概率统计中的反例

第四章　随机变量的数字特征

1. 随机变量的数学期望未必都存在
2. 随机变量的方差未必都存在
3. 数学期望存在但方差不存在
4. X 的函数的期望是否等于 X 的期望的函数
5. X 的各阶矩都存在也不能确定 X 的分布函数
6. 满足 E(XY)=E(X)E(Y)的 X,Y 未必独立

第五章　参数估计与假设检验

1. 矩估计是否有唯一性
2. 矩估计不具有“不变性”
3. 极大似然估计是否有唯一性
4. 似然方程的解未必是极大似然估计
5. 参数估计的无偏性与一致性有无关系
6. 无偏估计是否唯一
7. 零假设与备择假设是否处于对等的地位

前　　　言

数学是由两个大类——证明和反例组成

数学发现主要是提出证明和构造反例

从科学性来讲

反例就是推翻错误命题的有效手段

从教学上而言

反例能够加深对正确结论的全面理解

【美】B.R.盖尔鲍姆曾说

“一个数学问题用一个反例予以解决

给人的刺激犹如一出好的戏剧”

相信读了《概率统计中的反例》后

我们大家都会有这一同感

第一章　随机事件及其概率

1．同一问题的概型未必唯一

概型（Schema）是随机现象的数学形式，它不是实际本身，而是实际的数学抽象。对于现实世界中的随机现象，要想进入数学理论的研究，首先必须确定其概型。

由于我们的认识水平以及现实问题的复杂性，使得所选定的概型往往不是唯一的。

概率论中著名的“n 个球在 n 个盒子中的分布问题”（见王梓坤《概率论基础及其应用》P12-13 科学出版社）就说明了这一情况，这是一个典型概型的问题，内容是：设有 r 个球，每个都能以相同概率 1/n 落到 n 个盒子（n>=r）的每一个盒子中，求指定的某 r 个盒子中各有一个球的概率。

如果我们把 r 各球视作 r 个人，而把 n 个盒子视为一年的天数：n=365.这时上述问题就成为了概率论中一个颇为著名问题的概型。此问题是求参加某次集会的几个人中，没有 n 个人生日相同的概率。

众所周知，关于球彼此间可以认为是有区别，也可以认为无区别；一个盒子可以假定仅能容纳一个球，也可以允许它能容纳许多球，如此一来，就可以分为以下几种概型：

（1）　马克斯威尔－波尔茨曼　认为球彼此之间有区别，且对每盒中可容纳球数不加限制；

（2）　玻色－爱因斯坦　认为球彼此不能区别，且对每盒中可容纳球数不加限制；

（3）　费密－狄雷克　认为球彼此无区别，且限制每盒中不能同时容纳二个球。

后来，为了统一以上三种情况，又产生了第四种情况

（4）　布里龙　认为球彼此可以区别，且增加了一些其他条件限制（见杨宗磐《概率论入门》 P.13 科学出版社）

以上四种情况，形成了统计物理学中的四种统计：球可看作为质点，盒子看作状态。

再看一例：n 个人围成一个圆周，求其中甲、乙两人之间恰有 r(<n-2)个人的概率。（圆周排列时，仅考虑从甲到乙的顺时针方向）对此问题，至少可找到三种概型来处理即可以构造如下的三种随机试验：

（1）　n 个人的任意一种排列作为一个基本事件；

（2）　仅以甲、乙两人在 n 个人一行中的不同排法作为基本事件组；

（3）　可由甲与乙之间的间隔数来考虑。

不论取何种概型，本题的求概率均为 1/(n-1).

2．事件间的关系

(1). 由 $A-B=C$ 推不出 $A=B\cup C$

事实上，令 A={1,2,3,4},B={1,3,5},于是 C=A-B={2,4},而 $B\cup C=\{1,2,3,4,5\}\neq A$

注：但 $A\supset B$ 时，能由 $A-B=C\Rightarrow A=B\cup C$

(2). 由 $A=B\cup C$ 推不出 $A-B=C$

令 A={1,2,3,5},B={1,2},C={1,3,5},则 $A=B\cup C$ 但 $A-B=\{3,5\}\neq C$

注：当 $B\subset A,C\subset A$,且 $BC=\Phi$ 时可由 $A=B\cup C\Rightarrow A-B=C$

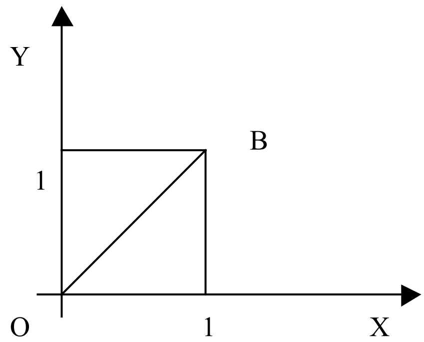
(3). 一般 $A\cup(B-C)\neq(A\cup B)-C$

令 A={1,2},B={2,3},C={2},则 $A\cup(B-C)=\{1,2,3\}\neq\{1,3\}=(A\cup B)-C$

注：当 $AC=\phi$ 时，$A\cup(B-C)=(A\cup B)-C$

3．概率为零的事件未必是不可能事件

不可能事件的概率必为零，反之却未必成立
当考虑的概型为古典概型时，概型为零的事件一定是不可能事件
当考虑的概型是几何概型时，概型为零的事件未必是一个不可能事件。例如：设试验 E 为“随机地向边长为 1 的正方形内投点”，事件 A 为“点投在正方形的一条对角线上”（见图）



此时 $\Omega=\{(x,y)\mid 0<x,y<1\}$

$A=\{x=y\mid 0<x,y<1\}$

尽管 $P(A)=\dfrac{\text{线段}OB\text{的面积}}{E\text{正方形的面积}}=\dfrac{0}{1}=0$

但 A 却可能发生，

另外，对于连续性随机变量，它在某固定点取值的概率为零，但它不是不可能发生。

发生上述情形的原因，在于概率是一个测度，有测度为 0 的不可数集存在，并且对于连续函数来说，在一点处的积分为零。

由对立事件知，概率为 1 的事件未必是必然事件。

4．由概率关系推不出事件间关系

概率中有这样的性质：若事件 A,B 有关系 $A\subset B$,则其相应的概率关系是 $P(A)\leq P(B)$,反之却不真。例如：

设 P(A)=0.1,P(B)=0.2, $P(A\cup B)=0.2$,此时 $A\subset B$ 不成立， 事实上， 由 $P(A\cup B)=P(A)+P(B)-P(AB)$ 得出 P(AB)=0.1.

于是 $P(A-AB)=P(A)-P(AB)=0$

由问题 3，这意味着可有 $A-AB\neq\phi$,从而未必有 $A\subset B$。

5．试验次数多概率就一定大吗

在概率论的萌芽时期，有一个著名的(Chevalier de Were)问题：一颗骰子掷 4 次至少得一个 1 点，与两颗骰子掷 24 次至少得两个 1 点，这两个事件究竟哪个概率大？曾引起很多人的注意。现在看来，利用独立试验概型容易求出它们的概率。

n 次独立重复试验中事件 A 至少发生一次的概率为$1-(1-p)^n$,其中 $p=P(A)$，

现考虑欲使$1-(1-p)^n\geq\frac{1}{2}$，则$n\geq-\frac{\lg 2}{\lg(1-p)}$，此式给出了 n 的下界，使问题得以解决。

以掷一颗骰子作试验，要连续掷 n 次使 1 点至少出现一次的概率大于等于 1/2,则$n\geq 3.8$.以掷两颗骰子作试验，要连续掷 n 次使两个 1 点至少出现一次的概率大于等于 1/2,则$n\geq 24.6$.由此得出，一颗骰子掷 4 次至少有一个 1 点的概率大于等于 1/2,而两颗骰子掷 24 次至少有一次得两个 1 点的概率小于 1/2.

本例说明试验次数很多，但概率不一定大。

6．概率与抽样方式是否有关

一般，所求事件的概率与抽样方式有关，常见的有放回抽样与不放回抽样两种。前者指同一个体可被重复抽取，后者指已被抽取的个体不再参与下一此抽样，每一个体至多被抽到一次。

例如有 n 件产品，其中有 m 件次品，现随机抽取l件产品。求其中恰有 k 件次品的概率。在抽样方式未定的情况下，此概率是不唯一的，事实上：

若取放回抽样，所求概率 $p_1=C_l^k(\frac{m}{n})^k(1-\frac{m}{n})^{l-k}$ （1）

若取不放回抽样 所求概率 $p_2=\frac{C_m^kC_{n-m}^{l-k}}{C_n^l}$ （2）

显然 $p_1\neq p_2$ $k=0,1,2,\cdots,\min(m,l)$

（1)，（2）分别称为二项分布与超几何分布。

当$n\to\infty$时$(2)\to(1)$即二项分布是超几何分布的极限分布。

由此得出如下结论：对于样本点个数较小的总体，在放回抽样与不放回抽样的场合，事件的概率会有较大的差别。当总体中样本点个数很大，样本容量不大时，

这两种抽样方式对所求事件的概率实际上影响不大。

7．事件概率与试验的先后次序是否有关

设有一口袋，内有 a 只黑球，b 只白球，他们除颜色不同外没有其它不同之处，现把球一只只地摸出，求第 k 次摸出的是黑球的概率 $(1 \le k \le a+b)$.

初看题目，很可能会认为所求概率与摸球次序有关，若那样的话，体育比赛中先后抽签者中签的机会就不均等了，这与我们日常生活中的经验不符，通过具体计算亦可看出所求概率与摸球次序无关。

按自然顺序给球编号，不妨先给黑球编号，再给白球编号，取样空间为第 k 次摸出的球的全部可能的结果，则$\Omega=\{\omega_1,\omega_2,\cdots,\omega_{a+b}\}$ ω_i表示第 k 次摸出第 I 号球，$i=1,2,\cdots,a+b$，于是要求的是事件 $A=\{\omega_1,\omega_2,\cdots,\omega_a\}$的概率。由古典概率$P(A)=\dfrac{a}{a+b}$，P(A)显然与 k 有关。

本题可用多种方法求解，这里介绍的是最简单的一种，本题存在多种解法的原因，在于一个随机现象有时可用不同的样本空间来描述，所以称本解法是最简单的，因为解法中的样本空间是最小的。

第二章　随机变量及其分布

1．离散型分布的最可能值是否唯一

离散型分布的最可能值指的是该随机变量取值中那些使概率达到最大的值，即若任意一个离散型分布$\begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$，若 $p_k=\sup(p_1,p_2,\cdots,p_n,\cdots)$，则称 x_k 为此分布的最可能值。

一般离散型分布的最可能值不唯一，比如：二项分布 B（n,p）中，当$(n+1)p$为非负整数时，恰有两个最可能值：$(n+1)p$与$(n+1)p-1$.如二项分布 B（8，1/3），其最可能值为 k=2 或 3.

可以证明，任何离散型分布的最可能值一定存在，而且至少有一个。证明见王梓坤《概率论基础及其应用》　科学出版社）

2．单调不降右连续是分布函数的必要条件

分布函数一定是单调不降（右）连续的函数，反之命题不成立。例如，取 $F(x)=\begin{cases}-1 & x<-1\\ \dfrac{x}{2} & -1\le x<1\\ 1 & x\ge 1\end{cases}$，$F(x)$ 显然是调调不降函数，且右连续，可是 $F(-\infty)=-1\ne 0$，所以 $F(x)$ 不可能是某个随机变量的分布函数。

因为只有当一个函数满足单调不见，非负有界（$F(-\infty)=0,F(+\infty)=1$，且右连续（或左连续）时，才能成为某个随机变量的分布函数。

3．既非离散型又非连续型的分布函数是否存在

如果一个分布函数 $F(x)$ 是连续的，并且其导函数几乎处处等于零（关于勒贝格测度而言），则称 $F(x)$ 为奇异型分布函数。如果随机变量 X 的分布函数是奇异型的，则称 X 为奇异型随机变量。

任何一个奇异型的分布函数都是一个既非离散型又非连续型的分布函数。

有没有非奇异型的分布函数属于既非离散型又非连续型的分布函数？

有，请看下例：设 $F(x)=\begin{cases}0 & x<0\\ \dfrac{1+x}{2} & 0\le x<1\\ 1 & x\ge 1\end{cases}$，由分布函数的定义又知 $F(x)$ 是分布函数，又 $F^{'}(x)=\dfrac{1}{2}\ne 0,x\in(0,1)$，故 $F(x)$ 不是奇异型的分布函数，与 $F(x)$ 对应的随机变量不是取有限个或可列多个值，故 $F(x)$ 不是离散型的分布函数，又

$$\int_{-\infty}^{\infty}f(x)dx=\int_{-\infty}^{\infty}F^{'}(x)dx=\int_0^1\frac{1}{2}dx=\frac{1}{2}\ne 1$$

故 $F(x)$ 也不是连续的分布函数。

4．具有无记忆性的离散型分布是否存在

设随机变量 X 服从某个分布，若它满足

$$P(X>s+t\mid X>s)=P(X>t)$$

则称概分布具有无记忆性。

对于连续型分布来说，指数分布是唯一的具有无记忆性的。（证明可见 复旦大学《概率论》 人们教育出版社 P.125-126）在可靠性问题中，把 X 理解为某元件的寿命，则无记忆

性表示某元件的寿命如果已知大于 5 年，则其寿命再延长七年的概率与年令无关。

具有无记忆性的离散型分布也是存在且唯一的，那就是几何分布

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

几何分布是一种等待分布，例如,在事件 A 发生的概率为 p 的贝努里试验之中，A 首次出现时的等待次数 X 的分布为几何分布。

5．不几乎相等的随机变量是否有相同的分布

若两个随机变量 X,,Y 满足 $P(X\neq Y)=0$,则称 X 与 Y 几乎相等。

可以证明：几乎相等的随机变量具有相同的分布，反之都不成立。

例如，设 X 与 Y 具有相同的分布

$$\begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

并设 X 与 Y 相互独立，据此可算得 $P(X=Y)=\frac{1}{2}\neq 1$，从而 $P(X\neq Y)=\frac{1}{2}\neq 0$，即 X 与 Y 不几乎相等。

所以不几乎相等得随机变量可以有相同的分布。

6．联合分布与其边缘分布未必是同类型分布

我们知道二维正态分布的边缘分布仍为正态分布，多项分布的边缘分布亦为多项分布。那么联合分布与边缘分布是否都是为同类型分布呢？答案是否定的。

例如二维均匀分布的边缘分布可以仍是均匀分布，也可以不是均匀分布。

边与坐标轴平行的矩形域上的二维均匀分布的边缘分布仍是均匀分布，而圆域上的二维均匀分布的边缘分布不再是均匀分布。

7．边缘分布不能决定联合分布

一般边缘分布由联合分布所决定，反之不真。

例如：$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$,则有 $X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$

反之，已知 $X\sim N(\mu_1,\sigma_1^2)$ ，$Y\sim N(\mu_2,\sigma_2^2)$ ，却得不出（X,Y）一定是二维正态分布的结论。若添加X与Y相互独立的条件，则可得 $(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;0)$

除连续型分布外，还可举出离散型分布的例子。

8. 不同的联合分布可具有相同的边缘分布

如下二个相异的联合分布：

Y \ X	0	1
0	0．1	0．2
1	0．2	0．5

Y \ X	0	1
0	0.15	0.15
1	0.15	0.55

它们的边缘分布完全相同

X	0	1
P	0.3	0.7

Y	0	1
P	0.3	0.7

由此可见边缘分布由联合分布唯一决定，反之不成立，除离散型分布外，还可举出连续型分布的例子。

9. 正态边缘分布可由非正态联合分布导出

正态分布具有许多好的性质，其中之一是：二维正态分布的边缘分布仍是正态分布。反之，两边缘分布都是正态分布，起联合分布未必是正态分布，例如：

设 $(X,Y) \sim f(x,y) = \dfrac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} (1+\sin x \sin y), -\infty < x, y < +\infty$

则 $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, -\infty < x < +\infty$

$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, -\infty < y < +\infty$

即 $X \sim N(0.1), Y \sim N(0.1)$,显然（X,Y）并不服从联合正态分布。

10. 均匀分布不具有可加性

若独立同分布的两随机变量之和仍服从原分布，则称该分布具有可加性。

可以证明二项分布，泊松（Poisson）分布，正态分布均具有可加性，而均匀分布不具有这个性质。

设 X,Y 相互独立，且都服从（a,b）上的均匀分布，令 Z=X+Y，则 Z 的密度函数为：

$$f(Z)=\begin{cases}0 & Z<2a,Z\geq 2b\\ \dfrac{Z-2a}{(b-a)^2} & 2a\leq Z<a+b\\ \dfrac{2b-Z}{(b-a)^2} & a+b\leq Z<2b\end{cases}$$

可见 Z 服从辛卜生（Simpson）分布，不再是均匀分布。

11. 分布函数之和不是分布函数

设 F(x),G(x)均为分布函数，其和$H(x)=F(x)+G(x)$，显然不是分布函数，

因为此时$H(+\infty)=F(+\infty)+G(+\infty)=2\neq 1$

若令$J(x)=\alpha F(x)+\beta G(x)$，当非负实数满足$\alpha+\beta=1$时，J(x)可作为某随机变量的分布函数。（证明见王梓坤《概率论基础及其应用》 P46）

顺便指出：分布函数之积必是分布函数.

第三章　独立性与相关性相容性

1. 两两独立但不相互独立

【例1】　设有一个均匀的正四面体，第一，二，三面分别涂上红，黄，兰一种颜色，第四面涂上红，黄，兰三种颜色。现以 A,B,C 分别记投一次四面体底面出现红，黄，兰颜色的事件，则

$$P(A)=P(B)=P(C)=\frac{1}{2},P(AB)=P(AC)=P(BC)=\frac{1}{4}$$

所以 A,B,C 两两独立，但

$$P(ABC)=\frac{1}{4}\neq\frac{1}{8}=P(A)P(B)P(C)$$

因而 A,B,C 不相互独立。

【例2】　设有四张形状，大小，质量完全一样的卡片，上面分别标有数字 112，121，211，222，现从四张卡片中任抽一张，以随机变量 X,Y,Z 分别表示抽到卡片上的第一，二，三位数字，则

$$P(X=1)=P(Y=1)=P(Z=1)=\frac{1}{2}$$

$$P(X=1,Y=1)=P(X=1,Z=1)=P(Y=1,Z=1)=\frac{1}{4}$$

所以 X,Y,Z 两两独立，但

$$P(X=1,Y=1,Z=1)=0\neq\frac{1}{8}=P(X=1)P(Y=1)P(Z=1)$$

因而 X,Y,Z 不相互独立。

2．P(ABC)=P(A)P(B)P(C)成立，但 A,B,C 不两两独立.

设有一均匀正八面体，其第 1，2，3，4 面涂有红色，第 1，2，3，5 面图黄色，第 1，6，7，8 面涂兰色。现以 A,B,C 分别表示投一次正八面体，底面出现红，黄，兰颜色的事件，则

$$P(A)=P(B)=P(C)=\frac{1}{2}$$

$$P(ABC)=\frac{1}{8}=P(A)P(B)P(C)$$

但是 $P(AB)=\frac{3}{8}\neq\frac{1}{4}=P(A)P(B)$

$$P(AC)=\frac{1}{8}\neq\frac{1}{4}=P(A)P(C)$$

$$P(BC)=\frac{1}{8}\neq\frac{1}{4}=P(B)P(C)$$

所以 A,B,C 不两两独立。

3．独立关系不具有传递性

设三事件 A,B,C，若由 A 与 B 独立，且 B 与 C 独立，得到 A 与 C 独立，我们就称 A,B,C 的独立关系具有传递性

考虑有两个孩子和家庭全体，假定生男生女是等可能的，因而样本空间 $\Omega=\{(b,b),(b,g),(g,b),(g,g)\}$，其中 b 为男孩，g 为女孩，每一对里的次序是指出生的次序.

现在从全体有两个孩子的家庭中随机地选择一个家庭，并考虑下面三个事件：

A 为“第一个孩子是男孩”，B 为“两个孩子不同性别”，C 为“第一个孩子是女孩”，则有

$$AB=\{(b,g)\},BC=\{(g,b)\},AC=\Phi$$

$$P(AB)=\frac{1}{4}=P(A)P(B),P(BC)=\frac{1}{4}=P(B)P(C)$$

即 A 与 B 独立，B 与 C 独立，但是

$$P(AC)=0\neq\frac{1}{4}=P(A)P(C)$$

因此 A 与 C 不独立.

顺便指出不独立关系也不具有传递性，即若 A,B 不独立，B,C 不独立，则 A,C 可以独立

考察掷三枚均匀硬币的试验

A 为“全正面或全反面”，B 为“至多两个正面”，C 为“至多一个正面”，试验的样本空间为

$$\Omega=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}$$

其中 H 表示正面，T 表示反正，容易算出：

$$P(A)=\frac{1}{4},P(B)=\frac{7}{8},P(C)=\frac{1}{2},P(AB)=\frac{1}{8},P(BC)=\frac{1}{2},P(AC)=\frac{1}{8}$$

于是有

$$P(AB)=\frac{1}{8}\neq\frac{7}{32}=P(A)P(B)$$

$$P(BC)=\frac{1}{2}\neq\frac{7}{16}=P(B)P(C)$$

$$P(AC)=\frac{1}{8}=P(A)P(C)$$

可见 A,B 不独立，B,C 不独立，A,C 却独立.

4．随机变量不独立，但其函数可以独立

正态分布有个特性：任何 n(>1)维正态随机变量，可由坐标轴的旋转转变为一组几个独立的正态随机变量.（参见丁寿田译的前苏联《概率论教程》 P157）

例如 n=2，即使 X,Y 不独立，当（X,Y）服从二维正态分布，令 $Z=X\cos\theta+Y\sin\theta$，$W=-X\sin\theta+Y\cos\theta$，则（Z,W）仍服从二维正态分布，其联合密度函数为：

$$f(Z,W)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}(AZ^2-2BZW+CW^2)}$$

只要适当地选择θ：$tg2\theta=\dfrac{2\rho\sigma_1\sigma_2}{\sigma_1^2-\sigma_2^2}$

则 B=0,此时 Z 与 W 独立

5．X 与 Y 不独立，但X^2与Y^2独立

若随机变量 X 与 Y 独立，则X^2与Y^2必相互独立，其逆不真

例如：设（X,Y）的联合密度函数为

$$f(x,y)=\begin{cases}\frac{1}{4}(1+xy) & |x|<1,|y|<1\\ 0 & \text{其他}\end{cases}$$

$$F_{X^2}(x)=P(X^2\le x)=\begin{cases}0 & x<0\\ \int_{-\sqrt{x}}^{\sqrt{x}}(\int_{-1}^{1}\frac{1+uy}{4}dy)du=\sqrt{x} & 0\le x<1\\ 1 & x\ge 1\end{cases}$$

$$F_{Y^2}(y)=\begin{cases}0 & y<0\\ \sqrt{y} & 0\le y<1\\ 1 & y\ge 1\end{cases}$$

$$F_{X^2Y^2}(x,y)=\begin{cases}0 & x<0\text{或}y<0\\ \sqrt{x} & 0\le x<1, y\ge 1\\ \sqrt{y} & x\ge 1, 0\le y<1\\ \sqrt{xy} & 0\le x<1, 0\le y<1\\ 1 & x\ge 1, y\ge 1\end{cases}$$

对于一切 x,y 恒有 $F_{X^2Y^2}(x,y)=F_{X^2}(x)F_{Y^2}(y)$

所以 X^2 与 Y^2 相互独立。

$$f_X(x)=\begin{cases}\frac{1}{2} & |x|<1\\ 0 & \text{其他}\end{cases}\qquad f_Y(y)=\begin{cases}\frac{1}{2} & |y|<1\\ 0 & \text{其他}\end{cases}$$

显然 $f(x,y)\ne f_x(x)f_Y(y)$，所以 X 与 Y 不独立.

6．X 与 Y 不独立，但有相同分布

观察如下的（X,Y）的联合分布及其边缘分布

Y \ X	0	1	2	Y 的分布
0	1/19	1/6	1/9	7/18
1	1/6	1/3	0	1/2
2	1/9	0	0	1/9
X 的分布	7/18	1/2	1/9	

由于 $P_{ij}\ne P_iP_j, i、j=1,2,3$

故 X 与 Y 不独立，但 X 与 Y 的分布显然相同.

7．既不相关也不独立的随机变量

若随机变量 X,Y 相互独立，则 X，Y 不相关，反之不真，这方面的反例很多，离散型与连续型各举一例.

例1． 设（X，Y）的分布为：

Y \ X	-1	0	1
-1	1/8	1/8	1/8
0	1/8	0	1/8
1	1/8	1/8	1/8

容易验证 $\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)=0 \Rightarrow X,Y$不相关

$$P_{ij} \neq P_i P_j, i,j=1,2,3 \Rightarrow X,Y\text{不独立}.$$

例2． 设（X,Y 服从单位圆域上的均匀分布，但其密度函数

$$f(x,y)=\begin{cases} \dfrac{1}{\pi} & x^2+y^2 \le 1 \\ 0 & x^2+y^2>1 \end{cases}$$

容易验证

$$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)=0 \Rightarrow X,Y\text{不相关}$$

$$f(x,y) \neq f_X(x) f_Y(y) \Rightarrow X,Y\text{不独立}.$$

8．随机变量独立但它们的函数未必独立

设$Z=f(X,Y),W=g(X,Y)$. X,Y 为相互独立的随机变量

（1） Z,W 独立的例子

设 X,Y 独立且有相同分布$N(0,\sigma^2)$，取$Z=\sqrt{X^2+Y^2}$，$W=\dfrac{X}{Y}$，

则（Z,W）的联合密度为

$$f(z,w)=\begin{cases} \dfrac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}\frac{1}{\pi(1+\omega^2)}} & z \ge 0 \\ 0 & z<0 \end{cases}$$

边缘密度为 $$f_Z(z)=\begin{cases} \dfrac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}} & z \ge 0 \\ 0 & z<0 \end{cases}$$

$$f_W(w)=\frac{1}{\pi(1+w^2)}, -\infty<w<\infty$$

则 $$f(z,w)=f_Z(z) f_W(w)$$

故 Z 与 W 独立

（2） Z,W 不独立的例子

设 X,Y 不独立，且都服从如下分布

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

取 $Z = X + Y, W = X - Y$,此时 Z 与 W 或者同为奇数或者同为偶数，所以 Z 与 W 不独立.

9．独立性与相容性

独立性是问题间的概率属性，相容性是事件间本身的关系。由第一章 4 可知，由概率关系推不出事件间关系，所以由独立性推不出不相容性。

看如下一个命题：

若 A,B 为两个独立事件，且 $P(A) > 0, P(B) > 0$,且 A,B 不能不相容，用反证法证明此命题。

若 $AB = \Phi$ ，则 $P(AB) = 0$ ，由 A,B 独立，且 $P(A) > 0, P(B) > 0$ 得 $P(AB) = P(A)P(B) > 0$,矛盾，因而 $AB \neq \Phi$

可见在题设条件下，A 与 B 独立同 A 与 B 互不相容不能同时成立。但若 A,B 中有一个概率为 0，则 A 与 B 独立同 A 与 B 互不相容可同时成立。

10． 独立同分布的随机变量是否必相等

设 X,Y 互相独立，且都服从两点分布

$$\begin{pmatrix} -1 & 1 \\ 1/2 & 1/2 \end{pmatrix}$$

则未必有 X=Y

事实上，由 X,Y 的独立性有

$$P(X = Y) = P(X = -1, Y = -1) + P(X = 1, Y = 1)$$
$$= P(X = -1)P(Y = -1) + P(X = 1)P(Y = 1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

显然 X=Y 不是必然事件.

11． 有函数关系的随机变量是否一定不独立

考验如下三个随机变量 X,Y 和 Z：假定 X 与 Y 独立且都服从参数为 p 的（0－1）分布，令 Z 为 X 与 Y 的函数

$$Z = \begin{cases} 0 & X + Y \text{为偶数} \\ 1 & X + Y \text{为奇数} \end{cases}$$

由于 X,Y 独立，可得（X,Y）的联合分布律

Y \ X	0	1
0	$(1-p)^2$	$p(1-p)$
1	$p(1-p)$	p^2

Z 的概率分布为：

$$\begin{pmatrix} 0 & 1 \\ p^2+(1-p)^2 & p^2 \end{pmatrix}$$

（X,Z）的联合分布率为

Z \ X	0	1	Z 的分布
0	$(1-p)^2$	p^2	$p^2+(1-p)^2$
1	$p(1-p)$	$p(1-p)$	$2p(1-p)$
X 的分布	$(1-p)$	p	

显然当 P=1/2,Z 与 X 相互独立

可见，尽管 Z 与 X 之间存在函数关系，但它们可以相互独立.

第四章　随机变量的数字特征

1．随机变量的数学期望未必都存在

在数学期望的定义中，要求级数绝对收敛或积分绝对可积，我们知道，绝对收敛的级数一定收敛，绝对可积的函数一定可积。反之都不真，故有数学期望不存在的随机变量存在。

（1）　离散的例子

设随机变量 X 取值 $x_k=(-1)^{k-1}\dfrac{2^k}{k},k=1,2,\cdots$，相应的概率为

$p_k=\dfrac{1}{2^k},k=1,2,\cdots$

由于 $\sum_{k=1}^{\infty}|x_k|p_k=\sum_{k=1}^{\infty}\dfrac{1}{k}=\infty$，所以 X 的数学期望不存在

然而 $\sum_{k=1}^{\infty}x_k p_k=\sum_{k=1}^{\infty}(-1)^{k-1}\dfrac{1}{k}=1-\dfrac{1}{2}+\dfrac{1}{3}-\dfrac{1}{4}+\cdots=\ln 2$

若把上式左边级数中的各项进行重排，会收敛到不同的数

例如：$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\cdots=\frac{3}{2}\ln 2$

$$1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}+\cdots=\frac{1}{2}\ln 2$$

一个随机变量的数学期望只能是一个数，因此数学期望定义中要求的绝对收敛是必要的，它们可以保证 x_k 顺序的变化不影响数学期望中级数的收敛性

（2） 连续的例子，见教材 P.141 例 5 柯西（Cauchy）分布

2．随机变量的方差未必都存在

按定义 $D(X)=E(X-E(X))^2$,由于方差被定义为一种特殊形式（即随机变量 X 的函数）的数学期望，而随机变量及随机变量函数的数学期望都未必存在，所以随机变量的方差也未必存在。

本章 1 中所举两例中的随机变量的方差都不存在.

3．数学期望存在但方差不存在

参数为 n 的 t 分布的密度函数是

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

设随机变量 $X\sim t(2)$，则其密度函数

$$f_2(x)=\frac{\sqrt{2}}{4}(1+\frac{x^2}{2})^{-\frac{3}{2}} \qquad E(X)=\int_{-\infty}^{+\infty}xf_2(x)dx=0$$

X^2 的数学期望不存在，所以 X 的方差不存在

关于 t 分布，其矩有一个特点，当 r<n 时，有矩 $E(X^r)$,但 $E(X^n)$ 不存在，而且当 n>2 时，$E(X)=0$，$D(X)=E(X^2)=\frac{n}{n-2}$，故在 n=2 时，$D(X)=\infty$.

4．X 的函数的期望是否等于 X 的期望的函数

一般 $E[f(X)]\neq f[E(X)]$

例如：设 X 服从如下分布：

$$\begin{pmatrix}-1 & 0 & 1 & 2\\ 1/4 & 1/4 & 1/4 & 1/4\end{pmatrix}$$

再设 $f(x)=x^3-x, x\in\{-1,0,1,2\}$ 于是随机变量函数

$$f(x)\sim\begin{pmatrix}0 & 6\\ 3/4 & 1/4\end{pmatrix}$$

故得 $E[f(X)]=\dfrac{3}{2}$,由于 $E(X)=1/2$，可得 $f[E(X)]=(\dfrac{1}{2})^3-\dfrac{1}{2}=-\dfrac{3}{8}$,所以 $E[f(X)]\neq f[E(X)]$

5．X 的各阶矩都存在也不能确定 X 的分布函数

若已经 X 属于某个分布，且又知道其某些矩，便可确定 X 的分布函数，比如知道一阶原点矩（即数学期望）就能确定（0－1）分布的泊松分布，指数分布的分布函数，若进一步还知道二阶中心矩，便能确定二项分布，均匀分布，正态分布的分布函数。由此是否能推断：知道 X 的各阶矩，就一定能确定 X 的分布函数呢？回答是否定的，事实上，存在着不同的分布函数，其各阶矩都一样。

例如 设随机变量 X 与 Y 的密度函数分别是

$$f_X(x)=\begin{cases}ce^{-x^a\cos a\pi} & x>0\\ 0 & x\le 0\end{cases}$$

$$f_Y(x)=\begin{cases}c[1+\sin(x^a\cos a\pi)] & x>0\\ 0 & x\le 0\end{cases}$$

其中 0<a<1/2, $c=\dfrac{a(\cos a\pi)^{\frac{1}{a}}}{\Gamma(\frac{1}{a})}$

显然 $f_X(x)\neq f_Y(x)$，从而 X 与 Y 各自的分布函数 $F_X(x)\neq F_Y(x)$，但它们却有相同的各阶矩：

$$\int_0^\infty x^n f_X(x)dx=\frac{\Gamma(\frac{n+1}{a})}{\Gamma(\frac{1}{a})}(\cos a\pi)^{-\frac{n}{a}}=\int_0^\infty x^n f_Y(x)dx\ \ n=1,2,\cdots$$

结论：由随机变量的分布函数可以确定随机变量的数字特征，反之不然.

6．满足 E(XY)=E(X)E(Y)的 X,Y 未必独立

若随机变量 X,Y 独立，且各自数学期望存在，则 $E(XY)=E(X)E(Y)$,反之不真。

例如：设随机变量 Z 服从 $[0,2\pi]$ 上的均匀分布

$$X=\cos Z\qquad\qquad Y=\cos(Z+\frac{\pi}{2})$$

$$E(X)=\frac{1}{2\pi}\int_0^{2\pi}\cos zdz=0 \quad , \quad E(Y)=\frac{1}{2\pi}\int_0^{2\pi}\cos(z+\frac{\pi}{2})dz=0$$

$$E(XY)=\frac{1}{2\pi}\int_0^{2\pi}\cos z\cos(z+\frac{\pi}{2})dz=\frac{1}{2}\cos\frac{\pi}{2}=0$$

于是$E(XY)=E(X)E(Y)$ ，即 X,Y 不相关，而 X,Y 满足$X^2+Y^2=1$,所以 X,Y 不独立.

第五章　参数估计与假设检验

1．矩估计是否有唯一性

以样本矩阵作为相应的总体矩的估计量，以样本矩的连续函数作为相应的总体矩的连续函数的估计量。这种估计的方法被称为矩估计法

我们知道矩有原点矩与中心矩之分，如果样本原点矩与样本中心矩都可作为相应总体矩的估计，则显然矩估计不具有唯一性

即使规定只用样本的 k 阶原点矩作为相应总体 k 阶矩的估计，那么参数的矩估计也不是唯一的社，例如

设总体 X 服从参数为λ的泊松（Poisson）分布,由于$\lambda=E(X)=D(X)$,故由矩估计法知：

$$\hat{\lambda}_1=\overline{X} \quad , \quad \hat{\lambda}_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

都为参数λ的矩估计量.

2．矩估计不具有“不变性”

所谓极大似然估计具有“不变性”，是指：若$\hat{\theta}$为θ的一个极大似然估计，则当函数$g(\theta)$具有单值反函数时，$g(\hat{\theta})$为$g(\theta)$的一个极大的似然估计，对矩估计而言，“不变性”不成立，例如：

对反射正态分布

$$f(x)=\begin{cases}\sqrt{\frac{2}{\pi}}\frac{1}{\sigma}e^{-\frac{x^2}{2\sigma^2}} & x>0\\ 0 & x\le 0\end{cases} \quad (\sigma>0)$$

用矩估计法分别对σ和σ^2作估计

$$E(X)=\int_0^{\infty}xf(x)dx=\sqrt{\frac{2}{\pi}}\sigma$$

$$E(X^2)=\int_0^{\infty}x^2 f(x)dx=\sqrt{\frac{2}{\pi}}\sigma^2\int_0^{\infty}t^2 e^{-\frac{t^2}{2}}dt=\sigma^2$$

设$(X_1,X_2,\cdots,X_n)$为取自反射正态总体的样本，由矩阵令

$$\overline{X}=E(X)=\sqrt{\frac{2}{\pi}}\sigma \quad , \quad \frac{1}{n}\sum_{i=1}^{n}X_i^2=E(X^2)=\sigma^2$$

得： $$\hat{\sigma}=\sqrt{\frac{2}{\pi}}\overline{X} \quad , \quad \hat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}X_i^2$$

由此可见 $(\hat{\sigma})^2=\frac{2}{\pi}\overline{X}\neq\frac{1}{n}\sum_{i=1}^{n}X_i^2=\hat{\sigma^2}$

所以矩估计不满足不变性.

3．极大的似然估计是否有唯一性

极大似然估计不具有唯一性，例如

设均匀分布的密度函数

$$f(x;\theta)=\begin{cases}1 & \theta-\frac{1}{2}\le x_{(1)}\le x_{(n)}\le\theta+\frac{1}{2}\\ 0 & \text{其他}\end{cases} \qquad \theta\text{为待估参数}$$

似然函数：

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

$$=\begin{cases}1 & \theta-\frac{1}{2}\le x_i\le\theta+\frac{1}{2}(1\le i\le n)\\ 0 & \text{其他}\end{cases}$$

$$=\begin{cases}1 & \theta-\frac{1}{2}\le x_{(1)}\le x_{(n)}\le\theta+\frac{1}{2}\\ 0 & \text{其他}\end{cases}$$

$$=\begin{cases}1 & x_{(n)}-\frac{1}{2}\le\theta\le x_{(1)}+\frac{1}{2}\\ 0 & \text{其他}\end{cases}$$

显然区间$[x_{(n)}-\frac{1}{2},x_{(1)}+\frac{1}{2}]$中的每一点均为$L(\theta)$的最大值点

于是$\hat{\theta}=X_{(n)}-\frac{1}{2}+t[X_{(1)}-X_{(n)}+1]$ $\quad 0\le t\le 1$， 均为θ的极大似然估计量

注： 当$X_{(n)}-\frac{1}{2}=X_{(1)}+\frac{1}{2}$时，$\theta$的极大似然估计唯一

当 $X_{(n)}-\frac{1}{2}>X_{(1)}+\frac{1}{2}$ 时，θ 的极大似然估计不存在.

4．似然方程的解未必是极大似然估计

设 $(X_1,X_2,\cdots,X_n)$ 是取自柯西（Cauchy）分布

$$f(x;\theta)=\frac{1}{\pi[1+(x-\theta)^2]} \qquad -\infty<x<\infty$$

的样本

当 $n=1$ 时，似然函数为 $L(\theta)=\frac{1}{\pi[1+(x_1-\theta)^2]}$

$$\ln L(\theta)=-\ln\pi-\ln[1+(x_1-\theta)^2]$$

似然方程为 $\frac{\partial \ln L}{\partial\theta}=\frac{2(x_1-\theta)}{1+(x_1-\theta)^2}=0$

故 $\hat{\theta}=X_1$ 是 θ 的极大似然估计量

当 $n=2$ 时，似然函数为

$$L(\theta)=\frac{1}{\pi[1+(x_1-\theta)^2][1+(x_2-\theta)^2]} \qquad (1)$$

要使 $L(\theta)$ 达到最大，只要（1）的分母最小，令其分母的变量部分为

$$f(\theta)=[1+(x_x-\theta)^2][1+(x_2-\theta)^2]$$

由 $$f'(\theta)=-2(x_1+x_2-2\theta)[\theta^2-(x_1+x_2)\theta+x_1x_2+1]=0 \qquad (2)$$

得三个解：$\theta_1=\frac{X_1+X_2}{2}$ ，$\theta_{2,3}=[(X_1+X_2)\pm\sqrt{(X_1-X_2)^2-4}]/2$

通过 $f''(\theta)$ 的正负可判得：当 θ_1 为 θ 的极大似然估计时，$\theta_{2,3}$ 不是 θ 的极大似然估计；反之，当 $\theta_{2,3}$ 为 θ 的极大似然估计时，θ_1 不是 θ 的极大似然估计。而无论发生什么情况，似然方程（2）的三个解不全是 θ 的极大似然估计

极大似然估计就是求似然函数的极大值点，似然函数是样本的联合分布密度，似然方程是似然函数的导函数为零的函数。我们知道，导函数为零的点是函数的驻点，而驻点未必是函数的极大值点，所以似然方程的解未必是极大似然估计.

5．参数估计的无偏性与一致性有无关系

两者间没有一定的关系，下表中所列的四种情况均可以发生

	无偏性	有偏性
一致性	√	√
非一致	√	√

以均匀分布为例 $(X_1, X_2, \cdots, X_n)$ 为如下均为总体的一个样本

$$X \sim f(x;\theta) = \begin{cases} \dfrac{1}{\theta} & 0 < x \le \theta \\ 0 & \text{其他} \end{cases}$$

由极大似然估计法可得 $\hat{\theta} = X_{(n)}$ 的总体参数 θ 的一个极大似然估计，且唯一

$X_{(n)}$ 的概率密度为 $f_{X_{(n)}}(x) = \begin{cases} \dfrac{nx^{n-1}}{\theta^n} & 0 < x \le \theta \\ 0 & \text{其他} \end{cases}$

$$E（X_{(n)}) = \int_{-\infty}^{\infty} x f_{X_{(n)}}(x)dx = \frac{n\theta}{n+1} \qquad E(\frac{n+1}{n}X_{(n)}) = \theta$$

可见 $X_{(n)}$ 不是 θ 的无偏估计，$\dfrac{n+1}{n}X_{(n)}$ 是 θ 的无偏估计

$$D(X_{(n)}) = E(X_{(n)}^2) - (E(X_{(n)}))^2 = \frac{n\theta^2}{n+2} - (\frac{n\theta}{n+1})^2 = \frac{n\theta^2}{(n+2)(n+1)^2}$$

由切贝雪夫不等式的推广（Bienayme 不等式的特例）

$$P(|X_{(n)} - \theta| \ge \varepsilon) \le \frac{E(X_{(n)} - \theta)^2}{\varepsilon^2} = E[(X_{(n)} - \frac{n\theta}{n+1}) - (\theta - \frac{n\theta}{n+1})]^2 / \varepsilon^2$$

$$= \frac{D(X_{(n)} + (\theta - \dfrac{n\theta}{n+1})^2}{\varepsilon^2} \to 0 \quad (n \to \infty)$$

由切贝雪夫不等式

$$P(|\frac{n+1}{n}X_{(n)} - \theta| \ge \varepsilon) \le \frac{(\dfrac{n+1}{n})^2 D(X_{(n)})}{\varepsilon^2} \to 0 \qquad (n \to \infty)$$

由此知有偏估计 $X_{(n)}$ 和无偏估计 $\dfrac{n+1}{n}X_{(n)}$ 都是 θ 的一致估计

再引进 θ 的两个估计 X_1 与 $2X_1$

$$E(X_1) = \int_0^{\theta} \frac{x}{\theta}dx = \frac{\theta}{2} \qquad E(2X_1) = \theta$$

可见 X_1 是 θ 的有偏估计，$2X_1$ 是 θ 的无偏估计

对任意$\varepsilon > 0$，有

$$P(|X_1-\theta|\geq\varepsilon)=\int_{|x-\theta|\geq\varepsilon} f(x;\theta)dx=\frac{1}{\theta}\int_0^{\theta-\varepsilon} dx=\frac{\theta-\varepsilon}{\theta}\not\to 0 \quad (n\to\infty)$$

$$P(|2X_1-\theta|\geq\varepsilon)=\int_{|2x-\theta|\geq\varepsilon} f(x;\theta)dx=\frac{1}{\theta}[\int_{\frac{\theta+\varepsilon}{2}}^{\theta} dx+\int_0^{\frac{\theta-\varepsilon}{2}} dx]=\frac{\theta-\varepsilon}{\theta}\not\to 0 \qquad (n\to\infty)$$

由此得有偏估计X_1和无偏估计$2X_1$都不是θ的一致估计.

6．无偏估计是否唯一的

仅由估计量的数学期望等于被估计的参数这一要求并不能保证估计量的唯一性

例如：设$(X_1, X_2, \cdots, X_n)$是来自参数为λ的泊松分布的总体的一个简单随即样本，则

$$\hat{\lambda}_1=\overline{X}=\frac{1}{n}\sum_{i=1}^{n} X_i \qquad \hat{\lambda}_2=S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

都是λ的无偏估计，事实上，由基本统计量的性质知

$$E(\hat{\lambda}_1)=E(\overline{X})=E(X)=\lambda \qquad E(\hat{\lambda}_2)=E(S^2)=D(X)=\lambda$$

又如，设总体$X\sim U(0,\theta]$的一个简单随机样本为$(X_1, X_2, \cdots, X_n)$，则θ的估计量$\hat{\theta}_1=2\overline{X}$与$\hat{\theta}_2=\frac{n+1}{n}\max\{X_1, X_2, \cdots, X_n\}$均为$\theta$的无偏估计，事实上，

$$E(\hat{\theta}_1)=E(2\overline{X})=2E(X)=2\cdot\frac{\theta}{2}=\theta \quad E(\hat{\theta}_2)=\frac{n+1}{n}\int_0^{\theta}\frac{n}{\theta^n}y^{n-1}ydy=\theta$$

其中$Y=\max\{X_1, X_2, \cdots, X_n\}$，其密度函数为

$$f_Y(y)=\begin{cases}\frac{n}{\theta^n}y^{n-1} & 0<y\leq\theta \\ 0 & \text{其他}\end{cases}$$

7．零假设与备择假设是否处于对等的地位

在假设检验中，首先要针对具体问题提出零假设H_0和备择假设H_1，由于零假设是作为检验的前提而提出来的，因此，零假设通常应受到保护，没有充足的证据是不能备拒绝的，而备择假设只有当零假设备被拒绝后，才能被接受，这就决定了零假设与备择假设不是处于对等的地位

下面举例说明交换零假设与备择假设可能会得出截然相反的检验结论

问题 某厂方断言，本产生产的小型电动机在正常负载条件下平均电流不会超过

0.8A,随机抽取该型号电动机 16 台，发现其平均电流为 0.92A，而由该样本求出的标准差是 0.32A，假定这种电动机的工作电流 X 服从正态分布，问根据这一抽样结果，能否否定厂方断言？（取显著性水平 $\alpha = 0.05$）

解：本题假定 $X \sim N(\mu, \sigma^2), \sigma^2$ 未知，以厂方断言作为零假设，则得假设检验问题： H_0： $\mu \le 0.8$ ； H_1： $\mu > 0.8$

此时 $n = 16, \bar{x} = 0.92, S = 0.32, \alpha = 0.05$，由 t 检验法可知拒绝域为 $\overline{X} > 0.8 + \frac{0.32}{\sqrt{16}} t_{0.05}(16-1) = 0.8 + \frac{0.32}{\sqrt{16}} \times 1.753 \approx 0.94$，由于 $\bar{x} = 0.92 < 0.94$，故不应该拒绝零假设 H_0，即在所给数据和检验水平下，没有充分理由否定厂方的断言

现在若把厂方断言的对立面（即 $\mu > 0.8$）作为零假设，则得假设检验问题：

H_0： $\mu > 0.8$ ； H_1： $\mu \le 0.8$

由 t 检验法，此时的拒绝域为 $\overline{X} \le 0.8 - \frac{0.32}{\sqrt{16}} t_{0.05}(16-1) = 0.66$，因为观测值 $\bar{x} = 0.92 > 0.66$，所以应接受零假设，即接受厂方断言的对立面

由此可见，随着问题提法的不同，得出了截然相反的结论，这一点会使初学者感到迷惑不解，实际上，这里有个着眼点不同的问题，当把“厂方断言正确”作为零假设时，我们根据该厂以往的表现和信誉，对其断言已有了较大的信任，只有很不利于它的观察结果才能改变我们的看法，因而一般难以拒绝这个断言，反之，当把“厂方断言不正确”作为零假设时，我们一开始就对该厂产品抱怀疑态度，只有很有利于该厂的观察结果，才能改变我们的看法，因此，在所得观察数据并非决定性地偏于一方时，我们的着眼点(即最初立场)决定了所得的结论

打一个通俗的比喻：某人是嫌疑犯，有些不利于他的证据，但并非是起决定性作用的，若我们要求“只有决定性的不利于他的证据才能判他有罪“，则他将被判为无罪，反之，若要”只有决定性的有利于他的证据才能判他无罪“，则他将被判有罪，在这里，也是着眼点的不同决定了看法，这类事件在日常生活中并不少见，原本不足为奇.