1° B_1，B_2，…，B_n两两互不相容，$P(Bi)$>0，$i=1$，2，…，n，

2° $A\subset\bigcup_{i=1}^{n}B_i$，$P(A)>0$，

则

$$P(B_i/A)=\frac{P(B_i)P(A/B_i)}{\sum_{j=1}^{n}P(B_j)P(A/B_j)}$$，i=1，2，…n。

此公式即为贝叶斯公式。

$P(B_i)$，（$i=1$，2，…，n），通常叫先验概率。$P(B_i/A)$，（$i=1$，2，…，n），通常称为后验概率。如果我们把A当作观察的“结果”，而B_1，B_2，…，B_n理解为“原因”，则贝叶斯公式反映了“因果”的概率规律，并作出了“由果朔因”的推断。

3、事件的独立性和伯努利试验

（1）两个事件的独立性

设事件A、B满足$P(AB)=P(A)P(B)$，则称事件A、B是相互独立的（这个性质不是想当然成立的）。

若事件A、B相互独立，且$P(A)>0$，则有

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

所以这与我们所理解的独立性是一致的。

若事件A、B相互独立，则可得到$\overline{A}$与B、A与$\overline{B}$、$\overline{A}$与$\overline{B}$也都相互独立。（证明）

由定义，我们可知必然事件Ω和不可能事件Ø与任何事件都相互独立。（证明）

同时，Ø与任何事件都互斥。

（2）多个事件的独立性

设ABC是三个事件，如果满足两两独立的条件，

P(AB)=P(A)P(B)；P(BC)=P(B)P(C)；P(CA)=P(C)P(A)

并且同时满足 P(ABC)=P(A)P(B)P(C)

那么 A、B、C 相互独立。

对于 n 个事件类似。

两两互斥→互相互斥。

两两独立→互相独立？

（3）伯努利试验

定义 我们作了n次试验，且满足

- ◆ 每次试验只有两种可能结果，A发生或A不发生；
- ◆ n次试验是重复进行的，即A发生的概率每次均一样；
- ◆ 每次试验是独立的，即每次试验A发生与否与其他次试验A发生与否是互不影响的。

这种试验称为伯努利概型，或称为n重伯努利试验。

用p表示每次试验A发生的概率，则$\overline{A}$发生的概率为$1-p=q$，用$P_n(k)$表示n重伯努利试验中A出现$k(0\le k\le n)$次的概率，

二. 随机变量及其分布

1、随机变量的分布函数

（1）离散型随机变量的分布率

设离散型随机变量X的可能取值为 X_k(k=1,2,…)且取各个值的概率，即事件($X=X_k$)的概率为

$P(X=x_k)=p_k$，k=1,2,…，

则称上式为离散型随机变量X的概率分布或分布律。有时也用分布列的形式给出：

$$\frac{X}{P(X=x_k)}\Big|\frac{x_1,x_2,\Lambda,x_k,\Lambda}{p_1,p_2,\Lambda,p_k,\Lambda}。$$

显然分布律应满足下列条件：

（1）$p_k\ge 0$，$k=1,2,\Lambda$，

（2）$\sum_{k=1}^{\infty}p_k=1$。

（2）分布函数

对于非离散型随机变量，通常有$P(X=x)=0$，不可能用分布率表达。例如日光灯管的寿命X，$P(X=x_0)=0$。

所以我们考虑用X落在某个区间$(a,b]$内的概率表示。

定义 设X为随机变量，x是任意实数，则函数

$$F(x)=P(X\le x)$$

称为随机变量 X 的分布函数。

$P(a<X\le b)=F(b)-F(a)$ 可以得到 X 落入区间$(a,b]$的概率。也就是说，分布函数完整地描述了随机变量 X 随机取值的统计规律性。

分布函数$F(x)$是一个普通的函数，它表示随机变量落入区间（–∞，x]内的概率。

$F(x)$的图形是阶梯图形，x_1,x_2,Λ是第一类间断点，随机变量X在x_k处的概率就是$F(x)$在x_k处的跃度。

分布函数具有如下性质：

1° $0\le F(x)\le 1,\quad -\infty<x<+\infty$；

2° $F(x)$是单调不减的函数，即$x_1<x_2$时，有$F(x_1)\le F(x_2)$；

3° $F(-\infty)=\lim_{x\to-\infty}F(x)=0$，$F(+\infty)=\lim_{x\to+\infty}F(x)=1$；

4° $F(x+0)=F(x)$，即$F(x)$是右连续的；

5° $P(X=x)=F(x)-F(x-0)$。

（3）连续型随机变量的密度函数

定义 设$F(x)$是随机变量X的分布函数，若存在非负函数$f(x)$，对任意实数x，有

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

则称X为连续型随机变量。$f(x)$称为X的概率密度函数或密度函数，简称概率密度。$f(x)$的图形是一条曲线，称为密度（分布）曲线。

由上式可知，连续型随机变量的分布函数$F(x)$是连续函数。

所以，

$$P(x_1\le X\le x_2)=P(x_1<X\le x_2)=P(x_1\le X<x_2)=P(x_1<X<x_2)=F(x_2)-F(x_1)$$

密度函数具有下面 4 个性质：

1° $f(x)\ge 0$。

2° $\int_{-\infty}^{+\infty}f(x)dx=1$。

$F(+\infty)=\int_{-\infty}^{+\infty} f(x)dx=1$ 的几何意义；在横轴上面、密度曲线下面的全部面积等于 1。

如果一个函数 $f(x)$ 满足 1°、2°，则它一定是某个随机变量的密度函数。

3° $P(x_1<X\le x_2)=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)dx$。

4° 若 $f(x)$ 在 x 处连续，则有 $F'(x)=f(x)$。

$P(x<X\le x+dx)\approx f(x)dx$

它在连续型随机变量理论中所起的作用与 $P(X=x_k)=p_k$ 在离散型随机变量理论中所起的作用相类似。

$E\to\omega,\Omega\to \quad A \quad \to P(A),$(古典概型，五大公式，独立性)

$X(\omega)\to X(\omega)\le x\to F(x)=P(X\le x)$

对于连续型随机变量 X，虽然有 $P(X=x)=0$，但事件 $(X=x)$ 并非是不可能事件 Ø。

$P(X=x)\le P(x<X\le x+h)=\int_{x}^{x+h} f(x)dx$

令 $h\to 0$，则右端为零，而概率 $P(X=x)\ge 0$，故得 $P(X=x)=0$。

不可能事件（Ø）的概率为零，而概率为零的事件不一定是不可能事件；同理，必然事件（Ω）的概率为 1，而概率为 1 的事件也不一定是必然事件。

2、常见分布

①0－1 分布

P(X=1)=p, P(X=0)=q

②二项分布

在 n 重贝努里试验中，设事件 A 发生的概率为 p。事件 A 发生的次数是随机变量，设为 X，则 X 可能取值为 $0,1,2,\Lambda,n$。

$P(X=k)=P_n(k)=C_n^k p^k q^{n-k}$，其中 $q=1-p,0<p<1,k=0,1,2,\Lambda,n$，

则称随机变量 X 服从参数为 n，p 的二项分布。记为 $X\sim B(n,p)$。

$$\frac{X}{P(X=k)}\left|\frac{}{q^n,npq^{n-1},C_n^2p^2q^{n-2},\Lambda,C_n^kp^kq^{n-k},\Lambda,p^n}\right.$$

容易验证，满足离散型分布率的条件。

当 $n=1$ 时，$P(X=k)=p^kq^{1-k}$，$k=0.1$，这就是(0-1)分布，所以（0-1）分布是二项分布的特例。

③泊松分布

设随机变量 X 的分布律为

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

则称随机变量 X 服从参数为 λ 的泊松分布，记为 $X\sim\pi(\lambda)$ 或者 P(λ)。

泊松分布为二项分布的极限分布（np=λ，n→∞）。

④超几何分布

$$P(X=k)=\frac{C_M^k\bullet C_{N-M}^{n-k}}{C_N^n},\quad \begin{array}{l}k=0,1,2\Lambda,l\\ l=\min(M,n)\end{array}$$

随机变量 X 服从参数为 n,N,M 的超几何分布。

⑤几何分布

$P(X=k)=q^{k-1}p,k=1,2,3,\Lambda$，其中 p≥0，q=1-p。

随机变量 X 服从参数为 p 的几何分布。

⑥均匀分布

设随机变量 X 的值只落在[a，b]内，其密度函数 $f(x)$ 在[a，b]上为常数 k，即

$$f(x)=\begin{cases}k, & a\le x\le b\\ 0, & 其他,\end{cases}$$

其中 k= $\frac{1}{b-a}$，

则称随机变量 X 在[a，b]上服从均匀分布，记为 X~U(a，b)。

分布函数为

$$\begin{cases}0, & x<a,\\ \frac{x-a}{b-a}, & a\le x\le b\end{cases}$$

$F(x)=\int_{-\infty}^{x} f(x)dx=$

1，　　　　x>b。

当 a≤x_1<x_2≤b 时，X 落在区间（x_1,x_2）内的概率为

P（$x_1<X<x_2)=\int_{x_1}^{x_2} f(x)dx=\int_{x_1}^{x_2}\frac{1}{b-a}dx=\frac{x_2-x_1}{b-a}$ 。

⑦指数分布

设随机变量 X 的密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0,\\ 0, & x<0,\end{cases}$$

其中 $\lambda>0$，则称随机变量 X 服从参数为 λ 的指数分布。X 的分布函数为

$$F(x)=\begin{cases}1-e^{-\lambda x}, & x\ge 0,\\ 0, & x<0。\end{cases}$$

记住几个积分：

$$\int_0^{+\infty} xe^{-x}dx=1,\qquad \int_0^{+\infty} x^2e^{-x}dx=2,$$

$$\int_0^{+\infty} x^{n-1}e^{-x}dx=(n-1)!$$

$$\Gamma(\alpha)=\int_0^{+\infty} x^{\alpha-1}e^{-x}dx,\qquad \Gamma(\alpha+1)=\alpha\Gamma(\alpha)$$

⑧正态分布

设随机变量 X 的密度函数为

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

其中 μ 、$\sigma>0$ 为常数，则称随机变量 X 服从参数为 μ 、σ 的正态分布或高斯（Gauss）分布，记为 $X\sim N(\mu,\sigma^2)$ 。

$f(x)$ 具有如下性质：

1° $f(x)$ 的图形是关于 $x=\mu$ 对称的；

2° 当 $x=\mu$ 时，$f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$ 为最大值；

3° $f(x)$ 以 ox 轴为渐近线。

特别当 σ 固定、改变 μ 时，$f(x)$ 的图形形状不变，只是集体沿 ox 轴平行移动，所以 μ 又称为位置参数。当 μ 固定、改变 σ 时，$f(x)$ 的图形形状要发生变化，随 σ 变大，$f(x)$ 图形的形状变得平坦，所以又称 σ 为形状参数。

若 $X\sim N(\mu,\sigma^2)$，则 X 的分布函数为

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$$ 。。

参数 $\mu=0$ 、$\sigma=1$ 时的正态分布称为标准正态分布，记为 $X\sim N(0,1)$，其密度函数记为

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

分布函数为

$\Phi(x)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$ 。$\Phi(x)$ 是不可求积函数，其函数值，已编制成表可供查用。

φ(x)和 Φ(x)的性质如下：

1° φ(x)是偶函数，φ(x)＝φ(−x)；

2° 当 x=0 时，φ(x)＝$\frac{1}{\sqrt{2\pi}}$ 为最大值；

3° Φ(−x)＝1−Φ(x)且 Φ(0)＝$\frac{1}{2}$ 。

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

所以我们可以通过变换将 $F(x)$ 的计算转化为 $\Phi(x)$ 的计算，而 $\Phi(x)$ 的值是可以通过查表得到的。

$$P(x_1<X\le x_2)=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right)。$$

分位数的定义

3、随机变量函数的分布

随机变量 Y 是随机变量 X 的函数 $Y=g(X)$，若 X 的分布函数 $F_X(x)$ 或密度函数 $f_X(x)$ 知道，则如何求出 $Y=g(X)$ 的分布函数 $F_Y(y)$ 或密度函数 $f_Y(y)$ 。

（1）X 是离散型随机变量

已知 X 的分布列为

$$\begin{array}{c|c} X & x_1,\ x_2,\ \Lambda,\ x_n,\ \Lambda \\ \hline P(X=x_i) & p_1,\ p_2,\ \Lambda,\ p_n,\ \Lambda \end{array},$$

显然，$Y=g(X)$ 的取值只可能是 $g(x_1),g(x_2),\Lambda,g(x_n),\Lambda$，若 $g(x_i)$ 互不相等，则 Y 的分布列如下：

$$\begin{array}{c|c} Y & g(x_1),\ g(x_2),\ \Lambda,\ g(x_n),\ \Lambda \\ \hline P(Y=y_i) & p_1,\ p_2,\ \Lambda,\ p_n,\ \Lambda \end{array},$$

若有某些 $g(x_i)$ 相等，则应将对应的 P_i 相加作为 $g(x_i)$ 的概率。

（2） X 是连续型随机变量

先利用 X 的概率密度 $f_X(x)$ 写出 Y 的分布函数 $F_Y(y)$，再利用变上下限积分的求导公式求出 $f_Y(y)$。

三. 二维随机变量及其分布

1、二维随机变量的基本概念

（1）二维连续型随机向量联合分布密度及边缘分布

对于二维随机向量 $\xi=(X,Y)$，如果存在非负函数 $f(x,y)(-\infty<x<+\infty,-\infty<y<+\infty)$，使对任意一个其邻边分别平行于坐标轴的矩形区域 D，即 D={(X,Y)|a<x<b,c<y<d}有

$$P\{(X,Y)\in D\}=\iint_D f(x,y)dxdy,$$

则称 ξ 为连续型随机向量；并称 f(x,y)为 ξ=（X，Y）的分布密度或称为 X 和 Y 的联合分布密度。

分布密度 f(x,y)具有下面两个性质：

（1） f(x,y)≥0;

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

一般来说，当（X，Y）为连续型随机向量，并且其联合分布密度为 f(x,y)，则 X 和 Y 的边缘分布密度为

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

注意：联合概率分布→边缘分布

（2）条件分布

当（X，Y）为离散型，并且其联合分布律为

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

在已知 $X=x_i$ 的条件下，Y 取值的条件分布为

$$P(Y=y_j|X=x_i)=\frac{p_{ij}}{p_{i\bullet}},$$

其中 $p_{i\cdot}$, $p_{\cdot j}$ 分别为 X，Y 的边缘分布。

当(X,Y)为连续型随机向量,并且其联合分布密度为 f(x,y)，则在已知 Y=y 的条件下，X 的条件分布密度为

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

在已知 X=x 的条件下，Y 的条件分布密度为

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

其中 $f_X(x)>0, f_Y(y)>0$ 分别为 X，Y 的边缘分布密度。

（3）常见的二维分布

①均匀分布

设随机向量（X，Y）的分布密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{S_D} & (x,y)\in D\\ 0, & \text{其他}\end{cases}$$

其中 S_D 为区域 D 的面积，则称（X，Y）服从 D 上的均匀分布，记为（X，Y）～U（D）。

②正态分布

设随机向量（X，Y）的分布密度函数为

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

其中 $\mu_1,\mu_2,\sigma_1>0,\sigma_2>0,|\rho|<1$，共 5 个参数，则称（X，Y）服从二维正态分布，

记为（X，Y）～N（$\mu_1,\mu_{2,}\sigma_1^2,\sigma_2^2,\rho$).

由边缘密度的计算公式，可以推出二维正态分布的两个边缘分布仍为正态分布，反推则错。

即 X～N（$\mu_1,\sigma_1^2),Y\sim N(\mu_{2,}\sigma_2^2)$.

（5）二维随机向量联合分布函数及其性质

设（X，Y）为二维随机变量，对于任意实数 x,y,二元函数

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

称为二维随机向量（X，Y）的分布函数，或称为随机变量 X 和 Y 的联合分布函数。

分布函数是一个以全平面为其定义域，以事件 $\{(\omega_1,\omega_2)|-\infty<X(\omega_1)\le x,-\infty<Y(\omega_2)\le y\}$ 的概率为函数值的一个实值函数。分布函数 F(x,y)具有以下的基本性质：

（1）$0\le F(x,y)\le 1;$

（2）F（x,y）分别对 x 和 y 是非减的，即
当 $x_2>x_1$ 时，有 $F(x_2,y)\geq F(x_1,y)$；当 $y_2>y_1$ 时，有 $F(x,y_2)\geq F(x,y_1)$；
（3）F（x,y）分别对 x 和 y 是右连续的，即

$$F(x,y)=F(x+0,y),F(x,y)=F(x,y+0);$$

（4）

$$F(-\infty,-\infty)=F(-\infty,y)=F(x,-\infty)=0,F(+\infty,+\infty)=1.$$

2、随机变量的独立性

（1）连续型随机变量

$f(x,y)=f_X(x)f_Y(y)$
联合分布→边缘分布→$f(x,y)=f_X(x)f_Y(y)$

直接判断，充要条件：
①可分离变量
②正概率密度区间为矩形

（2）二维正态分布

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

ρ=0

（3）随机变量函数的独立性

若 X 与 Y 独立，h,g 为连续函数，则：h（X）和 g（Y）独立。

四. 随机变量的数字特征

（1）一维随机变量及其函数的期望

①设 X 是离散型随机变量，其分布律为 $P(X=x_k)=p_k$，k=1,2,…,n，

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

期望就是平均值。
②设 X 是连续型随机变量，其概率密度为 f(x)，

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

③数学期望的性质
（1） E(C)=C
（2） E(CX)=CE(X)
（3） E(X+Y)=E(X)+E(Y)，$E(\sum_{i=1}^{n}C_iX_i)=\sum_{i=1}^{n}C_iE(X_i)$
（4） E(XY)=E(X) E(Y)，充分条件：X 和 Y 独立；
充要条件：X 和 Y 不相关。
（5） Y=g(X)

离散：$E(Y)=\sum_{i=1}^{n}g(x_k)p_k$

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

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

（2）方差

$D(X)=E[X-E(X)]^2$，方差

$\sigma(X)=\sqrt{D(X)}$，标准差

①离散型随机变量

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

②连续型随机变量

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

③方差的性质
（1） D(C)=0；E(C)=C
（2） $D(aX)=a^2D(X)$； E(aX)=aE(X)
（3） $D(aX+b)=a^2D(X)$； E(aX+b)=aE(X)+b
（4） $D(X)=E(X^2)-E^2(X)$
（5） D(X+Y)=D(X)+D(Y)，充分条件：X 和 Y 独立；
充要条件：X 和 Y 不相关。
D(X±Y)=D(X)+D(Y)±2E[(X-E(X))(Y-E(Y))]，无条件成立。
E(X+Y)=E(X)+E(Y)，无条件成立。

（3）常见分布的数学期望和方差

分布名称	符号	均值	方差
0-1 分布	$B(1,p)$	p	$p(1-p)$
二项分布	$B(n,p)$	np	$np(1-p)$
泊松分布	$P(\lambda)$	λ	λ
几何分布	$G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$

超几何分布	$H(n,M,N)$	$\frac{nM}{N}$	$\frac{nM}{N}\left(1-\frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$
均匀分布	$U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布	$e(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
正态分布	$N(\mu,\sigma^2)$	μ	σ^2

①0－1 分布

X	0	1
	q	p

E(X)=p，D(X)=pq

②二项分布　X~B(n,p)，$P_n(k)=C_n^k p^k q^{n-k}$，(k=0,1,2…n)

E(X)=np，D(X)=npq

③泊松分布　P(λ)　P(X=k)=$\frac{\lambda^k e^{-x}}{k!}$，k=0,1,2…

E(X)= λ，　D(X)= λ

④超几何分布　$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$

E(X)=$\frac{nM}{N}$

⑤几何分布　$P(X=k)=pq^{k-1}$，k=0,1,2…

E(X)=$\frac{1}{p}$，　D(X)=$\frac{q}{p^2}$

⑥均匀分布　X~U[a,b]，f(x)=$\frac{1}{b-a}$，[a, b]

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

⑦指数分布　f(x)= $\lambda e^{-\lambda x}$，(x>0)

E(X)=$\frac{1}{\lambda}$，　D(X)=$\frac{1}{\lambda^2}$

⑧正态分布　X~N(μ，σ²)，$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

E(X)= μ，　D(X)= σ²

2、二维随机变量的数字特征

（1）协方差和相关系数

对于随机变量 X 与 Y，称它们的二阶混合中心矩μ_{11}为 X 与 Y 的协方差或相关矩，记为σ_{XY}或$\operatorname{cov}(X,Y)$，即

$$\sigma_{XY}=\mu_{11}=E[(X-E(X))(Y-E(Y))].$$

与记号σ_{XY}相对应，X 与 Y 的方差 D（X）与 D（Y）也可分别记为σ_{XX}与σ_{YY}。

协方差有下面几个性质：

(i)　cov (X, Y)=cov (Y, X);

(ii)　cov(aX,bY)=ab cov(X,Y);

(iii)　cov(X_1+X_2, Y)=cov(X_1,Y)+cov(X_2,Y);

(iv)　cov(X,Y)=E(XY)-(E(X))(E(Y)).

对于随机变量 X 与 Y，如果 D（X）>0, D(Y)>0，则称

$$\frac{\sigma_{XY}}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为 X 与 Y 的相关系数，记作ρ_{XY}（有时可简记为ρ）。

|ρ|≤1，当|ρ|=1 时，称 X 与 Y 安全相关：

$$完全相关\begin{cases}正相关，当\rho=1时，\\负相关，当\rho=-1时，\end{cases}$$

而当$\rho=0$时，称 X 与 Y 不相关。

与相关系数有关的几个重要结论

（i）　若随机变量 X 与 Y 相互独立，则$\rho_{XY}=0$；反之不真。

（ii）　若（X，Y）～N（$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$），则 X 与 Y 相互独立的充要条件是$\rho=0$，即 X 和 Y 不相关。

（iii）　以下五个命题是等价的：

①$\rho_{XY}=0$；

②cov(X,Y)=0；

③E(XY)=E(X)E(Y)；
④D(X+Y)=D(X)+D(Y)；
⑤D(X-Y)=D(X)+D(Y).

（2）二维随机变量函数的期望

$$E[G(X,Y)]=\begin{cases}\sum_i\sum_j G(x_i,y_j)p_{ij},(X,Y)\text{为离散型；}\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}G(x,y)f(x,y)dxdy,(X,Y)\text{为连续型。}\end{cases}$$

（3）原点矩和中心矩

①对于正整数 k，称随机变量 X 的 k 次幂的数学期望为 X 的 k 阶原点矩，记为 v_k,即

u_k=E(X^k)， k=1,2, ….

于是，我们有

$$u_k=\begin{cases}\sum_i x_i^k p_i & \text{当}X\text{为离散型时，}\\ \int_{-\infty}^{+\infty}x^k p(x)dx, & \text{当}X\text{为连续型时.}\end{cases}$$

②对于正整数 k，称随机变量 X 与 E（X）差的 k 次幂的数学期望为 X 的 k 阶中心矩，记为 μ_k，即

$$\mu_k=E(X-E(X))^k,k=1,2,\Lambda .$$

于是，我们有

$$u_k=\begin{cases}\sum_i (x_i-E(X))^k p_i & \text{当}X\text{为离散型时，}\\ \int_{-\infty}^{+\infty}(x-E(X))^k p(x)dx, & \text{当}X\text{为连续型时.}\end{cases}$$

③对于随机变量 X 与 Y，如果有 $E(X^kY^l)$ 存在，则称之为 X 与 Y 的 *k+l* 阶混合原点矩，记为 u_{kl}，即

$$u_{kl}=E[(X-E(X))^k(Y-E(Y))].$$

五. 大数定律和中心极限定理

1、切比雪夫不等式

设随机变量 X 具有数学期望 E（X）=μ，方差 D（X）=σ^2，则对于任意正数 ε，有下列切比雪夫不等式

$$P(|X-\mu|\ge\varepsilon)\le\frac{\sigma^2}{\varepsilon^2}$$

切比雪夫不等式给出了在未知 X 的分布的情况下，对概率

$$P(|X-\mu|\ge\varepsilon)$$

的一种估计，它在理论上有重要意义。

2、大数定律

（1）切比雪夫大数定律

（要求方差有界）

设随机变量 X_1，X_2，…相互独立，均具有有限方差，且被同一常数 C 所界：D（X_i）<C(i=1,2,…),则对于任意的正数 ε，有

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

特殊情形：若 X_1，X_2，…具有相同的数学期望 E（X_I）=μ，则上式成为

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

或者简写成：

$$\lim_{n\to\infty}P(|\overline{X}-\mu|<\varepsilon)=1.$$

切比雪夫大数定律指出，n 个相互独立，且具有有限的相同的数学期望与方差的随机变量，当 n 很大时，它们的算术平均以很大的概率接近它们的数学期望。

（2）伯努利大数定律

设μ是 n 次独立试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对于任意的正数 ε，有

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

伯努利大数定律说明，当试验次数 n 很大时，事件 A 发生的频率与概率有较大判别的可能性很小，即

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

这就以严格的数学形式描述了频率的稳定性。

（3）辛钦大数定律

（不要求存在方差）

设 X_1，X_2，…，X_n，…是相互独立同分布的随机变量序列，且 E（X_n）=μ，则对于任意的正数 ε 有

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

3、中心极限定理

（1）列维－林德伯格定理

设随机变量 X_1，X_2，…相互独立，服从同一分布，且具有相同的数学期望和方差：$E(X_k)=\mu, D(X_k)=\sigma^2\neq 0(k=1,2,\Lambda)$，则随机变量

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

的分布函数 $F_n(x)$ 对任意的实数 x，有

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

或者简写成：$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\xrightarrow[n\to\infty]{}N(0,1)$

此定理也称为**独立同分布**的中心极限定理。

（2）棣莫弗－拉普拉斯定理

设随机变量 X_1，…X_n 均为具有参数 n，p(0<p<1)的二项分布，则对于任意实数 x,有

$$=\lim_{n\to\infty}P\left\{\frac{X_n-np}{\sqrt{np(1-p)}}\leq x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt.$$

4、二项定理和泊松定理

（1）二项定理

若当 $N\to\infty\text{时},\frac{M}{N}\to p(n,k\text{不变})$，则

$$\frac{C_M^K C_{N-M}^{n-k}}{C_N^N}\to C_n^k P^k(1-p)^{n-k}\qquad (N\to\infty).$$

可见，超几何分布的极限分布为二项分布。

（2）泊松定理

若当 $n\to\infty\text{时},np\to\lambda>0$，则

$$C_n^k P^k(1-p)^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda}\qquad (n\to\infty).$$

其中 k=0，1，2，…，n，…。

六．数理统计的基本概念

1、总体、个体和样本

（1）总体与样本

总体　　　在数理统计中，常把被考察对象的某一个（或多个）指标的全体称为总体（或母体）；而把总体中的每一个单元称为样品（或个体）。在以后的讨论中，我们总是把总体看成一个具有分布的随机变量（或随机向量）。

（2）样本函数与统计量

设 x_1,x_2,Λ,x_n 为总体的一个样本，称

$$\varphi=\varphi\quad(x_1,x_2,\Lambda,x_n)$$

为样本函数，其中 φ 为一个连续函数。如果 φ 中不包含任何未知参数，则称 φ（x_1,x_2,Λ,x_n）为一个统计量。

2、统计量

（1）常用统计量

样本均值　　$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i.$

样本方差

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

（与概率论中的方差定义不同）

样本标准差　　$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}.$

样本 k 阶原点矩　　$M_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k,k=1,2,\Lambda.$

样本 k 阶中心矩

$$M'_k=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^k,k=2,3,\Lambda.$$

（二阶中心矩 $S^{*2}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 与概率论中的方差定义相同）

（2）统计量的期望和方差

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

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

其中$S^{*2}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$，为二阶中心矩。

3、三个抽样分布（χ²、t、F 分布）

（1）χ²分布

设 n 个随机变量$X_1,X_2,\Lambda\ ,X_n$相互独立，且服从**标准正态分布**，可以证明：它们的平方和

$$W=\sum_{i=1}^{n}X_i^2$$

的分布密度为

$$f(u)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}u^{\frac{n}{2}-1}e^{-\frac{u}{2}} & u\ge 0,\\ 0, & u<0.\end{cases}$$

我们称随机变量 W 服从自由度为 n 的κ^2分布，记为 W～κ^2(n)，其中

$$\Gamma\left(\frac{n}{2}\right)=\int_0^{+\infty}x^{\frac{n}{2}-1}e^{-x}dx.$$

所谓自由度是指独立正态随机变量的个数，它是随机变量分布中的一个重要参数。

κ^2 分布满足可加性：设

$$Y_i\ -\kappa^2(n_i),$$

则

$$Z=\sum_{i=1}^{k}Y_i\sim\kappa^2(n_1+n_2+\Lambda\ +n_k).$$

注意两个结果：E(χ²)=n，D(χ²)=2n

（2）t 分布

设 X，Y 是两个相互独立的随机变量，且

$$X\sim N(0,1),Y\sim\kappa^2(n),$$

可以证明：函数

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

的概率密度为

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

我们称随机变量 T 服从自由度为 n 的 t 分布，记为 T～t(n)。

注意两个结果：E(T)=0，D(T)=$\frac{n}{n-2}$ (n>2)

（3）F 分布

设$X\sim\kappa^2(n_1),Y\sim\kappa^2(n_2)$，且 X 与 Y 独立，可以证明：

$F=\frac{X/n_1}{Y/n_2}$的概率密度函数为

$$f(y)=\begin{cases}\dfrac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\dfrac{n_1}{n_2}\right)^{\frac{n_1}{2}}y^{\frac{n_1}{2}-1}\left(1+\dfrac{n_1}{n_2}y\right)^{-\frac{n_1+n_2}{2}}, \\ 0,\end{cases}$$

我们称随机变量 F 服从第一个自由度为 n_1，第二个自由度为 n_2 的 F 分布，记为 F～f(n_1，n_2)。

正态分布$\mu_{1-\alpha}=-\mu_\alpha$，

$t_{1-\alpha}(n)=-t_\alpha(n)$，

$$F_{1-\alpha}(n_1,n_2)=\frac{1}{F_\alpha(n_2,n_1)}$$

4、正态总体下统计量的分布和性质

注意一个定理：$\overline{X}$与S^2独立。

（1）正态分布　　设$x_1,x_2,\Lambda\ ,x_n$为来自正态总体$N(\mu,\sigma^2)$的一个样本，则样本函数

$$u\overset{def}{=}\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1).$$

（2）t-分布　　设$x_1,x_2,\Lambda\ ,x_n$为来自正态总体

$N(\mu,\sigma^2)$的一个样本，则样本函数

$$t \overset{def}{=} \frac{\overline{x}-\mu}{S/\sqrt{n}} \sim t(n-1),$$

其中 t(n-1)表示自由度为 n-1 的 t 分布。

（3）κ^2 分布　　设 x_1,x_2,Λ,x_n 为来自正态总体 $N(\mu,\sigma^2)$的一个样本，则样本函数

$$w \overset{def}{=} \frac{(n-1)S^2}{\sigma^2} \sim \kappa^2(n-1),$$

其中$\kappa^2(n-1)$表示自由度为 n-1 的κ^2分布。

（4）F 分布　设 x_1,x_2,Λ,x_n 为来自正态总体 $N(\mu,\sigma^2)$的一个样本，而 y_1,y_2,Λ,y_n 为来自正态总体 $N(\mu,\sigma_2^2)$的一个样本，则样本函数

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

其中

$$S_1^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_i-\overline{x})^2, \qquad S_2^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(y_i-\overline{y})^2;$$

$F(n_1-1,n_2-1)$表示第一自由度为n_1-1，第二自由度为n_2-1的 F 分布。

七. 参数估计

1、点估计的两种方法

（1）矩法

所谓矩法就是利用样本各阶原点矩与相应的总体矩，来建立估计量应满足的方程，从而求得未知参数估计量的方法。

设总体 X 的分布中包含有未知数$\theta_1,\theta_2,\Lambda,\theta_m$，则其分布函数可以表成$F(x;\theta_1,\theta_2,\Lambda,\theta_m)$. 显示它的 k 阶原点矩 $v_k=E(X^k)(k=1,2,\Lambda,m)$ 中也包含了未知参数 $\theta_1,\theta_2,\Lambda,\theta_m$，即 $v_k=v_k(\theta_1,\theta_2,\Lambda,\theta_m)$。又设 x_1,x_2,Λ,x_n为总体 X 的 n 个样本值，其样本的 k 阶原点矩为

$$\hat{v}_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k \qquad (k=1,2,\Lambda,m).$$

这样，我们按照“当参数等于其估计量时，总体矩等于相应的样本矩”的原则建立方程，即有

$$\begin{cases} v_1(\hat{\theta}_1,\hat{\theta}_2,\Lambda,\hat{\theta}_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i, \\ v_2(\hat{\theta}_1,\hat{\theta}_2,\Lambda,\hat{\theta}_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i^2, \\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda\ \Lambda \\ v_m(\hat{\theta}_1,\hat{\theta}_2,\Lambda,\hat{\theta}_m)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i^m. \end{cases}$$

由上面的 m 个方程中，解出的 m 个未知参数$(\hat{\theta}_1,\hat{\theta}_2,\Lambda,\hat{\theta}_m)$即为参数（$\theta_1,\theta_2,\Lambda,\theta_m$）的矩估计量。

（2）最大似然法

所谓最大似然法就是当我们用样本的函数值估计总体参数时，应使得当参数取这些值时，所观测到的样本出现的概率为最大。

当总体 X 为连续型随机变量时，设其分布密度为$f(x;\theta_1,\theta_2,\Lambda,\theta_m)$，其中$\theta_1,\theta_2,\Lambda,\theta_m$为未知参数。又设$x_1,x_2,\Lambda,x_n$为总体的一个样本，称

$$L_n(\theta_1,\theta_2,\Lambda,\theta_m)=\prod_{i=1}^{n}f(x_i;\theta_1,\theta_2,\Lambda,\theta_m)$$

为样本的似然函数，简记为 L_n.

当总体 X 为离型随机变量时，设其分布律为$P\{X=x\}=p(x;\theta_1,\theta_2,\Lambda,\theta_m)$，则称

$$L(x_1,x_2,\Lambda,x_n;\theta_1,\theta_2,\Lambda,\theta_m)=\prod_{i=1}^{n}p(x_i;\theta_1,\theta_2,\Lambda,\theta_m)$$

为样本的似然函数。

若似然函数 $L(x_1,x_2,\Lambda,x_n;\theta_1,\theta_2,\Lambda,\theta_m)$ 在 $\hat{\theta}_1,\hat{\theta}_2,\Lambda,\hat{\theta}_m$处取到最大值，则称$\hat{\theta}_1,\hat{\theta}_2,\Lambda,\hat{\theta}_m$分别为

$\theta_1,\theta_2,\Lambda,\theta_m$ 的最大似然估计值，相应的统计量称为最大似然估计量。我们把使 L_n 达到最大的 $\hat{\theta}_1,\hat{\theta}_2,\Lambda,\hat{\theta}_m$ 分别作为 $\theta_1,\theta_2,\Lambda,\theta_m$ 的估计量的方法称为最大似然估计法。

由于 ln$x$ 是一个递增函数，所以 L_n 与 lnL_n 同时达到最大值。我们称

$$\left.\frac{\partial \ln L_n}{\partial \theta_i}\right|_{\theta_i=\hat{\theta}_i}=0, i=1,2,\Lambda,m$$

为似然方程。由多元微分学可知，由似然方程可以求出 $\hat{\theta}_i=\hat{\theta}_i(x_1,x_2,\Lambda,x_n)(i=1,2,\Lambda,m)$ 为 θ_i 的最大似然估计量。

容易看出，使得 L_n 达到最大的 $\hat{\theta}_i$ 也可以使这组样本值出现的可能性最大。

2、估计量的评选标准

（1）无偏性

设 $\hat{\theta}=\hat{\theta}(x_1,x_2,\Lambda,x_n)$ 为求知参数 θ 的估计量。若 E（$\hat{\theta}$）=θ，则称 $\hat{\theta}$ 为 θ 的无偏估计量。

若总体 X 的均值 E（X）和方差 D（X）存在，则样本均值 $\bar{x}$ 和样本方差 S^2 分别为 E（X）和 D（X）的无偏估计，即

E（$\bar{x}$）=E（X），E（S^2）=D（X）。

（2）有效性

设 $\hat{\theta}_1=\hat{\theta}_1(x_1,x_2,\Lambda,x_n)$ 和 $\hat{\theta}_2=\hat{\theta}_2(x_1,x_2,\Lambda,x_n)$ 是未知参数 θ 的两个无偏估计量。若 $D(\hat{\theta}_1)<D\hat{\theta}_2$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。

（3）一致性（相合性）

设 $\hat{\theta}_n$ 是 θ 的一串估计量，如果对于任意的正数 ε，都有

$$\lim_{n\to\infty}P(|\hat{\theta}_n-\theta|>\varepsilon)=0,$$

则称 $\hat{\theta}_n$ 为 θ 的一致估计量（或相合估计量）。

3、区间估计

（1）置信区间和置信度

设总体 X 含有一个待估的未知参数 θ。如果我们从样本 x_1,x_2,Λ,x_n 出发，找出两个统计量 $\theta_1=\theta_1(x_1,x_2,\Lambda,x_n)$ 与 $\theta_2=\theta_2(x_1,x_2,\Lambda,x_n)$ $(\theta_1<\theta_2)$，使得区间 $[\theta_1,\theta_2]$ 以 $1-\alpha(0<\alpha<1)$ 的概率包含这个待估参数 θ，即

$$P\{\theta_1\le\theta\le\theta_2\}=1-\alpha,$$

那么称区间 $[\theta_1,\theta_2]$ 为 θ 的置信区间，$1-\alpha$ 为该区间的置信度（或置信水平）。

（2）单正态总体的期望和方差的区间估计

设 x_1,x_2,Λ,x_n 为总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，在置信度为 $1-\alpha$ 下，我们来确定 μ 和 σ^2 的置信区间 $[\theta_1,\theta_2]$。具体步骤如下：

（i）选择样本函数；

（ii）由置信度 $1-\alpha$，查表找分位数；

（iii）导出置信区间 $[\theta_1,\theta_2]$。

下面分三种情况来讨论。

① 已知方差，估计均值

（i）选择样本函数

设方差 $\sigma^2=\sigma_0^2$，其中 σ_0^2 为已知数。我们知道 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ 是 μ 的一个点估计，并且知道包含未知参数 μ 的样本函数。

$$u=\frac{\bar{x}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1).$$

(ii) 查表找分位数

对于给定的置信度 $1-\alpha$，查正态分布分位数表，找出分位数 λ，使得

$$P(|u|\le\lambda)=1-\alpha。$$

即

$$P\left(-\lambda \le \frac{\bar{x}-\mu}{\sigma_2/\sqrt{n}} \le \lambda\right)=1-\alpha.$$

（iii）导出置信区间

由不等式

$$-\lambda \le \frac{(\bar{x}-\mu)\sqrt{n}}{\sigma_2} \le \lambda$$

推得

$$\bar{x}-\lambda\frac{\sigma_0}{\sqrt{n}} \le \mu \le \bar{x}+\lambda\frac{\sigma_0}{\sqrt{n}},$$

这就是说，随机区间

$$\left[\bar{x}-\lambda\frac{\sigma_0}{\sqrt{n}}, \bar{x}+\lambda\frac{\sigma_0}{\sqrt{n}}\right]$$

以$1-\alpha$的概率包含μ。

② 未知方差，估计均值

（i）选择样本函数

设x_1, x_2, Λ, x_n为总体$N(\mu,\sigma^2)$的一个样本，由于σ^2是未知的，不能再选取样本函数 u。这时可用样本方差

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

来代替σ^2，而选取样本函数

$$t=\frac{\bar{x}-\mu}{S/\sqrt{n}} \sim t(n-1).$$

(ii)查表找分位数

对于给定的置信度$1-\alpha$，查 t 分位数表，找出分位数λ，使得

$$P(|u| \le \lambda)=1-\alpha。$$

即

$$P\left(-\lambda \le \frac{\bar{x}-\mu}{S/\sqrt{n}} \le \lambda\right)=1-\alpha.$$

（iii）导出置信区间

由不等式

$$-\lambda \le \frac{(\bar{x}-\mu)\sqrt{n}}{\sigma_2} \le \lambda$$

推得

$$\bar{x}-\lambda\frac{S}{\sqrt{n}} \le \mu \le \bar{x}+\lambda\frac{S}{\sqrt{n}},$$

这就是说，随机区间

$$\left[\bar{x}-\lambda\frac{S}{\sqrt{n}}, \bar{x}+\lambda\frac{S}{\sqrt{n}}\right]$$

以$1-\alpha$的概率包含μ。

③ 方差的区间估计

（i）选择样本函数

设x_1, x_2, Λ, x_n为来自总体$N(\mu,\sigma^2)$的一个样本，我们知道$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$是$\sigma^2$的一个点估计，并且知道包含未知参数$\sigma^2$的样本函数

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

（ii）查表找分位数

对于给定的置信度$1-\alpha$，查κ^2分布分位数表，找出两个分位数λ_1与λ_2，使得由于κ^2分布不具有对称性，因此通常采取使得概率对称的区间，即

$$P(\lambda_1 \le \omega \le \lambda_2)=1-\alpha.$$

于是有

$$P\left(\lambda_1 \le \frac{(n-1)S^2}{\sigma^2} \le \lambda_2\right)=1-\alpha.$$

（iii）导出置信区间

$$\lambda_1 \le \frac{(n-1)S^2}{\sigma^2} \le \lambda_2$$

由不等式

$$\frac{(n-1)S^2}{\lambda_2} \le \sigma^2 \le \frac{(n-1)S^2}{\lambda_1}$$

以$1-\alpha$的概率包含σ^2，而随机区间

$$\left[\sqrt{\frac{n-1}{\lambda_2}}S, \sqrt{\frac{n-1}{\lambda_1}}S\right]$$

以$1-\alpha$的概率包含σ。