Wahrscheinlichkeitstheorie für Informatikstudien (Kapitel #10)

Siegfried Hörmann

TU Graz

WS 2020/21

Kursaufbau

1. Wahrscheinlichkeitsräume.
2. Laplace-Experimente.
3. Wichtige diskrete Verteilungen.
4. Zufallsvariablen.
5. Zufallsvektoren.
6. Momente.
7. Bedingte Wahrscheinlichkeiten.
8. Unabhängigkeit.
9. Konvergenz von Zufallsvariablen.
10. Der zentrale Grenzwertsatz.

Normalverteilung und der zentrale Grenzwertsatz.

Konvergenz in Verteilung

Wir befassen uns jetzt mit einem weiteren Konzept der Konvergenz, das sich von den beiden anderen sehr unterscheidet.

Sei $(X_n : n \geq 1)$ eine Folge von ZVen mit Verteilungen F_n. Sei X eine ZV mit Verteilung F. Wir nehmen an, dass F stetig ist. Wir sagen, dass X_n zu X in Verteilung konvergiert falls

$$F_n(x) \to F(x) \quad \forall x.$$

Kurz: $X_n \xrightarrow{d} X$.

Beachte: Die Variablen X und X_n sind sich nicht notwendig nahe!

Konvergenz in Verteilung

Beispiel

Sei $(X_n : n \geq 0)$ eine Folge von i.i.d. ZVen. Zeige $X_n \xrightarrow{d} X_0$.

Beispiel

Seien X_i i.i.d. ZVen mit $X_i \sim \text{Exp}(\lambda)$ und man definiere $M_n = \max\{X_1, \ldots, X_n\}$ das Maximums einer Zufallsstichprobe. Zeige

$$Z_n := M_n - \frac{\log n}{\lambda} \xrightarrow{d} Z,$$

wobei Z Verteilung $G(x) := e^{-e^{-\lambda x}}$ besitzt. Man G *Gumbel* Verteilung und sie wird zur Modellierung von Extremereignissen verwendet.

Der zentrale Grenzwertsatz

Nun sind wir bereit, eines der wichtigsten Resultate in der Wahrscheinlichkeitstheorie zu formulieren.

Satz (Der zentrale Grenzwertsatz (CLT))

Sei (X_n) *eine Folge von i.i.d. ZVen mit* $E(X_n) = \mu$ *und* $\mathrm{Var}(X_n) = \sigma^2$*. Dann gilt für das empirische Mittel* $\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$*, dass*

$$\sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{d} N(0,1).$$

In anderen Worten: für $\forall a < b \in \mathbb{R}$ gilt

$$P\left(\sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \in (a,b]\right) \to \int_a^b \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Der zentrale Grenzwertsatz

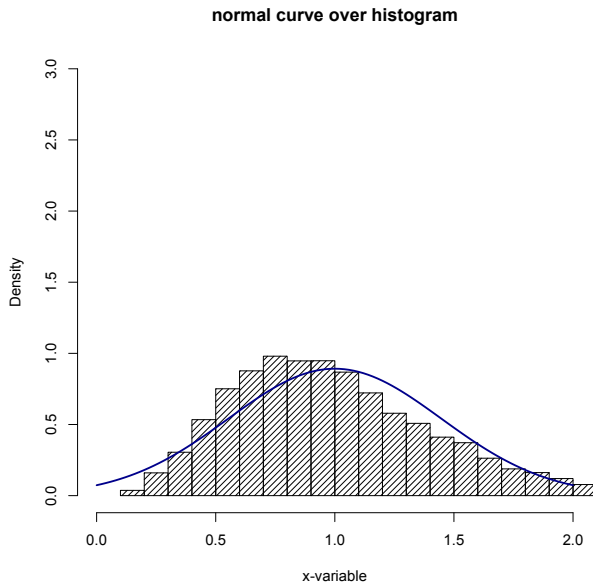
Zu Illustrationszwecken simulieren wir n ($n = 5, 10, 20, 50$) Exp(1)-verteilte ZVen und berechnen den Mittelwert. Wir wiederholen dieses Experiment 10000 mal und plotten das Histogramm dieser Mittelwerte. Man sieht dann sehr schön, wie das Histogramm mit wachsendem n immer mehr einer Normaldichte mit Parametern

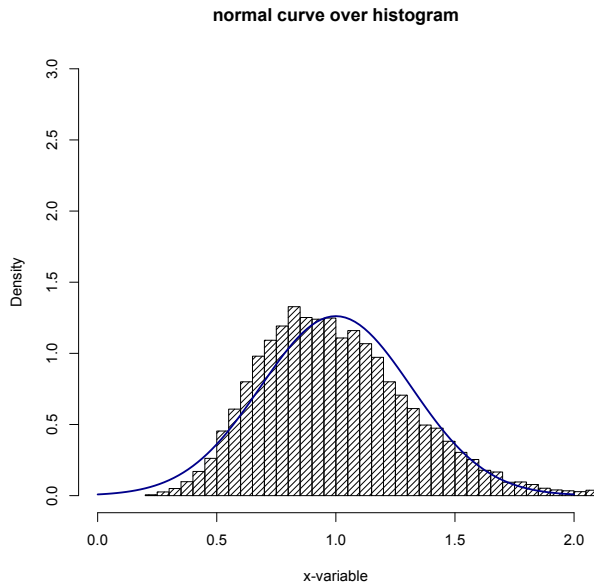
$$\mu = 1/\lambda = 1 \quad \text{und} \quad \sigma^2 = \frac{1}{n\lambda^2} = 1/n$$

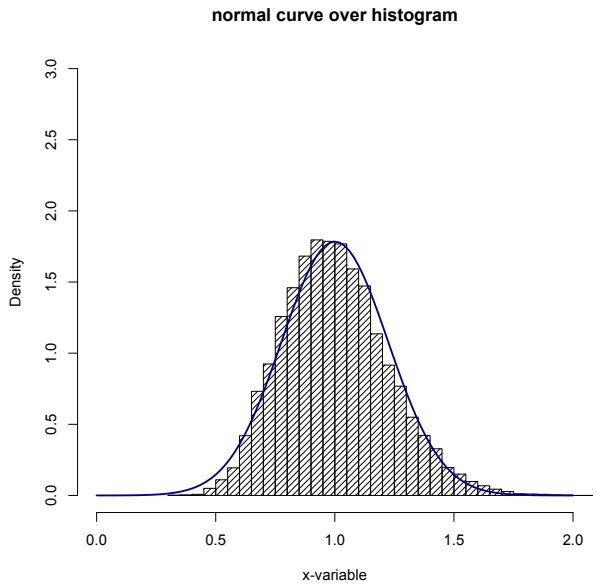
ähnelt.

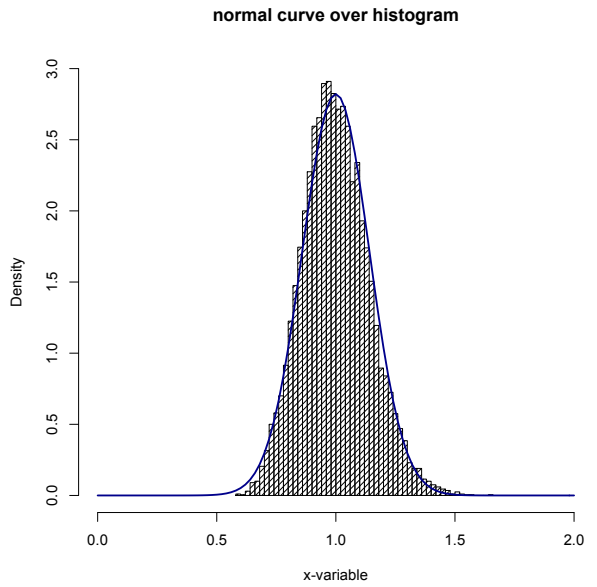
```
CLT = function(n){
  exp.means=c()
  for(i in 1:10000){ exp.means=c(mean(rexp(n,1)),exp.means) }
  exp.means
}
n=5
#n=10
#n=20
#n=50
g = CLT(n)
hist(g, density=20, breaks=40, prob=TRUE,
     xlab="x-variable", ylim=c(0, 3), xlim=c(0,2),
     main="normal curve over histogram")
curve(dnorm(x, mean=1, sd=sqrt(1/n)),
      col="darkblue", lwd=2, add=TRUE, yaxt="n")
```

Der zentrale Grenzwertsatz


normal curve over histogram


normal curve over histogram


normal curve over histogram

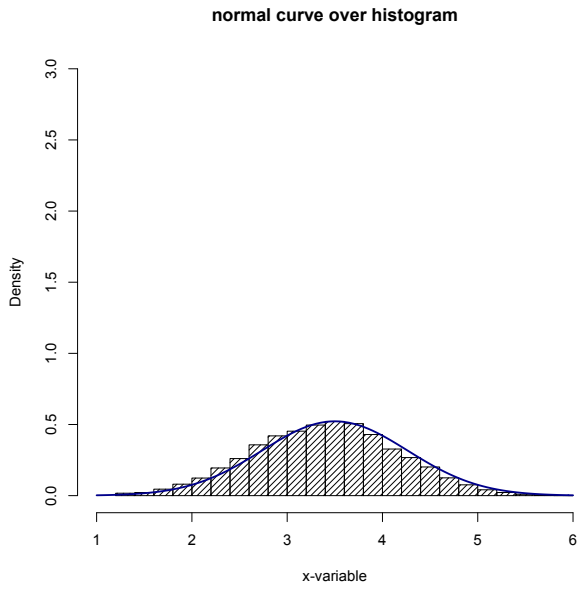

normal curve over histogram

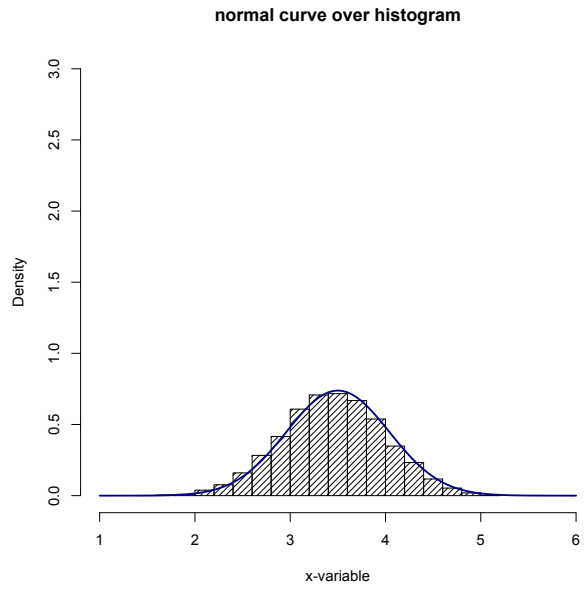
Der zentrale Grenzwertsatz

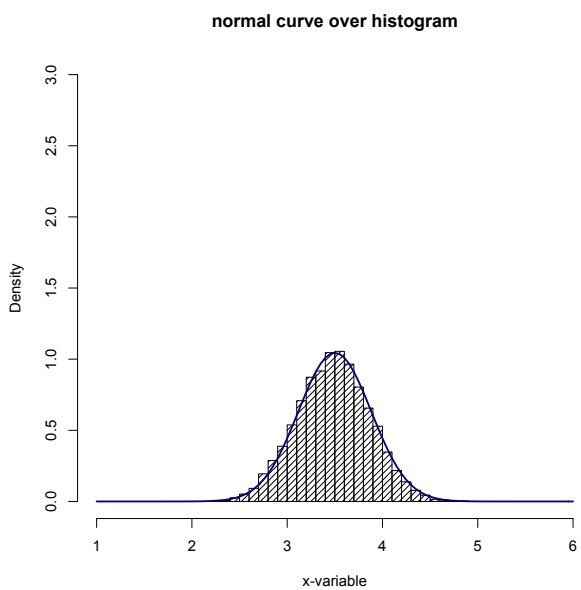
Noch besser ist die Anpassung bei einer symmetrischen Verteilung, wie z.B. die diskrete Gleichverteilung. Hier die Verteilung vom Mittelwert von n Unif($\{1, \ldots 6\}$)-verteilten ZVen mit $n = 5, 10, 20, 50$:

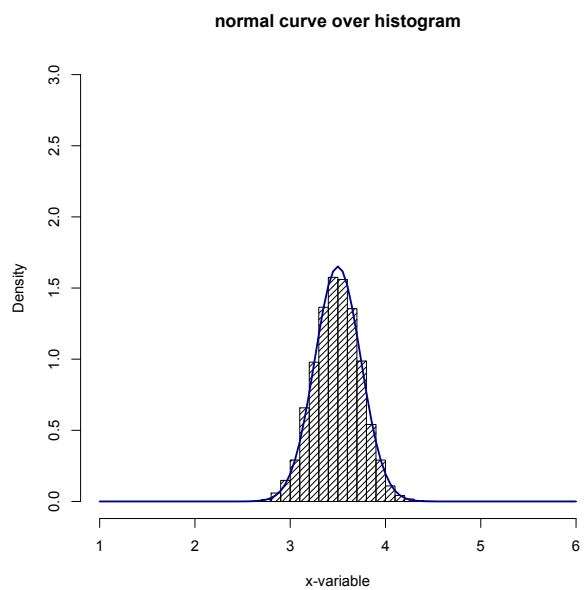
```
CLT = function(n){
  dice.means=c()
  for(i in 1:10000){ dice.means=c(mean(sample(1:6,n,replace=TRUE)),dice.means) }
  dice.means
}
n=5
#n=10
#n=20
#n=50
g = CLT(n)
hist(g, density=20, breaks=40, prob=TRUE,
     xlab="x-variable", ylim=c(0, 3), xlim=c(0,2),
     main="normal curve over histogram")
curve(dnorm(x, mean=3.5, sd=sqrt(35/12/n)),
      col="darkblue", lwd=2, add=TRUE, yaxt="n")
```

Der zentrale Grenzwertsatz









Der zentrale Grenzwertsatz

Wie interpretieren wir dieses Resultat?

Das Gesetz der großen Zahlen sagt uns: für X_k i.i.d. mit Erwartungswert μ und Varianz σ^2 gilt

$$\overline{X}_n := \frac{1}{n}(X_1 + \cdots + X_n) \xrightarrow{P} \mu.$$

Wird n größer, so sollte $\overline{X}_n$ nahe μ sein.

Aber wie nah sind sich $\overline{X}_n$ und μ für ein grosses aber festes n?

Der zentrale Grenzwertsatz

Seien X_i iid mit $E(X_i) = \mu$ und $\mathrm{Var}(X_i) = \sigma^2$. Bestimme mit dem CLT approximativ t derart, dass

$$P\Big(|\overline{X}_n - \mu| \leq t\Big) = 0.95.$$

Beispiel

Wir ziehen 100 mal mit Zurücklegen aus einer Urne mit schwarzen und weißen Kugeln. Wir kennen allerdings nicht den Anteil θ der schwarzen Kugeln. Um diesen zu schätzen nehmen wir die relative Häufigkeit $\overline{X}_n$, wobei $X_i = 1$ ist, falls die i-te Kugel schwarz war und $X_i = 0$ falls nicht. Bestimme

$$P\Big(|\overline{X}_n - \theta| \leq t\Big).$$

Der zentrale Grenzwertsatz

Eine andere Möglichkeit den CLT zu formulieren ist diese. Sei $S_n = X_1 + \cdots + X_n$, dann gilt

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} = \frac{S_n - n\mu}{\sqrt{n}\sigma}.$$

Daher gilt:

$$\begin{aligned} P(S_n \leq k) &= P(S_n - n\mu \leq k - n\mu) \\ &= P\Big(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{k - n\mu}{\sigma\sqrt{n}}\Big) \\ &\approx \Phi\left(\frac{k - n\mu}{\sigma\sqrt{n}}\right). \end{aligned}$$

Der zentrale Grenzwertsatz

Beispiel

Eine Dame behauptet Hellseherin zu sein. Wir wollen diese Aussage testen. Dazu werfen wir 100 mal eine Münze und bitten die "Hellseherin" vorherzusagen, ob wir Kopf oder Zahl werfen. Von den 100 Würfen hat sie 65 richtig vorhergesagt. Sei nun S_n die Anzahl der korrekten Vorhersagen in n Versuchen. Bestimmen $P(S_{100} \geq 65)$, falls sie nur geraten hat.

Der zentrale Grenzwertsatz

Dieses Beispiel lässt sich auch ohne die Normalapproximation lösen!

Eigentlich gilt unter der Hypothese dass sie nur rät, dass die Anzahl der richtigen Antworten S_{100} einer Binomialverteilung $B_{100,1/2}$ folgt.

$$P(S_{100} \geq 65) = \sum_{k=65}^{100} \binom{100}{k} 1/2^{100}.$$

Der zentrale Grenzwertsatz

Beispiel

Wir gehen ins Casino und spielen Roulette. Wir setzen jedes Mal 1 € auf rot und spielen n mal. Wie groß ist die Wahrscheinlichkeit, dass wir am Ende Geld gewonnen haben?

Beweisidee für den CLT

Zum Abschluss der VL noch die wesentliche Beweisidee.

1. Wir nehmen an, dass $E(X_i) = 0$ und $E(X_i^2) = 1$. Sonst führe den Beweis für die standardisierte Variable $Z_i = \frac{X_i - \mu}{\sigma}$.
2. Wir haben im Kapitel 8 die momenterzeugende Funktion (MEF) definiert und rechnen nun die MEF von

$$\frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

aus. Wir zeigen, dass die MEF von $\frac{X_1 + \cdots + X_n}{\sqrt{n}}$ gegen die MEF einer Standardnormalverteilung konvergiert. Die MEF der Standardnormalverteilung lautet $e^{t^2/2}$.

Beweisidee für den CLT

Nun gilt

$$
\begin{aligned}
M_{\frac{X_1+\cdots+X_n}{\sqrt{n}}}(t) &= E(e^{t\,\frac{X_1+\cdots+X_n}{\sqrt{n}}}) \\
&= \prod_{i=1}^{n} E(e^{t\,\frac{X_i}{\sqrt{n}}}) \quad [X_i \quad \text{unabhängig}] \\
&= \left(M_{X_1}(t/\sqrt{n})\right)^n \quad [X_i \quad \text{identisch verteilt}] \\
&= \Big(\underbrace{E(1+\frac{t}{\sqrt{n}}X_1+\frac{t^2}{2n}X_1^2+\mathit{Rest})}_{\text{Taylorentwichklung von } e^{tX_1/\sqrt{n}}}\Big)^n \\
&\approx \left(1+\frac{t^2}{2n}\right)^n \to e^{t^2/2}. \quad \square
\end{aligned}
$$