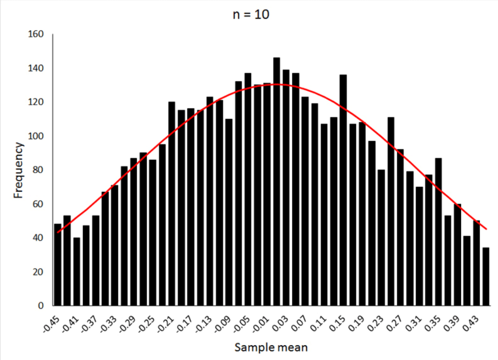
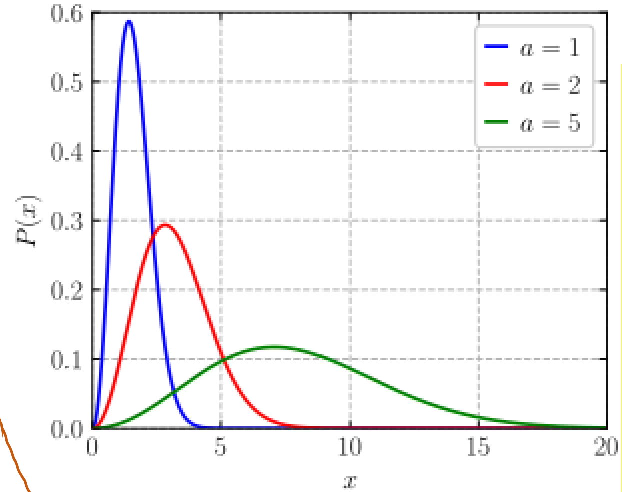
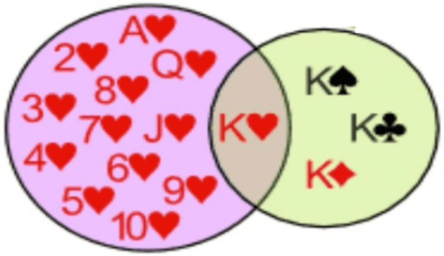
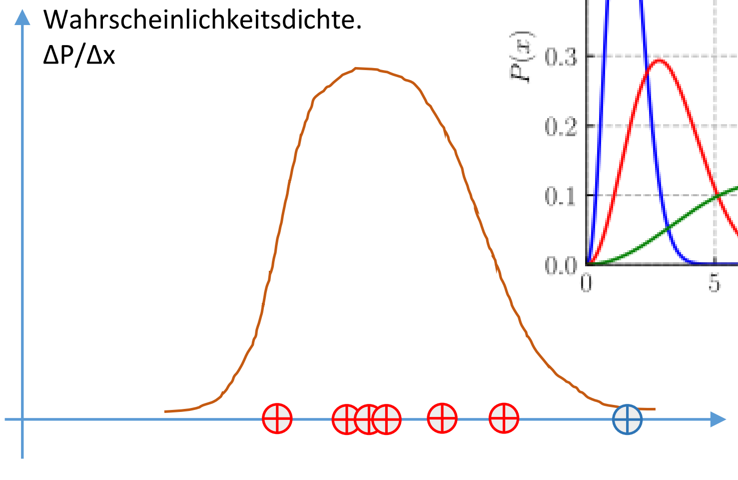
Ereigniss, Wahrscheinlichkeit, Verteilung.

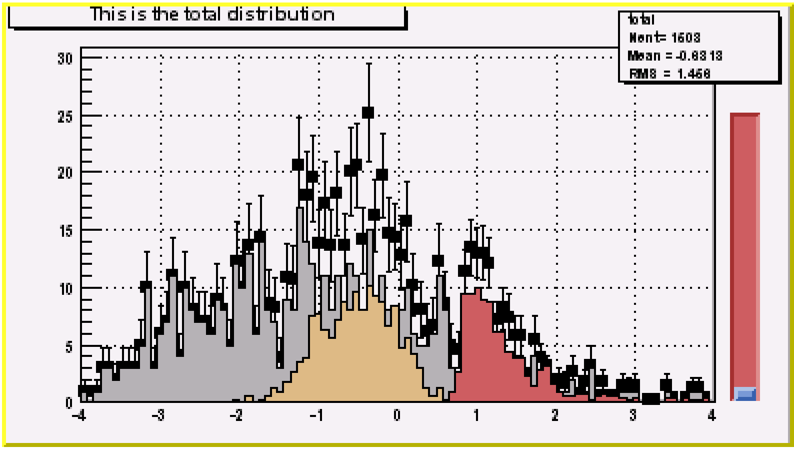
G. Schay

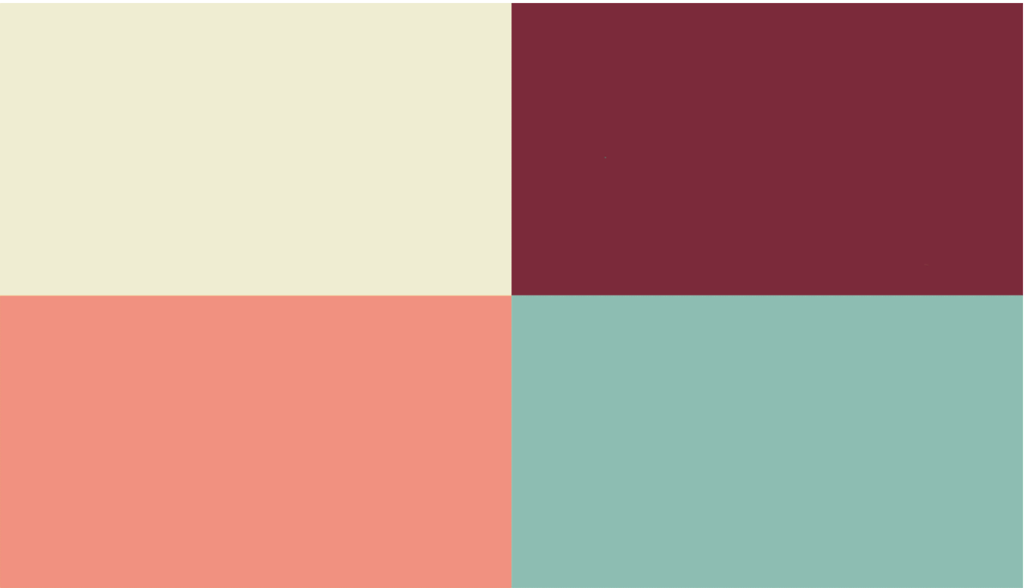


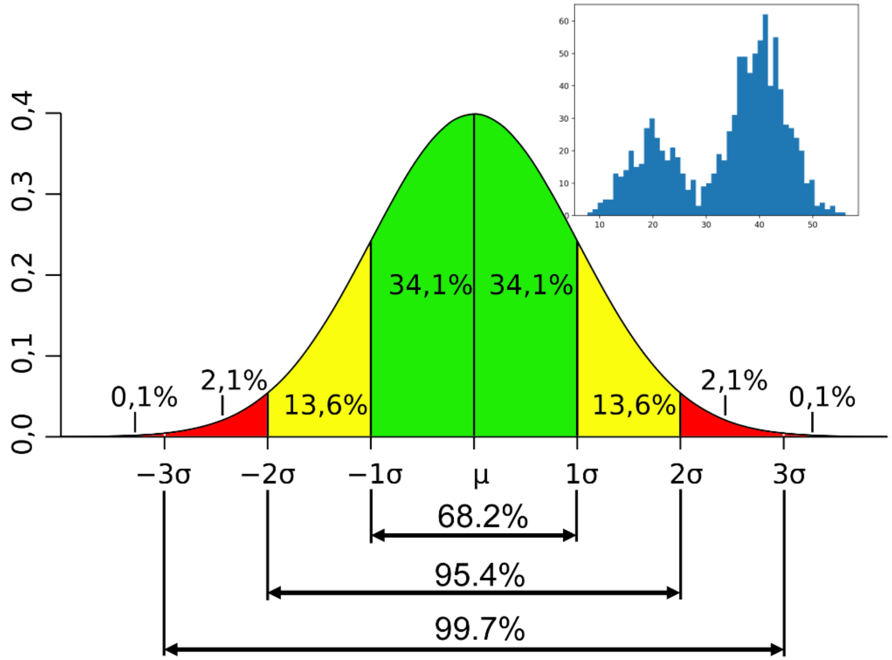












Ereigniss:

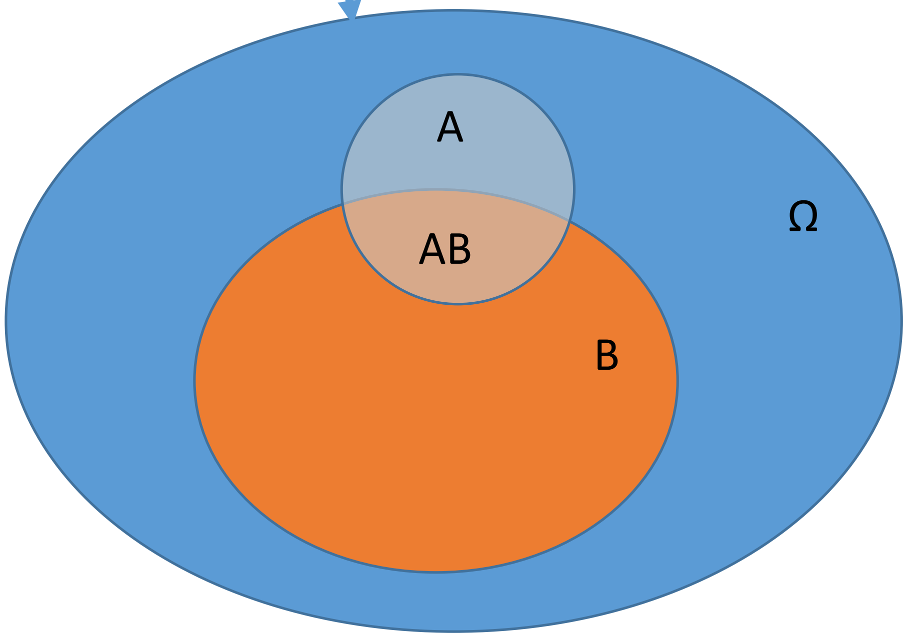
was wir beobachten können.

wir können immer genau bestimmen ob ein Ereigniss stattgefunden hat oder nicht!

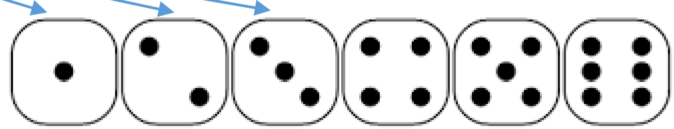
© Dreamworks: The Road to El Dorado – Zaragoza's dice

Ereignissraum Ereignisse in dem Raum.

Elementarereignisse
oder
zusammengesetzte E.



Wenn wir alle Elementarereignisse kennen, dann haben wir den ganzen Ereignissraum abgedeckt



$= \mathbf{\Omega}$

Ereignisse haben einen **Frequenz oder Häufigkeit**

aus N Versuche wie viel mal etwas passiert.
z.B.

-> 5 Mal Würfel werfen, darunter wie viele “6”-er.
-> wie oft müssen wir operieren in einem Nachtschift im Krankenhaus.

Anzahl der Vorkommen aus N versuche ist die absolute Häufigkeit

oft wir benutzen k als symbol, aber auch n, etc.
also k_A ist der Frequenz von Ereigniss A während der Versuchsreihe.

Relative Häufigkeit

k_A/N (auch in % möglich)

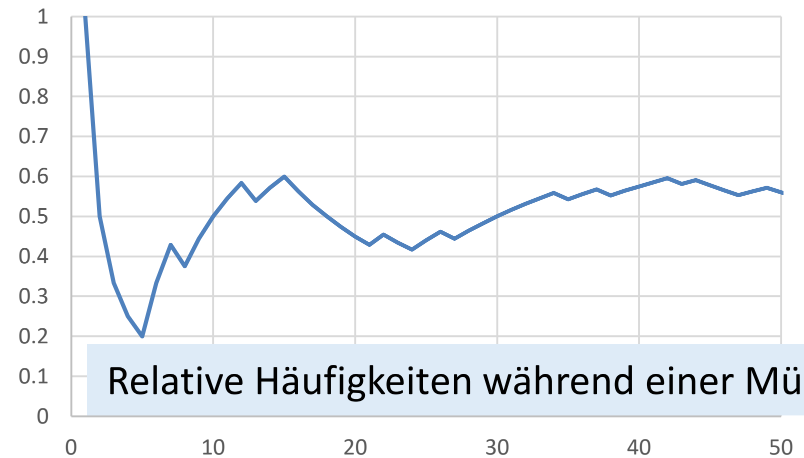
Regel von großen Zahlen

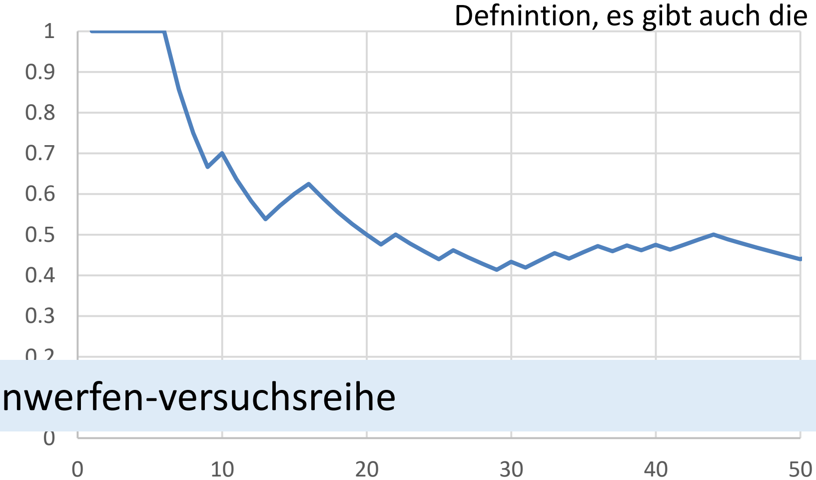
Der relativer Frequenz ist nicht immer ganz gleich wenn wir mehrere Versuchsserien vergleichen.

Aber wird um so stabiler je mehr Versuche wir haben.

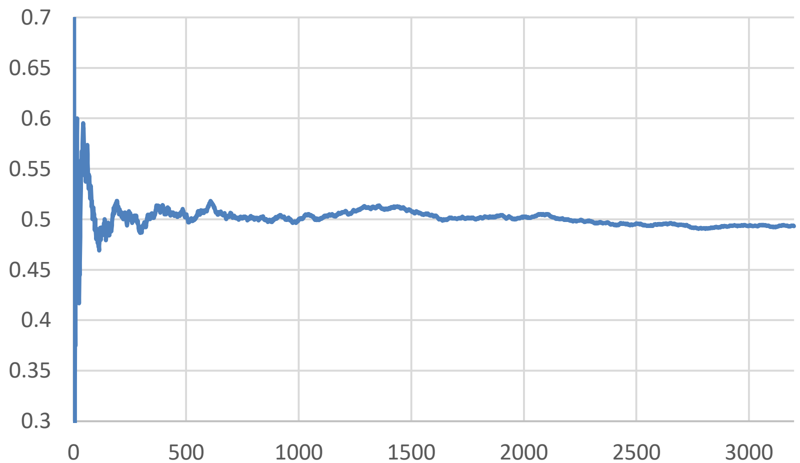
Wahrscheinlichekit ist der Zahl um welchen die relative Häufigkeiten* sich stabilisieren.

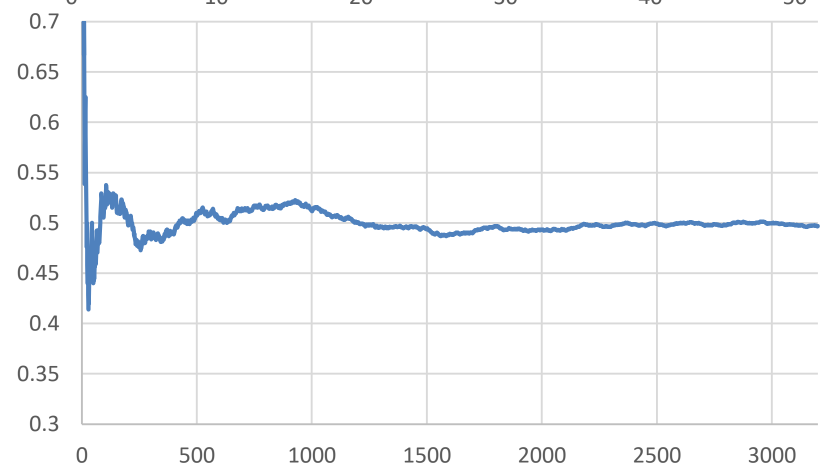
*: Das ist die Frequenzbestimmte Defnintion, es gibt auch die Bayes-methode



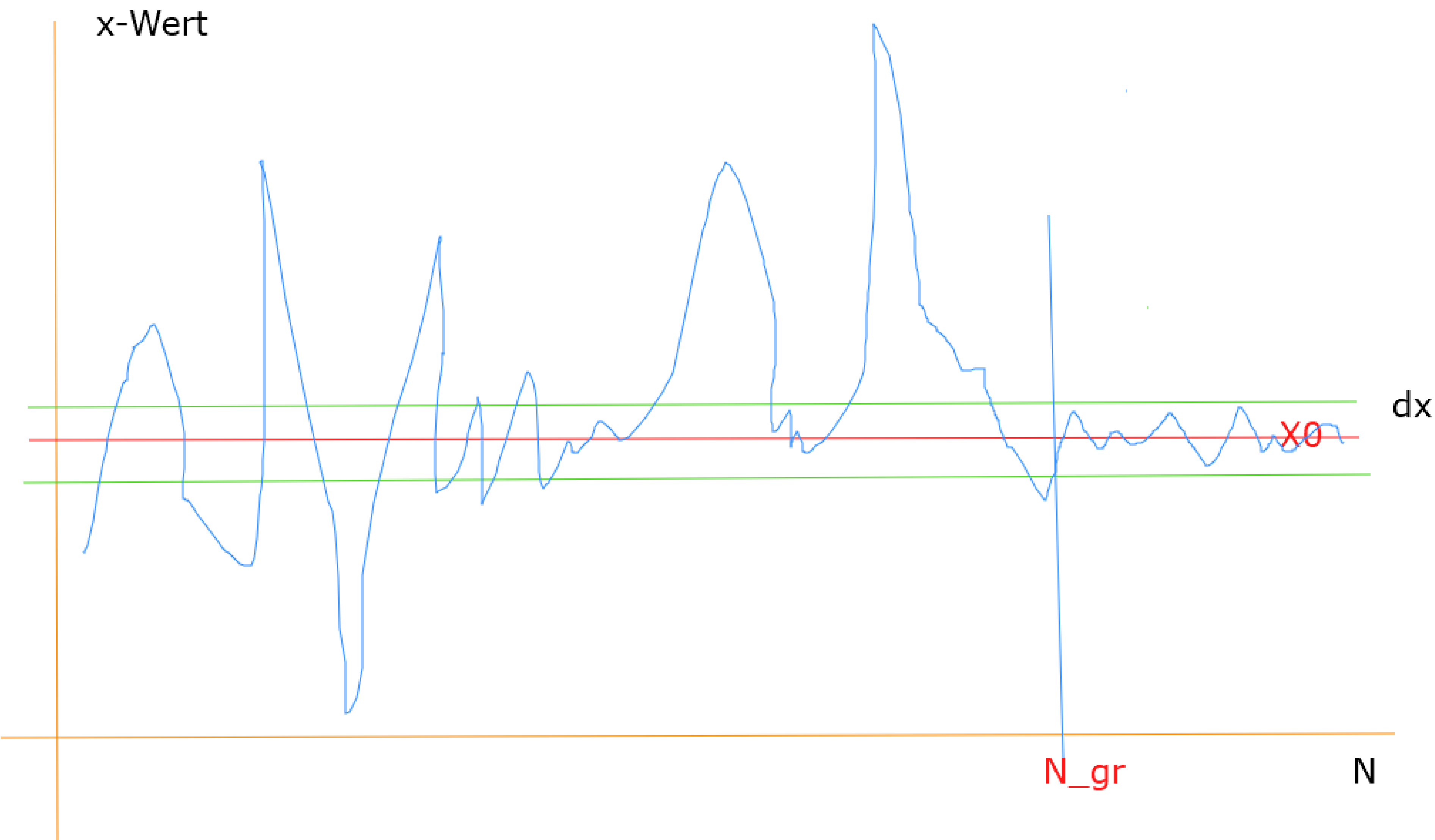


Relative Häufigkeiten während einer Münzenwerfen-versuchsreihe



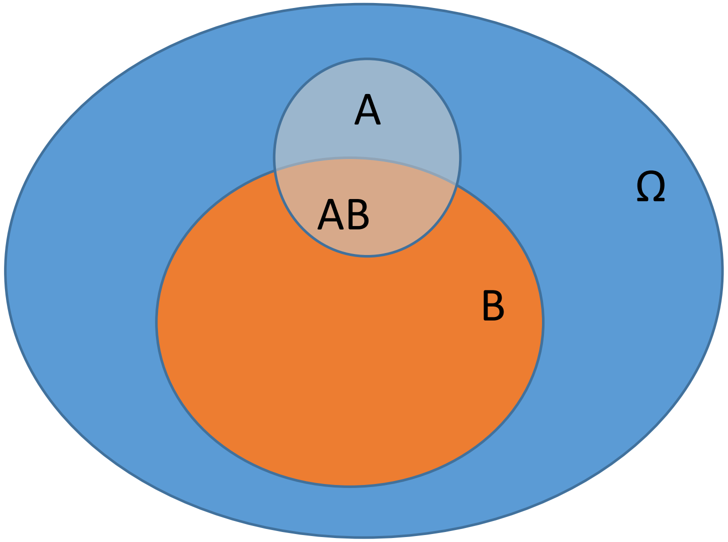


Ergänzungsmaterial: P ist eigentlich mit limes (Grenzwert) definiert, also für irgeneinen dx>0 es ist möglich einen N_gr Schwellenwert zu geben so dass wenn N>N_gr dann ist X-X0 <= dx. hier X ist die rel. Häufigkeit



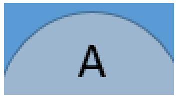
Ereignisse und ihre Zusammenhänge

Ereigniss ist etwas was wir beobachten können (hängt von uns ab was das sei)

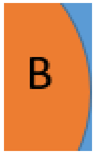


zusammengesetzte, oder Kombinationsereignisse
- wir haben einen geraden Zahl geworfen
- wir sehen eine Veränderung in der Blutprobe

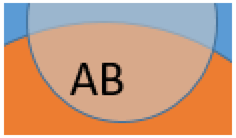
Verhältnisse



oder

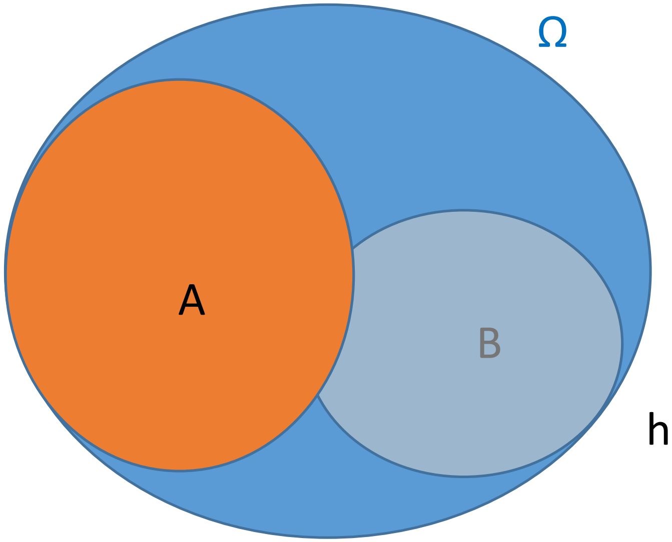


und



sich gegenseitig ausschließende Ereignisse

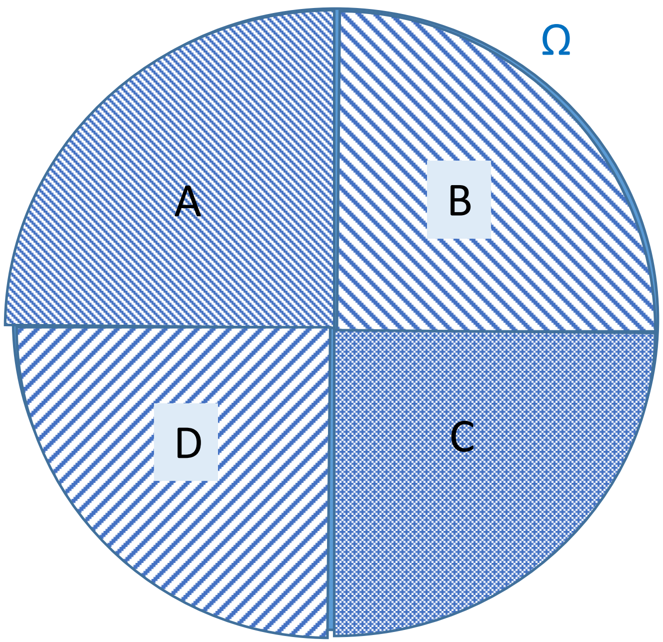
Gegenereginiss: Nicht-A



sich gegenseitig ausschließende Ereignisse

entweder oder

hier P(A oder B) = P(A)+P(B)



Elementarereignisse:
Es ist möglich den ganzen Ereignissraum abzudecken, und diese sind nicht mehr zu spalten.

übliche Symbologie

UND	P(A und B) = $P(A*B) = P(A \cap B) = P(A \cdot B) = P(AB)$
ODER	P(A oder B) = $P(A + B) = P(A \cup B) = P(A \vee B)$
NICHT	P(nicht A) = $P(!A) = P(\sim A) = P(\bar{A})$

Kolmogorovsche Axiome:

$0 <= P(A) <= 1$

$P(A \cdot \bar{A}) = P(\emptyset) = 0$ unmöglichkeit

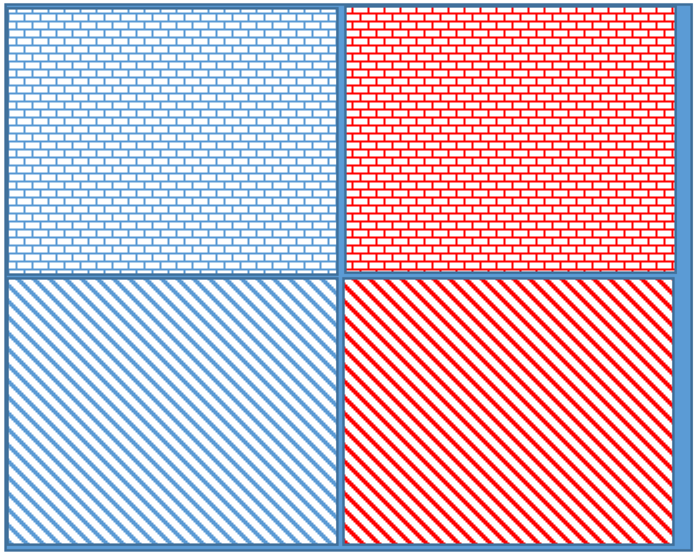
$P(A+\bar{A})=P(\Omega) = 1$ sicherheit

$\emptyset$: leerheit

Ω : alles

Stochastische Unabhängigkeit

zwei Ereignisse sind unabhängig wenn die gegenseitig keinen Einfluss auf einander (an die Häufigkeiten) haben.



P(gitter) = 2/4 = 0.5
P(schraffiert) = 2/4 = 0.5
P(blau) = 2/4 = 0.5
P(rot) = 2/4 = 0.5

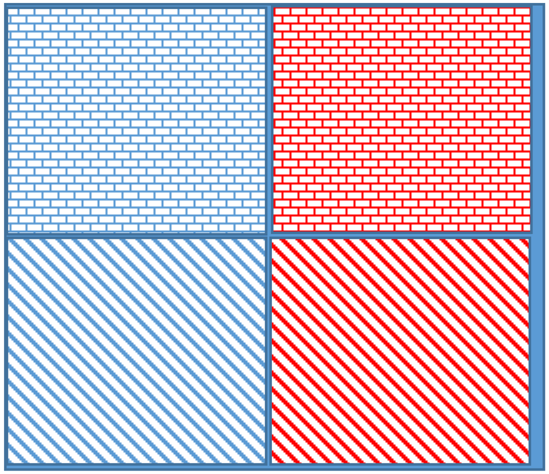
P(gitter und rot) = ¼ = ½ * ½ = P(gitter)*P(rot)

Zwei Ereignisse A and B sind unabhängig nur wenn P(A und B) = P(A)*P(B), und wenn es wahr ist dann sind A und B unabhängig.

Bedingte Wahrscheinlichkeit

P(A | B) = wahrscheinlichkeit von A ***wenn*** die Bedingung B ist wahr.

z.B. : unser Patient(inn) hat Fieber *wenn* Er/Sie COVID-19 angesteckt ist.
ich bekomme 5 in Statistik *wenn* ich alle Vorlesungen verstanden habe.



Wir sind nur an einen Unterraum von Ω interessiert.

P(blau | gitter) = wie viele blaue wir unter gitterte haben
= 1 blau UND gitter / 2 gitter = ½

Allgemein: P(A|B) = P(A·B)/P(B)

Bemerkungen:
1. für unabhängige: P(A|B=P(A)
2. für irgendwelche A,B: P(AB) = P(A|B)*P(B) Bayes-Regel oder Multiplikationsregel

-> mehr in Bayes-Methoden Vorlesung

Beispiele:

- P diabetes (B) sei 15%. die Wahrscheinlichkeit von Typ II. (A) sei 10%.

dann die Wahrscheinlichkeit das jemand Typ II hat WENN diabeteskrank ist

$P(A|B) = P(AB)/P(B) = P(A)/P(B) = 10\%/15\% = 66.7\%$

Bemerkung: wenn A ist ein Teil von B dann wenn A passiert, dann B ist automatisch auch passiert, also $P(AB)=P(A)$.

- unabhängigkeit: was ist die Wahrscheinlichkeit von Brilletragen (A), WENN die Patientin davor war eine Frau (B)?

$P(A|B) = \mathit{P(AB)}/P(B)$, aber unabhängig, also = $\mathit{P(A)*P(B)}$ / $P(B) = P(A)$

wir erwarten aus Vernunft dass B sollte gar nicht erscheinen, und gerade das haben wir bekommen. ☺

Ergänzungsmaterial

es gibt andere Parameter, die man aus P herleiten kann.

Odds: $O = \frac{P}{1-P}$

wenn Ω dichotomisch ist (also entweder etwas passiert oder nicht) dann ist Odds vorteilhaft: wie öfter passiert etwas als nicht passiert?

Risiko: Hier haben wir ZWEI Ereignisse, der eine (Risiko) vermutlich beeniflusst den anderen (Krankheit).

K R

P(Krankheit|Risiko) = „Wahrscheinlichkeit der Krankheit wenn Risiko vorhanden ist" = P(K·R)/P(R)

P(Krankheit|kein Risiko) = $P(K^*\bar{R})/P(\bar{R})$

$$\textbf{Relativ Risiko} \; (RR) = \frac{P(K|R) = P(K \cdot R)/P(R)}{P(K|\bar{R}) = P(K \cdot \bar{R})/P(\bar{R})}$$

siehe später in klinische Studien!

Oddsverhältniss: Verhältniss der Odds mit und ohne Risiko.
(OR)

$$OR = \frac{O_{K|R}}{O_{K|\bar{R}}} = \frac{\frac{P(K|R)}{1-P(K|R)}}{\frac{P(K|R)}{1-P(K|R)}}$$

Dafür brauchen wir 4 bedingte Wahrscheinlichkeiten zu kennen

wichtige spezielle Verteilungen

Verteilung: wenn Ω aus mehreren Ereignisse besteht, dann können wir zu allen Ereignisse einzeln die Wahrscheinlichkeiten angeben. Dies ist die Wahrscheinlichkeitsverteilung des Ω.

In manchen Fällen es ist möglich die Verteilung mit einem Formel zu definieren. Es gibt auch die *kumulative* Vert.: $P(\xi<x)$ (hier ξ ist der Wert der Zufallzahl, und x ist ein gegebener Wert)

in der Medizin benutzte:

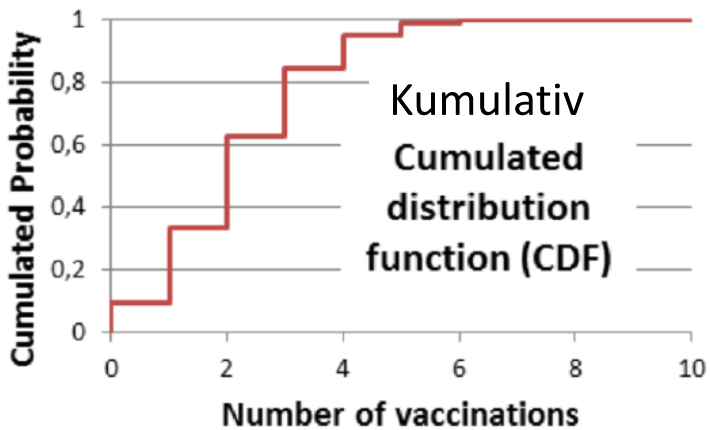
Nicht zu memorisieren!

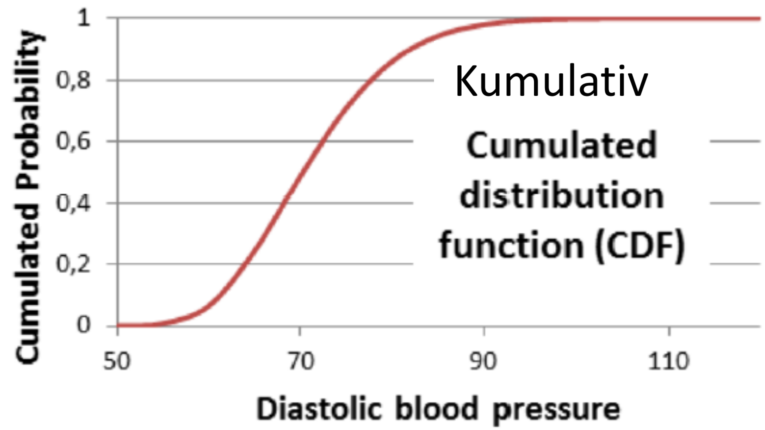
- normal oder Gauss-sche Verteilung
- Student t-Verteilung
- Gleichverteilung
- Exponential
- Binomial
- $\chi 2$
- Geometrisch
- log-Normal

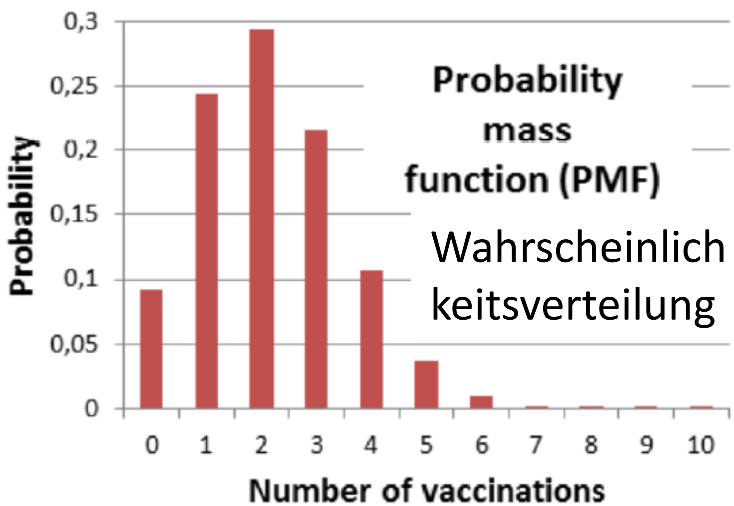
grafische Darstellungmöglichkeiten

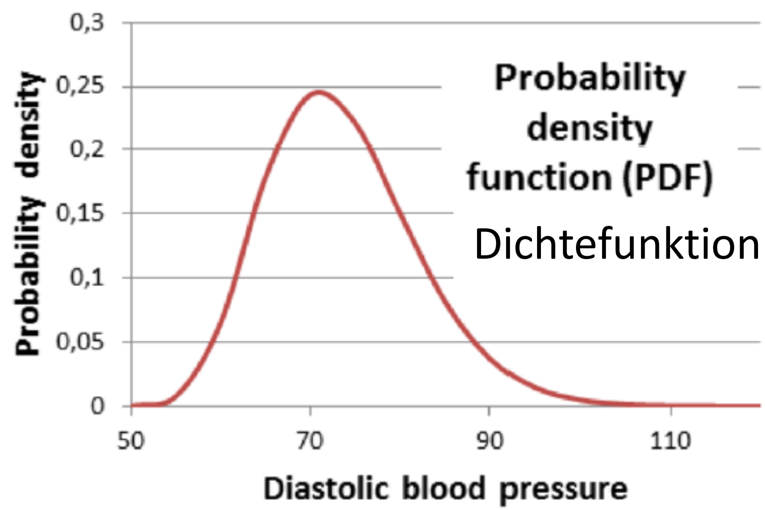
Diskret

Kontinuierlich



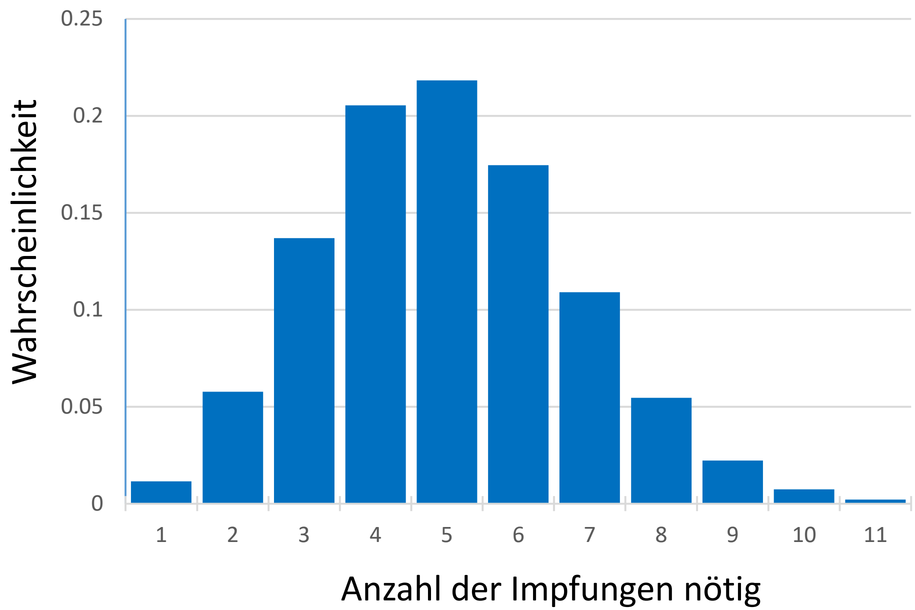


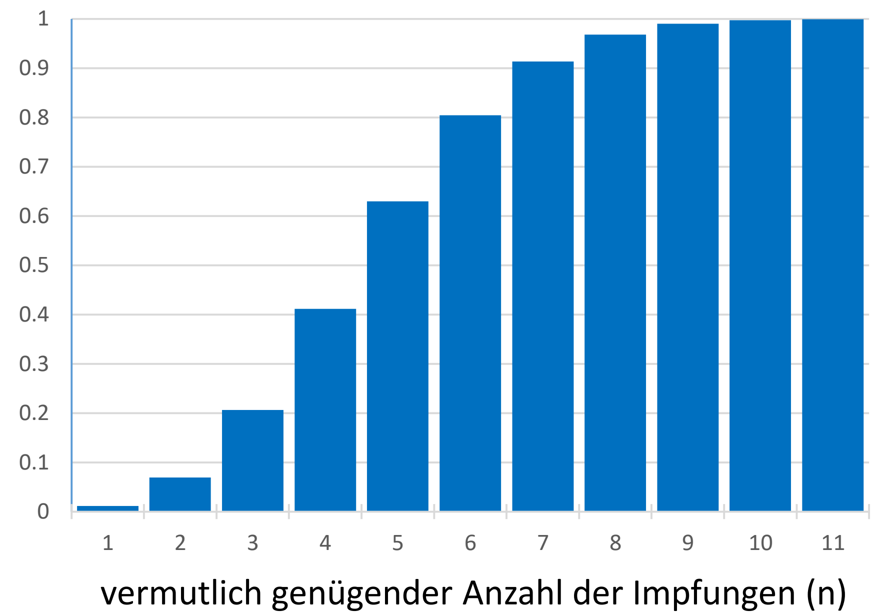




Binomial (Bernoulli) Verteilung

aus N Versuche mit welcher Wahrscheinlichkeit wird Ereigniss „A” k-mal vorkommen.



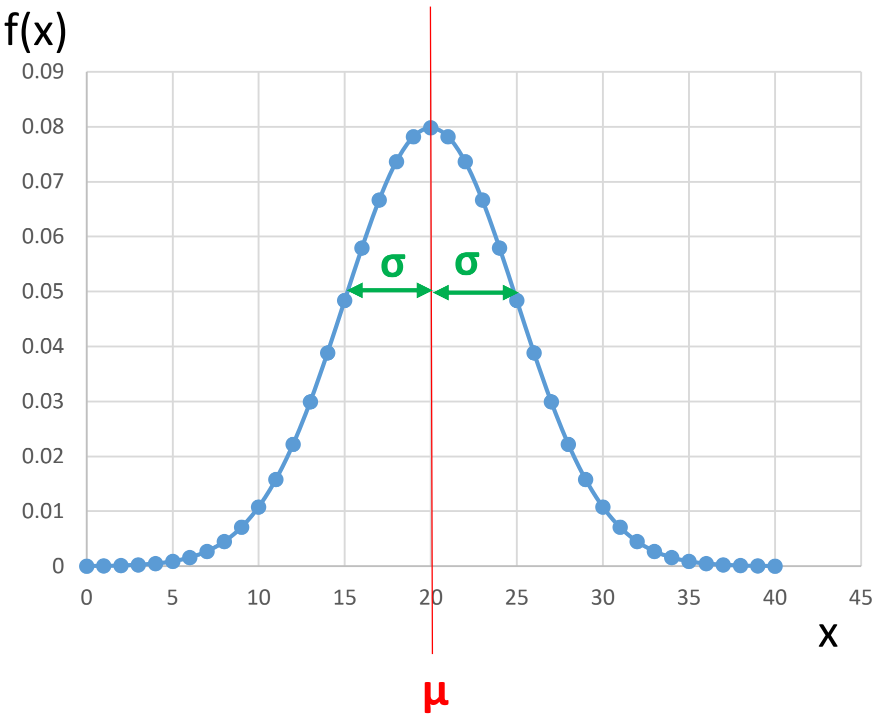


P(x)
(Wahrsch.verteilung)

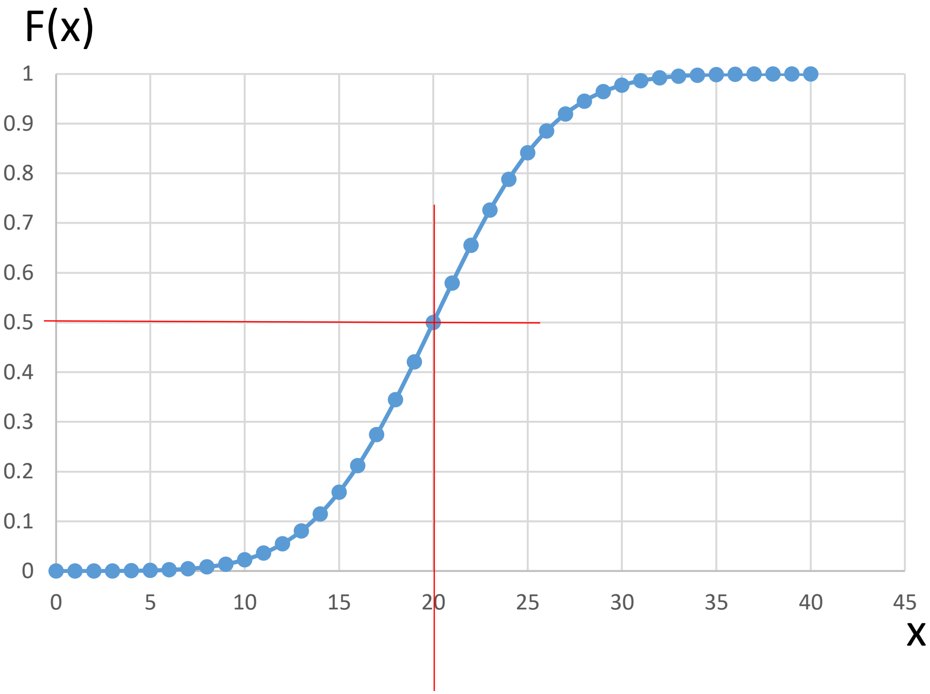
P(x<n)
(Kumulativ)

Beispiel: aus 3000 Fälle in der letzten Monat wir hatten 12 akute Operationen. Heute nacht erwarten wir 20 Patienten, wie viele akute OP werden wir haben, und mit welcher Wahrsch.?

Normal (Gauss-sche)



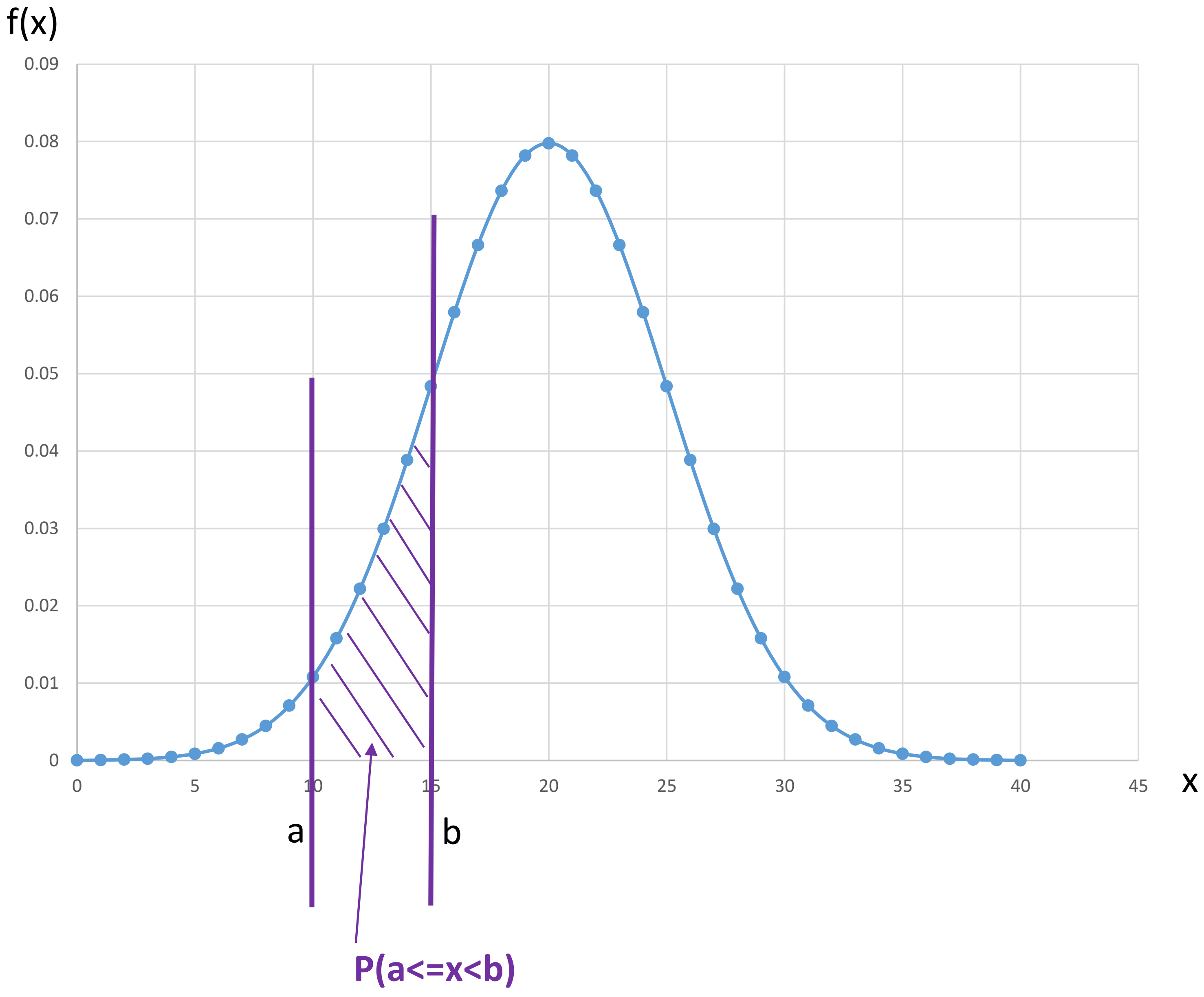
Dichtefunktion



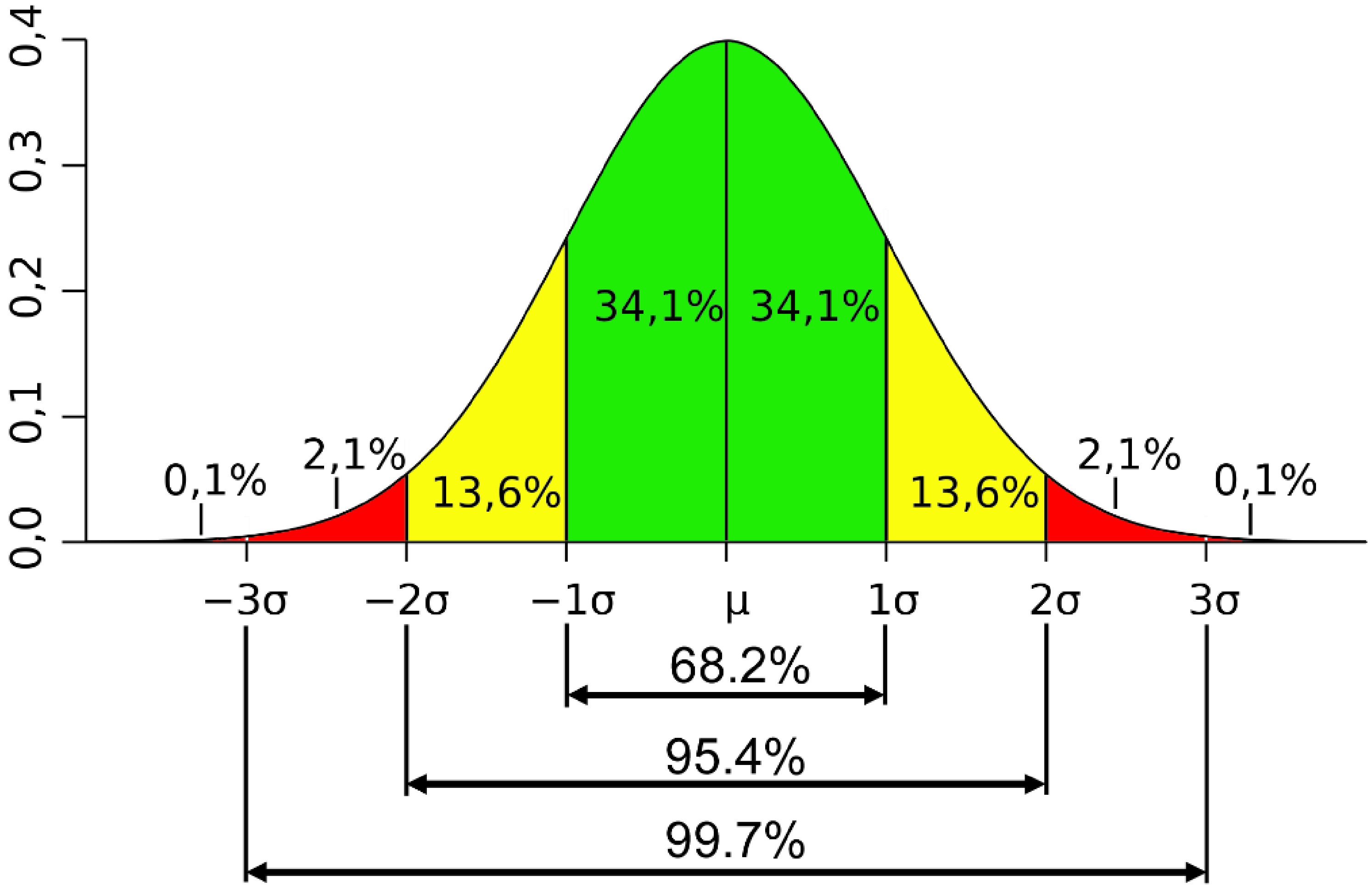
Kumulative Verteilung

für kontinuierliche Daten, der Erwartungswert ist in der mitte **μ** , und die Streuung (Standardabweichung) **σ**.

Vorsicht! Die Dichtefunktion gibt die Wahrscheinlichkeit nicht direkt an, sondern nur die Oberfläche unter der Kurve bedeutet eine Wahrscheinlichkeit.

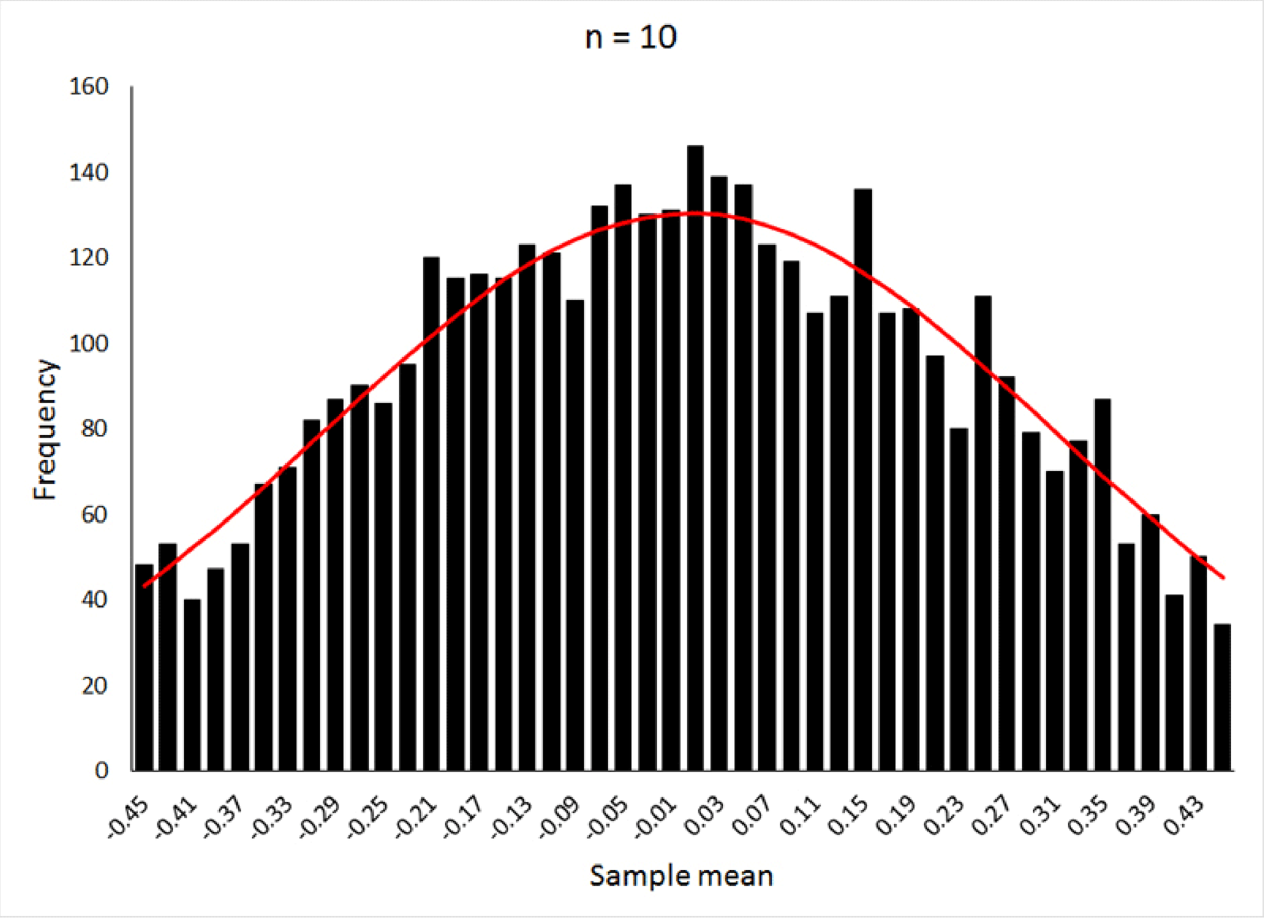


Wichtige Bereiche unter der Kurve



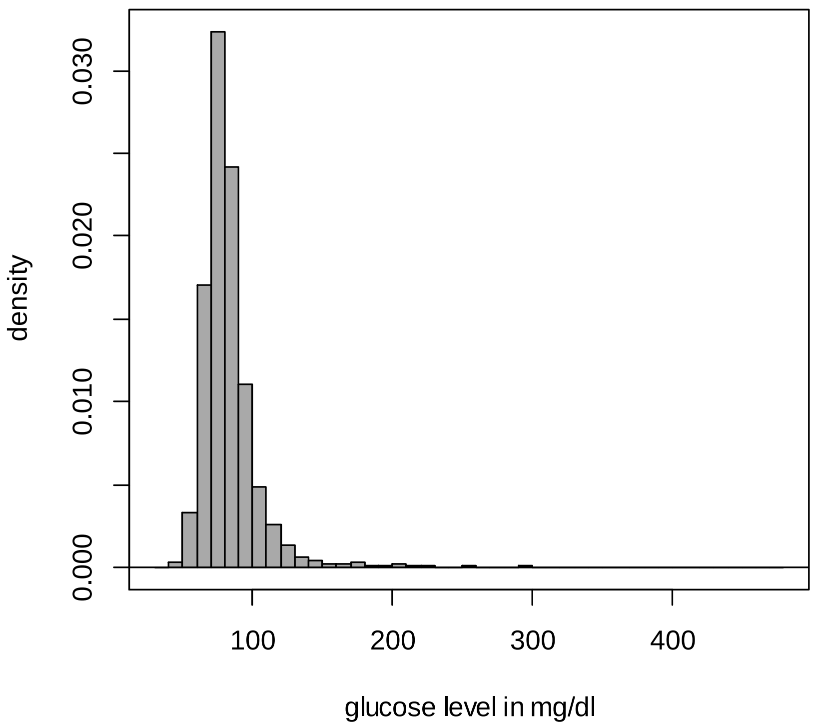
Oft folgen Daten eine Normalverteilung relativ gut/genau.

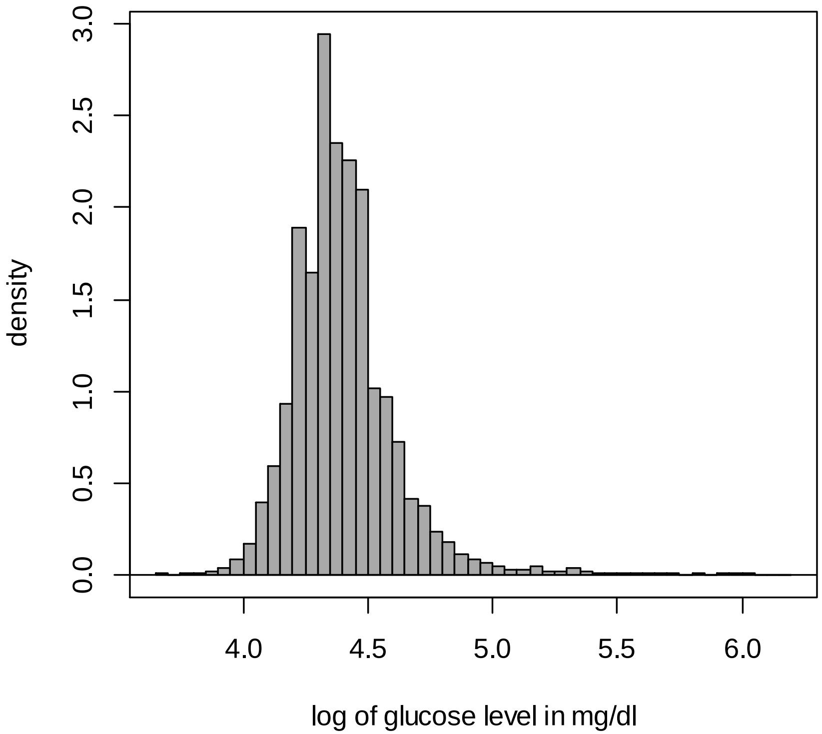
oft wollen wir (μ,σ) schätzen aus der Stichprobe.

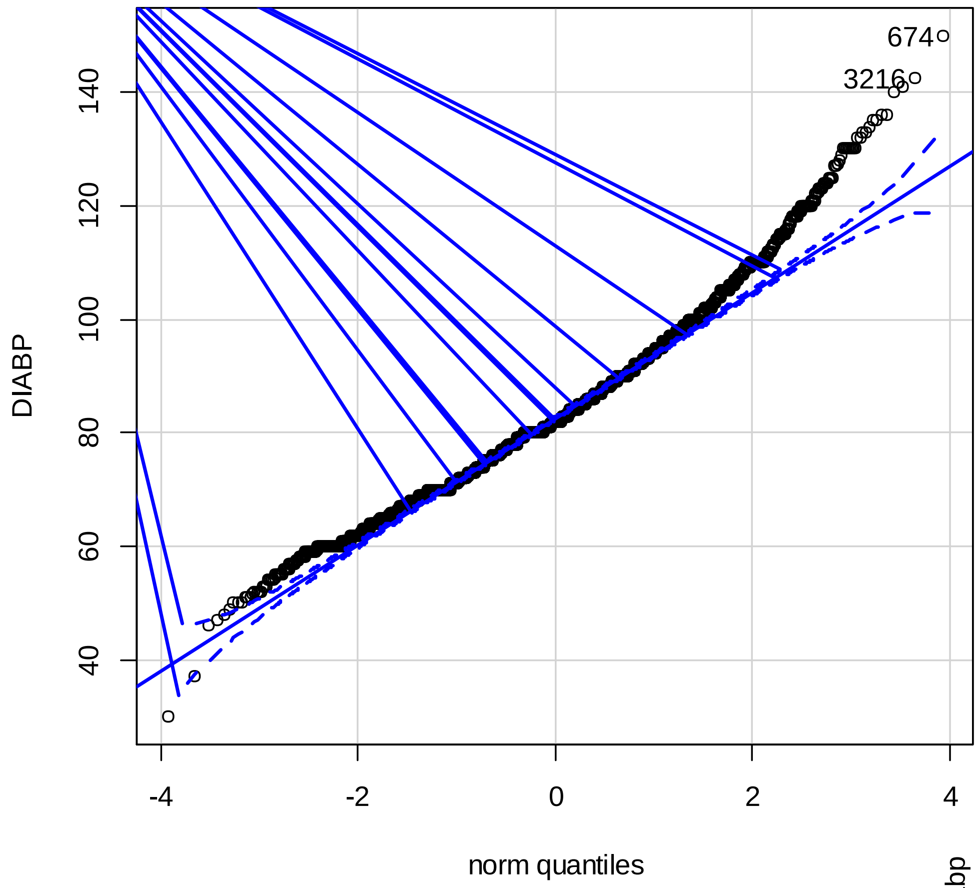


-> nächste Vorlesung

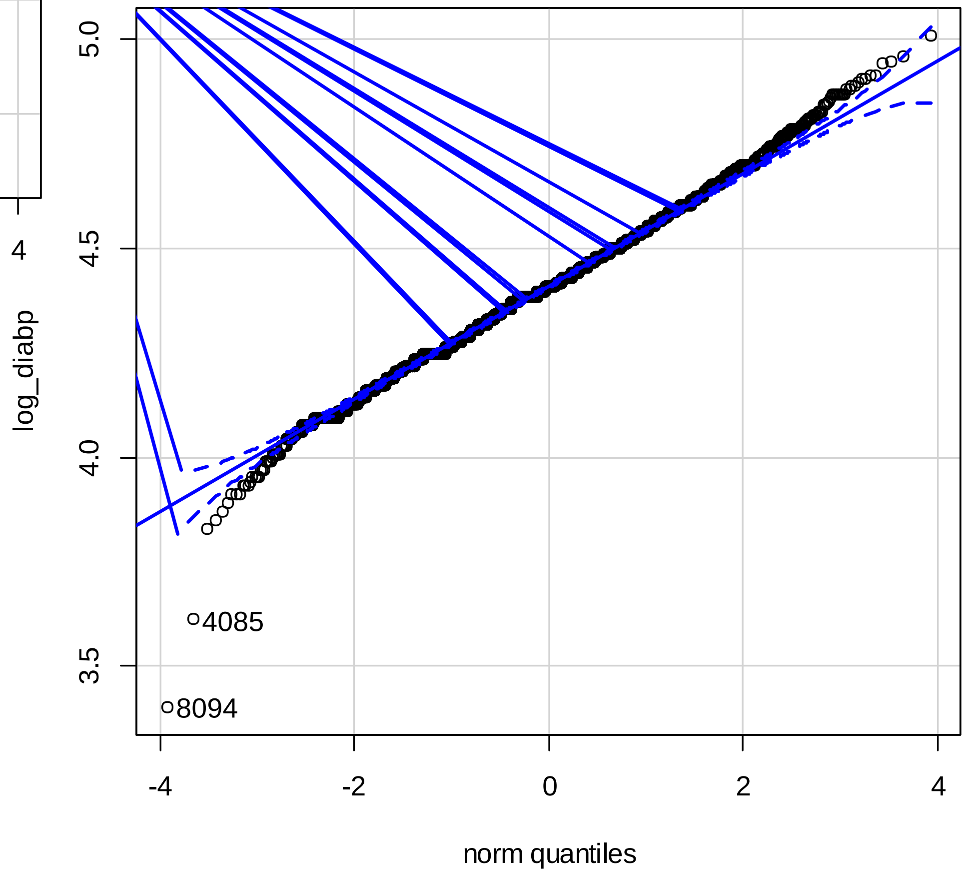
oft sehen wir das ein Merkmal eine schiefe Verteilung folgt, dann ist es möglich das es sich um eine *log-normale Verteilung* handelt, wo die log(x) Werte aber normalverteilt sind.







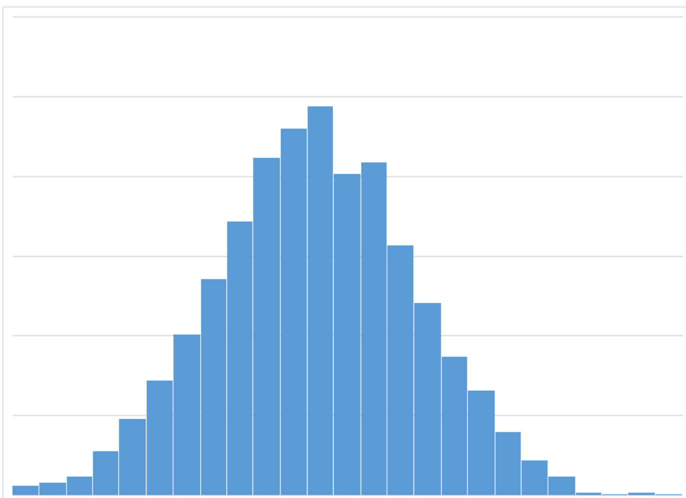
QQ plot
(quantile-quantile plot)



wir vergelichen die Quantile aus der Stichprobe mit denen aus einer bekannten Verteilung, um zu sehen ob die Stichprobe eine bestimmte Verteilung folgt.

Zentraler Grenzwertsatz

Wenn ein Zufallszahl gleichzeitig durch mehrere unabhängige Variablen beeinflusst wird, dann um so mehr Variablen Einfluss auf den Beobachtungswert haben desto besser werden die Beobachtete Werte eine Normalverteilung folgen.

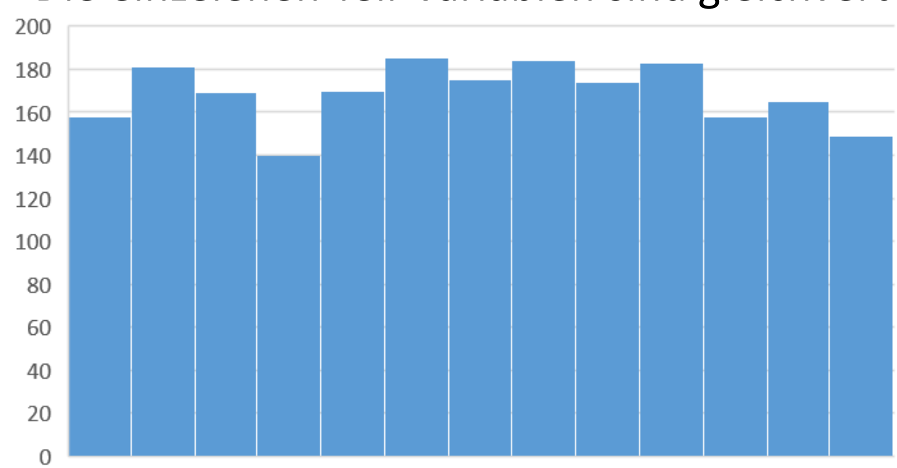


Mittelwert ist normalverteilt

Verteilung der Mittelwerte aus 30 Variablen

Beispiel

Die einzelenen Teil-Variablen sind gleichverteilt



Die spezielle Verteilungen sind benutzbar bei Entscheidungen ☺

-> Hypothesenprüfungen

feedback...