6. Ergebnisauswertung

Fortsetzung Kap. 4; Erinnerung:

Beobachtung (der Exekution) stochastischen Simulators entspricht (mathematisches Modell)
Beobachtung (der Realisierung) stochastischen Prozesses
endliche "Länge" Realisierung:
Beobachtung ist "Zeitreihe"

Beispiele aus Kap. 4:

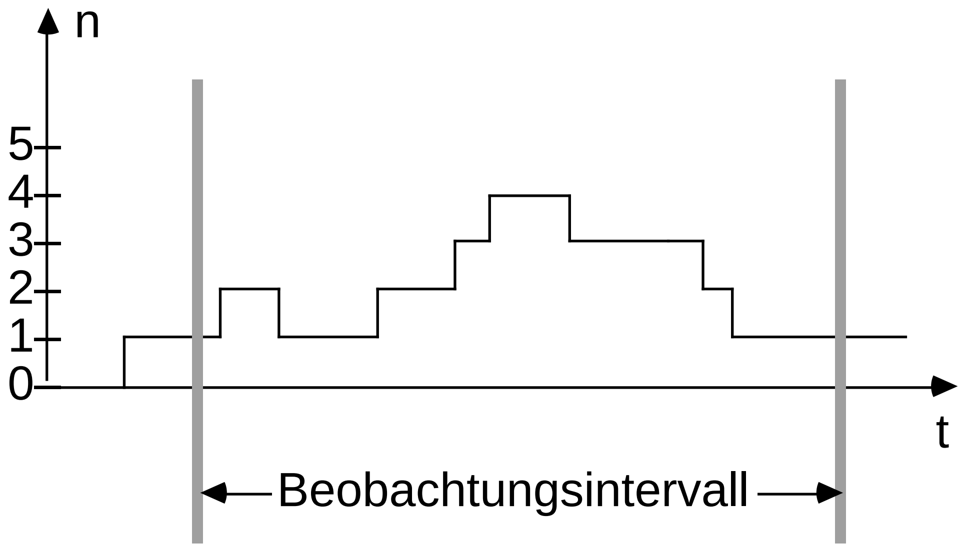


Abbildung 6.0.1: Beobachtung Zustand Bankschalter

Realität: Jeder Betriebsablauf "etwas anders", statistisch schwankend

Simulator: Jede Simulatorexekution (potentiell) "etwas anders", statistische Schwankung erreicht durch Änderung Saat-Wert(e) ZV-Generator(en)

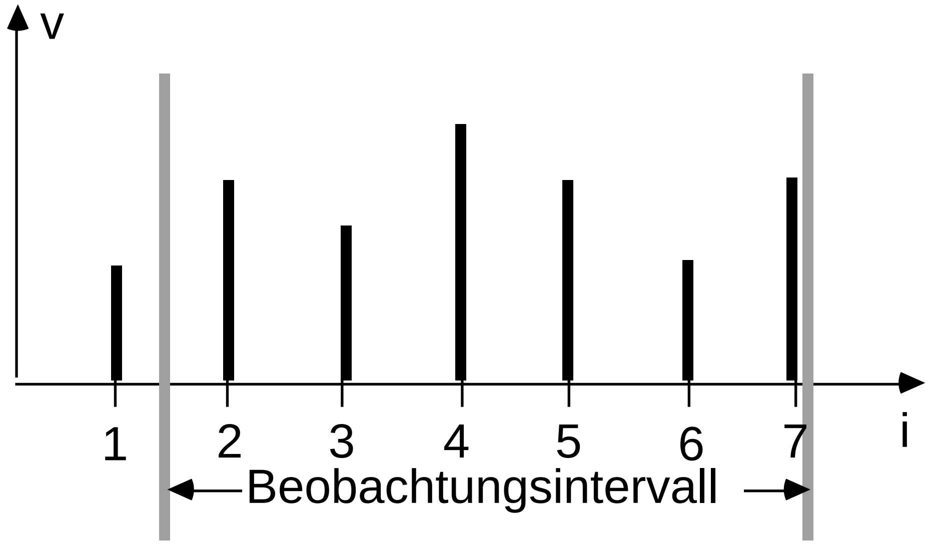


Abbildung 6.0.2: Beobachtung Verweilzeit Kunden

Merke: Beobachtungsintervall,
Anfangszustand u.U. wesentlich

Verfolgte Auswertungsfrage war:
" Wie **ist** Beobachtung für festes t* (bzw. festes i*) ? "

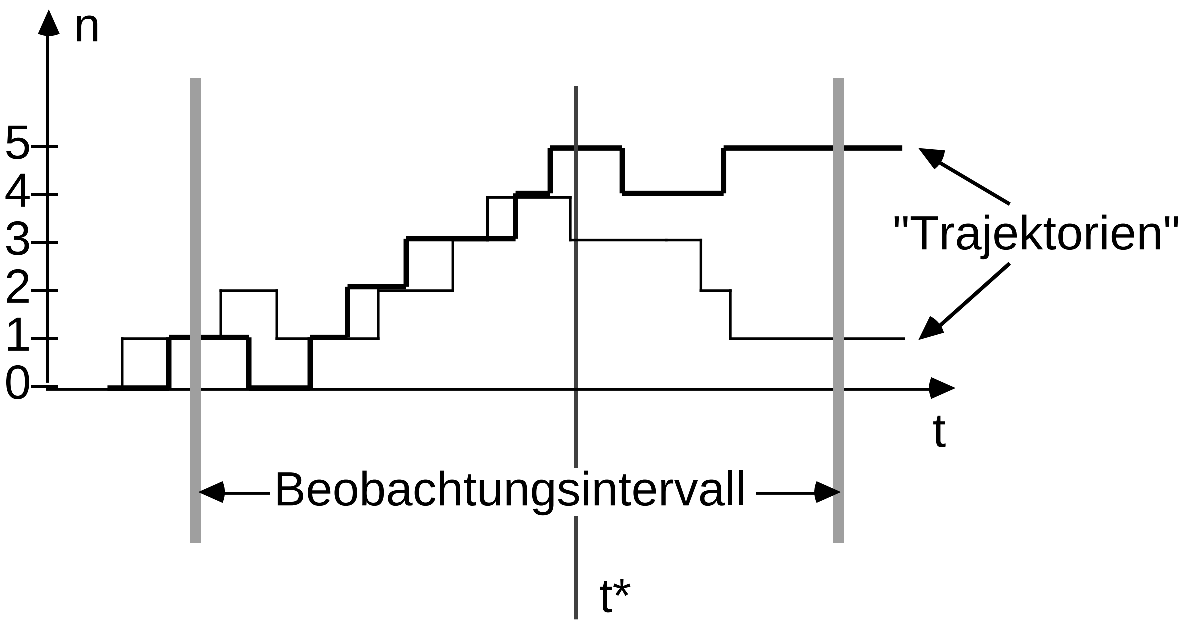


Abbildung 6.0.3: Prinzip Ensemble-Analyse

- Antwort i.allg. abhängig von "Startzustand" z_0 Simulator

- auch enthalten:
  Fragen / Antworten
  bzgl. gesamten berücksichtigten Intervalls (0,t*) bzw. [1,i*]
  wie etwa: mittlere Kundenzahl in (0,t*)
  mittlere Verweilzeit in [1,i*]

  Mittelwerte
  längs Beobachtung / **quer** zu Beobachtungen
  nicht verwechseln

- bisheriges Vorgehen:
  Antworten aus "Replikationen"
  –> Stichprobe für Y(t*)
  –> statistische Auswertung
  benannt: **statische Analyse,**
  Ensemble-Analyse,
  Analyse mit festem Horizont

Ahnungen:

Für bestimmte stochastische Prozesse könnte gelten

$Y(t|z_0) \to_{t \to \infty} Y$ mit Verteilung von Y
unabhängig von t,
unabhängig von z_0

also: Einfluß des Anfangszustands "ausgestorben"
Verhalten des Prozesses "bleibt so"

Falls Ahnungen zutreffend,
(also ab hinreichend großem t)
heißt
Zeitintervall ab t: **stationäre Phase** des Prozesses
Verhalten ab t: **eingeschwungener Zustand** d. Prozesses

Dann könnte u.U., für ein t*>t,

(da ja Verhalten und Beobachtungen (!) "so bleiben")

Ensemble-Analyse (quer zu Replikationen)
ersetzt werden durch
Zeitreihen-Analyse (längs einer Simulatorexekution)

Zunächst aber Wiederholung (Kap.4) und Ausbau von

6.1 Ensemble-Analyse

Modell der Beobachtungsgröße für festes t* (i*): ZV Y

Ensemble-Analyse liefert (hoffentlich!):
unabhängige Stichprobe $(y_1,y_2,\ldots,y_n)$

Dazu Forderungen an "seeding"
(widersprüchlich!, s. Mihr72, Jain91, LaKe91):
unbestritten: seeds nie wiederholen,
nie unerlaubte seeds ("0", gerade Zahlen, …)
aber: "zufällig" wählen (ZZ-Tafel, extra ZZ-Generator)
vs. "bewußt" wählen (nicht-überlapp'de "streams")
und: "genau ein stream" (Korrelationsgefahr)
vs. "ein stream je Zweck" (andere Korrelationsgefahr)
Daraus Forderungen an Funktionalität ZZ-Generatoren
(Eröffnung bewußter stream-Wahl)

Aus Stichprobe zu schätzen:
Charakteristika der Verteilung von Y(t*)
(vgl. auch Abschn. 5;
warum nur Charakteristika, nicht ganze Verteilung?)

Es ergaben sich:

erwartungstreuer Mittelwertschätzer (4.1.4)

(6.1.1a) $\tilde{\mu}_1 := \frac{1}{n} \sum_1^n Y_i$

sowie dessen Schätzwert

(6.1.1b) $\mu_1^* := \frac{1}{n} \sum_1^n Y_i$

erwartungstreuer Varianzschätzer (4.1.6)

(6.1.2) $\tilde{\sigma}^2 := \frac{1}{n-1} \sum_1^n \left(Y_i - \tilde{\mu}_1\right)^2$

ferner Varianz des Mittelwertschätzers (4.1.7)

(6.1.3) $V\left[\tilde{\mu}_1\right] = \frac{\sigma^2}{n} \qquad \left(= \frac{V[Y]}{n} \right)$

Wir diskutierten Konfidenzintervalle im Prinzip (4.1.8):

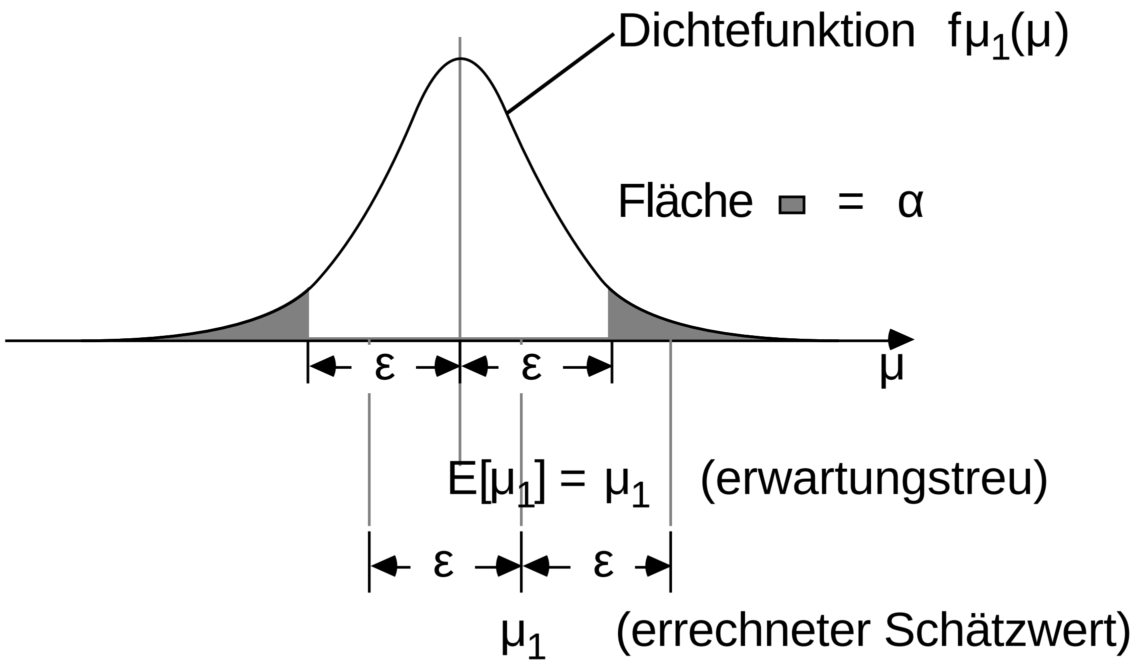


Abbildung 6.1.4: Prinzipskizze Konfidenzintervall

und ermittelten (mithilfe Tschebyscheff-Ungleichung) konkrete $\tilde{\mu}_1$-Konfidenzintervalle (4.1.12)

(6.1.5) $P[\, |\tilde{\mu}_1 - \mu_1| \geq \varepsilon \,] \leq V[Y]/(n \cdot \varepsilon^2)$

wo auf Basis des "pessimistischen" Charakters dieses Konfidenzintervalls auch $\tilde{\sigma}^2$ für V[Y] substituiert werden konnte

Wir erwähnten schon dort, daß

- weniger pessimistische (d.h. engere) KIs zu erhoffen

- falls Informationen über Verteilung von Y (oder Verteilung der Schätzer, z.B. des $\tilde{\mu}_1$) vorliegen und genutzt werden können

(i) Informationen über Verteilung von $\tilde{\mu}$ (ab jetzt statt: $\tilde{\mu}_1$)

aus (6.1.1)

$$\tilde{\mu} := \frac{1}{n} \sum_{1}^{n} Y_i$$

läßt sich auf Basis **zentraler Grenzwertsatz** vermuten, daß $\tilde{\mu}$ für große n approximativ normalverteilt

Zentraler Grenzwertsatz in verschiedenen Fassungen bekannt. Hier konkret (vgl. z.B. LaKe82):

(6.1.6a) zentraler Grenzwertsatz
Seien $Y_1, Y_2, \ldots, Y_n$
u.i.v. ZV mit endlichen $\mu = E[Y]$ und $\sigma^2 = V[Y]$
Sei eine ZV Z definiert zu

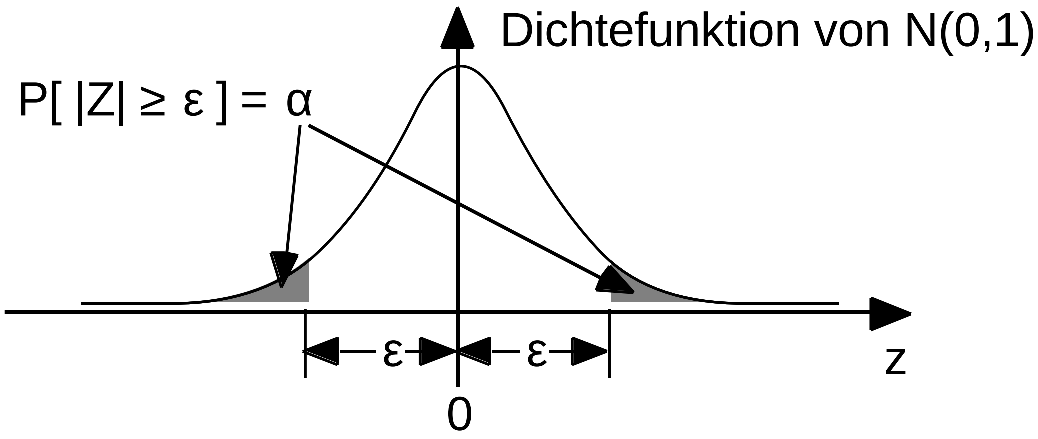
$$Z := \frac{\tilde{\mu} - \mu}{\sigma / \sqrt{n}}$$

Dann approximiert die Verteilung von Z für große n die N(0,1)-Verteilung (die "Standard-Normal-Verteilung") mit Verteilungsfunktion $\Phi(z)$ beliebig genau, d.h.

$$FZ(z) \rightarrow \Phi(z) \text{ für } n \rightarrow \infty$$

(6.1.6b) Zusatz
(6.1.6a) bleibt gültig, wenn σ ersetzt wird durch den erwartungstreuen Schätzer $\tilde{\sigma}$ gemäß (6.1.2)

Damit neues (approximatives, für große n gültiges) Konfidenzintervall
$$P[\,|Z| \geq \varepsilon\,] = \alpha$$
für Z gemäß Skizze (und daraus folgend für $\tilde{\mu}$)



mit ε aus (zweiseitiger) N(0,1)-Tafel

Beispiel: Bsp. 4.1.13 fragte nach 90%-KI, also α=0.1, bei Stichprobenumfang 10

Auf Basis Tschebyscheff wurde gewonnen: $\tilde{\mu} \pm \tilde{\sigma}$

ε-Wert hier aus N(0,1)-Tafel: 1.645
unser Interesse: KI für $\tilde{\mu}$
wo mit (6.1.6a,b)

$$\tilde{\mu} = Z\,\tilde{\sigma}/\sqrt{n} + \mu$$

wir erhalten - mit ε_α aus N(0,1)-Tafel -

(6.1.7a) $$P\left[\,\left|\tilde{\mu} - \mu\right| \geq \varepsilon_\alpha \; \tilde{\sigma}/\sqrt{n}\,\right] = \alpha$$

und als Konfidenzintervall für $\tilde{\mu}$ (bei Berechnung: μ*)

(6.1.7b) $$\tilde{\mu} \pm \varepsilon_\alpha\, \tilde{\sigma}/\sqrt{n}$$

im Beispiel (n=10):

$$\mu \pm 0.520 \;\; \tilde{\sigma}$$

Approximationsgüte umso schlechter,
je "nicht-normaler" Y verteilt: dann große n benötigt

(ii) Informationen über Verteilung von Y
speziell: falls Y normalverteilt, dann … ?

(6.1.8) Einschub t-Verteilung

- Erinnerung:
  Seien $Y_1, Y_2, \dots, Y_k$
  k unabhängige, identisch N(0,1)-verteilte ZV, d.h.

$$f_{Y_i}(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \qquad i = 1,2,\dots,k$$

  Dann hat

$$Y = \sum_{i=1}^{k} Y_i^2$$

  eine χ^2**-Verteilung mit k Freiheitsgraden**

- Sei X eine (unabhängig von allen Y_i)
  zusätzliche N(0,1)-Verteilte
  Dann ist

$$T = X \Big/ \sqrt{\sum_{i=1}^{k} Y_i^2 \Big/ k}$$

  eine ZV, deren Verteilung
  t-Verteilung mit k Freiheitsgraden
  (auch: t_k-Verteilung) genannt wird

- Familien der χ^2_k- und der t_k-Verteilungen sind tabelliert
  (keine explizite funktionale Form für V'Fkt. angebbar)

- für große n strebt die t-Verteilung der N(0,1)-Vert'g zu
  ("groß"-Anhaltspunkt: n>1000)

- Frage: Was ist (unter Annahme: Y normalverteilt) exakte Verteilung einer Testgröße T (anal. 6.1.6b)

$$T := \frac{\tilde{\mu}-\mu}{\tilde{\sigma}/\sqrt{n}}$$

- Es gilt

$$T = \frac{\tilde{\mu}-\mu}{\sigma/\sqrt{n}} \Big/ \sqrt{\frac{\tilde{\sigma}^2/n}{\sigma^2/n}}$$
$$= \frac{\tilde{\mu}-\mu}{\sigma/\sqrt{n}} \Big/ \sqrt{\frac{\tilde{\sigma}^2/(n-1)}{\sigma^2/(n-1)}}$$

- $\tilde{\mu}$ enthält Summe id. unabh. $N(\mu,\sigma^2)$-verteilter Y_i
  Summe wieder normalverteilt, gemäß $N(\mu \cdot n, \sigma^2 \cdot n)$
  und auch $\tilde{\mu}$ normalverteilt, gemäß $N(\mu,\sigma^2/n)$
  und somit **Zähler** von T **N(0,1)-verteilt**

- zum Nenner von T:

$$(n-1)\,\frac{\tilde{\sigma}^2}{\sigma^2} = \frac{n-1}{n-1}\sum_1^n \frac{(Y_i-\tilde{\mu})^2}{\sigma^2} = \sum_1^n \left(\frac{Y_i-\tilde{\mu}}{\sigma}\right)^2$$

  Summe mit μ statt $\tilde{\mu}$ ist - vgl. (6.1.8) - χ^2_n-verteilt;
  für Summe (mit $\tilde{\mu}$) läßt sich zeigen, daß sie χ^2_{n-1}-verteilt
  damit **Nenner** von T $\boldsymbol{\chi^2_{n-1}}$**-verteilt**

- Insgesamt - vgl. (6.1.8) -
  T ist $\boldsymbol{t_{n-1}}$-verteilt

Als Konfidenzintervall für (ii) damit

(6.1.9a) $P\left[\,\left|\tilde{\mu}-\mu\right| \geq \varepsilon_\alpha\, \tilde{\sigma}/\sqrt{n}\,\right] = \alpha$

(6.1.9b) $\tilde{\mu} \pm \varepsilon_\alpha\, \tilde{\sigma}/\sqrt{n}$

mit ε_α (hier) aus t_{n-1}-Tafel (zweiseitiger Test,
t-Verteilung ist symmetrisch)

Im Beispiel: $\varepsilon_{0.1}=1.833$
und KI: $\tilde{\mu} \pm 0.580 \cdot \tilde{\sigma}$ **immer** etwas breiter als (6.1.7)

Facit:
- (6.1.7) approximatives (für große n gültiges) KI für beliebige Y-Verteilung
- (6.1.9) exaktes KI für normalverteiltes Y (i.allg. nicht gegeben)
- o.B.: (6.1.9) konvergiert mit n gegen (6.1.7)
- da (6.1.9) "konservativer" (breiter) als (6.1.7), "sicherheitshalber" (6.1.9) verwenden; ist deutlich "besser" (schmaler) als Tschebyscheff-(6.1.5), dennoch hinreichend sicher

Soweit: **Minimalprogramm** !

Erwünscht sicher zusätzlich:
weitere Momente der Y-Verteilung,
Quantile der Y-Verteilung,
Y-Verteilung selbst … vgl. Abschn. 5,
aber: Stichproben aus Aufwandsgründen (leider) relativ klein

6.2 Zeitreihen-Analyse

Bisher diskutiert
Charakteristika der Verteilung von: $Y(t|z_0)$
(Beobacht'g zum Zeitpunkt t, falls im Zustand z_0 gestartet)

Bei Zeitreihenanalyse Schwerpunkt
Charakteristika der Verteilung von Y für Fälle, wo

(6.2.1) $Y(t|z_0) \rightarrow_{t \rightarrow \infty} Y$

gesichert (oder zu vermuten) ist
(Bsp.: Zustand Bankschalter, Verweilzeit Kunde)

Aber: Interessenlage kann durchaus anders sein
(Verhalten bei "Systemanlauf",
bei sprunghaften Betriebsänderungen)
Und: Konvergenz zu stationärer / eingeschwungener
Phase nicht notwendig gegeben
("überlastetes" System)

Stationäre / eingeschwungene Phase war "intuitiv":

(6.2.2) wo Einfluß Anfangszustand "verschwunden",
wo absoluter Zeitpunkt "keine Rolle spielt"

Dazu jetzt präziser:

(6.2.3a) Definitionen

- Stochastischer Prozeß Y(t) heißt
  stationär k-ter Ordnung
  wenn gemeinsame Verteilung von
  $Y(t_1),Y(t_2),\ldots,Y(t_k)$
  und gemeinsame Verteilung von
  $Y(t_1+h),Y(t_2+h),\ldots,Y(t_k+h)$
  identisch
  für jede Folge $t_1<t_2< \ldots <t_k$
  und jedes $h>0$
- Stochastischer Prozeß Y(t) heißt
  stationär im strengen Sinne
  wenn Stationarität k-ter Ordnung gilt
  für jedes $k\geq 1$
- Stochastischer Prozeß Y(t) heißt
  schwach stationär
  (im weiten Sinne stationär, Kovarianz-stationär)
  wenn
  $E[Y(t_i)] \equiv E[Y]$ für alle t_i
  und
  $COV[Y(t_i),Y(t_j)] := E[\,(Y(t_i)-E[Y])\,(Y(t_j)-E[Y])\,]$
  nur von Abstand $|t_i-t_j|$ abhängig
  (und unabhängig von "Lage" t_i,t_j)

(6.2.3b) Es gilt (leicht nachrechenbar):

Stationarität 2. Ordnung impliziert schwache Stationarität

Frage: Wie könnten, bzgl. dieser Definitionen,
"Beobachtungen Simulator"
ausfallen ??

Plausibilitätsbetrachtungen:

- Wir "wissen" i.allg., ob Abhängigkeiten
  zwischen Beobachtungsgrößen $Y(t_i)$ und $Y(t_j)$
  von "Lage" Intervall (t_i,t_j), d.h. von absoluter Zeit t,
  beeinflußt:

  Falls ja, wäre Einfluß von uns selbst
  "in Dynamik eingebaut"
  (z.B. Kundenankunftsprozeß von absolutem t abhängig)

  Wenn nicht eingebaut, ist
  $COV[Y(t_i),Y(t_j)] = f(|t_i-t_j|)$
  sinnvolle Annahme (zumindest, wenn Y Zustandsgröße)

- Wie haben i.allg. "Gefühl" dafür,
  ob etwaige Abhängigkeiten zwischen $Y(t_i)$ und $Y(t_j)$
  mit wachsendem $|t_i-t_j|$ abnehmen sollten
  z.B. Wartezeit 500. Bankkunde
  von der des 10. Kunden abhängig?
  Gefühl: Kaum! (bei fehlender Überlastung)

  Falls ja, sollte auch
  in den ZV $Y(t|z_0)$
  Abhängigkeit von z_0 mit großem t "aussterben"

  d.h. könnte Hinweis vorliegen auf
  asymptotisch stationären Prozeß

  : "nach hinreichend langer Zeit wird Prozeß
  eine stationäre Phase approximativ
  (und mit wachsender Zeit immer besser)
  erreichen"

"könnte" Hinweis vorliegen,
"muß" aber nicht:
andere Möglichkeiten für wachsendes t
- "Trends",
- "zyklisches (periodisches, saisonales) Verhalten"

Aufdeckung (und **stochastische Beschreibung**)
- linearer und nichtlinearer **Trends**
- **Zyklen**, u.U. in überlagerter Form

sind wesentlicher Inhalt **statistischer Zeitreihenanalyse**
Hier: Nicht weiter verfolgt

Hier Annahme:

ausgewählter Beobachtungs-"Prozeß" Y(t)
ist asymptotisch Kovarianz-stationär
mit genau einer (z_0-unabhängigen) **Grenzverteilung** PY

("nähert sich mit wachsendem t eindeutig
einem stationären / eingeschwungenen Verhalten,
einer stationären / eingeschwungenen Phase")

Sicherlich

haben wir laufend Gegenbeispiele "in der Hand":
- Kundenanzahl in überlastetem Bankschalter
- Gesamtzahl der Kundenankünfte bis zu Zeitpunkt

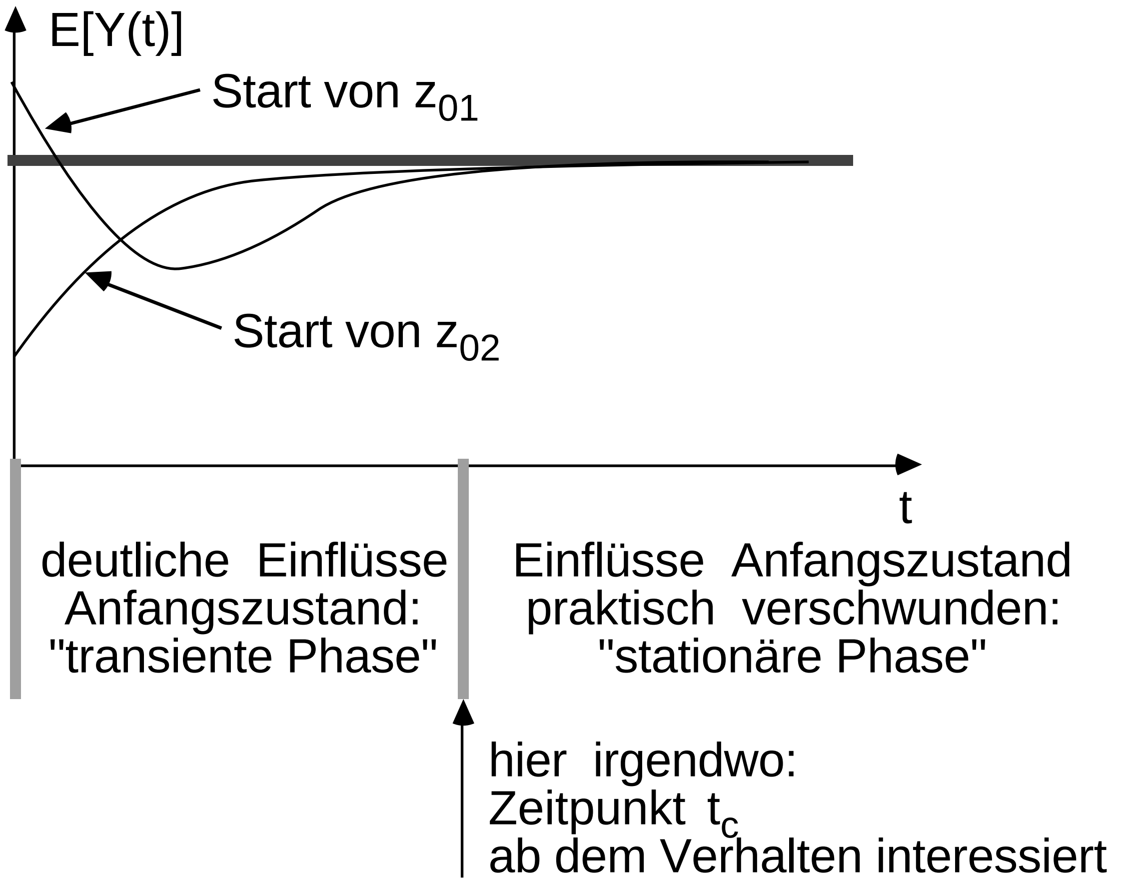


Abbildung 6.2.4: Skizze transiente / stationäre Phase für Erwartungswert Beob'größe, zwei Startzustände

bekanntes Vorgehen:
  Ensemble-Analyse
wäre einsetzbar zur Analyse stationäre Phase
  für irgendein $t^* > t_c$
  wo $Y(t^*)$ "in Verteilung annähernd gleich" Y

Aber: **wo liegt t_c** ??
- Sicherheitsargument: Möglichst großes t_c wählen
- Aufwandsargument: Für jede Replikation transiente Phase "durchzusimulieren": Möglichst kleines t_c wählen

Zusätzliche Schwierigkeit:

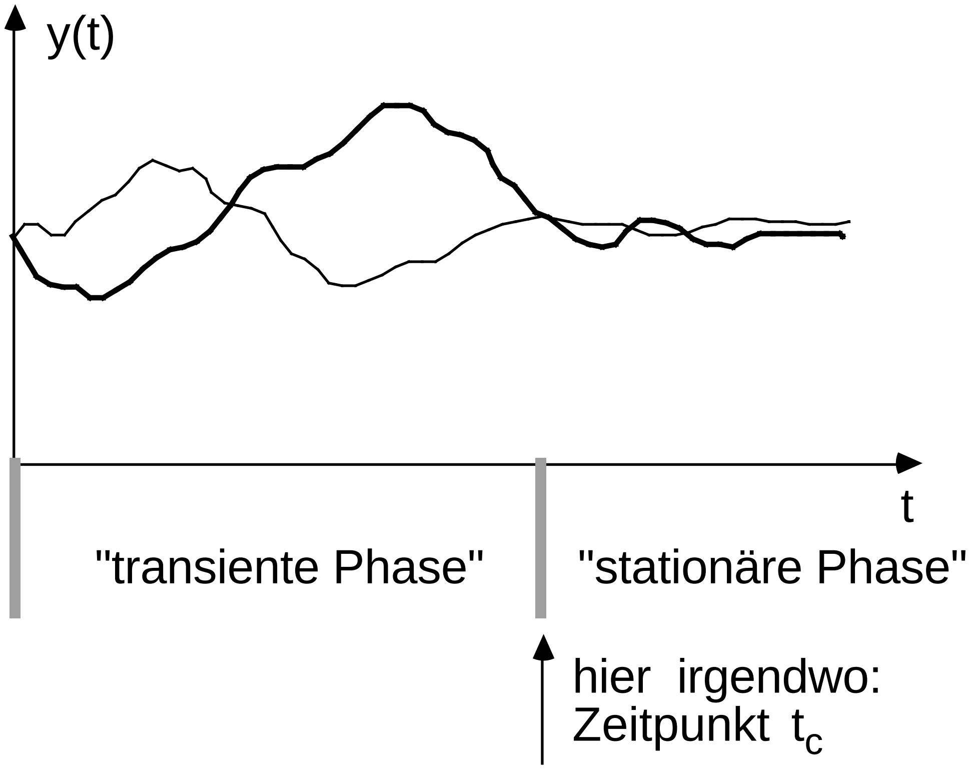


Abbildung 6.2.5: Replikationen bei festem z_0

Bei anderem z_0 wird t_c u.U. früher (oder später) erreicht !

Probleme also:

- wie z_0 wählen ?
- wie t_c wählen ?

zur (Fernziel:) Beobachtung stationärer Phase

Dazu einige (teils **sehr** heuristische) Hinweise
aus einschlägiger Literatur

(6.2.6) Tips zur Wahl von z_0

- start model "empty and idle"
  - nur für gewisse Modelle definiert
  - Grund unklar
  - aber: oft einfach zu bewerkstelligen,
    naheliegender "Standardfall"

- start model "at steady state mode"
  - soll heißen:
    in Zustand, der in stationärer Phase mit
    höchster Wahrscheinlichkeit / Häufigkeit
    eingenommen wird
  - Motivation:
    in "sehr typischem" Zustand,
    der in stationärer Phase häufig eingenommen
    (unter diesem Gesichtspunkt
    ist "empty and idle" ggf. sehr ungünstig !)
  - aber:
    woher (diesen Zustand) nehmen ?
    und, wenn bekannt (d.h. stat. Zust'Verteil'g bekannt)
    wozu dann noch simulieren ??

- start model "at steady state mean"
  - soll heißen:
    in mittlerem Zustand (im Sinne Erwartungswert)
    der stationären Phase
  - Motivation: wie oben
  - aber: woher bloß nehmen ? und … s.oben
    und: was ist "mittlerer" Zustand bei kompl. System?
    Mehrdimensionaler Zustandsraum:
    alle Komponenten bei E[Marginalzustand]?
    Ganzzahliger Z'raum:
    "physikalisch" sinnvoll ?
    (z.B. Schalter mit 3.71 Kunden ??)

- verbleibt als Tip:
  - z_0 so, daß einfach implementierbar ("Standardfall")
  - z_0 "rekurrent", d.h.
    kein Zustand, der ("offensichtlich") in station. Phase gar nicht auftritt (wäre sog. "transienter Zustand")
    exakter: rekurrent: P[Rückkehr in Zustand] = 1
    positiv rekurrent: E[Zeit bis Rückkehr] < ∞

- NB: start model "at steady state"
  wäre korrekt (und sogar t_c = "Beginn Simulation")
  in folgendem Sinn:

  Sei {P[z]; z Z} z: Zustand, Z: Zustandsraum
  stationäre Zustandsverteilung Modell;
  dann wähle (bei Replikationen)
  z_0 entsprechend {P[z]}

  nur erneut: woher nehmen ??

(6.2.7) Tips zur Wahl von t_c

- t_c so, daß $y(t_c)$ weder Minimum noch Maximum
  der Beobachtungswerte $\{y(t_i); t_i<t_c\}$
  (also der "vorherigen")

- t_c so, daß $y(t_c)$ weder Minimum noch Maximum
  der Beobachtungswerte $\{y(t_i); t_i>t_c\}$
  (also der "folgenden, zu berücksichtigenden")

Empirische Erkenntnis:
Bei beiden Vorschlägen t_c (z.T. gravierend) zu klein

- t_c so, daß "batch means"
  der k letzten "batches"   (später: der "berücksichtigten")
  innerhalb Intervall des Umfangs ε liegen

  k und ε sind dabei Parameter

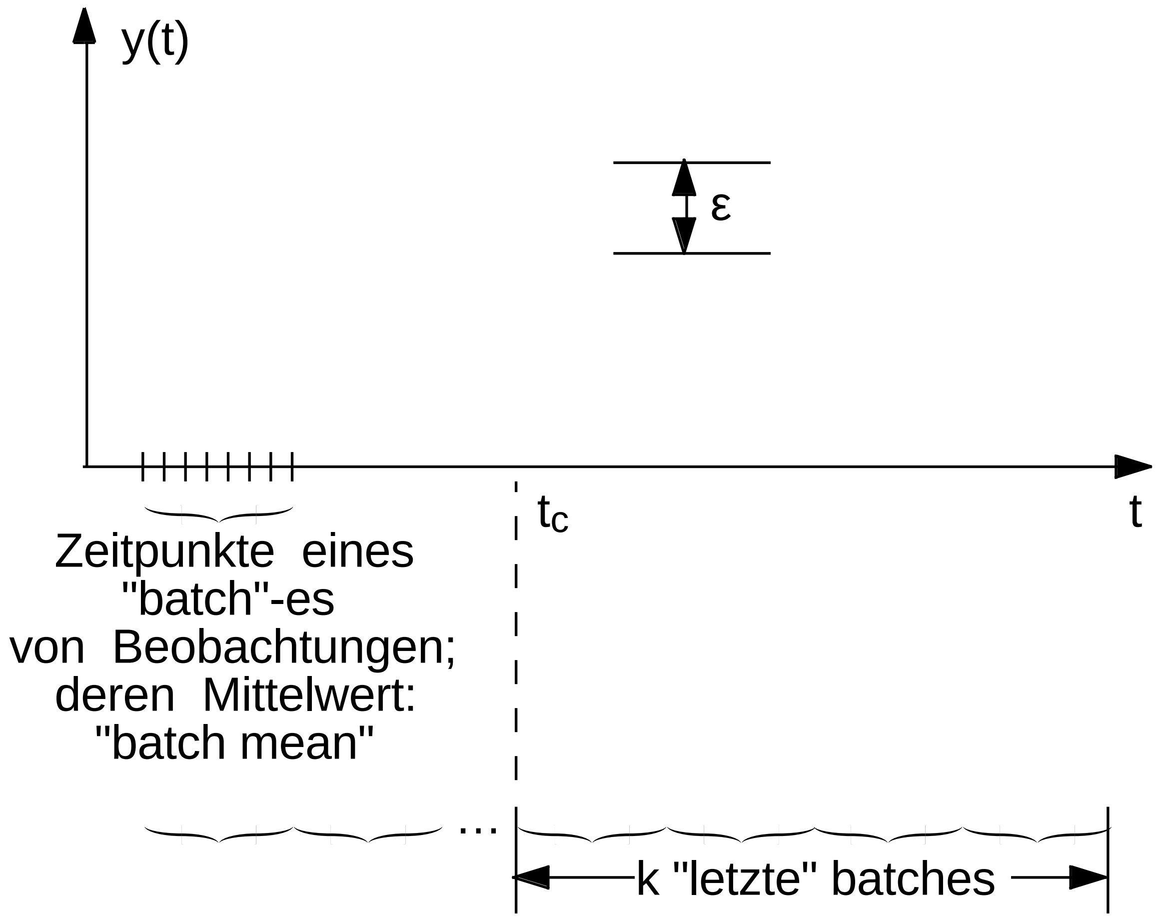


Abbildung 6.2.8: Illustration "batch means" Idee

Bis auf Wahl von k und ε (woher nehmen?)
kommt dies Prüfung $E[Y_t] \equiv E[Y]$? nahe.
(obwohl: Mit endlicher Wahrscheinlichkeit wird auch in stationärer Phase ein batch mean außerhalb jedes echten Teilintervalls des Wertebereichs von y liegen)

Wird bei "Zeitreihenanalyse" sinnvoller (Information !)

- t_c mittels **Pilotläufen** ermitteln

  Aufzeichnen Beobachtungsfolgen über der Zeit,
  für mehrere Läufe (Replikationen)
  "Visuell" t_c wählen

  Vorgang auch formalisierbar:

  z.B.: schätze (quer zu Replikationen)
  Folge der Mittel
  $\tilde{\mu}_t$ (für $E[Y_t]$) $t=t_1,t_2,\ldots$
  wähle t_c so, daß sich Schätzwerte μ_t^*
  von da ab "nur noch unwesentlich" ändern

  oder: schätze (längs zu Replikationen)
  jeweils Mittel der Beobachtungen $\tilde{\mu}_{1\ldots t}$ $(t=t_1,t_2,\ldots)$
  schätze (quer zu Replikationen)
  Folge der Varianzen der Mittel
  $V[\tilde{\mu}_{1\ldots t}]$
  wähle t_c so, daß Schätzwerte für $V[\tilde{\mu}_{1\ldots t}]$
  von da ab "wie $1/\#t_i$" fallen Begründung:(6.1.3)

- Insgesamt:

  Simple Tips unbefriedigend !
  Pilotläufe aufwendig !

Zur Ermittlung / Beurteilung "stationäre Phase"
bisher verfolgte Linie:

aus Replikationen
  Stichprobe $(y_1,y_2,\ldots,y_n)$
  für Y(t) mit $t>t_c$
so daß Stichprobe Y-Vert'g in station. Phase repräsentiert

Jetzt (wie laufend vorbereitet) weiter:

- ab t_c Verteilungen $Y(t_i)$ voraussetzungsgemäß identisch

- Idee naheliegend,
  Stichprobe $(y_1,y_2,\ldots,y_n)$
  aus über t_c fortgesetzter Simulation gewinnen,
  also in Form der Beobachtungen
  $(y(t_1),y(t_2),\ldots,y(t_n))$ $\quad t_i>t_c \quad i=1,2,\ldots,n$
  $\quad t_{i+1}>t_i \quad i=1,2,\ldots,n-1$
  denn: die ZV $Y(t_1),Y(t_2),\ldots,Y(t_n)$ sind identisch verteilt,
  die sequentiell gewonnene Stichprobe ist ebenfalls
  für "Y-Vert'g in station. Phase" repräsentativ

- Vorgehen bezeichnet als:
  dynamische Analyse,
  Zeitreihen-Analyse,
  (stationäre Analyse)

- zu erwartender Vorteil:
  transiente Phase nicht "immer wieder" zu durchlaufen;
  entspräche auch Tip "in stationärem Zustand" starten
  (der ja nicht direkt durchführbar)
- aber Vorsicht bei Auswertung sequentieller Stichprobe:
  Verfahren aus Abschn. 6.1 setzten (wiederholt)
  unabhängige Stichprobe voraus;
  hier gerechtfertigt ??
  (z.B. Folge Kunden-Verweilzeiten !!)

Überlegungen aus Abschn. 4 / 6.1 nochmals aufgerollt:

- aus Stichprobe $(y_1, y_2, \ldots, y_n)$
  wurde ermittelt erwartungstreuer Schätzer

  (6.2.9) $\tilde{\mu} = \frac{1}{n} \sum_{1}^{n} Y_i$

  für $\mu = E[Y]$

- Erwartungstreue dieses Schätzers auch bei sequentieller (potentiell nicht unabhängiger) Stichprobe gegeben (Ableitung von (4.1.5) nicht berührt)

- Varianz $V[\tilde{\mu}]$ des Schätzers
  (wesentlich für Gewinnung Konfidenzintervalle)
  hing nach (4.1.7) mit
  Varianz $V[Y]=\sigma^2$ der Beobachtungsgröße
  zusammen gemäß

  $$V[\tilde{\mu}] = \sigma^2/n$$

  Zugehörige Ableitung beruhte auf Unabhängigkeit:
  neu zu betrachten!

- Es wird sich zeigen,
  daß fälschliche Verwendung von (4.1.7)
  zu gefährlichen Fehlern führt

- $V[\tilde{\mu}] = E[(\tilde{\mu} - \mu)^2]$
  $= E[(\Sigma Y_i/n - \mu)^2]$
  $= E[\{1/n \ \Sigma(Y_i-\mu)\}^2]$

$$= \frac{1}{n^2} \ E\left[\left\{\sum_{i=1}^{n}(Y_i-\mu)\right\}^2\right]$$

$$= \frac{1}{n^2} \ \left\{\sum_{i=1}^{n} E\left[(Y_i-\mu)^2\right] + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E\left[(Y_i-\mu)(Y_j-\mu)\right]\right\}$$

dabei ΣΣ… bereits bekannt als Kovarianzen,
hier (wegen Gewinnung aus **einer** Prozeßrealisierung)
Auto-Kovarianzen;
sollten bei potentieller Abhängigkeit Stichprobe ≠ 0 sein,
dürften in realistischen Fällen > 0 sein ("Warten"!);

- damit
  $V[\tilde{\mu}] = \sigma^2/n$ **+ ">0"**
  und bisher verwendeter Wert **zu klein**,
  damit alle (zu $V[\tilde{\mu}]$ proportionalen) Konfidenzintervalle
  ebenfalls **zu klein**
  Vortäuschung falscher Sicherheit!

- Verfolgung Term ΣΣ…

$$\sum\sum \ldots = \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} COV\left[Y_i,Y_j\right]$$

nach Definition ist

$$COV[Y_i,Y_j] = COV[Y_j,Y_i]$$

damit

$$\sum\sum \ldots = 2 \sum_{i=1}^{n} \sum_{j>i}^{n} COV\left[Y_i, Y_j\right]$$

$$= 2 \sum_{i=1}^{n} \sum_{j>i}^{n} \sqrt{V[Y_i]\,V[Y_j]} \quad \frac{COV\left[Y_i, Y_j\right]}{\sqrt{V[Y_i]\,V[Y_j]}}$$

wo Term
bezeichnet als **Korrelation**
hier: **Auto-Korrelation**

vorausgesetzt war: **schwache Stationarität**
also $V[Y_i] \equiv V[Y] = \sigma^2$
und $COV[Y_i, Y_j] \equiv$ COV_F(Y,|j-i|)
eine "Kovarianzfunktion", die für
sinnvolle Beobachtungspunkte t_i, t_j
z.B. diskrete (etwa ganzzahlige) Zeit
mit konstanten Abständen $t_{i+1}-t_i$
nur von Differenz j-i abhängt
(und nicht von "Lage" t_i, t_j)

Man nennt
$\rho(s) = \rho(|j-i|) :=$ COV_F(Y,|j-i|) / σ^2
Autokorrelationsfunktion

damit

$$\frac{\sum\sum \ldots}{n^2} = \frac{2\,\sigma^2}{n^2} \sum_{i=1}^{n} \sum_{j>i}^{n} \rho\left(|j-i|\right)$$

$$= \frac{2\,\sigma^2}{n^2} \sum_{s=1}^{n-1} (n-s) \quad \rho(s) = \frac{2\,\sigma^2}{n} \sum_{s=1}^{n-1} \frac{n-s}{n} \quad \rho(s)$$

und insgesamt

$$\textbf{(6.2.10)}\quad V[\tilde{\mu}] = \frac{\sigma^2}{n}\left(1 + 2\sum_{s=1}^{n-1}\frac{n-s}{n}\ \rho(s)\right)$$

Um $V[\tilde{\mu}]$ zu ermitteln (Zweck: KI für $\tilde{\mu}$ gewinnen),
müßten demnach
zusätzlich zu σ^2
auch alle $\rho(s)$, $s=1,2,\ldots,n-1$
ermittelt (aus Stichprobe geschätzt) werden

Schätzer $\tilde{\rho}(s)$ sind zwar bekannt,
aber relativ aufwendig zu berechnen
und vor allem wechselseitig abhängig

Weitere Methoden, die $\rho(s)$ zu ermitteln
- Spektralanalyse
- Schätzung autoregressiver Modelle

nicht "blind" für alle Fälle anwendbar;
hier nicht weiter verfolgt

Festzuhalten:

- Bei unabhängigen Stichproben (korrekterweise) verwendetes σ^2/n unterschätzt $V[\tilde{\mu}]$ bei positiv autokorrelierten Stichproben.

- Unterschätzung gefährlich, da
  zu kleine KI's resultieren,
  zu hohe Sicherheit vorgespiegelt

  (z.B. kleinere Stichproben als notwendig)

Ausweg unter zusätzlicher (plausibler!) **Annahme**:

ρ(s) bzw. |ρ(s)| ist monoton mit s fallende Funktion

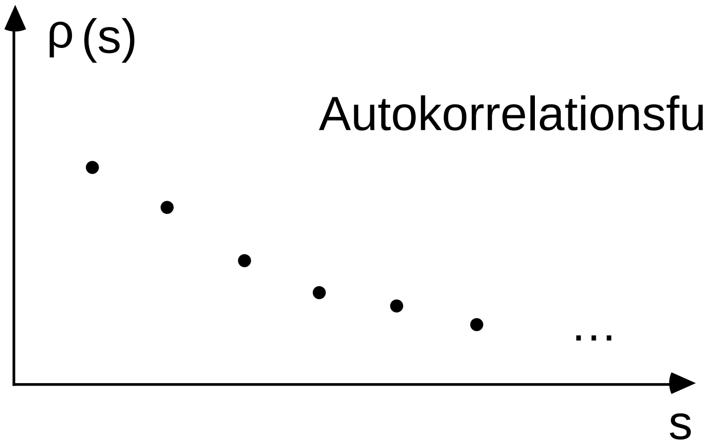


Abbildung 6.2.11: Prinzipskizze
fallende Autokorrelationsfunktion

Plausibilität: Es sollte fallender Einfluß mit wachsendem Abstand der Beobachtungen vorliegen; war für Erreichen stationärer Phase ohnehin vorausgesetzt

Ausnutzung dieser Annahme bei sog.
batch means Methode

(6.2.12) Annahme:

$$\rho(s) \approx 0 \quad \text{für } s > s_c$$

(6.2.13) batch-means-Methode

- Beobachtungen $(y_1,y_2,\ldots,y_N)$ aus einem Simulationslauf in k+1 "Blöcke" (batches) aufgeteilt;
  Umfang erster Block: n_0
  Umfang weitere Blöcke konstant: n
  so daß $N = n_0 + k \cdot n$
  zur weiteren Benennung $n_1 := k \cdot n$

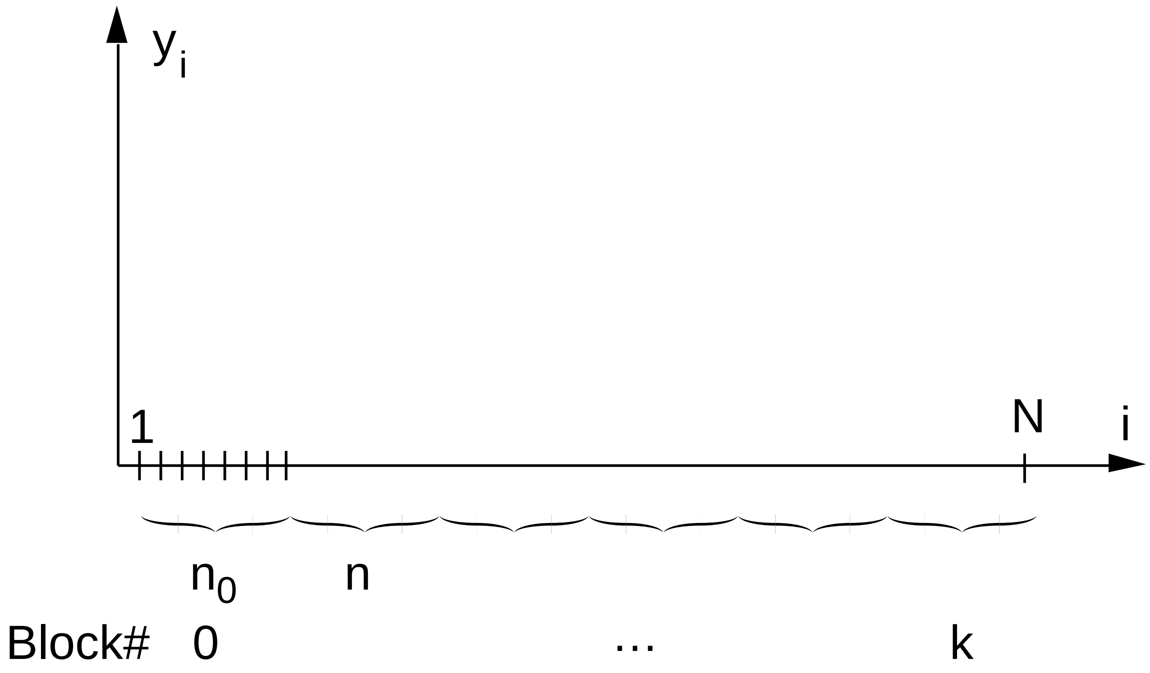


Abbildung 6.2.14: Veranschaulichung batch-means-Methode

- Erste Beobachtungen: transiente Phase, zunächst unberücksichtigt
  Umbenennungen zur Vereinfachung:
  y_{ij} ist j-te Beobachtung in i-tem Block,
  $i = 0,1,\ldots,k$
  für i>0: $j = 1,2,\ldots,n$
  für i=0: $j = 1,2,\ldots,n_0$

- Bildung der sog. "batch means"

$$\bar{y}_i := \frac{1}{n} \cdot \sum_{j=1}^{n} y_{ij} \qquad i = 1,2,\dots,k$$

und des Gesamtmittels ("grand mean")

$$\mu^* := \frac{1}{k} \cdot \sum_{k=1}^{k} \bar{y}_i$$

$$\left(= \frac{1}{k\,n} \sum_{i=1}^{k} \sum_{j=1}^{n} y_{ij} \right.$$

$= \mu^*$ bei Nichtbeachtung transiente Phase
"wie gehabt", vg. z.B. (6.2.9);
daher nach wie vor erwartungstreu
d.h. $E[\tilde{\mu}] = \mu = E[Y]$
für Mittelwert stationäre Phase)

- sei n so gewählt, daß entsprechend Annahme (6.2.12) Korrelation zwischen Blöcken (annähernd) verschwindet, also (zumindest)

$$n > s_c$$

und batch-means-Schätzer $\bar{Y}_i$ (approximativ) unkorreliert

- damit wird Beziehung (6.1.3) bzw. (4.1.7)

$$V[\tilde{\mu}] = \frac{V[\bar{Y}]}{k}$$

wieder verwendbar, da (Begründung)
die $\bar{Y}_i$ unkorrelierte, identisch verteilte ZV
sind

- erwartungstreue Schätzer für $V[\bar{Y}]$
"wie gehabt": (6.1.2)

$$\tilde{\sigma}_B^2 := \frac{1}{k-1} \sum_{i=1}^{k} \left(\bar{Y}_i - \tilde{\mu}\right)^2$$

und die (angestrebten) Konfidenzintervalle:
- entsprechend (6.1.7) "approximativ für große k"

$$\tilde{\mu} \pm \varepsilon_\alpha \, \tilde{\sigma}_B / \sqrt{k}$$

mit ε_α aus N(0,1)-Tafel
galt wegen zentralem Grenzwertsatz (6.1.6);
Zusatzbemerkung war:
Approximation bei sehr "nicht-normalen"
Beobachtungen erst für große k gut.
hier: batch means $\bar{Y}_i$ sind selbst Summen über ZV,
daher "normaler" als ursprüngliche Y_i;
wieder Grenzwertsatz
(andere Version: abhängige ZV)
- somit auch (6.1.8) "normalverteilte Beobachtungen"
(besser) rechtfertigbar

$$\tilde{\mu} \pm \varepsilon_\alpha \, \tilde{\sigma}_B / \sqrt{k}$$

mit ε_α aus t_{k-1}-Tafel
- da letzteres KI pessimistischer (breiter)
wieder Tip (Sicherheit): letzteres verwenden

soweit, sogut (alles sehr freundlich)

nur: Länge transiente Phase ? (t_c –> n_0) ?
Aussterben Autokorrelation? (s_c –> n) ?

zu n_0:

- n so definiert, daß für größere Abstände als n Autokorrelationen vernachlässigbar;
  falls rekurrenter (also nicht: transienter) Zustand als Anfangszustand gewählt,
  sollte auch Einfluß Anfangszustand ab n-ter Beobachtung verschwunden sein;
  rechtfertigbar demnach:

  $$n_0 \approx n$$

- des weiteren, da

  $$\tilde{\mu} := \frac{1}{k} \sum_{i=1}^{k} Y_i$$

  von (Einschluß 0-ter, transienter Block)

  $$\tilde{\mu} := \frac{1}{k+1} \sum_{i=0}^{k} Y_i$$

  zwar sicher verschieden, aber bei großem k (viele Blöcke) u.U. vernachlässigbar "stark", auch Mitberücksichtigung trans. Phase (0-ter Block) vertretbar!

zu n:

- Genaueres über (problematische) Schätzung der Autokorrelationsfunktion $\tilde{\rho}(s)$
  s. Literatur (z.B. Fish73/78, LaKe82)

- Überlegungen und heuristische Tips:

  (i) da Autokorrelation KI-Schätzung gefährlich verfälscht, vorsichtiges Vorgehen:
  n eher groß

  (ii) mit (i) wird bei festem Stichprobenumfang n der batch-means-Stichprobenumfang klein und daher die KI's (~1/k) groß; somit
  k eher groß, n eher klein

  (iii) allerdings bei größerem n (daher kleinerem k) auch
  $\tilde{\sigma}_B^2 \sim V[\bar{Y}]$
  kleiner; und (ii) nicht ganz so stichhaltig

  (iv) Tip:
  k > 10 nicht verletzen
  bei n eher "auf Nummer sicher" gehen
  außer bei gutem Grund: "Tests"

- n so, daß Unkorreliertheit der batch means gesichert; Unsicherheiten durch geeignete Tests mindern; z.B. durch folgenden Test

Einfach zu handhabender Hinweis
auf "Zufälligkeit" einer Stichprobe
- anzuwenden auf ($\overline{y}_i$; 1=1,2,…,k) -

(6.2.15) One-Sample-Runs-Test
Iterationstest (vgl. Lit.: Sieg56, BüTr78)

- liege vor Folge zweiwertiger Beobachtungen,
  Wertevorrat = {+,-};
  z.B.
  ++---+----++-+
  (Transformationen aus mehrwertigen Beobachtungen:
  über/unter Mittelwert, Median etc.)

- zähle Anzahl der "runs" (der Iterationen) r,
  d.h. Anzahl von Folgen identischer Werte;
  im Beispiel
  $\underline{++}$ $\underline{---}$ $\underline{+}$ $\underline{----}$ $\underline{++}$ $\underline{-}$ $\underline{+}$

  1 2 3 4 5 6 7 Zählung: r=7

- es sind
  bei positiver Korrelation: wenige runs
  bei negativer Korrelation: viele runs
  zu erwarten,
  wogegen bei zufälliger Stichprobe Wert r
  "zwischen den Extremen" zu erwarten ist

- Verteilung der ZV R (unter Zufälligkeits-Hypothese)
  ist bekannt,
  kritische (hohe, niedrige) Werte sind vertafelt

- zur Anwendung (im Beispiel)

  n_1 := Anzahl "+" = 6
  n_2 := Anzahl "-" = 8
  N := n_1+n_2 = 14
  r := Anzahl "runs" = 7

  Hypothese der Zufälligkeit zu verwerfen,
  falls (zweiseitiger Test,
  Tafelwerte für 5%-Typ 1-Fehlerwahrscheinlichkeit)

  $r \leq \text{Tafel}_{\text{untere kritische Werte}}$ (= 3)
  $r \geq \text{Tafel}_{\text{obere kritische Werte}}$ (= 12)

  im Beispiel also: Hypothese nicht verwerfen

- für große Stichproben (n_1 oder n_2 > 20)
  ist R in guter Näherung normalverteilt mit

$$\mu_r = \frac{2\, n_1\, n_2}{n_1+n_2} + 1$$

$$\sigma_r^2 = \frac{2\, n_1\, n_2 \, \left(2\, n_1\, n_2 - n_1 - n_2\right)}{\left(n_1+n_2\right)^2 \, \left(n_1+n_2-1\right)}$$

  und Transformierte
  $Z := (R-\mu_r)/\sigma_r$
  kann mithilfe N(0,1)-Tafel (zweiseitig!) getestet werden

- praktisches Vorgehen:
  entweder alle Beobachtungen y_i aufzeichnen,
  dann mit wachsendem batch-Umfang n bestimmen
  oder Pilotläufe bis n ermittelt,
  dann ernsthafter (langer) Simulationslauf

Ausklang:

- batch means arbeitete mit approximativ unkorrelierten batches fester Länge;
  formal schöner wäre, mit gesichert unabhängigen batches potentiell variabler Länge zu arbeiten;

  siehe Literatur: **regenerative Verfahren**
  (z.B. Igle78, Shed87)

- anstatt bei festem Stichprobenumfang n (Simulationsdauer!) von diesem abhängige KI's zu bestimmen,
  wäre sicher erwünschter,
  Stichprobenumfang n in Abhängigkeit der angestrebten KI-Breite zu wählen;

  etwa Wunsch: Breite relatives α-%-KI := $2 \cdot \varepsilon_\alpha / \mu^* < c$

  siehe Literatur: **sequentielle Verfahren**