

3 Wahrscheinlichkeit

Markus Lippitz

2. Mai 2022

Ziele Sie können eine Sequenz von 'Zufallszahlen' in Julia *erzeugen*, die einer gegebenen Verteilung genügen.

- diskrete Verteilungen: Binomial, Poisson
- stetige Verteilungen: Normal, Log-Normal
- Satz von Bayes
- zentraler Grenzwertsatz

Weitere Aufgaben

- Erzeugen sie Paare von 'Zufallszahlen', deren 2D Verteilung ein Elefant beschreibt. Betrachten sie das Mittel über N dieser Paare. Vergleichen sie es mit ihrer Erwartung.

Literatur Stahel Kap. 5 & 6, Bevington Kap. 2, Meyer Kap. 22–25

Überblick

Im letzten Kapitel haben wir die Stichprobe beschrieben. In diesem Kapitel geht es um die Verteilung, aus der die Stichprobe gezogen wurde. Das letzte Kapitel beschreibt also die Messwerte, diese die Wirklichkeit. Im nächsten geht es dann darum, wie man von den Messwerten auf die Wirklichkeit kommt ('schätzt'). Die Kennzahlen sind sehr ähnlich. Im folgenden wird ein Hut über die Kennzahlen gesetzt, wenn es um die Stichprobe geht.

Sichprobe	Verteilung (diskret)	Verteilung (kont.)
relative Häufigkeit r_k	Wahrscheinlichkeit $P\langle X = k \rangle$	Wahrscheinlichkeitsdichte $f\langle k \rangle$
Mittelwert $\bar{x} = \sum k r_k$	Erwartungswert $\mathcal{E}\langle X \rangle = \mu = \sum k P\langle X = k \rangle$	Erwartungswert $\mathcal{E}\langle X \rangle = \mu = \int k f\langle k \rangle dk$
Varianz $\hat{\text{var}} = \frac{n}{n-1} \sum (k - \bar{x})^2 r_k$	Varianz $\text{var}\langle X \rangle = \sum (k - \mu)^2 P\langle X = k \rangle$	Varianz $\text{var}\langle X \rangle = \int (k - \mu)^2 f\langle k \rangle dk$
Standardabweichung $\hat{\sigma} = \sqrt{\hat{\text{var}}}$	Standardabweichung $\sigma = \sigma_X = \sqrt{\text{var}}$	Standardabweichung $\sigma = \sigma_X = \sqrt{\text{var}}$

gekürzt aus Stahel, Tabelle 5.3.b

Zwei Punkte sind hierbei relevant: beim Übergang von diskreten zu kontinuierlichen Wahrscheinlichkeitsverteilungen muss man zu einer Wahrscheinlichkeits**dichte** übergehen. Und der (bislang) etwas merkwürdige Faktor $\frac{n}{n-1}$ tritt nur bei der Beschreibung der Stichprobe auf, nicht aber bei der von Verteilungen. Das wird sich im nächsten Kapitel klären.

Grundbegriffe

Wir müssen zunächst einige Begriffe in Analogie zur Mengenlehre definieren. Für Details siehe Stahel, Kapitel 4.

Symbol	Mengenlehre	WK-Rechnung
ω	Element	Elementarereignis
A, B, \dots	Mengen	Ereignisse
Ω	Grundmenge	sicheres Ereignis
\emptyset	leere Menge	unmögliches Ereignis
$A \cup B$	Vereinigung	A oder B (oder beides)
$A \cap B$	Schnitt-M.	A und B

gekürzt aus Stahel, Tabelle 4.2.d

Wir betrachten Wahrscheinlichkeiten als relative Häufigkeiten von Ereignissen. Das ist die Laplace'sche Definition der Wahrscheinlichkeiten

$$\text{Wahrscheinlichkeiten} = \frac{\text{Anzahl günstige Fälle}}{\text{Anzahl mögliche Fälle}}$$

Damit ergeben sich für die Wahrscheinlichkeit P einige grundlegende Beziehungen. Die Wahrscheinlichkeit eines Ereignisses kann nicht negativ sei:

$$P\langle A \rangle \geq 0 \quad \text{für alle } A$$

Die Wahrscheinlichkeit, dass irgend was eintritt, ist eins:

$$P\langle \Omega \rangle = 1$$

Wahrscheinlichkeiten addieren sich, falls die Ereignisse disjunkt sind:

$$P\langle A \cup B \rangle = P\langle A \rangle + P\langle B \rangle \quad \text{falls } A \cap B = \emptyset$$

Daraus folgt

$$P\langle \emptyset \rangle = 0 \quad ; \quad P\langle A \rangle \leq 1 \quad ; \quad P\langle A^C \rangle = 1 - P\langle A \rangle$$

Beispiel Würfel

Für einen gewöhnlichen Würfel ist die Grundmenge

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Falls A die Menge der gerade Würfeln-Augen ist, dann ist $P(A) = 0.5$.

Die Beobachtung eines Elementarereignisses ω bzw. das Resultat X von einmal würfeln wird geschrieben als $X = \omega$. Die Variable X nennt man **Zufallsvariable**. Beim Spiel 'Mensch ärgere dich' ist $X = 6$ von zentraler Bedeutung, zumindest am Anfang des Spiels. Die Wahrscheinlichkeit davon wird geschrieben als $P(X = 6) = 1/6$ bzw. oft auch kurz als $p_6 = 1/6$.

Zufallszahlen

Im Computer können wir 'würfeln', indem wir (Pseudo-)Zufallszahlen erzeugen. Diese Zahlen werden deterministisch erzeugt, entsprechen aber in allen Kriterien (beinahe) echtem Zufall.

In Julia geht das mit

```
1 using Random
```

Danach bekommt man bei jedem Aufruf eine neue, quasi-zufällige Zahl, beispielsweise im Intervall $(0, 1)$

```
0.049415811293434087
```

```
1 rand()
```

Probieren Sie einmal aus, obige Zelle mehrmals auszuführen!

Wir können auch die Grundmenge vorgeben. Ein Würfel bekommt man mit

```
1
```

```
1 rand( (1,2,3,4,5,6) )
```

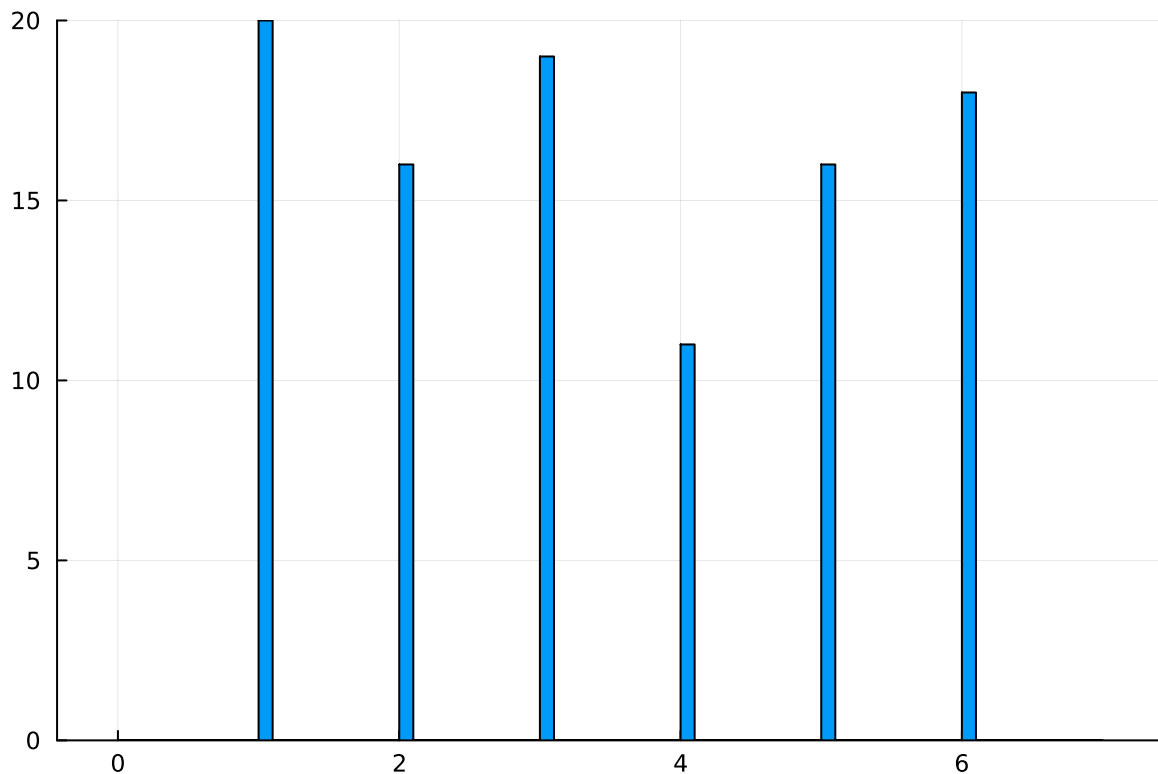
100 mal würfeln entspricht

```
hundert =
```

```
► [1, 3, 2, 2, 6, 3, 5, 4, 1, 3, 5, 2, 1, 1, 1, 1, 3, 5, 4, 4, ... more ,6, 6, 6, 5, 5, 1, 5
```

```
1 hundert= rand( (1,2,3,4,5,6) , 100)
```

```
1 using Plots, StatsBase
```

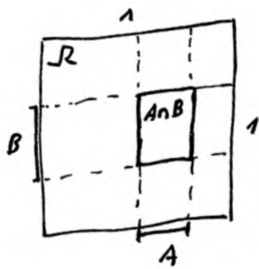


```
1 plot(StatsBase.fit(Histogram, hundert, range(0,7; step=0.1)),leg=false)
```

Wie man sieht ist der Würfel ein Beispiel für eine diskrete Zufallsvariable, die nur endlich viele Werte annehmen kann. Die Temperatur im letzten Kapitel würde man als kontinuierliche Zufallsvariable modellieren, da sie (vor der digitalen Erfassung) unendlich viele Werte annehmen kann.

Unabhängige Ereignisse

Zwei Messungen, Versuche oder Ereignisse nennt man unabhängig, wenn der Ausgang des einen keinen Einfluss auf den anderen hat. Wenn man mit zwei Würfeln würfelt, ist das Ergebnis des einen Würfels unabhängig vom anderen. Wenn man aber Kugeln aus einer Urne zieht, ohne diese wieder zurück zu legen, dann beeinflusst die erste Ziehung die Wahrscheinlichkeitsverteilung der zweiten Ziehung.



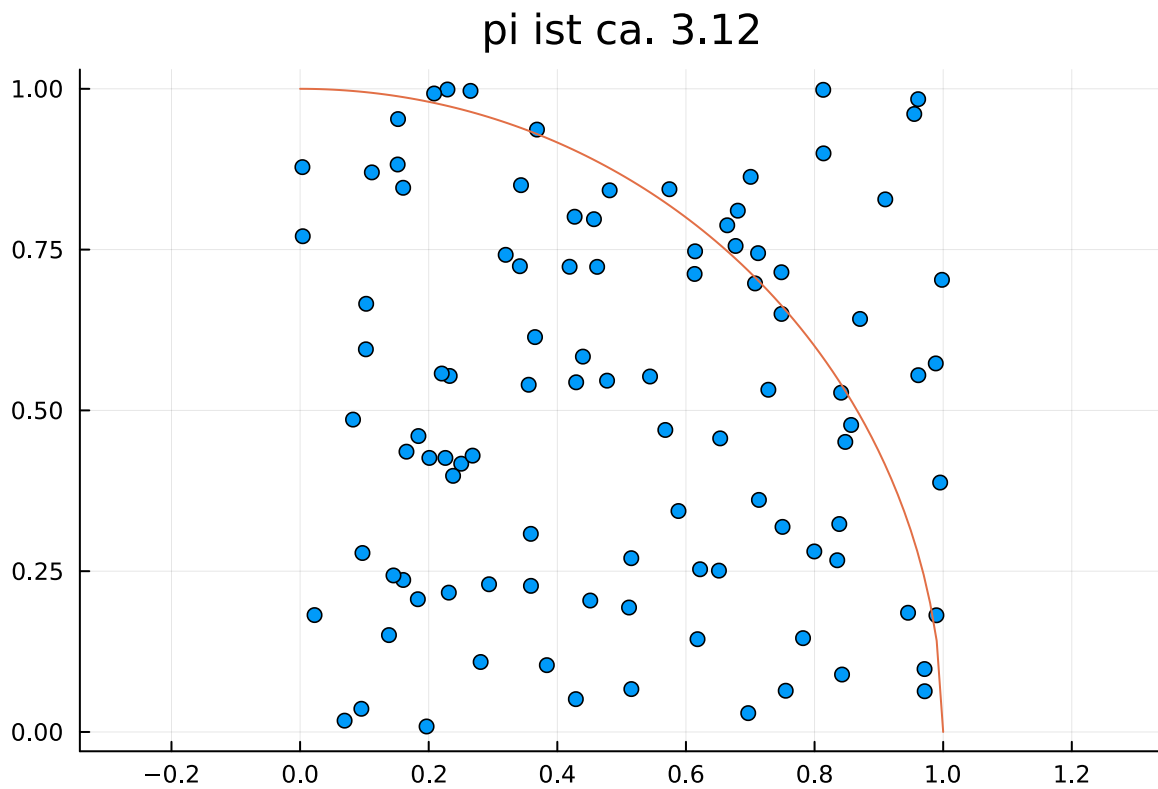
Schnittmenge zweier **unabhängiger** Ereignisse

Formal schreibt sich das als

$$P(A \cap B) = P(A) P(B)$$

Die Wahrscheinlichkeit, dass A und B eintritt, ist das Produkt der Wahrscheinlichkeiten für A und B .

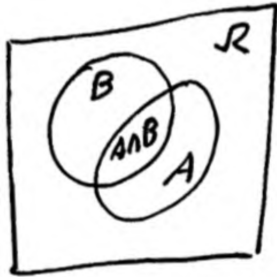
Wir können das benutzen, um π zu bestimmen: wir zeichnen zwei Zufallszahlen und verstehen sie als Koordinate in einem Einheits-Quadrat. Der Anteil der Punkte mit Abstand kleiner 1 zum Ursprung ist $\pi/4$.



```
1 begin
2   n = 100
3   x_coord = rand(Float64, n)
4   y_coord = rand(Float64, n)
5   pi_guess = 4 / n * count(x -> x < 1 , @. sqrt(x_coord^2 + y_coord^2))
6
7   scatter(x_coord, y_coord, leg=false, aspect_ratio=:equal)
8   plot!( range(0,1;length=100), sqrt.(1 .- range(0,1;length=100).^2),
9         title="pi ist ca. $(pi_guess)")
10 end
```

Bedingte Wahrscheinlichkeit

Wir betrachten zwei Ereignisse A und B aus dem gleichen Wahrscheinlichkeitsraum Ω .



Schnittmenge zweier **abhängiger** Ereignisse

Wir kennen die einzelnen Wahrscheinlichkeiten $P(A)$ und $P(B)$. Wie ändert sich die Wahrscheinlichkeit für das Ereignis B , wenn wir wissen, dass Ereignis A eingetreten ist? Dies beantwortet die bedingte Wahrscheinlichkeit $P(B|A)$ ('B unter der Voraussetzung A'). Aus geometrischen Überlegungen ergibt sich

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

falls $P(A) \neq 0$.

Wir lösen das nach $P(A \cap B)$ auf und erhalten den *allgemeinen Multiplikationssatz*

$$P(A \cap B) = P(B|A) P(A) = P(A|B) P(B)$$

Das gilt auch, falls $P(A)$ oder $P(B)$ Null sind.

Für unabhängige Ereignisse ist ja $P(A \cap B) = P(A) P(B)$, so dass

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = P(B)$$

Das Eintreten von A ändert nichts an der Wahrscheinlichkeit von B . Die bedingte Wahrscheinlichkeit ist gleich der unbedingten.

Satz von Bayes

Wir stellen den allgemeinen Multiplikationssatz um und erhalten den *Satz von Bayes*

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

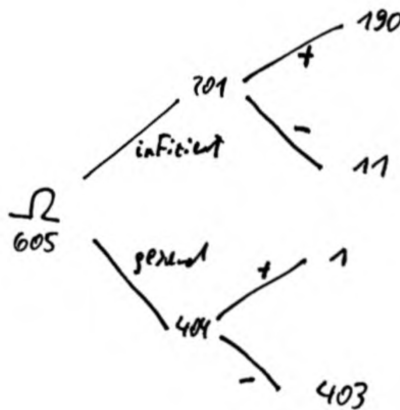
den man auch schreiben kann als

$$P(A|B) = \frac{P(B|A) P(A)}{P(B|A) P(A) + P(B|A^C) P(A^C)}$$

Die Bedeutung des Satzes von Bayes liegt darin, die bedingte Wahrscheinlichkeit umzukehren. Man kennt $P(B|A)$, ist aber eigentlich an $P(A|B)$ interessiert. Diese sind nur dann identisch, wenn A und B gleich wahrscheinlich sind.

Beispiel Corona-Test

Bei einem medizinischen Test (binärer Klassifikator) beobachtet man ein Ereignis B = 'Test positiv' und ist daran interessiert, ob jemand beispielsweise mit Corona infiziert ist (Ereignis A = 'wirklich infiziert'). Im Labor gemessen hat man das jedoch andersrum: Man wendete den Test bei vielen Leuten an, von den man wusste, ob sie infiziert sind oder nicht. Man hat also $P(B|A)$ bestimmt.



Baumdiagramm zum Corona-Test

Lassen sie uns als Beispiel den mir gerade vorliegenden Test betrachten. Im Beipackzettel findet sich das Ergebnis der Studie

	infiziert	gesund	Summe
Test +	190	1	191
Test -	11	403	414
Summe	201	404	605

Die *Sensitivität* gibt die Wahrscheinlichkeit an, mit der ein infizierter Patient korrekt als infiziert (positiv) erkannt wird (= richtig-positiv), also

$$P(\text{Test positiv}|\text{infiziert}) = \frac{P(A \cap B)}{P(B)} = \frac{190}{201} = 94.5\%$$

Die *Spezifität* gibt die Wahrscheinlichkeit an, mit der ein gesunder Patient korrekt als nicht infiziert (negativ) erkannt wird (= richtig-negativ) , also

$$P(\text{Test negativ}|\text{gesund}) = \frac{P(A \cap B)}{P(B)} = \frac{403}{404} = 99.8\%$$

Die Zahlen sehen ja alle gut aus, allerdings waren im Labor auch 201 von 605 Personen infiziert. Das sind viel mehr als in Wirklichkeit. Selbst eine kleine falsch-positiv-rate wird dominant, wenn die Infektion sehr selten ist (siehe Prävalenzfehler). Eigentlich interessiert uns ja $P(\text{infiziert} | \text{Test positiv})$. Dazu benutzen wir den Satz von Bayes und brauchen zunächst eine Annahme über die *a priori* (vorab) Wahrscheinlichkeit, infiziert zu sein. Also

$$P(\text{infiziert} | \text{positiv}) = \frac{P(\text{positiv}|\text{infiziert})P(\text{infiziert})}{P(\text{positiv}|\text{infiziert})P(\text{infiziert}) + P(\text{positiv}|\text{gesund})P(\text{gesund})}$$

Mit $P(\text{positiv}|\text{gesund}) = 1 - P(\text{negativ}|\text{gesund})$ und $P(\text{gesund}) = 1 - P(\text{infiziert})$ haben wir alles was wir brauchen.

 300

```
1 @bind inzidenz Slider(0:1000; show_value=true, default=300 )
```

0.5870780699937873

```
1 begin
2   sens = 0.945
3   spez = 0.998
4   P_infiziert = inzidenz / 1e5
5
6   # für weiter unten
7   P_infiziert_negativ = (1-sens) * P_infiziert / ((1-sens) * P_infiziert + spez*
   (1-P_infiziert))
8
9   P_infiziert_positiv = sens * P_infiziert / (sens * P_infiziert + (1-spez)*(1-
   P_infiziert))
10 end
```

Bei niedrigen Inzidenzen ist ein positiver Test also nicht gleichbedeutend mit einer Infektion. Bei hohen Inzidenzen aber schon.

Bei einer Inzidenz von 300 pro 100 000 Einwohner sind von 10 000 Personen mit positiven Testresultaten 5871.0 Personen infiziert. Gleichzeitig sind 2.0 Personen infiziert, obwohl der Test negativ ausgefallen ist.

Diskrete Wahrscheinlichkeitsverteilungen

Binomialverteilung

Wir betrachten eine binäre Zufallsvariable X , die also nur zwei verschiedene Werte annehmen kann. Ein Beispiel könnte der Wurf einer Münze sein. Die Wahrscheinlichkeit für eine Eins sein p , also $P(X = 1) = p$ und somit $P(X = 0) = 1 - p$. Dies nennt man die *Bernoulli-Verteilung*.

Wenn wir die Münze n mal werfen und die Summe der 'Augen' bilden, so können Werte zwischen 0 und n auftreten. Die Wahrscheinlichkeit für den Wert x ist

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \mathcal{B}(n, p)$$

Um die Summe x zu erhalten, müssen wir x mal die Eins ziehen und $n - x$ mal die Null. Da die Reihenfolge keine Rolle spielt geht der Binomialkoeffizient ein. Diese Verteilung nennt man daher Binomialverteilung. Sie beschreibt beispielsweise, wie viele Personen einer Gruppe in diesem Monat Geburtstag haben, wie viele von uns heute mit den linken Fuß aufgestanden sind, oder wie viele Fahrräder heute morgen einen Platten hatten. Wir zählen also 'positive' Ergebnisse in einer festgelegten Anzahl von Versuchen.

Der Erwartungswert der Binomialverteilung ist np . Die Varianz $np(1 - p)$. Die Summe von zwei Binomialverteilungen mit gleichem p ist wieder eine Binomialverteilung mit der Summe der n_i . Falls sich die p_i unterscheiden ist die Summe keine Binomialverteilung.

In Julia können wir beispielsweise auf das Distributions-Paket zurückgreifen. Zur graphischen Darstellung zeigen wir die probability density function (pdf).

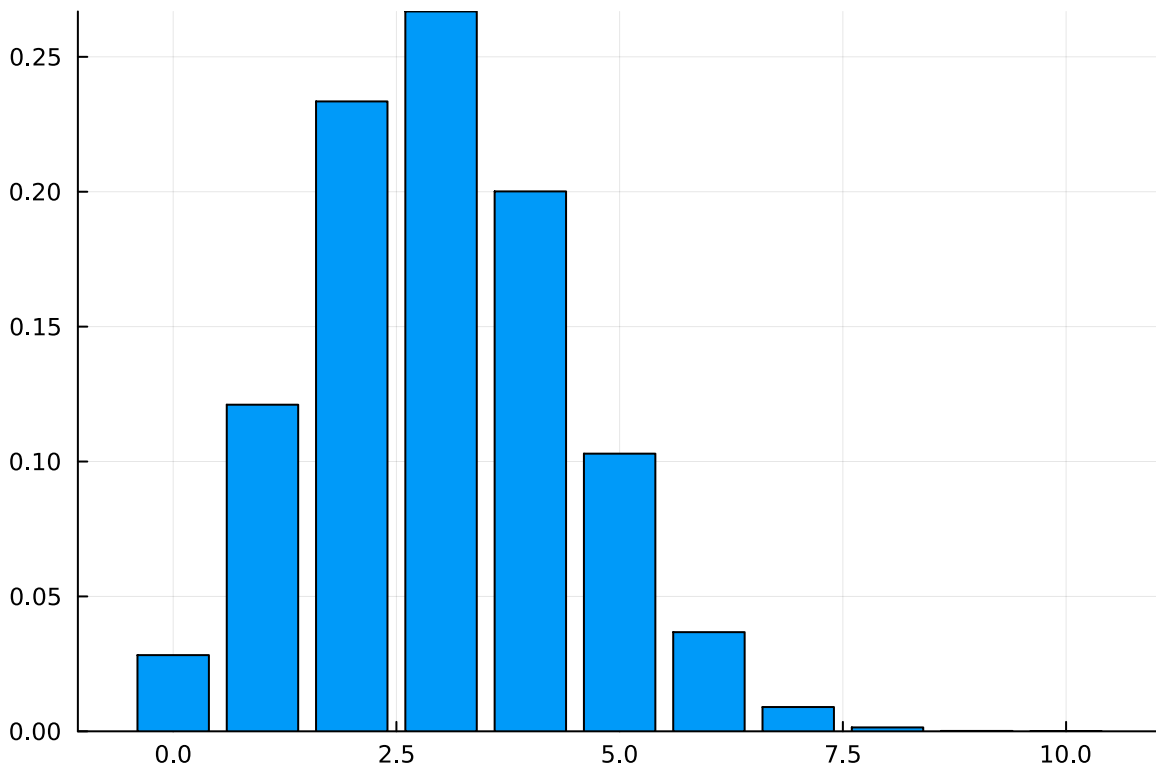
```
1 using Distributions
```

```
▶ [0.00604662, 0.0403108, 0.120932, 0.214991, 0.250823, 0.200658, 0.111477, 0.0424673, 0
```

```
1 Distributions.pdf(Distributions.Binomial(10, 0.4))
```

0.3

```
1 @bind p_binomial Slider(0:0.01:1; default=0.3, show_value=true)
```

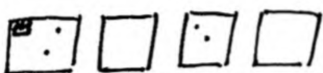


```
1 Plots.bar((0:10), pdf(Binomial(10, p_binomial)), leg=false)
```

Poisson-Verteilung

In der Poisson-Verteilung zählen wir auch Ereignisse, allerdings ohne dass die Anzahl der Versuche n der Binomialverteilung eine Rolle spielt. Ein Beispiel sind Klicks eines Geigerzählers pro Zeitintervall oder Regentropfen pro Fläche.

Wir folgen hier dem Regentropfen-Beispiel von Stahl und betrachten eine Kachel der Größe Eins. Auf dieser Kachel markieren wir einen Bereich der Fläche $\lambda < 1$. Der erste Regentropfen fällt auf diese Kachel. Die Wahrscheinlichkeit, dass er im markieren Bereich landet, beträgt λ .



Kacheln und Regentropfen

Nun gehen wir zu n Kacheln über und lassen insgesamt genauso viele Tropfen auf die Kacheln fallen. Die Anzahl Treffer im markierten Bereich der ersten Kachel ist wie oben binomialverteilt, nämlich

$$\mathcal{B}(n, \lambda/n) = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Die Idee hinter ' n Tropfen auf n Kacheln' ist, dass wir im Mittel immer gleich lang warten müssen, bis ein Tropfen den markierten Bereich trifft.

Die *Poisson-Verteilung* ist der Grenzwert davon für $n \rightarrow \infty$

$$\mathcal{P}(\lambda) = \lim_{n \rightarrow \infty} \mathcal{B}(n, \lambda/n) = \frac{\lambda^x}{x!} e^{-\lambda}$$

Beispiele sind alle vom Typ 'viele (n groß) seltene (λ/n klein) Ereignisse', z.B. Anzahl der Staubpartikel in der Luft oder Anzahl der Photonen in einem Laserstrahl.

Der Erwartungswert der Poissonverteilung und die Varianz sind beide λ . Die Summe von zwei Poissonverteilungen mit den Parametern λ_1 und λ_2 ist wieder einer Poissonverteilung mit dem Parameter $\lambda_1 + \lambda_2$.

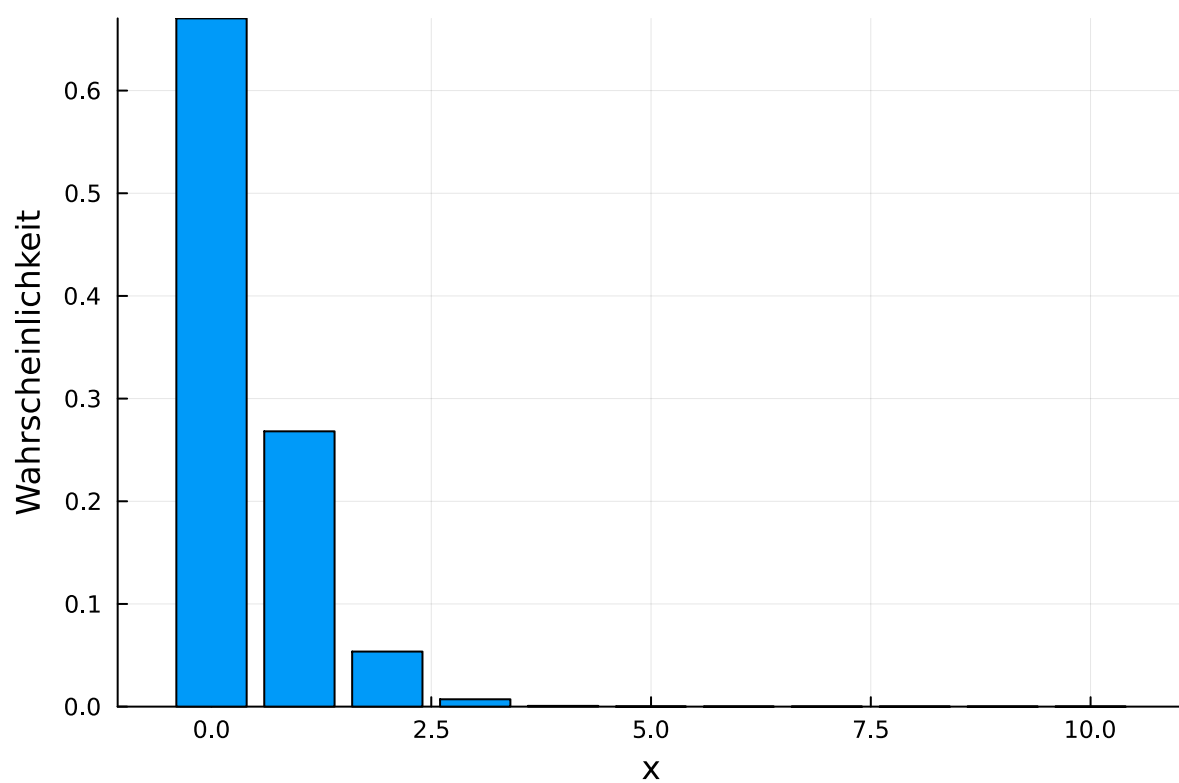
Die Poisson-Verteilung reicht bis ins Unendliche, ist also für jeden ganzzahligen Wert x definiert. Daher müssen wir bei der Berechnung der probability density function ein Intervall mit angeben. Für große λ (ab ca. 10) geht sie in die Normalverteilung über (siehe unten).

```
► [0.272532, 0.354291, 0.230289, 0.0997921, 0.0324324, 0.00843243]
```

```
1 @. pdf(Poisson(1.3), (0:5))
```

 0.4

```
1 @bind λ_poisson Slider(0:0.2:5; default=0.4, show_value=true)
```



```
1 let
2   x = (0:10)
3   h = @. pdf(Poisson( $\lambda_{\text{poisson}}$ ), x)
4   bar(x,h; legend=false, xlabel="x", ylabel="Wahrscheinlichkeit")
5 end
```

Kontinuierliche Verteilungen

Bei kontinuierlichen Verteilungen macht die Beschreibung durch die Wahrscheinlichkeit für ein bestimmtes Ergebnis x wenig Sinn, da der Wert x nie wirklich exakt erreicht wird. $P(X = x)$ ist nicht definiert oder Null. Wir betrachten daher die *kumulative Verteilungsfunktion*

$$F(x) = P(X \leq x)$$

und es gilt natürlich

$$P(a < X \leq b) = F(b) - F(a)$$

Damit können wir aber eine *Wahrscheinlichkeitsdichte* f definieren mit

$$f(x) = F'(x)$$

die an die Stelle der alten Wahrscheinlichkeit P tritt.

Exponentialverteilung

Die Exponentialverteilung beschreibt die Wartezeit zwischen Ereignissen, die der Poisson-Verteilung folgen. Beispiele sind also der Abstand zwischen zwei Klicks des Geigerzählers oder zwischen zwei Photonen eines Laserstrahls. Sie hat die Form

$$F(x) = 1 - e^{-\lambda x} \quad \text{für } x \geq 0 \quad \text{sonst} \quad = 0$$

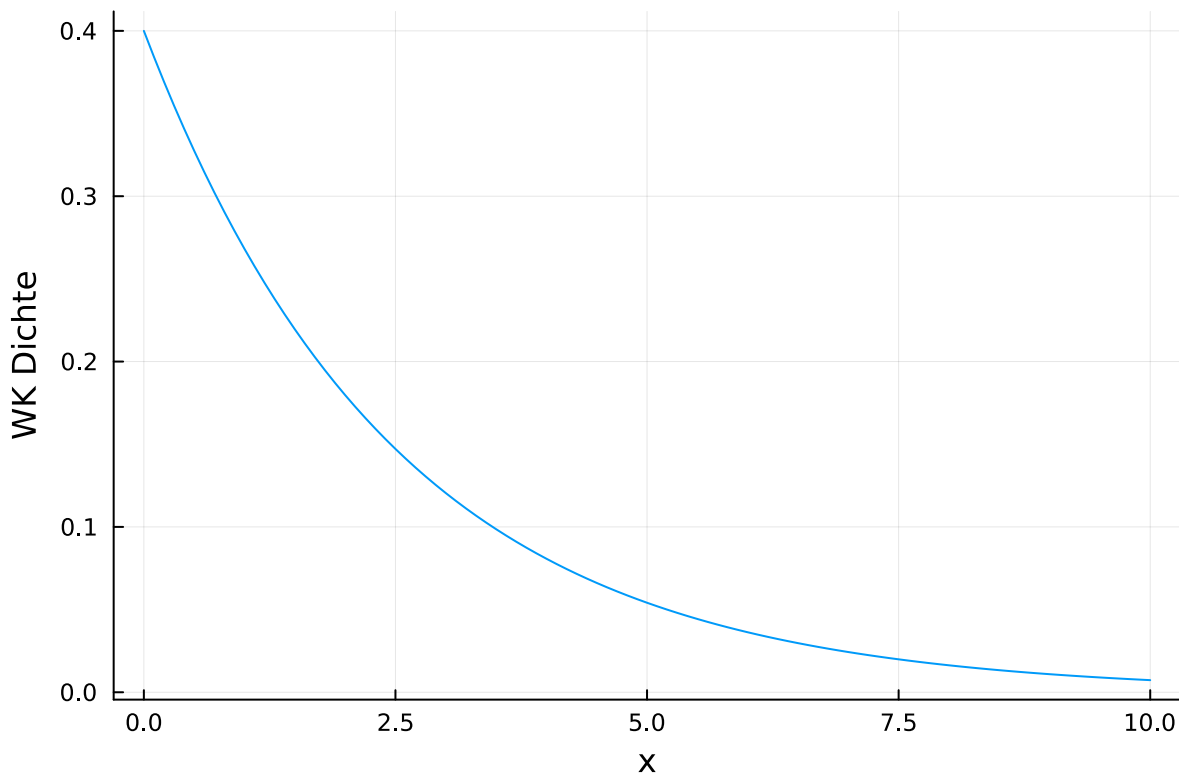
Ihre Dichte ist

$$f(x) = \lambda e^{-\lambda x} \quad \text{für } x \geq 0$$

Der Erwartungswert der Exponentialverteilung ist $1/\lambda$, die Varianz $1/\lambda^2$.

 0.4

```
1 @bind p_exponential Slider(0.1:0.1:1 ; default=0.4, show_value=true)
```

```
1 let
2   x = range(0, 10; step=0.1)
3   # the parameter is 1/λ
4   plot(x, pdf.(Exponential(1/p_exponential), x), legend=false, xlabel="x",
5         ylabel="WK Dichte")
6 end
```

Normalverteilung

Die Normalverteilung hat die Form einer Gauss-Funktion. Für die *Standard-Normalverteilung* ist die Wahrscheinlichkeitsdichte um die Null zentriert mit der Varianz 1, also

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Andere Formen der Normalverteilung erhält man durch lineare Transformation. Sei Z eine Zufallsvariable, die der Standard-Normalverteilung folgt, dann transformiert

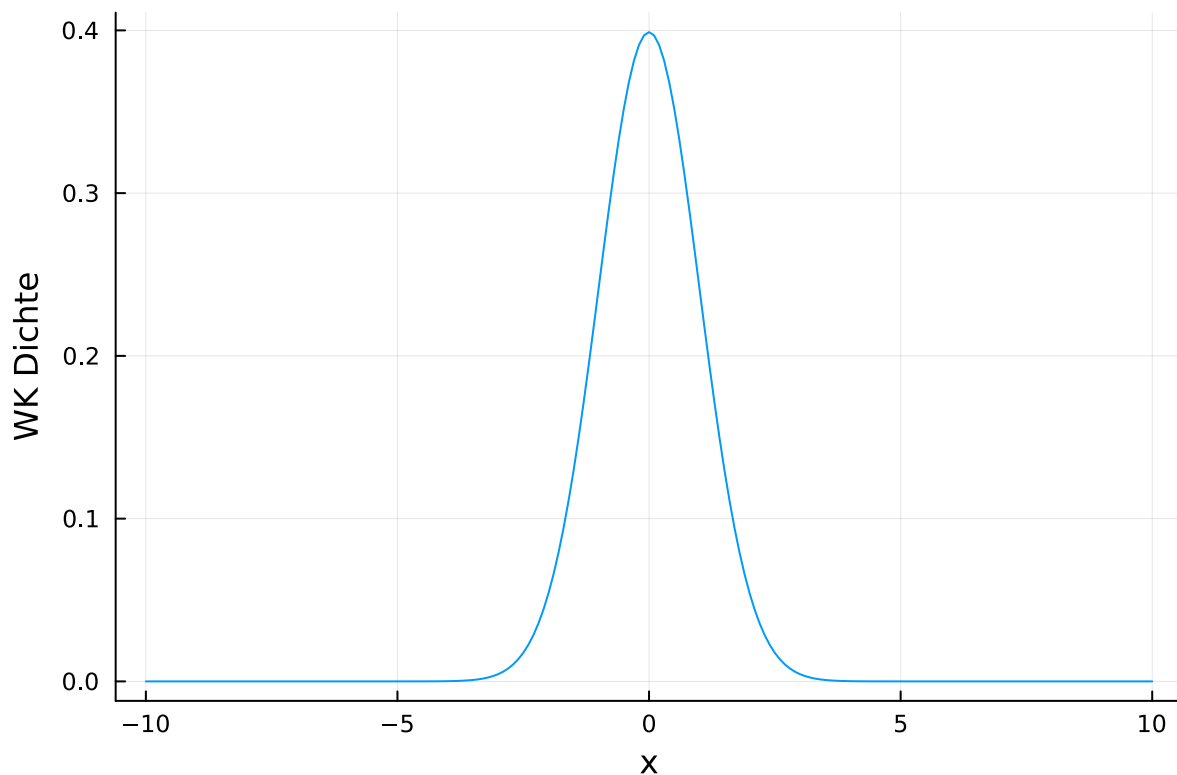
$$X = \mu + \sigma Z$$

hin zu einer Normalverteilung mit dem Erwartungswert μ und der Standardabweichung σ . Die Wahrscheinlichkeitsdichte ist dann also

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

 1.0

```
1 @bind sigma_normal Slider(0.5:0.5:4, ; default=1, show_value=true)
```



```
1 let
2   x = range(-10, 10; step=0.1)
3   plot(x, pdf.(Normal(0, sigma_normal), x), legend=false, xlabel="x",
4         ylabel="WK Dichte")
5 end
```

Die kumulative Verteilungsfunktion F der Normalverteilung, also

$$F(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy$$

nennt sich *error function*. Das Integral lässt sich nicht analytisch lösen. Wir brauchen es allerdings häufig, insbesondere um zu sagen, mit welcher Wahrscheinlichkeit ein Messwert innerhalb eines gewissen Intervalls liegt.

In Intervall $\mu \pm 1, 2$ oder 3σ liegen 68.3%, 95.5% oder 99.7% der Ereignisse. Oder in Julia

► (0.682689, 0.9545, 0.9973)

```
1 let
2     sigmas = (1,2,3)
3     @. cdf(Normal(0, 1), sigmas) - cdf(Normal(0, 1), - sigmas)
4 end
```

Summen von Normalverteilungen sind wieder eine Normalverteilung. Dabei addieren sich die Mittelwerte und die Varianzen (nicht die Standardabweichungen), also

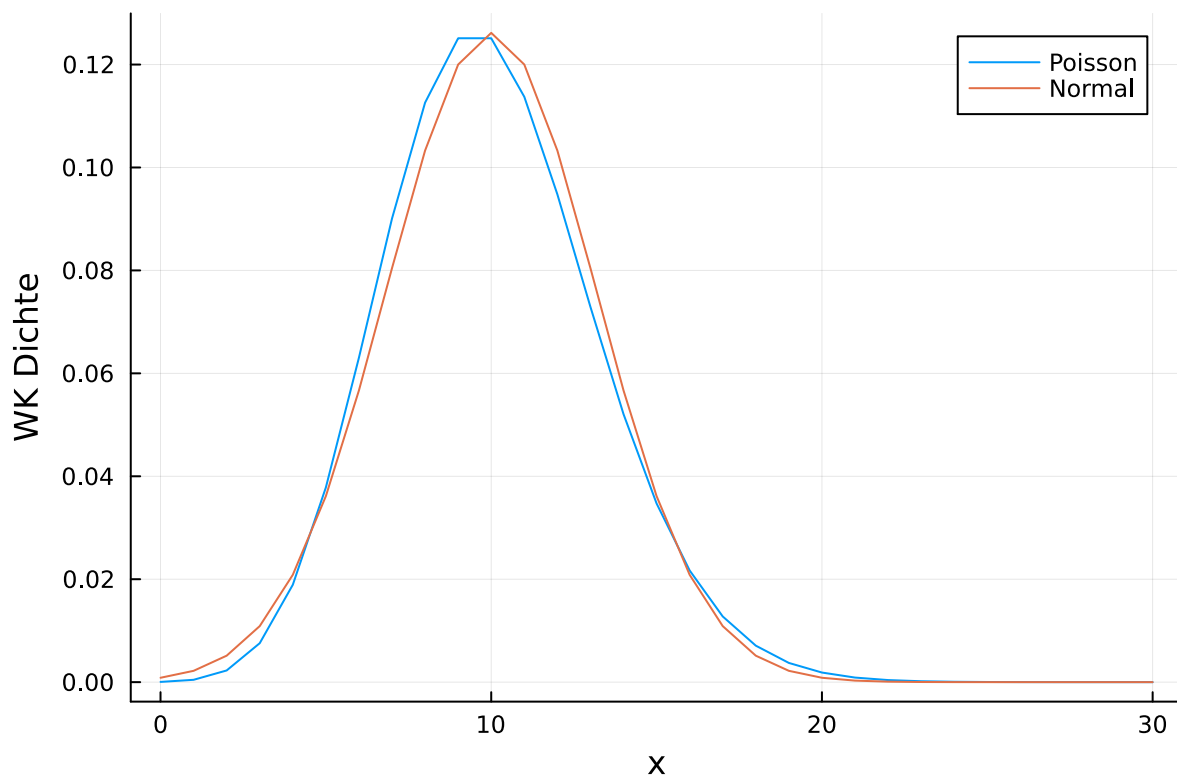
$$\mu_{sum} = \mu_1 + \mu_2 + \dots \quad \text{und} \quad \sigma_{sum}^2 = \sigma_1^2 + \sigma_2^2 + \dots$$

Beziehung zur Poisson-Verteilung

Die Poisson-Verteilung besitzt nur einen Parameter λ . Für große Werte von λ geht sie in eine Normalverteilung mit dem Mittelwert $\mu = \lambda$ und der Standard-Abweichung $\sigma = \sqrt{\lambda}$ über.

 10

```
1 @bind lambda_vgl Slider(0:30; show_value=true, default=10)
```



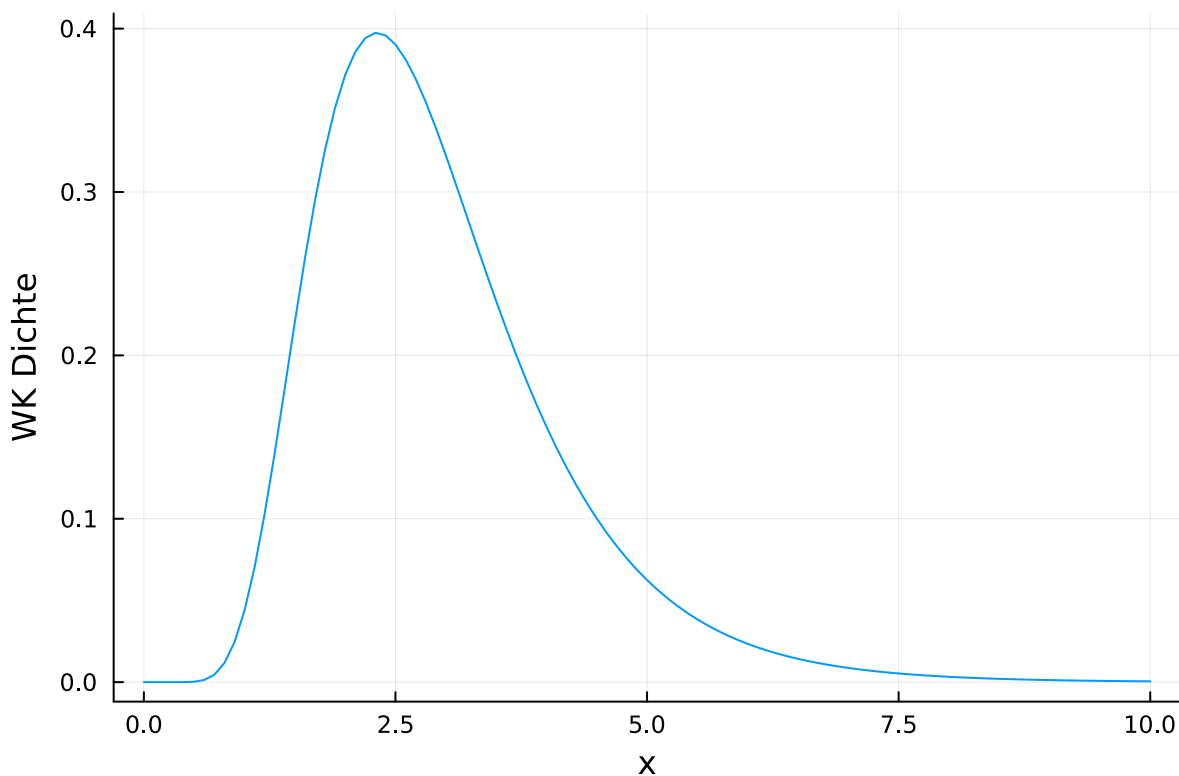
```

1 let
2   x = range(0, 30)
3   plot(x, pdf.(Poisson(lambda_vgl), x) , label="Poisson")
4   plot!(x, pdf.(Normal(lambda_vgl, sqrt(lambda_vgl)), x), label="Normal",
5         xlabel="x", ylabel="WK Dichte")
6 end

```

Log-Normalverteilung

Bei der Log-Normalverteilung ist der Logarithmus der Werte normalverteilt. Negative Werte treten also nicht auf und die Verteilung wird (auf einer linearen Achse) asymmetrisch.



```
1 let
2   x = range(0, 10; step=0.1)
3   plot(x, pdf.(LogNormal(1, 0.4), x), legend=false, xlabel="x", ylabel="WK
4   Dichte")
end
```

Funktionen von Zufallsvariablen

Nehmen wir an, dass wir wissen, eine Zufallsvariable X entspringe einer der obigen Verteilungen. Welchen Aussagen können wir dann über die Verteilung der Werte von $g(X)$ machen? Wir wenden also die Funktion g auf jede gezogenen Wert an und bilden daraus eine neue Verteilung. Die Log-Normalverteilung oben ist ein Beispiel: wir ziehen Zahlen aus einer Normalverteilung und wenden die Exponentialfunktion an.

Der Erwartungswert von $g(X)$ ist näherungsweise die Funktion g ausgewertet am Erwartungswert von X , also

$$\mathcal{E}(g(X)) \approx g(\mathcal{E}(X))$$

Die Beschreibung der Varianz von $g(X)$ ist Inhalt der **Gauss'schen Fehlerfortpflanzung**. Es gilt

$$\text{var}(\alpha + \beta X) = \beta^2 \text{var}(X) \quad \text{bzw} \quad \sigma_{\alpha + \beta X} = \beta \sigma_X$$

bzw. allgemein für eine Variable

$$\sigma_{g(X)} \approx |g'(\mu)| \sigma_X$$

Für Funktionen mehrere *unabhängiger* Variablen gilt

$$\sigma_{g(X,Y)} \approx \sqrt{\left(\frac{\partial g}{\partial X} \sigma_X\right)^2 + \left(\frac{\partial g}{\partial Y} \sigma_Y\right)^2}$$

Im nächsten Kapitel werden wir darauf genauer eingehen.

Zentraler Grenzwertsatz

Warum ist die Normalverteilung 'normal'? Die Normalverteilung ist dadurch ausgezeichnet, dass sie im Grenzwert der Summe von vielen Zufallsvariablen entspricht, egal welcher Verteilung diese Zufallsvariablen entstammen.

Seien X_1, X_2, X_i, \dots Zufallsvariablen aus der gleichen Verteilung mit dem Erwartungswert μ und der Standardabweichung σ . Diese Verteilung kann quasi beliebige Form haben. Wir betrachten als neue Zufallsvariable den Mittelwert über die ersten n X_i und nennen diese \bar{X} . Aus der Gauss'schen Fehlerfortpflanzung kennen wir Erwartungswert $\mathcal{E}(\bar{X})$ und Standardabweichung $\sigma_{\bar{X}}$, nämlich

$$\mathcal{E}(\bar{X}) = \mu \quad \text{und} \quad \sigma_{\bar{X}} = \sigma_X / \sqrt{n}$$

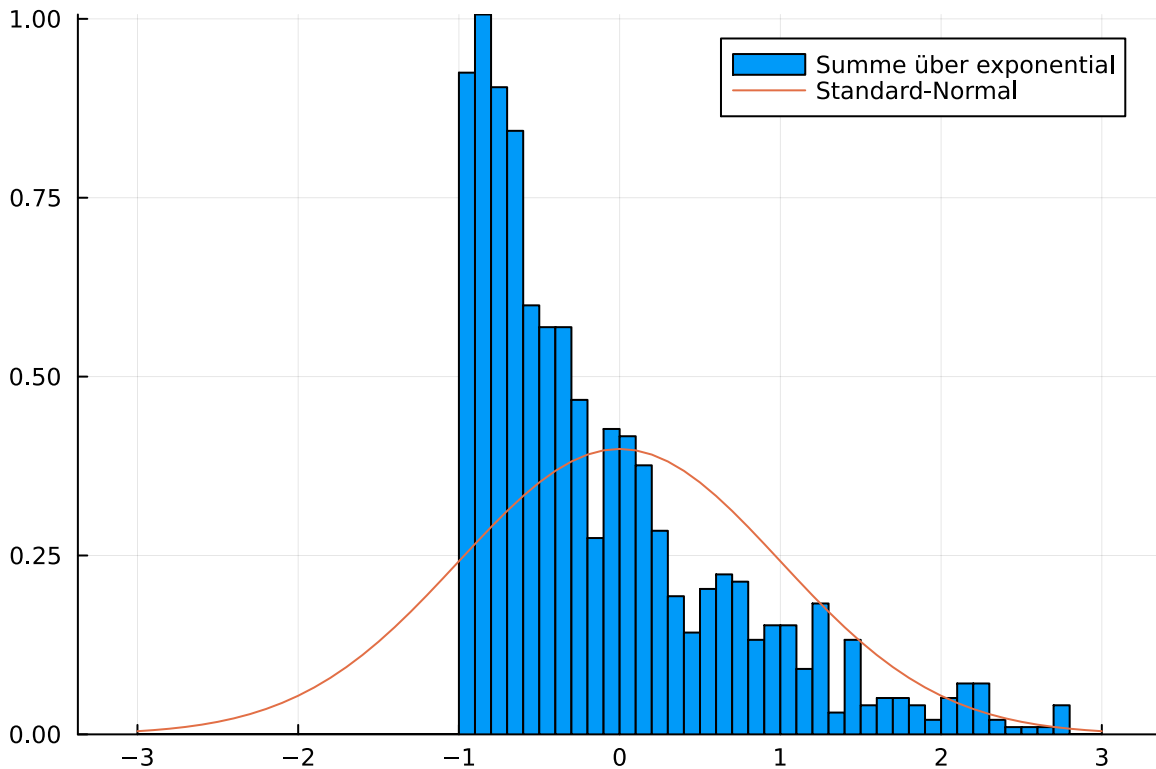
Jetzt interessiert uns die Form der Verteilung unserer neuen Zufallsvariable \bar{X} . Dazu standardisieren wir diese, skalieren also so zu einer dritten Zufallsvariablen Z , dass Erwartungswert Null und Standardabweichung Eins werden

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Der Zentrale Grenzwertsatz besagt, dass sich die Verteilung von Z mit steigender Mittelungszahl n immer weiter einer Standard-Normalverteilung annähert.

1

```
1 @bind n_mittel Slider(1:10; show_value=true)
```



```
1 let
2   N_Z = 1000 # Anzahl Punkte in Z-Verteilung
3   x = range(-3,3; step=0.1)
4
5   zufall = rand(Exponential(1), N_Z, n_mittel) # exponentialverteilt, λ=1
6   Xbar = sum(zufall; dims=2) # sum along n_mittel
7   Z = (Xbar .- n_mittel) ./ sqrt(n_mittel) # hier ist σ=1 und μ=n_mittel
8
9   h1 = StatsBase.fit(Histogram, vec(Z), x) # histogram
10  h1 = LinearAlgebra.normalize(h1, mode=:pdf) # normieren auf pdf
11  plot(h1, label="Summe über exponential") # plotten
12
13  plot!(x, pdf.(Normal(),x), label="Standard-Normal") # Normalverteilung zum
14  Vgl
end
```

```
1 using PlutoUI, LinearAlgebra
```

```
1 aside(s) = PlutoUI.ExperimentalLayout.aside(s);
```

☰ Inhalt

Überblick

Grundbegriffe

Beispiel Würfel

Zufallszahlen

Unabhängige Ereignisse

Bedingte Wahrscheinlichkeit

Satz von Bayes

Beispiel Corona-Test

Diskrete Wahrscheinlichkeitsverteilungen

Binomialverteilung

Poisson-Verteilung

Kontinuierliche Verteilungen

Exponentialverteilung

Normalverteilung

Beziehung zur Poisson-Verteilung

Log-Normalverteilung

Funktionen von Zufallsvariablen

Zentraler Grenzwertsatz

1 `TableOfContents(title="Inhalt")`