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.

Messsystemanalyse (gekreuzt)

Übersicht

Messsystemanalysen werden buchstäblich in jedem Fertigungszweig durchgeführt, um Produktionsprozesse ordnungsgemäß zu überwachen und zu verbessern. In einer typischen Messsystemanalyse führen mehrere Prüfer mit einem Messgerät wiederholte Messungen für ausgewählte Teile durch. In derartigen Analysen werden häufig zwei Komponenten der Messsystemstreuung generiert: Wiederholbarkeit und Reproduzierbarkeit. Die Wiederholbarkeit stellt die Streuung dar, die auftritt, wenn mit dem Messgerät dasselbe Teil durch denselben Prüfer gemessen wird. Die Reproduzierbarkeit hingegen bezieht sich auf die Streuung, die auftritt, wenn verschiedene Prüfer dasselbe Teil messen. Daher werden solche Studien auch als Untersuchung zur Wiederholbarkeit und Reproduzierbarkeit des Messsystems bezeichnet.

Mit einer Messsystemanalyse soll in erster Linie bestimmt werden, welcher Anteil der Streuung in den Daten auf das Messsystem zurückzuführen ist und ob das Messsystem in der Lage ist, die Prozessleistung angemessen auszuwerten. Ausführliche Erläuterungen zu Messsystemanalysen finden Sie im MSA-Handbuch (2003), in Montgomery und Runger (1993) sowie in Burdick, Borror und Montgomery (2005).

Mit dem Befehl „Messsystemanalyse (gekreuzt)" im Assistenten können Daten aus typischen Messsystemuntersuchungen analysiert werden. Dabei kommt der am häufigsten verfolgte Ansatz zur Anwendung, bei dem ein ANOVA-Modell an die Messdaten angepasst wird und mit Hilfe der Varianzkomponenten im Modell die verschiedenen Streuungsquellen im Messsystem geschätzt werden.

Wenn Sie die typischen Richtlinien zur Menge der Daten beachten, die für Messsystemanalysen zu erfassen sind, können die Varianzkomponenten u. U. nicht genau geschätzt werden (Montgomery und Runger, 1993a, 1993b; Vardeman und VanValkenburg, 1999). Der Assistent gibt an, ob die Anzahl der Teile und die Anzahl der Prüfer kleiner als gewisse Werte sind, was sich auf die Genauigkeit der Schätzwerte für die Streuung zwischen Teilen und Prüfern auswirken kann. Wir haben Simulationen durchgeführt, um die notwendige Anzahl der Teile, Prüfer und Replikationen für genaue Schätzwerte zu bestimmen.

Unter Verwendung dieser Simulationsergebnisse und allgemein anerkannter Verfahren für die Messsystemanalyse wurden die folgenden Datenprüfungen für die gekreuzte Messsystemanalyse entwickelt. Diese Datenprüfungen werden vom Assistenten automatisch durchgeführt, und die Ergebnisse werden in der Auswertung ausgegeben.

- Umfang der Daten
    - Prozessstreuung
    - Streuung des Messsystems

In diesem Artikel wird untersucht, in welcher Beziehung diese Datenprüfungen zur Messsystemanalyse in der Praxis stehen. Außerdem wird erläutert, wie wir die Richtlinien für die einzelnen Datenprüfungen aufgestellt haben.

Datenprüfungen

Umfang der Daten

Typischerweise wird in Richtlinien für Messsystemanalysen die Verwendung von 10 Teilen, 2 oder 3 Prüfern sowie 2 oder 3 Replikationen empfohlen (AIAG, 2003; Raffaldi und Ramsier, 2000; Tsai, 1988). Der empfohlene Stichprobenumfang ist jedoch nicht groß genug, um die Streuung zwischen Teilen mit hinreichender Präzision zu schätzen, so dass er u. U. keine gute Grundlage für die Beurteilung liefert, ob ein bestimmtes Messgerät verwendet werden sollte (Montgomery und Runger, 1993a, 1993b; Vardeman und VanValkenburg, 1999).

Beim Entwickeln der Richtlinien für die angemessene Menge von Daten konzentrierten wir uns darauf, wie viele Teile ausgewertet werden müssen, um Schätzwerte für die Streuung zwischen den Teilen mit unterschiedlichen Präzisionsniveaus zu erhalten. Zudem haben wir untersucht, wie viele Prüfer eingesetzt werden müssen, um einen präzisen Schätzwert der Streuung des Messsystems zu erhalten. Schließlich haben wir untersucht, wie viele Beobachtungen benötigt werden, um Schätzwerte der Wiederholbarkeit eines Messgeräts mit unterschiedlichen Präzisionsniveaus zu erhalten.

Anzahl der Teile zum Schätzen der Streuung zwischen Teilen mit unterschiedlichen Präzisionsniveaus

Zielstellung

Es sollte bestimmt werden, wie viele Teile ausgewertet werden müssen, um Schätzwerte der Streuung zwischen den Teilen mit unterschiedlichen Präzisionsniveaus zu erhalten.

Methode

Es wurde eine Simulationsstudie mit 5.000 Stichproben durchgeführt. Für alle Stichproben wurde die Standardabweichung der Teile geschätzt, und das Verhältnis der geschätzten Standardabweichung zur tatsächlichen Standardabweichung wurde berechnet. Die Verhältnisse wurden in absteigender Reihenfolge sortiert. Mit dem 125. Verhältnis und dem 4875. Verhältnis wurde das 95%-Konfidenzintervall definiert, und mit dem 250. Verhältnis und dem 4750. Verhältnis wurde das 90%-Konfidenzintervall definiert. Auf der Grundlage dieser Konfidenzintervalle wurde ermittelt, wie viele Teile zum Schätzen der Streuung zwischen den Teilen mit unterschiedlichen Präzisionsniveaus benötigt werden.

Ergebnisse

Auf der Grundlage der Simulationsstudie wurden folgende Schlussfolgerungen gezogen:

- Bei Verwendung von 10 Teilen, 3 Prüfern und 2 Replikationen beträgt das Verhältnis des 90%-Konfidenzintervalls zur tatsächlichen Standardabweichung etwa (0,61; 1,37) bei einer Fehlerspanne von 35 % bis 40 %. Bei einem 95%-Konfidenzniveau ist das

Intervall ca. (0,55; 1,45) bei einer Fehlerspanne von 45 %. Daher sind 10 Teile nicht ausreichend, um einen präzisen Schätzwert für die Streuung zwischen den Teilen zu erhalten.

- Es werden etwa 35 Teile benötigt, um mit einer Sicherheit von 90 % eine Streuung zwischen den Teilen zu schätzen, die innerhalb von 20 % vom tatsächlichen Wert liegt.
- Es werden etwa 135 Teile benötigt, um mit einer Sicherheit von 90 % eine Streuung zwischen den Teilen zu schätzen, die innerhalb von 10 % vom tatsächlichen Wert liegt.

Zudem haben wir festgestellt, dass diese Ergebnisse gleichermaßen für akzeptable, grenzwertige und nicht akzeptable Messgeräte gelten.

Eine ausführliche Erläuterung der Simulation und ihrer Ergebnisse finden Sie in Anhang A.

Anzahl der Prüfer zum Schätzen der Streuung zwischen Teilen mit unterschiedlichen Präzisionsniveaus

Zielstellung

Es sollte bestimmt werden, wie viele Prüfer Teile bewerten müssen, um Schätzwerte der Streuung zwischen den Prüfern mit unterschiedlichen Präzisionsniveaus zu erhalten.

Methode

Sowohl die Standardabweichung für Teile als auch die Standardabweichung für Prüfer werden anhand des ANOVA-Modells geschätzt. Daher gilt die in der Simulation für die Anzahl der Teile zum Schätzen der Streuung zwischen den Teilen verwendete Methode auch für die Simulation für die Anzahl der Prüfer zum Schätzen der Streuung zwischen den Prüfern.

Ergebnisse

Zwei oder drei Prüfer sind nicht ausreichend, um einen präzisen Schätzwert der Reproduzierbarkeit zu erhalten. Das Problem ist jedoch weniger schwerwiegend, wenn der Umfang der Streuung zwischen den Teilen viel größer als die Streuung zwischen den Prüfern ist – ein wahrscheinliches Szenario in vielen Anwendungen.

Eine ausführliche Erläuterung der Simulation und ihrer Ergebnisse finden Sie in Anhang A.

Anzahl der Beobachtungen zum Schätzen der Wiederholbarkeit mit unterschiedlichen Präzisionsniveaus

Zielstellung

Es sollte bestimmt werden, wie die Anzahl der Beobachtungen den Schätzwert der Wiederholbarkeit beeinflusst und ob mit 10 Teilen, 3 Prüfern und 2 Replikationen ein hinreichend präziser Schätzwert für die Streuung der Wiederholbarkeit erlangt werden kann.

Methode

Das Verhältnis der geschätzten Standardabweichung der Wiederholbarkeit zum tatsächlichen Wert folgt einer Chi-Quadrat-Verteilung. Zum Bestimmen der Anzahl von Beobachtungen, die erforderlich sind, um einen hinreichend präzisen Schätzwert der Wiederholbarkeit zu erhalten, wurden die Untergrenze und die Obergrenze des Verhältnisses für eine Wahrscheinlichkeit von 90 % berechnet, und die Ergebnisse wurden grafisch dargestellt.

Ergebnisse

In einer typischen Messsystemanalyse (z. B. Anzahl der Teile = 10, Anzahl der Prüfer = 3 und Anzahl der Replikationen = 2) sind die Freiheitsgrade für Fehler gleich 30. Damit kann mit einer Sicherheit von 90 % eine Wiederholbarkeit geschätzt werden, die innerhalb von 20 % vom tatsächlichen Wert liegt. Mit den typischen Einstellungen ist der Schätzwert der Wiederholbarkeit hinreichend präzise. Weitere Einzelheiten hierzu können Sie Anhang B entnehmen.

Gesamtergebnisse

Unsere Untersuchungen zeigen eindeutig, dass die in einer Messsystemanalyse verwendeten typischen Einstellungen nicht ausreichen, um präzise Schätzwerte für die Streuung zwischen den Teilen und die Streuung der Reproduzierbarkeit zu liefern, die sich auf das Verhältnis der Messsystemstreuung zur Prozessstreuung insgesamt und letztendlich auf die Entscheidung auswirken, ob das Messgerät akzeptabel ist. Typischerweise ist die Streuung zwischen den Teilen größer als die Streuung der Reproduzierbarkeit, so dass sich ihre Genauigkeit stärker darauf auswirkt, ob ein Messgerät als akzeptabel einzustufen ist. In vielen Anwendungen ist es jedoch u. U. nicht praktikabel, 35 oder mehr Teile auszuwählen und diese von mehreren Prüfern zweimal messen zu lassen.

Unter Berücksichtigung der in der Praxis verwendeten typischen Einstellungen für die Messsystemanalyse und der vorliegenden Simulationsergebnisse verfolgt der Assistent die folgenden Ansätze, um Benutzern das Erlangen präziser Schätzwerte für die Varianzkomponenten zu ermöglichen:

1. Es gibt eine Option im Dialogfeld, mit der Benutzer einen Schätzwert der Prozessstreuung aus einem umfassenden historischen Datensatz eingeben können. In den meisten Fällen weist der Schätzwert aus einem umfassenden historischen

Datensatz eine bessere Präzision als der Schätzwert auf, der anhand von Stichprobendaten gewonnen wurde.

2. Wenn der historische Schätzwert nicht verfügbar und die Anzahl der Teile klein ist, wird eine Meldung angezeigt, in der Benutzer darauf hingewiesen werden, dass mehr als 10 Teile ausgewählt werden müssen, um einen genaueren Schätzwert zu erhalten.

In der Auswertung des Assistenten werden entsprechend dem Umfang der Daten Informationen zur Prozessstreuung und zur Streuung des Messsystems angezeigt. Wenn Sie beispielsweise 10 Teile und 3 Prüfer verwenden und eine historische Standardabweichung angeben, wird in der Auswertung die folgende Angaben zur Datenprüfung angezeigt:

Status	Bedingung
i	Damit Sie feststellen können, ob mit einem Messsystem die Prozessleistung bewertet werden kann, benötigen Sie gute Schätzwerte der Prozessstreuung und der Streuung des Messsystems. Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine historische Standardabweichung eingegeben, so dass beide Schätzwerte verfügbar sind. Sie können beide Werte vergleichen, um zu ermitteln, wie gut sie übereinstimmen. Obwohl die Anzahl der Teile in dieser Untersuchung (10) der typischen Anforderung von 10 entspricht, sollte der historische Wert eine genauere Schätzung der Prozessstreuung liefern. Streuung des Messsystems: Wird anhand der Teile geschätzt und in Reproduzierbarkeit und Wiederholbarkeit aufgegliedert. Die Anzahl der Teile (10) und Prüfer (3) entspricht der typischen Anforderung von 10 Teilen und 3 Prüfern. Dies ist in der Regel ausreichend, um die Wiederholbarkeit zu schätzen. Die Schätzung der Reproduzierbarkeit ist jedoch weniger genau. Wenn die Schätzung von %Prozess für Reproduzierbarkeit hoch ist, sollten Sie möglicherweise die Differenzen zwischen den Prüfern untersuchen und ermitteln, ob diese Differenzen auch auf andere Prüfer übertragen werden können.

Unten sind sämtliche Meldungen für die verschiedenen Konfigurationen von Teilen, Prüfern und Replikationen aufgeführt.

PROZESSSTREUUNG

Historische Standardabweichung (Teile < 10)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine historische Standardabweichung eingegeben, so dass beide Schätzwerte verfügbar sind. Sie können beide Werte vergleichen, um zu ermitteln, wie gut sie übereinstimmen. Die Anzahl der Teile in dieser Untersuchung ist gering. Daher sollte der historische Wert eine genauere Schätzung der Prozessstreuung liefern.

Historische Standardabweichung (Teile ≥ 10, ≤ 15)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine historische Standardabweichung eingegeben, so dass beide Schätzwerte verfügbar sind. Sie können beide Werte vergleichen, um zu ermitteln, wie gut sie übereinstimmen.

Obwohl die Anzahl der Teile in dieser Untersuchung der typischen Anforderung von 10 entspricht, sollte der historische Wert eine genauere Schätzung der Prozessstreuung liefern.

Historische Standardabweichung (Teile > 15, < 35)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine historische Standardabweichung eingegeben, so dass beide Schätzwerte verfügbar sind. Sie können beide Werte vergleichen, um zu ermitteln, wie gut sie übereinstimmen. Die Anzahl der Teile in dieser Untersuchung ist wesentlich höher als die typische Anforderung von 10. Wenn die ausgewählten Teile die typische Prozessstreuung darstellen, sollte diese Schätzung der Prozessstreuung deutlich besser sein als bei der Verwendung von 10 Teilen.

Historische Standardabweichung (Teile ≥ 35)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine historische Standardabweichung eingegeben, so dass beide Schätzwerte verfügbar sind. Sie können beide Werte vergleichen, um zu ermitteln, wie gut sie übereinstimmen. Die Anzahl der Teile in dieser Untersuchung ist wesentlich höher als die typische Anforderung von 10. Wenn die ausgewählten Teile die typische Prozessstreuung darstellen, sollte diese Schätzung der Prozessstreuung ausreichen.

Keine historische Standardabweichung (Teile < 10)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine Schätzung anhand der Teile in der Untersuchung gewählt. Es sind jedoch weniger Teile als die typische Anforderung von 10 vorhanden. Die Genauigkeit dieser Schätzung ist möglicherweise nicht ausreichend. Wenn die ausgewählten Teile nicht die typische Prozessstreuung abbilden, sollten Sie möglicherweise eine historische Schätzung eingeben oder mehr Teile verwenden.

Keine historische Standardabweichung (Teile ≥ 10, ≤ 15)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine Schätzung anhand der Teile in der Untersuchung gewählt. Obwohl die Anzahl der Teile der typischen Anforderung von 10 entspricht, ist die Schätzung möglicherweise nicht genau. Wenn die ausgewählten Teile nicht die typische Prozessstreuung abbilden, sollten Sie möglicherweise eine historische Schätzung eingeben oder mehr Teile verwenden.

Keine historische Standardabweichung (Teile > 15, < 35)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine Schätzung anhand der Teile in der Untersuchung gewählt. Die Anzahl der Teile ist wesentlich höher als die typische Anforderung von 10. Wenn die ausgewählten Teile die typische Prozessstreuung darstellen, sollte diese Schätzung der Prozessstreuung deutlich besser sein als bei der Verwendung von 10 Teilen.

Keine historische Standardabweichung (Teile ≥ 35)

- Prozessstreuung: Besteht aus der Streuung zwischen den Teilen und der Streuung des Messsystems. Sie kann aus einer großen Stichprobe historischer Daten oder aus den Teilen in der Untersuchung geschätzt werden. Sie haben eine Schätzung anhand der Teile in der Untersuchung gewählt. Die Anzahl der Teile ist wesentlich höher als die typische Anforderung von 10. Wenn die ausgewählten Teile die typische Prozessstreuung darstellen, sollte diese Schätzung der Prozessstreuung ausreichen.

STREUUNG DES MESSSYSTEMS

Prüfer ≤ 2 oder Teile < 10

- Streuung des Messsystems: Wird anhand der Teile geschätzt und in Reproduzierbarkeit und Wiederholbarkeit aufgegliedert. Die Anzahl der Teile oder Prüfer entspricht nicht der typischen Anforderung von 10 Teilen und 3 Prüfern. Die Schätzungen für die Streuung des Messsystems sind möglicherweise nicht genau. Sie sollten die Schätzungen als Hinweise auf allgemeine Tendenzen und nicht als genaue Ergebnisse betrachten.

Prüfer ≥ 3 und ≤ 5 sowie Teile ≥ 10

- Streuung des Messsystems: Wird anhand der Teile geschätzt und in Reproduzierbarkeit und Wiederholbarkeit aufgegliedert. Die Anzahl der Teile und Prüfer entspricht der typischen Anforderung von 10 Teilen und 3 Prüfern. Dies ist in der Regel ausreichend, um die Wiederholbarkeit zu schätzen. Die Schätzung der Reproduzierbarkeit ist jedoch weniger genau. Wenn die Schätzung von %Prozess für Reproduzierbarkeit hoch ist, sollten Sie möglicherweise die Differenzen zwischen den Prüfern untersuchen und ermitteln, ob diese Differenzen auch auf andere Prüfer übertragen werden können.

Prüfer > 5 und Teile ≥ 10

- Streuung des Messsystems: Wird anhand der Teile geschätzt und in Reproduzierbarkeit und Wiederholbarkeit aufgegliedert. Die Anzahl der Teile und Prüfer entspricht der typischen Anforderung von 10 Teilen und 3 Prüfern und ist in der Regel ausreichend, um die Wiederholbarkeit zu schätzen. Die zusätzlichen Prüfer verbessern die Genauigkeit der Schätzung für die Reproduzierbarkeit.

Literaturhinweise

Burdick, R. K., Borror, C. M. und Montgomery, D.C. (2005). *Design and analysis of gauge R&R studies: Making decisions with confidence intervals in random and mixed ANOVA models.* Philadelphia, PA: Society for Industrial Applied Mathematics (SIAM).

Automotive Industry Action Group (AIAG) (2003). *Measurement systems analysis (MSA) manual* (*3rd edition*). Southfield, MI: Chrysler, Ford, General Motors Supplier Quality Requirements Task Force.

Montgomery, D. C. (2000). *Design and analysis of experiments*. New York, NY: Wiley.

Montgomery, D. C. und Runger, G. C. (1993 a). Gage capability and designed experiments. Part I: Basic methods. *Quality Engineering*, *6 (1993/1994)*, *115 – 135*.

Montgomery, D. C. und Runger, G. C. (1993 b). Gage capability analysis and designed experiments. Part II: Experimental design models and variance component estimation. *Quality Engineering, 6 (1993/1994)*, *289-305*.

Raffaldi, J. und Ramsier, S. (2000). 5 ways to verify your gages. *Quality Magazine*, *39 (3)*, *38-42*.

Tsai, P. (1988). Variable gage repeatability and reproducibility study using the analysis of variance method. *Quality Engineering, 1(1)*, *107-115*.

Vardeman, S. B. und VanValkenburg, E. S. (1999). Two-way random-effects analyses and gage R&R studies. *Technometrics, 41 (3)*, *202-211*.

Anhang A: Bewerten des Effekts von Teilen auf die Streuung zwischen den Teilen

Da keine genaue Formel zum Berechnen des Konfidenzintervalls für die Standardabweichung zwischen den Teilen verfügbar ist, haben wir eine Simulation zum Schätzen des Intervalls durchgeführt. Damit der Schwerpunkt der Simulation auf dem Einfluss der Anzahl der Teile auf die Genauigkeit der geschätzten Streuung zwischen den Teilen liegt, haben wir das Verhältnis des geschätzten Konfidenzintervalls für die Standardabweichung der Teile zur tatsächlichen Standardabweichung der Teile untersucht. Mit zunehmender Anzahl der Teile wird das Intervall schmaler. Anschließend wurde die Anzahl der Teile bestimmt, bei der die Fehlerspanne des Verhältnisses 10 % bzw. 20 % beträgt. Das Intervall für eine Fehlerspanne von 10 % ist (0,9; 1,1), das Intervall für eine Fehlerspanne von 20 % (0,8; 1,2).

Einrichtung der Simulation

Für eine Messsystemanalyse wird angenommen, dass die k-te Messung des i-ten Teils durch den j-ten Prüfer, dargestellt als Y_{ijk}, dem folgenden Modell entspricht:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

Dabei gilt Folgendes:

$i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, K$, und

$\alpha_i; \beta_j; \gamma_{ij}$ und ε_{ijk} sind unabhängig normalverteilt mit dem Mittelwert 0 und den Varianzen $\sigma_P^2; \sigma_O^2; \sigma_{OP}^2$ und σ_e^2. Hierbei stellen $\alpha_i; \beta_j; \gamma_{ij}$ und ε_{ijk} Teile, Prüfer, Teile x Prüfer sowie Fehlerterme dar.

Sei r das Verhältnis der Gesamt-Standardabweichung des Messsystems zur Gesamt-Standardabweichung des Prozesses. Dann gilt Folgendes:

$$r = \frac{\sqrt{Varianz\ der\ Wiederholbarkeit + Varianz\ der\ Reproduzierbarkeit}}{\sqrt{Varianz\ der\ Teile + Varianz\ der\ Wiederholbarkeit + Varianz\ der\ Reproduzierbarkeit}}$$
$$= \frac{\sqrt{\sigma_e^2 + \sigma_O^2 + \sigma_{PO}^2}}{\sqrt{\sigma_P^2 + \sigma_e^2 + \sigma_O^2 + \sigma_{PO}^2}}$$

Typischerweise wird mit der folgenden Regel bestimmt, ob ein Messsystem akzeptabel ist:

$r \leq 0{,}1$ (10 %): akzeptabel

$0{,}1 < r \leq 0{,}3$: grenzwertig

$0{,}3 < r$: nicht akzeptabel

Wir wählen r = 0,1 (akzeptabel), r = 0,25 (grenzwertig) und r = 0,35 (nicht akzeptabel), um drei Bereiche zu definieren. Für die Zwecke der Simulation wird angenommen, dass die Streuung der Wiederholbarkeit der Streuung der Reproduzierbarkeit entspricht; dies ergibt:

$$\frac{\sqrt{\sigma_e^2+\sigma_e^2}}{\sqrt{\sigma_P^2+2\sigma_e^2}}=r \Rightarrow \sigma_P=\frac{\sqrt{(2-2r^2)}}{r}\sigma_e$$

Um zu untersuchen, wie sich die Anzahl der Teile auf die Standardabweichung der Teile auswirkt, werden die Beobachtungen mit $\sigma_{e=}0{,}001$ und 1; $\sigma_O^2=\sigma_{PO}^2=0{,}5\sigma_e^2$ und $\sigma_P=\frac{\sqrt{(2-2r^2)}}{r}\sigma_e$ generiert, und es wird angenommen, dass 3 Prüfer jedes Teil zweimal messen.

Die folgenden Simulationsschritte wurden für die einzelnen Teileanzahlen, r, und σ_e durchgeführt:

1. 5.000 Stichproben anhand des oben aufgeführten Modells generieren.
2. Die Standardabweichung der Teile schätzen, und das Verhältnis der geschätzten Standardabweichung zur tatsächlichen Standardabweichung für alle 5.000 Stichproben berechnen.
3. Die 5.000 Verhältnisse in aufsteigender Reihenfolge sortieren. Von den 5.000 sortierten Verhältnissen stellen das 125. und das 4.875. Verhältnis die Untergrenze bzw. die Obergrenze des Intervalls beim 95%-Konfidenzniveau dar, während das 250. und das 4.750. Verhältnis die Untergrenze bzw. die Obergrenze des Intervalls beim 90%-Konfidenzniveau darstellen.
4. Die Intervalle untersuchen, um die Anzahl der Teile zu ermitteln, bei der die Fehlerspanne 10 % bzw. 20 % beträgt. Das Intervall für eine Fehlerspanne von 10 % ist (0,9; 1,1). Das Intervall für eine Fehlerspanne von 20 % ist (0,8; 1,2).

Simulationsergebnisse

Die Ergebnisse in den Tabellen 1-6 zeigen die Simulationsergebnisse bei jedem Konfidenzniveau für unterschiedliche Teileanzahlen, wobei jede Tabelle einer spezifischen Kombination von Werten für r und σ_e entspricht. Insgesamt zeigen diese Ergebnisse Folgendes:

- Bei der Verwendung von 10 Teilen, 3 Prüfern und 2 Replikationen beträgt das Verhältnis des 90%-Konfidenzintervalls zur tatsächlichen Standardabweichung etwa (0,61; 1,37) bei einer Fehlerspanne von 35 % bis 40 %. Beim 95%-Konfidenzniveau ist das Intervall ca. (0,55; 1,45) bei einer Fehlerspanne von 45 %. Daher reichen 10 Teile nicht aus, um einen präzisen Schätzwert für die Streuung zwischen den Teilen zu erhalten.
- Es werden etwa 35 Teile benötigt, um mit einer Sicherheit von 90 % eine Streuung zwischen den Teilen zu schätzen, die innerhalb von 20 % vom tatsächlichen Wert liegt.
- Es werden etwa 135 Teile benötigt, um mit einer Sicherheit von 90 % eine Streuung zwischen den Teilen zu schätzen, die innerhalb von 10 % vom tatsächlichen Wert liegt.

Beachten Sie, dass diese Zusammenfassung der Ergebnisse nicht spezifisch für eine bestimmte Kombination von r und σ_e. ist. Die Zeilen, die sich auf die oben aufgeführten Ergebnisse beziehen, sind unten in den Tabellen 1, 2, 3, 4, 5 und 6 hervorgehoben.

Tabelle 1 Akzeptables Messgerät (r = 0,1), $\sigma_e = 0{,}001$, tatsächliche Standardabweichung der Teile = 0,014071247

	Verhältnis der Konfidenzintervalle für geschätzte Standardabw./tatsächliche Standardabw. der Teile	
Anzahl der Teile	95 %-Konfidenz	90 %-Konfidenz
3	(0,15295; 1,93755)	(0,22195; 1,73365)
5	(0,34415; 1,67035)	(0,41861; 1,53873)
10	(0,55003; 1,44244)	(0,60944; 1,36992)
15	(0,63295; 1,36927)	(0,68721; 1,30294)
20	(0,68532; 1,31187)	(0,7295; 1,25701)
25	(0,7123; 1,27621)	(0,75578; 1,23251)
30	(0,74135; 1,24229)	(0,77645; 1,20841)
35	(0,76543; 1,23033)	(0,80066; 1,19706)
50	(0,79544; 1,20337)	(0,82636; 1,16595)
100	(0,85528; 1,13696)	(0,88063; 1,11635)
135	(0,87686; 1,12093)	(0,89448; 1,09760)
140	(0,88241; 1,11884)	(0,90130; 1,09974)

Tabelle 2 Akzeptables Messgerät (r =0,1), $\sigma_e = 1$, tatsächliche Standardabweichung der Teile = 14,071247

	Verhältnis der Konfidenzintervalle für geschätzte Standardabw./tatsächliche Standardabw. der Teile	
Anzahl der Teile	95 %-Konfidenz	90 %-Konfidenz
5	(0,34656; 1,68211)	(0,42315; 1,5588)
10	(0,55496; 1,45382)	(0,61319; 1,38233)
15	(0,63484; 1,36949)	(0,68767; 1,30505)
35	(0,76233; 1,23513)	(0,79749; 1,19623)
40	(0,77256; 1,21518)	(0,81224; 1,18121)
135	(0,88017; 1,12345)	(0,89883; 1,10249)
140	(0,88004; 1,11725)	(0,89787; 1,09713)
145	(0,88281; 1,11886)	(0,89966; 1,09583)

	Verhältnis der Konfidenzintervalle für geschätzte Standardabw./tatsächliche Standardabw. der Teile	
Anzahl der Teile	95 %-Konfidenz	90 %-Konfidenz
150	(0,88302; 1,11132)	(0,90096; 1,09296)

Tabelle 3 Grenzwertiges Messgerät (r = 0,25), $\sigma_e = 0{,}001$, tatsächliche Standardabweichung der Teile = 0,005477225575

	Verhältnis der Konfidenzintervalle für geschätzte Standardabw./tatsächliche Standardabw. der Teile	
Anzahl der Teile	95 %-Konfidenz	90 %-Konfidenz
30	(0,73879; 1,25294)	(0,77982; 1,21041)
35	(0,75881; 1,24383)	(0,79848; 1,20068)
40	(0,77281; 1,22813)	(0,80369; 1,18788)
135	(0,87588; 1,11910)	(0,89556; 1,10093)
140	(0,87998; 1,12001)	(0,89917; 1,09717)
145	(0,88100; 1,11812)	(0,89852; 1,09710)
150	(0,88373; 1,11563)	(0,90345; 1,09706)

Tabelle 4 Grenzwertiges Messgerät (r = 0,25), $\sigma_e = 1$, tatsächliche Standardabweichung der Teile = 5,477225575

	Verhältnis der Konfidenzintervalle für geschätzte Standardabw./tatsächliche Standardabw. der Teile	
Anzahl der Teile	95 %-Konfidenz	90 %-Konfidenz
30	(0,74292; 1,25306)	(0,78159; 1,20872)
35	(0,76441; 1,24391)	(0,79802; 1,20135)
40	(0,77525; 1,21339)	(0,80786; 1,17908)
135	(0,87501; 1,11711)	(0,89512; 1,09758)
140	(0,87934; 1,11756)	(0,89881; 1,09862)
145	(0,88308; 1,11530)	(0,90056, 1,09806)

Tabelle 5 Nicht akzeptables Messgerät (r = 0,35), $\sigma_e = 0{,}001$, tatsächliche Standardabweichung der Teile = 0,00378504

	Verhältnis der Konfidenzintervalle für geschätzte Standardabw./tatsächliche Standardabw. der Teile	
Anzahl der Teile	95 %-Konfidenz	90 %-Konfidenz
30	(0,74313; 1,25135)	(0,77427; 1,20568)
35	(0,75409; 1,24332)	(0,79444; 1,19855)
40	(0,76582; 1,22289)	(0,80599; 1,18615)
135	(0,87641; 1,12043)	(0,89507; 1,09820)
140	(0,87635; 1,11539)	(0,89651; 1,09368)
145	(0,88339; 1,11815)	(0,89772; 1,09591)

Tabelle 6 Nicht akzeptables Messgerät (r = 0,35), $\sigma_e = 1$, tatsächliche Standardabweichung der Teile = 3,78504

	Verhältnis der Konfidenzintervalle für geschätzte Standardabw./tatsächliche Standardabw. der Teile	
Anzahl der Teile	95 %-Konfidenz	90 %-Konfidenz
30	(0,73750; 1,26100)	(0,77218; 1,21285)
35	(0,74987; 1,23085)	(0,79067; 1,18860)
40	(0,77187; 1,22270)	(0,80648; 1,18329)
135	(0,87572; 1,11877)	(0,89409; 1,09827)
140	(0,87798; 1,11634)	(0,89590; 1,09695)
145	(0,87998; 1,11513)	(0,89683; 1,09534)

Anzahl der Prüfer

Sowohl die Standardabweichung für Teile als auch die Standardabweichung für Prüfer werden anhand des ANOVA-Modells geschätzt. Daher gelten die Simulationsergebnisse für Teile auch für die Streuung der Reproduzierbarkeit. Zwei oder drei Prüfer sind nicht ausreichend, um einen präzisen Schätzwert der Reproduzierbarkeit zu erhalten. Das Problem ist jedoch weniger schwerwiegend, wenn der Umfang der Streuung zwischen den Teilen viel größer als die Streuung zwischen den Prüfern ist – ein wahrscheinliches Szenario in vielen Anwendungen.

Angenommen, die Standardabweichung zwischen den Teilen beläuft sich auf das 20-fache der Standardabweichung zwischen den Prüfern. Die Standardabweichung zwischen den Teilen beträgt 20, und die Standardabweichung zwischen den Prüfern beträgt 1.

Angenommen, die Wiederholbarkeit ist gleich der Reproduzierbarkeit. In diesem Fall ist das tatsächliche Verhältnis der Streuung des Messsystems zur Gesamt-Prozessstreuung:

$$\sqrt{\frac{1+1}{400+1+1}} = 0{,}0705$$

Es sei nun angenommen, dass die Fehlerspanne beim Schätzen der Standardabweichung zwischen den Prüfern 40 % beträgt (hoch). Das heißt, die geschätzte Standardabweichung zwischen den Prüfern könnte 1,4 betragen. Daher wird das Verhältnis der Streuung des Messsystems zur Gesamt-Prozessstreuung:

$$\sqrt{\frac{1{,}4^2+1{,}4^2}{400+1{,}4^2+1{,}4^2}} = 0{,}0985$$

Da dieser Wert kleiner als 0,10 ist, wirkt sich eine große Streuung der Reproduzierbarkeit nicht auf die Annahme des Messgeräts aus, wenn als Trennwert 10 % festgelegt ist.

Wenn die Streuung zwischen den Prüfern und die Streuung zwischen den Teilen nahezu gleich sind, benötigen Sie eine größere Anzahl von Prüfern, um das Messsystem darzustellen und das Messgerät genau auszuwerten.

Anhang B: Schätzen der Wiederholbarkeit

Einrichtung der Berechnungen

Im Unterschied zu den Konfidenzintervallen für die Standardabweichung zwischen den Teilen, denen eine Approximation zugrunde liegt, folgt das Verhältnis der geschätzten Standardabweichung der Wiederholbarkeit zum tatsächlichen Wert einer Chi-Quadrat-Verteilung. Daher können die Untergrenze und die Obergrenze des Verhältnisses für eine Wahrscheinlichkeit von 90 % berechnet werden, und anschließend kann untersucht werden, wie sich die Grenzen mit zunehmender Anzahl der Teile, Anzahl der Prüfer und Anzahl der Replikationen 1 annähern.

Unter Verwendung der in Anhang A definierten Notation wird die Streuung der Wiederholbarkeit geschätzt durch:

$$S^2 = \sum (Y_{ijk} - \bar{Y}_{ij.})^2 / IJ(K-1)$$

Daher: $\frac{IJ(K-1)S^2}{\sigma_e^2}$ folgt einer Chi-Quadrat-Verteilung mit IJ(K-1) Freiheitsgraden (df), wobei *I* die Anzahl der Teile, *J* die Anzahl der Prüfer und *K* die Anzahl der Replikationen darstellt.

Basierend auf diesem Ergebnis erfüllt das Verhältnis der geschätzten Standardabweichung zu ihrem tatsächlichen Wert die folgende Wahrscheinlichkeitsgleichung:

$$\text{Wahrscheinlichkeit}\left(\sqrt{\frac{\chi^2_{df,\alpha/2.}}{df}} \leq \frac{S}{\sigma_e} \leq \sqrt{\frac{\chi^2_{df,1-(\alpha/2)}}{df}}\right) = 1 - \alpha$$

Hierbei ist df = IJ(K-1) = Anzahl der Teile * Anzahl der Prüfer * (Anzahl der Replikationen – 1). Wenn die Anzahl der Replikationen gleich 2 ist, sind die Freiheitsgrade gleich dem Produkt aus der Anzahl der Teile und der Anzahl der Prüfer.

Mit dieser Formel werden für jeden angegebenen Wert der Freiheitsgrade die Untergrenze und die Obergrenze des Verhältnisses $\frac{S}{\sigma_e}$ bei einer Wahrscheinlichkeit von 90 % berechnet. Anschließend werden die Freiheitsgrade identifiziert, bei denen die geschätzte Standardabweichung innerhalb von 10 % und 20 % von ihrem tatsächlichen Wert liegt. Das entsprechende Intervall ist (0,9; 1,1) für eine Fehlerspanne von 10 % und (0,8; 1,2) für eine Fehlerspanne von 20 %.

Ergebnisse der Berechnungen

Das Diagramm in Abbildung 1 zeigt die Untergrenze und die Obergrenze des Verhältnisses $\frac{S}{\sigma_e}$ bei einer Wahrscheinlichkeit von 90 % im Vergleich zu den Freiheitsgraden, wobei sich der Bereich der Freiheitsgrade von 1 bis 200 erstreckt.

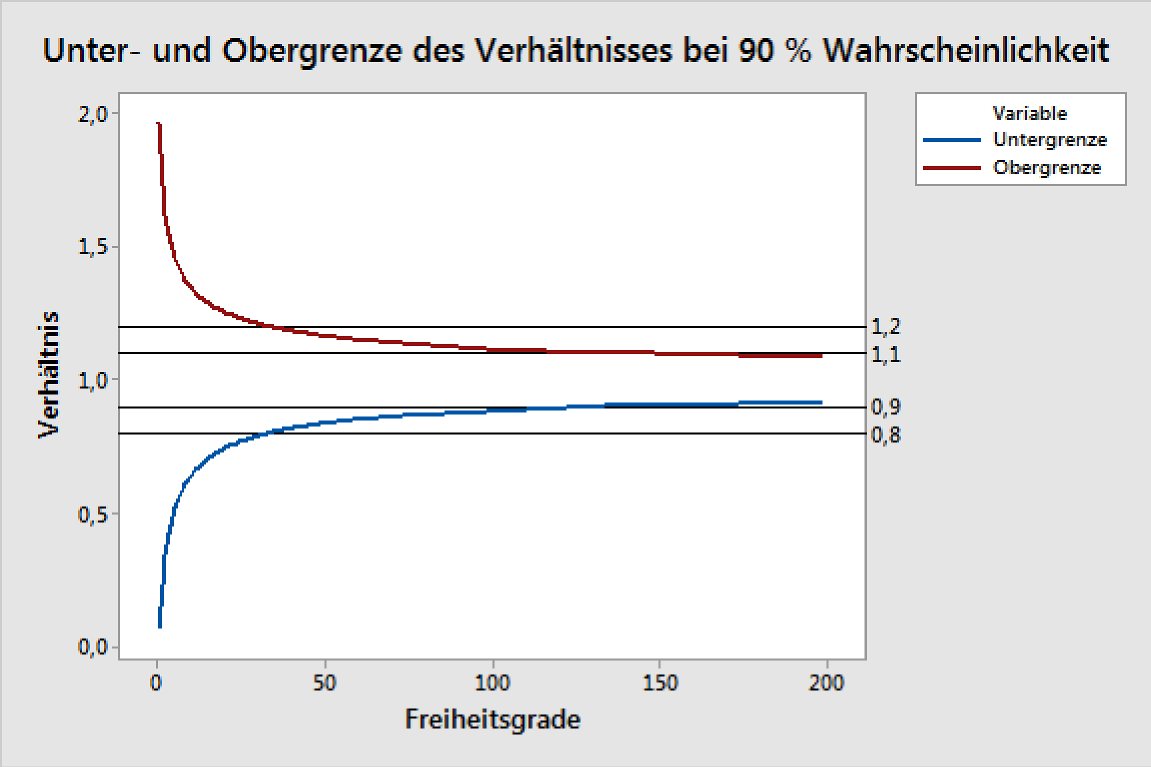

Abbildung 1 Untergrenze und Obergrenze von $\frac{s}{\sigma_e}$ bei einer Wahrscheinlichkeit von 90 % im Vergleich zu den Freiheitsgraden (1 bis 200)

Beachten Sie, dass das von der Untergrenze und der Obergrenze gebildete Intervall mit Zunahme der Freiheitsgrade schmaler wird. Die Breite des Intervalls nimmt bei Zunahme der Freiheitsgrade von 1 bis 50 erheblich ab. Dies wird im vergrößerten Diagramm in Abbildung 2 deutlicher sichtbar. Diese Abbildung veranschaulicht die Ergebnisse für die Freiheitsgrade von 1 bis 50.

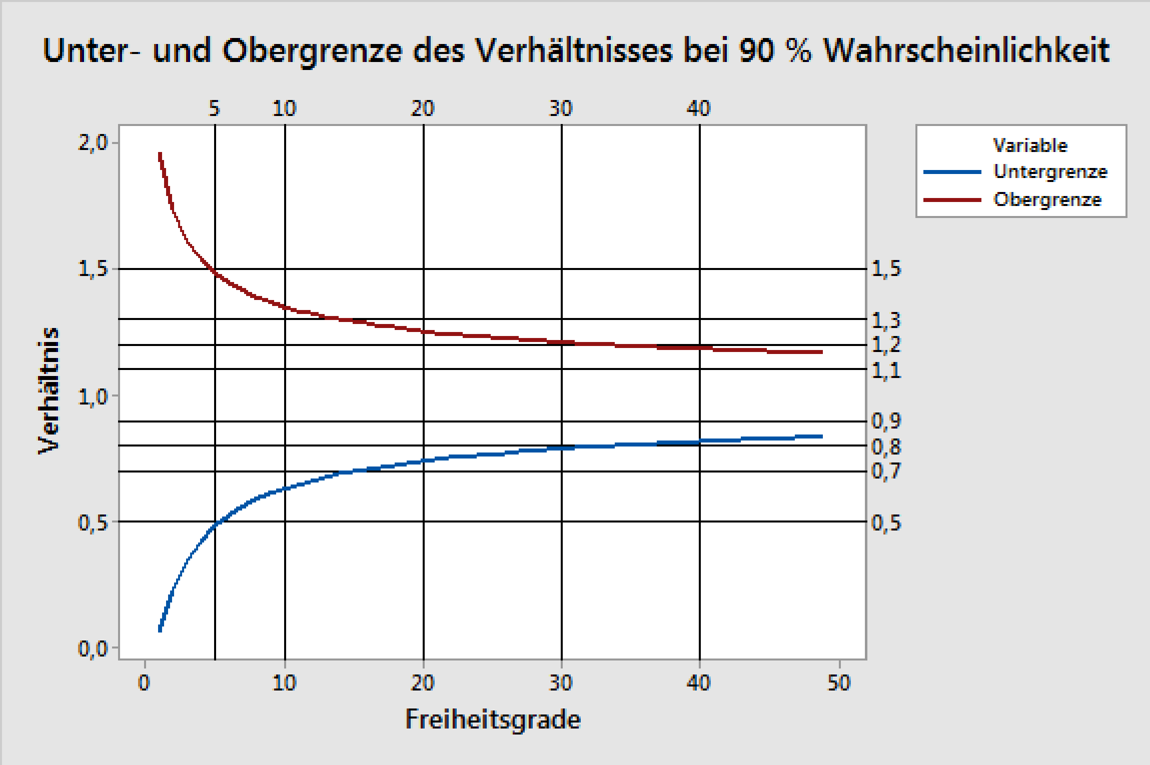

Abbildung 2 Untergrenze und Obergrenze von $\frac{S}{\sigma_e}$ bei einer Wahrscheinlichkeit von 90 % im Vergleich zu den Freiheitsgraden (1 bis 50)

Wie in Abbildung 2 gezeigt, ist das Intervall bei einer Anzahl von Freiheitsgraden unter 10 breiter als (0,63; 1,35). Mit Zunahme der Freiheitsgrade wird das Intervall schmaler, wie aus den Werten in der nachfolgenden Tabelle 7 ersichtlich.

Tabelle 7 Freiheitsgrade sowie Untergrenze und Obergrenze bei einer Wahrscheinlichkeit von 90 %

Freiheitsgrade	Von Untergrenze und Obergrenze gebildetes Intervall
5	(0,48; 1,49)
10	(0,63; 1,35)
15	(0,70; 1,29)
20	(0,74; 1,25)
25	(0,76; 1,23)
30	(0,79; 1,21)
35	(0,80; 1,19)
40	(0,81; 1,18)

Daher benötigen Sie bei einer Wahrscheinlichkeit von 90 % ca. 35 Freiheitsgrade, um eine Fehlerspanne von 20 % für die geschätzte Standardabweichung der Wiederholbarkeit zu erzielen. Wie bereits ausgeführt, sind die Freiheitsgrade gleich der Anzahl der Teile * Anzahl der Prüfer * (Anzahl der Replikationen – 1). Somit liefert die typische Empfehlung von 10 Teilen, 3 Prüfern und 2 Replikationen Freiheitsgrade (30), die diese Anforderung nahezu erfüllen. Um eine Fehlerspanne von 10 % bei einer Wahrscheinlichkeit von 90 % zu erhalten, benötigen Sie etwa 135 Freiheitsgrade (siehe Abbildung 1).