Mathematik, Studium MINT (MA 8802)

Tutorübung (15.-19.5.2017)

T 3.1. (**Regressionsgerade**)

Gegeben sei der folgende Datensatz:

x	2	3	4	5
y	1,4	1,6	1,9	2,1

a) Bestimmen Sie mittels Satz 5.1 (VL 4) und eines Taschenrechners die Regressionsgerade.

b) Skizzieren Sie Daten und Regressionsgerade.

c) Bestimmen Sie den relativen Standardfehler f (siehe VL 5).

T 3.2. (**Eine Eigenschaft der Regressionsgerade**)
Gegeben sei ein beliebiger Datensatz von $x-$Werten $x_1, ..., x_n$ und zugehörigen $y-$Werten $y_1, ..., y_n$. Die entsprechende Regressionsgerade sei $f(x) = a \cdot x + b$.

a) Zeigen Sie: Der Mittelwert der Fehler $r_i = y_i - (a \cdot x_i + b)$ ist gleich Null.

b) Ist es möglich, dass alle Datenpunkte oberhalb der Gerade (oder alle unterhalb der Gerade) liegen?

T 3.3. (**Minimalitätseigenschaft des Mittelwertes**)
Zeigen Sie: der Mittelwert $\overline{x}$ von n Zahlen $x_1, ..., x_n$ minimiert den mittleren quadratischen Fehler

$$R(\overline{x}) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2.$$

Hinweis: Leiten Sie R nach $\overline{x}$ ab und setzen Sie die Ableitung gleich 0.

T 3.4. (**Regressionskurve: Geschwindigkeit chemischer Reaktionen**)
Die Reaktionzeit chemischer Prozesse hängt von der Konzentration der Reaktanden ab. Als Beispiel betrachten wir die Bildung von Schwefeldioxid aus einer Natriumthiosulfatlösung der Konzentration c und Salzsäure. Gemessen wurde die Reaktionszeit t in Sekunden bis zu einem gewissen Grad der Trübung (die die Bildung von Schwefeldioxid anzeigt) in Abhängigkeit der Konzentration c. Die Messwerte sind in der folgenden Tabelle gegeben:

c [mol/l]	0,16	0,128	67	127
t [sec]	35	45	0,096	0,064

a) Skizzieren Sie die Daten.

b) Bestimmen Sie die Zehnerlogarithmen der Daten und skizzieren Sie Daten einfach-logarithmisch mit skalierter y-Achse, einfach-logarithmisch mit skalierter x-Achse, und

Bitte wenden!

doppelt-logarithmisch. Welche funktionale Abhängigkeit vermuten Sie?

c) Bestimmen Sie für die geeigneteste Darstellung die Regressionsgerade.

d) Bestimmen Sie durch Rücktransformation diejenige Funktion, die die ursprünglichen Daten näherungsweise darstellt. Skizzieren Sie Funktion und ursprüngliche Daten im selben Koordinatensystem.

T 3.5. (**Zusatzaufgabe (freiwillig, schwieriger): Toleranzgrenze für lineare Regression**)
Zeigen Sie: für Datenpunkte einer *quadratischen* Funktion, $x_j = j$, $y_j = j^2$ $(j = 1, ..., n)$, konvergiert der relative Standardfehler f der Regressionsgerade im Grenzwert einer grossen Anzahl Datenpunkte $(n \to \infty)$ gegen $1/4 = 0.25$. (Wie in VL5 erläutert: eine Regressionsgerade ist also "wertlos", wenn der relative Standardfehler diese Grenze erreicht: statt der durch den Ansatz $f(x) = a \cdot x + b$ postulierten näherungsweise affin-linearen Abhängigkeit könnte eine quadratische Abhängigkeit vorliegen.)

Hinweis: Bestimmen Sie zunächst die Mittelwerte $\overline{x}$ und $\overline{y}$, dann die Parameter a und b der Regressionsgerade, und schliesslich die Summen, die in Zähler bzw. Nenner der Formel für f auftreten. Versuchen Sie hierbei *nicht*, Differenzen wie $x_j - \overline{x}$ auszurechnen, sondern multiplizieren Sie Ausdrücke wie z.B. $(x_j - \overline{x})(y_j - \overline{y})$ aus,

$$(x_j - \overline{x})(y_j - \overline{y}) = x_j y_j - \overline{x} y_j - x_j \overline{y} + \overline{x}\overline{y},$$

und summieren Sie die vier Terme auf der rechten Seite einzeln. Sie dürfen die folgenden Summenformeln für Potenzen der ersten n natürlichen Zahlen verwenden:

$$1 + ... + n = \tfrac{n(n+1)}{2}, \quad 1^2 + ... + n^2 = \frac{n(n+1)(2n+1)}{6},$$
$$1^3 + ... + n^3 = \tfrac{n^2(n+1)^2}{4}, \quad 1^4 + ... + n^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}.$$

Aktuelle Informationen und Materialien zur Vorlesung finden Sie auf der Vorlesungsseite

`http://www-m7.ma.tum.de/bin/view/Analysis/MINT17`