《统计与数据分析》
Statistics & Data Analysis

§6 推断统计学的假设检验

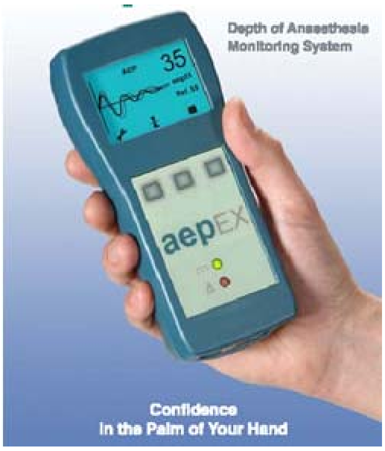
Zhu Huaiqiu
@Peking University

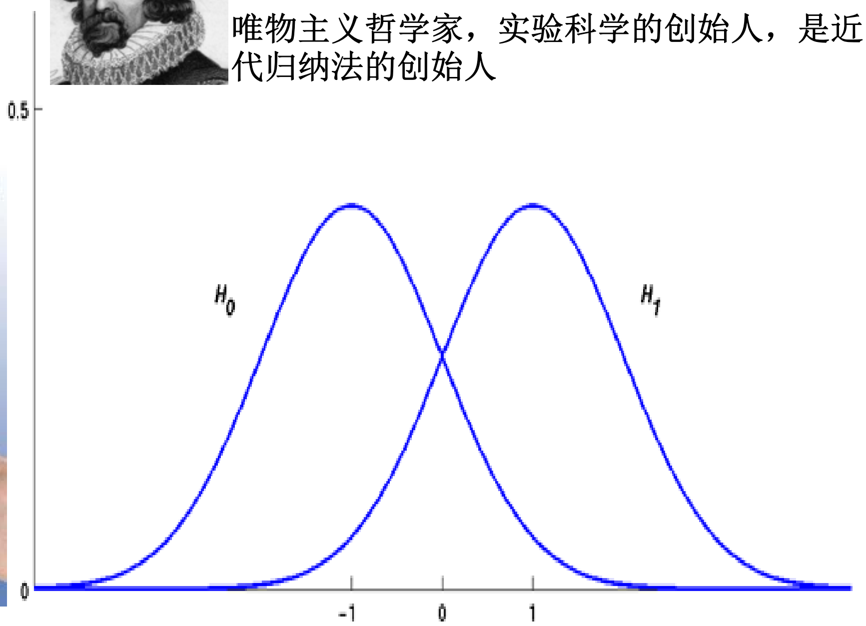
The strongest arguments prove nothing so long as the conclusions are not verified by experience. Experimental science is the queen of sciences and the goal of all *speculation*.

(Francis Bacon (1561一1626)

"英国唯物主义和整个现代实验科学的真正始祖"

唯物主义哲学家，实验科学的创始人，是近代归纳法的创始人



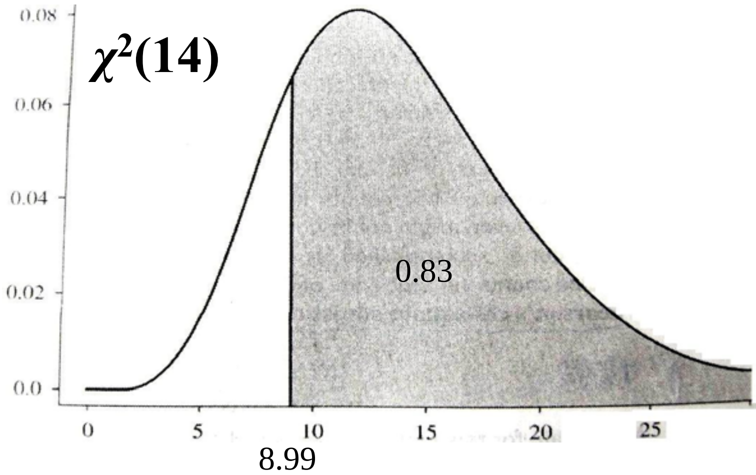


α粒子数的Poisson分布研究

n	0-2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17+
Observed	18	28	56	105	126	146	164	161	123	101	74	53	23	15	9	5
Expected	12.2	27.0	56.5	94.9	132.7	159.1	166.9	155.6	130.6	99.7	69.7	45.0	27.0	15.1	7.9	7.1
χ^2	2.76	0.04	0.01	1.07	0.34	1.08	0.05	0.19	0.44	0.02	0.27	1.42	0.59	0.00	0.57	0.57

问题 估计参数：根据观察的数据来估计Poisson分布的参数λ

讨论：
1. 估计值的概率分布性质
2. 参数估计的数学方法
3. 总体分布的假设检验
4. 总体分布的拟合检验



假设检验（hypothesis test）：

根据（简单随机抽样）抽取的（简单随机）样本信息来判别总体是否具有某种性质

——总体X

——（简单随机抽样）样本$X_1, X_2, \ldots, X_n$

——原假设（null hypothesis）：总体X的某一种概率分布性质（参数、分布类型）

——对立假设（alternative hypothesis）：总体X的另一种概率分布性质（参数、分布类型）

§6.1 假设检验概述

§6.1.1 问题的提出

【Example 6.1】某药品生产车间用粉剂定量自动包装机包装粉剂药品，每袋标准重量为50mg。长期实践表明该设备包装的这一药品重量服从正态分布，且标准差为1.5mg。
现从某天的包装产品中随机抽取9袋，精确秤得它们的重量分别为
49.5, 50.6, 51.8, 52.1, 49.3, 51.1, 52.0, 51.5, 50.0
它们的平均值为50.9mg。
问当日该包装机工作是否正常?

问题：假定方差不变，根据样本数据的平均值50.9来判断正态分布$N(50, 1.5^2)$是否成立?

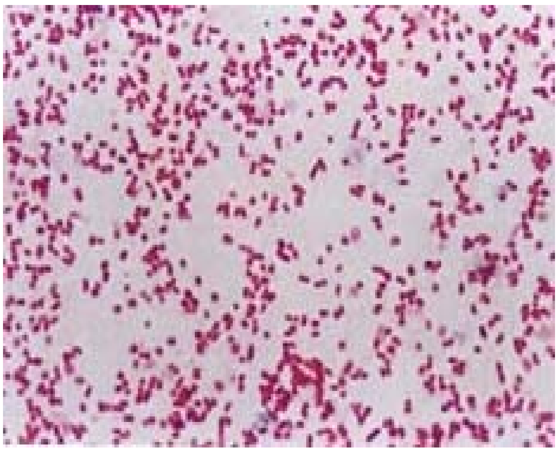
【Example 6.2】将0.1ml受细菌污染的牛奶均匀涂在1cm²的切片上，用显微镜观察切片每个小网格内的细菌菌落数目。根据400（20×20）个小网格的计数结果，统计出如下表格:

菌落数:	0	1	2	3	4	5	6	7	8	9	10	19
频　数:	56	104	80	62	42	27	9	9	5	3	2	1

试问单位面积内的菌落数是否服从Poisson分布?

问题:

根据样本信息，判别随机变量X是否服从Poisson分布。



（Bliss C. and Fisher R. A., (1953). Fitting the negative binomial distribution to biological data. *Biometric*, 9: 174-200）

参数假设检验问题（【例6.1】）

——已知总体的分布函数类型，对分布函数中的未知参数提出某种假设
——要求利用样本的信息对所提出的假设进行检验，根据检验结果作出接受或者拒绝所提假设的判断

非参数假设检验问题（【例6.2】）

——总体的分布类型未知
——要求根据样本的信息来对分布类型的假设进行检验，对总体的分布类型作出判断

§ 6.1.2 假设检验的基本思想与 N-P理论

【Example 6.1】某药品生产车间用粉剂定量自动包装机包装粉剂药品，每袋标准重量为50mg。长期实践表明该设备包装的这一药品重量服从正态分布，且标准差为1.5mg。
现从某天的包装产品中随机抽取9袋，精确秤得它们的重量分别为
49.5，50.6，51.8，52.1，49.3，51.1，52.0，51.5，50.0
其平均值为50.9mg。
问当日该包装机工作是否正常？

假设方差不变，根据样本的信息，正态分布$N(\mu, \sigma^2)$的$\mu = \mu_0=50.0$是否成立?

$N(\mu, \sigma^2)$:
——Given σ^2
——$\mu = \mu_0=50.0$?

Yes

Observe

样本
观察值

No

原假设
(null hypothesis)

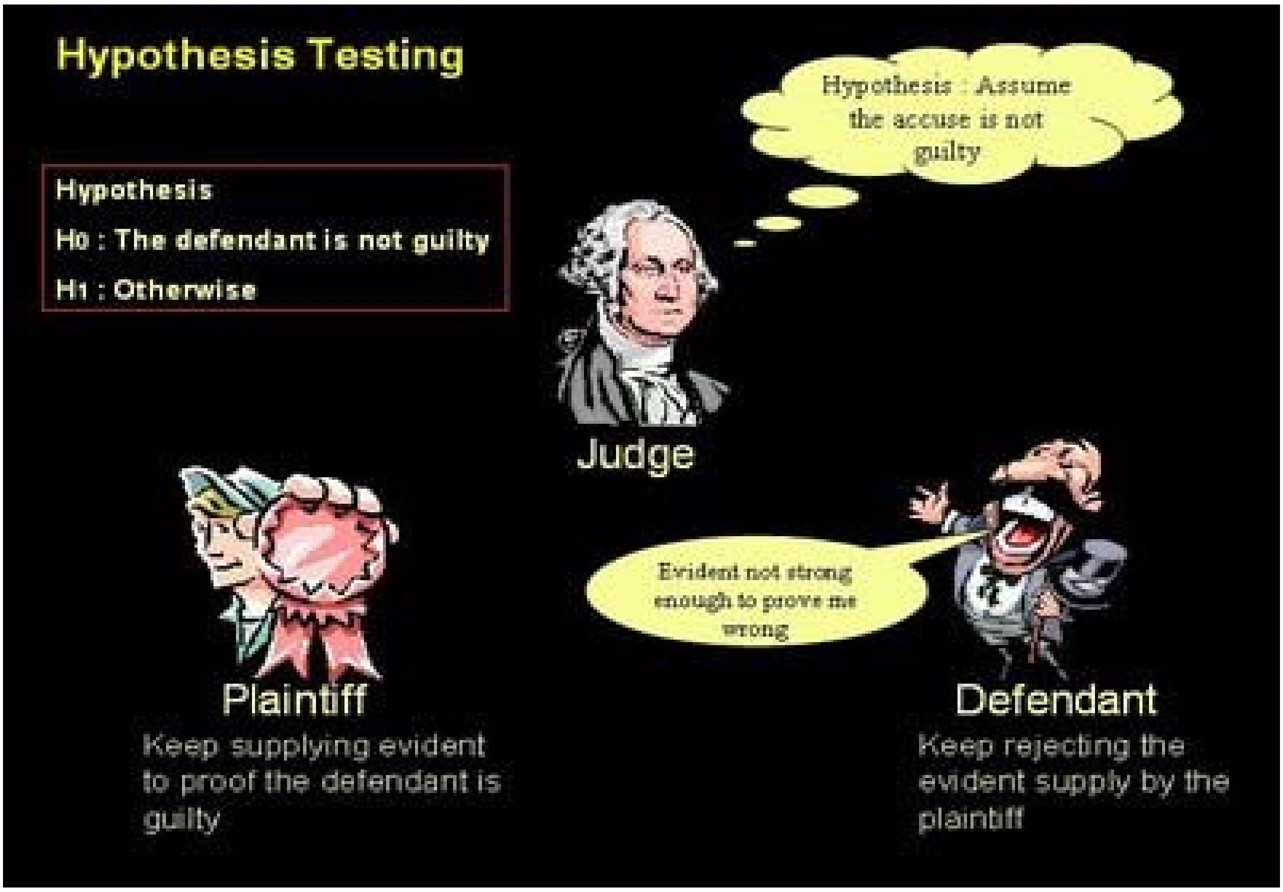
H_0: $\mu = \mu_0$?

对立假设
(alternative hypothesis)

H_1: $\mu \neq \mu_0$?

H_0 and H_1: which can we accept based on the sampling evidence?

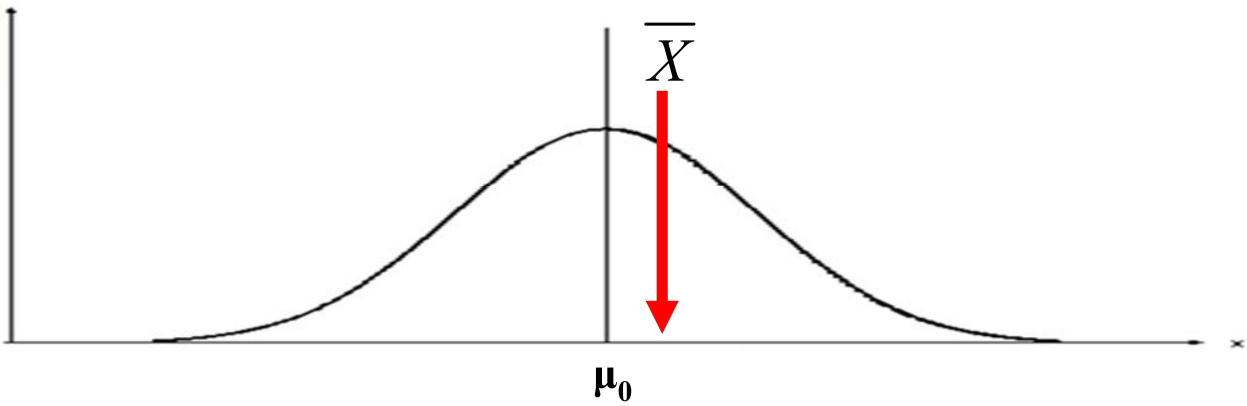
Jury system



【Example 6.1】某药品生产车间用粉剂定量自动包装机包装粉剂药品，每袋标准重量为50mg。长期实践表明该设备包装的这一药品重量服从正态分布，且标准差为1.5mg。
现从某天的包装产品中随机抽取9袋，精确秤得它们的重量分别为
49.5，50.6，51.8，52.1，49.3，51.1，52.0，51.5，50.0
其平均值为50.9mg。
问当日该包装机工作是否正常?

问题：给定$\sigma^2=1.5$，H_0: $\mu=\mu_0=50.0$；H_1: $\mu\neq\mu_0$

$$H_0 \Rightarrow E(\overline{X}) = \mu_0$$



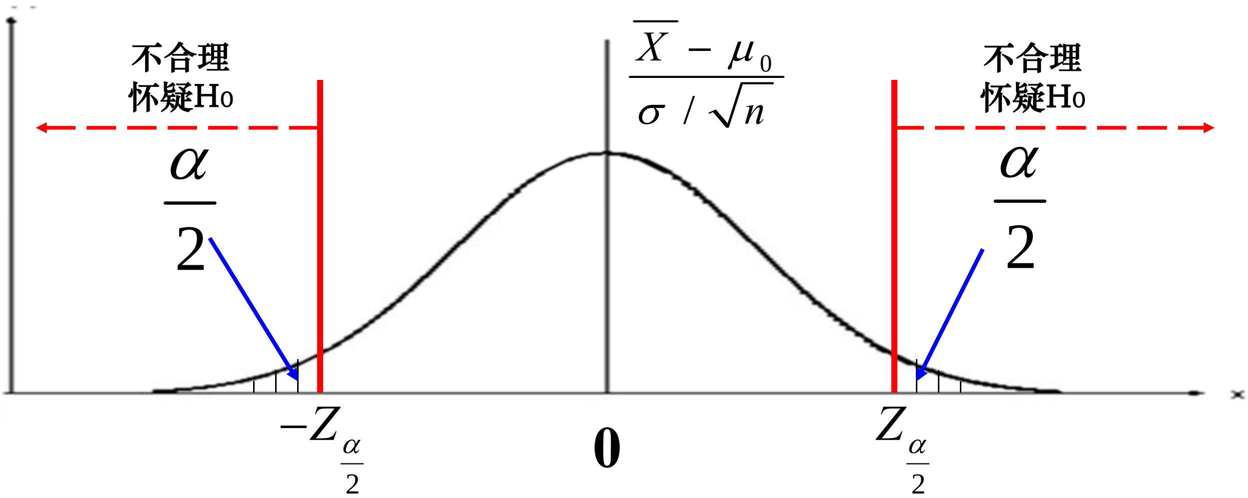
若H_0成立，则有：

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

中心极限定理

构造概率为α的小概率事件为不合理的后果（如 α =0.001, 0.01, 0.05, 0.10）：

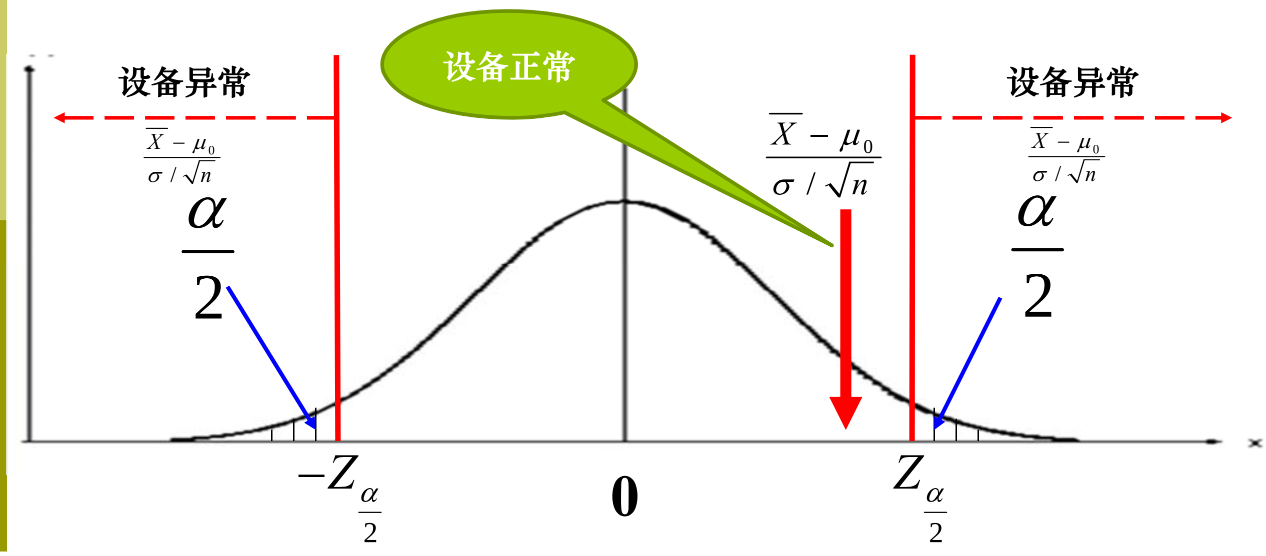
$$P\left\{\left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| > Z_{\frac{\alpha}{2}}\right\} = \alpha$$



令 $\alpha = 0.05$ 有 $Z_{\frac{\alpha}{2}} = 1.96$

$\overline{X} = 50.9, \quad \mu_0 = 50.0$

$$\left| \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \right| = \left| \frac{50.9 - 50.0}{1.5 / \sqrt{9}} \right| = 1.8 < Z_{\frac{\alpha}{2}} = 1.96$$



若H_0成立：

U检验or Z检验 $\longrightarrow$ $$\boxed{\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}} \sim N(0,1)$$

检验统计量
test statistic

构造不合理的小概率事件（如小概率α=0.001, 0.01, 0.05, 0.1)：

$$P\left\{ \left| \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \right| > Z_{\frac{\alpha}{2}} \right\} = \alpha$$

显著性水平
significance level

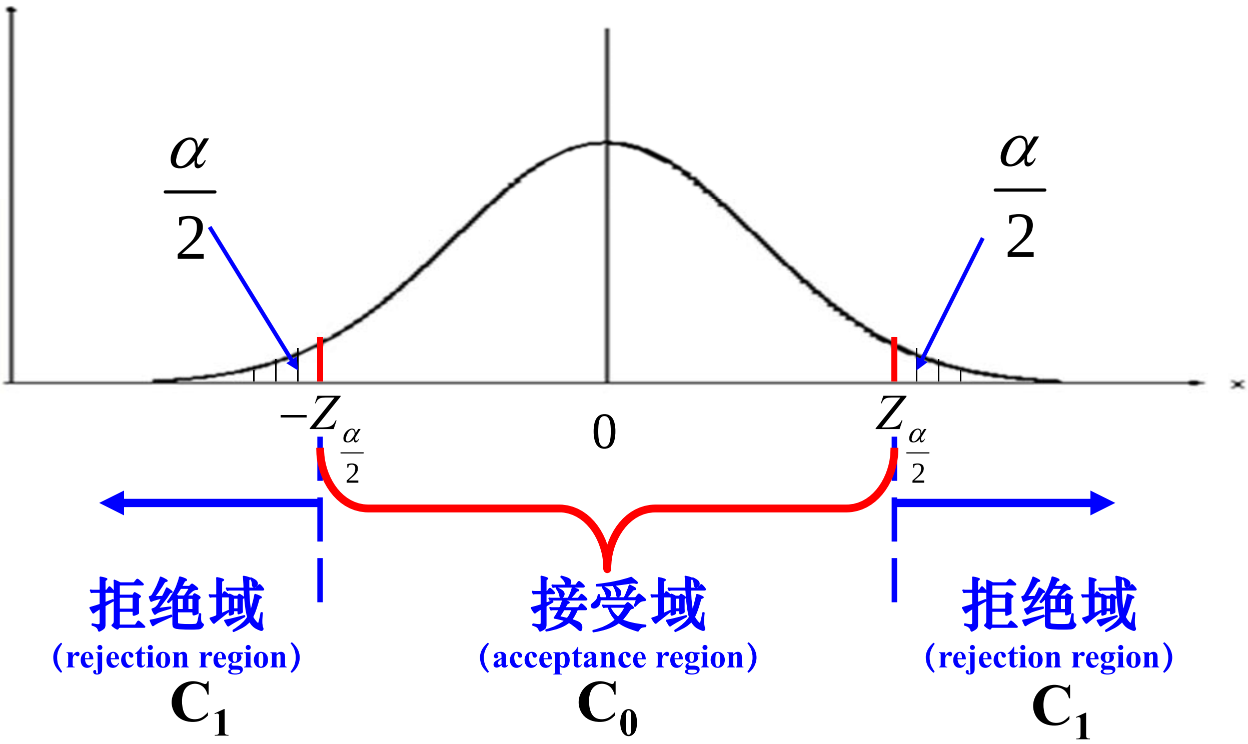
检验统计量（test statistic）

为检验原假设和对立假设，根据样本（随机变量）、样本容量和总体参数构的随机变量，具有较明确的、或渐进的概率分布性质

显著性水平（significance level）

根据检验统计量的（渐进）概率分布性质，定义不合理的小概率事件来推断拒绝原假设发生的条件

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$$



假设检验的检验错误

检验错误：抽样的随机性、实际推断中小概率事件仍然有可能发生

第一类错误（Type I error）（弃真错误）：

原假设H_0为真，但经过检验统计量观察值的判断，H_0被拒绝。

$$P\{\text{拒绝}H_0|H_0\text{为真}\}=\alpha$$

第二类错误（Type II error）（纳伪错误）：

原假设H_0非真，但经过检验统计量观察值的判断，H_0被接受。

$$P\{\text{接受}H_0|H_0\text{非真}\}=\beta$$

或者 $$P\{\text{拒绝}H_0|H_0\text{非真}\}=1-\beta$$

假设检验的两类错误

α： 显著性水平

$(1-\beta)$：检验功效（Power of test）

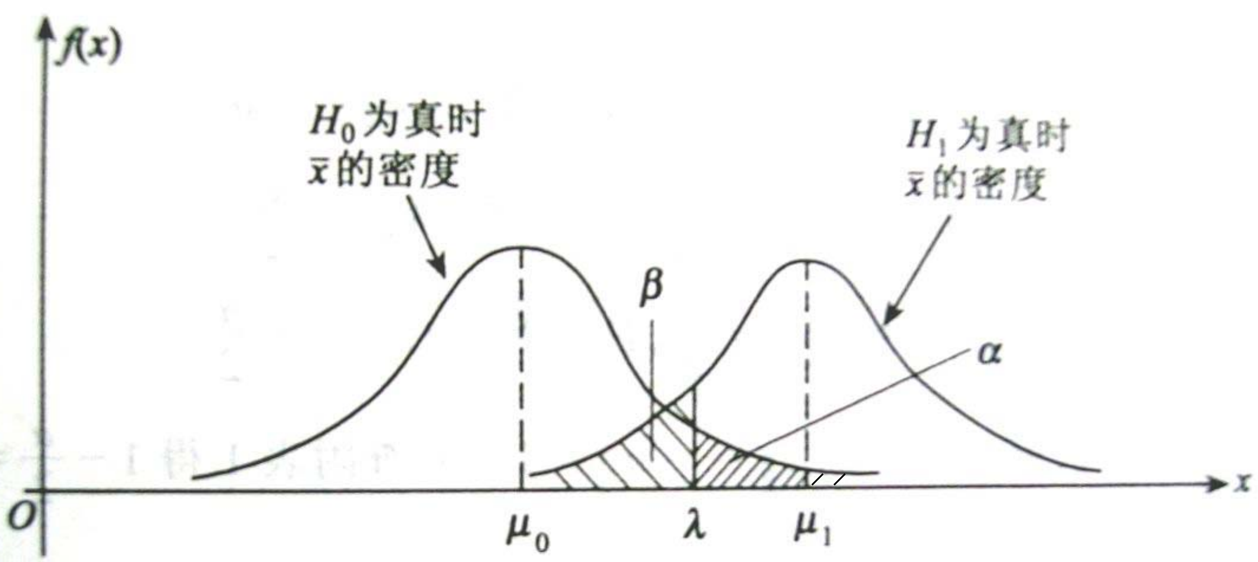
$(1-\alpha)$：置信水平（Confidence Level）

第一类错误α与第二类错误β的关系

【Example 6.3】设总体$X \sim N(\mu, \sigma^2)$，σ^2已知，而μ只能取两个值μ_0、μ_1，且$\mu_0<\mu_1$，现有总体X中抽取的容量为n的样本$X_1, X_2, \ldots, X_n$，在显著性水平α下检验假设：

$$H_0: \mu=\mu_0; \quad H_1: \mu = \mu_1(>\mu_0)。$$

请讨论第一类错误与第二类错误的关系。



	H_0 is true	H_0 is false *(H_1 is true)*
Accept H_0	**Correct Decision** $1-\alpha$: **Confidence Level**	**Type II Error** β
Reject H_0	**Type I Error** α : **Significance level**	**Correct Decision** $1-\beta$: **Power of a test**

Sensitivity (Sn), Specificity (Sp), PPV (Positive Predictive Value), and NPV (Negative Predictive Value)

The Truth

Test Score:	(H_0 is true) Has the disease	(H_0 is false) Does not have the disease	
Positive (Accept H_0)	True Positives (TP) a	b False Positives (FP) **Type II Error**	**PPV =** $\frac{TP}{TP + FP}$
Negative (Reject H_0)	c False Negatives (FN) **Type I Error**	d True Negatives (TN)	**NPV =** $\frac{TN}{TN + FN}$

Sensitivity (Sn)

$$\frac{TP}{TP + FN}$$

Or, $$\frac{a}{a + c}$$

Specificity (Sp)

$$\frac{TN}{TN + FP}$$

$$\frac{d}{d + b}$$

注意：
统计检验的阳性(positive)、阴性(negative)、假阳性(false positive)与假阴性(false negative)与医学诊断的概念有时不一致！

Sensitivity (Sn):

A/(A + C) × 100

20/30 × 100 = 66.7%

Specificity (Sp):

D/(D + B) × 100

37/70 × 100 = 52.9%

Positive Predictive Value (PPV):

A/(A + B) × 100

20/53 × 100 = 37.7%

Negative Predictive Value (NPV):

D/(D + C) × 100

37/47 × 100 = 78.7%

Truth

Test Result	Disease (number)	Non Disease (number)	Total (number)
Positive (number)	20 A (True Positive)	33 B (False Positive)	53 $T_{Test\ Positive}$
Negative (number)	10 C (False Negative)	37 D (True Negative)	47 $T_{Test\ Negative}$
	30 $T_{Disease}$	70 $T_{Non\ Disease}$	100 Total

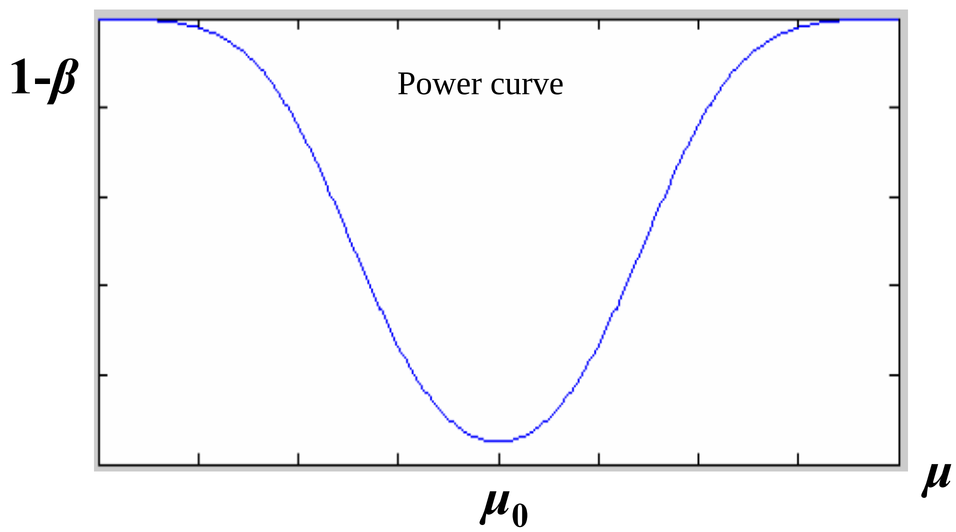
讨论:

（1）Sn与PPV分别对应控制Type I error与控制Type II error的水平或精度

（2）Sp与NPV分别对应控制Type I error与控制Type II error的水平或精度

关于检验功效(1-β)的讨论

双边检验的功效函数呈现“U”字形，“U”展得越宽，检验的功效越差。



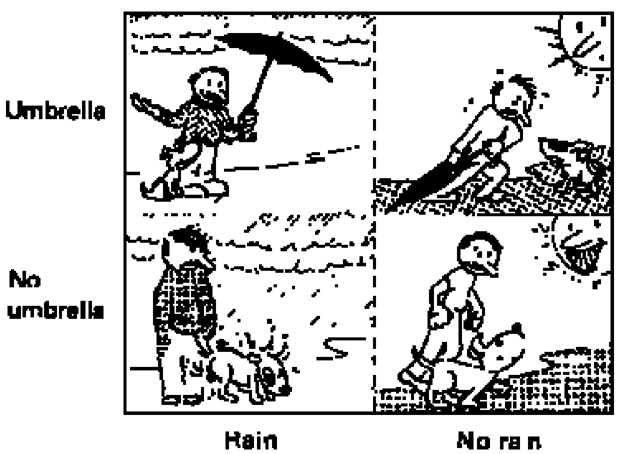
假设检验的Neyman-Pearson理论

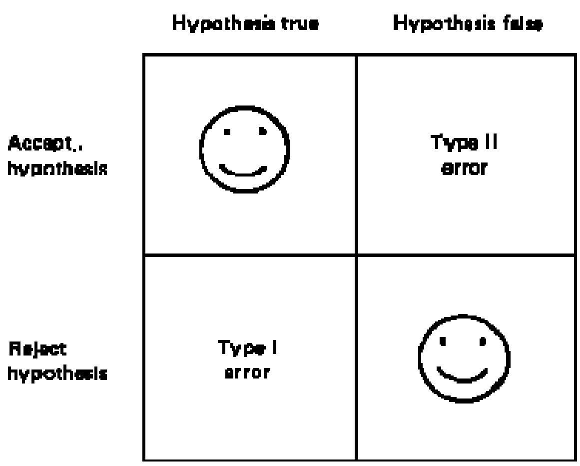
在控制出现第一类错误α的条件下，寻求使出现第二类错误β尽可能小（或检验功效$1-\beta$尽可能大）的检验。

在实际操作中，当样本容量n给定时，优先考虑对出现第一类错误的概率并加以控制，再适当考虑出现第二类错误的概率，即显著性检验优先的基本原则。

Two types of error

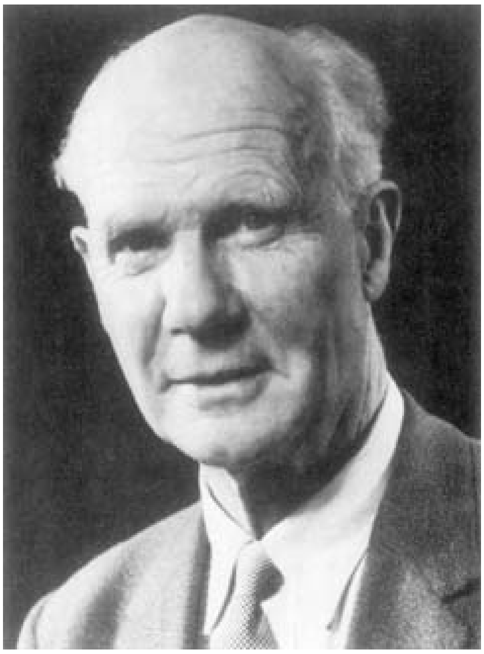
It's cloudy this morning, and I wonder whether or not I should take my umbrella. There are two possible kinds of error: either I take my umbrella and it doesn't rain, or I don't take it and it does rain. We can't prevent these errors, but we can control them by weighing the *relative costs* of one type versus the other. Most people would say that carting around an unused umbrella is less of a cost than getting soaked in a downpour, but each individual has to decide.





Jerzy Neyman
(1894-1981)
美国波兰裔统计学家

Egon Pearson
(1895-1980)
英国统计学家



1928—1938年期间合作发表一系列文章，建立了假设检验的一种严格的数学理论。

原假设的分类

简单假设（Simple hypothesis）

——指定了总体X的分布形式和参数具体的取值，

如：$X\sim N(\mu, \sigma^2)$，σ^2已知，$H_0: \mu=\mu_0$（点信息）

——简单假设的参数自由度为0；

复合假设（Composite hypothesis）

——对于总体X的分布形式未知，或者虽然指定了分布形式但参数值没有完全具体指定，即复合假设的参数自由度≥1。

如：

（1）$X\sim N(\mu_0, \sigma^2)$，μ_0已知，σ^2未指定（区间信息）。

（2）X～某种Poisson分布，其参数λ未指定（区间信息）；

（3）$X\sim B(100, p)$， $p>0.5$ （区间信息）。

对立假设的分类

单边对立假设（one-sided alternative hypothesis）

在原假设指定或部分指定参数时，原假设参数的互补区域（即对立假设的参数区间）由一个连续区间构成

单边假设对应的假设检验问题称为单边检验。

双边对立假设（two-sided alternative hypothesis）

在原假设指定或部分指定参数时，原假设参数的互补区域（即对立假设的参数区间）由两个不相交（不连通）的连续区间构成

双边假设对应的假设检验问题称为双边检验。

假设检验分类通常以单边、双边分类，关心的是
（1）检验统计量的分布关于不合理事件的定义；
（2）给定显著性水平下检验功效的问题。

§6.1.3 假设检验小结

一、假设检验过程中两个重要的经典统计学思想

（1）反证法思想

保护原假设的思想，没有充分的理由就不能拒绝原假设。

（2）小概率原理

小概率事件在一次试验中几乎是不会发生的，如果小概率事件发生了，就认为出现了不合理现象。

二、假设检验的一般步骤

（1）根据实际问题提出原假设H_0和对立假设H_1；

（2）选择合理的检验统计量，根据原假设H_0和对立假设H_1来确定拒绝域的表达式；

（3）选取适当的显著性水平α，求出使出现第一类错误的概率不超过α的临界值、拒绝域C_1（和接受域C_0）；

（4）根据样本观察值计算检验统计量的观察值，判断属于拒绝域或接受域，从而拒绝或接受原假设H_0。

§ 6.1.4 置信区间与接受域的对偶性

置信区间

以一定的置信度（$100(1-\alpha)\%$）给出未知参数的所在范围 。

接受域

在某一显著性水平α下判定未知参数是否接受原假设给定的取值。

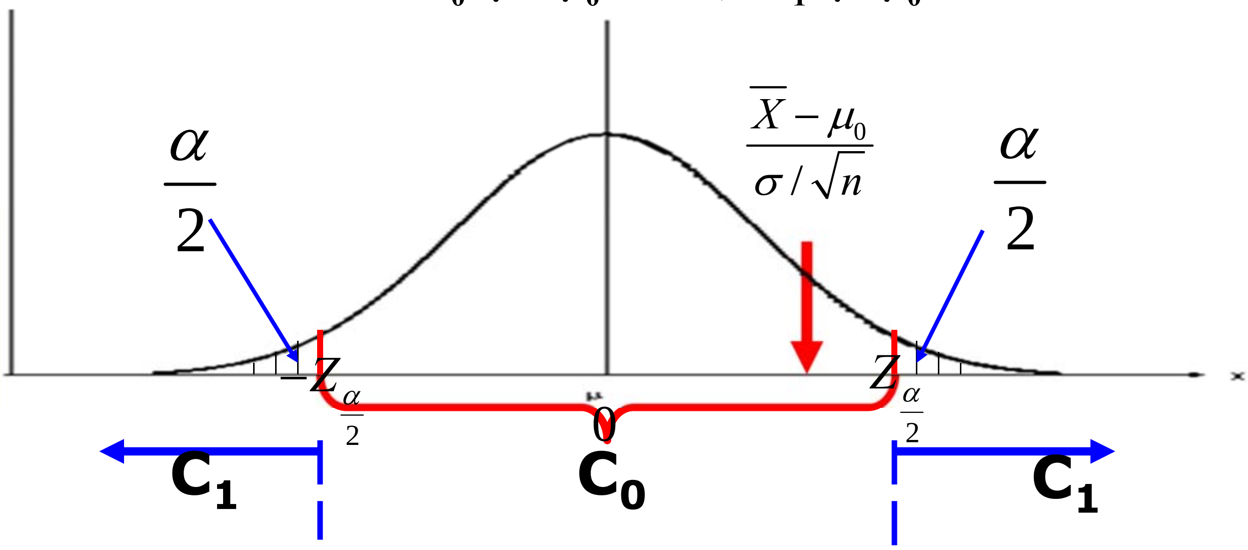
——二者的解决途径和方法都相通

【Example 6.1】某药品生产车间用粉剂定量自动包装机包装粉剂药品，每袋标准重量为50mg。长期实践表明该设备包装的这一药品重量服从正态分布，且标准差为1.5mg。
现从某天的包装产品中随机抽取9袋，精确秤得它们的重量分别为

49.5, 50.6, 51.8, 52.1, 49.3, 51.1, 52.0, 51.5, 50.0

其平均值为50.9mg。
问当日该包装机工作是否正常？

问题：给定$\sigma^2=1.5$，H_0: $\mu=\mu_0=50.0$；H_1: $\mu\neq\mu_0$



定理

设Θ为随机变量X的分布函数$f(X\,|\theta)$的参数空间，对任一$\theta_0\in\Theta$，都存在显著性水平α下的对假设H_0: $\theta=\theta_0$的检验，且接受域为$A(\theta_0)$。则集合$C(X)=\{X\in A(\theta)\}$为关于θ的$100(1-\alpha)\%$置信区间。

定理

设Θ为随机变量X的分布函数$f(X\,|\theta)$的参数空间，若$C(X)$为关于θ的$100(1-\alpha)\%$置信区间，即

$$P\left\{\theta_0\in C(X)\,|\,\theta=\theta_0\right\}=1-\alpha$$

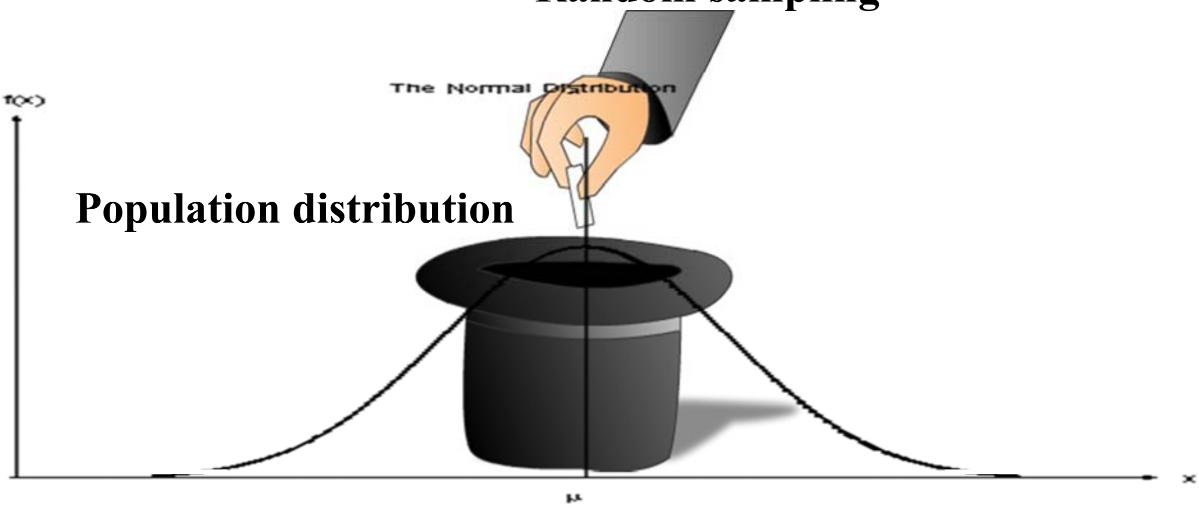
则在显著性水平α下的对假设H_0: $\theta=\theta_0$的检验，接受域为

$$A(\theta_0)=\left\{X\,|\,\theta_0\in C(X)\right\}$$

§6.2 正态总体的参数假设检验

Sample: $X_1, X_2, \ldots, X_n$

Random sampling



§6.2.1 单个总体均值μ的检验

一、均值μ的双边检验

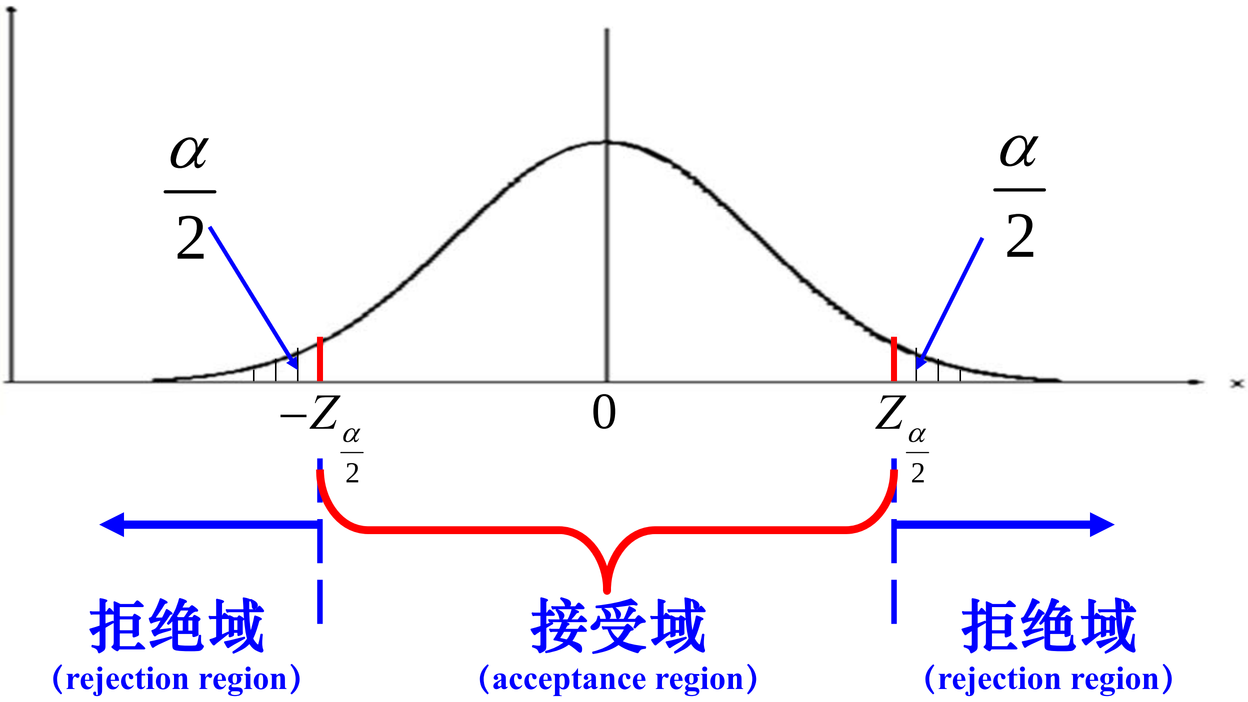
1. 已知方差σ^2，检验H_0: $\mu=\mu_0$，H_1: $\mu\neq\mu_0$

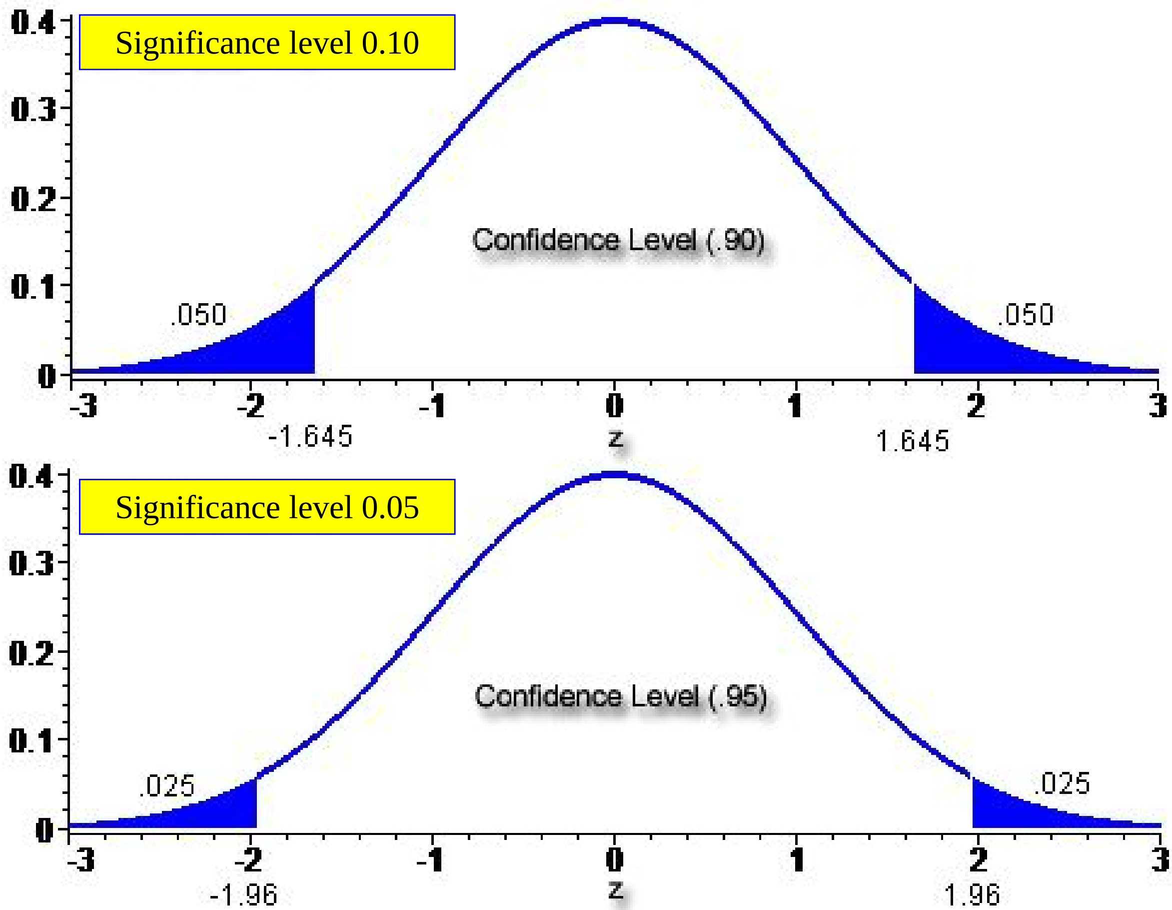
设总体$X\sim N(\mu,\sigma^2)$，σ^2已知，$X_1, X_2, \ldots, X_n$是来自总体X的（简单随机抽样）样本。

采用U检验，检验统计量为：

$$U=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$$

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$$





2. 未知方差σ^2，检验$H_0: \mu=\mu_0$，$H_1: \mu \neq \mu_0$

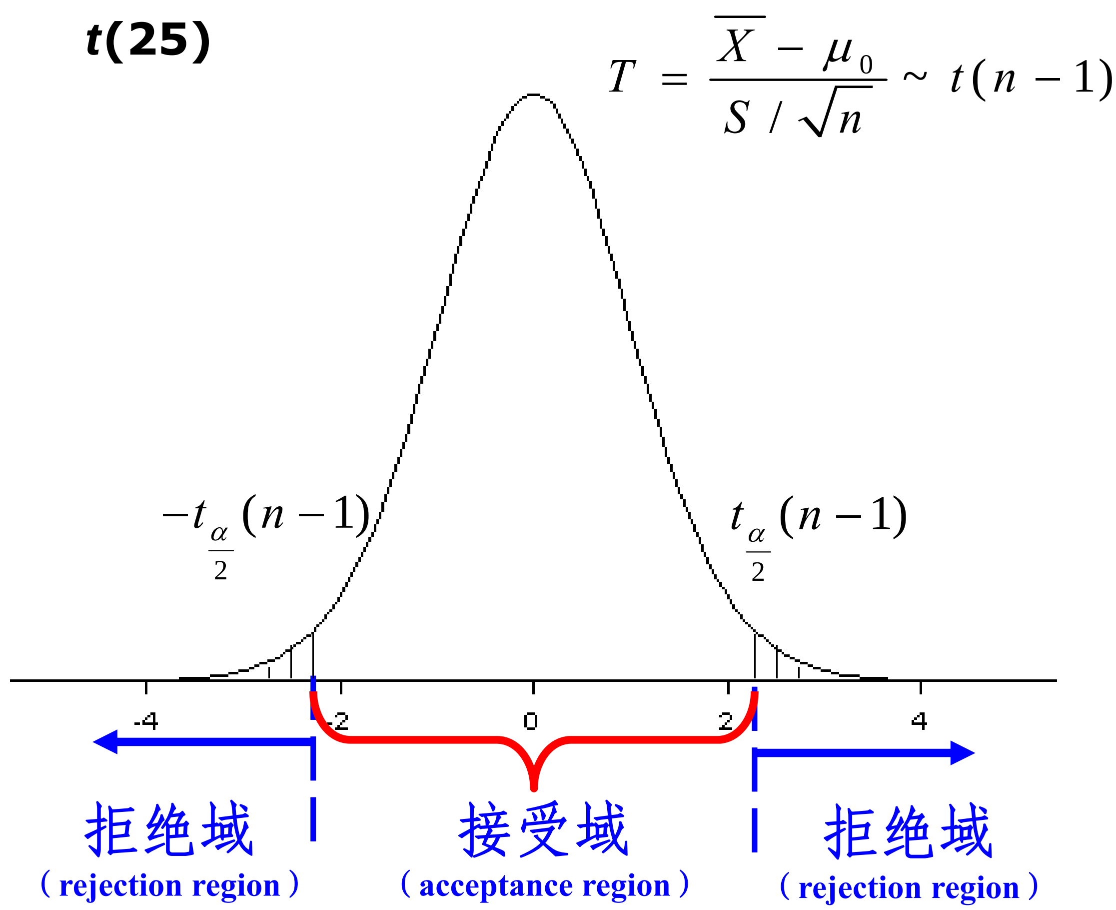
设总体$X \sim N(\mu, \sigma^2)$，σ^2未知，$X_1, X_2, ..., X_n$是来自总体X的样本。

选取的检验统计量为（即**T检验**）：

$$T=\frac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$$

显著性水平α下的拒绝域形式和接受域形式:

$$\left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right| \geq t_{\frac{\alpha}{2}}(n-1) \qquad \left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right| < t_{\frac{\alpha}{2}}(n-1)$$



二、均值μ的单边检验

1. 已知方差σ^2，检验H_0: $\mu=\mu_0$，H_1: $\mu<\mu_0$

设总体$X\sim N(\mu,\sigma^2)$，σ^2已知，$X_1, X_2, ..., X_n$是来自总体X的样本。

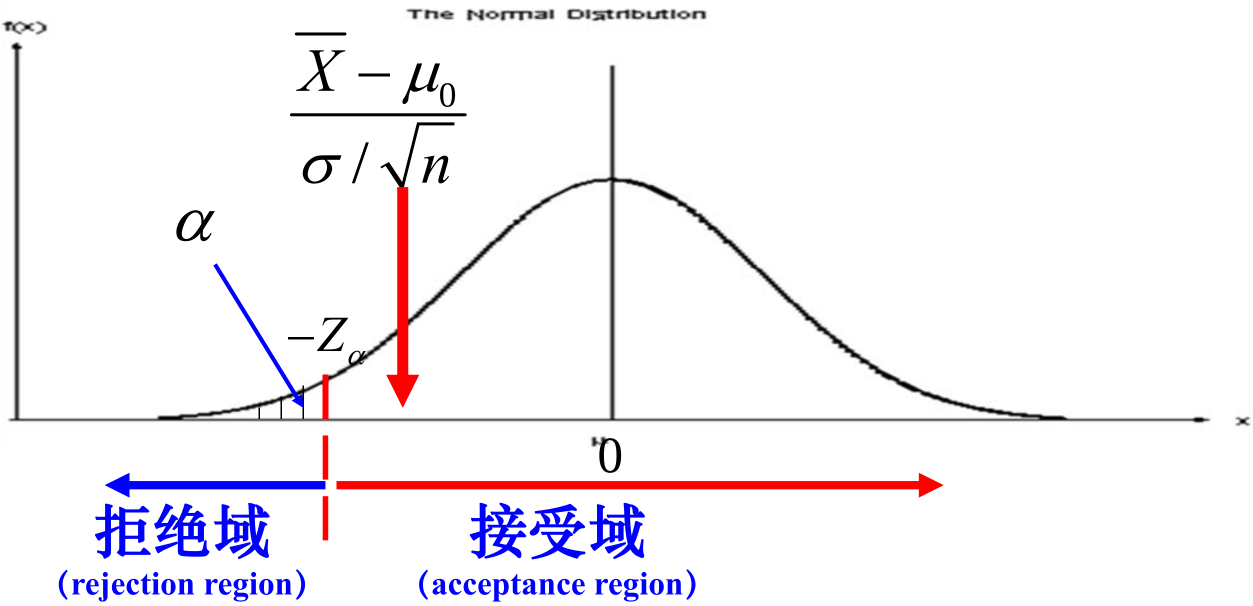
采用U检验，检验统计量为：

$$U=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$$

若拒绝原假设H_0: $\mu=\mu_0$，意味着接受对立假设H_1: $\mu<\mu_0$，只有当样本均值的观察值比μ_0小很多时，才有理由拒绝原假设。

给定显著性水平α的拒绝域和接受域形式：

$$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}<-Z_\alpha \qquad\qquad \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\geq -Z_\alpha$$



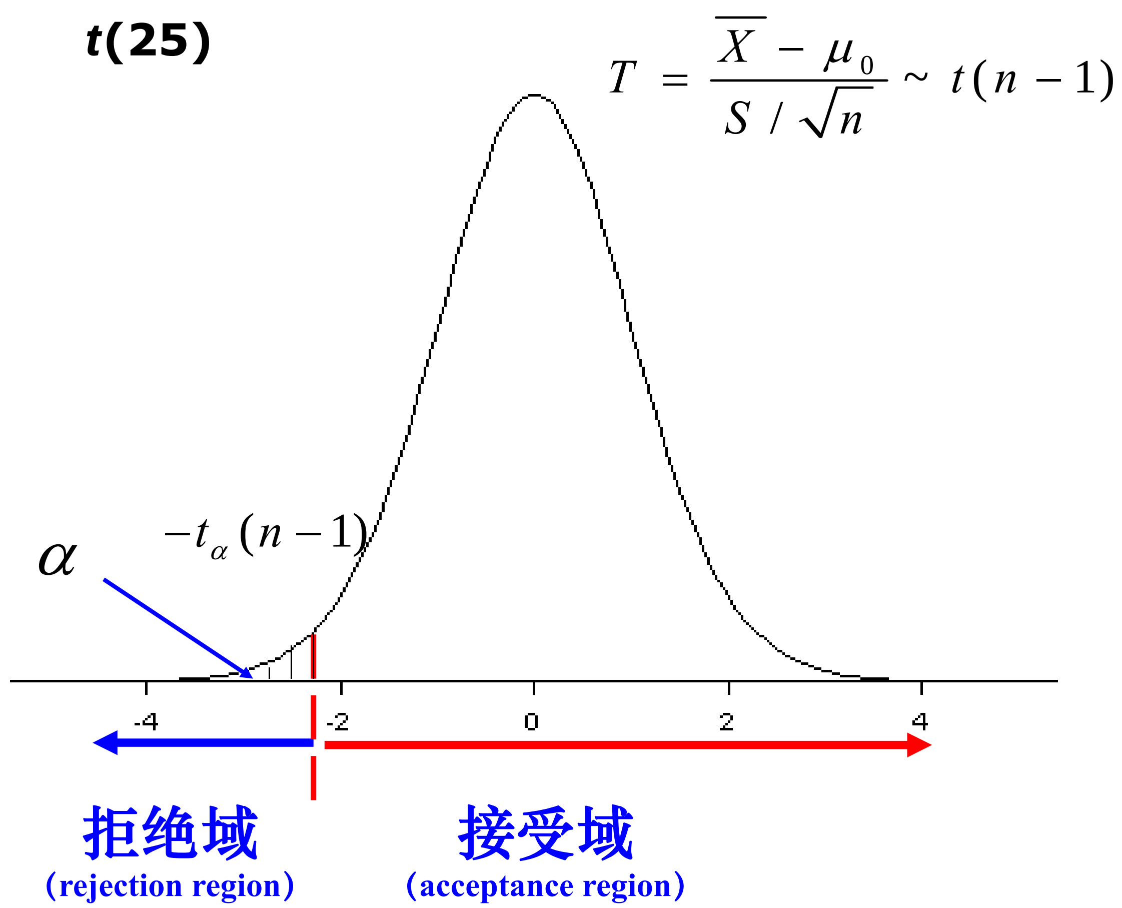
2. 未知方差σ^2，检验$H_0: \mu = \mu_0$，$H_1: \mu < \mu_0$

设总体$X \sim N(\mu, \sigma^2)$，σ^2未知，$X_1, X_2, ..., X_n$是来自总体X的样本。

采用**T检验**，给定显著性水平α的拒绝域和接受域形式：

$$\frac{\overline{X} - \mu_0}{S / \sqrt{n}} < -t_\alpha(n-1)$$

$$\frac{\overline{X} - \mu_0}{S / \sqrt{n}} \geq -t_\alpha(n-1)$$



3. 未知方差σ^2，检验H_0: $\mu\le\mu_0$，H_1: $\mu > \mu_0$

设总体$X\sim N(\mu, \sigma^2)$，σ^2未知，$X_1, X_2, ..., X_n$是来自总体X的样本。

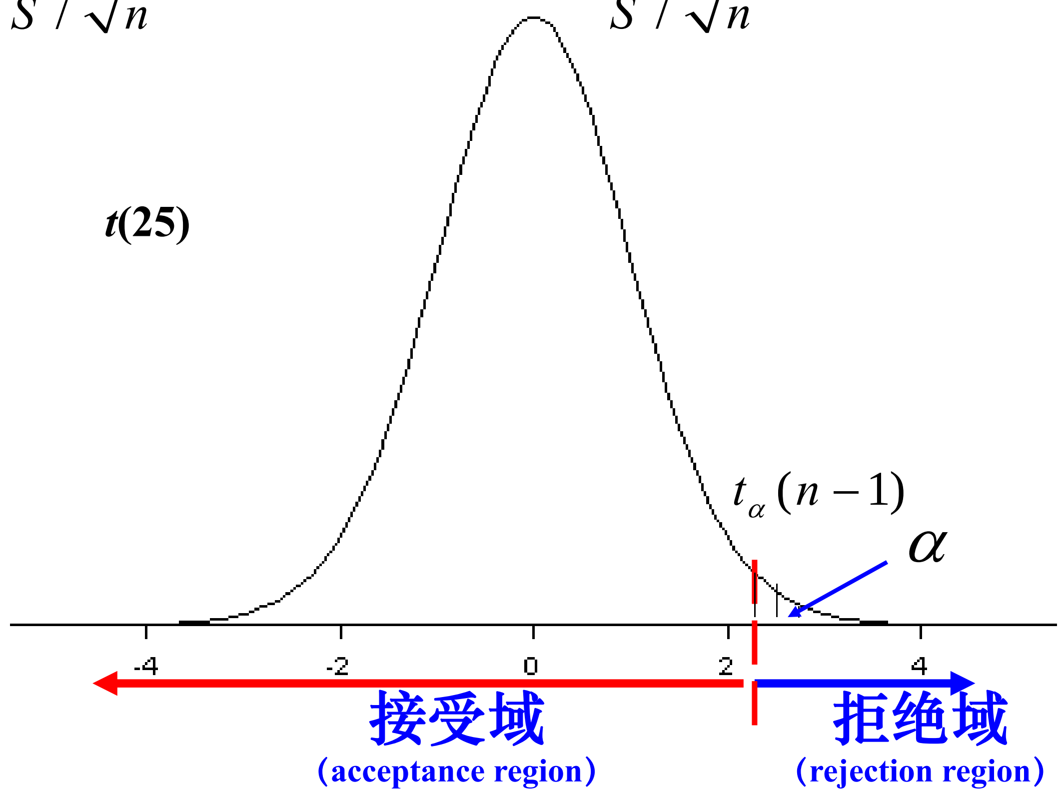
采用T检验，将原假设的复合假设分解为两种更简单的情况：

（1）若$\mu=\mu_0$成立

（2）若$\mu<\mu_0$成立

给定显著性水平α的拒绝域和接受域形式：

$$\frac{\overline{X}-\mu_0}{S/\sqrt{n}} > t_\alpha(n-1) \qquad \frac{\overline{X}-\mu_0}{S/\sqrt{n}} \le t_\alpha(n-1)$$



讨论:

（1）假设检验的关键:
——对立假设是否在小概率条件下成立
——对立假设是否显著地满足

（2）显著地满足对立假设的区域 ⇨ 拒绝域
显著的程度：显著性水平α

（3）对立假设 ⇨ 拒绝域 ⇨ 接受域 ⇨ 原假设

§ 6.2.2 两正态总体均值差的检验

设总体$X \sim N(\mu_1, \sigma_1^2)$和$Y \sim N(\mu_2, \sigma_2^2)$相互独立，$X_1, X_2, ..., X_{n1}$是来自总体X的样本，$Y_1, Y_2, ..., Y_{n2}$是来自总体Y的样本。

1. 已知方差σ_1^2和σ_2^2，检验
$H_0: \mu_1 = \mu_2$，$H_1: \mu_1 \neq \mu_2$

$$E\left(\overline{X}-\overline{Y}\right)=\mu_1-\mu_2$$

$$\left(\overline{X}-\overline{Y}\right) \sim N\left(\mu_1-\mu_2, \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$$

$$U=\frac{\left(\overline{X}-\overline{Y}\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

采用**U检验**，选取U为检验统计量，即：

$$U=\frac{\left(\overline{X}-\overline{Y}\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

给定显著性水平α的拒绝域和接受域形式为：

$$\frac{\left|\overline{X}-\overline{Y}\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}>Z_{\frac{\alpha}{2}} \qquad \frac{\left|\overline{X}-\overline{Y}\right|}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\le Z_{\frac{\alpha}{2}}$$

2. 未知方差σ_1^2和σ_2^2，但已知$\sigma_1^2=\sigma_2^2$，检验H_0: $\mu_1=\mu_2$，H_1: $\mu_1\neq\mu_2$

$$E\left(\overline{X}-\overline{Y}\right)=\mu_1-\mu_2$$

$$\Rightarrow \quad T=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$$

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

其中：

$$S_1^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2,$$

$$S_2^2=\frac{1}{n_2-1}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2$$

采用**T检验**，选取T为检验统计量，即：

$$T=\frac{(\overline{X}-\overline{Y})}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$$

给定显著性水平α的拒绝域和接受域形式为：

$$\frac{\left|\overline{X}-\overline{Y}\right|}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}>t_{\frac{\alpha}{2}}(n_1+n_2-2)$$

$$\frac{\left|\overline{X}-\overline{Y}\right|}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\le t_{\frac{\alpha}{2}}(n_1+n_2-2)$$

3. 已知方差σ_1^2和σ_2^2，检验H_0: $\mu_1=\mu_2$，H_1: $\mu_1>\mu_2$

采用**U检验**，选取U为检验统计量，即：

$$U=\frac{\left(\overline{X}-\overline{Y}\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

给定显著性水平α的拒绝域和接受域形式为：

$$\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}>Z_\alpha \qquad \frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\le Z_\alpha$$

§ 6.2.3 单个正态总体方差σ^2的检验

一、方差σ^2的的双边检验

设总体$X \sim N(\mu, \sigma^2)$，$X_1, X_2, ..., X_n$是来自总体X的样本。

1. 未知均值μ，检验$H_0: \sigma^2 = \sigma_0^2$，$H_1: \sigma^2 \neq \sigma_0^2$

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

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

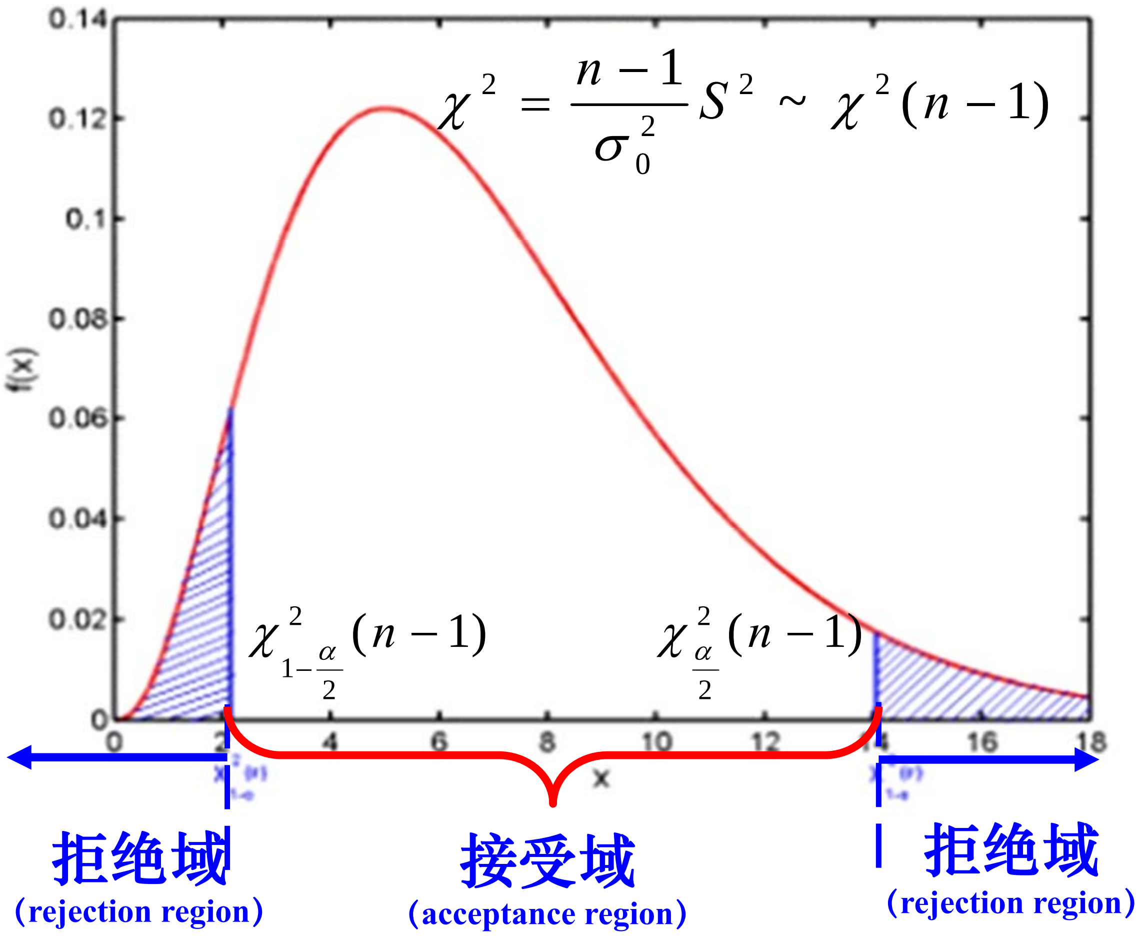
选取检验统计量（称为χ^2检验）：

$$\chi^2 = \frac{n-1}{\sigma_0^2}S^2 \sim \chi^2(n-1)$$

给定显著性水平α的拒绝域和接受域形式为:

$$\left(\frac{n-1}{\sigma_0^2}S^2 < \chi^2_{1-\frac{\alpha}{2}}(n-1)\right) \cup \left(\frac{n-1}{\sigma_0^2}S^2 > \chi^2_{\frac{\alpha}{2}}(n-1)\right)$$

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) \leq \frac{n-1}{\sigma_0^2}S^2 \leq \chi^2_{\frac{\alpha}{2}}(n-1)$$



2. 已知均值μ，检验H_0: $\sigma^2=\sigma_0^2$，H_1: $\sigma^2\neq\sigma_0^2$

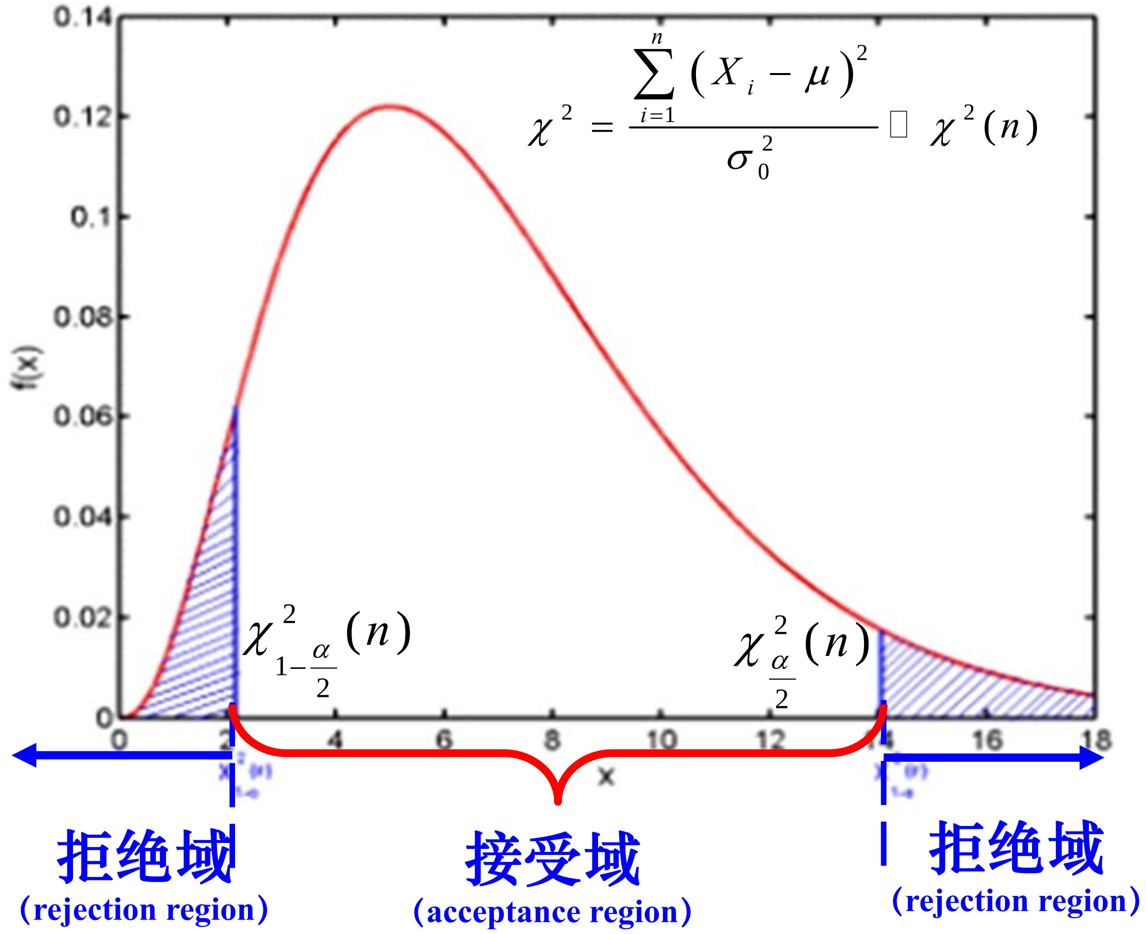
采用χ^2检验，选取检验统计量为:

$$\chi^2=\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n)$$

给定显著性水平α的拒绝域和接受域形式为:

$$\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}<\chi^2_{1-\frac{\alpha}{2}}(n)\right)\cup\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}>\chi^2_{\frac{\alpha}{2}}(n)\right)$$

$$\chi^2_{1-\frac{\alpha}{2}}(n)\le\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\le\chi^2_{\frac{\alpha}{2}}(n)$$



二、方差σ^2的的单边检验

设总体$X \sim N(\mu, \sigma^2)$，$X_1, X_2, \ldots, X_n$是来自总体X的样本。

1. 未知均值μ，检验H_0: $\sigma^2 \le \sigma_0^2$，H_1: $\sigma^2 > \sigma_0^2$

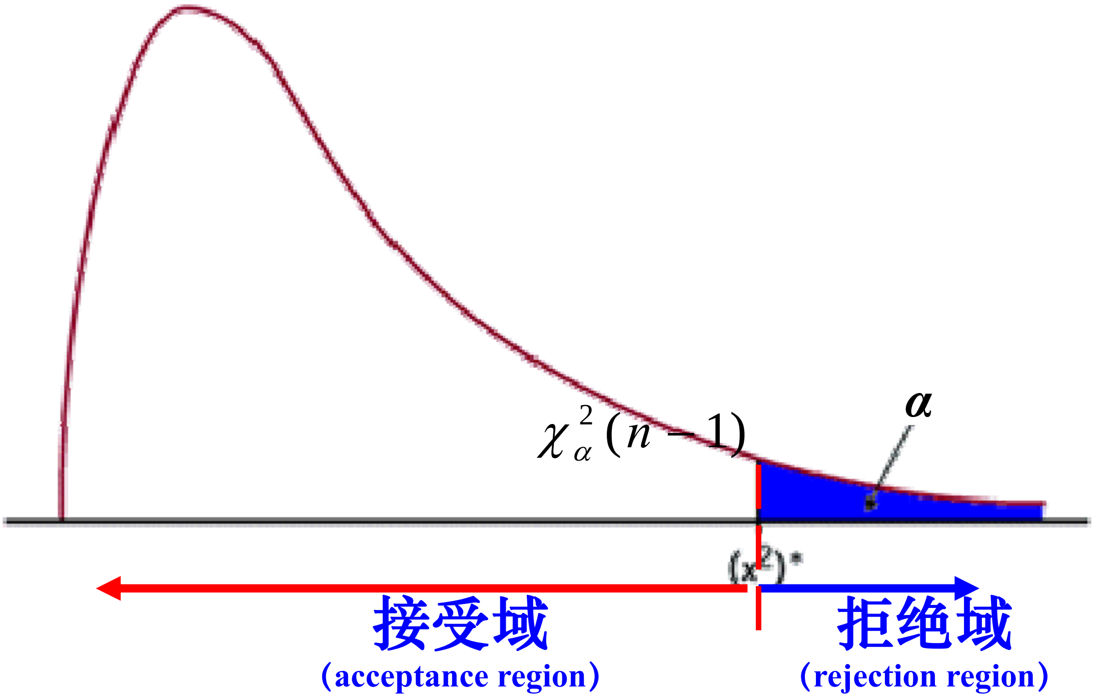
将原假设分解为更为简单的两种情况。采用χ^2检验，选取检验统计量为：

$$\frac{n-1}{\sigma_0^2}S^2 \sim \chi^2(n-1)$$

给定显著性水平α的拒绝域和接受域形式为：

$$\frac{n-1}{\sigma_0^2}S^2 > \chi^2_\alpha(n-1) \qquad \frac{n-1}{\sigma_0^2}S^2 \le \chi^2_\alpha(n-1)$$

$$\frac{n-1}{\sigma_0^2}S^2 \sim \chi^2(n-1)$$



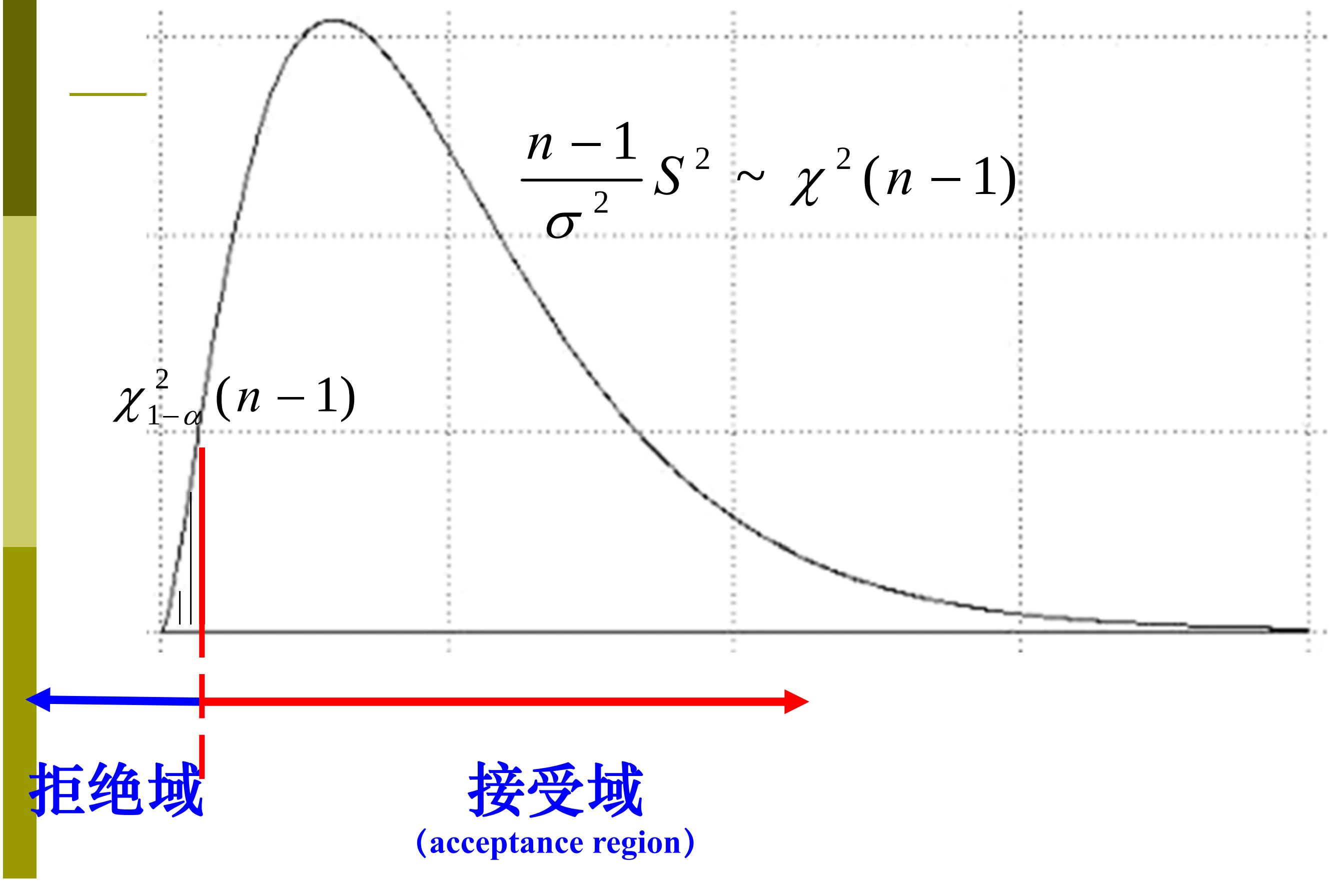
2. 未知均值μ，检验H_0: $\sigma^2 \geq \sigma_0^2$，H_1: $\sigma^2 < \sigma_0^2$

采用χ^2检验，选取检验统计量为：

$$\frac{n-1}{\sigma_0^2}S^2 \sim \chi^2(n-1)$$

给定显著性水平α的拒绝域和接受域形式为：

$$\frac{n-1}{\sigma_0^2}S^2 < \chi_{1-\alpha}^2(n-1) \qquad \frac{n-1}{\sigma_0^2}S^2 \geq \chi_{1-\alpha}^2(n-1)$$



§ 6.2.4 两正态总体方差比较的检验

设总体X~N(μ_1, σ_1^2)和Y~N(μ_2, σ_2^2)相互独立，$X_1, X_2, ..., X_{n1}$是来自总体X的样本，$Y_1, Y_2, ..., Y_{n2}$是来自总体Y的样本。

1. 未知均值μ_1、μ_2，检验
H_0: $\sigma_1^2 = \sigma_2^2$，H_1: $\sigma_1^2 \neq \sigma_2^2$

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$$

$$S_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2, \quad S_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2$$

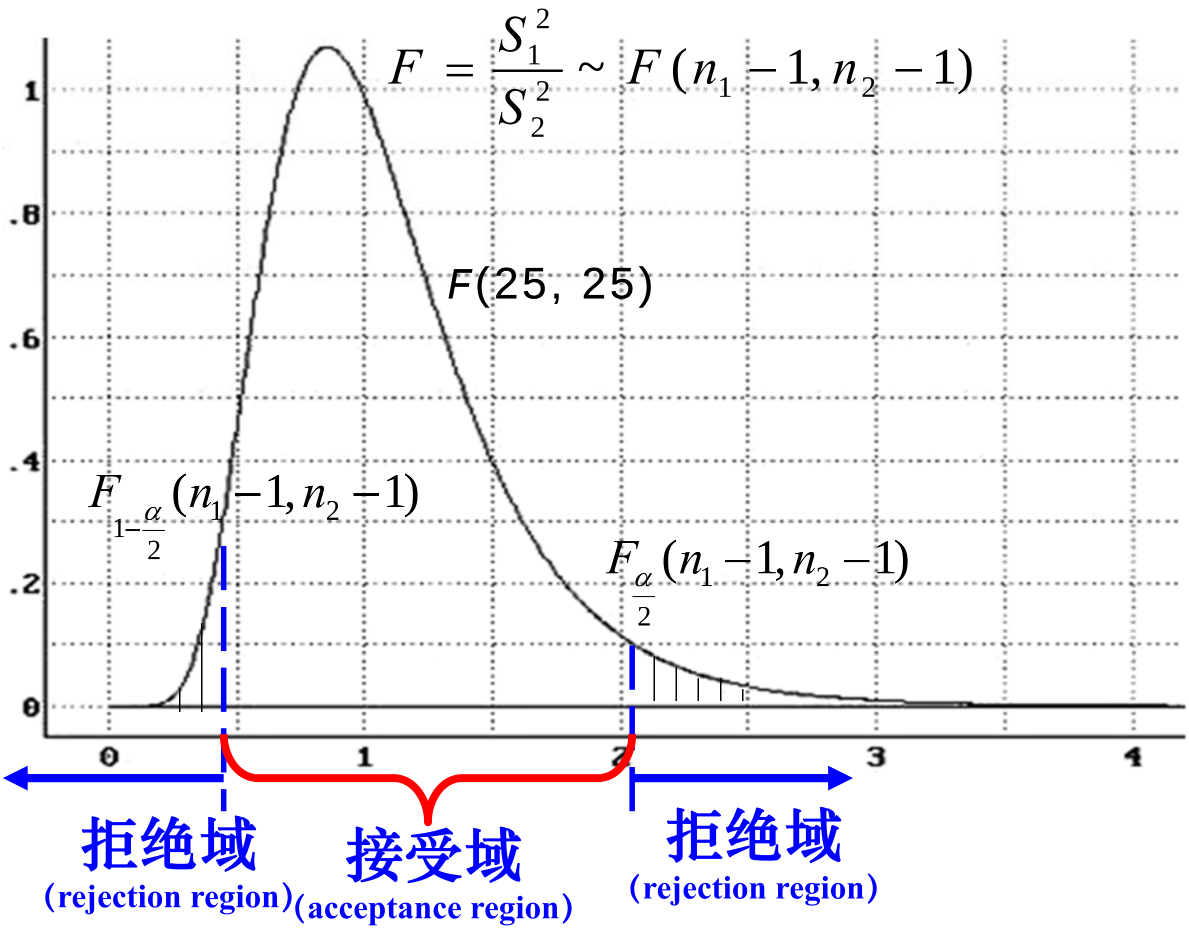
采用**F检验**，选取检验统计量为：

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

给定显著性水平α的拒绝域和接受域形式为：

$$\left\{\frac{S_1^2}{S_2^2}<F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}\cup\left\{\frac{S_1^2}{S_2^2}>F_{\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}$$

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



2 未知均值μ_1、μ_2，检验
$H_0: \sigma_1^2 = \sigma_2^2$，$H_1: \sigma_1^2 > \sigma_2^2$

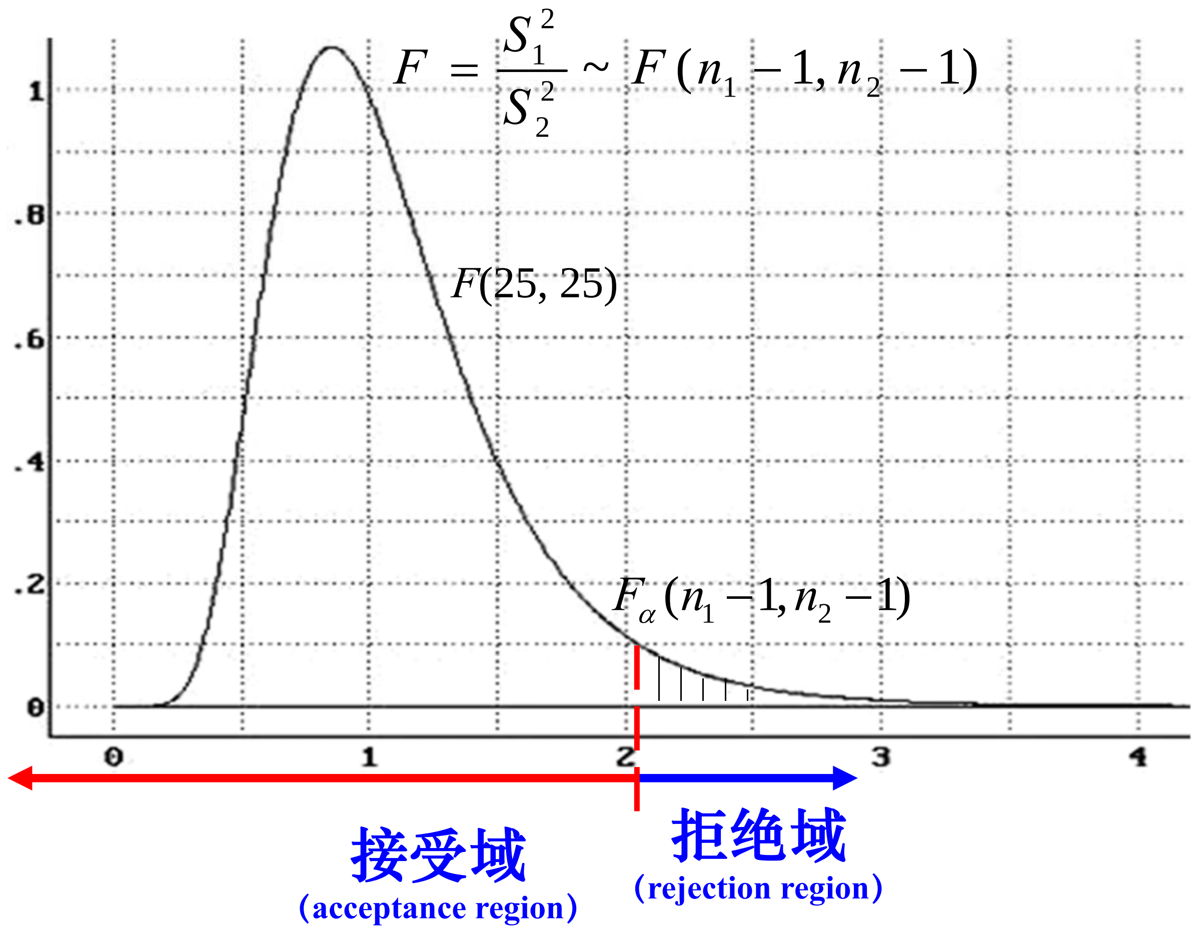
采用F检验，选取检验统计量为：

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

给定显著性水平α的拒绝域和接受域形式为：

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

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



§6.3 似然比假设检验

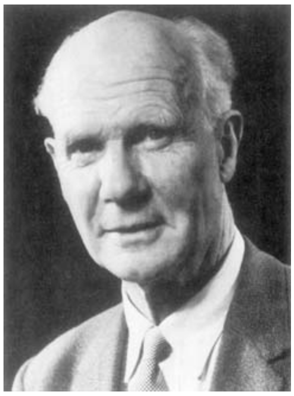
Neyman-Pearson（1928）

——更一般条件下的、统一的似然比（Likelihood ratio）检验方法

——检验统计量在一般的条件下有统一的渐近分布

——包括非正态的总体分布

Jerzy Neyman
（1894-1981）
美国波兰裔统计学家



Egon Pearson
（1895-1980）
英国统计学家

§6.3.1 N-P引理：似然比检验方法

N-P引理（Neyman-Pearson Lemma）

设原假设H_0和对立假设H_1均为简单假设，它们分别给出随机变量总体X的概率分布密度函数（或分布律）$f_0(x|\theta_0)$和$f_1(x|\theta_1)$。对于$X_1, X_2, ..., X_n$是来自总体X的（简单随机抽样）样本，定义似然比（LR, Likelihood Ratio）为检验统计量：

$$LR = \frac{lik(f_0)}{lik(f_1)} = \frac{\prod_{i=1}^{n} f_0(X_i|\theta_0)}{\prod_{i=1}^{n} f_1(X_i|\theta_1)}$$

（1）似然比检验（LR test）：在给定显著性水平α条件下，当$LR < c$满足（c为临界值），则可拒绝原假设H_0。

（2）对于上述似然比检验，任给显著性水平$\alpha^* \leq \alpha$的假设检验，均有：

$$(1-\beta^*) \leq (1-\beta)$$

其中，β^*和β均为第二类错误的概率。

N-P引理讨论：

（1）当似然比取值越大，表明随机变量的分布越趋近原假设，反之，越小越趋近对立假设。

（2）在给定显著性水平α的条件下，似然比检验发生的第二类错误概率β最小，即似然比检验方法是最大功效的检验方法。

（3）适用条件：原假设和对立假设均为简单假设。

【Example 6.4】设总体$X\sim N(\mu, \sigma^2)$，σ^2已知，而μ只能取两个值μ_0、μ_1，且$\mu_0<\mu_1$，$X_1, X_2, \ldots, X_n$为来自总体X的容量为n的简单随机抽样样本，试在显著性水平α下运用似然比检验假设：$H_0: \mu=\mu_0$；$H_1: \mu = \mu_1$。

基本步骤：

——写出LR的表达式

——根据LR取值趋小的形式找到拒绝原假设的统计量形式

——根据抽样分布求出临界点的值

§6.3.2 广义似然比检验方法

——广义似然比检验（GLR test, Generalized Likelihood Ratio test）

已知$X_1, X_2, \ldots, X_n$是来自总体X的（简单随机抽样）样本，根据总体分布函数可得到它们的联合概率密度形式为：

$$\prod_{i=1}^{n} f\left(X_i \mid \theta\right), \quad \theta \in \Theta$$

Θ为参数空间，其维数可能大于1。

设原假设H_0定义的参数范围为：$\theta\in\omega_0\subset\Theta$，对立假设$H_1$定义的参数范围为：$\theta\in\omega_1\subset\Theta$，且$\omega_0\cap\omega_1=\varnothing$，记$\Theta'=\omega_0\cup\omega_1$，显然有 $\Theta' \subseteq \Theta$。

定义**广义似然比**（GLR , Generalized Likelihood Ratio）为：

$$GLR^* = \frac{\max\limits_{\theta\in\omega_0}[lik(\theta)]}{\max\limits_{\theta\in\omega_1}[lik(\theta)]}, \qquad lik(\theta) = \prod_{i=1}^{n} f\left(X_i \mid \theta\right)$$

以GLR*为检验统计量，显然有GLR*越小，原假设越有可能被拒绝。

通常更多地采用下列广义似然比的定义：

$$GLR = \frac{\max\limits_{\theta\in\omega_0}[lik(\theta)]}{\max\limits_{\theta\in\Theta'}[lik(\theta)]}$$

讨论：

（1）当原假设H_0为简单假设时，有：$GLR = \dfrac{lik(\theta_0)}{\max\limits_{\theta\in\Theta'}[lik(\theta)]}$

（2）$GLR = \min(GLR^*, 1)$

【Example 6.5】设总体X~N(μ, σ²)，σ²已知，$X_1, X_2, \ldots, X_n$为来自总体X的容量为n的简单随机抽样样本，试在显著性水平α下运用广义似然比检验假设：H_0: μ=μ_0；H_1: μ≠μ_0。

$$GLR = e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(X_i-\mu_0)^2-\sum_{i=1}^{n}(X_i-\overline{X})^2\right]}$$

$$GLR' = -2\log GLR = \frac{n\left(\overline{X}-\mu_0\right)^2}{\sigma^2} \sim \chi^2(1)$$

定理

在概率密度函数或分布律函数充分光滑的条件下，当样本容量n趋近无穷大时，推广到非正态分布总体的检验统计量$GLR' = (-2\log GLR)$ 的抽样分布趋近$\chi^2(k)$分布，其自由度k为：

$$k = \dim\left(\Theta'\right) - \dim\left(\omega_0\right)$$

§ 6.3.3 多项分布的广义似然比检验

多项分布拟合检验问题：设$X_1, X_2, \ldots, X_m$为分布在m个格子中、满足多项分布的采样计数（即样本）且 $X_1+X_2+\ldots+X_m=n$ 。令落入每个格子的概率分别为$p_1, p_2, \ldots, p_m$，$p_1+p_2+\ldots+p_m=1$ 。

原假设H_0：给出多项分布概率密度矢量$p=p(\theta)\in\omega_0$，其中θ为本问题的参数或待定未知参数。

对立假设H_1：所有满足 $p_1+p_2+\ldots+p_m=1$ 的概率密度矢量$p\in\Theta'$。

$$GLR=\frac{\max\limits_{\theta\in\omega_0}[lik(\theta)]}{\max\limits_{\theta\in\Theta'}[lik(\theta)]}$$

分子:

$$\max_{p\in\omega_0}\left\{\frac{n!}{\prod_{i=1}^{m}X_i!}\prod_{i=1}^{m}p_i(\theta)^{X_i}\right\}\xrightarrow{\text{MLE}}\frac{n!}{\prod_{i=1}^{m}X_i!}\prod_{i=1}^{m}p_i(\hat{\theta})^{X_i}$$

分母（MLE）：

$$\frac{n!}{\prod_{i=1}^{m}X_i!}\prod_{i=1}^{m}\hat{p}_i^{X_i},\qquad \hat{p}_i=\frac{X_i}{n}$$

$$GLR = \frac{\max\limits_{\theta \in \omega_0} [lik(\theta)]}{\max\limits_{\theta \in \Theta'} [lik(\theta)]}$$

χ^2分布

$$-2\log GLR = -2n\sum_{i=1}^{m} \hat{p}_i \log\left(\frac{p_i(\hat{\theta})}{\hat{p}_i}\right) = 2\sum_{i=1}^{m} O_i \log\left(\frac{O_i}{E_i}\right)$$

□ $$\chi^2(m-k-1)$$

记： $O_i = n\hat{p}_i, \quad E_i = np_i(\hat{\theta})$

$$\dim(\Theta') - \dim(\omega_0) = (m-1) - k$$

定理

对于同一多项分布的拟合假设检验问题，在原假设H_0成立的条件下，当样本容量n趋近无穷大时，GLR检验统计量

$$-2\log GLR = -2n\sum_{i=1}^{m} \hat{p}_i \log\left(\frac{p_i(\hat{\theta})}{\hat{p}_i}\right) = 2\sum_{i=1}^{m} O_i \log\left(\frac{O_i}{E_i}\right)$$

与Pearson χ^2检验统计量

$$X^2 = \sum_{i=1}^{m} \frac{\left[X_i - np_i(\hat{\theta})\right]^2}{np_i(\hat{\theta})} = \sum_{i=1}^{m} \frac{[O_i - E_i]^2}{E_i}$$

近似等价，且二者的抽样分布趋近自由度为 $\dim(\Theta') - \dim(\omega_0)$ **的χ^2分布。**

§6.4 非参数假设检验简介

参数假设检验问题（【例6.1】）

——已知总体的分布函数类型，对分布函数中的未知参数提出某种假设
——要求利用样本的信息对所提出的假设进行检验，根据检验结果作出接受或者拒绝所提假设的判断

非参数假设检验问题（【例6.2】）

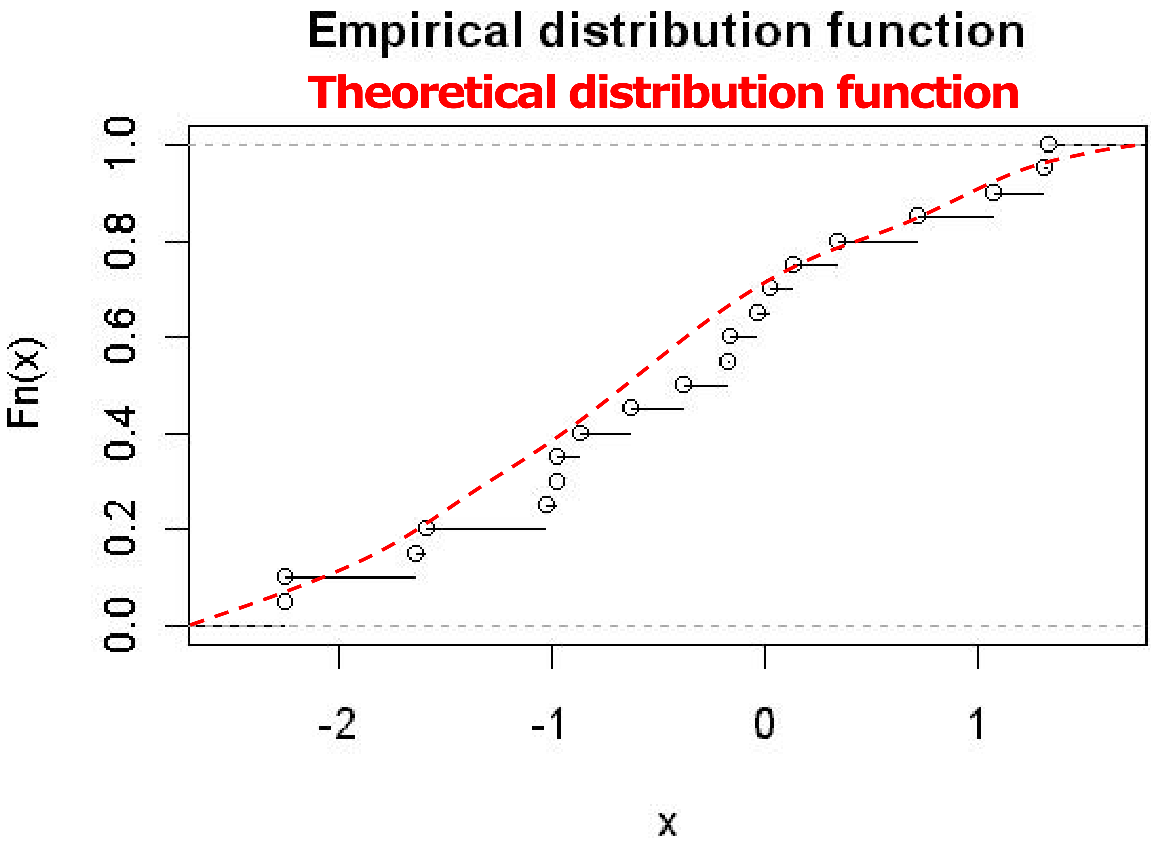
——总体的分布类型未知
——要求根据样本的信息来对分布类型的假设进行检验，对总体的分布类型作出判断

§6.4.1 K-S检验（Kolmogorov–Smirnov test）

定义 设$X_1, X_2, \ldots, X_n$为来自总体X的样本，$x_1, x_2, \ldots, x_n$为其观察值，将它们按照观察值的大小递增顺序排列生成顺序统计量$x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$，并作函数：

$$F_n(x) = \begin{cases} 0, & x < x_{(1)} \\ \dfrac{k}{n}, & x_{(k)} \le x < x_{(k+1)}, (k = 1, 2, \ldots, n-1) \\ 1, & x \ge x_{(n)} \end{cases}$$

称$F_n(x)$为总体X的经验分布函数（empirical distribution function）。对应地，总体X的分布函数F(x)称为理论分布函数（theoretical distribution function）。



讨论：

（1）对于给定的样本值顺序统计量$x_1, x_2, \ldots, x_n$，经验分布函数$F_n(x)$ 符合分布函数的三个条件

（2）由于样本的随机性，经验分布函数$F_n(x)$也是随机变量。且n $F_n(x)$服从二项分布$B(n, F(x))$

定理

设总体X的理论分布函数为F(x)，经验分布函数为$F_n(x)$，则当n→∞时，有：

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

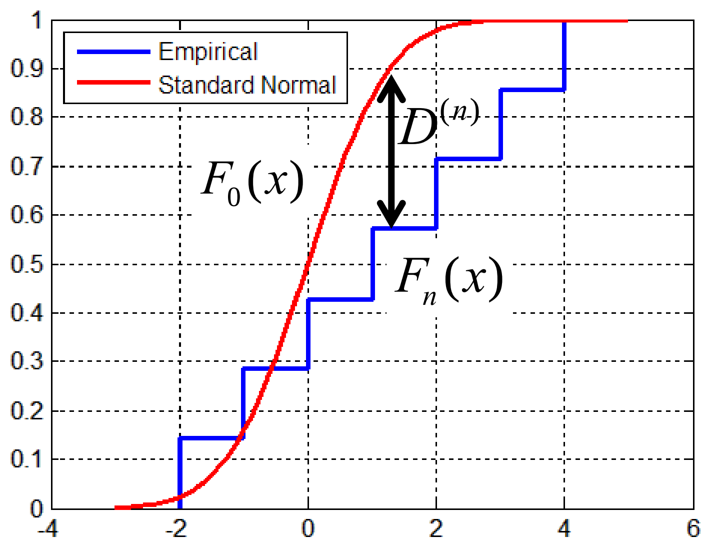
——当样本容量n很大时，可以用经验分布函数去估计总体X的理论分布函数。

K-S检验 （K检验）

设总体X的分布函数为F(x)，对一个给定的分布$F_0(x)$，考虑原假设H_0: F(x) = $F_0(x)$。当样本容量n充分大时，样本的经验分布函数$F_n(x)$是总体X分布F(x)的很好的近似，因此当原假设H_0为真时，$F_n(x)$与$F_0(x)$之间的差应该是一个小量。

Kolmogorov统计量:

$$D^{(n)}=\max_{-\infty<x<+\infty}\left|F_n(x)-F_0(x)\right|$$



定理

设总体X的分布函数F(x)连续，$X_1, X_2, \ldots, X_n$为来自总体X的样本，则当原假设H_0: F(x) = F_0(x)成立时，有：

$$\lim_{n\to\infty} P\left(D^{(n)} < \frac{\lambda}{\sqrt{n}}\right) = K(\lambda)$$

其中，

$$K(\lambda)=\begin{cases} 0, & \lambda \le 0 \\ \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2}, & \lambda > 0 \end{cases}$$

满足这一分布函数的分布称为**K分布**。

Kolmogorov统计量的良好性质：“分布无关性”

K-S检验（Smirnov，1939）

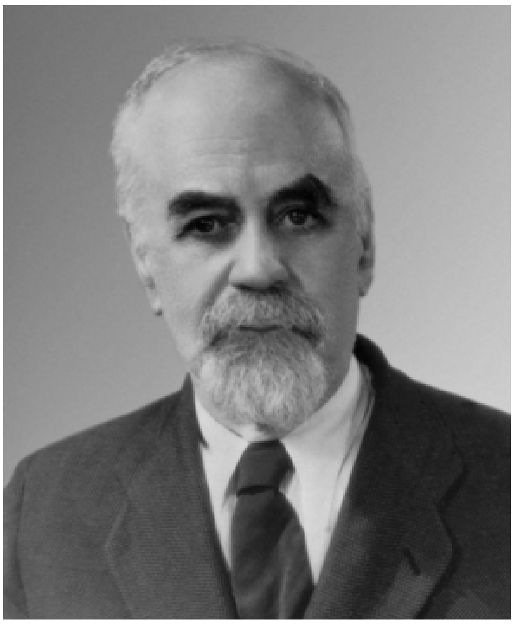
设总体X的分布为F(x)，考虑原假设H_0: F(x) = F_0(x)，在显著性水平α下的拒绝域形式为：

$$P\left(D^{(n)} \ge D^{(n)}_{1-\alpha}\right) = \alpha$$

$$D^{(n)}_{0.90} \approx \frac{1.23}{\sqrt{n}}$$

$$D^{(n)}_{0.95} \approx \frac{1.36}{\sqrt{n}}$$

$$D^{(n)}_{0.99} \approx \frac{1.63}{\sqrt{n}}$$



Vladimir Smirnov
(1887-1974)

Russian mathematician

K-S检验的基本步骤

（1）将样本观察值$x_1, x_2, \ldots, x_n$生成顺序统计量$x_{(1)} \le x_{(2)} \le \ldots, \le x_{(n)}$，并求出经验分布函数：

$$F_n(x) = \begin{cases} 0, & x < x_{(1)} \\ \dfrac{k}{n}, & x_{(k)} \le x < x_{(k+1)}, (k = 1, 2, \ldots, n-1) \\ 1, & x \ge x_{(n)} \end{cases}$$

（2）计算Kolmogorov统计量$D^{(n)}$：

$$D^{(n)} = \max_{-\infty < x < +\infty} \left| F_n(x) - F_0(x) \right| = \max_{1 \le i \le n} \left\| \frac{i}{n} - F_0(x_{(i)}) \right|, \left| \frac{i-1}{n} - F_0(x_{(i)}) \right\|$$

（3）若$D^{(n)} \ge D^{(n)}_{1-\alpha}$则拒绝原假设$H_0$。

$$D^{(n)}_{0.90} \approx \frac{1.23}{\sqrt{n}} \qquad D^{(n)}_{0.95} \approx \frac{1.36}{\sqrt{n}} \qquad D^{(n)}_{0.99} \approx \frac{1.63}{\sqrt{n}}$$

讨论

（1）K-S检验适合对单个总体的连续型分布进行假设检验；

（2）具有较高的灵敏度；

（3）缺点：无法检验离散型总体的分布。

§ 6.4.2 χ^2检验（Chi-square test）

——K. Pearson（1900）；

设总体X分布F(x)未知，已知$X_1, X_2, \ldots, X_n$为来自总体X的样本。

$$H_0: F(x) = F_0(x);\quad H_1: F(x) \neq F_0(x)$$

将总体X的样本空间Ω划分为m个互不相容的子空间$A_1, A_2, \ldots, A_m$，使

$$\bigcup_{i=1}^{m} A_i = \Omega \qquad A_i \cap A_j = \varnothing, i \neq j$$

于是当原假设H_0成立时，可计算离散型分布律：

$$p_i = P\left(X \in A_i\right), i = 1, 2, \ldots, m$$

在n次试验中实际频数为n_i，则n_i/n为事件$\{X \in A_i\}$的频率。

构造Pearson χ^2检验统计量

$$X^2 = \sum_{i=1}^{m} \frac{\left[n_i - np_i\right]^2}{np_i} = \sum_{i=1}^{m} \frac{\left[O_i - E_i\right]^2}{E_i}$$

Pearson χ^2检验统计量的渐近分布为$\chi^2(m-r-1)$，其中r是分布函数参数的维数。

讨论

（1）χ^2检验统计量的定义与样本空间Ω的划分有关。只有合适的划分才能使得构造的离散型分布律能合理地近似总体分布；

（2）样本容量n应该足够大，一般$n \geq 50$；

（3）一般要求$np_i>5$，否则应适当合并A_i以满足要求。

（4）当总体理论分布$F_0(x)$是连续型时，K-S检验要比Pearson χ^2检验效率高。

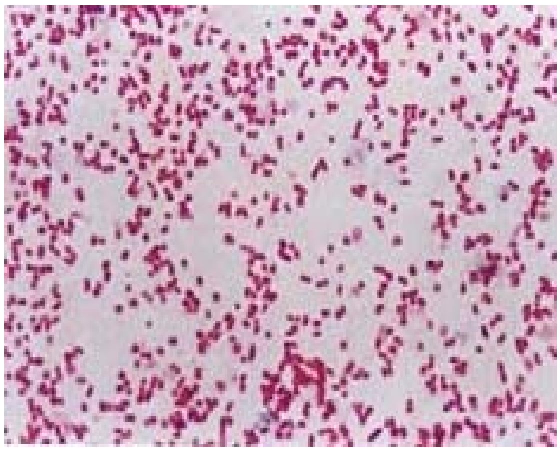
【Example 6.2】将0.1ml受细菌污染的牛奶均匀涂在$1cm^2$的切片上，用显微镜观察切片每个小网格内的细菌菌落数目。根据400（20 × 20）个小网格的计数结果，统计出如下表格:

菌落数:	0	1	2	3	4	5	6	7	8	9	10	19
频　数:	56	104	80	62	42	27	9	9	5	3	2	1

试问单位面积内的菌落数是否服从Poisson分布?

问题:

根据样本信息，判别随机变量X是否服从Poisson分布。



（Bliss C. and Fisher R. A., (1953). Fitting the negative binomial distribution to biological data. *Biometric*, 9: 174-200）

【Example 6.2】将0.1ml受细菌污染的牛奶均匀涂在1cm²的切片上，用显微镜观察切片每个小网格内的细菌菌落数目。根据400（20×20）个小网格的计数结果，统计出如下表格:

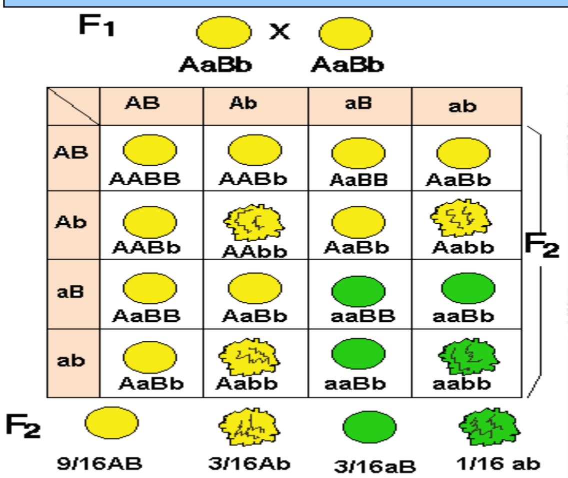
菌落数:	0	1	2	3	4	5	6	7	8	9	10	19
频　数:	56	104	80	62	42	27	9	9	5	3	2	1

试问菌落数是否服从Poisson分布？

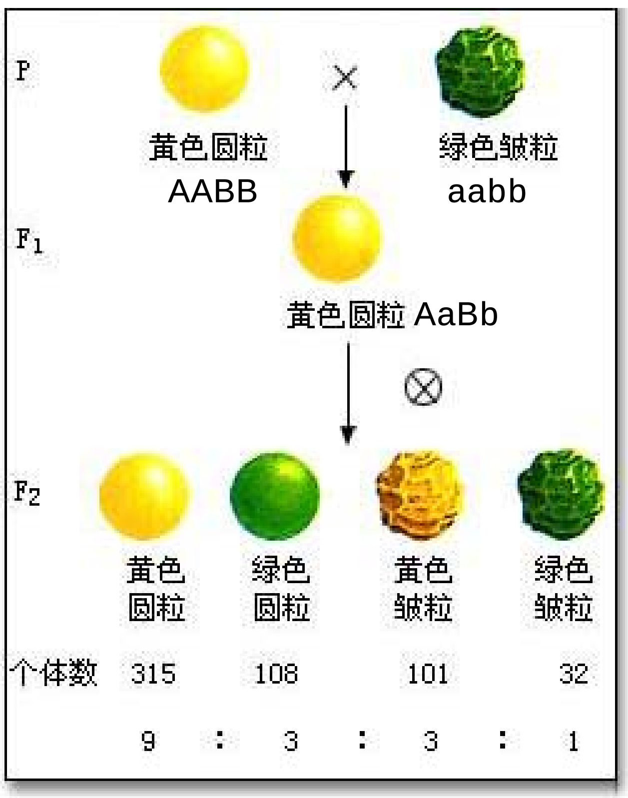
$$\tilde{\lambda} = \overline{X} = \frac{0\times 56 + 1\times 104 + \ldots + 19\times 1}{400} = 2.44$$

菌落数:	0	1	2	3	4	5	6	≥7
实际频数:	56	104	80	62	42	27	9	20
期望频数:	34.9	85.1	103.8	84.4	51.5	25.1	10.2	5.0
$\frac{[O_i - E_i]^2}{E_i}$:	12.8	4.2	5.5	6.0	1.8	0.14	0.14	45.0

【Example 6.6】Mendel遗传试验数据的假设检验



孟德尔曾做过杂交试验的园地



【Example 6.7】调查339名50岁以上有吸烟习惯者与慢性气管炎病的关系。试问吸烟习惯与慢性气管炎是否有关系？

即检验假设H_0：吸烟习惯与慢性气管炎相互独立（给定显著性水平$\alpha = 0.01$）。

	患病人数	未患病人数	合计	患病率
吸烟人数	43	162	205	21.0%
不吸烟人数	13	121	134	9.7%
合计	56	283	339	16.5%

§6.4.3 两总体之间的独立性检验

设有二元总体(X, Y)，x, y的取值范围分别划分成m个和k个互不相交的区间$A_1, A_2, \ldots, A_m$和$B_1, B_2, \ldots, B_k$。从总体中抽取容量为n的样本，样本值记为$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$，记$n_{ij}$为样本值的x落入$A_i$，y落入$B_j$的个数（i=1, 2, …, m; j=1, 2, .., k），又记

$$n_{i\bullet} = \sum_{j=1}^{k} n_{ij}, \quad n_{\bullet j} = \sum_{i=1}^{m} n_{ij}$$

——列联表

需要检验假设 H_0: X与Y相互独立。这种假设检验问题称为独立性检验。

列联表

		y				$n_{i\bullet}=\sum_{j=1}^{k} n_{ij}$
		B_1	B_2	…	B_k	
x	A_1	n_{11}	n_{12}	…	n_{1k}	$n_{1\cdot}$
	A_2	n_{21}	n_{22}	…	n_{2k}	$n_{2\cdot}$
	…	…	…	…	…	…
	A_m	n_{m1}	n_{m2}	…	n_{mk}	$n_{m\cdot}$
$n_{\bullet j}=\sum_{i=1}^{m} n_{ij}$		$n_{\cdot 1}$	$n_{\cdot 2}$	…	$n_{\cdot k}$	

$$p_{ij}=P\left(X\in A_i, Y\in B_j\right)$$

$$p_{i\cdot}=P\left(X\in A_i\right)$$

$$p_{\cdot j}=P\left(Y\in B_j\right)$$

$$p_{i\cdot}=\sum_{j=1}^{k} p_{ij},\quad p_{\cdot j}=\sum_{i=1}^{m} p_{ij},$$

$$\sum_{i=1}^{m} p_{i\cdot}=\sum_{j=1}^{k} p_{\cdot j}=\sum_{i=1}^{m}\sum_{j=1}^{k} p_{ij}=1$$

在原假设H_0成立的条件下，有：

$$p_{ij}=p_{i\cdot}\cdot p_{\cdot j},\quad i=1,2,\ldots,m;\ j=1,2,\ldots,k$$

假设检验就是要检验这一关系。

（1）当p_{ij}已知时

构造检验统计量：

$$\chi^2=\sum_{i=1}^{m}\sum_{j=1}^{k}\frac{\left(n_{ij}-np_{ij}\right)^2}{np_{ij}}=\sum_{i=1}^{m}\sum_{j=1}^{k}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}$$

其渐近分布为 $\chi^2(mk-1)$ 。

给定显著性水平α的拒绝域形式：

$$\chi^2=\sum_{i=1}^{m}\sum_{j=1}^{k}\frac{\left(n_{ij}-np_{ij}\right)^2}{np_{ij}}\geq\chi^2_{1-\alpha}\left(mk-1\right)$$

（2）当p_{ij}未知时

用p_{ij}的最大似然估计来代替。

$$\hat{p}_{i\cdot}=\frac{n_{i\cdot}}{n},\quad \hat{p}_{\cdot j}=\frac{n_{\cdot j}}{n}\quad\Longrightarrow\quad \hat{p}_{ij}=\hat{p}_{i\cdot}\cdot\hat{p}_{\cdot j}=\frac{n_{i\cdot}\cdot n_{\cdot j}}{n}=\frac{\hat{n}_{ij}}{n}$$

检验统计量：

$$\chi^2=\sum_{i=1}^{m}\sum_{j=1}^{k}\frac{\left(n_{ij}-n\hat{p}_{ij}\right)^2}{n\hat{p}_{ij}}=\sum_{i=1}^{m}\sum_{j=1}^{k}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}$$

服从χ^2分布，其自由度为 $\dim\left(\Theta'\right)-\dim\left(\omega_0\right)=(m-1)(k-1)$

给定显著性水平α的拒绝域形式为：

$$\chi^2=\sum_{i=1}^{m}\sum_{j=1}^{k}\frac{\left(n_{ij}-n\hat{p}_{ij}\right)^2}{n\hat{p}_{ij}}\geq\chi^2_{1-\alpha}\left((m-1)(k-1)\right)$$

【Example 6.7】调查339名50岁以上有吸烟习惯者与慢性气管炎病的关系。试问吸烟习惯与慢性气管炎是否有关系？

即检验假设H_0：吸烟习惯与慢性气管炎相互独立（给定显著性水平$\alpha = 0.01$）。

	患病人数	未患病人数	合计	患病率
吸烟人数	43	162	205	21.0%
不吸烟人数	13	121	134	9.7%
合计	56	283	339	16.5%