

6 Statistische Auswertung von Simulationen

6.1 Dauer von Simulationen: Benötigte Stichprobengröße (LK 9.4.1)

Durch Simulationen gewonnene Mittelwerte sind nur Schätzungen und insofern mit Fehlern behaftet. Der wahre Wert befindet sich mit Wahrscheinlichkeit $1-\alpha$ im Konfidenzintervall um den empirischen Mittelwert. Je kleiner das Konfidenzintervall ist, desto genauer ist also die Schätzung. Für kleine Konfidenzintervalle braucht man allerdings viele Stichproben, d.h. man muss lange simulieren. Das führt zur Frage, wie lange man simulieren sollte um hinreichend genaue Schätzungen für den Mittelwert zu erhalten. Herangehensweise: Kontrolle der Simulationsdauer durch den absoluten bzw. relativen Fehler des Konfidenzintervalls.

6.1.1 Absoluter Fehler

- $\bar{X}(n)$ ist der geschätzte und μ ist der tatsächliche Mittelwert
- Absoluter Fehler der Schätzung $|\bar{X}(n) - \mu|$ (5.1)
- Mit Wahrscheinlichkeit von mindestens $1-\alpha$ gilt:
 - Der wahre Mittelwert μ liegt im Konfidenzintervall $[\bar{X}(n) - \delta(n, \alpha); \bar{X}(n) + \delta(n, \alpha)]$
 - Der absolute Fehler ist kleiner als die Halblänge des Konfidenzintervalls $\delta(n, \alpha)$
 - Algorithmus für die Simulation eines maximalen absoluten Fehlers δ^*
 - Inkrementiere die Stichprobenmenge der Größe n
 - Berechne neu: $\bar{X}(n), S^2(n), \delta(n, \alpha)$
 - Solange bis $\delta(n, \alpha) \leq \delta^*$
 - Abschätzung der Anzahl benötigter Stichproben während der Laufzeit
 - Nutzen: Manche Simulationen brauchen enorm viel Rechenzeit um eine Stichprobe zu berechnen. Für sehr lange Simulationen ist eine ungefähre Abschätzung der Gesamtdauer der Simulation von Interesse.
 - Gegeben: Stichprobenmenge mit n Werten, $S^2(n), \delta(n, \alpha)$

$\Rightarrow \bar{X}(n)$ hat einen *absoluten Fehler* von höchstens

$$\delta(n, \alpha) = t_{n-1, 1-\alpha/2} \cdot \sqrt{\frac{S^2(n)}{n}}$$

- Frage: Wie viele Stichproben n_{\min} werden etwa benötigt, damit ein absoluter Fehler von höchstens δ^* (mit Wahrscheinlichkeit $1-\alpha$) nicht überschritten wird?

- Antwort

➤ Annahme: $S^2(m) \approx S^2(n)$

$$\Rightarrow n_{\min} = \min_{m \geq n} \left(t_{m-1, 1-\alpha/2} \cdot \sqrt{\frac{S^2(n)}{m}} \leq \delta^* \right) \quad (5.2)$$

6.1.2 Relativer Fehler

- Relativer Fehler: $\gamma = \frac{|\bar{X}(n) - \mu|}{|\mu|} \cdot 100\%$ (oder auch Prozentfehler) (5.3)

- Mit Wahrscheinlichkeit von mindestens $1-\alpha$ gilt:

- $|\bar{X}(n) - \mu| \leq \delta(n, \alpha)$ und $|\mu| \geq |\bar{X}(n) - \delta(n, \alpha)|$
- Falls das Konfidenzintervall nicht die 0 enthält, ist der maximale

$$\text{relative Fehler } \gamma_{\max} = \frac{\delta(n, \alpha)}{|\bar{X}(n) - \delta(n, \alpha)|} \geq \frac{|\bar{X}(n) - \mu|}{|\mu|} = \gamma$$

(5.4)

- Achtung: Der maximale relative Fehler ist nicht $\frac{\delta(n, \alpha)}{|\bar{X}(n)|}$. Vielmehr gilt

$$\frac{\delta(n, \alpha)}{|\bar{X}(n)|} \leq \frac{\gamma_{\max}}{1 + \gamma_{\max}}, \text{ d.h. der intuitive Ausdruck unterschätzt den}$$

maximalen relativen Fehler.

- Algorithmus für die Simulation eines maximalen relativen Fehlers γ_{\max}

- Inkrementiere die Stichprobenmenge der Größe n

- Berechne neu: $S^2(n)$, $\delta(n, \alpha)$

- Solange bis $\frac{\delta(n, \alpha)}{|\bar{X}(n) - \delta(n, \alpha)|} \leq \gamma_{\max}$

- Abschätzung der Anzahl benötigter Stichproben während der Laufzeit

- Gegeben: Stichprobenmenge mit n Werten, Stichprobenvarianz $S^2(n)$

$\Rightarrow \bar{X}(n)$ hat mit Wahrscheinlichkeit $1-\alpha$ einen *absoluten Fehler*

$$\text{von höchstens } \delta(n, \alpha) = t_{n-1, 1-\alpha/2} \cdot \sqrt{\frac{S^2(n)}{n}}$$

- Frage: Wie viele Stichproben n_{\min} werden etwa benötigt, damit ein relativer Fehler von höchstens γ_{\max} (mit Wahrscheinlichkeit $1-\alpha$) nicht überschritten wird?

- Antwort

➤ Annahme: $S^2(m) \approx S^2(n)$

$$\Rightarrow n_{\min} = \min_{m \geq n} \left(\frac{t_{m-1, 1-\alpha/2} \cdot \sqrt{\frac{S^2(n)}{m}}}{|\bar{X}(n)|} \leq \frac{\gamma_{\max}}{1 + \gamma_{\max}} \right) \quad (5.5)$$

6.2 Arten von Simulationen

Die statistische Datenerhebung ist von der Art der Simulation abhängig. Es gibt

- Terminierende Systeme / Simulationen
- Nicht-terminierende Systeme / Simulationen

6.2.1 Terminierende Simulationen (LK 9.4)

- Vorgegebene natürliche Simulationsdauer
- Beispiele
 - Begrenzte Simulationszeit auf Grund von Modellannahmen (z.B. 8-Stunden-Tag)
 - Simulation der Zeit bis zum Eintreffen eines bestimmten Ereignisses („first passage time“, z.B. beim „random walk“)

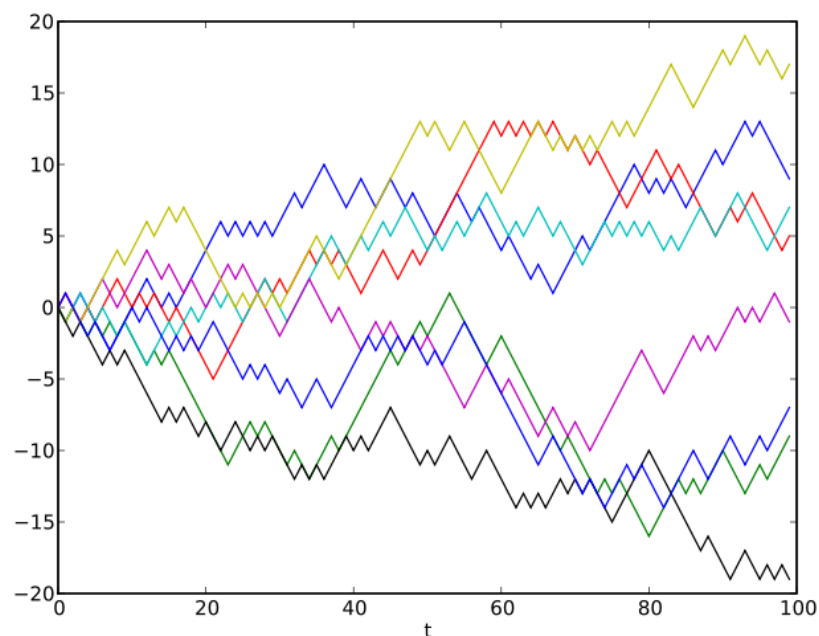


Abbildung 1: Beim „random walk“ startet ein Prozess in Zustand 0 und verändert in jedem Zeitschritt seinen Zustand um ± 1 mit Wahrscheinlichkeit 50%. Varianten existieren. Die Zeit bis zum ersten Erreichen einer Schwelle wird „first passage time“ genannt. (Bildquelle: http://de.wikipedia.org/wiki/Random_Walk)

- Art der Aussagen, die in einer Simulation gewonnen werden
 - Maße bezogen auf die Dauer eines gesamten Simulationslaufes (z.B. Anzahl bedienter Kunden pro Tag)
 - First passage time
 - Systemgrößen, die während der Simulation beobachtet werden (z.B. Warteschlangenlängen, Server-Auslastungen ...)

- Durchführung
 - Sinnvolle Wahl für die Initialisierung des Systemzustands der Simulation
 - Mehrmalige Durchführung von Simulationsläufen zur Generierung von Stichproben X_1, \dots, X_n
 - Unabhängigkeit der Stichproben durch unterschiedliche Zufallszahlen für die Simulationsläufe
 - Konstruktion von Konfidenzintervallen
 - Feste Anzahl n von Simulationsläufen
 - Simulation der Stichproben X_1, \dots, X_n
 - Wahl eines Signifikanzniveaus (z.B. $\alpha=5\%$)
 - Berechnung des Konfidenzintervalls durch einmalige Anwendung von Formel (4.8) auf Basis der Normal- bzw. der Student-t-Verteilung
 - Vorgegebener maximaler relativer oder absoluter Fehler
 - Sukzessive Simulation von Stichproben und Berechnung von Konfidenzintervallen auf Grund der bisherigen Stichprobenmenge bis relativer oder absoluter Fehler klein genug ist.

6.2.2 Nicht-terminierende Simulationen (LK 9.5)

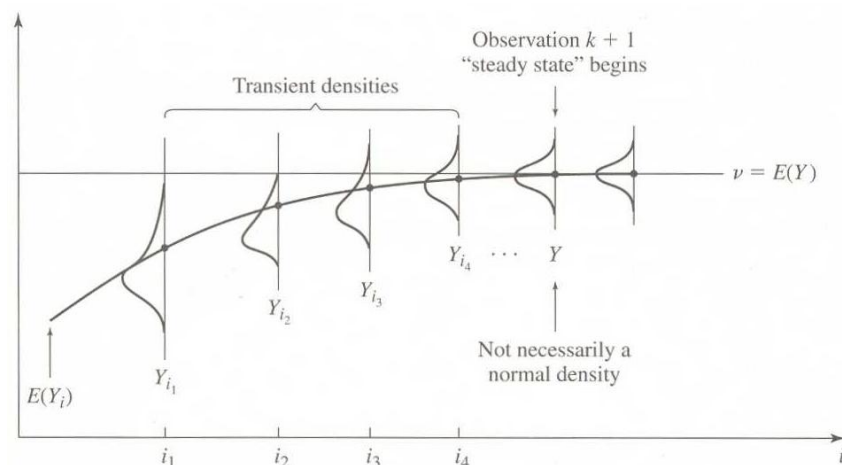
- Technische Systeme, bei denen keine charakteristischen Zeiteinheiten zu erkennen sind, die sich auch in der Praxis als abgeschlossene Operation wiederholen und keinen Einfluss auf das zukünftige Geschehen des Systems haben.
- Beispiele: Langzeitverhalten von
 - Fabrikanlagen
 - Rechnernetzen
- Art der Aussagen
 - Verteilungen über das Langzeitverhalten von Systemgrößen wie Serverbelegungen, Durchsatz, Warteschlangenlängen, Wartezeiten, Verlustwahrscheinlichkeiten, etc.
 - Anfangsphase nicht relevant, da untypisch
 - Beschreibung durch die Verteilung der beobachteten Größen im „eingeschwungenen“ System (steady state distribution)
- Probleme bei der Auswertung
 - Anfangsphase liefert untypische Ergebnisse
 - Beispiel
 - Leere Fabrik zu Beginn
 - Wartezeiten erster fertiger Produkte sind gering
 - Wartezeiten von Produkten steigen anfangs mit zunehmender Simulationszeit bis Fabrik einen typischen Betriebszustand erreicht hat
 - Aufeinanderfolgende Stichproben sind oft positiv korreliert, bisherige Stichprobentheorie verlangt aber unabhängige Stichproben
 - Beispiel
 - Typische Zeitreihe von Wartezeiten:
0, 0, 1, 1, 2, 2, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 6, 5, 5, 4, 3, 3, 2, 1, 1, ...
 - Autokorrelation
 - Berechnung von Konfidenzintervall für Mittelwert mit konventioneller Stichprobentheorie nicht möglich

Diese Probleme werden in den nächsten beiden Abschnitten behandelt.

6.3 Methoden zur Bestimmung der transienten Phase

6.3.1 Transiente und stationäre Phase eines Systems

- Beispiel: M/M/1- ∞ Wartesystem
 - Beobachtete Zufallsvariablen: Wartezeit Y_i des Kunden i
 - Wartezeiten zu Beginn hängen vom Anfangszustand des Systems ab
 - Bei leerem Anfangszustand hat erster Kunde keine Wartezeit
 - Zweiter Kunde hat geringe durchschnittliche Wartezeit, da er höchstens auf einen anderen Kunden warten muss
 - Mit zunehmender Kundenzahl verschimmt diese Abhängigkeit vom Anfangszustand des Systems
- Betrachtung der simulierten Zufallsvariablen als zeitdiskreten stochastischen Prozess
 - Aufeinanderfolgende Zufallsvariablen Y_i einer konkreten Simulation sind voneinander abhängig und eine Ausprägung dieses Prozesses.
 - Betrachte viele solcher Ausprägungen; dann kann für jedes Y_i eine Verteilung angegeben werden:



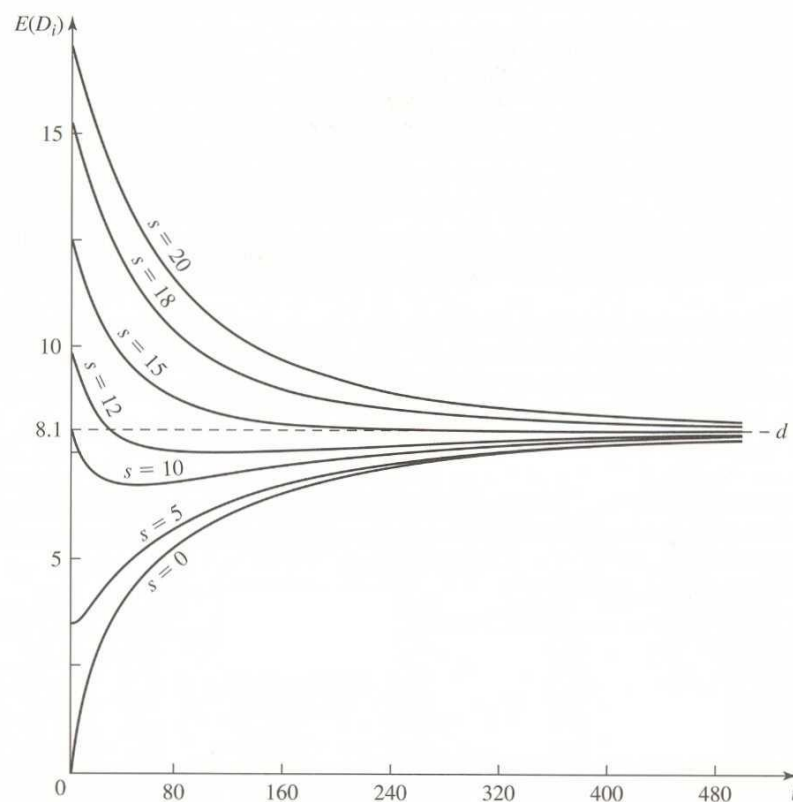
Transient and steady-state density functions for a particular stochastic process Y_1, Y_2, \dots and initial conditions I .

Abbildung 5.1: Die Verteilung der Zufallsvariablen Y_i (z.B. Wartezeit bei M/M/1- ∞) konvergiert in der transienten Phase gegen die Verteilung der Zufallsvariablen im stationären Zustand, in dem sich die Verteilungen aufeinanderfolgender Zufallsvariablen nicht mehr ändern. (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 499)

- Verteilung von Y_i ist stationär, falls sie der Verteilung von Y_k gleicht für alle $k > i$
- Gesucht: stationäre Verteilung von Y um Aussagen über das Langzeitverhalten des Systems zu machen
 - Die untypische Anfangsphase vor dem „stationären bzw. eingeschwungenen Zustand“ bzw. dem Erreichen des statistischen Gleichgewichts der Simulation wird transiente Phase, Aufwärmphase

oder Einschwingphase (warm-up period, initial transient, startup problem) genannt.

- Datenerhebung für statistische Aussagen über Y nur im eingeschwungenen Zustand
- Lösungen
 - Verwerfen der Stichproben in der Aufwärmphase der Simulation: Wann ist die zu Ende?
 - Solange simulieren bis der Einfluss der Aufwärmphase verschwindet: Wie lange?
 - Wahl eines geschickten Anfangszustandes (z.B. eines sehr wahrscheinlichen bzw. typischen Zustands): Welcher ist das?
- Probleme beim Erkennen des Endes der transienten Phase
 - Gleichgewichtszustand durch eine Verteilung gekennzeichnet, nicht durch einen konkreten Wert
 - Länge der transienten Phase stark von Bedingungen bei Simulationsbeginn abhängig



$E(D_i)$ as a function of i and the number in system at time 0, s , for the $M/M/1$ queue with $\rho = 0.9$.

Abbildung 5.2: Der Mittelwert einer Zufallsvariablen D_i (z.B. Wartezeit bei $M/M/1-\infty$) konvergiert in der transienten Phase gegen den Mittelwert d der stationären Verteilung. Dabei hängt er vom Anfangszustand s (z.B. Anzahl der Kunden in der Warteschlange zu Simulationsbeginn) ab. (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 500)

6.3.2 *Naive Methode zur Erkennung des eingeschwungenen Systems*

- Direkter Erkennungsversuch über Verteilungen für Y_i
- Parallele Simulationen erzeugen Ausprägungen der Y_i
- Erstellung von Histogrammen für alle Y_i
- Statistisches Gleichgewicht ist erreicht, wenn sich das Histogramm von Y_i von dem seiner Nachfolger nicht mehr unterscheidet
- Wegen des enorm hohen Simulationsaufwandes nicht durchführbar; Zufallsvariablen können auch mehrdimensional sein!
- Sobald man die stationäre Phase erkannt hätte, bräuchte man möglicherweise keine weiteren Simulationen mehr, da die stationäre Zustandsverteilung oft schon das gesuchte Ziel ist.

6.3.3 Graphisches Verfahren von Welch (LK 9.5.1)

- Heuristisches, graphisches Verfahren zur Bestimmung der Länge der transienten Phase
 - Reduktion des naiven Ansatzes auf die Mittelwertbetrachtung
 - Einfachere Betrachtung als naive Methode
 - Schnellere Konvergenz als naive Methode
- Messwerte aus n relativ langen Testläufen (n=10 oder 20, X_i^j ist Messwert i im Lauf j)

- Reihe der gemittelten Messreihen bilden $\bar{X}_i = \frac{\sum_{j=1}^n X_i^j}{n}$ (5.5)

- Bestimme das i, ab dem die gemittelte Messreihe konvergiert und setze $l=i$. Index l markiert das Ende der initialen Phase.
- Bei den „richtigen“ Simulationsläufen werden dann die ersten l Werte nicht berücksichtigt, d.h. bei den neuen Messwerten X_0, X_1, \dots, X_m schätzt man den

Mittelwert durch $\bar{X}(m,l) = \frac{\sum_{i=l+1}^m X_i}{m-l}$ (5.6)

- Beispiel 1: Durchlaufzeit der Kunden in einem M/M/1- ∞ Wartesystem

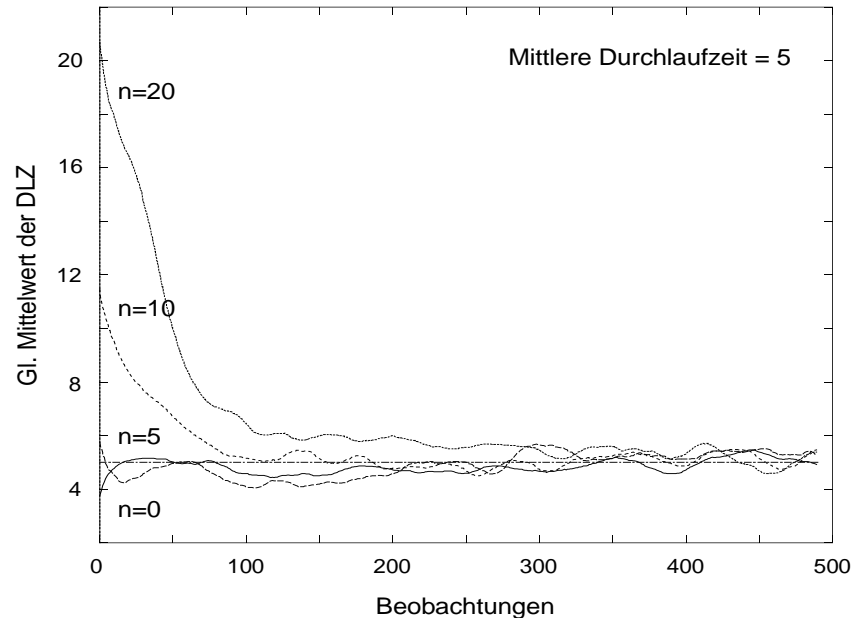
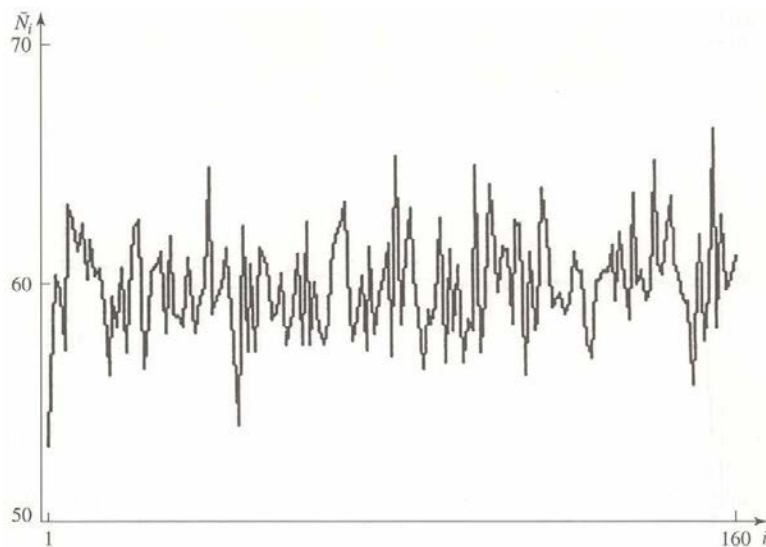


Abbildung 5.3: M/M/1-Wartesystem mit 80% Auslastung und $E[B]=1$: Mittelwert der Durchlaufzeiten bei 100 Replikationen ohne Glättung durch gleitenden Durchschnitt.

- Beispiel 2: Stündlicher Durchsatz einer kleinen Fabrik



Averaged process for hourly throughputs, small factory.

Abbildung 5.4: Das Ende der transienten Phase ist aus der gemittelten Reihe mehrerer Simulationsläufe noch nicht zu erkennen. (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 523)

- An Beispiel 2 sehen wir, dass die Reihe der gemittelten Zeitreihen nicht immer eine initiale Phase erkennen lässt.
- Darum wende anschließend einen gleitenden Mittelwert (moving average) mit halber Fensterbreite w zur Beseitigung hochfrequenter Schwingungen an:

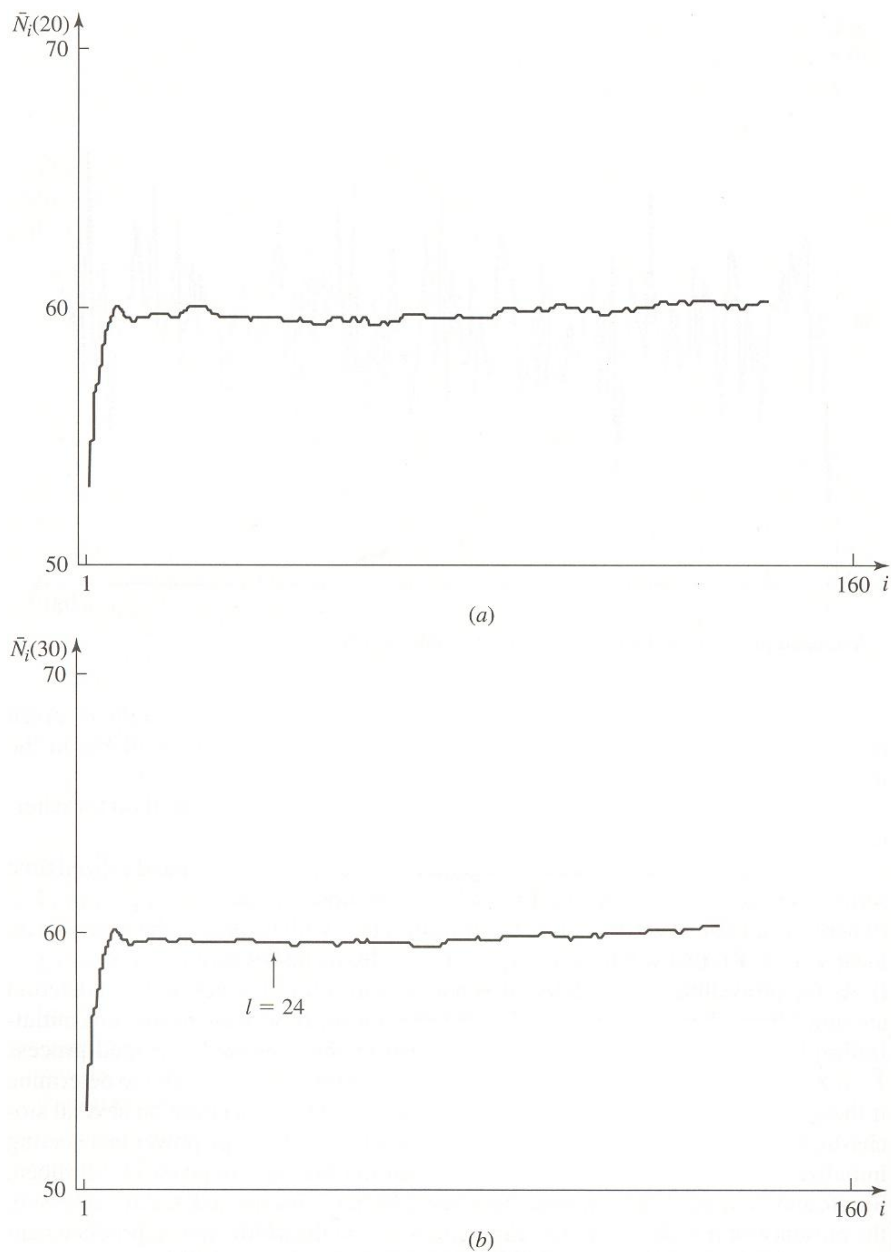
$$\bar{X}_i(w) = \begin{cases} \frac{\sum_{s=-w}^w \bar{X}_{i+s}}{2w+1} & \text{falls } i > w \\ \frac{\sum_{s=-(i-1)}^{i-1} \bar{X}_{i+s}}{2i+1} & \text{falls } 1 \leq i \leq w \end{cases} \quad (5.7)$$

- Beispiel:

X_i :	1.0	2.0	2.0	1.0	2.0	3.0	3.0	4.0	5.0	2.0	6.0	5.0	2.0	3.0	3.0	1.0	2.0
\bar{X}_i :	1.7	1.7	1.7	2.0	2.7	3.3	4.0	3.7	4.3	4.3	4.3	3.3	2.7	2.3	2.0		

Abbildung 2: Originale und gemittelte Zeitreihe (halbe Fensterbreite $w=1$).

- Verschiedene Werte w testen, geglättete Mittelwertreihe $\bar{X}_i(w)$ mit geeignetem Parameter w wählen, bei der sich eine Konvergenz der Reihe deutlich feststellen lässt, um Index l zu bestimmen.
- Falls keine brauchbaren Kurven zu finden, 10 bis 20 weitere Läufe durchführen und Prozedur wiederholen.



Moving averages for hourly throughputs, small factory: (a) $w = 20$; (b) $w = 30$.

Abbildung 5.5: Das Einschwingen des Systems ist zu Beginn des Prozesses gut durch die Glättung der gemittelten Messreihe zu erkennen. (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 524)

Als Ende der Einschwingphase wird $l=24$ Stunden gewählt. Dieser Wert ist zur Sicherheit deutlich größer angenommen als der Wert, den man in der Graphik als Grenze erkennt (z.B. $l=6$ Stunden).

6.3.5 Elegante Berechnung des Endes der Transienten Phase

- Beispiel: M/M/n Verlustsystem
 - System mit n parallelen Servern ohne Warteraum
 - Modell zum Dimensionieren von Leitungskapazitäten
 - Angenommen, zu Beginn der Simulation wäre schon ein Kunde im System gewesen.
 - Die Wahrscheinlichkeit, dass er sich nach t Zeit noch immer im System befindet ist $\exp(-\mu \cdot t)$, wobei $1/\mu$ die mittlere Bedienzeit ist.
 - Falls das System zu Beginn der Simulation voll besetzt gewesen wäre, ist die Anzahl der zur Zeit t noch aktiven Nutzer vom Beginn der Simulation nach $\text{Binom}(n, \exp(-\mu \cdot t))$ verteilt.
 - Nach welcher Zeit t ist die Wahrscheinlichkeit, dass kein Benutzer vom Beginn der Simulation noch aktiv ist, größer als p?

$$(1 - \exp(-\mu \cdot t))^n \geq p$$

$$t \geq \frac{\ln(1 - \sqrt[n]{p})}{-\mu} = -\ln(1 - \sqrt[n]{p}) \cdot \frac{1}{\mu} \quad (5.10)$$
 - Zahlenbeispiel: p=0.99, n=100:
Nach 9.21-facher mittlerer Bedienzeit ist die Wahrscheinlichkeit, dass kein Benutzer seit Systemstart noch aktiv ist, größer als 99%
- Anwendbarkeit stark modellabhängig

6.3.6 Sinnvolle Initialisierung des Systems

- Beispiel: M/M/n-∞ Wartesystem
 - System mit n parallelen Servern mit Warteraum
 - Mittlere Anzahl von Kunden im System: $a = \lambda / \mu$ wobei
 - die mittlere Zwischenankunftszeit $1/\lambda$ und
 - die mittlere Bedienzeit $1/\mu$ ist.
 - Initialisiere die Simulation mit a Kunden und ermittle ihre Restarbeitszeit
 - Wegen Gedächtnislosigkeit der Exponentialverteilung kann einfach eine volle zufällige Bedienzeit dafür genommen werden.
- Anwendbarkeit ebenfalls stark modellabhängig

6.3.7 Test auf Schätzwertverfälschung (Bias) durch transiente Phase

Siehe später bei Signifikanztests von Hypothesen!

6.4 Stichprobentheorie für korrelierte Zeitreihen (LK 4.4)

- Ziel: Konstruktion von Konfidenzintervallen für den Mittelwert einer Zeitreihe X_i
- Problem: Bisherige Stichprobentheorie nicht anwendbar
 - Theorie fordert unabhängige Stichproben
 - Werte einer Zeitreihe (aus Simulation) sind in der Regel stark korreliert

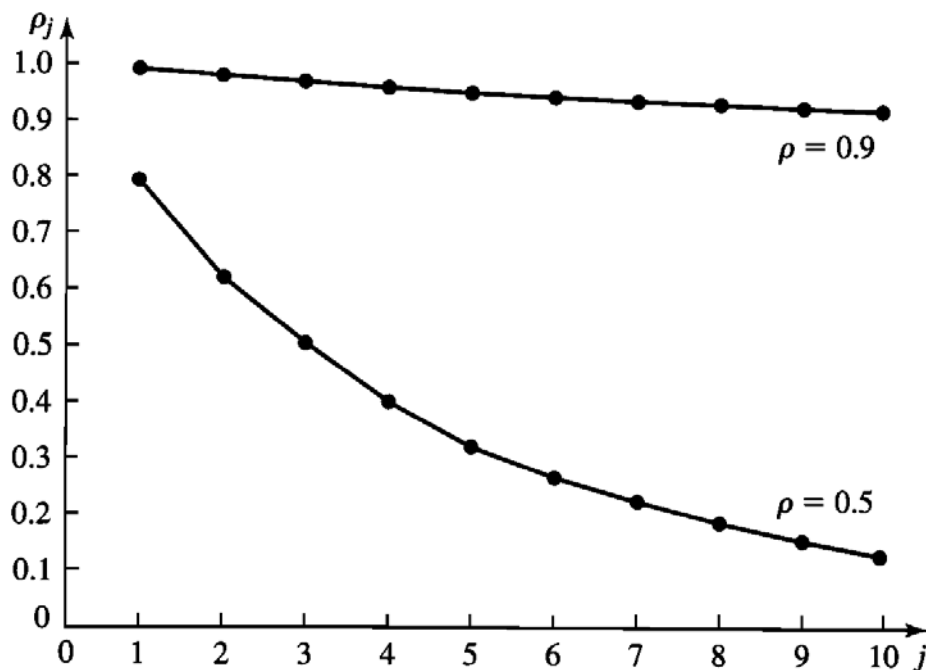


Abbildung 5.6: Korrelation aufeinanderfolgender Stichproben bei einem M/M/1-Wartesystem mit Auslastung ρ . (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 248)

6.4.1 Stichprobentheorie für korrelierte Zeitreihen (LK 4.4)

- Es sei $c = 1 + 2 \cdot \sum_{0 < j < n} (1 - j/n) \cdot \rho_j$ (5.11)
 - n ist die Größe der Stichprobe X_i und es gilt $0 \leq i < n$
 - ρ_j sind die Korrelationskoeffizienten der Zeitreihe X_i und es gilt $0 < j < n$
 - Bei positiv korrelierten Zeitreihen gilt $c > 1$.
- Problem: bei korrelierten Stichproben schätzt $S^2(n)$ die Varianz σ^2 einer Stichprobenmenge falsch, vielmehr gilt:

$$E[S^2(n)] = \text{VAR}[X] \cdot \frac{n-c}{n-1} \quad (5.12)$$

Folge: Stichprobenvarianz wird bei positiver Korrelation unterschätzt.

- Zusätzlich berechnet sich die Varianz des Mittelwertes anders als bei unabhängigen Stichprobenwerten (siehe Gl. 4.1):

$$\text{VAR}[\bar{X}(n)] = \frac{\text{VAR}[X]}{n} \cdot c \quad (5.13)$$

Folge: auch sie wird mit Gl. 4.1 unterschätzt, so dass davon abgeleitete Konfidenzintervalle zu klein wären.

- Aus Kombination von (5.12) und (5.13) ergibt sich

$$\text{VAR}[\bar{X}(n)] = \frac{E[S^2(n)]}{n} \cdot \frac{n-1}{n-c} \cdot c \quad (5.14)$$

Für unabhängige Stichproben erhalten wir mit $c=1$ genau Gl. 4.1

- Beispiel aus Abbildung 5.6:

- Annahme: exponentieller Abfall der Korrelationswerte: $\rho_n = \rho_1 \cdot p^{n-1}$

(Wert für p kann mit $p = \sqrt[n]{\frac{\rho_{10}}{\rho_1}}$ aus Abbildung 5.6 ermittelt werden.)

- Sei $c(n)$ der für c errechnete Wert bei einer Stichprobengröße von n .

Dann bezeichnen wir mit $\text{bias}(n) = \frac{n-1}{n-c(n)} \cdot c(n)$ den multiplikativen Bias bezüglich der konventionell berechneten Varianz.

- Beispiel 1: Korrelationen $\rho_1 = 0.88, \rho_{10} = 0.22 \Rightarrow p=0.8572$

n	c(n)	bias(n)
10	2.95	3.78
100	5.58	5.86
1000	5.96	5.99
10000	6.00	6.00

- Beispiel 2: Korrelationen $\rho_1 = 0.99, \rho_{10} = 0.9 \Rightarrow p=0.9895$

n	c(n)	bias(n)
10	5.24	9.91
100	36.04	55.78
1000	85.01	92.81
10000	93.04	93.90
100000	93.84	93.93
1000000	93.93	93.93

- Konventionelle Konfidenzintervallhalblänge: $\delta(n, \alpha) = z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2(n)}{n}}$

Korrektur: $\delta^*(n, \alpha) = z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2(n)}{n} \cdot \text{bias}(n)} = \delta(n, \alpha) \cdot \sqrt{\text{bias}(n)}$

6.5 Robuste Strategien für die Berechnung von Konfidenzintervallen für korrelierte Zeitreihen

Wir entwickeln zwei Strategien zur Berechnung eines Konfidenzintervalls für den Mittelwert einer korrelierten Zeitreihe.

- Naiver, direkter Ansatz
 - Aufwändig, weil zur Berechnung von c (5.11) sehr viele Autokorrelationen ρ_j benötigt werden
 - Schlecht, weil es keine guten Schätzer für die Autokorrelationen ρ_j gibt
- Indirekter Ansatz basierend auf Mittelwertstichproben
 - Problem: Einzelmesswerte sind korreliert und allgemeine Stichprobentheorie ist nicht anwendbar zur Berechnung von Konfidenzintervallen
 - Idee: Bilde Mittelwerte $\bar{X}(m)_j$ aus m aufeinanderfolgenden Einzelmesswerten
 - Die Reihe der Mittelwerte $\bar{X}(m)_j$ hat denselben Mittelwert wie die Reihe der Einzelmesswerte.
 - Mittelwerte sind i.d.R. normalverteilt
 - Unkorreliertheit der Reihe der Mittelwerte $\bar{X}(m)_j$ kann empirisch überprüft werden.
 - Keine oder nur schwache Korrelation dieser Mittelwerte, falls m groß genug, so dass annähernd Unabhängigkeit angenommen werden kann.
 - Berechnung von approximativen Konfidenzintervallen auf Basis dieser Mittelwerte möglich

Neben dem indirekten Ansatz basierend auf Mittelwertstichproben gibt es noch weitere Verfahren zur Berechnung von Konfidenzintervallen für Mittelwerte von Einzelmesswerten. Sie alle versuchen die Varianz der Mittelwerte $\text{VAR}[\bar{X}(n)]$ auf unterschiedliche Arten zu schätzen (siehe LK 9.5.3).

Es gibt zwei vorrangige Verfahren zur Erzeugung der Mittelwertstichproben: Batch-Means und Replicate-Delete.

6.5.1 Batch-Means Verfahren (LK 9.5.3)

- Variante A: die Stichproben liegen schon vor
 - Entfernung der transienten Phase
 - Einteilung der Messwerte in n gleichgroße Intervalle (Batches) der Größe m ; vermeide dabei Korrelationen zwischen benachbarten Mittelwerten; Verfahren für die Wahl von m :
 - Bestimmung der empirischen Autokorrelationskoeffizienten $\hat{\rho}_k$ der Messwerte (ohne jene der transienten Phase)
 - Wahl von k_{\max} so, dass die $\hat{\rho}_k$ für $k > k_{\max}$ praktisch verschwinden
 - Randbedingungen
 - $m \geq 10 \cdot k_{\max}$ (m : Größe des Batches)
 - $n \geq 10$ (n : Anzahl der Batches)
 - Berechnung der Konfidenzintervalle basierend auf den Mittelwerten der Messwerte in den n Batches

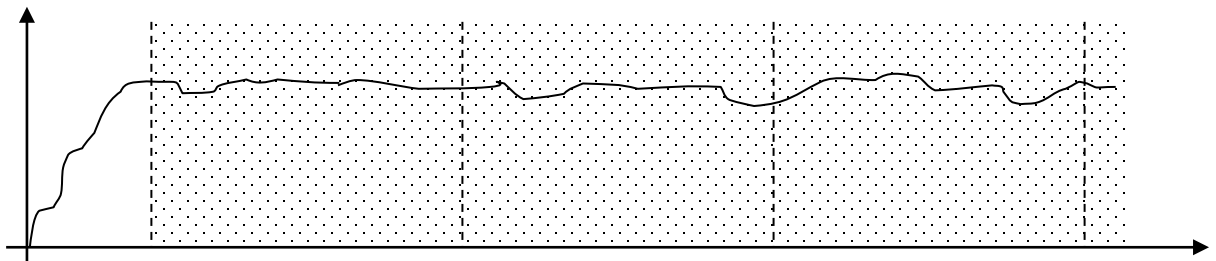


Abbildung 5.8: Beim Batch-Means Verfahren wird ein sehr langer Simulationslauf gestartet, dessen transiente Phase verworfen und der Restlauf wird in n Teile zerlegt, deren Mittelwerte unabhängige Stichproben zur Ermittlung von Konfidenzintervallen ergeben.

- Variante B: die Stichproben müssen noch erzeugt werden
 - Ziel: Simulation von Stichproben bis Konfidenzintervalle klein genug sind hinsichtlich des relativen oder absoluten Fehlers (siehe 5.1)
 - Bestimmen der transienten Phase
 - Wahl eines geeigneten m
 - Simulation so vieler Batches bis Konfidenzintervalle klein genug sind
 - Überprüfung der Mittelwertstichproben auf Korrelationen um Fehler durch zu kleine Batches auszuschließen
- Vorteile
 - Geringer Aufwand durch einmalige transiente Phase
 - Fehler bei der Schätzung der transienten Phase haben geringen Einfluss

- Nachteile
 - Bestimmung von m (bzw. von n und m) schwierig
 - Korrelationen möglich \Rightarrow Berechnung der Korrelationskoeffizienten zur Vermeidung korrelierter Intervallmittelwerte notwendig

6.5.2 Replicate/DeleteVerfahren (LK 9.5.2)

- Durchführung
 - Transiente Phase bestimmen
 - Läufe durchführen (Replicate), Länge: Vielfaches der transienten Phase
 - Messwerte der transienten Phase abschneiden (Delete)
- Berechnung der Konfidenzintervalle basierend auf den Mittelwerten der Einzelläufe bzw. Anwendung der Verfahren aus 5.1

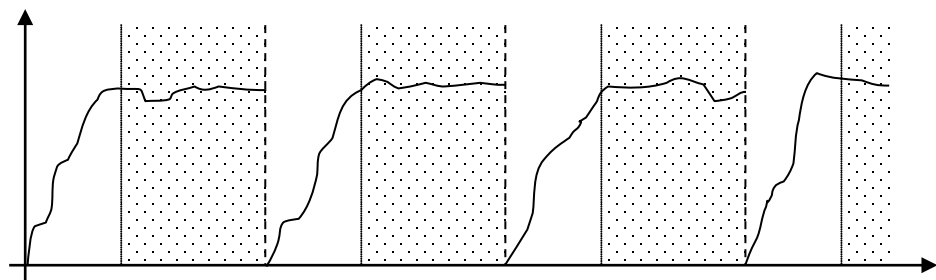


Abbildung 5.7: Beim Replicate/Delete Verfahren werden n Läufe hintereinander gestartet, die transiente Phase wird verworfen und die Mittelwerte der Restläufe ergeben unabhängige Stichproben zur Ermittlung von Konfidenzintervallen.

- Vorteile
 - Einfach
 - Nicht anfällig gegen Korrelationen
 - Von allen gängigen Simulationspaketen unterstützt
- Nachteile
 - Korrekte Bestimmung des Endes der transienten Phase hat großen Einfluss auf die Güte der Ergebnisse
 - Hoher Aufwand durch die Simulation vieler ungenutzter transienter Phasen

6.5.3 Konfidenzintervalle für andere Leistungskriterien als Mittelwerte (LK 9.5.4)

- Wahrscheinlichkeiten
 - Prinzipiell wie bei Mittelwerten (vgl. 4.1.5)
 - Verteilungswerte $P(X \leq x)$ sind auch Wahrscheinlichkeiten
 - Falls solche Wahrscheinlichkeiten für mehrere Werte von x berechnet werden, kann man eine Verteilungsfunktion mit Konfidenzintervallen approximieren
- Quantile
 - Wiederholung der Definition des Quantils: $Q_p = \inf(x \in W(X) : P(X \leq x) \geq p)$ ist ein Wert aus dem Wertebereich der Stichprobenmenge, evtl. mit Einheiten behaftet
 - p ist gegeben, Q_p ist gesucht
 - Konfidenzintervalle sind schwierig anzugeben, kein allgemein anerkanntes Verfahren

- Quotienten

- Beobachtungsgrößen X_i und Y_i (iid)
- Gesucht $\phi = \frac{E[X]}{E[Y]}$, Achtung: nicht $\phi = E\left[\frac{X}{Y}\right]!$
- Problem: einfacher Schätzer $\hat{\phi}_C(n) = \frac{\bar{X}(n)}{\bar{Y}(n)}$ ist nicht erwartungstreu und hat keine Konfidenzintervalle
- Lösung: Jackknife Estimator $\hat{\phi}_J(n)$
 - zwar kein erwartungstreuer Schätzer für $\phi = \frac{E[X]}{E[Y]}$
 - aber ein besserer als $\hat{\phi}_C(n)$.
- Idee: Konstruiere aus den X_i und Y_i neue Zufallsvariablen θ_i um durch diese neuen Stichproben Konfidenzintervalle zu erzeugen:

- $\theta_i = n \cdot \hat{\phi}_C(n) - (n-1) \cdot \frac{\sum_{0 \leq j < n, j \neq i} Y_j}{\sum_{0 \leq j < n, j \neq i} X_j}$ für $0 \leq i < n$

- Mittelwert: $\hat{\phi}_J(n) = \frac{1}{n} \cdot \sum_{0 \leq i < n} \theta_i$

- Varianz: $\hat{\sigma}_J^2(n) = \frac{\sum_{0 \leq g < n} (\theta_g - \hat{\phi}_J(n))^2}{n-1}$

- Es gilt: $\frac{\hat{\phi}_J(n) - \phi}{\sqrt{\hat{\sigma}_J^2(n)/n}} \xrightarrow{\mathcal{D}} N(0,1)$ (Konvergenz der Verteilung)

- Konfidenzintervall für das Konfidenzniveau $1-\alpha$:

$$\hat{\phi}_J(n) \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\sigma}_J^2(n)/n}$$

- Nachteil: für große Stichproben aufwändig

6.6 Simulation nicht-ergodischer Prozesse

6.6.1 Ergodische und nicht-ergodische Prozesse

6.6.1.1 Definitionen

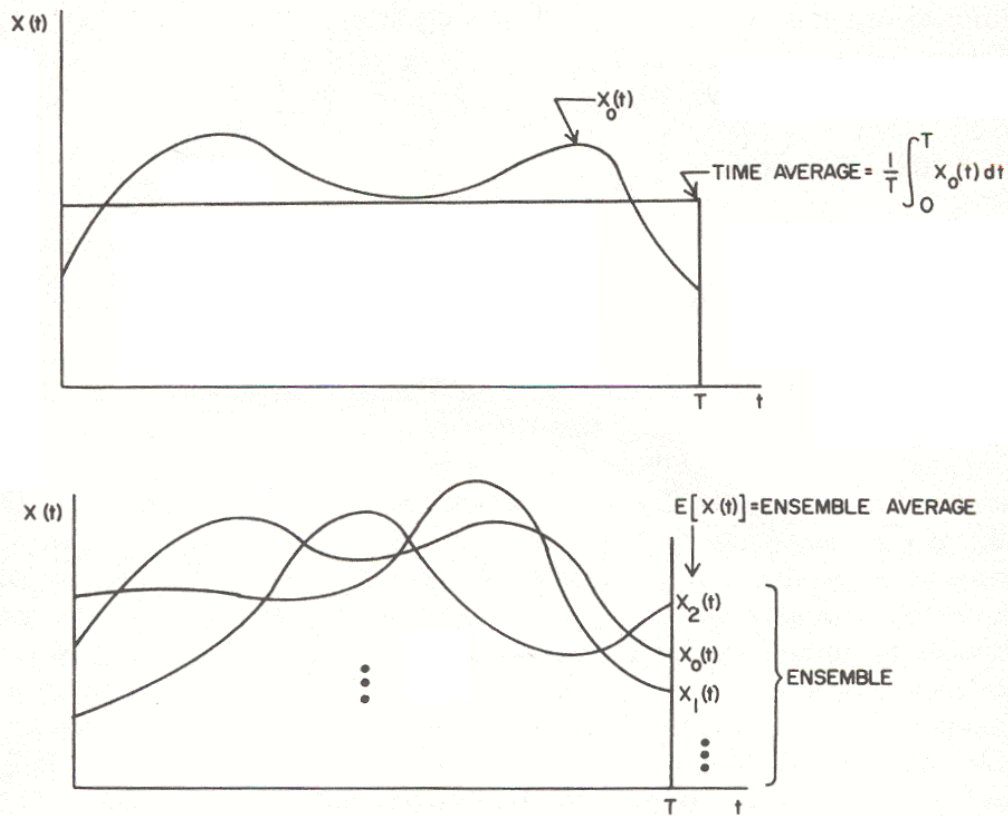


Abbildung 5.9: Zeitmittel und Ensemble-Mittel eines stochastischen Prozesses. (aus: Fundamentals of Queuing Theory, Gross/Harris, 1985, 2nd edition)

Definition: Zeit-Mittel eines stochastischen Prozesses über den Zeitraum T :

$\overline{X_T^k} = \frac{1}{T} \cdot \int_0^T [X_0(t)]^k dt$, wobei mit $X_0(t)$ nur eine einzige Realisierung des stochastischen Prozesses beobachtet wird.

Definition: Ensemble-Mittel eines stochastischen Prozesses zum Zeitpunkt t :

$E[X^k(t)] = \lim_{n \rightarrow \infty} \frac{\sum_{0 \leq i < n} [X_i(t)]^k}{n} = m_k(t)$, wobei $X_i(t)$ unterschiedliche Realisierungen des stochastischen Prozesses zum Zeitpunkt t sind.

Definition: Ein stochastischer Prozess ist ergodisch für das k-te Moment, falls gilt:

$$\lim_{t \rightarrow \infty} m_k(t) = \lim_{T \rightarrow \infty} \overline{X_T^k} = \overline{X^k}$$

Definition: Ein stochastischer Prozess ist ergodisch, falls er für alle Momente ergodisch ist.

Folgerung: Falls ein stochastischer Prozess Mittelwert-ergodisch ist, kann man ihn mit der Batch-Means-Methode simulieren.

6.6.1.2 Beispiel eines Nicht-Ergodischen Prozesses

Zu Beginn des Prozesses wird eine Zahl gewürfelt, die der Prozess für immer als Zustand annimmt.

- Die Zeitmittelwerte $\overline{X_T^1}$
 - haben den Wertebereich $\{1, 2, 3, 4, 5, 6\}$.
 - hängen zwar nicht von T ab aber von der Realisierung des Prozesses.
 - sind somit nicht eindeutig.
- Der Ensemble-Mittelwert $m_1(t)$ für beliebige Zeitpunkte ist 3,5.
- Wegen $m_1(t) \neq \overline{X_T^1}$ ist dieser Prozess nicht Mittelwert-ergodisch.

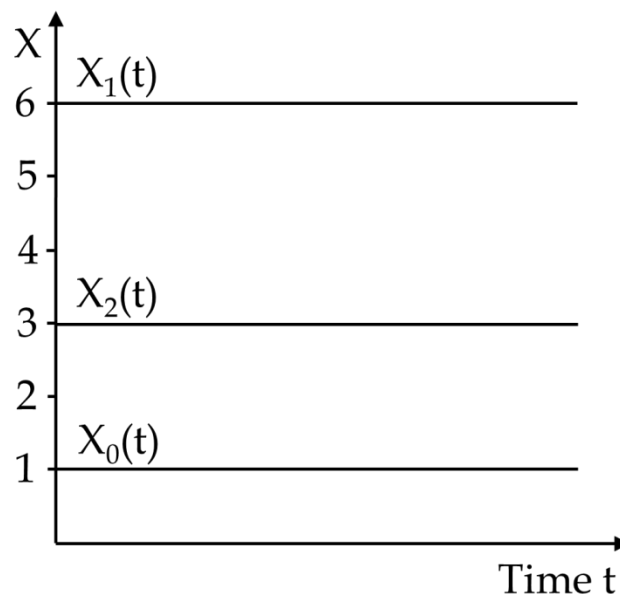
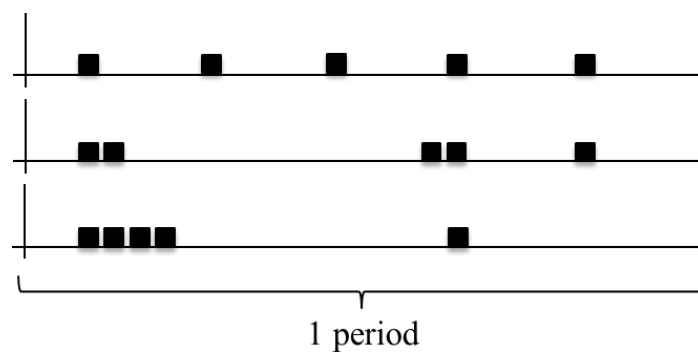


Abbildung 3: Drei Realisierungen eines einfachen nicht-ergodischen Prozesses.

6.6.1.3 Weitere Beispiele für Ergodische und Nicht-Ergodische Prozesse

- Ergodische Prozesse
 - Wartesysteme (GI/GI/n- ∞)
 - Warteverlustsysteme (GI/GI/n-S)
 - Charakteristisch
 - Zufällige und unabhängige Ankunfts- und Bedienzeiten

- Nicht-ergodische Prozesse
 - Wartesysteme ($n \cdot D / D / m - \infty$)
 - Warteverlustsysteme ($n \cdot D / D / m - S$)
 - Charakteristisch
 - n verschiedene Quellen senden Kunden in die Warteschlange
 - Die Zwischenankunftszeit zwischen Kunden einer Quelle ist konstant \Rightarrow periodische Ankunftszeiten bzw. -muster
 - Kunden erfahren konstante Bedienzeit



- **Abbildung 4: Unterschiedliche Sendemuster (Phasen der Sendezeitpunkte innerhalb einer Periode).**
 - Dieselben Systeme mit Varianten bezüglich Paketgrößen
 - On/off-Modulation (On/off-Quellen); erzeugt von Sprach-Codecs mit Silence Detection
 - Variable Paketgrößen gemäß einer Verteilung, wobei aufeinanderfolgende Paketgrößen i.d.R. korreliert sind; erzeugt von Sprach-Codecs für den Mobilfunk oder Skype

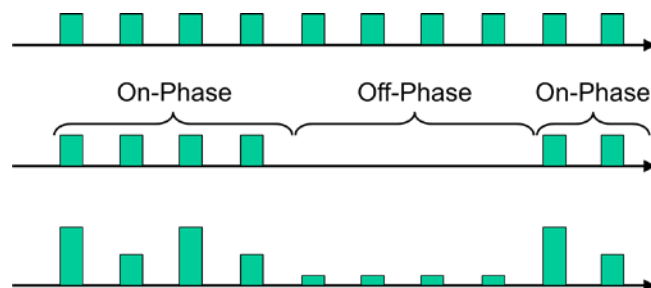


Abbildung 5: Konstante Paketgrößen, on/off-modulierte Paketgrößen und korrelierte, unterschiedliche Paketgrößen. Pakete mit unkomprimierten Sprachproben, die alle 20 ms gesendet werden, enthalten $\frac{1 \text{ byte}}{\text{sample}} \cdot \frac{8000 \text{ samples}}{s} \cdot 20 \text{ ms} = 160 \text{ bytes}$ Nutzlast.

6.6.2 Simulation von periodischen Systemen

- Anwendungsbeispiel: Echtzeitverkehr
 - Datenquelle generiert mit einer festen Periode A Sprachproben, Videoframes, etc.
 - Datenpakete
 - (i) können gleich groß sein,
 - (ii) können aber auch gemäß einer empirischen Verteilung iid sein oder on/off-moduliert sein
 - Wenn die Daten von n solchen Quellen mit gleicher Periode A auf ein gemeinsames Übertragungsmedium gemultiplext werden, dann wiederholen sich die Ankunfts Muster in einem Zeitfenster der Länge A
 - Die Paketverzögerungen sind für jedes Intervall der Länge A für die Verbindungen einer Quelle
 - (i) gleich bzw.
 - (ii) ähnlich,
 da sie hauptsächlich durch die Phasenlage der Quellen innerhalb einer Periode (Ankunfts Muster) bestimmt sind.
- Folgerung für Simulationen
 - Problem: Wartezeit hängt von Ankunfts Muster ab
 - Lösung
 - Simuliere viele Ankunfts Muster und
 - (i) Berechne die Wartezeit deterministisch für eine Periode
 - (ii) Simuliere die Wartezeit stochastisch für mehrere Perioden
 - Die Mittelwerte der Wartezeit für unterschiedliche Ankunfts Muster bilden die Stichprobenmenge für
 - die Ensemble-Mittelung für die Gesamtstatistik
 - die Berechnung von Konfidenzintervallen
- Im Fall (i) handelt es sich um eine typische Monte-Carlo-Simulation
 - Unabhängige Muster werden ausgewürfelt.
 - Auf Basis eines Musters wird eine Stichprobe rein deterministisch berechnet, was durch Discrete Event Simulation geschehen kann, aber andere Verfahren sind auch möglich.
 - Auf Basis einer Stichprobenmenge wird eine Statistik erstellt.
- Im Fall (ii) benutzt man Discrete Event Simulation (dynamisch) und die Replicate/Delete Methode
- Batch-Means Methode nicht anwendbar wegen fehlender Mittelwert-Ergodizität des Systems

6.7 Vergleich alternativer Systemkonfigurationen (LK 10)

- Testkriterium: Der Unterschied zwischen zwei Systemen ist statistisch signifikant, falls 0 nicht mehr im Konfidenzintervall der Differenz liegt.
- Zwei Methoden zur Bestimmung der Konfidenzintervalle
 - t-Confidence Interval (t-C.I.) für Paare
 - erfordert gleiche Anzahl von Stichproben der zu vergleichenden Systeme
 - Modifizierte t-C.I. für zwei Stichproben
 - erfordert Unabhängigkeit der Stichproben

6.7.1 t-Konfidenzintervall für Paare (Paired-t C.I., LK 10.2.1)

- Gegeben
 - (Mittelwert)stichproben zweier Systeme A und B: X_i^A und X_i^B , $0 \leq i < n$
- Berechne
 - Folge der Differenzen $Z_i = X_i^A - X_i^B$
 - Konfidenzintervall mit Signifikanz α auf Basis der Student-t-Verteilung für den empirischen Mittelwert \bar{Z} (t-Konfidenzintervall)
 - Falls aufeinanderfolgende Differenzen Z_i korreliert sind, dürfen Konfidenzintervalle nur auf Basis von Mittelwertstichproben berechnet werden.
- Falls das t-Konfidenzintervall die Null nicht enthält, dann sind die wahren Mittelwerte für X^A und X^B mit Wahrscheinlichkeit $1-\alpha$ unterschiedlich.
- Vorteil: Annahmen über Unabhängigkeit beider Messreihen nicht notwendig
- Nachteil: Beide Messreihen müssen gleich lang sein
 - Wenn das nicht der Fall ist, können nicht alle Messwerte genutzt werden.
 - Könnte nicht erfüllt sein, wenn z.B. Simulationsdaten und reale Messdaten verglichen werden.

6.7.2 Modifiziertes t-Konfidenzintervall für zwei Stichproben (LK 10.2.2)

- Problem beim paired t-confidence interval: gleichgroße Stichprobenmengen
- Lösung: Welch-Konfidenzintervall (Welch, 1938)

- Annahme

- Unabhängige Stichprobenreihen X_{1j} und X_{2j}
- Normalverteilte Stichproben

- Berechne $\bar{X}_i(n_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$ und $S_i^2(n_i) = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} [X_{ij} - \bar{X}_i(n_i)]^2$

für $i = 1, 2$.

- Schätzung der Freiheitsgrade:

$$\hat{f} = \frac{[S_1^2(n_1)/n_1 + S_2^2(n_2)/n_2]^2}{[S_1^2(n_1)/n_1]^2/(n_1-1) + [S_2^2(n_2)/n_2]^2/(n_2-1)}$$

- Approx. $100(1-\alpha)\%$ C.I.: $\bar{X}_1(n_1) - \bar{X}_2(n_2) \pm t_{[\hat{f}], 1-\alpha} \sqrt{\frac{S_1^2(n_1)}{n_1} + \frac{S_2^2(n_2)}{n_2}}$

- Vorteil: $n_1 = n_2$ nicht nötig
 - Gute Ausnutzung aller vorhandener Daten
- Nachteile: Vorausgesetzte Annahmen oft nicht erfüllt

6.7.3 Gleichzeitiges Schätzen von mehreren Konfidenzintervallen (LK 10.3)

- Gegeben
 - System, bei dem n unterschiedliche Parameter, z.B. Mittelwerte μ_i , $0 \leq i < n$, geschätzt werden mit Konfidenzintervallen für ein Signifikanzniveau von α_i
- Frage: Wie groß ist die Wahrscheinlichkeit, dass die wahren Werte für die Parameter alle in den Konfidenzintervallen enthalten sind?
 - Antwort: $\prod_{0 \leq i < n} (1 - \alpha_i)$

6.7.3.1 Bonferroni-Ungleichung

- Bonferroni-Ungleichung: $\prod_{0 \leq i < n} (1 - \alpha_i) \geq \max(1 - \sum_{0 \leq i < n} \alpha_i, 0)$
- Beispiel 1
 - $n=10$, gewünschtes Signifikanzniveau für Gesamtvergleich: $\alpha=10\%$
 - Wie groß muss Signifikanzniveau für Einzelintervalle sein?
 - Mit Bonferroni: $\alpha_i=1,000\%$
 - Exakt: 1,048%
- Beispiel 2
 - $n=10$, Signifikanzniveau für Einzelintervalle: $\alpha_i=0.1$
 - Signifikanzniveau für Gesamtvergleich (Wahrscheinlichkeit, dass nicht alle 10 Parameter in ihrem Konfidenzintervall enthalten sind)
 - Mit Bonferroni: 100%
 - Exakt: gut 65%
 - Keine hilfreiche Aussage

6.7.3.2 Vergleich mit einem Referenzsystem (LK 10.3.1)

- k Systeme, eines davon Referenzsystem $\Rightarrow k-1$ Vergleiche
- Mögliches Ziel: Referenzsystem ist besser als die $k-1$ Vergleichssysteme
- Gewünschtes Signifikanzniveau für Gesamtvergleich α
- Mit Bonferroni: $\alpha_i = \frac{\alpha}{k-1}$

6.7.3.3 Paarweise Vergleiche (LK 10.3.2)

- k Systeme $\Rightarrow k(k-1)/2$ Vergleiche
- Mögliches Ziel: Sortierung von k Systemen nach Leistung
- Gewünschtes Signifikanzniveau für Gesamtvergleich α
- Mit Bonferroni: $\alpha_i = \frac{2 \cdot \alpha}{k \cdot (k-1)}$

6.8 Simulationsbeschleuniger

Zwei Beispiele

- Geeignete Modellierung von Systemen mit abhängigen Zufallsvariablen
- Vergleich von Systemen unter vergleichbaren Bedingungen

6.8.1 Geeignete Modellierung von Systemen mit abhängigen Zufallsvariablen (LK 11.6 Conditioning)

- Gegeben
 - Zwei Zufallsvariablen $X(t)$ und $Y(t)$ in einer Simulation
 - Beispiel
 - X : Örtliche Verteilung von Nutzern in einer Mobilfunkzelle
 - Y : Interferenz in einer Mobilfunkzelle in Abhängigkeit von X ; für diese kann ein Mittelwert $\bar{Y}(X)$ analytisch berechnet werden.
- Gesucht: \bar{Y}
- Naiver Ansatz
 - X kann simuliert werden in Abhängigkeit der Zeit
 - Batch-Means oder Replicate/Delete Methode für Erzeugung von Konfidenzintervallen
 - Hoher Aufwand um den Wertebereich von X hinreichend gut abzudecken, wenn die Werte von X zeitlich korreliert sind (=sich nur langsam verändern)
- Cleverer Ansatz
 - Erzeuge Samples (Snapshots) X_i mit Hilfe einer Monte-Carlo-Simulation
 - Berechne Mittelwert und Konfidenzintervall von Y auf Basis der Stichproben $\bar{Y}(X_i)$

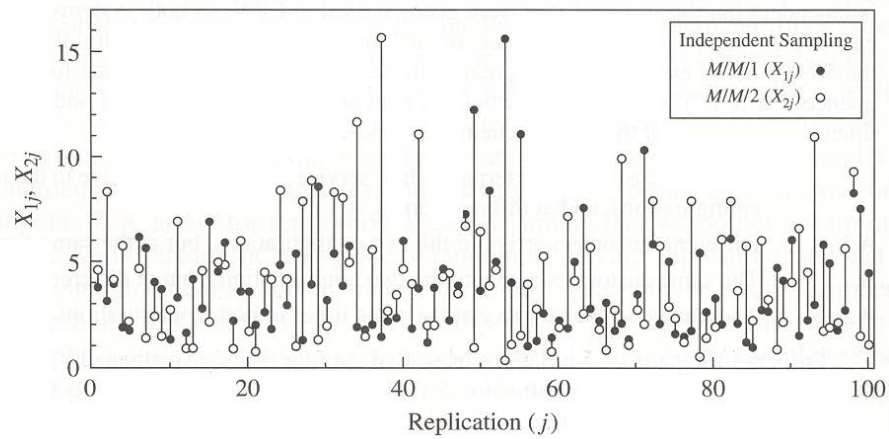
6.8.2 Vergleich von Systemen unter vergleichbaren Bedingungen (LK 11.2 Common Random Numbers)

6.8.2.1 Herstellung von vergleichbaren Bedingungen

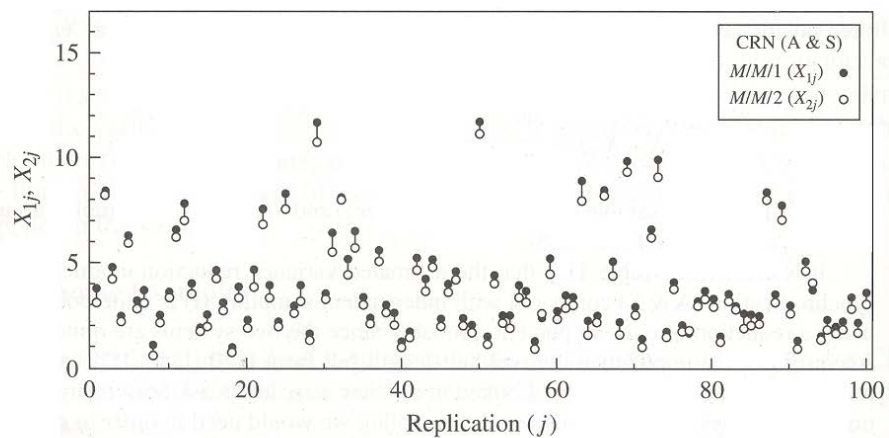
- Beispiel: M/M/n-Wartesystems
 - Vergleiche zwei Systemalternativen, welches führt zu kürzeren Wartezeiten?
 - M/M/1 mit einfacher Bedienzeit
 - M/M/2 mit doppelter Bedienzeit
- Verwende bei der Simulation aller Systeme dieselben Zufallszahlen zur Erzeugung von Zufallsvariablen
 - aufeinanderfolgender Zwischenankunftszeiten
 - aufeinanderfolgender Bedienzeiten
- Wichtig: entsprechende Zufallszahlen müssen für korrespondierende Zwecke verwendet werden
 - Nutze separate Zufallszahlenströme für Zwischenankunftszeiten und Bedienzeiten
 - Achtung bei Vergleich von Systemen mit unterschiedlichen Verteilungen: nicht anwendbar, wenn variable Anzahl von Zufallszahlen zur Erzeugung entsprechender Zufallsvariablen benötigt wird

6.8.2.2 Beispiel

Führt ein M/M/1 oder ein vergleichbares M/M/2-Wartesystem zu kürzeren Wartezeiten?



(a)



M/M/1 vs. M/M/2: individual replications. (b)

Abbildung 5.10: Differenz der Wartezeiten in vergleichbaren M/M/1 und M/M/2 Systemen. Der Vorteil des M/M/2 Systems ist bei der Nutzung gemeinsamer Zufallszahlen viel leichter ersichtlich. (aus Law/Kelton: "Simulation Modeling and Analysis", 3rd Edition, S. 592)

6.8.2.3 Mathematische Erklärung der Simulationsbeschleunigung durch vergleichbare Bedingungen

Systemvergleiche unter vergleichbaren Bedingungen scheinen vernünftig – aber warum muss man weniger simulieren?

- Unter vergleichbaren Bedingungen führen ähnliche Systeme zu tendenziell größeren bzw. kleineren Ergebnissen X_i^A und X_i^B (z.B. mittlere Wartezeiten beim i-ten Simulationsdurchlauf), d.h. sie sind positiv korreliert, so dass $\text{COV}(X^A, X^B) > 0$
- Achtung: durch die Korrelation von X_i^A und X_i^B wird nicht die Korrelation aufeinander folgender Zufallsvariablen $Z_i = X_i^A - X_i^B$ erhöht! Also kann die übliche Stichprobentheorie noch angewandt werden. Im obigen Beispiel sind die X_i^A und X_i^B Mittelwertstichproben einzelner Simulationsläufe.

- Allgemein gilt: $\text{VAR}[a \cdot X + b \cdot Y] = a^2 \cdot \text{VAR}[X] + 2 \cdot a \cdot b \cdot \text{COV}[X, Y] + b^2 \cdot \text{VAR}[Y]$

Beweis:

Es gilt:

$$E[a \cdot X + b \cdot Y] = a \cdot E[X] + b \cdot E[Y]$$

$$\begin{aligned} E[(a \cdot X + b \cdot Y)^2] &= E[a^2 \cdot X^2 + 2 \cdot a \cdot b \cdot X \cdot Y + b^2 \cdot Y^2] = \\ &= a^2 \cdot E[X^2] + 2 \cdot a \cdot b \cdot E[X \cdot Y] + b^2 \cdot E[Y^2] \end{aligned}$$

Daraus folgt:

$$\begin{aligned} \text{VAR}[a \cdot X + b \cdot Y] &= E[(a \cdot X + b \cdot Y)^2] - E[(a \cdot X + b \cdot Y)]^2 = \\ &= a^2 \cdot E[X^2] + 2 \cdot a \cdot b \cdot E[X \cdot Y] + b^2 \cdot E[Y^2] - (a \cdot E[X] + b \cdot E[Y])^2 = \\ &= a^2 \cdot E[X^2] + 2 \cdot a \cdot b \cdot E[X \cdot Y] + b^2 \cdot E[Y^2] - a^2 \cdot E[X]^2 - 2 \cdot a \cdot b \cdot E[X] \cdot E[Y] - b^2 \cdot E[Y]^2 = \\ &= a^2 \cdot \{E[X^2] - E[X]^2\} + 2 \cdot a \cdot b \cdot \{E[X \cdot Y] - E[X] \cdot E[Y]\} + b^2 \cdot \{E[Y^2] - E[Y]^2\} = \\ &= a^2 \cdot \text{VAR}[X] + 2 \cdot a \cdot b \cdot \text{COV}(X, Y) + b^2 \cdot \text{VAR}[Y] \end{aligned}$$

- Anwendung auf Vergleich zweier Systeme mittels $Z = X - Y$:

$$\text{VAR}[Z] = \text{VAR}[X] - 2 \cdot \text{COV}[X, Y] + \text{VAR}[Y]$$

- Verringerung der Varianz von Z durch positive Korrelation von X und Y

- Berechnung des Konfidenzintervalls: $Z \pm t_{n-1, 1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2(n)[Z]}{n}}$

- Reduktion der empirischen Varianz $S^2(n)[Z]$ von Z führt zu kleinerer Anzahl benötigter Stichproben n um Konfidenzintervalle mit gewünschter Maximalgröße zu erreichen
 - Daraus ergibt sich die Simulationsbeschleunigung