

## Physikalisch-Chemisches Praktikum I

### **Protokoll zu Versuch 30: Statistik**

Bei der Auswertung von Meßreihen jeglicher Art ist es wichtig, sich über die Genauigkeit eines Ergebnisses im Klaren zu sein. So ist zum Beispiel eine aus mehreren Punkten gewonnene Ausgleichsgerade durch die Streuung der Punkte um diese herum mit einem gewissen Fehler behaftet. Das mathematische Verfahren, aus den Meßwerten eine Gerade zu erhalten nennt man lineare Regression.

Die allgemein bekannte Gleichung für eine Gerade lautet :

$$y = a + x \cdot b$$

Die Abweichung der Meßpunkte  $x_i, y_i$  ist eine Funktion von  $a$  und  $b$ . Um das Minimum dieser Abweichung zu finden, leitet man den Ausdruck für die Summe der Quadrate der Abweichungen partiell nach  $a$  und nach  $b$  ab. Die beiden sich ergebenden Gleichungen liefern die Ausdrücke für die Steigung der Geraden  $b$  und den  $y$  - Achsenabschnitt  $a$ , wobei  $N$  die Anzahl der Meßwerte darstellt :

$$b = \frac{\sum_i^N x_i \cdot y_i - (\sum_i^N x_i \sum_i^N y_i) / N}{\sum_i^N x_i^2 - (\sum_i^N x_i)^2 / N} \qquad a = (\sum_i^N y_i - b \cdot \sum_i^N x_i) / N$$

Ziel des ersten Teilversuches war es nun, mit Hilfe eines Computers in der Programmiersprache BASIC ein Programm zu schreiben, das der Reihe nach die Meßwerte abfragt und danach die Werte für  $a$  und  $b$  berechnet. Das Programm ruft am Ende noch das Programm LRPLOT auf, welches die Meßpunkte und die erhaltene Gerade graphisch darstellt (s. Anhang).

Hierbei sind  $A$  und  $B$  die oben erwähnten Werte  $a$  und  $b$ , die über die Aufsummierung in der Schleife und anschließende Verrechnung entstehen.

Der zweite Teilversuch beschäftigt sich mit der Bernoulli- und Gauß-Verteilung sowie der Standardabweichung und dem Meßprobenumfang. Auch diese Verfahren sind von Bedeutung, um eine Aussage

über die Verlässlichkeit eines Ergebnisses machen zu können oder um einzuschätzen nach wie vielen Messungen das Ergebnis eine gewisse Stabilität erreicht hat.

Mathematische Grundlage bilden hier stochastische und kombinatorische Berechnungen.

Bezeichnet man die Wahrscheinlichkeit für das Eintreffen eines Ereignisses mit  $p$  und die seines Nichttreffens mit  $q$ , dann sagt die Bernoulli - Verteilung für die Wahrscheinlichkeit  $W$  des Eintreffens von  $n$  Ereignissen bei  $N$  Versuchen voraus :

$$W_{\text{Bernulli}} = \binom{N}{n} \cdot p^n \cdot q^{N-n} \quad \text{wobei der Binominalkoeffizient } \binom{N}{n} \text{ durch } \binom{N}{n} = \frac{N!}{n!(N-n)!} \text{ gegeben ist}$$

Für den Fall  $N \rightarrow \infty$ , also einer unendlichen Anzahl von Versuchen geht die Binominalverteilung in die Gaußverteilung über :

$$W_{\text{Gauß}} = \frac{1}{\sqrt{2\pi \cdot N \cdot p \cdot q}} \cdot e^{-\frac{(n-N \cdot p)^2}{2N \cdot p \cdot q}}$$

Bei einer verhältnismäßig geringen Anzahl an Meßwerten sollte also die Bernoulliverteilung die bessere Anpassung die wahre Wahrscheinlichkeit liefern als die Gaußverteilung. Der unumstrittene Vorteil der Gaußverteilung liegt jedoch in der Handhabbarkeit : Exponentialfunktionen sind wesentlich weniger schwierig abzuleiten oder zu integrieren als Fakultäten.

Für die Mittelwerte dieser beiden Verteilungen ergibt sich :

$$\bar{n} = \frac{1}{N} \cdot \sum_i^N n_i = N \cdot p$$

Die Streuung ist gegeben durch :

$$\sigma = \sqrt{N \cdot p \cdot q}$$

Die Menge der endlich vielen Meßwerte aus der Grundgesamtheit bildet eine Stichprobe . Für die geschätzte ergibt sich für die Standardabweichung :

$$\sigma = \sqrt{\frac{1}{N-1} \cdot \sum_i^N (n_i - \bar{n})^2}$$

Vor diesem Hintergrund wurde ein zweites Programm geschrieben, das mit Hilfe des Zufallszahlengenerators  $K$  Würfe mit  $N$  Würfeln simulieren soll. Es interpretiert eine zwischen 0 und  $1/6$  liegende Zahl als eine gewürfelte Sechs und die Zahlen zwischen  $1/6$  und 1 als „keine Sechs“ (s. Anhang).

Die folgende Tabelle zeigt die Schätzungen für Mittelwert und Standardabweichung der Gauß-Verteilung für die verschiedenen Versuchsdurchläufe verglichen mit denen der Grundgesamtheit:

| N (Anzahl der Würfel) | K (Anzahl der Würfe)       | Mittelwert | Standardabweichung |
|-----------------------|----------------------------|------------|--------------------|
| 50                    | $\infty$ (Grundgesamtheit) | 8 1/3      | 2,6352             |
|                       | 1000000                    | 8,333397   | 2,6350             |
|                       | 100000                     | 8,3278     | 2,6390             |
|                       | 10000                      | 8,3177     | 2,6233             |
|                       | 1000                       | 8,343      | 2,6424             |
|                       | 100                        | 7,91       | 2,4041             |
|                       | 10                         | 7,7        | 2,4060             |
|                       | $\infty$                   | 1 2/3      | 1,1785             |
| 10                    | 1000000                    | 1,6649     | 1,1775             |

Es ist deutlich erkennbar, dass sich der gefundene Mittelwert dem Mittelwert der Grundgesamtheit mit steigender Anzahl an Würfeln immer mehr nähert. So kann also in der Praxis durch häufiges Messen der zufällige Fehler minimiert werden.

Weiterhin ist sehr gut erkennbar, dass die Standardabweichung / Streuung für die geringere Anzahl an Würfeln wesentlich kleiner geworden ist, was auch aus obiger Formel für die Streuung hervorgeht. Bei einer Minimierung der äußeren Einflüsse (Fehler) ist also ein Meßwert mit einer größeren Genauigkeit bestimmbar. Das leuchtet ein, denn die Streuung ist graphisch gesehen für die Breite der Gauß-Kurve verantwortlich, wobei eine sehr schmale Kurve eine sehr geringe Abweichung der Meßwerte vom realen Wert bedeutet.

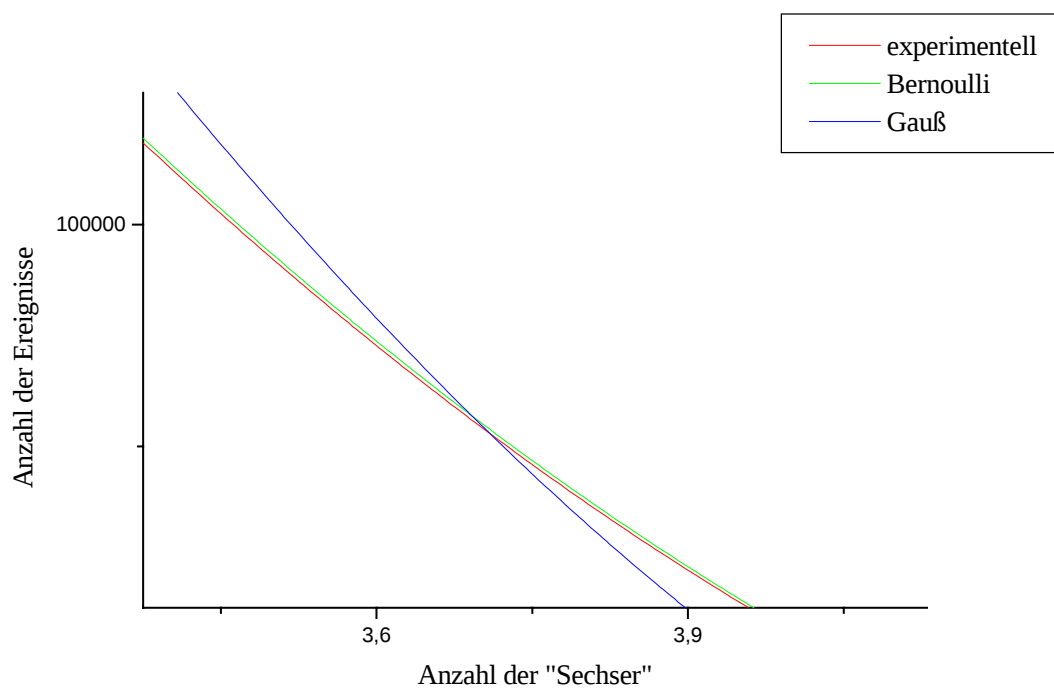
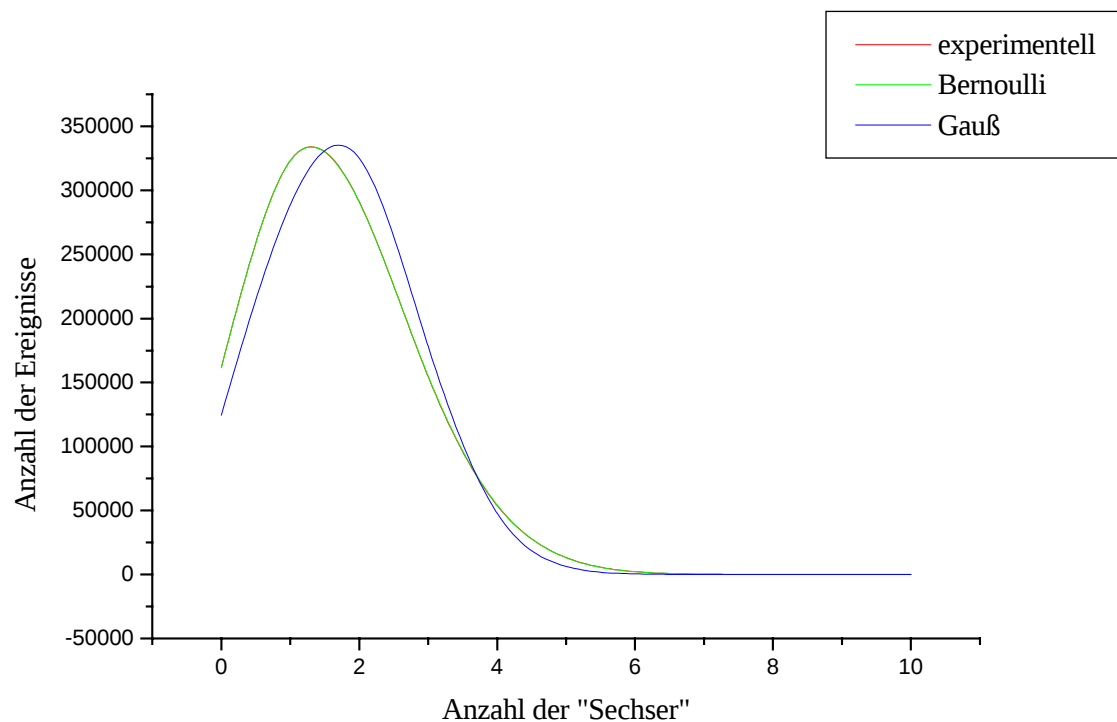
Anhand der Ergebnisse des Versuches mit 10 Würfeln und einer Million Würfe, soll die Anpassung durch die Gauß- und die Bernoulli-Verteilung an die wahren Werte untersucht werden.

Die nachstehende Tabelle zeigt die Ergebnisse des Versuches.

| Anzahl der „Sechser“ | Verteilung im Experiment | Bernoulli-Verteilung | Gauß-Verteilung |
|----------------------|--------------------------|----------------------|-----------------|
| 0                    | 161722                   | 161505,60            | 124544,00       |
| 1                    | 323009                   | 323011,20            | 288489,60       |
| 2                    | 291224                   | 290710,00            | 325271,10       |
| 3                    | 154630                   | 155045,30            | 178512,60       |
| 4                    | 53905                    | 54265,87             | 47687,01        |
| 5                    | 13084                    | 13023,81             | 6200,681        |
| 6                    | 2105                     | 2170,63              | 392,452         |
| 7                    | 247                      | 248,07               | 12,09043        |
| 8                    | 24                       | 18,605               | 0,18            |
| 9                    | 0                        | 0,83                 | 0               |
| 10                   | 0                        | 0                    | 0               |

Die deutliche Abweichung der Gauß-Verteilung von den anderen beiden zeigt, dass diese eine mit Vorsicht zu genießende Näherung ist. Zwar sind eine Million Würfe sehr viel, aber da die Kurven der Theorie nach erst im Unendlichen zusammenfallen ist die Divergenz immer noch merklich spürbar.

Die experimentelle und die Bernoulli-Verteilung liegen so nahe beieinander, dass sie im folgenden Graphen nicht mehr voneinander zu unterscheiden sind. Erst die Vergrößerung zeigt die (geringe) Abweichung.





```
FOR B=1 TO N
  X=RND
  IF X<1/6 THEN I=I+1
NEXT B

H(I+1)=H(I+1)+1
NEXT A

H(N+1)=1*exp(-20)

B(1)=(5/6)^n
FOR I=1 TO N+1
  B(I+1)=B(I)*(N-(I-1))/I*(1/5)
NEXT I

FOR I=1 TO N+1
  G(I)=1/SQR(2*3,141*N*5/36)*exp(-((I-1)-N/6)^2/N*3.6)
NEXT I

FOR I=1 TO N+1
  PRINT I-1,H(I),K*B(I),K*(G(I))
NEXT I

CALL RNDPLOT (H())
CALL RNDPLOT (B())
CALL RNDPLOT (G())

END
```

{würfelt mit N Würfeln und zählt die  
geworfenen Sechsen}

{Zählt die jeweiligen Ereignisse}

{Skalierung der x-Achse durch sehr kleinen  
Wert am Ende der Verteilung}

{Schleife für die Bernoulli-Verteilung mit  
Rekursionsformel:  
$$W(I+1, N, p) = W(1, N, p) \cdot \frac{N - (I-1)}{I \cdot 5}$$
}

{Schleife, die die Gauß-Verteilung berechnet}

{Schleife, die die Ereignisse anzeigen läßt}

{Ruft das Programm für die graphische Dar-  
stellung auf}