Trainingsblatt 11: **Lineare oder quadratische Regression**

Was man weiß, was Schüler/in wissen sollte. Check-Up zu Begriffen und Schreibweisen.

Eine Gerade ist durch Angabe zweier Punkte eindeutig bestimmt/berechenbar. Bei drei oder mehr Punkten – die Grafik ergibt ein Streudiagramm – kann man mit GTR oder Computer eine lineare Funktion berechnen lassen, die sehr gut „passt“. Die grundlegende Idee geht auf C. F. Gauss zurück (Methode der kleinsten Quadrate), das Verfahren heißt hier „lineare Regression“. Jede quadratische Funktion ist durch die Angabe dreier Punkte eindeutig festgelegt. Die zugehörige Ausgleichsparabel bei vier oder mehr Punkten erhält man durch „quadratische Regression“.

Aufgaben mit Minimalanforderungen

Zeichne ein Streudiagramm (≙ Punktdiagramm) und die Ausgleichsgerade.

a) durch (3|5), (4|7) u. (5|12) mittels Peilung.
b) durch (0|0), (21|0,3), (36|0,5), (51|0,7) u. (66|1,0) mittels Peilung als Gerade durch (0|0).
c) durch (2|6), (4|5), (6|4) u. (8|3).
d) durch (6|1,5), (7|1) u. (7|0).
e) durch (6|7), (6|8) u. (7|7,5).

Aufgaben mit Standardanforderungen

Zeichne ein Streudiagramm (≙ Punktdiagramm) und die Ausgleichsgerade.

a) Schokotafeln werden an eine Feder gehängt und (Anz. d. T.| Feder-Länge in cm) gemessen: (1|12,7), (2|13,6), (4|16,3). Welche Startlänge liegt vor und was erwartest du bei 8 Tafeln?
b) Prüfe, ob die Punkte A(1|0), B(-2|-4,5) und C(3|3) auf einer Geraden liegen?
c) Bestimme die Regressionsgerade und das Volumen eines Tropfens in cm³ für die Messdaten (Anz.|Volumen in cm³): (50|40), (100|90), (150|125), (200|160) u. (300|245).
d) Wähle P(9|z) so, dass der Punkt mit A(1|0,5), B(-2|-5,5), C(5|8,5) auf einer Geraden liegt.

Aufgaben mit höherem Schwierigkeitsgrad

a) Die Tabelle zeigt Daten zur Belastbarkeit von Drahtseilen.

Seildurchmesser in mm	12	14	18	20	28
Reissfestigkeit in kg	815	1050	1500	2000	8100

Um den Seildurchmesser für eine Last von 3 t abzuschätzen, wird so argumentiert: 2 mm mehr Durchmesser bringen bei 2 t Last etwa 500 kg an Reissfestigkeit, deshalb müssten es ca. 24 mm sein. Nimm begründet Stellung dazu.

b) Die Tabelle zeigt Siedetemperaturen in Abhängigkeit von der Höhenlage des Ortes.

Ort	Nordsee	Freiburg (Ch)	Zugspitze	Mt. Blanc	Mt. Everest
Höhe h in m	0	591	2963	4807	8848
Siedepunkt ϑ in °C	100	98,3	90	83	73,5

Modelliere die Höhenabhängigkeit und bestimme die Siedetemperaur auf dem 6273 m hohen Chimborasso in Ecuador.

c) Modelliere die Datenpunkte (x|y) und (y|x) im selben Graphen: (1|1), (2|3), (4|7), (6|10,5). Vergleiche die beiden Regressionsgleichungen und trage zum Vergleich y = x mit ein. Beschreibe und erläutere das Ergebnis.

Falls nicht anders vorgegeben, löst du die Aufgaben so wie in den Unterrichtsbeispielen. Durch die Dokumentation deines Lösungsweges wird immer deutlich, wie du selbst vorgegangen bist. Natürlich muss auch die Verwendung des eingeführten Rechners ersichtlich sein. Du findest hier deshalb nur die Endergebnisse oder mal einen Tipp für einen besonderen Lösungsweg.

LÖSUNGEN

Aufgaben mit Minimalanforderungen

Zeichne ein Streudiagramm ($\hat{=}$ Punktdiagramm) und die Ausgleichsgerade.

a) $y \approx 3{,}5x - 6$.
b) $y \approx 0{,}015x$.
c) Die 4 Punkte liegen exakt auf $y = -0{,}5x + 7$.
d) $y \approx -1{,}167x + 8{,}33$.
e) $y = 0x + 7{,}5$.

Aufgaben mit Standardanforderungen

Zeichne ein Streudiagramm ($\hat{=}$ Punktdiagramm) und die Ausgleichsgerade.

a) Start-Federlänge ca. 11,4 cm, wegen (0|11,35) und ca. 21,1 cm bei 8 Tafeln, wegen (8|21,1).
b) Ja, auf $y = 1{,}5x - 1{,}5$
c) $y \approx 0{,}8x + 1{,}35$. Ein einzelner Tropfen hat ca. 0,8 cm³.
d) Wähle b = 16,5.

Aufgaben mit höherem Schwierigkeitsgrad

a) Lineare Regression ist hier über den vollen Bereich nicht sinnvoll. Sie liefert einen Wert, der mit 24 mm lokal zu groß – also auf der sicheren Seite – ist.
   Bei quad. Regression $y \approx 36{,}8x^2 - 1028x + 8014$ käme man mit ca. 22 mm aus.
b) $\vartheta(h) \approx 99{,}5 - 0{,}00306 \cdot \vartheta$. Auf dem Chimborasso ist sie ca. 80,3 °C, wegen $\vartheta(6273) \approx 80{,}3$.
c) Es sind Umkehrfunktionen zueinander. Man sieht es an der Spiegelung bzgl. $y = x$ und dem Fixpunkt (1|1) sowie an den Steigungen der Reg.-Geraden, die Kehrwerte voneinander sind.

Hinweise zum Teacher-Tool "REGRESS"
Das Utility bietet den schnellen Zugriff auf die Lösung des Standardproblems:
„Wie lautet die beste lineare oder quadratische Funktion durch mindestens 3 Datenpunkte?“.
Durch die einfache Dateneingabe kann das vorliegende Tool drei Dinge erleichtern:

- Die Erstellung eigener Arbeitsblätter
- Die innere Differenzierung
- Die Erstellung von Tests und Klassenarbeiten

Hinweise zum Einsatz des Arbeitsblattes
Um die Selbstkontrolle für die Schüler/innen zu ermöglichen, sind lediglich die Endlösungen angegeben. Damit wird kein irgendwie gearteter Lösungsgang bevorzugt und das Arbeitsblatt bleibt universell einsetzbar.

Die grafische Darstellung der Daten ist i.d.R. mindestens dann notwendig sobald ein praktisches Problem dahinter steht. Zudem ist das Finden einer sinnvollen Skalierung stets als gute Übung zu sehen, denn der Modus „Auto“ für Grafiken im STAT-Modus ermöglicht häufig keinen guten Überblick.

Die Vielfalt der Lösungswege (klassisch händisch: Punkte in ein KoSy eintragen, Gerade per „Peilung“ bestimmen/einzeichnen und anschließend m und b ablesen, den Mittelwert geeigneter Grenzgeraden bestimmen, bei großer Datenanzahl kann man i.d.R. 2 Punkte hernehmen und aus ihnen die Regressionsgleichung bestimmen) wird durch den Rechnereinsatz nochmals erweitert (mit Rechner: Plot der Daten mittels Listen und anschließender linearer Regression sowie Darstellung der Gleichung, umfängliche Lösungsoptionen per G-Solv im Grafikmodus, insbesondere Interpolationen und Hochrechnungen werden stark vereinfacht).
Die Datenplots sowie die linearen Regressionen werden die Schüler in der Regel mittels Rechnereinsatz und geeigneten Datenlisten durchführen. Zur Dokumentation ihrer Ergebnisse sind dann letzlich wieder klassisch händische Fertigkeiten – wie z.B. aussagekräftige Skizzen mit passender Skalierung und Beschriftung – gefordert. Die Regressionsgeraden sind mit sinnvoller Genauigkeit (3 bis 4 gültige Stellen) anzugeben.