

Ich verstehe nicht, was ist der Unterschied zwischen Borel-Sigmaalgebra und Sigmaalgebra. Und in welcher Situation benutzen wir Borel-Sigmaalgebra?

Die Borel-Sigmaalgebra ist eine spezielle Sigmaalgebra und zwar die, die vom Mengensystem der halboffenen Intervalle erzeugt wird. In der Borel-Sigmaalgebra sind z.B. Intervalle wie $[7,8]$ drin, aber auch Mengen wie $\{-5,4,13\}$ oder Vereinigungen von Intervallen. Wir benutzen die Borel-Sigmaalgebra als häufigste Sigmaalgebra über den reellen Zahlen. Wichtig zu wissen ist noch, dass die Borel-Sigmaalgebra nicht alle Teilmengen der reellen Zahlen enthält, also kleiner als die Potenzmenge ist.

Wofür wird die t-Verteilung verwendet, wenn sie sich für größer werdende t's eh an die Normalverteilung annähert?

Die t-Verteilung wird z.B. für geringen Stichprobenumfänge ($n < 30$) beim t-Test benötigt. Sie gibt dort die Verteilung der Prüfgröße unter der Nullhypothese an (kommt alles in Induktive Statistik). Für kleine Stichprobenumfänge unterscheidet sich die t-Verteilung deutlich von der Std.normalverteilung.

Wie kann man prüfen ob die Verteilungsannahme für eine ZV gerechtfertigt ist?

Das macht man mit Verteilungstests. Die lernen wir in induktiver Statistik.

Ich verstehe noch nicht ganz wie man die Quantile und die Zufallszahlen der in R berechnet und ob es dafür Merkwörter gibt um sie zu berechnen ohne im Internet nach Hilfe zu suchen .

Das geht mit den Kürzeln `q` (für Quantil) und `r` (für random number), z.B. `qnorm(0.99)` für das 99%-Quantil der Standardnormalverteilung und `rnorm(100)` zur Generierung von 100 Zufallszahlen.

Außerdem habe ich bei Normalverteilungen noch kein Gefühl dafür wie man die Wahrscheinlichkeiten für normalverteilte Zufallsvariablen bestimmt, da man hier die Dichtefunktion ja nicht integrieren darf.

Die Dichtefunktion darf man schon integrieren, aber man findet keine Stammfunktion. Wahrscheinlichkeiten berechnet man am besten mit R, z.B. so:

$$N_{\{\mu, \sigma^2\}}(10;15] = \text{pnorm}(15,\mu,\sigma) - \text{pnorm}(10,\mu,\sigma)$$

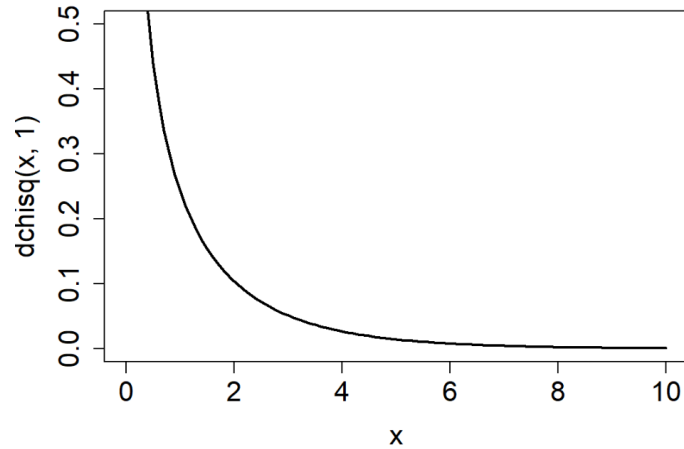
Wie ist es möglich, dass Zufallszahlen so einfach durch eine Uniformverteilung ($U[0,1]$) erzeugt werden können, obwohl in Vorlesungen oft betont wird, wie schwierig und kostspielig es ist, wirklich zufällige Zahlen zu generieren?

Sie dürfen bitte zwei Dinge nicht verwechseln:

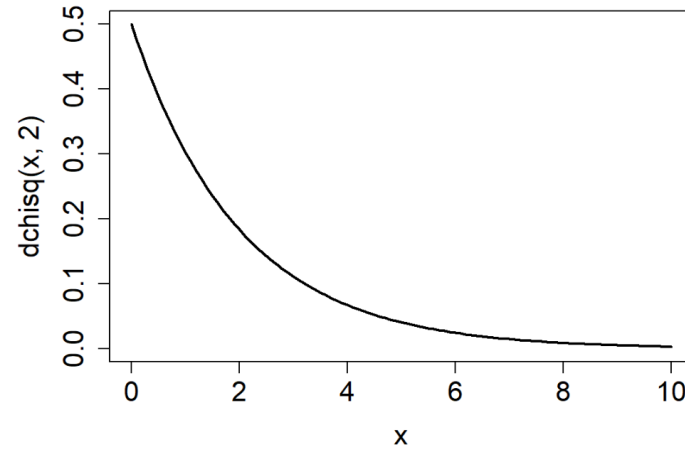
- 1) Wenn man $U[0,1]$ verteilte Zufallszahlen hat, kann man (über die Quantilfunktion) recht leicht Zufallszahlen beliebiger Verteilungen erzeugen. Das ist im Buch gemeint.
- 2) Die Erzeugung von echten Zufallszahlen ist schwierig, auch von $U[0,1]$ -verteilten Zufallszahlen. Für nahezu alle praktische Zwecke reichen aber die im Computer erzeugten Zufallszahlen.

Wie beeinflusst die Anzahl der Freiheitsgrade ν die Form der χ^2 -Dichtefunktion, und warum nähert sich die χ^2 -Verteilung mit zunehmenden Freiheitsgraden der Normalverteilung an?

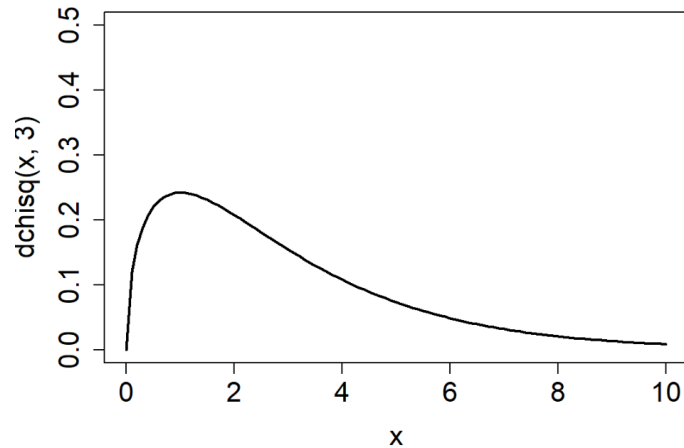
Chisq_1



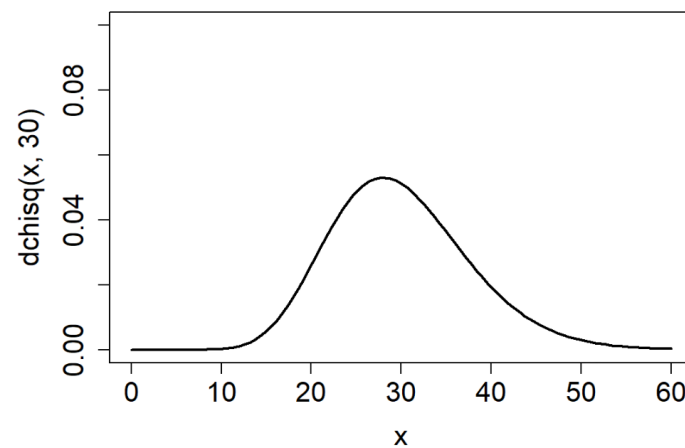
Chisq_2



Chisq_3



Chisq_30



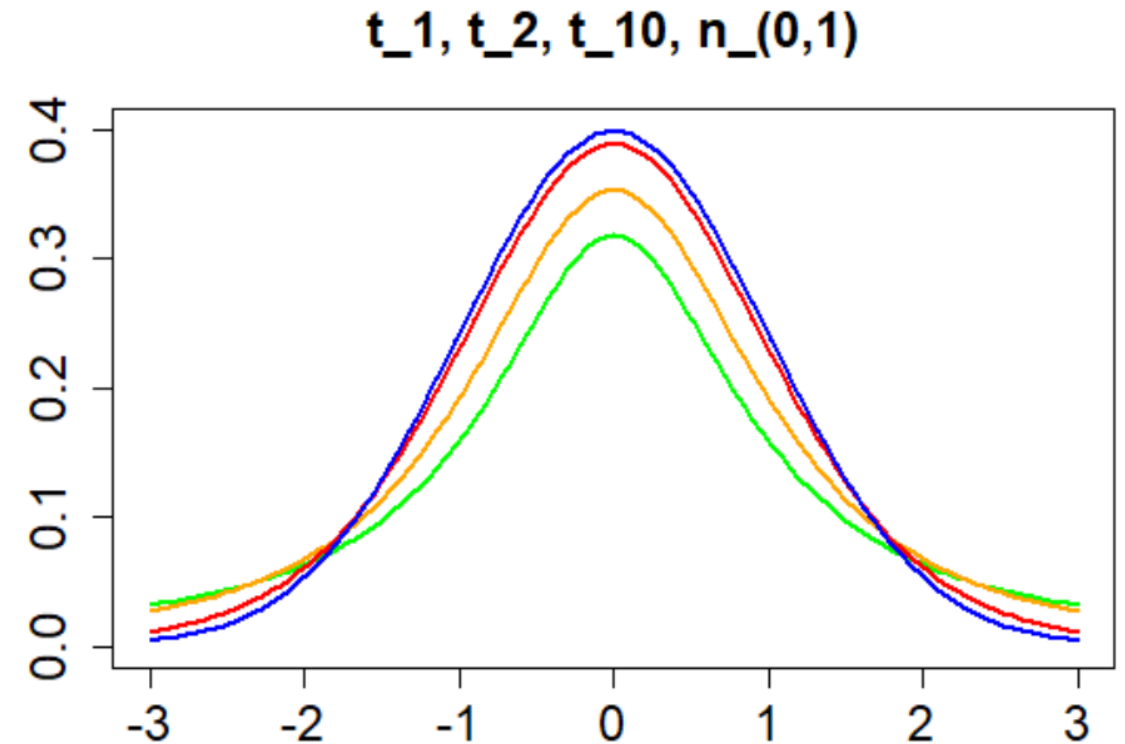
Das liegt am ZGWS.

Worin liegt der Unterschied zw. Normalvtlg. und t-vtlg.?

Die Graphen der Dichtefunktion der Standardnormalverteilung und der t-Verteilung schauen beide aus wie eine Glocke, die symmetrisch zur y-Achse ist. Allerdings sind die Ränder der t_n -Dichte weiter oben und dafür die Mitte der Glocke niedriger als die der Standardnormalverteilung und zwar um so weiter weg von der Std.normalverteilung, je kleiner der Freiheitsgrad n ist.

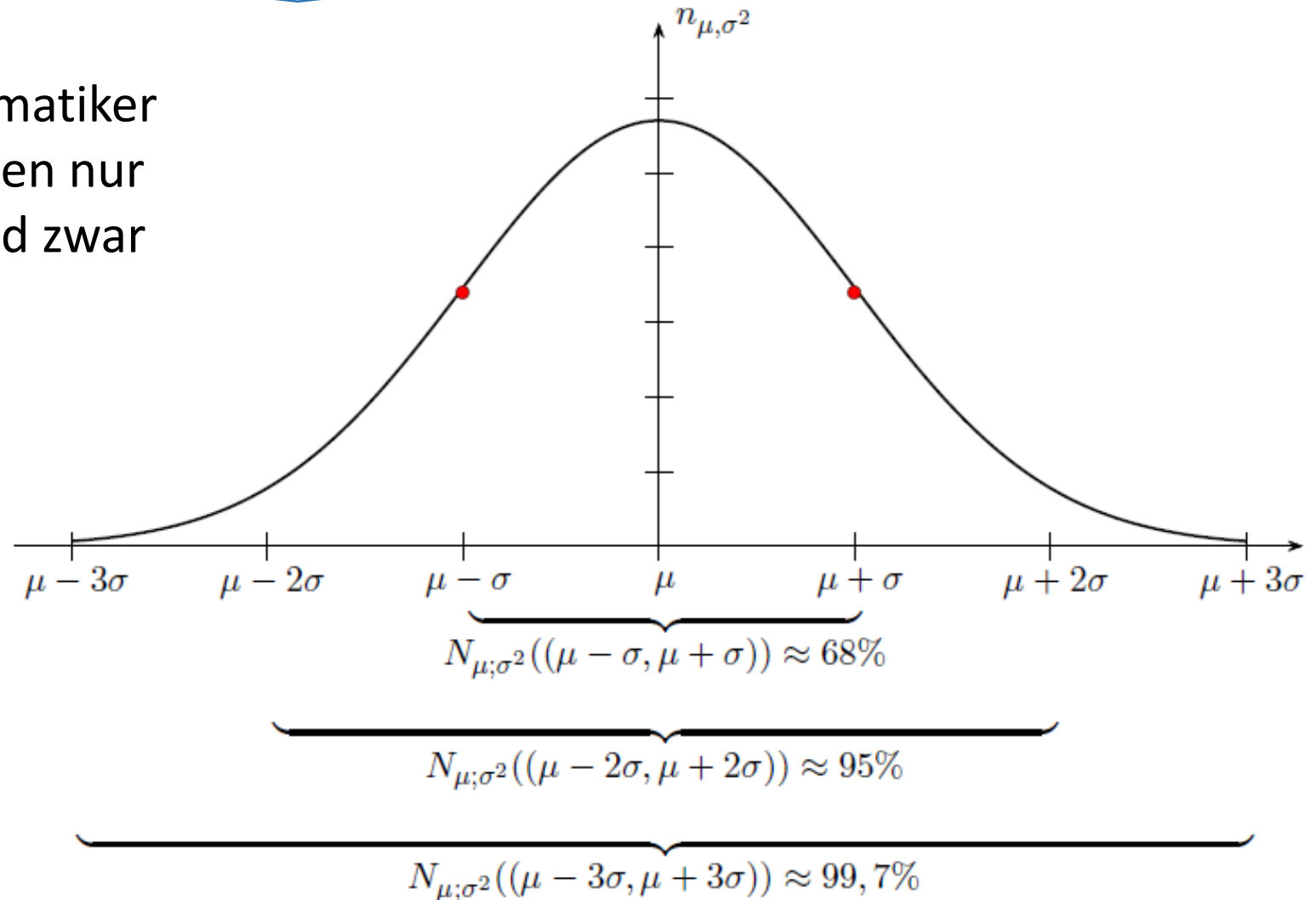
Das bedeutet, dass „extremere“ Werte (z.B. -4 oder 5) unter der t-Verteilung eine höhere Dichte besitzen.

Für $n \rightarrow \infty$ nähert sich die Dichte der t_n -Verteilung der Dichte der Standardnormalverteilung an.



Wie komme ich auf Schätzwerte bei der Normalverteilung?

Das haben schlaue Mathematiker für Sie bestimmt. Sie müssen nur noch auswendig lernen und zwar das:



Wie funktioniert die Umwandlung einer nicht standardnormalverteilte Zufallsvariablen ?

Sei $X \sim N_{\mu; \sigma^2}$:

Standardisierung einer ZUV bedeutet, dass wir den Erwartungswert auf 0 und die Varianz auf 1 bekommen wollen. Ein Satz aus der Statistik besagt, dass $a(X+b)$ wieder normalverteilt ist und zwar mit Erwartungswert $E[X]=a(\mu+b)$ und $\text{Var}[X]=a^2 \sigma^2$.

Daher muss man $b := -\mu$ und $a := 1/\sigma$ wählen.

Insgesamt ist also

$(X - \mu)/\sigma$ standardnormalverteilt.

Hier fehlt mir auch das Verständnis, was die Abbildung 6.12 aussagt...

6.3 Normalverteilung

175

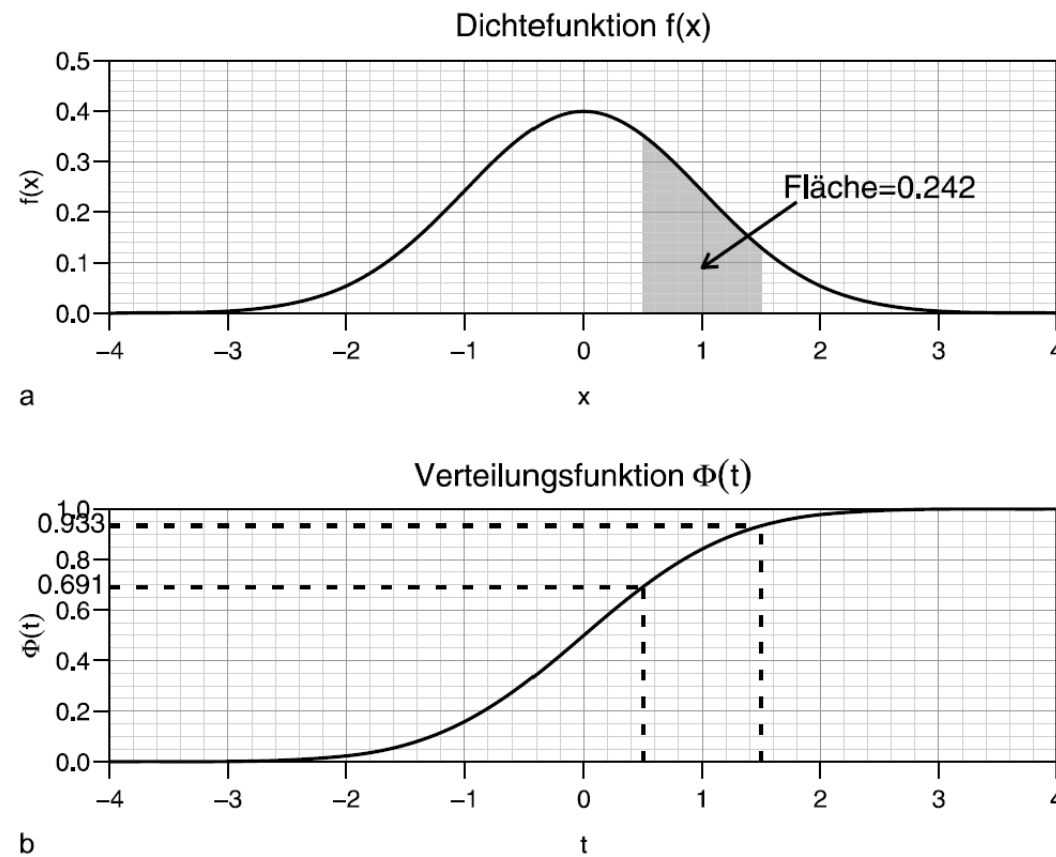


Abb. 6.12 Berechnung von $P(0.5 < X < 1.5)$ mit Hilfe der Verteilungsfunktion der Standardnormalverteilung

Mir ist es teilweise noch nicht ganz klar, wie man für ein Problem entscheidet was für eine Verteilung ein sinnvolles Modell ist. Im Buch wird ja als Beispiel für die Exponentialverteilung die „Lebensdauer eines Bauteils“ genommen.

Wieso ist das aber nicht (auch) Normalverteilt? Also nach dem Prinzip: die erwartete Lebensdauer ist z.B. 8 Jahre mit Varianz $\sigma^2 = 4$, manche halten also (deutlich) kürzer & manche (deutlich) länger.

Zur Modellierung braucht es Wissen, Übung und Erfahrung. Genau das versuche ich Ihnen über das Skript, die ABCs, die Moodle-Tests und die Übungen beizubringen.

Zum konkreten Beispiel: Die Normalverteilung ist allein schon deshalb nicht gut für Wartezeiten geeignet, weil bei der N.vtlg. auch negative Werte eine Wsk. haben.

Kann man immer einer Aufgabe, in der man eine Wahrscheinlichkeit berechnen soll, genau einer der Verteilungen zuordnen?

Nein.

Weshalb sind die cdf's der Chiquadrat und t-Verteilung laut Buch nicht integrierbar? Aufgrund der komplizierten Verteilungsfunktionen?

Die Verteilungsfunktionen existieren, aber sie lassen sich – genau wie bei der Normalverteilung – nicht in analytischer Form hinschreiben.

Könnten Sie bitte nochmal den Unterschied zwischen einer Dichtefunktion und einer Verteilungsfunktion erklären.

$$F(x) = \int_{-\infty}^x f(t) dt$$

Warum berechnet man die Verteilungsfunktion von
z.B. $P(a < x < b)$ [mit $a < b$] mit dem Integral von a bis b
und nicht immer, wie bei der Standard
Normalverteilung mit $F(b) - F(a)$?

Weil es oft einfacher ist, mit der Dichte statt der Verteilungsfunktion zu rechnen.

Ist die Exponentialverteilung immer gedächtnislos?

Ja.

Sollen wir die Dichten der einzelnen Verteilungen auch in der Klausur rechnen können?

Ja.

Wie unterscheide ich sofort
Wahrscheinlichkeitsfunktion und Verteilungsfunktion?

$$F(x) = \sum_{x_i \leq x} f(x_i)$$

Wie rechne ich komplizierte Erwartungswerte und Varianzen aus? Mit Integralen?

Bei kontinuierlichen Zufallsvariablen mit Integralen,
bei diskreten Zufallsvariablen mit Summen.

Wie komme ich von dem O mit dem Strich in der Mitte und (0.5) dahinter auf das Ergebnis 0.691? Also wie komme ich rechnerisch auf diesen Wert?

Gar nicht, weil es keine Stammfunktion zur Dichte der Std.normalverteilung gibt. Auf den Wert kommt man z.B. mit R durch den Befehl `pnorm(0.5)` oder durch Tabellen (früher).

Ich habe mir rausgeschrieben wann man Dichtefunktionen integrieren kann oder nicht, aber wie die Varianz mit der Dichtefunktion zusammenhängt ist mir noch nicht ganz klar.

$$\text{Var}[X] = E[X^2] - E[X]^2 = \int x^2 f(x) dx - \left(\int x f(x) dx \right)^2$$

Erhalten wir für die Klausur die Tabellen um die Wahrscheinlichkeit der Standardnormalverteilung abzulesen oder muss hier kein genaues Ergebnis in Zahlen angegeben werden?

Entweder Sie müssen die R-Befehle angeben oder ich drucke ein paar R-Befehle mit Ergebnissen mit auf die Klausur oder ich drucke eine Ausschnitt einer Tabelle auf die Klausur oder Sie müssen mit den Sigma-Bereichen arbeiten.

Was ist der Unterschied zwischen kontinuierlicher und stetiger Verteilung?

Das ist das gleiche.