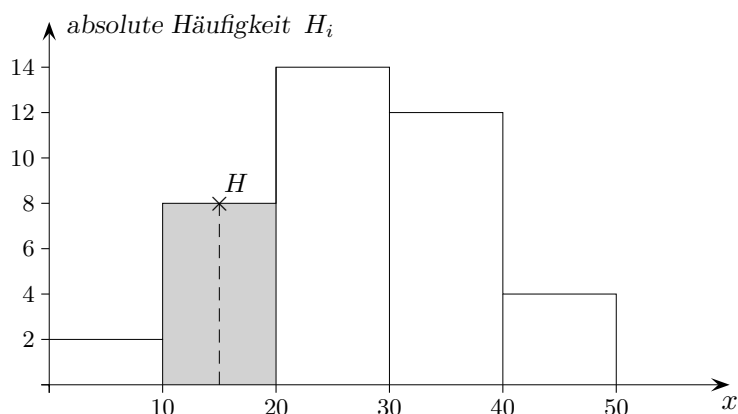


Mittelwert einer in Klassen eingeteilten Stichprobe

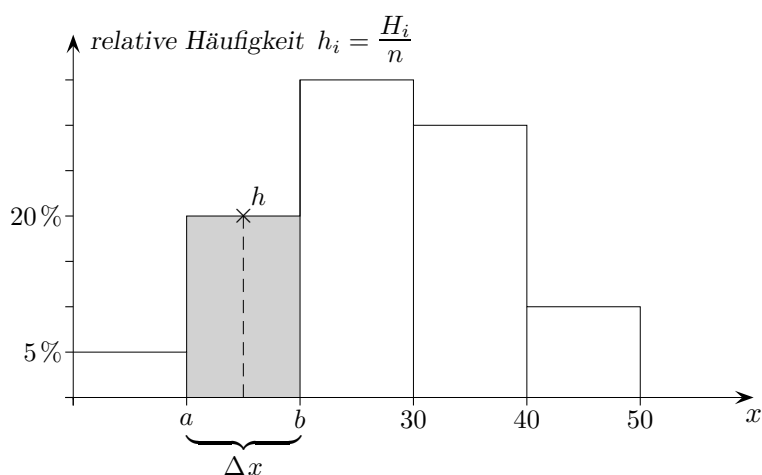
Mit einer Stichprobe sollen die Längen von Fischen eines Gewässers ermittelt werden. Um die Daten übersichtlich darzustellen, wird eine Klasseneinteilung vorgenommen. Der Mittelwert (der Klassenmitten) beträgt 27.

Länge (cm)	$x < 10$	$10 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 40$	$40 \leq x < 50$
absolute Häufigkeit	2	8	14	12	4



Um die Wahrscheinlichkeiten abzuschätzen, mit denen Fische einer bestimmten Länge gefangen werden können, werden die relativen Häufigkeiten berechnet. Wie wird mit ihnen der Mittelwert bestimmt?

Länge (cm)	$x < 10$	$10 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 40$	$40 \leq x < 50$
relative Häufigkeit	5%	20%	35%	30%	10%



Wichtige Begriffe:

statistische Erhebung, Grundgesamtheit,

Merkmal, Merkmalsausprägung, Stichprobe,

Ausreißer, Stichprobenumfang, repräsentative Stichprobe,

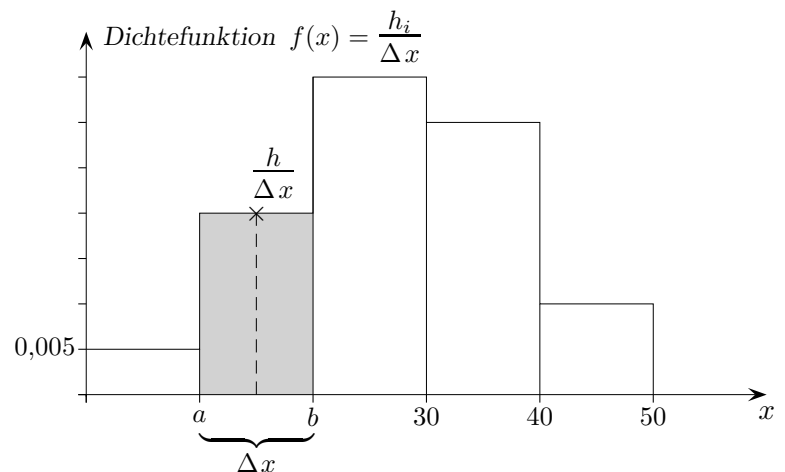
absolute und relative Häufigkeit, Klasseneinteilung, Histogramm,

Mittelwert (arithmetisches Mittel), Zentralwert (Median), Streuung um den Mittelwert

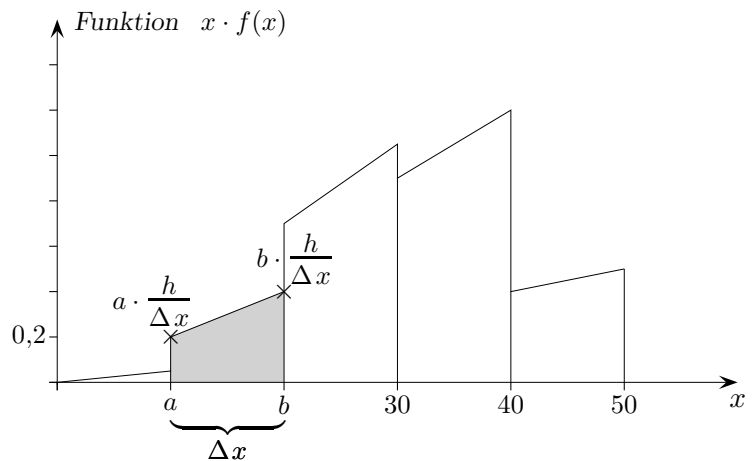
Erwartungswert einer stetigen Zufallsvariablen

Der Bezug zu stetigen Zufallsvariablen wird durch die Dichtefunktion hergestellt.
Mit ihr können Wahrscheinlichkeiten durch die Berechnung von Flächen ermittelt werden.

Länge (cm)	$x < 10$	$10 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 40$	$40 \leq x < 50$
Dichte	0,005	0,020	0,035	0,030	0,010



Der Erwartungswert kann als Inhalt der Fläche unter dem Graphen der Funktion $x \cdot f(x)$ berechnet werden.



Es gilt:
$$A_{\text{Trapez}} = \frac{a \cdot \frac{h}{\Delta x} + b \cdot \frac{h}{\Delta x}}{2} \cdot \Delta x = \frac{a+b}{2} \cdot h$$

Der letzte Term ist von der Mittelwertsberechnung mit relativen Häufigkeiten bekannt.

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx$$
 sollte nun verständlich sein.

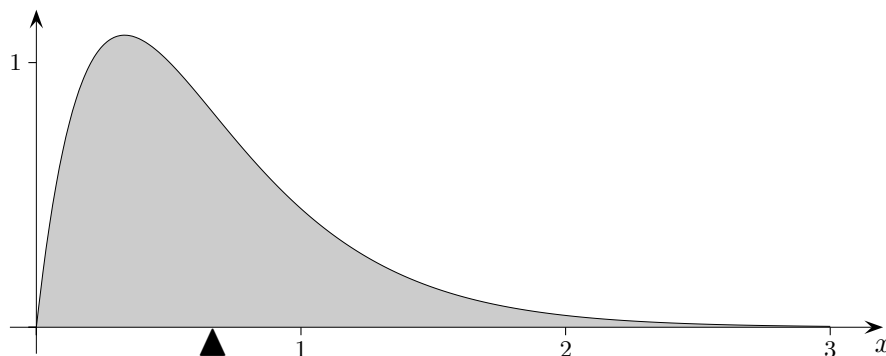
Erwartungswert $E(X)$ einer stetigen Zufallsvariablen

Die Formel

$$E(X) = \int x \cdot f(x) dx$$

liegt nahe (das Integral erstreckt sich über den Definitionsbereich der Dichtefunktion).

Hierzu sei an den Zusammenhang von Erwartungswert und Schwerpunkt einer Massenverteilung bei einer diskreten Zufallsvariablen erinnert. Eine Approximation der Fläche unter der Dichtefunktion f durch Rechtecke und eine Grenzwertbetrachtung führen zu der genannten Formel.



Für die x -Koordinate eines Flächenschwerpunkts gilt allgemein:

$$x_s = \frac{\int x \cdot f(x) dx}{\int f(x) dx}$$

Bei Dichtefunktionen entfällt der Nenner (warum)?

Zusammengefasst:

Der Erwartungswert einer diskreten Zufallsvariable X (anschaulich Durchschnittsgewinn), die die Werte a_i annimmt, lautet:

$$\mu = E(X) = \sum_{i=1}^m a_i \cdot P(X = a_i) \quad (\text{Jeder Wert } a_i \text{ wird mit seiner Wahrscheinlichkeit gewichtet.})$$

Für eine stetige Zufallsvariable mit der Dichte f ist der Erwartungswert

$$\mu = E(X) = \int_B x \cdot f(x) dx \quad (\text{auf dem Bereich } B \text{ ist } f(x) > 0)$$

Die Varianz (Maß für die Abweichung vom Erwartungswert) einer diskrete Zufallsvariable X , lautet:

$$\sigma^2 = V(X) = \sum_{i=1}^m (a_i - \mu)^2 \cdot P(X = a_i) \quad (\text{Die Quadrate der Abstände zum Erwartungswert werden gewichtet.}) \quad \text{Es gilt: } V(X) = E(X^2) - \mu^2$$

Für eine stetige Zufallsvariable ist die Varianz

$$\sigma^2 = V(X) = \int_B (x - \mu)^2 \cdot f(x) dx \quad \text{Zeige: } V(X) = \int_B x^2 \cdot f(x) dx - \mu^2$$