Versuch 001

Behandlung statistisch verteilter Messwerte

Aufgaben:

1. Führen Sie am PC Experimente mit dem Galtonschen Nagelbrett durch. Lassen Sie mehrmals Kugelmengen durchlaufen und registrieren Sie die Häufigkeitsverteilung der durchlaufenen Kugeln. Zeigen Sie, dass die Häufigkeitsverteilung eine Binomialverteilung ist.
2. Messen Sie die natürliche Radioaktivität im Labor mit Hilfe eines geeigneten Zählgerätes (in dem Fall: Szintillationszähler). Nehmen Sie insgesamt 4x 100 Messwerte auf. Diese Werte bilden den Datenbestand. Zu bestimmen sind: Mittelwert, Varianz und Standartabweichung. Als Auswertung wird ein Histogramm für je 100 und eins für 400 Messwerte benötigt.
3. Welche Schlussfolgerungen lassen sich aus dem Vergleich zwischen Varianz und Mittelwert bilden?
4. Tragen Sie die gemessenen Summenhäufigkeiten auf Wahrscheinlichkeitspapier auf und schlussfolgern Sie!

Grundlagen

Mittelwert, Standartabweichung, Varianz

Um eine Datenmenge einer Messung hinsichtlich der Genauigkeiten der einzelnen Messwerte darzustellen, benötigt man einen Leitwert, an dem sich alle Messwerte relativieren. Diesen Wert nennt man Mittelwert. Je nach der Art der Datenreihen, für die ein Mittelwert benötigt wird, verwendet man

das arithmetische Mittel (für lineare Reihen) $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$,

das geometrische Mittel $\overline{x} = \sqrt[n]{\prod_{i=1}^{n} x_i}$,

und das harmonische Mittel $\overline{x} = n\sum_{i=1}^{n}\frac{1}{x_i}$.

Da unsere Messwerte alle linear sind, verwenden wir den arithmetische Mittelwert. Vorteile: Der Wert gibt den aus der geometrischen Anschauung bekannten Mittel- oder Schwerpunkt zwischen den auf einer Achse abgetragenen Messwerten an. Mit anderen Worten: Er gleicht die positiven und negativen Messungenauigkeiten aus.

Aus der Gaussschen Statistiktheorie (Theorie der kleinsten Quadrate) entnehmen wie die Formel für die Varianz: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\overline{x} - x_i)^2$ und aus ihr abgeleitet die Standartabweichung: $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\overline{x} - x_i)^2}$

Häufigkeitsverteilung, Binomialverteilung

Für den Zufallsversuch mit n voneinander unabhängigen Stufen mit jeweils gleichen Wahrscheinlichkeiten gilt für den Weg A: $P(A) = \prod_{i=1}^{n} p_i$

Damit haben in $z \cdot P(A)$ (z ist die Anzahl der gestarteten Zufallsversuche) die ideale, absolute Häufigkeit für das Durchlaufen des Weges A, während die ideale relative Häufigkeit $P(A)$ ist. (Ideal bedeutet für sehr große z, respektive z geht gegen Unendlich.) Für das Galtonsche Brett mit n Reihen gilt auf Grund der Wegadditionen für die relative Häufigkeit des Ergebnisses: $P(k) = \binom{n}{k}\frac{1}{2}^n$

Das Galtonsche Brett stellt wegen seiner Einzelwahrscheinlichkeit von 50% einen Sonderfall der Binomialverteilung dar: $P(k) = \binom{n}{k}p^k(1-p)^{n-k}$, es gilt ferner: $\overline{x} = n \cdot p$ und $s^2 = \overline{x}(1-n)$

Poissonverteilung, Normalverteilung

Abgeleitet daraus kommt man für sehr kleine p bei konstanten pn (d.h. n muss gegen Unendlich gehen) zur Poissonverteilung

$P(x,\bar{x}) = \frac{\bar{x}^x}{x!} e^{-\bar{x}}$ mit $s^2 = \bar{x}$.

Gauss wiederum entwickelte im achtzehnten Jahrhundert eine Funktion zur Normalverteilung direkt gemessener Größen in Abhängigkeit von dem Messwert, den Mittelwert und die Varianz.

$$G(x,\bar{x},s^2) = \frac{1}{s\cdot\sqrt{2\Pi}} \cdot e^{-\frac{(x-\bar{x})^2}{2s^2}}$$

Histogramm

Graphische Darstellung der Häufigkeitsverteilung um den Mittelwert unter Zuhilfenahme einer selbst gewählten Unschärfe (Einteilung der Abszisse in Intervalle) und der daraus resultierende Summenabtrag auf der Ordinate. → siehe: Ergebnisse

Versuchsdurchführung

Versuchsteil 1: Galtonsches Nagelbrett

Versuchsobjekt:
Nagelbrett- Simulation am PC mit n- Zufalls- Ebenen

Mögliche systematische Fehler:
Keine wirklichen Zufallswerte, da Messwerte durch eine mathematische Funktion des PCs Zufallswerte nur simulieren.

Zu erwartendes Ergebnis:
Eine Binomialverteilung mit gleichen Einzelwahrscheinlichkeiten für Links und Rechts. Die Anzahlen der Kugeln in dem k- ten Ausgang müsste nach den Gesetzen der Statistik sich annähernd nach folgender Gleichung ergeben:

$$M(k) = M_{Ges}\binom{n}{k}\frac{1}{2}^n$$

Versuchsablauf:
- Messung von M(k)
- Vergleich mit den theoretischen Werten

Versuchsteil 2: Messen der natürlichen Radioaktivität des Labors

Versuchsobjekt:
Szintillationszähler, PC zur Aufnahme von Messwerten

Mögliche systematische Fehler:
- falsche Kalibrierung des Sensors (für den Versuch aber unbedeutend)

Formeln für den Mittelwert, die Varianz und die Standartabweichung sind in den *Grundlagen* zu finden.

Fehlerabschätzung:
Relativer Fehler von s: $\frac{\Delta s}{s} = \frac{1}{\sqrt{2(n-1)}}$,

Standartfehler des Mittelwertes: $s_m = \frac{s}{\sqrt{n}}$

Versuchsablauf:
- Aufnahme der Daten
- Umsetzen in Histogramm und graphische Darstellung in ein Wahrscheinlichkeitsnetz am PC

Messwerte

Versuchsteil 1: Galtonsches Nagelbrett

Anzahl der Kugeln:

Kanal:	100	1000	10000	Summe
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	1	1	2
4	0	0	5	5
5	0	0	22	22
6	0	10	81	91
7	1	23	185	209
8	5	46	458	509
9	7	89	779	875
10	10	114	1194	1318
11	12	143	1485	1640
12	19	159	1596	1774
13	15	156	1459	1630
14	15	123	1207	1345
15	11	66	760	837
16	1	46	447	494
17	3	16	208	227
18	0	6	83	89
19	1	1	25	27
20	0	0	5	5
21	0	1	0	1
22	0	0	0	0
23	0	0	0	0
24	0	0	0	0
Summe	100	1000	10000	

Versuchsteil 2: Messen der natürlichen Radioaktivität des Labors

Pos	Intervall	1. Messreihe	2. Messreihe	3. Messreihe	Summe:
1	[100;104]	0	1	1	2
2	[105;109]	1	2	3	6
3	[110;114]	2	4	5	11
4	[115;119]	10	10	20	40
5	[120;124]	16	13	22	51
6	[125;129]	21	20	33	74
7	[130;134]	15	20	38	73
8	[135;139]	13	9	29	51
9	[140;144]	10	15	21	46
10	[145;149]	5	2	15	22
11	[150;154]	3	2	8	13
12	[155;159]	4	1	3	8
13	[160;164]	0	0	1	1
14	[165;169]	0	0	1	1
15	[170;174]	0	1	0	1
	Summe:	100	100	200	400

Die abgezeichneten Messwerte sind im Anhang zum Protokoll zu finden.

Auswertung

Versuchsteil 1: Galtonsches Nagelbrett

Nach den in den Grundlagen näher beleuchteten Verteilungsvorschriften können wir die theoretische Verteilung mittels der Binomialkoeffizienten für die 3 Messreihen sowie für deren Summe die Häufigkeitsverteilungen errechnen (Ergebnisse gerundet):

Kanal:	$\binom{n}{k}$	p(Kanal)	H(100,Kanal)	H(1000,Kanal)	H(10000,Kanal)	H(1110,Kanal)
0	1	0,00000006	0	0	0	0
1	24	0,00000143	0	0	0,01	0,02
2	276	0,00001645	0	0,02	0,16	0,18
3	2024	0,00012064	0,01	0,12	1,21	1,34
4	10626	0,00063336	0,06	0,63	6,33	7,03
5	42504	0,00253344	0,25	2,53	25,33	28,12
6	134596	0,00802255	0,8	8,02	80,23	89,05
7	346104	0,02062941	2,06	20,63	206,29	228,99
8	735471	0,04383749	4,38	43,84	438,37	486,6
9	1307504	0,07793331	7,79	77,93	779,33	865,06
10	1961256	0,11689997	11,69	116,9	1169	1297,59
11	2496144	0,14878178	14,88	148,78	1487,82	1651,48
12	2704156	0,16118026	16,12	161,18	1611,8	1789,1
13	2496144	0,14878178	14,88	148,78	1487,82	1651,48
14	1961256	0,11689997	11,69	116,9	1169	1297,59
15	1307504	0,07793331	7,79	77,93	779,33	865,06
16	735471	0,04383749	4,38	43,84	438,37	486,6
17	346104	0,02062941	2,06	20,63	206,29	228,99
18	134596	0,00802255	0,8	8,02	80,23	89,05
19	42504	0,00253344	0,25	2,53	25,33	28,12
20	10626	0,00063336	0,06	0,63	6,33	7,03
21	2024	0,00012064	0,01	0,12	1,21	1,34
22	276	0,00001645	0	0,02	0,16	0,18
23	24	0,00000143	0	0	0,01	0,02
24	1	0,00000006	0	0	0	0
Summe		1	100	1000	10000	

Wir vergleichen Theorie und Praxis in graphischer Darstellung (→ siehe: Abb. I.1)

Versuchsteil 2: Messen der natürlichen Radioaktivität des Labors

Zur Ermittlung des Mittelwertes müssen wir unser arithmetisches Mittel für den Sachverhalt modifizieren:

$$\bar{x} = \frac{\sum_{i=1}^{pos}[\overline{pos(i)} \cdot n(i)]}{\sum_{i=1}^{pos} n(i)}$$

, wobei $pos = 15$ (Anzahl der Intervalle), $\overline{pos(i)}$ das Intervallmittel des i- ten Intervalls, sowie $n(i)$ die Anzahl der registrierten Messwerte im i-ten Intervall ist.

ferner berechnet sich Varianz und Standartabweichung folgender Maßen:

$$s^2 = \frac{1}{n-1}\sum(x-\bar{x})^2 = \frac{1}{\sum_{i=1}^{pos} n(i) - 1}\sum_{i=1}^{pos} n(i)(\overline{pos(i)} - \bar{x})^2$$

Damit ergibt sich folgende Wertetabelle (auf 3 Stellen gerundet):

Messwerte		Mittelwerte				Varianz und Standartabweichung:				
Pos	Intervall	1. MR	2. MR	3. MR	Ges.		1. MR	2. MR	3. MR	Ges.
1	[100;104]	0	102	102	204		0	778,41	897,0025	1711,125
2	[105;109]	107	214	321	642		585,64	1048,82	1867,5075	3528,375
3	[110;114]	224	448	560	1232		737,28	1281,64	1990,0125	4076,1875
4	[115;119]	1170	1170	2340	4680		2016,4	1664,1	4470,05	8122,5
5	[120;124]	1952	1586	2684	6222		1354,24	811,33	2178,055	4363,6875
6	[125;129]	2667	2540	4191	9398		370,44	168,2	808,5825	1336,625
7	[130;134]	1980	2640	5016	9636		9,6	88,2	0,095	41,0625
8	[135;139]	1781	1233	3973	6987		437,32	453,69	739,5725	1686,1875
9	[140;144]	1420	2130	2982	6532		1166,4	2196,15	2121,0525	5315,875
10	[145;149]	735	294	2205	3234		1248,2	584,82	3397,5375	5457,375
11	[150;154]	456	304	1216	1976		1297,92	976,82	3216,02	5597,3125
12	[155;159]	628	157	471	1256		2662,56	734,41	1882,5075	5304,5
13	[160;164]	0	0	162	162		0	0	903,0025	945,5625
14	[165;169]	0	0	167	167		0	0	1228,5025	1278,0625
15	[170;174]	0	172	0	172		0	1772,41	0	1660,5625
	Mittelwert:	131,2	129,9	131,95	131,25	s^2:	120,061	126,859	129,143	126,378
						s:	10,957	11,263	11,364	11,242

Damit erhalten wir einen ungefähren Standartfehler des Mittels und einen Fehler von s bei:

$$s_1=\pm1{,}10;\ s_2=\pm1{,}13;\ s_3=\pm0{,}81;\ s_4=\pm0{,}57$$
$$\Delta s_1=\pm0{,}779;\ \Delta s_2=\pm0{,}801;\ \Delta s_3=\pm0{,}570;\ \Delta s_4=\pm0{,}400$$

Eine graphische Auswertung findet in den Abbildungen Abb II.1 bis Abb II.4 statt.

Ergebnisse

Versuchsteil 1: Galtonsches Nagelbrett

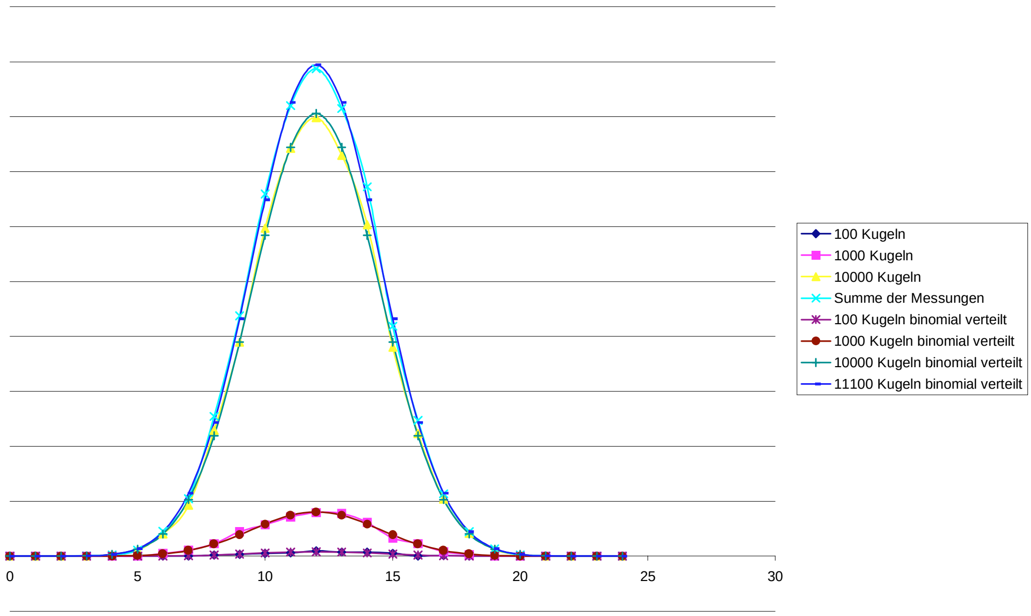

Abb. I.1: Die Kurven der Messergebnisse und theoretischen Voraussagen sind hier übereinandergelegt. Man erkennt deutlich, dass sich die Messungen in guter Näherung wie die Berechnungen verhalten. Somit kann man von binomial verteilten Messgrößen sprechen.

Versuchsteil 2: Messen der natürlichen Radioaktivität des Labors

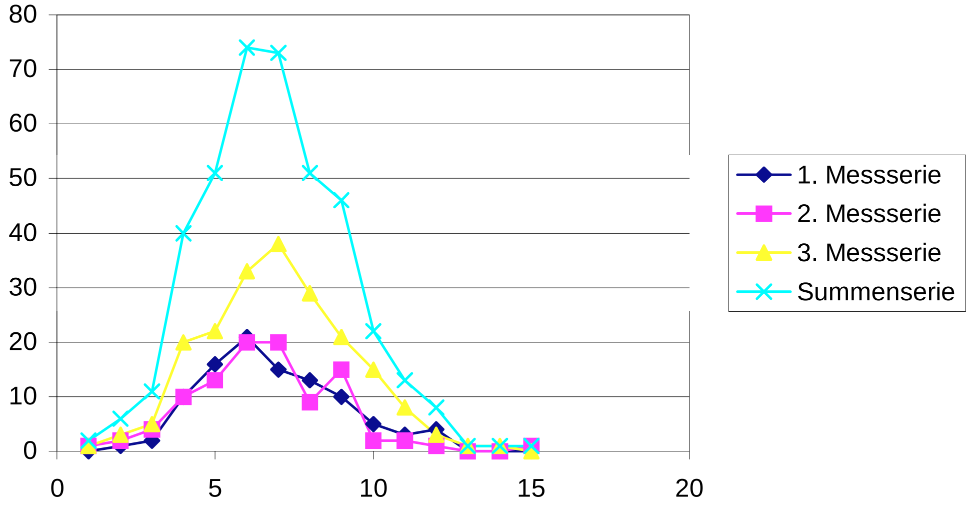

Abb. II.1: Die einzelnen Messwertserien als Fieberkurve in einem Diagramm übereinander gelegt.

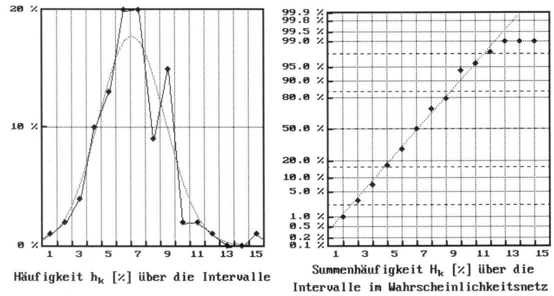

Abb. II.2: Die erste Messserie aufgetragen auf einer Relativ- und einer Wahrscheinlichkeitsskala. Zum Vergleich ist die Gaußkurve links mit eingetragen für 100 Messwerte (2. Messserie),...

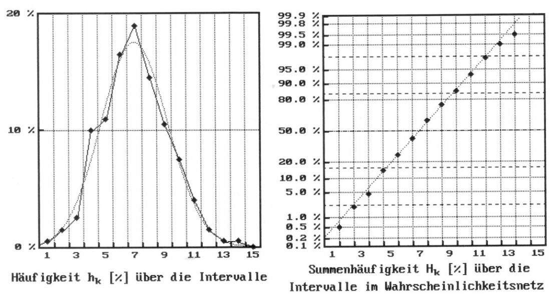

Abb. II.3: ... für 200 Messwerte...

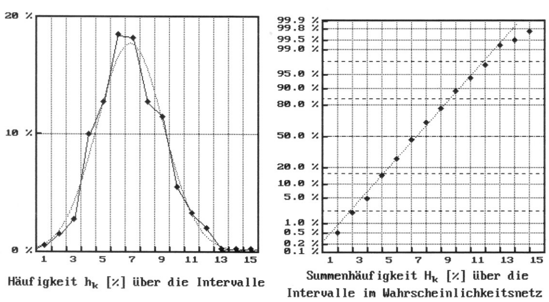

Abb. II.4: ... und für die Messwertsumme.

Diskussion:

Die Vermutungen betreffs des Versuchsteils I konnten durch die Anschauung bestätigt werden: Der Versuch mit dem Galtonsche Nagelbrett führt im Unendlichen Durchlauf zu einer Binomialverteilung. Damit sind wir zu einem wichtigen Schluss gekommen: Je größer die Anzahl der Zufallsversuche mit gleicher Versuchsapparatur ist, desto genauer werden die Ergebnisse, d.h. desto weniger weichen diese von den theoretischen Messwerten ab. Ergo: Je mehr Messergebnisse man aufnimmt, desto geringer ist die Gefahr, „neben" der theoretischen Voraussage zu liegen.

Die Versuchsreihe des zweiten Versuchs hingegen zeigt deutlich, dass bei Quantisierung der Messergebnisse Einbußen der Genauigkeit hinzunehmen sind (erkennbar an den unförmigen Kurven im Diagramm). Trotz dessen können wir mit hoher Wahrscheinlichkeit von einer Normalverteilung sprechen, da wir mit relativ wenigen Messwerten agierten. Zur Bestätigung wäre ein Versuch mit mindestens 10000 Messwerten zu empfehlen. Für diesen Versuch können wir feststellen, die graphische Auswertung (Abb. II.2 bis II.4) deuten auf eine Normalverteilung hin.
Als Bemerkung zum Versuch wäre noch anzubringen, dass sich unser Arbeitsplatz grundlegend von den der anderen Gruppen unterschied: wir nahmen nicht die Radioaktivität des Labores, sondern im Inneren eines Präperatetresors auf. Damit sind die Werte natürlich um ein Vielfaches kleiner (Abschätzung: etwa ein Viertel der Laborwerte).

Die aufgenommenen Messwerte befinden sich auf den folgenden zwei Seiten als Anhang.

Jena, 02.11.2001