Erwartungswert einer Zufallsvariablen

Definition des Erwartungswerts

Die Definition des Erwartungswerts (einer Zv.) lässt sich auf einfache Weise plausibel machen, wenn man vom arithmetischen Mittel ausgeht.

■ Arithmetisches Mittel einer Stichprobe

Wir denken uns eine Stichprobe aus lauter quantitativen Merkmalen gegeben. Ihr arithmetisches Mittel errechnet sich dadurch, dass man die Summe ihrer Elemente durch die Stichprobenlänge dividiert.

Das arithmetische Mittel einer Stichprobe lässt sich auch auffassen als <u>das mit den relativen Häufigkeiten ihrer Merkmale gewichtete Mittel aller Merkmale</u>.

Man fasse in der Stichprobensumme gleiche Merkmale zusammen. Jedes Merkmal wird dann mit seiner absoluten Häufigkeit multipliziert. Aus dem jeweiligen Faktor wird nach Division durch die Stichprobenlänge die relative Häufigkeit des Merkmals.

Der Sachverhalt soll noch formal hingeschrieben werden: Sei X eine quantitative Zufallsvariable mit dem Merkmalraum $\Omega = \{x_1, \ldots, x_m\}$ und S eine Stichprobe der Länge s. Wir betrachten die relativen Häufigkeiten h_i aller in Betracht kommenden Merkmale:

$$h_i := H_S(x_i) \quad (i = 1, 2, \ldots, m)$$

Merkmal x_i kommt in S offenbar $s \cdot h_i$-mal vor ($i = 1, 2, \ldots, m$). Daher ist $s\,h_1\,x_1 + \ldots + s\,h_m\,x_m$ gleich der Summe aller Merkmale, die in S aufgelistet sind. Dividiert man diese Summe durch die Länge s der Stichprobe, so erhält man ihr arithmetisches Mittel μ:

$$\mu = h_1\,x_1 + \ldots + h_m\,x_m$$

■ Der Erwartungswert als Verallgemeinerung des arithmetischen Mittels

Ersetzen wir in der zuletzt gewonnenen Darstellung von μ die relativen Häufigkeiten durch die Wahrscheinlichkeiten der W-Verteilung von X – genauer: durch $p_i = P(X = x_i)$ – so erhalten wir den **Erwartungswert von** X:

$$E(X) = p_1\, x_1 + \ldots + p_m\, x_m$$

Auch für den Erwartungswert $E(X)$ ist die Abkürzung μ gebräuchlich.

N.B.: Bei abzählbar-unendlichem Merkmalraum $\Omega = \{x_1, x_2, x_3, \ldots\}$ ist eine unendliche Summe (Reihe) $E(X) = \sum_{i=1}^{\infty} p_i\, x_i$ auszuwerten (vgl. mittlere Wartezeit).

Zwei einfache Beispiele

■ Mittlere Augenzahl beim Spielwürfel

Die Augenzahl X eines guten Spielwürfels hat den Merkmalraum $\Omega = \{1, 2, \ldots, 6\}$ und die W-Verteilung $p_i = \frac{1}{6}$ $(i = 1, 2, \ldots, 6)$. Für den Erwartungswert ergibt sich:

$$E(X) = 1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{6} + \ldots + 6 \cdot \tfrac{1}{6} = 3.5$$

Offenkundig ist der Erwartungswert einer Zvn. nicht notwendig auch ein Wert, den die Zv. tatsächlich annehmen kann.

Ein Massenversuch (500 Würfe) möge das Ergebnis von der frequentistischen Seite her beleuchten:

```
<< Statistics`DescriptiveStatistics`
<< Modellbildung`Zufallsversuche`
```

```
sprobe = ZieheMitZuruecklegen[{1, 2, 3, 4, 5, 6}, 500];
Mean[sprobe] // N
```

```
3.436
```

■ Mittlerer Gewinn beim Münzwurf

Beim Wurf einer Laplace-Münze werde 1 € gesetzt. Erscheint "Wappen", erhält man den doppelten Einsatz zurück (andernfalls ist der Einsatz verloren).

Aus der Sicht eines Spielers ist die Zv. $X = Gewinn$ von Interesse. Wir berechnen ihren Erwartungswert. "Gewinn" meint hier Reingewinn; bei "Wappen" erhält der Spieler 2 €, nach Abzug des Einsatzes bleibt ihm daher 1 € als Gewinn. Daher gilt:

$$E(X) = 1 \cdot \tfrac{1}{2} + (-1) \cdot \tfrac{1}{2} = 0$$

Für Spieler (und ebenso Gegenspieler) entsteht bei dieser Wette keine positive (oder negative) Gewinnerwartung. Man spricht von einem fairen Spiel.

Da in einer Versuchsreihe das aus den tatsächlich realisierten Häufigkeiten gebildete arithmetische Mittel um den (theoretischen) Erwartungswert herum schwankt, können für einzelne Spielteilnehmer sehr wohl Gewinn- oder Verlustsituationen eintreten (auch bei nicht-fairen Spielen). Darin liegt der Reiz von Glücksspielen.

```
sprobe = ZieheMitZuruecklegen[{-1, 1}, 500];
Mean[sprobe] // N
```

```
-0.028
```

Grundlegende Eigenschaften

Wir beweisen im Folgenden zwei grundlegende Eigenschaften des Erwartungswerts von reellwertigen Zufallsvariablen über demselben Merkmalraum.

■ *E* ist homogen

Für alle reellen c gilt: $E(c\,X) = c\,E(X)$

Beweis: Sei $p_i = P(X = x_i)$ für $i = 1, 2, \ldots, m$. Auch das Ereignis $c\,X = c\,x_i$ hat die Wahrscheinlichkeit p_i. Mithin gilt: $E(c\,X) = p_1(c\,x_1) + \ldots + p_m(c\,x_m) = c(p_1\,x_1 + \ldots + p_m\,x_m) = c\,E(X)$. ■

■ *E* ist additiv

$$E(X + Y) = E(X) + E(Y)$$

Beweis: Als elementare Merkmale von $X + Y$ wählen wir sämtliche Summen $x_i + y_i$ von Merkmalen von X und von Y, die einen Wert von $X + Y$ darstellen ($i = 1, 2, \ldots, n$). Sei p_i die Wahrscheinlichkeit, mit der die i-te Summe realisiert wird. Dann ergibt sich:

$$E(X + Y) = p_1(x_1 + y_1) + \ldots + p_n(x_n + y_n) = (p_1\,x_1 + \ldots + p_n\,x_n) + (p_1\,y_1 + \ldots + p_n\,y_n)$$

In dem Ausdruck $p_1\,x_1 + \ldots + p_n\,x_n$ sind nun eventuell noch Summanden mit gleichen Merkmalwerten x_i zusammenzufassen, um einzusehen, dass $E(X) = p_1\,x_1 + \ldots + p_n\,x_n$. Entsprechendes gilt für die zweite Klammer: $E(Y) = p_1\,y_1 + \ldots + p_n\,y_n$. ■

Eine homogene und additive Operation heißt **linear**. Der Erwartungswert ist ein linearer Operator.

Mittlere Trefferzahl im Bernoulli-Versuch

Ein N-stufiger Versuch mit einem (p, q)-Glücksrad werde durchgeführt. Die Zv. X_i stehe für das Ergebnis beim i-ten Teilversuch ($i = 1, 2, \ldots, N$). Die Trefferzahl wird dann durch $X = X_1 + \ldots + X_N$ dargestellt. Ihre W-Verteilung ist die Binomialverteilung. Wir berechnen den Erwartungswert von X (die "mittlere" Trefferzahl).

Wegen der Additivität von E hat man zunächst

$$E(X) = E(X_1 + \ldots + X_N) = E(X_1) + \ldots + E(X_N)$$

In allen Teilversuchen haben Treffer dieselbe Wahrscheinlichkeit p, d.h. es gilt $P(X_i = 1) = p$ und damit für die Erwartungswerte:

$$E(X_i) = 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) = 0 \cdot q + 1 \cdot p = p$$

Insgesamt erhalten wir daraus:

$$\mu = E(X) = N \cdot p$$

Zugleich ist dies der Erwartungswert einer binomialverteilten Zvn.

Die mittlere Trefferzahl μ liegt in der Nähe des wahrscheinlichsten Wertes $x_{\max}$ von X. Genauer gilt: $\mu - q \leq x_{\max} \leq \mu + p$. – (Beweis als Übung)

Mittlere Wartezeit

Beim Wartezeiten-Versuch wird ein (p, q)-Glücksrad mit $p > 0$ solange gedreht, bis ein Treffer erzielt ist. Für die Wartezeit Z gilt: $P(Z = n) = q^{n-1}\, p$. Wir berechnen die mittlere Wartezeit $E(Z)$:

$$E(Z) = \sum_{n=1}^{\infty} n \cdot P(Z = n) = \sum_{n=1}^{\infty} n\, q^{n-1}\, p = p \sum_{n=1}^{\infty} n\, q^{n-1}$$

Die rechterhand verbliebene Summe $\sum_{n=1}^{\infty} n\, q^{n-1}$ ist noch auszuwerten:

$$1 + 2\,q + 3\,q^2 + \ldots = \frac{1}{(1-q)^2}$$

Man sieht dies ein, indem man die Gleichung mit $1 - q$ multipliziert. Links vom Gleichheitszeichen bleibt dann die unendliche geometrische Reihe $1 + q + q^2 + \ldots$ stehen, die wegen $q < 1$ zum Wert $\frac{1}{1-q}$ konvergiert. Beachtet man $1 - q = p$, so ergibt sich insgesamt:

$$\mu = E(Z) = \frac{1}{p}$$

Für eine Münze (Warten auf "Wappen", $p = \frac{1}{2}$) gilt: $\mu = 2$; für einen Spielwürfel (Warten auf eine Sechs, $p = \frac{1}{6}$) gilt: $\mu = 6$.