

MINITAB-ASSISTENT – WHITE PAPER

Dieses White Paper ist Teil einer Reihe von Veröffentlichungen, welche die Forschungsarbeiten der Minitab-Statistiker erläutern, in deren Rahmen die im Assistenten der Minitab Statistical Software verwendeten Methoden und Datenprüfungen entwickelt wurden.

Regelkarten für Variablen

Übersicht

Regelkarten werden verwendet, um regelmäßig einen Prozess zu überwachen und somit zu ermitteln, ob dieser unter Kontrolle ist. Der Minitab-Assistent enthält zwei der am häufigsten verwendeten Regelkarten für stetige Daten:

- X-quer/R- und X-quer/S-Karten. Diese Regelkarten werden verwendet, wenn Daten in Teilgruppen erfasst werden. Minitab verwendet die zusammengefasste Standardabweichung, um die Standardabweichung innerhalb der Teilgruppen zu schätzen. Die R-Karte ermöglicht eine effektive Schätzung der Streuung für Teilgruppen mit einer Größe von bis zu ungefähr 10 (AIAG, 1995; Montgomery, 2001). Bei größeren Teilgruppengrößen ermöglicht eine S-Karte eine bessere Schätzung der individuellen Standardabweichung innerhalb der Teilgruppen. Um früheren Minitab-Konventionen zu folgen und konservativ vorzugehen, empfehlen wir eine S-Karte, wenn die Teilgruppengröße größer als 8 ist. Für Teilgruppengrößen kleiner oder gleich 8 bieten die R- und die S-Karte ähnliche Ergebnisse.
- Einzelwertkarte mit gleitender Spannweite (I/MR). Diese Regelkarte wird verwendet, wenn keine Teilgruppen vorhanden sind. Minitab verwendet eine Methode der Durchschnitte der gleitenden Spannweiten der Länge 2, um die Standardabweichung zu schätzen.

Die Eingriffsgrenzen für eine Regelkarte werden normalerweise in der Kontrollphase eines Six Sigma-Projekts festgelegt. Eine gute Regelkarte sollte ausreichend empfindlich sein, um das Vorhandensein einer Ausnahmebedingung schnell zu signalisieren. Diese Empfindlichkeit kann ausgewertet werden, indem die durchschnittliche Anzahl der Teilgruppen berechnet wird, die erforderlich ist, um eine Ausnahmebedingung zu signalisieren. Außerdem signalisiert eine gute Regelkarte selten einen „falschen Alarm“, wenn der Prozess unter Kontrolle ist. Die Rate falscher Alarme kann ausgewertet werden, indem der Prozentsatz von Teilgruppen berechnet wird, die als „außer Kontrolle“ betrachtet werden, wenn der Prozess unter Kontrolle ist.

Im Allgemeinen funktionieren Regelkarten optimal, wenn jede Beobachtung aus einer Normalverteilung stammt und unabhängig ist und in den Teilgruppen ausschließlich eine Streuung durch gewöhnliche Ursachen vorhanden ist. Daher werden für die Auswertung des Assistenten automatisch die folgenden Datenprüfungen durchgeführt, um diese Bedingungen zu untersuchen:

- Vorliegen einer Normalverteilung
- Stabilität
- Umfang der Daten
- Vorliegen von korrelierten Daten

Im vorliegenden White Paper wird untersucht, wie sich eine Regelkarte für Variablen verhält, wenn diese Bedingungen variieren. Außerdem wird beschrieben, wie wir einen Satz von Richtlinien aufgestellt haben, um die Anforderungen für diese Bedingungen auszuwerten.

Datenprüfungen

Vorliegen einer Normalverteilung

Regelkarten basieren nicht auf der Annahme, dass die Prozessdaten normalverteilt sind, aber die in den Tests auf Ausnahmebedingungen verwendeten Kriterien basieren auf dieser Annahme. Wenn die Daten eine starke Schiefe aufweisen oder zu viele Daten an den Enden der Verteilung liegen („stärker besetzte Randbereiche“), sind die Testergebnisse möglicherweise nicht genau. Die Regelkarte könnte dann zum Beispiel eine höhere Rate falscher Alarme signalisieren als erwartet.

Zielstellung

Wir haben die Auswirkung nicht normalverteilter Daten auf die X-quer-Karte und die I-Karte untersucht. Es sollte ermittelt werden, wie sich eine fehlende Normalverteilung auf die Rate falscher Alarme auswirkt. Insbesondere sollte festgestellt werden, ob die Regelkarte bei nicht normalverteilten Daten signifikant häufiger auf Punkte hinweist, die außer Kontrolle sind, wenn der Prozess tatsächlich unter Kontrolle ist (falsche Alarme).

Methode

Es wurden Simulationen mit 10.000 Teilgruppen und verschiedenen starken Abweichungen von der Normalverteilung durchgeführt, und der Prozentsatz falscher Alarme wurde aufgezeichnet. Simulationen ermöglichen es, verschiedene Bedingungen zu testen, um die Auswirkungen einer fehlenden Normalverteilung zu ermitteln. Hierfür wurden die Verteilung mit Rechtsschiefe und symmetrische Verteilungen mit stärker besetzten Randbereichen gewählt, da es sich hierbei um Nicht-Normalverteilungen handelt, in der Praxis häufig auftreten. Weitere Informationen finden Sie in Anhang A.

Ergebnisse

X-QUER-KARTE (TEILGRUPPENGROÖBE ≥ 2)

Die Simulation hat gezeigt, dass sich die Rate falscher Alarme nicht signifikant erhöht, wenn die Daten nicht normalverteilt sind, sofern die Teilgruppengröße 2 oder mehr beträgt. Aufgrund dieses Ergebnisses wird das Vorliegen einer Normalverteilung für die X-quer/R-Karte und die X-quer/S-Karte nicht überprüft. Selbst wenn die Daten eine starke Schiefe oder stärker besetzte Randbereiche aufweisen, beträgt die Rate falscher Alarme für Test 1 und Test 2 weniger als 2 % und liegt somit nicht wesentlich über der Rate falscher Alarme von 0,7 % bei Vorliegen einer Normalverteilung.

I-KARTE (TEILGRUPPENGROÖBE = 1)

Die Simulation hat gezeigt, dass die I-Karte empfindlich auf nicht normalverteilte Daten reagiert. Wenn die Daten nicht normalverteilt sind, verursacht die I-Karte eine Rate falscher Alarme, die 4 bis 5 Mal höher ist als bei normalverteilten Daten. Um die Empfindlichkeit der I-Karte gegenüber nicht normalverteilten Daten zu berücksichtigen, geht der Assistent wie folgt vor:

- Es wird ein Anderson-Darling-Test durchgeführt, wenn die Daten eine starke Nicht-Normalverteilung aufweisen könnten, worauf durch eine größere Zahl von Punkten außer Kontrolle als erwartet hingewiesen wird (d. h., 2 oder mehr Punkte und 2 % oder mehr der Punkte befinden sich außerhalb der Eingriffsgrenzen).
- Wenn der Anderson-Darling-Test darauf hinweist, dass die Daten nicht normalverteilt sind, transformiert der Assistent die Daten mit Hilfe des optimalen Box-Cox-Lambdas. Für die transformierten Daten wird ein Anderson-Darling-Test durchgeführt. Wenn der Prozess normalerweise nicht normalverteilte Daten erzeugt und der Test die Nullhypothese (dass die Daten normalverteilt sind) nicht zurückweist, schlägt der Assistent vor, die transformierten Daten zu verwenden.

Die Box-Cox-Transformation ist nur für nicht normalverteilte Daten mit Rechtsschiefe wirksam. Wenn die Transformation für die nicht normalverteilten Daten nicht wirksam ist, müssen Sie möglicherweise andere Optionen in Betracht ziehen. Da sich darüber hinaus extreme Beobachtungen auf den Anderson-Darling-Test und die Box-Cox-Transformation auswirken, sollten Sie Punkte mit bekannten Ausnahmebedingungen auslassen, bevor Sie die Daten transformieren.

Für die Prüfung auf das Vorliegen einer Normalverteilung werden in der Auswertung des Assistenten für die I-Karte die folgenden Statusindikatoren angezeigt:

Status	Bedingung
	Wenn < 2 Punkte oder < 2 % der Punkte außerhalb der Eingriffsgrenzen auf der I-Karte liegen oder Der Anderson-Darling-Test auf Normalverteilung für die Daten oder die mit der Box-Cox-Transformation transformierten Daten weist die Nullhypothese (dass die Daten normalverteilt sind) bei Alpha = 0,01 nicht zurück.
	Wenn die oben aufgeführten Bedingungen nicht zutreffen.

Stabilität

Für Regelkarten für Variablen können acht Tests durchgeführt werden, um die Stabilität des Prozesses auszuwerten. Durch die gleichzeitige Verwendung aller Tests wird die Empfindlichkeit der Regelkarte gesteigert. Es ist jedoch wichtig, den Zweck und Zusatznutzen jedes Tests abzuwägen, da mit der Anzahl unterschiedlicher Tests auch die Rate falscher Alarme ansteigt.

Zielstellung

Wir wollten ermitteln, welche der acht Tests auf Stabilität in die Regelkarten für Variablen im Assistenten aufgenommen werden sollten. Das erste Ziel war, die Tests zu bestimmen, mit denen die Empfindlichkeit gegenüber fehlender Beherrschung signifikant gesteigert wird, ohne gleichzeitig zu einem signifikanten Anstieg der falschen Alarme zu führen. Das zweite Ziel war, die Regelkarte einfach und praktisch in der Handhabung zu gestalten. Unsere Untersuchung hat sich auf die Tests für die X-quer-Karte und die I-Karte konzentriert. Für die

R-, S- und MR-Karten verwenden wir nur Test 1, der darauf hinweist, wenn ein Punkt außerhalb der Eingriffsgrenzen liegt.

Methoden

Es wurden Simulationen durchgeführt und die Fachliteratur ausgewertet, um herauszufinden, wie sich Kombinationen von Tests auf Stabilität, die Empfindlichkeit und die Rate falscher Alarme der Regelkarten auswirken. Außerdem wurde ausgewertet, wie gut Ausnahmebedingungen bei den Tests erkannt werden. Einzelheiten zu den Methoden, die für jeden Test verwendet wurden, finden Sie im Abschnitt „Ergebnisse“ unten und in Anhang B.

Ergebnisse

Wir haben ermittelt, dass Tests 1, 2 und 7 am nützlichsten zum Auswerten der Stabilität der X-quer-Karte und der I-Karte sind:

TEST 1: IDENTIFIZIERT PUNKTE AUßERHALB DER EINGRIFFSGRENZEN

Test 1 identifiziert Punkte > 3 Standardabweichungen von der Mittellinie entfernt. Test 1 ist allgemein als notwendig zum Erkennen einer fehlenden Beherrschung anerkannt. Die Rate falscher Alarme beträgt lediglich 0,27 %.

TEST 2: IDENTIFIZIERT SHIFTS DES MITTELWERTS

Test 2 gibt ein Signal aus, wenn 9 aufeinander folgende Punkte auf der gleichen Seite der Mittellinie liegen. Es wurde eine Simulation mit 4 unterschiedlichen Mittelwerten durchgeführt, die jeweils auf ein Vielfaches der Standardabweichung festgelegt waren, und die Anzahl von Teilgruppen ermittelt, die erforderlich waren, um ein Signal zu erkennen. Dabei wurden die Eingriffsgrenzen entsprechend der Normalverteilung festgelegt. Wir haben festgestellt, dass die Einbeziehung von Test 2 die Empfindlichkeit der Regelkarte gegenüber kleinen Shifts im Mittelwert signifikant steigert. Wenn Test 1 und Test 2 zusammen verwendet werden, sind signifikant weniger Teilgruppen erforderlich, um einen kleinen Shift im Mittelwert zu erkennen, als bei Test 1 allein. Daher ist das Einbeziehen von Test 2 hilfreich, um gängige Situationen zu erkennen, die außer Kontrolle sind, und die Empfindlichkeit ausreichend zu steigern, um einen geringfügigen Anstieg der Rate falscher Alarme zu rechtfertigen.

TEST 7: IDENTIFIZIERT EINGRIFFSGRENZEN, DIE ZU WEIT AUSEINANDER LIEGEN

Test 7 gibt ein Signal aus, wenn 12–15 aufeinander folgende Punkte innerhalb 1 Standardabweichung von der Mittellinie liegen. Test 7 wird nur für die X-quer-Karte verwendet, wenn die Eingriffsgrenzen aus den Daten geschätzt werden. Wenn dieser Test fehlschlägt, ist die Ursache häufig eine systemische Streuungsursache (Schichtung) in einer Teilgruppe, was häufig daran liegt, dass keine rationalen Teilgruppen gebildet wurden. Da rationale Teilgruppen unerlässlich sind, damit eine fehlende Beherrschung mit der Regelkarte akkurat erkannt werden kann, verwendet Minitab einen modifizierten Test 7, wenn die Eingriffsgrenzen aus den Daten geschätzt werden. Test 7 signalisiert ein Fehlschlagen, wenn die Anzahl der aufeinander folgenden Punkte je nach Anzahl der Teilgruppen zwischen 12 und 15 liegt:

k = (Anzahl der Teilgruppen) x 0,33	Erforderliche Punkte
$k < 12$	12
$k \geq 12$ und $k \leq 15$	Ganze Zahl $\geq k$
$k > 15$	15

Nicht im Assistenten enthaltene Tests

TEST 3: K AUFEINANDER FOLGENDE PUNKTE, ALLE ZU- ODER ABNEHMEND

Test 3 soll Drifts im Prozessmittelwert erkennen (Davis and Woodall, 1988). Wenn Test 3 jedoch zusätzlich zu Test 1 und Test 2 verwendet wird, steigert es die Empfindlichkeit der Regelkarte zum Erkennen von Drifts im Prozessmittelwert nicht signifikant. Da wir bereits auf der Grundlage der Simulationsergebnisse entschieden haben, Test 1 und Test 2 zu verwenden, würde Test 3 die Regelkarte nicht signifikant nützlicher machen.

TEST 4: K AUFEINANDER FOLGENDE PUNKTE, ABWECHSELND AUF- UND ABWÄRTS

Obwohl dieses Muster in der Praxis auftreten kann, empfehlen wir, dass Sie nach ungewöhnlichen Trends oder Mustern suchen, statt auf ein spezifisches Muster zu testen.

TEST 5: K VON K+1 PUNKTEN > 2 STANDARDABWEICHUNGEN VON DER MITTELLINIE ENTFERNT

Um die Regelkarte einfach zu halten, haben wir diesen Test ausgenommen, weil er nicht eindeutig Ausnahmebedingungen identifizieren kann, die in der Praxis häufig auftreten.

TEST 6: K VON K+1 PUNKTEN > 1 STANDARDABWEICHUNG VON DER MITTELLINIE ENTFERNT

Um die Regelkarte einfach zu halten, haben wir diesen Test ausgenommen, weil er nicht eindeutig Ausnahmebedingungen identifizieren kann, die in der Praxis häufig auftreten.

TEST 8: K AUF EINANDER FOLGENDE PUNKTE > 1 STANDARDABWEICHUNG VON DER MITTELLINIE ENTFERNT (BEIDE SEITEN)

Um die Regelkarte einfach zu halten, haben wir diesen Test ausgenommen, weil er nicht eindeutig Ausnahmbedingungen identifizieren kann, die in der Praxis häufig auftreten.

Für die Prüfung auf Stabilität werden in der Auswertung des Assistenten die folgenden Statusindikatoren angezeigt:

Status	Bedingung
	<p>Keine Testfehler auf der Regelkarte für den Mittelwert (I-Karte oder X-quer-Karte) und der Regelkarte für die Streuung (MR-, R- oder S-Karte).</p> <p>Für die einzelnen Karten werden die folgenden Tests verwendet:</p> <ul style="list-style-type: none">• I-Karte: Test 1 und Test 2.• X-quer-Karte: Test 1, Test 2 und Test 7. Test 7 wird nur durchgeführt, wenn die Eingriffsgrenzen aus den Daten geschätzt werden.• MR-, R- und S-Karten: Test 1.
	<p>Wenn die oben genannte Bedingung nicht zutrifft.</p>

Umfang der Daten

Wenn für die Eingriffsgrenzen keine bekannten Werte vorliegen, müssen diese aus den Daten geschätzt werden. Um genaue Schätzungen der Eingriffsgrenzen zu erhalten, müssen genügend Daten vorliegen. Wenn der Umfang der Daten nicht ausreicht, können die Eingriffsgrenzen aufgrund von Streuung zwischen den Stichproben stark von den „echten“ Eingriffsgrenzen abweichen. Sie können die Anzahl der Beobachtungen erhöhen, um die Genauigkeit der Eingriffsgrenzen zu verbessern.

Zielstellung

Wir haben untersucht, wie viele Beobachtungen benötigt werden, um präzise Eingriffsgrenzen zu erhalten. Es sollte ermittelt werden, welchen Umfang die Daten aufweisen müssen, um sicherzustellen, dass die Rate falscher Alarme aufgrund von Test 1 mit einer Konfidenz von 95 % nicht über 1 % liegt.

Methode

Wenn die Daten normalverteilt sind und kein Fehler aufgrund der Stichprobenstreuung auftritt, beträgt der Prozentsatz der Punkte über der oberen Eingriffsgrenze 0,135 %. Um zu ermitteln, ob die Anzahl der Beobachtungen ausreichend ist, wurde die von Bischak (2007) beschriebene Methode angewandt, um sicherzustellen, dass die Rate falscher Alarme aufgrund von Punkten über der oberen Eingriffsgrenze mit einer Konfidenz von 95 % nicht höher als 0,5 % ist. Aufgrund der Symmetrie der Eingriffsgrenze führt diese Methode zu einer Gesamtrate falscher Alarme von 1 % aufgrund von Test 1. Weitere Informationen finden Sie in Anhang C.

Ergebnisse

Es wurde ermittelt, dass für fast alle Teilgruppengrößen insgesamt 100 Beobachtungen ausreichen, um genaue Eingriffsgrenzen zu erhalten. Obwohl die Teilgruppengrößen von 1 und 2 eine geringfügig höhere Anzahl von Beobachtungen erfordert haben, lag die Rate falscher Alarme bei 100 Beobachtungen weiterhin relativ niedrig (1,1 %). Aus Gründen der Einfachheit wird daher der Trennwert von 100 Gesamtbeobachtungen für alle Teilgruppengrößen verwendet.

Auf der Grundlage dieser Ergebnisse zeigt der Assistent für die Prüfung des Umfangs der Daten in der Auswertung die folgenden Statusindikatoren an:

Status	Bedingung
	Anzahl der Beobachtungen ≥ 100 .
	Anzahl der Beobachtungen < 100 .

Vorliegen von korrelierten Daten

Die Autokorrelation ist ein Maß für die Abhängigkeit zwischen den Datenpunkten, die im Verlauf eines Zeitraums erfasst werden. Die meisten Prozessdaten weisen zumindest einen geringen Grad an Autokorrelation auf. Eine mäßige oder hohe Autokorrelation kann zu falschen Testergebnissen führen. In der Regel weisen autokorrelierte Daten eine positive Autokorrelation auf, die die Streuung innerhalb der Teilgruppen reduzieren und zu einer höheren Rate falscher Alarme führen kann.

Zielstellung

Wir haben die Beziehung zwischen der Autokorrelation und der Rate falscher Alarme untersucht. Als Zielstellung sollte das Niveau von Autokorrelation ermittelt werden, das zu einer nicht akzeptablen Rate falscher Alarme führt. Aus Gründen der Einfachheit wurde die Autokorrelation für Lag 1 berücksichtigt, da diese wahrscheinlich größer als die Autokorrelation für Lags ≥ 2 ist.

Methode

Unter Verwendung eines Standardmodells für einen autokorrelierten Prozess wurden Simulationen mit $\phi = 0,2; 0,4; 0,5; 0,6$ und $0,8$ (ϕ ist die Autokorrelation für Lag 1) für drei Teilgruppengrößen ($n = 1, 3$ und 5) durchgeführt. Es wurde ein erster Satz von 10.000 Teilgruppen verwendet, um die Eingriffsgrenzen festzulegen. Dann wurde der Prozentsatz falscher Alarme für weitere 2.500 Teilgruppen aufgezeichnet. Es wurden 10.000 Iterationen durchgeführt, und es wurde der durchschnittliche Prozentsatz falscher Alarme aufgezeichnet. Weitere Informationen finden Sie in Anhang D.

Ergebnisse

Die Simulationen haben gezeigt, dass selbst mäßige Niveaus von Autokorrelation die Rate falscher Alarme signifikant erhöhen. Bei einer Autokorrelation $\geq 0,4$ ist die Rate falscher Alarme sehr hoch, und die Regelkarte verliert ihre Aussagekraft. Um dieses Problem zu beheben, führt der Assistent einen Test auf Autokorrelation durch, wenn die Daten autokorreliert sein könnten, worauf durch eine größere Zahl von Punkten außer Kontrolle als erwartet hingewiesen wird (2 oder mehr Punkte *und* 2 % oder mehr der Punkte liegen außerhalb der Eingriffsgrenzen). In diesem Fall testet der Assistent zunächst, ob die Autokorrelation zwischen aufeinander folgenden Datenpunkten (Lag = 1) signifikant größer als 0,2 ist. Wenn die Autokorrelation signifikant größer als 0,2 ist, testet der Assistent, ob die Autokorrelation zwischen aufeinander folgenden Datenpunkten (Lag = 1) signifikant größer als 0,4 ist.

Für die Prüfung auf korrelierende Daten werden in der Auswertung des Assistenten die folgenden Statusindikatoren angezeigt:

Status	Bedingung
	<p>Die Anzahl der Punkte außerhalb der Eingriffsgrenzen ist nicht höher als erwartet, d. h., < 2 Punkte oder $< 2\%$ der Punkte liegen außerhalb der Eingriffsgrenzen.</p> <p>Die Anzahl der Punkte außerhalb der Eingriffsgrenzen ist höher als erwartet, aber bei einem Test auf Autokorrelation = 0,2 im Vergleich zu einer Autokorrelation $> 0,2$ wird die Nullhypothese bei Alpha = 0,01 nicht zurückgewiesen. Daher liegen nicht genügend Anzeichen für die Schlussfolgerung vor, dass mindestens eine mäßige Autokorrelation vorhanden ist.</p>
	<p>Wenn die oben aufgeführten Bedingungen nicht zutreffen.</p> <p>Hinweis: Wenn die Nullhypothese (Autokorrelation = 0,2) zurückgewiesen wird, wird ein Folgetest auf Autokorrelation = 0,4 im Vergleich zu einer Autokorrelation $> 0,4$ durchgeführt. Wenn der Test auf Autokorrelation = 0,4 zurückgewiesen wird, wird der Schweregrad der Warnmeldung erhöht.</p>

Weitere Informationen zum Hypothesentest für die Autokorrelation finden Sie in Anhang D.

Literaturhinweise

AIAG (1995). *Statistical process control (SPC) reference manual*. Automotive Industry Action Group.

Bischak, D.P. und 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. und O'Connell, R.T. (1979). *Forecasting and time series: An applied approach*. Belmont, CA: Duxbury Press.

Chan, L. K., Hapuarachchi K. P. und Macpherson, B.D. (1988). Robustness of \bar{X} and R charts. *IEEE Transactions on Reliability*, 37, 117–123.

Davis, R.B. und 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. und 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. und Zimmer, W.J. (1992). Non-normality and the design of control charts for averages. *Decision Sciences*, 23, 1099–1113.

Anhang A: Vorliegen einer Normalverteilung

Simulation A1: Auswirkungen der Nicht-Normalverteilung auf die Rate falscher Alarme

Um zu untersuchen, wie sich nicht normalverteilte Daten auf die Leistung der I-Karte und der X-quer-Karte auswirken, wurde eine Simulation durchgeführt, um die Rate falscher Alarme im Zusammenhang mit nicht normalverteilten Daten auszuwerten. Der Schwerpunkt lag auf Verteilungen mit Rechtsschiefe und symmetrische Verteilungen mit stärker besetzten Randbereichen, da es sich hierbei um Nicht-Normalverteilungen handelt, in der Praxis häufig auftreten. Insbesondere wurden 3 schiefe Verteilungen (Chi-Quadrat mit $df = 3, 5$ und 10) und 2 Verteilungen mit stärker besetzten Randbereichen (t mit $df=3$ und 5) untersucht.

Die Eingriffsgrenzen wurden unter Verwendung eines anfänglichen Satzes von 10.000 Teilgruppen festgelegt. Der Prozentsatz falscher Alarme für weitere 2.500 Teilgruppen wurde aufgezeichnet. Anschließend wurden 10.000 Iterationen durchgeführt, und es wurde der durchschnittliche Prozentsatz falscher Alarme bei Verwendung von Test 1 und Test 2 auf Ausnahmebedingungen berechnet. Die Ergebnisse finden Sie in Tabelle 1.

Tabelle 1 Prozentsatz falscher Alarme für Test 1, Test 2 für die I-Karte ($n=1$) und die X-quer-Karte ($n=2, 3, 4$ oder 5)

Verteilung	Teilgruppengröße				
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
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

Jede Zelle in Tabelle 1 gibt jeweils die berechnete Rate falscher Alarme für Test 1 und Test 2 an. Zum Beispiel gibt die Zelle für die Standardnormalverteilung (Normal (0; 1)) und die I-Karte ($n=1$) an, dass die I-Karte eine Rate falscher Alarme von 0,27 % für Test 1 und von 0,39 % für Test 2 aufweist, wenn die Daten normalverteilt sind.

Wenn die Raten falscher Alarme für die X-quer-Karte ($n=2, 3, 4, 5$) bei einer Normalverteilung und einer Nicht-Normalverteilung miteinander verglichen werden, zeigt sich, dass die Rate falscher Alarme sich nicht signifikant steigert, wenn gegen die Annahme

der Normalverteilung verstoßen wird. Selbst bei Verteilungen mit starker Schiefe (Chi-Quadrat (3)) und Verteilungen mit stärker besetzten Randbereichen (t (3)) ist die kombinierte Rate falscher Alarme für Test 1 und 2 bei Teilgruppengrößen, die lediglich 2 betragen, niedriger als 2,0 % (verglichen mit 0,7 % bei einer Normalverteilung). Somit haben wir aus praktischen Gründen geschlussfolgert, dass die X-quer-Karte gegenüber Abweichungen von der Normalverteilung robust ist.

Für die I-Karte zeigt Tabelle 1 eine Rate falscher Alarme von annähernd 3,2 % für Test 1 und Test 2 zusammen, wenn die Verteilung eine starke Schiefe aufweist (Chi-Quadrat (3)); diese Rate falscher Alarme ist fast 5 Mal höher als die erwartete Rate falscher Alarme, wenn die Daten normalverteilt sind. Die kombinierte Rate falscher Alarme für Test 1 und 2 für Verteilungen mit stärker besetzten Randbereichen (t (3)) beträgt 2,6 % und ist somit fast 4 Mal höher als die erwartete Rate, wenn die Daten normalverteilt sind.

Daher sollten die Daten für die I/MR-Karte zunächst auf Normalverteilung getestet werden. Wenn die Nullhypothese des Anderson-Darling-Tests auf Güte der Anpassung für Normalverteilungen zurückgewiesen wird (es liegen statistisch signifikante Anzeichen vor, dass die Daten keiner Normalverteilung folgen), kann eine Box-Cox-Transformation verwendet werden, um die Daten zu normalisieren. Die Box-Cox-Transformation ist jedoch nur für Daten mit Rechtsschiefe wirksam. Bei anderen nicht normalverteilten Daten sind weitere Untersuchungen erforderlich, um zu entscheiden, welche Maßnahmen am besten ergriffen werden sollten. Da sich darüber hinaus extreme Beobachtungen auf den Anderson-Darling-Test und die Box-Cox-Transformation auswirken, sollten Sie Punkte mit bekannten Ausnahmebedingungen aus den Berechnungen ausschließen, bevor Sie die Analyse durchführen.

Anhang B: Stabilität

Simulation B1: Auswirkungen der Ergänzung von Test 1 durch Test 2 auf die Empfindlichkeit

Test 1 identifiziert Punkte, die außer Kontrolle sind, indem signalisiert wird, wenn ein Punkt mehr als 3 Standardabweichungen von der Mittellinie entfernt liegt. Test 2 erkennt Shifts beim Mittelwert, indem signalisiert wird, wenn 9 aufeinander folgende Punkte auf der gleichen Seite der Mittellinie liegen.

Um abzuschätzen, ob die Verwendung von Test 2 mit Test 1 die Empfindlichkeit der Mittelwertkarten (I-Karte und X-quer-Karte) verbessert, wurden Eingriffsgrenzen für eine Normalverteilung (0; StdAbw) festgelegt. Der Mittelwert der Verteilung wurde um ein Vielfaches der Standardabweichung verschoben, und es wurde die Anzahl der Teilgruppen aufgezeichnet, die benötigt wird, um für jede von 10.000 Iterationen ein Signal zu erkennen. Die Ergebnisse finden Sie in Tabelle 2.

Tabelle 2 Durchschnittliche Anzahl von Teilgruppen bis zu einem Fehler bei Test 1 (Test 1), Fehler bei Test 2 (Test 2) oder einem Fehler bei Test 1 oder Test 2 (Test 1 oder 2). Der Shift im Mittelwert entspricht einem Vielfachen der Standardabweichung (StdAbw), und die Simulation wurde für die Teilgruppengrößen (n) = 1, 3 und 5 durchgeführt.

Shift	n=1			n=3			n=5		
	Test 1	Test 2	Test 1 oder 2	Test 1	Test 2	Test 1 oder 2	Test 1	Test 2	Test 1 oder 2
0,5 StdAbw	154	84	57	60	31	22	33	19	14
1 StdAbw	44	24	17	10	11	7	4	10	4
1,5 StdAbw	15	13	9	3	9	3	1,6	9	1,6
2 StdAbw	6	10	5	1,5	9	1,5	1,1	9	1,1

Wie in den Ergebnissen für die I-Karte (n=1) ersichtlich ist, werden bei Verwendung beider Tests (Spalte *Test 1 oder 2*) durchschnittlich 57 Teilgruppen benötigt, um einen Shift von 0,5 bei der Standardabweichung zu erkennen, verglichen mit durchschnittlich erforderlichen 154 Teilgruppen, um einen Shift von 0,5 bei der Standardabweichung zu erkennen, wenn nur Test 1 verwendet wird. Ebenso erhöht die Verwendung beider Tests die Empfindlichkeit für die X-quer-Karte (n=3, n=5). Zum Beispiel werden für die Teilgruppengröße 3 durchschnittlich 22 Teilgruppen benötigt, um einen Shift von 0,5 bei der Standardabweichung zu erkennen, wenn sowohl Test 1 als auch Test 2 verwendet werden, jedoch 60 Teilgruppen, um einen Shift von 0,5 bei der Standardabweichung zu erkennen, wenn nur Test 1 verwendet wird. Daher wird die Empfindlichkeit beim Erkennen kleiner Shifts im Mittelwert durch Verwendung beider Tests signifikant gesteigert. Bei größeren Shifts wird die Empfindlichkeit durch das Hinzufügen von Test 2 nicht signifikant erhöht.

Simulation B2: Wie effektiv erkennt Test 7 eine Schichtung (mehrere Streuungsquellen in Teilgruppen)?

Test 7 signalisiert in der Regel einen Fehler, wenn 12 bis 15 aufeinander folgende Punkte innerhalb 1 Standardabweichung von der Mittellinie liegen. Der Assistent verwendet eine modifizierte Regel, die die erforderliche Anzahl von Punkten auf der Grundlage der Anzahl von Teilgruppen in den Daten anpasst. Wir haben $k = (\text{Anzahl der Teilgruppen} \cdot 0,33)$ definiert und die erforderliche Anzahl der aufeinander folgenden Punkte festgelegt, die einen Fehler bei Test 7 auslösen, wie in Tabelle 3 gezeigt.

Tabelle 3 Anzahl der aufeinander folgenden Punkte, die einen Fehler bei Test 7 auslösen

k = (Anzahl der Teilgruppen) x 0,33	Erforderliche Punkte
$k < 12$	12
$k \geq 12$ und $k \leq 15$	Ganze Zahl $\geq k$
$k > 15$	15

Mit Hilfe von gängigen Szenarien für das Festlegen von Eingriffsgrenzen wurde eine Simulation durchgeführt, um die Wahrscheinlichkeit zu ermitteln, dass Test 7 unter Verwendung der oben genannten Kriterien einen Fehler signalisiert. Insbesondere sollte die Regel zum Erkennen einer Schichtung während der Phase ausgewertet werden, in der Eingriffsgrenzen anhand der Daten geschätzt werden.

Dabei wurden m Teilgruppen der Größe n zufällig aus einer Normalverteilung mit der Standardabweichung StdAbw ausgewählt. Die Hälfte der Punkte in den einzelnen Teilgruppen wies einen Mittelwert gleich 0 auf, die andere Hälfte einen Mittelwert gleich dem StdAbw-Shift (0 StdAbw, 1 StdAbw oder 2 StdAbw). Es wurden 10.000 Iterationen ausgeführt und der Prozentsatz der Regelkarten aufgezeichnet, die mindestens ein Signal bei Test 7 zeigten, wie in Tabelle 4 dargestellt.

Tabelle 4 Prozentsatz der Regelkarten, die mindestens ein Signal bei Test 7 aufweisen

Anzahl der Teilgruppen		m = 50	m = 75	m = 25	m = 38	m = 25
Teilgruppengröße		n = 2	n = 2	n = 4	n = 4	n = 6
Test		15 aufeinander folgend	15 aufeinander folgend	12 aufeinander folgend	13 aufeinander folgend	12 aufeinander folgend
Shift	0 StdAbw	5 %	8 %	7 %	8 %	7 %
	1 StdAbw	23 %	33 %	17 %	20 %	15 %
	2 StdAbw	83 %	94 %	56 %	66 %	50 %

Wie in der ersten „Shift“-Zeile der Tabelle gezeigt (Shift = 0 StdAbw), weist ein relativ kleiner Prozentsatz der Regelkarten mindestens einen Fehler bei Test 7 auf, wenn keine Schichtung vorliegt. Bei einer Schichtung (Shift = 1 StdAbw oder Shift = 2 StdAbw) weist hingegen ein deutlich größerer Prozentsatz der Regelkarten – 94 % – mindestens einen Fehler bei Test 7 auf. Auf diese Weise kann Test 7 eine Schichtung in der Phase identifizieren, in der die Eingriffsgrenzen geschätzt werden.

Anhang C: Umfang der Daten

Formel C1: Anzahl der erforderlichen Beobachtungen auf Basis einer 95%-KI für die obere Eingriffsgrenze

Wenn die Daten normalverteilt sind und kein Fehler aufgrund der Stichprobenstreuung auftritt, beträgt der Prozentsatz der Beobachtungen über der oberen Eingriffsgrenze für einen Prozess, der unter Kontrolle ist, 0,135 %. Um zu ermitteln, ob genügend Beobachtungen vorliegen, um sicherzustellen, dass die Rate falscher Alarme relativ niedrig bleibt, wurde gemäß Bischak (2007) wie folgt eine untere Konfidenzgrenze von 95 % für die obere Eingriffsgrenze berechnet:

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

Dabei gilt Folgendes:

$\hat{\mu}$ = Schätzung der Mittellinie

$h = 3 \sigma$

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

c_4 = erwartungstreue Konstante für eine X-quer-Karte

n = Anzahl der Beobachtungen

k = Teilgruppengröße

z_p = inverse kumulative Verteilungsfunktion bei $p = 0,95$ für die Normalverteilung mit Mittelwert = 0 und Standardabweichung = 1

Es wird $\frac{h}{w_p} = 2,65 \sigma$ festgelegt, was zu einer Rate falscher Alarme von 0,5 % über der oberen Eingriffsgrenze führt, und nach n aufgelöst.

Für die I-Karte wird σ mit Hilfe eines Durchschnitts der gleitenden Spannweite der Länge 2 geschätzt; daher wird einfach die Anzahl der Beobachtungen aus der X-quer-Karte mit einer Teilgruppengröße 2 verwendet, um den erforderlichen Umfang der Daten zu ermitteln. Basierend auf diesen Berechnungen wird die Anzahl der für verschiedene Teilgruppengrößen erforderlichen Beobachtungen in Tabelle 5 aufgeführt.

Tabelle 5 Anzahl der für verschiedene Teilgruppengrößen erforderlichen Beobachtungen

Teilgruppengröße	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Anzahl der Beobachtungen	114	114	84	72	70	66	63	64	63	60	66	60	65	56	60

Hinweis Die Anzahl der Beobachtungen sollte sich verringern, wenn sich die Teilgruppengröße erhöht. Tabelle 8 zeigt jedoch Ausnahmen zu dieser Regel. Diese Ausnahmen treten auf, da die Anzahl der Teilgruppen auf die nächste ganze Zahl gerundet wurde, bevor sie mit der Anzahl der Beobachtungen in jeder Teilgruppe multipliziert wurde, um die Gesamtanzahl der erforderlichen Beobachtungen zu berechnen.

Die Ergebnisse in Tabelle 5 zeigen, dass die Gesamtanzahl der erforderlichen Beobachtungen für alle Teilgruppengrößen kleiner oder gleich 100 ist, es sei denn, die Teilgruppengröße ist 1 oder 2. Doch selbst bei einer Teilgruppengröße von 1 oder 2 liegt die Rate falscher Alarme bei 100 Beobachtungen nur bei annähernd 1,1 %. Daher stellen 100 Beobachtungen einen wirksamen Trennwert für alle Teilgruppengrößen dar.

Bei der oben ausgeführten Analyse wird angenommen, dass jede Teilgruppe denselben Umfang an Streuung durch gewöhnliche Ursachen aufweist. In der Praxis können Daten, die zu verschiedenen Zeitpunkten erfasst werden, unterschiedliche Umfänge an Streuung durch gewöhnliche Ursachen aufweisen. Daher kann es ratsam sein, zu mehr Zeitpunkten als erforderlich Stichproben des Prozesses zu erfassen, um die Wahrscheinlichkeit zu erhöhen, eine repräsentative Schätzung der Prozessstreuung zu erhalten.

Anhang D: Autokorrelation

Simulation D1: Auswirkungen der Autokorrelation auf die Rate falscher Alarme

Unter Verwendung des folgenden Modells wurde ein autokorrelierter Prozess definiert:

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

Dabei gilt Folgendes:

x_t = Beobachtung zum Zeitpunkt t

μ = Mittelwert des Prozesses

ϕ = Korrelation zwischen zwei aufeinander folgenden Punkten

ε_t Fehlerterm zum Zeitpunkt t; $\varepsilon \sim$ Normalverteilung (0; 1)

Aus Gründen der Einfachheit wurde nur die Autokorrelation für Lag 1 berücksichtigt, da diese wahrscheinlich größer ist als die Autokorrelation für Lags ≥ 2 .

Unter Verwendung des oben gezeigten Modells wurde eine Simulation mit $\phi = 0,2; 0,4; 0,5; 0,6$ und $0,8$ und mit drei Teilgruppengrößen ($n=1, 3$ und 5) durchgeführt. Es wurde ein anfänglicher Satz von 10.000 Teilgruppen verwendet, um die Eingriffsgrenzen festzulegen, und der Prozentsatz falscher Alarme für weitere 2.500 Teilgruppen wurde aufgezeichnet. Es wurden 10.000 Iterationen durchgeführt, und es wurde der durchschnittliche Prozentsatz falscher Alarme unter Verwendung von Test 1 und Test 2 für jede Kombination von Autokorrelation und Teilgruppengröße aufgezeichnet. Die Ergebnisse finden Sie in der nachfolgenden Tabelle 6.

Tabelle 6 Durchschnittlicher Prozentsatz falscher Alarme für Test 1, Test 2 für einen Prozess mit Autokorrelation (ϕ) und Teilgruppengröße (n)

	ϕ				
Teilgruppengröße	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

Beachten Sie, dass sich die Rate falscher Alarme bei zunehmendem ϕ erhöht. Selbst eine mäßige Autokorrelation ($\phi = 0,4$) führt zu einer signifikanten Steigerung der Rate falscher Alarme. Bei einer Autokorrelation $\geq 0,4$ weist die Regelkarte einen so hohen Prozentsatz falscher Alarme auf, dass die Ergebnisse der Regelkarte ihre Aussagekraft nahezu vollständig verloren haben.

Formel D1: Test auf Autokorrelation

Um das Niveau der Autokorrelation ϕ in den Daten zu ermitteln, wird gemäß Bowerman (1979) ein Test durchgeführt, um die Nullhypothese $\phi = r$ im Vergleich zur Alternativhypothese $\phi > r$ unter Verwendung der folgenden Teststatistik auszuwerten:

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

Dabei gilt Folgendes:

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

m = Anzahl der Beobachtungen

$\hat{\sigma}$ = Stichprobenvarianz

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

Bei der Berechnung der Autokorrelation der Daten auf X-quer-Karten wird die Abhängigkeit zwischen dem letzten Punkt einer Teilgruppe und dem ersten Punkt der nächsten Teilgruppe nicht berücksichtigt, da die Einbeziehung dieser Vergleiche zu einem zu geringen Schätzwert für die Autokorrelation führen würde. Gemäß Bowerman (1979) wird der p-Wert auf Basis einer Standardnormalverteilung berechnet.

© 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. All other marks referenced remain the property of their respective owners.