

剔除计量检测中异常数据的方法研究

胡晓萍

(龙口市检验检测中心 山东 龙口 265701)

摘要: 计量检测是一项非常重要的工作, 计量检测结果将直接影响到计量检测工作的整体效果。只有保证计量检测数据的准确性, 才能为后期各项工作开展提供非常重要的参考数据。通常情况下, 因为计量检测环节涉及的内容较多, 所以会产生比较多的数据信息, 但是, 因为受到多方面因素的影响, 所以经常会产生一些异常数据。因此, 在开展相应的计量检测工作时, 一定要对这些异常数据进行全面剔除, 才能在更大程度上保障计量检测数据的质量。本文主要针对计量检测工作的实际情况, 对可能出现的异常数据进行探究, 提出一些合理的剔除方法, 希望能在一定程度上提升计量检测工作的效率与检测质量。

关键词: 计量检测; 异常数据; 剔除方法

0 引言

计量检测工作主要是对被计量对象中所涉及的各项参数进行准确检测。在计量检测过程中会产生非常多的测试数据, 对这些数据进行统计之后, 可以获取准确的数据信息, 然后将异常的数据进行剔除, 从而实现对检测对象的合理应用。通常情况下, 计量检测对于数据准确性有着非常高的要求, 因此, 需要对计量检测中出现的异常数据进行深入分析, 及时找出异常数据产生的具体原因, 根据实际情况采取有效的剔除措施, 从而实现对异常数据的有效剔除。

1 计量检测的重要作用

对于计量检测工作而言, 为了更好地满足使用要求, 需要为生产过程中比较常见的检测器提供数据, 并加强对数据的统计和分析工作。结合检测器所提供的相关信息, 对产品使用价值进行准确判定, 同时对计量测试中所产生的异常数据进行有效处理。通常情况下, 在开展计量检测工作时, 对于计量检测数据的准确性有着非常高的要求, 因此, 在实际的计量检测工作中, 一旦出现异常数据, 将会对计量检测结果的准确性带来非常严重的负面影响, 也会给后续各项工作的正常开展带来非常严重的阻碍。因此, 加强计量检测工作非常重要, 因为这与生产过程的安全性和最终的生产质量都有着非常密切的联系。

好的计量检测数据可以为每一个生产阶段提供合理的引导, 有助于生产流程的优化与完善。比如: 在对温度开关开展计量检测工作时, 当机电运行状态一切正常时, 对于双金属片而言, 可以判定状态比较自由; 在对电器温度进行检测时, 当满足设计方案的要求后, 就可

以判定金属片处于运行的状态; 可以在各种不同的电机电磁炉中进行使用, 同时可以对变压器以及电气设备起到非常重要的保护作用。在对感温开关开展相应的计量检测工作时, 在完成相应的计量检测工作基础上, 将其中存在的异常数据剔除之后, 对计量检测数据进行深入、全方位地分析, 看温度开关是否能满足实际的工作需求, 然后看是否可以发挥出一定的保护效果。通过采取以上方式, 对感温开关而言, 在整个生产过程中能够起到非常重要的保护效果, 并提升生产过程的精准性。因此, 对计量检测数据结果准确性提出严格要求, 是企业生产经营活动可以顺利开展的重要基础条件。同样, 如果在计量检测数据当中出现了异常数据, 不但会对影响到最终计量检测结果准确性, 同时也会影响到企业生产工作质量。因此, 在开展相应的计量检测工作时, 一定要对其中存在的异常数据进行全面剔除, 这也是整个检测工作中的重点内容。在对异常数据准确剔除的基础上, 不但能够促进数据精准性实现明显提升, 同时也能促进各项生产工作可以实现顺利开展。

2 计量检测中异常数据产生的原因分析

结合实际情况可以了解到, 在计量检测工作开展中, 不但涉及的工作量非常大, 同时还应用到了非常多的计量检测设备。然而, 最终所获取到的计量检测结果经常会受到一些外界因素的影响, 从而导致在计量检测结果中, 经常会面临一些异常数据的产生, 而导致该问题产生的原因主要体现在了以下几个方面。

2.1 因外部环境因素产生的计量检测异常现象

针对计量检测结果中存在的异常数据, 一项非常重要的原因是受到外部环境的影响。通常情况下, 在开展计量检测工作时, 对于所处检测环境温度以及湿度等方

面有着非常严格的要求,如果温度或者是湿度超出相应的标准范围时,就很容易导致计量检测结果异常数据问题的产生,从而将会对最终的计量检测结果产生非常严重的负面影响。结合相关的调查结果可以了解到,在实施相应的计量检测工作时,一般需要将温度控制在 20℃ 左右,将湿度控制在 70% 左右,只有温度与湿度处于这一标准范围内,才有利于保证最终计量检测数据质量。

2.2 计量检测工作人员专业性不足

在计量检测工作中,计量检测工作人员在其中发挥着非常重要的执行作用,其自身工作能力与最终计量检测工作质量之间有着非常密切的联系。但是,目前在具体的计量检测工作中,针对不同的计量检测工作人员而言,在检测方法以及工作能力等方面都具有非常明显的差异,存在很多计量检测工作人员专业能力达不到相应的标准要求,从而在对各类计量检测工具进行使用时,其操作流程不能满足一定的规范性,最终导致计量检测结果出现异常数据。另外,在开展相应的计量检测工作时,计量检测工作人员对于计量检测环境控制也明显不同,从而导致不同人员所获取到的计量检测结果存在一定的差异,进而增加了计量检测结果异常数据发生概率。同时,在保证工作环境以及额定电压正常的情况下,即便工作人员已经掌握了相应的技术知识与操作方法,也有可能就会导致异常数据的产生,但是可以在很大程度上降低数据发生概率。

2.3 计量检测设备故障问题

在实施相应的计量检测工作时,因为受到一些因素的影响,所以不可避免会面临故障问题的出现,因为计量检测设备属于整个计量检测工作中非常重要的辅助设备,同时也是能对检测对象进行有效统计。因此,在计量检测工作中,如果计量检测设备出现了故障问题,同时计量检测工作人员没有及时对问题进行解决,那么将会直接导致异常数据问题的出现,从而影响到最终计量检测结果准确性与真实性。

2.4 受额定电压的影响

在计量仪器正常使用时,一定额定电压都会处于一种稳定的状态,在保证不变的基础上,计量仪器才能准确地完成数据传输工作。同样,如果是在不能保证电压稳定性的情况下,会影响到数据输出质量,从而在计量检测工作中,导致异常数据问题的产生。异常数据的出现除受到以上几点原因影响之外,还会受到产品本身的影响。在排除以上几点因素之外,对于整个过程中出现的异常数据,应该进行深入分析,同时与产品数据进行融合,从而做出更加准确性的判定。

3 异常值判断

在对异常值进行判断的过程中,其判断准则主要涉

及 3α 准则、格拉布斯准则、肖维勒准则以及 t 检验法等几个方面。在对相关准则与公式全面应用的基础上,可以对计量测试数据开展相应的计算分析工作,从而对异常值进行准确认定。其次,了解产品在使用过程中所能发挥的作用与要求,在此基础上对异常值进行判定,最终完成对异常值的剔除工作。在计量检测工作开展中,针对其中存在的异常数据,需要对异常值做好准确地判定工作,同时分析计量检测中的异常数据是否需要进行剔除。在对产品实际使用性能综合考虑的基础上,对异常值进行准确判断,从而可以将计量检测中的异常数据进行剔除,从而保证最终计量检测结果的准确性,之后,对产品的科学性以及实用性等作出准确判定。

以上几种判定标准在应用时存在一定的差异,在实际的判断过程中,通常会使用到置信概率。置信概率主要指的是置信区间要高出随机变量的概率,使用 X 来表示置信区间测量数据取值范围。

3.1 3α 准则

将最终测量的数据代入到表达式当中,如果与判定值一致的情况下,可以将测量结果判定为异常值,从而完成对异常值的剔除工作。

3.2 t 检验法

对于数据集总体而言,主要指的是检验方法为可疑数值 x_i 之外的数值,将该数据集假设为一种正态分布模式,将可疑数值 x_i 当做样本容量特殊总值,然后用“1”进行表示。对两个总体之间进行对比分析,最终结果显示隶属于相同的总体,那么最终得出的结论为,两者之间没有存在明显的不同,因此,最终的统计量计算公式为:

$$K = \frac{x_i - \bar{x}}{s}$$

结合式中的内容进行分析,其中可疑值之外的样本数据算术平均值使用 $\bar{x}$ 来表示,可疑值之外的样本数据标准差使用 s 来表示。经过系统地对对比分析之后,如果得出的样本数据算术平均值与样本数据标准差超出了 t 检验值范围,那么 x_i 发生的概率就会非常低,可以将其看作一项异常值,然后将该值进行剔除。或者也可以不看作异常值,因为发生的概率非常低,所以即便出现之后,也不会对测量结果产生太大的影响。

3.3 肖维勒准则

将公式与相互独立测量的数据进行有效结合,当最终所测得的结果满足公式标准之后,可以将最终的数值判定为异常值,从而对其进行有效剔除。

3.4 格布拉斯准则

在公式当中,将最终计量检测结果用 x_d 表示,测量数据平均值用 $\bar{x}$ 表示。白塞尔公式当中的标准差用 S

表示,通过查表之后,可以得到 $G(a,n)$ 。

4 计量检测中异常数据剔除方法

通常情况下,在计量检测工作开展过程中,异常数据剔除属于其中非常重要的内容,与最终计量检测结果准确性之间有着非常密切的联系,在整个计量检测工作中占据着非常重要的核心地位。结合实际情况可以了解到,目前应用比较广泛的计量检测异常数据剔除方法,还涉及了其他一些有效的剔除法。在对以上几种剔除方法应用的基础上,能够对计量检测结果当中存在的异常数据进行准确剔除,从而为计量检测结果准确性提供非常重要的保障作用。

第一,计量检测中异常数据的 3α 数据剔除法。在开展相应的计量检查工作时,当发现其中存在异常数据时, 3α 数据剔除法有着非常普遍的应用,将计量检测数据带入到 $|x_d - \bar{x}| > 3\alpha$, 在具体的计量过程中,如果面临异常数据,就需要在第一时间采取有效的方法,完成对异常数据的剔除工作,通过这种方式来保证最终数据的准确性。

第二,计量检测中异常数据的 t 值检测数据剔除法。在对这种剔除方式进行应用的过程中,首先在其中选出一个参考数值,然后对数据进行有效收集。在整个过程中应用正态分布的方法,对于选择出的参照数据,作为相应的样本,在此基础上完成对比分析工作,从而可以得出计量检测数据总体,确定出一个异常数据。通常情况下,计量检测样本数据平均值、可疑数据以外的数据部分与样本数据形成的数据值表超出 t 值检测范围之后,那么就可以判定在计量检测数据中存在异常现象,从而对其进行有效剔除。

第三,肖维勒数据剔除法。该剔除方法在实际的应用过程中,主要采取独立性的方法,对其中所获取到的数据进行对比,在保证能够满足相应的标准要求之后,在对异常数据进行准确判定,进而完成相应的剔除过程。

第四,格布拉斯数据剔除法。在计量检测异常数据剔除方法中,该方法目前也实现了比较广泛的应用,并且受到了相关人员的高度重视。主要是以 x_d 计量检测的测量数据,对于最终的计量检测数据平均值,以 $\bar{x}$ 进行表示,计量检测数据的标准差通过 S 进行表示。

5 综合分析

对计量检测异常数据进行准确判定之后,同时还需要对异常值做好相应的分析工作。对异常值产生的具体原因进行准确判定,如果是因为仪器操作设备或者

是操作不当而引起的异常值,可以对异常数据进行排除。如果是因为产品本身所引起的问题,需要对异常值进行深入分析,同时站在整体角度上对问题进行考量,从而判定异常值是否需要进行排除。比如,如果一个产品对于精准度方面要求比较高,对该产品计量测试中产生的数据进行分析,其数据分别为:1.2326, 1.2325, 1.2331, 1.2329, 1.2112, 1.2335, 1.2326, 1.2324, 1.2330, 1.2328, 1.2331, 1.2329, 1.2326, 1.2324, 1.2330。

结合公式异常值进行判定,可以分析出 1.2112 为异常数据。对其进行综合分析,该精准度要求较高的产品是电气设备开关中的连接器,在使用过程中起到了非常重要的连接作用。因为受到连接器工作环境以及工作要求等多方面因素的影响,所以需要对其进行综合考量。最终计算出异常值为 1.112, 所以需要对该数据进行剔除。对于其他有效的计量检测数据进行对比分析,同时针对连接器应用过程中的科学性以及实用性等进行准确评价。

6 结语

综上所述,基于现代社会经济发展背景下,在产品生产过程中,对于质量以及精准度提出了更高的要求,而通过开展计量检测工作,不但可以提升工作效率,同时还能对产品检测结果进行准确判定,保证产品在投入使用之后,可以将本身的价值作用充分发挥出来。通常情况下,在计量检测工作中经常会存在一些异常数据,而导致异常数据发生的原因非常多,这就需要对异常数据产生的具体原因进行深入分析,在此基础上结合公式计算结果,将异常数据进行准确剔除,从而提升计量检测结果的准确性。

参考文献:

- [1] 周雷. 剔除计量检测中异常数据的方法初探[J]. 质量与市场, 2020(12): 79-81.
- [2] 卢云忠. 浅析如何剔除计量检测中的异常数据[J]. 科学技术创新, 2020(08): 165-166.
- [3] 黄志恒. 计量检测中异常数据剔除的有效方法[J]. 科技风, 2019(33): 23.
- [4] 王莹, 刘佳, 王欣, 等. 计量检测中异常数据剔除的有效方法[J]. 电子世界, 2016(24): 93+95.
- [5] 卢钧. 赵道文. 装备计量测试实验室信息管理系统研究与设计[J]. 计量与测试技术, 2021, 5(S2): 53-55.

作者简介: 胡晓萍(1983.11-), 女, 汉族, 山东龙口人, 本科, 工程师, 研究方向: 计量校准。