Lösung Aufgaben Seite 127

- Zeigen Sie, dass eine gemäß vorstehendem Algorithmus erzeugte Zufallszahl der gewünschten Verteilung gehorcht!

  Bew: Sei Y eine gemäß des Alg. erzeugte Zufallszahl. Dann gilt für $i = 1, \ldots, k$:

$$P(Y = x_i) = P(\sum_{x_j < x_i} p_j < u \leq \sum_{x_k \leq x_i} p_k) =$$

$$P(u \leq \sum_{x_k \leq x_i} p_k) - P(u \leq \sum_{x_k < x_i} p_j) = p_i = P(X = x_i).$$

- Implementieren Sie einen Zufallsgenerator für einen Münzwurf, bei dem Kopf, Zahl und Kante mit vorzugebenden Wahrscheinlichkeiten realisiert werden!

```
muenzwurf <- function(n, kopf=0.5, zahl=0.5, kante=0){
  randoms <- runif(n)
  result <- rep(NA, n)
  result[randoms <= kopf] <- "Kopf"
  result[(randoms > kopf) & (randoms <= kopf + zahl)]
                                                    <- "Zahl"
  result[randoms > kopf + zahl] <- "Kante"
  result
}
> muenzwurf(10, kopf=0.4, zahl=0.1, kante=0.5)
 [1] "Kopf"  "Kante" "Zahl"  "Kante" "Kopf"  "Kante" "Kante"
 [8] "Kante" "Kopf" "Kante"
```

Lösung Inversionsmethode

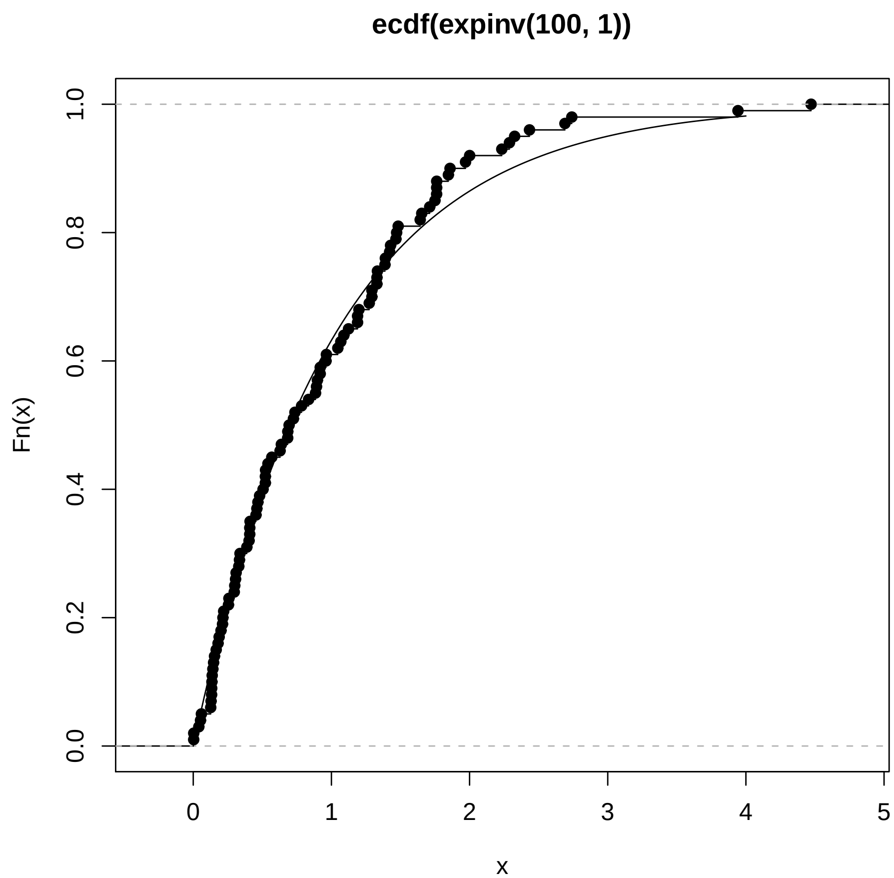
- Herleitung der Inversen $F^{-1}(y)$:

$$y = 1 - \exp(-\lambda x) \Leftrightarrow 1 - y = \exp(-\lambda x) \Leftrightarrow -\frac{\ln(1-y)}{\lambda} = x = F^{-1}(y).$$

- In R

```
expinv <- function (n, lambda ){
 -log(1-runif(n))/lambda
}
### Test:
plot(ecdf(expinv(100,1))) ; curve(pexp, 0, 4, add=TRUE)
```

Inversionsmethode: ECDF und Verteilung

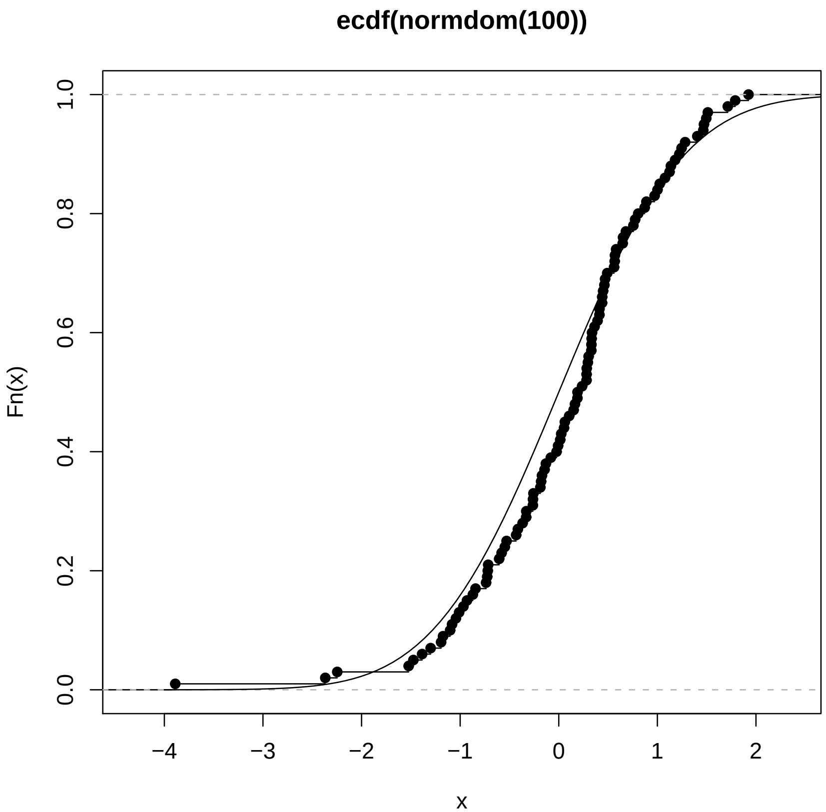


Lösung ARM für Normalverteilung

```
### direkt Alg von S. 136 übersetzen nach R
normdom <- function(n=1, mean=0, sd=1) {
    f <- function(x){(1/sqrt(2*pi)) * exp(-0.5 * x^2 )  }
    g <- function(x){0.5*exp(-abs(x))}
    result <- rep(NA, n)
    const <- 1.32
    for (i in 1:n) {
      repeat {
        x <- -log(runif(1))
        if (runif(1) <= f(x)/(const*g(x))) break
      }
      if (runif(1) < 0.5) x <- -x
      result[i] <- x
    }
    result*sd + mean }
```

ARM: ECDF und Verteilung

`plot(ecdf(normdom(100)));`curve(pnorm, -4, 4, add=TRUE)



Monte-Carlo Integration

- Literatur für diesen Abschnitt:
  Stochastische Simulation
  Grundlagen, Algorithmen und Anwendungen
  Michael Kolonko
  elektronisch verfügbar.

Monte-Carlo Integration

- Monte-Carlo Methoden versuchen, durch die geschickte Nutzung von Zufallszahlen Approximationen für “schwierige” analytische Probleme zu finden.

- Bemerkenswert ist hierbei der Übergang von der exakten analytischen Lösung zur Lösung mittels einer Zufallsvariablen.

- Die Lösungen von Monte-Carlo Ansätzen besitzen folglich Erwartungswert, Varianz und es lassen sich Konfidenzintervalle für die gesuchten Größen konstruieren.

- Im Gegensatz zu “exakten” numerischen Approximationen gibt es keine exakten Fehlerschranken, sondern Konfidenzniveaus.

Geschichte der Monte-Carlo Methoden

- Die erste Nutzung der Monte-Carlo Methode geht auf Georges-Louis Leclerc, Comte de Buffon, zurück (1737). Allerdings noch ohne Computer.

- Seine Überlegung war, wenn man ein Gitter mit Linienabstand t hat und wirft eine Nadel der Länge $l < t$ wiederholt "zufällig" auf dieses Gitter, so kann man aus der relativen Häufigkeit des Auftretens von Schnittpunkten der Nadel mit den parallelen Linien einen Schätzer für π herleiten, denn es gilt (ohne Herleitung):

$$P(\text{"Nadel schneidet Linie"}) = \int_{\theta=0}^{\frac{\pi}{2}} \int_{x=0}^{(l/2)\sin\theta} \frac{4}{t\pi} dx d\theta = \frac{2l}{t\pi}.$$

- Seien nun bei n Würfen der Nadel h Mal die Nadel so gefallen, dass eine

Linie geschnitten wird. Dann kann man schätzen:

$$\hat{P} = \frac{h}{n} = \frac{2l}{t\pi}.$$

- Stellt man damit nun die Gleichung nach π um, ergibt sich

$$\hat{\pi} = \frac{2ln}{th}!$$

- Bemerkenswerterweise wurden diese Experimente tatsächlich durchgeführt, z.B. 1901 von einem Italiener namens Mario Lazzarini, der 3408 Nadeln warf ... (Animation: http://www.cut-the-knot.org/ctk/August2001.shtml)

- Natürlich gibt es sofort mehrere Fragen:

  - Ist ein solcher Schätzer erwartungstreu?
  - Wie groß ist die Varianz?
  - Kann man ein solches Experiment durch geschickte Wahl der Zufallszahlen gut bzw. schlecht durchführen?
  - Welche Genauigkeit erlange ich durch n-fache Wiederholung?

- Buffons Nadelproblem ist ebenfalls das erste Problem der *stochastischen Geometrie*.

- In der Neuzeit wurde die Methode in Los Alamos wieder belebt. Die Hauptakteure waren dabei Ulam und von Neumann.

- Ulam wollte wissen, mit welcher Wahrscheinlichkeit seine Patiencen aufgehen, mit denen er sich während einer Krankheit die Zeit vertrieb. Von ihm stammt auch der Name, da sein Vater wohl in Monte-Carlo die Zeit mit diversen Glückspielen verbrachte.

- Er versuchte es erst kombinatorisch zu lösen, scheiterte aber (!) und verteilte daraufhin einfach 100 Mal die Karten, um den Anteil der aufgehenden Patiencen zu schätzen.

- Da in Los Alamos auch die ersten Computer standen, sah er sofort das Potential, solche Experimente von Rechner durchführen zu lassen und schrieb an von Neumann, der seinerseits die Methode so weiterentwickelt, dass man die Prozesse, die bei der Kernspaltung ablaufen, besser verstehen konnte. (1947)

- Die Differentialgleichungen zur Beschreibung der Strahlung in einem Kernreaktor waren zu kompliziert geworden, um sie numerisch oder exakt zu lösen!

- Seitdem ständige Weiterentwicklung!

Einführung: *hit or miss* zur Schätzung von π

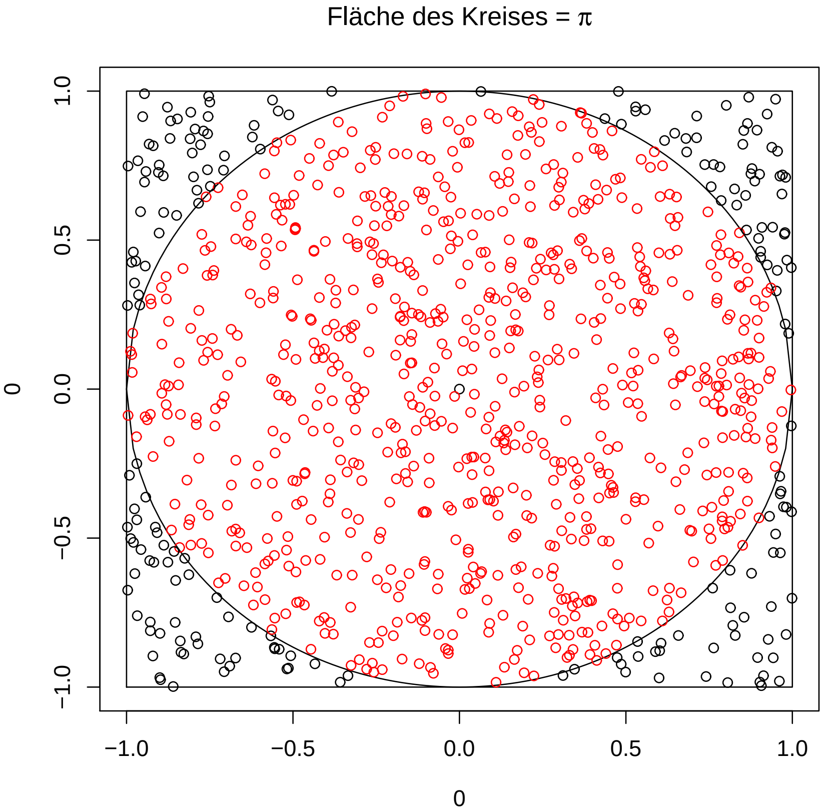
- Ein Kreis mit Radius r hat die Fläche πr^2.

- Folgender Algorithmus schätzt nun die Fläche π eines Einheitskreises ($r = 1$) über den Anteil an der Fläche eines umschreibenden Quadrats mit Kantenlänge 2.

  1. Erzeuge eine auf dem Einheitsquadrat gleichverteilte Zufallszahl $\mathbf{x} = (x, y)$. Erhöhe den Zähler n um 1.
  2. Wenn $x^2 + y^2 < 1$, liegt der Punkt im Kreis. Erhöhe Zähler i um 1.
  3. Schätze π mit $\hat{\pi} = 4 * i/n$.
  4. Gehe zu Schritt 1 oder breche ab, wenn n groß genug.

- Der Name des Verfahrens ist sofort klar: Entweder man trifft die gesuchte Fläche oder nicht. Das Verfahren ist äquivalent zur Schätzung der Erfolgswahrscheinlichkeit eines Bernoulli-Experiments.

pi **mit** ***hit or miss***



Multivariate Gleichverteilung

- Def: Die Verteilung U mit der Dichtefunktion

$$f(x_1, x_2, \dots, x_p) = \frac{1}{\int_A 1\, dx_1 dx_2 \dots dx_p} \text{ für } \mathbf{x} \in A,\ 0 \text{ sonst},$$

heißt p-dim stetige Gleichverteilung auf A.

- Ist A ein p-dim Quader Q der Form $[l_1, r_1] \times [l_2, r_2] \times \dots [l_p, r_p]$, so kann man Zufallszahlen der p-dim Gleichverteilung erzeugen, indem man komponentenweise unabhängige Zufallszahlen aus $U_{[l_i, r_i]}$, für alle i zieht. Beweis?

- Wie würde man gleichverteilten Zufallszahlen auf einem anderen Gebiet als einem Quader erzeugen?

Live-Aufgabe *hit or miss*

- Implementieren Sie diesen Algorithmus zur Schätzung von π! Ausgabe der aktuellen Schätzung von π nach jedem Schritt!
- Lösung hier nach der Vorlesung!

Formalisierung von *hit or miss*

- Der *hit-or-miss* Ansatz ist sehr universell anwendbar. Er ist verwandt mit ARM.

- Im obigen Beispiel wurde mit dieser Methode das 2-dim "Volumen" der Kreisscheibe in Relation zum Volumen des Quadrates mit Kantenlänge 2 geschätzt.

- Ganz allgemein kann man das p-dim Volumen eines beliebiegen beschränkenten Gebiets $G \subset \mathcal{R}^p$, schätzen, sofern eine Menge von Intervallen $\{I_i = [l_i, r_i], i = 1, \ldots, p\}$ exisitiert mit $G \subset I_1 \times I_2 \times \ldots \times I_p =: Q$. Sprich: G ist vollständig in einem p-dim (Hyper-)Quader Q enthalten.

- Betrachtet man den Vektor $U = (U_1, \dots, U_p)$, wobei für jedes $U_i, i = 1, \dots, p$ gilt, dass $U_i \sim U_{[l_i, r_i]}$, so ist $U \sim U_Q$.

- Definiert man nun eine Bernoulli ZV X mit "Erfolg" oder 1, wenn eine Realisierung u aus U_Q in G liegt, 0 sonst, so gilt

$$X \sim \text{Ber}(p = Vol(G)/Vol(Q)).$$

- Wiederholt man dieses Experiment n mal für ZV $X_1, \dots, X_n$, so gilt:

$$n\bar{X} := \sum_1^n X_i \sim \text{Bin}(n, p), E(Y) = np, Var(Y) = np(1-p).$$

- Aus einer Stichprobe $x_1, \dots, x_n$ ergibt sich somit als Schätzer für Vol(G)

$$\hat{\text{Vol}}(G) = \text{Vol}(Q) \cdot \bar{x}.$$

Stochastische Integration I

- Der *hit-or-miss* Ansatz kann auch der stochastischen Integration dienen.

- Gegeben sein eine Funktion $g : G \subset \mathcal{R}^p \to \mathcal{R}$, G beschränkt. Sei weiterhin $|g| < c \in \mathcal{R}$. Dann exisitiert ein Quader $Q \subset \mathcal{R}^{p+1}$ für den gilt $G \times [-c, c] \subset Q$ und der beschriebene Ansatz für *hit-or-miss* greift.

- Analog zu ARM gilt, dass je enger man den Quader Q um das Gebiet G legt, dessen Volumen man sucht, desto effizienter wird das Verfahren. Je kleiner Vol(Q), desto kleiner die Varianz in diesem Ansatz.

Stochastische Integration II

- Es gibt zur stochastischen Integration einen alternativen Ansatz, der ausnutzt, dass für eine ZV $X \sim U_{[a,b]}$ und eine Funktion $f : [a,b] \to \mathcal{R}$ gilt:
$$\frac{1}{b-a}\int_a^b f(x)dx = E(f(X)))$$
- Damit kennt man einen erwartungstreuen Schätzer für $I = \int_a^b f(x)dx$:
$$\hat{I} = (b-a)f(\bar{X}) = \frac{(b-a)}{n}\sum_1^n f(x_i).$$
für eine Stichprobe $x_1, \ldots, x_n$ aus $U_{[a,b]}$.

- Diese Definition entspricht im Ansatz den Überlegungen bei der Definition des Rieman-Integrals, nur dass die Stützstellen zufällig gewählt sind und nicht in einem gleichmäßigen Gitter angeordnet werden.

- Für niedrige Dimensionen ist die Monte-Carlo Integration nicht sehr effizient. Spezielle numerische Verfahren sind schneller und genauer.

- Vorteilhaft werden Monte-Carlo Integrationen für hohe Dimensionen (etwa $p > 6$) oder wenn die zu integrierende Funktion keine schöne Gestalt besitzt.

Aufgabe zur Monte-Carlo Integration

Berechnen Sie

$$\int_0^1 \int_0^1 \int_0^1 \exp(x + y + z) dx dy dz$$

mittels einer geeigneten Monte-Carlo Simulation!