Anhang 8A: Zusammengesetzte Ereignisse

- Zweistufiges Experiment mit Ereignissen

F_i, $p(F_i) = f_i$ und

G_j, $p(G_j) = g_j$

- Ereignis E_{ij} zusammengesetztes Ereignis, g_j nach Eintreten von F_i

$$p(E_{ij}) = p_{ij} = p(i)\,p(j|i) \qquad (A8.1)$$

$p(i) = f_i$ marginale W

$p(j|i)$ bedingte W

p_{ij}, $p(j|i)$ Wahrscheinl. im Sinne der Def. (S. 8.19)

Beispiel Münze, Würfel

p_{K6} Wahrscheinlichkeit, daß nach "Kopf" die Zahl "6" gewürfelt wird

- Falls G_j, $j = 1 \ldots J$ Ereignisse, die gegenseitig ausschließend und vollständig bzgl. g, muß offensichtlich gelten

$$p(i) = \sum_{j=1}^{J} p_{ij} \qquad (A8.2)$$

Beispiel: $p_{K1} + p_{K2} + p_{K3} + p_{K4} + p_{K5} + p_{K6}$

ist W, daß Kopf gewürfelt wurde

$$\Rightarrow p_{ij} = p_{ij}\left(\frac{\sum^{J} p_{ij}}{\sum p_{ij}}\right) = p_{ij}\left(\frac{p(i)}{\sum p_{ij}}\right) =$$

$$= p(i)\left(\frac{p_{ij}}{\sum p_{ij}}\right) \overset{(A8.1)}{=} p(i)\,p(j|i)$$

ALSO

$$p(j|i) = \frac{p_{ij}}{\sum_{k=1}^{7} p_{ik}} \quad , \qquad (A8.3)$$

falls G_j vollst., gegenseitig ausschließend.

BEISPIEL: F_i, G_j <u>unabhängig</u>

Einerseits : $p(j|i) = p(j) = g_j$ unabhängig, vgl. S 8.15

$$\Rightarrow p_{ij} = p(i)\, p(j|i) = f_i \cdot g_j$$

Andererseits $p_{ij} = p_i \cdot p_j = f_i \cdot g_j$

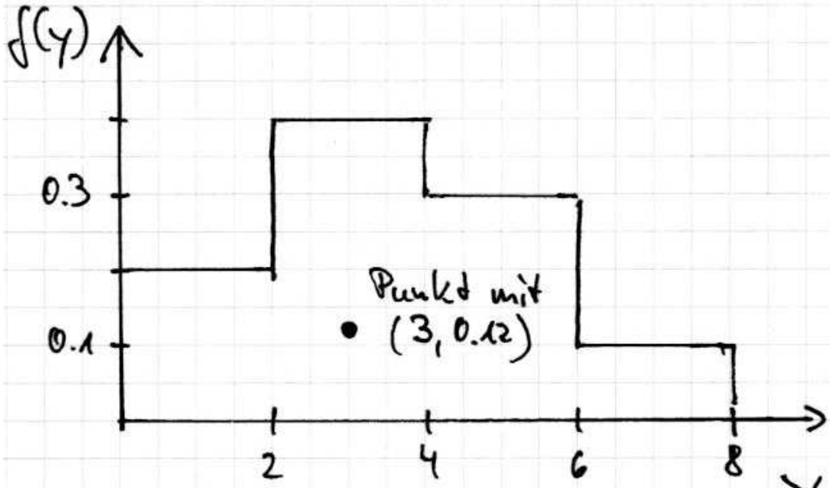
$$\Rightarrow p(j|i) = \frac{p_{ij}}{\sum_{k}^{7} p_{ik}} = \frac{f_i \cdot g_j}{\sum_k f_i g_k} = \frac{f_i \cdot g_j}{f_i \underbrace{\sum_{k} g_k}_{1 \text{ Aufgrund Ann.}}} = g_j$$

Anhang 8B: von Neumannsche Verwerfungsmethode

Frage: Wie $Y(x)$, wenn $F^{-1}(x)$ nur schwer oder gar nicht berechenbar? (Entweder Integral oder Inversion problematisch)

Idee: Erzeuge (auf geschickte Weise) gleichmäßig verteilte Punkte mit Zufallskoordinaten (x_p, y_p) in der Fläche unterhalb der gegebenen pdf $f(y)$

Dann sind die x_p-Werte die gesuchten Zufallsvariablen Y, die gemäß $f(y)$ verteilt sind.



$\hat{=} x_p$

$f(y)\,dy$

$=$ W., daß Y in $[y, y+dy]$

20% aller Punkte liegen bei $0 \le y < 2$
40% " " $2 \le y < 4$
30% " " $4 \le y \; 6$
10% " " $6 \le y < 8$

- Umgekehrt folgt also
Wenn Fläche unterhalb $f(y)$ gleichmäßig mit Punkten belegt wird, müssen....

20%	aller Punkte	im Bereich	$0 \le y < 2$
40%	"	"	$2 \le y < 4$
30%	"	"	$4 \le y < 6$
10%	"	"	$6 \le y < 8$

liegen. Demzufolge gibt es 20% aller Punkte, deren x_p- Koordinaten ($\hat{=}$ y-Werten) im Bereich zwischen

$0 \le y < 2$ liegen, 40% aller Punkte haben x_p-Koordinaten im Bereich $2 \le y < 4$ etc.

<u>ALSO:</u> Falls Fläche unter $f(y)$ gleichmäßig mit Punkten ausgefüllt, sind die dazugehörigen x_p-Werte entsprechend $f(y)$ verteilt!

<u>Frage also:</u> Wie wird diese gleichmäßige Verteilung auf geschickte Weise erzeugt?

- Schritt 1: Wähle Vergleichsfunktion $h(y)$ mit

  $$h(y) \ge f(y) \quad \forall y \qquad (A\ 8.4)$$

  (immer möglich, da $f(y)$ pdf, ≥ 0

Schritt 2: Erzeuge Zufallspunkte, die in der Fläche <u>unterhalb $h(y)$</u> gleichmäßig verteilt sind

Schritt 3: Überprüfe, ob diese Punkte in der Fläche unterhalb $f(y)$ liegen; [aufgrund Schritt 2 und $h(y) \ge f(y)$, sind auch die Punkte unterhalb $f(y)$ gleichmäßig verteilt]. Falls ja, akzeptiere solch einen Punkt;

Falls nein, <u>verwerfe</u> ihn und wähle neuen Punkt

Die akzeptierten Punkte liegen dann gleichmäßig in der Fläche unterhalb $f(y)$, und ihre x_p-Werte sind somit die gesuchten y-Werte, die $f(y)$ gemäß verteilt sind.

<u>Letzte Frage</u>: Wie erzeuge ich also gleichmäßig verteilte Punkte unterhalb $h(y)$??

geschickte Wahl von h entscheidend!

- h soll f möglichst gut approximieren
- $F = \int_{-\infty}^{y} h(y')\,dy'$ muß analytisch berechenbar sein
- F muß analytisch invertierbar sein
- h muß NICHT normiert sein

<u>Zunächst</u>: gleichmäßig verteilte Punkte unterhalb $h(y)$

a) x_p- Koordinate $= y$

d.h. Bestimme y, wenn y $h(y)$-verteilt; mit Inversionsverfahren;

Bestimme zunächst Normierungsfaktor A,

$$A = \int_{y_{min}}^{y_{MAX}} h(y')\,dy'$$

Inversionsverfahren: $x_p \hat{=} y = F^{-1}(A \cdot \rho_1)$,

$$F(y) = \int_{y_{min}}^{y} h(y')\,dy'$$

wenn ξ_1 aus Zufallszahlengenerator; Mit dieser Vorgehensweise haben wird dem Faktum Rechnung getragen, daß h nicht normiert war;

Test: $\xi = 1 \Rightarrow y = F^{-1}(A) = y_{MAX}$ ✓

b) y_P - Koordinate

Erzeuge 2. Zufallszahl y_P zwischen 0 und $h(x_P)$, d.h. $y_P = h(x_P) \cdot \xi_2$, ξ_2 aus Zufallszahlengen.

$\Rightarrow$ Erzeugte Koordinaten x_P, y_P füllen Fläche unterhalb $h(y)$ gleichmäßig aus

<u>Final</u>: Wollen jetzt diejenigen Punkte haben, die Fläche unterhalb $f(y)$ gleichmäßig füllen

$\Rightarrow$ Falls $y_P \leq f(x_P)$, dann akzeptiere Versuch

Falls $y_P > f(x_P)$, dann verwerfe ihn, "würfle" erneut.

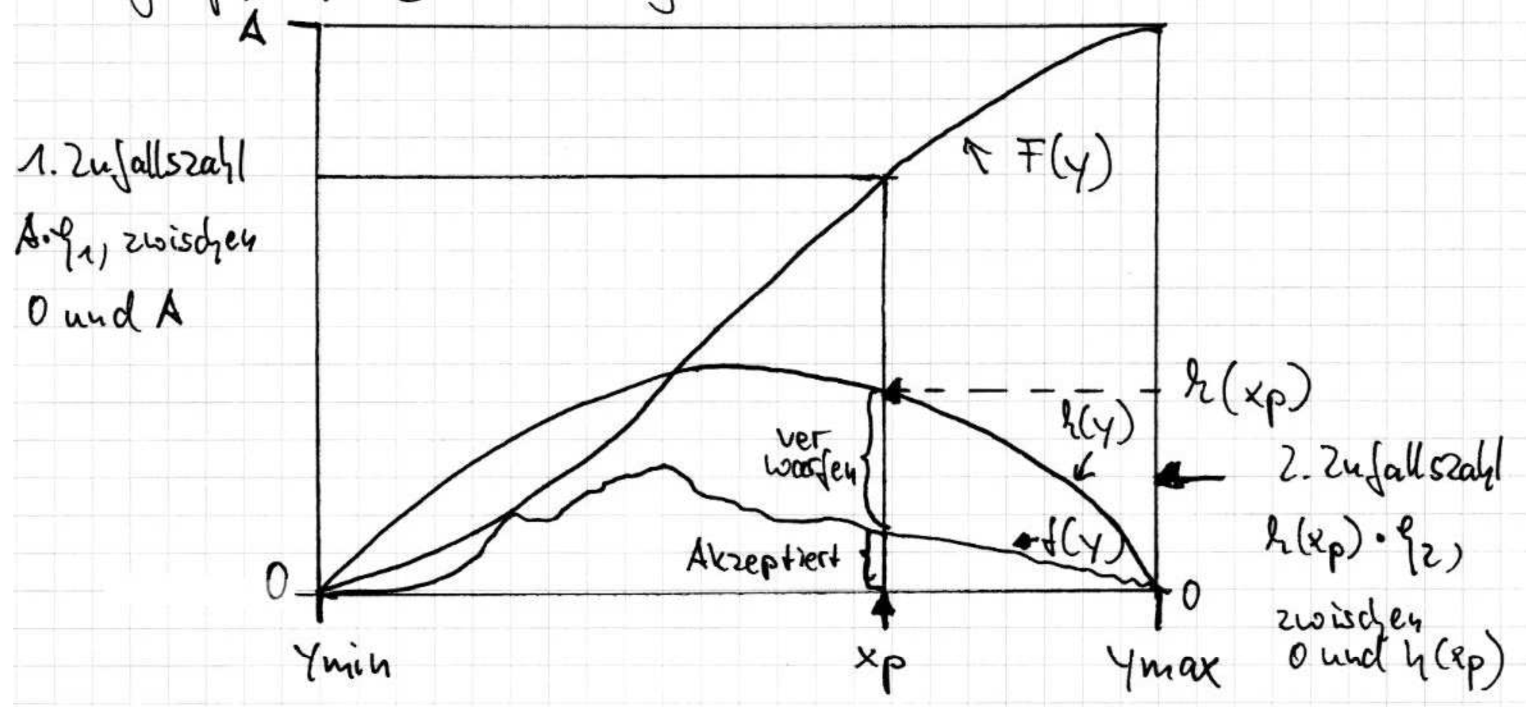
<u>Letztendlich</u> ergeben die x_P-Werte aller akzeptierten Versuche die gesuchte Z.V. y, die gemäß $f(y)$ verteilt ist!

BEACHTE: Je besser die Vgl. Funktion h die tatsächliche pdf $f(y)$ nähert, umso weniger Punkte müssen verworfen werden

Verhältnis akzeptierte / verworfene Punke

$$\hat{=} \frac{\text{Fläche unterhalb } f}{\text{Fläche zwischen } h \text{ und } f}$$

• graphische Darstellung



$$F(y) = \int_{y_{min}}^{y} h(y')\,dy' \quad , \quad F(y_{max}) = A$$

Bsp: (Vgl. Programm S. 8.45, start_angle)

gegeben : pdf $\quad f(\mu) = \frac{6}{3-u}\,(\underbrace{1-u+u\cdot\mu}_{<1})\,\mu$

Schwierig (aber möglich) zu invertieren, da $F(\mu)$ Polynom 3. Ordnung!

Hier: von Neumann

$\mu_{min} = 0, \ \mu_{MAX} = 1$

$\Rightarrow$ Vgl. Funktion $\quad h(\mu) = \frac{6}{3-u}\,\mu > f(\mu)$

$$F(\mu) = \int_0^{\mu} \frac{6}{3-u}\,\mu'\,d\mu' = \frac{3}{3-u}\,\mu^2$$

$$F(\mu_{MAX}) = F(1) = A = \frac{3}{3-u}$$

Algorithmus

(A) 1. Zufallszahl

$$A \cdot \xi_1 = F(\mu) = \frac{3}{3-u}\mu^2 = A \cdot \mu^2$$

$$\Rightarrow \mu = \sqrt{\xi_1} \quad (\hat{=} x_p)$$

2. Zufallszahl

$$h(x_p) = \frac{6}{3-u} \cdot x_p \Rightarrow y_p = h(x_p) \cdot \xi_2$$

$$f(x_p) = \frac{6}{3-u}(1-u+u \cdot x_p)$$

TEST

IF $(y_p \leq f(x_p))$ THEN

$\mu = x_p$ FERTIG

ELSE

GOTO (A) neu würfeln

ENDIF

8.C Varianzreduktion mit "Importance Sampling"

(→ "stratified sampling" : Literatur)

- Hatten bei MC Integration

$$I = \int g\,dV \approx V\left[\langle g\rangle \pm \left(\frac{1}{N}\left(\langle g^2\rangle - \langle g\rangle^2\right)\right)^{\frac{1}{2}}\right]$$

(vgl. 8.34),

d.h. Fehler $\sim \frac{1}{\sqrt{N}}$, aber auch

$\sim \left(\langle g^2\rangle - \langle g\rangle^2\right)$

- An $\frac{1}{\sqrt{N}}$ Gesetz läßt sich nichts ändern, aber 2. Faktor läßt sich ggf. minimieren

IDEE : $I = \int g\,dV = \int \frac{g}{p}\, p\,dV$

"Integriere" nicht mit gleichmäßiger Verteilung, sondern mit Verteilung p (unter der Nebenbedingung $\int p\,dV = 1$), um auf diese Weise Varianz zu minimieren!

$$\Rightarrow I = \left\langle \frac{g}{p}\right\rangle \pm \frac{1}{\sqrt{N}}\left(\left\langle\left(\frac{g}{p}\right)^2\right\rangle - \left\langle\frac{g}{p}\right\rangle^2\right)^{\frac{1}{2}} \qquad (A8.5)$$

wobei $\left\langle \frac{g}{p}\right\rangle$ jetzt das arithmetische Mittel bzgl. der pdf p bedeutet!

A.8.5 ist die verallgemeinerte MC - Integration; (Volumenfaktor taucht nicht mehr - explizit - auf)

BEACHTE: i) p ist pdf, d.h. ≥ 0 immer

ii) $g/p \approx$ const $\Rightarrow$ Varianz klein

Beispiel Integration mit gleichmäßiger Verteilung ist Spezialfall dieses Theorems, da

$$\int p \, dV = 1 \quad \text{mit} \quad p = \text{const} \Rightarrow p = \frac{1}{V} \quad (A\ 8.6)$$

$$I = \int g \, dV \approx \left\langle \frac{g}{1/V} \right\rangle \pm \frac{1}{\sqrt{N}} \left(\left\langle \frac{g}{1/V} \right\rangle^2 - \left\langle \frac{g}{1/V} \right\rangle^2 \right)^{\frac{1}{2}}$$

$$= V\left[\langle g \rangle \pm \frac{1}{\sqrt{N}} \left(\langle g^2 \rangle - \langle g \rangle^2 \right)^{\frac{1}{2}} \right], \text{ q.e.d.}$$

- Wähle nun p derart, daß Varianz minimal und p Verteilung auch realisiert werden kann!

(d.h. z.B. über Inversionsverfahren)

- ALSO $\left\langle \frac{g^2}{p^2} \right\rangle - \left\langle \frac{g}{p} \right\rangle^2 \stackrel{!}{=} \text{Min}$, und

"Klammern" sind MC-Integrale, d.h. eigentlich

$$\int \frac{g^2}{p^2} p \, dV - \left(\int \frac{g}{p} p \, dV \right)^2 = \text{Min}$$

$$= \int \frac{g^2}{p} dV - \left(\int g \, dV \right)^2$$

zu minimieren mit Nebenbedingung $\int p \, dV = 1$

$$\Rightarrow \frac{\partial}{\partial p} \left[\int \frac{g^2}{p} dV - \left(\int g \, dV \right)^2 + \lambda \left(\int p \, dV - 1 \right) \right] = 0, \quad (A\ 8.7)$$

mit Lagrange multiplikator λ

$$-\frac{g^2}{p^2} + \lambda = 0 \quad , \text{d.h.}$$

$$p_{opt} = \frac{|g|}{\sqrt{\lambda}} \; ; \qquad \text{Nebenb.} \int p\,dV \Rightarrow \int \frac{|g|}{\sqrt{\lambda}}\,dV = 1,$$

$$\text{d.h.} \; \sqrt{\lambda} = \int |g|\,dV$$

$$p_{opt} = \frac{|g|}{\int |g|\,dV} \qquad (A8.8)$$

- Varianz (eigentlich MC-Schätzwert der Varianz) war

$$\frac{1}{N}\left(\left\langle \frac{g^2}{p^2}\right\rangle - \left\langle \frac{g}{p}\right\rangle^2\right)$$

Sei nun $p = p_{opt}$.

a) g hat im Integrationsgebiet ein Vorzeichen,

$$\text{d.h.} \quad p_{opt} = \frac{g}{\int g\,dV}$$

$$\Rightarrow Var = \frac{1}{N}\left(\left\langle \left(\int g\,dV\right)^2\right\rangle - \left\langle \int g\,dV\right\rangle^2\right)$$

$$= \frac{1}{N}\left(\frac{N\cdot\left(\int g\,dV\right)^2}{N} - \left(\frac{N\cdot\int g\,dV}{N}\right)^2\right) = 0 \,!$$

Allerdings: zur Berechnung von p_{opt} muß das unbekannte Integral $\int g\,dV$ (das wir ja eigentlich bestimmen wollen) gelöst werden.
.... da beißt sich die Katze in den Schwanz!

b) g wechselt im Integrationsgebiet das Vorzeichen
⇒ (nach kurzer Rechnung, analog zu Fall a)

$$Var_{min} = \frac{1}{N}\left[\left(\int |g|\,dV\right)^2 - \left(\int g\,dV\right)^2\right] \neq 0 \quad (A8.9)$$

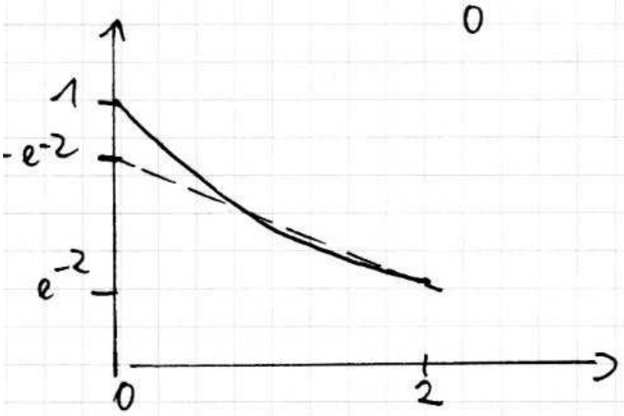
c) Addiere $g \to g + \text{const}$, so daß g überall <u>ein</u> Vorzeichen hat

$$\Rightarrow \underline{I} = I' - \text{const} \cdot V,$$

aber Integrand $\frac{g + \text{const}}{p}$ erschwert die Bestimmung von p

d) <u>in PRAXIS</u> wähle p so, daß g/p möglichst konstant und Stichprobe bzgl. p-Verteilung einfach erzeugt werden kann
(über Inversion oder v. Neumann)

Beispiel $\quad I = \int_0^2 e^{-x}\,dx \qquad$ (vgl. S. 8.31)



$$g(x) = \begin{cases} e^{-x} & 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

einfachster Fall: sei $p(x)$ linear, d.h.

$$p(x) = \begin{cases} a \cdot x + b & 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

1) Normierung

$$\int_0^2 p(x)\,dx = 1 \quad \Rightarrow \quad \frac{a}{2}x^2 + bx \Big|_0^2 = 1 \quad \Rightarrow a = \frac{1}{2} - b$$

Nur ein Parameter läßt sich wählen!

2) Z.V. aus Stichprobe bzgl p

Mit Inversionsverfahren

$$\underset{\uparrow}{\xi} = \int_0^x p(x')\,dx' = \frac{a}{2}x^2 + bx \quad \left(\equiv F(x)\right)$$

aus Zufallszahlengen.

$$a/2\, x^2 + bx - \xi = 0$$

$$x_{12} = \frac{-b \pm \sqrt{b^2 + 2a\xi}}{a} \quad , \text{ Vorzeichen so, daß } x_i \in [0,2)$$

3) Wie a, b

i) Falls a, b so gewählt, daß $p(2) = 0$, dann

$\langle \frac{g}{p} \rangle$ am "rechten" Rand stark variabel und groß

$\Rightarrow$ ii) $p(2) = e^{-2} = g(2)$

$$b = 1 - e^{-2} \quad ; \quad a = -\frac{1}{2} + e^{-2} \quad \Rightarrow p(0) = 1 - e^{-2}$$

4) <u>Algorithmus</u>

```
a = - 0.5 + exp(-2.)
b = 1 - exp(-2.)
SUM1 = 0.
SUM2 = 0.
DO I = 1,N
XI = RANDOM
x = (-b + SQRT(b·b + 2·a·XI))/a   ← Z.V. entspr. p
g = exp(-x)
p = a·x + b
```

A 8.13

$$f = g/p$$

$$SUM1 = SUM1 + f \quad \leftarrow \quad \left\langle \frac{g}{p} \right\rangle$$

$$SUM2 = SUM2 + f \cdot f \qquad \langle g^2/p^2 \rangle$$

$\left.\begin{matrix} \langle g/p \rangle \\ \langle g^2/p^2 \rangle \end{matrix}\right\}$ wird akkumuliert

ENDDO

$$\text{Integral} = SUM1/N$$

$$SIGMA = \left(\left(SUM2 - SUM1 \cdot SUM1\right)/N\right)^{\frac{1}{2}}$$

5. Vergleich mit einfacher MC-Integration bei gleicher Sequenz von Zufallszahlen (S. 8.31/32

	$N = 5$	$N = 25$	$N = 250000$
f:	1.3252 ± 0.300	0.85496 ± 0.10466	$0.86476 \pm 0.966 \cdot 10^{-3}$
Red:	0.9814 ± 0.0818	0.87461 ± 0.02885	$0.86481 \pm 0.277 \cdot 10^{-3}$

z. Vgl. "exakt" = 0.86467

<u>Ergebnis</u> Fehler bei gleicher Punktzahl ca Faktor 3.5 besser

$\Rightarrow$ Um gleiche Präzision zu erreichen, Faktor 12 weniger Punkte benötigt!
(aber auch mehr Operationen...)

8.1 Quasi-Zufallszahlen

DAS Problem bei Monte-Carlo Verfahren ist die sehr schwache Abhängigkeit des Fehlers von der Größe der Stichprobe:

$$\delta A \propto \sqrt{\frac{\sigma^2(A)}{N}}$$

Bei Verwendung von UNABHÄNGIGEN, gleichverteilten Zufallszahlen ist diese Beziehung immer gültig.

Bei "klassischen" Integratiosverfahren mit einem Gitter von N Stützstellen gilt dagegen:

- Fehler wird mindestens wie $1/N$ kleiner,
- Grund: Hier wird der betrachtete Raum mit wenigen Punkten gleichmäßiger ausgefüllt als bei unkorrelierten Punkten.
- Nachteil: es muß vorher klar sein, wie fein das Gitter für eine bestimmte Genauigkeit sein muß. Abbrechen nach einem Teil der Gitterpunkte ist nicht möglich.

Monte-Carlo Verfahren rechnen oft solange, bis eine bestimmte Genauigkeit erreicht ist:

$\Rightarrow$ QUASI- oder SUB-random Zufallszahlen. Diese sind sehr gut gleichverteilt, aber NICHT unabhängig: Die Sequenz ist so konstruiert, daß JEDER Schritt die "Lücken" optimal füllt.

$\delta A \sim O[(\log N)^d / N]$

- **Historischer Vertreter: HALTON-Sequenz H_j:**

  1. **Schreibe j als Zahl zur Basis b (b Primzahl).**
  2. **Vertausche die Reihenfolge der "Ziffern".**
  3. **Schreibe "0." (Basis-b) vor diese Ziffern.**

  Das Resultat ist H_j. In n Dimensionen wird für jede Komponente eine andere Primzahl b verwendet.

- **Modernes Verfahren: SOBOL'-Sequenzen:**

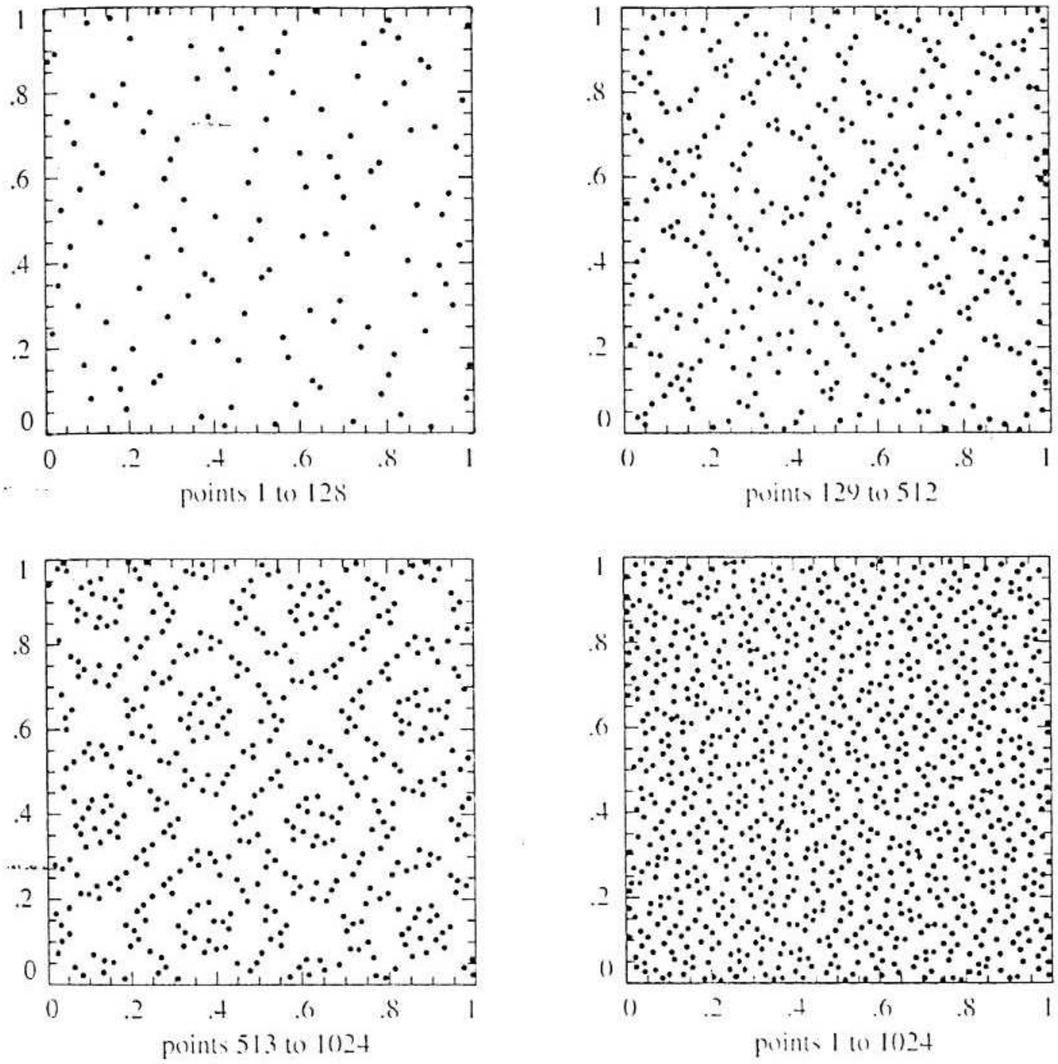


Figure 7.7.1. First 1024 points of a two-dimensional Sobol' sequence. The sequence is generated number-theoretically, rather than randomly, so successive points at any stage "know" how to fill in the gaps in the previously generated distribution.

A.8.16

Halton Sequenz zur Basis 3

$j = 1$	$j\underline{\|_3} = 1$	$\Rightarrow$ $0.1\underline{\|_3}$	$\hat{=}$ $1/3$	
2	2	0.2	2/3	
3	10	0.01	1/9	0/3 + 1/9
4	11	0.11	4/9	1/3 + 1/9
5	12	0.21	7/9	2/3 + 1/9
6	20	0.02	2/9	0/3 + 2/9
7	21	0.12	5/9	1/3 + 2/9
8	22	0.22	8/9	2/3 + 2/9
9	100	0.001	1/27	0/3 + 1/27
10	101	0.101	10/27	1/3 + 1/27
11	102	0.201	19/27	2/3 + 1/27
12	110	0.011	4/27	0/3 + 1/9 + 1/27
13	111	0.111	13/27	1/3 + 1/9 + 1/27
14	112	0.211	22/27	2/3 + 1/9 + 1/27
15	120	0.021	7/27	⋮
16	121	0.121	16/27	
17	122	0.221	25/27	
18	200	0.002	2/27	
			⋮	

Schnelle Variation der letzten Stelle $\Rightarrow$ schnelle Variation der führenden Stelle