CNAS 技术报告

测量不确定度在符合性判定中的应用

中国合格评定国家认可委员会

发布日期：2019 年 4 月 1 日

前 言

符合性判定是根据测量结果判断合格评定对象的特定属性是否满足规定要求的活动，是延伸测量结果的服务，也是实验室及其他合格评定机构经常从事的活动。测量不确定度表征赋予了被测量量值的分散性，是测量结果的一部分，也是判定规则考虑的主要内容。ISO/IEC 17025:2017《检测和校准实验室能力的通用要求》明确要求实验室“当作出与规范或标准符合性声明时，实验室应考虑与所用判定规则相关的风险水平（如错误接受、错误拒绝以及统计假设），将所使用的判定规则制定成文件，并应用判定规则”。

本文件主要依据 ISO/IEC 指南 98-4:2012《测量不确定度-第 4 部分：测量不确定度在合格评定中的应用》制定，提出了在符合性判定中考虑测量不确定度及风险评估的方法，包括常见的判定规则、合格概率的计算、基于合格概率确定接受区间、消费者和生产商风险的计算方法等内容，为合格评定机构选择和制定判定规则提供了指导。

本文件由中国合格评定国家认可委员会提出并归口。

本文件主要起草单位：中国合格评定国家认可中心、中国测试技术研究院、北京理工大学。

本文件主要起草人：安平、王阳、林志国、张明霞、刘浩峰、华广胜、陈凌峰。

测量不确定度在符合性判定中的应用

1 目的和范围

本文件为以下两类符合性判定活动提供了指南和实施步骤：

1）判断合格评定对象是否符合规定限值要求；

2）合理设置接受限，使合格评定对象的合格率达到期望值。

本文件所指的合格评定对象应具有可测量的属性，且测量结果应满足以下条件：

1）可用单一的标量表示；

2）容许区间由一个或两个容许限值确定；

3）表述方式与 GUM 规定的原则一致，其值的信息可以通过概率密度函数、概率分布函数、两种函数的数值近似或带有包含区间和相应包含概率的被测量估计值等表述。

本文件为指导性文件，供合格评定机构参考使用。

2 引用文件

下列文件对于本文件的应用是必不可少的。凡是标注日期的引用文件，仅注日期的版本适用于本标准；凡是不注日期的引用文件，其最新版本（包括所有的修改）适用于本文件。

2.1 ISO/IEC 指南 98-4:2012《测量不确定度-第 4 部分：测量不确定度在合格评定中的应用》

2.2 GB/T 27418《测量不确定度评定与表示》（修改采用 GUM）

2.3 GB/T 27419《测量不确定度评定与表示 补充文件 1：基于蒙特卡洛方法的分布传播》

2.4 GB/T 27000《合格评定 词汇和通用原则》

2.5 ISO/IEC 指南 99《国际计量学词汇 基础和通用概念及相关术语》（VIM）

2.6 JJF1001《通用计量术语及定义》

2.7 CNAS-CL01《检测和校准实验室能力认可准则》

2.8 RB/T 197《检测和校准结果及与规范符合性的报告指南》

2.9 IEC Guide 115《测量不确定度在电气领域合格评定活动中的应用》

3 术语和定义

VIM、JJF1001、GB/T 27418 和 GB/T 27419 中界定及下列术语和定义适用于本

文件。

3.1 合格评定对象（objects of conformity assessment）

能通过可测量的属性进行区别的事物。

注 1：本定义仅适用于本文件。

注 2：合格评定对象可以是实物量具、标准物质或样品等，例如量块、数字多用表或工业废水样品。

3.2 规定要求（specified requirement）

明示的需求或期望。

注：可在诸如法规、标准和技术规范这样的规范性文件中对规范要求做出明确说明。

注 1：特定要求里的术语“期望”并非随机变量的“期望”；

注 2：在本文件中，典型的特定要求表现为事物可测量属性的允许值区间的形式。

例如，工业废水样品（合格评定对象）中的溶解水银（属性）的质量浓度不高于 10ng/L；食品店用秤（合格评定对象）在称 1kg 的标准重量时，示值 R（属性）需满足[999.5g<R<1000.5g]。

3.3 容许限（tolerance limit）

规定限

可测量属性允许值的规定上限和下限。

3.4 容许区间（tolerance interval）

可测量属性允许值的区间

注 1：在没有其他说明的情况下，容许限在容许区间里。

注 2：符合性判定中的术语“容许区间”和统计学中的“容许区间”涵义不一样；

注 3：容许区间有时也称作规范区域。

3.5 容差（tolerance）

规定容差

容许上限和下限之间差值。

3.6 合格概率（conformance probability）

事物满足规定要求的概率。

3.7 接受限（acceptance limit）

测得值的允许上限或下限。

3.8 接受区间（acceptance interval）

测得值的允许区间。

注 1：在没有其他说明的情况下，接受限值在接受区间里。

注 2：接受区间有时也称作接受区域或合格区间。

3.9 拒绝区间（rejection interval）

测得值的不允许区间。

注：拒绝区间有时也称作拒绝区域或不合格区间。

3.10 保护带（guard band）

容许限和接受限之间的区间。

3.11 判定规则（decision rule）

当声明测量结果与规定要求的符合性时，描述如何考虑测量不确定度的规则。

3.12 特定消费者风险（specific consumer’s risk）

特定不合格事物被判断为合格的概率。

3.13 特定生产商风险（specific producer’s risk）

特定合格事物被判断为不合格的概率。

3.14 全局消费者风险（global consumer’s risk）

任何一个不合格的事物在后续的符合性判定中判断为合格的概率。

也称为消费者风险。

3.15 全局生产商风险（global producer’s risk）

任何一个合格的事物在后续的符合性判定中判断为不合格的概率。

也称为生产商风险。

3.16 测量能力指数（measurement capability index）

容差除以事物属性测得值的标准测量不确定度的倍数。

注：测量能力指数在有些文件中也叫测试不确定度比（Test Uncertainty Ratio，TUR）。

4 概述

4.1 符合性判定中的测量活动

判断合格评定对象的某一属性是否符合规定要求，通常包含三个步骤：

a）测量目标属性；

b）将测量结果与规定要求相比较；

c）做出下一步决定。

测量是为了获得足够的信息，使判定结果的风险在可接受的范围内。合理的测量方案应在降低测量不确定度所需的成本和获得更准确的被测量真值信息所带来的益处之间做出折衷考虑，具有适当的测量不确定度和足够的真值信息，以便在可接受的风险水平上做出合格与否的判定。

为易于理解，本文件中用于演示的符合性判定是二元决策的符合性判定，即判定结论只有合格与不合格两种，这也是一种常见判定情况。

4.2 容许限和容许区间

被测量（目标属性）的规定要求通常由限值组成，此限值称为容许限，它将被测量的允许值区间和不允许值区间分隔开。允许值区间也叫容许区间，分为两类：

a） 含一个容许上限或一个容许下限的单侧容许区间；

b） 同时含有容许上限和容许下限的双侧容许区间。

在以上任意情况中，当被测量的真值位于容许区间中则称该事物符合规定要求，反之则不符合要求，上面两种容许区间如图 1 所示。

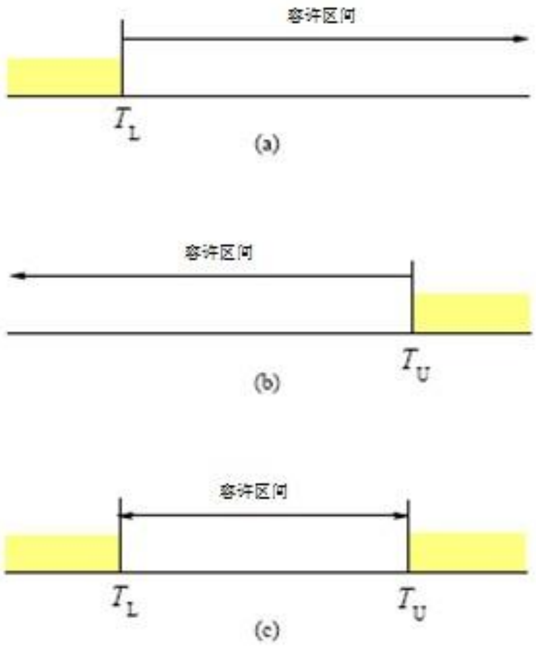

图 1 容许区间：（a）含单一容许下限 T_L 的单侧区间；（b）含单一容许上限 T_U 的单侧区间；（c）同时含有容许上限 T_U 和容许下限 T_L 的双侧区间，差值 $T_U - T_L$ 称为容差。

由于某些物理或理论原因，单侧容许区间通常具有隐含的附加限值，这些限值并未明确规定。这样的容许区间实际上是双侧的，包含一个规定的限值和一个隐含的限值，例如下文的例 4 和例 5。

例 1 单一容许上限

规定某种稳压二极管的击穿电压 V_b 不高于-5.4V，对于合格的二极管，V_b 值落在开放区间 $V_b \leq -5.4\text{V}$ 内。

例 2 单一容许下限

规定某种饮料金属容器的爆炸压力 B 不低于 490kPa，B 的合格值落在开放区间内 $B \geq 490\text{kPa}$。

例 3 明确的容许上限和下限

规定 OIML 的 E1 级 1 千克砝码质量的最大允许误差为$500\mu\text{g}$。也就是说一个砝码的质量 m 按规定不得低于 0.9999995kg，不得高于 1.0000005kg。合格的 1kg 砝码的质量误差$E = m - m_0$（其中$m_0 = 1\text{kg}$）应为$-500\mu\text{g} \le E \le 500\mu\text{g}$。

例 4 明确的容许上限和隐含的容许下限

某环境法规要求工业废水中水银的质量浓度 X 不高于$10\text{ng}/\text{L}$，这是一个明确的容许上限。由于质量浓度不可能低于 0，因此有一个隐含的容许下限$0\text{ng}/\text{L}$。遵守该法规的废水水样中的水银含量应为$0\text{ng}/\text{L} \le X \le 10\text{ng}/\text{L}$。

例 5 明确的容许下限和隐含的容许上限

规定食品防腐剂粉末状苯甲酸钠的纯度 P 不低于 99%（以干基的质量百分含量计），这是一个明确的容许下限。实际上纯度是不可能高于 100%的，这是隐含的容许上限。因此合格的苯酸钠样本的纯度应为$99\% \le P \le 100\%$。

4.3 判定规则与接受限、接受区间

当需要进行符合性判定时，直接将测量结果与容许区间相比较进行判定，会有以下 5 种情况（针对单侧容许区间，双侧容许区间类似）：

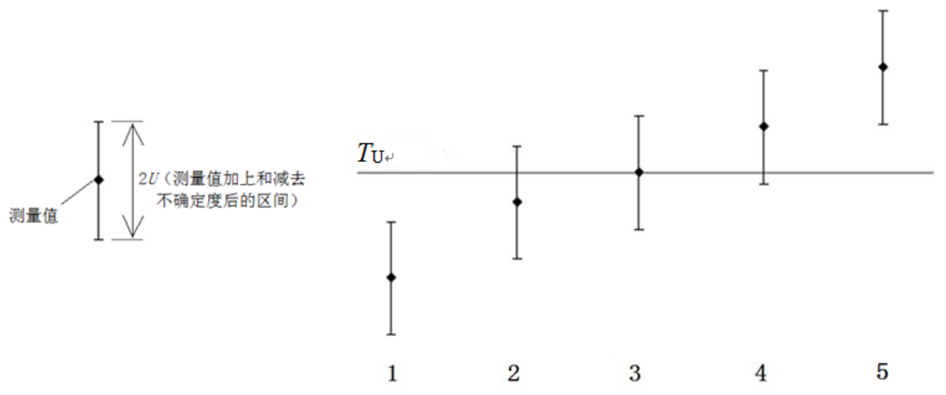

图 2 符合性判定的 5 种情况

对图 2 中的情况直接进行判定（不考虑测量不确定度），会有以下 4 种结果：

有效合格（正确接受）：测得值在容许区间内，真值也在容许区间内，如图 3（a）；

无效合格（错误接受）：测得值在容许区间内，但真值在容许区间外，如图 3（b）；

有效不合格（正确拒绝）：测得值在容许区间外，真值也在容许区间外，如图 3（d）；

无效不合格（错误拒绝）：测得值在容许区间外，但真值在容许区间内，如

图 3（c）。

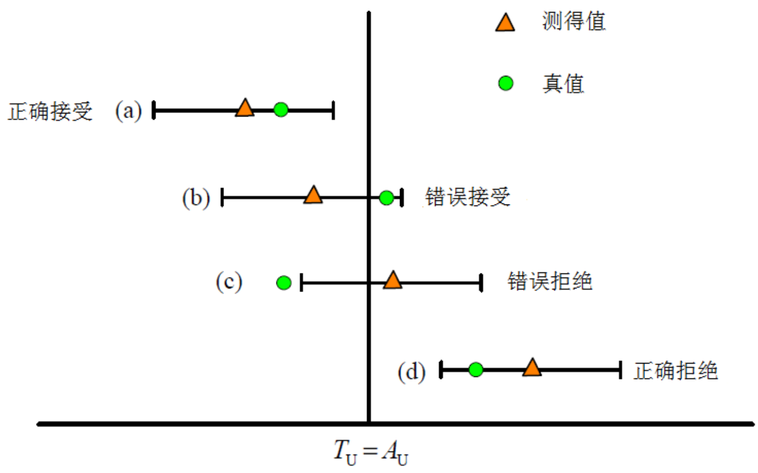

图 3 直接判定的四种情况

从图中可以看出，图 2 中的 1 和 5 两种情况是可以直接判断为合格或者不合格，而在 2、3、4 等三种情况，在考虑测量不确定度的情况下，不能直接判断是否合格（详见 RB/T 197-2015 《检测和校准结果及与规范符合性的报告指南》)，需要选择合理的判定规则。判定规则规定了如何考虑测量不确定度，由此确定可接受的测得值的区间，即接受区间，该区间的上限和/或下限，就是接受限。只要测得值出现在接受区间内，均可判定为合格。

4.4 选择判定规则的一般流程

当客户要求针对测量结果做出符合性声明时，合格评定机构应在合同评审时选择合适的判定规则并征得客户同意。需要注意的是，没有一种判定规则适用于所有的符合性判定活动，选择判定规则时应综合考虑被测属性的特点、所用的标准或技术规范要求、测量结果、双方风险等多方面的因素。图 4 是选择判定规则的一般流程图。

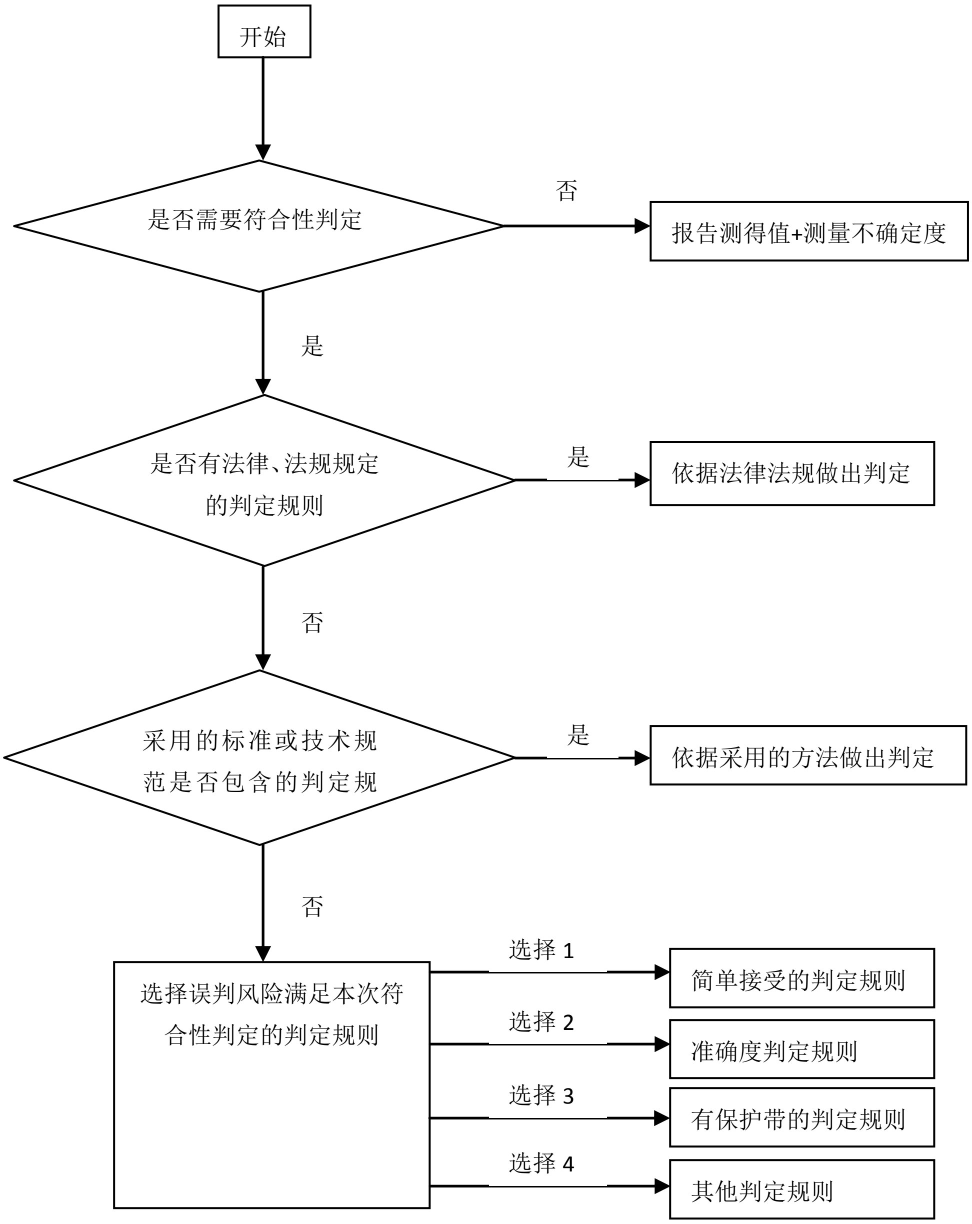

图 4 选择判定规则的通用流程图

5 几种常见的判定规则

5.1 简单接受（风险共担）判定规则

一种主要且应用广泛的判定规则叫做简单接受或者风险共担判定规则。这种判定规则不考虑测量不确定度的影响，被测属性的测得值落在容许区间时判定为合格，由实验室和客户共同承担误判的风险。

下列两种情况可用简单接受判定规则：

1）依据的标准或规范中没有明确要求符合性判定时考虑测量不确定度的影响；

2）客户和实验室之间有协议声明符合性判定时不需考虑测量不确定度的影

响。

实际应用中，一般假设测量方法的不确定度是可以接受的，而且其不确定度在必要时是可以评定的。对于双侧容许区间，测量不确定度与容差的一半之比通常应小于或等于 1: 3。

例 6 测量仪器特性评定

在 JJF 1094《测量仪器特性评定》中，在对测量仪器特性进行符合性判定时，需要仪器的示值误差落在最大允许误差区间$[-E_{\max}, E_{\max}]$内。此时如果仪器示值误差测得值e的扩展不确定度U_{95}与最大允许误差之比满足$U_{95} \le \frac{E_{\max}}{3}$，就可以忽略测量不确定度的影响。只要被判定测量仪器的示值误差测得值e满足$e \le E_{\max}$，即判为合格，反之则不合格。

5.2 "准确度法"判定规则

"准确度法"是通过严格控制测量时的人员、设备、环境、程序等影响测量不确定度的因素，将不确定度的变化控制在可以接受的小范围内，在符合性判定时可忽略测量不确定度的影响。

IEC Guide 115《电气领域合格评定活动测量不确定度的应用》中描述的"准确度法"，要求电气实验室使用先进的测量设备和完善的检测方法，并通过如下方式控制测量不确定度影响因素的变化：

（a）测量仪器的最大允许误差在规定限值内；

（b）环境条件变化在规定限值内；

（c）文件化的测试流程；

（d）有技术资质的人员。

如果能满足以上要求，则符合性判定时可以不考虑测量不确定度的影响。

这种判定方法也常见于我国计量检定领域，计量检定属于法定计量活动，需要明确给出测量仪器是否合格。因此，我国检定规程通常都有计量器具控制、检定项目、检定方法、检定周期等章节，这些内容将影响测量不确定度的可变量来源控制在规定限值内。执行检定规程的测量活动，其测量不确定度认为是可以忽略的，在符合性判定中不需考虑，将被测仪器的示值误差与其最大允许误差作比较，就可以判定是否合格。

例 7 供电电源输出电压测量（IEC Guide 115）

试验方法：将电源连接到额定电压（最大允许误差为±2%）、额定频率的试验源。使用无感电阻作为负载，当电流值达到额定电流（最大允许误差为±2%）时，测量电源的输出电压。测量时的环境温度为 23℃±2℃。测量用表的准确度符合 CTL 决议 251A 中的要求（如表 4 所示）。当供电电源输出电压在其额定电压的±5%范围（容许区间）内时，判定为合格。

例如，试验源额定输入 240V，50Hz；供电电源额定输出电压 DC 5V ，2A，用如下仪表实施测量：

测量用表	满量程准确度	CTL 决议 251A 允许的最大值
温度表	±1.0℃	±2.0℃
电压表	±0.5%	±1.5%
频率表	±0.2%	±0.2%
电流表	±0.5%	±1.5%

测量条件和结果如下表，供电电源输出电压的测量结果为 5.1V，容许区间为 $[4.75\text{V}, 5.25\text{V}]$，因此可判定合格。

输入		输出	
电压/V	频率/Hz	电流/A	电压/V
242	50	2.01	5.1
测量环境温度		24℃	

5.3 考虑测量能力指数的判定规则

5.3.1 测量能力指数

对被测量 Y 进行测量后，测得值 $y=\eta_{\mathrm{m}}$，标准测量不确定度 $u=u_{\mathrm{m}}$。容许上限为 T_{U},容许下限为 T_{L},容差 $T=T_{\mathrm{U}}-T_{\mathrm{L}}$，根据 3.16 的定义，测量能力指数表示为：

$$C_{\mathrm{m}}=\frac{T_{\mathrm{U}}-T_{\mathrm{L}}}{4u_{\mathrm{m}}}=\frac{T}{4u_{\mathrm{m}}}=\frac{T}{2U} \qquad (1)$$

其中 $U=2u_{\mathrm{m}}$ 是扩展不确定度，包含因子 $k=2$。标准测量不确定度的倍数之所以选为 4，是因为在报告测量结果时通常采用的包含区间为 $[\eta_{\mathrm{m}}-2u_{\mathrm{m}},\eta_{\mathrm{m}}+2u_{\mathrm{m}}]$。在被测量 Y 为正态概率密度函数的情况下，该区间的包含概率接近 95%。

5.3.2 考虑测量能力指数的判定规则

在5.1中，符合性判定忽略测量不确定度的影响时，（对于双侧容许区间）测量不确定度与容差的一半之比通常应小于或等于1: 3，此时测量能力指数为：

$$C_m=\frac{T}{2U}=3$$

区间（$T_{\mathrm{L}}+U$，T_{U}-U）占区间（T_{L}，T_{U}）的66.7%（如图5所示）。被测量 Y 为正态分布时，η_{m} 落在区间（T_{L}，T_{U}）内时的概率约为72%，即误判风险约

为28%（此处为粗略估算）。如果按η_{m}落在区间（$T_{\mathrm{L}}+U$，$T_{\mathrm{U}}-U$）才判为合格，则合格概率等于扩展不确定度U的置信水平，误判风险小于5%。（当被测量Y为其他分布时，可根据其概率密度函数计算合格概率，详见第6.2节）。

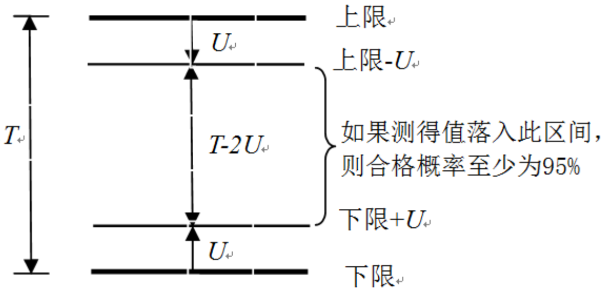

图5 测量能力指数$C_m=3$时的示意图

由于在大多数测量活动中，被测量Y服从正态分布，因此本文件以正态分布为例讨论测量能力指数与误判风险的关系。

以d代表η_m在容许区间内的位置，定义：

$$d=\frac{(\eta_{\mathrm{m}}-T_{\mathrm{U}})+(\eta_{\mathrm{m}}-T_{\mathrm{L}})}{T} \qquad (2)$$

当$\eta_{\mathrm{m}}=T_{\mathrm{U}}$时，$d=1$；$\eta_{\mathrm{m}}=T_{\mathrm{L}}$时，$d=-1$。

根据第6.2节计算合格概率的方法，可计算在不同测量能力指数（1~6）下，η_{m}在容许区间内的不同位置与误判风险的关系，如图6所示。

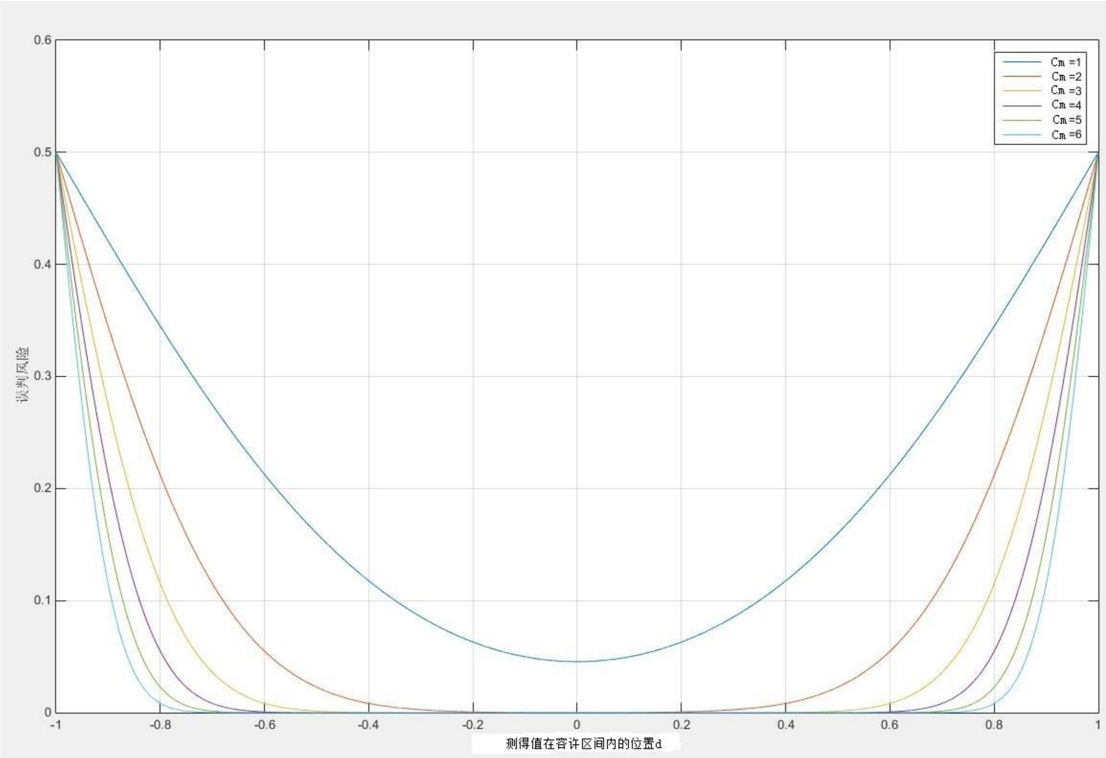

图6 被测量Y服从正态分布时，测得值η_{m}在容许区间内的位置与误判风险的关系

由图6可知，对于同一η_{m}值，测量能力指数越大，误判的风险越低。因此，在合格概率未知的情况下，判定规则可考虑用提高测量能力指数的方式，降低测得值的误判风险。但需要注意的是，在容许区间一定的情况下，提高测量能力指数意味着采用准确度等级更高的测量设备和/或更精密的测量程序，这都增加了测量的成本，因此实际应用中，要在权衡测量能力指数和误判风险的基础上，制定或选择合理的判定规则。

5.4 有保护带的判定规则

相对于风险共担判定规则，有保护带的判定规则带有风险偏好，根据出现误判后果的严重程度，在容许区间的基础上设置保护带，确定接受区间，减小其中一方的误判风险。需要注意的是，风险不能消除，当减少其中一方的误判风险时，会大大增另一方的误判风险。

具体而言，有保护带的判定规则又分为有保护带的接受和有保护带的拒绝。

5.4.1 有保护带的接受

通过在容许区间里设置接受限A_{U}可以降低无效合格的风险（即消费者风险）。如图 7 所示，由T_{U}和A_{U}确定的区间叫做保护带，A_{U}确定的区间为接受区间（也称为合格区间），落在接受区间内的测得值均判为合格。有保护带的接受也叫可靠接受、严格接受或积极符合接受。

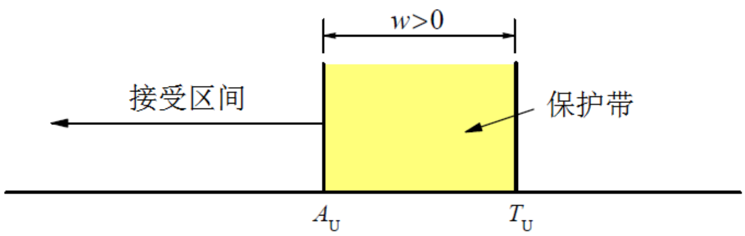

图 7 单侧容许区间有保护带接受的判定规则。接受上限 A_U 位于容许上限 T_U 之内，确定了接受区间，降低了无效合格的概率。长度参数 w 习惯上取正：$w = T_U - A_U > 0$。

容许限值和对应的接受限值之间的差值确定了保护带的长度参数 w：$w = T_U - A_U$，对于有保护带的接受，$w > 0$。

在实际应用中，长度参数 w 一般取扩展不确定度（包含因子 $k = 2$，$U = 2u$）的倍数：

$$w = rU \tag{3}$$

通常选择 $w = U$，$r = 1$，此时有效合格概率至少为 95%，这种保护带也叫 U_{95} 保护带。

对于双侧容许区间，接受上限和下限是对应的容许限值分别偏移一个保护带（长度参数 $w = U$），如图 8 所示。A_L 和 A_U 确定的区间为接受区间（图中合格区间）。

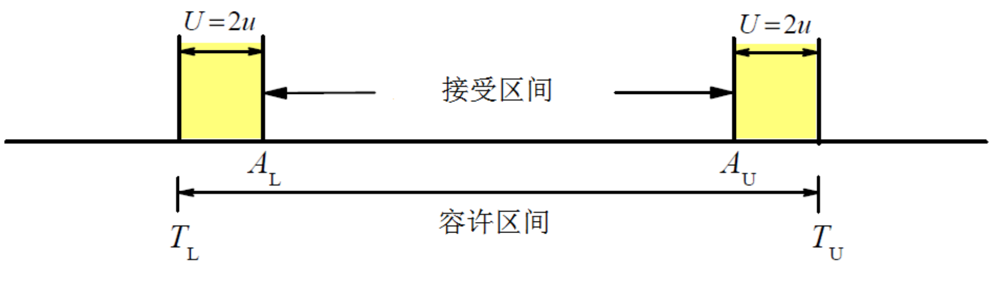

图 8 通过将容许区间的两侧各缩小一个扩展不确定度 U 的长度，确定的双侧接受区间。

5.4.2 有保护带的拒绝

通过在容许区间之外设置接受限 A_U 可以降低无效不合格的概率（即生产商风险），如图 9 所示。当需要获得超过限值的确凿证据时，一般使用这种有保护带拒绝的判定规则。有保护带的拒绝也叫可靠拒绝、严格拒绝、积极不符合拒绝。

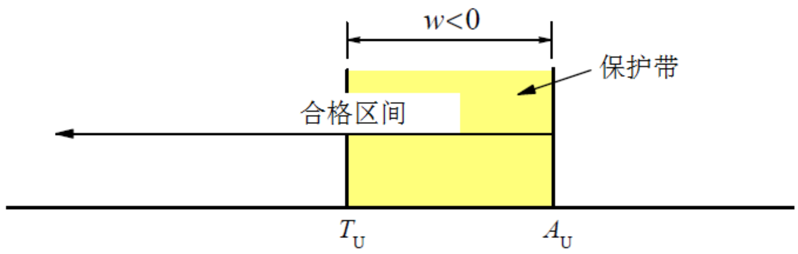

图 9 单侧容许区间有保护带拒绝的判定规则。在容许上限 T_U 之外的接受上限 A_U 确定了接受区间，

降低了无效不合格的概率。长度参数 $w = T_U - A_U < 0$。

双侧容许区间的情况类似。

当长度参数 $w = U$ 时，有效不合格的概率至少为 95%。

事实上，采用 U_{95} 保护带降低一方误判风险的同时，会显著增加另一方的误判风险，因此在实际应用中，可根据本文第 6 章介绍的方法计算合格概率，确定合理的保护带长度，也可根据历史测量数据、法律法规要求、双方协商结果等因素确定保护带长度。

例 8 国家标准 GB 6675.4-201《玩具安全 第 4 部分：特定元素的迁移》中采用的校正系数就是一种特殊的保护带形式。该标准考虑了测试方法精确度、结果有效性、误判风险、不同实验室间测试结果一致性等因素，要求对实验室的分析结果进行校正后再进行判定。

该标准规定，按标准第 7 至 9 章方法对玩具材料中可迁移有害元素含量进行测试，测试结果应减去按下表计算的校正值，得到校正后的分析结果，如果该分析结果不超过对应的最大限量要求，即认为符合标准要求。

表 1 各元素的分析校正系数

元素	锑	砷	钡	镉	铬	铅	汞	硒
分析校正系数%	60	60	30	30	30	30	50	60

规定标准规定：测试得到的分析结果应减去上表计算的校正值，以得到校正后的分析结果。凡玩具材料经校正的分析结果低于或等于 GB 6675.4 中规定的最大限量要求时，认为符合标准要求。

例如，某样品中铅元素的分析结果为 120mg/kg，表 1 中铅元素的分析校正系数为 30%，则校正后的分析结果=120-120×30%=120-36=84（mg/kg），标准规定的可迁移元素铅的最大限量要求为 90mg/kg，因此该样品的铅元素含量符合标准要求。

这个例子的保护带长度是测得值乘以校正系数，接受区间是在容许区间再加上保护带，实际上是有保护带的拒绝，降低了错误拒绝的概率。这种方法省去了复杂的数据处理过程，易于合格评定机构使用。

6 合格概率

5.1~5.3 节介绍的判定规则较为简便，在特定条件下，可不考虑测量不确定度，将被测量的测得值直接与容许区间比较得出判定结论，但对某一测得值的合格概率仍然是未知的。5.4 节引入了保护带的概念，当保护带长度 $w = U$ 时，一方的误判风险会降低至 5%以下，而另一方的误判风险会显著增加，如果权衡考虑双方风险，则需要根据合格概率选择其他长度的保护带。因此，量化的合格概率将有助于实验室、消费者和生产商准确评估风险，选择合理的判定规则。

6.1 被测量的相关知识

6.1.1 概率和信息

在符合性判定的测量中，被测量的数学模型是通过概率密度函数（PDF）建立的，该函数描述了被测量的可能取值，其形式取决于已知的被测量信息。这种信息通常包括两部分，测量前已知信息（先验信息）和测量得到的信息（后验信息）。

对于符合性判定的规定要求而言，信息量较少的被测量的概率密度函数通常较为平缓，意味着被测量可能值的范围较宽，通过测量获得被测量更多的信息，使概率密度函数变陡，则被测量可能值的范围变窄。

因此，测量的作用就是对被测量的（先验）信息进行更新，产生测量后的包含测量数据的（后验）信息。这种变换规则称为贝叶斯定理，基本的数学理论叫做贝叶斯概率论。

6.1.2 贝叶斯定理

被测量 Y 可以看作是可能值为 η 的随机变量。在测量 Y 之前，其合理可信的可能取值由先验概率密度函数 $g_0(\eta)$ 描述，该函数形式与测量系统无关。先验概率密度函数 $g_0(\eta)$ 的配置一般基于类似被测量在以前的测量中得到的知识，配置方法可参考 ISO/IEC Guide 98-4:2012 附录 B。

利用测量系统对 Y 进行测量，其输出可以看作是可能值为 η_{m} 的随机变量 Y_{m}。对 Y 的测量产生了测得值 η_{m}，由此得到基于新信息的 Y 的后验概率密度函数，表示为：

$$g(\eta \mid Y_{\mathrm{m}} = \eta_{\mathrm{m}}) =: g(\eta \mid \eta_{\mathrm{m}}) \tag{4}$$

先验概率密度函数和后验概率密度函数通过贝叶斯定理相关联：

$$g(\eta \mid \eta_{\mathrm{m}}) = C g_0(\eta) h(\eta_{\mathrm{m}} \mid \eta) \tag{5}$$

其中对于给定测得值η_{m}，C是常数，满足$\int_{-\infty}^{\infty} g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta = 1$。给定被测量为某一特定值$Y=\eta$，式（5）中的$h\left(\eta_{\mathrm{m}} \mid \eta\right)$是$\eta_{\mathrm{m}}$的概率密度函数。

将测得值η_m表示为η的函数，概率密度函数$h\left(\eta_{\mathrm{m}} \mid \eta\right)$称作给定$\eta_m$时$\eta$的似然函数，表示为：

$$h\left(\eta_{\mathrm{m}} \mid \eta\right) =: L\left(\eta : \eta_{\mathrm{m}}\right) \tag{6}$$

测量可以看作是一个输入-输出过程。按照这种观点，似然函数$L\left(\eta : \eta_{\mathrm{m}}\right)$描述了能产生可观测输出（测得值$\eta_{\mathrm{m}}$）的输入（$\eta$的值）的分布。

以数学模型表示的似然函数的形式一般取决于具体的测量原理和测量系统。在很多实际情况中，似然函数可用正态（高斯）分布描述。

在很多情况下，使用测量系统是为了给被测量相对较少的先验信息补充准确的测量信息。此时后验（测量后）概率密度函数实质上可由似然函数（包含测量信息）极度近似：

$$g\left(\eta \mid \eta_{\mathrm{m}}\right) = Ch\left(\eta_{\mathrm{m}} \mid \eta\right) \tag{7}$$

其中C是常数。

6.1.3 被测量的估计值和标准不确定度

测量结果通常可以由被测量的估计值和表征该估计值分散性的参数表示。本文中，被测量Y的估计值y就是期望$E\left(Y \mid \eta_{\mathrm{m}}\right)$。相应的分散性参数$u\left(y\right)=u$（称作标准不确定度）是$Y$的标准差，即方差$V\left(Y \mid \eta_{\mathrm{m}}\right)$的正平方根。

$$y = E\left(Y \mid \eta_{\mathrm{m}}\right) = \int_{-\infty}^{\infty} \eta g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta \tag{8}$$

$$u^2 = V\left(Y \mid \eta_{\mathrm{m}}\right) = \int_{-\infty}^{\infty} \left(\eta - y\right)^2 g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta \tag{9}$$

标准不确定度u描述了Y对于估计值y的分散性。当Y的概率密度函数是单峰并对称的函数（分布方式）时，估计值y是Y最可能的值。

测量活动产生了被测量的测得值η_{m}和相应的标准不确定度u_{m}。如果先验信息非常少，此时后验概率密度函数$g\left(\eta \mid \eta_{\mathrm{m}}\right)$可由测得值$y=\eta_{\mathrm{m}}$和相应的标准不确定度$u=u_{\mathrm{m}}$表示。

6.1.4 包含区间

在测量后，Y 不大于给定值 a 的概率为：

$$\Pr\left(Y \le a \mid \eta_{\mathrm{m}}\right) = G\left(a\right) = \int_{-\infty}^{a} g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta \tag{10}$$

其中 $G\left(z\right) = \int_{-\infty}^{z} g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta$ 是 Y（对于给定值 η_{m}）的分布函数。

由此 Y 落在区间 $\left[a,b\right]$（$a<b$）内的概率 p 为：

$$p = \Pr\left(a \le Y \le b \mid \eta_{\mathrm{m}}\right) = \int_{b}^{a} g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta = G\left(b\right) - G\left(a\right) \tag{11}$$

$\left[a,b\right]$ 的区间称为 Y 的包含区间，p 是对应的包含概率。

当 Y 的概率密度函数是对称单峰函数时，包含区间一般取中点为测得值 y，长度与标准不确定度的倍数相同。标准不确定度的倍数就是扩展不确定度 $U = ku$（其中 k 为包含因子）。

包含因子 k 取决于包含区间 $\left[y-U, y+U\right]$ 内期望的包含概率 p，二者之间的关系由 Y 的概率密度函数确定。

注 1：包含区间 $\left[y-U, y+U\right]$ 有时也称为不确定度区间。

注 2：如果 Y 的概率密度函数是非对称的，合理做法是在给定的包含概率下确定最短包含区间。

6.2 符合规定要求的合格概率

6.2.1 计算合格概率的一般原则

被测量 Y 的真值落在容许区间内时，可判为符合规定要求。Y 的信息是通过概率密度函数 $g\left(\eta \mid \eta_{\mathrm{m}}\right)$ 表达的，因此符合性声明是一定概率正确的推断。以 C 表示 Y 的允许（合格）值的集合，以 p_c 表示的合格概率为：

$$p_{\mathrm{c}} = \Pr\left(Y \in C \mid \eta_{\mathrm{m}}\right) = \int_{C} g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta \tag{12}$$

那么被测量 Y 的双侧容许区间（例如上限为 T_{U}，下限为 T_{L}，$C = \left[T_{\mathrm{L}}, T_{\mathrm{U}}\right]$），合格概率为：

$$p_{\mathrm{c}} = \int_{T_{\mathrm{L}}}^{T_{\mathrm{U}}} g\left(\eta \mid \eta_{\mathrm{m}}\right) d\eta \tag{13}$$

如果只有合格和不合格两种情况，那么不合格的概率为：

$$\overline{p_{\mathrm{c}}} = 1 - p_{\mathrm{c}} \tag{14}$$

6.2.2 正态概率密度函数的合格概率

合格概率取决于由后验概率密度函数 $g(\eta|\eta_{\mathrm{m}})$ 表达的被测量 Y 的信息。在多数情况下，正态分布可以合理表征 Y 的信息。如果先验分布是正态的并且似然函数也可用正态分布描述，那么后验概率密度函数 $g(\eta|\eta_{\mathrm{m}})$ 也是正态分布；如果似然函数可用正态分布描述且先验信息很少，那么后验（测量后）概率密度函数 $g(\eta|\eta_{\mathrm{m}})$ 也是近似正态的，该正态分布的期望和标准差就是测得值 y 和标准不确定度 u。

注：正态概率密度函数的特性见 ISO/IEC Guide 98-4:2012 附录 A。

由于多数被测量 Y 是正态分布的，本文以正态概率密度函数为例计算合格概率，并且本文这种方法也可以用于被测量为 t 分布的情况。

正态分布完全由其期望（均值）和标准差确定。假设被测量 Y 的概率密度函数 $g(\eta|\eta_{\mathrm{m}})$ 是由测得值（期望）y 和标准不确定度（标准差）u 确定的正态分布（或者极度逼近），则 $g(\eta|\eta_{\mathrm{m}})$ 表示为：

$$g(\eta|\eta_{\mathrm{m}})=\frac{1}{u\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\eta-y}{u}\right)^2\right]=:\varphi(\eta;y,u^2) \qquad (15)$$

测得值 y 一般由 η_{m} 确定，即 $y=y(\eta_{\mathrm{m}})$。当测量前 Y 的先验信息很少时，通常 $y\approx\eta_{\mathrm{m}}$。

从式（11）和概率密度函数（15）可以得出，

$$\Pr(a\le Y\le b|\eta_{\mathrm{m}})=\int_a^b g(\eta|\eta_{\mathrm{m}})d\eta=\frac{1}{u\sqrt{2\pi}}\int_a^b exp\left[-\frac{1}{2}\left(\frac{\eta-y}{u}\right)^2\right]d\eta \qquad (16)$$

令 $z=(\eta-y)/u$，$dz=d\eta/u$，且

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\exp\left(-t^2/2\right)dt=\int_{-\infty}^{z}\varphi_0(t)dt \qquad (17)$$

其中 $\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\exp\left(-t^2/2\right)dt$ 为标准正态分布函数。

将（17）代入（16）中，则 Y 落在区间 $[a,b]$ 的概率为：

$$\Pr(a\le Y\le b|\eta_{\mathrm{m}})=\int_{(a-y)/u}^{(b-y)/u}\varphi_0(z)\,dz=\Phi\left(\frac{b-y}{u}\right)-\Phi\left(\frac{a-y}{u}\right) \qquad (18)$$

其中 $y = y(\eta_{\mathrm{m}})$。

6.3 单侧容许区间正态概率密度函数的合格概率

6.3.1 含单一容许下限的单侧容许区间

图 10 展示了含单一容许下限 T_{L} 的单侧容许区间。被测量 Y 的合格取值落在 $\eta \geq T_{\mathrm{L}}$ 的区间内。测量后，Y 的信息通过测得值 y 和标准不确定度 u 共同确定的正态概率密度函数描述。容许区间和 PDF 同时表示在图 10 上。测得值 y 落在容许区间里；T_{L} 左侧的阴影部分表示不合格的概率。根据式（18），此处 $a = T_{\mathrm{L}}$，$b \to \infty$，且 $\Phi(\infty) = 1$，合格概率为：

$$p_{\mathrm{c}} = \Phi(\infty) - \Phi(\frac{T_{\mathrm{L}} - y}{u}) = 1 - \Phi(\frac{T_{\mathrm{L}} - y}{u}) \qquad (19)$$

因 $\Phi(t) + \Phi(-t) = 1$，所以概率式（19）也可以表示为：

$$p_{\mathrm{c}} = \Phi(\frac{y - T_{\mathrm{L}}}{u}) \qquad (20)$$

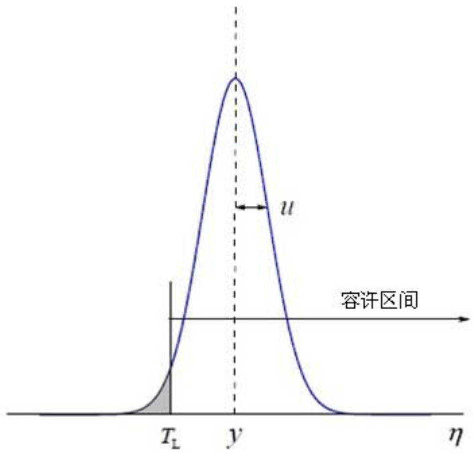

图 10 含单一容许下限的容许区间。Y 的合格值落在区间 $\eta \geq T_{\mathrm{L}}$ 内。

6.3.2 含单一容许上限的单侧容许区间

图 11 显示了单一容许上限 T_{U} 的单侧容许区间和正态概率密度函数。Y 的合格值落在区间 $\eta \leq T_{\mathrm{U}}$ 内。在这种情况下，T_{U} 右侧的阴影部分表示不合格的概率。根据式（18），此处 $a \to \infty$，$b = T_{\mathrm{U}}$，且 $\Phi(-\infty) = 0$，合格概率为：

$$p_{\mathrm{c}} = \Phi(\frac{T_{\mathrm{U}} - y}{u}) - \Phi(-\infty) = \Phi(\frac{T_{\mathrm{U}} - y}{u}) \qquad (21)$$

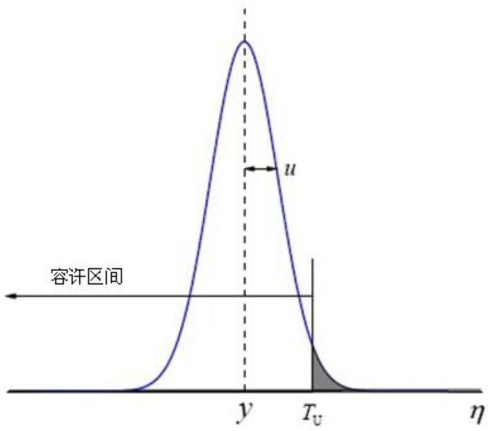

图 11 含单一容许上限的容许区间， Y 的合格值落在区间 $\eta \le T_U$ 内。

6.3.3 单侧容许区间的一般计算方法

概率式（20）和（21）可以用相同的形式表示：

$$p_{\mathrm{c}} = \Phi(z) \tag{22}$$

其中单下限情况 $z = (y - T_{\mathrm{L}})/u$，单上限情况 $z = (T_{\mathrm{U}} - y)/u$。两种情况中，$p_c$ 在测得值 y 处于容许区间（$z \ge 0$）时大于或等于 1/2，反之则小于 1/2。表 2 是合格概率 p_c 的几种取值下 z 的值。

表 2 带有单侧容许区间的正态概率密度函数的合格概率 p_{c} 和不合格概率 $\overline{p_{\mathrm{c}}}$

p_{c}	$\overline{p_{\mathrm{c}}}$	z
0.80	0.20	0.84
0.90	0.10	1.28
0.95	0.05	1.64
0.99	0.01	2.33
0.999	0.001	3.09

注：其他 p_{c} 与 z 的对应关系可查询标准正态分布表。

例 9 测量某稳压二极管的击穿电压 V_{b}，得到测得值 $v_{\mathrm{b}} = -5.47\mathrm{V}$，标准不确定度 $u = 0.05\mathrm{V}$，稳压二极管击穿电压的合格区间为 $V_{\mathrm{b}} \le -5.47\mathrm{V}$，即容许上限为 $-5.47\mathrm{V}$。由此得到 $z = [-5.40 - (-5.47)]/0.05 = 1.40$，且根据式（22），

$p_c = \Phi(1.40) = 0.92$。因此该二极管的合格概率为 92%。

例 10 用高压水对某金属罐进行破坏性试验，测得爆破力为 B。测得值为 $b = 509.7\text{kPa}$，标准不确定度为 $u = 8.6\text{kPa}$。金属罐的合格要求（即爆破力的下限要求）为 $B \geq 490k\text{Pa}$。由此得到 $z = (509.7 - 490)/8.6 = 2.3$，$p_c = \Phi(2.3) = 0.99$。因此该金属罐的合格概率为 99%。

6.4 双侧容许区间正态概率密度函数的合格概率

图 12 展示了容许限为 T_L 和 T_U 的双侧容许区间，区间长度 $T = T_U - T_L$ 定义了容差 T。如前面所述，认为被测量 Y 的知识是由正态概率密度函数表达的。测得值 y 落在容许区间内，并且可以看到一部分概率在容许上限之外 $\eta > T_U$ 的区域内。

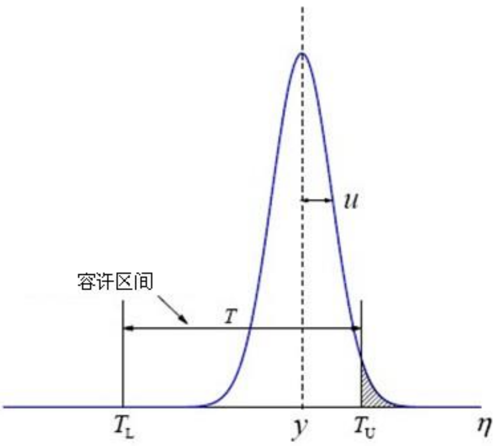

图 12 双侧容许区间，区间长度 $T_U - T_L$ 等于容差 T。Y 的合格值落在区间 $T_L \leq \eta \leq T_U$ 内。

由式（18）及 $a = T_L, b = T_U$ 得到合格概率：

$$p_c = \Phi(\frac{T_U - y}{u}) - \Phi(\frac{T_L - y}{u}) \tag{23}$$

例 11 美国汽车工程学会要求 40 级发动机润滑油样品 100℃时的运动粘度 Y 不低于 12.5mm²/s 且不高于 16.3mm²/s。在 100℃时样品运动粘度的测得值为 $y = 13.6\text{mm}^2/\text{s}$，标准不确定度为 $u = 1.8\text{mm}^2/\text{s}$。根据式（23）得出：

$(T_U - y)/u = (16.3 - 13.6)/1.8 = 1.5$，$(T_L - y)/u = (12.5 - 13.6)/1.8 = -0.6$。

因此，$p_c = \Phi(1.5) - \Phi(-0.6) = 0.93 - 0.27 = 0.66$

该发动机润滑油样品合格概率为 66%。

例 12 高速公路测速

高速公路测速执法中，警察通过雷达或激光测速仪测量机动车的速度。如果不考虑测量不确定度，贸然根据测得的速度值出具超速罚单可能会引起诉讼，因此必须要有确凿的证据表明驾驶员确实超速。

使用多普勒雷达进行现场测速时，在 50km/h 到 150km/h 的范围内，速度测量的相对标准不确定度为 2%。被测速度的知识假设为一个期望为υ，标准差为 0.02υ 的正态概率密度函数。

那么在这种情况下，对于速度限值 $\upsilon_0 = 100\,\mathrm{km/h}$，需要设置一个速度阈值 $\upsilon_{\max}$（接受限）可以使得被测速度 $\upsilon \geq \upsilon_{\max}$ 的概率不低于 99%。

这个数学问题等同于计算一个有单侧容许限的合格概率问题。此问题需要计算 $z = (\upsilon_{\max} - \upsilon_0)/(0.02\upsilon_{\max})$，使得 $V \geq \upsilon_0$ 的概率为 99.9%。通过表 2 $p_{\mathrm{c}} = 0.999$ 时得到 $z = 3.09$，因此确凿的超速速度为：

$$\upsilon_{\max} = \frac{\upsilon_0}{1-0.02\upsilon} = \frac{100}{1-0.02\times 3.09}\,\mathrm{km/h} \approx 107\,\mathrm{km/h}$$

区间 $[100\,\mathrm{km/h} \leq v \leq 107\,\mathrm{km/h}]$ 实际上是一个保护带，确保被测速度大于等于 107km/h 时，超速的概率至少为 99.9%。

例 13 活畜和畜产品中的药物检测

合成代谢类固醇诺龙是禁止给食品动物使用的生长促生剂。这种物质会天然的存在于某些活畜体内，因此为此物质设定了阈值（容许限）$T = 2.00\,\mu\mathrm{g/L}$。

在诺龙的筛查检测中，被测质量浓度超过阈值的概率大于等于 95%时，认为是可疑的并且需要进一步确认。

在筛查检测中，实验室希望接受限 A 为：

$$A = T + g$$

其中 $g = |w|$ 是保护带（见图 11），使得质量浓度的测得值 $y \geq A$ 时，$Y \geq T$ 的概率不小于 95%。由 5.3 节可知，这个保护带实际上就是 U_{95} 保护带。

实验室将接近阈值质量浓度的物质加到十个空白样品中，以确认其测量程序。然后在实验室内复现性条件下检测样品，产生可观测的复现性标准差 $s = 0.20\,\mu\mathrm{g/L}$。

实验室从加标试验中得出其检测结果没有系统误差的结论。测量不确定度主

要来自于复现性，因此认为诺龙质量浓度Y的概率密度函数是自由度$\nu=9$的t分布。

由t分布（单侧，自由度$\nu=9$）分位数表查得$t_{0.95}=1.83$，计算保护带为：

$$g=t_{0.95}\times s=1.83\times 0.20\mu\text{g}/\text{L}=0.37\mu\text{g}/\text{L}$$

样品诺龙质量浓度的测得值y大于等于A时，认为样品是可疑的，其中

$$A=\left(2.00+0.37\right)\mu\text{g}/\text{L}=2.37\mu\text{g}/\text{L}\text{。}$$

6.5 合格概率和包含区间

对于被测量Y而言，如果用带有包含概率p的包含区间描述测量结果，那么包含区间和容许区间、合格概率p_c的关系如图 13 所示。如果包含区间完全落在容许区间内，那么$p_{\text{c}}\geq p$（如图中 b），可以做出合格的判断，反之则不能确定p_{c}（如图中 a），不能判断是否合格，这与 4.3 节所述内容是一致的。

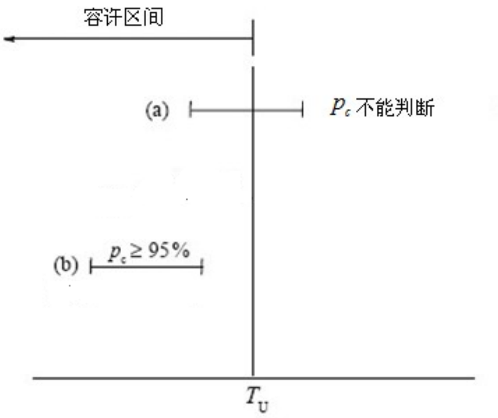

图 13 在单侧容许上限附近的两个包含概率为 95%的包含区间

一般而言，如果$\left[\eta_{\text{low}},\eta_{\text{high}}\right]$是$Y$包含概率为$p$的包含区间，那么

对于单一容许上限，如果$\eta_{\text{high}}\leq T_{\text{U}}$，则$p_{\text{c}}\geq p$；

对于单一容许下限，如果$\eta_{\text{low}}\geq T_{\text{L}}$，则$p_{\text{c}}\geq p$；

对于容许上限为T_{U}，容许下限为T_{L}，如果$\eta_{\text{high}}\leq T_{\text{U}}$且$\eta_{\text{low}}\geq T_{\text{L}}$，则$p_{\text{c}}\geq p$。

注 1：将包含区间和容许区间作比较是被测量符合性判定的基础。

注 2：给定Y的概率密度函数，一般可以计算出合格概率。被测量的概率密度函数所包含的信息要多于带有包含概率的包含区间。

注 3：当进行测量仪器的符合性判定时，特别是依据特定标准进行的评价，

被测量的定义及其测量不确定度的评估可能不明确并且需要特别关注的。

7 消费者和生产商风险

7.1 总则

符合性判定采用二元决策判定准则时，对可测量属性进行测量，如果测得值位于接受区间内，符合性评价结果为合格。反之若该值位于接受区间之外，则符合性评价结果为不合格。图 14 再现图 1 中表示了（合格值的）容许区间和（允许测得值的）接受区间。

图 14 基于测得值的二元符合性判定，规定可测量属性（被测量）的真值需落在由容许区间 (T_L, T_U) 内，当属性测得值位于接受区间 (A_L, A_U) 内时判定为合格，反之为不合格。

保护带的使用降低了根据测量信息进行判定时的误判概率。在本文第 5 节和第 6 节的分析中，均假设被测量 Y 的先验信息很少，后验概率密度函数 $g(\eta|\eta_m)$ 完全由测量得到的信息确定，误判风险取决于后验概率密度函数和测得值，这样的误判风险实际上是特定风险。本节将考虑被测量 Y 的先验分布，针对生产商生产过程的误判概率进行精确分析。通常，这一概率取决于生产程序和测量系统这两个因素。

如果测量系统绝对准确，所有的符合性判定结果将完全正确且风险为零。但测量都会有测量不确定度，且随着测量不确定度的增大，误判的概率也会增大，当测得值接近容许限值时，误判的概率最大。

误判风险还取决于生产程序的特性。若生产出的产品可测量属性极少有位于容许限值附近的值，将大大减少误判的风险。反之如果这些值总是很接近容许限值，则需要将测量活动产生的测量结果不确定度引入符合性判定活动。本节将介绍如何评估这两个因素对符合性评价判定的影响。

7.2 生产过程和测量系统的概率密度函数

对于一个生产某一系列产品的生产程序，每一件产品都有可测量属性 Y，对

应可能值为η。例如这个程序是生产（标称）值为 10kΩ 电阻的生产线。在这个程序中随机抽取一件产品，开展测量活动前，Y 的信息表达为先验概率密度函数 $g_0(\eta)$。概率密度函数可 $g_0(\eta)$ 描述生产程序的特征，有时也称为过程概率密度。$g_0(\eta)$ 的形式是根据测量产品的目标属性进行测量获得的信息确定的。

通过对产品目标属性进行测量，对产品实施符合性判定。测量系统输出一个可测量的随机变量Y_{m}，对应的可能值为η_{m}。假定已知输入值$Y=\eta$，该值分布服从概率密度函数$h(\eta_{\mathrm{m}}|\eta)$。$h(\eta_{\mathrm{m}}|\eta)$ 的形式由测量系统、仪器设备校准、环境参数、材料特性等相关影响因素确定。

7.3 二元决策的符合性评价可能出现的结果

设C和$\tilde{C}$分别表示Y的合格值和不合格值。对应的A和$\tilde{A}$分别表征Y_{m}的接受（合格）区间和不接受（不合格）区间。以图 14 为例，C对应的Y处于区间$T_{\mathrm{L}} \le Y \le T_{\mathrm{U}}$，$A$对应的$Y_{\mathrm{m}}$处于区间$A_{\mathrm{L}} \le Y_{\mathrm{m}} \le A_{\mathrm{U}}$。

采用二元决策判定，类似于 4.3 节，对测得值η_m的符合性评价结论有：

有效合格：产品目标属性的符合性判定结论合格（$Y_{\mathrm{m}}=\eta_{\mathrm{m}} \in A$）并符合客观事实（$Y \in C$）。这是希望得到的符合性判定结论，可判定产品该属性符合要求。

无效合格：产品目标属性的符合性判定结论合格（$Y_{\mathrm{m}}=\eta_{\mathrm{m}} \in A$），但不符合客观事实（$Y \in \tilde{C}$）。这种错误结论将残次品判断为合格品并出售给消费者，带来的损失通常由消费者承担，因此这种误判称为消费者风险。

在实际符合性判定活动中依据被测量的测得值$Y_{\mathrm{m}}=\eta_{\mathrm{m}} \in A$得到的满意结果时，得到无效合格结论的概率称为特定消费者风险，符号为R_{C}^{*}。根据合格概率的定义，R_C^{*}为：

$$R_{\mathrm{C}}^{*}=1-p_{\mathrm{c}} \tag{24}$$

此时被测量值η_{m}位于合格区间内。生产程序中随机选取产品的情况下，测量活动导致出现无效合格结论的概率称为全局消费者风险，符号为R_{C}，其计算方法见 7.5。

有效不合格：产品目标属性的符合性判定不合格 $Y_{\mathrm{m}} = \eta_{\mathrm{m}} \in \tilde{A}$ 且符合客观事实（$Y \in \tilde{C}$）。这是希望得到的符合性判定结果，可判定产品该属性不符合要求。

无效不合格：产品目标属性的符合性判定不合格（$Y_{\mathrm{m}} = \eta_{\mathrm{m}} \in \tilde{A}$），但不符合客观事实（$Y \in C$）。这种错误结论将合格品误判为残次品，使生产商无法售出合格品，带来的损失由生产商承担，因此这种误判称为生产商风险。

实际符合性评价中依据被测量的值 $Y_{\mathrm{m}} = \eta_{\mathrm{m}} \in \tilde{A}$ 得到不满意结果时，得到无效不合格结论的概率称为特定生产商风险，表示为 R_{P}^{*}。根据合格概率的定义，R_{P}^{*} 为：

$$R_{\mathrm{P}}^{*} = p_{\mathrm{c}} \tag{25}$$

此时被测得值 η_{m} 位于合格区间外。在生产程序中随机选取产品的情况下，测量活动导致出现无效不合格结论的概率称为全局生产商风险，符号为 R_{P}，R_{P} 的计算方法见 7.5。

7.4 Y 和 Y_{m} 的联合概率密度函数

如 7.3 节所示，给定测量结果，特定消费者风险 R_{C}^{*} 和特定生产商风险 R_{P}^{*} 仅与特定被测产品测量结果的合格概率有关。被测量 Y 处于容许区间之外同时测量结果 Y_{m} 位于合格区间内，则存在消费者风险。这两个条件同时出现的概率就是发生全局消费者风险的概率，由生产程序和测量系统确定的联合概率分布确定。

联合概率密度函数可以写成已知概率密度函数的乘积。被测量 Y 的值位于容许区间外且测得值 Y_{m} 位于接受区间内的概率，是生产程序生产 Y 的真值位于容许区间之外的产品的概率和已知被测量 Y 位于容许区间外但测量系统输出测得值 Y_{m} 位于接受区间内的概率的乘积。这就是全局消费者风险。

同样的，全局生产商风险也有一个类似的联合概率密度函数。若容许区间，生产程序和测量系统均固定不变，则全局消费者风险和全局生产商风险均由接受限确定，可以通过合理设置接受限来平衡两种风险。但是，通常设置接受限不可能同时最小化两种风险，在降低一种风险的同时必然会使另一种风险上升。

对于给定的生产程序和测量系统，随机抽取产品进行符合性判定，其可能输出结果的信息由联合概率密度函数描述。对该随机选取的产品，满足情况：（a）被测量 Y 的值位于区间 $\eta \le Y \le \eta + d\eta$（b）对 Y 的测量产生的测得值 Y_{m} 位于区间

$\eta_{\mathrm{m}} \le Y_{\mathrm{m}} \le \eta_{\mathrm{m}} + d\eta_{\mathrm{m}}$ 的概率为

$$\Pr(\eta \le Y \le \eta + d\eta \text{且} \le \eta_{\mathrm{m}} \le Y_{\mathrm{m}} \le \eta_{\mathrm{m}} + d\eta_{\mathrm{m}}) = f(\eta, \eta_{\mathrm{m}}) d\eta d\eta_{\mathrm{m}} \qquad (26)$$

这里 $f(\eta, \eta_{\mathrm{m}})$ 是 Y 和 Y_{m} 的联合概率密度函数。

根据概率的乘法规则，联合密度 $f(\eta, \eta_{\mathrm{m}})$ 可通过以下两种方式进行因式分解：

$$f(\eta, \eta_{\mathrm{m}}) = g_0(\eta) h(\eta_{\mathrm{m}} | \eta) \qquad (27a)$$

和

$$f(\eta, \eta_{\mathrm{m}}) = h_0(\eta_{\mathrm{m}}) g(\eta | \eta_{\mathrm{m}}) \qquad (27b)$$

式（27a）中等号右边的表达式是 7.2 节中描述的两个概率密度函数。已知这些概率密度函数的形式就能计算出式（27b）右侧的两个概率密度函数。详细计算过程可见 ISO/IEC Guide 98-4:2012 的附录 A。

7.5 全局风险的计算

7.5.1 历史背景

下文使用公式计算全局风险。计算全局风险的传统做法是测量具有同一标称值的庞大样本空间，通过各种测量结果出现的频率分布计算全局风险。采用这种计算方式的全局消费者风险等于庞大测量样品中通过符合性判定但实际上不满足要求的那部分产品。对于特定的不合格产品，需要在事后通过独立测量进行证明，且测量系统的不确定度要远远小于符合性判定中使用的测量系统不确定度。

本文中采用概率密度函数而不是频率分布计算全局风险，因此不需要考虑实际上可能不存在的样本总体。由此计算的概率在数值上与测得的样本频率是一致的。因此合理选择接受限，可以降低产品符合性判定中的全局误判风险。

7.5.2 通用公式

已知联合概率密度函数（27a）和两个概率密度函数 $g_0(\eta)$ 和 $h(\eta_{\mathrm{m}} | \eta)$，可以计算出 7.3 节中讨论的四种情况出现的概率。简单的讲，这些概率就是四种情况对应的四个区域在联合概率密度函数 $f(\eta, \eta_{\mathrm{m}})$ 上的积分值占比。

被测属性的全局消费者风险和全局生产商风险计算如下：

测得值位于接受区间内且 Y 的值位于容许区间外的情况，全局消费者风险为：

$$R_{\mathrm{C}}=\int_{\tilde{C}}\int_{A}g_0(\eta)h(\eta_{\mathrm{m}}|\eta)d\eta_{\mathrm{m}}d\eta \tag{28}$$

测量值位于接受区间外且Y的值位于容许区间内，全局生产商风险为：

$$R_{\mathrm{P}}=\int_{C}\int_{\tilde{A}}g_0(\eta)h(\eta_{\mathrm{m}}|\eta)d\eta_{\mathrm{m}}d\eta \tag{29}$$

式（28）、（29）为计算全局消费者风险和全局生产商风险的通用公式。根据概率密度函数$g_0(\eta)$和$h(\eta_{\mathrm{m}}|\eta)$的具体形式，可以计算出$R_{\mathrm{C}}$和$R_{\mathrm{P}}$。

7.5.3 二元决策的全局风险

如图 14 的二元决策符合性判定中，式（28）和（29）转变为：

$$R_{\mathrm{C}}=\left(\int_{-\infty}^{T_{\mathrm{L}}}+\int_{T_{\mathrm{U}}}^{\infty}\right)\int_{A_{\mathrm{L}}}^{A_{\mathrm{U}}}g_0(\eta)h(\eta_{\mathrm{m}}|\eta)d\eta_{\mathrm{m}}d\eta \tag{30}$$

和

$$R_{\mathrm{P}}=\left(\int_{-\infty}^{A_{\mathrm{L}}}+\int_{A_{\mathrm{U}}}^{\infty}\right)\int_{T_{\mathrm{L}}}^{T_{\mathrm{U}}}g_0(\eta)h(\eta_{\mathrm{m}}|\eta)d\eta d\eta_{\mathrm{m}} \tag{31}$$

在下面的例子中，联合概率密度函数服从正态分布，使用式（30）和（31）计算全局风险。

例 14 高精度电阻生产

某一电器元件的生产商，产品为标称值为 1500Ω 的高精度绕线电阻。每个绕线电阻（产品）的电阻值Y（可测量属性）可以确定位于定义的容许区间$T_L=1499.8\Omega$和$T_U=1500.2\Omega$。为评估产品生产线的性能，采用高精度欧姆表测量电阻样品的电阻值，测量不确定度可忽略不计。测得值的柱状图呈期望为标称值的正态分布，对应标准差$\sigma=0.12\Omega$。根据已知信息，生产程序的概率密度分布服从正态分布$g_0(\eta)=\varphi(\eta;y_0,u_0^2)$，对应$y_0=1500\Omega$，$u_0=\sigma=0.12\Omega$。

因此，该型号电阻生产线生产出合格产品的概率为：

$$p_{\mathrm{c}}=\int_{T_{\mathrm{L}}}^{T_{\mathrm{U}}}g_0(\eta)d\eta=\int_{1499.8}^{1500.2}\varphi(\eta;1500,0.12^2)d\eta\approx0.90=90\%$$

也就是说，如果生产商将生产线上下来的电阻直接出厂，那么其中大概会有10%的不合格品。这种结果在现代商业中是不能接受的。因此生产商有必要投入更多成本，购买更加精密的生产设备，提高生产出的电阻合格率。但是高精度的生产设备往往非常昂贵，投入大量成本提高不到 10%的产品合格率，显然不是经济的做法。对于生产商来说，更经济的做法是采用现有生产设备，并制定合理的产品检验程序来剔除不合格电阻。一般做法是在容许区间内设置合格区间（接受区间）。

本例中，电阻的合格检验采用校准过的高速欧姆表。根据该测量系统模型和

测量不确定度的评估，包括欧姆表校准的不确定度，$Y=\eta$ 的电阻样品的测得值服从正态概率密度函数 $h\left(\eta_{\mathrm{m}}|\eta\right)=\varphi\left(\eta_{\mathrm{m}};\eta,u_{\mathrm{m}}^{2}\right)$，测量不确定度 $u_{\mathrm{m}}=0.04\Omega$。

为了降低出厂电阻产品的不合格率（消费者风险），接受限设置在容许区间内，$A_{\mathrm{L}}=1499.82\Omega$，$A_{\mathrm{U}}=1500.18\Omega$。

长度参数 $w=(1500.2-1500.18)\Omega=0.02\Omega=0.25U$。

根据正态分布的概率密度函数的定义，先验概率密度函数为：

$$g_0(\eta)=\varphi\left(\eta;y_0,u_0^2\right)=\frac{1}{u_0\sqrt{2\pi}}exp\left[-\frac{1}{2}\left(\frac{\eta-y_0}{u_0}\right)^2\right]$$

$Y=\eta$ 时 Y_{m} 的概率密度函数 $h\left(\eta_{\mathrm{m}}|\eta\right)$ 为：

$$h\left(\eta_{\mathrm{m}}|\eta\right)=\varphi\left(\eta_{\mathrm{m}};\eta,u_{\mathrm{m}}^2\right)=\frac{1}{u_{\mathrm{m}}\sqrt{2\pi}}exp\left[-\frac{1}{2}\left(\frac{\eta_{\mathrm{m}}-\eta}{u_{\mathrm{m}}}\right)^2\right]$$

则联合概率密度函数为：

$$f\left(\eta,\eta_{\mathrm{m}}\right)=g_0(\eta)h\left(\eta_{\mathrm{m}}|\eta\right)=\frac{1}{2\pi u_0u_{\mathrm{m}}}exp\left\{-\frac{1}{2}\left[\left(\frac{\eta-y_0}{u_0}\right)^2+\left(\frac{\eta_{\mathrm{m}}-\eta}{u_{\mathrm{m}}}\right)^2\right]\right\}$$

令 $v=\left(\eta_{\mathrm{m}}-\eta\right)/u_{\mathrm{m}}$，$dv=d\eta_{\mathrm{m}}/u_{\mathrm{m}}$，$z=\left(\eta-y_0\right)/u_0$，$dz=d\eta/u_0$

对应的全局消费者风险为：

$$R_{\mathrm{C}}=\int_{-\infty}^{\frac{T_{\mathrm{L}}-y_0}{u_0}}F(z)\varphi_0(z)dz+\int_{\frac{T_{\mathrm{U}}-y_0}{u_0}}^{\infty}F(z)\varphi_0(z)dz$$

全局生产商风险为：

$$R_{\mathrm{P}}=\int_{\frac{T_{\mathrm{L}}-y_0}{u_0}}^{\frac{T_{\mathrm{U}}-y_0}{u_0}}\left(1-F(z)\right)\varphi_0(z)dz$$

其中：

$$\varphi_0(z)=\left(1/\sqrt{2\pi}\right)exp\left(-z^2/2\right)$$

$$F(z)=\Phi\left(\frac{A_{\mathrm{U}}-y_0-u_0z}{u_{\mathrm{m}}}\right)-\Phi\left(\frac{A_{\mathrm{L}}-y_0-u_0z}{u_{\mathrm{m}}}\right)=\Phi(4.5-3z)-\Phi(-4.5-3z)$$

通过计算得到：

$R_{\mathrm{C}}=\int_{-\infty}^{-1.667}F(z)\varphi_0(z)dz+\int_{1.667}^{\infty}F(z)\varphi_0(z)dz=0.01=1\%$，且

$R_{\mathrm{P}}=\int_{-1.667}^{1.667}\left[1-F(z)\right]\varphi_0(z)dz=0.07=7\%$

将这个过程用于判定该生产线的 100 个电阻样品的合格与否情况。根据该生产线的生产能力，真实情况为 90 个电阻合格，10 个电阻不合格，但需要通过产品检验程序判定哪些为合格电阻：

- 90 个合格电阻经过测量，将合格判为不合格的全局生产商风险为 7%，即有 7 个判断为不合格淘汰，有 83 个判断为合格并出厂；
- 10 个不合格电阻经过测量，将不合格判为合格的全局消费者风险为 1%，即有 9 个判断为不合格淘汰，1 个判断为合格出厂；
- 那么最终有 84 个电阻判断为合格出厂，其中错判率为 $(84-83)/84\approx 1\%$。这就是检验程序的目的，将原本不合格件出厂的概率从 10%降低为 1%。
- 剩余的电阻，其中有 $7/16\approx 44\%$ 为误判，这是为了降低不合格件出厂概率所必须付出的代价。

7.5.4 设置接受限（合格限）

上面生产电阻的例子中，根据已知的接受限 A_{L} 和 A_{U} 计算全局消费者风险和生产商风险。而多数实际应用是基于成本分析选定希望的风险水平，再计算出合理的接受限值，确保能达到所需的风险水平。下面通过实例说明这种方法。

例 15 滚珠轴承生产

某一生产球轴承的生产商，合格的产品要求每个球轴承的径向运动误差（可测量属性）不大于 $2\mu\mathrm{m}$。球轴承的径向运动是垂直于球轴承旋转轴方向（轴向）的运动，是产品不需要的运动。径向运动误差为零的球轴承是完美的球轴承，但事实上球轴承总是有径向运动误差的。

为了评估生产线的性能，使用测量不确定度可以忽略不计的高精度测量设备对大量球轴承的径向运动误差进行测量。例如，测量得到平均径向运动误差 $\tilde{y}=1\mu\mathrm{m}$，对应标准偏差 $s=0.5\mu\mathrm{m}$。出厂前，对所有轴承进行符合性判定，使用校准过的检测设备测量径向误差。测量系统的概率密度函数为 $\varphi\left(\eta_{\mathrm{m}};\eta,u_{\mathrm{m}}^{2}\right)$ 对应标准不确定度 $u_{\mathrm{m}}=0.25\mu\mathrm{m}$。出于经济原因的考虑，不合格产品出厂的概率（全局消费者风险）不得大于 0.1%，那么，如何选择接受区间才能满足这个要求呢？

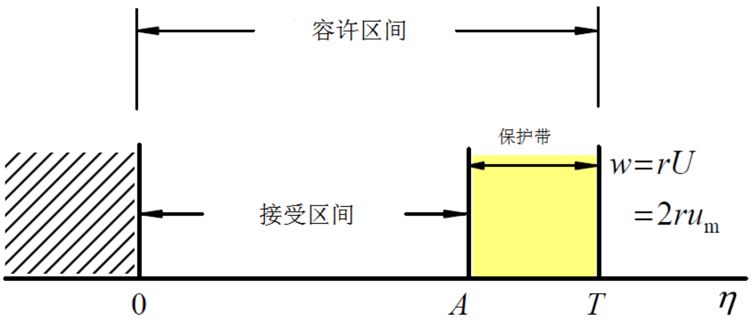

图 15 球轴承轴的容许区间和接受区间。合格的径向运动误差Y 的值位于区间$0<\eta<T$ 内，接受限A和容许限T之间的保护带长度$w=rU=2ru_{\mathrm{m}}$。

如图 15 所示，合格的径向运动误差Y 的值位于区间$0<\eta<T$ 内。由于径向运动误差恒为正，且测得值接近于 0，因此径向运动误差先验概率密度函数模型为 gamma 概率密度分布。基于样本测量，先验概率密度函数的期望值和标准不确定度分别为$y_0=\tilde{y}=1\mu\mathrm{m}$，$u_0=s=0.5\mu\mathrm{m}$。

正参数为α、λ的 gamma 概率密度函数定义为：

$$\mathrm{gamma}(\eta;\alpha,\lambda)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\eta^{\alpha-1}e^{-\lambda\eta},\eta\geq 0$$

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

期望$E(Y)=y_0=\dfrac{\alpha}{\lambda}$，方差$V(Y)=u_0^2=\dfrac{\alpha}{\lambda^2}$，且$\eta=\dfrac{\alpha-1}{\lambda}$的分布概率最大。

由此计算参数α、λ为：

$$\alpha=\frac{y_0^2}{u_0^2}=\frac{1^2}{(0.5)^2}=4,\quad \lambda=\frac{y_0}{u_0^2}=\frac{1}{(0.5)^2}=4$$

将各参数代入，得到径向运动误差Y的先验概率密度函数为：

$$g_0(\eta)=\mathrm{gamma}(\eta;4,4)=\frac{128}{3}\eta^3e^{-4\eta},\eta\geq 0$$

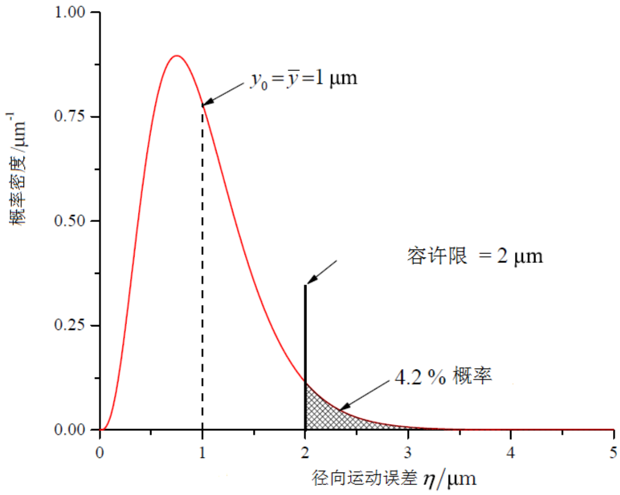

图 16 对应上式的 gamma 先验概率密度函数，由球轴承样品的被测径向运动误差的频率分布确定。容许区间为$0\le\eta\le 2\mu\mathrm{m}$。该先验分布的期望值$y_0=1\mu\mathrm{m}$，对应标准不确定度$u_0=0.5\mu\mathrm{m}$。由于该分布函数不对称，$Y$ 最可能的值不等于先验期望值y_0，而等于$0.75\mu\mathrm{m}$。

如图 16 的概率密度函数所示，生产线上随机取一个球轴承，其径向运动误差大于$2\mu\mathrm{m}$的概率为图中阴影部分，计算得到这个概率为：

$$\bar{p}_\mathrm{c}=\int_2^\infty \mathrm{gamma}(\eta;4,4)\,d\eta=0.042$$

也就是说，如果所有生产的球轴承不测量而直接出厂，那么其中将有 4.2%为不合格件。因此需要设计检验程序对生产的球轴承进行检验，且选取的接受限需要将全局消费者风险降低到 0.1%以下。对于图 15 所示的符合性判定，容许区间为$0<Y<T$，接受区间为$0<Y_\mathrm{m}<A$。采用类似的步骤，通过式（30）和（31）计算出全局消费者风险和全局生产商风险为：

$$R_\mathrm{C}=\int_T^\infty\int_0^A g_0(\eta)h(\eta_\mathrm{m}|\eta)\,d\eta_\mathrm{m}d\eta,\quad R_\mathrm{P}=\int_0^T\int_A^\infty g_0(\eta)h(\eta_\mathrm{m}|\eta)\,d\eta_\mathrm{m}d\eta$$

测量系统的特征可用正态概率密度函数$h(\eta_\mathrm{m}|\eta)=\varphi(\eta_\mathrm{m};\eta,u_\mathrm{m}^2)$描述，令$z=(\eta_\mathrm{m}-\eta)/u_\mathrm{m}$，$dz=d\eta_\mathrm{m}/u_\mathrm{m}$，对 z 积分，表达式变为：

$$R_C=\int_T^\infty\left[\Phi\left(\frac{A-\eta}{u_\mathrm{m}}\right)-\Phi\left(-\frac{\eta}{u_\mathrm{m}}\right)\right]g_0(\eta)\,d\eta,\quad R_P=\int_0^T\left[1-\Phi\left(\frac{A-\eta}{u_\mathrm{m}}\right)\right]g_0(\eta)\,d\eta$$

根据图 15 可见$A=T-2ru_\mathrm{m}$，这里$T=2\mu\mathrm{m}$，$u_\mathrm{m}=0.25\mu\mathrm{m}$。$g_0(\eta)$设置为式（37）中的 gamma 概率密度函数，得到结果：

$$R_C(r)=\frac{128}{3}\int_2^{\infty}\left[\Phi(8-2r-4\eta)-\Phi(-4\eta)\right]\eta^3 e^{-4\eta}d\eta$$
$$R_P(r)=\frac{128}{3}\int_0^{2}\left[1-\Phi(8-2r-4\eta)\right]\eta^3 e^{-4\eta}d\eta$$

根据上面的函数式 $R_C(r)$ 可得出 r 与 R_C 的曲线图，如图 17，展示了 $-1<r<1$ 时的全局消费者风险曲线。

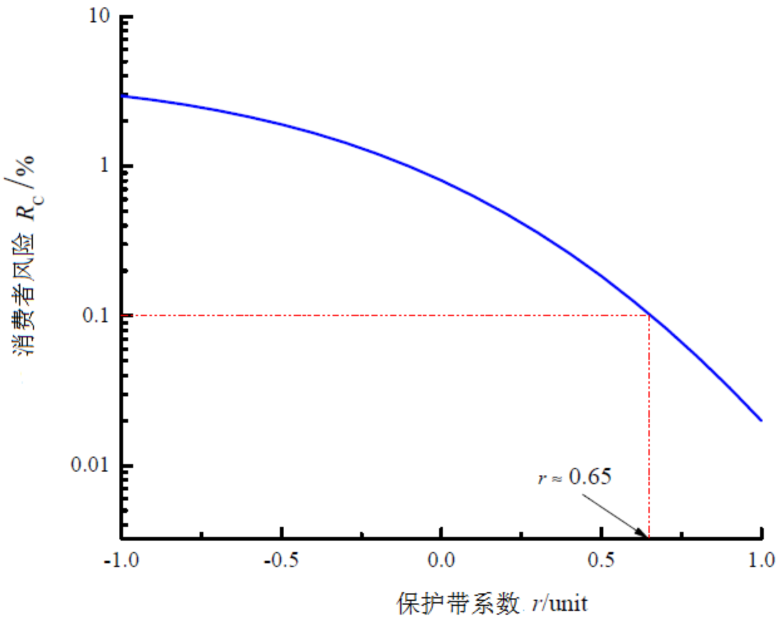

图 17 保护带系数 r 与全局消费者风险的关系。当 $r\approx 0.65$ 时，接受限 $A=T-2\times(0.65)u_m=1.7\mu m$，可实现 $R_C=0.1\%$。

图 17 中，$A<T$（保护的接受限）时 r 为正，$A>T$ 时 r 为负。$r=0$ 时没有保护带（$A=T$），这种情况就是简单接受或风险共担的判定规则。期望的风险水平 $R_C=0.1\%$ 对应保护带乘数为 $r\approx 0.65$，此时接受限选择为：

$$A=T-2ru_m=(2-2\times 0.65\times 0.25)\mu m\approx 1.7\mu m$$

也就是选择 $A=1.7\mu m$ 可以解决球轴承的判定问题。

在二元决策的判定规则中，降低消费者风险的行为往往会增加生产商风险。在球轴承的例子中，根据函数式 $R_C(r)$ 和 $R_P(r)$ 得到二者的关系如图 18 所示。保护带系数 $r\approx 0.65$ 时对应全局生产商风险 R_P 约为 7.5%，这意味着每 1000 个滚珠轴承中有 75 个合格产品判断为不合格产品而不能出厂，这使得生产商损失了这 75 个合格品带来的收益。

有保护带的接受增加了合格产品的废弃率，但是也降低了不合格产品的出厂

率。在实际应用中，生产商需要在图 18 所示曲线上找到平衡风险的工况点，以获得最佳收益。这个工况点的选择是在对符合性判定问题进行经济分析的基础上做出的商业决策。

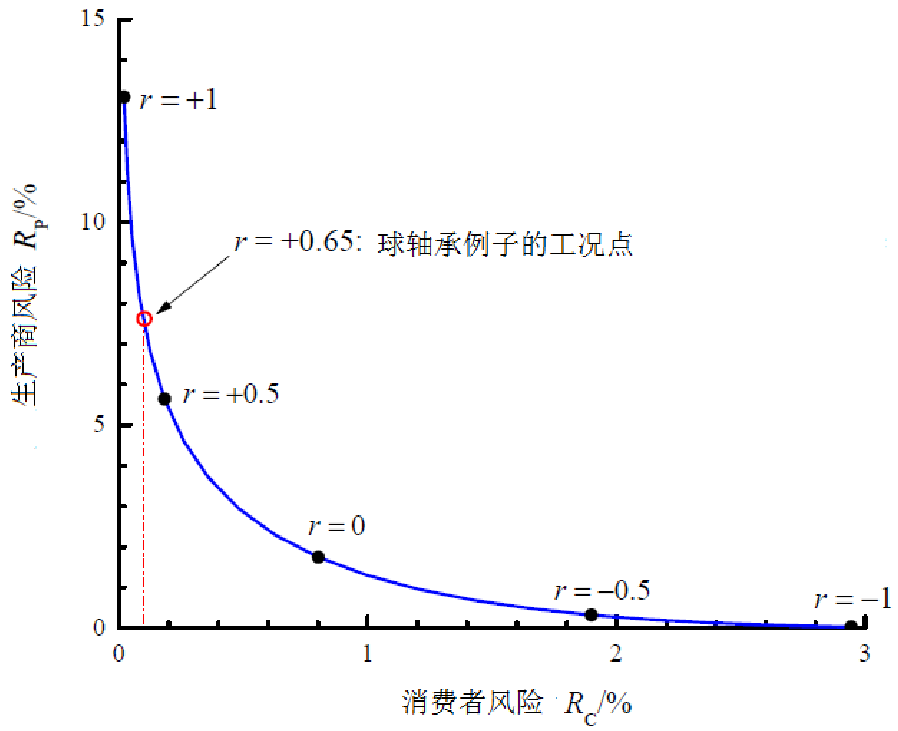

图 18　球轴承例子中全局消费者风险和全局生产商风险之间的关系图。曲线中每一个点均对应一个 r 值（即保护带系数），图中标出了几个典型值。增加保护带的长度使接受限向容许区间内移动（增大 r），会增加将合格轴承判定为不合格轴承的风险。最优判定规则的选择需要建立在经济分析的基础上。图中红色圆圈为本例采用的工况点。

7.5.5 通用图解法

对于已知容许限 T 、正态先验概率密度函数 $g_0(\eta)=\varphi(\eta;y_0,u_0^2)$ 和正态分布的测量系统概率密度函数 $h(\eta_m|\eta)=\varphi(\eta_m;\eta,u_m^2)$，可用图 19 的图解法来设置接受限。

本节假设被测量 Y 的先验信息很少，其后验概率密度函数完全由测量得到的信息确定，因此测得值 $y\approx\eta_m$ 且对应标准不确定度 $u\approx u_m$ 。

图 19 显示了为 $u_0=T/6$ 时 R_C 和 R_P 的关系。图中 5 条曲线分别对应测量能力指数 $C_m=T/(4u_m)$ 取值从 2 到 10。每一条曲线上的点均对应不同保护带长度参数，从 $w=-U$ 到 $w=U$ ，扩展不确定度 $U=2u$ 。

使用图 19 需要注意以下问题：

- 假设生产程序是中心对称的，因此被测量的先验期望值 y_0 位于容许区间的中心；
- 假设上、下保护带长度参数的绝对值相等（对称的接受区间）；
- R_C 和 R_P 计算的前提为生产过程和测量系统均可用正态概率密度函数描述；
- 还可以在图中增补 5 个测量能力值之外值的曲线；
- 也可以沿着曲线增加保护带的值。

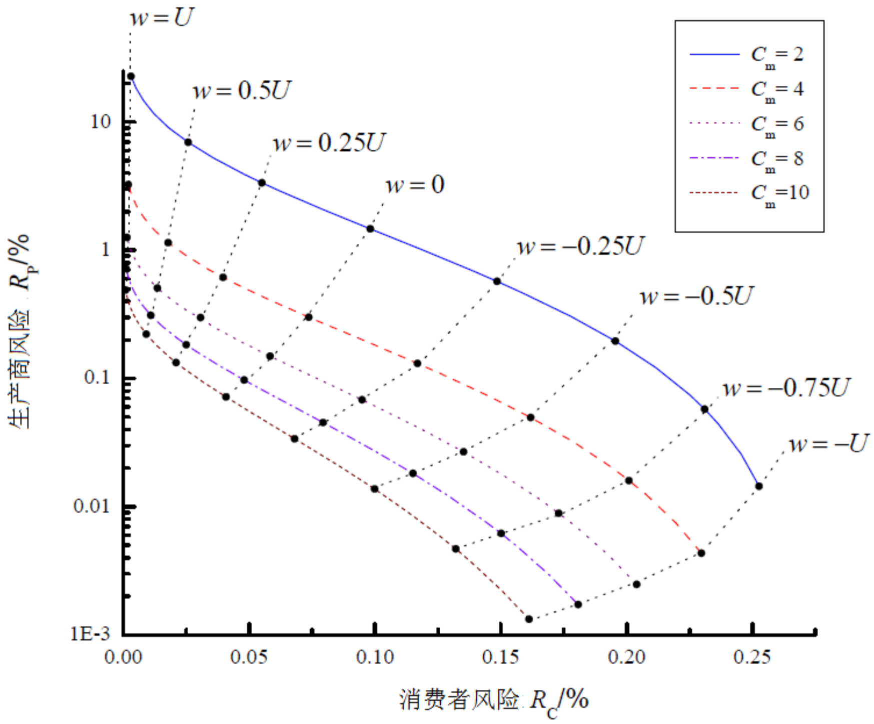

图 19 先验标准不确定度 $u_0 = T/6$ 的二元决策判定规则的全局生产商风险和全局消费者风险的关系。图中 5 条曲线分别对应测量能力指数 $C_m = T/(4u_m)$ 的值从 2 到 10 的情况。黑点对应保护带长度参数 w 从 U 到 $-U$ 的情况（$U = 2u$）。正的 w 对应有保护带的接受，接受限位于容许限内，如图 12 所示。

7.5.6 减小测量不确定度

减小符合性判定中测量结果的不确定度（提高测量能力指数），将会降低做出错误判定结论的概率。图 19 中标记各个保护带的虚线显示了这一趋势。

以简单接受的判定规则（ $w=0$ ）为例，若测量不确定度满足 $C_m = T/(4u_m) = 2$，则消费者风险 $R_C \approx 0.1\%$，对应的生产商风险 $R_P \approx 1.5\%$。

如果采用更先进的测量系统，测量能力指数 $C_{\mathrm{m}}=10$，此时的消费者风险降为 $R_{\mathrm{C}}\approx 0.04\%$，对应的生产商风险 $R_{\mathrm{P}}\approx 0.07\%$。减小测量不确定度在经济上是否可行降低，取决于改进测量系统投入的成本和降低错判风险所节约成本之间的权衡。

改进生产程序（降低先验标准不确定度 u_0）也会同样的降低消费者和生产商的风险，其经济性同样取决于成本/收益分析。