

이 문서는 Minitab Statistical Software의 보조 도구에서 사용되는 방법과 데이터 검사를 개발하기 위해 Minitab 통계 학자들이 실시한 연구에 대해 설명하는 전체 백서 중 하나입니다.

계량형 관리도

개요

관리도는 공정을 정기적으로 모니터링하여 공정이 관리 상태에 있는지 여부를 확인하기 위해 사용됩니다. Minitab 보조 도구에는 연속형 데이터에 널리 사용되는 두 가지 관리도가 포함되어 있습니다.

- Xbar-R 또는 Xbar-S 관리도. 이 관리도는 데이터가 부분군으로 수집되는 경우 사용됩니다. Minitab에서는 합동 표준 편차를 사용하여 군내 표준 편차를 추정합니다. R 관리도는 대략 최대 크기 10의 부분군에 대해 효과적인 변동 추정치를 제공합니다(AIAG, 1995; Montgomery, 2001). 부분군 크기가 더 큰 경우에는 S 관리도가 더 효과적인 군내 표준 편차의 추정치를 제공합니다. Minitab의 과거 규칙을 따르고 보수적인 상태를 유지하기 위해 부분군 크기가 8보다 클 때는 S 관리도를 사용하는 것이 좋습니다. 부분군 크기가 8보다 작거나 같을 때는 R 및 S 관리도가 비슷한 결과를 제공합니다.
- 개체 및 이동 범위(I-MR) 관리도. 이 관리도는 부분군이 없을 때 사용됩니다. Minitab에서는 길이 2의 평균 이동 범위 방법을 사용하여 표준 편차를 추정합니다.

관리도의 관리 한계는 일반적으로 Six Sigma 프로젝트의 관리 단계에서 설정됩니다. 특수 원인이 존재할 때 신속하게 신호를 보낼 수 있을 만큼 충분히 민감한 관리도가 좋은 관리도입니다. 이 민감도는 특수 원인에 대한 신호를 보내는 데 필요한 평균 부분군 수를 계산하여 평가할 수 있습니다. 좋은 관리도는 또한 공정이 관리 상태에 있을 때 "잘못된 경고"를 보내는 일이 거의 없습니다. 잘못된 경고 비율은 공정이 관리 상태에 있을 때 "관리 이탈" 상태에 있는 것으로 간주되는 부분군의 비율을 계산하여 평가할 수 있습니다.

일반적으로, 관리도는 각 관측치가 정규 분포를 따르고, 각 관측치가 독립적이면서 부분군 내에 일반 원인 변동만 존재하는 경우 최적화됩니다. 따라서 보조 도구 보고서 카드에서는 다음과 같은 데이터 검사를 자동으로 수행하여 이러한 조건을 평가합니다.

- 정규성
- 안정성
- 데이터 양
- 상관된 데이터

이 문서에서는 이러한 조건이 달라질 때 계량형 관리도가 어떻게 동작하는지 조사하고 이러한 조건에 대한 요구 사항을 평가하기 위해 어떠한 가이드라인을 설정했는지 설명합니다.

데이터 검사

정규성

관리도는 공정 데이터가 정규 분포를 따른다는 가정을 바탕으로 하지 않지만 이 특수 원인 검정에 사용되는 기준은 이 가정을 바탕으로 합니다. 데이터가 심하게 치우쳐 있거나 데이터 분포의 끝이 너무 두꺼운("두꺼운 꼬리") 경우 검정 결과가 정확하지 않을 수 있습니다. 예를 들어, 관리도가 예상보다 높은 비율로 잘못된 경고 신호를 보낼 수 있습니다.

목적

Minitab에서는 Xbar 관리도 및 I 관리도에서 비정규 데이터의 영향을 조사했습니다. 또한 비정규성이 잘못된 경고 비율에 어떻게 영향을 미치는지 확인하고자 했습니다. 특히 공정이 실제로 관리 상태에 있는데 관리도에서는 점들이 관리 이탈 상태에 있다고(잘못된 경고) 나타내는 비율이 비정규 데이터로 인해 유의하게 증가하는지 여부를 확인하고자 했습니다.

방법

Minitab에서는 10,000개의 부분군 및 여러 수준의 비정규성을 사용하여 시뮬레이션을 수행하고 잘못된 경고의 백분율을 기록했습니다. 시뮬레이션을 통해 다양한 조건을 검정하고 비정규성의 영향을 확인할 수 있습니다. Minitab에서는 실제로 일반적인 비정규 분포인 오른쪽으로 치우친 분포 및 두꺼운 꼬리를 갖는 대칭 분포를 선택했습니다. 자세한 내용은 부록 A를 참조하십시오.

결과

XBAR 관리도(부분군 크기 ≥ 2)

시뮬레이션 결과를 보면 부분군 크기가 2 이상이면 데이터가 정규 분포를 따르지 않을 때 잘못된 경고 비율이 유의하게 증가하지 않는다는 것을 알 수 있습니다. 이 결과를 바탕으로 Minitab에서는 Xbar-R 또는 Xbar-S 관리도에 대해 정규성을 검사하지 않습니다. 데이터가 심하게 치우치거나 극히 두꺼운 꼬리를 가진 경우에도 검정 1과 검정 2에 대한 잘못된 경고 비율은 2%보다 작으며, 정규 분포의 잘못된 경고 비율 0.7%보다 아주 많이 높지 않습니다.

I 관리도(부분군 크기 = 1)



시뮬레이션 결과를 보면 I 관리도가 비정규 데이터에 민감한 것을 알 수 있습니다. 데이터가 정규 분포를 따르지 않으면 I 관리도는 데이터가 정규 분포를 따를 때보다 4 - 5배 높은 잘못된

경고 비율을 생성합니다. 비정규 데이터에 대한 I 관리도의 민감도를 해결하기 위해 보조 도구에서는 다음과 같은 작업을 수행합니다.

- 데이터의 비정규성이 심한 경우, 즉 관리 한계를 벗어나는 점의 수가 예상보다 많은 경우(2개 이상의 점 또는 점의 2% 이상이 관리 한계를 벗어나는 경우) Anderson-Darling 검정을 수행합니다.
- Anderson-Darling 검정 결과 데이터가 정규 분포를 따르지 않는 경우 보조 도구에서는 최적의 Box-Cox 람다를 사용하여 데이터를 변환합니다. 변환 데이터에 대해 Anderson-Darling 검정이 수행됩니다. 검정이 (데이터가 정규 분포를 따른다는) 귀무 가설을 기각하지 못하면 보조 도구에서는 공정이 원래 비정규 데이터를 생성하는 경우 변환 데이터를 사용할 것을 제안합니다.

Box-Cox 변환은 오른쪽으로 치우친 비정규 데이터에 대해서만 효과적입니다. 변환이 비정규 데이터에 대해 효과적이지 못할 경우 다른 옵션을 고려해야 할 수도 있습니다. 또한 Anderson-Darling 검정 및 Box-Cox 변환은 극단 관측치의 영향을 받기 때문에 데이터를 변환하기 전에 특수 원인을 가진 것으로 알려진 점은 제외해야 합니다.

정규성을 확인하는 경우 I 관리도에 대한 보조 도구의 보고서 카드에는 다음과 같은 상태가 표시됩니다.

| 상태 | 조건 |
|---|--|
|  | I 관리도에서 2개 미만의 점 또는 점의 2% 미만이 관리 한계를 벗어나는 경우 또는 데이터 또는 Box-Cox 변환 데이터에 대한 Anderson-Darling 정규성 검정이 알파 = 0.01에서 (데이터가 정규 분포를 따른다는) 귀무 가설을 기각하지 못하는 경우 |
|  | 위의 조건이 유지되지 않는 경우 |

안정성

계량형 관리도의 경우 8개의 검정을 수행하여 공정의 안정성을 평가할 수 있습니다. 이러한 검정을 동시에 사용할 경우 관리도의 민감도가 증가합니다. 그러나 관리도에 검정이 추가될수록 잘못된 경고 비율이 증가하기 때문에 각 검정의 목적 및 증가되는 유용성을 확인하는 것이 중요합니다.

목적

Minitab에서는 8개의 안정성 검정 중 보조도구의 계량형 관리도에 포함시킬 검정을 결정하고자 했습니다. 첫 번째 목표는 잘못된 경고 비율을 유의하게 증가시키지 않고 관리 이탈 조건에 대한 민감도를 유의하게 증가시키는 검정을 식별하는 것이었습니다. 두 번째 목표는 관리도의 단순성과 실용성을 보장하는 것이었습니다. Minitab의 연구는 Xbar 관리도 및 I 관리도에 대한 검정에 초점을 맞추었습니다. R, S 및 MR 관리도의 경우에는 점이 관리 한계를 벗어날 때 신호를 보내는 검정 1만 사용합니다.

방법

Minitab에서는 안정성 검정 조합을 사용할 경우 관리도의 민감도 및 잘못된 경고 비율에 어떤 영향을 미치는지 확인하기 위해 시뮬레이션을 수행하고 문헌을 검토했습니다. 또한 검정과 연관된 특수 원인의 출현율을 조사했습니다. 각 검정에 사용된 방법에 대한 자세한 내용은 아래 결과 항목과 부록 B를 참조하십시오.

결과

Minitab에서는 검정 1, 2 및 7이 Xbar 관리도 및 I 관리도의 안정성을 평가하는 데 가장 유용하다는 것을 알았습니다.

검정 1: 관리 한계를 벗어난 점 식별

검정 1은 중심선으로부터 3 표준 편차 밖에 있는 점들을 식별합니다. 검정 1은 보편적으로 관리 이탈 상황을 탐지하는 데 필요한 것으로 간주됩니다. 검정 1의 잘못된 경고 비율은 0.27%에 불과합니다.

검정 2: 평균의 변동 식별

검정 2는 9개의 연속된 점이 중심선으로부터 같은 쪽에 있을 때 신호를 보냅니다. Minitab에서는 표준 편차의 배수로 설정된 4개의 다른 평균을 사용하여 시뮬레이션을 수행하고 신호를 탐지하기까지의 부분군의 수를 확인했습니다. 그리고 정규 분포를 기반으로 관리 한계를 설정했습니다. 검정 2를 추가하면 관리도의 민감도가 유의하게 증가하여 평균의 작은 변동을 탐지할 수 있다는 것을 알았습니다. 검정 1과 검정 2를 함께 사용할 경우 검정 1만 사용할 경우보다 훨씬 더 적은 수의 부분군으로 평균의 작은 변동을 탐지할 수 있습니다. 따라서 검정 2를 추가하면 일반적인 관리 이탈 상황을 탐지하는 데 도움이 되며 약간의 잘못된 경고 비율 증가를 용인할 수 있을 만큼 민감도가 증가합니다.

검정 7: 너무 넓은 관리 한계 식별

검정 7은 12-15개의 연속된 점이 중심선으로부터 1 표준 편차 내에 있을 때 신호를 보냅니다. 검정 7은 데이터로부터 관리 한계를 추정하는 경우 XBar 관리도에만 사용됩니다. 이 검정이 실패하는 경우 원인은 일반적으로 합리적인 부분군을 구성하지 않은 결과 발생하는 부분군 내 계층적 변동(계층화)입니다. 관리도에서 관리 이탈 상황을 정확히 탐지하기 위해서는 합리적인 부분군을 구성하는 것이 중요하기 때문에 Minitab에서는 데이터에서 관리 한계를 추정할 때 수정된 검정 7을 사용합니다. 검정 7은 부분군의 수에 따라 연속된 점의 수가 12와 15 사이일 때 실패 신호를 보냅니다.

| $k = (\text{부분군 수}) \times 0.33$ | 필요한 점의 수 |
|----------------------------------|-------------|
| $k < 12$ | 12 |
| $k \geq 12$ 및 $k \leq 15$ | 정수 $\geq k$ |
| $k > 15$ | 15 |

보조 도구에 포함되지 않은 검정

검정 3: K개의 연속된 점이 모두 상승 또는 하락

검정 3은 공정 평균의 점진적 이동을 탐지하기 위해 설계되었습니다(Davis and Woodall, 1988). 그러나 검정 1 및 검정 2 외에 검정 3을 사용해도 공정 평균의 점진적 이동을 탐지하는 관리도의 민감도는 유의하게 증가되지 않습니다. 이미 시뮬레이션 결과를 기반으로 검정 1과 2를 사용하기로 결정했기 때문에 검정 3을 포함시키더라도 관리도가 유의하게 더 좋아지는 것은 아닙니다.

검정 4: K개의 연속된 점이 교대로 상승 또는 하락

실제로 이 패턴이 발생할 수 있지만 특정한 하나의 패턴에 대한 검정 대신 비정상적인 추세나 패턴을 찾아보는 것이 좋습니다.

검정 5: K+1개의 점 중에서 K개 점이 중심선으로부터 2 표준 편차 범위 밖에 있음

이 검정은 일반적인 특수 원인 상황을 고유하게 식별하지 않았기 때문에 관리도의 단순성을 보장하기 위해 제외했습니다.



검정 6: K+1개의 점 중에서 K개 점이 중심선으로부터 1 표준 편차 범위 밖에 있음

이 검정은 일반적인 특수 원인 상황을 고유하게 식별하지 않았기 때문에 관리도의 단순성을 보장하기 위해 제외했습니다.

검정 8: K개의 연속된 점이 중심선으로부터 1 표준 편차 범위 밖에 있음(양쪽)

이 검정은 일반적인 특수 원인 상황을 고유하게 식별하지 않았기 때문에 관리도의 단순성을 보장하기 위해 제외했습니다.

안정성을 확인하는 경우 보조 도구 보고서 카드에는 다음과 같은 상태가 표시됩니다.

| 상태 | 조건 |
|--|--|
|  | <p>평균에 대한 관리도(I 관리도 또는 Xbar 관리도) 및 변동에 대한 관리도(MR, R 또는 S 관리도)에 검정 실패 없음.</p> <p>각 관리도에 사용되는 검정은 다음과 같습니다.</p> <ul style="list-style-type: none">• I 관리도: 검정 1 및 검정 2.• Xbar 관리도: 검정 1, 검정 2 및 검정 7. 검정 7은 관리 한계가 데이터로부터 추정되는 경우에만 수행됩니다.• MR, R 및 S 관리도: 검정 1. |
|  | <p>위의 조건이 유지되지 않는 경우.</p> |

데이터 양

관리 한계에 대해 알려진 값이 없을 경우 데이터로부터 추정해야 합니다. 정밀한 한계 추정치를 얻으려면 충분한 데이터가 있어야 합니다. 데이터가 부족한 경우 표본 추출 변동성으로 인해 한계 추정치가 “참” 한계와 거리가 멀 수도 있습니다. 관측치의 수를 늘리면 한계의 정밀도를 개선할 수 있습니다.

목적

Minitab에서는 정밀한 관리 한계를 얻는데 필요한 관측치의 수를 조사했습니다. 목표는 검정 1로 인한 잘못된 경고가 95% 신뢰 수준에서 1%를 넘지 않도록 하는 데 필요한 데이터의 양을 결정하는 것이었습니다.

방법



데이터가 정규 분포를 따르고 표본 추출 변동성으로 인한 오차가 없을 경우 관리 상한 위에 있는 점의 백분율은 0.135%입니다. 관측치의 수가 적절한지 여부를 확인하기 위해 Minitab에서는 Bischaj(2007)에 요약된 방법에 따라 관리 상한 위에 있는 점으로 인한 잘못된

경고 비율이 95% 신뢰 수준에서 0.5%를 넘지 않도록 유지했습니다. 관리 한계의 대칭성으로 인해 이 방법을 사용할 경우 검정 1로 인한 잘못된 경고 비율이 1%가 됩니다.

결과

Minitab에서는 거의 모든 부분군 크기에 대해 정밀한 관리 한계를 얻기 위해 총 100개의 관측치가 적절하다는 것을 확인했습니다. 부분군 크기가 1이나 2인 경우 약간 더 많은 관측치가 필요했지만, 관측치가 100개인 경우 잘못된 경고 비율은 여전히 상당히 낮았습니다. 따라서 간단히 모든 부분군 크기에 대해 총 관측치 수 100개를 컷오프 값으로 사용합니다.

이러한 결과를 바탕으로 데이터 양을 확인하는 경우 보조 도구에는 다음과 같은 정보가 표시됩니다.

| 상태 | 조건 |
|--|-------------------|
|  | 관측치 수 \geq 100. |
|  | 관측치 수 $<$ 100. |

상관된 데이터

자기 상관은 시간별로 수집되는 데이터 점 간의 종속성에 대한 척도입니다. 대부분의 공정 데이터는 최소한 약간의 자기 상관을 나타냅니다. 자기 상관이 중간 정도이거나 높으면 검정 결과가 잘못될 수 있습니다. 일반적으로, 자기 상관된 데이터는 양의 자기 상관을 나타내며, 이에 따라 군내 변동이 감소하고 잘못된 경고 비율이 증가할 수 있습니다.

목적

Minitab에서는 자기 상관과 잘못된 경고 비율 간의 관계를 조사했습니다. 목적은 허용되지 않는 잘못된 경고 비율을 생성하는 자기 상관 수준을 확인하는 것이었습니다. 간단하게 하기 위해, 시차 1에 대한 자기 상관이 시차 \geq 2에 대한 자기 상관보다 클 가능성이 높기 때문에 시차 1에 대한 자기 상관을 고려했습니다.

방법



자기 상관된 공정의 표준 모형을 사용하여 Minitab에서는 세 가지 부분군 크기($n = 1, 3, 5$)에 대해 $\phi = 0.2, 0.4, 0.5, 0.6, 0.8$ (ϕ 은 시차 1 자기 상관임)로 자기 상관을 수행했습니다. 처음에 10,000개의 부분군을 사용하여 관리 한계를 설정했습니다. 그런 다음 추가 2,500개 부분군에

대한 잘못된 경고의 백분율을 기록했습니다. 10,000번의 반복을 수행하고 잘못된 경고의 평균 백분율을 계산했습니다. 자세한 내용은 부록 D를 참조하십시오.

결과

시뮬레이션 결과를 보면 자기 상관 수준이 중간 정도인 경우에도 잘못된 경고 비율이 유의하게 증가한다는 것을 알 수 있습니다. 자기 상관 ≥ 0.4 인 경우 잘못된 경고 비율이 매우 높고 관리도가 무의미해집니다. 이 문제를 해결하기 위해 보조 도구에서는 관리 한계를 벗어나는 점의 수가 예상보다 많으면(즉, 2개 이상의 점 또는 점의 2% 이상이 관리 한계를 벗어나는 경우) 데이터에 자기 상관이 있는지 자기 상관 검정을 수행합니다. 이 경우 보조 도구에서는 먼저 연속된 데이터 점(시차 = 1) 간의 자기 상관이 0.2보다 유의하게 큰지 여부를 검정합니다. 자기 상관이 0.2보다 유의하게 큰 경우 보조 도구에서는 연속적인 데이터 점(시차 = 1) 간의 자기 상관이 0.4보다 유의하게 큰지 여부를 검정합니다.

상관된 데이터를 확인하는 경우 보조 도구의 보고서 카드에는 다음과 같은 상태가 표시됩니다.

| 상태 | 조건 |
|---|---|
|  | <p>관리 한계를 벗어나는 점의 수가 예상보다 많지 않습니다. 즉, 2개 미만의 점 또는 점의 2% 미만이 관리 한계를 벗어납니다.</p> <p>관리 한계를 벗어나는 점의 수가 예상보다 많지만 자기 상관 = 0.2 대 자기 상관 > 0.2의 검정이 알파 = 0.01에서 귀무 가설을 기각하지 못합니다. 따라서 중간 수준 이상의 자기 상관이 존재한다는 결론을 내리기에 충분한 증거가 없습니다.</p> |
|  | <p>위의 조건이 유지되지 않는 경우</p> <p>참고: 자기 상관 = 0.2의 귀무 가설이 기각되는 경우 Minitab에서는 자기 상관 = 0.4 대 자기 상관 > 0.4의 후속 검정을 수행합니다. 자기 상관 = 0.4 검정이 기각되면 주의 메시지의 심각도를 높입니다.</p> |

자기 상관의 가설 검정에 대한 자세한 내용은 부록 D를 참조하십시오.

참고 문헌

- 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.
- Montgomery, D. (2001). *Introduction to statistical quality control, 4th edition*. John Wiley & Sons.
- 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 Company.
- Wheeler, D.J. (2004). *Advanced topics in statistical process control. The power of Shewhart's charts, 2nd 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.

부록 A: 정규성

시뮬레이션 A1: 비정규성이 잘못된 경고 비율에 미치는 영향

비정규 데이터가 I 관리도 및 Xbar 관리도에 어떤 영향을 미치는지 조사하기 위해 Minitab에서는 시뮬레이션을 수행하여 비정규 데이터 분포와 연관된 잘못된 경고 비율을 평가했습니다. Minitab에서는 실제로 일반적인 비정규 분포인 오른쪽으로 치우친 분포 및 두꺼운 꼬리를 갖는 대칭 분포에 주목했습니다. 구체적으로, 3개의 치우친 분포(df=3, 5 및 10의 카이-제곱 분포) 및 2개의 두꺼운 꼬리를 갖는 분포(df=3 및 5의 t 분포)를 조사했습니다. Minitab에서는 초기 10,000개의 부분군 집합을 사용하여 관리 한계를 설정했습니다. 추가 2,500개 부분군에 대한 잘못된 경고의 백분율을 기록했습니다. 그런 다음 10,000번의 반복을 수행하고 특수 원인 검정 1과 검정 2를 사용한 잘못된 경고의 평균 백분율을 계산했습니다. 결과는 표 1에 나와 있습니다.

표 1 I 관리도(n=1) 및 Xbar 관리도(n=2, 3, 4 또는 5)에 대한 검정 1, 검정 2의 잘못된 경고 비율(%)

| 분포 | 부분군 크기 | | | | |
|--------------|------------|------------|------------|------------|------------|
| | n = 1 | n = 2 | n = 3 | n = 4 | n = 5 |
| Normal (0,1) | 0.27, 0.39 | 0.27, 0.39 | 0.27, 0.39 | 0.27, 0.39 | 0.27, 0.39 |
| Chi (3) | 2.06, 1.17 | 1.18, 0.79 | 0.98, 0.62 | 0.86, 0.57 | 0.77, 0.53 |
| Chi (5) | 1.54, 0.83 | 0.93, 0.60 | 0.77, 0.53 | 0.67, 0.50 | 0.61, 0.47 |
| Chi (10) | 1.05, 0.60 | 0.67, 0.50 | 0.56, 0.46 | 0.50, 0.44 | 0.45, 0.43 |
| t (3) | 2.18, 0.39 | 1.30, 0.39 | 1.22, 0.39 | 1.16, 0.39 | 1.11, 0.39 |
| t (5) | 1.43, 0.39 | 0.93, 0.39 | 0.80, 0.39 | 0.71, 0.39 | 0.66, 0.39 |

표 1의 각 셀은 각각 검정 1과 검정 2에 대해 계산된 잘못된 경고 비율을 나타냅니다. 예를 들어, 표준 정규 분포(Normal (0, 1)) 및 I 관리도(n=1)와 연관된 셀은 데이터가 정규 분포를

따르는 경우 I 관리도의 잘못된 경고 비율이 검정 1에 대해 0.27%, 검정 2에 대해 0.39%라는 것을 나타냅니다.

정규 및 비정규 분포 간에 \bar{X} 관리도($n=2, 3, 4, 5$)에 대한 잘못된 경고 비율을 비교해 보면 정규성 가정이 위반될 경우 잘못된 경고 비율이 유의하게 증가하지 않는다는 것을 알 수 있습니다. 심하게 치우친 분포(카이-제곱, 3) 및 두꺼운 꼬리를 갖는 분포($t, 3$)에서 부분군 크기가 2인 경우에도 검정 1과 검정 2에 대한 잘못된 경고 비율을 합한 값이 (정규 분포의 0.7%에 비교하여) 2.0%보다 작습니다. 따라서 실제로 \bar{X} 관리도가 정규성 가정의 위반에 대해 크게 영향을 받지 않는다고 결론을 내립니다.

I 관리도의 경우 표 1을 보면 분포가 심하게 치우친 경우(카이-제곱, 3) 검정 1과 검정 2에 대한 잘못된 경고 비율을 합하면 약 3.2%로, 이 잘못된 경고 비율은 데이터가 정규 분포를 따르는 경우 예상되는 잘못된 경고 비율보다 거의 5배 더 높습니다. 두꺼운 꼬리를 갖는 분포($t, 3$)의 경우 검정 1과 검정 2에 대한 잘못된 경고 비율을 합하면 2.6%로 데이터가 정규 분포를 따르는 경우 예상되는 잘못된 경고 비율보다 거의 4배 더 높습니다.

따라서 I-MR 관리도에 대한 데이터의 정규성을 먼저 검정해야 합니다. 정규성에 대한 Anderson-Darling 적합도 검정의 귀무 가설이 기각되는 경우(데이터가 정규 분포를 따르지 않는다는 통계적으로 유의한 증거가 있는 경우) Box-Cox 변환을 사용하여 데이터를 정규화할 수 있습니다. 그러나 Box-Cox 변환은 데이터가 오른쪽으로 치우친 경우에만 효과적입니다. 다른 비정규 데이터의 경우에는 최적의 방법을 결정하기 위해 추가 조사가 필요합니다. 또한 Anderson-Darling 검정 및 Box-Cox 변환은 극단 관측치의 영향을 받기 때문에 분석을 수행하기 전에 특수 원인을 가진 것으로 알려진 점은 제외해야 합니다.

부록 B: 안정성

시뮬레이션 B1: 검정 1에 검정 2를 추가할 경우 민감도에 미치는 영향

검정 1은 한 점이 중심선으로부터 3 표준 편차 범위 밖에 있을 때 신호를 보내 관리 이탈 상태에 있는 점을 탐지합니다. 검정 2는 9개의 연속된 점이 중심선으로부터 같은 쪽에 있을 때 신호를 보내 평균의 이동을 탐지합니다.

Minitab에서는 검정 2를 검정 1과 함께 사용하면 평균 관리도(I 관리도 및 Xbar 관리도)의 민감도가 개선되는지 여부를 확인하기 위해 정규 (0, SD) 분포에 대한 관리 한계를 설정했습니다. 분포의 평균을 표준 편차의 배수 단위로 이동한 후 10,000번의 각 반복 횟수에 대한 신호를 탐지하기 위해 필요한 부분군의 수를 기록했습니다. 결과는 표 2에 나와 있습니다.

표 2 검정 1 실패(검정 1), 검정 2 실패(검정 2) 또는 검정 1 및 검정 2 실패(검정 1 또는 2)까지 평균 부분군의 수. 평균은 표준 편차(SD)의 배수로 이동되었으며 부분군 크기 (n) = 1, 3 및 5에 대해 시뮬레이션이 수행되었습니다.

| 이동 | n=1 | | | n=3 | | | n=5 | | |
|--------|------|------|-----------|------|------|-----------|------|------|-----------|
| | 검정 1 | 검정 2 | 검정 1 또는 2 | 검정 1 | 검정 2 | 검정 1 또는 2 | 검정 1 | 검정 2 | 검정 1 또는 2 |
| 0.5 SD | 154 | 84 | 57 | 60 | 31 | 22 | 33 | 19 | 14 |
| 1 SD | 44 | 24 | 17 | 10 | 11 | 7 | 4 | 10 | 4 |
| 1.5 SD | 15 | 13 | 9 | 3 | 9 | 3 | 1.6 | 9 | 1.6 |
| 2 SD | 6 | 10 | 5 | 1.5 | 9 | 1.5 | 1.1 | 9 | 1.1 |

I 관리도(n= 1)에 대한 결과에 표시된 대로 0.5 표준 편차 이동을 탐지하기 위해 검정 1만 사용하는 경우 평균 154개의 부분군이 필요한 것에 비해 두 검정 모두 사용하는 경우(검정 1 또는 2 열)에는 평균 57개의 부분군이 필요합니다. 이와 유사하게 두 검정을 모두 사용하면 Xbar 관리도(n = 3, n = 5)에 대한 민감도가 증가합니다. 예를 들어, 부분군 크기가 3인 경우, 검정 1과 검정 2를 모두 사용하면 0.5 표준 편차를 탐지하기 위해 평균 22개의 부분군이

필요하지만, 검정 1만 사용하는 경우에는 0.5 표준 편차를 탐지하기 위해 60개의 부분군이 필요합니다. 따라서 두 검정을 모두 사용하면 평균의 작은 이동을 탐지하기 위한 민감도가 유의하게 증가합니다. 이동 크기가 증가하는 경우에는 검정 2를 추가해도 민감도가 유의하게 증가하지 않습니다.

시뮬레이션 B2: 검정 7이 계층화(부분군의 여러 변동 원인)를 얼마나 효과적으로 탐지하는가?

검정 7은 일반적으로 12-15개의 연속된 점이 중심선으로부터 1 표준 편차 내에 있을 때 실패 신호를 보냅니다. 보조 도구에서는 데이터 내 부분군의 수를 바탕으로 필요한 점의 수를 조정하는 수정된 규칙을 사용합니다. Minitab에서는 $k = (\text{부분군 수} * 0.33)$ 을 설정하고 표 3에 표시된 대로 검정 7 실패에 필요한 연속된 점의 수를 정의합니다.

표 3 검정 7의 실패에 필요한 연속된 점의 수

| $k = (\text{부분군 수}) \times 0.33$ | 필요한 점의 수 |
|----------------------------------|-------------|
| $k < 12$ | 12 |
| $k \geq 12$ 및 $k \leq 15$ | 정수 $\geq k$ |
| $k > 15$ | 15 |

관리 한계를 설정하는 일반 시나리오를 사용하여, 위의 기준에 따라 검정 7에서 실패 신호를 보낼 확률을 확인하기 위해 시뮬레이션을 수행했습니다. 구체적으로는 데이터에서 관리 한계를 추정하는 동안 계층화를 탐지하는 규칙을 평가하고자 했습니다.

Minitab에서는 표준 편차(SD)의 정규 분포에서 크기 n의 부분군 m개를 랜덤하게 선택했습니다. 각 부분군 내 점 중 절반의 평균은 0이고 나머지 절반의 평균은 SD(0 SD, 1 SD 또는 2 SD)만큼 이동했습니다. 10,000번 반복을 수행하고 표 4에 표시된 대로 1번 이상 검정 7 신호를 보여주는 관리도의 비율을 기록했습니다.

표 4 검정 7에서 1번 이상 신호를 보낸 관리도의 비율

| 부분군 수 | | m = 50 | m = 75 | m = 25 | m = 38 | m = 25 |
|--------|------|---------|---------|---------|---------|---------|
| 부분군 크기 | | n = 2 | n = 2 | n = 4 | n = 4 | n = 6 |
| 검정 | | 연속된 15개 | 연속된 15개 | 연속된 12개 | 연속된 13개 | 연속된 12개 |
| 이동 | 0 SD | 5% | 8% | 7% | 8% | 7% |
| | 1 SD | 23% | 33% | 17% | 20% | 15% |
| | 2 SD | 83% | 94% | 56% | 66% | 50% |

표의 첫 번째 이동(이동 = 0 SD) 행에 표시된 대로, 계층화가 없는 경우 검정 7 실패가 1번 이상 있는 관리도의 비율은 상대적으로 작습니다. 그러나 계층화(이동 = 1 SD 또는 이동 = 2 SD)가 있는 경우 검정 7 실패가 1번 이상 있는 관리도의 비율이 훨씬 더 높습니다. 이러한 방식으로, 검정 7은 관리 한계가 추정되는 단계에서 계층화를 식별할 수 있습니다.

부록 C: 데이터 양

공식 C1: 관리 상한에 대한 95% CI를 기준으로 필요한 관측치의 수

데이터가 정규 분포를 따르고 표본 추출 변동성으로 인한 오차가 없을 경우 관리 상태에 있는 공정에 대해 관리 상한 위에 있는 관측치의 백분율은 0.135%입니다. 관측치의 수가 잘못된 경고 비율을 낮게 유지하기에 충분한지 확인하기 위해 Minitab에서는 Bischak(2007)을 따라 다음과 같이 관리 상한에 대한 95% 신뢰 하한을 계산합니다.

$$UCL_L = \hat{\mu} + \frac{h}{w_p}$$

여기서

$\hat{\mu}$ = 중심선의 추정치

$h = 3 \sigma$

$$w_p = 1 - \frac{z_p \sqrt{1 - c_4^2}}{c_4 \sqrt{\frac{n}{k}}}$$

c_4 = Xbar 관리도에 대한 불편화 상수

n = 관측치 수

k = 부분군 크기

z_p = 평균=0이고 표준 편차=1인 정규 분포의 경우 $p = 0.95$ 에 대한 역누적확률(inverse cdf)

Minitab에서는 관리 상한 위의 잘못된 경고 비율이 0.5%가 되도록 $\frac{h}{w_p} = 2.65 \sigma$ 를 설정하고 n 을 구합니다.

I 관리도의 경우 길이 2의 평균 이동 범위를 사용하여 σ 를 추정하며, 따라서 부분군 크기가 2인 Xbar 관리도에서 관측치 수를 사용하여 필요한 데이터의 양을 결정합니다. 이러한 계산을 바탕으로 다양한 부분군 크기에 필요한 관측치의 수가 표 5에 나와 있습니다.

표 5 다양한 부분군 크기에 필요한 관측치의 수

| 부분군 크기 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|--------|-----|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 관측치 수 | 114 | 114 | 84 | 72 | 70 | 66 | 63 | 64 | 63 | 60 | 66 | 60 | 65 | 56 | 60 |

참고 관측치의 수는 부분군 크기가 증가함에 따라 감소해야 합니다. 그러나 표 8에 이 규칙에 대한 예외가 나와 있습니다. 이러한 예외는 필요한 전체 관측치의 수를 계산하기 위해 각 부분군의 관측치 수를 곱하기 전에 부분군의 수를 다음 정수로 올림하기 때문에 발생합니다.

표 5의 결과를 보면 필요한 전체 관측치의 수가 부분군 크기가 1이거나 2인 경우를 제외하고, 모든 부분군 크기에 대해 100보다 작거나 같다는 것을 알 수 있습니다. 그러나 부분군 크기가 1이나 2더라도 관측치 수가 100개인 경우 잘못된 경고 비율이 약 1.1%에 지나지 않습니다. 따라서 모든 부분군 크기에 대해 100개의 관측치가 효과적인 컷오프 값입니다.

위의 분석에서는 각 부분군의 일반 원인 변동의 양이 같다고 가정합니다. 실제로는 여러 시점에서 수집된 데이터의 일반 원인 변동성의 양이 다를 수 있습니다. 따라서 공정 변동의 대표적인 추정치를 얻을 확률을 높이기 위해 필요한 것보다 많은 시점에서 공정으로부터 표본을 추출할 수도 있습니다.

부록 D: 자기 상관

시뮬레이션 D1: 자기 상관이 잘못된 경고 비율에 미치는 영향

Minitab에서는 다음과 같은 모형을 사용하여 자기 상관된 공정을 정의했습니다.

$$x_t = \mu + \phi (x_{t-1} - \mu) + \varepsilon_t$$

여기서

x_t = 시간 t의 관측치

μ = 공정의 평균

ϕ = 연속된 두 점 간의 상관

ε_t 시간 t의 오차항, $\varepsilon \sim \text{Normal}(0, 1)$

간단하게 하기 위해, 시차 1 자기 상관이 시차 ≥ 2 에 대한 자기 상관보다 클 가능성이 높기 때문에 시차 1 자기 상관만을 고려했습니다.

위의 모형을 사용하여 Minitab에서는 $\phi = 0.2, 0.4, 0.5, 0.6, 0.8$ 및 세 가지 부분군 크기($n=1, 3, 5$)에서 시뮬레이션을 수행했습니다. 초기 10,000개의 부분군을 사용하여 관리 한계를 설정한 후 추가 2,500개 부분군에 대한 잘못된 경고의 백분율을 기록했습니다. 10,000번의 반복을 수행하고 각 자기 상관 및 부분군 크기의 조합에 대해 검정 1과 검정 2를 사용하여 잘못된 경고의 평균 백분율을 기록했습니다. 결과는 아래 표 6에 나와 있습니다.

표 6 자기 상관(ϕ) 및 부분군 크기(n)의 공정에 대한 검정 1, 검정 2의 잘못된 경고의 평균 백분율

| | ϕ | | | | |
|--------|------------|------------|------------|-------------|--------------|
| 부분군 크기 | 0.2 | 0.4 | 0.5 | 0.6 | 0.8 |
| n = 1 | 0.73, 1.08 | 2.01, 2.99 | 3.38, 4.98 | 5.77, 8.34 | 17.94, 23.83 |
| n = 3 | 1.48, 0.57 | 5.35, 0.98 | 9.07, 1.39 | 14.59, 2.19 | 33.91, 7.91 |
| n = 5 | 1.49, 0.50 | 5.48, 0.68 | 9.40, 0.86 | 15.24, 1.18 | 35.52, 3.84 |

ϕ 값이 증가함에 따라 잘못된 경고 비율이 증가합니다. 실제로, 자기 상관 수준이 중간 정도인 경우에도($\phi = 0.4$) 잘못된 경고 비율이 유의하게 증가합니다. 자기 상관 ≥ 0.4 인 경우 관리도에 관리도의 결과가 거의 무의미할 정도로 높은 비율의 잘못된 경고가 표시됩니다.

공식 D1: 자기 상관 검정

데이터 내 자기 상관의 수준 ϕ 을 결정하기 위해 Minitab에서는 Bowerman(1979)에 따라 다음과 같은 검정 통계량을 사용하여 귀무 가설 $\phi = r$ 대 대립 가설 $\phi > r$ 를 평가하기 위한 검정을 수행합니다.

$$z = \frac{(\hat{\phi} - r)}{se_{\hat{\phi}}}$$

여기서

$$\hat{\phi} = \frac{1}{m \hat{\sigma}} \sum_{i=1}^{m-1} (x_i - \mu)(x_{i+1} - \mu)$$

m = 관측치 수

$\hat{\sigma}$ = 표본 분산

$$se_{\hat{\phi}} = \sqrt{m}$$

Xbar 관리도에서 데이터의 자기 상관을 계산할 때는 부분군의 마지막 점과 다음 부분군의 첫 번째 점 간의 종속성을 포함하지 않습니다. 그 이유는 이러한 비교가 자기 상관의 과소 추정으로 이어지기 때문입니다. Minitab에서는 Bowerman(1979)에 따라 표준 정규 분포를 기반으로 p-값을 계산합니다.

© 2015, 2017 Minitab Inc. All rights reserved.

Minitab®, Quality. Analysis. Results.® and the Minitab® logo are all registered trademarks of Minitab, Inc., in the United States and other countries. See minitab.com/legal/trademarks for more information.