Anhang A

Verteilungs-Datenblätter

Die folgenden Abschnitte stellen die wichtigsten Eigenschaften der verschiedenen in früheren Kapiteln behandelten Verteilungen zusammen.

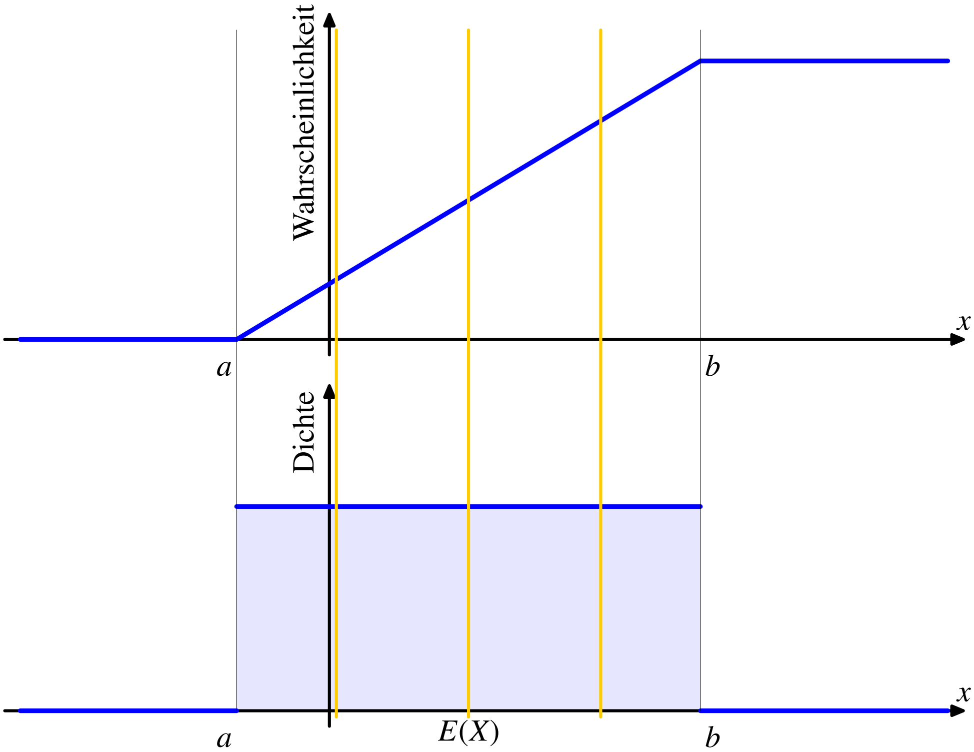
A.1 Gleichverteilung

A.1.1 Steckbrief

Name	Gleichverteilung
Dichtefunktion	$\begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{sonst} \end{cases}$
Verteilungsfunktion	$\begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & x \le a \le b \\ 1 & x > b \end{cases}$
Erwartungswert	$\dfrac{a+b}{2}$
Varianz	$\dfrac{(b-a)^2}{12}$
Median	$\dfrac{a+b}{2}$
$P(\lvert X - E(X)\rvert > \varepsilon)$	$1 - \dfrac{2\varepsilon}{b-a}$ für $\varepsilon < \frac{b-a}{2}$
Anwendungen	• Verteilung von Zufallszahlen • Keine bevorzugten Werte

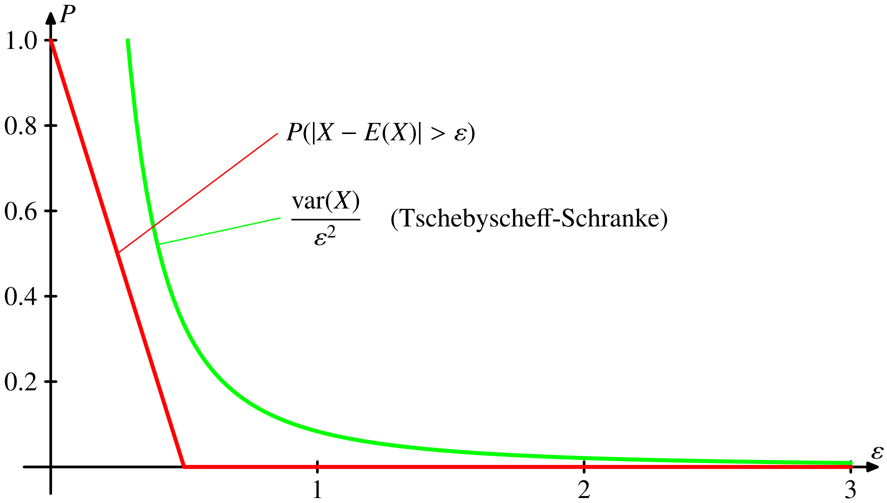
A.1.2 Verteilungsfunktion und Wahrscheinlichkeitsdichte

Verteilungsfunktion (oben) und Wahrscheinlichkeitsdichte (unten) der Gleichverteilung:



A.1.3 Wahrscheinlichkeit einer grossen Abweichung

Wahrscheinlichkeit einer Abweichung vom Mittelwert einer in $[0, 1]$ gleichverteilten Zufallsvariable (rot) im Vergleich mit der oberen Schranke aus dem Satz von Tschebyscheff (grün):



A.1.4 Parameter schätzen

Die Parameter a und b der Gleichverteilung können mit den erwartungstreuen Schätzern

$$\hat{a}(x_1, \dots, x_n) = \frac{n+1}{n} \min(x_1, \dots, x_n)$$
$$\hat{b}(x_1, \dots, x_n) = \frac{n+1}{n} \max(x_1, \dots, x_n)$$

geschätzt werden.

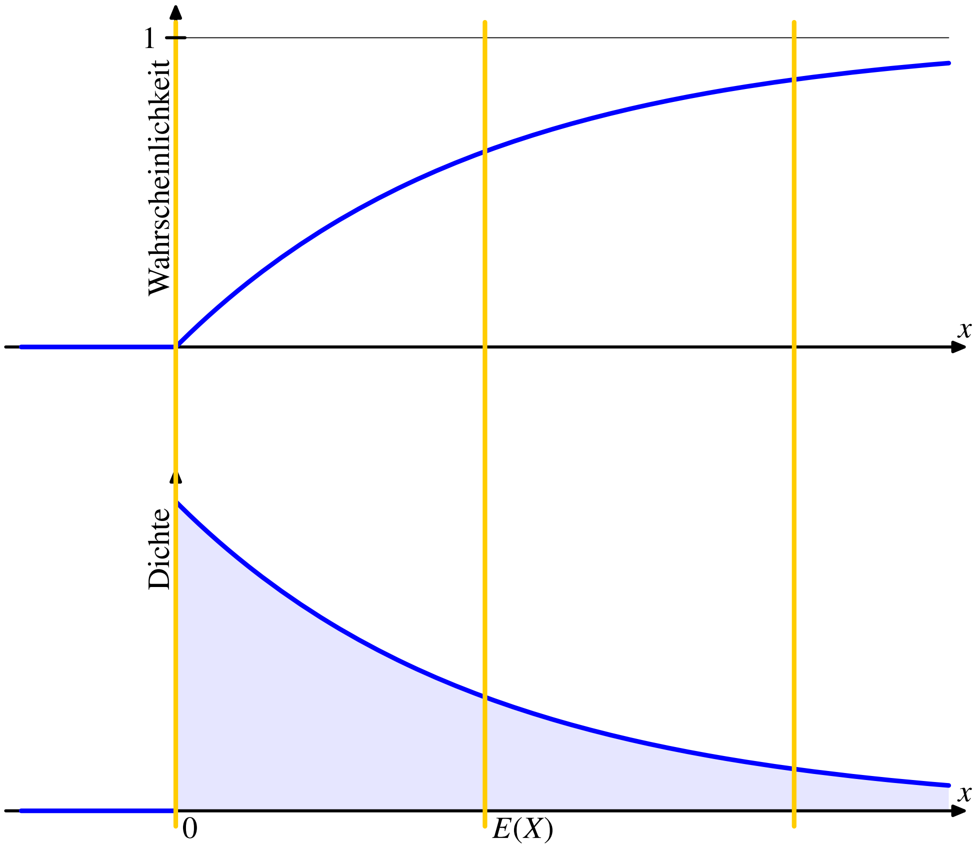
A.2 Exponentialverteilung

A.2.1 Steckbrief

Name	Exponentialverteilung
Dichtefunktion	$ae^{-ax}, \quad a > 0$
Verteilungsfunktion	$1 - e^{-ax}$
Erwartungswert	$\frac{1}{a}$
Varianz	$\frac{1}{a^2}$
Median	$\frac{1}{a} \log 2$
$P(\|X - E(X)\| > \varepsilon)$	$\begin{cases} e^{-a\varepsilon-1} & \text{für } \varepsilon > \frac{1}{a} \\ 1 - e^{a\varepsilon-1} + e^{-a\varepsilon-1} & \text{für } \varepsilon \leq \frac{1}{a} \end{cases}$
Anwendungen	• Prozess ohne Erinnerungsvermögen • Radioaktivität

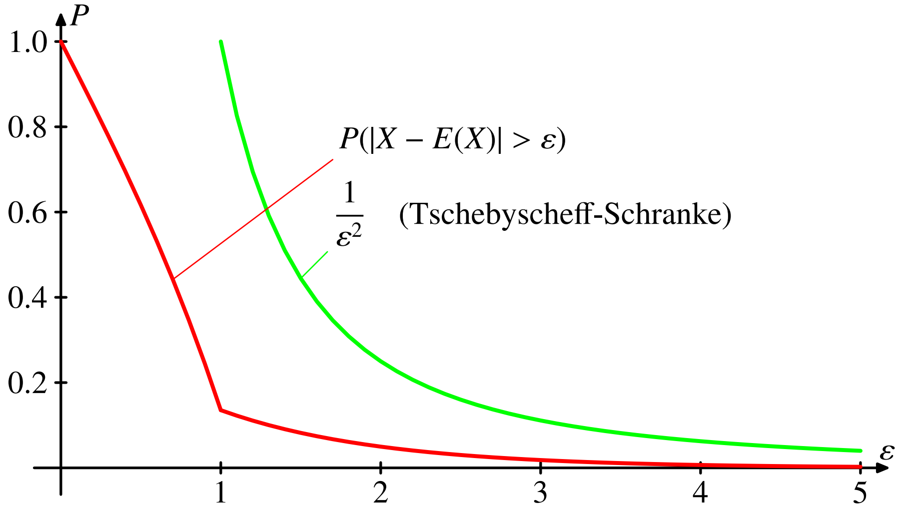
A.2.2 Verteilungsfunktion und Wahrscheinlichkeitsdichte

Verteilungsfunktion (oben) und Dichtefunktion (unten) der Exponentialverteilung:



A.2.3 Wahrscheinlichkeit grosser Abweichungen

Wahrscheinlichkeit für eine grosse Abweichung bei einer Exponentialverteilten Zufallsvariable, oben die durch den Satz von Tschebyscheff gegebene Schranke (grün), unten die exakte Rechnung mit Hilfe der Exponentialvereteilung (rot):



A.2.4 Parameter schätzen

Der Parameter $1/a$ kann mit dem erwartungstreuen Schätzer

$$\frac{1}{\hat{a}(x_1, \dots, x_n)} = \frac{x_1 + \cdots + x_n}{n}$$

geschätzt werden.

A.2.5 Erlang-Verteilungen

Seien X_i unabhängige, identisch mit Parameter a exponentialverteilte Zufallsvariablen. Dann ist $X_1 + \cdots + X_n$ Erlang-verteilt. Die Erlang-Verteilungen wurden für die Analyse von Telefonzentralen erfunden, und sind allgemein in der Queueing-Theorie nützlich. Die Wahrscheinlichkeitsdichte und die Verteilungsfunktionen sind

$$F_{X_1+\cdots+X_k}(x) = \begin{cases} 1 - e^{-ax} \sum_{i=0}^{k-1} \frac{(ax)^i}{i!} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\varphi_{X_1+\cdots+X_k}(x) = \begin{cases} a^k \frac{x^{k-1}}{(k-1)!} e^{-ax} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

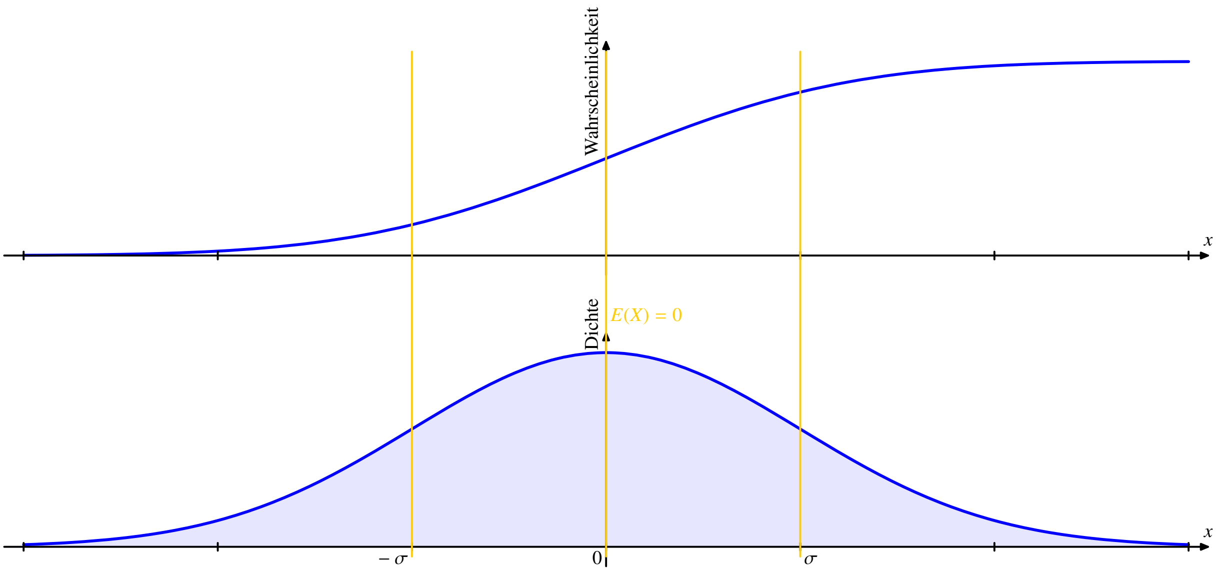
A.3 Normalverteilung

A.3.1 Steckbrief

Name	Normalverteilung
Dichtefunktion	$\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
Verteilungsfunktion	keine elementare Funktion
Erwartungswert	μ
Varianz	σ^2
Median	μ
$P(\lvert X - E(X)\rvert > \varepsilon)$	keine einfache Formel
Anwendungen	• Messwerte • Summe vieler kleiner Einflüsse vergleichbar grosser Varianz (Zentraler Grenzwertsatz) • Approximation der Binomialverteilung

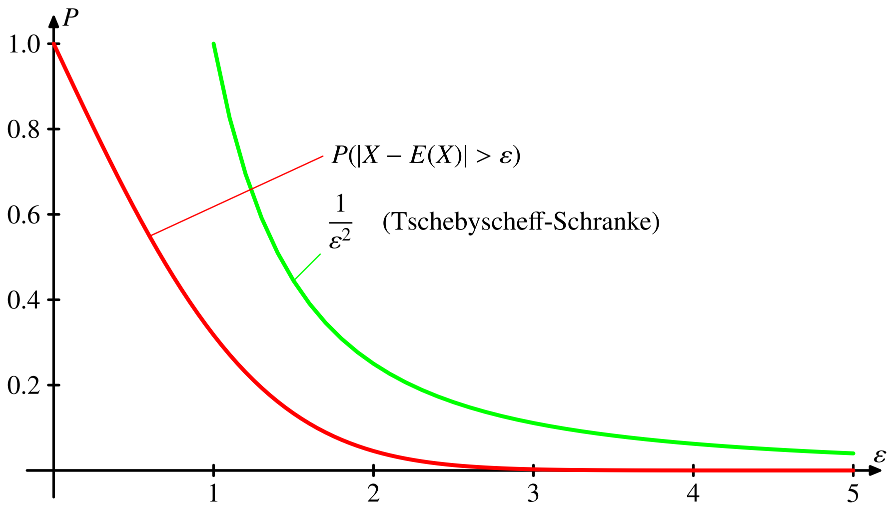
A.3.2 Verteilungsfunktion und Wahrscheinlichkeitsdichte

Verteilungsfunktion (oben) und Dichtefunktion (unten) der Normalverteilung:



A.3.3 Wahrscheinlichkeit einer grossen Abweichung

Vergleich der Wahrscheinlichkeit für eine grosse Abweichung für die Normalverteilung (rot) und die Schranke von Tschebyscheff (grün):



A.3.4 Parameter schätzen

Die Parameter μ und σ können mit den erwartungstreuen Schätzern

$$\begin{aligned}\hat{\mu}(x_1,\dots,x_n) &= \frac{x_1+\dots+x_n}{n}\\ \hat{\sigma}(x_1,\dots,x_n)^2 &= \frac{1}{n-1}\biggl(\sum_{i=1}^n x_i^2 - \frac{1}{n}\Bigl(\sum_{i=1}^n x_i\Bigr)^2\biggr)\end{aligned}$$

geschätzt werden.

Der Mittelwert ist $\hat{\mu}(x_1,\dots,x_n)$ ist normalverteilt mit Erwartungswert μ und Varianz $\frac{1}{n}\sigma^2$. Die Stichprobenvarianz $\hat{\sigma}(x_1,\dots,x_n)^2/\sigma^2$ ist χ^2-verteilt mit $n-1$ Freiheitsgraden.

A.3.5 Zentraler Grenzwertsatz

Der zentrale Grenzwertsatz besagt, dass die Verteilungsfunktion einer Summe einer grossen Zahl von Zufallsvariablen unter milden Voraussetzungen gegen die Verteilungsfunktion einer Normalverteilung konvergiert. Dies rechtfertigt den Einsatz der Normalverteilung als Modell für Prozesse, in denen eine grosse Zahl von vergleichbar grossen Einflüssen zu einem Effekt beitragen, zum Beispiel bei Messwerten.

A.3.6 Standardisierung

Ist X eine normalverteilte Zufallsvariable, dann ist

$$Z = \frac{X-\mu}{\sigma}$$

eine normalverteilte Zufallsvariable mit Erwartungswert 0 und Varianz 1, d. h. eine standardnormalverteilte Zufallsvariable. Für die Standardnormalverteilung ist im Tabellenanhang eine Tabelle der Verteilungsfunktion sowie einzelner Quantilen zu finden.

Man beachte, dass die Zufallsvariable Z nicht mehr normalverteilt ist, wenn man μ und σ durch Schätzwerte ersetzt, die resultierende Verteilung ist dann eine t-Verteilung.

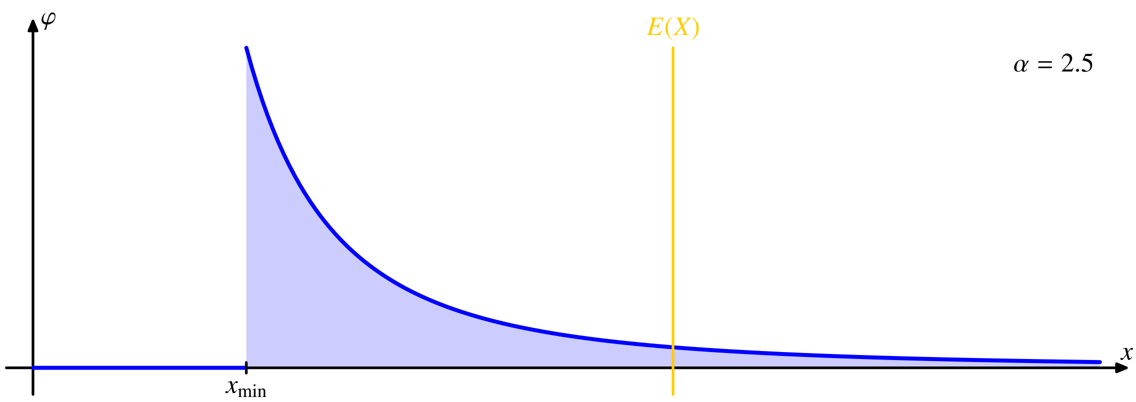
A.4 Potenzgesetz, Pareto-Verteilung

A.4.1 Steckbrief

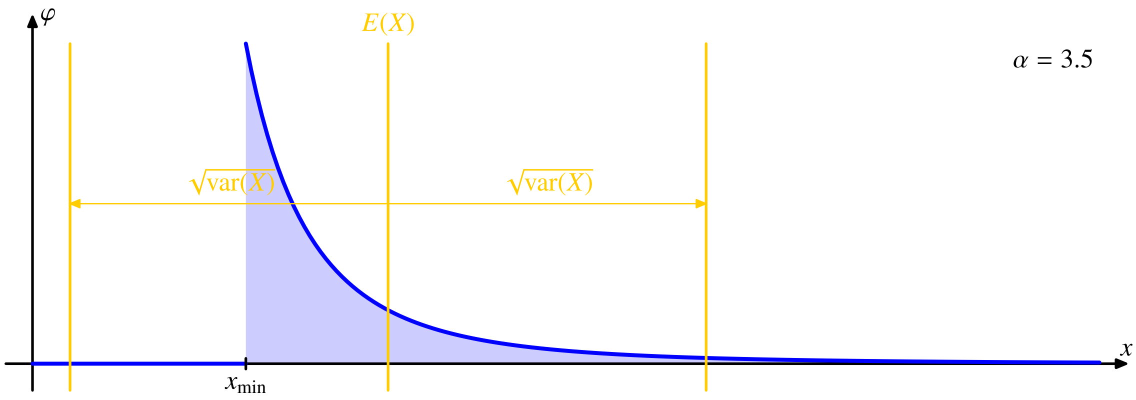
Name	Potenzverteilung, Pareto-Verteilung
Dichtefunktion	$\begin{cases} \frac{\alpha-1}{x_{\text{min}}}\left(\frac{x}{x_{\text{min}}}\right)^{-\alpha} & x > x_{\text{min}} \\ 0 & \text{sonst} \end{cases}$
Verteilungsfunktion	$\begin{cases} 1-\left(\frac{x}{x_{\text{min}}}\right)^{1-\alpha} & x > x_{\text{min}} \\ 0 & \text{sonst} \end{cases}$
Erwartungswert	$\dfrac{\alpha-1}{\alpha-2}x_{\text{min}}$, undefiniert für $\alpha \le 2$
Varianz	$\left(\dfrac{\alpha-1}{\alpha-3}-\left(\dfrac{\alpha-1}{\alpha-2}\right)^2\right)x_{\text{min}}^2$, undefiniert für $\alpha \le 3$
$P(\lvert X-E(X)\rvert > \varepsilon)$	
Median	$2^{\frac{1}{\alpha-1}}\,x_{\text{min}}$
Anwendungen	• Häufkeitsverteilung für skaleninvariante Prozesse • Einkommensverteilung • Grösse und Häufigkeit von Mondkratern • Verkaufszahlen von Büchern • Einwohnerzahlen von Städten

A.4.2 Wahrscheinlichkeitsdichte

Wahrscheinlichkeitsdichte einer Potenzverteilung mit $\alpha = 2.5$, diese Verteilung hat keine Varianz:



Wahrscheinlichkeitsdichte einer Potenzverteilung mit $\alpha = 3.5$:



A.4.3 80/20-Regeln

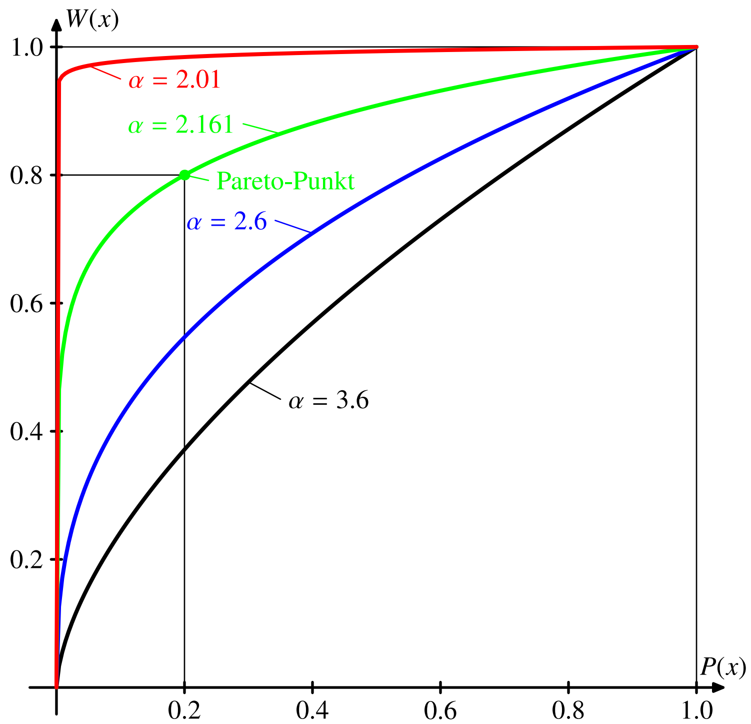
Potenzgesetze geben Anlass zu 80/20-Regeln. Wir bezeichnen die Wahrscheinlichkeit des Auftretens von Werten $x > 0$ mit $P(x) = 1 - F(x)$ und den Anteil des Erwartungswertes dieser Wert am gesamten Erwartungswert mit

$$W(x) = \frac{\int_x^\infty \xi\varphi(\xi)\,d\xi}{\int_{x_{\text{min}}}^\infty \xi\varphi(\xi)\,d\xi}.$$

Man kann $W(x)$ als den Anteil des "Wertes" interpretieren, den die Werte oberhalb von x beisteuern. Es ist klar, dass $P(x) = 1$ auch bedeutet, $W(x) = 1$: 100% der Werte tragen 100% des Wertes bei. Die Definitionen besagen, dass für beliebiges x der obere $P(x)$-Anteil der Werte den Anteil $W(x)$ des gesamten Wertes beitragen. Es gilt

$$W(x) = P(x)^{\frac{\alpha-2}{\alpha-1}}.$$

Kurven $(P(x), W(x))$ für verschiedene Werte von α:



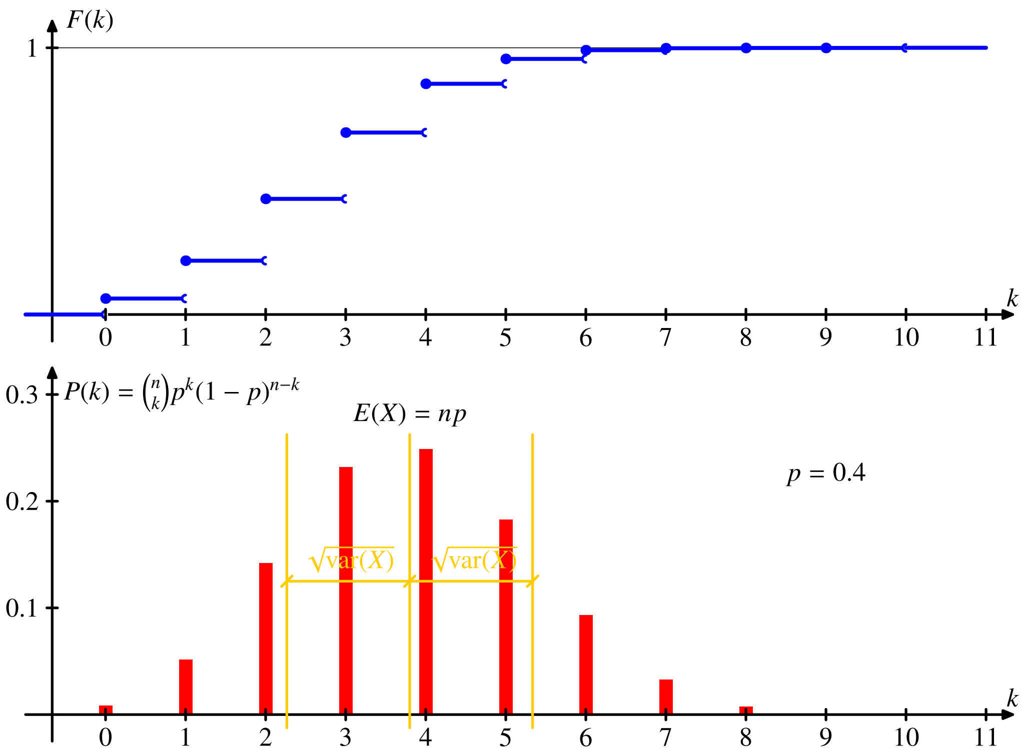
A.5 Binomialverteilung

A.5.1 Steckbrief

Name	Binomialverteilung
Wahrscheinlichkeit	$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$
Verteilungsfunktion	$F(k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$
Erwartungswert	np
Varianz	$np(1-p)$
Anwendungen	• Anzahl Eintreten eines Bernoulliexperimentes

A.5.2 Verteilungsfunktion und Wahrscheinlichkeitsverteilung

Wahrscheinlichkeitsverteilung und Verteilungsfunktion einer Binomialverteilung mit $p = 0.4$ und $n = 10$:



A.5.3 Parameter schätzen

Wir ein Bernoulliexperiment mit m Wiederholungen n mal wiederholt, erhält man n Werte $k_1, \dots, k_n$. Die Wahrscheinlichkeit p kann daraus mit Hilfe des Mittelwertes erwartungstreu geschätzt werden:

$$\hat{p}(k_1, \dots, k_n) = \frac{k_1 + \cdots + k_n}{n}.$$

Als Schätzer für den Parameter m liegt das Maximum nahe, da aber grosse Werte sehr selten sind, ist dieser Schätzer von unbrauchbarer Qualität.

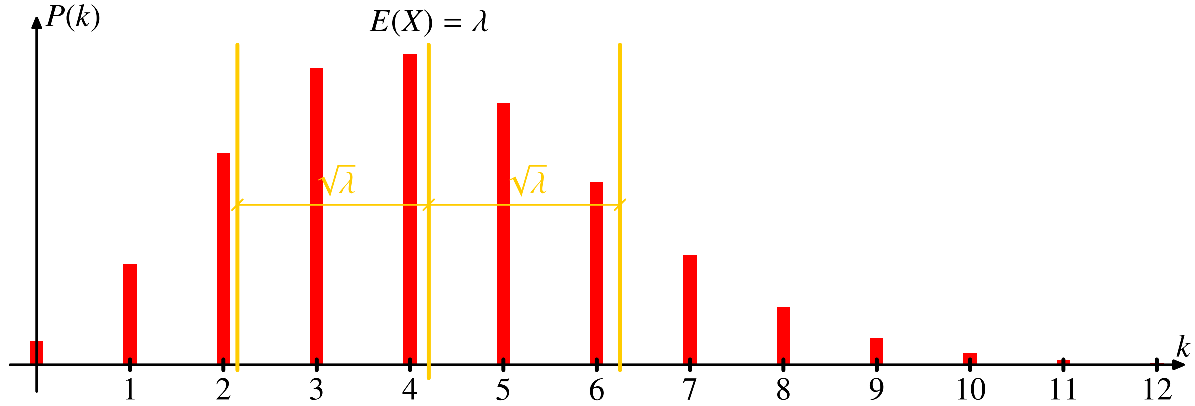
A.6 Poisson-Verteilung

A.6.1 Steckbrief

Name	Poissonverteilung
Wahrscheinlichkeit Erwartungswert Varianz	$P_\lambda(k) = \dfrac{\lambda^k}{k!}e^{-\lambda}$ λ λ
Anwendungen	• Anzahl Ereignisse mit exponentialverteilten Intervallen • Approximation der Binomialverteilung für seltene Ereignisse, die mit Rate λ eintreten

A.6.2 Wahrscheinlichkeitsverteilung

Wahrscheinlichkeitsverteilung der Poisson-Verteilung für $\lambda = 4.2$, $E(X) = \lambda$ und $\operatorname{var}(X) = \sqrt{\lambda}$.



A.6.3 Schätzung des Parameters λ

Der Parameter λ ist der Erwartungswert einer Poisson-Verteilung, er lässt sich daher mit dem Mittelwert

$$\hat{\lambda}(k_1, \dots, k_n) = \frac{k_1 + \cdots + k_n}{n}$$

erwartungstreu schätzen.