



RICUP

MATHEMATICS



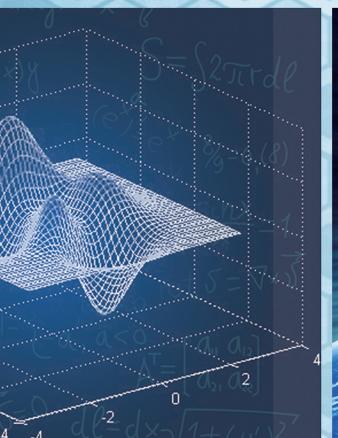
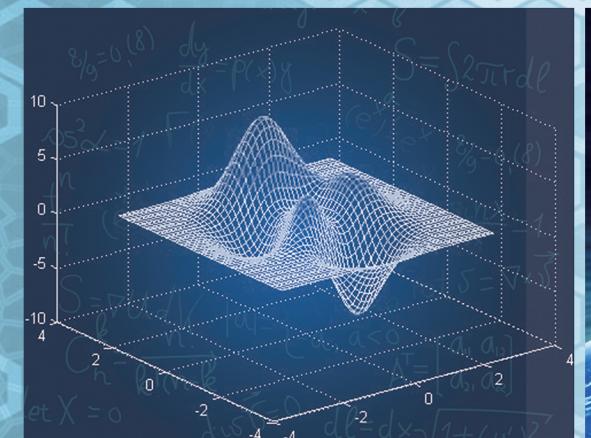
主办 数学科学学院
顾问 谢云荪
主编 荆燕飞
副主编 杨春
编委 陈绍刚 高建 龚丽莎 何军华 李晓东 蒲和平 覃思义
王博 王志勇

2019.12



电子科技大学
University of Electronic Science and Technology of China

成电大学生 数学园地



成电大学生

数学园地

2019.12 第10期

RICUP
MATHEMATICS

数学科学学院
SCHOOL OF MATHEMATICAL SCIENCES

成电大学生数学园地

2019 年 12 月

第十期

目 录

基于抽样分布定理的对于正态总体下的考试分数波动的研究	何世图	(1)
探究利用独立同分布极限定理近似计算二项分布随机变量序列的误差以及校正方法	刘明康	(4)
有关假设检验的两类错误和检验水平的研究	陈思佳 孟翔晨	(8)
有关样本均值的分布的研究	舒 婷 胡智成 苏 力	(13)
导数运算 $(fg)' = f'g'$ 及 $(f/g)' = f'/g'$ 的条件	张杰瑞	(17)
借助反射矩阵快速求解反射光线方程的方法	蔡思源	(20)
初探 CRAMER 法则	郭 好 邓远哲	(23)
第一类曲面积分的两种简单计算方法	肖健雄	(26)
积分因子法在求解微分方程中的应用	肖健雄	(31)
留数定理在 LAPLACE 反变换上的计算	刘思雨	(35)
绕任意直线旋转的旋转体体积公式的探究	张晋豪 何蔚阳	(38)
最速下降算法的简单研究及其应用	肖健雄 陈思佳	(41)
坐标系旋转及矩阵变化下的二维随机变量	文彦博	(45)
不同相关系数定义辨析及应用举例	甘 玲	(48)
关于最短置信区间的讨论	侯付豪	(51)
以频率估计概率的误差分析	胡嘉仪	(54)
线性回归方法与相关系数在足球分析中的应用	金嘉明	(56)
小概率事件在实际中的应用	李兆欣	(60)
一元线性回归中 F 检验法与相关系数检验法的联系	刘华砾	(62)
二项分布对正态分布的逼近模拟分析	卢舒云	(65)
一元非线性回归问题的线性化处理	王子涵	(70)
资深赌徒脑中的概率应用	杨 博	(77)
贝叶斯公式的简要分析及推广	张 虎	(79)
二维随机变量差和积的分布及特殊分布的拓展	张家豪	(85)

浅谈小概率事件原理在实际中的应用	张新豫	(90)
单侧假设检验矛盾性的探究	张益恺	(93)
成电讲坛次数概率模型的参数假设检验及解释	陈圣坤	(97)
利用最小二乘法拟合物理实验数据	张新宇	(101)
伴随矩阵的特征值与特征向量	刘淏楠 陈 阳	(105)
求函数在特殊区间的最值	明安杰 袁 茵	(108)
积分型余项的泰勒公式的证明及应用	梁宇浩	(110)
求函数零点的特殊方法	王 池 彭 博 孙乙赫	(113)

基于抽样分布定理的对于正态总体下的考试分数波动的研究

何世图 (2018021409023)

指导教师：陈绍刚

【摘要】 本文根据抽样分布定理，研究了题目分数波动为正态总体下的个体试卷分数波动分布规律，并且基于此规律，给出了处于不同分数波动的考生人数，并且根据研究结论，提出了合理化建议。

【关键词】 分数波动；抽样分布

引言

在大型考试中，由于老师的判卷时间短，判卷工作量大，因此，即便每位老师只对某几道题给出分数，仍然存在受主观因素而出现的给分波动情况。基于此种情况，本文建立模型，探究了对某一场考试，任一学生出现分数波动的情况的概率；以及全年级中，处于不同分数波动区间的同学的人数。并根据结论，给出一些简化查分工作量的建议。

1 模型假设

1. 假设一张卷子中，共有 n 道主观题
2. 假设每道主观题所占分数均相等，为 S_n 分。
3. 假设每位判卷老师在对待每一道题时，给出分数的波动率互相独立且同分布。
4. 假设每位判卷老师在对待每一道题时，给出分数的波动率服从正态分布 $N(\mu, \sigma)$ 。
5. 假设客观题不会有分数波动。

2 名词解释及变量说明

变量符号	变量含义
X_i	第 i 道主观题的分数波动
\bar{X}	一位同学总分波动的平均值
n	主观题数目
S_n	一道主观题的分数
μ	正态分布中的均值
σ	正态分布中的方差
E_r	一张试卷中的总分数波动

3 模型建立及求解

3.1 单一主观题的分数波动模型

我们知道，对于一位经验丰富的老师来说，一道题准确给分的概率很大，而波动给分的概率很小，且绝大多数的波动给分均在正确分数附近徘徊。

基于此，同时考虑到模型的简化，因此，我们假设每位判卷老师在对待每一道题时，给出分数的波动 X_i 互相独立且同分布，且 X_i 均服从正态分布 $N(\mu, \sigma)$ 。

并且，根据 μ, σ 的参数性质，如果一道主观题的分数为 S_n ，则我们假设：

$$\begin{cases} \mu = 0 \\ \sigma = \frac{S_n}{10} \end{cases}$$

即 X_i 的期望为 0, $X_i > 3\sigma$ 是小概率事件, 换句话说, 即判卷老师可以将一道题的分数波动控制在这道题的分数的±30%以内。

3.2 主观题的总分波动模型

根据数理统计的基本概念, 可以知道,

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

即

$$\bar{X} \sim N(0, \frac{S_n}{10n}) \quad (2)$$

因为

$$\bar{X} = \frac{1}{n} \sum_1^n X_i \quad (3)$$

所以一张试卷中的主观题分数波动 $Er = \sum_1^n X_i$ 有:

$$Er = n\bar{X} \quad (4)$$

有

$$P\{e_1 < Er < e_2\} = P\left\{\frac{e_1}{n} < \bar{X} < \frac{e_2}{n}\right\} \quad (5)$$

所以, 对于给定区间 (e_1, e_2) , 首先求出对于一道题的平均波动区间 $(e_1/n, e_2/n)$, 再根据

$$\bar{X} \sim N(0, \frac{S_n}{10n}) \quad (6)$$

算出

$$P\left\{\frac{e_1}{n} < \bar{X} < \frac{e_2}{n}\right\} \quad (7)$$

4 基于模型对于各种问题的分析

4.1 分数波动区间的概率

假设, 一张卷子中, 共有十道主观题, 且主观题总分为 60 分, 且不会发生错误计算分数的情况。

同时, 考虑到, 主观题内, 也有一部分题目类似于客观题, 不容易出现主观上导致的分数波动。因此, 我们假设受判卷老师主观因素影响的总分为 30 分, 且平均分布。

根据模型, 如果分数波动在 $(-3, 3)$ 区间内, 那么它出现的概率为: $P\left\{\frac{-3}{10} < \bar{X} < \frac{3}{10}\right\}$, 且 $\bar{X} \sim N(0, 0.3)$ 。因此根据公式可以算出 $P = 68.269\%$ 。以此类推, 可以得到更多分数波动区间出现的概率。

则可得出如下结论:

一名学生的分差在正负 3 分以内的概率为: 68.269 %

一名学生的分差在正负 5 分以内的概率为: 90.441 %

一名学生的分差在正负 7 分以内的概率为: 98.037 %

一名学生的分差在正负 10 分以内的概率为: 99.914 %

一名学生的分差在正负 15 分以内的概率为: 100 %

4.2 分数波动区间的人数

假设一个年级共有 5000 人参加考试, 根据 4.1 结论, 则有:

分差在正负 3 分以内的人数为: 3413

分差在正负 5 分以内的人数为: 4522

分差在正负 7 分以内的人数为: 4902

分差在正负 10 分以内的人数为：4996

分差在正负 15 分以内的人数为：5000

可以看出，90.4%的人的分数波动在正负 5 分以内，对成绩有一些影响。

4.3 建议：

1. 在核对分数的时候，应对分数在 80 分左右以及 55 分左右的同学给予更大关注，看其分数是否发生较大的主观因素波动导致未满绩或者未及格。

因为对于这两类同学，有满绩以及及格的需求，对分数较为看重。而根据结论，90.4%的人的分数波动在 5 分以内，因此，若真实成绩在 85 分、60 分，且其成绩发生了波动，其波动成绩最可能在 90-80 分以及 65-55 分这两个区间。

而对于 90-85，以及 65-60 分数区间的同学，分数波动对其造成的损失远没有分数在 85-80，以及 60-55 区间的同学大。因此，在时间、精力有限的情况下，应该对分数区间在 85-80，以及 60-55 分的同学给予更大关注，努力消除他们卷面上因为判卷老师主观因素造成的分数波动。

2. 由于学校目前并不开放针对题目是否判改正确的查分，建议未来对分数区间在 85-80，以及 60-55 分的同学开放请求重新核查试卷的渠道。

3. 每次考试判卷工作完成之后，应对总体试卷进行抽样调查。如果分数波动情况符合 4.1 中的结论，那么，此次判卷工作是较为成功的；如果分数波动情况较 4.1 中的结论有较大增加，则说明判卷过程中出现了较大的主观波动，总体成绩并不客观，要考虑判卷工作中是否出现了一些问题；反之，如果分数波动情况较 4.1 中结论有明显减少，则说明此次判卷过程中，分数受主观因素影响较少，比较客观，可以反映出学生的真实水平，应总结经验，加以推广。

参考文献

[1] 王梓坤.论随机性[J].数学通报.2004(03)

[2] 张树美.正态常相关样本的抽样分布和方差分析[J].广西师院学报(自然科学版).1994(01)

探究利用独立同分布极限定理近似计算二项分布随机变量序列的误差以及校正方法

刘明康 (2018140601023)

指导教师：陈绍刚

【摘要】 利用 matlab 进行模拟和估计利用独立同分布极限定理近似计算二项分布随机变量序列的误差以及在此基础上探究更好的估计方法

【关键词】 极限同分布定理；二项分布

1 引言

由于在平时的作业中有利用独立同分布极限定理估计二项分布的题目，在估计过程中由于题目给出的时间次数比较小，所以产生了较大的误差，本文意在探究误差大小与二项分布参数的关系，以及通过改进近似方法来提高估算准确度。

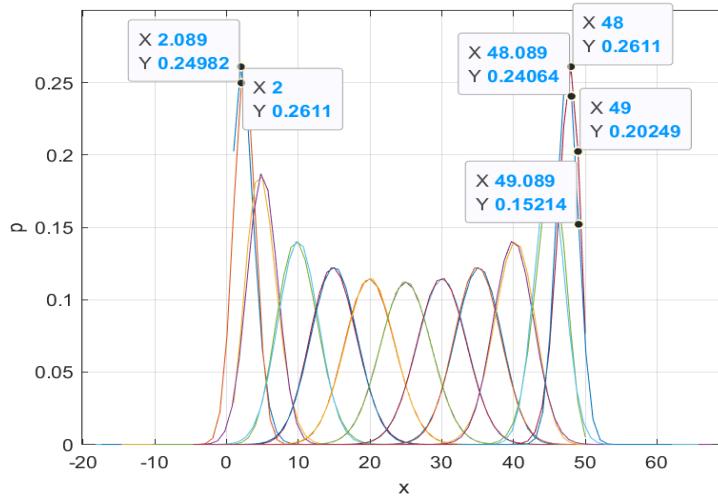
2 正文

由于二项分布序列实际上具有上限 n 和下限 0，而在近似后的正态分布中不能很好的体现这个上限和下限的实际意义，在计算正态分布的过程中势必由于在 $(-\infty, 0)、(n, +\infty)$ 部分的概率导致一些误差。

--误差分析

设：随机事件序列共 n 项；每个随机事件发生的概率为 p ；

则对于随机变量的分布情况进行 matlab 模拟得（代码见附录 1）：

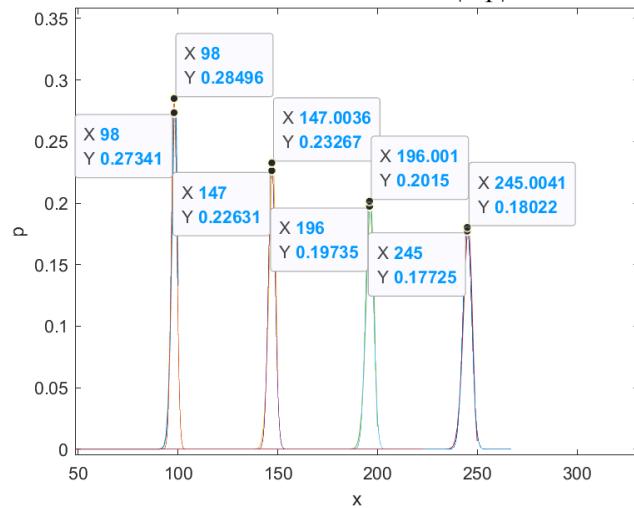


自左向右每组分别为 $n=50$ 的情况下， p 分别等于 0.05、0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、0.95 时的二项分布与其模拟得到的正态分布的图像。可以看到 p 越接近 0.5，则估计的正态分布曲线和二项分布的计算结果的曲线重合得越好；越接近 0 或越接近 1，由于二项分布序列具有下限 0，和上限 n ，带来的误差就越明显。在 $p=0.05$ 以及 $p=0.95$ 时，可见估计得到的概率密度的误差可达 0.06.

由于二项分布的概率为 $P\{X=k\}=C_n^k \times p^k \times (1-p)^{n-k}$ ，所以 $P\{p, X=k\}=P\{1-p, X=n-k\}$ 。所以两个 n 相同， p 之和等于 1 的二项分布概率曲线关于 $x=np$ 对称。

由于正态分布在方差相同，期望不同时，概率密度图像形状不变，又 $D=n \times p \times (1-p)$ ，所以两个和为 1 的二项分布的拟合得到的两个正态分布的曲线形状相同，因为关于 np 对称。

由二项分布和正态分布上述的对称特性，得拟合的精确性与 $|1-p|$ 成正相关



自左向右每组分别为 $p=0.98$ 的情况下， n 分别等于 100、150、200、250 时的二项分布与其模拟得到的正态分布的图像。可以看到 n 越大，则估计的正态分布曲线和二项分布的计算结果的曲线重合得越好。由图像可知两概率密度曲线在最高点的相对误差分别为

n	100	150	200	250
概率密度最高点 相对误差	4.2%	2.8%	2.1%	1.7%

可见正态分布拟合的精确度与 n 成正相关。

--改进方法

1. 在已经学习过的分布中，泊松分布去除了小于下限 0 的概率的影响。当 $|0.5-p|$ 十分接近 0.5 时，可以利用 $p \rightarrow 0$ ，或 $(1-p)$ 构造泊松分布进行近似。

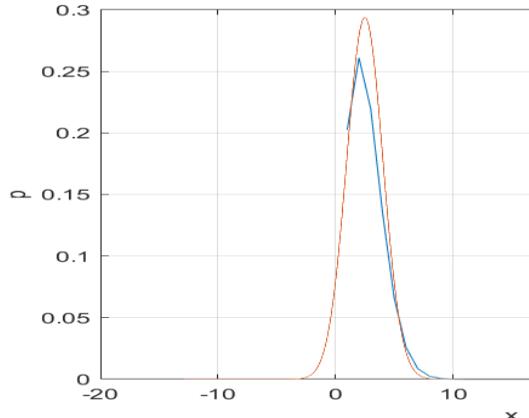
2. 计算过程中如果 $|0.5-p|$ 十分接近 0.5，则利用条件概率公式对正态分布的概率密度进行修正：
 $f'(x)=f(x)/\Phi(t)$

【 $p \rightarrow 0$ 】时 $t=[E(x)]/(D(x))^{0.5}$

【 $p \rightarrow 1$ 】时 $t=[n-E(x)]/(D(x))^{0.5}$

对于 $n=50, p=0.05$

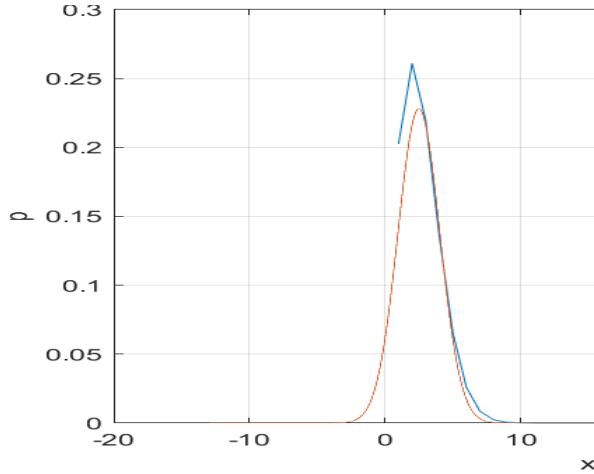
图像为：（蓝色为二项分布图像，红色为模拟图像）



可见新建拟合函数在 $x < np$ 时拟合效果较好，而在 $x > np$ 时误差较大。

于是根据概率密度的归一性进行进一步拟合，由于而正态分布概率密度对于 $x=np$ 对称，所以若 $f'(x)=f(x) \times \Phi(t)$ ($x < np$) 则要使归一性条件达到需要 $f'(x)=f(x)/\Phi(t)$ ($x > np$)

于是 $f'(x)=f(x)/\Phi(t)$ 的图像与二项分布结果对比如下：(蓝色为二项分布图像，红色为模拟图像)



得出在 $x > np$ 时具有良好的拟合效果。

于是得到最终的校正后的函数如下：

$$f'(x)=f(x)/\Phi(t) \quad (x < np)$$

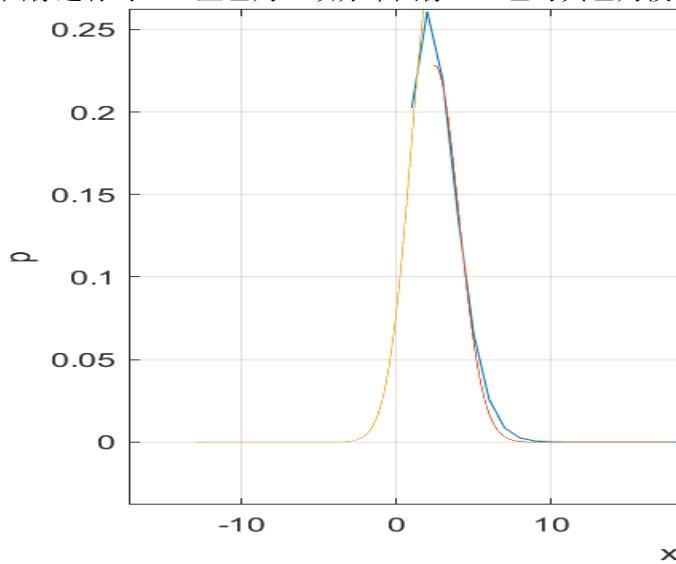
$$f'(x)=f(x)/\Phi(t) \quad (x > np)$$

【 $p \rightarrow 0$ 】 时 $t=[E(x)]/(D(x))^{0.5}$

【 $p \rightarrow 1$ 】 时 $t=[n-E(x)]/(D(x))^{0.5}$

并对 $x=np$ 一定邻域范围内，取两者的平均值进行平滑处理。(代码见附录 2)

绘图与二项分布图像进行对比 (蓝色为二项分布图像，红色与黄色为模拟图像)



得到拟合效果较好的正态分布图像。

3 结论

1. 通过同分布极限定理拟合二项分布的精确度与 $|1-p|$ 成正相关，与 n 成正相关。

2. 对于二项分布的 p 接近 0 或接近 1 时，可以根据下面的 $f'(x)$ 对距离 $x=np$ 不太近的概率密度。

$$f'(x)=f(x)/\Phi(t) \quad (x < np)$$

$$f'(x)=f(x)/\Phi(t) \quad (x > np)$$

【 $p \rightarrow 0$ 】时 $t = [E(x)]/(D(x))^{0.5}$

【 $p \rightarrow 1$ 】时 $t = [n - E(x)]/(D(x))^{0.5}$

参考文献

- [1] 万祥兰.二项分布的近似计算与应用举例[J].科技视界,2019(23):135-137.

4 附录

1

Matlab 运行代码:

```
n=250;
p=0.98;
fer=[];
fer(1)=n*p*(1-p)^(n-1)
for i=2:n
    fer(i)=nchoosek(n,i)*p^(i)*(1-p)^(n-i); %计算二项分布各个可行性的概率
end
plot(fer);
hold on;
e=n*p;
dk=(n*p*(1-p))^(0.5);
x = (e-10*dk:0.01:e+10*dk);
y = normpdf(x,e,dk); %拟合期望为 n*p, 标准差为(n*p*(1-p))^(0.5) 的正态分布曲线
plot(x,y);
grid;
xlabel('x'); ylabel('p')
```

2

Matlab 运行代码:

```
n=50;
p=0.05;
fer=[];
fer(1)=n*p*(1-p)^(n-1)
for i=2:n
    fer(i)=nchoosek(n,i)*p^(i)*(1-p)^(n-i);
end
plot(fer);
hold on;
e=n*p;
dk=(n*p*(1-p))^(0.5);
x1 = (e:0.01:e+10*dk);
y1= normpdf(x1,e,dk)*0.8810; %前半段拟合 0.8810=Φ(t)
x2=(e-10*dk:0.01:e);
y2= normpdf(x2,e,dk)/0.8810; %后半段拟合
plot(x1,y1);
hold on;
plot(x2,y2);
grid;
xlabel('x'); ylabel('p')
```

有关假设检验的两类错误和检验水平的研究

陈思佳 (2018040701026) 孟翔晨 (2018021412013)

指导老师: 高晴

【摘要】 在充分理解和分析假设检验两类错误的基础上, 对二者的异同进行合理分析, 研究两类错误分别对总体预测的影响, 并讨论第一类错误优先级高于第二类错误的原因, 同时进一步分析讨论使犯两类错误概率都尽可能小的决策方法。

【关键词】 假设检验; 弃真; 纳伪; 小概率原理

1 引言

数理统计的核心内容是采用参数估计和假设检验的方法, 利用样本推断总体。对于假设检验而言, 可以实现根据某种实际需要, 预先对未知总体作出假设, 再根据样本实测的结果检验假设的合理性, 最后决定对该假设的取舍。但由于以样本预测总体时, 采用小概率原理, 而我们知道小概率事件并非不会发生, 因此会出现两种情况: 一为“弃真”, 即原假设 H_0 正确, 但由于样本本身的随机性致使其落入拒绝域; 二为“纳伪”, 即原假设 H_0 错误, 但由于样本本身的随机性致使其落入接受域[1]。由于假设检验结果非此即彼, 即不是接受原假设 H_0 (意味着存在“纳伪”的可能性), 就是拒绝原假设 H_0 (意味着存在“弃真”的可能性), 因此错误概率不可避免, 只能够使其尽可能小, 因此我们对二者关系与影响条件进行了研究。

2 基本概念

2.1 小概率原理

所谓“小概率原理”即指: 当某个事件在随机试验中发生的概率很小时, 可以认为该事件在一次随机试验中出现概率极小, 以至于看作单次试验中不会出现[2]。因此利用小概率原理可以辅助判断假设检验中原假设是否可信, 若抽取样本值与假设结果相差较大(落入拒绝域), 则拒绝原假设, 但这种判断并非绝对正确, 犯这种错误的可能性大小取决于事先规定的小概率事件“小”的程度。这种判断小概率的标准是统计假设检验中所谓的显著性水平。

2.2 第一类错误・弃真

设原假设为 H_0 , 对立假设为 H_1 。

当原假设 H_0 确实为真, 但此时的所得的检验统计量位于计算所得拒绝域内, 我们做出拒绝 H_0 的判断, 此时发生第一类错误, 将其概率记为 α , α 被称为假设检验的显著性水平或检验水平, 满足以下概率等式:

$$\alpha = P(\text{拒绝 } H_0 | H_0 \text{ 为真}) = P(|u| > u_{\alpha/2} | H_0 \text{ 为真}) [3]$$

选取正态总体 $X \sim N(\mu, \sigma^2)$, 方差 σ^2 已知的分析示例, 对总体均值 μ 进行检验, 显著性水平为 α 。

设原假设为 $H_0: \mu = \mu_0$, 对立假设为 $H_1: \mu > \mu_0$ 。

当原假设 H_0 成立时, 构建检验统计量为 $U = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$, 由于在原假设前提下, μ 越大越不利, 因此为单侧检验, $\bar{X} - \mu_0$ 若偏大为小概率事件, 根据文献[4]可以得出下式:

$$P \left\{ \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq u_{\alpha/2} \right\} = \alpha \quad (1)$$

$$P\left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq \frac{\sigma}{\sqrt{n}} u_{\alpha/2}\right\} = \alpha \quad (2)$$

得到原假设 H_0 的拒绝域为 $u > u_{\alpha/2}$, 或 $\bar{X} - \mu_0 > \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$, $A = \mu_0 + \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$ 如图 1 所示:

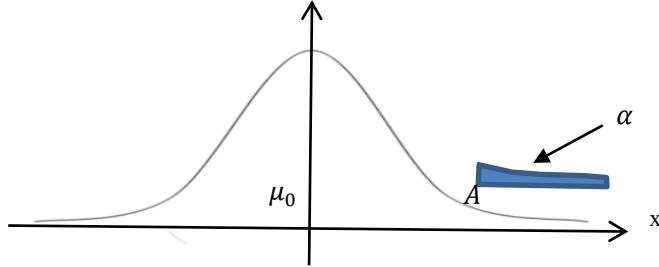


图 1 第一类错误概率示意图

2.3 第二类错误 · 纳伪

设原假设为 H_0 , 对立假设为 H_1 。

当原假设 H_0 确实为假, 但此时的所得的检验统计量位于计算所得接受域内, 我们做出接受 H_0 的判断, 此时发生第二类错误, 将其概率记为 β , 满足以下概率等式[3]:

$$\beta = P(\text{接受 } H_0 | H_0 \text{ 为假}) = P(|u| \leq u_{\alpha/2} | H_0 \text{ 为假})$$

选取正态总体 $X \sim N(\mu, \sigma^2)$, 方差 σ^2 已知的分析示例, 对总体均值 μ 进行检验, 显著性水平为 α 。

设原假设为 $H_0: \mu = \mu_0$, 对立假设为 $H_1: \mu > \mu_0$ 。

当原假设 H_0 不成立时, 即 $\mu = \mu_1 > \mu_0$, 构建检验统计量为 $U = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$, 由抽样可知, $\bar{X} \sim N\left(\mu_1, \frac{\sigma^2}{n}\right)$, 由于在原假设不成立前提下, μ 越小越不利, 因此为单侧检验, $\bar{X} - \mu_0$ 若偏小为小概率事件, 根据文献[4]可以得出下式:

$$P\left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq u_{\alpha/2}\right\} = \beta \quad (3)$$

$$P\left\{\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} \leq \frac{\mu_0 + \frac{\sigma}{\sqrt{n}} u_{\alpha/2} - \mu_1}{\sigma/\sqrt{n}}\right\} = \beta \quad (4)$$

$$u_{\beta/2} = \frac{\mu_0 + \frac{\sigma}{\sqrt{n}} u_{\alpha/2} - \mu_1}{\sigma/\sqrt{n}} \quad (5)$$

$$\mu_1 - \mu_0 = \frac{\sigma}{\sqrt{n}}(u_{\alpha/2} - u_{\beta/2}) \quad (6)$$

得到 H_1 的拒绝域为 $\bar{X} - \mu_1 < \mu_0 + \frac{\sigma}{\sqrt{n}} u_{\alpha/2} - \mu_1$, $A = \mu_0 + \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$, 如图 2 所示:

(6)

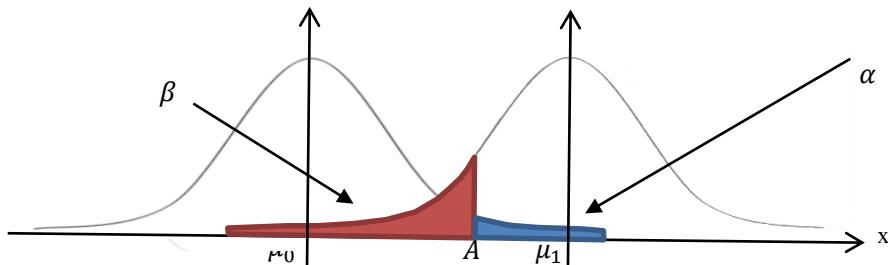


图 2 第二类错误概率示意图

3 两类错误的关系及产生影响因素

3.1 两类错误的关系

第一类错误与第二类错误之间存在着紧密的联系。当样本的容量确定后，是不可能同时控制犯第一类错误的概率和第二类错误的概率。犯第一类错误的概率 α 越小。拒绝域越小，从而拒绝 H_0 的可能性也越小，犯第二类错误的概率便会增大。两类错误关系为相互制约。两类错误关系总结见表格 1。

表格 1 两类错误关系示意表格

判断	判断正误 \ 真实情况	H_0 为正确	H_0 为错误
拒绝 H_0		犯第一类错误	正确
		正确	犯第二类错误

现举例具体理解并说明二者关系：

若现在对 A（坏人）进行判决，出现以下两种假设： H_0 : A 是好人 H_1 : A 是坏人。

情况 1：如果 A 确实是好人，但是判决结果为坏人，则犯了第一类错误“弃真”；

情况 2：如果 A 不是好人而是坏人，但是判决结果为好人，则犯了第二类错误“纳伪”。

于是法官请来了 100 个人作为陪审团，并问了他们如下 2 个问题：

问题 1：A 是不是好人？

问题 2：A 是不是坏人？

有一些陪审员坚定的认为 A 是好人，不是坏人，另一些人认为 A 是坏人，不是好人，但是还有一些人说不准，觉得 A 应该不是好人，但是又不能说 A 是坏人。现在对两个问题的投票结果如下显示：

表格 2 陪审团 1 投票结果

	好人	坏人
是	3	45
不是	97	55

由此可以得出，100 个人里，有 97 个说他不是好人，但是还有 3 个人坚定的说 A 是好人。因此对于情况 1，可以得出误判概率 α 为 0.03，可以理解为：A 确实是好人，但误判犯错的概率是 0.03。另一方面，100 个人里面却只有 45 个人认为，他应该是坏人。所以对于情况 2，可以得出误判概率 β 是 0.55，可以理解为：A 是坏人，但误判犯错的概率是 0.55。

根据两个问题的投票结果，分别得出了犯两类错误的概率，那么这时法官如何而进行判决？

此时给定一些前提条件：

如果判断意向为：宁愿“错杀”三千，也不愿放过一个。这时就需要尽可能减少把坏人当作好人的概率，让第二类错误“纳伪”的概率尽可能的小，只要陪审团有足够的认为他是坏人，那么就判定他是坏人。可以从此组投票中看出，对 A 是否为坏人的判决投票情况鉴别程度不明显，误判概率极大。

如果此时选取判断意向为：疑罪从无，无法证明其有罪则以无罪判决。此时如果没有足够多的人说明 A 并非好人的话，则可以将其以好人看待。若判决前提的置信水平比 0.03 大，则可以看作好人，翻制若置信水平比 0.03 小，则看作坏人。

那么如何看待二者相互制约的关系呢？从上述投票中可以看出，由于对第二个问题的投票情况含糊不明，使得结论有些偏颇，若我们抽取另外 100 人作为第二陪审团，同样对两个问题进行投票，得出表 3 结果：

表格 3 陪审团 2 投票结果

	好人	坏人
是	3	90
不是	97	10

可以看出，此时犯第一类错误的概率是 $\alpha = 0.03$ ，犯第二类错误的概率为 $\beta=0.1$ 。两组陪审团相比较可以发现，犯第一类错误概率相同，但陪审团 2 犯第二类错误的概率更小。可以表示为，控制犯第一类错误概率相同条件下，使得犯第二类错误的概率更小的抽样结果，此时两种错误共同发生对判决结果影响较小。

3.2 两类错误产生的影响因素

根据示例可以分析得出

第一类错误产生原因：

1、样本中存在极端数值：在 σ^2 较小，总体分布较集中的情况下，检测到的样本为 μ_1 且 $\mu_1 \gg \mu$ 或 $\mu_1 \ll \mu$ ，此时会出现第一类错误。

2、采用决策标准较宽松：总体的均值为 μ ，而 H_0 的标准应为 $|\mu_1 - \mu| < \sigma_1$ ，而实际上 $|\mu_1 - \mu| < \sigma$, $\sigma_1 \ll \sigma_2$ ，此时会出现第一类错误。

第二类错误产生原因：

1、实验设计不灵敏： H_1 设计的范围过大，使得样本值变化时无法准确判断检验统计量所在范围。

2、样本数据变异性过大：样本的方差 σ^2 过大，使得总体分布不均匀。

3、处理效应本身比较小：数据容量 n 本身过小，造成不确定因素过多，最终会导致第二类错误的产生。

4 决策方法

由上述分析结合解题示例可知，通常情况下对于一总体而言，标准差 σ 为已知条件，而均值 μ 未知， α 与 β 相互制约，但由所画正态分布函数图像以及计算所得式可以看出，在其他量不变的条件下，减小 α ，使拒绝域变小的同时，上侧分位数向右移动，导致 β 变大；增大 α ，使拒绝域变大的同时，上侧分位数向左移动，导致 β 变小，因此在同一抽样结果下，犯两类错误概率不可同时减小。示意图见图 3。

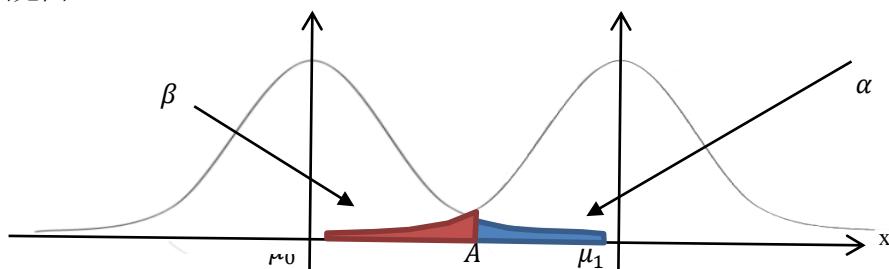


图 3 两类错误概率示意图

在其他量不变的条件下，增大 μ_1 ，可减小 β 。

同时根据式子可得：

$$\mu_1 - \mu_0 = \frac{\sigma}{\sqrt{n}}(u_{\alpha/2} - u_{\beta/2}) \quad (7)$$

在其他量不变的条件下，若样本容量 n 增大，则可以同时减小 $u_{\alpha/2}$ 和 $u_{\beta/2}$ ，即减小 α 与 β ，可以理解为，当样本容量增大时，图像会向对称轴集中，因此靠近对称轴部分面积会更大，靠近尾部部分面积会减小。

5 结论

通过上述分析可以得出，对于假设检验的两种错误犯错概率的决策控制方法应该采取：控制犯第一类错误概率相同条件下，使得犯第二类错误的概率更小，此时的抽样结果中两种错误对判决结果影响更小。由于第一类错误“弃真”影响效果更为突出，应该考虑“疑罪从无”原则，即小心接受，大胆拒绝。因此为了使假设检验结论可信度更高，应该巧妙设定原假设条件，以拒绝原假设 H_0 为目的，证明 H_0 位于拒绝域中，可以将传统的、问题中原始的条件设定为原假设 H_0 ，同时原假设中包括等号，将影响后果更为严重的错误设定为第一类错误 0。而在同一次实验中，两种错误共同发生的概率互相制约，若想要同时减小犯错概率，则可以采用选取合适假设检验条件或适当增大样本容量的方法。

参考文献

- [1] 徐全智, 吕怒, 概率论与数理统计, 高等教育出版社, 2010.
- [2] 吴启富, 张玉春, 统计假设检验中小概率原理的辨析, 统计与决策, 2012, 第 17 期, 70-71.
- [3] 黄艳华, 郑芳琳, 参数假设检验中两类错误的剖析, 钦州学院学报, 2018, 第 33 卷第 3 期, 22-24.
- [4] 高发玲, 刘春霞, 孙建英, 郭萍, 单正态总体假设检验的两类错误概率关系分析, 重庆工商大学学报(自然科学版), 2015, 第 32 卷第 9 期, 58-60.

有关样本均值的分布的研究

舒婷 (2018151202006) 胡智成 (2018021404013) 苏力 (2018021404020)

指导教师：高晴

【摘要】 在充分认识样本均值和中心极限定理基础上，利用 MATLAB 软件，研究总体分布为正态分布的样本均值的分布，同时进一步分析讨论总体分布不服从正态分布的样本均值的分布。

【关键词】 总体分布；样本均值；正态分布；MATLAB 仿真

1 引言

假设 X_1, X_2, \dots, X_n 是一个服从正态分布的总体 $N(\mu, \sigma^2)$ 的样本， \bar{X}, S^2 分别是样本均值和样本方差[1]。则有：

(1) \bar{X} 服从 $N\left(\mu, \frac{\sigma^2}{n}\right)$ 分布；

(2) $\frac{\bar{X} - \mu}{S/\sqrt{n}}$ 服从 $t(n-1)$ 分布

然而，当总体 X 不服从正态分布时，其样本均值的分布往往复杂得多，但根据中心极限定理，我们有理由猜想，当 n 充分大时， \bar{X} 近似服从 $N(\mu, \frac{\sigma^2}{n})$ 分布， $\frac{\bar{X} - \mu}{S/\sqrt{n}}$ 近似服从 $N(0,1)$ 分布[2]。

2 实验过程

由理论知识我们不难发现，当总体服从正态分布的时候，其样本均值 \bar{X} 服从 $N(\mu, \frac{\sigma^2}{n})$ 在本节我们用 MATLAB 仿真来验证这一结论；当总体不服从正态分布时，其样本均值的分布往往复杂的多。

但根据中心极限定理，我们大胆进行猜想，此时的样本均值 \bar{X} 仍然服从 $N(\mu, \frac{\sigma^2}{n})$ ，在这一节里，我们用 MATLAB 仿真来验证我们这个猜想是否正确，并采用正态分布检验图（点分布在一条直线上则符合正态分布），以及 Jarque-Bera 检验， h 输出值为 0 即为正态分布。

2.1 数据来源

根据独立同分布中心极限定理，随机变量序列依分布收敛于标准正态分布随机变量[3]。因此可知，现实中许多随机变量可以表示成为大量相互独立的随机变量之和，而且其中每一个随机变量对总和起到微小的影响，这类随机变量往往服从或近似服从正态分布[4]。因此我们采用下面所述数据进行实验探究。

实验采用未知分布来源于 2019 年电子科技大学美赛模拟赛 B 题，其数据量为 4908，代表天球坐标系中星空坐标系所探测的恒星的赤经数值，取值范围为 $[0, 360]$ ，其分布未知。其可以看成是一个未知分布的总体，本文以该数据为代表来计算未知分布的样本均值的分布。

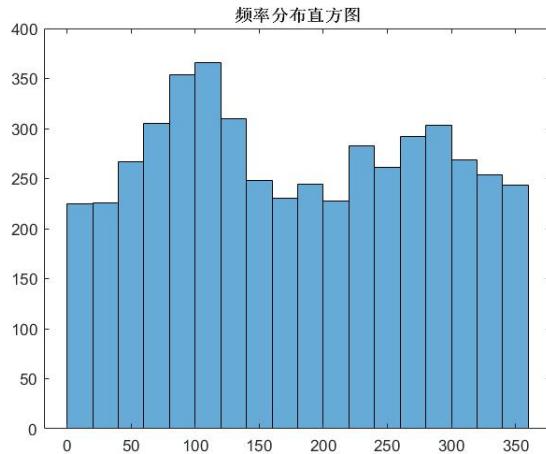


图 1 未知分布总体的频率分布直方图

2.2 当总体服从正态分布时，样本均值服从的分布的验证

本节利用 matlab 画出当总体分布为正态分布时，样本均值服从的分布密度函数图，并画出正态分布检验图对样本均值的分布函数进行检验。

MATLAB 示例代码如附件代码一所示：

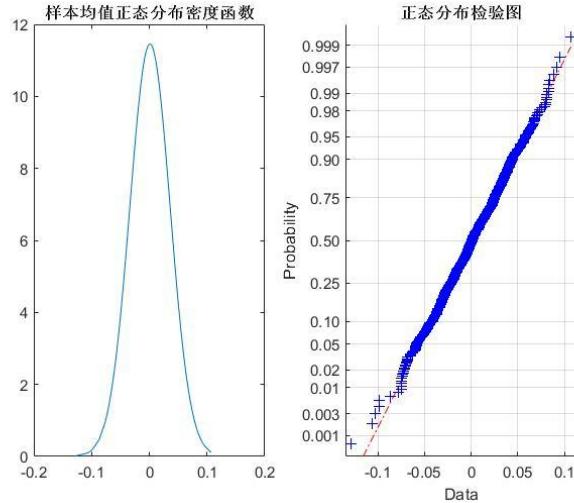


图 2 当总体服从正态分布时，样本均值服从的分布密度函数和正态分布检验图

表 1 $n=800$ 时，当总体服从正态分布，样本均值的均值和方差和正态分布检验结果

	样本均值的 \bar{X}	样本均值的 S^2	h
$n=800$	0.0044	0.0012	0

从图中和表中可以看出，当总体服从正态分布时，其分布密度函数与正态分布的密度函数非常相近，对其进行正态分布检验，其正态分布检验图近似为一条直线，且 $h=0$ ，可知其通过正态分布检验，因此可将其看成是近似服从正态分布。

此时根据样本均值公式： $\bar{X} = \frac{1}{n} \sum_{i=0}^n X_i$

样本方差公式： $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

计算得样本均值的均值为 0.0044，样本均值的方差为 0.0012，理论上样本均值的均值为 0，样本均值的方差为 0.00125，两个数据吻合完好。即当总体服从正态分布时，其样本均值也服从正态分布。

2.3 当总体服从均匀分布时，样本均值服从的分布的验证

与上一小节手法类似，本小节研究当总体服从均匀分布时，将样本均值服从的分布密度函数图和对其进行的正态分布检验图画出来，对其是否继续服从正态分布进行研究。

MATLAB 示例代码如附件代码二所示：

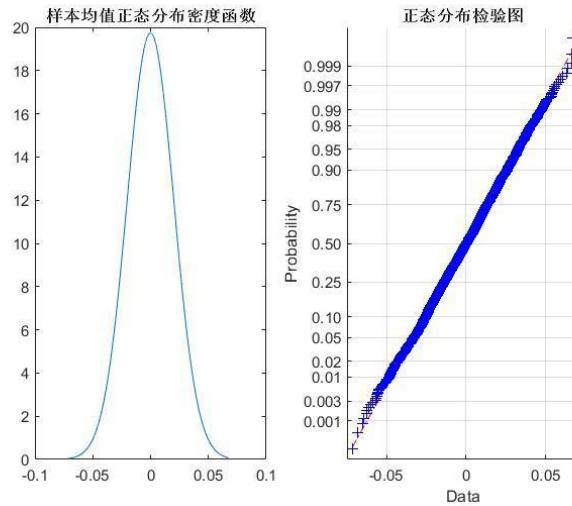


图 3 当总体服从均匀分布 $U(-1,1)$ 时，样本均值的分布密度函数和正态分布检验图

表 2 当总体服从均匀分布 $U(-1,1)$ 时，样本均值的均值和方差及正态分布检验结果

	样本均值的 \bar{X}	样本均值的 S^2	h
$n=800$	6.9154×10^{-4}	4.1985×10^{-4}	0

从图中可以看出，当总体服从均匀分布 $U(-1,1)$ 时，样本均值的分布密度函数与正态分布函数十分近似，对其进行正态分布检验，得出来的曲线接近一条直线，且 $h=0$ ，由此可以说明这个时候样本均值通过正态分布检验。且样本均值的均值为 6.9154×10^{-4} ，样本均值的方差为 4.1985×10^{-4} ，猜想中样本均值的均值应该为 0，样本均值的方差应该为 0.00041667 ，即当总体分布是均匀分布时，其样本均值也可近似为正态分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$ 。

2.4 当总体服从未知分布时，样本均值服从的分布的验证

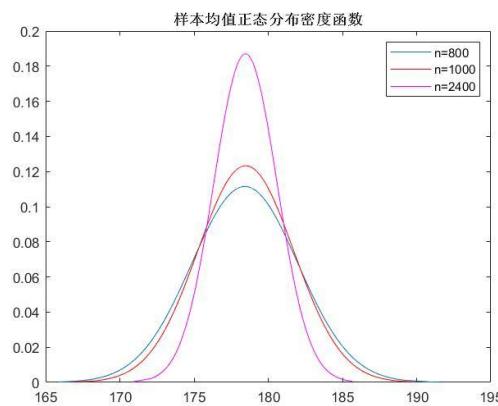


图 4 当总体为未知分布时，样本均值服从的分布密度函数曲线 ($n=800, 1000, 2400$)

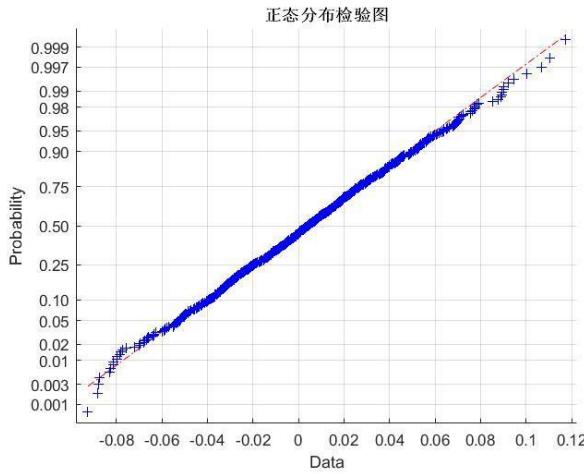


图 5 当总体为未知分布时，样本均值的正态分布检验图

表 3 当总体服从位置分布,样本均值的均值和方差及正态分布检验结果($n=800、1000、2400$)

	样本均值的 \bar{X}	样本均值的 S^2	h
$n=800$	178.4372	12.3738	0
$n=1000$	178.4900	10.4114	0
$n=2400$	178.4752	4.2797	0

验证代码见附件代码三，从图中可以看出，当总体服从未知分布， $n=800、1000、2400$ 时，样本均值的分布密度函数与正态分布函数近似。进而对其进行正态分布检验，得出曲线接近一条直线，且 $h=0$ 。由此，可以说此时样本均值通过正态分布检验。

当 $n=800$ 时，样本均值的均值为 178.4372，样本均值的方差为 12.3738，可以猜想样本均值的均值为 178.4531，样本均值的方差为 13.003265，因此理论值与实验值符合完好。

当 $n=1000$ 时，样本均值的均值为 178.4900，样本均值的方差为 10.4114，猜想中样本均值的均值为 178.4531，样本均值的方差为 10.4002612，理论值与实验值符合完好。

当 $n=2400$ 时，猜想中样本均值的均值为 178.4531，样本均值的方差为 4.33344217，同样，理论值与实验值符合完好。

即当总体分布是未知分布时，其样本均值也可近似为正态分布 $N(\mu, \frac{\sigma^2}{n})$ ，且可以看出当 n 越大时，样本均值的分布越来越接近于正态分布，进一步猜想当 $n \rightarrow \infty$ 时，样本均值的分布将非常近似于正态分布。

3 结论

通过上述理论猜想和 MATLAB 仿真，我们可以得到以下结论：

当总体服从正态分布时，样本均值服从正态分布 $N(\mu, \frac{\sigma^2}{n})$ 。

当总体服从未知分布时，样本均值近似服从正态分布 $N(\mu, \frac{\sigma^2}{n})$ ，且 n 越大，近似程度越高。

无论当总体服从何种分布，当 n 取很大时，样本均值近似服从正态分布 $N(\mu, \frac{\sigma^2}{n})$ ，且 n 越大，近似程度越高。

参考文献

- [1] 徐全智, 吕恕. 概率论与数理统计, 高等教育出版社, 2010.
- [2] 易秀英, 王三宝. 基于 Mathematica 的中心极限定理的实验分析[J]. 湖北理工学院学报, 2010, 26(1):29-32.
- [3] 韩明. 《概率论与数理统计》中借助数学实验理解几个极限定理[J]. 大学数学, 2013, 29(4):127-131.
- [4] 徐彬. 中心极限定理的形象化模拟及其应用[J]. 湖北汽车工业学院学报, 2016, 30(2):77-80.

导数运算 $(fg)' = f'g + fg'$ 及 $(f/g)' = f'/g - f \cdot g'/g^2$ 的条件

张杰瑞 (2019040401006)

指导老师: 李红

【摘要】 本文讨论了何时 $(fg)' = f'g + fg'$ 这一问题, 采用两种方法从两个方面得出了结论; 并且对上述问题进行了推广, 即何时 $(f/g)' = f'/g - f \cdot g'/g^2$; 最后针对这两个问题给出了应用实例。

【关键词】 导数; 函数

1 问题描述

设 f, g 是定义在区间 I 上关于 x 的可导函数, 分别求满足

$$(fg)' = f'g + fg' \text{ 和 } (f/g)' = f'/g - f \cdot g'/g^2$$

的条件。

2 求解过程

2.1 该部分用两种方法讨论了 $(fg)' = f'g + fg'$ 的条件

解法一: 由导数四则运算法则可得 $(fg)' = f'g + fg'$, 若 $(fg)' = f'g$, 则 $f'g + fg' = f'g$, 即

$$f'(g' - g) = fg' \quad (1)$$

下面进行分类讨论:

(1) 若 $f \equiv 0$, (1)式显然成立。

(2) 若 $f \neq 0$ 且 $g \neq g'$, (1) 式可整理为 $f'/f = g'/(g' - g)$, 取不定积分得:

$$\int f'dx/f = \int g'dx/(g' - g), \text{ 即 } \ln|f| - \ln|C| = \int g'dx/(g' - g), \text{ 可得}$$

$$f = Ce^{\int g'dx/(g' - g)} (C \neq 0)$$

(3) 若 $f \neq 0$ 且 $g \equiv g'$, 即

$$g = dg/dx \quad (2)$$

若 $g \equiv 0$, (2)式显然成立; 若 $g \neq 0$, (2)式可整理为 $dg/g = dx$, 并取不定积分得: $\int dg/g = \int dx$,

即

$$g = Ce^x \quad (3)$$

将(3)式代入原式 $(fg)' = f'g + fg' = f'g$ 中, 得 $f' \cdot Ce^x + f \cdot Ce^x = f' \cdot Ce^x$, 即 $f = 0$.

综上所述:

1) 当 $f = 0$ ($g = 0$)时, $(fg)' = f'g$ 对任意 g (f)都成立。

2) 当 $f \neq 0$ 且 $g \neq g'$ ($g \neq 0$ 且 $f \neq f'$)时, $(fg)' = f'g$ 成立的条件为:

$$f = Ce^{\int g'dx/(g' - g)} (C \neq 0) \quad (g = Ce^{\int f'dx/(f' - f)} (C \neq 0)) .$$

3) 当 $f \neq 0$ 且 $g \equiv g'$ 时 ($g \neq 0$, $f \equiv f'$), $(fg)' = f'g$ 不可能成立。

解法二: $(fg)' = f'g$, 变形可得 $f/f' + g/g' = 1$, 则两个函数与其导数的商只相差一个常数。令 $f/f' = \varphi$, $g/g' = \psi$ (φ, ψ 是关于 x 的不恒为 0 的函数, 且 $\varphi + \psi = 1$), 则我们有:

$$f = C_1 e^{\int dx/\varphi}, \quad g = C_2 e^{\int dx/\psi}.$$

注：本解法更适合用于两个符合条件的具体函数。

结论

- (1) 当 $f = 0$ ($g = 0$) 时, $(fg)' = f'g'$ 对任意 g (f) 都成立。
- (2) 当 $f \neq 0$ 且 $g \neq g'$ 时, $(fg)' = f'g'$ 成立的条件为: $f = Ce^{\int g'dx/(g'-g)}$ ($C \neq 0$)。
- (3) 当 $f \neq 0$ 且 $g \equiv g'$ 时 ($g \neq 0, f \equiv f'$), $(fg)' = f'g'$ 不可能成立。
- (4) $(fg)' = f'g'$ 成立的条件是 $f = C_1 e^{\int dx/\varphi}$, $g = C_2 e^{\int dx/\psi}$ (φ, ψ 是关于 x 的不恒为 0 的函数, 且 $\varphi + \psi = 1$)。

2.2 该部分讨论了 $(f/g)' = f'/g'$ 的条件

由导数四则运算法则可得 $(f/g)' = (f'g - fg')/g^2$, 若 $(f/g)' = f'/g'$, 则 $(f'g - fg')/g^2 = f'/g'$, 即

$$f'g(g' - g) = fg'^2 \quad (4)$$

下面进行分类讨论:

- (1) 若 $f \equiv 0$, (4) 式显然成立。
- (2) 若 $f \neq 0$ 且 $g \neq g'$, (4) 式可整理为 $f'/f = g'^2/g(g' - g)$, 取不定积分得: $\int f'dx/f = \int g'^2 dx/g(g' - g)$, 即 $\ln|f| - \ln|C| = \int g'^2 dx/g(g' - g)$, 可得

$$f = Ce^{\int g'^2 dx/g(g' - g)} (C \neq 0)$$

- (3) 若 $f \neq 0$ 且 $g \equiv g'$, 即 $g = dg/dx$ 时, 参考 (2) 式, 又 $g \neq 0$ 有

$$g = Ce^x (C \neq 0) \quad (5)$$

将 (5) 式代入原式 $(f/g)' = (f'g - fg')/g^2 = f'/g'$ 中, 即 $(f' \cdot Ce^x - f \cdot Ce^x)/(Ce^x)^2 = f'/(Ce^x)$, 易得 $f = 0$.

结论:

- (1) 当 $f = 0$ 时, $(f/g)' = f'/g'$ 对任意 f 都成立。
- (2) 若 $f \neq 0$ 且 $g \neq g'$ 时, $(f/g)' = f'/g'$ 成立的条件为 $f = Ce^{\int g'^2 dx/g(g' - g)}$ ($C \neq 0$)。
- (3) 若 $f \neq 0$ 且 $g \equiv g'$ 时, $(f/g)' = f'/g'$ 不可能成立。

3 应用举例

例 1 令 $g(x) = \sin x$, 由 $f = Ce^{\int g'dx/(g'-g)}$ 计算可得

$$f = Ce^{\int \cos x dx / (\cos x - \sin x)} = Ce^{[x - \ln(\cos x + \pi/4)]/2}$$

所以当 $f = Ce^{[x - \ln(\cos x + \pi/4)]/2}$ 时, 满足求导关系式 $(fg)' = f'g'$.

例 2 令 $g(x) = x^\alpha$ ($\alpha \neq 0$), 由 $f = Ce^{\int g'dx/(g'-g)}$ 计算可得

$$f = Ce^{\int \alpha x^{\alpha-1} dx / (\alpha x^{\alpha-1} - x^\alpha)} = Ce^{\alpha \int dx / (\alpha - x)} = C|\alpha - x|^{-\alpha}$$

所以当 $f = C|\alpha - x|^{-\alpha}$ 时, 满足求导关系式 $(fg)' = f'g'$.

例 3 令 $f(x) = e^{\lambda x}$ ($\lambda \neq 0$ 且 $\lambda \neq 1$), 由 $g = Ce^{\int f'dx/(f'-f)}$ 计算可得 $g = Ce^{\int \lambda e^{\lambda x} dx / (\lambda e^{\lambda x} - e^{\lambda x})} = Ce^{\lambda x / (\lambda - 1)}$, 所以当 $g = Ce^{\lambda x / (\lambda - 1)}$ 时, 满足求导关系式 $(fg)' = f'g'$.

例 4 令 $f/f' = \cos x$, $g/g' = 1 - \cos x$, 由 $f = C_1 e^{\int dx/\varphi}$, $g = C_2 e^{\int dx/\psi}$ 计算可得
 $f = C_1 e^{\ln|\tan x + \sec x|} = C_1 \ln|\tan x + \sec x|$, $g = C_2 e^{\int dx/[1 - (2\cos^2(x/2) - 1)]} = C_2 e^{-\cot(x/2)}$, 所以当
 $f(x) = C_1 \ln|\tan x + \sec x|$, $g = C_2 e^{-\cot(x/2)}$ 时, 满足求导关系式 $(fg)' = f'g'$.

例 5 令 $g(x) = e^{\lambda x}$ ($\lambda \neq 0$ 且 $\lambda \neq 1$), 由 $f = Ce^{\int g'^2 dx/g(g'-g)}$ 计算可得

$$f = Ce^{\int (\lambda e^{\lambda x})^2 dx / [e^{\lambda x}(\lambda e^{\lambda x} - e^{\lambda x})]} = Ce^{\int \lambda^2 dx / (\lambda - 1)} = Ce^{\lambda^2 x / (\lambda - 1)}$$

所以当 $f = Ce^{\lambda^2 x / (\lambda - 1)}$ 时, 满足求导关系式 $(f/g)' = f'/g'$.

4 总结

《微积分》(上)课本, 导数四则运算中, 加减的导数等于导数的加减, 但是乘积以及商的导数公式则比较复杂。本文则分别讨论了 $(fg)' = f'g'$ 和 $(f/g)' = f'/g'$ 成立的条件, 并给出了应用实例。

参考文献

- [1] 电子科技大学数学科学学院, 微积分(上), 高等教育出版社, 2018

借助反射矩阵快速求解反射光线方程的方法

蔡思源 (2019091602004)

指导教师：李厚彪

【摘要】 光线通常按直线传播，给出一条或若干条入射光线，怎样快速计算出反射光线？本文具体讨论了关于求一条光线经某一平面反射后的反射方程的问题，从 4 个不同角度，给出了四种具有一般性的求解方法。

【关键字】 镜面反射矩阵；向量

问题 1：一条光线沿直线 $l: \frac{x+3}{5} = \frac{y-1}{1} = \frac{z-4}{3}$ 入射，经平面 $\pi: 2x - y - z + 5 = 0$ 反射，求反射光线 l_1 。

设 l 与 l_1 的方向向量分别为 \vec{s} 和 \vec{s}_1 ，设平面 π 的法向量为 \vec{n} ，单位法向量为 \vec{N} 。

解法一：设 l 于 π 交点为 P ，若能找到 l 上一点 $M(-3,1,4)$ ，即可求出点 M 关于 π 的对称点 M' ，则 $\vec{s}_1 = \overrightarrow{PM'}$ ，然后由点向式即可求得反射光线 l_1 。

解：取点 $M(-3,1,4)$ ，令 $t = \frac{x+3}{5} = \frac{y-1}{1} = \frac{z-4}{3}$ ，得 l 的参数方程

$$\begin{cases} x = 5t - 3 \\ y = t + 1 \end{cases} \quad (1)$$

$$\begin{cases} z = 3t + 4 \end{cases} \quad (2)$$

$$\begin{cases} \end{cases} \quad (3)$$

将 (1) (2) (3) 代入平面方程 $\pi: 2x - y - z + 5 = 0$ 得 $t = 1$ ，所以 P 点坐标为 $(2,2,7)$ 。

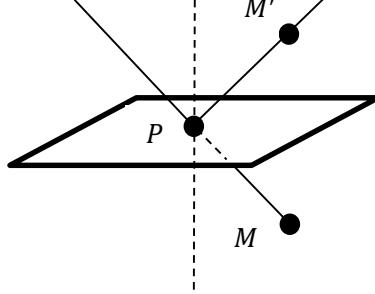


图 1 解法一示意图

过 M 点且垂直于 π 平面的直线参数方程为

$$\begin{cases} x = 2t - 3 \\ y = -t + 1 \end{cases} \quad (4)$$

$$\begin{cases} z = -t + 4 \end{cases} \quad (5)$$

$$\begin{cases} \end{cases} \quad (6)$$

将 (4) (5) (6) 代入平面方程 $\pi: 2x - y - z + 5 = 0$ 得 M 在平面上的投影点 $M_0(-1,0,3)$ ，利用中点坐标公式可得 $M'(1,-1,2)$ ，则 $\vec{s}_1 = \overrightarrow{PM'} = (-1, -3, -5)$ ，所以 $l_1: \frac{x-2}{1} = \frac{y-2}{3} = \frac{z-7}{5}$ 。

解法二：如图 2，设 \vec{n} ， \vec{s} ， \vec{s}_1 共面，且 \vec{n} 与 \vec{s} ， \vec{s}_1 夹角相等，若取 $\|\vec{s}\| = \|\vec{s}_1\|$ ，则 $(\vec{s} + \vec{s}_1) \parallel \vec{n}$ ，设 $\vec{s}_1 = (m, n, p)$ ，易知 $\vec{s} = (5, 1, 3)$ ， $\vec{n} = (2, -1, -1)$ ，有

$$\begin{cases} m^2 + n^2 + p^2 = 5^2 + 1^2 + 3^2 = 35 \\ \frac{m+5}{2} = \frac{n+1}{-1} = \frac{p+3}{-1} \end{cases},$$

解得 $\vec{s}_1 = (-1, -3, -5)$ ，通过“解法一”求出交点 $P(2,2,7)$ ，则可得

$$l_1: \frac{x-2}{1} = \frac{y-2}{3} = \frac{z-7}{5}.$$

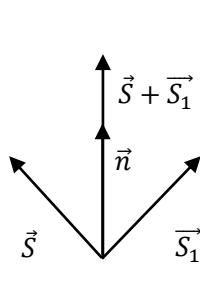


图 2 解法二示意图

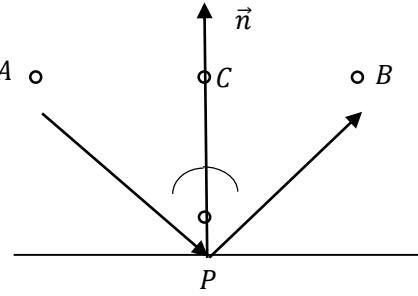


图 3 解法三示意图

解法三：如图 3，设入射光线为 \overrightarrow{AP} ，反射光线为 \overrightarrow{PB} ，且 \vec{n} 与 \overrightarrow{PA} ， \overrightarrow{PB} 夹角相等，均为 θ 。取 $\|\overrightarrow{AP}\| = \|\overrightarrow{PB}\|$ ， $\|\overrightarrow{PC}\|$ 为 \overrightarrow{AP} 在 \vec{n} 上的投影。有 $\overrightarrow{PB} = \overrightarrow{AB} - \overrightarrow{AP}$ ， $\overrightarrow{AB} = 2\overrightarrow{AC} = 2(\overrightarrow{AP} + \overrightarrow{PC})$ ，所以 $\overrightarrow{PB} = \overrightarrow{AP} + 2\overrightarrow{PC}$ （转化为求 \overrightarrow{PC} 的问题）。

由图 3 可得 $\|\overrightarrow{PC}\| = \|\overrightarrow{AP}\| \cos \theta$ ，则 $\overrightarrow{PC} = \|\overrightarrow{AP}\| \cos \theta \cdot \vec{n}$ ，又 $\overrightarrow{PA} \cdot \overrightarrow{PC} = \|\overrightarrow{AP}\| \|\overrightarrow{PC}\| \cos \theta$ 。

所以 $\overrightarrow{PC} = \frac{\overrightarrow{PA} \cdot \overrightarrow{PC}}{\|\overrightarrow{PC}\|} \cdot \vec{n} = (\overrightarrow{PA} \cdot \vec{n}) \cdot \vec{n}$ ，所以 $\overrightarrow{PB} = \overrightarrow{AP} + 2\overrightarrow{PC} = \overrightarrow{AP} - 2(\overrightarrow{AP} \cdot \vec{n}) \cdot \vec{n}$ 。

由题知 $\vec{n} = (2, -1, -1)$ ，所以

$$\vec{n} = \frac{1}{\sqrt{6}}(2, -1, -1), \quad \overrightarrow{AP} = \vec{S} = (5, 1, 3), \quad \overrightarrow{AP} \cdot \vec{n} = \sqrt{6},$$

$$\overrightarrow{PB} = \overrightarrow{AP} - 2\sqrt{6}\vec{n} = (1, 3, 5), \text{ 所以 } \overrightarrow{S}_1 = \overrightarrow{PB} = (1, 3, 5), \text{ 所以}$$

$$l_1: \frac{x-2}{1} = \frac{y-2}{3} = \frac{z-7}{5}.$$

解法四：利用镜面反射矩阵快速计算

定义 1：若向量 u 满足 $\|u\|_2 = 1$ ，即 $u^T u = 1$ 时，矩阵 $H = I - 2uu^T$ 称为镜面反射矩阵。

几何意义：

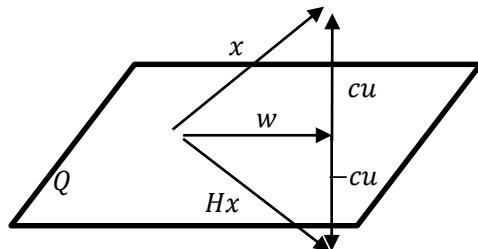


图 4 镜面反射矩阵的几何意义

如图 4， $\forall x = cu + w$ ， $w^T u = 0$ （ c 为常数），

$$\begin{aligned} Hx &= (I - 2uu^T)x = x - 2uu^T x \\ &= cu + w - 2u[u^T(w + cu)] \\ &= cu + w - 2cu = -cu + w. \end{aligned}$$

其几何意义是：将 n 维空间分成两个直交的子空间 Q 和 $v, u \in v, v$ 为一维空间， Q 是 $n-1$ 维子空间且由向量 u 唯一决定，将 Q 看作镜面，则 Hx 恰好是向量 x 关于镜面 Q 的像，故称 H 为镜面反射矩阵。由上述得： $y_{\text{反射光线}} = Hx_{\text{入射光线}}$ 。

在本题中，

$$u = N^T = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}, \quad u^T = \frac{1}{\sqrt{6}} [2 \quad -1 \quad -1],$$

则

$$H = I - 2uu^T = \frac{1}{3} \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{bmatrix},$$

$$y_{\text{反射光线}} = Hx_{\text{入射光线}} = \frac{1}{3} \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix},$$

所以 $\vec{S}_1 = y_{\text{反射光线}}^T = (1, 3, 5)$, 故 $l_1: \frac{x-2}{1} = \frac{y-2}{3} = \frac{z-7}{5}$.

参考文献

- [1] 罗芳. 镜面反射矩阵及其应用[J]. 吕梁高等专科学校学报, 2000, 02:65-66.
- [2] 黄廷祝, 成孝予. 线性代数与空间解析几何. 第五版. 北京: 高等教育出版社, 2018.

初探 Cramer 法则

郭好 (2019091604028) 邓远哲 (2019081304005)

指导教师: 李厚彪

【摘要】 本文落脚于克莱姆 (Cramer) 法则的定义及多种证明方式, 从多个角度出发, 特别是从几何的角度揭示了克莱姆法则的几何意义, 对克莱姆法则有了新的理解。

【关键词】 克莱姆法则; 证明; 几何展示

1 克莱姆法则的表述

定理 1^[1]. 设矩阵 A 是一个 $n \times n$ 的方阵, 向量 $x = (x_1, x_2, \dots, x_n)^T$, $b = (b_1, b_2, \dots, b_n)^T$, 有
$$Ax = b$$

如果 A 是可逆矩阵, 则该方程有唯一解 $x = (x_1, x_2, \dots, x_n)^T$, 其中

$$x_i = \frac{\det(A_i)}{\det(A)}, i = 1, 2, \dots, n$$

A_i 是 A 中的第 i 列被 b 取代了之后的矩阵.

2 克莱姆法则的 4 种证明

沿用上面表述中的符号和字母, 我们有对克莱姆法则的证明如下:

1) 证 (利用伴随矩阵):

由于 A 可逆, 解的唯一性是显然的, 下面证明解的表示的正确性。

$$x = (x_1, x_2, \dots, x_n)^T = A^{-1}b = \frac{1}{\det(A)}A^*b = \frac{1}{\det(A)} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \ddots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} b,$$

相应地, $x_i = \frac{1}{\det(A)}(b_1 A_{1i} + b_2 A_{2i} + \cdots + b_n A_{ni}) = \frac{\det(A_i)}{\det(A)}$.

2) 证 (利用行列式的线性性):

令 $A = (a_1, a_2, \dots, a_n)$, 则有:

$$Ax = \sum_{k=1}^n a_k x_k = b$$

那么

$$\begin{aligned} \det(A_i) &= \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots) \\ &= \det\left(a_1, \dots, a_{i-1}, \sum_{k=1}^n a_k x_k, a_{i+1}, \dots\right) \\ &= \det(a_1, \dots, a_{i-1}, a_1 x_1 + a_2 x_2 + \cdots + a_i x_i + \cdots + a_n x_n, a_{i+1}, \dots) \\ &= 0 + 0 + \cdots + \det(\dots, a_{i-1}, a_i x_i, a_{i+1}, \dots) + \cdots + 0 + 0 \\ &= x_i \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots) \\ &= x_i \det(A), \end{aligned}$$

所以有: $x_i = \frac{\det(A_i)}{\det(A)}, i = 1, 2, \dots, n$.

3) 证 (构造辅助矩阵):

$$\text{构造矩阵 } X_1 = \begin{pmatrix} x_1 & 0 & 0 & \cdots & 0 \\ x_2 & 1 & 0 & \cdots & 0 \\ x_3 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & 0 & \cdots & 1 \end{pmatrix} = (x, e_2, e_3, \dots, e_n), \text{ 有 } \det(X_1) = x_1.$$

令 $A = (a_1, a_2, \dots, a_n)$, $Ae_i = a_i$, 则 $e_i = A^{-1}a_i$, 所以
 $X_1 = (A^{-1}b, A^{-1}a_2, A^{-1}a_3, \dots, A^{-1}a_n) = A^{-1}A_1$

两边取行列式得:

$$x_1 = \det(X_1) = \det(A^{-1})\det(A_1) = \frac{\det(A_1)}{\det(A)}$$

同理其他 x_i 也可以这样得到。

4) 证 (二维几何展示):

对于一个二元一次方程组 $\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$, 我们用(*)式
 $x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ (*)

来表示。

首先给出: 如果二维平面直角坐标系中的一个以原点为顶点的平行四边形的两条边, 分别是向量 $a_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$, $a_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$, (如下图 1)

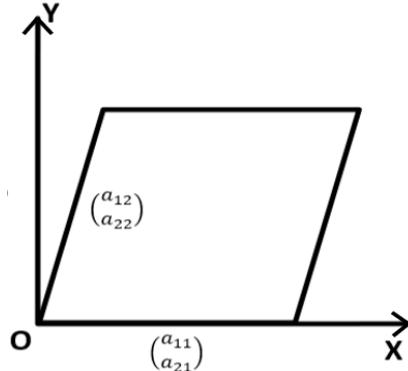


图 1 系数向量

那么其面积

$$S = a_{11} \times a_{22} - a_{12} \times a_{21} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

该图形坐标经过(*)式中的变换后的结果如下图:

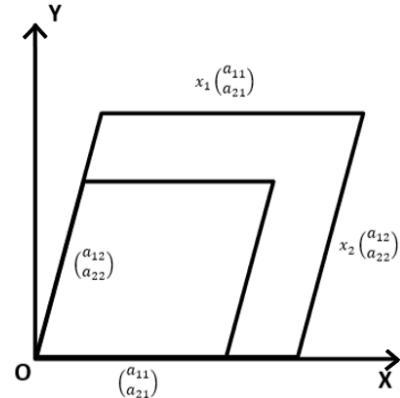


图 2 系数向量的伸缩变换

现将原图形由 a_2 表示的边扩大 x_2 倍,

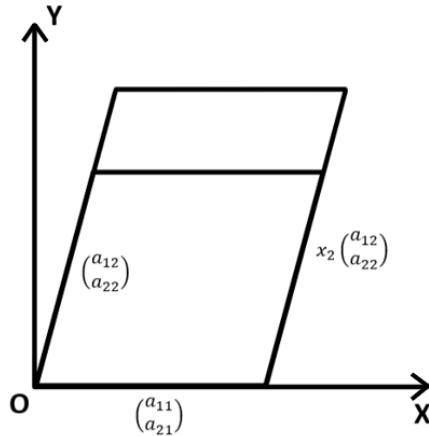


图 3 a_2 的伸缩变换

那么现在面积变为

$$S_1 = x_2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

将这个平行四边形的顶边进行平移，是左上方的顶点落于 (b_1, b_2) 处，

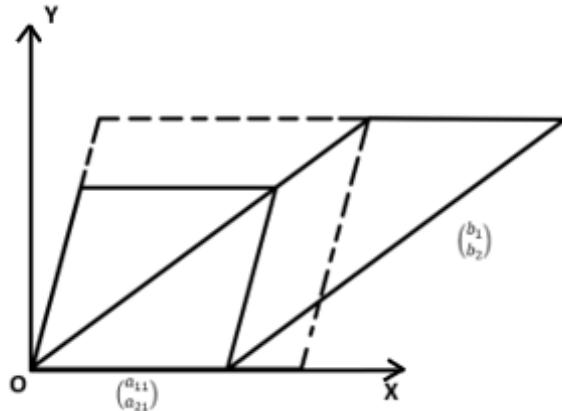


图 4 平行四边形的顶边平移变换

则箭头所指的边可以用向量 $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ 表示，由于底和高都不变，该平行四边形的面积

$$\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = S'_1 = S_1 = x_2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

即

$$x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{11} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{\det(A_2)}{\det(A)}$$

同理可以证明

$$x_1 = \frac{\det(A_1)}{\det(A)}.$$

实际上，在更高维的空间中也可以类似的对克莱姆法则进行展示，本文不再赘述。

参考文献

- [1] 黄廷祝, 成孝予.线性代数与空间解析几何.第五版. 北京: 高等教育出版社, 2018.
- [2] 维基百科编者. 克莱姆法則[G/OL]. 维基百科, 2019(20191205)[2019-12-05].
- [3] Wikipedia contributors. (2019, October 14). Cramer's rule. In Wikipedia, The Free Encyclopedia. Retrieved 02:44, December 22, 2019.

第一类曲面积分的两种简单计算方法

肖健雄

指导教师：刘碧森

【摘要】 柱坐标系和球坐标系在微积分中有着广泛的应用。本文从第一类曲面积分的计算公式出发，根据第一类曲面积分的“对面积的曲面积分”这一概念，将第一类曲面积分在柱坐标和球坐标下转化，从而简化计算。

【关键词】 第一类曲面积分；柱坐标；球坐标；坐标系变换

1 柱坐标下的第一类曲面积分

在直角坐标系下，第一类曲面积分可以转化为二重积分，计算公式为

$$\iint_S f(x, y, z) dS = \iint_D f(x, y, z) \sqrt{1 + z_x^2 + z_y^2} dx dy$$

其中，曲面 S 的方程为 $z = z(x, y)$ ，曲面的面积微元

$$dS = \sqrt{1 + z_x^2 + z_y^2} dx dy$$

由直角坐标系和柱坐标系的坐标关系

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \\ z = z \end{cases}$$

可以将函数 z 看作复合函数， z 为中间变量 x, y 的函数， x, y 为自变量 r, θ 的函数。根据复合函数求导的链式法则，有

$$\begin{cases} \frac{\partial z}{\partial r} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial r} = z_x \cos \theta + z_y \sin \theta \\ \frac{\partial z}{\partial \theta} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial \theta} = -z_x r \sin \theta + z_y r \cos \theta \end{cases}$$

由 Cramer 法则，有

$$\begin{cases} z_x = z_r \cos \theta - \frac{1}{r} z_\theta \sin \theta \\ z_y = z_r \sin \theta + \frac{1}{r} z_\theta \cos \theta \end{cases}$$

另一方面，由二重积分的换元法，有

$$d\sigma = dx dy = \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| dr d\theta = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} dr d\theta = r dr d\theta$$

因此，曲面 S 在柱坐标下的面积微元

$$dS = \sqrt{1 + z_r^2 + \frac{1}{r^2} z_\theta^2} \cdot r dr d\theta$$

柱坐标系下的第一类曲面积分的计算公式为

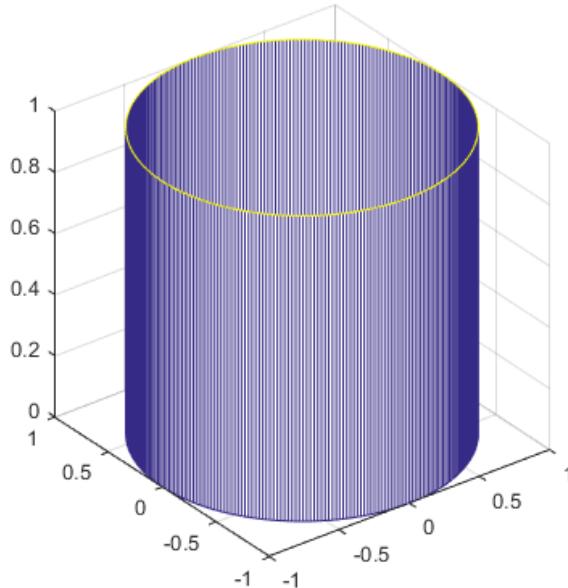
$$\iint_S f(x, y) dS = \iint_D f(r \cos \theta, r \sin \theta) \sqrt{1 + z_r^2 + \frac{1}{r^2} z_\theta^2} \cdot r dr d\theta$$

若曲面 S 的方程可表示为 $z = z(r)$, 即曲面方程与 θ 无关, 则计算公式为

$$\iint_S f(x, y) dS = \iint_D f(r \cos \theta, r \sin \theta) \sqrt{1 + z_r^2} \cdot r dr d\theta$$

曲面 S 为柱面的情形下, z 不是 r, θ 的单值函数, 故以上公式不再成立。但考虑到曲面为柱面, 则可以利用微元法解决这一种情形的第一类曲面积分。设柱面

$$S = \{(r, \theta, z) | r = R, 0 \leq \theta \leq 2\pi, z_1 \leq z \leq z_2\}$$



则柱面的面积微元为

$$dS = R d\theta dz$$

柱面上的第一类曲面积分为

$$\iint_S f(R, \theta, z) dS = \iint_S f(R, \theta, z) R d\theta dz$$

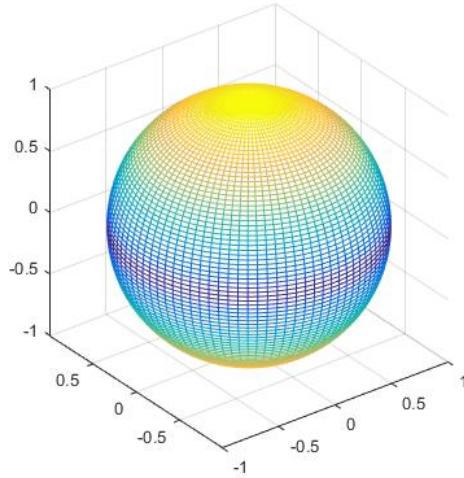
2 球坐标下的第一类曲面积分

在球坐标系下, 由直角坐标系和球坐标系的坐标关系

$$\begin{cases} x = \rho \sin \varphi \cos \theta \\ y = \rho \sin \varphi \sin \theta \\ z = \rho \cos \varphi \end{cases}$$

仅考虑球面上的第一类曲面积分, 则可以利用微元法解决。设球面

$$S = \{(\rho, \varphi, \theta) | \rho = R, 0 \leq \varphi \leq 2\pi, 0 \leq \theta \leq 2\pi\}$$



则球面的面积微元为

$$dS = R^2 \sin \varphi d\varphi d\theta$$

球面上的第一类曲面积分

$$\iint_S f(R, \varphi, \theta) dS = \iint_S f(R, \varphi, \theta) R^2 \sin \varphi d\varphi d\theta$$

3 应用实例

(1) 计算 $\iint_S (x^2 + y^2) dS$, 其中 S 是曲面 $z = \sqrt{x^2 + y^2}$ 包含在柱面 $x^2 + y^2 = R^2$ 内的部分。

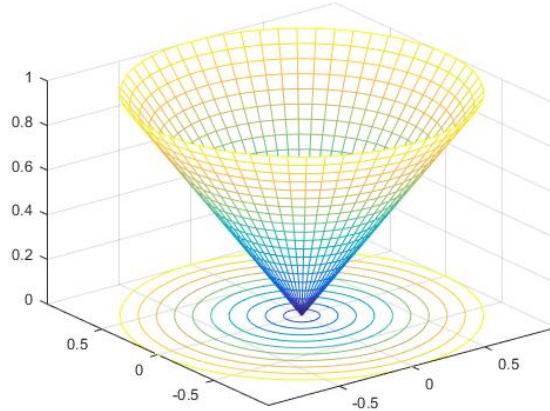
解: 由直角坐标系和柱坐标系的坐标关系, 有 $r^2 = x^2 + y^2$,

曲面 S 的方程为 $z = r$ ($0 \leq r \leq R, 0 \leq \theta \leq 2\pi$), 面积微元

$$dS = \sqrt{1 + z_r^2} r dr d\theta = \sqrt{2} r dr d\theta$$

曲面积分:

$$\iint_S (x^2 + y^2) dS = \iint_S \sqrt{2} r^2 dr d\theta = \sqrt{2} \int_0^{2\pi} d\theta \int_0^R r^2 dr = \frac{2\sqrt{2}}{3} R^3$$



(2) 计算 $\iint_S \frac{1}{x^2 + y^2 + z^2} dS$, 其中 S 是柱面 $x^2 + y^2 = R^2$ 在第 I 卦限内位于平面 $z = 0$ 与

$z = H$ ($H > 0$) 之间的部分。

解：由直角坐标系和柱坐标系的坐标关系，有 $r^2 = x^2 + y^2$ ，曲面 S 的方程为

$$r = R(0 \leq \theta \leq \frac{\pi}{2}), \text{ 面积微元}$$

$$dS = R d\theta dz$$

曲面积分：

$$\begin{aligned} & \iint_S \frac{1}{x^2 + y^2 + z^2} dS \\ &= \iint_S \frac{1}{R^2 + z^2} R d\theta dz = R \int_0^{\frac{\pi}{2}} d\theta \int_0^H \frac{1}{R^2 + z^2} dz \\ &= \frac{\pi R}{2} \int_0^H \frac{1}{R^2 + z^2} dz = \frac{\pi}{2R} \arctan\left(\frac{z}{R}\right) \Big|_0^H \\ &= \frac{\pi}{2R} \arctan H \end{aligned}$$

(3) 计算 $\iint_S (x^2 + y^2 + z^2) dS$ ，其中 S 是球面 $x^2 + y^2 + z^2 = R^2$ 在第 I 半限内位于平面 $z = 0$ 与 $z = H$ ($H > 0$) 之间的部分。

解：由直角坐标系和球坐标系的坐标关系，有 $\rho^2 = x^2 + y^2 + z^2$ ，

曲面 S 的方程为 $\rho = R(0 \leq \cos \varphi \leq \frac{H}{R}, 0 \leq \theta \leq \frac{\pi}{2})$ ，面积微元

$$dS = R^2 \sin \varphi d\varphi d\theta$$

① $H \geq R$ 时，曲面 S 的方程为 $\rho = R(0 \leq \varphi \leq \frac{\pi}{2}, 0 \leq \theta \leq \frac{\pi}{2})$ ，

曲面积分：

$$\begin{aligned} & \iint_S (x^2 + y^2 + z^2) dS \\ &= \iint_S R^4 \sin \varphi d\varphi d\theta \\ &= R^4 \int_0^{\frac{\pi}{2}} d\theta \int_0^{\frac{\pi}{2}} \sin \varphi d\varphi = \frac{\pi R^4}{2} \end{aligned}$$

② $H < R$ 时，曲面 S 的方程为 $\rho = R(\arccos \frac{H}{R} \leq \varphi \leq \frac{\pi}{2}, 0 \leq \theta \leq \frac{\pi}{2})$ ，

曲面积分：

$$\begin{aligned} & \iint_S (x^2 + y^2 + z^2) dS \\ &= \iint_S R^4 \sin \varphi d\varphi d\theta \\ &= R^4 \int_0^{\frac{\pi}{2}} d\theta \int_{\arccos \frac{H}{R}}^{\frac{\pi}{2}} \sin \varphi d\varphi \\ &= \frac{\pi R^4}{2} \int_{\arccos \frac{H}{R}}^{\frac{\pi}{2}} \sin \varphi d\varphi = \frac{\pi R^3 H}{2} \end{aligned}$$

因此，曲面积分

$$\iint_S (x^2 + y^2 + z^2) dS = \frac{\pi R^3}{2} \min\{H, R\}$$

4 结论

第一类曲面积分是对面积的曲面积分，因此计算的基本方法是被积函数对面积微元积分。对于不同类型的曲面，选定恰当的坐标，就可以达到简化计算的目的。

参考文献

- [1] 傅英定 谢云荪 电子科技大学应用数学学院 微积分（第二版）上册 [M] 北京：高等教育出版社，2009.6
- [2] 王强 蒋青 同济大学数学系 高等数学（第七版）上册 [M] 北京：高等教育出版社，2014.7

积分因子法在求解微分方程中的应用

肖健雄

指导教师：刘碧森

【摘要】 积分因子法是一类求解常微分方程的重要方法。本文从全微分方程出发，给出了积分因子法的一般性求解思路，并给出了特定形式的积分因子。

【关键词】 积分因子；全微分方程；偏微分方程

1 全微分方程和积分因子法

对于一个如下形式的一阶常微分方程

$$P(x, y)dx + Q(x, y)dy = 0$$

若存在某个函数 $u(x, y)$ ，使得

$$du = Pdx + Qdy$$

则此方程为全微分方程。方程的通解为 $u(x, y) = C$ 。

当 P, Q 在某单连通区域内具有一阶连续偏导数时，方程为全微分方程的充分必要条件是在此连通区域内恒有

$$\begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ P & Q \end{vmatrix} = 0$$

当方程 $Pdx + Qdy = 0$ 不是全微分方程，若存在连续可微函数 $\mu(x, y) \neq 0$ ，使得方程

$$\mu Pdx + \mu Qdy = 0$$

为全微分方程，则函数 $\mu(x, y)$ 为方程的一个积分因子。

2 积分因子的一般求解方法

积分因子的一般性求解方法是求解微分方程。

对于方程 $Pdx + Qdy = 0$ ，记

$$\begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ P & Q \end{vmatrix} = D$$

方程 $\mu Pdx + \mu Qdy = 0$ 为全微分方程，则有

$$\begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ \mu P & \mu Q \end{vmatrix} = 0$$

上式就是一个微分方程。展开并整理，有

$$P \frac{\partial \mu}{\partial y} - Q \frac{\partial \mu}{\partial x} = \mu \cdot D$$

记为 (*) 式。或整理成下式，记为 (**).

$$P \frac{\partial \ln \mu}{\partial y} - Q \frac{\partial \ln \mu}{\partial x} = D$$

这是一个一阶线性偏微分方程。

一阶偏微分方程的求解较为复杂，但在特定的情形下，可以较为简单地求解之。

3 几种特定形式的积分因子

以下考虑几种特定的形式。在以下几种形式下，积分因子的求解还是较为简单的。

a) 积分因子是 x 的一元函数，即 $\mu = \mu(x)$ ：此时，(*) 式可化简为

$$-Q \frac{d\mu}{dx} = \mu \cdot D$$

令 $-\frac{D}{Q} = \Phi$ ，则当且仅当 Φ 是 x 的一元可积函数时，即 $\Phi = \varphi(x)$ ，积分因子存在，通解为

$$\mu = C e^{\int \varphi(x) dx}$$

b) 积分因子是 y 的一元函数，即 $\mu = \mu(y)$ ：此时，(*) 式可化简为

$$P \frac{d\mu}{dy} = \mu \cdot D.$$

令 $\frac{D}{P} = \Psi$ ，则当且仅当 Ψ 是 y 的一元可积函数时，即 $\Psi = \psi(y)$ ，积分因子存在，通解为

$$\mu = C e^{\int \psi(y) dy}$$

c) 积分因子是 $\xi = \varphi(x, y)$ 的一元函数，即 $\mu = f(\xi) = f[\varphi(x, y)]$ ：根据链式法则，

$$\frac{\partial \mu}{\partial x} = f' \cdot \frac{\partial \varphi}{\partial x}, \frac{\partial \mu}{\partial y} = f' \cdot \frac{\partial \varphi}{\partial y}$$

此时，(*) 式可化简为

$$\left(P \frac{\partial \varphi}{\partial y} - Q \frac{\partial \varphi}{\partial x} \right) \cdot \frac{d\mu}{d\xi} = \mu \cdot D$$

记

$$\frac{D}{\left(P \frac{\partial \varphi}{\partial y} - Q \frac{\partial \varphi}{\partial x} \right)} = \Omega$$

则当且仅当 Ω 是 $\xi = \varphi(x, y)$ 的一元可积函数时，即 $\Omega = \omega(\xi)$ ，积分因子存在，通解为

$$\mu = C e^{\int \omega(\xi) d\xi}$$

4 应用实例

(1) 求解 $ydx - xdy = 0$ 。

解：此方程不是全微分方程，假设方程有积分因子 $\mu = \mu(y)$ ，则函数

$$\psi(y) = \frac{D}{P} = -\frac{2}{y}$$

方程的一个积分因子

$$\mu(y) = e^{\int_1^y \psi(y) dy} = e^{\int_1^y -\frac{2}{y} dy} = \frac{1}{y^2}$$

因此，与原方程同解的全微分方程

$$\frac{ydx - xdy}{y^2} = 0$$

全微分的一个原函数

$$u(x, y) = \int_{(0,1)}^{(x,y)} \frac{ydx - xdy}{y^2} = \int_1^y 0 dy + \int_0^x \frac{dx}{y} = \frac{x}{y}$$

因此，方程的通解是

$$\frac{x}{y} = C$$

$$(2) \text{ 求解 } \frac{dy}{dx} = \frac{x^2 + y^2}{xy}.$$

解：原方程可以写成

$$(x^2 + y^2)dx - xydy = 0$$

此方程不是全微分方程，假设方程有积分因子 $\mu = \mu(x)$ ，则函数

$$\varphi(x) = -\frac{D}{Q} = -\frac{3}{x}$$

方程的一个积分因子

$$\mu(x) = e^{\int \varphi(x) dx} = e^{\int \frac{-3}{x} dx} = \frac{1}{x^3}$$

因此，与原方程同解的全微分方程

$$\frac{x^2 + y^2}{x^3} dx - \frac{y}{x^2} dy = 0$$

全微分的一个原函数

$$u(x, y) = \int_{(1,0)}^{(x,y)} \frac{x^2 + y^2}{x^3} dx - \frac{y}{x^2} dy = \int_1^x \frac{dx}{x} + \int_0^y -\frac{ydy}{x^2} = \ln x - \frac{y^2}{2x^2}$$

因此，方程的通解是

$$y^2 = x^2 \ln x^2 + Cx^2$$

$$(3) \text{ 求解一阶线性微分方程 } \frac{dy}{dx} + p(x)y = f(x).$$

解：原方程可以写成

$$[p(x)y - f(x)]dx + dy = 0$$

此方程不是全微分方程，假设方程有积分因子 $\mu = \mu(x)$ ，则函数

$$\varphi(x) = -\frac{D}{Q} = p(x)$$

方程的一个积分因子

$$\mu(x) = e^{\int \varphi(x) dx} = e^{\int p(x) dx}$$

因此，与原方程同解的全微分方程

$$[p(x)y - f(x)]e^{\int p(x) dx} dx + e^{\int p(x) dx} dy = 0$$

全微分的一个原函数

$$u(x, y) = \int [p(x)y - f(x)]e^{\int p(x) dx} dx + e^{\int p(x) dx} dy = ye^{\int p(x) dx} - \int f(x)e^{\int p(x) dx} dx$$

因此，方程的通解是

$$y = e^{-\int p(x)dx} \left[\int f(x)e^{\int p(x)dx} dx + C \right]$$

5 结束语

积分因子法是一类求解一阶常微分方程的重要的方法。对于给定形式的积分因子，可以通过简单的计算得到；而积分因子的一般性求解则是偏微分方程中需要研究的课题。

参考文献

- [1] 傅英定 谢云苏 电子科技大学应用数学学院 微积分（第二版）上册 [M] 北京：高等教育出版社，2009.6.

留数定理在 Laplace 反变换上的计算

刘思雨

指导老师：刘碧森

【摘要】 运用复变函数与积分变换中的留数及其应用对求解信号系统中 Laplace 反变换的计算具有很大的帮助，能够很好的解决信号系统中有关 Laplace 变换的问题，

【关键词】 留数定理；Laplace 反变换；方法；简化

序言：

由于信号的时域表示和 S 域表示是一一对应的，当由信号的的 Laplace 变换 $X(s)$ 求解信号的时域表示 $x(t)$ ，即为 Laplace 反变换，在信号系统中计算 Laplace 反变换的方法主要是留数法和部分分式展开法，前者根据 Laplace 反变换的定义入手，利用复变函数中的留数定理得到时域信号，后者是将 S 域表示式分解成许多简单的表示式之和，然后分别得到原时域信号。

(1) 留数的定义：设 Z_0 为 $f(z)$ 的孤立奇点，那么 $f(z)$ 在 Z_0 的留数 $\text{Res}[f(z), Z_0] = C - 1 = \frac{1}{2\pi i} \oint_C f(z) dz$ ，其中 C 为去心邻域 $0 < |Z - Z_0| < R$ 内的任意一条正向简单闭曲线。如果 $Z = \infty$ 为

$f(z)$ 的孤立奇点，那么 $f(z)$ 在 $Z = \infty$ 的留数 $\text{Res}[f(z), \infty] = \frac{1}{2\pi i} \oint_{C_\infty} f(z) dz$ ，其中 C 为 R 内绕原点的

任意一条正向简单闭曲线。

(2) 留数定理：设函数 $f(z)$ 在区域 D 内除有限个孤立奇点 z_1, z_2, \dots, z_n 外处处解析， C 为 D 内包围诸奇点的一条正向简单闭曲线，那么 $\oint_C f(z) dz = 2\pi i \sum_{k=1}^n \text{Res}[f(z), z_k]$ ，这个定理把求沿封

闭曲线 C 的积分，转化为求被积函数在 C 中的各孤立奇点处的留数。

(3) 根据留数的定义及留数定理对 Laplace 反变换的计算可以直接从其定义，即 $x(t) \xleftarrow{L} X(s) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} X(s) e^{st} ds$ ，上式为一复变积分，积分路径是 s 平面上平行于虚轴的直线 $\sigma = C > \sigma_0$ 。为了应用留数定理，必须补上一个半径充分大的圆弧，使圆弧与直线构成闭合围线，用围线积分来代替线积分。由 Jordan (约当) 引理，若满足条件 $\lim_{|s|=R \rightarrow \infty} |X(s)| = 0$ ，则

$\lim_{R \rightarrow \infty} \int_{C_{R_1}} X(s) e^{st} ds = 0, t > 0, \lim_{R \rightarrow \infty} \int_{C_{R_2}} X(s) e^{st} ds = 0, t < 0$ ，因此 Laplace 反变换积分等于围线积分

乘以 $\frac{1}{2\pi j}$ ，即

$$x(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} X(s) e^{st} ds = \frac{1}{2\pi j} [\int_{c-j\infty}^{c+j\infty} X(s) e^{st} ds + \int_{C_{R_1}} X(s) e^{st} ds], t > 0$$

或

$$\frac{1}{2\pi j} [\int_{c-j\infty}^{c+j\infty} X(s) e^{st} ds + \int_{C_{R_2}} X(s) e^{st} ds], t < 0$$

由留数定理，复平面上任意闭合围线积分等于围线内被积函数所有极点的留数之和。举例如下：

例 (1): 已知信号 $x(t)$ 的 Laplace 变换为 $X(s) = \frac{s+2}{s(s+3)(s+1)^2}$, $\text{Re}\{s\} > 0$, 试用留数法求 $x(t)$.

解: $X(s)$ 具有两个单极点 $P_1 = 0, P_2 = -3$ 和一个二阶极点 $P_3 = -1$. 则分别求出相应极点的留数为

$$\begin{aligned}\text{Res}_{s=0}\{X(s)e^{st}\} &= s \frac{s+2}{s(s+3)(s+1)^2} e^{st} \Big|_{s=0} = \frac{2}{3} \\ \text{Res}_{s=-3}\{X(s)e^{st}\} &= (s+3) \frac{s+2}{s(s+3)(s+1)^2} e^{st} \Big|_{s=-3} = \frac{1}{12} e^{-3t} \\ \text{Res}_{s=-1}\{X(s)e^{st}\} &= \frac{1}{(2-1)!} \frac{d}{ds} (s+1)^2 X(s)e^{st} \Big|_{s=-1} = \frac{d}{ds} \frac{s+2}{s(s+3)} e^{st} \Big|_{s=-1} = -\frac{1}{2} te^{-t} - \frac{3}{4} e^{-t}\end{aligned}$$

所以: $x(t) = \left(\frac{2}{3} + \frac{1}{12} e^{-3t} - \frac{1}{2} te^{-t} - \frac{3}{4} e^{-t} \right) u(t)$. 即得解

例 (2): 利用留数法对信号 $X(s) = \frac{se^{-2s}}{(s+1)(s+2)^2}$, $\text{Re}\{s\} > -1$, 进行 Laplace 反变换, 求 $x(t)$.

解: 由 Jordan 引理, $0 \leq t < 2$ 范围内, Laplace 反变换可表示为

$$x(t) = \frac{1}{2\pi j} \left[\int_{c-j\infty}^{c+j\infty} \frac{se^{s(t-2)}}{(s+1)(s+2)^2} ds + \int_{c_R}^{\infty} \frac{se^{s(t-2)}}{(s+1)(s+2)^2} ds \right],$$

由于围线内无极点, 所以 $x(t) = 0$. 当 $t \geq 2$ 时, Laplace 反变换可表示为

$$x(t) = \frac{1}{2\pi j} \left[\int_{c-j\infty}^{c+j\infty} \frac{se^{s(t-2)}}{(s+1)(s+2)^2} ds + \int_{c_R}^{\infty} \frac{se^{s(t-2)}}{(s+1)(s+2)^2} ds \right] = \sum_{k=1}^2 \text{Res}_{s=p_k} [X(s)e^{st}],$$

由于 $X(s)$ 有一个单极点 $P_1 = -1$ 和一个二重极点 $P_2 = -2$, 其相应的留数为

$$\begin{aligned}\text{Res}_{s=-1}[X(s)e^{st}] &= \frac{se^{-2s}}{(s+2)^2} e^{st} \Big|_{s=-1} = -e^{-(t-2)} \\ \text{Res}_{s=-2}[X(s)e^{st}] &= \frac{d}{ds} [(s+1)^2 X(s)e^{st}] \Big|_{s=-2} = \frac{d}{ds} \left[\frac{s+2}{s(s+3)} e^{st} \right] \Big|_{s=-1} = -3e^{-2(t-2)} + 2te^{-2(t-2)}\end{aligned}$$

所以: $x(t) = [-e^{-(t-2)} - 3e^{-2(t-2)} + 2te^{-2(t-2)}] u(t-2)$. 即得解。

例 3: 已知信号 $x(t)$ 的 Laplace 变换为 $X(s) = \frac{s+2}{s^2+4s+3}$, $\text{Re}\{s\} > -1$, 利用部分分式法求 Laplace 反变换。

解: $X(s)$ 为有理真分式, 极点均为一阶, 因此有

$$\begin{aligned}X(s) &= \frac{s+2}{s^2+4s+3} = \frac{s+2}{(s+1)(s+3)} = \frac{k_1}{s+1} + \frac{k_2}{s+3} \\ k_1 = (s+1)X(s) \Big|_{s=-1} &= \frac{s+2}{s+3} \Big|_{s=-1} = \frac{1}{2}, \quad k_2 = (s+3)X(s) \Big|_{s=-3} = \frac{s+2}{s+1} \Big|_{s=-3} = \frac{1}{2}\end{aligned}$$

故 Laplace 反变换为

$$x(t) = \frac{1}{2} e^{-t} u(t) + \frac{1}{2} e^{-3t} u(t)$$

分析比较: 从以上典型例题可以看出: 运用复变函数与积分变换中的留数及其应用对求解信号系统中 Laplace 反变换的计算具有很大的帮助, 能够很好的解决信号系统中有关 Laplace 变换的问题, 与部分分式法相比, 虽然比较复杂, 但留数法适用的范围却比较广, 能够更好的辅助信号系统的学习, 对信号系统有很大的促进作用。这也是这篇论文的目的所在。

参考文献

- [1] 信号与系统（第二版）奥本海姆；
- [2] 复变函数（第四版）西安交通大学高等数学教研室编； 余家荣

绕任意直线旋转的旋转体体积公式的探究

张晋豪 (2016220204026) 何蔚阳 (2016220205031)

指导教师: 刘碧森

【摘要】 教材上已经给出了函数关于坐标轴或者平行于坐标轴的直线旋转的体积公式, 作为拓展, 我将在回顾以往知识的基础上, 借助线性代数中矩阵的工具, 进行函数绕任意直线旋转的旋转体体积公式推导。

【关键词】 直线; 旋转体体积; 矩阵; 定积分; 函数图像

1 函数 $f(x)$ 绕 x 轴旋转的旋转体体积

在 x 轴上取微元 $[x, x+dx]$, 将每一段绕 x 轴上的旋转体积近似为圆柱, 由圆柱体积公式, 有 $dV = \pi f^2(x)dx$, 积分, 计算从 $x=a$ 到 $x=b$ 的体积, 则有:

$$V = \pi \int_a^b f^2(x)dx \quad (1)$$

2 函数 $f(x)$ 绕 y 轴旋转的旋转体体积

在 x 轴上取微元 $[x, x+dx]$, 将每一段绕 y 轴上的旋转体关于 y 轴展开近似看做以圆柱柱底的周长 $2\pi x$ 为长, dx 为宽, $f(x)$ 为高的长方体, 由长方体体积公式, 有 $dV = 2\pi x f(x)dx$, 积分, 计算从 $x=a$ 到 $x=b$ 的体积, 则有:

$$V = 2\pi \int_a^b x f(x)dx \quad (2)$$

3 函数 $f(x)$ 绕 $y = y_0$ 旋转的旋转体体积

通过坐标变换, 将 $y = y_0$ 视为 x 轴, 则 $y = f(x)$ 对应变换为 $y = f(x) - y_0$, 类似于第一种情况, 计算从 $x=a$ 到 $x=b$ 的体积, 则有:

$$V = \pi \int_a^b (f(x) - y_0)^2 dx \quad (3)$$

4 函数 $f(x)$ 绕 $x = x_0$ 旋转的旋转体体积

通过坐标变换, 将 $x = x_0$ 视为 y 轴, 则 $y = f(x)$ 对应变换为 $y = f(x - x_0)$, 类似于第二种情况, 计算从 $x=a$ 到 $x=b$ 的体积, 则有:

$$V = 2\pi \int_a^b (x - x_0) f(x) dx \quad (4)$$

5 函数 $f(x)$ 绕 $y = kx$ 旋转的旋转体体积

设 $k = \tan \theta$, 由坐标变换矩阵, 函数绕原点顺时针 (原公式为逆时针, 故原公式中 θ 相应变换为 $-\theta$) 旋转 θ 度角, 新坐标轴为:

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta \\ y' &= -x \sin \theta + y \cos \theta \end{aligned}$$

即:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (5)$$

∴ 得

$$x' = x \cos \theta + y \sin \theta = x \cos \theta + f(x) \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta = -x \sin \theta + f(x) \cos \theta$$

$x = a, x = b$ 对应类似于 x 对应变换, 由公式 (1), 计算从 $x = a$ 到 $x = b$ 的体积, 则有:

$$V = \pi \int_{a \cos \theta + f(a) \sin \theta}^{b \cos \theta + f(b) \sin \theta} f^2(x') dx' = \pi \int_a^b f^2(x \cos \theta + f(x) \sin \theta) d(x \cos \theta + f(x) \sin \theta) \quad (6)$$

6 函数 $f(x)$ 绕 $y = kx + b$ 旋转的旋转体体积 (为区分 $y = kx + b$ 中的 b , 故从 $x = A$ 到 $x = B$)

法一 (类似 5, 两步坐标变换)

$$\text{第一步: } \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ -b \end{pmatrix} = \begin{pmatrix} x \\ y - b \end{pmatrix}$$

$$\text{第二步: } \begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

∴ 得

$$x'' = x \cos \theta + y' \sin \theta = x \cos \theta + (y - b) \sin \theta = x \cos \theta + (f(x) - b) \sin \theta$$

$$y'' = -x \sin \theta + y' \cos \theta = -x \sin \theta + (y - b) \cos \theta = -x \sin \theta + (f(x) - b) \cos \theta$$

由公式 (1), 计算从 $x = A$ 到 $x = B$ 的体积, 则有:

$$V = \pi \int_{A \cos \theta + f(A) \sin \theta}^{B \cos \theta + f(B) \sin \theta} f^2(x'') dx'' = \pi \int_A^B f^2(x \cos \theta + (f(x) - b) \sin \theta) d(x \cos \theta + (f(x) - b) \sin \theta)$$

法二 (类似 3, 一步坐标变换)

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$y = kx + b$ 中, $b' = b \cos \theta$, 由公式 (3), 视已变换的 x 轴, y 轴仅为旋转变换, 但中心即坐标原点并未变化, 则新坐标系中的 $y = kx + b$ 即变换为 $y' = b' = b \cos \theta$, 即 $y = kx$ 已经变换为 x' 轴。

∴ 由公式 (3), 计算从 $x = A$ 到 $x = B$ 的体积, 则有:

$$V = \pi \int_{A \cos \theta + f(A) \sin \theta}^{B \cos \theta + f(B) \sin \theta} (f(x) - b')^2 dx' = \pi \int_A^B (f(x \cos \theta + y \sin \theta) - b')^2 dx (b' = b \cos \theta)$$

7 二维直角坐标系坐标变换公式证明:

设在二维直角坐标系中有点 (x, y) , 其与坐标原点距离为 r , 它与 x 轴的夹角定义为从 x 轴绕逆时针旋转至原点到该点位矢所转过的角度 α , 则

$$x = r \cos \alpha$$

$$y = r \sin \alpha$$

现在将该点绕原点逆时针旋转 θ 度角, 新的该点坐标应为:

$$x' = r \cos(\theta + \alpha) = r \cos \theta \cos \alpha - r \sin \theta \sin \alpha = x \cos \theta - y \sin \theta$$

$$y = r \sin(\theta + \alpha) = r \sin \theta \cos \alpha + r \cos \theta \sin \alpha = x \sin \theta + y \cos \theta$$
$$\text{即} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

总结

通过线性代数中的坐标变换矩阵很好的将旋转体体积推广到绕二维平面中的任意直线旋转的旋转体体积。这也充分说明数学的不同分支之间有着千丝万缕的联系，不同分支之间的知识点可以相互帮助更好更快也更简洁地解决相关实际问题。

参考文献

- [1] 傅英定 谢云荪 电子科技大学应用数学学院 微积分（第二册）上册[M] 北京：高等教育出版社，2009.6
- [2] 黄廷祝 成孝予 电子科技大学数学科学学院 线性代数与空间解析几何(第四版)[M] 北京：高等教育出版社，2015.9

最速下降算法的简单研究及其应用

肖健雄 陈思佳 (2018 级)

指导教师: 刘碧森

【摘要】 最速下降算法是一个求解区域最小值的算法。本文从二元函数的方向导数和梯度出发，对最速下降算法的基本原理进行了简单的研究，并结合一些常见的曲面，对最速下降算法进行了一些简单的应用。

【关键词】 最速下降算法；方向导数；梯度

1 方向导数、梯度和极小值问题

从平面任一点出发，有无限多个方向。为研究二元函数 $z = f(x, y)$ 沿某个方向 l 的变化率，引入方向导数：

$$\frac{\partial f}{\partial l} = \lim_{\rho \rightarrow 0} \frac{f(x + \Delta x, y + \Delta y) - f(x, y)}{\rho}$$

记方向 l 的单位方向向量为 $\mathbf{l} = (\cos \alpha, \cos \beta)$ ，引入 $z = f(x, y)$ 的梯度：

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j}$$

则 $\frac{\partial f}{\partial l} = \nabla f \cdot \mathbf{l}$ 。因此，当方向 l 为负梯度方向时，方向导数取得最小值，即沿负梯度方向，函

数值的下降速率最大。为求解无条件极值中的极小值问题，可以沿负梯度方向进行搜索，找到极值点。

函数 $z = f(x, y)$ 在极值点的梯度 $\nabla f = 0$ ，这往往可以作为极值点的判断条件（事实上，这只是驻点的定义，驻点是否成为极值点还应该通过 Hesse 矩阵的行列式判断）。

2 最速下降算法的原理

对于二元函数 $f(X)$ ，可从一给定点 X_0 出发，沿负梯度方向搜索，步长为 λ_0 ，则搜索后得到第一个近似点 X_1 。若 $\|\nabla f\|$ 足够小，则可近似将得到的近似点看作所求的极值点。若给定精度 $\varepsilon > 0$ ，则可以通过判断 $\|\nabla f\|^2 \leq \varepsilon$ 是否成立，以确定近似点是否可以看作所求的极值点。若 $\|\nabla f\|^2 \leq \varepsilon$ ，则搜索结束，返回当前近似点；否则，从当前近似点出发，搜索下一个近似点。

基于以上分析，可设计算法如下：

- ① 给定初始点 X_0 及精度 $\varepsilon > 0$ ；
- ② 若 $\|\nabla f(X_k)\|^2 \leq \varepsilon$ ，则 X_k 即为所求近似点；若 $\|\nabla f(X_k)\|^2 > \varepsilon$ ，则求 λ_k ，确定下一个近似点：

$$X_{k+1} = X_k - \lambda \nabla f(X_k)$$

对于以上方法而言，确定最优的步长是比较复杂的。为保证搜索的次数尽可能少，步长应该尽可能大；然而，步长过大，又可能导致搜索时越过极值点。因此，步长应该控制在一定的范围内。考虑关于 λ 的函数

$$\varphi(\lambda) = f(X_k - \lambda \nabla f(X_k))$$

为保证搜索的次数尽可能少， $\varphi(\lambda)$ 应该尽可能小（因为在寻找极小值点）。最优步长 λ_k 即为使得 $f(X_k - \lambda \nabla f(X_k))$ 取得最小值的 λ 。因此可将 $\varphi(\lambda)$ 对 λ 求导，令 $\varphi'(\lambda) = 0$ ，求解得到的 λ 即为最优步长。以下是计算步长的一种方法：

若 $f(X)$ 具有二阶连续偏导数，则 Hesse 矩阵

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$$

在 X_k 点用 Taylor 公式展开 $f(X_k - \lambda \nabla f(X_k))$ ：

$$f(X_k - \lambda \nabla f(X_k)) \approx f(X_k) - \lambda \nabla f(X_k)^T \nabla f(X_k) + \frac{1}{2} \lambda^2 \nabla f(X_k)^T H(X_k) \nabla f(X_k)$$

上式对 λ 求导，令之为 0，得近似最优步长：

$$\lambda_k = \frac{\nabla f(X_k)^T \nabla f(X_k)}{\nabla f(X_k)^T H(X_k) \nabla f(X_k)}$$

以上算式较为复杂，故可利用计算机进行计算。注意到 Hesse 矩阵：若 $X \in R^n$ ，则算法的空间复杂度为 $\Theta(n^2)$ ，时间复杂度为 $\Omega(n^2)$ 。

3 最速下降法算法的应用实例

$$(1) f(x, y) = x^2 + y^2$$

给定初始点 $X = (1, 1)$ 及精度 $\varepsilon = 1 \times 10^{-4}$ ，则有：

计算 $\nabla f(X_0) = (2, 2)$ ， $\|\nabla f(X_0)\|^2 > \varepsilon$ ，进入第一次搜索；

$X_1 = (1 - 2\lambda, 1 - 2\lambda)$ ， $f(X_1) = 2(1 - 2\lambda)^2$ ，为使得 $f(X_1)$ 最小，对 λ 求导，令导函数为 0，解得 $\lambda_1 = \frac{1}{2}$ ，故 $X_1 = (0, 0)$ ；

计算 $\nabla f(X_1) = (0, 0)$ ， $\|\nabla f(X_1)\|^2 \leq \varepsilon$ ，结束搜索， $X_1 = (0, 0)$ 为所求极值点。

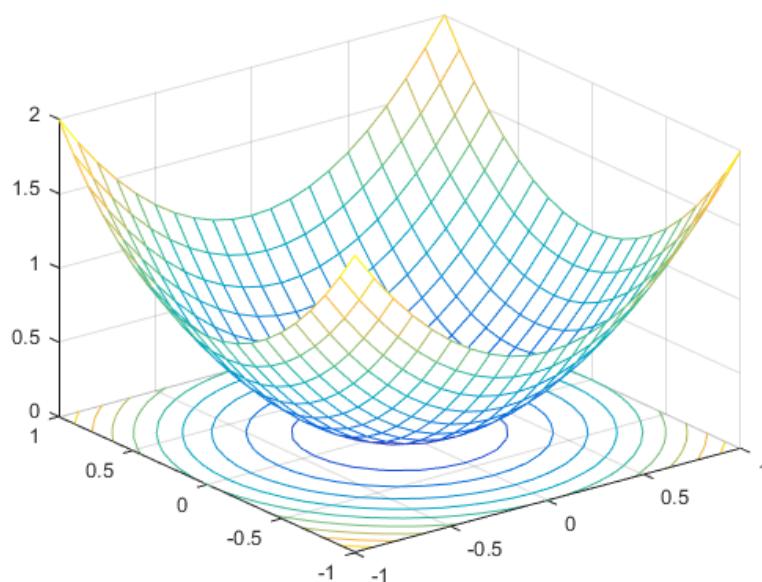


图 3 旋转抛物面

事实上，对于旋转抛物面而言，经过一次搜索即可精确地找到极小值点。

设旋转抛物面

$$z = f(x, y) = a[(x - x_0)^2 + (y - y_0)^2] + z_0$$

初始点 $X_0(x^*, y^*)$, 则:

$$\nabla f = 2a(x - x_0, y - y_0), \quad \nabla f(X_0) = 2a(x^* - x_0, y^* - y_0) \neq \vec{0}$$

进入第一次搜索:

$$X_1 = (2a\lambda x_0 - (2a\lambda - 1)x^*, 2a\lambda y_0 - (2a\lambda - 1)y^*) \\ f(X_1) = a(2a\lambda - 1)^2[(x_0 - x^*)^2 + (y_0 - y^*)^2] + z_0$$

对 λ 求导, 令导函数为 0, 解得 $\lambda_1 = \frac{1}{2a}$, 故 $X_1 = (x_0, y_0)$; $\nabla f(X_1) = 0$, 故 $X_1 = (x_0, y_0)$ 为

所求极值点。

$$(2) \quad f(x, y) = x^2 + ky^2 (k \neq 0)$$

① $k = 1$, $z = f(x, y)$ 为旋转抛物面, 等值线为圆, 见 (1)。

② $k > 0$, $k \neq 1$, $z = f(x, y)$ 为椭圆抛物面, 等值线为椭圆, 搜索路径为锯齿形; 经过有限次搜索, 总能够找到近似极小值点。以下是计算机模拟的结果:

(i) K=2;X=(1,1),EPS=0.01:

表格 4 计算机模拟结果

i	x[i]	y[i]
0	+1.0000	+1.0000
1	+0.4444	-0.1111
2	+0.0741	+0.0741
3	+0.0329	+0.0082

(ii) K=2;X=(2,1),EPS=0.01:

表格 5 计算机模拟结果

i	x[i]	y[i]
0	+2.0000	+1.0000
1	+0.6667	-0.3333
2	+0.2222	+0.1111
3	+0.0741	-0.0370
4	+0.0247	+0.0123

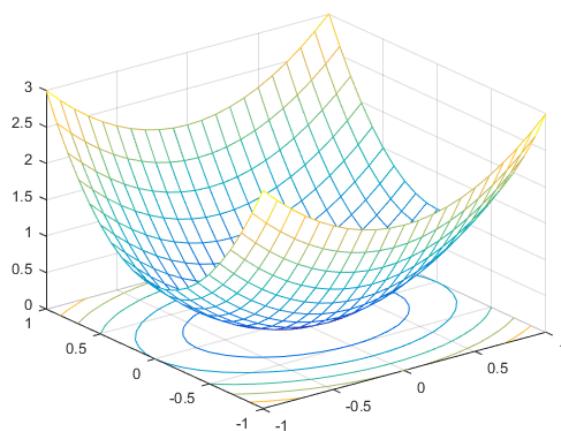


图 4 所得三维曲面

(iii) $K=4; X=(1,1), EPS=0.01$:

表格 6 计算机模拟结果

i	x[i]	y[i]
0	+1.0000	+1.0000
1	+0.0739	-0.0462
2	+0.1108	+0.1108
3	+0.0818	-0.0051
4	+0.0123	+0.0123
5	+0.0091	+0.0006

(iv) $K=4; X=(2,1), EPS=0.01$:

表格 7 计算机模拟结果

i	x[i]	y[i]
0	+2.0000	+1.0000
1	+1.4118	-0.1765
2	+0.5294	+0.2647
3	+0.3737	-0.0467
4	+0.1401	+0.0701
5	+0.0989	-0.0124
6	+0.0371	+0.0185
7	+0.0262	-0.0032

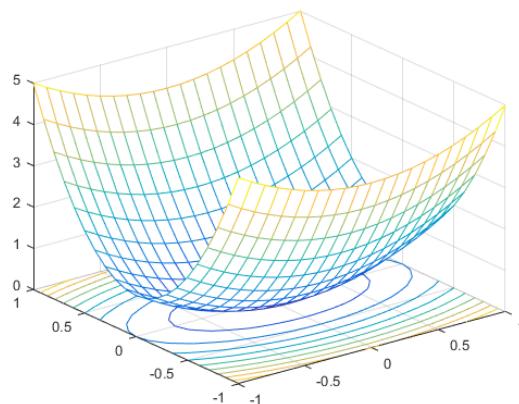


图 5 模拟所得图像

③ $k < 0$ ， $z = f(x, y)$ 为双曲抛物面，无极值。用计算机模拟，将执行无限多次。

4 结论

本文从二元函数的方向导数和梯度出发，对最速下降算法的基本原理进行了简单的研究，并结合一些常见的曲面，对最速下降算法进行了一些简单的应用。

参考文献

- [1] 袁亚湘, 孙文瑜.最优化理论与方法[M].高等教育出版社,2008.7
- [2] 英定,谢云荪.微积分下册[M].高等教育出版社,2009.7

坐标系旋转及矩阵变化下的二维随机变量

文彦博 (2018011203023)

指导教师：骆艳

【摘要】 在处理二维随机变量相关问题时，我们可以通过坐标系的旋转，使得二维随机变量服从的概率密度函数在新坐标系下有简易的表达形式，从而简化概率的求解，进一步的，在旋转坐标系不再适用时，我们可以考虑利用矩阵，进一步探究二维正态随机变量的线性组合所具有的性质。

【关键词】 坐标系旋转；二维随机变量；二维正态随机变量；矩阵

1 特殊的二维正态分布的应用

首先，我们通过一道题进行说明。

设 $(X, Y) \sim N(0, 1; 0, 1; \rho)$ ，计算概率 $P\{-\sqrt{2} < X + Y < 2\sqrt{2}\}$ 。

对于此题，在不知服从正态分布的二维随机变量的线性组合关系时，常规方法是不易求解的，但通过旋转坐标系，可以达到简化的效果。

记 $Z = X + Y = \sqrt{2} \left(\cos \frac{\pi}{4} X + \sin \frac{\pi}{4} Y \right) = \sqrt{2} Y'$, 其中
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \theta = \frac{\pi}{4}$$
 (1)

x', y' 是原坐标系旋转 45° 后的新坐标系下的横纵坐标。

又 $f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right]$, 则其在新坐标系下的形式为：

$$\begin{aligned} f(x', y') &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}((1-\rho)x'^2 + (1+\rho)y'^2)\right] \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\left(\frac{x'^2}{1+\rho} + \frac{y'^2}{1-\rho}\right)\right] \end{aligned} \quad (2)$$

由此可见 $(X', Y') \sim N(0, 1 + \rho; 0, 1 - \rho; 0)$,

又 $P\{-\sqrt{2} < X + Y < 2\sqrt{2}\} = P\{-\sqrt{2} < \sqrt{2}Y' < 2\sqrt{2}\} = P\{-1 < Y' < 2\}$,

再由 X', Y' 相互独立可知 $Y' \sim N(0, 1 - \rho)$.

从而可得 $P = \Phi\left(\frac{2}{\sqrt{1-\rho}}\right) - \Phi\left(-\frac{1}{\sqrt{1-\rho}}\right)$.

由此可见，对特殊的二维正态分布的随机变量的线性组合仍为正态分布，事实上，我们已知对一般的二维正态分布，其随机变量的线性组合仍为正态分布，具体讨论留在第 3 节中。

2 利用坐标系的旋转证明卷积公式

设 (X, Y) 的联合概率密度函数是 $f(x, y)$, 则和 $Z = X + Y$ 的分布函数是

$$F_Z(z) = P\{X + Y \leq z\} = \iint_{x+y \leq z} f(x, y) dx dy \quad (3)$$

同样，通过将原坐标系旋转 45° ，得

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \theta = \frac{\pi}{4} \quad (4)$$

$$X + Y = \sqrt{2} \left(\cos \frac{\pi}{4} X + \sin \frac{\pi}{4} Y \right) = \sqrt{2} Y' \quad (5)$$

记 (X', Y') 的联合概率密度函数为 $g(x', y')$, 则

$$g(x', y') = f\left(\frac{1}{\sqrt{2}}(x' + y'), \frac{1}{\sqrt{2}}(y' - x')\right) \quad (6)$$

记 Y' 的概率密度函数为 $t(y')$, 则

$$t(y') = \int_{-\infty}^{+\infty} g(x', y') dx' \quad (7)$$

于是

$$F_z(z) = P\{X + Y \leq z\} = P\left\{Y' \leq \frac{z}{\sqrt{2}}\right\} \quad (8)$$

$$\int_{-\infty}^{\frac{z}{\sqrt{2}}} \int_{-\infty}^{+\infty} g(x', y') dx' dy' = \int_{-\infty}^{\frac{z}{\sqrt{2}}} \int_{-\infty}^{+\infty} f\left(\frac{1}{\sqrt{2}}(x' + y'), \frac{1}{\sqrt{2}}(y' - x')\right) dx' dy' \quad (9)$$

左右两式求导, 得

$$f_z(z) = \frac{1}{\sqrt{2}} \int_{-\infty}^{+\infty} f\left(\frac{1}{\sqrt{2}}\left(x' + \frac{z}{\sqrt{2}}\right), \frac{1}{\sqrt{2}}\left(\frac{z}{\sqrt{2}} - x'\right)\right) dx' \quad (10)$$

再令 $\frac{1}{\sqrt{2}}\left(x' + \frac{z}{\sqrt{2}}\right) = x$

得

$$f_z(z) = \int_{-\infty}^{+\infty} f(x, z - x) dx \quad (11)$$

卷积公式得证。

3 二维正态随机变量线性组合性质

首先考虑二维情况, $(X, Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$, 考虑其线性组合

$$Z = l_1 X + l_2 Y, (l_1 \neq 0, l_2 \neq 0)$$

在第 1 节中, 我们通过旋转原坐标系解决了问题, 但不幸的是, 对于一般的二维正态随机变量, 旋转坐标系已经不再适用, 故而此时我们考虑利用适当的矩阵解决问题, 具体思路为, 寻找适当的变化矩阵, 使得 z 与 x, y 的另一线性组合 v 服从二维正态分布。

此时 x, y 满足函数关系式:

$$\varphi(x, y) = \frac{1}{2\pi(\det \mathbf{C})^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{X} - \boldsymbol{\mu})^T \mathbf{C}^{-1}(\mathbf{X} - \boldsymbol{\mu})\right\} \quad (12)$$

$$\mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad (13)$$

考虑变量 (Z, V) , 其满足如下关系:

$$\mathbf{Z} = \begin{pmatrix} Z \\ V \end{pmatrix} = \begin{pmatrix} l_1 & l_2 \\ s & t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\text{其中 } s, t \text{ 满足关系: } tl_1 - sl_2 = 1, \text{ 即 } \left| \begin{array}{cc} l_1 & l_2 \\ s & t \end{array} \right| = 1 \quad (14)$$

于是可得关系式:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t & -l_2 \\ -s & l_1 \end{pmatrix} \begin{pmatrix} Z \\ V \end{pmatrix}$$

进一步的, 可以构造矩阵 $\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$, 其满足如下关系:

$$(\mathbf{X} - \boldsymbol{\mu}) = \begin{pmatrix} t & -l_2 \\ -s & l_1 \end{pmatrix} (\mathbf{Z} - \boldsymbol{\varepsilon})$$

其中 $\boldsymbol{\varepsilon}$ 由 $\boldsymbol{\mu}, l_1, l_2, s, t$ 确定。

现将 (z, v) 带入原方程:

$$\varphi(z, v) = \frac{1}{2\pi(\det \mathbf{C})^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\mathbf{Z} - \boldsymbol{\varepsilon})^T \begin{pmatrix} t & -s \\ -l_2 & l_1 \end{pmatrix} \mathbf{C}^{-1} \begin{pmatrix} t & -l_2 \\ -s & l_1 \end{pmatrix} (\mathbf{Z} - \boldsymbol{\varepsilon}) \right\} \quad (15)$$

考虑矩阵

$$\mathbf{C}'^{-1} = \begin{pmatrix} t & -s \\ -l_2 & l_1 \end{pmatrix} \mathbf{C}^{-1} \begin{pmatrix} t & -l_2 \\ -s & l_1 \end{pmatrix} \quad (16)$$

当如下关系式成立时

$$\rho\sigma_1\sigma_2(sl_2 + tl_1) + tl_2\sigma_2^2 + sl_1\sigma_1^2 = 0$$

可知 \mathbf{C}'^{-1} 为对角矩阵.

于是, 由关于 s 与 t 的二元一次方程组

$$\begin{cases} \rho\sigma_1\sigma_2(sl_2 + tl_1) + tl_2\sigma_2^2 + sl_1\sigma_1^2 = 0 \\ tl_1 - sl_2 = 1 \end{cases}$$

可确定唯一的一组 s 与 t, 使得

$$\mathbf{Z} = \begin{pmatrix} z \\ v \end{pmatrix} = \begin{pmatrix} l_1 & l_2 \\ s & t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

又根据 (14) 和 (16) 知 $\det \mathbf{C}' = \det \mathbf{C}$.

进一步的

$$\varphi(z, v) = \frac{1}{2\pi(\det \mathbf{C}')^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\mathbf{Z} - \boldsymbol{\varepsilon})^T \mathbf{C}'^{-1} (\mathbf{Z} - \boldsymbol{\varepsilon}) \right\}$$

由此可见 (z, v) 服从二维正态分布, 于是 $Z = l_1 X + l_2 Y$ 是一维正态随机变量, 命题得证.

参考文献

- [1] 黄廷祝, 蒲和平, 线性代数与空间解析几何, 高等教育出版社, 2015
- [2] 徐全智, 吕怒, 概率论与数理统计, 高等教育出版社, 2017

不同相关系数定义辨析及应用举例

甘玲 (2018120706004)

指导教师：吕恕

【摘要】 通过对 Pearson, Spearman, Kendall 相关系数定义的探究，归纳总结几种不同定义方式下对相关性描述的优缺点，并举例其基本应用。

【关键词】 相关系数；定义方法；应用

1 相关系数定义

1.1 Pearson 相关系数

为消除协方差存在量纲的问题，Karl Pearson 在 20 世纪初提出了相关系数的概念^[1]。面对总体和样本，Pearson 相关系数有不同的计算公式：

总体(X,Y 相互独立)：

$$\rho_{XY} = \frac{n \sum XY - \sum X \sum Y}{\sqrt{n \sum X^2 - (\sum X)^2} \sqrt{n \sum Y^2 - (\sum Y)^2}} = \frac{COV(X,Y)}{\sqrt{D(X)D(Y)}} \quad (1)$$

样本：

$$\rho_{XY} = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2(y-\bar{y})^2}} = \frac{l_{xy}}{\sqrt{l_{xx}l_{yy}}} \quad (2)$$

Pearson 相关系数用来衡量两个连续变量之间的相关关系。取值在[-1,1]之间，其绝对值 r 表示两变量间相关关系程度的强弱，越接近 1，表明两变量间相关程度越高，它们之间的关系越密切。r > 0 表示正相关，r < 0 表示负相关，r = 0 表示不相关，r = 1 表示完全相关。如果 r 是样本统计量，其取值会受到抽样波动性的影响，因此一般可通过 r 进行统计检验判断两个变量之间线性关系是否显著^[2]。

1.2 Spearman 相关系数^[3]

斯皮尔曼相关系数被定义成等级变量之间的皮尔逊相关系数。对于样本容量为 n 的样本，n 个原始数据被转换成等级数据，相关系数 ρ 为

$$\rho = 1 - \frac{6 \sum D_i^2}{n(n^2-1)} \quad (3)$$

式中， ρ 表示等级相关系数； D_i 为每对观测值的等级差；n 为样本容量。计算过程可理解为首先对两个变量(X, Y)的数据进行排序，然后记下排序以后的位置(X', Y')，(X', Y')的值就称为秩次， D_i 是秩次的差值，n 是变量中数据的个数，最后根据公式求解结果。 ρ 的取值范围为 [-1,1]，当一个变量随另一个变量单调递增的时候， $\rho = 1$ ；反之，当一个变量随另一个变量单调递减的时候， $\rho = -1$ ^[4]。

1.3 Kendall 相关函数^[5]

Kendall 相关系数的计算也是以变量 X 和 Y 的等级数据进行，根据配对等级顺序排列的位置是否颠倒或换位，得出等级换位的次数，进而进行计算。其计算公式为：

$$\rho = \frac{C-D}{\sqrt{(N_3-N_1)(N_3-N_2)}} \quad (4)$$

其中 C 表示 XY 中拥有一致性的元素对数（两个元素为一对）；D 表示 XY 中拥有不一致性的元素

对数。

$$N_3 = \frac{1}{2}N(N-1); N_1 = \frac{1}{2}\sum_{i=1}^s U_i(U_i-1); N_2 = \frac{1}{2}\sum_{i=1}^s V_i(V_i-1) \quad (5)$$

N_1 、 N_2 分别是针对集合 X 、 Y 计算的，现在以计算 N_1 为例，给出 N_1 的由来(N_2 的计算可以类推)：将 X 中的相同元素分别组合成小集合， s 表示集合 X 中拥有的小集合数（例如 X 包含元素：1 2 3 4 3 3 2，那么这里得到的 s 则为 2，因为只有 2、3 有相同元素）， U_i 表示第 i 个小集合所包含的元素数。

如果两个属性排名是相同的，系数为 1，两个属性正相关；如果两个属性排名完全相反，系数为 -1，两个属性负相关；如果排名是完全独立的，系数为 0。

2 三种定义方式的辨析与使用范围

因为 Pearson 相关系数公式的分母是变量的标准差，所以计算 Pearson 相关系数时，变量的标准差不能为 0（分母不能为 0），即两个变量中任何一个的值不能都是相同的。在这个基础上我们提出了 Spearman 相关系数。

相比于 Pearson 相关系数，Spearman 相关系数对统计数据没有那么多的条件要求，只要两个变量的观测值是成对的等级评定资料，或者是由连续变量观测资料转化得到的等级资料，不论两个变量的总体分布形态、样本容量的大小如何，都可以用 Spearman 等级相关来进行研究^[6]。即使统计数据出现异常值，由于异常值的秩次通常不会有明显的变化（比如过大或者过小，要么排第一，要么排最后），对 Spearman 相关系数的影响也非常小。但是在能用 Pearson 相关系数解决问题的情况下，Spearman 相关系数的精确度会变低，这对于来说是不想见到的。

之后 Kendall 提出了精确度更高的 Kendall 相关系数，Kendall 相关系数的置信区间更容易解释，可靠性也更高。但是 Kendall 相关系数只是说明两组数据相关性高低，并不反映线性相关程度的大小。

3 应用举例

a) Pearson 相关系数

将一枚硬币抛 n 次，想知道他出现正反面次数的相关系数。首先假设变量 X 表示 n 次试验中出现正面的次数， Y 表示 n 次试验中出现反面的次数。由于 $X + Y = n$ ，则 $Y = -X + n$ ，根据相关系数的性质推论^[2]，得 $\rho_{XY} = -1$ 。

b) Spearman 相关系数

假设当前有 5 个人的视觉、听觉反应时（单位：毫秒），数据如下表。请问视觉、听觉反应时是否具有一致性？

表一

被试	听觉反应时	视觉反应时	X	Y	D	D^2	XY
1	170	180	3	4	-1	1	12
2	150	165	1	1	0	0	1
3	210	190	5	5	0	0	25
4	180	168	4	2	2	4	8
5	160	172	2	3	-1	1	6
Σ	870	875	15	15		6	52

因为被测试对象只有 5 人，属于小样本，不知是否为正态分布，其中， x 为听觉反应时间按大小排序， y 为视觉反应时间按大小排序，用 Spearman 相关系数公式(3)，计算得： $\rho = 0.7$ 。

c) 实际举例

想要了解空气湿度与体感温度的相关程度, 我们在中国天气网找到成都 2019 年天气数据如(表二)。

表二

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
平均温度	7°C	10°C	15°C	23°C	23°C	26°C	26°C	28°C	23°C	19°C	14°C	11°C
体感温度	6°C	9°C	15°C	23°C	23°C	28°C	29°C	32°C	24°C	20°C	14°C	11°C
平均温度-体感温度	1°C	1°C	0°C	0°C	0°C	2°C	3°C	5°C	1°C	1°C	0°C	0°C
平均湿度	6 6%	62%	56%	56%	62%	64%	74%	69%	74%	70%	68%	57%

用变量 X 表示平均温度和体感温度差的绝对值, Y 表示平均湿度。计算得 $\bar{X} = \frac{7}{6}$, $\bar{Y} = \frac{389}{6}$, 由 Pearson 公式 (2) 计算可得 $\rho_1 = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2(y-\bar{y})^2}} = \frac{l_{xy}}{\sqrt{l_{xx}l_{yy}}} \approx 0.5474$

将表格中得 X, Y 排序, 相同位次的数据序号取平均值 (如表三)。

表三

.	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
X_i	5.5	5.5	10	10	10	3	2	1	5.5	5.5	10	10
Y_i	6	8.5	11.5	11.5	8.5	7	1.5	4	1.5	3	5	10
D_i^2	0.25	9	2.25	2.25	2.25	16	0.25	9	16	6.25	25	0

$\sum D_i^2 = 88.5$, 应用 Spearman 相关系数得 $\rho_2 = 1 - \frac{6\sum D_i^2}{n(n^2-1)} \approx 0.6906$ 。

计算第三种相关系数中系数得: $C = 39$, $D = 9$, $N_3 = 66$, $N_1 = 16$, $N_2 = 3$, 代入公式 (4) 得

$$\rho_3 = \frac{C - D}{\sqrt{(N_3 - N_1)(N_3 - N_2)}} = 0.5345$$

根据获得的相关系数得值, 我们可以认为空气湿度与体感温度有相关关系。

比较得到的三种相关系数, Pearson 相关系数虽然适用的范围有局限, 但是能体现一组数据更多性质, 有更高的可靠度。后两种是等级相关系数, 不能表现数据的线性相关性。

4 总结

在对几种相关系数的探索中, 了解到关于统计量的提出是可以从不同角度入手, 只要面对原定义更好的一方面, 那它就有存在的必要, 告诉我们在学习数理统计部分的时候, 不要只死板的盯着书上那一种统计量的给定办法, 也可以从其他角度构造统计量, 提高我们的创新思维和加深对统计的认识。

参考文献

- [1] Pearson K. Notes on the History of Correlation[J]. Biometrika, 1920, 13(1).
- [2] 徐权智, 吕恕. 概率论与数理统计(第二版)[M]. 高等教育出版社, 2012, 210-213.
- [3] Spearman C. The Proof and Measurement of Association Between Two Things[J]. The American Journal of Psychology, 1904, 15(1).
- [4] 樊嵘, 孟大志, 徐大舜. 统计相关性分析方法研究进展[J]. 数学建模及其应用, 2014, 3(1).
- [5] Kendall M G. A New Measure of Rank Correlation[J]. Biometrika, 1938, 30(1/2).
- [6] 金林, 李研. 几种相关系数辨析及其在 R 语言中的实现[J]. 统计与信息论坛, 2019 年 4 月, 第 34 卷第 4 期 (3-11).

关于最短置信区间的讨论

侯付豪 (2018040710008)

指导教师：吕恕

【摘要】 在概率论与数理统计的学习中，我们常会发现书中对于一些问题并未给出详细的解释或讨论，例如区间估计这一章中，利用枢轴变量所求得的置信区间在正态分布和t分布时是对称的，而卡方分布和F分布并不对称，为什么得出这一结果？本文将会在一个总体和两个总体的情况下讨论几种常见枢轴变量（正态分布、卡方分布、t分布、F分布等）最位置信区间的估计问题，进而给出一个普遍性结论。

【关键词】 区间估计；枢轴变量；置信区间

1 区间估计的研究背景

我们知道，区间估计是给出一个范围，并在要求的可信程度下保证这个范围包含所估计的未知参数，如今比较流行的一种区间估计方法是统计学家 Neyman 创立的区间估计理论。

设 θ 是总体分布中的一个未知参数， X_1, X_2, \dots, X_n 是来自总体的样本。对 θ 的区间估计，就是给出两个统计量 $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$ 、 $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$ ，并且恒有 $\hat{\theta}_1 \leq \hat{\theta}_2$ ，以它们为端点构成一个区间 $[\hat{\theta}_1, \hat{\theta}_2]$ ，则区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 是一个随机区间。对于区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 有两个要求：

① $P\{\hat{\theta}_1 \leq \theta \leq \hat{\theta}_2\}$ 要尽可能大，此概率反映了区间估计的可靠程度。

② 区间长度 $\hat{\theta}_2 - \hat{\theta}_1$ 要尽可能小，即精度要尽可能高。

在样本容量不改变的前提下，很难同时满足这两个要求，于是，一般的区间估计原则是：先确定一个可靠程度，在这个可靠度的前提下是精度尽可能提高。为此，有如下定义：

定义 1^[1] 给定一个很小的数 $\alpha > 0$ ，如果对参数 θ 的所有可能取值，都有

$$P\{\hat{\theta}_1(X_1, X_2, \dots, X_n) \leq \theta \leq \hat{\theta}_2(X_1, X_2, \dots, X_n)\} = 1 - \alpha$$

则称随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 是 θ 的置信度为 $1 - \alpha$ 的置信区间。

对于构造置信区间，我们常使用枢轴变量法，即：找出一个关于待估参数 θ 的良好估计量；构造一个包含位未知参数 θ 及其估计量的函数 W ，函数 W 称为枢轴变量；对于任意参数 $a < b$ ，不等式 $a \leq W \leq b$ 可以改写为等价的形式 $A \leq \theta \leq B$ ，且满足 $P\{A \leq \theta \leq B\} = 1 - \alpha$ ， A 与 B 是不包含未知参数的统计量；区间 $[A, B]$ 为参数 θ 的置信度为 $1 - \alpha$ 的置信区间。

2、置信区间的最短条件

在给出一个总体参数的置信区间的最短条件之前，需要给出相关的定义及定理。

定义 2^[2] 设连续型随机变量 X 的概率密度为 $f(x)$ ，在 $(-\infty, +\infty)$ (或 $(0, +\infty)$) 内有唯一极大值点 x_0 ，则称 X 的分布为单峰分布，点 x_0 称为其峰点。

最短置信区间的问题可以转化为：在 $P\{A \leq \theta \leq B\} = 1 - \alpha$ 条件下，求 $B - A$ 的最小值。对于枢轴变量 W 来说，就是在 $P\{a \leq W \leq b\} = 1 - \alpha$ 条件下，求 $b - a$ 的最小值。那么，根据所学过的条件极值问题，我们可以得到如下定理：

定理 1 设 θ 是总体分布中的一个未知参数， X_1, X_2, \dots, X_n 是来自总体的样本。枢轴变量 W 的概率密度函数 $W(x)$ 是单峰一阶可微函数，在约束条件 $P\{a \leq W \leq b\} = 1 - \alpha$ 下，枢轴变量 W 的最短置信区间 $[a, b]$ 满足， $W(a) = W(b)$ 。

证明：证明此定理可以用拉格朗日乘数法，即目标函数为 $f(a,b) = b-a$, 约束条件为 $P\{a \leq W \leq b\} = 1-\alpha$ ，由于 $W(x)$ 为单峰一阶可微函数，则约束条件可化简为 $\int_a^b W(x)dx - (1-\alpha) = 0$ 。

作拉格朗日函数：

$$F(a,b, \lambda) = (b-a) + \lambda [\int_a^b W(x)dx - (1-\alpha)]$$

求 F 关于 a, b 的偏导数，并令其等于零：

$$F_a = -1 - \lambda W(a) = 0$$

$$F_b = 1 + \lambda W(b) = 0$$

解得：

$$W(a) = W(b) = \frac{1}{\lambda}$$

所以，当 $[a,b]$ 为最短置信区间时， a, b 应满足 $W(a) = W(b)$ 。

此时，我们可以发现，当枢轴变量 W 为正态分布、 t 分布等概率密度关于 y 轴对称的分布时，置信区间 $[a,b]$ 也关于 y 轴对称的，这和我们课本上所提到的置信区间为对称区间相吻合。

例如^[3] 设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，其中 σ^2 已知，求未知参数 μ 的置信度为 $1-\alpha$ 的置信区间。

解 取枢轴变量 $W = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ ，由标准正态分布的上侧分位数可知，对于 $P\{a \leq W \leq b\} = 1-\alpha$ ，有 $P\{|W| \leq u_{\frac{\alpha}{2}}\} = 1-\alpha$ ，即 $P\{-\frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}\} = 1-\alpha$ ，所以 μ 的置信度为 $1-\alpha$ 的置信区间为 $[\bar{x} - \frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}, \bar{x} + \frac{\sigma}{\sqrt{n}}u_{\frac{\alpha}{2}}]$ ，关于 y 轴对称。

从这道题中我们可以看出，当求最短置信区间时，我们选取的上侧分位数满足了 $W(u_{\frac{\alpha}{2}}) = W(-u_{\frac{\alpha}{2}})$ ，这也就印证了定理 1 的结论。

那么，如果枢轴变量 W 的分布不对称（卡方分布， F 分布等）时，置信区间 $[a,b]$ 的情况又如何呢？

课本上所给出的卡方分布， F 分布的枢轴变量的置信区间分别为 $[\chi^2_{1-\frac{\alpha}{2}(n-1)}, \chi^2_{\frac{\alpha}{2}(n-1)}]$ 和 $[F_{1-\frac{\alpha}{2}(n_1-1, n_2-1)}, F_{\frac{\alpha}{2}(n_1-1, n_2-1)}]$ ，此时的置信区间并不满足 $W(a) = W(b)$ 的条件（也可以看做等高条件），而是满足左右侧等面积的条件（也可以看做等尾条件）。那么根据定理 1，这两个置信区间并不是最短置信区间。通过计算机模拟可以明显看出置信区间分别满足等高条件和等尾条件时的差异：

表 2.1^[4] 常见枢轴变量的等尾置信区间和等高置信区间比较

枢轴变量分布	置信水平	等尾置信区间长度	等高置信区间长度	长度差
N(0,1)	0.10	3.290	3.290	0.000
	0.05	3.920	3.920	0.000
	0.01	5.152	5.152	0.000
t(16)	0.10	3.492	3.492	0.000
	0.05	4.240	4.240	0.000
	0.01	5.842	5.842	0.000
$\chi^2_{(10)}$	0.10	14.367	13.693	0.674
	0.05	17.236	16.447	0.789
	0.01	23.032	22.035	0.997
F(9,14)	0.10	2.315	1.995	0.320
	0.05	2.946	2.549	0.397
	0.01	4.553	3.997	0.556

从表中可以得到如下结论：

- ①当枢轴变量满足对称分布（正态分布、t 分布等）时，等尾置信区间与等高置信区间无差异，即均为最短置信区间；
- ②当枢轴变量不满足对称分布（卡方分布，F 分布等）时，等尾置信区间的精确度明显低于等高置信区间，此时等尾置信区间并非最短置信区间；
- ③平时为了计算方便，常用等尾置信区间代替等高置信区间。

3 总结

本文从最短置信区间的求取出发，利用了拉格朗日乘数法，得到了当枢轴变量置信区间为等高置信区间时，此区间为最短置信区间的结论。这个结论对于概率密度为单峰一阶可微函数的枢轴变量具有适用性。但是，在实际求解过程中，当枢轴变量的分布不对称时，由于求解等高置信区间较为困难，所以一般用等尾置信区间来代替等高置信区间。

参考文献

- [1] 徐全智,吕恕.概率论与数理统计(第三版) [M].北京：高等教育出版社，2017.
- [2] 刘瑞香.枢轴量为单峰分布的最短区间估计[J].统计与决策，2011,17(164-165).
- [3] 徐全智,吕恕.概率论与数理统计(第三版) [M].北京：高等教育出版社，2017.
- [4] 康文倩.枢轴量服从单峰分布情形下最佳置信域的等高性质探讨及其数值模拟[D].云南：云南师范大学，2015.

以频率估计概率的误差分析

胡嘉仪 (2018051403025)

指导教师：吕恕

【摘要】 在实际生活中，我们常用事件发生的频率估计其概率。本文利用德莫弗-拉普拉斯中心极限定理得到了以频率估计概率的近似误差，并给出了该误差的部分应用。

【关键词】 频率；概率；误差；伯努利大数定律；德莫弗-拉普拉斯中心极限定理

1 引言

伯努利大数定律证明了，在大量重复的试验中，随着试验次数地无限增加，随机事件发生的频率稳定在概率附近上下摆动，频率的稳定性为用频率估计概率提供了理论依据。

但是在实际生活中，用频率估计概率通常不够精确，存在误差。以下提供一种利用德莫弗-拉普拉斯中心极限定理近似计算该误差的方法。

2 误差计算公式推导

在伯努利试验中，设 f_n 是 n 次独立重复试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对于任意正整数 $\varepsilon > 0$ ，由伯努利大数定理，有：

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{f_n}{n} - p \right| < \varepsilon \right\} = 1 \quad (1)$$

即当 n 趋近于无穷的时候， $\frac{f_n}{n}$ 依概率收敛于 p，即频率收敛于概率。

事实上，n 不可能取到无穷，只能尽可能大，而此时由德莫弗-拉普拉斯中心极限定理可知：

$$\lim_{n \rightarrow \infty} P \left\{ \frac{f_n - np}{\sqrt{np(1-p)}} \ll x \right\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \phi(x) \quad (2)$$

定理表明，n 很大的时候， $\frac{f_n - np}{\sqrt{np(1-p)}}$ 可以近似为正态分布，据此可以将二项分布做近似，即若 $X \sim B(n, p)$ ，则 $\frac{X - np}{\sqrt{np(1-p)}} \sim N(0, 1)$ ，且有：

$$P\{X \ll a\} = P \left(\frac{X - np}{\sqrt{np(1-p)}} \ll \frac{a - np}{\sqrt{np(1-p)}} \right) \approx \phi \left(\frac{a - np}{\sqrt{np(1-p)}} \right) \quad (3)$$

则通过将 $P \left\{ \left| \frac{f_n}{n} - p \right| < \varepsilon \right\}$ 变形可以得到：

$$\begin{aligned} P \left\{ \left| \frac{f_n}{n} - p \right| < \varepsilon \right\} &= P \left\{ -\varepsilon < \frac{f_n}{n} - p < \varepsilon \right\} = P \left\{ -\frac{n\varepsilon}{\sqrt{np(1-p)}} < \frac{f_n - np}{\sqrt{np(1-p)}} < \frac{n\varepsilon}{\sqrt{np(1-p)}} \right\} \\ &= P \left\{ -\varepsilon \sqrt{\frac{n}{p(1-p)}} < \frac{f_n - np}{\sqrt{np(1-p)}} < \varepsilon \sqrt{\frac{n}{p(1-p)}} \right\} \\ &\approx \phi \left(\varepsilon \sqrt{\frac{n}{p(1-p)}} \right) - \phi \left(-\varepsilon \sqrt{\frac{n}{p(1-p)}} \right) = 2\phi \left(\varepsilon \sqrt{\frac{n}{p(1-p)}} \right) - 1 \end{aligned} \quad (4)$$

该式子为当 n 有限大时，用频率估计概率存在的误差计算方法，即误差小于 ε 的概率可用正态分布估计，近似为 $2\phi \left(\varepsilon \sqrt{\frac{n}{p(1-p)}} \right) - 1$ 。

3 误差计算的应用拓展

例 1. 在 n 次独立重复实验中, 事件 A 发生的概率 p 未知。问: 至少进行多少次试验才能使事件发生的频率与 p 相差不超过 0.01 的概率达到 95%。

方法一: 由切比雪夫不等式计算

令 X 表示 n 重伯努利试验中事件 A 出现的次数, 则 $X \sim B(n, p)$

$$E(X) = np, D(X) = np(1-p), \text{ 频率为 } \frac{f_n}{n},$$

$$\begin{aligned} P\left\{\left|\frac{f_n}{n} - p\right| < 0.01\right\} &= P\{|f_n - np| < 0.01n\} \\ &\geq 1 - \frac{D(X)}{(0.01n)^2} = 1 - \frac{np(1-p)}{(0.01n)^2} \end{aligned} \quad (5)$$

$$1 - \frac{np(1-p)}{(0.01n)^2} \geq 0.95 \quad (6)$$

$$np(1-p) \leq (0.01n)^2 \times 0.05 \quad (7)$$

$$n \geq \frac{p(1-p)}{0.01^2 \times 0.05} = 200000p(1-p) \quad (8)$$

方法二: 利用频率估计概率的误差公式计算

$$P\left(\left|\frac{f_n}{n} - p\right| < 0.01\right) \geq 0.95 \quad (9)$$

$$P\left(\left|\frac{f_n}{n} - p\right| < 0.01\right) \approx 2\phi\left(0.01\sqrt{\frac{n}{p(1-p)}}\right) - 1 \geq 0.95 \quad (10)$$

$$\phi\left(0.01\sqrt{\frac{n}{p(1-p)}}\right) \geq 0.975 \quad (11)$$

查标准正态分布表, 得到:

$$0.01\sqrt{\frac{n}{p(1-p)}} \geq 1.96 \quad (12)$$

$$n \geq 196^2 p(1-p) \quad (13)$$

即得到在某一确定置信度下, 以频率估计概率所产生的误差与试验次数之间关系式。

显然可得, 方法二所需的试验次数少于方法一所需试验次数, 即方法二更好。

例 2. 在 n 次独立重复试验中, 事件 A 发生的概率为 p , 求频率与概率之差不超过 0.01 的概率。

若分布或方差已知, 即可用切比雪夫不等式计算;

若方差未知, 则可利用以上结论:

$$P\left(\left|\frac{f_n}{n} - p\right| < 0.01\right) \approx 2\phi\left(0.01 \times \sqrt{\frac{n}{p(1-p)}}\right) - 1 \quad (14)$$

得到所求概率。

参考文献

- [1] 宁如云, 李德清, 棟莫弗-拉普拉斯中心极限定理应用中的一些问题[J], 科技信息, 2010,(25):542.
- [2] 郭向荣, 以频率估计概率的误差估计及其应用[J], 廊坊师范学院学报(自然科学版), 2009,9(03):9-10 +15.

线性回归方法与相关系数在足球分析中的应用

金嘉明 (2018051410004)

指导教师：吕恕

【摘要】 在概率论与数理统计的学习中，我们明白了一元线性回归的作用。但是这种方法有明显的不足，它不能很好地解决更为实际普遍存在的问题。在此文中将给出一些方法使这种分析手段更加实际可靠。并且运用相关性分析、假设检验等方法对实际的模型进行分析。

【关键词】 一元线性回归；假设检验方法；相关性分析；Pearson 相关；Spearman 相关

1 引例：15/16 赛季西班牙足球甲级联赛进失球之比与联赛积分的关系

设有两个随机变量 X、Y 具有某种统计关系，当 X 变化时，Y 也相应的变化。通常称为 X 自变量，Y 为因变量，这种研究 X 与 Y 之间关系的统计方法称为回归分析。回归分析即找到 X 与 Y 之间的某种函数关系 $Y = f(X, \beta)$ (β 为待估计的参数或者一系列参数的组合)，使其最能代表 X 与 Y 之间的统计关系。我们常常通过假定的模型来寻找这种变量的依存关系，当用线性模型进行参数估计时，称为线性回归；选用其他模型进行回归时则称为非线性回归。最小二乘法是进行回归分析时最常用的方法。

以一个实例开始分析，下面给出的是 15/16 赛季西班牙足球甲级联赛进失球之比与联赛积分关系的表格。（数据均真实，可靠）

表一 15/16 赛季西班牙足球甲级联赛进失球之比与联赛积分

球队得分	91	90	88	64	62	60	52	48	48	45	44	44	43	43	42	39	39	38	36	32
球队（进球/失球）	112 / 29	110 / 34	63 / 18	44 / 35	58 / 45	51 / 59	51 / 50	38 / 35	45 / 48	34 / 52	45 / 53	46 / 48	40 / 74	49 / 61	45 / 61	46 / 69	40 / 62	52 / 73	37 / 67	37 / 70

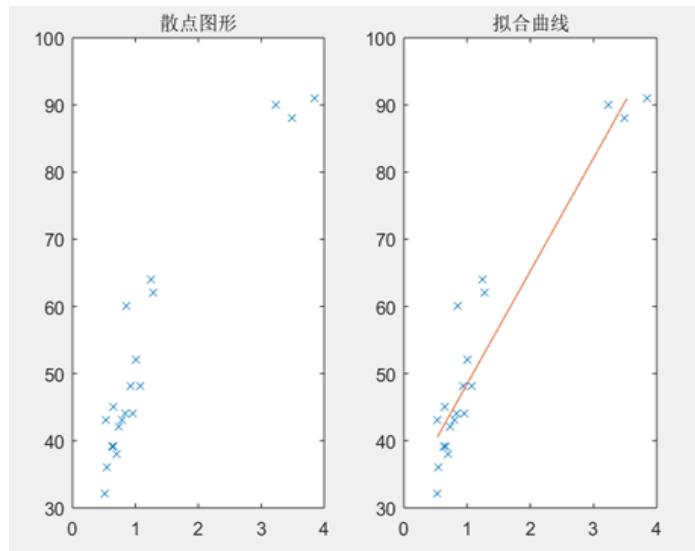
一般我们可以认为一个球队的成绩优秀与否很大程度取决于球队的进球与失球，进球与失球的比值若大于一，则可认为该球队常常取胜。通过直觉我们也可以察觉到，球队赛季的最终得分，应该与进失球之比正相关。采用一元线性回归的方法， Y 代表求队的得分， x 代表得失球之比：

$$Y = a + bx + \varepsilon \quad (1)$$

通过最小二乘法得出 a, b 的估计量： a' , b' 。

$$b' = l_{xy} / l_{xx} ; \quad a' = \bar{y} - b' \cdot \bar{x} \quad (2)$$

可以求出 a, b 的估计量，并拟合出曲线如下：



图一 直接段线性回归拟合图

通过计算可以得到：

$$a' = 31.64;$$

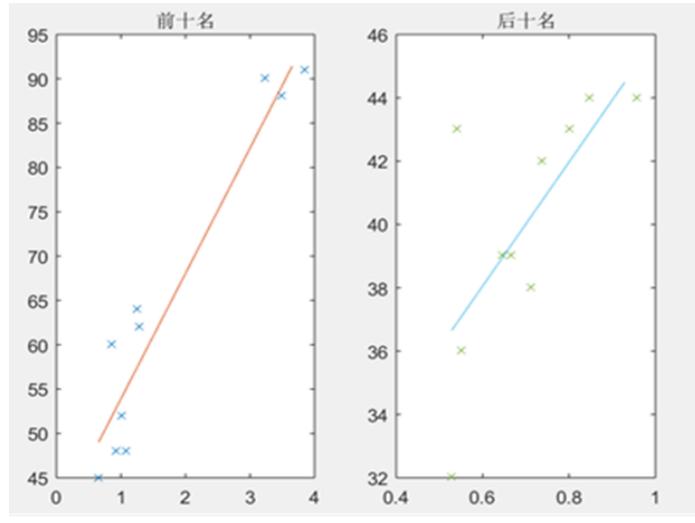
$$b' = 16.81;$$

可以通过图形直观发现这条直线没有那么尽如人意，因为该线对于中下球队以及顶级球队（前三名）的符合度较好，而对中等球队、最后几名的球队的预测偏差较大。

通过计算得出此时的总体方差的无偏估计值为：D=36.44。

2 根据情况分段线性回归

下面我们改进一元线性回归的方法。在散点图形中可以发现，上游球队与下游球队存在两条不一样的拟合曲线。运用同样的一元线性回归的拟合办法，分别对前 10 名球队，后 10 名球队进行拟合。得出拟合直线如下：



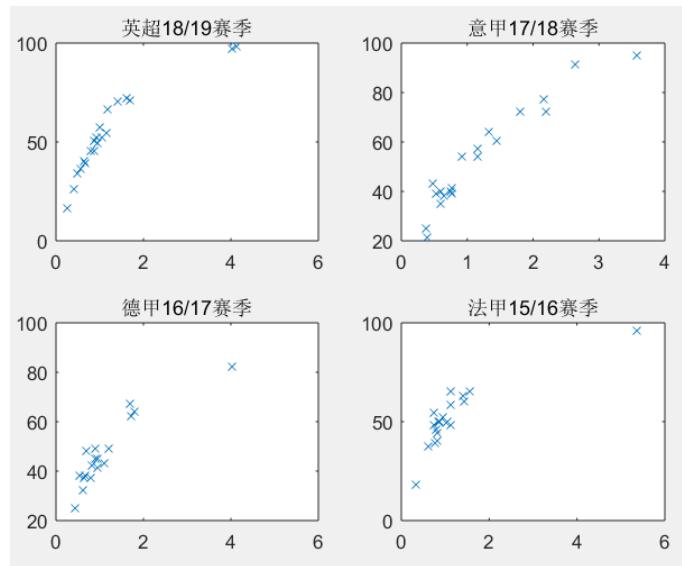
图二 分阶段线性回归拟合图

前十名： $a' = 39.76 \quad b' = 14.14$ 总体方差的无偏估计值 $D = 31.48$ 。

后十名： $a' = 26.31 \quad b' = 19.58$ 总体方差的无偏估计值 $D = 8.78$ 。

由 D 的值可见，后面一种方法的拟合程度更好。

通过对不同联赛（英超，西甲，德甲，法甲，意甲），不同赛季的情况进行总结如下图：



图三 不同联赛不同赛季参数拟合图

可以发现不同联赛的不同赛季都存在着着上下游球队拟合曲线斜率不一的情况，故对于足球俱乐部的比赛的一般赛季，前述模型都可采用。可以通过球队的当前排名，进失球比例综合考虑，带入上一赛季的拟合曲线，就可大致推算出当前赛季结束后球队的最终得分。

在实际解决问题时，一元线性回归方法不一定能很良好地拟合出符合总体的情况，这时可根据实际情况对拟合线进行分段拟合，这样得出的曲线更加准确。

3 对当前赛季（19-20）西甲西班牙人的保级形势进行数学预测

i) 当前赛季刚过半，西班牙人保级形势严峻，我们可以用得失球比来大致分析西班牙人赛季末的得分情况，并与上赛季西甲联赛的保级线作对比得出结论。

ii) 由于西班牙人出于联赛排名的后半区，现使用前述的建模结果。可以计算出西班牙人这赛季末尾的得分：

$$Y = a + bx \quad (3)$$

计算结果为 33.22 分。

对照前一赛季的降级线（37 分），可以认为西班牙人的降级可能性很大。

4 进失球比率与赛季得分的相关性分析

i) 两种相关性分析：Pearson 相关性分析、Spearman 相关性分析的异同及用途

相关系数是最早由统计学家卡尔·皮尔逊设计的统计指标，是研究变量之间线性相关程度的量。

Pearson 相关：双变量正态分布资料的直线相关性一般用 Pearson 相关系数 $r(\rho)$ 来进行描述。

Pearson 相关系数又称积差相关系数，它是说明具有直线关系的两个变量间相关关系的密切程度与相关方向的指标。公式如下：

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (4)$$

Spearman 等级相关：在实际工作中，常遇到有些资料并不呈正态分布，或总体分布类型未知，或等级资料的情形。对于这些资料，分析变量间的关系就不宜用 Pearson 相关系数来描述变量间的

相关关系，而常选用 Spearman 等级相关进行统计推断。Spearman 等级相关又称为秩相关，是将原始数据资料经过秩转换后计算等级相关系数 r_s (ρ)，说明两个变量间直线相关关系的密切程度和相关方向。这类方法对原变量分布不作要求，属于非参数统计方法。公式如下：

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \quad (5)$$

通俗来讲， Spearman 相关系数是在 Pearson 相关系数的基础上，利用两个集合中元素在各自集合的等级(排名)来计算他们之间的相关性，可以用于对数据进行分析。斯皮尔曼相关系数用来估计两个变量 X、Y 之间的相关性，其中变量间的相关性可以使用单调函数来描述。如果两个变量取值的两个集合中均不存在相同的两个元素，那么，当其中一个变量可以表示为另一个变量的很好的单调函数时（即两个变量的变化趋势相同），两个变量之间的相关系数可以达到+1 或-1。假设两个随机变量分别为 X、Y (也可以看作两个集合)，它们的元素个数均为 N，两个随机变量取的第 i (1 <= i <= N) 个值分别用 x_i 、 y_i 表示。对 X、Y 进行排序 (同时为升序或降序)，得到两个元素排行集合 x、y，其中元素 x_i 、 y_i 分别为 x_i 在 X 中的排行以及 y_i 在 Y 中的排行。将集合 x、y 中的元素对应相减得到一个排行差分集合 d，其中 $d_i=x_i-y_i$ ， $1 <= i <= N$ 。随机变量 X、Y 之间的斯皮尔曼相关系数可以由 x、y 或者 d 计算得到。

表二 两种相关的对比

系数	使用场景	备注
Pearson	定量数据，数据满足正态性	可用散点图查看数据关系
Spearman	定量数据，数据不满足正态性	可用散点图查看数据关系

ii) 选择相关性系数对进失球比与球队最终得分的相关性分析

通过 1、2 两节中画出的球队得分-进失球关系的散点图可以发现：不同赛季不同联赛的足球比赛的数据是定量数据，并且不满足正态性，故一般的基于线性相关性分析的 Pearson 相关系数不如 Spearman 相关性系数分析。故下面的分析选用 Spearman 相关性分析。

利用 matlab 的 corr 函数：可得相关性系数 $\rho = \text{corr}(\text{x}, \text{y}, \text{'type'}, \text{'Spearman'})$ 得到 15-16 赛季西班牙足球甲级联赛得分-进失球的相关性系数如下。

```
>> x=[91 90 88 64 62 60 52 48 48 45 44 44 43 43 42 39 39 38 36 32];
>> y=[112/29 110/34 63/18 44/35 58/45 51/59 51/50 38/35 45/48 34/52 45/53 46/48 40/74 49/61 45/61 46/69 40/62 52/73 37/67 37/70];
>> pearson=corr(x, y, 'type', 'Pearson'); spearman=corr(x, y, 'type', 'Spearman');
>> [pearson spearman]
```

ans =

0.9461	0.9014
--------	--------

图四 15-16 赛季西班牙足球甲级联赛相关性分析的例子

用相同的方法，可以发现几乎所有联赛的所有赛季，赛季得分与进失球比均存在一定关系，其 Pearson，Spearman 相关系数均在 0.9 附近，并且 Pearson 相关系数均略大于 Spearman 相关系数。到这里我们可以得出结论，赛季末球队的得分与球队进失球比很相关。在一支球队处于困难时，不要一心只想着进攻或者只想着防守，要将两者统一起来，也就是球迷口中的更具有“韧性”的球队。

参考文献

- [1] 徐全智，吕恕. 概率论与数理统计，高等教育出版社，2017
- [2] 尚敏，廖芬. 基于一元线性回归模型八字门滑坡累积位移预测，工程地质学报，2019
- [3] 舒林，万天军. 基于线性回归的数据处理方法研究，2019 年（第四届）中国航空科学技术大会论文集，2019
- [4] 易东，陈军，刘岭. 相关系数的正确理解和表达，重庆医学 2011 年 12 月第 40 卷第三十四期，2011

小概率事件在实际中的应用

李兆欣 (2018040705012)

指导教师: 吕恕

【摘要】 小概率事件在实际生活中有着非常广泛的应用背景, 本文深入浅出的剖析了小概率事件的原理和含义以及其在产品检测、保险业务和各类彩票中的应用。期望通过对生活中小例子的解释令读者对小概率事件有一个初步的认识。

【关键词】 小概率事件; 应用; 产品质量检测; 保险; 彩票

1 小概率事件的含义

小概率事件在生活中常有发生, 生产生活中的许多原理都离不开小概率事件。为了便于我们在下文的论述, 我们在这里给出其定义, 首先小概率事件是某一次试验中概率很接近于 0 的事件, 但是应注意, 小概率事件并不是不可能事件, 在一定条件下不可能发生的事件为不可能事件, 不可能事件的概率一定为 0, 但概率为 0 的事件未必就是不可能事件, 小概率事件并不代表不可能事件。由连续型随机变量的知识可知, 连续型随机变量取某个值的概率恒等于 0 (这可以看作小概率事件的极限情况), 但是这一事件仍旧是可能发生的, 比如向靶纸射击, 弹着点刚好落在靶纸边缘上的概率为零, 但其的确有发生的几率, 只是该几率较低。对于离散型随机变量也是如此, 在泊松分布中, 当试验次数 n 趋近于无穷时, 稀有事件发生的次数趋近于 $np_n(\lambda)$, 其中 p 代表一次独立重复试验中稀有事件发生的次数, 且由

$$\lim_{n \rightarrow \infty} np_n = \lambda \quad (\lambda > 0) \quad (1)$$

可知 p_n 是 $1/n$ 的同阶无穷小, 即 p_n 接近于 0, 但是由泊松分布可知, 当 n 很大时, 该稀有事件的确可能发生。

根据贝努力大数定律: 设 $\frac{m}{n}$ 是 n 次独立重复试验中事件 A 发生的频率, p 是事件 A 在每次试验中发生的概率, 则对于 $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{m}{n} - p \right| < \varepsilon \right\} = 1 \quad (2)$$

即 A 发生的频率依概率收敛于事件 A 的概率。当 A 为小概率事件时, p 非常小, 则 A 发生的频率也极可能很小, 从而 A 发生的次数也极可能很小。日常生活中往往无法做很多次的试验以获得较大的样本数据量, 所以常常忽略这些概率很小的随机事件发生的可能性, 在一次试验中将其看作不可能事件。

2 小概率事件在产品质量检测中的应用

对于稳定连续生产的生产线, 其产品质量 X 服从正态分布, 即 $X \sim N(\mu, \sigma^2)$, 其中

$$P\{|X-\mu| \leq \sigma\} = 0.6286$$

$$P\{|X-\mu| \leq 2\sigma\} = 0.9544$$

$$P\{|X-\mu| \leq 3\sigma\} = 0.9974$$

可以认为事件 $|X-\mu| \geq 3\sigma$ 为小概率事件, 其概率为 $p=0.0026$, 所以说对于生产车间的控制图, 其产品质量特性值的上下限 UCL 和 LCL 应分别设定为 $\mu+3\sigma$ 和 $\mu-3\sigma$, 当探测器检测到产品质量特性值超过这个范围, 即小概率事件发生了, 那么我们可以认为生产线有很大的可能出现了故障, 应该停车检修生产线, 以防止更大的损失产生。

3 小概率事件在保险业务中的应用

保险公司业务范围主要为小概率且一旦发生会对客户有较大影响的事件，比如车祸，重大疾病等等。据数据统计，中国交通事故的死亡人数约为 11 万人，如果除去极偏远地区和高龄老人，按照参与交通人数为 11 亿人计算，则交通事故死亡率约为 0.01%，目前交通意外险的保费约为每人每年 200 元，一旦投保人因交通事故意外死亡则可获赔 50 万，我们认为每个人出交通意外是彼此独立的， X 可近似看作服从 $\lambda=1$ 的泊松分布 ($X \sim P(1)$) 则可以由此计算保险公司获利的概率。假设有 10000 人投保，保险公司一年的收入为 200 万元，设 X =当年出险人数，若使其亏本，则 $50X > 200$ ，即 $X > 4$ ，其概率为

$$P\{X > 4\} = \sum_{k=5}^{+\infty} \frac{e^{-1}}{(k)!} \approx 0.0061$$

所以某一年度公司亏本对与该保险公司来说可能性极小，可近似看作不可能事件，所以即使赔付的额度很高，公司也几乎不可能亏损，而高额的赔付又吸引了更多的顾客投保，进一步提升了保险公司的利润，这便是保险类业务的营业模式。

4 小概率事件在彩票中的应用

各类彩票在生活中可谓层出不穷，其总是以其头奖高额的奖金为诱惑吸引人们购买，下面我们以双色球为例介绍小概率事件在彩票中的应用。已知购买双色球中头奖的概率为约 1772 万分之一，这显然是一个小概率事件，但是当购买的人说相当多时（实际中正是如此），则中奖人数近似服从泊松分布。设 X =一次开奖时中头奖的彩票号码数，若某次购买双色球的彩民投注总数为一千万注，则 $X \sim P(1.772)$ ，其期望 $E(X)=1.772$ ，由独立同分布大数定律：

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

当开奖次数足够多的时，每次的中奖号码数依概率收敛于其期望值 1.772，故每次开奖时经常可以看到有人中到头奖，而数百万甚至上千万的头奖人们来说非常有诱惑力，所以会有更多的彩民去购买彩票，这又进一步增大了中奖号码数的数学期望值，但是对我们每个人而言，夺得头奖的概率小之又小，可近似看作不可能事件，可见，希望依靠彩票而一夜暴富的想法是几乎不可能实现的。

5 总结

小概率事件是概率论中的一个简单的概念，但是在生产生活中有着广泛的应用，本文仅举出三个拙例以带读者认识小概率事件的含义及其在生活中的实际应用，小概率事件的原理告诉我们在面对极大的样本容量时，不要被某一事件出现的次数蒙蔽了双眼，不要盲目投资，不要过度自信，要关注实验结果背后的概率。也要学会利用小概率事件在小样本容量中的几乎不可能性来检测系统故障。掌握并学会科学地利用小概率事件是每个知识分子都应具备的基本科学素养。

参考文献

- [1] 徐全智、吕怒.概率论与数理统计
- [2] 张洁,杨文国.浅谈小概率事件的内涵[J].科技风,2018(25):74-76.
- [3] 王德斌.小概率在生活中的几个简单应用
- [4] 于雪梅.小概率事件特点、原理及其应用[J].科技风,2017(16):256-261.

一元线性回归中 F 检验法与相关系数检验法的联系

刘华烁 (2018100403007)

指导教师：吕恕

【摘要】 从统计量判断、数学用表两个方面论证了一元线性回归中 F 检验法与相关系数检验法在本质上的联系，并得出结论与总结。

【关键词】 F 检验法；相关系数检验法；一元线性回归；假设检验；F 分布表；相关系数临界值表

对任何的两个变量 X 和 Y，无论它们之间是否存在线性相关关系，我们都可以根据实验数据 (x_i, y_i) ($i=1, 2, \dots, n$)，用最小二乘法写出一个线性回归方程。但是如果 X 与 Y 的相关关系很弱，那么求出的回归方程是没有价值的。因此我们需要对回归变量 X 和 Y 的线性相关性进行检验。常用的方法是 F 检验法和相关系数检验法。下面将对两种方法的联系加以讨论。

1 统计量判断结果的一致性

F 检验法的基本思路是：定义 Q_T 为总离差平方和，反映变量 Y 的观测值的总的离散强度， Q_E 为残差平方和， Q_R 为回归离差平方和，就有：

$$Q_T = \sum_{i=1}^n (y_i - \bar{y})^2 = l_{yy} \quad (1)$$

$$Q_R = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \hat{b}^2 l_{xx} \quad (2)$$

$$Q_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = l_{yy} - \hat{b}^2 l_{xx} \quad (3)$$

$$Q_T = Q_E + Q_R \quad (4)$$

于是对回归系数 \hat{b} 进行检验，提出原假设：

$$H_0: b = 0 \quad (5)$$

对于给定的显著性水平 α ，如果拒绝 H_0 ，则自变量 X 与因变量 Y 之间线性相关关系显著；如果接受 H_0 ，所选定的自变量 X 与因变量 Y 之间的线性相关关系不明显。

当假设 H_0 成立时，统计量

$$F = \frac{Q_R}{Q_E/(n-2)} \sim F(1, n-2) \quad (6)$$

原假设 H_0 的拒绝域为 $F > F_{\alpha}(1, n-2)$ 。

而所谓的相关系数检验法就是通过其矩估计量样本相关系数 R 进行检测

$$R = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{l_{xy}}{\sqrt{l_{xx}} \sqrt{l_{yy}}} \quad (7)$$

对于给出的显著性水平 α ，当 $|R| > R_\alpha(n-2)$ 时，认为 X 与 Y 的线性关系显著，否则认为线性关系不显著。

根据上面推算出来的 R 、 Q_R 、 Q_E 的定义式可以发现：

$$\frac{R^2}{1-R^2} = \frac{Q_R}{Q_E}; \quad F = \frac{Q_R}{Q_E / (n-2)}; \quad (8)$$

整理消去 Q_R 与 Q_E ：

$$\frac{(n-2) R^2}{1-R^2} = F; \quad (9)$$

所以由上式与线性相关关系显著性我们可以推出：

R 增大 \Leftrightarrow 显著性增加

F 增大 \Leftrightarrow 显著性增加

R 增大 $\Leftrightarrow F$ 增大

得证：虽然 F 检验法与相关系数检验法的具体做法不同，但最后判断的结果是一致的；换言之，具体做法不影响检验结果的一致性。

2 相关系数临界值表与 F 分布表的联系

F 检验法和相关系数检验法使用了不同的检验表，分别是 F 分布表和相关系数临界值表。

在 F 检验法中，根据给定的显著性水平 α 和自由度 $(n-2)$ ，可以在 F 分布表中查到相对应的值 $F_\alpha(1, n-2)$ ，用字母 f 来表示：

$$P(F > f) = \alpha, \quad (10)$$

当 F 的观测值大于 f 的时候，认为 X 与 Y 的线性关系显著。相关系数检验法同理，当 $R > R_{0.01}(n-2)$ 时，认为 X 与 Y 的线性关系显著。

因，

$$\frac{(n-2) R^2}{1-R^2} = F, \quad (11)$$

故，

$$\frac{(n-2) R^2}{1-R^2} > f \quad (12)$$

经过计算得

$$|R| > \sqrt{\frac{1}{1 + \frac{(n-2)}{f}}} = r \quad (13)$$

下面分别以显著性系数 $\alpha = 0.05$ 和 $\alpha = 0.01$, 第一自由度为 1, 第二自由度 $(n-2) = 1, 2, 3, 4, 5$ 为例, 通过查找相关系数临界值和 F 检验表, 在表格中列出 f 、 r 、 $R_\alpha(n-2)$ 的值:

$\alpha = 0.05$

$n-2$	1	2	3	4	5
f	161.4	18.51	10.13	7.71	6.61
r	0.9969	0.9500	0.8784	0.8114	0.7545
$R_{0.05}$	0.997	0.950	0.878	0.811	0.754

$\alpha = 0.01$

$n-2$	1	2	3	4	5
f	4052	98.50	34.12	21.20	16.26
r	0.9999	0.9900	0.9587	0.9172	0.8745
$R_{0.05}$	1.000	0.990	0.959	0.917	0.847

由表可得,

$$r = R_\alpha \quad (14)$$

在任何情况下, 都存在:

$$F > F_\alpha(1, n-2) \Leftrightarrow R > R_\alpha(n-2) \quad (15)$$

因此, 在任何情况下, 无论使用相关系数检验法, 还是使用 F 检验法, 显著性检验的结果一定是相同的。

3 结论与总结

综合上述两个方面的讨论结果, 我们得出结论:

F 检验法与相关系数检验法从本质上讲是一致的, 都是通过对方差进行分析来求出因变量与自变量的线性相关性, 所以其检验结果也必然是一致的。二者的区别在于, F 检验法的用途更广, 并不仅限于线性回归的显著性检验, 但 F 检验法依赖于 F 分布表, 在使用时需要考虑第一、第二自由度, 所以过程相对比较繁琐; 而相关系数检验法是一种专门用于检验线性相关关系的方法, 而且有属于他自己的“相关关系显著性检验相关系数临界值表”, 不需要考虑自由度这一问题, 所以在处理相关问题时, 相关系数检验法要比 F 检验法好用得多。

参考文献

- [1] 徐全智, 吕恕. 概率论与数理统计 (第 3 版) [M]. 高等教育出版社, 2017.
- [2] 周玉珠, 姜奉华. 实验数据的一元线性回归分析及其显著性检验 [J]. 大学物理实验, 2001(04): 43-46.
- [3] 徐安. 参数估计与假设检验的最小二乘估计——相关系数检验方法 [J]. 济南交通高等专科学校学报, 2001(03): 1-5.
- [4] 曹俊忠. 一元线性回归显著性检验方法分析 [J]. 西安工程科技学院学报, 1988(Z1): 78-82.

二项分布对正态分布的逼近模拟分析

卢舒云 (2018051405008)

指导教师：吕恕

【摘要】 当实验次数较大时，二项分布将会趋近于正态分布，在将二项分布的分布函数与正态分布的分布函数比较的过程中，得到了二项分布的趋近程度随着试验次数的增加而变好的结果，验证了棣莫弗-拉普拉斯中心极限定理的正确性。

【关键词】 二项分布；正态分布；棣莫弗-拉普拉斯中心极限定理；离差均值

1 引言

若一次试验的结果只有两种情况，A 发生或 A 不发生，则该试验称为贝努利试验。若 A 发生的概率为 p，则 n 次重复独立贝努利试验中事件 A 发生的次数 X 服从二项分布 $X \sim B(n, p)$ 。由棣莫弗-拉普拉斯中心极限定理可知，当 n 很大时，X 可以认为近似服从正态分布，本文通过 matlab 的仿真模拟计算，对 n, p 取不同值时计算标准化后的二项分布的分布函数，并与标准正态分布曲线进行比较，从而验证棣莫弗-拉普拉斯中心极限定理的正确性。

2 棣莫弗-拉普拉斯中心极限定理

棣莫弗-拉普拉斯中心极限定理是独立同分布中心极限定理的一个推论，假设 $X_1, X_2, \dots, X_n, \dots$ 是相互独立的随机变量序列，由于二项分布满足独立同分布的条件，根据独立同分布中心极限定理的公式

$$\lim_{n \rightarrow \infty} P\left\{ \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \leq x \right\} = \Phi(x) \quad (1)$$

其中 μ, σ^2 分别是随机变量 X_i 的期望和方差，假设 X_i 服从两点分布 $X \sim B(1, p)$ ，则 $\mu=p$, $\sigma^2=p(1-p)$ ，由二项分布可加性可知 $Y_n = \sum_{i=1}^n X_i \sim B(n, p)$ ，式 (1) 可改写为

$$\lim_{n \rightarrow \infty} P\left\{ \frac{Y_n - np}{\sqrt{np(1-p)}} \leq x \right\} = \Phi(x) \quad (2)$$

此式就是棣莫弗-拉普拉斯中心极限定理，此式表明在 n 足够大的情况下，服从二项分布的随机变量可以近似看作服从正态分布。

3 参数 p 为 0.5 的情形——抛硬币问题

抛硬币问题可以看作是一种简单的二项分布问题，假设 X 为抛 n 次硬币时正面向上的硬币数，则 $X \sim B(n, 0.5)$ ，我们取 $n=10$ ，得到一个抛 10 次硬币的场景。根据二项分布概率的计算公式 $P\{X=k\} = C_n^k p^k (1-p)^{n-k}$ ，我们取 $X=0, 1, \dots, 10$ 计算 X 的分布律，如表 1 中所示。

表1 $X \sim B(10, 0.5)$ 的分布律

X	0	1	2	3	4
$P\{X=x_i\}$	0.0010	0.0098	0.0439	0.1172	0.2051
5	6	7	8	9	10
0.2461	0.2051	0.1172	0.0439	0.0098	0.0010

由公式 (2) 可知, 当 n 非常大时, 随机变量 $U = \frac{X - np}{\sqrt{np(1-p)}}$ 趋近于标准正态分布, 我们称 U 为标准化的二项分布随机变量, $n=10$, $p=0.5$ 时, U 的分布律如表 2 所示。

表2 $n=10$, $p=0.5$ 时, U 的分布律

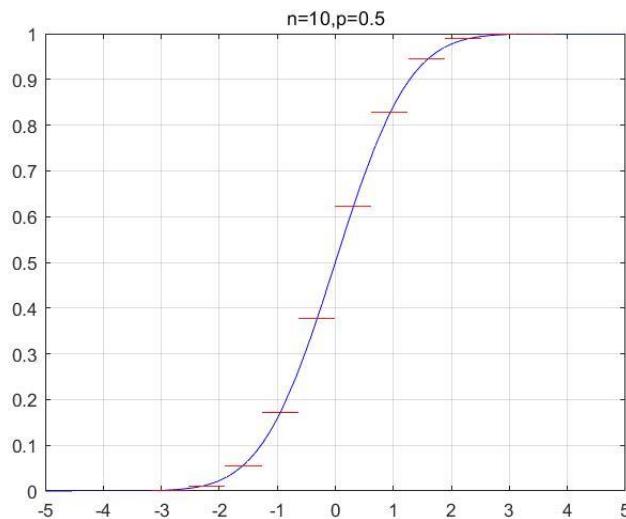
U	-3.16	-2.53	-1.90	-1.26	-0.63
$P\{U=u_i\}$	0.0010	0.0098	0.0439	0.1172	0.2051
0	0.63	1.26	1.90	2.53	3.16
0.2461	0.2051	0.1172	0.0439	0.0098	0.0010

由于分布律是离散的, 无法与概率密度函数进行直接比较, 我们接下来的比较分析都会基于标准化后的二项分布的分布函数与标准正态分布的分布函数。

根据分布函数的定义 $F(u) = P\{U \leq u\}$, 我们得到 $n=10$, $p=0.5$ 时, 随机变量 U 的分布函数

$$F(u) = \begin{cases} 0, & u < -3.16 \\ 0.0010, & -3.16 \leq u < -2.53 \\ \dots & \dots \\ 1, & u \geq 3.16 \end{cases} \quad (3)$$

我们将 U 的分布函数 $F(u)$ 与标准正态分布的分布函数 $\Phi(u)$ 放在同一张图里对比, 得到如下结果, 其中一系列与 u 轴平行的红色直线段表示 $F(u)$ 的图像, 蓝色曲线表示 $\Phi(u)$ 的图像

图1 二项分布函数与正态分布函数的对比($n=10$, $p=0.5$)

可见, 当 $n=10$, $p=0.5$ 时, $F(u)$ 中的每一段直线都位于 $\Phi(u)$ 的曲线上, 如果我们把二项分布函数中每一段直线近似看作是一个点, 我们取直线段中点坐标为其坐标, 则这一系列点的坐标就非常的接近于在正态分布函数的曲线上。为了方便描述二项分布对正态分布的趋近程度, 我们做如下

定义：

定义 1 若假设 u_i 、 u_i^* 分别是第 i 段线段的中点横坐标和其所在直线与正态分布的分布函数交点的横坐标，其中 $i=1,2,\dots,n$ ，则称 $\varepsilon_i = |u_i - u_i^*|$ 第 i 段线段的离差。

定义 2 $\bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^n \varepsilon_i$ 称为离差均值， $\bar{\varepsilon}$ 可以用来刻画二项分布与标准正态分布的趋近程度。由于 $\bar{\varepsilon}$ 值计算复杂，本文只做定性分析，而不做定量计算。

接下来分别取 $n=36$, $n=100$, p 不变，计算随机变量 U 的分布律，得到的 $F(u)$ 与 $\Phi(u)$ 进行对比，得到如下结果

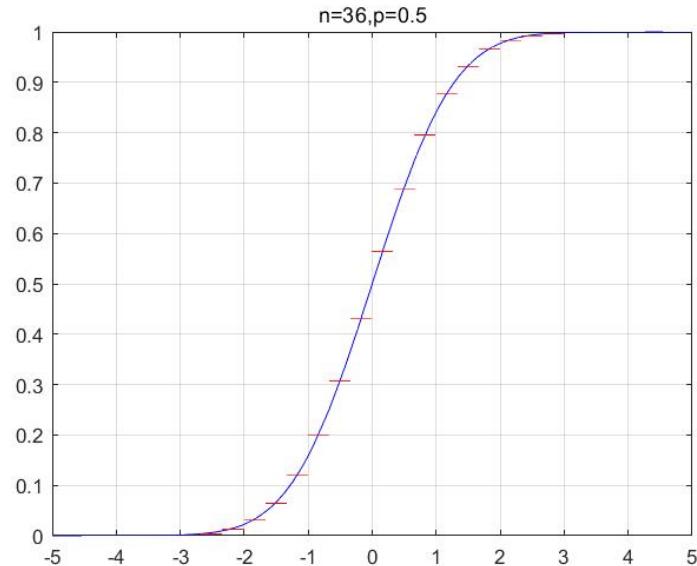


图 2 二项分布函数与正态分布函数的对比($n=36$, $p=0.5$)

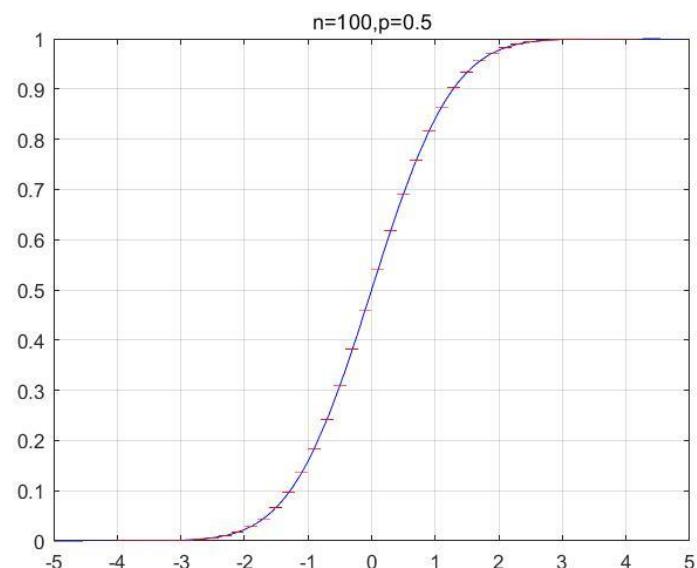


图 3 二项分布函数与正态分布函数的对比($n=100$, $p=0.5$)

根据对图像的观察，当 $p=0.5$ 时，随着 n 的增大，每一段线段的离差值 ε_i 越来越小，故 $\bar{\varepsilon}$ 值也越来越小，二项分布对正态分布的趋近程度越来越好。

4 参数 p 不为 0.5 的情形

上述分析只针对了 $p=0.5$ 时的特殊情况，如果 p 取其他值时情况会不会发生改变呢？我们取 $p=0.25$ 时，重复一遍上述的分析方法，分别取 $n=10$, $n=36$, $n=100$ ，得到随机变量 U ，再画出 $F(u)$ 与 $\Phi(u)$ 比较，结果如下

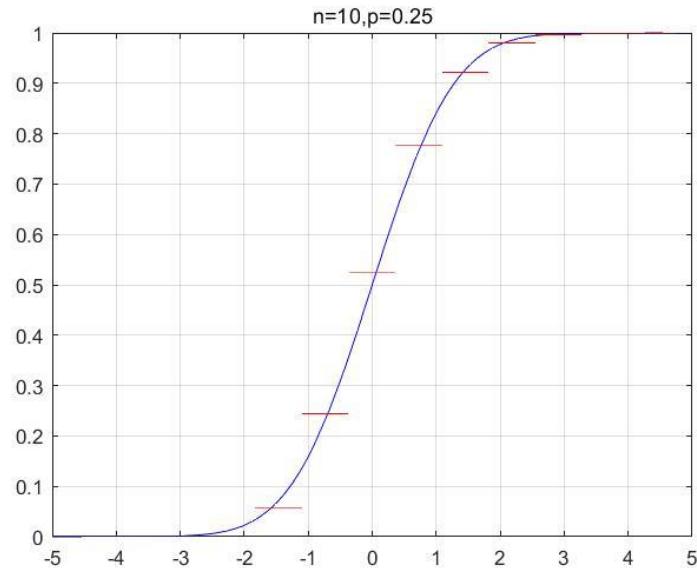


图 4 二项分布函数与正态分布函数的对比($n=10$, $p=0.25$)

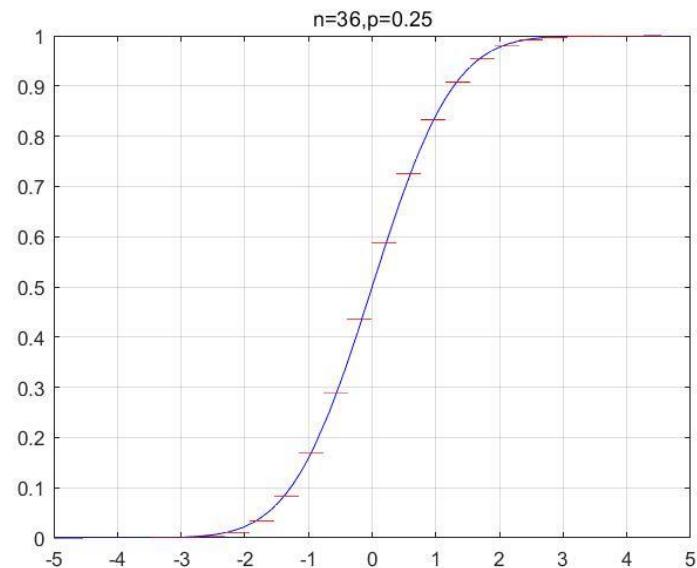


图 5 二项分布函数与正态分布函数的对比($n=36$, $p=0.25$)

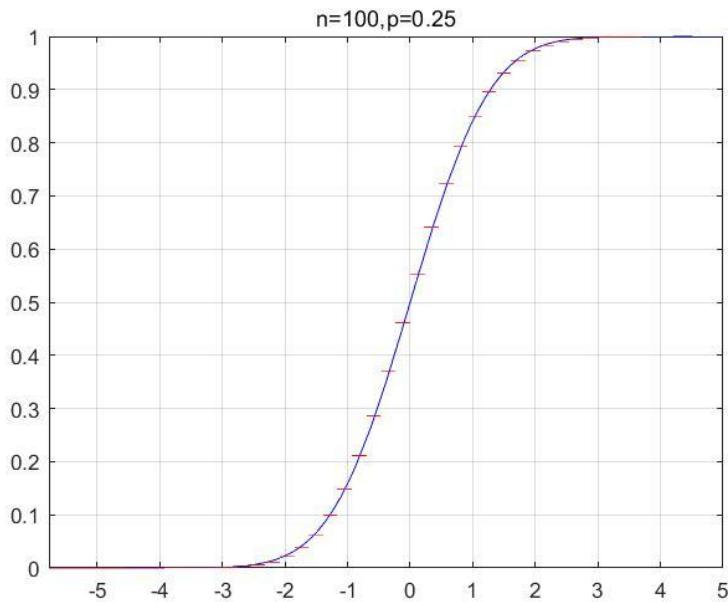


图 6 二项分布函数与正态分布函数的对比($n=100$, $p=0.25$)

根据对图像的观察，随着 n 的增大，每一段线段的离差值 ε_i 逐渐减小，离差均值 $\bar{\varepsilon}$ 也随之减小，趋近情况越来越好，与 $p=0.5$ 的情况横向对比，同为 $n=10, 36, 100$ 时， $p=0.25$ 的情况下离差均值 $\bar{\varepsilon}$ 要比 $p=0.5$ 的情况下更大，趋近效果更差。

笔者还对 $p=0.1, p=0.75, p=0.9$ 的情况进行了绘图分析，由于篇幅有限，这里只给出结果：不论 p 取何值时，只要随着 n 的增大，离差均值 $\bar{\varepsilon}$ 都会随之减小，而在 n 值相同的情况下，随着 p 值与 0.5 的差距的增大，离差均值 $\bar{\varepsilon}$ 也随之增大。

5 总结

本文以离差均值 $\bar{\varepsilon}$ 为定性指标，对于随机变量服从不同 n, p 值的二项分布时导出的标准化后的分布函数与标准正态分布函数进行比较，得到如下结论：

- ① 对于固定的 p 值， n 值越大，二项分布对正态分布的趋近效果越好，验证了棣莫弗-拉普拉斯中心极限定理的正确性；
- ② 当 n 值不变时， $|p-0.5|$ 越小，二项分布对正态分布的趋近效果越好，可以猜测，当 p 的取值为 0.5 时，二项分布对正态分布的趋近速率最快。

可见，对于所有的服从二项分布的随机变量，随着试验次数 n 的增加，最终都会趋近于正态分布，但是对于不同的事件发生概率 p ，二项分布对正态分布的趋近速率有所不同。

参考文献

- [1] 徐全智,吕恕.概率论与数理统计.3 版.北京:高等教育出版社,2017
- [2] 杜鸿飞,陈绍刚.二项分布近似计算的误差量化分析[J].大学数学,2019,35(04):108-114.

一元非线性回归问题的线性化处理

王子涵 (2018060401010)

指导教师：吕恕

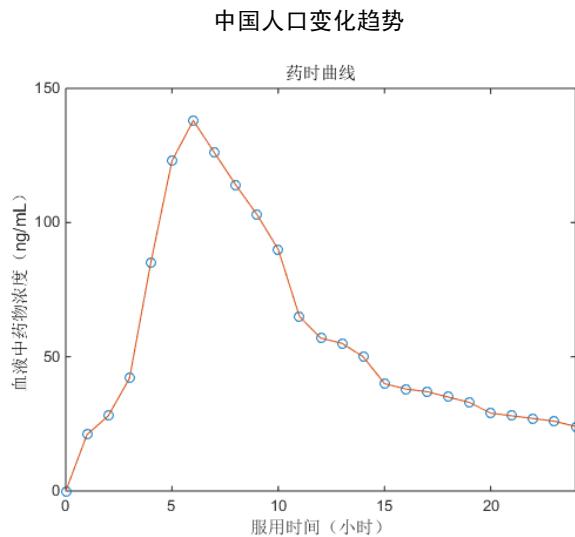
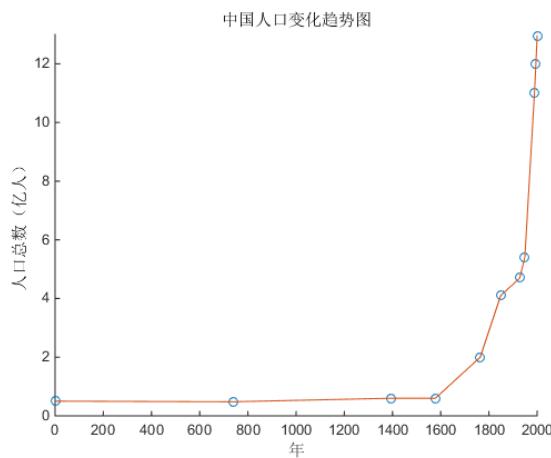
【摘要】本文主要讨论一元非线性回归问题的线性化处理，以对呈非线性相关关系的变量进行科学预测，列出了6种常见的典型曲线，以这6种典型曲线为基础，结合变量替换的方法，展开对一元非线性回归问题线性化处理方法的讨论，并结合线性回归分析的方法，给出了评判回归方程拟合优度的有关指标，以确定最优拟合方式。

【关键词】一元非线性回归；典型曲线；拟合优度；最优拟合方式

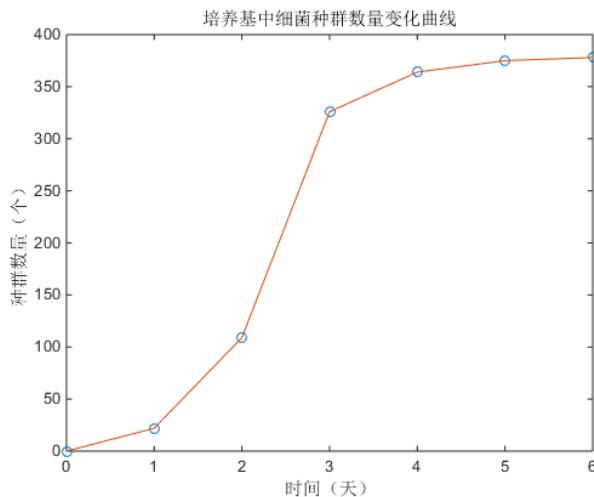
1. 引言和引例

在许多实际问题中，变量之间存在的相关关系不一定都是线性的，可能存在着某种复杂的非线性的相关关系。本文将对一元非线性回归问题进行线性化处理，并对线性化处理后的回归方程的拟合优度提出比较与评优的指标。

下面给出有关非线性相关问题的具体实例：



口服药物后血液中药物浓度随服用时间的变化（药时曲线）



培养基中细菌种群数量的变化趋势

2 用变量替换将一元非线性回归问题线性化

由上述实例可以看出，在日常生活和生产活动中，变量之间的非线性相关关系无处不在。

分析变量之间相关关系的类型，一般有两条途径：一是根据所讨论问题的专业知识或以往积累的经验确定；二是对于仅有一个自变量的一元回归问题，可根据实际观测数据描出的散点图确定。

下面给出常见的几种典型曲线方程及其图形，以便于在实际应用时进行选择。

1) 双曲线函数

函数形式为

$$(2.1.1) \quad \frac{1}{Y} = a + b \cdot \frac{1}{x}$$

$$\text{令 } Z = \frac{1}{Y}, \omega = \frac{1}{X}$$

则可得到

$$(2.1.2) \quad Z = a + b\omega$$

其曲线图形如图 2.1 所示

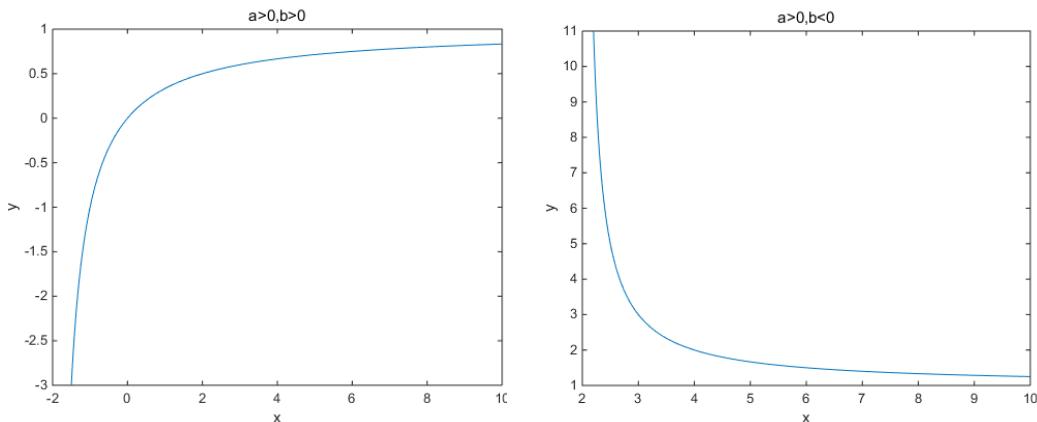


图 2.1

2) 对数曲线函数

函数形式为

$$Y = a + b \cdot \ln x \quad (2.2.1)$$

令 $\omega = \ln x$

则可得到

$$Y = a + b \cdot \omega \quad (2.2.2)$$

其曲线图形如图 2.2 所示

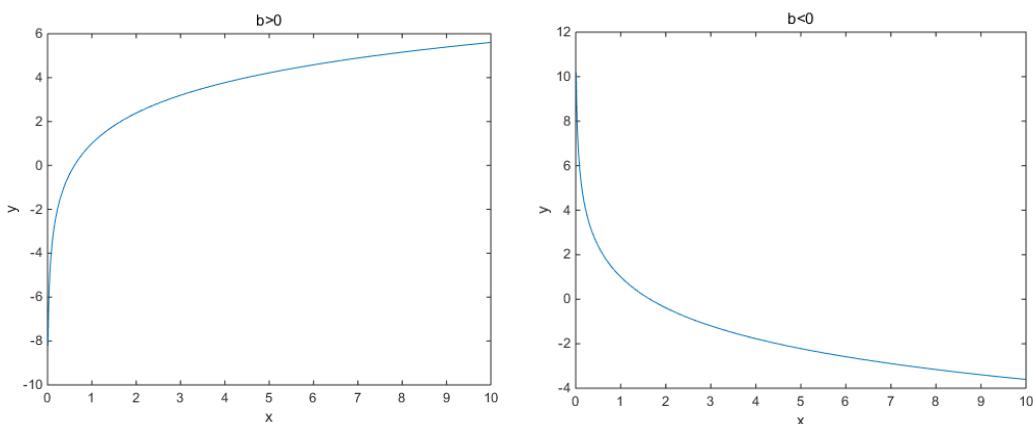


图 2.2

3) 幂函数

函数形式为

$$Y = a \cdot x^b \quad (2.3.1)$$

令 $Z = \ln Y \quad \nu = \ln a \quad \omega = \ln x$

则可得到

$$Z = \nu + b \cdot \omega \quad (2.3.2)$$

其曲线图形如图 2.3 所示

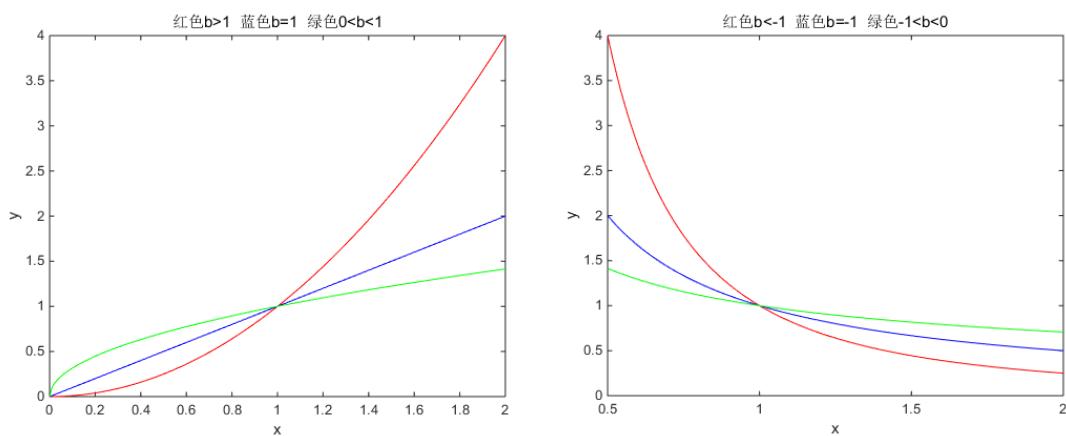


图 2.3

4) S 形曲线函数

函数形式为

$$Y = \frac{1}{a + b \cdot e^{-x}} \quad (2.4.1)$$

令 $Z = \frac{1}{Y}$ $\omega = e^{-x}$
则可得到

$$Z = v + b \cdot \omega \quad (2.4.2)$$

其曲线图形如图 2.4 所示

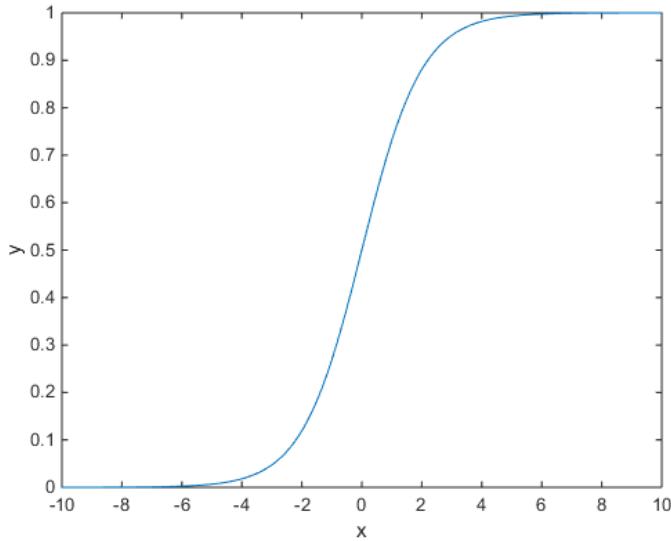


图 2.4

5) 指数函数

函数形式为

$$Y = a \cdot e^{bx} \quad (2.5.1)$$

令 $Z = \ln Y$ $v = \ln a$ $\omega = x$

则可得到

$$Z = v + b \cdot \omega \quad (2.5.2)$$

其曲线图形如图 2.5 所示

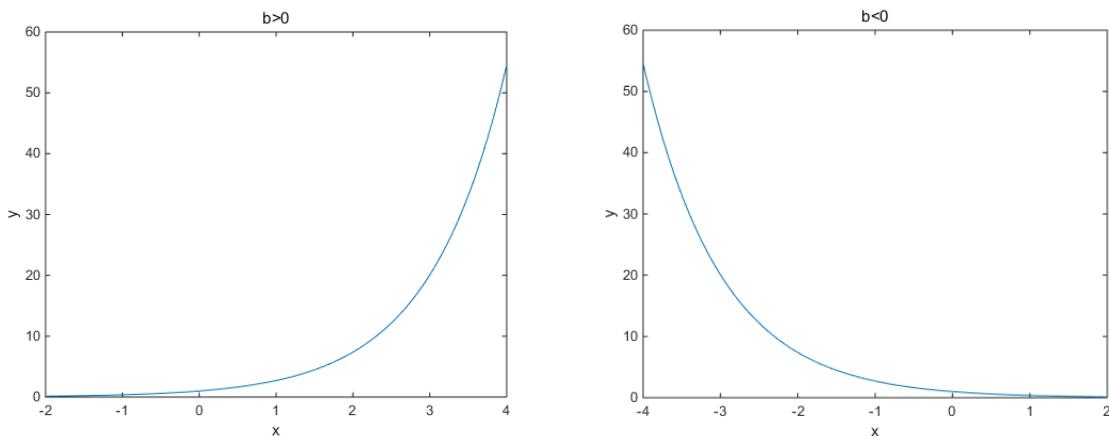


图 2.5

6) 倒数指数函数

函数形式为

$$Y = a \cdot e^{\frac{b}{x}} \quad (2.6.1)$$

$$\text{令 } Z = \ln Y \quad \nu = \ln a \quad \omega = \frac{1}{x}$$

则可得到

$$Z = \nu + b \cdot \omega \quad (2.6.2)$$

其曲线图形如图 2.6 所示

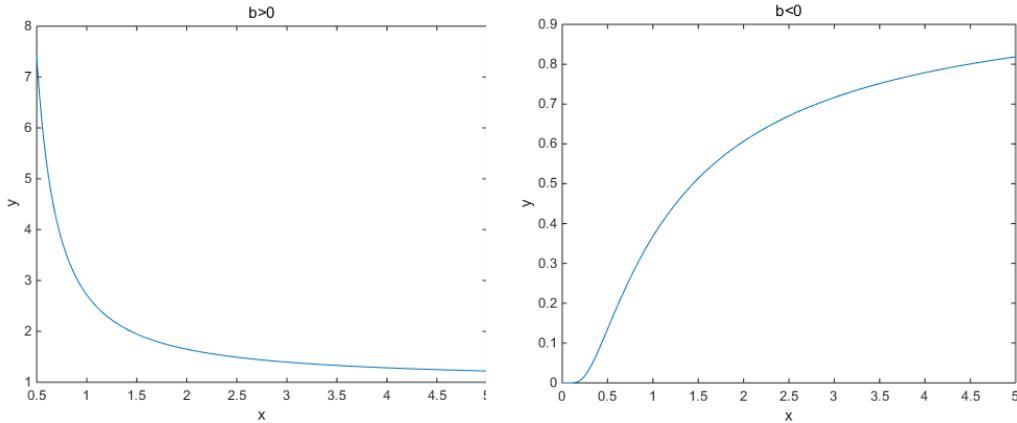


图 2.6

对回归方程选择一种合适的函数形式，必须事先对散点图进行认真的分析，此时对上述各种函数类型所描述的曲线形状有充分的了解是十分必要的。

3 不同回归方程的拟合优度的比较与评优指标

一般来说，对同一种散点图所呈现的 Y 与 x 之间的关系，可以选择不同的函数形式来描述回归方程。那么如何判断并比较不同回归方程的拟合程度呢？

通常使用的比较指标有下面两个：

1) 相关指数 R

R 的定义式为

$$R = \sqrt{1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3.1)$$

上式与一元线性回归模型中定义的样本相关系数是一致的。

显然，如果回归曲线对样本点拟合优度好，那么 R 值较大，因此对于不同的回归曲线方程通常选择 R 值较大的一个。

2) 剩余标准差 S

S 的定义式为

$$S = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-2}} \quad (3.2)$$

剩余标准差 S 反映了样本点偏离回归曲线的平均大小，显然 S 值越小，回归曲线方程的拟合优

度越好。

从相关指数 R 和剩余标准差 S 的定义式可以看出，上述两个比较准则是一致的，即 R 越大，则 S 越小；反之亦然。

4 应用一元非线性回归问题的线性化处理对实际问题进行分析

经过调查，得到 8 个厂家对同种类型的产品年新增加投资额和年利润额的数据资料，如下表所示。

厂家	1	2	3	4	5	6	7	8
年新增投资额 x (万元)	4	6	10	11	15	17	18	20
年利润额 Y (万元)	6	7	9	10	17	24	23	26
lnY	1.79	1.95	2.20	2.30	2.83	3.18	3.14	3.26

Y 关于 x 的散点图如图 4.1 所示

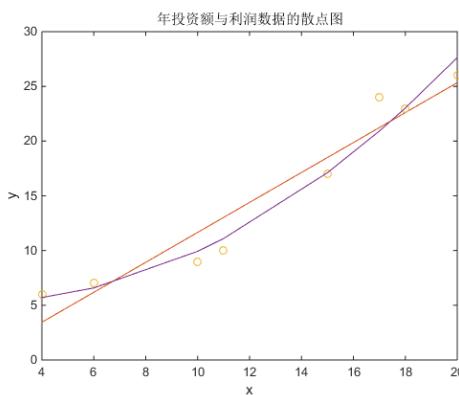


图 4.1

lnY 关于 x 的散点图如图 4.2 所示

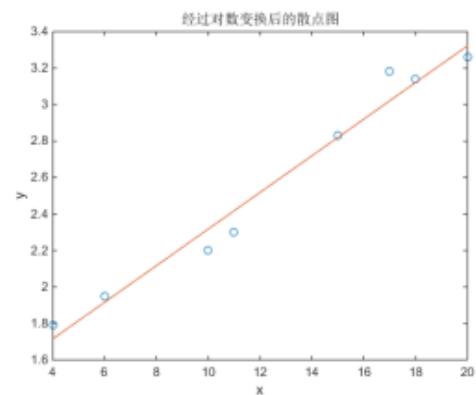


图 4.2

图 4.1 给出了年利润额 Y 与年新增加投资额 x 的散点图，从图中可以清楚地看出来，随着 x 的增大，Y 也有明显的增加趋势，因此两者之间存在着相关关系，但是这种相关关系与其用一条直线来描述倒不如用曲线描述更为合适，因此 Y 与 x 之间更加倾向于被认为是一种非线性关系。回归方程也需要用一些非线性函数来刻画，比如

$$\hat{Y} = \hat{\beta}_0 \cdot e^{\hat{\beta}_1 x} \quad (4.1)$$

或者

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot x^2 \quad (4.2)$$

等等。

图 4.2 给出的是变量 lnY 与变量 x 的散点图，从中可以看出这些点基本上是围绕一条直线波动，说明变量 lnY 与 x 之间近似是一种线性关系，从而也印证了回归方程取 (4.1) 形式的合理性。

①一元线性回归方程

代入表格中已知数据，易求得 $\hat{b} = 1.37$ $\hat{a} = -2.02$

所以求得回归方程

$$\hat{Y} = -2.02 + 1.37x \quad (4.3)$$

②指数曲线拟合

首先对样本数据 $(x_i, Y_i) (i=1, 2, \dots, n)$ 做对数变换

$$Z_i = \ln Y_i, i = 1, 2, \dots, n \quad (4.4)$$

然后利用最小二乘法求出变量 Z 对 x 的回归方程

$$\hat{Z} = \hat{\alpha}_0 + \hat{\alpha}_1 \cdot x \quad (4.5)$$

结合式 (4.1) 可知 $\hat{\beta}_0 = e^{\hat{\alpha}_0}$, $\hat{\beta}_1 = \hat{\alpha}_1$

根据表中数据可以算得 $\ln Y$ 对 x 的线性回归方程为

$$\hat{Z} = 1.3139 + 0.1003x \quad (4.6)$$

因此, Y 对 x 的回归方程为

$$\hat{Y} = 3.7208 \cdot e^{0.1003x} \quad (4.7)$$

③抛物线 (幂函数曲线) 拟合

若采用形如式 (4.2) 的抛物线形回归方程, 做变量替换 $\omega = x^2$, 则式 (4.2) 可表示为变量 Y 对 ω 的线性回归方程

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \omega \quad (4.8)$$

利用表格中的数据可求出式 (4.8) 中的系数估计值为

$$\hat{\beta}_0 = 4.413, \hat{\beta}_1 = 0.057$$

所以 Y 对 x 的回归方程为

$$\hat{Y} = 4.413 + 0.057x^2 \quad (4.9)$$

④对上述三种拟合方法的拟合优度进行比较

根据式 (3.1) 和式 (3.2) 的定义式可求得

曲线形式	回归方程	相关指数 R	剩余标准差 S
直线	$\hat{Y} = -2.02 + 1.37x$	0.9637	2.3762
指数曲线	$\hat{Y} = 3.7208 \cdot e^{0.1003x}$	0.9804	1.7522
抛物线	$\hat{Y} = 4.413 + 0.057x^2$	0.9842	1.5746

因此, 针对本例而言, 相比之下抛物线形曲线方程拟合优度最好, 是最可取的。

参考文献

- [1] 徐全智, 吕怒. 概率论与数理统计 (第三版), 高等教育出版社, 2017
- [2] 龙永红. 概率论与数理统计 (第三版), 高等教育出版社, 2009

资深赌徒脑中的概率应用

杨博 (2018100403020)

指导教师：吕恕

【摘要】 在学习了概率论之后，我们已经具备了将概率论应用到日常生活中的所有理论知识，本文将就概率在赌博方面的应用，分析一位能在零和游戏甚至是负收益游戏里赚的盆满钵满的赌徒是怎样应用概率论知识的。本文将从利用概率论克服直觉陷阱，利用概率论计算最优下注方案以及针对小概率事件的最优应对。

【关键词】 克服直觉陷阱；最优化资金分配方案；小概率事件；切比雪夫大数定理；伯努利大数定律；海恩法则

1 概率论与直觉陷阱

当人们面临一些选择的时候，往往会依赖于直觉来给出自己的答案，但是对于一位资深赌徒来说，如果他不能放弃自己的直觉，那么他一定不能满载而归。比如：对于一台很多次没有中奖的老虎机来说，人们的直觉是远离它，因为怕自己沾上霉运。但是我们知道由全概率公式：

$$P(B) = \sum_{i=1}^n P(A_i)P(B|A_i) \quad (1)$$

之前未中奖的次数恰恰是增大了下一次中奖的概率，所以这个时候正是我们中奖概率较大的时候。

但是这对于一个资深赌徒来说还远远不够，因为当你选取了错误的对象的时候，有时全概率公式会得到错误的结论。比如：当赌场的很多台老虎机都一天无所获得时候，由上文可知我们会选择去其中一台进行下注，在成功中奖之后，如果我们凭借着直觉，以自身为参考对象应用全概率公式，那么可以得出在我们身上已经发生了一次小概率事件，直觉会告诉你该收手了，这实际上是直觉带来的又一次骗局；这个时候我们在应用全概率公式时应该选取的参考对象应该是不同的老虎机而不是自身，因为不同老虎机的中奖之间相互独立，所以对于很多台老虎机都没中奖的情况下，即使你在这台老虎机上中奖了，另一台老虎机上中奖的概率其实依然很高，所以这个时候该做出的选择是换一台老虎机继续将大奖拿走。

2 最优化资金分配方案

当我们刚进入赌场之时，常常会在各种各样收益不同，中奖几率也不相同的项目之间来回徘徊，犹豫自己该怎样做才能获得最大的回报，下面本文将详细论述：

假设赌客的本金为 N ，投注比例为 f ，游戏每局有 n 种结果，第 i 种结果的净收益率为 r_i ，发生的概率为 p_i 。

容易得出

$$N = \prod_{i=1}^n (1 + r_i f)^{p_i} \quad (2)$$

我们想求极值，但是直接求导不方便，于是借鉴最大似然估计法的做法，对 $\ln N$ 来求导可以化乘为加得

$$\ln N = \sum_{i=1}^n p_i \ln(1 + r_i f) \quad (3)$$

令导数等于 0，求解我们便能得到著名的凯利公式[1]

$$\sum_{i=1}^n \frac{p_i}{(\frac{1}{r_i} + f)} \quad (4)$$

如果赌局只有 2 种结果，赢或者输，且赢得回报率为 a ，那么可以将凯利公式简化为

$$f^* = p - \frac{1-p}{a-1} \quad (5)$$

按照这样得比例投资可以最大化资金的长期增长率

3 小概率事件的最优应对

对于一些小概率事件（以千分之三为分界），由于我们下注的次数很少，可以近似的视为每次的成功概率变化很小，那么这个时候我们该采取怎样的策略呢？首先我们明白对于小概率事件我们绝对不能作为先锋，可能先锋会成功，但是他们取得成功的概率实在是微乎其微，那么我们该什么时候下手呢，这将会是本文论述的重点，首先由大数定律可知，频率会慢慢收敛于概率，但是对于小概率事件我们该在什么时间点下手则取决于你要求的成功把握是多少，而成功把握的计算方法则是分为两种，对于只有一种最高奖励的呢，我们可以用伯努利大数定律来估计[2]，即

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right) = 1 \quad (6)$$

而如果对于那些有不同阶段奖励，但是我们只对最高阶段奖励感兴趣，则要回到切比雪夫大数定理[3]上来，即

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n x_k - \frac{1}{n} \sum_{k=1}^n E x_k\right| < \varepsilon\right\} = 1 \quad (7)$$

这样我们在进行小概率搞回报下注之前对于实际的成功概率心里已经有了一个底。但是这种估计方法对信息量要求较高，需要准确的之前发生的信息积累，并且对于外行人有较大的门槛。有没有牺牲一定精度但是更简便的方法呢？本文在这里提供海恩准则；即：每一起严重事故的背后，必然有 29 次轻微事故和 300 起未遂先兆以及 1000 起事故隐患。法则强调了事故的发生是量的积累的结果。以此类推，对于有 4 种不同阶段奖励的游戏，将一等奖视为小概率事件，我们可以在 1000 次四等奖，300 次三等奖，29 次二等奖之后开始出手追逐一等奖。但是海恩准则的精度并不及切比雪夫大数定律，因此只作为参考，实际上我们还可以通过海恩准则来印证大数定律。

参考文献

- [1] The Kelly Criterion, Money Management. Investopedia
- [2] 盛骤，谢式千，潘承毅 编 . 概率论与数理统计(第 4 版): 高等教育出版社，2008
- [3] 孙静娟主编；杨光辉，杜婷副主编. 统计学 第 3 版: 清华大学出版社，2015.08

贝叶斯公式的简要分析及推广

张虎 (2018120702012)

指导教师：吕恕

【摘要】 贝叶斯公式是概率论与数理统计中的一大基础工具，其应用十分广泛。本文从贝叶斯公式出发，探讨了公式中的先验概率和似然概率（调整因子），并结合例子分析了它们对结果（后验概率）的影响。最后推广了贝叶斯公式并尝试用其解释“最大似然估计”和“最大先验概率估计”，以此加深对贝叶斯公式的理解。

【关键词】 贝叶斯公式；先验概率；似然概率；后验概率；最大似然估计

1 贝叶斯公式及其说明

1.1 贝叶斯公式

贝叶斯公式是概率论中求“逆向概率”的基础工具，其基本内容如下：

定理（贝叶斯公式）：设随机试验 E 的样本空间为 Ω , $A \subset \Omega, B_i(i = 1, 2, \dots, n)$ 是 Ω 的一个有限划分，且 $P(A) > 0$, $P(B_i) > 0$, 则有：

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

由全概率公式可将(1)式改写为：

$$P(B_i|A) = P(B_i) \times \frac{P(A|B_i)}{P(A)} \quad (2)$$

(2)式所示贝叶斯公式常用于现实工程中的推断。

1.2 贝叶斯公式的说明

①韦恩图示下的直观意义

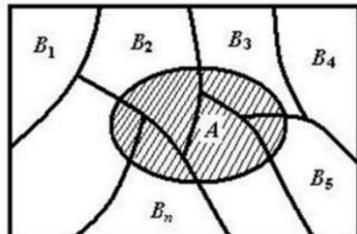


图 1 韦恩图示

如图 1，样本空间 Ω , 事件 A 、 $B_i(i = 1, 2, \dots, n)$ 。 B_i 是 Ω 的一个有限划分。

因概率的公理化定义中非负性、可列可加性与几何测度对应性质类似，故又称 $P(A)$ 为 A 的概率测度。于是不妨在韦恩图中将事件的概率与其在图示中所占面积（几何测度）相对应。样本空间 Ω 下，贝叶斯公式则可理解为是去求 A 面积中属于 B_i 的那部分面积在 A 总面积中的比重，即 $P(B_i|A)=B_i$ 与 A 重合的面积/ A 的面积。

②三个概念

在贝叶斯公式中将 $P(B_i)$ 为先验概率, $P(A|B_i)$ 称为似然概率(或调整因子), $P(B_i|A)$ 称为后验概率。下面分别对它们进行阐述。

先验概率 $P(B_i)$ 。事件 B_i 可在所研究的样本空间 Ω 下直接发生, 不需要发生的理由或前提条件, 且每一个 B_i 都可以导致 A 的发生, 这样的事件 B_i 所对应概率 $P(B_i)$ 称为“先验概率”。依照伯努利大数定律“频率稳定于概率”的结果, B_i 发生的频率 $f_n(B_i)$ 稳定于 $P(B_i)$, 则可通过历史统计数据求得 $f_n(B_i)$ 来代替 $P(B_i)$ 。

似然概率(调整因子) $P(A|B_i)$ 。任一 B_i ($i = 1, 2, \dots, n$)都能导致 A 的发生, 为使贝叶斯公式有意义, 对任一 B_i 有 $P(A|B_i) \neq P(A)$, 即事件 B_i 与 A 不相互独立, B_i 对 A 是有影响的, 似然概率 $P(A|B_i)$ 在一定程度上能反映这种影响, 其值也能由历史统计数据求频率获得。

后验概率 $P(B_i|A)$ 是在事件 A 已发生后, 去寻找其发生的原因, 于是分析每一个原因 B_i 对这一结果的“贡献”, 其值由先验概率、似然概率而定, 不能由统计数据获得。

③对“后验概率”的剖析

贝叶斯公式是去求后验概率, 前提是事件 A 已发生。 A 可看作独立存在的, 其概率 $P(A)$ 也能通过统计数据获得。既然 A 已发生, $P(A)$ 既定, 我们就不能主宰 $P(A)$, 因此抛开常数 $P(A)$ 。

则

$$P(B_i|A) \propto \{P(B_i) \times P(A|B_i)\} \quad (3)$$

据此可通过调整 $P(B_i)$ 和 $P(A|B_i)$ 来调整后验概率 $P(B_i|A)$ 的值, 使 $P(B_i|A)$ 表达的含义更具可靠性(将在后文通过例子的形式阐明)。

2 “先验概率”对贝叶斯公式的贡献

先验概率 $P(B_i)$ 是计算 $P(B_i|A)$ 的前提, 似然概率 $P(A|B_i)$ 的作用也通常要在先验概率作用后才能体现出来。对应于1.2中①的“测度”来说, 即要先存在 B_i 的面积, 讨论 B_i 与 A 重合的面积才有意义, 进而才能讨论面积的比(面积的比对应于条件概率)。

以例题的形式来进一步说明先验概率的作用。

2.1 应用例——假病问题

设某医院通过某方法诊断某种疾病, 由于各种原因, 被诊断出患有该病的就诊者未必患有该病, 记 $A=\{\text{就诊者被检测患有该病}\}$, $B=\{\text{就诊者确实患有该病}\}$, 现有以下数据: $P(B)=0.0004$ (患者在人群中占比很低), $P(A|B)=0.95$ (对患有此病的就诊者诊断准确率很高), $P(\bar{A}|\bar{B})=0.9$ (对未患此病的就诊者诊断准确率很高)。

下面用贝叶斯公式计算出某人去医院用该方法“被诊断出患此病且确实患此病”的概率。

A 已发生, 且 B 和 \bar{B} 构成样本空间 Ω 的有限划分, 于是运用贝叶斯公式:

$$\begin{aligned} P(B|A) &= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} \\ &= \frac{0.95 \times 0.0004}{0.95 \times 0.0004 + (1 - 0.9) \times (1 - 0.0004)} \\ &= 0.0038 \end{aligned} \quad (4)$$

2.2 结果分析

事件 $B|A$ (就诊者“被诊断出患此病且确实患此病”)的概率不到0.4%, 是一个小概率事件。

而由小概率实际推断原理，一次检查（即一次试验）中， $B|A$ 这一事件几乎不可能发生，说明在检测方法正确率很高的情况下一次检测中诊断正确的结果却基本不会发生，这与我们的认知相悖。分析(4)式，导致 $P(B|A)$ 很小的原因是分子项 $P(B)=0.0004$ 很小，此处用其作为“先验概率”并不恰当，因为就诊者已经主动到医院问诊，其患病的概率应远远超过一般人群。

下面将先验概率 $P(B)$ 看作变量，其他数值不变，绘制出后验概率 $P(B|A)$ 与 $P(B)$ 的曲线图(图2)，

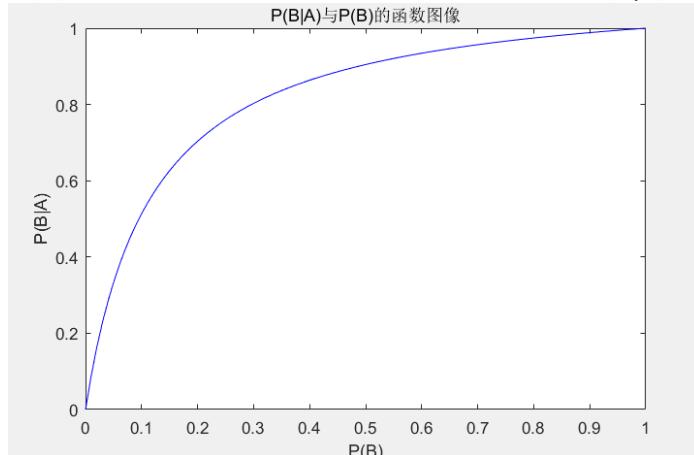


图 2

可看出函数图像前期变化迅速，后期趋于平稳。因此当先验概率 $P(B)$ 较小时对后验概率结果的可靠性有很大影响， $P(B)$ 较小的变化都可能使 $P(B|A)$ 变化显著。而当 $P(B)$ 达到一定值时，再调整 $P(B)$ 结果 $P(B|A)$ 变化也不显著，这时就需考虑调整“似然概率”了。

因此若选取的先验概率不恰当，特别是当 $P(B)$ 本身较小时，计算所得后验概率表达的含义就不再可靠。更进一步，当 $P(B)$ 极小时，无论如何调整 $P(A|B)$ 也无多大意义，结果 $P(B|A)$ 也会很小。

这启示我们：在医疗诊断的过程中，对于人群中患病率极低的病，即使是准确度很高的检测方法，若对全民检测进行病例排查，也无多大意义，检查者“被检测出患病但却未患病”的情况会普遍发生，只会造成众人虚惊一场；而对发病率极高的病（如传染病等），就很有必要进行全民监测预防。

2.3 “先验概率”的调整

综上，先验概率是贝叶斯公式结果可靠的基础。为确保实际应用中后验概率结果的可靠性，必须调整先验概率 $P(B_i)$ 使之最适合于事件的背景。

以新冠肺炎的诊断过程为例，简单说明“先验概率”的调整。设 $A=\{\text{就诊者被检测患有新冠肺炎}\}$, $B=\{\text{就诊者患新冠肺炎}\}$ 。计算中用 $f_n(B)$ 代替 $P(B)$ ，而 $f_n(B) = m/n$ (m 是 n 重伯努利试验中事件 B 出现的次数)，提醒我们一方面要选取合适的“每一次伯努利试验 E ”，另一方面要使样本容量 n 尽可能大。

若 $E=\{\text{全国人口范围内任取一人检测其患新冠肺炎与否}\}$ ，计算所得 $f_{n1}(B)$ ，该值代表的是新冠肺炎的国民发病率。但当一位居民出现了疑似症状并主动网上问诊时，若用 $f_{n1}(B)$ 代替 $P(B)$ 显然会削弱 $P(B|A)$ 结果的表达能力。此时应令 $E=\{\text{全国所有网上问诊者中任取一人检测其患新冠肺炎与否}\}$ ，利用已有统计数据计算出 $f_{n2}(B)$ 代替 $P(B)$ 。因此，调整先验概率 $P(B)$ 要结合所研究事件的背景，据此选取合理的 $P(B)$ 。

此外，对新冠肺炎这样 $P(B)$ 较高且在初发阶段历史统计数据不可观的新发病，更需对 n 与 m 进行实时更新，并随着样本容量的扩大不断合理调整 $P(B)$ 。

当先验概率调整好后，似然概率的作用也就开始发挥。在 $P(B)$ 确定后，医生通过经验知识等初步断定该就诊者是否患病，再通过科学的检测方法进行最终的检测，对检测结果进行最终调整。而检测方法的准确度就体现在 $P(B|A)$ 上，此时 $P(B|A)$ 当然越高越好。

3 “似然概率（调整因子）”对贝叶斯公式的贡献

3.1 先验概率的局限性、贝叶斯公设

先验概率是贝叶斯公式可靠的基础，但有时试验前无法通过统计数据获得 $P(B_i)$ 。为运用先验概率进行计算，在完全不知道任何先验信息时，可假设完备事件组中的事件 B_1, B_2, \dots, B_n 具有等可能性，这一假设称之为贝叶斯公设。

如 2.1 在贝叶斯公设的前提下，设 $P(B) = 0.5$ ，代入到(4)中，解得 $P(B|A) = 0.9048$ ，则 $P(\bar{B}|A) = 1 - 0.9048 = 0.0952$ ，即就诊者“被检测出患病但并未患病”为一小概率事件，一次检测中几乎不可能发生，符合我们的认知。

下面再通过“拼写纠正”的例子来说明“似然概率”的作用。

3.2 应用例——拼写纠正

贝叶斯公式是 AI、深度学习等应用的一大基础，这里以拼写纠正为例来简述其应用体现，并以此分析似然概率对后验概率的“贡献”。

设 $A=\{\text{拼写者输入拼错单词}W\}$, $B_i=\{\text{拼写者想输入单词}W_i(i=1,2,3,\dots,n)\}$ (假设 W 的形近词为 W_i ，其数目总是有限的)。要从这 n 个形近词里猜测出拼写者最可能想输入的单词，即求使 $P(B_i|A)$ 最大的事件 B_i 。 B_i 构成样本空间 Ω 的有限划分，所以 $P(B_i|A)$ 可由(2)式得出，且总 $\exists k \in \{1,2,3,\dots,n\}$ ，使 $P(B_k|A)$ 具有最大值 P_{max} ，则 B_k 所对应单词 W_k 即为猜测的拼写者最想输入的单词。

这里先验概率 $P(B_i)$ 表示单词 W_i 出现的概率，似然概率 $P(A|B_i)$ 则表示在拼写者想输入单词 W_i 的条件下误拼为 W 的概率，它们都能由历史统计数据获得。由(3)式： $P(B_i|A) \propto \{P(B_i) \times P(A|B_i)\}$ ，因此只需计算出每个单词 W_i 所对应的 $P(B_i) \times P(A|B_i)$ ，找出使 $P(B_i) \times P(A|B_i)$ 最大的事件 B_k ，则 B_k 对应的 $P(B_k|A)$ 即为 P_{max} ，相应单词 W_k 即为所求单词。

而计算过程中可能存在 $P(B_i)$ 相等的情况，这时候只需找出所有 $P(B_i)$ 相等的单词中 $P(A|B_i)$ 最大的一项进行后续比较，这便是“似然概率”此处对“后验概率”的“贡献”。当先验概率确定后，根据应用背景合理调整或选择恰当的似然概率，也能使结果的可靠性更高。

3.3 小结

“似然概率”(调整因子)在先验概率已合理求出或贝叶斯公设下(各事件 B_i 等可能的前提下)，其作用就开始发挥。 $P(A|B_i)$ 反映的是 B_i 对 A 的一种影响，当这种“影响”对我们有益时，需采取手段提高其数值，否则降低其数值。

而 $P(A|B_i) = P(AB_i)/P(B_i)$ ，反映在韦恩图上即为 A 与 B_i 重合的面积与 B_i 的面积之比，也说明了 $P(A|B_i)$ 要以 $P(B_i)$ 为基础才能使 $P(B_i|A)$ 的值的意义更具有可靠性。

4 贝叶斯公式的推广及简单应用

4.1 贝叶斯公式的推广

综上，贝叶斯公式的结果 $P(B_i|A)$ 表示 A 发生后 B_i 这一原因在所有原因中的比重。而这一比重取决于“该原因本身独立的可能性大小”(先验概率 $P(B_i)$)和“这一原因下导致结果发生的可能性

大小”(似然概率 $P(A|B_i)$)。上面的讨论都是基于事件的讨论，下面对贝叶斯公式进行推广来解决数理统计中的一些问题。

用参数 θ_i 代替 B_i ，用 x 代替 A ，将 x 看作是一个总体或一种输出，参数 θ_i 和 x 间能相互影响。用 $P(x|\theta_i)$ 替换 $P(A|B_i)$ ， $P(\theta_i|x)$ 替换 $P(B_i|A)$ ， $P(\theta_i)$ 替换 $P(B_i)$ 。

$P(x|\theta_i)$ 表示在给定参数 θ_i 时，随机变量 $X = x$ 的概率，当 X 为离散型随机变量时，则 $P(X = x_k|\theta_i), k = 1, 2, \dots$ 为 X 的分布律，为方便公式的推导，直接将 $P(x|\theta_i)$ 看作 X 的分布律；同理，将 $P(\theta_i|x)$ 看作输出 x 下参数 θ_i 的分布律， $P(\theta_i)$ 表示 θ_i 的分布律。

则贝叶斯公式可改写为：

$$P(\theta_i | x) = \frac{P(x|\theta_i)P(\theta_i)}{\sum_{i=1}^n P(x|\theta_i)P(\theta_i)} \quad (5)$$

该式表示输出 x 时参数 θ_i 的分布律的求法。类似 1.2 中的讨论， x 已知时分母不是我们能主宰的，则：

$$P(\theta_i | x) \propto \{P(x|\theta_i) \times P(\theta_i)\} \quad (6)$$

(5)(6) 两式是 θ_i 为离散型随机变量、 x 在参数 θ_i 下也为离散型随机变量的结果，而 θ_i 和 $x|\theta_i$ 也可能为连续型随机变量，故实质上有四种组合，当其都为连续型随机变量时。则：

$$f(\theta_i | x) = \frac{f(x|\theta_i)f(\theta_i)}{\int_{\Omega} f(x|\theta_i)f(\theta_i)d\theta_i} \quad (7)$$

其余两种情况不再赘述，对式 (7)，同理应有：

$$f(\theta_i | x) \propto \{f(x|\theta_i) \times f(\theta_i)\} \quad (8)$$

将 $P(x|\theta_i)$ 和 $f(x|\theta_i)$ 称作似然函数，记作 $L(x; \theta)$ 。

4.2 最大似然估计与最大后验概率估计

现用样本推断总体分布中的某参数 θ 。设 X_1, X_2, \dots, X_n 为来自总体的样本， x_1, x_2, \dots, x_n 为观测值，将这组样本看成一个输出 x 。每个样本都对应一似然函数 $L(x_i; \theta)$ ($i = 1, 2, \dots, n$)。为充分利用样本信息，令 $L(x_1, x_2, \dots, x_n; \theta) = \prod_1^n L(x_i; \theta)$ ，则 $L(x_1, x_2, \dots, x_n; \theta)$ 可代表该样本构成的总体的似然函数。

现在从 (6) 式的角度进行对 θ 的估计，假定 θ 为一个数，将其估计值记为 $\hat{\theta}$ ， $\hat{\theta}$ 为一随机变量。我们想让由样本 x_1, x_2, \dots, x_n 推出的 $\hat{\theta}$ 最接近 θ ，则使 $P(\hat{\theta} = \theta | x_1, x_2, \dots, x_n)$ 最大即可，由 (6) 式，则使 $P(x_1, x_2, \dots, x_n | \hat{\theta} = \theta) \times P(\hat{\theta} = \theta)$ 最大。理论上希望 $\hat{\theta} = \theta$ ，故假定 $P(\hat{\theta} = \theta)$ 为一接近 1 的常数，于是直接使 $P(x_1, x_2, \dots, x_n | \hat{\theta})$ 最大即可，则 $L(x_1, x_2, \dots, x_n; \theta) = \prod_1^n P(x_i | \theta)$ 也同步具有最大值。因此只需求出使 $L(x_1, x_2, \dots, x_n; \theta) = \prod_1^n P(x_i | \theta)$ 最大时的 θ ，即为估计值 $\hat{\theta}$ ，这一做法也能推广到 (8) 式中。总之，去寻找使 $L(x_1, x_2, \dots, x_n; \theta) = \prod_1^n L(x_i; \theta)$ 最大时的 θ 即为所求估计的 $\hat{\theta}$ 。这就是我们所熟悉的“最大似然估计法”。

但 θ 也可能是随机变量，即 θ 具有先验分布，在明确 θ 的分布律或概率密度下，为充分利用所给信息，需求出使后验分布 $P(\theta | x_1, x_2, \dots, x_n)$ 或 $f(\theta | x_1, x_2, \dots, x_n)$ 最大的 θ 即为 $\hat{\theta}$ ，转化为求使 (6)(8) 两式中乘积部分最大时的 θ 值，此方法称作“最大后验概率估计法”。

5 结语

本文从贝叶斯公式出发，阐述了公式中每一项的含义，并用“假病例”和“拼写纠正”两个例子简单叙述了“先验概率”和“似然概率”对贝叶斯公式结果“后验概率”的影响。最后对贝叶斯公式进行了推广并尝试用推广后的式子解释两种“估计”方法。但贝叶斯公式的魅力远不止于此，希望能以此引发出更多对贝叶斯公式的认识和思考。

参考文献

- [1] 徐金智, 吕恕.概率论与数理统计 (第三版) [M].北京: 高等教育出版社, 2017
- [2] 傅莺莺, 田振坤.疾病检测中的概率问题——一例读懂贝叶斯公式[J].中国统计, 2017 (11), 26-28
- [3] 孟红岩.对 AI 的基础——贝叶斯公式的研究 [J]. 现代经济信息, 2018 (3), 430
- [4] 刘娟, 胡桂武.人工智能思维利器——贝叶斯公式的教学探究[J].教育教学论坛, 2019 (9), 38, 141-143
- [5] 刘未鹏.数学之美番外篇: 平凡而又神奇的贝叶斯方法,
<http://mindhacks.cn/2008/09/21/the-magical-bayesian-method>

二维随机变量差和积的分布及特殊分布的拓展

张家豪 (2018051403015)

指导教师：吕恕

【摘要】本文着重讨论了已知 (X, Y) 的联合分布，两个随机变量差的分布，积的分布，以及两个随机变量和，差，积，商的分布的拓展

【关键词】随机变量；特殊分布；随机变差的分布；积的分布

1 差的分布

设随机变量 (X, Y) 的联合概率密度为 $f(x,y)$ ，根据分布函数的定义，则它们差 $Z=X-Y$ 的分布函数是

$$F_Z(z) = P\{Z \leq z\} = \iint_{x-y \leq z} f(x,y) dx dy = \int_{-\infty}^{+\infty} dx \int_{x-z}^{+\infty} f(x,y) dy \quad (1)$$

积分区域如下图红色箭头部分。做变量替换，令 $y = x - t$, $dy = -dt$ 说明：这里要固定 x , z ，变量替换的目的是将 $x - z$ 替换为我们需要的 z ，以希望凑成积分上限函数的形式，便于对分布函数求导直接将积分符号去掉，不改变其他形式，从而得到概率密度。也即是 $t=z$ 时， $y=x-z$ ，带入进去，从而变量替换的具体关系

$$y = x - t \quad (2)$$

$$\text{原式} = \int_{-\infty}^{+\infty} dx \int_z^{-\infty} -f(x, x-t) dt = \int_{-\infty}^{+\infty} dx \int_{-\infty}^z f(x, x-t) dt$$

$$\text{交换积分次序后} \quad = \int_{-\infty}^z \int_{-\infty}^{+\infty} f(x, x-t) dx dt \quad (3)$$

$$\text{对 } F_Z(z) \text{ 求导得} \quad f_Z(z) = \int_{-\infty}^{+\infty} f(x, x-z) dx \quad (4)$$

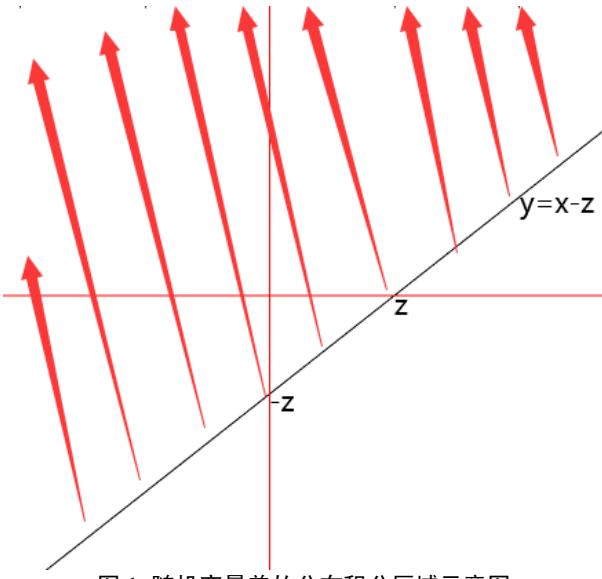


图 1 随机变量差的分布积分区域示意图

2 积的分布

设随机变量 (X, Y) 的联合概率密度为 $f(x,y)$ ，根据分布函数的定义，则它们积 $Z=X*Y$ 的分布函数是

$$FZ(z) = P\{Z \leq z\} = \iint_{xy \leq z} f(x, y) dx dy \quad (5)$$

$xy \leq z$ 等价于当 $x > 0$ 时 $y \leq z/x$, 当 $x < 0$ 时 $y \geq z/x$

① 当 $z > 0$ 时图像如下, 红色箭头所在区域为积分区域

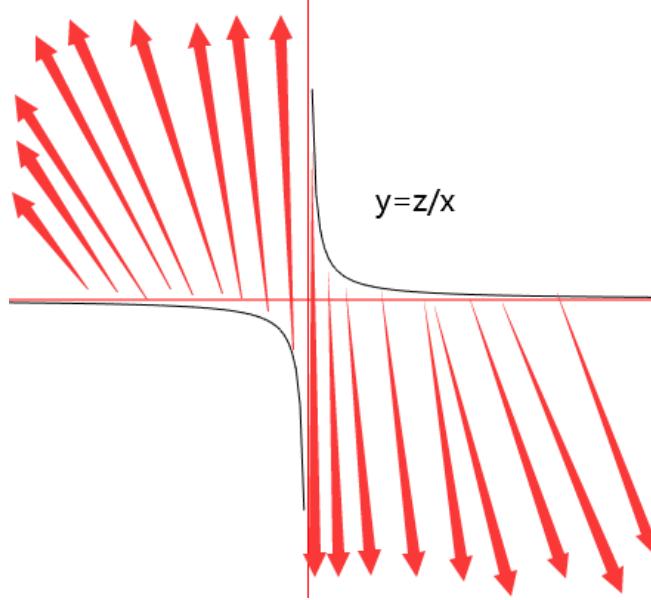


图 2 $z > 0$ 时积分区域示意图

② 当 $z < 0$ 时图像如下, 红色箭头区域为积分区域

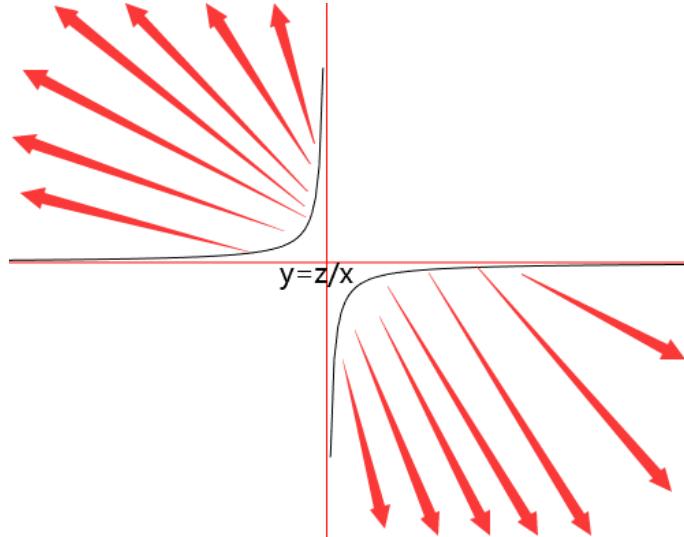


图 3 $z < 0$ 积分区域示意图

$$\begin{aligned} FZ(z) &= P\{Z \leq z\} = \iint_{xy \leq z} f(x, y) dx dy \\ &= \int_0^{+\infty} dx \int_{-\infty}^{z/x} f(x, y) dy + \int_{-\infty}^0 dx \int_{z/x}^{+\infty} f(x, y) dy \end{aligned} \quad (6)$$

做变量替换, 令 $y = \frac{t}{x}$, $dy = \frac{1}{x} dt$ 方法与前面相同

$$\begin{aligned} \text{原式} &= \int_0^{+\infty} dx \int_{-\infty}^z \frac{1}{x} f\left(x, \frac{t}{x}\right) dt + \int_{-\infty}^0 dx \int_z^{-\infty} \frac{1}{x} f\left(x, \frac{t}{x}\right) dt \\ &= \int_0^{+\infty} dx \int_{-\infty}^z \frac{1}{x} f\left(x, \frac{t}{x}\right) dt + \int_{-\infty}^0 dx \int_{-\infty}^z -\frac{1}{x} f\left(x, \frac{t}{x}\right) dt \\ &= \int_0^{+\infty} dx \int_{-\infty}^z \frac{1}{|x|} f\left(x, \frac{t}{x}\right) dt + \int_{-\infty}^0 dx \int_{-\infty}^z \frac{1}{|x|} f\left(x, \frac{t}{x}\right) dt \\ &= \int_{-\infty}^{+\infty} dx \int_{-\infty}^z \frac{1}{|x|} f\left(x, \frac{t}{x}\right) dt \end{aligned}$$

交换积分次序后

$$= \int_{-\infty}^z \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{t}{x}\right) dx dt \quad (7)$$

对 $F_Z(z)$ 求导得

$$f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{z}{x}\right) dx \quad (8)$$

3 二维随机变量和, 差, 积, 商的分布的拓展

利用已学和, 商的公式以及上述差, 积的公式固然可以做到简便的求解, 然而存在一些问题不是简单的二维随机变量和, 差, 积, 商的分布, 而是由其变化而来.

例如: 设已知 (X, Y) 的联合概率密度为 $f(x, y)$, 则 $Z=2(X+Y)+1$, $Z=2(X-Y)+1$, $Z=2X*Y+1$, $Z=2X/Y+1$ 的概率密度又该如何求解呢?

固然同样可以利用 1, 2 中的根据分布函数定义方法来一步步进行求解, 但运算难免会比较繁琐, 容易出错。我们考虑形如 $Z=a(X-Y)+b$ 这样的分布函数 (Z 与 $X-Y$ 有着线性的关系) 是不是也有着类似的结论。

我们可以考虑借助中间变量 $U=X-Y$, 通过上面的结论容易求出 $U=X-Y$ 的概率分布 $f_U(u) = \int_{-\infty}^{+\infty} f(x, x-u) dx$, 再找出 $Z=aU+b$ 的概率分布 $f_Z(z)$ 与 $f_U(u)$ 之间的关系便能求出 $f_Z(z)$

转化为已知 $f_U(u)$, 求 $Z=aU+b$ 的概率密度

当 $a>0$ 时

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{aU+b \leq z\} = P\{U \leq (z-b)/a\} \\ &= FU((z-b)/a) \end{aligned} \quad (9)$$

两边求导得

$$\begin{aligned} f_Z(z) &= f_U((z-b)/a)(z-b)/a' \\ &= f_U((z-b)/a) \frac{1}{a} \end{aligned} \quad (10)$$

当 $a<0$ 时

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{aU+b \leq z\} \\ &= P\{U \geq (z-b)/a\} \\ &= 1 - FU((z-b)/a) \end{aligned} \quad (11)$$

两边求导得

$$\begin{aligned} f_Z(z) &= -f_U((z-b)/a)(z-b)/a \\ &= -f_U((z-b)/a) \frac{1}{a} \end{aligned} \quad (12)$$

综合可以写成

$$f_Z(z) = f_U((z-b)/a) \frac{1}{|a|} \quad (13)$$

即是 $Z=aU+b$ 的概率密度

1. 对于差的分布 $U=X-Y$ 的概率分布

$$f_U(u) = \int_{-\infty}^{+\infty} f(x, x-u) dx \quad (14)$$

则对于分布 $Z=a(X-Y)+b$ 的概率密度

$$f_Z(z) = f_U((z-b)/a) \frac{1}{|a|} = \frac{1}{|a|} \int_{-\infty}^{+\infty} f(x, x-(z-b)/a) dx \quad (15)$$

2. 同理对于和的分布的拓展

$$Z=a(X+Y)+b=aU+ \quad (16)$$

有 $U=X+Y$ 的概率密度

$$f_U(u) = \int_{-\infty}^{+\infty} f(x, u-x) dx \quad (17)$$

$Z=a(X+Y)+b=aU+b$ 的概率密度

$$f_Z(z) = f_U((z-b)/a) \frac{1}{|a|} = \frac{1}{|a|} \int_{-\infty}^{+\infty} f(x, \frac{z-b}{a}-x) dx \quad (18)$$

3. 对于积的分布的拓展

$$Z=aX*Y+b=aU+b \quad (19)$$

有 $U=X*Y$ 的概率密度

$$f_U(u) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{u}{x}\right) dx \quad (20)$$

$Z=aX+b=aU+b$ 的概率密度

$$f_z(z) = f_U((z-b)/a) \frac{1}{|a|} \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{z-b}{a}\right) dx \quad (21)$$

4. 对于商的分布的拓展

$$Z=aX/Y+b=aU+b \quad (22)$$

有 $U=X/Y$ 的概率密度

$$f_U(u) = \int_{-\infty}^{+\infty} |y| f(uy, y) dy \quad (23)$$

$Z=aX/Y+b=aU+b$ 的概率密度

$$f_z(z) = f_U((z-b)/a) \frac{1}{|a|} \int_{-\infty}^{+\infty} |y| f\left(\frac{z-b}{a}, y\right) dy \quad (24)$$

4 二维随机变量和, 差分布的进一步拓展

由前面我们得到了 $Z=aX+b$, $Z=aX/Y+b$, $Z=a(X-Y)+b$, $Z=a(X+Y)+b$ 的概率分布函数, 对积和商的分布我们得到了较为一般的结论, 而和, 差的分布虽然进行了拓展, 但能能求解的形式还是具有特殊性。我们进一步考虑更一般的形式 $Z=aX+bY+c$ 的概率密度。

在这里 $a \neq 0$ 且 $b \neq 0$, 当 a, b 中出现等于 0 的情况时, 由上面的结论可直接求出。

1. 设随机变量 (X, Y) 的联合概率密度为 $f(x,y)$, 根据分布函数的定义, 则 $Z=aX+bY+c(a, b$ 取值

不同积分区域不同, 先讨论 $a>0, b>0$ 的情形, 其他情况方法相同)

分布函数是

$$F_Z(z) = P\{Z \leq z\} = \iint_{ax+by+c \leq z} f(x, y) dx dy = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{\frac{z-ax-c}{b}} f(x, y) dy \quad (25)$$

做变量替换, 令 $y = \frac{t-ax-c}{b}$, $dy = \frac{1}{b} dt$ 说明: 这里要固定 x, z , 变量替换的目的是将 $\frac{z-ax-c}{b}$ 替换为我们需要的 z , 以希望凑成积分上限函数的形式, 便于对分布函数求导直接将积分符号去掉, 不改变其他形式, 从而得到概率密度。也即是 $t=z$ 时, $y=\frac{z-ax-c}{b}$, 带入进去, 从而变量替换的具体关系

$$y = \frac{t-ax-c}{b} \quad (26)$$

原式

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{+\infty} dx \int_{-\infty}^z \frac{1}{b} f\left(x, \frac{t-ax-c}{b}\right) dt \\ &= \frac{1}{b} \int_{-\infty}^z \int_{-\infty}^{+\infty} f\left(x, \frac{t-ax-c}{b}\right) dx dt \end{aligned} \quad (27)$$

$$\text{对 } F_Z(z) \text{ 求导得 } f_z(z) = \frac{1}{b} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax-c}{b}\right) dx \quad (28)$$

同理, 当 $a<0, b>0$ 时,

$$f_z(z) = \frac{1}{b} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax-c}{b}\right) dx \quad (29)$$

当 $a>0, b<0$ 时,

$$f_z(z) = -\frac{1}{b} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax-c}{b}\right) dx \quad (30)$$

当 $a<0, b<0$ 时,

$$f_z(z) = -\frac{1}{b} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax-c}{b}\right) dx \quad (31)$$

总结为

$$f_z(z) = \frac{1}{|b|} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax-c}{b}\right) dx \quad (32)$$

若对固定 y, z ; 对 x 进行变量替换可以得到对称的结构 $f_z(z) = \frac{1}{|a|} \int_{-\infty}^{+\infty} f\left(\frac{z-by-c}{a}, y\right) dy$

利用得到的公式对 $Z=a(X-Y)+b$, $Z=a(X+Y)+b$ 的情形进行检验

$$\text{对于 } Z=a(X-Y)+b; \text{ 得到 } f_z(z) = \frac{1}{|a|} \int_{-\infty}^{+\infty} f\left(x, x - \frac{z-b}{a}\right) dx \quad (33)$$

$$\text{对于 } Z=a(X-Y)+b; \text{ 得到 } f_z(z) = \frac{1}{|a|} \int_{-\infty}^{+\infty} f\left(x, \frac{z-b}{a} - x\right) dx \quad (34)$$

这与 3 中通过另一种算法得到的结果一致

5 结论的应用

利用这些结论能够很简便迅速的求解出一般题目中的概率密度

例如：(05 年考研试题) 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 1 & , 0 < x, 0 < y < 2x \\ 0 & others \end{cases}$$

$Z = 2X - Y$, 求 Z 的概率密度 $f_z(z)$

一般方法是根据分布函数的定义对 z 的不同取值导致的不同积分区间进行讨论，先求出 Z 的分布函数，再对其求导从而得出 $f_z(z)$ ，较为麻烦

而我们直接用上面的结论 $f_z(z) = \frac{1}{|b|} \int_{-\infty}^{+\infty} f\left(x, \frac{z-ax-c}{b}\right) dx$ 式(32)就可以求解，其中 $a=2$, $b=-1$, $c=0$ 带入得

$$f_z(z) = \int_{-\infty}^{+\infty} f\left(x, \frac{z-2x}{-1}\right) dx \quad (35)$$

当仅当 $0 < x < 1, 0 < \frac{z-2x}{-1} < 2x$ 时，即 $0 < x < 1, z > 0, x > z/2$ 时为非零，画出积分区域

图容易得出

$$\begin{aligned} f_z(z) &= \begin{cases} \int_{z/2}^1 f\left(x, \frac{z-2x}{-1}\right) dx, & 0 < z < 2 \\ 0 & others \end{cases} \\ &= \begin{cases} 1 - z/2 & 0 < z < 2 \\ 0 & others \end{cases} \end{aligned} \quad (36)$$

参考文献

- [1] 徐全智, 吕怒, 概率论与数理统计, 高等教育出版社, 2017.8
- [2] 杨永发, 概率论与数理统计教程, 南开大学出版社, 2000.11
- [3] 刘洋, 何璐, 孙丽英, 徐贵桥, 随机变量积的分布及其应用, 大学数学, 第 30 卷第 5 期, 112-115

浅谈小概率事件原理在实际中的应用

张新豫 (2018040704028)

指导教师：吕恕

【摘要】 小概率事件原理是概率论与数理统计中一个简单、基本的原理,是概率论与数理统计的精髓,实用价值比较高、应用范围也很广。文章主要阐述小概率事件的原理及其证明、特点和应用,帮助人们正确认识、对待小概率事件,同时更好地了解小概率事件原理在生活实践中的应用。

【关键词】 小概率事件；原理；贝努利大数定律；特点；应用

1 小概率事件的原理简介

在概率论中,我们把发生概率几乎接近于 0 的事件称为小概率事件。所谓小概率事件原理是说,“概率很小的事件在一次实验中几乎不会发生”。那么多大的概率值算小概率呢?一般多采用 0.01、0.05 这两个值作为衡量标准:即事件发生的概率在 0.01 或 0.05 以下的事件称为小概率事件,0.01、0.05 这两个值称为小概率标准。而在某些重要的试验或场合,当事件一旦发生,后果不堪设想时,小概率事件的阀值应选得比上述两个值更小一些。小概率事件在日常生活中也是时有发生的,例如:某人买彩票中头奖,飞机失事等。虽然这些事件本身发生的概率较小,但往往具有一定的影响力,因此小概率事件是不可以忽视的,研究小概率事件原理很有必要。

2 小概率事件原理的证明

贝努利大数定律:在 n 次独立重复试验中,记事件 A 发生的次数为 n_A , p 是事件 A 发生的概率,则对于任意正数 $\varepsilon < 0$, 有

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0 \quad (1)$$

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1 \quad (2)$$

根据贝努利大数定律,事件 A 发生的频率 n_A/n 依概率收敛于事件 A 发生的概率 p 。就是说,当 n 很大时,事件 A 发生的频率与概率有较大偏差的可能性非常小,二者在数值上基本相等。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近似代替概率。如果某事件 A 发生的概率 α 很小,那么它在大量重复试验中出现的频率也应该很小。例如,若 $\alpha=0.001$,则大体上在 1000 次试验中, A 才能出现 1 次。因此,概率很小的事件在一次试验中不大可能发生。在概率论的应用中,称这样的事件为实际不可能事件。实际不可能事件在一次试验中实际上是不可能发生的。这就是小概率事件原理,也称为小概率的实际不可能性原理。

3 小概率事件的特点

3.1 小概率事件总有可能发生

虽然小概率事件是发生可能性很小的事件,它发生的概率很小,但是这不能说明小概率事件永远不发生,无论其发生概率多小,只要存在概率值不为 0,小概率事件都有可能发生。根据概率的非负性,其概率值总是小于 1 的正数。小概率事件在一次试验中几乎不可能发生,但是只要独立的试验次数无限增多,小概率事件总有可能发生。

3.2 小概率事件具有双重性

假如小概率事件在一次试验中偶然发生了，那么对于这次试验，它就是“必然事件”。反之，假如小概率事件在很多次试验中都没有发生，那么对于这些试验，它就是“零概率事件”或“不可能事件”。由此可见，小概率事件的界定标准不是唯一的、一成不变的，它是在实际生活中随着客观现象发生而变化的。因此，小概率事件实际上存在着发展和消亡的双重性质。

3.3 小概率事件存在突发性

小概率事件平常几乎不会发生，所以它一旦发生，往往使人措手不及。而小概率事件大多数都具有极端性：要么破坏性很大，如突发的重大疫情、造成严重后果的地震、难以预料的飞机失事等等；要么惊喜感很强，如购买彩票或抽奖中头奖等。当然对于中头奖这样的好事我们希望它发生，但是对于天灾人祸这样的坏事我们是不可能完全避免其发生的，因此，研究破坏性强的小概率事件的发生可能性，提前准备预先的应急办法和应对方案，是非常有必要的。

3.4 小概率事件不同于不可能事件

由于发生概率比较小，人们常常会把小概率事件与不可能事件混为一谈。但事实上，小概率事件和不可能事件是有本质区别的。不可能事件，就是指绝对不可能发生、概率为零的事件。而小概率事件是有一定可能性发生，只不过发生的可能性很小的事件。小概率事件发生概率是一个趋近于0的正数，如果将某一小概率试验不断独立地重复进行，不论它发生的概率有多小，只要概率值是不为0的正数，它终究必然会发生一次，继续不断独立地重复进行下去，也必然会多次发生。而对于不可能事件，无论重复试验多少次它都不可能发生。由此可见，小概率事件不同于不可能事件，二者完全是两种概念，差别很明显。

4 小概率事件原理的应用

4.1 在统计学中的应用

概率论与数理统计中假设检验的方法是建立在小概率事件原理上的概率反证法。假设检验是指在给定显著性水平之下，判断某一假设的正确与否。在显著性假设检验理论中，一般把小概率 α 称为显著性水平，假设检验的结论与选取的显著性水平 α 有密切关系，不同的显著性水平下做出的假设检验结论可能不同。某种假设 H_1 需要检验，首先假设 H_1 是正确的，在此假设下构造某一事件A，在 H_1 为正确的前提条件下，以小概率 α 为显著性水平，事件A发生的概率很小，可以达到小概率标准。那么在一次试验中，如果小概率事件A发生了，不满足小概率事件原理，则可以怀疑假设 H_1 的正确性，因此否定 H_1 ；如果小概率事件A没有发生，满足小概率事件原理，则表明假设 H_1 成立，不能拒绝 H_1 。然而，假设检验也可能犯“弃真”和“纳伪”这两类错误，在实际应用中应该通过增加样本容量等方法尽量减少犯两类错误的概率。

4.2 在保险业中的应用

一直以来，保险行业的发展都比较平稳，很少出现保险公司重大亏损的事件。随着经济的快速发展，人民的生活水平不断提高，人们越来越注重生活质量、关注身体健康，因此，各种重大疾病保险层出不穷，生活中几乎每个人都至少买过一份保险。一般保险都属于意外伤害险，是以被保险人因遭受意外伤害造成死亡、残废为给付保险金条件的人身保险。投保人向保险人交纳一定的保险费，如果被保险人在保险期限内遭受意外伤害并以此为直接原因或近因，在自遭受意外伤害之日起

的一定时期内造成死亡、残废、支出医疗费或暂时丧失劳动能力，则保险人给付被保险人或其受益人一定量的保险金。通常情况下，赔付的保险金远高于保险费，一旦投保人发生意外事故需要保险公司赔付保险金，保险公司就会面临亏本甚至破产的危险。但是，为什么保险行业的发展能经久不衰呢？这其中便运用了小概率事件原理。要知道，人们在日常生活中遭遇重大意外事件的概率微乎其微，可以将发生意外需要保险公司赔付的事件看作小概率事件，因此保险公司每年的保险费收入远高于赔付的保险金，几乎稳赚不赔，所以保险行业作为小概率事件原理催生的行业才能一直平稳发展。

4.3 在气象学中的应用

在气象服务业，用户关心的暴雨、强降雹、雷雨大风、龙卷风等许多灾害性天气现象，都属于发生次数较少的小概率事件。提高这些小概率天气的预报成功率对于提升气象服务质量、提前应对不良天气状况有重大意义。气象学家认为，变小概率事件为条件概率下的大概率事件，是提高小概率天气预报成功率的一条有效途径。以 2008 年北京奥运会期间极端天气预测为例，气象学家利用 1954-2003 年北京、青岛、秦皇岛、天津、沈阳和上海 6 座城市的高温、沙尘、雾、雷暴、大风和暴雨等日值资料，统计了几种不利天气近 50 年在上述 6 城市的发生频率，通过数据统计分析得出，极端天气事件在近 50 年中发生的频率非常低，符合发生频率接近于 5% 的小概率事件标准。再通过概率统计方法，得出了各个城市发生不利于比赛的天气事件的理论概率，为赛事组织部门安排赛事提供了一定的参考。

4.4 在日常生活中的应用

在日常生活中，很多人爱买彩票，希望能中头奖而一夜暴富。人们明知道买彩票中奖是小概率事件，还愿意买来碰运气。我们来看一个实例：据报道，河南省安阳市一位彩民用 172 元购买了 2 注 44 倍投注的“6+1”双色球彩票，竟然一次中 88 注 409.07 万（每注一等奖），共获奖金 3.599 亿！有人计算过，中双色球一等奖的概率为 0.0000000564，二等奖的概率为 0.0000008464，三等奖的概率为 0.0000091417。可见，买彩票中一等奖的概率几乎接近于零，属于典型的小概率事件。彩票行业正是利用小概率事件原理得以长久发展。

5 结束语

小概率事件原理是概率论与数理统计中的一个简单、基本而且颇为实用的原理，在各学科领域及我们的日常生活中有着非常广泛的应用。本文主要介绍了小概率事件的原理及其证明、特点和应用，希望读者朋友能对小概率事件原理有更加清晰、全面的认识和了解。在日常生活中，我们既不应该将“小概率事件”与“不可能事件”混为一谈而对其过分忽视，也不应该过度追求或避免“小概率事件”，而是应该正确对待“小概率事件”，利用小概率事件原理指导我们的生产生活，帮助人们处理工作和生活问题。

参考文献

- [1] 安国玲.浅谈小概率事件的原理及应用[J].河南机电高等专科学校学报,2010,18(04):106-108.
- [2] 杨兴民,刘华巧.小概率原理及其应用[J].中国科技信息,2006(04):34+37.
- [3] 于雪梅.小概率事件特点、原理及其应用[J].科技风,2017(16):256+261.
- [4] 徐全智,吕恕.概率论与数理统计[M].高等教育出版社,2017,8
- [5] 彭治班,吴宝俊,江剑民,王淑静.提高小概率事件预报成功率的一条途径[J].气象,2000(02):3-5+10.
- [6] 李庆祥,朱燕君,熊安元.北京等 6 城市奥运期间不利天气的概率统计[J].应用气象学报,2006(S1):42-47.

单侧假设检验矛盾性的探究

张益恺 (2018051404019)

指导教师：吕恕

【摘要】 在单侧假设检验中，常常会出现相反的矛盾性结论，为此，本文根据假设检验原理分析了矛盾出现的原因并给出了合理的解决方法。

【关键词】 单侧假设检验；矛盾性；原假设；备择假设

1 单侧假设检验基本原理

假设检验是对总体参数或总体分布作出假设，依据小概率事件实际不发生原理来判断是否拒绝假设的检验方法，通常可分为单侧检验和双侧检验，本文所讨论的单侧假设检验是作出单边原假设 H_0 和备择假设 H_1 ，然后构造检验统计量，代入数据看其是否落入拒绝域内，最后得出结论。

本文将对单侧检验中设立假设方向不同所导致的矛盾性问题进行探究。

2 引例

现有一种电子元件，要求其平均使用寿命不低于 $1000h$ 才算合格，现从一批这类元件中随机抽取 25 件，测得其使用寿命的样本平均值 $\bar{x} = 970h$ ，样本标准差 $s = 100h$ ，已知这种元件的寿命服从正态分布，请问是否可以认为这批元件合格 ($\alpha = 0.05$)？

解 (1) 由题意需假设：

$$H_0: \mu \geq 1000, \quad H_1: \mu < 1000 \quad (\mu_0 = 1000)$$

因为总体方差 σ^2 未知，所以采用 t 检验法， $T = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$ ，根据假设可知拒绝域为

$$(-\infty, -t_\alpha(n-1))$$

查表得， $t_{0.05}(24) = 1.7109$

又已知 $n = 25, \bar{x} = 970, s = 100$ ，因此

$$t = \frac{970-1000}{100/\sqrt{25}} = -1.5 > -1.7109,$$

因此不拒绝 H_0 ，认为这批元件合格。

(2) 若将假设变为

$$H_0: \mu \leq 1000, \quad H_1: \mu > 1000 \quad (\mu_0 = 1000)$$

拒绝域变为

$$t = \frac{970-1000}{100/\sqrt{25}} = -1.5 < 1.7109$$

不在拒绝域内，故接受 H_0 ，认为这批元件不合格。

可以看出两种假设方法所得到的结果产生了矛盾。

3 对矛盾的分析

3.1 矛盾产生的原因

第一种假设对应的拒绝域为

$$(-\infty, -t_{0.05}(24))$$

第二种假设对应的拒绝域为

$$(t_{0.05}(24), +\infty)$$

很显然，不会出现同时拒绝的情况，但是，同时接受的情况却可能出现，即当 t 取值在

$$(-t_{0.05}(24), t_{0.05}(24))$$

内时，同时接受两种假设的原假设 H_0 ，便出现了引例中的矛盾情况。

类比可得，单侧检验矛盾性普遍存在于标准正态分布， χ^2 分布， t 分布和 F 分布中，如下表所示：

表 1 正态分布下拒绝域的分布

U	$(-\infty, -u_\alpha)$	$(-u_\alpha, u_\alpha)$	$(u_\alpha, +\infty)$
左侧检验	拒绝	不拒绝	不拒绝
右侧检验	不拒绝	不拒绝	拒绝

表 2 χ^2 分布下拒绝域的分布

χ^2	$(0, \chi^2_{1-\alpha})$	$(\chi^2_{1-\alpha}, \chi^2_\alpha)$	$(\chi^2_\alpha, +\infty)$
左侧检验	拒绝	不拒绝	不拒绝
右侧检验	不拒绝	不拒绝	拒绝

表 3 t 分布下拒绝域的分布

t	$(-\infty, -t_\alpha)$	$(-t_\alpha, t_\alpha)$	$(t_\alpha, +\infty)$
左侧检验	拒绝	不拒绝	不拒绝
右侧检验	不拒绝	不拒绝	拒绝

表 4 F 分布下拒绝域的分布

F	$(0, F^2_{1-\alpha})$	$(F^2_{1-\alpha}, F^2_\alpha)$	$(F^2_\alpha, +\infty)$
左侧检验	拒绝	不拒绝	不拒绝
右侧检验	不拒绝	不拒绝	拒绝

3.2 对矛盾的解释

假设检验依据的是小概率事件实际不发生的推断原理，因此拒绝原假设比较可信，而接受原假设需要谨慎。

对于上述两种假设，不妨这样理解：

假设（1）中，原假设 $H_0: \mu \geq 1000$ ，因为没有进入拒绝域，所以结论是：没有充分的理由拒绝平均寿命大于 $1000h$ 的假设，即认定这批元件的平均寿命不大于 $1000h$ 的理由不充分。

假设（2）中，原假设 $H_0: \mu \leq 1000$ ，因为没有进入拒绝域，所以结论是：没有充分的理由拒绝平均寿命小于 $1000h$ 的假设，即认定这批元件的平均寿命不小于 $1000h$ 的理由不充分。

这种解释并没有简单地将原假设 H_0 跟备择假设 H_1 看作对立事件，其实不拒绝原假设仅仅说明

在现有的条件下不能轻易拒绝它，我们常常在这种情况下接受它，是为了做出选择，得出结论，而且是不严谨的（除非原假设本身就非常可信，这将在后面提到）。上述解释立足于假设检验的基本原理，落脚于“拒绝比较可信，接受需要谨慎”的思想，避免了结论的矛盾性，且这样的表述更为严谨。

4 矛盾的解决方法

4.1 调整显著性水平

在实际情况允许的情况下，可以选用调整显著性水平 α 的方式来解决矛盾性问题。

对标准正态分布而言，矛盾产生的区间为 $(-u_\alpha, u_\alpha)$ ，增大 α ， u_α 减小，产生矛盾性区间长度缩短，有利于消除矛盾。

对 χ^2 分布而言，增大 α ， $\chi^2_{1-\alpha}$ 增大， χ^2_α 减小， $(\chi^2_{1-\alpha}, \chi^2_\alpha)$ 区间长度缩短。

对 t 分布而言，增大 α ， t_α 减小， $-t_\alpha$ 增大， $(-t_\alpha, t_\alpha)$ 区间长度缩短。

对 F 分布而言，增大 α ， F^2_α 减小， $F^2_{1-\alpha}$ 增大， $(F^2_{1-\alpha}, F^2_\alpha)$ 区间长度缩短。

综上所述，在调整显著性水平时，需要增大显著性水平 α ，才能减小矛盾的影响。值得一提的是，不可以随意改变 α 。增大 α 则会增加犯第一类错误“弃真”的概率，减小 α 则会增加犯第二类错误“纳伪”的概率，因此，在改变 α 的时候，要结合实际情况对“弃真”和“纳伪”两类错误的后果进行评估，例如，一家制药厂在生产药品的时候，如果“弃真”最多会承受一定的经济损失，相比之下，如果“纳伪”，则有可能危机他人的生命，因此，在这种情况下，万万不可减小 α 。

4.2 增大样本容量

增大样本容量可以获得更加逼近真实统计量的样本数据，这种方法同时还能减小犯两类错误的概率，如果有条件的话应为首选。

4.3 科学地设定原假设和备择假设（适用于不同的情况）

4.3.1 针对题目没有单侧结论的明确要求

当题目没有明确要求得出单侧结论时，可以认为在 $(-u_\alpha, u_\alpha)$ 内， $\mu = \mu_0$ （以标准正态分布为例）。对应假设为：

$$H_0: \mu = \mu_0 \quad H_1: \mu \geq \mu_0 \quad H_2: \mu \leq \mu_0$$

表 5 H_0 、 H_1 、 H_2 三种假设下的拒绝和接受情况

U	$(-\infty, -u_\alpha)$	$(-u_\alpha, u_\alpha)$	$(u_\alpha, +\infty)$
$H_1: \mu \geq \mu_0$	拒绝 H_1	接受 H_0	接受 H_1
$H_2: \mu \leq \mu_0$	接受 H_2	接受 H_0	拒绝 H_2

可以看出，这样规定可以有效避免矛盾性问题，但是需要指出的是，如果使用双边检验，设立假设：

$$H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0$$

规定显著性水平为 α ， $(-\frac{u_\alpha}{2}, \frac{u_\alpha}{2})$ 内接受 H_0 ，这时犯第一类错误“弃真”的概率为 α 。参照上文， $(-u_\alpha, u_\alpha)$ 内接受 H_0 ，说明上述三种假设的情况下对应的显著性水平为 2α ，犯第一类错误“弃真”的概率为 2α ，也就是说，这样做，如果接受 H_0 ，则增加了“弃真”的概率。因此在实际情况中，如果“弃真”后果十分严重的话，不建议采用这种方法。

4.3.2 针对题目有单侧结论的明确要求

(1) 在对两种情况没有明显的可信度倾向时, 根据题目信息, 把不易发生的情况作为原假设(小概率事件原理), 这样与拒绝假设的可能性会大。具体操作是, 根据已知信息, 将与已知相矛盾的假设设为原假设, 这样一来就充分利用了已知信息, 也将原假设往“可信的拒绝”方向推进。

如上述例题中, 已知样本平均值 $\bar{x} = 970h$, 故令原假设为 $H_0: \mu \geq 1000$, 与已知相矛盾, 倾向于拒绝它, 得到可信的结论。

对此进行推广, 可得:

表 6 与已知相矛盾的原假设设法

已知情况	原假设 H_0
$\bar{X} > \mu_0$	$\mu \leq \mu_0$
$\bar{X} < \mu_0$	$\mu \geq \mu_0$
$S^2 > \sigma^2_0$	$\sigma^2 \leq \sigma^2_0$
$S^2 < \sigma^2_0$	$\sigma^2 \geq \sigma^2_0$
$\bar{X}_1 - \bar{X}_2 > 0$	$\mu_1 - \mu_2 \leq 0$
$\bar{X}_1 - \bar{X}_2 < 0$	$\mu_1 - \mu_2 \geq 0$
$\frac{S^2_1}{S^2_2} > 1$	$\frac{\sigma^2_1}{\sigma^2_2} \leq 1$
$\frac{S^2_1}{S^2_2} < 1$	$\frac{\sigma^2_1}{\sigma^2_2} \geq 1$

(2) 把不能轻易否定的情况设为原假设, 由于接受原假设的结论需要谨慎对待, 因此把不能轻易否定的情况放在原假设, 如果没有充分的理由拒绝它, 那么接受它时也会显得比较可信(本身就不易拒绝)。

例如, 某厂商生产的产品合格率多年以来总是维持在 95%以上, 今年与往年相比没有什么大的不同, 故今年对产品进行假设检验的时候就可以把原假设设为 $H_0: \mu \geq 95\%$, 这样如果拒绝了 H_0 , 依据假设检验的小概率事件原理, 得出今年产品合格率低于 95%的结论比较可信, 如果未能拒绝 H_0 , 由于多年来的经验, 再加上外部因素未改变, 接受 H_0 也是比较可信的。

(3) 把等待验证的情况设为备择假设, 因为这类情况通常要求把犯错的概率控制在非常小的区间内, 由于拒绝原假设是依据小概率事件原理, 因此“拒绝比较可信”, 备择假设为真的结论可信度就比较高, 这样一来, 验证的结论才不容易出错。

5 总结

假设检验的矛盾性在日常生活中是普遍存在的, 对此, 我们要正确地理解它, 还要科学地设定原假设跟备择假设来合理地解决矛盾性, 从而得到更可信的结论。

参考文献

- [1] 徐全智, 吕怒 概率论与数理统计, 高等教育出版社, 2017
- [2] 龚辉锋, 单侧检验中假设的提法研究, 沈阳大学学报, 2006, 第 18 卷第 2 期: 102-106
- [3] 黄艳华, 单侧假设检验结论的矛盾性及其解决方法, 数学学习与研究, 2018 (03): 150
- [4] 黄发贵, 单侧假设检验中备择假设的设定依据, 理论新探, 2006 年第 4 期: 12-14
- [5] 王凌峰, 陈灿, 假设检验中的 H_0-H_1 交换矛盾及其解决, 统计与咨询, 2013.3: 19-20
- [6] 徐兆强, 单侧检验中的零假设应当“反设”, 工科数学, 1993.10.1: 90-91

成电讲坛次数概率模型的参数假设检验及解释

陈圣坤 (2018051410012)

指导教师：彭江艳

【摘要】 概率论是研究随机现象数量规律的数学分支。运用概率论，我们可以解决很多现实生活中遇到的问题。本篇主要利用 t 分布和参数假设检验对“2018-2019 学年大一学生成电讲坛观看次数概率模型”进行参数假设检验和模型分析。

1 t 分布

在概率论和统计学中，t-分布 (t-distribution) 用于根据小样本来估计呈正态分布且方差未知的总体的均值。如果总体方差已知（例如在样本数量足够多时），则应该用正态分布来估计总体均值。

t 分布曲线形态与 n (确切地说与自由度 df) 大小有关。与标准正态分布曲线相比，自由度 df 越小，t 分布曲线愈平坦，曲线中间愈低，曲线双侧尾部翘得愈高；自由度 df 愈大，t 分布曲线愈接近正态分布曲线，当自由度 $df=\infty$ 时，t 分布曲线为标准正态分布曲线。

由于在实际工作中，往往 σ 是未知的，常用 s 作为 σ 的估计值，为了与 u 变换区别，称为 t 变换，统计量 t 值的分布称为 t 分布。

假设 X 服从标准正态分布 $N(0,1)$ ， Y 服从 $\chi^2(n)$ 分布，那么 $z = \frac{X}{\sqrt{Y/n}}$ 的分布称为自由度为 n 的 t 分布，记为 $Z \sim t(n)$ 。分布密度函数为 $f_z(x) = \frac{\text{Gam}\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \text{Gam}\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$ 。其

中， $\chi^2(n)$ 为伽马函数。

2 参数假设检验

参数假设检验 (parametric hypothesis test) 简称假设检验，是数理统计学中根据一定假设条件由样本推断总体的一种方法。

其基本原理是先对总体的特征作出某种假设，然后通过抽样研究的统计推理，对此假设应该被拒绝还是接受作出推断。

【关键词】 t 分布;参数假设检验

1 引言

成电讲坛作为每名成电学子的必修课，其完成质量往往反映了学生对于大学生活的适应程度，以及其对于学习的态度。因此，我们对“2018-2019 学年大一学生成电讲坛观看次数”进行概率建模分析，希望借此了解大一学生的学习态度、对大学生活适应程度，为学校引导学生适应大学生活，不断提升自己提供可靠参考。

2 成电讲坛规则简述

(1) 本科学生至少选修 1 个成电讲坛学分，至多选修 2 个成电讲坛学分。且学校鼓励本科生在大一学年修够至少 1 个该项学分。

(2) 一学年内观看 6 场成电讲坛, 可以兑换 1 个学分; 观看 10 场及以上成电讲坛, 可以兑换 2 个学分。

3 “2018-2019 学年大一学生成电讲坛观看次数” 数据

表 1 2018-2019 学年大一学生成电讲坛观看次数统计表

打卡次数	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
人数	435	922	457	235	109	720	494	152	40	289	206	81	19	9	4	2

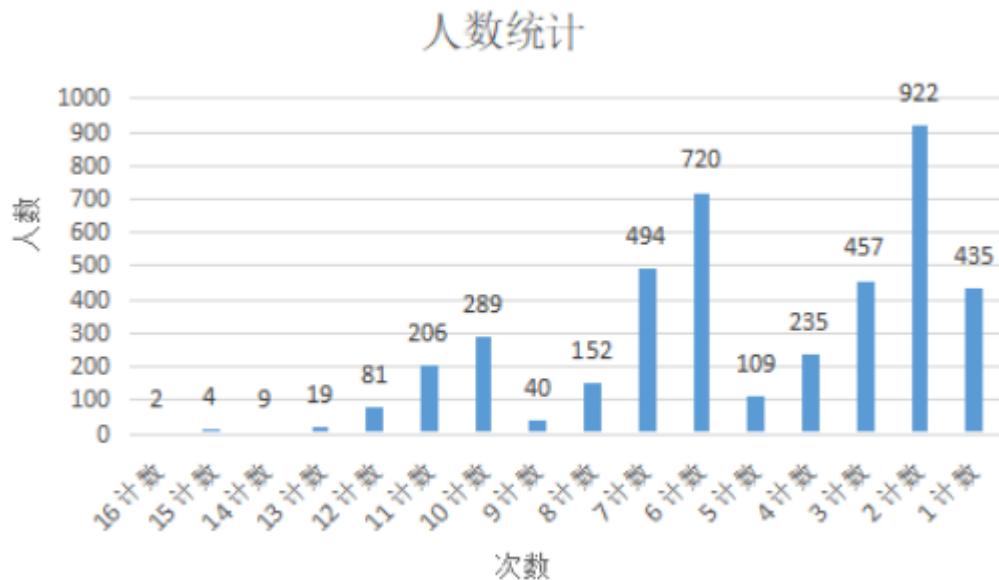


图 1 2018-2019 学年大一学生成电讲坛观看次数统计直方图

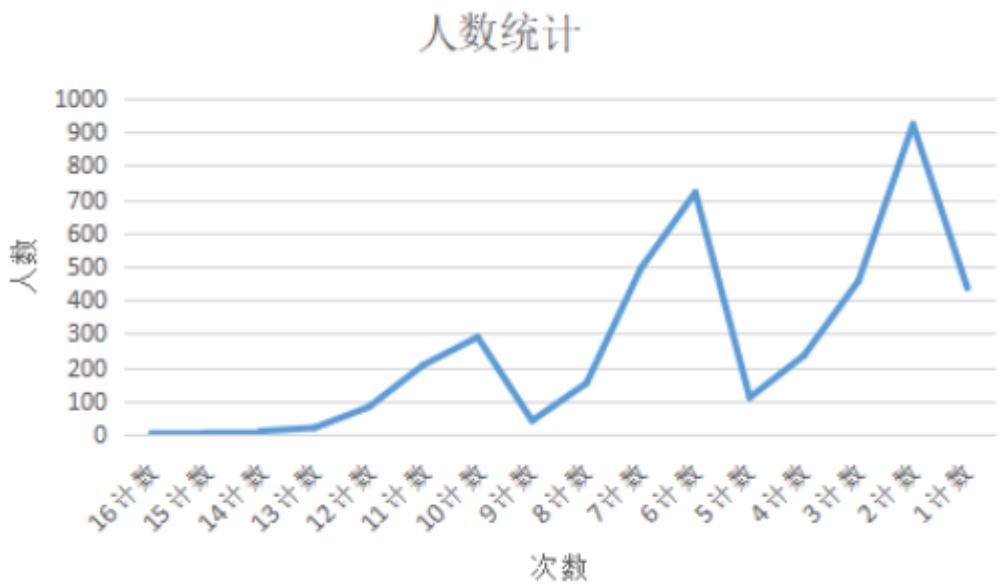


图 2 2018-2019 学年大一学生成电讲坛观看次数统计折线图

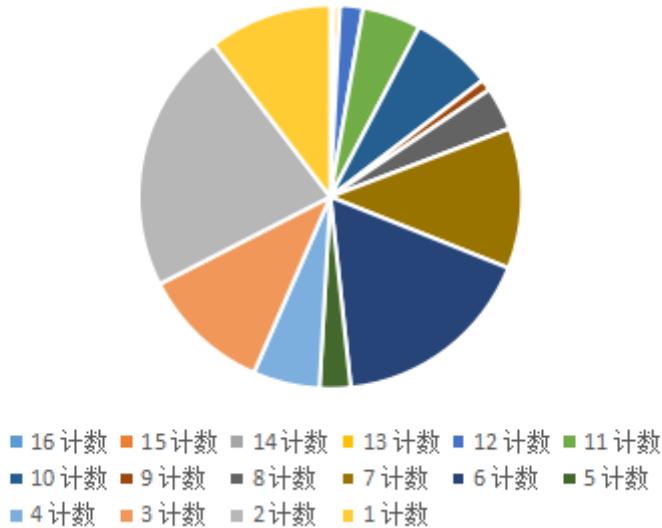


图 3 2018-2019 学年大一学生成电讲坛观看次数统计饼状图

4 分析及建模

由数据及图表可知，人数集中在“观看 2 次”、“观看 6 次”和“观看 10 次”。不妨假设该模型由三个 t 分布组成：

- (1) “观看 1-4 次” 服从： $\mu=2.24$ 的 t 分布
- (2) “观看 5-8 次” 服从： $\mu=6.47$ 的 t 分布
- (3) “观看 9-16 次” 服从： $\mu=10.7$ 的 t 分布

5 参数假设检验

(1) 观看 1-4 次时： $H_0: \mu=2.24$; $H_1: \mu \neq 2.24$ 。

$$\bar{x} = (1 \times 435 + 2 \times 922 + 3 \times 457 + 4 \times 235) / (435 + 922 + 457 + 235) = 2.24011,$$

$$n = 435 + 922 + 457 + 235 = 2049,$$

$$S^2 = (1 \times 1 \times 435 + 2 \times 2 \times 922 + 3 \times 3 \times 457 + 4 \times 4 \times 235 - 2049 \times 2.24011 \times 2.24011) / (2049 - 1) = 0.8369.$$

当 H_0 成立时： $T = (\bar{x} - \mu) / [S / \sqrt{n-1}] \sim t(2048)$, $2048 \geq 45$, 近似服从标准正态分布。

此时： $|T| = 0.0058 \leq 1.96$, $u_{0.025} = 1.965$, 所以可以在 $\alpha=0.05$ 的情况下接受 H_0 。

(2) 观看 5-8 次时： $H_0: \mu=6.47$; $H_1: \mu \neq 6.47$ 。

$$\bar{x} = (5 \times 109 + 6 \times 720 + 7 \times 494 + 8 \times 152) / (109 + 720 + 494 + 152) = 6.46712,$$

$$n = 109 + 720 + 494 + 152 = 1475,$$

$$S^2 = (5 \times 5 \times 109 + 6 \times 6 \times 720 + 7 \times 7 \times 494 + 8 \times 8 \times 152 - 1475 \times 6.46712 \times 6.46712) / (1475 - 1) = 0.6032.$$

当 H_0 成立时： $T = (\bar{x} - \mu) / [S / \sqrt{n-1}] \sim t(1474)$, $1474 \geq 45$, 近似服从标准正态分布。

此时： $|T| = 0.1424 \leq 1.96$, $u_{0.025} = 1.965$, 所以可以在 $\alpha=0.05$ 的情况下接受 H_0 。

(3) 观看 9-16 次时： $H_0: \mu=10.7$; $H_1: \mu \neq 10.7$ 。

$$n = 40 + 289 + 206 + 81 + 19 + 9 + 4 + 2 = 650,$$

$$\bar{x} = (9 \times 40 + 10 \times 289 + 11 \times 206 + 12 \times 81 + 13 \times 19 + 14 \times 9 + 15 \times 4 + 16 \times 2) / 650 = 10.69692,$$

$$S^2 = (9 \times 9 \times 40 + 10 \times 10 \times 289 + 11 \times 11 \times 206 + 12 \times 12 \times 81 + 13 \times 13 \times 19 + 14 \times 14 \times 9 + 15 \times 15 \times 4 + 16 \times 16 \times 2) / (650 - 1) = 10.69692.$$

$$10.69692 \times 10.69692 \times 650) / (650-1) = 1.14223.$$

当 H_0 成立时: $T = (\bar{x} - \mu) / [S / \sqrt{(n-1)}] \sim t(649)$, $649 \geq 45$, 近似服从标准正态分布。

此时: $|T| = 0.0734 \leq 1.96$, $u_{0.025} = 1.965$, 所以可以在 $\alpha=0.05$ 的情况下接受 H_0 。

6 模型分析

(1) “观看 6 次”与“观看 10 次”分别是取得“1 个学分”和“2 个学分”的“最小观看次数”，在次数 1-16 中，“6”和“10”是最优的次数（达到相同目的所消耗的最少成本），在认为每名同学都是理性决策者的前提下，每个人的观看讲座次数都会尽力向这两个次数靠拢，所以得到 $\mu=6.47$ 和 $\mu=10.7$ 的 t 分布。

(2) 受限于时间和学习能力，观看讲座次数分组（观看 1-4 次，观看 5-8 次，观看 9-16 次）人数依次递减。

(3) 根据数据我们还得到了 $\mu=2.24$ 的 t 分布。我们可以认为，这部分（观看 1-4 次）同学对学习缺乏持续的热情，对大学生活的适应程度有待提高。

7 总结

根据“2018-2019 学年大一学生成电讲坛观看次数统计数据”，我们可以认为其大致由 $\mu=6.47$, $\mu=10.7$ 和 $\mu=2.24$ 这三个 t 分布组成。由此看出：18 级电子科技大学同学大一阶段观看成电讲坛情况并不理想，有半数同学成电讲坛次数未达到最低学分要求，这部分同学将会占用之后成电讲坛的名额，导致新一届大一学生无法在大一选修完成电讲坛学分。同时我们可以看出：18 级电子科技大学学生对于大学生活的适应程度还有待加强，学习态度也有待改进。建议学校加大对 18 级本科生的管理、引导，开设更多的成电讲坛，设立针对 18 级同学的专场成电讲坛，帮助 18 级同学尽早修满成电讲坛学分，并在这个过程中找到适合自己的生活学习节奏，不断提升自己。

参考文献

- [1] 徐全智,吕怒.概率论与数理统计, 高等教育出版社, 2016 年
- [2] 关于印发《电子科技大学“成电讲坛”及“成电舞台”认定通识教育学分办法的通知》, 电子科技大学, 2016 年 7 月
- [3] 成电讲坛次数统计 (20180501-20190430), 电子科技大学, 2019 年 5 月

利用最小二乘法拟合物理实验数据

张新宇 (2018120702008)

指导教师：彭江艳

【摘要】 实验数据的处理十分重要，利用概率论与数理统计中的“最小二乘法”对需要得到数值变化规律的数据进行处理，可以得到数据结果直观的变化规律，并从最小二乘法得到的回归方程预测因变量随自变量的变化规律，从而对实验数据的处理带来许多便利。

【关键词】 最小二乘法；线性回归方程；异常值处理

1 引言

实验是检验真理的唯一标准，在学习理论知识之余，进行恰当且符合的实验时验证理论并帮助掌握理论的有效手段。以大学物理实验为例，除了要求完善的实验装置、步骤设计和正确的实验过程以外，恰当的实验数据处理也十分重要。由于实验的误差范围较大，在处理数据的过程中往往遇到许多困难，如果我们将概率论与数理统计的知识运用到这里，实验数据的处理可能会方便许多。

2 实际问题及分析

本人正在参加中国大学生物理学术竞赛 (CUPT)，该竞赛中第七题：

将当把底部有孔的碗放在水中时，碗会往下沉。撒克逊人曾用这个装置来进行计时。研究确定下沉时间的参数。

经过粗略的理论分析（涉及物理知识此处不详细列举），将该碗放入水中后，该碗应呈加速，再缓慢减速的下沉趋势。整个下沉过程中，下沉速度很小，减速过程的加速度也很小。

3 实验数据结果

为了验证第二段减速过程，我们设计了相关实验，并获得如下数据。同时得到两次间隔时间的时间差 Δt ：

表 1 第一次实验时间差

实验序号	第一次实验时间差/s
1	2.27
2	1.97
3	2.30
4	2.10
5	1.97
6	1.93
7	2.50
8	2.40
9	2.20
10	2.40
11	2.33
12	2.13

实验序号	第一次实验时间差/s
13	2.90
14	2.40
15	2.30
16	2.53
17	2.67
18	2.60
19	2.50
20	2.83

根据以上数据在 excel 中绘制图表，得到下图的 $\Delta t-n$ 的散点图

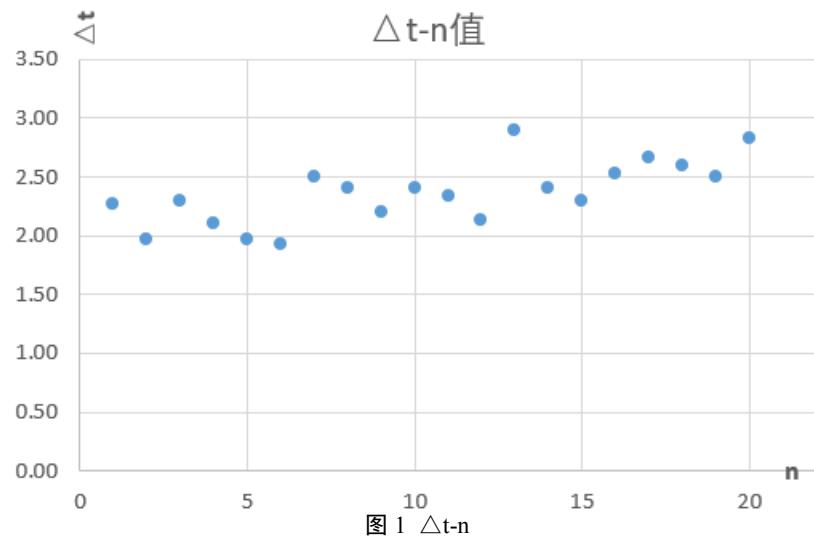


图 1 $\Delta t-n$

考虑到实验具有偶然性，我们重复了四次以上实验，并取每次实验 Δt 的平均值，同时按照以上绘图方法绘制 $\Delta t-n$ 散点图，如下所示

表 2 水面上涨 1mm 所需平均时间

实验序号	第一次实验时间差/s	第二次实验时间差/s	第三次实验时间差/s	水面上涨 1mm 所需平均时间(s)/s
1	2.27	2.07	2.23	2.19
2	1.97	2.03	2.10	2.03
3	2.30	1.93	2.33	2.19
4	2.10	2.70	2.07	2.29
5	1.97	1.87	2.10	1.98
6	1.93	2.50	2.27	2.23
7	2.50	2.07	1.97	2.18
8	2.40	2.60	2.27	2.42
9	2.20	2.43	2.63	2.42
10	2.40	2.30	2.33	2.34
11	2.33	2.17	1.97	2.16
12	2.13	2.07	2.37	2.19
13	2.90	2.80	2.40	2.70
14	2.40	2.63	2.57	2.53

实验序号	第一次实验时间差/s	第二次实验时间差/s	第三次实验时间差/s	水面上涨 1mm 所需平均时间(s)/s
15	2.30	2.40	2.37	2.36
16	2.53	2.17	2.50	2.40
17	2.67	2.63	2.13	2.48
18	2.60	2.43	2.17	2.40
19	2.50	2.27	2.87	2.54
20	2.83	2.40	1.93	2.39

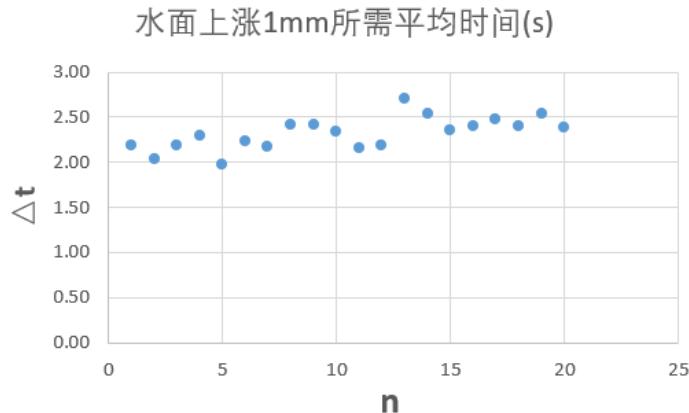


图 2 $\Delta t(\text{平均})-n$ 图像

4 利用最小二乘法对数据进行拟合

根据最小二乘法的定义：

设已知变量 X,Y 的 n 对观测值 $(x_i, y_i), i=1, 2, \dots, n$, 我们用最小二乘法求参数 a、b 的估计值.

记 a,b 的点估计为 \hat{a}, \hat{b} , 则回归函数 $a+bx$ 被估计为 $\hat{a}+\hat{b}x$, 在 $X=x_i$ 处, 因变量 Y 的值被估计为 $\hat{y}_i=\hat{a}+\hat{b}x_i, i=1, 2, \dots, n$, 而实际观测已得到 $X=x_i$ 处 Y 的值 y_i , 这样就有偏离 $y_i - \hat{y}_i, i=1, 2, \dots, n$, 称为在 $X=x$ 处的残差, 作残差平方和:

$$Q = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i)^2$$

残差平方和越小越好, 于是考虑选择使 $Q=\sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i)^2$ 取最小值的 \hat{a}, \hat{b} 作为 a,b 的估计值, 该方法成为最小二乘法。

记 $\hat{y}=\hat{a}+\hat{b}x$, 该方程称为 Y 对 X 的经验回归方程。

通过上述方法对 3 中得到的数据进行最小二乘法拟合。过程如下:

$$\begin{aligned} n &= 20, \text{计算出 } \bar{x} = 10.5, \quad \sum_{i=1}^{20} x_i^2 = 2870; \quad \bar{y} = 2.32, \quad \sum_{i=1}^{20} y_i^2 = 108.3632 \\ \sum_{i=1}^{20} x_i y_i &= 500.8586 \end{aligned}$$

故得:

$$l_{xx} = \sum_{i=1}^{20} x_i^2 - n\bar{x}^2 = 2870 - 20 \times 10.5^2 = 665$$

$$l_{xy} = \sum_{i=1}^{20} x_i y_i - n\bar{x}\bar{y} = 500.8586 - 20 \times 10.5 \times 2.32 = 13.6586$$

$$l_{yy} = \sum_{i=1}^{20} y_i^2 - n\bar{y}^2 = 108.3632 - 20 \times 2.32^2 = 0.7152$$

于是有

$$\hat{b} = \frac{l_{xy} - l_{yy}}{l_{xx}} = \frac{13.6586}{665} \approx 0.0205$$

而 $\hat{a} = \bar{y} - \hat{b} * \bar{x} = 2.32 - 0.0205 \times 10.5 = 2.1048$

得到 $\hat{y} = 0.0205x + 2.1048$

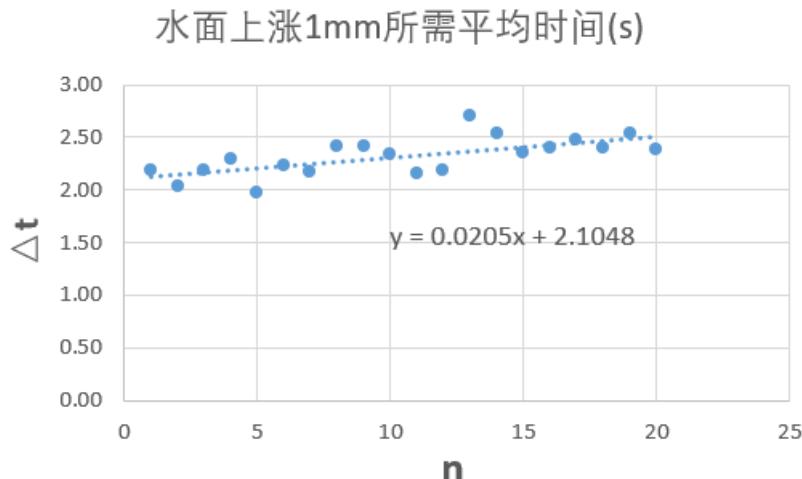


图 3 $\Delta t-n$ 线性拟合图像

在图像中拟合得到上图的图线。

可以看到图像的斜率为正值， Δt 随 n 呈增加趋势，根据物理中的速度-位移公式($x=v \cdot \Delta t$)，易知速度随 n 呈减小趋势，即此部分速度减小，与理论计算结果相符。

5 拟合图线的显著性分析

下面利用相关系数检验法对该回归方程的显著性进行检验：

$$R = \frac{l_{xy}}{\sqrt{l_{xx}}\sqrt{l_{yy}}} = \frac{13.6586}{\sqrt{665}\sqrt{0.7152}} \approx 0.6263, n=20, \text{查表得 } R_{0.01(18)}=0.561, \text{因为 } R > R_{0.01(18)}, \text{所以得到的}$$

Δt 与 n 的关系显著。

6 总结

通过以上分析，我们得到了实验结果的线性回归方程，并通过方程画出了对应的一元函数图像，得以直观地判断 Δt 与 n 的关系，与理论计算结果相符合。

对于散点和回归方程的图像，我们观察到有少数离散的点与图线相隔较远，在网上查阅文献后，发现这类差异较大，或出现“跳变”的点，叫做异常点。考虑此次实验的实际性，并询问物理老师后，得知只要得出速度在逐渐变小即可。

综上所述，利用数理统计中的“最小二乘法”对实验数据进行分析，得到了与理论计算结果符合的回归方程图线。

参考文献

- [1] 徐全智, 吕恕. 概率论与数理统计, 高等教育出版社, 2004
- [2] 王玉梅, 线性回归模型中的异常点分析, 2008, 2-8

伴随矩阵的特征值与特征向量

刘淏楠 (2019021407027) 陈阳 (2019021407018)

指导教师: 王博

【摘要】 我们给出了伴随矩阵的特征值和特征向量的规律, 此结论可以作为工科大学线性代数教学的一个合理补充。

【关键词】 特征值; 特征向量; 伴随矩阵; 秩

1 引言

当计算伴随矩阵特征值和特征向量时, 按照教材^[1]中的方法, 我们需要先计算出对应的伴随矩阵, 再计算其特征值和特征向量, 此计算方法较为复杂。本文在工科线性代数的知识范畴内, 给出了当矩阵的秩在不同情况下对应伴随矩阵的特征值与特征向量的规律。在已有规律的基础上解决上述问题时, 可以先求出已知矩阵的特征值和特征向量, 再通过规律得到对应伴随矩阵的特征值和特征向量, 大大减少了计算量。

2 伴随矩阵的特征值和特征向量

在已知矩阵的特征值和特征向量的情况下, 可以通过伴随矩阵和特征值的相关知识推导出如下定理:

定理: 设 A 为 n 阶方阵, 且 $A\alpha = \lambda\alpha (\alpha \neq 0)$, 则对于伴随矩阵 A^* 有

(1) 当 $R(A) = n$ 时, $A^*\alpha = \frac{\det(A)}{\lambda}\alpha$ 。

(2) 当 $R(A) < n - 1$ 时, A^* 的特征值为0, 特征向量为任意非零列向量。

(3) 当 $R(A) = n - 1$ 时, A^* 有 $n - 1$ 重特征值0, 特征向量为 A 列向量的最大无关组; 1重特征值 $\text{tr}(A^*)$, 即 $A^*\beta = \text{tr}(A^*)\beta$, 其中特征向量 β 为 A^* 的任意一个非零列向量。

证明: 我们对三种情况进行逐一证明。

(1) 当 $R(A) = n$ 时, 因为 A 可逆且 $\lambda \neq 0$, 则有

$$A^{-1}\alpha = \frac{1}{\lambda}\alpha,$$

根据 A 与 A^* 的关系 $A^*A = \det(A)I$, 可以推出

$$A^*\alpha = \det(A)/\lambda\alpha.$$

(2) 当 $R(A) < n - 1$ 时, 方阵 A 的所有 $n - 1$ 阶子式的皆为0, 则 $A^* = 0$ 。所以 A^* 有 n 重特征值0, 特征向量为任意非零列向量。

(3) 当 $R(A) = n - 1$ 时, 方阵 A 为降秩矩阵, $\det(A) = 0$ 。

由 $A^*A = \det(A)I = 0$, 可知, A 的列向量是方程 $A^*X = 0$ 的解。又因为 A 列向量最大无关组有 $n - 1$ 个列向量, 且 $A^* \neq 0$, 所以 A^* 有 $n - 1$ 重特征值均为0, 特征向量为 A 列向量的最大无关组。而由

$$AA^* = \det(A)I = 0,$$

可知, A^* 的列向量为 $Ax = 0$ 的解。因为 $n - R(A) = 1$, 所以 $R(A^*) \leq n - R(A) = 1$ 。又因为 $A^* \neq 0$, 所以 $R(A^*) \geq 1$ 。故有

$$R(A^*) = 1.$$

所以，可设 A^* 的一个非零列为 β ，且存在一个列向量 γ 使得 $A^* = \beta\gamma^T$ 。进而有，

$$A^*\beta = (\beta\gamma^T)\beta = (\gamma^T\beta)\beta.$$

而 $\gamma^T\beta = \text{tr}(A^*)$ ，代入上式可得

$$A^*\beta = \text{tr}(A^*)\beta.$$

即， $\text{tr}(A^*)$ 为 A^* 的特征值， β 为对应的特征向量。

综上，定理证毕。

我们需要注意，如果需要求出 A^* 来找 β ，计算量仍然较大。而 β 是 A 的任意一行元素对应子式不全为零时，取这一行的所有元素对应的子式组成的列向量。找出 A 中该行简便的方法是找到一个 A 的 $n-1$ 阶非零子式，若该非零子式是 a_{ij} 的子式，则 A 的第 i 行即为所求行。

3 例题：以四个例子对定理加以验证

例 1 当 $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ ，求解 A^* 的特征值和特征向量

解：由 $|\lambda I - A| = 0$ 解得 A 的特征值

当 $\lambda_1 = 1$ 时， $I - A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ ；所以方程 $(\lambda I - A)\alpha = 0$ 该方程解得
 $\alpha_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ ，同理可知 $\alpha_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ ， $\alpha_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ 。

根据 $AA^* = \det(A)I$ 可以解得

$$A^* = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

此时显然

$$\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 2,$$

再由 $(\lambda I - A^*)\alpha = 0$ ，可以解得

$$\alpha_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \alpha_5 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \alpha_6 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

即该结果与上述定理(1)所得结论一致。

例 2 当 $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ，求解 A^* 的特征值和特征向量

解：由 $|\lambda I - A| = 0$ 解得特征值

$$\lambda_1 = 1, \lambda_2 = 0(\text{二重});$$

根据 $AA^* = \det(A)I$ ，得

$$A^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

再次使用 $|\lambda I - A^*| = 0$ ，计算得到 A^* 的特征值

$$\lambda = 0(\text{三重})。$$

即该结果与上述定理(2)所得结论一致。

例 3 当 $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ，求解 A^* 的特征值。

解：由 $|\lambda I - A| = 0$ 解得 $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 0$ ，由 $(\lambda I - A)\alpha = 0$ 解得

$$\alpha_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

再利用 $AA^* = \det(A)I$ ，解得

$$A^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

解出 A^* 的特征值为

$$\lambda_1 = 2, \lambda_2 = 0(\text{二重})。$$

即该结果与上述定理(3)所得结论一致。

例 4 当 $A = \begin{bmatrix} 1 & 3 & 2 & 5 & 2 \\ 2 & 7 & 8 & 9 & 4 \\ 5 & 1 & 3 & 3 & 0 \\ 6 & 2 & 4 & 1 & 1 \\ 3 & 10 & 10 & 14 & 6 \end{bmatrix}$ ，求解 A^* 的特征值和特征向量。

该矩阵为五阶方阵，若使用教材中提供的方法求解较为复杂，而使用上文中所提到得定理III可大大减小计算量： A^* 的特征值 $\lambda_1 = \text{tr}(A^*) = -278$ ，而 a_{11} 所对应得子式 $A_{11} = 4$ 不等于零，所以 A 的

第一行每一个元素所对应的子式所构成的列向量 $\alpha_1 = \begin{bmatrix} 4 \\ 184 \\ -54 \\ -32 \\ -144 \end{bmatrix}$ 是 λ_1 对应的特征向量。而 $\lambda_2 = 0$ (四重

)对应特征向量为 A 的列向量最大无关组，为

$$\alpha_2 = \begin{bmatrix} 1 \\ 2 \\ 5 \\ 6 \\ 3 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} 3 \\ 7 \\ 1 \\ 2 \\ 10 \end{bmatrix}, \quad \alpha_4 = \begin{bmatrix} 2 \\ 8 \\ 3 \\ 4 \\ 10 \end{bmatrix}, \quad \alpha_5 = \begin{bmatrix} 5 \\ 9 \\ 3 \\ 1 \\ 14 \end{bmatrix}$$

4 总结

本文通过伴随矩阵和特征值的相关知识，回答了工科大学生在学习特征值和特征向量时经常提出了一个疑问，即已知矩阵的伴随矩阵特征值和特征向量如何简易求解。证明了所有情况下已知矩阵特征值和特征向量与其对应伴随矩阵特征值和特征向量之间的关系：

若 $R(A) = n$ 时， $A^*\alpha = \frac{\det(A)}{\lambda}\alpha$ 。

若 $R(A) < n - 1$ 时， A^* 的特征值为0，特征向量为任意非零列向量

若 $R(A) = n - 1$ 时， A^* 有 $n - 1$ 重特征值0，特征向量为 A 列向量的最大无关组；1重特征值 $\text{tr}(A^*)$ ，即 $A^*\beta = \text{tr}(A^*)\beta$ ，其中 β 为 A^* 的任意一个非零列向量，即 A 的任意一行元素对应子式不全为零时，该行所有元素对应子式所组成的列向量。

已知上述规律后，可以先计算已知矩阵特征值和特征向量，再通过上述规律得到对应伴随矩阵的特征值和特征向量，大大减少了计算量。

参考文献

- [1] 黄廷祝, 成孝予, 线性代数与空间解析几何, 高等教育出版社, 2018

求函数在特殊区间的最值

明安杰 (2019091615015) 袁茵 (2019091615026)

指导教师: 张利琼

【摘要】 教材中介绍的最值定理要求区间是有限闭区间, 而有限闭区间是连续函数取得最值的充分非必要条件。本文讨论在有限开区间及无穷区间上连续函数取得最值的条件及结论。

【关键词】 有限闭区间; 有限开区间; 无穷区间; 最值

教材中的最值定理内容如下: 有限闭区间上的连续函数一定存在最大值与最小值。

最值定理中的有界闭区间是充分而非必要条件, 在有限开区间及无限区间上的连续函数也可以取到最值。下面分别进行讨论。

1 有限开区间上连续函数的最值

引理 1 设 $\lim_{x \rightarrow x_0} f(x) = a$, $\lim_{x \rightarrow x_0} g(x) = b$, 若 $a < b$, 则存在 x_0 的某个去心邻域, 在此邻域内恒有 $f(x) < g(x)$.

引理 1 称为极限的局部保序性.

定理 1 设函数 $f(x)$ 在 (a, b) 内连续, 且 $\lim_{x \rightarrow a^+} f(x) = A$, $\lim_{x \rightarrow b^-} f(x) = B$, 在 (a, b) 内 $f(x)$ 能取到比 A, B 都大的数, 则 $f(x)$ 在 (a, b) 上能取到最大值.

证明 由已知可得, $\exists x_0 \in (a, b)$, 使得 $f(x_0) > A, f(x_0) > B$

由引理 1 知, $\exists \delta_1 > 0, a + \delta_1 < x_0$, 对 $\forall x \in (a, a + \delta_1)$, 都有 $f(x) > f(x_0)$ (1)

$\exists \delta_2 > 0, b - \delta_2 > x_0$, 对 $\forall x \in (b - \delta_2, b)$, 都有 $f(x) > f(x_0)$ (2)

又因为 $[a + \delta_1, b - \delta_2] \subset (a, b)$, 所以 $f(x)$ 在 $[a + \delta_1, b - \delta_2]$ 上连续,

由教材中的最值定理可知, $f(x)$ 在 $[a + \delta_1, b - \delta_2]$ 上必存在最大值.

不妨设该最大值点为 x_1 , 则对 $\forall x \in [a + \delta_1, b - \delta_2]$, 有 $f(x) \leq f(x_1)$, (3)

而 $x_0 \in [a + \delta_1, b - \delta_2]$, 故 $f(x_0) \leq f(x_1)$, (4)

则由(1)(2)(3)(4)可得, $f(x_1)$ 就是函数 $f(x)$ 在 (a, b) 内的最大值.

定理 2 设函数 $f(x)$ 在 (a, b) 内连续, 且 $\lim_{x \rightarrow a^+} f(x) = A$, $\lim_{x \rightarrow b^-} f(x) = B$, 在 (a, b) 内 $f(x)$ 取到比 A, B 都小的数, 则 $f(x)$ 在 (a, b) 上能取到最小值.,

定理 2 的证法与定理 1 类似, 此处不再证明.

例 1 求函数 $y = 1 + xe^{-x}$ 在 $(0, 2)$ 上的最大值和最小值.

解 $f(x)$ 在 $(0, 2)$ 内连续, 且 $\lim_{x \rightarrow 0^+} f(x) = 1$, $\lim_{x \rightarrow 2^-} f(x) = 1 + 2e^{-2}$,

取 $x=1$, $f(1) = e^{-1} + 1$,

有 $f(1) > \lim_{x \rightarrow 0^+} f(x)$, $f(1) > \lim_{x \rightarrow 2^-} f(x)$,

由定理 1 可知, $f(x)$ 在 $(0, 2)$ 上存在最大值,

$f'(x) = e^{-x} - xe^{-x} = e^{-x}(1-x)$,

令 $f'(x) = 0$, 得 $x=1$, 又 $f(1) = e^{-1} + 1$,

故 $f(x)$ 在 $x=1$ 处取得最大值 $e^{-1} + 1$,
由于驻点的唯一性, 函数在 $(0,2)$ 无最小值.

2 无穷区间上连续函数的最值

定理 3 $f(x)$ 在 $(-\infty, +\infty)$ 上连续, 且 $\lim_{x \rightarrow +\infty} f(x) = A$, $\lim_{x \rightarrow -\infty} f(x) = B$,

(1) 若存在 x_0 , 使 $f(x_0) > \max\{A, B\}$, 那么 $f(x)$ 在 $(-\infty, +\infty)$ 上存在最大值;

(2) 若存在 x_1 , 使 $f(x_1) < \min\{A, B\}$, 那么 $f(x)$ 在 $(-\infty, +\infty)$ 上存在最小值;

(3) 若 (1) (2) 同时成立, 则 $f(x)$ 在 $(-\infty, +\infty)$ 上存在最小值, 也存在最大值.

下面我们证明 (1)

证明 不妨设 $f(x_0) > \max\{A, B\}$, $x_0 \in \mathbb{R}$,

由引理 1 知, $\exists x_1 (x_1 > x_0)$, 当 $x > x_1$ 时, 有 $f(x) < f(x_0)$,

同理, $\exists x_2 (x_2 < x_0)$, 当 $x < x_2$ 时, 有 $f(x) < f(x_0)$,

$\because f(x)$ 在 $[x_2, x_1]$ 上连续, \therefore 由教材中的最值定理知, 存在 $x_3 \in [x_2, x_1]$, $f(x)$ 在 $[x_2, x_1]$ 的最大值为 $f(x_3)$,

且 $f(x_3) \geq f(x_0)$

又 $\because f(x)$ 在 $(-\infty, x_2)$, $(x_1, +\infty)$ 上小于 $f(x_0)$, $\therefore f(x)$ 在 $(-\infty, +\infty)$ 上取得最大值 $f(x_3)$

(2) (3) 的证明与 (1) 类似, 此处不再赘述.

例 2 讨论 $f(x) = xe^{-x^2}$ 在 $(-\infty, +\infty)$ 上的最值情况.

解 $f(x)$ 在 $(-\infty, +\infty)$ 内连续, 且 $\lim_{x \rightarrow \infty} xe^{-x^2} = 0$

取 $x=1$, 则 $f(1) = e^{-1} > 0$; 取 $x=-1$, 则 $f(-1) = -e^{-1} < 0$

由定理 3 可知, $f(x)$ 在 $(-\infty, +\infty)$ 存在最大值和最小值

$$f'(x) = (1 - 2x^2)e^{-x^2},$$

令 $f'(x) = 0$, 得驻点 $x_1 = -\frac{1}{\sqrt{2}}$, $x_2 = \frac{1}{\sqrt{2}}$,

$$f(x_1) = f(x_2) = -\frac{e^{-0.5}}{\sqrt{2}}$$

又因为 $\lim_{x \rightarrow \infty} xe^{-x^2} = 0$, 所以 $f(x)$ 在 x_2 处取得最大值 $\frac{e^{-0.5}}{\sqrt{2}}$,

在 x_1 处取得最小值 $-\frac{e^{-0.5}}{\sqrt{2}}$.

3 总结

最值是现实生活中大量存在的问题, 如何求解最大值与最小值非常重要! 最值定理要求区间是有限闭区间 (很多情况下并不满足此条件), 使得其应用非常有限。本文阐述了在开区间及无穷区间函数取得最值的条件及结论, 并将其应用于求解最值。

参考文献

- [1] 电子科技大学, 微积分第三版上册, 高等教育出版社, 2018
- [2] 教育部考试中心, 全国硕士研究生入学统一考试数学考试大纲解析, 高等教育出版社, 2007

积分型余项的泰勒公式的证明及应用

梁宇浩 (2019091201012)

指导教师：张利琼

【摘要】 关于泰勒公式，大部分数学教材中一般只介绍了皮亚诺型余项和拉格朗日型余项，对积分型余项的泰勒公式则阐述较少。本文将具体介绍积分型余项泰勒公式的证明，并给出其在定积分计算中的具体应用。

【关键词】 分部积分法；积分型余项；泰勒公式

1 定理及其证明

定理 1 若函数 $f(x)$ 在 x_0 的某邻域 $u(x_0)$ 内有连续的 $n+1$ 阶导数，则对 $\forall x \in u(x_0)$ ，有

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t)(x - t)^n dt$$

下面用数学归纳法证明其正确性

证明 $f(x) = f(x_0) + f(x) - f(x_0)$

$$\begin{aligned} &= f(x_0) + \int_{x_0}^x f'(t)dt \\ &= f(x_0) - \int_{x_0}^x f'(t)d(x-t) \\ &= f(x_0) - [f'(t)(x-t)]_{x_0}^x + \int_{x_0}^x (x-t)f''(t)dt \\ &= f(x_0) + f'(x_0)(x-x_0) - \frac{1}{2} \int_{x_0}^x f''(t)d(x-t)^2 \\ &= f(x_0) - f'(x_0)(x-x_0) - \frac{1}{2} [f''(t)(x-t)^2]_{x_0}^x + \frac{1}{2} \int_{x_0}^x (x-t)^2 f^{(3)}(t)dt \\ &= f(x_0) - f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + \frac{1}{2} \int_{x_0}^x (x-t)^2 f^{(3)}(t)dt \\ &\quad = \dots \\ &= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!} + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t)(x-t)^n dt \end{aligned}$$

下面用数学归纳法证明其正确性。

$$\text{记 } R_n(x) = \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t)dt$$

当 $n=0$ 时，所证泰勒公式即为 N-L 公式，

假设当 $n=k$ 时，积分性余项的泰勒公式成立，即

$$\begin{aligned} R_k(x) &= \frac{1}{k!} \int_{x_0}^x (x-t)^k f^{(k+1)}(t)dt \\ &= \frac{1}{k!} \int_{x_0}^x f^{(k+1)}(t) d\left[-\frac{(x-t)^{k+1}}{k+1}\right] \\ &= -\frac{1}{(k+1)!} (x-t)^{k+1} f^{(k+1)}(t) \Big|_{x_0}^x + \frac{1}{(k+1)!} \int_{x_0}^x (x-t)^{k+1} f^{(k+2)}(t)dt \end{aligned}$$

$$= \frac{1}{(k+1)!} (x-x_0)^{k+1} f^{(k+1)}(x_0) + R_{k+1}(x)$$

可见，当 $n=k+1$ 时，积分型余项的泰勒公式成立。

由数学归纳法，可知定理 1 成立。

证明的关键在于两点：一.如何合理地构造出定积分，使其与泰勒公式产生联系，二.分部积分法的运用，由此泰勒公式可以继续向下展开。

2 积分型余项的泰勒公式的应用

例 求解下列积分

$$(1) \int_0^1 (1-x)^n e^x dx, n \in N^+$$

解 设 $f(x) = e^x$, $f^{(n+1)}(x) = e^x$, 由定理 1 可得

$$\begin{aligned} \text{原式} &= \int_0^1 f^{(n+1)}(x)(1-x)^n dx, \\ &= (e^1 \cdot e^0 - e^0 \cdot 1 - \frac{e^0}{2!} \cdot 1^2 - \dots - \frac{e^0}{n!} \cdot 1^n) n! \\ &= (e - 2 \cdot \frac{1}{2!} - \dots - \frac{1}{n!}) n! \end{aligned}$$

$$(2) \int_a^b \frac{1}{x^{n+2}} (b-x)^n dx \quad (n \in N^+, a < x < b)$$

$$\text{解 设 } f(x) = \frac{1}{x}, \quad \text{则 } f^{(n+1)}(x) = \frac{(-1)^{(n+1)}(n+1)}{x^{n+2}},$$

由定理 1 可得

$$\begin{aligned} \text{原式} &= \frac{(-1)^{n+1}}{(n+1)!} \int_a^b f^{(n+1)}(x)(b-x)^n dx \\ &= \frac{(-1)^{n+1} n!}{(n+1)!} \left[\frac{1}{b} - \frac{1}{a} + \frac{1}{a^2} (b-a)^2 - \frac{2!}{a^3} (b-a)^3 + \dots + \frac{(-1)^n n!}{a^{n+1}} (b-a)^n \right] \\ &= \frac{(-1)^{n+1} (1 - \frac{b}{a})^{n+1}}{(n+1)b} \end{aligned}$$

$$(3) \int_0^1 (1-x^2)^n dx, n \in N^+$$

$$\text{解 } \int_0^1 (1-x^2)^n dx$$

$$\begin{aligned} &= \int_0^1 \left[\frac{(1+x)^{2n+1}}{(2n+1)2n\cdots(n+1)} \right]^{(n+1)} (1-x)^n dx \\ &= n! \left[\frac{2^{2n+1} - 1^{2n+1}}{(2n+1)2n\cdots(n+1)} - \frac{1}{2n(2n-1)\cdots(n+1)} - \frac{1}{2!(2n-1)\cdots(n+1)} - \frac{1}{n!(n+1)} \right] \\ &= n! \left[\frac{2^{2n+1} n! - C_{2n+1}^0 n!}{(2n+1)!} - \frac{C_{2n+1}^1 n!}{(2n+1)!} - \frac{C_{2n+1}^2 n!}{(2n+1)!} - \frac{C_{2n+1}^n n!}{n!(n+1)} \right] \\ &= (n!)^2 \left[\frac{2^{2n+1} - \frac{1}{2}(1+1)^{2n+1}}{(2n+1)!} \right] \end{aligned}$$

$$= \frac{4^n (n!)^2}{(2n+1)!}$$

3 小结

$\int_0^1 e^x (1-x)^n dx$ ($n \in N^+$) 这种以 n 为参变量的积分具有一种简洁美, 解决起来也需要一定的方法. 在计算此类比较复杂的积分时, 用本文所讲的积分型余项的泰勒公式具有相当的优越性。

参考文献

- [1] 黄军华.带积分型余项的泰勒公式在定积分计算中的应用,玉林师范学院学报,2006年,第27卷,15-17
- [2] 赵文强,丁宣浩.不同形式的泰勒定理及其应用,重庆工商大学学报:自然科学版,2012年,第29卷,9-13
- [3] 任春丽,王剑锋,杨有龙.浅谈泰勒公式的积分型余项,高等数学研究,2002年,第20卷,26-27

求函数零点的特殊方法

王池 (2019091201008) 彭博 (2019091201004) 孙乙赫 (2019091201003)

指导教师: 张利琼

【摘要】 零点定理只适用于有界闭区间，在开区间、半开闭区间上不一定成立。本文讨论在开区间、半开区间上函数存在零点的条件，并将其用于求解函数的零点。

【关键词】 闭区间；开区间；半开区间；零点

1 闭区间的零点存在定理

在《微积分（上册）》第一章 1.6 节里的定理 11 推论 2，我们学习了闭区间连续函数的零点存在定理，定理如下：

定理 1 若 $f(x) \in C[a, b]$ ，且 $f(a) \cdot f(b) < 0$ ，则 $\exists \xi \in (a, b)$ ，使 $f(\xi) = 0$.

此定理在证明方程根的存在性时有着非常重要的作用。

我们来看以下例题。

例 1 在 $(0, 2)$ 内，函数 $f(x) = \frac{e^x - 2x - 1}{x}$ 是否存在零点.

此题要求函数在区间内是否存在零点，则我们先想到使用零点存在定理。但是对比定理

可以看出，题目中的函数 $f(x)$ 并不符合 $f(x) \in C[0, 2]$ 这一条件，因为显然 $f(x)$ 在 $x=0$ 时无定义。那么此时我们该怎么办呢？在半闭开区间内是否也有与闭区间的零点存在定理相似的推论呢？

2 极限存在时的零点存在定理

定理 2 若 $f(x) \in C(a, b)$ ，且 $\lim_{x \rightarrow a^+} f(x) \cdot \lim_{x \rightarrow b^-} f(x) < 0$ ，则 $\exists \xi \in (a, b)$ ，使 $f(\xi) = 0$.

证明 1 由于此函数在开区间连续，想要利用闭区间的零点存在定理，我们需要构造出闭区间上的连续函数：

$$\text{令 } F(x) = \begin{cases} \lim_{x \rightarrow a^+} f(x), & x = a \\ f(x), & a < x < b \\ \lim_{x \rightarrow b^-} f(x), & x = b \end{cases}$$

$$\text{则 } \because F(a) = \lim_{x \rightarrow a^+} F(x), \quad F(b) = \lim_{x \rightarrow b^-} F(x),$$

$$F(x) \text{ 在 } C[a, b] \text{ 上连续，且 } F(a) \cdot F(b) = \lim_{x \rightarrow a^+} f(x) \cdot \lim_{x \rightarrow b^-} f(x) < 0$$

则 $\exists \xi \in (a, b)$ ，使 $F(\xi) = 0$.

\Leftrightarrow 则 $\exists \xi \in (a, b)$ ，使 $f(\xi) = 0$.

此外，利用极限的局部保号性，同样可证此结论。

证明 2 由于 $\lim_{x \rightarrow a^+} f(x) \cdot \lim_{x \rightarrow b^-} f(x) < 0$ ，利用极限的局部保号性，

$\exists a_1 > a$ ，当 $a < x < a_1$ 时， $f(x)$ 与 $\lim_{x \rightarrow a^+} f(x)$ 正负号相同，

可取 a_2 ($a < a_2 < a_1$)， $f(a_2)$ 与 $\lim_{x \rightarrow a^+} f(x)$ 正负号相同。

同理, $\exists b > b_1$, 当 $b_1 < x < b$ 时, $f(x)$ 与 $\lim_{x \rightarrow b^-} f(x)$ 正负号相同,

可取 $b_2 (b_1 < b_2 < b)$, $f(b_2)$ 与 $\lim_{x \rightarrow b^-} f(x)$ 正负号相同.

显然, $a < a_2 < b_2 < b$, 又因为 $\lim_{x \rightarrow a^+} f(x) \cdot \lim_{x \rightarrow b^-} f(x) < 0$

故 $f(a_2) \cdot f(b_2) < 0$, 即 $f(x)$ 在 $[a_2, b_2]$ 上满足零点定理1,

则 $\exists \xi \in (a_2, b_2) \subset (a, b)$, 使 $f(\xi) = 0$.

推论1 若 $f(x) \in C[a, b]$, 且 $f(a) \cdot \lim_{x \rightarrow b^-} f(x) < 0$, 则 $\exists \xi \in (a, b)$, 使 $f(\xi) = 0$.

显然推论1在 $f(x) \in C(a, b]$, $\lim_{x \rightarrow a^+} f(x) \cdot f(b) < 0$ 时也成立.

通过以上讨论得知, 在满足定理2, 或推论1的条件时, 函数在开区间, 半开区间也存在零点。

例1 证明函数 $f(x) = \frac{e^x - 2x - 1}{x}$ 在 $(0, 2)$ 内一定有零点.

解 通过计算易得

$$\lim_{x \rightarrow 0^+} f(x) = -1 < 0, f(2) = \frac{e^2 - 5}{2} > 0,$$

则 $\lim_{x \rightarrow 0^+} f(x) \cdot f(2) < 0$, 且 $f(x)$ 在 $C(0, 2]$ 上连续,

由推论1知 $\exists \xi \in (0, 2)$, 使 $f(\xi) = 0$.

通过这个结论, 我们把零点存在定理的使用条件从闭区间扩大到了开区间及半开区间。由此结论, 我们可以解决更多闭区间零点存在定理无法解决的问题。定理2同样有局限性, 思考下面一个问题。

例2 当 $x \in (0, 2)$ 时, 函数 $\frac{\ln x + \cos x}{x^2 + 6}$ 是否存在零点.

看到零点是否存在的问题, 我们首先思考闭区间零点存在定理。但是可以发现, 题目中的函数 $f(x)$ 并不符合 $f(x) \in C[0, 2]$ 这一条件, 因为显然 $f(x)$ 在 $x = 0$ 时无定义。从而我们联想到开区间的零点存在定理, 而题目中的函数 $f(x)$ 在左端点 0 无极限, 不满足定理2或推论1对于极限不存在的情况。我们又该怎么办呢?

3 极限不存在时的零点存在定理

当函数在两端点 a, b 的极限都不存在时, 函数在 (a, b) 有无零点呢? 有以下结论:

定理3 若 $f(x) \in C(a, b)$, 且 $\lim_{x \rightarrow a^+} f(x) = +\infty$, $\lim_{x \rightarrow b^-} f(x) = -\infty$ (或者 $\lim_{x \rightarrow a^+} f(x) = -\infty$,

$\lim_{x \rightarrow b^-} f(x) = +\infty$), 则 $\exists \xi \in (a, b)$, 使 $f(\xi) = 0$.

证明 由于 $\lim_{x \rightarrow a^+} f(x) = +\infty$, 由正无穷大的定义可知

任取 $M > 0$, $\exists \delta_1 > 0$, 当 $a < x < a + \delta_1$ 时, 有 $f(x) > M$,

同理, 由 $\lim_{x \rightarrow b^-} f(x) = -\infty$, 任取 $M_1 > 0$, $\exists \delta_2 > 0$, 当 $b - \delta_2 < x < b$ 时, 有

$$f(x) < -M_1,$$

取 $a < a_1 < a + \delta_1$, $b - \delta_2 < b_1 < b$,

则 $f(a_1) > M > 0$, $f(b_1) < -M_1 < 0$.

显然, $f(x)$ 在闭区间 $[a_1, b_1]$ 满足零点定理1, 根据定理1可得

$$\exists \xi \in (a_1, b_1) \subset (a, b), \text{ 使 } f(\xi) = 0.$$

(当 $\lim_{x \rightarrow a^+} f(x) = -\infty$, $\lim_{x \rightarrow b^-} f(x) = +\infty$ 时, 同理可证.)

4 其他形式的零点存在定理

当函数在一个端点存在极限，另一个端点无极限（但为无穷大时）也可以得到类似结论。

定理4 若 $f(x) \in C(a, b)$ ，且 $\lim_{x \rightarrow a^+} f(x) = +\infty$, $\lim_{x \rightarrow b^-} f(x) < 0$ ，则 $\exists \xi \in (a, b)$ ，使 $f(\xi) = 0$.

定理4的证明方法与上述证明一致，这里就不再证明。

定理4的条件 $\lim_{x \rightarrow a^+} f(x) = +\infty$, $\lim_{x \rightarrow b^-} f(x) < 0$ 可改为下述情形中的一种：

或 $\lim_{x \rightarrow a^+} f(x) = -\infty$, $\lim_{x \rightarrow b^-} f(x) > 0$

或 $\lim_{x \rightarrow a^+} f(x) < 0$, $\lim_{x \rightarrow b^-} f(x) = +\infty$

或 $\lim_{x \rightarrow a^+} f(x) > 0$, $\lim_{x \rightarrow b^-} f(x) = -\infty$

这里也不再证明。

由此结论，可以解决例2。

例2 当 $x \in (0, 2)$ 时，函数 $f(x) = \frac{\ln x + \cos x}{x^2 + 6}$ 是否存在零点。

解 $\lim_{x \rightarrow 0^+} f(x) = -\infty$, $f(2) = \frac{\ln 2 + \cos 2}{10} > 0$,

$\lim_{x \rightarrow 0^+} f(x)$ 与 $f(2)$ 异号，且 $f(x)$ 在 $(0, 2)$ 上连续，

\therefore 利用定理4知， $\exists \xi \in (0, 2)$ ，使 $f(\xi) = 0$.

例3 证明函数 $f(x) = \frac{x}{1 + \frac{1}{x+1}} - \frac{x}{1 - \frac{1}{x-1}} - \frac{1}{3}$ 在 $(-1, 1)$ 内至少存在两个零点。

证明 $\lim_{x \rightarrow (-1)^+} f(x) = (\frac{x(x+1)}{x+2} - \frac{x(x-1)}{x-2} - \frac{1}{3}) = \frac{1}{3} > 0$

$\lim_{x \rightarrow (+1)^-} f(x) = (\frac{x(x+1)}{x+2} - \frac{x(x-1)}{x-2} - \frac{1}{3}) = \frac{1}{3} > 0$

$f(0) = -\frac{1}{3} < 0$

$f(x)$ 在区间 $(-1, 0]$ 及 $[0, 1)$ 满足推论1，故在 $(-1, 0)$ 内，函数 $f(x)$ 至少有一个零点。

在 $(0, 1)$ 内，函数 $f(x)$ 至少有一零点。

综上，函数在 $x \in (-1, 1)$ 内至少有两个零点。

5 总结

通过对闭区间零点存在定理的分析，我们将其拓展为了极限存在时的开区间及半开闭区间的零点存在定理，极限不存在时的开区间零点存在定理，以及其他形式的零点存在定理这几种情况，使定理的使用范围得到了拓宽。但是我们不难发现，在这几种结论中，最为重要的就是转换法。在面对证明零点存在一类的题目时，一方面，我们要牢记结论；另一方面，在面对结论之外的题目时，我们要学会主动将其转换为定理，或是由定理推导出的结论，比如极限存在时的或是极限趋于无穷时开区间零点存在定理等。如果能够做到这两点，在面对这一类的题型时，我们的解题速度就会变得更快，思路更加清晰！

参考文献

- [1] 李娅, 薛玉梅, 教育教学论坛, 河北教育出版社, 2017
- [2] 叶国炳, 陕西工学院学报, 陕西理工大学出版社, 2000