Wahrscheinlichkeitstheorie für Informatikstudien (Kapitel #4)

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.

Zufallsvariablen

4.1. Zufallsvariablen.

4.2. Wahrscheinlichkeitsfunktion.

4.3. Verteilungsfunktion.

4.4. Dichten.

Zufallsvariablen

Ein wichtiges Konzept in der Wahrscheinlichkeitstheorie sind Zufallsvariablen. Wir beginnen mit einem

Beispiel

Angenommen wir spielen n am Roulette:

$$\Omega = \{0, \ldots, 36\}^n, \quad \mathcal{A} = \mathcal{P}(\Omega), \quad P(\omega) = \frac{1}{37^n}.$$

Wenn wir an der Anzahl der Nullen in den n Spielen interessiert sind, dann kann dies formal durch die Funktion

$$X(\omega) = \sum_{k=1}^{n} I\{\omega_k = 0\}, \quad \omega = (\omega_1, \ldots, \omega_n),$$

beschrieben werden. Wir sagen:

"Sei X die Anzahl der Nullen in n Roulette-Spielen".

Zufallsvariablen

Generell ist eine Zufallsvariable (ZV) eine Funktion

$$X : \Omega \to \Omega', \quad \omega \mapsto X(\omega).$$

Meistens ist $\Omega' = \mathbb{N}$ oder $\mathbb{R}$.

Andere wichtige Beispiele sind $\Omega' = \mathbb{R}^d$ (Zufallsvektoren) oder $\Omega' = \mathbb{R}^\infty$ (Zufallsprozesse).

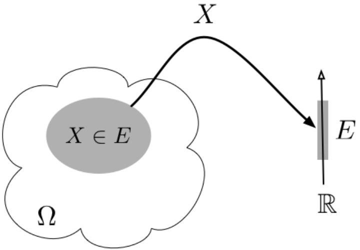
Zufallsvariablen

Eine Zufallsvariable

1. "entnimmt/erfasst" Information aus Elementarereignissen;

2. erzeugt Ereignisse:

$$\{X \in E\} = \{\omega \colon X(\omega) \in E\} = \underbrace{X^{-1}(E)}_{\text{das Urbild von } E} .$$



Zufallsvariablen

Offensichtlich haben wir dieses Konzept bereits benutzt!

Beispiel (Würfeln)

Wir würfeln 2x und interessieren uns für die Summe der Augenzahlen. Beschreibe ein geeignetes Ω und eine ZV X.

Beispiel (Binomialmodell)

Wir ziehen $5x$ aus einer Urne (w weiße und s schwarze Kugeln) mit Zurücklegen. Wir wählen das übliche Ω für ein Laplace-Experiment. Beschreibe formal die ZV X, welche die Anzahl der weißen Kugeln angibt.

Zufallsvariablen

Beispiel (Roulette)

Wir spielen Roulette. Sei X der Gewinn wenn 5 € auf eine ungerade Zahl gesetzt wird. Sei $\Omega = \{0, 1, 2, \ldots, 36\}$. Dann ist

$$X(\omega) = \begin{cases} 5 & \text{für } \omega = 1, 3, 5, ... \\ -5 & \text{für } \omega = 0, 2, 4, 6, ... \end{cases}$$

Also $\{X = 5\} = \{1, 3, \ldots, 35\}$.

Beispiel (Würfel)

Wir würfeln mit 2 Würfel. Sei X der Mittelwert und $\Omega = \{1, \ldots, 6\}^2$. Dann ist

$$X(\omega) = \frac{1}{2}(\omega_1 + \omega_2).$$

Also $\{X \geq 5\} = \{(5, 5), (5, 6), (6, 5), (6, 6), (4, 6), (6, 4)\}$.

Beispiel (Binomiale Zufallsvariable)

Wir spielen Roulette und setzen auf rot. Sei X die Anzahl der Erfolge bei 3 Spielen. Sei $\Omega = \{0, 1, 2, \ldots, 18, 19, \ldots, 36\}^3$. Dann ist

$$X(\omega) = \sum_{k=1}^{3} I\{\omega_k \in \{1, \ldots, 18\}\}.$$

Wir haben z.B.

$$\{X > 1\} = \{\omega \colon X(\omega) > 1\} = \{(1,1,1), (1,1,2), \ldots, (36,18,18)\}.$$

Insbesondere sehen wir, dass alle Ereignisse die wir in den Modellen von Kapitel 3 untersucht haben, sich ganz einfach durch Zufallsvariablen beschreiben lassen.

Beachte: Es gibt die Konvention Zufallsvariablen mit Großbuchstaben zu bezeichnen aber den Wert, den die Zufallsvariable annimmt (Realisation) klein zu schreiben:

“Wie groß ist die Wahrscheinlichkeit, dass $X = x$?”

Wahrscheinlichkeitsfunktion

Wenn X nur abzählbar viele Werte $x_1, x_2, \ldots$ annimmt, dann heißt die Zufallsvariable diskret.

Ist X eine diskrete ZV mit Realisierungen x_k, dann nennt man die Funktion $p_k = P(X = x_k)$ die Wahrscheinlichkeitsfunktion (PMF = probability mass function) von X.

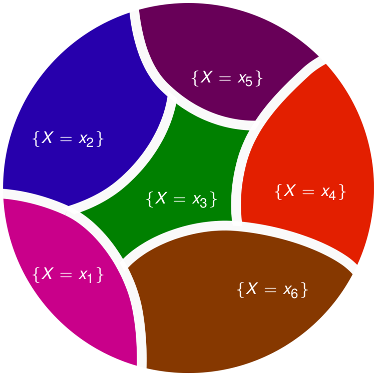
Beispiel

Wir ziehen $5x$ aus einer Urne (10 weiße und 6 schwarze Kugeln) mit Zurücklegen. Sei X die Anzahl der weißen Kugeln. Wie lautet die PMF von X?

Wahrscheinlichkeitsfunktion

Eine diskrete Zufallsvariable partitioniert den Stichprobenraum Ω in abzählbar viele Ereignisse $A_k = \{X = x_k\}$.

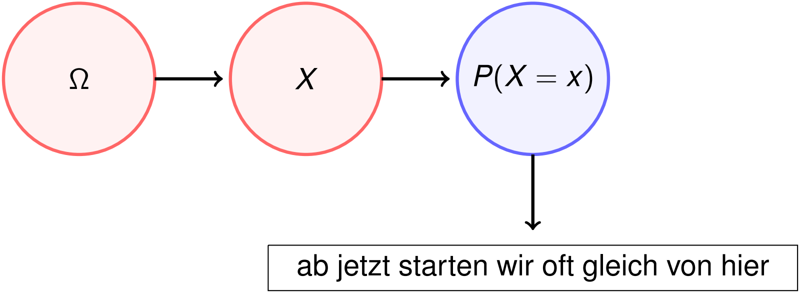
Die Ereignisse A_k haben Wahrscheinlichkeit p_k.



Wahrscheinlichkeitsfunktion

Für die Modelle in Kapitel 2 und 3 wissen wir jetzt wie man geeignete Wahrscheinlichkeitsräume und Zufallsvariablen definieren kann.

Außerdem haben wir deren Verteilungen bestimmt.



Jetzt wo wir ein paar wichtige PMFs kennen (wie zB jene von zuvor) und wo wir wissen wie und in welchem Kontext wir sie verwenden können, interessiert uns oft der zugrunde liegende W-Raum nicht mehr.

Wir arbeiten dann nur mehr mit der PMF und starten die Modellierung auf dem Level von PMFs.

Wir sagen dann: Sei X eine ZV mit PMF xyz ...

Beispiel

> Sei X binomalverteilt mit $p = 0.3$ und $n = 4$. Bestimme $P(X \text{ ist eine gerade Zahl})$.

Wahrscheinlichkeitsfunktion

In Kurzform:

$X \sim \mathrm{Unif}(\Omega)$	Gleichverteilung auf Ω.
$X \sim H_{n,w,s}$	hypergeom. Vert. mit Parametern w und s.
$X \sim B_{n,p}$	Binomialverteilung mit Parametern n und p.
$X \sim G_p$	geometrische Verteilung mit Parametern p.
$X \sim P_\lambda$	Poisson-Verteilung mit Parametern λ.

Beispiel

Sei $X \sim G_{0.5}$. Man bestimme $P(X > 2)$.

Beispiel

Sei $X \sim P_2$. Man bestimme $P(2 < X \leq 4)$.

Verteilungsfunktion

Als nächstes möchten wir mit stetigen ZVen arbeiten, also ZVen die überabzählbar viele Werte annehmen können. Dazu brauchen wir das Konzept der kumulative Verteilungsfunktion (CDF = cumulative distribution function).

> Die Funktion $F^X(x) = P(X \leq x)$, $x \in \mathbb{R}$, nennt sich *kumulative Verteilungsfunktion* von X.

Beachte: Es gilt stets, dass

$$P(X \in (a, b]) = F(b) - F(a),$$

denn wir verwenden das Axiom (A) für

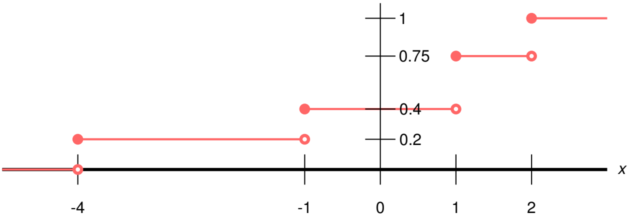
$$\{X \leq b\} = \underbrace{\{X \leq a\} \cup \{X \in (a, b]\}}_{\text{disjunkt}}.$$

Verteilungsfunktion

Die CDF einer Zufallsvariable X mit PMF

$$P(X = k) = \begin{cases} 0.2 & \text{falls} \quad k = -4 \\ 0.2 & \text{falls} \quad k = -1 \\ 0.35 & \text{falls} \quad k = 1 \\ 0.25 & \text{falls} \quad k = 2 \end{cases}$$

ist



Verteilungsfunktion

Beispiel (Würfeln)

Angenommen wir würfeln und X ist die Augenzahl. Was ist die PMF und wie sieht die CDF von X aus?

Beispiel (Gleichverteilung)

Skizziere die CDF von X, wenn X auf $\{-3, -1, 0, 1\}$ gleichverteilt ist.

Beispiel (Binomialverteilung)

Skizziere die CDF von X, wenn $X \sim B_{1,1/3}$.

Verteilungsfunktion

Bei diskreten Zufallsvariablen haben wir also wir folgende Relation zwischen PMF und CDF:

- $F^X(x) = \sum_{x_k \leq x} p_k$.
- Wenn $\forall k \; x_{k-1} < x_k$ dann ist $F^X(x_k) - F^X(x_{k-1}) = p_k$.
- F^X ist eine Treppenfunktion und konstant auf $[x_{k-1}, x_k)$.

Verteilungsfunktion

Ganz allgemein gilt folgende Aussage:

Lemma

Sei F eine CDF. Dann gilt

- *F ist monoton steigend;*
- *F ist rechtsstetig;*
- $\lim_{x\to-\infty} F(x) = 0$ *und* $\lim_{x\to\infty} F(x) = 1$.

Jede Funktion, die diese drei Eigenschaften erfüllt, nennt man eine Verteilungsfunktion (DF = distribution function).

Es gilt folgende wichtige “Umkehrung” zum Lemma:

Satz

Sei F eine DF. Dann existiert ein W-Raum $(\Omega, \mathcal{A}, P)$ und auf diesem eine ZV X sodass $P(X \leq x) = F(x)$.

Verteilungsfunktion

Bis jetzt haben wir nur ZV und Verteilungen von diskreten Experimenten gesehen (X nimmt also Werte einer abzählbaren Menge an). Was ist ein Beispiel für eine stetige ZV?

Beispiel

Angenommen wir fahren zu einer Ampel. Wir wissen, dass die rote Phase 2 Minuten dauert. Wenn die Ampel bei Ankunft rot ist, wie sieht die Verteilung der Wartezeit aus?

Beispiel

Angenommen wir sitzen an einem Teich und fischen. Wir wissen aus Erfahrung, dass es im Durchschnitt 30 Minuten dauert, bis man einen Fisch fängt. Wie groß ist die Wahrscheinlichkeit, dass wir weniger als x Minuten warten müssen?

Verteilungsfunktion

Beispiel

Angenommen wir müssen eine Garantie für die Lebenszeit eines Produkts geben (zB Glühbirne). Wie groß ist die Wahrscheinlichkeit, dass die Glühbirne länger als x Zeiteinheiten funktioniert?

Beispiel

Was groß ist die Wahrscheinlichkeit, dass ein Schadensfall mit mehr als x € eintritt? (Versicherungen)

Verteilungsfunktion

Wir betrachten die ersten beiden Beispiele und versuchen ein Modell zu konstruieren.

Ampelbeispiel. Wenn wir zufällig an der Ampel ankommen, so scheint jede Wartezeit zwischen 0 und 2 Minuten gleich wahrscheinlich. Das klingt nach einem Laplace-Experiment, aber jetzt ist $\Omega = [0, 2]$ und daher $|\Omega| = \infty$ (nicht zählbar)!

Analog zum Laplace-Experiment setzen wir:

$$P(X \in A) = \frac{\text{Länge}(A)}{\text{Länge}(\Omega)}.$$

Dann ist die CDF

$$F^X(x) = P(X \leq x) = \begin{cases} 0 & \text{wenn} \quad x < 0; \\ x/2 & \text{wenn} \quad 0 \leq x \leq 2; \\ 1 & \text{wenn} \quad x > 2. \end{cases}$$

Beispiel (Stetige Gleichverteilung)

In einem Bergwerk verwendet man schwere 10m lange Ketten zum Heben von Schlacke. Im Prinzip würde auch schon eine Länge von 8m reichen. Wenn eine Kette reißt (an einer komplett zufälligen Stelle), wie groß ist die Wahrscheinlichkeit, dass man Sie weiter verwenden kann?

Verteilungsfunktion

Fischbeispiel. Im diskreten Fall ist die Wartezeit bis zum ersten Erfolg mit einer geometrischen Verteilung modelliert:

$$P(X = k) = pq^k, \quad k = 1, 2, \ldots$$

Die CDF ist

$$F^X(k) = P(X \leq k) = 1 - q^k = 1 - e^{-\lambda k}, \quad \lambda = -\log(q).$$

Eine stetige Version davon könnte so sein

$$F^X(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad \text{für ein } \lambda > 0 \text{ und } x \geq 0.$$

Zeige, dass das für $\lambda > 0$ eine CDF ist. Damit bekommt man eine parametrische (mit Parameter λ) Familie von Verteilungsfunktionen $F_\lambda(x)$.

Verteilungsfunktion

Wie sollen wir im letzten Beispiel λ wählen?

Eine naive Idee wäre es, $P(X = 30)$ in Abhängigkeit von λ zu maximieren.

Allerdings ist $P(X = x) = 0$ für jedes x und jedes λ!

Denn: sei $x_n \nearrow x$. Dann ist

$$P(X = x) = \lim_n P(X \in (x_n, x]) = \lim_n (F_\lambda(x) - F_\lambda(x_n)) = 0.$$

Verteilungsfunktion

Alternativ:

Wähle jenes λ wo $f_\lambda(30) := \frac{\partial}{\partial x} F_\lambda(x)|_{x=30}$ maximal wird.

Es gilt

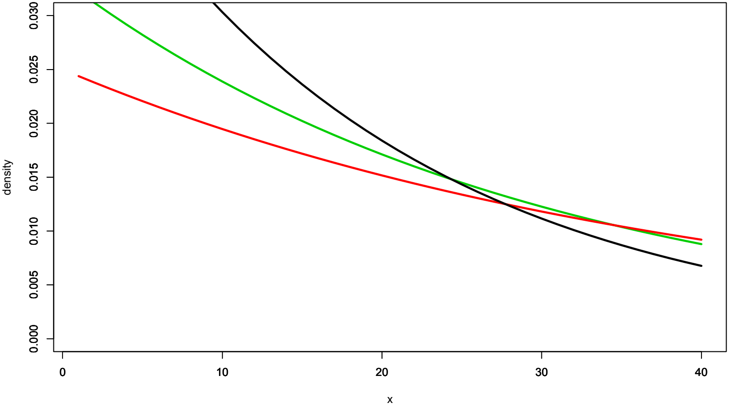
$$f_\lambda(30) = \lambda e^{-30\lambda}.$$

> Wähle λ derart, dass $f_\lambda(30)$ maximal ist und zeige dass dann $\lambda = 1/30$.

Dichten

Für differenzierbare CDFs ist die Funktion $f^X(x) = \frac{d}{dx}F^X(x)$ sehr wichtig. Es ist das stetige Analogon der PMF.

Der Graph zeigt $f_\lambda(x)$ für $\lambda = 1/20$, $\lambda = 1/40$, $\lambda = 1/30$.



Dichten

Stetige Zufallsvariablen haben ein sehr einfaches Konzept mit dem man CDFs generieren kann.

Sei $f : \mathbb{R} \to [0, \infty)$ integrierbar und man nehme an,

$$\int_{-\infty}^{\infty} f(t)dt = 1.$$

Dann ist f eine Dichtefunktion (PDF).

Es folgt dann: $F(x) = \int_{-\infty}^{x} f(t)dt$ ist eine Verteilungsfunktion.

Wenn X Dichte f hat, dann ist

$$P(X \in (a, b]) = \int_{a}^{b} f(t)dt.$$

Je größer $f(t)$, umso wahrscheinlicher ist X in der Nähe von t.

Dichten

Beispiel (Exponential-Verteilung mit Parameter λ)

Sei $\lambda > 0$ (beliebig aber fix) und

$$f(t) = \begin{cases} 0, & t < 0; \\ \lambda e^{-\lambda t}, & t \geq 0. \end{cases}$$

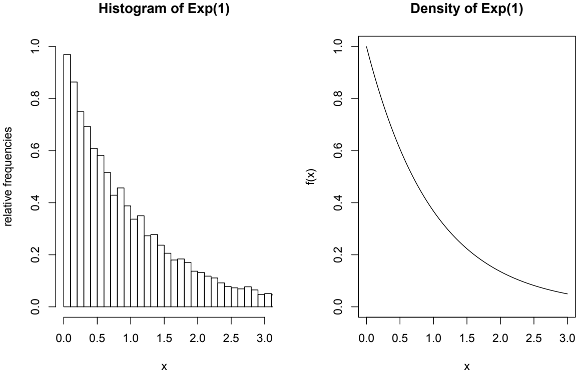
Dann

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, & x < 0; \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases}$$

Diese Verteilung wird oft benutzt um Wartezeiten zu modellieren.

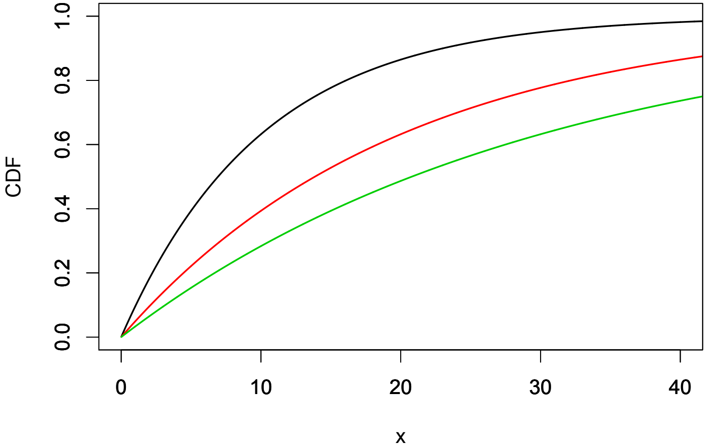
Folgt X dieser Verteilung, so schreiben wir $X \sim \mathrm{Exp}(\lambda)$.

Dichten


Histogram of Exp(1)
relative frequencies
0.0
0.2
0.4
0.6
0.8
1.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0
x
Density of Exp(1)
f(x)
0.0
0.2
0.4
0.6
0.8
1.0
0.0 0.5 1.0 1.5 2.0 2.5 3.0
x

Dichten

Dieser Graph zeigt die Verteilungsfunktion einer Exp(λ) Variable für $\lambda = 1/10$, $\lambda = 1/20$, $\lambda = 1/30$.



Beispiel (Stetige Gleichverteilung auf $[a, b]$)

Sei

$$f(t) = \begin{cases} \frac{1}{b-a}, & t \in [a, b]; \\ 0, & \text{sonst.} \end{cases}$$

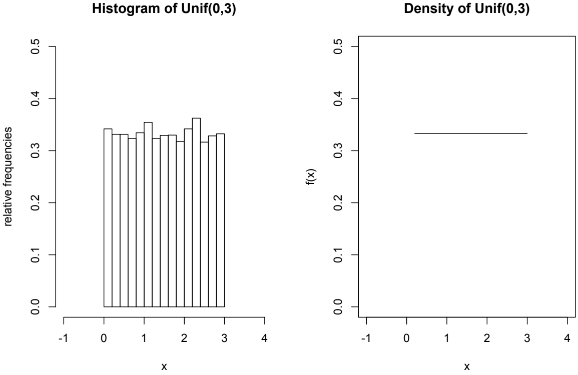
Dann

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & x \in [a, b] \\ 1, & x > b. \end{cases}$$

Die stetige Gleichverteilung ist das stetige Analogon des Laplace-Experiments.

Folgt X dieser Verteilung, schreiben wir $X \sim \mathrm{Unif}([a, b])$.

Dichten


Histogram of Unif(0,3)
relative frequencies
0.0
0.1
0.2
0.3
0.4
0.5
-1
0
1
2
3
4
x
Density of Unif(0,3)
f(x)
0.0
0.1
0.2
0.3
0.4
0.5
-1
0
1
2
3
4
x

Beispiel (Normalverteilung (μ, σ^2))

Sei
$$\phi_{\mu,\sigma^2}(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2}.$$

Dann ist
$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \phi_{\mu,\sigma^2}(t)dt$$

die Normalverteilungsfunktion.

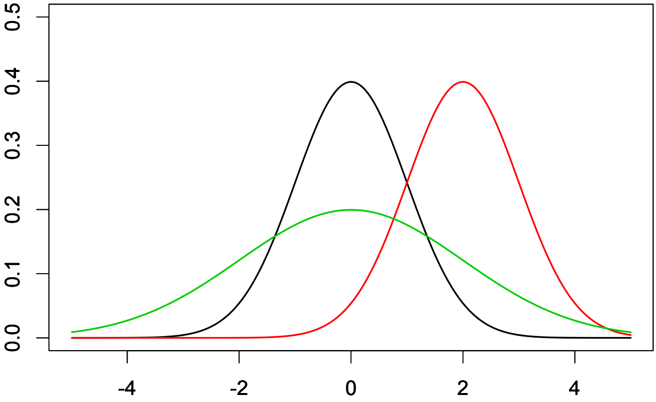
Es gibt keine geschlossene Form für das Integral. Die Werte sind tabelliert für $\mu = 0$ und $\sigma^2 = 1$: Standardnormalverteilung.

Die Normalverteilung ist fundamental aus mehreren Gründen. Sie wird später noch vorkommen.

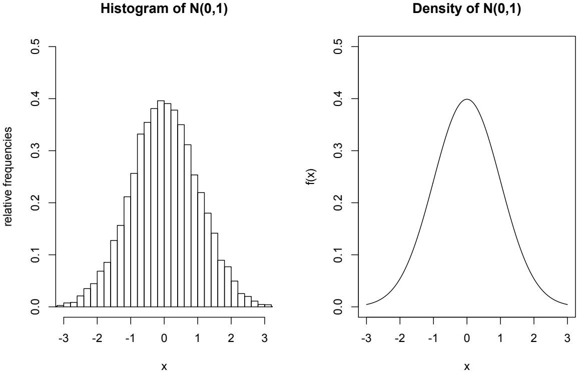
Folgt X dieser Verteilung, schreiben wir $X \sim N(\mu, \sigma^2)$.

Dichten

Der Graph zeigt die Dichtefunktion einer $N(\mu, \sigma^2)$ Variable für $(\mu, \sigma) = (0, 1)$, $(\mu, \sigma) = (2, 1)$, $(\mu, \sigma) = (0, 2)$.



Dichten


Histogram of N(0,1)
relative frequencies
0.0
0.1
0.2
0.3
0.4
0.5
-3
-2
-1
0
1
2
3
x
Density of N(0,1)
f(x)
0.0
0.1
0.2
0.3
0.4
0.5
-3
-2
-1
0
1
2
3
x

Dichten

Wichtige Beobachtung: Wenn $X \sim N(\mu, \sigma^2)$, dann gilt

$$P(X \leq \sigma z + \mu) = \int_{-\infty}^{\sigma z + \mu} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

$$\underbrace{=}_{\text{Substitution } s = \frac{t-\mu}{\sigma}} \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds.$$

Es gilt also

$$P\left(\frac{X-\mu}{\sigma} \leq z\right) = P(Z \leq z) =: \Phi(z)$$

wobei $Z \sim N(0, 1)$.

Außerdem:

$$P(Z \leq z) \underbrace{=}_{\text{Symmetrie}} P(Z > -z) \underbrace{=}_{\text{Gegenwahrscheinlichkeit}} 1 - P(Z \leq -z).$$

Dichten

Beispiel (Normalverteilung)

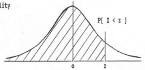
Es sei $X \sim N(3, 4)$. Bestimme $P(X \leq 4)$ und $P(X \leq 2)$.

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z i.e.

$$P[\,Z < z\,] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\, dZ$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7854
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8804	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000