Statistik

Wahrscheinlichkeiten

Zufallszahlen

Monte-Carlo Methoden

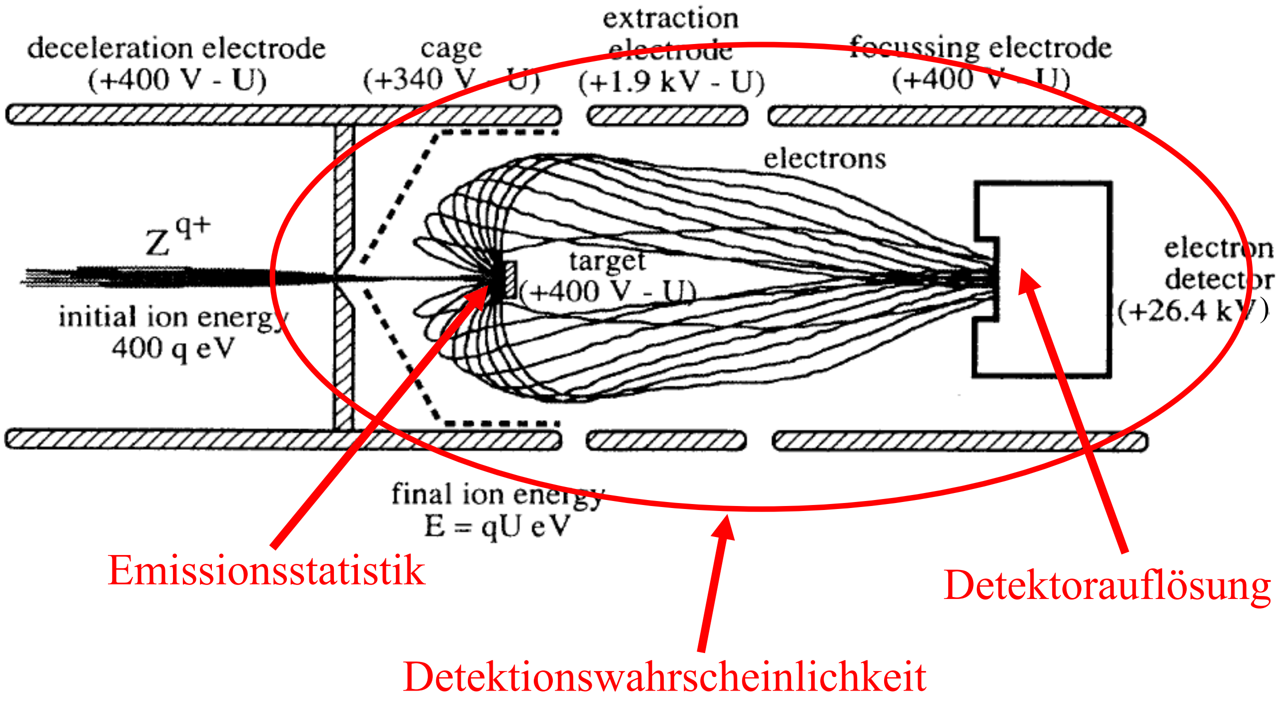
Wozu Beschäftigung mit Statistik?

Quellen von Ungewißheiten:

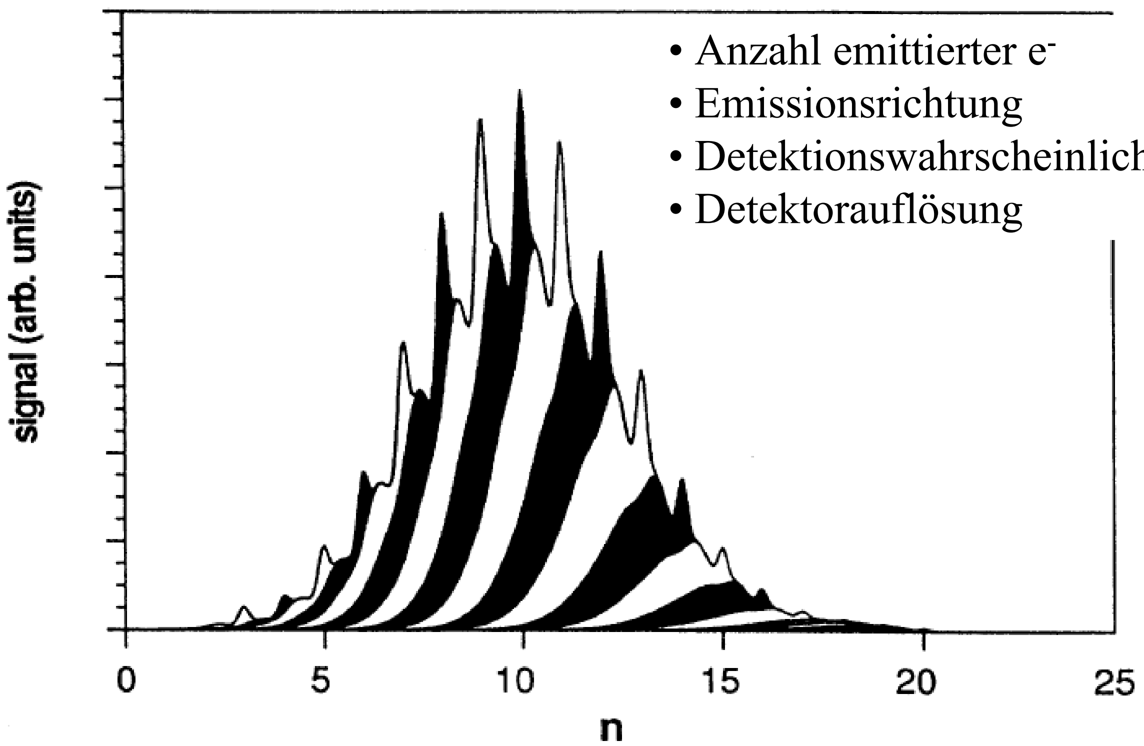
- experimentelle Gründe (experimenteller Aufbau, begrenzte Detektorauflösung, Rauschen, ...)
- Physik (Quantenmechanik, Vielteilchensysteme, ...)
- Simulationsmethode (Monte-Carlo-Integration)

Statistik durch Experiment

Beispiel: **Elektronenemission in Stößen von Ionen mit Festkörperoberflächen**



Meßergebnisse



- Anzahl emittierter e^-
- Emissionsrichtung
- Detektionswahrscheinlichkeit
- Detektorauflösung

Numerik: Ermittlung der Wahrscheinlichkeiten

Wahrscheinlichkeitsverteilungen:

Rekonstruktion einer Wahrscheinlichkeitsverteilung aus Meßdaten:

a) Berechne die n-ten Momente $\langle X^n \rangle$ des Datensatzes

b) Berechne $\phi_X(k) = \sum_{n=0}^{\infty} \frac{(ik)^n \langle X^n \rangle}{n!} = \langle e^{ikX} \rangle$

c) $\phi_X(k)$ ist die Fouriertransformierte von $f(x)$

andere Methode: Kumulantenentwicklung (bessere Konvergenz)

$$\phi_X(k) = \sum_{n=0}^{\infty} \frac{(ik)^n \langle X^n \rangle}{n!} = \exp\left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \right]$$

Entwicklung und Koeffizientenvergleich → $C_n(X)$
Vorteil: meist nur wenige C_n erforderlich zur Rekonstruktion von $f(x)$

Berechnung der Momente:

a) Verteilungsfunktion ist bekannt

$$\langle X^n \rangle = \int x^n f_X(x)\, dx$$

b) Verteilungsfunktion ist nicht bekannt

$$\langle X^n \rangle = \frac{1}{N} \sum_i \left(x_i - \langle x \rangle \right)^n$$

benannte Größen:

Mittelwert: $\langle X \rangle = \overline{x}$
Varianz: $\langle X^2 \rangle = \sigma^2$

Schiefe: $\langle X^3 \rangle / \sigma^3 = \gamma_1$
Überhöhung: $\langle X^4 \rangle / \sigma^4 = \gamma_2$
Exzeß: $\gamma_2 - 3$

Physikalische Gründe

Beispiel Thermodynamik (Vielteilchensysteme):

zentraler Grenzwertsatz: Mittelwerte vieler Messungen bilden Gaußverteilung um Mittelwert der Verteilungsfunktion

Druckmessung: „Mittelwert“ der Impulsänderungen

$\Delta p_\perp = 2 p_\perp \rightarrow$ Kraftstoß: Summe der Stöße/Fläche = Druck

Geschwindigkeitsverteilung: Maxwell-Boltzmann

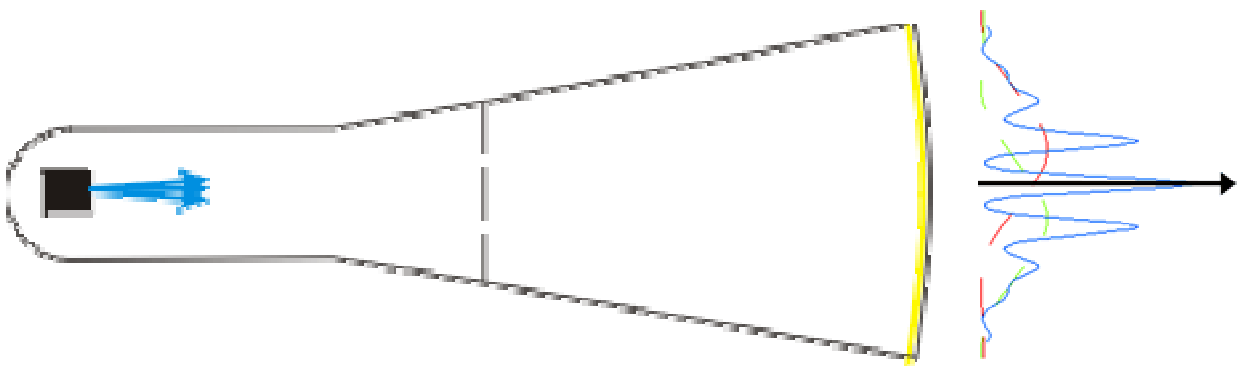
$N \sim 10^{26}$ Teilchen → Mittelwert über SEHR viele Kraftstöße

Meßfehler: $\dfrac{\Delta P}{P} \propto \dfrac{1}{\sqrt{N}}$!!!

Physikalische Gründe

Beispiel Quantenmechanik:

Wellenfunktion hat keine physikalische Bedeutung, aber Wahrscheinlichkeitsinterpretation für $|\psi|^2$



Numerik: Lösung der Differentialgleichungen (demnächst)
Simulation mithilfe statistischer Methoden

notwendige Zutat für Simulation: Zufallszahl-Generator

Wunsch:

- **<u>unkorrelierte</u>** Sequenz von Zahlen
- darf nicht vom aufrufenden Programm abhängen

aber:

- **<u>niemals</u>** auf Zufallszahl-Generator verlassen (Tests!)
- **<u>niemals</u>** auf sich selbst verlassen (Initialisierung!)

```
iseed = -int(time())
test = ran2(iseed)
```

- nur in Ausnahmefällen selbst programmieren

Erzeugung von Zufallszahlen:

Lineare Kongruenz Methode:

$$\boxed{a_{i+1} = \left(a_i \cdot b + c\right) \operatorname{mod} m}$$

Fibonacci-Generator:

$$\boxed{a_{i+1} = \left(a_i + a_{i-1}\right) \operatorname{mod} m}$$

von Neumann-Generator (Quadratmittengenerator):

$$\boxed{a_{i+1} = \tilde{a}_i \cdot \tilde{a}_i}$$

$\tilde{a}_i \ldots$ Zahlenfolge aus der Mitte von a_i

Kombinationsmethoden: Kombinationen der obigen

Häufigkeitsverteilungen von Zufallszahlen:

- auf bestimmten Bereich beschränkte ganzzahlige Zufallszahlen:

Generator liefert Zufallszahlen aus Bereich $\left[0,\ldots,(M-1)\right]$

gesucht: z.B. Würfel $\left[1,\ldots,6\right]$ $\rightarrow$ $\boxed{\tilde{R} = 1 + R(M) \bmod 6}$

Vorsicht: führt (wahrscheinlich) zu gezinktem Würfel!!!

- gleichverteilte Zufallszahlen $\tilde{R} \in \left[0, x\right)$:

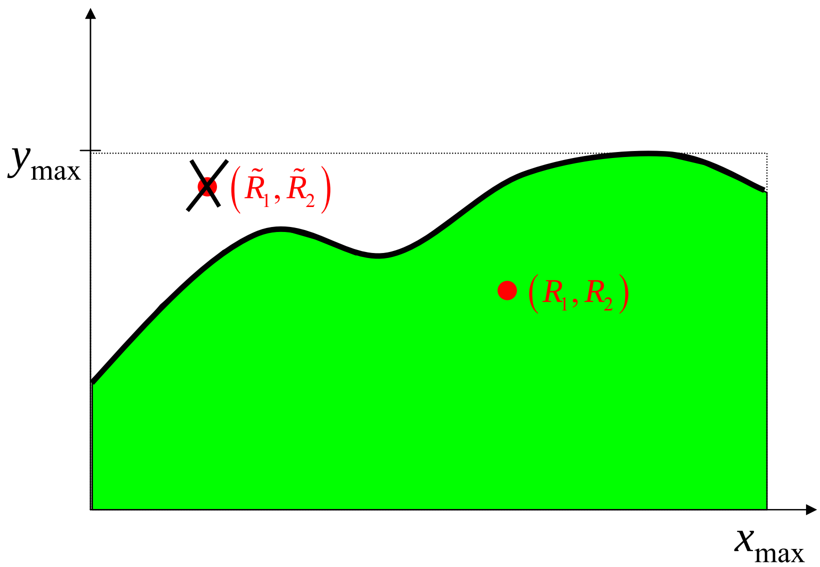
Generator liefert Zufallszahlen aus Bereich $\left[0,\ldots,(M-1)\right]$

$$\boxed{\tilde{R} = x \cdot \frac{R(M)}{M}}$$

Vorsicht: Integerdivision vermeiden!!!

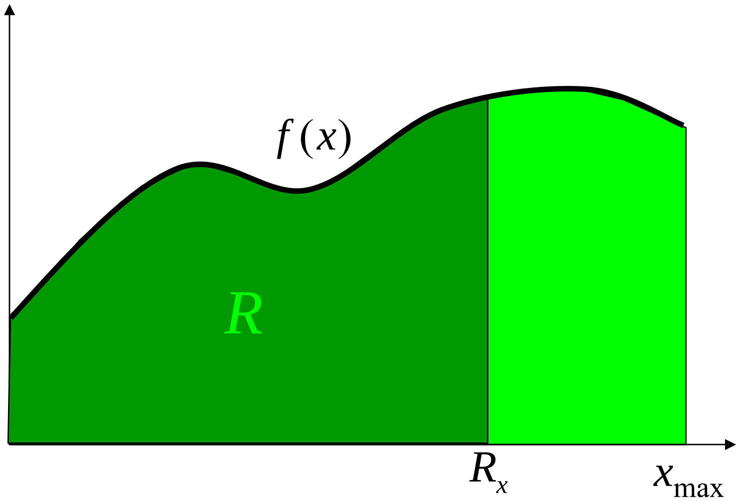
- Zufallszahlen gemäß einer vorgegebenen Verteilungsfunktion:

a) Zurückweisungsmethode:



1) generiere ein Paar Zufallszahlen $R_1 \in \left[0, x_{\max}\right); R_2 \in \left[0, y_{\max}\right)$
2) liegt der Punkt $\left(R_1, R_2\right)$ im grünen Bereich akzeptiere R_1

b) Transformationsmethode:



$$R \in [0,1) = \frac{\int_{\bar{x}=0}^{R_x} f(\bar{x})\, d\bar{x}}{\int_{\bar{x}=0}^{x_{\max}} f(\bar{x})\, d\bar{x}} = \frac{F(R_x) - F(0)}{F(x_{\max}) - F(0)} \quad \rightarrow \quad F(R_x) = F(0) + R \cdot \left(F(x_{\max}) - F(0)\right)$$

suche $F^{-1}(x)$ $\rightarrow$ $\boxed{R_x = F^{-1}\left(F(0) + R\left(F(x_{\max}) - F(0)\right)\right)}$

z.B. Exponentialverteilung $P(t) = \lambda e^{-\lambda t}$ $\rightarrow$ $t = -\frac{1}{\lambda}\ln(1-R) = -\frac{1}{\lambda}\ln R$

c) gaußverteilte Zufallszahlen:

Box-Muller-Algorithmus:

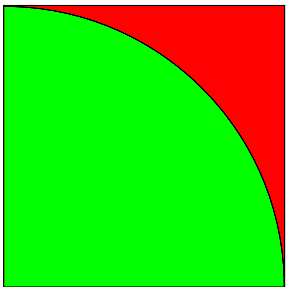
- generiere $R_1 \in [0,1);\ R_2 \in [0,1)$
- berechne $w = 2\pi R_1;\ \rho = -2\ln R_2$
- berechne $\tilde{R}_1 = \sqrt{\rho}\cos w;\ \tilde{R}_2 = \sqrt{\rho}\sin w$

Polarmethode:

- generiere $R_1 \in [0,1);\ R_2 \in [0,1)$
- berechne $S = 2R_1 - 1;\ T = -2R_2 - 1$
- falls $W^2 = S^2 + T^2 > 1$ $\rightarrow$ Neubeginn
- sonst $\tilde{R}_1 = (S/W)\sqrt{-2\ln W^2}\ ;\ \tilde{R}_2 = (T/W)\sqrt{-2\ln W^2}$

Verwendung von Zufallszahlen in Monte-Carlo Simulationen:

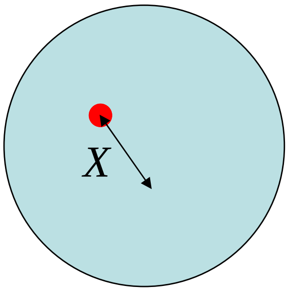
a) Berechnung von π



- generiere $R_1 \in [0,a)$; $R_2 \in [0,a)$
- Anzahl der Punkte im Quadrat $N_{\square} = N_{\square} + 1$
- falls (R_1, R_2) im Kreis $\rightarrow \quad N_{\bigcirc} = N_{\bigcirc} + 1$

$$\left.\begin{array}{l}\text{Fläche 1: } a^2 \\ \text{Fläche 2: } \pi * a^2/4\end{array}\right\} \rightarrow \quad \pi \approx 4 \cdot \frac{N_{\bigcirc}}{N_{\square}}$$

b) Brownsche Bewegung (random walk in 2D)



- generiere $R \in [0, 2\pi)$
- $\tilde{x} = x + \Delta \cos R$; $\tilde{y} = y + \Delta \sin R$

$$\langle X^2 \rangle \propto N$$

Übungsbeispiel 1

- Programmieren Sie einen „Zufallszahlen-Generator“ für den Zahlenbereich (0,1) gemäß der Methode der linearen Kongruenz mit selbstgewählten Konstanten *a*, *c*, *m*. Führen Sie folgenden Qualitätstest durch: Schreiben Sie Tripel aufeinanderfolgender Zufallszahlen in eine Datei und plotten Sie das Ergebnis in `gnuplot` (eine schlechte Wahl der Konstanten ist $a = 65539$, $c = 0$, $m = 2^{31} - 1 = 2147483647$; wurde lange auf IBM-Mainframes verwendet). Wie macht sich die mangelnde Qualität des Generators bemerkbar?

Übungsbeispiel 2

- Führen Sie eine Monte-Carlo-Integration der Funktion $f(x) = x^2$ in den Grenzen [0,1] aus und stellen Sie das Integrationsergebnis als Funktion der Anzahl der „Schüsse“ dar. Wie schnell konvergiert das Ergebnis gegen das korrekte Resultat?

Übungsbeispiel 3

- Programmieren Sie einen Würfel unter Verwendung eines beliebigen Zufallszahl-Generators (z.B. die Routine `ran.f` aus dem Verzeichnis `/home/EDV2/Unterlagen`). Berechnen Sie nach 60, 600, 6000, …, $6*10^7$ Würfen die Häufigkeit, mit der 1, 2, 3, …, 6 gewürfelt wurde. Zeigen Sie mithilfe von gnuplot, daß die mittlere relative Abweichung vom Erwartungswert $N/6$ mit steigender Wurfzahl wie $1/\sqrt{N}$ sinkt. Z.B. Würfelzahl n wurde I mal gewürfelt:

$$\rightarrow \quad \Delta I_n = \left| I_n - \frac{N}{6} \right|; \quad \Delta N = \frac{1}{6} \sum_{n=1}^{6} \Delta I_n$$

$$\rightarrow \quad \frac{\Delta N}{N} \propto \frac{1}{\sqrt{N}} \quad \text{(siehe Thermodynamik!)}$$