贝叶斯估计及其在抽样调查中的应用

- （Bayes，Thomas）(1702—1761)
- 贝叶斯是英国数学家.1723年生于伦敦；1761年4月17日卒于坦布里奇韦尔斯.
- 贝叶斯是一位自学成才的数学家.曾助理宗教事务，后来长久担任坦布里奇韦尔斯地方教堂的牧师.1742年，贝叶斯被选为英国皇家学会会员.
- 如今在概率、数理统计学中以贝叶斯姓氏命名的有贝叶斯公式、贝叶斯风险、

贝叶斯措施(Bayesian approach)

贝叶斯措施是基于贝叶斯定理而发展起来用于系统地论述处理统计问题的措施(Samuel Kotz和吴喜之,2023)。

贝叶斯推断的基本措施是将有关未知参数的先验信息与样本信息综合，再根据贝叶斯定理，得出后验信息，然后根据后验信息去推断未知参数(茆诗松和王静龙等,1998年)。

“贝叶斯提出了一种归纳推理的理论(贝叶斯定理)，后来被某些统计学者发展为一种系统的统计推断措施，称为贝叶斯措施.”——摘自《中国大百科全书》（数学卷）

第一章先验分布与后验分布

统计学有两个主要学派:频率学派与贝叶斯学派.它们之间有异同,贝叶斯统计是在与经典统计的争论中发展起来,主要的争论有:

1.未知参数可否作为随机变量?

2.事件的概率是否一定的频率解释?

3.概率是否可用经验来拟定?

……….

§1.1 先简介三种信息的概念

经典统计学派要求统计推断使用两种信息:

总体信息　　样本信息

而贝叶斯学派以为是三种信息:

总体信息　　样本信息　　先验信息

总体信息

即总体分布或总体所属分布族给我们的信息。譬如，“总体是正态分布”就给我们带来诸多信息：密度函数是一条钟形曲线；一切一阶距都存在；有关正态变量（服从正态分布随机变量）的某些事件的概率能够计算；由正态分布能够导出分布，分布和分布等主要分布，还有许多成熟的点估计、区间估计和假设检验措施可供我们选用。总体信息是很主要的信息，为了取得此信息，往往耗资巨大。

样本信息

- **从总体中抽取的样本给我们提供的信息。**
- 这是最“新鲜”的信息，而且愈多愈好。人们希望对样本的加工和处理对总体的某些特征作出较为精确的统计推断。没有样本就没有统计学可言。这是大家都了解的事实。

基于上述两种信息进行的统计推断称为经典统计学，它的基本观点是把数据（样本）看成是具有一定概率分布的总体，所研究的对象是这个总体而不局限于数据本身。这方面最早的工作是高斯(Gauss,C.F.1777~1855）和勒让德（Legendre,A.M.
3）的误差分析，正态分布和最小二乘法。从十九世纪末到二十世纪上半叶，经皮尔逊（Pearson,K.185
6）、费歇（Fisher,R.A.1890~1962）奈曼（Neyman.J.）等人的杰出工作创建了经典统计学。伴随经典统计学的连续发展与广泛的应用，它本身的缺陷也逐渐暴露出来了。

先验信息

即在抽样之前有关统计问题的某些信息，一般说来，先验信息主要起源于经验和历史资料。

例1：有一英国妇女，对奶茶能辨别出先倒进茶还是先倒进奶，做十次试验她都正确说出。

若H_0:每次成功概率$P=0.5$,那么十次猜中的概率为

$P_{10}(10)=0.5^{10}=0.0009766$，这是几乎不可能发生的

小概率事件,可见应拒绝H_0,$P>>0.5$.是经验在起作用.

某学生**第一次**看到他的数学老师,即有反应:老师30岁到40之间,极可能35岁左右(左右可了解为正负3岁,极可能可了解为90%的可能).

P(32≤X≤38)=0.90

例2:"免检产品"的确定

工厂每天都抽取几件产品,以估计不合格率θ,根据历史资料,

对过去的不合格率构造一个分布(先验分布)

$$P(\theta=\frac{i}{n})=\pi_i, i=1,2,\ldots,n$$

若这个分布的概率绝大部分在$\theta=0$附近,那么该产品为信得过产品,

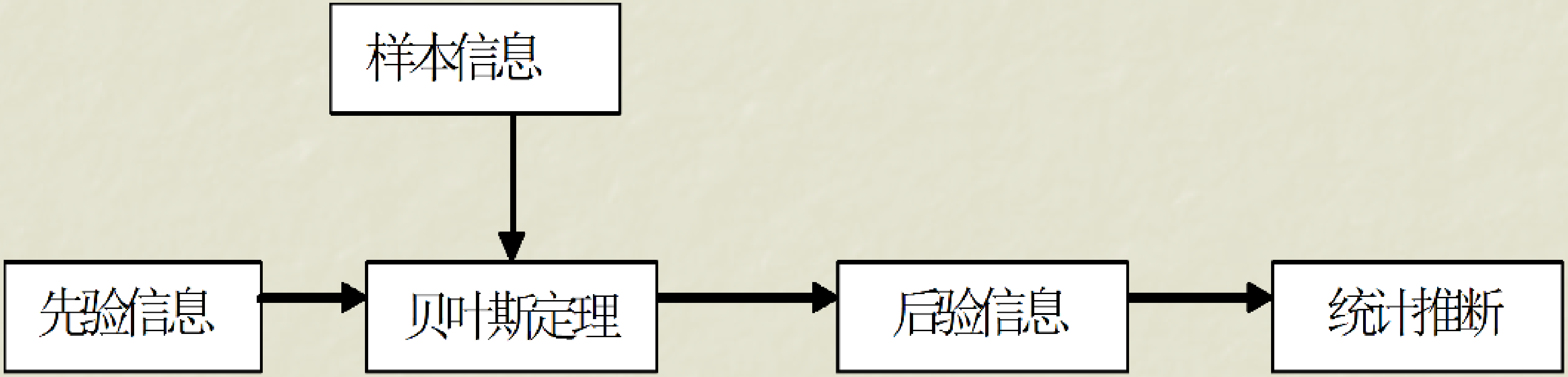
可见假定以后每天都抽取几件产品与历史资料的不合格率分布一致,

使用单位就可以确认为"免检产品".

三种信息

基于上述三种信息（总体信息、样本信息和先验信息）进行的统计推断被称为贝叶斯统计学。它与经典统计学的主要差别在于是否利用先验信息。**贝叶斯统计学派把任意一种未知参数都看成随机变量，应用一种概率分布去描述它的未知情况，该分布称为先验分布。**

贝叶斯的信息处理途径



- **后验分布是三种信息的综合,先验分布反应人们在抽样前对参数的认识,后验分布反应人们在抽样后对参数的认识**
- **Bayes统计推断原则:**对参数 θ 所作任何推断(参数估计,假设检验等)都必须建立在后验分布基础上.

共轭分布法

大家知道，在区间 $(0,1)$ 上的均匀分布是贝塔分布 $beta(1,1)$，我们从上例中看到一

$b(n,\theta)$ θ 的先验分布若取 $beta(1,1)$，则其后验

$beta(x+1,n-x+1)$ x n 为 次独立试验中成功出现的次数。先验分布与后验分布同属于一个贝塔分布族，只是参数不同而已。这一现象不是偶然的，假如把

θ 的先验分布换成一般的贝塔分布 $beta(\alpha,\beta)$，其中 $\alpha>0,\beta>0$，经过类似的计算可以看出，

θ 的后验分布仍是贝塔分布 $beta(\alpha+x,\beta+n-x)$，此种先验分布被称为

θ

后验分布和先验分布是同一种类型

定义：设是总体分布中的参数（或参数向量），是的先验密度函数，假如由抽样信息算得的后验密度函数与有相同的密度函数形式，则称是的（自然）共轭先验分布。

应该着重指出，共轭先验分布是对某一分布中的参数而言的。如正态均值、正态方差、泊松均值等。离开指定参数及其所在的分布去谈论共轭先验分布是没有意义的。

正态均值（方差已知）的共轭先验分布是正态分布

设 $x_1, \square, x_n$ 是来自正态分布 $N(\theta, \sigma^2)$ 的一个样本的观察值，其中

σ^2 已知，此样本的似然函数为

$$p(x \mid \theta) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2\right\}$$

$N(\mu, \tau^2)$ 作为正态均值 θ

验分布，即

$$\pi(\theta) = \frac{1}{\sqrt{2\pi}\tau}\exp\left\{-\frac{(\theta-\mu)^2}{2\tau^2}\right\}, -\infty < \theta < +\infty$$

其中

μ 与 τ^2 为已知，

由此可以写出样本 x 与参数 θ 的联合密度函数

其中

$$h(x,\theta)=k_1\exp\left\{-\frac{1}{2}\left[\frac{n\theta^2-2n\theta\bar{x}+\sum_{i=1}^{n}x_i^2}{\sigma^2}+\frac{\theta^2-2\mu\theta+\mu^2}{\tau^2}\right]\right\}$$

$$k_1=(2\pi)^{-(n+1)/2}\tau^{-1}\sigma^{-n},\bar{x}=\sum_{i=1}^{n}\frac{x_i}{n}$$

其中

$$则有\ h(x,\theta)=k_1\exp\left\{-\frac{1}{2}[A\theta^2-2\theta B+C\right\}=k_2\exp\left\{-\frac{(\theta-B/A)^2}{2/A}\right\}$$

由此容易算得样本

$$k_2=k_1\exp\{-\frac{1}{2}(C-B^2/A)\}$$

x 的边缘分布

上面两式相除，即得

$$m(x)=\int_{-\infty}^{+\infty}h(x,\theta)d\theta=k_2\left(\frac{2\pi}{a}\right)^{\frac{1}{2}}$$

θ 的后验分布

$$\pi(\theta/x)=\left(\frac{2\pi}{A}\right)^{-\frac{1}{2}}\exp\left\{-\frac{(\theta-B/A)^2}{2/A}\right\}$$

这是正态分布，其均值 u_1 与方差 τ_1^2 分别为

$$\mu_1=\frac{B}{A}=\frac{\overline{x}\sigma_0^{-2}+\mu\tau^{-2}}{\sigma_0^{-2}+\tau^{-2}}\qquad\frac{1}{\tau_1^2}=\frac{1}{\sigma_0^2}+\frac{1}{\tau^2}$$

这说明了正态均值（方差已知）的共轭先验分布是正态分布。

常用共轭先验分布

总体分布	参数	共轭先验分布
二项分布	成功概率	贝塔分布 $beta(\alpha,\beta)$
		$Ga(\alpha,\lambda)$
		$Ga(\alpha,\lambda)$
		$N(\mu,\tau^2)$
		$IGa(\alpha,\lambda)$

共轭先验分布的优点

它有两个优点

1. 计算方便
2. 后验分布中的一些参数可以得到很好的解释

在“正态均值 θ 的共轭先验分布为正态分布”的例题中，

其后验均值可改写为

$$\mu_1=\frac{\sigma_0^{-2}}{\sigma_0^{-2}+\tau^{2}}\bar{x}+\frac{\tau^{-2}}{\sigma_0^{-2}+\tau^{-2}}\mu=\gamma\bar{x}+(1-\gamma)\mu$$

其中

$$\gamma=\sigma_0^{-2}/(\sigma_0^{-2}+\tau^{-2})$$

是用方差倒数组成的权，于是后验均值 μ_1 是样本均值 $\bar{x}$ 和先验均值 μ 的加权平均。这表明后验均值是在先验均值与样本均值间采取折衷方案。