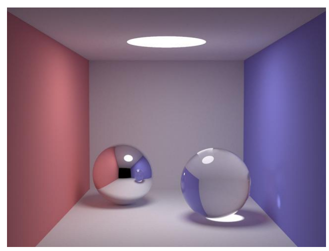
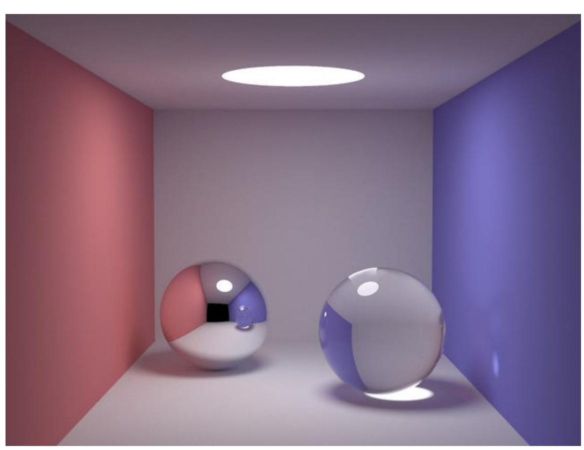
Angewandte Mathematik
Zufallszahlen

Dr. Marcel Ritter
Univ.-Prof. Dr. Matthias Harders
Sommersemester 2022



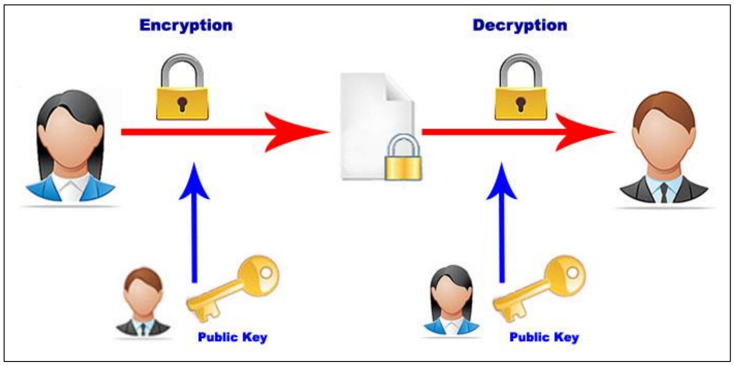
Motivation

- Zufallszahlen im Path Tracing für Global Illumination



Motivation

- Zufallszahlen für Schlüssel in der Kryptographie



Inhalt

- Einführung und Beispiele
- Arten von Zufallszahlen
- Pseudozufallszahlengeneratoren
- Umrechung gleichverteilter Zufallszahlen

Inhalt

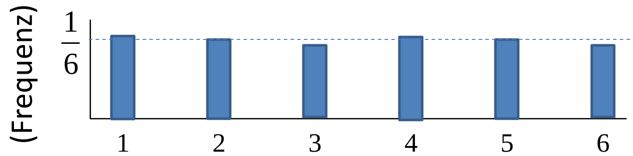
- **Einführung und Beispiele**
- Arten von Zufallszahlen
- Pseudozufallszahlengeneratoren
- Umrechung gleichverteilter Zufallszahlen

Zufallszahlen

- Eine Zufallszahl ist eine Zahl, die zufällig gemäß z.B. einer Vorschrift oder einem Experiment erzeugt wird
- (Echte) Zufallszahlen folgen dabei einer bestimmten Verteilung
- Zum Beispiel tritt jedes Ergebnis eines (fairen) Münz- oder Würfelwurfs mit gleicher Wahrscheinlichkeit auf
- Zufallszahlen werden benötigt u.a. in numerischen Berechnungen, in der Kryptographie, zur Simulation physikalischer Phänomene, zum Testen von Algorithmen, in künstlicher Intelligenz, etc.

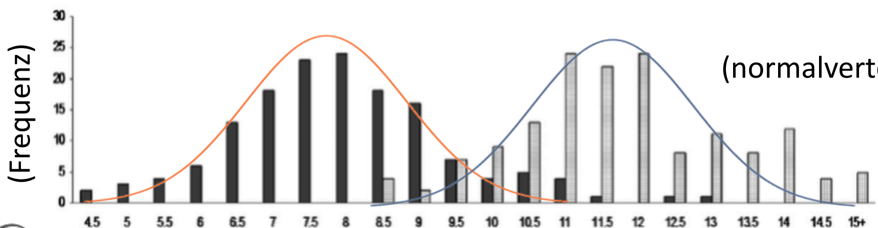
Verteilungen

- Würfe mit (fairem) sechsseitigen Würfel (N groß)



(gleichverteilt)

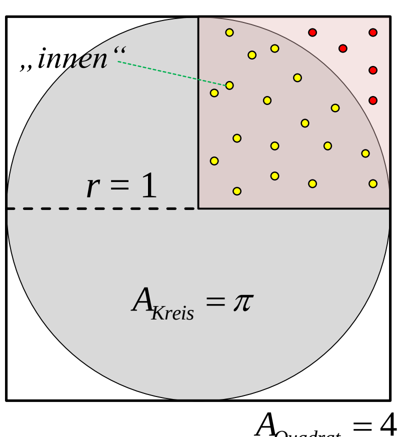
- Schuhgrößen (w (154): ■ / m (158): ■), US-Maße



(normalverteilt)

Beispiel – Annäherung von π

- Zufällige Tupel in $\mathbb{R}^2$ als Annährung von Flächen



$$\frac{A_{Kreis}}{A_{Quadrat}} = \frac{\pi}{4} \Rightarrow \pi = 4\frac{A_{Kreis}}{A_{Quadrat}}$$

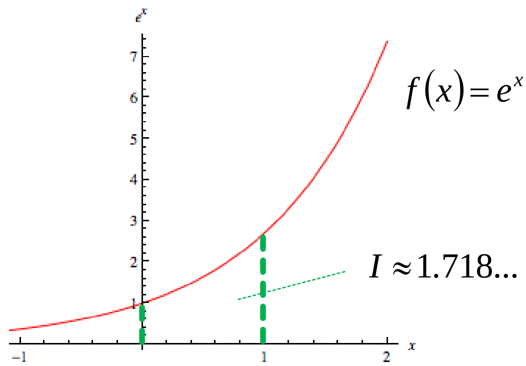
$$\pi = 4\frac{A_{Kreis}}{A_{square}} = 4\frac{4A_{Viertelkreis}}{4A_{Viertelquadrat}}$$

$$\pi \approx \frac{4 \cdot n_{innen}}{n_{gesamt}} = \frac{4 \cdot 18}{23} = 3.1304\ldots$$

Beispiel – Berechnung eines Integrals

- Zufallszahlen für Monte-Carlo Integration

$$I = \int_0^1 e^x dx = e^x \Big|_0^1 = e - 1 \approx 2.718281\ldots - 1 \approx 1.718\ldots$$



Beispiel – Berechnung eines Integrals

- Zufallszahlen für Monte-Carlo Integration

$$I = \int_0^1 e^x dx \approx \frac{1}{N} \sum_{i=1}^{N} e^{\xi_i} = \hat{I} \qquad \xi_i \in [0,1]$$

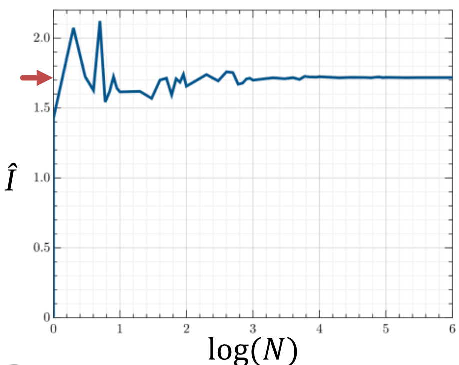
(gleichverteilte Zufallszahlen)

```
function MonteCarloExp( N )
    R = rand(N)
    xi = exp.(R)
    sum(xi) / N
end
```

Beispiel – Berechnung eines Integrals

- Resultat einer zufälligen Auswertung:

(Standard-
abweichung)



N	$\hat{I}$	σ
10	1.5490	0.1353
100	1.6995	0.0540
1 000	1.7254	0.0163
10 000	1.7245	0.0049
100 000	1.7214	0.0015
1 000 000	1.7185	0.0004

Inhalt

- Einführung und Beispiele
- **Arten von Zufallszahlen**
- Pseudozufallszahlengeneratoren
- Umrechung gleichverteilter Zufallszahlen

Arten von Zufallszahlen

- Echte Zufallszahlen resultieren aus realen, nicht (bzw. schwer) vorhersagbaren physikalischen Prozessen
- Beispiele: Würfel-/Münzwürfe, radioaktiver Zerfall, Lottoziehungen, quantenphysikalische Effekte, etc.
- Für viele Anwendungen sind diese nicht praktikabel (Kosten, Zeitaufwand, Komplexität)
- Im Computer werden statt dessen Pseudozufallszahlen verwendet, die deterministisch berechnet werden
- Je nach Algorithmus erfüllen die erzeugten Zahlen nur zum Teil die Eigenschaften echter Zufallszahlen

Echte Zufallszahlen

- Bespiele:
  - hotbits:
    Verwendet eine eigene Hardware bestehend aus Geigerzähler und Strahlungsquelle Cæsium-137.
  - random.org:
    Mehrere Radios liefern Rauschen aus der Atmosphäre
  - Lavarand:
    Bilder von Lavalampen erzeugen eine Zufallszahl

- Gut für Abbildung echter Zufallsprozesse
- Anzahl an Zufallszahlen pro Sekunde limitiert

Historisches Beispiel – Buch mit Zufallszahlen

- “A Million Random Digits with 100,000 Normal Deviates” (RAND Cooperation 1955)

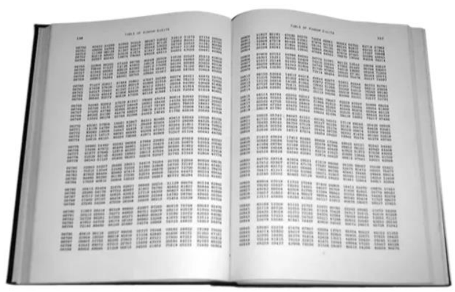
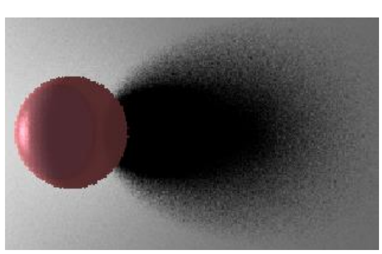


TABLE OF RANDOM DIGITS

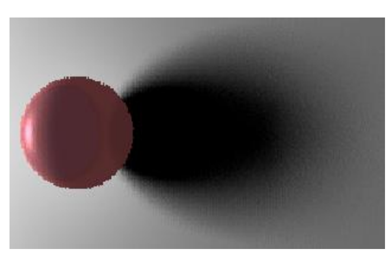
93716	16894	66083	24653	84609
32886	59780	08355	60860	29735
92052	46215	55121	29281	59076
95819	06831	00911	98936	76355
39510	35905	14060	40619	29549
27699	06494	14845	46672	61958
92962	61773	41839	55382	17267
10274	12202	39685	23309	10061
75867	20717	74416	53166	35208
85783	47619	53152	67433	35663
35075	33949	42614	29297	01918
56623	34442	34994	41374	70071
36409	83232	99385	41600	11133
57620	52606	66497	68646	78138
07399	37408	48509	23929	27482

Beispiel – Schattenberechnung im Ray Tracing

- Monte Carlo Sampling → sehr viele Zufallszahlen möglichst schnell
- Visuelle Qualität der Schatten abhängig von der Wahl der Methode zur Erzeugung der Zufallszahlen



36 Pseudozufall Samples



36 Halton-Sequenz Samples

Pseudozufallszahlengeneratoren

- Ziel: Erzeugen einer Zahlenfolge, üblicherweise im Intervall $[0,1]$, die möglichst genau den idealen Eigenschaften von echten Zufallszahlen entspricht
- Zufallszahlen meistens gleichverteilt; andere Verteilungen und/oder Intervalle durch Transformation
- Erzeugung sollte möglichst schnell/effizient sein
- Manchmal ist es von Interesse (z.B. für Tests), dass diese reproduzierbar sind
- Oft ist ein Startwert (engl. „seed“) erforderlich, von welchem Zahlen deterministisch erzeugt werden

Inhalt

- Einführung und Beispiele
- Arten von Zufallszahlen
- **Pseudozufallszahlengeneratoren**
- Umrechung gleichverteilter Zufallszahlen

Mittquadratmethode

- Erstes Beispiel eines Zufallszahlengenerators (von Neumann und Metropolis, 1940er Jahre)
- Algorithmus:
  - Start mit vierstelliger, positiver ganzer Zahl z_0
  - Berechne $z_i \cdot z_i$ (Zahl mit bis zu acht Ziffern)
  - Mittlere vier Ziffern ergeben Zufallszahl r_i im Intervall $[0,1]$ (als Nachkommastellen), sowie die neue vierstellige Zahl z_{i+1}

i	z_i	$z_i \cdot z_i$	r_i
0	1234	01**5227**56	0.5227
1	5227	27**3215**29	0.3215
2	3215	09**7656**25	0.7656

Mittquadratmethode – Nachteile

- Zahlenfolgen haben evtl. Fixwerte oder kurze Zyklen

i	z_i	$z_i \cdot z_i$
0	2176	04734976
1	7349	54007801
2	0078	00006084
3	0060	00003600
4	0036	00001296
5	0012	00000144
6	0001	00000001
7	0000	00000000

i	z_i	$z_i \cdot z_i$
0	3792	14379264
1	3792	…

i	z_i	$z_i \cdot z_i$
0	5030	25300900
1	3009	09054081
2	0540	00291600
3	2916	08503056
4	5030	…

Lineare Kongruenzgeneratoren

- Erzeugen rekursiv eine Folge ganzer Zahlen $z_1, z_2, \ldots$ im Intervall $[0,m-1]$ mittels

$$z_{i+1} = (a \cdot z_i + c) \bmod m, \qquad i = 0,1,2,\ldots$$

  mit Modulus $m = 2,3,4,\ldots$, Startwert $z_0 \in \{0,1,\ldots,m-1\}$, Faktor a und Inkrement c, mit $a,c \in \{1,2,\ldots,m-1\}$
- Zufallszahlen in $[0, (m-1)/m]$ ergeben sich dann als

$$r_i = \frac{z_i}{m}$$

- Wahl der Werte z_0, a, c, m hat großen Einfluss auf die statistischen Eigenschaften der Zahlenfolge

Lineare Kongruenzgeneratoren

- Je nach Wahl der Parameter ergeben sich evtl. niedrige Periodenlängen (maximal möglich: m)
- Beispiel: $a = 13$, $c = 0$, $m = 64$ (mit verschiedenen z_0)

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
z_i	1	13	41	21	17	29	57	37	33	45	9	53	49	61	25	5	1
z_i	2	26	18	42	34	58	50	10	2								
z_i	3	39	59	63	51	23	43	47	35	7	27	31	19	55	11	15	3
z_i	4	52	36	20	4												

z_0

Periodenlängen

- Ein Linearer Kongruenzgenerator besitzt die volle Periodenlänge m, wenn gilt:
  - m und c sind teilerfremd (d.h. der einzige positive ganze Teiler ist 1)
  - Wenn 4 ein Teiler von m ist, dann auch von a–1
  - Wenn Primzahl q ein Teiler von m ist, dann dividiert q auch a–1
- Sind diese Bedingungen erfüllt, ist die Periodenlänge auch unabhängig vom Startwert

Maximale Periodenlänge – Beispiel

- Zahlenfolge für $a = 41$, $c = 23$, $m = 32$, $z_0 = 0$

$$z_{i+1} = (41 \cdot z_i + 23) \bmod 32, \qquad i = 0, 1, 2, \ldots$$

0	23	6	13	12	3	18	25	24
	15	30	5	4	27	10	17	16
	7	22	29	28	19	2	9	8
	31	14	21	20	11	26	1	**0**

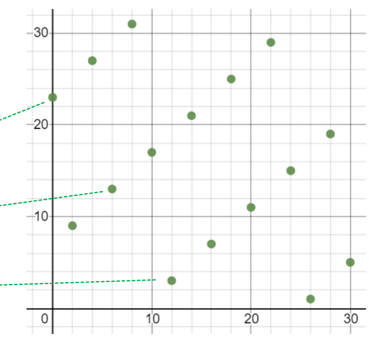
Satz von Marsaglia

- Weist man die durch einen Linearen Kongruenz-generator erzeugten Zufallszahlen Tupeln in $\mathbb{R}^k$ zu, dann liegen diese auf maximal $\sqrt[k]{m \cdot k!}$ parallelen Hyperebenen
- Beispiel in $\mathbb{R}^2$ (mit Werten der vorherigen Folie)

$$(z_0, z_1) = (0, 23)$$

z.B.: $$(z_2, z_3) = (6, 13)$$

$$(z_4, z_5) = (12, 3)$$



Halton-Sequenzen

- Zufällig erscheinende Punkte in $\mathbb{R}^2$, die einen Bereich gleichmäßig abdecken, können deterministisch über Halton-Sequenzen erzeugt werden
- Startpunkt sind dabei van-der-Corput Sequenzen, gegeben zur Basis b
- Eine natürliche Zahl $n \geq 1$ ist gegeben, für Basis b mit L Ziffern d_k, als:

$$n = \sum_{k=0}^{L-1} d_k(n) b^k \qquad 0 \leq d_k(n) < b$$

- Beispiel ($L = 4$): $13_{10} = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 1101_2$

Halton-Sequenzen

- Die n-te Zahl der van-der-Corput Sequenz zur Basis b ist damit gegeben als:

$$g_b(n) = \sum_{k=0}^{L-1} d_k(n) b^{-(1+k)}$$

- Es werden somit die Ziffern der Zahl n, in Basis b, invertiert nach dem Dezimalpunkt geschrieben
- Durch Kombination von Zahlen aus solchen Sequenzen, jeweils erzeugt mit teilerfremden Basen, ergeben sich pseudozufällige Tupel (die sogenannten Halton-Sequenzen)

Halton-Sequenzen

- Beispiel: van-der-Corput Sequenz zur Basis 2

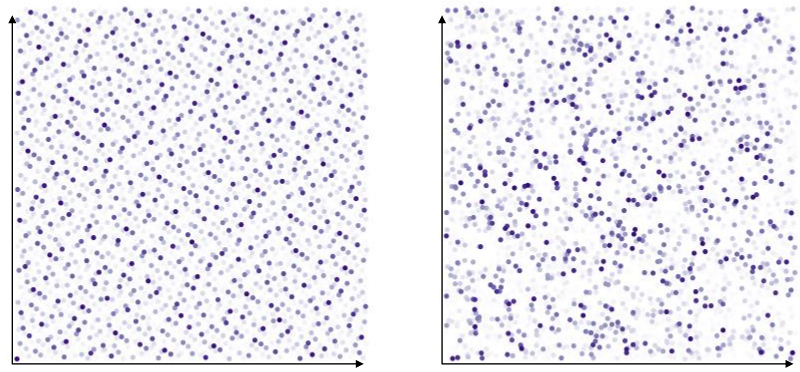
n Dezimal	0	1	2	3	4	5	6	7
n Binär	0000	0001	0010	0011	0100	0101	0110	0111
Invertiert	0000	1000	0100	1100	0010	1010	0110	1110
Kommazahl	0.0000	0.1000	0.0100	0.1100	0.0010	0.1010	0.0110	0.1110
Dezimal	0	$1/2$	$1/4$	$3/4$	$1/8$	$5/8$	$3/8$	$7/8$

- Beispiel: van-der-Corput Sequenz zur Basis 3

$1/3,\ 2/3,\ 1/9,\ 4/9,\ 7/9,\ 2/9,\ 5/9,\ 8/9,\ 1/27,\ \dots$

Halton-Sequenzen

- Halton-Sequenz in $\mathbb{R}^2$ mit Basis 2 und 3
  $(\frac{1}{2}, \frac{1}{3}), (\frac{1}{4}, \frac{2}{3}), (\frac{3}{4}, \frac{1}{9}), \ldots$



(Halton-Sequenz) (Pseudozufalls-Tupel)

Generieren von (Pseudo) Zufallszahlen

- Betriebssystem oder Bibliotheken, zum Beispiel
- Linux:
  - dev/urand
- Julia:
  - Random.rand(), .bitrand(), .randexp()
- C++11 random:
  - uniform_real_distribution, normal_distribution, chi_squared…
- Python:
  - random.uniform(), .gauss(), .expovariate(), .
- …

Generieren von (Pseudo) Zufallszahlen

- Kann mit einem spezifischen ‚seed' initialisiert werden
- Man erhält dann immer die gleiche Abfolge

```
using Random

Random.seed!(1234)
rand(1)

Random.seed!(666)
rand(1)
```

0.5908446386657102

0.013871894106960436

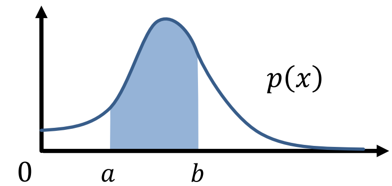
Inhalt

- Einführung und Beispiele
- Arten von Zufallszahlen
- Pseudozufallszahlengeneratoren
- **Umrechung gleichverteilter Zufallszahlen**

Wahrscheinlichkeitsdichtefunktion

- WDF (PDF) beschreibt die Wahrscheinlichkeit einer Zufallsvariable X sich in einem Intervall [a,b] zu befinden:

$$P(a \leq X \leq b) = \int_a^b p(x)dx$$

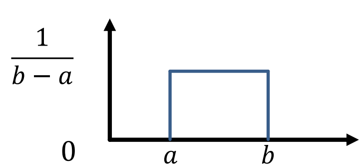


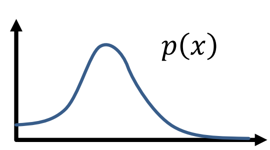
- Normiert, sodass

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

Inversionsmethode

- Gleichverteilten Zahlen können mit bekannter WDF umgerechnet werden





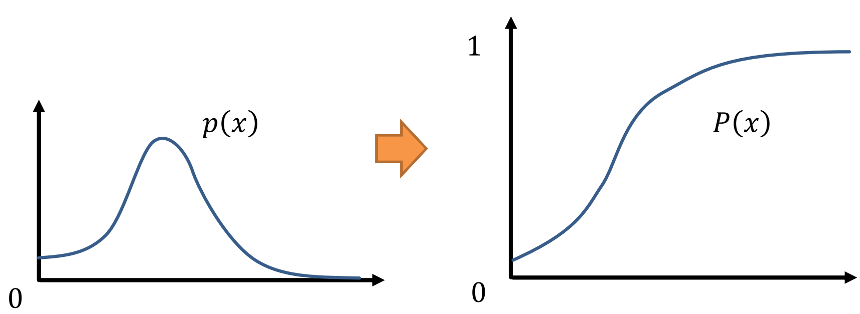
- Zahlengeneratoren für Gleichverteilungen können verwendet werden andere Verteilungen zu erreichen
  - Generatoren oft normiert auf das Intervall $[a = 0, b = 1]$

Inversionsmethode

- Methode:
  1. Berechnung der kumulierten WDF (Verteilung):
     $$P(X) = \int_{-\infty}^{X} p(x)dx$$
  2. Inverse Funktion finden (nicht immer möglich)
  3. Generieren einer Gleichverteilen Zahl ξ
  4. Transformation in Bezug auf die Verteilung
     $$X = P^{-1}(\xi)$$

Inversionsmethode

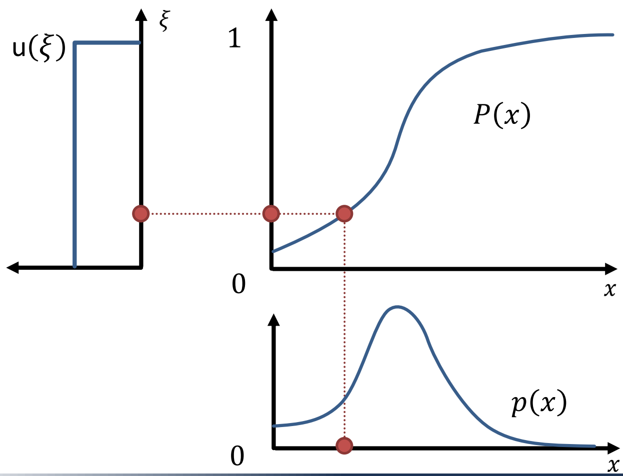
- Kumulierte WDF:



$$P(X) = \int_{-\infty}^{X} p(x)dx$$

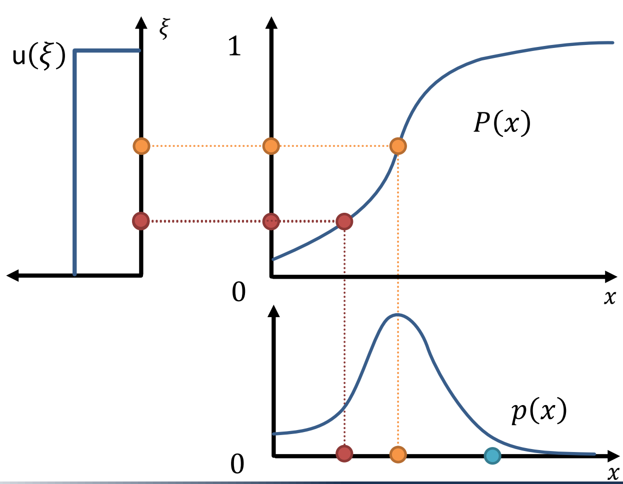
Inversionsmethode

- Umrechnung graphisch:



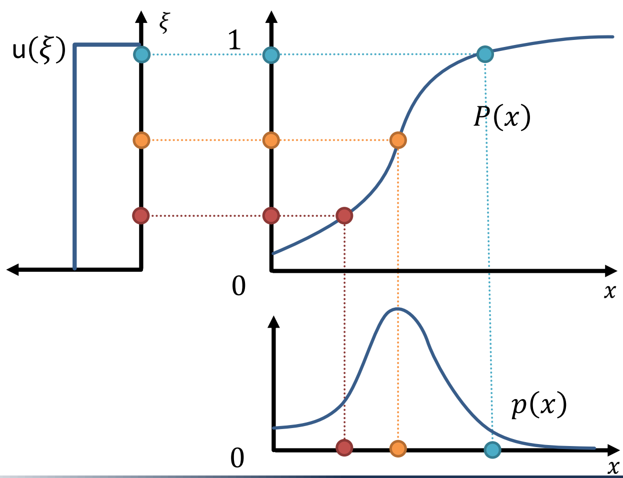
Inversionsmethode

- Umrechnung graphisch:



Inversionsmethode

- Umrechnung graphisch:



Inversionsmethode

- Beispiel:
  - Zufallszahlen im Intervall [0,π] und Verlauf: $\sin(x)/2$

$$P(X) = \int_0^X \frac{\sin(x)}{2} dx = \frac{-\cos(x)}{2}\bigg|_0^X = \frac{-\cos(X)+1}{2}$$

$$Y = (-\cos(X)+1)/2$$

$$X = (-\cos(Y)+1)/2$$

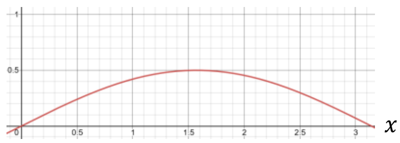
$$\cos(Y) = 1 - 2X$$

$$Y = \underbrace{\cos^{-1}(1-2X)}_{P^{-1}(\xi)}$$

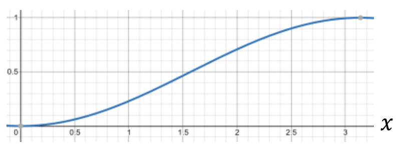
Inversionsmethode

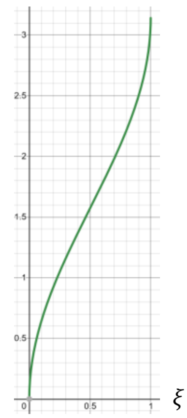
- Beispiel:

$$p(x) = \frac{\sin(x)}{2}$$



$$P(x) = \frac{-cos(x) + 1}{2}$$

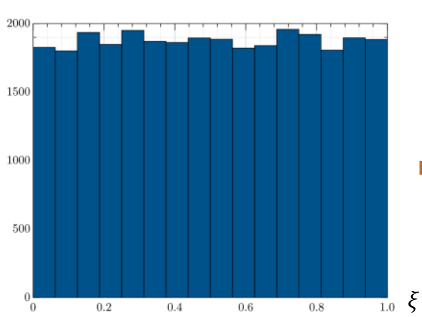




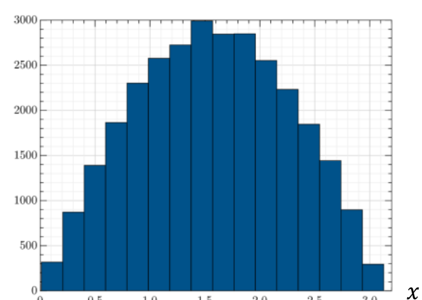
$$P^{-1}(\xi) = \arccos(1 - 2\xi)$$

Inversionsmethode

- Beispiel:
  - Histogramme von 30000 Zufallszahlen



`rand(N)`

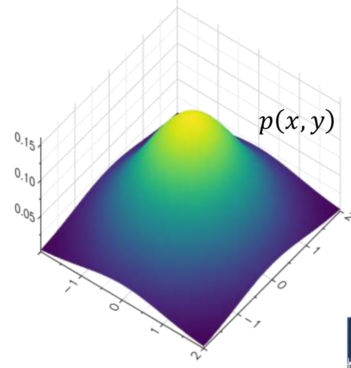


`acos.((1 .- 2.*rand(N) ))`

Zweidimensionale Verteilung

- Wahrscheinlichkeitsverteilungen können multivariat erweitert werden
- Verteilungsfunktion einer 2D Zufallsvariable

$$P(X,Y) = \int_{-\infty}^{Y} \int_{-\infty}^{X} p(x,y)\, dx\, dy$$



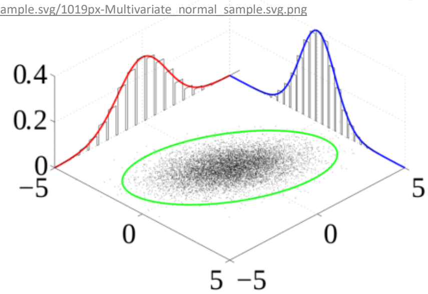
Bedingte Dichte und Randdichte

https://upload.wikimedia.org/wikipedia/commons/thumb/9/95/Multivariate_normal_sample.svg/1019px-Multivariate_normal_sample.svg.png

- Randdichten:

$$p(x) = \int p(x,y)\, dy$$

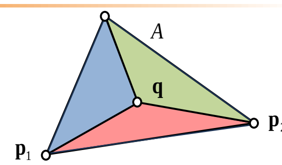
$$p(y) = \int p(x,y)\, dx$$



- Bedingte WDF:

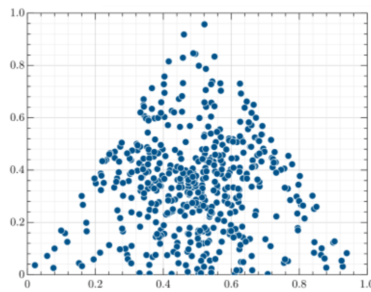
$$p(y|x) = \frac{p(x,y)}{p(x)}$$

Homogenes Sampling in Baryzentr. Koord.



- Problembeispiel:
  - Homogenes Sampling eines Dreiecks in Baryzentrischen Koordinaten

    $\mathbf{q} = \lambda_0 \mathbf{p}_0 + \lambda_1 \mathbf{p}_1 + \lambda_2 \mathbf{p}_2$ $\lambda_0 + \lambda_1 + \lambda_2 = 1$
  - λ_i als unabhängige Zufallszahlen (normiert durch Summe)



Verteilung nicht gleichförmig!

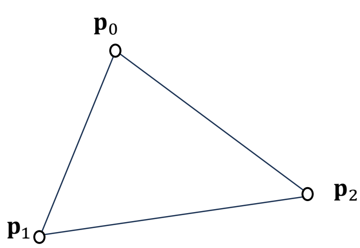
Homogenes Sampling in Baryzentr. Koord.

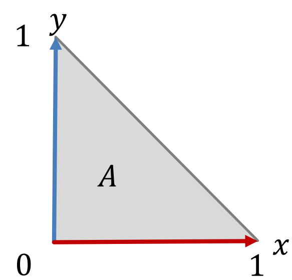
- Methode:
  - Berechne zwei abhängige Koordinaten (λ_0, λ_1) über bedingte Verteilungen und $\lambda_2 = 1 - \lambda_0 - \lambda_1$

1. Berechne Randdichte $p(x) = \int p(x,y)\, dx$
2. Berechne bedingte Dichte $p(y|x) = \dfrac{p(x,y)}{p(x)}$
3. Samplen der Randdichteverteilung für λ_0
4. Samplen der bedingten Verteilung für λ_1

Homogenes Sampling in Baryzentr. Koord.

- Annahme Einheitsdreieck mit Flächeninhalt $A = \frac{1}{2}$
- Verallgemeinert durch Baryzentrische Koordinaten



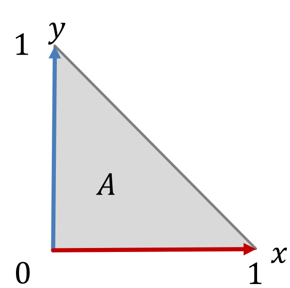


Homogenes Sampling in Baryzentr. Koord.

- Aufstellen der 2D Verteilung
  - Konstante Dichte
  - Normiertes Integral

$$\int_A p(x,y)\,dA = p(x,y)\overbrace{\int_A dA}^{\frac{1}{2}} = 1$$

$$p(x,y) = 2$$



Homogenes Sampling in Baryzentr. Koord.

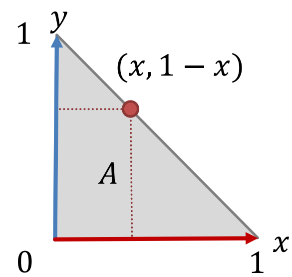
- Berechnen der Randdichte:

$$p(x) = \int p(x,y)\,dy = \int_0^{1-x} p(x,y)dy =$$

$$= 2\int_0^{1-x} dy = 2 - 2x$$

- Randverteilung:

$$P(X) = \int_0^X p(x)dx = 2X - X^2$$



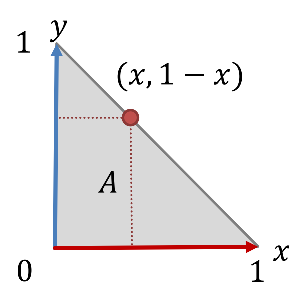
Homogenes Sampling in Baryzentr. Koord.

- Berechnen der bedingten WDF:

$$p(y|x) = \frac{p(x,y)}{p(x)} = \frac{2}{2-2x} = \frac{1}{1-x}$$

- Bedingte Verteilung:

$$P(Y) = \int_0^Y p(y|x)dy = \frac{Y}{1-X}$$



Homogenes Sampling in Baryzentr. Koord.

- Inversion der Verteilungsfunktion zum Steuern der Samplinghäufigkeit

$$\check{\xi}_0 = 2X - X^2$$

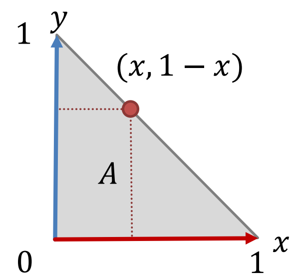
$$1 - \xi_0 = 2X - X^2$$

$$\xi_0 = 1 - 2X + X^2$$

$$\xi_0 = (1 - X)^2$$

$$\sqrt{\xi_0} = 1 - X$$

$$X = 1 - \sqrt{\xi_0}$$



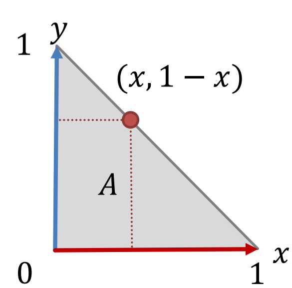
Homogenes Sampling in Baryzentr. Koord.

- Inversion der Verteilungsfunktion zum Steuern der Samplinghäufigkweit

$$\xi_1 = \frac{Y}{1 - X}$$

$$\xi_1 = \frac{Y}{\sqrt{\xi_0}}$$

$$\mathrm{Y} = \xi_1 \sqrt{\xi_0}$$

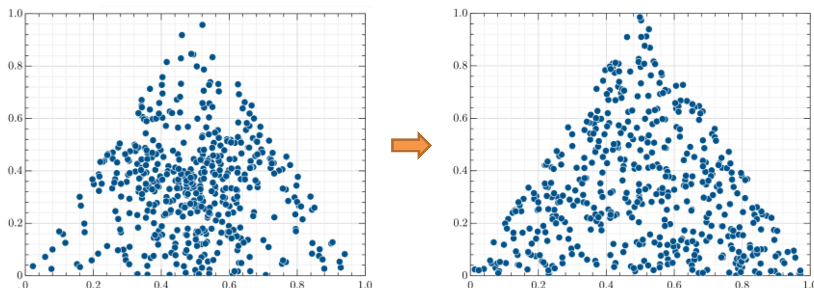


Homogenes Sampling in Baryzentr. Koord.

- Zufallspunkte im Dreieck mittels zweier gleichförmig verteilten Zufallszahlen ξ_0 und $\xi_1 \epsilon [0,1]$

$$\lambda_0 = 1 - \sqrt{\xi_0} \qquad \lambda_2 = 1 - \lambda_0 - \lambda_1 = \sqrt{\xi_0}(1 - \xi_1)$$

$$\lambda_1 = \xi_1 \sqrt{\xi_0} \qquad \mathbf{q} = \lambda_0 \mathbf{p}_0 + \lambda_1 \mathbf{p}_1 + \lambda_2 \mathbf{p}_2$$



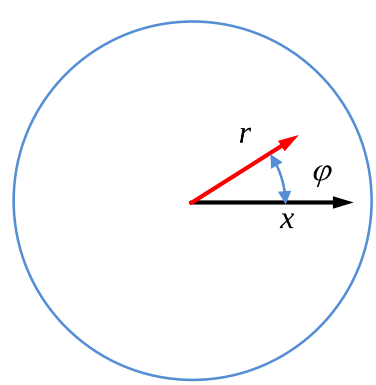
Homogenes Sampling in Polarkoordinaten

- Kreisscheibe mit Radius $r = 1$

$$A_{disk} = r^2 \pi = \pi$$

$$\int_A p(x,y) dA = 1$$

$$\rightarrow p(x,y) = 1/\pi$$



- Polarkoordinaten

$$p(x,y) = p(r,\varphi)/r$$

$$p(r,\varphi) = r\, p(x,y) = \frac{r}{\pi}$$

Homogenes Sampling in Polarkoordinaten

- Randdichte:

$$p(r) = \int_0^{2\pi} p(r,\varphi)d\varphi = \int_0^{2\pi} \frac{r}{\pi} d\varphi = 2r$$

- Randverteilung:

$$P(R) = \int p(r)dr = 2\frac{R^2}{2} = R^2$$

- Inversion:

$$\hat{\xi}_0 = R^2 \qquad \rightarrow R = \sqrt{\hat{\xi}_0}$$

Homogenes Sampling in Polarkoordinaten

- Bedingte WDF:

$$p(\varphi|r) = \frac{p(r,\varphi)}{p(r)} = \frac{r/\pi}{2r} = \frac{1}{2\pi}$$

- Verteilung:

$$P(\Phi|r) = \int p(r,\varphi)\, d\varphi = \frac{\Phi}{2\pi}$$

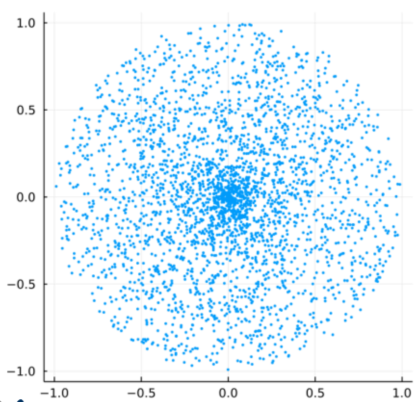
- Inversion:

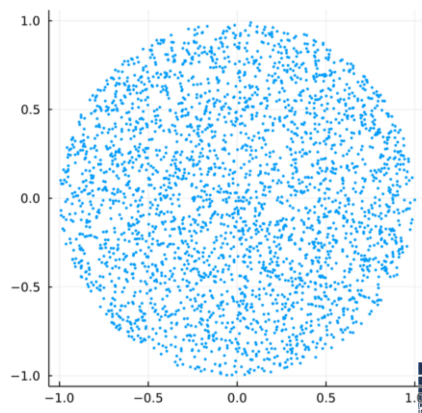
$$\hat{\xi}_1 = \frac{\Phi}{2\pi} \qquad \rightarrow \Phi = \hat{\xi}_1\, 2\pi$$

Homogenes Sampling in Polarkoordinaten

$$x = r\cos(\varphi) = \xi_0 \cdot \cos(\xi_1\, 2\pi)$$
$$y = r\sin(\varphi) = \xi_0 \cdot \sin(\xi_1\, 2\pi)$$

$$x = r\cos(\varphi) = \sqrt{\xi_0} \cdot \cos(\xi_1\, 2\pi)$$
$$y = r\sin(\varphi) = \sqrt{\xi_0} \cdot \sin(\xi_1\, 2\pi)$$





Einige Hilfreiche Weblinks

- Interaktive Beispiele zu Halton-Sequenzen
  https://observablehq.com/@jrus/halton
- Multidimensionales Sampling
  https://www.pbr-book.org/3ed-2018/Monte_Carlo_Integration/2D_Sampling_with_Multidimensional_Transformations

Vorlesungsplan

Datum	Thema	Proseminar
11.03.22	Einführung, Grundlagen, Funktionen	*(Beginn zuvor am 8.3.)*
18.03.22	Differentialrechnung	
25.03.22	Integralrechnung	
01.04.22	Differentialgleichungen	
08.04.22	Weitere Funktionen	
	Osterferien	
29.04.22	Reihen und Folgen	
06.05.22	Numerische Auswertung von Funktionen	
13.05.22	Lösung von Gleichungssystemen	
20.05.22	Interpolation	
27.05.22	Zufallszahlen	
03.06.22	Komplexe Zahlen	
10.06.22	Klausurvorbereitung	
01.07.22	**Klausur**	