随机变量的数字特征

- 与随机变量有关的某些数值，虽然不能完整地描述随机变量，但能描述随机变量在某些方面的重要特征。这些数字特征在理论和实践上都具有重要的意义。下面将介绍随机变量的常用数字特征：数学期望、方差、相关系数和矩

离散型随机变量的数学期望

定义:设离散型随机变量X的分布率为

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

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛，则称 $\sum_{k=1}^{\infty} x_k p_k$

变量X的数学期望（或均值），记为$E(X)$。即

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

92

连续型随机变量的数学期望

定义:设连续型随机变量X的概率密度为$f(x)$，若积分

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

绝对收敛,则称积分$\int_{-\infty}^{\infty} xf(x)dx$ 的值为随机变量X的

数学期望（或均值），记为$E(X)$。即

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

例：随机变量X服从正态分布$N(\mu,\sigma^2)$，求数学期望。

$$解:\int_{-\infty}^{\infty} xf(x)dx=\int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2}dx$$

$$(令 z=\frac{x-\mu}{\sigma}) \qquad =\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma z+\mu)e^{-z^2/2}dz$$

$$=\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-z^2/2}dz+\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}ze^{-z^2/2}dz$$

$$=\frac{\mu}{\sqrt{2\pi}}\times\sqrt{2\pi}=\mu$$

方差

设X是一个随机变量，则

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

称为X的方差。$\sqrt{D(X)}=\sigma(X)$称为均方差或标准差。

若X是离散性随机变量，则

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2\cdot p_k$$

若X是连续性随机变量，概率密度为$f(x)$，则

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

随机变量X的方差与数学期望有如下关系：

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

证明

$$\begin{aligned} D(X) &= E\{[X-E(X)]^2\} \\ &= E\{X^2-2XE(X)+[E(X)]^2\} \\ &= E(X^2)-2E(X)E(X)+[E(X)]^2 \\ &= E(X^2)-[E(X)]^2 \end{aligned}$$

方差的性质

若 $X_i \sim N(\mu_i, \sigma_i^2)$ $i = 1, 2, \cdots, n$ 且它们互相独立

那么,它们的线性组合$C_1X_1 + C_2X_2 + \cdots + C_nX_n$, (其中

$C_1, C_2, \cdots, C_n$ 不全为零), 仍服从正态分布,且

$$C_1X_1 + C_2X_2 + \cdots + C_nX_n \sim N\left(\sum_{i=1}^{n} C_i\mu_i,\ \sum_{i=1}^{n} C_i^2\sigma_i^2\right)$$

- 对于二维随机变量，除了讨论变量X与Y的数学期望及方差外，还需要讨论描述X与Y这间相互关系的数字特征.

- 当随机变量X与Y相互独立时，有:

$$E\{[X-E(X)][Y-E(Y)]\}=0$$

- 也就是说，当 $E\{[X-E(X)][Y-E(Y)]\}\neq 0$ 时，X与Y不相互独立，即：可能存在某种关系.

协方差与相关系数

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

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

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

称为随机变量X与Y的相关系数。

显然 $$Cov(X,Y)=Cov(Y,X),Cov(X,X)=D(X)$$

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

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

独立 $\longrightarrow$ 不相关

不相关 $\not\longrightarrow$ 独立

矩的概念

设X为随机变量，若

$$a_k = E(X^k), k = 1, 2, \cdots$$

存在,称它为X的k阶原点矩。数学期望是一阶原点矩。

若$\mu_k = E\{[X - E(X)]^k\}$，$k = 2, 3, \cdots$

存在,称它为X的k阶中心矩。方差是二阶中心矩。

矩的概念

设(X,Y)是二维随机变量，若

$$a_{kl}=E(X^kY^l),k,l=1,2,\cdots$$

存在，称它为X和Y的$k+l$阶混合原点矩。

若$\mu_{kl}=E\{[X-E(X)]^k[Y-E(Y)]^l\}$，$k,l=1,2,\cdots$

存在，称它为X和Y的$k+l$阶混合中心矩。

二维随机变量协方差矩阵的概念

二维随机变量(X_1,X_2)有四个二阶中心矩（设它们都存在），分别记为

$$c_{11}=E\{[X_1-E(X_1)]^2\}$$

$$c_{12}=E\{[X_1-E(X_1)][Y_1-E(Y_1)]\}$$

$$c_{21}=E\{[Y_1-E(Y_1)][X_1-E(X_1)]\}$$

$$c_{22}=E\{[X_2-E(X_2)]^2\}$$

将它们排成矩阵的形式：

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

这个矩阵称为随机变量(X_1,X_2)的协方差矩阵。

n维随机变量协方差矩阵的概念

N维随机变量$(X_1, X_2, \cdots, X_n)$的二阶混合中心矩

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

$$i, j = 1, 2, \cdots, n$$

都存在,则称矩阵

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

为**随机变量的协方差矩阵.**

显然,上述矩阵是一个对称矩阵.

数理统计

105

- 引言：**数理统计学**是一门关于数据收集、整理、分析和推断的科学。在概率论中已经知道，由于大量的随机试验中各种结果的出现必然呈现它的规律性，因而从理论上讲只要对随机现象进行足够多次观察，各种结果的规律性一定能清楚地呈现，但是实际上所允许的观察永远是有限的，甚至是少量的。

  例如：若规定灯泡寿命低于1000小时者为次品，如何确定次品率？由于灯泡寿命试验是破坏性试验，不可能把整批灯泡逐一检测，只能抽取一部分灯泡作为样本进行检验，以样本的信息来推断总体的信息，这是数理统计学研究的问题之一。

106

§1 总体和样本

- 总体：研究对象的全体。如一批灯泡。
- 个体：组成总体的每个元素。如某个灯泡。
- 抽样：从总体X中抽取有限个个体对总体进行观察的取值过程。
- 随机样本：随机抽取的n个个体的集合$(X_1, X_2, \cdots, X_n)$，n为样本容量
- 简单随机样本：满足以下两个条件的随机样本$(X_1, X_2, \cdots, X_n)$称为简单随机样本。

  1. 每个X_i与X同分布
  2. $X_1, X_2, \cdots, X_n$是相互独立的随机变量

[说明]：后面提到的样本均指简单随机样本，由概率论知，若总体X具有概率密度f(x)，

则样本（$X_1, X_2, \cdots, X_n$）具有联合密度函数：

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

107

- 统计量：样本的不含任何未知参数的函数。
- 常用统计量：设（$X_1, X_2, \cdots, X_n$）为取自总体X的样本

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

2. 样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，$S$为样本标准差

3. 样本矩 k阶矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k \quad (k=1,2,\cdots)$

   k阶中心矩：$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k \quad (k=1,2,\cdots)$

108

§2 常用的分布

- χ^2分布

定义：设随机变量$X_1, X_2, \cdots X_n$相互独立，$X_i \sim N(0,1)$ $(i=1,2,\cdots,n)$

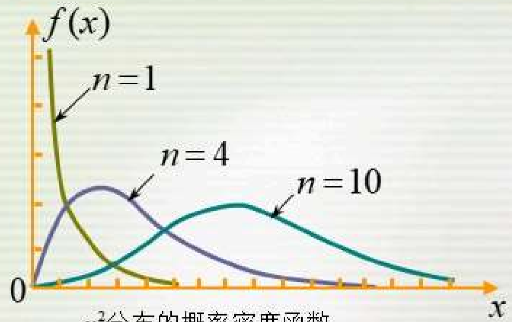
则称 $\chi_n^2 = \sum_{i=1}^{n} X_i^2$ (1)

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

自由度指(1)式右端包含的独立变量的个数

定理6.3: $\chi^2(n)$分布的概率密度为：$f_n(y) = \begin{cases} \dfrac{1}{2\Gamma(n/2)}\left(\dfrac{y}{2}\right)^{\frac{n}{2}-1} e^{-\frac{y}{2}} & y>0 \\ 0 & y \le 0 \end{cases}$

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



χ^2分布的概率密度函数

109

χ^2分布的一些重要性质：

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

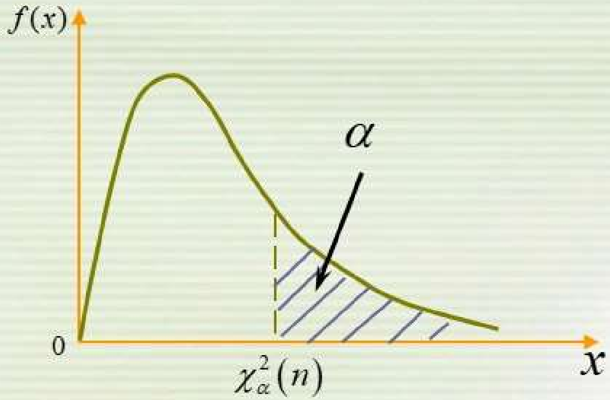
2. 设$Y_1 \sim \chi^2(n_1), Y_2 \sim \chi^2(n_2)$,且$Y_1,Y_2$相互独立，则有$Y_1+Y_2 \sim \chi^2(n_1+n_2)$

性质2称为χ^2分布的可加性，可推广到有限个的情形：

设$Y_i \sim \chi^2(n_i)$,且$Y_1,Y_2,\cdots,Y_m$相互独立，则$\sum\limits_{i=1}^{m} Y_i \sim \chi^2\left(\sum\limits_{i=1}^{m} n_i\right)$

对给定的概率$\alpha, 0<\alpha<1$,称满足条件$\int_{\chi_\alpha^2(n)}^{\infty} f_n(y)dy=\alpha$的点$\chi_\alpha^2(n)$为

$\chi^2(n)$分布的上α分位数,上α分位数$\chi_\alpha^2(n)$的值可查χ^2分布表



χ^2分布的分位数

110

$t-$分布

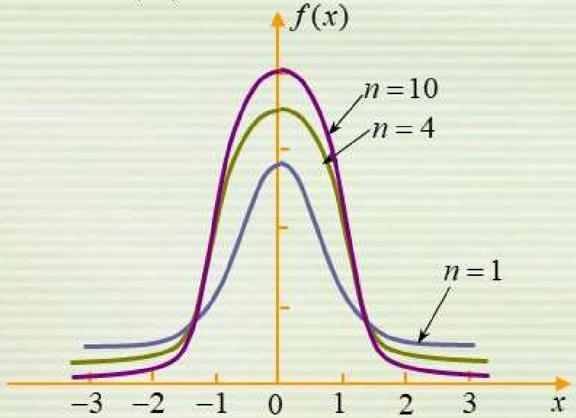
定义：设$X\sim N(0,1)$，$Y \sim \chi^2(n)$，并且X,Y相互独立，

则称随机变量$T=\dfrac{X}{\sqrt{Y/n}}$服从自由度为n的t分布，记为$T \sim t(n)$

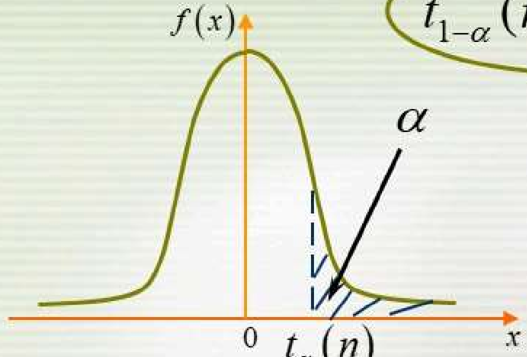
$\boxed{\text{定理6.4：}t(n)\text{分布的概率密度为：}f(t,n)=\dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\dfrac{t^2}{n}\right)^{-\frac{n+1}{2}},-\infty<t<+\infty}$

对给定的α，$0<\alpha<1$，称满足条件$\int_{t_\alpha(n)}^{\infty} f(t,n)dt=\alpha$的点$t_\alpha(n)$

为$t(n)$分布的上α分位数。t分布的上α分位数可查t分布表



t分布的密度函数



t分布的分位数

$t_{1-\alpha}(n)=-t_\alpha(n)$

111

F分布

定义：设$X\sim\chi^2(n_1),Y\sim\chi^2(n_2)$，且$X,Y$独立，

则称随机变量$F=\dfrac{X/n_1}{Y/n_2}$服从自由度(n_1,n_2)的F分布，记为$F\sim F(n_1,n_2)$

其中n_1称为第一自由度，n_2称为第二自由度

性质：$F\sim F(n_1,n_2)$,则$F^{-1}\sim F(n_2,n_1)$

定理6.5: $F(n_1,n_2)$分布的概率密度为：

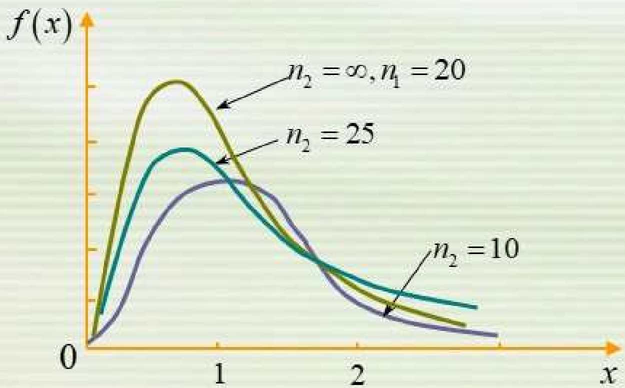
$$f(x;n_1,n_2)=\begin{cases}\dfrac{1}{B\left(\frac{n_1}{2},\frac{n_2}{2}\right)}n_1^{\frac{n_1}{2}}n_2^{\frac{n_2}{2}}x^{\frac{n_1}{2}-1}(n_2+n_1x)^{-\frac{n_1+n_2}{2}} & x>0\\ 0 & x\le 0\end{cases}$$

其中$B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}dx=\dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

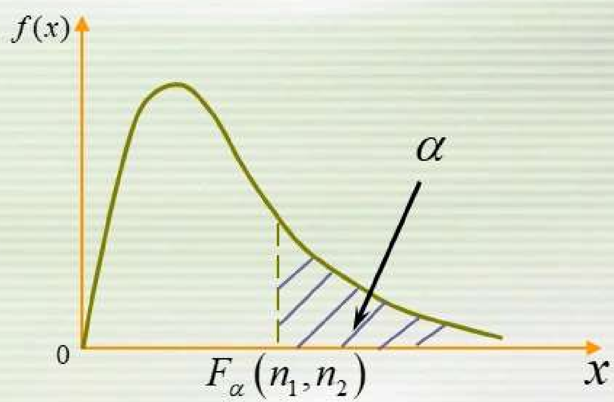
112

对于给定的α, $0<\alpha<1$,称满足条件$\int_{F_\alpha(n_1,n_2)}^{\infty} f(x;n_1,n_2)\,dx=\alpha$的点$F_\alpha(n_1,n_2)$为$F(n_1,n_2)$分布的上$\alpha$分位数。$F_\alpha(n_1,n_2)$的值可查$F$分布表

$$\boxed{F_{1-\alpha}(n_1,n_2)=[F_\alpha(n_2,n_1)]^{-1}}$$



F分布的密度函数



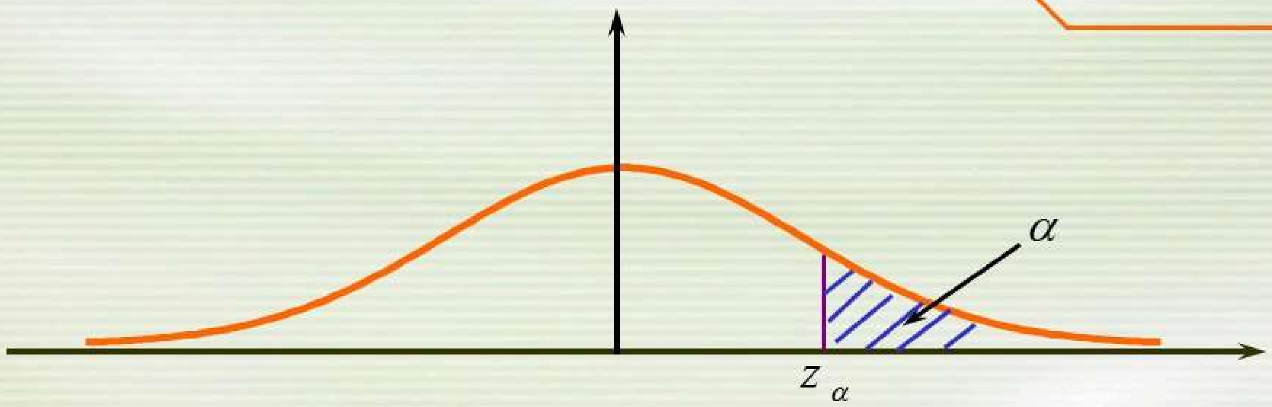
F分布的分位数

113

此外，设$X \sim N(0,1)$，若Z_α满足条件$P\{X > Z_\alpha\} = \alpha, 0 < \alpha < 1$

则称点Z_α为标准正态分布的上α分位数。

$$\boxed{Z_{1-\alpha} = -Z_\alpha}$$



正态总体样本均值和方差的分布

定理6.6: 设$(X_1,X_2,\cdots,X_n)$是总体$N(\mu,\sigma^2)$的样本，$\bar{X},S^2$分别是样 本均值和样本方差，则有：

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

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

3. $\bar{X}$和S^2相互独立

定理6.7: 设$(X_1,\cdots,X_n)$是总体$N(\mu,\sigma^2)$的样本，$\bar{X}$和 S^2分别是样本

均值和样本方差，则有：$\dfrac{\sqrt{n}(\bar{X}-\mu)}{S}\sim t(n-1)$

证明：由定理6.6知，$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，$\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$

且两者独立，由t分布定义得：

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\Big/\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}=\frac{\sqrt{n}(\bar{X}-\mu)}{S}\sim t(n-1)$$

115

定理6.8: 设样本$(X_1,\cdots,X_{n_1})$和$(Y_1,\cdots,Y_{n_2})$分别来自总体$N(\mu_1,\sigma_1^2)$和$N(\mu_2,\sigma_2^2)$并且它们相互独立，其样本方差分别为 S_1^2,S_2^2，

则：1° $F=\dfrac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} \sim F(n_1-1,n_2-1)$

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

3° 当$\sigma_1^2=\sigma_2^2=\sigma^2$时，$$\frac{(\bar X-\bar Y)-(\mu_1-\mu_2)}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$$

其中$S_W^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}, S_W=\sqrt{S_W^2}$

116

问题的提出：

参数估计是统计推断的基本问题之一，实际工作中碰到的总体X，它的分布类型往往是知道的，只是不知道其中的某些参数，

例如：产品的质量指标X服从正态分布，其概率密度为：

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

但参数μ,σ^2的值未知，要求估计μ,σ^2，有时还希望以一定的可靠性来估计μ值是在某个范围内或者不低于某个数。

参数估计问题就是要求通过样本估计总体分布所包含的未知参数的值。

参数估计的两种方法：点估计法和区间估计法