Wahrscheinlichkeit

Mathematische Bissen zu Kursvorlesung Theoretische Physik

Im Gegensatz zum Ideal der Newton'schen Mechanik oder der Maxwellschen Elektrodynamik können in Wirklichkeit Ergebnisse von Messungen nur mit bestimmten Wahrscheinlichkeiten (kurz: W'keiten) vorhergesagt werden. Man verfügt schlicht nicht über die Mittel alle Freiheitsgrade, die das interessierende System beeinflussen, kontrollieren zu können. Schlimmer noch: selbst wenn wir über diese Mittel verfügten, macht uns die Quantenmechanik einen Strich durch die Rechnung: die Quantenmechanik ist eine probabilistische Theorie, und jegliches Bemühen den Zufall aus quantenmechanischen Experimenten zu eliminieren, ist fruchtlose Zeitverschwendung. Zeit also, den W'keitsbegriff kurz zu erläutern.

1 W'keitsraum

In der W'keitstheorie wird jedem Experiment (mit einem Würfel, oder einem Glücksrad) eine Menge Ω zugeordnet, genannt *Elementarereignisraum*, oder *Stichprobenraum* (engl. sample space). Im einfachsten Fall ist Ω eine endliche Menge, beim einfachen Wurf eines Würfel etwa $\Omega = \{$⚀,⚁,⚂,⚃,⚄,⚅$\}$ – die Menge der möglichen oben liegenden Seiten nach dem Wurf.[1]

Das Experiment (der Wurf) wird weiterhin durch die Angabe eine Wahrscheinlichkeit $p(\omega)$ für jedes Elementarereignis $\omega \in \Omega$ charakterisiert. Ein fairer Würfel, beispielsweise, ist charakterisiert durch $p(\omega) = 1/6$ für alle $\omega \in \Omega$, aber es gibt bekanntlich auch gezinkte Würfel. Die einzige Einschränkung an die $p(\omega)$ sind (i) $p(\omega) \geq 0$, und (ii) $\sum_{\omega\in\Omega} p(\omega) = 1$. Experimente mit fairen Würfeln nennt man *Laplace-Experimente*. Laplace-Experimente sind in der W'keittheorie von hervorragneder Bedeutung.

Eine *Observable* wird beschrieben durch eine Funktion $F : \Omega \to \mathbb{R}$. Die Quadrat-Zahl der Augen, beispielsweise, ist eine Observable, also $F($⚀$) = 1, F($⚁$) = 4$ usw. In der Sprache der W'keitstheorie heißen die F *Zufallsvariable*, zuweilen stochastische Variable. Der Wert $F(\omega)$ wird als derjenige Messwert interpretiert, den man abläse wenn das Elementarereignis ω einträte. Der Mittelwert ist definiert

$$\langle F\rangle = \sum_{\omega\in\Omega} p(\omega)F(\omega)\,, \tag{1}$$

wobei die spitzen Klammern $\langle\cdot\rangle$ *Erwartungswert* genannt werden, zuweilen auch Mittelwert.

Ganz wichtig sind Zufallsvariable, hier genannt χ, die nur die Werte 1 oder 0 annehmen können. Solche binäre Zufallsvariable sind eindeutig charakterisiert durch die Angabe eine Teilmmenge $S \subset \Omega$ auf der sie den Wert 1 annehmen,

$$\chi_S(\omega) = \begin{cases} 1 & \omega \in S \\ 0 & \text{sonst}\,. \end{cases} \tag{2}$$

[1]Auch wenn Sie versucht sind, die oben liegenden Seiten eines Würfels mit Zahlen zu identifizieren – tun Sie es nicht: Schliesslich kann man die oben liegenden Seiten eines Würfels nicht addieren, man denke an Farbenwürfel.

Funktionen $\chi_S(\cdot)$ heißen *Indikator-Funktionen.* Eine Teilmenge $S \subset \Omega$ nennt man auch ein *Ereignis* (nicht *Elementar*ereignis), und die Menge aller Teilmengen von Ω, die sog. Potenzmenge von Ω – bezeichnet $\mathcal{P}(\omega)$ heisst in der W'keitstheorie *Ereignisraum* (nicht *Elementar*ereignisraum). Die Teilmenge $S = \{$⚁,⚃,⚅$\}$, beispielsweise, steht für das Ereignis “die Zahl der gewürfelten Augen ist gerade”. Würfelt man beispielsweise eine ⚃, ist dieses Ereignis eingetreten. Bei einer ⚁ ist es nicht eingetreten.

Die mit Abstand wichtigsten Ereignisse sind das Ereignis Ω – interpretiert “irgendein Ereignis (egal welches)” – und $\emptyset$, interpretiert “das unmögliche Ereignis (das niemals eintritt)” – beim Wrfeln z.B. die Augenzahl 7.

Der Erwartungswert $\langle \chi_S \rangle$

$$P(S) := \langle \chi_S \rangle = \sum_{\omega} p(\omega) \chi_S(\omega) = \sum_{\omega \in S} p(\omega)\,. \tag{3}$$

ist die W'keit für das Ereignis S. Für disjunkte Teilmengen $S_i \subset \Omega$ gilt offensichtlich $P(\cup_i S_i) = \sum_i P(S_i)$, also ist P ein Maß. Und da obendrein $P(S) \geq 0$ sowie $P(\emptyset) = 0$ und $P(\Omega) = 1$ ist P ein *Wahrscheinlichkeitsmaß.*

Man kann das bisher gesagt zusammenfassen. Ein W'keitsexperiment ist charakterisiert durch seinen Elementarereignisraum Ω, einer darüber gebildeten σ-Algebra $\mathcal{A}$ – im Falle des Würfels ist das genau die Potenzmenge $\mathcal{P}(\Omega)$ – und ein auf $\mathcal{A}$ definiertes W'keitsmaß P. Das Tupel $(\Omega, \mathcal{A}, P)$ heißt in der Mathematik auch ein *W'keitsraum.*

Beim Glücksrad (ohne Teilungsnägel) ist die Sache schon schwieriger. Der Stichprobenraum darf hier mit dem Intervall $\Omega := [0, 2\pi]$ identifiziert werden: Elementarereignisse sind reellwertige Winkeln $\varphi \in \Omega$, und derer gibt es bekanntlich überabzählbar unendlich viele. Das Wahrscheinlichkeitsmaß ist nun sicher nicht schon durch die W'keiten $p(\varphi) = P(\{\varphi\})$ bestimmt: die W'keit, irgendeinen Winkel, beispielsweise $\varphi = \pi/e$, genau zu treffen, ist eben Null. Allerdings – und das ist der Grund für die Unterteilung mit Nägeln – ist die W'keit, in irgendeinem “endlichen” Intervall zu landen nicht Null. Ein faires Glücksrad, beispielsweise, ist definiert durch das W'keitsmaß $P(d\varphi) = d\varphi/(2\pi)$ wobei $d\varphi$ hier das Ereignis (offene Teilmenge) $(\varphi, \varphi + d\varphi) \subset \Omega$ bezeichnet. Offensichtlich $\int P(d\varphi) = 1$ – irgendwo auf dem Kreis wird man sicherlich landen.

Das faire Glücksrad vermittelt ein Beispiel für ein absolut stetiges Maß auf $[0, 2\pi] \subset \mathbb{R}$. Allgemein sind solche Maße gegeben durch eine *(W'keits)Dichte* $\varrho : \Omega \to \mathbb{R}$, wobei $\varrho(\varphi) \geq 0$ und $\int d\varphi \varrho(\varphi) = 1$. Damit W'keitsmaß $P(d\varphi) = \varrho(\varphi) d\varphi$, und Erwartungswert einer Observablen $F : \Omega \to \mathbb{R}$ entsprechend $\langle F \rangle = \int d\varphi F(\varphi) \varrho(\varphi)$.

Wahrscheinlichkeiten erfahren eine natürliche Interpretation durch sog. *relative Häufigkeiten.* Würfelt man in einer Versuchsreihe N mal, und wertet ein bestimmtes Ereignis ω als “S-Treffer” falls $\omega \in S$, so ist die relative Häufigkeit die in der t-ten Versuchreihe aus den Messdaten (gewürfelten Augenzahlen) ermittelt wird

$$r(S; N, t) = \frac{[\text{Zahl der } S\text{-Treffer}]_t}{N}\,. \tag{4}$$

Das große Versprechen der Natur ist nun, daß im Grenzfall $N \to \infty$ die $r(S; N, t)$ mit Sicherheit einen von t unabhängigen (also objektiven) Wert $P(S)$ annehmen,

$$\lim_{N \to \infty} r(S; N, t) = P(S)\,. \tag{5}$$

Wir glauben der Natur, müssen aber darauf hinweisen daß diese Konstruktion philosophisch durchaus umstritten ist. "Mit Sicherheit" heisst genauer "mit an Sicherheit grenzender W'keit", und damit bewegt sich der Versuch, W'keiten über relative Häufigkeiten zu definieren im Kreis.

Auch sind W'keitsaussagen grundsätzlich weder verifizierbar noch falsifizierbar. Das Versprechen "Dieser Würfel ist nicht gezinkt" behaupet unter anderem $p(5) = 1/6$. Haben Sie Pech (oder Glück?), und würfeln permanent eine 5, könnte man meinen, Sie hätten die behauptete Fairness des Würfels widerlegt. Haben Sie aber nicht. Schließlich haben Sie sicherlich nicht unendlich oft gewürfelt. Finden Sie andererseits bei einer langen Messreihe eine relative Häufigkeit $r = 1/6$ ist damit auch nichts bewiesen: schließlich könnte der Würfel trotzdem gezinkt sein.

Die Mathematik braucht sich um derartige Einwände nicht zu kümmern. "W'keitstheorie" ist aus ihrer Sicht angewandte Maßtheorie, die durch eine Liste spezieller Axiome, den sog. Kolmogoroffschen Axiomen, genauestens präzisiert wird. Für den Physiker bleibt bei der probabilistischen Quantenmechanik allerdings ein Rest Unbehagens, muss er sich doch aus dem Paradies unverrückbarer Gewissheiten verabschieden. Trotzdem – auch der mit Unbehagen geplagte Physiker tut gut daran, bei einer beoachteten relativen Häufigkeit $r = 1/6$ auf einen fairen Würfel zu setzen. Eine gegenteilige Wette könnte seine bürgerliche Existenz schnell ruinieren. Mit Sicherheit.

2 Binomialverteilung

Ein Teilchen hüpft mit W'keit p nach rechts (Elementarereignis $\mathtt{R}$) und mit W'keit $q := 1-p$ nach links (Elementarereignis $\mathtt{L}$). Stichprobenraum ist hier $\Omega = \{\mathtt{R}, \mathtt{L}\}$, der Ereignisraum $\mathcal{P}(\Omega) = \{\emptyset, \{\mathtt{R}\}, \{\mathtt{L}\}, \{\mathtt{R}, \mathtt{L}\}\}$, und das W'keitsmaß auf $\mathcal{P}(\Omega)$ ist durch die Angabe von p bereits vollständig charakterisiert,

$$P(\emptyset) = 0\,, \quad P(\{\mathtt{R}\}) = p\,, \quad P(\{\mathtt{L}\}) = 1-p\,, \quad P(\{\mathtt{R}, \mathtt{L}\}) = 1\,. \tag{6}$$

Sind dem Teilchen zwei Schritte erlaubt – ist das Experiment also auch ein anderes – gibt es offensichtlich 4 verschiedene Resultate. Beispielsweise kann das Teilchen beide Schritte nach rechts machen mit dem Resultat $\mathtt{RR}$, oder den ersten Schritt nach Rechts, den zweiten aber nach links mit dem Resultat $\mathtt{RL}$ usw. Der Stichprobenraum ist in diesem Fall das kartesische Produkt $\Omega^2 := \Omega \times \Omega$ – die Menge aller geordneten Sequenzen $s_1 s_2$ worin s_i einer der Schritte $\mathtt{L}$ oder $\mathtt{R}$.

Für ein Teilchen das N Schritte machen kann ist der Stichprobenraum das N-fache kartesiche Produkt $\Omega^N := \Omega \times \cdots \times \Omega$. Jedes Elementarereignis $\omega \in \Omega^N$ kann mit einem bestimmten Pfad identifiziert werden, beispielsweise

$$\underbrace{\mathtt{LLRL} \cdots \mathtt{RL}}_{\text{N Schritte}} \tag{7}$$

Da die einzelnen Schritte hier zufällig genommen werden, heißt so ein Pfad auch *Zufallspfad*. Unter den Annahme, dass jeder Schritt unabhängig von der vorangegangenen Schrittfolge "ausgewürfelt" wird, und dass die $\mathtt{R}$-W'keit immer die Gleiche für jeden einzelnen Schritt, ist die W'keit, einen bestimmten Pfad ω zu realisieren, gegeben $p^n(1-p)^{N-n}$, worin n die Gesamtzahl der Schritte `R`, und $N-n$ entsprechend die Gesamtzahl der Schritte `L`.

Meistens interessiert man sich aber nicht für einen bestimmten Pfad, sondern für eine Klasse von Pfade. Eine solche Klasse wäre zum Beispiel die Klasse S_n^N aller Pfade mit insgesamt n Schritten nach rechts und dementsprechend $N-n$ Schritten nach links, wobei Reihenfolge der Schritte beliebig ist. Jeder einzelne Pfad in dieser Klasse tritt mit der gleichen W'keit $p^n(1-p)^{N-n}$ auf, und da die Klasse insgesamt $\binom{N}{n}$ Mitglieder umfasst ist die W'keit, dass ein Pfad in die Klasse S_n^N fällt, gegeben

$$P_N(n) = \binom{N}{n} p^n (1-p)^{N-n} \,. \tag{8}$$

Die hier eingeführte Verteilung $P_N(n)$ heißt *Binomialverteilung* und spielt in der W'keitstheorie eine wichtige Rolle. Der binomische Lehrsatz

$$\sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} = (p+q)^N \tag{9}$$

in Verbindung mit $q = 1-p$ beweist, dass die Binomialverteilung korrekt normiert ist, $\sum_{n=0}^N P_N(n) = 1$.[2]

Die Binomialverteilung ist die W'keitsverteilung für die Zufallsgröße "Gesamtzahl der Schritte nach rechts", hier bezeichnet $\hat{n}$ (der Hut hat nichts mit Quantenmechanik zu tun). Momente der Binomialverteilung

$$\langle \hat{n}^\nu \rangle := \sum_{n=0}^{N} n^\nu \binom{N}{n} p^n (1-p)^{N-n} \tag{10}$$

lassen sich mit einem "Ableitungstrick" leicht berechnen

$$\langle \hat{n}^\nu \rangle = \sum_{n=0}^{N} n^\nu \binom{N}{n} p^n q^{N-n} \bigg|_{q=1-p} = \left[\left(p \frac{\partial}{\partial p} \right)^\nu \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n} \right]_{q=1-p} \tag{11}$$

$$= \left[\left(p \frac{\partial}{\partial p} \right)^\nu (p+q)^N \right]_{q=1-p} \tag{12}$$

Die ersten beiden Momente berechnen sich zu

$$\langle \hat{n} \rangle = Np \,, \qquad \langle \hat{n}^2 \rangle = N^2 p^2 + Np(1-p) \,, \tag{13}$$

und also die *mittlere quadratische Schwankung*

$$\langle (\hat{n} - \langle \hat{n} \rangle)^2 \rangle \equiv \langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2 = Np(1-p) \,. \tag{14}$$

Die relative Häufigkeit das Symbol `R` in einer gegebenen Sequenz anzutreffen ist eine Zufallsvariable, definiert $\hat{r} := \frac{\hat{n}}{N}$. Offensichtlich $\langle \hat{r} \rangle = p$: der relative Häufigkeit für einen

[2] Im Anschluss an den vorhergehenden Abschnitt wäre die Binomialverteilung zu notieren $P(S_n^N)$. Der Nachweis der korrekten Normierung wäre dann mengentheoretisch zu führen: Für $n \neq m$ ist $S_n^N \cap S_m^N = \emptyset$, und da $\cup_{n=0}^N S_n^N = \Omega^N$ ist $\sum_{n=0}^N P(S_n^N) = 1$.

Schritt nach rechts ist im Mittel genau die W'keit, einen Schritt nach rechts zu machen! Für die Varianz ergibt sich $\mathrm{Var}(r) = \sqrt{\frac{p(1-p)}{N}}$. Für lange Sequenzen $N \to \infty$ geht die Varianz nach Null $\sim 1/\sqrt{N}$, sprich: bei langen Sequenzen gleicht die relative Häufigkeit der `R`-Schritte mit hoher W'keit ihrem Mittelwert p. Kurz und knackig: W'keit (hier: p) ist relative Häufigkeit (hier: $\hat{r}$). Sie dürfen sich das ruhig so merken, sollten aber die genaue Bedeutung verstanden haben ...