概率统计笔记

臧亦驰

2021 年 12 月 3 日

目录

1 概率

1.1 随机事件概率

1.1.1 排列组合

定义式：

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

常用排列组合公式：

$$\begin{cases} C_n^m = C_n^{n-m} \\ C_n^m + C_n^{m-1} = C_{n+1}^m \\ \sum_{k=0}^{m} C_{n_1}^k C_{n_2}^{m-k} = C_{n_1+n_2}^m \\ \sum_{k=0}^{n} C_n^k = 2^n \end{cases}$$

1.1.2 概率模型

古典概型：满足样本空间只有有限个样本点，且每个样本点出现的可能性相等的随机试验。

超几何概型：N个物体，分成s类，每一类分别有N_i个物体，从中不放回地抽取n个，恰好抽到n_i个第i类物体的概率为

$$P(A) = \frac{\prod_{i=1}^{s} C_{N_i}^{n_i}}{C_N^n}$$

几何概型：样本空间是一个可以度量的几何区域，向其上随意抛掷一点，该点落在区域A中的可能性只与A的度量$\mu(A)$成比例而与位置和形状无关的随机试验。

1.1.3 条件概率

事件B发生的前提下事件A发生的概率称作**条件概率**，其定义式为

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

显然从中可以推出

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

这称为**贝叶斯公式**或后验概率公式。

对于一组两两互斥且包含所有可能事件的事件组A_i，有

$$P(B) = \sum_i P(A_i)P(B|A_i)$$

这称为**全概率公式**。

1.2 一维随机变量

1.2.1 分布函数

设X为随机变量，x是任意实数，则X的**分布函数**为

$$F(x) = P\{X \leqslant x\}$$

1.2.2 概率密度

对于连续型随机变量X，它的**概率密度函数**是这样的$f(x)$，使得

$$F(x) = \int_{-\infty}^{x} f(u)\mathrm{d}u$$

1.2.3 常见离散随机变量

两点分布（伯努利分布）：记作$X \sim B(1,p)$

$$P\{X=0\} = p, \quad P\{X=1\} = 1-p$$

二项分布：独立重复n次伯努利试验，得到指定结果的次数满足的概率分布，记作$X \sim B(n,p)$

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

泊松分布：记作$X \sim P(\lambda)$

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

几何分布：独立重复伯努利试验，第一次得到指定结果的次数满足的概率分布，记作$X \sim Ge(p)$

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

负二项分布（帕斯卡分布）：独立重复伯努利试验，得到n次指定结果时的试验次数满足的概率分布，记作$X \sim Nb(r,p)$

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

超几何分布：N个物品不放回取件n次，得到指定种类物品（共M件）的次数满足的概率分布，记作$X \sim h(n,N,M)$

$$P\{X=k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \quad k=1,2,\cdots,\min\{M,n\}$$

1.2.4 常见连续随机变量

均匀分布：记作$X \sim U[a,b]$

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

指数分布：记作$X \sim E(\lambda)$

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

正态分布：由大量微小的独立随机因素叠加而成的现象通常近似符合正态分布，记作$X \sim N(\mu,\sigma^2)$

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

伽马分布：记作$X \sim Ga(\alpha,\lambda)$

$$f(x)=\begin{cases}\dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

其中$\Gamma(\alpha)$为**伽马函数**，定义为

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x$$

贝塔分布：记作$X \sim Be(a,b)$

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

其中$B(a,b)$为**贝塔函数**，有如下定义和性质

$$\begin{aligned}B(a,b)=B(b,a)&=\int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x\\ &=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}\end{aligned}$$

1.3 多维随机变量

1.3.1 分布函数

设$(X_1,X_2,\cdots,X_n)$是多维随机变量，$x_1,x_2,\cdots,x_n$是任意实数，则随机变量的**联合分布函数**（简称分布函数）为

$$F(x_1,x_2,\cdots,x_n)=P\{\forall k, X_k\leqslant x_k\}$$

关于X_k的**边缘分布函数**为

$$F_{X_k}(x)=P\{X_k\leqslant x\}$$

1.3.2 概率密度

对于多维连续型随机变量$(X_1,X_2,\cdots,X_n)$，它的**联合概率密度**是这样的函数$f(x_1,x_2,\cdots,x_n)$，使得

$$F(x_1,x_2,\cdots,x_n)=\int_{-\infty<y_k<x_k}f(y_1,y_2,\cdots,y_n)\mathrm{d}^nV$$

X_k的**边缘密度函数**$f_{X_k}(x)$，使得

$$F_{X_k}(x)=\int_{-\infty}^{x}f_{X_k}(u)\mathrm{d}u$$

1.3.3 常见多维随机变量

多项分布：二项分布的推广，r维随机变量，记作$M(n, p_i, p_2, \cdots, p_r)$

$$P\{\forall k, X_k = x_k\} = \frac{n!}{\prod\limits_{i=1}^{r}(x_i!)}\prod_{j=1}^{r} p_j^{x_j}$$

多维均匀分布：记作$U(D)$

$$f(x_1, x_2, \cdots, x_n) = \begin{cases} \dfrac{1}{\mu(D)}, & (x_1, x_2, \cdots, x_n) \in D \\ 0, & others \end{cases}$$

多维正态分布：记作$N(\mu, \Sigma)$，其中$\Sigma = (\sigma_{ij})$，$\mu = (\mu_1, \mu_2, \cdots, \mu_n)^T$，并且记$x = (x_1, x_2, \cdots, x_n)^T$

$$f(x) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\}$$

1.3.4 条件分布

定义条件概率为

$$P\{X_i = x_i | X_j = x_j\} = \frac{P\{X_i = x_i, X_j = x_j\}}{P\{X_j = x_j\}}$$

或

$$P\{X_i \leqslant x_i | x_j - \epsilon < X_j \leqslant x_j + \epsilon\} = \frac{P\{X_i \leqslant x_i, x_j - \epsilon < X_j \leqslant x_j + \epsilon\}}{P\{x_j - \epsilon < X_j \leqslant x_j + \epsilon\}}$$

如果$\epsilon \to 0$时上式极限存在，则将其记为$F_{X_i|X_j}(x_i|x_j)$。条件密度函数是这样的函数$f_{X_i|X_j}(x_i|x_j)$，使得

$$F_{X_i|X_j}(x_i|x_j) = \int_{-\infty}^{x_i} f_{X_i|X_j}(u|x_j)\mathrm{d}u$$

1.3.5 随机变量的独立性

若$\forall x_i$，都有

$$F(x_1, x_2, \cdots) = \prod_i F_{X_i}(x_i)$$

则称这些随机变量互相独立。亦可以等价地写为

$$f(x_1, x_2, \cdots) = \prod_i f_{X_i}(x_i)$$

1.3.6 随机变量函数的分布

对于二维情况，假设随机变量X, Y独立，随机变量函数$Z = X + Y$在离散情况下的分布律为

$$P\{Z = z\} = \sum_{k=0}^{z} p(k)q(z-k)$$

这是**一维离散卷积公式**。在连续情况下概率密度为

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-y)\mathrm{d}x = \int_{-\infty}^{\infty} f_X(z-x)f_Y(y)\mathrm{d}y$$

这是**一维连续卷积公式**。

随机变量函数$Z=\max\{X,Y\}$和$Z=\min\{X,Y\}$的分布函数分别为

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

1.4 随机变量数字特征

1.4.1 数学期望

若对于离散型随机变量X，级数$\sum_i x_iP(X=x_i)$绝对收敛，则称其为X的**数学期望**，即

$$E(X)=\sum_{i=1}^{\infty}x_iP(X=x_i)$$

若对于连续型随机变量X，积分$\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$绝对收敛，则称其为$X$的**数学期望**，即

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

数学期望对不同随机变量是线性的。

1.4.2 方差

对于一个随机变量X，定义它的**方差**如下

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

可以将其写成

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

1.4.3 协方差

两个随机变量的和/差的方差为

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

其中$Cov(X,Y)$称为**协方差**，定义为

$$Cov(X,Y)=E[(X-E(X))(Y-E(Y))]$$

可以将其写成

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

随机变量独立是协方差为零的充分不必要条件。

1.4.4 相关系数

对于两个方差大于零的随机变量，定义

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

为X和Y的**相关系数**。相关系数的绝对值反映了两个随机变量之间存在线性关系的概率。

1.5 极限定理

1.5.1 切比雪夫不等式

有随机变量X，则对于任意正数ϵ，有

$$P\{|X-E(X)| \geqslant \epsilon\} \leqslant \frac{D(X)}{\epsilon^2}$$

1.5.2 大数定律

若对于随机变量序列$\{X_n\}$，存在一个常数μ，使得对于任意正数ϵ，都有

$$\lim_{n\to\infty} P\{|X_n-\mu|<\epsilon\}=1$$

则称该随机变量序列**依概率收敛**于μ，记为$X_n \xrightarrow{P} \mu$。

设随机变量序列$\{X_n\}$互相独立，且$\{D(X_i)\}$存在上确界，则对于任意正数ϵ，有

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

这被称为**切比雪夫大数定律**，可以通过对随机变量序列的平均数应用切比雪夫不等式得到。

设μ_n是n次伯努利试验中事件A发生的次数，p是事件A发生的概率，则对于任意正数ϵ，有

$$\frac{\mu_n}{n} \xrightarrow{P} p$$

这被称为**伯努利大数定律**，它为频率的稳定值作为概率提供了理论依据。

对于服从相同分布且相互独立的随机变量序列$\{X_i\}$，其数学期望为μ，则对于任意正数ϵ，有

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

这被称为**辛钦大数定律**。

1.5.3 中心极限定理

对于服从相同分布且相互独立的随机变量序列$\{X_i\}$，其数学期望为μ，方差为σ^2，则对于任意实数x，有

$$\lim_{n\to\infty} P\{\frac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leqslant x\} = \Phi(x)$$

这称为**林德伯格-列维中心极限定理**，其中$\Phi(x)$是标准正态分布的分布函数。这表明，若$\bar{X}=\frac{1}{n}\sum\limits_{i=1}^{n} X_i$，则当$n$充分大时

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

对于相互独立的随机变量序列$\{X_i\}$，有$E(X_i)=\mu_i$，$D(X_i)=\sigma_i^2$，记$B_n=\sqrt{\sum\limits_{i=1}^{n}\sigma_i^2}$，如果存在正数$\delta$，使得$n\to\infty$时

$$\frac{1}{B_n^{2+\delta}}\sum_{i=1}^{n} E(|X_i-\mu_i|^{2+\delta}) \to 0$$

则对于任意x，有

$$\lim_{n\to\infty} P\{\frac{1}{B_n}\sum_{i=1}^{n}(X_i-\mu_i)\leqslant x\}=\Phi(x)$$

这被称为**李雅普诺夫中心极限定理**。这个定理说明无论各个随机变量服从什么分布，只要符合定理条件，那么n足够大时，他们的和近似地服从正态分布，即

$$\sum_{i=1}^{n}X_i \overset{approx}{\sim} N(\sum_{i=1}^{n}\mu_i, B_n^2)$$

2 统计

2.1 抽样分布理论

2.1.1 样本

从总体中抽取n个个体$X_1, X_2, \cdots, X_n$称为总体的一个**样本**，样本是随机变量，但一旦具体抽取后，就能得到一组确定的数值，用x_i表示，成为**样本观测值**。一般抽样时采取简单随机抽样方法，特点是样本的每一个分量X_i与总体X具有相同的概率分布，且样本之间互相独立。采用简单随机抽样方法得到的样本称为**简单随机样本**。

2.1.2 统计量

若关于样本的一个函数其中不含任何未知参数，则称这个函数是一个**统计量**。以下是一些常见统计量

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

称为**样本均值**。

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

称为**样本方差**。

$$R=\max_{1\leqslant i\leqslant n}\{X_i\}-\min_{1\leqslant i\leqslant n}\{X_i\}$$

称为**样本极差**。

$$C_v=\frac{S}{|\bar{X}|}$$

称为**样本变异系数**。

2.1.3 经验分布函数

通过样本观测值可以估计总体的分布函数，将样本观测值按从小到大的顺序排列，则函数

$$F_n(x)=\begin{cases}0, & x<x_1\\ \dfrac{k}{n}, & x_k\leqslant x<x_{k+1}\\ 1, & x\geqslant x_n\end{cases}$$

称为样本的**经验分布函数**。有

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

2.1.4 分位数

对于总体和给定正数α，若存在实数x_α，满足

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

则称x_α为X的**上侧α分位数**。若存在实数$x_{\alpha/2}$，满足

$$P\{|X| \geqslant x_{\alpha/2}\}=\alpha$$

则称x_α为X的**双侧α分位数**。

2.1.5 常见抽样分布

统计量的分布称为抽样分布。以下抽样分布是以标准正态分布为基石而构造的。

χ^2分布：来自标准正态分布的n个样本，它们平方和Y所服从的分布为自由度为n的χ^2分布，记为$Y\sim\chi^2(n)$，密度函数为

$$f(y)=\begin{cases}\dfrac{1}{2^{n/2}\Gamma(\frac{n}{2})}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}, & y\geqslant 0\\ 0, & y<0\end{cases}$$

可以发现这是一种特殊的伽马分布，$\chi^2(n)=Ga(\frac{n}{2},\frac{1}{2})$。$\chi^2$分布的期望值和方差分别为$E(Y)=n, D(Y)=2n$。

t分布：设随机变量$X\sim N(0,1), Y\sim\chi^2(n)$，则随机变量

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

服从的分布为自由度为n的t分布，记为$t\sim t(n)$，密度函数为

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

n大于一时期望值为$E(t)=0$，$n=1$时期望值不存在，n大于二时方差为$D(t)=\frac{n}{n-2}$，$n=1,2$时方差不存在。

F分布：设随机变量$Y_1\sim\chi^2(n_1), Y_2\sim\chi^2(n_2)$，则随机变量

$$F=\frac{Y_1/n_1}{Y_2/n_2}$$

服从的分布为第一自由度为n_1第二自由度为n_2的F分布，记作$F\sim F(n_1,n_2)$，密度函数为

$$f(x)=\begin{cases}\dfrac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})}(\dfrac{n_1}{n_2})^{\frac{n_1}{2}}x^{\frac{n_1}{2}-1}(1+\dfrac{n_1}{n_2}x)^{-\frac{n_1+n_2}{2}}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

2.1.6 样本均值和样本方差的分布

对于来自均值为μ，方差为σ^2的某总体的一组样本，当n充分大时，近似有

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

对于来自正态分布的样本，则有

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

且样本均值与样本方差独立。

对于分别来自正态分布$N(\mu_1, \sigma^2)$和$N(\mu_2, \sigma^2)$且相互独立的样本X_i, Y_j，有

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

其中

$$S = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

又有

$$\frac{S_1^2}{S_2^2} \sim F(n_1-1, n_2-1)$$

2.2 参数估计

2.2.1 参数估计

假设总体的分布函数中有一个未知参数θ，通过样本构造统计量$\hat{\theta}(X_i)$来估计参数θ，称为参数的**估计量**，将样本观测值带入得到一个具体的值称为估计值。

构造一个统计量作为未知参数的估计量，就称其为参数的**点估计**，构造两个统计量而用他们的区间作为未知参数可能取值范围的一种估计，称为这个参数的**区间估计**，也称为**置信区间**。

2.2.2 参数的点估计

样本数字特征：样本均值和样本方差分别可以用来估计总体均值和总体方差，即

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

矩估计法：对于随机变量X，若$E(|X|^k)$存在，则将$a_k = E(X^k)$称为总体的k**阶原点矩**，若$E(|X - E(X)|^k)$存在，则将$b_k = E((X-E(X))^k)$称为总体的k**阶中心矩**。对于样本而言，同样有**样本原点矩**和**样本中心矩**。

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

对于原点矩和中心矩，有$\hat{a}_k = A_k, \hat{b}_k = B_k$。若用未知参数表示总体的矩，再带入矩的估计值，则可以求出对应参数的估计值，这称为**矩估计**。

最大似然法：设总体的密度函数中有一未知参数θ，则样本的联合密度函数为

$$f(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

对于一组确定的观察值，它是θ的函数，记为

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

称为θ的**似然函数**。似然函数取到最大值时的$\hat{\theta}$能使得样本落在当前观察值的概率最高，因此将其作为θ的估计值。

2.2.3 估计量的优良准则

无偏性：若未知参数估计量的期望值等于该未知参数，则称这样的估计为**无偏估计**。样本数字特征得到的估计量都是无偏估计。

有效性：某一未知参数的两个估计量中方差更小的更有效，同一个未知参数的所有无偏估计中方差最小的估计称为**一致最小方差无偏估计**

均方误差准则：将估计量与参数值的距离平方用来度量估计的好坏，称为**均方误差**

$$MSE(\hat{\theta}) = E((\hat{\theta} - \theta)^2)$$

可以将其化为

$$MSE(\hat{\theta}) = D(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2$$

可以看出，如果是无偏估计，那么均方误差就等于方差，而对于有偏估计的考察，还需要看其偏差大小。

2.2.4 区间估计

若通过样本确定两个统计量$\hat{\theta}_1, \hat{\theta}_2$，使得$P(\hat{\theta}_1 \leqslant \theta \leqslant \hat{\theta}_2) = \alpha$，则称区间$[\hat{\theta}_1, \hat{\theta}_2]$为参数的置信度为$\alpha$的**置信区间**。$\hat{\theta}_1$为（双侧）置信下界，$\hat{\theta}_2$为（双侧）置信上界。可以通过构建随机变量并通过其分位数计算置信区间上下界。

2.3 假设检验

2.3.1 假设检验

一般地，统计假设要针对研究的问题提出一对假设，其中我们较为关心的一个称为**原假设**，记为H_0，与之对立的一个假设称为**备择假设**，记为H_1。作出假设后要根据数据提供的信息对假设做出检验，并作出接受或拒绝原假设的判决。这样的判断通常是依据一个统计量完成，它称为**检验统计量**，使得原假设被拒绝的样本观察值所在区域称为该假设的**拒绝域**，记作W，而$\bar{W}$称作**接受域**。接受域与拒绝域的临界点称为**临界值**。

2.3.2 错误

假设检验的基本原理是小概率事件原理，即在一次实验中小概率事件是不可能发生的。可以想见由此作出的判决有可能是错误的，错误分为两类：第I类错误和第II类错误。

第I类错误：原假设正确，而样本观测值落在了拒绝域，作出了拒绝原假设的判断。这称为**弃真错误**，有发生这种错误的概率α

$$P\{W|H_0\ is\ true\} = \alpha$$

概率α称为检验的**显著性水平**。

第II类错误：原假设错误，而样本观测值落在了接受域，作出了接受原假设的判断。这称为**取伪错误**，有发生这种错误的概率β

$$P\{\bar{W}|H_0\ is\ false\} = \beta$$

可以证明，α减小时，β就会增大，反之亦然。通常采取“保护零假设原则”，即先限制第I类错误的概率不超过显著性水平，并在此基础上使第II类错误的概率尽可能小。

3 附录

3.1 区间估计

单个正态分布：方差已知时均值的$1-\alpha$置信区间为

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

其中$z_{\alpha/2}$是标准正态分布的双侧分位数。

方差未知时均值的$1-\alpha$置信区间为

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

其中$t_{\alpha/2}(n-1)$时$t(n-1)$分布的双侧分位数。

均值已知时方差的$1-\alpha$置信区间为

$$[\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}, \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}]$$

其中$\chi^2_{\alpha/2}(n), \chi^2_{1-\alpha/2}(n)$是$\chi^2(n)$分布的双侧分位数。

均值未知时方差的$1-\alpha$置信区间为

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

两个正态分布：方差已知时两个均值之差的$1-\alpha$置信区间为

$$[\bar{X} - \bar{Y} - u_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \bar{X} - \bar{Y} + u_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}]$$

其中$u_{\alpha/2}$是标准正态分布的双侧分位数。

方差相等但未知时两个均值之差的$1-\alpha$置信区间为

$$[\bar{X}-\bar{Y}-t_{\alpha/2}(n_1+n_2-2)S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\bar{X}-\bar{Y}+t_{\alpha/2}(n_1+n_2-2)S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}]$$

其中

$$S^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

S_1^2,S_2^2为样本方差。

两个方差之比的$1-\alpha$置信区间为

$$[\frac{S_1^2}{S_2^2}\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}]$$

3.2 假设检验

单个正态分布：方差已知时均值的检验

原假设	检验统计量	临界值	拒绝域
$H_0:\mu=\mu_0$	$U=\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{n}}$	$u_{\alpha/2}$	$\{\lvert U\rvert\geqslant u_{\alpha/2}\}$
$H_0:\mu\geqslant\mu_0$	$U=\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{n}}$	u_α	$\{U\leqslant -u_\alpha\}$
$H_0:\mu\leqslant\mu_0$	$U=\frac{\bar{X}-\mu_0}{\sigma_0/\sqrt{n}}$	u_α	$\{U\geqslant u_\alpha\}$

方差未知时均值的检验

原假设	检验统计量	临界值	拒绝域
$H_0:\mu=\mu_0$	$T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$t_{\alpha/2}(n-1)$	$\{\lvert T\rvert\geqslant t_{\alpha/2}(n-1)\}$
$H_0:\mu\geqslant\mu_0$	$T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$t_\alpha(n-1)$	$\{T\leqslant -t_\alpha(n-1)\}$
$H_0:\mu\leqslant\mu_0$	$T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$t_\alpha(n-1)$	$\{T\geqslant t_\alpha(n-1)\}$