

# 属性控制图

## 概述

控制图用于定期监控过程，以确定过程是否在控制限值内。当无法通过连续数据来衡量产品或服务的质量时，通常会收集属性数据，以评估其质量。Minitab “协助”包括两个广泛使用的控制图，通过它们可以根据属性数据监控过程：

- P 控制图：此图在产品或服务为不良品或良品时使用。P 控制图标出子组不良品率。收集到的数据是每个子组中的不良品数，假定遵循比例参数 ( $p$ ) 未知的二项式分布。
- U 控制图：此图在产品或服务有多个缺陷时使用并计算缺陷数。U 控制图标出单位缺陷数。收集到的数据是子组缺陷总数，假定遵循子组缺陷平均数未知的 Poisson 分布。

控制图的控制限值通常在六西格玛项目的控制阶段设置。好的控制图应该具有足够的灵敏度，能够在出现特殊原因时快速发出信号。这种灵敏度可以通过计算发出特殊原因信号所需的子组平均数进行评估。当过程处于控制范围时，好的控制图很少发出“误报”信号。当过程处于控制范围时，误报率可以通过计算被认为“失控”的子组比例来评估。

为了帮助评估控制图的执行效果，“协助报告卡”将自动执行下列数据检查：

- 稳定性
- 子组数
- 子组大小
- 预期变化

在本文中，我们探讨了当这些情况各不相同属性控制图的表现，还介绍了如何建立一套准则来评估这些情况的要求。

我们还介绍了当观测到的数据变化不符合预期变化以及 Minitab 检测到过度离散或过度集中时，推荐使用的 Laney P' 控制图和 U' 控制图。

注意 P 控制图和 U 控制图取决于无法检查或难以检查的其他假设。有关详细信息，请参见附录 A。

# 数据检查

## 稳定性

可以对属性控制图进行四种检验，以评估过程的稳定性。使用这些检验还会提高控制图的灵敏度。不过，一定要确定每种检验的目的和附加值，因为随着更多的检验被添加到控制图中，误报率会增加。

### 目标

我们想要确定四种检验的稳定性，以将属性控制图添加到“协助”中。我们的首要目标是确定检验是否会在失控情况下显著提高灵敏度，而不会显著提高误报率，并确保控制图的简单性和实用性。

### 方法

属性控制图的四种稳定性检验分别对应变量控制图的 1-4 种特殊原因检验。使用足够的子组大小，不良品率（P 控制图）或单位缺陷数（U 控制图）遵循正态分布。其结果是，模拟同样基于正态分布的变量控制图将产生相同的灵敏度和误报率检验结果。因此，我们采用模拟结果并翻阅了变量控制图的相关资料，来评估四种稳定性检验如何影响属性控制图的灵敏度和误报率。此外，我们还评估了与检验相关的特殊原因的普遍性。有关每种检验使用的方法的详细信息，请参见下面的“结果”部分和附录 B。

### 结果

在用于评估属性控制图中的稳定性的四种检验中，我们发现，检验 1 和检验 2 最有用：

#### 检验 1：确定控制限值之外的点

检验 1 发现距离中心线大于 3 个标准差的点。检验 1 是公认的必要检验，用于发现失控情况。误报率仅 0.27%。

#### 检验 2：确定不良品率偏差（P 控制图）或单位缺陷平均数（U 控制图）

检验 2 在连续 9 个点在中心线的同一侧时发出信号。我们进行了模拟，以确定检测到不良品率偏差（P 控制图）或单位缺陷平均数偏差（U 控制图）信号所需的子组数。我们发现，添加检验 2 会显著提高控制图检测不良品率或单位缺陷平均数中的小偏差的灵敏度。当检验 1 和检验 2 一起使用时，检测小偏差所需的子组数将比仅使用检验 1 时少得多。因此，添加检验 2 有助于检测正常的失控情况，提高灵敏度，充分保证误报率略有增加。

### “协助”中不提供检验



#### 检验 3：连续 K 个点，全部递增或全部递减

检验 3 旨在检测不良品率或单位缺陷平均数中的偏差（Davis and Woodall, 1988）。但是，当在使用检验 1 和检验 2 的同时使用检验 3 时，并不会显著提高控制图的灵敏度。因为我们已经根据模拟结果决定使用检验 1 和检验 2，包括检验 3 不会将任何有效值添加到控制图中。

#### 检验 4：连续 K 个点，上下交错

虽然这种模式会在实际工作中发生，但是我们建议您寻找任何不寻常的趋势或模式，而不是检验一个特定的模式。

因此，“协助”只使用检验 1 和检验 2 来检查属性控制图中的稳定性并在“报告卡”中显示以下状态指标：

状态	条件
	控制图上无检验 1 和检验 2 错误。
	如果上述条件不成立。

## 子组数

如果没有已知的控制限值，则必须使用数据估计这些值。为了获得精确的限值估计，您必须拥有足够的数量。如果数据量不足，则控制限值可能会因抽样的可变性偏离“真”限值。为了提高限值的精准度，可以增加子组数。

### 目标

我们研究了获得精确的 P 控制图和 U 控制图控制限值所需的子组数。我们的目的是确定为确保检验 1 的误报率不超过 95% 置信度的 2% 而所需的子组数。我们没有评估子组数对中心线（检验 2）的影响，因为中心线的估计值比控制限值的估计值更精确。

### 方法

在提供足够的子组大小且抽样可变性不出错的前提下，超出控制上限的点比例为 0.135%。为了确定子组数是否足够，我们按照 Trietsch (1999) 中概述的方法，通过高于控制上限的点不超过 95% 置信度的 1% 来确保误报率。由于控制限值的对称性，这种方法会导致检验 1 的误报率为 2%。有关详细信息，请参见附录 C。

### 结果

#### P 控制图

下面显示为了确保检验 1 的误报率不超过 2%，对于各个子组大小 (n) 和比例 ( $\bar{p}$ )，P 控制图所需的子组数。

子组大小 (n)	$\bar{p}$				
	0.001	0.005	0.01	0.05	0.1
10	1881	421	228	60	35
50	425	109	64	23	16
100	232	65	41	17	13



	$\bar{p}$				
150	165	49	32	14	11
200	131	41	27	13	10
500	65	24	18	10	9

### U 控制图

下面显示为了确保检验 1 的误报率不超过 2%，对于每个子组的缺陷平均数的各个给定值 ( $\bar{c}$ )，U 控制图所需的子组数 (m)。

$\bar{c}$	0.1	0.3	0.5	0.7	1.0	3.0	5.0	10.0	30.0	50.0
子组数	232	95	65	52	41	22	18	14	10	9

基于这些结果，检查属性控制图中的子组数时，“协助报告卡”会显示以下状态指标：

状态	条件
	子组数 $\geq$ 所需子组数。 子组数足够大，从而确保检验 1 的误报率不超过 95% 置信度的 2%。
	如果上述条件不成立。

## 子组大小

正态分布可用于近似估计 P 控制图中的不良品率分布 ( $\hat{p}$ ) 和 U 控制图中的单位缺陷数分布 ( $\hat{u}$ )。当子组大小增加时，这种近似估计的精确度会随之提高。因为每个控制图中所用的检验标准基于正态分布，增加子组大小可获得更好的正态近似估计，从而提升控制图准确识别失控情况的能力并降低误报率。当不良品率或单位缺陷数较低时，需要较大的子组，才能确保结果准确。

### 目标

Minitab 研究了为确保正态近似估计足以获得精确的 P 控制图和 U 控制图测量结果而所需的子组大小。

### 方法



我们进行了模拟，以评估 P 控制图的各种子组大小和不同比例 (p) 的误报率以及 U 控制图的子组缺陷平均数 ( $\bar{c}$ )。为了确定子组大小是否足够大，以获得足够的正态近似估计，从而获得足够低的误报率，我们将这些结果与正态假设下的预期误报率（检验 1 为 0.27%，检验 2 为 0.39%）进行了比较。有关详细信息，请参见附录 D。

## 结果

### P 控制图

我们的研究表明，P 控制图所需的子集大小取决于不良品率 ( $p$ )。 $p$  的值越小，所需的子组大小 ( $n$ ) 越大。当产品  $np$  值大于或等于 0.5 时，检验 1 和检验 2 的组合误报率低于 2.5% 左右。但是，当产品  $np$  值小于 0.5 时，检验 1 和检验 2 的组合误报率会更高，正好达到 10% 以上的水平。因此，基于这一标准，当值  $np \geq 0.5$  时，P 控制图的功效足够。



检查 P 控制图的子组大小时，“协助报告卡”会显示以下状态指标：

状态	条件
	$n_i \bar{p} \geq 0.5$ 适用于所有 $i$ 其中 $n_i$ = 第 $i$ 个子组的子组大小 $\bar{p}$ = 不良品率的平均值
	如果上述条件不成立。

### U 控制图

我们的研究表明，U 控制图所需的子组大小取决于子组缺陷数 ( $c$ )，相当于子组大小 ( $n$ ) 乘以单位缺陷数 ( $u$ )。当缺陷数  $c$  较小时，误报率最高。当  $c = nu$  大于或等于 0.5 时，检验 1 和检验 2 的组合误报率小于 2.5% 左右。然而，当  $c$  值小于 0.5 时，检验 1 和检验 2 的组合误报率要高得多，正好达到 10% 以上的水平。因此，基于这一标准，当值  $c = nu \geq 0.5$  时，U 控制图的功效足够。

检查 U 控制图的子组大小时，“协助报告卡”会显示以下状态指标：

状态	条件
	$n_i \bar{u} \geq 0.5$ 适用于所有 $i$ 其中 $n_i$ = 第 $i$ 个子组的子组大小 $\bar{u}$ = 单位缺陷平均数
	如果上述条件不成立。

## 预期变化

传统的 P 控制图和 U 控制图假设数据变化遵循适用于不良品的二项式分布或适用于缺陷的 Poisson 分布。这些控制图还假设不良品率或缺陷率始终保持不变。当数据变化大于或小于预期变化时，您的数据可能会过度离散或过度集中，控制图可能无法达到预期效果。

过度离散

当数据变化超出预期变化时，就会出现过度离散。通常情况下，随着时间的推移，受外部噪音因素等非特殊原因的影响，不良品率或缺陷率会发生某种程度的变化。在这些控制图的大多数应用中，子组统计数据的采样变化足够大，导致相关不良品率或缺陷率变化并不明显。然而，当子组大小增大时，采样变化会变得越来越小，在某些时候，相关缺陷率变化会变得比采样变化还大。最终将导致控制图具有极窄的控制限值和极高的误报率。

### 过度集中

当数据变化不如预期变化时会出现过度集中。当相邻的子组彼此相关时会发生过度集中，也称为“自相关”。例如，当某种工具磨损时，缺陷数就会增加。子组中的缺陷数增加时，子组相似的几率将更高。当数据表现出过度集中时，传统的 P 控制图或 U 控制图上的控制限值可能会过宽。如果控制限值过宽，控制图很少发出信号，这意味着您可以忽略特殊原因变化，误以为是常见原因变化。

如果过度离散或过度集中严重到一定程度，Minitab 建议使用 Laney P' 或 U' 控制图。有关详细信息，请参见下面的 Laney P' 或 U' 控制图。

## 目标

我们想要确定一种方法，以便检测数据中是否存在过度离散和过度集中情况。

## 方法

我们查阅了相关资料，发现了过度离散和过度集中的几种检测方法。我们选择 Jones and Govindaraju (2001) 中发现的一种诊断方法。如果数据来自适用于不良品数据的二项式分布或适用于缺陷数据的 Poisson 分布，这种方法使用概率图来确定预期变化量。然后，我们对预期变化量和观测到的变化量进行了比较。有关详细的诊断方法，请参见附录 E。

在检查过度离散时，Minitab 还确定有多少点位于传统的 P 控制图和 U 控制图的控制限值之外。因为过度离散问题会导致高误报率，如果只有一小部分点失控，则过度离散不太可能成为一个问题。

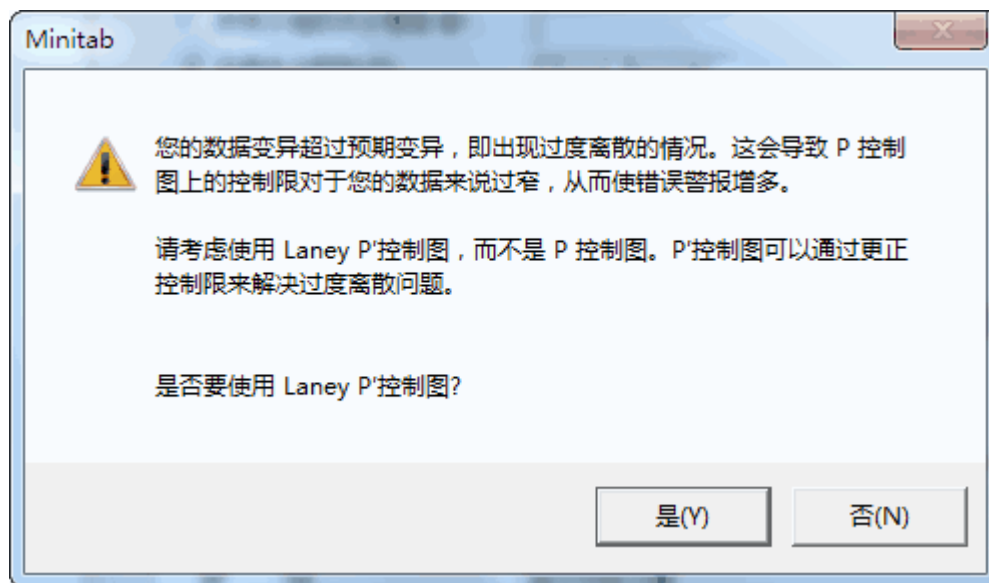
## 结果

当用户在 P 或 U 控制图的对话框中选择“确定”之后，Minitab 将对过度离散和过度集中执行诊断检查。

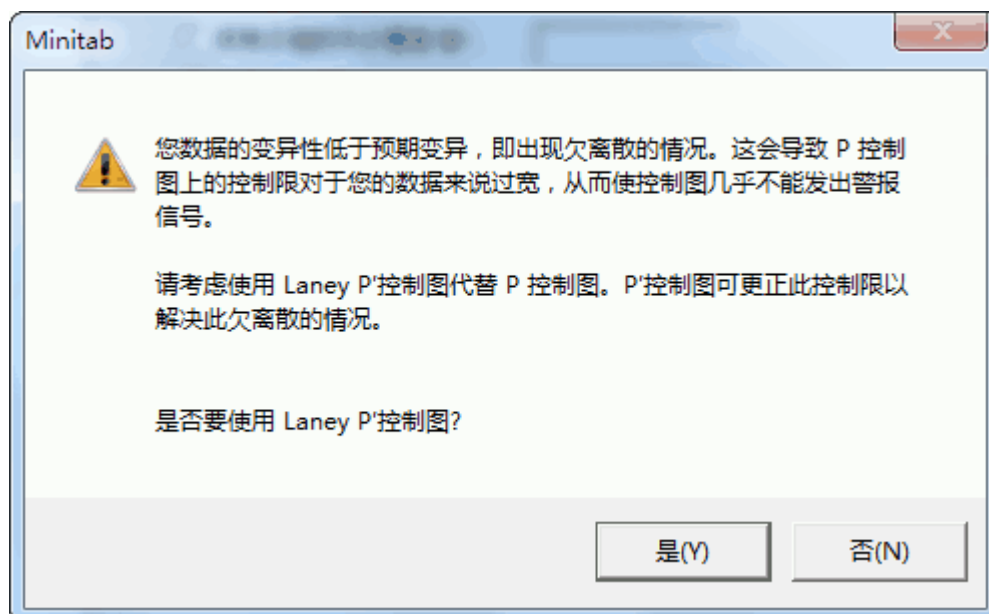
满足下列条件时会发生过度离散：

- 观测到的变化与预期变化的比率大于 130%。
- 超过 2% 的点在控制限值外。
- 超出控制限值的点数大于 1。

如果检测到过度离散，Minitab 将显示一条消息，询问用户是否要显示 Laney P' 或 U' 控制图。下面显示的是 P' 控制图的消息：



当观测到的变化与预期变化的比率小于 75% 时会发生过度集中。如果检测到过度集中，Minitab 会显示一条消息，询问用户是否要显示 Laney P' 或 U' 控制图。下面显示的是 P' 控制图的消息：



如果用户选择使用 Laney 控制图，则 Minitab 将在总结报告中显示 Laney 控制图。如果用户选择不使用 Laney 控制图，Minitab 将在总结报告中显示传统的 P 或 U 控制图。不过，传统的控制图和 Laney 控制图都将显示在诊断报告中。显示两个控制图可以让用户看到过度离散或过度集中对传统的 P 或 U 控制图的影响，并确定 Laney 控制图是否更适合其数据。

此外，检查过度离散或过度集中时，“协助报告卡”会显示以下状态指标：

状态	条件
	<p>离散率 &gt; 130%，不到 2% 的点超出控制限值或超出控制限值的点数 = 1</p> <p>离散率 &gt; 75% 且 &lt;= 130%</p> <p>离散率 &gt; 130%，超过 2% 的点超出控制限值，超出控制限值的点数 &gt; 1 且用户选择使用 Laney P' 或 U' 控制图</p> <p>离散率 &lt; 75%，且用户选择使用 Laney P' 或 U' 控制图</p> <p>其中</p> <p>离散率 = <math>100 * (\text{观测到的变化}) / (\text{预期变化})</math></p>
	<p>离散率 &gt; 130%，超过 2% 的点超出控制限值，超出控制限值的点数 &gt; 1 且用户未选择使用 Laney P' 或 U' 控制图</p> <p>离散率 &lt; 75% 且用户未选择使用 Laney P' 或 U' 控制图</p>



# Laney P' 和 U' 控制图

传统的 P 控制图和 U 控制图假设数据变化遵循适用于不良品数据的二项式分布或适用于缺陷数据的 Poisson 分布。这些控制图还假设不良品率或缺陷率始终保持不变。Minitab 会执行相关检查，以确定数据变化大于或小于预期变化，从而表明数据可能过度离散或过度集中。请参见上面的“预期变化”和“数据检查”。

如果数据存在过度离散或过度集中，传统的 P 控制图和 U 控制图可能无法达到预期效果。过度离散会导致控制限值过窄，从而导致高误报率。过度集中可能导致控制限值过宽，从而导致忽视特殊原因变化，并误以为是常见原因变化。

## 目标

我们的目标是当数据中检测到过度离散或过度集中时，找出传统的 P 控制图和 U 控制图的备选方案。

## 方法

我们翻阅了相关资料，确定了处理过度离散和过度集中的最佳办法是 Laney P' 和 U' 控制图 (Laney, 2002)。Laney 方法使用常见原因变化的修正定义，修正了太窄（过度离散）或太宽（过度集中）的控制限值。

在 Laney 控制图中，常见原因变化不仅包括子组变化中的常见短期变化，还包括连续子组之间的平均短期变化。Laney 控制图的常见原因变化通过规范数据，并使用相邻子组的平均移动范围（在 Laney 控制图上被称为西格玛 Z）来计算，以调整标准 P 或 U 控制图的控制限值。当子组中的数据变化因相关缺陷率发生波动或缺少数据随机性而大于或小于预期变化时，包括连续子组之间的变化有助于修正功效。

西格玛 Z 计算完毕后，数据将变回原来的单位。使用原始数据单位的好处是，如果子组大小不一样，则可以更改控制限值，就像在传统的 P 控制图和 U 控制图中一样。有关 Laney P' 和 U' 控制图的详细信息，请参见附录 F。

## 结果

Minitab 执行过度离散或过度集中检查，如果检测到任何一种情况，Minitab 建议使用 Laney P' 或 U' 控制图。

# 参考书

- AIAG (1995). Statistical process control (SPC) reference manual. Automotive Industry Action Group.
- Bischak, D.P., & Trietsch, D. (2007). The rate of false signals in  $\bar{X}$  control charts with estimated limits. *Journal of Quality Technology*, 39, 55 - 65.
- Bowerman, B.L., & O'Connell, R.T. (1979). *Forecasting and time series: An applied approach*. Belmont, CA: Duxbury Press.
- Chan, L. K., Hapuarachchi K. P., & Macpherson, B.D. (1988). Robustness of  $\bar{X}$  and R charts. *IEEE Transactions on Reliability*, 37, 117 - 123.
- Davis, R.B., & Woodall, W.H. (1988). Performance of the control chart trend rule under linear shift. *Journal of Quality Technology*, 20, 260 - 262.
- Jones, G., & Govindaraju, K. (2001). A Graphical Method for Checking Attribute Control Chart Assumptions, *Quality Engineering*, 13(1), 19-26.
- Laney, D. (2002). Improved Control Charts for Attributes. *Quality Engineering*, 14(4), 531-537.
- Montgomery, D.C. (2001). *Introduction to statistical quality control*, 4<sup>th</sup> edition. New York: John Wiley & Sons, Inc.
- Schilling, E.G., & Nelson, P.R. (1976). The effect of non-normality on the control limits of  $\bar{X}$  charts. *Journal of Quality Technology*, 8, 183 - 188.
- Trietsch, D. (1999). *Statistical quality control: A loss minimization approach*. Singapore: World Scientific Publishing Co.
- Wheeler, D.J. (2004). *Advanced topics in statistical process control. The power of Shewhart's charts*, 2<sup>nd</sup> edition. Knoxville, TN: SPC Press.
- Yourstone, S.A., & Zimmer, W.J. (1992). Non-normality and the design of control charts for averages. *Decision Sciences*, 23, 1099 - 1113.

# 附录 A：属性控制图的其他假设

P 控制图和 U 控制图需要使用不经过数据检查评估的其他假设：

P 控制图	U 控制图
<ul style="list-style-type: none"><li>• 数据由 <math>n</math> 个不同的项目组成，每个项目分成不良品或良品两类。</li><li>• 子组内的每个项目的不良品率相同。</li><li>• 前面的产品是否有缺陷不会影响不良品率。</li></ul>	<ul style="list-style-type: none"><li>• 计数为离散事件计数。</li><li>• 离散事件在一些明确界定的有限空间、时间或者产品区域内发生。</li><li>• 事件独立发生，事件的可能性与机会区域的大小成比例。</li></ul>

对于每个控制图，前两个假设是数据收集过程固有的一部分；数据本身不能被用来检查这些假设是否成立。第三个假设只能用详尽、高级的数据分析进行验证，此过程不在“协助”中进行。

# 附录 B：稳定性

## 模拟 B1：在检验 1 中增加检验 2 对灵敏度有何影响

当一个点距离中心线大于 3 个标准差时，检验 1 通过发出信号来检测失控点。当连续 9 个点在中心线的同一侧时，检验 2 通过发出信号来检测不良品率或单位缺陷数中的偏差。

为了评估是否同时使用检验 2 和检验 1 来提高属性控制图的灵敏度，我们根据正态 ( $p$ ,  $\sqrt{\frac{p(1-p)}{n}}$ ) ( $p$  是不良品率,  $n$  是子组大小) 分布确定了 P 控制图的控制限值, 根据正态 ( $u$ ,  $\sqrt{u}$ ) ( $u$  是单位缺陷平均数) 分布确定了 U 控制图的控制限值。我们按标准差 (SD) 的倍数来设置分布的位置偏差 ( $p$  或  $u$ )，然后记录下检测到 10,000 次迭代中的每一次迭代信号所需的子组数。结果显示在表 1 中。

表 1 检验 1 失败 (检验 1)、检验 2 失败 (检验 2)，或者检验 1 或检验 2 失败 (检验 1 或 2) 前的子组平均数。偏差等于标准差 (SD) 的倍数。

偏差	检验 1	检验 2	检验 1 或 2
0.5 SD	154	84	57
1 SD	44	24	17
1.5 SD	15	13	9
2 SD	6	10	5

如表中所示，同时使用两种检验 (检验 1 或 2 列) 时，平均需要 57 个子组，才能检测到 0.5 的标准位置偏差，相比而言，若单独使用检验 1，平均需要 154 个子组，才能检测到 0.5 的标准平均值偏差。因此，使用这两种检验可显著提高检测到不良品率或单位缺陷平均数中的小偏差的灵敏度。但是，当偏差量增加时，增加检验 2 不会显著提高灵敏度。

# 附录 C：子组数

## 公式 C1：基于控制上限的 95% 置信度的 P 控制图所需的子组数

为了确定是否有足以确保误报率保持相当低的水平的子组，我们按照 Bischak (1999) 确定了子组数，从而确保检验 1 的误报率不高于 95% 置信度的 2%。

首先，我们发现  $p_c$ ，从而

$$p_c + 3 \sqrt{\frac{p_c(1-p_c)}{n}} = \bar{p} + z_{0.99} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

其中

$p_c$  = 产生高于控制上限的 1% 误报率的比例，假设  $\bar{p}$  为  $p$  的真值。由于控制限值的对称性，同时考虑控制上限和下限时，总误报率变为 2%。

$n$  = 子组大小（如果子组大小不同，则使用平均子组大小）

$\bar{p}$  = 平均不良品率

$z_p$  = 对根据平均值 = 0，标准差 = 1 的正态分布的  $p$  值估出的 cdf 求倒数

为了确定子组数，我们将计算控制上限的 95% 置信度下限，并将其设置为等于  $p_c$ ，

$$p_c = \bar{p} - z_{0.95} \sqrt{\frac{\bar{p}(1-\bar{p})}{nm}}$$

同时求得  $m$  的值，结果如下：

$$m = \frac{\bar{p}(1-\bar{p})}{n(\bar{p}-p_c)^2 z_{0.95}^2}$$

使用这个公式，我们可以针对各种比例和子组大小，确定为确保高于控制上限的误报率始终低于 95% 置信度的 1% 而所需的子组数，如表 2 中所示。由于控制限值的对称性，需要相同的子组数，以确保针对 P 控制图的检验 1 的总误报率始终低于 95% 置信度的 2%。

表 2 各种子组大小 ( $n$ ) 和比例 ( $\bar{p}$ ) 的子组数 ( $m$ )

子组大小 ( $n$ )	$\bar{p}$				
	0.001	0.005	0.01	0.05	0.1
10	1881	421	228	60	35
50	425	109	64	23	16
100	232	65	41	17	13
150	165	49	32	14	11

200	131	41	27	13	10
500	65	24	18	10	9

注意：对于变量控制图，我们将检验 1 的总误报率限制为 1%。对于属性控制图，我们出于实际原因将标准放宽到了 2%。在许多情况下，P 控制图上的不良品率较小，从而需要一个非常大的子组数来实现精度，如表 2 中所示。

## 公式 C2：基于 95% 置信度的控制上限的 U 控制图所需的子组数

我们使用如上述 P 控制图所示的相同方法。我们按照 Trietsch (1999) 来确定将确保检验 1 的总误报率不超过 95% 置信度的 2% 的子组数。

首先，我们发现  $c_c$ ，从而

$$c_c + 3\sqrt{c_c} = \bar{c} + z_{0.99}\sqrt{\bar{c}}$$

其中

$c_c$  = 产生高于控制上限的 1% 误报率的子组缺陷平均数，假设  $\bar{c}$  为  $c$  的真值。由于控制限值的对称性，同时考虑控制上限和下限时，检验 1 的总误报率变为 2%。

$\bar{c}$  = 子组缺陷平均数（如果子组大小各不相同，则使用平均子组大小）

$z_p$  = 对根据平均值 = 0，标准差 = 1 的正态分布的  $p$  值估出的 cdf 求倒数

为了确定子组数，我们将计算控制上限的 95% 置信度下限，并将其设置为等于  $c_c$ ，

$$c_c = \bar{c} - z_{0.95} \sqrt{\frac{\bar{c}}{m}}$$

同时求得  $m$  的值，结果如下：

$$m = \frac{\bar{c}}{\left(\frac{\bar{c} - c_c}{z_{0.95}}\right)^2}$$

根据上述计算得出的一些结果显示在表 3 中。

表 3 子组缺陷平均数的各个值 ( $\bar{c}$ ) 对应的子组数 ( $m$ )

$\bar{c}$	0.1	0.3	0.5	0.7	1.0	3.0	5.0	10.0	30.0	50.0
子组数	232	95	65	52	41	22	18	14	10	9

注意 对于变量控制图，我们将检验 1 的总误报率限制为 1%。对于属性控制图，我们出于实际原因将标准放宽到了 2%。在许多情况下，子组缺陷数较小，从而需要一个非常大的子组数来实现精度，如表 3 中所示。

# 附录 D：子组大小

中心极限定理表明，正态分布可以近似估计独立、恒等分布的随机变量的平均值分布。对于 P 控制图， $\hat{p}$ （子组率）是独立、恒等分布的 Bernoulli 随机变量的平均值。对于 U 控制图， $\hat{u}$ （子组率）是独立、恒等分布的 Poisson 随机变量的平均值。因此，在这两种情况下，正态分布可以用作近似估计。

当子组大小增加时，近似估计的精确度会随之提高。这种近似估计也会随着不良品率的增高（P 控制图）或单位缺陷数的增高（U 控制图）而提高。无论是子组大小较小还是 p 值（P 控制图）或 u 值（U 控制图）较小， $\hat{p}$  和  $\hat{u}$  都呈右偏分布，从而将增加误报率。因此，我们可以通过查看误报率来评估正态近似估计的准确度，我们也可以确定为获得足够的正态近似估计而所需的最小子组大小。

为此，我们进行了模拟，以评估 P 控制图和 U 控制图的各种子组大小的误报率并将结果与正态假设下的预期误报率进行比较（检验 1 为 0.27%，检验 2 为 0.39%）。

## 模拟 D1：P 控制图的子组大小、比例以及误报率之间的关系

使用 10,000 个初始子组集，我们确定了各种子组大小（n）和比例（p）的控制限值。我们还记录了其他 2,500 个子组的误报率。然后，我们进行了 10,000 次迭代并计算了检验 1 和检验 2 的平均误报率，如表 4 中所示。

表 4 检验 1、检验 2 中各种子组大小（n）和比例（p）的误报率（np）

子组大小 (n)	p				
	0.001	0.005	0.01	0.05	0.1
10	0.99, 87.37 (0.01)	4.89, 62.97 (0.05)	0.43, 40.14 (0.1)	1.15, 1.01 (0.5)	1.28, 0.42 (1)
50	4.88, 63.00 (0.05)	2.61, 10.41 (0.25)	1.38, 1.10 (0.5)	0.32, 0.49 (2.5)	0.32, 0.36 (5)
100	0.47, 40.33 (0.10)	1.41, 1.12 (0.5)	1.84, 0.49 (1)	0.43, 0.36 (5)	0.20, 0.36 (10)
150	1.01, 25.72 (0.15)	0.71, 0.43 (0.75)	0.42, 0.58 (1.5)	0.36, 0.42 (7.5)	0.20, 0.36 (15)
200	1.74, 16.43 (0.2)	1.86, 0.50 (1.00)	0.43, 0.41 (2)	0.27, 0.36 (10)	0.34, 0.36 (20)
500	1.43, 1.12 (0.5)	0.42, 0.50 (2.5)	0.52, 0.37 (5)	0.32, 0.37 (25)	0.23, 0.36 (50)

表 4 中的结果表明，当比率（p）较小（例如 0.001 或 0.005）时，或当样本量较小（n=10）时，误报率通常最高。因此，当产品值 np 较小时，误报率最高，当 np 值较大时，误报率最

低。当  $np$  值大于或等于 0.5 时，检验 1 和检验 2 的组合误报率低于 2.5% 左右。然而，当  $np$  值小于 0.5 时，检验 1 和检验 2 的组合误报率要高得多，正好达到 10% 以上的水平。因此，基于这一标准，当值  $np \geq 0.5$  时，P 控制图的功效足够。由此，子组大小应至少为  $\frac{0.5}{\bar{p}}$ 。

## 模拟 D2: U 控制图的子组大小、单位缺陷数以及误报率之间的关系

使用 10,000 个初始子组集，我们确定了各种子组大小 ( $n$ ) 和子组缺陷数 ( $c$ ) 的控制限值。我们还记录了其他 2,500 个子组的误报率。然后，我们进行了 10,000 次迭代并计算了检验 1 和检验 2 的平均误报率，如表 5 中所示。

表 5 检验 1、检验 2 中各个子组缺陷数的误报率 ( $c = nu$ )

c	0.1	0.3	0.5	0.7	1.0	3.0	5.0	10.0	30.0	50
误报率	0.47, 40.40	3.70, 6.67	1.44, 1.13	0.57, 0.39	0.36, 0.51	0.38, 0.40	0.54, 0.38	0.35, 0.37	0.29, 0.37	0.25, 0.37

表 5 中的结果表明，当子组产品大小 ( $n$ ) 与单位缺陷数 ( $u$ ) 相乘的结果（相当于子组缺陷数）较小时，误报率最高。当  $c$  值大于或等于 0.5 时，检验 1 和检验 2 的组合误报率低于 2.5% 左右。然而，当  $c$  值小于 0.5 时，检验 1 和检验 2 的组合误报率要高得多，正好达到 10% 以上的水平。因此，基于这一标准，当值  $c = nu \geq 0.5$  时，U 控制图的功效足够。由此，子组大小应至少为  $\frac{0.5}{\bar{u}}$ 。



# 附录 E：过度离散/过度集中

假设  $d_i$  是子组  $i$  中的缺陷计数， $n_i$  是子组大小。

首先，规范缺陷计数。考虑到子组大小可能不同，使用调整后的缺陷计数 (adjd<sub>i</sub>)：

adjd<sub>i</sub> = 调整后的缺陷计数，子组  $i = \frac{d_i}{n_i}(\bar{n})$ ，其中

$\bar{n}$  = 平均子组大小

$$X_i = \sin^{-1} \sqrt{\frac{adjd_i + 3/8}{\bar{n} + 0.75}}$$

规范计数 ( $X_i$ ) 的一个标准差等于  $\frac{1}{\sqrt{4*\bar{n}}}$ 。这意味着 2 个标准差等于  $\frac{1}{\sqrt{\bar{n}}}$ 。

然后，利用规范的数据计数生成标准的正态概率图。只使用中间 50% 的绘图点的回归线拟合。查找 25% 和 75% 的已转换计数数据，并使用所有  $\geq 25\%$  且  $\leq 75\%$  的 X-Y 对。这条线用于获得对应于 -1 和 +1 的 Z 值的预计已转换计数值。此回归分析中的“Y”数据是已转换计数的正态计分，“X”数据是已转化计数。

所观测到的变化计算如下：

假设  $Y(-1)$  是  $Z = -1$  的预计已转换计数

假设  $Y(+1)$  是  $Z = +1$  的预计已转换计数

2 个标准差的观测估计值 =  $Y(+1) - Y(-1)$ 。

预期变化计算如下：

2 个标准差的预期估计值 =  $\frac{1}{\sqrt{\bar{n}}}$

计算观测到的变化与预期变化之比，并转换为一个百分比。如果该比率  $> 130\%$ ，则超过 2% 的点超出控制限值，超出控制限值的点数  $> 1$ ，表明过度离散。如果该比率  $< 75\%$ ，则表明过度集中。

# 附录 F: Laney P' 和 U' 控制图

Laney P' 控制图和 U' 控制图的设计理念是，当子组数据来自一个缺陷率或不良品率固定不变的随机过程时，考虑观测到的子组间变化不符合预期变化的情况。相关缺陷率或不良品率中的小变动通常在每个过程中都会发生。当子组大小相对较小时，该子组中的采样变化足够大，导致这些小变动并不明显。当子组大小增加时，采样变化将减小，而且相关缺陷率或不良品率中的小变动会变得足够大，误报率随之增大，从而对标准的 P 控制图和 U 控制图产生不利影响。一些示例显示，误报率高达 70%。这种情况被称为过度离散。

为了解决这个问题，我们研制出了另一种方法，它将规范子组 p 值或 u 值，并在 I 控制图中绘制规范数据。I 控制图使用规范值的移动范围来确定其控制限值。因此，I 控制图方法将通过将一个子组的不良品率或缺陷率变化增加到下一个子组中来改变常见原因变化的定义。

Laney 方法会将数据变回原来的单位。这样做的好处是，如果子组大小各不相同，控制限值将不固定，因为它们采用的是 I 控制图方法。

P' 和 U' 控制图根据不同子组大小的变量控制限值，重新定义了常见原因变化。因此，这些控制图的关键假设是，常见原因变化的定义已改变，它包括子组中存在的常见短期变化，以及连续子组间存在的平均短期变化。

## Laney P' 控制图

假设

$X_i$  = 子组 i 中的缺陷数

$n_i$  = 子组 i 的子组大小

$p_i$  = 子组 i 的不良品率

$$\bar{p} = \frac{\sum X_i}{\sum n_i}$$

$$\sigma p_i = \sqrt{\frac{\bar{p} * (1 - \bar{p})}{n_i}}$$

首先，将  $p_i$  值转换为 z 值：

$$Z_i = \frac{p_i - \bar{p}}{\sigma p_i}$$

然后，使用长度为 2 的移动范围评估 Z 值变化，并计算西格玛 Z ( $\sigma z$ ) 值。

$$\sigma z = \frac{\overline{MR}}{1.128}$$

其中 1.128 是无偏常量。

将数据变回原来的比例：

$$p_i = \bar{p} + \sigma p_i * \sigma z$$

因此， $p_i$  的标准差是：

$$sd(p_i) = \sigma p_i * \sigma z$$

控制限值和中心线计算如下：

$$\text{中心线} = \bar{p}$$

$$UCL = \bar{p} + 3 * sd(p_i)$$

$$LCL = \bar{p} - 3 * sd(p_i)$$

## Laney U' 控制图

假设

$X_i$  = 子组  $i$  中的缺陷数

$n_i$  = 子组  $i$  的子组大小

$u_i$  = 子组  $i$  的不良品率

$$\bar{u} = \frac{\sum X_i}{\sum n_i}$$

$$\sigma u_i = \sqrt{\frac{\bar{u} * (1 - \bar{u})}{n_i}}$$

首先，将  $p_i$  转换为  $z$  值：

$$Z_i = \frac{u_i - \bar{u}}{\sigma u_i}$$

然后，使用长度为 2 的移动范围评估  $Z$  值变化，并计算西格玛  $Z$  ( $\sigma z$ ) 值。

$$\sigma z = \frac{\overline{MR}}{1.128}$$

其中 1.128 是无偏常量。

将数据变回原来的比例：

$$u_i = \bar{u} + \sigma u * \sigma z$$

因此， $p_i$  的标准差是：

$$sd(u_i) = \sigma u_i * \sigma z$$

控制限值和中心线计算如下：

$$\text{中心线} = \bar{u}$$

$$\text{UCL} = \bar{u} + 3 * sd(u_i)$$

$$\text{LCL} = \bar{u} - 3 * sd(u_i)$$

© 2020 Minitab, LLC. All rights reserved. Minitab®, Minitab Workspace™, Companion by Minitab®, Salford Predictive Modeler®, SPM®, and the Minitab® logo are all registered trademarks of Minitab, LLC, in the United States and other countries. Additional trademarks of Minitab, LLC can be found at [www.minitab.com](http://www.minitab.com). All other marks referenced remain the property of their respective owners.